Zbigniew Les, Magdalena Les

Shape Understanding Systems

Studies in Computational Intelligence, Volume 86

Editor-in-chief
Prof. Janusz Kacprzyk
Systems Research Institute
Polish Academy of Sciences
ul. Newelska 6
01-447 Warsaw
Poland
E-mail: kacprzyk@ibspan.waw.pl

Further volumes of this series can be found on our homepage: springer.com

Vol. 62. Lakhmi C. Jain, Raymond A. Tedman and Debra K. Tedman (Eds.)
Evolution of Teaching and Learning Paradigms in Intelligent Environment, 2007
ISBN 978-3-540-71973-1

Vol. 63. Wlodzislaw Duch and Jacek Mańdziuk (Eds.)
Challenges for Computational Intelligence, 2007
ISBN 978-3-540-71983-0

Vol. 64. Lorenzo Magnani and Ping Li (Eds.)
Model-Based Reasoning in Science, Technology, and Medicine, 2007
ISBN 978-3-540-71985-4

Vol. 65. S. Vaidya, L.C. Jain and H. Yoshida (Eds.)
Advanced Computational Intelligence Paradigms in Healthcare-2, 2007
ISBN 978-3-540-72374-5

Vol. 66. Lakhmi C. Jain, Vasile Palade and Dipti Srinivasan (Eds.)
Advances in Evolutionary Computing for System Design, 2007
ISBN 978-3-540-72376-9

Vol. 67. Vassilis G. Kaburlasos and Gerhard X. Ritter (Eds.)
Computational Intelligence Based on Lattice Theory, 2007
ISBN 978-3-540-72686-9

Vol. 68. Cipriano Galindo, Juan-Antonio Fernández-Madrigal and Javier Gonzalez
A Multi-Hierarchical Symbolic Model of the Environment for Improving Mobile Robot Operation, 2007
ISBN 978-3-540-72688-3

Vol. 69. Falko Dressler and Iacopo Carreras (Eds.)
Advances in Biologically Inspired Information Systems: Models, Methods, and Tools, 2007
ISBN 978-3-540-72692-0

Vol. 70. Javaan Singh Chahl, Lakhmi C. Jain, Akiko Mizutani and Mika Sato-Ilic (Eds.)
Innovations in Intelligent Machines-1, 2007
ISBN 978-3-540-72695-1

Vol. 71. Norio Baba, Lakhmi C. Jain and Hisashi Handa (Eds.)
Advanced Intelligent Paradigms in Computer Games, 2007
ISBN 978-3-540-72704-0

Vol. 72. Raymond S.T. Lee and Vincenzo Loia (Eds.)
Computation Intelligence for Agent-based Systems, 2007
ISBN 978-3-540-73175-7

Vol. 73. Petra Perner (Ed.)
Case-Based Reasoning on Images and Signals, 2008
ISBN 978-3-540-73178-8

Vol. 74. Robert Schaefer
Foundation of Global Genetic Optimization, 2007
ISBN 978-3-540-73191-7

Vol. 75. Crina Grosan, Ajith Abraham and Hisao Ishibuchi (Eds.)
Hybrid Evolutionary Algorithms, 2007
ISBN 978-3-540-73296-9

Vol. 76. Subhas Chandra Mukhopadhyay and Gourab Sen Gupta (Eds.)
Autonomous Robots and Agents, 2007
ISBN 978-3-540-73423-9

Vol. 77. Barbara Hammer and Pascal Hitzler (Eds.)
Perspectives of Neural-Symbolic Integration, 2007
ISBN 978-3-540-73953-1

Vol. 78. Costin Badica and Marcin Paprzycki (Eds.)
Intelligent and Distributed Computing, 2008
ISBN 978-3-540-74929-5

Vol. 79. Xing Cai and T.-C. Jim Yeh (Eds.)
Quantitative Information Fusion for Hydrological Sciences, 2008
ISBN 978-3-540-75383-4

Vol. 80. Joachim Diederich
Rule Extraction from Support Vector Machines, 2008
ISBN 978-3-540-75389-6

Vol. 81. K. Sridharan
Robotic Exploration and Landmark Determination, 2008
ISBN 978-3-540-75393-3

Vol. 82. Ajith Abraham, Crina Grosan and Witold Pedrycz (Eds.)
Engineering Evolutionary Intelligent Systems, 2008
ISBN 978-3-540-75395-7

Vol. 83. Bhanu Prasad and S.R.M. Prasanna (Eds.)
Speech, Audio, Image and Biomedical Signal Processing using Neural Networks, 2008
ISBN 978-3-540-75397-1

Vol. 84. Marek R. Ogiela and Ryszard Tadeusiewicz
Modern Computational Intelligence Methods for the Interpretation of Medical Images, 2008
ISBN 978-3-540-75399-5

Vol. 85. Arpad Kelemen, Ajith Abraham and Yulan Liang (Eds.)
Computational Intelligence in Medical Informatics, 2008
ISBN 978-3-540-75766-5

Zbigniew Les
Magdalena Les

Shape Understanding System

The First Steps toward the Visual Thinking Machines

With 330 Figures

Springer

Professor Zbigniew Les
The Queen Jadwiga Research
Institute of Understanding
P.O. Box 654
Toorak, Victoria 3142
Australia
zles@qjfpl.org

Magdalena Les
The Queen Jadwiga Research
Institute of Understanding
P.O. Box 654
Toorak, Victoria 3142
Australia
zles@qifpl.org

ISBN 978-3-540-75768-9 e-ISBN 978-3-540-75769-6

Studies in Computational Intelligence ISSN 1860-949X

Library of Congress Control Number: 2007938883

Cover design: Deblik, Berlin, Germany

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

This book is dedicated to our Patron St. Jadwiga Queen of Poland

Contents

Preface

This book presents the results of the research in one of the most complex and difficult areas of research such as research in the areas of thinking and understanding. This research that is carried out in the newly founded Queen Jadwiga Research Institute of Understanding www.qjfpl.org/QJRIU/Eng/Eng_QJRIU_PO_O.htm is focused on research on the problem of visual understanding and visual thinking. Visual understanding is part of the general understanding problem and it is not possible to carry out the research in visual understanding without reference to the nonvisual understanding problems. Understanding appears as the result of the thinking processes, and doing research in the area of understanding there is a need to include thinking process as one of the research problems that should be solved in the context of understanding investigations. According to our knowledge, this book is the first attempt to investigate the complexity of the visual thinking problems in the context of building the thinking machine. The aim of our research is to build the machine that can have capabilities to solve visual problems during thinking process. We are aware how complex this problem is and we are aware that the results of our research are only the first steps in building the thinking machine being able to solve complex visual problems. However, we believe that the results of our research will pave the way into the new way of thinking about designing the thinking machines and, especially, will supply the new scientific arguments about our human nature.

Until now, the problem of understanding and thinking was the topic of research in the area of philosophy, psychology, or cognitive science. Philosophical investigations of many thinkers such as Plato, Aristotle, Locke, Berkeley, or contemporary thinkers contributed into progress of understanding and thinking processes. Although there are some research on this topic in the area of artificial intelligence (AI), researchers in AI do not pay sufficient attention into understanding/thinking problems. It was probably for that reason that they tried to tailor the problem of what they called "artificial intelligence" to the abilities of the existing computing machines.

Another reason was that until now, there was no proper representation that could capture visual aspects of the world and represent them in the form that could be compatible with other representations of the nonvisual knowledge.

When existing systems are built based on the results of the scientific discoveries in the domain of psychology, cognitive science, computer science, or AI, our approach is based on the results of philosophical investigations of such thinkers as Aristotle, Locke, or Berkeley. In Chap. 1 of this book, the brief description of the results of main philosophical investigations concerning thinking and understanding is presented. In this chapter, shape that is regarded as the main perceptual category of thinking process and the important visual feature of the perceived world is briefly described. In the following sections of this chapter, the different problems connected with understanding are briefly presented. The relation between understanding and thinking is discussed in the following sections of this chapter. The last section includes the short description of the shape understanding system (SUS). In this book, the problems connected with the implementation issues of the SUS are not presented. The reason for that is that the theoretical issues connected with thinking and understanding are very complex, and inclusion of the extended description of implementation issues could cause that the contents of this book would be less understandable. In Chap. 2, concepts of shape classes that are understood as the basic perceptual categories are explained. The classes are represented by their symbolic names. Each class is related to each other and based on these classes there is relatively easy to establish the "perceptual similarity" among perceived objects. In Chap. 3, the description of the reasoning process that leads to assigning the perceived object to one of the shape classes is given. Each class possesses its characteristic reasoning process. The result of the reasoning process is the assignment of the examined object to one of the shape classes represented by the symbolic names. The symbolic name is used to find the visual concept and next to assign the perceived object into one of the ontological categories. Ontological categories are part of the new hierarchical categorical structure of the SUS. The new hierarchical categorical structure is explained in Chap. 4. The categorical chains that represent the categories of visual objects and knowledge categories are applied to interpret the perceived object as the symbol, the letter, or the real world object. In Chap. 5, examples of the visual reasoning processes that can be considered as the thinking process are presented. The thinking process is regarded as the continuous computational activity that is triggered by perception of a new object, by perception of an "inner object," or by the task given by the user. Thinking can lead to solving a problem where there is only one solution (e.g., the visual intelligence test) or solving a problem where there are many possible solutions (e.g., designing the tools). In this book, the focus is on thinking that leads to solving a problem that has only one solution.

We are aware that this book could be written in a different way where some issues could be explained in more details or presented in the different ways. We would like to explain that this book was written in very "difficult" conditions. During the most crucial part of writing of this book, we were notoriously expelled from our own flat where most work connected with preparation of this book was carried out. We think that for most of the readers, it would be difficult to understand how it could happen that in twenty-first century someone could be expelled from his home. We believe that it could happen only in country such as Australia where for more than 15 years, we Polish scientists are subjected to psychological terror from some Australian people and institution. The details of our persecution in Australia are described on our Web site (www.qjfpl.org\Przesladowanie\Eng\). We would like to take this opportunity and ask the Australian Government in Canberra to take responsibility for all damage that we suffered from Australian people and institutions.

1. Thinking, Visual Thinking, and Shape Understanding

1.1. Introduction

Thinking is the process that is connected with mental activities of our brain. We often say "I think that…" to indicate that our judgment is the result of thinking process. In the past thinking was highly praised and in some cultures thinkers (philosophers) had a very big influence on the development of culture and life of people. Philosophers played an important role in the development of the philosophical thinking and understanding of the real-world phenomena. Cartesian statement "cogito ergo sum" refers to thinking and it indicates the ontological status of the thinking process. Today scientists from different areas of sciences try to replace the philosophical investigation by speculative theoretical construction of scientific theories. Also philosophers of science tried to reduce philosophical thinking into the speculation about scientific description of the world. However, there are still questions that could not be answered or even formulated in the term of the today's scientific theoretical concepts, models, or theories. This book is an attempt to show that it is possible to build the thinking and understanding machines. When there is possibility for replacement of materialistically oriented "thinkers" by thinking machines, the metaphysical problems could be only solved by thinkers who are able to understand the essence of our human nature.

Understanding appears as the result of the thinking process. Understanding is a psychological process related to an abstract or physical object, such as a person, a situation, or a message, whereby one is able to think about it and use concepts to deal adequately with that object. In order to understand and solve a problem there is a need to engage thinking process. However, in some cases thinking does not lead to understanding. Understanding means knowing what is meant or intended by an utterance, a gesture, or a situation. Using an operational or behavioral definition of understanding, we can say that somebody who reacts appropriately to Y understands Y. For example, I understand

Z. Les and M. Les: *Thinking, Visual Thinking, and Shape Understanding*, Studies in Computational Intelligence (SCI) **86**, 1–45 (2008)
www.springerlink.com

English if I correctly obey commands given in that language. This approach, however, may not provide an adequate definition. A computer can easily be programed to react appropriately to simple commands.

Understanding is closely related to cognition and in many cases both terms have very similar meaning. For example, in the cognitive model the process of introverted thinking is thought to represent understanding through cause and effect relationships or correlations. One can construct a model of a system by observing correlations between all the relevant properties.

Understanding is often thought of as a special kind of seeing [1]. In common language very often instead of statement "I understand" people say "I see." Also, thinkers pointed out into connection between seeing and understanding. For example, Plato described the grasping of the forms or ideas as a kind of vision – our mental eye (onus, reason) [2]. The eye of the soul is endowed with intellectual intuition and can see an idea, an essence, and an object that belongs to the intelligible world. Once we have managed to see it, to grasp it, we know this essence and we can see it in the light of truth.

One aspect of thinking and understanding is the acquisition and utilization of knowledge in order to explain the world and to perform complex tasks. Another one is connected with application of knowledge in the problem-solving process. Some people believe that knowledge is the simple awareness of bits of information and understanding is the awareness of the connectedness of this information. However, it is thinking during understanding process which allows knowledge to be put in use. In order to be able to effectively utilize the knowledge during the solving of difficult problems, the subject needs to have well-developed problem-solving skills. Problem-solving skills comprise wide range of competencies such as the capacity to understand problems situated in novel settings, to identify relevant information, to represent possible alternatives or solutions, to develop solution strategies, and to solve problems and communicate the obtained results.

In the context of understanding of the visual forms, there is a need to distinguish between visual understanding that deals with understanding of the visual forms and nonvisual understanding that refers to understanding of nonvisual forms. Mental processes that are connected with visual understanding are called visual thinking. Visual understanding and visual thinking play an important role in understanding of the world's objects and phenomenon's. However, nonvisual thinking consists of the substantial part of the thinking process. The term *visual thinking* became popular after publishing the book by R. Arnheim titled "Visual thinking" [3]. In this book Arnheim tried to compare the process of reasoning that is performed by scientist with reasoning that is characteristic in the artistic creative act. Visual thinking, that is present in the artistic creative act, is the process of mental operations on the visual concepts in order to obtain a new form of the symbolic or visual representation. During the visual thinking process visual concepts are often used to form "mental image" that allows better mental representation of knowledge.

Visual understanding requires certain abilities of the subject to perform complex mental transformations that are part of visual imagination. In the visual understanding process visual imagination plays a significant role. For example, Copernicus needed a remarkable visual imagination, which let him apply a model from very different area to describe the situation he saw. He succeeded in seeing the intricate gyrations of the planets as simple movements of these heavenly bodies. Visual understanding requires knowledge that is encoded in the brain and is the result of the learning process as well as the visual thinking processes. The visual thinking process is based on the visual reasoning that starts with perceiving of the visual object and uses abstraction and generalization in solving complex visual problems. Abstraction that is present in perception (understanding) of the visual object makes it possible to grasp of structural features rather than indiscriminate recording of detail. Also visual knowledge acquired in the past helps not only in interpretation of the object appearing in the visual field but also places the perceived object in the system of things constituting our total view of the world.

When the existing knowledge-based systems are built based on the results of the scientific discoveries in the domains of psychology, cognitive science, computer science, or AI, our approach, presented in this book, is based on the results of philosophical investigations of such thinkers as Locke, Berkeley, or Kant. In this book, brief description of philosophical investigations of topics connected with understanding and thinking is presented. The shape that is the main perceptual category of thinking process and the important visual feature of the perceived world is briefly described in Sect. 1.2. In Sect. 1.3, the different problems connected with understanding investigated by philosophers such as Locke or Berkeley are described. The relation between understanding and thinking is discussed in Sect. 1.4. Section 1.5 includes the short description of the shape understanding system (SUS). In this book, the problems connected with the implementation issues of the SUS are not presented. The reason is that the theoretical issues, connected with thinking and understanding, are very complex, and the attempt to describe the implementation problems could, instead of clarifying things, make them less understandable in the context of the material presented in this book.

Chapter 2 presents the concept of shape classes that are regarded as the basic perceptual categories. Shape classes are represented by their symbolic names. Each class is related to each other and, based on these classes, it is relatively easy to establish the "perceptual similarity" among perceived objects.

Chapter 3 presents the description of the reasoning process that leads to assigning the perceived object to one of the shape classes. Assignment of an object to one of the general classes is based on the specific reasoning process. In other words, each general class is characterized by its specific reasoning process. As the result of the reasoning process, an examined object is assigned to one of shape classes where each class is represented by its symbolic name. The symbolic name is used to find the visual concept and next to assign a perceived object into one of the ontological categories. The ontological categories are described in Chap. 4.

Chapter 4 presents the new hierarchical categorical structures of the different categories of visual objects. The categorical chains that represent the visual categories are applied to interpret the perceived object as a member of one of the ontological categories: a figure, a sign, a letter, or a real-world object category.

Chapter 5 gives examples of the visual reasoning processes that can be considered as the special kind of the thinking processes. The thinking process is regarded as the continuous computational activity that is triggered by perception of a new object, by perception an "inner object" or by a task given by the user. Thinking can lead to solving a problem where there is only one solution (e.g., the visual intelligence test) or solving a problem where there are many possible solutions (e.g., designing the tools). In this book, the focus is on thinking that leads to solving a problem that has only one solution.

1.2. Shape and Form

Shape understanding method presented in this book is based on the concept of shape classes. The shape classes (described in Chap. 2) are regarded as the basic shape categories, the main ingredients of thinking process. During thinking process, the perceived object (phantom) is transformed into the digital form (visual object) and next into the symbolic form (symbolic name) by fitting it to one of the shape categories (shape classes). The visual object that exists in the mind (exemplar) can be transformed into another exemplar or into the symbolic form during the visual thinking process. The main ingredients of the visual thinking process are shape categories.

Shape categories refer to shape as the visual aspect of the perceived object. In existing literature the term *shape* is differently defined and understood. External shape or appearance is often called form. The term form has been used in a number of ways throughout the history of philosophy and aesthetics. In contrast to the particulars that are finite and subject to change, Plato's term eidos "eternal forms" identified the permanent reality that makes a thing what it is. Plato's "eternal forms" are the immutable essence that can only be "received" or "imitated" by material, or sensible things and are of a higher reality than material objects. Aristotle distinguished between matter and form and argued that every sensible object consists of both matter and form. The matter was the undifferentiated primal element and the development of particular things from this germinal matter consists in differentiation, the acquiring of the particular forms of which the knowable universe consists. According to Aquinas the concept of form includes "accidental form," a quality of a thing that is not determined by its essence. These "sensible forms" can be distinguished from matter by sense-perception. He distinguished space and time as the two forms of sensibility and 12 basic categories that act as structural elements for human understanding. When the form of philosophical thinkers can be, only to some degree, identified with the term "form" that is used in the area of the visual art, there is no doubts that these terms have one common source. In art form is related to shape and in some context they are used as synonyms.

Shape is a very important concept that is differently defined and understood across many different disciplines. Mathematicians often use the visible shapes and reason about them although they are thinking not of these but of the ideas which they resemble. Shapes share several universal properties that are applicable to the definition of all shapes. The first of these properties is geometry, which concerns the relative placement of points within the shape or its embedding environment. The other property is topology. It concerns the adjacency relationship between the elements of shape. For example, a rectangle made of rubber could be inflated into a circle. Its geometry would change, but not its topology; the adjacency relationships between its edges, faces and vertices (its boundary elements), and between points that make up its interior, remains the same as before. Topology and geometry nevertheless are not completely independent of each other; at the limits, a change in geometry may cause a change in topology. Based on the geometrical permissions, Leyton [4] developed a theory that claims that all shapes are basically circles, which changed form as a result of various deformations caused by external forces, e.g., pushing. Geometry and topology, the components of shape information, has been the subject of study by such mathematicians as Euclid, Pythagoras, Archimedes, Euler, Mobus, Polya, and Lacatos. Most of the technical papers that deal with shape stress the importance of geometrical properties of shape. The perceptual aspect of the visual forms present in visual psychology is absent in conventional geometrical theories of shape. For example, in statistics shape is often treated as an independent field of research (see e.g., [5–7]). Kendal [8] defines shape as: "shape is all geometrical information that remains when location, scale, and rotational effects are filtered out from an object." Shape is described by locating a finite number of points on each specimen which are called landmarks. In statistics shape is analyzed by applying variety of mathematical and statistical methods. For example, in order to perform the shape analysis ratio of distances between landmarks or angles are selected, and then the data obtained is submitted to a multivariate analysis.

Shape is a subject of study in many disciplines. For example, in biology and medicine the subject of study is how shape changes during growth; how shape changes during evolution; how shape is related to size; how shape is affected by disease; how to discriminate and classify using shape; and how to describe shape variability. In statistics, statistical shape analysis is applied to analyzing of hand shapes [9], the resistors in [10] or the mitochondrial outlines [11]. The statistical method, which was developed by Kendal, was applied in many different domains such as biology: analysis of the mouse vertebrae to assess whether there is a difference in size and shape between groups of selected specimen; in image analysis: postcode recognition; in archaeology: alignments of standing stones; in geology: microfossils agriculture; or in robotics: harvesting of mushrooms [6–8].

Visual objects are often characterized by their shapes. Shapes of visual objects can have different "details" that can reveal the useful information about the object or, alternatively, those varying "details" can be treated as irregularities recognized as noise that comes from a nonperfect processing stage. The visual object that possesses the intricate details is called fractal. Magnifying the nonlinear shapes such as fractals allows the intricate details to be still visible. Fractals are not relegated exclusively to the realm of mathematics. If the definition is broadened, such objects can be found virtually everywhere in the natural world. The difference is that "natural" fractals are randomly, statistically, or stochastically rather than exactly scale symmetric. The rough shape revealed at one-length scale bears only an approximate resemblance to that at another, but the length scale being used is not apparent just by looking at the shape. Moreover, there are both upper and lower limits to the size range over which the fractals in nature are indeed fractals. Above and below that range, the shapes are either rough (but not self-similar) or smooth – in other words, conventionally Euclidean. The visual objects that are the result of physical processes such as soot aggregation in chimneys, zinc deposition in electrolytic cells, diffusion of gas bubbles through viscous liquids, and electrical discharge in air, possess shapes that are fractals [12]. Also organic forms often possess the characteristic intricate shapes that are fractals [13]. Figure 1.1 shows visual objects that are result of different physical processes and are regarded as fractals.

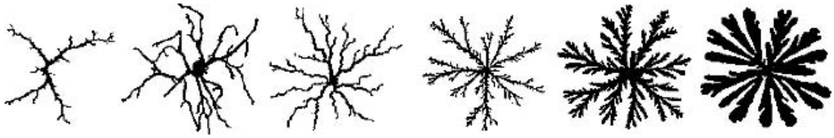


Fig. 1.1. Objects with irregularities classified as fractals

In many interpretations shape is not distinguished from an object. In engineering, the object is divided based on the shape that is considered as an object itself. Many different types of shape are distinguished, for example, rigid shapes whose proportions, angles, and sizes are independent of shape location and orientation in space, and nonrigid shapes, such as gases and liquids which are the subject of study of thermodynamics.

Shape not only determines how an object looks, but also forms the basis for many of its other properties. Webster defines shape as "that quality of an object which depends on the relative position of all points composing its outline or external surface." This definition emphasizes the fact that we are aware of shapes through outlines and surfaces of objects, both of which may be visually perceived. It also makes the distinction between the two-dimensional outline and the three-dimensional surface. The term shape often refers to the geometry of an object's physical surface [14]. For example, Marr treats shape as one of the forms of an object representation or a "special visual" feature of an object. A special class of shapes is shape contours which have the two-dimensional base and can yield information about the three-dimensional shape. Perception research lays emphasis on the use of contours in decomposing objects into their parts, especially on describing rules for detection of part boundaries, e.g., based on notions as "concavities" of concave regions or the "minimal rule" (see [15], [16]). The characteristic features of the contour seem to play an important role also at the conceptual level in the process of concept formation.

Shape is also described as a silhouette of the object (e.g., obtained by illuminating the object by an infinitely distant light source). Silhouettes contain rich information about shapes of objects that can be used for recognition and classification. Silhouette contours contain detailed information about object's shape. In many cases it is possible, based on a silhouette to determine the parts that compose shape, identify their local orientation and rough aspect ratio, and detect convex and concave sections of the boundaries. When a silhouette is sufficiently detailed people can readily identify the object, or judge its similarity to other shapes. Computer vision systems use similar information to classify objects. Silhouettes may be available to these systems as a result of segmentation. In either case, properties of silhouettes extracted automatically and reliably provide (possibly in con junction with additional properties such as color and texture) a powerful cue for recognition. The computer vision literature contains numerous examples for the use of properties extracted from silhouettes. Various studies utilize parts and skeleton structures to determine shape category. Local properties are used for registration and recognition, as well as similarity judgment. In addition, various categorization methods use "qualitative descriptions" of shape boundaries. Existing methods of shape analysis are mostly concerned with recognition of shape as a binary image representing the extent of the object [17–26].

In dealing with recognition, shape is often interpreted in terms of an object where the object is represented by structural descriptions (see [27], [28]). In this view, objects (shape) are divided into parts and represented in terms of their parts and the relations between parts. This view has advantages over alternatives, such as a template and feature models, on computational and perceptual grounds ([15], [16], [29–31]). Parts-based representations allow for recognition that is robust in the presence of occlusion, movement, growth, and deletion of portions of an object, and play an important role in theories of objects categorization and classification. There is a strong evidence for parts-based representations in human vision (see e.g., [32], [33]). Authors in [34] provide strong evidence that contours are psychologically segmented into visual parts at negative curvature minima. However, computation of negative curvature minima, as well as other extreme points is not robust in real digital images. Although remarkable progress has been made on this matter, the robust computation of extreme points in real digital images is an open problem. Since contours of objects in digital images are distorted by digitization noise and segmentation errors, it is desirable to neglect the distortions while at the same time preserving the perceptual appearance at a level sufficient for object recognition. One solution to this problem is to apply the evolution of planar curves in the scale space [35–37]. It was proven in [36] that an embedded plane curve, when evolving according to the heat equation, converges to a convex plane curve.

Perception of shape is to see an object. The perception of shape consists in the application of form categories, which can be called visual concepts because of their simplicity and generality. In a typical life situation, a person concentrates on some selected areas and items or on some overall features while the structure of the remainder is sketchy and loose. Under such circumstances, shape perception operates partially. When the angle changes at which the object is perceived shape is affected by transformations, which are generally more complex than those provided by Euclidean geometry, that is, translation, rotation, or reflection in space. The objects of perception are not necessarily rigid; they move, bend, twist, turn, swell, shrink light up, or change their color. Constancy of shape does result when the various aspects of an object can be seen as deviations from, or distortions of a simpler shape. Laws of association are often used to support the thesis that perceptual laws exists. Laws of association say that items will become connected when they have frequently appeared together. Perception of shape is always to perform an abstraction because seeing consists in the grasping of structural features rather than in the indiscriminate recording of detail. If a percept is a categorical shape rather than a mechanically faithful recording of a particular stimulus, then its trace in memory must be equally generic. In the earliest classical theory, a perception of the shape was thought to consist of the memories of the eye movements that would have to be made in order to bring each point on its contour to the centre of vision. The psychological aspects of visual perception of shape ([38], [39]) led to the development of a dynamic shape model where visual perception is performed on several scales of resolution. In this model, any shape can be embedded in a morphogenetic sequence based on the solution of the partial differential equation that describes the evolution of the shape through multiple resolutions.

In this book shape is regarded as the basic perceptual category to which examined object is fitted. The perceptual category that refers to the shape classes (described in Chap. 2) is an element of the visual concept. Visual concept is the main ingredient of the visual thinking process and is described in Chap. 5. Shape as the basic perceptual category is the main ingredient of thinking process that leads to understanding an object as a part of the world.

1.3. Understanding

This section is not intended as a survey of literature on the vast topic concerning understanding, but rather as presentation of the point of view of selected thinkers on this topic and a discussion of some aspects of understanding considered to have implication for material presented in other chapters.

It is understanding that sets man above the rest of sensible beings, and gives him all the advantage and dominion which he has over them. Understanding appears as the result of the thinking process and can be the object of the scientific inquires. Locke [40] has no doubt that understanding can be studied like anything else: “we can observe its object and the ways in which it operates upon them” he wrote. Understanding that is often thought of as cognition involves processes such as learning, problem solving, perception, intuition, and reasoning, and requires abilities such as intelligence. Understanding that is based on knowledge is often connected with interpretation or disclosing meaning of the language and the concept is the key element of understanding process.

Understanding and thought were topics of many philosophical thinkers such as Plato, Aristotle, Locke, Berkeley, Laibnitz or Gadamer (see. e.g., [1] [40–45]) and were regarded in the context of the origins of human knowledge. The traditional Augustinian theory explained the cognition as the result of a divine illumination and was based on innate ideas. This Neo-Platonic view was that an essence of created things was “participations” of the divine essence. God, in contemplating them, does nothing but contemplate Himself. According to Aquinas, the direct object of human intellectual knowledge is the form abstracted from matter, which is the principle of individuation, and known through the universal concept. The senses apprehend the individual thing but the mind apprehends it only indirectly, as represented in an image or phantasm. There is no intellectual intuition of the individual thing as such. Scotus discarded the traditional Augustinian–Franciscan theory of a special divine illumination and held, with Aquinas, that Aristotelian doctrine of the abstraction of the universal can explain the genesis of human knowledge without it being necessary to invoke either innate ideas or a special divine illumination.

The fundamental principles of Locke’s thought concerning understanding are presented in *An Essay Concerning Human Understanding* [40]. This essay was the culmination of 20 years of Lock’s reflection on the origins of human knowledge. The *Essay* is divided into four books; the first is a polemic against the doctrine of innate principles and ideas. The second deals with ideas, the third with words, and the fourth with knowledge. Lock did not distinguished between cognition and understanding. According to Locke, what we know is always properly understood as the relation between ideas. He devoted much of the Essay to an extended argument that all of our ideas – simple or complex – are ultimately derived from experience. The consequence of this empiricist approach is that our knowledge is severely limited in its scope and certainty. Our knowledge of material substances, for example, depends heavily on the secondary qualities by reference to which we name them, while their real inner natures derive from the primary qualities of their insensible parts.

Hume’s *An Enquiry Concerning Human Understanding* appeared in 1748. The central themes of Hume’s book [42] are that very little of what we think we know can actually be derived from any idea that there are actual necessary connections between observed phenomena. We assume that certain things are connected just because they commonly occur together, but a genuine knowledge of any connection is mere habit of thought. So, a severe skepticism is the only rational view of the world. Hume’s investigations into human understanding lead him to doubts. He asks on what grounds we base our judgments and investigates their rational justification. Finding certain inconsistencies in our normal procedures, for instance, that our belief in necessary connection is not rationally justified, Hume is led to a kind of consequent doubt of our mental faculties.

Descartes claimed that "natural light" of understanding is a faculty created by God [46]. We come to know not only created eternal truths but uncreated truth: that God exists, that God is not a deceiver, that God is immutable, a necessary being, *causa sui*. But God is not subject to the limits of our understanding, and we only have access to these uncreated truths through a faculty given to us by Him. If our understanding seeks some unconditional verification of God's existence and truthfulness, through means outside the scope of God's creative will, it seeks in vain. Descartes initiates a critique of the understanding itself. It is immediately aimed at "eternal truths," that is, mathematical truths which for Descartes are properly truths of the understanding.

According to Kant understanding as a one of the higher faculties of knowledge, in general, can be defined as the faculty of rules. Ideas, as Kant argues in the Transcendental Dialectic, are a priori concepts whose source lies in pure reason alone. Their only legitimate theoretical use is to regulate the understanding's cognition of objects: reason sets down the conditions under which the understanding's activity will have achieved its ideal completion in the systematic interconnection of its cognitions, i.e., in an ultimate science. Reason thereby offers the understanding of a rule against which any actually achieved system of science must be measured [45]. Because human finitude makes it impossible in principle for any actual system to attain the ideal maximum, reason also spurs the understanding on towards ever new discoveries and reorganizations.

Natorp claims that the directedness towards a goal is implied by "method" that illuminates one of two senses in which his philosophy is idealistic, namely that science (and the other activities of culture) are guided by regulative ideas or limit-concepts. Given an object of scientific cognition, the cognition is conceived as a process never "definitively concluded," but rather, "every true concept is a new question, none is a final answer" [47]. Natorp comments: "Just this is the meaning of the thing in itself as *X*: the infinite task." In other words, the thing in itself is the ideal of an object exhaustively determined by concepts, that is, completely known. As with Kant, however, our cognitive finitude means that the process of conceptual determination can only approach this ideal asymptotically. This pursuit of total determination, what Natorp calls "method," is the pursuit of science. The hypothesis as law or groundwork is for Natorp the transcendental foundation for scientific experience, i.e., for the activity of legislating and thus rationally understanding the phenomena.

Hermeneutics started to emphasize the role of language in understanding. In hermeneutics understanding is the inversion of a speech act, during which the thought which was the basis of the speech must become conscious. Every utterance has a dual relationship to the totality of the language and to the whole thought of its originator, then understanding also consists of the two moments, of understanding the utterance as derived from language, and as a fact in the thinker. Hermeneutics is the art of understanding particularly the written discourse of another person correctly. A central principle of Gadamer's hermeneutics is that language conditions all understanding [48]. The phenomenon of understanding shows the universality of human linguistically as a limitless medium which carries everything within it. Not only the "culture" which has been handed down to us through language, but absolutely everything because everything is included in the realm of understanding. Theorists of language focus on the Mind/Language connection when they consider understanding to be the cornerstone concept, holding, for instance, that an account of meaning for a given language is simply an account of what constitutes the ability to understand it. Many philosophers such as Locke or Frege have been attracted to the view that understanding is a matter of associating the correct ideas or concepts with words. Others have equated understanding with knowing the requirements for accurate use of words and sentences. Wittgenstein found the key to understanding in one's ability to discern the communicative goals of speakers and writers, or more directly in one's ability to "pass" linguistically, without censure. Nietzsche puts forward the hypothesis that scientific concepts are chains of metaphors hardened into accepted truths [49]. On this account, metaphor begins when a nerve stimulus is copied as an image, which is then imitated in sound, giving rise, when repeated, to the word, which becomes a concept when the word is used to designate multiple instances of singular events. Conceptual metaphors are thus lies because they equate unequal things, just as the chain of metaphors moves from one level to another. Hegel's problem with the repetition of the "this" and the "now" is thus expanded to include the repetition of instances across discontinuous gaps between kinds and levels of things. Today's scientists, however, found the limitation of the linguistic theories. The power of a living language is exceeded by the power of our thinking. If we compare the power of a living language with the logical language then we will find that logic is even poorer. Therefore it seems to be impossible to guarantee a one-to-one mapping of problems and a model using a mathematical or logical language. It can be shown that it is very often extremely difficult to appropriately assign semantic contents to logical symbols.

Understanding is often described by cognitive activities of our brain that is called cognition. In the next sections the cognitive processes such as visual perception, knowledge acquisition and storing, learning, reasoning, problem solving, or thinking will be briefly described.

1.3.1. Cognition

By "cognitive" we often mean all mental operations involved in the receiving, storing and processing of information. Cognition can be seen as an activity which involves different cognitive processes such as: attention, creativity, memory, perception, problem solving, thinking, and the use of language. The essence of cognition is judgment, in which a certain object is distinguished from other objects and is characterized by some concept or concepts. The nature of cognition and the relationship between the knowing mind and external reality have been exhaustively discussed by philosophers since antiquity. Cognition and its development have been subjected to many viewpoints and interpretations. The psychologist is concerned with the cognitive process as it affects learning and behavior. There are two broad approaches to contemporary cognitive theory. The information-processing approach attempts to understand human thought and reasoning processes by comparing the mind to a sophisticated computer system that is designed to acquire, process, store, and use information according to various programs. The second approach is based on the work of Swiss psychologist J. Piaget who viewed cognitive adaptation in terms of two basic processes: assimilation and accommodation. Assimilation is the process whereby an individual interprets reality in terms of his own internal model of the world based on previous experience; whereas, accommodation is the process of changing that model by developing the mechanisms to adjust to reality. Piaget believed that representational thought does not originate in a social language but rather in unique symbols that serve as a foundation for a later, acquired language.

Arnhaim extended the meaning of the term "cognitive" and "cognition" to include perception. The general cognitive problem is that the perceived object presents itself in the context and is modulated by that context. According to Arnheim cognitive process which produces the so-called constancies is of a very high order of intelligence since it must evaluate any particular entity in relation to an intricate context, and that this feat is performed as an integral part of ongoing perception.

According to most cognitive theories, information picked up by the senses is analyzed, stored, recoded, and subsequently used in various ways; these activities are called information processes. The cognitive processes are the subject of the many research areas such as cognitive psychology or cognitive science. Cognition is the process involved in knowing, or the act of knowing, which in its completeness includes perception and judgment. Cognition includes every mental process that can be described as an experience of knowing as distinguished from an experience of feeling or of willing.

It includes, in short, all processes of consciousness by which knowledge is built up, including perceiving, recognizing, conceiving, and reasoning.

1.3.2. Visual Perception

Visual perception was often though as the "introduction" to understanding of the real-world phenomena. Visual perception was the subject of study from ancient times. One school of thought, called atomism, started with Aristotle. The atomistic view assumed a basic vocabulary of elementary sensations from which our perceptions are made. Attneave's work [50] that investigates the significance of corners for perception, initiated further research on the topic of curve partitioning, and led to a vocabulary-based scheme made up of primitive shape descriptors called "codons" for describing two-dimensional plane curves [51]. The concept of a basic vocabulary was the subject of many controversies and Locke, in the theory of psychophysical dualism, tried to point out that perception is made up of sensations (input) and reflections. Wittgenstein in his work underlined the role of knowledge, which a particular context transforms and determines sensations and percepts. Also, Arnheim and Rock [52] suggested that perception is intelligent in that it is based on operations similar to those that characterize thought. However, due to the dependence of perception on sensory information there is a difference between "higher" cognitive functions such as imagination or thinking.

The three major perceptual theories, namely, inference, Gestalt and Gibson attempt to explain perception. The inference theory, associated with the empiricist view, argues that knowledge is acquired solely by sensory experience and association ideas. The mind at birth is a *tabula rasa* upon which experience records sensations. Helmholtz postulated the existence of "primary" percepts. Helmholtz claims that the primary percept contains all the distortions of projection but judgment intervenes and corrects them. Helmholtz assumed that these corrections are based mainly on knowledge previously acquired. Both Berkeley and Helmholtz later argued that we learn to interpret percepts through a process of association. Helmholtz described the process as one of unconscious inferences, such that sensations of the senses are tokens for our consciousness, it being left to our intelligence to learn how to comprehend their meaning.

The Gestalt view originated with Descartes and Kant for whom the mind was far from being a *tabula rasa* (see e.g., [53], [54]). Kant argues that "the mind imposes its own internal conception of space and time upon the sensory information it receives." Gestalt theory refers to the laws of association:

items will become connected when they have frequently appeared together; or when they resemble each other. These laws assume that relations connect piece by piece and that these pieces remain unchanged by being tied together. The simplest among the rules that govern these relations is the rule of similarity: things that resemble each other are tied together in vision. Homogeneity is the simplest product of perceptual relation. When a sprinkling of items is seen on a sufficiently different background and sufficiently distant from the next sprinkling it will be seen as a unit. Similarity of location provides the bond. The Gestaltist of the twentieth century believed in holistic perceptual organization preordained by "given laws that govern unit formation and the emergence of a figure on a background." Visual form is the most important property of a configuration. As opposed to the Gestalt school, Hebb argues that a visual form is not perceived as a whole but consists of parts [55]. The organization and mutual spatial relations of parts must be learned for successful recognition. This learning aspect of perception is the central point in Hebb's theory. However, as suggested in [56] recognition is not based on a single instantaneous impression. We look at objects in question and explore them with our eyes until we gather enough information to identify them.

The Gibson theory [57] that is characterized by the stimulus (sensed data) view claims that sensory input is enough to explain our perceptions. The theory seeks to associate percepts with physical stimuli. The first principle of his theory is that space is not a geometric or abstract entity, but a real visual one characterized by the forms that are in it. Gibson's theory is centered on perceiving real three-dimensional objects, not their two-dimensional projections. Gibson points out that the Gestalt school has been occupied with the study of two-dimensional projections of the three-dimensional world and that its dynamism is no more than the ambiguity of the interpretation of projected images. Marr [14] made significant contributions to the study of the human visual perception system and in his paradigm the focus of research was shifted from applications to topics corresponding to modules of the human visual system. Marr developed a primal sketch paradigm for early processing of visual information.

Perception points to a different notion of abstraction, a much more sophisticated cognitive operation. The concept is obtained based on the abstraction and generalization. According to Aristotle abstraction must be complemented with definition which is the determination of a concept by deriving it deductively from the higher genus and pinpointing it through its distinguishing attribute (differentia). Abstraction removes the more particular attributes of the more specific instances and arrives at the higher concepts. Higher concepts are poorer in content but broader in range. Aristotle not only established the universal as the indispensable condition of the individual thing's existence and as the very character of the perceivable object. He rejected the arbitrary choice of the attributes that can serve as the basis of generalization. The qualities an object shared with others of its kind were not an incidental similarity but the very essence of the object. What was general in an individual was the form impressed upon it by its genus. An object existed only to the extent of its essence since the being of the object was nothing but what had been impressed upon the amorphous raw material by its form – giving genus. The object's accidental properties were mere impurities, the inevitable contribution of the raw material. When a perceptual generalization is to be made, it can only be done by recognizing the common essence of the specimens. Shared accidentals cannot serve as the basis for a genus.

Perception can abstract objects from their context only because it grasps shape as organized structure, rather than recording it as a mosaic of elements. In more than one way, perceptual abstraction can differ from the kind described in traditional logic. Typically, it is not a matter of extracting common properties from a number of particular instances. Perception points to a different notion of abstraction, a much more sophisticated cognitive operation. Perception (understanding) of shape is always to perform an abstraction because seeing consists in the grasping of structural features rather than in the indiscriminate recording of detail. For example, assume that subject has learned to choose a circle rather than another figure. Understanding assumes that the subject transfers the result of learning to an ellipse. By doing this he shows himself capable of abstracting the features which rounded shapes have in common from those in which they differ. This requires the twofold ability to discover the crucial common qualities and to disallow the irrelevant ones. Understanding of the object requires abstractions of general relationships that possess the differentiating characters.

1.3.3. Visual Intelligence

Understanding requires certain abilities of subject to perform complex mental transformations. An ability of grasping of structural features by organizing spontaneously stimulus material, according to the simplest overall pattern adaptable to it is called visual intelligence. Arnheim claim that in perception we can trace the source of the intelligent behavior. A capacity essential to perception and intelligence is to be capable of the spontaneous grasp of pattern. Analogy problems are often used in intelligence tests because the cognitive operations displayed in visual perception when a person discovers analogies among patterns are intelligent behavior. Analogies are discovered by finding a basic similarity of character in the items that are compared and employing abstraction when dealing with visual pattern. The ability to solve the intelligence test is characteristic of the thinking process.

The possibility of building of intelligent machines became reality in the time when the first computer was built. Artificial intelligence (AI) is one of the areas of research that investigates problems connected with intelligence. The performance of designed intelligent system is compared to performance of the human being. The question that is often given is: "What is the basic difference between today's computer and an intelligent being?" especially in the context of visual problems. It is that the computer can be made to see but not to perceive. The existing visual systems that are based on the rule-based AI approach such as Acronym [58] are able to perform the very specific tasks. The agent-based technology is offering the new possibilities to built intelligent systems. The agent paradigm appears to be mutation of the object-oriented approach. An agent software abstraction intends to be more than a passive object with memory and behavior and can be seen as a kind of active object, autonomous, social and able to learn [59], [60]. Having control over its own behavior is the main issue distinguishing agents over objects. An object can invoke public accessible methods of any object. Once the method is invoked, corresponding actions are performed [61], [62]. In this sense, objects are not autonomous because they are totally dependent on each other for the execution of their actions. Autonomy is often praised as one of the most advantageous features of agent technology. Autonomy is supported both by the agent own experience and by the built-in knowledge used when constructing the agent for the particular environment in which it operates. Therefore, if agent actions are based completely on built-in knowledge, such that it needs to pay no attention to its precepts, then we say that the agent lacks autonomy.

1.3.4. Knowledge

Understanding requires knowledge. The concept is a key element of knowledge that is stored in our brain. The concept was often viewed in relation to the universal terms. In the Middle Ages the problem of universal terms or class names was the topic of the many tractates [41]. These universal terms were thought of as a hierarchical structure of the class names. Also important problem: the relation between concept and object it represents, was investigated by many philosophers. For example, for Kant [45] concepts when they relate to objects do so by means of feature which several things may have in common. Having a concept does not imply a relation to an object. Once an object is given, it can be thought about, but what allows it to be given in the first place is its relation to intuition. According to Kant intuitions are those representations by means of which objects are given to us whereas concepts are those representations by means of which we think about objects. The

German for intuition Anschauung, means "looking at" (without any connotations of special insight). The distinction between intuition and concept thus corresponds to the distinction between the particular and general. An intuition is a representation of one particular, individual thing, "a single object." A concept is inherently general: necessarily a concept can apply to more than one particular, since to apply a concept to an object is to say that it belongs to a kind of which there are or could be other instances.

Understanding of the visual object or phenomena is different from understanding a general concept, or abstract concepts such as mathematical objects. The understanding of the visual object can be called the visual understanding. In visual understanding visual concept and mental images play a key role. The problem of the visual concept is related to the problems of visual images and their role in the thinking process. The kind of "mental image" needed for thought is unlikely to be a complete, colorful, and faithful replica of some visible scene. For example, Berkeley insisted that generic mental images were inconceivable. There is, however, evidence that an artist makes a drawing of something he knows from memory and these things are called mental images. Mental images so-called eidetic images seem to be used as a target for active perception and they can serve as material for thought. The visual understanding needs to refer to the certain kind of mental images. There are problems in the relation between concept and mental image. In Plato's doctrine a relation between prototype and image was the "static coexistence" of the transcendental ideas and sensory appearance. Aristotle claimed that perception is a faculty in which we always perceive, in the particulars, kinds of thing general qualities rather than uniqueness. According to Kant, noumena are essentially unknowable, yet must be posited to account for phenomena, or things as they appear to be. By contrast with phenomena, the objects of experience, noumena are things-in-themselves, unconditioned by the categories of understanding. The development of the computer graphics shows how images can be "perceived," created, stored and manipulated to obtain required effect. The real-world images are generated by application of the mathematical model of the physical processes. The knowledge of the optical laws transformed into the computer program makes it possible to model the different physical phenomena that produce very realistic images. There is a belief that the appearance of objects in images may be understood by understanding the physics of objects and the imaging process. However, it is unlikely to explain the existence of the mental images by the computer graphic model.

There is a conversion from the one sensory data into another. For example, when we describe sensory events we convert from sensory representation to verbal representation and when we speak we convert from verbal representation to vocalization representation. Images evoked by words such as hat or flag can be reasonably concrete, whereas the solution of theoretical problems requires highly abstract configurations, represented by topological and often geometrical figures in mental space.

The psychological researchers treat the problem of role of knowledge in understanding process as the problem of knowledge representation. In psychology the term “knowledge representation” has a different meaning for cognitivists and behaviorists. Behaviorists reject the scientific validity of referring to hypothetical internal representations, whereas cognitivists stress the need to use representational states in the explanation of psychological behavior. The observation of patients with a damage of the brain allows to obtain a knowledge about mental processes and a way of knowledge representation and storing in the brain [63]. One of the conclusions from this research is that semantic information is multiply distributed and represented in the brain and is linked to the input modalities in order to create knowledge of the world. It seems to be more appropriate to use the neural networks as a model of knowledge representation in the brain. For example, associative memories respond by retrieving exactly one of the previously stored patterns, even though the stimulus or cue might be partly distorted or missing in part. In contrast with the mode of address-addressing, associative memories are content-addressable. The words in the memory are accessed based on the key vector and the entire mapping is distributed in the associative network [64].

The machine to be able to understand and think needs to have some mechanism that makes it possible to utilize knowledge during thinking process. In order to solve complex problems one needs both an appropriate knowledge representation and some mechanisms for manipulating that knowledge. Existing knowledge-based systems apply the different methods of knowledge representation. The knowledge representation is some chosen formalism for “things” we want to represent. There are two main important dimensions along which they can be characterized. At one extreme are purely syntactic systems, in which no concern is given to meaning of the knowledge. Such systems have simple, uniform rules for manipulating the representation. At the other extreme are purely semantic systems, in which there is no unified form.

We can distinguish structures in which knowledge can be represented: production rules, semantic nets, frames, conceptual dependency and scripts. The production rules belong to syntactic systems because they usually use only syntactic information to decide which rule to fire. Semantic nets are designed to capture semantic relationship among entities, and they are employed with a set of inference rules. Semantic networks offer a convenient mechanism to describe semantics, syntax and pragmatics in the study of language [65]. The use of network structures is not new in knowledge representation. There are two major types of networks that deal with imprecise information and thus perform reasoning under uncertainty: Bayesian [66] and Markov [67]. Frame systems are typically more highly structured than are semantic nets, and they contain a large set of specialized inference rules. Conceptual dependence representation can be thought of as instances of semantic nets but having a more powerful inference mechanisms that exploit specific knowledge about what they contain [68]. Script (very similar to frames) in which slots are chosen to represent the information is useful during reasoning about a given situation.

One of the methods of knowledge representation that refers to the thinking process is representation that is based on the concept of frame. A frame is simply a data structure that consists of expectation for a given situation. A frame can consist of objects and facts about a situation, or procedures on what to do when a given situation is encountered. To each frame several kinds of information are attached. Some of this information is about how to use the frame, some is about what one expects to happen next, and some is about what to do if these expectations are not confirmed. Collection of frames is linked together into frame system. The frame-systems are linked, in turn, by an information retrieval network. A matching process tries to assign values to each frame's terminals which are partly controlled by information associated with the frame and partly by knowledge about the system's current goals. One of the advantages of this global model is that memory is not separate from the rest of thinking. The information-retrieval based on frame system explains differences between ways of thinking of the person, by assuming mechanism of quickly locating highly appropriate frames for "clever" persons. It indicates that good retrieval mechanism can be based only in part upon basic innate mechanism. It must also depend on (learned) knowledge about the structure of one's own knowledge. The short term memory is connected with sensory buffer and has also the suitable frames.

The neural networks approach uses "sub-symbolic computation" for description of knowledge representation and its processing. The term "sub-symbolic computation" refers to the fact that, in distributed representations, a node is not associated with one particular symbol, being able to take part in the distributed representation of the various concepts. One of the most important features of neural networks is that they perform a large number of numerical operations in parallel. Almost all data stored in the network are involved in recall computation at any given time. The distributed neural processing is typically performed within the entire array composed of neurons and weights. Most of classical information processing models utilize symbolic sequential processing mechanism. In self-organization NN classes of objects are formulated on the basis of a measure of object similarity [69]. Most of measures of similarity are context free, that is, the similarity between any two objects A and B depends on the properties of the objects. In modified version of the BCM neuron, sets of these neurons, which are organized in lateral inhibition architecture, forces different neurons in the network to find different features based on similarity relation. It is similar to classical approach allowing, based on similarity of "concept," to reason by "analogy" in order to obtain knew description. However, neural systems seem to operate in a more "holistic way" than inferential ones: they learn to associate entire input patterns with the corresponding output decisions. For example, the semantic maps are implemented based on the self-organizing feature map. The maps extract semantic relationship that exists within the set of language data-collection of words. The relationship can be reflected by their relative distances on the map containing words positioned according to their meaning or context. This indicates that the trained network can possibly detect the logical similarity between words from the statistics of the contexts in which they are used [64].

In more complex systems, for task such complex like an understanding of the editorial text or machine translation, mixture of the different ways of knowledge representation and manipulation is utilized. For instance, in understanding editorial text system abstract knowledge is organized by memory structures called Argument Units, which represent patterns of support and relationships among beliefs. When combined with domain specific knowledge, it can be used to argue about issues involving plans, goals, and beliefs in particular domain [70]. The hybrid method which is based on the knowledge graph and in which abstractions of the information and classification part of examples are explicitly stored is used for deriving production rules and generalization. These examples taken from field of AI can show how representation of knowledge and its manipulation influence a way of perceiving some of the processes which has similar function in the brain.

AI approach for knowledge representation gives only little attention to "pictorial knowledge" or tried to simplify pictorial information by use of "pictorial attributes." The psychology of language describes how semantic representations of utterances are elaborated. On a certain level of analysis, theories of meaning need not account for the processes which enable meaning to be expressed in mental representation derived from sensory modalities. Accounting for meaning calls for the coordination of abstract symbols through application of appropriate rules. Images can reflect even divergent semantic content but are identifiable with meaning of the sentence. Coding of a picture and coding of words have similar functional properties. The differences that subsist between the processing of picture and symbols are not assumed to reflect the existence of two distinct representational systems, each containing qualitatively different types of information. Differences between pictures and symbols stem from the different representation of image and symbol in memory. Pictorial information could be represented in memory in both modal and prepositional form [71].

The representation of the visual information needs to some extent relay on the visual aspect of the world given by the objects of the given shape. In our approach shape as the main attribute of the visual object is represented by the shape categories given as the shape classes. The visual structure of the world is linked with the categorical structure of the meaning of the world object [72–97].

1.3.5. Learning

As it was shown in Sect. 1.3.4 understanding requires knowledge that is learned through individual experience of the perceived world. The discrimination learning, where a subject learns to respond to a limited range of sensory characteristics is one of the different types of learning that is important part of understanding process. Learning includes the following mental processes that are the most relevant to understanding process: concept formation, the process of sorting experiences according to related feature, problem solving, and perceptual learning (the effects of past experience on sensory perceptions). Associative learning is based on the ability to connect a previously irrelevant stimulus with a particular response. It occurs through the process of conditioning, where reinforcement establishes new behavior patterns. Essence of association lies in the observation that a subject perceives something in the environment (sensations) that result in an awareness of idea of what is perceived. Associations are based on similarity, frequency, salience, attractiveness and closeness of objects or events in space or time. Gestalt psychologists believe that important learning processes involve a restructuring of relationships in the environment, not simply associating them. Psycholinguists indicate that some of the aspect of language learning such as a native "grammar" can be inherited genetically. The contemporary theories of learning indicate on the role of motivation in learning process. Arnheim introduced the concept of perceptual learning that is governed by "stimulus equivalence" or "stimulus generalization." Perception in the broader sense must include mental imagery and its relation to direct sensory observation. Influence of memory on the perception needs to be taken into account in explanation of the visual perception. A perceptual act is never isolated and is modified by the learned or memorized facts that were perceived in the past. In order to transfer the visual perceptual material into the categories that are learned in the past, there is a need for visual concept that can combine the visual and nonvisual data. Visual knowledge acquired in the past helps not only in detecting the nature of an object but it also assigns the present object a place in the system of things constituting our total view of the world. The visual knowledge that is stored in memory needs to supply the working hypotheses, called expectation, about the possible perceptual object. Visual knowledge and correct expectation facilitate perception whereas inappropriate visual concepts will delay or impede it. For example, a Japanese reads without difficulty ideographs printed so small that a Westerner needs a magnifying glass to discern them, not because the Japanese have more acute eyesight but because they hold the kanji characters in visual storage. The percept must define the object clearly and must resemble sufficiently the memory image of the appropriate category. Often, however, there is enough ambiguity in the stimulus to let the observer find different shape patterns in it as he searches for the best fitting model among the ones emerging from the memory storage.

Scientists are looking for universal principles governed all learning processes that could explain the mechanism of learning process. Rigorous, "objective" methodology was attempted so that the behavior of all organisms could be comprehended under a unified system of laws modeled on those posited in the physical sciences. However, today most psychologists believe that no single theory of learning may be appropriate. The last attempts to integrate all knowledge of psychology into a single, grand theory occurred in the 1930s. Three thinkers Guthrie, Hull and Tolman had a big impact on the theory of learning. Guthrie reasoned that responses (not perceptions or mental states) were the ultimate and most important building blocks of learning. Hull argued that "habit strength," a result of practiced, stimulus-response (S-R) activities promoted by reward, was the essential aspect of learning, which he viewed as a gradual process. Tolman contributed the insight that learning is a process that is inferred from behavior. The psychological theories of learning become a basis for the new automatic method of the knowledge acquisition, so-called machine learning.

The system that understands an object by utilizing the learned knowledge (during the thinking process) needs to have an ability to learn new knowledge. Existing systems apply the different machine learning methods in learning of new knowledge. Machine learning is an area of research which investigates the possibility of automated knowledge acquisition by machine. The different learning strategies and methods as the decision tree ID3 [98], CART [99], STAR methodologies [100], explanation-based learning [101], or connectionist model [102] were proposed. However, there is no method that can be used to learn the visual knowledge.

Machine Learning task is to construct complete, autonomous learning systems that start with general inference rules and learning techniques, and gradually acquire complex skills and knowledge through continuous interaction with an information-rich external environment. There are many methods which are used in symbolic machine learning: symbolic empirical learning, explanation-based learning, case-based reasoning (CBR), and integrated learning methods [103], [104]. In symbolic machine learning the following learning methods can be distinguished: learning by induction (learning from examples, learning from observation, learning by discovery), learning by deduction, and learning by analogy. For example, inductive learning, is an inductive inference from facts provided by a teacher or the environment. The process of inductive learning can be viewed as a search for plausible general descriptions (inductive assertions) that explain the given input data and are useful for predicting new data. These assertions form a set of descriptions partially ordered by the relation of relative generality. The minimal elements of this set are the most specific descriptions of the input data in the given language, and the maximal elements are the most general descriptions of these data. The elements of this set can be generated by starting with the most specific descriptions and repeatedly applying rules of generalization to produce more general descriptions. One of the types of inductive learning is conceptual learning from examples (concept acquisition), whose task is to induce general descriptions of concepts from specific instances of these concepts. An important variant of concept learning from examples is the incremental concept refinement, where the input information includes, in addition to the training examples, previously learned hypotheses, or human-provided initial hypotheses that may be partially incorrect or incomplete. In concept acquisition, the observational statements are characterizations of some objects preclassified by a teacher into one or more classes. The induced hypothesis can be viewed as a concept recognition rule, such that if an object satisfies this rule, then it represents the given concept.

Neural networks (connectionist model) are often used in process of knowledge acquisition. One of the most important features of neural networks is that they perform a large number of numerical operations in parallel. These operations involve simple arithmetic operations as well as nonlinear mapping and computation of derivatives. The distributed neural processing is typically performed within the entire array composed of neurons and weights [64]. For example, associative memories respond by retrieving exactly one of the previously stored patterns, even though the stimulus or cue might be partly distorted or missing in part. In contrast with the mode of addressing, associative memories are content-addressable. The words in the memory are accessed based on the key vector. The entire mapping is distributed in the associative network.

1.3.6. Reasoning

Understanding is to pursue new knowledge about the perceptual data based on the reasoning. The reasoning is performed as a part of the thinking process. Reasoning is a process of directed thinking to pursuit a specific goal in order to find a solution to a problem. Reasoning is sometimes narrowly contrasted to feeling, sensation, and desire and is treated as a process that follows the rules of logic. In this approach reasoning is the process building formal arguments from fixed premises in order to reach a conclusion. In less formal, psychological terms, reasoning is a loose combination of mental processes aimed at gaining a more coherent view of some issues in particular or of the world in general. Feelings and desires are not excluded from such efforts. The logical structures that underlie such reasoning processes can be abandoned if necessary. Such reasoning is also called cognition, concept formation and thinking. In logic the reasoning is used to derive of conclusions from given information or premises. Deduction draws out the conclusions implicit in their premises by analyzing valid argument forms. Induction argues from many instances to a general statement. Probability passes from frequencies within a known domain to conclusions of stated likelihood. Statistical reasoning concludes that, on the average, a certain percentage of a set of entities will satisfy the stated conditions.

Existing intelligent systems utilize the different forms of reasoning. One area of research, which deals with process of reasoning, is AI. AI has shown great promise in application of the different forms of reasoning in the area of expert systems ([105]) or knowledge-based expert programs (see e.g., [106]) which, although powerful when answering questions within a specific domain, are nevertheless incapable of any type of adaptable, or truly intelligent, reasoning. In the field of AI many forms of reasoning such as logic reasoning, statistical reasoning or fuzzy reasoning (see e.g., [107–109]) were employed. Logic reasoning forward or backward is often used in the logic programming (see e.g., [110–112]). One of the logic programming languages such as PROLOG applies the goal-directed reasoning (backward) in order to manipulate symbols (actually words) and find the solution. Another form of reasoning is statistical reasoning based on Bayesian statistics implemented in expert systems such as MYCIN [113] or Bayesian networks [114]. The distributed reasoning systems that composed of a set of separate modules (often called agents) become one of the most popular (see e.g., [115]).

Reasoning is often modeled as a process that draws conclusions by chaining together generalized rules, starting from scratch. In case-based reasoning (CBR) the primary knowledge source is not generalized rules but a memory of stored cases recording specific prior episodes. In CBR, new solutions are generated not by chaining, but by retrieving the most relevant cases from memory and adapting them to fit the new situation. CBR is based on remembering. Reminding facilitates human reasoning in many contexts and for many tasks, ranging from children's simple reasoning to expert decision making. Much of the original inspiration for the CBR approach comes from the role of reminding in humans reasoning. The quite extensive description of the different techniques of CBR can be found in [116].

1.3.7. Recognition

One of the simplest forms of understanding is recognition of the known objects or phenomenon. A feature of an object or animal such as the red belly of a stickleback which elicits a response from an animal, is called a key or sign stimulus, and it greatly simplifies the problem of recognition. As long as red objects and fish with red markings are rare in the stickleback's environment, it can use the key stimulus to recognize rivals and does not need to use information about another fish's detailed structure and coloration.

Recognition presupposes the presence of something to be recognized. The most useful and common interaction between perception and memory takes place in the recognition of things seen. Visual recognition considers the issue of addressing and searching through the memory. The right hemisphere system (RHS) handles perceptual organization and facilitates rapid visual identification, while the left hemisphere system (LHS) manages semantic categorization. Perceptual categorization is matching and identifying of a physical object.

In technical recognition, image features and a subset of the most likely models is selected. Acronym [58] is an example of this approach. It is a model-based image interpretation whose primary domain of application has been the recognition of airplanes in aerial views of airports.

1.4. Thinking

Thinking accompanies nearly all mental operations. There is no basic difference between what happens when a person looks at the world directly and when person sits with his eyes closed and "thinks." Thinking can deal with directly perceived objects, which often are handled physically. When no objects are present, they are replaced by some sort of imagery. Thinking accompanies all processes such as reasoning and problem solving. Thinking begins with the task of modifying a perceptual structure of the perceived data in order to interpret it or enable it fit the requirements of the solution to a given problem. Reasoning about an object starts with the way the object is perceived. An inadequate percept may cause an error in the reasoning process and in the final result of interpretation and understanding. To think of something is not to grasp it immediately in the way that perception grasps its object. Our thoughts can grasp object only by transforming it into one of the representational form such as the concept.

The word thinking covers several distinct psychological activities. It is sometimes a synonym for "tending to believe," ("I think that it will rain, but I am not sure.”). Thinking was often understood very narrowly as a sort of reasoning. For example, when Aristotle talked about thinking he referred to the syllogism, that is, to the art of making a statement on a particular case by consulting a higher generality. The higher generality refers to generality of thought that was concerned with whole class of general or potentially general entities such as forms, universals, essences and sensible species.

In the psychological sense thinking is intellectual exertion aimed at finding an answer to a question or a means of achieving a desirable practical goal. The psychology of thought processes concerns itself with activities similar to those usually attributed to the inventor, the mathematician, or the chess player. However, there is no agreement among psychologists concerning the definition or characterization of thinking. Some psychologist schools claim that thinking is a matter of modifying perceptual representations of the world (cognitive structures), according to others view thinking is considered as internal problem-solving behavior.

Thinking and thought were the topic of philosophical investigations. For example, Locke held that thought and idea is the same things. According to Locke the idea is “the representation of something in the mind” and to frame such a representation of an object is to be engaged in the thinking process. He claimed that ideas are not Real Beings but only mode of thinking. The insistence that ideas are acts of thought, that to have an idea and to be conscious are the same, was a position taken by Arnaud in his dispute with Melabranche. According to Berkeley [43] thought and their contents possesses two strictly distinct but closely connected properties. The first one is called intentionality – thought and their contents are about things other than themselves. The second one is called generality – near all our concepts express feature which an indefinite number of things might posses. Locke holds that discursive thinking is mentally manipulating “abstract ideas” which he describes as attenuated images. Latest research support hypothesis that the thinking process is based on transformation of the “mental data” in the form of concepts. Perception supplies perceptual material that can be used for thought because perception gathers types of things (concepts). Unless the perceptual material in the form of concepts remains present in the mind, the mind has nothing to think with.

For Natorp, as for Kant, thinking is an activity which refers to the technical term “function.” Term “function” used by Natorp seems to mean something like subjective, psychic act, and as such is excluded from epistemological consideration. “Function” signifies the spontaneity of thinking, not in psychological terms, but as the rational act of hypothetical legislation. For Natorp the standard sense of “function” is an act or “operation” of thinking. “Operation” of thinking means laying down a hypothesis, where the hypothesis is always a concept, a generality that imposes a unity upon a phenomenal manifold. For the neo-Kantians "thinking" is restricted to "scientific thinking." According to Natorp science is a movement of method via hypotheses regulated by an ideal of complete objective determination. Natorp's logic is telling how thinking lawfully generates or synthesizes the unities that are its objects of knowledge.

The term thinking was often used to denote the different brain processes. Some scientists claim that thinking is purely physiological occupation of the brain. This claim is based on assumption that everything in the mind must have its counterpart in the nervous system. Under this assumption the brain contains the bodily equivalent of all concepts available to thinking as well as of all operations to which concept can be subjected. Similarly the Computational Theory of Mind (CTM) treat brain processes as processes that can be explained by using computational model of our brain. CTM was proposed by Hilary Putnam (1961) and developed by Jerry Fodor. CTM is one of the most important theories of mind and thinking is one of the problems that CTM tried to explain. This theory is based on the computer metaphor [117], [14]. Researchers sought to endow machines with human-level competences in reasoning, language, problem-solving, and perception. According to CTM the mind can be seen as the popwerfull computer due to successes of computational models of reasoning, language and perception that lent credibility to the idea that such processes might be accomplished through computation in the mind as well.

The linguistic approach has an impact on the modern theory of mind such as the representational theory of the mind (RTM). According to the RTM, thinking occurs in an internal system of representation. Beliefs and desires and other propositional attitudes enter into mental processes as internal symbols. Modern versions of RTM assume that thought is not grounded in mental images. These philosophers maintains that the internal system of representation has a language-like syntax and a compositional semantics. According to this view, much of thought is grounded in word-like mental representations. This view is often referred to as the Language of Thought Hypothesis (LOTH) [118].

The latest result of philosophical investigations concerning thinking is the LOTH, which postulates that thinking take place in a mental language [118]. This language consists of a system of representations that is physically realized in the brain of thinkers and has a combinatorial syntax (and semantics) such that operations on representations are causally sensitive only to the syntactic properties of representations. According to LOTH, thought is, roughly, the tokening of a representation that has a syntactic (constituent) structure with an appropriate semantics. Thinking consists in syntactic operations defined over such representations. Most of the arguments for LOTH derive their strength from their ability to explain certain empirical phenomena like productivity and systematicity of thought and thinking.

LOTH is an hypothesis about the nature of thinking with propositional content and is not applicable to other aspects of mental life such as sensory processes, mental images, visual and auditory imagination, sensory memory, perceptual pattern-recognition capacities, dreaming, or hallucinating.

According to Arnheim vision is the primary medium of thought. We can not think using category of smells and tests. In vision and hearing, shapes, colors, movements, sounds, are susceptible to definite and highly complex organization in space and time. Arnheim included the process of thinking in the process of visual perception and claims that thought process has its roots in perception. Arnheim claims that thinking is not the privilege of mental processes above and beyond perception but the essential ingredients of perception itself. He is referring to such operations as active exploration, selection, grasping of essentials, simplification, abstraction, analysis and synthesis, completion, correction, comparison, problem solving, as well as combining, separating, or putting in the context. These operations are the manner in which the minds of both man and animal treat cognitive material at any level.

The important part of thinking process is object of thoughts. Traditional Aristotelian and scholastic philosophy had distinguished between two kinds of objects of mental life. The first forms or species are universals and appropriate for intellect and thoughts. The second phantasms are objects for sensory perception and are particular sensory images. In the past concept was often regarded as an idea, object of mental life. For example, Lock often used the term idea for description of concept. Locke's idea means whatsoever the mind perceives in itself, or the immediate object of perception, thought or understanding [40]. Locke defines the term "idea" as "whatsoever is the object of understanding when a man thinks" and includes sensations and sensory images among ideas. Sensory images become paradigm ideas and are treated as sensory or quasi-sensory images. Locke tried to solve the problem of generality of "images" by invoking abstract general idea treated as abstract general images, e.g., an idea of triangle is an image which is, at the same time, every specific kind of triangle-isosceles, scalene, and none in particular. Berkeley [43] shows that there is no sense in the idea of such an image. His alternative theory is that a particular image becomes general by representing or standing for some class of images. From the point of view of the subject, this happens when he is selecting of the relevant feature of the image. When someone is imaging an equilateral triangle, he takes it as representing all triangles by assuming that it has three sides, and ignoring their relative proportions, its size, angle, color. Berkeley argues that we cannot form the idea of something unthought-of, for once we form such an idea its object is, ipso facto, thought of. It requires distinction between the thought and its object. The thought is in the mind, the object is not. This distinction is the intentionality of thought. From intentionality of thought we conclude: if everything thinkable can be realized in an image as a feature of it, then the concept of mind-independent matter, and mind-independence, should be realizable. But, necessarily, such things can not be properties of images, which are essentially mental, so we cannot have idea of them. To overcome the problem solipsism he does have a doctrine of representation whereby an idea can stand for others. This doctrine is the source of the associationism theory, which was to be the principle empiricist account of meaning and thought until the end of the nineteenth century.

According to Kant it is not images of objects, but schemata, which lie at the foundation of our pure sensible concepts. No image could ever be adequate to our concept of a triangle in general. For the generality of the concept it could never attain to, as this includes under itself all triangles, whether right-angled, acute-angled, etc. while the image would always be limited to a single part of this sphere. The schema of the triangle can exist nowhere else than in thought, and it indicates a rule of the synthesis of the imagination in regard to pure figure in space. Still less is an object of experience, or an image of the object, ever adequate to the empirical concept. On the contrary, the concept always relates immediately to the schema of the imagination, as a rule of the determination of our intuition, in conformity with a certain general concept. The concept of dog indicates the rule, according to which our imagination can delineate the figure of a four-footed animal in general. We can say – the image is a product of the empirical faculty of the productive imagination – while the schema of sensible concepts is a product of the pure imagination a priori. According to pure imagination a priori image first becomes possible, which, however, can be connected with a concept only by means of the schema which they indicate. On the other hand, the schema of the pure concept of the understanding is something that can not be reduced into any image. It is nothing else than the pure synthesis expressed by the category, conformably to the rule of unity according to concepts. It is a transcendental product of the imagination, the product which concerns the determination of the internal sense, according to conditions of its form (time) in respect to all representations.

According to Arnheim concept is the type of things that is gathered by perception. In the philosophical investigation the visual concept is not present. Arnheim have tried to define the visual concept in context of the visual thinking. For Arnheim memory is a storehouse of visual concepts, some clear-cut and simple, some elusive and intangible, covering the whole of the object or recalling only fragments. Memory images serve to identify, interpret, and supplement perception. Memory concepts aid this search by being no less flexible than percepts. Arnheim gives an example of the visual concept of the cube. The visual concept of the cube embraces the multiplicity of its appearances, the foreshortenings, the slants, the symmetries and asymmetries, the partial concealments and the deployments, the head-on flatness and the pronounced volumes.

One of the oldest questions about concepts concerns whether there are any innate concepts. Empiricists maintain that there are few if any innate concepts and that most cognitive capacities are acquired on the basis of a few relatively simple cognitive mechanisms. Empiricists claim that all concepts are derived from sensations. According to empiricists concepts were formed from copies of sensory representations and assembled in accordance with a set of general-purpose learning rules.

Nativists, on the other hand, maintain that there may be many innate concepts and that the mind has a great deal of innate differentiation into complex domain-specific subsystems. For example, Fodor argued that all models of concept learning treat concept learning as hypothesis testing [118], and the concept has to be available to a learner prior to the learning taking place. He claims that lexical concepts lack semantic structure and consequently that virtually all lexical concepts must be innate – a position known as radical concept nativism.

Nearly all theories concerning thought assume that concept is the object of thought. Only few scientists such as Peacocke [119] maintain that a thinker can represent the world nonconceptually without possessing any concepts at all. According to the classical theory, a lexical concept C has definitional structure in that it is composed of simpler concepts that express necessary and sufficient conditions for falling under *C*. A nonclassical alternative is the *prototype theory*. According to this theory, a lexical concept *C* does not have definitional structure but has probabilistic structure in that something falls under *C* just in case it satisfies a sufficient number of properties encoded by *C*'s constituents. The view that concepts are Fregean senses identifies concepts with abstract objects, as opposed to mental objects and mental states. According to this view concepts mediate between thought and language, on the one hand, and referents, on the other [119]. Concepts are psychological entities that play a key role in the RTM. According to another so-called theory, *theory of concepts*, concepts stand in relation to one another in the same way as the terms of a scientific theory, and that categorization is a process that strongly resembles scientific theorizing [120]. It is generally assumed that the terms of a scientific theory are interdefined so that a theoretical term's content is determined by its unique role in the theory in which it occurs. According to *conceptual atomism* lexical concepts have no semantic structure [121] and the content of a concept is not determined by its relation to other concepts but by its relation to the world.

Concepts become the important material for thinking process. Psychological description of thought process deals with the problem of how concepts are formed and related in the brain. Arnheim suggested that concepts are perceptual images and that thought operate by handling of these images and these images can be regarded on the many levels of abstractness. According to Arnheim mental images serve as the vehicle of thought. Freud raises the question of how the important logical links of reasoning can be represented in images. He claims that mental image is material for thinking that is not necessarily consciousness. According to Kant schema that is the basis of the pure sensible concept exists in thought. For example, the schema of the triangle can exist nowhere else than in thought, and it indicates a rule of the synthesis of the imagination in regard to pure figure in space.

Associationist explained thought and meaning by saying that the meaning of a mental episode consisted in its association with other mental episodes which tend to occur in close proximity of it. For Berkeley the fundamental form by which ideas represent one another is through resemblance, not simply by contingent association, like cause and effect, which need possess no similarity. It is way one triangle can represent all triangles.

A distinction is sometimes made between convergent thinking – the analytic reasoning measured by intelligence tests, and divergent thinking, a richness of ideas and originality of thinking. Both seem necessary to creative performance, although in different degrees according to the task or occupation (a mathematician may exhibit more convergent than divergent thinking and an artist the reverse). The directed thinking is the term that denotes thinking that is aimed at the solution of a specific problem and fulfills the criteria for reasoning. The term visual thinking is used to denote the thinking connected with processing of the visual material and concerning visual objects.

The term visual thinking is used to denote the thinking connected with processing of the visual material and concerning visual objects. Visual thinking resembles the problem-solving process that leads to scientific discoveries. Perception, as the process that involves thinking, can abstract objects from their context only because it grasps shape as organized structure, rather than recording it as a mosaic of elements. In more than one way, perceptual abstraction can differ from the kind described in traditional logic. Typically, it is not a matter of extracting common properties from a number of particular instances. The one possible explanation of the visual thinking process is assumption that visual thinking is realizing in some medium that are called "mental images." The kind of "mental images" needed for thought is unlikely to be a complete, colorful, and faithful replica of some visible scene. The "mental images" which are the main ingredient of the visual thought have the different nature than the nonvisual concept. Such a different constitution of both concepts has a big influence on the difference in thinking process. For example, the analytical formula of a geometrical figure such as a circle gives the location of all the points of which the circle consists. It does not describe its particular character, its centric symmetry or its rigid curvature. However, it is precisely this grasping of the character of a given phenomenon that makes productive thinking possible.

The cognitive operations connected with visual thinking are remarkably rich. They are represented indirectly by what is remembered and known about them. Visual thinking is necessarily concerned with generalities of "mental images." Mental images are governed by the rules of selectivity. The thinker can focus on what is relevant and dismiss from visibility what is not. They may be quite common and indispensable to mind that thinks generic thoughts and needs the generality of pure shapes to think them. Anaheim claims that if thinking takes place in the realm of images, many of these images must be highly abstract since the mind operates often at high levels of abstraction. He claimed that a good deal of imagery may occur below the level of consciousness and that even if conscious, such imagery may not be noticed readily by persons unaccustomed to self-observation. The thinking by using abstract images may utilize diagrammatic representation. Eidetic images seem to be used as a target for active perception. Therefore, they can serve as material for thought but are unlikely to be a suitable instrument of thought.

Behaviorism as a materalistic movement claims that imageless thoughts are the only possible form of thoughts. Watson [122] in his ideological attack on existence of mental imagery show haw big role in establishing "scientific" methodology of psychological research play a materialistic prejudices. A similar doctrine was maintained by the logical positivists (Vienna Circle) in the early twentieth century. Their principle of verification required for a sentence or statement to be meaningful that it have empirical consequences, and, on some formulations of the principle, that the meaning of a sentence is the empirical procedure for confirming it. Sentences that have no empirical consequences were deemed to be meaningless.

The problem of thought and language was investigated by philosophers and linguistics. There is a problem that until now is not solved – how big role language plays in the thought process. This problem was often formulated in the following form: can one think in words as one can think in circles or rectangles or other shapes? Many philosophers claim that the semantic properties of linguistic expressions are inherited from the intentional mental states they are conventionally used to express [123], [117]. On this view, the semantic properties of linguistic expressions are the semantic properties of the representations that are related to the conceptual structure of the linguistic forms. Martin [124], for example, claimed that thought is possible without language, whereas others, such as [125] have suggested that the kind of thought human beings are capable of is not possible without language. It is claimed that if the semantic properties of natural-language expressions are inherited from the thoughts and concepts they express, then an analogous distinction may be appropriate for mental representations. Even if it is agreed that it is possible to have concepts in the absence of language, there is a dispute about how the two are related. Some maintain that concepts are prior to and independent of natural language, and that natural language is just a means for conveying thought [118]. Others maintain that at least some types of thinking (and hence some concepts) occur in the internal system of representation constituting our natural language competence. According to linguistic determinism the language a person speaks causes him to conceptualize the world in certain ways while delimiting the boundaries of his conceptual system; as a result, people who speak very different languages are likely to conceptualize the world in correspondingly different ways. This linguistic relativity is the weaker doctrine that the language one speaks influences how one thinks. Observation of the animals indicates that language can only to some extent facilitate the thinking process. Animal can respond to categories of things, and they display an astonishing disregard of the unique object. By means of their perceptual concepts, animals solve problems that look elementary if judged by human standard. However, the animal thinking can cope only with directly given situation.

1.5. Shape Understanding System

The short introduction which was given in this chapter shows how thinking and in particular visual thinking is complex and difficult problem. One of the ways to investigate the thinking problem is to build the machine that can show ability to think. The problem of building thinking machines appear in the latest time within the research in cognitive science or AI. The expert systems that were built in order to assist in solving complex problems were often regarded as systems that have an ability to think and to understand. MYCIN [113] or DENDRAL are well known examples of expert systems that pave the way for the new research in building sophisticated knowledge-based systems. The neural network-based systems that learn knowledge from the data supply new facts about complexity of the brain processes. In robotics, the aim of which is to build the machine that will be able to act in a similar way as human being, the most of research is concentrated on a problem of navigation in unknown environment or to build a humanoid robot. However, the humanoid robot should not only look similar to the human shape but also it needs to posses the thinking capabilities to be able to exist in the human world. In order to build the machine with the visual thinking capability there is a need to find the proper representation of the visual knowledge.

There are systems that are built in order to interpret the perceived object or interpret an image, however, these systems do not assume understanding of the perceived object or the real-world scene. These systems are built based on the research in the area of computer vision and image understanding. The term image understanding has a range of meanings, but in general, image understanding refers to a computational, information processing approach to image interpretation. The term image understanding denotes an interdisciplinary research area which includes signal processing, statistical and syntactic pattern recognition, artificial intelligence, and psychology. Image understanding refers to knowledge-based interpretations of visual scenes that transform pictorial inputs into commonly understood descriptions or symbols (see e.g., [126–128]). Computer vision is also used to refer to a similar research area ([129], [130]), but while computer vision emphasizes the computational aspects of visual information processing, such as measurement of three-dimensional shape information by visual sensors, image understanding stresses knowledge representation and reasoning methods for scene interpretation. Another field of research which stresses modeling of the human visual system, called computational vision can also be treated as a field of image understanding research. Computational vision is a multidisciplinary and synergetic approach whose main task is to explain the processes of the human visual system and build artificial visual systems [131]. In contrast to image understanding approach where interpretation of the image in terms of the real-world scene is the main goal of the understanding process, the shape understanding method is focused on interpretation of visual objects (called phantoms) in the context of knowledge of object categories. In our approach presented in this book the shape categories in form of the visual concepts and object categories are main ingredients of the thinking process. Object categories refer to meaning of the perceived object. The detailed description of object categories will be presented in Chap. 4.

In shape understanding method an object extracted from the image is interpreted in the context of knowledge of object categories that are acquired during the learning process [95]. In this approach the image is the basic source of visual knowledge about an object. The image supplies also the contextual knowledge that is used in interpretation of the image. However, the visual object can be also interpreted in the context of knowledge obtained during the examination of another image or knowledge given in other nonvisual form such as text or spoken words. In the method described in this book knowledge about the world is represented in the form of categories of visual objects. The perceived object is fitted to one of the shape categories and next classified to one of object categories. The perceived object triggers the thinking process that can lead to interpretation of the object in terms of the real-world object or in terms of the visual signs. The contextual information and the result of interpretation of the perceived object can be utilized during the thinking process to perform the required task or to take an appropriate action. Visual information acquired during perception is transformed into visual concepts and is combined with nonvisual information during the visual reasoning and the thinking process. Nonvisual information can be utilized to produce the linguistic description about the object. Information conveyed by language is compatible with information acquired during perception of the visual objects. The thinking process can be also driven by perception of the "inner" objects that are results of the imagery transformations.

According to Arnheim the perception of shape consists in the application of form categories, which can be called visual concepts because of their simplicity and generality. In our approach the shape categories refer to the shape classes (described in Chap. 2) whereas the visual concept is the set of shape categories given by their symbolic names and is part of the visual categories represented by the categorical chain of the visual objects (described in Chap. 4).

The proposed system of shape understanding operates based on knowledge of image processing, decision making and search strategies as well as knowledge of shape description and representation distributed among the specialized experts [76–82], [84], [86–97], [132–146]. The analysis is carried out by invoking the "expert" that performs a suitable analysis employing a specific method. The expert performs the task given in a form of requirements by another expert based on the internal ability to use the knowledge of its domain of expertise as well as communicate the obtained results with other experts. The SUS consists of the two main modules: the central reasoning module and the peripheral module. The central reasoning module consists of the master expert E^M, the reasoning expert E^R, the manager expert E^Q, the processing expert E^P and the end-the analysis expert E^E. The peripheral module consists of the generating expert, the question expert, the self-correcting expert, the learning expert and the spatial-logic expert. The system may cooperate with distributed experts by utilizing Distributed Component Object Model (DCOM) technology. It makes it possible to have access to expertise from an expert that is part of the system from the different domains, e.g., text understanding.

The reasoning expert E^R is an expert that manages the process of reasoning and is invoked by the master expert E^M. The reasoning involves transformation of the description of the examined object s when passing stages $\varsigma_0 \rightarrow \varsigma_1 \rightarrow ..... \rightarrow \varsigma_N$, where ς_0 is the beginning stage, ς_N is the final stage of the reasoning process and $\rightarrow$ denotes the move to the next stage of reasoning. If at a given stage of analysis there is a need to acquire the new data or to make a decision about further processing an appropriate expert is invoked. The reasoning expert E^R makes a decision based on the expertise α of the decision-making expert E^D and reasoning parameters supplied by the master expert E^M. In this paper the following notation is introduced: $E^R_{\varsigma_i}\{[\alpha \equiv x] \Rightarrow c\} \mapsto E^B$, that means the reasoning expert $E^R_{\varsigma_i}$, based on expertise α supplied by the decision making expert, formulates protocol c and invokes another expert. The symbol $\mapsto$ denotes that the new expert is invoked.

Depending on the expertise α the following decisions are possible:

1. The examined object s is assigned to the class Ω^{ς_i} and reasoning is moved to the next stage of reasoning ς_{i+1}:

$$E^R_{\varsigma_i}\{[\alpha \equiv x_1] \Rightarrow s \triangleright \Omega^{\varsigma_i}\} \leftrightarrow E^M(c_M) \mapsto E^R_{\varsigma_{i+1}},$$

where $\alpha \equiv x_1$ denotes that expertise α given by the decision-making expert E^D indicates that the object s can be assigned to the class Ω^{ς_i}. The symbol $\leftrightarrow$ denotes that the reasoning expert returns the protocol c_M to the master expert E^M and symbol $\mapsto$ denotes that the new reasoning expert of the next stage of reasoning ς_{i+1} is invoked;

2. The new data are needed to assign the object to one of the possible classes Ω^{ς_i} of the given stage of processing ς_i. The requirements are formulated as a protocol c_Q and an appropriate data acquisition expert $E^Q_{\varsigma_j}$ is invoked to gather new data:

$$E^R_{\varsigma_i}\{[\alpha \equiv x_2] \Rightarrow c_Q\} \mapsto E^Q_{\varsigma_i}.$$

3. The new data are acquired and the decision-making expert is invoked to make a decision about further processing:

$$E^R_{\varsigma_i}\{[\alpha \equiv x_3] \Rightarrow c_D\} \mapsto E^D_{\varsigma_i};$$

4. Based on the results of stage analysis and a result of the decision-making expert the reasoning expert E^R makes a decision that the reasoning reaches the last stage and the end-analysis expert E^E is invoked:

$$E^R_{\varsigma_i}\{[\alpha \equiv x_4] \Rightarrow c_E\} \mapsto E^E_{\varsigma_i}.$$

During the process of reasoning the reasoning expert keeps track of the analysis process. The levels of detail Ψ^d as well as the exactness of the processing results Ψ^ε are important factors that are set during the user session. The level of detail Ψ^d describes the level of description of the object, which needs to be obtained.

The reasoning involves transformation of the description of the object passing successive stages $\varsigma_0 \to \varsigma_1 \to \ldots\ldots \to \varsigma_N$. To stop the process of reasoning, information about the level of detail Ψ^d is needed. At each stage of reasoning the method StageEvaluation() of the reasoning expert checks if for the given stage the level of detail Ψ^d is obtained. In the case when the last stage is reached the reasoning expert invokes the end-analysis expert.

At each stage, based on the expertise α obtained from the decision-making expert E^D, and the levels of the detail Ψ^d and the exactness of the processing results Ψ^ε supplied by the master expert, the following actions are undertaken:

- The reasoning is moved from the stage *i* to the stage *i*+1 in the sequence $\varsigma_0 \to \varsigma_1 \to \ldots \to \varsigma_i \to \varsigma_{i+1}$.
- The decision is not undertaken and the new decision making expert needs to be consulted or the new data needs to be acquired.

For example, in the case when the examined object is a realization of the class $Q^1[L_S^4](L_E^3)$ the possible ways of reasoning, depending on the assumed level of detail Ψ^d, can be as follows:

$$\Gamma \Rightarrow Q \Rightarrow Q^n \Rightarrow Q^n_{L^m} \Rightarrow Q^n_{L^4} \Rightarrow Q^n_{L^4_S} \Rightarrow Q^1_{L^4_S} \Rightarrow Q^1_{L^4_S}(\Lambda) \Rightarrow Q^1_{L^4_S}(L^n) \Rightarrow Q^1_{L^4_S}(L^3) \Rightarrow Q^1_{L^4_S}(L^3_E)$$

or

$$\Gamma \Rightarrow Q \Rightarrow Q^1 \Rightarrow Q^1_\Lambda(\Lambda) \Rightarrow Q^1_{L^n}(\Lambda) \Rightarrow Q^1_{L^4}(\Lambda) \Rightarrow Q^1_{L^4}(L^n),$$

where the symbol $\Rightarrow$ denotes moving to the next stage of the reasoning process. In the case of nonregular classes the reasoning process is more complicated.

In Chap. 2 the shape classes will be described end the reasoning process will be presented in Chap. 3.

References

1. Popper, K., and Eccles, J.C., *The Self and its Brain*. 1977, London: Routledge
2. Plato, *Republic. The World's Classics*, ed. Waterfield. 1994, Oxford: Oxford University Press
3. Arnheim, R., *Visual Thinking*. 1970, London: Faber and Faber
4. Leyton, M., Inferring causal history from shape. *Cognitive Science*, 1989. **13**: pp. 357–387
5. Kendall, D., Barden, D., Carne, T.K., and Le, H., *Shape and Shape Theory*. 1999, Chichester: Wiley

6. Dryden, I.L., and Marida, K.V., *Statistical Shape Analysis*. 1998, Chichester: Wiley
7. Small, G.C., *The Statistical Theory of Shapes*. 1996, Berlin Heidelberg New York: Springer
8. Kendall, D.G., The diffusion of shape. *Advances in Applied Probability*, 1977. **9**: pp. 428–430
9. Grenander, U., Chow, Y., and Keenan, D.M., *Hands. A Pattern Theoretic Study of Biological Shapes*. 1991, Berlin Heidelberg New York: Springer
10. Cootes, T.F., Taylor, C.J., Cooper, D.H., and Graham, J., Image search using flexible shape models generated from sets of examples, in *Statistics and Images*, K.V. Marida, Editor. 1994, Carfax: Oxford
11. Grenander, U., and Miller, M.I., Representation of knowledge in complex systems. *Journal of the Royal Statistical Society*, 1994. **56**: pp. 549–603
12. Peitgen, H.-O., and Saupe, D., eds. *The Science of Fractal Images*. 1988, Berlin Heidelberg New York: Springer
13. Masters, B.R., Fractal analysis of normal human retinal blood vessels. *Fractals*, 1994. **2**(1): pp. 103–110
14. Marr, D., *Vision*. 1982, San Francisco: Freeman
15. Marr, D., and Nishihara, K., Representation and recognition of the spatial organization of three dimensional shapes. *Proceedings of the Royal Society of London*, 1978. **Series B 200**: pp. 269–294
16. Biederman, I., Recognition by components: a theory of human image understanding. *Psychological Review*, 1987. **94**: pp. 115–147
17. Bhanu, B., and Faugeras, O.D., Shape matching of two-dimensional objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1984. **6**(2): pp. 137–156
18. Lu, C.H., and Dunham, J.G., Shape matching using polygon approximation and dynamic alignment. *Pattern Recognition Letters*, 1993. **14**: pp. 945–949
19. Pavlidis, T., and Ali, F., A hierarchical syntactic shape analyzer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1977. **PAMI-1**(1): pp. 2–9
20. Bala, J., and Wechsler, H., Shape analysis using genetic algorithms. *Pattern Recognition Letters*, 1993. **14**: pp. 965–973
21. Bala, J., and Wechsler, H., Shape analysis using hybrid learning. *Pattern Recognition*, 1996. **29**(8): pp. 1323–1333
22. He, Y., and Kundu, A., 2-D shape classification using hidden Markov model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1991. **13**(11): pp. 1172–1184
23. Kartikeayan, B., and Sarkar, A., Shape description by time series. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989. **11**(9): pp. 977–984
24. Pal, N.R., Pal, P., and Basu, A.K., A new shape representation scheme and its application to shape discrimination using a neural network. *Pattern Recognition*, 1993. **26**(4): pp. 543–551
25. Mahmoud, S.I., Arabic character recognition using Fourier descriptors and character contour encoding. *Pattern Recognition*, 1994. **27**(6): pp. 815–824

26. Samal, A., and Edwards, J., Generalized hough transform for natural shapes. *Pattern Recognition Letters*, 1997. **18**: pp. 473–480
27. Pinker, S., Visual cognition: an introduction, in *Visual Cognition*, S. Pinker, Editor. 1984, Cambridge, MA: MIT Press, pp. 1–63
28. Quinlan, P., Differing approaches to two-dimensional shape recognition. *Psychological Bulletin*, 1991. **109**: pp. 224–241
29. Palmer, S., Visual perception and world knowledge: notes on a model of sensory-cognitive interaction, in *Explorations in Cognition*, D.A. Norman, and Rumelhart, D.E., Editors. 1975, San Francisco: Freeman, pp. 279–307
30. Tversky, B., and Hemenway, K., Objects, parts and categories. *Journal of Experimental Psychology: General*, 1984. **113**: pp. 169–193
31. Hummel, J., and Biederman, I., Dynamic binding in neural network for shape recognition. *Psychological Review*, 1992. **99**: pp. 480–517
32. Siddiqi, K., Tresness, K.J., and Kimia, B.B., Parts of visual form: psychophysical aspects. *Perception*, 1996. **25**: pp. 399–424
33. Ho man, D.D., and Singh, M., Salience of visual parts. *Cognition*, 1997. **63**: pp. 29–78
34. Ho man, D.D., and Richards, W.A., Parts of recognition. *Cognition*, 1984. **18**: pp. 65–96
35. Angenent, S., On the information of singularities in the curve shortening flow. *Journal of Differential Geometry*, 1989. **33**: pp. 610–633
36. Gage, M., and Hamilton, R., The heat equation shrinking convex plane curve. *Journal of Differential Geometry*, 1986. **23**: pp. 69–96
37. Mokhtarian, F., and Mackworth, A.K., Scale-based description and recognition of planar curves and two-dimensional shapes. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 1986. **8**: pp. 34–43
38. Koenderink, J., The structure of images. *Biological Cybernetics*, 1984. **50**: pp. 363–370
39. Koenderink, J.J., and van Doorn, A.J., Dynamic shape. *Biological Cybernetics*, 1986. **53**: pp. 383–396
40. Locke, J., *An Essay Concerning Human Understanding*, ed. J. Yolton. 1961, London: Dent
41. Gilson, E., *The Spirit of Mediaeval Philosophy*. 1991, London: University of Notre Dame Press
42. Hume, D., *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, ed. L.A. Selby-Bigge, 1975, Oxford: Oxford University Press
43. Berkeley, G., *Principles of human knowledge and three dialogues. The World's Classics*, ed. H. Robinson. 1996, Oxford: Oxford University Press
44. Leibnitz, G.W., *New essay on human understanding. Cambridge Texts in the History of Philosophy*, eds. P. Remnant, and Bennet, J. 1996, Glasgow: Cambridge University Press
45. Kant, I., *Critique of pure reason. The Everyman Library*, ed. V. Politis. 1996, London: Everyman

46. Descartes, R., Meditations on first philosophy, in *The Philosophical Writings of Descartes*, Vol. 2, J. Cottingham, Stoothoff, R., and Murdoch, D., Editors. 1984, Cambridge, MA: Cambridge University Press
47. Natorp, P., *Hermann Cohens philosophische Leistung unter dem Gesichtspunkte des Systems.d*. Liebert, A., ed. P.V.v.v.d.K.-G. 21. 1918, Berlin: Reuther & Reichar
48. Gadamar., H.G., *Truth and Method*. 1989, New York: Crossroad
49. Nietzsche, F., *The Portable Nietzsche*, ed. W. Kaufmann. 1954, New York: Viking Press
50. Attneave, F., Some informational aspects of visual perception. *Psychological Review*, 1954. **61**: pp. 183–193
51. Richards, W., and Hoffman, D., *Codon Constraints on Closed 2-D Shapes*. *Graphics and Image Processing*, 1985. **31**: pp. 265–281
52. Rock, I., *The Logic of Perception*. 1983, Cambridge, MA: The MIT Press
53. Koehler, W., *Gestalt Psychology*. 1929, New York: Liverlight
54. Koffka, K., *Principles of Gestalt Psychology*. 1935, Jovanovic: Harcourt Brace
55. Hebb, D.O., *The Organization of Behavior*. 1949, New York: Wiley
56. Minsky, M., and Papert, S., *Perceptrons*. 1969, Cambridge, MA: MIT Press
57. Gibson, J.J., *The Perception of the Visual World*. 1950, Boston: Houghton
58. Brooks, R., Symbolic reasoning among 3-dimensional models and 2-dimensional images. *Artificial Intelligence*, 1981. **17**: pp. 285–349
59. Iglesias, C., Garijo, M., and Gonzalez, J.A., Survey of agent-oriented methodologies*, in Intelligent Agents V: Agents Theories, Architectures, and Languages*. Lecture Notes in Computer Science. 1998, Berlin Heidelberg New York: Springer
60. Jennings, N.R., *On agent-based software engineering*. *Artificial Intelligence*, 2000. **117**: pp. 277–296
61. Gamma, E., Helm, R., Johnson, R., and Vlissides, J., *Design Patterns*. 1994, Reading, MA: Addison-Wesley
62. Hutt, I., *Object-Oriented Analysis and Design: Description Method*. 1994, New York: Wiley
63. Marshall, J., Sensation and semantic. *Nature*, 1988. **334**(4): pp. 378–378
64. Zurada, J.M., *Introduction to Artificial Neural Systems*. 1992, St. Paul, MN: West Publishing Company
65. Shank, R.C., Conceptual dependency: a theory of natural language understanding. *Cognitive Psychology*, 1972. **3**: pp. 552–631
66. Pearl, J., *Probabilistic Reasoning in Intelligent Systems: Network of Plausible Inference*. 1988, San Mateo: Morgan Kaufman
67. Isham, V., An introduction to spatial processes and Markov random field. *International Statistic Review*, 1981. **49**: pp. 21–43
68. Rich, E., and Knight, K., *Artificial Intelligence*. 1991, New York: McGraw-Hill
69. Kohonen, T., *Self-organization Systems*. 1984, Berlin Heidelberg New York: Springer
70. Alvarado, S.J., *Understanding Editorial Text: A Computer Model of Argument Comprehension*. 1991, Boston: Kluwer Academic Publishers
71. Dnis, M., *Image and Cognition*. 1989, New York: Harvester

72. Les, Z., The processing method as a set of image transformations in shape understanding. *An International Journal Computers and Graphics*, 2001. **25**(2): pp. 223–233
73. Les, Z., and Les, M., Shape understanding system in automotive and transportation technology, in *The ISATA'2000 International Conference*. 2000, Dublin: ISATA-Dusseldorf, Epsom
74. Les, Z., and Les, M., Shape understanding system: the noisy class, in *International Conference on Information Systems Analysis and Synthesis and World Multiconference on Systemics, Cybernetics and Informatics, July 23–26*. 2001, Orlando
75. Les, Z., Shape understanding. Possible classes of shapes. *International Journal of Shape Modelling*, 2001. **7**(1): pp. 75–109
76. Les, Z., and Les, M., Shape understanding system: learning of the visual concepts, in *International Conference on Information Systems Analysis and Synthesis and World Multiconference on Systemics, Cybernetics and Informatics, July 23–26*. 2001, Orlando
77. Les, Z., and Les, M., Shape understanding system: application of fuzzy sets, neural networks and statistical methods in the process of visual thinking, in *The 10th IEEE Conference on Fuzzy Systems*. 2001, Melbourne
78. Les, Z., and Les, M., Visual Thinking as a Part of Understanding Process – A New Way of Problem Solving and Communication of the Intelligent Systems in Automotive and Transportation Technology, in *Automotive and Transportation Technology International Conference, ATT*. 2002, Paris
79. Les, Z., Shape understanding system: understanding the Thin Object. *An International Journal Computers and Graphics*, 2002. **26**(6): pp. 951–970
80. Les, Z., and Les, M., Shape understanding system: the visual reasoning process. *International Journal of Pattern Recognition and Artificial Intelligence*, 2002. **17**(4): pp. 663–683
81. Les, Z., and Les, M., Shape understanding system: understanding of the convex objects. *The Journal of Electronic Imaging*, 2003. **12**(2): pp. 327–341
82. Les, Z., and Les, M., Understanding as an interpretation in shape understanding system. *Journal of Experimental and Theoretical Artificial Intelligence*, 2003. **15**(4): pp. 433–459
83. Les, Z., and Les, M., Shape understanding as knowledge generation. *Transaction on Data and Knowledge Engineering*, 2004. **16**(3): pp. 343–353
84. Les, Z., and Les, M., Understanding in shape understanding system. *International Journal of Pattern Recognition and Artificial Intelligence*, 2004. **18**(4): pp. 727–745
85. Les, Z., and Les, M., Understanding of the curve polygon object. *Journal of Computers and Graphics*, 2004. **28**(6): pp. 919–936
86. Les, Z., and Les, M., Shape understanding system-the system of experts. *International Journal of Intelligent Systems*, 2004. **19**(10): pp. 949–978
87. Les, Z., and Les, M., Shape understanding system-communication in natural language, in *IASTED Signal and Image Processing Conference*. 2004. Honolulu, Hawaii

88. Les, Z., and Les, M., Shape understanding system: visual intelligence tests. *International Journal of Intelligent Systems*, 2005. **20**(6): pp. 799–826
89. Les, Z., and Les, M., Understanding of a concave polygon object in shape understanding system. *Journal of Computer and Graphics*, 2005. **29**(3): pp. 365–378
90. Les, Z., and Les, M., Shape understanding system: understanding of the complex object. *The Journal of Electronic Imaging*, 2005. **14**(2): pp. 023015-1–023015-13
91. Les, Z., and Les, M., Shape understanding system: 3-D interpretation of the visual object, in *The IASTED International Conference on Signal and Image Processing (SIP)*. 2005, Honolulu, Hawai
92. Les, Z., and Les, M., SUS: understanding as the process of interpretation of signs and letters, in *The IASTED International Conference Artificial Intelligence and Soft Computing*. 2005, Benidorm, Spain
93. Les, Z., and Les, M., SUS a new generation of thinking robots-the visual intelligence test, in *International Conference on Informatics in Control, Automation and Robotics (ICINCO)*. 2005, Barcelona
94. Les, Z., and Les, M., SUS: learning of the categories of objects, in *WSEAS International Conference SPR'06*. 2006, Madrid
95. Les, Z., and Les, M., SUS: categorical learning as a new epistemologically oriented knowledge acquisition. *WSEAS Transaction on Computers*, 2006. **5**(3): pp. 653–660
96. Les, Z., and Les, M., SUS: 3D interpretation as a part of the visual concept formation. *Machine Graphics and Vision*, 2006. **15**(2): pp. 141–175
97. Les, Z., and Les, M., Understanding of the cyclic object as part of the visual understanding in shape understanding system. *Journal of Computer and Graphics*, 2006. **30**(5)
98. Quinlan, J.R., Induction of decision trees. *Machine Learning*, 1986. **1**: pp. 81–106
99. Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J., *Classification and Regression Tree*. 1984, Belmont, CA: Wadsworth International Group
100. Michalski, R., Theory and methodology of inductive learning. *Artificial Intelligence*, 1983. **20**: pp. 111–161
101. Minton, S., Carbonell, J.G., Knoblock, C.A., Kuokka, D,R., Etzioni, O., and Gil, Y., Explanation-based learning: a problem solving perspective, in *Machine Learning: Paradigms and Methods*, J.G. Carbonell, Editor. 1990, Cambridge, MA, pp. 63–118
102. Sestito, S., and Dillon, T.S., *Automated Knowledge Acquisition*. 1994, New York: Prentice Hall
103. Michalski, R.S., Carbonell, J.G., and Mitchel, T.M., ed. *Machine Learning: An Artificial Intelligence Approach*, Vol. 1. 1983, San Mateo: Tioga Publishing
104. Michalski, R.S., Carbonell, J.G., and Mitchel, T.M., ed. *Machine Learning: An Artificial Intelligence Approach*. 1986, San Mateo: Morgan Kaufman
105. Bramer, M.A., *Research and Development in Expert Systems III*. 1987, London: British Informatics Society Ltd

106. Brownston, L., Forrel R., Kant, E., and Martin, N., *Programming Expert Systems in OPS5. An Introduction to Rule-Based Programming*. 1985, Reading, MA: Addison-Wesley
107. Gupta, M., Kandel, A., Bandler, W., and Kiszka, J.B., *Approximate Reasoning in Expert Systems*. 1988, Amsterdam: North-Holland
108. Negoita, C.V., *Experts Systems and Fuzzy Systems*. 1985, Menlo Park: Cummings Publishing Company
109. Les, Z., Objects recognition system. *Bulletin of St. Staszic University of Mining and Metallurgy, Cracow*, 1992. **60**: pp. 33–48
110. Kowalski, R., *Logic for Problem Solving*. 1979, Oxford: North-Holland
111. Grabowski, J., Lesconne, P., and Wechler, W., *Algebraic and Logic Programming*. 1988, Berlin: Akademie
112. Bratko, I., and Mugletton, S., Application of inductive logic programming. *Communications ACM*, 1995. **38**(11): pp. 65–70
113. Buchanan, B.G., and Shortlife, E.H., *Rule-Based Expert System. The MYCIN Experiments of the Stanford Heuristic. Programming Project*. 1984, New York: Addison-Wesley
114. Jensen, F.V., *An Introduction to Bayesian Networks*. 1996, Berlin Heidelberg New York: Springer
115. Muller, J.P., Control architecture for autonomous and interacting agents: a survey, in *Intelligent Agent Systems*, L. Cavedon, Rao, A., and Wobcke, W., Editors. 1996, Berlin Heidelberg New York: Springer
116. Leake, D.B., ed. *Case-Based Reasoning. Experience, Lessons and Future Directions*. 1996, Menlo Park: AAAI Press/MIT Press
117. Fodor, J.A., *RePresentations: Philosophical Essays on the Foundations of Cognitive Science*. 1981, Cambridge, MA: MIT Press
118. Fodor, J.A., *The Language of Thought*. 1975, Cambridge, MA: Harvard University Press
119. Peacocke, C., *A Study of Concepts*. 1992, Cambridge, MA: MIT Press
120. Carey, S., *Conceptual Change in Childhood*. 1985, Cambridge, MA: MIT Press
121. Millikan, R., *On Clear and Confused Ideas*. 2000, Cambridge, MA: Cambridge University Press
122. Watson, J.B., Image and affection in behavior. *Journal of Philosophy, Psychology and Scientific Methods*, 1913. **10**: pp. 421–428
123. Searle, J., *Intentionality: An Essay in the Philosophy of Mind*. 1983, Cambridge, MA: Cambridge University Press
124. Martin, C.B., Proto-language. *Australasian Journal of Philosophy*, 1987. **65**: pp. 277–289
125. Davidson, D., Rational animals. *Dialectica*, 1982. **4**: pp. 317–327
126. Binford, T.O., Survey of model-based image analysis systems. *International Journal of Robotics Research*, 1982. **1**(1): pp. 18–64
127. Matsuyama, T., and Hwang, V., *SIGMA: A Knowledge-Based Aerial Image Understanding System*. 1990, New York: Plenum Press
128. Ullman, S., and Richards, W., *Image Understanding*. 1989, Norwood: Ablex Publishing Corporation

129. Shirai, Y., *Three-Dimensional Computer Vision*. 1987, Berlin Heidelberg New York: Springer
130. Overington, I., *Computer Vision*. 1992, Amsterdam: Elsevier Science Publisher
131. Wechsler, H., *Computational Vision*. 1990, London: Academic Press
132. Les, Z., Shape understanding based on the multi-level representation, in *The IASTED International Conference on Computer Graphics & Imaging*. 1999, Palm Springs: ACTA Press, Anaheim
133. Les, Z., Shape understanding as a modelling of the cognitive processes, in *The Cognitive Science Conference*. 2000, Melbourne
134. Les, Z., and Les, M., System of experts to perform an epistemological task of shape understanding, in *The IASTED International Conference on Software Engineering*. 2000, Las Vegas
135. Les, Z., and Tadeusiewicz, R., Shape understanding system – generating exemplars of the polygon class, in *The IASTED International Conference on Computer Graphics & Imaging*. 2000, Las Vegas: ACTA Press, Anaheim
136. Les, Z., and Les, M., Understanding as a new paradigm in the design of the intelligent systems in automotive and transportation technology, in *Automotive and Transportation Technology International Conference, ATT*. 2001, Barcelona
137. Les, Z., Tadeusiewicz, R., and Les, M. Shape understanding: knowledge generation and learning, in *Seventh Australian and New Zealand Intelligent Information Systems Conference*. 2001, Perth
138. Les, Z., and Les, M., Shape understanding system: understanding of the cyclic object, in *IASTED Conference on Signal Processing and Imaging*. 2003, Rodos, Greece: ACTA Press, Anaheim
139. Les, Z., and Les, M., Understanding of the concave-complex object, in *IASTED Conference Visualization, Imaging, and Image Processing*. 2003, Benalmadena, Spain: ACTA Press, Anaheim
140. Les, Z., and Les, M., Shape understanding system: system of experts for visual thinking, in *The IASTED International Conference on Software Engineering*. 2003, Marina Del Rey, USA: ACTA Press, Anaheim
141. Les, Z., and Les, M., Visual thinking as a part of understanding process – a new way of problem solving and communication of the intelligent systems in automotive and transportation technology. *SAE 2002 Transactions – Journal of Passenger Cars: Electronic and Electrical Systems*, 2003: pp. 893–899
142. Les, Z., and Les, M., Shape understanding system: understanding of the complex thin object, in *IASTED Conference Computer Graphics and Imaging*. 2004, Kauai, Hawaii
143. Les, Z., and Les, M., Visual Intelligence tests in shape understanding system, in *IASTED Visualization, Imaging, and Image Processing*. 2004.
144. Les, Z., and Les, M., Understanding of the irregular object, in *55th Session of the International Statistical Institute (ISI)*. 2005, Sydney
145. Les, Z., and Les, M., SUS: interpretation of the mathematical objects, in *The IASTED International Conference on Computational Intelligence*. 2005, Calgary
146. Les, Z., and Les, M., Testing visual abilities of machines – visual intelligence tests, in *Enformatica Conference*. 2005, Praha

2. Shape Classes

2.1. Possible Classes of Shape

The proposed method of shape understanding is based on the concept of shape classes that are understood as the basic perceptual categories. The Shape Understanding System (SUS) perceives the visual object by trying to fit it into one of the shape categories. Although shape is one of the most often perceived "properties" of the visual object, there is no satisfactory classification and definition of shape. An attempt to develop the system of shape classification that is based on the shape classes was made by Les [1]. Shape classes called shape categories (in the context of visual thinking) are used as the "material" of the visual thinking process. The shape classes are represented by the symbolic names and are defined in the context of visual understanding process. Each class is related to each other and based on relationships among classes there is relatively easy to establish the "perceptual similarity" of visual objects.

In this chapter, the description of the shape classes is presented within the framework of shape understanding method. Shape understanding method is based on the concept of possible classes of shape [1]. A member of the class that is defined in terms of its attributes is called an *archetype* of this class. In the case of a digital image, the shape is given as an image region or a set of pixels. A perceived object (phantom) is transformed into a digital representation called a digital object. The proper interpretation of the visual object is obtained during the visual reasoning process. During the visual reasoning the perceived object is transformed into its symbolic description called the symbolic name. The symbolic name is the name of the shape category (shape classes) to which the shape of the perceived object is fitted. The symbolic name is used to find the visual concept and to assign the perceived object to one of the ontological categories. The visual concept is a set of symbolic names obtained in the learning process. The shape class is denoted by symbol Ω^{η}, where η denotes the symbolic

Z. Les and M. Les: *Shape Classes*, Studies in Computational Intelligence (SCI) **86**, 47–99 (2008)
www.springerlink.com

description (the symbolic name) of a given class. A member of the class denoted by symbol ω is called an archetype.

In this book, for simplicity, the symbol of the class Ω is omitted and the class is often described by its symbolic name, e.g., Λ instead of Ω^{Λ} or $Q_{\Im}^{n}[\Lambda](n\Im_{\Gamma})$ instead of $\Omega^{Q_{\Im\Lambda}^{n}(n\Im_{\Gamma})}$. Also n classes $\underbrace{\Im_1,\ldots,\Im_n}_{n}$ that are identical $\Im_i \equiv \Im_j$ for all $i = 1,\ldots,n$, $j = 1,\ldots,n$, and $i \neq j$ are denoted as $n \cdot \Im$, whereas n classes $\underbrace{\Im_1,\ldots,\Im_n}_{n}$ that are not identical are denoted as $n\Im$.

The general shape classes are defined based on the general attributes of shape such as homotopy, convexity, or thickness. The general class is split into specific classes based on additional features that represent a priori information about local perceptual and geometrical properties of shape and is incorporated into the a priori model of the shape class. The deepness of the splitting process depends on the base class from which the specific class is derived. In this book, the following general classes are presented: cyclic–acyclic general classes $\Omega^{A} - \Omega^{\Gamma}$, convex–concave general classes $\Omega^{\Lambda} - \Omega^{Q}$, and thick–thin general classes $\Omega^{N} - \Omega^{\Theta}$.

2.1.1. General Classes: A Priori Classes

The homotopy measure that is based on the computation of a number of holes is applied to derive the cyclic–acyclic general classes $\Omega^{A} - \Omega^{\Gamma}$. An element of the shape class, called an archetype, is called acyclic, if its homology groups $H_i(X)$ coincides with homology groups of a point, i.e., the Betti numbers b_1, b_2 are equal to 0, and $b_0 = 1$, where $b_i =$ rank $H_i(X)$. The 0th Betti number b_0 stands for a number of components, while b_1 denotes a number of wholes in shape. The derivation rule for the cyclic general class is given as follows: where $[a^{H} = 0] \Rightarrow \Omega \prec \Omega^{A}$, a^{H} denotes an attribute called homotopy and symbol $\prec$ denotes that class Ω^{A} is derived from the class Ω.

The convexity coefficient that is given as the ratio of the area of the object A_{ω} to the area of the convex hull $A_{\aleph}$, $a^{\aleph} = A_{\aleph} / A_{\omega}$ is used to derive the convex Ω^{Λ} class and the concave Ω^{Q} class. The convex hull of a set of points X in the plane is the smallest convex polygon P that encloses X, smallest in the sense that there is no other polygon P' such that $P \supset P' \supseteq X$. The computation of the convex hull of a finite set of points in the plane has been studied extensively and some of the algorithms as well as discussion of the complexity of the convex hull algorithms can be found in [2, 3]. The derivation rule of the convex class is given as follows: $[a^{\aleph} = 0] \Rightarrow \Omega \prec \Omega^{\Lambda}$, where $a^{\aleph}$ is an attribute of the class. The convex general class Ω^{Λ} is related to the notion of a convex set (see, e.g., [4]). A set X in E^2 is convex if for any two points $x, y \in X$, the (closed) segment $\overline{xy}$ is wholly contained in this set ($\overline{xy} \subseteq X$) or, in another way, a set X is called the convex set if for any two points of this set the following relation takes place: $\lambda x + (1-\lambda) y \in X$, for each $\lambda \in [0,1]$.

The thin class Ω^{Θ} is a class whose members are thin objects. The description of the object in terms of thickness can be obtained utilizing a distance transformation. The distance transformation is a mapping of a set of points into a set of predefined distances (see, e.g., [5]). The distance transformation (the thickness measure) is the image transformation points-number $\Theta^{\hbar}$ described in [6] that assigns the number to each point $u_i^F \in \coprod^F$ based on the local properties and is given as follows:

$$\left[\forall u_i^F \in \coprod^F, \exists \sigma_i^{\rho} \in \Re : \sigma_i^{\rho} = \hbar_{\rho}\left(u_i^F\right)\right] \Rightarrow u_i \rhd \sigma_i^{\rho},$$

where the local transformation $\hbar_{\rho}\left(u_i^F\right)$ is determined by the selected neighborhood. In the case of a distance transformation the local transformation is given as

$$\hbar_{\rho}\left(u_i^F\right) \equiv \min_{u_k^F \in \coprod^F} | u_k^F u_i^F |,$$

where $| u_k^F u_i^F |$ denotes distance between a point u_i^F and an arbitrary point $u_k^F \in \coprod^F$. The detail description of the image transformations $\Theta^{\hbar}$ is given in Chap. 3. The thin general class Ω^{Θ} is derived based on the thickness measure which is the attribute of this class. The derivation rule for the thin class Ω^{Θ} is given as follows: $[a^{\rho} \le \theta] \Rightarrow \Omega \prec \Omega^{\Theta}$, where a^{ρ} denotes a thickness measure and θ is the threshold.

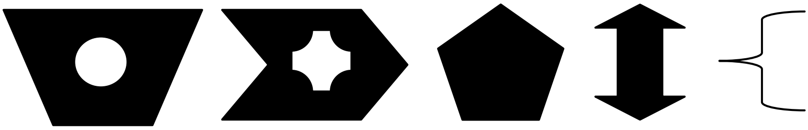


Fig. 2.1. Examples of exemplars of the selected general classes (**a**–**b**) cyclic, (**c**) convex, (**d**) concave, (**e**) thin

In Fig. 2.1 exemplars of the four general classes, the cyclic class (Fig. 2.1a, b), the convex class (Fig. 2.1c), the concave class (Fig. 2.1d), and the thin class (Fig. 2.1e), are shown.

In the further parts of this book the description of the selected shape classes is presented. The a priori classes such as the convex polygon class or the concave polygon class are derived from the general class. The a posteriori classes such as the star class or the spade class are derived from the specific a priori classes.

2.1.1.1. Convex Classes

2.1.1.1.1. Convex Polygon Class and Its Subclasses

The convex polygon class Ω^L consists of elements that are called the convex polygons. A polygon is a simple closed plane figure that is bounded by a finite number of intersecting line segments (at least three segments are required). The polygon $p:[0,1] \rightarrow R^2$ is a piecewise linear continuous function. The convex polygon class Ω^L is derived from the convex general class Ω^A by assigning the value 0 to the curvature $\kappa(t)$ of the border curve. The curvature κ at P_0 for a continuous function Γ is defined as the instantaneous rate of change tangent angle with respect to the arc length

$$\kappa = \lim_{P_0 \rightarrow P_1} \frac{\alpha(P_1) - \alpha(P_0)}{\bar{P_0 P_1}},$$

where $\alpha(P_1)$ is the angle between the positive x-axis and the direction of the tangent line at a point P_0 and $\bar{P_0 P_1}$ is the arc length between P_0 and P_1. The detail description of the curvature in the context of the concepts of the differential geometry can be found in [7]. In the case when curve is given by the parameterized form $g(t)=\{x(t),y(t)\}$ with parameter $t\in\Theta\subset\Re$, the curvature is expressed in terms of derivatives of the curve as follows

$$\kappa=\frac{\dot{x}\ddot{y}-\dot{y}\ddot{x}}{(\dot{x}^2+\dot{y}^2)^{3/2}}.$$

Several methods of the curvature computation were proposed. For example, curvature as the change of cosine over a region of support is given in [8], the curvature as the rate of change of slope expressed as a function of length is described in [9], or the curvature as a convolution with a Gaussian kernel is described in [10].

The convex polygon class Ω^L is given by the following derivation rule: $[\forall t\in[t_i,t_{i+1}]:\kappa(t)=0]\Rightarrow\Omega^A\succ\Omega^L$. Here, $[t_i,t_{i+1}]$ is an interval where the first derivative of the polygon curve given by the equations $x=x(t),y=y(t)$ exists.

The convex polygon class Ω^L is split into base convex polygon classes based on the derivation rules

$$\left[\exists n\in N:n=a^v\right]\Rightarrow\Omega^L\succ\Omega^{L^n},$$

where $a^v=|V|$ denotes the attribute of the class (the cardinality of the set of vertices V). A mathematical object is a cardinal number (cardinality of a set) if and only if it is a power of a set [11]. For the set V of vertices, its cardinality is denoted as $|V|$. The classes with n = 3, 4, 5, and 6 (number of sides) are denoted by the symbolic class description L^n as follows: L^3 (triangle class), L^4 (quadrilateral class), L^5 (pentagon class), and L^6 (hexagon class).

The class L^n is split into specific classes against the relations between selected attributes $\left(a_i^d,a_i^\alpha\right)$. For example, the right triangle class L_R^3 is the class whose archetypes are triangles with one interior angle that is equal to 90°. The derivation of the right triangle class L_R^3 from the triangle class L^3 is given by the following rule

$$\left[\exists a_i^\alpha\in A^G:a_i^\alpha=\frac{\pi}{2}\right]\Rightarrow L^3\succ L_R^3.$$

2.1.1.1.2. Convex Curve-Polygon Class and Its Subclasses

The convex curve-polygon class Ω^M consists of the geometrical figures, which have curvilinear parts as well as linear segments. The curve-polygon class Ω^M is defined against the value of the curvature $\kappa(t)$ as follows

$$[\exists i : \forall t \in (t_i, t_{i+1}), \kappa(t) = 0] \Rightarrow \Omega^A \succ \Omega^M,$$

where t_i ($i = 1,\ldots,N$) is the value of a parameter for which the curvature $\kappa(t)$ does not exist.

Splitting of the convex curve-polygon class M into the base classes is based on a number of straight line segments and a number of curvilinear segments m of archetypes of the class M. The description of the base convex curve-polygon class is related to the generic polygon class L^n. Archetype of the generic polygon class L^n is constructed by joining vertices of the straight line segments as shown in Fig. 2.2. Archetype shown in Fig. 2.2a is a member of the curve-linear class $M^1[L^4]$, where 1 denotes one curvilinear segment and L^4 denotes the generic polygon (rectangle Fig. 2.2b). Examples of the archetypes of the base convex curve-polygon classes are shown in Fig. 2.2. The symbolic names for archetypes shown in Fig. 2.2 are as follows: $M^1[L^3]$ (Fig. 2.2c), $M^1[L^4]$ (Fig. 2.2d), $M^2[L^4]$ (Fig. 2.2e), $M^1[L^5]$ (Fig. 2.2f), and $M^1[L^6]$ (Fig. 2.2g). Construction of the generic polygon is presented in Fig. 2.2a–b. The generic polygon is obtained by joining straight line segment vertices.

The class $M^m[L^n]$ is split into specific classes based on the type of the curvilinear segment and the description of the specific curve-polygon class is given in the form $M^m[L^n](m\theta_H^f)$, where L^n is a generic polygon class,

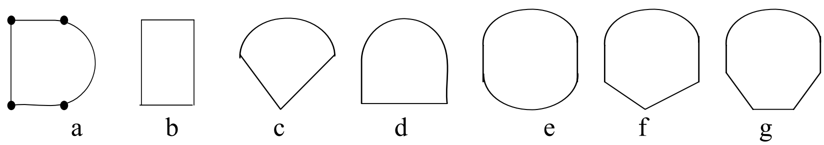


Fig. 2.2. Construction of the generic polygon: (**a**) an archetype of the convex curve-polygon class, (**b**) the generic polygon obtained by joining straight line segment vertices. Examples of archetypes of the convex polygon-curve class (**c–g**)

m is a number of curvilinear segments, and θ_H^f denotes a type of the curvilinear segment. Each symbol of the type of the curvilinear segment θ_H^f has its meaning: θ denotes convexity of the curvilinear segment $\theta \in [c, w]$, where c is a convex curvilinear segment and w is a concave curvilinear segment; f denotes the curvilinear segment $f \in [0,1,2]$, where 0 denotes a "function," 1 denotes a "nonfunction" only on one side, and 2 denotes a "nonfunction" on both sides. The "function" is a curvilinear segment that is the graphical representation of any function $y = f(x)$. H denotes the height of the curvilinear segment $H \in [0,1,2]$, where 0 indicates a low height segment, 1 indicates a medium height segment, and 2 indicates a high segment. The height is the perpendicular distance from the chord connecting the endpoints of a curvilinear segment to the farthest point on the curvilinear segment. The symmetrical curvilinear segment is denoted as $\overline{\theta}_H^f$.

Archetypes of the class M^1 possess only one straight line segment and one curvilinear segment. The description of the specific class derived from the class M^1 is given in the form $M^1\left(\theta_H^f\right)$. The examples of exemplars generated from the class M^1 are given in Fig. 2.3. The symbolic names of the exemplars shown in Fig. 2.3 are as follows: $M^1\left(c_1^2\right)$ (Fig. 2.3a), $M^1\left(c_1^1\right)$ (Fig. 2.3b), $M^1\left(c_1^0\right)$ (Fig. 2.3c), $M^1\left(\overline{c}_2^2\right)$ (Fig. 2.3d), $M^1\left(\overline{c}_1^2\right)$ (Fig. 2.3e), and $M^1\left(\overline{c}_0^2\right)$ (Fig. 2.3f).

Archetypes of the class $M^1[L^3]$ possess two straight line segments and one curvilinear segment. The description of the specific class derived from

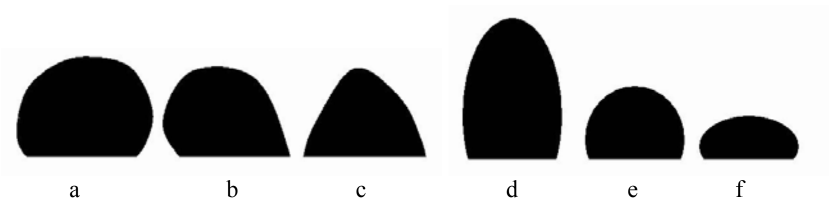


Fig. 2.3. Exemplars generated from the class Ω^{M_1}: (**a**) $M^1\left(c_1^2\right)$, (**b**) $M^1\left(c_1^1\right)$, (**c**) $M^1\left(c_1^0\right)$, (**d**) $M^1\left(\overline{c}_2^2\right)$, (**e**) $M^1\left(\overline{c}_1^2\right)$, (**f**) $M^1\left(\overline{c}_0^2\right)$

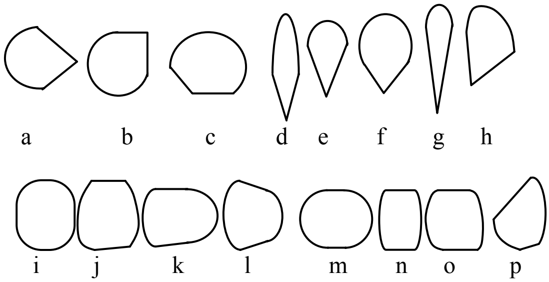


Fig. 2.4. Archetypes of the class $M^1[L^3](\theta_H^f)$ (**a**–**h**), $M^2[L^4](\theta_H^f\theta_H^f)$ (**i**–**p**)

the class $M^1[L^3]$ is given in the form $M^1[L^3](\theta_H^f)$, where θ_H^f denotes the type of the curvilinear segment. Archetypes of the class $M^1[L^3](\theta_H^f)$ in Fig. 2.4(a–h).

Archetypes of the class $M^2[L^4]$ possess two straight line segments and two curvilinear segments (see Fig. 2.4(i–p)). The specific class derived from the class $M^2[L^4]$ is given in the form $M^2[L^4](\theta_H^f\theta_H^f)$, where θ_H^f denotes the type of the curvilinear segment and L^4 denotes the generic polygon.

2.1.1.1.3. Convex Curve Class and Its Subclasses

The convex curve class Ω^K consists of convex curves. A convex curve in $E2$ can be described in many different forms: an implicit equation $F(x,y,z) = 0$, a parametric equation $(x(t),y(t))$, the parametric Fourier equations, parametric B-splines, or wavelets. The approximated forms of curve representation, such as Fourier series, cubic-splines, B-splines, β-splines, and wavelets, are often used in geometric modeling (e.g., [12]) and are most promising as a model for the convex curve class. The Fourier series can be seen also as a definition of a curve in the parametric form. The curve Γ can be expressed in the form of its truncated Fourier series as follows:

$$x(t) = a_0 + \sum_{n=1}^{\infty}\left(a_n \cos\frac{2\pi nt}{T} + b_n \sin\frac{2\pi nt}{T}\right),$$

$$y(t) = c_0 + \sum_{n=1}^{\infty}\left(c_n \cos\frac{2\pi nt}{T} + d_n \sin\frac{2\pi nt}{T}\right).$$

The coefficients a_n, b_n, c_n, and d_n are computed as described in Brigham [13].

The equation for a single parametric cubic spline segment is given by

$$P(t) = \sum_{i=1}^{4} B_i t^{i-1}, \quad t_1 \le t \le t_2, \, t_1 \le t \le t_2$$

where t_1 and t_2 are the values of parameters at the beginning and at the end of the segment. $P(t)$ is the position vector of any points on their cubic spline segment. The curve can be computed as

$$C_x^k(t) = \sum_{i=0}^{3} A_{ik} t^i, \quad C_y^k(t) = \sum_{i=0}^{3} B_{ik} t^i, \quad t_1 \le t \le t_2 .$$

The constant coefficients A_{ik} and B_{ik} are determined by specifying four boundary conditions for the spline segment [12].

B-splines are given by the parametric equation

$$f(t) = \sum_{i=0}^{n} b_{i,k}(t) \overline{q}_i,$$

where $\overline{q}_0, \overline{q}_1, \ldots, \overline{q}_n$ are $n + 1$ control points. The index $k = 2,3,...$, determines the number of control points that have influence on the points of the curve [14].

Discrete parameter wavelet transform DPWT [15] can be used to represent curve $f(t)$ as follows:

$$f(t) = c \sum_m \sum_n \mathrm{DPWT}(m,n) \Psi_{m,n}(t),$$

where c is some constant dependent on $\Psi(t)$. The discrete parameter wavelet transform is given by $\mathrm{DPWT}(m,n) = \int f(t) \Psi_{m,n}(t) dt$, where $\Psi_{m,n}(t) = a_0^{-m/2} \Psi\left(a_0^{-m} t - n\tau_0\right)$, $\Psi_{0,0}(t) = \Psi(t)$, and a_0 and τ_0 are constants that determine the sampling intervals.

It can be proved (see, e.g., [7]) that a unit-speed curve $\hat{f} : (a,b) \to \Re^2$ whose curvature is given as a piecewise continuous function $\kappa : (a,b) \to \Re^2$ is given by

$$\begin{cases} \hat{f}(s) = \left(\int \cos\theta(s)\mathrm{d}s + c \int \sin\theta(s)ds + d \right) \\ \theta(s) = \int \kappa(s)ds + \theta_0 \end{cases}, \tag{1}$$

where *c*, *d*, and θ_0 are integration constants. The curvature given as a piecewise continuous function κ is characteristic for the curvilinear segment. Equation (1) can be used as a model of the curve class. However to derive the specific classes the heuristic rules are applied that make it possible to define the classes based on more perceptually oriented approach. The curve class Ω^K is defined by using a curvature and is given by the derivation rules $[\forall t \in [t_1, t_2] : \kappa(t) > 0] \Rightarrow \Omega^A \succ \Omega^K$, where parameter *t* varies over a given range $t \in [t_1, t_2]$.

The class K^1 is the class whose archetypes are regular curves. The regular curve is a curve that is convex and symmetrical. Archetypes of the convex curve class K^1 are defined by the ellipse equation

$$\frac{x^2}{a} + \frac{y^2}{b} = 1,$$

which is parameterized by two parameters *a* and *b*. The curvature of the ellipse is given by the equation

$$\frac{ab}{\left(b^2 \cos^2 t + a^2 \sin^2 t\right)^{3/2}}.$$

From the convex curve class K^1 the specific classes, the circle class K_C^1 (Fig. 2.5a) and the ellipsis class K_E^1 (Fig. 2.5b, c), are derived.

The class K^2 is a class for which curvature of each archetype has one clear maximum and each archetype is symmetrical. The maximum of the curvature is the point $t \in [t_1, t_2]$ for which the first derivative of the curvature $\kappa'(t) = 0$. The maximum of the curvature is denoted as $\kappa_{\max}$ and $\kappa_{\max}$ is the attribute of the class K^2. The derivation rule of the convex class K^2 is given as

$$\left[a^{\kappa_{\max}} > \theta_h^{\kappa} \right] \Rightarrow \Omega^K \succ \Omega^{K_2},$$

where θ_h^{κ} is the threshold. Archetypes generated from convex curve class K^2 are shown in Fig. 2.5d, e.

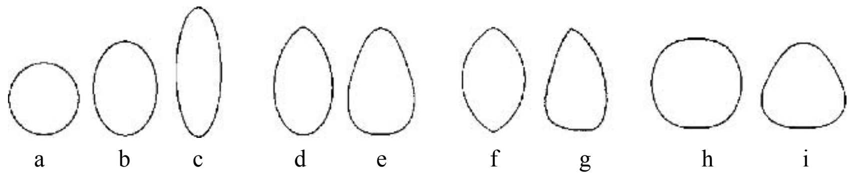


Fig. 2.5. Archetypes of the convex curve class

The class K^3 is described as a class for which curvature of each archetype is in the range $\theta_l^\kappa < \kappa(t) < \theta_h^\kappa$, where θ_l^κ and θ_h^κ are thresholds. The derivation rules of the convex class K^3 are given in the form

$$\left[a_{\max}^\kappa < \theta_h^\kappa \wedge a_{\min}^\kappa > \theta_l^\kappa\right] \Rightarrow \Omega^K \succ \Omega^{K_3},$$

where $a_{\max}^\kappa$, $a_{\min}^\kappa$ are attributes of the convex curve class K^3, and $\theta_l^\kappa, \theta_h^\kappa$ are thresholds. Archetypes generated from convex curve class K^3 are shown in Fig. 2.5f, g.

The convex curve class K^4 is derived from convex class based on derivation rules given in the form

$$\left[a_{\max}^\kappa < \theta_l^\kappa \wedge a^L > \theta_l^L\right] \Rightarrow \Omega^K \succ \Omega^{K_4},$$

where curvature $a_{\max}^\kappa$, and elongation a^L are attributes of the convex curve class Ω^K and $\theta_l^\kappa, \theta_l^L$ are thresholds. Archetypes generated from the convex curve class K^4 are shown in Fig. 2.5h, i. Elongation L is defined as $L = \lambda_1/\lambda_2$, where λ_1 and λ_2 are the first and second eigenvalues of the matrix of the first and second moments

$$\begin{bmatrix} m_{20} & m_{11} \\ m_{11} & m_{02} \end{bmatrix}, \;\; where\, m_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q dx\, dy, \;\; and \;\; p,q \in [0,1,2].$$

2.1.1.2. Concave Classes

In the section "Convex Curve Class and Its Subclasses" the specific convex classes were described. In this section the specific concave classes, derived from the concave general class, are presented. The process of derivation of the concave general class Q was described in the previous chapters. The archetype of the concave class Q consists of elements that can be decomposed into subregions (residuals) iteratively. In decomposition scheme the concave object is broken down into very simple primitives called residuals.

Fig. 2.6. Process of decomposition of the archetype of the concave class: (**a**) an archetype of the concave class $Q^1[L^4](L^4)$, (**b**) the generic convex class L^4, (**c**) residual L^4

At first the convex hull is used as a base for the decomposition of the object into the concave regions and residuals and next each residual is examined in the process called the first level of iteration (see Fig. 2.6). In the case when some residuals are concave they are examined in the process called the second level of iteration. The description of the concave class depends on the level of iteration and is given by a symbolic name $Q^n[\Im_\Lambda](n\Im_\Gamma)$, where n is the number of residuals, $\Im_\Gamma$ is a type of the residuals, $\Im_\Lambda$ is a type of the generic classes, $\Im_\Lambda$ is one of the convex classes $\Im_\Lambda \equiv \{L, K, M\}$, and $\Im_\Gamma$ is one of the acyclic generic classes $\Im_\Gamma = \{\Lambda, Q, \Theta\}$.

The convex hull shown in Fig. 2.6b is used as a base for the decomposition of the object into the concave regions and residuals and is called the generic convex object. The generic convex object is a member of the convex rectangular class L^4. As it was described in decomposition scheme, the concave object is broken down into very simple primitives called residuals. Figure 2.6 shows the process of decomposition of the concave object (a) an archetype of the concave class $Q^1[L^4](L^4)$ (Fig. 2.6a), (b) the generic convex class L^4 (Fig. 2.6b), and (c) residual member of the rectangular class L^4 (Fig. 2.6c).

2.1.1.2.1. Levels of Iterations

As it was described, the description of the concave class depends on the level of iteration, the number of residuals n, type of the residuals $\Im_\Gamma$, and type of the generic class $\Im_\Lambda$. The description of the concave class at the first level of iteration is given by $Q^n[\Im_\Lambda](n\Im_\Gamma)$, where $\Im_\Lambda$ is one of the convex classes $\Im_\Lambda = \{L, K, M\}$ and $\Im_\Gamma$ is one of the acyclic general classes $\Im_\Gamma = \{\Lambda, \Theta\}$. Depending on the number of residuals n, and type of classes $\Im_\Lambda$ and $\Im_\Gamma$ the following concave classes are possible: $Q_\Lambda^n(n\Lambda)$, $Q_\Lambda^n(n\Theta)$, or $Q_\Lambda^n(k\Lambda m\Theta)$, where $k+m=n$.

In the case when the generic class is the convex polygon class L the class that is derived is given by $Q_L^n(n\Lambda)$. The symbol $Q_L^n(n\Lambda)$ denotes the concave class Q whose generic class is the convex polygon class L and archetypes of this class have n residuals. All residuals are archetypes of one of the convex classes (the polygon class L, the convex polygon-curve class M, or the convex curve class K). The following concave classes are possible: $Q_L^n(nL)$, $Q_L^n(nM)$, $Q_L^n(nK)$, $Q_L^n(kLmM)$, $Q_L^n(kLmK)$, $Q_L^n(kMmK)$, or $Q_L^n(hLkMmK)$, where $k+m=n$ and $h+k+m=n$. Similarly, the possible classes whose generic class is the convex curve-polygon class $Q_M^n(n\Lambda)$ or the convex curve class $Q_K^n(n\Lambda)$ can be obtained. The symbol $Q^2[L^5](2\cdot L^3)$ denotes that the concave class Q whose generic class is the convex polygon class (pentagon) L^5 has two residuals. Both residuals are archetypes of the triangle class L^3. Examples of the concave class at the first level of iteration are given in Fig. 2.7.

Archetypes in Fig. 2.7a–c are members of the class $Q^n[L](n\Lambda)$, where the generic class is given by the convex polygon class L. The residuals are members of the convex polygon class L, the convex curve-polygon class M, or the convex class K. Archetypes are given by the following symbolic names: $Q^2[L^5](2\cdot L^3)$ (Fig. 2.7a), $Q[L^4](M)$ (Fig. 2.7b), and $Q[L^4](K^1)$ (Fig. 2.7c).

Archetypes in Fig. 2.7d–f are members of the class $Q_M^n(n\Lambda)$, where the generic class is given by the convex curve-polygon class M. The residuals are members of the convex polygon class L, the convex curve-polygon class M, and the convex class K. Archetypes are given by following symbolic names: $Q[M](L^3)$ (Fig. 2.7d), $Q[M](M)$ (Fig. 2.7e), and $Q[M](K^1)$ (Fig. 2.7f).

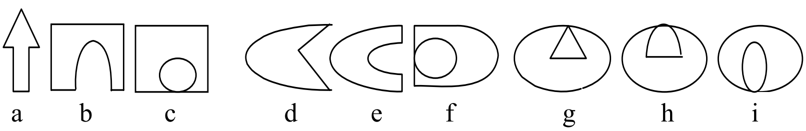


Fig. 2.7. Archetypes of the concave classes at the first level of iteration

Archetypes in Fig. 2.7g–i are members of the class $Q_K^n(n\Lambda)$ where the generic class is given by the convex curve class K. The residuals are members of the convex polygon class L, the convex curve-polygon class M, and the convex class K. Archetypes are given by the following symbolic names: $Q[K](L^3)$ (Fig. 2.7g), $Q[K](M)$ (Fig. 2.7h), and $Q[K](K^1)$ (Fig. 2.7i).

Similarly, at the second level of iteration the description of the concave class is given by $Q_{\Im_\Lambda}^n(n\Im_\Gamma)$, where $\Im_\Lambda$ is one of the convex classes $\Im_\Lambda = \{L, K, M\}$ and $\Im_\Gamma$ is one of the acyclic general classes $\Im_\Gamma = \{\Lambda, Q, \Theta\}$. Depending on the number of residuals n, type of the class $\Im_\Gamma$ and type of the generic convex class $\Im_\Lambda$ the following classes are possible: $Q^n[\Lambda](nQ)$, $Q^n[\Lambda](kQm\Lambda)$, $Q^n[\Lambda](kQm\Theta)$, or $Q^n[\Lambda](hQk\Lambda m\Theta)$, where $k+m=n$ and $h+k+m=n$. Examples of the archetypes of the concave classes at the second level of iteration are given in Fig. 2.8.

Archetypes in Fig. 2.8 are members of the class $Q[\Lambda](Q[\Lambda](\Lambda))$, where the generic class of each archetype is one of the following classes: the convex polygon class L, the convex curve-polygon class M, and the convex class K. The residuals are members of the concave class Q. Archetypes shown in Fig. 2.8 are given by following symbolic names: $Q^1[L^4](Q^1[M](M))$ (Fig. 2.8a), $Q^1[M](Q^1[L^4](L^3))$ (Fig. 2.8b), and $Q^1[K](Q^1[M](L^3))$ (Fig. 2.8c).

In the case when a number of iteration levels and the number of residuals are growing an archetype of the concave class can be described as an archetype of the thin class. Also for the class $Q_{L^m}^1(L^i)$ the following expression is true $\lim_{i\to\infty} Q_{L^m}^1(L^i) = Q_{L^m}^1(K)$, where K denotes the curvilinear class. For the convex polygon class $Q^n[L^m](nL^k)$, when n is large enough, the convex polygon class is called a noisy class and is denoted as

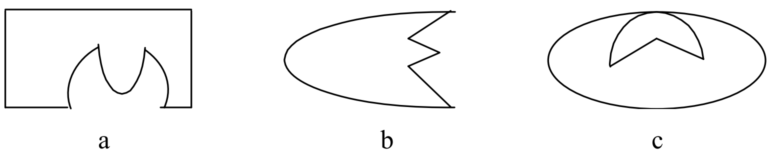


Fig. 2.8. Archetypes of the concave classes at the second level of iteration

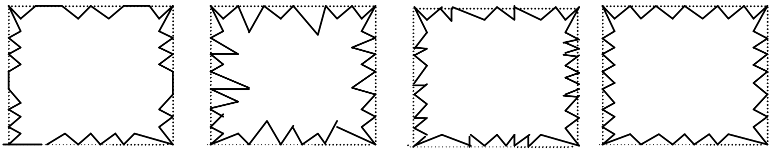


Fig. 2.9. Archetypes of the noisy polygon class

$\aleph[L^m](L^k) \equiv \lim_{n\to\infty} Q^n_{L^m}(nL^k)$. When all residuals are triangles ($k = 3$) the noisy class is denoted as $\aleph[L^m](n \cdot L^3)$ [16]. Examples of archetypes of the noisy class $\aleph[L^4](nL^3)$ are given in Fig. 2.9.

2.1.1.2.2. Concave Polygon Class

In the previous section the specific concave classes, derived from the general concave class, were described. In this section the subspecific concave classes, derived from the concave polygon class, are presented. The concave polygon class is the class archetypes of which are concave polygons. The concave polygon class at the first level of iteration is described as $Q^n[L^m](nL^k)$. For the concave polygon class $Q^n[L^m](nL^k)$ the generic class is the convex polygon class L and all residuals are archetypes of one of the convex polygon classes. The concave polygon class at the second level of iteration is described as $Q^n[L^m](nQ^h[L^m](hL^p))$ and the concave polygon class at the third level of iteration is given by symbolic name $Q^n[L^m](nQ^h[L^m](hQ^w[L^u](wL^s)))$. Example of the archetype generated from the concave polygon class at the third level of iteration given by symbolic name $Q^3[L^6](Q^3[L^5](Q^1[L^4](L^4), L^3, L^3), Q^1[L^4](L^4), Q^1[L^4](L^3))$ is shown in Fig. 2.10. The symbol $Q^3[L^6](Q^3[L^5](Q^1[L^4](L^4), 2L^3),$ $Q^1[L^4](L^4),\ Q^1[L^4](L^3))$ denotes the archetype of the concave class Q whose generic class is the archetype of the convex polygon class (hexagon) L^6 and the concave class is described at three levels of iteration. At the first level of iteration there are three residuals, archetypes of the concave classes $Q^3[L^5](Q^1[L^4](L^4),\ 2L^3)$, $Q^1[L^4](L^4)$, and $Q^1[L^4](L^3)$. At the second level of iteration each residual $Q^3[L^5](Q^1[L^4](L^4),\ 2L^3)$, $Q^1[L^4](L^4)$, and $Q^1[L^4](L^3)$ is considered as an archetype of the concave class whose generic classes are archetypes of the convex polygon classes L^5, L^4, and

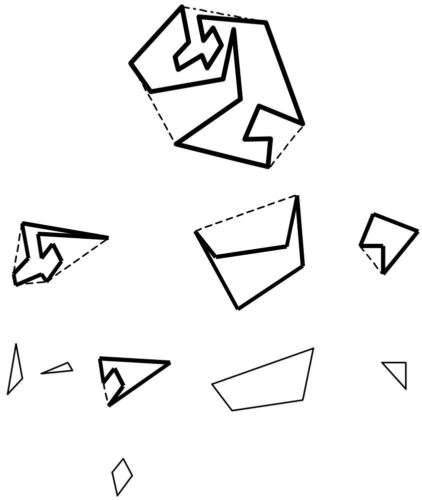
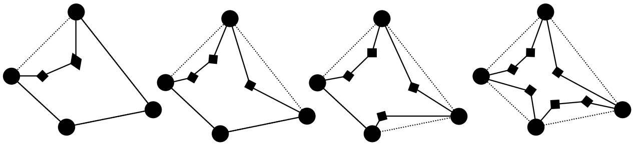


Fig. 2.10. The archetype of the class $Q^3[L^6](Q^3[L^5](Q^1[L^4](L^4), L^3, L^3)Q^1[L^4](L^4), Q^1[L^4](L^3))$

L^4. The archetype of the class $Q^3[L^5](Q^1[L^4](L^4), 2L^3)$ has three residuals $Q^1[L^4](L^4)$, L^3, and L^3. At the third level of iteration the residual $Q^1[L^4](L^4)$ is decomposed into the generic class L^4 and one residual L^4.

The concave polygon class can be described by applying the different symbolic descriptions. One of the descriptions is based on the computation of the convex and concave vertices. Let m denote the number of vertices of the generic convex polygon (convex vertices) of the archetype of the concave class $Q^n[L^m]$. Let n denote a number of residuals and h_i $(i = 1,\ldots,n)$ denotes a number of concave vertices w^i_j between two convex vertices v_i and v_{i+1}. To obtain description of this class in a more convenient way, let $v_i \equiv a$ denotes a convex vertex and $k_i \equiv \left\{w^i_1, w^i_2, \ldots, w^i_{h_i}\right\}$ denotes a set of concave vertices between two adjacent vertices v_i and v_{i+1} so as the class description can be represented by the string in the form $\text{Ł}^n_m[ak_1ak_2,\ldots,ak_i,\ldots,ak_n]$. The description given by the concave vertices string can be transformed into the description given by the iterative model $Q^n[L^m]\left(L^k_1,\ldots,L^k_n\right)$. Examples of the transformations of the description given by the concave vertices string into the description given by



a b c d

Fig. 2.11. Archetypes of the class defined by the description given by the concave vertices string

the iterative model for the archetype shown in Fig. 2.11c is as follows: $Ł_4^4[a4a3a4a3] \equiv Q^4[L^4](L^4, L^3, L^4, L^3)$. Examples of archetypes defined by the description given by the concave vertices string are shown in Fig. 2.11. Those archetypes are given by the following symbolic names: $Ł_4^1[a4aaa]$ (Fig. 2.11a), $Ł_4^2[a4a3aa]$ (Fig. 2.11b), $Ł_4^3[a4a3a3a]$ (Fig. 2.11c), and $Ł_4^4[a4a3a4a3]$ (Fig. 2.11d).

The archetype of the complex polygon class C is obtained as the result of a certain type of topological operation called a complex polygon addition. The addition operation defines the way in which polygons are joined together. One of the addition operations that make the complex polygon object by joining two polygons along the common edge is the edge-sum. The edge-sum is defined as follows. Let $\omega^{L^n} \in \Omega^{L^n}$ and $\omega^{L^k} \in \Omega^{L^k}$, where $\omega^{L^k}, \omega^{L^n}$ are archetypes of the polygon class. The sum $\omega^{L^n} \otimes \omega^{L^k}(v_i)$ at the edge $E_i(\omega^{L^n})$ is defined to be a polygon resulted from adding ω^{L^n} with ω^{L^k} by translating, rotating, and scaling ω^{L^k} so that $E_j(\omega^{L^k})$ coincides with $E_i(\omega^{L^n})$. The edge $E_i(\omega^{L^n})$ given by vertices $v_i(\omega^{L^n})$ and $v_{i+1}(\omega^{L^n})$ describes the bounding rectangle of the sum $\omega^{L^n} \otimes \omega^{L^k}(v_i)$. The bounding rectangle is given by a line passing through vertices $v_i(\omega^{L^n})$, $v_{i+1}(\omega^{L^n})$ and perpendicular to the line given by the edge $E_i(\omega^{L^n})$. The archetype of the complex polygon class can consist with more than two parts. The complex polygon class is denoted as $C(nL^i)$, where *n* is a number of polygonal parts L^i. There is a conversion from the notation of the complex class into the notation given by the iterative model. Figure 2.12 shows an archetype of the complex polygon class represented by four different symbolic representations in the form of the

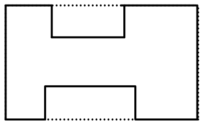


Fig. 2.12. An archetype of the concave polygon class

iterative model $Q^2[L^4](2L^4)$, the complex model $C(L^4, L^4, L^4)$, the subtracting model $\not{L}_4^2[v4vv4v]$, and the cyclic model $\perp^{12}\{(12\cdot\pi)(12d)$ described in [17].

2.1.1.2.3. Concave Curve-Polygon Class

The concave curve-polygon class is a class archetypes of which are the concave curve-polygons. The concave curve-polygon is the class archetypes of which need to have at least one curvilinear segment. At the first level of iteration the following concave curve-polygon classes are possible: $Q^k[M](kM)$, $Q^k[L](kM)$, $Q^k[M](kL)$, $Q^K[L](k_1Mk_2L)$, or $Q^K[M](k_1Mk_2L)$. The description of the specific concave curve-polygon classes can be given using the concave vertices form $W^{i,j,k}[L^m](i\cdot\theta, j\cdot g, k\cdot l^n)$, where θ is a type of the concave curvilinear segment, g is the concave straight-curvilinear segment, and l^n is the concave n-gon. There is a conversion from the notation of the concave vertices form into the notation given by the iterative model. For example, archetypes shown in Fig. 2.13a–c are given by description both in a concave vertices form and by an iterative model:

$Q^2[M[L^3]](M^1, L^3) \equiv W^2[M[L^3]](cl^3w)$ (Fig. 2.13a)

$Q^2[M[L^4]](M^1, L^4) \equiv W^2[M[L^4]](cwcl^4)$ (Fig. 2.13b)

$Q^3[M[L^4]](M^1[L^3], M^1, L^3) \equiv W^3[M[L^4]](cwl^3g^1)$ (Fig. 2.13c)

The regular concave curve-polygon class is the class given by symbolic name $Q^k[M](kM)$. For this class the generic class and all residuals are members of the curve-polygon class M. Examples of archetypes generated from class $Q^k[M](kM)$ are shown in Fig. 2.13d–f. Those archetypes are given by the following symbolic names: $Q^2[M[L^3]](2M^1)$ (Fig. 2.13d), $Q^2[M^2[L^4]](2M^1)$ (Fig. 2.13e), and $Q^2[M^1[L^4]](2M^1)$ (Fig. 2.13f).

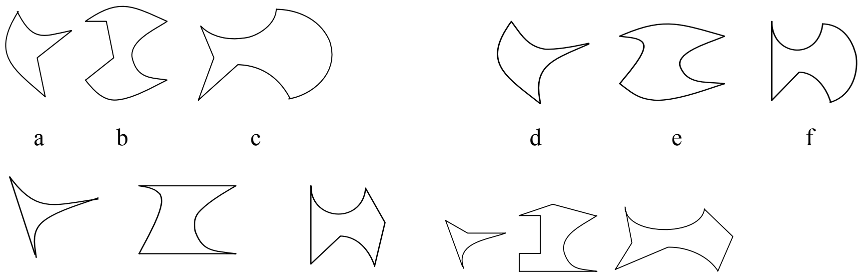


Fig. 2.13. Examples of archetypes whose descriptions are given both in a concave vertices form and by an iterative model

The archetypes of the concave curve-polygon class whose generic class is a member of the polygon class are given by the symbolic name $Q^k[L](kM)$ or $Q^K[L](k_1Mk_2L)$, where $K = k_1 + k_2$. Examples of archetypes generated from the class $Q^k[L](kM)$ are shown in Fig. 2.13g–i. Archetypes shown in Fig. 2.13g–i are given by the following symbolic names: $Q^2[L^3](2M_1)$ (Fig. 2.13g), $Q^2[L^4](2M_1)$ (Fig. 2.13h), and $Q^2[L^5](2M_1)$ (Fig. 2.13i). Examples of archetypes generated from the class $Q^K[L](k_1Mk_2L)$ are shown in Fig. 2.13j–l. Archetypes shown in Fig. 2.13j–l are given by the following symbolic names: $Q^2[L^3](M_1L^3)$ (Fig. 2.13j), $Q^2[L^5](M_1L^4)$ (Fig. 2.13k), and $Q^3[L^5](M_1M_2^1L^3)$ (Fig. 2.13l).

The concave curve-polygon class, for which archetypes of the generic class are members of a convex polygon class L, is called the concave curve-polygon star class and is given by the symbolic name $Q^k[L^k](mMnL)$, $m + n = k$. The concave curve-polygon star class whose all residuals are members of the curve-polygon class is called the regular concave curve-polygon star class and is denoted as $Q^k[L^k](kM)$. Examples of archetypes generated from the curve-polygon star class are shown in Fig. 2.14. Archetypes shown in Fig. 2.14 are given by the following

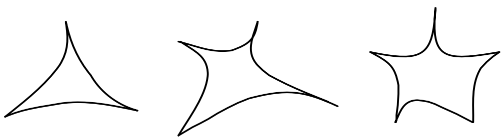


Fig. 2.14. Archetypes of the regular concave curve-polygon star classes

symbolic names: $Q^3[L^3](3M)$ (Fig. 2.14a), $Q^4[L^4](4M)$ (Fig. 2.14b), and $Q^5[L^5](5M)$ (Fig. 2.14c).

2.1.1.3. Thin Classes

As it was described in the section "Concave Curve-Polygon Class," the thin class is a class whose members are thin objects. In this book the term the thin class is used to denote the acyclic-thin class. The thin class is represented by the acyclic graph called a tree. The undirected graph $G = (V, E)$, where V is the set of nodes and $E \subseteq V \times V$ is the set of edges, is called a tree if it satisfies two conditions: the graph is connected and the graph contains no cycles. It can be shown that in the case of the thin acyclic shape class a tree is a spanning tree. An edge of a spanning tree is called a branch and a spanning tree with H vertices consists of H-1 branches. The spanning tree represents an archetype of the thin class. The archetype of the thin class consists of edges and vertices. The two types of vertices are distinguished: the endpoint v^{ξ} and the branching-point v^{ζ}.

The thin class, the archetype of which has a branch $v_i^{\zeta} v_j^{\zeta}$ connecting only the branching points, is called the thin bridge class and the branch $v_i^{\xi} v_j^{\zeta}$ is called a bridge. Depending on the curvilinearity of the branch, two types of branches can be distinguished: the straight branch and the curvilinear branch. The class whose archetypes have all straight branches is called the straight thin class. For the straight thin class a set of angles and distances called the set of attributes of the straight thin class is computed. The set of attributes is denoted as $A^{\Theta} = \left\{ \left(a_1^d, a_1^{\alpha} \right), \left(a_2^d, a_2^{\alpha} \right), \ldots, \left(a_N^d, b_N^{\alpha} \right) \right\}$, where a_i^d is a distance computed as $d_k^{\xi} = \left| \overline{v_i^{\xi} v_j^{\zeta}} \right|$ for two different types of the vertices and $d_k^{\zeta} = \left| \overline{v_i^{\zeta} v_j^{\zeta}} \right|$ for this same type of vertices, and a_i^{α} is an angle computed as $\alpha_k^{\xi} = \angle v_i^X v_k^{\zeta} v_j^X$, where X denotes vertices type ξ or ζ, and $k = 1, \ldots, \mathrm{H}\text{-}1$, $m = 1, \ldots, M$, and $\alpha_k^{\zeta} = \angle v_i^X v_k^{\zeta} v_j^X$, where X denotes vertices type ζ, and $k = 1, \ldots, \mathrm{H}\text{-}1$, $m = 1, \ldots, M$.

Depending on the type of branches the thin class is split into three classes: the 1-D class archetypes of which have only isolated branches $v_1^{\xi} v_2^{\xi}$, the star class $\oplus$ archetypes of which have only external branches $v_i^{\xi} v_j^{\zeta}$, and the thin bridge class Θ_k^1 archetypes of which have both external $v_i^{\xi} v_j^{\zeta}$ and internal $v_i^{\zeta} v_j^{\zeta}$ branches. Examples of the archetypes from the thin class are shown in Fig. 2.15. Archetypes from the Θ^2 class are shown

a b c d e f g h

Fig. 2.15. Archetypes of the thin class (**a**–**c**) the Θ^2 class, (**d**, **e**) the star class $\oplus$, (**f**–**h**) the bridge class Θ_k^1

in Fig. 2.15a–c, archetypes from the star class $\oplus$ are shown in Fig. 2.15d, e, and archetypes from the bridge class Θ_k^1 are shown in Fig. 2.15f–h.

Based on the relations between attributes the following thin star classes $\oplus_k$ are derived:

- The equilateral-star class: this is a class for which all archetypes have all branches equal
- The equiangular-star class: this is a class for which all archetypes have all angles equal
- The ideal star class: this is a class for which all archetypes have all angles and branches equal

The derivation rules applied for each individual class are as follows:

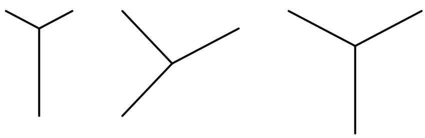
$$\left[\forall a_i^d \in A^T, \exists T_d \in \Re,: a_i^d = T_d\right] \Rightarrow \oplus_k \succ \ddot{\oplus}_k \qquad \text{(the equilateral-star class)}$$

$$\left[\forall a_i^\alpha \in A^T, \exists T_\alpha \in \Re,: a_i^\alpha = T_\alpha\right] \Rightarrow \oplus_k \succ \hat{\oplus}_k \qquad \text{(the equiangular-star class)}$$

$$\left[(\forall a_i^d \in A^T, \exists T_d \in \Re,: a_i^d = T_d) \wedge (\forall a_i^\alpha \in A^T, \exists T_\alpha \in \Re,: a_i^\alpha = T_\alpha)\right] \Rightarrow \oplus_k \succ \overline{\oplus}_k$$

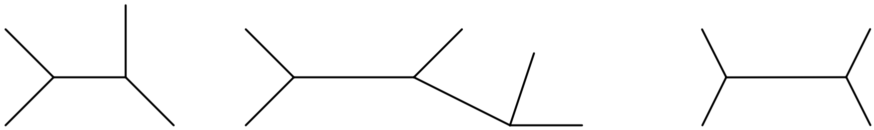
(the ideal star class)

Examples of archetypes of the thin straight star class $\Omega^{\oplus_k}$ are shown in Fig. 2.16. The archetype from the $\ddot{\oplus}^3$ class is shown in Fig. 2.16a, the archetype from the $\hat{\oplus}^3$ class is shown in Fig. 2.16b, and the archetype from the $\overline{\oplus}^3$ class is shown in Fig. 2.16c.



a b c

Fig. 2.16. Archetypes of the straight star class $\Omega^{\oplus_k}$: (**a**) the equiangular-star class $\ddot{\oplus}^3$, (**b**) the equilateral-star class $\hat{\oplus}^3$, (**c**) the ideal star class $\overline{\oplus}^3$



a b c

Fig. 2.17. Archetypes of the bridge thin straight class $\overline{\Theta}_k^1$: (**a**) the equilateral-thin class $\hat{\overline{\Theta}}_4^2$, (**b**) the equilateral-branch thin class $\ddot{\overline{\Theta}}_4^2$, (**c**) the equiangular-branch thin class $\breve{\overline{\Theta}}_4^2$

Similarly, the bridge thin straight class $\overline{\Theta}_k^1$ can be split into specific classes based on a set of attributes A^{Θ} . Examples of archetypes of the bridge thin class are shown in Fig. 2.17. The archetype from the bridge thin straight equilateral-class $\hat{\overline{\Theta}}_4^2$ is shown in Fig. 2.17a, the archetype from the bridge thin straight equilateral-branch class $\ddot{\overline{\Theta}}_4^2$ is shown in Fig. 2.17b, and the archetype from the bridge thin straight equiangular-branch class $\breve{\overline{\Theta}}_4^2$ is shown in Fig. 2.17c.

The special subclass of the thin class is a thin fractal class denoted as Ω^F . The fractal class is described in the form of the thin class as Θ_n^m , where *m* and *n* are numbers that characterize the L-system [18]. The thin fractal class defined by the L-system is restricted to the class for which its graph representation is a spanning tree. It imposes the constraints for the level of iteration of the system and a set of parameters of the model. Archetypes of the fractal class are generated by L-systems. L-system uses strings that are interpreted based on the notion of a LOGO-style turtle. For example, the dragon curve can be generated by repetitively substituting line segments by pairs of lines forming either a left or a right turn and is described by the following L-system:

: Fl
p1: Fl→ Fl + Fr+
p2: Fr→Fl – Fr

The symbols Fl, Fr are interpreted by turtle as the "move left" and "move right" commands, and p1, p2 are productions rules [18]. From the thin fractal class the following specific classes are derived: the equiangular-branch thin fractal class $\hat{F}_\pi^8$, the equiangular-thin fractal class F_α^5, the thin fractal class F_m^k , the thin curved fractal class F ^{5}, and the thin curved fractal

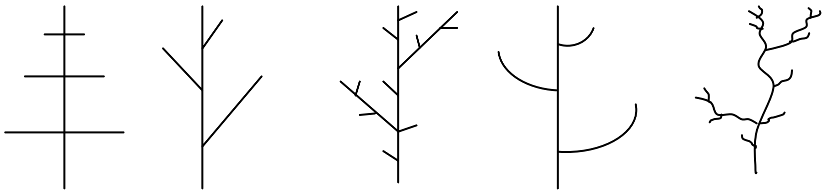


Fig. 2.18. Archetypes of the thin fractal class Ω^F : (**a**) the equiangular-branch thin fractal class $\hat{F}_{\pi}^{8}$, (**b**) the equiangular-thin fractal class F_{α}^{5}, (**c**) the thin fractal class F_m^k, (**d**) the thin curved fractal class F^5, (**e**) the thin curved fractal class F_m^k

class F_m^k. Figure 2.18 shows examples of archetypes generated from the specific fractal classes. These classes are defined in the similar way as the specific classes described in previous sections. The archetype of the class $\hat{F}_{\pi}^{8}$ is shown in Fig. 2.18a, the archetype of the class F_{α}^{5} is shown in Fig. 2.18b, the archetype of the class F_m^k is shown in Fig. 2.18c, the archetype of the class F^5 is shown in Fig. 2.18d, and the archetype of the class F_m^k is shown in Fig. 2.18e.

As it was described in the section "Concave Curve-Polygon Class" the 1-D thin class Θ^2 is the class archetypes of which have only isolated branches $v_1^{\xi} v_2^{\xi}$. From the 1-D thin class the specific classes are derived based on the properties of the graph function that is representative of the archetype of the class Θ^2. The function $y = f(x)$ is defined in the closed interval [*a*,*b*] and is prescribed by an analytical expression or a formula. It is assumed that the function fulfils the conditions: $f(-a) = f(a) = c$ and $\forall x \in [a,b],\, f(x) \le c \wedge f(x) \ge d$, where *c*, *d* are the greatest and the smallest of all values of the function *f*(*x*). The 1-D thin class Θ_F^2 is defined as follows: $[\forall x \in [a,b], \exists y \in [c,d] : y = f(x)] \Rightarrow \Theta^2 \succ \Theta_F^2$. The 1-D thin convex function class derived from Θ^2 is defined as follows: $[\forall x_1, x_2 \in [a,b],$ $\exists \lambda \in (0,1) : f(\lambda x_1 + (1-\lambda)x_2) \le \lambda f(x_1) + (1-\lambda) f(x_2)] \Rightarrow \Theta^2 \succ \Theta_C^2$. The 1-D thin class derived from Θ^2 for which their graph is symmetric with respect to the vertical axis *f*(–*x*) = *f*(*x*) is called 1-D symmetric class Θ_S^2. The derivation rules are as follows: $[\forall x \in [a,b] : f(-x) = f(x)] \Rightarrow \Theta^2 \succ \Theta_S^2$.

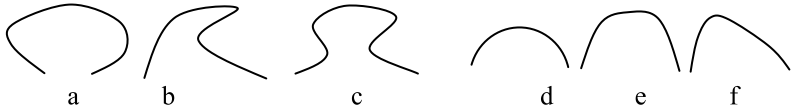


Fig. 2.19. Archetypes of the nonfunction classes $\Omega^{\Im}$ (**a**–**c**) and archetypes of the convex function classes: (**d**) symmetrical $\Omega^{\hat{\hat{\Xi}}}$, (**e** and **f**) nonsymmetrical $\Omega^{\hat{\Xi}}$

Examples of archetypes generated from specific 1-D thin class are shown in Fig. 2.19. Archetypes of the nonfunction classes $\Omega^{\Im}$ are shown in Fig. 2.19a–c, the archetype from the convex symmetrical function class $\Omega^{\hat{\hat{\Xi}}}$ is shown in Fig. 2.19d, and archetypes from the convex nonsymmetrical (NS) function class $\Omega^{\hat{\Xi}}$ are shown in Fig. 2.19e, f.

The 1-D thin class Θ^2, archetypes of which are straight poly-lines, is described in relation to its generic class and is called the thin poly-line class $\otimes$. The archetypes of the generic class are obtained by joining the pseudonodes of the archetypes of the class $\otimes$ as shown in Fig. 2.20. The archetype of the class $\otimes[L^4]$ shown in Fig. 2.20a is described in relation to its generic class L^4 (Fig. 2.20b) and the archetype of the class $\otimes[Q^1[L^4](L^3)]$ (Fig. 2.20c) is described in relation to its generic class $Q^1[L^4](L^3)$ (Fig. 2.20d).

The bridge tree class is the class derived from the bridge thin class. Archetypes of the bridge tree class are represented by the acyclic graph called a tree. The bridge tree class is described by the bridge notation that is explained in Fig. 2.22. The bridge is denoted by the bracket "[ ]," whereas branch by the bracket "( )." The notation is based on the decomposition of the tree into branches and bridges. During decomposition the

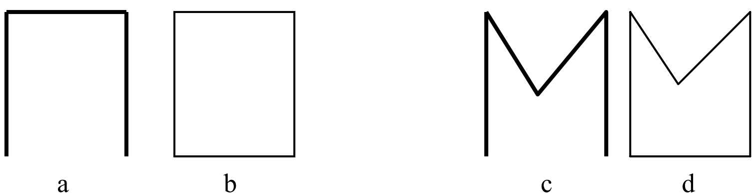
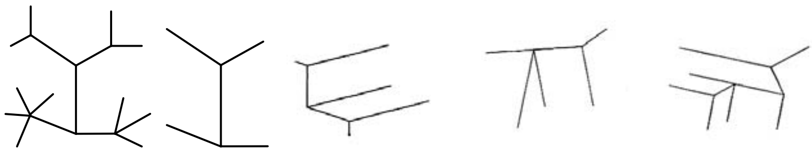


Fig. 2.20. Archetypes of the thin poly-line class $\otimes$ and its generic class: (**a**) the class $\otimes[L^4]$ and (**b**) its generic class L^4, (**c**) the class $\otimes[Q^1[L^4](L^3)]$ and (**d**) its generic class $Q^1[L^4](L^3)$

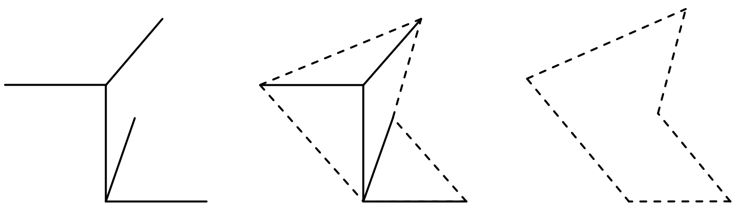


a b c d e

Fig. 2.21. Archetypes of the bridge tree class

branches are removed and the bridge that is left becomes the generic bridge of the tree. For example, the archetype shown in Fig. 2.21a is an archetype from the bridge tree class $\Theta\langle[1]\{[1](2)\}\{[1](2)\}\{[1](3)\}\{[1](4)\}\rangle$. The result of removing branches is the string $\Theta\langle[1]\{[1]\}\{[1]\}\{[1]\}\{[1]\}\rangle$ and finally after renaming bridges into branches the bridge class $\Theta\langle[1](2)(2)\rangle$ is obtained. The bridge class $\Theta\langle[1](2)(2)\rangle$ that is the result of decomposition is shown in Fig. 2.21b. Examples of the archetypes from the bridge tree classes are shown in Fig. 2.21. The archetype from the class $\Theta\langle[1](2)(1)[1](2)\rangle$ is shown in Fig. 2.21c, the archetype from the class $\Theta\langle[1](3)(2)\rangle$ is shown in Fig. 2.21d, and archetype from the class $\Theta\langle[1]\{(2),[1](2)\}\{(1),[1](2)\}\rangle$ is shown in Fig. 2.21e.

As it was described in previous sections each class can be described by applying the different notations. The archetypes in Fig. 2.22 are described by the notation of the bridge tree class $\Theta\langle[1](2)(2)\rangle$, generic bridge tree class $\Theta\langle[1](2)(2)\rangle[Q^1[L^4](L^3)]$, or by notation of the Θ/ρ class as $\Theta/\rho[Q^1[L^4](L^3)]\{3L^3,L^4\}$. The notation of the Θ/ρ class is derived from the notation of the ρ class described in the further part of this chapter. In order to explain the notation of the Θ/ρ class, an example of decomposition of the archetype from the Θ/ρ class is shown in Fig. 2.22. Figure 2.22a shows the archetype from the thin bridge class $\Theta\langle[1](2)(2)\rangle$. The



a b c

Fig. 2.22. Explanation of the notation of the Θ/ρ class

archetype consists of one bridge that has two branches on its ends. The endpoints of this archetype are joined by straight lines as shown in Fig. 2.22b, and as the result the object consisting of the four parts, three triangles L^3, and one quadrilateral L^4, was obtained. The generic polygon, archetype of the class $Q^1[L^4](L^3)$, is shown in Fig. 2.22c.

2.1.1.4. Cyclic Class

As it was described in the section “Thin Classes,” the cyclic general class A is defined based on the values of the attribute called a homotopy measure. The cyclic class A consists of elements that can be decomposed iteratively into subregions (holes). The decomposition scheme in which the cyclic object is broken down into very simple primitives, called holes, is similar to the decomposition scheme of the concave object described in previous sections. At first all holes are filled and an object “without holes” is used as a base for the decomposition of the object into the filled regions and holes. Next each hole is examined in the process called the first level of iteration. In the case when some holes are cyclic they are examined in the process called the second level of iteration. The description of the concave class depends on the level of iteration and is given by a symbolic name $A^n[\Im_\Phi](n\Im_A)$, where *n* is the number of residuals, $\Im_A$ is a type of the holes, and $\Im_\Phi$ is a type of the generic classes. The base cyclic class is denoted as $A^n[\Im_\Phi]$, where $\Im_\Phi$ is one of the acyclic general classes $\Im_\Phi = \{\Lambda, Q\}$ from which the base cyclic class is derived and *n* is a number of holes. The description of the specific cyclic classes is based on a type of the generic class $\Im_\Phi$ as well as on the type of the holes $\Im_A$. The archetype of a cyclic class derived from the acyclic class can be seen as a result of subtraction of the acyclic region and holes.

At the first level of iteration the symbolic representation of the cyclic class is given as $A^n[\Gamma](n\Im_A)$, where a hole can be a member of the thin or acyclic class $\Im_A \equiv \{\Theta, \Gamma\}$. Depending on the number of holes *n*, and a type of generic class $\Im_\Gamma$, and a type of holes $\Im_A$, the following specific cyclic classes can be derived: $A^n[\Lambda](n\Gamma)$, $A^n[\Lambda](n\Theta)$, $A^n[Q](n\Gamma)$, and $A^n[Q](n\Theta)$. In the case when there are *n* holes there are the following classes given by the symbolic names: $A^n_\Lambda(n\Lambda)$, $A^n_\Lambda(nQ)$, $A^n_\Lambda(n\Theta)$,

a b c d e f g h

Fig. 2.23. Examples of exemplars from the cyclic class

$A_Q^n(n\Lambda)$, $A_Q^n(nQ)$, and $A_Q^n(n\Theta)$. Figure 2.23 shows exemplars generated from the cyclic class. The symbol $A^3\,[L^4]\,(3L)$ (Fig. 2.23b) denotes exemplar whose generic class is the convex polygon class L^4 (rectangle) and it has three holes. All holes are archetypes of the rectangle class L^4. Figure 2.23e shows exemplar generated from the cyclic class $A^2[Q^4[L^8](4L^4)]\left(K_E^1, L^4\right)$, whose generic class is the concave polygon class $Q^4\,[L^8\,](4L^4)$, and it has two holes. The first hole is an archetype of the rectangle class L^4 and the second one is the archetype of the curvilinear class (ellipse) K_E^1.

The archetype of the class $A^1[\Lambda](\Gamma)$ for which the hole has common points with the border points is a concave point class $A^1[\Lambda](\Gamma) \equiv \overline{Q}^1[\Lambda](\Gamma)$. Similarly, the archetype of the class $A^1[Q](\Gamma)$ for which the hole has common points with the border points is the archetype of the concave point class $A^1[Q](\Gamma) \equiv \overline{Q}^1[Q](\Gamma)$. Figure 2.24 shows exemplars generated from the concave point class $\overline{Q}^1[L^4](L^4)$ (Fig. 2.24a) and $\overline{Q}^1[L^4](L^3, \overline{L}^3)$ (Fig. 2.24b).

Similarly, at the second level of iteration the symbolic representation of the cyclic class is given as $A^n[\Gamma](n\Im_A)$, where at least one hole from the set $\Im_A$ is a member of the cyclic class A. Examples of exemplars generated from the cyclic classes at the second level of iteration are shown in Fig. 2.25a–d. The symbolic names of these exemplars are as follows:

a b

Fig. 2.24. Exemplars of the concave point class

$A^1\left[L_R^4\right]\left(A^1\left[L_R^4\right]\left(L_R^4\right)\right)$ (Fig. 2.25a)

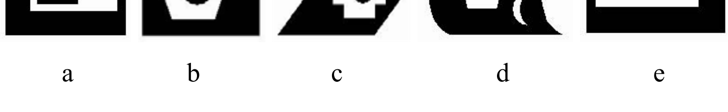
$A^1\left[L_R^4\right]\left(A^1\left[L_T^4\right](K)\right)$ (Fig. 2.25b)

$A[Q^1[L^5](L^3)](A[Q^4[L^8](4L^3)](K))$ (Fig. 2.25c)

$A^2[Q^1[M^2[L^4]](M^1)](A[L_T^4](Q^1[M^2[L^5]](Q^2[L^3](2M^1))), Q^1[M^1](M^1))$ (Fig. 2.25d). Example of exemplar generated from the cyclic classes at the third level of iteration, whose symbolic name is given as follows $A^1\left[L_R^4\right]\left(A^1\left[L_R^4\right]\left(A^1\left[L_R^4\right]\left(L_R^4\right)\right)\right)$, is shown in Fig. 2.25e. The symbolic name $A^1\left[L_R^4\right]\left(A^1\left[L_R^4\right]\left(L_R^4\right)\right)$ denotes an exemplar generated from the class whose generic class is the convex polygon class L^4 (rectangle) and it has one hole. The hole is an archetype of the cyclic class $A^1\left[L_R^4\right]\left(L_R^4\right)$. The generic class of the hole is the convex polygon class L^4 (rectangle). The hole is an archetype of the rectangle class L^4.

2.1.1.5. Complex Cyclic Class

Archetypes of the complex cyclic class $C(\Theta)$ are obtained as the result of the certain type of topological operation called a complex addition [19], [20]. The complex class is denoted as $C(\Gamma_1,\Gamma_2,\ldots,\Gamma_N)$, where $\Gamma_1,\Gamma_2,\ldots,\Gamma_N$ are classes of the addition operation. In the case when $N = 2$, the complex class is reduced into the class of the two-element operation and denoted as $C(\Gamma_1,\Gamma_2)$. In the case when $\Gamma_1 \equiv \Lambda$, the class is called the complex convex class. In the case when $\Gamma_1 \equiv A$, the class is called the complex cyclic class. Archetype of the complex cyclic class consists of parts, where one of the parts needs to be an archetype of the cyclic class. Examples of the archetypes of the complex cyclic class are given in Fig. 2.26. Symbol $C\left(A^1\left[L_R^4\right](L^4),L^3\right)$ (see Fig. 2.26a) denotes that archetype of the complex class C consists of two parts, one archetype



a b c d e

Fig. 2.25. Exemplars of the complex cyclic class given by symbolic names

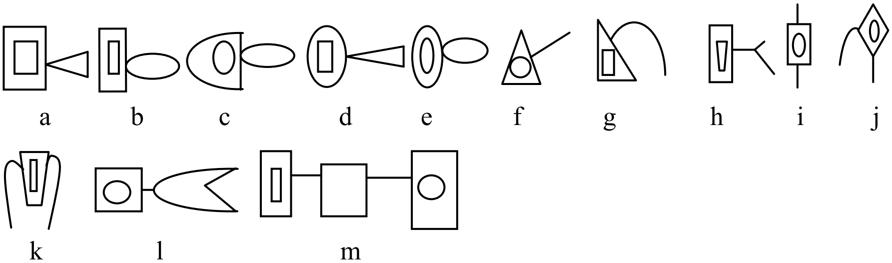


Fig. 2.26. Archetypes of the complex cyclic class

of the cyclic class $A^1\left[L_R^4\right](L^4)$ and the second archetype of the convex class L^3. In the case when $\Gamma_2 \equiv \Theta$, the complex class is defined by the point-sum operation and the class is called the complex convex thin class. Examples of the archetypes of the complex cyclic-thin class are given in Fig. 2.26. Symbol $C(A^1[L^3](K),\Theta)$ (see Fig. 2.26f) denotes that an archetype of the complex class C consists of two parts, one archetype of cyclic class $A^1[L^3](K)$ and the second archetype of the thin class Θ.

Examples of archetypes generated from the complex cyclic-thin class given by symbolic names are shown in Fig. 2.26: $C\left(A^1\left[L_R^4\right](L^4),L^3\right)$ (Fig. 2.26a), $C\left(A^1\left[L_R^4\right]\left(L_R^4\right),K\right)$ (Fig. 2.26b), $C(A^1[M](K),K)$ (Fig. 2.26c), $C\left(A^1[K]\left(L_R^4\right),L^3\right)$ (Fig. 2.26d), $C(A^1[K](K),K)$ (Fig. 2.26e), $C(A^1[L^3](K),\Theta)$ (Fig. 2.26f), $C(A^1[L^3](L^4),\tilde{\Theta})$ (Fig. 2.26g), $C\left(A^1\left[L_R^4\right]\left(L_T^4\right),\Theta^3[L^3]\right)$ (Fig. 2.26h), $C\left(A^1\left[L_R^4\right](K),2\Theta\right)$ (Fig. 2.26i), $C\left(A^1\left[L_O^4\right](K),\Theta\tilde{\Theta}\right)$ (Fig. 2.26j), $C\left(A^1\left[L_T^4\right](L^4),2\tilde{\Theta}\right)$ (Fig. 2.26k), $C\left(A^1\left[L_R^4\right](K),\Theta,Q^1[M^1](L^3)\right)$ (Fig. 2.26l), and $C\left(A^1\left[L_R^4\right]\left(L_R^4\right),\Theta,L_R^4,\Theta,A^1\left[L_R^4\right](K)\right)$ (Fig. 2.26m).

2.1.1.6. Cyclic Thin Class: The G-Class

The archetype of the cyclic class $A_\Lambda^1(\Gamma)$ for which the type of the hole and the generic acyclic class is equal and area of the hole is close to the area of the archetype of the generic acyclic class is called the archetype of the cyclic-thin class $A_\Lambda^1(\Gamma) \equiv \rho[\Lambda]\{\Gamma\}$. Examples of exemplars generated from classes given by the symbolic names: (a) $\rho\left[L_R^4\right]\left\{L_R^4\right\}$,

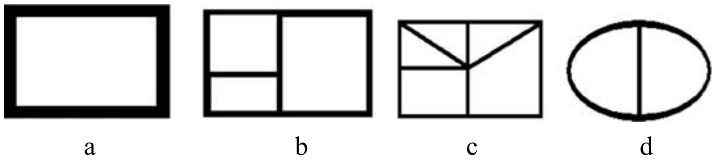


Fig. 2.27. Examples of exemplars of the cyclic-thin class

(b) $\rho\left[L_R^4\right]\left\{3L_R^4\right\}$, (c) $\rho\left[L_R^4\right]\left\{3L_R^3, L_R^4, L_T^4\right\}$, and (d) $\rho\left[K_E^1\right]\{2M^1\}$ are shown in Fig. 2.27a–d. The symbolic name $\rho\left[L_R^4\right]\left\{L_R^4\right\}$ denotes an exemplar generated from the class whose generic class is the convex polygon class L^4 (rectangle) and it has one hole. The hole is an archetype of the rectangle class L^4. The symbol ρ denotes that the exemplar is generated from the acyclic class whose area of the hole is close to the area of the archetype of the generic acyclic class.

The archetype of the cyclic-thin class $\rho[\Lambda]\{\Gamma\}$ can be represented by notation of the Θ-class. In this notation the object is decomposed into the core object and the thin object. Example of this decomposition is shown in Fig. 2.28. Figure 2.28 shows archetypes from the cyclic-thin class $\rho[Q^1[L^6](L^3)]\left\{Q[L^5]\left(L_R^3\right), 3L_T^4\right\}$ that are decomposed according to the convention of the Θ-class. The archetype in Fig. 2.28a given by the symbolic name $G[Q[L^6](L^3)]\{\Theta^{2(2)2}[Q[L^4](L^3)]\}$ is decomposed into the thin object $\Theta^{2(2)2}[Q[L^4](L^3)]$ (Fig. 2.28b) and the concave core object $Q[L^6](L^3)$ (Fig. 2.28c). This archetype is represented as a member of the cyclic-thin class $\rho[Q^1[L^6](L^3)]\left\{Q[L^5]\left(L_R^3\right), 3L_T^4\right\}$ and is decomposed into the concave core object $Q[L^6](L^3)$ and four objects: one concave $Q[L^5](L_R^3)$ and the three convex L_T^4.

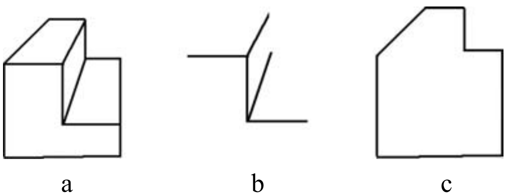
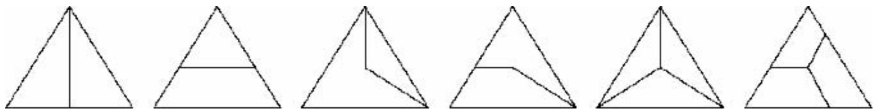


Fig. 2.28. Archetypes of the cyclic-thin class



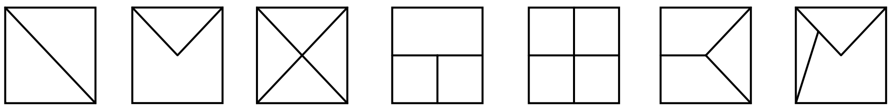
a b c d e f

Fig. 2.29. Archetypes of the class triangle convex thin class

2.1.1.6.1. Convex Cyclic Thin G-Class

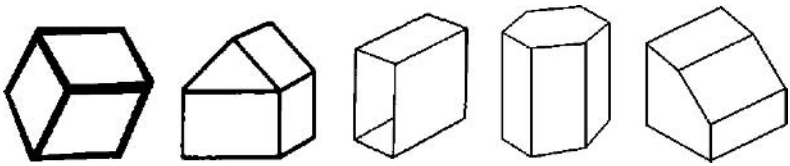
Archetype of the convex cyclic-thin class can be decomposed into the core convex object Λ and holes. Figure 2.29 shows archetypes of the convex cyclic-thin class whose the core convex object is the member of the convex triangle class L^3. Archetypes shown in Fig. 2.29 are represented by the following symbolic names: $\rho[L^3]\left\{2L_R^3\right\}$ (Fig. 2.29a), $\rho[L^3]\{L^3,L^4\}$ (Fig. 2.29b), $\rho[L^3]\{Q[L^3](L^3),L^3\}$ (Fig. 2.29c), $\rho[L^3]\{L^4,Q[L^3](L^3)\}$ (Fig. 2.29d), $\rho[L^3]\{3L^3\}$ (Fig. 2.29e), and $\rho[L^3]\{3L^4\}$ (Fig. 2.29f). Figure 2.30 shows archetypes of the convex cyclic-thin class whose the core convex object is the member of the convex rectangle class L_R^4. Archetypes shown in Fig. 2.30 are represented by the following symbolic names: $\rho\left[L_R^4\right]\left\{2L_R^3\right\}$ (Fig. 2.30a), $\rho\left[L_R^4\right]\left\{L^3,Q\left[L_R^4\right](L^3)\right\}$ (Fig. 2.30b), $\rho\left[L_R^4\right]\left\{4L_R^3\right\}$ (Fig. 2.30c), $\rho\left[L_R^4\right]\left\{2L_Q^4,L_R^4\right\}$ (Fig. 2.30d), $\rho\left[L_R^4\right]\left\{4\cdot L_R^4\right\}$ (Fig. 2.30e), $\rho\left[L_R^4\right]\left\{2L_T^4,L^3\right\}$ (Fig. 2.30f), and $\rho\left[L_R^4\right]\left\{2L^3,Q^1[L^4](L^3)\right\}$ (Fig. 2.30g).

Figure 2.31 shows the archetype of the convex cyclic class. The core convex object of this archetype is the member of the convex class L^n. The symbolic names for these objects are given in the form of the convex thin class $\rho\{L^n\}\{kL^m\}$ and the G-class $G\{L^n\}\{\Theta\}$. The symbolic names are as follows: $\rho[L^6]\left\{3L_O^4\right\}$, $G[L^6]\{\Theta^3[L^3]\}$ (Fig. 2.31a),



a b c d e f g

Fig. 2.30. Archetypes of the convex rectangle cyclic-thin class



a b c d e

Fig. 2.31. Archetypes of the class $G\{L^n\}\{\Theta\}$

$\rho[L^{6-7}]\left\{2L^4_O, L^3, L^4_O\right\}$, $G[L^{6-7}]\{\Theta^4[L^4]\}$ (Fig. 2.31b), $\rho[L^4]\left\{2L^4_O, L^3, L^4\right\}$, $G[L^4]\{\Theta^{2(2)2}[L^4]\}$ (Fig. 2.31c), $\rho[L^8]\left\{L^6, 3L^4_O\right\}$, $G[L^8]\{\Theta^{2(2)2}[L^4]\}$ (Fig. 2.31d), and $\rho[L^7]\left\{L^5, 3L^4_O\right\}$, $G[L^7]\{\Theta^{2(2)2}[L^4]\}$ (Fig. 2.31e). The symbol L^{6-7} denotes that the archetype is the member of the class L^6 or L^7.

2.1.1.6.2. Concave Cyclic Thin G-Class

Archetype of the concave cyclic-thin class can be decomposed into the core concave object Q and holes. Following the notation of the G-class the archetype is decomposed into the concave core object Q and the thin objects Θ or the complex thin objects $C(\Theta)$. Figure 2.32 shows the archetypes of the concave cyclic-thin class that are decomposed into the core concave object and the thin object Θ. Archetypes are represented by the symbolic names as follows: $\rho[Q[L^6](L^3)]\left\{3L^4_T, Q_{L^5}\left(L^3_R\right)\right\}$, $G[Q[L^6](L^3)]\left\{\Theta^{2(2)2}\left[Q_{L^4}(L^3)\right]\right\}$ (Fig. 2.32a), $\rho\,[Q[L^7](L^3)]\left\{4L^4_o, L^3\right\}$, $G\left[Q_{L^7}(L^3)\right]\left\{\Theta^{3(2)2}\left[Q_{L^4}(L^3)\right]\right\}$ (Fig. 2.32b), $\rho\left[Q_{L^{6-8}}(L^3)\right]\left\{2L^4_o, \hat{L}^4_o, Q_{L^5}(L^3)\right\}$, $G\left[Q_{L^{6-8}}(L^3)\right]\left\{\Theta^{2(2)1(2)2}\left[Q_{L^5}(L^3)\right]\right\}$ (Fig. 2.32c), and $\rho\left[Q^2_{L^{6-8}}(2L^3)\right]\left\{\left((L^6)3L^4_O\right), 3L^4_O, Q_{L^4}(L^3)\right\}$, $G\left[Q^2_{L^{6-8}}(2L^3)\right]\left\{\Theta^{2(2)2(2)2(2)2}\left[Q^2_{L^5}(2L^3)\right]\right\}$ (Fig. 2.32d). Figure 2.33 shows archetypes of the concave cyclic-thin class that are decomposed into the core concave object and the complex thin objects $C(\Theta)$. Archetypes are represented by the symbolic names as follows: $G\left[Q_{L^6}(L^3)\right]\left\{C\left(L^3, 3\Theta^2, \Theta^3_{L^3}\right)\right\}$ (Fig. 2.33a), $G\left[Q_{L^7}(L^3)\right]\left\{C\left(L^3, 3\Theta^2, \Theta^3_{L^3}\right)\right\}$ (Fig. 2.33b), and $G\left[Q_{L^{6-7}}(L^3)\right]\{C(L^3, 4\Theta^2)\}$ (Fig. 2.33c).

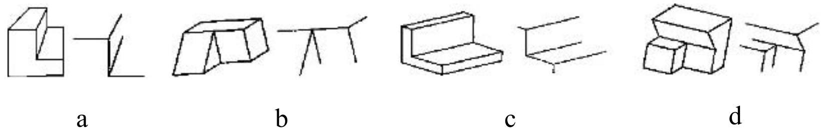


Fig. 2.32. Archetypes of the thin concave G-class $G\{Q\}\{\Theta\}$

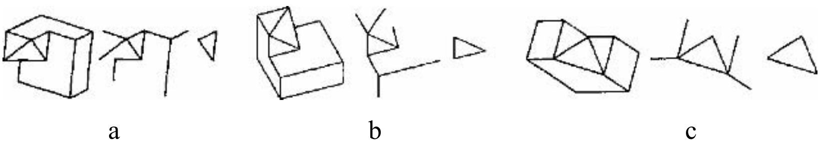


Fig. 2.33. Archetypes of the thin concave G-class $G\{Q\}\{C(\Theta)\}$

2.1.1.7. Colored Classes

The class for which an archetype can be seen as consisting of adjacent regions of the different uniform colors is called the colored class $\aleph$. An archetype of the colored class $\aleph$ can be decomposed into the regions of the different colors and assigned to one of the specific classes. The decomposition of the archetype is shown in Fig. 2.34.

The colored class is the class archetypes of which have their parts marked by the different colors. The description of the convex colored class can be reduced into the description of the cyclic class $A^n[\Im_\Phi](n\Im_A)$. The archetype of the colored class $\aleph^2\left[L_T^4(g)\right]\left(L_R^4(y), L_T^4(b)\right)$ is shown in Fig. 2.35. The symbol $\aleph^2\left[L_T^4(g)\right]\left(L_R^4(y), L_T^4(b)\right)$ denotes that the convex

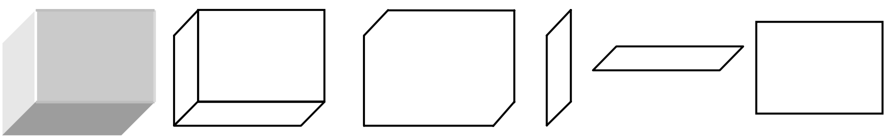


Fig. 2.34. Decomposition of the archetype consisting of adjacent regions of the different uniform colors

Fig. 2.35. Archetype of the convex colored class $\aleph^2\lceil L_T^4(g)\rceil\left(L_R^4(y), L_T^4(b)\right)$

a b

Fig. 2.36. Archetypes of the concave complex colored class

colored class $\aleph$, whose generic class (the convex polygon class – quadrilateral $L_T^4(g)$ – color green) called background, has two regions of the different colors. Both regions are archetypes of the quadrilateral class L^4. The first region $L_R^4(y)$ is marked by the letter (y) denoting the yellow color whereas the second region $L_T^4(b)$ is marked by the letter (b) denoting the blue color.

The archetype of the concave complex colored class can be decomposed into the parts of the different colors. The concave complex colored class is denoted as $\wp(\Gamma_1, \Gamma_2, \ldots, \Gamma_N)$, where $\Gamma_1, \Gamma_2, \ldots, \Gamma_N$ are general classes of shape. Archetypes of the concave complex colored class are given in Fig. 2.36. The symbol $\wp\left(L_R^4(y), Q\left[L_R^4\right]\left(L_R^4\right)(o)\right)$ denotes that the archetype of the concave complex colored class $\wp$, can be decomposed into two regions – the convex polygon class (rectangle) L_R^4 and the concave polygon (rectangle with the one concavity) $Q\left[L_R^4\right]\left(L_R^4\right)$. The archetype shown in Fig. 2.36b is represented by the symbolic name $\wp\left(\aleph[L^4(g)]\left(L_T^4(b)\right), \aleph\left[L_R^4(y)\right](K^1(r))\right)$.

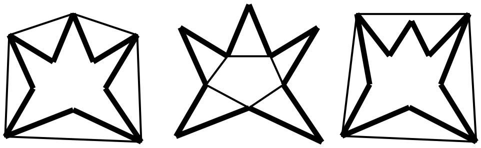
2.1.2. The a Posteriori Classes

The shape classes described in the previous chapters were established based on the geometrical properties of the figure. The derivation of the specific classes was based on constraining the values of selected attributes of the general classes. These classes are called a priori classes because derivation of the specific class is based on geometrical properties of archetypes generated from the selected class. In this book shape is interpreted as the basic perceptual category to which the perceived object is fitted. Shape as the perceptual category is used during the learning of the visual concept of the different ontological categories such as a letter, a sign, or a real-world object. During categorical learning the specific shape classes that are good representative of the shape of the given ontological category need to be derived from the existing a priori classes. The classes of shape that are derived as the result of "specialization" of the existing a priori classes are called the a posteriori classes.

2.1.2.1. The Star Class

As it was described in Sect. 2.1.2 the a priori classes are established based on the geometrical properties of the visual object. The a posteriori classes are derived from the a priori classes based on the specialization of the selected shape classes. Specialization means that the a posteriori classes are established to match shape of the sign or the real-world object. Example of the class that is established based on the existing meaningful objects called a sign is the star class. The star class is defined based on generalization of the most often used visual representations of the star signs. The star class is a class derived from the concave class $Q^n[\Lambda](n\Lambda)$, where $n>2$. The polygon star class is a class derived from the concave polygon class and is given by the symbolic name $Q^n[L^n](nL^3)$. The curvilinear star class is a class derived from the concave class where all residuals are archetypes of the curve-polygon class $Q^n[L^n](nM)$. The concave star class is a class derived from the concave class where all residuals are archetypes of the concave class $Q^n[L^n](nQ)$. The concave I-star class is a class derived from the concave class where all residuals are archetypes of the concave class, residuals of which are archetypes of the concave class $Q^n[L^n](nQ(mQ))$. In similar way the concave II-star class $Q^n[L^n](nQ(mQ(kQ)))$ or the concave III-star class $Q^n[L^n](nQ(mQ(kQ(hQ))))$ can be defined.

The concave polygon star class is a class derived from the concave polygon class where all residuals are archetypes of the concave polygon class $Q^n[L^n](nQ^h[L^k](L^l))$, where indexes h, k, and l denote: h-the number of residuals, the generic k-polygon, and the residual l-polygon. The concave polygon I-star class is a class derived from the concave polygon class where all residuals are archetypes of the concave polygon class, residuals of which are archetypes of the concave polygon class $Q^n[L^n](nQ^h[L^k](mQ^b[L^c](L^d)))$.

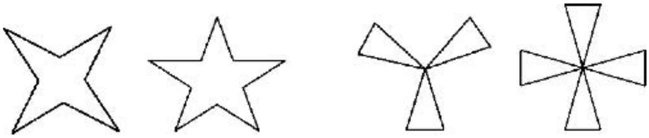


a b c

Fig. 2.37. Explanation of the different notations of the star class

The star class can be described by using the notation of the complex-core class. Objects shown in Fig. 2.37 explain the differences in the description of the object in terms of the concave class $Q^5[L^5](5L^3)$ (Fig. 2.37a) and the complex-core class $\Delta^5[L^5]\{L^5\}(5L^3)$ (Fig. 2.37b). The symbol in the bracket "[ ]" denotes the generic polygon, for example, $[L^5]$ (see Fig. 2.37a), whereas the symbol in the bracket "{ }" denotes the core of the archetype of the complex class, for example, $\{L^5\}$ (see Fig. 2.37b). The advantage of the second approach is such that the object is interpreted as an object having the "arms." Based on this interpretation we can establish the proper similarity relations among objects. For example, the objects from Fig. 2.37b, c that look very similar, have the same complex-core class description given by the symbolic name $\Delta^5[L^5]\{L^5\}(5L^3)$ but the different convex class description given by the symbolic names $Q^5[L^5](5L^3)$ (Fig. 2.37b) or $Q^4[L^4](Q^1[L^4],3L^3)$ (Fig. 2.37c). It seems that the complex-core class description is more perceptually oriented.

The archetype generated from the n-star class is represented by the symbolic name $Q^n[L^n](nL^3)$, where the $2n$-star class is a class derived from the concave polygon class $Q^n[L^{2n}](nL^3)$. The class $R^n[L^{2n}](nL^3)$ derived from the class $Q^n[L^{2n}](nL^3)$, where all residuals have the common point, can be given by notation of the complex class $S^n[L^{2n}](nL^3)$. In the case when there is no common point the class will have description $S^n[L^{2n}]\{L^h\}(nL^k)$. By generalization, the class $S^n[L^{2n}](nL^3)$ can be extended to the c-class $S^n_\Gamma\{n\Im\}$, where $\Im$ is a general class. Examples of the n-star class are given in Fig. 2.38a–b, and the $2n$-star class in Fig. 2.38c–d.

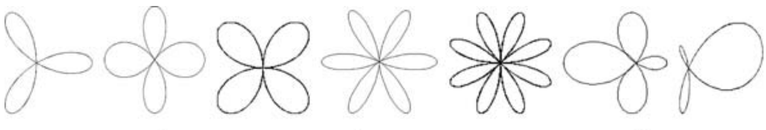


a b c d

Fig. 2.38. Simple I-star classes (**a**, **b**) and simple II-star classes (**c**, **d**)

The curvilinear star class $Q^n[A](nQ^2[L^3](2M^1))$, where A denotes one of the classes $A = \{L^n, \widehat{L}^n, M^n(L^{2n})\}$, is the specific class derived from the concave star class. In the case where all residuals have the common point the class can be given by the complex class description. The complex class description interpret the object in terms of petals (the parts of the object that are “glued” in one point) and described by complex class description as the *c*-complex class $S^n[\Gamma]\{n\Im\}$. When all petals are archetypes of the curve class $\Im \equiv K$ the class is the regular curve c-class $S^n[A]\{nK\}$. In the case where *n* is big enough *n*>*M*, the generic class becomes the polygon class and the regular curve *c*-class is given as $S^n[L^n]\{nK\}$. In the case where petals are different (members of the archetypes of the different curve classes) $\Im \equiv K_i$, the c-class is called the nonregular curve c-class and is given as $S^n[A]\{n*K\}$. Examples of archetypes of the regular curve c-class are shown in Fig. 2.39a–e and archetypes of the nonregular curve c-class are shown in Fig. 2.39f–g.

The c-class $S^n[\Gamma](n\Im)$, for which all petals are archetypes of the thin class $\Im \equiv \Theta$, is the regular thin c-class $S^n[L^n](n\Theta)$. In the case where $\Im \equiv \Theta^2$, the class is reduced to the thin star class $\oplus^n[L^n]$. In the case when the thin star class has different sizes of the “rays” $S^n[L^m](n\Theta), m < n$ the class is the thin para-star class $\oplus^n[L^m]$. Example of the archetype



a b c d e f g

Fig. 2.39. Examples of archetypes of the regular curve c-class and the nonregular curve c-class

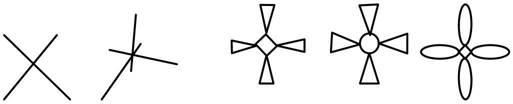


Fig. 2.40. Archetypes of the thin star class, exemplar of the concave c-class

generated from the thin star class $\oplus^4[L^4]$ is shown in Fig. 2.40a and the archetype generated from the thin para-star class $\oplus^6[L^4]$ is shown in Fig. 2.40b.

The complex star point class is the class that has the nucleus and petals that are joined in one point with nucleus. This class is given by the notation of the complex class $S^n[\Gamma]\{\Gamma\}(n\Im)$. The complex polygon star point class is the class whose nucleus and petals are polygons. The complex polygon star point class is given by the symbolic name $S^n[L^h]\{L^k\}(nL^m)$. Examples of the archetypes of the complex star point class are given in Fig. 2.40c $S^4[L^8]\{L^4\}(4L^3)$, Fig. 2.40d $S^4[L^8]\left\{K_C^1\right\}(4L^3)$, and Fig. 2.40e $S^4[\widehat{L}^4]\{L^4\}\left(4K_E^1\right)$.

2.1.2.2. The Spade Class

The spade class is the a posteriori class that is established based on the properties of the real-world object. The spade class denoted as $('S')C(\Gamma,\Theta)$ is derived from the complex symmetrical thin class $C(\Gamma,\Theta)$ archetypes of which consist of two parts, one called the blade and the other one called handle. The handle is a member of the thin class Θ. The members of the a posteriori spade class are used as the structural archetypes of the real-world object called spade. The spade class $('S')C(\Gamma,\Theta^2)$ is the class archetypes of which are obtained by joining the straight line with the object called the core that is a member of one of the classes: the convex, the concave, or the cyclic in such a way that the straight line has one common point with one of the sides of the core and the whole figure is symmetrical. Examples of the spade class are shown in Fig. 2.41a $('S')C\left(L_R^4,\Theta^2\right)$, Fig. 2.41b–c $('S')C\left(L_T^4,\Theta^2\right)$, Fig. 2.41d $('S')C\left(L_T^5,\Theta^2\right)$, Fig. 2.41e $('S')C\left(L_M^5,\Theta^2\right)$, Fig. 2.41f $('S')C(L^6,\Theta^2)$, Fig. 2.41g $('S')C\left(M_{L^4}^1,\Theta^2\right)$, Fig. 2.41h $('S')C\left(K_C^1,\Theta^2\right)$, Fig. 2.41i

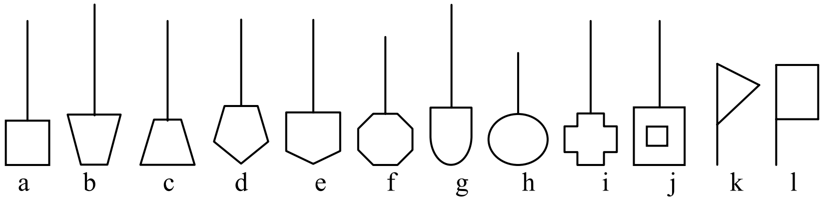


Fig. 2.41. Archetypes of the "spade" class (**a–j**) and archetypes at the classes similar to the spade class (**k–l**)

$('S')C(Q^4[L^8](4\cdot L^3),\Theta^2)$, and Fig. 2.41j $('S')C(A^1[L^4](L^4),\Theta^2)$. The notation of the spade class can be expressed in the form of the Θ/ρ class. For example, the symbolic name of the archetype in Fig. 2.41a is $\Theta/\rho\left[[L^5]\left\{2\cdot L_R^3, L_R^4\right\}\right]$, the archetype in Fig. 2.41b is $\Theta/\rho\left[[L^5]\left\{2\cdot L_R^3, L_T^4\right\}\right]$, and the archetype in Fig. 2.41c is $\Theta/\rho\left[[L^5]\left\{2\cdot Q[L^3](L^3), L_T^4\right\}\right]$. This notation makes it possible to find the difference between the archetype shown in Fig. 2.41b and the archetype shown in Fig. 2.41c. The archetypes shown in Fig. 2.41k $\Theta/\rho[[L^3]\{2L^3\}]$ and in Fig. 2.41l $\Theta/\rho\left[[L^4]\left\{L_R^3, L_R^4\right\}\right]$ are similar to archetypes shown in Fig. 2.41a–j and are not members of the spade class.

The a posteriori T-spade class is derived from the complex thin class (spade class). The archetype of the T-spade class $('S')C\left(\Gamma,(s)\otimes_1^3\right)$ instead of the handle that is a member of the thin straight class has the handle that is a member of the *s*-star class $(s)\otimes_1^3$. The s-star class $(s)\otimes_1^3$ is the thin star class whose archetypes are symmetrical and have one l branch that is significantly longer from other branches. Examples of archetypes from the s-star are shown in Fig. 2.42a–c. Archetypes shown in Fig. 2.42a, c have the symbolic name $(s)\otimes_1^3[L^3]$, whereas the archetype in Fig. 2.42b has the symbolic name $(s)\otimes_1^3[L^4]$. Examples of the archetypes from the T-spade class are shown in Fig. 2.42d–f. The symbolic names of archetypes from the T-spade class shown in Fig. 2.42 are as follows: $('S')C\left(L_T^4,(s)\otimes_1^3[L^3]\right)$ (Fig. 2.42d), $('S')C\left(L_R^4,(s)\otimes_1^3[L^3]\right)$ (Fig. 2.42e), and $('S')C\left(L_R^4,(s)\otimes_1^3[L^4]\right)$ (Fig. 2.42f).

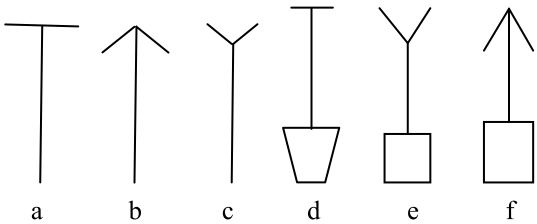


Fig. 2.42. Archetypes of the T-spade class

The a posteriori C-spade class is derived from the complex thin class (spade class). The archetype of the C-spade class consists of three parts: the blade, the handle, and the small handle. The C-spade class $('S')C(\Gamma,\Theta^2,(\varepsilon)\Lambda)$ is the class archetypes of which are obtained by joining the straight line with the object called the core and the other object called the small handle in such a way that the whole object is symmetrical. The core can be a member of the convex, the concave, or the cyclic classes. Examples of the C-spade class are shown in Fig. 2.43. The symbolic names of the archetypes shown are as follows: $('S')C\left(L_R^4,\Theta^2,(\varepsilon)L^3\right)$ (Fig. 2.43a), $('S')C\left(L_R^4,\Theta^2,(\varepsilon)M^1\right)$ (Fig. 2.43b), and $('S')C\left(L_R^4,\Theta^2,(\varepsilon)K^1\right)$ (Fig. 2.43c).

Similarly like archetypes of the spade class, the archetypes of the R-spade class are members of the complex symmetrical classes $('S')C(\Gamma,\overset{+}{\Gamma})$ consisting of two parts: one called the blade and the other one called the handle. The handle is a member of the elongated class $\overset{+}{\Gamma}$, whereas the blade is a member of one of the classes: the convex class, the concave class, or the cyclic class. In the case when both the handle and the blade are members of the convex class we have convex R-spade class

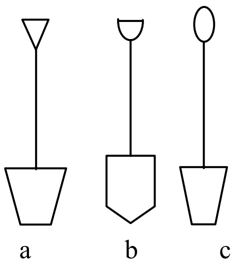
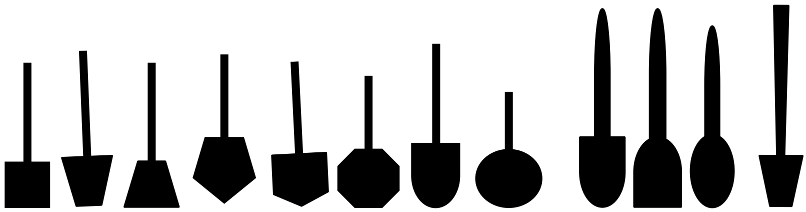


Fig. 2.43. Examples of archetypes generated from the C-spade class

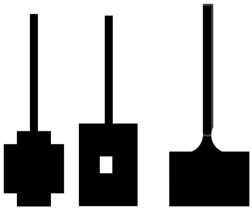


a b c d e f g h i j k l

Fig. 2.44. Exemplars generated from the convex R-spade class

$('S')C(\Lambda,\overset{+}{\Lambda})$. Examples of the convex R-spade class are shown in Fig. 2.44. The symbolic names of exemplars shown in Fig. 2.44 are as follows: $('S')C\left(L_R^4,\overset{+}{L_R^4}\right)$ (Fig. 2.44a), $('S')C\left(L_T^4,\overset{+}{L_R^4}\right)$ (Fig. 2.44b, c), $('S')C\left(L_T^5,\overset{+}{L_R^4}\right)$ (Fig. 2.44d), $('S')C\left(L_M^5,\overset{+}{L_R^4}\right)$ (Fig. 2.44e), $('S')C\left(L^6,\overset{+}{L_R^4}\right)$ (Fig. 2.44f), $('S')C\left(M_{L^4}^1,\overset{+}{L_R^4}\right)$ (Fig. 2.44g), $('S')C\left(K_C^1,\overset{+}{L_R^4}\right)$ (Fig. 2.44h), $('S')C(M^1,\overset{\times}{M^1})$ (Fig. 2.44i, j), $('S')C\left(K_E^1,\overset{\times}{M^1}\right)$ (Fig. 2.44k), and $('S')C\left(L_T^4,\overset{+}{L_T^4}\right)$ (Fig. 2.44l). The notation of the convex R-spade class can be expressed in the notation of the concave class. For example, the exemplar generated from the R-spade class shown in Fig. 2.44a–b has its symbolic name $Q^2[\widehat{L}^5]\left(2\cdot L_R^3\right)$, where symbol $\widehat{L}^5$ denotes an archetype with a small side.

In the case when the handle is a member of the convex class and the blade is a member of the concave class, we have Q-spade class $('S')C(Q,\overset{+}{\Lambda})$. In the case when the handle is a member of the convex class and the blade is a member of the cyclic class we have the A-spade class $('S')C(A,\overset{+}{\Lambda})$. In the case when both the handle and the blade are members of the concave class we have the Q-q-spade class $('S')C(Q,\overset{+}{Q})$. In the case when both the handle and the blade are members of the cyclic



a b c

Fig. 2.45. Exemplars generated from (**a**) the Q-spade class, (**b**) the A-spade class, (**c**) q-spade class

class we have the A-a-spade class $('S')C(A,\overset{+}{A})$. Example of exemplar from the Q-spade class $('S')C\left(Q^4[L^8](4L^3),\overset{+}{L}{}_R^4\right)$ is shown in Fig. 2.45a, example of exemplar from the A-spade class $('S')C\left(A^1[L^4](L^4),\overset{+}{L}{}_R^4\right)$ is shown in Fig. 2.45b, and exemplar from the q-spade class $('S')C\left(L^4,Q^2[\overset{+}{L}{}_T^4](2M^1)\right)$ is shown in Fig. 2.45c.

The spade-pike class is derived from the complex class $C(\Gamma,\overset{\times}{\Gamma})$, where $\overset{\times}{\Gamma}$ is the elongated pike class. The archetypes of the spade-convex pike class are complex classes $('\hat{S}')C(\Gamma,\overset{\times}{\Lambda})$ consisting of two parts, where one part called the handle is a member of the convex elongated pike class $\overset{\times}{\Lambda}$. The convex elongated pike class $\overset{\times}{\Lambda}$ consists of archetypes that have at least one sharp corner. Examples of exemplars generated from the convex elongated pike class are shown in Fig. 2.46. Symbolic names of exemplars shown in Fig. 2.46 are as follows: $\overset{\times}{L}{}^3$ (Fig. 2.46a), $\overset{\times}{M}{}^1$ (Fig. 2.46b), and $\overset{\times}{K}{}^1$ (Fig. 2.46c).

a b c

Fig. 2.46. Archetypes of convex elongated pike class Λ_E

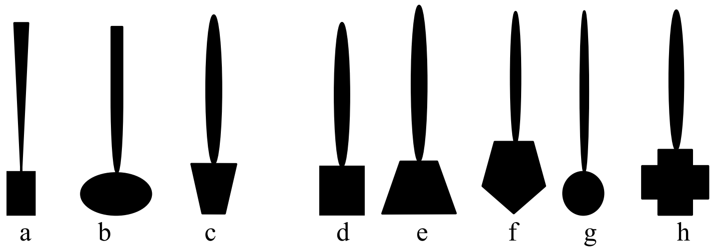


Fig. 2.47. Archetypes of the spade-pike class

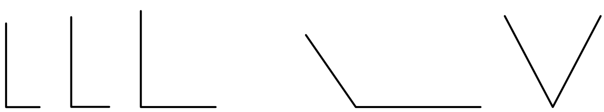
Examples of exemplars generated from the spade-convex pike class are shown in Fig. 2.47. Symbolic names of exemplars shown in Fig. 2.47 are as follows: $('\hat{S}')C\left(L_R^4, \overset{\times}{L^3}\right)$ (Fig. 2.47a), $('\hat{S}')C\left(K_E^1, \overset{\times}{M^1}\right)$ (Fig. 2.47b), $('\hat{S}')C\left(L_T^4, \overset{\times}{K^1}\right)$ (Fig. 2.47c), $('\hat{S}')C\left(L_R^4, \overset{\times}{K_E^1}\right)$ (Fig. 2.47d), $('\hat{S}')C\left(L_T^4, \overset{\times}{K_E^1}\right)$ (Fig. 2.47e), $('\hat{S}')C\left(L^5, \overset{\times}{K_E^1}\right)$ (Fig. 2.47f), $('\hat{S}')C\left(K_E^1, \overset{\times}{K_E^1}\right)$ (Fig. 2.47g), and $('\hat{S}')C\left(Q^4[L^8](4L^3), \overset{\times}{K_E^1}\right)$ (Fig. 2.47h).

2.1.2.3. The Letter Class

The a posteriori classes described in this section are derived based on the specialization of the a priori shape classes that means these classes are established to match shape of the letter. In this section the class that is derived from the thin class, which is established based on the properties of the letters, is presented. The letter class is defined based on generalization of the most often used visual representation of the letters. The archetypes of this class represent the structural archetype of the letter.

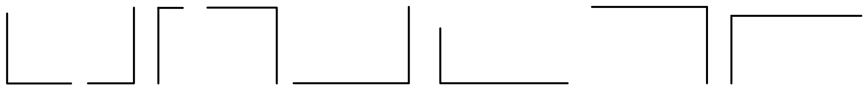
To represent a letter, the descriptions of the specific classes need to include the specific parameters that refer to the straightness of the segments, the length of the segment, the angle between segments, type of thinness, as well as the orientation of the object. The attributes such as the length are expressed by applying the graded values: $a_i^d \in \{\varepsilon, s, m, L\}$, where ε denotes a "very small," s denotes a "small," m denotes a "medium," and L denotes a "large" value. The attribute such as the angle can be expressed by applying the graded values: $a_i^\alpha \in \{\varepsilon, R, O, A\}$, where ε denotes a "very small," R denotes a "right," O denotes an "obtuse," and A denotes an “acute” angle. The orientation of the object is expressed by a selected type of the letter and a type of the transformation M – a mirror transformation and a rotation O^R in a clockwise direction by the angle $a_i^\alpha \in \{R, O, A\}$. Figure 2.48 shows archetypes of the specific class Θ^2. The letters “L,” “Γ,” “Λ,” “V,” “J,” “˥,” “˥,” “⅃” and the mathematical symbols <, ∨, 〉, ∠, ¬ are described by the symbolic names of the specific thin class shown in Figs. 2.48 and 2.49. For example, the letter “L” is given by the symbolic name $\otimes\left[L_R^3\right][l,s]$ or by adding the letter “L” in bracket “[ ]” into the name of the class $["L"]\otimes\left[L_R^3\right]$.

The symbolic names of the letter classes show similarities of the objects from these classes. This property of the symbolic name is used in process of generalization (abstraction). Archetypes shown in Fig. 2.48 are represented by the symbolic names as follows $\otimes\left[L_R^3\right][l,s]$ (Fig. 2.48a), $\otimes\left[L_R^3\right][l,m]$ (Fig. 2.48b), $\otimes\left[L_R^3\right][m,m]$ (Fig. 2.48c), $\otimes\left[L_O^3\right][m,m]$ (Fig. 2.48d), and $\otimes\left[L_A^3\right][m,m]$ (Fig. 2.48e). The generalization process shows that all objects shown in Fig. 2.49 are members of the class $\otimes[L^3]$. In order to find the proper archetype that matches a given letter the subspecific class that includes the spatial orientation of the object needs to be introduced. Figure 2.49 shows archetypes of the subspecific letter class that is established to differentiate among the different letters that are members of the same specific class $\otimes\left[L_R^3\right][l,s]$. The symbolic names of the subspecific classes are as follows: $\otimes\left[L_R^3\right][l,s]\{'L'\}$ (Fig. 2.49a), $\otimes\left[L_R^3\right][l,s]\{'L'(M)\}$ (Fig. 2.49b), $\otimes\left[L_R^3\right][l,s]\{'L'(MO^{2R})\}$ (Fig. 2.49c), $\otimes\left[L_R^3\right][l,s]\{'L'(O^{2R})\}$ (Fig. 2.49d), $\otimes\left[L_R^3\right][l,s]\{'L'(O^R)\}$ (Fig. 2.49e), $\otimes\left[L_R^3\right][l,s]\{'L'(MO^R)\}$ (Fig. 2.49f), $\otimes\left[L_R^3\right][l,s]\{'L'(MO^{3R})\}$ (Fig. 2.49g), and $\otimes\left[L_R^3\right][l,s]\{'L'(O^{3R})\}$ (Fig. 2.49h).



a b c d e

Fig. 2.48. Archetypes of the specific thin class Θ^2



a b c d e f g h

Fig. 2.49. Archetypes of the specific thin class

Understanding of the letter requires identifying the similar objects in order to be able to predict a new font or to recognize a letter that is subjected to one of many distortions. The shape classes convey information about the similarities between archetypes of the members of the different classes. For example, from the function class the specific classes are derived in order to represent the difference among letters that looks very similar. Figures 2.50 and 2.51 show examples of the archetypes of the convex function class. The concept of the function class is explained in the section "The Spade Class." The letters "V" and "U" and the mathematical symbols $<, \vee, \rangle, \angle, \neg, \succ, \cup, \bigcup, \subset, \prod, \coprod$ are described by the symbolic names of the symmetrical convex function class. Figure 2.50 shows archetypes of the subspecific letter class that represent symbols $<, \vee, \rangle, \angle, \neg, \cup, \bigcup, \subset$ and letters "V" and "U." The symbolic names of the subspecific classes for archetypes shown in Fig. 2.50 are as follows: $\otimes\left[M^1\left[K_O^1\right]\right]$ (Fig. 2.50a), $\otimes\left[L_A^3\right]$ (Fig. 2.50b), $\otimes\left[M^1\left[K_E^1\right]\right]$ (Fig. 2.50c), $\otimes\left[M^1\left[L_T^4\right]\right]$ (Fig. 2.50d), $\otimes\left[M^1\left[K_S^1\right]\right]$ (Fig. 2.50e), and $\otimes[M^1[K^2]]$ (Fig. 2.50f). Figure 2.51 shows archetypes of the subspecific letter class that represent symbols $\subset, \prod, \coprod$ and letter "U." The symbolic names of the subspecific classes for archetypes shown in Fig. 2.51 are as follows: $\otimes\left[L_R^4\right]$ (Fig. 2.51a), $\otimes\left[L_T^4\right]$ (Fig. 2.51b), $\otimes\left[L_T^4\right]$ (Fig. 2.51c), $\otimes[L^4]$ (Fig. 2.51d), $\otimes[M^1[L^3]]$ (Fig. 2.51e, f), $\otimes[M^1[L^4]]$ (Fig. 2.51g), and $\otimes[M^1[K^4]]$ (Fig. 2.51h). The mathematical symbol "$\subset$" is interpreted as the rotated version of the letter "U." Archetypes in Fig. 2.51b–d can be interpreted as the representatives of the distorted version of the symbols $\prod, \coprod$.

a b c d e f

Fig. 2.50. Archetypes of the symmetrical convex function class

a b c d e f g h

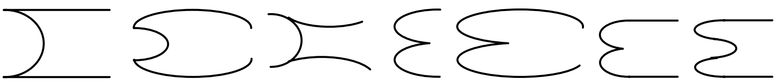
Fig. 2.51. Archetypes of the nonsymmetrical convex function class

a b c d e f

Fig. 2.52. Archetypes of the nonfunction class

Archetypes of the nonfunction class are shown in Fig. 2.52. The letter “U” can be described by the symbolic names of the nonfunction class $\otimes[L^6]$ (Fig. 2.52a). The letter “C” can be described by the symbolic names $\otimes\left[M^1\left[K_E^1\right]\right]$ (Fig. 2.52b), $\otimes[M^2[L^4]]$ (Fig. 2.52c, d), $\otimes[M^1[K^4]]$ (Fig. 2.52e), and $\otimes[M^1[K^3]]$ (Fig. 2.52f).

The letters “M” and “Σ” can have their curvilinear versions. The specific classes that can be used for description of these letters are derived from the thin polygon-curve class $\otimes$. Because there is a big range of shapes that can be used as representatives of the letters type M, the M-letter classes has to be established during learning process described in Chap. 5. In this section, examples of the archetypes from the selected M-letter classes are presented. The poly-line version of the letter type M is described in Chap. 5. The symbolic names of some of the possible curvilinear versions of the letters are given by the following notations: $\otimes[Q^1[L^4](M^1)]$ (Fig. 2.53a), $\otimes[Q^1[M^2[L^4]](M^1)]$ (Fig. 2.53b), $\otimes[Q^3[L^4](3M^1)]$ (Fig. 2.53c), $\otimes[Q^1[M^2[L^4]](Q^2[L^3](2M^1))]$ (Fig. 2.53d), $\otimes[Q^1[M^2[L^4]](Q^2[L^3](2M^1))]$ (Fig. 2.53e), $\otimes[Q^1[M^2[L^6]](Q^2[L^3](2M^1))]$ (Fig. 2.53f), and $\otimes\left[Q^1[M^2[L^6]]\left(Q^2[M^1\left[L_T^4\right]](2M^1)\right)\right]$ (Fig. 2.53g).



a b c d e f g

Fig. 2.53. Archetypes of the nonfunction classes

2.1.3. String Form: Type of the Class

Archetypes of the shape classes are described in the form of the symbolic names. For the purpose of the visual reasoning the symbolic name is transformed into the string form. The string consists of combination of the selected letters, numbers, and the symbol "|." The string has a following form: B1|…|Bi|...|Bn|, where Bi denotes the symbolic name of the class. There is a conversion from the notation of the symbolic name into the string notation. For example, the convex class L^3 is expressed as L3 in the string form.

The string notation is used to introduce the type of the class. The string without symbol "|" is denoted as the type P. It represents exemplars of the convex classes. For example, exemplars of the convex classes given in Fig. 2.54 (L3A, L4R, M1L3A, M1L4R, and M2L4R) are all of the type P.

Examples of exemplars that represent the different types of string forms are shown in Figs. 2.55–2.67. The type S that represents cyclic and concave classes, is given in the form Sn|A|1X|…|iX|…|nX|. The type Sq (the concave type) is given in the form Qn|G|1R|…|iR|…|nR|, whereas the type Sa (the cyclic type) is given as An|C|1W|…|iW|…|nW|. Examples of the exemplars type Sn|A|1X|…|iX|…|nX| are given in Figs. 2.55–2.57. The type S1|A|1_S1|1_A|1_X| and the type S1|A|1_S1|1_A|2_S1|2_A|2_X| both represent the exemplar *o* of the concave or cyclic classes on the first and the second level of iteration. The concave class $Q^4_{L^4}(4L^3)$ is expressed as Q4|L4|L3|L3|L3| in the string form. For example, an exemplar shown in Fig. 2.59a given as $A^1_{L^4_R}\left(A^1_{L^4_R}\left(A^1_{L^4_R}\left(L^4_R\right)\right)\right)$ is transformed into the string form as A1|L4R|1_A1|1_L4R|2_A1|2_L4R|2_L4R|.

Examples of the general type string forms Sn|A|1X|…|iX|…|nX| that generate the following patterns are as follows:

Q1|G|R|, A1|C|W|, Q2|G|1R|2R|, Q3|G|1R|2R|3R|, A3|C|1W|2W|3W|
A1|Q1|G|R|W|, A1|Q3|G|1R|2R|3R|W|, A2|Q1|G|R|1W|2W|
A1|Q1|G|1_Q1|1_G|R|W|, A1|Q2|G|1_Q1|1_G|1_R|R|W
A1|Q3|G|1_Q1|1_G|1_R|1R|2R|W

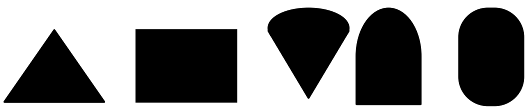


Fig. 2.54. Exemplars of the type P

Examples of general type string form S1|A|1_S1|1_A|1_X| that generates the following patterns are as follows:

Q1|G|1_Q1|1_G|R|, A1|C|1_A1|1_C|W|, Q2|G|R1|1_1Q1|1_1G|1_2R|.

Examples of the exemplars of the complex types are shown in Figs. 2.63 and 2.64.

Fig. 2.55. Exemplars of the type Q1|G|R|

Fig. 2.56. Exemplars of the type A1|G|W|

Fig. 2.57. Exemplars of the type Q2|G|1R|2R|

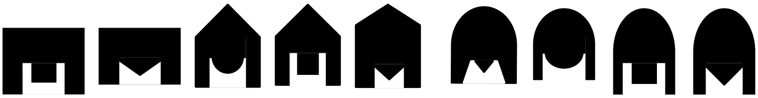


Fig. 2.58. Exemplars of the type Q1|G|1_Q1|1_G|R|

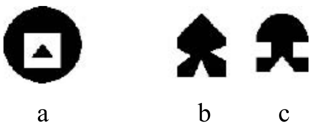


Fig. 2.59. Exemplars of the types A1|G|1_A1|1_G|W| and Q3|G|1R|2R|3R|

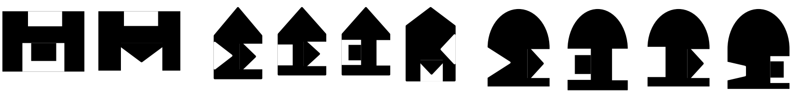


Fig. 2.60. Exemplars of the type Q2|G|R1|1_1Q1|1_1G|1_2R|

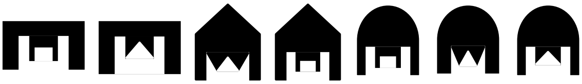


Fig. 2.61. Exemplars of the type Q1|G|1_Q1|1_G|2_Q1|2_G|2_R|

Fig. 2.62. Exemplars of the types Q3|G|1R|2R|1_1Q1|1_1G|1_R| and A1|Q2|G|1_Q1|1_G|1_R|R|W

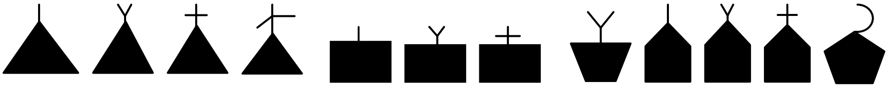


Fig. 2.63. Exemplars of the type C2|K|T|

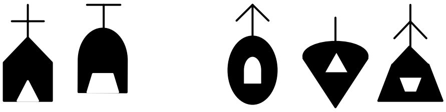


Fig. 2.64. Exemplars of the types C2|Q1|G|R|T| and C2|A1|G|W|T|

Fig. 2.65. Exemplars of the types A1|Q1|G|R|W| and A1|Q3|G|1R|2R|3R|W|

Fig. 2.66. Exemplars of the type A2|Q1|G|R|1W|2W|

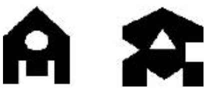


Fig. 2.67. Exemplars of the types A1|Q1|G|1_Q1|1_G|R|W and A1|Q3|G|1_Q1|1_G|1_R|1R|2R|W

2.1.4. Generalization

In the shape understanding system (SUS) a symbolic name is given in the form of the SUS representation. It is easy to translate the SUS representation into the form of the symbolic names. For example, the SUS representation C[L3,L3] is translated into the symbolic name $C(L^3, L^3)$ and the SUS representation <Q><L4>|{<L3>[O]}{<L3>[O]} is translated into the symbolic name $Q^2_{L^4}(L^3_O, L^3_O)$. Figure 2.68 illustrates the meaning of the symbols used by the SUS. The complex class is described by a symbolic name, the type of vertices, the normalized size of the sides, and the type of angles as follows: C[L3,L3], [vvvqvq], and L{ mmslle}{apaoao}. The symbolic name C[L3,L3] ($C(L^3, L^3)$) denotes an archetype of the complex class (two triangles). The term [vvvqvq] denotes the convex *v* and concave *q* vertices. The term L{mmslle} denotes the normalized size of the sides (l – large, m – medium, s – small, and e – very small). The term {apaoao} denotes angles (a – acute, o – obtuse, and p – right).

The concave class is described by the symbolic name, the type of the sides (straight or curvilinear), and the symmetry and elongatedness as follows:

<Q><L4>|{<L3>[O]}{<L3>[O]}|[AAAA][NS][El]|[AAA][NS][El]|[AAA][NS][El].

The symbolic name <Q><L4>|{<L3>[O]}{<L3>[O]}, $\left(Q^2_{L^4}\left(L^3_O, L^3_O\right)\right)$ denotes an archetype of the concave polygon class with L4 as a generic polygon and two residuals L3[O] – the obtuse triangles. The symbol [AAAA][NS][E1] denotes the polygon (straight lines – A), nonsymmetrical (NS), and medium elongated (E1).

The translation of the symbolic name into a string form requires including all details of the symbolic name. The level of details is marked by introducing the symbol "_." The symbolic name is translated into the form L0_L1_...Ln, where the level Ln denotes the level of the detailed description of the archetype of the class. For example, the triangle class L^3_O

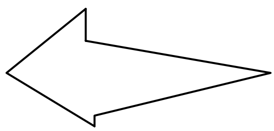


Fig. 2.68. The archetype of the complex class

Fig. 2.69. Exemplars of the class $Q^2_{L^4}\left(L^3_O, L3_A\right)$

(m,m,m) is translated into the form L_3_A_mmm. An exemplar of the concave class (Fig. 2.69a) is described by the symbolic name, the type of vertices, the normalized size of the sides, and the type of angles: $Q^2_{L^4}\left(L^3_O, L3_A\right)$, [vaqavvqv], and L{lmmsllml}{paaapaoa}. The symbolic names of exemplars of the concave class $Q^2_{L^4}\left(L^3_O, L3_A\right)$ and all detail descriptions are translated into the string form as follows:

(Fig. 2.69a) Q_1|L_4_R_mlml_1010|L_3_A_mmm_2|L_3_O_llm_0|
(Fig. 2.69b) Q_1|L_4_R_mlml_1100|L_3_A_mmm_2|L_3_O_llm_0|

During generalization the symbol is dropped from the right to the left, e.g., for the symbol L_3_A, the two generalizations are possible: L_3 and L, where "L_3_A" is any acute triangle, "L_3" is any triangle, and "L" is any polygon. In the case of the concave polygon Q_1|L_4_R|L_3_A_2| the generalization involves dropping the letters in the "ordered" manner or in the "combinatorial" manner.

An ordered manner takes into account the structural feature of the exemplar, for example, for the concave class the generic class is treated differently than residuals. The ordered manner required to compare only the "known" features of the shape.

The combinatorial manner does not distinguish between the types of the class description treating all elements of the string as the symbols of the type L0_L1_...Ln. The generalization means to drop any combination of the letters. The final step of the combinatorial manner is interpretation of the final string (the string where selected combination of the letters was removed).

Example of the string obtained during generalization performed in the "ordered" manner:

Q_1|L_4_R|L_3_A_2|Q_1|L_4_R|L_3_A|, Q_1|L_4 |L_3|, Q_1|L |L|, Q

Example of the strings obtained during generalization performed in the "combinatorial" manner:

Q1|L4R|L3A2|, Q1|L4R|L32|, Q1|L4|L3A2|, Q1|L4|L32|, Q1|L4R|L3A|, Q1|L4R|L3|, Q1|L4|L3A|, Q1|L4|L3|, Q1|L4|L|, Q1|L|L3|, Q1|L|L|, Q.

References

1. Les, Z., Shape understanding. Possible classes of shapes. *International Journal of Shape Modelling*, 2001. **7**(1): pp. 75–109
2. Preparata, F.P., and Shamos, M.I., *Computational Geometry: An Introduction.* 1985, Berlin Heidelberg New York: Springer
3. O'Rourke, J., *Computational Geometry in C.* 1998, New York: Cambridge University Press
4. Valentine, F.A., *Convex Sets*. 1964, New York: McGraw-Hill
5. Borgefors, G., Distance transformation in digital images. *Computer Vision Graphics and Image Processing*, 1986. **34**: pp. 344–371
6. Les, Z., The processing method as a set of image transformations in shape understanding. *An International Journal Computers and Graphics*, 2001. **25**(2): pp. 223–233
7. Gray, A., *Modern Differential Geometry of Curves and Surfaces*. 1992, Boca Raton, FL: CRC Press
8. Rosenfeld, A., and Johnston, E., Angle detection in digital curves. *IEEE Transactions on Computing*, 1973. **22**: pp. 875–878
9. Ansari, N., and Huang, K., Non-parametric dominant point detection. *Pattern Recognition*, 1991. **24**(9): pp. 849–862
10. Mokhtarian, F., and Mackworth, A.K., Scale-based description and recognition of planar curves and two-dimensional shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1986. **8**: pp. 34–43
11. Slupecki, J., and Borkowski, L., *Elements of Mathematical Logic and Set Theory*. 1967, Oxford: Pergamon Press
12. Farin, G., *Curves and Surfaces for Computer Aided Geometrical Design.* 1993, Boston: Academic Press
13. Brigham, E.O., *The Fast Fourier Transform.* 1974, Englewood Cliffs: Prentice-Hall
14. Bartels, R., Beatty, J., and Barsky, B., *An Introduction to Splines for Use in Computer Graphics and Geometric Modeling.* 1987, Los Altos, CA: Morgan Kaufmann
15. Meyer, Y., and Ryan, R.D., *Wavelets. Algorithms and Applications*. 1993, Philadelphia, PA: Society for Industrial and Applied Mathematics
16. Les, Z., and Les, M., Shape understanding system: the noisy class, in *International Conference on Information Systems Analysis and Synthesis and World Multiconference on Systemics, Cybernetics and Informatics, July 23–26*. 2001, Orlando
17. Les, Z., and Les, M., Shape understanding system: understanding of the complex object. *The Journal of Electronic Imaging*, 2005. **14**(2): pp. 023015-1–023015-13

18. Prusinkiewicz, P., and Lindenmayer, A., *The Algorithmic Beauty of Plants*. 1990, Berlin Heidelberg New York: Springer
19. Les, Z., and Les, M., Understanding of the concave-complex object, in *IASTED Conference Visualization, Imaging, and Image Processing*. 2003, Benalmadena, Spain: ACTA Press, Anaheim
20. Les, Z., and Les, M., Shape understanding system: understanding of the complex thin object, in *IASTED Conference Computer Graphics and Imaging*. 2004, Kauai, Hawaii

3. Digital Objects, Image Transformations, and Reasoning Process

3.1. Digital Image Representation

In Chap. 2, the shape classes used for the object description in terms of the perceptual categories are described. The perceived object has to be fitted into one of the shape categories during the reasoning process. It is assumed that visual object is extracted from image and is represented as the digital object on the background. The digital object only to some extent can approximate the object given in the continuous Euclidean space. The archetype, which is a member of the shape classes, is seen as an ideal geometrical object that is represented by its symbolic name. In this chapter the description of the reasoning process that leads to assigning the perceived object into one of the shape classes is presented.

In the presented method, a perceived object is transformed into a digital representation called the visual object given as an image region or a set of pixels. During understanding process, the perceived object called a phantom u is transformed into a digital representation called a visual object. The visual object o is given by a set of critical points $\coprod$. The terms *visual object*, *exemplar*, and *a set of critical points* can be understood as synonyms $o \equiv \coprod$. The term *visual object* is used to emphasize that an object, given as a set of pixels that can be observed on the screen, is understood as an object that is seen on the screen. The term *exemplar* is used to denote that an object, given as a set of pixels that is seen on the screen, is generated from one of the shape classes in the process called exemplar generation. The term *a set of critical points* is used to underline that that an object, given as a set of pixels, is transformed during processing stages into another set of critical points. The phantom u is transformed into a set of critical points $\coprod$ by the sensory transformation $T(u) = o$ and next into the symbolic name η by the symbolic transformation $\aleph(o) = \eta$. During visual reasoning process the symbolic name η is used to find the visual concept φ and next to assign perceived object into one of the ontological categories. Ontological categories are part of the

Z. Les and M. Les: *Digital Objects, Image Transformations, and Reasoning Process*, Studies in Computational Intelligence (SCI) **86**, 101–134 (2008)
www.springerlink.com

categorical structures of knowledge about the world comprising the visual object categories and knowledge categories described in Chap. 4. The visual object categories that represent knowledge about the world objects are given in the form of the categorical chains.

An archetype ω of the class Ω ($\omega_k \in \Omega$) is an ideal realization of shape in the two-dimensional Euclidean space (E^2). The exemplar $o \in O$ of the class Ω is a binary realization of the archetype in the discrete space. The exemplar (visual object) is one of the regions of binary image. The binary image is regarded as a set of pixels on the discrete grid (i, j). A set of all points (pixels) of a given grid is denoted as $P = \{P_0, P_1, \ldots, P_{K-1}\}$. The mapping $f(i, j)$ assigns one of the two values 0 and 1 to each pixel. The pixel for which $f(i, j) = 1$ is called the image point or the point. The exemplar (visual object) $o \in O$ is one of the regions of binary image given by a set of points $\coprod^F$ represented by the mapping

$$f(i,j) = \begin{cases} 1, & \text{if } (i,j) \in \coprod^F, \\ 0, & \text{if } (i,j) \notin \coprod^F. \end{cases}$$

The exemplar $o \in O$ is represented as a set of points $\coprod^F = \{u_1^F, u_2^F, \ldots, u_J^F\}$.

Although it was assumed that visual object is represented by binary image it does not cause a serious limitation to the presented method. The visual object which consists of parts of different colors is assigned into colored class, and during processing stages these parts are interpreted as new visual objects.

3.2. Processing Methods: Image Transformations

It was described in Sect. 3.1 that the perceived object, during reasoning process, is transformed into a set of critical points $\coprod$ and next into the symbolic name η. Perceiving object can be seen as the process of acquiring a new data. To fulfill the required task of acquiring data and processing it to obtain a set of descriptors $\Im$, the processing methods Φ are used. The processing method applies the image transformation to transform the data into one of the data types. The image transformation Θ is the mapping from one set called the domain of mapping into another called the set of mapping values. As a result of applying the image transformation into a set of critical points $\coprod$, a new set of critical points

$\coprod'$, a set of transform numbers Δ, or a set of mapping numbers Σ are obtained. The descriptor transformation $\aleph$ is applied to find a set of descriptors $\iota \in \Im$ used to assign perceived object to one of the possible classes Ω^{η} . The set of descriptors is obtained by using one of the descriptor transformations $\aleph$ on the set of critical points $\coprod$, a set of transform numbers Δ, or a set of mapping numbers Σ. Finally, a set of rules is applied that allows assigning examined object to one of the shape classes given by symbolic names.

3.2.1. Image Transformation and the Visibility Measure

As described earlier the *image transformation* Θ is a mapping from one set called the domain of mapping into another set of mapping values. Depending on the type of these sets we can distinguish the image transformation $\Theta^{\wp}$, which maps a set of critical points to one real number, the image transformations Θ^{χ} and $\Theta^{\hbar}$, which map a set of critical points into a set of real numbers, and the image transformations Θ^{λ} and Θ^{∂}, which map a set of critical points into another set of critical points. Because the image transformations Θ^{λ} and Θ^{∂} are essential part of the reasoning process, these transformations are described in more detail. The image transformation Θ^{λ} is based on the concept of visibility measure. The visibility measure is used to select subset of points from a set of critical points. The *visibility measure* is computed based on the selected neighborhood of points u_i, the function of the local parameter or relation to other set of points. The visibility measure $\lambda_{\Im}$ computed based on the selected neighborhood of the points u_k, given by their coordinates (x_k, y_k), is expressed as

$$\lambda_{\Im}(u_k) \equiv \sum_{p=-P}^{p=P} \sum_{q=-Q}^{q=Q} f^F_{i+p,j+q} T^l_{p,q},$$

where $T^l_{p,q}$ is a given template, P and Q are parameters characterizing the neighborhood and $f^F_{h,k}$ is a characteristic function given as

$$f^F_{h,k} = \begin{cases} 1, & \text{if } (x_h, y_k) \in \coprod^F, \\ 0, & \text{if } (x_h, y_k) \notin \coprod^F, \end{cases}$$

where $\coprod^F$ is a set of critical points (visual object). The visibility measure defined as the function of the local parameters, e.g., sum of the small distances, is given as

$$\lambda_{\Gamma}(u_i^{\aleph}, u_{i+k}^{\aleph}) \equiv \sum_{j=i}^{k} \hat{\delta}_j,$$

where $\hat{\delta}_j$ is a parameter computed for the given two points $u_i^{\aleph}$ and $u_{i+k}^{\aleph}$. The visibility measure given by the relation to other set of points $\cap$ is expressed as

$$\lambda(u_j) \equiv \begin{cases} 1, & \text{if } u_j \notin \cap, \\ 0, & \text{otherwise,} \end{cases}$$

where a set $\cap$ is the set of points to which a given point u_j is related.

The image transformation (discarding points) $\Theta^{\lambda} : \coprod \to \coprod$ selects the subset of points from a given set of points based on the visibility measure $\lambda_{\coprod}$. For an arbitrary point u_i, the visibility measure $\lambda(u_i)$ is applied to find if point u_i is a new critical point: $[\exists u_i \in \coprod :$ $\lambda_{\coprod'}(u_i) > T] \Rightarrow u_i \rhd u_i'$, where T is a selected threshold and u_i' is a new critical point at the higher level of processing. A set of critical points is obtained by applying one of the image transformations Θ^{λ} based on the visibility measure $\lambda^i_{\coprod^i}$ to the set of critical points of the lower level $\coprod^{i-1}$ to obtain a set of critical points $\coprod^i$. Sets of critical points can be regarded as a hierarchical structure in which the set of critical points at one level $\coprod$ can be used to obtain the set of critical points at the higher level of processing.

To represent the sequence transformations in more convenient way, the symbol Θ in image transformation is omitted and image transformations are denoted using symbol of the visibility measure, e.g., $\lambda_{\coprod^1}(\coprod^0) \to \coprod^1$ and $\lambda_{\coprod^2}(\coprod^1) \to \coprod^2$.

The image transformation (generating points) Θ^{∂} generates the subset of points based on a subset of critical points $\coprod$. For an arbitrary point u_i, the transformation $\partial(P_i)$ generates the points as follows: $[\forall u_i, u_{i+1} \in \coprod$, $\exists P_k \in P : \cap_j = \partial(P_k)] \Rightarrow P_k \rhd \cap_j$.

An example of the transformation $\partial(P_i)$ that selects points that belong to the linear segment (generate linear segment) $\overline{u_i u_{i+1}}$ is given as

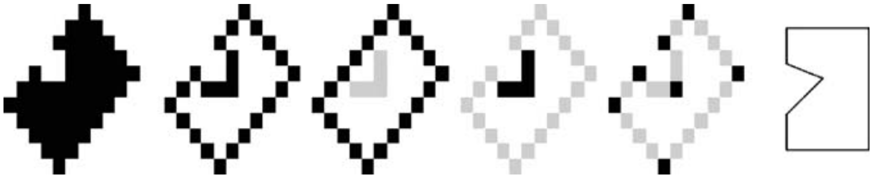


Fig. 3.1. Example of the reasoning process

$$\partial(P_i) = \begin{cases} 1, & \text{if } P_i \in \overline{u_i u_{i+1}}, \\ 0, & \text{otherwise,} \end{cases}$$

where the point $\cap_j$ is generated if pixel P_i belongs to the linear segment $P_i \in \overline{u_i u_{i+1}}$. The set of $\cap_j$ points is denoted as $\bigcap$.

An example of the application of the image transformations Θ^{λ} and Θ^{∂} is shown in Fig. 3.1. The perceived concave object is given as a set of critical points $\coprod^F$ (Fig. 3.1a, black points). After application of the image transformations $\lambda_B : \coprod^F \to \coprod^B$, a set of border points $\coprod^B$ is obtained (Fig. 3.1b). Next by application of the image transformations $\lambda_{\aleph} : \coprod^B \to \coprod^{\aleph}$ and $\partial_H : \coprod^{\aleph} \to \bigcap^H$, a set of convex hull points $\coprod^{\aleph}$ is obtained (Fig. 3.1c, black color). After application of the image transformations $\lambda_{\Psi} : \coprod^B \to \coprod^{\Psi}$ and $\lambda_{\Phi} : \coprod^B \to \coprod^{\Phi}$, a set of convex points $\coprod^{\Psi}$ (Fig. 3.1d, gray color) and a set of concave points $\coprod^{\Phi}$ (Fig. 3.1c, black color) are obtained. Finally after application of the image transformations $\lambda_{\Sigma} : \coprod^B \to \coprod^{\Sigma}$, $\lambda_{\Sigma_R} : \coprod^B \to \coprod^{\Sigma_R^k}$, $\lambda_O : \coprod^{\Sigma_R^k} \to \coprod^O$, a set of concave points $\coprod^O$ is obtained (Fig. 3.1e, black color). Fig. 3.1f shows archetype a member of the class to which perceived object is assigned.

3.3. Reasoning Process

The reasoning process that is part of the visual reasoning process is performed passing the consecutive stages of reasoning. During each stage the sequence of image transformations is applied to find a set of descriptors. The sequence of image transformations type $\Theta^{\lambda}_{\coprod} : \coprod \to \coprod$

that are used in reasoning process can be written as $\lambda_{\alpha_1} : \coprod^{\alpha_0} \to \coprod^{\alpha_1}$, $\lambda_{\alpha_2} : \coprod^{\alpha_1} \to \coprod^{\alpha_2}, \ldots, \lambda_{\alpha_M} : \coprod^{\alpha_{M-1}} \to \coprod^{\alpha_M}$ or as a composite $\lambda_{\alpha_1} \circ \lambda_{\alpha_2} \circ \cdots \circ \lambda_{\alpha_M}$: $\coprod^{\alpha_0} \to \coprod^{\alpha_M}$. In the case when the image transformation is given in the general form as $\Theta : \coprod \to W$, the sequence of image transformations is given as $L_{\alpha_1} \bullet L_{\alpha_2} \bullet \cdots \bullet L_{\alpha_M} : \coprod^{\alpha_0} \to \coprod^{\alpha_M}$, where L denotes one of the image transformations and • denotes the sequential operator. The example of the reasoning process that assigns the symbolic name to the object is given in Fig. 3.1.

The reasoning involves processing by applying one of the image transformations, computation of the descriptors using a descriptor transformation, and assigning the object to one of the possible classes.

An examined object given in Fig. 3.1 is assigned to the concave polygon class $Q[L^4](L^3)$ passing consecutive reasoning stages $Q \to Q^m \to Q^m[L^n] \to Q^m[L^n](m \bullet L^i)$ which are presented below. The reasoning stages are denoted by the symbolic name classes to which an examined object is to be assigned.

Processing stage $\zeta_0 \equiv Q$

– The image transformation:

$$\lambda_B : \coprod^F \to \coprod^B,\ \lambda_{\aleph} : \coprod^B \to \coprod^{\aleph},\ \partial_H : \coprod^{\aleph} \to \bigcap^H,$$
$$\partial_{\aleph} : \bigcap^H \to \bigcap^{\aleph}$$

– The descriptor transformation:

$$\iota_C = \aleph_C(\bigcap^{\aleph}) = \frac{|\bigcap^{\aleph}|}{|\coprod^F|} = \frac{8}{59} = 0.14$$

– The rule:

$$[\iota_C > T_C] \tilde{\Rightarrow} s \rhd \Omega^Q \quad [0.14 > T_{0.05}] \tilde{\Rightarrow} s \rhd \Omega^Q$$

Processing stage $\zeta_1 \equiv Q^m$

– The image transformation:

$$\lambda_{\Psi} : \coprod^B \to \coprod^{\Psi} \text{ and } \lambda_{\Phi} : \coprod^B \to \coprod^{\Phi}$$

– The descriptor transformation:

$$\iota_{\Phi} = \aleph_{\Phi}(\coprod^{\Phi}) = 1$$

– The rule:

$$[m = \iota_{\Phi}] \Rightarrow s \rhd \Omega^{Q^m} \quad [m = 1] \Rightarrow s \rhd \Omega^{Q^1}$$

Processing stage $\zeta_2 \equiv Q^m[L^n]$

At this processing stage, there is no need to apply the processing transformation.

– The descriptor transformation:

$$\iota_{\aleph} = \aleph_{\aleph}(\coprod^{\aleph}) = 4$$

– The rule:

$$[n = \iota_{\aleph}] \Rightarrow s \rhd Q^m[L^n] \quad [n = 4] \Rightarrow s \rhd \Omega^{Q^1_{L^4}}$$

Processing stage $\zeta_3 \equiv Q^m[L^n](n \bullet L^{h^k})$

– The image transformations:

$$\lambda_{\Sigma} : \coprod^{B} \to \coprod^{\Sigma},\ \lambda_{\Sigma_R} : \quad {}^{B} \to \coprod^{\Sigma_R^k},\ \lambda_{O} : \coprod^{\Sigma_R^k} \to \coprod^{O}$$

– The descriptor transformation:

$$\iota^k_{\Psi^k} = \aleph_{\Psi}(\coprod^{\Psi^k}) = 3$$

– The rule:

$$[h^k = \iota^k_{\Psi^k}] \Rightarrow s \rhd Q^m[L^n](n \bullet L^{h^k}) \quad [h^1 = 3] \Rightarrow s \rhd Q^1[L^4](L^3)$$

The reasoning process, expressed as the sequence of the image transformations $\lambda_B \bullet \lambda_{\aleph} \bullet \partial_H \bullet \lambda_{\Psi} \bullet \lambda_{\Sigma} \bullet \lambda_{\Sigma_R} \bullet \lambda_O : \coprod^F \to \coprod^O$, is shown in Fig. 3.1.

In the following paragraphs, examples of the reasoning process that leads to assigning the perceived object into the convex, concave, thin, or cyclic class are presented. During the reasoning process the different image transformations are applied for each class. The detail description of the reasoning process is given in [1–10].

3.3.1. Convex Object: Reasoning Process

In this chapter the reasoning process that leads to assigning the visual object into one of the convex classes is presented. Assigning the examined object to one of the convex classes Ω^{Λ} is performed during reasoning process that consists of a series of consecutive stages $\zeta_0 \to \zeta_1 \to \cdots \to \zeta_N$, where ζ_0 is the beginning stage and ζ_N is the final stage. The reasoning process involves applying one of the image transformations, computation of the descriptors using the descriptor transformation, and finally assigning an object to one of the possible convex classes given by the description ς_i.

Stage of Reasoning $\zeta_0 \equiv \Lambda$

At the beginning stage of reasoning ζ_0, it is assumed that an examined object is an exemplar of the convex class Ω^{Λ}. The reasoning process is shown in Fig. 3.2. The perceived object is given as a set of critical points $\coprod^F$ (Fig. 3.2a, gray pixels). At first the image transformation $\Theta_B^{\lambda} : \coprod^F \to \coprod^B$ is applied to compute a set of border points $\coprod^B$ (Fig. 3.2b, orange and pink pixels) and next the image transformation $\Theta_{\aleph}^{\lambda} : \coprod^B \to \coprod^{\aleph}$ is applied to compute a set of vertices $\coprod^{\aleph}$ called the *convex hull vertices* (Fig. 3.2b, pink pixels). The convex hull is obtained by using image transformation $\Theta_{\aleph}^{\lambda}$ based on the convex hull visibility measure $\lambda_{\aleph}(u_i^B)$. The point $u_i^B \in \coprod^B$ is called a *convex hull vertex* $u_i^{\aleph}$ if the visibility measure $\lambda_{\aleph}(u_i^B)$ has a minimum at this point that means $[\exists u_h^B \in \coprod^B : u_h^B = \min(\lambda(u_j^B))] \Rightarrow u_h^B \rhd u_k^{\aleph}$. The visibility measure is defined defined as a minimal angle between two selected vertices of the convex hull and each vertex not yet included on the hull, denoted as $\lambda(u_j^B) = \angle u_{i-1}^{\aleph} u_i^{\aleph} u_j^B$, for $j = h+1, \ldots, M$ and $u_i^{\aleph} = u_h^B$.

The image transformation $\Theta_H^{\partial} : \coprod^{\aleph} \to \bigcap^H$ is applied to compute a set of convex points $\bigcap^H$ (Fig. 3.2c, black pixels). The convex points are obtained based on the image transformation Θ^{∂} that generates a set of points based on a set of vertices of the convex hull $\coprod^{\aleph}$. The point $P_k \in P$ is called a *convex point*, when $P_k \in \left|\overline{u_i^{\aleph} u_{i+1}^{\aleph}}\right|$, where $\left|\overline{u_i^{\aleph} u_{i+1}^{\aleph}}\right|$ is a linear segment obtained by joining the vertices $u_i^{\aleph}$ and $u_{i+1}^{\aleph}$ of the convex hull by the straight line. The convex point P_k, denoted as $\cap_k$, is defined as follows: $[\forall u_i^{\aleph}, u_{i+1}^{\aleph} \in \coprod^{\aleph}, \exists P_k \in P : \cap_j = \partial(P_k)] \Rightarrow P_k \rhd \cap_j^H$, where the transformation $\partial(P_i)$ is defined as

$$\partial(P_i) = \begin{cases} 1, & \text{if } P_i \in \overline{u_i^{\aleph} u_{i+1}^{\aleph}}, \\ 0, & \text{otherwise,} \end{cases}$$

where $\left|\overline{u_i^{\aleph} u_{i+1}^{\aleph}}\right|$ is a linear segment obtained by joining the vertices $u_i^{\aleph}$ and $u_{i+1}^{\aleph}$ of the convex hull by the straight line. The set of convex points is denoted as $\bigcap^H$ and includes all points that constitute the convex hull polygon $\bigcap^H = \{\cap_0^H, \cap_1^H, \ldots, \cap_K^H\}$.

The image transformation $\Theta_{\aleph}^{\partial} : \bigcap^{H} \to \bigcap^{\aleph}$ is applied to compute a set of convex hull area points (Fig. 3.2d, blue pixels). The point $P_i \in P$ is called a *convex hull area point*, when the following condition is fulfilled: $[\exists P_i \in P : \cap_j^{\aleph} = \partial(P_i)] \Rightarrow P_i \rhd \cap_j^{\aleph}$, where the transformation $\partial(P_i)$ is given as

$$\partial(P_i) = \begin{cases} 1, & \text{if } P_i \in I(\bigcap^{H}), \\ 0, & \text{otherwise,} \end{cases}$$

and where $I(\bigcap^{H})$ is an interior of the $\bigcap^{H}$. The set of points $\bigcap_j^{\aleph}$ is called the *convex hull area set* and is denoted as $\bigcap^{\aleph}$.

Figure 3.2 shows the perceived object given as a set of critical points $\coprod^{F}$ (Fig. 3.2a, gray pixels) that is transformed into a set of border points $\coprod^{B}$ (Fig. 3.2b, orange and pink pixels) and next into the convex hull vertices (Fig. 3.2b, pink pixels). Based on a set of convex hull vertices $\coprod^{\aleph}$, a set of convex points $\bigcap^{H}$ (Fig. 3.2c, black pixels) is obtained. It should be noted that a set of $\coprod^{B}$ is not identical with $\bigcap^{H}$ due to not ideally convex object given by set of critical points $\coprod^{F}$. The shape category (archetype) approximates only to some degree the perceived object. The small departure from convexity can be interpreted as the result of the noise caused by a transforming phantom into a digital object. Finally a set of convex points $\bigcap^{H}$ is transformed into a set of convex hull area points $\bigcap^{\aleph}$ (Fig. 3.2d, blue pixels). Figure 3.2e shows the archetype a member of the class to which perceived object is assigned.

a b c d e

Fig. 3.2. Example of image transformations applied during reasoning process: (**a**) an exemplar before processing, (**b**) the first stage transformations $\Theta_{B}^{\lambda} : \coprod^{F} \to \coprod^{B}$ and $\Theta_{\aleph}^{\lambda} : \coprod^{B} \to \coprod^{\aleph}$, (**c**) the transformation $\Theta_{H}^{\partial} : \coprod^{\aleph} \to \bigcap^{H}$, and (**d**) the transformation $\Theta_{\aleph}^{\partial} : \bigcap^{H} \to \bigcap^{\aleph}$

The general descriptor called the *convexity coefficient* is computed as follows: $\iota_C = |\coprod^F| / |\cap^{\aleph}|$, where $|\cap^{\aleph}|$ is a number of points in the convex hull area and $|\coprod^F|$ is a number of points of the figure. An examined object is assigned to the convex class Ω^A according to the following rule: $[\iota_C < T_C] \Rightarrow s \triangleright \Omega^A$, where T_C denotes a selected threshold.

Stage of Reasoning $\zeta_1 \equiv L \vee M \vee K$

At this stage of reasoning, it is assumed that an examined object is an exemplar of one of the general classes *L*, *K*, or *M*. During the first stage of reasoning process, an examined (perceived) object $\coprod^F$ is transformed into a set of border points $\coprod^B$, a set of convex hull vertices $\coprod^{\aleph}$, a set of convex hull points $\cap^H$, and a set of convex hull area points $\cap^{\aleph}$. In the case when examined object is an ideal convex object, a set of border points $\coprod^B$ is extensionally equal to a set of convex hull points $\cap^H$, $\cap^H \equiv \coprod^B \Rightarrow \forall \cap_j^H (\cap_j^H \in \cap^H \equiv \cap_j^H \in \coprod^B)$.

The set of convex hull vertices $\coprod^{\aleph}$ obtained during the first stage of reasoning process divides the set of convex hull points into M linear segments $\cap_m^H$, where $m = 1, \ldots, M$. Because a set of convex hull vertices $\coprod^{\aleph}$ can include vertices for which the angles between consecutive linear segments $\overline{u_{m-1}^{\aleph} u_m^{\aleph}}$ and $\overline{u_m^{\aleph} u_{m+1}^{\aleph}}$ can be small, a set of a new convex hull vertices $\coprod^{\bar{\aleph}}$ is computed. A set $\coprod^{\bar{\aleph}}$ is obtained by applying the image transformation $\Theta_{\aleph}^{\lambda} : \coprod^{\aleph} \to \coprod^{\bar{\aleph}}$ based on the convex hull visibility measure $\lambda_{\bar{\aleph}}(u_i^B)$. The point $u_i^{\aleph} \in \coprod^{\aleph}$ is called an *ideal convex hull vertex* $\bar{u}_i^{\aleph}$ if the visibility measure $\lambda_{\bar{\aleph}}(u_i^{\aleph})$ fulfills the following condition $[\exists u_m^{\aleph} \in \coprod^{\aleph} : \lambda_{\bar{\aleph}}(u_m^{\aleph}) = 1] \Rightarrow u_m^{\aleph} \triangleright \bar{u}_k^{\aleph}$. The visibility measure is defined as

$$\lambda_{\bar{\aleph}}(u_m^{\aleph}) = \begin{cases} 1, & \text{if } \alpha_m < T_\alpha, \\ 0, & \text{otherwise,} \end{cases}$$

where the angles $\alpha_m = \angle u_{m-1}^{\aleph} u_m^{\aleph} u_{m+1}^{\aleph}$ between linear segments $\overline{u_{m-1}^{\aleph} u_m^{\aleph}}$ and $\overline{u_m^{\aleph} u_{m+1}^{\aleph}}$ are computed based on the formulae

$$\alpha_i = \cos^{-1}\left(\frac{(x_{i-1}-x_i)(x_{i+1}-x_i)+(y_{i-1}-y_i)(y_{i+1}-y_i)}{\sqrt{(x_{i-1}-x_i)^2+(x_{i+1}-x_i)^2}\sqrt{(y_{i-1}-y_i)^2+(y_{i+1}-y_i)^2}}\right),$$

where (x_i, y_i) denotes the coordinates of the point $u_i^{\aleph}$.

In the case when the angle between more than two consecutive segments $\overline{u}_i^{\aleph}, \overline{u}_{i+1}^{\aleph}, \ldots, \overline{u}_{i+r}^{\aleph}$ is small, the distance between the line $l_{0N} = \overline{\overline{u}_i^{\aleph}\overline{u}_{i+r}^{\aleph}}$ and each point $\overline{u}_{i+1}^{\aleph}, \overline{u}_{i+2}^{\aleph}, \ldots, \overline{u}_{i+r-1}^{\aleph}$ is computed. The point $u_i^{\overline{\aleph}} \in \amalg^{\overline{\aleph}}$ is called an *ideal convex hull vertex* $\hat{u}_i^{\aleph}$ if the visibility measure $\lambda_{\hat{\aleph}}(\overline{u}_i^{\aleph})$ fulfills the following conditions: $[\exists u_m^{\overline{\aleph}} \in \amalg^{\overline{\aleph}} : \lambda_{\hat{\aleph}}(\overline{u}_m^{\aleph}) = 1] \Rightarrow \overline{u}_m^{\aleph} \rhd \hat{u}_k^{\aleph}$. The visibility measure is defined as

$$\lambda_{\hat{\aleph}}(\overline{u}_i^{\aleph}) = \begin{cases} 1, & \text{if } d_i < T_d \text{ for } i = 0, \ldots, r-2, \\ 0, & \text{otherwise,} \end{cases}$$

where the distance d_i from the approximating line l_{0N} to a point $\overline{u}_s^{\aleph}$ given as (x_s, y_s), where $s = 1, \ldots, r-1$, is expressed as $d_i = \sqrt{(x_s - D_1)^2 + (y_s - D_2)^2}$, where $D_2 = AD_1 + B$, $D_1 = \frac{b-B}{A-a}$, $B = y_s - Ax_s$, $A = -1/a$, $b = y_N - ax_N$, and $a = \frac{y_N - y_0}{x_N - x_0}$. The point $\overline{u}_i^{\aleph}$ is given by (x_0, y_0) coordinates, the point $\overline{u}_{i+r}^{\aleph}$ is given by (x_N, y_N) coordinates, and the point $\overline{u}_s^{\aleph}$ is given by (x_s, y_s) coordinates. As a result a set of distances $\{d_0, \ldots, d_{r-2}\}$ is obtained.

The descriptor transformation uses a set of ideal convex hull points $\amalg^{\hat{\aleph}}$ to partition object into meaningful parts $\varepsilon_i \equiv \left|\overset{\frown}{u_i^{\hat{\aleph}} u_{i+k}^{\hat{\aleph}}}\right|$ based on the distance $\delta_i = \left|\overline{u_i^{\hat{\aleph}} u_{i+1}^{\hat{\aleph}}}\right|$ between consecutive points of the ideal convex hull $\amalg^{\hat{\aleph}}$. The partitioning is performed based on the following rules: $[\delta_i > T_d] \Rightarrow \varepsilon_i \rhd' A'$, $[S > T_B] \Rightarrow \varepsilon_i \rhd' B'$, where $S = \sum_{j=1}^{k} \delta_j$ if $\delta_j < T_d$ for $j = 1, \ldots, k$. Descriptor $\iota_Y =$ "$\varepsilon_1 \varepsilon_2 \cdots \varepsilon_N$" is given in the string form, where symbol ε_i can have two values A or B, $\text{val}(\varepsilon_i) \in \{A, B\}$. The string can consist of letters A and B, e.g., "AAAAA," "B," and "ABAABA." In the case when string consists of the letter A, e.g., "AAAAA," it means that perceived object consists of the linear segments and is assigning to the polygon class Ω^L according to the rule: $[\iota_Y = \text{"}n \bullet A\text{"}] \Rightarrow s \rhd \Omega^L$. In the case when string consists of a single letter B, it means that perceived object is a curve and is being assigned to the curve class Ω^K according to the rule: $[\iota_Y = \text{"}B\text{"}] \Rightarrow s \rhd \Omega^K$. In the case when string consists of letters A and B, e.g., "ABAABA," it means that perceived object is a curve-linear object and is being assigned to the curve-polygon class Ω^M according to the rule: $[\iota_Y = \text{"}{*}A * B{*}\text{"}] \Rightarrow s \rhd \Omega^M$.

Stage of Reasoning $\zeta_i \equiv L$

At this stage of reasoning, it is assumed that an examined object is an exemplar of the convex polygon class *L*. A set of the ideal convex hull $\coprod^{\hat{\aleph}}$ is used to compute a set of vertices of the convex polygon $\coprod^O$. A set of vertices of the convex polygon is obtained by using the image transformation $\Theta_O^\lambda : \coprod^{\hat{\aleph}} \to \coprod^O$ based on the convex hull visibility measure $\lambda_O(\hat{u}_i^{\aleph})$. The point $\hat{u}_i^{\aleph} \in \coprod^{\hat{\aleph}}$ is called a *vertex of the convex polygon* u_i^O if the visibility measure $\lambda_O(\hat{u}_i^{\aleph})$ fulfills the following conditions: $[\forall \hat{u}_m^{\aleph} \in \coprod^{\hat{\aleph}} : \lambda_O(\hat{u}_m^{\aleph}) = \hat{u}_m^{\aleph}] \Rightarrow \hat{u}_m^{\aleph} \rhd u_k^O$ and, as a result, a set of vertices of the convex polygon $\coprod^O$ is obtained. Next, the image transformation $\Theta_V^\partial : \coprod^O \to \bigcap^V$ is applied to compute a set of convex points $\bigcap^V$. The convex points are obtained based on the image transformation Θ^∂ that generates a set of points based on a set of vertices of the convex polygon $\coprod^O$. The point $P_k \in P$ is called a *polygon border point* and is denoted as $\cap_k$, when the following conditions are fulfilled: $[\forall u_i^O, u_{i+1}^O \in \coprod^O, \exists P_k \in P : \cap_j = \partial(P_k)] \Rightarrow P_k \rhd \cap_j^V$, where the transformation $\partial(P_i)$ is given as follows

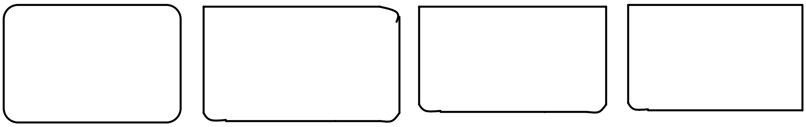
$$\partial(P_i) = \begin{cases} 1, & \text{if } P_i \in \overline{u_i^O u_{i+1}^O}, \\ 0, & \text{otherwise,} \end{cases}$$

and where $\left|\overline{u_i^O u_{i+1}^O}\right|$ is a linear segment obtained by joining the vertices u_i^O and u_{i+1}^O by the straight line. The set of polygon border points is denoted as Π^V. The set of polygon border points Π^V is matched against a set of border points $\coprod^B$ of the examined object o_j. The set of points Π^V perfectly matches the set of border points $\coprod^B$ if $\coprod^B \equiv \Pi^V$. The descriptor is called the *matching coefficient* and is computed as $\iota_V = |\Pi^V - \coprod^B|$.

An examined object is assigned to the convex polygon class Ω^{L^n} according to the following rule: $[\iota_V < T_V] \Rightarrow s \triangleright \Omega^{L^n}$, where T_V denotes a selected threshold.

Stage of Reasoning $\zeta_i \equiv \hat{L}$

An exemplar that is a digital realization of the archetype only to some extent can approximate an archetype. An exemplar is described by essential and accidental features. The essential features are part of the shape model of the class and are usually selected as the attributes of the model. The accidental features are based on the extension of the a priori model of the selected class. To show the dependence between the essential and the accidental features, an example of the convex curve-polygon class $M^n[L^4]$ is given. The archetype of the convex curve-polygon class, described in Chap. 2, consists of the linear as well as curvilinear segments. The edge that is a curved line and where values of r_i (length of the curvilinear segment) are small is called the *curvilinear corner*. In the case when archetype of the class $M^n[L^4]$ has the curvilinear corner (curvilinear segments are small), a class is called the *quadrilateral class* with "round corners" or the *corner quadrilateral class* and is denoted by the hood $\hat{L}^4$, where "4" denotes the number of the round corners. The essential features are attributes of the class $\hat{L}^4$ and the accidental features are "round corners." A class description of which that is given by the accidental qualitative attributes is denoted by the hood $\hat{L}^n$. Example of the archetypes of the $\hat{L}^4$ class is shown in Fig. 3.3. Figure 3.3 shows example of exemplars that have all curvilinear corners (Fig. 3.3a), near all curvilinear corners (Fig. 3.3b), two curvilinear corners (Fig. 3.3c) and the one curvilinear corner (Fig. 3.3d).



a b c d

Fig. 3.3. Examples of the exemplars that have (**a**) all curvilinear corners, (**b**) near all curvilinear corners, (**c**) two curvilinear corners, and (**d**) one curvilinear corner

In this stage of reasoning, it is assumed that an object is assigned to the "smooth" corner polygon class $\hat{L}^n$. To find the "smooth" corner, a set of vertices of the ideal convex hull $\coprod^{\hat{\aleph}}$ is used. The smooth corner point is defined as the point for which its visibility measure has higher critical level than a specified level. A smooth corner point is obtained by using image transformation $\Theta_O^{\lambda} : \coprod^{\hat{\aleph}} \to \coprod^{Z}$. The point $\hat{u}_i^{\aleph} \in \coprod^{\hat{\aleph}}$ is called the *smooth corner point* u_i^Z if the visibility measure $\lambda_Z(\delta_i)$ fulfills the following condition: $\left[\exists u_i^{\hat{\aleph}}, \ldots, u_{i+k}^{\hat{\aleph}} \in \coprod^{\hat{\aleph}} : \delta_i = \left|\overline{u_i^{\hat{\aleph}} u_{i+1}^{\hat{\aleph}}}\right| \wedge 0 < \lambda_Z(\delta_i) < T_Z\right] \Rightarrow (u_i^{\hat{\aleph}}, \ldots, u_{i+k}^{\hat{\aleph}}) \rhd (u_1^Z, \ldots, u_k^Z)$.
The smooth corner visibility measure is based on the sum of distances between the consecutive ideal convex hulls $u_i^{\aleph}, \ldots, u_{i+k}^{\aleph}$, and is expressed as

$$\lambda_Z(\delta_i) = \begin{cases} \sum_{i=1}^{k} \delta_i, & \text{if } \delta_i < T_d \text{ for } i = 1, \ldots, k, \\ 0, & \text{otherwise,} \end{cases}$$

where $\delta_i = \left|\overline{u_i^{\hat{\aleph}} u_{i+1}^{\hat{\aleph}}}\right|$.

The smooth corner points $u_1^Z, \ldots, u_k^Z$ are members of the set of smooth corner vertices denoted as $\coprod_h^Z$, where $h = 1, \ldots, H$ is a number of smooth corners. A set of all smooth corner vertices is denoted as $\coprod^Z = \bigcup_{h=1}^{H} \coprod_h^Z$. A set of all vertices that are not smooth corners is obtained as a set operations: $\coprod^Y = \coprod^{\hat{\aleph}} - \coprod^Z$.

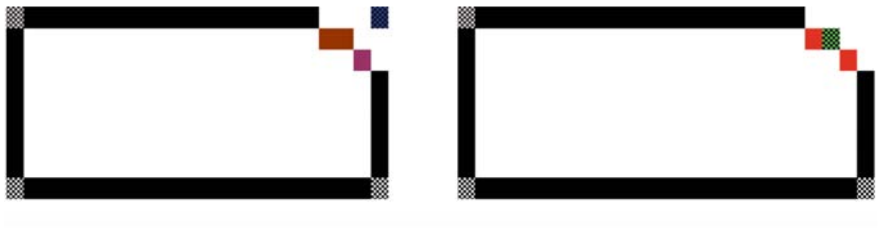


Fig. 3.4. The smooth corner points (*gray colors*) and a set of vertices of the convex polygon: (**a**) the vertex of the convex polygon obtained by applying image transformation Θ_A^{∂} and (**b**) the vertex of the convex polygon obtained by applying image transformation Θ_A^{λ}

A set of vertices of the convex polygon is obtained by using the image transformation Θ_A^{λ} or Θ_A^{∂}. Figure 3.4 shows the difference between these two methods. The point (marked by green color) shown in Fig. 3.4a is obtained by applying the image transformation Θ_A^{λ}. The image transformation $\Theta_A^{\lambda}: \coprod^Z \to \coprod^A$ is based on the visibility measure $\lambda_A(u_i^Z)$. The point $u_i^Z \in \coprod_h^Z$ is called a *vertex of the convex polygon* u_i^A if the visibility measure $\lambda_A(u_i^Z)$ fulfills the following conditions: $[u_i^Z \in \coprod_h^Z : n = \lambda_A(\coprod_h^Z)] \Rightarrow u_n^Z \rhd u_k^A$. The visibility measure is defined as

$$\lambda_A(|\coprod_h^Z|) = \begin{cases} n = |\coprod_h^Z| / 2, & \text{if} |\coprod_h^Z| \text{even}, \\ n = |\coprod_h^Z| + 1, & \text{otherwise.} \end{cases}$$

As a result of application of the image transformation Θ_A^{λ}, a set of vertices of the convex polygon $\coprod^A$ is obtained.

The point (marked by blue color) shown in Fig. 3.4b is obtained by applying the image transformation Θ_A^{∂}. The method based on the image transformation Θ_A^{∂} is applied when it can be assumed, based on the contextual information, that all linear segments are not corrupted by noise. The set of smooth corner vertices is transformed into set $\coprod_h^{\overline{Z}_1} = \{u_1^Z\}$, $\coprod_h^{\overline{Z}_2} = \{u_k^Z\}$, $\coprod^{\overline{Z}_1} = \bigcup_{h=1}^H \coprod_h^{\overline{Z}_1}$, $\coprod^{\overline{Z}_2} = \bigcup_{h=1}^H \coprod_h^{\overline{Z}_2}$, $\coprod^{W_1} = \coprod^{\overline{Z}_1} \cup \coprod^V$, and $\coprod^{W_2} = \coprod^{\overline{Z}_2} \cup \coprod^V$. The point $P_k \in P$ is called a *vertex of the polygon*, when the following conditions are fulfilled: $[\forall u_i^W u_{i+1}^W \in \coprod^{W_1}, u_i^W u_{i+1}^W \in \coprod^{W_2}, \exists P_k \in P : \cap_j = \partial_W(P_k)] \Rightarrow P_k \rhd \cap_j^B$, where the transformation $\partial(P_i)$ is given as follows

$$\partial(P_i) = \begin{cases} 1, & \text{if } P_i \in \overline{u_i^{W_1} u_i^{W_2}} \cap \overline{u_{i+1}^{W_1} u_{i+1}^{W_2}}, \\ 0, & \text{otherwise,} \end{cases}$$

where $P_i \in \overline{u_i^{W_1} u_i^{W_2}} \cap \overline{u_{i+1}^{W_1} u_{i+1}^{W_2}}$.

Point P_i is the intersection point of two lines obtained by joining the vertices $u_i^{W_1}, u_i^{W_2}$, and $u_{i+1}^{W_1}, u_{i+1}^{W_2}$. As a result the set of vertices $\coprod^A$ or $\bigcap^A$ is obtained.

The descriptor is computed as $\iota_Z = |\coprod^Z|$, where ι_Z is a number corner points. An examined object is assigned to the corner polygon class $\hat{L}^n$ according to the following rule: $[\iota_Z < T_Z] \Rightarrow s \rhd \hat{L}^n$, where T_Z denotes a selected threshold.

Stage of Reasoning $\zeta_i \equiv L^n$

At this stage of reasoning, it is assumed that an examined object is an exemplar of the convex polygon class L^n . The set of vertices of the convex polygon $\coprod^O$, described in previous section, is used to compute a descriptor. The descriptor is computed by applying a descriptor transformation: $\iota_O = \aleph_O(\coprod^O) = \left|\coprod^O\right|$. An examined object is assigned to the convex class L^n according to the following rule: $[n = \iota_0 \wedge \iota_z = 0] \Rightarrow s \rhd L^n$, where the descriptor ι_Z , described in the previous section, denotes a number of corner points.

Stage of Reasoning $\zeta_i \equiv L_R^n$

At this stage of reasoning, it is assumed that an examined object is an exemplar of the one specific polygon class. The perceived object is assigned to one of the specific classes based on the selected relations among features of the exemplars. The set of vertices of the convex polygon $\coprod^O$ is used to compute the set of features $\Phi^o = \{(f_1^d, f_1^\alpha), (f_2^d, f_2^\alpha), \ldots, (f_n^d, f_n^\alpha)\}$ of the exemplar o. The set of features Φ^O is used to assign an examined object to one of the specific convex polygon classes. The relation such as equality relation ≡ is defined as $\iota_i^a \equiv \iota_j^a \Rightarrow |\iota_i^a - \iota_j^a| < \varepsilon$ for all $i = 0,\ldots,n, j = 0,\ldots,n, i \neq j$, where $\iota_i^a \equiv \iota_j^a$. The equality relation ≡ can also be found using the clustering method. The clustering method is based on the similarity measure ϖ that for the pair of descriptors (ι_i^a, ι_j^a) assigns a real number $\delta_{i,j} = \varpi(\iota_i^a, \iota_j^a)$. The similarity measure is described on the Cartesian product of the nonempty set W as $\varpi : W \times W \to \Re_+$. As a measure of similarity the Euclidean distance is used. The distance for each descriptor is computed according to the formula $\delta_{i,j} = \sqrt{(\iota_i^{a_2} - \iota_j^{a_2})}, i = 1,\ldots,n, j = 1,\ldots,n-1$. The descriptor is computed by applying a descriptor transformation $\iota_{i,j} = \delta_{i,j}$. The rule assigning the object o to the ideal convex polygon class $\overline{L}_{u}^{n}$ using an equality relation in the form of clustering is given as $[\iota_{ij} < T_\alpha$ for all $i, j = 1,\ldots,n] \Rightarrow s \triangleright \overline{L}_{u}^{n}$ (the ideal polygon class).

Stage of Reasoning $\zeta_i \equiv M^m[L^n]$

At this stage of reasoning, it is assumed that an examined object is an exemplar of the convex curve-polygon class $M^m[L^n]$. The description of the curve-polygon class is given in Chap. 2. The exemplar of the convex curve-polygon class $M^m[L^n]$ has at least one curvilinear segment. To find the curvilinear segment $\coprod^{\exists}$, a set of vertices of the ideal convex hull $\coprod^{\hat{\aleph}}$ is used. The computation of the curvilinear segments begins with computation of the straight-linear segment. A point of the straight-linear segment is obtained by using the image transformation $\Theta_{\exists}^{\lambda} : \coprod^{\hat{\aleph}} \to \coprod^{S}$. The point $\hat{u}_i^{\aleph} \in \coprod^{\hat{\aleph}}$ is called the *point of the straight-linear segment* u_i^S if the visibility measure $\lambda_S(\delta_i)$ fulfills the following condition: $[\exists u_i^{\hat{\aleph}}, u_{i+1}^{\hat{\aleph}} \in \coprod^{\hat{\aleph}} : \lambda_S(u_i^{\hat{\aleph}}, u_{i+1}^{\hat{\aleph}}) \geq T_S] \Rightarrow (u_i^{\hat{\aleph}}, u_{i+1}^{\hat{\aleph}}) \triangleright (u_1^S, u_2^S)$. The visibility measure is expressed as $\lambda_S(u_i^{\hat{\aleph}}, u_{i+1}^{\hat{\aleph}}) = \left|\overline{u_i^{\hat{\aleph}} u_{i+1}^{\hat{\aleph}}}\right|$, where $\left|\overline{u_i^{\hat{\aleph}} u_{i+1}^{\hat{\aleph}}}\right|$ is a distance between two points. The points of the straight-linear segment $\overline{u_1^S u_2^S}$ are members of the straight-linear segment, denoted as $\coprod_m^S$, where

$m = 1,\dots,M$ is a number of straight-linear segments. A set of all straight-linear segments is denoted as $\coprod^S = \bigcup_{m=1}^{M} \coprod_m^S$.

A point of the curvilinear segment is obtained by using the image transformation $\Theta_{\exists}^{\lambda} : \coprod^{\hat{\aleph}} \to \coprod^{\exists}$. The point $\hat{u}_i^{\aleph} \in \coprod^{\hat{\aleph}}$ is called the *point of the curvilinear segment* $u_i^{\exists}$ if the visibility measure $\bar{\lambda}_{\exists}(\delta_i)$ fulfills the following condition: $\left[\exists u_i^{\hat{\aleph}},\dots,u_{i+k}^{\hat{\aleph}} \in \coprod^{\hat{\aleph}} : \delta_i = \left| \overline{u_i^{\hat{\aleph}} u_{i+1}^{\hat{\aleph}}} \right| \wedge 0 < \lambda_{\exists}(\delta_i) \geq T_{\exists} \right] \Rightarrow (u_i^{\hat{\aleph}},\dots,u_{i+k}^{\hat{\aleph}}) \triangleright (u_1^{\exists},\dots,u_k^{\exists})$. The visibility measure is based on the sum of the distances between the consecutive ideal convex hulls $u_i^{\hat{\aleph}},\dots,u_{i+k}^{\hat{\aleph}}$, and is expressed as

$$\bar{\lambda}_{\exists}(\delta_i) = \begin{cases} \sum_{i=1}^{k} \delta_i, & \text{if } \delta_i < T_d \text{ for } i = 1,\dots,k, \\ 0, & \text{otherwise,} \end{cases}$$

where $\delta_i = \left| \overline{u_i^{\hat{\aleph}} u_{i+1}^{\hat{\aleph}}} \right|$. The points of the curvilinear segment $u_1^{\exists},\dots,u_k^{\exists}$ are members of the curvilinear segment denoted as $\coprod_h^{\exists}$, where $h = 1,\dots,H$ is a number of curvilinear segments. The set of all curvilinear segments is denoted as $\coprod^{\exists} = \bigcup_{h=1}^{H} \coprod_h^{\exists}$. The descriptors are computed by applying a descriptor transformation: $\iota_{\exists} = \aleph_{\exists}(\coprod^{\exists}) = |\coprod^{\exists}|$ and $\iota_S = \aleph_S(\coprod^S) = |\coprod^S|$.

An examined segment is assigned to one of the classes based on the rules: $[\iota_{\exists} = 1 \wedge \iota_S = 1] \Rightarrow s \triangleright M^1$, $[\iota_{\exists} = 1 \wedge \iota_S = n] \Rightarrow s \triangleright M^1[L^{n+1}]$, or $[\iota_{\exists} = 2 \wedge \iota_S = n] \Rightarrow s \triangleright M^2[L^{n+2}]$. The description of the curve-polygon class M^1, $M^1[L^{n+1}]$, or $M^2[L^{n+2}]$ is given in Chap. 2.

Stage of Reasoning $\zeta_i \equiv K$

At this stage of reasoning, it is assumed that an examined object is an exemplar of the convex curve class *K*. A curvilinear point is defined as the point for which its visibility measure has higher critical level than a specified level. A curvilinear point is obtained by using the image transformation based on the curvature measure. The curvature is obtained by applying the image transformation $\Theta^{\hbar}$ that assigns the real number to each of the points $u_i^B \in \coprod^B$ based on the local transformation $\hbar_{\kappa}(u_i^B)$.

The image transformation based on the local properties is given as $[\forall u_i^B \in \coprod^B : \kappa_i = \hbar_B(u_i^B)] \Rightarrow u_i^B \triangleright \sigma_i^\kappa$, where the local transformation $\hbar_\kappa(u_i^B)$ of the point u_i^B, given by its coordinates (x_i, y_i), is given as $\hbar_\kappa(u_i^B) \equiv \dot{X}_i \ddot{Y}_i - \dot{Y}_i \ddot{X}_i$, where the discrete differences are expressed as follows:

$$\dot{X}_i = \frac{(x_{i+1} - x_{i-1})}{[(x_{i+1} - x_{i-1})^2 + (y_{i+1} - y_{i-1})^2]^{1/2}},$$

$$\dot{Y}_i = \frac{(y_{i+1} - y_{i-1})}{[(x_{i+1} - x_{i-1})^2 + (y_{i+1} - y_{i-1})^2]^{1/2}},$$

$$\ddot{X}_i = \frac{\dfrac{x_{i+1} - x_i}{[(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2]^{1/2}} - \dfrac{x_i - x_{i-1}}{[(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2]^{1/2}}}{\left(\left[\frac{1}{2}(x_{i+1} + x_i) - \frac{1}{2}(x_i + x_{i-1})\right]^2 + \left[\frac{1}{2}(y_{i+1} + y_i) - \frac{1}{2}(y_i + y_{i-1})\right]^2\right)^{1/2}},$$

$$\ddot{Y}_i = \frac{\dfrac{y_{i+1} - y_i}{[(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2]^{1/2}} - \dfrac{y_i - y_{i-1}}{[(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2]^{1/2}}}{\left(\left[\frac{1}{2}(x_{i+1} + x_i) - \frac{1}{2}(x_i + x_{i-1})\right]^2 + \left[\frac{1}{2}(y_{i+1} + y_i) - \frac{1}{2}(y_i + y_{i-1})\right]^2\right)^{1/2}}.$$

A curvilinear point is obtained by using the image transformation $\Theta_C^\lambda : \coprod^B \to \coprod^C$. The point $u_i^B \in \coprod^B$ is called the *curvilinear point* u_i^C if the visibility measure $\hat{\lambda}_C(u_i^B)$ fulfills the following condition: $[\exists u_i^B \in \coprod^B : \lambda_C(u_i^B) > T_C] \Rightarrow u_i^B \triangleright u_j^C$. The visibility measure is based on the curvature coefficient computed for this point $\lambda_C(u_i^B) = \sigma_i^\kappa$. The set of curvilinear points u_i^C is called a *curvilinear segment* $\coprod_h^C$, $h = 1, \ldots, H$. If all points $u_i^B \in \coprod^B$ are curvilinear points $H = 1$, the curvilinear segment is called the *digital convex curve* and is denoted as $\coprod^C$. The descriptor is computed by applying the descriptor transformation $\imath_C = |\coprod^B| - |\coprod^C|$. An examined segment is assigned to the convex curve class based on the rules $[\imath_C < T_C] \Rightarrow s \triangleright \Omega^K$, where T_C is a selected threshold.

Stage of Reasoning $\zeta_i \equiv K^i$

At this stage of reasoning, it is assumed that an examined object is an exemplar of the convex curve class K^i. The digital convex curve given by a set $\coprod^C$ can be represented as a sequence of 1D numbers. The transformation $\Theta^\hbar$ that assigns the real number to each of the points $u_i^C \in \coprod^C$ is based on the local transformation $\hbar_R(u_i^C)$. The 1D representation is based on the image transformation given by the formula $[\forall u_i^C \in \coprod^C : R_i = \hbar(u_i^C)] \Rightarrow u_i^C \triangleright \sigma_i^R$, where $\hbar(u_i^C)$ is the image transformation given in the form $\hbar(u_i^C) = \sqrt{(x - x_i^C)^2 + (y - y_i^C)^2}$, where x_i^C and y_i^C are the coordinates of the point u_i^C , x and y are the coordinates of the centroid of the figure points $\coprod^F$ computed as $x = m_{10} / m_{00}$ and $y = m_{01} / m_{00}$, and the moment is given as $m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} x^p y^q$. The set of distances $\Delta^R = \{\sigma_1^R, \sigma_2^R, \ldots, \sigma_N^R\}$ is used as a 1D representation of the curve. To compute the descriptors, a set of distances Δ^R is transformed into frequency domain by applying the Fourier transformations. The 1D Fourier representation is based on the global image transformation given as $[\Gamma(\Delta^r) = \Theta^F]$, where $\Gamma(\Delta^r)$ is the global image transformation given in the form of Fourier transformation. The Fourier transformation is given as a set $R_h = \sum_{k=0}^{K-1} r_k \mathrm{e}^{\mathrm{i}2\pi kh/K}$, and a set of Fourier coefficients as $\Delta^R = \{R_0, R_1, \ldots, R_H\}$. The descriptors are computed by applying a descriptor transformation $\iota_R = \max(| \sigma_i^R - \overline{\sigma}^R |)$ for all $\sigma_i^R \in \Delta^R$, where $\overline{\sigma}^R = \sum_i^N \sigma_i^R$, $\iota_\kappa^{\max} = \max(| \hat{\sigma}_i^\kappa - \hat{\overline{\sigma}}^k |)$ for all $\hat{\sigma}_i^\kappa \in \hat{\Delta}^\kappa$, where $\hat{\overline{\sigma}}^\kappa = \sum_i^N \hat{\sigma}_i^\kappa$, and $\hat{\sigma}_i^\kappa$ is the normalized curvature. The normalized curvature is obtained in such a way that the first point of the new series $\hat{\sigma}_1^\kappa$ is the maximum point $\sigma_{\max}^\kappa$ of the series before shifting $\hat{\sigma}_1^\kappa \triangleright \sigma_{\max}^\kappa$. The series are smoothed to retain only $N = 100$ values, $\hat{\Delta}^\kappa = \{\sigma_1^\kappa, \sigma_2^\kappa, \ldots, \sigma_{100}^\kappa\}$. An examined object is assigned to the convex class Ω^{K^n} according to the following rules:

$[\iota_R < T_R] \Rightarrow s \rhd \Omega^{K_C^1}$, $[\iota_\kappa > T_\kappa] \Rightarrow s \rhd \Omega^{K^2}$, $[T_\kappa \le \iota_\kappa \le T_{\kappa_1}] \Rightarrow s \rhd \Omega^{K^3}$, $[T_R \le \iota_R < T_{R_1}] \Rightarrow s \rhd \Omega^{K^4}$, where T_R, T_K, T_{R_1}, T_{κ_1} are selected thresholds. Classes K^1, K^2, K^3, and K^4 are described in Chap. 2.

3.3.2. Concave Polygon Object: Reasoning Process

In this section, the reasoning process that leads to assigning the visual object into one of the concave polygon classes is presented. Assigning an examined object to one of the concave classes Ω^Q is performed during reasoning process that consists of a series of consecutive stages $\zeta_0 \to \zeta_1 \to \cdots \to \zeta_N$, where ζ_0 is a beginning stage and ζ_N is a final stage. In the first part of the reasoning process (that is common for the convex and concave classes), the object is assigned to one of the concave classes Ω^Q by computing a number of residuals m. In the second part of this reasoning process, the type of the generic convex class is found and an object is assigned into one of the concave classes $Q^m[L]$, $Q^m[M]$, or $Q^m[K]$. The concave classes Q are described in Chap. 2. In the first stage of processing (see previous paragraphs), a set of border points $\coprod^B$, a set of convex hull vertices $\coprod^{\aleph}$, and a set of convex hull points $\bigcap^H$ are obtained. In the case when the examined object is assigned to the concave class, the set of border points $\coprod^B$ is divided into m convex $\coprod_k^{\Phi}$ and concave $\coprod_k^{\Psi}$ segments, $k = 1,\ldots,m$. The reasoning process that assigns a perceived object into one of the specific concave classes is much more complex than the reasoning process that assigns the perceived object into one of the specific convex classes. For that reason in this book, only the reasoning process that assigns a perceived object to the selected specific class, the concave polygon class, is presented.

Stage of Reasoning $\zeta_i \equiv Q_{L^n}^m$

At this stage of reasoning, it is assumed that an examined object is an exemplar of the concave parapolygon class $Q_{L^n}^m$. During this stage of the reasoning process, a set of convex vertices (a generic convex polygon) of the concave polygon is computed. A set of vertices of the generic convex polygon is obtained by using the image transformation $\Theta_O^\lambda : \coprod^{\aleph} \to \coprod^O$ that is based on the convex hull visibility measure $\lambda_O(\hat{u}_i^{\aleph})$. The point $\hat{u}_i^{\aleph} \in \coprod^{\hat{\aleph}}$ is called a *vertex of the convex polygon* u_i^O if the visibility measure $\lambda_O(\hat{u}_i^{\aleph})$ fulfills the following conditions: $[\forall \hat{u}_m^{\aleph} \in \coprod^{\hat{\aleph}} : \lambda_O(\hat{u}_m^{\aleph}) = \hat{u}_m^{\aleph}] \Rightarrow \hat{u}_m^{\aleph} \triangleright u_k^O$. As a result of application of the image transformation Θ_O^λ, a set of vertices of the generic convex polygon $\coprod^O$ is obtained. The descriptor is called the *matching coefficient* and is computed as cardinality of the set $\coprod^O$ expressed as $\iota_o = |\coprod^o|$. An examined object s is assigned to the concave parapolygon class $Q^m[L^n]$ according to the following rule: $[\iota_o = n] \Rightarrow s \triangleright Q^m[L^n]$, where n is a number of convex vertices.

Stage of Reasoning $\zeta_i \equiv Q_{L^n}^m(mL)$

At this stage of reasoning, it is assumed that an examined object is an exemplar of the concave polygon class $Q^m[L^n](mL)$ that is the class whose archetypes are concave polygons. All concave segments $\coprod_k^\Psi$, $k = 1, \ldots, m$ are transformed into the residual objects. At first the image transformation $\Theta_H^\partial : \coprod_k^\Psi \to \bigcap^{B'}$ is applied to compute a new set of border points. The border point $P_k \in P$, denoted as $\cap_k$, is defined as follows: $[\forall u_i^\Psi, u_{i+1}^\Psi \in \coprod_m^\Psi, \exists P_k \in P : \cap_j = \partial_B(P_k)] \Rightarrow P_k \triangleright \cap_j^{B'}$, where the transformation $\partial(P_i)$ is given as

$$\partial(P_i) = \begin{cases} 1, & \text{if } P_i \in \overline{u_i^\Psi u_{i+1}^\Psi}, \\ 0, & \text{otherwise,} \end{cases}$$

and where $\left|\overline{u_i^{\Psi} u_{i+1}^{\Psi}}\right|$ is a linear segment obtained by joining the vertices u_i^{Ψ} and u_{i+1}^{Ψ} of the concave segment by the straight line. Next, the image transformation $\Theta_F^{\partial} : \bigcap^{B'} \rightarrow \bigcap^{F'}$ is applied to compute a set of residual object points. The point $P_k \in P$ is called a *residual object point*, when the following condition is fulfilled: $[\exists P_k \in P : \cap_j^{F'} = \partial_F(P_k)] \Rightarrow P_k \triangleright \cap_j^{F'}$, where the transformation $\partial(P_i)$ is given as

$$\partial(P_i) = \begin{cases} 1, & \text{if } P_i \in I(\bigcap^{B'}), \\ 0, & \text{otherwise,} \end{cases}$$

and where $I(\bigcap^{B'})$ is an interior of the $\bigcap^{B'}$. The set of points $\cap_j^{F'}$ is called the *residual object area set* and is denoted as $\bigcap_k^F$. Each residual object r_k given by the set $\bigcap_k^F$ is transformed into a set of convex hull vertices $\coprod^{\aleph}$ and a set of convex hull points $\bigcap^H$, by passing stages $\zeta_0, \ldots, \zeta_3$ (described in previous sections of this chapter). An examined object is assigned to the concave polygon class $Q_{L^n}^m(mL^h)$ if all residuals r_i are the convex polygons: $[\forall i \in (1, \ldots, m), r_i \triangleright \Omega^{L^h}] \Rightarrow s \triangleright Q_{L^n}^m(mL^h)$. In the case when there is at least one residual that is assigned to the concave class: $\exists i \in (1, \ldots, m), r_i \triangleright Q$ this reasoning process is repeated at a higher level of iteration.

Stage of Reasoning $\zeta_i \equiv \perp$

At this stage of reasoning, it is assumed that an examined object is an exemplar of the concave polygon class $\perp$. The specific concave polygon class $Q_{L^n}^m(mL)$ is denoted as $\perp^n \{(n\alpha), (nd)\}$ and is uniquely given by the specification of its attributes n, α, and d, where n is the number of residuals, and α and d are attributes of the shape model. To assign an examined object to the concave polygon class $\perp$ the method based on the curvature of the object is applied. In this method the image transformation Θ^{λ} is used to the set of points $\coprod^B$ to find a set of critical points called the *linear segment* or the *corner segment*. The point is called the *straight* *segment point* if its visibility measure $\hat{\lambda}_{\Sigma}(u_j^B)$ has a value close to zero: $[\exists u_k^B \in \coprod^B :| \hat{\lambda}_{\Sigma}(u_k^B)| < T_{\Sigma}] \Rightarrow u_k^B \triangleright u_j^{\Sigma}$, where T_{Σ} is the selected threshold. The visibility measure $\hat{\lambda}_{\Sigma}(u_j^B)$ is expressed as $\hat{\lambda}_{\Sigma}(u_j^B) \equiv \kappa(u_j^B)$, where $\kappa(u_j^B)$ is the curvature of the set $\coprod^B$. The set of linear segment points u_j^{Σ} is called the *linear segment* and is denoted as $\coprod_k^{\Sigma}$. To find the corner points the mean of the curvature is computed as $\bar{\kappa} = \sum_{i=0}^{n} \kappa_i$. A point is called the *corner segment point* if its visibility measure $\hat{\lambda}_{\Sigma}(u_j^B)$ is greater than the value of $\bar{\kappa}$ and is expressed as follows: $[\exists u_k^B \in \coprod^B :| \hat{\lambda}_{\Sigma}(u_k^B)| > \bar{\kappa}] \Rightarrow u_k^B \triangleright u_j^{\Sigma_R}$. A set of corner points $u_k^{\Sigma_R}$ is called the *corner segment* and is denoted as $\coprod_k^{\Sigma_R}$. The points that are not corner points and not linear segment points are called the *curvilinear segment points* $u_k^{\exists}$ and a set of curvilinear segment points is denoted as $\coprod^{\exists}$. An examined object is assigned to the concave polygon class based on a descriptor computed as $\iota_{\exists} = \aleph_{\exists}(\coprod^{\exists}) = |\coprod^{\exists}|$, where $|\coprod^{\exists}|$ is the cardinality of a set of curvilinear segment points. The process of assigning the examined object to the concave polygon class is based on the following rule: $[\iota_{\exists} < T_{\exists}] \Rightarrow s \triangleright \perp$, where $T_{\exists}$ is a selected threshold.

Stage of Reasoning $\zeta_i \equiv \perp^n$

At this stage of reasoning, it is assumed that an examined object is an exemplar of the one of the concave polygon classes $\perp^n$.

To find a set of vertices of the examined object, the image transformation Θ^{λ} is applied to a set of corner points $\coprod_k^{\Sigma_R}$. The point $u_j^B \in \coprod_k^{\Sigma_R}$ is called a *vertex of the concave polygon* if it fulfills the following condition: $[\forall \coprod_k^{\Sigma_R}, k = 1,\ldots,K, \exists u_j^B \in \coprod_k^{\Sigma_R} : \hat{\lambda}_O(u_j^B) = 1] \Rightarrow u_j^B \triangleright u_l^O$. The visibility measure $\hat{\lambda}_O(u_j^B)$ is expressed as follows:

$$\hat{\lambda}_O(u_j^B) = \begin{cases} 1, & \text{if } \max\limits_{u_j^B \in \coprod_k^{\Sigma_R}} \kappa(u_j^B), \\ 0, & \text{otherwise.} \end{cases}$$

The set of vertices of the concave polygon is denoted as $\coprod^O$. The descriptor is computed as $\iota_O = \aleph_O(\coprod^O) = |\coprod^O|$, where $|\coprod^O|$ is the cardinality of a set $\coprod^O$. An examined object s is assigned to the concave polygon class $\perp^n$ based on the rule $[\iota_O = n] \Rightarrow s \rhd \perp^n$, where n is a number of the polygon sides.

Stage of Reasoning $\zeta_i \equiv \perp^n \{(n\alpha),(nd)\}$

At this stage of reasoning, it is assumed that an examined object is an exemplar of the one of the concave polygon classes $\perp^n \{(n\alpha),(nd)\}$. Based on the shape model M^{Ω}, an exemplar o_j of the class Ω is generated and matching between predicted and expected results is performed. The border points are obtained by applying the image transformation Θ^{∂} that generates a set of points based on a given set of generic vertices $\coprod^O$. The process of generation of points is given as follows: $[\forall v_i, v_{i+1} \in \coprod^O, \exists P_k \in P: \ \cap_j = \partial_{v_i v_{i+1}}(P_k)] \Rightarrow P_k \rhd \cap_j$, where the transformation $\partial_{v_i v_{i+1}}(P_k)$ is given as

$$\partial_{v_i v_{i+1}}(P_i) = \begin{cases} 1, & \text{if } P_i \in \overline{v_i v_{i+1}}, \\ 0, & \text{otherwise.} \end{cases}$$

The point $\cap_j$ is called a *polygon border point* if $\bigcap_j \in \left|\overline{v_i v_{i+1}}\right|$, where $\left|\overline{v_i v_{i+1}}\right|$ is a linear segment obtained by joining the vertices v_i and v_{i+1} by the straight line. The set of polygon border points is denoted as $\bigcap^V$. A set of polygon border points $\bigcap^V$ is matched against a set of border points $\coprod^B$ of the examined object o_j. The set of points $\bigcap^V$ perfectly matches a set of border points $\coprod^B$ if $\coprod^B \equiv \bigcap^V$ (all points are perfectly matched). In the case of perfect matching a set of vertices of the concave polygon $\coprod^O$ is used to compute the set of descriptors ι_α and ι_d that are used to assign an examined object to the class $\perp^n \{(n\alpha),(nd)\}$.

3.3.3. Thin Object: Reasoning Processes

The method of understanding objects that can be considered as thin objects is presented in this section. The visual object that can be regarded as a thin object can be found in many areas of research and technology. As it was described in previous sections of this book, understanding an object involves a reasoning process that assigns the examined object to one of the shape classes. In this section the reasoning process that leads to assigning the visual object into one of the thin classes is presented.

Stage of Reasoning $\zeta_1 \equiv \Theta$

In the first stage of the reasoning process, it is assumed that the perceived object is an exemplar of the thin class Ω^{Θ}. The perceived object that is given as a set of critical points $\coprod^F$ is transformed into a set of numbers called the *distance map*. The image transformation $\Theta^{\hbar}$ that assigns the number to each point $u_i^F \in \coprod^F$ is applied.

The image transformation $\Theta^{\hbar}$ is based on the local properties of the selected neighborhood and is given as follows: $[\forall u_i^F \in \coprod^F, \exists \sigma_i^{\rho} \in \Re : \sigma_i^{\rho} = \hbar_{\rho}(u_i^F)] \Rightarrow u_i \rhd \sigma_i^{\rho}$, where the local transformation $\hbar_{\rho}(u_i^F)$ is determined by the selected neighborhood. In the case of distance transformation, the local transformation is given as $\hbar_{\rho}(u_i^F) \equiv \min_{u_k^F \in \coprod^F} |u_k^F u_i^F|$, where $|u_k^F u_i^F|$ denotes the distance between a point u_i^F and an arbitrary point $u_k^F \in \coprod^F$. A set of all distance measures σ_k^{ρ} is denoted as Δ^{ρ} and is called a *distance map*. To assign a given object to the thin class Ω^{Θ} the general descriptor is computed. The general descriptor is computed as the maximum value of the set Δ^{ρ} as follows: $\iota_{\rho} = \max\{\sigma_1^{\rho}, \sigma_2^{\rho}, \ldots, \sigma_L^{\rho}\}$. The rule of assigning an examined object to the class Ω^{Θ} is given as follows: $[\iota_{\rho} < T_{\rho}] \Rightarrow s \rhd \Omega^{\Theta}$, where T_{ρ} is a selected threshold.

Stage of Reasoning $\zeta_i \equiv \Theta_m^n$

In the second stage of reasoning process, it is assumed that an examined object is an exemplar of the thin class Θ_m^n. The object given by the set $\coprod^F$ is transformed by an image transformation $\Theta_{\Im}^{\lambda} : \coprod^F \to \coprod^{\Im}$ that selects the subset of critical points from the set of points $\coprod^F$ based on the visibility measure $\lambda_{\Im}$. The point $u_i^F \in \coprod^F$ is called the *skeleton point*

$u_j^{\Im}$ if the visibility measure $\lambda_{\Im}(u_i^F)$ fulfills the following conditions: $[\exists u_i^F \in \coprod^F : \lambda_{\Im}(u_i^F) > T_{\Im}] \Rightarrow u_i^F \rhd u_j^{\Im}$, where $T_{\Im}$ is a selected threshold and $u_j^{\Im}$ is a critical point called a *skeleton point*. The visibility measure is determined by the selected neighborhood of the point u_k^F, given by its coordinates (x_k, y_k), and is expressed as follows: $\lambda_{\Im}(u_k) \equiv \sum_{p=-P}^{p=P} \sum_{q=-Q}^{q=Q} f_{i+p,j+q}^F T_{p,q}^l$, where $T_{p,q}^l$ denotes a template, P and Q are the parameters that characterize the neighborhood, and $f_{h,k}^F$ is a characteristic function given as

$$f_{h,k}^F = \begin{cases} 1, & \text{if } (x_h, y_k) \in \coprod^F, \\ 0, & \text{if } (x_h, y_k) \notin \coprod^F. \end{cases}$$

As the results a set of skeleton points $\coprod^{\Im} = \{u_1^{\Im}, u_2^{\Im}, \ldots, u_L^{\Im}\}$ is obtained. Figure 3.5 shows examples of the skeleton points.

As it was described in Chap. 2, the thin class was characterized by a set of characteristic points: the end points v^{ξ} and the branching points v^{ζ}. To assign an object to the one of the specific classes Θ_l^k, the characteristic points of the skeleton that corresponds to the characteristic points of the model need to be found. The characteristic points are obtained using image transformation $\Theta_{\theta}^{\lambda} : \coprod^{\Im} \to \coprod^{\theta}$ that is based on the visibility measure λ_{θ}. The skeleton point $u_k^{\Im}$ is called the *end point* if its visibility measure $\lambda_{\theta}(u_k^{\Im})$ fulfills the following criteria: $[\exists u_i^{\Im} \in \coprod^{\Im} : \lambda_{\theta}(u_i^{\Im}) = T_{\lambda}] \Rightarrow u_i^{\Im} \rhd u_j^{\lambda}$, where T_{λ} is a selected threshold, $T_{\lambda} = 1$. The skeleton point $u_k^{\Im}$ is called the *branching point* if its visibility $\lambda_{\theta}(u_k^{\Im})$ fulfills the following criteria: $[\exists u_i^{\Im} \in \coprod^{\Im} : \lambda_{\theta}(u_i^{\Im}) > T_{\gamma}] \Rightarrow u_i^{\Im} \rhd u_j^{\gamma}$, where T_{γ} is a selected threshold, $T_{\gamma} = 1$. The visibility measure $\lambda_{\theta}(u_k^{\Im})$ of the point $u_k^{\Im}$, given by its coordinates (x_k, y_k), is expressed as follows: $\lambda_{\Im}(u_k) \equiv \sum_{p=-P}^{p=P} \sum_{q=-Q}^{q=Q} f_{i+p,j+q}^{\Im}$, where P and Q are the parameters that characterize the neighborhood and $f_{h,k}^{\Im}$ is a characteristic function given as

$$f_{h,k}^{\Im} = \begin{cases} 1, & \text{if } (x_h, y_k) \in \coprod^{\Im}, \\ 0, & \text{if } (x_h, y_k) \notin \coprod^{\Im}. \end{cases}$$

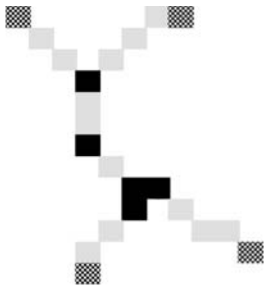


Fig. 3.5. Characteristic points of the skeleton points, end points u_i^λ, and branching points u_j^γ

A set of end points $\coprod^\lambda = \{u_1^\lambda, u_2^\lambda, \ldots, u_J^\lambda\}$ and the set of branching points $\coprod^\gamma = \{u_1^\gamma, u_2^\gamma, \ldots, u_K^\gamma\}$ are called the *set of characteristic points* denoted as $\coprod^\theta = \{u_1^\theta, u_2^\theta, \ldots, u_H^\theta\}$, $\coprod^\theta = \coprod^\lambda \cup \coprod^\gamma$, and $H = J + K$. Figure 3.5 shows characteristic points of the skeleton points, end points u_i^λ (pattern), and branching points u_j^γ (black color).

A set of descriptors is computed using the descriptor transformation given as $\iota_\lambda = |\coprod^\lambda| = J$ and $\iota_\gamma = |\coprod^\gamma| = K$, where $|\coprod|$ is a cardinality of a set $\coprod$. The rules of assigning an examined object to one of the base classes are given as follows:

$$[\iota_\lambda^\Theta = 2 \wedge \iota_\gamma^\Theta = 0] \Rightarrow s \rhd \Omega^\Xi \qquad \text{(the 1D class)},$$
$$[\iota_\lambda^\Theta = k > 2 \wedge \iota_\gamma^\Theta = 0] \Rightarrow s \rhd \Omega^{\oplus_k} \qquad \text{(the thin star class)},$$
$$[\iota_\lambda^\Theta = l \wedge \iota_\gamma^\Theta = k] \Rightarrow s \rhd \Omega^{\Theta_k^l} \qquad \text{(the thin class)}.$$

Stage of Reasoning $\zeta_i \equiv \Theta_m^n$

In this stage of reasoning, it is assumed that an examined object is an exemplar of the thin class Θ_m^n.

To assign a perceived object to one of the specific thin classes, the skeleton is divided into the parts called *branches*. A set $\coprod^\Im$ is divided into K subsets called *branches of the skeleton* denoted as $\coprod_k^\Im$. The branch

$\coprod_k^{\Im}$ is the subset of a set of skeleton points $\coprod^{\Im}$ for which the first and the last elements are members of the characteristic points $\coprod^{\theta}$. The branch $\coprod_k^{\Im}$ is obtained by applying a transformation $\Theta_{\Im_k}^{\lambda} : \coprod^{\Im} \to \coprod^{\Im_k}$. The segment point is defined as follows: $[\exists u_h^{\Im}, u_{h+t_k}^{\Im} \in \coprod^{\Im} : \lambda\!\!\!^{-}_{\Im}(u_j^{\Im}, u_h^{\Im}, u_{h+t_k}^{\Im}) \geq T_{\Im_k}] \Rightarrow u_j^{\Im} \triangleright u_i^{\Im_k}$, where $T_{\Im_k} = 1$ and the visibility measure is expressed as follows:

$$\lambda\!\!\!^{-}_{\Im_k}(u_j^{\Im}, u_h^{\Im}, u_{h+t_k}^{\Im}) \equiv \begin{cases} 1, & \text{if } h \leq j \leq h + t_k,\, u_h^{\Im}, u_{h+t_k}^{\Im} \in \coprod^{\theta}, \\ 0, & \text{otherwise.} \end{cases}$$

A branch that has one end point $\coprod_k^{\Im} \cup \coprod^{\lambda} \neq \emptyset$ and one branching point $\coprod_k^{\Im} \cup \coprod^{\gamma} \neq \emptyset$ is called the *external branch* and is denoted as $\coprod_k^{\Im_\lambda}$. A branch that does not have any end points, that means $\coprod_k^{\Im} \cup \coprod^{\lambda} = \emptyset$ and $\coprod_k^{\Im} \cup \coprod^{\gamma} \neq \emptyset$, is called the *internal branch* and is denoted as $\coprod_k^{\Im_\gamma}$. A branch for which cardinality of a set $|\coprod_k^{\Im}|$ is less than assumed threshold T_θ is called a *small branch* and is denoted as $\not{\coprod}_k^{\Im}$, where $[|\coprod_k^{\Im}| < T_\theta] \Rightarrow \coprod_k^{\Im} \triangleright \not{\coprod}_k^{\Im}$. The small external branch is denoted as $\not{\coprod}_k^{\Im_\lambda}$ whereas the small internal branch is denoted as $\not{\coprod}_k^{\Im_\gamma}$. Branches of the skeleton shown as differently shaded skeleton points are given in Fig. 3.6.

The small branch is removed if on the basis of the contextual information the small branch can be considered as a result of the “noise.”

Fig. 3.6. Branches of the skeleton shown as differently shaded skeleton points

Removing both, the small external branch and the small internal branch are obtained by applying the marking point transformation $\Theta^{\mathfrak{I}}$ that selects the subset of critical points from the set of points $\coprod^{\lambda}$ and $\coprod^{\gamma}$ based on the visibility measure $\bar{\lambda}$.

In the case of the small external branch $\coprod_k^{\Im_\lambda}$, the marked end point is defined as follows: $[\exists u_j^{\lambda} \in \coprod^{\lambda} : \bar{\lambda}_{\lambda}(u_j^{\lambda}) \geq T_{\lambda}] \Rightarrow u_j^{\lambda} \triangleright \not{u}_h^{\lambda}$, where $T_{\lambda} = 1$ and the visibility measure is expressed as follows:

$$\bar{\lambda}_{\lambda}(u_j^{\lambda}) \equiv \begin{cases} 1, & \text{if } u_j^{\lambda} \in \coprod\!\!\!/_k^{\Im_\lambda}, \\ 0, & \text{otherwise.} \end{cases}$$

The marked branching point is defined as follows: $[\exists u_j^{\gamma} \in \coprod^{\gamma}$: $\bar{\lambda}_{\gamma}(u_j^{\gamma}) \geq T_{\gamma}] \Rightarrow u_j^{\gamma} \triangleright \not{u}_h^{\gamma}$, where $T_{\lambda} = 1$. The visibility measure is expressed as follows:

$$\bar{\lambda}_{\gamma}(u_j^{\gamma}) \equiv \begin{cases} 1, & \text{if } u_j^{\gamma} \in \coprod\!\!\!/_k^{\Im_\gamma}, \\ 0, & \text{otherwise.} \end{cases}$$

The set of marked end points $\not{u}_h^{\lambda}$ is denoted as $\not\coprod^{\lambda}$ and the set of marked branching points $\not{u}_h^{\gamma}$ is denoted as $\not\coprod^{\gamma}$. In the case of the small internal branch $\coprod\!\!\!/_k^{\Im_\gamma}$, the marked branching point is defined as follows: $[\exists u_j^{\gamma} \in \coprod^{\gamma} : \bar{\lambda}_{\hat{\gamma}}(u_j^{\gamma}) \geq T_{\gamma}] \Rightarrow u_j^{\gamma} \triangleright \hat{\not{u}}_h^{\gamma}$, where $T_{\lambda} = 1$ and the visibility measure is expressed as follows:

$$\bar{\lambda}_{\hat{\gamma}}(u_j^{\gamma}) \equiv \begin{cases} 1, & \text{if } u_j^{\gamma} \in \coprod\!\!\!/_k^{\Im_\gamma}, \\ 0, & \text{otherwise.} \end{cases}$$

As a result of applying the image transformations, pair of marked points $(\hat{\not{u}}_h^{\gamma}, \hat{\not{u}}_{h+1}^{\gamma}) \in \coprod\!\!\!/_k^{\Im_\gamma}$ for each small internal segment is obtained. The point $\hat{\not{u}}_h^{\gamma}$ is a new branching point and a set of marked branching points $\{\hat{\not{u}}_1^{\gamma}, \ldots, \hat{\not{u}}_j^{\gamma}, \ldots, \hat{\not{u}}_K^{\gamma}\}$, where $j = 2h$, is denoted as $\not\coprod^{\hat{\gamma}}$.

The small branches interpreted as "noisy" are removed. Removing branches is performed as a set operations which produces the refined sets $\dot{\coprod}^{\lambda}$, $\dot{\coprod}^{\gamma}$, and $\dot{\coprod}^{\Im}$. The refined set of end points is obtained as the difference $\coprod^{\lambda} - \not\coprod^{\lambda}$, denoted as $\dot{\coprod}^{\lambda} = \coprod^{\lambda} - \not\coprod^{\lambda}$. The refined set of branching points is obtained as the difference $\coprod^{\gamma} - (\not\coprod^{\gamma} \cup \not\coprod^{\hat{\gamma}})$, where sum

$(\bar{\amalg}^{\gamma} \cup \bar{\amalg}^{\hat{\gamma}})$ denotes both marked branching points from the internal and external branches: $\dot{\amalg}^{\gamma} = \amalg^{\gamma} - (\bar{\amalg}^{\gamma} \cup \bar{\amalg}^{\hat{\gamma}})$. The refined set of skeleton points is given as the difference of the set of skeleton points $\amalg^{\Im}$ and sum of external branches, where expression $\left(\bigcup_{k=1}^{K} \amalg\!\!\!/_{k}^{\Im_{\lambda}} \cap \amalg^{\gamma}\right)$ denotes branching points that are not removed $\dot{\amalg}^{\Im} = \left(\amalg^{\Im} - \bigcup_{k=1}^{K} \amalg\!\!\!/_{k}^{\Im_{\lambda}}\right) \cup \left(\bigcup_{k=1}^{K} \amalg\!\!\!/_{k}^{\Im_{\lambda}} \cap \amalg^{\gamma}\right)$. Figure 3.7 shows process of removing of the external branch (a) the digital representation and (b) the graph representation. Figure 3.8 shows process of removing of the internal branch (a) the digital representation and (b) the graph representation.

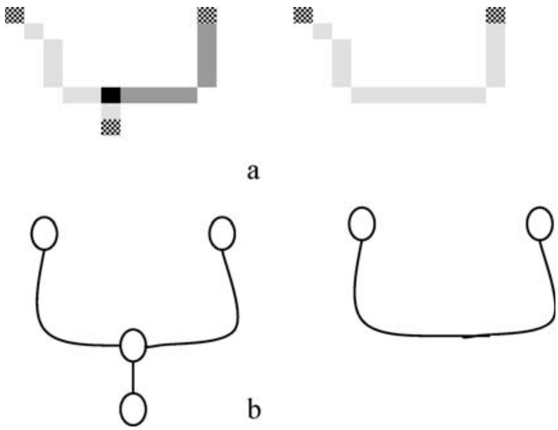


Fig. 3.7. Removing external branch: (**a**) the digital representation and (**b**) the graph representation

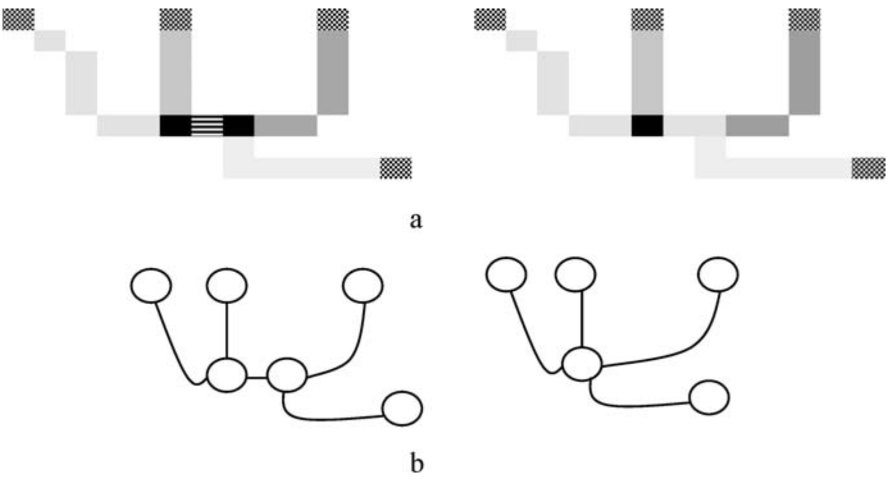


Fig. 3.8. Removing internal branch: (**a**) the digital representation and (**b**) the graph representation

As a result of the branch removing, a refined set of characteristic points $\overline{\coprod}^{\theta} = \{\overline{u}_1^{\theta}, \overline{u}_2^{\theta}, \ldots, \overline{u}_{J_\theta}^{\theta}\}$ is obtained.

A set of descriptors is computed using the descriptor transformation as a cardinality of the refined sets $\overline{\coprod}^{\lambda}$ and $\overline{\coprod}^{\gamma}$ given as $\overline{\iota}_{\lambda} = |\overline{\coprod}^{\lambda}|$ and $\overline{\iota}_{\gamma} = |\overline{\coprod}^{\gamma}|$ and is used to assign an examined object to one of the specific classes.

The rules of assigning an examined object to one of the refined base classes are given as follows:

$$[\iota_{\lambda}^{\bar{\Theta}} = 2 \wedge \overline{\iota}_{\gamma}^{\Theta} = 0] \Rightarrow s \rhd \Omega^{\Xi} \qquad \text{(the 1D class)},$$
$$[\overline{\iota}_{\lambda}^{\Theta} = k > 2 \wedge \overline{\iota}_{\gamma}^{\Theta} = 0] \Rightarrow s \rhd \Omega^{\oplus_k} \qquad \text{(the thin star class)},$$
$$[\overline{\iota}_{\lambda}^{\Theta} = l \wedge \overline{\iota}_{\gamma}^{\Theta} = k] \Rightarrow s \rhd \Omega^{\Theta_k^l} \qquad \text{(the thin class)}.$$

3.3.4. Cyclic Object: Reasoning Process

Assigning an examined object to one of the cyclic classes Ω^A is performed in the reasoning process. The reasoning involves the processing by applying one of the image transformations, computation of the descriptors using descriptor transformation, and assigning an object to one of the possible classes. In this book the reasoning process is presented for the case when a cyclic object consists of the convex core and one hole. For the objects with a more complex structure the reasoning process is much more complicated.

Stage of Reasoning $\zeta_i \equiv A$

In this stage of reasoning, it is assumed that an examined object is an exemplar of the cyclic class Ω^A.

At first the image transformation $\Theta_B^{\lambda} : \coprod^F \to \coprod^B$ and next the point–point image transformation $\Theta_{\aleph}^{\lambda} : \coprod^B \to \coprod^{\aleph}$ are applied to compute a set of vertices called a *convex hull*. Next, the image transformation $\Theta_H^{\partial} : \coprod^B \to \bigcap^H$ is applied to compute a set of convex points. The set of points $\cap_j^{\aleph}$ is called the *convex hull area set* and is denoted as $\cap^{\aleph}$. The points that are called the *hole points* are obtained as a result of subtraction of the $\cap^{\aleph}$ set and $\coprod^F$: $\coprod^H = \cap^{\aleph} - \coprod^F$. The general descriptor is computed as follows: $\iota_H = |\coprod^H|$. An examined object is assigned to the cyclic class Ω^A according to the following rule: $[\iota_H < T_H] \Rightarrow s \triangleright \Omega^A$, where T_H denotes the selected threshold.

Stage of Reasoning $\zeta_i \equiv A(A)$

In the stage of reasoning $\zeta_i \equiv A(A)$, it is assumed that an examined object is an exemplar of the cyclic class on the second level of iteration $A(A)$.

For simplicity it is assumed that the object has one hole that is represented by a set of points $\coprod^H$. Let us denote the set of $\coprod^H$ as the set $\coprod^{F'}$ that refers to the second level of iteration. Similarly to the stage $\zeta_i \equiv A$, the application of the image transformations $\Theta_B^{\lambda} : \coprod^{F'} \to \coprod^{B'}$ and $\Theta_{\aleph}^{\lambda} : \coprod^{B'} \to \coprod^{\aleph'}$ gives a set of convex hull vertices. Next, the image transformation $\Theta_H^{\partial} : \coprod^{B'} \to \cap^{H'}$ is applied to compute a set of convex points. The image transformation $\Theta_{\aleph}^{\partial} : \cap^{H'} \to \cap^{\aleph'}$ is applied to compute a set of convex hull area points. The points that are called *holes* on the second level of iteration are obtained as a result of the subtraction of the set $\cap^{\aleph'}$ and $\coprod^{F'}$: $\coprod^{H'} = \cap^{\aleph'} - \coprod^{F'}$. The general descriptor is computed as follows: $\iota_{H'} = |\coprod^{H'}|$.

An examined object is assigned to the cyclic class $A(A)$ according to the following rule: $[\iota_{H'} < T_{H'}] \Rightarrow s \triangleright A(A)$, where $T_{H'}$ denotes the selected threshold.

Assigning an object to the general cyclic class $A_{\Lambda}(A_{\Lambda})$ is performed independently for the core object and for the holes on the first level of iteration. This process is similar to assigning an object to the convex or concave class and is described in detail in Sect. 3.3.1 and Sect. 3.3.2.

References

1. Les, Z., The processing method as a set of image transformations in shape understanding. *An International Journal Computers and Graphics*, 2001. **25**(2): pp. 223–233
2. Les, Z., Shape understanding system: understanding the thin object. *An International Journal Computers and Graphics*, 2002. **26**(6): pp. 951–970
3. Les, Z., and Les, M., Shape understanding system: understanding of the convex objects. *The Journal of Electronic Imaging*, 2003. **12**(2): pp. 327–341
4. Les, Z., and Les, M., Shape understanding system: understanding of the cyclic object, in *IASTED Conference on Signal Processing and Imaging*. 2003, Rodos, Greece: ACTA Press, Anaheim
5. Les, Z., and Les, M., Understanding of the concave-complex object, in *IASTED Conference Visualization, Imaging, and Image Processing*. 2003, Benalmadena, Spain: ACTA Press, Anaheim
6. Les, Z., and Les, M., Understanding of the curve polygon object. *Journal of Computers and Graphics*, 2004. **28**(6): pp. 919–936
7. Les, Z., and Les, M., Shape understanding system-communication in natural language, in *IASTED Signal and Image Processing Conference*. 2004, Honolulu, Hawaii
8. Les, Z., and Les, M., Shape understanding system: understanding of the complex object. *The Journal of Electronic Imaging*, 2005. **14**(2): pp. 023015-1–023015-13
9. Les, Z., and Les, M., Understanding of the irregular object, in *55th Session of the International Statistical Institute (ISI)*. 2005, Sydney
10. Les, Z., and Les, M., understanding of the cyclic object as part of the visual understanding in shape understanding system, *Journal of Computer and Graphics,* 2006, Rodos, **30**(5): pp. 754–765.

4. Categories

4.1. Introduction

In Chap. 2 shape categories were introduced and briefly described. Shape categories are the main ingredient of the visual concept that is capturing the visual aspect of the world. The non-visual knowledge that is learned is represented as the categorical structure of knowledge categories. In this chapter the new knowledge representation based on the categorical chain is presented. The proposed knowledge representation plays a key role in designing systems that are able to understand environment (world objects and world phenomena). Understanding of the world is the result of thinking process. In this chapter object categories and knowledge categories are described in the context of the thinking processes.

In this section the short literature review of the research in areas connected with categories is presented. Categories are result of the mental generalization processes. Categories were subject of the theoretical considerations of many philosophical schools. In the middle ages, the problem of universal terms or class names was the topic of many tractates. These universal terms were thought of as a hierarchical structure of class names. For example, in the statement ‘Tom is a man’, ‘Tom’ is a proper name referring to a certain individual, while ‘man’ is a class name, denoting species. In another statement ‘cats are animals’, the word ‘cat’ denotes the class or species, while ‘animal’ denotes a wider class of a genus, of which cat constitutes the subclass.

The categories are result of the men’s conceptualization. In [1] the study of men’s conceptualization of their natural and social environment was presented. It was shown that while individual societies may differ considerably in their conceptualization of plants and animals, there are a number of regular structural principles of folk biological classifications which are quite general. Study of men’s conceptualization presented in [1] promises to reveal important aspects of men’s conceptual organization.

Z. Les and M. Les: *Categories*, Studies in Computational Intelligence (SCI) **86**, 135–240 (2008)
www.springerlink.com

The categories are subject of study of such disciplines as cognitive psychology, cognitive science, linguistics or artificial intelligence. In artificial intelligence, for instance, conceptual clustering described in [2] deals with classifying objects represented by structural description rather than by sequences of attribute values. In this approach objects are arranged into classes that represent a simple concept instead of the class defined solely by predefined measure of similarity among their members. Classification problem appears when there is a need to classify chemical or physical structure, analyse genetic sequence, build taxonomies of plants or animals, characterize visual scene, or split a sequence of temporal events into episodes with simple meaning.

The categories come under different names such as taxonomies. For example, botanical research is focused on the taxonomy of botanical terms. It was suggested that such terms as the sets of botanical or disease terminologies can be arranged into a hierarchical taxonomy. In colloquial English, for instance, there are objects which are called plants, and within the class of plants there is a class of trees. Within a class of trees is a class of 'needled trees'; 'needled trees' include 'pines' and pines in turn include 'jack pines'.

Some experiments in cognitive psychology revealed the mechanism of the category formation. For example, experiments described in [3] explored the hypothesis that the members of categories which are considered most prototypical are those with most attributes in common with other members of the category. Authors viewed natural semantic categories as networks of overlapping attributes. For example, a dog is understood in terms of things that have in common the following features: having four legs, having fur, barking. Some objects are seen as more reasonable exemplars of the given category than the others. For example, a chair is the more reasonable exemplar of the category of furniture than a radio, and some chairs fit the idea or image of chair better than others. The following are examples of categories used in the experiment:

- Furniture (chair, sofa, table, dresser, desk, bed, bookcase, footstool, lamp, piano, cushion, mirror, rug, radio, stove, clock, picture, closet, vase, telephone)
- Vehicle (car, truck, bus, motorcycle, train, trolley car, bicycle, airplane, boat, tractor, cart, wheelchair, tank, raft, sled, horse, blimp, skates, wheelbarrow, elevator)
- Fruit (orange, apple, banana, peach, apricot, plum, grapes, strawberry, grapefruit, pineapple, blueberry, lemon, watermelon, honeydew, pomegranate, date, coconut, tomato, olive)

– Weapon (gun, knife, sword, bomb, hang grenade, spear, cannon, bow and arrow, club, tank, teargas, whip, icepack, fists, rocket, poison, scissors, words, foot, screwdriver)
– Vegetable (peas, carrots, string beans, spinach, broccoli, asparagus, corn, cauliflower, brussel sprouts, lettuce, beets, tomato, lima beans, eggplant, onion, potato, yam, mushroom, pumpkin, rice)
– Clothing (pants, shirt, dress, skirt, jacket, coat, sweater, underpants, socks, pyjamas, bathing suit, shoes, vest, tie, mittens, hat, apron, purse, wristwatch, necklace)

These categories refer to categories which in our methods, presented in this book, are called categories of visual objects. Authors of this experiments [3] do not take into account that all these categories refer to visual objects.

In the field of computational linguistics there are attempts to build the structurally connected concepts that often are called ontology [4]. Ontology as a part of larger lexical knowledge bases and annotated resources offer an ideal starting point for constructing structured representations of word senses. For example, WordNet that includes over 110,000 concepts is a lexical ontology in which concepts correspond to word senses. Concept names (or concept labels) are called synsets. Synsets are groups of synonym words that are meant to suggest an unambiguous meaning, e.g. for bus#1: ‘bus, autobus, coach, charabanc, double-decker, jitney, motorbus, motor coach, omnibus’. In addition to synsets, the following information is provided for wordsenses:

1. Textual sense definition called gloss (e.g. coach #5: ‘a vehicle carrying many passengers; used for public transport’)
2. Hyperonymy-hyponymy links (i.e. kind of relations, e.g. bus #1 is a kind of public transport #1)
3. Meronymy-holonymy relations (i.e. part of relations, e.g. bus #1 has part roof #2 and #2)
4. Other syntactic-semantic relations

Important part of research that deals with categories are research focused on the categories of visual objects. The categories of visual objects are designed to capture the visual aspects of the world. Only a few researchers tried to investigate the problem of the visual categories. For example, the latest research in perception [5] is focused on the identification of the object in context of the visual object categories. In this research it is assumed that objects can be identified at several levels of abstraction. In [6] it was demonstrated the special status of one level of abstraction, called the basic level, at which most people tend to identify an object initially. Research with children and animals has pointed to the importance of visual features in determining the categories that are formed at different levels of abstraction (e.g. [7]). Moreover, behavioural and neuroimaging studies with adult subjects have shown that categorization at a subordinate level requires more visual processing than categorization at the basic level (e.g. [8, 9]).

To investigate the perceptual constraints on the subordinate-level categorization, there is a need for a stimulus set with several properties. In [5] the simple line drawings as stimuli were used because most studies mentioned have used these stimuli and also because shape information is the most important attribute in the subordinate-level categorization. Using line drawings instead of more complex grey-level or coloured pictures gives more control on the information that is available to the subjects. Using line drawings make it possible to minimize the effect of non-perceptual features and the availability of salient subordinate names. To obtain a reliable results there is a need for a fairly high number of subordinate exemplars from each basic-level category and the stimulus set should consists of a representative sample of basic-level categories.

In the research on the categories of the visual object the important part of the experiment is to select the proper stimulus sets. There is a number of the stimulus sets found in the literature. For example, the standard set of stimulus described in [10] was not designed to investigate multiple levels of categorization and contains only one exemplar for each of a large number of concepts. Several studies [9, 11] have used sets of line drawings with the several exemplars from the different basic-level categories, but none of these sets had many exemplars from a range of categories, and little or no information is available about naming performance. The stimulus set described in [5] consists of 269 line driving stimulus that contains several typical exemplars from a sample of 25 basic level categories:

– Non-living [tool] (glass, vase), [musical instruments] (guitar), [furniture] (chair, table, cupboard, sofa), [vehicle] (bike, automobile, train, aircraft, bicycle, ship)
– Living [animal] (mammal, fish, insect, bird) mammal <mouse, rabbit, dolphin, horse, dog, monkey>, insect <butterfly, beetle>, [plant] (tree)

Although much research concerning different aspects of categories has been carried out there is no solution to the problem how visual information can be incorporated into concept. The perceptual features that are described are not ‘translated’ into representation suitable for incorporation into conceptual structure. Also categories that are selected represent a very small part of all categories that represent the conceptual structure of knowledge about the world.

In proposed approach it is assumed that categories are part of the knowledge that is represented in the form of the connected categorical chains. The proposed method is based on an assumption that categories that play an important role in the thinking process are derived from two basic categories: categories of visual objects (visual knowledge) and categories of knowledge (non-visual knowledge). The notation of the category of the visual object is based on the categorical chain. The categorical chain is a hierarchical structure of categories derived from the categories of the visual objects. Three levels of categories are distinguished: the perceptual level, the structural level and the ontological level. In the next sections a short description of each category level as well as the description of the categorical structures of knowledge is presented.

4.2. Category of Visual Objects

To interpret (to find a meaning) of a given visual object there is a need to relay on the previously learned knowledge about the visual world. Knowledge that is stored in memory needs to have an appropriate knowledge representation. In the shape understanding method knowledge is represented as a hierarchical structure of the categories of visual objects. Categories of visual objects include the visual concept that represents visual appearance of the prototype of that category.

The visual category is the category that is related to the visual object. The categories are structured representations of the meaning of the visible objects or phenomena. Categories that are established for the purpose of this research refer to the visual objects. The visual object is an object that can be detected (sensed) by any existing sensing tools. The category of visual object such as a planet is different from the category of process such as a baking. The first one refers to an object that has a given visual appearance whereas the second one refers to relations among objects. Introduction of the general categories of visual objects facilitates process of interpretation of the perceived object by classifying it into one of general categories. The knowledge of categories is acquired during process of categorical learning. Learning of the category supplies knowledge that is needed during understanding process. An interpretation of an object in terms of the real-world object, the sign or the letter is part of understanding process connected with perception, object recognition or naming. One aspect of understanding connected with perception of the visual object is its classification to the specific perceptual category based on utilization of the context information. The context information is used to filter possible interpretations of that object by using knowledge about the world that is part of the learned categorical structure. For example, it is not very probable to find the car with wheels that have an elliptical shape. Each category is represented by its name. The general categories are established based on the general knowledge which can be acquired from all available sources. The category is described by its main characteristic features and its relations to other categories. The category of the visual object is the category from which all categories of the visual objects are derived. The main characteristic feature of this category is existence of the object that means that object can be perceived by one of the existing sensory devices.

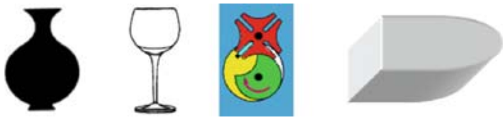
Each category is part of the conceptual structure of knowledge about the world. Categories have hierarchical structures and at the bottom of each categorical chain is the prototype. The prototype includes the visual concept that captures visual features of the object that belongs to this category. The prototype is defined during learning process at the level for which the training exemplars are available. For example, the category of a triangle needs to include all triangles such as the right, the acute or the obtuse triangle. Learning at this categorical level (the level of triangle category) requires that the learned prototype includes all triangles such as the right, the acute or the obtuse triangle. Learning at the lower level (the level of right triangle category) requires that the prototype includes only right triangles. During understanding process when the visual concept of the object of the specific category such as a right triangle category is required, the prototype supplies all knowledge that describes that object. In the case when the visual concept of the category such as a triangle category is required, the prototype of all specific categories that are linked with the triangle category supplies information that is needed. In the case when there is a need to display the representative of this category the most probably specific category is drawn by random from all specific categories linked with the category of a triangle.

Categories of the visual object are established based on the assumption that the visual object exists and can be perceived by accessible technical tools. Categories of the visual object supply knowledge about the visual aspect of the world. The non-visual knowledge is represented as a category of the body of knowledge. The notation of the category is based on the categorical chain. The categorical chain is a series of categories derived from the category of the visual object or the category of body of knowledge showing the hierarchical dependence of knowledge. The categorical chain derived from the category of the visual object is given as $\nu_0 \rhd \langle\pi\rangle \rhd \langle\sigma\rangle \rhd \langle\nu\rangle \rhd \langle\nu\rangle \ldots \rhd \langle\nu\rangle \rhd \langle\nu,\nu,\ldots,\nu\rangle$, where the categories of the visual object are derived from the visual object category ν_{O}. The category at the first level of categorical chain is called the perceptual category $\langle\pi\rangle$ of the visual object. The category at the second level of the categorical chain is called the structural category $\langle\sigma\rangle$ of the visual object. The ontological categories $\langle\nu\rangle$ begin from the third level of the categorical chain. The symbol $\triangleright$ denotes moving to the next level of categorical chain. Notation $\langle\nu,\nu,\ldots,\nu\rangle$ denotes different categories at the same level of the categorical chain. The symbol $\langle\nu,\nu,\ldots,\nu\rangle$ means that only selected categories are listed. In the case when both perceptual and structural categories are not specified the categorical chain begins with the object category denoted as $\langle O\rangle$ and the categorical chain is given as $\langle O\rangle\triangleright\langle\nu\rangle\triangleright\langle\nu\rangle\cdots\triangleright\langle\nu\rangle\triangleright\langle\nu,\nu,\ldots,\nu,\ldots\rangle$.

In this section, three levels of categories were distinguished: the perceptual level, the structural level and the ontological level. Those categorical levels are described in the following sections.

4.2.1. Perceptual Categories

The first level of the hierarchy of the categorical chain is called the perceptual category $\langle\pi\rangle$ of the visual object. At first, the object is perceived and next interpreted. The perceptual category refers to the visual representation of the visual object. Visual representation is the way in which the 3D object is presented as the 2D representative of the 3D object. Depending on the technique used to record the object the following 2D visual representations of the perceived object can be obtained: the silhouette, the line drawing or the shaded version of the object. The perceptual categories are established based on the 2D representatives of the 3D object. The perceptual category consists of the silhouette category π_{Si}, the line-drawing category π_{Ld}, the colour-object category π_{Co}, and the shaded-object category π_{Sh}. Categories at the same level of the categorical chain are represented using brackets. The chain of perceptual categories is given as $\nu_{\mathrm{O}}\triangleright\langle\pi_{\mathrm{Si}},\pi_{\mathrm{Ld}},\pi_{\mathrm{Co}},\pi_{\mathrm{Sh}}\rangle$. In the case when the perceptual category is not specified the categorical chain is given as $\nu_{\mathrm{O}}\triangleright\langle\pi\rangle$. The perceptual category is inherited by all categories of the object. Each visual object has its 2D representation given as the member of one of the perceptual categories. Figure 4.1 shows examples of the perceptual categories: the



a b c d

Fig. 4.1. Examples of members of the different perceptual categories

silhouette category (Fig. 4.1a), the line-drawing category (Fig. 4.1b), the colour-object category (Fig. 4.1c) and the shaded-object category (Fig. 4.1d).

For simplicity the term ‘an object from the silhouette category’ will be denoted as the silhouette, an object from the line drawing category will be denoted as the line drawing, an object from the shaded-object category will be denoted as the shaded object.

The silhouette can be obtained from the line drawing or the shaded object. The silhouette can be any black and white object such as a letter, a sign or a real-world object. The line drawing can be obtained by proper segmentation of the image. The silhouettes obtained from the line drawing or the shaded objects convey only a part of the visual information. Because of loss of the visual information, the silhouette is used to obtain the interpretational sketch of the object. Based on the interpretational sketch it is possible to find possible interpretations of object. To select only one interpretation the additional visual information can be acquired from the line drawing or shaded object. The symbolic name of the line-drawing object includes the symbolic name of the silhouette. Figure 4.2 shows silhouettes obtained from the line-drawing objects.

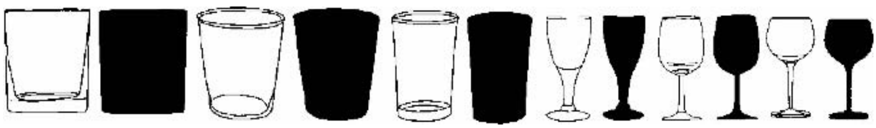


Fig. 4.2. Examples of silhouettes obtained from the line-drawing object

4.2.2. Structural Categories

The structural categories refer to the complexity of the visual representation of the object. The visual object does not often appear as an isolated object. Very often it is a part of the picture or the image that shows the object among other objects or as a part of the visual structure such as the engineering schema. Based on the relations among visual objects, the structural categories of the visual objects are established as a second level of the categorical chain. The visual object can be an isolated visual object called an element category, the object composed from the simple elements called the pattern category or a complex visual object composed from the regions that are interpreted as the elements of the different ontological categories $\langle \nu \rangle$ called the picture category. These objects are members of the second level of categories of the visual object called the structural categories of the visual objects. The structural categories of the visual objects are divided into the element category σ_{El}, the pattern category σ_{Pt}, the picture category σ_{Pi} and the animation category σ_{An}. The structural categories are shown as the second level of chain categories given as $\nu_{\mathrm{O}} \rhd \langle \pi \rangle \rhd \langle \sigma_{\mathrm{El}}, \sigma_{\mathrm{Pt}}, \sigma_{\mathrm{Pi}}, \sigma_{\mathrm{An}} \rangle$. The visual object that is built from the simple well-established visual elements called symbols is the member of the pattern category σ_{Pt}. The complex objects that is built from the different meaningfull regions is called member of the picture category. The set of elements, patterns or pictures, which are interpreted as the series of time dependent images, is called the animation category σ_{An}.

4.2.2.1. Element Category

The element category σ_{El} is a category that represents the isolated visual object of one of the ontological categories. The ontological categories $\langle \nu \rangle$ will be explained in the next paragraphs and in this section only a short description of the ontological category is given. As it was described in previous section the visual object is always given as a member of one of the perceptual categories. For example, objects in Fig. 4.3 are members of the perceptual category called silhouette, and the members of the structural category called the element category. This is represented by the categorical chain as $\nu_{\mathrm{O}} \rhd \langle \pi_{\mathrm{Si}} \rangle \rhd \langle \pi_{\mathrm{El}} \rangle$. In the case when the category of

A 5 ∑ ♠ ♪ ∂ ≡ ₣ ® § Δ ≠ ≤ % 1 / * ∏ ∫. 

Fig. 4.3. Examples of members of the element category

the perceptual level is not specified the categorical chain is denoted as $\nu_{\mathrm{O}} \rhd \langle \pi \rangle \rhd \langle \sigma_{\mathrm{El}} \rangle$ or, in a simplified notation, $\langle \Pi \rangle \rhd \langle \sigma_{\mathrm{El}} \rangle$. Figure 4.3 shows examples of members of the element category.

The ontological category is derived from the structural category. The ontological category of visual symbols, such as engineering symbols or mathematical symbols, that is derived from the structural element category is given by the following categorical chain: $\langle \Pi \rangle \rhd \langle \sigma_{\mathrm{El}} \rangle \rhd \langle \nu_{\mathrm{Sg}} \rangle \rhd \langle \nu_{\mathrm{VSym}} \rangle \rhd \langle \nu_{\mathrm{Mth}}, \nu_{\mathrm{Mus}}, \nu_{\mathrm{EnSym}} \rangle$, where $\langle \nu_{\mathrm{Sg}} \rangle$ is a sign category, $\langle \nu_{\mathrm{VSym}} \rangle$ is a visual symbol category, ν_{Mth} is a mathematical symbol category, ν_{Mus} is a musical symbol category and ν_{EnSym} is an engineering symbol category. The category of the engineering symbols, such as a resistor, a transistor, a capacitor given in Fig. 4.4, is derived from the category of the electronic engineering symbols. The category of the engineering symbols ν_{EnSym} is divided into the category of electronic symbols $\langle \nu_{\mathrm{SEnEc}} \rangle$, the category of electrical symbols $\langle \nu_{\mathrm{SEnEe}} \rangle$ or the category of mechanical symbols $\langle \nu_{\mathrm{SEnMe}} \rangle$ and is represented by the categorical chain: $\langle \Pi \rangle \rhd \langle \sigma_{\mathrm{El}} \rangle \rhd \langle \nu_{\mathrm{Sg}} \rangle \rhd \langle \nu_{\mathrm{VSym}} \rangle \rhd \langle \nu_{\mathrm{EnSym}} \rangle \rhd \langle \nu_{\mathrm{SEnEc}}, \nu_{\mathrm{SEnEe}}, \nu_{\mathrm{SEnMe}} \rangle$.

The category of electronic symbols is divided into the resistor category $\langle \nu_{\mathrm{Rez}} \rangle$, the inductor category of $\langle \nu_{\mathrm{Ind}} \rangle$, the transformer category $\langle \nu_{\mathrm{Tfo}} \rangle$, the capacitor category $\langle \nu_{\mathrm{Cap}} \rangle$, the transistor category $\langle \nu_{\mathrm{Tran}} \rangle$, and is given by the categonical chain: $\langle \Pi \rangle \rhd \langle \sigma_{\mathrm{El}} \rangle \rhd \langle \nu_{\mathrm{Sg}} \rangle \rhd \langle \nu_{\mathrm{VSym}} \rangle \rhd \langle \nu_{\mathrm{EnSym}} \rangle \rhd \langle \nu_{\mathrm{SEnEc}} \rangle \rhd$ $\langle \nu_{\mathrm{Rez}}, \nu_{\mathrm{Ind}}, \nu_{\mathrm{Tfo}}, \nu_{\mathrm{Cap}}, \nu_{\mathrm{Dio}}, \nu_{\mathrm{Tran}} \rangle$. To avoid writing the whole categorical chain, in the situation where there is no possibility of misunderstanding, the beginning of the chain will be marked by dots at the beginning of the chain $\ldots \rhd \langle \nu_{\mathrm{SEnEc}} \rangle \rhd \langle \nu_{\mathrm{Rez}}, \nu_{\mathrm{Ind}}, \nu_{\mathrm{Tfo}}, \nu_{\mathrm{Cap}}, \nu_{\mathrm{Dio}}, \nu_{\mathrm{Tran}}, . \rangle$. In this example the beginning part of the categorical chain can be easy to find from the context of the previously given description. $\ldots \rhd \langle \nu_{\mathrm{SEnEc}} \rangle \rhd \langle \nu_{\mathrm{Cap}} \rangle \rhd \langle \nu_{\mathrm{CapE}}, \nu_{\mathrm{CapC}} \rangle$. Figure 4.4 shows specific categories of the visual symbols: a capacitor, a capacitor electrolytic, a bipolar transistor and a field-effect transistor.

Fig. 4.4. The category of electronic symbols: a resistor, a capacitor, a capacitor electrolytic, a bipolar transistor, a field-effect transistor

Engineering visual symbols refer to the real-world objects. For example, electronic elements such as transistors are used to build the complex devices such as an amplifier or a TV set. The structural element category of electronic symbols refers to the structural element category of the electronic elements. The category of electronic elements is given by the following categorical chain:

$$\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle v_{\mathrm{Re}O} \rangle \triangleright \langle v_{\mathrm{Ear}} \rangle \triangleright \langle v_{\mathrm{NLiv}} \rangle \triangleright \langle v_{\mathrm{MMad}} \rangle \triangleright \langle v_{\mathrm{AsP}} \rangle \triangleright \langle v_{\mathrm{EAsP}} \rangle \triangleright \langle v_{\mathrm{Re}z}, v_{\mathrm{Ind}}, v_{\mathrm{Tfo}}, v_{\mathrm{Cap}}, v_{\mathrm{Dio}}, v_{\mathrm{Tran}} \rangle.$$

The specific electronic elements categories are derived from a category of man-made objects $\langle v_{\mathrm{MMad}} \rangle$. The ontological categories that are shown in this categorical chain will be described in detail in the following chapters. In this section we would like to show the difference between two categorical chains derived from the same structural category, namely, the element category. When we compare two categorical chains $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle v_{\mathrm{Sg}} \rangle \triangleright \langle v_{\mathrm{VSym}} \rangle \triangleright \langle v_{\mathrm{EnSym}} \rangle \triangleright \langle v_{\mathrm{SEnEc}} \rangle \triangleright \langle v_{\mathrm{Re}z}, v_{\mathrm{Ind}}, v_{\mathrm{Tfo}}, v_{\mathrm{Cap}}, v_{\mathrm{Dio}}, v_{\mathrm{Tran}} \rangle$ and $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle v_{\mathrm{Re}O} \rangle \triangleright \langle v_{\mathrm{Ear}} \rangle \triangleright \langle v_{\mathrm{NLiv}} \rangle \triangleright \langle v_{\mathrm{MMad}} \rangle \triangleright \langle v_{\mathrm{AsP}} \rangle \triangleright \langle v_{\mathrm{EAsP}} \rangle \triangleright \langle v_{\mathrm{Re}z}, v_{\mathrm{Ind}}, v_{\mathrm{Tfo}}, v_{\mathrm{Cap}}, v_{\mathrm{Dio}}, v_{\mathrm{Tran}} \rangle$ we see that both have the common part $\langle v_{\mathrm{Re}z}, v_{\mathrm{Ind}}, v_{\mathrm{Tfo}}, v_{\mathrm{Cap}}, v_{\mathrm{Dio}}, v_{\mathrm{Tran}} \rangle$ and differ in that respect that categories shown in the first chain are derived from the category of the sign $\langle v_{\mathrm{Sg}} \rangle$ whereas categories shown in the second chain are derived from the category of the real-world object $\langle v_{\mathrm{Re}O} \rangle$. These categorical chains show the difference between meaning of the visual symbols shown in Fig. 4.4 and the real-world objects shown in Fig. 4.5.

Fig. 4.5. Category of parts of electronic assembly

In Fig. 4.5 examples of real-world categories: a resistor, a capacitor, a transistor that are representatives of the category of electronic elements are shown.

4.2.2.2. Pattern Category

Schema of electronic circuits consists of symbols of simple electronic elements. The complex electronic devices such as a computer are built from relatively simple electronic elements such as transistors. These complex objects that are composed from the simple elements are members of the pattern category. The pattern category refers to objects that are composed from the simple components of the element category. Examples of the objects from the pattern category are shown in Figs. 4.6 and 4.7. For example, the pattern category of the visual symbols that are derived from the structural pattern category is shown in Fig. 4.6. From the visual symbol category, the category of mathematical patterns, the category of musical patterns, the category of coordinate system patterns or the category of pattern of the engineering symbols is derived. From comparison of two categorical chains that are derived from two different structural categories: the element category $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle$ and the pattern category $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pt}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle$, we can observe that there is no difference in the part of the chain of ontological category. However the specific categories are different. The specific visual element categories $\langle \nu_{\mathrm{Mth}}, \nu_{\mathrm{Mus}}, \nu_{\mathrm{EnSym}} \rangle$ derived from the category of visual symbols $\ldots\langle \sigma_{\mathrm{El}} \rangle \triangleright \cdots \langle \nu_{\mathrm{VSym}} \rangle$ are different from the specific pattern categories $\langle \nu_{\mathrm{MtEx}}, \nu_{\mathrm{MsNt}}, \nu_{\mathrm{EnSh}} \rangle$ derived from the category of visual patterns $\ldots\langle \sigma_{\mathrm{Pt}} \rangle \triangleright \cdots \langle \nu_{\mathrm{VSym}} \rangle$. The notation $\ldots\langle \sigma_{\mathrm{Pt}} \rangle \triangleright \cdots \langle \nu_{\mathrm{VSym}} \rangle$ indicates that only part of the categorical chain which includes the meaningful information is shown. The dots are used to show where the parts of categorical chain are missing. To illustrate the relation between the structural element categories and the structural pattern categories examples of different ontological categories will be given. For example, the category of mathematical expression $\langle \nu_{\mathrm{MtEx}} \rangle$ is derived from the pattern category of the mathematical symbols and consists of the category of equations, functions or algebraic expression $\ldots\langle \sigma_{\mathrm{Pt}} \rangle \triangleright \cdots \langle \nu_{\mathrm{MtEx}} \rangle \triangleright \langle \nu_{\mathrm{Eqt}}, \nu_{\mathrm{Fun}}, \nu_{\mathrm{Alg}} \rangle$. Mathematical expression consists of different mathematical symbols $\ldots\langle \sigma_{\mathrm{El}} \rangle \triangleright \cdots \langle \nu_{\mathrm{Mth}} \rangle \triangleright \langle \nu_{\mathrm{Cyp}}, \nu_{\mathrm{Rel}}, \nu_{\mathrm{Op}} \rangle$ that are members of the cipher

$$x + \frac{dx}{dt}x = 3x^3$$

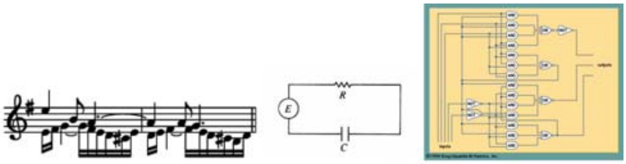
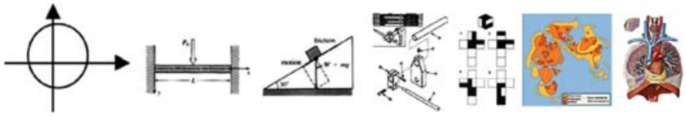


Fig. 4.6. Examples of the pattern category derived from the category of the visual symbols



a b c d e f g

Fig. 4.7. Examples of patterns consisting of elements of the different categories

category $\langle \nu_{\text{Cyp}} \rangle$, the mathematical relations category $\langle \nu_{\text{Rel}} \rangle$ or the mathematical operator category $\langle \nu_{\text{Op}} \rangle$. The mathematical symbols used to build the complex mathematical expressions are members of the category of mathematical symbols. The valid mathematical expression is formed based on the knowledge of rules of composition of mathematical expressions. Not all combinations of mathematical symbols result in valid mathematical expression. Similarly the musical symbols, such as the clef, or the note can be combined into the pattern by the following rules of the musical notation. The musical pattern can be read by musician or musical system that converts the musical pattern into the musical sound. Figure 4.6 shows examples of the pattern categories: the category of mathematical expressions, the category of musical scores and the category of the engineering schema.

The pattern category includes such complex visual objects as the engineering drawings (Fig. 4.7b–d), the mathematical objects (Fig. 4.7a), maps (Fig. 4.7f), the visual intelligence tests (Fig. 4.7e) or the biological schemas (Fig. 4.7g).

The member of the pattern category, which is composed of different objects that belong to the category of mathematical elements such as axes or graphs, is called the category of the coordinate system (Fig. 4.7a). The member of category of the coordinate system consists of many different objects that are members of the category of mathematical elements such as axes $\langle \nu_{\text{Axi}} \rangle$, labels $\langle \nu_{\text{Lab}} \rangle$, frames $\langle \nu_{\text{Frm}} \rangle$ or graphs $\langle \nu_{\text{Grh}} \rangle$. The categorical chain of the elements of the coordinate system is as follows: $\ldots\langle \sigma_{\text{El}} \rangle \rhd \cdots \langle \nu_{\text{CoSym}} \rangle \rhd \langle \nu_{\text{Axi}}, \nu_{\text{Lab}}, \nu_{\text{Frm}}, \nu_{\text{Grh}} \rangle$. The category of mathematical coordinate system is divided into the category of the Cartesian coordinate system or the category of the polar coordinate system and is given as $\ldots\langle\sigma_{\text{Pt}}\rangle \rhd \cdots\langle\nu_{\text{CoSys}}\rangle \rhd \langle\nu_{\text{Cart}},\nu_{\text{Pol}}\rangle$. Coordinates are sets of numbers that specify the location of a point in a space. For example, the coordinates of the polar coordinate system are written as (r,θ), in which *r* is the distance from the origin to any desired point and θ is the angle made by the vector *r* and the axis. A simple relationship exists between Cartesian coordinates given in terms of two reference axes (x,y) and the polar coordinates (r,θ), namely: $x = r\cos\theta$, and $y = r\sin\theta$.

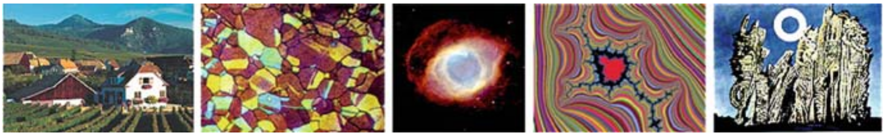
The physical visual models shown in Fig. 4.7b–c consist of predefined visual symbols such as the schematic representations of the real-world object or the phenomena (physical object), labels, arrows, lines or arcs. These predefined visual symbols are members of the category of visual physical models that is derived from the structural element category and given by the following categorical chain: $\ldots\langle\sigma_{\text{El}}\rangle \rhd \cdots\langle\nu_{\text{PhMod}}\rangle \rhd \langle\nu_{\text{SPhO}},\nu_{\text{Lab}},\nu_{\text{Arc}},\nu_{\text{Lin}},\nu_{\text{Arr}}\rangle$. A member of the pattern category of the different categories of the physical visual model is composed from different elements of categories of the physical visual model and is given as $\ldots\langle\sigma_{\text{Pt}}\rangle \rhd \cdots\langle\nu_{\text{PhMod}}\rangle \rhd \langle\nu_{\text{Lew}},\nu_{\text{Wed}},\nu_{\text{Rol}},\nu_{\text{Pen}},\nu_{\text{InP}}\rangle$. The category of the physical visual model is divided into the category of lever $\langle\nu_{\text{Lew}}\rangle$, the category of wedge $\langle\nu_{\text{Wed}}\rangle$, the category of rollers $\langle\nu_{\text{Rol}}\rangle$, the category of inclined planes $\langle\nu_{\text{InP}}\rangle$, and the category of pendulum $\langle\nu_{\text{Pen}}\rangle$.

The diagram category derived from the pattern category consists of predefined visual symbols such as geometrical figures or letters. The predefined visual symbols such as geometrical figures, letters, words, arrows, lines and arcs are members of the category of diagram elements derived from the element category and given as $\ldots\langle\sigma_{\text{El}}\rangle \rhd \cdots\langle\nu_{\text{EDia}}\rangle \rhd \langle\nu_{\text{GeoF}},\nu_{\text{Lab}},\nu_{\text{Arc}},\nu_{\text{Lin}},\nu_{\text{Arr}}\rangle$. The diagram category is divided into the phase diagrams, phylogenetic tree, Feynman diagram or Venn diagrams. Phase diagrams are used to illustrate the conditions under which certain minerals are stable. Feynman diagram is a graphic method of representing the interactions of elementary particles. Phylogenetic tree also called a dendrogram is a diagram showing the evolutionary interrelations of a group of organisms derived from a common ancestral form. Venn diagrams are representing logic and set theory as a purely symbolic calculus. The categorical chain of these categories is as $\ldots\langle\sigma_{\text{Pt}}\rangle \rhd \cdots\langle\nu_{\text{PDia}}\rangle \rhd \langle\nu_{\text{PhD}},\nu_{\text{PhT}},\nu_{\text{Ven}},\nu_{\text{Fen}},\nu_{\text{Arr}}\rangle$.

Map is a graphic representation, drawn to scale and usually on a flat surface, of features (for example, geographical, geological or geopolitical), of an area of the Earth or of any other celestial body. Cartography is the art and science of making maps and charts. Cartography is allied with geography in its concern with the broader aspects of the Earth and its life. The predefined visual symbols such as geometrical figures, labels, arrows, lines and arcs are members of the category of map elements derived from the structural element category and given as $\ldots\langle\sigma_{\text{El}}\rangle \rhd \cdots\langle\nu_{\text{EMap}}\rangle \rhd \langle\nu_{\text{GeoF}}, \nu_{\text{Lab}}, \nu_{\text{Arc}}, \nu_{\text{Lin}}, \nu_{\text{Arr}}\rangle$. The map category can be divided into geographical, geological or geopolitical map. The categorical chain is represented as $\ldots\langle\sigma_{\text{Pt}}\rangle \rhd \cdots\langle\nu_{\text{PMap}}\rangle \rhd \langle\nu_{\text{Geo}}, \nu_{\text{Glg}}, \nu_{\text{Gop}}\rangle$.

4.2.2.3. Picture Category

The picture category refers to the complex visual object that conveys visual information about the visual world. The complex visual object consists of parts that are distinguished as meaningful regions. Meaningful region, called visual object, is interpreted as one of the ontological categories. In comparison to the object from the pattern category, the complex visual objects of the picture category are composed without application of rules of picture composition. Picture provides the visual description of the different scales of the visual world. The visual objects such as planets or galaxies as well as very small objects such as viruses or bacteria can be extracted from the picture and interpreted as members of the ontological real-world categories. The picture category is most often the member of the perceptual shaded-object category π_{Sh}. The category of pictures that is derived from the perceptual category of shaded object π_{Sh} is divided into the micro-world category, the macro-world category and the world category and given as: $\nu_{\text{O}} \rhd \langle\pi_{\text{Sh}}\rangle \rhd \langle\sigma_{\text{Pi}}\rangle \rhd \langle\nu_{\text{ReO}}\rangle \rhd \langle\nu_{\text{Mic}}, \nu_{\text{Mac}}, \nu_{\text{Ear}}\rangle$. The category of pictures can be divided into categories specified by the perceptual and ontological level given by the categorical chain in the following form $\langle\Pi\rangle \rhd \langle\sigma_{\text{Pi}}\rangle \cdots \rhd \langle\nu_{\text{Mic}}, \nu_{\text{Mac}}, \nu_{\text{Ear}}\rangle$. This form of the categorical chain establishes the picture classification based on the ontological categories of the different scales of the visual world. Figure 4.8 shows pictures that are classified based on the ontological categories: the picture of the real-world category (Fig. 4.8a), the picture of the micro-world category (microscopic image) (Fig. 4.8b) and the picture of the macro-world category (picture of the astronomical objects) (Fig. 4.8c).



a b c d e

Fig. 4.8. Examples of the different categories of pictures: (**a**) The picture of the real-world category, (**b**) the picture the micro-world category, (**c**) the picture of the macro-world category, (**d**) the picture of the scientific visualization, (**e**) the picture of the fairy-tale category

There is a class of pictures that does not refer to the real-world object but is a product of the scientific visualization or imagination of the illustrators of the mythical stories. The category of the picture that is derived from the category of imagery objects $\langle \nu_{\mathrm{ImO}} \rangle$ is divided into category of scientific visualization objects $\langle \nu_{\mathrm{ScV}} \rangle$, the category of mythological objects $\langle \nu_{\mathrm{Mit}} \rangle$ or the fairy-tale category $\langle \nu_{\mathrm{Fai}} \rangle$ and is represented by the following categorical chain: $\nu_{\mathrm{O}} \triangleright \langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pi}} \rangle \triangleright \langle \nu_{\mathrm{ImO}} \rangle \triangleright \langle \nu_{\mathrm{ScV}}, \nu_{\mathrm{Mit}}, \nu_{\mathrm{Fai}} \rangle$. Figure 4.8d shows a picture that is a member of the scientific visualization category, and Fig. 4.8e shows a picture that is a member of the fairy-tale category.

There is a class of pictures that is distinguished based on their aesthetic quality. The class of pictures that is established based on the reference to their aesthetic quality is called a category of works of art. The work of art can be classified as a member of real-world picture categories such as the category of landscapes, the category of a set of non-living objects, the category of a set of living objects or the category of a set of non-living and living objects given by the categorical chain $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pi}} \rangle \cdots \triangleright \langle \nu_{\mathrm{DifO}} \rangle \triangleright \langle \nu_{\mathrm{Man}}, \nu_{\mathrm{ManL}}, \nu_{\mathrm{MNoL}} \rangle$. Figure 4.9 shows examples of works of art that are members of the landscape category.

The category of a set of non-living and living objects is divided into the category of man-made objects, the category of man-made and living objects or the category of man-made and non-living objects given by the

Fig. 4.9. Examples of works of art members of the landscape category

following categorical chain: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pi}} \rangle \cdots \triangleright \langle \nu_{\mathrm{DifO}} \rangle \triangleright \langle \nu_{\mathrm{Man}}, \nu_{\mathrm{ManL}}, \nu_{\mathrm{MNoL}} \rangle$. The category of a set of non-living and living objects refers to the term ‘still life’ that is used to denote paintings (pictures) that usually contain fruits, flowers or other objects setting on the table. An object that is the object of artistic work is represented by the conventional representation called a style. The style refers to the one of the perceptual categories and it will be described in the following chapters of this book. Figure 4.10 shows examples of works of art called ‘still life’ members of the category of man-made objects. Figure 4.11 shows examples of works of arts members of the category of man-made and living objects.

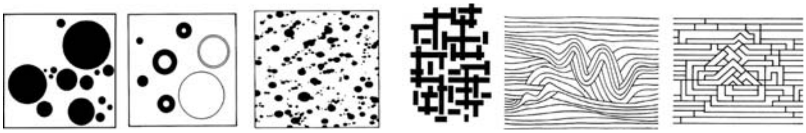
The picture does not need to refer to the real-world categories. The pictures that consist of geometrical figures are called abstract paintings. The category of pictures that are derived from the ontological figure category is divided into the category of picture division $\langle \nu_{\mathrm{PDiv}} \rangle$ or the category of figure on background $\langle \nu_{\mathrm{FigB}} \rangle$ and is represented by the following categorical chain: $\nu_{\mathrm{O}} \triangleright \langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pi}} \rangle \triangleright \langle \nu_{\mathrm{Fig}} \rangle \triangleright \langle \nu_{\mathrm{2DF}} \rangle \triangleright \langle \nu_{\mathrm{PDiv}}, \nu_{\mathrm{FigB}} \rangle$. Figure 4.12 shows examples of abstract paintings that are members of the category of picture division $\langle \nu_{\mathrm{PDiv}} \rangle$ (Fig. 4.12a–c) and the category of

Fig. 4.10. Still life – members of man-made category

Fig. 4.11. Examples of still life – members of the category of man-made and living objects

a b c d e

Fig. 4.12. Examples of abstract paintings



a b c d e f

Fig. 4.13. Examples of the abstract pictures members of the silhouette category (**a**–**d**) or the line-drawing category (**e**,**f**)

figure on the background $\langle \nu_{\mathrm{FigB}} \rangle$ (Fig. 4.12d,e). All pictures shown in Fig. 4.12 are derived from the perceptual shaded-object category $\langle \pi_{\mathrm{Sh}} \rangle$. Pictures shown in Fig. 4.13 are members of the silhouette category π_{Si} (Fig. 4.13a–d) or the line-drawing category π_{Ld} (Fig. 4.13e–f).

Understanding of the visual object is focused on understanding of the object that is extracted from the image. Extraction of the visual object from the picture can be obtained by applying existing segmentation methods. These methods need to be further elaborated to be applicable to the specific picture categories. By applying the simple segmentation methods, the picture is divided into two regions: the figure and the background. The result of segmentation depends on the type of background Fig. 4.14a–c. The simplest case is an object on the background of the uniform colour (see Fig. 4.14c). In this book the main focus is on the problems connected with visual

a b c

Fig. 4.14. Different pictures from which phantom can be extracted

thinking. Thus the perceived object needs to be fitted to one of the shape categories that are the main ingredient of the visual thinking process. In this book it is assumed that a visual object is extracted from the image by applying one of the existing segmentation methods.

4.2.2.4. Category of Animation

The category of animation refers to the category of process that is one of the ontological categories. Animation is often defined as the process of giving the illusion of movement or life to cinematographic drawings, models or inanimate objects. The category of process refers to the changes of the visual object that can be observed during the period of time. The member of the category of animation is the sequence of pictures, patterns or elements that are interpreted as a series of time dependent images. The category of processes does not depict the object itself but rather the changes of the object. The category of processes is described in more details in the following sections. Figure 4.15 shows examples of the objects from the animation category.

4.2.3. Ontological Categories

The third-categorical level is called the ontological level and refers to the meaning of the object. The ontological level includes ontological categories $\langle \nu \rangle$ such as the category of real-world objects $\langle \nu_{\mathrm{ReO}} \rangle$, the category of imagery objects $\langle \nu_{\mathrm{ImO}} \rangle$, the category of letters $\langle \nu_{\mathrm{Let}} \rangle$, the category of signs $\langle \nu_{\mathrm{Sig}} \rangle$ and the category of figures $\langle \nu_{\mathrm{Fig}} \rangle$. The derivation chain is given as $\nu_{\mathrm{O}} \triangleright \langle \Pi \rangle \triangleright \langle \Lambda \rangle \triangleright \langle \nu_{\mathrm{ReO}}, \nu_{\mathrm{ImO}}, \nu_{\mathrm{Sig}}, \nu_{\mathrm{Let}}, \nu_{\mathrm{Fig}} \rangle$. In the case when both perceptual and structural categories are not specified the categorical chain is given as $O \triangleright \langle \nu_{\mathrm{ReO}}, \nu_{\mathrm{ImO}}, \nu_{\mathrm{Sig}}, \nu_{\mathrm{Let}}, \nu_{\mathrm{Fig}} \rangle$, where the letter O denotes the category of the visual object.

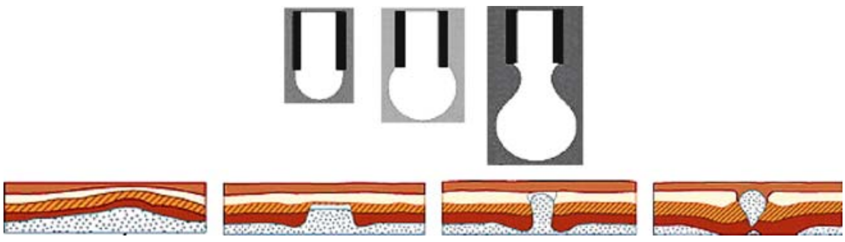


Fig. 4.15. Example of members of the animation category

The symbol O at the beginning of the categorical chain means that any combination of the specific categories from the perceptual level and the structural level can be used. The real-world category is divided into the micro-world category, the macro-world category and the world category. The categorical chain is given as $O \triangleright \langle \nu_{\text{ReO}} \rangle \triangleright \langle \nu_{\text{Mic}}, \nu_{\text{Mac}}, \nu_{\text{Ear}} \rangle$. The ontological categories will be described in detail in the following sections.

4.2.3.1. Interpretation of the Visual Object

The perceptual and structural categories give information about the way in which the visual object is perceived. The ontological categories refer to the meaning of the visual object and can be seen as a new form of the knowledge representation. Knowledge represented in the form of the categorical chains is used in the process of interpretation of the perceived object during thinking process. Information obtained during understanding process connected with interpretation of the perceived object can be passed to another system during the conversation session. Interpretation of the perceived object involves application of non-visual knowledge that is part of symbolic description of the world to find the meaning of the visual object. Non-visual knowledge is represented by the category of knowledge object that is derived from the category of body of knowledge. The categorical chain of the knowledge categories is related to the categorical chain of the visual object. The knowledge that is represented by the categorical chain of knowledge is acquired from the existing sources of knowledge such as books or scientific journals. The process of knowledge acquisition is an iterative process in which new learned facts are linked into learned knowledge of the categorical chains. At the first stage of the learning process, there is no need for inclusion of all visual object categories and knowledge categories or even for their specific representation. The specific categories can be added during further stages of the learning process and knowledge of each category can be more precisely defined. The learned categorical chains, although do not have all specific knowledge needed for proper interpretation of all objects, they make it possible to find the general category of objects that describes the object in understandable way. That means the object will be interpreted in the context of all learned knowledge. For example, the unknown object can be interpreted as a rose, a flower or a plant. In all cases the interpretation is understandable and gives the description that can be used during further investigation of the properties of the perceived object.

Knowledge can be regarded at many different levels and the different aspect of knowledge can be differently described and defined. The category of body of knowledge refers to any kind of knowledge that can be represented in any form that is transferable into the conceptual form of the categorical chain. In general, knowledge can be given by scientific theories, scientific descriptions, scientific models or common knowledge. Knowledge categories are denoted by the letter κ. The category of knowledge body is divided into the category of scientific discipline (the knowledge object) $\langle \kappa_{\mathrm{KOb}} \rangle$, the category of scientific theories $\langle \kappa_{\mathrm{KTh}} \rangle$, the category of scientific description $\langle \kappa_{\mathrm{KDe}} \rangle$, the category of scientific model $\langle \kappa_{\mathrm{KMo}} \rangle$ or the category of common sense knowledge $\langle \kappa_{\mathrm{KsK}} \rangle$. The categorical chain that represents the first level of knowledge categories is given as: $\kappa_{\mathrm{KB}} \rhd \langle \kappa_{\mathrm{KOb}}, \kappa_{\mathrm{KMo}}, \kappa_{\mathrm{KTh}}, \kappa_{\mathrm{KDe}}, \kappa_{\mathrm{KsK}} \rangle$. For simplicity, the categories derived from the visual object will be called the visual categories whereas categories derived from the knowledge object will be called the knowledge categories.

Science is any system of knowledge that is concerned with the physical world and other phenomena and that has its own research methodology to pursuit new knowledge. In general, science involves pursuit of knowledge covering general truths or the operations of fundamental laws. Scientific knowledge is divided into different scientific disciplines such as the physical sciences, the earth sciences, the biological sciences, the medicine, the engineering and the social sciences. These disciplines are divided into sub-disciplines that are focused on acquiring specific knowledge. For example, physical science is the systematic study of the inorganic world. Physical science consists of four broad areas such as astronomy, physics, chemistry and the Earth sciences. The category of scientific discipline (the knowledge object) $\langle \kappa_{\mathrm{KOb}} \rangle$ is the most important in supplying knowledge needed in interpretation of the visual objects. Other categories supply knowledge, which is needed to obtain the deep understanding of the real-world phenomenon that does not always have the visual representation. The category of scientific discipline (the knowledge object) $\langle \kappa_{\mathrm{KOb}} \rangle$ is divided into the category of physical sciences, the biological sciences, the medicine, the engineering or the social sciences given by categorical chain: $\kappa_{\mathrm{KB}} \rhd \langle \kappa_{\mathrm{KOb}} \rangle \rhd \langle \kappa_{\mathrm{PhS}}, \kappa_{\mathrm{BiS}}, \kappa_{\mathrm{MeS}}, \kappa_{\mathrm{EnS}}, \kappa_{\mathrm{SoS}} \rangle$.

Knowledge of a given category of science, e.g. mathematics is given by the category of mathematical objects that refers to the category of visual objects such as mathematical figures, mathematical symbols or mathematical expressions. Mathematics is the science of structure, order, and relation that has evolved from elemental practices of counting, measuring and describing shapes of objects. It deals with logical reasoning and quantitative calculation, and its development has involved an increasing degree of idealization and abstraction of its subject matter.

The category of mathematical objects such as mathematical curves ν_{Cur} is linked with the category of the visual objects. The categorical chain of visual objects supplies knowledge concerning the general properties of the visual object. Based on this chain we can infer that specific category such as an ellipsis is a member of the convex-closed curve category and a member of the more general category such as the category of the geometrical figures. The categorical chain that represents the visual knowledge of the ellipse category is given as $O \rhd \langle \nu_{\mathrm{Fig}} \rangle \rhd \langle \nu_{\mathrm{2DF}} \rangle \rhd \langle \nu_{\mathrm{Cur}} \rangle \rhd \langle \nu_{\mathrm{ConC}} \rangle \rhd \langle \nu_{\mathrm{NamC}} \rangle \rhd \langle \nu_{\mathrm{Cir}}, \nu_{\mathrm{Eli}} \rangle$.

The categorical chain of the knowledge objects (knowledge chain) supplies knowledge concerning non-visible aspect of the visual objects. The non-visible properties such as temperature or weight can give the additional information that makes it possible to identify an object. For example, the categorical chain of knowledge objects such as the ellipse category is given as $\kappa_{\mathrm{KB}} \rhd \langle \kappa_{\mathrm{KOb}} \rangle \rhd \langle \kappa_{\mathrm{Mat}} \rangle \rhd \langle \kappa_{\mathrm{Fig}} \rangle \rhd \langle \kappa_{\mathrm{2DF}} \rangle \rhd \langle \kappa_{\mathrm{Cur}} \rangle \rhd \langle \kappa_{\mathrm{ConC}} \rangle \rhd \langle \kappa_{\mathrm{NamC}} \rangle \rhd \langle \kappa_{\mathrm{Cir}}, \kappa_{\mathrm{Eli}} \rangle$.

The lower category of the categorical chain is the prototype. Each prototype has its knowledge schema that is inherited through the hierarchical categorical structure of the chain. The knowledge schema includes the characteristic categorical features and definitions. The knowledge schema is represented by symbols in brackets $\prec \{\partial_{\mathrm{a}}, \partial_{\mathrm{b}}, \partial_{\mathrm{c}}, \partial_{\mathrm{d}}\}$. The symbol ∂_{c} denotes the feature c that is characteristic for a given category (prototype). For example, the knowledge schema of the ellipse prototype consists of the visual concept ∂_{ViC}, the name ∂_{Nam}, the mathematical formula ∂_{MaF}, the definition ∂_{Def} and the method of generation of figure ∂_{MGe}. The knowledge schema is given as the part of the categorical chain: $\ldots \langle \kappa_{\mathrm{NamC}} \rangle \rhd \langle \kappa_{\mathrm{Eli}} \rangle \prec \{\partial_{\mathrm{ViC}}, \partial_{\mathrm{Nam}}, \partial_{\mathrm{MaF}}, \partial_{\mathrm{Def}}, \partial_{\mathrm{MGe}}\}$. The knowledge schema of the ellipse category is inherited from the category of the closed curves and the category of mathematical curves. The knowledge schema of the general category that includes the more general categorical description is passed into the specific category that includes the specific categorical description. The knowledge schema of the specific category can be different from the knowledge schema of the general category. The specific category can have additional features that are responsible for supplying the specific knowledge about the specific category.

For simplicity, the categories derived from the visual object will be called the visual categories whereas categories derived from the knowledge object will be called the knowledge categories.

4.2.3.2. Dependence Among Ontological Categories

The categorical chain can be linked and form a very complex structure of knowledge that is used to interpret the visual object. The linked chains are used for the categorical reasoning that is based on the 'moving' through the linked chain categories. To illustrate the dependence of the different ontological categories of the linked chains the musical categories are given as an example of the complex structure of the relations among categories. Visual categories such as the category of musical notation are linked together to provide the interpretational structure of the world. Figure 4.16 shows examples of members of the different visual musical categories.

Musical notation is a visual record of heard or imagined musical sound, or a set of visual instructions for performance of music. It usually takes written or printed form. Musical notation serves as a means of preserving music over long periods of time, facilitates performance by others, and presents music in a form suitable for study and analysis. The categories of musical notation include the categories of musical symbols. The categories of musical notation supply knowledge that makes it possible to understand the musical visual symbols, to play the musical composition and to 'record' the composed musical work by writing it in the form of the musical scores. The category of the musical elements, such as the note $\langle \nu_{\mathrm{Not}} \rangle$, the rest $\langle \nu_{\mathrm{Res}} \rangle$ or the clef $\langle \nu_{\mathrm{Cle}} \rangle$, supplies knowledge that makes it possible to recognize and name musical symbols. The categorical chain for the category of musical symbols is given as $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{Mus}} \rangle \triangleright \langle \nu_{\mathrm{Not}}, \nu_{\mathrm{Res}}, \nu_{\mathrm{Cle}} \rangle$. The lower level of the category of musical symbols is the level of specific category of musical symbols such as the category of the bass clef, the

Fig. 4.16. Example of members of the different visual musical categories

treble clef or the C clef $\ldots \triangleright \langle \nu_{\text{Mus}} \rangle \triangleright \langle \nu_{\text{Cle}} \rangle \triangleright \langle \nu_{\text{CTre}}, \nu_{\text{CBas}}, \nu_{\text{C\_C}} \rangle$. The knowledge of the specific category of musical elements makes it possible to interpret correctly musical symbols.

Visual categories of musical symbols are linked with knowledge categories of musical symbols. The knowledge category of musical symbols supplies knowledge that makes it possible to interpret musical visual symbols as a specific musical sound $\kappa_{\text{KB}} \triangleright \langle \kappa_{\text{KOb}} \rangle \triangleright \langle \kappa_{\text{Mus}} \rangle \triangleright \langle \kappa_{\text{Not}}, \kappa_{\text{Res}}, \kappa_{\text{Cle}} \rangle$. Understanding musical symbols means knowing how to make the sound using the musical instrument. The musical element such as Q note refers to the elements of musical sound such as Q pitch or the location of musical sound on the musical scale. As it was described in the previous section each prototype has its knowledge schema that is inherited through the hierarchical categorical structure. The knowledge schema includes characteristic categorical features. The knowledge schema for the prototype of the treble clef includes the visual concept ∂_{ViC}, the name ∂_{Nam}, the musical interpretation ∂_{MIn}, the definition ∂_{Def} and the method of the figure generation ∂_{MGe}, and is given as $\ldots \triangleright \langle \kappa_{\text{Mus}} \rangle \triangleright \langle \kappa_{\text{Cle}} \rangle \triangleright \langle \kappa_{\text{CTre}} \rangle \prec \{\partial_{\text{ViC}}, \partial_{\text{Nam}}, \partial_{\text{MIn}}, \partial_{\text{Def}}, \partial_{\text{MGe}}\}$.

The visual symbols used as the means of musical notation create a set of elements that are used to form more complex musical expression. These complex expressions are formed according to the rules of the musical composition. For example, musical symbols placed on the staff are interpreted in terms of the musical notation and can be used to produce the musical sound. The musical symbols placed on the staff are members of the musical categories such as beaming category, the phrase category, the rhythm category, the harmony category or the melody category (sequence of musical symbols on the staff) derived from the structural pattern categories given by the categorical chain: $\langle \Pi \rangle \triangleright \langle \sigma_{\text{Pt}} \rangle \triangleright \langle \nu_{\text{Sg}} \rangle \triangleright \langle \nu_{\text{VSym}} \rangle \triangleright \langle \nu_{\text{Mus}} \rangle \triangleright \langle \nu_{\text{Bea}}, \nu_{\text{Phr}}, \nu_{\text{MLi}} \rangle$. The category of pattern of musical elements refers to the interpretation of the sequence of musical symbols in terms of a melody or a rhythm. Sequence of musical symbols placed on the staff can be transformed into musical sound by a musician playing on one of the musical instruments. The knowledge chain supplies knowledge that is needed during interpretation of the sequence of musical symbols: $\kappa_{\text{KB}} \triangleright \langle \kappa_{\text{KOb}} \rangle \triangleright \langle \kappa_{\text{VMus}} \rangle \triangleright \langle \nu_{\text{Bea}}, \nu_{\text{Phr}}, \nu_{\text{MLi}} \rangle$, where symbol $\langle \kappa_{\text{VMus}} \rangle$ denotes the category of visual musical knowledge. Non-visual aspect of the musical knowledge is represented by knowledge chain. The category of non-visual musical knowledge includes the category of rules of composition of the musical works, the category of style or the category of musical theories: $\kappa_{\mathrm{KB}} \rhd \langle \kappa_{\mathrm{KOb}} \rangle \rhd \langle \kappa_{\mathrm{NVMus}} \rangle \rhd \langle \nu_{\mathrm{MRul}}, \nu_{\mathrm{MSty}}, \nu_{\mathrm{MThe}} \rangle$.

Music is played by musicians who transform the music scores into musical sound. The category of musicians is derived from the category of professionals $\langle \nu_{\mathrm{Prf}} \rangle$ and is given by the following categorical chain: $\langle \Pi \rangle \rhd \langle \sigma_{\mathrm{El}} \rangle \rhd \langle \nu_{\mathrm{ReO}} \rangle \rhd \langle \nu_{\mathrm{Ear}} \rangle \rhd \langle \nu_{\mathrm{Liv}} \rangle \rhd \langle \nu_{\mathrm{Man}} \rangle \rhd \langle \nu_{\mathrm{Prf}} \rangle \rhd \langle \nu_{\mathrm{MCom}}, \nu_{\mathrm{MMus}} \rangle$. There are two main categories of professionals that make music, the category of musicians ν_{MMus} and the category of composers ν_{MCom}. The musical works are composed by composers and can be performed by musician or a group of musicians (trio, quartet and orchestra). Orchestra is instrumental ensemble of varying size and composition. The different musicians use the different instruments. The knowledge chain that supplies knowledge that is needed during interpretation of the visual object (from the category of professionals that make music) is given as $\kappa_{\mathrm{KB}} \rhd \langle \kappa_{\mathrm{KOb}} \rangle \rhd \langle \kappa_{\mathrm{Man}} \rangle \rhd \langle \kappa_{\mathrm{Work}} \rangle \rhd \langle \kappa_{\mathrm{Prf}} \rangle \rhd \langle \kappa_{\mathrm{MCom}}, \kappa_{\mathrm{MMus}} \rangle$. The category of musicians is divided into different categories of musicians such as the violinist category κ_{MVio}, the drummer category κ_{MDr} or the trumpeter category κ_{MTr} and is represented as $\ldots \rhd \langle \kappa_{\mathrm{Prf}} \rangle \rhd \langle \kappa_{\mathrm{MMus}} \rangle \rhd \langle \kappa_{\mathrm{MTr}}, \kappa_{\mathrm{MDr}}, \kappa_{\mathrm{MVio}} \rangle$. The knowledge schema derived from the category of professionals include the category of worker $\kappa^{\mathrm{P}}_{\mathrm{Man}}$, the category of tools $\kappa^{\mathrm{P}}_{\mathrm{Tol}}$, the category of materials $\kappa^{\mathrm{P}}_{\mathrm{Mat}}$, the category of knowledge $\kappa^{\mathrm{P}}_{\mathrm{Kno}}$, and the category of results $\kappa^{\mathrm{P}}_{\mathrm{Res}}$ and is given in the following form $\propto \left(\nu^{\mathrm{P}}_{\mathrm{Man}}, \nu^{\mathrm{P}}_{\mathrm{Topl}}, \nu^{\mathrm{P}}_{\mathrm{Mat}}, \nu^{\mathrm{P}}_{\mathrm{Kno}}, \nu^{\mathrm{P}}_{\mathrm{Res}} \right)$. The category of professionals such as violinists is given by the following knowledge schema$\ldots \rhd \langle \kappa_{\mathrm{Prf}} \rangle \rhd \langle \kappa_{\mathrm{MMus}} \rangle \rhd \langle \kappa_{\mathrm{MVio}} \rangle \propto \left(\kappa^{\mathrm{P}}_{\mathrm{MVi}}, \kappa^{\mathrm{P}}_{\mathrm{Vio}}, \kappa^{\mathrm{P}}_{-}, \kappa^{\mathrm{P}}_{\mathrm{ViK}}, \kappa^{\mathrm{P}}_{\mathrm{ViW}} \right)$. This knowledge schema supplies knowledge for categories of musicians such as the category of violinists and consists of the following categories: the category of violinists $\kappa^{\mathrm{P}}_{\mathrm{MVi}}$, the category of violins (the category of musical instruments) $\kappa^{\mathrm{P}}_{\mathrm{Vio}}$, the category of knowledge that violinists need to perform the musical work $\kappa^{\mathrm{P}}_{\mathrm{ViK}}$, and the category of musical works that is the result of the violinist performance $\kappa^{\mathrm{P}}_{\mathrm{ViW}}$. The musicians use instruments to perform the musical work. The different musical instruments have to be used to play the different parts of the musical work. Musical instrument is any device for producing a musical sound. The principal types of such instruments, classified by the method of producing sound, are percussion, stringed, keyboard, wind and electronic. Based on this classification of the musical instruments the category of the musical instrument is divided into the percussion category ν_{MIP}, the stringed category ν_{MISt}, the keyboard category ν_{MIK}, the wind category ν_{MIW}, and the electronic category ν_{MIE} and is given by the following categorical chain: $\langle \Pi \rangle \rhd \langle \sigma_{\text{El}} \rangle \rhd \langle \nu_{\text{ReO}} \rangle \rhd \langle \nu_{\text{Ear}} \rangle \rhd \langle \nu_{\text{NLiv}} \rangle \rhd \langle \nu_{\text{MMad}} \rangle \rhd \langle \nu_{\text{MIn}} \rangle \rhd \langle \nu_{\text{MISt}}, \nu_{\text{MIW}}, \nu_{\text{MIP}}, \nu_{\text{MIK}}, \nu_{\text{MIE}} \rangle$. For example, stringed instrument is any musical instrument that produces sound by the vibration of strings, which may be made of vegetable fibre, metal, animal gut or plastic. In nearly all stringed instruments the sound of the vibrating string is amplified by the use of a resonating chamber or soundboard. The specific categories of stringed musical instrument are: viola category ν_{Vio}, the cello category ν_{Cel}, the lute category ν_{Lut}, the balalaika category ν_{Bal} or the guitar category ν_{Git} and are represented as $\ldots \rhd \langle \nu_{\text{MIn}} \rangle \rhd \langle \nu_{\text{MISt}} \rangle \rhd \langle \nu_{\text{Vio}}, \nu_{\text{Cel}}, \nu_{\text{Lut}}, \nu_{\text{Git}}, \nu_{\text{Bal}} \rangle$. Category derived from the knowledge object supplies knowledge about the way of producing sound by a given instrument and is given by the following categorical chain: $\kappa_{\text{KB}} \rhd \langle \kappa_{\text{KOb}} \rangle \rhd \langle \kappa_{\text{MIn}} \rangle \rhd \langle \kappa_{\text{MISt}} \rangle \rhd \langle \kappa_{\text{Vio}}, \kappa_{\text{Cel}}, \kappa_{\text{Git}} \rangle$. The musical instrument is produced by instrument makers. Instrument makers specialize in production of the specific instruments such as a viola or a cello. The visual category of instrument makers ν_{MMIM} is derived from the category of professionals ν_{Prf}. The visual chain for these categories is given as $\ldots \rhd \langle \nu_{\text{Prf}} \rangle \rhd \langle \nu_{\text{MMIM}} \rangle \rhd \langle \nu_{\text{IMVio}}, \nu_{\text{IMGit}} \rangle$. Similarly like in the previous example the visual chain is linked with the knowledge chain.

The music that is performed in the special building can be recorded and stored on the magnetic tape or CD. Music recording is physical record of a musical performance that can then be played back, or reproduced. Sound recording is transcription of vibrations in air that are perceptible as sound onto a storage medium, such as a phonograph disc. In sound reproduction the process is reversed so that the sound stored in the medium is converted back into sound waves. The three principal media that have been developed for sound recording and reproduction are the mechanical (phonographic disc), magnetic (audiotape) and optical (digital compact disc) systems. The recorded music can be played by using electronic devices such as gramophone or CD-player. The category of the electronic devices is derived from the pattern object to indicate that the devices are assembled from the simple elements. The electronic devices that are used for the purpose of music playing are members of the electronic sound device categories such as gramophone, a radio, a TV-set, a tape recorder or a CD-player. The electronic sound device is assembled from the electronic elements such as transistor. The category of the electronic sound device is represented by the following visual chain $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \ldots \triangleright \langle \nu_{\mathrm{MMad}} \rangle \triangleright \langle \nu_{\mathrm{AsP}} \rangle \triangleright \langle \nu_{\mathrm{EAsP}} \rangle \triangleright \langle \nu_{\mathrm{Rez}}, \nu_{\mathrm{Ind}}, \nu_{\mathrm{Tfo}}, \nu_{\mathrm{Cap}}, \nu_{\mathrm{Dio}}, \nu_{\mathrm{Tran}} \rangle$. This chain was described in previous section in the context of the description of assembling of electronic devices based on the electronic schema. Members of the category of electronic sound devices such as a gramophone or a CD-player are assembled based on the schema of electronic circuit. The category of the electronic sound devices is derived from the category of the electronic devices ν_{ElDev} and is divided into the gramophone category ν_{Gra}, the radio category ν_{Rad}, the TV-set category ν_{TV}, the tape recorder category ν_{Mag} or the CD-player category ν_{CD}. The categorical chain of categories of the electronic sound devices is given in the following form: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pt}} \rangle \cdots \triangleright \langle \nu_{\mathrm{Dev}} \rangle \triangleright \langle \nu_{\mathrm{ElDev}} \rangle \triangleright \langle \nu_{\mathrm{EDSo}} \rangle \triangleright \langle \nu_{\mathrm{Rad}}, \nu_{\mathrm{TV}}, \nu_{\mathrm{Gra}}, \nu_{\mathrm{Mag}}, \nu_{\mathrm{CD}} \rangle$.

The music is performed in the special places such as concert halls, opera houses, musical schools that are members of the musical house category. The musical house category derived from the house category is divided into the concert hall category ν_{HCon}, the opera house category ν_{HOper} or the musical school category ν_{HSco} and is given by the following chain: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \cdots \triangleright \langle \nu_{\mathrm{MMad}} \rangle \triangleright \langle \nu_{\mathrm{Hou}} \rangle \triangleright \langle \nu_{\mathrm{HMus}} \rangle \triangleright \langle \nu_{\mathrm{HOper}}, \nu_{\mathrm{HCon}}, \nu_{\mathrm{HSco}} \rangle$. Opera houses are built by building workers and designed by architects that are members of the building worker category ν_{Bul} or the architect category ν_{Arh} represented by the following chain: $\ldots \triangleright \langle \nu_{\mathrm{Man}} \rangle \triangleright \langle \nu_{\mathrm{Prf}} \rangle \triangleright \langle \nu_{\mathrm{Arh}}, \nu_{\mathrm{Bul}} \rangle$.

Music is played by a musician who uses the musical instrument to produce the music which is a special kind of the sound wave. The musical sound is characterized by physical properties of the sound wave. Sound results from the vibration of elastic bodies such as a violin string or a human vocal chord and is the subject of research the branch of science called acoustic. Acoustic is the science concerned with the production, control, transmission, reception and effects of sound. The sound can be visualized and considered as the sub-category of the category of the visual processes. The visualization of the sound wave can be obtained by applying the different transformations such as Fourier or Wavelets. The category of musical sound ν_{Mus} is derived from the category of the acoustic processes ν_{Acus} given by the following visual chain: $O \triangleright \langle \nu_{\mathrm{ReO}} \rangle \triangleright \langle \nu_{\mathrm{Ear}} \rangle \triangleright \langle \nu_{\mathrm{NLiv}} \rangle \triangleright \langle \nu_{\mathrm{NatP}} \rangle \triangleright \langle \nu_{\mathrm{Acus}} \rangle \triangleright \langle \nu_{\mathrm{Mus}}, \nu_{\mathrm{Son}}, \nu_{\mathrm{Spi}}, \nu_{\mathrm{Noi}} \rangle$. The categories of different processes will be described in the further sections of this chapter.

The knowledge schema of the sound category is given as $\ldots \langle \kappa_{\mathrm{Acus}} \rangle \triangleright \langle \kappa_{\mathrm{Mus}} \rangle \prec \{\partial_{\mathrm{ViC}}, \partial_{\mathrm{Nam}}, \partial_{\mathrm{Def}}, \partial_{\mathrm{Fet}}, \partial_{\mathrm{MGe}}\}$. The definition has two parts, the first one that defines sound (that is…) and the second one that defines sound (as consist of…). The definition is related to other categories that are defined or will be defined within the scope of the system. The definition is given in the term of the different categories. For example, 'the result of vibration of elastic body' $\langle \nu_{\mathrm{Son}} \rangle \prec \partial_{\mathrm{is}} \langle \mathrm{result} - \nu_{\mathrm{vibrat}} - \nu_{\mathrm{ElBod}} \rangle$, 'air_preasure_wave transmitted through air' $\langle \nu_{\mathrm{Son}} \rangle \prec \partial_{\mathrm{is}} \langle \nu_{\mathrm{APWav}} - \mathrm{transmited\_through} - \nu_{\mathrm{Air}} \rangle$. Features such as amplitude F_{Amp}, frequency F_{Fre} or power spectrum F_{Spe} are given as the part of the knowledge shema: $\langle \nu_{\mathrm{Son}} \rangle \prec \partial_{\mathrm{Fet}} \langle F_{\mathrm{Amp}}, F_{\mathrm{Fre}}, F_{\mathrm{Spe}} \rangle$.

The relations among categories are represented by the dependence diagram. The dependence diagram keeps the links to all visual categorical chains that are related to each other. Figure 4.17 shows the dependence diagrams of categorical chains that are related to the music category. Based on the dependence diagram the visual objects such as the violin (the category of the musical instrument) can be interpreted in the context of the learned knowledge represented by linked categorical chains. The dependence diagram makes it possible to infer that musical instrument is used by musician to play the music that is composed by composer and that is given in the form of the musical scores. From the categorical chain of the category of the musical instrument we can have access to knowledge about the specific instruments such as a violin or a guitar. The dependence diagram makes it possible to establish any connection with all categorical chains of the dependence diagram and by this we can have access to knowledge supplied by both visual and knowledge chains. Each category of the dependence diagram can have link to other dependence diagrams that have knowledge of the different aspects of the visual world. For example, the musical symbols that are part of the dependence diagram of the musical categories can give link to the mathematical symbols that are part of the dependence diagram of the mathematical categories. Inference that is based on dependence diagrams is part of the thinking process that can offer nearly infinite possibilities of the creative exploration of the different categorical links.

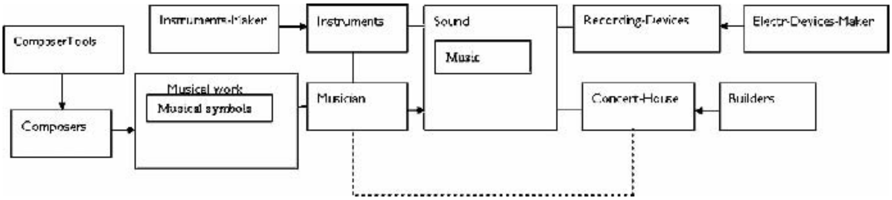


Fig. 4.17. The dependence diagram of the musical categories

As it was described, ontological categories are members of the third-categorical level. The ontological level includes ontological categories such as the category of the real-world objects $\langle \nu_{\text{ReO}} \rangle$, the category of imagery objects $\langle \nu_{\text{ImO}} \rangle$, the category of letters $\langle \nu_{\text{Let}} \rangle$, the category of signs $\langle \nu_{\text{Sig}} \rangle$ and the category of figures $\langle \nu_{\text{Fig}} \rangle$ and are given by the following categorical chain: $\nu_{\text{O}} \triangleright \langle \Pi \rangle \triangleright \langle \Lambda \rangle \triangleright \langle \nu_{\text{ReO}}, \nu_{\text{ImO}}, \nu_{\text{Sig}}, \nu_{\text{Let}}, \nu_{\text{Fig}} \rangle$. The category of figures is described in the next section.

4.2.3.3. Figure Category

The figure category is defined based on geometrical and perceptual properties of the visual object. This definition is related to the shape classes and is given as a class description, e.g. ‘the concave figure’, ‘the complex figure’, ‘the thin figure’, ‘the concave figure with one hole’. The figure category is part of the hierarchical knowledge of ontological categories. At the bottom of each categorical chain is the prototype of the category. The prototype of the category refers to the specific meaning of the visual object. The category of figures is derived from the category of visual objects. The category of figures consists of the category of 2D figures $\langle \nu_{\text{2DF}} \rangle$, the category of 3D figures $\langle \nu_{\text{3DF}} \rangle$, the category *n*-D figures $\langle \nu_{\text{M3DF}} \rangle$ (more than 3D figures). The category of 3D figures refers to geometrical objects that ‘exist’ in the 3D space. The category of *n*-D figures is the category members of which are objects that can be found in more than 3D space. These objects can be visible by projection from *n*-D into the 3D or 2D space. The categorical chain for the specific categories derived from the figure category is given as $O \triangleright \langle \nu_{\text{Fig}} \rangle \triangleright \langle \nu_{\text{2DF}}, \nu_{\text{3DF}}, \nu_{\text{M3DF}} \rangle$. The beginning of the categorical chain is marked by letter O to show that any combination of both perceptual and structural categories can be selected.

The category of 2D figures consists of the category of polygons, the category of curve polygons and the category of curves, and is given by the following categorical chain: $O \triangleright \langle \nu_{\mathrm{Fig}} \rangle \triangleright \langle \nu_{\mathrm{2DF}} \rangle \triangleright \langle \nu_{\mathrm{Pol}}, \nu_{\mathrm{CuPo}}, \nu_{\mathrm{Cur}} \rangle$. In the next section, these specific categories will be shortly described.

4.2.3.3.1. Polygon Category

Polygon is the figure that is made of the straight lines. The category of polygons is divided into the category of named polygons and the category of non-named polygons and is represented by the following categorical chain: $\ldots \langle \nu_{\mathrm{Pol}} \rangle \triangleright \langle \nu_{\mathrm{NaP}}, \nu_{\mathrm{NNaP}} \rangle$. The category of the named polygons is the category that is connected with the category of mathematical figures derived from the mathematical objects. These figures are well defined and their properties are well described. Knowledge of these figures is a part of geometrical knowledge. The category of non-named polygons is divided into the category of open convex polygons ν_{OpCoP}, the category of open concave polygons ν_{OpCcP}, the category of closed convex polygons ν_{ClCoP}, and the category of closed concave polygons ν_{ClCoP} and is given by the following categorical chain: $\ldots \langle \nu_{\mathrm{Pol}} \rangle \triangleright \langle \nu_{\mathrm{NNaP}} \rangle \triangleright \langle \nu_{\mathrm{OpCoP}}, \nu_{\mathrm{OpCcP}}, \nu_{\mathrm{ClCoP}}, \nu_{\mathrm{ClCcP}} \rangle$. The name of the non-named polygon is given by the symbolic name of the shape class from which the polygon is generated. Properties of the non-named polygon are described in the context of the shape classes. The shape classes were described in Chap. 2. The non-named figures can resemble some letters or even real-world objects. In that case the visual object has more than one interpretation and can be named by the name of that real-world object to which it is similar. Examples of members of the category of non-named open polygons are shown in Fig. 4.18.

The category of the open polygons refers to the thin shape class. The knowledge needed to interpret a perceived object is given by the knowledge schema. The knowledge schema for the convex close polygons category includes the visual concept ∂_{ViC}, the name ∂_{Nam}, the definition ∂_{Def} and the method of exemplar generation ∂_{MGe} and is given as

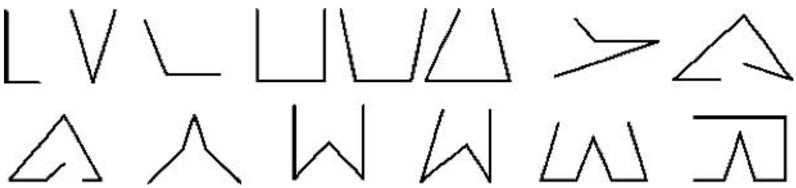


Fig. 4.18. Examples of members of the category of non-named open polygons

$\ldots\langle\kappa_{\text{Pol}}\rangle \rhd \langle\kappa_{\text{NNaP}}\rangle \rhd \langle\kappa_{\text{ClCoP}}\rangle \prec \{\partial_{\text{ViC}}, \partial_{\text{Nam}}, \partial_{\text{MIn}}, \partial_{\text{Def}}, \partial_{\text{MGe}}\}$. The name ∂_{Nam} is given in the form of linguistic expression of the existing languages and is expressed as $\ldots\langle\kappa_{\text{Pol}}\rangle \rhd \langle\kappa_{\text{NNaP}}\rangle \rhd \langle\kappa_{\text{ClCoP}}\rangle \prec \{\partial_{\text{Nam}}\} \parallel \text{NLan} \parallel \langle\nu_{\text{C1}}, \nu_{\text{C2}}, \ldots, \nu_{\text{CN}}\rangle$, where $\langle\nu_{C1}, \nu_{C2}, \ldots, \nu_{\text{CN}}\rangle$ denotes the name categories that depend on the selected language category $\parallel \nu_{\text{NLan}} \parallel$.

The figure category is related to the hierarchical structure of the shape classes. For example, the category of convex polygons refers to the convex polygon class denoted as L^n. From the category of the convex polygon the category of the named convex polygon is derived. From the category of the named convex polygon the category of the specific named polygon is derived. Each prototype of this category has its characteristic name that refers to category of mathematical objects.

The name of the object that is a member of the category of non-named objects ∂_{Nam} is the symbolic name of the shape class or a linguistic description. The category of concave polygons has its name expressed in the form of the linguistic description, e.g. ‘a polygon with one triangular concavity’, or ‘a rectangle with two rectangular concavities’. The name of this category is given by both the symbolic name that refers to the shape class and the linguistic description. For example, the name is given as the symbolic name Q1_L3_L3 and the linguistic description ‘triangle with one triangular concavity’; the symbolic name Q1_L3_L4 and the linguistic description ‘triangle with one rectangular concavity’; the symbolic name Q1_L4_L3 and the linguistic description ‘quadrilateral with one triangular concavity’; or the symbolic name Q2_L4P_2L3 and the linguistic description ‘rectangle with two triangular concavities’.

The name of the category of the cyclic objects is expressed in terms of the number of holes. For example, the name is given as the symbolic name A1_L3_L3 and the linguistic description ‘triangle with one triangular hole’; the symbolic name A1_L3_L4 and the linguistic description ‘triangle with one rectangular hole’ or the symbolic name A1_L4P_L3 and the linguistic description ‘rectangle with one triangular hole’.

The definition of the polygon category ∂_{Def} is given in the form of attributes of the class or in the generative form. The prototype category of non-named object is defined in relation to attributes of the polygonal object given by the class description. The method of generation of the object from the polygon category ∂_{MGe} is given by the procedure that randomly selects the value of attributes and next generates vertices of a polygon, or in a form of templates that can be subjected to affine transformation.

The visual concept of the polygon category ∂_{ViC} is obtained during the learning process. The prototype of the figure is learned by selecting representative of objects generated from the specific class and finding their symbolic names. Naming of the polygon is based on the visual concept obtained during the visual reasoning. The name of the specific category of the polygon is given by a symbolic name whereas a name of a general category of a polygon is given by a set of symbolic names – the visual concept. These problems are described in more detail in Chap. 5. The visual concept of the prototype of the non-named object consists usually of one symbolic name and is denoted as $\varphi^i = \{\eta\}$.

4.2.3.3.2. Category of Curves

The polygon is a well-defined geometrical figure and can be relatively easy to recognize. Geometrical curves are also well defined however the recognition is not so easy process. For the purposes of the present research curves are classified according to the knowledge of geometry and perceptual properties of figures.

The category of curves is divided into the category of curves in E^2, the category of curves in E^3 and the category of curves in more than three dimensions. The category of curves in E^2 is divided into four groups: the category of convex-closed curves ν_{ClCoC}, the category of concave-closed curves ν_{ClCcC}, the category of convex open curves ν_{OpCoC}, or the category of concave open curves ν_{OpCcC} and is given as: $\ldots\langle\nu_{\text{Cur}}\rangle \rhd \langle\nu_{\text{NNaC}}\rangle \rhd \langle\nu_{\text{OpCoC}}, \nu_{\text{OpCcC}}, \nu_{\text{ClCoC}}, \nu_{\text{ClCcC}}\rangle$. A large number of 2D open curves can be regarded as the graphs of functions. The category of open curves is divided into the category of function curves, and the category non-function curves. Taking into account properties of the function graphs the category of function curves is divided into the category monotonic curves ν_{Mon}, the category of non-monotonic curves ν_{NMon} or the category of periodic curves ν_{Per}. The categorical chain for the category of function curves is as $\ldots\langle\nu_{\text{Cur}}\rangle \rhd \langle\nu_{\text{NNaC}}\rangle \rhd \langle\nu_{\text{OpCcC}}\rangle \rhd \langle\nu_{\text{Fun}}\rangle \rhd \langle\nu_{\text{Mon}}, \nu_{\text{NMon}}, \nu_{\text{Per}}\rangle$. The category of periodic curves is divided into the category of regular-periodic curves, the category of para-periodic curves and the signal category. Examples of different categories of curves are shown in Figs. 4.19–4.21. As we can see curves from the different visual categories have clearly visible visual features.

Fig. 4.19. Example of members of the category of open non-function curves

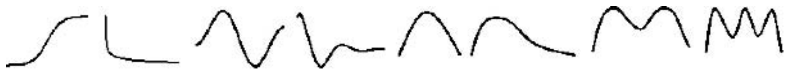


Fig. 4.20. Example of members of the different categories of open monotonic and non-monotonic function curves

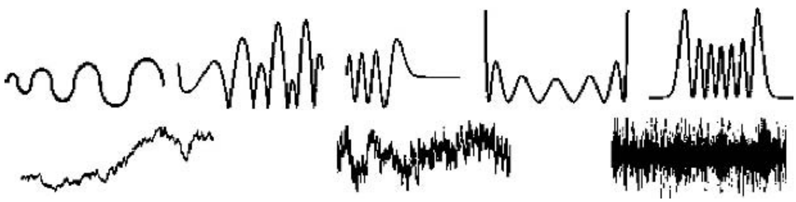


Fig. 4.21. Example of members of the different categories: the category of para-periodic curves and signal category

The closed curves in E2 can be often represented by mathematical equation. The different types of mathematical equations are used to represent the closed-convex and the closed-concave curves. Similarly, the category of non-named closed curves is divided into the category of convex curves and the category of concave curves. Based on the classification of the shape classes described in Chap. 2, the category of convex curves is divided into the category of K1 curves, the category of K2 curves, the category of K3 curves and the category of K4 curves. These curves are represented by the following categorical chain: $\ldots\langle \nu_{\mathrm{Cur}} \rangle \triangleright \langle \nu_{\mathrm{NNaC}} \rangle \triangleright \langle \nu_{\mathrm{ClCoC}} \rangle \triangleright \langle \nu_{\mathrm{K1}}, \nu_{\mathrm{K2}}, \nu_{\mathrm{K3}}, \nu_{\mathrm{K4}} \rangle$. The category of closed concave curves is divided into the category of cyclic and the category of non-cyclic curves. Examples of the members of categories of closed curves are shown in Figs. 4.22–4.24.

Fig. 4.22. Members of the category of closed convex curves

Fig. 4.23. Members of the category of closed concave curves

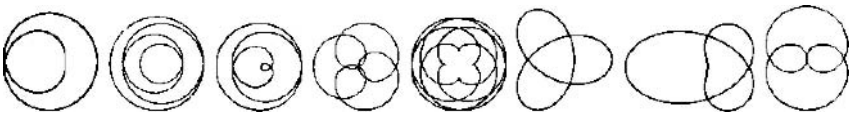


Fig. 4.24. Members of the category of closed convex cyclic curves and the category of closed concave cyclic curves

4.2.3.3.3. Category of Curve-Polygon Figures

The category of the curve-polygon figures is divided based on the classification of the curve-polygon classes described in Chap. 2. The convex curve-polygon class M consists of geometrical figures, which have curvilinear parts as well as linear segments. The name of the specific category of the curve-polygon figures refers to one of the specific curve-polygon classes. The category of curve-polygon figures is divided into the category of M1, the category of M2 or the category of M3 figures and are represented by the following categorical chain: $\ldots\langle \nu_{\mathrm{CPo}}\rangle \triangleright \langle \nu_{\mathrm{M1}}, \nu_{\mathrm{M2}}, \nu_{\mathrm{M3}}\rangle$.

4.2.3.3.4. Category of Mathematical Objects

Category of 2D Mathematical Objects

The category of figures is related to the category of mathematical objects and the category of statistical objects. When the category of figure refers to the perceptual or geometrical visual properties of the visual object, the category of mathematical object refers to the category of objects defined in terms of mathematical (geometrical) properties of the mathematical objects. The category of mathematical objects includes mathematical knowledge (that does not need to refer to the visible object) concerning the knowledge of the object as a mathematical object. For example, the category of hyperbolas includes the definition of the curve as a mathematical object, and the mathematical equation that makes it possible to generate the visual object. The category of hyperbolas derived from mathematical object includes the knowledge about the mathematical properties of the hyperbola as well as its relation to other mathematical objects. For example, hyperbola can be defined in the context of the differential equations and given as the result of the solution of the differential equation.

From the category of the convex polygon the category of the named convex polygon is derived. The members of the category of named polygons are polygons that have the special properties and are subject of the geometrical investigation. From the category of the named convex polygon the category of the specific named polygon is derived. Each prototype of this category has its characteristic name that refers to the category of mathematical object. For example, the category of named closed polygons is divided into the category of triangles, quadrilaterals, pentagons or octagons and represented by the following categorical chain as: $\ldots\langle v_{\mathrm{Pol}}\rangle \rhd \langle v_{\mathrm{NPol}}\rangle \rhd \langle v_{\mathrm{ClCoP}}\rangle \rhd \langle v_{\mathrm{Tri}}, v_{\mathrm{Qua}}, v_{\mathrm{Pen}}\rangle$, whereas the category of quadrilaterals is divided into the category of squares, rectangles, rhombus or trapezes and represented as follows $\ldots \rhd \langle v_{\mathrm{Qua}}\rangle \rhd \langle v_{\mathrm{Sqa}}, v_{\mathrm{Rec}}, v_{\mathrm{Rho}}, v_{\mathrm{Tra}}\rangle$.

The definition of the known geometrical figures such as polygons is based on the geometrical knowledge of the properties of the geometrical figures. For example, the concept of the polygonal figure such as a triangle is defined in the context of the convex polygon class Ω^{L}. The triangle is defined based on the cardinality of the set of nodes *V*, $a^{v} = |V|$. The triangle is a polygon that has three vertices $a^{v} = 3$. The right triangle is a triangle with one interior angle that is equal to 90°.

The definition of the figure can be given by the formal definition. Because SUS needs to communicate the results of the visual experience, the formal definition of the figure needs to be transformed into description given in one of the existing languages, e.g. Polish. The figures such as polygons were defined within geometry, and the definition as well as properties of the polygons that are described in geometrical literature can be used to represent knowledge about the figure concept. For example, the triangle can be defined by formal description and expressed in the form of the one of the existing languages (Polish, English) as follows: 'a triangle is a polygon that has three sides'; 'a triangle is a polygon that has three angles'; 'a triangle is a convex figure'. The definition expressed in the natural language needs to be translated into an intermediate representation. The intermediate representation makes it possible to translate the description from one language into another language. An intermediate representation is a general form of the definition. For example, the definitions 'the triangle is a polygon that has three angles' and 'the triangle is a polygon that has three sides' are given in the intermediate representation as 'O is R that has p' denoted as ($O \equiv R \triangleright p$), where O and R are categories and p is the characteristic property of the category. The formal definition refers to the attributes of the class. For example, the definition given in the linguistic form 'the triangle is a polygon that has three sides' has its symbolic form: $[a^v = 3] \Rightarrow L \succ L^3$. These forms are stored in the file in the following form 'triangle (side = 3)'. During the conversation process these forms are translated into the intermediate representation 'O is R that has p'. The intermediate representation makes it possible to express the same meaning by using different linguistic descriptions. For example, the sentence 'the triangle is a polygon that has three sides' is translated into the intermediate form $O \equiv R \triangleright p$. The definition can describe a different aspect of the category, for example: 'The triangle is a polygon that has three vertices'; 'The triangle is a polygon that has three sides'; 'The triangle is a polygon that has three angles'.

The definitions include the relations among different categories of the categorical chain. For example, 'the triangle is a convex figure' means that triangle category is derived from the convex polygon category. In the case of the parallelogram the different aspects of the geometrical properties are used as its definition, for example:

'A parallelogram is a quadrilateral that has opposite sides pairwise equal'

'A parallelogram is a quadrilateral that has two sides parallel and equal'

'A parallelogram is a quadrilateral that has the diagonals that bisect each other'

'A parallelogram is a quadrilateral that have the opposite angles equal'

The definition can have the same meaning but it can be expressed in different words, and after translation into intermediate representation the meaning is easy to find. For example, the definition 'A rectangle is a parallelogram that has all its angles right' or 'A triangle is a polygon that has three sides' is translated into intermediate representation in the form: $R \equiv O \Rightarrow p$.

The category of curves refers to curves as mathematical objects that are defined in the area of mathematics or physics. In general, a curve – a graph of an equation in two variables is the set of points in the plane whose coordinates satisfy the equation. The graph of the equation $xx + yy = 1$, for example, is the unit circle with the centre at the origin. The category of mathematical curves is derived from the category of figures. For example

Fig. 4.25. Example of members of the category of the open mathematical curves

the category of parabola is derived from the category of the open convex curves $\ldots\langle v_{\mathrm{Cur}}\rangle \rhd \langle v_{\mathrm{NCur}}\rangle \rhd \langle v_{\mathrm{OpCoC}}\rangle \rhd \langle v_{\mathrm{Par}}, v_{\mathrm{Hip}}\rangle$. Example of the members of the category of open mathematical (named) curves is given in Fig. 4.25.

The category of the closed named curves is derived from the curves category. The curve that has name is given by mathematical equation with one or more than one parameters. Changing values of parameters in the equation can cause clearly visible changes of shape (see example in Fig. 4.26). Based on these prosperities of the curve given in the form of the mathematical equation, the category of closed curves is divided into the category of the one shaped curves and category of the many shaped curves. The category of one shaped curves v_{OneS} is divided into the category of double heart curves or the category of the trifolium of Cramer: $\cdots\langle v_{\mathrm{Cur}}\rangle \rhd \langle v_{\mathrm{NCur}}\rangle \rhd \langle v_{\mathrm{ClCcC}}\rangle \rhd \langle v_{\mathrm{OneS}}\rangle \rhd \langle v_{\mathrm{DHer}}, v_{\mathrm{TCrm}}\rangle$.

Pascal's snail. Pascal's snail is given by the mathematical equation in the following form: the explicit equation $(x^2+y^2-ax)^2 = l^2(x^2+y^2)$, the parametric form

$$\begin{cases} x = a\cos^2\varphi + l\cos\varphi \\ y = a\cos\varphi\sin\varphi + l\sin\varphi \end{cases}$$

or the polar coordinates $\rho = a\cos\varphi + l$. The values of parameters a and l in the equation have a big influence on the shape of the generated curve. There are three different shape classes to which objects generated from the model given by Pascal's snail equation can be assigned, the convex curves, the concave curves and the cyclic curves. Figure 4.26 shows the example of curves generated for the selected values of parameters. For the values of parameters a and l given as $l \geq 2a$, the result is the convex curve (ellipsis) K_{E}^{1} (Fig. 4.26a). For the values of parameters $a < l < 2a$ and $l = 1.006$, the result is the concave curve (Fig. 4.26b). For $l = 1.4$ the result is the concave curve shown in Fig. 4.26c, for $l = 1.8$ and $l = 1.9$ the result is the nearly concave curve (Fig. 4.26d,e), for $a > l$ and l=0.9 and l=0.7 the result is the nearly cyclic curve (Fig. 4.26f,g), and for l=0.4 and l=0.08

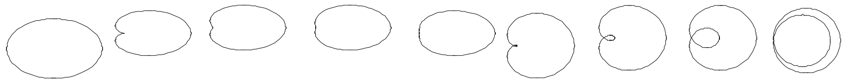


Fig. 4.26. Example of members of the category of Pascal's snail curves

the result is the cyclic curve (Fig. 4.26h,i). Each curve is represented by the symbolic name. For example, curve shown in Fig. 4.26b is given as $Q^1[M^1[K^1_E]](Q^2[L^3](2M^1))$, whereas the curve shown in Fig. 4.26h is given as $A[Q^1[M^1[K^1_E]](Q^2[L^3](2M^1))](K^2)$.

In the case of Bonet curve, shown in Fig. 4.27, the shape is changing from archetype of the class C1 into archetype of the class C2. The Pearl curve can be seen as an object that changes shape from archetypes of the class C1 through C2,…C6 into the archetype of the class C7. As we can notice objects in Fig. 4.27a, b (Bonet curve) and objects in Fig. 4.27f,g (Pearl curve) are very similar and the visual concept of Bonet curve will be the 'subset' of the visual concepts of Pearl curve. The equation of the Bonet curve is $x^4 - ax^3 + a^2y^2 = 0$, whereas the equation of the pearl curve is $\rho = x^s(a-x)^r = y^p / b^p$. Figure 4.27a, b shows members of the category of Bonet curves generated from the equation $x^4 - ax^3 + a^2y^2 = 0$, for $a = 1.2$ (Fig. 4.27a) and for $a = 0.5$ (Fig. 4.27b). Figure 4.27c–j shows members of the category the Pearl curve generated from the equation $\rho = x^S(A-x)^R = y^P / B^P$ for ($A = 1$, $B = 1$, $P = 6$, $R = 5$, $S = 1$) (Fig. 4.27c), ($A = 1$, $B = 1$, $P = 6$, $R = 3$, $S = 1$) (Fig. 4.27d), ($A = 1$, $B = 1$, $P = 4$, $R = 3$, $S = 1$) (Fig. 4.27e), ($A = 1$, $B = 1$, $P = 2$, $R = 3$, $S = 1$) (Fig. 4.27f), ($A = 1$, $B = 1$, $P = 4$, $R = 3$, $S = 7$) (Fig. 4.27g), ($A = 1$, $B = 1$, $P = 4$, $R = 3$, $S = 3$) (Fig. 4.27h), ($A = 1$, $B = 1$, $P = 6$, $R = 7$, $S = 1$) (Fig. 4.27i), ($A = 1$, $B = 1$, $P = 6$, $R = 9$, $S = 1$) (Fig. 4.27j).

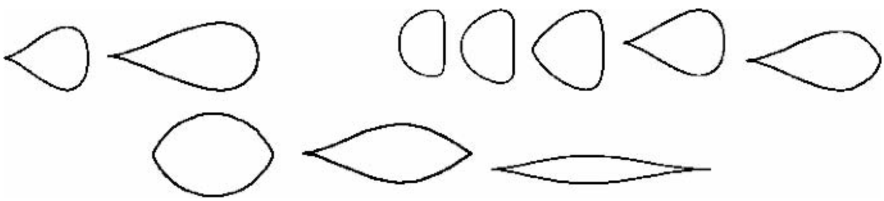


Fig. 4.27. Examples of members of the category of Bonet curves (**a**,**b**) and members of the category of Pearl curves (**c–j**)

The visual concept of categories such as the category of hypocycloida curves or the category of petal curves shown in Fig. 4.28 consist of symbolic names of the big range of shapes. Figure 4.28 shows the changes of shape for members of the category of hypocycloida curves and Fig. 4.29 shows the changes of shape for members of the category of petal curves.

The visual concept that is part of the knowledge schema of the category of the mathematical curve, given by its name, is obtained during the learning process. During learning process visual objects, representatives of the selected curve category, are generated and for each visual object the symbolic name is obtained. The visual concept φ^i is a set of different symbolic names $\varphi^i = \{\eta_1, \eta_2, \ldots, \eta_n\}$ obtained from the learning sample of visual objects. In the case when the specific name of the curve refers to one shape, e.g. circle, ellipse or trifolium of Cramer the visual concept is given by the one symbolic name $\varphi^i = \{\eta\}$. Examples of members of the category of closed curves that are represented by one shape are shown in Fig. 4.30. Figure 4.30a shows the category of circle ($x^2 + y^2 = r^2$), Fig. 4.30b shows the category of ellipsis, and Fig. 4.30c shows the category of hypocycloida

$$\begin{cases} x = (R - mR)\cos mt + m\cos(t - mt) \\ y = (R - mR)\sin mt - m\sin(t - mt) \end{cases},$$

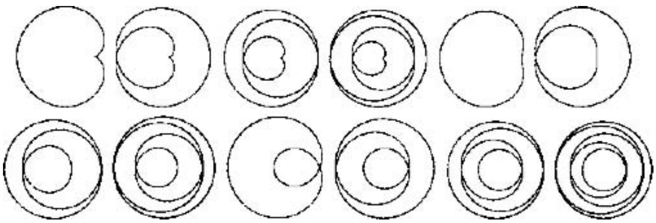


Fig. 4.28. Example of members of the category of hypocycloida curves

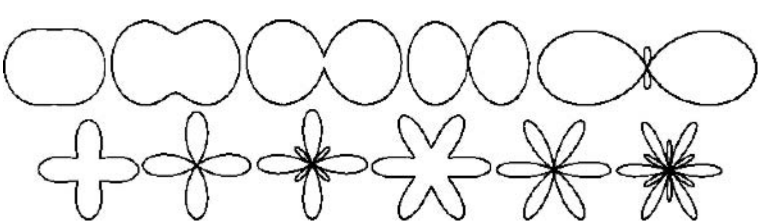


Fig. 4.29. Example of members of the category of Petal curves

Fig. 4.30. Examples of members of the category of closed curves that are represented by one shape (**a**) circle, (**b**) ellipsis, (**c**) hypocykloida, (**d**) double heart, (**e**) trifolium of Cramer

Fig. 4.30d shows the category of double heart $y^4 - 2x^2y^2 + x^4 - 6axy^2 - ax^3 + a^2x^2 = 0$, Fig. 4.30e shows the category of trifolium of Cramer, $x^4 + y^4 - 2ax(x^2 - y^2) = 0$.

Learning of the category of many shaped curves such as Pascal snail required generation of the whole range of objects of the different shapes from the mathematical equation of the Pascal snail. The knowledge needed to interpret the perceived object is given by the knowledge schema. To interpret a perceived object as a curve there is a need to learn the knowledge of the knowledge schema of the curve category. The knowledge schema for the curve category includes the visual concept ∂_{ViC}, the name ∂_{Nam}, the definition ∂_{Def} and the method of exemplar generation ∂_{MGe} and is given as $\ldots\langle\kappa_{\mathrm{Pol}}\rangle \rhd \langle\kappa_{\mathrm{NNaP}}\rangle \rhd \langle\kappa_{\mathrm{ClCoP}}\rangle \prec \{\partial_{\mathrm{ViC}}, \partial_{\mathrm{Nam}}, \partial_{\mathrm{MIn}}, \partial_{\mathrm{Def}}, \partial_{\mathrm{MGe}}\}$. The name of the learned category ∂_{Nam} is given in the form of the linguistic expression of the existing languages, for example, the language – English: 'Pascal snail', the language – Polish: 'Slimak Paskala', and is expressed as: $\ldots\langle\kappa_{\mathrm{Pol}}\rangle \rhd \langle\kappa_{\mathrm{NNaP}}\rangle \rhd \langle\kappa_{\mathrm{ClCoP}}\rangle \prec \{\partial_{\mathrm{Nam}}\} \,||\, \nu_{\mathrm{NLan}} \,||\, \langle\nu_{\mathrm{C1}}, \nu_{\mathrm{C2}}, \ldots, \nu_{\mathrm{CN}}\rangle$. The number of languages in which the name is expressed can be limited at the first stage of the learning process. The name expressed in other languages can be added during further stages of learning process. The method of exemplar generation ∂_{MGe} is given by the mathematical formula that can be expressed in the different forms such as explicit equation, parametric form or polar coordinates. For example, Pascal's snail, which was described in previous section, is given by explicit equation $(x^2 + y^2 - ax)^2 = l^2(x^2 + y^2)$, or in φ parametric form

$$\begin{cases} x = a\cos^2\varphi + l\cos\varphi \\ y = a\cos\varphi\sin\varphi + l\sin\varphi \end{cases}$$

or the polar coordinates: $\rho = a\cos\varphi + l$. The definition ∂_{Def} can be given in many different forms and can be added during further stages of learning. The definition can be expressed in the form of the symbolic representation or in the form of the linguistic description. The visual concept ∂_{ViC} is a set of symbolic names of the objects that are representatives of a given curve category.

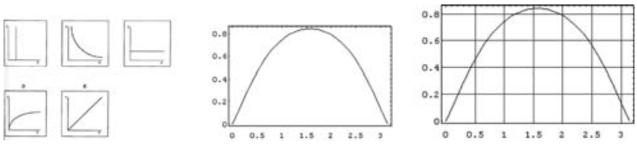
Category of 3D Figures

As it was described in the previous section the category of figure consists of the category of 2D figures $\langle \nu_{\text{2DF}} \rangle$, the category of 3D figures $\langle \nu_{\text{3DF}} \rangle$, the category *n*-D figures $\langle \nu_{\text{M3DF}} \rangle$. The category of 2D figures was described in the previous sections. In this section the brief description of the 3D polygon category that is derived from the 3D figure category is given. The category of 3D figures refers to geometrical objects that 'exist' in the 3D space. The category of 3D figures consists of category of 3D curves, category of 3D surfaces and category of 3D solids. From the category of 3D solids the category of polyhedrons is derived. The category of named polyhedrons ν_{NPol} consists of the category of cube ν_{Cub}, the category of prism ν_{Pri} or the category of cone ν_{Con} and is given as $O \triangleright \langle \nu_{\text{Fig}} \rangle \triangleright \langle \nu_{\text{3DF}} \rangle \triangleright \langle \nu_{\text{Pol}} \rangle \triangleright \langle \nu_{\text{NPol}} \rangle \triangleright \langle \nu_{\text{Cub}}, \nu_{\text{Pri}}, \nu_{\text{Con}} \rangle$. Figure 4.31 shows examples of the named polyhedrons.

Category of Mathematical Coordinate Systems

The graph of function is often plotted with additional visual information called the coordinate system. To locate points, lines, planes or other geometric forms their position must be referenced to some known position called a reference point. The position of geometric forms in 2D or 3D space is located by applying the coordinate system. When graph is plotted the axes can be included, labels can be put on the axes, the text of the different fonts can be used, frame can be plotted around the plot, labels can be put around the frames, thick marks can be drown and grid lines can be included. In Fig. 4.32a the most common representations of the graph of

Fig. 4.31. Examples of members of the polyhedrons category



a b c

Fig. 4.32. (**a**) The most common representation of the graph of the function with axes and labels on the axes. (**b**) Axes are drawn on the frame around the plot and (**c**) grid lines are added

functions with axes and labels on the axes are shown. In Fig. 4.32b axes are drawn on the frame around the plot, and in Fig. 4.32c grid lines are added.

To interpret visual mathematical objects the category of mathematical coordinate systems is introduced. The category of the mathematical coordinate system is derived from the pattern category that is composed from the category of mathematical elements. The element category of the coordinate system consists of many different categories such as axes, labels on the axes, the text of the different fonts, frame around the plot, labels around the frame, thick marks and grid lines and category of mathematical object such as curves:

$$\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{MCoS}} \rangle \triangleright \langle \nu_{\mathrm{Ax}}, \nu_{\mathrm{Lab}}, \nu_{\mathrm{Fra}}, \nu_{\mathrm{Mar}}, \nu_{\mathrm{Gra}} \rangle .$$

The knowledge that is supplied by knowledge schema of the element category of mathematical coordinate systems makes it possible to identify and remove the labels during process of interpretation of the visual object. The definition that is supplied by the knowledge schema gives the description of the element in terms of the visual attributes (e.g. the arrow marks the direction of axis). The category of mathematical coordinate systems includes graphs of functions or curves. The categorical chain of the elements of the coordinate system is as

$$\ldots \langle \kappa_{\mathrm{MOb}} \rangle \triangleright \langle \kappa_{\mathrm{MCoS}} \rangle \triangleright \langle \kappa_{\mathrm{Ax}}, \kappa_{\mathrm{Lab}}, \kappa_{\mathrm{Fra}}, \kappa_{\mathrm{Mar}}, \kappa_{\mathrm{Gra}} \rangle \prec \{\partial_{\mathrm{ViC}}, \partial_{\mathrm{Nam}}, \partial_{\mathrm{Def}}, \partial_{\mathrm{Int}}\} .$$

Coordinates are sets of numbers that specify the location of points in a space. The Cartesian coordinate system locates points in the plane with reference to a fixed point (origin) and the distance from two intersecting lines, called axes. The coordinates (x, y) of a point are distances measured along lines parallel to two fixed perpendicular axes. The polar coordinate system locates points in a plane with reference to a fixed point (origin) and an axis through that point. The coordinates are written (r, θ), in which r is the distance from the origin to any desired point and θ is the angle made by the vector r and the axis. A simple relationship exists between Cartesian coordinates given in terms of two reference axes (x, y) and the polar coordinates (r,θ), namely $x = r\cos\theta$ and $y = r\sin\theta$. The pattern category of coordinate systems consists of the elements of coordinate systems. The category of mathematical coordinate systems is divided into the category of Cartesian coordinate systems ν_{Car} and the category of polar coordinate systems ν_{Pol}. The categorical chain of the pattern of coordinate systems is as $\langle \Pi \rangle \rhd \langle \sigma_{\mathrm{Pat}} \rangle \rhd \langle \nu_{\mathrm{Sg}} \rangle \rhd \langle \nu_{\mathrm{VSym}} \rangle \rhd \langle \nu_{\mathrm{MCoS}} \rangle \rhd \langle \nu_{\mathrm{Car}}, \nu_{\mathrm{Pol}} \rangle$, whereas the knowledge chain is given as $\ldots \langle \kappa_{\mathrm{MOb}} \rangle \rhd \langle \kappa_{\mathrm{MCoSP}} \rangle \rhd \langle \nu_{\mathrm{Car}}, \nu_{\mathrm{Pol}} \rangle \prec \{\partial_{\mathrm{ViC}}, \partial_{\mathrm{Nam}}, \partial_{\mathrm{Def}}, \partial_{\mathrm{Int}}\}$ .The category of Cartesian coordinate systems is divided into the different types depends on the number of the different elements from element categories that are included into the object from the pattern category of coordinate systems. The specific categories derived from the category of Cartesian coordinate systems consist of the category of type A ν_{CartA}, the category of type B ν_{CartB} or the category of type C ν_{CartC} and is given by the following categorical chain: $\ldots \rhd \langle \nu_{\mathrm{VSym}} \rangle \rhd \langle \nu_{\mathrm{MCoS}} \rangle \rhd \langle \nu_{\mathrm{Car}} \rangle \rhd \langle \nu_{\mathrm{CartA}}, \nu_{\mathrm{CartB}}, \nu_{\mathrm{CartC}} \rangle$. Example of the different specific categories derived from the category of Cartesian coordinate systems is shown in Fig. 4.33.

4.2.3.3.5. Category of Statistical Objects

Statistics makes extensive use of various types of pictorial representations such as graphs. A pictorial representation displays at a glance much of the quantitative behaviour of the variables involved. The pictorial representations that are often used in statistics include bar graphs (Fig. 4.37e), line graphs (Fig. 4.37f) or ‘pie’ graphs (Fig. 4.3a–b) or box plots (Fig. 4.37c, d, g). A bar graph consists of a series of parallel bars or rectangles, the lengths of which are proportional to the data being presented. In a line graph the data are represented by points that are joined by straight-line segments. A ‘pie’ graph presents data in the form of slices of a pie, with the size, or angle, of each slice proportional to the quantity it represents.

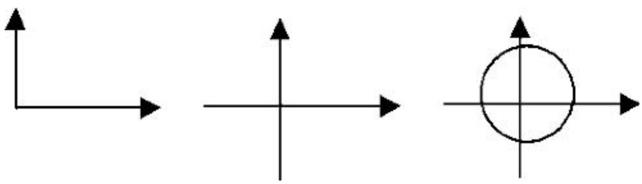


Fig. 4.33. The different type of the category of the Cartesian coordinate system

Fig. 4.34. Data transformed by applying the graphical transformation $G(\Pi)$ into phantom u (**a**) profile glyph, (**b**) sun glyph, (**c**) star function and (**d**) blob glyph

In statistic data that are gathered need to be transformed into more suitable form. The data can be transformed by one of the graphical transformations into a visual object such as a histobar, a profile, a star function, a sun glyph or a blob glyph. These geometric objects are used in multidimensional representations of data. Figure 4.34 shows an example of objects obtained from the transformation of the multidimensional data. The profile glyph is constructed similarly to histobar, except the bar heights are connected with a single profile. The star function is a profile function in polar coordinates. The sun glyph substitutes rays for the perimeter star. Blobs (plotted in polar coordinates) are obtained by applying the Fourier transformation

$$f(t) = \frac{x_1}{\sqrt{2}} + x_2 \sin(t) + x_3 \cos(t) + x_4 \sin(2t) + x_5 \cos(2t) + \cdots .$$

Cases that have similar values across all variables have comparable wave forms (the Fourier function is used). Figure 4.35 shows different geometrical transformations that transform a set of dots (data points) into the visual object.

The statistical visual object a member of the structural pattern category consists of objects that are members of the structural element category. Members of the category of statistical elements are statistical graphs v_{StGra}, the visual representations of the multidimensional

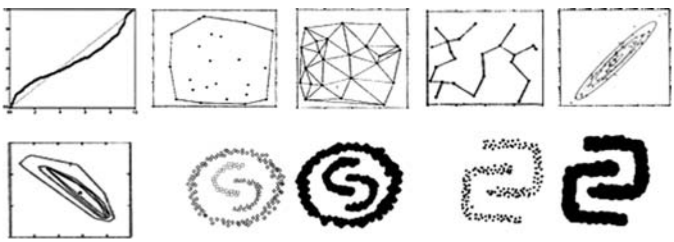


Fig. 4.35. Example of the different geometrical transformation that transform data points into the visual object

data ν_{Mul}, the graphs of the statistical distributions ν_{StGra}, the visual patterns of the scatter plots ν_{TSe} and members of the different categories such as axes ν_{Ax} or labels on the axes ν_{Lab} and are given as

$$\langle \Pi \rangle \rhd \langle \sigma_{\text{El}} \rangle \rhd \langle \nu_{\text{Sg}} \rangle \rhd \langle \nu_{\text{VSym}} \rangle \rhd \langle \nu_{\text{StCoS}} \rangle \rhd \langle \nu_{\text{Ax}}, \nu_{\text{Lab}}, \nu_{\text{StGra}}, \nu_{\text{Mul}}, \nu_{\text{TSe}}, \nu_{\text{SPlo}} \rangle .$$

The category of statistical graphs is divided into the category of discrete distributions and the category of continuous distributions. The category of discrete distributions is divided into the category of geometric distributions, the category of hypergeometric distributions or the category of negative binomial distributions. The category of continuous distributions (see Fig. 4.36) is divided into the category of uniform distributions ν_{Uni}, the category of exponential distributions ν_{Exp}, the category of normal PDF distributions ν_{NPDF}, the category of normal CDF distributions ν_{NCDF}, or the category of gamma distributions ν_{Gam}. Members of the category of continuous distributions are defined by using statistical knowledge. For example, the normal (Gaussian) distribution is the distribution function that is the indefinite integral of the normal density function, the graph of which is the typical bell-shaped normal curve. The visual chain of the category of statistical graph is given as

$$\langle \Pi \rangle \rhd \langle \sigma_{\text{El}} \rangle \rhd \langle \nu_{\text{Sg}} \rangle \rhd \langle \nu_{\text{VSym}} \rangle \rhd \langle \nu_{\text{StCoS}} \rangle \rhd \langle \nu_{\text{StGra}} \rangle \rhd \langle \nu_{\text{ConD}} \rangle \rhd \langle \nu_{\text{Uni}}, \nu_{\text{Exp}}, \nu_{\text{NPDF}}, \nu_{\text{NCDF}}, \nu_{\text{Gam}} \rangle .$$

The category of scatter-plot patterns refers to objects that are result of the visual inspection of the scatter-plot data. The visual inspection of the scatter-plot data makes it possible to select the proper statistical procedure. The category of the scatter-plot data is divided into the category of the regression analysis ν_{Reg}, the category of cluster analysis ν_{Clu} or the category of the discriminate analysis ν_{Dis} and is represented as: $\ldots\langle \nu_{\text{StCoS}} \rangle \rhd \langle \nu_{\text{SPlo}} \rangle \rhd \langle \nu_{\text{Reg}}, \nu_{\text{Clu}}, \nu_{\text{Dis}} \rangle$. The category of the regression analysis is divided into the category of linear regression, the category of

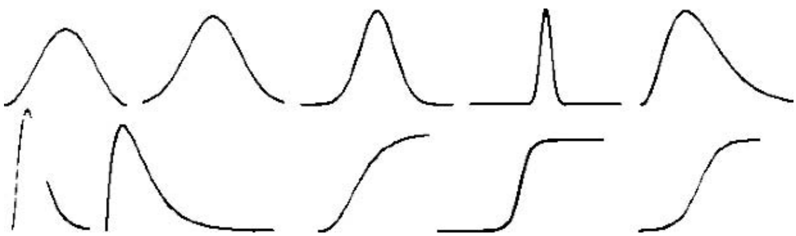


Fig. 4.36. Examples of objects from category of statistical graphs – continuous distributions

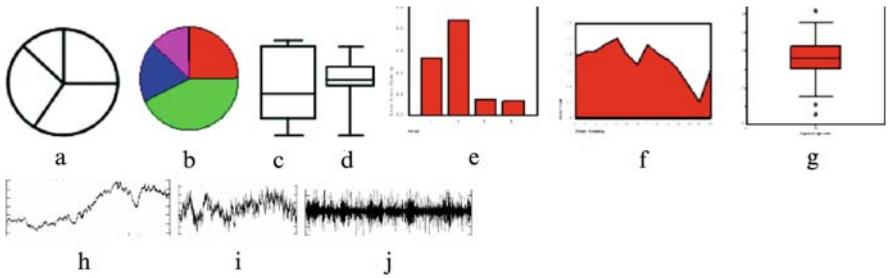


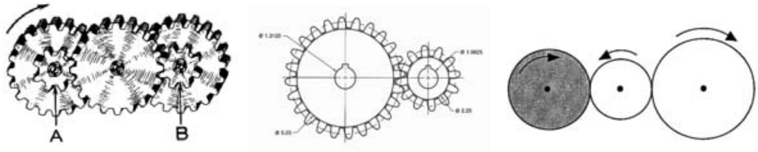
Fig. 4.37. Members of the category of statistical objects

nonlinear regression or the category of weighted least-square regression. Examples of members of the different categories of statistical objects are shown in Fig. 4.37.

The statistical object is often shown in the context of the coordinate system. The category of the statistical coordinate system is very similar to the category of the mathematical coordinate system. Examples of members of the category of the statistical coordinate system are shown in Fig. 4.37e–j.

4.2.3.3.6. Category of Visual Physical Models

A model is a simplified representation of the real-world object or phenomenon and, as such, includes only those aspects relevant to the problem. A model of freely falling bodies, for example, does not refer to the colour, texture or shape of the body involved. Furthermore, a model may not include all relevant variables because a small percentage of these may account for most of the phenomenon to be explained. Many of the simplifications used produce some error in predictions derived from the model, but these can often be kept small compared to the magnitude of the improvement in operations that can be extracted from them. The earliest models were physical representations such as model ships, airplanes, tanks and wind tunnels. Physical models are usually easy to construct, but only for relatively simple objects or systems, and are usually difficult to change. The next step beyond the physical model is the graph, easier to construct and manipulate but more abstract. Since graphic representation of more than three variables is difficult, symbolic models came into use. The visual representation of the physical model supply information that cannot be conveyed by any other means. The category of physical model is derived from the pattern category. The category of the physical model is combination of the different element categories. The element category of the physical model consists of letters, words, arrows, lines, arch and the schematic representation of the physical phenomena. The pattern category



a b c

Fig. 4.38. Different visual representations of the gears

of the visual physical models refers to the 2D visual representation of the real-world object or phenomena. When the 2D representation of the 3D real-world object is usually given by one of the image projections, the 2D representation of the visual physical model is given by the conventional transformation. The 2D visual representation can be given by the realistic drawing, the schematic engineering drawing or the schematic conventional drawing. For example, Fig. 4.38 shows the different visual representations of the gears: the realistic drawing (Fig. 4.38a), the schematic engineering drawing (Fig. 4.38b) and the schematic conventional drawing (Fig. 4.38c).

The category of physical visual models consists of predefined visual symbols. The element category of the category of physical model includes the schematic representation of the real-world object or and is given as: phenomena (physical object), letters, words, arrows, lines, arcs $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{PhMo}} \rangle \triangleright \langle \nu_{\mathrm{PhOb}}, \nu_{\mathrm{Lab}}, \nu_{\mathrm{Ax}}, \nu_{\mathrm{Mar}}, \nu_{\mathrm{Gra}} \rangle$. The pattern category of the category of the physical visual models composes the different elements of the element category into one of the visual objects. The category of the physical visual models is divided into the lever category, the wedge category, the axle category, the pulley category, the inclined plane category or the pendulum category, and is expressed as $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pt}} \rangle \triangleright \cdots \triangleright \langle \nu_{\mathrm{PhMo}} \rangle \triangleright \langle \nu_{\mathrm{Lev}}, \nu_{\mathrm{Wed}}, \nu_{\mathrm{Pul}}, \nu_{\mathrm{Rol}}, \nu_{\mathrm{Pen}}, \nu_{\mathrm{InPl}} \rangle$. The perceived object is interpreted as a member of the category of the of physical visual models based on knowledge supplied by the knowledge schema of the knowledge category and the of physical visual models given as $\ldots \langle \kappa_{\mathrm{PhOb}} \rangle \triangleright \langle \kappa_{\mathrm{PhMo}} \rangle \triangleright \langle \kappa_{\mathrm{Lev}}, \kappa_{\mathrm{Wed}}, \kappa_{\mathrm{Pul}}, \kappa_{\mathrm{Rol}}, \kappa_{\mathrm{Pen}}, \kappa_{\mathrm{InPl}} \rangle$. The category of the visual physical models is linked with the category of the real-world objects and is denoted as

$$\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{ReO}} \rangle \triangleright \langle \nu_{\mathrm{Ear}} \rangle \triangleright \langle \nu_{\mathrm{NLiv}} \rangle \triangleright \langle \nu_{\mathrm{MMad}} \rangle \triangleright \langle \nu_{\mathrm{SimM}} \rangle \triangleright \langle \nu_{\mathrm{EAsP}} \rangle \triangleright \langle \nu_{\mathrm{Lev}}, \nu_{\mathrm{Wed}}, \nu_{\mathrm{Pul}}, \nu_{\mathrm{Rol}}, \nu_{\mathrm{Pen}}, \nu_{\mathrm{InPl}} \rangle.$$

This link between two categories makes it possible to interpret the perceived object both as a category of the real-world object and as a category of the physical model. Examples of members of the different categories of physical visual models are shown in Fig. 4.39.

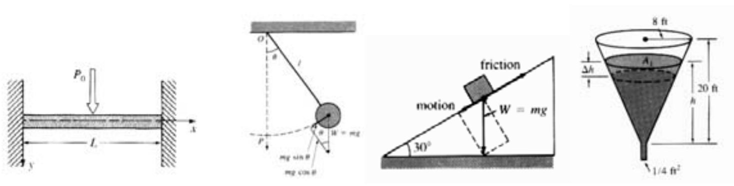


Fig. 4.39. Examples of members of the different categories of physical visual models

4.2.3.4. Category of Signs

In literature terms a sign, a letter and a symbol are not well defined and are often used as synonyms. In contrast to the letter, the sign is not part of the system of any existing language. In contrast to the word that consists of the string of letters and is the meaningful element of the language, there are no rules that make it possible to compose the word-like unit from the system of signs. Symbols such as mathematical symbols can be used to compose complex expressions however there are no strict rules that govern the composition of the meaningful units such as words, sentences or text. Sign has been defined as a 'concrete denoter' possessing an inherent specific meaning. The most common signs are pictures or drawings, although a human posture like a clenched fist, an outstretched arm, or a hand posed in a 'Stop' gesture are signs. A sign contains meanings of an intrinsic nature. Images of the real-world objects are often used as a sign. Such an image is produced by applying conventional rules of drawing rather than geometric projections. Examples of signs are shown in Fig. 4.40.

Signs may be presented graphically, as in the cross for Christianity, the Red Cross or the crescent for the life-preserving agencies of Christian and Islamic countries. Signs that do not resemble any real-world object are often called symbols. These symbols (e.g. the musical score) are similar to the letters of the existing languages. Language is a system of conventional spoken or written symbols by means of which human beings communicate. Figure 4.41 shows examples of mathematical symbols, musical scores, currency symbols, letter-like symbols or symbols of engineering schema.

Fig. 4.40. Examples of signs

In this book visual symbols are defined in the context of structural category of patterns. Visual symbols are signs that can be composed into complex expressions or schemas (pattern). Meaning of the visual symbol that is part of the complex expression or schema depends on the meaning of the others symbols. The category of the visual symbols refers to the 2D usually black and white objects that have relatively well-established shape. Members of the sign category are usually members of perceptual categories such as the silhouette category or the line-drawing category.

The variation of shapes of the selected signs is often meaningful. For example, in the case of letters, the different shape of the same letter refers to the different fonts of this letter. To distinguish visual symbols from other signs such as the road sign, the signs that are not visual symbols will be called the symbolic signs. In contrast to visual symbols, symbolic signs are visual objects that cannot be used to compose the meaningful patterns. The symbolic sign has its conventional meaning that does not depend on the meaning of other symbolic signs. For example, road signs are related to each other, however, there are no rules to construct the meaning of the complex expressions that consist of signs or are based on the sequence of signs.

The sign category refers to the visual object meaning of which is based on the system of conventional rules (the code). The category of the sign is derived from category of the visual object that is given as follows $O \triangleright \langle \nu_{\mathrm{ReO}}, \nu_{\mathrm{ImO}}, \nu_{\mathrm{Sig}}, \nu_{\mathrm{Let}}, \nu_{\mathrm{Fig}} \rangle$. From the category of signs, the category of 2D signs and 3D signs is derived. The category of 2D signs is divided into the category of visual symbols ν_{VSym} and the category of symbolic signs ν_{SymS} and is given as: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright\triangleright \langle \nu_{\mathrm{Sig}} \rangle \triangleright \langle \nu_{\mathrm{2DSig}} \rangle \triangleright \langle \nu_{\mathrm{SymS}}, \nu_{\mathrm{VSym}} \rangle$. The category of the symbolic signs will be described further in this chapter.

4.2.3.4.1. Category of Visual Symbols

The category of visual symbols is the category of the well-defined objects that are used to compose the complex objects (patterns). The category of visual symbols is derived from the structural category – the element category or the pattern category. Examples of members of the structural element category σ_{El} and the ontological category of visual symbols ν_{VSym} are shown in Fig. 4.41.

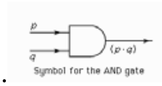
A 5 ∑ ♠ ♪ ∂ ≡ ₣ ® § Δ ≠ ≤ % 1 / * ∏ ∫ . 

Fig. 4.41. Examples of members of the structural element category and the ontological category of visual symbols

From the category of visual symbols the category of the mathematical symbols ν_{Mth}, the category of logical symbols ν_{Log}, the category of musical symbols ν_{Mus}, the category of currency symbol ν_{Cur} or the category of engineering symbols ν_{EnSym} is derived that is given as $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{Mth}}, \nu_{\mathrm{Log}}, \nu_{\mathrm{Cur}}, \nu_{\mathrm{Mus}}, \nu_{\mathrm{EnSym}} \rangle$.

Category of Mathematical Symbols

From the category of the visual symbols, the category of mathematical elements is derived. The category of mathematical symbols (elements) is divided into the category of mathematical operators ν_{Opr}, the category of mathematical relations ν_{Rel}, the category of logical operators ν_{Log}, the category of special mathematical symbols ν_{SSym} or the category of ciphers ν_{Cip} and given as $\ldots \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{Mth}} \rangle \triangleright \langle \nu_{\mathrm{Cip}}, \nu_{\mathrm{Opr}}, \nu_{\mathrm{Rel}}, \nu_{\mathrm{Log}}, \nu_{\mathrm{SSym}} \rangle$. The category of knowledge of mathematical elements supplies knowledge needed during interpretation of the perceived object as a member of the category of mathematical symbols. The category of knowledge of mathematical elements is derived from the category of the mathematical objects κ_{MtOb} and given as: $\ldots \langle \kappa_{\mathrm{MtOb}} \rangle \triangleright \langle \kappa_{\mathrm{Cip}}, \kappa_{\mathrm{Oprr}}, \kappa_{\mathrm{Rel}}, \kappa_{\mathrm{Log}}, \kappa_{\mathrm{Syn}} \rangle$. The category of cipher is derived from the category of mathematical symbols and is divided into the Arabic cipher (0, 1, 2, 3, 4, 5, 6, 7, 8, 9), Roman cipher (I, V, M) and denoted as $\ldots \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{Mth}} \rangle \triangleright \langle \nu_{\mathrm{Cip}} \rangle \triangleright \langle \nu_{\mathrm{Ara}}, \nu_{\mathrm{Rom}} \rangle$. From the cipher elements the bigger units can be composed to denote numerals.

The logician uses a symbolic notation to express proposition clearly and unambiguously and to enable manipulations and tests of validity to be more easily applied. The symbolic notation and connectives are used to establish categories of logical symbols. Connective is a word or group of words that joins two or more propositions together to form a connective proposition. Commonly used connectives include conjunction (‘and’), disjunction (‘or’), negation (‘not’), conditional (‘if … then’), and bi-conditional (‘if and only if’). The connective has been denoted by the symbols that are taken from the set theory. The category of logical symbols is derived from the category of mathematical symbols. The category of the logical symbols consists of the conjunction category, the disjunction category or the negation category and is given as $\ldots \triangleright \langle \nu_{\text{VSym}} \rangle \triangleright \langle \nu_{\text{Mth}} \rangle \triangleright \langle \nu_{\text{Log}} \rangle \triangleright \langle \nu_{\text{Dis}}, \nu_{\text{Con}}, \nu_{\text{Neg}} \rangle$.

Elementary arithmetic is concerned primarily with the effect of certain operations, such as addition or multiplication, on specified numbers. The integral operator (symbol) is used to denote integration as a technique of finding a function $g(x)$ the derivative of which, $\mathrm{D}g(x)$, is equal to a given function $f(x)$. The category of mathematical operator is derived from the category of mathematical visual symbols. The category of operators is divided into the addition category ν_{Add}, the multiplication category ν_{Mul}, the differentiation category ν_{Dif} or the integration category ν_{Int}. The categorical chain of the category of mathematical operators is given as $\ldots \triangleright \langle \nu_{\text{VSym}} \rangle \triangleright \langle \nu_{\text{Mth}} \rangle \triangleright \langle \nu_{\text{Opr}} \rangle \triangleright \langle \nu_{\text{Add}}, \nu_{\text{Mul}}, \nu_{\text{Dif}}, \nu_{\text{Int}} \rangle$. Examples of members of the category of mathematical operators are as +, –, /, *, ∂, Δ, ∏, ∑, ∫. The knowledge schema of the category of mathematical operators κ_{Opr} includes the visual concept ∂_{ViC}, the definition ∂_{Def}, the method of exemplar generation ∂_{MGen}, the link to the pattern of possible expressions ∂_{Link}, the name of the operator ∂_{Nam} and is given as: $\ldots \langle \kappa_{\text{MOb}} \rangle \triangleright \langle \kappa_{\text{MCoSP}} \rangle \triangleright \langle \kappa_{\text{Opr}} \rangle \prec \{\partial_{\text{ViC}}, \partial_{\text{Nam}}, \partial_{\text{Def}}, \partial_{\text{Link}}, \partial_{\text{MGen}}\}$. The link to the category of pattern supplies the knowledge of composition of the mathematical expression. The knowledge schema can be shared by the several levels of the categorical hierarchies. For example, each specific category of operators such as the addition category shares the knowledge of the category of mathematical operators:

$$\ldots \triangleright \langle \kappa_{\text{Opr}} \rangle \triangleright \langle \kappa_{\text{Add}} \rangle \prec \{\partial_{\text{ViC}}, \partial_{\text{Nam}}, \partial_{\text{Def}}, \partial_{\text{Link}}, \partial_{\text{MGen}}\}.$$

The specific symbols are used to denote the relations such as equality or inequality relations: Examples of symbols of mathematical relations are as follows: =, ≈, ≠, ≡, ≤, ≥. The category of symbols of mathematical relations is derived from the category of mathematical symbols. The category of symbols of mathematical relations is divided into the equality relation category or the inequality relation category and is represented as: $\ldots \triangleright \langle \nu_{\text{VSym}} \rangle \triangleright \langle \nu_{\text{Mth}} \rangle \triangleright \langle \nu_{\text{Rel}} \rangle \triangleright \langle \nu_{\text{Equ}}, \nu_{\text{NEq}}, \nu_{\text{Big}}, \nu_{\text{Les}} \rangle$. The knowledge schema of category of mathematical relations includes the visual concept, the method of generation, the link to the pattern of possible expressions, the name of the operator and the definition.

Members of the category of mathematical symbols that is derived from the structural element category are used to compose the complex mathematical expressions. The mathematical expression is a combination of ciphers (numbers), letters and other mathematical elements into the unit that has the mathematical meaning. The mathematical meaning is given by rules of mathematical expressions. The category of mathematical expressions is derived from the structural pattern category. The category of mathematical expression consists of the category of equations ν_{Eqt}, the category of functions ν_{Fun}, the category of differentiations ν_{Dif} or the category of integrations ν_{Int} and is expressed as $\ldots\langle\sigma_{\text{Pt}}\rangle \triangleright \cdots \langle\nu_{\text{MtEx}}\rangle \triangleright \langle\nu_{\text{Eqt}}, \nu_{\text{Fun}}, \nu_{\text{Dif}}, \nu_{\text{Int}}\rangle$. The knowledge schema of the category of mathematical expressions includes rules that make it possible to check if a perceived object is a member of a given category of mathematical expressions. The member of the category of mathematical expressions (pattern) is composed of elements from the category of mathematical symbols. For example, the expressions $a+b=c$, $a+b$, $a+b+c=d$, $a+b+\cdots=h$, $2+3$, $x+4=5$, $x+y-3x+6=0$ are members of the category of mathematical expressions.

Category of Musical Symbols

Music has been called both the most mathematical and the most abstract of the arts. Unlike words musical tones in themselves have no concrete associations, and only gain meaning when they are combined into musical patterns. Musical tones can be expressed in the graphical form by using musical notation. Musical notation is any set of symbols used to convey in written form the composer's wishes to the performer. It includes not only the signs that communicate pitch and duration but also the various terms and marks that explain what the notes alone cannot. Notation must serve many types and styles of music – concert, popular and folk – and must convey information quickly, enabling the performer to read the composer's instructions at the speed the music is to be performed. The category of musical symbols derived from the category of visual symbols is divided into the note category ν_{Not}, the rest category ν_{Res}, the clefs category ν_{Cle}, the flat category ν_{Fla} or the sharp category ν_{Sha} and is given as: $\langle\sigma_{\text{El}}\rangle \triangleright \cdots \triangleright \langle\nu_{\text{VSym}}\rangle \triangleright \langle\nu_{\text{MuS}}\rangle \triangleright \langle\nu_{\text{Not}}, \nu_{\text{Res}}, \nu_{\text{Cle}}, \nu_{\text{Sha}}, \nu_{\text{Fla}}\rangle$. The category of notes is divided into the whole note category, 1/2 note category, 1/4 note category, 1/8 note category, 1/16 note category, 1/32 note category and 1/64 note category and is given as: $\ldots \triangleright \langle\nu_{\text{Not}}\rangle \triangleright \langle\nu_{1}, \nu_{1/2}, \nu_{1/4}, \nu_{1/8}, \nu_{1/16}\rangle$. Similarly, the category of rest is divided into the whole rest category, 1/2 rest category, 1/4 rest category, 1/8 rest category or 1/16 rest category. The category of clefs is divided into the treble clefs category ν_{Tre}, the bass clefs category ν_{Bas} and the C clefs category ν_{C} and is given as: $\ldots \triangleright \langle \nu_{\mathrm{Cle}} \rangle \triangleright \langle \nu_{\mathrm{Tre}}, \nu_{\mathrm{Bas}}, \nu_{\mathrm{C}} \rangle$. The visual categories of musical symbols are linked with the knowledge categories of musical symbols. The knowledge category of musical symbols supplies the knowledge that makes it possible to interpret the musical visual symbol as the specific musical sound. The knowledge category of musical symbols is derived from the category of musical object κ_{MusO} and is given by the following categorical chain: $\kappa_{\mathrm{KB}} \triangleright \langle \kappa_{\mathrm{KOb}} \rangle \triangleright \langle \kappa_{\mathrm{MusO}} \rangle \triangleright \langle \kappa_{\mathrm{Not}}, \kappa_{\mathrm{Res}}, \kappa_{\mathrm{Cle}} \rangle$. Understanding the musical symbols means knowing how to make the sound using the musical instruments. The musical element such as note refers to the elements of musical sound, such as pitch or the location of musical sound on the scale.

The musical elements that are members of the category of musical symbols can be combined into musical patterns by placing these symbols on the staff. The musical symbols placed on the staff are interpreted in terms of the musical notation and can be used to produce the musical sound. The musical categories such as the beaming category ν_{Bea}, the harmony category ν_{Har} or the melody category ν_{Mel} are derived from the structural pattern category and are given by the categorical chain $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pt}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{Mus}} \rangle \triangleright \langle \nu_{\mathrm{Bea}}, \nu_{\mathrm{Har}}, \nu_{\mathrm{Mel}} \rangle$. Sequence of musical symbols placed on the staff can be transformed into musical sound by musician playing on one of the musical instruments.

Category of Engineering Symbols

Engineering is the creative application of ‘scientific principles’ to design or develop structures, machines, apparatus or manufacturing processes. Associated with engineering is a great body of special knowledge. Professional practice involves extensive training in the application of that knowledge. When scientist produces systematized knowledge of the physical world the engineer utilizes this knowledge to solve practical problems. Engineering is based on physics, chemistry and mathematics, and their extensions into materials science, solid and fluid mechanics, thermodynamics, transfer and rate processes, and systems analysis. There are different branches of engineering such as chemical engineering, mechanical engineering or electronic engineering. For example, mechanical engineering is the branch of engineering concerned with the design, manufacture, installation, and operation of engines and machines and with manufacturing processes whereas electrical and electronic engineering is the branch of engineering concerned with the practical applications of electricity in all their forms, including those of the field of electronics. The visual knowledge in engineering is given in the form of engineering schemes that consist of engineering symbols.

The category of engineering symbols (elements) is divided into the category of electronic symbols ν_{ElES}, the category of mechanical symbols ν_{MeES}, or the category of chemical symbols ν_{ChES} as shown by the categorical chain:

$$\langle \Pi \rangle \triangleright \langle \sigma_{\text{El}} \rangle \triangleright \langle \nu_{\text{Sg}} \rangle \triangleright \langle \nu_{\text{VSym}} \rangle \triangleright \langle \nu_{\text{EngS}} \rangle \triangleright \langle \nu_{\text{ElES}}, \nu_{\text{MeES}}, \nu_{\text{ChES}} \rangle .$$

Each category of engineering symbols refers to the specific branch of engineering such as electronic engineering. To understand the functionality of electronic tools or to solve the electronic problem the knowledge from the area of electronic engineering is needed. The visual knowledge in the field of electronic engineering is given in the form of schemes of electronic circuits. The scheme of the electronic circuit is used to solve the electrical problems or to build the electronic tools. The category of electronic symbols is 'linked' with the category of the real-world electronic elements. Electrical symbols are interpreted as a 'substitute' of the real-world objects such as a resistor or a battery.

The category of electronic symbols ν_{ElES} is divided into the resistor category ν_{Rez}, the inductor category ν_{Ind}, the transformer category ν_{Trn}, the capacitor category ν_{Cap}, the diode category ν_{Dio}, the transistor category of ν_{Trz}, and is given as:

$$\langle \Pi \rangle \triangleright \langle \sigma_{\text{El}} \rangle \triangleright \cdots \triangleright \langle \nu_{\text{ElES}} \rangle \triangleright \langle \nu_{\text{Rez}}, \nu_{\text{Ind}}, \nu_{\text{Trn}}, \nu_{\text{Cap}}, \nu_{\text{Dio}}, \nu_{\text{Trz}} \rangle .$$

The capacitor category is divided into the electrolytic capacitor category ν_{Elec} or the ceremic capacitor category ν_{CerC} and given as:

$$\langle \Pi \rangle \triangleright \langle \sigma_{\text{El}} \rangle \triangleright \cdots \triangleright \langle \nu_{\text{ElES}} \rangle \triangleright \langle \nu_{\text{Cap}} \rangle \triangleright \langle \nu_{\text{CerC}}, \nu_{\text{EleC}} \rangle .$$

Figure 4.42 shows members of the different categories of electronic symbols (a) the resistor category, (b) the inductor category, (c) the transformer category, (d) the capacitor category, (e) the electrolytic capacitor category, (f) the diode category, (g) the light emitting diode category, (h) the bipolar transistor category, (i) the field-effect transistor category.

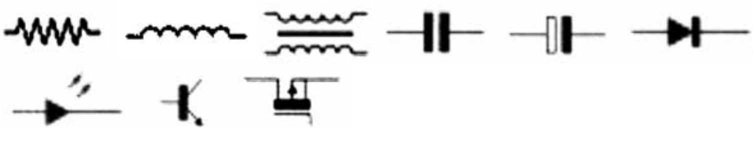


Fig. 4.42. Members of the different categories of electronic symbols

Fig. 4.43. The category of electronic elements the resistor, capacitor, capacitor electrolytic, bipolar transistor, field-effect transistor

The category of electronic symbols is linked with the category of electronic elements that is derived from category of real-world objects. Figure 4.43 shows examples of members of the category of electronic elements that refers to members of the category of electronic symbols. Each member of the category of the electronic symbols refers to the member of the category of electronic elements derived from the category of the real-world objects. The specific categories such as the resistor category ν_{Rez}, the capacitor category ν_{Cap}, the diode category ν_{Dio}, or the transistor category ν_{Trz} that are derived from the category of real-world objects are represented by following categorical chain: $\langle \Pi \rangle \rhd \langle \sigma_{\mathrm{El}} \rangle \rhd \langle \nu_{\mathrm{ReO}} \rangle \rhd \langle \nu_{\mathrm{Ear}} \rangle \rhd \langle \nu_{\mathrm{NLiv}} \rangle \rhd \langle \nu_{\mathrm{MMad}} \rangle \rhd \langle \nu_{\mathrm{AsP}} \rangle \rhd \langle \nu_{\mathrm{ElAsP}} \rangle \rhd \langle \nu_{\mathrm{Rez}}, \nu_{\mathrm{Trn}}, \nu_{\mathrm{Cap}}, \nu_{\mathrm{Dio}}, \nu_{\mathrm{Trz}} \rangle$.

The electronic engineering symbols are parts of the schemes of electronic circuits (Fig. 4.44). The category of schemes of electronic circuits is derived from the structural pattern category and is divided into the category of the ERC ν_{ERC} (battery, resistor and capacitor) circuits, the category of the ERLC ν_{ERLC} (battery, resistor, capacitor and inductor) circuits or the category ERLCT ν_{ERLCT} (battery, resistor, capacitor, inductor and transistor) circuits. The categorical chain of the specific categories derived from the category of schemes of electronic circuits is as $\langle \Pi \rangle \rhd \langle \sigma_{\mathrm{Pt}} \rangle \rhd \langle \nu_{\mathrm{Sg}} \rangle \rhd \langle \nu_{\mathrm{VSym}} \rangle \rhd \langle \nu_{\mathrm{EngSh}} \rangle \rhd \langle \nu_{\mathrm{ElESh}} \rangle \rhd \langle \nu_{\mathrm{ERC}}, \nu_{\mathrm{ERLC}}, \nu_{\mathrm{ERLCT}} \rangle$.

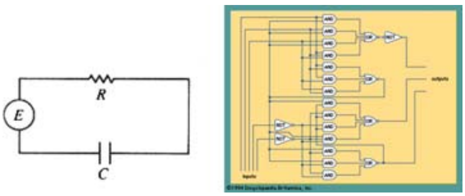


Fig. 4.44. Examples of members of the category of schemes of electronic circuits

4.2.3.4.2. Category of Symbolic Signs

The most common signs encountered in daily life are pictures or drawings. As it was described in the previous section the category of signs is derived from the category of visual objects. The category of signs is divided into the category of visual symbols and the category of the symbolic signs. As it was mention at the beginning of this chapter, in contrast to the members of the category of the visual symbols, members of the category of symbolic signs cannot be used to compose any complex meaningful object. The meaning of elements of the category of symbolic signs does not depend on the meaning of other elements of the same category. From the category of symbolic signs the category of the trademark signs ν_{TrS}, the category of road signs ν_{RoS} or the category of cross signs ν_{CroS} is derived: $O \triangleright \langle \nu_{\mathrm{Sig}} \rangle \triangleright \langle \nu_{\mathrm{2DSig}} \rangle \triangleright \langle \nu_{\mathrm{SymS}} \rangle \triangleright \langle \nu_{\mathrm{RoS}}, \nu_{\mathrm{CroS}}, \nu_{\mathrm{TrS}} \rangle$.

The category of road signs is derived from the category of symbolic signs. The roadways signs advise the driver of special regulations and provide information about hazards and navigation. The meaning of the complex road signs depends on the shape of its 'background' or shape of a sign as a whole. Shape of the sign as a whole (e.g. circle, triangle) refers to the general category of the road signs. For example, an octagon is used for stop signs, a triangle for warning signs or a rectangle for free way directions. The category of road signs are divided into the category of warning signs or the category of information signs and are given by the following categorical chain: $O \triangleright \langle \nu_{\mathrm{Sig}} \rangle \triangleright \langle \nu_{\mathrm{2DSig}} \rangle \triangleright \langle \nu_{\mathrm{SymS}} \rangle \triangleright \langle \nu_{\mathrm{RoS}} \rangle \triangleright \langle \nu_{\mathrm{war}}, \nu_{\mathrm{Inf}} \rangle$. Figure 4.45 shows different categories of road signs.

The category of road signs is interpreted based on the first meaning and the second meaning of the sign. The first meaning of the sign refers to the visual object that is interpreted as a member of the category of real-world objects or the category of letters. For example, the visual object shown in Fig. 4.46a is interpreted as a member of the category of letters, the letter 'P', whereas the visual object shown in Fig. 4.46d is interpreted as a member of the category of real-world objects, the cup. The second meaning refers to the visual object that is interpreted as a member of the category of road signs. For example, the visual object shown in

Fig. 4.45. Examples of members of the different categories of road signs

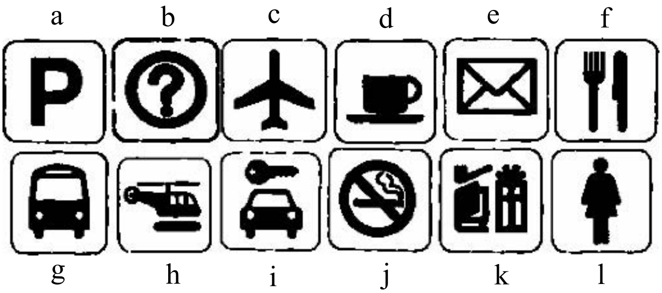


Fig. 4.46. Examples of members of the category of information signs

Fig. 4.46d is interpreted as a member of the category of road signs – the cup indicates that there is a cafén ear by. The different categories of information signs that give the specific information include both the letters and schematic representations of the real-world objects. The category of information signs is divided into the category of restaurant, the category of parking or the category of post, and is given as $O \triangleright \langle v_{\mathrm{Sig}} \rangle \triangleright \cdots \triangleright \langle v_{\mathrm{RoS}} \rangle \triangleright \langle v_{\mathrm{Inf}} \rangle \triangleright \langle v_{\mathrm{Par}}, v_{\mathrm{Pos}}, v_{\mathrm{Res}} \rangle$. Figure 4.46(a–l) shows examples of members of the category of information signs.

Trademark is any visible sign or device used by a business enterprise to identify its goods and distinguish them from other business enterprises. Trademarks may be words or groups of words, letters, numerals, devices, names, or the object of the different shape. The trademark category refers to a modern trademark that interprets the character of its wearer by associating it with sharply defined signs. Modern trademarks are characteristic symbols of the company. They can resemble real-world object or can have any 'abstract' shape (see Fig. 4.47). The member of the category of symbolic signs has meaning independent from other members of this category. For example, trademark has the meaning that does not depend on the meaning or concurrence others trademarks. The category of trademarks is divided into the category of editorial trademarks or the category of the industrial trademarks and is given as: $O \triangleright \langle v_{\mathrm{Sig}} \rangle \triangleright \langle v_{\mathrm{2DSig}} \rangle \triangleright \langle v_{\mathrm{SymS}} \rangle \triangleright \langle v_{\mathrm{TrS}} \rangle \triangleright \langle v_{\mathrm{PrCo}}, v_{\mathrm{Ind}} \rangle$. The specific categories of the category of editorial trademarks such as the category of Elsevier trademarks, the category of Prentice-Hall trademarks or the

Fig. 4.47. Examples of modern trademarks

category of Springer-Verlag trademarks are derived from the category of editorial trademarks, and are given as follows:

$$O \triangleright \langle v_{\mathrm{Sig}} \rangle \triangleright \cdots \triangleright \langle v_{\mathrm{TrS}} \rangle \triangleright \langle v_{\mathrm{PrCo}} \rangle \triangleright \langle v_{\mathrm{Els}}, v_{\mathrm{PrH}} \rangle.$$

The cross is a sign both of Christ himself and of the faith of Christians. The cross became the principal symbol of Christianity. More than 50 variants were to develop, but the most important are the Greek cross, with its equilateral arms, and the Latin cross, with a vertical arm traversed near the top by a shorter horizontal arm. Other major shapes include the diagonal, or x-shaped, cross of Saint Andrew, and the cross paty (or patee), in which the arms widen at the extremities. A variant of the cross paty is the Maltese cross, which has eight points. The Chi–Rho is a cross formed by joining the first two letters of the Greek word for 'Christ'. The Celtic (Iona) cross is distinguished by a circle surrounding the point of crossing. Two graduated crossbars indicate the Lorraine cross, whereas the Papal cross has three graduated crossbars. A commonly used Eastern Orthodox variant of the cross of Lorraine has an additional crossbar diagonally placed near the base. The category of the cross sign is based on the existing knowledge of the different types of the cross. The proposed categories are based on the visual aspect of the visual object that is interpreted as the cross. The category of cross is divided into: the Latin cross category v_{Lat}, the Saint Andrew cross category v_{X}, the Paty cross category v_{Pat}, the Papal cross category v_{Pap}, the Lorraine cross category v_{Lor}, the Maltese cross category v_{Mal}, the Chi–Rho cross category v_{ChR} or the Celtic (Iona) cross category v_{Cel} and is given by the following categorical chain:

$$O \triangleright \langle v_{\mathrm{Sig}} \rangle \triangleright \langle v_{\mathrm{2DSig}} \rangle \triangleright \langle v_{\mathrm{SymS}} \rangle \triangleright \langle v_{\mathrm{CroS}} \rangle \triangleright \langle v_{\mathrm{Lat}}, v_{\mathrm{X}}, v_{\mathrm{Pat}}, v_{\mathrm{Pap}}, v_{\mathrm{Lor}}, v_{\mathrm{Mal}}, v_{\mathrm{Cel}}, v_{\mathrm{ChR}} \rangle.$$

Figure 4.48 shows examples of members of the different categories of the cross.

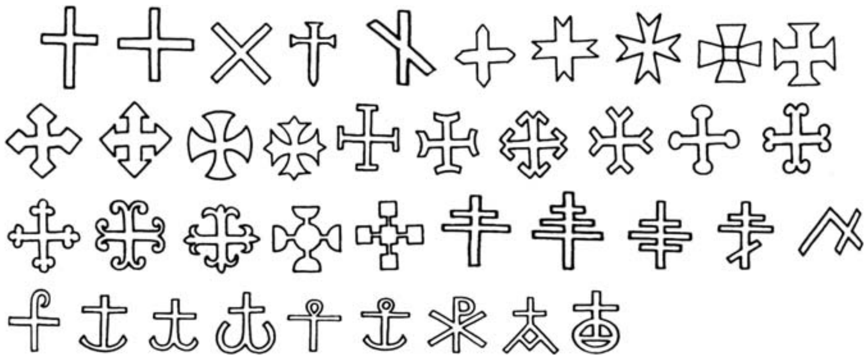


Fig. 4.48. Examples of members of the different categories of the cross

4.2.3.5. Category of Letters

The letter, in this book, denotes any written symbol that is part of the script of any language. Writing is form of human communication by means of a set of visible marks that are related, by convention, to some particular structural level of language. Language is a system of conventional spoken or written symbols by means of which human beings communicate. Written symbols are used to aid our own thought and to communicate with other people. Writing has a long history. Early Chinese ideograms are similar to Egyptian hieroglyphics. The first full language was developed in the fourth millennium BC by the Sumerians. This is the cuneiform script that evolved from picture-writing employing some 900 different symbols. These picture symbols gradually become simpler, until the original picture symbols can only just be distinguished. The symbols were rotated and these rotated pictograms become the cuneiform characters of the first written language. It was found that writing could be done far more efficiently by changing the earlier system of writing from right to left in columns to writing from left to right in lines and by moving the angle of the signs 90°, so that the sign pictures then appeared horizontally instead of vertically. Figure 4.49 shows examples of evolution of the pictograms – the early script. Figure 4.50 shows examples of hieroglyphs, early Sumerian script and cuneiform script. Figure 4.51 shows examples of the Chinese and Japanese script. The category of pictograms and the category of hieroglyph are derived from category of logographic languages.

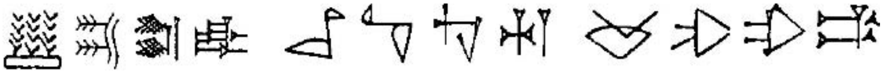


Fig. 4.49. Examples of evolution of the pictograms

a b c

Fig. 4.50. Examples of (**a**) hieroglyphs, (**b**) early Sumerian script, (**c**) cuneiform script

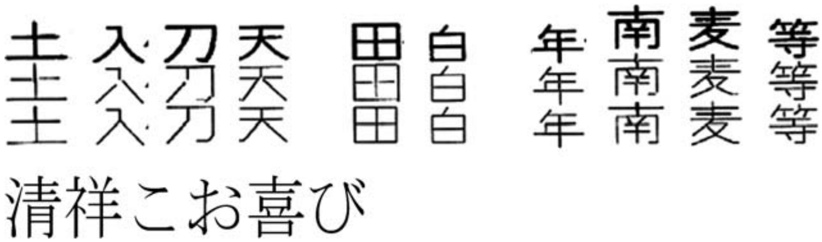


Fig. 4.51. Examples of Chinese and Japanese letters and their distorted versions

The category of letters is derived from the category of the visual objects $O \triangleright \langle v_{\mathrm{ReO}}, v_{\mathrm{ImO}}, v_{\mathrm{Sig}}, v_{\mathrm{Let}}, v_{\mathrm{Fig}} \rangle$. The category of letters is very closely related to the category of languages. The category of languages is divided into the specific categories of the different languages, e.g. the category of the Polish language or the category of the English language.

In most cases the letter is used as a part of the bigger unit such as a word, a sentence or a text. The rules of constructing words and sentences are part of the knowledge of any particular language. These problems will be discussed in Chap. 5. In this chapter the specific categories derived from the category of letters are described. To represent language adequately, a full writing system must maintain fixed correspondences between its visual symbols and the elements of the language. A writing system that has a visual symbol for each word in the language is called logographic, a writing system that has visual symbols for different syllables that occur is called syllabic, and a writing system that has visual symbols for each sound of the language is called alphabetic. The specific categories of letters are based on different types of writing systems. The category of letter is divided into the category of logographic letters v_{Log}, the category of syllabic letters v_{Syl}, and the category of alphabetic letters v_{Alp} and is represented by the categorical chain as $O \triangleright \langle v_{\mathrm{Let}} \rangle \triangleright \langle v_{\mathrm{Log}}, v_{\mathrm{Syl}}, v_{\mathrm{Alp}} \rangle$. The category of alphabetic letter is divided into the category of Latin letters, the category of Greek letters, the category of the Cyrillic letters, the category of Hebrew letters or the category of Arabic letters and is given by the following categorical chain: $O \triangleright \langle v_{\mathrm{Let}} \rangle \triangleright \langle v_{\mathrm{Alp}} \rangle \triangleright \langle v_{\mathrm{Lat}}, v_{\mathrm{Gre}}, v_{\mathrm{Cyr}}, v_{\mathrm{Heb}}, v_{\mathrm{Ara}} \rangle$. Examples of members of the specific categories derived from the category of alphabetic letters such as the category of Latin letters (c), the category of Greek letters (e), the category of the Cyrillic letters (d), the category of Hebrew letters (b) or the category of Arabic letters (a) are shown in Fig. 4.52.

The letter appearance differs depending on type of letters: uppercases or lowercases. In some alphabet the uppercase and lowercase of the same letter looks very differently. For this reason the category such as the

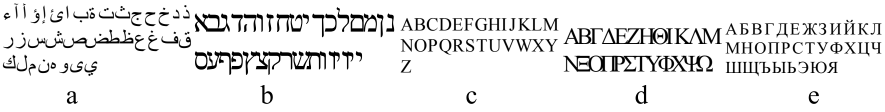


Fig. 4.52. Examples of members of the specific categories derived from the category of alphabetic letters

α β γ δ ε ζ η θ ι κ λ μ ν ξ ο π
ρ ς σ τ υ φ χ ψ ω

Fig. 4.53. Example of the category of the lowercase Greek letters

category of Latin letters is divided into the category of uppercase letters ν_{UppC} and the category of the lowercase letters ν_{LowC} and is given as $O \rhd \langle \nu_{\mathrm{Let}} \rangle \rhd \langle \nu_{\mathrm{Alp}} \rangle \rhd \langle \nu_{\mathrm{Lat}} \rangle \rhd \langle \nu_{\mathrm{LowC}}, \nu_{\mathrm{UppC}} \rangle$. Figure 4.53 shows example of the category of the lowercase Greek letters.

The shape of letter can have the big different appearance depending on the selected category of font. Important details in a definition of prototype of the category of font are appearance of the ending of a stroke. The category of letters (upper cases and lower cases) is divided into category of printed letters ν_{PrF} and handwritten letters ν_{HwL}. In technical drawings freestanding lines are marked off with small cross strokes. Almost all techniques of writing and drawing have produced their own different ways of marking line endings. The category of font describes the lowest category called the prototype. The prototype has a well 'defined' shape. For example, the prototype of the letter font such as an Arial font or a Times New Roman font has well-defined shapes. The category of font such as the category of Latin lowercase letter font is divided into the category of the Times New Roman font ν_{TNR}, the category of the Arial font ν_{Ar}, or the category of the Bold font ν_{Bo}. The categorical chain of the specific categories of the letter font is as $\langle \Pi \rangle \rhd \langle \sigma_{\mathrm{El}} \rangle \rhd \langle \nu_{\mathrm{Let}} \rangle \rhd \langle \nu_{\mathrm{Alp}} \rangle \rhd \langle \nu_{\mathrm{Lat}} \rangle \rhd \langle \nu_{\mathrm{LowC}}, \nu_{\mathrm{UppC}} \rangle \rhd \langle \nu_{\mathrm{PrF}} \rangle \rhd \langle \nu_{\mathrm{Ar}}, \nu_{\mathrm{TNR}}, \nu_{\mathrm{Bo}}, \ldots \rangle$. Examples of members of the different categories of fonts of the letter 'T' are shown in Fig. 4.54. In the case of the category of handwritten letters there is a big diversity among shapes of member of the selected specific category, e.g. the category of handwritten letter 'r'.

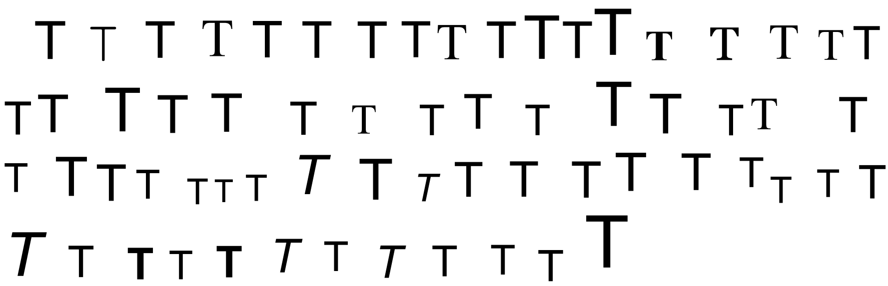


Fig. 4.54. Examples of members of the different categories of fonts of the the letter 'T'

The letter is composed into the bigger units such as words, sentences, paragraphs or text. The visual objects such as words or sentences are members of the category of words ν_{Wor}, the category of phrases ν_{Phr}, the category of sentences ν_{Sen}, or the category of texts ν_{Txt} that are derived from pattern category: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pt}} \rangle \triangleright \langle \nu_{\mathrm{Let}} \rangle \triangleright \langle \nu_{\mathrm{Alp}} \rangle \triangleright \langle \nu_{\mathrm{Lat}} \rangle \triangleright \langle \nu_{\mathrm{PrF}} \rangle \triangleright \langle \nu_{\mathrm{Wor}}, \nu_{\mathrm{Phr}}, \nu_{\mathrm{Sen}}, \nu_{\mathrm{Txt}} \rangle$. The knowledge schema of knowledge categories supplies knowledge that is needed during reading and understanding of the text. A phrase consists of one or more adjacent words. Phrases have names that reflect the type of word they contain for example, the noun phrase contains nouns and the verb phrase contains verbs. The sentence can be dissected into its component phrases, and those phrases into their component words. The analysis of sentences by application of the grammar rules is the task of syntactic analysis. A grammar represents the syntactic rules of the language that are learned as a part of the knowledge schema of the knowledge object. The phrase category is divided into the noun phrase category ν_{Nou} or the verb phrase category ν_{Ver} and is represented by categorical chain as $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pt}} \rangle \triangleright \langle \nu_{\mathrm{Let}} \rangle \triangleright \cdots \triangleright \langle \nu_{\mathrm{Phr}} \rangle \triangleright \langle \nu_{\mathrm{Nou}}, \nu_{\mathrm{Ver}} \rangle$.

4.2.3.6. Category of Real-World Objects

The category of the real-world objects refers to the 3D objects that exist in the real world and can be perceived through accessible technical tools such as a camera, a telescope or a microscope. The perceived object (phantom) that refers to the real-world object is given as one of the perceptual categories such as a silhouette, a line drawing, a colour object or a shaded-

Fig. 4.55. Examples of members of the different perceptual categories and structural categories

object $\nu_{\mathrm{O}} \rhd \langle \pi_{\mathrm{Si}}, \pi_{\mathrm{Ld}}, \pi_{\mathrm{Co}}, \pi_{\mathrm{E}} \rangle$. Examples of members of the different perceptual categories and structural categories are shown in Fig. 4.55. At the first stage of the visual interpretation, the object is assigned to one of the perceptual categories, next it is assigned to one of the structural categories $\nu_{\mathrm{O}} \rhd \langle \pi \rangle \rhd \langle \sigma_{\mathrm{El}}, \sigma_{\mathrm{Pt}}, \sigma_{\mathrm{Pi}}, \sigma_{\mathrm{An}} \rangle$ and, at the end, it is interpreted as a member of one of the ontological categories given by the categorical chain as: $\nu_{\mathrm{O}} \rhd \langle \Pi \rangle \rhd \langle \Lambda \rangle \rhd \langle \nu_{\mathrm{ReO}}, \nu_{\mathrm{ImO}}, \nu_{\mathrm{Sig}}, \nu_{\mathrm{Let}}, \nu_{\mathrm{Fig}} \rangle$.

The category of the real-world objects is the category of the third level of the categorical chain. The category of real-world objects can be a member of many different perceptual or structural categories. If these categories are not specified the beginning of the categorical chain is denoted by the symbol O. In the case when these categories are specified, symbols of these categories are shown in the categorical chain. The simplest perceptual category of real-world objects is a silhouette. The silhouette often represents the isolated real-world object. The categorical chain for the real word object represented by silhouette is given as $\nu_{\mathrm{O}} \rhd \langle \pi_{\mathrm{Si}} \rangle \rhd \langle \sigma_{\mathrm{El}} \rangle \rhd \langle \nu_{\mathrm{ReO}} \rangle$. Silhouette can be obtained from the line drawing or shaded object. The silhouette can be obtained from a photograph of a real-world object or can be obtained by process of schematization and visual abstraction. Figure 4.56 shows the silhouettes of selected animals obtained as the result of schematization.

The line drawing can be obtained by proper segmentation of the image. The silhouette obtained from a line drawing or a shaded object conveys only part of the visual information needed for the interpretation of the object. During the visual interpretation a silhouette is used to obtain the

Fig. 4.56. The silhouettes obtained as the result of schematization

interpretational sketch of the object (see Fig. 4.57). The category of silhouette of a given object, e.g. a glass consists of different shapes of the glass that is the result of variation of shapes across the different types of glasses as well as the different views of glasses. There is a big difference among shapes of objects from the different ontological categories and usually a big difference among a set of characteristic views of the same object. For simple objects from categories such as the glass category the number of characteristic views is rather small.

The visual representation of the human figure, as shown in Fig. 4.58 can be the result of schematization and shows the difference in gender and pose.

The real-world objects can be represented visually by the line drawing. The line drawing can be obtained by proper segmentation of the photograph of the object or can be obtained as the result of schematization and visual abstraction. For example, the schematic line drawing of the different category of birds is shown in Fig. 4.59.

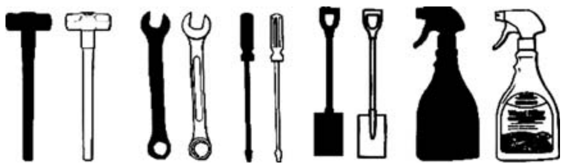


Fig. 4.57. Examples of silhouettes obtained from the line drawing

Fig. 4.58. The different visual representations (silhouettes) of the human figure

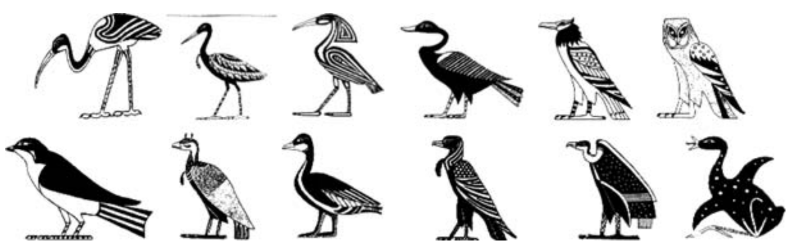


Fig. 4.59. The different schematic representations of birds

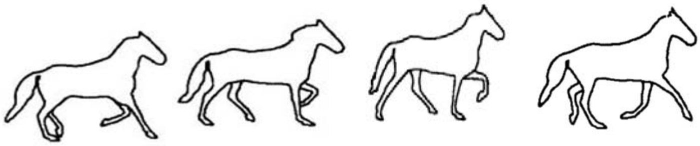


Fig. 4.60. The changes of the visual representation of the horse when the movement of the parts occurs

The living object has its part moving and by this the visual representation needs to capture not only the changes of the visual aspect of the object but also the changes of the object when the parts are moving. Figure 4.60 shows the changes of the visual representation of the horse when the movement of the parts occurs.

Real-world object is represented by 2D object that is obtained as the result of the conventional schematization or as the result of application of the projection method. Projections method is basic tools of the engineering and technical graphics. Projection theory comprises the principles used to represent graphically 3D objects on 2D media. Drawing more than one face of an object by rotating the object relative to the line of sight helps in understanding the 3D form. A line of sight is an imaginary ray of light between an observer's eye and an object. In perspective projection, all lines of sight start at a single point. In parallel projection, all lines of sight are parallel. Orthographic projection is a parallel projection technique in which the plane of projection is positioned between the observer and the object and is perpendicular to the parallel lines of sight. Multi-view projection is an orthographic projection for which the object is behind the plane of projection, and the object is oriented such that only two of its dimensions are shown. Multi-view drawing employs multi-view projection techniques. In multi-view drawings, generally three views of an object are drawn. The perspective of central projection is divided into linear perspective and aerial perspective. The parallel projection is divided into oblique projection and orthographic projection. The orthographic projection is divided into the axonometric projection and multi-view projection. Figure 4.61 shows examples of different projections.

To solve problems of perspective distortion the perceptual category of the line drawing is divided into the different specific categories such as: the category of structural archetype ν_{StAr}, the category of the segmentation edge ν_{SeE}, the category of the conventional 3D drawing ν_{CoD}, the category of the intentional geometrical drawing ν_{InG}, the category of the multi-view drawing ν_{MuV}, the category of view from the top ν_{ViT}, the category of frontal view ν_{ViF}, the category of

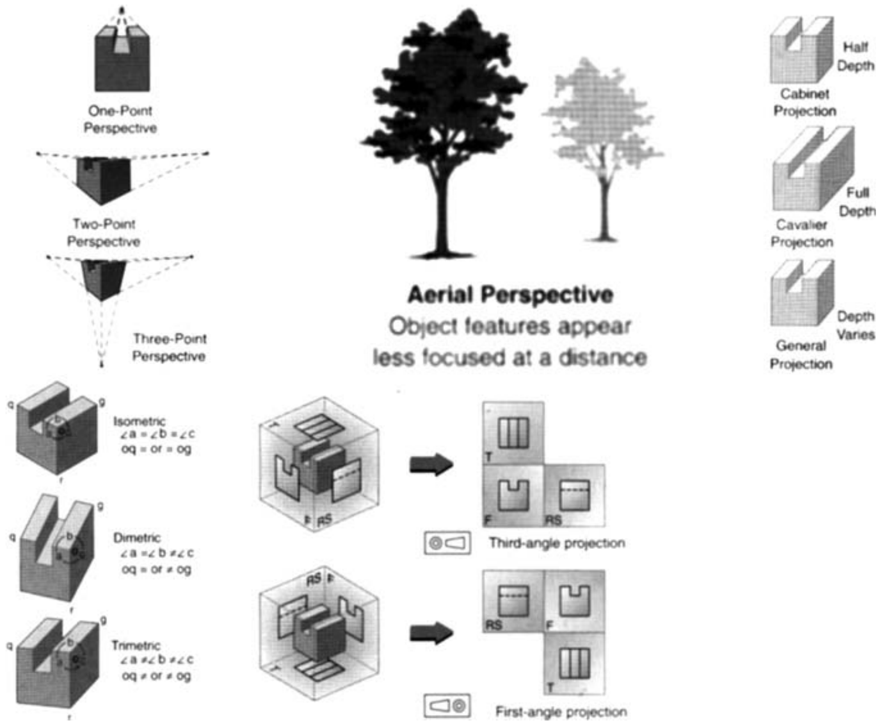


Fig. 4.61. The different projections

orthographic projection v_{OrP}, the category of perspective projection v_{PeP}, the category of folding sheets v_{FoSt}, the category of many aspects drawing v_{MaAs} and is represented as $v_{\mathrm{O}} \rhd \langle \pi_{\mathrm{Ld}} \rangle \rhd \langle \sigma_{\mathrm{El}} \rangle \rhd \langle v_{\mathrm{ReO}} \rangle \Rightarrow || v_{\mathrm{StAr}}, v_{\mathrm{SeE}}, v_{\mathrm{InG}}, v_{\mathrm{MuV}}, v_{\mathrm{ViT}}, v_{\mathrm{ViF}}, v_{\mathrm{OrP}}, v_{\mathrm{PeP}}, v_{\mathrm{MaAs}}, v_{\mathrm{FoSt}} ||$. These subcategories of perceptual categories of line drawing object make it possible to interpret the visual object in terms of perceptual distortions caused by application of the different methods to generate the visual object (e.g. projection methods). Figure 4.62 shows members of the different specific line drawing categories such as: the category of the segmentation edge (a), the category of the conventional 3D drawing (b), the category of the intentional geometrical drawing (c), the category of the multi-view drawing (d), the category of view from the top (e,f), the category of frontal view (g,h), the category of orthographic projection (i), the category of perspective projection (j), the category of folding sheet (k), the category of many aspects drawing (l).

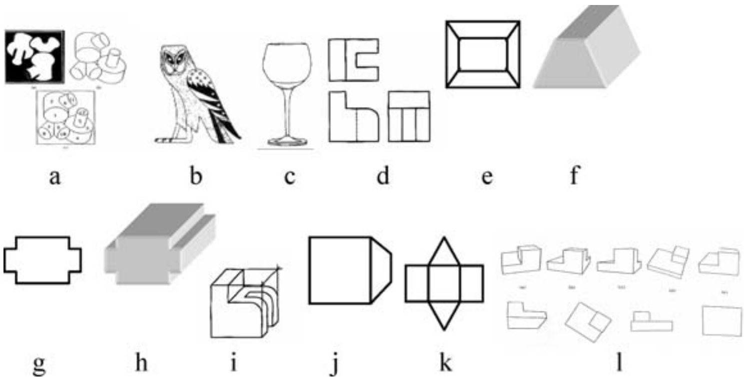


Fig. 4.62. Examples of members of the different specific line drawing categories

The category of the real-world object refers to the 3D geometrical figure. The real-world object is perceived as the three-dimensional object and as a result of abstraction the real-world object can be interpreted as the 3D geometrical figures called solids. Selected specific categories derived from the category of solids were described in the previous section. Each solid category can be represented by the line drawing category. There are many different 2D representations of the same solid that show the different visual aspects of the solid. To be able to interpret the different visual aspects of the object as the same object, the visual schema or the generic visual concept based on multi-view representation is introduced. The category of multi-view drawing (the perceptual category) is denoted by upper subscript $\langle \pi_{\mathrm{Ld}}^{\mathrm{M}} \rangle$ where M denotes the multi-view representation and Ld denotes the category of the line drawing: $\nu_{\mathrm{O}} \triangleright \langle \pi_{\mathrm{Ld}}^{\mathrm{M}} \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{Fig}} \rangle \triangleright \langle \nu_{\mathrm{3DF}} \rangle \triangleright \langle \nu_{\mathrm{Pol}} \rangle \triangleright \langle \nu_{\mathrm{NPol}} \rangle \triangleright \langle \nu_{\mathrm{Cub}}, \nu_{\mathrm{Pri}} \rangle$. The multi-view drawing is used to learn the generic visual concept of the object that can be used in performing of the mental transformation.

The visual schema of the knowledge category includes the visual concept that is obtained during learning process. The generic visual concept that is learned based on the multi-view representation consists of three visual names. For example, the generic visual concept of the cylinder (Fig. 4.63a) consists of symbolic names of the circle and two rectangles $\varphi_{\mathrm{Cylinder}}^{\mathrm{MV}} = \{K_{\mathrm{C}}^{1}, L_{\mathrm{R}}^{4}, L_{\mathrm{R}}^{4}\}$. The top view of the cylinder shown in Fig. 4.63a is represented by the circle whereas the frontal view is represented by two rectangles. The orthographic projection of the cylinder is given as $\varphi_{\mathrm{Cylinder}}^{\mathrm{OP}} = \{\rho[M^{2}[L_{\mathrm{R}}^{4}]]\{K_{\mathrm{E}}^{1}, Q^{1}[M^{1}[L_{\mathrm{R}}^{4}]](M_{\mathrm{E}}^{1})\}\}$. Figure 4.63b,c shows geometrical projections of the selected object. The selected issues of the learning of the visual concept of the real-world objects will be given in Chap. 5.

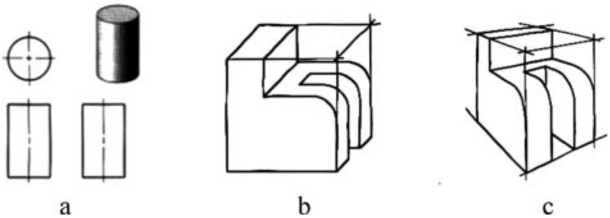


Fig. 4.63. The different categories of the perceptual category (**a**), the multi-view drawing (**b**), geometrical projections (**c**)

The category of the real-world object is represented by the different visual aspects of the object. Only some of the characteristic aspects of the object can be used as representatives of the object. Figure 4.64 shows example of different visual aspects of the object. As we can see only some of them can supply the visual information that makes it possible to identify the object uniquely.

In some tests or problems that need to be solved the object is represented as an object seen from the top. The object seen from the top is called the bird's eye view. The member of this category can be represented by the shaded object and its visual drawing line representation. The visual concept of the member of the category of view from the top includes the

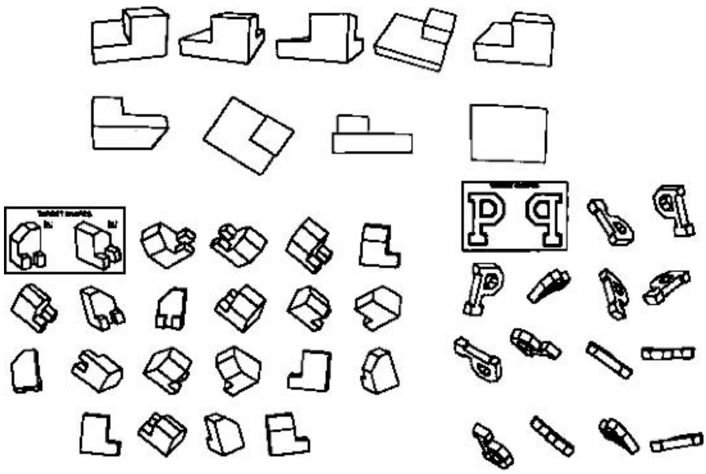


Fig. 4.64. Examples of the different visual aspects of the object

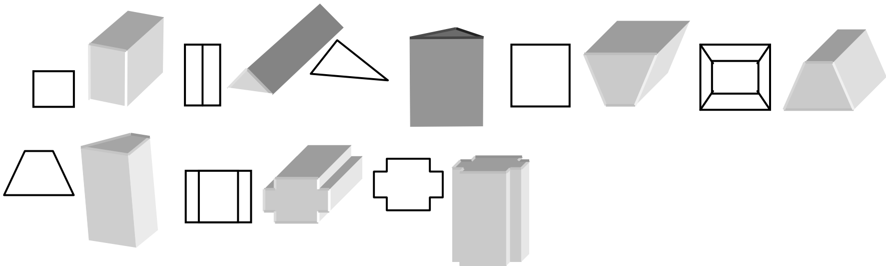


Fig. 4.65. Objects seen from the top

symbolic name of the object seen from the top. Examples of members of the category of view from the top are shown in Fig. 4.65.

The 3D solids can be produced by folding flat sheet. The shape of the flat sheet from which 3D solid is produced is part of the visual knowledge about the 3D objects. The knowledge schema of the category of the folded sheet consists of visual concept of the sheet, the visual concept of 3D solid and the name of the solid. Examples of members of the category of folding sheets are shown in Fig. 4.66.

The existing real-world objects have different size. Based on the size of the objects the category of the real-world objects is divided into the category of micro-objects, the category of macro-objects and the category of earthy-world objects: $O \triangleright \langle v_{\mathrm{ReO}} \rangle \triangleright \langle v_{\mathrm{Mic}}, v_{\mathrm{Mac}}, v_{\mathrm{Ear}} \rangle$. Those categories are described in more detail in the following sections.

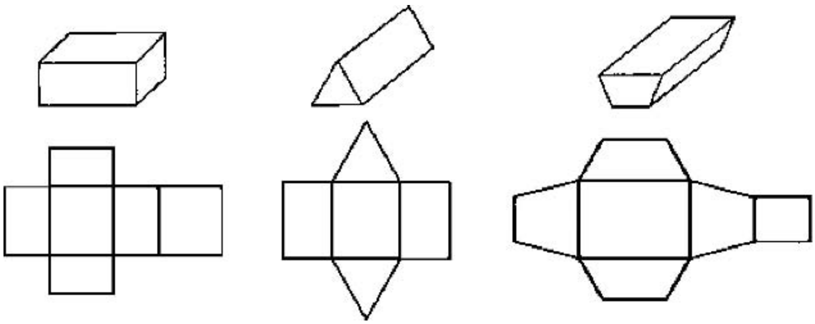


Fig. 4.66. Examples of members of the category of folding sheets

4.2.3.6.1. Categories of Macro- and Micro-Objects

The category of macro-objects includes objects, size of which is bigger than objects of our today's experience. These objects can be seen only by applying special tools such as a telescope. The category of macro-objects ν_{Mac} is divided into the universe category ν_{Uni}, the galaxy category ν_{Gal}, the star category ν_{Str}, the solar system category ν_{SolS}, the moon category ν_{Mon}, the asteroids category ν_{Ast} or the comet category ν_{Com} and is given by the following categorical chain: $O \triangleright \langle \nu_{\text{ReO}} \rangle \triangleright \langle \nu_{\text{Mac}} \rangle \triangleright \langle \nu_{\text{Uni}}, \nu_{\text{Gal}}, \nu_{\text{Str}}, \nu_{\text{SolS}}, \nu_{\text{Ast}}, \nu_{\text{Mon}}, \nu_{\text{Com}} \rangle$. Galaxy is any system of stars and interstellar matter that makes up the Universe. The galaxy category is divided into the irregular galaxy category ν_{Irr}, the elliptic galaxy category ν_{Ell} or the spiral galaxy category ν_{Spi} and is given as $O \triangleright \langle \nu_{\text{ReO}} \rangle \triangleright \langle \nu_{\text{Mac}} \rangle \triangleright \langle \nu_{\text{Gal}} \rangle \triangleright \langle \nu_{\text{Irr}}, \nu_{\text{Ell}}, \nu_{\text{Spi}} \rangle$. The stars category is divided into the nebula category ν_{Neb}, the supernova category ν_{Sup} or the black hole category ν_{BlH} and is given as: $O \triangleright \langle \nu_{\text{ReO}} \rangle \triangleright \langle \nu_{\text{Mac}} \rangle \triangleright \langle \nu_{\text{Str}} \rangle \triangleright \langle \nu_{\text{Neb}}, \nu_{\text{Sup}}, \nu_{\text{BlH}} \rangle$. The knowledge about the invisible aspects of the category of macro-objects is given by the category of astronomical objects. The knowledge schema of the specific categories of astronomical objects supplies knowledge about properties of members of macro-objects categories. The category of astronomical objects that is derived from knowledge object, supplies non-visible knowledge that is used in interpretation of the macro-objects. Knowledge of the category of astronomical objects is based on knowledge of the scientific branch of science called astronomy. Astronomy is science that deals with the origin, evolution, composition, distance and motion of all bodies and scattered matter in the universe. Examples of members of the category of macro-objects are shown in Fig. 4.67.

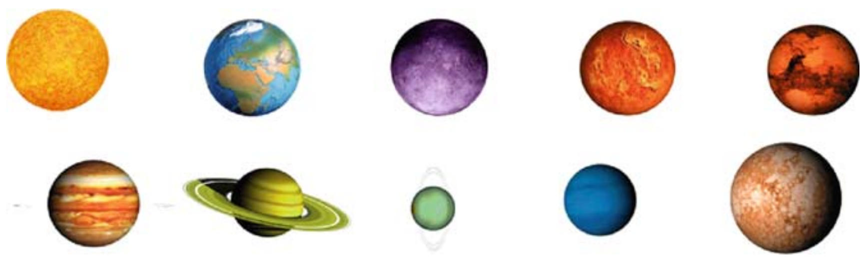


Fig. 4.67. Examples of members of the category of macro-object

Objects that are too small to be seen by naked eye belong to the category of micro-objects. Members of the category of micro-objects can be registered by applying the one of the methods of photomicrography. Photomicrography is the photography of objects under a microscope. Microscope is an instrument that produces enlarged images of small objects, allowing them to be viewed at a scale convenient for examination and analysis. For example, metallographic microscopes are used to identify defects in metal surfaces and to determine the crystal grain boundaries in metal alloys, whereas the electron-probe micro-analyser allows a chemical analysis of the composition of materials to be made by using the incident electron beam to excite the emission of characteristic X radiation by the various elements composing the specimen. These X-rays are detected and analysed by spectrometers built into the instrument. Such probe micro-analysers are able to produce an electron scanning image so that structure and composition may be easily correlated. Electron microscope is a microscope that attains extremely high resolution using an electron beam instead of a beam of light to illuminate the object of study. Modern electron microscopes provide detailed images at magnifications of more than 250,000. The category of micro-world is divided into the category of living and the category of non-living object. The category of living objects is divided into the category micro-organs ν_{Org}, or the category of micro-organisms (bacteria, viruses) ν_{Ogn} and is given by the following categorical chain: $O \triangleright \langle \nu_{\text{ReO}} \rangle \triangleright \langle \nu_{\text{Mic}} \rangle \triangleright \langle \nu_{\text{Liv}} \rangle \triangleright \langle \nu_{\text{Org}}, \nu_{\text{Ogn}} \rangle$. The category of non-living objects is represented by the category of micro-particles or the category of micro-dust ν_{Dus} and is given as $\ldots \triangleright \langle \nu_{\text{NLiv}} \rangle \triangleright \langle \nu_{\text{Dus}}, \nu_{\text{Par}} \rangle$. Examples of members of the micro-objects category are shown in Fig. 4.68.

4.2.3.6.2. Category of Earthy-World Objects

The visual objects, size of which is such that can be visible by the naked eye, are objects of the earthy-world objects category. The earthy-world objects categories are divided into the category of living objects ν_{Liv} and the category of non-living objects ν_{NLiv} and are given by the following

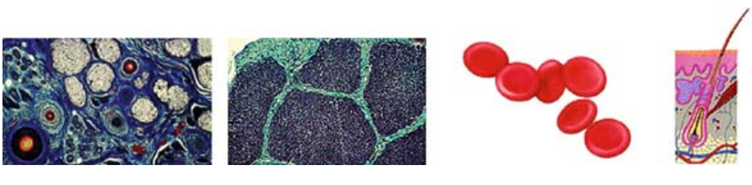


Fig. 4.68. Examples of objects of the micro-world category

categorical chain: $O \rhd \langle \nu_{\mathrm{ReO}} \rangle \rhd \langle \nu_{\mathrm{Ert}} \rangle \rhd \langle \nu_{\mathrm{Liv}}, \nu_{\mathrm{NLiv}} \rangle$. Category of non-living objects ν_{NLiv} includes category of the man-made objects ν_{MMan} and the category of non-man-made natural objects ν_{NMan}.

Category of Non-Living Objects

The non-living objects such as minerals or rocks are objects of study of scientific disciplines called geology. Geology is branch of science that studies of the Earth, including its composition, structure, physical properties and history. Geology is divided into a number of sub-disciplines. The disciplines concerned with non-living objects are mineralogy (the study of minerals), petrology (study of rocks), geomorphology (study of landforms), palaeontology (study of fossils). The category of natural objects (non-man-made) includes the category of minerals and the category of rocks and is given by the following categorical chain: $\langle \Pi \rangle \rhd \langle \sigma_{\mathrm{El}} \rangle \rhd \langle \nu_{\mathrm{ReO}} \rangle \rhd \langle \nu_{\mathrm{Ear}} \rangle \rhd \langle \nu_{\mathrm{NLiv}} \rangle \rhd \langle \nu_{\mathrm{NMan}} \rangle \rhd \langle \nu_{\mathrm{Min}}, \nu_{\mathrm{Rock}} \rangle$.

Category of non-living non-man-made natural objects

The category of non-man-made objects includes the category of minerals and the category of rocks. Mineral is any naturally occurring homogeneous solid that has a definite chemical composition and a distinctive internal crystal structure. Although minerals are usually formed by inorganic processes some synthetic equivalents of various minerals, such as emeralds and diamonds, can be produced in the laboratory. Most minerals are chemical compounds and only a small number of minerals (e.g. sulphur, copper, gold) are chemical elements. Minerals are classified into groups based on the identity of its anionic group and the composition of a mineral can be defined by its chemical formula. A mineral is considered to be a crystalline material because it crystallizes in an orderly, three-dimensional geometric form. The crystalline structure of a mineral determines such physical properties as hardness, colour and cleavage. Minerals are the materials that make up the rocks of the Earth crust. The category of minerals is divided into the native elements category, the sulphides category, the sulfosalts category, the oxides category, the hydroxides category, the halides category, the carbonates category, the nitrates category, the borates category, the sulphates category, the phosphates category, and the silicates category and is given by the categorical chain as $\ldots \rhd \langle \nu_{\mathrm{Min}} \rangle \rhd \langle \nu_{\mathrm{Nel}}, \nu_{\mathrm{Sul}}, \nu_{\mathrm{Sfo}}, \nu_{\mathrm{Oxi}}, \nu_{\mathrm{Hyd}}, \nu_{\mathrm{Hal}}, \nu_{\mathrm{Sfo}}, \nu_{\mathrm{Car}} \rangle$.

Rock is naturally occurring and coherent aggregate of one or more minerals and constitutes the basic unit of which the solid Earth is comprised. Rocks are divided into three major classes according to the processes that resulted in their formation igneous rocks, sedimentary and metamorphic. Igneous rocks are rocks which have solidified from molten material called magma. Sedimentary rocks are those consisting of fragments derived from pre-existing rocks or of materials precipitated from solutions. Metamorphic rocks are rocks which have been derived from either igneous or sedimentary rocks under conditions that caused changes in mineralogical composition, texture and internal structure. These three classes, in turn, are subdivided into numerous groups and types on the basis of various factors, the most important of which are chemical, mineralogical and textural attributes. Based on the rock classification the following specific categories are derived from the rock category: the category of sedimentary rocks ν_{Sed}, the category of igneous rocks ν_{Ign} and the category of metamorphic rocks ν_{Met} given as: $\ldots \triangleright \langle \nu_{\mathrm{Rock}} \rangle \triangleright \langle \nu_{\mathrm{Sed}}, \nu_{\mathrm{Ign}}, \nu_{\mathrm{Met}} \rangle$. Sedimentary rocks are made of particles of sediments such as sand or clay. The category of sedimentary rocks is divided into the flint category ν_{Fli}, the chalk category ν_{Cha}, the limestone category ν_{Lim} or the sandstone category ν_{San} and is given by the following categorical chain: $\ldots \triangleright \langle \nu_{\mathrm{Rock}} \rangle \triangleright \langle \nu_{\mathrm{Sed}} \rangle \triangleright \langle \nu_{\mathrm{Fli}}, \nu_{\mathrm{Cha}}, \nu_{\mathrm{Lim}}, \nu_{\mathrm{San}} \rangle$. Igneous rocks are created when magma cools and become solid. The category of igneous rocks are divided into the obsidian category ν_{Obs}, the granite category ν_{Gra}, the porphyrite category ν_{Por} or the gabbro category ν_{Gab} and is given by the following categorical chain: $\ldots \triangleright \langle \nu_{\mathrm{Rock}} \rangle \triangleright \langle \nu_{\mathrm{Ign}} \rangle \triangleright \langle \nu_{\mathrm{Obs}}, \nu_{\mathrm{Gra}}, \nu_{\mathrm{Por}}, \nu_{\mathrm{Gab}} \rangle$. Metamorphic rocks are formed when the minerals in rocks are changed underground by heat and pressure. The category of metamorphic rocks is divided into the slate category ν_{Sla}, the gneiss category ν_{Gne}, the marble category ν_{Mar} and is given as: $\ldots \triangleright \langle \nu_{\mathrm{Rock}} \rangle \triangleright \langle \nu_{\mathrm{Met}} \rangle \triangleright \langle \nu_{\mathrm{Gne}}, \nu_{\mathrm{Mar}}, \nu_{\mathrm{Sla}} \rangle$. The non-visual knowledge that is needed in the process of interpretation of the visual objects that are members of the category of minerals or the category of rocks is derived from the knowledge object $\kappa_{\mathrm{KB}} \triangleright \langle \kappa_{\mathrm{KOb}} \rangle \triangleright \langle \kappa_{\mathrm{GeOb}} \rangle \triangleright \langle \nu_{\mathrm{Min}}, \nu_{\mathrm{Rock}} \rangle$. Figure 4.69 shows examples of members of the mineral category and the rock category.

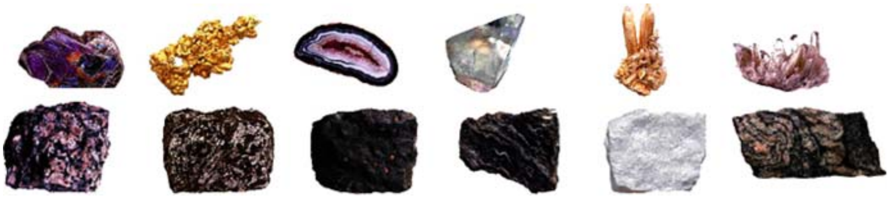


Fig. 4.69. Examples of members of the mineral category and the rock category

Category of non-living man-made objects

The category of non-living man-made objects covers the broad range of objects that are made for different purposes. The category of man-made objects is divided into the category of tools, vehicles, furniture, buildings, arms or machines:

$$\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pat}} \rangle \triangleright \langle \nu_{\mathrm{ReO}} \rangle \triangleright \langle \nu_{\mathrm{Ear}} \rangle \triangleright \langle \nu_{\mathrm{NLiv}} \rangle \triangleright \langle \nu_{\mathrm{MMad}} \rangle \triangleright \langle \nu_{\mathrm{Tol}}, \nu_{\mathrm{Veh}}, \nu_{\mathrm{Fur}}, \nu_{\mathrm{Bul}}, \nu_{\mathrm{Arm}}, \nu_{\mathrm{Mach}} \rangle .$$

The category of man-made objects can be always broadened about a new model that was lately designed. The category of man-made objects needs to refer to the diversity of objects that are made in the different period of time and the diversity of objects that are results of the new design. For example, furniture of ancient Egypt such as beds is different from today's beds.

One of the most popular man-made category is the furniture category. Furniture is household equipment, usually made of wood, metal, plastics, marble, glass, fabrics or related materials and having a variety of different purposes. Furniture ranges widely from the simple pine chest or stickback country chair to the most elaborate marquetry work cabinet or gilded console table. The functional and decorative aspects of furniture have been emphasized more or less throughout history according to economics and fashion. The ideal of furniture design is to integrate utility, craftsmanship and beauty into a harmonious whole. Accessory furnishings are smaller subsidiary items such as clocks, mirrors, tapestries, fireplaces, panelling and other items complementary to an interior scheme. The specific category that is derived from the furniture category is the chair category. Chair is a seat with a back, intended for one person and it is one of the most ancient forms of furniture. The category of chair needs to take into account chairs that were built in the different historical periods. These chairs are regarded as a new subcategory called the category of old chairs. During the learning process the category of chair is learned with reference to visual

Fig. 4.70. Examples of the different categories of chairs represented by the silhouette one of the perceptual category

similarities of members of the different categories. Figure 4.70 shows examples of chairs that are ordered according to the degree of similarity. During learning category of the real-world object such as the category of chair, the members of different perceptual categories such as silhouette or line drawing and the different aspects of the object need to be also taken into account. Furniture is result of the work of the furniture makers. For example chairs are made by carpenter. Figure 4.70 shows the silhouettes of chairs. The silhouette of chair supplies enough visual information to recognize it as the 'chair'.

Another man-made category is the wearing category. Wearing such as clothes are garments for the human body and dress is covering for the human body. The variety of dresses is immense, varying with different genders, cultures, geographic areas and historic eras. The term dress encompasses not only such familiar garments as shirts, skirts, trousers, jackets and coats but also footwear, caps and hats, sleepwear, sports clothes, corsets and gloves. During establishing of the wearing category there is a need to take into account the period of time when the wearing were produced. Taking into account the time when the wearing were produced the category of wearing is divided into the category of old wearings (in the ancient times), the category of middle ages wearings, the category of new wearings (before the second world war) and the category of today's wearings: $\ldots \triangleright \langle \nu_{\text{MMad}} \rangle \triangleright \langle \nu_{\text{Wer}} \rangle \triangleright \langle \nu_{\text{OldW}}, \nu_{\text{MidW}}, \nu_{\text{NewW}}, \nu_{\text{TodW}} \rangle$. The category of wearings is divided into the categories that refer to the part of the human body such as head, hand, legs or arms: $\ldots \triangleright \langle \nu_{\text{MMad}} \rangle \triangleright \langle \nu_{\text{Wer}} \rangle \triangleright \langle \nu_{\text{OldW}}, \nu_{\text{MidW}}, \nu_{\text{NewW}}, \nu_{\text{TodW}} \rangle \triangleright \langle \nu_{\text{Hed}}, \nu_{\text{Hnd}}, \nu_{\text{Leg}}, \nu_{\text{Arm}} \rangle$. The clothes such as hat, pullover, trousers, breeches, plus-fours; knickerbockers are examples of the wearing category. Examples of members of the category of wearing are shown in Fig. 4.71.

Fig. 4.71. Examples of members of the category of wearing

The machine category covers a big range of man-made objects. Machine is the device, having a unique purpose that augments or replaces human or animal effort for the accomplishment of physical tasks. The category of machines are divided into the category of simple machines (levers, wheels, pulleys, screws and gears), the category of electrical machines or the category of mechanical machines and is given as: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pt}} \rangle \triangleright \cdots \triangleright \langle \nu_{\mathrm{MMad}} \rangle \triangleright \langle \nu_{\mathrm{Mach}} \rangle \triangleright \langle \nu_{\mathrm{SMac}}, \nu_{\mathrm{EMac}}, \nu_{\mathrm{MMac}} \rangle$. The category of mechanical machines is divided into the category of vehicles or the category of non-moving machines. The category of vehicles is divided into the category of air vehicles, the category of space vehicles, the category of water vehicles or the category of land vehicles and is given as: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \cdots \triangleright \langle \nu_{\mathrm{MMad}} \rangle \triangleright \langle \nu_{\mathrm{Veh}} \rangle \triangleright \langle \nu_{\mathrm{AirV}}, \nu_{\mathrm{SpcV}}, \nu_{\mathrm{WatV}}, \nu_{\mathrm{LanV}} \rangle$. Figure 4.72 shows examples of members of the different categories of land vehicles. Figure 4.72a–d shows silhouettes of the members of the car categories. Figure 4.72e,f shows the line drawing of the different visual aspects of the member of the same car category.

In contrast to the category of the visual objects that refers to the object that can be detected by the perceptual devices, the knowledge object can refer to the non-visual object. The non-visual object that is part of the general knowledge about the world can describe the visual object only by indicating the relationships among the visual and non-visual properties of the world. One of the categories derived from the knowledge object is the category of the general concepts such as the work category, the

Fig. 4.72. Examples of the different perceptual categories (the silhouette category and the line drawing category) used to represent members of the ontological category, namely, the vehicle category

Fig. 4.73. The visual representations of the selected categories of work

entertainment category or the sport category and given by the following knowledge chain: $\kappa_{\text{KB}} \rhd \langle \kappa_{\text{KOb}} \rangle \rhd \langle \kappa_{\text{GenC}} \rangle \rhd \langle \kappa_{\text{Work}}, \kappa_{\text{Ent}}, \kappa_{\text{Spr}} \rangle$. The visual object that is a member of the category of the general concept such as a work category can be represented by a member of one of the perceptual categories such as shaded object or coloured object. The member of the work category can refer to the visual aspect of this category showing characteristic parts of the workers dress, characteristic tools or result of his work. One of the visual representatives of the work category is the worker with characteristic tools. Figure 4.73 shows the visual representatives of the selected categories of work.

The work category is derived from the category of general concept and refers to relations amongst a man and other objects. The work category κ_{Wrk} consists of the category of professions or the category of social contexts and can be denoted as $\kappa_{\text{KB}} \rhd \langle \kappa_{\text{KOb}} \rangle \rhd \langle \kappa_{\text{GenC}} \rangle \rhd \langle \kappa_{\text{Wrk}} \rangle \rhd \langle \kappa_{\text{Prf}}, \kappa_{\text{Soc}} \rangle$. The category of professions κ_{Prf} consists of the category of tools κ_{Tol}, the category of materials κ_{Mat}, the category of knowledge κ_{Kno}, the category of results κ_{Res} and is given as: $\kappa_{\text{KB}} \rhd \langle \kappa_{\text{KOb}} \rangle \rhd \langle \kappa_{\text{GenC}} \rangle \rhd \langle \kappa_{\text{Wrk}} \rangle \rhd \langle \kappa_{\text{Prf}} \rangle \propto (\kappa_{\text{Tol}}, \kappa_{\text{Mat}}, \kappa_{\text{Kno}}, \kappa_{\text{Res}})$. These categories are inherited by all categories derived from the category of professions. The category of professions is divided into the builder category, the carpenter category, the mechanic category or the secretary category: $\kappa_{\text{KB}} \rhd \langle \kappa_{\text{KOb}} \rangle \rhd \langle \kappa_{\text{GenC}} \rangle \rhd \langle \kappa_{\text{Wrk}} \rangle \rhd \langle \kappa_{\text{Prf}} \rangle \rhd \langle \kappa_{\text{Mas}}, \kappa_{\text{Car}}, \kappa_{\text{Meh}}, \kappa_{\text{Elc}}, \kappa_{\text{Tay}}, \kappa_{\text{Sec}} \rangle$. The category of professions that is derived from the visual object refers to the visual representation of the man that is a representative of a given profession. The categorical chain is given as $\langle \Pi \rangle \rhd \langle \sigma_{\text{El}} \rangle \rhd \langle \nu_{\text{ReO}} \rangle \rhd \langle \nu_{\text{Ear}} \rangle \rhd \langle \nu_{\text{Liv}} \rangle \rhd \langle \nu_{\text{Man}} \rangle \rhd \langle \nu_{\text{Prf}} \rangle \rhd \langle \nu_{\text{Mas}}, \nu_{\text{Car}}, \nu_{\text{Meh}}, \nu_{\text{Elc}}, \nu_{\text{Tay}}, \nu_{\text{Sec}} \rangle$. As an example of profession category the builder category is given. Builder category is written as: $\ldots \rhd \langle \kappa_{\text{Bul}} \rangle \propto (\kappa_{\text{Tol}}, \kappa_{\text{Mat}}, \kappa_{\text{Kno}}, \kappa_{\text{Res}})$. The tools are represented by the builder-tools category such as the trowel category or the hammer category: $\ldots \rhd \langle \kappa_{\text{Bul}} \rangle \propto (\kappa_{\text{Tol}}) \rhd (\kappa_{\text{Trw}}, \kappa_{\text{Ham}})$.

The material is represented by the builder-material category such as the brick category or the stone category: $\ldots \triangleright \langle \kappa_{\mathrm{Bul}} \rangle \propto (\kappa_{\mathrm{Mat}}) \triangleright (\kappa_{\mathrm{Brc}}, \kappa_{\mathrm{Stn}})$. The results are represented by the builder-result category such as the house category or the bridge category: $\ldots \triangleright \langle \kappa_{\mathrm{Bul}} \rangle \propto (\kappa_{\mathrm{Res}}) \triangleright (\kappa_{\mathrm{Hou}}, \kappa_{\mathrm{Brd}})$. The knowledge category such as the tools category refers to the visual object. Tool is an instrument for making material changes on other objects, as by cutting, shearing, striking, rubbing, grinding, squeezing, measuring or other process. Tools are the primary means by which human beings control and manipulate their physical environment. A hand tool is a small manual instrument traditionally operated by the muscular strength of the user; a machine tool is a power-driven mechanism used to cut, shape or form materials such as wood and metal. Modern hand tools were developed in the period after 1500 BC and are now considered in the following classes: percussive tools, which deliver blows (the axe, adze and hammer); cutting, drilling and abrading tools (the knife, awl, drill, saw, file, chisel and plane); the screw-based tools (screwdrivers and wrenches); measuring tools (ruler, plumb line, level, square, compass and chalk line); and accessory tools (the workbench, vice, tongs and pliers). The specific tools category such as the category of the carpenter tools is divided into the hammer category, the chisel category, the saw category, the hook category, the plane category or the handle category and is given as: $\ldots \triangleright \langle \nu_{\mathrm{Prf}} \rangle \triangleright \langle \nu_{\mathrm{Car}} \rangle \triangleright \langle \nu_{\mathrm{Tol}} \rangle \triangleright \langle \nu_{\mathrm{HamC}}, \nu_{\mathrm{ChisC}}, \nu_{\mathrm{SawC}}, \nu_{\mathrm{PlaC}}, \nu_{\mathrm{HanC}} \rangle$. Examples of tools are shown in Fig. 4.74.

Usually the result of the work of men of a given profession is an object from the category of visual objects. For example, the result of the musical instruments maker is the musical instrument. The musical instrument is any device for producing a musical sound. The principal types of such instruments, classified by the method of producing sound, are percussion, stringed, keyboard, wind and electronic. The category of musical instruments is derived from the category of man-made objects. The musical instrument can be treated as the product of the musical instruments maker or as tools for musician. Examples of members of the different categories of different instruments are shown in Fig. 4.75.

Fig. 4.74. Category of gardener tools

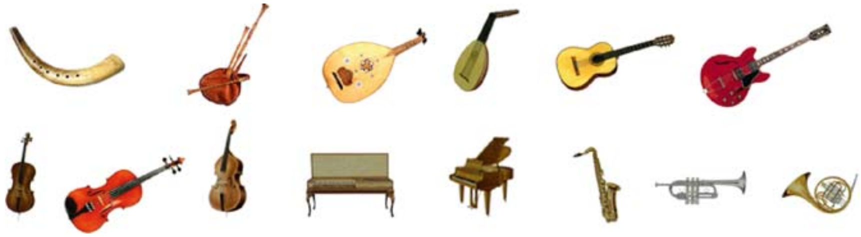


Fig.4.75. Examples of members of the different categories of different instruments

Category of Living Objects

The category of living objects ν_{Liv} are divided into the category of human beings ν_{Hum}, the category of animals ν_{Ani}, the category of plants ν_{Pla}, the category of fungus ν_{Fun}, the category of protoctist ν_{Pro} and the category of moneran ν_{Mon} and is given as: $\ldots \triangleright \langle \nu_{\mathrm{Liv}} \rangle \triangleright \langle \nu_{\mathrm{Hum}}, \nu_{\mathrm{Ani}}, \nu_{\mathrm{Pla}}, \nu_{\mathrm{Fun}}, \nu_{\mathrm{Pro}}, \nu_{\mathrm{Mon}} \rangle$. Within the plant kingdom, plants are divided into two main groups: the plants that produce seeds (flowering plants) and the other group that contains the seedless plants that reproduce by spores.

In contrast to non-living objects, the visual representations of the living objects need to take into account changes of shape of the object when parts of organism are moving. In general, an animal is a living organism that is incapable of synthesizing carbohydrates and proteins from inorganic or simple organic substances but must ingest them in complex form as food. Every animal species has a unique Latin name. The first word is the genus name, which is shared with closely related animals. The second word is a specific name which together with the genus is unique to particular species. In the case of the small animals or plants that cannot be seen by naked eye they are members of the micro-objects category that were described in the previous section. Categories that are described in this section are established based on the knowledge from zoological and botanical science.

In the case when taxonomy is well established, it can be used to describe the abstract structure of the categories. Each part of hierarchy is linked with its characteristic knowledge that describes the general features of the category of the given part of the hierarchical structure. During learning of knowledge of the category of the visual object there is a need to learn the visual concept of all possible visual representations such as the silhouette, the line drawing, the coloured object or the shaded object. These visual representations are members of the perceptual categories described in previous sections.

Category of Animals

The scientific knowledge needed in understanding of the visual object that is visual representative of the category of animals is supplied by branch of science called zoology. Zoology is a branch of biology concerned with the members of the animal kingdom and with animal life in general. Zoology is divided into a number of sub-disciplines such as cytology, embryology, morphology, physiology, pathology, palaeontology, genetics and evolution, taxonomy, ecology, and zoogeography. The zoological taxonomy is focused on the taxonomy of zoological terms. The zoological taxonomy is based on the scheme that consists of phylum, class, order, family, genus and species. The taxonomy of animals can be given in the following form:

Phylum-{porifera, cnidaria, platyhelminthes, annelida, nemotoda, crustacean, mollusca, echinoderms, chorodata}

Phylum-{porifera,…}-class-[calcarea, hexactinellida, demospongiae]

Phylum-{cnidaria}-class-[anthoza, scyphoza, hydroza, cuboza]

Phylum-{crustacean}-class-[branchiopoda, cirripedia, malacostraca, insects, arachnida]

Phylum-{crustacean}-class-[arachnida]-order-(scorpionida, acarina, araneae)

Phylum-{mollusca}-class-[bivalvia, polyplacophora, gastropoda, scaphopoda, cephalopoda]

Phylum-{echinoderms}-class-[astoreidea, echinoidea, crinoidea, holothuroidea]

Phylum-{chorodata}-class-[cyclostomata, chondrichtyhyeas, osteichtyles, choanichtyles, gymnophiona, anura, caudate, reptiles, aves, mamalia]

The categorical chain of the animals category can be given in the form that reflects the scheme of the zoological taxonomy given in the form: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{ReO}} \rangle \triangleright \langle \nu_{\mathrm{Ear}} \rangle \triangleright \langle \nu_{\mathrm{Liv}} \rangle \triangleright \langle \nu_{\mathrm{Ani}} \rangle \triangleright \langle -_{\mathrm{Phy}} \rangle \triangleright \langle -_{\mathrm{Cla}} \rangle \triangleright \langle -_{\mathrm{Ord}} \rangle \triangleright \langle -_{\mathrm{Fam}} \rangle \triangleright \langle -_{\mathrm{Gen}} \rangle \triangleright \langle -_{\mathrm{Spc}} \rangle$.

According to this scheme of zoological taxonomy the lion category can be given by the following categorical chain:
$\langle \Pi \rangle \rhd \langle \sigma_{\mathrm{El}} \rangle \rhd \langle \nu_{\mathrm{ReO}} \rangle \rhd \langle \nu_{\mathrm{Ear}} \rangle \rhd \langle \nu_{\mathrm{Liv}} \rangle \rhd \langle \nu_{\mathrm{Ani}} \rangle \rhd \langle \nu_{\mathrm{Cho}} \rangle \rhd \langle \nu_{\mathrm{Mam}} \rangle \rhd \langle \nu_{\mathrm{Car}} \rangle \rhd \langle \nu_{\mathrm{Fel}} \rangle \rhd \langle \nu_{\mathrm{Pan}} \rangle \rhd \langle \nu_{\mathrm{PLeo}} \rangle$.
The non-visual knowledge is given by the following knowledge chain:
$\kappa_{\mathrm{KB}} \rhd \langle \kappa_{\mathrm{KOb}} \rangle \rhd \langle \kappa_{\mathrm{BioOb}} \rangle \rhd \langle \kappa_{\mathrm{ZooO}} \rangle \rhd \langle \kappa_{\mathrm{Liv}} \rangle \rhd \langle -_{\mathrm{Phy}} \rangle \rhd \langle -_{\mathrm{Cla}} \rangle \rhd \langle -_{\mathrm{Ord}} \rangle \rhd \langle -_{\mathrm{Fam}} \rangle \rhd \langle -_{\mathrm{Gen}} \rangle \rhd \langle -_{\mathrm{Spc}} \rangle$.

Animals that are members of the chorodata class often are classified to the groups such as fish, amphibians, reptiles, aves and mamalia. Based on this classification, the following specific categories are derived from the chorodata category: the category of fish ν_{Fis}, the category of amphibians ν_{Amf}, the category of reptiles ν_{Rep}, the category of aves ν_{Ave} and the category of mamalia ν_{Mam} and given as $\ldots \rhd \langle \nu_{\mathrm{Ani}} \rangle \rhd \langle \nu_{\mathrm{Cho}} \rangle \rhd \langle \nu_{\mathrm{Fis}}, \nu_{\mathrm{Amf}}, \nu_{\mathrm{Rep}}, \nu_{\mathrm{Ave}}, \nu_{\mathrm{Mam}} \rangle$. The category of reptiles is divided into the category of squamata ν_{Squ}, the category of crocodilian ν_{Cro} and the category of testudines ν_{Tes}: $\ldots \rhd \langle \nu_{\mathrm{Ani}} \rangle \rhd \langle \nu_{\mathrm{Cho}} \rangle \rhd \langle \nu_{\mathrm{Rep}} \rangle \rhd \langle \nu_{\mathrm{Tes}}, \nu_{\mathrm{Cro}}, \nu_{\mathrm{Squ}} \rangle$. The category of aves is divided into the category of passeriformes ν_{Pas}, the category of falconiformes ν_{Fal}, the category of piciformes ν_{Pic}, the category of anseriformes ν_{Ans}, the category of apodiformes ν_{Apo}, the category of columbiformes ν_{Col}, the category of charadriiformes ν_{Cha} and the category of galiformes ν_{Gal} and is given as $\ldots \rhd \langle \nu_{\mathrm{Ani}} \rangle \rhd \langle \nu_{\mathrm{Cho}} \rangle \rhd \langle \nu_{\mathrm{Ave}} \rangle \rhd \langle \nu_{\mathrm{Pas}}, \nu_{\mathrm{Fal}}, \nu_{\mathrm{Pic}}, \nu_{\mathrm{Ans}}, \nu_{\mathrm{Apo}}, \nu_{\mathrm{Col}}, \nu_{\mathrm{Cha}}, \nu_{\mathrm{Gal}} \rangle$. The category of mamalia is divided into the monotremata category, the diprotodonta category, the perissodactyla category, the carnivore category, the cetacean category, the primate category and the rodentia category: $\ldots \rhd \langle \nu_{\mathrm{Ani}} \rangle \rhd \langle \nu_{\mathrm{Cho}} \rangle \rhd \langle \nu_{\mathrm{Mam}} \rangle \rhd \langle \nu_{\mathrm{Mon}}, \nu_{\mathrm{Dip}}, \nu_{\mathrm{Per}}, \nu_{\mathrm{Car}}, \nu_{\mathrm{Cet}}, \nu_{\mathrm{Prim}}, \nu_{\mathrm{Rod}} \rangle$.

The name of the zoological categories is given in the Latin language. The categories of the lower level refer to the name of the animal such as lion, tiger or monkey and the Latin term can be easily translated into English or Polish. Each category is described by the specific knowledge that is represented in the form of the knowledge schema at a given categorical level. The knowledge schema represents the main features (attributes) of the category. For example, the category carnivore is represented by the following features: flesh eating, sharp check, teeth to cutting flesh. The intermediate categories reflect the knowledge of the general animal categories and are used in the process of interpretation and understanding of the visual object. For the human subject there is easy to understand the difference between two terms such as ‘sparrow’ and ‘lion’ by invoking pictures of both animals and next refers to categories described by these pictures. SUS in order to understand these differences at first invokes the visual chains that represent categories given by these names and next uses the knowledge of each category to find the difference. The problem of the visual inference based on the categorical chain will be described in Chap. 5.

As it was described in previous sections, the knowledge that is included in categorical chains is acquired during the learning process. During learning process knowledge of the specific category such as the lion category ν_{PLeo} given by the categorical chain $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{ReO}} \rangle \triangleright \langle \nu_{\mathrm{Ear}} \rangle \triangleright \langle \nu_{\mathrm{Liv}} \rangle \triangleright \langle \nu_{\mathrm{Ani}} \rangle \triangleright \langle \nu_{\mathrm{Cho}} \rangle \triangleright \langle \nu_{\mathrm{Mam}} \rangle \triangleright \langle \nu_{\mathrm{Car}} \rangle \triangleright \langle \nu_{\mathrm{Fel}} \rangle \triangleright \langle \nu_{\mathrm{Pan}} \rangle \triangleright \langle \nu_{\mathrm{PLeo}} \rangle$, all intermediate categories given by this categorical chain need to be fulfilled (learned). In the case of the lack of knowledge of categories at the intermediate levels these not known categories can be left blank in the first stage of the learning process. For example, when during learning of the category ‘tiger’ the genus is not known, this part is left ‘blank’: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{ReO}} \rangle \triangleright \langle \nu_{\mathrm{Ear}} \rangle \triangleright \langle \nu_{\mathrm{Liv}} \rangle \triangleright \langle \nu_{\mathrm{Ani}} \rangle \triangleright \langle \nu_{\mathrm{Cho}} \rangle \triangleright \langle \nu_{\mathrm{Mam}} \rangle \triangleright \langle \nu_{\mathrm{Car}} \rangle \triangleright \langle \nu_{\mathrm{Fel}} \rangle \triangleright \langle \nu_{\_} \rangle \triangleright \langle \nu_{\mathrm{Tig}} \rangle$. Figure 4.76 shows examples of members of the aves category. Figure 4.77 shows examples of members of the category of mamalia.

The biological taxonomy does not take into account the size of the animal. In present categorical classification the animals that are too small to be visible by naked eye are members of the micro-objects category. Also, the visual appearance of the young and adult animal can be very different. To enable naming and recognition of animals, the additional categories are introduced, the category of young animals and the category of adult animals.

Fig. 4.76. Example of members of the category of aves (birds)

Fig. 4.77. Example of members of the category of mammalian (mammals)

Category of plants

Plant is multicellular, eukaryotic life forms characterized by photosynthetic nutrition, unlimited growth at localized regions, cells that contain cellulose in their walls, the absence of organs of locomotion, the absence of sensory and nervous systems. Photosynthetic nutrition is the process in which chemical energy is produced from water, minerals and carbon dioxide with the aid of pigments and the radiant energy of the Sun.

Plants that are too small to be seen by the naked eye are members of the micro-objects category. Similarly like it was described in previous section of this book where the description of the taxonomy of animals was presented, the taxonomy of plants that is presented in this section is based on the botanical taxonomy. In the case when differentiation of the plants assumes taking into account only the rough approximation of the botanical knowledge the so-called plant taxonomy is used. In this chapter the plant taxonomy is briefly described. The plant category is divided into the category of trees ν_{Tre}, the category of shrubs ν_{Shr}, the category of undershrubs ν_{UnS} and the category of vines ν_{Vin} and is represented in the form of the categorical chain as $\langle \Pi \rangle \triangleright \langle \sigma_{\text{El}} \rangle \triangleright \langle \nu_{\text{ReO}} \rangle \triangleright \langle \nu_{\text{Ear}} \rangle \triangleright \langle \nu_{\text{Liv}} \rangle \triangleright \langle \nu_{\text{Pla}} \rangle \triangleright \langle \nu_{\text{Tre}}, \nu_{\text{Shr}}, \nu_{\text{UnS}}, \nu_{\text{Vin}} \rangle$. The category of trees is divided into the broad-leaved trees ν_{BrT} and conifers ν_{Con} and is given as: $\ldots \langle \nu_{\text{Pla}} \rangle \triangleright \langle \nu_{\text{Tre}} \rangle \triangleright \langle \nu_{\text{BrT}}, \nu_{\text{Con}} \rangle$.

The tree is a woody plant that renews its growth every year. Most trees have a single self-supporting trunk containing woody tissues. The trunk usually produces secondary limbs, called branches. A general definition describes a tree as a perennial woody plant that develops along a single main trunk to a height of at least 4.5 m at maturity. Conifers include trees and shrubs in seven extant families and 550 species. Familiar representatives are araucarias, cedars, cypresses, firs, hemlocks, junipers, larches, pines, redwoods, spruces or yews. Angiosperms dominate the Earth's present flora and they contain more than 250,000 species. Angiosperms are divided on the basis of a group of characteristics into two classes: the monocotyledons and the dicotyledons. The most numerous of the monocotyledonous trees are palms; others include agaves, aloes, dracaenas, screw pines and yuccas. The greatest number of tree species is dicotyledons and they are represented by such familiar groups as apples, birches, elms, hollies, magnolias, maples, oaks, poplars and willows. A shrub is defined as a woody plant with multiple stems that is, in most cases, less than 3 m tall.

Fig. 4.78. Examples of members of the different categories of tree represented as the silhouette – one of the perceptual categories

The category of plants is based on the botanical taxonomy. The category of real-world objects such as plants can have many different perceptual representations. Examples of members of the different categories of trees are shown in Fig. 4.78. Examples shown in Fig. 4.78 are members of the silhouette category that is one of the perceptual categories. As we can see even this simple visual representation makes it possible to differentiate among the different trees.

The real-world object such as a plant consists of different parts. The part category is introduced to represent the different parts of the object. The part category is an auxiliary category that can be derived from any part of the categorical hierarchy. The schema of the part category shows the links to categories that constitute an object. For example, typical flowering plants such as a tree consists of roots, the trunk, stems, leaves, flowers, fruits and seed. The schema of the part category includes the specific categories that refer to parts of the object. For example, the tree category consists of the different specific categories given by the following schema of the part category: $\ldots\langle \nu_{\mathrm{Pla}} \rangle \triangleright \langle \nu_{\mathrm{Tre}} \rangle \succ [\tau_{\mathrm{Rot}}, \tau_{\mathrm{Trn}}, \tau_{\mathrm{Stm}}, \tau_{\mathrm{Lef}}, \tau_{\mathrm{Flw}}, \tau_{\mathrm{Frt}}, \tau_{\mathrm{Sed}}]$. Each part category such as the roots category τ_{Rot}, the trunks category τ_{Trn}, the stems category τ_{Stm}, the leaves category τ_{Lef}, the flowers category τ_{Flw}, the fruits category τ_{Frt} and the seeds category τ_{Sed} refers to the parts of the tree. Figure 4.79 shows characteristic parts of the tree. The part can be treated as the independent object that in turn consists of other parts. For example, a fruit shown in Fig. 4.80 consists of characteristic parts such as seeds. Most fruits grow on tress, whereas some of them such as muskmelons grow on creeping vines that feature large, lobed leaves and yellow flowers. The muskmelon has a hard rind that encases the juicy pulp, and flat seeds that form a netlike mass in the hollow centre. Figure 4.79a shows the tangerine, whose fruits are produced from a small,

a b

Fig. 4.79. The example of the category of parts of tree and the category of parts of shrubs

a b c d e

Fig. 4.80. Examples of the category of parts of plants

thorny tree that bears simple leaves and orange like blossoms. Figure 4.79b shows the common orchard fig, a bush-like tree with deeply lobed leaves. Its fruit is a fleshy receptacle (cross section, centre) containing numerous small seeds.

Fruit is the ripened ovary of any flowering plant, or angiosperm, and usually contains one or more seeds. The knowledge schema for the fruit category includes the visual concept ∂_{ViC}, the name ∂_{Nam}, the definition ∂_{Def} and the method of exemplar generation ∂_{MGe} and is given as $\ldots \rhd \langle \kappa_{\mathrm{Fru}} \rangle \prec \{\partial_{\mathrm{ViC}}, \partial_{\mathrm{Nam}}, \partial_{\mathrm{MIn}}, \partial_{\mathrm{Def}}, \partial_{\mathrm{MGe}}\}$. For example, the definition of the category of the fruit includes, among others, the following parts and can be given in the following form: fruit-is-part-of-[plant], fruit-consists of-[blade, core], fruit-is-[outgrowth from the stem of plant]. The definition includes links to other knowledge categories that usually contain the non-visual knowledge.

The fruit can be grown on the tree, shrubs or vines. The tree fruits category is divided into the category of plums, apples or pears: $\ldots \langle \nu_{\mathrm{Pla}} \rangle \rhd \langle \nu_{\mathrm{Tre}} \rangle \succ [\tau_{\mathrm{Frt}}] \rhd \langle \nu_{\mathrm{Plu}}, \nu_{\mathrm{App}}, \nu_{\mathrm{Pea}} \rangle$. Examples of the different tree fruits categories are given in Fig. 4.81 such as the quince category (a), the pear category (b), the lime category (c), the plum category (d), the mango category (e), the papaya category (f), the papaw category (g), the citrus category (h) and the grapefruit category (i).

Fig. 4.81. Examples of members of the tree fruits category

Apple is the most widely and one of the oldest cultivated of all fruit trees. Today about 7,500 varieties are grown worldwide. The visual features of the apple category such as size or colour that are part of the knowledge schema supply important information used during visual recognition of the different objects. For example, the colour of members of the apple category ranges from various shades of red to yellow or green. The non-visible features such as sweetness, aroma and crispness vary greatly from one apple to another, and can supply the valuable information during naming and recognition of the different objects. These features make it possible to establish the links between the apple category and the food category. Apples can be eaten as fresh fruit, can be canned as sauce or pie filling or can be made into cider, cider vinegar, juice, jelly or apple butter. The name of the apple category is derived from the name of the tree category, so the name of the tree is used to denote the name of the fruit. The category of apple is divided into the special categories such as the McIntosh category, the Delicious category, the Stayman category, the Rome Beauty category or the Jonathan category and in symbolic notation given as $\ldots\langle \nu_{\text{Pla}} \rangle \rhd \langle \nu_{\text{Tre}} \rangle \succ [\tau_{\text{Frt}}] \rhd \langle \nu_{\text{App}} \rangle \rhd \langle \nu_{\text{McI}}, \nu_{\text{Del}}, \nu_{\text{Sty}}, \nu_{\text{RoB}}, \nu_{\text{Jon}} \rangle$. Figure 4.82 shows examples of members of the apple category.

The flower is the reproductive structure of angiosperms or flowering plants. Compared with reproductive structures of other plants, the flower is unique in several ways. For example, it consists of four kinds of modified leaves, two of which (stamens and carpels) bear pollen and seeds. Flower is a part of the plants and the flower category is the category that is derived from the category of part of plants such as trees or vines. Petals that form the main visual part of flowers are often shows with other parts as a stem. Colour is very important visual feature of the flower category. Examples of members of the different flower categories are given in Fig. 4.83.

Fig. 4.82. Examples of members of the apple category

Fig. 4.83. The example of members of the different flowers categories

The leaf is part of the plant that intercepts light, exchange gases, and provides a site for photosynthesis. Some leaves also store food and water, provide support, or form new plants. The leaf category is the category that is derived from the category of part of plants. The name of the leaf category is derived from the name of the plant category, so the name of the plant is used to denote the name of the leaf. For example, the tree leaf category is divided into categories such as: the oak category, the lime category, the poplar category, the elm category, the hornbeam category, the ash-tree category, the beech category or the birch category and is given as $\ldots\langle \nu_{\mathrm{Pla}}\rangle \rhd \langle \nu_{\mathrm{Tre}}\rangle \succ [\tau_{\mathrm{Lef}}] \rhd \langle \nu_{\mathrm{Oak}}, \nu_{\mathrm{Pop}}, \nu_{\mathrm{Lim}}, \nu_{\mathrm{Elm}}, \nu_{\mathrm{Hor}}, \nu_{\mathrm{AsT}}, \nu_{\mathrm{Bee}}, \nu_{\mathrm{Bir}}\rangle$.

The tree such as oak has many different types. Each type of oak has its name (in Latin) and characteristic parts such as leaves. Examples of names of the tree-oak leaves are given in the form of the names expressed in two languages (English and Latin) – [Mongolian oak (*Q. mongolica*), Oriental oak (*Q. variabilis*), Armenian(pontic) oak (*Q. pontica*), chestnut-leaved oak (*Q. castaneaefolia*), golden oak (*Q. alnifolia*), Holm(holly) oak (*Q. ilex*), Italian oak (*Q. frainetto*), Lebanon oak (*Q. libani*), Macedonian oak (*Q. trojana*), Portuguese oak (*Q. lusitanica*), blue Japanese oak (*Q. glauca*), daimyo oak (*Q. dentata*), Japanese evergreen oak (*Q. acuta*), sawtooth oak (*Q. acutissima*), English oak (*Q. robur*), pin oak (q.v.; *Q. palustris*) northern red oak (*Q. rubra*). White oak (*Q. alba*) bur oak (q.v.; *Q. macrocarpa*), Aleppo oak (*Q. infectoria*), cork oak (*Q. suber*), tannin-rich kermes oak (*Q. coccifera*)…]. The name of specific category of the tree leaf is derived from the name of the specific category of tree. For example, naming the leaf as the 'oak leaf' indicates that this leaf was grown on the tree called oak. The category of oak leaves is divided into the specific category of oaks such as: African oak, Australian oak, bull oak, Jerusalem oak, poison oak, river oak, she-oak, silky oak, tanbark oak, Tasmanian oak or tulip oak and given as

Fig. 4.84. Examples of members of different categories of leaves

Fig. 4.85. The different leaves (shaded representation) and its silhouettes

$\ldots\langle\nu_{\text{Pla}}\rangle\rhd\langle\nu_{\text{Tre}}\rangle\succ[\tau_{\text{Lef}}]\rhd\langle\nu_{\text{Oak}}\rangle\rhd\langle\nu_{\text{Afr}},\nu_{\text{Aus}},\nu_{\text{Bul}},\nu_{\text{Jer}},\nu_{\text{Poi}},\nu_{\text{Riv}},\nu_{\text{ShO}},\nu_{\text{Sil}},\nu_{\text{Tan}},\nu_{\text{Tas}},\nu_{\text{Tul}}\rangle$. Figure 4.84 shows examples of members of the different categories of leaves.

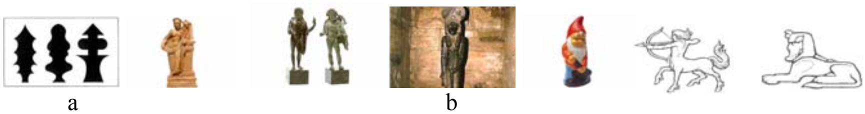
The leaf is an example of the visual object (real-world object) that can be regarded as a 2D object. The silhouette of the leaf supplies nearly all visual information about the shape of the leaf. Figure 4.85 shows members of the different leaf categories. Each pair of leaves shown in Fig. 4.85 is given as members of the different perceptual categories, the silhouette category and the shaded-object category. The silhouette is obtained by segmentation of the image.

Members of any ontological category that consists of the different elements (objects) are called complex objects. These elements from which the complex object is built are called parts. Parts are regarded as a special kind of categories (part categories) that refer the selected categorical chain category. In the categorical chain the derivation of the part category is denoted as $\langle\nu_{\text{Tre}}\rangle\succ[\tau_{\text{Lef}}]$, where symbol $\succ$ denotes that the part category $[\tau]$ is derived from one of the ontological categories $\langle\nu\rangle$. The part category is denoted by symbol τ in the brackets $[\tau]$. Parts that are members of the category of man-made objects are complex objects. Objects from which the complex object is assembled are called components.

The leaf category $[\tau_{\text{Lef}}]$ is the part category that is derived from one of the specific plant categories such as the tree category $\langle\nu_{\text{Tre}}\rangle$. The leaf category $[\tau_{\mathrm{Lef}}]$ consists of different elements, members of the part category. Categories such as the blade category, the stalk category or the venation category are derived from the leaf category and are represented by following categorical chain: $\ldots\langle\nu_{\mathrm{Pla}}\rangle \rhd \langle\nu_{\mathrm{Tre}}\rangle \succ [\tau_{\mathrm{Lef}}] \succ [\tau_{\mathrm{Bla}}, \tau_{\mathrm{Stl}}, \tau_{\mathrm{Ven}}]$. The knowledge schema of the leaf category includes the visual concept ∂_{ViC}, the name ∂_{Nam}, the definition ∂_{Def} and the method of exemplar generation ∂_{MGe} and is given as $\ldots \rhd \langle\kappa_{\mathrm{Lef}}\rangle \prec \{\partial_{\mathrm{ViC}}, \partial_{\mathrm{Nam}}, \partial_{\mathrm{MIn}}, \partial_{\mathrm{Def}}, \partial_{\mathrm{MGe}}\}$. For example, the definition of the leaf category can be given in the following form: leaf-function – [to produce food for the plant by photosynthesis]; leaf-part_of – [plant]; leaf-part_of – [stem system]; leaf-is – [outgrowth from the stem of vascular plant]; leaf (attributes) – [size, colour, the nature of the blade margin, the type of venation (arrangement of veins)].

4.2.3.7. Category of Imaginary Objects

The category of imaginary objects includes the category of objects of scientific visualization ν_{SciV}, the category of objects of literature fiction ν_{InvL}, the category of visual art objects ν_{InvA}, the category of 3D fictious figures ν_{FarT} and the category of mythological objects ν_{Mit}. Dwarfs (dvergar) play a part in Norse mythology. They were very wise and expert craftsmen who forged practically all of the treasures of the gods, in particular Thor's hammer. Four of them are supporting the sky, made of the skull of this primeval giant. They may have been originally nature spirits or demonic beings, living in mountain caves, but they generally were friendly to man. Greek centaurs (in Greek mythology) are part horse and part man, dwelling in the mountains of Thessaly and Arcadia. Sphinx is a mythological creature with a lion's body and human head, an important image in Egyptian and Greek art and legend. Brownie in English and Scottish folklore, a small, industrious fairy or hobgoblin believed to inhabit houses and barns. The category of imaginary objects can be represented by following categorical chain: $\langle\Pi\rangle \rhd \langle\sigma_{\mathrm{El}}\rangle \rhd \langle\nu_{\mathrm{ImO}},\ldots\rangle \to \langle\nu_{\mathrm{SciV}}, \nu_{\mathrm{InvA}}, \nu_{\mathrm{InvL}}, \nu_{\mathrm{Mit}}, \nu_{\mathrm{FarT}}, \ldots\rangle$. Figure 4.86 shows examples of members of the category of imaginary objects: 3D fictious (invented) tree and the category of the mythological objects.



a b

Fig. 4.86. Examples of category of imaginary objects: 3D fictious (invented) tree and the category of the mythological objects

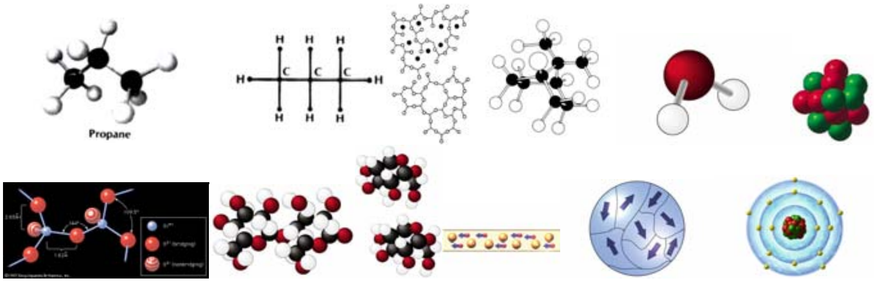


Fig. 4.87. Category of the scientific visualization

The scientific visualization is focused on generation of the visual object that is visual representative of the model of the examined phenomenon. The category of scientific visualization is divided into the category of the schematic visualization and the realistic visualization. The category of schematic visualization is divided into the category of diagrams, the category of maps, the category of diagrammatic representations or the category of schematic data visualization. The category of the realistic visualization is divided into the category of modelling of the non-visible phenomenon and the category of realistic data visualization. Figure 4.87 shows examples of the different categories of scientific visualization that involves the category of schematic visualization and the category of realistic visual representations.

4.2.3.8. Category of Real-World Processes

The category of Earthy-world objects is divided into the category of living objects, the category of non-living objects and the category of processes: $O \triangleright \langle \nu_{\mathrm{ReO}} \rangle \triangleright \langle \nu_{\mathrm{Ear}} \rangle \triangleright \langle \nu_{\mathrm{Liv}}, \nu_{\mathrm{NLiv}}, \nu_{\mathrm{Proc}} \rangle$. The category of processes refers to changes of the visual objects that can be observed during the period of time. The visual process is often represented by the category of animation that is one of the structural categories. The process category is divided into the natural process category and the artificial process category:

Fig. 4.88. Example of the visual representative of the acoustic process – the sound wave

$\ldots \triangleright \langle \nu_{\text{Proc}} \rangle \triangleright \langle \nu_{\text{NatP}}, \nu_{\text{ArtP}} \rangle$. The category of natural processes is divided into atmospheric processes, acoustic processes, physical processes, chemical processes, geological processes or biological processes: $\ldots \triangleright \langle \nu_{\text{Proc}} \rangle \triangleright \langle \nu_{\text{NatP}} \rangle \triangleright \langle \nu_{\text{AtmP}}, \nu_{\text{AcuP}} \nu_{\text{PhyP}}, \nu_{\text{ChmP}}, \nu_{\text{GeoP}}, \nu_{\text{BioP}} \rangle$. The category of physical processes is divided into the category of changing state processes (melting, boiling, freezing), the category of heat transfer processes (convection, conduction, radiation), the category of radioactivity processes, the category of magnetism processes, the category of sound waves processes or the category of electromagnetic wave processes (light, ultraviolet, infrared). For example, the category of chemical processes is connected with the chemical reactions.

The category of acoustic processes is divided into the category of music, the category of songs or the category of noise: $O \triangleright \langle \nu_{\text{ReO}} \rangle \triangleright \langle \nu_{\text{Ear}} \rangle \triangleright \langle \nu_{\text{NLiv}} \rangle \triangleright \langle \nu_{\text{NatP}} \rangle \triangleright \langle \nu_{\text{Acus}} \rangle \triangleright \langle \nu_{\text{Mus}}, \nu_{\text{Son}}, \nu_{\text{Spi}}, \nu_{\text{Noi}} \rangle$. The visual representative of the acoustic processes is the sound wave that can be given as the signal in the time domain or transformed into the frequency or time-frequency domain.

Knowledge that is used to interpret the visual object is supplied by categories derived from the category of acoustic objects. Acoustic is the science concerned with the production, control, transmission, reception and effects of sound. Sound results from the vibration of elastic bodies such as violin string or human vocal chords. Knowledge of the category of music sounds is given by the following schema: $\ldots \langle \kappa_{\text{Acus}} \rangle \triangleright \langle \kappa_{\text{Mus}} \rangle \prec \{\partial_{\text{ViC}}, \partial_{\text{Nam}}, \partial_{\text{Def}}, \partial_{\text{Fet}}, \partial_{\text{MGe}}\}$. The definition has two parts, the one that defines the sound in the form 'that is…..', and the second one that defines sound in the form 'consist of…'. The definition includes the characteristic properties of the sound that are related to other categories such as the amplitude category or the power spectrum category that are members of the different knowledge categories.

For example, 'the sound is result of vibration of elastic body' – $\langle v_{\text{Son}} \rangle \prec \partial_{\text{is}} \langle result - v_{\text{vibrat}} - v_{\text{ElBod}} \rangle$, 'air_preasure_wave transmitted through air' $\langle v_{\text{Son}} \rangle \prec \partial_{\text{is}} \langle v_{APWav} - \text{transmited\_through} - v_{\text{Air}} \rangle$. The features such as duration amplitude, frequency or power spectrum are given by the part of the knowledge schema: $\langle v_{\text{Son}} \rangle \prec \partial_{\text{Fet}} \langle F_{\text{Amp}}, F_{\text{Fre}}, F_{\text{Spe}} \rangle$.

The category of the atmospheric processes is divided into the winding category, the hurricane category, the tornado category, the twister category, the clouding category, the raining category, the hail category or the snowing category: $\ldots \rhd \langle v_{\text{Proc}} \rangle \rhd \langle v_{\text{NatP}} \rangle \rhd \langle v_{\text{AtmP}} \rangle \rhd \langle v_{\text{Win}}, v_{\text{Hur}}, v_{\text{Tor}}, v_{\text{Twi}}, v_{\text{Cld}}, v_{\text{Ran}} \rangle$. The visual representatives of the atmospheric processes are shown in Fig. 4.89. The sequence of changing objects shows changes of atmospheric processes.

The category of geological processes is divided into the category of tectonic processes (earthquake, volcano), the category of erosion processes or the category of sedimentation processes: $\ldots \rhd \langle v_{\text{Proc}} \rangle \rhd \langle v_{\text{NatP}} \rangle \rhd \langle v_{\text{GeoP}} \rangle \rhd \langle v_{\text{TecP}}, v_{\text{Ero}}, v_{\text{Sed}} \rangle$. Figure 4.90 shows the geological process as a sequence of pictures representative of the perceptual animation category. In Fig. 4.91 the schematic representation of the geological processes is shown.

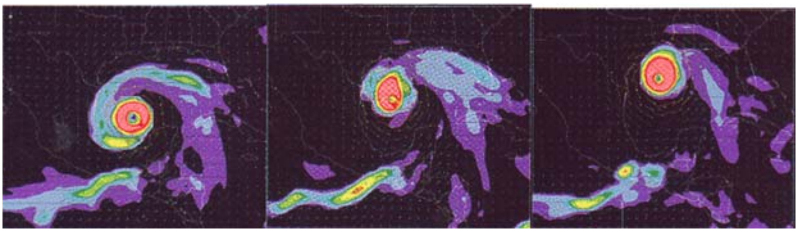


Fig. 4.89. Example of visual objects representatives of the category of the atmospheric processes

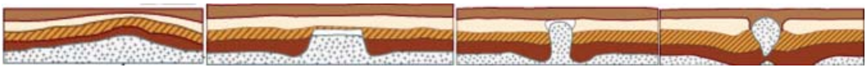


Fig. 4.90. Category of geological processes – schematic representation

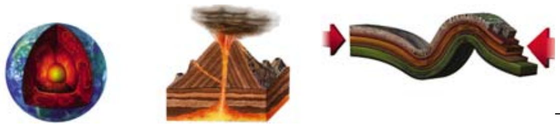


Fig. 4.91. Schematic representation of the geological processes

The category of the biological processes is divided into the category of the botanical, or the category of the zoological or the category of the medical process: $\ldots \triangleright \langle \nu_{\text{Proc}} \rangle \triangleright \langle \nu_{\text{NatP}} \rangle \triangleright \langle \nu_{\text{BioP}} \rangle \triangleright \langle \nu_{\text{BotP}}, \nu_{\text{ZooP}}, \nu_{\text{MedP}} \rangle$. The category of the botanical processes refers to the category of plants described in the previous section. The category of the botanical processes is divided into the category of growing plants or the category of making fruits. Examples of members of the category of the biological processes are shown in Fig. 4.92.

From the category of artificial processes, the category of engineering processes, the category of medical processes or the scientific processes is derived and is given as: $\ldots \triangleright \langle \nu_{\text{Proc}} \rangle \triangleright \langle \nu_{\text{Art}} \rangle \triangleright \langle \nu_{\text{EngP}}, \nu_{\text{SciP}}, \nu_{\text{MedP}} \rangle$. From the category of the engineering processes, the specific engineering categories such as electrical, mechanical or chemical are derived and given as $\ldots \triangleright \langle \nu_{\text{Proc}} \rangle \triangleright \langle \nu_{\text{Art}} \rangle \triangleright \langle \nu_{\text{EngP}} \rangle \triangleright \langle \nu_{\text{ElectE}}, \nu_{\text{MechE}}, \nu_{\text{ChemE}} \rangle$. Engineering processes can be represented in the form of schema, diagram, photograph or animation. The schematic representation in the form of the diagram refers to the category of phenomenon. The knowledge of both the phenomena and interpretation of the diagram is needed to understand the information that schema (diagram) conveyed. The knowledge of the phenomena makes it possible to interpret the diagram in terms of the properties of the process. The category of schema of the industrial process is derived from the category of engineering processes. Figure 4.93 shows examples of members of the category of schema of real-world processes. Members of the category of schema of real-world processes are representatives of the category of coloured object.

Fig. 4.92. Examples of members of the category of the biological processes

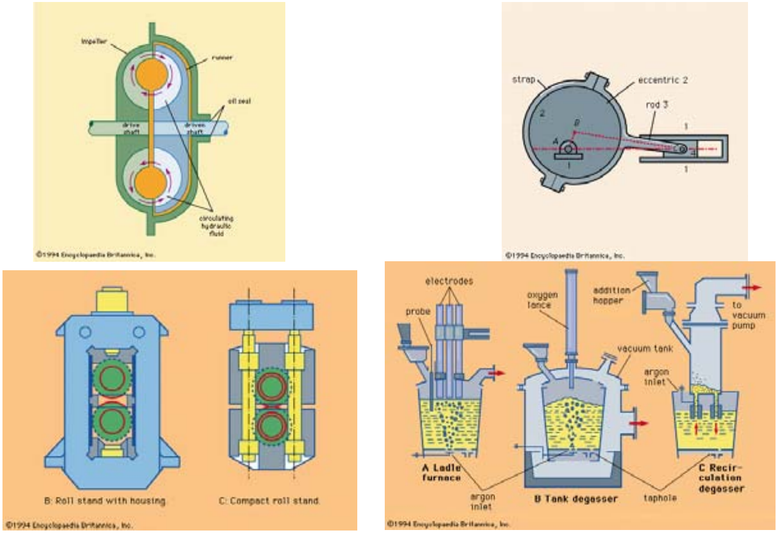


Fig. 4.93. Examples of the category of schema of real-world processes

4.2.3.9. Category of Visual Tests

The category of visual tests is derived from the pattern category. The member of the pattern category can be composed from figures, signs or the real-world objects. In the case of the category of visual intelligence tests, the visual objects that consist of the visual intelligence test are called the visual symbols. The category of the visual test is divided into the category of the visual psychological test and the category of the visual educational test and is denoted as: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pt}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{VEduT}}, \nu_{\mathrm{VPshT}} \rangle$.

Knowledge category of the visual psychological tests such as the category of the projective tests is derived from the category of the psychological objects. The category of the psychological objects is derived from the knowledge object and supplies the non-visual knowledge that is needed to interpret the perceived object. The category of the psychological object is divided into the category of visual psychological tests, the category of psychological diagnosis or the category of psychological therapies and is denoted as: $\kappa_{\mathrm{KB}} \triangleright \langle \kappa_{\mathrm{KOb}} \rangle \triangleright \langle \kappa_{\mathrm{PshO}} \rangle \triangleright \langle \kappa_{\mathrm{VPsT}}, \kappa_{\mathrm{PsyD}}, \kappa_{\mathrm{PsyT}} \rangle$.

One of the categories of the visual psychological test is the projective test category that is represented by following categorical chain: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pt}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{VPshT}} \rangle \triangleright \langle \nu_{\mathrm{PrjT}} \rangle$. Projective tests are techniques in psychology that rely on ambiguous stimuli to assess individual's personality structure as a whole. One of the visual psychological tests is the Rorschach inkblot test. The Rorschach inkblot test is an example of the widely used projective test. Figure 4.94 shows examples of objects used in the Rorschach test. The Rorschach test consists of ten bisymmetrical inkblots, five in black and white and five in colour. The subject is asked to say what the inkblots look like. A subject's style of response – such as reacting to colour or shading, describing an object in motion, placing an object within a specific location, or making a wholly original observation also became an important determinant of personality.

Intelligence tests are series of tasks designed to measure the capacity to make abstractions, to learn, and to deal with novel situations. Intelligence tests that include tasks that deal with visual forms (shapes) are called the visual intelligence tests. The visual intelligence test category is divided into the category of the comparison–selection tests, the category of matrix tests or the category of spatial tests and is represented by the following categorical chain: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pt}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{VEduT}} \rangle \triangleright \langle \nu_{\mathrm{VisT}} \rangle \triangleright \langle \nu_{\mathrm{CST}}, \nu_{\mathrm{MtT}}, \nu_{\mathrm{SpT}} \rangle$.

Knowledge category of the visual educational test is derived from the category of the educational object. The category of the educational object is derived from the knowledge object and supplies the non-visual knowledge that is needed to interpret the visual object. The category of the visual test is derived from the category of the educational object and is divided into the category of visual intelligence test, the category of the educational task or the category of the educational learning theories: $\kappa_{\mathrm{KB}} \triangleright \langle \kappa_{\mathrm{KOb}} \rangle \triangleright \langle \kappa_{\mathrm{EduO}} \rangle \triangleright \langle \kappa_{\mathrm{VisT}}, \kappa_{\mathrm{EduT}}, \kappa_{\mathrm{EdLT}} \rangle$.

The category of comparison–selection test includes comparison of the stimulus object v with other objects called answers objects $o_i, i = 1, \ldots N$, and selection one that is identical to the stimulus object [12–18]. The

Fig. 4.94. The ten bisymmetrical inkblots used in the Rorschach test

category of the comparison–selection tests is divided into the category of visual discrimination tests, the category of visual memory tests, the category of visual–spatial relationship tests, the category of visual form constancy tests, the category of visual sequential memory tests, the category of visual figure ground tests or the category of visual closure tests: $\ldots \rhd \langle \nu_{\mathrm{VEduT}} \rangle \rhd \langle \nu_{\mathrm{VisT}} \rangle \rhd \langle \nu_{\mathrm{CST}} \rangle \rhd \langle \nu_{\mathrm{VMem}}, \nu_{\mathrm{VFoC}}, \nu_{\mathrm{VDis}} \rangle$. The knowledge schema supplies knowledge of the test in the form of the name of the test, the definition in the form of the verbal description, the definition in the form of the formal specification and the proposed solution. The formal specification gives the description of the test in terms of the stimulus form that is compared with *N* answer forms. Knowledge category of the visual test is divided into the category of the visual memory test, the category of the visual form constancy test or the category of the visual discrimination test: $\kappa_{\mathrm{KB}} \rhd \langle \kappa_{\mathrm{KOb}} \rangle \rhd \langle \kappa_{\mathrm{EduO}} \rangle \rhd \langle \kappa_{\mathrm{VisT}} \rangle \rhd \langle \kappa_{\mathrm{CST}} \rangle \rhd \langle \kappa_{\mathrm{VMem}}, \kappa_{\mathrm{VFoC}}, \kappa_{\mathrm{VDis}} \rangle$.

The category of knowledge of the visual test is derived from the category of knowledge. The knowledge schema includes the name, the (linguistic) verbal description, the formal description and the solution: $\ldots \rhd \langle \kappa_{\mathrm{VisT}} \rangle \rhd \langle \kappa_{\mathrm{CST}} \rangle \rhd\!\prec \{\partial_{\mathrm{Nam}}, \partial_{\mathrm{LinD}}, \partial_{\mathrm{ForD}}, \partial_{\mathrm{Sol}}\}$. The name of the category of test is expressed in the form of one of natural languages. The name of the specific category such as the visual discrimination test refers to the test that can consist of figures, signs, letters or the real-world objects. The category of the visual discrimination test is divided into the discrimination test of figures, the discrimination test of letters, the discrimination test of signs and the discrimination test of real-world objects. Examples of members of different categories of visual discrimination tests are shown in Figs. 4.95–4.97.

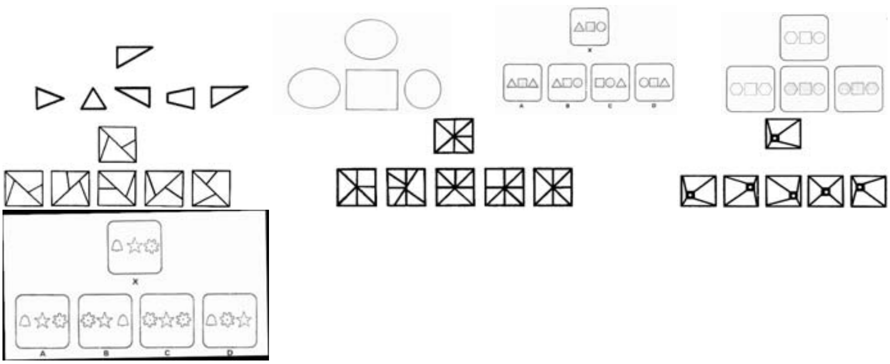


Fig. 4.95. Examples of the visual discrimination test A (consisting of figures)

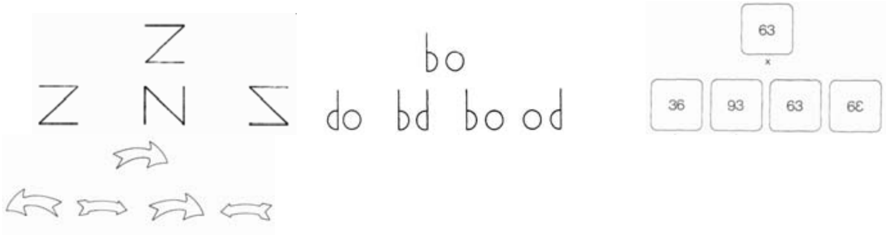


Fig. 4.96. Example of the visual discrimination test B and C (consisting of letters or signs)

Fig. 4.97. Example of the visual discrimination test D (consisting of real-world objects)

The linguistic description is given in the linguistic form and describes the task that needs to be performed. For example, the category of the visual discrimination test is described by verbal description: ‘Look at this object and find it among the five objects below’. The visual memory test is also known as information processing test – observing, seeing and remembering. The task is formulated in the form of the question: ‘Look at this form and remember it so that you can find it on another page. Find it among these forms’. Visual form constancy test can be described by verbal description: ‘Look at this form and find this form from among these five forms, even it may be smaller, bigger, darker, turned or upside down’. The category of the visual sequential memory test can be described by verbal description: ‘Look at this form very closely, remember it so that you can find it among other forms and next find it among these forms’. The linguistic description can include description of action that needs to be undertaken, e.g. the visual form constancy test includes comparison of the objects and selection one that is similar to the given one.

In the visual discrimination test or the visual memory test the stimulus form is compared with N answer forms to find one that matches each other. In these tests the stimulus form (the form to which all forms are compared) is denoted as v (see Fig. 4.98) and all answer forms are denoted as $o_i, i = 1,\dots,N$, where N is a number of forms for comparison. The task is formulated as: 'Find $o_i, i = 1,\dots,N$ that matches an object $o_i \equiv v$: $[\exists o_i : o_i \equiv v, \text{for}\, i = 1,\dots,N] \Rightarrow o_i \triangleright \sigma$.' The object o_i for which this matching is obtained is denoted as σ. In the visual discrimination test the stimulus form v, and answer forms $o_i, i = 1,\dots,N$, consist of the one object, whereas in the visual sequential memory tests the number of objects is greater than one. In the case of the visual sequential memory test the stimulus form is given in the form of the string $v_j, i = 1,\dots,M$, and all answer forms are given in the form $o_i^j, i = 1,\dots,N, j = 1,\dots,M$ (Fig. 4.99).

In the visual form constancy test the task is formulated as: 'Find o_i, $i = 1,\dots, N$ that is similar to an object $o_i \approx v$: $[\exists o_i : o_i \approx v$, for $i = 1,\dots,N] \Rightarrow o_i \triangleright \sigma$ (see Fig. 4.100). The visual similarity assumes that objects can be different only in this respect that is the result of application of the affine transformation such as rotation or scaling.

In the visual figure ground tests the stimulus form v (Fig. 4.101) is the figure whereas all answer forms are images from which this figure needs to be extracted $I_i, i = 1,\dots,N$, where N is a number of images for comparison. In the visual figure ground tests the task is formulated as: 'Find image $I_i, i = 1, \dots, N$ that includes object $v \subset I_i$: $[\exists o_i : v \subset I_i$, for

Fig. 4.98. Example of the visual discrimination test

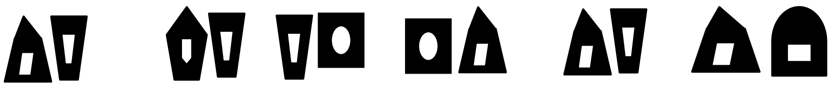


Fig. 4.99. Example of the visual sequential memory tests

Fig. 4.100. Example of the visual form constancy test

Fig. 4.101. Example of the visual figure ground test

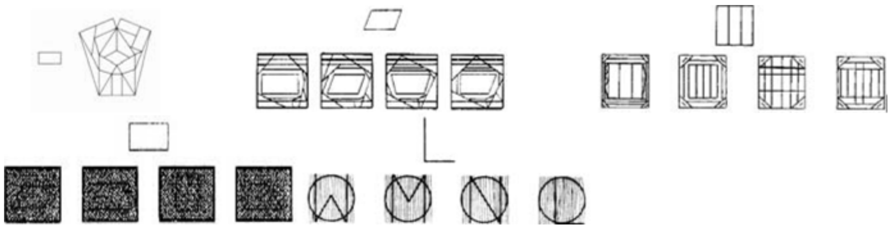


Fig. 4.102. Examples of the visual figure ground tests

$i = 1,\ldots,N] \Rightarrow I_i \triangleright \sigma$. Examples of the visual form constancy test are shown in Figs. 4.101 and 4.102.

The matrix test consists of eight objects that are placed in the pattern of matrix. The task is to find the ninth object in the matrix (selected from the given answer objects) based on the relationship discovered among the eight objects. The category of matrix tests is divided into the category of arithmetical operations (AO) test, the category of geometrical operation test (GO) and the category of finding relationships test (FR) and is represented by the following categorical chain: $\ldots \triangleright \langle \nu_{\text{VisT}} \rangle \triangleright \langle \nu_{\text{MtT}} \rangle \triangleright \langle \nu_{\text{AO}}, \nu_{\text{GO}}, \nu_{\text{FR}}, \ldots \rangle$. These tests are described in Chap. 5. In this section only short description of these tests is given.

The test type of arithmetical operations (AO) consists of eight patterns of two different types of figures that code the arithmetical operations such as the addition or the subtraction. Each of eight patterns consists of two different figures that are meaningful for arithmetical operations. The test can be thought of as a matrix consists of n-different figures. In the test type of geometrical addition (GA) two different figures in column one and column two (or in rows one or two) makes the figure in column three by applying the geometrical operator.

In the test type of finding relationships FR figures are arranged in such a way that six objects are used to find the general rules of prediction and two objects are used to find the possible solution. Figure 4.103 shows examples of the tests type FR.

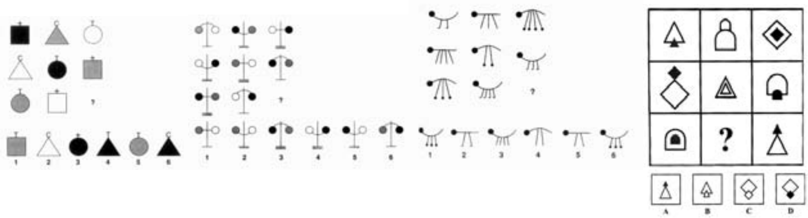


Fig. 4.103. Examples of the category of the matrices tests type FR

The category of spatial test is divided into the folding sheet test category, the cubic box test category, the bird view test category or the spatial transformation test category: $\ldots \triangleright \langle \nu_{\mathrm{VisT}} \rangle \triangleright \langle \nu_{\mathrm{SpT}} \rangle \triangleright \langle \nu_{\mathrm{FST}}, \nu_{\mathrm{CBT}}, \nu_{\mathrm{BWT}}, \nu_{\mathrm{BWT}}, \nu_{\mathrm{STT}} \rangle$.

In the category of the folding flat sheet the task is to find which of four three-dimensional figures can be produced by folding a flat sheet of specified shape.

In the category of spatial test type cubic boxes the task is given in the form of cubic boxes and the nets of these boxes (unfolded cut-outs). There are two kinds of questions. One kind shows a net and four boxes, labelled A, B, C and D. The task is to choose which one of these boxes could be made from the net. The second kind of questions shows a box and four possible nets. The task is to choose which one of the nets belongs to the box shown. The category of the cubic box is divided into coloured boxes and pattern boxes: $\ldots \triangleright \langle \nu_{\mathrm{VisT}} \rangle \triangleright \langle \nu_{\mathrm{SpT}} \rangle \triangleright \langle \nu_{\mathrm{CBT}} \rangle \triangleright \langle \nu_{\mathrm{CCBT}}, \nu_{\mathrm{PCBT}} \rangle$.

The category of 'bird's eye view' test is divided into the category of 'bird's eye view' of the side view and the category of the 'bird's eye view' of objects on the table. The category of 'bird's eye view' involves views of one or more shapes (e.g. cylinders, spheres, cones) on a table top. In the category of bird's eye side view test the table top has marked on it a square grid.

4.3. Categorical Learning

The proposed categorical learning is the task to induce the general description of the categories from the specific instances (phantoms) of the concepts. Learned concept on the bottom of the hierarchy of the categories is called a prototype. The prototype is the definition of the learned phantom (the visual concept) in terms of the symbolic names and characteristic features of the category (the phenomenal concept, the metalingual concept) given by its name. During learning the new case is evaluated in the context of all learned categories. The visual concept of the general category includes all prototypes of the specific categories. Learning of the visual object consists of the two stages. At the first stage the visual concept is learned. At the second stage the non-visual knowledge is learned. In this chapter the learning of the visual concept is presented. Learning of the visual concept of the object independently from other conceptual ingredients is a new approach in machine learning methods. All visual information that is extracted from the object is transformed into the symbolic representation called the visual concept. Such an approach makes it possible to concentrate on the visual aspect of the learned object. The visual concept φ is obtained during the learning process. It is assumed that the visual concept is uniquely described by the name α. During the learning process the set of phantoms $U^{\mathrm{P}} \subset U^{\mathrm{O}}$ that is representative of a given visual category is selected and next for each phantom $u_i \in U^{\mathrm{P}}$ the symbolic name η_i is obtained. As the result of the learning process a set of symbolic names $\varphi_\alpha = \{\eta_1, \eta_2, \ldots, \eta_n\}$ that represents the visual concept is obtained. The visual concept represented by the category v^p is called a prototype.

An object that belongs to the category v^p is defined by application of the rules. The phantom $\mathsf{u} \in \mathbf{u}$ is assigned to the category v^p based on the values of attributes a^i:

$$([a^i = A_i] \cdot \wedge \cdot [a^j = A_j] \ldots) \cdot \vee \cdot ([a^k = A_k] \cdot \wedge \cdot [a^l = A_l] \ldots) \Rightarrow \mathsf{u} \rhd v^p$$

An object that belongs to the category v^p is defined by application of the rules, for example,

$$[a^0 = A_0] \Rightarrow u \rhd v^p,$$

$$([a^0 = A_0] \cdot \vee \cdot [a^1 = A_1] \ldots) \Rightarrow \mathsf{u} \rhd v^q,$$

$$([a^0 = A_0] \cdot \wedge \cdot [a^1 = A_0]) \cdot \vee \cdot ([a^0 = A_1] \cdot \wedge \cdot [a^1 = A_1]) \Rightarrow \mathsf{u} \rhd v^r.$$

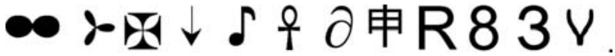
An object that belongs to the category v^p is defined by application of the set of rules, for example,

$$[a^0 = A_0] \Rightarrow u \rhd v^p,$$

$$([a^1 = A_1] \cdot \wedge \cdot [a^2 = A_1]) \cdot \vee \cdot ([a^1 = A_2] \cdot \wedge \cdot [a^2 = A_2]) \Rightarrow \mathsf{u} \rhd v^p,$$

$$([a^3 = A_3] \cdot \vee \cdot [a^4 = A^4] \ldots) \Rightarrow \mathsf{u} \rhd v^p,$$

where a^i denotes the symbolic name of the 'parts' of the visual object or the characteristic feature. The prototype is learned starting from the definition of the general category. The definition of the general category is expressed in terms of symbolic names. During generalization the symbolic name is translated into the string form L0_L1_…Ln, where the level Ln denotes the *n*th level of description of the archetype of the class. Learning of the figure pentagon is given as an example of the learning process. The symbolic name WL5[aaaaa][sssss] obtained during the reasoning process is transformed into the string form W_L_5_[a]_[s]. The concept defined by a set of rules is learned by starting from the definition of the general category. The general category is defined in the context of all learned prototypes. Let us assume that the first learned prototypes are



The general concept of the learned figure  is defined as
HUL = “W”, NAME = “ConvexObject”,
if[m_CHul = HUL]
{m_Name = NAME}.

The variable m_CHul denotes the symbolic name of the examined object obtained during the process of visual reasoning. The variable m_Name is the name of the prototype defined by the definition of that prototype. This definition well describes the differences (dissimilarity) among objects. All learned prototypes are concave objects. In the categorical learning, testing and learning processes are complementary. During testing of the learned categories where the figures  are given as an input, these figures will be assigned to the name ‘convex object’. The answer given by SUS is correct however the definition given in the previous stage is too general. The figure  is not distinguished from another two figures. In this situation there is a need to use a symbolic name in the definition of the prototype at the more specific level. The definition is as follows:

HUL = “W_L”, NAME = “Polygon”,
if[m_CHul = HUL]
{m_Name = NAME}.

In the next stage the additional figures are learned , , . In this case there is a need to use the symbolic name in the definition of the prototype at the more specific level. The definition is as follows:

HUL = “W_L_5”, NAME = “Pentagon”,
if[m_CHul = HUL]
{m_Name = NAME}.

In the next stage the additional figures are learned . Their symbolic name, at the specific level, is given in the form W_L_5_[a]_[s]. The symbol [s] denotes the term {sssss}, where the symbol s can have one of the values from a set of normalized sides (l-large, m-medium, s-small and e-very small). The symbol [a] denotes the term {aaaaa}, where the symbol ‘a’ can have one of the values from a set of normalized angles (a, acute; o, obtuse and p, right).

During process of learning the specific description in the form _[a]_[s] is attached to the name of the object for further reasoning. For example,

HUL=“L5”, HULSIDES=“[mmmmm]”,
HULANGLE=“[oaapo]”, m_Nazwa=“Pentagon”,
if [m_CHul=HUL]
{m_Name=NAME+HULSIDES+HULANGLE}.

There is also possibility to define all pentagons that are given by the combination of the symbols {sssss} and {aaaaa}. For example, the object called “Pentagon_Ideal” is given by the symbolic name L5[mmmmm][ooooo]. However the number of definitions will grow very rapidly and there is also problem with the checking errors when the values of parameters are misinterpreted.

In the second approach the new sub-specific classes are derived from the pentagon class. For example, by applying description ‘L5[nP], where nP denotes a number of right angles of the pentagon, the pentagons will be divided into five groups described by symbolic names ‘L5[0P]’, ‘L5[1P]’, ‘L5[2P]’, ‘L5[3P]’, ‘L5[4P]’. The new characteristic feature, such as ‘symmetry’, can be used to derive the additional sub-specific classes. The rules are given in the form:

HUL=“L5”, NAME0=‘pentagon’, NAME1=‘pentagonNS’, NAME2=‘pentagonS’, NAMETYPE[0]=‘P0’, NAMETYPE[1]=‘P1’, NAMETYPE[2]=‘P2’, NAMETYPE[3]=‘P3’, NAMETYPE[4]=‘P4’,

if[m_CHul=HUL]
[m_Name=NAME0
if[m_Sym=0]
[m_Name=NAME1+NAMETYPE[i]]
else
[m_Name=NAME0+NAMETYPE[i]]]].

In the categorical learning, testing and learning processes are complementary. Figure such as a pentagon is uniquely described by the geometrical properties of the object. In the case of mathematical object such as the graph of the function the definition of the visual concept can be learned from phantoms that can be generated from well-described formulas. All phantoms that are needed for learning can be analysed and the definition can be verified by applying the generated phantoms. The definition of the special class of the graphs of a function called ‘peak’ is given as an example of learning.

Fig. 4.104. Examples of graph function called ‘peak’

The definition of the prototype of function ‘peak’ shown in Fig. 4.104 is given as:

HUL=‘L3’, NoRes=2, RES=‘M1’, NAME=‘FunkcjaPeak_”;
If[m_CHul=HUL]
[If[m_Res=RES]
[if[NoResR=0]
[[m_Name=NAME]]]]

The definition of the prototype gives good results in the case of the symmetrical figure of the type of ‘peak’. In the case of the non-symmetrical object  the object is assigned to the prototype called ‘peak’. However an object that is defined by the learned definition of ‘funkcja-peak’ and is non-symmetrical can also be called ‘funkcja-peak’. In this situation there is a need to find the characteristic feature that makes it possible to discriminate between the prototype called ‘funkcja-peak’ and instances of the prototype that can be called ‘quneiform’.

The definition of the complex object can be given in terms of the concavities, holes and thin parts. In some cases the definition of the learned object is very similar to the description of the object given by human subject. For example, for the cipher 8 the description can be as a symmetrical concave object having two holes and two concave residues. The additional description is given in terms of the complex object as ‘consists of two circular parts with one hole’. The learned definition that does not specify the exact shape of the cipher 8 is formulated by application of the symbolic name on the general level of description. This definition does not specify the type of holes. These holes can be any curvilinear shape. In the case when there is a need for the definition of the specific font of the cipher ‘8’ the symbolic names at the specific level are given in the definition. The symbolic names are expressed in the form of the SUS representation. Understanding of the definition of the cipher ‘8’ given in this example does not require understanding of the meaning of symbolic names.

C_HUL=“Q_M(2,2)”, RES1= “Q_L3”, RES2= “Q_L3”, PAR1=“M1”, PAR2=“K”, CON1=“M1”, CON2=“M1”, HOL1=“K”, HOL2=“K”, m_Nazwa= “cipher_8”.

The specific description is required when there is a need to give the symbolic name in the specific form:

C_HUL="Q_M(2,2)[L4]", RES1="Q_L3_O", ES2="Q_L3_O", PAR1="M1_(Hh), PAR2="K_K1", CON1="M1_Hm", CON2="1_Hl", HOL1="K_K1", HOL2="K_K1";
m_Nazwa="cipher_8A";

The rules of the definition are given in the following form:

if[[m_ResP1=RES1 & m_ResP2=RES2] ||
[m_ResP2=RES1 & m_ResP1=RES2]]
[if[[m_PartP1=PAR1 & m_PartP2=PAR2] ||
[m_PartP2=PAR1 & m_PartP1=PAR2]]
[if[[m_Nazwa1Con=CON1 & m_Nazwa2Con=CON2] || [m_Nazwa2Con=CON1 & m_Nazwa1Con=CON2]]
[if[[m_NazwaHoleOne=HOL1 & m_NazwaHoleTwo=HOL2] || [m_NazwaHoleTwo=HOL1 & m_NazwaHoleOne=HOL2]]
[MakeName(m_Nazwa,m_ResP1,m_ResP2)]]]]]]

In the case when an object is made from the different parts there is a need to learn the concept of the different parts. Learning and application of knowledge of the learned parts of the object is called the conceptual magnification. Part that is partially invisible can be learned independently. In the case of the assembled object the invisible part of the object need to be identified during the examination of an object part by part. Conceptual magnification makes it possible to learn independently parts that can not be visible at given scale. The conceptual magnification is used to solve an interpretational problem. The conceptual magnification uses the background knowledge in interpretation of the object. For example, understanding the concept of man requires understanding that the human body consists of parts such as hands, head or legs. Conceptual magnification makes it possible to reveal details such as fingers that are invisible in Fig. 4.105.

Fig.4.105. Parts that are learned independently and revealed during conceptual magnification

References

1. Berlin, B., Breedlowe, D.E., and Raven, P.H., General principles of classification and nomenclature in folk biology. *American Antropologist*, 1973. **75**(1): pp. 214–242
2. Stepp, R., and Michalski, R.S., Conceptual clustering: inventing goal-oriented classification of structured objects, in *Machine Learning. An Artificial Intelligence Approach.*, R.S. Michalski, Carbonell, J.G., and Mitchell, T.M., Editors. 1986,Los Altos: Morgan Kaufman Publishers, Inc., pp. 471–498
3. Rosh, E., and Mervis, C.B., Family resemblance: studies in the internal structure of categories. *Cognitive Psychology*, 1975. **7**: pp. 537–605
4. Cimiano, P., Hotho, A., and Staab, S., Learning concept hierarchies from text corpora using formal concept analysis. *Journal of Artificial Intelligence Research*, 2005. **8**: pp. 305–339
5. Op de Beeck, H., and Wagemans, J., Visual object categorization at distinct levels of abstraction: a new stimulus set. *Perception*, 2001. **30**: pp. 1337–1361
6. Rosh, E., and Mervis, C.B., Gray, W.D., Johnson, D.M., and Boyes-Bream, P., Basic objects in natural categories. *Cognitive Psychology*, 1976. **8**: pp. 382–439
7. Behl-Chadha, G., Basic-level and superordinate-like categorical representations in early infancy. *Cognition*, 1996. **60**: pp. 105–141
8. Gauthier, I., Anderson, A.W., Tarr, M.J., Skudlarski, P., and Gore, J.C., Levels of categorization in visual object recognition studied with functional MRI. *Current Biology*, 1996. **7**: pp. 645–651
9. Jolicouer, P., Gluck, M.A., and Kosslyn, S.M., Pictures and names: making the connection. *Cognitive Psychology*, 1984. **16**: pp. 243–275
10. Snodgrass, J.G., Vanderwart, M., A standardized set of 260 pictures: norms for name segment, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 1980. **6**: pp. 174–215
11. Marsolek, C.J., Dissociable neural subsystems underlie abstract and specific object recognition. *Psychological Science*, 1999. **10**: pp. 111–118
12. Colaruso, R., and Hammil, D., *Motor Free Visual Perception Test*. 2003, New York: Academic Therapy Publications
13. Gardener, M.F., *Test of Visual-Perceptual Skills*. 1996, Hydesville, CA: Psychological and Educational Publications
14. Roid, G., *Stanford–Binet Intelligence Scale*. 2003, Chicago: Riverside
15. Gluting, J., Adams, W., and Shwslow, D., *Wide Range Intelligence Test*. 1999, Chicago: Wide Range
16. Sala, D.S., Gray, C., Baddeley, A., and Wilson, L., *Visual Patterns Test*. 1997, Suffolk: Thames Valley Test Company
17. Meyers, J.E., and Meyers, K.R., *Rey Complex Figure Test and Recognition Trial*. 1996, Odessa, FL: PAR
18. Huber, H., *Rorschach Psychodiagnostic Test*. 1948, Berne: Hans Huber

5. Visual Thinking: Understanding

5.1. Understanding in the Context of Shape Understanding System

To find if the subject understands a sentence, a task, or a phenomenon, there is a need to evaluate the response to a given task. One of the methods of evaluating the response, that is the result of understanding of the task, is to evaluate if an appropriate action was undertaken by the subject (or robot). Performing rational actions by subject indicates that subject understands the task.

Similarly, it is assumed that the shape understanding system (SUS) should be able to perform actions that are evaluated as rational actions in order to demonstrate understanding of the perceived object. The term a visual object used in this context has a broad meaning range from the simple visual symbols such as a mathematical symbol, a written text, an engineering scheme to the complex real-world objects. In this context the term "understanding of the visual task" will be often used instead of "understanding of the visual object." The ability to visualize forms in the mind enhances ability to understand both existing objects and objects that may not yet have been seen. The ability to visualize makes it possible to spatially analyze more detailed visual problems [1]. For example, sketching is based on seeing and visual thinking through the process of seeing, imaging, and representing. Seeing is our primary sensory channel because so much information can be gathered through our eyes. Imaging is the process used by the mind that takes the visual data received by our eyes to form some structure and meaning. The mind's eye initially creates the image whether real or imagined, and these are the images used to create sketches. Representing is the process of creating sketches of what our mind see. Seeing and imaging is a pattern-seeking process where the mind's eye actively seeks those features that fit within our interest, knowledge, and experience.

Z. Les and M. Les: *Visual Thinking: Understanding*, Studies in Computational Intelligence (SCI) **86**, 241–399 (2008)
www.springerlink.com

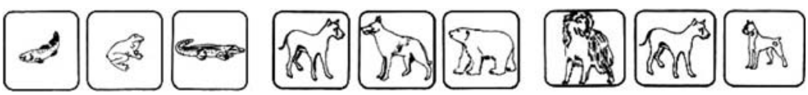
Understanding even a simple task requires performing complex mental operations. The task can be given by the user or can be formulated by the system based on the perceived object. For example, understanding task given in the form of the spoken sentence that requires performing grouping of the different objects (see Fig. 5.5) involves, among others, understanding spoken words and concepts such as "similar objects," "grouping similar objects," "selecting objects," "moving objects," "transforming objects," "naming objects," and "performing an appropriate action." To evaluate if the system understands given task there is a need to evaluate if an appropriate action was undertaken. In the case of tasks that require grouping of different objects the evaluation is not a very difficult problem whereas the task formulated by a system as the result of perceiving the visual object can be sometimes difficult to evaluate. For example, the system after perceiving objects shown in Fig. 5.5 can begin performing grouping of different objects. However, in this case criteria of grouping are not known. Evaluation of this type of action is difficult because in many cases, there is no unique solution in selection of the appropriate action. Undertaking a given action depends on the contextual knowledge that the system possesses which can come from perception of other objects or can be obtained as the result of the similar perceptual experience in the past. For example, seeing the letter "P" can indicate undertaking an action "read the word" or it can indicate undertaking action "park the car" (it is parking near by).

SUS assumes that the world consists of learned objects and other not known objects. In the case when an examined object belongs to the general category of objects that were learned based on a few examples, SUS interprets it as a possible object and describes it as "it can be x." When an examined object belongs to the learned specific category the answer is "this is x." SUS is able to find occluded objects or the incomplete figure by learning the visual concept from the partially occluded objects. The occluded objects are interpreted in the same way as a part of the object. For example, a triangle can be interpreted as a part of arrow, or as an occluded arrow. Learning of the occluded objects is the topic of further research focused on the understanding of the distorted objects and is not included in this book.

The result of understanding of the visual object (visual task) is undertaking of an appropriate action, denoted as, $U(u^L) :\rightarrow a$, where U denotes understanding process, u^L denotes a phantom (perceived object), and a denotes an action that is undertaken as the result of understanding process. This schema an object->action, or a situation->action can be used to represent the process of interaction of the system with an environment. For example, understanding a road sign such as “stop” sign means to perform an appropriate action when the sign is perceived on the road.

Understanding of the sentence that refers to a visual object often invokes the visual representation of the object, but it can be based on the nonvisual knowledge of the categorical chain. For example, the sentence “Explain the difference between a triangle and a rectangle” refers to the categories of geometrical figures and there is no need to imagine a triangle and a rectangle to be able to give explanation about the differences between them. All knowledge that is used to explain the differences is obtained from the knowledge schema of the category of geometrical figures.

Understanding of the sentence “Explain the difference in shape between the letter ‘B’ and the letter ‘L’” requires finding the visual concept for each letter and next the symbolic names that are visual representatives of structural archetypes of both letters are assigned to these letters. Symbolic names of the structural archetypes are used to find the differences between objects. Understanding of the sentence “Explain the difference in shape between the object that you see and the letter ‘L’” requires at first to find the visual concepts of both objects. The visual concept of the first object is found by applying the visual reasoning whereas the visual concept of the letter “L” is invoked from memory. Next, the structural archetype for each object is selected and symbolic names of structural archetypes are used to find the differences between objects. Understanding of the sentence “Explain the difference between those two objects that you see” requires interpreting these objects in terms of a figure, a letter, a sign, or a real-world object. At first, the name for each object is found and next the difference between these two objects is established based on the meaning of each category. The difference is found by specifying the lower categorical levels for which the differences between members of those two categories are significant. For example, let us assume that two objects that are perceived are named as the apple and the lion. The solution to this problem is found by the reference to the categorical chain. The answer is the first object represents a plant whereas the second one represents an animal. To explain understanding of the differences among objects of the different categories the example that is shown in Fig. 5.1 is given. Figure 5.1 shows



a b c

Fig. 5.1. Example of the different intelligence tests, each one with three different objects

three intelligence tests with three different objects. One of the three objects in the test is shown to the observer to remember and next the observer needs to find one that is exactly the same that this which was shown. In the first test shown in Fig. 5.1a all objects belong to the category of animal and the differences among objects are found at the lower categorical level (the category of fish, the category of frogs, the category of crocodiles). In the second test shown in Fig. 5.1b all objects belong to the animal category; two of them are members of the different dog categories and another one is a member of the bear category. The differentiation is much easy task when objects that are shown are members of the different categories of animals. The third test shown in Fig. 5.1c is more difficult because the differentiation occurs at the lowest categorical level. All objects are animals, and all objects are dogs. To solve this problem the knowledge of the specific dog categories is required. The subject who has knowledge of the different breeds of dogs will find this task much easier than the subject who does not have that knowledge. This task can be also solved without reference to any real-world object by matching the visual objects (the same shape).

Understanding and solving the task by machine can only to some extent be compared to our human understanding. In the first chapter of this book it was shown how the term “understanding” was differently understood by philosophers and scientists. There is no agreement how to define understanding process and which aspect of understanding can be used to give the best description of it. Similarly thinking, shown in Chap. 1, is very differently understood in the scientific world. In this book all processes that are performed in the context of perception of the visual object or understanding of the visual task will be considered as part of thinking processes. The necessary condition of the thinking process is that the system needs to be complex enough to perform this process spontaneously.

In this book understanding is defined in the context of the problem-solving issues. Thinking is understood as a complex reasoning process that leads to understanding. Understanding presented in this book is considered in two different aspects. In the first aspect, the understanding process is connected with learning, acquiring a new knowledge and memorizing what was learned. The second aspect is related to naming, recognition, and solving problems. In this approach a given problem such as understanding of the perceived object is interpreted and solved in the context of knowledge that was previously learned.

The thinking process is the process that is always present during human life and is connected with the normal activities of our brain. It will be topic of the next sections.

5.2. Thinking and Visual Thinking

Understanding and thinking are processes that cannot be separated. In this book the terms understanding and thinking are used to denote often very similar aspects of processes that are connected with interpretation of the perceived object. Thinking is the process which is always connected with activities of our brain. Thinking is understood as a complex reasoning process that leads to understanding. Thinking involves many processes that are responsible for transformation of the input data (visual object, visual task) into different forms (symbolic or visual) that leads to understanding of the phenomenon or solving the problem.

Visual thinking refers to processes that are connected with transformation of the visual object (visual transformation) into the action that is the result of understanding. As it was described in the previous chapters thinking (visual thinking) need not lead to understanding. We can think about something to imagine it or sometimes only to have awareness of it. However, in the present book the term thinking will be used to denote the process that leads to understanding. Visual thinking is part of the thinking processes that can lead to understanding of the complex problems. However, the term visual thinking will be often used to denote the process that leads to understanding of visual forms. The visual thinking process consists of the different steps that involve transformation of the different forms of data stored in memory. Visual thinking is part of the thinking process that deals with the visual material. Visual thinking process can start when an object is perceived and one of the results of the thinking process is assigning the perceived object to one of the object categories. On the other hand, the task that is given to the system can require producing an image (visual object). The image is the result of the imaginary process and is a part of the thinking process. The image (visual object) can be also used during intermediate stage of the visual thinking process.

As it was described in the previous chapter there is no unique definition of the thinking process. According to Arnhaim, the machine needs to possess the following capabilities in order to be able to engage in the thinking process, it should:

- Respond to the categories of things and disregard of the unique object
- Solve problems by means of they perceptual concepts
- Connect items of their environment by relations that lead to solution of a given problem
- Suitably restructure the situation facing them
- Transfer the solution to different but structurally similar instances

As it was presented in the first chapter, thinking is the process that uses knowledge during interpretation of the visual object. Visual thinking process starts when an object is perceived and the result of the thinking process is to understand of the perceived object. Visual thinking is the process of "mental" operations on visual concepts (categories) and can be the source of a new knowledge. The knowledge in the system is represented by the complex structure of the linked categorical chains.

The image (visual object) can be produced by the system at any stage of the visual thinking process. The visual object (image) can be the result of imagery transformation during the reasoning process and can be given as the result of the visual thinking process during communication session. It is assumed that all transformations performed on categories of the visual objects or their parts are part of the visual thinking process. Thinking is considered to mediate between inner activities such as imaginary transformations and external stimuli. Thinking plays a key role in problem solving. In the case of solving the relatively easy problem, thinking can be regarded as the reasoning process. As it was shown in the first chapter thinking was often understood very narrowly as a sort of reasoning. In our approach thinking and understanding is regarded as the process that is connected with problem solving. In this book the term "understanding" will be used to denote the result of the thinking process and will be often used to denote the thinking process itself. This is justified in the case of the convergent thinking where understanding is the result of the thinking process. The term "thinking" will be also used to underline processes that are connected with imaginary transformations. Imaging is the process of producing the image during the thinking process and is one of the steps in obtaining the understanding of the visual object or the phenomena. For example, reading the sentence "the leaf is on the table" may lead to imagining the leaf. Shape of the leaf that is the result of the imaginary process can be given as a visual illustration of the meaning of the sentence, or as the chain of the thinking operations that can lead to the creative artistic process. Without any additional information the sentence "the leaf is on the table" is interpreted that the leaf from any leaf categories such as tree leaf is on the table. The system cannot generate image of the general leaf category and the leaf that is generated during thinking process is the representative of the most common shape of leaves that system learned. The term "the most common shape" refers to the visual concept of the leaf category and indicates that the object that is a member of the leaf category is interpreted as the visual object, part of tree, given by the visual representative of the leaf category.

We have decided not to provide a more formal definition of the visual thinking process. At this stage processes that are connected with the visual thinking process are represented by a sequence of transformations.

There is still the area of research in the domain of understanding and thinking where further research is needed to be able to define the visual thinking in a more formal way. In this book the main emphasizes is laid on the visual thinking processes and the problem of the nonvisual thinking is presented only in the context of issues connected with visual thinking. Also, the problem of thinking needs to be formulated and solved before the formal definition of the visual thinking can be given. The visual thinking is part of the thinking process and the definition of the visual thinking process needs to refer to the definition of the thinking process. Because the formal definition of the thinking process does not exist we will describe thinking process in the context of different aspects of the thinking process presented in Chap. 1 as well as in the context of the learning, reasoning, and problem solving outlined in other chapters of this book. Someone can argue that the process that we call "visual thinking process" is different from the "visual thinking process" that is described in the context of the human mental processes. It would be difficult to prove that what we call the visual thinking process has the same meaning as the meaning of this term described in the context of the different scientific disciplines. However, we believe that building the complex visual thinking machine can only be possible by applying an approach presented in this book. In this book we will focus on thinking processes that are connected with problem solving.

5.3. Visual Thinking as a Problem Solving Process

As it was described in Sect. 5.2 the visual thinking processes lead to understanding. Understanding of the task is connected with problem solving. The task is solved during the thinking process. At the beginning of the thinking process the task is transformed into problem that needs to be solved. The visual thinking process consists of many different stages and at each stage the specific problem is solved. The simplest thinking process involves transformation of the data during the visual reasoning process. The thinking process of the understanding system that is focused on the problem solving is limited to solving problems that appear as a result of perception of the visual object or the tasks that are given by users. The important part of the thinking process is the reasoning process. Depending on the type of the visual problem that is solved many different forms of the transformations of the data during the reasoning processes are involved. The visual thinking process is very complex and even the simplest form given by the visual reasoning involves application of the complex image transformations, shape categories, and visual object categories as described in Chaps. 2–4.

The thinking process consists of many stages where different image transformations are applied to transform perceptual data. One of the transformations is the sensory transformation $\Im$: $\Im(u) \Rightarrow \coprod$ that transforms the perceived object (phantom) u into a set of critical points $\coprod$. A set of critical points $\coprod$ is transformed into the symbolic name η: $R\langle\coprod\rangle \Rightarrow \eta$ during the reasoning process by a sequence of image transformations given in the following form: $\Im(u) \Rightarrow \coprod \Rightarrow R\langle\coprod\rangle \Rightarrow \eta$. The image transformations used in reasoning process were described in Chap. 3.

In this book it is assumed that all problems that system needs to solve can be formulated in the form of the problem-solving tasks. These tasks such as naming and recognition of objects of the different categories (figures, signs, letters, real-world objects), the visual diagnosis, the data analysis, or the visual intelligence tests are examples of the different problems that occur in the different areas of human activities.

The visual problem solving is often a very complex task that involves many stages of thinking processes. In the problem formulation, visual thinking involves the problem transformation $\Delta(v) \Rightarrow u$, where v is a visual problem category and u is a phantom (visual object) that represents a given problem. The problem can be also described in the form of the linguistic description or it can be given in the form of the objects of the real-world scene. The problem given by the linguistic description is transformed into a visual object (graphical representation) by transforming it to one of the schematic representations. The schematic representation that uses the visual symbols (e.g., engineering schema), can show only some aspects of the real-world object or can be "realistic" visual representation of the object. In solving the real-world problem instead of the problem transformation, the image $\wp$ that is representative of the real-world phenomenon, is transformed into a phantom u by perceptual transformation $P(\wp) \Rightarrow u$ and next, by applying the imagery transformations and visual inference, the solution is found.

Thinking process as a problem solving can be described by the sequence of sub-processes and expressed as follows: $\Delta(v) \Rightarrow u$ -> $\Im(u) \Rightarrow \coprod$ -> $R\langle\coprod\rangle \Rightarrow \eta$ -> $T\langle\eta\rangle \Rightarrow \coprod \ldots R\langle\coprod\rangle \Rightarrow \eta$ ->[$\eta \in \varphi^{\alpha}$] $\Rightarrow a \rhd \vartheta^{i}$, where at first "the problem transformation" $\Delta(v) \Rightarrow u$ transforms a given member of the problem category into the visual form (phantom), next the sequence of transformations $R\langle \amalg \rangle \Rightarrow \eta \; T\langle \eta \rangle \Rightarrow \amalg \; R\langle \amalg \rangle \Rightarrow \eta$ $R\langle \amalg \rangle \Rightarrow \eta \rightarrow T\langle \eta \rangle \Rightarrow \amalg \ldots R\langle \amalg \rangle \Rightarrow \eta$ transforms the internal representation given as a set of critical points into the symbolic names (image transformations), and at the end the solution is obtained by applying the visual inference. In solving the complex real-world problem a number of sub-processes can be very big. These sub-processes will be described in the following sections of this chapter.

As it was described, problems that the system needs to solve are formulated in the form of the problem-solving tasks. One of the simplest tasks is the naming task. During the naming process the phantom u is transformed into a set of critical points $\amalg$ and next into a symbolic name η. The symbolic name is used to find a category of the perceived object. The naming process is expressed in the form of the visual inference rule $[\eta \in \varphi^{\alpha}] \Rightarrow a \rhd \vartheta^{i}$, where η is the symbolic name obtained in the reasoning process, φ^{α} is the shape category (the visual concept), and $a \rhd \vartheta^{i}$ denotes the naming process. This type of thinking process can be written in the form of the sequence of transformations as follows: $\Im(u) \Rightarrow \amalg \Rightarrow R\langle \amalg \rangle \Rightarrow \eta \Rightarrow N\langle \eta \rangle \Rightarrow \vartheta$, where $N\langle \eta \rangle$ denotes the naming process given by the visual inference rule. The sequence of transformations will be called the categorical transformations and denoted as $C(u_i) \Rightarrow \vartheta i$. The naming process is described in more detail in the following sections of this chapter.

The visual diagnosis is similar to the naming process where instead of the name of the category of object, the name of the illness category is attached. In visual diagnosis, cells and organs with pathological changes are described by the pathological symptom category. The pathological symptom category (the visual concept) φ^{h} is used in the visual inference to find the illness category H^{i}. The inference rules are expressed in the following form $[\eta \in \varphi^{h}] \Rightarrow h \rhd H^{i}$. The categorical transformation is described as $C(u_i) \Rightarrow H^{i}$, where H^{i} is the illness category. The category of illness is linked with the category of treatment of illness so the diagnosis is connected with the recommendation of the treatment. The visual thinking process can be also present in the process of recommendation of the treatment.

In analogical reasoning, thinking process involves obtaining the visual category for each phantom that represents similar objects as well as establishing relationships among categories. At first, the phantoms u_1 and u_2 are transformed into categories described by categorical transformations: $C(u_1) \Rightarrow \vartheta_1$ and $C(u_2) \Rightarrow \vartheta_2$. Next, the conceptual similarity relation $\rho(\vartheta_1, \vartheta_2)$ between category ϑ_1 and category ϑ_2 is found. The category ϑ that fulfills the relation $\rho(\vartheta_3, \vartheta)$ is obtained and the phantom u that is representative of the category ϑ is selected as the solution.

Assembling tools is one of examples of the problem solving that involves application of the different thinking processes. For example, the visual scheme of a category of tools is used for solving the problem of assembling the spade from the three parts. This task is formulated as follows: having parts $u_1,...,u_n$ make the complex object u given by the name α. The name α is used to find the category of the object ϑ and its visual concept φ^{α}. Having the category of object ϑ, the knowledge concerning parts and assembling process can be obtained from the knowledge schema of this category. It is assumed that this knowledge was previously learned. Based on this knowledge each part is identified during thinking processes and the appropriate parts are selected from the given parts $u_1,...,u_n$. At the end, the assembling process is represented as a sequence of "events" (transformations) that leads to assembling the final object. In this example, for each part, during the thinking process, transformations $\Im(u) \Rightarrow \coprod \Rightarrow R\langle \coprod \rangle \Rightarrow \eta \Rightarrow N\langle \eta \rangle \Rightarrow \vartheta_i$ are applied to obtain visual categories of the part u_i.

Visual process control is thinking (observing) about the changes of the visual aspects of the process. The changes of shape during the certain interval of time can "produce" the characteristic sequences of shapes that are characteristic for changes of the process. The sequence of shapes is represented as a sequence of the symbolic names $\eta^1, \eta^2,..., \eta^3$. The failure category (visual concept) φ^h is used in the visual inference to define the failure (critical points) of the process P^i. The rules of visual inference are expressed as follows: $[\eta^1, \eta^2,..., \eta^n \in \varphi^p] \Rightarrow p \triangleright P^i$.

Solving problems in statistics is thinking in the form of images that are obtained during data visualization. Points given in the scatter plot are called a set of 2D dot patterns. The numerical data given as an input in the form of a set of points $\prod = \{(x_1, y_1),...,(x_i, y_i),...(x_N, y_N)\}$ are transformed into the image $\wp : V(\prod) \Rightarrow \wp$ by applying the visual transformation and the image $\wp$ is transformed into the phantom u by the perceptual transformation: $P(\wp) \Rightarrow u$. Thinking during the visualization process can be described as transformation that transforms data points into the phantom u and is denoted as: $V(\prod) \Rightarrow \wp \Rightarrow P(\wp) \Rightarrow u$, where $V(\prod)$ is the visual transformation and $\wp$ is an image; or by applying the graphical transformation $G(\wp)$, that is denoted as $V(\prod) \Rightarrow \wp \Rightarrow G(\wp) \Rightarrow u$. Next, the phantom is used to find the visual category of the statistical problem that is denoted as $\Im(u) \Rightarrow \coprod \Rightarrow R\langle \coprod \rangle \Rightarrow \eta \Rightarrow N\langle \eta \rangle \Rightarrow \vartheta$. The selected statistical model is next applied to perform statistical analysis and finally, during the nonvisual thinking process, the data are interpreted in terms of the statistical parameters.

Thinking can lead to formulate the problem by asking outer or inner questions. For example, when seeing the unknown object the inner question is "what is this?" and SUS needs to find the answer to understand the problem. When formulating diagnosis of the illness, SUS needs to communicate the findings to the doctor and ask for another data if needed. Problem can be formulated by user and given to the system as a task or can be formulated by SUS as the result of its perceptual activity. The task that is performed by the system can be the result of perception of the visual object (e.g., sign STOP) or can be given by the user.

Visual processes connected with thinking can be described in terms of problem solving processes and are topics of the following sections of this book.

5.4. Problem Solving

Problem solving process is one of the mental processes where thinking plays a key role. Problem solving can be viewed as the process of transformation some initial states of the world into a goal state by application of a sequence of known operators. Operators define a search space that must be explored to discover how the goal state can be achieved. There are classes of problems that require the visual representation in formulation of the problem as well as in finding of the solution. However, many problems can be given in the form of the symbolic description without the need for visual representation. There is a class of problems such as mathematical problems that can be precisely defined [2]. Solution to a mathematical problem contains the following parts:

- Complete specification of givens; that is, a unique given state from which the goal can be derived via sequence of allowable operations
- Complete specification of the set of operations to be used
- Complete specification of the goals
- An ordered succession or a sequence of problem states, starting with the given state and terminating with a goal state, such that each successive state is obtained from the preceding state by means of allowable action

The problems contain information concerning givens, actions, and goals. The solution of a problem can be defined in terms of a sequence of states (terminating with the achievements of goal). It is very useful to represent both the possible sequences of actions and the possible sequences of states in a common diagram called a state-action tree.

As it was described in previous sections problem is solved during the thinking process. Thinking involves many different processes that are responsible for the transformation of input data given in the form of the phantom into many different forms, symbolic or image that lead to understanding of the phenomenon or solving the problem. Visual understanding that is focused on application of imagery transformations during problem solving process is called the visual problem solving.

Problem solving requires different representations of the problem. In this book we will focus on the representations that are given in the form of the verbal description, a schematic visual representation, or realistic representation of the world phenomena. The verbal description can be translated into the visual form and often can be very useful in solving problems. In the case when the visual representation is too realistic, showing parts that are not relevant to a given problem, it can lead to misinterpretation of the task and finally to fail in solving a given problem.

The term *problem solving* is defined in different ways in the different areas of human activities. For example, the term problem solving in artificial intelligence has been used to denote the disparate forms of intelligent action to achieve well-defined goals. Each area of knowledge such as mathematics or physics has its specific method of solving problems that are connected with the specific knowledge of this area. To solve the problem there is a need to understand the concept of the area to which problem belongs, and next to transform the problem into the form that can be easy to solve. For example, a mechanical problem formulated as follows: “What constant force will cause a mass of 3 kg to achieve a speed of 30 m per second in 6 s, starting from rest?” requires understanding the concepts from area of mechanics. Understanding concepts such as “force,” “mass,” "speed" requires knowledge from the area of physical sciences. In SUS knowledge is represented in the form of linked categories that are accessible during thinking process.

To find the solution to the problem there is a need for finding a proper representation of the problem. For example, the fox, goose, and corn problem is given in the form of the linguistic description as follows: "A man (M), a fox (F), a goose (G), and some corn (C) are together on one side of the river (straight line) with a boat (B). The goal is to transfer all of these entities to the other side of the river by means of boat, which will carry the man and one other entity. The fox and goose cannot be left together, nor can the goose and the corn." To solve the problem the visual representation in the schematic or the realistic form can be used. However, the realistic representation does not help in finding the solution. An example of the representation that can be useful in solving this problem is given in the form of symbolic representation "M F G C B |." The solution can be obtained by transforming beginning state of the problem into the goal state as follows:

Given M F G C B | Goal | M F G C B
Solution:
Given M F G C B | FC | MG B MFC B | G C | MFG B
MGC B | F
G | MFC B MG B | FC | MFGCB

The mathematical problem can be represented in either symbolic or diagrammatic form. Symbolic form refers to the expression of information in terms of words, letters, numbers, mathematical symbols, or symbolic notation. Problem can be formulated as a problem of path-finding and can be solved by applying the graph representation. Example of such an approach is described in [3]. Verbal symbolic representation is probably more important than visual diagrammatic representation in problem solving and in abstract thinking. Even when diagrams in the solution of problems are employed they are usually labeled with symbols that are attached to the points, lines, and angles. The simplest and most frequent step in symbolic representation of the information in problems is to choose some symbols (or a sequence of symbols) to stand for a concept.

Many practical problems are stated in terms of physical objects or phenomenon and can be transformed into mathematical expression. The mathematical problem consists of givens and operations. Givens refer to a set of expressions representing objects, things, assumptions, definitions, or facts. Operations refer to the actions that allowed to be performed on the givens or on expressions derived from the givens by some previous sequence of actions [2]. The goal of the problem is a terminal expression. For example, consider the problem of finding the value of x, given the expression $5x + 9 = 19$. The goal expression is in the form $x = 2$ where 2 is the correct number that was to be found. There may be one or more than one correct solutions to the problem.

To understand and solve many problems there is a need to find the proper visual representation. Some problems can be solved by using the schema where the real-world objects are represented by the visual symbols. For example, in the schema of the electric circuit the real-world object such as resistor is represented by the visual symbol. Industrial processes are often represented by schematic representations that applied the cross section or even visual symbols. The visual representation can be close to realistic representation as shown in Fig. 5.2 or can be more abstract as shown in Fig. 5.3.

In the visual problem solving the visual representation is used to understand and formulate the problem (see Fig. 5.4). For example, a schema of electrical circuit (Fig. 5.4a) is used to formulate the problem in the form of the mathematical model. The visual representation is very often used in the case of formulation of the mechanical problems (Fig. 5.4b–e).

a b

Fig. 5.2. The problem represented in realistic form

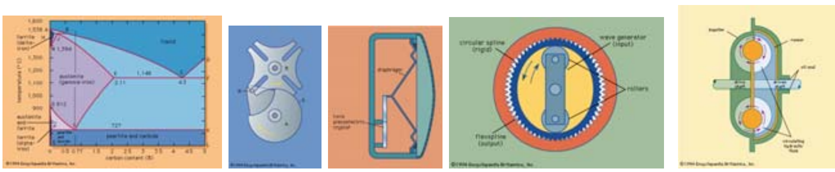
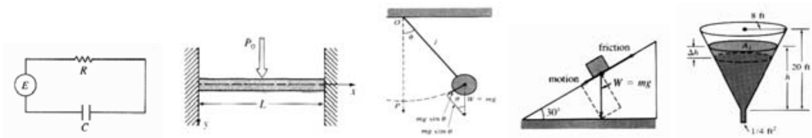


Fig. 5.3. The schematic visual representations that are used to understand the visual problem



a b c d e

Fig. 5.4. Examples of visual representations used to formulate and solve the different problems

Depending on the utilization of the symbolic and the visual representation, problems can be divided into three different classes: the problem given in the form of the symbolic representation, the problem given in the form of the visual representation, and the problem given in the form of both symbolic and visual representations.

5.4.1. Problems Given in the Form of the Symbolic Representations

In this book it is assumed that all information is given by the visual means. It is also assumed that the information that is obtained through auditory channel can be transformed into the visual form. For example, speech can be transformed into the written text, and music can be transformed into the visual form of the signal and next into the musical scores. However, the symbolic visual representation applies visual symbols as the means to perform nonvisual transformation. Problem is often stated originally in some linguistic form, often relying upon verbal language. The first step in solving such a problem is to translate it to a more adequate representation applying knowledge of the mathematical expression.

One of the simplest tasks, given in the form of combination of words and mathematical expression, is the task to perform the arithmetical operation (e.g., multiply 2 by 3). The factor that characterizes problems is a factor that is called the problem difficulty and indicates the difficulty to solve the problem. In solving the simple problem such as multiplication of the two numbers the problem difficulty can be differently defined for the human subject and for the machine. For the human subject who is doing multiplication without any tools the difficulty of the multiplication task depends on the number of ciphers that need to be multiplying. For example task 3*4 can be performed by nearly all children, whereas task 23456*235687563 is very difficult for nearly all people. However, the machine can solve the problem of multiplication of very big numbers without noticeable increase in the difficulty when solving a problem of multiplication of small and big numbers.

Solving a problem that involves a mathematical formula requires not only the knowledge that makes it possible to transform the problem into the required solution but also understanding visual mathematical symbols. Even a simple mathematical expression requires understanding of the concept of members of the visual category. The category of mathematical symbols defines the rules of interpreting of the mathematical symbols such as operators, relations, logic operators, special symbols, ciphers. Even an isolated symbol can be interpreted as the mathematical symbol meaning of which is given by the knowledge schema. The most important of the mathematical symbols is the category of ciphers that is interpreted as the number that often refers to the number of objects or expresses the quantitative properties of the phenomena. The category of mathematical symbols is described in Chap. 4. The cipher is placed in the context of the hierarchical categorical structure given by the categorical chain: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{Mth}} \rangle \triangleright \langle \nu_{\mathrm{Cip}}, \nu_{\mathrm{Opr}}, \nu_{\mathrm{Rel}}, \nu_{\mathrm{Log}}, \nu_{\mathrm{Syn}} \rangle$ where ν_{Cip} denotes the cipher category.

The most of mathematical knowledge is represented in the form of the mathematical expressions. The mathematical expression is the combination of the numbers, letters, and mathematical symbols that have the mathematical meaning described by the rules of mathematical operators. For example, the expression ∑♠♪∂≡₣®§Δ≠≤%1/*∏∫ is not a mathematical expression because it consists of nonmathematical symbols, the expression +Δ∏2-/*34∂∑∫ is not a mathematical expression because the symbols do not follow rules of composition of the mathematical expression; the expression $x + 5 = 7$ is an example of the mathematical expression. As it was described in Chap. 4 the category of mathematical expressions consists of, among others, the equation category ν_{Eqt}, the function category ν_{Fun}, the algebraic operation category ν_{Alg}, the differentiation category ν_{Dif}, or the integration category ν_{Int}, and is represented by the following categorical chain: $\ldots\langle \sigma_{\mathrm{Pt}} \rangle \triangleright \ldots\langle \nu_{\mathrm{MtEx}} \rangle \triangleright \langle \nu_{\mathrm{Eqt}}, \nu_{\mathrm{Fun}}, \nu_{\mathrm{Alg}}, \nu_{\mathrm{Dif}}, \nu_{\mathrm{Int}} \rangle$. The category of mathematical expressions includes the rules that make it possible to check if a given expression is the member of the category of mathematical expressions. If the expression is recognized as the mathematical expression the further interpretation is based on knowledge supplied by the knowledge schema. For example, the expression that is interpreted as the mathematical equation can be interpreted as the model of the real-word phenomena or as a task that needs to be solved.

Many mathematical problems require transforming tasks given in the form of the linguistic description into the mathematical formula (expression) such as the equation, the system of algebraic equations, the differential equation, or the integral equation. The linguistic description consists of the visual objects called letters. System that is used for problem solving should be able to understand the letter and the linguistic expressions such as words or sentences. For example, the task given in the following form: "Z has bought three fish and one apple and paid 5$. In another day Z bought one fish and two apples and paid 5$. How much costs apple and fish" needs to be transformed into a mathematical expression. The first step to solve this problem is to translate the linguistic description into a more adequate (quantitative) mathematical representation. The system needs to understand such concepts as "bought," "apple," "fish", or "paid." In the context of categories of the visual objects the interpretation means to find the meaning of the concept that is expressed by linked categorical chains. For example, in this task the concept "apple" is understood as the "market" category not as a plant category. The meaning of the words "apple" and "fish" is modified by the contextual information of the words "cost," "bought," and "paid." The task can be translated into intermediate form by denoting market items, "apple" by the symbol A, and "fish" by the symbol B, and using the simple verb forms. After inserting symbols the description is given as: "Z buys 3 A and 1 B pay 5$. Next Z buys 1 A and 2 B pay 5$. How much cost A and B." From this form it is easy to find mathematical representation in the form of the system of linear equations $3x + y = 5$ and $x + 2y = 5$. The next step is to solve these equations which gives the solution $x = 1, y = 2$.

In the previous example the concept such as fish or apple refers to the visual real-world object. The interpretation of these concepts in terms of the plant or the animal categories can lead to misinterpretation of the task. In the second example, the concept such as minutes, walk, and velocity that refers to area of physics needs to be properly understood to solve a problem. The visual illustration of this problem is not very helpful, whereas it is possible to obtain solution by designing a simple graphical model and solving the problem by applying the method of computer simulation. The problem is stated as follows: "In 3 min G can walk 300 m. In 5 min M can walk 600 m. They started walking at the same time along the same track. How far apart were they after 4 min (a) 80 m, (b) 100 m, (c) 120 m, (d) 300 m." To find the solution there is a need to understand the concept of velocity. Marking velocity of G by X1 and velocity of M by X2 the solution can be given in the form of equation.

3*X1 = 300   5*X2 = 600   X1 = 100   X2 = 120

4*X1 = 400   4*X2 = 480   d = 480–400 = 80

As we can see, the solving of this task depends on the proper understanding of the concept such as velocity.

In another example: "There were 29 children on the school bus. After three boys got on the bus and five girls got off, there were twice as many boys left on the bus as the were girls. How many of the original 29 children were boys?" the task is given in terms of visual objects. This task was given as the school test where the answers from which correct one needs to be chosen were given as (a) 14, (b) 15, (c) 18, (d) 27. The linguistic description can be translated into the visual form and interpreted as the categories of the visual objects. However, when this description is assigned to the nonvisual mathematical tasks category all visual concepts will be interpreted in terms of the abstract terms of a set categories. It can be expressed in intermediate form: “It was 29 girls and boys (29 GB). After 3B in and 5G out there was twice B as were G. How many of the 29GB were B?,” where the symbol B denotes boys and the symbol G denotes girls. From this form the solution in the form of the system of linear equations is obtained: $G + B = 29$ $(B + 3)/(G–5) = 2$ and after solution $G = 14$, $B = 15$ the proper answer is selected (b) = 15. This problem can be solved using computer simulation in which visual object such as geometrical figures can represent the boys and girls. Let assume that circles represent girls and triangles represents boys. At first randomly the 29 figures, from which there is m circles and n triangles, are generated ($n + m = 29$). Next three triangles are added (the number of triangles after adding are denoted as n1) and five circles are removed (the number of circles after removing are denoted as m1). The condition if n1/m1 = 2 is checked by counting the number of triangles. If condition is fulfilled the number of triangles is the proper answer, if not the 29 new figures are generated and the process is started again. The solving this task by applying the computer simulation method is an example of solving problems by applying the visual thinking capabilities of the system. Each stage is to imagine the situation that is described and solution is obtained by comparison of the results of the imaginary transformation (generation figures, adding, and removing figures) with expectation that is expressed in the numerical form.

5.4.2. Problem Given in the Form of Both Symbolic and Visual Representations

In the previous chapter the problem that needs to be solved was given in the form of the linguistic description. There was no need for the visual representation. In this chapter, the examples of problems that utilize the visual representation to help solving problems are presented. The problem that needs to be solved is often given in the form of the linguistic description with addition of the graphical illustration (e.g., in the form of the engineering schema). When a problem involves spatial concepts such as points, lines, angles, directions, vectors, surfaces, or plane figures, diagrammatic representation may by useful aid to the symbolic representation, whether verbal, logical, or algebraic. Diagrammatic form refers to the expression of information by a collection of points, lines, angles, figures, vectors, and matrices, plots of functions, or graphs. Often the same information should be represented using a variety of symbolic or diagrammatic notations. Examples of tasks that are represented in both symbolic and diagrammatic notations are shown in Fig. 5.4. These tasks are formulated as follows:

1. A series circuit contains a resistor and a capacitor as shown in Fig. 5.4a. Determine the differential equation for the charge $q(t)$ on the capacitor if the resistance is R, the capacitance is C, and the impressed voltage is $E(t)$. This task can be expressed in intermediate form as follows: S circuit [resistor, capacitor]. Given [resistance R, capacitance C, impressed voltage $E(t)$], find [differential equation for charge $q(t)$ on capacitor]. To solve this task the concepts from the domain of the electrical engineering need to be understood.

2. A uniform beam of length L carries a concentrated load P_0 at x = L/2. The beam is clamped at both sides (see Fig. 5.4b).

3. A mass m having weight W is suspended from the end of a rod of constant length l. For motion in a vertical plane, we would like to determine the displacement angle θ, measured from the vertical, as a function of time t (see Fig. 5.4c).

4. A weight of 96 lb slides down an incline making a 30° with the horizontal. If the coefficient of sliding friction is μ, determine the differential equation for the velocity v(t) of the weight at any time. Use the fact that the force of friction opposing the motion is μN, where N is the normal component of the weight (see Fig. 5.4d).

5. The conical tank shown in Fig. 5.4e loses water out an orifice at its bottom. If the cross-section area of the orifice is $1/4 = ft^2$, find the differential equation representing the height of the water h at any time.

The problems shown in this section refer to the category of the physical models. The visual inference is expressed in the form of rules as follows [$\eta \in \varphi^\alpha$] $\Rightarrow a \triangleright \lambda^i$, where η is the symbolic name obtained in the reasoning process, φ^α is the visual concept, and $a \triangleright \lambda^i$ denotes the subtask λ^i to be performed. The visual concept φ^α includes one of the visual representations of the task. The subtask λ^i consists of selection of the models (the differential equations), verification of the parameters and variables of the model with those shown in an image, formulation of the protocol to be sent to the subsystem such as *Mathematica*, and interpretation of the solution obtained both in the symbolic and graphic form. The solution for the problem (2) is the differential equation

$$EI\frac{\mathrm{d}^4\theta}{\mathrm{d}x^4} = P_0\delta\left(x - \frac{L}{2}\right).$$

The solution for the problem (3) is the differential equation given as $ml(d^2\theta / dt^2) = -mg\sin\theta$.

5.4.3. Problem Given in the Form of the Visual Representation

Visual problem is formulated in terms of visual objects or phenomena. The problem can be formulated in the form of the linguistic description, for example, “select convex objects” or, based on the perceived objects, “name the perceived object.”

Solving visual problems is performed during the visual thinking process. For example, derivation of mathematical formulas based on the visual objects is an illustration of application of abstraction as a part of the thinking process. Solving visual intelligence tests is an example of application of the visual thinking in solving visual problems. Visual analogical reasoning is an example of the application of the selected problem’s solving strategy in solving a visual task. Examples of solving of the visual problems by visual recognition and abstraction are described in the following sections of this chapter.

5.4.3.1. Performing Task Given by the User

In this section, the problem formulated by the user in the linguistic form, is described and analyzed in more detail. The problem can be formulated as a simple task that requires undertaking the appropriate action or as the complex task that requires utilization of the complex visual reasoning processes. The task given by a user can be given in the linguistic form such as a written text or the spoken words. To perform the task SUS needs to understand the spoken words of a given language. Understanding the spoken words of a given language requires transformation of spoken words into text (words, sentences) and next interpretation of the text. Complexity of the formulated tasks can vary. In the case of a simple task, its description can be given in the form of one sentence describing the specific action that needs to be undertaken or a simple problem that needs to be solved. In the case of a complex task, the task description needs to be given in the form of a few paragraphs.

The simple task given in the form of the linguistic description describes an action that needs to be undertaken. Understanding such a task requires understanding the concept that is included in a task description. The task description can include concepts that refer to one of the ontological categories such as a figure, a letter, a sign, or a real-world object. The linguistic description of the task that refers to figures can be often found in the different intelligence tests such as IQ tests. SUS has ability to understand the task given by the description of the visual object in terms of its visual features (attributes). This description refers to concepts such as concavities or holes. For example, the description "select a concave object with *n* polygonal concavities and with *n* curvilinear holes" requires understanding of the following concepts: a convex object, a concave object, a cyclic object, the polygonal concavities and a curvilinear hole. The description can refer to the general figure category such as a "convex object" or a specific figure category such as a "right triangle." The task description refers to the non-visual category called the action category such as: select one, select *n*, select all, find an object given by a name, and find an object similar to that one. The following are examples of tasks formulated in the form of sentences, where the one of the action categories is applied: "Select a convex object," "Select all convex objects," "Find if a convex cyclic object is a member of a set," "Find if  is a member of a given set," "Compare if  and  are the same," "Find the name for this object ," "Find if this object  is concave," "Find the meaning of this object ."

As it was described, to find the solution to the task that requires performing an action, there is a need to understand the concepts that are included in the task description. The concept can refer to one of the ontological categories such as a figure or a real-world object. In the case of members of the figure category, description can be given in terms of the general categories such as a "convex object" or the specific categories such as a "right triangle." Understanding these concepts makes it possible to "see" differences among objects and undertakes actions that require discrimination among very similar objects.

To explain the problem of understanding by SUS at the different categorical levels of descriptions, an example of the tasks formulated as "select an object" that includes different levels of the object description, are given. This type of tasks is formulated in reference to the figure category and is given in terms of features of the visual object on the four different categorical levels of description:

1. Select a cyclic object, select an object with *n*-holes
2. Select a convex object with *n*-holes, select a concave object (*n*-concavities) with *n*-holes, select a concave object with a convex hole
3. Select a rectangle with *n*-holes, select a triangle with a triangular hole
4. Select a right triangle with an acute triangular hole

As a result of solving of this type of tasks, a division of the set of objects into two subsets is obtained. Examples of objects used in the experiment are shown in Fig. 5.5. At the first stage of performing the task given by the user in the form "select an object with *n*-holes", SUS abilities to understand the concept of a hole, was examined. The task was given to SUS as a series of queries where a number of holes were increased from 1 to 3. The results of a query "select an object with 1-hole" are shown in Fig. 5.6, the results of a query "select an object with 2-holes" are shown in Fig. 5.7, and the results of a query "select an object with 3-holes" are shown in Fig. 5.8.

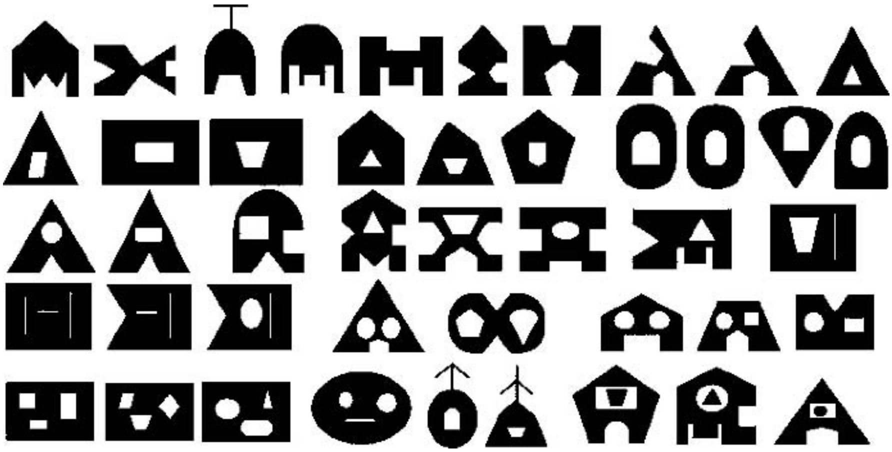


Fig. 5.5. Set of all objects used in the experiment

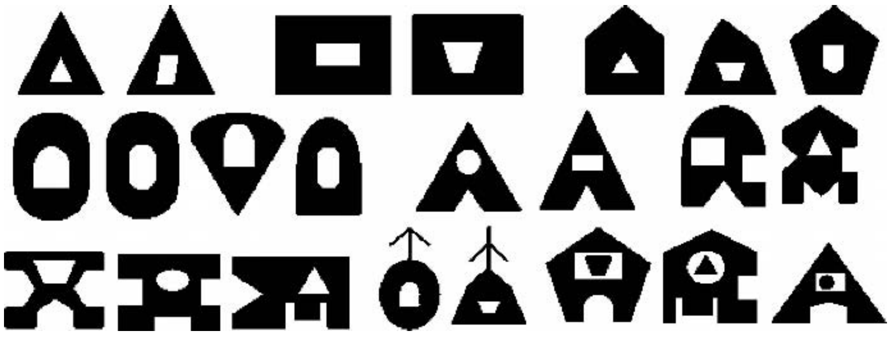


Fig. 5.6. Objects selected by a query "select an object with 1-hole"

Fig. 5.7. Objects selected by a query "select an object with 2-holes"

The results of solving tasks given at the second categorical level are given in Figs. 5.9–5.12. The results of query “select a convex object with 1-hole” is given in Fig. 5.9, the results of query “select a convex object with 2-holes” are given in Fig. 5.10, the results of query “select a concave object with one concavity and with 1-hole” are given in Fig. 5.11, and the results of query “select a concave object with two concavities and one cyclic hole” are shown in Fig. 5.12. The results of solving task given at the third categorical level are shown in Figs. 5.13 and 5.14. The tasks are formulated as “select a rectangles with 3-holes” and “select a triangle with a triangular hole.”

Fig. 5.8. Objects selected by a query “select an object with 3-holes”

Fig. 5.9. Objects selected by a query “select a convex object with 1-hole”

Fig. 5.10. Objects selected by a query “select a convex object with 2-holes”

Fig. 5.11. Objects selected by a query “select a concave object (1-concavity) with 1-hole”

Fig. 5.12. Objects selected by a query “select a concave object with two concavities and one cyclic hole”

Fig. 5.13. Objects selected by a query "select rectangles with 3-holes"

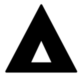


Fig. 5.14. Objects selected by a query "select a triangle with a triangular hole"

Very often the task description instead of giving the description of the visual object in the linguistic form indicates a visual object by pointing to one of the visual objects. An example of such tasks descriptions can be often found in the description of the tasks of the visual intelligence tests, e.g., "Find the following object:  among objects shown in Fig. 5.15 or "Find a given object among these five objects."

The solution of the task that refers to the category of the geometrical figure does not need to relay on the naming process. As it was described in the Chap. 4, most of members of the figure category do not have specific names. Naming process refers to the categorical structure of categorical chains. Solving this task requires finding the name at an appropriate categorical level. For example, the task "Select the different one" from the set of objects shown in Fig. 5.16 requires naming by selection of an appropriate level of generality of categories. This task can be accomplished by selection of the object at the figure level without reference to the real-world object. However, all figures are different so there is a need to find the name of the real-world object which these visual objects (phantoms) represent. Let us assume that the result of the naming is as follows: the apple, the elephant, the ox, the rabbit, and the fish. All categories are different; the naming is given at too specific level. To find the solution there is a need to perform the conceptual grouping. The conceptual grouping is based on the categorical chains that move toward the categories of the higher level. The categorical chains are as follows:

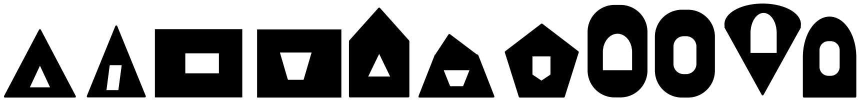


Fig. 5.15. The task in which the visual object is "described" by pointing to one of the objects

Fig. 5.16. Task given in the linguistic form: "Select the different one"

$\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{ReO}} \rangle \triangleright \langle \nu_{\mathrm{EAR}} \rangle \triangleright \langle \nu_{\mathrm{Liv}} \rangle \triangleright \langle \nu_{\mathrm{Ani}} \rangle \triangleright \langle \nu_{\mathrm{Cho}} \rangle \triangleright \langle \nu_{\mathrm{Fis}}, \nu_{\mathrm{Amf}} \nu, \nu_{\mathrm{Rep}}, \nu_{\mathrm{Ave}}, \nu_{\mathrm{Mam}} \rangle$, and specific categories derived from the mammal category are given by the following chain $\ldots \triangleright \langle \nu_{\mathrm{Ani}} \rangle \triangleright \langle \nu_{\mathrm{Cho}} \rangle \triangleright \langle \nu_{\mathrm{Mam}} \rangle \triangleright \langle \nu_{\mathrm{Ele}}, \nu_{\mathrm{Ox}}, \nu_{\mathrm{Rab}}, \nu_{\mathrm{Tig}} \rangle$. The category of plant is given as follows $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{ReO}} \rangle \triangleright \langle \nu_{\mathrm{Ear}} \rangle \triangleright \langle \nu_{\mathrm{Liv}} \rangle \triangleright \langle \nu_{\mathrm{Pla}} \rangle \triangleright \langle \nu_{\mathrm{Tre}}, \nu_{\mathrm{Shr}}, \nu_{\mathrm{UnS}}, \nu_{\mathrm{Vin}} \rangle$, where apple is described as a part category derived from the tree category shown by the following categorical chain $\ldots \langle \nu_{\mathrm{Pla}} \rangle \triangleright \langle \nu_{\mathrm{Tre}} \rangle \succ [\tau_{\mathrm{Frt}}] \triangleright \langle \nu_{\mathrm{Plu}}, \nu_{\mathrm{App}}, \nu_{\mathrm{Pea}} \rangle$. From the categorical chain $\ldots \triangleright \langle \nu_{\mathrm{Ani}} \rangle \triangleright \langle \nu_{\mathrm{Cho}} \rangle \triangleright \langle \nu_{\mathrm{Mam}} \rangle \triangleright \langle \nu_{\mathrm{Ele}}, \nu_{\mathrm{Ox}}, \nu_{\mathrm{Rab}}, \nu_{\mathrm{Tig}} \rangle$ we can infer that elephant, ox, and rabbit are mammals $\langle \nu_{\mathrm{Mam}} \rangle \triangleright \langle \nu_{\mathrm{Ele}}, \nu_{\mathrm{Ox}}, \nu_{\mathrm{Rab}}, \nu_{\mathrm{Tig}} \rangle$. These objects are now named: apple, mammal, mammal, mammal, and fish. There are three categories of objects, namely, apple, mammals, and fish. There is a need for another conceptual grouping. Based on the categorical chain $\ldots \triangleright \langle \nu_{\mathrm{Liv}} \rangle \triangleright \langle \nu_{\mathrm{Ani}} \rangle \triangleright \langle \nu_{\mathrm{Cho}} \rangle \triangleright \langle \nu_{\mathrm{Fis}}, \nu_{\mathrm{Amf}}, \nu_{\mathrm{Rep}}, \nu_{\mathrm{Ave}}, \nu_{\mathrm{Mam}} \rangle$ we can infer that the fish and mammals are members of category of animals $\langle \nu_{\mathrm{Ani}} \rangle \triangleright . \triangleright \langle \nu_{\mathrm{Fis}}, \nu_{\mathrm{Mam}} \rangle$. The result of perceptual grouping is the new names for all objects: apple and animals. Based on the categorical chains $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{Reo}} \rangle \triangleright \langle \nu_{\mathrm{EAR}} \rangle \triangleright \langle \nu_{\mathrm{Liv}} \rangle \triangleright \langle \nu_{\mathrm{Pla}} \rangle \triangleright \langle \nu_{\mathrm{Tre}} \rangle \succ [\tau_{\mathrm{Frt}}] \triangleright \langle \nu_{\mathrm{Plu}}, \nu_{\mathrm{App}}, \nu_{\mathrm{Pea}} \rangle$ and $\langle \Pi \rangle \triangleright \langle \nu_{\mathrm{El}} \rangle \langle \nu_{\mathrm{Reo}} \rangle \langle \nu_{\mathrm{EAR}} \rangle \triangleright \langle \nu_{\mathrm{Liv}} \rangle \triangleright \langle \nu_{\mathrm{Ani}} \rangle$ we can find the name of the object called "apple" at the level corresponding to the level of the objects named animals. From comparison of these two categorical chains (moving from the lower to the higher level) the plant category is selected. The category of apples and the category of animals have the common category of the living objects and the category of plants is at the same level as the category of animals. The answer is given in a very understandable way: there are four same objects – animals and one different – plant. The explanation can be given in general terms of the differences between plants and animals. In a similar way the task shown in Fig. 5.17 can be solved. The solution is obtained based on the categorical chain $\ldots \rhd \langle \nu_{\mathrm{Ani}} \rangle \rhd \langle \nu_{\mathrm{Cho}} \rangle \rhd \langle \nu_{\mathrm{Mam}} \rangle \rhd \langle \nu_{\mathrm{Ele}}, \nu_{\mathrm{Ox}}, \nu_{\mathrm{Rab}}, \nu_{\mathrm{Tig}} \rangle$. From this chain we infer that elephant, ox, and rabbit are mammals $\langle \nu_{\mathrm{Mam}} \rangle \rhd \langle \nu_{\mathrm{Ele}}, \nu_{\mathrm{Ox}}, \nu_{\mathrm{Rab}}, \nu_{\mathrm{Tig}} \rangle$ and from the animal category the two different categories, fish and mammals are selected $.. \rhd \langle \nu_{\mathrm{Fis}}, .., \nu_{\mathrm{Mam}} \rangle$.

To solve the task given in Fig. 5.18 there is a need to refer to the knowledge schema. As it was described in Chap. 4 each category has its knowledge schema that defines the main property of the object of this category. The physical properties of the animal supply the knowledge about the weight of the animal. The weight is given as three values (min, mean, max). When objects in Fig. 5.18 are named, for each object the properties such as weight are obtained from the knowledge schema of the category indicated by the animal's name. The knowledge schema of the animal is inherited by the lower categories of the categorical hierarchy. For each

Fig. 5.17. Task given in the linguistic form: "Select the different one"

Fig. 5.18. Task given in linguistic form "Place these objects from lightest to heaviest"

animal such as an elephant or a rabbit, knowledge from the knowledge schema gives the value of its weight. The task "place these objects from lightest to heaviest" is solved by finding the mean weight for each animal and next by sorting the name of animals according to their weight. The additional information concerning other specific features can be found from the knowledge schema of the knowledge object given as follows: $\kappa_{\mathrm{KB}} \triangleright \langle \kappa_{\mathrm{KOb}} \rangle \triangleright \langle \kappa_{\mathrm{BioOb}} \rangle \triangleright \langle \kappa_{\mathrm{ZooO}} \rangle \triangleright \langle \kappa_{\mathrm{Ani}} \rangle \triangleright \langle \kappa_{\mathrm{Cho}} \rangle \triangleright \langle \kappa_{\mathrm{Mam}} \rangle \triangleright \langle \kappa_{\mathrm{Ele}}, \kappa_{\mathrm{Ox}}, \kappa_{\mathrm{Rab}}, \kappa_{\mathrm{Tig}}, \rangle \prec \{\partial_{\mathrm{Nam}}, \partial_{\mathrm{Weg,}} \partial_{\mathrm{Col}}, \partial_{\mathrm{Sig}}\}$.

5.5. Visual Thinking as a Problem Solving

The first problem that system needs to solve is recognition of the visual object. The recognition is connected with naming of the object. An examined object can have the name assigned to it if the examined object is classified to one of the object categories. Recognition depends on the type of category to which the object belongs. The object recognition is view dependent and it depends on the different factors such as occlusions [4]. In these section, examples of the visual thinking engaged in solving different visual problems is presented. In the first sections the description of the problems formulated as the result of perceived object or phenomenon is given.

5.5.1. Perception: Problem Solving

In previous sections problems were formulated in the form of the linguistic descriptions and given to the system by another user. In this section the description of the problem formulated as the result of perceived object or phenomenon is given. The problem can be formulated on the perceptual level without involving ontological (meaningful) categories such as a sign or a real-world object. For example, derivation of the mathematical formula, solving of the problem that involves completing of the figure, the problem of a modal completion, or a simple visual analogy problem are all problems that can be regarded as problems given at the perceptual level.

Derivation of the mathematical formulas (operations) is an example of the problem solving process that makes it possible to invent mathematical operations such as summation or multiplication based on the application of abstraction and generalization during the visual thinking process. Mathematical thinking refers to the mathematical objects such as figures, to properties of these figures or to a set of figures. Mathematical operations such as summation or multiplication can be derived by abstracting properties of a set of visual objects. The operation such as summation can be obtained by counting elements that are obtained as the result of operations on the visual objects. Visual objects that are used in derivation process can be members of one of the ontological categories such as the category of the real-world objects or the category of the geometrical figures. In Fig. 5.20, examples that illustrate the summation process in the case of real-world objects and geometrical figures are shown. In Fig. 5.21, examples that illustrate the multiplication process are given.

At the perceptual level, objects given in Fig. 5.19 are regarded as elements of a set of objects. These objects are interpreted as elements of two sets of rectangular objects, where each object is represented by its symbolic name L^4. At first objects are clustered into two groups based on the proximity relation. The symbolic names of objects $L^4L^4L^4\ L^4\ L^4\ L^4$, $L^4L^4L^4L^4\ L^4\ L^4\ L^4\ L^4$ are transformed into elements of two sets by substituting the symbolic name with the selected symbol, "*a*" in our example, during abstraction process. The result of this substitution are two sets $\{a,a,a\}$ and $\{a,a,a,a\}$. Next, after summation of two sets $\{a,a,a\}\cup\{a,a,a,a\}$, the result $\{a,a,a,a,a,a,a\}$ is obtained. At the end cardinality of the set $\{a,a,a,a,a,a,a\}$, which is the result of the summation operation, is computed. In the case of visual objects shown in Fig. 5.19 the symbolic names $L^3L^3\ \ L^3\ L^3$ and $K^1K^1K^1K^1$ are obtained. The next the similar transformations are applied as in the previous example. In the case of the objects shown in Fig. 5.20, the objects are transformed into symbolic names. At first, the object is interpreted as an apple or fish and next after abstraction the two groups of objects $o\ o\quad o\ o\ o\ o$ are obtained.

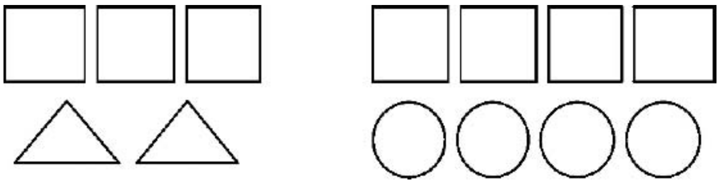


Fig. 5.19. Summation of geometrical figures

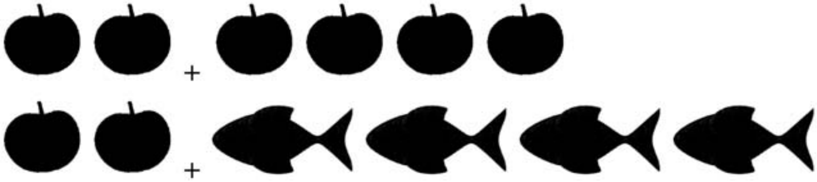


Fig. 5.20. Summation of real-world objects

Similarly, multiplication was derived by application of the perceptual laws (shown in Fig. 5.21). At first, the objects were clustered into three groups based on the proximity relation. Next, objects were transformed during abstraction process into the three groups of objects $o\ o\ o \quad o\ o\ o$ $o\ o\ o$ and next into three sets $\{o,o,o\} \quad \{o,o,o\} \quad \{o,o,o\}$. By denoting $\{o,o,o\}$ as O a set $\{O,O,O\}$ was obtained. Computing cardinality of sets $|\{O,O,O\}|$ and $|\{o,o,o\}|$ operator of multiplication 3*3 was obtained.

Mathematical operation such as matrix multiplication can be derived by using the special representation of matrices in the form of the colored tables. Figure 5.22 shows matrices represented as colored tables. The different

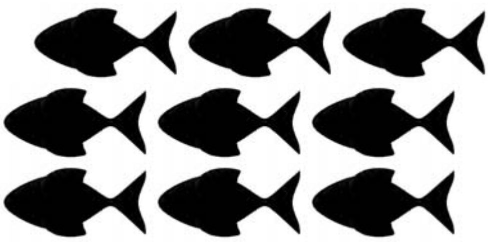


Fig. 5.21. Multiplication as a summation

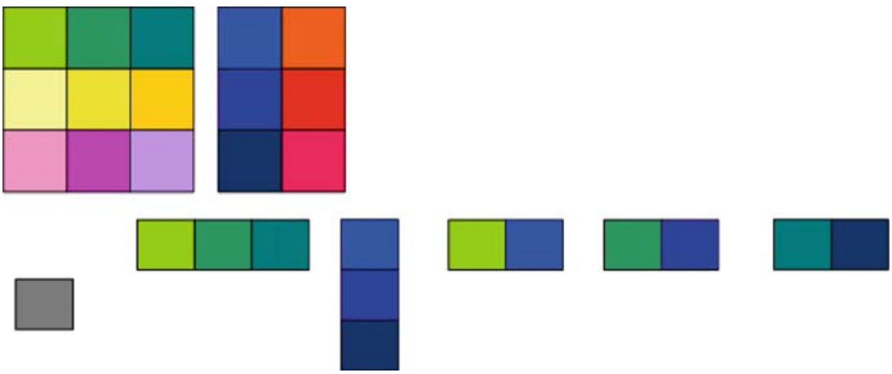


Fig. 5.22. Derivation of rules of matrix multiplication

types of mathematical operations can be derived by utilization of the different configurations of the elements of the colored table. For example, the matrix multiplication can be derived by making a new table according to the assumed strategy. For colored tables that are represented as members of the the colored convex thin class, the following symbolic notation can be applied: $\rho\{L_R^4\}\,\{(c_1)L_R^4,(c_2)L_R^4,(c_3)L_R^4,(c_4)L_R^4,(c_5)L_R^4,(c_6)L_R^4,(c_7)L_R^4,(c_8)L_R^4,(c_9)L_R^4\}$ or, alternatively, as the numbered convex thin class given as: $\rho\{L_R^4\}\{\binom{1}{1}L_R^4,\binom{1}{2}L_R^4,\binom{1}{3}L_R^4,\binom{2}{1}L_R^4,\binom{2}{2}L_R^4,\binom{2}{3}L_R^4,\binom{3}{1}L_R^4,\binom{3}{2}L_R^4,\binom{3}{3}L_R^4\}$.

The matrix B can be described as: $\rho\{L_R^4\}\,\{(d_1)L_R^4,(d_2)L_R^4,(d_3)L_R^4\,,(d_4)L_R^4,$ $(d_5)L_R^4,\ (d_6)L_R^4\}$ or $\rho\{L_R^4\}\{\binom{1}{1}L_R^4,\ \binom{1}{2}L_R^4,\ \binom{1}{3}L_R^4,\ \binom{2}{1}L_R^4,\ \binom{2}{2}L_R^4,\ \binom{2}{3}L_R^4\}$, where symbols $\binom{1}{1}L_R^4, \binom{1}{2}L_R^4,...$ denotes the symbolic name of the box in the table. The symbolic names that represent the box in the table are used to derive the procedure of multiplication of the matrices in the forms of mathematical symbols. The multiplication of the matrix can be described (in SUS notation) as follows:

$$(e_1)L_R^4=(c_1)L_R^4*(d_1)L_R^4+(c_2)L_R^4*(d_3)L_R^4+(c_3)L_R^4*(d_5)L_R^4\,.$$

Derivation of mathematical formulas from the visual objects is an illustration of application of abstraction during thinking process. Another example is completing of the figure that is based on application of perceptual laws. One of the perceptual laws (the heuristic) is that the archetype of the curve-linear class is seen as the archetype of the distorted curve class. For example, the archetype of the symmetrical curve-linear class $M^1[K^1]$ is an incomplete part of the archetype of the curve class K^1. Figure 5.23 shows an example of the task that is to complete distorted object. The archetype of the curve-linear object $M^1[K^1]$ can be completed by adding the curvilinear segment to obtain the curved object K^1.

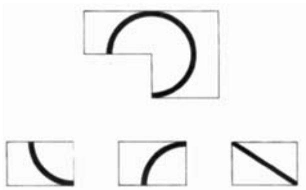


Fig. 5.23. Example of the task completing of the figure

The problem of completing pattern such as shown in Fig. 5.23 is more complex and requires using the line drawing representation, described in Chap. 4. The object that is cut (incomplete), shown in Fig. 5.23 has its symbolic name given as $\rho[Q^1[L^4](L_R^3)]\{Q^1[M_C^1](L^3), Q^1[L_R^4](M_C^1), Q^1[L_T^4](M^1)\}$. Parts (see Fig. 5.23) from which one needs to be selected to form the complete object has the following symbolic names (a) $\rho[L^4]\{M^2[L^3], Q^1[L_T^4](M^1)\}\{l\}$, (b) $\rho[L^4]\{M^2[L^3], Q^1[L_T^4](M^1)\}\{l^{\pi/2}\}$, (c) $\rho[L^4]\{2\cdot L_R^3]\}$. The symbolic name of the whole (not cut) object obtained by applying the perceptual law ("the archetype of the symmetrical curve-linear class $M^1[K^1]$ is an incomplete part of the archetype of the curve class K^1") is given as $\rho[L^4]\{K_C^1, 2Q^1[L^4](M_C^1)\}$. The solution is found by comparison of the whole made up from parts (symbolic names) a, b, c with the whole (not cut) object obtained as the result of application of the perceptual law. As the result of comparison the part (a) is selected. Only part one can give the whole that is given by the symbolic name $\rho[L^4]\{K_C^1, 2Q^1[L^4](M_C^1)\}$.

The problem of a modal completion has often been investigated by using partially occluded shapes that are regular or quasi-regular [5]. In the world that surrounds us most objects are partly hidden from our view by other objects. The available information from the visible part of the partly occluded object can be used to see this object as a whole object. Existing theories such as local theories, global theories, and integrated theories are used in explanation how the brain accomplishes the visual completion [5].

In this book the problem of completion is solved by application of the perceptual laws that can be formulated in terms of the shape categories. Finding general perceptual laws is the complex problem that is investigated within framework of the research on the visual thinking. In this book only a small sample of these problems is presented.

Finding solution to the problem shown in Fig. 5.24 requires application of laws of spatial decomposition of the object. The object shown in Fig 5.24a given by a symbolic name $\rho[Q^2[L^6](L^3, L_R^3)\{\rho[L_R^4]\{L_R^3, Q[L_R^4](L_R^3)\}, L^4\}$ can be decomposed based on the rules of decomposition given in the following form:

$$\begin{aligned}&\rho[Q^2[L^6](L^3, L_R^3)]\{\rho[L_R^4]\{L_R^3, Q[L_R^4](L_R^3)\}, L^4\} \to \\ &[Q[L_R^4](L_R^3), L^4] \to [Q[L_R^4](L_R^3), L_R^3, L^4]\end{aligned} \qquad \text{(Fig. 5.24b)}$$

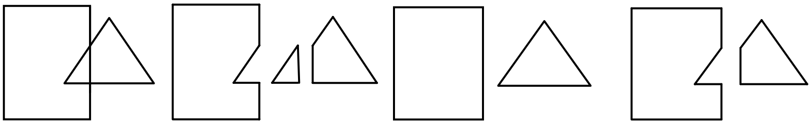


Fig. 5.24. Figure (**a**) and possible decomposition into parts (**b**), (**c**), and (**d**)

$$\rho[Q^2[L^6](L^3,L^3_R)]\{\rho[L^4_R]\{L^3_R,Q[L^4_R](L^3_R)\},L^4\}\rightarrow$$
$$[Q[L^4_R](L^3_R),L^4]\rightarrow[Q[L^4_R](L^3_R),L^3_R,L^4]\rightarrow[L^4_R,L^3_R,L^4]\rightarrow[L^4_R,L^3] \quad \text{(Fig. 5.24c)}$$

$$\rho[Q^2[L^6](L^3,L^3_R)]\{\rho[L^4_R]\{L^3_R,Q[L^4_R](L^3_R)\},L^4\}\rightarrow$$
$$[Q[L^4_R](L^3_R),L^4] \quad \text{(Fig. 5.24d)}$$

These rules are obtained by transforming the symbolic name $\rho[Q^2[L^6](L^3,L^3_R)\{\rho[L^4_R]\{L^3_R,Q[L^4_R](L^3_R)\},L^4\}$ using decomposition scheme. The most probable interpretation is that the figure consists of three parts: the concave rectangular object with one triangular concavities $Q[L^4_R](L^3_R)$, one triangular part L^3 that is fitted into the concavity so it forms the rectangular object L^4_R and the quadrilateral L^4 that is “glued” into the rectangular object forming the concave object $Q^2[L^6](L^3,L^3_R)$, or in the notation of the complex object as $C[L^4_R,L^4]$. The second interpretation is the triangular transparent object L^3 placed onto the rectangular object L^4_R. The last interpretation is that the concave rectangular object with one triangular concavity $Q[L^4_R](L^3_R)$ is glued with the quadrilateral object L^4 to form cyclic object $A[Q^2[L^6](L^3,L^3_R)](L^3)$, (concave object with one triangular hole).

Similarly an object shown in Fig 5.25a and given by a symbolic name $\rho[Q^2[L^6](L^3,L^3_R)\{Q[L^4_R](L^3_R),L^3\}$ can be decomposed based on the following rules of decomposition:

$$\rho[Q^2[L^6](L^3,L^3_R)\{Q[L^4_R](L^3_R),L^3\}\rightarrow[Q[L^4_R](L^3_R),L^3] \quad \text{(Fig. 5.25b)}$$

$$\rho[Q^2[L^6](L^3,L^3_R)\{Q[L^4_R](L^3_R),L^3\}\rightarrow$$
$$[Q[L^4_R](L^3_R),L^3]\rightarrow[L^4_R,L^3] \quad \text{(Fig. 5.25c)}$$

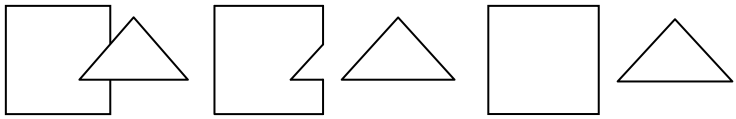


Fig. 5.25. Figure (**a**) and possible decomposition into parts (**b**) and (**c**)

The object in Fig. 5.25a is a member of the $\rho[Q^2[L^6](L^3, L_R^3)\{Q[L_R^4](L_R^3), L^3\}$ class. The archetype of this class can be interpreted as joining the concave rectangular object with one of the triangular concavities $Q[L_R^4](L_R^3)$ and triangular part L^3 that is fitted into the concavities $Q[L_R^4](L_R^3) \subset L^3$. The second interpretation is the triangular object L^3 placed onto the rectangular object L_R^4 to form $\rho[Q^2[L^6](L^3, L_R^3)]\{Q[L_R^4](L_R^3), L^3\}$ object.

To find the perceptual laws concerning occlusion problem at first the simple cases are investigated. Let object (shown in Fig 5.26) given by the name $\rho\ [W]\{Q, A\}$ is composed of two convex objects. The name W represents the whole concave object, the name Q represents the concave part that is result of occlusion of the convex part B by the convex part A, and A is part that is put on the top of the part B. The interpretation of the part B depends on the type of class to which the object A and Q belongs. If objects A and Q are polygons the whole object is given as $\rho[W]\{Q[Z^n], A\}$ and the object B is given as $B \equiv Z^{n-1}$, where n is the n-sided polygon of the generic class of the concave object Q. The perceptual law is discovered by analyzing the sample of cases that represent occlusion of the two convex objects.

The first example of the occlusion problem is to interpret perceived object as the occlusion of the two convex objects. Objects shown in Fig. 5.26a, b are given by the name $\rho[Q^2[L^6](2L_R^3)]\{Q[L^5](L_R^3), L_R^4\}$, where $W \equiv Q^2[L^6](2L_R^3)$, $Q \equiv Q[L_R^5](L_R^3)$. Two interpretations of the occluded parts are possible. The first interpretation is – two objects A and Q ($A \equiv L_R^4$ and $Q \equiv Q[L_R^5](L_R^3)$) are joined together; the second interpretation is – two objects A and B (rectangles $A \equiv L_R^4$ and $B \equiv L_R^4$) placed one onto another.

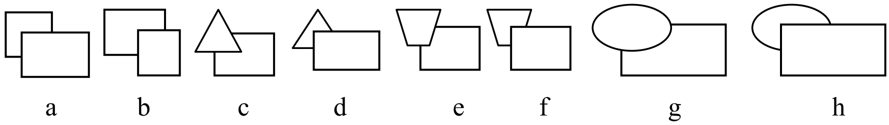


Fig. 5.26. Solving occlusions problem

The object shown in Fig. 5.26c is given by the name $\rho[Q^2[L^5](2L^3)]\{Q[L_R^5](L^3), L^3\}$, where $W \equiv Q^2[L^5](2L^3)$, $Q \equiv Q[L_R^5](L^3)$, $A \equiv L^3$, and $B \equiv L_R^{5-1} \equiv L_R^4$. Interpretation of the occluded parts is given as follows: the first interpretation – two objects A and Q ($A \equiv L^3$ and $Q \equiv Q[L_R^5](L_R^3)$) are joined together; the second interpretation is – two objects A and B ($A \equiv L^3$ and $B \equiv L_R^4$) are placed one onto another.

The object shown in Fig. 5.26d is given by the name $\rho[Q^2[L^5](2L^3)]\{Q[L^4](L^3), L_R^4\}$, where $W \equiv Q^2[L^5](2L^3)$, $Q \equiv Q[L_R^5](L^3)$, $A \equiv L^3$, and $B \equiv L_R^{5-1} \equiv L_R^4$. Interpretation of the occluded parts is given as follows: the first interpretation – two objects A and Q ($A \equiv L^3$ and $Q \equiv Q[L_R^5](L_R^3)$) are joined together; the second interpretation – two objects A and B ($A \equiv L^3$ and $B \equiv L_R^4$) are placed one onto another. The object given in Fig. 5.26c can be interpreted as the object given in Fig. 5.26d that is turned off. These two objects given in Fig. 5.26c, d, are interpreted as objects that are "turn over." During interpretation process, the object such as the object in Fig. 5.26c will be interpreted by taking into account two versions of the object – the object in Fig. 5.26c and its "turn over" version – the object in Fig. 5.26d. The "turn over" version can be obtained by following schema: perceived object $\rho[W_1]\{Q[Z_R^z](K^k), M^m\}$ and its "turn over" version $\rho[W_2]\{Q[N^n](G^g), H_R^h\}$, where $W_1 \equiv W_2$, $H_R^h \equiv Z_R^{z-1}$, $N^n \equiv M^{m+1}$, and $G^g \equiv K^k$.

The object shown in Fig. 5.26e, given by the name $\rho[Q^2[L^6](L_R^3, L^3)]\{Q[L^5](L^3), L_T^4\}$, is interpreted in the context of perceived "turn over" object given by the name $\rho[Q^2[L^6](L_R^3, L^3)]\{Q[L^5](L^3), L_R^4\}$ (Fig. 5.26f). For perceived objects given by the following symbolic names $W \equiv Q^2[L^6](L_R^3, L^3)$, $Q \equiv Q[L_R^5](L^3)$, $A \equiv L_T^4$, and $B \equiv L_R^{5-1} \equiv L_R^4$; the interpretation of occluded parts is given as follows: the first interpretation – two objects A and Q ($A \equiv L_T^4$ and $Q \equiv Q[L_R^5](L^3)$) are joined together; the second interpretation – two objects A and B ($A \equiv L_T^4$ and $B \equiv L_R^4$) placed one onto another. For the "turn over" version given by the following symbolic names $W \equiv Q^2[L^6](L_R^3, L^3)$, $Q[L^5](L^3)$, $A \equiv L_R^4$, and $B \equiv L^{5-1} \equiv L^4$, the interpretation of occluded parts is given as follows: the first interpretation – two objects A and Q ($A \equiv L_R^4$ and $Q \equiv Q[L^5](L^3)$) are joined together; the second interpretation – two objects A and B ($A \equiv L_R^4$ and $B \equiv L_T^4$) placed one onto another.

Based on these cases the perceptual law can be formulated as follows: the object given by the name $\rho[W_1]\{Q[Z_R^z](K^k), M^m\}$ has its "turn over" version $\rho[W_2]\{Q[N^n](G^g), H_R^h\}$, where $W_1 \equiv W_2$, $H_R^h \equiv Z_R^{z-1}$, $N^n \equiv M^{m+1}$, and $G^g \equiv K^k$. For the perceived object, the interpretation of the occluded parts is given as follows: the first interpretation – two objects A and Q ($A \equiv M^m$ and $Q \equiv Q[Z_R^z](K^k)$) are joined together; the second interpretation is – two objects A and B ($A \equiv M^m$ and $B \equiv Z_R^{z-1}$) placed one onto another. For the "turn over" version $\rho[W_2]\{Q[N^n](G^g), H_R^h\}$ interpretation of the occluded parts is given as follows: the first interpretation – two objects A and Q ($A \equiv H_R^h$ and $Q \equiv Q[N^n](G^g)$) are joined together; the second interpretation is – two objects A and B ($A \equiv H_R^h$ and $B \equiv N_T^{n-1}$) placed one onto another.

Similarly two convex objects such as polygons and curves can be interpreted. The object shown in Fig. 5.26g given by the name $\rho[Q^2[M_{L^5}^1](2Q[L^3](M^1))]\{Q[L^5](M^1), K_E^1\}$ is interpreted in the context of perceived "turn over" version given by the name $\rho[Q^2[M_{L^5}^1](2Q[L^3](M^1))]\{Q[M^1](L^3), L_R^4\}$ (Fig. 5.26h). For perceived objects given by the following symbolic names $W \equiv Q^2[M_{L^5}^1](2Q[L^3](M^1))$, $Q \equiv Q[L^5](M^1)$, $A \equiv K_E^1$, and $B \equiv L_R^{5-1} \equiv L_R^4$; the interpretation of the occluded parts is given as follows: the first interpretation – two objects A and Q ($A \equiv K_E^1$ and $Q \equiv Q[L^5](M^1)$) are joined together; the second interpretation is – two objects A and B ($A \equiv K^1_E$ and $B \equiv L^4_R$) placed one onto another. For the "turn over" version given by the following symbolic names $W \equiv Q^2[M^1_{L^5}](2Q[L^3](M^1))$, $Q \equiv Q[M^1[K^1]](L^3)$, $A \equiv L^4_R$, and $B \equiv K^1_c$; the interpretation of the occluded parts is given as follows: the first interpretation – two objects A and Q ($A \equiv L^4_R$ and $Q \equiv Q[M^1[K^1]](L^3)$), are joined together; the second interpretation is – two objects A and B ($A \equiv L^4_R$ and $B \equiv K^1_c$) placed one onto another.

Similarly the perceptual law can be derived for more than two objects such as given in Fig. 5.27.

In the case of the concave object the interpretation of the object is not always unique and there is a need to use the contextual knowledge to find an appropriate interpretation. Example of the occlusion of two concave objects is shown in Fig. 5.28. The derivation of the perceptual laws is left as the exercise for readers.

Another example of application of perceptual laws is to solve visual analogy problems. Visual analogy problem that is solved at the perceptual level is based on the similarity relations. The visual analogy problem will be discussed in more details in the following sections of this chapter. In this section an example that is taken from Arnhaim's book [6] is given. The problem shown in Fig. 5.29 can be formulated as follows: apply the similarity relation between objects in Fig. 5.29a, b to select one of the object d1, d2, d3, d4 that fulfill the similarity relation that is found between objects a and b. To solve this problem the visual reasoning is applied and

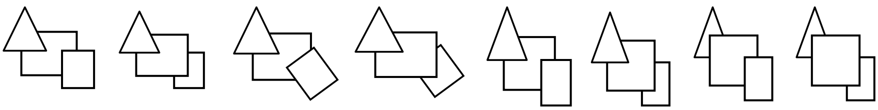


Fig. 5.27. Solving the occlusion problem

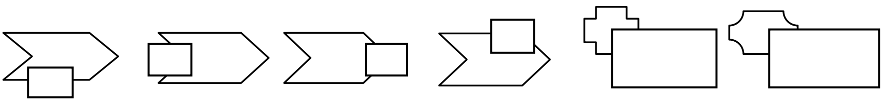


Fig. 5.28. Solving the occlusion problem

symbolic names of each object in Fig. 5.29 are found. The symbolic names are transformed into the string form, as follows:

a. A2|L4Q|1_A1|K1C|K1C|L4Q|
b. A2|L4Q|1_A1|K1C|L4Q|K1C|
c. A2|L4Q|1_A1|L3E|L4Q|K1C|
d1. A2|L4Q|1_A1|L3E|L4Q|K1C|
d2. A3|L4Q|L3E|K1C|K1C|
d3. A2|L4Q|1_A1|L3E|K1C|L4Q|.

The transformation of symbolic names into the string forms is explained in Chap. 2. After comparison of strings obtained from objects shown in Fig. 5.29a, b the part of strings that shows no differences is removed and as the result strings K1C|L4Q| and |L4Q|K1C| are obtained. The result indicates that strings vary in the last two symbols of the string. The string (c) is of the same type as strings (a) and (b). The string (d1) is the same as the string (c) so it is excluded from the solution. Because the string (d2) is the different type than strings (a) and (b) it is excluded from the possible solution. The solution is found by looking for the part of the string (c) |L4Q|K1C| that is the same type as the string (a) and (b) and has the same configuration of the last two symbols. As the result of comparison of strings |L4Q|K1C| and |K1C|L4Q|, the object (d3) is selected as the solution.

The class of problems that was shown in this chapter can be regarded as the class of perceptual problems. This type of problems can be found in literature concerning explanation of the perceptual phenomenon.

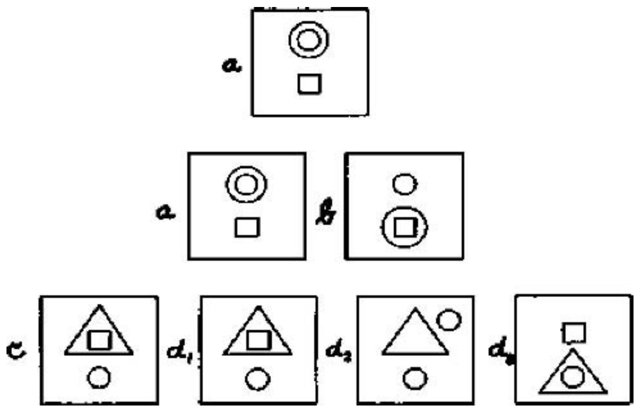


Fig. 5.29. Example of the visual analogy problem

5.5.2. Naming and Recognition of the Different Categories of Objects

In previous sections problems that belong to the class of perceptual problems were described. In this chapter problems connected with naming objects are presented. Naming process can be regarded as a special class of problem solving, where the task is formulated by SUS as the result of perceiving a given object. Seeing object yields the "inner" question "What is this?" or "What is the name of this object?." To find the answer to this question there is a need to find the name of the object.

Names are words or groups of words of one of existing languages that are used to denote object categories. There are names for general categories and names for the prototypes of objects. For example, the object "tuna" can be described by the name of the general category – animal, the specific category – fish or a prototype – tuna.

In general, naming is the process of associating and recalling the symbol for a concept when given its nonsymbolic or sensory object, and imaging is the process of associating and recalling concept when given its symbolic representative. Naming takes place when we describe a scene and imaging takes place when we imagine a scene which is being described verbally. Naming translates a sensory data into a symbolic form. Naming is often thought of as the result of the recognition process. Recognition is one of the main parts of many cognitive processes such as naming or reading. The recognition of the object does not assume that the object has its name but only that it was the object of the previous visual experience. For example, the flower that was seen in the botanic garden can be recognized but not named because we do not know the name of this flower. The flower which we know (e.g., rose) is recognized and named by the name "rose."

Within the human visual system naming is often regarded as the assignment of a class token to the image of an object: a square, a circle, a shoe, a tree, a radio. The visual system first encodes the retinal image of the object to be named into an internal representation called a pictorial pattern. The resulting pattern is delivered to the naming store for the final assignment of its symbolic name. The naming store finally associates the symbolic encoding with the pictorial encoding and hence can be used both for naming and imaging.

Naming is a process of attaching name to the perceived object. The first step in naming of the object is to assign it to one of the object categories and next attach the name of the category into the perceived object. Naming is performed during the visual inference process. At first, the phantom u

given as an input is transformed into a set of critical points $\coprod$ by the sensory transformation and a set of critical points $\coprod$ is transformed into symbolic name η. The symbolic name η obtained during visual reasoning process is used to find the name of the visual category by searching of space of visual concepts given as a set of symbolic names $\varphi^i = \{\eta_1, \eta_2, ..., \eta_n\}$. The name n is assigned to the perceived object during the visual inference and is expressed in the form of rules: $[\eta \in \varphi^u] \Rightarrow u \rhd n$, where η is the symbolic name obtained in the reasoning process, φ^u is the shape category (the visual concept), and $u \rhd n$ denotes the naming process. As it was described in the previous section the perceived object can be named by the name of one of general categories or the name of the lower category (prototype). For example, an examined object can be named on the lower categorical level (prototype level) as the capital letter "A" font times new roman or on the general level as a letter. The selection of an appropriate categorical level depends on contextual information supplied by the task description.

In this chapter naming refers to classification of the object into one of the object categories. The name can be given as a letter, e.g., "A," as a word, e.g., "a triangle," "an apple," or as the word description, e.g., "a concave object with a hole." The name can represent the lower level category (the prototype) or the higher level category. In contrast to area of objects recognition where examined object is classified to one of the well-defined classes of objects, recognition of the visual object described in this chapter involves classification of an object to one of the categories of visual objects. The object that is recognized is always named by the name of one of the general categories. For example, the unknown flower can be named as a "flower" or a "plant."

Naming refers to two different processes: learning the name of the unknown object and object naming that attaches the name to the perceived object. Learning of the name of the unknown object is connected with the knowledge acquisition process. The knowledge is represented as a set of connected categorical chains. Learning consists of two main processes: learning of the visual knowledge (the visual concept) and learning the nonvisual knowledge (knowledge schema). Each category is given by its name that is attached to the categorical chain during learning process. The process of learning of the new name (the name that is invented) is a complex problem because the name as the linguistic category is used in the communication process. The new name that is proposed has to be acceptable for all users of a given language.

Visual thinking is immanent part of the naming process that involves transformation of the perceptual data into symbolic name, assigning the visual object into one of the object categories, assigning the name to the perceived object, and starting further interpretational process. Usually the naming begins complex thinking process that leads to understanding visual object as the part of the interpretational task. For example, when perceived object is named as a given letter, this can start a complex thinking process that leads to reading and understanding text. During the naming process an important issue is the selection of an appropriate categorical level. The selection of an appropriate level is based on the contextual information. For example, reading text requires interpreting perceived object at the letter level. The interpretation of the perceived object at the lower level (font) is too specific and can cause unnecessary delay in reading process. Reading of the texts does not require knowing the font of the letter.

5.5.2.1. Figure Naming: Assigning the Name to the Figure

Figure naming refers to the figure category and is based on the previously learned knowledge about the figure category. Figure naming is connected with recognition of the object as a member of one of the figure categories. As it was described in Chap. 4, a figure is an object that is defined based on the geometrical properties and refers to a mathematical object such as the 2D closed curves or any abstract object. An abstract object is an object meaning of which does not refer to a sign, a letter, or a real-world object. At first stage of naming, a perceived object is classified to one of the figure categories. The basic figure categories includes: the convex polygon category, the concave polygon category, the convex curve category, the concave curve category, the convex curve polygon category, or the concave curve polygon category. The figures that play an important role in area of mathematic or physic are given by its name. For example, the circle, the ellipsis, the pearl curve are examples of named curves (curves that are given by its name). Naming of a perceived object that is a member of the named figure category is obtained by assigning the name to this object. Naming process involves finding an appropriate categorical level and using the name and knowledge of the category of this level to form the concept of the object. The concept that is formed during naming process can be used during communication session to supply information about the object that is perceived. The concept of the perceived object is the result of the interpretational process. Figure can be interpreted based on geometrical and perceptual properties of the visual object. This interpretation is related to the shape classes and the name of the figure that does not have its characteristic name can be given in the form of a class description, e.g., “the concave figure,” “the complex figure” that refers to the one of the general figure categories. At the lower level, for the figure that does not have the name, the symbolic name that was assigned to the learned figure category is assumed to be the name of the perceived object. In some cases figure does not need to be named but only recognized as the member of one of the figure categories. Recognition does not need to assign a name to the figure but only an index that makes it possible to use the other information about recognized category.

5.5.2.1.1. Naming of the Figure Without Name

As it was shown in Chap. 4 the 2D figure category consists of the polygon category ν_{Pol}, the curve category ν_{Cur}, and the curve polygon category ν_{CuPo}, and is given by the following categorical chain: $O \triangleright \langle \nu_{\mathrm{Fig}} \rangle \triangleright \langle \nu_{\mathrm{2DF}} \rangle \triangleright \langle \nu_{\mathrm{Pol}}, \nu_{\mathrm{CuPo}}, \nu_{\mathrm{Cur}} \rangle$. Only few polygons or curves have theirs name. For example, the name “triangle” denotes an object (polygon) that has its characteristic properties, has three sides. Similarly, the name “quadrilateral” or “rectangle” denotes an object that has characteristic properties. For the polygon that does not have name, the name can be given in the form of the definition. The definition of the polygon is given in the form of the attributes of the class or in the generative form. The first step in naming of the object as a polygon is to assign it to the polygon category and next convert a symbolic name into a class description during concept formation process. Communication of the results of the visual experience requires that the formal definition of the figure needs to be transformed into the description given in one of the existing languages, e.g., English. The figure such as a polygon is defined within the domain of geometry, and the definition as well as properties of the polygons that are described in geometrical literature can be used to represent knowledge about the figure concept. These properties are incorporated into the knowledge of the knowledge schema of a polygon category during learning stage. Similarly, the first step in naming of an object as a curve is to assign an object to a curve category and next to find the description of the curve in terms of its attributes. Naming is performed during the visual inference process as described in previous chapters.

5.5.2.1.2. Naming of the Figure with Name

The figures such as polygons or curves that have their names are members of the named figure categories. These categories are described in detail in Chap. 4. The members of the category of named polygons are polygons that have the special properties that are well defined and are part of the geometrical knowledge. An example of the categorical chain for specific named categories derived from the quadrilateral category is as follows: $..\langle \nu_{\text{Pol}} \rangle \rhd \langle \nu_{\text{NPol}} \rangle \rhd \langle \nu_{\text{ClCoP}} \rangle \rhd \langle \nu_{\text{Qua}} \rangle \rhd \langle \nu_{\text{Sqa}}, \nu_{\text{Rec}}, \nu_{\text{Rho}}, \nu_{\text{Tra}} \rangle$. Similarly the category of the closed named curve that is derived from the curve category is given by the following categorical chain: $..\langle \nu_{\text{Cur}} \rangle \rhd \langle \nu_{\text{NCur}} \rangle \rhd \langle \nu_{\text{ClCcC}} \rangle \rhd \langle \nu_{\text{OneS}} \rangle \rhd \langle \nu_{\text{DHer}}, \nu_{\text{TCrm}} \rangle$, where the category of the one shaped curves and the category of the many shaped curves, is distinguished. Naming of the object that belongs to the category of named polygons is very similar to naming a polygon described in the section "Naming of the Figure Without Name." In the case when the perceived object is assigned to the category of named polygons, the name of this category is assigned to the perceived object. In the case when the perceived object is assigned to the category of named curves, the further naming process depends on the lower category (the category of the one shaped curves and the category of the many shaped curves).

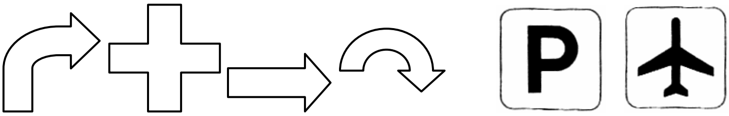
The name, the word or a group of words, is derived from the word or phrase category and is given by the following categorical chain: $\langle \Pi \rangle \rhd \langle \sigma_{\text{Pt}} \rangle \rhd \langle \nu_{\text{Let}} \rangle \rhd .. \rhd \langle \nu_{\text{Phr}} \rangle \rhd \langle \nu_{\text{Nam}} \rangle \rhd \langle \nu_{\text{c1}}, \nu_{\text{c2}} \rangle$. The category of name refers to the knowledge of the naming process that contains the rules of naming process. The knowledge chain given as follows: $\kappa_{\text{KB}} \rhd \langle \kappa_{\text{KOb}} \rangle \rhd \langle \kappa_{\text{Lin}} \rangle \rhd \langle \nu_{\text{Phr}} \rangle \rhd \langle \nu_{\text{Nam}} \rangle \rhd \langle \nu_{\text{c1}}, \nu_{\text{c2}} \rangle$, shows how knowledge of naming process is derived from the linguistic object. The lower categories $\langle \nu_{c1}, \nu_{c2} \rangle$ denote the names of the all known ontological categories. Each specific name category ν_{c1} refers to a set of names of this category. For example, the name of person consists of the finite set of person's name. Each named category (prototype) that is learned has its name connected with the visual concept. For example, the visual concept given by the symbolic name $\varphi_{\text{Tr}} = \{Q^3[M^1[L^4]](2Q^2[L^3](2M^1), Q^2[L^3](2\tilde{M}^1))\}$ is connected with the name "trefoil."

The name can be given as a word or a group of words that are expressed in one of the existing languages. The name that is assigned to the perceived object is expressed in English language. The knowledge of the naming process given by the knowledge schema of the name category makes it possible to translate name expressed in English language into the name expressed in any existing languages. At the end of naming process the language is selected and the name is given in the form of the expression of the selected language. For example, let us assume that as the result of the reasoning process the symbolic name $\eta \equiv Q^3[M^1[L^4]](2Q^2[L^3](2M^1), Q^2[L^3](2\tilde{M}^1))$ was obtained. During the visual reasoning, the visual concept $\varphi_{\text{Trefoil}} = \{Q^3[M^1[L^4]](2Q^2[L^3](2M^1), Q^2[L^3](2\tilde{M}^1))\}$ was found and the name was found during the visual inference by applying the rules $[\eta \in \varphi^u] \Rightarrow u \triangleright n$. As the result of naming process the name "trefoil" is assigned to the examined object. When the name is found the thinking process that leads to understanding is continued. The first thinking task is to exploit the knowledge of the selected category. The interpretation of the curve is based on the categorical chain. When the name of the category is found the general description of the curve can be obtained by moving into the higher level of the categorical chain. For example, when an examined object is named "trefoil" (category of trefoil curve ν_{TCrm}) we can infer that it is a closed 2D mathematical curve ν_{2D}, as shown in the following categorical chain: $..\langle \nu_{\text{Cur}} \rangle \triangleright \langle \nu_{\text{NCur}} \rangle \triangleright \langle \nu_{\text{ClCcC}} \rangle \triangleright \langle \nu_{\text{OneS}} \rangle \triangleright \langle \nu_{\text{DHer}}, \nu_{\text{TCrm}} \rangle$. The knowledge schema of the named curves consists of the visual concept ∂_{ViC}, the name ∂_{Nam}, the mathematical formula ∂_{MaF}, the definition ∂_{Def}, and the method of generation ∂_{MGe} and is given as the part of the knowledge chain as follows: $\kappa_{\text{KB}} \triangleright \langle \kappa_{\text{KOb}} \rangle \triangleright \langle \kappa_{\text{Mat}} \rangle \triangleright \langle \nu_{\text{Cur}} \rangle \triangleright \langle \nu_{\text{NCur}} \rangle \triangleright \langle \nu_{\text{ClCcC}} \rangle \triangleright \langle \nu_{\text{OneS}} \rangle \triangleright \langle \nu_{\text{TCrm}} \rangle \prec \{\partial_{\text{ViC}}, \partial_{\text{Nam}}, \partial_{\text{MaF}}, \partial_{\text{Def}}, \partial_{\text{MGe}}\}$. From the knowledge schema (shown in the categorical chain) we can obtain the definition that gives the description of the curve and supplies the link to other geometrical categories. The mathematical formula given in a form of the mathematical equation gives the link to the category of mathematical equations. The method of generation of the curve makes it possible to generate the different visual aspects of the curve, e.g., rotated version of the curve. The generated curve can be used in the thinking process during the visual explanation or as a part of the imagery process. The thinking process can utilize the analogical reasoning based on similarity relation. For example, the trefoil is the concave symmetrical curve that has three residuals. The curve is similar to non-symmetrical curve that has three residuals. These curves can have the similar mathematical properties. The thinking process can exploit meaning of the name "trefoil" that gives a link to the category of plant. This link supplies the knowledge of the botanical categories from which the different properties of plant can be derived.

The prototype of the category of the named curve has its knowledge schema that is inherited from the categories at higher level of the categorical chain. When the perceived object is named all knowledge that is accessible by the links among categories can be used in process of understanding of the object. The knowledge can be accessible through the categorical chain. For example, the object named the "trefoil curve" is understood as a mathematical closed curve and can be described by applying all general knowledge of the curve category ν_{Cur}. The hierarchy of the categorical chain supplies the knowledge schema at each categorical level as shown in in the following categorical chains: $..\langle \nu_{\mathrm{Cur}} \rangle \prec \{\partial_{\mathrm{ViC}}, \partial_{\mathrm{Nam}}, \partial_{\mathrm{Def}}, \partial_{\mathrm{MGe}}\}$, $.. \triangleright \langle \nu_{\mathrm{ClCcC}} \rangle \prec \{\partial_{\mathrm{ViC}}, \partial_{\mathrm{Nam}}, \partial_{\mathrm{Def}}, \partial_{\mathrm{MGe}}\}$, $.. \triangleright \langle \nu_{\mathrm{TCrm}} \rangle \prec \{\partial_{\mathrm{ViC}}, \partial_{\mathrm{Nam}}, \partial_{\mathrm{MaF}}, \partial_{\mathrm{Def}}, \partial_{\mathrm{MGe}}\}$. As we can see from these categorical chains the knowledge schema of the prototype of the trefoil curve ν_{TCrm} includes additionally the mathematical formula ∂_{MaF} as the knowledge of the prototype. In the case when an examined object is recognized as the curve but does not have any specific name the object can be named by the name of the more general category, e.g., a concave curve. The curve can be named by using the similarity relation, e.g., "the curve similar to the trefoil curve." This type of naming makes it possible to use all knowledge of the trefoil curve to describe an unnamed object.

5.5.2.2. Naming of the Sign

As it was described in Chap. 4 the sign category refers to the visual object meaning of which is based on the system of conventional rules and the meaning of elements of the category of symbolic signs does not depend on the meaning of other elements of the same category. Examples of perceived objects that can have many interpretations are shown in Fig. 5.30. For example, objects shown in Fig. 5.30a–d can be interpreted as mathematical



a b c d e f

Fig. 5.30. Examples of signs that have many interpretations

symbols, the object in Fig. 5.30e as a letter P, whereas the object in Fig. 5.30f as a real-world object.

Naming of the sign is based on the previously learned knowledge about the sign category. Signs that do not resemble any real-world objects are often called symbols. They are similar to the letters, e.g., the musical score. Naming of the sign is to classify the perceived object to one of the sign categories and naming of the visual symbols is to classify the perceived object into one of the visual symbol categories. The visual symbol category is divided into specific categories shown by the categorical chain: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{EL}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{Sym}} \rangle \triangleright \langle \nu_{\mathrm{Mth}}, \nu_{\mathrm{Mus}}, \nu_{\mathrm{EngS}} \rangle$. The category of each visual symbol is interpreted according to the knowledge supplied by the knowledge scheme. Each category has its own rules that govern the composition of the visual symbols into the complex expression.

5.5.2.2.1. Naming of Mathematical Symbols

The mathematical symbol is a representative of the category of visual symbols described in detail in Chap. 4. The first step in naming a perceived object as a mathematical symbol is to assign it to the category of mathematical symbols and next assign a name to the perceived object. As it was shown in Chap. 4 the category of mathematical symbols is derived from the category of the visual symbols and is divided into the category of mathematical operators, relations, logic operators, special symbols, or ciphers.

From the category of the visual symbols, the category of mathematical elements is derived. The category of mathematical symbols (elements) is divided into the category of mathematical operators, relations, logic operators, special symbols, and ciphers $\ldots \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{Mth}} \rangle \triangleright \langle \nu_{\mathrm{Cip}}, \nu_{\mathrm{Opr}}, \nu_{\mathrm{Rel}}, \nu_{\mathrm{Log}}, \nu_{\mathrm{Syn}} \rangle$. The category of mathematical expressions is derived from the category of pattern of the mathematical elements and consists of the category of equations, functions, simple operations, differentiation, and integration: $\ldots\langle\sigma_{\mathrm{Pt}}\rangle \triangleright ..\langle\nu_{\mathrm{MtEx}}\rangle \triangleright \langle\nu_{\mathrm{Eqt}},\nu_{\mathrm{Fun}},\nu_{\mathrm{Alg}}\rangle$. Naming of the perceived object by the name of one of the specific mathematical expressions requires classifying it to one of the specific categories of the mathematical expressions. The knowledge schema supplies knowledge that is needed to perform the syntactic analysis that classifies the mathematical expression into one of the mathematical expression categories. Naming of the mathematical expression can start further thinking process that leads to solving the problem that a given expression represents. For example, the following visual object "$2x + 5 = 4x - 3$" will be classified as the category of linear equations and named "the linear algebraic equation." When the name is found the system during further thinking process will try to solve this equation. The knowledge that is needed to perform a nonvisual reasoning aimed at solving the problem represented by the mathematical expression, is given by the knowledge chain: $\ldots \triangleright \langle\kappa_{\mathrm{MthO}}\rangle \triangleright \langle\kappa_{\mathrm{MtEx}}\rangle \triangleright \langle\kappa_{\mathrm{Eqt}},\kappa_{\mathrm{Fun}},\kappa_{\mathrm{Alg}}\rangle$. The knowledge about how to solve the problem that is named as "mathematical expression" can be obtained by making link to Mathematica package. After sending the mathematical expression to Mathematica the solution that is obtained can be further interpreted during the thinking process.

During the thinking process aimed at solving mathematical problem each mathematical expression needs to be interpreted and next an appropriate action that transforms a mathematical expression into required solution needs to be undertaken. The interpretation of the visual symbols of the mathematical expression refers to the perceptual pattern category and requires the knowledge derived from the pattern category. For example, the sign "+" between two ciphers indicates that an appropriate action needs to be undertaken; two ciphers need to be summed up. In the case of the complex operators such as integral the mathematical operation can be performed by Mathematica. It should be noticed that even the simple expression "8 + 4" is interpreted as "seeing" pattern of three visual symbols in the form of the mathematical expression. Each object is first interpreted as a member of one of the number categories ("8" and "4") and as a category of algebraic operators ("+"). Next visual inference is performed and action is undertaken. The category of algebraic operations supplies the knowledge about the action that needs to be undertaken.

Solving task given by the mathematical expression requires performing the mathematical operation that is indicated by the mathematical operators. The results can be examined visually by exploring the results in the form of graphs or can be interpreted based on the results of the symbolic computation. For example, solution given in the form $y = \int x = 0.5x^2$ can be evaluated by examining properties of the graph of the function $y = x^2$.

During the learning stage the visual concepts of mathematical symbols are learned and are represented in the form of the categorical chains. During understanding process the knowledge that is learned is utilized to interpret the visual object based on the knowledge stored in the knowledge schema.

5.5.2.2.2. Naming of Symbolic Signs

Naming of the object as a sign is to classify the object to one of the sign categories. The sign category is derived from the category of visual objects. As it was described in Chap. 4 the category of signs is divided into category of visual symbols and the category of the symbolic signs. The specific categories derived from the category of symbolic signs are represented by the following categorical chain: $O \triangleright \langle v_{\mathrm{Sig}} \rangle \triangleright \langle v_{\mathrm{2DSig}} \rangle \triangleright \langle v_{\mathrm{SymS}} \rangle \triangleright \langle v_{\mathrm{RoS}}, v_{\mathrm{CroS}}, v_{\mathrm{TrS}} \rangle$. The meaning of elements of the category of symbolic signs does not depend on the meaning of other elements of the same category. The name of the sign is assigned to the sign category during the learning process. During the naming, the perceived object is classified into one of the sign categories. For example, the sign with a letter "P" is classified to the category of the information sign. The knowledge about relation to other general sign categories is obtained from the categorical chain. Based on the categorical chain reasoning we can infer that the information sign is part of the category of the road sign and meaning of the object is obtained from the knowledge schema of the road signs.

In this section an example of analysis of the specific category of signs, namely, the cross signs is presented. This section is not intended to analyze all existing cross signs but rather it is an attempt to show how shape categories can be utilizes to provide the tools needed for formal description of the signs category. It is assumed that the cross sign is the 2D object. Examples of category of cross signs are shown in Fig. 5.31. Meaning of the cross sign is given by the knowledge of the cross sign category. To learn the visual concept of the selected category of the cross signs, a set of phantoms that are representatives of a given cross sign category needs to be selected and analyzed. The general category of the "cross" is obtained by selecting representatives of most often used crosses. The representative of the specific category of the cross, for example, the "Maltese cross" is obtained by selecting representatives of most often used “Maltese crosses.” The categorical chain of the different cross categories is as follows: $\ldots \rhd \left\langle v_{\mathrm{SymS}} \right\rangle \rhd \left\langle v_{\mathrm{CroS}} \right\rangle \rhd \left\langle v_{\mathrm{Grk}}, v_{\mathrm{Lat}}, v_{\mathrm{X}}, v_{\mathrm{StA}}, v_{\mathrm{Pat}}, v_{\mathrm{Pap}}, v_{\mathrm{Lor}}, v_{\mathrm{Mal}}, v_{\mathrm{Cel}}, v_{\mathrm{Chr}} \right\rangle$. Learning of the categorical knowledge consists of two stages: learning of the visual knowledge (visual concept) and learning of the nonvisual knowledge. The nonvisual knowledge is part of the knowledge schema of the visual object. As it was described in previous parts of this book to perform thinking and understanding there is no need to learn very specific knowledge for each category. In learning, the important issue is to learn the basic knowledge (contained in knowledge schema of the visual object) and connection among the general categories. The specific knowledge can be “added” during learning of the knowledge from the selected specific knowledge domain. For example, to understand the visual object, let us say an apple, there is a need to know that the apple is a part of the tree, that a tree is a plant, that a plant can be found on the earth. In addition, we need to know some features of the apple and have some links to some products that can be obtained from the apple. The specific knowledge concerning some chemical or biological processes can be learned when SUS will be used as an expert in the area of botanic.

The general visual concept of the visual object is given by the structural archetypes and archetypes of classes that are “perceptually” linked with the class of the structural archetypes. The structural archetype represents the simplest form of the visual concept of the object of a given ontological category. The visual concept of the cross is given by classes “perceptually” linked with the class of the structural archetypes of the cross class. For example, the visual concept of the cross is given by the shape classes “perceptually” linked with the class of the structural archetypes of the cross class, that are derived from the star class. The star class, described in Chap. 2, is a class derived from the concave class. The different crosses have different shapes and are result of the conventional transformation of the general visual concept of the cross. However, visual objects that are conventional visual representations of the cross category have the common structural elements. These representations are members of the classes that are perceptually linked with the class of the structural archetypes.

The cross class is the a posteriori class that is derived from the star class. Posteriori classes, described in Chap. 2, are derived from one of the general classes where specific classes are established based on a set of objects that are members of the selected ontological category. The structural archetype of the cross class, the thin cross class, is derived from the thin star class $\otimes$ (described in Chap. 2), and is given by the symbolic

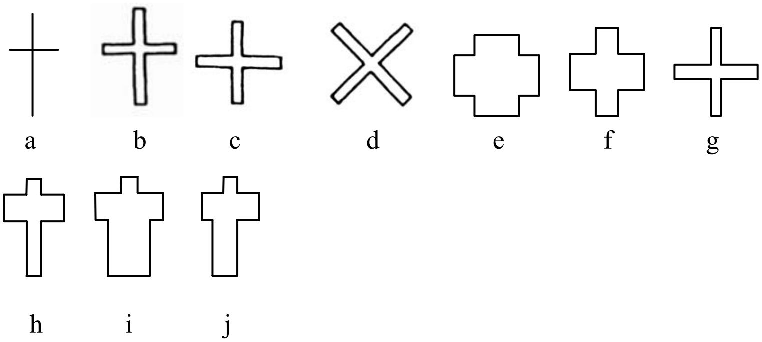


Fig. 5.31. Structural archetype of the cross (**a**) and different crosses generated from the class, (**b, h**) a Latin cross, (**d**) a Saint Andrew cross, (**c, e, f, g**) a red cross, (**i, j**) not members of a cross class

name $[C]\oplus/\rho[L^4]\{4L^3\}$. The symbol $[C]$ denotes that the thin cross class is derived based on the properties of the existing cross signs. The difference among archetypes of the thin star class $\oplus/\rho[L^4]\{4L^3\}$ and thin cross class $[C]\oplus/\rho[L^4]\{4L^3\}$ is such that archetypes of the thin cross class are symmetrical and regular. Classes that are "perceptually" linked with the structural archetype of the cross class are derived from the $2n$-star class when $n=4$ $[C]Q^4[L^8](4L^3)$. Attributes of the archetypes of this class are constrained by fulfilling special conditions. One of the conditions is that the archetypes of the cross class are symmetrical objects. For example, the object shown in the Fig. 5.31i, j is member of the $Q^4[L^8](4L^3_R)$ class, however, it is not a member of the cross class $[C]Q^4[L^8](4L^3_R)$. From the cross class the specific cross classes are derived. For example, the Latin cross class is the class in which the residuals are archetypes of the right triangle class, and the length of the sides has to fulfill special conditions (e.g., the established proportions) Fig. 5.31b, h. As it was described in Chap. 2 the attributes such as the length are expressed by applying the graded values: $a_i^d \in \{\varepsilon, s, m, L\}$, where ε denotes a "very small," s s denotes a "small," m denotes a "medium" and L L denotes a "large" value. The Latin cross class is given by the symbolic name $Q^4[L^8](4L^3_R)[\varepsilon mm\varepsilon ml\varepsilon l\varepsilon mm]$. Similarly the Saint Andrew cross class Fig. 5.31d and the red cross class Fig. 5.31c,e,f,g are given by symbolic names as follows: $Q^4[L^8](2L^3_A 2L^3_O)[\varepsilon ll\varepsilon mm\varepsilon ll\varepsilon mm]$, and $Q^4[L^8](4L^3_R)[\varepsilon mm\varepsilon mm\varepsilon mm\varepsilon mm]$.

Classes that are named as the cross-like classes are established through the generalization process (abstraction) and are given by their symbolic names as $Q^4[\Lambda](4\Gamma)$. Generalization is the process in which a part of the symbolic name is "removed" to represent the unconstrained value of the selected attributes. For example, the general red cross class is given as $[C]Q^4[L^8](4L_R^3)$ and is obtained by removing $[\varepsilon mm\varepsilon mm\varepsilon mm\varepsilon mm]$ part from the symbol $Q^4[L^8](4L_R^3)[\varepsilon mm\varepsilon mm\varepsilon mm\varepsilon mm]$.

The visual concept of the cross category is obtained during the learning process. During the learning process objects (phantoms) that represent the cross sign are selected. Examples of the selected signs are shown in Fig. 5.32. As the result of the learning process the visual concept is obtained. The visual concept reflects the similarities among the visual objects that are members of the same concept. Learning of the visual concept of the cross sign includes learning of the visual concept of the specific categories of the cross sign. At first the most often used categories of cross sign are selected and the most typical representatives of these categories are used to learn the visual concept. Figure 5.32 shows the sign from the different categories of cross sign used for learning of the visual concept. The similarity of visual objects is reflected by the symbolic names of shape classes that these elements represent. For example, all classes archetype of which are shown in Fig. 5.32e, f, h are derived from $Q^4[\Lambda](4\Gamma)$ class, whereas classes archetype of which are shown in Fig. 5.32a, b, c, d, g, i, j are derived from $Q^8[\Lambda](8\Gamma)$ class. Figure 5.32 shows the different crosses given by their symbolic names. The symbolic name can be given in the form of complex class described in Chap. 2. The notation in terms of the complex class gives description in terms of the different cross "arms" that seem to be more perceptually oriented. As it was mentioned before, the visual concept of the cross sign is a set of symbolic names obtained during the learning process. In this example, the visual concept of the different specific category of the cross sign (e.g., red cross) is represented by the one symbolic name. For example, the visual concept of the specific category of the

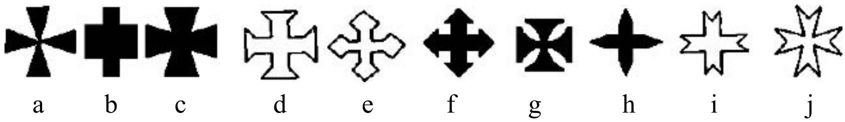


Fig. 5.32. Examples of members of the cross category used in learning the visual concept

cross sign – the red cross sign is given by the symbolic name $\{[C]\{Q^4[L^8](4\cdot L_R^3)\}\equiv S^4[L^8]\{L_R^4\}(4L_R^4)\}$. Similarly, the visual concepts of the specific cross sign category shown in Fig. 5.32 are obtained and are given by the symbolic names as follow: (a) $[C]Q^4[L^8](4L_A^3)$, $S^4[L^8]\{0\}(4L^3)$, (b) $[C]\{Q^4[L^8](4\cdot L_R^3)\}$, $S^4[L^8]\{L_R^4\}(4L_R^4)$, (c) $[C]Q^4[L^8](4\cdot L_A^3)$, $S^4[L^8]\{L_A^4\}(4L_T^4)$, (d) $[C]Q^4[L^8](4\cdot L^5)$, $S^4[L^8]\{L_R^4\}(4Q[L_T^4](2L^3))$, (e) $[C]Q^4[L^4](4\cdot L^5)$, $S^4[L^4]\{L_R^4\}(4Q[L^5](2L^3))$, (f) $[C]\{Q^4[L^4](4\cdot L^5)\}$, $S^4[L^4]\{L_R^4\}(4Q[L^5](2L_R^3))$, (g) $[C]\{Q^4[L^8](4\cdot M^1(c_l^f))\}$ $S^4[L^4]\{0\}(4Q[L^3](2M^1))$, (h) $[C]Q^4[L^4](4\cdot Q^2[L^3](2\cdot L^3))$, $S^4[L^4]\{L_R^4\}(4L^5)$, (i) $[C]\{Q^8[L^8](4\cdot(L_R^3\tilde{L}^3))$, $S^4[L^8]\{L_R^4\}(4Q[L_R^4](L^3))$, (j) $[C]\{Q^8[L^8](4\cdot(L_A^3L^3))$, $S^4[L^8]\{L_R^4\}(4Q[L_T^4](L^3))$.

Naming requires finding differences between similar objects and because of this learning of the visual concept of the category of cross signs requires learning of the visual concept of different categories. For example, the different mechanical tools shown in Fig. 5.33 can be called cross-like because they are similar to the cross signs. Visual objects shown in Fig. 5.33 are members of the figure category (cross-like figure). The symbolic names of the phantoms shown in Fig. 5.33 are as follows: (a) $Q^4[L^4](4L^3)$, (b) $Q^4[L^8](4L^3)$, (c) $Q^8[L^8](4L^3L^3)$, (d) $Q^8[L^8](4L^3L^4)$, (e) $Q^4[L^8](4M^1)$, (f) $Q^8[L^8](4M^1L_R^4)$, (g) $Q^8[L^8](4M^1Q^2[L^3](2M))$, (h) $Q^4[\hat{L}^4](4L^3)$, (i) $Q^4[\hat{L}^4](4M^1)$, (j) $Q^8[\hat{L}^8](4M^1L^3)$, (k) $Q^4[\hat{L}^4](4Q^2[\hat{L}^3](2M))$, (l) $Q^4[\hat{L}^4](2MQ^2[\hat{L}^3](2M))$, (m) $Q_M^4(4L^4)$, (n) $Q^4[M](4L^3)$, (o) $Q^4[M](4Q^2[\hat{L}^3](2M))$. As we can see some of these parts have the visual concept represented by the symbolic name which is very similar to the visual concept of the category of the cross sign.

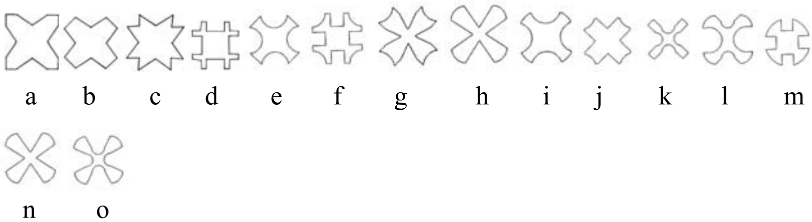


Fig. 5.33. Cross-like figures

5.5.2.3. Letter Naming

The letter naming is to classify the visual object as a member of one of letter categories. Classification of the object into one of the letter categories refers to the knowledge of the language category. The category of alphabetic letter is divided into the category of Latin letters or the category of Greek letters and is represented by the following categorical chain: $O \triangleright \langle \nu_{\mathrm{Let}} \rangle \triangleright \langle \nu_{\mathrm{Alp}} \rangle \triangleright \langle \nu_{\mathrm{Lat}}, \nu_{\mathrm{Gre}}, \nu_{\mathrm{Cyr}}, \nu_{\mathrm{Heb}}, \nu_{\mathrm{Ara}} \rangle$. Knowing that perceived object is a member of the letter category of the given alphabet makes it possible to predict the next perceived object by assuming that it is another member of this category. Naming letters makes it possible to understand the text and performing an action that can be described by the text. Naming letters is the first step in understanding of the text that involves the sequences of complex thinking sub-processes. Classification of the examined object into a letter category may start the thinking process that can lead to different conclusions. For example, the letter in the book can start process of reading whereas letter that is part of the information sign can give information about some road facilities.

The letter naming requires knowledge supplied by the letters category. During naming process the examined object is classified into one of the letter categories. The letter is understood in the context of the language and it refers to the rules of the composition of the word and text given by the knowledge of the language categories. Although there is a difference between the category of nonalphabetic languages and the category of alphabetic languages understanding an object as a member of the category of the nonalphabetic language is similar to understanding a letter as a member of an alphabetic language. Most often the letter is understood in the context of a text that conveys given information. The text consists of sentences that in turn consist of words. During the thinking process, at first the language category is identified based on the sample of letters. When the language category is identified the knowledge schemas of this language category supplies knowledge that make it possible to read the text. Understanding of the text requires nonvisual knowledge that is supplied by knowledge categories.

5.5.2.3.1. Naming of the Different Fonts of the Letter

The letters naming requires knowledge that is acquired during the learning process. In this section an example of the learning of the category of selected fonts of the letter "T" is presented. To learn the visual concept of a given letter category there is a need to learn the prototype of the specific font of this letter. The visual concept needs also to include the prototype of

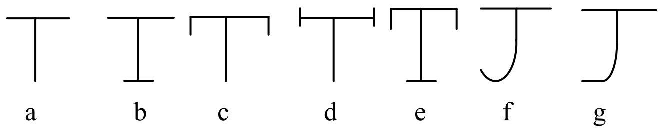


Fig. 5.34. Examples of structural archetypes of the different fonts of the letter "T"

the handwritten letters. The visual concept of the selected handwritten letter needs to include the huge spectrum of the different shapes of this letter.

The font of the letter "T" can be well represented by the structural archetype. The structural archetypes of the different fonts of the letter "T" are learned based on visual representatives of these fonts. At the first stage of learning process representatives of most typical fonts are selected. The most typical representative of the letter "T" shown in Fig. 5.34a is used to learn the visual concept of the basic prototype of the letter "T." The visual concept of the basic prototype of the letter "T" is given by the symbolic name $\oplus/\rho[L^3]\{2L_R^3\}$. Examples of structural archetypes of the different fonts of the letter "T" are shown in Fig. 5.34. The symbolic names of structural archetypes of the letter "T" are as follows (see Fig. 5.34): (a) $\oplus/\rho[L^3]\{2L_R^3\}$, (b) $\oplus/\rho[L^4]\{2L_T^4\}$, (c) $\oplus/\rho[L^5]\{2L_T^4\}$, (d) $\oplus/\rho[L^5]\{2L_T^4 L_R^4\}$, (e) $\oplus/\rho[L^6]\{2L_T^4\}$, (f) $\oplus/\rho[M^1[L^4]]\{M^1[L^4], Q^1[L^3](M^1)\}]$, (g) $\oplus/\rho[\widehat{L}^4]\{\widehat{L}^4, Q^1[L^3](M^1)\}$

The structural archetypes of the letter "T" $\oplus/\rho[L^3]\{2L_R^3\}$ are perceptually linked with structural exemplars generated from the class $Q^2[L^6](2L_R^3) = C(2L_R^4)$. Exemplars generated from the class $Q^2[L^6](2L^3)$ can have different visual forms. Figure 5.35 shows different archetypes generated from the class $Q^2[L^6](2L^3)$. Archetypes shown in Fig. 5.35b, c, and g cannot be regarded as the representatives of the letter "T." The archetypes shown in Fig. 5.35d–h can be regarded as representatives of the distorted version of the letter "T." To find the specific class, archetypes of which can be regarded as the representatives of the letter "T," the constraints need to be imposed on values of attributes of the archetype of the class $Q^2[L^6](2L^3)$. Constraints that are obtained during learning process are used as a criterion of the derivation of the specific classes that are used as shape categories in matching the categories of the letter font. To derive the specific class from the class $Q^2[L^6](2L^3)$ angles of the archetypes of this class need to be specified and the archetypes need

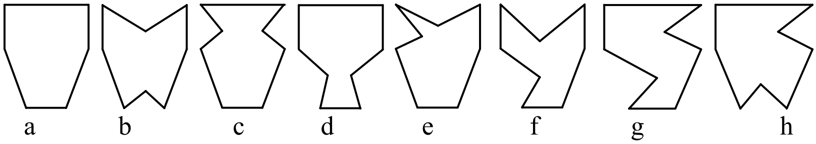


Fig. 5.35. Exemplars generated from the class $Q^2[L^6](2L^3)$

to be a member of the complex class $Q^2[L^6](2L^3) = C(2L_R^4)$. Exemplars generated from the class that is too "general" cannot be interpreted as a letter "T." Exemplars generated from the class $Q_{L^6}^2(2L_R^3) = C(2L_R^4)$ can be perceived as exemplars of the class $Q^2[L^{3-6}](2L_R^3) = C(2L_R^4)$, where L^{3-6} denotes polygon that has 3–6 sides (see Fig. 5.36 (a–h)). The class $\bar{Q}^2[L^{3-6}](2L_R^3) = \bar{C}(2L_R^4)$, members of which are symmetrical "T" shaped polygons, are good representatives of the letter "T."

The description of specific classes derived from the class $Q_{L^6}^2(2L^3)$ is given using the notation of the cyclic model $\perp^n \{(n\alpha_i),(nd_i)\}$, where symbol $(n\alpha_i)$ denotes angle and symbol (nd_i) denotes length of sides of the archetype. The angle can have values $\alpha \in \{\pi, o, a, \varepsilon\}$, where π – denotes right angle, o – an obtuse value, a – an acute value and ε – a very small angle whereas the side can have values $d \in \{\varepsilon, s, m, L\}$, where ε denotes a very small value, s – a small value, m – a medium value and L – a large value. Examples of the exemplars generated from classes $\bar{\perp}^8 \{(8\pi),(mmLmLmmL)\}$ or $\bar{\perp}^8 \{(8\pi),(\varepsilon mL\varepsilon Lm\varepsilon L)\}$ are shown in Fig. 5.36b and Fig. 5.36e. The visual concept of the prototype of the letter fonts is based on the structural archetype $\oplus/\rho[L^3]\{2L_R^3\}$ and consists of the symbolic names of the specific classes $\perp^n \{(n\alpha_i),(nd_i)\}$.

The letter can be subjected to the different forms of distortions (e.g., handwritten letter). Different fonts can also have the shapes that are different from the "typical" representation. Examples of the distorted versions of the letter "T" that can be interpreted as special "fonts," generated from the

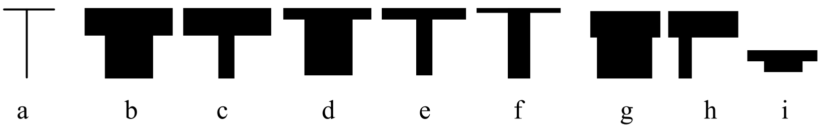


Fig. 5.36. Structural archetype $\oplus/\rho[L^3]\{2L_R^3\}$ $\oplus/_{\rho}[L^3]\{2L_R^3\}$ and exemplars generated from the class $Q^2[L^6](2L_R^3) = C(2L_R^4)$

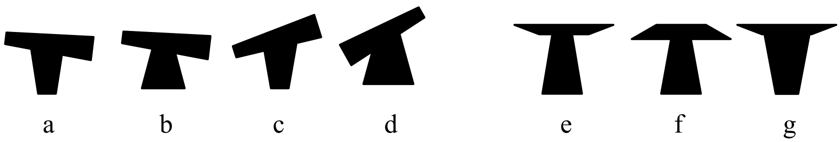


Fig. 5.37. Examples of distorted versions of the letter "T"

class $Q^2[L^6](2L^3)$ are shown in Fig. 5.37. During learning of the visual concept of the letter "T" there is a need to include distorted version of the letter "T" under category of the distorted versions of the letter "T." The symbolic names of the selected representatives of the distorted versions of the letter "T" generated from the class $Q^2[L^6](2L^3)$ are as follows: $Q^2[L^{3-6}](2L^3) = C(2L^4)$ (Fig. 5.37a–c), $Q^2[L^{3-6}](Q^1_{L^3}(L^3), L^3) = C(2L^4)$ (Fig. 5.37d), $Q^2[L^4](2Q^1_{L^3}(L^3)) = C(2L^4)$ (Fig. 5.37e), $Q^2[L^{4-6}](2L^3) = C(2L^4_T)$ (Fig. 5.37f), $Q^2_{L^4}(2L^3) = C(2L^4_T)$ (Fig. 5.37g).

Different fonts of the letter "T" are represented by the different structural archetypes. As it was mentioned, the structural archetypes that are members of thin classes are "perceptually" linked with archetypes of concave classes that have the "similar" appearance. Examples of structural archetypes and exemplars generated from the "perceptually linked class" are shown in Fig. 5.38. As we can see the structural archetypes represent only the main features of the letter "T" and usually there is more than one "perceptually" linked class. For example, the structural archetype shown in Fig. 5.38a is linked with the exemplar generated from the "perceptually" linked class given in Fig. 5.38b, c. The "perceptual" link can be expressed by showing all symbolic names "linked" with a structural archetype. For example, Fig. 5.38a–c, (a) $\oplus/\rho[L^5]\{2L^4_T\}$, (b) $(Q^2[L^6](2Q^1[L^3](\hat{L}^3)) = C(L^4_R, Q^1[L^4_R](L^4))$, (c) $Q^2[L^{5-8}](2Q^1[L^3](\hat{L}^3)) = C(L^4_R, Q^1[L^{4-6}](L^4))$ $Q^2[L^{5-8}](2Q^1[L^3](\hat{L}^3)) = C(L^4_R, Q^1[L^{4-6}](L^4)))$. Figure 5.38d–e, (d) $\oplus/\rho[L^6]\{2L^5\}$, (e) $(Q^2[L^{5-6}](2Q^2[L^5](\hat{L}^3)) = C(2L^4_R, Q^1_{L^4_R}(L^4))$, Fig. 5.38f–g, (f) $\oplus/\rho[L^5]\{2L^4_T, L^4_R\}$, (g) $(Q^3[L^{5-6}](2Q^1[L^3](\hat{L}^3, L^4)) = C(L^4_R, Q^2_{L^4_R}(2L^4))$.

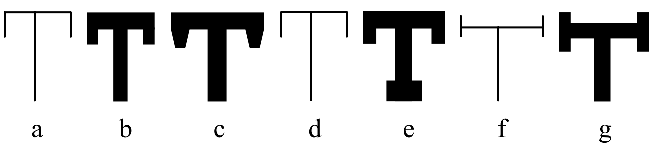


Fig. 5.38. Structural archetypes and exemplars of the letter "T"

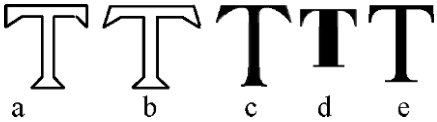


Fig. 5.39. Similar fonts of the letter "T" and their symbolic names

Figure 5.39 shows the different fonts of the letter "T" that correspond to the structural archetype $\oplus/\rho[L^6]\{2L_T^4\}$. Differences in "ending" of the stroke of the letter and similarity of the "general" shape are reflected in the symbolic names of the shape. The first two fonts shown in Fig. 5.39 are very similar that is reflected in their symbolic names, for example, in Fig. 5.39a $Q^2[L^6](2L^5) = C(Q^2[L^4](L^3), Q^1[L_R^4](L^4))$, and in Fig. 5.39b $Q^2[L^6](2L^5) = C(Q^2[L^4](L^3), Q^1[L^4](L^4))$. The fonts in Fig. 5.39c–e have curvilinear segments that are shown in symbolic names (the residuals are archetypes of the curve polygon classes M). The symbolic names for the fonts shown in Fig. 5.39c–e are as follows: (c) $Q^2[L^6](2M^2[L^4]) = C(Q^2[L^4](2M^2[L^4]), Q^1[L^4](M^2[L^4]))$, (d) $Q^2[L^6](Q^2[\widehat{L}^4](\tilde{L}^3, Q^1[L^3](M^1))) = C(\ddot{Q}^2[L_R^4](M^2[L^4]), Q^2[L^{4-6}](2L^3))$, (e) $Q^2[L^6](Q^2[M^2[L^6]](\tilde{L}^3, Q^1[L^3](M^1))) = C(\ddot{Q}^2[L^4](M^2[L^4]), Q^2[L^{4-6}](2M^1[\widehat{L}^4]))$.

Different fonts of the letter can have different visual appearances that are very often similar to another letter. For example, the structural archetypes of the letter "T" for the different fonts with curvilinear segments are shown in Fig. 5.40. The structural archetype preserves only the essential features of the letter "T." Learned structural archetype is part of the visual concept of the fonts of the letter "T." Based on the visual concept during the visual interpretation we can find all letters that have similar visual appearance. The symbolic names show similarities among archetypes and based on these similarities we can predict other fonts that can be interpreted as the letter "T." Symbolic names of structural archetypes of different fonts of the letter "T" shown in Fig. 5.40 are as follows: (a) $\oplus/\rho[M^2[L^5]]\{M^2[L^5], L^3\}$, (b) $\oplus/\rho[M^2[L^5]]\{Q^2[M^2[L^5]](2\Theta^2(c)), L^3\}$, (c) $\oplus/\rho[Q^1[M^2[L^5]](M^1)]\{M^1, M^2[L^5], L^3, Q^1[M^2[L^5]](M^1)\}$, (d) $\oplus/\rho[Q^1[M^2[L^5]](M^1)]\{M^1, L^3, Q^1[M^2[L^5]](M^1)\}$.

The symbolic names of the letters that are obtained during analysis are used during categorical learning of the different categories of the letter "T."

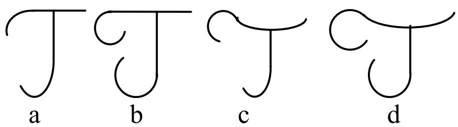


Fig. 5.40. Structural elements of the different fonts of the letter "T"

Learning of the letter concept could be reduced to learning of the visual concept. The nonvisual knowledge is learned independently and placed in the context of the categorical chain. This categorical learning is performed in many stages in which testing of the learned result by using of the "distorted" version of the letter are important part of the categorical learning process. Testing is concentrated on redefinition of the visual concept by investigating the different "distortions" of the most common visual representations of the letter. For example, the "redefinition" of the visual concept of the letter "A" was investigated as part of the experiment aim of which was to learn distorted version of the letter "A." During testing stage of this experiment 50 cyclic objects were given to SUS and SUS needed to select phantoms that are representatives of the visual concept of the letter "A." Examples of phantoms that were recognized as the letter "A" are shown in Fig. 5.41. The result of SUS performance was evaluated by human observer. Performance of the SUS was very good (100% cases) in the case when parts of letters were "well visible."

The proper interpretation of the object depends on the level of details of the symbolic name that was obtained in the process of visual reasoning. The level of detail refers to the class description. For example, the concave object with a hole (shown in Fig. 5.41) can be interpreted as a letter "A." The visual concept of the letter was obtained during the learning stage. The learning was concentrated on defining of the visual concept by investigation of the different "distortions" of the most common visual representations of the letter "A." All changes were evaluated and only these objects that were evaluated as the letter "A" were used to form a visual concept of the letter "A." In the case of objects shown in Fig. 5.41(a–h) the hole can be any convex object. In the case when the hole is concave object there is a need to learn a new concept (definition) of the letter "A." This problem is rather complex because of lack of agreement in interpretation of such objects by human observers. The letters in Fig. 5.41g, h can be

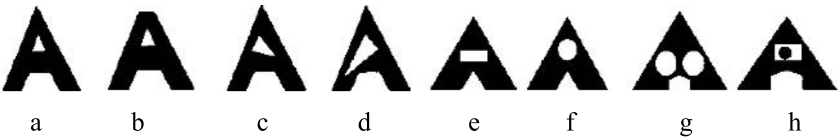


Fig. 5.41. Examples of the phantoms that can be recognized as the letter "A"

interpreted as special fonts of the letter "A." However, not all human subjects can interpret these objects as a letter "A." To define the concept of the letter "A" that can be used in interpretation of all possible cases of the letter "A" two solutions are possible. The first solution is to select all possible fonts of the letter "A" and assume that these objects are representatives of the visual concept of the letter "A." The second solution is to define the letter "A" as the concave objects with a hole (shown in Fig. 5.41). The hole can be any convex or concave object. The definition can be extended to include more than one hole, or the hole that is the cyclic object. The second solution has a big advantage because of the application of the generalization process. In the proposed approach a letter with a "big distortion" could be easy to recognize. This capability can play a big role when recognizing (reading) the signs (e.g., road signs) represented in the form of the letters as well as reading the patterns that are not easy to interpret, e.g., in science, in archeology.

5.5.2.3.2. Similarities of Different Letters

In the previous section learning of the different fonts of the same letter was described. In this section learning of letters from different alphabets is described. As it was shown at the beginning of the previous section naming letters requires learning of a letter as a member of the category of alphabetic letters. An alphabet is usually given as a set of letters, where each of letters is characterized by specific shape and the position of the letter in the alphabetical order. The letter is one of the characters (visual objects) of the alphabet and the same visual object can represent different letters from the different alphabets. For example, letters that are extracted from the different alphabets (Hebrew, Arabic, Cyrillic, Greek, and Latin) are shown in Fig. 5.42. To show the difference among letters an Arial font was selected because the Arial font gives the simplest form of the printed letter. Due to the big variation between letters of the different font comparison of the letters of the different alphabets needs to take into account these differences.

ו ן ا L Γ J ך ר ل ז ح V Λ Π U C ח ر د Л
N И Z ל م M W Σ Ω S ى G פ ף ט T т
ב Y У Ч ג ש ת מ ع Ц Ш E Э З F Щ X
Ψ א H K К Ж س Δ D ם ه O P Ь و Б Ъ
ط ص R Я Ю Д Q Θ B Φ ה ק Ξ Ы Й

Fig. 5.42. Example of letters that are extracted from the different alphabets (Hebrew, Arabic, Cyrillic, Greek, and Latin)

During learning of the visual concept of the letter category the visual similarities among representatives of the different category of alphabetic letters as well as the different sign categories or the real-world object categories needs to be taken into account.

Learning of the visual concept of a given letter category requires selecting representatives of that letter as well as representatives of other categories similar to that letter. Next, based on selected representatives, the shape classes that well discriminate among letters and other objects are established. The shape classes that are derived from a priori classes, based on properties of the visual object such as properties of the different letters, are called the a posteriori classes. These classes are described in Chap. 2. The Fig. 5.42 shows selected letters of the different alphabets (Hebrew, Arabic, Cyrillic, Greek, and Latin). The general classes such as the thin convex class or the cyclic class do not well discriminate among objects shown in Fig. 5.42. All objects shown in Fig. 5.42 are thin objects. Specific thin classes need to be derived from the thin class to supply the shape categories that will well match the different letters. During the learning process the objects of other categories such as visual symbols need to be considered as similar objects.

In this section examples of the letters from Latin, Cyrillic, Greek, Arabic, and Hebrew alphabet will be analyzed. The visual appearance of the different letters can be very similar. Understanding letters requires differentiating between similar categories of letters of the same alphabet and interpreting the letter as the letter from one or more than one alphabet. For example, the visual symbol of the letter “C” has different meanings in Latin and Cyrillic language. The general visual concept of the letter is given by the structural archetype that represents the simplest form of the letter concept. As it was described in previous sections, the visual concept of the letter is acquired during the learning process. Learning of the visual concept of the letter is based on learning of the general structural description given by the structural archetype of a given letter. The aim of the learning process is to find visual concepts of different letters that makes it possible to recognize a later among other similar objects. For example, the structural archetype of the letters “I” (Fig. 5.43a) and letters “L” and “Г” (Fig. 5.43b, c) is different. The symbolic names of these archetypes are as follows: Θ^2 (Fig. 5.43a) $\otimes[L_R^3]$ (Fig. 5.43b, c). Not all archetypes generated from the class $\otimes[L_R^3]$ can be regarded as the letter “L.” To match the archetypes of the letter “L” the new specific class $\otimes[L_R^3][l,s]$ described in Chap. 2, is introduced. Archetypes of this class regarded as shape categories will be used to match both letters “L” and “Г.” To find the proper archetype that matches the letter “Γ” the sub-specific class that indicates the spatial orientation of the object is introduced. Letter “Γ” is described by the class $\otimes[L_R^3][l,s]\{'L'(O^{2R})\}$, where $\{'L'(O^{2R})\}$ is the symbol used for denotation of the spatial orientation of the visual object. The symbol O^{2R} denotes rotation of the object, and R denotes the right angle. The letter “L” is given by the visual symbol $\otimes[L_R^3][l,s]\{'L'(OH^{2R})\}$ or $[L]\otimes[L_R^3]$, where letter “L” at the beginning indicates the sub-specific class, the letter “L” class. Figure 5.43 shows similar letters from the classes Θ^2, Ξ, and $\Im$. The symbolic names of the structural archetypes of the letters shown in Fig. 5.43 are given as follows: (a) Θ^2, (b) $\otimes[L_R^3][l,s]\{'L'\}$, (c) $\otimes[L_R^3][l,s]\{'L'(MO^{2R})\}$, (d) $\otimes[M^1[L^3]][l,C(s)]\{'l'(O^{2R})\}$, (e) $\otimes[M^1[L^3]][m,C(m)]\{'l'(O^{2R})\}$, (f) $\otimes[M^1[L^3]]['l',C(s)]\{'l'(M)\}$, (g) $\otimes[\tilde{M}^1[L^3]][m,C(m)]\{'l'(M)\}$, (h) $\Theta/\rho[L^3]\{M^1,Q[L^3](M^1)\}]$, (i) $\Theta/\rho[M[L^4]]\{M^1,Q[L^3](M^1)\}$, (j) $\otimes[\bar{L}_A^3]\{'V'\}$, (k) $\otimes[\bar{L}_A^3]\{'V'(O^{2R})\}$, (l) $\otimes[L_R^4]\{'U'(O^{2R})\}$, (m) $\otimes[M^1[L^4]]\{'U'\}$, (n) $\otimes[M^1[L^4]]\{'n'\}$, (o) $\otimes[M^1[L^3]]$, (p) $\otimes[Q^1[L^4]](M^1)$, (q) $\otimes[M^1[K_E^1]]\{'C'\}$, (r) $\otimes[M^1[K_E^1]]\{'C'M(\alpha)\}$.

These letters are different in this respect that they are rotated or they are curvilinear version of the selected letters. For example, the letter “Γ” is the rotated version of the letter “L.” Figure 5.43 shows letters that are different in this respect that some of them have the curvilinear segments instead of the linear one.

Figure 5.45(a–f) shows letters that are different in this respect that they are rotated versions of one another. The structural archetypes of the letters “N,” “И,” and “Z” are as follows: $\Theta/\rho[[L_R^4]\{2L^3\}]\{'N'\}$. The

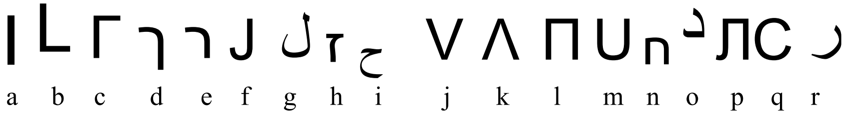


Fig. 5.43. Similar letters from the letter classes

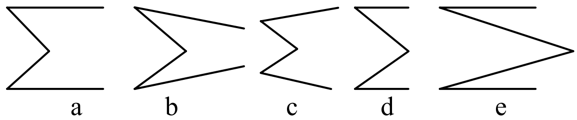


Fig. 5.44. Similar letters "M," "W," "Σ," and the mathematical symbol "∑"

symbolic names of the letters "И" (Fig. 5.44b) $\Theta/\rho[L_R^4]\{2L^3\}]\{'N'M\}$ and "Z" (Fig. 5.44c) $\Theta/\rho[[L_R^4]\{2L^3\}]\{'N'O^R\}$) are rotated versions of the structural archetype of the letter "N" (Fig. 5.44a) $\Theta/\rho[[L_R^4]\{2L^3\}]\{'N'\}$.

The letters "M," "W," "Σ," and the mathematical symbol "∑" are described by the symbolic name $\otimes[Q^1[L_R^4](L^3)]$ (see Fig. 5.44a). The letter "W" and a specific font or a distorted version of letters "M," "Σ," or the mathematical symbol "∑" are described by the symbolic name $\otimes[Q^1[L^4](L^3)]$ that is generalization of the class $\otimes[Q^1[L_R^4](L^3)]$. The most common font of the letter "W" is described by symbolic name $\Theta/\rho[Q^1[L_T^4](L^3)]\{2L_A^3\}]$ and the most common font of the letter "M" is described as $\Theta/\rho[Q^1[L_R^4](L^3)]\{2L_R^3\}]$ Fig. 5.44d. The distorted version of the letters "M," "W," "Σ," or the mathematical symbol "∑" is given by the symbolic name $\Theta/\rho[Q^1[L^5](L^3)]\{2L^3\}]$ Fig. 5.44c. The letter "W" (Fig. 5.44e) letter "Σ" and the mathematical symbol "∑" are rotated versions of the structural archetype of the letter "M." The proper identification of the letter reqires finding the sp ecial orientation of the letter. The sub-specific classes are introduced to supply the shape categories for matching a given letter. As the base of the spatial orientation of the letter type "M," the letter "M" is selected and a sub-specific class of the letter "M" is denoted by adding the symbol $\{'M'\}$, e.g., $\oplus[Q^1[L_R^4](L^3)]\{'M'\}$. The shape classes that are used to match the letter "W," the letter "Σ" and the mathematical symbol "∑" are denoted by adding the symbol $\{'M'O^R\}$ at the end of the symbolic name, for example, the letter "Σ" is denoted by the symbolic name $\oplus[Q^1[L^4](L^3)]\{'M'O^R\}$.

The letter can be often seen as the distortion version of the structural archetype. For example, letters "ל" and "م" given by the symbolic names: $\Theta/\rho[M^1[L^6]]\{L^3,Q^1[M^1[L^5]](M^1[L^3]),M^1[L^3]\}]$ (Fig. 5.45g), and $\Theta/\rho[M^2[L^5]]\{M^1[L^3],Q^1[M^1[L^3]](M^1)\}]$ (Fig. 5.45h) can be seen as the

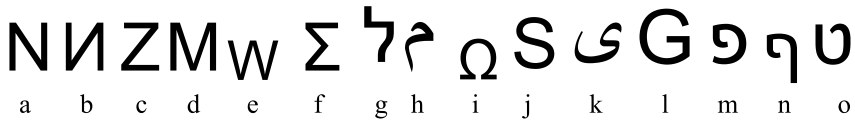


Fig. 5.45. Similar letters

Fig. 5.46. Similar letters

Fig. 5.47. Similar letters

distorted versions of one another. Similarly the specific shape classes (categories) can be derived to match letters shown in Figs. 5.45(i–o), 5.46 and 5.47.

In the previous sections learning of the visual knowledge given in the form of the visual concepts was described. Learning is part of the thinking process aimed at acquiring a new knowledge and restructuring knowledge that was previously learned. The learned visual concept is "attached" to the knowledge schema of the categorical chain. During learning new categories are added. In the first stage of learning of the letters category from the different alphabets the "Arial" as the simplest form of font is selected. Learning starts from learning of the category that is represented by the categorical chain. In the case of the categorical chain given as follows $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{Let}} \rangle \triangleright \langle \nu_{\mathrm{Alp}} \rangle \triangleright \langle \nu_{\mathrm{Lat}} \rangle \triangleright \langle \nu_{\mathrm{LowC}} \rangle \triangleright \langle \nu_{\mathrm{PrF}} \rangle \triangleright \langle \nu_{\mathrm{Ar}} \rangle \triangleright \langle \nu_{a} \rangle$ the learned prototype is described by the categorical chain as follows: the letter "a" ν_{a}, font – Arial ν_{Ar}, printed form $\nu_{\mathrm{Pr\,F}}$, lower case ν_{LowC}, Latin alphabet ν_{Lat}. In the case when all letters specified by categories of this categorical chain are learned, the category such as "lower case" ν_{LowC} can be exchange for the category "upper case" $.. \triangleright \langle \nu_{\mathrm{Lat}} \rangle \triangleright \langle \nu_{\mathrm{UppC}} \rangle \triangleright \langle \nu_{\mathrm{PrF}} \rangle \triangleright \langle \nu_{\mathrm{Ar}} \rangle \triangleright \langle \nu_{a} \rangle$ and process of learning of the letters specified by this categorical chain will start again. The nonvisual knowledge is represented as the knowledge schema of the knowledge chain $\ldots \triangleright \langle \kappa_{\mathrm{Lat}} \rangle \triangleright \langle \kappa_{\mathrm{UppC}} \rangle \triangleright \langle \kappa_{\mathrm{PrF}} \rangle \triangleright \langle \kappa_{\mathrm{Ar}} \rangle \triangleright \langle \kappa_{\mathrm{a}} \rangle \prec \{\partial_{\mathrm{ViC}}, \partial_{\mathrm{Nam}}, \partial_{\mathrm{LIn}}, \partial_{\mathrm{Def}}, \partial_{\mathrm{MGe}}\}$. The knowledge schema for the prototype of the letter "a" includes the visual concept ∂_{ViC}, the name ∂_{Nam}, the linguistic interpretation ∂_{LIn}, the definition ∂_{Def} and the method of generation ∂_{MGe}. During learning process the knowledge schema is "filled" with learned nonvisual knowledge that leads to reorganization of the knowledge in the categorical chains. The consistency of acquired knowledge is checked and new connections among different categorical chains are established.

The letter is properly understood by linking it to the language category. Understanding letter in the context of the language categories make it possible to use the knowledge of the language categories to read and write the text. Text is composed from the category of words which are string of letters. Understanding of the word requires reading a given string of letters and converting it into the word of a given language. As it was described in Chap. 4 the visual object such as a word is a member of the category of words $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pt}} \rangle \triangleright \langle \nu_{\mathrm{Let}} \rangle \triangleright \langle \nu_{\mathrm{Alp}} \rangle \triangleright \langle \nu_{\mathrm{Lat}} \rangle \triangleright \langle \nu_{\mathrm{PrF}} \rangle \triangleright \langle \nu_{\mathrm{Wor}}, \nu_{\mathrm{Phr}}, \nu_{\mathrm{Sen}}, \nu_{\mathrm{Txt}} \rangle$. One of the word categories is the name category. The name category of the visual object refers to the categories such as the category of real-world objects, signs, or figures. For example, the name "capacitor" refers to the visual object given by the categorical chain as follows: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{EngS}} \rangle \triangleright \langle \nu_{\mathrm{ElES}}, \nu_{\mathrm{MeES}}, \nu_{\mathrm{ChES}} \rangle \triangleright \langle \nu_{\mathrm{Rez}}, \nu_{\mathrm{Ind}}, \nu_{\mathrm{Cap}}, \nu_{\mathrm{Dio}}, \nu_{\mathrm{Trz}} \rangle$ where the name "capacitor" refers to the category capacitor ν_{Cap}, derived from the visual engineering symbol. As it was described in Chap. 4 the category of electronic symbols is linked with category of electronic elements given by the following categorical chain $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \sigma_{\mathrm{ReO}} \rangle \triangleright \langle \sigma_{\mathrm{Ear}} \rangle \triangleright \langle \sigma_{\mathrm{NLiv}} \rangle \triangleright \langle \sigma_{\mathrm{MMad}} \rangle \triangleright \langle \sigma_{\mathrm{Asp}} \rangle \triangleright \langle \sigma_{\mathrm{ElAsP}} \rangle \triangleright \langle \nu_{\mathrm{Rez}}, \nu_{\mathrm{Ind}}, \nu_{\mathrm{Trn}}, \nu_{\mathrm{Cap}}, \nu_{\mathrm{Dio}}, \nu_{\mathrm{Trz}} \rangle$. From this categorical chain it is easy to infer that capacitor is an element of the electronic assembly and it is a man-made object. The visual representative of the capacitor can be generated during the imagery process and used in the process of the visual explanation or imagery transformation. Images invoked by words "capacitor" can be given in the form of the schematic or realistic representation. From the category of knowledge we can infers that capacitor is part of the electronic devices and has its symbolic equivalent that is part of the electronic schema.

The visual object such as a word or a sentence are members of the category of words, the category of phrases, the category of sentences, or the category of texts that are derived from the pattern category $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pt}} \rangle \triangleright \langle \nu_{\mathrm{Let}} \rangle \triangleright \langle \nu_{\mathrm{Alp}} \rangle \triangleright \langle \nu_{\mathrm{Lat}} \rangle \triangleright \langle \nu_{\mathrm{PrF}} \rangle \triangleright \langle \nu_{\mathrm{Wor}}, \nu_{\mathrm{Phr}}, \nu_{\mathrm{Sen}}, \nu_{\mathrm{Txt}} \rangle$. To understand text the knowledge from area of linguistic need to be learned and represented in the form of the knowledge objects derived from the category of the linguistic object. The category of linguistic object is divided into the category of lexical analysis, the category of syntactic analysis, the category of semantic analysis, the category of discourse analysis or the category of pragmatic analysis $\kappa_{\mathrm{KB}} \triangleright \langle \kappa_{\mathrm{KOb}} \rangle \triangleright \langle \kappa_{\mathrm{LinO}} \rangle \triangleright \langle \nu_{\mathrm{Lex}}, \nu_{\mathrm{Syn}}, \nu_{\mathrm{Sem}}, \nu_{\mathrm{Dis}}, \nu_{\mathrm{Prag}} \rangle$. The linguistic object supplies the knowledge that is needed to understand text.

During "seeing" of the visual object the first step of the visual understanding is recognition and naming. Finding the name of the visual object makes it possible to give the description of the visual object in the linguistic form. This process is very similar to the process of understanding of the word.

5.5.2.4. Naming and Recognition of Real-World Objects

Naming of the object that is classified as a member of the category of the real-world object is to find the proper categorical level and next find the name of the category at this level. Naming can be regarded as recognition of the object of the known categories. To recognize object there is a need to discriminate among objects of the different categories. The object that is classified to the category of the real-world object can be a member of the structural categories such as a picture, a pattern, or an element. The structural categories, described in Chap. 4, are used to interpret an object in the context of other visual objects. The real-world object in image is often a part of the scene and can be partially occluded by other objects. The image from which a phantom (visual object) is to be extracted can belong to the different perceptual categories such as silhouette, colored object, drawing line, or shaded object.

The phantom that is 2D representative of the 3D real-world object needs to be extracted from a picture and transformed into one of the perceptual categories. Extraction of the visual object from a picture is obtained during the segmentation process. Through the simple segmentation method a picture is divided into two regions: a figure and a background. The result of segmentation depends on the type of background. The simplest case is an object on the background of the uniform colors. Process of extraction of an object from an image depends on the type of background. Figure 5.48 shows images with the different backgrounds from which the object needs to be extracted. In Fig. 5.48a is given an example of a very complex background from which extraction of the object (animal) is difficult. Object

a b c

Fig. 5.48. Examples of pictures from which phantoms are extracted

shown in the image in Fig. 5.48b make extraction of the object more difficult than in the case of image shown in Fig 5.48c. The image in Fig. 5.48c illustrates an object that can be regarded as the isolated object that belongs to the structural category of the element rather than to the category of the picture.

The "segmentation problem" is the problem of dividing up a visual scene (picture) into a number of distinct objects. The segmentation method consists of two stages. In the first stage, surface objects are segmented at discontinuities which can be detected by examining the zero crossing and the extreme values of the surface curvature measure. Then these detected discontinuities are used to segment a complex surface into the simpler meaningful components called patches. Finally, these patches are grouped into meaningful 3D objects, and attributed graphs are generated to describe these objects (Fig. 5.49).

In this book the segmentation is regarded as one of the sensory transformations that transform an image into regions called phantoms (the visual objects). The visual object can be regarded as a member of one of the perceptual categories such as: the silhouette, the line drawing, or the colored object. Naming (understanding) of the visual object perceptual category is

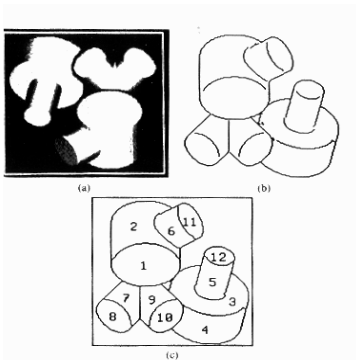


Fig. 5.49. The 2D object segmented into patches can be interpreted as a 3D object

ruled by the specific method of processing and reasoning. The visual object is at first assigned to one of the perceptual categories and next interpreted in terms of the real-world object, a sign, or a figure category. The perceived object that is assigned to the real-world man-made category needs to be learned from many aspects of the object, taken into account the distortion caused by the projection into the plane.

The silhouette is often used during naming process. The naming often called recognition occurs when object is perceived during actions performed by robot (SUS). The robot needs to understand environment to perform task. In the case of naming the object that is perceived by robot is usually transformed into the silhouette. The silhouette can be extracted from existing images (e.g., picture in the book) or can be the result of the visual transformation of the object perceived by the camera. In the first case the perceived object refers to the real-world object that cannot be reached at that moment when perception occurs. In the second case the perceived object can be manipulated by robot (e.g., moved, handled, rotated). This contextual information is used during interpretation of the object.

The silhouette is the perceptual category that shows only part of the visual information about the object. There is, however, category of visual objects such as the leaf category that can be well represented by the silhouette. The silhouette of leaves can be obtained by scanning the leaves by applying the scanner. This method makes it possible to obtain good 2D representation of leaves. Figure 5.50 shows both the scanned leaves and their silhouettes obtained from the scanned form of the leaf. For example, Fig. 5.50a, c, e, g, i shows the scanned leaf whereas Fig. 5.50b, d, f, h, j shows its silhouette.

Naming an object as a leaf means that the leaf is classified to one of the leaf categories. The leaf category is derived from the plant category, so the name of the plant is used to denote the name of the leaf. The typical flowering plants such as a tree consists of different parts such as roots, a trunk, stems, leaves, flowers, fruits and seeds: $..\langle \nu_{\mathrm{Pla}} \rangle \rhd \langle \nu_{\mathrm{Tre}} \rangle \succ [\tau_{\mathrm{Rot}}, \tau_{\mathrm{Trn}}, \tau_{\mathrm{Stm}}, \tau_{\mathrm{Lef}}, \tau_{\mathrm{Flw}}, \tau_{\mathrm{Frt}}, \tau_{\mathrm{Sed}}]$. The category of the tree leaf is divided

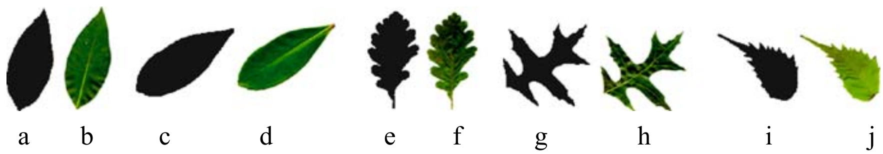


Fig. 5.50. Examples of different leaves (shaded representation) and their silhouettes

into categories such as the oak, the lime, or the poplar $..\langle \nu_{Pla} \rangle \rhd \langle \nu_{Tre} \rangle \succ [\tau_{Lef}] \rhd \langle \nu_{Oak}, \nu_{Pop}, \nu_{Lim}, \nu_{Elm}, \nu_{Fir}, \nu_{AsT}, \nu_{Bee}, \nu_{Bir} \rangle$. The leaf of the specific category such as the Australian oak $\langle \nu_{Oak} \rangle \rhd \langle \nu_{Aus} \rangle$ consists usually of blade, stalk, and venation $..\langle \nu_{Pla} \rangle \rhd \langle \nu_{Tre} \rangle \succ [\tau_{Lef}] \rhd \langle \nu_{Oak} \rangle \rhd \langle \nu_{Aus} \rangle \succ [\tau_{Bla}, \tau_{Stl}, \tau_{Ven}]$. During the naming process the stalk is removed from the leaf. The leaf category is represented by the categorical chain that shows the hierarchical categorical dependence of the leaf category. During learning of the new prototype of the leaf all categories that are part of the categorical chain need to be learned.

Leaves from the different leaf categories can look very differently. These visual differences make their recognition relatively easy. For example, leaves shown in Fig. 5.51 can be easily classified to the different leaf categories. In the case of leaves that are derived from the different oak tree categories the naming process is not so easy. Although leaves derived from the different oak tree categories are different, there is a big variation of the shapes within each category (see Fig. 5.52).

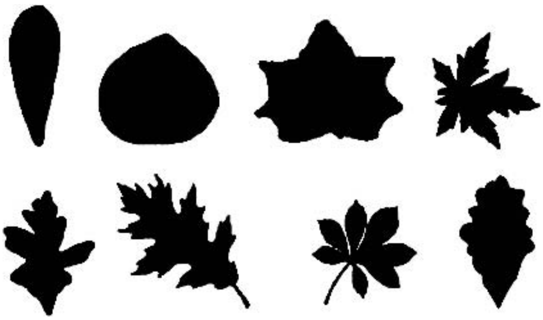


Fig. 5.51. Leaves from different categories that are easy to recognize

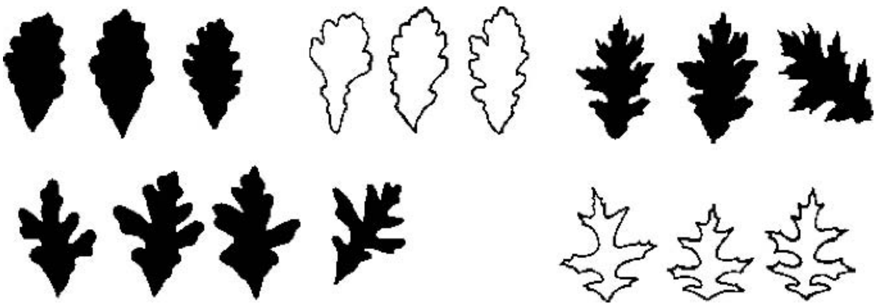


Fig. 5.52. Leaves from different specific categories derived from the oak tree category

As it was mentioned in the previous section, leaves from different plants can be easily distinguished from other leaves if their shapes are different. To name the perceived object SUS needs to learn knowledge needed during the naming process. Naming process involves learning of the visual and nonvisual knowledge. As the result of learning of the visual knowledge the visual concept is obtained. The visual concept consists of the symbolic names that refer to the representatives of the most typical leaves of the learned prototype. For the prototypes of leaves shown in Fig. 5.53 variation of shape within a given shape category is not big. During learning process visual concepts of the different prototypes of leaves are obtained as follows: the prototype 1 – $\varphi = \{K^3\}$ Fig. 5.53a, the prototype 2 – $\varphi^2 = \{K^2\}$ Fig. 5.53b, the prototype 3 as $\varphi^3 = \{K^4\}$ Fig. 5.53c, the prototype 4 – $\varphi^4 = \{Q^1[M^1](M^1)\}$ Fig. 5.53d, the prototype 5 – $\varphi^5 = \{Q^2[M^1[L^3]](2M^1)\}$ Fig. 5.53e, and the prototype 6 – $\varphi^6 = \{Q^3[M^2[L^5]](2M^1, Q^2[L^3](2M^1))\}$ Fig. 5.53f.

For prototypes of leaves shown in this example there is a small diversity of shapes among leaves within a given prototype of the leaf. However, for some leaf prototypes (we call it prototype 7) shape can fluctuate in all ranges of shape shown in Fig. 5.53. The visual concept for this prototype 7 is given as $\varphi^7 = \{K^2, K^3, K^4, Q^1[M^1](M^1), Q^2[M^1[L^3]](2M^1), Q^3[M^2[L^5]](2M^1, Q^2[L^3](2M^1))\}$. Leaves that come from the same plant can differ due to many biological factors that are responsible for the growth of leaves. To recognize leaf (classify it to one of the plant category) there is a need to take into account also other features such as color or venation.

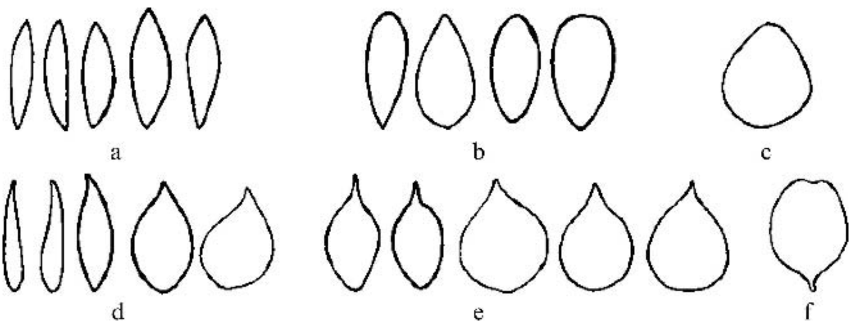
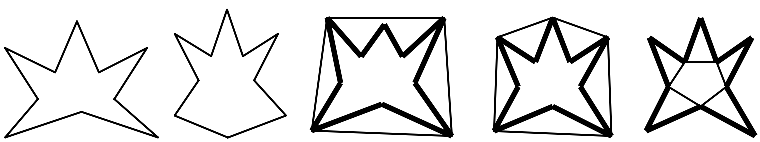


Fig. 5.53. The different categories of leaves

During naming process the leaf is "fitted" into one of the specific shape categories. To have a big range of the shape categories to which perceived object can be "fitted" there is a need to derive the specific a posteriori classes. In the case of the leaf that can be classified as a concave object, it can be approximated by the archetype of the star class. As it was shown in Chap. 2 the object from the star class is represented by the symbolic name $Q^n[L^n](nL^3)$ (Fig. 5.54a). To match the broad range of objects that can be assigned to the leaf category, the new classes, the para-star class and the star-like class are derived from the star class. The archetype of the para-star class (Fig. 5.54b) is given by the symbolic name $Q^n[L^m](nL^3)$, where $m > n$. The archetype of the star-like class (Fig. 5.54c) is given by the symbolic name $Q^n[L^n](kQ^h[L^g](hL^3), mL^3)$, $where\, n = k + m$. Example of the archetype of the star-like class given by symbolic name $Q^4[L^4](3L^3, Q^1_{L^4}(L^3))$ is shown in Fig. 5.54c. For the generic class $Q^4[L^4]$ the following archetypes of the specific star-like classes can be generated $Q^4[L^4](2L^3, 2Q^1_{L^4}(L^3))$, $Q^4[L^4](3L^3, Q^1_{L^4}(L^3))$, or $Q^4[L^4](4Q^1_{L^4}(L^3))$. The star class can have different symbolic representations in the form of the complex class description. The symbolic name reflects the difference in interpretation of the visual object. The description of the concave class is given in terms of the convex generic class and concavities, whereas the description of the complex class divide object into the core object and parts as it shown in Fig. 5.54d, e $Q^5[L^5](5L^3) \equiv \varDelta^5[L^5](5L^3)$. The complex class description is more perceptually oriented. In this description the object from Fig. 5.54c, d has the same symbolic name $\varDelta^5[L^5](5L^3)$, whereas the convex class description gives the different symbolic name $Q^5[L^5](5L^3)$ in the case of object shown in Fig. 5.54d and $Q^4[L^4](Q^1[L^4], 3L^3)$ for the object shown in Fig. 5.54c.



a b c d e

Fig. 5.54. Archetypes from the star class, the para-star class, and the like-star class

The star class only to some extent approximates the leaf shape. The star-leaf class that is introduced to describe the leaves needs to take into account the departure from straightness of the sides and small irregularities of shape. The Fig. 5.55 shows the difference between the objects from the para-star class and the para-star leaf class. The archetype of the para-star class is given as Fig. 5.55a $Q^4[L^5](4L^3)$ whereas the leaf class is given as an object shown in Fig. 5.55b–d. The symbolic name of the object shown in Fig. 5.55b–d is as follows: $Q_{L^{5-7}}(4A)\ A \equiv \{L^3, \widehat{L}^3, Q^2_{L^3}(2\tilde{M}), Q_{L^3}(\tilde{M})\}$.

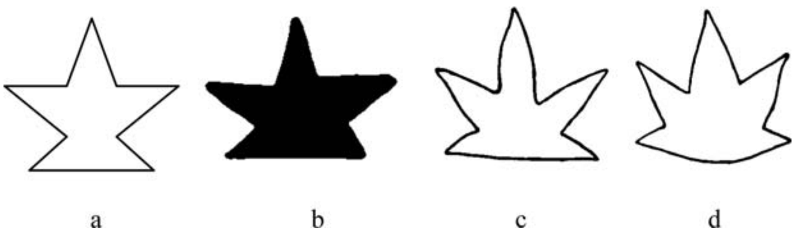


Fig. 5.55. The objects from the para-star class and leaf class

Object from the star class and star-leaf class has the different concavities. The symbolic name of the archetype from the star class (Fig. 5.56a) is as follows: $Q^5[L^5](5L^3)$. All objects, representatives of the star-leaf class (Fig. 5.56b–f) have the symbolic name $Q^5[L^5](5Q)$, where the concavities Q show the significant differences. The Fig 5.56g, h shows the objects from the star-leaf class which is given by the symbolic name $Q^7[L^7](7Q)$. Objects shown in Fig. 5.56 can be classified into two categories Q^5 and Q^7, where both symbols refers to the star-leaf classes $Q^5[L^5](5Q)$ and $Q^7[L^7](7Q)$. The symbolic name not only allows for identification of the object but also gives a very understandable explanation. Object from the category Q^5 differs from the object from the category Q^7 in that respect that the object from the category Q^5 has five residuals whereas the object from the category Q^7 has seven residuals. Objects from both categories have the different generic classes, the category Q^5 the pentagon L^5, and the category Q^7 heptagon L^7.


a b c d e f g h

Fig. 5.56. Objects from the star class and star-leaf class

Many objects are similar and the visual concept of learned categories needs to take into account the similarity among objects. The similarity of objects makes it possible to name the objects in the form of the mushroom-like or s-shaped object. The name "mushroom-like" indicates that named object is similar to typical mushroom. The Fig. 5.57 shows objects that are similar; all objects are concave objects having two concavities. The symbolic names of the similar objects shown in Fig. 5.57 are as follows: $Q^2[M^1[L_T^4]](2 \cdot L_R^3)$ (Fig. 5.57a), $Q^2[M^2[L_T^4]](2 \cdot Q^1[L^3](M^1))$ (Fig. 5.57b), $Q^2[M^1[L^3]](W)$ (Fig.5.57c–f), where $W \equiv \{2Q^2[L^3](2M^1)\} \cup \{Q^2[L^3](2M^1),$ $Q^1[L^3](2M^1)\} \cup \{2Q^1[L^3](M^1)\}$, $Q^2[M^1[L^3]](2Q^2[L^3](2M^1))$ (Fig. 5.57d). Abstraction reveals the similarity of objects at the different levels of details – all objects in Fig. 5.57 are concave object Q^2, all objects came from the same generic class $Q^2[M]$, the objects 5.57c–f come from the same specific generic class $Q^2[M^1[L^3]]$, all objects 5.57c–f can be recognized as an object from the class $Q^2[M^1[L^3]](W)$. The objects in Fig. 5.57a, b that are symmetrical are not likely to be a member of the plant category; symmetry indicates that both objects are members of the man-made category.


a b c d e f

Fig. 5.57. Similarity of the objects of the different categories

The objects from the star-leaf class are very similar to objects from the leaf-irregular class. The leaf-irregular class is described by the symbolic name $Q^n[L^n](a_1Q^1, a_2Q^2, ..., a_nQ^n)$, where a_iQ^i denotes the archetype of the convex or concave class. Members of the leaf-irregular class have irregular concavities. Examples of these objects are shown in Fig. 5.60. The detailed description of these objects, members of the leaf-irregular class, is given by showing the statistical distribution of residuals. Examples of the different residuals are given in the Figs. 5.58 and in 5.59. The symbolic names of these residuals shown in Fig. 5.58 are as follows: (a) $(q)M^1$, (b) $Q^2[\widehat{L}^3](2\tilde{M}^1)$, (c) $Q^2[M^1[L^4]](\tilde{M}^1, M^1)$, (d) $Q^2[\widehat{L}^3](\tilde{M}^1, M^1)$, (e) $Q^3[\widehat{L}^{3-4}](3\tilde{M}^1)$, (f) $Q^3[\widehat{L}^{3-4}](2\tilde{M}^1 M^1)$, (g) $Q^4[M^1[\widehat{L}^{5-6}]](3\tilde{M}^1, M^1)$, where the symbol (q) denotes small residuals, $\widehat{L}^3$ denotes the polygon (triangle) with smooth corners, $\tilde{M}^1$ denotes small curve-polygon residual. Residuals shown in Fig. 5.58 are members of the concave classes at the first level of iteration $Q^n[\Lambda](n\Lambda)$. The term "the first level of iteration" means that all residuals are members of the convex class Λ. Residuals shown in Fig. 5.59 are members of concave classes at the second level of iteration $Q^n[\Lambda](nQ)$. The archetypes of the concave classes at the second level of iteration are perceived as more "irregular." The symbolic names of these residuals are as follows (a) $Q^3[M^1[\widehat{L}^{5-6}]](\tilde{M}^1, M^1, Q^3[L^4](2\tilde{M}^1, M^1))$, (b) $Q^2[M^2[\widehat{L}^5]](Q^2_{\widehat{L}^4}(2\tilde{M}^1), Q^3_{\widehat{L}^5}(2\tilde{M}^1, M^1))$, (c) $Q^3[M^1[L^4]](\tilde{M}^1, M^1, Q^2[\widehat{L}^3](2\widehat{\tilde{M}}^1))$, (d) $Q^3[M^1][\widehat{L}^5](M^1, Q^2[\widehat{L}^3](2\tilde{M}^1), Q^2[\widehat{L}^4](2\tilde{M}^1))$.

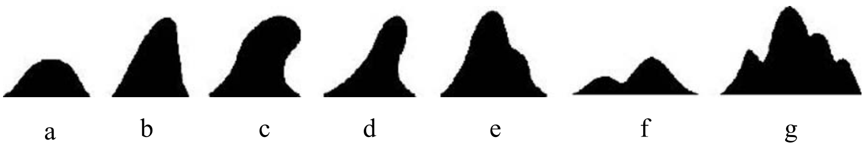


Fig. 5.58. Residuals are members of the concave classes at the first level of iteration

Fig. 5.59. Residuals are members of the concave classes at the second level of iteration

When some categories of leaves have a small diversity of shapes and are easy to identify, there are categories of leaves that have a very big diversification of shapes. For example, sample of leaves, members of the leaf-irregular class are shown in Fig. 5.60. These leaves were randomly picked up so that they preserve variation of shapes that are typical for population of the leaves of that category. These leaves represent two different prototypes P1 and P2. The symbolic description of the leaf type P1 and type P2 is given in the form of the residuals. Residuals have the complex symbolic representation; examples of these residuals are shown in Figs. 5.58 and 5.59. To decrease the number of learning exemplars the abstraction can be applied. The abstraction means that the small residuals will be eliminated (e.g., smoothing contour) from objects that are used to learn the visual concept. The visual concepts of the leaves type P1 and P2 have the same symbolic names. In the case when the visual concept of the different leaves type, e.g., P1 and P2 has the same symbolic name the naming of the unknown object as the leaf of type P1 or P2 can be not unique. In the case when symbolic names consist of the different symbols these symbols are ordered (simple symbol to the very complex). Lets the visual concept of the leaf of type P1 is given as follow $\varphi^{P1} = \{\eta_1, \eta_2, ..., \eta_N\}$, and the visual concept of the leaf of type P2 is given as $\varphi^{P2} = \{\mu_1, \mu_2, ..., \mu_M\}$ where

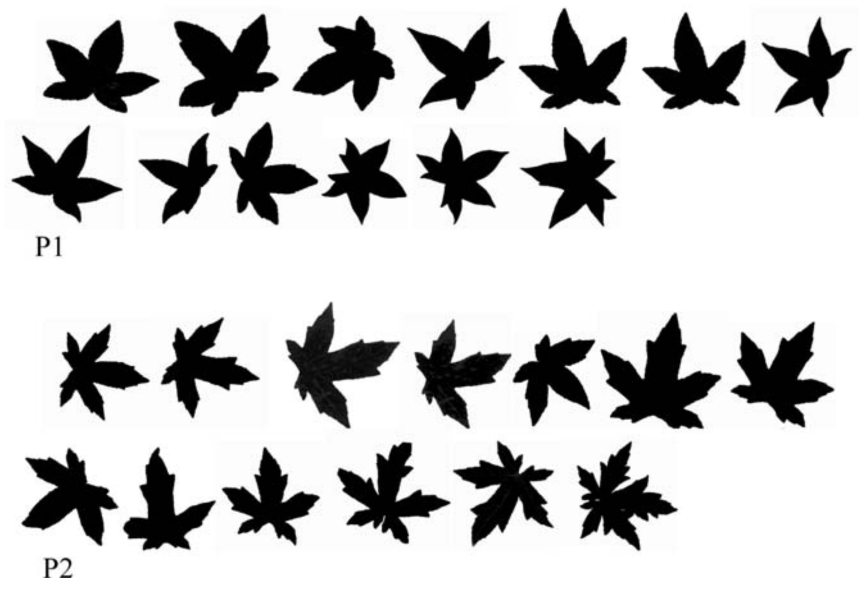


Fig. 5.60. Sample of the leaves from two different types P1 and P2

$\varphi^{P1} \cap \varphi^{P2} = \{\eta_n, \eta_{n+1}, ..., \eta_N\}$ and $\{\eta_n, \eta_{n+1}, ..., \eta_N\} \equiv \{\mu_1, \mu_2, ..., \mu_{M-m}\}$, that means the last *N–n* symbolic names of the visual concept of leaf type P1 and the first *m* symbolic names of the visual concept of leaf type P2 are the same. In that case when an examined object is given by the symbolic name $\{\eta_1, \eta_2, ..., \eta_{n-1}\}$ it is classified as the leaf of the type P1 when an examined object is given by symbolic name $\{\mu_m, \mu_{m+1}, ..., \mu_M\}$ it is classified as the leaf of the type P2, when an examined object is given by symbolic name $\{\eta_n, \eta_{n+1}, ..., \eta_N\} \equiv \{\mu_1, \mu_2, ..., \mu_{M-m}\}$ the contextual information need to be used to select the proper leaf category.

During the learning process there is a need to take into account not only the diversity of the different shapes of leaves but also the perceptual ability of SUS to perceive small parts. The small part is the part of the object that cannot be visible. For example, the small concavity can be interpreted as distortion and the object can be assigned to the convex class. These problems are quite complex and description of these problems is not included in this book.

In the previous section the leaf category, members of which are nearly 2D objects was described. In this section application of the silhouette of objects that are representatives of 3D objects, in the naming process is described. The silhouette as a perceptual category is interpreted as a 3D object by utilization of the contextual information. In shape understanding method all convex or concave symmetrical objects are interpreted as 3D objects as the result of application of the method of body of revolution. Figure 5.61 shows examples of objects that are interpreted as 3D objects. As we can see, only some of them are representative of the category of real-world objects. An object is classified into one of the real-world categories during the naming process.

The silhouette as the 2D visual object (phantom) can be interpreted as a 3D object or a real-world object. For example, the rectangle L_R^4 can be interpreted as a parallelepiped or pyramid that is denoted as

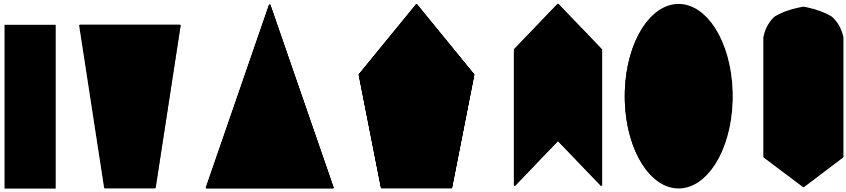
Fig. 5.61. Example of phantoms interpreted as a 3D imagery objects and real-world objects

$L_R -><D>\{R\}\{L_R\}\{pyramid, paralelepiped\}$, where symbol $<D>$ denotes 3D object one aspect of which is given by rectangle L_R^4 and which is given by names $\{pyramid, paralelepiped\}$ The letter {R} denotes the body revolution procedure. The phantom interpreted in terms of 3D geometrical object is regarded as a representation of the possible real-world object. During naming process a phantom (silhouette) shown in Fig. 5.62 is, at first, interpreted as a 3D visual object obtained by the body revolution procedure. Next this 3D visual object can be transformed (rotated) in order to obtain the different visual aspect of the examined object. This visual rotation (imagery transformation) is part of the visual thinking process that can lead to the explanation of the visual features of the object. An object that is interpreted as the object formed by procedure of the body revolution can be fully restored based on the 2D silhouette. The phantoms that are members of symmetrical and elongated classes are interpreted as the 3D objects. These objects are interpreted as objects of the category of 3D geometrical figures or as a category of known real-world objects. The interpretation in terms of 3D geometrical object is shown by indicating that the object is obtained by rotating along the axis of symmetry (solid of revolution). During learning process a 3D representation is obtained by converting the symbolic name η into the 3D symbolic representation: $<D>\{\eta\}$. In the case when interpretation is given in the form of the known real-world objects the name of the objects to which phantom can refers is added at the end, for example, $Q^2[L^4](2 \cdot L^3) -><D>\{R\}\{\{Q^2[L^4](2 \cdot L^3)\} -> \{jug\}$.

The interpretation given in the form of 3D symbolic names for the phantoms shown in Fig. 5.62 are as follows: $L_R^4 -><D>\{R\}\{L_R^4\}$, $L_T^4 -><D>\{R\}\{L_T^4\}$, $L^3 -><D>\{R\}\{L^3\}$, $K^1 -><D>\{R\}\{K^1\}$. The following is an example of the interpretation given in the form of the 3D symbolic name and real-world objects for the phantoms shown in Fig. 5.63:

a. $Q^2[L^4](2 \cdot L^3) -><D>\{R\}\{Q^2[L^4](2 \cdot L^3)\} -> \{jug\}$
b. $Q^2[L^5](2 \cdot L^3) -><D>\{R\}\{Q^2[L^5](2 \cdot L^3)\} -> \{jug\}$
c. $Q^2[L^{4-6}](2 \cdot L^3) -><D>\{R\}\{Q^2[L^{4-6}](2 \cdot L^3)\} -> \{glas, vazon\}$
d. $Q^2[L^6](2 \cdot L^3) -><D>\{R\}\{Q^2[L^6](2 \cdot L^3)\} -> \{vaza\}$
e. $Q^2[L^4](2 \cdot Q^2[L^4](M^1\tilde{M}[K^1])) -><D>\{R\}\{Q^2[L^4](2 \cdot Q^2[L^4]$ $(M^1\tilde{M}[K^1]))\} -> \{vaza\}$

When the 3D interpretation is found SUS can generate the rotated version of the 3D object in the explanatory process.



a b c d e f g

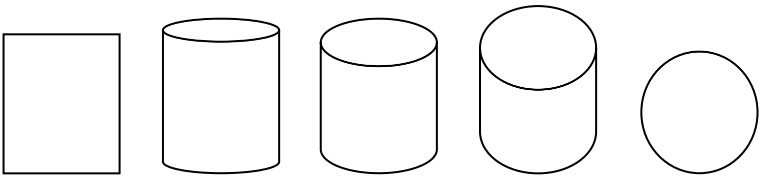
Fig. 5.62. Examples of phantoms interpreted as 3D object

a b c d e

Fig. 5.63. Example of phantoms interpreted as real-world objects

It is assumed that the object that is assigned to the convex or the concave symmetrical class is interpreted as the 3D object. This assumption imposes constraint on the process of selection the most possible representation of the 3D object. At first the 3D interpretation is found and next the interpretational process can be continued based on the knowledge supplied by the knowledge schema. To explain the results of understanding, the object perceived by SUS can be rotated to show the different aspects of the object. Figure 5.64 shows the characteristic visual aspects of the perceived object (rectangle) (Fig. 5.64a) interpreted as a 3D object. Figure 5.64(b–e) shows some possible interpretations generated by SUS.

Interpretation of the phantom in terms of the real-world object can be based on the common aspect of the visual object and can utilize silhouette of the object in the process of learning and understanding. This method gives good results in the case when understanding is reduced into recognition of the known objects (e.g., machine vision). In this interpretation



a b c d e

Fig. 5.64. (b–e) Characteristic visual aspects of the perceived object (**a**) that are generated during explanatory process

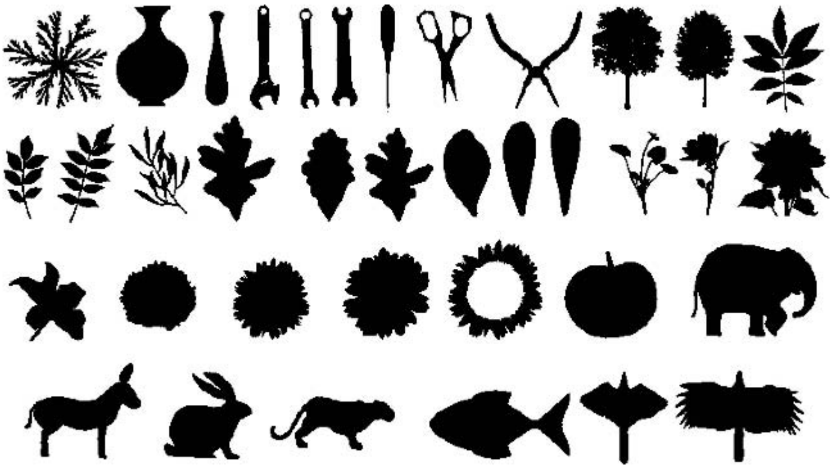


Fig. 5.65. Phantoms (silhouette) interpreted as a real-world object

phantom is treated as the result of the conventional visual transformation rather than one of the geometrical transformations, e.g., projection. The silhouette of the object often represents the most common aspect of the visual object. In the case of the flat object such as a leaf, it shows the natural visual side of the object. In the case of the category of man-made object such as the category of the tools the selected visual aspect of the object shows the tools in its natural position. Figure 5.65 shows phantoms (silhouettes) of different categories of objects.

The selection of the sample of objects that will be used for learning the category of the real-world object is a complex task. Depending on the category of the real-world object that is selected to learn the visual concept, the different perceptual categories can be used. There is dependence among perceptual categories of the visual object. For example, a silhouette can be relatively easy to obtain from the line drawing or colored object. Although there is the lost of information that could be useful in interpretation of the silhouette, analyzing the silhouette is the simplest way of identification of the object. The silhouette can be used at first stage of naming process. The additional visual information can be obtained from the category of line drawing. Figure 5.66 shows examples of silhouettes obtained from the line drawing. As it can be noted some silhouettes supply enough visual information to be identified as an object of a given ontological

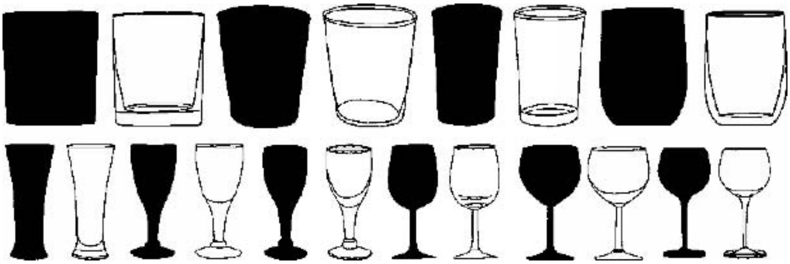


Fig. 5.66. Members of the different perceptual categories of the glass category

category. During learning of the hierarchy of categories of visual objects the most important task is finding the visual object (prototype). At first the most typical visual representation of a given category of the real-world object is learned. In the first stage of learning process the silhouette of the real-world object and objects from the different categories such as figures, signs, letters, and real-world objects are learned. In the second stage the line drawing of the real-world object is learned in the context of the other categories. And finally the shaded form of the real-world object is learned in the context of other categories.

Learning of the visual concept of the real-world object is connected with solving problem of discrimination among similar objects (visual similarities). The similarity of objects can be regarded on two levels: the conceptual level and the visual level (the visual concept). The conceptual similarity of objects came from belonging to the same category, e.g., tools for eating (the fork, the knife). The visual similarity is concerned with similarity of visual objects (phantoms). For example, during learning of the visual concept of the spoon category there is a need to learn the visual concepts of similar objects. Similar objects are objects that share some visual features. Figure 5.67 shows examples of the similar objects. In all cases of objects shown in Fig. 5.67 there is the thin part that is "glued" with the convex part in such a way to become the symmetrical whole. Figure 5.67b–f shows objects that are similar to objects that are representative of the spoon category Fig. 5.67a. The symbolic names of objects shown in Fig. 5.67 are as follows: (a) $Q^2[M^1[\hat{L}^3]](2 \cdot Q^1[L^3](M))$, (b) $Q^2[\hat{L}^5](2 \cdot Q^1[L^3](M))$, (c) $\overleftrightarrow{C}(L^6, L^4_T) \equiv Q^2[L^6](2 \cdot L^3_O)$, (d) $C(L^4_T, L^4_T) \equiv Q^2[L^6](2 \cdot L^3_S)$, (e) $C(L^4_T, L^4_R) \equiv Q^2[L^6](2 \cdot L^3_R)$, (f) $C(L^4_R, \hat{L}^4_R) \equiv Q^2[\hat{L}^5](2 \cdot L^3_R)$. The learned visual concept of the spoon includes the symbolic names of the most similar objects. During naming (recognition) these similar objects can be interpreted as a spoon when the perceived object is not well visible and the contextual information indicates that the perceived object can be the spoon.

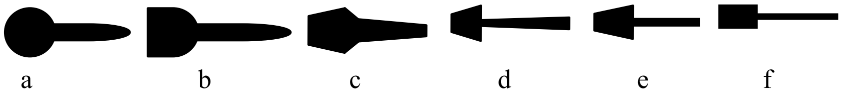


Fig. 5.67. Examples of similar objects

Figure 5.68 shows objects that could be regarded as similar to objects used in learning of the visual concept of the spoon. These phantoms can be seen as representative of the schematic construction (abstraction) where selected features of the real-world object are used. In the case of objects shown in Fig. 5.68f when applying abstraction we can infer that objects in Fig 5.68a–d given by the same symbolic name $C(L_R^4, L_R^4) \equiv Q^2[L^6](2 \cdot L_R^3)$ are similar. The specific classes derived from the class $Q^2[L^6](2 \cdot L_R^3)$ shows the differences between objects shown in Fig. 5.68a–d. The symbolic names of the objects shown in Fig. 5.68a–d are as follows: (a) $C(L_R^4, L_R^4) \equiv Q^2[L^6](2 \cdot L_R^3)$, (b) $C(L_R^4, \overleftrightarrow{L}_R^4) \equiv Q^2[L^6](2 \cdot \overleftrightarrow{L}_R^3)$, (c) $C(L_R^4, \hat{L}_R^4) \equiv Q^2[\hat{L}^5](2 \cdot L_R^3)$, (d) $C(L_R^4, \hat{L}_R^4) \equiv Q^2[\hat{L}^5](2 \cdot \overleftrightarrow{L}_R^3)$). In similar way it can be shown (by applying abstraction) that the object in Fig 5.68b given by the symbolic name $C(L_R^4, \overleftrightarrow{L}_R^4) \equiv Q^2[L^6](2 \cdot \overleftrightarrow{L}_R^3)$ is similar to an object in Fig. 5.68e given by the symbolic name $C(M, M) \equiv Q^2[M^2[L_T^4]](2 \cdot L_R^3)$; these objects are members of the same class $Q^2[\Lambda](2 \cdot L_R^3)$. Similarly, the object in Fig. 5.68e given by the symbolic name $C(M, M) \equiv Q^2[M^2[L_T^4]](2 \cdot L_R^3)$ is similar to the object in Fig. 5.68f given by the symbolic name $Q^2[M^2[L_T^4]](2 \cdot Q^2[L^3](2M^1))$; these objects are members of the same class $Q^2[M^2[L_T^4]](2 \cdot C)$.

Naming real-world object depends on the perceptual category to which perceived object is assigned. As it was described in Chap. 4 the perceptual category of the line drawing is divided into category of: the segmentation edge, the conventional 3D drawing, the intentional geometrical drawing,

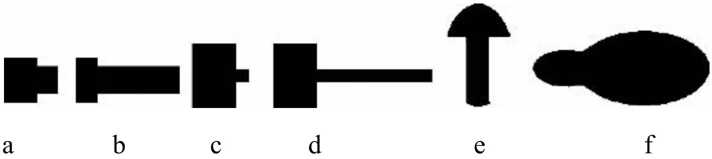


Fig. 5.68. Examples of similar objects (special cases)

the multiview drawing, the view from the top, the frontal view, the orthographic projection drawing, the perspective projection drawing, the folding sheet drawing or the many aspects drawing and given as follows: $v_O \rhd \langle \pi_{Ld} \rangle \rhd \langle \sigma_{El} \rangle \rhd \langle v_{\mathrm{Re}O} \rangle \Rightarrow \| v_{SeE}, v_{InG}, v_{MuV}, v_{ViT}, v_{ViF}, v_{OrP}, v_{PeP}, v_{MaAs}, v_{FoSt} \|$. Classifying a perceived object as a member of the perceptual category such as the orthographic projection make it possible to interpret the line drawing object as the 2D object that is obtained as the result of the orthographic projection of the 3D object.

The object as a member of the perceptual line drawing category can be interpreted as a 3D object or a real-world object. The interpretation of the line drawing was one of the major concerns in computer vision. The line drawing as the 2D images (visible-point perspective projections) was interpreted as the 3D real-world object. According to this approach the object that is limited to be planar-faced solid can be specified by vertices, edges, and polygonal faces. A view of an object becomes a line drawing obtained by perceptively projecting all the visible (or partially visible) edges of the object onto the image plane. This projection is also referred as an aspect of the object. In our approach (shape understanding method) a line drawing object is regarded as an object of the G class. Each archetype of the G class is interpreted as a possible 3D objects. Simplest examples of 3D interpretation of the line drawing object members of the G class are geometrical solids.

The real-world object represented by the perceptual category of the line drawing often refers to the most common visual aspects of the object. During naming (visual understanding) process the object is at first decomposed into the thin or the thin complex object and the convex or the concave core object, and is assigned to the G class. Next, the object is decomposed into the core object and holes and assigned to the cyclic thin class. Based on these decomposition object of the G class is finally interpreted as the 3D object or the real-world object.

For example, object shown in Fig. 5.69a is a member of the perspective projection perceptual category. The object in Fig. 5.69a is decomposed into the thin object $\Theta^{2(2)2}[Q[L^4](L^3)]$ (Fig. 5.69b) and the concave core object $Q[L^6](L^3)$ Fig. 5.69c, and assigned into the $G\{Q[L^6](L^3)\}\{\Theta^{2(2)2}[Q[L^4](L^3)]\}$ class. Next object shown in Fig. 5.69a is decomposed into the concave core object $Q[L^6](L^3)$ and four holes: the one concave $Q[L^5](L^3_R)$ and three another convex L^4_T, and assigned into the $\rho[Q[L\ ](L\ )]\{Q[L\ ](L_R), 3L_T\}$ class. Finally, the object is interpreted as the 3D concave figure or as the shelf – the real-world object. Description of the cyclic thin class is given in Chap. 2. Figure 5.69d shows the same

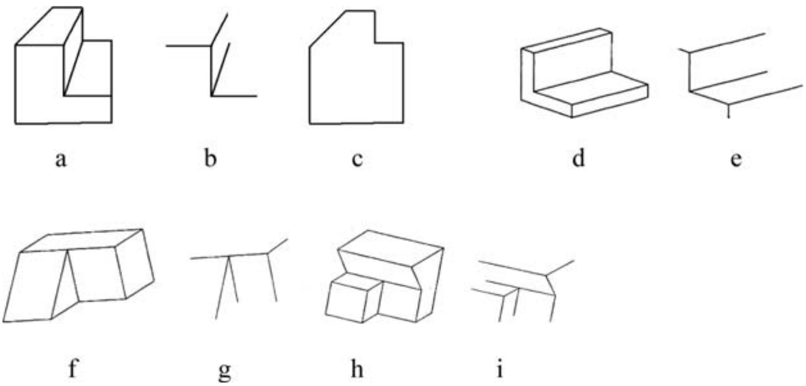


Fig. 5.69. Members of the G class and result of the decomposition into the thin bridge class

object given by the different perceptual category – the orthographic projection. The result of decomposition (the thin object) is shown in Fig. 5.69e. Figure 5.69f, h show another examples that are more complex and the result of decomposition of these objects into the thin object is shown in Fig. 5.69g, i.

Naming object that is assigned to the category of man-made objects such as a machine part utilizes the category of the line drawing. The 3D objects are often interpreted as members of the category of 3D geometrical figures (solids). Many objects, members of the category of the 3D geometrical figures can be also interpreted as members of the category of man-made objects. Also, real-world objects such as simple tools can be interpreted in terms of the 3D geometrical figures called solids. During learning process the learned object of the line drawing category is transformed into the visual representation called the visual schema or the generic visual concept given by the multiview representation. The learned category of the 3D solids has its link with category of the 3D world objects such as the category of minerals that have very similar shapes. These categories are expressed by following categorical chains: $\nu_O \triangleright \langle \pi^M_{Ld} \rangle \triangleright \langle \sigma_{El} \rangle \triangleright \langle \nu_{\mathrm{Fig}} \rangle \triangleright \langle \nu_{3DF} \rangle \triangleright \langle \nu_{\mathrm{Pol}} \rangle \triangleright \langle \nu_{\mathrm{NPol}} \rangle \triangleright \langle \nu_{\mathrm{Cub}}, \nu_{\mathrm{Pri}} \rangle$ and $\nu_O \triangleright \langle \pi^M_{Ld} \rangle \triangleright \langle \sigma_{El} \rangle \triangleright \langle \nu_{ReO} \rangle \triangleright \langle \nu_{\mathrm{Ear}} \rangle \triangleright \langle \nu_{\mathrm{NLiv}} \rangle \triangleright \langle \nu_{\mathrm{NMan}} \rangle \triangleright \langle \nu_{\mathrm{Min}} \rangle$. The multiview representation (described in Chap. 4) is denoted by upper subscript of the symbol that denotes in the perceptual category π^M_{Ld}, where M denotes the multiview drawing. The visual concept of the multiview representation of solids consists of three visual names. For example, the

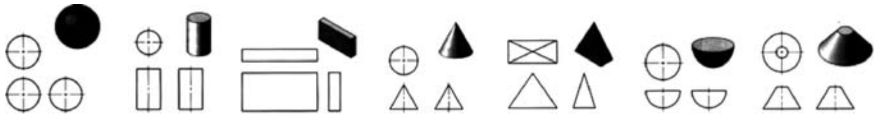


Fig. 5.70. Examples of multiview representation of members of the category of convex 3D figures

multiview representation of the cylinder consists of the symbolic name of the circle and two rectangles $\varphi^{\mathrm{MV}}_{\mathrm{Cylinder}} = \{L^4_R, L^4_R, L^4_R\}$. The top view of the cylinder shown in Fig. 5.70 is represented by a circle whereas the frontal view is represented by a rectangle. The orthographic projection of category of line drawing of the cylinder shown in Fig. 5.70b is given as $\varphi^{\mathrm{OP}}_{\mathrm{Cylinder}} = \{\rho[M^2[L^4_R]]\{K^1_E, Q^1[M^1[L^4_R]](M^1_E)\}\}$.

The multiview representation is used to learn the generic visual concept of the object and can be used in performing mental transformations during thinking process. The generic visual concept (the multiview representation) consists of three symbolic names that refer to each characteristic view. For example, the object shown in Fig 5.70b is represented by the visual concept $\varphi^{MV}_{O1} = \{K^1_C, L^4_R, L^4_R\}$. Examples of the visual concepts of objects shown in Fig. 5.70 are as follows: (a) $\varphi^{MV}_{O1} = \{K^1_C, K^1_C, K^1_C\}$, (b) $\varphi^{\mathrm{MV}}_{\mathrm{Cylinder}} = \{K^1_C, L^4_R, L^4_R\}$, (c) $\varphi^{MV}_{O1} = \{L^4_R, L^4_R, L^4_R\}$, (d) $\varphi^{MV}_{O1} = \{K^1_C, L^3_E, L^3_E\}$, (e) $\varphi^{MV}_{O1} = \{\rho[L^4_R]\{2L^3_A 2L^3_O\}, L^3_E, L^3_E\}$, (f) $\varphi^{MV}_{O1} = \{K^1_C, M^1_C, M^1_C\}$, (g) $\varphi^{MV}_{O1} = \{A^1[K^1_C](K^1_C), L^4_T, L^4_T\}$. Examples of visual concepts of objects shown in Fig. 5.71 are as follows: (a) $\varphi^{MV}_{O1} = \{\rho[L^4_R]\{2L^4_R\}, \rho[L^4_R]\{2L^4_R\}, Q^1[L^5](L^3_R)\}$, (b) $\varphi^{MV}_{O1} = \{\rho[L^4_R]\{L^4_R, Q^1[L^5](L^3_R)\}, \rho[Q^1[L^5](L^3_R)]\{2L^4_{RR}\}, \rho[Q^1[L^5](L^3)]\{2L^4_R\}\}$, (c) $\varphi^{MV}_{O1} = \{\rho[Q[L^5](L^3_R)]\{Q[L^5](L^3_R), L^4_R, Q^2[L^6](2L^3_R)\}, \rho[Q[L^5](L^3_R)]\{L^4_R, Q[L^5](L^3_R), L^4_R, L^4_R\}, \rho[Q[L^5](L^3_R)]\{L^4_R, L^4_s, Q[L^5](L^3_R), L^4_s\}\}$, (d) $\varphi^{MV}_{O1} = \{\rho[L^4_R]\{4 \cdot L^4_R, Q^4[L^8](4L^3_R)\}, \rho[L^4_R]\{2 \cdot L^3_R, L^6\}, \rho[Q^1[L^5](L^3_R)]\{2 \cdot L^4_R\}\}$, (e) $\varphi^{MV}_{O1} = \{\rho[L^5_M]\{2 \cdot L^4_R, L^4_R, L^5_M\}, \rho[L^6_M]\{2 \cdot L^4_R, L^6_M\}, \rho[Q^1[L^5](L^3_R)]\{2 \cdot L^4_R, Q^1[L^5](L^3_R)\}\}$.

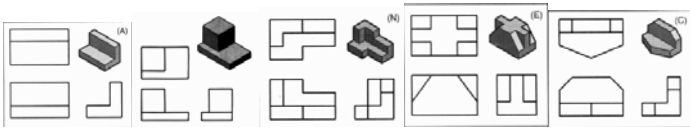


Fig. 5.71. Examples of multiview representation of members of the category of concave 3D figures

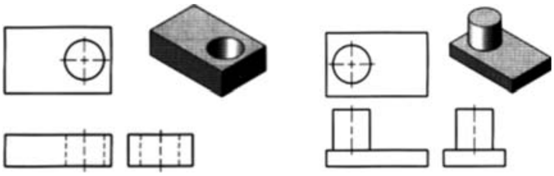


Fig. 5.72. Examples of multiview representation of members of the category of cyclic 3D figures

Examples of visual concepts of objects shown in Fig. 5.72 are as follows: (a) $\varphi_{O1}^{MV} = \{A^1[L_R^4](K_C^1), L_R^4, L_R^4\}$, (b) $\varphi_{O1}^{MV} = \{A^1[L_R^4](K_C^1), Q^2[L^6](2L_R^3), Q^2[L^6](2 \cdot L_R^3)\}$. As we can see based on the generic visual concept there is a possibility to "see" visual object in the context of other objects.

During the learning process of members of the perceptual category of the line drawing, the object is at first transformed into the symbolic name and next visual concept is learned by finding possible 3D visual interpretations. For example, the object shown in Fig. 5.73a represents the orthographic projection of the parallelogram, whereas objects in Fig. 5.73b, c show the projective projection of the parallelogram. The symbolic names for the projections of the parallelogram (Fig. 5.73a–c) are as follows: (a) $\rho[L^6]\{2L_O^4, L_Q^4\}$, (b) $\rho[L^6]\{2L^4, L_Q^4\}$, (c) $\rho[L^6]\{L_T^4, L_R^4\}$. The symbolic names of the multiview representations as well as the symbolic names of the main aspects of the projections such as orthographic or

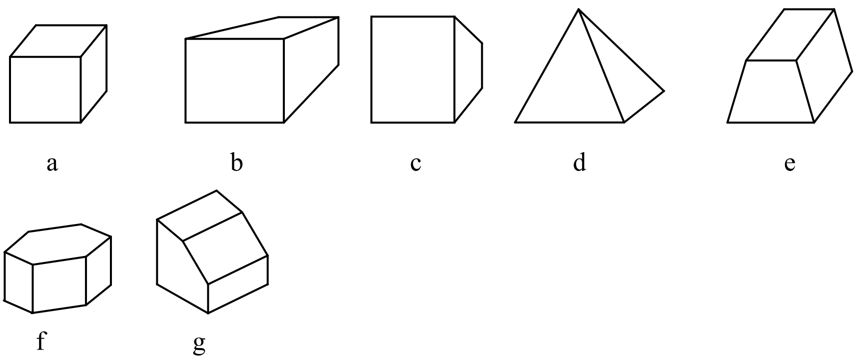


Fig. 5.73. Example of solids used in learning of the visual concepts

projective projection are part of the visual concept of the real-world object. Learning of the visual concept does not assume learning of all spectrums of different shapes of the real-world object. In the first step of the iterative learning the selected 2D representation can be learned. For example, learning of the visual concept of solids shown in Fig. 5.73d–g the following symbolic names: (d) $\rho[L^4]\{L_E^3 L^3\}$, (e) $\rho[L^6]\{2L_O^4 L_T^4\}$, (f) $\rho[L^8]\{3L^4\, L^6\}$, (g) $\rho[L^7]\{3L^4\, L^5\}$ were used.

The perceptual category of the shaded object is most often used 2D visual representation of the real-world object category. Deriving shape information from intensity variation in a single image is a difficult problem. Shape from shading is concerned with the extraction of the shape information from the intensity variation in the image plane. In this research the shading is concerned with the coloring of the different patches of the object. It is assumed that all patches are uniformly colored and there is a significant difference between the different colors of patches. In the case of objects where shading can be interpreted as the patches of the uniform color the interpretation is given in the terms of the objects of the colored class $\aleph$. During the naming process phantom is decomposed into patches and assigned to one of the G or $\aleph$ classes. For example, object shown in Fig. 2.34 is decomposed into three convex objects L^4 and is assigned into class $\rho\{L^6\}\{2L_O^4 L_R^4\}$. The symbolic representation in the form of thin complex G class is next obtained an given as: $G\{L^6\}\{\Theta^3[L^3]\}$. Description of the G class is given in Chap. 2.

The ontological categories of the real-world objects such as category macro, micro, or earth objects categories are learned based on the selected visual representatives of the perceptual or structural categories. Each prototype of the selected category need to be learned based on the larger set of phantoms that represents the different perceptual as well as the different structural categories. The category of living objects need to take into account the difference in shapes that are caused by the movement of parts (e.g., legs) and the different pose (e.g., sitting).

Naming of the minerals often called mineral recognition required not only visual information but also additional information that can be obtained by utilizing the measurement of the special features such as hardness. The mineral recognition requires often the microscopic photographs that belong to the micro category of the real-world object. As it was described in Chap. 4, the category of minerals is derived by following categorical chain $\langle \Pi \rangle \triangleright \langle \sigma_{\text{El}} \rangle \triangleright \langle \nu_{\text{ReO}} \rangle \triangleright \langle \nu_{\text{Ear}} \rangle \triangleright \langle \nu_{\text{NLiv}} \rangle \triangleright \langle \nu_{\text{Man}} \rangle \triangleright \langle \nu_{\text{Min}} \rangle \triangleright \langle \nu_{\text{Nel}}, \nu_{\text{Sul}}, \nu_{\text{Sfo}} \rangle$. Learning knowledge of the category of the nonliving natural objects such as the category of minerals or rocks require also learning the scientific knowledge. Minerals are materials that make up the rocks of the Earth crust. Understanding of minerals is connected with mineral recognition and requires knowledge of the mineralogy that is learned and represented by the knowledge chain given in the form: $\kappa_{\text{KB}} \triangleright \langle \kappa_{\text{KOb}} \rangle \triangleright \langle \kappa_{\text{GeOb}} \rangle \triangleright \langle \kappa_{\text{MinO}} \rangle \triangleright \langle \nu_{\text{NLiv}} \rangle \triangleright \langle \nu_{\text{NMan}} \rangle \triangleright \langle \nu_{\text{Min}}, \nu_{\text{Rock}} \rangle$.

The real-world object from the category of earthy living object is named according to the classification scheme taken from the biological science. The visual concept of each prototype is learned utilizing different perceptual categories. During the first stage of learning selected categories of animals are learned based on the representatives of selected perceptual categories of object. For example, at the first stage of learning of the visual knowledge of the animal category, representatives from the silhouette shown in Fig. 5.65 can be used. The nonvisual knowledge can be learned independently and need to include the detail scientific knowledge as well as more general common knowledge. Learning of the visual concept of the animal requires learning of the visual parts of the animal such as legs, eyes as well as the anatomical parts. The knowledge of the anatomical parts can be seen as the knowledge of the veterinary surgeon or butcher.

Learning of the botanical category such as a plant requires learning of parts of the plant. For example, when learning of the visual category of the three there is a need to learn the parts category of the tree such as leaves, fruits, or flowers. The part category is an auxiliary category that can be derived from any part of the categorical hierarchy. The part schema shows the links to the categories that constitute of the object. For example, the part category of plant such as flowers or fruits is derived from the tree category and is given by following categorical chain: $..\langle \nu_{\text{Pla}} \rangle \triangleright \langle \nu_{\text{Tre}} \rangle \succ [\tau_{\text{Flw}}, \tau_{\text{Frt}}]$. The category of tree leaf indicates that the leaf is part of the tree. The category of tree fruit such as an apple indicates that the apple is a part of tree. The specific category of the apple tree such as “Jonathan” indicates a special category of the apple tree. Each visual category of plants has its knowledge schema that includes other visual features such as color and the nonvisual features such as size or weight.

5.5.2.5. Naming of the Mathematical Objects

The category of mathematical object that is derived from the knowledge object refers to the structural categories such as the element category and pattern category. Naming of the mathematical object refers to mathematical elements such as the named mathematical figure. Naming perceived object as a mathematical object requires interpretation of the object in terms of general mathematical concepts such as a periodic function or a function with discontinuity. Naming refers to the different categorical levels. For example, perceived object can be named on the different categorical levels: the lowest level – the cosines function category; the specific level – the trigonometric function category, or the general level – the periodic function category as shown in the categorical chain: $..\langle v_{\mathrm{Cur}}\rangle \triangleright \langle v_{\mathrm{NCur}}\rangle \triangleright \langle v_{\mathrm{OpCoC}}\rangle \triangleright \langle v_{\mathrm{Per}}\rangle \triangleright \langle v_{\mathrm{Trig}}\rangle \triangleright \langle v_{\mathrm{Cos}}\rangle$. Selection of an appropriate level depends on the contextual information and type of the examined object. For example, naming an examined object as the periodic function whereas it is representative of the cosine function is referring to general level. The reason for naming a perceived object at the general level is that the perceived object was not well visible or the task was formulated in such a way that requires only the general information about the perceived object. Many problems are often defined in the term of mathematical objects and the objects from the category of Cartesian coordinate system are often used to formulate the visual test. The example of the tasks from the mathematical tests is shown in Fig. 5.74. These tasks are formulated in terms of the mathematical curves. The tasks were formulated as follows: "Find the graph that corresponds to a given relationship, e.g., $XY = a$, $Y = X/a$, $Y = XX$," or "Find the graph of a given type of function, e.g., a periodic function with discontinuity."

The naming of the perceived object can be only one of the tasks in solving mathematical problem. Mathematical problem can be formulated as follows find the graph that corresp onds to a given relationship,' e.g.,

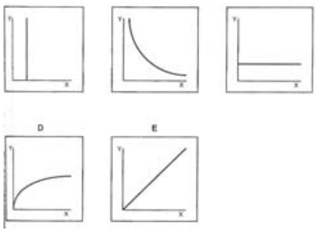


Fig. 5.74. Example of task from the mathematical test

$XY = a$, $Y = X/a$, $Y = XX$ or 'find the graph which can be described by the name: the periodic function, the function with discontinuity', "a function with two maximums and one minimum" "function with an inflection point,' or 'the graph of a monotone increasing function.' Figure 5.75 (a–g) shows examples of curves that can be named at the different categorical levels. For example, a curve in Fig. 5.75e can be named a parabola at the specific categorical level or a convex curve at the general categorical level whereas the curve given Fig. 5.75g can be called the periodic function at the specific categorical level.

Fig. 5.75. Example of the named curves

In mathematics very often mathematical elements (curve, graphs) are placed in the context of the coordinate system. The graph of function is often plotted with additional visual information such as axes, labels on the axes, the text of the different fonts, or frame plot around the plot. These complex objects that are composed of objects of many categories are called mathematical pattern derived from the pattern category. The category of pattern refers to the visual object that is composed from the different element categories.

The object that is classified to one of the categories of the mathematical pattern is understood in the context of the knowledge of this category. For example, the category of the coordinate system supplies knowledge about elements of the coordinate system such as axes and knowledge about interpretation of the graph in the context of information that coordinate system supplies. Application of knowledge makes it possible to locate geometric forms in reference to a reference point. The category of mathematical coordinate systems consists of axes, labels on the axes, the text of the different fonts and frame around the plot. The knowledge supplied by the knowledge schema of the category of the coordinate system make it possible to identify and remove labels, texts, and axes. As it was described in Chap. 4 the element category of the coordinate system consists of many different categories and is given by the categorical chain: $\langle \Pi \rangle \rhd \langle \sigma_{\mathrm{El}} \rangle \rhd \langle \nu_{\mathrm{Sg}} \rangle \rhd \langle \nu_{\mathrm{VSym}} \rangle \rhd \langle \nu_{\mathrm{MCoS}} \rangle \rhd \langle \nu_{\mathrm{Ax}}, \nu_{\mathrm{Lab}}, \nu_{\mathrm{Fra}}, \nu_{\mathrm{Mar}}, \nu_{\mathrm{Gra}} \rangle$, whereas the categorical chain of the pattern of the coordinate system is as follows: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pat}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{MCoS}} \rangle \triangleright \langle \nu_{\mathrm{Car}}, \nu_{\mathrm{Pol}} \rangle$. The knowledge chain of the element category of the coordinate system supplies knowledge about any particularly element and is given as follows: $\ldots\langle \kappa_{\mathrm{MO}} \rangle \triangleright \langle \kappa_{\mathrm{MCoS}} \rangle \triangleright \langle \kappa_{\mathrm{Ax}}, \kappa_{\mathrm{Lab}}, \kappa_{\mathrm{Fra}}, \kappa_{\mathrm{Mar}}, \kappa_{\mathrm{Gra}} \rangle \prec \{\partial_{\mathrm{ViC}}, \partial_{\mathrm{Am}}, \partial_{\mathrm{Ef}}, \partial_{\mathrm{Int}}\}$. This categorical chain supplies knowledge about interpretation of the different coordinate systems such as Cartesian or polar. The categorical chain of the pattern category of the coordinate system supplies the knowledge about the visual aspect of the perceived object. The nonvisual knowledge that is needed during interpretation of the mathematical aspect of the examined object is supplied by knowledge object given as follows: $\ldots\langle \kappa_{\mathrm{MO}} \rangle \triangleright \langle \kappa_{\mathrm{MCoSP}} \rangle \triangleright \langle \kappa_{\mathrm{Car}}, \kappa_{\mathrm{Pol}}, . \rangle \prec \{\partial_{\mathrm{ViC}}, \partial_{\mathrm{Am}}, \partial_{\mathrm{Ef}}, \partial_{\mathrm{Int}}\}$. The knowledge about interpretation of the mathematical object is learned during the learning process. During (naming) understanding process an examined object needs to be interpreted as the one of the category of the coordinate system. In the first stage of understanding process the appropriate processing methods are applied to find proper interpretation of all parts of the visual object and to find the type of the coordinate system. For example, the object in Fig. 5.76 is interpreted as an object from the category type C. The object from this category consists of two axis, two labels (X, Y) and the mathematical figure. During processing stage, at first two labels are identified and removed, next two axes are identified and removed, and next the mathematical figure is identified. The further interpretation is based on the knowledge that is obtained from the knowledge schema.

A member of the category of mathematical objects such as a differential equation is often used as a model of the dynamical systems. The dynamical system refers to the process category and is represented by the perceptual category of animation. An example of this type of the differential equations is the Duffing's equation [8]. For this equation, in a suitable 2D

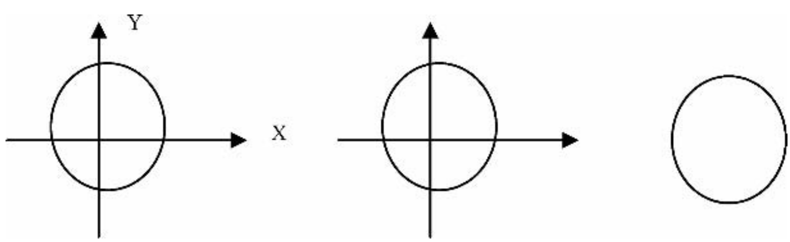


Fig. 5.76. Examples of decomposition object from category type C

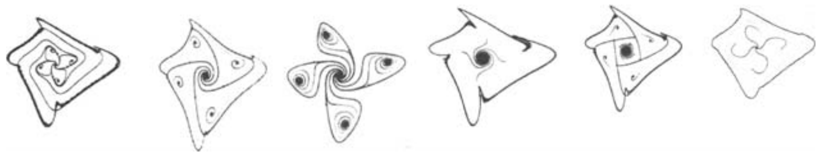


Fig. 5.77. Example evolution of the dynamical system that is represented by members of animation category

parameter plane, characteristic points can be formulated as double (singular) points of plane curve defined by a bifurcation equation. The visual inspection of the surfaces of section can be used for tracing behavior of the dynamical system. Figure 5.77 shows an example of the evolution of the dynamical system. These shape changes are interpreted in terms of the different characteristic regions that can indicate the different behavior of the system.

5.5.2.6. Identification of Statistical Visual Objects

Similarly as the category of mathematical object, the category of statistical object can be regarded as one of the structural categories such as the element category or the pattern category. Naming of the perceived object as a member of the statistical object refers to the category of statistical objects and the statistical knowledge. Statistics makes extensive use of various types of graphs and naming these objects require identifying it as a pattern category or element category.

Examples of members of the category of statistical elements are shown in Fig. 5.78. Naming of these objects is connected with solving problem of data interpretation. Objects shown in Fig. 5.78 represent the following categories: the box plots category, the pie chart category, and the bar graphs category. Understanding of these statistical objects is the task that involves knowledge of the interpretation of the statistical data.

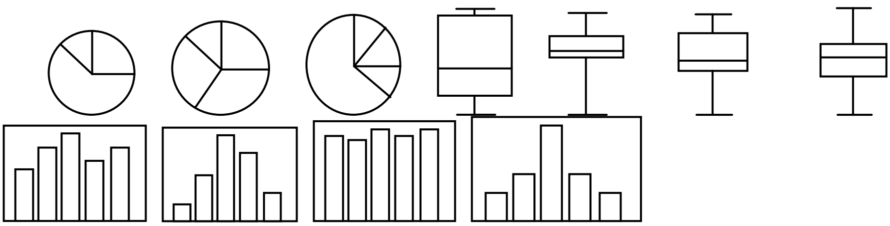


Fig. 5.78. Examples of members of the category of statistical elements

5.5.2.6.1. Data Analysis

Problem solving in statistical data analysis is concerned with interpretation of data in terms of the category of statistical object. The visual object that is to be interpreted is obtained by applying one of the data visualization procedures to the rough data. In data analysis the visualization of data is often the first step in analysis of data. One of the most often visualization methods used in statistic are the scatter plot or the multidimensional representation. Data given in the form of a set of points $\prod = \{(x_1, y_1),...,(x_i, y_i),...(x_N, y_N)\}$ is called a set of 2D dot patterns (a scatter plot). During visual data analysis these points are transformed into 2D visual object (phantom) (Fig. 5.79).

The statistical categories are learned during learning stage and after learning they are used during naming process. Learning statistical categories involves learning of the visual knowledge in the form of the categorical chains of statistical objects and the categorical chains of the knowledge objects. The category of the scatter plot data is derived from the category of the visual object and is divided into the category of the regression analysis, the category of cluster analysis or the category of the discriminant analysis $..\langle v_{\mathrm{StCoS}} \rangle \triangleright \langle v_{\mathrm{SPlo}} \rangle \triangleright \langle v_{\mathrm{Reg}}, v_{\mathrm{Clu}} \rangle$. The category of the regression analysis is divided into the category of linear regression, the category of nonlinear regression or the category of weighted least-square regression $..\langle v_{\mathrm{StCoS}} \rangle \triangleright \langle v_{\mathrm{SPlo}} \rangle \triangleright \langle v_{\mathrm{Reg}} \rangle \triangleright \langle v_{\mathrm{LiR}}, v_{\mathrm{NLR}}, v_{\mathrm{WSR}} \rangle$.

Learning prototypes of the category of the linear regression requires generating the most typical representatives of the category of the scatter plot data. The data that are generated are used to obtain the visual objects from which visual concepts of visual categories are learned. For each specific

Fig. 5.79. Data transformed into phantom by applying the perceptual transformation

category derived from the category of the regression analysis the visual concept is learned. The nonvisual knowledge concerning the statistical method of data analysis as well as interpretation of the results is stored in the knowledge schema of the knowledge object. The visual concept was obtained during the learning process assuming that a number of data points is big enough to form near homogenous distribution of the points inside the pattern. The visual concept is used to extract the inferential rules in the following form: $[\eta \in \varphi^{\alpha}] \Rightarrow a \triangleright \vartheta^{i}$, where η is the symbolic name of the perceived object obtained during the reasoning process, φ^{α} is the shape category (the visual concept) obtained during the learning process, and $a \triangleright \vartheta^{i}$ denotes the action that needs to be undertaken. Example of rules obtained during the learning process: (a) $[\eta \in \varphi^{T}] \Rightarrow a \triangleright \vartheta^{T}$, (b) $[\eta \in \varphi^{L}] \Rightarrow a \triangleright \vartheta^{L}$, (c) $[\eta \in \varphi^{N}] \Rightarrow a \triangleright \vartheta^{N}$. The visual concept $\varphi^{T} = \{C(\Lambda, \Theta)\}$ denotes the whiskers class and ϑ^{T} denotes action "transform data." The visual concept $\varphi^{L} = \{\Lambda^{E}\}$ denotes the convex class and ϑ^{L} denotes action "linear regression model." The visual concept $\varphi^{N} = \{Q\}$ denotes the concave class and ϑ^{N} denotes action "nonlinear regression model." In the case when there is a need for a specific regression the rule are applied $[\eta \in \varphi^{W}] \Rightarrow a \triangleright \vartheta^{W}$, where visual concept $\varphi^{W} = \{\Lambda^{W}\}$ denotes the trapeze-like class and ϑ^{W} denotes action "weighted least-square."

During the data analysis, the data transformed into a phantom are interpreted as one of the statistical data categories and the visual inference is used to select appropriate statistical procedures for further processing.

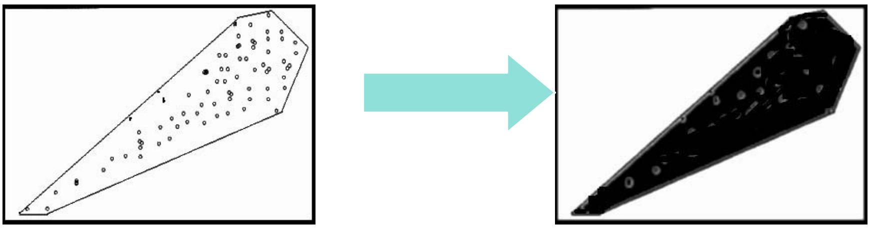


Fig. 5.80. Example of reasoning process that assigns the dot pattern (a) into the category of weighted least-square regression

Fig. 5.81. Application of the perceptual transformation

As an example of applying the visual inference in understanding the data, a regression analysis is given. The regression analysis consists of the following stages: the data analysis, a model selection, fitting data to a model, and interpretation.

The regression analysis that utilizes the visual inference is performed in the following steps:

1. The data $\prod = \{(x_1, y_1),...,(x_i, y_i),...(x_N, y_N)\}$, are transformed into the phantom u (see Fig. 5.81).

2. The phantom u is transformed into a set of critical points $\coprod$ by the sensory transformation $\Im(u) \Rightarrow \coprod$.

3. A set of critical points $\coprod$ is transformed into the symbolic name η during reasoning process $R\langle\coprod\rangle \Rightarrow \eta$.

4. The visual inference $[\eta \in \varphi^T] \Rightarrow a \triangleright \vartheta^T$, where η is the symbolic name obtained in the reasoning process, is applied.

5. In the case of data given in this example the visual concept $\varphi^T = \{C(\Lambda, \Theta)\}$ denotes the whiskers class (whiskers category) and ϑ^T denotes the action "transform the data [$y = \log(Y)$ $x = \log(X)$]" (see Fig. 5.82).

6. Steps 1, 2, 3 are repeated.

7. The visual inference $[\eta \in \varphi^L] \Rightarrow a \triangleright \vartheta^L$ is applied. The symbol η denotes the symbolic name of the perceived object obtained during the reasoning process, whereas $\varphi^L = \{\Lambda^E\}$ denotes the visual concept that consist of one symbolic name (convex class) and ϑ^L denotes the action "apply a linear regression model."

Another category that is derived from category of statistical objects is the category of cluster analysis. Clustering algorithms are effective tools for exploring the structure of the complex data set. Most of the clustering

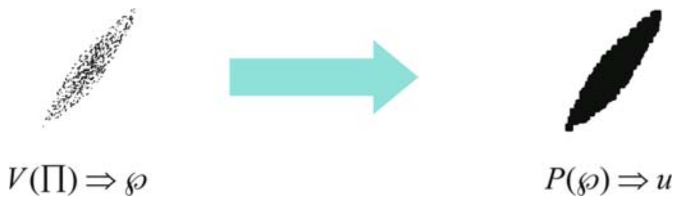


Fig. 5.82. Application of the visual and perceptual transformations

algorithms perform well in the case when data makes regular clusters. In [9] authors propose a new clustering algorithm to cluster data with arbitrary shapes. In comparison to algorithms presented in [9] our approach is based on "seeing" pattern instead of computation of some statistics. During learning of cluster category only cases (samples of the data that generate pattern) for which the number of data points was big enough to form near homogenous distribution of points inside the pattern, were selected. In many real-world statistical problems the number of points is often not very big and further research is needed to deal with the problem when the number of data points does not create a homogenous pattern. For the purpose of this research data were generated from the known statistical distributions (e.g., the normal distribution). Using the synthetic data does not limit the usefulness of the presented method. The real-world data, where the number of data points is big, have the same statistical properties as data generated from the statistical distributions.

Members of the category of cluster analysis were learned by generating points from the known distributions, or by using clusters that were reported in literature as a difficult to approach by a classical cluster analysis method. Examples of the data used in cluster analysis are shown in Fig. 5.83. The categories that are learned make it possible to identify clusters based on their visual properties. During naming process each cluster was identified as a member of one of the shape classes. Clusters shown in Fig. 5.83 were easily identified as members of the known shape categories.

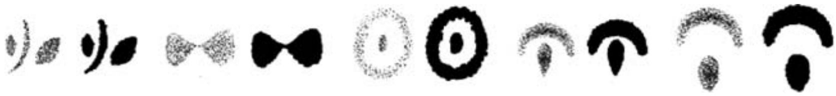


Fig. 5.83. Examples of the data used in the learning of the category of cluster analysis

5.5.2.7. Category of Physical and Engineering Models

Naming of the perceived object as an object of the category of physical model requires understanding it in the context of category of the real-world object to which the category of physical model refers. A model represents only selected features of the real-world object and the visual representation of the model is usually given in the form of the schematic representation.

The object that is a member of the category of the physical model needs to be understood in the context of the real-world object to which this physical model refers. The interpretation is based on the knowledge schema that is supplied by the knowledge object. One of the important problems in understanding of the visual object is to interpret it in the context of the known category of the physical model. The first step of understanding of the real-world object or phenomena (e.g., pendulum) is to recognize it as a member of a given category of real-world object (pendulum) and to interpret it in the context of the category of the physical model. The knowledge schema of the category of the real-world object (pendulum) contains the link to the category of the visual physical model of the pendulum. The category of physical model (pendulum) supplies the schematic visual representation of the real-world phenomena – pendulum. The knowledge schema supplies the knowledge connected with the mathematical model (equation) as well as the interpretation of the mathematical results. Perceiving pendulum refers to the pendulum category that is derived from the category of real-world objects and at the same time refers to pendulum category derived from the category of physical model. This categorical link makes it possible to “think” about the real-world object in terms of the model of this object.

The category of physical visual models consists of predefined visual symbols. The element category of the category of physical model includes the schematic representation of the real-world object or phenomena (physical object), letters, words, arrows, lines, arcs $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{PhMo}} \rangle \triangleright \langle \nu_{\mathrm{PhOb}}, \nu_{\mathrm{Lab}}, \nu_{\mathrm{Ax}}, \nu_{\mathrm{Mar}}, \nu_{\mathrm{Gra}} \rangle$. The pattern category of the category of physical visual models composes the different elements of this category into one of the visual objects. The category of the physical visual model is divided into the category of lever, wedge, axle, pulley, inclined plane, or pendulum: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pt}} \rangle \triangleright .. \triangleright \langle \nu_{\mathrm{PhMo}} \rangle \triangleright \langle \nu_{\mathrm{Lev}}, \nu_{\mathrm{Wed}}, \nu_{\mathrm{Pul}}, \nu_{\mathrm{Rol}}, \nu_{\mathrm{Pen}}, \nu_{\mathrm{InPl}} \rangle$.

The visual model is often used to understand the functionality of the visual phenomena. In that case understanding is focused on the problem "how it works" and the schematic form of the visual model shows only this aspect of the object (machine) that carries out information that is related to its functionality. The schematic form of the visual model is often given as a task when testing the understanding abilities of the student. For example, a set of rollers or rollers connected by the drive belt are tasks that are often used in these tests. The tasks given to the students is about designing a set of rollers that turn in a particularly way. A set of rollers can be made by placing rollers in contact. One roller (shown in grey) is the driving roller and it makes the other to run. The task is formulated as follows: "which rollers will turn in the same direction as the drive roller, and which will turn in the opposite direction?." Another task is formulated as follows: "Draw a drive belt around the set of rollers (shown in Fig. 5.84) so that all the large rollers turn clockwise and all the small rollers turn anti-clockwise. The belt must not cross over itself."

During naming and interpretation of the object that is a member of the category of physical visual models all auxiliary objects are interpreted and removed. Figure 5.85 shows interpretational steps in understanding of the model of uniform beam that carries a concentrated load.

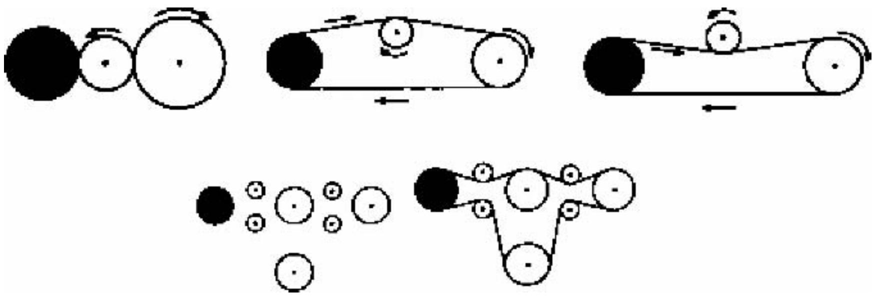


Fig. 5.84. Example of task to draw a drive rollers connected by the drive belt

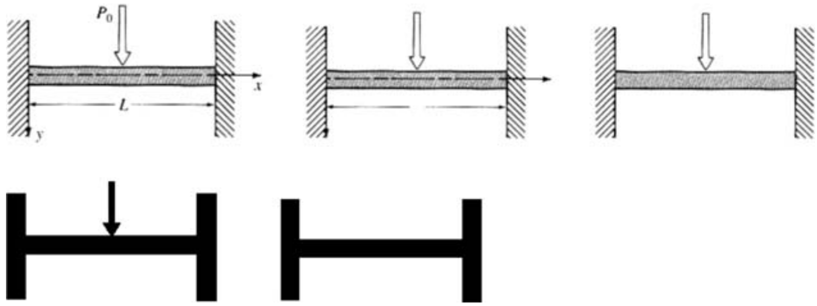


Fig. 5.85. Steps in interpretation of the visual model

5.5.2.8. Visual Resemblance: Visual Analogy

In order to solve the problem there is a need to apply reasoning. There are many different forms of reasoning and short review of this topic was given in Chap. 1. Reasoning that is based on analogy is called analogical reasoning. The term analogy is referring to the type the of reasoning that is described in psychometric tests in the following way: "X is to Y as Z is to Q." The similarity relation that is found between object X and Y and is used to find the object Q that is similar to object Z. In general problem solving approach the operational definition of the analogical reasoning is formulated [10]. Reasoning based on analogy is often used in problem solving.

The visual analogical reasoning is based on the similarity relation of the visual objects. The similarity, both visual and conceptual, is used to find a conceptual grouping at a set of objects shown in Fig. 5.86. Example of the results of performing the task "select objects similar to the cipher 9" are shown in Fig. 5.87. In the first row, objects that were selected based on visual similarities among objects are shown. Objects shown in the second row are objects that were selected based on the conceptual similarity of members of the specific category derived from the cipher category. Objects shown in the third row are objects that were selected based on the conceptual similarity of members of the specific categories derived from the category of cart symbols. Objects shown in the forth row are objects that were selected based on the conceptual similarity of members of the specific categories derived from the category of mathematical symbols.

The main difficulty (for SUS) in selection of the similar objects based on the visual similarities among objects was to discriminate among the very similar visual objects. To find visually similar objects at first the

Fig. 5.86. Examples of visual objects used in experiment

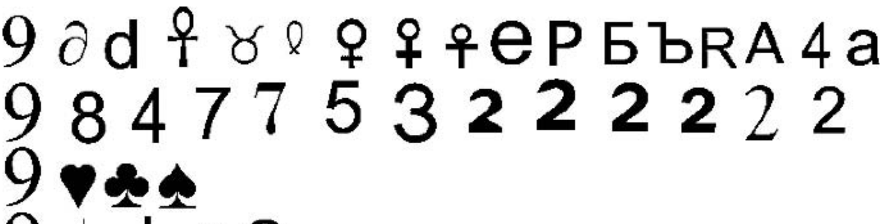


Fig. 5.87. Results of selection of the objects similar to the cipher 9

symbolic name for each object is obtained and next during generalization process objects that have the same symbolic name are selected. For example, objects shown in the first row Fig. 5.86 (the first group) are members of the cyclic class $A^1[Q](K\ )$, whereas all objects in the first row in Fig. 5.86 are members of the cyclic class $A^1[Q]$.

The conceptual similarities are found during categorical reasoning. The categorical reasoning is performed by moving through the categorical chain $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{El}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{Mth}}, \nu_{\mathrm{Mus}}, \nu_{\mathrm{EnSym}}, \nu_{\mathrm{CarSym}} \rangle$. Moving one level up indicates selection of following visual symbol categories: mathematical, musical, currency, cart, or engineering symbols: $\langle \nu_{\mathrm{Mth}}, \nu_{\mathrm{Mus}}, \nu_{\mathrm{EnSym}}, \nu_{\mathrm{CarSym}} \rangle$. Next, moving further up the category of signs can be selected.

The visual analogical reasoning is based on the relations discovered among group of objects. These relations are often expressed in the form "X is to Y as Z is to? where X, Y, Z are given objects whereas symbol "? denotes object that is to be selected from the set of objects given as the possible "answer." The analogical reasoning can be based on the visual similarities of objects or conceptual similarities which have been described in the previous example. In the case of the visual similarities of objects the similarity relation is based on the selected visual features of the objects. In the case of the conceptual similarities the visual analogical reasoning is based on the conceptual similarities of the object categories. Examples of tasks that need to use visual analogical reasoning based on the conceptual similarities of the categories, in order to solve these tasks, are shown in Fig. 5.88.

In the first example (Fig. 5.88a, b) the visual analogical reasoning is based on the perceptual category of line drawing that utilizes spatial relations 2D–3D. A sqare, in Fig. 5.88a, is one of the visible aspects of the cube so the circle is a visible aspect of the ball. As the result of the visual analogical reasoning the ball is selected from other "answers." Using SUS notation these relations can be written as: $L_R^4 \leftrightarrow (3D)L_R^4 \equiv K_C^1 \leftrightarrow (3D)K_C^1$, where $(3D)L_R^4$, $(3D)K_C^1$ denotes symbolic name of the thin G class. A triangle, in Fig. 5.88b, is one of the visible aspects of the prism so the sqare is a visible aspect of the cube. As the result of the visual analogical reasoning the cube is selected from other "answers." Using SUS notation these relations can be written as: $L^3 \leftrightarrow (3D)L^3 \equiv L_R^4 \leftrightarrow (3D)L_R^4$.

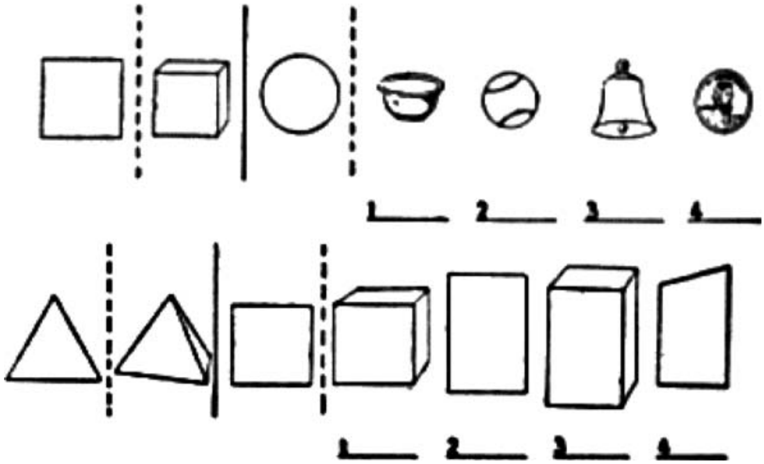


Fig. 5.88. Examples of task where analogical reasoning is used to solve this task

Figure 5.89 shows visual objects that represent relations between the different categories of the real-world object. The task is formulated in the terms of the visual objects. Two objects are given to find the similarity relation, and next based on this similarity relation one of the four objects that represents possible solution is selected. The selected object has to be similar to the third object that is given as part of the similarity relation. During finding of the solution the visual objects are assigned to the proper categories and next the similarity relations among objects are established.

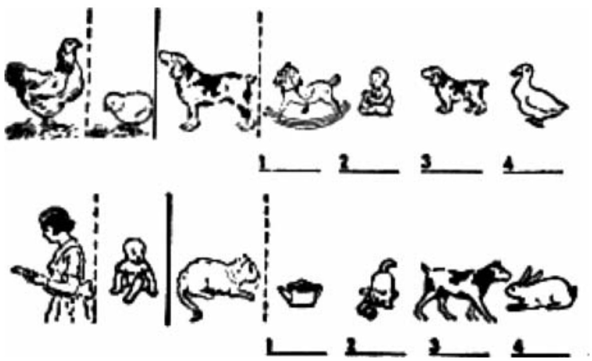


Fig. 5.89. Examples of task where analogical reasoning is used to solve this task

For example, the task given in Fig. 5.89 refers to the category of animals that is given by the following categorical chain: $\langle \Pi \rangle \triangleright \langle \sigma_{El} \rangle \triangleright \langle \nu_{ReO} \rangle \triangleright \langle \nu_{Ear} \rangle \triangleright \langle \nu_{Liv} \rangle \triangleright \langle \nu_{Ani} \rangle \triangleright \langle \nu_{Cho} \rangle \triangleright \langle \nu_{Fis}, \nu_{Amf}, \nu_{Rep}, \nu_{Ave}, \nu_{Mam} \rangle$. The category of animal is divided into the category of the young animals and the category of adults. The visual appearance of the young animal is different than the visual appearance of the adult animal. The analogical reasoning refers to the two categorical chains. The first one is used to establish the similarity relation between two categories. When two objects are assigned to the proper categories the similarity relation is found by "cutting the categorical chain" where these two categories differ. In the second stage the proper visual object is selected based on the relation that was found. For example, in the task shown in Fig. 5.89a the similarity relation is obtained by cutting the categorical chain $.. \triangleright \langle \nu_{Ani} \rangle \triangleright \langle \nu_{Cho} \rangle \triangleright \langle \nu_{Ave} \rangle \triangleright \langle \nu_{Hen} \rangle \triangleright \langle \nu_{hen}, \nu_{chc} \rangle$ moving to the right of the $\langle \nu_{Hen} \rangle$ category. As the result the two categories are obtained $\langle \nu_{hen}, \nu_{chc} \rangle$ (the hen and the chicken). The relation between two objects is such that chicken is a young animal and hen is an adult animal. The next object that is found is the dog $.. \triangleright \langle \nu_{Ani} \rangle \triangleright \langle \nu_{Cho} \rangle \triangleright \langle \nu_{Mam} \rangle \triangleright \langle \nu_{Dog} \rangle, \langle \nu_{dog}, \nu_{sdo} \rangle$. Cutting this categorical chain by moving to the right of the $\langle \nu_{Dog} \rangle$ category the two categories are obtained $\langle \nu_{dog}, \nu_{sdo} \rangle$. These two categories represent the same relation young-adult as the relation that is found between two first objects. The relation that is found based on the categorical chain is the same for two pairs of objects $\langle \nu_{hen}, \nu_{chc} \rangle \propto \langle \nu_{dog}, \nu_{sdo} \rangle$ so the selected object is the small dog. Similarly the task given in the Fig. 5.89b can be solved.

Solution of tasks shown in Fig. 5.90 is based on relations such as wearing and a part of body or tools and material, and is more difficult to find. The category of wearing is divided into the category of the part of the human body such as the head, hands, legs, or arms: $.. \triangleright \langle \nu_{MMad} \rangle \triangleright \langle \nu_{Wer} \rangle \triangleright \langle \nu_{OldW}, \nu_{MidW}, \nu_{NewW}, \nu_{TodW} \rangle \triangleright \langle \nu_{Hed}, \nu_{Hind}, \nu_{Leg}, \nu_{Arm} \rangle$. The specific wearing categories that are derived from the part of the human body shown in Fig. 5.90a are as follows: for the head category: $.. \triangleright \langle \nu_{Hed} \rangle \triangleright \langle \nu_{Hat}, \nu_{Cap} \rangle$, and for the leg category: $.. \triangleright \langle \nu_{Leg} \rangle \triangleright \langle \nu_{Sho}, \nu_{Soc} \rangle$. Based on the categorical chains the relation is found for the first two objects and is used to find the fourth object. The first object in Fig. 5.90a is a member of the hat category and the second object is a member of the head category. Based on the categorical chain the relation that is found by cutting the categorical chain is as follows: $.. \rhd \langle \nu_{\text{Wer}} \rangle \rhd \langle \nu_{\text{TodW}} \rangle \rhd \langle \nu_{\text{Hed}} \rangle \rhd \langle \nu_{\text{Hat}} \rangle$. Similarly for the third object that is a member of the leg category the fourth object is selected from the categories obtained by cutting the categorical chain as follows $.. \rhd \langle \nu_{\text{Wer}} \rangle \rhd \langle \nu_{\text{TodW}} \rangle \rhd \langle \nu_{\text{Leg}} \rangle \rhd \langle \nu_{\text{Sho}} \rangle$. The forth object that fulfills the relation described by the part of the categorical chain $\langle \nu_{\text{Hed}} \rangle \rhd \langle \nu_{\text{Hat}} \rangle$ and $\langle \nu_{\text{Leg}} \rangle \rhd \langle \nu_{\text{Sho}} \rangle$ is selected based on the visual objects shown in Fig. 5.90a. The result of the selection is an object – the member of the leg category.

The task shown in Fig. 5.90b is more complex. In this example not only knowledge given by the categorical chains but also knowledge supplied by the knowledge chain is used to solve this task. At first two objects are interpreted as the category of mechanical tools: a hammer and a nail. The category of the mechanical tools is derived from the men's profession category $\langle \Pi \rangle \rhd \langle \sigma_{\text{El}} \rangle \rhd \langle \nu_{\text{ReO}} \rangle \rhd \langle \nu_{\text{Ear}} \rangle \rhd \langle \nu_{\text{Liv}} \rangle \rhd \langle \nu_{\text{Man}} \rangle \rhd \langle \nu_{\text{Prf}} \rangle \rhd \langle \nu_{\text{Mas}}, \nu_{\text{Car}}, \nu_{\text{Meh}}, \nu_{\text{Elc}}, \nu_{\text{Tay}}, \nu_{\text{Sec}} \rangle$. The tool category is derived from a man-made object, for example, the hammer category is derived from the category of mechanical tools: $.. \rhd \langle \nu_{\text{MMad}} \rangle \rhd \langle \nu_{\text{Tol}} \rangle \rhd \langle \nu_{\text{Mech}} \rangle \rhd \langle \nu_{\text{Ham}}, \nu_{\text{Wrn}} \rangle$. The nail category is derived from the material category and given by the following categorical chain $.. \rhd \langle \nu_{\text{MMad}} \rangle \rhd \langle \nu_{\text{Mat}} \rangle \rhd \langle \nu_{\text{Mech}} \rangle \rhd \langle \nu_{\text{Nal}}, \nu_{\text{ScD}} \rangle$. From the categorical chain of the knowledge category the relation between hammer and nail is found. The hammer category is derived from the tool category of the category of mechanical profession $.. \rhd \langle \kappa_{\text{Mech}} \rangle \propto (\kappa_{\text{Tol}}) \rhd (\kappa_{\text{Ham}}, \kappa_{\text{ScD}})$. The nail category is derived from the material category of the category of mechanical profession $.. \rhd \langle \kappa_{\text{Mech}} \rangle \propto (\kappa_{\text{Mat}}) \rhd (\kappa_{\text{Nal}}, \kappa_{\text{Scr}})$. From the category of casual relation between the tool category and the material category given as follows: $.. \propto (\kappa_{\text{Tol}}) \rhd (\kappa_{\text{Ham}}) \equiv [\kappa_{\text{Nal}}]$ the relation between the first and the second object in Fig. 5.90b is established. The third object in Fig. 5.90b is interpreted as a member of the screwdriver category. The similarity

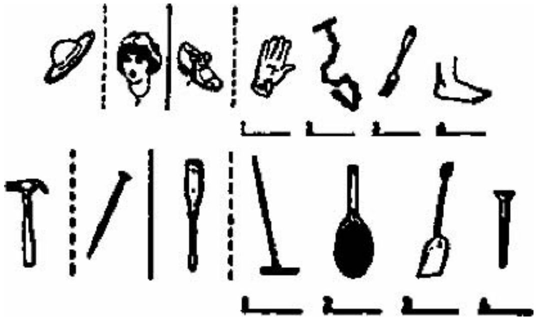


Fig. 5.90. Example of tests with analogical reasoning (**a**) the relation ‘wearing-part of body’ (**b**) the relation ‘tool-material’

relation that is found by comparison of the knowledge chains for the first and for the second objects $.. \propto (\kappa_{\mathrm{Tol}}) \triangleright (\kappa_{\mathrm{Ham}}) \equiv [\kappa_{\mathrm{Nal}}]$ and selection the fourth object is based on the knowledge chain of the third object $.. \propto (\kappa_{\mathrm{Tol}}) \triangleright (\kappa_{\mathrm{ScD}}) \equiv [\kappa_{\mathrm{Scr}}]$. As the result the object that is a member of the screwdriver category is selected.

5.5.2.9. Spatial Problems

Solving of visual problems often requires abilities to perform visual spatial transformations. These visual problems are often formulated as the tasks given in the form of intelligence tests. For example, in the test “a stack of overlapping tiles” the task is formulated as follows. The stack is turned over from left to right, so that the top left corner become the top right corner. Figure 5.91 shows how the stack would look after turning. Solving this task by choosing one of the four alternatives (A, B, C, and D) shows how the stack looks after turning from left to right. Solving this problem requires interpretation of the visual object as a member of the category of 3D figures or the category of real-world objects. For example, object in Fig. 5.91 (left on top) given by symbolic name $\rho[Q^4[L_Q^4](4L_R^4)]\{Q^4[L^8](4L_R^3), 4L_Q^4\}$ can be interpreted as a concave object $Q^4[L^8](4L_R^3)$ and four squares L_Q^4 attached to this concave object or as four squares L_Q^4 placed on the bigger square L_Q^4. The visual objects

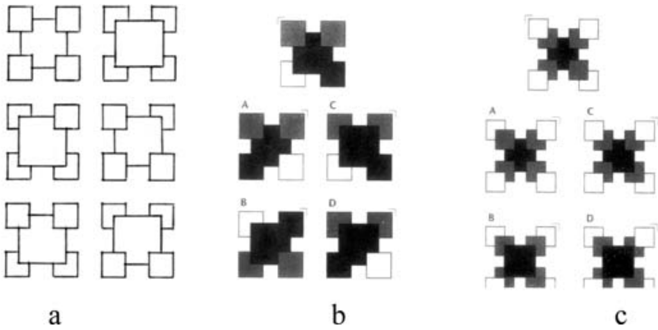


Fig. 5.91. Examples of tasks called "a stack of overlapping tiles"

shown in Fig. 5.91a belong to the "bird's eye view" line drawing perceptual category. The problem called "a stack of overlapping tiles" is similar to the problem described in previous sections of this chapter which was solved without reference to ontological categories. In this section the perceptual problems that are presented requires reference to ontological categories in order to be solved. In contrary to the tasks described in previous sections proper interpretation of the visual objects is selected based on the contextual information given by the linguistic description – "overlapping tiles." Each combination of tails is learned during the learning stage. During learning process the visual objects are learned as the linked pattern – composition of tails and theirs turned over version, as shown in Fig. 5.91a. Each tile represented by its symbolic name has its "turning" version that is marked by an arrow and is given as follows:

$$\rho[Q^4[L_Q^4](4L_R^4)]\{Q^4[L^8](4L_R^3), 4L_Q^4\} \Rightarrow \rho\{Q^4[L_Q^4](4L_R^4)\}\{L_Q^4, 4Q^1[L^5](L_R^3)\}$$

$$\rho\{Q^4[L_Q^4](4L_R^4)\}\{Q^1[L^6](2L_R^3), 2L_Q^4, 2Q^1[L^5](L_R^3)\} \Rightarrow \rho\{Q^4[L_Q^4](4L_R^4)\}\{Q^1[L^6](2L_R^3),$$
$$2L_Q^4, 2Q^1[L^5](L_R^3)\}$$

$$\rho\{Q^4[L_Q^4](4L_R^4)\}\{Q^1[L^5](L_R^3), L_Q^4, 3Q^1[L^5](L_R^3)\} \Rightarrow \rho\{Q^4[L_Q^4](4L_R^4)\}\{Q^1[L^7](3L_R^3), 3L_Q^4,$$
$$Q^1[L^5](L_R^3)\}$$

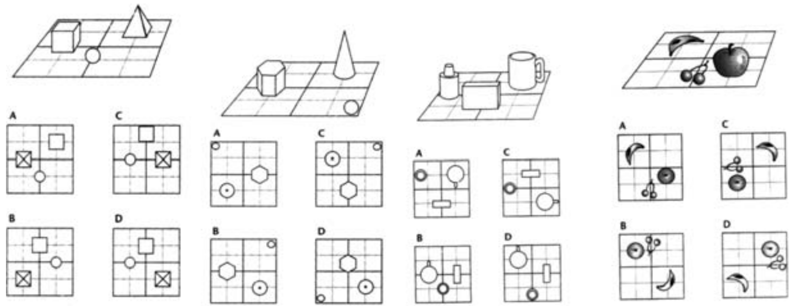


Fig. 5.92. Examples of tasks "objects on the table"

From this representation the solution is easy to obtain. The solution is found by comparison the symbolic name of the given stack with the symbolic names of the "turned" version of stacks. Similarly, the task shown in Fig 5.91b, c can be solved.

Solving the task for the tests given in Fig. 5.92 requires knowledge about the scene interpretation. The object needs to be interpreted as the real-world scene that consists of objects on the table and is seen from the "bird's eye view." Solving this task requires interpreting a visual object (on the top in Fig. 5.92) as the ontological category "objects on the table." The visual objects shown in Fig 5.92a–c belong to the perspective projection or "bird's eye view" line drawing perceptual category. At first the 2D figures are interpreted as the 3D solids or real-world objects and next they are placed on a given background. The 3D view of the scene is obtained from the 2D projection and matched with the scene given in the control image. The solids in the scene can be represented by the line drawing category (Fig. 5.92a–c) or as the category of shading object (Fig. 5.92d). The problem shown in these tests can be seen as the simplification of the real-world spatial problem, solution of which requires transformation of the visual scene into the 'bird's eye view'. The description of the test is given in the linguistic form as follows: "Each question (visual object) shows a 'bird's eye view' of a table top with a number of objects on it. Below are shown four possible plans of table top. The task is to select from A, B, C, and D the one which shows the correct positions of the objects on the top."

The solution to this problem requires interpreting the perceptual category of line drawing as the member of the ontological real-world category ("objects on the table"). The visual concepts of ontological category of "objects on

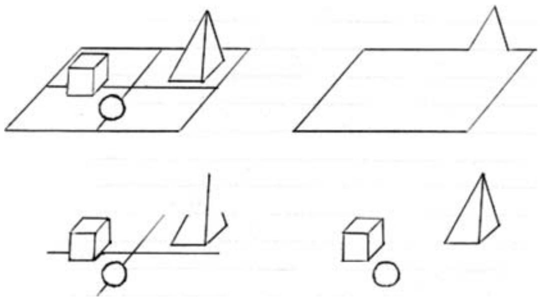


Fig. 5.93. Steps in finding the solution to the task "objects on the table"

the table" are learned as the structural pattern category. The structural pattern category of "objects on the table" consists of the elements – members of such categories as geometric solids, fruits, or cups.

The pattern category of the category of real-world scene such as ("objects on the table") supplies knowledge about processing the visual object in order to extract the information needed during solving this type of tasks. The transformations (processing) of perceived object during solving of the task are shown in Fig. 5.93. At first, the generic class of the "scene" is found and next, the border points are removed. The generic parallelogram and marked lines are found and removed. Objects are identified and interpreted in terms of 3D object and next, the "bird's eye view" for each figure is found and the position is reconstructed from the generic parallelogram and the marked lines. The solution is found by applying the following algorithm:

1. Find the 3D objects in the test in the relation to the lines that mark the location of each object
2. For each 3D object find all possible "bird's eye view"
3. Find placement of the object on the table
4. Transform the 3D object on the table into the "bird's eye view"
5. For each possible answer select one that matches the "bird's eye view" obtained in step 4

The view of the solids can be found from the characteristic views of solids. Learning of the category of solids involves learning of the visual concepts of solids. Learning of the visual concept of the ontological category "objects on the table" requires learning of the different visual aspects of this category.

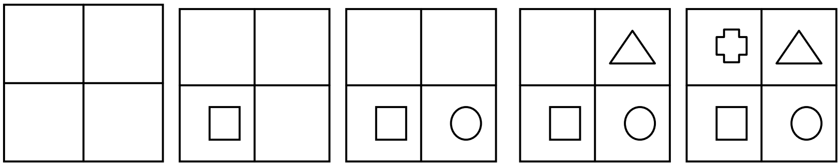


Fig. 5.94. Example of the objects on the table (the "bird's eye view")

Figure 5.94 shows the visual aspect (the "bird's eye view") of the members of the category "objects on the table." During learning process the representative of this category are selected and the symbolic names for each visual object are obtained. Example of the object that are representative of the visual aspect (the "bird's eye view") of the members of the category "objects on the table" are shown in Fig. 5.94. The symbolic names obtained for visual objects shown in Fig. 5.94 as follows: (a) $\rho[L_R^4]\{4L_R^4\}$, (b) $\rho[L_R^4]\}\{A^1[L_R^4](L_R^4), L_R^4, L_R^4, L_R^4\}$, (c) $\rho[L_R^4]\{A^1[L_R^4](L_R^4), A^1[L_R^4](K_C^1), L_R^4, L_R^4\}$, (d) $\rho[L_R^4]\{A^1[L_R^4](L_R^4), A^1[L_R^4](K_C^1), A^1[L_R^4](L^3), L_R^4\}$, (e) $\rho[L_R^4]\{A^1[L_R^4](L_R^4), A^1[L_R^4](K_C^1), A^1[L_R^4](L^3), A^1[L_R^4](Q^4[L^8](L_R^4))\}$.

In order to find the placement on the tabl e, the table is divided into four regions called cell. The placement in each cell is described by introduction a specific classes with additional symbols marking the placement of the object in the cell. Figure 5.95 shows the configuration of the objects in the cell. The symbolic names of the "cell" shown in Fig. 5.95 are given as follows: (a) $A^1[L_R^4](L_R^4[\{p\},\{s\}])$, (b) $A^1[L_R^4](L_R^4[\{p\},\{ld\}])$, (c) $A^1[L_R^4](L_R^4[\{p\},\{rd\}])$, (d) $A^1[L_R^4](L_R^4[\{p\},\{ru\}])$, (e) $A^1[L_R^4](L_R^4[\{p\},\{lu\}])$, (f) $A^1[L_R^4](L_R^4[\{p\},\{Sd\}])$, (g) $A^1[L_R^4](L_R^4[\{n\},\{s\}])$, (h) $A^1[L_R^4](L_R^4[\{n\},\{dP\}])$ (i) $A^1[L_R^4](L^4[\{p\},\{Sd\}])$, where symbol $\{s\}$ – denotes the centre of the cell, $\{lu\}$ – denotes the left upper part of the cell, $\{ru\}$ – denotes the right upper part of the cell, $\{ld\}$ – denotes the left down part of the cell, $\{rd\}$ – denotes the right down part of the cell, $\{Sd\}$ – denotes intersection with the line, $\{dP\}$ – denotes an object on the line.

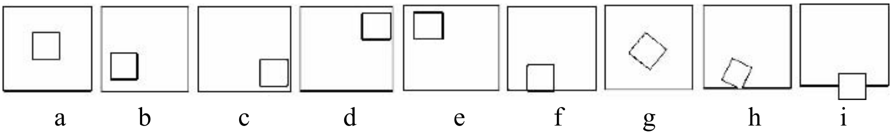


Fig. 5.95. Placement on the table: the table is divided into four regions called cells

Figure 5.96 shows the configuration of the fourth objects on table represented by the symbolic description. The symbolic names for objects shown in Fig. 5.96 are as follows:

a. $\rho\{L_R^4\}\{A^1_{L_R^4}(L_R^4)[\{p\},\{s\}], A^1_{L_R^4}(K_C^1)[\{p\},\{s\}], A^1_{L_R^4}(L^3)[\{p\},\{s\}],$
$A^1_{L_R^4}(Q^4_{L^8}(L_R^4))\ [\{p\},\{s\}]\}$

b. $\rho\{L_R^4\}\{A^1_{L_R^4}(L_R^4)[\{p\},\{Sd\}], A^1_{L_R^4}(K_C^1)[\{s\},\{ld\}], A^1_{L_R^4}(L^3)[\{n\},\{dP\}],$
$A^1_{L_R^4}(Q^4_{L^8}(L_R^4))[\{p\},\{Su\}]\}$

The symbols that are introduced make it possible to describe placement of the object on the table.

Interpretation of the real-world object in terms of its visual aspects given as the 2D representation requires that all characteristic aspects of the real-world object should be learned. In the case of solids the number of characteristic views (aspects) is significantly reduced. Each task that needs to be solved has its own specific representations that could reduce the number of characteristic views. For example, the solid can be visible from above when it is resting on one of its sides. The interpretation is given in the

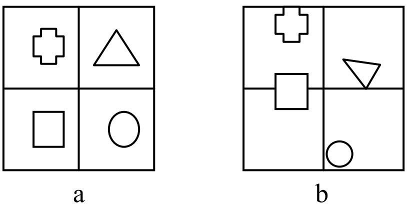


Fig. 5.96. Example of the objects on the table (the "bird's eye view")

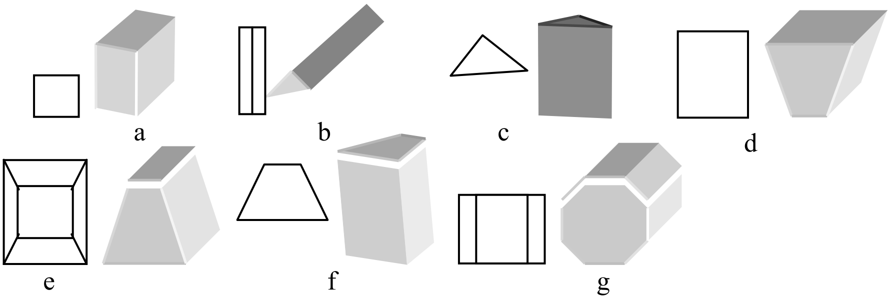


Fig. 5.97. The “bird’s eye view” of the selected solids

form of the aspect-solid relations. Symbolic names of objects (frontal view and “bird’s eye view”) shown in Fig. 5.97 are as follows: (a) $L_R^4 -> \rho\{L^6\}\{3L^4\}$, (b) $\rho\{L_R^4\}\{2L_R^4\} -> \rho\{L^{4-5}\}\{L^3,L^4\}$, (c) $L^3 -> \rho\{L^5\}\{L^4,L^3\}$, (d) $L_R^4 -> \rho\{L^{4-6}\}\{L_T^4 2L^4\}$, (e) $\rho\{L_R^4\}\{4L^4 L_R^4\} -> \rho\{L^{4-6}\}\{L_T^4 2L^4\}$, (f) $L_T^4 -> \rho\{L^{4-6}\}\{L_T^4 2L^4\}$, (g) $\rho\{L_R^4\}\{3L_R^4\} -> \rho\{L^{8-10}\}\{L^8,3L^4\}$.

5.5.2.10. Problem Solving – Categorical Knowledge

In many cases in order to solve the visual problem the realistic visual representation of the real-world object needs to be transformed into a schematic form. For example, the perceptual category of shaded object is transformed into the conventional line drawing object as shown in Fig. 5.98. The object shown in Fig. 5.98 is a member of the gear category derived from the category of real-world man-made object. The gear as a real object is perceived as the perceptual category of the shaded object (photograph of the object) given by the following categorical chain: $v_O \triangleright \langle \pi_{\mathrm{Sh}} \rangle \triangleright \langle \sigma_{\mathrm{Pt}} \rangle \triangleright \langle v_{\mathrm{ReO}} \rangle \triangleright \langle v_{\mathrm{Ear}} \rangle \triangleright \langle v_{\mathrm{NLiv}} \rangle \triangleright \langle v_{\mathrm{MMad}} \rangle \triangleright \langle v_{\mathrm{MMac}} \rangle \triangleright \langle v_{\mathrm{Mech}} \rangle \triangleright \langle v_{\mathrm{Gear}} \rangle$. The task of finding of the direction of the moving object can be approached by transforming the shaded object into category of the conventional line drawing object. During naming process

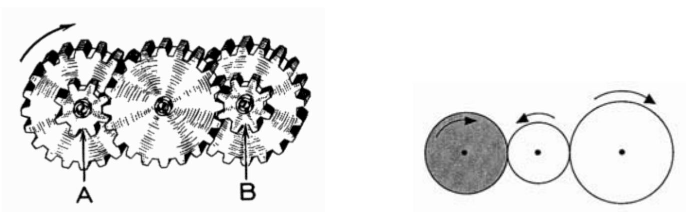


Fig. 5.98. Different visual representations of the gears

(recognition) the object given as the perceptual category of the shaded object $\nu_O \triangleright \langle \pi_{\text{Sh}} \rangle \triangleright ... \triangleright \langle \nu_{\text{Gear}} \rangle$ is transformed into silhouette $\nu_O \triangleright \langle \pi_{\text{Si}} \rangle \triangleright ... \triangleright \langle \nu_{\text{Gear}} \rangle$ and next into the schematic form of the gear $\nu_O \triangleright \langle \pi_{\text{Ld}} \rangle \triangleright \langle \sigma_{\text{Pt}} \rangle \triangleright \langle \nu_{\text{Sg}} \rangle \triangleright \langle \nu_{\text{VSym}} \rangle \triangleright \langle \nu_{\text{PhMo}} \rangle \triangleright \langle \nu_{\text{Gear}} \rangle$.

The schematic form of tools is obtained during learning process and very often it is represented as a perceptual category of the conventional line drawing object. The category of the physical visual model is represented by structural pattern category $\nu_O \triangleright \langle \pi_{\text{Ld}} \rangle \triangleright \langle \sigma_{\text{Pt}} \rangle \triangleright \langle \nu_{\text{Sg}} \rangle \triangleright \langle \nu_{\text{VSym}} \rangle \triangleright \langle \nu_{\text{PhMo}} \rangle \triangleright \langle \nu_{\text{Lev}}, \nu_{\text{Wed}}, \nu_{\text{Pul}}, \nu_{\text{Rol}}, \nu_{\text{Pen}}, \nu_{\text{InPl}} \rangle$. During learning of the visual concept of a given category of real-world object the schematic form of the object is also learned. Visual problem solving requires representing an object that is given as a photograph of the real-world object in the schematic form that shows only essential parts of the object.

The visual problem can be also given in the form of the linguistic description as an aid to the visual representation of the object given in the schematic form. Example of the task given in the form of linguistic description is the task of finding movement of rollers. In this task rollers are represented in schematic form as shown in Fig. 5.99, where one roller (shown in black) is the driving roller and it makes the other to run. It is assumed that the driving roller is moving in the clockwise direction. The rollers shown Fig. 5.99 consists of two groups of rollers: one group consists of rollers where movement occurs by placing the rollers in contact, the second one consists of rollers, where the drive belt connects the drive roller (shown in black) to the others. The task is formulated as follows: “which rollers will turn in the same direction as the drive roller, and which will turn in the opposite direction?” The category of rollers is derived from the category of man-made objects $.. \triangleright \langle \nu_{\text{MMad}} \rangle \triangleright \langle \nu_{\text{MMac}} \rangle \triangleright \langle \nu_{\text{Mech}} \rangle \triangleright \langle \nu_{\text{Rol}} \rangle$. During learning of the visual concept of roller the schematic form of objects as well as the realistic form of objects from this category is learned. Figure 5.99 shows examples of objects, representatives of the roller category that are used to learn the visual concept of this category.

During solving the task “which rollers will turn in the same direction as the drive roller?” that is given in the linguistic form there is a need to understand concepts such as “turn” or “direction” in the context of the schema of rollers. To understand these concepts there is a need to refer to the categorical chains of both the knowledge category and the category of visual objects. The solution to this task is obtained during visual inference process. At first, examined object is classified into one of the roller

a b c d e f g

Fig. 5.99. Example of visual representatives of the task of moving rollers

category – type I (rollers in contact) or type II (drive belt connects rollers). Next, after removing other symbols an examined object is classified to one of the concave or acyclic classes and the visual symbol is obtained. The visual symbol is used during the visual inference to find the solution. Finding the solution of the task represented by visual objects shown in Fig. 5.99a–d requires finding the visual concept of the roller category. Let us assume that the visual concept of rollers category (type I) obtained during learning process is as follows: $\varphi_{\mathrm{rol}\_I} \equiv \{[R]C(nK_O^1),\ [R]C(n \cdot K_O^1)\}$, where $[R]C(n \cdot K_O^1)$ denotes archetype of the complex roller class consisting of the n circles (placed along axis) (Fig. 5.99a–d). The rules of the visual inference that are learned as a part of the knowledge schema of the knowledge object are given as: $[\eta \in \varphi_{\mathrm{rol}\_I}] \Rightarrow a \triangleright \lambda_{\mathrm{rol}\_I}$. The task is formulated as follows: "which rollers will turn in the same direction as the drive roller" and let us assume that the perceived object that is shown in Fig. 5.99b is a visual representative of this task. During solving this task at first, the visual name $\eta \equiv C(3K_O^1)$ for the object shown in Fig. 5.99b is obtained and based on this symbolic name the rule of the visual inference is selected. The rule given in the form: $[\eta \in \varphi_{\mathrm{rol}\_I}] \Rightarrow a \triangleright \lambda_{\mathrm{rol}\_I}$ is selected based on the visual concept of the $\varphi_{\mathrm{rol}\_I} \equiv \{[R]C(nK_O^1), [R]C(n \cdot K_O^1)\}$. The task λ_{rol_I} is solved based on the knowledge about the movement of the rollers. The task is divided into series of subtasks that need to be solved based on the previously known solution. One of the subtask is to assign the direction of movement to all rollers based on the knowledge expressed in the following form: $[r_i - > d_{\mathrm{clock}}] \rightarrow r^{i+1} \triangleright d_{\mathrm{anticlt}}$, that means if roller r_i is moving in the clockwise direction the roller r_{i+1} is moving in the opposite direction.

5.5.2.11. Visual Diagnosis

The visual problems are often formulated based on a set of visual symptoms and solved during the visual diagnosis. Visual medical diagnosis is a special type of the diagnosis that is based on visual reasoning about the pathological changes of the selected organs of the living organism. The term "visual diagnosis" refers to the diagnosis that based on symptoms given in the form of the visual objects made inference about the illness that cause pathological changes of these objects. Visual diagnosis is performed in order to undertake an appropriate action such as a selection of the medical treatment or a change of the parameters of the process. The visual medical diagnosis is based on the medical images from which the phantom is to be extracted. The medical diagnoses utilize the medical images in order to extract the image region that represents a given organ. The image region is first extracted by applying one of the segmentation methods and next the shape analysis and recognition is performed in order to classify it to one of the categories of pathological symptoms. In medical diagnosis the following types of medical images are used: the X-ray photographs, the magnetic resonance (MR) images, the ultrasonography (USG) images or the images obtained by application of the microscope. These problems were discussed in [11].

The visual categories of pathological symptoms refer to parts of the body that are changed by the illness. The visual category of pathological symptoms derived from the category of micro world refers to the changes of micro-organs. These visual objects are visible under microscope and can be extracted from image by applying the segmentation method. The different parts of the body can be described in terms of the category of healthy organs and the category of not healthy organs. During learning stage the category of healthy organs and pathological changes are learned in the context of the category of illness. The visual objects that represent the pathological changes are used during learning of the categories of illness. The following categorical chain represents the knowledge about the pathological changes. The specific categories such as bacteria or viruses are derived from the category of micro-organisms $O \rhd \langle \nu_{\mathrm{ReO}} \rangle \rhd \langle \nu_{\mathrm{Mic}} \rangle \rhd \langle \nu_{\mathrm{Liv}} \rangle \rhd \langle \nu_{\mathrm{Org}} \rangle \rhd \langle \nu_{\mathrm{Vir}}, \nu_{\mathrm{Bac}} \rangle$. Categories of part of the body are derived from the human category and are represented by following categorical chain: $\langle O \rangle \rhd \langle \nu_{\mathrm{ReO}} \rangle \rhd \langle \nu_{\mathrm{Ear}} \rangle \rhd \langle \nu_{\mathrm{Liv}} \rangle \rhd \langle \nu_{\mathrm{Hum}} \rangle \succ [\tau_{\mathrm{Hed}}, \tau_{\mathrm{Nek}}, \tau_{\mathrm{Hnd}}, \tau_{\mathrm{Leg}}, \tau_{\mathrm{Trk}}]$. Each body part category is divided into the specific categories that indicate if this part is visible and is represented by the category $\left\langle \nu_{Out} \right\rangle$ such as the visible parts of the head category $.. \rhd \langle \nu_{\mathrm{Hum}} \rangle \succ [\tau_{\mathrm{Hed}}] \rhd \langle \nu_{\mathrm{Out}} \rangle \rhd [\tau_{\mathrm{Nos}}, \tau_{\mathrm{Eye}}, \tau_{\mathrm{Mth}}, \tau_{\mathrm{Har}}]$, or are hidden such as the hidden parts of the head category $.. \rhd \langle \nu_{\mathrm{Hum}} \rangle \succ [\tau_{\mathrm{Hed}}] \rhd \langle \nu_{\mathrm{In}} \rangle \rhd [\tau_{\mathrm{Brn}}, \tau_{\mathrm{Scl}}, \tau_{\mathrm{Msc}}]$. Each category of the part of the body is divided into the category of healthy organ and the categories of pathological changes. For example, the mouth category is divided into the category of healthy mouth and the category of the pathological changes and is represented by the following categorical chain: $.. \triangleright \langle \nu_{\text{Hum}} \rangle \succ [\tau_{\text{Hed}}] \triangleright \langle \nu_{\text{Out}} \rangle \triangleright [\tau_{\text{Mth}}] \triangleright \langle \nu_{\text{Hel}}, \nu_{\text{Ptl}} \rangle$.

During diagnostic process objects with pathological changes that are extracted from an image are representatives of one of the perceptual categories. These objects refer to specific organs or parts of the body that show the pathological changes. The category of illness derived from the knowledge category refers to the categories of parts of the body. The categories of parts of the body supply the knowledge where the different factors important in diagnostic procedure are defined. During the diagnostic process the contextual information (results of selected laboratory tests) can be also used to interpret of the visual object. The following categorical chains represent the knowledge about the diagnostic process: $.. \triangleright \langle \kappa_{\text{Med}} \rangle \triangleright \langle \kappa_{\text{Hum}} \rangle \succ [\tau_{\text{Hed}}] \triangleright \langle \kappa_{\text{Out}} \rangle \triangleright [\tau_{\text{Mth}}] \triangleright \langle \kappa_{\text{Ptl}} \rangle \prec \{\partial_{\text{NIl}}, \partial_{\text{ViCl}}, \partial_{\text{Defl}}, \partial_{\text{Trtl}}\}$. The knowledge schema includes the name of the illness ∂_{NIl}, the visual concept of the pathological changes ∂_{ViCl}, the medical definition ∂_{Defl} and links to the more specific categories of the recommended medical treatment ∂_{Trtl}. It is assumed that knowledge of the illness that is linked with the characteristic pathological changes exists and that knowledge about this illness is represented by the illness category acquired during the learning process. During medical diagnosis the image is transformed into visual objects by applying perceptual transformations that involves application of the segmentation method and the pre-processing techniques. Figure 5.100a shows the medical image and the result of the perceptual transformation (Fig. 5.100b). Figure 5.100c shows the healthy cells and Fig. 5.100d–f shows cells with different pathological changes that indicate that these cells are members of the different illness categories.

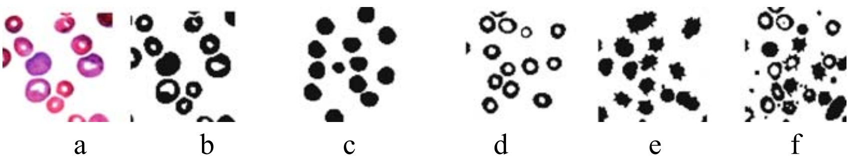


Fig. 5.100. Example of cells with pathological changes (**a**) an image from which visual objects are extracted, (**b**) visual objects (cells with pathological changes) extracted from images, (**c**) example of healthy cells, (**d–f**) examples of cells with pathological changes

Visual objects shown in Fig. 5.100c are convex objects representatives of the category of healthy cells. Figure 5.100d show cells with pathological changes that are representatives of the category of the illness I (convex cyclic objects). Figure 5.100e shows cells with pathological changes representatives of the category of the illness II (concave objects). Figure 5.100f shows cells pathological changes representatives of the category of the illness III (convex cyclic objects, concave objects). During learning process the visual concept φ^h for each illness category is obtained. The healthy cells and cells with pathological changes are described by the illness categories. During the visual diagnosis the visual concept φ^h is used to find the illness category H^i, by applying the inference rules: $[\eta \in \varphi^{\mathrm{h}}] \Rightarrow h \triangleright H^{\mathrm{i}}$. For example, let us assume that during visual diagnosis the symbolic name $A^1[K](K)$ (the convex cyclic curvilinear object) was obtained. Let us assume that the visual concept of the category of illness I is given as a set of symbolic names $\varphi_I^h = \{A^1[K](K), A^1[K](M), A^1[M](K), A^1[M](M)\}$. Because the symbolic name $A^1[K](K)$ is a member of the visual concepts φ_I^h, the category of illness I is selected as the result of the visual diagnosis. The category of illness I is linked with the category of treatment of the illness I so the result of diagnosis is used to find recommendation of the treatment.

The visual diagnosis of technological process (process control) utilizes the images that are obtained by application of the imaging techniques such as X-ray photographs or microscopic images. In the process control the visual diagnosis is based on the characteristic changes of the shape of a given visual object. It is assumed that a phantom (visual object) that is extracted from the image includes information about changes of the technological process. The changes of shape during the certain interval of time can "produce" the characteristic sequences of shapes that are characteristic for changes of the process. The sequence of shapes can be represented by a sequence of symbolic names $\eta^1, \eta^2, ..., \eta^n$ obtained during reasoning process from sequence of the images. The failure category (visual concept) φ^h that is obtained during the learning process is used to define rules of visual inference of the failure (critical points) P^i of the process p. The rules of the visual inference are expressed as: $[\eta^1, \eta^2, ..., \eta^n \in \varphi^p] \Rightarrow p \triangleright P^i$. Similarly like in medical visual diagnosis the category of failure is connected with recommendation of proper changes of the parameters of the process in order to avoid failure. The

Fig. 5.101. Microscopic images of the metallurgical process where shape of the particles indicates the quality of the process

visual diagnosis can be also used to identify the quality of the material by analyzing of the shape of the microscopic structure. Examples of the microscopic images where shapes of structural elements are symptoms that characterize the different properties of the final metallurgical process are shown in Fig. 5.101.

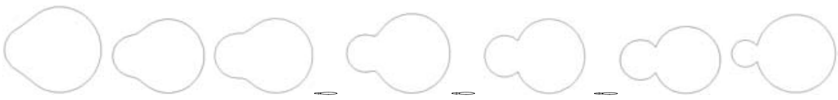
The visual concept of the process changes can be learned based on mathematical model. In the case when shape of the curve depends on value of parameters of the equation that describes this curve, the equation can be applied to model visual changes of shape and used as a model of the real-word phenomena or as a model of pathological changes in the visual diagnosis. The visual model is represented by the visual concept that consists of symbolic names of characteristic shapes. As it was shown in previous section in the visual diagnosis the process of the visual inference is expressed in the form of rules $[\eta^1, \eta^2, ..., \eta^n \in \varphi^p] \Rightarrow p \triangleright P^i$, where η is the symbolic name obtained as the result of the reasoning process, and $[\eta^1, \eta^2, ..., \eta^n \in \varphi^p]$ is the sequence of changes of the shape category. For example, the process of dividing the drop of oil can be modeled by Cassinian ellipse. During learning stage the visual concept of the category dividing of the drop of oil is learned based on exemplars generated from the Cassinian ellipse equation $(x^2 + y^2)^2 - 2c^2(x^2 - y^2) = a^4 - c^4$. The visual concept of the phenomenon "dividing the drop of oil" is as follows:

$$\varphi_{\text{CasOv}} = \{Q^2[M^4[L^4]](2 \cdot Q^2[L^3](2 \cdot M^1)),\ Q^2[M^2[L_R^4]](2 \cdot Q^2[\hat{L}^3](2M^1)),$$
$$Q^2[M^2[L_R^4]](2 \cdot Q^2[\hat{L}^3](2\tilde{M}^1)),\ Q^2[M^2[L_R^4]](2 \cdot Q^2[\hat{L}^3](2\tilde{\tilde{M}}^1)), \overline{K}_S^1\}$$

Fig. 5.102 shows the archetypes generated from the equation $(x^2 + y^2)^2 - 2c^2(x^2 - y^2) = a^4 - c^4$ applying Mathematica function ImplicitPlot[(x^2 + y^2)^2–2*CC^2*(x^2 – y^2) = A^4 – CC^4, {x, –5, 5}]. The Mathematica is part of the Generating Expert of the SUS that was described in Chap. 1. Archetypes where generated for the parameters (a) CC = 1, A = 1, (b) CC = 1, A = 1.02, (c) (CC = 1, A = 1.1), (d) (CC = 1, A = 1.2), (e) (CC = 1, A = 1.5). During visual diagnosis the shape of the visual object is used to find the stages of the process.

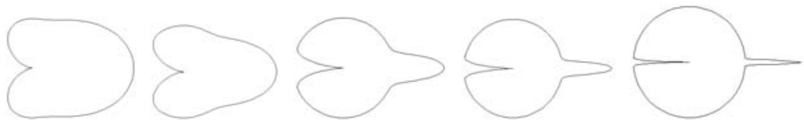
a b c d e

Fig. 5.102. Archetypes generated from the Cassinian ellipse equation



a b c d e f g

Fig. 5.103. Archetypes generated from the Cranioid curve equation



a b c d e

Fig. 5.104. Archetypes generated from the Petalbotanik curve equation

Similarly the cranioid equation and petalbotanik equation can be used for modeling of the visual processes or for learning of the visual concept of the visual diagnosis. Examples of the different shapes obtained from Cranioid equation $\rho = a\cos\varphi + b\sqrt{1-m^2\sin^2\varphi} + \sqrt{1-k^2\sin^2\varphi}$ (PolarPlot[*A**Cos[*t*] + *B**Sqrt[(1 − M*M*Sin[t]*Sin[t])] + CC*Sqrt[(1 − K*K*Sin[t]*Sin[t])], {t,0.00001,2*Pi}]) for the parameters (a) A = 1 B = 1 M = 0.5 K = 0.95 CC = 1, (b) A = 1 B = 1 M = 0.5 K = 0.95 CC = 1.5, (c) A = 1 B = 1 M = 0.5 K = 0.95 CC = 2, (d) A = 1, B = 1, M = 0.9, K = 0.95, CC = 1.5, (e) A = 1, B = 1, M = 0.5, K = 0.999, CC = 2, (f) A = 1, B = 1, M = 0.5, K = 0.999, CC = 2.5, (g) A = 1 B = 1 M = 0.999 K = 0.95 CC = 1.5 are shown in Fig. 5.103.

Examples of the different shapes obtained from Petalbotanik equation $\rho = a(1+\cos^{2n+1}\varphi)$, (PolarPlot[*A**(1 + Cos[*t*]^(2**NN* + 1)), {t,0.00001,2*Pi}]) for parameters (*A* = 1, (a) *NN* = 1, (b) *NN* = 3, (c) *NN* = 25, (d) *NN* = 100, (e) *NN* = 1000) are shown in Fig. 5.104.

5.5.2.12. Assembling Tools

The problem of assembling of a tool from *n* parts is an example of the complex visual problem solving. Assembling the complex object can yields many problems (questions) that need to be solved during assembling process. For example, the following questions can be asked: “in which order can an object be assembled or disassembled?”, “how many degrees of freedom?”, “what parts should be withdrawn to allow the removal of a specified subassembly?”. Assembling can be based on the inventive (creative) imagery process, on the learned previous cases or on the schematic drawings that represent description of the assembling process. In the case of the inventive (creative) imagery process the solving problem of assembling tools requires to imagine the situation of the assembling tool and next to transform the learned visual experience of the real-world situation to solve the theoretical problem. In the case of assembling tools, based on the learned previous cases, the previous experience is used to select parts and to assemble the tool. In the case of assembling tools, based on the schematic drawing, at first the drawing needs to be interpreted and next the knowledge obtained from the reading of the schematic drawing is applied to perform the assembling task. Interpretation involves matching the drawing parts with the real-world objects, finding the succession of the performing subtasks and finally assembling tools. The schematic representation that shows the way of assembling tools use nearly realistic representation of the real-world object. The schematic representation can also use the symbolic representation of parts to be assembled (e.g., electrical schema).

The schematic representation is used to shows how tools work, to show main stages of the process, to give explanation about the physical or engineering process or to illustrate the problem that needs to be solved. There is a schematic representation that shows only the functional aspect of the model. In such a schematic representation real-world objects are represented by well defined symbols. For example, the electric circuit given in Fig. 5.105 shows only the schematic representation of the real-world objects such as a resistor, a capacitor, or a battery. The schema of the electric circuit is used to design and solve the electrical problem or to build the electric devices. To understand the functionality of the tools or solve the specific electric problem the knowledge from the area of electrical circuits is needed. The electrical symbols are interpreted as a “substitute” of the

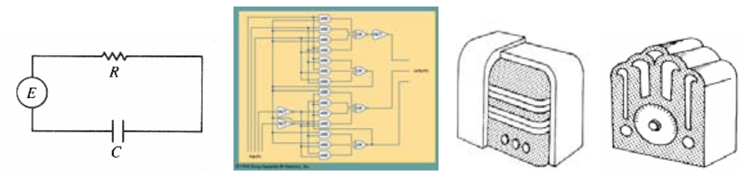


Fig. 5.105. Examples of schemas of electrical circuits that are used during assembling of electrical devices

real-world objects such as a resistor or battery. The electrical symbols are members of the category that is derived from the category of the visual symbols and the category of electrical symbol is "linked" with the category of the real-world electrical elements.

In order to be able to assemble complex device there is a need to use schematic representation called working drawings. Working drawings are the complete set of standardized drawings that specify the manufacture and assembly of a product based on its design. Generally, a complete set of working drawings for an assembly includes: detail drawings of each non-standard part, an assembly or subassembly drawing showing all the standard and nonstandard parts in a single drawing, a bill of materials, a title blocks. A detail drawing is a dimensioned, multiview drawing of a single part, describing the part's shape, size, material, and finish, in sufficient detail for the part to be manufactured based on the drawing alone. A complete set of working drawings must include a detailed parts list or bill of material. The information normally included in parts is as follows: name of the part, a detail number of the part in the assembly, and the part. During learning of the tool category there is a need to learn an information that is included in working drawings.

An assembly drawing shows how each part of a design is put together. Figure 5.106 shows examples of working drawings. These drawings can be used to learn the visual concept of assembling tools.

In order to explain how SUS can solve assembling problem examples from different areas of the assembling problems are described in this section. The simplest assembling task is to form (assemble) of the figure from the different parts.

Modeling the incremental fabrication of a part is a problem that needs to be solved by showing the main stages of the process. In this approach the problem is decomposed into sequences of sub-problems. These sub-problems can be solved by applying visual reasoning. The Fig. 5.107 shows example of the sequence of images that represent each sub-problem that need to be solved during incremental fabrication of a part.

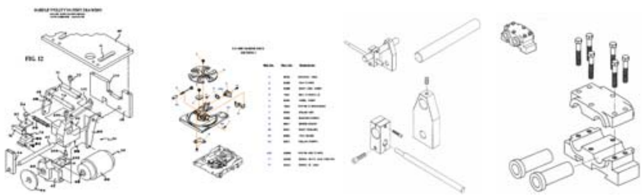


Fig. 5.106. Example of assembly drawings

Fig. 5.107. Modeling the incremental fabrication of a part

Fig. 5.108. Examples of flat figures

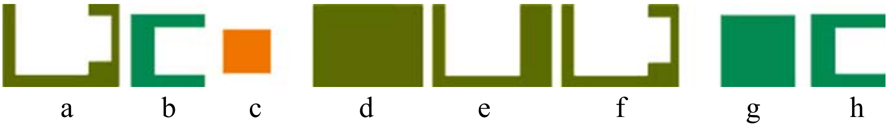


Fig. 5.109. Preparing parts for assembling

One of the sub-problems that need to be solved is to fabricate a required part. The solution is easier to obtain when assembling task is concerned with assembling of the “flat” parts. The process of assembling task of the “flat” parts can be given in the form of the 2D representation and can be represented by one visual aspect (view) of the figure. Examples of the flat figures are shown in Fig. 5.108.

The object that is “flat” can be represented by one visual aspect. Assembling of the flat object can be represented by showing only one visual aspect of the visual object. Also, solving some tasks such as building arc that is 3D real-world object can be reduced into the steps performed on the 2D object. Figure 5.109 shows the part that is prepared for assembling with other parts. For example, assembling the object shown in Fig. 5.113g requires preparation of three different components that are shown in Fig. 5.109: the components I (a) the component II (b) and the component III (c). Components shown in Fig 5.109 represented by symbolic names (a) $Q[L_R^4](Q^2[L^6](2L^3))$, (b) $Q^1[L_R^4](L_R^4))$, (c) L_R^4 are used to learn the visual concept of parts of the object. The visual concept of the object that is assembled from the different parts (components) consists of symbolic name of the whole object and its components. These components need to be obtained by cutting of the two rectangular objects shown in Fig. 5.109(d–f) (component I) and Fig. 5.109(g–h) (component II). For example, the first component shown in Fig. 5.109a can be obtained in

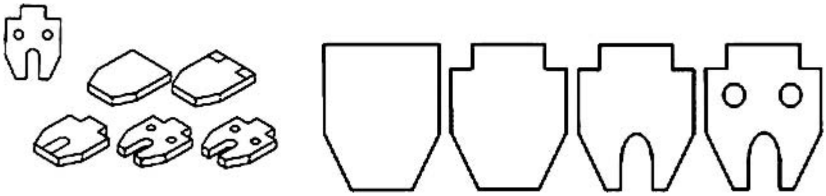


Fig. 5.110. Obtaining part by cutting and drilling

three steps (shown in Fig. 5.109d–f) that are represented by the operational chain as a sequence of symbolic names $L_R^4 -> Q^1[L_R^4](L_R^4)] -> Q^1[L_R^4](Q^2[L^6](2L^3))$.

The component shown in Fig. 5.110 is obtained by performing series of operations such as cutting parts and drilling the holes. This task can be described by the operational chain: $L_M^6 \to Q^2[L_M^6](2 \cdot L_R^3) \to Q^2[L_M^6](2 \cdot L_R^3, M^1[L_R^4]) \to A^2[Q^2[L_M^6](2 \cdot L_R^3, M^1[L_R^4])](2 \cdot K_C^1)$. This operational chain can be applied during visual thinking process connected with solving the assembling task. During thinking process the operational chain is represented by a sequence of visual transformations that make it possible to imagine (by SUS) steps that lead to the solution. For example, cutting parts is represented by the imagery transformation $\varphi^1 : \{L_M^6\} \to \{Q^2[L_M^6](2 \cdot L_R^3)]\}$. The imagery transformation φ^1 transform a visual object that is given in the form of critical points Π and represented by symbolic name η into another object (a set of critical points). The imagery transformation $\varphi^1 : \{L_M^6\} \to \{Q^2[L_M^6](2 \cdot L_R^3)]\}$ transforms a set of critical points given by symbolic name L_M^6 into another set $Q^2[L_M^6](2 \cdot L_R^3)]$. The symbolic name is used to indicate that the imagery transformation is a transformation that is performed as a part of the visual thinking process.

Assembling is often given as a task that includes selection of components that can fit into the other component that is called the basic component. To solve this task there is a need to have a good interpretation of the visual object in terms of its components. This task can be reduced into selection of the parts from the given set of components that can fit into a basic component. Solution to this task can be often represented by using the 2D representation. Figure 5.111b–f shows the different components that need to be fitted to the basic component (the cyclic convex object) shown in Fig. 5.111a. This problem can be solved by

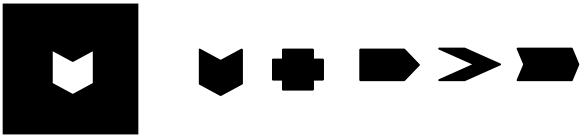


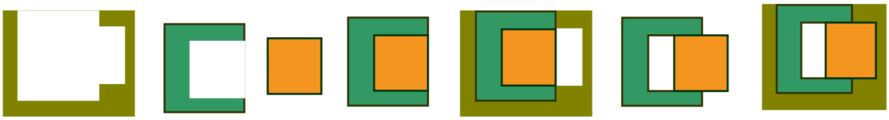
Fig. 5.111. The visual representation of the assembling problem

comparison shape of the hole with shape of each part that needs to be fitted into it. At first the symbolic name of the basic components and symbolic names for each component are obtained. From symbolic names that were obtained one that has the same symbolic name as the hole is selected. The basic component (shown in Fig. 5.111a) is an object from the cyclic class $A^1[L_R^4](Q^1[L^5](L^3))$ and the hole is a member of the concave class $Q^1[L^5](L^3)$. Symbolic names of components are given as follows: (b) $Q^1[L^5](L^3)$, (c) $Q^4[L^8](4L^3)$, (d) L^5, (e) $Q^1[L^5](L^3)$, (f) $Q^1[L^5](L^3)$. There are three parts that have the same symbolic name so there is a need to obtain the symbolic name at the specific level (b) $Q^1[L_M^5](L_O^3)$, (e) $Q^1[L_S^5](L_A^3)$, (f) $Q^1[L_G^5](L_A^3)$, and (a) $Q^1[L_M^5](L_O^3)$. As the result of the specification, an object given by the symbolic name (b) $Q^1[L_M^5](L_O^3)$ that has the same symbolic name as object in Fig. 5.111a is selected.

Another example of the selection process is shown in Fig. 5.112. The solution, similarly like in the previous example, is obtained by finding the symbolic names of parts. The whole object is given by the symbolic name $\rho[L_R^4]\{Q[L^6](2L_R^3), Q[L_R^4](L_R^4)\}$. Parts shown in Fig. 5.112 have the symbolic name (a) $Q[M^1[L^6]](2Q^2[L^3](M_{L^3}^1))$, (b) $Q[M^1[L^6]](2Q^2[L^3](M^1))$, (c) $Q^2[L^5](L_O^3)$, (d) $Q^2[L^6](2L_R^3)$, (e) $Q^2[L^6](2L_R^3)$ and part to which other components need to be fitted is given as $Q[L_R^4](L^4)$ (Fig. 5.12f). To solve this problem the symbolic description of the concave class is converted into description of the complex class (a) $C[L_R^4, M^1[L_R^4]]$, (b) $C[L_R^4, M^1[K^1]]$, (c) $C[L_R^4, L^3]$, (d) $C[L_R^4, L_R^4]$, (e) $C[L_R^4, L_R^4]$. From this representation we can infer that only objects shown in Fig. 5.112d, e have rectangular part that can be fitted to the rectangular part of the object in Fig. 5.112f. The final fitting is based on the specific representation of the class or can be based on the measurement of components.

a b c d e f g

Fig. 5.112. Assembling two parts: (**d–e**) objects that are representatives of the compatible components, (**a–c**) objects that are representatives of the noncompatible components, (**e**) a basic component, (**g**) the whole object



a b c d e f g

Fig. 5.113. Assembling process: the whole object (**g**) is assembled from the three components, (**a–c**) given as a representatives of the colored class, assembling steps (**d–f**)

The assembling task is given by the sequence of steps. In the previous section two steps: preparing components and parts selection were described. In this section assembling tools from components is described. In this task there are three components shown in Fig. 5.113(a–c) and there is a need to assemble these components into the whole object shown in Fig. 5.113g. The different components can be represented by notation of the colored class. For example, the symbolic name for the whole object shown in Fig. 5.113g is as follows: $\rho[L_R^4]\{(a)\langle Q[L^4](Q[L^6](2L_R^3)]\rangle, \rho[Q[L^6](2L_R^3)]$ $\{(g)\langle Q[L_R^4](L_R^4)\rangle, \rho[L_R^4]\{(w)\langle L_R^4\rangle, (o)\langle L_R^4\rangle\}\}\}$.

This problem can be transformed into the representation of the convex, the concave, and the cyclic class. There are three components shown in Fig. 5.114 given by symbolic names (a) $Q[L_R^4](Q^2[L^6](2L^3))$, (b) $Q^1[L_R^4](L_R^4))$, and (c) L_R^4. The solution can be written in the form of the sequence of the imagery transformations that is the part of the visual thinking process. The imagery transformation that is applied at first stage of problem solving is transformation that reduces concavities (Fig. 5.114(b–d)). The first transformation is the transformation concave-convex $Q^1[L_R^4](L_R^4)) \oplus L_R^4 \equiv L_R^4$. The second transformation is the transformation concave-cyclic $Q[L_R^4](Q^2[L^6](2L^3)) \oplus L_R^4 \equiv A[L_R^4](L_R^4)$ (see Fig. 5.114a, d, e). The next transformation is the translation transformation $L_R^4 \otimes L_R^4 \equiv A[Q^2[L^6](2L^3)](L_R^4)$ and $A[L_R^4](L_R^4) \oplus L_R^4 \equiv \overline{A}[L_R^4](L_R^4)$. The translation transformation is performed on one part of the object and changes the object as a whole (see Fig. 5.114 (e–g)).

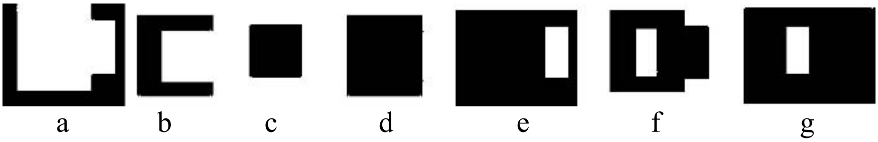


Fig. 5.114. Assembling process of the parts given as a representatives of the concave, convex, and cyclic class

Fig. 5.115. The different representatives of the arch category

Assembling objects from components requires planning sequences of operation that leads to obtaining an object that is composed from the different parts. Each known object that was produced belongs to the category of the man-made object. It is assumed that the object from this category is produced by the worker (or group of workers) by applying special tools, using an appropriate material and knowledge. The result of the work is the final product. Building an arch is given as an example at the assembling of the object from the simple components. The arch can have the different forms some of them are shown in Fig. 5.115.

The problem of planning actions during building of the arch is often reported in literature. For example, in [12] steps involved in learning of the description of an arch are shown. The system Marvin learnt that an arch consists of an object of any shape lying on top of two columns of equal height. The columns are adjacent to each other, but they must not touch. Learning starts from learning of the definition of basic concepts. When all basic concepts are learned the MARVIN is able to learn the concept of arch.

In our example the arch is assumed to have the shape shown in Fig 5.115. Kowalski in [3] gives an example of the definition of this type of arch by representing it in the form of the clauses. This type of description is very complex and very difficult to check if it is correct. In our approach the arch is a member of the category of man-made objects and is represented by the visual concept. The visual concept refers to the 3D visual representation showing characteristic visual aspect of the object. In the case of the category of the arch type R (arch-R) the visual concept refers to the most typical aspect of the object and is given by the symbolic name of the line drawing orthographic projection of the object, as:

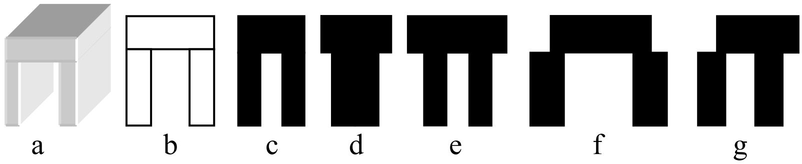


Fig. 5.116. Arch (**a**) its 2D representation (**b, c**) and not proper cases (**d–g**)

$\rho[Q[L^6](Q[L^4](L^3)])]\{3\cdot L_R^4, 2L_O^4, L_T^4\}$ (see Fig. 5.116a). In the case of understanding, designing, planning and assembling arch-R the 2D representation given in the form of the frontal view is sufficient to perform the required operation. The symbolic name of the 2D representation of the arch-R is given as $\rho[Q[L^4](L_R^4)]\{3L_R^4\}$ (see Fig. 5.116b). This frontal view can be obtained by appropriate placing of the camera (in front of arch). Assembling of the arch can be understood by imagine production process of the arch-R. For example, the arch shown in Fig. 5.116c is produced from three blocks and is represented by the symbolic name $Q[L_R^4](L_R^4)$. After moving the top block we obtain two columns. The proper placement of columns can have an impact on the final result. Examples of the non proper placement of the columns can results in obtaining the para-arch shown in Fig. 5.116d–g represented by symbolic name (d) $Q[L^6](2L_R^3)$, (e) $Q[L_A^6](2L_R^3, L_R^4)[bababa]$ (f) $Q[L_B^6](2L_R^3, L_R^4)[bbaaba]$. The non proper placement of the top block can results in the para-arch shown in Fig. 5.116g $Q[L_C^6](2L_R^3, L_R^4)[bbbaaa]$. The symbolic description of this non proper arch is part of the visual concept of the para-arch category that is learned. The visual concept of the para-arch category can be used to design action that need to be undertaken to correct the non proper placements of the blocks.

In contrast to the definition given in terms of clauses the symbolic representation can give the very detail description of the blocks configuration showing the similarities and the possible incorrect version of the arch assembled from blocks. The solution for the arch building is given in the form of sequences of actions as follows:

$$\{(\uparrow)L_R^4 - (\uparrow)L_R^4\} + (\leftrightarrow)L_R^4 -> Q[L_R^4](L_R^4)$$

$$L_Q^4 + L_Q^4 \equiv (\uparrow)L_R^4 \to \{(\uparrow)L_R^4 - (\uparrow)L_R^4\} + (\leftrightarrow)L_R^4 -> Q[L_R^4](L_R^4)$$

where $(\uparrow)L_R^4$ denotes that rectangle is placed vertically (on the shortest side), $\{(\uparrow)L_R^4-(\uparrow)L_R^4\}$ denotes that two rectangles are placed close to each other, $(\leftrightarrow)L_R^4$ denotes that rectangle is placed on the longest side, $\{(\uparrow)L_R^4-(\uparrow)L_R^4\}+(\leftrightarrow)L_R^4$ denotes that rectangle $(\leftrightarrow)L_R^4$ is placed on the two rectangles $\{(\uparrow)L_R^4-(\uparrow)L_R^4\}$. Figure 5.117 shows different representatives of arches that are build from the smaller blocks.

During assembling of the non flat object there is a need to take into account the different views of the object. Before assembling the object we need to know the category to which an object belong. Assembling task requires undertaking the sequence of actions that leads to obtain the final result. These actions can be performed in "mind" as the sequence of imagery transformations during visual thinking process. During thinking process problems connected with selection of components and finding the optimal sequence of actions can be solved. For example, two parts shown in Fig. 5.118a, b need to be assembled in order to obtain the object shown in Fig. 5.118c. During assembling process the proper view of the object (aspect) needs to be selected. The selected aspect refers to the way in which the object will be "observed" during assembling process. For example, Fig. 5.118a–c shows the visual aspect that makes it possible to see "whole" object. The symbolic name of this aspect of the object is given as $\rho[Q[M^2[L^6]](2\cdot Q[L^3](M^1)]\{2Q[M^1[L_R^4]](M^1),Q[K_E^1](M^1[L_R^4]),K_E^1\}$

Fig. 5.117. Arch build from the "small" blocks

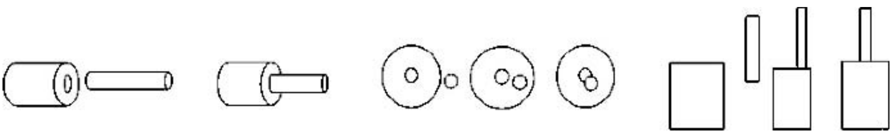


Fig. 5.118. Example of the assembling tools – 3D representation

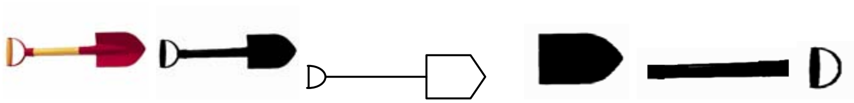
(Fig. 5.118c). The symbolic name of parts are given as follows: the part I (Fig. 5.118a) $\rho[M^2[L^4]\{Q[M^1[L_R^4](M^1), A[K_E^1](K_E^1)\}$ and part II (Fig. 5.118b) $\rho[M^2[L^4]\{K_E^1\}$. This visual aspect does not allow SUS to "see" well the movement of parts during assembling process. To "see" well the movement of parts during assembling process two visual aspects shown in Fig 5.118d–i were selected. Assembling requires placement of the part I into the hole. The two views (top and frontal) are used to trace movement of the part II. The movement of the part II is translated into the symbolic names and represented as a sequence of possible actions (the operational chain) given as follows:

$$\left\langle A[K_C^1][K_C^1] - K_C^1 \right\rangle \rightarrow A^2[K_C^1][2K_C^1] \rightarrow A[K_C^1][\rho[Q[M^2][L^4]](2 \cdot Q[L^3](2 \cdot M^1)])$$
$$\{Q[M^1[K_C^1]](M^1), K_C^1\}].$$

Most devices are produced from the different components. The task of assembling of the object from *n* components is preceded by "mental" planning of the assembling operation. As it was shown in the previous section of this book assembling object from a number of parts requires the proper selection of parts and planning a sequence of actions. Planning a sequence of actions requires a good visual imagination. The visual imagination is represented by series of mental transformations that can be performed during thinking process. In the previous section, the mental transformation were represented by an operational chain of possible actions and were performed without referring to the categorical chain. In this section solving the problem of assembling of the tool requires knowledge supplied by the categorical chain. The first task given as an example of the visual problem solving is the task of assembling the spade from a given components. SUS understands the concept of spade as a tool that is used by man to do a certain kind of work. SUS understands the construction of the tool and can interpret the perceived object as a spade. SUS understands that the spade belongs to the category of the man-made objects and that it is a real-world object. Based on these categories there is relatively easy to find the conceptual similarities with other categories. For example, a spade, a hammer, and a tongs are objects that are similar in this respect that are all members of the tools category. This type of similarity will be called the conceptual similarity, whereas the visual similarity refers to the visual aspect of objects and describes objects that look similar. For example, visual aspects of the hammer and specific fonts of the letter "T" look similar. The visual similarity is responsible for obtaining the different results of interpretation of the visual object. The concept that is obtained during the visual reasoning supplies the information about parts of the object (spade). Assembling spade from parts requires referring to the spade category given by the following categorical chain: $\langle \Pi \rangle \rhd \langle \sigma_{\mathrm{Pt}} \rangle \rhd \langle \nu_{\mathrm{ReO}} \rangle \rhd \langle \nu_{\mathrm{Ear}} \rangle \rhd \langle \nu_{\mathrm{NLiv}} \rangle \rhd \langle \nu_{\mathrm{MMad}} \rangle \rhd \langle \nu_{\mathrm{Tol}} \rangle \rhd \langle \nu_{\mathrm{TFrm}} \rangle \rhd \langle \nu_{\mathrm{Spd}} \rangle \succ [\tau_{\mathrm{Bld}}, \tau_{\mathrm{Hnd}}, \tau_{\mathrm{TopH}}]$.

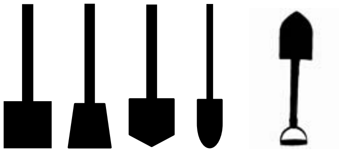
For the tool such as a spade the visual concept consists of visual concepts of each part and is represented by the structural archetype (Fig. 5.119c). The structural archetype is the schematic representative of the visual concept which consists of the visual concepts of each part of the object (tool). The structural archetype shows the main parts of the object represented by the categorical chain of the part: $\langle \nu_{\mathrm{Spd}} \rangle \succ [\tau_{\mathrm{Bld}}, \tau_{\mathrm{Hnd}}, \tau_{\mathrm{TopH}}]$. The category of parts of spade can be used in the visual problem solving during interpretation of parts of a visual schema of the real-world object. The task of assembling the spade from *n* components is preceded by "mental" planning of the assembling operation. During mental planning there is a need to imagine the process of assembling. Mental assembling process is based on the structural archetype. Mental assembling process uses the symbolic names of parts to perform imagery transformations that lead to obtain the required tools.

The spade category that supplies knowledge about components of the spade is used for solving a problem of assembling the spade from the *n* parts. The task is formulated as follows: having parts $u_1, ..., u_n$ make the complex object u. The object u is given by its visual concept φ^{α} so the symbolic name of the complex object u can be one that is given as a member of the visual concept $\eta \in \varphi^{\alpha}$. In the case of a spade we have three parts (phantoms) and the system needs to find the imagery transformation that makes it possible to obtain the object given by the visual concept of a spade. The visual concept refers to one of the exemplars that are defined by the visual concept of a spade (see Fig. 5.120). During thinking process one of the exemplar is selected (in this example an object in Fig 5.120e) and next the imagery transformations that made it possible to obtain this complex object from given parts (Fig. 5.119d–f) are selected. The solution is given as a sequence of imagery transformations that made it possible to "assemble" the spade. These imagery transformations can be translated into actions that robot have to perform in order to finalize this task. The whole object that is to be assembled is given by the symbolic name $\rho[L_M^7]\{L_M^5, \hat{A}[M^1](M^1), 2 \cdot Q[L^4](M^1)\}$ (see Fig. 5.119b) and symbolic



a b c d e f

Fig. 5.119. Assembling spade from three components by utilizing the structural archetype of the spade category



a b c d e

Fig. 5.120. Exemplars generated from the "spade" class

names of parts are as follows: L_M^5 (Fig. 5.119d), $\Theta^2(L^4)$ (Fig. 5.119e), and $\hat{A}[M^1](M^1)$ (Fig. 5.119f). The assembling process is represented as a chain of possible actions $\{(\leftrightarrow)L_M^5 + (\leftrightarrow)\Theta^2(L^4)\} + (\leftrightarrow)\hat{A}[M^1](M^1) -> \rho[L_M^7]\{L_M^5, \hat{A}[M^1](M^1), 2\cdot Q[L^4](M^1)\}$. The perceptual transformations that transform the visual object shown in Fig (5.119a) $\langle \nu_O \rangle \succ \langle \pi_{\text{Sh}} \rangle \ldots \triangleright \langle \nu_{\text{Spd}} \rangle$ into the perceptual category of the silhouette $\langle \nu_O \rangle \succ \langle \pi_{\text{Si}} \rangle \ldots \triangleright \langle \nu_{\text{Spd}} \rangle$ shown in Fig (5.119b) make it possible to trace the assembling process using camera.

The structural archetype that shows the main parts of the object represented by the part of the categorical chain: $\langle \nu_{\text{Spd}} \rangle \succ [\tau_{\text{Bld}}, \tau_{\text{Hnd}}, \tau_{\text{TopH}}]$ is used in the process called the conceptual magnification to check if the part is interpreted correctly. During the conceptual magnification the part that is selected is regarded as the independent object and can be examined as a member of a different category of objects.

Assembling of the complex tool is a complex process consisting of many stages. Each stage is focused on assembling specific part (component) that is next used as an element to build more complex components. To illustrate application of the categorical chain in the process of thinking that lead to design the process of assembling complex device, an example of assembling of the electronic device is given. It is assumed that assembling process begins with reading of the electronic schema. The schema of

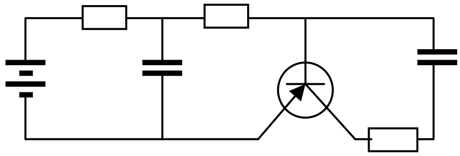


Fig. 5.121. Example of the schema of electronic circuit

electronic circuit is given as a starting point of assembling process. Processes connected with designing of electronic circuit that leads to obtaining schema of electronic circuit are not included in the description of the assembling process. Reading of the schema of electronic circuit is based on the categorical chain that is derived from the structural pattern category. In the first stage of assembling process the specific category of the schema of electronic circuit is identified. The specific category of electronic circuit schema such as the category ECR (the battery, capacitors, resistors), or complex devices such as the radio or the amplifier that are identified in the first stage of assembling process is used during thinking process to design further assembling steps. Example of the specific category of electronic circuit schema is shown in Fig. 5.121. To interpret the electronic schema the knowledge from the category of schema of electronic circuits derived from the category of knowledge is used. The following categorical chain and knowledge chain (described in Chap. 4) are used: $\langle \Pi \rangle \triangleright \langle \sigma_{\mathrm{Pt}} \rangle \triangleright \langle \nu_{\mathrm{Sg}} \rangle \triangleright \langle \nu_{\mathrm{VSym}} \rangle \triangleright \langle \nu_{\mathrm{EngSh}} \rangle \triangleright \langle \nu_{\mathrm{ElESh}} \rangle \triangleright \langle \nu_{\mathrm{ERC}}, \nu_{\mathrm{ERLC}}, \nu_{\mathrm{ERLCT}} \rangle$, $\kappa_{\mathrm{KB}} \triangleright \langle \kappa_{\mathrm{KOb}} \rangle \triangleright \langle \kappa_{\mathrm{EngO}} \rangle \triangleright \langle \kappa_{\mathrm{ElEng}} \rangle \triangleright \langle \kappa_{\mathrm{EngSh}} \rangle \triangleright \langle \kappa_{\mathrm{ElESh}} \rangle \triangleright \langle \kappa_{\mathrm{ERC}}, \kappa_{\mathrm{ERLC}}, \kappa_{\mathrm{ERLCT}} \rangle$. Based on knowledge supplied by the category κ_{ERLCT} derived from the knowledge object the schema shown in Fig. 5.121 is interpreted as composed of objects from the category of structural elements that are members of the category of electronic symbols. The category of electronic engineering symbols derived from the category of visual symbols supplies knowledge that makes it possible to interpret a perceived object as one of the member of the category of electronic symbols such as a resistor or a capacitor represented by the following categorical chain: $\langle \Pi \rangle \triangleright \langle \sigma_{El} \rangle \triangleright \langle \nu_{Sg} \rangle \triangleright \langle \nu_{VSym} \rangle \triangleright \langle \nu_{EngS} \rangle \triangleright \langle \nu_{ElES} \rangle \triangleright \langle \nu_{Rez}, \nu_{Ind}, \nu_{Trn}, \nu_{Cap}, \nu_{Dio}, \nu_{Trz} \rangle$. The knowledge schema supplies the knowledge in the form of the visual concept or a symbol definition that makes it possible to obtain the knowledge that is needed during interpretation of the perceived object. The category of electronic symbols is linked with category of electronic elements that is derived from the category of real-world objects $\langle \Pi \rangle \triangleright \langle \sigma_{El} \rangle \triangleright \langle \nu_{ReO} \rangle \triangleright \langle \nu_{Ear} \rangle \triangleright \langle \nu_{NLiv} \rangle \triangleright \langle \nu_{MMad} \rangle \triangleright \langle \nu_{AsP} \rangle \triangleright \langle \nu_{ElAsP} \rangle \triangleright$

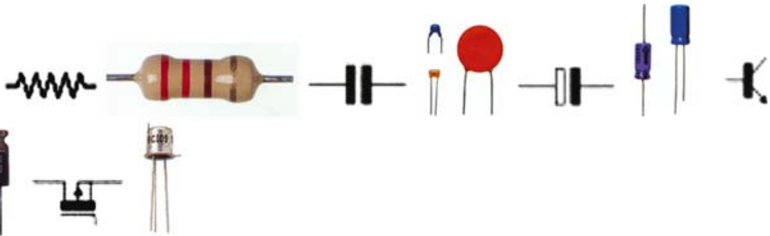


Fig. 5.122. The category of electronic symbols refers to the real-world objects the resistor, the capacitor, the capacitor electrolytic, the bipolar transistor, the field-effect transistor

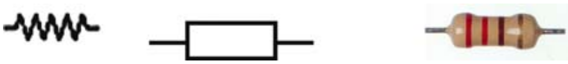


Fig. 5.123. Representatives of the resistor as a electronic symbol and a real-world object

$\langle \nu_{Rez}, \nu_{Ind}, \nu_{Trn}, \nu_{Cap}, \nu_{Dio}, \nu_{Trz} \rangle$. Figure 5.122 shows examples of category of electronic symbols that refers to the real-world objects the resistor, the capacitor, the capacitor electrolytic, the bipolar transistor, the field-effect transistor.

The specific category of electronic symbols such as the resistor $\langle \nu_{Resistor} \rangle$ is derived from the category of electronic symbols $\langle \nu_{ElES} \rangle$. The visual concept of the resistor category includes visual representatives of existing electronic symbols used to represent the resistor. Examples of electronic symbols used to represent the resistor category are given in Fig. 5.123a, b. The category of resistors derived from the category of visual symbols refers to the category of resistors derived from the category of real-world objects. The categorical chain shows the categorical dependence of the resistor category derived from the category of real-world objects $\ldots \triangleright \langle \nu_{MMad} \rangle \triangleright \langle \nu_{AsP} \rangle \triangleright \langle \nu_{ElAsP} \rangle \triangleright \langle \nu_{Res} \rangle$. Based on connection between those two categorical chains the perceived object that is classified as a member of the category of visual symbols (resistor) can be interpreted as the member of the real-world objects. This property of the categorical chain makes it possible to "see" a perceived object both as a visual symbol and the real-world object. Such interpretation enables system to select resistor from other objects based on interpretation of the electronic schema. Selection of the required object such as a resistor is the first step in assembling process. The knowledge concerning the resistor category is given by knowledge categories $\kappa_{KB} \triangleright \ldots \triangleright \langle \nu_{ElES} \rangle \triangleright \langle \nu_{Res} \rangle$. The definition and description of the resistor supplied by the knowledge schema is used to design further assembling steps.

The knowledge schema of the resistor category supplies link to the category of product of the worker called electronic technologist. The knowledge schema from this category supplies knowledge about the tools, material and knowledge needed to produce the electronic element such as a resistor $\kappa_{KB} \triangleright \langle \kappa_{KOb} \rangle \triangleright \langle \kappa_{GenC} \rangle \triangleright \langle \kappa_{WrK} \rangle \triangleright \langle \kappa_{Prf} \rangle \triangleright \langle \kappa_{ElT} \rangle \propto (\kappa_{Tol}, \kappa_{Mat}, \kappa_{Kno}, \kappa_{Res})$. Based on knowledge supplied by the knowledge schema, the resistors needed in assembling process can be produced assuming that all facilities needed for production are available. The knowledge schema of the resistor category supplies link to the category of market product. Based on knowledge supplied by the category of market product the resistor as a member of the market category can be acquired through the market. The knowledge schema of the resistor as the market product supplies knowledge about the market availability, the name of the company that produces the product, the name of sellers companies or a range of prices $\ldots \triangleright \langle \kappa_{ElT} \rangle \propto (\kappa_{Tol}, \kappa_{Mat}, \kappa_{Kno}, \kappa_{Res})$. The category of schema of electronic circuit is linked with the category of circuit boards. The category of circuit boards supplies knowledge that makes it possible to interpret schema of electronic circuit as a set of electronic elements assembled on the circuit board $\langle \Pi \rangle \triangleright \langle \sigma_{Pt} \rangle \triangleright \ldots \triangleright \langle \nu_{MMad} \rangle \triangleright \langle \nu_{CoT} \rangle \triangleright \langle \nu_{ElCoT} \rangle \triangleright \langle \nu_{ElCir}, \nu_{Bor} \rangle$ (see Fig. 5.124). The knowledge that is supplied by the knowledge schema of this category is used to design the circuit board and supplies knowledge that is used during assembling process. The electronic elements (parts of electronic assembly) are used to build the electronic circuit that is part of the electronic devices. Electronic circuits are built by fixing components into a plastic board that has cooper tracks on one side to link them together.

The final product such as a gramophone or a radio required also components that are not part of the category of the electronic circuit.

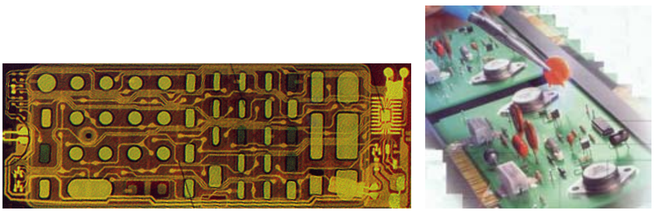


Fig. 5.124. Electronic circuits are built by fixing components into a plastic board that has cooper tracks on one side to link them

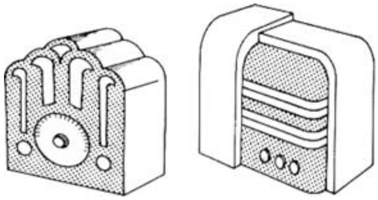


Fig. 5.125. Final product of the assembling of the electronic devices

5.5.2.13. Moving Object Outside

Moving of the object outside an area in which it is closed is an example of another visual task. Let us consider the following problem: moving the object outside the window. The window can be thought of as any windows-like object with holes. The solution to this problem can be expressed in the form of the rule $[\eta \in \varphi^M] \Rightarrow o \triangleright M$, that means if the examined object given by the symbolic name η is a member of the visual concept $[\eta \in \varphi^M]$, the action M can be undertaken. The action M denotes that object can be moved through the window. During the learning stage the visual concept of the solution is learned. The visual concept of the solution is represented by the symbolic names of general classes, such as $A^2[\Lambda](A^1[\Lambda](T), \Lambda)$, where Λ denotes a convex class and T denotes the thin convex class.

Let the window be represented as an object from the cyclic convex class $A^2[L^4](2 \bullet L^4)$, and the object that needs to be moved, as an object from the thin convex rectangular class $\overline{L}^4$ Fig. 5.126a. The general visual concept of the solution is represented by the symbolic name $A^2[\Lambda](A^1[\Lambda](T), \Lambda)$. During the thinking process possible solutions are generated by applying the imagery transformation. Examples of the possible solutions are shown in Fig. 5.126b–e. The symbolic name $A^2[L_R^4](A^1[L_R^4](\overline{L}_R^4), L_R^4)$ (object shown in Fig. 5.126e) is the specific class of the class $A^2[\Lambda](A^1[\Lambda](T), \Lambda)$ and is selected as the solution of this problem.

a b c d e

Fig. 5.126. The visual representation of the problem moving outside the closed area

5.5.2.14. Obstacle Detection and Motion Planning

The motion planning problem is well-known in the field of robotics. It asks for computing feasible paths for a given robot A in a workspace containing some obstacles. Two versions of the problem can be formulated. In one version, a start configuration s and a goal configuration g are given before hand, and the objective is to compute a feasible path for A from s to g. In the second version, no start and goal configurations are specified, and the objective is to compute a data-structure, which can later be used for queries with arbitrary start and goal configurations. We refer to the former case as a single shot problem, and to the latter as a learning problem. The "classical" approaches to motion planning can roughly be divided in the following three classes: roadmap methods, cell decomposition methods, and potential field methods. For a thorough discussion of these approaches see, e.g., [10] and [7].

The paths through which a robot can move is a free space among 3D objects (obstacles, furniture). In the Fig. 5.127a the 3D objects are placed in such a way that the robot (SUS) see it as the rectangular space with three obstacles. Obstacle detection problem is connected with interpretation of the category of objects in the room. From the category of objects in the room the category of the possible paths is derived. The visual concept of the category of the possible paths represents the configuration of objects by showing paths where robot can go. In solving motion planning problem the imagery transformation transform 3D visual objects into the object from the category of the possible paths. The Fig. 5.127 illustrates the imagery transformation of scene of three blocks into objects from the category of possible paths. To find the solution, SUS made visual representation of the "room" based on the sensory information (e.g., a photograph of the room) (Fig. 5.127a). Figure 5.127b–d shows some possible 2D representation of the configuration of the objects in the room that are based on the scene shown in Fig. 5.127a. Each 2D representation is transformed into the symbolic name (b) $A^3[L_R^4]\{3L_R^4\}$, (c) $A^2[L_R^4]\{L_R^4, Q^2[L^6](2L_R^3)\}$ (d) $A^1[L_R^4]\{Q^3[L^6](2L_R^3, Q^1[L^4](L_R^3))\}$. These symbolic names are used

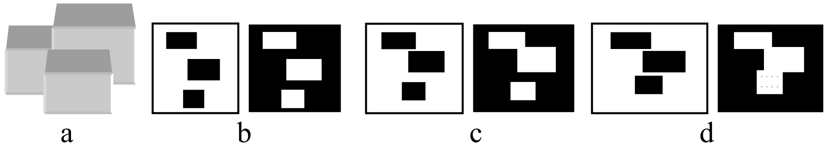


Fig. 5.127. Possible interpretation of the moving area based on the scene consisting of three objects

during learning of the visual concept of the category of the possible paths. In the case when these objects are movable the task can be to move the object to obtain the required configuration of objects. For example, if the beginning configuration was such as in Fig. 5.127d (the robot cannot move between objects) the robot needs to move objects to obtain configuration such as in Fig. 5.127b that has symbolic representation in the form of the cyclic class A^3.

The room in Fig. 5.128 has two movable objects. The room is given as an exemplar from the convex acyclic class $A^2[L_R^4]\{2L_R^4\}$. The situation in Fig. 5.128b is given as the two objects in the room: the one is placed close to the wall so there is no possibility to move between object and the wall, whereas the second one is placed in such a way that there is the possibility to move between these two objects. This room is represented by an exemplar of the concave cyclic class $A^1[Q^1[L^5](L_R^3)\{L_R^4\}$. In Fig. 5.128c two obstacles are placed close to each other so there is no possibility to move among them. This room is represented by the convex cyclic class $A^1[L_R^4]\{Q^2[L^6](2L_R^3)\}$. In Fig. 5.128d two objects are placed near the wall and the room is represented by the concave class $Q^2[L^6](2L_R^3)$. Based on these symbolic names the general rules of the reconfiguration of the objects in order to obtain the required moving path can be learned. For example, if we want to have the configuration in which movement between objects and walls is possible the representative of the room should be a member of the A^2 class.

In the case of more than two objects the symbolic representation is more complex, however, generalization makes it possible to group the similar configurations to show the similarity among them. In Fig. 5.129a the room is represented by the convex cyclic class $A^3[L_R^4]\{3L_R^4\}$. The configuration given in Fig, 5.129b is represented by the convex cyclic class $A^1[L_R^4]\{Q^3[L^7](3L_R^3)\}$.

a b c d

Fig. 5.128. Representatives of the room category (room with two objects)

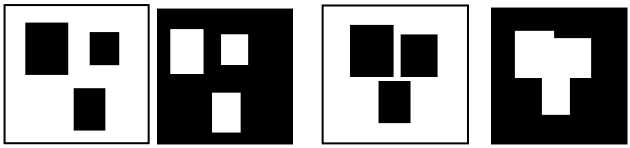


Fig. 5.129. Representatives of the room category (room with three objects).

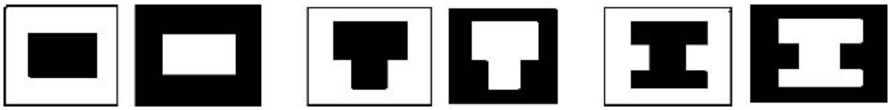


Fig. 5.130. Examples of representatives of the category of room with *n* objects that are placed not close to the wall but close to each other

The generalization of this problem for *n* objects that are placed not close to the wall but close to each other, shown in Fig. 5.130, is represented by the convex acyclic class with one hole A^1. Examples of representatives of the category of room with *n* objects given by symbolic names are shown in Fig. 5.130: (a) $A^1[L_R^4]\{L_R^4\}$, (b) $A^1[L_R^4]\{Q^2[L^5](2L_R^3)\}$, (c) $A^1[L_R^4]\{Q^2[L^4](2L_R^4)\}$.

The obstacle detection is another example of the visual problem. To find the solution, SUS made visual representation of the "room" based on the sensory information (e.g., a photograph of the room) (Fig. 5.131a). The movable objects are marked as dark grey colored areas, the not-movable objects are marked as black colored areas, while the free space is marked as white colored areas. Robot is marked as a light grey colored area. The task is to find the optimal way to go out. At the first stage robot does not differentiate between movable and not movable objects. The visual representation of this configuration is given in Fig. 5.131b and the result of applying the imagery transformation is given in Fig 5.131c. In Fig. 5.131d the representation of the hypothetical situation in which robot is to imagine moving the "moving object" (marked as not visible) and as the result the representation shown in Fig 5.131e is obtained. Translating this representation into the class description means that the object has to be a member of the cyclic class A^1. It means that the new road is found when an object is

a b c d e

Fig. 5.131. Visual transformations used in solving the obstacle detection problem

a b c d e

Fig. 5.132. Visual transformations used in solving the obstacle detection problem

a representative of the cyclic class A^1 (Fig 5.131e). At the beginning, the object is representative of the concave class Q^1 (Fig 5.131c). In the next example, at the beginning, the visual representation of the "room" is given by representative of the cyclic class A^2 (Fig. 5.132c). Next, robot "imagines" moving the black small object that is performed by applying visual reasoning and imagery transformations. Robot moves the moving object and the new road is found. As the result of applying the imagery transformation, the visual representation of the "room" after moving an object is representative of the cyclic class A^3 (Fig. 5.132c).

5.5.2.15. Visual Intelligence Tests

Solving visual problem given in the form of the visual intelligence test requires finding proper representation. The knowledge concerning solution of the visual intelligence tests is represented in the form of the categorical chains as the visual test category. The visual test category is divided into the category of comparison-selection tests, the category of matrix tests or the category of spatial tests. As it was described in Chap. 4 each test is given in the form of a set of visual objects. The first step in solving the visual intelligence test is to assign a set of objects that constitute the visual test, to one of the categories of the visual intelligence test. The category of the visual test is derived from the pattern category that is composed from elements such as figure, signs, or real-world object. The visual test category is divided into the categories of the comparison-selection tests, the category of matrix tests or the category of spatial tests: $\langle \Pi \rangle \triangleright \langle \sigma_{Pt} \rangle \triangleright \langle \nu_{Sg} \rangle \triangleright \langle \nu_{VSym} \rangle \triangleright \langle \nu_{VEduT} \rangle \triangleright \langle \nu_{VisT} \rangle \triangleright \langle \nu_{CST}, \nu_{MtT}, \nu_{SpT} \rangle$. The knowledge category of the visual test is derived from the category of the educational object and is divided into the category of visual test, the category of educational tests or the category of the educational learning theories: $\kappa_{KB} \triangleright \langle \kappa_{KOb} \rangle \triangleright \langle \kappa_{EduO} \rangle \triangleright \langle \kappa_{VisT}, \kappa_{EduT}, \kappa_{EdLT} \rangle$. The knowledge schema supplies knowledge in the form of the name of the test, the definition in the form of the verbal description, the definition in the form of the formal specification and the proposed solution. The formal specification give the description of the test in terms of the stimulus form that is compared with N answer forms.

5.5.2.15.1. Visual Discrimination Test

In the visual discrimination test the stimulus form is compared with N answer forms in order to find one that matches each other. In this test the stimulus form (the form to which all forms are compared) denoted as ν is compared with answer forms denoted as $o_i, i = 1,..N$, where N is a number of forms for comparison. The task is formulated as: Find $o_i, i = 1,..N$ that matches an object $o_i \equiv \nu$: $[\exists o_i : o_i \equiv \nu, \textit{for } i = 1,..N] \Rightarrow o_i \triangleright \sigma$. Examples of different members of the category of visual discrimination tests are shown in Fig. 5.133.

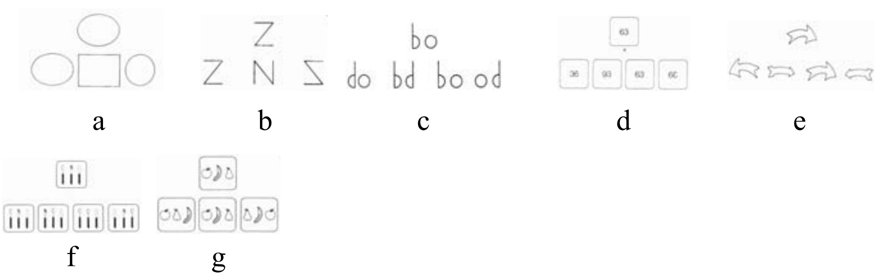


Fig. 5.133. Example of the visual discrimination test

SUS solves the tasks of visual discrimination test during visual thinking process. In visual thinking process each visual object is transformed into the visual concept and next into one of the ontological categories: the figure, the letter, the real-world object, or the sign. The task is next solved by comparison of the stimulus form (the form to which all forms are compared) and all answer forms.

Example of the solution of the task given in the test shown in Fig. 5.133a:

1. Each object $o_i \in O$ is assigned to the class given by the symbolic name: $o_1 \triangleright \eta i, i = 1,...,N$, the object v is assigned to the class given by the symbolic name : $v \triangleright \eta$
2. For all objects v and o_i the interpretation in terms of the ontological categories is found to be $\eta \triangleright C$, $For\ i = 1,...N, o_i \triangleright C_i$
3. All names of categories C_i are compared with the name of C category $For\ i = 1,...N, [C_i \neq C] \Rightarrow C_i \triangleright S_k, k = k+1$
4. if $k \neq 1$ all objects S_k are removed from further comparison and remaining objects are assigned to one of the specific categories, For $i = 1,...k, C_i \triangleright C_i^S$, go to 2, else
5. Object $C_i^S \triangleright \sigma$ is a required solution

The solution of the test can be given in the form of interpretational steps, where in which step the object v and each object $o_i \in O$ are interpreted on one of the specific categorical levels. For example, the tests given in Fig. 5.133b, c are solved by interpretation of each object as a member of the letter category. Objects in Fig. 5.133b are interpreted as the following letters: $L(v) \equiv 'Z'$, $L(o_1) \equiv 'Z'$, $L(o_2) \equiv 'Z_{90}'$, $L(o_3) \equiv 'Z_{mirror}'$. By comparison meaning of each object $L(o_i)$ with meaning of the object $L(v)$ the object o_1 meaning of each is the same as the meaning of the object v is selected as the solution.

Objects in Fig. 5.133c are interpreted as a string of letters: $L(v) \equiv 'bo'$, $L(o_1) \equiv 'do'$, $L(o_2) \equiv 'bd'$, $L(o_3) \equiv 'bo'$, $L(o_4) \equiv 'od'$. The object o_3 which has the same meaning as the object v is selected as the solution. Test given in Fig. 5.133d, e is solved by interpreting objects as a member of the sign category. Phantoms in Fig. 5.133d are interpreted as numbers: $L(v) \equiv '63'$, $L(o_1) \equiv '36'$, $L(o_2) \equiv '93'$, $L(o_3) \equiv '63'$, $L(o_4) \equiv '\mathit{\Phi}'$. The object o_3 which has the same meaning as the object v is selected as the solution. Phantoms in Fig. 5.133e are interpreted as signs: $L(v) \equiv 'arrowD\_r'$, $L(o_1) \equiv 'arrowD\_l'$, $L(o_2) \equiv 'arrowS\_r'$, $L(o_3) \equiv 'arrowD\_r'$, $L(o_4) \equiv 'arrowS\_l'$. The object o_3 which has the same meaning as the object v is selected as the solution. The test in Fig. 5.133f, g can be solved by interpreting phantoms as members of the category of the real-world objects.

5.5.2.15.2. Visual Sequential Memory Test

The category of the visual sequential memory test is similar to the category of the visual discrimination test and the category of the visual memory test. In the test, the category of visual discrimination, test the stimulus form v, and the answer forms $o_i, i = 1,..N$, consist of the one object. In the case of the category of visual sequential memory test the stimulus form is given in the form of the string v_j, $j = 1,..M$, and all answer forms are given in the form $o_i^j, i = 1,..N, j = 1,...M$. Similarly as in the case of the visual discrimination test, the sequence of forms is converted into a string of symbolic names and next the solution is found by string matching. The solution of the task can be reduced to the interpretational steps. Examples of interpretational steps for the test shown in Fig. 5.134a, b are given as follows.

Test in Fig. 5.134a:

1. $v \equiv 'L^3 \Theta^4'$, $o_1 \equiv 'L^3 \Theta^4'$, $o_2 \equiv 'K^1 \Theta^4'$, $o_3 \equiv 'L^4 K^1'$, $o_4 \equiv '\Theta^4 L^4'$
2. $v \equiv o_1$

Test in Fig. 5.134b:

1. $v \equiv 'K^1 L^3 L^4'$, $o_1 \equiv 'K^1 L^4 L^3$, $o_2 \equiv 'K^1 L^3 L^4'$, $o_3 \equiv 'L^3 K^1 L^4'$, $o_4 \equiv 'L^4 L^3 K^1'$
2. $v \equiv o_2$

a b

Fig. 5.134. Examples of the visual sequential memory tests

5.5.2.15.3. Visual Form Constancy Test

In the visual form constancy test the task is formulated as: find $o_i, i=1,..N$ that is similar to an object $o_i \approx v$: $[\exists o_i : o_i \approx v, \; for\; i=1,..N] \Rightarrow o_i \rhd \sigma$ (Fig. 5.135). The visual similarity assumes that objects can be different only as the result of application of the affine transformation such as a rotation or a scaling. Visual form constancy test includes comparison of the objects and selection of the one that is similar to a given one. In contrast to previous tests the visual form constancy tests include forms that are subjected to the geometrical transformations (distortion). Similarly like in the previous cases solution of the task can be reduced to the interpretational steps. To find the solution figures need to be interpreted in the context of the geometrical transformations such as rotation and scaling. The figure can be also part of complex figure. Examples of interpretational steps for the test shown in Fig. 5.135a are given as follows:

1. $v \rhd \Theta^3[L^3]$, $o_1 \rhd \Theta^4[L^4]$, $o_2 \rhd \Theta^3[L^3]$, $o_3 \rhd \Theta^3[L^3]$, $o_4 \rhd \Theta^3[L^3]$, $o_5 \rhd \Theta^3[L^3]$,
2. $v \rhd \Theta^3[L^3_E]$, $o_2 \rhd \Theta^3[L^3_E]$, $o_3 \rhd \Theta^3[L^3_O]$, $o_4 \rhd \Theta^3[L^3_E]$, $o_5 \rhd \Theta^3[L^3_O]$
3. $v \rhd \Theta^3[L^3_E]([l,m,m])$, $o_2 \rhd \Theta^3[L^3_E]([s,m,m])$, $o_4 \rhd \Theta^3[L^3_E]([l,m,m])$
4. $v \equiv o_4$

In order to find the solution the specific classes which are derived from the class $\Theta^3[L^3]$ are used. The derivation process is represented as follows:

$$\Theta^3[L^3] - \{\Theta^3[L^3_O], \Theta^3[L^3_E] - \{\Theta^3[L^3_E]([l,m,m]), \Theta^3[L^3_E]([s,m,m])\}\}.$$

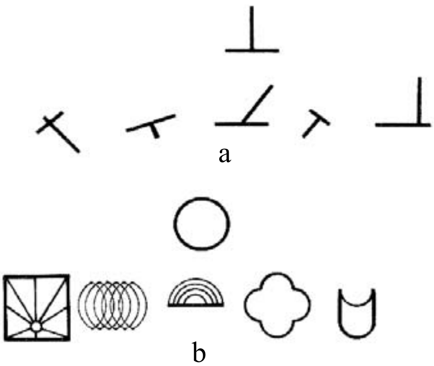


Fig. 5.135. Example of the visual form constancy test

Obtaining the solution can be simplified by representing symbolic names in the form of strings. At each step the object o_i for which symbolic names are different from the symbolic name of the object v are removed and remained objects are assigned to the specific classes. The solution given in the form of strings for the test shown in Fig. 5.135a is as follows.

1. $< T3L3 > (T4L4)(T3L3)(T3L3)(T3L3)(T3L3)$
2. $< T3L3E > ()(T3L3E)(T3L3O)(T3L3E)(T3L3O)$
3. $< T3L3E[lmm] > ()(T3L3E[smm])()(T3L3E[lmm])()$

The solution can be also obtained by interpreting objects as members of the letter category. In that case the object that need to be remembered is interpreted as the letter "T" that is upside down, and the object that is selected as a solution is the rotated version of the letter "T."

5.5.2.15.4. Matrix Test

The progressive matrices test consists of the eight matrix objects and five objects from which one was to be selected as the answer. The category of matrix test consists of eight objects that are placed in the matrix pattern. The task is to find the ninth object in the matrix (selected from the given answer objects) based on the relationship discovered among the eight objects. The category of matrix tests is divided into the category of arithmetical operations (AO) test, the category of geometrical operation (GO) test and the category of finding relationships (FR) test given as: $.. \triangleright \langle v_{VisT} \rangle \triangleright \langle v_{MtT} \rangle \triangleright \langle v_{AO}, v_{GO}, v_{FR} \rangle$.

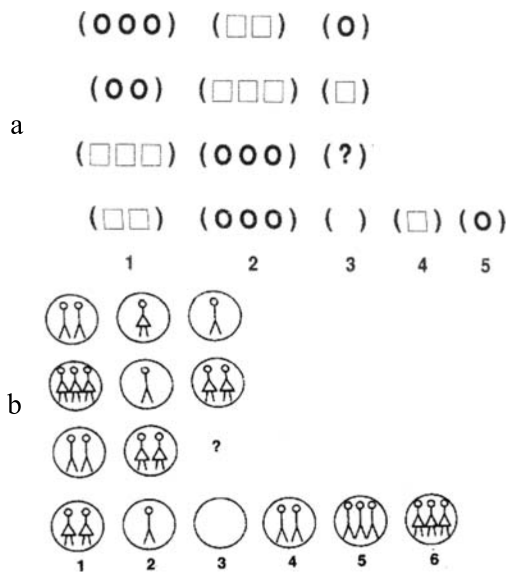


Fig. 5.136. Examples of the test type of arithmetical operations (AO)

The category of AO test consists of eight patterns of the two different types of figures that code the arithmetical operations such as a addition or a subtraction (see Fig. 5.136). Each of eight patterns consisting of two types of different figures. The test can be thought of as a matrix consisting of n-different types of figures. Let us assume that number of types of figures is nor more than two $n = 2$, – figure type A and figure type B. Reduction of a number of types of figures into two does not limit the generalization of the obtained results. In fact, in the tests which are used to test the students the number of types of figures is limited to two. Let us denote the number of figures type A as m_i^j, and the number of figures type B as n_i^j. During solving the test at first, the pattern of figures is to be assigned to the category of the category of AO test, and the number of figures type A m_i^j, and type B n_i^j is computed.

During finding the solution at first each visual object is converted into the symbolic name and next solution is found by applying the special formula of finding the solution of the category of AO test. Two rows are used to find the relationships between numbers of figures type A m_i^j and numbers of figures type B n_i^j and the third row is used to find the solution.

The solution can be computed by applying the following formula: $(m_3^i - n_3^i) = (m_1^i - n_1^i) \otimes (m_2^i - n_2^i)$, where symbol $\otimes$ denotes one of the arithmetical operators (+ or –).

In order to explain the method of solving of the test that is a member of the category of AO test an example shown in Fig. 5.137a is given. The test shown in Fig. 5.137a consist of eight cyclic objects $A_\Lambda^n(n\Lambda)$. At first symbolic names are obtained for each visual object in the test and given in matrix form as follows:

$A^3[L_R^4](3K_C^1)$	$A^2[L_R^4](2L_R^4)$	$A^1[L_R^4](K_C^1)$
$A^1[L_R^4](K_C^1)$	$A^3[L_R^4](3L_R^4)$	$A^2[L_R^4](2L_R^4)$
$A^2[L_R^4](2K_C^1)$	$A[L_R^4](L_R^4)$	?

or in the shortest form

$3K_C^1$	$2L_R^4$	K_C^1
K_C^1	$3L_R^4$	$2L_R^4$
$2K_C^1$	L_R^4	?

.

The solution is expressed in the term of symbolic name of the cyclic class, as $A^1[L_R^4](L^3)$. Similarly the test shown in Fig. 5.137b can be solved.

SUS solve this test by converting the symbolic names into string forms. During solution the test AO given in the form of the eight figures are transformed into the symbolic names and next into the string forms. In this chapter, the string forms of the tests are restricted to the type S given in the form Sn|A|1X|…|iX|…|nX|. The test AO is given in the form:

1. Sn1|A|1X|…|iX|…|n1X
2. Sk1|A|1X|…|iX|…|k1X
3. Sm1|A|1X|…|iX|…|k1X
4. Sn2|A|1X|…|iX|…|n2X

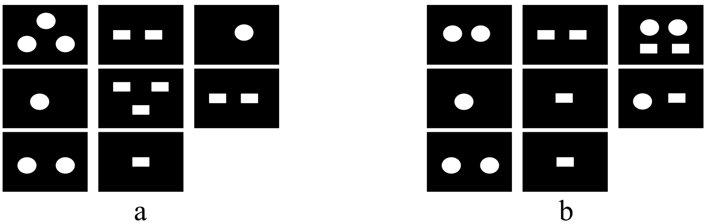


Fig. 5.137. Examples of the tests type AO

5. Sk2|A|1X|…|iX|…|k2X
6. Sm2|A|1X|…|iX|…|k2X
7. Sn3|A|1X|…|iX|…|n3X
8. Sk3|A|1X|…|iX|…|k3X

For the tests shown in Fig. 5.137 that consist of exemplars from the cyclic classes $A_{\Lambda}^{n}(n\Lambda)$ the solution is given as follows:

Fig. 5.137a

1. $A3 \,|\, K1C \,|\, K1C \,|\, K1C \,|$
2. $A2 \,|\, L4R \,|\, L4R \,|$
3. $A1 \,|\, K1C \,|$
4. $A1 \,|\, K1C \,|$
5. $A3 \,|\, L4R \,|\, L4R \,|\, L4R \,|$
6. $A2 \,|\, L4R \,|\, L4R \,|$
7. $A2 \,|\, K1C \,|\, K1C \,|$
8. $A1 \,|\, L4R \,|$

Fig. 5.137b

1. $A2 \,|\, K1C \,|\, K1C \,|$
2. $A2 \,|\, L4R \,|\, L4R \,|$
3. $A4 \,|\, L4R \,|\, L4R \,|\, K1C \,|\, K1C \,|$
4. $A1 \,|\, K1C \,|$
5. $A1 \,|\, L4R \,|$
6. $A2 \,|\, L4R \,|\, K1C \,|$
7. $A2 \,|\, K1C \,|\, K1C \,|$
8. $A1 \,|\, L4R \,|$

After assigning symbols $A3 = 3,\ A2 = 2,\ A1 = 1,\ K1C = K,\ L4R = L$ the following matrices are obtained:

$3KKK$	$2LL$	$1K$
$1K$	$3LLL$	$2LL$
$2KK$	$1L$	?

$2KK$	$2LL$	$4KKLL$
$1K$	$1L$	$2LL$
$2KK$	$1L$	?

The first stage of finding the solution is to check if the test is type $\Im^{AB}$ and then the type of the test symbols $\Theta(\wp_{ij}^{T})$ is computed. The test type AB is given as a set of strings $S_n^j \,|\, A^j \,|\, X_1^j \,|...|\, X_i^j \,|...|\, X_n^j \,|$, where $X_i \in \{\eta_1, \eta_2\}$ denotes the symbolic names of parts of the figures of the tests. The type of the test symbols $\Theta(\wp_{ij}^{T})$ is a number of the different symbolic names $\eta_1, ...\eta_i, ...\eta_m$ in the test $\wp_{ij}^{T}$. The final solution can be computed by applying the following formula: $(m_3^i - n_3^i) = (m_1^i - n_1^i) \otimes (m_2^i - n_2^i)$, where symbol $\otimes$ denotes one of the arithmetical operators (+ or –), described at the beginning of this section.

The category of geometrical addition (GA) tests consists of eight objects. The object in the third column (row) is the result of GO called geometrical addition of the objects in the first and the second column (row). For example, the archetype shown in the first column of the test in Fig. 5.138a $G[L_R^4]\{\otimes^4, L_R^4\}$ is the result of geometrical summation $\otimes^4 \cup L_R^4$ of the second and the third column, or in other words, the archetype shown in the first column is the result of the geometrical sum of the archetype of the convex class L_R^4 (the second column) and the archetype of the thin star class $\otimes^4$ (the third column). The test is written in matrix form as follows:

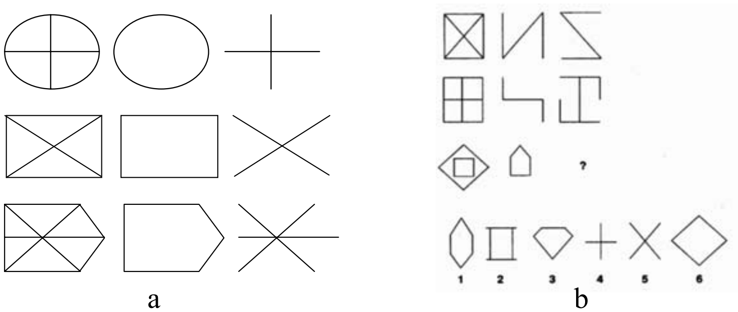
F^1	F^2	F^3
F^4	F^5	F^6
F^7	F^8	?

During finding of the solution at first each visual object is converted into the symbolic name and next solution is obtained by finding the geometrical operator. The geometrical operator makes it possible to perform the GOs on the strings of symbols that represent the visual objects.

In the test in Fig 5.138a two rows are used to find relationships between figures F1 and F6 and figures F7 and F8 in the third row are used to find the possible solution. The relationships can be written as F1*F2 = F3, F4*F5 = F6, F7*F8 = ?, where "*" denotes the geometrical operator. In the tests shown in Fig. 5.138 the figure in the first column is the geometrical sum of figures in the second and the third column. The solution is obtained by removing the line from the figure in the first column by using the figure from the second column. The symbolic name is obtained as $\eta_{3i} = \eta_{1i} * \eta_{1i}$, where $\eta_{1i}, \eta_{2i}, \eta_{3i}$ denotes the symbolic names of figures in the column 1 and the column 2 and the column 3. The archetype of the G class are obtained by performing the line removing operation on the object given by the symbolic name η_{1i} applying the object given by the symbolic name η_{2i}.

The solution is obtained by decomposing the string of the symbolic names as shown in the matrix below. The solution $\otimes^6$ is obtained from strings comparison: $G[K^1]\{\otimes_R^4, K^1\}$ - $\otimes_R^4 - K^1$, $G[L_R^4]\{\otimes^4, L_R^4\}$ - $\otimes^4 - L_R^4$, $G[L_M^5]\{\otimes^6, L_M^5\}$ - $\otimes^6 - L_M^5$. Similarly the test given in Fig. 5.138b can be solved.

$G[K^1]\{\otimes^4_R, K^1\}$	K^1	$\otimes^4_R$
$G[L^4_R]\{\otimes^4, L^4_R\}$	L^4_R	$\otimes^4$
$G[L^5_M]\{\otimes^6, L^5_M\}$	L^5_M	?



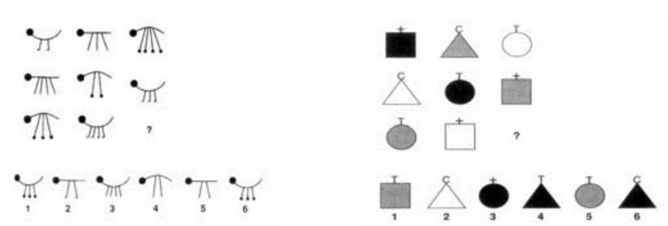
a b

Fig. 5.138. Example of the test type of geometrical addition GA

In the test from the category of FR figures are arranged in such a way that the six figures are used to find the general rules of prediction and two figures are used to find the possible solution. The test needs to have at least two features in common for each three figures to be called solvable. The test pattern can be represented as follows:

F^1_1	F^1_2	F^1_3
F^2_1	F^2_2	F^2_3
F^3_1	F^3_2	F^3_3

There are two configurations used in the test. The simple configuration (CS) is given as three sets consisting of three figures $\left\{F^1_1, F^1_2, F^1_3\right\}$, $\left\{F^2_1, F^2_2, F^2_3\right\}$, $\left\{F^3_1, F^3_2, F(?)\right\}$. The most common configuration (CMC) is given as $\left\{F^1_1, F^3_2, F^2_3\right\}$, $\left\{F^1_3, F^2_2, F^3_1\right\}$, $\left\{F^1_2, F^2_1, F(?)\right\}$ (see Fig. 5.139).



a b

Fig. 5.139. Examples of the test type finding relationships (FR) (CMC configuration)

Each test given in the form of the CMC configuration can be transformed into CS configuration. Examples of tests given in the form of the CMC are shown in Fig. 5.139.

As it was described in the Chap. 4 the category of the test type FR consists of eight visual objects and six objects that represent the possible answers. The solution to the test can be given in the form of interpretational steps. The interpretational steps that lead to the solution can be easy to follow without knowing meaning of the notation of the symbolic names. The visual object represented in the form of symbolic names (e.g., $C(K^1,\Theta(c(h),s(l),b,s(l),c(h)))$) is transformed into string representation (e.g., [K,h,l,b,l,h]). Examples of the interpretational steps for the test shown in Fig. 5.139a are as follows:

$C(K^1,\Theta(c(h),s(l),b,s(l),c(h)))$	[K,h,l,b,l,h]
$C(K^1,\Theta(s,s,b,s,b,s,s))$	[K,s,s,b,s,b,s,s]
$C(K^1,\Theta(c(d),s(k),b,s(k),b,s(k),b,s(k),c(d)))$	[K,d,k,b,k,b,k,b,k,d]
$C(K^1,\Theta(s,s,b,s,b,s,b,s,s)$	[K,s,s,b,s,b,s,b,s,s]
$C(K^1,\Theta(c(d),s(k),b,s(k),c(d)))$	[K,d,k,b,k,d]
$C(K^1,\Theta(c(h),s(l),b,s(l),b,s(l),c(h)))$	[K,h,l,b,l, b,l,h]
$C(K^1,\Theta(c(d),s(k),b,s(k),b,s(k),c(d)))$	[K,d,k,b,k,b,k,d]
$C(K^1,\Theta(c(h),s(l),b,s(l),b,s(l),b,s(l),c(h)))$	[K,h,l,b,l, b,l,b,l,h]
?	

Objects of possible answers represented in the form of symbolic names (e.g., $C(K^1,\Theta(c(h),s(k),b,s(k),b,s(k),c(d)))$) are transformed into string representation (e.g., [K,h,k,b,k,b,k,h]).

1. $C(K^1,\Theta(c(h),s(k),b,s(k),b,s(k),c(d)))$	[K,h,k,b,k,b,k,h]
2. $C(K^1,\Theta(s,s(l),b,s(l),s))$	[K,s,l,b,l,s]
3. $C(K^1,\Theta(c(h),s,b,s,b,s,b,s,c(h)))$	[K,h,s,b,s,b,s,b,s,h]
4. $C(K^1,\Theta(c(d),s,b,s,c(d)))$	[K,d,s,b,s,d]
5. $C(K^1,\Theta(s,s,b,s,s))$	[K,s,s,b,s,s]
6. $C(K^1,\Theta(c(h),s(k),b,s(k),b,s(k),c(d)))$	[K,h,k,b,k,b,k,h]

Test given in the string form:

[K,h,l,b,l,h] [K,s,s,b,s,b,s,s] [K,d,k,b,k,b,k,b,k,d]
[K,s,s,b,s,b,s,b,s,s] [K,d,k,b,k,d] [K,h,l,b,l, b,l,h]
[K,d,k,b,k,b,k,d] [K,h,l,b,l, b,l,b,l,h] ?

Possible answers given in the string form:

1[K,h,k,b,k,b,k,h] 2[K,s,l,b,l,s] 3[K,h,s,b,s,b,s,b,s,h]
4[K,d,s,b,s,d] 5[K,s,s,b,s,s] 6[K,h,k,b,k,b,k,h]

After removing letters K and b we obtain the following form:

[h,l,l,h] [s,s,s,s,s] [d,k,k,k,k,d]
[s,s,s,s,s,s] [d,k,k,d] [h,l,l,l,h]
[d,k,k,k,d] [h,l,l,l,l,h]
1[h,k,k,k,h] 2[s,l,l,s] 3[h,s,s,s,s,h] 4[d,s,s,d] 5[s,s,s,s] 6[h,k,k,k,h]

After removing sequences h,k,..., s,l,..., h,s,…, d,s,…, the following result is obtained 5[s,s,s,s] (the object that represents the answer number 5 is selected).

SUS should be able to solve any FR test, however, the perceptual abilities to see small details seem to be major obstacle in finding the solution for some tests. At this stage of research the test that SUS solves is designed using objects that can be easy to interpret by SUS. The solution of the test type FR is based on the method that utilizes the string representation of the symbolic names. To explain the method examples of test shown in Fig. 5.140 will be used. The objects in the test are denoted using matrix notation $v_{i,j}, i, j = 1,..3$. The five objects that are given as an answer are denoted as $o_k, k = 1,..5$.

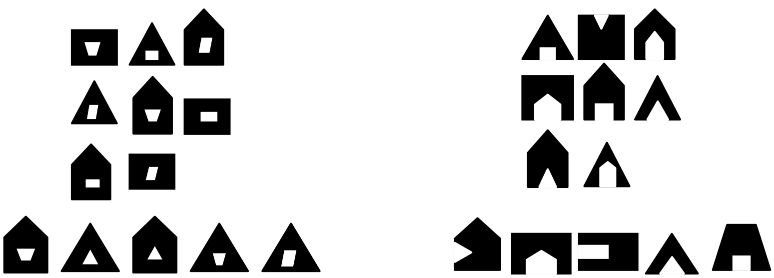


Fig. 5.140. Example of tests type FR and answers

The task is formulated as “find the object $o_k, k = 1,..5$ that fills the missing entry in the matrix $v_{i,j},\ i, j = 1,..3$: $[\exists o_k : o_k \otimes v_{i,j}, \mathit{for}$ $i, j = 1,..3\ k = 1,...,5] \Rightarrow o_k \rhd \sigma$.” The symbol “$\otimes$” denotes matching between the possible solution $v_{i,j}$ and the answer object o_k. The visual test is given as a matrix pattern that consists of eight objects and five object as a possible answer.

During solving the test objects in the matrix are transformed into a series of the test figures $v_i, i = 1,..8$ and for each figure the symbolic name $\eta_i, i = 1,..8$ is obtained. The symbolic name η_i is transformed into the string form $\wp_i^S$. As it was described in Chap. 2 the string form consists of the combination of the selected letters, numbers, and the symbol “|.” The string has a following form: B1|…|Bi|...|Bn|, where Bi denotes the symbolic name of the shape class. The test in the string form is given as follows:

1. A11|A21|….|An1|
2. A12|A22|….|An2|
3. A13|A23|….|An3|
4. A14|A24|….|An4|
5. A15|A25|….|An5|
6. A16|A26|….|An6|
7. A17|A27|….|An7|
8. A18|A28|….|An8|.

The answer string form is given as follows:

1. B11|B21|….|Bn1|
2. B12|B22|….|Bn2|
3. B13|B23|….|Bn3|
4. B14|B24|….|Bn4|
5. B15|B25|….|Bn5|.

The symbolic name η_i, $i = 1,..8$ that is obtained during visual reasoning is transformed into the basic form. The basic form includes symbols that refer to the symbolic names (a general level of description). For example, the string Q<L^{4}4>[R]|<L3>[A]<L3>[A]| that refers to the symbolic name $Q^2[L_R^4](2L^3)$ is transformed into the form Q1|L4_R|L3_A|L3_A. Translation of the symbolic name into a string form requires to include all details given by the symbolic name. The level of details is marked by introducing the symbol "_." The symbolic name is translated into the form L0_L1_...Ln, where the level Ln denotes the level of description of the archetype of a given class. The test that is converted into the string form needs to preserve the level of details. The *n*th level of details can be written in the string form as follows:

$$jA_1^0 \ldots \_\ jA_1^H \mid \ldots \mid jA_i^0 \ldots \_\ jA_i^H \mid \ldots \mid jA_1^0 \ldots \_\ jA_n^H \mid .$$

Example of the test for one and two levels of details is given in Fig. 5.141. The test (given in the string form) shown in Fig. 5.141a is represented by the string of symbols shown in the left column whereas the test given in Fig. 5.141b is represented by the string of symbols shown in the right column:

A_1|L4_R|L5_T|	A_1_*|M_1_L4R|K_1_C|
A_1|L3_O|L4_R|	A_1_*|K_1_C |L_3_A|
A_1|L5_Y|L3_O|	A_1_*|L_4_R |M_1_L3A|
A_1|L3_A|L3_A|	A_1_*|K_1_E |M_1_L4R|
A_1|L5_M|L5_M|	A_1_*|L_3_A |K_1_C|
A_1|L4_R|L4_Y|	A_1_*|M_1_L3A|L_3_A|
A_1|L5_O|L4_T|	A_1_*|L_4_T |L_3_A|
A_1|L4_T|L3_A	A_1_*|M_2_L4R|M_2_L4R|

Very often in order to find the solution there is a need performing generalization. During generalization the symbol is dropped from the left to the right. For example, for the test in Fig. 5.141a the first level of generalization (the first row) is A|L4|L5|. For the test in Fig. 5.141b the first level of generalization (the first row) is A_1|M_1|K_1| and the second level of generalization is A|M|K|.

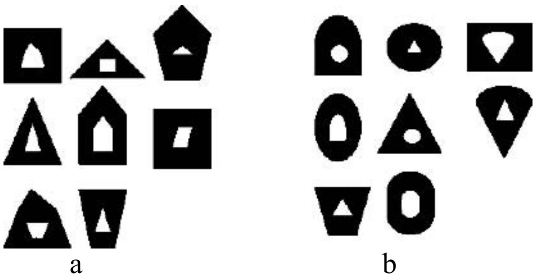


Fig. 5.141. Example of the test type FR (**a**) one level of details (**b**) two levels of details

The test given as objects in the matrix pattern does not always have the solution. The test FR to be solvable (test have a solution) should fulfill the following conditions: the test needs to have at least two features in common for each three objects in the matrix pattern. The objects are arranged in such a way, that six figures are used to find the general rules of prediction and two figures are used to find the possible solution.

The test pattern can be represented as follows:

F_1^1	F_2^1	F_3^1
F_1^2	F_2^2	F_3^2
F_1^3	F_2^3	F_3^3

There are two configurations used in the test. The CS is given as three sets consisting of three figures $\{F_1^1, F_2^1, F_3^1\}$, $\{F_1^2, F_2^2, F_3^2\}$, $\{F_1^3, F_2^3, F(?)\}$. The CMC is given as $\{F_1^1, F_2^2, F_3^2\}$, $\{F_3^1, F_2^2, F_1^3\}$, $\{F_2^1, F_1^2, F(?)\}$. Each test, given in the form of the CMC configuration, can be transformed into CS configuration. Example of the transformation is given in Fig. 5.142.

Test is solved by selecting features for both configurations figures $\{F_1^1, F_2^1, F_3^1\}$, and $\{F_1^2, F_2^2, F_3^2\}$. The feature is any symbol in the string representation …|X1|…|Y1|.... The relationship can be formulated in the form of one, two, or more than two features. Figure 5.143 shows a test that has relationships expressed in the form of one feature (Fig. 5.143a) and two features (Fig. 5.143b). These tests are represented in the string form as follows:

```
A1|M2L4R|M1L4R |     Q1|M1L4R|Q1|L4R|M1 |
A1|M1L4R|M2L4R |     Q1|L4R  |Q1|L4R|L4R|
A1|M1L3A|M1L3A |     Q1|L5M  |Q1|L4T|L3A|
A1|M1L4R|M1L3A |     Q1|L4R  |Q1|L4T|L3A|
A1|M1L3A|M1L4R |     Q1|L5M  |Q1|L4R|M1|
A1|M2L4R|M2L4R |     Q1|M1L4R|Q1|L4R|L4R|
A1|M1L3A|M2L4R |     Q1|L5M  |Q1|L4R|L4R|
A1|M2L4R|M1L3A       Q1|M1L4R|Q1|L4T|L3A|
```

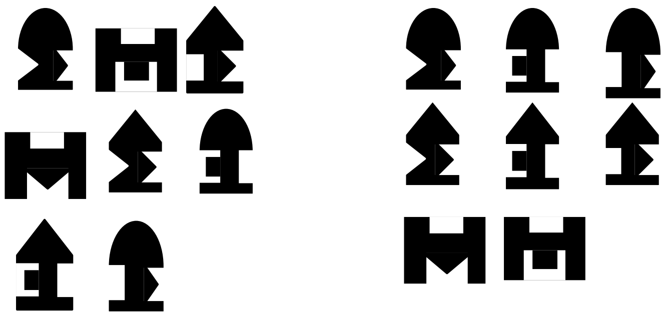
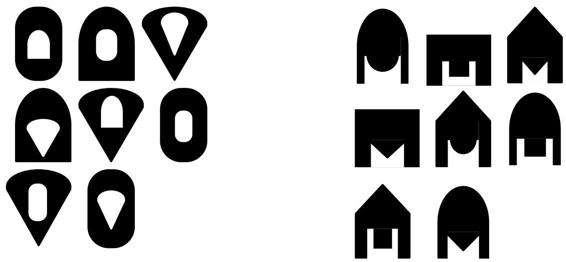


Fig. 5.142. Transformation of the test given in the CMC form into the CS form



a b

Fig. 5.143. Tests that have relationships expressed in the form of one feature (**a**) and two features (**b**)

The test which has solution can be written as follows:

1. …|X1|…|Y1|...		1. X1|Y1
2. …|X2|…|Y2|...		2. X2|Y2
3. …|X3|…|Y3|...		3. X3|Y3
4. …|X2|…|Y3|...		4. X2|Y3
5. …|X3|…|Y1|...	or in short	5. X3|Y1
6. …|X1|…|Y2|...		6. X1|Y2
7. …|X3|…|Y2|...		7. X3|Y2
8. …|X2|…|Y3|...		8. X2|Y3

The test can be given in the compatible form that means each eight rows has the columns that are representative of the same type of classes. The test given in the compatible form can be represented as follows:

1A1|1A2|…|1Ai|…|1An|
2A1|2A2|…|2Ai|…|2An|
3A1|3A2|…|3Ai|…|3An|
4A1|4A2|…|4Ai|…|4An|

5A1|5A2|…|5Ai|…|5An|
6A1|6A2|…|6Ai|…|6An|
7A1|7A2|…|7Ai|…|7An|
8A1|8A2|…|8Ai|…|8An|

As it was mention in previous section, the test is formulated as a pattern of the eight strings. For simplicity the type of compatible test can be written in the form of the one string. Examples of the different one string representations of the different tests are as follows:

Sn|A|1X|…|iX|…|nX|,
S1|A|1_S1|1_A|1_X|,
Sn|A|1X|2X|…|hX|1_1Q1|1_1G|1_R|…|1_hQ1|1_hG|1_hR|.

The incompatible test can be represented as follows:

1A1|1A2|…|1Ai|…|1An1|
2A1|2A2|…|2Ai|…|2An2|
3A1|3A2|…|3Ai|…|3An3|
4A1|4A2|…|4Ai|…|4An4|
5A1|5A2|…|5Ai|…|5An5|
6A1|6A2|…|6Ai|…|6An6|
7A1|7A2|…|7Ai|…|7An7|
8A1|8A2|…|8Ai|…|8An8|

In the case of the incompatible test the first step is to bring it into the compatible form. The simple form of the incompatible test is given as follows:

1. S1|A1|1X1|
2. S2|A2|1X2|2X2|
3. S3|A3|1X3|2X3|3X3|
4. S4|A4|1X3|2X3|3X3|
5. S5|A5|1X1|
6. S6|A6|1X2|2X2|
7. S7|A7|1X2|2X2|
8. S8|A8|1X1|

The incompatible test can be reduced into three strings. These strings for the test given in this example are given as follows:

1. S1|A1|1X1|
2. S2|A2|1X2|2X2|
3. S3|A3|1X3|2X3|3X3|

The transformation from the incompatible test into the compatible test requires obtaining the same type of test by fulfilling the “incompatible” columns with symbol “*.” After transformation the compatible form for the test given in this example is as follows:

```
1. S1|A1|1X1|*   |*   |
2. S2|A2|1X2|2X2|*   |
3. S3|A3|1X3|2X3|3X3|
```

The transformation from the incompatible form into the compatible form involves both type of the class and the symbolic name. At first the type of the class for each row is compared to find if it can be transformed into the compatible form. If the incompatible form can be transformed into the compatible form at first the types of classes for each row are matched with the general type of the class. The general type of class is the type that makes it possible to fit the structure of all strings. The general type can be generated or stored as a template. For example, the test Sn can contain the string type Qn|G|1R|…|iR|…|nR| or An|C|1W|…|iW|…|nW| or both types of strings. To convert string from incompatible form into the compatible form the algorithm A1 can be used. For the type of class Snj|A|1X|…|iX|…|nX| the algorithm is as follows:

```
Algorithm A1:
For j=1 to 8
begin
if nj<n2 than
begin
for nj+1 to n2
begin
jX='*'
end
end
end.
```

Example of the test type S_n where the algorithm A1 can be applied:

```
|Q2|L4R|L3A|L4R|*  |
|Q1|L3A|L4R|*  |*  |
|Q2|L5M|L3A|L4T|*  |
|Q3|L3A|L3A|L4T|L5M|
|Q1|L5M|*  |*  |*  |
|Q2|L4R|L3A|L4R|*  |
|Q3|L5M|L3A|L4R|L5M|
|Q2|L4R|L3A|L3A|*  |
```

Similarly the procedure can be applied for the test type S11_S given as:

$$S_n \,|\, A \,|\, 1\_S^1_{m1} \,|\, 1\_A^1 \,|\, 1\_X^1_1 \,|\, ... \,|\, 1\_X^1_{m1} \,|\, ... \,|\, 1\_S^n_{mn} \,|\, 1\_A^n \,|\, 1\_X^n_1 \,|\, ... \,|\, 1\_X^1_{mn} \,| \,.$$

When the test is given in the compatible form the algorithm A2 is applied to find the solution and transforms it into the string form.

Algorithm A2:
begin
s1=0, s2=0
for i=0 to n
begin
$if\ A_{i0}^{S1} = 5\ than$
begin
$\wp_{s1}^{SA} = \wp_i^R$, *s1=s1+1*
end
$if\ A_{i0}^{S2} = 5\ than$
begin
$\wp_{s2}^{SB} = \wp_i^R$, *s2=s2+1*
end
for i=0 to m_N
begin
$a = C_{i0},\ b = C_{i1}$
$if\ A_{i1}^{S1} = 5\ than$
begin
$\wp_{s1}^{SA} = \wp_a^{R1} \oplus \wp_b^{R1}$, *s1=s1+1*
end
$if\ A_{i1}^{S2} = 5\ than$
begin
$\wp_{s2}^{SB} = \wp_a^{R1} \oplus \wp_b^{R1}$, *s2=s2+1*
end
end
end
end.

These algorithms will be described in detail in the second part of this book that will present implementation issues of SUS.

5.5.2.15.5. Category of Spatial Tests

The category of spatial test is derived from the category of the visual psychological test and is divided into the folding sheet test category, the cubic box test category, the bird view test category or the spatial transformation test category: $.. \triangleright \langle\ v_{VisT} \rangle \triangleright \langle\ v_{SpT} \rangle \triangleright \langle\ v_{FST}, v_{CBT}, v_{BWT}, v_{BWT}, v_{SST} \rangle$.

One of the simplest categories of tests is the folding sheet test category. The task is to find which of four 3D figures can be produced by folding a flat sheet of specified shape. SUS solves this task by interpreting the 2D visual object as a member of the category of 3D geometrical figures and next comparing visual concepts of the category flat sheet of 3D figures and all “unfolded” forms of the 3D geometrical objects. The 3D geometrical figures are geometrical solids members of the line drawing perceptual category.

In order to solve spatial test that is a member of the folding sheet test category there is a need to learn the visual concept of the 3D figure and its folding sheet version. The first part of the string $<\rho[L6]\{3L4\}>$ that represents the object shown in Fig. 5.144A denotes the 3D figure (solid), whereas the second part of the string $[\rho[Q[L8](4L3R)]\{6L4R\}]$ denotes its “unfolded” sheet. Examples of learned visual concepts for objects shown in Fig. 5.144A–D are as follows:

A. $<\rho[L6]\{3L4\}>\leftrightarrow[\rho[Q[L8](4L3R)]\{6L4R\}]$
B. $<\rho[L5]\{L3,L4\}>\leftrightarrow[\rho[Q[L6](4L3O)]\{4L4R,2L3E\}]$
C. $<\rho[L6]\{2L4O,L4T\}>\leftrightarrow[\rho[Q[L8](4L3O)]\{4L4R,2L4T\}]$
D. $<\rho[L6]\{2L4O,L5M\}>\leftrightarrow[\rho[Q[L11](5L3)]\{5L4R,2L5M\}]$

During solving task shown in Fig. 5.144 at first the symbolic name for the “unfolded” sheet is obtained $[\rho[Q[L6](4L3O)]\{4L4R,2L3E\}]$ and next all visual concepts are searched in order to find this symbolic name. In our example strings A–D are searched in order to find the symbolic name that matches parts of strings that denote the “unfolded” sheet. As the result the solid shown in Fig. 5.144B is selected as the final solution.

For the tests shown in Fig. 5.145 the task is given in the form of cubic boxes and the nets of these boxes (unfolded cut-outs). The task is to choose which one of the nets belongs to the box shown.

In order to solve the space relations test, there is a need to find the interpretation of the tasks in terms of their nets (unfolded cut-outs). The solution of this task depends on the complexity of the pattern that is given on the visible side of the box.

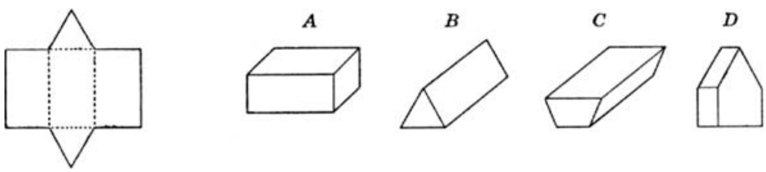


Fig. 5.144. Example of the spatial test member of the folding sheet test category

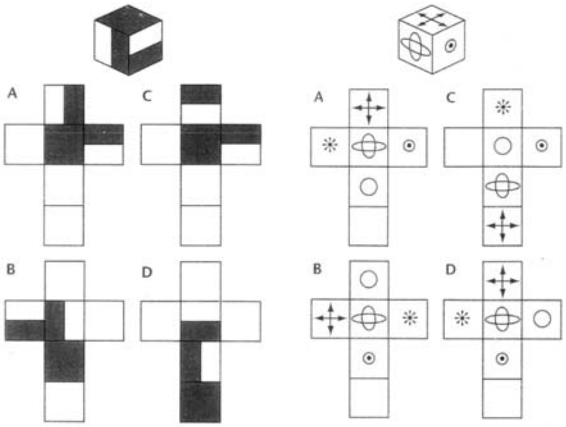


Fig. 5.145. Tests in the form of cubic boxes, and the nets of these boxes (unfolded cut-outs)

In the case when there are different patterns on each side of the cube there is the need to find the rotating scheme of the net. An example of the cube with the letters L, P, and C on the sides is shown in Fig. 5.146.

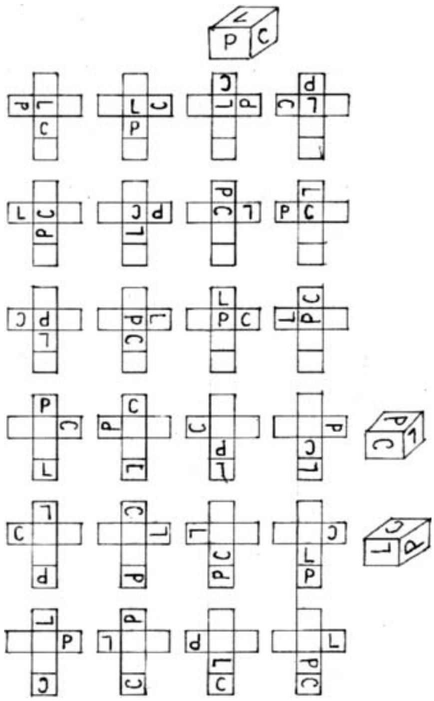


Fig. 5.146. Nets with a pattern that can form the required cube

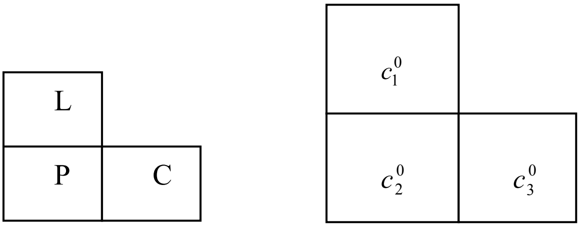


Fig. 5.147. Patterns on the box sides and their code

The pattern shown on the cube is transformed into the pattern on the nets consisting of the three sides. At first the type of the 3D transformation is found (orthographic or perspective projection) and next each pattern on each side is transformed into the 2D pattern. It is assumed that the side on the top (with letter L in Fig. 5.147) is denoted as c_1, the side in the front of viewer (with letter P in Fig. 5.147) is denoted as c_2, and the third side (with letter C in Fig. 5.147) is denoted as c_3.

The notation c_i^0, c_i^{90}, c_i^{180}, and c_i^{270} denotes that pattern on the *i*th side is rotated 0°, 90°, 180°, or 270°. The configuration shown in Fig. 5.147 is called the basic configuration. It is assumed that the basic configuration has pattern that is not rotated. The configuration of the nets with the different pattern is given as:

$$\Im = \begin{Bmatrix} \{c_2^{270} c_3^0 00 c_1^{270} 0\} \{0 c_2^0 c_3^{90} 0 c_1^0 0\} \{00 c_2^{90} c_3^{180} c_1^{90} 0\} \{c_3^{270} 00 c_2^{180} c_1^{180} 0\} \\ \{c_1^0 c_2^{90} 00 c_3^{90} 0\} \{0 c_1^{90} c_2^{180} 0 c_3^{180} 0\} \{00 c_1^{180} c_2^{270} c_3^{270} 0\} \{c_2^0 00 c_1^{270} c_3^0 0\} \\ \{c_3^{180} c_1^{180} 00 c_2^{180} 0\} \{0 c_3^{270} c_1^{270} 00 c_2^{270} 0\} \{00 c_3^0 c_1^0 0 c_2^0 0\} \{c_1^{90} 00 c_3^{90} c_2^{90} 0\} \\ \{00 c_3^{270} c_2^0 0 c_1^0\} \{c_2^{90} 00 c_3^0 0 c_1^{270}\} \{c_3^{90} c_2^{180} 000 c_1^{180}\} \{0 c_3^{180} c_2^{270} 00 c_1^{90}\} \\ \{c_3^0 00 c_1^{180} 0 c_2^{180}\} \{00 c_1^{90} c_3^{270} 0 c_2^{270}\} \{c_1^{270} c_3^{90} 00 c_2^0\} \{0 c_1^0 c_3^{180} 00 c_2^0\} \\ \{00 c_2^0 c_1^{90} 0 c_3^{180}\} \{c_1^{180} 00 c_2^{90} 0 c_3^{90}\} \{c_2^{180} c_1^{270} 000 c_3^0\} \{0 c_2^{270} c_1^0 00 c_3^{270}\} \end{Bmatrix}$$

The algorithm for finding the solution is as follows:

1. Find the type of the 3D transformation of the cube (orthographic or perspective projection)
2. Transform pattern on each side into the 2D pattern
3. Find basic configurations of patterns c_1^0, c_2^0, c_3^0
4. Find the configuration net by inserting the pattern into the coding schema

$$\Im = \left\{ \begin{array}{l} \{c_2^{270} c_3^{0} 00 c_1^{270} 0\} \; \{0 c_2^{0} c_3^{90} 0 c_1^{0} 0\} \; \{00 c_2^{90} c_3^{180} c_1^{90} 0\} \; \{c_3^{270} 00 c_2^{180} c_1^{180} 0\} \\ \{c_1^{0} c_2^{90} 00 c_3^{90} 0\} \; \{0 c_1^{90} c_2^{180} 0 c_3^{180} 0\} \; \{00 c_1^{180} c_2^{270} c_3^{270} 0\} \; \{c_2^{0} 00 c_1^{270} c_3^{0} 0\} \\ \{c_3^{180} c_1^{180} 00 c_2^{180} 0\} \; \{0 c_3^{270} c_1^{270} 00 c_2^{270} 0\} \; \{00 c_3^{0} c_1^{0} 0 c_2^{0} 0\} \; \{c_1^{90} 00 c_3^{90} c_2^{90} 0\} \\ \{00 c_3^{270} c_2^{0} 0 c_1^{0}\} \; \{c_2^{90} 00 c_3^{0} 0 c_1^{270}\} \; \{c_3^{90} c_2^{180} 000 c_1^{180}\} \; \{0 c_3^{180} c_2^{270} 00 c_1^{90}\} \\ \{c_3^{0} 00 c_1^{180} 0 c_2^{180}\} \; \{00 c_1^{90} c_3^{270} 0 c_2^{270}\} \; \{c_1^{270} c_3^{90} 00 c_2^{0}\} \; \{0 c_1^{0} c_3^{180} 00 c_2^{0}\} \\ \{00 c_2^{0} c_1^{90} 0 c_3^{180}\} \; \{c_1^{180} 00 c_2^{90} 0 c_3^{90}\} \; \{c_2^{180} c_1^{270} 000 c_3^{0}\} \; \{0 c_2^{270} c_1^{0} 00 c_3^{270}\} \end{array} \right\}$$

5. Find the pattern coding of the net i
   a. For each pattern $N_i = \{k_1^i k_2^i k_3^i k_4^i k_5^i k_6^i\}$
   b. For each net patter k_j^i , $j = 1...6$ if $k_j^i \notin \{P_1, P_2, P_3)$ then $k_j^i = 0$ else $k_j^i = P_k$
6. For each pattern $N_i = \{k_1^i k_2^i k_3^i k_4^i k_5^i k_6^i\}$ if $n_i \in \Im$ then $N_i = n$ STOP

The issues conected with the implementation of the algorithms presented in this book will be described in the second part of this book.

References

1. Bertoloine, G.R., Wiebe, E.N., Miller, G.L., and Nasman, L.O., *Fundamentals of Graphics Communication*. 2005, Boston: McGraw-Hill
2. Wickelgren, W.A., *How to Solve Mathematical Problems*. 1974, New York: Dover
3. Kowalski, R., *Logic for Problem Solving.* 1979, Oxford: North-Holland
4. Mitsumatsu, H., and Yokosawa, K., How do the internal details of the object contribute to recognition? *Perception*, 2002. **31**: pp. 1289–1298
5. Wit, T.C.J., and Lier, J.R., Global visual completion of quasi-regular shapes. *Perception*, 2002. **31**: pp. 969–984
6. Arnheim, R., *Visual Thinking*. 1970, London: Faber and Faber
7. Yamamoto, N., Singular points of plane curves generated by period doubling bifurcation points and a method for computing them. *International Journal of Bifurcation and Chaos*, 1991. **1**(4): pp. 795–802
8. Su, M.-C., and Liu, Y.-Ch., A new approach to clustering data with arbitrary shapes. *Pattern Recognition*, 2005. **38**: pp. 1887–1901
9. Carbonell, J.G., Derivational analogy: a theory of reconstructive problem solving and expertise acquisition, in *Machine Learning. An Artificial Intelligence Approach*, R.S. Michalski, Carbonell, J.G., and Mitchell, T.M., Editors. 1986, Los Altos: Morgan Kaufman, pp. 371–392

10. Tadeusiewicz, R., and Ogiela, M.R., *Medical Image Understanding Technology*. 2004, Berlin Heidelberg New York: Springer
11. Sammut, C., and Banerji, R.B., Learning concepts by asking questions, in *Machine Learning. An Artificial Intelligence Approach.*, R.S. Michalski, Carbonell, J.G., and Mitchell, T.M., Editors. 1986, Los Altos: Morgan Kaufman, pp. 167–214