

I The Logic of Practical Reason

Practical reason, we are usually told, is reasoning about what to do, and theoretical reason is reasoning about what to believe. But if this is so, it ought to seem puzzling to us that we do not have a generally accepted account of the deductive logical structure of practical reason in a way that we apparently do for deductive theoretical reason. After all, the processes by which we figure out how to best achieve our goals seem to be just as rational as the processes by which we figure out the implications of our beliefs, so why do we seem to have such a powerful logic for the one and not for the other? Aristotle more or less invented the theoretical syllogism and, though generally it has been less influential, he also invented the practical syllogism. Why is there no accepted theory of the practical syllogism in the way there is an accepted theory of the theoretical syllogism and a theory of deductive logic generally?

To see what the problem is, let us review how it is apparently solved for theoretical reason. We need to

distinguish questions of logical relations from questions of philosophical psychology. Great advances in deductive logic were made when, in the nineteenth century, Frege separated questions of philosophical psychology (the “laws of thought”) from those of logical relations. After Frege, it has seemed that if you get the logical relations right the philosophical psychology should be relatively easy. For example, once we understand the relations of logical consequence between propositions then many of the corresponding questions about belief seem fairly simple. If I know that the premises “all men are mortal” and “Socrates is a man” jointly entail the conclusion “Socrates is mortal” then I already know that someone who *believes* those premises is *committed* to that conclusion; that someone who *knows* the premises to be true is *justified* in *inferring* the truth of the conclusion, etc. There seems in short to be a fairly tight set of parallels within theoretical reason between such “logical” notions as premise, conclusion, and logical consequence on the one hand and such “psychological” notions as belief, commitment, and inference on the other. The reason for this tight set of parallels is that the psychological states have propositional contents and therefore inherit certain features of the logical relations between the propositions. Because logical consequence is truth-preserving, and belief is a commitment to truth, the features of logical consequence can be mapped onto the commitments of belief. If q is a logical consequence of p , and I believe p , then I am committed to the truth of q . The tacit principle that has worked so well in assertoric logic is that if you get the logical relations right, then most of the philosophical psychology will take care of itself.

Now, supposing we accept this distinction between the logical relations and the philosophical psychology, how is it supposed to work for practical reason? What are the logical relations in practical reason and how do they bear on the philosophical psychology? Some of the questions about logical relations would be: What is the formal logical structure of practical argument? In particular, can we get a definition of formal validity for practical reason in the way that we can for deductive "theoretical" reason? Does practical logic exhibit the same or does it require different rules of inference than assertoric logic? The questions about the philosophical psychology of deliberation would concern many of the issues that we have been discussing in this book, especially the character of the intentional states in practical reasoning, their relation to the logical structure of deliberation, their relations to action, and their relations to reasons for action generally. What sorts of intentional states figure in deliberation and what are the relations between them? What sorts of things can be reasons for action? What is the nature of motivation, and how does deliberation actually motivate action?

In light of our distinction between logical theory and philosophical psychology, the question we are asking is, "Are there formal patterns of practical validity, such that the *acceptance* of the premises of a valid practical argument commits one to the acceptance of the conclusion, in the way that is characteristic of theoretical reason?" We have seen that in theoretical reason belief in the premises of a valid argument commits you to a belief in the conclusion. Could we get similar commitments to desires and intentions as conclusions in practical reason? The aim of a formal logic of practical reason, it seems to me, would

have to be to get a set of valid forms of practical inference; and a test for any such a project would be whether the agent who accepted the premises of a putatively valid practical inference would be committed to desiring or intending the conclusion, in a way that the agent who accepts the premises of a valid theoretical inference is committed to believing the conclusion.

II Three Patterns of Practical Reason

To begin, let us consider some attempts to state a formal logical structure of practical reason. I will confine the discussion to so-called means-ends reasoning, since most authors on the subject are in the tradition of the Classical Model and think that practical reason is a matter of deliberating about means to achieve ends. Oddly enough it is not at all easy or uncontroversial to state the formal structure of means-ends reasoning, and there is no general agreement on what it is. In the philosophical literature there is a bewildering variety of formal models of such reasoning, and even fundamental disagreements over what its special elements are supposed to be—are they desires, intentions, fiats, imperatives, norms, *noemata*, actions, or what?¹ I think the reason for this variety is that the authors in question are coping with the fact that the elements in reasoning are factitives, and factitives can come in different forms. Many philosophers speak rather glibly about the belief-desire model of explanation and deliberation, but what exactly is the structure of this

1. For a good survey of the literature up to the mid-1970s, see Bruce Aune, *Reason and Action*, Dordrecht-Holland: D. Reidel Publishing Company, 1977, ch. 4, pp. 144–194.

model supposed to be? Anthony Kenny suggests that the structure of practical reason is quite different from theoretical reason. He gives the following example:

I'm to be in London at 4.15

If I catch the 2.30 I will be in London at 4.15

So I'll catch the 2.30.²

Because the premises exhibit both directions of fit, we can represent the form of the argument generally with the following symbolism, using "↑" and "↓" for the upward and downward directions of fit respectively, and using "E" and "M" for ends and means:

↑ (E).

↓ (If M then E).

Therefore, ↑ (M).

In the case where one has beliefs and desires as "premises" this pattern of inference can be represented as follows:

DES (I achieve E).

BEL (If I do M I will achieve E).

Therefore, DES (I do M).

But it seems this could not be right, because two premises of this form simply do not commit one to accepting the conclusion. You do not get a commitment to a desire, much less an intention, as the conclusion of this form of argument. To see this note that a lot of the Es one can think of are quite trivial and many Ms are ridiculous. For

2. Anthony Kenny, *Will, Freedom and Power*, New York: Barnes and Noble, 1976, p. 70.

instance I want this subway to be less crowded and I believe that if I kill all the other passengers it will be less crowded. This does not commit me to desiring to kill the other passengers. Of course one *might* form a homicidal desire on a crowded subway, but it seems absurd to claim that rationality *commits* me to a desire to kill just on the basis of my other beliefs and desires. The most that this pattern could account for would be *possible* motivations for forming a desire. Someone who has the appropriate beliefs and desires has a possible motive for desiring M. But there is no *commitment* to such a desire.

It is sometimes said that this pattern fails because there is no entailment relationship between the propositional contents of the premises and the conclusion. Indeed, if we just look at the propositional contents, the inference is guilty of the fallacy of affirming the consequent. Some philosophers think the standard form of practical reason is to be found in cases where the means is a necessary condition of achieving the end. Thus they endorse the following (or variations on it):

↑ (I achieve end E).

↓ (The only way to achieve E is by means M) (sometimes stated as "M is a necessary condition of E", or "to achieve E, I must do M").

Therefore, ↑ (I do M).

In this case the satisfaction of the premises guarantees the satisfaction of the conclusion, but the acceptance of the premises still does not commit one to a desire or intention in the conclusion. If you think about this pattern in terms of real life examples it seems quite out of the question as a general account of practical reason. In general there are

lots of means, many of them ridiculous, to achieve any end; and in the rare case where there is only one means, it may be so absurd as to be out of the question altogether. Suppose that you have any end you care to name: you want to go to Paris, become rich, or marry a Republican. Well, in the Paris case, for example, there are lots of ways to go. You could walk, swim, take a plane, ship, kayak, or rocket; you could tunnel through the earth or go via the moon or the North Pole. In very rare cases there may be only one means to an end. As far as I know there is no quick way to get rid of flu symptoms short of death. Therefore, on the above model, if I desire to get rid of my flu symptoms immediately, and I believe the only way to do it is death, I am committed to desiring my death. This model, like the first one, has very little application. Most means-ends reasoning is not about necessary conditions, and even when it is, desiring the end does not commit me to desiring the means.³

In the first of these examples there was no entailment relation between the propositional contents of the premises and the conclusion; but in the second there was. The fact that entailment relations do not generate a commitment to a secondary desire reveals an important contrast between the logic of beliefs alone and the logic of belief-desire combinations. If I believe both p and (if p then q), then I am committed to the belief that q . But if I want p and believe that (if p then q), I am not committed to wanting q . Now why is there this difference? When we understand that, we will go a long way toward understanding why there is no plausible logic of practical reason.

3. Aune, *Reason and Action*, who sees that the first model is inadequate for reasons similar to those I have suggested, nonetheless fails to see that the same sorts of objections seem to apply to the second model.

Let's try again to construct a formal logical model of practical reason. Generally when you have a desire, intention, or goal you seek not just *any* means; nor do you search for the *only* means; you seek the *best* means (as Aristotle says you seek the "best or easiest" means). And if you are rational, when there isn't any good or at least reasonable means you give up on the goal altogether. Furthermore, you don't just have a goal, but if you are rational, you appraise and select your own goals in the light of—well, what? We will have to come back to this point later. In the meantime let us suppose you have seriously selected a goal and appraised it as reasonable. Suppose you seriously want to go to Paris, that is, you have "made up your mind," and you try to figure the best way to get there and conclude that the best way is to go by plane. Is there a plausible formal model of the logic of means-ends reasoning for such a case?

The form of the argument seems to be:

Des (I go to Paris).

Bel (the best way, all things considered, is to go by plane).

Therefore, Des (I go by plane).

If we separate the questions of logical relations from the questions of philosophical psychology—as I have been urging—we see that from a logical point of view this argument, as it stands, is enthymematic. To be formally valid it would require an extra premise of the form:

Des (If I go to Paris I go by the best way, all things considered).

If we add this premise, the argument is valid by the standards of classical logic. Let P = I go to Paris, Q = I go by the best way, and R = I go by plane. Then its form is:

$$\begin{array}{l} P \\ P \rightarrow Q \\ Q \leftrightarrow R \\ \hline \therefore R \end{array}$$

And though the argument is not truth-preserving because two of its premises and its conclusion don't have truth values, this doesn't really matter since the argument is satisfaction-preserving, and truth is just a special case of satisfaction. Truth is satisfaction of representations with the word-to-world direction of fit.

But once again, as in the earlier examples, it seems the logical relations don't map onto the philosophical psychology in the right way. It is by no means obvious that a rational person who has all those premises must have, or be committed to having, a desire to go by plane. Furthermore, to make it plausible, we had to introduce a fishy-sounding premise, about wanting to do things "by the best way all things considered." Indeed it looks as if any attempt to state formally the structure of a practical argument of this sort would in general require such a premise, but it is not at all clear what it means. What is meant by "the best way," and what is meant by "all things considered"? Notice furthermore that such premises have no analogue in standard cases of theoretical reason. When one reasons from one's belief that all men are mortal and that Socrates is a man to the conclusion that Socrates is mortal, one does not need any premise about what is the best thing to believe, all things considered.

I have tried to make a sympathetic attempt to find a formal logical model of the traditional conception of means-ends reasoning, the conception that goes back to Aristotle, and this is the best that I can come up with. I

have also tried to give a statement of its formal structure that seems to me an improvement on other versions I have seen. But I think it is still hopelessly inadequate. After various unsuccessful tries I have reluctantly come to the conclusion that it is impossible to get a formal logic of practical reason that is adequate to the facts of the philosophical psychology. To show why this is so, I now turn to the discussion of the nature of desire. The essential feature of desire for the present discussion is that it has the upward direction of fit. Many of the features that I will specify as features of desire are also features of other factitives with the upward direction of fit such as obligations, needs, commitments, etc. However, for the sake of simplicity I will state most of the discussion in terms of desire and generalize it to other upward direction-of-fit factitives later.

III The Structure of Desire

In order to understand the weaknesses in my revised logic for practical reasoning, and in order to understand the general obstacles to a formal logic of practical reasoning, we have to explore some general features of desire and especially explore the differences between desires and beliefs. I will use the general account of intentionality that I gave in chapter 2, as well as other features of the theory of intentionality that I presented in the book of that name.⁴ Specifically, I am going to assume that contrary to the surface grammar of sentences about desire, all desires have whole propositions as intentional contents (thus “I want your car” means something like “I want that I have

4. John R. Searle, *Intentionality: An Essay in the Philosophy of Mind*, Cambridge: Cambridge University Press, 1983.

your car"); that desires have the world-to-mind direction of fit, whereas beliefs have the mind-to-world direction of fit; and that desires do not have the restrictions on intentional contents that intentions have. Intentions must be about future or present actions of the agent and must have causal self-referentiality built into their intentional content. Desires have no such causal condition, and they can be about anything, past, present, or future. Furthermore, I am going to assume that the usual accounts of the *de re/de dicto* distinction are hopelessly muddled, as is the view that desires are intensional-with-an-s. The *de re/de dicto* distinction is properly construed as a distinction between kinds of *sentences* about desires, not between kinds of desires. The claim that all desires, beliefs, etc. are in general intensional is just false. *Sentences about* desires, beliefs, etc. are in general intensional. Desires and beliefs themselves are not in general intensional, though in a few oddball cases they can be.⁵

Where a state of affairs is desired in order to satisfy some other desire, it is best to remember that each desire is part of a larger desire. If I want to go to my office to get my mail, there is indeed a desire the content of which is simply: I want that (I go to my office). But it is part of a larger desire whose content is: I want that (I get my mail by way of going to my office). This feature is shared by intentions. If I intend to do *a* in order to do *b*, then I have a complex intention whose form is I intend (I do *b* by means of doing *a*). I will say more about this point later.

The first feature about desiring (wanting, wishing, etc.), in which it differs from belief is that it is possible for an agent consistently and knowingly to want that *p* and want

5. For a discussion of these points about intensionality-with-an-s and the *de re/de dicto* distinction, see Searle, *Intentionality*, chaps. 7 and 8.

that not p in a way that it is not possible for him consistently and knowingly to believe that p and believe that not p . And this claim is stronger than the claim that an agent can consistently have desires that are impossible of simultaneous satisfaction because of features he doesn't know about. For example, Oedipus can want to marry a woman under the description "my fiancée" and want not to marry any woman under the description "my mother" even though in fact one woman satisfies both descriptions. But I am claiming that he can consistently both want to marry Jocasta and want not to marry Jocasta, under the same description. The standard cases of this are cases where he has certain reasons for wanting to marry her and reasons for not wanting to. For example, he might want to marry her—because, say, he finds her beautiful and intelligent, and simultaneously not want to marry her—because, say, she snores and cracks her knuckles. Such cases are common, but it is also important to point out that a person might find the same features simultaneously desirable and undesirable. He might find her beauty and intelligence exasperating as well as attractive, and he might find her snoring and knuckle-cracking habits endearing as well as repulsive. (Imagine that he thinks to himself: "It is wonderful that she is so beautiful and intelligent, but at the same time it is a bit tiresome; her sitting there being beautiful and intelligent all day long. And it is exasperating to hear her snoring and cracking her knuckles, but at the same time there is something endearing about it. It is so human.") Such is the human condition.

The possibility of rationally and consistently held inconsistent desires has the unpleasant logical consequence that desire is not closed under conjunction. Thus if I desire that p and desire that not p , it does not follow that I desire

that (p and not p). For example I want to be in Berkeley right now and I want to be in Paris right now but knowing that these are inconsistent desires, it is not the case that I am rationally committed to wanting: (I am simultaneously in Berkeley and Paris right now).

In order to understand the possibility of rationally and consistently held inconsistent desires and its consequences for practical reason, we need to probe a bit deeper. It is customary, and I think largely correct to distinguish, as the Classical conception does, between primary and secondary or derived desires. It is literally true to say to my travel agent, "I want to buy a plane ticket." But I have no lust, yearning, yen, or passion for plane tickets—they are just "means" to "ends." A desire that is primary relative to one desire may be secondary relative to another. My desire to go to Paris is primary relative to my desire to buy a plane ticket, secondary relative to my desire to visit the Louvre. The primary/secondary desire distinction will then always be relative to some structure whereby a desire is motivated by another desire or some other motivator. This is precisely the picture that is incorporated in the classical conception of practical reason. In such cases, as I just noted, the complete specification of the secondary desire makes reference to the primary desire. I don't just want to buy a ticket, I want to buy a ticket in order to go to Paris.

Once we understand the character of secondary desires we can see that there are at least two ways in which fully rational agents can form conflicting desires. First, as noted earlier, an agent can simply have conflicting inclinations. But second, he can form conflicting desires from consistent sets of primary desires together with beliefs about the best means of satisfying them. Consider the example of

the man who reasons that he wants to go to Paris by plane. Such a man has a secondary desire to go by plane, motivated by a desire to go to Paris, together with a belief that the best way to go is by plane. But the same man might have constructed a practical inference as follows: I don't want to do anything that makes me nauseated and terrified, but going anywhere by plane makes me nauseated and terrified, therefore I don't want to go anywhere by plane, therefore I don't want to go to Paris by plane. It is easy enough to state this according to the pattern of practical reasoning I suggested above: all things considered the best way to satisfy my desire to avoid nausea and terror is *not* to go to Paris by plane. Since this can be stated as a piece of practical reasoning, it seems that *the same person, using two independent chains of practical reason, can rationally form inconsistent secondary desires from a consistent set of his actual beliefs and a consistent set of primary desires.* A consistent set of "premises" will generate inconsistent secondary desires as "conclusions." This is not a paradoxical or incidental feature of reasoning from beliefs and desires; rather, it is a consequence of certain essential differences between practical and theoretical reason.

Let's probe these differences further: in general it is impossible to have any set of desires, even a consistent set of primary desires, without having, or at least being rationally motivated to having, inconsistent desires. Or, to put this point a bit more precisely: if you take the set of a person's desires and beliefs at any given point in his life, and work out what secondary desires can be rationally motivated from his primary desires, assuming the truth of his beliefs, you will find inconsistent desires. I don't know how to demonstrate this, but any number of examples can be used to illustrate it. Consider the example of going to

Paris by plane. Even if planes do not make me nauseated and terrified, still I don't want to spend the money; I don't want to sit in airplanes; I don't want to eat airplane food; I don't want to stand in line at airports; I don't want to sit next to people who put their elbow where I am trying to put my elbow. And indeed, I don't want to do a whole host of other things that are the price, both literally and figuratively, of satisfying my desire to go to Paris by plane. The same line of reasoning that can lead me to form a desire to go to Paris by plane can also lead me to form a desire *not* to go to Paris by plane.

A possible answer to this, implicit in at least some of the literature, is to invoke the notion of preference. I prefer going to Paris by plane and being uncomfortable to not going to Paris by plane and being comfortable. But this answer, though acceptable as far as it goes, mistakenly implies that the preferences are given *prior* to practical reasoning, whereas, it seems to me, they are typically the product of practical reasoning. And since ordered preferences are typically products of practical reason, they cannot be treated as its universal presupposition. Just as it is a mistake to suppose that a rational person must have a consistent set of desires, so it is a mistake to suppose that rational persons must have a rank ordering of (combinations of) their desires prior to deliberation.

This points to the following conclusion: even if we confine our discussion of practical reasoning to means-ends cases, it turns out that practical reason essentially involves the adjudication of conflicting desires and other sorts of conflicting motivations (i.e., factitives with upward direction-of-fit) in a way that theoretical reason does not essentially involve the adjudication of conflicting beliefs. Practical reasoning is typically about adjudicating between

conflicting desires, obligations, commitments, needs, requirements, and duties. That is why in our attempt to give a plausible account of the Classical conception of practical inference we needed a step about wanting to go by "the best way, all things considered." Such a step is characteristic of any rational reconstruction of a process of means-ends reason, because "best" just means the one that best reconciles all of the conflicting desires and other motivators that bear on the case. However, this also has the consequence that the formalization of the classical conception I gave is essentially a trivialization of the problem, because the hard part has not been analyzed: how do we arrive at the conclusion that such and such is "the best way to do something all things considered" and how do we reconcile the inconsistent conclusions of competing sets of such valid derivations?

If all one had to go on were the Classical conception of reasoning about means to ends, then in order to reach a conclusion of the argument that could form the basis of action one would have to go through a whole set of other such chains of inference and then find some way to settle the issue between the conflicting reasons. *The Classical conception works on the correct principle that any means to a desirable end is desirable at least to the extent that it does lead to the end. But the problem is that in real life any means may be and generally will be undesirable on all sorts of other grounds, and the model has no way of showing how these conflicts are adjudicated.*

The matter is immediately seen to be worse when we consider another feature of desires, which we already noticed in passing. A person who believes that p and that (if p then q) is committed to the truth of q ; but a person who desires that p and believes that (if p then q) is not

committed to desiring that q .⁶ You can want that p and believe that (if p then q) without being committed to wanting that q . For example, there is nothing *logically* wrong with a couple who want to have sexual intercourse and who believe that if they do she will get pregnant but who do not want her to get pregnant.

We can summarize these points about desire and the distinction between desire and belief as follows: desires have two special features that make it impossible to have a formal logic of practical reason parallel to our supposed formal logic of theoretical reason. The first feature we might label "the necessity of inconsistency." Any rational being in real life is bound to have inconsistent desires and other sorts of motivators. The second we might label "the nondetachability of desire." Sets of beliefs and desires as "premises" do not necessarily commit the agent to having a corresponding desire as "conclusion" even in cases where the propositional contents of the premises entail the propositional content of the conclusion. These two theses together go a long way to account for the fact that there is in the philosophical literature no remotely plausible account of a deductive logical structure of practical reason.

The moral is: as near as I can tell, the search for a formal deductive logical structure of practical reason is misguided. Such models either have little or no application, or, if they are fixed up to apply to real life, it can only be by trivializing the essential feature of practical

6. Of course you are not committed to a belief in the sense that you must actually have *formed* the belief that q . You might believe that p and that (if p then q) without having thought any more about it. (Someone might believe that 29 is an odd number and that it is not evenly divisible by 3, 5, 7, or 9 and that any number satisfying these conditions is prime, without ever having actually drawn the conclusion, i.e., formed the belief, that 29 is prime.)

deliberation: the reconciliation of conflicting desires and conflicting reasons for action generally and the formation of rational desires on the basis of the reconciliation. We can always construct a deductive model of any piece of reasoning; but where an essential feature of the reasoning contains both p and not p —as in I want that p and I want that not p , or I am under obligation to bring it about that p and under an obligation to bring it about that not p —deductive logic is unilluminating, because it cannot cope with such inconsistencies. The models either have to pretend that the inconsistencies do not exist or they have to pretend that they have been resolved (“by the best way all things considered”). The first route is taken by the models I criticized at the beginning, the second route is taken by my revised version. The possibility, indeed the inevitability, of contradictory desires, obligations, needs, etc. renders the Classical conception unilluminating as a model of the structure of deliberation. Furthermore even if you do fudge to the extent of trivializing the problem you still do not get a commitment to a desire as the conclusion of the argument. Modus ponens simply doesn’t work for desire/belief combinations to produce a commitment to desiring the conclusion.

Does modus ponens work for desire/desire combinations? This is not the standard subject matter of means-ends reasoning, but it is worth considering the question. It seems to me that if you want that p and want that if p then q , you are committed to wanting that q , but you may still rationally also want that not q . Thus I might want for me to be very rich, and as a matter of public policy I want the very rich to be very heavily taxed, and logically speaking this commits me to the desire that if I become rich I should be very heavily taxed. I am indeed committed to such a

desire, but at the same time I do not want for me to be very heavily taxed. Thus I have a commitment to a desire that is inconsistent with another desire I also have.

IV Explanation of the Differences between Desire and Belief

Now why should there be these differences? What is it about the philosophical psychology of desire that makes it logically so unlike belief? Well, any answer to that has to be tautological, and so disappointing, but here goes anyhow.

Both desires and beliefs have propositional contents, both have a direction of fit, both represent their conditions of satisfaction, and both represent their conditions of satisfaction under certain aspects. So, what is the difference that accounts for the different logical properties of desires and beliefs? The difference derives from two related features, the difference in direction of fit and the difference in commitment. The job of beliefs is to represent how things are (downward direction of fit) and the holder of a belief is committed to its truth. To the extent that the belief does this or fails to do it, it will be true or false respectively. The job of desires is not to represent how things are, but how we would like them to be. And desires can succeed in representing how we would like things to be even if things don't turn out to be the way we would like them to be. In the case of belief, the propositional content represents a certain state of affairs as actually existing. But in the case of desire, the propositional content does not function to represent an actual state of affairs, but rather a *desired* state of affairs, which may be actual, nonexistent, possible, impossible, or what have you. And the propositional

content represents the state of affairs under the aspects that the agent finds desirable. There is nothing wrong with unsatisfied desires, *qua* desires, whereas there is something wrong with unsatisfied beliefs, *qua* beliefs, namely, they are false. They fail in their job of representing how things are. Desires succeed in their job of representing how we would like things to be even in cases where things are not the way we would like them to be, that is, even in cases where their conditions of success are not met. Roughly speaking, when my belief is false, it is the belief that is at fault. When my desire is unsatisfied, it is the world that is at fault.

The two logical features of desire, inconsistency and nondetachability, both derive from this underlying feature of desire: desires are inclinations toward states of affairs (possible, actual, or impossible) under aspects. There is no necessary irrationality involved in the fact that one can be inclined and disinclined to the same state of affairs under the same aspect; and the fact that one is inclined to a state of affairs under an aspect together with knowledge about the consequences of the existence of that state of affairs does not guarantee that, if rational, one will be inclined to those consequences.

But if you try to state parallel points about belief it doesn't work. Beliefs are convictions that states of affairs exist under aspects. But one cannot rationally be convinced both that a state of affairs exists and does not exist under the same aspect. And the fact that one is convinced of the existence of a state of affairs under an aspect together with knowledge about the consequences of the existence of that state of affairs does guarantee that, if one is rational, one will be convinced of (or at least committed to) those consequences. It is important to emphasize that

these features of belief follow from two of its characteristics: downward direction of fit and commitment. If you just have downward direction of fit that is not enough. Thus hypotheses that one may form about how things might be also have the downward direction of fit. But one can consistently and rationally entertain inconsistent hypotheses in a way that one cannot consistently and rationally hold inconsistent beliefs, and this is because beliefs, unlike hypotheses, though both involve the downward direction of fit, have the additional feature of commitment.

These features of desire are characteristic of other sorts of representations with the world-to-word direction of fit. The features of inconsistency and nondetachability apply to needs and obligations as well as desires. I can consistently have inconsistent needs and obligations and I do not necessarily need the consequences of my needs, nor am I obligated to achieve the consequences of my obligations. Examples of all these phenomena are not hard to find: I might need to take some medicine to alleviate one set of symptoms, but I need to avoid that medicine because it aggravates another set of symptoms. I have an obligation to meet my class at the university, but I also have an obligation to give a lecture at another university, because I promised to do so a year earlier. I need to take aspirin to avoid heart ailment, but aspirin upsets my stomach, and so I need to avoid aspirin. Jones has an obligation to marry Smith because she made a promise, but marrying Smith will make her parents unhappy, and she does not have an obligation to make her parents unhappy. It is amazing, by the way, how much the referential opacity of all of these concepts, "obligation," "need," etc., is neglected in the literature.

In objecting to this account, one might say, "Look, when I believe something, what I believe is that it is true. So, if I believe something and know that it can't be true unless something else is true, then my belief and knowledge must commit me to the truth of that other thing as well. But now why isn't it the same for desire? When I want something what I want is that something should happen or be the case, but if I know that it can't happen or be the case unless something else happens or is the case then surely I must be committed to wanting that something else." But the analogy breaks down. If I want to drill your tooth to fill your cavity and I know that drilling the tooth will cause pain it simply does not follow that I am in any way committed to causing pain, much less committed to wanting to cause pain. And the proof of this distinction is quite simple: if I fail to cause pain one of my beliefs is thereby false, but none of my desires is thereby unsatisfied.

When I want something, I want it only under certain aspects. "Yes, but when I believe something I believe it only under certain aspects as well. Sentences about belief are just as opaque as sentences about desire." Yes, but there is this difference: when something is desired under certain aspects it is, in general, the aspects that make it desirable. Indeed the relation between the aspects and the reasons for desiring are quite different from the case of belief, since *the specification of the reasons for desiring something is, in general, already a specification of the content of the desire*; but the specification of the evidence on the basis of which I hold a belief is not in general itself part of the specification of the belief. The reasons for believing stand in a different relation to the propositions believed than the contents of reasons for wanting do to the proposition that is the content of the desire, because in general the state-

ments of the reasons for wanting state part of what one wants. If one wants something for a reason then that reason is part of the content of one's desire. For example, if I want it to rain in order to make my garden grow, then I both want that it should rain and that my garden should grow. If I believe it will rain and I believe that the rain will make my garden grow, then I both believe that it will rain and that my garden will grow. But there is still a crucial difference. If I want it to rain *in order to* make my garden grow, then my reason for wanting it to rain is part of the whole content of the entire complex desire. My reason for believing both that it will rain and that the rain will make my garden grow, on the other hand, has to do with a lot of evidence about meteorology, the reliability of weather predictions, and the function of moisture in producing plant growth. All of these considerations count as evidence for the truth of my belief, but they are not themselves the content of that very belief. But in the case of my desire, the role of reasons is not at all like that of evidence, for the reasons state the aspects under which the phenomenon in question is desired. The reasons, in short, are part of the content of the complex desire.

In sum: beliefs have the mind-to-world direction of fit, and the holder of the belief is committed to the fit actually existing, that is, he is committed to the truth of the belief. Desires have the world-to-mind direction of fit, and the holder of a desire need not be committed to its ever being satisfied. The job of desire is not to represent how things are, but how we would like them to be. It is the notion of "the commitment to how things are" that blocks the simple possibility of consciously held contradictory beliefs, and that requires a commitment to the consequences of one's beliefs, but there is no such block and no such requirement when it is a question of how we would like

things to be. In spite of certain formal similarities, belief is really radically unlike desire in both its logical and its phenomenological features.

For these reasons, it is misleading to think of theoretical reason as reasoning about what to believe in the way that we think of practical reason as reasoning about what to do. What one should believe is dependent on what is the case. Theoretical reasoning, therefore, is only derivatively about what to believe. It is primarily about what is the case—what must be the case given certain premises. Furthermore, we can now see that it is misleading to think even that there is a “logic” of theoretical reason. There is just logic—which deals with logical relations between, for example, propositions. Logic tells us more about the rational structure of theoretical reason than it does about the rational structure of practical reason, because there is a close connection between the rational constraints on belief and the logical relations between propositions. This connection derives from the fact that, to repeat, beliefs are meant to be true. But there is no such close connection between the structure of desire and the structure of logic. Because of the upward direction of fit of desires, I both can and do have conflicting desires even after all the facts are in.

V Some Special Features of Intentions

I have been concentrating on desires, but intentions are in important respects different from desires. Like desires, intentions have the upward direction of fit, but unlike desires, they are always about the agent as subject matter and they are causally self-referential. My intention is carried out only if I act by way of carrying out the intention. For this reason intentions have a logical constraint

quite unlike desire. It is logically inconsistent to have inconsistent intentions in a way that it is not logically inconsistent to have inconsistent desires. Intentions are designed to cause actions, and for that reason they cannot function if they are inconsistent. This prohibition against inconsistency is shared by other causally self-referential motivators, such as orders and promises, even though they also have the world-to-mind direction of fit. It's okay—up to a point—for a speaker to say reflectively "I both wish you would go and wish you would stay." But he is irrational if he says simultaneously "Go!" and "Stay!" and you are equally irrational if you form the simultaneous intentions to go and to stay, or make simultaneous promises both to go and to stay. One cannot consistently have inconsistent intentions or make inconsistent promises and issue inconsistent orders, because intentions, orders, and promises are designed to cause actions, and there cannot be inconsistent actions. For the same reason intentions, orders, and promises commit the agent to the belief that the action is possible, but it is not possible to carry out both of two inconsistent actions. Desires and obligations in general have no such condition. One can hold inconsistent desires and be under inconsistent obligations.

Does this feature give us the possibility of a principle of detachment for intentions? If I intend that p and I believe that if p then q , am I committed to intending that q ? I think not; however, the question is trickier than it might appear at first sight, and because it ties in with Kant's famous principle, I now turn to a discussion of Kant.

VI "He Who Wills the End Wills the Means"

No discussion of the logic of practical reason would be complete without at least some mention of Kant's famous

doctrine that he who wills the end wills the means. Does this give us a deductive logical principle of practical reason? That is to say, does the statement "I will end *E*" logically commit me to "I will means *M*" at least in cases where *M* is a necessary condition of achieving *E*? Is it analogous to the way "I believe *p*" commits me to "I believe *q*," in cases where *q* is a logical consequence of *p*?

Well, it all depends on what we mean by "will." On a perfectly natural interpretation the doctrine is just false, for reasons that I have stated earlier. If willing is a matter of having a very strong desire or pro-attitude toward some future course of action that I am capable of engaging in, then it is simply not the case that when I will the end I am logically committed to willing the means. As I suggested earlier, it may be the case that the means are out of the question for one reason or another. I very much want to eliminate my flu symptoms, but the only way to eliminate the symptoms is to commit suicide, there being no known cure, but all the same, I am not committed to willing suicide.

So if we interpret "will" as desire, Kant's principle comes out false. But suppose we interpret it as intention, both prior intention and intention-in-action. Suppose I have a prior intention to do *E* and I believe that doing *M* is a necessary condition of doing *E*. Am I committed to the intention to do *M*? It seems to me we need to distinguish between having a commitment to doing something that I know will involve doing *M* and having a commitment to doing *M* intentionally. Trivially it follows from the fact that I intend to do *E* and I know that doing *E* necessarily involves doing *M* that I have a commitment to doing something intentionally that will involve *M*. But I need not thereby have any commitment at all to doing *M*

intentionally. Thus consider our earlier example of my intention to fix your tooth. We have as premises:

Intend (I fix your tooth).

Bel (If I fix your tooth I cause you pain).

But I am not thereby committed to the conclusion

Intend (I cause you pain).

An intention commits me to a course of action, but it does not commit me to doing all of the things that I know are involved in carrying out the original intention. So the fact that I have an intention to bring it about that p and I have a belief that if p then q does not commit me to having the intention to bring it about that q . The argument for this claim, using the above example, is that when I cause you pain, I do not do so intentionally, but only as a by-product of my intentional action. And the argument for that point, in turn, is that causing you pain is not part of the conditions of satisfaction of my intention, nor is it implied by the conditions of satisfaction of my intention, because if I fail to cause you pain, I do not fail in what I was trying to do. When I fix your tooth I may have a firm belief to the effect that fixing your tooth will cause you pain, but I am not thereby committed to the intention to cause you pain. And the conclusive proof is given if we ask what counts as succeeding or failing. If I fail to cause you pain, it is not my original intention that has failed; rather one of my beliefs has turned out to be false. So it is simply not the case in general that anybody who wills the end (in the sense of having an intention to achieve that end) thereby wills everything that occurs as a known part of carrying out that intention.

However, there is a type of case in which Kant's principle is true. Suppose I have the intention-in-action to fix your tooth, and suppose that I also have the belief that the necessary condition of fixing your tooth is that I *intentionally* drill your tooth. This case differs from the previous case because drilling your tooth is not a collateral part of fixing your tooth in the way that causing you pain is a collateral part of fixing your tooth. Rather, it is a *means* that must be intended in order that the original intention can be carried out. So, there is a natural interpretation of Kant's principle where it turns out to be correct, and that interpretation is as follows:

If I intend an end *E*, and I know that in order to achieve *E* I must intentionally do *M*, then I am committed to intending to do *M*. In that sense it does seem to me that "he who wills the end" is committed to willing the means.

VII Conclusion

The moral of this discussion can be stated quite briefly. Deductive logic deals with logical relations between propositions, predicates, sets, etc. In the strict sense there is no such thing as a deductive logic of practical reason, but then in the strict sense there is no such thing as a deductive logic of theoretical reason. Because of the combination of commitment and direction of fit of beliefs, it is possible to get a mapping of the logical relations occurring in theoretical reason onto deductive logic of a sort that is not possible for practical reason. Why the difference? In two important respects desire is unlike belief. Desire has the upward direction of fit, and a person with a desire is not committed to the satisfaction of that desire in the way

that a person who holds a belief is committed to the truth of the belief. This allows for the two features of desire we noted earlier, the necessity of inconsistency and the non-detachability of desire. Intentions are a bit more like belief because they do involve a commitment to the satisfaction of the intention. Nonetheless, the person who has an intention is not committed to intending to achieve all of the consequences of the achievement of his intention. He is committed only to those means that are necessarily intended in order to achieve his ends. For these reasons there will not be a "deductive logic of practical reason" even in the limited sense in which we found that it is possible to have a deductive logic of theoretical reason.