DISCUSSION

COMPATIBILISTIC REFLECTIONS

Peter van Inwagen

There are four possible positions one might take about the logical relations that obtain among free will, determinism, and indeterminism: (1) Free will is compatible with determinism and incompatible with indeterminism (*sc.* of human actions); (2) Free will is incompatible with determinism and compatible with indeterminism; (3) Free will is incompatible with determinism and incompatible with indeterminism; (4) Free will is compatible with determinism and compatible with indeterminism.

Positions (1) and (2) are the historically important ones. Position (3) has, to my knowledge, been taken only by C. D. Broad.[1] Position (4) has, to my knowledge, been taken by no one.

The adherents of positions (1) and (2) spend a good deal of time accusing each other of confusion and lack of insight. While I should not want to deny that one of these positions *is* a repository of confusion and lack of insight, I think that positions (1) and (2) are a lot more similar than is usually supposed. Each is, in a way I hope to make evident, a sort of mirror-image of the other. In the present paper, I will lay out what seems to me to be a deep symmetry between what adherents of position (1) have often said in defence of the thesis 'Free will is compatible with determinism' and something that adherents of position (2) might with equal justification say in defence of the thesis 'Free will is compatible with indeterminism'—though, to their credit, none of them ever *has* said it.

The words of Section I are spoken by an imaginary defender of position (1). (But he is typical, for all he is imaginary.) The words of Section II are spoken by a wholly imaginary defender of position (2).

I

Free will is compatible with determinism. Many philosophers have denied this (among writers in the present century, one might cite C. D. Broad, A. C. Campbell, Roderick M. Chisholm, and Richard Taylor), but we may prove them wrong as follows. Free will is the ability to act otherwise than one in fact does. That the possession of this ability is consistent with determinism may be seen from the following analysis of 'can':

Analysis I   X can do A $=_{df}$ If X decided to do A, X would do A.

---

[1] 'Determinism, Indeterminism and Libertarianism', in *Ethics and the History of Philosophy* (New York: 1952).

It is evident that even if it is determined by past events and the laws of nature that X is not going to do A, the conditional 'If X decided to do A, X would do A' may very well be true.

II

Free will is compatible with indeterminism. Many philosophers have denied this (among writers in the present century, one might cite C. D. Broad, R. E. Hobart, A. J. Ayer, and J. J. C. Smart), but we may prove them wrong as follows. Free will is the ability to act otherwise than one in fact does. That the possession of this ability is consistent with indeterminism may be seen from the following analysis of 'can':

Analysis II   X can do A $=_{df}$ If X decided to do A, X might do A.[2]

It is evident that even if it is not determined by past events and the laws of nature whether X is going to do A, the conditional 'If X decided to do A, X might do A' may very well be true.

III

It might be objected that the symmetry I allege to hold between the argument of Section I and the argument of Section II is contrived and merely verbal, owing to the fact that 'X can do A' is *not* equivalent to 'If X decided to do A, X might do A'. These statements are not equivalent (it will be said) because the conditional could be true even if it were a *mere matter of chance* whether a decision by X to do A would be followed by X's actually doing A; but (the argument continues) if it were a mere matter of chance whether a decision by X to do A would be followed by his actually doing A, then 'X can do A' would be false, since it would not be up to X whether a decision of his to do A would actually issue in his doing A.

I will concede that this argument shows that Analysis II is incorrect. This concession does not entail that the symmetry I have alleged is merely verbal, however, since Analysis I is also incorrect. This is well known: to suppose that X *decides* to do A may well be to suppose him to be endowed with powers in respect of doing A that he does not in fact possess,[3] and this possibility generates a rather diverse class of counter-examples to Analysis I. Analysis I, for example, entails that a man who is in a coma—but who is otherwise unimpaired as regards rising and walking—*can* rise and walk: if he decided to rise and walk he would. (To imagine him deciding to arise and walk is to imagine him as having emerged from his coma, and is therefore to imagine him as endowed with powers he does not in fact possess.)

---

[2] Note that the 'would' counterfactual of Analysis I and the 'might' counterfactual of Analysis II are intimately connected. As David Lewis has observed (*Counterfactuals*, Cambridge, Mass.: 1973, p. 2), 'would' and 'might' counterfactuals are interdefinable:

    If it were the case that A, it might be the case that B $=_{df} \sim$ (if it were the case that A, it would be the case that $\sim$ B)

    If it were the case that A, it would be the case that B $=_{df} \sim$ (if it were the case that A, it might be the case that $\sim$ B)

[3] Cf. Keith Lehrer, " 'Can' in Theory and Practice: A Possible Worlds Approach", in M. Brand and D. Walton, eds, *Action Theory* (Dordrecht: 1976).

Let us take this line of thought a step further. Those who think that Analysis I is on the right track, even if it is wrong as it stands, often propose adding to its analysans a conjunct saying that X can decide to do A:

Analysis Ia   X can do A = $_{df}$ If X decided to do A, X would do A, and X can decide to do A.

The 'can' in the analysans is normally treated as having a different sense from the 'can' in the analysandum (Analysis Ia could hardly be called an *analysis* if this were not so), and its presence in the analysans is held to constitute a promissory note: when the sense of *that* 'can' has been spelled out, the analysis of 'can' in the primary sense will be complete.

But if there were any philosophers who thought that Analysis II was on the right track, even if it was wrong as it stood, why couldn't they likewise add a conjunct and issue a note? What would be wrong with their proposing the following revision of Analysis II?

Analysis IIa   X can do A = $_{df}$ If X decided to do A, X might do A, and X can act on a decision to do A.

IV

Objection

The added conjunct in Analysis IIa is equivalent to 'X can do A'—that is, to the analysandum itself, while the added clause in Analysis Ia is not equivalent to 'X can do A'.

Reply

It is not true that 'X can act on a decision to do A' is equivalent to 'X can do A'. For it might be that X lacks the power to do A, though he *would* have the power to do A *if* he decided to do A. In other words, it might be that it is not within X's power to decide to do A, though it is within X's power to act on a decision to do A. We have considered just such a case in Section III: the case of the comatose man.

Objection

If human acts are undetermined, then it is not *only* true that if X decided to do A, X might do A; it is also true that if X decided to do A, X might *not* do A. And these two conditionals together are inconsistent with 'X can act on a decision to do A'. Thus, the analysans of Analysis IIa could not be true if human acts were undetermined.

Reply

Perhaps. That depends on whether 'can' in the analysans of Analysis IIa is being used in a sense that entails that 'X can act on a decision to do A' is inconsistent with the proposition that there would be only an indeterministic connection between a decision by X to do A and his subsequent action (his doing or not doing A, as the case may be).

Compare this exchange with the following exchange:

If human decisions to act are determined, and if X does not decide to do A, then it is determined that X not decide to do A. But the proposition that it is determined that X not decide to do A is inconsistent with 'X can decide to do A'. Therefore, the analysans of Analysis Ia could not be true if human decisions to act were determined and X did not decide to do A.

Perhaps. That depends on whether 'can' in the analysans of Analysis Ia is being used in a sense that entails that 'X can decide to do A' is inconsistent with the proposition that it is determined that X not decide to do A.

V

There have been various attempts to spell out 'X can decide to do A' in a way that entails that this proposition is compatible with its being determined that X not decide to do A. (And which are not themselves conditionals; a conditional analysis of 'can decide' would presumably invite its critics to raise the question whether the truth of the antecedent of the conditional mightn't augment X's powers in respect of *deciding* to do A, and the advocates of the analysis would find themselves with the problem of analysis not solved but postponed.) For example, here is an adaptation of a proposal by Wilfrid Sellars:

X can (at *t*) decide to do A $=_{df}$ There obtains at *t* no state of affairs that is incompossible with X's deciding at *t* to do A.[4]

But what is sauce for Analysis Ia is sauce for Analysis IIa:

X can (at *t*) act on a decision $=_{df}$ There obtains at *t* no state of affairs
to do A   that in conjunction with X's deciding at *t* to do A would determine that X not do A

Note that if this analysis is correct, then 'X can act on a decision to do A' is compatible with its being the case that a decision by X to do A would have no determinate connection with his subsequent action (with his doing or his not doing A, as the case may be).

VI

In parts I through V of this paper, two lines of argument are presented. One of them is an argument for the compatibility of free will and determinism. The other is an argument for the compatibility of free will and indeterminism. The former seems to be regarded by many philosophers as an adequate defence of the thesis that free will is compatible with determinism. The latter

---

[4] 'Fatalism and Determinism', in K. Lehrer, ed., *Freedom and Determinism* (New York: 1966). This, of course, is a very sketchy adaptation of Sellars' proposal. It is meant merely to provide an example and is not supposed to be an accurate representation of what Sellars actually says. It is perhaps worth noting (a) that 'incompossible' refers to physical, as opposed to logical, incompossibility, and (b) that Sellars' actual proposal involves not decisions but 'volitions' in a certain technical sense.

is regarded by no one as an adequate defense of the thesis that free will is compatible with indeterminism, for I have made it up and it certainly does not strike *me* as an adequate defence of that thesis. Though I in fact do think that free will is compatible with indeterminism, the argument for that conclusion that I have presented in this paper has, in my judgement, no merit whatever. What I should like to know is: Why does anyone suppose that the other argument is any better?

Syracuse University