THE JOURNAL OF PHILOSOPHY

VOLUME LXVIII, NO. 10, MAY 20, 1971

CONSISTENCY OF BELIEF *

PHILOSOPHERS have traditionally assumed that the set of beliefs of a rational man should be *belief-closed* and *belief-consistent,* which I define as follows:

(1) The set of a man's beliefs at time t is *belief-closed* if and only if whenever the man believes at t sentences $S_1,S_2,\ldots,S_n$ and also believes at t that $S_1,S_2,\ldots,S_n$ deductively imply sentence S, then at t the man believes S.

(2) The set of a man's beliefs is *belief-consistent* at time t if and only if, whenever a man believes at t that sentences $S_1,S_2,\ldots,S_n$ deductively imply both sentence S and sentence not-S, then at t the man does not believe all of $S_1,S_2,\ldots,S_n$.

In the above definitions, $S_1,S_2,\ldots,S_n$ deductively imply S just in case S follows from $S_1,S_2,\ldots,S_n$ by any of the usual systems of natural deduction without the addition of any suppressed premises.

Of late, some philosophers have been moved by considerations involving probabilities to question whether our beliefs are belief-closed.† I wish to offer a more radical critique: I will argue that the set of beliefs of a rational man is neither belief-closed nor belief-consistent. My argument will not depend on any probabilistic considerations, but will start from an analysis of the way beliefs function in seemingly deductive-nomological (hereafter, D-N) explanations and predictions. The adjective 'seemingly' is necessary here, for it will follow from the argument that there are no pure D-N explanations or predictions.

* I would like to thank Gilbert Harman and James Rachels for their comments and criticisms.

† See Isaac Levi, "Deductive Cogency in Inductive Inference," this JOURNAL, LXII, 3 (Feb. 4, 1965): 68–77, for a discussion and further references.

In saying that our beliefs are neither closed nor consistent, I do not wish to deny the importance of the laws of logic in explanation and prediction. In the last part of this paper I will offer a constructive, though extremely sketchy, account of the place of logic in the ordering of our beliefs. Such a positive account will partially explain our strong disinclinations to abandon the principles of belief-closedness and belief-consistency.

A simple tale: One day Betty is unable to start her car; when she turns the ignition key, nothing happens. She asks her friend Archie what he thinks the trouble is, and after making certain tests (he tries horn, lights, etc.) Archie informs her that the car's battery is dead. Betty, uncertain whether Archie has stated symptom or cause, asks if the dead battery is the source of her difficulty. "Yes," replies Archie, "A car won't start if its battery is dead." Archie now suggests that Betty's car be given a push. Betty, unsure whether that will start the car or whether Archie simply wants the car pushed to a service station, asks whether a push will start the car. "Yes," says Archie, "A car with a dead battery will start if given a push." Another car is found, a push given, and Betty's car starts.

According to the D-N theory, in the first part of the story Archie explained why Betty's car refused to start. He did this by giving a general law ("A car won't start if its battery is dead") and a statement of initial conditions ("Your car's battery is dead") from which could be deduced a sentence describing the phenomenon to be explained ("Your car will not start"). In the second part of the story, Archie made a prediction about how the car would behave under certain conditions. The form of the prediction is the same as in the preceding explanation: A general law ("A car with a dead battery will start if given a push") and initial conditions ("Your car's battery is dead" and "Your car is pushed") were stated or implied, from which a sentence describing the phenomenon predicted ("Your car will start") could be deduced. The prediction differs from an explanation only in that the sentence deduced is in the future tense. The explanation was important because it provided the statement of one of the initial conditions occurring in the prediction. Thus, the explanation was instrumental in helping Archie and Betty get the car started.

The above analysis cannot be accepted as it stands. To see this, let us assume that Archie believes:

(A1) A car will not start if its battery is dead.
(B1) A car with a dead battery will start if given a push.
(C1) Cars with dead batteries have been pushed.

and that Archie does not believe

(X) A car with a dead battery will not start if given a push.

The story gives us good reason to say that Archie believes (A1) and (B1). Common sense tells us that Archie might naturally be expected to believe (C1), and not to believe (X). But now it is obvious that there are substantial problems here. First, it would seem that Archie was able to make the correct prediction only because he used (B1) as a premise—had he used (A1) instead, with the same statement of initial conditions, he could have deduced (predicted) "Your car will not start"; the argument that yields this prediction is, except for tense, identical with the argument that yielded the original explanation of why the car refused to start. Secondly, it is obvious that Archie's beliefs as given are neither belief-closed nor belief-consistent; the believed (A1) deductively implies (X), which is not believed, and (A1), (B1), and (C1) taken together imply that a certain car—one that has a dead battery and is pushed—will both start and not start.

One might argue that whereas Archie believes (A1), (B1), and (C1), he does not believe that (A1) implies (X), or that (A1), (B1), and (C1) together entail a contradiction. Certainly, Archie need not have had these entailments in mind while he was helping Betty start her car. But if we restrict beliefs in entailments to those of which one is actually conscious, then the principles of belief-closedness and belief-consistency become perfectly trivial.

I will not speculate on what form a theory of belief in entailment would take. Nonetheless, I think it can be said that, as a minimal condition, a rational person usually believes that $S_1, S_2, \ldots, S_n$ deductively imply S if the entailment does in fact hold and if the argument (in most of the standard systems) is very short. In the story above, these conditions are fulfilled.

If Archie's beliefs are to be consistent, he cannot believe all of (A1), (B1), and (C1). Yet he obviously believes something about cars and batteries and pushes. Any attempt to save belief-consistency must, therefore, deny Archie either (A1) or (B1)—I assume (C1) is not in question—while offering a replacement belief or beliefs that will allow Archie to get Betty's car started.

A simple way out may seem to present itself: Suppose Archie doesn't believe (A1)—"A car will not start if its battery is dead"—but instead believes

(A2) If a car's battery is dead and the car is not being pushed, then the car will not start.

If we add the statement that the car is not being pushed, we can explain the car's not starting. (B1) still remains; so the prediction still goes through. And (A2), (B1), and (C1) are consistent. We might say that (A2) was what Archie meant to say in uttering (A1).

I would not deny that Archie may believe (A2). But if he does so, I imagine that he believes *both* (A1) *and* (A2). Simply replacing (A1) by (A2) seems ad hoc, and for good reason. Suppose Archie, instead of proposing to start the car by giving it a push, had proposed instead that the dead battery be wired to a live battery. "Will that start the car?" Betty asks. "Yes," replies Archie, "A car with a dead battery will start if wired to a live battery." Referring to this last sentence as (B2), and letting (C2) be the unproblematic "Cars with dead batteries have been wired to live batteries," we can see that if Archie believes (A2), (B2), and (C2), then his beliefs again fail to be belief-consistent.

The problem is that if we are to replace (A1) by a sentence like (A2) so as to render Archie's beliefs consistent, it must be by some sentence where the antecedent of the conditional explicitly denies that the car with the dead battery is in any of those configurations where Archie believes a car with a dead battery will start—such as being pushed, wired to a live battery, hand-cranked, rolled down hill, etc. The sentence will thus be of the form

(A3) If a car's battery is dead and the car is not [disjunctive list of all ways Archie believes a car with a dead battery can be started], then the car will not start.

Assuming that such a list can be given, there still remains an awkward problem: Assume Archie is rational enough to be modest—he realizes that he does not know everything that is known. In particular, he believes that other, more knowledgeable people have used methods of starting cars unknown to him. If he can believe a sentence of the form (A3), I do not see how we can deny that he might believe a sentence of the form

(D3) Cars with dead batteries and which weren't [disjunctive list of all ways Archie believes a car with a dead battery can be started] have started.

If Archie believes both (A3) and (D3), then his beliefs are not belief-consistent.

Of course, it might be that starting cars is one of the passions of Archie's life and that, owing to a lifetime of studying car starting, Archie may rationally believe that he knows of every method of starting a car with a dead battery that has ever proved successful.

In such a case, Archie would not and need not believe (D3). But life is short, and if one is passionate about car starting, one must perforce be less passionate about other matters. Thus, there will always be some areas in which Archie, if he is rational, will be able to make explanations and predictions while believing that others know more than he. So even for a passionate car starter, the solution of replacing (A1) by (A3) cannot be duplicated in other areas.

The attempt to save belief-consistency by modifying the antecedent of (A1) proved futile; perhaps the trouble with (A1) lies in the consequent. Suppose we replace (A1) by

(A4) If a car's battery is dead, it will not start in the usual way.

Leaving (B1) and (C1) unchanged, we see that the contradiction (X) no longer follows. Instead we simply get the consequence that there is a car which starts, but not in the usual way.

As long as we are speculating about what Archie meant to say in uttering (A1), we might as well inquire what Archie means by "starts in the usual way." If he simply means "starts by turning the key," we are in trouble, for then Archie presumably believes

(B4) If a car with a dead battery is wired to a live battery, then it will start in the usual way.

(A4), (B4), and (C2)—"Cars with dead batteries have been wired to live batteries"—now yield a contradiction.

But it may be that what Archie means by "starts in the usual way" is that the car starts by turning the key, but without the help of various external aids, such as extra live batteries wired to the car's battery. If this is the case, then Archie presumably believes a sentence of the form

(A5) If a car's battery is dead, and the car is not [disjunctive list of all ways Archie believes a car with a dead battery can be started by turning the key], then the car will not start by turning the key.

(A5) represents, however, only a very slight improvement on the defective (A3), in that Archie is perhaps a bit less likely to believe the following parallel to (D3):

(D5) Cars with dead batteries and which weren't [disjunctive list of all ways Archie believes a car with a dead battery can be started by turning the key], have been started by turning the key.

Still, Archie may very well believe (D5), in which case we are again served with a contradiction. And even if Archie does not believe

(D5), the proposed solution, like the previous one, cannot be duplicated for other cases.

Suppose it is suggested that Archie doesn't really believe (A1), but instead believes

(A6) All other things being equal, a car will not start if its battery is dead.

and that (A6) was "what Archie meant" when he uttered (A1). The intuitive appeal of this emendation is obvious: If Archie had thought about it for a few moments, he would have realized that he didn't really believe (A1), but did believe the more carefully hedged (A6). Similar reasoning will replace (B1) by

(B6) All other things being equal, a car with a dead battery will start if given a push.

I shall refer to this answer to the problem raised above as the *ceteris paribus solution.* The ceteris paribus solution saves belief-consistency—(A6), (B6), and (C1) are logically compatible. But the "solution" also raises some serious difficulties.

The "insight" that caused us to replace (A1) and (B1) by (A6) and (B6) will hold for virtually every universal generalization: At first we think that we believe that swans are white, but on reflection we see that there can be painted swans, mutations, etc., and that what we *really* believe is that all other things being equal, swans are white. With the possible exceptions of analytic truths and a few laws of physics, all generalizations may be seen to be believed only ceteris paribus.

If this is the case, then there will be very few interesting logical connections among our beliefs. For virtually nothing follows from the statement that all other things being equal, swans are white. Even the weak, "If something is a swan, something is white" need not follow, for it may be that all other things have never been equal.

If there exist few interesting logical connections among our beliefs, a way must be found to explain the apparent importance of logical connections in the business of explanation and prediction. Consideration of the story of Betty's balky car suggests that when giving explanations we employ sentences which we don't really believe but which do possess interesting logical connections. These sentences are like our beliefs except that the ceteris paribus condition has been deleted. Presumably we can do this in certain situational contexts because we judge that "all other things are equal." Since a great many of our everyday utterances are treated in logic courses as universal generalizations, we seem forced to say that we

do not believe, or at least do not quite believe, a great many of our everyday utterances. More distressing, these utterances include most of those great universal truths of which professors attempt to convince students.

These objections are not fatal. But they are serious, and would give one cause for rejecting this answer if a better one were available. I shall now sketch what I consider to be such an answer.

It seems that Archie was only able to explain the car's failure to start, and make arrangements for getting it started, because he ignored the deductive consequences of some of his beliefs while emphasizing the deductive consequences of others. He originally stressed the deductive consequences of (A1) in order to explain the car's not starting. Then, in trying to get the car started, he ignored (A1) and its consequences, laying stress instead on (B1) and its consequences.

Let us assume that what seemed true of Archie in the story is generally true of ourselves—that is, in various situations, we accept some of the deductive consequences of some of our beliefs, but not others. If this is true, then explanations and predictions which seem to be D-N must involve not only the deduction of a sentence describing the phenomenon to be explained or predicted from statements of general laws and initial conditions, but also a decision whether to accept the deduction. Such a decision is already implicit in our choosing to start from certain beliefs and not others. Thus Archie, in giving his explanation, chose to start from (A1) rather than (B1), but then in making the prediction chose to start from (B1) rather than (A1).

Why do we accept some deductions and not others? Or, which is the same, why do we accept certain premises at some times and not others? I do not know how to answer this question in adequate detail, but I can suggest the form of an answer. We can imagine that many, and perhaps all, the sentences we use may only be used in certain situational contexts. That is, they are considered acceptable for explaining and predicting only if certain conditions, involving the disposition of certain physical objects and the states of mind of certain people, are satisfied.[1] Acceptable explanations and predic-

[1] I tend to believe that the situational contexts in which a sentence is acceptable are invariant for different speakers of the language. If I am wrong about this, some interesting consequences follow. If a sentence believed by two different people need not be acceptable to each of them in all the same contexts, then people who in one sense "believe all the same things" may still have important functional differences related to their beliefs—differences which will, among other things, affect their behavior.

tions can involve only those premises acceptable in that particular situational context. This would explain why we get so excited about some inconsistent beliefs while remaining blasé about others; when believed sentences $S_1,\ldots,S_n$ deductively imply both sentences S and not-S, there is no cause for concern unless there is some situational context in which $S_1,\ldots,S_n$ are all simultaneously acceptable. If there is such a situational context, the inconsistency is no longer a trifling matter. (There may, of course, be some contexts in which only some nonstandard logic is acceptable.) Applying this concept to the story of Archie and the balky car, we can say that there was a context of explanation and a context of prediction, these contexts differing because of certain changes in the mental make-up of the participants. In the context of explaining the car's failure to start, (A1) was acceptable. Once the explanation was given, the context changed, and, in the new context of predicting how the car would react to being pushed, (A1) was no longer acceptable. Archie's beliefs were inconsistent, but because there was no single context in which all of (A1), (B1), and (C1) were acceptable, the inconsistency was unimportant and probably unnoticed.

Assuming that sentences are acceptable only in certain contexts may help explain certain other phenomena. Suppose that, in the story, Archie had verified that there was gasoline in the car and had made this fact known to Betty and to Archie's friend, Jughead. Now suppose that when Archie says that a car with a dead battery will start if given a push, Jughead interrupts to tell Archie that he remembers Archie once saying that a car with a dead battery and an empty gas tank will *not* start if given a push. We may expect Archie to respond (perhaps with an edge in his voice) "But the gas tank *isn't* empty, Jughead!" Jughead's point—that Archie believes

(A7) A car with a dead battery and an empty gas tank will not start if given a push.

and presumably also

(C7) Cars with dead batteries and empty gas tanks have been pushed.

and that these contradict (B1)—is quite correct. But so is Archie's rather natural-sounding reply: The car's gas tank is not empty; and in such a context, (A7) is not acceptable. There is no contradiction of any importance here, for there is no situational context in which (A7), (B1), and (C7) are all acceptable. If there are no situational contexts in which (A7), (B1), and (C7) are all simultaneously acceptable, then Archie's beliefs, while formally inconsistent, are not inconsistent in any important sense.

If sentences used as premises in explanations and predictions are acceptable in some situational contexts but not others, then each seeming D-N explanation or prediction must be preceded by a decision whether the premise is acceptable in the particular context at hand. This is the reason for my claim that there are no pure D-N explanations. It may be, of course, that there are certain sentences that are acceptable as premises in *all* contexts. I doubt that there could be many such sentences, and perhaps there aren't any. But in any case, the existence of such sentences would not force me to alter my claim. For the fact that a sentence is acceptable in all contexts is another preliminary fact that must be discovered, and its discovery is subsidiary to (and must presumably precede) the explanation in which it figures. Even if virtually all the sentences used in explanations were acceptable in all contexts, the mere existence of some that were acceptable in some contexts but not others would still cause *every* explanation to involve certain preliminary decisions about the acceptability of the sentences in the particular context at hand.

The abandonment of belief-closedness and belief-consistency, combined with the modified theory of D-N explanation sketched above, has substantial trouble-making potential. One can imagine someone refusing to give up a theory in the face of a demonstrated inconsistency by asserting, contrary to fact, that there has been a shift in context. Without a specification of how to determine those contexts in which a sentence is acceptable—a specification which I am unable to give—his position may seem unassailable. All I can say is that I suspect that generally we have strong intuitions about acceptability, and that there is a high order of agreement among different individuals concerning these intuitions. But there probably are some borderline cases, and at the borderline my sketch of a theory seems to offer the charlatan plenty of room for maneuver. I hope that we may some day have a detailed systematic theory of the relations between sentences and the contexts in which they are acceptable, but I am not optimistic about the prospects for obtaining such a theory in any short or middle run.

Despite its present trouble-making potential, I would claim that the theory confirms most of our intuitions about cases. In our story, and on my theory, the explanations and predictions offered were the correct ones. On my theory, Archie's rather natural-sounding objection to Jughead's kibitzing went directly to the point. And lastly, on my theory, Archie's beliefs turn out to be just what they appear to be. In this last point lies the superiority of my answer to the ceteris paribus solution considered earlier.

There is an evident parallel between my answer and the ceteris paribus solution. On my theory, there are certain situational contexts in which a sentence is acceptable. The contexts in which a sentence is acceptable on my theory will presumably be just those contexts in which the ceteris paribus condition is satisfied on the other theory. Both theories will, therefore, allow the same explanations and predictions. The problem of deciding whether a certain sentence is acceptable in a certain situational context is identical for the two answers. Thus, the charlatan has as much room for maneuver in the ceteris paribus solution as on my theory. On both views, the sentences used in various explanations and predictions are, taken together, logically inconsistent. And on both views, legitimate doubts may be expressed about our understanding of what it is a person really believes; on my theory, these are doubts about whether we really understand a belief unless we know how it is used, i.e., the situational contexts in which it is considered acceptable. On the ceteris paribus view, these are doubts about our understanding of the phrase "All other things being equal, ..."

My theory, I would argue, possesses the considerable advantage of being closer to colloquial usage. On my theory, the sentences we sincerely utter in explanation and prediction (and instruction, casual conversation, etc.) are sentences we believe; our beliefs are relatively scrutable. On the ceteris paribus theory, our beliefs become more shadowy entities much further from the surface of our lives and the tips of our tongues.

My answer separates the problem of explicating belief from the problem of delineating the relationship between sentences and the situational contexts in which they are judged acceptable. The latter problem seems to me to be much the harder, and I doubt that it can be solved by purely philosophical techniques. But the former problem seems solvable, and I flatter myself that I have made some progress with it.[2] Naturally, any theory of belief worthy of the name must give us some insight into at least some of those philosophical tangles which involve belief. Whether a theory that ignores situational contexts can accomplish this, I do not know (though I have some optimistic suspicions). But it seems to be the only type of theory we are likely to have for the forseeable future; so I suspect that philosophers will have to reconcile themselves to it.

HOWARD DARMSTADTER

New York University

[2] "A Model for Belief" (forthcoming).