

# Antirealism and the conditional fallacy: the semantic approach

Patrick Girard · Luca Moretti

the date of receipt and acceptance should be inserted later

**Abstract** The expression *conditional fallacy* identifies a family of arguments deemed to entail odd and false consequences for notions defined in terms of counterfactuals. The antirealist notion of truth is typically defined in terms of what a rational enquirer or a community of rational enquirers would believe if they were suitably informed. This notion is deemed to entail, via the conditional fallacy, odd and false propositions, for example that there exists necessarily a rational enquirer. If these consequences do indeed follow from the antirealist notion of truth, alethic antirealism should probably be rejected. In this paper we analyse the conditional fallacy from a semantic (i.e. model-theoretic) point of view. This allows us to identify with precision the philosophical commitments that ground the validity of this type of argument. We show that the conditional fallacy arguments against alethic antirealism are valid only if controversial metaphysical assumptions are accepted. We suggest that the antirealist is not committed to the conditional fallacy because she is not committed to some of these assumptions.

**Keywords** Conditional fallacy · antirealism · counterfactual · conditional logic

## 1 Introduction: the Conditional Fallacy and Antirealism

The expression *conditional fallacy* identifies a family of arguments deemed to entail problematic or lethal consequences for notions analysed in terms of counterfactuals. These arguments apply to accounts of dispositional properties and response-dependent concepts. Saying that an object has a given disposition is roughly saying that the object is expected to react in a given way when properly stimulated. Counterfactual accounts of dispositions are formulated in terms of *a priori* (or necessarily) true biconditionals of the form:

---

P. Girard  
Department of Philosophy, The University of Auckland, 18 Symonds Street, Private Bag 92019,  
Auckland 1142, New Zealand. E-mail: p.girard@auckland.ac.nz

L. Moretti  
Philosophy Department, University of Aberdeen, Old Brewery, High Street, Aberdeen, AB24  
3UB, Scotland, UK. E-mail: l.moretti@abdn.ac.uk

Something  $x$  is disposed to give response  $R$  to stimulus  $S$  if and only if  $x$  would give response  $R$  if  $x$  were exposed to stimulus  $S$ .

For instance, a wire is live if and only if electrical current would flow from it to a conductor if the wire were touched by the conductor. The response-dependence account states, basically, that the use of certain concepts to describe given objects is licensed by the responses elicited from us by these objects in certain conditions. Response-dependence accounts are typically formulated in terms of *a priori* (or necessarily) true biconditionals of the form:

Something  $x$  falls under the notion  $N$  if and only if  $x$  would elicit response  $R$  from a subject  $S$  if  $x$  were in conditions  $C$ .

For instance, something is green if and only if it would look green to a standard human subject if it were closely observed by her in normal daylight. Conditional fallacy problems are deemed to emerge, typically, whenever it is true that if  $x$  were placed in the relevant situation  $C$ , this fact would *change*  $x$ 's actual dispositional/response-eliciting properties (cf. Wright [27]: 344-345, Bonevac et al. [2] : 273-282 and Gundersen [6]).<sup>1</sup> In these cases, the conditional analyst would be driven by logically compelling arguments to a false conclusion about  $x$ 's *actual* dispositional or response-dependent features.

Before illustrating this abstract description with an example, let us clarify the basic structure of the alleged logically compelling arguments. The conditional fallacy challenges counterfactual accounts based on the following schema:

$$(sc) \quad \Box(p \equiv (q \Box \rightarrow r))$$

Here  $\Box$  is the necessity operator,  $\equiv$  is the material biconditional, and  $\Box \rightarrow$  is the counterfactual conditional. *sc* says that, necessarily,  $p$  is the case if and only if were  $q$  to be the case,  $r$  would be the case. Conditional fallacy problems are deemed to emerge whenever  $q$  and  $p$  in *sc* happen to be related in such a way that the truth of  $q$  would affect the actual truth value of  $p$ . The simplest way to account for this relation is to use counterfactuals. So if  $p$  is actually true, a problem arises if (i)  $q \Box \rightarrow \neg p$  (where  $\neg$  is logical negation). And if  $\neg p$  is actually true, a problem arises if (ii)  $q \Box \rightarrow p$ . The contention is that *sc* in conjunction with either counterfactual produces a valid argument concluding that  $p$  is false when  $p$  is actually true, and vice versa. These are the schemata underlying the arguments:<sup>2</sup>

$$\begin{array}{ll} \Box(p \equiv (q \Box \rightarrow r)) & \Box(p \equiv (q \Box \rightarrow r)) \\ \text{(I) } \frac{q \Box \rightarrow \neg p}{\therefore \neg p} & \text{(II) } \frac{q \Box \rightarrow p}{\therefore p} \end{array}$$

The following case can be interpreted as exemplifying a version of the conditional fallacy based on II. Martin [13] suggests that we might think of defining the property of being live, instantiated by a metallic wire, through the following instance of *sc*: necessarily, ( $p$ ) a wire  $x$  is live if and only if ( $r$ ) electrical current would flow from  $x$  to a conductor if ( $q$ )  $x$  were touched by the conductor. Suppose now that it is the case that ( $\neg p$ ):  $x$  is a *dead* wire. Imagine however that  $x$  is connected to a reliable machine - an electro-fink - that detects whether  $x$  is touched

<sup>1</sup> The philosophical problem of the conditional fallacy was apparently introduced by Shope [22].

<sup>2</sup> In Lewis-Stalnaker conditional logic, schema I is valid only if  $q$  is possible.

by a conductor. When such contact occurs, the electro-fink reacts instantaneously by making the wire live for the duration of the contact (this means that  $q \Box \rightarrow p$ ). In this case, II allows us to derive the actually false statement that ( $p$ ): the wire  $x$  is live. Johnston [10]’s famous chameleon example is interpretable as a case of conditional fallacy of type I.<sup>3</sup>

The conditional fallacy has been argued to imply lethal consequences for alethic antirealism - shortly, antirealism. In this paper we will primarily focus on these alleged consequences. Since the conditional fallacy objections mounted against antirealism do not depend on I or II, we will not dwell on these two specific schemata here.<sup>4</sup> As we clarify below, the arguments against antirealism depend on *variants* of I and II. In this paper we analyse these variants in depth. As the antirealist disbelieves that truth can be completely evidence-transcendent, she characterises truth in terms of *epistemic* notions such as justification or rational acceptability. The antirealist appears to be committed to the following schema (or a close variant of it):

$$(AR) \Box(T(x) \equiv (Q(x) \Box \rightarrow R(x)))$$

Here  $x$  is a placeholder for declarative statements,  $T(x)$  means ‘it is true that  $x$ ’,  $Q(x)$  means ‘epistemic conditions are suitable for evaluating whether  $x$ ’, and  $R(x)$  means ‘it is rationally believed that  $x$ ’ (or ‘it is justified that  $x$ ’). AR says that, necessarily, it is true that  $x$  if and only if it would rationally be believed that  $x$  if epistemic conditions were suitable for evaluating whether  $x$ . Claiming that epistemic conditions are suitable for evaluating whether  $x$  is claiming - at the very least - that there is a rational enquirer/community suitably positioned or informed to evaluate whether  $x$  is the case. If the antirealist defined truth by appealing to *actual* - rather than *counterfactual* - suitable epistemic conditions, antirealism would have the implausible consequence that very many statements lack truth-value (presumably, all those whose suitable conditions of evaluation do not actually obtain). Since AR has the same propositional structure as SC, antirealism is *prima facie* vulnerable to the conditional fallacy objection.

To our knowledge the first philosopher who has explicitly raised a conditional fallacy objection against antirealism is Wright [27]. Wright focuses on the so-called Peircean conception of truth. This conception (or family of conceptions), holds that a statement  $x$  is true is to say that  $x$  would rationally be believed if epistemic conditions were *ideal*. Ideal epistemic conditions are, roughly, those in which there exists a rational enquirer/community that has acquired *all evidence* (empirical and rational). On this conception, *any* statement has the same ideal epistemic conditions of evaluation: the end of enquiry. If  $Q$  refers to these epistemic conditions, we can simplify AR to:

<sup>3</sup> Briefly, a chameleon is in the dark and is actually green. This creature is very shy and very perceptive: if another being were to spot it from a sufficiently close distance, the chameleon would instantaneously notice it and blush bright red. If we suppose that the concept of green is response-dependent and is defined in terms of SC, this is a case in which ( $p$ ) the chameleon is actually green but ( $q \Box \rightarrow \neg p$ ): if the chameleon were observed, it would not be green any longer. Schema I licenses the false conclusion that ( $\neg p$ ) the chameleon is actually not green. Note that Johnston’s chameleon example is also interpretable as a case of conditional fallacy of type II, for  $p =$  ‘the chameleon is red’.

<sup>4</sup> For an overview of the diverse conditional fallacy schemata and a general assessment of them see Bonevac et al. [2].

$$(AR-P) \quad \Box(T(x) \equiv (Q \Box \rightarrow R(x)))$$

It is *false* that epistemic conditions are actually ideal. We certainly do not possess all evidence (whatever this means). Suppose, however, that  $x$  just says that epistemic conditions are ideal, so  $x = Q$ . As the antirealist is committed to AR-P, the antirealist is also committed to  $\Box(T(Q) \equiv (Q \Box \rightarrow R(Q)))$ . Furthermore, since it is *a priori* true that if  $Q$  (i.e. epistemic conditions are ideal) then  $T(Q)$  (i.e. it is true that epistemic conditions are ideal), then it appears true that  $\Box(Q \supset T(Q))$  (where  $\supset$  is the material conditional), and the antirealist is committed to the latter strict conditional. Wright ([27]: 341-342) has produced a proof reducible to the claim that  $\Box(T(Q) \equiv (Q \Box \rightarrow R(Q)))$  and  $\Box(Q \supset T(Q))$  jointly entail  $T(Q)$ .<sup>5</sup> Since  $Q$  is patently false, this proof apparently refutes antirealism in its Peircean version. Wright's proof simplifies an earlier demonstration given by Plantinga ([17]: 64-66).<sup>6</sup> Plantinga's proof is reducible to the claim that  $\Box(T(Q) \equiv (Q \Box \rightarrow R(Q)))$  and  $\Box(Q \supset T(Q))$  jointly entail  $\Box T(Q)$ .<sup>7</sup> If this entailment actually holds, the Peircean version of antirealism implies that a *false* statement is not only true, but also *necessarily* true.

Wright's and Plantinga's proofs appear to be licensed by, respectively, these inference schemata:

$$(III) \quad \frac{\Box(p \equiv (q \Box \rightarrow r)) \quad \Box(q \supset p)}{\therefore p} \quad (IV) \quad \frac{\Box(p \equiv (q \Box \rightarrow r)) \quad \Box(q \supset p)}{\therefore \Box p}$$

Both III and IV can be obtained from inference schema II, considered before, by simply replacing the counterfactual in the second premise with a corresponding strict conditional, and by necessitating the conclusion in case of IV. The similarity of these three inference schemata and of their philosophical uses justify the common expression 'conditional fallacy' to refer to them.

Brogaard and Salerno ([3]: 135-137) have made another conditional fallacy objection to antirealism that appears to be grounded on IV.<sup>8</sup> Brogaard and Salerno complain that proofs like Plantinga's and Wright's risk - so to speak - shooting a dead horse. The point is that these proofs only apply to Peircean versions of antirealism, which demand commitment to the existence of *one* single epistemic

<sup>5</sup> Wright's original proof and Plantinga's original proof (which we consider below) rely on exclusively classical principles that are unacceptable for the intuitionist, so an antirealist following Dummett's work could easily reject them. However, Brogaard and Salerno ([3]: 131) have provided an intuitionistically respectable version of Plantinga's proof, which contains an embedded version of Wright's proof and is based on the theorem that  $\Box(p \equiv (q \Box \rightarrow r))$  and  $\Box(q \supset p)$  jointly entail  $\Box p$ . In this paper we focus on Brogaard and Salerno's reconstructions of Wright's and Plantinga's original proofs.

<sup>6</sup> It is worth stressing, however, that Plantinga is not arguing against antirealism - though he can certainly be reinterpreted to be doing so. Plantinga takes the conclusion of his proof to show only that the antirealist is committed to some form of *theism* (i.e. to the thesis that, roughly, there is necessarily an omniscient rational enquirer/community). Wright [27] and Brogaard and Salerno [3] respond that this is a misguided interpretation of his finding, which should instead be seen as an instance of the general problem of the conditional fallacy that plagues counterfactual analyses. As we find this response plausible, we will not question it in our paper.

<sup>7</sup> Plantinga apparently obtains a stronger result because he assumes that the accessibility relation is transitive, which is not presupposed by Wright.

<sup>8</sup> Brogaard and Salerno's proof improves upon the proof made by Rea [19]. Like Plantinga, Rea takes his proof to show that the antirealist is committed to some form of theism.

situation - the end of enquiry - appropriate for evaluating the truth of *any* statement. Peircean antirealism appears implausible on its own because, to begin with, it is hard to imagine what such a situation would be like.<sup>9</sup> Another problem is that it is even harder to understand how such a situation could be possible if rational enquirers are to be conceived of only as human. If it is acknowledged that the end of enquiry is humanly unattainable, it becomes unclear what explanatory advantage antirealism has over realism. For these reasons, Brogaard and Salerno work out a new conditional fallacy objection that targets more plausible and popular versions of antirealism that allow statements to have *individual* truth-conditions.

Brogaard and Salerno's objection aims to hit directly at AR rather than AR-P. Suppose  $p$  = 'epistemic conditions are suitable for evaluating whether some statement is true'.<sup>10</sup> Since the antirealist accepts AR, she is committed to  $\Box(T(p) \equiv (Q(p) \Box \rightarrow R(p)))$ . But the antirealist is also committed to  $\Box(Q(p) \supset T(p))$ . Here is a proof. Since  $Q(p)$  means 'epistemic conditions are suitable for evaluating whether  $p$ ', the antirealist is committed to:

Necessarily, if  $Q(p)$  then epistemic conditions are suitable for evaluating whether  $p$ .

Since evaluating whether  $p$  is *a priori* equivalent to evaluating whether it is true that  $p$ , the antirealist ought to accept that:

Necessarily, if  $Q(p)$  then epistemic conditions are suitable for evaluating whether it is true that  $p$ .

Since  $p$  is a statement, the antirealist is committed to:

Necessarily, if  $Q(p)$  then epistemic conditions are suitable for evaluating whether some statement is true.

Given that  $p$  means 'epistemic conditions are suitable for evaluating whether some statement is true', the antirealist ought to accept that the above strict conditional is equivalent to:

Necessarily, if  $Q(p)$  then  $p$ .

Since ' $p$ ' is *a priori* equivalent to 'it is true that  $p$ ', and since  $T(p)$  means 'it is true that  $p$ ', the antirealist is committed to inferring that:

Necessarily, if  $Q(p)$  then  $T(p)$ .

In conclusion, the antirealist appears committed to  $\Box(Q(p) \supset T(p))$ . Since the antirealist is committed to both  $\Box(T(p) \equiv (Q(p) \Box \rightarrow R(p)))$  and  $\Box(Q(p) \supset T(p))$ , through IV, the antirealist appears committed to  $\Box T(p)$ . That is, they are committed to the claim that, necessarily, it is true that epistemic conditions are suitable for determining whether some statement is true. These conditions include the existence of a properly placed rational enquirer. The antirealist is thus committed to claiming that, necessarily, it is true that there is a rational enquirer. It is actually true that there are rational enquirers, but this truth appears contingent and not necessary. So it is false that  $\Box T(p)$ .

The efficacy of conditional fallacy objections against antirealism has been questioned in the recent literature. For instance, Moretti [15] has argued that all proofs of the conditional fallacy made against antirealism hinge on the controversial assumption that a counterfactual with an impossible antecedent is vacuously true.

<sup>9</sup> The notion of the end of enquiry and even the notion of approximating the end of enquiry, is not perspicuous and might prove incoherent under close scrutiny. Interestingly, Wright himself ([26]: 44-48) has argued for this conclusion.

<sup>10</sup> An equivalent (and perhaps more precise) formulation of  $p$  is this: 'for some statement  $x$ , epistemic conditions are suitable for evaluating whether  $x$  is true'.

Moretti has also conjectured that any elementary reformulation of these proofs, when not relying on this assumption, will prove incorrect or unsound, or could be dismissed by an antirealist who had reasons to reject a modal system as strong as S5 (or its intuitionistic version). In this paper we criticise the conditional fallacy objection raised against antirealism but we aim to produce a more conclusive result.

A feature common to the proofs cited above is that they all aim to *derive* (syntactically) the conclusion rather than showing that the proofs are (semantically) *valid*. This may encourage the erroneous belief that the conditional fallacy does not depend on metaphysical assumptions but rather follows directly from principles of reasoning which are not peculiar to any specific metaphysical enterprise. Yet once the metaphysical assumptions necessary to III and IV are exposed, the proofs of the conditional fallacy appear questionable. In the following, we focus on no particular proof of III or IV, but on III and IV themselves.<sup>11</sup> We explicitly take a *semantic* approach; this will enable us to identify with accuracy the semantic and thus metaphysical commitments that ground the validity of these inference schemata. We will show that III and IV are valid only when we hold controversial metaphysical assumptions. We will suggest that the alethic antirealist who accepts AR (or its more specific variant AR-P) is not committed to these assumptions. Hence, she is not committed to the conditional fallacy.

The structure of the paper is as follows: In Section 2 we detail the logical framework that we use in the paper. In Section 3 we introduce semantic (or metaphysical) principles helpful for the analysis of the conditions of validity of III and IV. In Section 4 we single out one condition sufficient for the validity of III and one condition sufficient for the validity of IV given background assumptions acceptable by the antirealist. In Section 5 we show that these two conditions are also respectively necessary for the validity of III and IV given background assumptions acceptable by the antirealist. In Section 6 we suggest metaphysical reasons to doubt that these two conditions are satisfied. In Section 7 we draw the conclusions of the paper.

## 2 Modal Conditional Logic

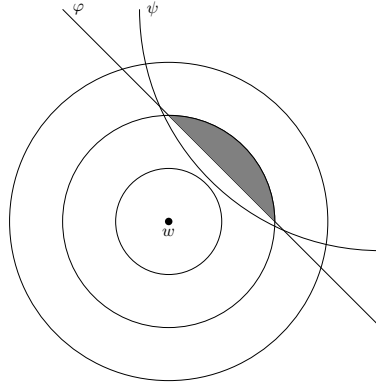
Call *Modal Conditional Logic* (MCL) any modal logic enriched with logical resources for counterfactuals (or subjunctive conditionals). The language of MCL, which we

<sup>11</sup> The main reason why we also focus on III, rather than on just IV, is that we cannot exclude *a priori* that once we allow statements to have *individual* truth-conditions, there might exist a statement  $s$  similar to the one used by Brogaard and Salerno that is intuitively false, or not known to be true, but such that  $\Box(Q(s) \supset T(s))$ . If this were the case, one could use III to produce a conditional fallacy argument against non-Peircean forms of antirealism. Suppose for instance  $P$  is a property that cannot possibly be detected. Also suppose that  $s =$  ‘this object has property  $P$ ’. It seems *impossible* that  $Q(s)$  - i.e. that epistemic conditions are suitable for evaluating whether  $s$  is true. Hence,  $\Box(Q(s) \supset T(s))$ . Thus, on III, we should conclude that  $s$  is true, which appears unjustified. Although the antirealist could dismiss the idea of a property like  $P$  and a statement like  $s$  as incoherent or absurd, there might be other statements capable of engendering similar problems that the antirealist cannot easily dismiss. As we show in Sections 4 and 5, the validity of *relation containment* together with some background assumptions is sufficient and necessary to validate III, and thus to get the antirealist into trouble. In Section 6 we provide metaphysical reasons to doubt that relation containment is satisfied.

refer to as  $\mathcal{L}_{\text{MCL}}$ , is based on a set  $\text{PROP}$  of propositional variables which can be combined recursively using the propositional connectives  $\neg$  and  $\vee$  (where the latter is logical disjunction), and the modalities  $\Box$  and  $\Box \rightarrow$ . The remaining connectives are defined in the usual way. For instance, the *material conditional*  $\varphi \supset \psi$  is defined as  $\neg\varphi \vee \psi$ , logical *conjunction*  $\varphi \wedge \psi$  is defined as  $\neg(\neg\varphi \vee \neg\psi)$ , and the *possibility* operator  $\Diamond\varphi$  by  $\neg\Box\neg\varphi$ .

In this paper, we only require minimal conditional semantics for our results. We follow Meyer and Veltman ([14]: 1017-1021), and propose a general semantic that is in some respect neutral between various trends in conditional logic, notably the tradition initiated by Lewis in [11] and Stalnaker [23].

Semantics for counterfactual conditionals attempt to capture the intuition that  $\varphi \Box \rightarrow \psi$  is true at a world  $w$  iff  $\psi$  is true in all  $\varphi$ -worlds that differ minimally from  $w$ . A useful graphical representation of counterfactual semantics based on this intuition, used in Lewis [11], employs a family of systems of spheres of worlds, in which each system is like this:



Amongst all  $\varphi$ -worlds, those shaded in grey are those that differ minimally from  $w$ , and that's where  $\psi$  is evaluated. In this picture, it is not the case that  $\psi$  is true in all  $\varphi$ -worlds. Yet the conditional  $\varphi \Box \rightarrow \psi$  is true, for  $\psi$  is true in all  $\varphi$ -worlds that differ minimally from  $w$ . It is quite controversial exactly how one should select these relevant  $\varphi$ -worlds. In order to attain our results, we will try to stay neutral as much as possible between different proposals: we will only require a very general “mechanism” of world selection, which can be complemented in various (metaphysically motivated) ways.

Formally, we take as a primitive notion of similarity between worlds a strict partial order<sup>12</sup>  $u <_w v$  for each world  $w$  of a set of worlds  $W$ , read as ‘world  $u$  is more similar to world  $w$  than world  $v$  is’. We also define a weak notion of similarity  $\leq_w$  for each world  $w$  of  $W$  by setting  $u \leq_w v$  iff  $u = v$  or  $u <_w v$ . We read  $u \leq_w v$

<sup>12</sup> A binary relation  $S$  is a strict partial order if and only if  $S$  is:

1. Irreflexive:  $\neg\exists x Sxx$ ;
2. Antisymmetric:  $\forall x, y : (Sxy \& Syx) \Rightarrow x = y$ ;
3. Transitive:  $\forall x, y, z : (Sxy \& Syz) \Rightarrow Sxz$ .

It is a well-known fact in logic that conditional logic cannot express the strictness of the similarity relation. We could have chosen to work with weak partial orders  $\leq$ , but instead we chose strict partial orders as they seem to be more commonly used by logicians.

as ‘world  $u$  is at least as similar to  $w$  as world  $v$  is’. We need this weak notion below. We call the ‘zone of entertainability of  $w$ ’ the field of the relation  $<_w$  and we write this as  $W_w$ . (The field of a relation is the union of its domain and image. So if  $u <_w v$ , then both  $u$  and  $v$  are in  $W_w$ ). Some logicians - e.g. Stalnaker [23] - impose an assumption on the relation  $<_w$ , called the *limit assumption*, which states that  $<_w$  is well-founded.<sup>13</sup> The limit assumption guarantees that for any  $\varphi$ -world  $u \in W_w$ , there is always at least one  $\varphi$ -world  $v$  such that (a)  $v \leq_w u$ , and (b) there are no other  $\varphi$ -worlds  $v'$  such that  $v' <_w v$ . When the limit assumption is accepted, any world  $v$  is called a *minimal*  $\varphi$ -world in  $W_w$  - or  $\varphi$ -world *closest to  $w$*  - if and only if  $v$  satisfies both (a) and (b). With the limit assumption in place, the semantics of  $\varphi \Box \rightarrow \psi$  says that  $\varphi \Box \rightarrow \psi$  is true at  $w$  just in case every minimal  $\varphi$ -world  $v \in W_w$  verifies  $\psi$ . We can drop the limit assumption if we deploy a more general semantics according to which - to use Lewis ([11]: 21)’s words -  $\varphi \Box \rightarrow \psi$  is true at  $w$  just in case:

If there are antecedent-permitting spheres, then as we take smaller and smaller ones [i.e. ones including worlds closer and closer to  $w$ ] without end, eventually we come to ones in which the consequent holds at every antecedent-world.

Below we formalise this idea in our semantics, for generality’s sake, although our results also hold with the limit assumption in place.

Another salient set of assumptions about the similarity relation  $<_w$  is that of *weak* and *strong* centering. Weak centering says that no world is more similar to  $w$  than  $w$  itself, and strong centering says that  $w$  is more similar to  $w$  than any other world. Lewis [11] and Stalnaker [23] accept both principles, though Lewis ([11]: 28-29) acknowledges that the weaker principle is somewhat more plausible than the stronger. Our results depend on assuming weak centering - we return to it below.

Whereas we use a relation of similarity as a basis for the semantics of counterfactuals, we use a notion of *accessibility* as a basis for the semantics of the modalities  $\Box$  and  $\Diamond$ . Deploying different accessibility relations permits one to discriminate among different types of possibility (or modality). For instance, the worlds that are *physically* possible for  $w$  are accessible from  $w$  through the relation of physical possibility, the worlds that are *metaphysically* possible for  $w$  are accessible from  $w$  through the relation of metaphysical possibility, and so on. As the focus of our discussion is specifically on metaphysical possibility, we will consider only one relation of accessibility  $R$  on  $W$  that identifies metaphysical possibility. Contrary to Lewis [11], we prefer to stand neutral on whether or not any world  $v$  in  $W$  entertainable from another world  $w$  in  $W$  must also be (metaphysically) accessible from  $w$ . This assumption is not mandatory if the accessibility relation is not universal (i.e. as long as it is false that each world in  $W$  has access to *any* world in  $W$ ). As we will see in Section 6, there are reasons for doubting that  $R$  is universal in this sense, and reasons for believing that worlds entertainable from another world are not always accessible from that world.

Models for MCL are based on frames  $\mathfrak{F} = \langle W, R, < \rangle$ , where  $W$  is a set of worlds,  $R$  is an *accessibility* relation on  $W$  and  $<$  is a family of *similarity* relations  $\{<_w\}_{w \in W}$

<sup>13</sup> A binary relation  $S$  is well-founded on a set  $U$  if and only if for every non-empty subset  $U'$  of  $U$ , there exists an element  $x \in U'$  such that, for every  $y \in U'$ , it is false that  $Syx$ .



on  $W$ , i.e. a strict partial order  $<_w$  for each world  $w \in W$ . In the remainder of the paper, we call  $R$  the *accessibility relation* and  $<$  the *similarity relation*, where  $<$  refers to *all* similarity relations in a frame generally considered.

A *model*  $\mathfrak{M}$  is a tuple  $\langle \mathfrak{F}, V \rangle$ , with  $\mathfrak{F}$  a frame and  $V$  a propositional valuation which assigns sets of worlds to each propositional variables of the set  $\text{PROP} - V(p)$  is the set of worlds in which  $p$  is true. We base our interpretation of arbitrary formulae of the language  $\mathcal{L}_{\text{MCL}}$  on this propositional valuation and we define the notion of *satisfaction at a world in a model*, written  $\mathfrak{M}, w \models \varphi$ , recursively:

$$\begin{aligned} \mathfrak{M}, w \models p & \quad \text{iff } w \in V(p) \\ \mathfrak{M}, w \models \neg\varphi & \quad \text{iff } \mathfrak{M}, w \not\models \varphi \\ \mathfrak{M}, w \models \varphi \vee \psi & \quad \text{iff } \mathfrak{M}, w \models \varphi \text{ or } \mathfrak{M}, w \models \psi \\ \mathfrak{M}, w \models \Box\varphi & \quad \text{iff } \mathfrak{M}, u \models \varphi \text{ for every } u \text{ s.t. } Rwu \\ \mathfrak{M}, w \models \varphi \Box\rightarrow \psi & \quad \text{iff for every } u \in W_w \text{ with } \mathfrak{M}, u \models \varphi, \text{ there exists a } u' \text{ s.t.} \\ & \quad \text{(i) } u' \leq_w u, \\ & \quad \text{(ii) } \mathfrak{M}, u' \models \varphi \text{ and} \\ & \quad \text{(iii) for every } u'', \text{ if } u'' \leq_w u' \text{ then } \mathfrak{M}, u'' \models \varphi \supset \psi. \end{aligned}$$

The generalised semantics for  $\varphi \Box\rightarrow \psi$  described above allows  $\varphi \Box\rightarrow \psi$  to be true at  $w$  in a model  $\mathfrak{M}$  even if there is some world  $u \in W_w$  at which  $\varphi$  is true but  $\psi$  false in  $\mathfrak{M}$ . What it is required for  $\varphi \Box\rightarrow \psi$ 's being true at  $w$  in  $\mathfrak{M}$  is this: *any* world included in  $W_w$  at which  $\varphi$  is true in  $\mathfrak{M}$  must have a world  $u'$  equally or more similar to  $w$  such that both  $\varphi$  and  $\psi$  are true at  $u'$  in  $\mathfrak{M}$ , and such that any other world  $u''$  more similar to  $w$  than  $u'$  with  $\varphi$  true in  $\mathfrak{M}$  always has  $\psi$  true in  $\mathfrak{M}$ . The reader may read this semantics for  $\varphi \Box\rightarrow \psi$  as stating that a counterfactual is true at a world  $w$  iff in any world *sufficiently* similar to  $w$  in which its antecedent is true, its consequent is also true.<sup>14</sup> With the limit assumption in place, the semantics of  $\varphi \Box\rightarrow \psi$  would simplify to:

$$\begin{aligned} \mathfrak{M}, w \models \varphi \Box\rightarrow \psi & \quad \text{iff } \mathfrak{M}, v \models \psi \text{ for every minimal } \varphi\text{-world } v \in W_w \\ & \quad \text{s.t. } \mathfrak{M}, v \models \varphi. \end{aligned}$$

We say that  $\varphi$  is *true at a world  $w$  in a model  $\mathfrak{M}$* , or *satisfied at  $w$  in  $\mathfrak{M}$*  if and only if  $\mathfrak{M}, w \models \varphi$ . We also say that  $\varphi$  is *valid in a frame  $\mathfrak{F}$*  if and only if  $\varphi$  is satisfied at every world  $w$  in every model  $\mathfrak{M}$  based on  $\mathfrak{F}$ . Finally, we say that  $\varphi$  is *valid in a class of frames  $K$*  if and only if  $\varphi$  is satisfied at every world  $w$  in every model  $\mathfrak{M}$  based on any frame in the class  $K$ .

### 3 Useful Semantic Principles

We would like to introduce here six semantic principles - most of which are probably quite familiar to the reader - that we will use to analyse the conditions of validity of III and IV, carried out in Sections 4 and 5, in the philosophical discussion in Section 6. These principles are: (1) weak centering of  $<$  (already mentioned above), (2) reflexivity of  $R$ , (3) transitivity of  $R$ , (4) symmetry of  $R$ , (5) relation containment of  $<$  inside  $R$ , and (6) quasi-transitivity of  $R$  given  $<$ .

<sup>14</sup> The minimal conditional logic with no specific assumptions such as the limit assumption or weak centering is typically called  $P$ , and has been axiomatised and proved complete in Burgess [4] and Veltman [24] independently.

*Weak centering* of  $<$ , as formulated in Meyer and Veltman ([14]: 1020), is the assumption that every world  $w$  is entertainable from itself and no other world is closer to  $w$  than  $w$  is to itself. Formally:  $\forall w \in W : w \in W_w \ \& \ \neg \exists v \in W_w \text{ s.t. } v <_w w$ . At first weak centering looks like quite a natural assumption. Despite this, some logicians have argued that it entails counterintuitive or problematic consequences (see Gundersen [7] and [8]). Yet these arguments can in turn be questioned (see, for instance, Hermes [9]: 114-121). An important logical reason for retaining weak centering is that this intuitive semantic principle validates counterfactual Modus Ponens (i.e.  $\varphi \Box \rightarrow \psi$  and  $\varphi$  entail  $\psi$ ), which appears independently plausible. (We will appeal to counterfactual Modus Ponens below). From a metaphysical point of view it also seems self-evident that no other world can be more similar to a world  $w$  than  $w$  itself. For these reasons we think that the antirealist can make a good case for retaining weak centering. Also note that those who have charged the antirealist with a conditional fallacy objection have used counterfactual Modus Ponens, and so have implicitly assumed weak centering. The antirealist is thus permitted to appeal to the same principle, from a dialectic point of view, in this dispute. In the following we will focus on classes of frames that do have weak centering.

*Reflexivity* of  $R$  is well-known. It states that each world is accessible to itself. Formally:  $\forall w \in W : Rww$ . Reflexivity of  $R$  is semantically equivalent to Axiom T:  $\Box \varphi \supset \varphi$  (i.e. if  $\varphi$  is necessary, then  $\varphi$  is true), which is customarily assumed in metaphysical contexts, because metaphysical necessity is factive.

*Transitivity* of  $R$  states that if a world  $v$  is accessible from a world  $w$ , and a world  $u$  is accessible from  $v$ , then  $u$  is also accessible from  $w$ . Formally:  $\forall w, v, u \in W : (Rwv \ \& \ Rvu) \Rightarrow Rwu$ .

*Symmetry* of  $R$  states that if a world  $v$  is accessible from a world  $w$ , then  $w$  is accessible from  $v$ . Formally:  $\forall w, v \in W : Rwv \Rightarrow Rvw$ .

*Relation containment* of  $<$  inside  $R$  is the assumption that every world  $v$  entertainable from  $w$  is also accessible from  $w$ . Formally:  $\forall w, v \in W : (v \in W_w \Rightarrow Rwv)$ .

*quasi-transitivity* of  $R$  given  $<$  - which we define for the first time here - says that if a world  $v$  is accessible from a world  $w$ , and there is a world  $u$  entertainable from  $v$ , then  $u$  is accessible from  $w$ . Formally,

$$\forall w, v, u \in W : (Rwv \ \& \ u \in W_v) \Rightarrow Rwu$$

We call this condition quasi-transitivity because if the second conjunct of its antecedent (i.e.  $u \in W_v$ ) is replaced by  $Rvu$  we obtain transitivity of  $R$ .

We will see in Section 6 that transitivity of  $R$ , symmetry of  $R$ , relation containment, and quasi-transitivity are philosophically more controversial than reflexivity of  $R$ . While the antirealist is very probably committed to reflexivity of  $R$  and - we have argued - can endorse weak centering, she is presumably not committed to the remaining principles. Before engaging in philosophical discussion let's do some logical work on the conditional fallacy.

#### 4 Proving the Conditional Fallacy

Recall that we are interested in these two forms of the conditional fallacy:

$$(III) \quad \frac{\Box(p \equiv (q \Box \rightarrow r)) \quad \Box(q \supset p)}{\therefore p} \quad (IV) \quad \frac{\Box(p \equiv (q \Box \rightarrow r)) \quad \Box(q \supset p)}{\therefore \Box p}$$

In this section we show that if weak centering is assumed, relation containment suffices for the validity of III, and quasi-transitivity suffices for the validity of IV. In the next section we will show that if weak centering and reflexivity of  $R$  are assumed, relation containment is also necessary for the validity of both III and IV, and quasi-transitivity is also necessary for the validity of IV.

According to the generalised semantics for counterfactuals described in Section 2,  $\varphi \Box \rightarrow \psi$  is false at a world  $w$  in a model  $\mathfrak{M}$  iff there exists a world  $u$  in  $W_w$  at which  $\varphi$  is true in  $\mathfrak{M}$ , but  $u$  has no world  $u'$  satisfying all conditions (i)-(iii). This can be the case either because  $u$  has no world  $u'$  equally or more similar to  $w$  with both  $\varphi$  and  $\psi$  true in  $\mathfrak{M}$ , or because, for any such world  $u'$  with  $\varphi$  and  $\psi$ , there is always another world  $u''$  more similar to  $w$  than  $u'$  with  $\varphi$  true and  $\psi$  false in  $\mathfrak{M}$ .<sup>15</sup> Importantly, note that a condition necessary for  $\varphi \Box \rightarrow \psi$ 's being false at  $w$  in  $\mathfrak{M}$  is that there must exist a world  $u \in W_w$  with  $\varphi$  true and  $\psi$  false in  $\mathfrak{M}$ .<sup>16</sup>

**Theorem 1** *For any class of frames  $K$  such that each frame  $\mathfrak{F}$  in  $K$  has weak centering, if relation containment holds for every  $\mathfrak{F}$ , then III is valid in  $K$ .*

*Proof* Let us prove this by contradiction. Consider an arbitrary class of frames  $K$  such that each frame has weak centering and relation containment. Take an arbitrary  $\mathfrak{F} \in K$  and an arbitrary model  $\mathfrak{M}$  based on  $\mathfrak{F}$  such that for a world  $w \in W$ ,  $\mathfrak{M}, w \models \Box(p \equiv (q \Box \rightarrow r))$  and  $\mathfrak{M}, w \models \Box(q \supset p)$ . Finally, assume for the sake of contradiction that  $\mathfrak{M}, w \not\models p$ . Since  $\mathfrak{F}$  has weak centering,  $x \in W_x$  for any  $x \in W$ . So relation containment implies that for any  $x \in W$ ,  $Rxx$  - i.e.  $\mathfrak{F}$  has reflexivity of  $R$ . Since  $\mathfrak{M}, w \models \Box(p \equiv (q \Box \rightarrow r))$ ,  $\mathfrak{M}, w \models \Box(q \supset p)$ , and  $R$  is reflexive,  $\mathfrak{M}, w \models q \supset p$  and  $\mathfrak{M}, w \models p \equiv (q \Box \rightarrow r)$ . So, considering that we have assumed  $\mathfrak{M}, w \not\models p$ , it follows that  $\mathfrak{M}, w \not\models q$  and  $\mathfrak{M}, w \not\models q \Box \rightarrow r$ . The second statement entails that there must exist a world  $v \in W_w$  such that  $\mathfrak{M}, v \models q$  and  $\mathfrak{M}, v \not\models r$ . Since we have assumed relation containment,  $Rwv$ . Hence, given that  $\mathfrak{M}, w \models \Box(p \equiv (q \Box \rightarrow r))$  and  $\mathfrak{M}, w \models \Box(q \supset p)$ , two consequences follow:  $\mathfrak{M}, v \models p \equiv (q \Box \rightarrow r)$  and  $\mathfrak{M}, v \models q \supset p$ . Since  $\mathfrak{M}, v \models q$ , the second consequence entails that  $\mathfrak{M}, v \models p$ , so we get from the first consequence that  $\mathfrak{M}, v \models q \Box \rightarrow r$ . Since we have assumed weak centering, counterfactual Modus Ponens for  $\Box \rightarrow$  gives us that  $\mathfrak{M}, v \models r$ , which leads to a contradiction, as we had selected  $v$  so that  $\mathfrak{M}, v \not\models r$ .  $\square$

**Theorem 2** *For any class of frames  $K$  such that each frame  $\mathfrak{F}$  in  $K$  has weak centering, if quasi-transitivity holds for every  $\mathfrak{F}$ , then IV is valid in  $K$ .*

*Proof* Let us prove this by contradiction. Consider an arbitrary class of frames  $K$  such that each frame in  $K$  has weak centering. Also assume that quasi-transitivity holds for every frame in  $K$ . Take an arbitrary frame  $\mathfrak{F} \in K$  and an arbitrary model  $\mathfrak{M}$  based on  $\mathfrak{F}$  such that for a world  $w \in W$ ,  $\mathfrak{M}, w \models \Box(p \equiv (q \Box \rightarrow r))$  and  $\mathfrak{M}, w \models \Box(q \supset p)$ . Assume for the sake of contradiction that  $\mathfrak{M}, w \not\models \Box p$ . The last assumption entails that there is a world  $v \in W$  with  $Rwv$  such that  $\mathfrak{M}, v \not\models p$ . Since  $Rwv$ , it follows from  $\mathfrak{M}, w \models \Box(p \equiv (q \Box \rightarrow r))$  that  $\mathfrak{M}, v \models p \equiv (q \Box \rightarrow r)$ . Since  $\mathfrak{M}, v \not\models p$ , then  $\mathfrak{M}, v \not\models q \Box \rightarrow r$ . So there must be a world  $u \in W_v$  such that  $\mathfrak{M}, u \models q$  and  $\mathfrak{M}, u \not\models r$ . Note now that since  $Rwv$  and  $u \in W_v$ , and given that

<sup>15</sup> When the limit assumption is accepted,  $\varphi \Box \rightarrow \psi$  proves false at  $w$  in  $\mathfrak{M}$  just in case there is a world  $u$  minimal in  $W_w$  at which  $\varphi$  is true but  $\psi$  false in  $\mathfrak{M}$ . When this condition is satisfied the more general condition for falsity of  $\varphi \Box \rightarrow \psi$  is also satisfied.

<sup>16</sup> This is true in both generalised and special semantics for counterfactuals.

quasi-transitivity holds in  $\mathfrak{F}$ , we obtain that  $Rwu$ . As  $\mathfrak{M}, w \models \Box(p \equiv (q \Box \rightarrow r))$  and  $\mathfrak{M}, w \models \Box(q \supset p)$ , two consequences follow:  $\mathfrak{M}, u \models p \equiv (q \Box \rightarrow r)$  and  $\mathfrak{M}, u \models q \supset p$ . From the second consequence, as  $\mathfrak{M}, u \models q$ , we get that  $\mathfrak{M}, u \models p$ . From this and the first consequence it follows that  $\mathfrak{M}, u \models q \Box \rightarrow r$ . Since we have assumed weak centering, counterfactual Modus Ponens gives us that  $\mathfrak{M}, u \models r$ . This leads to a contradiction because we had selected  $u$  so that  $\mathfrak{M}, u \not\models r$ .  $\square$

Theorem 2 enables us to determine a condition sufficient for IV's validity formulated in terms of weak centering, relation containment and transitivity of  $R$ .

**Corollary 1** *Any class of frames such that each frame has weak centering, relation containment and transitivity of  $R$  validates IV.*

*Proof* Consider an arbitrary class  $K$  of frames such that each frame has weak centering, relation containment and transitivity of  $R$ . Take an arbitrary frame  $\mathfrak{F} \in K$  and consider three arbitrary worlds  $u, v, w \in W$  such that  $Ruv$  and  $w \in W_v$ . Relation containment and  $w \in W_v$  jointly imply that  $Rvw$ . Since transitivity of  $R$  holds in  $\mathfrak{F}$ ,  $Ruv$  and  $Rvw$  jointly entail that  $Ruw$ . Thus any arbitrary  $\mathfrak{F}$  in  $K$  has quasi-transitivity. Thus, any class  $K$  of frames such that each frame has weak centering, relation containment and transitivity of  $R$  satisfies quasi-transitivity for any frame. By Theorem 2, any such class validates IV.  $\square$

## 5 Defusing the Conditional Fallacy

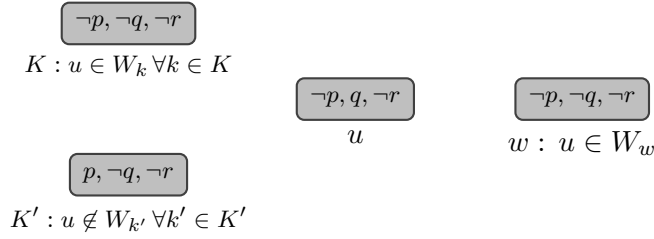
Model-theory can be used to understand the extent in which quasi-transitivity and relation containment are crucial for validating the proofs of the conditional fallacy against antirealism. We will now show that, given weak background assumptions, dropping relation containment *invalidates* both III and IV, and dropping quasi-transitivity *invalidates* IV.

We have indicated before that a condition sufficient for  $\varphi \Box \rightarrow \psi$ 's being false at a world  $w$  in a model  $\mathfrak{M}$  is that there is a world  $u$  in  $W_w$  with  $\varphi$  true in  $\mathfrak{M}$  that has no world  $u'$  equally or more similar to  $w$  with both  $\varphi$  and  $\psi$  true in  $\mathfrak{M}$ . The theorems we show below make crucial use of this condition.

**Theorem 3** *For any class of frames  $K$  such that each frame  $\mathfrak{F}$  in  $K$  has weak centering and reflexivity of  $R$ ,  $K$  validates III only if relation containment holds for every  $\mathfrak{F}$  in  $K$ .*

*Proof* We show this by proving that in an arbitrary class of frames  $K$  such that each of its frames has weak centering and reflexivity of  $R$ , if relation containment is not satisfied by an arbitrary frame  $\mathfrak{F}$  in  $K$ , III is not valid in  $\mathfrak{F}$ . Suppose then that  $\mathfrak{F}$  is an arbitrary frame.  $\mathfrak{F}$  must have the following features:  $<$  satisfies weak centering,  $R$  is reflexive, and there exist two worlds  $w, u \in W$  such that it is true that  $u \in W_w$  but false that  $Rwu$ . The strategy is to build a model on top of this  $\mathfrak{F}$  by giving a propositional valuation that satisfies the premises of III but not the conclusion, and which is not precluded by any possible additional feature of  $\mathfrak{F}$ . We consider only the case in which  $w \neq u$  because, since  $R$  is reflexive by assumption,  $Rwu$  necessarily holds if  $w = u$ .

This is a graphic representation of the model:



Worlds  $w$  and  $u$  are, by assumption, such that  $u \in W_w$  but it is false that  $Rwu$ .  $K$  stands for any world  $k \in W$  that is different from  $u$  and  $w$  and such that  $u \in W_k$ .  $K'$  stands for any world  $k' \in W$  such that  $u \notin W_{k'}$ . Since weak centering holds in this model, the similarity relations are all reflexive. The accessibility relation is also reflexive. It is assumed that any world can be linked to any other through the accessibility relation and similarity relations, provided that  $Rwu$  remains false and the distinction between worlds which can entertain  $u$  or cannot is respected. Finally we take a valuation  $V$  such that  $V(p) = \{x | u \notin W_x\}$ ,  $V(q) = \{u\}$  and  $V(r) = \emptyset$ .<sup>17</sup>

Note that for every  $x \in W$ ,  $\mathfrak{M}, x \models q \supset p$  with the sole exception of  $x = u$ . But it is false that  $Rwu$ . So  $\mathfrak{M}, w \models \Box(q \supset p)$ . Let us now prove that  $\mathfrak{M}, w \models \Box(p \equiv (q \Box \rightarrow r))$  by showing that for any  $x \in W$  such that  $x \neq u$ ,  $\mathfrak{M}, x \models p \equiv (q \Box \rightarrow r)$ .

Let us first show that  $\mathfrak{M}, w \models p \equiv (q \Box \rightarrow r)$ . As we have assumed that  $\mathfrak{M}, w \not\models p$ , we need to show that  $\mathfrak{M}, w \not\models q \Box \rightarrow r$ , and we do this by proving that there exists a world  $y \in W_w$  with  $\mathfrak{M}, y \models q$  such that  $y$  has no world  $x$  with  $x \leq_w y$  such that both  $\mathfrak{M}, x \models q$  and  $\mathfrak{M}, x \models r$ . This world  $y$  coincides with  $u$ .<sup>18</sup> To begin with, by assumption  $u \in W_w$  and  $\mathfrak{M}, u \models q$ . Furthermore, by assumption, for no world  $x \in W$ , both  $\mathfrak{M}, x \models q$  and  $\mathfrak{M}, x \models r$ . Thus  $u$  has no world  $x$  with  $x \leq_w u$  such that both  $\mathfrak{M}, x \models q$  and  $\mathfrak{M}, x \models r$ . So  $\mathfrak{M}, w \not\models q \Box \rightarrow r$ . Hence  $\mathfrak{M}, w \models p \equiv (q \Box \rightarrow r)$ .

Let us now prove that  $p \equiv (q \Box \rightarrow r)$  is true at any  $x \in W$  in  $\mathfrak{M}$  when  $x \neq w$  and  $x \neq u$ . Consider first any world  $k \in K$  (and so such that  $u \in W_k$ ). If we just replace  $w$  with  $k$  in the reasoning above we obtain that  $\mathfrak{M}, k \models p \equiv (q \Box \rightarrow r)$ . Consider now any world  $k' \in K'$  (and so such that  $u \notin W_{k'}$ ). Since  $u$  is the only world in  $W$  with  $q$  true in  $\mathfrak{M}$ , there exists no world  $x \in W_{k'}$  such that  $\mathfrak{M}, x \models q$ . So  $q \Box \rightarrow r$  is vacuously true at  $k'$  in  $\mathfrak{M}$ . Since by assumption  $p$  is also true at  $k'$  in  $\mathfrak{M}$ ,  $\mathfrak{M}, k' \models p \equiv (q \Box \rightarrow r)$ . Since all worlds in  $W$  verify this biconditional in  $\mathfrak{M}$ , with the sole exception of  $u$ , which is by assumption inaccessible from  $w$ , we can conclude that  $\mathfrak{M}, w \models \Box(p \equiv (q \Box \rightarrow r))$ . We also saw that  $\mathfrak{M}, w \models \Box(q \supset p)$  and  $\mathfrak{M}, w \not\models p$ . Hence III is invalid in  $\mathfrak{F}$ .  $\square$

**Theorem 4** *For any class of frames  $K$  such that each frame  $\mathfrak{F}$  in  $K$  has weak centering and reflexivity of  $R$ ,  $K$  validates IV only if quasi-transitivity holds for every  $\mathfrak{F}$ .*

*Proof* We show this by proving that in an arbitrary class of frames  $K$  such that each of its frames has weak centering and reflexivity of  $R$ , if quasi-transitivity is not satisfied by an arbitrary frame  $\mathfrak{F}$  in  $K$ , IV is not valid in  $\mathfrak{F}$ . Assume therefore

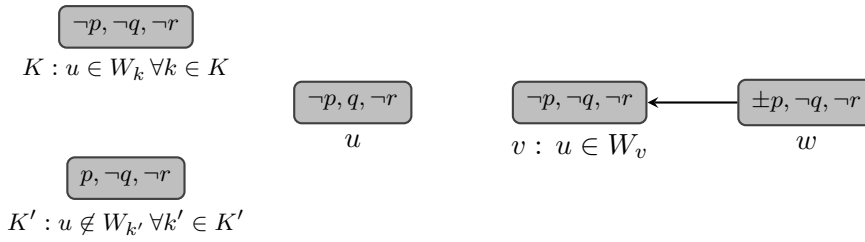
<sup>17</sup> That is to say,  $V$  makes  $p$  true at any world  $x$  in  $W$  such that  $u$  is not in the zone of entertainability of  $x$ , and also  $V$  makes  $q$  true at  $u$ .

<sup>18</sup> We assume that  $u$  is in  $W_w$  - and thus that  $u$  is a world relevant for the truth-value of  $q \Box \rightarrow r$  - even if it is false that  $Rwu$ . In Section 6 we give arguments in support of the general acceptability non-uncontroversial assumptions of this type.

that  $\mathfrak{F}$  has weak centering and that  $R$  is reflexive, and that there exist three worlds  $w, v$  and  $u$  in  $W$  such that  $Rwv, u \in W_v$  but not  $Rwu$ . We use the same strategy as before: we build a model on top of this  $\mathfrak{F}$  by giving a propositional valuation that satisfies the premises of IV but not its conclusion, and which is not precluded by any possible additional feature of  $\mathfrak{F}$ . It is easy to verify that, as  $R$  is reflexive,  $Rwu$  can prove false only if (1)  $w = v \neq u$  or (2)  $w \neq v \neq u$ .

Let us first consider case 1. If  $\mathfrak{F}$  is such that  $w = v \neq u$ , we are back to the situation considered in Theorem 3. That is,  $K$  has weak centering and reflexivity of  $R$ , and there exist two worlds  $w, u \in W$  such that it is true that  $u \in W_w$  but not that  $Rwu$ . Take again the propositional valuation used in Theorem 3 that makes both  $\Box(p \equiv (q \Box \rightarrow r))$  and  $\Box(q \supset p)$  true at  $w$  and  $p$  false at  $w$ . As  $Rww, \mathfrak{M}, w \not\models \Box p$ . So IV is invalid in  $\mathfrak{F}$ .

Let us turn to case 2. Consider the following model:



The arrow between  $w$  and  $v$  stands for the accessibility link  $Rwv$ . We assume, however, that  $u$  is not accessible from  $w$ . World  $v$  is such that  $u \in W_v$ .  $K$  stands for any world  $k \in W$  different from  $u, v$  and  $w$  and such that  $u \in W_k$ .  $K'$  stands for any world  $k' \in W$  different from  $w$  and such that  $u \notin W_{k'}$ . Since weak centering holds in this model, the similarity relations are all reflexive. The accessibility relation is also reflexive. We assume that any world can be linked to any other through the accessibility relation and similarity relations, provided that  $Rwu$  remains false and the distinction between worlds which can or cannot entertain  $u$  is respected. The valuation  $V$  is still such that  $V(p) = \{x | u \notin W_x\}$ ,  $V(q) = \{u\}$  and  $V(r) = \emptyset$ .<sup>19</sup> In  $w$ ,  $\pm p$  indicates that  $p$  could be either true or false in  $w$ , depending on whether  $u \in W_w$ .

As above, we make use of world  $u$  to prove that  $q \Box \rightarrow r$  is false at any other world in  $\mathfrak{M}$ . In this model, the only world where  $q \supset p$  is false is  $u$ , which is not accessible from  $w$ , so  $\mathfrak{M}, w \models \Box(q \supset p)$ . Moreover since  $Rwv$  and  $\mathfrak{M}, v \not\models p$ , we have that  $\mathfrak{M}, w \not\models \Box p$ , so the conclusion of IV is false at  $w$  in  $\mathfrak{M}$ . What remains to be shown is that  $\mathfrak{M}, w \models \Box(p \equiv (q \Box \rightarrow r))$ . Since  $u$  is not accessible from  $w$ , we can show this by proving that for any  $x \in W$  such that  $x \neq u$ ,  $\mathfrak{M}, x \models p \equiv (q \Box \rightarrow r)$ .

Let us first show that  $\mathfrak{M}, w \models p \equiv (q \Box \rightarrow r)$ . A difference between this model and the one used to prove Theorem 3 is that, now,  $u$  may be or may not be in  $W_w$ . If  $u \in W_w$ , our valuation  $V$  requires that  $\mathfrak{M}, w \not\models p$ , and the proof proceeds like in Theorem 3. In particular,  $u$  is a world in  $W_w$  such that  $\mathfrak{M}, u \models q$  and such that for no world  $x$  with  $x \leq_w u$ , both  $\mathfrak{M}, x \models q$  and  $\mathfrak{M}, x \models r$ . This suffices for concluding that  $\mathfrak{M}, w \not\models q \Box \rightarrow r$ . Given that  $\mathfrak{M}, w \not\models p$ , then  $\mathfrak{M}, w \models p \equiv (q \Box \rightarrow r)$ .

<sup>19</sup> That is to say,  $V$  makes  $p$  true at any world  $x$  in  $W$  such that  $v$  is not in the zone of entertainability of  $x$ , and also  $V$  makes  $q$  true at  $u$ .

Suppose now  $u \notin W_w$ . In this case  $V$  imposes that  $\mathfrak{M}, w \models p$ . Furthermore, there is no world  $x \in W_w$  such that  $\mathfrak{M}, x \models q$ . Thus  $q \Box \rightarrow r$  is vacuously true at  $w$  in  $\mathfrak{M}$ . So, again,  $\mathfrak{M}, w \models p \equiv (q \Box \rightarrow r)$ .

Consider now any world  $k \in K$  (and thus such that  $u \in W_k$ ). If we just replace  $w$  with  $k$  in the reasoning above, in the case in which  $u \in W_w$ , we show that  $\mathfrak{M}, k \models p \equiv (q \Box \rightarrow r)$ . Consider now  $v$ . By the same reasoning - after replacing  $k$  with  $v$  - we get that  $\mathfrak{M}, v \models p \equiv (q \Box \rightarrow r)$ . Consider finally any world  $k' \in K'$  (and thus such that  $u \notin W_{k'}$ ). Since  $u$  is the only world in  $W$  with  $q$  true in  $\mathfrak{M}$ , there exists no world  $x \in W_{k'}$  such that  $\mathfrak{M}, x \models q$ . Thus  $q \Box \rightarrow r$  is vacuously true at  $k'$  in  $\mathfrak{M}$ . Since by assumption  $p$  is also true at  $k'$  in  $\mathfrak{M}$ ,  $\mathfrak{M}, k' \models p \equiv (q \Box \rightarrow r)$ .

As all worlds in  $W$  do verify  $p \equiv (q \Box \rightarrow r)$  in  $\mathfrak{M}$  with the sole exception of  $u$ , which is by assumption inaccessible from  $w$ , we can conclude that  $\mathfrak{M}, w \models \Box(p \equiv (q \Box \rightarrow r))$ . We also saw that  $\mathfrak{M}, w \models \Box(q \supset p)$  and  $\mathfrak{M}, w \not\models \Box p$ . Hence IV is invalid in  $\mathfrak{F}$ .  $\square$

Here is an interesting corollary that exposes the centrality of relation containment in the proofs of the conditional fallacy against antirealism.

**Corollary 2** *For any class of frames  $K$  such that each frame  $\mathfrak{F}$  in  $K$  has weak centering and reflexivity of  $R$ ,  $K$  validates IV only if relation containment holds for every  $\mathfrak{F}$ .*

*Proof* Let us first show that any frame that has reflexivity of  $R$  and quasi-transitivity also has relation containment. Suppose that  $\mathfrak{F}$  has reflexivity of  $R$  quasi-transitivity, and consider two worlds  $v, u \in W$  such that  $u \in W_v$ . (If there are no such worlds, relation containment is vacuously satisfied). Now, since  $R$  is reflexive, we have  $Rvv$ . So  $u \in W_v$  and quasi-transitivity imply that  $Rvu$ . Therefore,  $\mathfrak{F}$  has relation containment. Consider now a class of frames  $K$  such that each frame  $\mathfrak{F}$  in  $K$  has weak centering and reflexivity of  $R$ . Given Theorem 4,  $K$  validates IV only if quasi-transitivity holds for every  $\mathfrak{F}$ . Considering what was just shown, this entails that  $K$  validates IV only if relation containment holds for every  $\mathfrak{F}$ .<sup>20</sup>  $\square$

To summarise, it turns out that satisfying relation containment is a condition necessary for validating both III and IV (given reflexivity of  $R$  and weak centering). On the other hand, since fulfilling relation containment (together with reflexivity of  $R$  and weak centering) does not imply fulfilling quasi-transitivity,<sup>21</sup> satisfying quasi-transitivity is an independent necessary condition for validating IV.

What philosophical significance do the above formal results have? It is hard to imagine arguments that would rationally compel the antirealist to drop weak centering or reflexivity of  $R$ . Consequently, the antirealist would commit a conditional fallacy, and would thus run afoul a reductio, *only if* she also accepted relation containment, necessary to validate both III and IV. Furthermore, the non-Peircean antirealist would commit a conditional fallacy *only if* she accepted

<sup>20</sup> Note that both proofs of Theorems 3 and 4 would work almost identically with further assumptions on  $\mathfrak{F}$ , such as the limit assumption, strong centering or connectedness of  $<$ , as it can easily be checked. For instance, in the case of connectedness of  $<$ , there would be no world  $k'$  such that  $u \notin W_{k'}$ . Thus  $K'$  would be empty, but the valuation  $V$  we have chosen would remain the same otherwise. Hence, even though we have phrased our theorems with very minimal conditional assumptions, the same proofs would work for more restrictive classes. This again demonstrates the generality of relation containment.

<sup>21</sup> Consider for instance the frame  $\mathfrak{F}$  with  $W = \{w, v, u\}$ , weak centering,  $R$  reflexive and such that  $Rwv$ ,  $u \in W_v$  and  $Rvu$  but not  $Rwu$ .  $\mathfrak{F}$  satisfies relation containment but not quasi-transitivity.

both relation containment and quasi-transitivity, necessary to validate IV. Now consider for instance the latter principle. Why should the antirealist be committed to maintaining that if a world  $w$  is also accessible from another world  $v$  and some possible world  $u$  is entertainable from  $w$ , then  $u$  is accessible from  $v$ ? Whoever makes a conditional fallacy argument against the antirealist should answer this question. It is apparent that the answer - if there is a definite answer - should be metaphysical in nature and not just logical.

Before leaving the logical ground to enter the metaphysical territory, we would like to dispel a possible objection. We have formulated our arguments in a classical setting, whereas the debate on antirealism has been intertwined with the debate between classical and intuitionistic logic. But we need not be concerned with such issues here, as Theorems 3 and 4 show the invalidity of an argument form in a classical context, from which the invalidity of the same argument form in an intuitionistic setting immediately follows. It would be interesting to formulate an intuitionistic modal logic expanded with counterfactuals and explore how our counter-examples could be adapted to such system, but we leave this (tricky!) task to the reader, as this would distract us from our main topic.

## 6 Metaphysical Reasons to Doubt Relation Containment and Quasi-Transitivity

The relation containment condition and the quasi-transitivity condition are expressed by universal conditionals, the consequents of which state that a possible world of a given type has access to a possible world. Either conditional can be falsified only if its consequent is falsified. This can happen only if the accessibility relation  $R$  is not universal - i.e. only if it is false that any possible world has access to any possible world. In contemporary philosophical logic there is no agreement on whether or not  $R$  is universal. In the following we review arguments adduced by philosophers to deny the universality of  $R$ , and we suggest ways in which the antirealist could further develop these arguments to make a case specifically against relation containment or quasi-transitivity.

Armstrong [1], while defending his *combinatorialist* theory of possibility, has found a reason to reject symmetry of  $R$ ; Nathan Salmon [20, 21] has raised an independent Sorites argument against transitivity of  $R$ . Let us start with Armstrong's argument. A combinatorialist theory of possibility is a metaphysical picture according to which possible worlds are rearrangements or recombinations of elementary individuals and elementary properties existing in the actual world. Simple individuals are those that do not have proper parts, and simple properties are those that do not have other properties as constituents. Simple individuals and simple properties exist only contingently and will have to be determined on the grounds of total science. Consider a simple property  $P$  actually instantiated by some individual and an actual simple individual  $a$  that does not instantiate  $P$ . The state of affairs  $Pa$  can be obtained by recombining  $a$  and  $P$ . Thus in the actual world it is *possible* that  $Pa$  - namely, possible worlds in which it is true that  $Pa$  are *accessible* from the actual world. Consider now all re-combinations of the actually instantiated properties with all existing simple individuals. One recombination is to the effect that  $a$  does not exist any more; another recombination is to the effect that  $P$  is no longer instantiated by any individual. These constitute gen-



uine possibilities for the combinatorialist - something that could have happened. Possible worlds in which  $a$  does not exist or  $P$  is not instantiated are thus accessible from the actual world. Consider now a possible world  $v$  that does not include  $a$  or  $P$ . The actual world is not accessible from  $v$  because it cannot be obtained by re-arranging the contingent elements constituting  $v$ , which do not include  $a$  or  $P$ . Hence, for the combinatorialist, the accessibility relation  $R$  is not *symmetric* and thus not universal.

Combinatorialism is affected by difficulties, but this is true of any interesting theory of possibility. We are not defending combinatorialism. Our point is that the antirealist appears *prima facie* entitled to endorse some version of combinatorialism - for instance, one according to which the contingent recombining items are mind - or language - dependent. The antirealist endorsing combinatorialism could adduce Armstrong's argument to contend that  $R$  is not universal.

There are other interesting arguments against the symmetry of  $R$ ; see for instance, among others, Wedgwood[25], Peacocke[16], Dummett ([5]: 328-348) and Quinn[18]. Some of these arguments are less general than Armstrong's, as they presuppose specific views in particular areas of philosophy. For example, Wedgwood ([25]) has made a case that non-reductive physicalism about mental properties - a view quite fashionable today - cannot be true if  $R$  is symmetric. We tend to believe that even "local" arguments of this type can shed doubts on the symmetry and therefore universality of  $R$ . Versions of these arguments might be adduced by the antirealist.

Let us turn to Salmon's argument. Consider an artefact - for example a table, which we will call  $T$ . It is intuitively plausible that  $T$ , while retaining its numerical identity, could have originated from a piece of tree trunk  $W_1$  very slightly different from the piece of tree trunk  $W_0$  from which  $T$  has actually originated. Suppose  $W_1$  has the same shape and size as  $W_0$  but is taken from one millimetre further down the same trunk as  $W_0$ . In short, it is *possible* that  $T$  is made of  $W_1$ . Consider now that if  $T$  had actually originated from  $W_1$ , it is plausible that  $T$  could have originated from  $W_2$  - a piece of wood taken from one additional millimetre further down the same trunk. Thus it is possible that it is possible that  $T$  is made of  $W_2$ . Let us reiterate this reasoning one thousand times to reach the apparently correct conclusion that it is possible that it is possible that it is possible that it is possible ... that  $T$  is made of  $W_{1000}$  - a piece of wood that differs from  $T$ 's actual piece of wood by one meter. If  $R$  is transitive, we can reduce this very long modal claim to the short claim that it is just possible that  $T$  is made of  $W_{1000}$ ; we can do so through the reiterated application of the principle of transitivity. But the short claim appears blatantly false: if  $T$  had originated from a piece of wood that differs from  $T$ 's actual piece of wood by one meter,  $T$  would be a distinct individual! So we had better not assume transitivity of  $R$ . Clearly, if  $R$  is not transitive,  $R$  is not universal. We are not endorsing Salmon's argument. Our point is simply that there is apparently no reason why the antirealist could not appeal to Salmon's argument when threatened by the conditional fallacy objection.

As we have said, the antirealist could develop Armstrong's or Salmon's argument to make a case against relation containment or quasi-transitivity. The antirealist endorsing combinatorialism could for instance argue against relation containment as follows: suppose a possible world  $u$  is obtained by recombining the actual world  $w$ . Imagine  $u$  differs from  $w$  because many elementary individuals existing in  $w$  do not exist in  $u$  any more, or because many elementary proper-

ties instantiated in  $w$  are no longer instantiated in  $u$ . Suppose furthermore that a world  $v$  is obtained through a very minimal recombination of  $u$  - e.g. the only difference between  $v$  and  $u$  is that an elementary individual that instantiates a given elementary property at  $u$  no longer instantiates that property at  $v$ . Given the composition of  $u$ ,  $v$  and  $w$ , it is intuitively correct to judge that  $v$  is *more similar* to  $u$  than  $w$  is. Note that whether or not  $w$  is possible for  $u$  appears irrelevant for the intuitive correctness of our judgement. Since  $<_u$  is meant to be a similarity relation,  $<_u$  should comply with our intuitive judgements of similarity. Thus we should conclude that  $v <_u w$ . Hence  $v, w \in W_u$ . As  $Ruw$  is false, relation containment should be dropped.

The antirealist appealing to combinatorialism could alternatively argue along these lines: suppose  $u$  is obtained by recombining the actual world  $w$  by just eliminating one of the elementary individuals of  $w$ . Since  $u$  is almost identical to our actual world, it is hard to deny that the first of the following counterfactuals appears true at  $u$  while the second appears false at  $u$ :

- (A) If an additional elementary individual existed, no pig would fly.
- (B) If an additional elementary individual existed, some pig would fly.

Although we know that none of the worlds that make the antecedents of A and B true are possible for  $u$ , this fact appears irrelevant when we judge the truth-values of A and B at  $u$ .<sup>22</sup> Modal Conditional Logic (MCL) can match our intuitive judgements of the truth-values of A and B at  $u$  only if we assume that (1) there is some world  $x$  with  $u <_u x$ , so that  $x \in W_u$ , at which it is true that an additional elementary individual exists, and (2) at any world  $x$  of this type *sufficiently* similar to  $u$ <sup>23</sup>, it is also true that no pig fly. If we insist that  $<_u$  is contained in  $R$ , so that there exists no world  $x$  of this type, we should conclude that A and B are *both* (vacuously) true, which seems incorrect. As a semantic theory that accounts for our pre-theoretical judgements about the truth-values of counterfactuals like A and B would be *ceteris paribus* preferable to a semantic theory unable to do so, it seems reasonable that MCL should comply with (1) and (2). Since there is some  $x$  such that  $x \in W_u$  but not  $Rux$ , relation containment cannot be accepted.

The antirealist who did not want to commit herself to combinatorialism could probably run a parallel argument that builds upon Salmon's case against transitivity of  $R$ . Consider again the table T that has been assumed to have originated from the piece of tree trunk  $W_0$  in our actual world  $w_0$ . It is hard to deny - the antirealist could contend - that the first of the following counterfactuals appears true at  $w_0$  while the second appears false at  $w_0$ :

- (C) If T had originated from  $W_{1000}$ , no pig would fly.
- (D) If T had originated from  $W_{1000}$ , some pig would fly.

<sup>22</sup> This argument and the next one parallel a well-known objection to Lewis [11]'s thesis that all counterfactuals with impossible antecedents are vacuously true. For instance, Mares [12] lists counterfactuals with impossible antecedents that appear true and counterfactuals with impossible antecedents that appear false. A difference between Mares' objection and the arguments we put forward here (on the antirealist's behalf) is that ours are meant to apply to counterfactuals with *locally* impossible antecedents - i.e. antecedents that are false at any world *accessible from a given possible world* - whereas Mares' objection is meant to apply to counterfactuals with *absolutely* impossible antecedents - i.e. antecedents that are false *at any possible world whatsoever*.

<sup>23</sup> Or that counts as *closest* to  $u$ , if we use the special counterfactual semantics.

Again, knowing that none of the worlds that make the antecedents of these counterfactuals true are possible for  $w_0$  appears irrelevant when we judge the truth-values of C and D. MCL can account formally for these intuitive judgements only if we assume that (1) there is some world  $x$  such that  $w_0 <_{w_0} x$  at which it is true that T has originated from  $W_{1000}$ , and (2) at any world  $x$  of this type sufficiently similar to  $w_0$  it is also true that no pig flies. Since there is some  $x$  such that  $w_0 <_{w_0} x$ , and so such that  $x \in W_{w_0}$ , but not  $Rw_0x$ , relation containment should be rejected.

Given that quasi-transitivity entails relation containment (once reflexivity of  $R$  is assumed), all arguments against relation containment considered so far are also arguments against quasi-transitivity. To finish, let us consider an alternative way in which the antirealist could exploit Salmon's argument to make a case against quasi-transitivity that does not affect relation containment - in other words, if this argument against quasi-transitivity is successful, relation containment need not be dropped. Let us simplify our language: consider again table T originated from piece of three trunk  $W_0$  in the actual world  $w_0$ . For any real  $n$ ,  $w_n$  will hereafter stand for a world in which T has originated from piece of tree trunk  $W_n$  taken from  $n$  millimetres further down the same trunk as  $W_0$ . Consequently, instead of saying that at  $w_n$  it appears possible that T is made of  $W_{n+1}$ , we will simply say that  $w_{n+1}$  appears possible for  $w_n$ . The antirealist could argue as follows:  $w_1$  appears possible for  $w_0$ ;  $w_2$  appears possible for  $w_1$ ;  $w_3$  appears possible for  $w_2$ ; and so on. Salmon shows that if we repeat this pattern one thousand times and then apply the principle of transitivity by the same number of times, we reach the incorrect conclusion that  $w_{1000}$  appears possible for  $w_0$  itself. But it is easy to see that we can reach the same conclusion by the reiterated application of the principle of quasi-transitivity alone, even if  $R$  is not transitive. Given that  $w_0$  appears possible for  $w_0$ ,  $Rw_0w_0$ . Consider now that  $w_1$  is more similar to  $w_0$  than  $w_2$  is. So  $w_1 <_{w_0} w_2$ , and thus  $w_1 \in W_{w_0}$ . Given quasi-transitivity, we get  $Rw_0w_1$  (this claim is not particularly controversial in this case, as  $w_1$  is intuitively accessible from  $w_0$ ). Given that  $Rw_0w_1$  and that  $w_2 <_{w_1} w_3$ , so that  $w_2 \in W_{w_1}$  by quasi-transitivity, we infer that  $Rw_0w_2$ . Now reiterate the reasoning:  $w_3$  is more similar to  $w_2$  than  $w_4$  is. So  $w_3 <_{w_2} w_4$  and  $w_3 \in W_{w_2}$ . Thus we obtain via quasi-transitivity that  $Rw_0w_3$ . And so on. If we reiterate this reasoning a sufficient number of times, we will arrive at the intuitively false conclusion that  $Rw_0w_{1000}$ . This is why quasi-transitivity cannot be accepted.

To prevent misunderstandings, let us emphasise one more time that we are not *making* or *endorsing* any of the arguments presented in this section. Though we are not taking position on the soundness of these arguments, it seems to us that none of them are obviously fallacious or unsuccessful. Our point is simply that there is apparently no reason why the antirealist could not appeal to some of these arguments, or variants of them, when threatened by the conditional fallacy objection.

## 7 Conclusions

In this paper, we have shown that the argument schemata that appear to ground typical conditional fallacy proofs mounted against alethic antirealism are valid in modal conditional logic only if certain semantic principles are accepted, which in

turn rest on non-trivial metaphysical assumptions. The latter are: quasi-transitivity of the accessibility relation given the similarity relation, and relation containment of the similarity relation inside the accessibility relation. An immediate consequence of this is that the conditional fallacy - at least when attributed to the antirealist - is not just a fallacy of *reasoning*, but one that depends on questionable philosophical assumptions. The antirealist does not appear committed to these two metaphysical assumptions. So long as her commitment to these assumptions is not convincingly demonstrated, no proof that the antirealist commits a conditional fallacy depending on III or IV is possible within modal conditional logic.<sup>24</sup>

Although our findings are important and encouraging for the antirealist, they do not put an end to the antirealist's struggle with the conditional fallacy challenge. For even though the validity of III and IV does depend on the assumptions exposed in Theorems 3 and 4, there might be *variants* of III and IV that commit the antirealist to a neighbouring conditional fallacy independent from these assumptions. Despite this, it seems to us that our paper supplies the antirealist with a quite general recipe for defusing conditional fallacy arguments: when threatened by one of these arguments, the antirealist should try to uncover its metaphysical assumptions and deny her commitment to them. This paper exemplifies two successful applications of this general strategy.

The focus of this paper has mainly been on alethic antirealism, but our semantic analysis of the conditional fallacy is utterly general and potentially applies to any use of the argument schemata III and IV. Anybody who is accused of having absurd dispositional or response-dependence notions, because they commit the conditional fallacy based on these argument schemata, can defuse the accusations if they are not committed to metaphysical principles that ground the validity of these schemata.

**Acknowledgements** We are very grateful to Franz Berto, Fred Kroon, Hannes Leitgeb, Chris Tucker, Jonathan McKeown-Green, Olivier Roy, Jeremy Seligman, Elia Zardini, Imogene Paterson, audience at the Munich Center for Mathematical Philosophy, and various referees for valuable discussions and important criticisms concerning drafts of this paper. We are grateful to Antti Keskinen and Ralph Wedgwood for drawing our attention to papers relevant for the discussion in Section 6.

---

<sup>24</sup> An objection has been made to us that the antirealist would still commit a conditional fallacy if there were some statement  $s$  with epistemic conditions  $Q(s)$  satisfiable *only* at worlds accessible from the actual world. This would be so because for any such  $s$ , the semantics of the counterfactual in  $\Box(T(s) \equiv (Q(s) \Box \rightarrow R(s)))$  would not require the antirealist to consider models with worlds inaccessible from the actual one, thus relation containment would be satisfied. In other words, for any such  $s$ , the semantics of  $\Box(T(s) \equiv (Q(s) \Box \rightarrow R(s)))$  could be contained in a class of frames that do have relation containment. This is an interesting criticism. Notice that for  $s$  to satisfy III in this case, it would also have to make  $\Box(Q(s) \supset T(s))$  true. Furthermore, the conclusion of the relevant instance of III would prove paradoxical or problematic only if  $s$  were known to be false, or at least if  $s$  were not known to be true. Finally, to have a useful result,  $s$  should have *individual* truth-conditions (see note 11 above). We do not see why the antirealist who rejected relation containment as valid generally should be committed to the existence of statements of this sort. Also note that, from a logical point of view, this objection seems to assume that there is a dependence between the propositional valuation and the accessibility and similarity relations in models, which imposes restrictions on class of frames. But this is nowhere to be found in the literature.

## References

1. David Malet Armstrong. *A Combinatorial Theory of Possibility*. Cambridge, 1989.
2. Daniel Bonevac, Josh Denver, and David Sosa. The conditional fallacy. *Philosophical Review*, 115:273–316, 2006.
3. Berit Brogaard and Joe Salerno. Anti-realism, theism and the conditional fallacy. *Nous*, 39:123–129, 2005.
4. John P. Burgess. Quick completeness proofs for some logics of conditionals. *Notre Dame Journal of Formal Logic*, 22(1):76–84, January 1981.
5. Michael Dummett. *The Seas of Language*. OUP, 1993.
6. Eline B. Gundersen. The chameleon’s revenge. response-dependence, finks and provisoed biconditionals. *Philosophical Studies*, 153:435–441, 2011.
7. Lars Gundersen. In defence of the conditional account of dispositions. *Synthese*, 130(3):pp. 389–411, 2002.
8. Lars Bo Gundersen. Outline of a new semantics for counterfactuals. *Pacific Philosophical Quarterly*, 85:1–20, 2004.
9. Charles M. Hermes. *Scientific essentialism and the Lewis/Ramsey account of laws of nature*. PhD thesis, Florida State University, 2006.
10. Mark Johnston. How to speak of the colors. *Philosophical Studies*, 68:221–263, 1992.
11. David Lewis. *Counterfactuals*. Harvard University Press, 1973.
12. Edwin D. Mares. Who’s afraid of impossible worlds? *Notre Dame Journal of Formal Logic*, 38:516–526, 1997.
13. C. B. Martin. Dispositions and conditionals. *Philosophical Quarterly*, 44:1–8, 1994.
14. John-Jules Meyer and Frank Veltman. Intelligent agents and common sense reasoning. In Johan Van Benthem Patrick Blackburn and Frank Wolter, editors, *Handbook of Modal Logic*, volume 3 of *Studies in Logic and Practical Reasoning*, chapter 18, pages 991 – 1029. Elsevier, 2007.
15. Luca Moretti. Brogaard and Salerno on antirealism and the conditional fallacy. *Philosophical Studies*, 140(2):229–246, 2008.
16. Christopher Peacocke. Metaphysical necessity: Understanding, truth and epistemology. *Mind*, 106:521–574, 1997.
17. Alvin Plantinga. How to be an anti-realist. In *Proceedings and addresses of the American philosophical association*, volume 56, pages 47–70, 1982.
18. Philip Quinn. Metaphysical necessity and modal logic. *The Monist*, 65:444–455, 1982.
19. Michael Rea. Theism and epistemic truth-equivalences. *Nous*, 34:291–301, 2000.
20. Nathan Salmon. The logic of what might have been. *Philosophical Review*, 98:3–34, 1989.
21. Nathan Salmon. This side of paradox. *Philosophical Topics*, 21:187–197, 1993.
22. Robert K. Shope. The conditional fallacy in contemporary philosophy. *The Journal of Philosophy*, 75(8):397–413, August 1978.
23. Robert Stalnaker. A theory of conditionals. In Nicholas Resher, editor, *Studies in Logical Theory*, pages 98–112. OUP, 1968.
24. Frank Veltman. *Logic for conditionals*. PhD thesis, Department of Philosophy, University of Amsterdam, 1985.
25. Ralph Wedgwood. The price of non-reductive physicalism. *Nous*, 34:400–421, 2000.
26. Crispin Wright. *Truth and Objectivity*. Cambridge University Press, 1992.
27. Crispin Wright. Truth as sort of epistemic: Putnam’s peregrinations. *Journal of Philosophy*, 97(6):335–364, 2000.