Notre Dame Journal of Formal Logic
Volume 33, Number 2, Spring 1992

Book Review

David H. Sanford. *If P, then Q: Conditionals and the Foundations of Reasoning*. Routledge, London, 1989. 265 pages.

This interesting book covers a great deal of ground: it is a survey of the history of philosophical discussion of conditionals from the Stoics to the twentieth century, an exposition and critical discussion of the contemporary literature on conditionals, and a sketch of an alternative approach to the problems—all of this in about 250 pages. The historical survey provides a useful background, though it is of course brief and sketchy and does not for the most part claim to be making new discoveries or interpretations. The discussion of the contemporary literature concentrates on the possible worlds analyses of conditionals. The expository chapters are clear, accurate, and sympathetic; in the critical chapters a good deal of skepticism is expressed about the usefulness and relevance of possible worlds to the understanding of the interpretation of conditionals. Many of the critical points are impressionistic and not developed in much detail. It is suggested that if the possible worlds analyses were fully specified (including specification of the similarity relations relative to which conditionals are interpreted), reference to possible worlds would be seen to be unnecessary—the analysis could be expressed entirely in terms of facts about the actual world. But little is done to clarify or support this suggestion. The sketch of an alternative approach in the last three chapters emphasizes the diversity in the patterns of dependence and independence that make conditionals true, describing a range of examples and facts that a theoretical account of dependence must account for, but giving the reader little idea what a positive theory that accounts for these facts might look like. In general, I think Sanford exaggerates the contrast between an analysis of conditionals in terms of possible worlds and an analysis in terms of facts (patterns of dependence and independence) about the actual world. Everyone—even a modal realist like David Lewis—agrees that contingent counterfactuals state facts about the actual world. The issue is how best to describe these facts.

I will comment in more detail on two issues that Sanford discusses: first on his criticisms of what has come to be called the Ramsey test for the evaluation of conditionals; second on the problem about validity posed in the last chapter and a notion called "circumstantial validity" developed in response to the problem.

The Ramsey test is a thesis about how conditional statements—both indicative and subjunctive—are evaluated. The claim is that we can decide whether we think a conditional is true by performing the following thought experiment: add the antecedent, hypothetically, to your stock of beliefs, make the minimal adjustments necessary to maintain consistency, and then consider whether or not the consequent follows from the hypothetical stock of beliefs that remains. If it does, then accept the conditional; if not, do not. Sanford argues that some acceptable conditionals fail the Ramsey test. Consider "If I have been on the moon, then I have been over ten thousand miles from Detroit." This statement is, he suggests, unproblematically true, but the Ramsey test is inapplicable, since "no minimal or even moderate revision of my beliefs can coherently accommodate the additional belief that I have been on the moon. . . . Some knots in the web of belief cannot be untied without unravelling the whole fabric" (p. 143). But I think this criticism is based on a misunderstanding of the Ramsey test. The proposed thought experiment does not ask you to consider what you would believe if you learned that the antecedent were true, and it does not ask you to consider a hypothetical situation in which you have certain beliefs. Rather, you are asked to consider a certain partial conception of the way things are that is defined in terms of what you believe but that can be abstracted from the facts about your state of mind, or anyone's state of mind. Begin with your actual conception of the way things are—your actual stock of beliefs. Now modify it by adding a certain proposition and adjusting the resulting set of propositions to maintain a consistent conception of the way things are. The resulting set of propositions characterizes a certain conception of the way things might be, but it need not be a conception that could be *your* conception of the way things are. To use another of Sanford's examples, suppose I am evaluating the counterfactual, "If I had died before age ten, I would never have heard of Tarski". What does the Ramsey test suggest I do? It does *not* suggest that I ask myself, "What would I conclude if I suddenly learned that I had been dead for forty years? Would I conclude that I have never heard of Tarski?" Nor does it ask me to consider a situation in which, however I got that way, I now believe that I died before age ten. Instead of this the Ramsey test tells me to add to the set of propositions that constitutes my actual beliefs the proposition that I die before age ten. I should then throw out from this set, for example, my beliefs about my experiences as a graduate student, since these propositions now conflict with my beliefs about what can happen to people after they die. I should, however, retain in this hypothetical stock of beliefs the proposition that at all ages before ten, I had never heard of Tarski, from which it will follow, given the antecedent, that I never heard of Tarski at all. I should also keep the proposition that Tarski is a great logician, since that is perfectly compatible with my dying before age ten. Now the most natural ways of carrying out this procedure will yield a hypothetical stock of beliefs that is not a candidate to be my beliefs, since it includes both the proposition that I have never heard of Tarski and the proposition that Tarski is a great logician. The procedure might in some cases even yield a hypothetical stock of beliefs that is not a candidate to be *anyone's* beliefs, but that won't prevent it from being a perfectly good partial conception of a way things might be, and that is all it needs to be for the applicability of the Ramsey test.

The Ramsey test is silent on the question of how one should choose among

the many ways of revising the stock of beliefs to maintain consistency; it is a modest proposal that does not contribute to a solution to Goodman's problem. But it does provide a test that can discriminate between different proposals for an abstract account of the truth conditions of conditionals. It is clear, for example, that the Ramsey test would not be an appropriate way to evaluate conditionals if either the material conditional analysis or an analysis in terms of a necessary connection of some kind between antecedent and consequent were correct. So the Ramsey test, insofar as it is intuitively plausible, provides evidence against such analyses even if it says little of substance about which counterfactuals to believe.

Now there is a second thesis that has sometimes confusedly been labeled "the Ramsey test", and this *is* a thesis about belief revision. The thesis is that to accept a conditional whose antecedent is compatible with what one accepts is to be disposed to accept the consequent of the conditional upon receiving (total) evidence whose content is the antecedent. This thesis follows from the conjunction of the Ramsey test with a principle of methodological conservatism that says that if one's total new evidence is compatible with one's prior beliefs, one should retain all of one's prior beliefs. But the thesis is severely limited in its scope, applying only when the antecedent is something that is compatible with what is accepted, and only when the antecedent is something that could be the content of one's total relevant new evidence. So Sanford's counterexamples do not apply to this thesis either, since their antecedents ("I have been on the moon," and "I died before age ten") are incompatible with the relevant person's beliefs.

But the first of these examples does raise some problems, since it is an *indicative* conditional. It seems plausible to hold that indicative conditional sentences are appropriately used only in a context in which the proposition expressed by the antecedent is compatible with what is presupposed in that context—which is to say only in a context in which it is an open question whether the antecedent is true. If I cannot coherently entertain a change in my beliefs in which I accept that I have been on the moon, how can I understand an indicative conditional which requires a context in which it is an open question whether I have been on the moon? All that is required is a context in which it is not mutually presumed by speaker and addressee that the antecedent of the conditional is false. The speaker may believe that the antecedent is false—may even be absolutely certain that it is and unable to imagine learning otherwise. But even if Sanford cannot coherently suppose that he learns that he has been on the moon, he will have no problem operating in a context in which it is an open question for others whether he has been on the moon (he might, for example, be at a dinner party where half the guests are astronauts and his dinner partner does not know whether he is one of them). It is hard to think of a natural context of this kind in which it would be informative and appropriate to say "If I have been on the moon, then I have been over ten thousand miles from Detroit"; but then this is a peculiar conditional, and it is less clear to me than it is to Sanford that it is true. We need a context to tell whether it would be acceptable, and in some contexts I think one might be inclined to assert a conditional that seems to conflict with it. ("And have you been to the moon, Mr. Sanford?" she asks. "Well, if I have," he replies, "it's a lot closer than I think, since I have never ventured more than a few hundred miles from Detroit.")

Perhaps, as Sanford argues, following Malcolm, we cannot coherently imagine learning that certain of our most securely held beliefs are false. I am skeptical even of this, but that is another issue. But this does not imply that we cannot coherently suppose those beliefs to be false, nor does it imply that we cannot coherently speak in a context in which those beliefs are called into question by others. This is all we need for the Ramsey test and for the assumption that indicative conditionals require a context in which the antecedent is compatible with what is presupposed in that context.

The last chapter is called "A problem about validity". I think the examples used to pose this problem do raise a problem, but I want to express some doubts both about the way Sanford characterizes the problem and about his proposals for its solution. I will make some suggestions about an alternative way of describing and explaining the phenomena.

The problem, Sanford says, "is the problem of accounting for valid substitution instances of invalid argument forms" (p. 227). As the discussion makes clear, there is no general mystery about the fact that some instances of invalid argument forms are valid. Particular arguments are always instances of many forms, and for *all* arguments, valid or invalid, there are invalid forms of which they are instances. For example, all two-premise arguments are instances of the invalid form *A, B, therefore C*, even those that also have the valid form *if P then Q, P, therefore Q*. But the problem Sanford is worried about does not concern instances of invalid forms that are valid because they are also instances of different forms all instances of which are valid. The problem is that for some argument forms that have both valid and invalid instances, the valid instances "seem to be valid in virtue of the argument form in question". Even though there are, for example, genuine counterexamples to contraposition—instances that show that the argument form is not in general valid—there are also instances that seem to be valid *and valid in virtue of being instances of contraposition*. Of course no argument can be valid in the classical sense in virtue of being an instance of a form that has genuine counterexamples. So, it is suggested, we need a new notion of validity—circumstantial validity—to explain the appearance of validity that these arguments have and to explain the fact that they are, despite being invalid in the classical sense, rationally compelling arguments.

So what is circumstantial validity? An argument is circumstantially valid if there is a related argument with different premises and the same conclusion that is classically valid. The premises of the related argument are statements that state the *grounds* of the premises of the given argument. "By 'grounds' here," Sanford says, "I mean 'reason why something is true' as opposed to 'reason why something is believed to be true'" (p. 232). So an argument is circumstantially valid if there is a classically valid argument from statements of the facts in virtue of which the premises of the original argument are true to the conclusion.

I have several worries about this definition, all deriving from the fact that circumstantial validity, defined this way, depends on facts on which the truth of the premises objectively depends. First, the definition seems to apply only to arguments that have true premises. Suppose I give an argument with a premise that is (perhaps unknown to me) false. It may be classically valid, and it may be rationally compelling in the sense that its premises provide conclusive reasons for its conclusion. But I don't see how it can be circumstantially valid, since there

are no grounds for the truth of the premises—no reason why they are true. Second, circumstantial validity, as I understand it, is an external notion in the sense that whether an argument is circumstantially valid will depend on facts that may be unknown both to the speaker giving the argument and to the audience to which the argument is addressed. Suppose I have been given a piece of disjunctive information, "*A or B*". The ground for its truth (assuming it is true) is either *A* or *B*, but I do not know which. Different arguments using this premise will be circumstantially valid, depending on which disjunct is true, but surely the rational force of such arguments will not depend on this. Third, there will be arguments that are circumstantially valid, in the sense defined, but that have no appearance of validity and that are in no ordinary sense good arguments. Suppose I argue as follows: "At least one U.S. President was born in Kentucky; at least one U.S. President was assassinated. Therefore, at least one U.S. President born in Kentucky was assassinated". I don't think we want a notion of validity according to which this argument is valid; but I believe it is circumstantially valid, since the fact that Lincoln was born in Kentucky and assassinated is part of the grounds for the truth of the premises, and the conclusion follows from a statement of this fact.

If I've got Sanford's notion of circumstantial validity right, I don't think it is an appropriate notion for evaluating arguments in context, but I do think that his examples point to a real problem about validity. Let me conclude by trying to say what I think it is.

Validity, in the classical sense, is an abstract semantic relation between a set of sentences or propositions and a sentence or proposition. Its definition makes no reference to reasons or reasoning. But we usually assume that validity has something to do with arguments in which premises are presented as reasons for accepting a conclusion. If one wanted to make the pragmatic function of argument explicit, one might define a good demonstrative argument as one whose premises provide conclusive reason for the conclusion. Slightly more precisely, an argument might be defined as good if and only if no ideally rational person would accept the premises without also accepting the conclusion. Now it seems reasonable to conclude that arguments that are logically valid in the classical sense are good arguments in the sense defined, but there are arguments that seem to be good compelling arguments, in context, even though they are not valid in the classical sense. Furthermore, it seems in some cases to be the form of the arguments that accounts for their being conclusive arguments. How can this be? I think the key to the explanation is in the context dependence of the premises and conclusion. Context constrains the interpretation of the sentences used to make assumptions or assertions, and the making of assumptions or assertions in turn alters the context in which subsequent assumptions or assertions are made. The pragmatic rules governing this dynamic process may give rise to systematic connections between the premises and conclusion of an argument. Those connections depend on the logical forms of the premises and conclusion but are not explicable in terms of the classical notion of validity, which abstracts away from the fact that the premises are premises, asserted or assumed, and the conclusion is a conclusion, inferred from what is asserted or assumed. Conditionals are highly context dependent—their interpretation depends on the presumed interests and mutual beliefs and presuppositions of the participants in a conver-

sation—so arguments involving conditional premises and conclusions are arguments that might be expected to manifest this kind of pragmatic structure. I think one can explain apparently valid instances of both the hypothetical syllogism and contraposition in terms of such structure. Using one of the examples Sanford discusses in this chapter, let me sketch how an explanation might go in the case of contraposition.

Suppose an indicative conditional, *if A then B*, is asserted as a premise. It must be an open question, in the context, whether the antecedent is true. (It need not have been an open question before the conditional statement was made, but by making it the speaker opens the question.) It must also be an open question whether the consequent is true, since otherwise the conditional will be either presumed false—incompatible with the context—or else already presumed true, and in either case it would be inappropriate to assert. In *most* cases, it will also be an open question whether both antecedent and consequent are false, since otherwise, in context, the speaker's assertion, *if A then B*, will have exactly the same force as the categorical assertion, *B*, and so the conditional qualification will have no point. (I say in most cases since there may in some cases be some explanation of the point of the redundant conditional antecedent. Perhaps the consequent is of interest only if the antecedent is true, as in J. L. Austin's example, "there are biscuits on the sideboard if you want them". Or perhaps the speaker wants to assure the addressee that he is taking account of the possibility that *A* be true, as in such examples as "even if it rains, the game will be played".) So in a normal context in which an indicative conditional *if A then B* is asserted, at least the three possibilities, $A\&B$, $A\&\sim B$, and $\sim A\&\sim B$, will be compatible with the context. It follows that if the assertion is then accepted in this context, the context that results will be one in which the contrapositive, *if not B then not A*, is also accepted. It does *not* follow, however, that the proposition expressed by the conditional premise semantically entails the proposition expressed by its contrapositive. The conclusion is only that if what the premise says in the initial context is accepted, then what the conclusion says in the resulting context must also be accepted.

Let me illustrate with an example from Sanford's discussion of circumstantial validity: "If I draw a flush, I will win at least one big pot tonight. Therefore, if I do not win at least one big pot tonight, I will not draw a flush". A natural context would be one in which, for all that is presupposed before the premise is asserted, the speaker might or might not draw a flush and might or might not win at least one big pot, and it is not excluded that the speaker might neither win a big pot nor draw a flush. If we accept what he says, we are then in a context in which it is presupposed that he does not draw a flush without winning at least one pot, but it is still not excluded that he neither wins a big pot nor draws a flush. So the contrapositive conclusion, "if I do not win at least one big pot tonight, I will not draw a flush", is accepted in the resulting context. But it does not follow that if the premise is *true* the conclusion must be *true*. Suppose the facts of the case are these: the speaker will in fact win the first hand, but not with a flush. Furthermore, the hand will be one on which everyone bets heavily, and the pot is big. But after that everyone will become extremely cautious (perhaps as a result of being burned on the first hand), and there will be no more big pots. I think the facts, as I have described them, make the conditional true. The

speaker will win at least one big pot whether or not he draws a flush sometime during the evening and so will win at least one big pot if he draws a flush. But the contrapositive, it seems to me, might well be false. If he didn't win that first hand—the only big pot—he wouldn't win at least one big pot, but he still might draw a flush.

If the facts of the example are as I have described them, the argument will not be circumstantially valid in Sanford's sense, since whereas the premise is true and the speaker believes it is true, the ground of its truth is different from the ground of the speaker's belief. The conclusion may be false, and in any case does not follow from the facts in virtue of which the premise is true. But that the facts turn out this way should not be relevant to the semantic and pragmatic evaluation of the argument.

As is customary, I have concentrated my attention on arguments and proposals with which I disagree, but I found much to agree with and learn from. The book has many interesting examples and stimulating arguments, as well as concise and reliable exposition of what has been said about conditionals over the years. This is a good place to learn about conditionals or to have one's ideas about them challenged.

Robert Stalnaker
Department of Linguistics and Philosophy
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139