Peter van Inwagen
University of Notre Dame
vaninwagen.1@nd.edu

Original Scientific Paper
UDC 123
124.6
Received: September 10, 2024
Revised: December 12, 2024
Accepted: December 19, 2024

GINET'S PRINCIPLE: OUR FREEDOM IS THE FREEDOM TO ADD TO THE GIVEN PAST

Abstract

This paper is a defense of the principle:

> *All the worlds that are might-worlds in the world x are worlds that share some of their momentary states with x*

(A (possible) world *y* is a might-world in the world *x* if it is true in *x* that if someone were to do something that person is able to do, *y* might be actual.) This principle is (given some reasonable definitions) a consequence of "Ginet's Principle": Our freedom is the freedom to add to the given past.

Keywords: the consequence argument · the past · momentary states · compatibilism · miracles

1.

We are going to need a statement of the Consequence Argument. My choice will be an elaboration of the Second Argument from the chapter "Three Arguments for Incompatibilism" of *An Essay on Free Will*.[1] The unelaborated argument was as follows.

We say that a possible world (or simply a world) *x* and a world *y* are the same just in the case that the distribution of truth-values over propositions is the same in *x* and *y*. The *actual* world is the world such that a proposition is true in that world if and only if it is true. A world is *deterministic* if no other world shares both its laws of nature and one or more of its momentary states.[2] Of course, both 'its laws of nature' and 'its momentary states' cry out

1 Oxford: at the Clarendon Press, 1983.

2 It may be that a deterministic world and an indeterministic world have the same laws. If every world that has the same laws as the world *x* is deterministic, we say that not only is *x* deterministic, but the *laws* of *x* are deterministic. The distinction between a world's being deterministic and its laws being deterministic may be important in some areas of philosophy—the philosophy of science, perhaps, or the metaphysics of modality—but it is of no relevance to the problem of the compatibility of free will and determinism.

for clarification, but I hope that each is intuitively clear enough for us to go on with. *Determinism* is the thesis that the actual world is deterministic.

Let *the free-will thesis* be this: There are agents who are trying to decide between incompatible courses of action and are able to perform each of them. (For example, a woman who must respond to some supremely important question by saying either "Yes" or "No", and is desperately trying to decide which response to make, is able to respond by saying "Yes" *and* able to respond by saying "No.")

I suppose I need to say something at some point in defense of my calling a thesis 'the free-will thesis'—or, indeed, in defense of my using the term 'free will' at all. I will explain why I see such a need. There is tightly knit group of philosophers who specialize in what *they* would call the problem of free will. I won't identify them and the topic of their researches by reference to the phrase 'free will', however. For my part, I'll say that they're the philosophers whose writings are full of phrases like 'semi-compatibilism' and 'transfer of powerlessness' and 'and 'alternative possibilities' and 'ultimate responsibility'. Let's call them the Group. I have lately taken to accusing the members of the Group of ruining the term 'free will'. And, because I think it *has* been ruined, I have sworn never to use it myself. I'm breaking that oath today simply because this issue of the *Belgrade Philosophical Annual* is devoted to "the philosophical analysis of the problem of free will," and, in this venue, I don't want to get into my quarrel with the Group. But understand this. 'Free will' is a technical term. (It occurs in everyday discourse almost entirely in the phrases 'of my own free will', 'of her own free will', and so on. And 'I did it of my own free will' means no more than that my doing it was uncoerced.) 'Free will' was once a *useful* technical term like 'Abelian group' and 'moment of inertia' and 'negotiable instrument'—useful because unambiguous. And what '*x* has free will' used to mean was that, sometimes at least, when *x* is in the process of deciding which of two incompatible courses of action to pursue, *x* is able to perform each of them. But at some point the members of the Group began to say things like this:

> Free will seems to require that *open alternatives* or *alternative possibilities* lie before us—a garden of forking paths—and it is "up to us" which of these alternatives we choose. But . . . free will also seems to require that the *sources* or *origins* of our actions lie "in us" rather than in something else (such as the decrees of fate, the foreordaining acts of God, or antecedent causes or laws of nature) outside us and beyond our control.[3]

That is, it became common for members of the Group to make assertions that began like this (more or less):

> Free will seems to require that *open alternatives* lie before us. But free will *also* seems to require that . . .

3 Robert Kane, "Libertarianism" in John Martin Fischer, Robert Kane, Derk Pereboom, and Manuel Vargas, *Four Views on Free Will* (Oxford: Blackwell Publishing Ltd, 2007), pp. 5-43. The quoted passage is on p. 14.

These assertions began like that, but the additional requirement that followed this common beginning varied, and, in consequence, 'free will' became in due course that useless thing, an ambiguous term of art. (The ruin was aggravated by the fact that many members of the Group insisted on saying that 'open alternatives' meant alternatives of the kind, whatever that kind might be, that are required by ascriptions of moral accountability.)

My contention is that it is a mistake to suppose that free will can *require* anything. If free will "seems to require" alternative possibilities, that is only because, in those days when it was still a useful technical term, that is how it was defined—that was its *stipulative* definition. Saying that free will seems to require alternative possibilities is like saying that locomotion seems to require change of position in space. And to say that free will seems to require not only alternative possibilities but something else as well is like saying that locomotion seems to require not only change of position in space but something else as well. That is, to say the former is at best to say something trivial, and to say the latter is to say something either (depending on how you look at it), either trivially false or entirely meaningless.

Whatever the justice of these remarks may be, you at any rate know what thesis the phrase 'the free-will thesis' refers to when *I* use it. Let us now return to the question whether determinism is compatible with that thesis.

2.

I would suppose that one should believe that determinism is incompatible with the free-will thesis only if one believes that it is possible formally to deduce the denial of the free-will thesis from determinism—or at least from determinism and some other premises that one believes to be both uncontroversial and necessarily true. But anyone who searches out such a deduction faces the problem of the vocabulary gap. The prospect of formally deducing thesis A from thesis B is dim if the two theses are stated in vocabularies whose overlap comprises only logical constants—and determinism is stated in words that refer to worlds and propositions and the free-will thesis in words that refer to agents and their abilities.

In constructing the second version of the Consequence Argument I attempted to bridge the vocabulary gap by treating agents' abilities to perform acts as their "having access to" a possible world in which they do perform them. So, for example, instead of saying, "Sally is able to be in Chicago by noon tomorrow", one might say, "Sally has access to a world in which she will be in Chicago at noon tomorrow", and instead of saying, "And Willie is able to be there before Sally", one might say "And Willie has access to a world in which he will be in Chicago, at a certain moment, and neither at that moment nor at any earlier future moment will Sally be in Chicago".

I should note that, in order to make my sentences as simple as possible, in the remainder of this talk I will suppress the time variable in almost all the statements about agents and their abilities that I am going to make or consider. I expect that anyone who is likely to read this essay know how to undo this suppression by sticking expressions like '$\forall t \forall t'$ (t is a time & t' is a time →' on at the left-hand side of my statements, and sprinkling them liberally with 'at t' and 'at t''.

The free-will thesis, stated in terms of access to worlds is this:

> There is an agent x who is trying to decide between the incompatible courses of action y and z, and x has access to a world w_y in which x performs y and to a world w_z which x performs z.

Now since y and z are incompatible, w_y and w_z must be distinct worlds. Therefore—since there is only one actual world—the free-will thesis entails the *minimal free-will thesis*,

> There are agents who have access to some non-actual world,

and the denial of the minimal free-will thesis entails the denial of the free-will thesis. To show that the denial of the free-will thesis can be deduced from determinism, it therefore suffices to show that the denial of the minimal free-will thesis can be deduced from determinism, and that was what the second version of the Consequence Argument promised to do. It will be instructive to see how the attempt worked. We start with a few abbreviations.

α	the actual world

(Note that 'α' is not a proper name of the actual world; it is an abbreviation of the definite description 'the actual world'.)

Hxy	(the agent) x has access to (the world) y
Lxy	the laws of nature are the same in (the world) x and (the world) y
Sxy	(the worlds) x and y share a momentary state.

Determinism is then the thesis

$$\forall x\, (\mathrm{L}x\alpha \;\&\; \mathrm{S}x\alpha \,.\!\rightarrow x = \alpha),$$

and the minimal free-will thesis is

$$\exists x \exists y(\mathrm{H}xy \;\&\; y \neq \alpha).$$

The denial of the minimal free-will thesis is not deducible from determinism, but it is deducible from determinism and these two premises—metaphysical assumptions, I'll call them.

MA*a*	$\forall x \forall y\, (\mathrm{H}xy \rightarrow \mathrm{L}y\alpha)$
MA*b*	$\forall x \forall y\, (\mathrm{H}xy \rightarrow \mathrm{S}y\alpha)$.

That is, any world that anyone has access to has the same laws as the actual world; any world that anyone has access to is in the same state as the actual world at some time.

The deduction is trivial. The denial of the minimal thesis is equivalent to

$$\forall x \forall y\ (\mathrm{H}xy \rightarrow y = \boldsymbol{\alpha}).$$

Assume 'H*xy*'. From the two metaphysical assumptions, we have 'L$y\boldsymbol{\alpha}$' and 'S$y\boldsymbol{\alpha}$'. And from determinism

$$\mathrm{L}y\boldsymbol{\alpha}\ \&\ \mathrm{S}y\boldsymbol{\alpha}\ .\rightarrow y = \boldsymbol{\alpha}.$$

And so we have '$y = \boldsymbol{\alpha}$', and the deduction—the second version of the Consequence Argument, or near enough—is complete. If one is willing to say that the two metaphysical assumptions are not only true but necessarily true, one will probably be willing to describe the deduction as a purported demonstration of the incompatibility of free will and determinism.

The most important reason for refusing to regard the deduction as really being a demonstration of the incompatibility of free will and determinism is that the minimal free-will thesis is predicated on the idea that our abilities can be described in terms of access to individual possible worlds. And, sadly, our abilities are not so fine-grained as that. That is to say, there is no world—or at any rate, no world but the actual world—and no action that anyone is able to perform, such that if someone did perform that action, that world would be actual[4]. No action any finite being is able to perform is such that, for some possible assignment of truth-values to propositions, its performance would result in that assignment being the actual assignment. Or at any rate, not unless that assignment *were* the actual assignment.

Very well then, it is at best highly implausible to suppose that any agent has access to any one non-actual world. And, therefore, it is highly implausible to suppose that the minimal free-will thesis (so-called) is true. But if it is highly implausible to say that Sally has access to any given world in which she, say, owns a Shiba Inu, it is at least *more* plausible to say that she has access to the *set* of worlds in which she owns a Shiba Inu—or to say simply that she has access to *the worlds* in which she owns a Shiba Inu. To say, that is, that there is something Sally is able to do, such that, if she did that

4 Suppose I am going to take a nap this afternoon. Let '**A**' be a proper name of the actual world (of the world that happens to be actual). The 'would' conditional 'If I were to take a nap this afternoon, **A** would be the actual world' is true (according to the Lewis semantics) on the assumption of "strong centering"—the assumption that, for every world *w*, *w* is closer to w than is any other world. (According to Lewis, that conditional is true just in the case that there is a 'nap this afternoon/**A** is actual' world that is closer to actuality than is any 'nap this afternoon/A is not actual' world. And, of course, that is "automatically" the case if I am going to take a nap this afternoon, A is the actual world, and **A** is closer to **A** than is any other world.) In my view, this result is an excellent reason to reject strong centering.

thing, *some* world in which she owns a Shiba Inu would be actual—which of course, does not imply that any of the presumably innumerably many worlds in which she owns a Shiba Inu is the one that would be actual if she did that thing.

More generally,

> x has access to the worlds in which $p =_{df}$ There is something x is able to do that is such that, if x did that thing, some world in which p would be actual.

Suppose, for example, that I am now able to raise my left arm. There is no world that would be actual if I were now to raise my left arm, but one of the worlds in which I now raise my left arm would be actual if I now raised my left arm. (For that matter, one of the worlds in which I either raise my left arm or bring about world peace would be actual if I now raised my left arm: if x has access to the worlds in which p, x has access to the worlds in which $p \vee q$).)

The free-will thesis, we remember, was this:

> There are agents who are trying to decide between incompatible courses of action and are able to perform each of them.

Or, stated in terms of access to pluralities of worlds:

> There is an agent x who is trying to decide between the incompatible courses of action y and z, and x has access to the worlds in which x undertakes y and x has access to the worlds in which x undertakes z.

Now since y and z are incompatible, no world is both one of the worlds in which x undertakes y and one of the worlds in which x undertakes z. Therefore—since there is only one actual world—the free-will thesis entails the *new minimal free-will thesis* (hereinafter, the minimal free-will thesis):

> There are agents and there are worlds, none of which is actual, such that those agents have access to those worlds.

Most theses involving access to worlds will be easier to state and to examine if they are stated with the use of plural variables— 'the *x*s', 'the *y*s', 'the *z*s', and so on. (Variables, that is, that stand to 'they'/'them' as '*x*', '*y*', and '*z*' stand to 'it'.) In this essay, plural variables range only over worlds. Using plural variables, we can define the notion of access to certain worlds as follows:

> x has access to the ys $=_{df}$ There is something x is able to do that is such that, if x did that thing, one of the ys would be actual.

Since the free-will thesis entails the minimal free-will thesis, to show that determinism is incompatible with the free-will thesis, it suffices to show that it's incompatible with the minimal free-will thesis. This we will proceed to do.

We begin with two more abbreviations:

H*x* *y*s $=_{df}$ (the agent) *x* has access to the *y*s
x o— *y*s $=_{df}$ *x* is one of the *y*s.

‘Is one of’ is in a sense the “plural analogue” of ‘is a member of’. And the plural analogue of ‘is a subset of’ is ‘are among’, which is abbreviated as ‘’:

*x*s ↲ *y*s $=_{df}$ ∀*z* (*z* o— *x*s → *z* o— *y*s)

Next, we introduce plural quantifiers to bind the plural variables:

‘λ’ (‘λ*x*s’ is pronounced ‘for any *x*s’)
‘E’ (‘E*x*s’ is pronounced ‘for some *x*s’ or ‘there are *x*s such that’)

Having these abbreviations and the plural existential quantifier at our disposal, we can write the minimal free-will thesis this way:

∃*x* E*y*s (H*x* *ys* & ~ α o— *y*s).

Now assume that I am at this moment able to raise my left arm *and* assume that there are worlds in which the proton-electron mass ratio is not the actual ratio; it follows from these two assumptions that

∃*x* E*y*s E*z*s (H*x* *y*s & *z*s ↲ *y*s & ~ H*x* *z*s)

—for I have access to the worlds in which either I now raise my left arm *or* the proton-electron mass-ratio is not the actual ratio, the worlds in which the proton-electron mass-ratio is not the actual ratio are among those worlds, and I do not have access to the worlds in which the proton-electron mass-ratio is not the actual ratio. (Or here is a less exotic example. Suppose I am able to toss a dart in such a way that, if I were to toss it that way, it would strike the target. And suppose that none of the ways in which I am able to toss a dart is such that, if I were to toss a dart *that* way, it would strike the bull’s-eye—although there are, of course, ways in which I am able to toss a dart that are such that, if I were to toss a dart that way, it *might* strike the bull’s-eye. The worlds in which I toss the dart and it strikes the bull’s-eye are properly among the worlds in which I toss it and it strikes the target, and I have access to the worlds in which I toss it and it strikes the target, and I do not have access to the worlds in which I toss it and it strikes the bull’s-eye.)

Let us say that a world *x* is a *might-world* for an agent *y* just in the case that, if *y* did a certain thing (which *y* is able to do) *x* might be actual. If an agent has access to certain worlds, it does not follow that those worlds are all might worlds for that agent—for if it did follow, then there would be something I am able to do such that, if I were to do it, the proton-electron mass ratio might be different from what it actually is. But it does follow that *some* of them must be might worlds for that agent: Whatever world would have turned out to be actual if I had just raised my left hand, it would have been one of the worlds that might have been actual if I had just raised my left hand.

And let us say that a world is a might-world (without qualification) if it is a might-world for someone. Now for some worlds it is a contingent matter whether those worlds are might-worlds. A world x that is not a might-world in actuality may well be a might-world in some world y that is closer to x than $\boldsymbol{\alpha}$ is. Obviously, if an agent has access to the xs, then some of the xs are might-worlds in the actual world—a principle I'll call the 'Might' Requirement. If we allow ourselves one final abbreviation:

$$\mathrm{M}xy =_{df} x \text{ is a might-world in the world } y$$

we can state the 'Might' Requirement this way:

$$\text{MR} \qquad \forall x\, \lambda y\mathrm{s}\, (\mathrm{H}x\, y\mathrm{s} \rightarrow \exists z\, (z \circ\!\!-\, y\mathrm{s}\ \&\ \mathrm{M}z\boldsymbol{\alpha})).$$

We now state two metaphysical assumptions—the plural analogues of MA*a* and MA*b*.

$$\text{MA}A \qquad \forall x \forall y\, (\mathrm{M}xy \rightarrow \mathrm{L}xy)$$

$$\text{MA}B \qquad \forall x \forall y\, (\mathrm{M}xy \rightarrow \mathrm{S}xy)$$

The denial of the minimal free-will thesis—that is

$$\forall x\, \lambda y\mathrm{s}\, (\mathrm{H}x\, y\mathrm{s} \rightarrow \boldsymbol{\alpha} \circ\!\!-\, y\mathrm{s})$$

is deducible from determinism, MR, MA*A*, and MA*B*. (The deduction, if not trivial, is at any rate straightforward.)

For assume '$\mathrm{H}x\, y\mathrm{s}$'. MR gives us

$$\mathrm{H}x\, y\mathrm{s} \rightarrow \exists z\, (z \circ\!\!-\, y\mathrm{s}\ \&\ \mathrm{M}z\boldsymbol{\alpha}),$$

and then we have, by MP and EI

$$z \circ\!\!-\, y\mathrm{s}\ \&\ \mathrm{M}z\boldsymbol{\alpha}.$$

'$\mathrm{M}z\boldsymbol{\alpha}$' and our two metaphysical assumptions yield

$$\mathrm{L}z\boldsymbol{\alpha}\ \&\ \mathrm{S}z\boldsymbol{\alpha},$$

and this together with determinism gives us

$$z = \boldsymbol{\alpha},$$

which, with '$z \circ\!\!-\, y\mathrm{s}$', yields

$$\boldsymbol{\alpha} \circ\!\!-\, y\mathrm{s},$$

and the deduction is complete. (A formal deduction with all the bells and whistles and dotted 'i's and crossed 't's is provided in the Appendix.)

But is the deduction a demonstration of the incompatibility of determinism and the minimal free-will thesis? Obviously it's not a demonstration of the *logical* incompatibility of determinism and the minimal free-will thesis in the strict sense of 'logical', owing to the fact that the deduction has three premises other than determinism—MR, MA*A*, and MA*B*. And it's not a demonstration of that proposition in the loose sense of 'logical', for, although MR can plausibly be said to be a conceptual or analytic truth, MA*A* and MA*B* are substantive metaphysical theses.

The possibility remains that the deduction demonstrates that determinism and the minimal free-will thesis are *metaphysically* incompatible—that in no deterministic world *x* is there a *y* and are there *z*s such that *y* has access to the *z*s, and *x* is not among the *z*s. This will be the case only if MA*A* and MA*B* are necessary truths. It seems evident, however, that any argument for the truth of either MA*A* or MA*B* will be an argument for its necessity, owing to the fact that each of these principles is necessarily true if it is true at all. For say that a binary relation on worlds is *internal* if whether it holds between two worlds does not depend upon which world (among all worlds) is actual. (The relations expressed by '*x* is actual and *y* is not identical with *x*' and '*x* is closer to actuality than *y* is' are examples of relations that are not internal.) MA*A* says of a certain internal relation *R* that it holds in **α** between any two worlds *x* and *y*. Since *R* is internal, if it holds between *x* and *y* in any world, it holds between *x* and *y* in any other world. (Essentially the same argument applies to MA*B*.)

But even if MA*A* and MA*B* are both true in all possible worlds, our deduction is not a *demonstration* of the metaphysical incompatibility of determinism and the minimal free-will thesis unless it is self-evident or provable that both theses have that modal status.

In the next part of this talk, I will present arguments for the truth—and hence the metaphysical necessity—of MA*B*. (A defense of MA*A* must be reserved for another occasion.) I do not suppose that I can convince those who reject MA*B* that they are mistaken. My only purpose in presenting these arguments is to ensure that those who reject MA*B* understand why some incompatibilists believe that MA*B* is true—and, in fact, self-evident.

3.

Carl Ginet has put a familiar thought in a very striking way: "Our freedom is the freedom to add to the given past."[5] But striking statements are not often statements that can figure in formal deductions. Still, anyone who finds Ginet's Principle plausible will almost certainly find the following statement (a statement that is not striking at all) plausible.

GP All the worlds that are might-worlds in the world *x* are worlds whose past is the past of *x*.

(The label 'GP' is not an abbreviation of 'Ginet's Principle'—that is, I do not maintain that GP is Ginet's Principle.) And, of course, all worlds in which the

5 *On Action* (Cambridge: Cambridge University Press, 1990), pp. 102–103.

past is the past of *x* are worlds that share some of their momentary states with *x*, and therefore,

> All the worlds that are might-worlds in the world *x* are worlds that share some of their momentary states with *x*

—that is, MA*B* is a consequence of GP. But why should we accept 'Ginet's Principle' or any of the principles that it suggests? Why should we suppose that all worlds that are might-worlds in a world *x* are worlds that share some of their momentary states with *x*?

Suppose Ingrid thinks that GP is false. Suppose, in fact, that she thinks that some worlds that are might-worlds for her share *no* state with the actual world. In that situation, Ingrid might well reason as follows. Reflecting on the thesis that some worlds whose past is not the actual past are might-worlds for her, she says to herself,

> There is something I'm able to do such that, if I did it, a world whose past is not the actual past might be the actual world.

Having had this thought, she indulges in further reflection:

> Let 'Clio' be a proper name of the actual past. Something I am able to do is such that, were I to do it, I might find myself in a situation in which I could say truly, "Clio is not the actual past"—or if I make allowance for the fact that that "something" will take a little time to do—"Clio is not a *part* of the actual past".

In my view, Ingrid's reasoning is correct; that conclusion does follow from the denial of GP. And so much the worse for the denial of GP, for that conclusion is impossible. Or so *I* say. But it is not my position that those who dispute this statement are denying the obvious.

There is more.

Ginet's Principle has a sister, something like this,

> Our freedom is the freedom to act in accordance with the laws of nature.

And GP has a sister as well:

> ML All the worlds that are might-worlds in the world *x* are worlds whose laws are the laws of *x*.

Suppose that Ingrid embraces both ML and soft determinism—the conjunction of determinism and the free-will thesis. (Compatibilism is of course the thesis that soft determinism is possibly true.) She must conclude that she has access to the worlds that share no momentary state with the actual world, and hence to the worlds in which the past is not the actual past. And she must accept the following first-person statement:

> Something I am able to do is such that, were I to do it, I *would* find myself in a situation in which I could say truly, "Clio is not (and is not a part of) the actual past".

It seems to me to be quite impossible to suppose that anyone has such an ability. But, again, I won't press that point. I will say once more that it's not my position that anyone who disputes this statement is denying the obvious. I will instead say something that seems to me to be right even if what I have said in defense of Ginet's principle is wrong.

The point I want to make also pertains to soft determinism. Soft determinism, I contend, implies that some of us at least—and presumably all of us—are able to work miracles. Suppose for example that it is now causally determined that Tim will have three martinis at the reception tonight. (That is, suppose that in every possible world in which all matters now stand exactly as they actually do, and in which the laws of nature are the actual laws, Tim has three martinis at the reception tonight.) Tomorrow, Sophie the soft determinist clicks her tongue and tells Tim that he ought to have stopped at two. Tim points out that it was determined from all eternity (or from the earliest moments of time) that he'd have three. Sophie, being a determinist, agrees—but being a *soft* determinist, she insists that he was nonetheless *able* to stop at two. She thereby ascribes to Tim the power to work miracles, for the power now to prevent or forfend that which is now causally determined is the power to work miracles.

So, at any rate, I maintain. It is sometimes said or implied that a sentence of the form 'I am able to bring it about that *p*' is a claim to miraculous powers only if the corresponding sentence of the form 'It is contrary to the laws of nature that *p*' is true. But there are simple and obvious counterexamples to this conditional. If I were to say—say *here*, *now*, in *these* circumstances—, "I am able to be in Jerusalem 30 seconds from now", I'd be claiming to be able to perform a miracle. Or if someone commanded me—commanded me *here*, *now*, in *these* circumstances—to be in Jerusalem thirty seconds from now, I'd be able to obey that command only if I were able to work miracles. And yet it is not contrary to the laws of nature that I should be in Jerusalem 30 seconds from now. There are many, many possible worlds that have the same laws as the actual world and in which I'm in Jerusalem 30 seconds from now. (Of course, in those worlds, if I wasn't in Jerusalem to start with, I had set out for Jerusalem some time ago.)

H. G. Wells wrote a story called "The Man Who Could Work Miracles". If I were asked to define the phrase 'human being who is able to work miracles', I should insist that, whatever the precise contours of the definition might be, it should have this consequence: A human being who is, at *t*, able to prevent the later occurrence of an event that is, at *t*, causally determined later to occur has to count as someone who is able to work miracles.

And yet again I say: So, at any rate, *I* maintain—for some philosophers have objected for various reasons to the use I have made of the word 'miracle' and the phrase 'able to work miracles'. I do not think these objections are good ones, but, rather than defend my choice of words, I'll do what I generally do in the face of such accusations, and drop the word and make my point using other language.

Let us say that an agent now has an *ancestrally contra-causal ability* if he or she is able to do a certain thing at some future time and that that agent will not do that thing unless at some point later than the present a causally impossible event occurs—that is, an event that follows *this* state of affairs only in worlds in which the laws of nature are different (or miraculously violated, if such worlds there be).

There is a logical point that needs to be made about this way of speaking. Suppose Sally will say "Yes" when the question is put to her at *t* only if a causally impossible event happens between now and *t*. That does not imply that if Sally says "Yes", she can go on to say, and to say truly, "My goodness—a causally impossible event just happened". For 'unless at some point later than the present a causally impossible event occurs' means 'unless at some point later than the present a certain event occurs—an event that is as things stand causally impossible'. Generally speaking, most causally impossible events would not be causally impossible if they happened.

But let us look at an example. It is now 5:00 p.m. and it is causally determined that Tim will have a third martini at 7:07 p.m. If I say, "It's now 5:00 p.m., and the past and the laws entail that Tim will have a third martini at 7:07 p.m. Therefore, Tim will not refrain from having a third martini at 7:07 p.m. unless some causally impossible event occurs between now and 7:07 p.m.," I put forward a sound argument: its premise is true, and its conclusion follows from its premise. And if I make the further inference, ". . . and, therefore, if Tom is able to refrain from having a third martini at 7:07 p.m., he is able to do a certain thing at 7:07 p.m. that he will not do unless some causally impossible event happens between now and 7:07 p.m.," my reasoning is valid.

And what are we to say of such an ability? Can it be that Tim is able to do something that he will not do unless some causally impossible thing happens first? (That is: . . .for some *x*, he will not do that thing unless *x* happens first—and *x* is *in fact* causally impossible. For if he did do that thing, *x* would not be causally impossible.) Well, suppose that on a certain occasion next week, you are going to be asked an important question—a question to which you must respond by saying either "Yes" or "No". You are a soft determinist, and for some reason you believe that you are going to say, "Yes". (Maybe you believe that a time-traveler from the future told you this; maybe you believe that God has told you this; maybe you believe that fortune sticks and hexagrams from the *I Ching* have told you this . . .) You therefore believe that you are not going to say "No" unless something causally impossible happens. Your best friend, whose life, fortune, and sacred honor depend on your answer comes to you and tearfully begs you to promise that you'll say "No". You reply, "But I can't honestly make that promise because I know I wouldn't be able to keep it. For I'm not going to say 'No' unless something causally impossible happens, and nothing causally impossible is going to happen. And I mean, nothing that is *actually* causally impossible is going to happen. I realize that if it *did* happen, it wouldn't be causally impossible." Your friend replies, "But you're a *soft* determinist like me. You therefore believe that the fact that you're not going to say 'No' unless something causally impossible happens does not imply that you're unable to say 'No'. So it follows from your own philosophy that when you said that you wouldn't be able to keep the promise I want you to make, you were mistaken—or at least there's no reason to suppose you were right."

And that's something to think about. Suppose you do think about what your friend has said, and your thoughts are as follows:

> Hmm. I know I'm a going to say "No", unless some causally impossible event occurs between the present moment and the moment at which I must say either "Yes" or "No". I realize, however, that my friend is right to say that my soft determinism implies that that statement is perfectly consistent with the statement that I am *able* to say "No". How shall I frame my explanation of why it is that I refuse to promise to say "No"?

That is a very good question, and I'm not at all sure what the best soft-determinist answer to it is.

Libertarianism—that is, the conjunction of incompatibilism and the free-will thesis—is sometimes described as affirming the reality of "contra-causal freedom". This is a canard: it is soft determinism that implies the reality of contra-causal freedom: every soft determinist believes that there are agents who are now able to do things that they will not do unless some causally impossible thing happens first. Libertarians, by contrast, believe that everything we do is consistent with the past and the laws of nature.

My advice to the compatibilist is not to dispute this consequence, but rather to accept it and attempt to show that it is harmless. For the compatibilist can certainly say the following, or something very much to the same purpose:

> All right, it's a consequence of compatibilism that free agents in a deterministic world have what you have called ancestrally contra-causal abilities. I suppose you could even say that they're able to work miracles. But it doesn't follow that, in a deterministic world, you could expect free agents to raise the dead or turn water into wine. For one thing, if the miracles the free agents would be able to work actually occurred, they wouldn't look like miracles. I mean they wouldn't be events that *obviously* couldn't have evolved naturally from Ginet's

"given past". They'd be more like the miracles Hume described in a well-known passage in *On Miracles*:

> A miracle may either be discoverable by men or not. This alters not its nature and essence. The raising of a house or ship into the air is a visible miracle. The raising of a feather, when the wind wants ever so little of a force requisite for that purpose, is as real a miracle, though not so sensible with regard to us.

More importantly, however, these events *won't* actually occur. For if someone has an ancestrally contra-causal ability, it's determined that that ability will not be exercised. If, for example, it's determined that Tim is going to have three martinis at the reception tonight and he's able to stop at two, then it's determined that he not exercise or act on that ability.

The Interlocutor asks, "But what would happen if he *did* act on his ability to stop at two?—given that it was determined that he was going to have three, I mean?"

And I reply, "Why, he'd have two. That's what it says on the label. In fact, the two counterfactual conditionals

> If Tim acted on his ability to stop at two martinis, he would not have three martinis
>
> If Tim acted on his ability to stop at two martinis, it would not be causally determined that he would have three martinis

are both true."

And that exchange comprises a no more puzzling or counter-intuitive series of statements than this one:

> Sally says, "It's causally determined that this pane of window glass will never break."
>
> Bertie responds, "But what would happen if it were struck by a thrown brick?"
>
> And Sally replies, "Why, it would break, of course. These two counterfactual conditionals
>
> > If this pane of glass were struck by a thrown brick, it would break
> >
> > If this pane of glass were struck by a thrown brick, it would not be causally determined that it will never break."
>
> are both true."

My *Doktorvater*, Richard Taylor, thought that the concept of an ability such that it's causally determined that that ability not be exercised was self-contradictory. He would say things along the following lines.

> An ability that is, of necessity, never exercised is no ability at all. I grant that if someone *were* able to prevent causally determined events from occurring, that would be an ability that it was causally determined would not be exercised. But if someone does claim to have that ability, it's perfectly in order for me to respond by saying, "All right, if you're able to do that, let's see you do it. Go ahead. Prevent the occurrence of something whose future occurrence is causally determined". It's evident that to be able to do *x* and to be able to obey a command to do *x* are the same ability. And no one is able to obey the command, "Prevent the occurrence of something whose future occurrence is causally determined".

This was not one of Richard's better arguments. For one thing, it proves too much. Anyone, compatibilist or incompatibilist, who accepts the free-will thesis believes that we are sometimes able to do things other than the things we are actually going to do. Imagine that someone said to a believer in free will

> All right, if you're able to do that, let's see you do it. Go ahead. Do something that isn't one of the things you're actually going to do.

I grant that the believer would be unable to obey the command, 'Do something that isn't one of the things that you're actually going to do'. But that inability does not imply that statements like the following can't be true:

> Alice is not actually going to teach baby logic in the spring, and Alice is able to obey the command, 'Teach baby logic in the spring'.
>
> Tim is actually going to have a third martini, and Tim is able to obey the command, 'Do not have a third martini'.
>
> Robin is not actually going to testify at James's trial, and Robin is able to obey the command, 'Testify at James's trial'.

And, similarly, although no one is able to obey the command, 'Prevent the occurrence of something whose future occurrence is causally determined', it doesn't follow that statements like

> It's causally determined that Alice is not going to teach baby logic in the spring, and Alice is able to obey the command, 'Teach baby logic in the spring'.

can't be true. (*If*, as I believe, such statements can't be true, *that* they can't be true doesn't follow from the true statement, 'No one is able to obey the command, "Prevent the occurrence of something whose future occurrence is causally determined."')

To sum up my final point, soft determinism implies contra-causal freedom, the ability to do things that one will do only if something causally impossible happens—the necessarily never-exercised ability to do things that follow upon the causally impossible but which, if done, would not follow upon the causally impossible. And compatibilism—that is, the thesis that soft determinism is true in some possible worlds—implies that *there* might be agents who were able to do things that they would do only if the causally impossible happened.

The compatibilist will reply—at any rate, this is reply I recommend to the compatibilist—"All right, yes, my position has those consequences. So sue me. They sound a little odd when they're put the way you've put them. But putting them that way is a rhetorical, not a philosophical achievement. I'm willing to put up with them".

And there's no resolving this issue. This is, after all, philosophy. Refutations of philosophical positions are rare indeed. (As David Lewis once said, "Gödel and Gettier may have done it".) What we can do is count the costs. That's what I see myself as having been doing in the final section of this paper: counting costs.[6]

6 This essay is based on a paper composed for and delivered at a conference on free will and determinism that was held at Rutgers University in May of 2022. I thank Adrian Kuźniar for valuable comments and criticisms.

APPENDIX
FORMAL DEDUCTION OF THE DENIAL OF THE FREE-WILL THESIS FROM DETERMINISM, MR, MA*A*, AND MA*B*

The format of the deduction is essentially the "box and cancel" format of Kalish and Montague, *Logic: Techniques of Formal Reasoning.*

In this deduction, the annotation *PUI* stands for 'the Rule of Plural Universal Instantiation'. This rule may be stated as follows:

> At any point in a derivation, an earlier line is said to be *accessible* if it is then unboxed and does not contain an uncanceled '*Show*'
> If α_S and β_S are plural variables and φ and ψ are formulas, ψ may occur as a line at any point in a derivation if ⌜ $\lambda\alpha_S\ \varphi$ ⌝occurs in an earlier line that is accessible line and β_S does not occur in φ and ψ is the result of substituting β_S for all free occurrences
> of α_S in φ

Example:

12. a *ys* (F *ys* → ~ (G *ys* ∨ H *zs*))

.

.

.

18. F *xs* → ~ (G *xs* ∨ H *zs*)) *12, PUI*

And the deduction is

1.	~~*Show*~~ ∀*x* λ*ys* (H*x ys* → **α** o— *ys*)	*Universal Proof*
2.	H*x ys*	*Asspt for Cond Proof*
3.	∀*x* λ*ys* (H*x ys* → ∃*z* (*z* o— *ys* & M*z***α**))	*Premise* [MR]
4.	H*x ys* → ∃*z* (*z* o— *ys* & M*z***α**)	*3 UI, PUI*
5.	*z* o— *ys* & M*z***α**	*2, 4, MP, EI*
6.	∀*x*∀*y* (M*xy* → L*xy*)	*Premise* [MA*A*]
7.	∀*x*∀*y* (M*xy* → S*xy*)	*Premise* [MA*B*]
8.	M*z***α** → **.** L*z***α** & S*z***α**	*6,7, UI, UI, p→q.& p→r:→ . p→.q&r*
9.	L*z***α** & S*z***α**	*5, Simp, 8 MP*
10.	∀*x* (L*x***α** & S*x***α** **.** → *x* = **α**)	*Premise* [Determinism]
11.	*z* = **α**	*9, 10 UI, MP*
12.	**α** o— *ys*	*5, Simp, 11 Euclid's Law*