Journal of Applied Logic 12 (2014) 235–251

Contents lists available at ScienceDirect

Journal of Applied Logic

www.elsevier.com/locate/jal

Probabilities of counterfactuals and counterfactual probabilities

Alan Hájek [1]

Australian National University, Canberra, ACT, 0200, Australia

ARTICLE INFO

Article history:
Available online 7 February 2014

Keywords:
Counterfactuals
Probability
Chance
Truth conditions
Edgington
Leitgeb
Similarity

ABSTRACT

Probabilities figure centrally in much of the literature on the semantics of conditionals. I find this surprising: it accords a special status to conditionals that other parts of language apparently do not share. I critically discuss two notable 'probabilities first' accounts of counterfactuals, due to Edgington and Leitgeb. According to Edgington, counterfactuals lack truth values but have probabilities. I argue that this combination gives rise to a number of problems. According to Leitgeb, counterfactuals have truth conditions-roughly, a counterfactual is true when the corresponding conditional chance is sufficiently high. I argue that problems arise from the disparity between truth and high chance, between approximate truth and high chance, and from counterfactuals for which the corresponding conditional chances are undefined. However, Edgington, Leitgeb and I can unite in opposition to Stalnaker and Lewis-style 'similarity' accounts of counterfactuals.

1. Introduction

My topic is the interaction of probabilities and counterfactuals. Can probabilities illuminate the semantics of counterfactuals: their truth conditions (or lack thereof), their approximate-truth conditions, and their logic?

It is natural to think that probabilities of conditionals are parasitic on their truth conditions. After all, that is the model that we find with the basic connectives: probabilities of conjunctions, disjunctions, and negations are determined by their respective truth conditions, rather than the other way round. Similarly, probabilities of quantified statements display the same order of dependence. And so it apparently goes with

[1] For valuable discussion and inspiration, I am grateful particularly to Sharon Berry, Rachael Briggs, Yoaav Isaacs, Leon Leontyev, Karen Lewis, Hanti Lin, Daniel Nolan, Paolo Santorio, Wolfgang Schwarz, Julia Staffel, Hlynur Stefánsson, and the audience at the Progic 2013 conference. Thanks again to Rachael Briggs and Hanti Lin for razor-sharp comments on a near-final draft. Special thanks to Hannes Leitgeb for his generous comments and most illuminating discussion. I am grateful also to Dorothy Edgington, David Etlin, and James Hawthorne for helpful e-mail correspondence, to an anonymous referee for the *Journal of Applied Logic*, to the organizers of the conference (especially Niki Pfeifer), and to the Munich Center for Mathematical Philosophy for their largesse and hospitality.

http://dx.doi.org/10.1016/j.jal.2013.11.001
1570-8683/

almost all of our language that is probability-apt: it's truth first, probability second.[2] But a striking theme in much of the conditionals literature is that this order of dependence is reversed for conditionals: their probabilities are regarded as primary, and consequences for their truth conditions—or lack thereof—are then drawn. Then there are accounts that write probabilities into the truth conditions themselves, thus still giving probabilities primacy.[3] Either way, I find it surprising that conditionals should have special status in having their semantics underpinned probabilistically. But if they should, then that fact is striking in itself.

In this paper I will discuss two notable 'probabilities first' accounts of counterfactuals: those presented in agenda-setting papers by Edgington [6] and by Leitgeb [16,17].

2. Edgington's account

2.1. Outline of the account

Edgington begins with a partial job description for counterfactuals:

> they have figured ... in accounts of causation, perception, knowledge, rational decision, action, explanation, and so on. And outside philosophy, in ordinary life, counterfactual judgements play many important roles, for instance in inferences to factual conclusions ... (1)

She has long been an influential advocate of a *suppositional* account of indicative conditionals.

> On this view, a conditional statement is not a categorical assertion of a proposition, true or false as the case may be; it is rather a statement of the consequent under the supposition of the antecedent. A conditional belief is not a categorical belief that something is the case; it is belief in the consequent in the context of a supposition of the antecedent. (2)

Enter probability theory. Edgington regards the strongest evidence for the suppositional account to come from the way uncertain conditional judgments work. Our best theory of uncertainty is probability theory, and our best understanding of conditional uncertainty is conditional probability. Putting these ideas together, a core part of her account is Adams' Thesis that the probability of an indicative conditional 'if A, then B' is the conditional probability of B, given A (where this is defined). But this probability is not to be understood as that of the conditional's *truth*. On the contrary, it is equally central to the account that indicative conditionals do not have truth values.

Various authors (including [1,3,4,8], and [22]) subscribe to this theory of indicative conditionals alongside Edgington. But most authors believe that counterfactuals must be treated differently—typically with some sort of 'similarity' semantics, à la Stalnaker [27] and Lewis [19]. Edgington insists, however, that counterfactuals and indicative conditionals should be given parallel treatment. She forcefully argues that "the easy transitition [sic.] between 'suppose' and 'if' is as evident for subjunctives as it is for indicatives." (4–5). And it is certainly welcome that her view gives a unified treatment of 'if'—especially so when indicatives and counterfactuals seem to coincide in the future tense. Above all, according to her they should be given unified *probabilistic*, rather than truth conditional, treatment.

However, as she is well aware, their treatment had better not be *too* unified: witness our divergent assessments of

[2] Or such is the dominant position in the literature. Exceptions include Field [7], Leblanc [15], and Harper [13], who offer probability-first semantics for classical logic. Of course, once we introduce the counterfactual connective into our language, our logic is no longer classical.

[3] The probabilities that are given primacy here might not be probabilities *of counterfactuals*, but rather probabilities of the counterfactuals' consequents, given their antecedents—as we will soon see.

if Oswald didn't shoot Kennedy, then someone else did

and

if Oswald hadn't shot Kennedy, then someone else would have.

(The former seems highly assertable, while the latter does not.) According to Edgington, while indicative conditionals such as the former are governed by Adams' Thesis, counterfactuals such as the latter go by a different conditional probability:

> ...conditional probability—the probability of C on the sup-position [sic.] that A—is, I claim, the key to how close to certain we are of a conditional of any kind. For subjunctive conditionals, this conditional probability does not (normally) represent your current degree of belief in C given A, but (most typically) your view about how likely it *was* that C *would have happened*, given that A had." (5)

The locution "your view about how likely it was ..." has two components: "your view", and "how likely it was". This combines a subjective and an objective element. It is clear from the preceding material that Edgington regards the objective element to be the *conditional chance* at the appropriate time, at least in typical cases.

If you have a single "view" as to what that conditional chance is, then it seems that that's the value you should assign to the counterfactual. However, you may instead entertain various such "views", spreading credences over various hypotheses about what the conditional chance is. Then the subjective element needs some massaging. I think it best to think of the required value as your *expectation* of the conditional chance at that time[4]: the weighted average of the conditional chance according to each hypothesis, the weights being your credences for each hypothesis that it is the correct one. (This brings us close to Skyrms' [26] account of counterfactuals.) And what is that time? The italicized words in the quoted passage are the key: it is not the present time, but rather, I take it, a past time just before the truth value of the antecedent was settled.[5]

I want to focus on the central pillars of Edgington's account: counterfactuals *have probabilities* (typically suitably related to conditional chances), but they *do not have truth values*.

2.2. Some questions about Edgington's account

Let me raise a number of questions about Edgington's account.

---

[4] Thanks here to Rachael Briggs.

[5] There is a complication. Edgington realizes that her account needs some tweaking because of cases like this:

> the car breaks down on the way to the airport and I miss my flight. 'If I had caught that plane I'd be half way to Paris by now', I remark to the mechanic, who has just been listening to the radio. 'You're wrong', he tells me. 'It crashed. If you had caught that plane you would be dead by now.' ...

> This example and others like it suggest that the conditional probability we are concerned to estimate, for counterfactuals (and in a sense the ultimate verdict on some forward-looking wills), is the chance, at a time when A still had some chance of coming about, of C given A and any relevant, *causally independent*, subsequent facts that have a causal bearing on C. (15, my italics)

The thought is something like this. The plane in fact crashed, and something caused it to do so. We are to imagine that whatever this cause was, it was independent of Edgington catching the flight. We should then conjoin the fact that specifies this cause to the antecedent to form the condition of the conditional expectation; we then take the expectation of *that* quantity, much as before.

There is something right about this idea, but it isn't exactly right as it stands. Let us suppose that the plane crashed because it was overloaded with people, and it was as a result too heavy. Now suppose that Edgington had caught the plane. It would have been *even heavier* (only slightly, to be sure!). Then, *all the more* the plane would have crashed, it seems. All the more we want to factor in the crashing of the plane, even though it is *not* causally independent of the antecedent, and it thus does not meet her condition.

Notice, furthermore, how naturally one understands the mechanic's statement "You're wrong" as telling her that her counterfactual was *false* (as opposed to having some other defect); and he goes on to offer one that he regards as *true*.

i) Are Edgington's "probabilities" of counterfactuals really probabilities?

Here and elsewhere I put Edgington's "probabilities" of counterfactuals in scare quotes, because for a number of reasons I think that they are not really *probabilities*—neither formally, nor in virtue of their functional role.

Formally, in order to be probabilities (on the standard understanding), they must obey the probability calculus. But what sense are we to make of the additivity rule as applied to them? It is an axiom that the probability of a disjunction of incompatible sentences is the sum of their individual probabilities. There are two problems here for truth-valueless sentences: *disjunction*, and *incompatibility*.

First, *disjunction*. It is a truth-functional operation; how should it be defined for sentences that lack truth values? To be sure, Edgington's account is solely about single counterfactuals, not more complex sentences in which they might appear, such as disjunctions. Yet such disjunctions may have perfectly determinate meaning. For example, I submit that the following is a tautology:

'if Oswald hadn't shot Kennedy, someone else would have **or**
it's not the case that if Oswald hadn't shot Kennedy, someone else would have'

As such, it should receive probability 1. I wonder how this can be if the disjuncts lack truth value. More tellingly, if disjunctions in which counterfactuals appear (such as this one) *do not* have probabilities, then the counterfactuals themselves fail to have probabilities for an even more basic reason: probabilities are defined over an *algebra*, closed under negation and disjunction—the two operations I employed to produce the tautology. In that case, at best counterfactuals might receive values that are probability-like—call them *quasi-probabilities*—but not genuine probabilities.[6]

Second, *incompatibility*. This is usually defined in terms of truth: two sentences are incompatible if it is not possible for both of the sentences to be *true*. Either this notion makes no sense for truth-valueless sentences, or it trivially applies to all pairs of such sentences (for one way to fail to be true is to lack a truth value).

Now, there may seem to be a natural way to understand 'incompatibility' for truth-valueless but probability-bearing sentences: not in terms of truth, but in terms of probability! For example, we might say that p and q are *probabilistically incompatible* just in case: for all probability functions P, $P(p \text{ and } q) = 0$. And it may seem natural for Edgington to understand 'incompatibility' for counterfactuals in these probabilistic terms. (Compare how no-truth-value theorists about conditionals cash out *validity* of arguments involving them in probabilistic terms—see [2].)

I have two concerns with this proposal. First, probabilistic incompatibility is defined *in terms of probability*. But what is probability? Well, it's characterized by the axioms of probability: additivity, ... Hang on—the problem was that it is unclear how to make sense of the additivity axiom for truth-valueless sentences, and thus unclear how to make sense of probability assignments to them; it is no response simply to appeal to *probability* assignments to them! Second, the definition I have suggested of probabilistic incompatibility requires a *conjunction* of sentences ('p and q'). However, Edgington's account applies to single counterfactuals; I am not sure how it applies to compounds such as the conjunction of a counterfactual with another sentence. Again, conjunction is a truth-functional operation; how does it operate on sentences that lack truth values?

Another tempting way to understand incompatibility for truth-valueless counterfactuals might be in terms of the suppositional account. Under the subjunctive supposition that Oswald hadn't shot Kennedy, 'someone else would have' and 'it's not the case that someone else would have' are straightforwardly incompatible: under that supposition, they *cannot both be true*. More generally, 'if it were that p, it would be that q' and 'if it were that p, it would be that r' are incompatible if q and r are incompatible in the usual sense. This works

[6] This recalls Lewis's [20] discussion of Adams' [1] theory of conditionals.

for a single supposition; but I don't see how to generalize it for different suppositions. What determines the incompatibility (or otherwise) of 'if it were that p, it would be that q' and 'if it were that s, it would be that t'? At best, we have an impoverished notion—call it *quasi-incompatibility*.[7] Moreover, whenever you believe that there is a (genuinely, not merely probabilistically) valid inference involving a counterfactual, you are committed to it figuring in an inconsistent set of sentences. For example, if you believe that modus ponens is valid for counterfactuals, then you are committed to the inconsistency of any set of the form {'if it were that p, it would be that q', 'p', and '*not-q*'}. But Edgington's suppositional account cannot do justice to this. Now, perhaps she would not *want* to do justice to this, insisting that this is not really an inconsistent set after all—quasi-incompatibility is the closest relation to incompatibility that counterfactuals can figure in. If so, that only underscores my point that they do *not* receive genuine probabilities, which are defined in terms of genuine incompatibility. Again, at best they can receive only quasi-probabilities.

I don't want to leave you with the impression that this is just a verbal dispute. After all, when Edgington speaks of the probability of a counterfactual, she simply *means* a corresponding conditional probability. Should I begrudge her this way of talking?[8] The issue is just how *quasi* these quasi-probabilities of counterfactuals are. It is questionable whether formally they are probabilities if they don't conform to the most fundamental requirements of the standard understanding of probability—if they are not defined on an algebra, and if they are not additive in the usual sense.

That said, perhaps we should not be so beholden to the standard formal understanding of probability.[9] Maybe Edgington's "probabilities" of counterfactuals deserve the name in virtue of their functional role—if they walk like probabilities, and talk like probabilities, then maybe they *are* probabilities, or close enough.

Well, do they? Arguably the most important functional role of one's subjective probabilities is the guidance they give to one's *actions*, as codified in decision theory. And the paradigm of this guidance is found in *betting behaviour*. However, there is the problem of making sense of betting on something that has no truth value: there are no conditions for determining whether the bet wins or loses. Note that this is not merely the old concern about the betting interpretation that it misrepresents one's attitudes to unverifiable propositions, which can never be settled favourably. There, at least, it is clear what it takes for a bet on such a proposition p to win: it does so just in case p is true. Whether or not anyone realizes or verifies that it is a winning bet is another matter. We might imagine God settling the matter. But if p has no truth value, there is nothing for God to settle: there are no winning conditions for the bet.

Now, it might seem that Edgington has an easy answer to this problem. Since she understands probabilities of counterfactuals in terms of corresponding conditional probabilities, surely she should understand betting on counterfactuals in terms of corresponding *conditional bets*—bets that are called off if their conditions are not met. Specifically, a conditional bet at stakes of \$1 on q, given p,

- pays \$1 if p & q
- pays 0 if p & $\neg q$
- is called off if $\neg p$

Conditional bets corresponding to counterfactuals will typically be called off—after all, the hallmark of typical counterfactuals is that their antecedents are *false*! And when one knows that the antecedent of a given counterfactual is false, as is also typical, one is in a position to know that the corresponding bet is called off—in which case one will presumably pay nothing for the bet. For example, suppose you know that this fair die was not tossed at noon. Nevertheless, I take it that Edgington would have you assign high

[7] Thanks here to Rachael Briggs.

[8] I am grateful to Hannes Leitgeb for pressing this point.

[9] Indeed, I have long questioned it myself, advocating instead an understanding in which conditional probability is primitive, à la Popper [23]. (See especially my [9].) However, Popper functions will offer no refuge for Edgington's account, since their domain is closed under the binary connective '&' and the monadic operator '∼'.

probability, 5/6, to the counterfactual: 'if the die had been tossed at noon, it would have landed greater than 1'. This probability does *not* guide your valuation of a conditional bet at stakes \$1 on the die landing greater than 1, given that it is tossed at noon. For knowing as you do that this bet has been called off, you value the bet at 0, not at \$5/6.

In sum, your probability for a counterfactual typically does *not* guide at all your betting behaviour regarding the counterfactual, by Edgington's lights. It does not guide your betting behaviour regarding an unconditional bet on the truth of the counterfactual—there is no such bet according to her. Nor, typically, does it guide your betting behaviour regarding a conditional bet on the consequent, given the antecedent, when you know that the antecedent is false, and hence that the conditional bet is called off—you value the bet at 0, when typically your probability for the counterfactual will be positive. The usual connection between probability and betting behaviour, a central part of probability's functional role, is absent.

To be sure, I take it that on Edgington's proposal, what matters is not one's *actual* behaviour regarding such a bet, based on one's *actual* valuation of it. As she writes, "the conditional probabilities relevant to the assessment of subjunctive conditionals do not (typically) represent your present actual distribution of belief, but those of a hypothetical belief state in a different context, normally that of an earlier time" (8). So what matters, rather, is a *hypothetical* valuation, one made from the standpoint of a *hypothetical* credal state, which guides the betting behaviour of a *hypothetical* agent—a differently-situated counterpart of you, as we might say. (Imagine projecting yourself into the shoes of a counterpart of yourself situated shortly before noon, for whom the toss of the die is a live possibility, and whose future-directed conditional probability for the die landing greater than 1, given that it is tossed at noon, is 5/6.) Notice how different this makes probabilities of counterfactuals from probabilities of other sentences, which go by one's *actual* valuations of corresponding bets. Nevertheless, perhaps this is connection enough between probability and betting behaviour to subserve this part of probability's functional role? I am not convinced that it is—your probabilities guide *you*, never mind some fictional agent. That said, the fact that they provide action-guidance at all is part of that role. So I do not claim that the matter has been settled.

I have questioned whether Edgington's "probabilities" of counterfactuals are properly so-called. However, let us suppose for the sake of the argument that they are. A further question then immediately arises.

*ii) Why **should** an agent's assignments of such "probabilities" to counterfactuals obey the probability calculus?*

Why *should* your subjective probabilities for counterfactuals obey the probability calculus? There are various arguments that credences assigned to sentences should do so, on pain of irrationality; but most of these arguments are parasitic on the sentences having truth values. I will focus on three such arguments.

Consider the *Dutch Book argument* for probabilism. (See [24].) We identify your credences with your betting prices, and show that if these are not probabilities, then you are susceptible to a Dutch Book: a set of bets each of which you will regard as acceptable, but which collectively guarantee your loss. However, again it is hard to make sense of betting on something that has no truth value, and never will: there are no conditions for determining whether the bet wins or loses. If we can make sense at all of such a bet, it is clear what its fair price is: *zero*. After all, it is a bet that cannot pay off.

Or consider the *calibration argument* for probabilism. (See [30].) Calibration is a measure of how well one's credences match the corresponding relative frequencies. Consider all your credences over a set of sentences. You are perfectly calibrated if for each p, among those sentences to which you assigned credence p, proportion p of them were *true*. More generally, we can measure how well calibrated your assignments are with respect to a set of sentences, where this comes in degrees that may fall short of perfection. We can then show that if your assignments violate the laws of probability, then there is a probability function whose assignments are guaranteed to better calibrated than yours; your credence function is *calibration-dominated*. However, it is difficult to make sense of calibration with respect to a set of sentences that don't all have truth values. Edgington apparently cannot appeal to the calibration argument in support of her probabilities of counterfactuals obeying probability theory, for by her lights there is nothing to which these probabilities can be calibrated.

Or consider the *accuracy* argument for probabilism. (See [14].) Every time you assign a credence to a proposition, there is a fact of the matter of the distance by which your credence missed the proposition's truth value (identifying 1 with truth and 0 with falsehood). The smaller this distance, the more accurate your credence is. We may give an overall accuracy score to your entire credence function, according to how well it fares accuracy-wise over all the propositions in its domain. We can show that if your credence function is not a probability function, then there is a probability function that is guaranteed to score better than your function; your credence function is *accuracy-dominated*. However, it is especially clear that no sense can be made of 'accuracy' with respect to sentences that lack truth values, and thus that do not express propositions. Edgington apparently cannot appeal to the accuracy argument in support of her probabilities of counterfactuals obeying probability theory, for by her lights they have no truth values by which accuracy can be defined.

Now, much as before, it may seem that Edgington has an obvious answer to these concerns: we all know that there are various arguments for why *conditional probabilities* should obey the probability calculus, and by her lights, probabilities of counterfactuals are conditional probabilities. What's the problem?

The problem is that the arguments show that *you* are somehow at fault if *your* conditional probabilities do not obey the probability calculus: *you* are Dutch bookable, *you* are calibration-dominated, *you* are accuracy-dominated. But on Edgington's proposal, the person who suffers these indignities is not *you*, but a non-actual counterpart of you.

Consider the Dutch Book argument. Suppose that your probability assignments to certain counterfactuals do not obey the probability calculus. It is not *you* who is susceptible to sure losses at the hands of a Dutch bookie. After all, *you* will not take conditional bets at prices in accordance your probability assignments. (Recall how *you* valued the conditional bet on the die landing greater than 1, given that it is tossed, at 0, which is not in accordance with your probability of 5/6 to the corresponding counterfactual.) Your probability assignments correspond to the betting prices of a *hypothetical* agent, a differently-situated counterpart of you. That's the agent who hypothetically suffers the sure losses, not you.

The calibration and accuracy arguments go much the same way. Suppose that your probability assignments to certain counterfactuals do not obey the probability calculus. It is not *you* who is calibration-dominated, or accuracy-dominated; it's a differently-situated counterpart to you. I think the challenge for Edgington, then, is to show how the ill fates of these counterparts of yours reflect ill on *you*.

Perhaps the challenge can be met. Perhaps their ill fates merely dramatize a defect in *your* mental economy: you are in some important sense endorsing their ill fates, and that redounds upon *your* head.[10] It is noteworthy, though, that this brings out a striking disanalogy with the original Dutch book, calibration, and accuracy arguments for probabilism, in which the ill fate of your violating the probability calculus fell directly upon *you*.[11] And one wonders whether counterfactuals are the only sentences of our language that have this rather peculiar property of shifting the burden of one's doxastic derelictions onto other, non-actual agents.

This brings me to my next question. Let's grant that it makes sense to speak of Edgington's "probabilities" of counterfactuals obeying the probability calculus, and that an agent's assignments of such "probabilities" should do so, while the counterfactuals themselves lack truth values. That still makes counterfactuals odd critters.

[10] Thanks to Rachael Briggs for suggesting a version of this point.

[11] In the case of the original Dutch Book argument, the ill fate may not literally be that of your losing money—there may not be any bookies on hand to milk you—but rather that of having inconsistent valuations of the very same betting arrangement, depending on how it is presented to you. (See Skyrms [25,26].) But even this ill fate falls upon your non-actual counterpart, not you, when it is that counterpart who is Dutch Bookable, not you.

iii) Are conditionals loners on the Edgington account?

Are counterfactuals and indicative conditionals to be the only sentences of our language that have probabilities, but not truth values? We are familiar with other sentences that lack truth values:

- imperatives, such as "Shut up!"
- questions, such as "Who is that in the gorilla suit?"
- optatives, such as "May your dress sense improve!"
- hortatives, such as "Go, Collingwood!"
- exclamations, such as "Yay!" (after a Collingwood win)
- expletives, such as "xxxx!" (after a Collingwood loss).

But they also lack probabilities, and probability-related locutions crash on them:

- "The probability of 'shut up!' is 0.3".
- "The chance of 'who is that in the gorilla suit?' is 0.8".
- "The conditional probability of 'may your dress sense improve!', given 'go, Collingwood!', is high.
- 'xxxx!' is three times as probable as 'yay!'. (Though admittedly I have been to faculty meetings where I was somewhat tempted to say something like that.)

And this pattern seems to be no accident. Probabilities attach to the *satisfaction* of their arguments, and a sentence gets to be satisfied by being *true*. Conditionals, on the Edgington view, seem to live in a strange halfway house: lacking truth values (like all these kinds of sentences), but *having* probabilities (unlike any of them).[12] Perhaps there is a reason for the house to stay empty.

That said, she is not alone in giving conditionals special treatment—recall my introductory observation that much of the conditionals literature reverses for them the usual order of dependence of probabilities on truth. And perhaps conditionals are not entirely alone in that house after all. Perhaps epistemic modals live there too. Consider 'the butler might have done it', said in a context in which I am not sure whether the evidence conclusively rules out his guilt or not.[13] Perhaps I may assign this intermediate probability, while on various accounts of the epistemic modal 'might', it lacks a truth value. Perhaps moral claims, attributions of predicates of taste, and claims about what is good, and what is rational, also live there on a Gibbardian expressivist account of them. Perhaps future contingents live there. If so, some of the puzzlement that I have voiced over the last page or so carries over to these other inhabitants of the halfway house; but at least they are not alone. Then again, perhaps *not*–all these putative examples of truth-valueless claims are controversial, in a way that my bulleted examples above are not. I am not aware of any *uncontroversial* residents of the house.

Anyway, let us suppose for the sake of the argument that counterfactuals are not so odd after all, enjoying some epistemic modal, expressivist, and future-contingent company. Still, much of that company is *subjectivist* in nature (future contingents are the exceptions), and where it is, the subjectivism seems crucial to the putative admission into the halfway house. Is that the sort of company that we want counterfactuals to keep?

iv) Does Edgington's account do justice to the objectivity of counterfactuals?

There is a subjectivist element in Edgington's account of counterfactuals. It appeared in her locution "*your* view". I glossed this in terms of *your* expectation (of a conditional chance). But it seems that many

[12] On Edgington's view, conditionals also have a *modal* status—for probability is a kind of modality—while not having a *truth* status. I also find this to be an odd combination.

[13] Thanks to Daniel Nolan for the example, and to him and Karen Lewis for the cases that immediately follow.

counterfactuals that matter to us have nothing to do with you, or me, or anyone else. Consider, for example, the counterfactuals that we find in science. To be sure, we may have various attitudes to them; but the counterfactuals themselves are not to be understood in terms of such attitudes. Whatever standing they have, they have objectively.

Moreover, if counterfactuals are to serve in accounts of entirely objective properties or relations in the world, such as causation or dispositions, then they had better be objective themselves. For grounding them in subjective attitudes such as credences renders mysterious how objectivity can emerge therefrom. And while other properties and relations on Edgington's initial job description list involve subjects—perception, knowledge, action, and perhaps explanation—the properties and relations themselves are objective. For it is an objective matter who perceives what, who knows what, what actions are performed, and so on. To be sure, claims about rational decision may be exceptional on this list, but even their subjectivity arises not from counterfactuals themselves, but from an agent's attitudes to them.

Now, perhaps Edgington is only giving an account of what it is to be more or less *confident* of a counterfactual—after all, she speaks of "how close to certain we are" in counterfactuals, which of course is a subjective matter. It is compatible with this that there is an objectively correct, fully-informed, 'God's eye view' on counterfactuals.[14] Then I look forward to the account being further developed. After all, we can hold many other attitudes towards counterfactuals besides having various degrees of belief in them—we can desire them, despise them, apprehend them, comprehend them, forget them, regret them, ... We can do many other things with them—we can state them, conflate them, reject them, retract them, explain them, ascertain them, ... And counterfactuals serve many other purposes, as the job description list makes clear, not to mention their place in science. Stalnaker's and Lewis's similarity-based accounts have no difficulty recognizing the many roles of counterfactuals. Can Edgington's probability-based account do likewise?

It is tempting to strip away from Edgington's account its subjective element, its cloak of confidence—"your view". What remains is "how likely it *was* that C *would have happened*, given that A had." When that is understood in terms of conditional chance, it is fully objective. It is also fit to be the content of various attitudes and to play various roles. And as it happens, it brings us to Leitgeb's account. While he solves some of the problems that I have raised for Edgington, he faces some others.

3. Leitgeb's account

3.1. Outline of the account

In his rich papers [16] and [17], Leitgeb offers a new semantics for counterfactuals. Like Edgington's, it is probabilistic; unlike Edgington's, it accords truth values to counterfactuals. An important motivation of his is admirably naturalistic: to "ground counterfactuals semantically in a scientifically respectable manner" (26). Relatedly, he hopes his probabilistic semantics will "state what the correct truth conditions for many occurrences of counterfactual statements in natural language ought to be like." (60) Thus, he offers more a Carnap-style explication than a conceptual analysis of counterfactuals as we find them in the wild. I can hardly do justice to his two papers in this short space. For example, I will not touch on his two kinds of pragmatic meanings of counterfactuals, his triviality result, or his representation theorems, though I commend them to you. Instead, I will concentrate on his truth conditions.

I take there to be three main steps leading to them. They proceed as follows. Step one:

> [consider] the subjunctive conditional

[14] Thanks here to an anonymous referee.

> (1) If the match were struck, it would light
> . . . As step one, it should be unproblematic to qualify the consequent of (1) in either of the following ways, while leaving the meaning of the conditional—completely or at least almost completely—unaffected:
>
> (2) If the match were struck, it would necessarily [definitely] light.
> (3) If the match were struck, it would be very likely to light.
>
> Without any bias from some philosophical theory, (2) and (3) simply seem to say pretty much the same as (1). (36)

Step two:

> Next, we intend to analyse (3) with the aim of reformulating it somewhat more transparently. A sentence such as (3) invites two obvious questions: Which kind of probability is expressed by the qualification 'very likely'? And what is the logical form of the whole conditional statement? . . . The probability in question is an ontic one—single-case chance . . .—and the logical form of (3) is the one of:
>
> (4) The conditional chance of the match lighting given that it is struck is very high. (37)

Generalizing from this particular case of Leitgeb's, I take it the main steps so far are:
Step 1. 'if p were the case, q would be the case' means almost the same as 'if p were the case, q would very probably be the case'.
Step 2. The latter is analysed as: $ch(q|p)$ is very high, where ch is the objective chance function (suitably relativized to a world and time). "Very high" is understood as "above a vaguely determined threshold" (113).

The third and final step is to give a precise mathematical characterization of such conditional chances. He wants to allow for them to be defined even when the condition p has chance 0. This can be achieved by taking conditional chances as primitive, and there is a framework to represent this:
Step 3. ch is a Popper function.

In a nutshell, this is a version of what [3] calls the "near-miss proposal", which we may symbolise as follows:

$$p\Box \rightarrow q \text{ is true iff } ch(q|p) \text{ is very high}$$

(remembering that for Leitgeb, $ch(\_|\_)$ is a Popper function).

Again, I elide over many details (e.g. the time to which the chance function should be indexed); but I think that steps 1–3 capture the essential features of Leitgeb's account. The appeal to chance secures the objectivity of the truth conditions for counterfactuals; their *conditionality* is captured by the appeal to *conditional* chance.

I will now discuss these steps in turn.

3.2. Questioning Leitgeb's account

3.2.1. Questioning Step 1

There may be close connections between the truth of a proposition and its having very high chance, but I do not think that *near synonymy* is one of them. "It will rain tomorrow" and "it will very probably rain tomorrow" usually have the same truth value, and believing one will typically have the same impact on one's behaviour as believing the other (though not always). However, their *meanings* are quite different. A sunny, rain-free day tomorrow falsifies the first, but not the second.

And so it is with counterfactuals, I would have thought. There is something defective about

"If the match were struck, it *would* light; but if the match were struck, it *might not* light."

Following Lewis [19], I think it is the defect of semantic inconsistency. But there is nothing defective whatsoever about

"If the match were struck, it would be very likely to light; but if the match were struck, it might not light."

There is no clash *at all* between 'would very probably' and 'might not'. On the contrary, 'would very probably' seems to *conversationally implicate* 'might not'. Our inner Gricean infers that if '... would very probably' is the strongest thing that can be said, '... might not' must be the case!

Leitgeb notes that "it is clear that might-conditionals should not be represented as the duals of the would-conditionals that obey our semantics" (111). He surmises that at best, one may turn to 'would' counterfactuals with possibility operators in their consequents. That is, 'if it were that p, it might be that q' has the form:

$$p\Box \rightarrow \Diamond q$$

But this seems to be a cost of his view, since again there is no clash *at all* between 'would' and 'would be possible that not', much as there is no clash at all between the claim that I am a philosopher and the claim that it is possible for me not to be a philosopher (at least for a non-epistemic sense of possibility, which is clearly operative in tandem with the ontic nature of objective chance).

Let us return to the putative near-synonymy of 'would' and 'would very probably'. Suppose a doctor truthfully tells your father: 'if you were to take this pill, you would very probably survive'; your father takes the pill and dies. Now suppose instead that the doctor had said: 'if you were to take this pill, you *would* survive'; your father takes the pill and dies. Your case for a law suit is much stronger in the latter case than in the former. This also shows how Leitgeb's semantics for counterfactuals allow modus ponens failures, which trouble me more than they appear to trouble him. 'Would' and 'would very probably' do not have almost the same meanings, because they license quite different inferences.

But even Leitgeb himself speaks of (3) being " ... "almost" synonymous with (1)" (37), with tell-tale scare quotes around "almost", as if he does not really *mean* "almost". (Perhaps "*almost* almost"?!) Indeed, by parallel reasoning,

(4) 'if p were the case, q would very probably be the case'

presumably means "almost" the same as

(5) 'if p were the case, it would very probably be the case that q would very probably be the case'.

And so on ad infinitum. This is difficult even to understand, unlike (4), and in any case a far cry from what (4) *means*. I suggest that Leitgeb's "almost" synonymy is an unreliable guide to truth conditions. This point will be important again later, when we come to his account of the "approximate truth" of counterfactuals. However, recalling that Leitgeb's project is one of explication rather than conceptual analysis, perhaps near enough is good enough for his purposes.

3.2.2. Questioning Steps 2 and 3

Is 'if p were the case, q would very probably be the case' to be analysed as '$ch(q|p)$ is very high'? Here are some reasons to think not.

On Leitgeb's proposal, many counterfactuals will correspond to conditional chances for which the condition has chance 0—notably, those whose antecedents never had a chance of being true. To be sure, Popper functions provide a mathematical framework for which such conditional chances may nonetheless be defined. But they may also be undefined. The Popper formalism comes with no guarantee that its conditional probability functions are endowed with a rich domain, still less a complete domain of all propositions that might appear as antecedents of counterfactuals. Moreover, whether or not the conditional *chance* function (at a time) is well defined for this or that condition is a *metaphysical* question, not one that can be settled by appealing to the axioms and theorems of the Popper calculus.

Indeed, there are good reasons for thinking that such conditional chances will *not* be defined for various counterfactuals. Consider counterfactuals about *the chances themselves* being different from what they actually are—e.g. 'if this (actually fair) coin had chance 1 of landing heads, the coin would repeatedly land heads'. It is quite unclear that the conditional chance of the consequent, given the antecedent, is well defined. Similarly, some counterfactuals do not explicitly mention chances, but tacitly involve them nonetheless—e.g. 'if the half-life of uranium were much longer, we would see more uranium around us', and perhaps even 'if the vase were less fragile, it would survive greater impacts than it actually does'. I doubt that such implicitly contrary-to-chance counterfactuals have truth conditions expressible in terms of the *actual* conditional chances.

And then there are counterfactuals for which chances just don't seem to be germane—for example, the very first counterfactual that Leitgeb discusses in his paper, 'if Oswald hadn't killed Kennedy, someone else would have'. There are reasons to think that there is no corresponding conditional chance for this counterfactual. The antecedent corresponds to a huge and miscellaneous disjunction of many ways for Oswald to fail to kill Kennedy: Oswald shot but only injured Kennedy; he shot and missed; he failed to shoot; he joined the Secret Service and became one of Kennedy's most dependable bodyguards; he never existed; no humans ever existed; ... Maybe each of these disjuncts corresponds to a chance process, with well-defined chances for its possible outcomes, and maybe not. (Actually, more likely *not*, since each disjunct is still highly disjunctive—e.g. there are so many disparate ways for Oswald never to have existed.) But in any case it is quite unclear that the huge *disjunction* corresponds to a chance process with well-defined chances for *its* possible outcomes. Furthermore, there is some controversy over whether macro-events such as these have chances in anything like the paradigmatic sense that quantum mechanical events do. Even Lewis-style 'best systems' analyses, which are friendly to chances being assigned to macro-events, may not accord chances, or conditional chances, to the requisite propositions—Oswald's shooting Kennedy or not is such a 'one-off' event that assigning a chance to it may not earn its keep in the best theory of the universe, costing too much in simplicity for too little gain in strength. And there is the further complication that these events involve free acts, which may not have chances at all for reasons arising from the metaphysics of freedom. (See [9] and [5].)

Or consider a counterfactual that one might well find in a physics textbook: 'if gravity hadn't obeyed an inverse square law, the planets would not follow elliptical orbits'. I see no reason to think that the corresponding conditional chance is high, or even defined.[15] After all, this counterfactual is a counterlegal, and the chances are intimately related to the laws of nature; God knows what they say about situations in which the laws are broken. Indeed, I wager that God knows that they don't say anything. The upshot is that there may be *chance gaps*—propositions that do not belong to the domain of the relevant conditional chance function—that nonetheless may appear in counterfactuals that I think Leitgeb would regard as true.[16]

While I laud Leitgeb's goal of naturalizing counterfactuals, I wonder whether he achieves it. I am not convinced that science supports his appeal to Popper functions, or their delivering conditional chances that align with the truth values of counterfactuals, in accordance with his semantics. Somewhat curiously, none of his examples of counterfactuals are drawn from science.

[15] I have similar concerns about Edgington's account.
[16] See my [9] and [10] for more arguments for and discussion of the existence of chance gaps.

3.3. Truth, or approximate truth?

To repeat, Leitgeb offers the *truth* condition for a counterfactual as the corresponding conditional chance being very high. It is important for me to emphasize this.

> In order for [‘if I had dropped the plate, it would have fallen to the floor’] to be **true** according to our probabilistic truth condition, it suffices for there to be a *very high* conditional probability of the plate falling to the floor if dropped. But that conditional probability *is* in fact high—exceptions to it having a probability close to 0—even by the lights of quantum physics. So the counterfactual comes out **true** in our semantics according to our ... interpretation of ‘$ch(B|A) = 1$’ in terms of ‘the probability of B given A is very high, or close to 1’... (111, my **bold.** I replace his notation for the conditional chance with my own.)

This exposes his account to the objections I raised in the Questioning Step 1 section. However, he tantalizes us with what looks like a hard-line truth condition that requires the conditional chance to be exactly 1: $ch(B|A) = 1$. I confess that I am puzzled by his interpretation of this as ‘the probability of B given A is very high, or close to 1’. To be sure, he justifies it with some formidable mathematical machinery (involving limits of infinite sequences of Popper functions). But for all that, ‘$ch(B|A) = 1$’ denotes an identity, not an approximate identity. Be that as it may, I think the best prospect for a probabilistic analysis of counterfactuals is the hard-line one.

That said, even the hard-line truth condition is not hard enough for my liking. Probability 1 is insufficient for truth. There are various *irregular* probability functions, which assign probability 1 to contingent propositions. And some of them assign probability 1 even to contingently *false* propositions. Moreover, the chance function (at any given time) is plausibly among them. The chance that a particular radium atom will decay at *exactly* time t is 0, for all t; yet when it decays, it does so at some exact time. The atom’s decaying at some *other* time has chance 1, but as it happens, this corresponds to a false proposition. Similarly, a conditional chance of 1 is insufficient for the truth of the corresponding counterfactual. The chance that a particular radium atom does not decay at time t_0, given that it is initially in a particular state S, is 1. But the counterfactual

> ‘if the atom were initially in state S, it would not decay at time t_0’

is false, or so I claim. After all, the counterfactual

> ‘if the atom were initially in state S, it *might* decay at time t_0’

is true. And I claim that this ‘might’ counterfactual is inconsistent with the ‘would not’ counterfactual, siding as I do with Lewis’s treatment of ‘might’ counterfactuals.

‘Might’ counterfactuals, and similarly ‘might not’ counterfactuals, are easily made true: the chances of their consequents, given their antecedents, do not even need to be positive. Correspondingly, I submit, ‘would’ counterfactuals are easily made false. This is one of several arguments of mine that most counterfactuals are false. I believe that Leitgeb would do better explicitly to endorse the hard-line truth condition, bringing him closer to my position. However, I admit that it is a cost of my position that it flies in the face of commonsense. For example, ‘If I had dropped the plate, it would have fallen to the floor’ is intuitively true, while I maintain that it is false.

3.4. Lotteries and plates

While Leitgeb seeks to give an explication of counterfactuals that is grounded in science, he does so in a way that recovers commonsense verdicts about them. Since the conditional chance of the plate falling to the floor, given I drop it, is very high, the counterfactual comes out true according to his truth condition. And so it goes with most ordinary counterfactuals that we take to be true. Nevertheless, he seems to equivocate on this when faced with another of my arguments. However high we set the bar for 'very high' (short of 1) in Leitgeb's truth condition, there will be a lottery with sufficiently many tickets to create a problem.

The problem turns on the following 'agglomeration' inference rule, common to all of the leading logics of counterfactuals:

$$\text{(Agglomeration)} \qquad \begin{array}{l} p\Box\rightarrow q \\ \underline{p\Box\rightarrow r} \\ \therefore\ p\Box\rightarrow (q\,\&\,r). \end{array}$$

Suppose, for example, that we set the bar at 0.999999. Consider a lottery with 1,000,001 tickets that in fact is never played. Then all of the following counterfactuals come out true, for their corresponding conditional probabilities clear the bar:

Lottery is played $\Box\rightarrow$ ticket 1 loses.
Lottery is played $\Box\rightarrow$ ticket 2 loses
...
Lottery is played $\Box\rightarrow$ ticket 1,000,001 loses.

Then by Agglomeration, we infer:

Lottery is played $\Box\rightarrow$ (ticket 1 loses & ticket 2 loses & ... & ticket 1,000,001 loses).

Given the set-up, this is equivalent to

Lottery is played $\Box\rightarrow$ each ticket loses.

But we also assume that

Lottery is played $\Box\rightarrow$ some ticket wins.

By Agglomeration again, we conclude

Lottery is played $\Box\rightarrow$ each ticket loses & some ticket wins.

The antecedent is contingent, but the consequent is a contradiction, so this conclusion is a contradiction. The mistake, I claim, is to regard each of the 'ticket i loses' counterfactuals to be true. But by symmetry, they all have the same truth value. So they are all false, *contra* the Leitgeb truth condition, with the bar for 'very high' set at 0.999999. And so it goes wherever we set the bar, short of 1. Moreover, most ordinary counterfactuals involve natural 'lotteries', as they are governed by chance processes. (If some are not, as I suggested in Section 3.2.2 may be the case with the 'Oswald' counterfactual, then Leitgeb's account runs aground on them immediately.) Even a dropped plate is not certain to fall; at best, it does so with very high chance. In fact, whatever that chance is, we can easily imagine a lottery with so many tickets that the chance of any given ticket losing is higher than *that*.

Another way to respond is to deny Agglomeration. If Leitgeb sets the bar any lower than 1, then he must do this. So, too, if he allows the bar to be vague—as we might say, 1-*ish*. For we may imagine a lottery so large that each ticket's chance of losing is 1-ish, but the chance of all the tickets losing is 0, and hence not 1-ish—for any reasonable tolerance of 'ish'. Each premise counterfactual is then adjudicated as true, but the agglomerated conclusion counterfactual is adjudicated as false. Either way, there is a significant cost, since

Agglomeration is so intuitive. To uphold it on Leitgeb's approach here,[17] the bar must be set *at* 1—and by that, I do *not* mean "very high, or close to 1". In fact, I think this even gives the intuitively correct verdicts for the individual lottery-ticket counterfactuals. Intuitively, it *is* false that ticket i would lose. What's true is that ticket i would *probably* lose, but that is something else.

We thus come to a key move of Leitgeb's, and we come full circle back to the difference between a proposition being true, and it being *probably* true, as I promised that we would. Recall Leitgeb's claim of the near-synonymy of 'would' and 'would very probably', which I disputed. Previously, he regarded the truth condition for 'if it were that p, it would be that q' to be '$ch(q|p)$ is very high', where this is understood to be above a vaguely determined threshold. In response to my lottery argument, he moves to regarding $ch(q|p)$ above an "exact threshold" to give the "semantics for approximate truth" (113). On this reading, each of the 'individual ticket' premises of the lottery argument is *approximately* true, but agglomeration fails, so the argument does not go through. Likewise, I take it that on this reading, 'if I had dropped the plate, it would have fallen to the floor' is *approximately* true—though strictly speaking false. Leitgeb takes his cue from science, which is full of such statements.

I would be delighted if this worked: my uncompromising view about most counterfactuals would be vindicated, while commonsense would be mostly vindicated too—the *truth* of counterfactuals is exceedingly hard to come by, but approximate truth is rather easier. Indeed, for years I said much the same thing myself.

However, I eventually came to have misgivings with this way of speaking. In the end I preferred a cagier way of putting the point, which Leitgeb quotes (on p. 113): "In the neighborhood of the ordinary but false counterfactuals that we utter, there are closely related counterfactuals that are true but not ordinary. They are counterfactuals with appropriate probabilistic or vague consequents." This is not to say that the false counterfactuals that we utter are *approximately true*, thanks to the truth of the related counterfactuals with appropriate probabilistic consequents. (Vague consequents are not relevant here.) *A high chance of p is not the same thing as p being approximately true.* Likewise, a high conditional chance of q, given p is not the same thing as 'if it were that p, it would be that q' being approximately true.

For think of a case of approximate truth, taken from science. The speed of light is approximately 3×10^8 m/s. Call its exact speed c. These two speeds are similar, and because of this, 'the speed of light is 3×10^8 m/s' is *approximately true*. A world in which light has a speed of (exactly) 3×10^8 m/s is ipso facto similar to our world in which it has speed c, with respect to light's speed. But a world in which a coin lands heads with chance 0.99 is not ipso facto similar to a world in which it is *true* that the coin lands heads. The worlds need not be similar with respect to the coin's chance of heads: the latter world might have a *low* chance of heads, but it turns out to be true nonetheless. Nor need the worlds be similar with respect to the outcome: suppose the coin lands tails in the former world and heads in the latter. Similarly, a world in which the coin has a high conditional chance of landing heads, given it is tossed, is not ipso facto similar to a world in which it is *true* that the coin would land heads if it were tossed.

A way of thinking about this is that *probability crosscuts similarity*. As it turns out, Leitgeb himself comes close to making this point when discussing "the differences between Lewis' [similarity-based] semantics and [his own] probabilistic semantics ... after all, degrees of similarity are not supposed to correspond to magnitudes of probability in any straightforward manner" (55).

4. Conclusion

At least this allows me to end on a happy note of rapprochement. In giving their probabilistic semantics for counterfactuals, Edgington and Leitgeb must disavow similarity-based semantics for counterfactuals, à

[17] In his more recent work [18] (forthcoming), Leitgeb suggests another conditional-chance semantics for counterfactuals, which I think is better in this regard. According to it, the truth of a counterfactual is given by its corresponding conditional chance being high, the appropriate threshold given by a real number, and yet agglomeration is valid. He achieves this by adopting a strong context-sensitivity for the truth values of counterfactuals.

la Stalnaker and Lewis. I have long argued against such similarity accounts. Worlds in which the plate is dropped and falls to the floor may well be more similar to ours than worlds in which it is dropped and does something else. But that doesn't make it true that if the plate were dropped, it would fall to the floor. That counterfactual is undermined by the fact that if the plate were dropped, it would have a positive chance of *not* falling to the floor. (See my MS [11].) This chance is indifferent to how similar is a world where this happens.[18]

The irrelevance of similarity of worlds to the truth conditions of counterfactuals, in turn, inclines me towards regarding them as *strict conditionals*. Other arguments that I offer in my (MS) triangulate to the same conclusion. No wonder so many counterfactuals are false—strict conditionals have demanding truth conditions! But I must leave a proper discussion of the strict conditional account for another occasion.

Leitgeb's project of naturalizing counterfactuals sits well with our opposition to similarity accounts. Science has no truck with a notion of similarity; nor does Lewis's [21] ordering of what matters to similarity have a basis in science.[19] Like Edgington and Leitgeb, I see a tight connection between chance—which *is* grounded in science—and counterfactuals. However, some counterfactuals have the truth values they do without such a connection (recall my various examples in Section 3.2.2). And the undermining connection that *I* see drives me to my radical conclusion of their widespread falsehood, where of course we part company. Leitgeb is more moderate regarding their truth values, while Edgington advocates their widespread truth-valuelessness. But at least she and I agree that most counterfactuals are *not true*!

References

[1] Ernest Adams, The Logic of Conditionals, Reidel, Dordrecht, 1975.
[2] Ernest Adams, A Primer of Probability Logic, CSLI, Stanford University, Stanford California, 1998.
[3] Jonathan Bennett, Conditionals, Oxford University Press, Oxford, 2003.
[4] Richard Bradley, Indicative conditionals, Erkenntnis 56 (3) (2002) 345–378.
[5] Lara Buchak, Why the rollback argument fails, Philos. Q. 63 (250) (2013) 20–28.
[6] Dorothy Edgington, Counterfactuals, in: Proceedings of the Aristotelian Society CVIII, Part I, 2008, pp. 1–21.
[7] Hartry Field, Logic, meaning, and conceptual role, J. Philos. 74 (1977) 379–409.
[8] Allan Gibbard, Two recent theories of conditionals, in [12].
[9] Alan Hájek, What conditional probability could not be, Synthese 137 (3) (December 2003) 273–323.
[10] Alan Hájek, Arguments for – or Against – probabilism?, Br. J. Philos. Sci. 59 (December 2008) 793–819, reprinted in: Franz Huber, Christoph Schmidt-Petri (Eds.), Degrees of Belief, Springer, 2009, pp. 229–251.
[11] Alan Hájek (MS), Most counterfactuals are false.
[12] W.L. Harper, R. Stalnaker, G. Pearce (Eds.), Ifs, Reidel, Dordrecht, 1981.
[13] William L. Harper, A conditional belief semantics for free quantification logic with identity, in: H. Leblanc, R.D. Gumb, R. Stern (Eds.), Essays in Epistemology and Semantics, Haven Publications, New York, 1983, pp. 79–94.
[14] James M. Joyce, A non-pragmatic vindication of probabilism, Philos. Sci. 65 (1998) 575–603.
[15] H. Leblanc, Probabilistic semantics for first-order logic, Z. Math. Log. Grundl. Math. 25 (1979) 497–509.
[16] Hannes Leitgeb, A probabilistic semantics for counterfactuals. Part A, Rev. Symb. Log. 5 (1) (2012) 16–84.
[17] Hannes Leitgeb, A probabilistic semantics for counterfactuals. Part B, Rev. Symb. Log. 5 (1) (2012) 85–121.
[18] Hannes Leitgeb, A lottery paradox for counterfactuals without agglomeration, Philos. Phenomenolog. Res., forthcoming.
[19] David Lewis, Counterfactuals, Blackwell and Harvard University Press, 1973.
[20] David Lewis, Probabilities of conditionals and conditional probabilities, Philos. Rev. 85 (3) (1976) 297–315.
[21] David Lewis, Counterfactual dependence and time's arrow, Noûs 13 (1979) 455–476.

[18] The probabilistic cornerstone of Edgington's suppositional account of counterfactuals traces back to [28], an early proponent of the thesis that the probability of a conditional equals the corresponding conditional probability. This has been called *Stalnaker's Hypothesis*:

$$P(A \to B) = P(B|A)$$

for a suitably large class of P, and for all A and B in the domain of P for which these quantities are defined. To be sure, he later [29] disavowed and even argued against the thesis. He was *also* a pioneer of similarity or 'closeness' accounts of counterfactuals. But the left-hand side of this equation, by Stalnaker's lights, is sensitive to the closeness ordering among worlds, while the right-hand side is blind to it. This, in turn, furnishes a new argument against Stalnaker's Hypothesis: changes in the similarity relation (e.g. due to a context shift) may change the left-hand side of the equation, but not the right-hand side. Or driving the argument the other way, as Edgington surely would: Stalnaker's Hypothesis furnishes a new argument against similarity accounts of counterfactuals.

[19] Thanks here to Hannes Leitgeb.

[22] Vann McGee, Conditional probabilities and compounds of conditionals, Philos. Rev. 98 (4) (1989) 485–541.
[23] Karl Popper, The Logic of Scientific Discovery, Basic Books, New York, 1959.
[24] F.P. Ramsey, Truth and probability, in: R.B. Braithwaite (Ed.), Foundations of Mathematics and Other Essays, Routledge & P. Kegan, 1931, reprinted in: D.H. Mellor (Ed.), Philosophical Papers, Cambridge University Press, Cambridge, 1990.
[25] Brian Skyrms, Higher order degrees of belief, in: D.H. Mellor (Ed.), Prospects for Pragmatism, Cambridge University Press, Cambridge, 1980, pp. 109–137.
[26] Brian Skyrms, Pragmatics and Empiricism, Yale University Press, New Haven, 1984.
[27] Robert Stalnaker, A Theory of Conditionals, Studies in Logical Theory, American Philosophical Quarterly Monograph Series, vol. 2, Blackwell, Oxford, 1968.
[28] Robert Stalnaker, Probability and conditionals, Philos. Sci. 37 (1970) 64–80, reprinted in [12].
[29] Robert Stalnaker, Letter to van Fraassen, in: Harper, Hooker (Eds.), Foundations of Probability Theory, Statistical Inference and Statistical Theories of Science, vol. I, Reidel, 1976, pp. 302–306.
[30] Bas van Fraassen, Belief and the Will, J. Philos. 81 (1984) 235–256.