finite ordinals.[16] As above, there are models of T + S in which the collection of finite ordinals is standard, but in which the model of T is non-standard.

In conclusion, if one wishes to use set theory to study or shed light on a particular theory (whose consistency can be proved in set theory) or structure (which can be modeled in set theory), then there is a trade-off between deductive conservativeness and applicability of the set theory in the usual or straightforward way. One cannot have both. The trade-off depends on the decision as to whether a first-order or a second-order axiomatization is employed for the original structure or theory. The theorems for the application of set theory—the existence of representing homomorphisms—are not possible unless all models of the original theory are "standards." This, in turn, requires a second-order axiomatization, and, in this case, deductive conservativeness does not hold.

STEWART SHAPIRO

Ohio State University/Newark

COMMENTS AND CRITICISM

LEVI AGAINST U-MAXIMIZATION

THE theory of *U*-maximization is a form of decision theory advocated by Allan Gibbard and William Harper, among others.[1] The distinctive feature of this theory is that when we speak of the probability of getting a certain outcome if a certain

[16] The policies concerning the axioms of T + S are analogous to those stated in note 14 above. In T + S, the instances of the axiom schemes of T do not include set-theoretic terminology, but instances of any axiom schemes in S may contain arithmetic terminology. Of course, as in note 14, if one did allow instances of the axiom schemes of T to contain set-theoretic terminology, then one could prove that ω is isomorphic to the set of natural numbers in the usual way. In this case, S is not conservative over T.

[1] Allan Gibbard and William Harper, "Counterfactuals and Two Kinds of Expected Utility," in C. A. Hooker, J. T. Leach, and E. F. McClennen, eds., *Foundations and Applications of Decision Theory* (Boston: Reidel, 1978), pp. 123-162; reprinted in W. L. Harper, R. Stalnaker, and G. Pearce, eds., *Ifs* (Boston: Reidel, 1981), pp. 153-190. Robert Stalnaker had proposed *U*-maximization somewhat earlier in correspondence and discussion; see *Ifs*, pp. 15-18 and 151-152. Very close relatives of *U*-maximization have been advocated by others: see Brian Skyrms, "The Role of Causal Factors in Rational Decision," in his *Causal Necessity* (New Haven: Yale University Press, 1980), pp. 128-139; Jordan Howard Sobel, *Probability, Chance and Choice: A Theory of Rational Agency* (unpublished; presented in part at a conference at the University of Western Ontario, May 1978); and my "Causal Decision Theory," *Australasian Journal of Philosophy*, LIX, 1 (March 1981): 5-30.

0022-302X/83/8006/0531$00.50

option is chosen, Gibbard and Harper understand this not as a conditional probability but as the probability of truth of an option-to-outcome conditional.

Isaac Levi has argued that this theory is in deep trouble: it leads to contradictory recommendations if applied in different ways to the same case.[2] He demands that *U*-maximizers produce further principles if they wish to disown some of these applications. I shall consider what further principles might serve.

Levi's first argument concerns what he calls a "pseudo Newcomb problem." He says that in this problem "*U*-maximizers appear to be committed to two conflicting recommendations pending further clarification of their views" (Levi, p. 341). The problem is to choose between the following alternatives.

A_1: Receive (only) the contents of a certain opaque box that has been drawn at random, in advance, from the good urn. The good urn contained 90 boxes containing a million dollars each, and 10 empty boxes.

A_2: Receive a thousand dollars, and receive also the contents of a different opaque box that has been drawn at random, in advance, from the bad urn. The bad urn contained only 10 boxes containing a million dollars each, and 90 empty boxes.

The sensible recommendation is to choose A_1, and Levi grants that the theory of *U*-maximization affords an argument for this recommendation.

But he claims that it also affords an argument for the conflicting recommendation to choose A_2. That argument, quoted in full, is that "the agent has no causal influence whatsoever over whether the opaque box he selects has a million in it or not. That was 'fixed and determined' well before he makes his decision, just as in the original Newcomb problem" (Levi, p. 340).

Distinguish:

(1) The opaque box the agent selects is such that: the agent has no causal influence over whether or not it has a million in it.

(2) The agent has no causal influence over whether or not: the opaque box he selects has a million in it.

Now, (1) is true in the case imagined; and (2) would imply that A_2 is the *U*-maximizing choice. But (2) is probably false. It is false, if, for instance, the box drawn from the good urn has a million in it and the box drawn from the bad urn is empty.

[2] "A Note on Newcombmania," this JOURNAL, LXXIX, 6 (June 1982): 337-342.

Levi doubts, and rightly, that *U*-maximizers "have built into their theories the resources needed to prohibit" the argument for choosing A_2 (Levi, p. 341). Quite so: they have been remiss in neglecting to explicitly endorse

Principle 1: Fallacies of equivocation are prohibited.

Levi's second argument concerns the original Newcomb problem, in which the alternatives are as follows.

B_1: Receive (only) the contents of a certain opaque box that may contain a million dollars or may be empty.

B_2: Receive a thousand dollars, and receive also the contents of the opaque box.

This time, it's the same opaque box in both cases; and it contains a million dollars iff it has been predicted in advance that the agent will choose B_1. Given certain further assumptions (which distinguish the Newcomb problem from the many non-problems it resembles) B_1 has the greater expected value. Yet it still seems to some of us that B_2 is the rational choice, since he who chooses B_2 gets a thousand more than he would have if he had chosen B_1. The theory of *U*-maximization was built to endorse this line of thought; and Levi does not deny that it affords an argument recommending the choice of B_2.

But he claims that it also affords an argument for the conflicting recommendation to choose B_1. "To obtain this result," says Levi, "I identified the 'possible outcomes' . . . differently than Newcombites are accustomed to do," *viz.*, as B_1 and B_2 themselves; "but since nothing in their formal principles or informal arguments prohibit this maneuver, it is entirely fair to conclude that pending further elucidation the theory of *U*-maximization is inconsistent" (Levi, p. 341). If it is fair to take B_1 and B_2 themselves as the outcomes (and if we grant Levi his identification of "desirabilities of outcomes" with expected values), then indeed the *U*-maximizing alternative is B_1, the one with the greater expected value.

So the *U*-maximizers need a principle that would prohibit taking B_1 and B_2 as "outcomes." Perhaps the trouble is that B_1 and B_2 are insufficiently specific, insufficiently homogeneous in value. Either one of them could come true in two different ways—million or no million—between which the agent is far from indifferent. So it would seem that the theory of *U*-maximization ought to incorporate a prohibition against such inhomogeneous outcomes; something along the lines of

Principle 2: An outcome is complete in the sense that any further specification of detail is irrelevant to the agent's concerns.[3]

Should we then add Principle 2 to the theory of *U*-maximization?

No. We should not add it, for we cannot add it, for what is already there cannot be *added.* I quoted Principle 2 *verbatim* from Gibbard and Harper, §3.

DAVID LEWIS

Princeton University

THE WRONG BOX

IN the pseudo Newcomb problem, does the agent have causal influence over whether or not the opaque box he selects has a million in it?

David Lewis alleges that my denying this derives from a fallacy of equivocation.[1] I am alleged to have confused the following two claims:

(1) The opaque box the agent selects is such that the agent has no causal influence over whether or not it has a million in it.

(2) The agent has no causal influence over whether or not the opaque box he selects has a million in it.

Lewis declares that (1) is true but that (2) is probably false. It is highly probable that the box selected from urn 1 (Lewis's bad urn) contains nothing and the box selected from urn 2 (Lewis's good urn) contains a million. In that case, Lewis declares that (2) is false; for, given that that case obtains, if the agent were to choose the option of selecting the transparent box together with the opaque box from urn 1, the opaque box he selected would be empty. On the other hand, if the agent were to choose the opaque box from urn 2, the opaque box he selected would contain a million.

Consider the Newcomb problem and contrast the following pair of claims:

(1′) The opaque box the agent selects is such that the agent has no causal influence over whether or not it has a million in it.

[3] Given Principle 2, Levi's identification of desirabilities of outcomes with expected values is correct; but if inhomogeneous outcomes were permitted, that identification would be unwarranted.

[1] This JOURNAL, this issue, 531-534; commenting on my "A Note on Newcombmania," *ibid.*, LXXIX, 6 (June 1982): 337-342.

0022-302X/83/8009/0534$00.80