Reversing the Consequence Argument[1]

Abstract: In this paper I present and evaluate van Inwagen's famous Consequence Argument, as presented in *An Essay on Free Will*. The grounds for the incompatibility of freewill and determinism, as argued by van Inwagen, is dependent on our actions being logical consequences of events outside of our control. Particularly, his arguments depend upon, in on guise or another, the transference of the modal property of not being possibly rendered false through the logical consequence relation, i.e. the β-principle. I argue that, due to the symmetric nature of determinism, van Inwagen is exposed to what I call "reversibility arguments" in the literature. Such arguments reverse the β-principle and start from our apparent control over our own actions to our control over the initial conditions. Since van Inwagen does not endorse a particular theory of laws or logical consequence, he is open to such counterarguments. The plausibility of such reversibility arguments depends on what would be called a Wittgensteinian conception of logical consequence. In the *Remarks on the Foundation of Mathematics,* one of Wittgenstein's main concerns is the normativity of logical inference i.e. proof. Such concerns with normativity and rule-following are generally a feature of his later philosophy. In the *Remarks* Wittgenstein resists a conception of logical deduction which places the source of normativity outside of human practice.

1. The Consequence Argument

In *An Essay on Freewill*, van Inwagen provides an influential argument for the incompatibility of determinism and freewill: the Consequence Argument. Here I quote his sketch of the argument (from which he develops several variations):

> "If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born, and neither is it up to us what the laws of nature are.

Therefore, the consequences of these things (including our present acts) are

not up to us." (1983: 16, 56)

What this argument depends on, I will argue, is that the notion of 'consequence' is such that it preserves the property denoted by 'is not up to us'.

Let me discuss the operative meaning of 'determinism'. Determinism, in bare form, is the thesis that given the laws of nature and given some sufficient amount of information about the universe, usually thought of as the complete set of facts given at a particular time, one could deduce the rest of the facts. Another way to put this is that a set of facts, within a time slice, or chunk of the world (depending on the form of the laws), and the natural laws uniquely identify a complete world history. Note that the notion of determinism that van Inwagen is concerned with is what may be called logical determinism as opposed to causal determinism or universalism (see Ibid. pp 1-5). Determinism is a logical relation between propositions, rather than a relation between objects or events in the world. Relatedly, van Inwagen aims to pitch his argument without any robust theory of laws of nature. He merely gives some desiderata for a theory of laws: (1) that 'is a law of nature' denotes a real predicate, (2) that laws of nature are truth-apt, and (3) that the status of lawhood is objective. However it is questionable that his focus on the logical determinism makes good on his claim to be neutral with respect to competing theories of lawhood. Beebee (2000, p 579) notes that arguments for incompatibilism often rely on an anti-Humean conception of laws, this question we will return to.

Essential to van Inwagen's argument is that the laws of nature and some set of facts, call them the boundary conditions, determine the rest of the facts. At least as currently understood, this is not supposed to bear on whether the laws of nature are somehow independent of the facts or not (i.e. Humeanism). What van Inwagen is committed to the laws of nature imposing limits upon our abilities and being such that no one can render them false (Ibid., 62-64). One further stipulation that van Inwagen makes about the laws: the consequences of any set of laws of nature are also laws of nature, let's call these *derivative laws*. From his remarks it is clear that the derivative laws follow by logical consequence from the primary laws and further that derivative laws can must be consequent from the sets of propositions which are exclusively laws. Due to the schematic nature of the laws, the truth of the laws of nature and the derivative laws do not determine

the rest of the facts, the boundary conditions are essential for the inference. Significantly, the consequent facts do not have law status or even derivative law status, themselves.

Let us return to consideration of the "not up to us" property through logical inference. What this assumption amounts to is a reduction of law status to truth in a set of possible worlds. This is distilled into van Inwagen's β-principle:

β: N(p → q), Np ⊢ Nq

Where N is the higher-order property of "not being up to us" in the sense that a law of nature is not up to us. The first inference rule for N, α, guarantees that the laws (and any other necessary truths) have this property, $\square$p entails Np. Some, like Spencer (2017) reject the α-principle and claim that we have the ability to perform impossibilities, but I will not question α here. More or less, the β-principle is defended by van Inwagen on grounds of intuition. I will let this direct defense alone for the moment, however note that van Inwagen has since admitted to the invalidity of β based on objections from McKay and others and modified N to N*, see Vihvelin (2017, section 5), the new argument structure is similar enough for me to continue with the original.

First, I want to question whether β is really the core principle at play here, consider B:

B: $\square$(p → q), $\square$p ⊢ $\square$q

This principle more straightforwardly seems to also fit in the argument for incompatiblism and indeed, N is supposed to be a modal operator, weaker than $\square$. This principle is also uncontroversially valid. A principle such as B shows up in van Inwagen's second, "simple" argument as a metaphysical assumption, which "asserts that no person has access to any world in which the laws of nature are different from what they are in the actual world" (Ibid., 92). The relevant accessibility relation is such that an agent has access to a world if and only if they could render some proposition in that world true (i.e. they have some unexercised ability which is exercised in that world) or they actually render some propositions true, which gives any agent access to the actual world (by definition). So why not work with B instead of β? One reason might be this, that it is possible "in the broadly logical sense" for the laws of nature to have been different (Ibid., 92). So then it may not be true that a law of nature, p, can be true while $\square$p is not, if $\square$ is interpreted as logical possibility. This just highlights that N really is just a reinterpretation of $\square$ so that it is weak

enough to apply to the physical laws but strong enough to represent the higher-order property of not being able to be rendered false by humans. However this will still not do, as noted before, the consequence argument does not result in all of the determined facts being laws of nature or necessary, only consequences of sets of proposition consisting of only laws of nature (and derivative laws) have law status. So β need not have the strength of B, what is transferred by such instances of logical consequence is not law status but the weaker property of being unable to be rendered false. So, since, van Inwagen is stuck with β rather than B, he cannot claim the validity of β as a mere theorem of logic as he would be able to with B.

The core of van Inwagen's Consequence Argument is that the higher-order property or modal operator explains how it is that laws of nature and the initial conditions of the world are outside of our control can be transferred to the rest of the facts, constraining all of our actions. The nature of this transference is put in terms of an inference rule, which leaves the van Inwagen merely gesturing at the "logical" nature of the relation between the two sets of propositions, the premises and the conclusion (as opposed to a causal relation). Focusing on the purely logical leads van Inwagen no recourse to make an in principle distinction between what should be taken as premise and what would be taken as conclusion.

2. Reversibility Arguments

In van Inwagen's discussion of β the boundary conditions taken to be the initial conditions of the universe, or at least some facts about the universe far enough in the past to be uncontroversially outside of our control. There is not much reason given by van Inwagen for believing that this temporal ordering (from past facts to future fact) is essential to either the determining or logical consequence relation. As noted by Hoefer (2002) and Ismael (2016), if a one-to-one determinism is true then this inference relation is perfectly symmetric. $N(p \rightarrow q), Np \vdash Nq$ is just as valid as $N(q \rightarrow p), Nq \vdash Np$, however this inference would appear to beg the question against the compatibilist, who would not accept their current actions as not up to them. This reminds one of the adage: "one philosopher's modus ponens is another's modus tollens," and in that spirit, I will call such arguments "reversibility arguments".

The compatibilist can present a reversed argument, showing that we have "freedom from the inside out" compatible with determinism. Consider RB:

RB: $\Box(q \rightarrow p), \Diamond q \vdash \Diamond p$

RB is valid in normal modal logics and a corresponding Rβ principle follows if we assume that $\sim N \sim p$ implies $\Diamond$, i.e. if ability implies possibility. RB states that given the laws are determinate from future to past and my current actions are up to me, then the initial conditions of the universe are also up to me. Van Inwagen has no recourse to the temporal ordering of the past facts and the current facts because the logical consequence relation is neither temporal nor causal, to do so would be to to renounce his focus on logical determinism as well as voiding his neutrality on the proper conception of laws. His argument depends on a notion of logical consequence which is not uncontroversial and would suggest an anti-Humean, governing conception of the laws, in which the laws make things happen. On the other hand, if what can be called a Wittgensteinian, constructivist account of rules of inference is correct, then the incompatibilist has no recourse to challenge this reversal, indeed then "we are perfectly justified in viewing our own actions *not* as determined by the past, *nor* as determined by the future, but rather as simply determined… *by ourselves, by our own wills*" (Hoefer 2002: 207).

3. Wittgensteinain Considerations Against the Governing Conception of Laws

One of the major threads in Wittgenstein's later work is a questioning of the grounds of logical inference. What he takes himself to be overturning is a picture of proof evident in the works of Russell and Frege, in which our proofs are considered tracings of something already there (RFM 1978: I, 21). From this sort of picture it is, perhaps, easy to see the normative force behind a *rule* of inference as grounded in the independent reality of this pattern. One might say that if the world is such that when there is two and two there is four, or that if it really is such that given the necessity of the laws and the unchangeability of the boundary conditions of the universe and determinism, that it follows that all of my actions are outside of my control. This is a picture in which the logical structure of the world is independent of our proofs, which holds great significance for the incompatibilist. Particularly, this independence must be what accounts for the primacy of particular boundary conditions which our intuitively outside of my control over those conditions which are intuitively within my control, e.g. the spacetime region directly following a

decision of mine which is complete enough for a determining of the complete past and future. If we take such "boundary" conditions as primary, then we can very well infer that the distant past is in some way in my control.

How do Wittgenstein's considerations of logical consequence make way for reversibility arguments? Wittgenstein's stance is distinctly constructivist, he claims that "mathematician[s] create essences" (1978: I, 32) and that a "proof serves as the picture of an experiment" (Ibid., 36). So then, in the first instance a proof is a pattern produced by a sequence of *actions* and constituted by a sequence of propositions (or sometimes images). These patterns become paradigms of varying procedures and in that way constitute rules. Two and two is four not because of some independent structure of reality (or at least not directly) but rather because it constitutes proper addition. If someone was to tell me that two and two is five, I would believe they do not mean what I mean by "two and two." Proofs deal with essences, which are conventional (Ibid., 74) and so the normative force of proofs is only as strong that of "other laws in human society" (Ibid., 116). So then, if van Inwagen aims to ground the normativity of the β-principle he must make reference either to a governing conception of laws, which would given the facts of the world an ordered structure, or he must make reference to the fact that we take the past to be either temporally or causally prior to our current actions (if temporal or causal structure is innate toe he laws, the first horn must be taken). In either case, he must take on some controversial metaphysical positions that he purports are independent of his argument.

<div align="center">Bibliography</div>

Beebee, H. (2000) "The Non-Governing Conception of Laws of Nature," in *Philosophy and Phenomenological Research,* in Vol. 61, No. 3, 571-594.

Hoefer, C. (2002) "Freedom From the Inside Out," in *Time, Reality, and Experience,* C. Callender ed., Cambridge: Cambridge University Press, 201-222.

Ismael, J. T. (2016) *How the Physics Makes Us Free,* Oxford: Oxford University Press.

Spencer, J. (2017) "Able to Do the Impossible," in *Mind,* Vol. 126, No. 502, 465 497.

Wittgenstein, L. (1978) *Remarks on the Foundations of Mathematics,* G. H. von Wright, R. Rhees, and G. E. M. Anscombe (eds.), G. E. M. Anscombe (tr.), Cambridge: MIT Press.

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

Vihvelin, K. (2017) "Arguments for Incompatibilism", *The Stanford Encyclopedia of Philosophy,* Edward N. Zalta (ed.).