The irrelevance of the Consequence Argument

Lynne Rudder Baker

Peter van Inwagen has offered two versions of an influential argument that has come to be called 'the Consequence Argument'. The Consequence Argument purports to demonstrate that determinism is incompatible with free will.[1]   It aims to show that, if we assume determinism, we are committed to the claim that, for all propositions p, no one has or ever had any choice about p. Unfortunately, the original Consequence Argument employed an inference rule (the β-rule) that was shown to be invalid. (McKay and Johnson 1996)   In response, van Inwagen revised his argument. I shall argue that the conclusion of the revised Consequence Argument is wholly independent of the premiss of determinism, and hence that the revised Consequence Argument is useless in showing that determinism is incompatible with free will.

1. From the Original to the Revised Consequence Argument

The first version of the Consequence Argument relies on an operator 'N', defined as follows:

$$\text{'N}p\text{'} =_{df} \text{'}p \text{ and no one has, or ever had any choice about } p\text{'}$$

Having a choice about p, if p is a true proposition, implies 'being able to ensure that p is false.' (Van Inwagen 2000: 8) The first version requires two rules of inference, the α -rule and the β-rule:

$$\alpha \quad \Box p \vdash Np$$

$$\beta \quad Np, N(p \supset q) \vdash Nq$$

Let 'P_0' represent the proposition describing the complete state of the world at some time in the distant past, t_0, and let 'L' represent the conjunction of the laws of nature. Let 'P' represent any true contingent proposition. Then:

---

[1] Determinism is the thesis that "there is at any instant exactly only physically possible future." (Van Inwagen 1983: 3)

(a) $\Box((P_0 \,\&\, L) \supset P)$ (follows from thesis of determinism)

(b) $\Box(P_0 \supset (L \supset P))$ (a), standard logic

(c) $N(P_0 \supset (L \supset P))$ (b), α -rule

(d) NP_0 premiss

(e) $N(L \supset P)$ (c),(d), β-rule

(f) NL premiss

$\therefore$ (g) NP (e),(f), β-rule

Since no one has had any choice about the distant past, and no one has had any choice about the laws of nature, then—assuming that the α - and β-rules are valid—if determinism is true, 'no one has any choice about anything.'[2]

Thomas McKay and David Johnson proposed a counterexample to the β-rule. (McKay and Johnson 1996) If 'Np, Nq |- N(p&q)' is invalid, so is the β-rule. McKay and Johnson's counterexample, roughly, is this: Suppose that there is a coin that was untossed yesterday although I had the power to toss it. If I had tossed it, I would have had no choice about whether it landed heads or tails. So, N(the coin did not land 'heads' yesterday) and N(the coin did not land 'tails' yesterday).[3] But the conjunctive proposition, N(the coin did not land 'heads' yesterday & the coin did not land 'tails' yesterday), is false: I had the power to have tossed the coin yesterday, and if I had tossed it, then it would have landed heads or it would have landed tails. So, the conjunctive proposition is false even though each conjunct is true. The counterexample shows that the original β-rule is invalid.

Van Inwagen accepted the counterexample,[4] and revised the β-rule by redefining 'N'. Begin by defining 'having access to a region in logical space,' where logical space is a space 'whose points are possible worlds.' (Van Inwagen 2000: 2) A region of logical space 'corresponds to a proposition, or to a set containing the proposition and all and only those propositions necessarily equivalent to it.' And '[t]o have access to a region of logical space is to

---

[2] Van Inwagen 2000: 2. However, for a challenge to 'NP$_0$' (line d), see Campbell 2007.

[3] This is so because 'the coin did not land "heads" yesterday' is true and 'the coin did not land "tails" yesterday' is true, and I was not able to ensure that either was false. "[I]f p is a true proposition, having a choice about the truth-value of p implies being able to ensure that p is false." (Van Inwagen 2000: 8)

[4] Van Inwagen 2000: 3-4, 8. Although van Inwagen has accepted McKay and Johnson's counterexample, not everyone has. E.g., see Crisp and Warfield 2000.

be able to ensure the truth of the proposition that corresponds to that region, or to be able to ensure that that region contains the actual world.' (Van Inwagen 2000: 4) So:

> (A) Someone x has *access* to a region of logical space p iff x can ensure that p is true (i.e., that p contains the actual world). (Van Inwagen 2000: 4)

The original definition of 'Np' is equivalent to 'p and every region to which anyone has, or ever had, access overlaps p,' where one region overlaps another iff they have a part in common.   To have access to p is to have access to the 'superregions' of p—those regions of which p is a subset. To have access to a region that does not overlap p is to be able to ensure that p is false. (Van Inwagen 2000: 5)

Instead of defining 'N' in terms of access, van Inwagen's revision defines 'N' in terms of exact access:

> (EA) Someone x has *exact access* to a region p iff x has access to p and x does not have access to any proper subregions of p.

If one has exact access to a region, then one has exact access to none of its proper superregions, by definition. Intuitively, to have exact access to p is to be able to ensure the truth of p, but nothing 'more definite.' (Van Inwagen 2000: 8)   If r is a region to which I have access, then I have access to the superregions of r; but if r is a region to which I have exact access, then (by definition) I have no exact access to the (proper) superregions of r—though I do have access to them.

The revised definition of 'Np' is 'p and every region to which anyone has, or ever had, exact access is a subregion of p.' The revision of 'Np' avoids the counterexample, and presents no challenge to the validity of the β-rule—Np, N(p ⊃ q) |- Nq. This is so 'for the simple reason that every set that is a subset of both p and p ⊃q (that is, of p & q) is a subset of q.' (Van Inwagen 2000: 9) Hence, if every region to which anyone has ever had exact access is within both p and p ⊃ q, then every region to which anyone has ever had exact access is within q. In that case, the β-rule is valid.

2. What Does the Conclusion of the Revised Consequence Argument Show?

The revised Consequence Argument replaces the conclusion of the original argument with the following:

> p and every region of logical space to which anyone has, or ever had, exact access, is a subregion of p.

Using only van Inwagen's definitions and the assumption that (EA) is satisfied, I shall argue that every region of logical space to which anyone has, or ever had, exact access is the region containing only the actual world. If that is correct, then—since any true proposition contains the actual world as a subregion—the conclusion of the revised Consequence Argument immediately follows, whether determinism is true or not.

There are two parts of the argument for the conclusion that every region of logical space to which anyone has or ever had exact access is the actual world: Part I shows that everyone has exact access to the actual world, A, and Part II shows that no one has exact access to any non-actual region, p.

> Let x be an existent person.
> Let R be the region of logical space containing only the actual world.
>
> Part I: Show that x has exact access to R (the region containing only the actual world).
>
> (1) x has exact access to R iff x has access to R and x does not have access to any proper subregions of R.   [(EA)]

Since x is an existent person, x is inside R, and '[i]f one is inside a region one *ipso facto* has access to that region.' (Van Inwagen 2000: 4-5) So,

> (2) x has access to R.

R was stipulated to be the region of logical space containing only the actual world. By 'logical space,' van Inwagen means 'a space whose points are possible worlds.' (Van Inwagen 2000: 2) A region (like R) containing only one world contains only one point. Hence, it has no proper

subregions.[5] Since R has no proper subregions, x does not have access to any proper subregions of R. So,

(3) x does not have access to any proper subregions of R.

∴ (4) x has exact access to R.   [1-3, conjunction; modus ponens]

So, everyone has exact access to the actual world.[6]

Part II: Show that nobody has or ever had exact access to a nonactual region p.

What remains to be shown is that -- *pace* van Inwagen -- x has no exact access to any nonactual region of logical space. Van Inwagen notes that it is 'impossible to give a plausible example of a nonactual region to which I have exact access.' (Van Inwagen 2000: 9)   For example, suppose that I join people playing darts, but, although I can hit the board, I do not throw a dart; it is within my power to hit the board, but it is not within my power to hit the bull's eye, or any particular place on the board. Yet, I do not have exact access to the nonactual region in which I hit the board. Access, yes; exact access, no. (By (A), I have access to the nonactual region p, corresponding to the proposition that I hit the board, because I can ensure the actuality of p.) But as van Inwagen points out, I do not have exact access to hitting the board because, presumably, I could have hit the board and said 'Ah'. Since I have access to a proper subregion of hitting the board, my access to hitting the board is not exact.[7]

Although he cannot give an example of anyone's having exact access to a nonactual region, van Inwagen is confident that one does have exact access to nonactual regions. He is confident 'simply because a human being's ability to ensure the truth of things, to "fine tune" his actions and their consequences, must come to an end somewhere.' (Van Inwagen, 2000: 9)   Van Inwagen gives necessary conditions for one to have exact access to a nonactual region: 'For one to have exact access to a nonactual region p, it must be the case that one can ensure the actuality

---

[5] An important difference between a set of worlds and a region of worlds is that there is no such thing as a null region. A region of logical space with no worlds would be like a line with no points.

[6] More precisely, every existing person has exact access to the region containing only the actual world. Van Inwagen on occasion speaks of the actual world as itself a region. E.g., "the only region I am inside and have exact access to is the actual world." (Van Inwagen 2000: 8)

[7] Note that the problem for exact access here is not that all that is within my power is *at least* to hit the board, without my being able to hit any particular place on it. Specifying a region using 'at least' is not what precludes exact access to it.

of p, but not the actuality of p and any logically independent region.' (Van Inwagen 2000: 9) So,

> (EAN) For any nonactual region p, x has exact access to p only if: (i) x can ensure that p is true, and (ii) $\sim\exists s$(s is a region logically independent of p & x can ensure that (p & s) is true.

Is (EAN) ever satisfied? Here's an example that suggests an in-principle strategy for answering in the negative: Let p be the nonactual region of logical space (i.e., the false proposition) of my throwing the dart and hitting the board. Let l_1 be the halfway point on the shortest path from the starting location l of my left foot to the line for throwing darts. Let s_1 be the region of logical space consisting of worlds in which my left foot traverses at least the path from l to l_1. I do not have exact access to p because there is a logically independent region, s_1, such that I can ensure that (p & s_1) is true—by moving my foot from l to l_1 while throwing the dart. It follows from (EAN) that I do not have exact access to the nonactual region, p, of my throwing the dart. Do I have exact access to the smaller region of the intersection of p and s_1? Well, no. Let l_2 the halfway point on the shortest path between l_1 and the line for throwing darts. Let s_2 be the region consisting of worlds in which my left foot traverses at least the path from l to l_2. Now I do not have exact access to (p & s_1) because s_2 is a region logically independent of (p & s_1), and I can ensure that the conjunction (p & s_1 & s_2) is true—by moving my foot from l to l_2 while throwing the dart. The region, s_2, in which I move my foot from l to l_2 is logically independent of the intersection of p & s_1: There are worlds in which I move my foot from l to l_2 and I do not throw a dart; and there are worlds in which I throw a dart and move my foot from l to l_1, but not to l_2. Repetition of the argument shows that I do not have exact access to (p & s_1 & s_2) either. And so on for s_3, s_4,....s_n..... to infinity.

Someone may object: "Zeno's paradoxes notwithstanding, we know that while throwing the dart, your left foot's traversing infinite segments of physical space can be completed in a short finite period of time. Some of the paths entailed by some of the 's_i's will be too short for you to move your left foot exactly the required distance. Hence, 'fine-tuning' your action of moving your foot does come to an end somewhere.'

No. What matters is only this: Whether the additional regions of possible worlds, s_i, are logically independent of the intersection of p & s_1 &...& s_{i-1} and whether you can make the conjunction (p & s_1 &...& s_{i-1} & s_i) true. Given the way that the 's_i's are constructed, any s_i is logically independent of the intersection of p & s_1 &...& s_{i-1}. Consider a region – call it 's_{67}' – of worlds in which your left foot moves at least from l to l_{67}, where l_{67} is one billionth of an inch past l_{66} -- the end point of your foot's movement in s_{66}. First, s_{67} is logically independent of the intersection of p & s_1 &...& s_{66}. This is so, because there are worlds in which the conjunction (p & s_1 &...& s_{66}) is true, but s_{67} is false (your foot happened to stop at l_{66}); and there are worlds in which s_{67} is true, but p -- and hence the conjunction (p & s_1 &...& s_{66}) -- is false. Moreover, you have access to the intersection of (p & s_1 &...& s_{66} & s_{67}): You can make the conjunction (p & s_1 &...& s_{66} & s_{67}) true by throwing the dart and moving your foot from l all the way up to the line for throwing darts. You have access, but not exact access, to the conjunction.[8]

Although I have only given one example in which (EAN) is, in principle, not satisfied, the strategy used here can be used, I believe, for any other case in which (EAN) is putatively satisfied. Any kind of event that requires a change of position, color, tone or anything else whose instances are dense will generate an infinite series.[9]   Here is the general strategy against satisfiability of (EAN): Let p be a nonactual region to which one putatively has exact access. Consider the huge class of events in which there is a change in spatial location, weight, color, musical tone – any kind of event that has a dense set of instances. For any event E in that huge class, if one can ensure the truth of the proposition that E obtains, then one has access to an infinite number of regions of logical space.   (If you can sing the continuous interval from tone 1 to tone 2, then you have access to worlds in which you sing tone 1, and to worlds in which you sing the interval from tone 1 to tone 1.5, and to worlds in which you sing the interval from tone 1 to tone 1.75, and so on forever.) If e_1 (the proposition that E obtains) is logically independent of p, and one can ensure that the conjunction (p & e_1) is true, then the region e_1 is a counterexample to one's having exact access to p. Now consider the intersection of p and e_1, there will be an e_2 such that e_2 logically independent of the intersection of p and e_1, and one can ensure that the conjunction of (p & e_1 & e_2) is true. And so on.

_____

[8] Anyone who is able to move something at least a foot is ipso facto able to move it at least an inch...or a billionth of an inch
[9] Spatial locations, e.g., are dense because between any two, there is another.

Although this is not a proof that no one has or ever has had exact access to a nonactual region of logical space, it does make credible the thesis that there are in-principle (not just a practical) blocks to satisfaction of (EAN).[10] Call the thesis that every claim of exact access to a nonactual region is susceptible to a counterexample by infinite series, 'the Infinite-Series claim.' The point of the Infinite-Series claim is to show that as one "squeezes" p down to smaller and smaller regions of logical space, one never gets to a region to which she has exact access. If, as it seems, the Infinite-Series claim is true, we can summarize Part II as follows:

(5) x has exact access to a nonactual region p → x can ensure that p is true & ~∃s(s is a
        region logically independent of p & x can ensure that (p & s) is true). [(EAN)]

(6)   x can ensure that a nonactual region p is true →∃s(s is a region logically
        independent of p & x can ensure that (p & s) is true).   [The Infinite-Series claim]

∴ (7) ~(x has exact access to a nonactual region p) [6, re-stated; 5,6, modus tollens]

If this is right, then a necessary condition for having exact access to a nonactual region r—that there be *no* r' that is logically independent of r and such that x can ensure that (r & r') is true—is never met. So, no one has or ever had exact access to a nonactual region—the conclusion of Part II.

Putting Part I and Part II together, we have

(8) Everyone has exact access to the actual world. [Part I]

(9) No one has exact access to any nonactual region of logical space.   [Part II]

(10) If 8 & 9, then every region of logical space to which any has or ever had access is
        the actual world. [logical truth]

∴ (11) Every region of logical space to which anyone has or ever had exact access is the
        actual world.[11] [8, 9, 10 conjunction, modus ponens]

---

[10] Moreover, if there were a nonactual region to which one had exact access, it would be such an unfathonable region that it would be of dubious relevance to issues of free will.

[11]This conclusion follows from line 9 alone. Although the conclusion (line 11) would still be true even if no one ever had exact access to any region of logical space, the claim that everyone has exact access to the actual world guarantees that line 11 is not true vacuously.

Now we have a simple argument that is independent of determinism for the conclusion of the revised Consequence Argument. Let p be a true proposition corresponding to an extended region of logical space:

    (12) p   [premiss]

    (13) p →the actual world is a proper subregion of p. [defn 'p is true']

    (14) Every region of logical space to which anyone has or ever had exact access is the
          actual world. (line 11 above)

    (15) p & every region of logical space to which anyone has or ever had exact access is
          the actual world. [12, 14, conjunction]

∴ (16) p & every region to which anyone has or ever had exact access is a proper
          subregion of p. [15, 13, substitution]

Line 16 is the conclusion of the revised Consequence Argument.

So, the conclusion of the revised Consequence Argument follows solely from any true proposition together with the assumption that the definition of 'exact access' is satisfied—without any recourse to determinism or to indeterminism, without use of the β-rule, and without any compatibilist construal of 'ensure'. What I have shown is that, on van Inwagen's definitions, for any p that turned out to be true, no one had been able to ensure that p turned out to be false—whether determinism is true or false. Hence, the revised Consequence Argument is irrelevant to any challenge to free will raised by determinism.

3. The Upshot

Although I am a compatibilist, my aim here is not to defend compatibilism. It is the more modest one of arguing that the revised Consequence Argument may be put aside in the debates about determinism and free will. I am not claiming that there is any fallacy in the Consequence

Argument, nor am I contesting the soundness of the revised Consequence Argument.[12]   If the revised Consequence Argument is sound, then, of course, its conclusion is true. What I have shown is that if its conclusion is true, then it remains true whether determinism is true or false.

Some philosophers may want to argue that the conclusion of the Consequence Argument by itself is incompatible with free will. If such philosophers believe in free will, they should give up either belief in free will or the conclusion of the Consequence Argument; determinism would not come into the picture at all.[13] Other philosophers may want to argue that the conjunction of determinism and the conclusion of the revised Consequence Argument is incompatible with free will.[14] Such a move would merely throw us back into the original debates about determinism and free will: The conclusion of the revised Consequence Argument would not add anything to the debate.

Since the original Consequence Argument turned out to depend on an invalid inference rule, and the revised Consequence Argument turned out to be irrelevant to the issue of the compatibility of determinism and free will, we may move on safely beyond the Consequence Argument in our debates about determinism and free will.[15]

University of Massachusetts Amherst
Amherst, MA 01003-9269 USA
lrbaker@philos.umass.edu

References

Campbell, J. K. 2007. Free will and the necessity of the past. *Analysis* 67: 105-111.

Crisp, T. M. and T. A. Warfield. 2000. The irrelevance of indeterministic counterexamples to Principle Beta. *Philosophy and Phenomenological Research* 61: 173-184.

---

[12] Ted Warfield has argued that the Consequence Argument commits a modal fallacy. See also Warfield 2000: 171.

[13] Van Inwagen may interpret my result as implying that free will is incompatible with both determinism and indeterminism, and hence is incoherent. (See Van Inwagen 2000:11.) I would not draw such a conclusion. Rather, I would reconsider van Inwagen's construal of 'free will'. Perhaps the conclusion of the Consequence Argument has nothing to do with free will, properly conceived.   See my "The Moral Irrelevance of Determinism," in preparation.

[14] See Campbell 2007: 110. Campbell, a compatibilist, does not endorse this move; he merely mentions it as a possibility. See also Warfield 2000:170.

[15] I am grateful to Edmund L. Gettier, Phillip Bricker, Hilary Kornblith, Gareth B. Matthews, and an anonymous referee for helpful comments on this paper.

McKay, T. and D. Johnson 1996. A reconsideration of an argument against compatibilism. *Philosophical Topics* 24: 113-122.

Van Inwagen, P. 1983. *An Essay on Free Will* (Oxford: Clarendon Press).

Van Inwagen, P. 2000. Free will remains a mystery. In *Philosophical Perspectives 14: Action and Freedom, 2000*, ed. J. E. Tomberlin (Malden, MA: Blackwell): 1-19.   Reprinted in *The Oxford Handbook of Free Will*, ed. R. Kane (Oxford: Oxford University Press, 2002): 158-177.

Warfield, T. A. 2000. Causal determinism and human freedoms are incompatible: a new argument for incompatibilism." In *Philosophical Perspectives: Action and Freedom, 2000,* ed. J. E. Tomberlin (Malden, MA: Blackwell Publishers): 167-180.