Various philosophers have formulated arguments for the incompatibility of freedom and causal determinism.\(^1\) Carl Ginet has recently presented such an argument in a particularly lucid way.\(^2\) It has been noted that this sort of argument can be readily transformed into an argument for the incompatibility of moral responsibility and determinism.\(^3\) I shall set forth in a compact way the structure of Ginet’s argument. I then show how the compatibilist might respond to both incompatibilist arguments—about determinism and freedom and determinism and responsibility.

1. GINET’S ARGUMENT

I shall discuss a reformulated version of the argument presented by Ginet in his article, ‘The conditional analysis of freedom’.\(^4\) In order to set out Ginet’s argument, some definitions are necessary.

Determinism (D) is the thesis that, for any given time, a complete description of the state of the world at that time, together with a complete statement of the laws of nature, entails every truth as to what happens after that time.

\[\Diamond_t^S p = \text{df. It was in } S\text{'s power at } t\text{ to make it the case that } p.\]

\(S\) made it the case that \(p\) just in case \(p\) is true and there was some action of \(S\)’s, \(S\)’s \(V\)-ing, such that either \(p\) entails that \(S\) \(V\)-ed or there is some proposition \(q\) such that \(S\)’s \(V\)-ing caused it to be the case that \(q\) and \(p\) entails \(q\).\(^5\)

\[\boxdot_t^S p = \text{df. } p \& \sim \Diamond_t^S \neg p\]
\[ [P]_t p = \text{df. } p \land (S) [P]_t p \text{ (i.e., it was in no one's power at } t \text{ to make it not the case that } p) \]

\[ [P] p = \text{df. } p \land (S)(t) [P]_t p \text{ (i.e., it was never in anyone's power to make it not the case that } p) \]

'\(a_t\)' is a variable that ranges over propositions as to what happens at or after \(t\).

'\(b_t\)' is a variable that ranges over propositions at to what happens before \(t\).

'\(g\)' is a variable that ranges over universal propositions.

'\(p\)' is a variable that ranges over propositions about what happens.

Now we introduce the following assumptions about the logic of the operator, '[P]'.

\[(\alpha) \quad \Box (\Box p \supset [P] p)\]
\[(\beta) \quad \Box \left[ [P] (p \supset q) \land [P] p \supset [P] q \right] ;
\quad (t) \Box \left[ [P]_t (p \supset q) \land [P]_t p \supset [P]_t q \right]\]
\[(\gamma) \quad \Box (g)(g \text{ is entailed by the laws of nature } \supset [P] g)\]
\[(\delta) \quad (t)(b_t) \Box (b_t \supset [P]_t b_t)\]

Assumption (\(\beta\)) is a reformulation of a crucial rule of inference Ginét adopts in his proof; Ginét calls the rule, "modus ponens for relative power necessity":

\[ [P]_t S (p \supset q) \land [P]_t S p \therefore [P]_t S q \]

Ginet defends this rule (and thus also (\(\beta\))) as follows:

Surely this is a valid form of inference, on any reasonable understanding of 'It was in S's power at \(t\) to make it the case that' (or 'At \(t\) it was open to S to make it the case that' or 'At \(t\) S could have made it the case that'). Suppose, for example, that (first premise) if \(p\) it rained this afternoon then \(q\) the fresh paint on the house is ruined, and it was not in S's power this morning \((t)\) to make it the case that it would rain this afternoon but the paint would not be ruined. Suppose also that (second premise) it did rain this afternoon and it was not in S's power this morning to make it the case that it would not rain this afternoon. Surely it follows that (conclusion) the paint is ruined and it was not in S's power this morning to make it the case that the paint would not be ruined.6

(\(\gamma\)) expresses our powerlessness to determine the laws of nature, and (\(\delta\)) expresses our powerlessness to determine the past.
Ginet's argument can now be presented succinctly as follows:

(I) \( \square \{ D \supset (t)(a_t) \ [ a_t \supset (\exists b_t)(\exists g) \} \) such that:
1. \( b_t \& \) the laws of nature entail that
   \( g \& \square (g \& b_t \supset a_t) \)
2. \( \therefore \square (g \supset b_t \supset a_t) \)  
3. \( \therefore \square (g \supset b_t \supset a_t) \)  
4. \( \therefore \square g \)  
5. \( \therefore \square (b_t \supset a_t) \)  
6. \( \therefore \square (b_t \supset a_t) \)  
7. \( \therefore \square (b_t \supset a_t) \)  
8. \( \therefore \square (\square, a_t) \} \)

(II) \( \square [ D \supset (t)(a_t)(a_t \supset [ \square_t a_t ] ) ] \)

(III) \( \square [ D \supset (t)(p)(p \supset [ P_t p ] ) ] \)

(IV) \( \square [ D \supset (p)(p \supset [ P_t p ] ) ] \)

II. EXTENSION OF THE ARGUMENT

The argument can be transformed into an argument for incompatibilism about determinism and moral responsibility as follows:

\[ \mathit{R} p = \text{df. } p \& \text{ no one is responsible for its being the case that } p. \]
\[ \mathit{R} S p = \text{df. } p \& S \text{ is not responsible for its being the case that } p. \]

'\( a_s \)' ranges over propositions as to what happens after \( S \) begins to exist.

'\( b_s \)' ranges over propositions as to what happens before \( S \) begins to exist.

The assumptions are easily transformed as follows:

\( \alpha' \) \( \square ( \square p \supset \mathit{R} p ) \)
\( \beta' \) \( \square ( \mathit{R} (p \supset q) \& \mathit{R} p \supset \mathit{R} q ) \)
\( \gamma' \) \( \square (g) (g \text{ is entailed by the laws of nature } \supset \mathit{R} g) \)
\( \delta' \) \( \square (S)(b_S) \square (b_S \supset \mathit{R} S b_S) \)

It is obvious how, given the reinterpretations just presented, Ginet's argument can be converted into an argument for incompatibilism about determinism
and responsibility; the argument has the same form as the argument for incompatibilism about determinism and freedom. I include the argument in an appendix.

III. THE COMPATIBILIST’S RESPONSE

Suppose that in possible world $W$ $S$ does not $V$ at $t_1$ ($t_1 > t_0$); in $W$ determinism is true and $\neg p(t_0)$ is also true. Imagine that there is a possible world $W'$ (suitably related to $W$) in which $S$'s at $t_1$ and in which $p(t_0)$ obtains. That is, we suppose that $V$-ing is something the compatibilist intuitively wants to say that $S$ can do at $t_1$. Perhaps $S$'s $V$-ing at $t_1$ would be $S$'s raising his arm at $t_1$ ($S$ is a normal agent in normal circumstances). We also imagine here that the causal laws which obtain in $W$ are such that $S$'s $V$'s at $t_1$ only if $p$ obtains at $t_0$.

Given (for the moment) that $S$ cannot determine the laws of nature, is it true that $S$ has the power at $t_1$ in $W$ to $V$ at $t_1$? That is, though $\neg p(t_0)$ obtains in $W$, is $\Diamond_s p(t_0)$ true in $W$? It follows from determinism that if $\Diamond_{t_1} (S$'s at $t_1)$ is true in $W$, then it is in $S$'s power at $t_1$ in $W$ to make it the case that $p(t_0)$.

But what is it to make it the case that a certain proposition is true? Ginet's notion of "making it the case that a proposition obtains" is ambiguous in a critical way; 'disambiguation' will be illuminating.

Suppose $p(t_0)$ is the proposition that event $e$ occurs at $t_0$. There are now two interpretations of "$S$ has it in his power at $t_1$ to make it the case that $p(t_0)$":

(i) $S$ has it in his power at $t_1$ to cause $e$'s occurrence at $t_0$, or
(ii) $S$ has it in his power at $t_1$ to perform some act $e^*$ such that if $e^*$ were to occur, then $e$ would have occurred at $t_0$.

Let $\Diamond_{t_1} p(t_0)$ symbolize (i), which is Ginet's intended interpretation, and let $\Diamond_{t_1} p(t_0)$ symbolize (ii).

Now the compatibilist will readily admit that $\Diamond_s p(t_0)$ is false in $W$, since one cannot 'determine' or causally affect the past; that is, the compatibilist can unite with the incompatibilist in rejecting this sort of backwards causation. But the compatibilist will insist that $\Diamond_{t_1} p(t_0)$ does not follow from the truth of $\Diamond_{t_1} (S$'s at $t_1)$ and determinism in $W$ (together with the unalterability of the laws of nature). All which follows is: $\Diamond_{t_1} p(t_0)$.9
But this can be true and unproblematic, according to the compatibilist. If $S$’s $V$-ing is in his power at $t_1$, then $S$ can so act at $t_1$ (perform $V$) that the past would have been different from what it was; but this involves an innocuous, non-causal counterfactual.

Thus, the compatibilist will accept the truth of

(a) $\square^s_{t_1} \neg p(t_0)$

and

(b) $\square^s_{t_1} [\neg p(t_0) \supset \neg (S \ V’s \ at \ t_1)]$

but deny that (a) and (b) entail

(c) $\square^s_{t_1} (S \ V’s \ at \ t_1)$.

The compatibilist must reject ($\beta$), but ($\beta$) can be replaced by a principle which the compatibilist believes underlies the plausibility of Ginet’s ($\beta$). Let ‘$\square^* p$’ abbreviate ‘$p \& \neg \square^* \neg p$’. The compatibilist will replace ($\beta$) with:

($\beta^*$)

\[
(t) \quad [\square^* (p \supset q) \& \square^* p \supset \square^* q] \supset [\square^* (p \supset q) \& \square^* p \supset \square^* q]
\]

The compatibilist argues that ($\beta$) robs it plausibility from ($\beta^*$); instances of ($\beta$) appear valid only because they fail to be distinguished from the corresponding instances of ($\beta^*$). Ginet’s own example can be understood as an instance of ($\beta^*$). If it’s not in my power this morning so to act that it wouldn’t rain this afternoon, and it’s not in my power so to act that if it did rain this morning, then it wouldn’t ruin the fresh paint on the house, then given that it does rain this morning, it follows that it is not in my power this morning so to act that the fresh paint on the house wouldn’t be ruined.

The compatibilist says that in attempting to justify ($\beta$), the incompatibilist is implicitly relying on ($\beta^*$). Since whenever $\square^* p$ is true, $\square^* p$ will also be true, Ginet is relying on cases in which $\square^* p$ is true in virtue of $\square^* p$. When

(d) $\square^* p$

and

(e) $\square^* (p \supset q)$

are both true, it follows that

(f) $\square^* q$
is true (by (β*)). Hence when (d) and (e) are true, the following will all be true:

\[(d') \quad \Box p\]
\[(e') \quad \Box (p \supset q),\]

and

\[(f') \quad \Box q.\]

But this will not be because (d') and (e') entail (f'). Further, sometimes '\[\Box p\]' is true while '\[\Box^* p\]' is false. In these cases, it might be that (d') and (e') are true but (f') is false.

Of course, in order to establish that (β) is unacceptable, the compatibilist must produce an example in which '\[\Box p\]' and '\[\Box (p \supset q)\]' are both true and '\[\Box q\]' is false. Consider a case where determinism is true and an ordinary agent A fails to scratch his back at time t; suppose also that he would have scratched his back at t if he had chosen to do so, and nothing prevented his so choosing. Imagine that some condition C obtained at \(t^+\) (prior to t) which causally necessitated the person's failure to scratch his back. In a case like this, the compatibilist might agree with the incompatibilist that

\[\Box_t A \cdot C(t^-)\]

is true, since no agent can determine the past, and that

\[\Box_t A \cdot [C(t^-) \supset (A \text{ fails to scratch his back})]\]

is true, since no agent can determine the laws of nature. But the compatibilist will nonetheless insist that

\[\Box_t A \cdot (A \text{ fails to scratch his back})\]

is false, since it is within A's power at t to scratch his back.

While the compatibilist claims that this very ordinary kind of case shows (β) to be invalid, the case is essentially controversial. This is what makes (β) such a focal point in the free-will debate; any counter-example to it appears to be essentially controversial (or perhaps question-begging). The compatibilist will need to rest his case on independent grounds, and then show how it is a consequence of such a theory that in the case discussed above, A is free at t to scratch his back. It would then follow that (β) is invalid. But this depends on the force of the compatibilist's independent reasons.

It might be argued that just as the incompatibilist trades on the ambiguity
of \((\delta)\) and \((\beta)\), he also trades on the ambiguity of \((\gamma)\). Consider the two propositions presented by van Inwagen:\(^{10}\)

\[(A) \quad \text{Nothing ever travels faster than light.} \]
\[(B) \quad \text{Jones, a physicist, can construct a particle accelerator that would cause protons to travel at twice the speed of light.} \]

van Inwagen claims that \((B)\)'s truth entails that \((A)\) does not express a law of nature. But the compatibilist can accept this claim without accepting \((\gamma)\), suitably interpreted. The compatibilist can agree that Jones doesn't have it in his power to cause a violation of a natural law; \((B)\) expresses the claim that Jones can cause protons to travel faster than light. In other words, \((B)\) says that Jones can build a machine at a particular time which will cause protons to travel faster than light (at some later time) and the protons' traveling faster than light will be a result of Jones' activity.

While this sort of conclusion may be unpalatable, it is not entailed by compatibilism. The compatibilist (who denies \((\gamma)\)) says that one can sometimes so act that a law which does hold wouldn't hold, but the violation of the law needn't be a result of one's activity. Indeed, it might be the case that one can so act that there would have been a violation of law immediately prior to one's activity.\(^{11}\)

Suppose now that \(p\) is the proposition that law \(l\) does not obtain. There are two interpretations of "\(S\) has it in his power at \(t_1\) to make it the case that \(p\)":

\[(i) \quad S\text{ has it in his power at } t_1 \text{ to cause a violation of } l, \text{ or} \]
\[(ii) \quad S\text{ has it in his power at } t_1 \text{ so to act that } l \text{ wouldn't obtain.} \]

While van Inwagen may be correct to say that \((B)\)'s truth requires that \((A)\) not express a natural law, this does not support \((\gamma)\) on interpretation (ii); it only supports \((\gamma)\) on interpretation (i). (van Inwagen himself, in contrast to Ginet, believes \((\gamma)\) should not use interpretation (i).) Rejection of \((\gamma)\) gives rise to what might be dubbed (borrowing David Lewis' phrase) "local-miracle compatibilism".

We can now produce a more general compatibilist response to Ginet's argument. The response is as follows: either (i) or (ii) is the correct kind of interpretation of Ginet's \(\Box_t S_{f_1} p\). Suppose first that (i), the causal interpretation, is correct. Then \((\gamma)\) and \((\delta)\) will be true, but not \((\beta)\). Hence, I (5) doesn't follow from I(3) and I(4); and I(8) doesn't follow I(6) and I(7). Suppose now that (ii),
the non-causal interpretation, is correct; it follows that \((\beta)\) interpreted as \((\beta^*)\) will be true, but either \((\gamma)\) or \((\delta)\) will be false. Hence, either I(4) doesn’t follow from I(1), or I(7) doesn’t follow from I(1). Some compatibilists say that one can sometimes so act that the past would have been different, while others say that one can sometimes so act that the laws would be different.12

IV. EXTENSION OF THE RESPONSE

It is useful to note that the compatibilist response sketched above can be transformed into a response to the incompatibilist about responsibility and determinism. Consider ‘\(\mathbf{R} \ S p\)’. There are two interpretations of ‘\(\mathbf{R} \ S p\)’, based on two interpretations of “\(S\) is not responsible for its being the case that \(p\)”. On the causal interpretation, \((\gamma')\) and \((\delta')\) will be accepted by the compatibilist, but not \((\beta')\). And on the non-causal interpretation, \((\beta')\) will be accepted, but either \((\gamma')\) or \((\delta')\) will be rejected.

Suppose \(p\) is the proposition that event \(f\) occurred at \(t_0\), some time before \(S\)’s birth. Now there are two interpretations of ‘\(\mathbf{R} \ S p\)’:

(i) \(f\) occurred at \(t_0\) and \(S\) is not responsible for causing the occurrence of \(f\) at \(t_0\),

or

(ii) \(f\) occurred at \(t_0\) and \(S\) is not responsible for any action (omission) which is such that were he not to perform it (it not to occur), then \(f\) would not have occurred at \(t_0\).

Consider interpretation (ii), symbolized by ‘\(\mathbf{R} \ S p\)’. Suppose \(S\) clenches his fist at \(t_1\). Imagine again that determinism is true and that \(p\) obtained, and that some suitable relation holds between \(S\)’s clenching his fist and \(p\) obtaining — \(p\) only if \(S\) clenches his fist at \(t_1\). It is open to the compatibilist to argue as follows. \(S\) is responsible for clenching his fist at \(t_1\) (since he is a normal agent in normal circumstances and certainly could have done otherwise at \(t_1\)); hence, \(S\) is responsible for some action which is such that were he not to perform it, event \(f\) wouldn’t have occurred at \(t_0\). Thus the compatibilist might argue that \(\mathbf{R} \ S p\) is false in this ordinary case and so \((\delta')\) is false (on interpretation (ii)). Of course, it is also open to the compatibilist to deny \((\gamma')\) on similar grounds. The general strategy of response is parallel to the strategy of response to the incompatibilist about determinism and freedom.
By adopting interpretation (i) instead of (ii) of *P*\[S\] \[t, p\], Ginet's argument for incompatibilism is vulnerable in a way in which van Inwagen's similar argument is not. While the compatibilist claims that on (ii), (β) [(β*)] is true while either (γ) or (δ) is false, the incompatibilist claims that (γ) and (δ) are both analytically true (on (ii)). I believe that one's assessment of the standard argument for incompatibilism depends on one's view about the incompatibilist's claim.

I want to discuss one final point about the incompatibilist's argument. It is important to note that the incompatibilist is not committed to the incoherence of backwards causation. Suppose we agree with Dummett that backwards causation is conceptually coherent. Suppose, that is, that in some possible worlds there exist chains of causation extending backward in time (as well as forward). We need to alter (D) to

\[(D^*)\] Determinism is the thesis that, for any given time, a complete description of the state of the world at that time, together with a complete statement of the laws of nature, entails every truth as to what happens at every time.

Now the incompatibilist will argue that with respect to forward-flowing causal chains, (δ) is true, but with respect to backward-flowing causal chains, (δ*) is true:

\[(δ^*)\qquad (t)(a_t) \Box (a_t \supset P_t a_t).\]

That is, when one's action is part of a backward-flowing causal chain, the incompatibilist will claim that one cannot so act that the future would be different from what it actually is (will be). The same intuitions which ground (δ) ground (δ*). Hence the incompatibilist result is achieved (if the argument is sound), even in a world with backwards causation. (δ), suitably interpreted, is supposed to be a necessary truth, even if the fact that there is no backwards causation is only a contingent truth.

Yale University
APPENDIX

The Incompatibility of Determinism and Moral Responsibility

I. \( \Box (D \supset (S)(a_S) \[a_S \supset (\exists b_S)(\exists g) \] s.t.

\begin{align*}
1. & b_S & \text{and the laws of nature entail that } g & \& \Box (g & \& b_S \supset a_S) & \text{df. } (D) \\
2. & \therefore \Box (g \supset b_S \supset a_S) & & & & 1 \\
3. & \therefore R (g \supset b_S \supset a_S) & & & & 2, \alpha' \\
4. & \therefore R g & & & & 1, \gamma' \\
5. & \therefore R (b_S \supset a_S) & & & & 3, 4, \beta' \\
6. & \therefore R S (b_S \supset a_S) & & & & 5, \text{dfs.} \\
7. & \therefore R S b_S & & & & 1, \delta' \\
8. & \therefore R S a_S \} & & & & 6, 7, \beta' \\
\end{align*}

II. \( \Box [D \supset (S)(a_S)(a_S \supset [R S a_S])] \)

III. \( \Box [D \supset (S)(p)(p \supset [R S p])] \)

IV. \( \Box [D \supset (p)(p \supset [R p])] \)

NOTES

1 Ginet, Carl: ‘Might we have no choice?’, in K. Lehrer (ed.): Freedom and Determinism (Random House, New York), pp. 87–104.


3 van Inwagen, Peter: 1975, ‘The incompatibility of free will and determinism’, Philosophical Studies 27, pp. 185–199.


7 Philosophers have disagreed about exactly what the relation must be between \( W \) and \( W' \) which suffices to show that \( S \) can \( V \) at \( t_1 \) in \( W \) (by virtue of \( S \)'s \( V \)-ing at \( t_1 \) in \( W' \)): see Lehrer, Keith: 1976, ‘“Can” in theory and practice: a possible Worlds Analysis’, in M. Brand and D. Walton (eds.): Action Theory (Reidel, Dordrecht), pp. 241–270; Horgan, Terence: 1979, ‘“Could”, possible worlds, and moral responsibility’, Southern Journal of Philosophy 17, pp. 345–358.

8 In fairness to Ginet, neither (i) nor (ii) captures exactly Ginet’s explication of “making it the case that a proposition obtains” presented in Ginet, 1980, p. 173. For the purposes of this discussion, however, (i) and (ii) will be adequate. Also, we can assume that “x
caused some part of $e$ to occur'' entails that $x$ caused $e$ to occur; this would bring (i) and (ii) into stricter conformity with Ginet’s explication.

This follows given our assumption about the relation between ‘$p(t_0)$’ and ‘$S V$’s at $t_1$’.
The general proposition lying behind this inference is: from ‘~(S $V$’s at $t_1$)’ and $\diamondsuit_{t_1} S (S V$’s at $t_1)$’ and ‘D’ there follows ‘$(\exists p) [~p(t_0) \& \Box_{t_1} p(t_0)]$’.

david lewis develops an account of counterfactuals which supports this position in Lewis, David: 1979, ‘Counterfactual dependence and time’s arrow’, Nous 13, pp. 455–476. I have been informed that Lewis independently develops a critique of van Inwagen’s example which is similar to mine in Lewis, David: forthcoming, ‘Are we free to break the laws?’, Theoria.


Lehrer’s criticism of van Inwagen in Lehrer, Keith: 1980, is unfair; it should rather be directed at Ginet.


I am deeply indebted throughout this paper to Carl Ginet, from whom I have borrowed liberally.