

# New Foundations for Counterfactuals

Franz Huber  
Department of Philosophy  
University of Toronto, Canada  
franz.huber@utoronto.ca  
<http://huber.blogs.chass.utoronto.ca/>

Penultimate version: please cite the final version from *Synthese*

July 7, 2014

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>The Principal Principle Without Humean Supervenience</b>	<b>10</b>
<b>3</b>	<b>Ranking Functions</b>	<b>15</b>
<b>4</b>	<b>The Royal Rule</b>	<b>20</b>
<b>5</b>	<b>Consequences</b>	<b>25</b>
<b>6</b>	<b>Appendix: Logical Considerations</b>	<b>30</b>

## Abstract

Philosophers typically rely on intuitions when providing a semantics for counterfactual conditionals. However, intuitions regarding counterfactual conditionals are notoriously shaky. The aim of this paper is to provide a principled account of the semantics of counterfactual conditionals. This principled account is provided by what I dub the Royal Rule, a deterministic analogue of the Principal Principle relating chance and credence.

The Royal Rule says that an ideal doxastic agent's initial grade of disbelief in a proposition  $A$ , given that the counterfactual distance in a given context to the closest  $A$ -worlds equals  $n$ , and no further information that is not admissible in this context, should equal  $n$ . Under the two assumptions that the presuppositions of a given context are admissible in this context, and that the theory of deterministic alethic or metaphysical modality is admissible in any context, it follows that the counterfactual distance distribution in a given context has the structure of a ranking function. The basic conditional logic  $\mathbf{V}$  is shown to be sound and complete with respect to the resulting rank-theoretic semantics of counterfactuals.

# 1 Introduction

Philosophers of language and philosophical logicians typically rely on intuitions when providing a semantics for counterfactual conditionals. However, intuitions regarding counterfactual conditionals are notoriously shaky.

As an example consider the debate between Stalnaker and Lewis about the validity of the so-called law of conditional excluded middle,  $(A \Box \rightarrow C) \vee (A \Box \rightarrow \neg C)$ . According to Stalnaker (1968) this principle is logically valid. Lewis (1973: 77ff) disagrees and presents the following counterexample:

- C It is not the case that if Bizet and Verdi were compatriots, Bizet would be Italian; and it is not the case that if Bizet and Verdi were compatriots, Bizet would not be Italian; nevertheless, if Bizet and Verdi were compatriots, Bizet either would or would not be Italian.

Stalnaker (1981: 91ff), according to whom the example originates with Quine (1950), defends an analysis which says that both

- C1 If Bizet and Verdi had been compatriots, Bizet would have been Italian.
- C2 If Bizet and Verdi had been compatriots, Verdi would have been French.

“are indeterminate – neither true nor false. It seems to me that the latter conclusion is clearly the more natural one. I think most speakers would be as hesitant to deny as to affirm either of the conditionals, and it seems as clear that one cannot deny them both as it is that one cannot affirm them both. Lewis seems to agree that unreflective linguistic intuition favors this conclusion.” (Stalnaker 1981: 92)

The reason for Stalnaker’s last claim is that Lewis (1973: 80) says: “I want to say [C], and think it probably true [...]. But offhand, I must admit, it does sound like a contradiction. Stalnaker’s theory does, and mine does not, respect the opinion of any ordinary language speaker who cares to insist that it is a contradiction.”

As Stalnaker (1981: 92) points out, “it would be arbitrary to require a choice of one of [C1 and C2] over the other, but [...] this is not at issue. What is at issue is what conclusion about the truth values of the counterfactuals should be drawn from the fact that such a choice would be arbitrary.”

The conclusion drawn by Lewis is that both C1 and C2 are false, and hence that conditional excluded middle is not logically valid. The conclusion drawn by Stalnaker is that both C1 and C2 are indeterminate, and that conditional excluded middle is logically valid. End of discussion.

An example from the more recent literature is the discussion between Gillies and Moss about where to draw the line between the semantics and the pragmatics of counterfactuals. Here is the relevant background. Lewis (1973: 10), referring to Sobel (1970), uses so-called Sobel sequences to argue against the analysis of the counterfactual conditional as a strict conditional. Sobel sequences are examples such as:

S1 If Sophia had a good offer from MIT, she would move to the East Coast.

S2 If Sophia had a good offer from MIT and if she had an even better offer from Caltech, she would not move to the East Coast.

Lewis assumes that counterfactuals such as these can be jointly true. This is not so if the counterfactual conditional is a strict, or necessarily true, material conditional. Lewis concludes that the counterfactual conditional is a *variably* strict conditional, a strict conditional whose strictness varies with the antecedent.

Gillies (2007) considers *reverse* Sobel sequences:

S2 If Sophia had a good offer from MIT and if she had an even better offer from Caltech, she would not move to the East Coast.

S1 If Sophia had a good offer from MIT, she would move to the East Coast.

For Gillies (2007: 332) “this sounds for all the world like a contradiction.” He then goes on to argue that the counterfactual conditional is a strict conditional, but one whose truth values interact with context in such a way that the order in which two counterfactual conditionals are asserted matters.

Moss (2012) defends Lewis’ analysis of the counterfactual conditional as a variably strict conditional. She admits that reverse Sobel sequences are “generally infelicitous”, though contrary to Gillies (2007) she does not consider them to be inconsistent. Moss first considers sentences other than counterfactual conditionals and notes that the order in which they are uttered matters:

M1 My car is around the corner.

M2 But cars get stolen in NYC all the time.

For her this conversation is felicitous, while the reversed one is not:

M2’ Cars get stolen in NYC all the time.

M1’ But my car is around the corner.

According to her “[o]ur intuitions [...] point towards a general principle governing assertability [... which ...] tells us that if a speaker cannot rule out a possibility made salient by some utterance, then it is irresponsible of her to assert a proposition incompatible with this possibility.” (Moss 2012: 568)

As before, “[w]hat is at issue is what conclusions about the truth values of the counterfactuals should be drawn” from the fact that reverse Sobel sequences are not felicitous. The conclusion drawn by Gillies is that the truth values of counterfactuals depend on the order in which they are uttered. The conclusion drawn by Moss is that the assertability conditions, but not the truth values, of counterfactuals depend on the order in which they are uttered. End of discussion.

Both examples<sup>1</sup> illustrate a common pattern. Two philosophers share a (linguistic) intuition – the arbitrariness of choosing between C1 and C2; the infelicity of reverse Sobel sequences – but disagree on the details of, or the best explanation for, this (linguistic) intuition. Where Lewis intuits falsity in addition to arbitrariness, or explains arbitrariness by falsity, Stalnaker intuits vagueness in addition to arbitrariness, or explains arbitrariness by vagueness. Where Gillies intuits inconsistency in addition to infelicity, or explains infelicity by inconsistency, Moss intuits unassertability in addition to infelicity, or explains infelicity by unassertability.

This is bad news. Our intuitions are supposed to be the evidence deciding between rival philosophical theories. If we cannot agree on what the data are, on whose intuitions to rely on, we cannot agree on which theory to believe. Nor can we agree on which theory to believe, if we cannot agree on which explanation is best. This is bad news, indeed. Besides discussions in the philosophy of language and in philosophical logic this affects other discussions involving counterfactuals: in epistemology knowledge is analyzed in terms of counterfactuals (Nozick 1981), in metaphysics causation is analyzed in terms of counterfactuals (Collins & Hall & Paul 2004), and in the philosophy of science dispositions are analyzed in terms of counterfactuals (Mumford 1998). Outside philosophy psychologists discuss regret and responsibility in counterfactual terms (Connolly & Ordóñez & Coughlan 1997). This list can be continued.

One reaction might be to go experimental (Knobe & Nichols 2008) and to see which intuitions, or subjective evaluations of the quality of explanations, are more widespread. On my view this would not help much, because truth is not a matter

---

<sup>1</sup>Another example is provided by the discussion between Lewis and Stalnaker versus Kratzer and Pollock about the semantic principle of Comparability according to which, roughly, any two worlds can be compared with respect to their similarity to the actual world. Cf. Lewis (1981: sect. 5) and references therein.

of democracy.<sup>2</sup> What we need, or at least what we should aim at, is a *principled* account of the semantics of counterfactuals.

I call an account principled if it is derived from a valid normative principle. A normative principle says that one ought to do something given that one has a certain goal. For instance, a normative principle might say that one should hold consistent beliefs given that one has the goal of holding only true beliefs. A normative principle is valid only if obeying the norm, doing what it says one should do, is a means to attaining the goal. The normative principle that one ought to hold consistent beliefs given that one has the goal of holding only true beliefs is valid only if holding consistent beliefs is a means to attaining the end of holding only true beliefs. One justifies a normative principle by establishing that obeying the norm is a means to attaining the goal. We may be able to justify the normative principle that one should hold consistent beliefs given that one has the goal of holding only true beliefs by proving that the consistency of one's beliefs is necessary for their joint truth. Of course, if one does not aim at true beliefs, that will not cut any ice. But that is besides the point. It is mistaking a hypothetical for a categorical imperative.

A principled account opens up the possibility of a justification that is not, or not exclusively, based on intuitions, or subjective evaluations of the quality of explanations, which vary across subjects and within subjects across contexts and over time. The reason is that a normative principle is justified by establishing that obeying the norm in question is a means to attaining the goal at issue. Whether obeying a norm is a means to attaining a goal is an objective matter of fact and not dependent on anyone's intuitions or subjective evaluations of the quality of explanations. What may not be an objective matter of fact is which cognitive goals we have or should have, or which cognitive goals an ideal doxastic agent has or should have.

Causal relevance and logical entailment are the prototypes of means-end relationships in science and philosophy. As an example consider classical logic. Interpreted as a set of norms for making inferences classical logic is justified relative to the goal of truth preservation in all logically possible worlds. The reason is that one's inferences attain the cognitive goal of truth preservation in all logically possible worlds if and only if they obey the norms of classical logic. The epistemological significance of the soundness of classical logic lies in establishing the

---

<sup>2</sup>I do not want to say that experimental philosophers hold that truth is a matter of democracy. If anyone is aware of the unreliability of intuitions, they are. What I want to say is that knowledge of how intuitions are distributed across various populations will not settle the philosophical issue at hand.

if-direction of this means-end relationship. The epistemological significance of the completeness of classical logic lies in establishing its only-if-direction. Further examples include non-monotonic logic which, again interpreted as a set of norms for making inferences, is justified relative to the different cognitive goal of truth preservation in all normal among the logically possible worlds (Kraus & Lehmann & Magidor 1990) and the probability calculus which, interpreted as a set of norms for organizing one's credences at a given moment in time, is justified relative to the cognitive goal of accuracy minimization (Joyce 1998; 2009). One important role played by logical and mathematical reasoning in philosophy, and by causal and statistical reasoning in science, is to establish these means-end relationships.

Let us return to counterfactual conditionals, or counterfactuals for short. Besides possible worlds their truth values depend on something else: selection functions in case of Stalnaker (1968) and spheres of similarity in case of Lewis (1973). I call this additional element *counterfactual distance*. Counterfactual distance figures as a primitive on my account. This does not mean it is void of content. Similar to Lewis' (1979: 472) "system of weights or priorities" we can say more about it when we study its relation to causality (Huber 2013), and we can even empirically confirm it (Huber ms). (For more on this see section 4.) Moreover, counterfactual distance can be put in relation to belief, just as objective chance can be put in relation to subjective credence. The central normative principle fixing the semantics of counterfactuals will do just that.

In this paper I am content with pointing out that a principled account in the sense explained is possible, and that it opens up the possibility for a justification of the semantics of counterfactuals that is not, or not exclusively, based on intuitions or subjective evaluations of the quality of explanations. The completion of this justification has to be left to another paper. More precisely, I will present two normative principles plus two assumptions about "admissibility" that together entail that the semantics of counterfactuals has a certain structure. I will present a means-end justification for the second normative principle according to which gradings of disbelief should obey the ranking calculus. The first normative principle is a deterministic analogue of Lewis' (1980) Principal Principle that is based on Spohn's (1988) theory of ranking functions. I dub it the *Royal Rule*.

The Royal Rule says that an ideal doxastic agent's prior subjective strength of disbelief in a proposition  $A$ , given that the counterfactual distance in a given context to the closest  $A$ -worlds equals  $n$ , and no further information that is not admissible in this context, should equal  $n$ . Contexts are represented as, roughly, functions from worlds to sets of worlds. Assume first that the presuppositions in

a given context are admissible in this context. Assume next that the theory of deterministic alethic or metaphysical modality is admissible in any context. Then it follows that the counterfactual distance distribution in a given context equals an ideal doxastic agent's prior subjective grading of disbelief conditional on the theory of deterministic alethic or metaphysical modality and the presuppositions in this context. Call this the first conclusion.

In Huber (2007) I argue for the thesis that all and only ranking functions are reasonable gradings of disbelief. The argument for this thesis is based on a theorem to the effect that an ideal doxastic agent's belief set is and will always be consistent and deductively closed just in case her gradings of disbelief obey the ranking calculus and she updates according to the update rules of ranking theory. Since a conditional ranking function is a ranking function the first conclusion implies that deterministic alethic or metaphysical modalities have the structure of ranking functions. Call this the second conclusion.

If we now substitute Spohn's ranking functions for Stalnaker's selection functions or Lewis' systems of spheres we get a rank-theoretic semantics of counterfactuals with respect to which the basic conditional logic  $\mathbf{V}$  is sound and complete. I thus side with Lewis against Stalnaker as far as conditional excluded middle is concerned, and with Moss against Gillies with regard to the question where to draw the line between the semantics and the pragmatics of counterfactuals.<sup>3</sup>

In three and a half companion papers I have explained how these rank-theoretic counterfactuals relate to causation (Huber 2011), how they fix a problem besetting Lewis' (1979) story about the relation between laws, counterfactuals, and particular facts that arises from to Arrow's impossibility result from social choice theory (Kroedel & Huber 2013), how they go beyond structural equations (Huber 2013), and how they can be empirically confirmed (Huber ms). In another companion paper I still have to provide a means-end justification of the Royal Rule (including the two assumptions about admissibility) with respect to a pertinent cognitive goal. This pertinent cognitive goal will be to stably hold true beliefs.<sup>4</sup>

---

<sup>3</sup>I side with Lewis and Stalnaker against Kratzer and Pollock regarding Comparability.

<sup>4</sup>Here is the idea. Once the ideal doxastic agent is certain of the relevant counterfactuals, the Royal Rule requires her to hold onto her conditional beliefs when she receives new, but admissible information. Given the relevant counterfactuals no additional information that is admissible may affect her conditional beliefs. Given the relevant counterfactuals the ideal doxastic agent's conditional beliefs are stable across all admissible information. Stable true beliefs are better than mere true beliefs in a way that is similar to how safety (Williamson 2000) makes knowledge better than mere true belief. My plan is to show that obeying the Royal Rule is a means to attaining the cognitive goal of stably holding true beliefs. This will justify the Royal Rule, and thereby also the derived semantics for counterfactuals, to the extent that one desires this cognitive goal.

Such a means-end justification is most important, for otherwise the Royal Rule would be on its own, and the principled account would merely relocate rather than reduce reliance on intuitions or subjective evaluations of the quality of explanations.

Here is the plan for the rest of this paper. In the next section I will give an unorthodox formulation of Lewis' Principal Principle that does not rely on Lewis' Humean supervenience assumption with respect to chances. Then I will introduce ranking functions. These sections pave the way for a formulation of the Royal Rule that is sensitive to the supervenience of deterministic alethic or metaphysical modalities on non-modal matters of particular fact. The latter supervenience on the metaphysical side will turn out to correspond to the iterability of the counterfactual conditional operator on the logical side. After deducing the first and second conclusion from above I will conclude by discussing some consequences. In an appendix I will prove that the basic conditional logic  $\mathbf{V}$  is sound and complete with respect to the rank-theoretic semantics of counterfactuals.

## 2 The Principal Principle Without Humean Supervenience

In his (1980) Lewis presents the following normative principle:

**Principal Principle 1** *Let  $\text{Pr}$  be an ideal doxastic agent's initial credence function, let  $t$  be a point of time, let  $x$  be a real number in the unit interval, let  $ch_t(A) = x$  be the proposition that the objective chance at time  $t$  that  $A$  holds equals  $x$ , and let  $E_t$  be a proposition that is admissible at  $t$ . Then*

$$\text{Pr}(A \mid ch_t(A) = x \cap E_t) = x.$$

The Principal Principle for probabilities says that an ideal doxastic agent's initial, or a priori, credence in a proposition  $A$ , given that the objective chance of  $A$  at time  $t$  equals  $x$ , and no further information that is not admissible at  $t$ , should equal  $x$ .

Lewis argues that the complete history of world  $w$  up to time  $t$ ,  $H_{wt}$ , is admissible at  $t$  as well as that  $w$ 's theory of indeterministic alethic modality,  $w$ 's theory of chance,  $T_w$ , is admissible at any time. This allows him to "reformulate" the Principal Principle as follows:

**Principal Principle 2** *Let  $\text{Pr}$  be an ideal doxastic agent's initial credence function. Then for any point of time  $t$ , world  $w$ , and proposition  $A$*

$$ch_{wt}(A) = \text{Pr}(A \mid H_{wt} \cap T_w).$$

Here  $ch_{wt}$  is the chance distribution of world  $w$  at time  $t$ , and  $A$  is a proposition in the domain of  $ch_{wt}$ . The thesis that an ideal doxastic agent's credence function should obey the probability calculus, and the theorem that conditional probability measures are probability measures, then allow Lewis to conclude that objective chances obey the probability calculus.

A mathematical oddity in this argument is that Lewis is working with non-standard probabilities to ensure that all propositions  $ch_{wt}(A) = x$  receive non-zero probability in the sense of  $\text{Pr}$  (Bernstein & Wattenberg 1969). Otherwise the conditional probabilities  $\text{Pr}(A \mid ch_t(A) = x \cap E_t)$  may not be defined. In order to avoid this problem, Spohn (2010) considers propositions of the form:  $ch_t(A) \in (a, b)$ , where  $a < b$ . Another option may be to refer to Popper-measures instead of classical probabilities (Popper 1955, Rényi 1955, Spohn 1986, Stalnaker 1970). Such manoeuvres will not be necessary in our case.

A philosophical oddity in this argument arises in connection with Lewis' assumption that chances supervene on local matters of particular fact (Lewis 1986). In order to get clear about what this assumption amounts to let  $W$  be the set of all *factual* worlds, where a factual world  $w$  completely specifies all *non-modal* matters of fact, all local matters of particular fact. Worlds completely specifying both all non-modal matters of fact as well as all modal matters of "fact" are called universes. The reason for this slightly deviating terminology is that it allows us to get clear about the implications of Lewis' Humean supervenience assumption with respect to chances for the Principal Principle. For the time being the only modalities considered will be indeterministic alethic modalities or chances.

A *universe*  $u$  is a pair  $(w_u, ch_u)$  consisting of a purely factual component  $w_u$  from the set of factual worlds  $W$  and a purely (indeterministic alethic) modal component  $ch_u$  from the set of potential chance measures  $CH$  defined on some field or algebra over  $W$ ,  $\mathcal{A}_W$ . The set of all universes  $U$  is a (not necessarily proper) subset of the Cartesian product  $W \times CH$  of the set of factual worlds  $W$  and the set of potential chance measures  $CH$  on  $\mathcal{A}_W$ . Let  $\mathcal{A}_U$  be a field or an algebra of propositions over  $U$ . The ideal doxastic agent's initial credence function  $\Pr$  is defined on  $\mathcal{A}_U$  and so assigns credences  $\Pr(B)$  to propositions  $B$  from  $\mathcal{A}_U$ .

Lewis' Humean supervenience assumption with respect to chance implies the following claim. For every<sup>5</sup> factual world  $w$  in  $W$  there is *exactly one* potential chance measure  $ch_w$  in  $CH$  such that  $(w, ch_w)$  is in  $U$ . Therefore, for any two universes  $u = (w_u, ch_u)$  and  $u' = (w_{u'}, ch_{u'})$  from  $U$ : if  $w_u = w_{u'}$ , then  $ch_u = ch_{u'}$ , and hence  $u = u'$ . This means that a potentially modal claim such as  $ch_t(A) = x$  amounts to a purely factual claim, and, more generally, that the algebra of potentially modal propositions  $\mathcal{A}_U$  reduces to the algebra of purely factual propositions  $\mathcal{A}_W$ . Please note that it does not make sense to speak of the chance distribution of a factual world  $w$  at a certain time  $t$ ,  $ch_{wt}$ , or a factual world's theory of chance,  $T_w$ , if the chances do not supervene on the totalities of non-modal, local matters of particular fact. Without Humean supervenience with respect to chance we have to speak of the chance distribution of universe  $u = (w_u, ch_u)$  at time  $t$ ,  $ch_{ut}$ , and of a universe's theory of chance,  $T_u$ .

Humean supervenience with respect to chance allows Lewis to iterate chances so that it makes sense to speak of the chance at an earlier time  $t_0$  of the chance at a later time  $t_1$  that some factual proposition is true,  $ch_{t_0}(ch_{t_1}(A) = y) = x$ .

---

<sup>5</sup>Here I am following Stalnaker (1996: §4) who suggests that Lewis' Humean supervenience assumption should not be viewed as a contingent thesis. If one did, one would have to restrict the scope of the universal quantifier to some proper subset of  $W$  (including the actual factual world).

This in turn leads to the problem of undermining futures (Hall 1994, Lewis 1994, Thau 1994; for criticism see Hoefer 1997), something to be shied away from. In a nutshell, the problem is that the Principal Principle implies that there are no propositions  $F$  with the following properties: the conjunction of  $F$  and the actual history up to a certain time  $t$ ,  $H_{wt} \cap F = \{v\}$ , yields  $t$ -chances different from what they actually are,  $ch_{vt} \neq ch_{wt}$  or  $H_{wt} \cap F \cap T_w = \emptyset$ , even though  $F$  has some non-zero  $t$ -chance in  $w$  of coming about,  $ch_{wt}(F) > 0$ . For a recent discussion see Briggs (2009).

While certainly attractive, I follow Spohn (2010) in rejecting Humean supervenience with respect to chance. I will also reject Humean supervenience with respect to deterministic alethic modality. The metaphysical picture I have allows there to be several possible universes that agree on the non-modal facts but disagree on the modalities. Factual claims say what is or is not the case. Modal claims, whether alethically (metaphysically) modal or doxastically (epistemically) modal, do not say *what* is the case. They say *how* something else is the case (for someone). They essentially rely on something else, and I take that something else to be factual – at least eventually, if modalities can be iterated or mixed as in the claim that I believe that the plate would be likely to break if it were dropped.

Consequently a possible universe is represented as a triple  $(w, ch, r)$  of a non-modal, factual component  $w$ , an indeterministic alethic modal component  $ch$ , and a deterministic alethic modal component  $r$ . The set of all possible universes  $\Omega$  is represented as a (not necessarily proper) subset of the Cartesian product  $W \times CH \times R$  of the set of factual worlds  $W$ , the set of potential indeterministic alethic modalities or chance measures  $CH$ , and the set of potential deterministic alethic modalities or counterfactual distances  $R$ . Potential chances  $ch$  are defined on an algebra  $\mathcal{A}_W$  of purely factual propositions  $A \subseteq W$ . Potential counterfactual distances  $r$  are defined on an algebra  $\mathcal{A}_U$  of propositions  $B \subseteq U \subseteq W \times CH$ . Doxastic modalities such as credences and beliefs are defined on an algebra  $\mathcal{A}_\Omega$  of propositions  $X \subseteq \Omega \subseteq W \times CH \times R$ .

If chances do not supervene on facts the chance that a given uranium atom decays within a certain interval can be different in two possible universes, even though those two universes agree on all matters of particular fact, including the (limiting) relative frequencies of decay.

If counterfactual distances do not supervene on facts it may be true in one universe  $u$  that Bizet would be Italian if Bizet and Verdi were compatriots. It may be true in another universe  $u'$  that Verdi would be French if Bizet and Verdi were compatriots. And this may be so, even though these two universes  $u$  and  $u'$  agree on all non-modal matters of fact such as whether Bizet is French and

whether Verdi is Italian and when they were born where. The difference between  $u = (w_u, ch_u, r_u)$  and  $u' = (w_{u'}, ch_{u'}, r_{u'})$  may only concern the modal components  $r_u$  and  $r_{u'}$ : among the worlds where Bizet and Verdi are compatriots,  $r_u$  may rank those closer to  $w_u = w_{u'}$  where Bizet is Italian, while  $r_{u'}$  may rank those closer to  $w_{u'} = w_u$  where Verdi is French. However, this difference in counterfactual distance will not be reflected in any difference between  $w_u$  and  $w_{u'}$ . Ann and Bob inhabit the same actual world  $w$ , but the way Ann conceives of alternative worlds may be different from the way Bob conceives of alternative worlds. What are important differences and far-fetched alternatives for Ann, may be differences and distances Bob does not even notice. The counterfactuals that are true in Ann's universe  $(w, ch, r_A)$  will be different from the counterfactuals that are true in Bob's universe  $(w, ch, r_B)$ . Only one of Ann and Bob can be right about the actual differences and distances of other worlds from  $w$ , but both their views may be possible, and I can suspend judgment about, or distribute my mass of credence equally among, those two possible views or universes.

We can, of course, also combine deterministic and indeterministic alethic modalities. It may be true in one universe that the plate would be likely to break, if it were dropped, but false in another one, even though the plate is dropped and breaks in both of these two universes, and they also agree on all other local matters of particular fact. And it may be necessarily true in one universe, and possibly false in another, that the chance of past events is zero, even though these two universes agree on all local matters of particular fact.

On Lewis' view chance or indeterministic alethic modality is a primitive. All we know about it is encapsulated in the Principal Principle. Traditional metaphysics has paid more attention to deterministic alethic modalities. Kripke (1959) and others clarified the qualitative notions of necessity and possibility, and then Stalnaker (1968) pointed out that they are intimately related to counterfactuals. Similarly, traditional epistemology has focused on the qualitative notion of belief rather than the quantitative notion of credence. Hintikka (1961) clarified the logic of unconditional belief, and Alchourrón & Gärdenfors & Makinson (1985) started to realize that there is more to the doxastic state than can be expressed by unconditional beliefs. However, as observed by Spohn (1988), AGM belief revision theory is a theory of belief revision by name only, and Hild & Spohn (2008) show why one has to go all the way from AGM's entrenchment orderings that characterize one-step revisions to Spohn's ranking functions, which characterize indefinitely iterable revisions.

The thesis I submit is that deterministic alethic modality constrains qualitative belief in much the same way that chance constrains credence. However, this

is something that cannot even be adequately formulated within the confines of Hintikka-style epistemic logic or AGM belief revision theory. The reason is that both accounts lack a notion of conditional belief. In contrast to this ranking functions are essentially conditional (and essentially numerical).

Before turning to ranking functions let me address a worry about the rejection of Humean supervenience with respect to deterministic alethic modalities. Is it not an utter mystery how we could possibly gain knowledge of non-trivial deterministic alethic modalities if the latter do not supervene on the facts? And consequently, is it not utterly mysterious how we can gain non-trivial philosophical knowledge? The worry arises because all our evidence is restricted to facts. This is not the place to explain how such knowledge is possible, but rather to point out that we do not find a similar mystery worrisome.

First, the mystery would remain even if Humean supervenience with respect to deterministic alethic modalities were to hold. For while it is true that all our evidence is restricted to facts, it is also true that all our evidence is restricted to a finite initial segment of the facts. In fact, all our evidence is at best a very incomplete finite initial segment of the facts. Second, consider the widespread view of scientific realism which holds that science aims at truth rather than mere empirical or observational adequacy. This view is interesting because it is uncontroversial that there are some (theoretical) facts that are in principle unobservable in the widest possible sense, and so will never be part of anyone's evidence. (This claim is agreed upon, even though some philosophers think there is no clearcut distinction between theory and observation.) Scientific realism may be challenged because it seems mysterious how we could gain knowledge of the unobservable if all our evidence is restricted to the observable. However, no anti-realist would claim that the unobservable supervened on the observable because otherwise it was unclear how we could gain knowledge of the unobservable. Similarly, the epistemological mystery of the possibility of philosophical and other modal knowledge should not be taken to support the metaphysical thesis that the modalities in question supervene on non-modal facts.

### 3 Ranking Functions

Ranking functions have been introduced by Spohn (1988) to represent qualitative conditional belief. They are discussed at booklength in Spohn (2012). The theory is numerical in the sense that ranking functions assign numbers to propositions, but these numbers are really only needed for the definition of conditional ranking functions representing conditional beliefs. Once these are defined everything can be read in purely qualitative, though conditional terms.

Here is the definition. Consider a space of possibilities  $P$  and an algebra of propositions  $\mathcal{A}_P$  over  $P$ . A function  $\varrho : \mathcal{A}_P \rightarrow \mathbb{N} \cup \{\infty\}$  is a *finitely / countably / completely minimitive ranking function* on  $\mathcal{A}_P$  just in case for all finite / countable / arbitrary sets of propositions  $\mathcal{B} \subseteq \mathcal{A}_P$ :

$$\varrho(P) = 0 \tag{1}$$

$$\varrho(\emptyset) = \infty \tag{2}$$

$$\varrho\left(\bigcup \mathcal{B}\right) = \min \{\varrho(B) : B \in \mathcal{B}\} \tag{3}$$

For a non-empty or consistent proposition  $A \neq \emptyset$  from  $\mathcal{A}_P$  the conditional ranking function  $\varrho(\cdot | A) : \mathcal{A}_P \setminus \{\emptyset\} \rightarrow \mathbb{N} \cup \{\infty\}$  based on the unconditional ranking function  $\varrho(\cdot) : \mathcal{A}_P \rightarrow \mathbb{N} \cup \{\infty\}$  is defined as

$$\varrho(\cdot | A) = \begin{cases} \varrho(\cdot \cap A) - \varrho(A), & \text{if } \varrho(A) < \infty, \\ 0, & \text{if } \varrho(A) = \infty. \end{cases}$$

Stipulating  $\varrho(\emptyset | A) = \infty$  ensures that every conditional ranking function is a ranking function on  $\mathcal{A}_P$ . A ranking function  $\varrho$  is *regular* if and only if  $\varrho(A) < \varrho(\emptyset)$  for all non-empty or consistent propositions  $A$  from  $\mathcal{A}_P$ . (As a terminological aside: a set of subsets of  $P$ ,  $\mathcal{A}_P \subseteq \wp(P)$ , is an algebra over  $P$  if and only if for all  $A, B$  from  $\mathcal{A}_P$ :  $\emptyset \in \mathcal{A}$ ,  $\overline{A} \in \mathcal{A}$ , and  $A \cup B \in \mathcal{A}$ . An algebra  $\mathcal{A}_P$  over  $P$  is a  $\sigma$ - / complete algebra over  $P$  if and only if for all countable / arbitrary  $\mathcal{B} \subseteq \mathcal{A}_P$ :  $\bigcup \mathcal{B} \in \mathcal{A}_P$ .)

Ranks are interpreted doxastically as grades of disbelief. A proposition  $A$  is disbelieved just in case it is assigned a positive rank  $A$ ,  $\varrho(A) > 0$ . A proposition  $A$  is believed just in case its negation  $\overline{A}$  is disbelieved,  $\varrho(\overline{A}) > 0$ . A proposition  $A$  is disbelieved conditional on a proposition  $C$  just in case  $A$  is assigned a positive rank conditional on  $C$ ,  $\varrho(A | C) > 0$ . A proposition  $A$  is believed conditional on a proposition  $C$  just in case the negation of  $A$ ,  $\overline{A}$ , is disbelieved conditional on  $C$ ,  $\varrho(\overline{A} | C) > 0$ . It takes getting used to read positive numbers in this “negative”

way, but mathematically this is the simplest formulation and so I will stick to it. Note that it follows from the definition of conditional ranks that an ideal doxastic agent should not disbelieve a proposition conditional on itself,  $\varrho(A | A) = 0$ .

In doxastic terms the first axiom says that an ideal doxastic agent should not disbelieve the tautological proposition  $W$ . The second axiom says that she should disbelieve the empty or contradictory proposition  $\emptyset$  with maximal strength  $\infty$ . Given the definition of conditional ranks, the second axiom can be read in purely qualitative, though conditional terms: it says that she should disbelieve the empty or contradictory proposition conditional on any proposition (with a finite rank). It follows that one should believe the tautological proposition with maximal strength or conditional on any proposition (with a finite rank).

Part of what the third axiom says is that an ideal doxastic agent should disbelieve a disjunction  $\bigcup \mathcal{B}$  just in case she disbelieves all its disjuncts  $B \in \mathcal{B}$ . Given the definition of conditional ranks, the third axiom extends this requirement to conditional beliefs. As just noted we may assume that the ideal doxastic agent should not disbelieve a proposition conditional on itself. Given this the third axiom says – in purely qualitative, though conditional terms – the following: she should conditionally disbelieve a disjunction  $\bigcup \mathcal{B}$  just in case she conditionally disbelieves all its disjuncts  $B \in \mathcal{B}$ .

Doxastically interpreted axioms (1)-(3) are synchronic norms for organizing an ideal doxastic agent's conditional beliefs at a given moment in time. They are supplemented by diachronic norms for updating her beliefs over time if new information of various formats is received. The first update rule is defined for the case where the new information comes in form a certainty and mirrors the update rule of strict conditionalization from probability theory.

**Update Rule 1 (Plain Conditionalization, Spohn 1988)** *If  $\varrho(\cdot) : \mathcal{A}_P \rightarrow \mathbb{N} \cup \{\infty\}$  is the ideal doxastic agent's ranking function at time  $t$ , and between  $t$  and  $t'$  she becomes certain of  $E \in \mathcal{A}_P$  and no logically stronger proposition (in the sense that  $E$  is the logically strongest proposition whose negation she assigns  $\infty$  as new rank at  $t'$ ), then her ranking function at time  $t'$  should be  $\varrho_E(\cdot) : \mathcal{A}_P \rightarrow \mathbb{N} \cup \{\infty\}$ ,  $\varrho_E(\cdot) = \varrho(\cdot | E)$ , where for all non-empty or consistent  $B$  in  $\mathcal{A}_P$ ,*

$$\varrho_E(B) = \varrho(B \cap E) - \varrho(E) \quad \text{and} \quad \varrho(\emptyset | E) = \infty.$$

*and, for  $n \in \mathbb{N}$ ,  $\infty - n = \infty$  and  $\infty - \infty = 0$ .*

The second update rule is defined for the case where the new information comes in form of new ranks for the elements of a partition and mirrors the update rule of Jeffrey conditionalization from probability theory (Jeffrey 1983).

**Update Rule 2 (Spohn Conditionalization, Spohn 1988)** *If  $\varrho(\cdot) : \mathcal{A}_P \rightarrow \mathbb{N} \cup \{\infty\}$  is the ideal doxastic agent's ranking function at time  $t$ , and between  $t$  and  $t'$  her ranks on the evidential partition  $\{E_i \in \mathcal{A}_P : i \in I\}$  directly change **to**  $n_i \in \mathbb{N} \cup \{\infty\}$  with  $\min\{n_i : i \in I\} = 0$  ( $n_i = \infty$  for  $E_i = \emptyset$  and  $n_i = 0$  for  $E_i = P$ ), and her finite ranks do not directly change on any finer partition, then her ranking function at time  $t'$  should be  $\varrho_{E_i \rightarrow n_i}(\cdot) : \mathcal{A}_P \rightarrow \mathbb{N} \cup \{\infty\}$ ,*

$$\varrho_{E_i \rightarrow n_i}(\cdot) = \min_{i \in I} \{\varrho(\cdot | E_i) + n_i\}.$$

The third update rule is defined for the case where the new information reports the differences between the old and the new ranks for the elements of a partition and mirrors the update rule of Field conditionalization from probability theory (Field 1978).

**Update Rule 3 (Shenoy Conditionalization, Shenoy 1991)** *If  $\varrho(\cdot) : \mathcal{A}_P \rightarrow \mathbb{N} \cup \{\infty\}$  is the ideal doxastic agent's ranking function at time  $t$ , and between  $t$  and  $t'$  her ranks on the evidential partition  $\{E_i \in \mathcal{A}_P : i \in I\}$  directly change **by**  $z_i \in \mathbb{N}$ , where  $\min\{z_i : i \in I\} = 0$ , and her finite ranks do not directly change on any finer partition, then her ranking function at time  $t'$  should be  $\varrho_{E_i \uparrow z_i}(\cdot) : \mathcal{A}_P \rightarrow \mathbb{N} \cup \{\infty\}$ ,*

$$\varrho_{E_i \uparrow z_i}(\cdot) = \min_{i \in I} \{\varrho(\cdot \cap E_i) + z_i - m\}, \quad m = \min_{i \in I} \{z_i + \varrho(E_i)\}.$$

Spohn conditionalizing the evidential propositions  $E$  and  $\bar{E}$  to 0 and  $n$ , respectively, keeps the relative positions of all possible worlds in  $E$  and all possible worlds in  $\bar{E}$  fixed. It improves the rank of  $E$  to 0 and changes the rank of  $\bar{E}$  to  $n$ . Shenoy conditionalizing  $E$  and  $\bar{E}$  by 0 and  $n$ , respectively, improves the possibilities within  $E$  by  $n$ , as compared to the possibilities in  $\bar{E}$ .  $m$  merely is a normalization parameter that ensures that at least one possible world is assigned rank zero so that the result is a ranking function.

In both cases the new information consists of a (partition of) proposition(s) together with a (list of) number(s). This reflects the fact that the quality of new information varies with the reliability of its source: it makes a difference if the weather forecast predicts that it will rain, if a friend the ideal doxastic agent trusts tells her so, or if she sees herself that it is raining. In each case the proposition she learns is that it is raining, but the effect of the new information on her old beliefs may be a different one in each case. The difference in the reliability of the sources of the new information is reflected in the number(s) accompanying the proposition(s). The effect the new information has on her old beliefs depends on those numbers.

Why should an ideal doxastic agent organize her conditional beliefs at a given moment in time according to axioms (1-3)? Why should an ideal doxastic agent update her beliefs over time according to update rules (1-3) if she receives evidence of the appropriate format? To answer these questions we need a little bit of terminology.

An ideal doxastic agent's *degree of entrenchment* for a proposition  $A$  is defined as the number of independent and minimally positively reliable information sources saying  $A$  that it takes for her to give up her disbelief in  $A$ . Independent and minimally positively reliable information sources for a proposition are idealized entities that are used to define an ideal doxastic agent's degrees of entrenchment. Information sources as we know them usually are neither independent nor minimally positively reliable.

The ideal doxastic agent's grades of disbelief, her ranks, are theoretical entities. Their relation to her idealized, but otherwise observable degrees of entrenchment is a delicate one: minimally degrees of entrenchment are used to measure grades of disbelief; maximally the latter can be defined in terms of the former.<sup>6</sup>

Let  $\rho$  be the ideal doxastic agent's entrenchment function, i.e. the function that summarizes her degrees of entrenchment for all the propositions under consideration. Then her *belief set*  $\mathcal{B}_\rho$  is defined as the set of propositions whose negations have a positive degree of entrenchment:

$$\mathcal{B}_\rho = \{A \in \mathcal{A}_P : \rho(\bar{A}) > 0\}$$

An ideal doxastic agent's belief set  $\mathcal{B}_\rho \subseteq \mathcal{A}_P$  is *consistent in the finite / countable / complete sense* if and only if for every finite / countable / arbitrary set of propositions  $\mathcal{B} \subseteq \mathcal{B}_\rho$ :  $\bigcap \mathcal{B} \neq \emptyset$ .

An ideal doxastic agent's belief set  $\mathcal{B} \subseteq \mathcal{A}_P$  is *deductively closed in the finite / countable / complete sense* if and only if for every finite / countable / arbitrary set of propositions  $\mathcal{B} \subseteq \mathcal{B}_\rho$  and all propositions  $A \in \mathcal{A}_P$ : if  $\bigcap \mathcal{B} \subseteq A$ , then  $A \in \mathcal{B}_\rho$ .

Now we can answer the question why an ideal doxastic agent should organize her conditional beliefs at a given moment in time according to axioms (1-3), and why she should update her beliefs over time according to update rules (1-3) if she receives evidence of the appropriate format. She should do so, because

---

<sup>6</sup>The situation is somewhat similar in Bayesianism, where *betting ratios* or *fair betting ratios* minimally are used to measure degrees of belief, and maximally are used to define them (Eriksson & Hájek 2007). Whatever the exact relation between (fair) betting ratios and degrees of belief, the former are the central notion in the best known argument in favor of Bayesianism, the Dutch Book Argument.

**Theorem 1** *An ideal doxastic agent's belief set  $\mathcal{B}_\rho$  based on her entrenchment function  $\rho$  is and will always be consistent and deductively closed in the finite / countable / complete sense, possibly conditional on some element of the evidential partition, if and only if  $\rho$  is a finitely / countably / completely minimitive ranking function and, depending on the format of the evidence, the ideal doxastic agent updates according to update rule 1, 2, or 3.*

For proof and discussion the reader is referred to Huber (2007).

As indicated in the introduction, this theorem can be used to establish the thesis that an ideal doxastic agent's beliefs should obey the synchronic and diachronic rules of the ranking calculus. It can be used to provide a means-ends justification for this thesis in the spirit of epistemic consequentialism (Percival 2002, Stalnaker 2002). The idea is that obeying the normative constraints of the ranking calculus is a (necessary and sufficient) means to attaining the end of being "eternally consistent and deductively closed." The latter end in turn is a (necessary, but insufficient) means to attaining the end of always having only true beliefs, and as many as possible thereof. Brössel & Eder & Huber (2013) discuss the importance of this result as well as its Bayesian role-model, Joyce's (1998; 2009) "non-pragmatic vindication of probabilism", for considering epistemic rationality a form of instrumental rationality and for means-ends epistemology in general.

## 4 The Royal Rule

There is one more thing to note before we can state how deterministic alethic modality constrains qualitative belief. While chances develop over time, counterfactual distances vary across contexts. In the literature (Stalnaker 1999: ch. 1, Gillies 2009) contexts are represented as relations between (centered) worlds or as functions from (centered) worlds to sets of (centered) worlds. Since I have a slightly more complicated picture I represent contexts as functions from universes to sets of factual worlds. This means that what can be presupposed is restricted to the realm of the non-modal.

Let a context be a function from universes to sets of factual worlds,  $c : \Omega \rightarrow \wp(W) \times \{R\}$ , where  $\Omega \subseteq W \times R$  is the set of all possible universes  $u = (w_u, r_u)$ . (Chances are ignored in order to avoid unnecessary complications.)  $c$  is *alethically respectable* just in case for all  $u$  in  $\Omega$ :  $u \in c(u)$ . In contrast to Gillies (2009), Stalnaker (1998) argues against this assumption as a definitional constraint on all contexts. I agree: not every context is alethically respectable. However, those contexts in which an ideal doxastic agent aims at the truth, and in which the Royal Rule is supposed to hold, are. Stalnaker (1998) merely requires all contexts  $c$  to be *deontically respectable* in the sense that for all  $u$  in  $\Omega$ :  $c(u) \neq \emptyset$ . Both Gillies (2009) and Stalnaker (1998) assume further that all contexts  $c$  are *doxastically respectable* in the sense that for all  $u$  and  $u'$  in  $\Omega$ :  $c(u) \subseteq c(u')$  if  $u' \in c(u)$ .<sup>7</sup> I do not make this assumption.

Now we can finally state the

**Royal Rule 1** *Let  $Ra : \mathcal{A}_\Omega \rightarrow \mathbb{N} \cup \{\infty\}$  be an ideal doxastic agent's initial grading of disbelief, and hence a regular ranking function on a complete algebra  $\mathcal{A}_\Omega$  over  $\Omega$ , where  $\Omega \subseteq W \times R$  is the set of all possible universes  $u = (w_u, r_u)$ . Let  $c : \Omega \rightarrow \wp(W) \times \{R\}$  be an alethically (and hence deontically) respectable context. Let  $n$  be a number from  $\mathbb{N} \cup \{\infty\}$ . For factual proposition  $A \subseteq W$ , let  $r_c(A) = n$  be the proposition that the counterfactual distance in context  $c$  to the closest  $A$ -worlds equals  $n$ . Finally, let  $E_c \in \mathcal{A}_\Omega$  be an arbitrary proposition that is admissible in  $c$ . Then*

$$Ra(A \times R \mid r_c(A) = n \cap E_c) = n.$$

<sup>7</sup>Strictly speaking Stalnaker (1998) is representing contexts by a Kripke semantics with an accessibility relation between worlds that is serial as well as transitive and Euclidian, but not necessarily reflexive. These requirements translate into deontic (serial), doxastic (transitive plus Euclidian), and alethic (reflexive) respectability.

The Royal Rule for ranks says that an ideal doxastic agent's initial strength of disbelief in a factual proposition  $A$ , given that the counterfactual distance to the closest  $A$ -worlds in context  $c$  equals  $n$ , and no further information that is not admissible in  $c$ , should equal  $n$ . It is the rank-theoretic counterpart of the Principal Principle. The latter principle says that, in the absence of further information (or, if only admissible information is present), objective chances should guide an ideal doxastic agent's subjective credences. More generally, the idea is that, in the absence of further information (or, if only admissible information is present), the objective alethic modalities should guide an ideal doxastic agent's subjective doxastic modalities.

This idea underlies both the Principle Principle as well as the Royal Rule. In the former case we focus on the alethic modality of objective chances, and on the doxastic modality of subjective credences. In the latter case we focus on the alethic modality of counterfactuality, and on the doxastic modality of conditional belief. In these terms the Royal Rule says that, in the absence of further information (or, if only admissible information is present), objective counterfactuals should guide the ideal doxastic agent's subjective conditional beliefs – and, more generally, in the absence of further information (or, if only admissible information is present), objective counterfactual distances should guide the ideal doxastic agent's subjective conditional grades of disbelief.

For purposes of illustration suppose you are certain of, or conditionalize on, the counterfactual that the plate would break if it were dropped, and that this is all you are certain of, or conditionalize on. The Royal Rule requires you to hold the conditional belief that the plate will break given that it is dropped. More generally, suppose an ideal doxastic is certain of, or conditionalizes on, the counterfactual that  $C$  would be the case if  $A$  were the case, and suspends judgment with respect to the antecedent  $A$ . Furthermore, suppose this is all the ideal doxastic agent is certain of, or conditionalizes on. Given these assumptions, the Royal Rule requires the ideal doxastic agent to hold the conditional belief in  $C$  given  $A$ .

If you have been trained in terms of Kripke-style accessibility relations, think of the counterfactual distances as stairways that take you from the factual world you are located at to other factual worlds. The counterfactual distance of a world can be thought of as the number of steps on the stairway to this world. The counterfactual distance of a proposition can be thought of as the number of steps on the shortest stairway into this proposition. If you rather prefer Lewis style similarity orderings, make the limit assumption so you can start counting, allow for multiple copies of spheres with the copies being pictured as larger circles without any additional dots (worlds) in them, and think of the counterfactual distance of

a world as the number of the first sphere in this ordering that contains the world. The counterfactual distance of a proposition can be thought of as the number of the first sphere in this ordering which properly overlaps with the proposition.

Suppose an ideal doxastic agent is certain that Bizet was French and that Verdi was Italian. For such an ideal doxastic agent the Royal Rule does not apply to the counterfactual that Bizet would have been Italian if Bizet and Verdi had been compatriots. The proposition that Bizet was French and Verdi was Italian contains too much information and so overrules the counterfactual: it is not admissible. Just as chances should guide one's credences only if one is not certain about too much else besides, counterfactual distances should guide an ideal doxastic agent's conditional beliefs only if she does not conditionalize on too much else besides. If the ideal doxastic agent is certain of, or conditionalizes on, the proposition that the coin lands heads, we can tell her whatever we want about how likely it is that the coin lands tails: it will not, and it should not, affect her credence in the proposition that the coin lands heads. If the ideal doxastic agent is certain of, or conditionalizes on, the fact that Bizet was French, no counterfactual will change her belief that he was. It is only when she has little or no factual information that she should let the objective modalities guide her doxastic modalities.

The clinically clean case in which an ideal doxastic agent should obey the Royal Rule is when she suspends judgment about all factual matters before she conditionalizes on any counterfactual. This clinically clean case is, of course, very different from our actual doxastic situations, and so it is difficult for us to evaluate the intuitive plausibility of the Royal Rule in this case. Most of us cannot foresee the future, though, and so forward looking counterfactuals may be a helpful approximation. If I am certain that my friends would bring along their children if the weather was nice tomorrow, and that is all I believe as far as their possible visit is concerned, then I should hold the conditional belief that my friends will bring along their children given that the weather is nice tomorrow. As will become clear below, for the purposes of this paper it is sufficient that the Royal Rule holds, and that counterfactual distances guide an ideal doxastic agent's conditional grades of disbelief, in this one single clinically clean case where she is suspends judgment with respect to all factual propositions before she conditionalizes on any counterfactual.<sup>8</sup>

---

<sup>8</sup>In this connection I may perhaps repeat that I do not intend the Royal Rule to be accepted on intuitive grounds and because it seems plausible, even in this one clinically clean case. As stressed in the introduction, like every normative principle the Royal Rule has to be justified by proving it to be the means to some end. Such a means-end relationship between the Royal Rule and some pertinent epistemological goal still has to be established.

Chances develop over time, and Lewis takes the complete history of world  $w$  up to a certain point of time  $t$ ,  $H_{wt}$ , to be admissible at  $t$ . Counterfactual distances vary across contexts instead, and so it is tempting to assume accordingly that the (intersection of all) presupposition(s) of universe  $u$  in context  $c$ ,  $c(u)$ , is admissible in  $c$ . However, when giving in to the temptation it is important to note that  $c(u)$  is the logically strongest proposition that is necessarily true at  $u$  in  $c$ . In this sense  $c(u)$  is already contained in, or implied by, any counterfactual that is true at  $u$  in  $c$ .  $c(u)$  merely delimits the set of factual worlds on which counterfactual distances are defined. The assumption that  $c(u)$  is admissible in  $c$  is not needed for the argument that counterfactual distances have the structure of ranking functions; it merely defines their domain.

The second assumption Lewis makes is that the theory of indeterministic alethic modality of world  $w$ ,  $w$ 's theory of chance,  $T_w$ , is admissible at any time. For Lewis  $T_w$  is the conjunction of all true history-to-chance conditionals of the form:  $H_t \rightarrow ch_t(A) = x$ , where  $H_t$  completely, but not necessarily truly, specifies the history of  $w$  up to  $t$  and, for true  $H_t$ ,  $H_t \rightarrow ch_t(A) = x$  is true at  $w$  just in case  $ch_t(A) = x$  is (otherwise  $\rightarrow$  is stronger than the material conditional  $\supset$ ).

On this view historical information<sup>9</sup> has to be admissible in order for information about chances to be admissible, and hence in order for the Principal Principle to entail that chances are probabilities. As explained in section 2, on my account it does not make sense to speak of a factual world's theory of indeterministic or deterministic alethic modality. It makes only sense to speak of a universe's theory of indeterministic or deterministic alethic modality. The theory of chance of universe  $u = (w_u, ch_u, r_u)$  is simply  $ch_u$ , and the theory of counterfactual distance of  $u$  is simply  $r_u$ . We can also formulate these theories as the following purely modal propositions  $T_u$  and  $D_u$ :<sup>10</sup>

$$T_u = \{(w, ch, r) \in W \times CH \times R : w \in W, ch = ch_u, r \in R\} = W \times \{ch_u\} \times R$$

$$D_u = \{(w, ch, r) \in W \times CH \times R : w \in W, ch \in CH, r = r_u\} = W \times CH \times \{r_u\}$$

However, it is important to note that  $T_u$  is not a factual proposition over  $W$ , and so is not in the domain of  $ch_u$ , as Hall (1994) has it (cf. Hoefer 1997). Consequently chances cannot be iterated and we avoid the problem of undermining

<sup>9</sup>As an aside, note that history up to an arbitrary time  $t$ , whether complete or not, is an alethically respectable context.

<sup>10</sup>As another side, note that Lewis' assumptions about admissibility, when translated into my account, entail but are not entailed by the assumption that  $T_u$  is admissible.

futures. Equally important,  $D_u$  is not a factual proposition over  $W$  and so is not in the domain of  $r_u$ .

The presupposition of universe  $u$  in context  $c$ ,  $c(u)$ , and  $u$ 's theory of counterfactual distance,  $D_u$ , combine to the conjunction

$$c(u) \cap D_u = \{(w, r) \in \Omega : \exists r' \in R : (w, r') \in c(u), r = r_u\} = f(c(u)) \times \{r_u\},$$

where  $f(B) = \{w \in W : \exists r \in R : (w, r) \in B\}$  is the *factual component* of  $B \subseteq \Omega$ .

In alethically respectable contexts this conjunction is never empty, because  $(w_u, r_u) \in c(u) \cap D_u$ . Since  $Ra$  is regular – which, in contrast to the probabilistic case, is always possible – this implies that  $Ra(c(u) \cap D_u) < \infty$ . Hence  $Ra(A \mid c(u) \cap D_u)$  is always well-defined, even if conditional ranking functions are defined in a way different from section 3 (Huber 2006: 464).

We have finally reached the heart of this paper and can adopt an argument from Lewis (1980: 276ff), who uses the Principal Principle to show that objective chances are probabilities.

Assume the presupposition of universe  $u$  in context  $c$ ,  $c(u)$ , as well as  $u$ 's theory of deterministic alethic modality or counterfactual distance,  $D_u$ , to be admissible in  $c$ . The Royal Rule yields, for every factual proposition  $A$  from  $\mathcal{A}_W$ ,

$$Ra(A \times R \mid r_c(A) = n \cap c(u) \cap D_u) = n,$$

where  $r_c(A) = n$  is the purely modal proposition that the counterfactual distance in context  $c$  to the closest  $A$ -worlds equals  $n$ :

$$\begin{aligned} r_c(A) = n &= \{(w, r) : w \in W, r(A) = n\} \\ &= W \times \{r \in R : r(A) = n\} \end{aligned}$$

$c(u)$  and  $D_u$  entail, for every factual proposition  $A$ , that  $r_c(A)$  is what it is in universe  $u$  in context  $c$ . Hence we can “reformulate” the Royal Rule as follows, where  $r_{uc}(A)$  is the counterfactual distance of the factual proposition  $A$  in universe  $u$  in context  $c$ :

**Royal Rule 2** *Let  $Ra : \mathcal{A}_\Omega \rightarrow N \cup \{\infty\}$  be an ideal doxastic agent's initial grading of disbelief, and hence a regular ranking function on a complete algebra  $\mathcal{A}_\Omega$  over  $\Omega$ , where  $\Omega \subseteq W \times R$  is the set of all possible universes  $u = (w_u, r_u)$ . Let  $c : \Omega \rightarrow \wp(W) \times \{R\}$  be an alethically (and hence deontically) respectable context. Then we have for every factual proposition  $A$  from  $\mathcal{A}_W$*

$$r_{uc}(A) = Ra(A \times R \mid c(u) \cap D_u).$$

In words: the counterfactual distance distribution of universe  $u$  in context  $c$ ,  $r_{uc}$ , comes from an ideal doxastic agent's initial grading of disbelief  $Ra$  by conditionalization on the presupposition of  $u$  in  $c$ ,  $c(u)$ , and  $u$ 's theory of deterministic alethic modality,  $D_u$ .

As we have seen in section 3 an ideal doxastic agent's grading of disbelief  $Ra$  should be a ranking function. Since, as we have also seen in section 3, everything that comes from a ranking function by conditionalization is itself a ranking function, it follows that the counterfactual distance distribution of a universe in a given context has the structure of a ranking function.

Note that one single case in which the Royal Rule holds suffices for this argument to work. The clinically clean case in which the ideal doxastic agent suspends judgment with respect to all factual matters before conditionalizing on any counterfactual, and in which she obeys the Royal Rule and lets counterfactual distances guide her conditional grades of disbelief, suffices for a principled account of the semantics of counterfactuals.

## 5 Consequences

For factual propositions  $A, C \subseteq W$  the proposition  $A \square \rightarrow C$  is true in universe  $u$  in context  $c$  – alternatively: at world  $w_u$  relative to ranking function  $r_u$  in context  $c$  – just in case every  $r_{uc}$ -minimal  $A$ -world is a  $C$ -world. As shown in the appendix, the logic that results from this semantics is the basic conditional logic **V**. **V** is the familiar Stalnaker-Lewis logic of counterfactuals, but without their controversial centering axioms discussed below, and without Stalnaker's controversial law of conditional excluded middle discussed in the introduction. In this concluding section I want to briefly discuss some consequences thereof.

Fix context  $c$  and let  $A^{r_{uc}}$  be the set of  $r_{uc}$ -minimal  $A$ -worlds,

$$\begin{aligned} A^{r_{uc}} &= \{w \in A : \forall B \in \mathcal{A}_W : \text{if } w \in B, \text{ then } r_{uc}(B) \leq r_{uc}(A)\} \\ &= \{w \in A : r_{uc}(\{w\}) \leq r_{uc}(A)\}, \end{aligned}$$

where the last equation presupposes that  $\{w\} \in \mathcal{A}_W$  for all  $w \in W$ , which is the case if  $\mathcal{A}_W$  is the powerset of  $W$ ,  $\wp(W)$ . Then, in context  $c$ ,  $A \square \rightarrow C$  is the following proposition:  $\{u \in \Omega : A^{r_{uc}} \subseteq C\}$ .

Therefore, for factual  $A$  and  $C$ ,  $A \square \rightarrow C$  is a proposition in  $\mathcal{A}_U$  (not in  $\mathcal{A}_W$ ), but it does not make sense to iterate the counterfactual conditional operator  $\square \rightarrow$ . (Counterfactual conditionals can be embedded by Boolean connectives, though.)

The reason is that, without Humean supervenience of the counterfactual distances  $r$  on the factual worlds  $w$ , this would require there to be second-order ranking functions  $r_2$  defined on some algebra  $\mathcal{A}_Y$  over the set of all possible pairs  $(w, r)$  of factual worlds  $w$  and first-order ranking functions  $r$ ,  $Y \subseteq W \times R$  (the first-order ranking functions are defined on some algebra  $\mathcal{A}_W$  over the set of all factual worlds  $W$ ). For two iterations of the counterfactual conditional operator we would need third-order ranking functions  $r_3$  defined over some algebra  $\mathcal{A}_Z$  over the set of all possible triples  $(w, r, r_2)$  of factual worlds  $w$ , first-order ranking functions  $r$ , and second-order ranking functions  $r_2$ ,  $Z \subseteq Y \times R_2$ .

Consequently the non-iterability of some conditional operator  $\rightarrow$  does not imply that  $A \rightarrow C$  is no proposition. This blocks the following fairly prominent argument in the literature on (mainly indicative, but in the case of Edgington also counterfactual) conditionals (Bennett 2003: ch. 7 and ch. 16, Edgington 1995; 2008, Gibbard 1981). With reference to natural language it is argued that it is hard if not impossible to iterate (indicative) conditionals. This is then taken as evidence for the thesis that (indicative) conditionals have no truth value. The present account shows that this inference is not valid. The purported non-iterability of the counterfactual conditional can also be explained by the failure of Humean supervenience of counterfactual distances on factual worlds.<sup>11</sup>

Let us briefly turn to contexts. For factual  $A \subseteq W$  define  $\Box A$  to be true in universe  $u$  in context  $c$  – alternatively: at world  $w_u$  relative to ranking function  $r_u$  in context  $c$  – just in case  $\neg A \Box \rightarrow A$  is true in  $u$  in  $c$ .

Suppose  $A$  is a presupposition of  $u$  in  $c$  in the sense that  $c(u) \subseteq A \times R$ . Then, since  $A \times R$  is non-empty (this is so because  $c(u)$  is non-empty),

$$r_{uc}(A) = Ra(A \times R \mid c(u) \cap D_u) = Ra(c(u) \cap D_u) - Ra(c(u) \cap D_u) = 0.$$

Conversely, if  $\neg A$  is a presupposition of  $u$  in  $c$  so that  $c(u) \cap (A \times R) = \emptyset$ , then

$$r_{uc}(A) = Ra(A \times R \mid c(u) \cap D_u) = Ra(\emptyset) - R(c(u) \cap D_u) = \infty,$$

because  $c(u) \cap D_u$  is non-empty and  $Ra$  is regular. Recall that  $c(u)$  is of the form  $F \times R$ , where  $F \subseteq W$ . For the factual component of  $c(u)$ ,  $f(c(u))$ , we thus have:  $r_{uc}(\overline{f(c(u))}) = \infty$ .

---

<sup>11</sup>To be sure, if we have Humean supervenience of deterministic alethic modalities on non-modal facts, a modal proposition such as  $A \Box \rightarrow C$  amounts to a factual proposition, and then  $\Box \rightarrow$  can be iterated indefinitely. In this case the above inference from non-iterability to non-propositionality is valid.

Since  $Ra$  is regular,  $f(c(u))$  is the weakest proposition with this property. That is: if  $F \subset f(c(u))$ , then  $r_{uc}(\overline{F}) = n < \infty$ . This means that  $f(c(u))$  is the set of all factual worlds that are accessible from  $u$  in  $c$  so that  $\Box A$  is true in  $u$  in  $c$  if and only if  $f(c(u)) \subseteq A$ . In other words, all and only the presuppositions of  $u$  in  $c$  are necessarily true in  $u$  in  $c$ .

Furthermore,  $w_u \in f(c(u))$  in alethically respectable contexts  $c$ , which means that, in alethically respectable contexts, what is necessarily true is true,  $\Box A \supset A$ . This, however, is not a consequence of the Royal Rule, but built into the definition of alethically respectable contexts, in which the Royal Rule is supposed to hold. Other assumptions about the contexts in which the Royal Rule is supposed to hold give rise to other modal logics. However, the logic of counterfactuals always remains the same basic conditional logic  $\mathbf{V}$  (contrast this with Williamson 2007: ch. 5 and app. 1).

Another consequence is the limit assumption (Lewis 1973: 19ff), which is necessary and sufficient for counterfactual consistency (Herzberger 1979): for any universe  $u$  and context  $c$  there is at least one world  $w$  in  $W$  such that  $r_{uc}(\{w\}) = 0$ . Call such a world a *center* of  $u$  in  $c$ . While there always is a center in any non-empty context, nothing guarantees that the actual world  $w_u$  is a center of  $u$ , let alone its only center. The latter two conditions are the contents of the principles of weak centering and strong centering, to which we turn now.

Strong centering has long been attacked, because it validates the following principle:  $(A \cap C) \supset (A \Box \rightarrow C)$ . Weak centering validates:  $(A \Box \rightarrow C) \supset (A \supset C)$ , which is Modus Ponens for the counterfactual conditional operator  $\Box \rightarrow$  and seems more reasonable. But arguing on the basis of seeming reasonableness is exactly the kind of intuition-based philosophy I try not to engage in. While we may not be able to do completely without intuitions, we can try to minimize reliance on intuitions.

Nozick (1981) and Iatridou (2000) and Gunderson (2004) and Menzies (2004) and Leitgeb (2012a, 2012b) are examples of an increasing number of philosophers who reject weak centering (as well as strong centering). The Royal Rule does not imply that  $r_{uc}(\{w_u\}) = 0$ , and so I do not subscribe to weak centering either. In fact, weak centering would trivialize the Royal Rule. For suppose we had  $r_{uc}(\{w_u\}) = 0$  for all  $u$  and  $c$ . The Royal Rule implies  $Ra(A \times R \mid r_c(A) = n) = n$ . But if  $r_{uc}(A) = n > 0$  and weak centering holds, then  $w_u \notin A$ , which means that  $A$  is false in  $w_u$  in  $c$  relative to  $r_u$ . Hence  $n = Ra(A \times R \mid r_c(A) = n) = Ra(A \times R \mid r_c(A) = n \cap \overline{A} \times R) = \infty$ , which means the Royal Rule would hold only if  $n = \infty$  or  $n = 0$ .

Strong centering is true, and weak centering is analytically true, on an interpretation of counterfactuals in terms of similarity. I reject this interpretation together with those principles (see also Hájek ms). Weak and strong centering are not plausible on an interpretation of counterfactuals in terms of atypicality or abnormality. As mentioned in the introduction, counterfactual distance figures as a primitive on my account. In particular, I do not want to suggest that counterfactual distance should be interpreted in terms atypicality or abnormality. The counterfactual  $A \Box \rightarrow C$  implies, but is not implied by the claim that: if  $A$ , then normally or typically  $C$ . However, as I show in (Huber ms) atypicality or abnormality<sup>12</sup>, in a precise sense, provides empirical evidence for counterfactual distance, and under certain assumptions allows us to reliably infer the truth values of various counterfactuals. Therefore it may be helpful to think of counterfactuals in terms of typicality when considering why weak and strong centering fail.

Despite my reservations about the reliability of intuitions, it is interesting to speculate about the reason for the intuitive pull of weak and strong centering as well as conditional excluded middle that some philosophers seem to feel. I will conclude by briefly indulging therein. A particularly important subclass of the class of all counterfactuals are those non-backtracking counterfactuals that enter into the analysis of causation. On a structural equations account of causation in terms of so-called causal models those causal counterfactuals satisfy weak and strong centering. If the causal model is acyclic, as is usually assumed, these counterfactuals additionally satisfy conditional excluded middle. The present rank-theoretic semantics for counterfactuals is a semantics for all counterfactuals, backtracking and other non-causal counterfactuals and causal counterfactuals alike. Weak and strong centering and conditional excluded middle do not hold for the class of all counterfactuals. However, these principles do hold for the class of all causal counterfactuals, as shown in Huber (2013). Maybe this can explain the intuitive pull of these principles that some philosophers feel.

Another possible explanation is suggested by Leitgeb (2012a, 2012b). He presents a different, probabilistic semantics for the same basic conditional logic **V** and points out that the following substitutes for weak and strong centering are logically valid.

**Lt 1**  $(A \Box \rightarrow C) \supset (W \Box \rightarrow (A \supset C))$

**Lt 2**  $(W \Box \rightarrow A \cap C) \supset (A \Box \rightarrow C)$

---

<sup>12</sup>Strictly speaking it is not typicality or normality, but rather what statisticians call the *mode(s)* of a sample that provide(s) evidence for counterfactuals. Since talk of the mode of a sample misleadingly suggests a connection to modalities, I have formulated the above in terms of typicality.

Whatever the best explanation for the intuitive pull of these principles, fortunately we do not have to rely on it.

## 6 Appendix: Logical Considerations

Let us first consider system **C2** from Stalnaker (1968). Let  $\mathcal{L}_0$  be the smallest set that contains a given countable set of propositional variables  $PV$  and is closed under the classical connectives  $\neg, \wedge, \vee$ , and  $\supset$ . Let  $\mathcal{L}_1$  be the smallest set containing  $\mathcal{L}_0$  and  $\alpha \Box \rightarrow \beta$  for any two  $\alpha, \beta$  from  $\mathcal{L}_0$ . Let  $\mathcal{L}$  be the smallest set that contains  $\mathcal{L}_1$  and is closed under the classical connectives. Finally, let  $\mathcal{L}^+$  be the smallest set containing  $PV$  (or any of the above mentioned languages) that is closed under the classical connectives  $\neg, \wedge, \vee, \supset$  plus  $\Box \rightarrow$ .

So the language  $\mathcal{L}$  is built up from a countable set of propositional variables in the usual way, with the only exception that  $\alpha \Box \rightarrow \gamma$  is a well-formed formula if and only if  $\alpha$  and  $\gamma$  are well-formed formulae and do not contain an occurrence of  $\Box \rightarrow$ .  $A, B, C, \dots$  are the sets of models in which  $\alpha, \beta, \gamma, \dots$  are true. Ranking functions are defined on some finitary /  $\sigma$ - / complete field over the set of models of  $\mathcal{L}$ ,  $Mod_{\mathcal{L}}$ .

The rules of inference of **C2** are Modus Ponens and Necessitation, both of which preserve validity.

$$\mathbf{RI\ 1} \quad \vdash \alpha, \quad \vdash \alpha \supset \beta \quad \Rightarrow \quad \vdash \beta$$

$$\mathbf{RI\ 2} \quad \vdash \alpha \quad \Rightarrow \quad \vdash \Box \alpha$$

The axiom schemata of **C2** are **St 1** to **St 7**, all of which are valid except for Conditional Excluded Middle **St 5** and Weak Centering **St 6**.

**St 1** *All tautologies are axioms.*

$$\mathbf{St\ 2} \quad \Box (\alpha \supset \gamma) \supset (\Box \alpha \supset \Box \gamma)$$

$$\mathbf{St\ 3} \quad \Box (\alpha \supset \gamma) \supset (\alpha \Box \rightarrow \gamma)$$

**St 1** is trivially validated, because classical logic is presupposed. As to **St 2** and **St 3**, suppose  $\Box (\alpha \supset \gamma)$  is true at  $w_u$  in  $c$  relative to  $r_u$ . This means that  $\neg (\alpha \supset \gamma) \Box \rightarrow (\alpha \supset \gamma)$  is true at  $w_u$  in  $c$  relative to  $r_u$ , i.e.  $(A \cap \bar{C})^{r_{uc}} \subseteq \bar{A} \cup C$ . Since  $(A \cap \bar{C})^{r_{uc}} \subseteq A \cap \bar{C}$ , this can only happen if  $\bar{A} \cup C = Mod_{\mathcal{L}}$ , i.e.  $A \subseteq C$ . For **St 2** suppose that  $\Box \alpha$  is true at  $w_u$  in  $c$  relative to  $r_u$ . This means  $\bar{A}^{r_{uc}} \subseteq A$  and since  $\bar{A}^{r_{uc}} \subseteq \bar{A}$ , this can only happen if  $A = Mod_{\mathcal{L}}$ . Consequently  $C = Mod_{\mathcal{L}}$  and hence  $\bar{C}^{r_{uc}} \subseteq \bar{C} = \emptyset$  and hence  $\bar{C}^{r_{uc}} \subseteq C$ , which means that  $\Box \gamma$  is true at  $w_u$  in  $c$  relative to  $r_u$ . For **St 3** we have  $A^{r_{uc}} \subseteq A \subseteq C$ , which means that  $\alpha \Box \rightarrow \gamma$  is true at  $w_u$  in  $c$  relative to  $r_u$ .

**St 4**  $\diamond\alpha \supset ((\alpha \Box\rightarrow \beta) \supset \neg(\alpha \Box\rightarrow \neg\beta))$

Suppose  $\diamond\alpha$  and  $\alpha \Box\rightarrow \beta$  are true at  $w_u$  in  $c$  relative to  $r_u$ , i.e.  $A^{r_{uc}} \not\subseteq \overline{A}$  and  $A^{r_{uc}} \subseteq B$ . It follows that  $\emptyset \neq A^{r_{uc}} \cap A \subseteq A^{r_{uc}} \subseteq B$ , and hence that  $A^{r_{uc}} \not\subseteq \overline{B}$ , i.e. that  $\neg(\alpha \Box\rightarrow \neg\beta)$  is true at  $w_u$  in  $c$  relative to  $r_u$ .

**St 5 (Conditional Excluded Middle)**  $(\alpha \Box\rightarrow (\beta \vee \gamma)) \supset (\alpha \Box\rightarrow \beta) \vee (\alpha \Box\rightarrow \gamma)$

Conditional Excluded Middle **St 5** is not valid. Consider the language  $\mathcal{L}$  over the two propositional variables  $p$  and  $q$ .  $Mod_{\mathcal{L}} = \{w_{pq}, w_{p\neg q}, w_{\neg pq}, w_{\neg p\neg q}\}$ . Let  $c(u) = Mod_{\mathcal{L}} \times \{R\}$  for all  $u \in Mod_{\mathcal{L}} \times R$ , and let  $u^* = (w_{u^*}, r_{u^*})$  with  $w_{u^*} = w_{pq}$  and  $r_{u^*c}(\{w\}) = 0$  for all  $w \in Mod_{\mathcal{L}}$ .  $p \Box\rightarrow q \vee \neg q$  is true at  $w_{u^*}$  in  $c$  relative to  $r_{u^*}$ , but both  $p \Box\rightarrow q$  and  $p \Box\rightarrow \neg q$  are false at  $w_{u^*}$  in  $c$  relative to  $r_{u^*}$ . This is so because  $P^{r_{u^*c}} = P$  and  $Q^{r_{u^*c}} = Q$  and  $\overline{P}^{r_{u^*c}} = \overline{P}$  and  $\overline{Q}^{r_{u^*c}} = \overline{Q}$ , and hence  $P^{r_{u^*c}} \subseteq Q \cup \overline{Q} = Mod_{\mathcal{L}}$  while  $P^{r_{u^*c}} \not\subseteq Q$  and  $P^{r_{u^*c}} \not\subseteq \overline{Q}$ .

**St 6 (Weak Centering)**  $(\alpha \Box\rightarrow \beta) \supset (\alpha \supset \beta)$

Weak Centering **St 6** is not valid. Let  $\mathcal{L}$ ,  $Mod_{\mathcal{L}}$ , and  $c$  be as before, and let  $u^* = (w_{u^*}, r_{u^*})$  with  $w_{u^*} = w_{p\neg q}$  and  $r_{u^*c}(\{w_{p\neg q}\}) = 1$  and  $r_{u^*c}(\{w\}) = 0$  for all  $w \in Mod_{\mathcal{L}}$  with  $w \neq w_{p\neg q}$ .  $p \Box\rightarrow q$  is true at  $w_{u^*} = w_{p\neg q}$  in  $c$  relative to  $r_{u^*}$ , but  $p \supset q$  is false at  $w_{u^*} = w_{p\neg q}$  in  $c$  relative to  $r_{u^*}$ . This is so because  $P^{r_{u^*c}} = \{w_{pq}\}$  and  $Q = \{w_{pq}, w_{p\neg q}\}$ , and hence  $P^{r_{u^*c}} \subseteq Q$  while  $\{w_{p\neg q}\} \notin P \supset Q = \overline{P} \cup Q$ .

**St 7**  $((\alpha \Box\rightarrow \beta) \wedge (\beta \Box\rightarrow \alpha)) \supset ((\alpha \Box\rightarrow \gamma) \supset (\beta \Box\rightarrow \gamma))$

Suppose  $(\alpha \Box\rightarrow \beta) \wedge (\beta \Box\rightarrow \alpha)$  is true at  $w_u$  in  $c$  relative to  $r_u$ , i.e.  $A^{r_{uc}} \subseteq B$  and  $B^{r_{uc}} \subseteq A$ . This implies  $A^{r_{uc}} = (A \cap B)^{r_{uc}}$  and  $B^{r_{uc}} = (A \cap B)^{r_{uc}}$ , and so  $A^{r_{uc}} = B^{r_{uc}}$ . It follows that  $B^{r_{uc}} \subseteq C$  if  $C^{r_{uc}} \subseteq C$ , i.e. that  $(\alpha \Box\rightarrow \gamma) \supset (\beta \Box\rightarrow \gamma)$  is true at  $w_u$  in  $c$  relative to  $r_u$ .

Strong Centering **St 8** is a consequence of **St 1-7** and **RI 1-2**.

**St 8 (Strong Centering)**  $(\alpha \wedge \gamma) \supset (\alpha \Box\rightarrow \gamma)$

Strong Centering **St 8** is not valid. Let  $\mathcal{L}$ ,  $Mod_{\mathcal{L}}$ ,  $c$ , and  $u^*$  be as in the countermodel to Conditional Excluded Middle **St 5**.  $p \wedge q$  is true at  $w_{u^*} = w_{pq}$  in  $c$  relative to  $r_{u^*}$ , but  $p \Box\rightarrow q$  is not, because  $P^{r_{u^*c}} = P \not\subseteq Q$ .

In Lewis' (1973) terminology **C2** is the system **VCS**. **VCS** results from the basic conditional logic **V** by adding **St 5**, **St 6**, and **St 8**. Lewis' "official logic of

counterfactuals” (Lewis 1973: 132) is the system **VC**, which results from **V** by adding Weak Centering **St 6** and Strong Centering **St 8**. The system **VW** results from **V** by adding Weak Centering **St 6**.

The system **V** consists of the rules of inference **R 1-3** and the axiom schemata **L 1-5**, where  $\equiv$  is the material biconditional,  $\Box\alpha$  is defined as  $\neg\alpha \Box\rightarrow \alpha$ , and  $\diamond\alpha$  is defined as  $\neg\Box\neg\alpha$ .

$$\mathbf{R\ 1} \quad \vdash \alpha, \quad \vdash \alpha \supset \beta \quad \Rightarrow \quad \vdash \beta$$

$$\mathbf{R\ 2} \quad \vdash \beta_1 \wedge \dots \wedge \beta_n \supset \gamma \quad \Rightarrow \quad (\alpha \Box\rightarrow \beta_1) \wedge \dots \wedge (\alpha \Box\rightarrow \beta_n) \supset (\alpha \Box\rightarrow \gamma)$$

$$\mathbf{R\ 3} \quad \vdash \alpha \equiv \beta \quad \Rightarrow \quad \vdash \gamma \equiv \gamma[\beta/\alpha],$$

where  $\gamma[\alpha/\beta]$  results from  $\gamma$  by substituting  $\alpha$  for some occurrence of  $\beta$  in  $\gamma$ .

**L 1** *Truth-functional tautologies.*

$$\mathbf{L\ 2} \quad \alpha \Box\rightarrow \alpha$$

$$\mathbf{L\ 3} \quad \Box\gamma \supset (\alpha \Box\rightarrow \gamma)$$

$$\mathbf{L\ 4} \quad \neg(\alpha \Box\rightarrow \neg\beta) \supset ((\alpha \wedge \beta \Box\rightarrow \gamma) \equiv (\alpha \Box\rightarrow (\beta \supset \gamma)))$$

**V** is sound and complete with respect to the rank-theoretic semantics of counterfactuals. To state and prove this result some technical terminology has to be introduced, though I will ignore contexts in order to avoid unnecessary complications.  $(W, \mathcal{A}_W, R, \Omega, \varphi)$  is a *rank-theoretic model* for  $\mathcal{L}$  just in case  $W$  is a non-empty set of factual worlds,  $\mathcal{A}_W$  is an algebra over  $W$  such that  $\varphi(\alpha) \in \mathcal{A}_W \times \{R\}$  for every  $\alpha$  from  $\mathcal{L}_0$ ,  $R$  is a set of ranking functions  $r : \mathcal{A}_W \rightarrow \mathbb{N} \cup \{\infty\}$ ,  $\Omega \subseteq W \times R$  is such that for each  $w \in W$  there is *at least one*  $r \in R$  such that  $(w, r) \in \Omega$ , and  $\varphi : \mathcal{L} \rightarrow \wp(\Omega)$  is an interpretation function such that for all  $\alpha, \beta$  from  $\mathcal{L}$ :

1.  $\varphi(p) \in (\wp(W) \times \{R\}) \cap \Omega$  if  $p \in PV$
2.  $\varphi(\neg\alpha) = \Omega \setminus \varphi(\alpha)$
3.  $\varphi(\alpha \wedge \beta) = \varphi(\alpha) \cap \varphi(\beta)$
4.  $\varphi(\alpha \Box\rightarrow \beta) = \{u = (w_u, r_u) \in \Omega : f(\varphi(\alpha))^{r_u} \times R \subseteq \varphi(\beta)\},$

where  $f(\varphi(\alpha)) = \{w \in W : \exists r : (w, r) \in \varphi(\alpha)\}.$

$(W, \mathcal{A}_W, R, \Omega, \varphi)$  is a *Humean* rank-theoretic model for  $\mathcal{L}^+$  just in case  $W$  is a non-empty set of factual worlds,  $\mathcal{A}_W$  is an algebra over  $W$  such that  $\varphi(\alpha) \in \mathcal{A}_W$  for every  $\alpha$  from  $\mathcal{L}^+$ ,  $R$  is a set of ranking functions  $r : \mathcal{A}_W \rightarrow N \cup \{\infty\}$ ,  $\Omega \subseteq W \times R$  is such that for each  $w \in W$  there is *exactly one*  $r = r_w \in R$  such that  $(w, r_w) \in \Omega$ , and  $\varphi : \mathcal{L}^+ \rightarrow \wp(W)$  is an interpretation function such that for all  $\alpha, \beta$  from  $\mathcal{L}^+$ :

1.  $\varphi(p) \in \wp(W)$  if  $p \in PV$
2.  $\varphi(\neg\alpha) = W \setminus \varphi(\alpha)$
3.  $\varphi(\alpha \wedge \beta) = \varphi(\alpha) \cap \varphi(\beta)$
4.  $\varphi(\alpha \sqsupset \beta) = \{w \in W : \varphi(\alpha)^{r_w} \subseteq \varphi(\beta)\}$

Every Humean rank-theoretic model  $\mathcal{M}^+ = (W^+, \mathcal{A}_{W^+}, R^+, \Omega^+, \varphi^+)$  for  $\mathcal{L}^+$  can be reduced to a rank-theoretic model  $\mathcal{M} = (W, \mathcal{A}_W, R, \Omega, \varphi)$  for  $\mathcal{L}$  as follows. Let  $W = W^+$  and  $\mathcal{A}_W = \mathcal{A}_{W^+} \cap \{\varphi^+(\alpha) : \alpha \in \mathcal{L}_0\}$ . Take any  $r_w^+ \in R^+$  and restrict it to  $\mathcal{A}_W$ , i.e.  $r_w(A) = r_w^+(A)$  if  $A \in \mathcal{A}_W$  and undefined otherwise. This gives us  $R$ . Let  $\Omega$  be the set of all pairs  $(w, r_w)$  with  $r_w$  resulting from  $r_w^+$  as indicated. Since for each  $w \in W^+ = W$  there is exactly one  $r_w^+ \in R^+$  it immediately follows that there is at least one  $r_w \in R$  for each  $w \in W$ . Finally, let  $\varphi(\alpha) = \varphi^+(\alpha) \times R$  for  $\alpha \in \mathcal{L}$  to obtain  $\varphi : \mathcal{L} \rightarrow \wp(\Omega)$ . It is routine to check that  $\alpha \in \mathcal{L}$  is true in the Humean rank-theoretic model  $\mathcal{M}^+$  just in case  $\alpha$  is true in the rank-theoretic model  $\mathcal{M}$ . Consequently each set  $L$  of sentences from  $\mathcal{L}$  that has a Humean rank-theoretic model also has a rank-theoretic model.

Soundness is easily checked – especially since we are ignoring contexts. **R 1** is **RI 1**. **R 3** is trivial, because classical logic is presupposed. As to **R 2**, suppose  $\vdash \beta_1 \wedge \dots \wedge \beta_n \supset \gamma$ , i.e.  $B_1 \cap \dots \cap B_n \subseteq C$ . If  $(\alpha \sqsupset \beta_1) \wedge \dots \wedge (\alpha \sqsupset \beta_n)$  is true at  $w_u$  relative to  $r_u$ , then  $A^{r_u} \subseteq B_1$ , and  $\dots$ , and  $A^{r_u} \subseteq B_n$  and so  $A^{r_u} \subseteq B_1 \cap \dots \cap B_n \subseteq C$ , which means that  $\alpha \sqsupset \gamma$  is true at  $w_u$  relative to  $r_u$ , for any universe  $u = (w_u, r_u)$ .

**L1** is **St1**. **L2** follows from **St1**, **RI2**, and **St3**. **L3** follows from **St3**. As to **L4**, suppose  $\neg(\alpha \neg \sqsupset \beta)$  is true at  $w_u$  relative to  $r_u$ . This means that  $A^{r_u} \not\subseteq \overline{B}$ . Suppose first that  $\alpha \wedge \beta \sqsupset \gamma$  is true at  $w_u$  relative to  $r_u$ . This means that  $(A \cap B)^{r_u} \subseteq C$ . If  $w_u \in A^{r_u} \cap \overline{B}$ , then  $w_u \in \overline{B} \cup C$ . If  $w_u \in A^{r_u} \cap B$ , then  $w_u \in (A \cap B)^{r_u} \subseteq C$ . Therefore  $A^{r_u} \subseteq \overline{B} \cup C$ , which means that  $\alpha \sqsupset (B \supset \gamma)$  is true at  $w_u$  relative to  $r_u$ . Now suppose  $\alpha \sqsupset (\beta \supset \gamma)$  is true at  $w_u$  relative to  $r_u$ . This means that  $A^{r_u} \subseteq \overline{B} \cup C$ . If  $w \in (A \cap B)^{r_u} \cap A^{r_u}$ , then  $w \in A^{r_u} \cap B$ , and so  $w \in C$ . If  $w \in (A \cap B)^{r_u} \cap \overline{A^{r_u}}$ , then  $r_u(A) < r_u(A \cap B)$  and so  $A^{r_u} \cap B = \emptyset$ , contradicting  $A^{r_u} \not\subseteq \overline{B}$ . Therefore  $(A \cap B)^{r_u} \subseteq C$ , which means that  $\alpha \wedge \beta \sqsupset \gamma$  is true at  $w_u$  relative to  $r_u$ .

This shows that the system **V** on the language  $\mathcal{L}$  is sound with respect to the rank-theoretic semantics of counterfactuals and, consequently, that the system **V** on the language  $\mathcal{L}^+$  is sound with respect to the Humean rank-theoretic semantics of counterfactuals.

Let us now establish completeness.<sup>13</sup> It follows from Lewis (1973: ch. 6) that each **V**-consistent set  $L^+$  of sentences from  $\mathcal{L}^+$  has a model  $\mathcal{M}^* = (W, (\$w)_{w \in W}, \varphi)$ , where for each  $w \in W$ ,  $\$w \subseteq \wp(W)$  is a set of spheres (around  $w$ ), i.e.  $\$w$  is nested and closed under arbitrary unions and intersections. Since **V** has the finite model property with respect to Lewis' semantics, we can assume  $W$  to be finite. For each  $w \in W$ ,  $\$w$  is of the form  $S_0 \cup S_1 \cup \dots \cup S_{w_n}$ , where any two  $S_i$  and  $S_j$  are disjoint and  $\bigcup_{0 \leq i \leq j} S_i$  is the  $j$ -th sphere (around  $w$ ). For  $w' \in W$ , define  $r_w(\{w'\}) = i$  if  $w' \in S_i$ , and let  $r_w(A) = \min\{r_w(\{w'\}) : w' \in A\}$ . Each  $r_w$  defined in this way is a regular ranking function on  $\wp(W)$ . Let  $\mathcal{A}_W = \wp(W)$ ,  $R = \{r_w : w \in W\}$ , and  $\Omega = \{(w, r_w) : w \in W\}$ .  $\mathcal{M}^+ = (W, \mathcal{A}_W, R, \Omega, \varphi)$  is a Humean rank-theoretic model. For  $\alpha \in \mathcal{L}^+$  let  $j_\alpha$  be the smallest number (if any) such that  $\varphi(\alpha) \cap S_{j_\alpha}$  is non-empty. By definition  $\varphi(\alpha)^{r_w} = S_{j_\alpha} \cap \varphi(\alpha)$ . Therefore  $\alpha \in L^+$  is true in  $w$  in  $\mathcal{M}^*$  in Lewis' sense just in case  $\alpha$  is true in  $w$  in  $\mathcal{M}^+$  in our sense.

Now suppose  $L$  is a **V**-consistent set of sentences from  $\mathcal{L}$ . Close  $L$  under the classical connectives  $\neg, \wedge, \vee, \supset$  plus  $\Box \rightarrow$  to obtain a set of sentences  $L^+$  from  $\mathcal{L}^+$ . By the above,  $L^+$  has a Humean rank-theoretic model  $\mathcal{M}^+$ . Reduce  $\mathcal{M}^+$  to obtain a rank-theoretic model  $\mathcal{M}$  for  $\mathcal{L}$  as shown above. This completes the proof for

**Theorem 2** *V restricted to  $\mathcal{L}$  is sound and complete with respect to the rank-theoretic semantics of counterfactuals. V on the full language  $\mathcal{L}^+$  is sound and complete with respect to the Humean rank-theoretic semantics of counterfactuals.*

Call a (Humean) ranktheoretic model  $(W, \mathcal{A}_W, R, \Omega, \varphi)$  *weakly/strongly centered* just in case each  $r_u \in R$  is such that  $r_u(\{w\}) = 0$  if/iff  $w = w_u$ . Then it holds that

**Theorem 3** *VW restricted to  $\mathcal{L}$  is sound and complete with respect to the weakly centered rank-theoretic semantics of counterfactuals. VW on the full language  $\mathcal{L}^+$  is sound and complete with respect to the weakly centered Humean rank-theoretic semantics of counterfactuals.*

**Theorem 4** *VC restricted to  $\mathcal{L}$  is sound and complete with respect to the strongly centered rank-theoretic semantics of counterfactuals. VC on the full language  $\mathcal{L}^+$  is sound and complete with respect to the strongly centered Humean rank-theoretic semantics of counterfactuals.*

<sup>13</sup>The idea to derive completeness from Lewis' completeness results and the finite model property is due to Leitgeb (2012a, 2012b).

## **Acknowledgments**

I am grateful to the participants of the Causaprobe Research Colloquium at the University of Konstanz, especially Thomas Kroedel, Brian Leahy, and Wolfgang Spohn, as well as to Alan Hájek, Jim Joyce, and Hannes Leitgeb for many helpful comments on earlier versions of this paper.

Part of my research was supported by the German Research Foundation through its Emmy Noether program and the Zukunftskolleg of the University of Konstanz.

## References

- [1] Alchourrón, Carlos E. & Gärdenfors, Peter & Makinson, David (1985), On the Logic of Theory Change: Partial Meet Contraction and Revision Functions. *Journal of Symbolic Logic* **50**, 510-530.
- [2] Bennett, Jonathan (2003), *A Philosophical Guide to Conditionals*. Oxford: Clarendon Press.
- [3] Bernstein, Allen R. & Wattenberg, Frank (1969), Non-standard Measure Theory. In W. Luxemburg (ed.), *International Symposium on the Applications of Model Theory to Algebra, Analysis, and Probability*. California Institute of Technology. New York: Holt, Reinhart and Winston, 171-185.
- [4] Briggs, Rachael (2009), The Big Bad Bug Bites Anti-Realists About Chance. *Synthese*, **167**, 81-92.
- [5] Brössel, Peter & Eder, Anna-Maria & Huber, Franz (2013), Evidential Support and Instrumental Rationality. *Philosophy and Phenomenological Research* **87**, 279-300.
- [6] Collins, John & Hall, Ned & Paul, L.A. (eds., 2004), *Causation and Counterfactuals*. Cambridge, MA: MIT Press.
- [7] Connolly, Terry & Ordóñez, Lisa D. & Coughlan, Richard (1997), Regret and Responsibility in the Evaluation of Decision Outcomes. *Organizational Behavior and Human Decision Processes* **70**, 73-85.
- [8] Edgington, Dorothy (1995), On Conditionals. *Mind* **104**, 235-329.
- [9] Edgington, Dorothy (2008), Counterfactuals. *Proceedings of the Aristotelian Society* **108**, 1-21.
- [10] Eriksson, Lina & Hájek, Alan (2007), What Are Degrees of Belief? *Studia Logica* **86**, 185-215.
- [11] Field, Hartry (1978), A Note on Jeffrey Conditionalization. *Philosophy of Science* **45**, 361-367.
- [12] Gibbard, Allan (1981), Two Recent Theories of Conditionals. In W. Harper & R. Stalnaker & G. Pearce (eds.), *Ifs*. Dordrecht: D. Reidel, 211-247.

- [13] Gillies, Anthony S. (2007), Counterfactual Scorekeeping. *Linguistics and Philosophy* **30**, 329-360.
- [14] Gillies, Anthony S. (2009), On Truth-Conditions for *If* (but not quite only *If*). *Philosophical Review* **118**, 325-349.
- [15] Gundersen, Lars Bo (2004), Outline of a New Semantics for Counterfactuals. *Pacific Philosophical Quarterly* **85**, 1-20.
- [16] Hájek, Alan (ms), Most Counterfactuals Are False.
- [17] Hall, Ned (1994), Correcting The Guide to Objective Chance. *Mind* **103**, 505-518.
- [18] Herzberger, Hans G. (1979), Counterfactuals and Consistency. *Journal of Philosophy* **76**, 83-88.
- [19] Hild, Matthias & Spohn, Wolfgang (2008), The Measurement of Ranks and the Laws of Iterated Contraction. *Artificial Intelligence* **172**: 1195-1218.
- [20] Hintikka, Jaakko (1961), *Knowledge and Belief. An Introduction to the Logic of the Two Notions*. Ithaca, NY: Cornell University Press.
- [21] Hoefer, Carl (1997), On Lewis's Objective Chance: 'Humean Supervenience Debugged'. *Mind* **106**, 321-334.
- [22] Huber, Franz (2006), Ranking Functions and Rankings on Languages. *Artificial Intelligence* **170**, 462-471.
- [23] Huber, Franz (2007), The Consistency Argument for Ranking Functions. *Studia Logica* **86**, 299-329.
- [24] Huber, Franz (2011), Lewis Causation is a Special Case of Spohn Causation. *British Journal for the Philosophy of Science* **62**, 207-210.
- [25] Huber, Franz (2013), Structural Equations and Beyond. *The Review of Symbolic Logic* **6**, 709-732.
- [26] Huber, Franz (ms), What Should I Believe About What Would Have Been the Case? Unpublished manuscript.
- [27] Iatridou, Sabine (2000), The Grammatical Ingredients of Counterfactuality. *Linguistic Inquiry* **31**, 231-270.

- [28] Jeffrey, Richard C. (1983), *The Logic of Decision*. 2nd ed. Chicago: University of Chicago Press.
- [29] Joyce, James M. (1998), A Non-Pragmatic Vindication of Probabilism. *Philosophy of Science* **65**, 575-603.
- [30] Joyce, James M. (2009), Accuracy and Coherence: Prospects for an Alethic Epistemology for Partial Belief. In F. Huber & C. Schmidt-Petri (eds.), *Degrees of Belief*. *Synthese Library* **342**. Dordrecht: Springer, 263-297.
- [31] Knobe, Joshua & Nichols, Shaun (eds.), *Experimental Philosophy*. Oxford: Oxford University Press, 2008.
- [32] Kraus, Sarit & Lehmann, Daniel & Magidor, Menachem (1990), Nonmonotonic Reasoning, Preferential Models, and Cumulative Logics. *Artificial Intelligence* **40**, 167-207.
- [33] Kripke, Saul A. (1959), A Completeness Theorem in Modal Logic. *Journal of Symbolic Logic* **24**, 1-14.
- [34] Kroedel, Thomas & Huber, Franz (2013), Counterfactual Dependence and Arrow. *Noûs* **47**, 453-466.
- [35] Leitgeb, Hannes (2012a), A Probabilistic Semantics for Counterfactuals. Part A. *Review of Symbolic Logic* **5**, 26-84.
- [36] Leitgeb, Hannes (2012b), A Probabilistic Semantics for Counterfactuals. Part B. *Review of Symbolic Logic* **5**, 85-121.
- [37] Lewis, David K. (1973), *Counterfactuals*. Cambridge, MA: Harvard University Press.
- [38] Lewis, David K. (1979), Counterfactual Dependence and Time's Arrow. *Noûs* **13**, 455-476.
- [39] Lewis, David K. (1980), A Subjectivist's Guide to Objective Chance. In R.C. Jeffrey (ed.), *Studies in Inductive Logic and Probability*. Vol. II. Berkeley: University of Berkeley Press, 263-293.
- [40] Lewis, David (1981), Ordering Semantics and Premise Semantics for Counterfactuals. *Journal of Philosophical Logic* **10**, 217-234.

- [41] Lewis, David K. (1986), Introduction. In D. Lewis (1986), *Philosophical Papers II*. Oxford: Oxford University Press, ix-xvii.
- [42] Lewis, David K. (1994), Humean Supervenience Debugged. *Mind* **103**, 473-490.
- [43] Menzies, Peter (2004), Difference-Making in Context. In J. Collins & N. Hall & L.A. Paul (eds.), *Causation and Counterfactuals*. Cambridge, MA: MIT Press, 139-180.
- [44] Moss, Sarah (2012), On the Pragmatics of Counterfactuals. *Noûs* **46**, 561-586.
- [45] Mumford, Stephen (1998), *Dispositions*. Oxford: Oxford University Press.
- [46] Nozick, Robert (1981), *Philosophical Explanations*. Oxford: Oxford University Press.
- [47] Percival, Philip (2002), Epistemic Consequentialism. *Supplement to the Proceedings of the Aristotelian Society* **76**, 121-151.
- [48] Popper, Karl R. (1955), Two autonomous axiom systems for the calculus of probabilities. *British Journal for the Philosophy of Science* **6**, 51-57.
- [49] Quine, Willard Van Orman (1950), *Methods of Logic*. New York: Holt, Rinehart, and Winston.
- [50] Rényi, Alfred (1955), On a New Axiomatic System for Probability. *Acta Mathematica Academiae Scientiarum Hungaricae* **6**, 285-335.
- [51] Shenoy, Prakash P. (1991), On Spohn's Rule for Revision of Beliefs. *International Journal of Approximate Reasoning* **5**, 149-181.
- [52] Sobel, Howard J. (1970), Utilitarianisms: Simple and General. *Inquiry* **13**, 394-449.
- [53] Spohn, Wolfgang (1986), On the Representation of Popper Measures. *Topoi* **5**, 69-74.
- [54] Spohn, Wolfgang (1988), Ordinal Conditional Functions: A Dynamic Theory of Epistemic States. In W.L. Harper & B. Skyrms (eds.), *Causation in Decision, Belief Change, and Statistics II*. Dordrecht: Kluwer, 105-134.

- [55] Spohn, Wolfgang (2010), Chance and Necessity: From Humean Supervenience to Humean Projection. In E. Eells & J. Fetzer (eds.), *The Place of Probability in Science. Boston Studies in the Philosophy of Science* **284**. Dordrecht: Springer, 101-131.
- [56] Spohn, Wolfgang (2012), *The Laws of Belief. Ranking Theory and its Philosophical Applications*. Oxford: Oxford University Press.
- [57] Stalnaker, Robert C. (1968), A Theory of Conditionals. In N. Rescher (ed.), *Studies in Logical Theory. American Philosophical Quarterly. Monograph Series* **2**. Oxford: Blackwell, 98-112.
- [58] Stalnaker, Robert C. (1970), Probability and Conditionality. *Philosophy of Science* **37**, 64-80.
- [59] Stalnaker, Robert C. (1981), A Defense of Conditional Excluded Middle. In W. Harper & R. Stalnaker & G. Pearce (eds.), *Ifs: Conditionals, Belief, Decision, Chance, and Time*. Dordrecht: D. Reidel, 87-104.
- [60] Stalnaker, Robert C. (1996), Varieties of Supervenience. *Philosophical Perspectives* **10**, 221-241.
- [61] Stalnaker, Robert C. (1998), On the Representation of Context. *Journal of Logic, Language, and Information* **7**, 3-19.
- [62] Stalnaker, Robert C. (1999), *Context and Content*. Oxford: Oxford University Press.
- [63] Stalnaker, Robert C. (2002), Epistemic Consequentialism. *Supplement to the Proceedings of the Aristotelian Society* **76**, 153-168.
- [64] Thau, Michael (1994), Undermining and Admissibility. *Mind* **103**, 491-503.
- [65] Williamson, Timothy (2000), *Knowledge and Its Limits*. Oxford: Blackwell.
- [66] Williamson, Timothy (2007), *The Philosophy of Philosophy*. Oxford: Blackwell.