The counterfactual structure of the consequence argument

Stefan Rummens (KU Leuven)

1. Introduction

Peter van Inwagen's (1975, 1983) famous consequence argument states that, in a deterministic world, the ability of an agent 'to do otherwise' presupposes an ability to either change the distant past or the laws of nature. Since we do not have these strange abilities, there are, thus, no alternative possibilities if determinism holds true.

This paper revisits a well-known rebuttal of van Inwagen's argument, which focuses on the counterfactual structure of alternative possibilities. According to this CS-rebuttal, as I shall call it, the consequence argument only succeeds in showing that the ability to do otherwise is such that if the agent had exercised it, the distant past and/or the laws of nature *would have been* different all along. On the counterfactual scenario, there is, therefore, no need for the agent to exercise an ability to *change* the past or the laws of nature.

The aim of the present argument is to provide and defend a version of the CS-rebuttal that is capable of countering the recent criticisms that have been raised against it. To that purpose, I first present van Inwagen's original version of the consequence argument (2). After exposing some difficulties with David Lewis' (1981) famous version of the CS-rebuttal (3), I proceed by explaining and defending an older and, in my view, superior version (4). I subsequently discuss a traditional incompatibilist rejoinder, which insists that the past and the laws of nature are fixed. Although this rejoinder delivers a valid argument against the existence of alternative possibilities, it relies on premises the compatibilist explicitly rejects (5). The outcome of the debate is therefore properly characterized as a genuine dialectical stalemate between

1

compatibilists and incompatibilists (6). In the final sections of the paper, I demonstrate that attempts by John Martin Fischer (1994) (7), by Wesley Holliday (2012) (8) and by John Martin Fischer and Garrett Pendergraft (2013) (9) to move beyond the stalemate in favor of the incompatibilist position all fail. I thereby show that the debate is marred by a misunderstanding of the semantics underlying the backtracking conditionals sometimes associated with the compatibilist position. In view of my arguments, the dialectical stalemate between compatibilists and incompatibilists regarding the counterfactual structure of the ability to do otherwise remains fully intact (10).

2. Van Inwagen's original argument

Peter van Inwagen (1975) presents his consequence argument on the basis of the example of a judge, *J*, who did not raise his hand at time *t*, whereas, if he had raised his hand, this would have prevented the execution of a criminal. The question before us is, therefore, whether or not *J* could have raised his hand at *t*. 'P_t' is used to denote the proposition expressing the state of the actual world, *w*, at time *t*, 'P_0' denotes the proposition expressing the state of the world, *w*, at a time in the distant past and '*L*' is used to denote the proposition expressing the laws of nature in *w*. Van Inwagen (1975: 191) spells out his argument in seven steps:

(1)   If determinism is true, then the conjunction of P_0 and *L* entails P_t.

(2)   If *J* had raised his hand at *t*, then P_t would be false.

(3)   If (2) is true, then if *J* could have raised his hand at *t*, *J* could have rendered P_t false.

(4)   If *J* could have rendered P_t false, and if the conjunction of P_0 and *L* entails P_t, then *J* could have rendered the conjunction of P_0 and *L* false.

(5)   If *J* could have rendered the conjunction of P_0 and *L* false, then *J* could have rendered *L* false.

(6)   *J* could not have rendered *L* false.

(7)   If determinism is true, *J* could not have raised his hand at *t*.

3. David Lewis' version of the CS-rebuttal

Although the debate on the consequence argument has been extensive, it seems fair to say that the most prominent attempt to refute it has focused on its underlying counterfactual structure.[1] Kadri Vihvelin's (2013: 155-166) recent formulation of this CS-rebuttal goes to the core of the problem. She explains that the consequence argument fails to show that an ability to do otherwise presupposes an ability to *change* the past or the laws of nature. Instead, the argument merely shows that either the past or the laws of nature would have *been* different on the counterfactual scenario in which the agent exercises his ability to do otherwise. This means that if determinism is true, the following counterfactual is also true (Vihvelin 2013: 164):

> (Diff P/L)   If an agent had done otherwise, the remote past and/or the laws of nature would have been different.

Although several authors (Horgan 1985, Vihvelin 2013) attribute the CS-rebuttal of the consequence argument mainly to David Lewis (1981), this attribution is somewhat problematic. The CS-rebuttal had, first of all, already been presented by several authors prior to Lewis' publication (Narveson 1977, Foley 1979, Lehrer 1980). Secondly, and more importantly, the argument presented by Lewis differs from these earlier versions in important respects.

All of the earlier versions directly target premise (4) of van Inwagen's argument. Jan Narveson, for instance, explains that

> "if determinism is true, then if *J* had raised his hand at *t*, then that would show that either P_0 or L is false; but it would not 'render' either of them false. It would show that one had been false all along: it would not 'make', 'cause', or, therefore, 'render' it false." (Narveson 1977: 85)

David Lewis' argument, in contrast, does not reject premise (4) (Lewis 1981: 296-297). He proceeds, instead, by making a distinction between a strong and a weak sense of rendering false a proposition.[2] In the strong sense, I render something false as a causal consequence of my action. So, for example, I render it false in the strong sense that the door is closed by opening it. In the weak sense, I render false a proposition by acting in a world in which that proposition is false. So, for example, all of my actions render it false in the weak sense that gravity is a repulsive rather than an attractive force or that the earth was destroyed by an asteroid one million years ago.

Lewis now assumes without further argument that the first four premises of van Inwagen's argument are true on both readings of 'rendering false' but then argues that premise (5) is false on the strong reading and premise (6) is false on the weak reading. Although Lewis does not extensively explain why he believes premise (5) to fail on the strong reading, Andre Gallois (1977: 100-101), Terence Horgan (1985: 341-342) and Danilo Šuster (2012: 82) all correctly point out that it is not generally the case that if an agent can render a conjunction false in the strong sense, this agent can then also render one of the conjuncts false in the strong sense. Although I can render it false that the coin does not fall heads and does not fall tails by tossing the coin, I cannot render it false that the coin does not fall heads nor that the coin does not fall tails because I have no control over the outcome of the toss. With regard to the falsity of premise (6) on the weak reading, in turn, Lewis argues that we 'are free to break the laws' (in the weak sense) because if we had acted otherwise, some law-breaking 'divergence miracle' would have occurred in the counterfactual world.

I find the version of the CS-rebuttal which rejects premise (4) superior to Lewis' version for several reasons. The primary reason is, of course, that I fail to see how premise (4) could be true on the strong reading. If, on the counterfactual scenario, the judge raises his hand, this action indeed renders P_t false in the strong sense. But for this to be possible, it is not necessary that he also renders the conjunction of P_0 and L false in the strong sense. The fact that it is merely necessary that this conjunction *is* false on the counterfactual scenario implies that it is merely necessary that the agent can render it false in the weak sense (Perry 2004: 249). Another reason for not siding with Lewis has to do with his rejection of premise (5) on the strong reading. Although it is true that an agent who can render a conjunction false in the strong sense cannot always render one of the conjuncts false in the strong sense, van Inwagen (1977a) himself has argued against Gallois (1977) that this conclusion does not hold for conjunctions in which one of the conjuncts only refers to states of affairs in the distant past. Although I do not wish to take sides in this debate, I believe that it nevertheless shows that, on Lewis' strong reading, the falsity of premise (5) is perhaps less secured than he supposes. A final reason for preferring the CS-rebuttal which rejects premise (4) over Lewis' account is that it more clearly shows that the problem with the consequence argument really lies with its underlying counterfactual structure and that this problem can easily be formulated without any reference at all to the distinction

between a weak and a strong sense of 'rendering false'. Narveson (1977: 85) had already recognized that premise (4), which he rejects, could be salvaged by stipulating that 'rendering false P' means 'doing something from which it may be inferred that P is false'. But he immediately added that such an approach would "merely illustrate the danger of stipulating peculiar meanings for fairly ordinary terms" (Narveson 1977: 85). Whereas Vihvelin (2013: 163) finds it mysterious that Lewis' argument has not met with wider acceptance, it seems plausible to me that his claim that we are 'free to break the laws' has not been very helpful in this regard. Although this claim is true on the 'peculiar meaning' of the term intended by Lewis, it is misleading. The fact that we are free to break the laws (or to change the past) simply refers to the fact that we could have acted in a counterfactual world in which these laws (or the past) would have *been* different all along, and in which we – in the 'fairly ordinary' usage of words – never would have had to *change* anything about them.

4. The counterfactual structure of the ability to do otherwise

In order to more fully explain the CS-rebuttal targeting premise (4), I propose, first, to explicate the counterfactual structure of the ability to do otherwise (AO) of an agent, S, as follows:

(AO)   *Ability to do Otherwise:* given P_t as a true proposition about the actual world, w, the ability to do otherwise (as the ability to 'render false P_t') is an ability such that all worlds, w', in which it is exercised, are worlds about which P_t is a false proposition.

I submit that (AO) is an analytic truth about human agency which holds in deterministic and indeterministic worlds alike. It simply expresses the fact that when we talk about an ability to do otherwise we are always comparing the actual course of events with an alternative course of events. Thereby, the counterfactual scenario necessarily plays out in a *different* world, w', since P_t cannot be true and false about the same world, w. In terms of van Inwagen's example, it is simply impossible for a judge to raise and not to raise his hand at the same time in the same world.

Although (AO) is straightforward, it is relevant to point out that it rules out a confused interpretation of what it means for an agent to 'render false P_t'. When a judge renders it false that his hand is not raised at time t by raising his hand at some time shortly

prior to t, the judge is *not* changing the truth value of P_t from true to false. On the counterfactual scenario, the agent is simply acting within a counterfactual world, w', about which P_t *has never been* a true proposition. Changing the truth value of time-indexed propositions such as P_t would require an agent to mysteriously perform some action such that, prior to the action, his actual world is w, whereas, after the action, his actual world is a different world w'. As both Perry (2004: 246) and Vihvelin (2008: 315-316) correctly point out, this sort of action is metaphysically impossible.

The core of the CS-rebuttal now consists in the rejection of premise (4) of van Inwagen's argument on the basis of (AO). Since (AO) implies that the ability to do otherwise is necessarily exercised in a different world, w', there are no grounds for assuming that the counterfactual agent has to *render false* the past or the laws of nature of his world. The only necessary condition that can be assumed is that the past and/or the laws *are* different on the counterfactual scenario that plays out in w'. This means that premise (4) can, at best, be replaced by the following premise

(4′)   If J could have rendered P_t false, and if the conjunction of P_0 and L entails P_t, then the conjunction of P_0 and L would have been false (in the counterfactual scenario).

Although this premise is true, it is too weak to rescue the consequence argument. The purpose of the original premise (4) is to provide a 'transfer principle' which connects our inability to change the past and the laws of nature with an inability to do otherwise in the present. Since premise (4′), in contrast, no longer refers to our inability to change the past or the laws of nature, it can no longer be combined with premises (5) and (6) (which express these inabilities) to yield the conclusion (7).

Notice that the CS-rebuttal of the soundness of the consequence argument only relies on (AO) for the rejection of premise (4) and therefore does *not* presuppose a commitment to (Diff P/L). What is true, however, is that (4′) makes clear that if the compatibilist wants to do more than reject the soundness of the consequence argument and, in addition, also reject the conclusion that alternative possibilities cannot exist in a deterministic world, she will necessarily have to accept the truth of (Diff P/L). Compared to (AO), (Diff P/L) imposes an additional requirement on any counterfactual world in which the ability to do otherwise is exercised. This means, more specifically, that compatibilists are committed to:

(AO-COMP)   *Ability to do Otherwise (Compatibilism):* given P_t as a true proposition about the actual world, w, the ability to do otherwise (as the ability to 'render false P_t') is an ability such that all worlds, w', in which it is exercised, are worlds about which P_t is a false proposition and in which the past and/or the laws of nature are different.

Hereby, (AO-COMP) obviously only explicates the counterfactual structure of the ability to do otherwise in a deterministic universe. It does not provide a comprehensive account of the necessary and/or sufficient conditions for its existence. If the judge's hand happens to be paralyzed or if his hand has been tied down, the judge will *not* have the ability to raise his hands. In the first case, he lacks some necessary *internal ability*, whereas in the second case he does not have the *external opportunity*. Therefore, a comprehensive compatibilist account of alternative possibilities will have to explicate the internal abilities and external opportunities that have to be present in the actual sequence of events in order to attribute an ability to do otherwise to an agent. Such a more comprehensive account, as found, for instance, in Lehrer (1976, 1980), Campbell (1997) and Vihvelin (2013), is, however, not needed for the present purpose of assessing the soundness of the consequence argument in view of its underlying counterfactual structure.

5. The incompatibilist rejoinder

The compatibilist commitment to (AO-COMP) could give rise to an obvious rejoinder by the incompatibilist. She could argue that the notion of the 'ability to do otherwise' underlying the CS-rebuttal fails to capture the gist of the idea of having alternative possibilities at stake in the consequence argument. Of course, it is true, as expressed by (AO), that the ability to do otherwise is exercised in a different counterfactual world, w'. At the same time, the rejoinder goes, it is essential that this ability is also exercised *within our actual world*, in the sense that – prior to the action – both worlds, w and w', should be identical and that, therefore, P_0 and L should also be true in w'. Arguably, it is this kind of concern which leads van Inwagen (1977b: 96) himself to stress, in his reply to Narveson (1977), that P_0 (and by extension presumably also L) is supposed to be a *true* proposition.[3]

This rejoinder amounts to a rejection of (AO-COMP) and the concomitant imposition of a competing condition on the set of worlds in which the ability to do otherwise is exercised. This alternative requirement expresses the typical 'forking paths' conception of the ability to do otherwise that incompatibilists generally endorse:

(AO-INCOMP) *Ability to do Otherwise (Incompatibilism):* given P_t as a true proposition about the actual world, w, the ability to do otherwise (as the ability to 'render false P_t') is an ability such that all worlds, w', in which it is exercised, are worlds about which P_t is a false proposition and which are identical to w at a time shortly before the ability is exercised.

Importantly, however, the commitment to (AO-INCOMP) cannot by itself restore the truth of the original premise (4). Since (AO-INCOMP) already presupposes (AO), it can also, at best, support premise (4′). Although the rejoinder therefore *cannot* rescue the consequence argument in its original form, it *can* seemingly save its conclusion. Since (AO-INCOMP) implies that the conjunction of P_0 and L needs to be true in the counterfactual world, w', it now follows from (4′) that the agent could not have rendered P_t false.

More directly still, (AO-INCOMP) could simply be combined with the definition of determinism to yield the same result. On the one hand, (AO-INCOMP) directly implies that P_t is false about w'. On the other hand, (AO-INCOMP) also implies that P_0 and L are true about w' and hence – assuming determinism – that P_t is true about w'. Since there cannot exist a counterfactual world about which P_t is false and true at the same time, the ability to do otherwise cannot exist in a deterministic universe.

A prime example of an incompatibilist argument with this more direct structure is provided by John Martin Fischer's (1994: 62-65) so-called 'conditional version' of the consequence argument. The argument takes as its basic premises the Fixity of the Past (FP) and the Fixity of the Laws (FL):

(FP)　an agent cannot do A if her doing A would imply that some fact about the past of the actual world would not be a fact

(FL)　an agent cannot do A if her doing A would imply that some law which obtains in the actual world would not obtain

These principles clearly amount to a rejection of (AO-COMP) and a commitment to (AO-INCOMP), since they jointly imply that an agent can only do otherwise if this ability is exercised in a world with the same past and the same laws. It is, therefore, no surprise that the proof of the conditional version proceeds very swiftly. Fischer (1994: 62-63) simply combines (FP) and (FL) with the definition of determinism in order to yield the desired result that there are no alternative possibilities in a deterministic world.

This kind of incompatibilist rejoinder will, however, not impress the advocate of the CS-rebuttal for two related reasons. The rejoinder, first of all, no longer relies on the same *rationale* as the original argument. The consequence argument intended to show that our inability to change the past and the laws of nature 'transfers' into an inability to change the present. In this new argument, in contrast, neither our inabilities to change the past and the laws of nature, expressed by premises (5) and (6) in van Inwagen's version, nor the transfer principle, expressed by premise (4), play any role in deriving the conclusion. Although Fischer presents his argument as an alternative ('conditional') version of the consequence argument, this characterization is seriously misleading as his rejoinder, in reality, amounts to a completely different argument.

The second and related reason is that this new and different argument relies on premises which the compatibilist explicitly rejects. Although the compatibilist agrees that we do not have the ability to change the past or the laws of nature – as expressed in premises (5) and (6) of van Inwagen's argument –, premises (FP) and (FL) express very different ideas. They state that the ability to do otherwise needs to be exercised in a world in which the past and the laws of nature are the same as in the actual world (AO-INCOMP). And although this premise indeed implies that there are no alternative possibilities in a deterministic universe, this merely means that there are no alternative possibilities *in the forking path sense* expressed by (AO-INCOMP). Compatibilists, of course, readily accept that forking paths do not exist in a deterministic universe, but maintain that this is *not* the sort of alternative possibilities they are concerned with. For them, the ability to do otherwise is necessarily exercised in a world with a different past and/or different laws (AO-COMP). The incompatibilist rejoinder, however, provides no argument whatsoever against the existence of *this* kind of alternative possibilities

6. A dialectical stalemate

In order to assess the dialectics of the CS-rebuttal, it is important to appreciate the fact that the rebuttal does *not* constitute an argument against incompatibilism and that it also does *not* constitute an argument in favor of compatibilism. The CS-rebuttal as defended here simply aims to deny the soundness of the consequence argument on the basis of (AO), while remaining neutral with regard to the truth of its conclusion. What the CS-rebuttal thereby reveals is that the debate about the consequence argument ultimately hinges on an apparently irresolvable disagreement between incompatibilists and compatibilists with regard to whether either (AO-INCOMP) or (AO-COMP) provides the appropriate characterization of the counterfactual structure of alternative possibilities.

The suggestion that the debate about the consequence argument is characterized by a dialectical stalemate has, famously, been made by John Martin Fischer (1983: 132, 1994: 83-85) himself. According to his account, a stalemate situation is one where one side defends some principle *P* in order to construct an argument for some conclusion *C*, and where the other side, in turn, either denies *P* or provides an alternative principle *P**, which does not support the conclusion *C* (Fischer 1994: 83). Importantly, the argument at stake, here, cannot be the original consequence argument. The CS-rebuttal, relying on the uncontested principle (AO) as its main premise, shows that this original consequence argument is unsound. It reveals, at most, that the compatibilist has to commit herself to (AO-COMP) in order to uphold the existence of alternative possibilities in a deterministic world. The subsequent incompatibilist rejoinder, which insists that (AO-COMP) should be rejected in favor of (AO-INCOMP), fails to restore the soundness of the original consequence argument and can, at most, endorse a 'conditional version' of the consequence argument which explicitly relies on (AO-INCOMP) as its main premise and which, therefore, fails to convince the compatibilist. What we are left with, in the end, is a stalemate between two competing views, (AO-COMP) and (AO-INCOMP), about the counterfactual structure of the ability to do otherwise, but no sound 'consequence argument' able to resolve the matter one way or the other.

In the next three sections, I investigate three different arguments which explicitly claim to move beyond this dialectical stalemate in favor of the incompatibilist position. My analysis reveals, however, that these claims are unwarranted. The

argument presented by Holliday (2012), which I discuss in section 8, still relies on the contested premise (AO-INCOMP); the argument presented by Fischer and Pendergraft (2013), discussed in section 9, relies on a mistaken interpretation of (Diff P/L); and the argument by Fischer (1994), with which we start in the next section, relies on both.

7. A critique of Fischer's interpretation of (Diff P/L)

In his seminal book on *The Metaphysics of Free Will*, John Martin Fischer (1994) does not rest content with the 'conditional version' of the consequence argument we discussed before. He recognizes that (FP) and (FL) can be challenged by compatibilists who insist that there are situations in which an agent has a genuine ability to do otherwise and in which some counterfactual conditional of the form (Diff P/L) is true. In our terminology, this type of challenge amounts to a rejection of (AO-INCOMP) in the name of some version of (AO-COMP). Fischer (1994: 83-85) acknowledges that this debate represents an example of a dialectical stalemate but, nevertheless, aims to avoid the stalemate by presenting yet another, so called 'basic version' of the consequence argument (Fischer 1994: 88-93).

This newest version does not, however, bring us closer again to the original consequence argument. It constitutes, in reality, an argument which is very similar to the 'conditional version' presented before. The 'basic version' relies, more specifically, on what Fischer calls the Principle of the Fixity of the Past and Laws (PFPL) (Fischer 1994: 88, Fischer and Pendergraft 2013: 577):

(PFPL) an agent can only do *A* if her doing *A* can be an extension of the actual past, holding the laws fixed

This principle essentially combines (FP) and (FL) and, thus, similarly comprises a commitment to (AO-INCOMP). As before, this implies that a simple reference to the definition of determinism suffices to prove that (PFPL) entails that an agent cannot do otherwise in a deterministic world (Fischer 1994: 88). But again, as before, the compatibilist will not be impressed. Since she explicitly rejects (AO-INCOMP) in favor of (AO-COMP), she cannot feel bound by an argument which relies on (AO-INCOMP) as its main premise.

The main difference between the conditional and the basic version of the consequence argument consists in the fact that Fischer now relies on a possible-world semantics to analyze the truth of can-claims (the ability to do otherwise) on the one hand and the truth of counterfactual conditionals of the form (Diff P/L) on the other. With regard to the former, Fischer explains that (PFPL) implies that a can-claim can be true only if there exists some otherwise suitably related possible world with the same past and laws as the actual world in which the agent exercises his ability to do otherwise. With regard to the latter, Fischer uses a Stalnaker/Lewis analysis of counterfactuals. This means that the conditional 'if *P* were the case, then *Q* would be the case' is true if in the possible worlds most similar to the actual world in which *P* is true, *Q* is also true (Fischer 1994: 91).

These possible-worlds semantics now suggest, according to Fischer, that the truth conditions of the can-claims and the counterfactuals become independent in the sense that "they point us to different possible worlds" (Fischer 1994: 91). As a result, the truth of (PFPL) can become consistent with the truth of (Diff P/L). So, returning to van Inwagen's example, it could be true that the judge can raise his hand, because he does so in some (indeterministic) world with the same past and laws as the actual world, while it could *also* be true that if the judge had raised his hand, the past or the laws would have been different. Both claims can be true at the same time, Fischer argues, if we assume that the possible world relevant for the assessment of the can-claim is *not* in the set of possible worlds in which the judge raises his hand which are most similar to our actual world. In other words, a departure from our actual world to a world with the *same* past and laws in which the judge raises his hand (relevant for the can-claim) would take us further away than a departure to a world with a *different* past and/or laws in which the judge raises his hand (relevant for the truth of (Diff P/L)).

If this account were convincing, it would avoid, as Fischer claims, the dialectical stalemate under discussion because it would imply that the compatibilist commitment to (Diff P/L) becomes compatible with the truth of (PFPL) as an appropriate basis for the analysis of can-claims. As a result, the compatibilist commitment to (Diff P/L) could now no longer be used to challenge either (PFPL) or the basic version of the consequence argument which builds on (PFPL).

In response, it should be pointed out that this argument fails because Fischer's claim regarding the independence of the can-claims and the counterfactual conditionals is based on a misunderstanding of the underlying semantics of the latter. Although Fischer is right that an analysis of can-claims refers to different possible worlds than the Stalnaker/Lewis assessment of (Diff P/L), this point is irrelevant because compatibilists do *not* claim or assume that (Diff P/L) is true in the Stalnaker/Lewis sense. It is, in this regard, an advantage of (AO-COMP) over (Diff P/L) that (AO-COMP) more clearly shows that the compatibilist simply means to say that the consequent of the conditional needs to be true *about all counterfactual worlds in which the ability to do otherwise is exercised.* This conditional therefore *does* refer to the possible worlds relevant for an analysis of can-claims. (AO-COMP) imposes, more specifically, the necessary condition that these worlds have to belong to the subset of possible worlds with a *different* past and/or laws. Since (PFPL), as Fischer recognizes, imposes the exact opposite condition, (PFPL) and (AO-COMP) remain clearly incompatible and the dialectical stalemate remains fully intact.

As a possible rejoinder, the incompatibilist could point out that the compatibilist does not get to choose her modal semantics.[4] If the widely endorsed Stalnaker/Lewis approach does not suit her purposes, she cannot simply opt out. This response, however, fails to convince. (Diff P/L) does not represent an ordinary conditional but constitutes, rather, an admittedly still somewhat intuitive philosophical attempt to clarify the counterfactual commitments of the compatibilist position. (Diff P/L) should therefore not be subjected to a Stalnaker/Lewis semantics, but should be seen, rather, as a semantical claim itself, explicating the modal structure of the ability to do otherwise. In order to avoid all confusion in this regard, I propose that the compatibilist would do well to refrain from committing to (Diff P/L) as it is usually formulated and simply commit to (AO-COMP) instead as a principle which expresses the same modal structure but which avoids the use of possibly misleading conditionals. I shall return to the issue of the role and the truth conditions of (Diff P/L) in section 9 below.

8. Holliday's unrealizable action-types

In his contribution to this debate, Wesley Holliday (2012: 181) claims that he can prove the truth of (FP) and, ipso facto, break the stalemate identified by Fischer (1994: 83).[5] Of course, in order to fully deal with the stalemate and demonstrate the truth of (AO-INCOMP), we would, in addition, also need a proof of (FL). For the sake of analyzing his argument, however, I shall join Holliday (2012: 184) in simply assuming that if an agent had done otherwise, none of our actual laws of nature would have been different.

Holliday's intricate argument makes use of action-types defined across possible worlds. Such an action-type is a function which maps a possible world, w, and a time, t, onto the set of actions of a specific type in that world at that time. Importantly, this set of actions does not necessarily only contain actions that are actually performed in that world. We can illustrate the idea by means of the action-type which takes center stage in Holliday's (2012: 191) argument:

$I(w, t) = $ the set of actions inconsistent with the past relative to t of w

Actions are thereby inconsistent with the past if it is true that if the action were performed, the past relative to t would be different from the past in w. We can think here of the example of the judge who does not raise his hand at time t in the actual world. If we assume that the actual world is deterministic, it follows that the action of the judge *raising* his hand at time t is an action which is inconsistent with the past and thus an example of an action of type I in the actual world at time t.

For purposes of clarity, it is useful to also define the concept of an *unrealisable action-type*. Holliday himself makes use of the concept, but does not give it an explicit name. An action-type is, by definition, unrealisable if there is no possible world in which an agent actually performs an action of this type. As Holliday (2012: 198-199) points out, the action-type I of actions that are inconsistent with the past provides an example of such an unrealizable action-type. Although there are possible worlds in which our judge raises his hand, these worlds either have a different past or are indeterministic, which means that the action of the judge raising his hand will *not* be inconsistent with the past in these counterfactual worlds.

With these conceptual tools in place, Holliday (2012: 194) now presents a clear sketch of his own argument:

"if there is no world in which an agent performs an action of a certain type, then no agent *can* perform an action of that type; and since there is no world in which an agent performs an action that is inconsistent with the past (an action of type *I*), it follows that no agent can perform an action of that type, which establishes (FP)"

In terms of our example, this means that since the judge raising his hand in the actual (deterministic) world is an action of type *I* and since *I* is an unrealizable action-type, the judge *cannot* perform this action.

Although this reasoning might sound plausible at first sight, I believe that there are serious issues with the first premise in Holliday's argument. He provides a first, tentative, explication of this premise as follows (Holliday 2012: 194):

(1)   An agent *cannot* perform an action that belongs to an unrealizable action-type.

He immediately recognizes, however, that there are some problems with this version of the premise. The most serious one, in my view, relates to the fact that it turns out to be way too strong for his purposes. This follows from considering the action-type defined by the 'actions that the agents in the world will not perform', which provides another example of an unrealizable action-type. Applying (1) to this action-type, however, yields the *fatalist conclusion* that an agent cannot perform the actions he will not perform. This conclusion follows irrespective of whether the worlds we are talking about are deterministic or not and so (1) seems to rule out the existence of alternative possibilities across the board (Holliday 2012: 197).

In order to solve this problem, Holliday provides a further specification of premise (1), which – if we look through the technicalities of his formal notation – essentially claims the following (Holliday 2012: 198):

(1*)   If it is settled at time t that an action at time t_0 (with $t < t_0$) belongs to an unrealizable action-type, then the action *cannot* be performed by the agent at time t.

Thereby, the idea of propositions being 'settled' refers to an idea introduced by John Perry (2004). Holliday (2012: 193) stipulates, more precisely, that a proposition is *settled* at time t if and only if it is entailed by true propositions that are either made true by events prior to t or belong to the sort of propositions that are not made true by events. If we return to van Inwagen's original example about the judge, this definition

implies that, if determinism is true, it is settled at all times prior to the judge's not raising his hand in the actual world at t_0 that the counterfactual action of the judge raising his hand at t_0 is an action that belongs to the unrealizable action-type of actions that he will not perform. On the basis of (1*), it therefore follows that, at all times prior to t_0, the judge *cannot* raise his hand at t_0 (Holliday 2012: 200). In an indeterministic world, in contrast, it is not yet necessarily settled at all times prior to t_0 that the judge will not raise his hand at t_0 and, therefore, the conclusion that he *cannot* raise his hand does not follow (Holliday 2012: 197).

In my view, compatibilists should firmly reject premise (1*) for at least two (related) reasons. The specification of (1) into (1*) is, first of all, completely ad hoc. Although Holliday at first justifies this specification in view of a rather contrived scenario about actions forbidden by the gods at some particular time in history, it immediately becomes clear that it mainly serves the purpose of avoiding fatalism (Holliday 2012: 197). (1*) is tailored, more specifically, to the task of avoiding fatalism just enough to open up alternative possibilities in indeterministic worlds but still bar them from deterministic ones. In order to make the distinction between deterministic and indeterministic worlds, the new requirement that it should be settled at *t* that an action belongs to a certain type has to do all the work. In the example of the judge, it is noteworthy that it is *true in both the deterministic and the indeterministic cases* that the action of the judge raising his hand belongs to the *unrealizable* action-type of 'actions that the agent does not perform'. It is now merely because it was *not yet settled* shortly before the action that the action belongs to this action-type that we may allow for the possibility, in the indeterministic case, that the judge could have raised his hand. In the deterministic case, in contrast, the fact that this was already *settled* now guarantees that the judge could *not* have raised his hand. Whereas the original premise (1) drew its intuitive appeal from the suggestion that counterfactual actions could not be performed because they belonged to an *unrealizable* action-type, the shift to (1*) implies that the assessment of whether counterfactual actions could be performed now completely turns on whether or not it is *settled* in advance that they will be performed. Although Holliday presents the move from (1) to (1*) as a form of specification, (1*), in reality, represents a very different principle relying on very different intuitions for its plausibility.

In view of the fact that, in a deterministic world, it is always necessarily settled in advance which actions will and will not be performed, the unwarranted shift from (1) to (1*) is, of course, highly convenient from the incompatibilist point of view. The second and related reason for compatibilists to reject (1*) is, therefore, that it essentially amounts to a highly intricate reformulation of the incompatibilist intuitions contained in (AO-INCOMP). Premise (1*) is a very general premise which makes use of a universal quantifier – not made explicit in our informal rendering – which quantifies over all (unrealizable) action-types.[6] Taking into account that the four action-types defined as 'actions inconsistent with the past', 'actions inconsistent with the laws of nature', 'actions agents will not perform' and 'actions agents are determined not to perform' are all unrealizable action-types, it is clear that the following four claims are immediate *instantiations* of (1*):

(i)   if it is settled at t that an agent's doing A at time t_0 (with $t < t_0$) is inconsistent with the past, then she cannot do A at t

(ii)   if it is settled at t that an agent's doing A at time t_0 (with $t < t_0$) is inconsistent with the laws of nature, then she cannot do A at t

(iii)   if it is settled at t that an agent will not do A at time t_0 (with $t < t_0$), then she cannot do A at t

(iv)   if it is settled at t that an agent is determined not to do A at time t_0 (with $t < t_0$), then she cannot do A at t

This shows that premise (1*) already contains a commitment to a version of (FP) (i) and a version of (FL) (ii), and, thus, combining (i) and (ii), to (AO-INCOMP). By making use of the additional premise that, in view of the definition of determinism, it is settled that an agent 'will not' and is 'determined not to' perform a counterfactual action A, we can now directly deduce from (iii) and (iv) that alternative possibilities do not exist in deterministic worlds (Holliday 2012: 202-205). Although this argument is a valid one, it fails – pace Holliday – to move beyond the dialectical stalemate under discussion. The core premise (1*) needed to make it work is intuitively no more compelling than (AO-INCOMP) and, in fact, comprises (AO-INCOMP) as part of its instantiations. The attempt by Holliday to provide an independent argument for (AO-INCOMP) on the basis of less contested premises therefore fails.

9. Can we prove that the past is fixed?

More recently, John Martin Fischer and Garrett Pendergraft (2013) have made a renewed attempt to move beyond the dialectical stalemate by providing a new independent argument in favor of (PFPL). Their argument purports to show, more specifically, that the rejection of (PFPL) leads to unpalatable consequences for practical reasoning. In the deliberative perspective of an agent facing a choice, the rejection of (PFPL) would commit him to reasons for acting which are manifestly *unreasonable* (Fischer and Pendergraft 2013: 586-587). Throughout their argument, they are assuming the truth of (FL). Although this means that they are already assuming that a rejection of (PFPL) can only take the form of a rejection of (FP) rather than (FL), I shall join them in presupposing (FL) for the sake of my analysis.

The practical irrationality involved in the rejection of (PFPL) is demonstrated on the basis of an example containing a backtracking conditional already discussed in Fischer's earlier work (Fischer 1994: 95).

> "Consider the example of the Icy Patch. Sam saw a boy slip and fall on an icy patch on Sam's sidewalk on Monday. The boy was seriously injured, and this disturbed Sam deeply. On Tuesday, Sam must decide whether to go ice-skating. Suppose that Sam's character is such that if he were to decide to go ice-skating at noon on Tuesday, then the boy would not have slipped and hurt himself on Monday." (Fischer and Pendergraft 2013: 587)

Fischer and Pendergraft now focus on the backtracking conditional in the final sentence:

(BC) if Sam were to go ice-skating, the accident would not have happened

They argue that its truth seems to commit us to the following Irrational Conclusion:

(IC) It is open to Sam on Tuesday, by deciding to go ice-skating, to make the world contain the accident's not occurring on Monday.

Assuming that Sam is a nice guy who wants to help the boy, (IC) seems to provide him with a good reason to go ice-skating. Fischer and Pendergraft now claim that a commitment to (PFPL) allows us to reject this reason. Since Sam can only add to the existing past, all reasons flowing from the non-occurrence of the accident yesterday

are irrelevant for his practical reasoning. The rejection of (PFPL), in contrast, implies that Sam can do things such that if he did them, the world would have been different, and, therefore, the rejection of (PFPL) fails to explain why Sam should not commit to the manifestly irrational conclusion that he can bring about the non-occurrence of the accident (Fischer and Pendergraft 2013: 587-588).

In our terminology, the acceptance of (PFPL) amounts to the acceptance of (AO-INCOMP) and its rejection, therefore, to a commitment to (AO-COMP): when analyzing can-claims, we can either assume that the counterfactual world is identical to the actual world prior to the action (AO-INCOMP), or we can assume that it differs (AO-COMP). Fischer and Pendergraft's claim that (AO-COMP) implies a commitment to (IC) should, however, be firmly rejected for at least two different and independently sufficient reasons.

First of all, (AO-COMP) does not imply that we can undo anything about the actual past. As Holliday (2012: 186), for instance, recognizes, compatibilists do *not* claim that agents have "an incredible power to *change* the past, to undo events that [have] already occurred in history." Throughout this paper we have always assumed that compatibilists *agree* that we cannot render false true propositions about the past (the idea captured by premise (5) of van Inwagen's original argument). As a result, compatibilists can unhesitatingly side with the incompatibilists in rejecting all reasons following from the non-occurrence of the accident yesterday as irrelevant for Sam's reasoning.

In this context, it is important to bring to mind a distinction between the non-causal and the causal interpretation of the Fixity of the Past which Fischer (1994: 79) himself has introduced. The non-causal interpretation is, in fact, identical to (FP) as introduced before and is the interpretation we have been using throughout. It says that an agent cannot do something such that if he were to do it, the past would have *been* different. The causal version, (FPc), in contrast, reads as follows:

(FPc)   an agent cannot do something such that if she were to do it, she would thereby initiate a causal sequence issuing in the non-occurrence of some past event.

(FPc) is, however, a principle which is not under discussion in this debate since it is *accepted* by compatibilists and incompatibilists alike. Compatibilists merely reject (FP), in the sense that they are committed to the claim that the counterfactual scenario

in which Sam goes ice-skating on Tuesday, plays out in a world with a different past from ours. They are, thereby, however, *not* committed to the claim that, on the counterfactual scenario, Sam *causes* the past to be different.

Fischer and Pendergraft (2013: 588) seem to be aware of the problem and therefore explicitly deny that they are supposing that Sam can "initiate a backward-flowing causal chain issuing in the accident's not happening yesterday". Instead they are merely assuming that he can "make the world contain the accident's not happening" or "bring it about that the world did not contain the accident". The distinction they are trying to make here seems moot. But assume, for the sake of the argument, that we are prepared to introduce a third, 'making contain' version of the Fixity of the Past as follows:

(FPmc) an agent cannot do something such that if he were to do it, he would thereby make the world contain some different past event

Even then, however, it is still the case that compatibilists readily *accept* (FPmc) and thus *accept* that Sam *cannot* 'make the world contain' a different past. The denial of either (FPc) or (FPmc) would lead to the confused interpretation of what it means to render a proposition false which we mentioned when first introducing (AO). This denial would imply that an action could be such that, before the action, the agent is in the actual world, *w*, whereas, after the action, he is now in the counterfactual world, *w'*, with a different past. As compatibilists like Perry (2004: 246) and Vihvelin (2008: 315-316) recognize, this type of action, which changes the truth-value of a time-indexed proposition ('the accident happened on Monday') from true to false, is metaphysically impossible. Since (AO) excludes the metaphysically impossible types of action expressed by (FPc) and (FPmc), Fischer and Pendergraft misrepresent the denial of (PFPL) as a denial of (AO). Such a denial would, indeed, be very unreasonable. In reality, however, the compatibilist is firmly committed to (AO) and merely rejects (AO-INCOMP) – and, thus, (PFPL) – in favor of (AO-COMP).

In the same context, it is useful to point out that Fischer and Pendergraft in a similar manner misrepresent the meaning of the 'accessibility' relationship which compatibilists like Lehrer (1976) and Campbell (1997) use to analyse the ability to do otherwise. Lehrer, for instance, assumes that an agent *can* do something if there is at least one possible world minimally different from ours in which some condition

obtains which does not obtain in the actual world and which implies that the agent performs the action (1976: 247-248, 253-254). In order to ensure that agents also have the necessary internal abilities and external opportunities, Lehrer imposes the additional restriction that this possible world should not be different from the actual one in the sense of removing physical or psychological obstacles which prevent the agent from performing the action in the actual world (Lehrer 1976: 257-258). The set of worlds that meet all of these criteria is now defined by Lehrer as the set of worlds that are *accessible* to the agent. In the terminology we have been using, this means that the set of accessible worlds thus understood coincides with the set of worlds in which the agent exercises her ability to do otherwise. The condition (AO-COMP) thus expresses the compatibilist assumption that the set of accessible worlds is a subset of the worlds in which the past and/or the laws of nature are different compared to the actual world.

Fischer and Pendergraft now claim that this accessibility relation implies that "the agent has access to those worlds – he can get to those worlds from the actual world. Less metaphorically, he can actualize those worlds." (Fischer and Pendergraft 2013: 589). This claim is wrong, however, in the same sense as before. It mistakenly assumes that having access to a counterfactual world means that the agent can go from the actual world (in which the boy slips on the ice) into a different world (in which the boy does not slip on the ice). As emphasized, all compatibilists agree that this would be metaphysically impossible since all agents necessarily always act within their own actual world. The analysis of the set of accessible worlds only purports to clarify what it means to ascribe the ability to do otherwise to an agent in the actual world, it does *not* purport to clarify the set of worlds that he could 'go to' or 'actualize' in some metaphysically mysterious manner.

The second and independently sufficient reason for denying that the rejection of (PFPL) leads to a commitment to (IC) is that the truth of the back-tracking conditional (BC) does not imply, as Fischer and Pendergraft claim, that the set of accessible worlds (i.e. the set of worlds in which Sam exercises his ability to go ice-skating) contains a world in which the boy does not slip. Somewhat ironically, it was Fischer (1994: 91) himself, as we have seen, who has pointed out that an analysis of accessible worlds, i.e. an analysis of can-claims, points to different possible worlds than the Stalnaker/Lewis truth conditions of backtracking conditionals. Although

Lehrer's analysis of can-claims is based on a standard possible world semantics[7], these semantics are used to assess conditionals which are very different from the backtracking conditionals of the type (BC). In the case of Sam, the claim that he *can* go ice-skating could be true, for instance, because it is true that 'if Sam remembers that his therapist advised him to try to get over his anxieties by confronting his fears, then he goes ice-skating'. On the Stalnaker/Lewis semantics, this means that in the world or worlds most similar to the actual world (and meeting the appropriate additional restrictions intended by Lehrer) in which Sam remembers his therapist's advice, Sam goes ice-skating. Now, even if – for the sake of the argument – we join Fischer and Pendergraft in assuming that the story of the icy patch can be fine-tuned to simultaneously ensure the truth of (BC) in the Stalnaker/Lewis sense, this truth would *not* imply that the set of worlds relevant for Lehrer's can-claim are also contains a world in which the boy does not slip. Indeed, it is very well possible that the worlds ensuring the truth of the can-claim (i.e. the worlds most similar to ours in which Sam remembers the advice) are further removed from the actual world than the worlds ensuring the truth of (BC) (i.e. the worlds most similar to ours in which Sam goes ice-skating). In the case of the can-claims, the proposition that Sam goes ice-skating is part of the consequent of the conditional; in the case of the back-tracking conditional, that same proposition is part of the antecedent. And this, of course, makes a world of difference.

The overall argument of Fischer and Pendergraft reminds us again to be very cautious about the interpretation of (Diff P/L). Incompatibilists sometimes assume that (Diff P/L) implies that compatibilists are committed to backtracking conditionals of the type (BC) and that this commitment leads to 'irrational conclusions'. In reality, (Diff P/L) has no such implication whatsoever. Applied to this case, (Diff P/L) states that 'if Sam had gone ice-skating, the past and or the laws would have been different'. As we have emphasized before, the conditional formulation of this semantic rule is misleading as it mistakenly invites an interpretation on the basis of the Stalnaker-Lewis semantics. In order to avoid that confusion, the compatibilist should exchange her commitment to (Diff P/L) for a commitment to the more precise formulation of the same idea by means of (AO-COMP). What the compatibilist claims is simply that the set of worlds in which Sam exercises his ability to go ice-skating (i.e. the set of accessible worlds), only contains worlds with a different past and/or different laws. In

this case, some of these worlds are worlds in which Sam remembers the advice of his therapist.[8] On Lehrer's analysis, this can-claim commits the compatibilist to the truth of the conditional claim that 'if Sam remembers the advice of his therapist, he goes ice-skating'; it does *not* – pace Fischer – commit her to the truth of any backtracking conditional: *not* to the claim that 'if Sam were to go ice-skating, he would have remembered the advice of his therapist' and *not* to (BC). If we are clear about the counterfactual structure of the compatibilist analysis of the ability to do otherwise, the claim that compatibilism leads to irrationality remains unsubstantiated.

10. Conclusion

In this paper I have defended the CS-rebuttal of van Inwagen's consequence argument. An appreciation of the general counterfactual structure of alternative possibilities (AO) reveals that the argument is unsound and that it cannot be proven that the ability to do otherwise presupposes an ability to change either the past or the laws of nature. The argument, at best, forces the compatibilist to accept that the ability to do otherwise is exercised in a world in which the past and/or the laws of nature simply *are* different (AO-COMP). Here, the incompatibilist disagrees and claims that the ability to do otherwise needs to be exercised in a world with the same past and the same laws (AO-INCOMP). Although a commitment to (AO-INCOMP) allows the incompatibilist to prove that alternative possibilities do not exist in a deterministic world, this proof will not convince the compatibilist, who explicitly rejects (AO-INCOMP) in favor of (AO-COMP).

My analysis of attempts by Fischer (1994), by Holliday (2012) and by Fischer and Pendergraft (2013) to go beyond the dialectical stalemate in favor of the incompatibilist position reveals that they all fail in providing independent reasons in support of (AO-INCOMP). Their arguments either still rely on (AO-INCOMP) as one of their key premises or they are based on a misinterpretation of the commitments of the compatibilists expressed by (Diff P/L). In order to avoid this confusion, I recommend that compatibilists exchange their commitment to (Diff P/L) for a commitment to the more precise formulation of the same idea provided by (AO-COMP).

References

Campbell, Joseph K. (1997). A Compatibilist Theory of Alternative Possibilities. *Philosophical Studies*, 88, 319–330.

Fischer, John M. (1983). Incompatibilism. *Philosophical Studies*, 43, 127–137.

Fischer, John M. (1994). *The Metaphysics of Free Will: An Essay on Control*. Oxford: Blackwell Publishers.

Fischer, John M. and Garrett Pendergraft (2013). Does the Consequence Argument Beg the Question? *Philosophical Studies*, 166, 575–595.

Foley, Richard (1979). Compatibilism and Control over the Past. *Analysis*, 39, 70–74.

Gallois, Andre (1977). Van Inwagen on Free Will and Determinism. *Philosophical Studies*, 32, 99–105.

Holliday, Wesley H. (2012) Freedom and the Fixity of the Past. *Philosophical Review*, 121, 179–207.

Horgan, Terence (1985). Compatibilism and the Consequence Argument. *Philosophical Studies*, 47, 339–356.

Kapitan, Tomis (2002). A Master Argument for Incompatibilism? In Robert Kane (Ed.) *The Oxford Handbook of Free Will* (pp. 127–157). Oxford: Oxford University Press.

Lehrer, Keith (1976). "Can" in Theory and Practice: A Possible Worlds Analysis. In Myles Brand and Douglas Walton (Eds.) *Action Theory* (pp. 241–270). Dordrecht: Reidel.

Lehrer, Keith (1980). Preferences, Conditionals, and Freedom. In Peter van Inwagen (Ed.). *Time and Cause: Essays Presented to Richard Taylor* (pp. 187–201). Dordrecht: Reidel.

Lewis, David (1981). Are We Free To Break the Laws? In David Lewis. *Philosophical Papers. Volume II* (pp 291–298). Oxford: Oxford University Press.

Narveson, Jan (1977). Compatibilism Defended. *Philosophical Studies*, 32, 83–87.

Perry, John (2004). Compatibilist Options. In Joseph K. Campbell, Michael O'Rourke and David Shier (Eds.) *Freedom and Determinism* (pp. 231–254). Cambridge: The MIT Press.

Šuster, Danilo (2012). Lehrer and the Consequence Argument. *Philosophical Studies*, 161, 77–86.

Tognazzini, Neal A. and John M. Fischer (2017). Symposium on the Fixity of the Past. Incompatibilism and the Fixity of the Past. In John A. Keller (Ed.). *Being, Freedom and Method: Themes from the Philosophy of Peter van Inwagen* (pp. 140-148). Oxford: Oxford University Press.

van Inwagen, Peter (1975). The Incompatibility of Free Will and Determinism. *Philosophical Studies*, 27, 185–199.

van Inwagen, Peter (1977a). Reply to Gallois. *Philosophical Studies*, 32, 107–111.

van Inwagen, Peter (1977b). Reply to Narveson. *Philosophical Studies*, 32, 89–98.

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

Vihvelin, Kadri (2008). Compatibilism, Incompatibilism and Impossibilism. In Theodore Sider, John Hawthorne and Dean W. Zimmerman (Eds.) (2008). *Contemporary Debates in Metaphysics* (pp. 303–318). Oxford: Blackwell.

Vihvelin, Kadri (2013). *Causes, Laws, and Free Will: Why Determinism Doesn't Matter.* Oxford: Oxford University Press.

[1] Overviews of this debate can be found, for instance, in Kapitan (2002) and Vihvelin (2013: 155-166). Much of the debate has focused on the so-called 'modal versions' of the consequence argument. A more comprehensive account of the CS-rebuttal would also have to deal with these versions, but that task is beyond the scope of my present argument.

[2] In this regard, Lewis' argument resembles an earlier argument by Gallois (1977).

[3] 'If there is some state of affairs that entails the falsity of some *true* proposition about the way the world was before I was born, then I can't bring about (and never could have brought about) this state of affairs.' (van Inwagen 1977: 96) [italics in the original]

[4] I would like to thank an anonymous referee of this journal for pressing me to clarify this point.

[5] A criticism of Holliday's argument somewhat similar to the one presented here can be found in Tognazzini and Fischer (2017).

[6] An already somewhat stricter rendering of Holliday's (2012: 198) formal version yields '(1*): it is true for every action type X that if X is unrealisable, then if it is settled at time t that an action y at time t^0 $(t < t^0)$ by an agent s belongs to the action type X, then s cannot perform y at time t'.

[7] Lehrer (1976: 247) himself uses the possible worlds semantics introduced by Pollock, but indicates that a Stalnaker/Lewis approach would yield the same results.

[8] There could be many other scenarios in the set of accessible worlds. In one of them Sam's friend could, for instance, send him an invitation to go ice-skating together, which convinces Sam to go. This means that it would then also be true that 'if Sam receives an invitation of his friend, then he goes ice-skating'.