

STEFAN KAUFMANN\*

## CONDITIONAL PREDICTIONS

*A probabilistic account*

**ABSTRACT.** The connection between the probabilities of conditionals and the corresponding conditional probabilities has long been explored in the philosophical literature, but its implementation faces both technical obstacles and objections on empirical grounds. In this paper I first outline the motivation for the probabilistic turn and Lewis' *triviality* results, which stand in the way of what would seem to be its most straightforward implementation. I then focus on Richard Jeffrey's 'random-variable' approach, which circumvents these problems by giving up the notion that conditionals denote propositions in the usual sense. Even so, however, the random-variable approach makes counterintuitive predictions in simple cases of embedded conditionals. I propose to address this problem by enriching the model with an explicit representation of causal dependencies. The addition of such causal information not only remedies the shortcomings of Jeffrey's conditional, but also opens up the possibility of a unified probabilistic account of indicative and counterfactual conditionals.

**KEY WORDS:** conditionals, counterfactuals, probability, causality

## INTRODUCTION

Most current theories of conditionals are inspired by Ramsey's (1929) paraphrase of the process involved in their evaluation:

- (RT) If two people are arguing 'If  $p$  will  $q$ ?' and are both in doubt as to  $p$ , they are adding  $p$  hypothetically to their stock of knowledge and arguing on that basis about  $q$  . . . We can say they are fixing their degrees of belief in  $q$  given  $p$ .

This suggestion can be made precise in a number of ways, depending on what is assumed about beliefs and belief update. One often-made pair of assumptions is the following:

- (i) Degrees of belief are measured by probabilities.
- (ii) The hypothetical addition of the antecedent proceeds by conditionalization.

---

\* Author for correspondence: E-mail: kaufmann@northwestern.edu

With these assumptions, the Ramsey Test comes down to the following doctrine, which I will refer to as “the Thesis”:<sup>1</sup>

- (T) The probability of a conditional ‘*if A then C*’ is the conditional probability of *C*, given *A*.

All of these notions will be explained below. As we will see, (T) as stated here requires further elaboration regarding the interpretation of the probability measure and the exact role of the conditional probability. I am going to propose a detailed unified account of simple and right-nested *predictive* conditionals and their counterfactual counterparts. Further extensions, such as an application to epistemic indicatives (see below) and conditionals with conditional antecedents, are possible but left for future occasions.

The probabilistic approach, its motivation, its merits and its specific problems with respect to conditionals have all been studied extensively, but this work went largely unnoticed in the linguistic literature (but see Cohen, 2003 for some discussion). Part of this paper is therefore devoted to introducing these various strands that come together in the proposal. In Section 1, I explain my choice to deal with one class of indicative conditionals at the exclusion of others. Some of the merits of the probabilistic account and its relationship to the standard quantificational framework are discussed in Section 2. In Section 3, I introduce basic definitions and discuss the relationship between truth and probability for truth-functional sentences. I then show why the same strategy is not available for conditionals if (T) is to hold in general, and how this problem is avoided by making the interpretation of the conditional dependent upon the probability distribution. This move is motivated and formalized in Section 4 on truth and chance in time. In Sections 5 and 6, I examine the predictions that this account makes about the probabilities of right-nested indicative conditionals and counterfactuals, and argue that reference to *causal dependencies* is required to correct certain counterintuitive consequences. The effect of such causal information in the interpretation of the examples is discussed in Section 7. The paper ends with an outlook on some further issues in Section 8.

---

<sup>1</sup> Also known as “Stalnaker’s Thesis” after Stalnaker (1970) (Stalnaker, 1976, disavowed it); “Adams’ Thesis” after Adams (1965, 1975); or the “Conditional Construal of Conditional Probability (CCCP)” following Hájek and Hall (1994) and Hall (1994).

## 1. CLASSIFICATION

Within the landscape of conditionals, all three of (1a–c) are *direct* and *specific* in the terms of Quirk et al. (1985). I will focus on *predictive* conditionals and their counterfactual counterparts, illustrated in (1a,b). I will not deal with the *epistemic* conditional in (1c).

- (1) a. If you strike the match, it will light. [*predictive*]  
       b. If you had struck the match, it would have lit. [*counterfactual*]  
       c. If you struck the match, it lit. [*epistemic*]

The decision to restrict attention in this way is not motivated by any assumption of a fundamental semantic difference between these classes.<sup>2</sup> On the contrary, I believe that the treatment I propose here for predictive conditionals is just as applicable, *mutatis mutandis*, to epistemic ones. What differs is the interpretation of the probability measure, roughly corresponding to the difference between various accessibility relations, or “modal bases”, in non-probabilistic approaches. In this paper I will focus on objective or “metaphysical” readings and the attendant interpretation of probability as objective chance. It is this limitation which places epistemic conditionals beyond the scope of the paper, since an adequate account of the latter must involve reference to epistemic uncertainty.

Even granting that there are good reasons for limiting coverage in the interest of exposition, however, my restricting it to what under some analyses is not considered a natural class of conditionals deserves some explanation. In the remainder of this section, I will give some relevant arguments.

1.1. *Indicative vs. Counterfactual*

According to one classification, (1a,c) are *indicative* and (1b) is *counterfactual*. *Subjunctive* is sometimes used instead of the latter and usually meant to be coextensive with it.<sup>3</sup> A similar line is drawn between “open” and “hypothetical” conditionals by Quirk et al. (1985).

This classification has wide currency despite the well-known fact that both “counterfactual” and “subjunctive” are misnomers. Anderson’s

<sup>2</sup> Such a fundamental difference has been postulated in connection with otherwise somewhat similar proposals; see for instance Gibbard’s (1981) distinction between “nearness” and “epistemic” readings.

<sup>3</sup> Opinions vary on sentences like *If you were to strike the match, it would light*, which Lewis (1973) claims are subjunctives that “appear to have the truth conditions of indicative conditionals” (p. 4).

(1951) example (2) shows that counterfactuals can be used to argue *for* the truth of the antecedent (see also, Stalnaker 1975; Adams 1976; Karttunen and Peters 1979; Barwise 1986; Comrie 1986; von Stechow 1998).

- (2) If he had taken arsenic, he would have shown just these symptoms. *[those which he in fact shows]*

“Subjunctive”, as a morphological category, is useless in drawing the distinction. The use of the subjunctive mood in conditionals is optional, restricted to formal style, and possible in sentences of both types (Quirk et al. 1985; Dudman 1988; Bennett 1988; Edgington 1995). Thus (3a) is a subjunctive non-counterfactual, whereas (3c) (unlike 3b) is an indicative counterfactual.<sup>4</sup>

- (3) a. If any person be found guilty, he shall have the right to appeal.  
 b. If I were rich, I would buy you anything you wanted.  
 c. If I was rich, I would buy you anything you wanted.

More pertinent than these terminological issues, however, is the question of how deep a *semantic* division ought to be drawn between the two classes. The minimal pair in (4), due to Adams (1970), is often cited in this connection.

- (4) a. If Oswald did not kill Kennedy, someone else did.  
 b. If Oswald had not killed Kennedy, someone else would have.

From the observation that “[4a] is probably true while [4b] may very well be false”, Lewis (1973) concludes that “there really are two different sorts of conditional; not a single conditional that can appear as indicative or as counterfactual depending on the speaker’s opinion about the truth of the antecedent” (p. 3).

Not everyone agrees. Strawson (1986) presents a pair like the following (using different sentences):

- (5) a. Remark made on November 21, 1963:  
 “If Oswald does not kill Kennedy, someone else will”.  
 b. Remark made on November 23, 1963:  
 “If Oswald had not killed Kennedy, someone else would have.”

<sup>4</sup> One reviewer questioned the grammaticality of (57c), but many informants do find it acceptable.

In Strawson's opinion, "[i]t seems obvious that about the least attractive thing one could say about the *difference* between these two remarks is that it shows that, or even that it is partly accounted for by the fact that, the expression 'if... then...' has a different meaning in one remark from the meaning which it has in the other" (p. 230). Edgington (1995) agrees and suggests that the difference may be "more like the difference between mature cheddar and freshly-made cheddar than the difference between chalk and cheese" (p. 239).

Notice that (4a) is an epistemic conditional whereas (5a) is predictive. It is perhaps no coincidence that the pairs in (4a,b) and (5a,b) have been used to argue against and for a unified account, respectively. Both of 4(a,b) are conditional *predictions* made from the perspective of different times, referring not merely to (non-actual) *facts* but to alternative *courses of events*, which depart from the actual history at a point in time prior to that referred to by the antecedent.<sup>5</sup> This opinion is held by many authors and implemented in a variety of accounts (Downing, 1959; Adams, 1975; Ellis, 1978; Thomason and Gupta, 1981; Tedeschi, 1981; Dudman, 1994; Strawson, 1986; Bennett, 1988; Mellor, 1993; Edgington, 1995; Dahl, 1997; and others).

Aside from the intuitions that motivate these unified accounts, the close semantic affinity between counterfactuals and indicatives is demonstrated, perhaps even more forcefully, by their similar behavior in inference. In particular, the familiar inadequacies of the material and strict conditionals – the fact that they validate inference patterns for which there are linguistic counterexamples, which I discuss in Section 2.1 – can be shown equally well with examples from either class (Veltman, 1985). But if inference patterns are to be the semanticist's primary source of empirical data, then the claim that this parallelism arises in spite of a deep semantic difference is hard to substantiate.

## 1.2. *Predictive vs. Non-Predictive*

Another tradition draws the major dividing line between predictive and counterfactual conditionals on the one hand, and epistemic ones, on the other (Dancygier, 1998). Dudman (1984, 1994 and elsewhere) motivates this decision with the observation that while the antecedent of (1c) retains its interpretation when used in isolation as in (6b), that of (1a), when used

<sup>5</sup> The fact that the departure from the actual course of events lies in the past from the perspective of the reference time of the antecedent has been exploited by Dudman, Dahl and others to give an explanation of *backshift*, the addition of a layer of past morphology in the antecedents of counterfactuals. For alternative views, see James (1982), Fleischman (1989), Iatridou (2000).

as in (6a), is only felicitous under a “scheduling” interpretation (cf. also Palmer, 1983; Comrie, 1985).

- (6) a. You strike the match.
- b. You struck the match.

Funk (1985) arrives at a similar distinction on different grounds, relating the difference between (1a) and (1c) to one in the status of the facts that determine the truth or falsehood of the antecedent: They are not yet “manifested” in the case of (1a), and “manifested” but unknown in the case of (1c). In Funk’s words, “the meaning of the conditioning frame can be said to vary from ‘if it happens that ...’ to ‘if it is true that ...’” (p. 376). He goes on to discuss examples from languages in which the distinction has clearer morphological reflexes than it does in English.

The relevant difference, as I see it, is directly related to the reasons given above for excluding epistemic conditionals from the scope of this paper. Indicative conditionals are generally used under uncertainty as to the truth of the antecedent. The distinction Funk appeals to concerns the nature of that uncertainty. In the epistemic conditional (1c), it is purely *subjective*: The speaker is ignorant of the relevant facts, but objectively, those facts are settled. In the case of the predictive conditional (1a), the world itself may yet follow different future courses, and whether the addressee ends up striking the match cannot be determined until the facts are in. The uncertainty is *objective*.

Objective uncertainty implies subjective uncertainty under the reasonable assumption that what is objectively not yet determined cannot already be known. On the other hand, many settled facts, for instance whether the match was struck at the time in question, remain unknown to most epistemic subjects.

This difference has clearly felt consequences for the question of whether there is something objectively “correct” to be believed about the conditional. It is appropriate to say that (7a) is true or false depending on whether the match is dry or wet; likewise in retrospect for the corresponding counterfactual (7b).

- (7) a. If the match is struck, it will light.
- b. If the match had been struck, it would have lit.
- c. If the match was struck, it lit.

Things are not so clear for the epistemic conditional (7c). If the match was not struck, there does not seem to be anything objectively ‘correct’ to

be believed about it. Nor do other facts, such as whether the match was dry or not, have any bearing on the objective value of (7c). However, a speaker who knows that the match was wet but does not know whether it was struck may still consistently believe that the conditional is false. More generally, Gibbard (1981) shows that when the antecedent is false, it is possible for two speakers with different sets of correct but partial beliefs to arrive at diametrically opposed conclusions about the conditional, neither of which is proven wrong by the facts (although from the fact that they disagree one can infer that the antecedent is false). Edgington (1995) shows that this possibility arises not only with past reference, but whenever the antecedent at speech time has no (objective) chance of being true.

This suggests that in a unified account of conditionals, the two kinds of uncertainty should be kept separate and some care must be taken in setting up the relationship between them. In focusing on predictive conditionals, I avoid the complexities this would require. Thus I will make two assumptions throughout the paper: (i) Probabilities are interpreted as objective chance, and (ii) the antecedents of the conditionals discussed have some chance of turning out true.

## 2. WHY PROBABILITY?

The idea that the analysis of conditionals would benefit from a probabilistic perspective has a long history. Ramsey's original proposal can be, and often has been read as suggesting that it is what he had in mind. Later, (RT) was taken up and developed in a more explicitly probabilistic setting by Jeffrey (1964), Adams (1965), Stalnaker (1970), among others, and has since inspired a vast and varied field of inquiry (for recent overviews, see Edgington, 1995, and the papers in Eells and Skyrms, 1994; also Adams, 1998). The purpose of this paper is not to reiterate those points, but to propose a particular implementation which avoids certain technical problems and dubious predictions. Nevertheless, in this section I will mention some reasons for exploring this direction in the first place.

### 2.1. *Classical Logic*

Minimal assumptions imply that if '*if... then ...*' is to be represented by a truth-functional connective alongside conjunction, disjunction and negation, then that connective has to be the material conditional (Gibbard 1981; Edgington 1986). However, the material conditional is hardly a suitable translation of '*if... then ...*' familiar from everyday uses of English. It is true "too easily": It follows from premises from which one would not in-

fer the corresponding ‘*if . . . then . . .*’-sentence. Instances of such inference patterns are often used to illustrate a parallel problem with counterfactual conditionals, which however arises with indicatives as well. Following are examples of such violations of Strengthening of the Antecedent (8), Contraposition (9), Vacuous truth (10), and Hypothetical Syllogism (11).<sup>6</sup>

- (8) a. “If I install a better alarm system”, Griliches said, “that is an improvement in the quality of my life, and therefore a decline in inflation.  
 b. But if the burglars learn how to trick this alarm system, that is a rise in price, because the quality advantage will be eroded.”  
 nyt961217.0474
- (9) a. If you’re a high achiever, it takes a long time to get recognition.”  
 nyt961130.0128  
 b. ??If it takes a short time to get recognition, you’re not a high achiever.
- (10) a. You won’t strike the match.  
 b. ??Therefore, if you strike the match, it will/won’t light.
- (11) a. If I quit my job, I won’t be able to afford my apartment.  
 b. If I win a Million, I’ll quit my job.  
 c. ??If I win a Million, I won’t be able to afford my apartment.

Adams (1975) noted that (12) is not as clear a counterexample to Hypothetical Syllogism as (11). Depending on the order in which the premises are given, it is harder to even ponder the case that both are true. Adams attributes this to a tendency to interpret (12b) in the given context as (12c).

- (12) a. If I win a Million, I’ll quit my job.  
 b. ??If I quit my job, I won’t be able to afford my apartment.  
 c. ??If I *win a Million and* quit my job, I won’t be able to afford my apartment.

This phenomenon falls under *modal subordination* in the linguistic literature (Roberts, 1989). With the modally subordinated reading of the

<sup>6</sup> Examples with labels of the form ‘nyt—.’ are attested in the New York Times Corpus. The numbers refer to their locations in the corpus.



second premise, the example instantiates, as a non-counterexample, a related pattern which does seem to be valid not only with the material conditional, but also with its natural-language counterpart, *viz.* the inference from ‘*if A then B*’ and ‘*if AB then C*’ to ‘*if A then C*’.

Aside from this case, the above inference patterns demonstrate that conditionals are not well represented by the material conditional. Since the latter is the only truth function that could plausibly represent them, it follows that conditionals are not truth-functional.

An alternative is the *strict* conditional, first discussed by C.I. Lewis,<sup>7</sup> which is true just in case the corresponding material conditional is true at all possible worlds. This rules out certain cases of truth “by accident”, such as arise with conditionals both of whose constituents happen to have the same truth value. Nor is the falsehood of the antecedent sufficient for the truth of the strict conditional. Aside from these merits, however, the strict conditional is afflicted by many of the same problems as the material conditional. Contradictory antecedents still lead to vacuous truth; Strengthening of the Antecedent, Contraposition and Hypothetical Syllogism are valid; at the same time, the strict conditional is too strong in stating in effect that the antecedent *entails* the consequent.

A somewhat more well-behaved variant that has taken hold in the linguistic literature is the *variably* strict conditional. Here the domain of the quantifier is restricted by a contextually given parameter, representing the speaker’s beliefs or some other background that the conditional is evaluated against. This contextual restriction in effect opens up a continuum of readings with the material conditional at one extreme and strict implication at the other, according as the domain of quantification is the singleton set containing the world of evaluation, or the set of all worlds, or something in between. Thus the truth of the conditional no longer implies that the antecedent entails the consequent. However, since under this approach the conditional is true if and only if the corresponding material conditional is true at all worlds in the contextually given domain, the inference patterns that were problematic for both material and strict conditional continue to be so.

## 2.2. *Similarity*

The invalidity of the classically valid inference patterns illustrated above with indicative conditionals carries over to counterfactuals; this is easy to check by substituting the corresponding forms. Counterfactuals can be, and typically are used under the assumption that the antecedent is false. Lest

---

<sup>7</sup> See Hughes and Cresswell (1996) for a historical discussion.

they be predicted vacuously true in all such cases, the speaker's belief state is not suitable here to provide the domain of the variably strict conditional.

Instead, in the accounts of Stalnaker (1968) and Lewis (1973), the domain is set up with the aid of a relation of *similarity* between worlds: (13a) is true despite the fact that there are possible worlds where the floor is covered with layers of blankets, for those worlds are less similar to ours than those at which, as in ours, it is made of concrete. Thus it is possible for (13a) to be true while (13b) is false.<sup>8</sup>

- (13) a. If I had dropped the vase, it would have broken.  
 b. If I had dropped the vase and the floor had been covered with layers of blankets, the vase would have broken.

Stalnaker's and Lewis' theories differ in the details of the similarity relation, but both agree that each world is most similar to itself, with the consequence that at worlds at which the antecedent is true, the conditional is equivalent to its consequent. In the general case, according to Lewis a counterfactual '*If A, would have been C*' is true at  $w$  just in case there is no  $A\bar{C}$ -world that is not surpassed in similarity to  $w$  by an  $AC$ -world. According to Stalnaker, the conditional is true at  $w$  if and only if  $C$  is true at the nearest  $A$ -world.

The similarity-based approach correctly invalidates the problematic inference patterns for counterfactuals, but not for indicatives. For the latter, Stalnaker's account comes down to the simpler quantificational account discussed earlier.<sup>9</sup>

However, the failure to get the inference patterns right should not be held against the idea that the interpretation of conditionals involves *some* quantifier. It is the universal force of the quantification that re-introduces the problem. Thus the question arises whether a more suitable quantificational account can be found.

<sup>8</sup> Restricting the accessibility of worlds in this way corresponds to adding to the antecedent a set of true sentences that are consistent with it. In this form, the idea goes back to Goodman (1947) and is formalized in *premise semantics* (Veltman 1976, 1985; also Kratzer 1979, 1991). For a comparison, see Lewis (1981).

<sup>9</sup> Stalnaker (1975) assumes that the selection of nearest  $A$ -worlds is restricted to those in the contextually given domain if there are any. If there are, each  $A$ -world in the domain selects itself, and each non- $A$ -world in the domain selects an  $A$ -world in the domain. The conditional is then true iff the material conditional is true at all worlds in the domain. If there are no  $A$ -worlds in the domain, the selection of alternatives is allowed to "reach" outside of it. This relaxation, Stalnaker argues, is the essential semantic correlate of the difference between indicative and counterfactual conditionals.

Lewis did not believe in a unified account. Instead, he followed Jackson (1979, 1987) in saying that indicatives have the truth conditions of the material conditional, combined with pragmatic "assertibility" conditions of the probabilistic kind discussed below.

2.3. *Conditional Probability: The Thesis*

To see how a probabilistic analysis in terms of the Thesis (T) addresses the problems raised above, it is helpful to consider how it is related to the quantificational account. There the truth of a conditional ‘*if A then C*’ relative to a set of worlds depends on whether or not all *A*-worlds in that set are *C*-worlds. The conditional is true if this is the case and false if there are any  $A\bar{C}$ -worlds. In a probabilistic framework, the question is not whether or not, but *to what extent* the conditional is supported. Roughly speaking, this support is high if *AC* is almost as likely as *A*, or much more likely than  $A\bar{C}$ .

The first argument for this notion comes from simple intuitions about beliefs under uncertainty. What do speakers believe about a conditional ‘*if A then C*’ if they are not entirely sure that  $A\bar{C}$  can be ruled out? Van Fraassen (1976) noted:

[T]he English statement of a conditional probability sounds exactly like that of the probability of a conditional. What is the probability that I throw a six if I throw an even number, if not the probability that: if I throw an even number, it will be a six? (pp. 272–273)

Most agree that the Thesis “sounds right”, or at least more plausible than the logical alternative. Aside from this ‘gut-feeling’, however, a more convincing argument in favor of the probabilistic account comes from the use of conditionals in inference. Most of the relevant work in this area was carried out by Ernest Adams (1965, 1975, 1998). The two basic assumptions are the following:

- (14) a. The Thesis (T): The probability of a conditional is the corresponding conditional probability;
- b. What is preserved in everyday reasoning is not truth, but (high) probability.

Adams devised a formal system of probability-preserving inference based on (14a,b) and centered around the following notion of validity:

- (15) An inference is probabilistically valid (*p-valid*) if and only if it is impossible for the premises to be highly likely while the conclusion is not highly likely.<sup>10</sup>

---

<sup>10</sup> The formal definition is as follows: A sentence  $\varphi$  follows from a set of sentences  $\Gamma$  if and only if for all  $\epsilon > 0$  there is a  $\delta > 0$  such that for all conditional probability functions  $\Pr$  such that  $\Pr(B) \geq 1 - \delta$  for all  $B \in \Gamma$ ,  $\Pr(\varphi) \geq 1 - \epsilon$ .

In this system,  $p$ -valid inferences are also classically valid, but some classically valid ones are not  $p$ -valid. In particular, the following inferences with conditional conclusions are not  $p$ -valid (writing  $A \rightarrow C$  for the probabilistic conditional):

1. Strengthening of the Antecedent:  
 $A \rightarrow C$  does not  $p$ -entail  $AB \rightarrow C$ .
2. Contraposition:  
 $A \rightarrow C$  does not  $p$ -entail  $\bar{C} \rightarrow \bar{A}$ .
3. Vacuous Truth:  
 $\bar{A}$  does not  $p$ -entail  $A \rightarrow C$ .
4. Hypothetical Syllogism:  
 $A \rightarrow B$  and  $B \rightarrow C$  do not  $p$ -entail  $A \rightarrow C$ . *However:*  $A \rightarrow B$  and  $AB \rightarrow C$  do jointly  $p$ -entail  $A \rightarrow C$ .

A detailed introduction to these results is found in Adams (1998). I end this brief overview by concluding that a theory of conditionals in which (T) plays a central role is a promising candidate for an interpretation of everyday uses of conditionals. What remains to be seen is what exactly that role should be.

#### 2.4. *Interim Summary*

Two recurring themes in quantificational accounts of conditionals are the *force* of the quantification and its *domain*. Between the two main ingredients of this paper, the labor is divided accordingly: The use of probabilities addresses the former, the appeal to causal relations the latter. The case for adopting a probabilistic perspective was outlined above. The details will be developed in Sections 3 and 4.

In Section 5, I turn to arguments for making the interpretation sensitive to causal relations. Among the accounts discussed above, this solution is closest to the *set selection* approach, a special case of the Stalnaker/Lewis similarity-based framework (Lewis 1973 attributed the idea to John Vickers and Peter Woodruff) built on the assumption that there is always a non-empty, but not necessarily singleton set of closest antecedent worlds. The use I propose to make of causal relations can be thought of as defining the analog of a set selection function – one which is not based on similarity, but on what I consider the less obscure notion of causal independence.

Similarity between worlds has also been put to use in probabilistic accounts, where it gives rise to the update operation of *imaging*, as opposed to

conditionalization (Stalnaker 1970; Lewis 1976; Gärdenfors 1982, 1988). Skyrms (1984, 1988, 1994) made extensive use of the analog of set selection in a probabilistic setting, also referring to causality for motivation. Skyrms’ account thus is closest to the one I am going to propose. Despite this similarity, however, there are several important differences in the use of causal relations, details of the implementation, and the predictions about the probabilities of conditionals, especially counterfactuals. These differences deserve a detailed comparison, but although I will briefly return to them at the end, I leave a full discussion for another occasion.

### 3. IMPLEMENTATION

#### 3.1. *Probabilities of Propositions*

Uncertainty is modeled as usual, using sets of possible worlds to represent “live” possibilities. The quantifiers of modal logic allow us to express statements such as whether something is necessary or possible with respect to such a set of worlds. With probabilities, we can populate the space between 0 and 1 with degrees of “support”. Starting with  $W$ , the set of all worlds, a probability distribution  $\text{Pr}$  is defined as follows.

DEFINITION 1 (Probability model). A *probability model* is a structure  $\langle W, \text{Pr} \rangle$ , with  $W$  a non-empty set of worlds and  $\text{Pr}: \wp(W) \mapsto [0, 1]$  a function such that for all  $X, Y \subseteq W$ :

$$\begin{aligned} \text{Pr}(W) &= 1 \\ 0 &\leq \text{Pr}(X) \leq 1 \\ \text{Pr}(X \cup Y) &= \text{Pr}(X) + \text{Pr}(Y) \text{ if } X \text{ and } Y \text{ are disjoint}^{11} \end{aligned}$$

The conditional probability of  $Y$ , given  $X$ , is defined as the “amount” of  $Y$ -worlds within the set of  $X$ -worlds, whenever the latter has positive probability.

DEFINITION 2 (Conditional probability). The *conditional probability* of  $Y$ , given  $X$ , for all  $X, Y \subseteq W$ , is:

$$\text{Pr}(Y|X) = \begin{cases} \frac{\text{Pr}(X \cap Y)}{\text{Pr}(X)} & \text{if } \text{Pr}(X) \neq 0 \\ \text{undefined} & \text{otherwise} \end{cases}$$

---

<sup>11</sup> The third condition must hold in general for the limits of countable unions of pairwise disjoint propositions. This is important for the coherence of the measure  $\text{Pr}$ , but I will ignore it here for simplicity.

The conditional probability is undefined if  $X$  has zero probability. This is above all an artifact of the mathematical definition; one may make use of or dispose of it in various ways. Some authors stipulate a value in those cases, such as 1 (Adams 1965; McGee 1989),  $\Pr(Y)$  (Milne 1997), or any arbitrary value (Skyrms 1988).

Alternatively, with an eye toward the application to conditionals, one may consider the undefined value a virtue rather than a shortcoming. Linguists typically account for the observation that there is nothing “correct” to be believed about a conditional whose antecedent is definitely false by stipulating that conditionals *presuppose* that there are antecedent-worlds to quantify over. Assuming that presupposition failure results in lack of a truth value, this may be seen as the logical analog of the probabilistic fact that the probability of a conditional is undefined if its antecedent has zero probability.

### 3.2. Probabilities of Truth-functional Sentences

The last section defined probabilities of propositions. The probabilities of sentences are defined in terms of those.

DEFINITION 3 (Language  $\mathcal{L}_{\mathcal{A}}^1$ ). Given a set  $\mathcal{A}$  of propositional letters, the *language*  $\mathcal{L}_{\mathcal{A}}^1$  is the smallest set containing  $\mathcal{A}$  and closed as follows: If  $\varphi, \psi \in \mathcal{L}_{\mathcal{A}}^1$ , then  $\bar{\varphi}, \varphi\psi \in \mathcal{L}_{\mathcal{A}}^1$ .

Note that  $\overline{\varphi\psi}$  is the negation of  $\varphi\psi$ , whereas  $\bar{\varphi}\bar{\psi}$  is the conjunction of  $\bar{\varphi}$  and  $\bar{\psi}$ . Disjunction and material conditional may be defined as usual:  $\varphi \vee \psi$  and  $\varphi \supset \psi$  are  $\overline{\bar{\varphi}\bar{\psi}}$  and  $\overline{\varphi\bar{\psi}}$ , respectively. I refer to the material conditional using the symbol  $\supset$ . We already know, of course, that the material conditional is not a suitable connective to model the natural-language conditional. Below, I will use  $\rightarrow$  for the latter.

Truth values are assigned to sentences pointwise at individual worlds.

DEFINITION 4 (Interpretation for  $\mathcal{L}_{\mathcal{A}}^1$ ). An *interpretation* of the language  $\mathcal{L}_{\mathcal{A}}^1$  in a Probability model  $\langle W, \Pr \rangle$  is a function  $V^1: \mathcal{L}_{\mathcal{A}}^1 \mapsto \{0, 1\}^W$  satisfying the following conditions:

$$\begin{aligned} \text{For } A \in \mathcal{A} : V^1(A)(w) &\in \{0, 1\} \\ V^1(\bar{\varphi})(w) &= 1 - V^1(\varphi)(w) \\ V^1(\varphi\psi)(w) &= V^1(\varphi)(w) \cdot V^1(\psi)(w) \end{aligned}$$

For each  $\varphi$ ,  $V^1(\varphi)$  is the characteristic function of the set of worlds at which  $\varphi$  is true. In a probabilistic context, functions from possible worlds

to truth values are a special case of *random variables*.<sup>12</sup> By convention, the range of a random variable is the set of real numbers or a subset thereof, such as  $\{0, 1\}$  or  $[0, 1]$ . Random variables may in general be continuous. I will ignore that case throughout. To generalize to the continuous case, summations are substituted with integrals.

As we will see, there is a certain tension between the probabilistic account and the idea that conditionals denote (characteristic functions of) propositions, as other sentences do. By way of preview, I will not make this latter assumption below: The denotations of conditionals will be random variables, but not ones whose range can in general be restricted to  $\{0, 1\}$ .

I will have occasion to refer to the set of worlds at which a sentence takes a particular value; for instance, the set  $\{w \in W | V^1(A)(w) = 1\}$  is the proposition “that  $A$  is true.” In the interest of readability, I abbreviate this expression as ‘ $V^1(A) = 1$ ’; thus  $\Pr(\{w \in W | V^1(A)(w) = 1\})$  is alternatively written  $\Pr(V^1(A) = 1)$ . Furthermore, I will write  $\Pr(V^1(A) = 1, V^1(B) = 1)$  instead of  $\Pr(\{w \in W | V^1(A)(w) = 1\} \cap \{w \in W | V^1(B)(w) = 1\})$  for the proposition that both  $A$  and  $B$  are true.

There is a straightforward relationship between the values of the characteristic function of a proposition and that proposition’s probability: The latter equals the *expectation* of the former – i.e., the weighted sum of its values, where the weights are the probabilities that the function takes those values. This notion, as well as the more general one of *conditional expectation*, will be important below.

DEFINITION 5 (Expectation). For random variables  $X, Y$ , expectation and conditional expectation are defined as follows:

$$E[X] = \sum_{x \in \text{range}(X)} x \cdot \Pr(X = x)$$

$$E[X|Y = y] = \sum_{x \in \text{range}(X)} x \cdot \Pr(X = x|Y = y)$$

Here  $X$  stands for any random variable, such as the denotation  $V^1(\varphi)$  of a sentence  $\varphi$ . The probability of a sentence is defined as the expectation of the characteristic function it denotes.

---

<sup>12</sup> In statistical jargon, the set of possible worlds is the *sample space*, in which each world is an *outcome*. A proposition is an *event*; its characteristic function is an *indicator function*. The statistical notion of an event is not to be confused with the “events” often dealt with in natural-language semantics. Instead, it is best thought of as the “event” of the English phrase ‘*in the event that ...*’.

DEFINITION 6. (Probabilities of sentences). Given a probability model  $\langle W, \text{Pr} \rangle$  and an assignment  $V^1$  for  $\mathcal{L}_{\mathcal{A}}^1$ , a probability distribution  $P$  on the sentences in  $\mathcal{L}_{\mathcal{A}}^1$  is defined as follows: For all  $\varphi, \psi \in \mathcal{L}_{\mathcal{A}}^1$ ,

$$\begin{aligned} P(\varphi) &= E[V^1(\varphi)] \\ P(\psi|\varphi) &= E[V^1(\psi)|V^1(\varphi) = 1] \end{aligned}$$

Since all sentences in  $\mathcal{L}_{\mathcal{A}}^1$  denote functions with range  $\{0, 1\}$ , the summation in Definition 5 is trivial and the relationship between  $P$  and  $\text{Pr}$  is straightforward: For all  $\varphi, \psi$  in  $\mathcal{L}_{\mathcal{A}}^1$ ,  $P(\varphi)$  is the probability that  $\varphi$  is true,<sup>13</sup> and  $P(\psi|\varphi)$  is the conditional probability that  $\psi$  is true, given that  $\varphi$  is true.<sup>14</sup>

### 3.3. Probabilities of Conditionals

The probability of the sentence *you strike the match* is the probability that you strike the match. The Thesis asserts that the probabilities of conditionals are conditional probabilities. The goal is to extend the truth assignment  $V^1$  to conditionals in such a way that the expectation  $P(A \rightarrow C)$  of the truth values of  $A \rightarrow C$  equals the conditional probability  $P(C|A)$  whenever the latter is defined, regardless of the underlying probability distribution  $\text{Pr}$ .

Unfortunately, it is impossible, aside from certain special cases, to assign truth values to conditionals in this way. A conditional probability is not the probability that a proposition is true. Consequently, if conditionals are to obey the Thesis, they cannot denote propositions like atomic and truth-functional sentences. This is the lesson from Lewis' (1976, 1986) *triviality results* and a large body of subsequent work.

#### 3.3.1. Triviality

Lewis' results and their various extensions and generalizations are introduced in a number of detailed discussions (Gibbard 1981; Hájek and Hall

$$\begin{aligned} \text{13} \quad P(\varphi) &= E[V^1(\varphi)] = \sum_{x \in \{0,1\}} x \cdot \text{Pr}(V^1(\varphi) = x) \\ &= 0 \cdot \text{Pr}(V^1(\varphi) = 0) + 1 \cdot \text{Pr}(V^1(\varphi) = 1) = \text{Pr}(V^1(\varphi) = 1) \end{aligned}$$

$$\begin{aligned} \text{14} \quad P(\psi|\varphi) &= E[V^1(\psi)|V^1(\varphi) = 1] \\ &= \sum_{x \in \{0,1\}} x \cdot \text{Pr}(V^1(\psi) = x | V^1(\varphi) = 1) \\ &= 0 \cdot \text{Pr}(V^1(\psi) = 0 | V^1(\varphi) = 1) + 1 \cdot \text{Pr}(V^1(\psi) = 1 | V^1(\varphi) = 1) \\ &= 1 | V^1(\varphi) = 1 = \text{Pr}(V^1(\psi) = 1 | V^1(\varphi) = 1) \end{aligned}$$



1994; Hájek 1994; Edgington 1995, and elsewhere). I will only give an informal description of one of them in the framework introduced in Section 3. This will illustrate an essential part of the problem, and it will help the reader see how the solution I adopt below avoids it.

The assumption that conditionals denote propositions in the usual sense implies that their values are, at each individual world, (i) either 0 or 1 and (ii) independent of the probability distribution. In particular,  $A \rightarrow C$  will retain its interpretation under conditioning, so we may ask what  $P(A \rightarrow C|D)$  should be when  $P(D)$  is not zero. Lewis assumes that this probability should equal  $P(C|AD)$  whenever the latter is defined. Consider the special case in which  $D$  is either  $C$  or  $\bar{C}$ . Then the assumption implies that

- (16) a.  $P(A \rightarrow C|C) = P(C|AC) = 1$
- b.  $P(A \rightarrow C|\bar{C}) = P(C|A\bar{C}) = 0$

This is plausible. In case the coin is rigged to come up tails no matter what, (17a) is certainly true and (17b) is certainly false.

- (17) a. If I bet on tails, the coin will come up tails.
- b. If I bet on tails, the coin will come up heads.

To see the formal consequences of this, recall that conditional probabilities of sentences equal the underlying conditional expectations. So (16a,b) amount to (18a,b):

- (18) a.  $E[V(A \rightarrow C)|V(C) = 1] = 1$
- b.  $E[V(A \rightarrow C)|V(C) = 0] = 0$

Thus with probability 1, the conditional is equivalent to its consequent. Recall also that the above should hold for *any* underlying probability distribution  $\text{Pr}$  in which the conditional probabilities are defined. As a consequence, (19a) must be the case for any corresponding distribution  $P$  over sentences. Together with the Thesis, this implies (19b).

- (19) a.  $P(A \rightarrow C) = P(C)$
- b.  $P(C|A) = P(C)$

It follows that the consequent must be stochastically independent of the antecedent. This is absurd: Conditionals are typically used to assert that such a dependence *does* hold. In addition, it implies that the Thesis only

holds in full generality – i.e., for *all* distributions in which the probabilities are defined – if the language distinguishes at most two non-empty propositions, for (19a) is inconsistent with the Thesis whenever three mutually disjoint propositions  $AC$ ,  $A\bar{C}$  and  $\bar{A}$  have non-zero probability.<sup>15</sup> Lewis calls such a language *trivial*.

This first triviality result is that the conditional connective  $\rightarrow$  of a non-trivial language cannot be a “universal probability conditional”; that is, its interpretation cannot be independent of the probability distribution. It is possible, however, to start out with a given probability distribution and define a conditional in such a way that the Thesis holds within that particular distribution. In his remaining triviality results, Lewis considers what happens if such a conditional that is “tailored” to a given probability distribution retains its interpretation while the probabilities are altered by various kinds of update operations. He shows that except for special trivial cases, it is impossible to guarantee that the Thesis holds both before and after the update. Lewis’ results have since been extended and generalized in a number of ways (see the references mentioned above).<sup>16</sup>

### 3.3.2. Context Dependence

The triviality results show that the Thesis is incompatible with the assumption that conditionals denote propositions in the usual sense. It does not follow that the Thesis cannot be upheld. One way to avoid triviality is to assign values to the conditional that in turn depend on the probability distribution, thus making its denotation dependent upon the context in which it is interpreted.

Various authors have found good arguments for such a position, both in probabilistic and non-probabilistic settings (Harper 1976; Stalnaker 1975; Gibbard 1981; van Rooy 1997). How best to make sense of it depends on the relevant notion of “context”. Some of the mentioned authors explicitly made their arguments with regard to epistemic conditionals and subjective belief states. I adopt it here for predictive conditionals and objective chance. The question of what constitutes the context in this case and in what sense it is objective will be taken up in Section 4.

Specifically, I will adopt, but modify below, a proposal made by Jeffrey (1991; cf. also Stalnaker and Jeffrey 1994). Jeffrey, in a sense, put the cart in front of the horse: Instead of deriving the probability of the conditional

<sup>15</sup> Consider  $A \rightarrow AC$ . (i) By (19a),  $P(A \rightarrow AC) = P(AC)$ . (ii) By the Thesis, the probabilistic calculus and propositional logic,  $P(A \rightarrow AC) = P(AC|A) = P(ACA)/P(A) = P(AC)/P(A)$ . Hence by (i) and (ii),  $P(A) = 1$  and  $P(\bar{A}) = 0$ .

<sup>16</sup> Aside from the probabilistic setting with which these elaborations are mostly concerned, an interesting related result was given by Gärdenfors (1988) for the non-probabilistic case.

from considerations of what its truth values at individual worlds should be, he started out with a given amount of probability mass – the conditional probability – and proposed a systematic way of distributing it over the set of possible worlds. His proposal rests on the following simple fact about the probabilistic calculus:

$$(20) \quad P(\psi|\varphi) = P(\psi|\varphi)[P(\varphi) + P(\bar{\varphi})]$$

$$= P(\varphi\psi) + P(\psi|\varphi)P(\bar{\varphi})$$

$$(21) \quad = 0 \cdot P(\varphi\bar{\psi}) + 1 \cdot P(\varphi\psi) + P(\psi|\varphi) \cdot P(\bar{\varphi})$$

Each of the three terms on the right-hand side of (21) can be read as the value of a random variable, multiplied by the probability that it takes that value. The sentences determining the weights are mutually incompatible and jointly exhaust all possibilities. Conditionals, Jeffrey suggests, denote random variables defined in just this way. Recalling that  $P(\psi|\varphi) = E[V(\psi)|V(\varphi) = 1]$ , the corresponding value assignment would be (22).

$$(22) \quad V(\varphi \rightarrow \psi)(w) = \begin{cases} V(\psi)(w) & \text{if } V(\varphi)(w) = 1 \\ E[V(\psi)|V(\varphi) = 1] & \text{if } V(\varphi)(w) = 0 \end{cases}$$

This is, of course, only one among countless ways of distributing values over worlds in such a way that their expectation is the conditional probability. It is appealing, however, in that it assigns to the conditional a definite truth value (that of the material conditional) at worlds at which the antecedent is true, in accordance with the intuition that in that case the truth or falsehood of the conditional can be determined by simple inspection of the facts. Where the antecedent is false, the conditional takes values that may lie between 0 and 1 and depend, via the expectation, on the probability distribution.

Thus the assignment in (22) ensures by its very definition that the Thesis will be upheld.<sup>17</sup> It is not immediately clear, however, in what sense

---

<sup>17</sup> 
$$P(\varphi \rightarrow \psi) = E[V(\varphi \rightarrow \psi)] = \sum_{x \in [0,1]} x \cdot \Pr(V(\varphi \rightarrow \psi) = x)$$

$$= 0 \cdot \Pr(V(\varphi \rightarrow \psi) = 0) + 1 \cdot \Pr(V(\varphi \rightarrow \psi) = 1)$$

$$+ E[V(\psi)|V(\varphi) = 1] \cdot \Pr(V(\varphi \rightarrow \psi) = E[V(\psi)|V(\varphi) = 1])$$

$$= 0 \cdot \Pr(V(\varphi) = 1, V(\psi) = 0) + 1 \cdot \Pr(V(\varphi) = 1, V(\psi) = 1)$$

$$+ E[V(\psi)|V(\varphi) = 1] \cdot \Pr(V(\varphi) = 0)$$

$$= 0 \cdot P(\varphi\bar{\psi}) + 1 \cdot P(\varphi\psi) + P(\psi|\varphi) \cdot P(\bar{\varphi})$$

$$= P(\varphi\psi) + P(\psi|\varphi) \cdot P(\bar{\varphi}) = P(\psi|\varphi)[P(\varphi) + P(\bar{\varphi})] = P(\psi|\varphi)$$

this “rather weird three-valued entity” (Edgington 1995) expresses what conditionals *mean*. Jeffrey did not offer much in the way of an intuitive motivation, but other authors have argued that such a motivation can be gleaned from the set-selection semantics in the tradition of Stalnaker and Lewis (van Fraassen 1976; Stalnaker and Jeffrey 1994; Edgington 1995).

Recall that in Stalnaker’s theory, the value of  $A \rightarrow C$  at worlds at which  $A$  is false depends on the nearest  $A$ -world, which is uniquely identified in the model by a selection function. Definition (22) does not require a single nearest  $A$ -world; instead, an  $A$ -world is chosen *at random*. The conditional is guaranteed by (22) to have a definite truth value there. The expectation of that truth value is just the conditional probability of the consequent, given the antecedent. This is the value assigned to the conditional at the non-antecedent world of evaluation.

Jeffrey’s proposal is my point of departure. In the next section, I will flesh out the justification of the assignment function in (22). Two questions need to be addressed: First, as defined above,  $V$  is a function of *three* arguments: a sentence, a world, and a probability distribution. The value at one and the same world will change if the probability distribution changes. Is there a sense, then, in which this value can be thought of as “objective”, like truth values? Secondly, what does it *mean* for the values at non-antecedent worlds – in the absence of any uncertainty about the facts – to be intermediate?

#### 4. TRUTH AND CHANCE IN TIME

Predictive conditionals are used to make conditional predictions. Their formal treatment requires a representation of time. In this section I will enrich the model accordingly.

First, however, the notion of “truth” must be made precise when it comes to future reference. Since I defined the probabilities of sentences as the expectations of their truth values, in order to speak of the *current* probabilities of predictions, it must be assumed that they already have truth values. Some speakers find this counterintuitive and maintain that except in cases of predetermination, sentences about the future are generally “not yet true” (nor false) at the time they are used. This judgment is legitimate, but the notion of truth that it rests on is not the “Ockhamist” one I have in mind.

It is well-known that the past is settled in a sense in which the future is not. This asymmetry has been discussed at least since Aristotle’s remarks on future sea battles (On Interpretation 1:9), and more recently by Ramsey (1929), Reichenbach (1956), Prior (1967), Thomason (1970), Burgess

(1979), Lewis (1979, 1980), and others. The difference is that at their respective evaluation times, (23a) is either necessarily true or necessarily false, whereas (23b) is neither necessarily true nor necessarily false.

- (23) a. The coin landed heads.  
 b. The coin will land heads.

The Ockhamist view maintains that (23b) does have the weaker property of being necessarily either true or false: It is true at the time it is used if and only if it *turns out* true at the relevant later time. Accordingly, it may be true without there being any way of *knowing* (already) that it is true.

On the alternative, “Peircean” view, there is no use for truth values that exist but elude us when they are needed most (namely when the sentence is used); instead, for the sentence to be true, it must be inevitable, i.e., true in *all* possible continuations of history. The Peircean notion of “truth” coincides with Ockhamist “settledness”. This is the notion behind the intuition that sentences like (23b) have no truth values at utterance time.

The model theory to be introduced in this section will ensure that a sentence with a future reference time  $t$  is now Ockham-true if and only if it is Peirce-true at  $t$ . Thus the expectation of its Ockhamist truth values is the probability that it will be Peirce-true at  $t$ . This establishes a straightforward connection between truth and probability which should satisfy adherents to both views, at least as long as it is assumed that a definite reference time  $t$  is given.<sup>18</sup>

#### 4.1. Models

The basic structure I adopt is what Thomason (1984) calls a “ $T \times W$ -frame”.<sup>19</sup> Worlds are now “world-lines”, complete trajectories through time. A *history* is a decreasing sequence of sets of such trajectories, recording at each time  $t$  the past facts accumulated up to  $t$  as well as the continuations possible at  $t$ . Formally, this is embodied in the condition that all worlds included in the time slice of a history at  $t$  be indistinguishable at all times up to  $t$  (but may come apart thereafter).

In Figure 1, three histories are represented by narrowing gray areas. The black areas enclose worlds which are alike up to some particular time.

<sup>18</sup> Things are a bit more complex when this last assumption is dropped. See Kaufmann (2004) for a proposal.

<sup>19</sup> Some of the definitions in this section are adapted from Thomason (1984), where the reader is also referred for a discussion of the relationship between  $T \times W$ -frames and models of branching time.

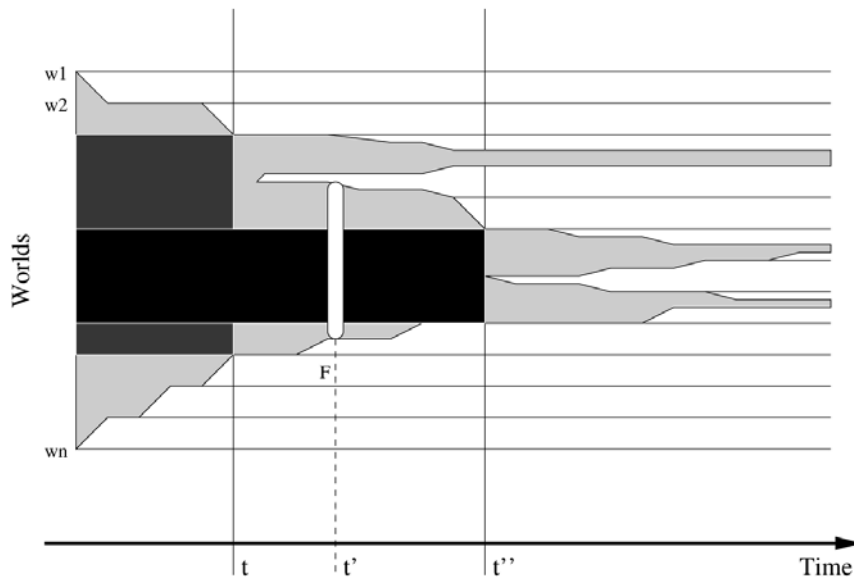


Figure 1.  $T \times W$ -frame.

These are each other's *historical alternatives* at that time. It is with respect to such sets of alternatives that time-variant notions which refer to quantities of worlds, such as objective chance, are encoded.

DEFINITION 7 ( $T \times W$ -frames (Thomason, 1984)). A  $T \times W$ -frame is a quadruple  $\langle W, T, <, \approx \rangle$ , where (1)  $W$  and  $T$  are disjoint nonempty sets; (2)  $<$  is a transitive relation on  $T$  which is also irreflexive (for all  $t \in T$ ,  $t \not< t$ ) and linear (for all  $t, t' \in T$ , either  $t < t'$  or  $t' < t$  or  $t = t'$ ); and (3)  $\approx$  is a three-place relation in  $T \times W \times W$ , such that (3a) for all  $t$ ,  $\approx_t$  is an equivalence relation; and (3b) for all  $w, w' \in W$  and  $t, t' \in T$ , if  $w \approx_t w'$  and  $t' < t$ , then  $w \approx_{t'} w'$ .

Each world evolves by shedding alternative futures. I will refer to the equivalence class  $\{w' | w \approx_t w'\}$  of historical alternatives of world  $w$  at time  $t$  as  $[w]_t^\approx$ . Clause (3b) in the definition is intended to ensure that the progressive loss of alternatives is not reversible. It is not possible for worlds to “become” historical alternatives at some point in time, or for alternatives “not to have been” alternatives at earlier times.

The assignment function  $V^1$  is now defined in such a way that each atomic sentence at each time is assigned (the characteristic function of) the set of worlds in which it is true at that time.

DEFINITION 8 ( $T \times W$  interpretation). An *interpretation* of the language  $\mathcal{L}_{\mathcal{A}}^1$  in a  $T \times W$ -frame is a function  $V^1: \mathcal{L}_{\mathcal{A}}^1 \mapsto (T \mapsto (W \mapsto \{0, 1\}))$  from expressions in  $\mathcal{L}_{\mathcal{A}}^1$  to functions from members of  $T$  to (characteristic functions of) propositions, provided that for all atomic sentences  $A$  in  $\mathcal{A}$  and all times  $t$ , if  $w \approx_t w'$ , then  $V^1(A)(t)(w) = V^1(A)(t)(w')$ .

Propositions, sets of worlds as before, are now sets of world-lines. The added temporal dimension affords a richer set of means by which to *identify* them, using a variety of linguistic means. I will not be concerned with the details of this process, assuming instead that sentences are somehow given their reference times. For convenience, I superscript sentences with elements of  $T$ , the times in the model. It would be cleaner to use a set of distinct constants interpreted as referring to those times, but no confusion is likely to arise from the shortcut.

DEFINITION 9 (Language  $\mathcal{L}_{\mathcal{A}}^{T1}$ ). The language  $\mathcal{L}_{\mathcal{A}}^{T1}$  is the smallest set containing  $\mathcal{A}$  and such that for all  $\varphi, \psi \in \mathcal{L}_{\mathcal{A}}^{T1}$  and  $t \in T, \bar{\varphi}, \varphi\psi, \varphi^t \in \mathcal{L}_{\mathcal{A}}^{T1}$ .

The assignment of truth values is now as given in Definition 10. Below I will drop the superscripts from  $V^{T1}$  and  $\mathcal{L}_{\mathcal{A}}^{T1}$ .

DEFINITION 10 (Interpretation for  $\mathcal{L}_{\mathcal{A}}^{T1}$ ). An interpretation  $V^1$  for  $\mathcal{L}_{\mathcal{A}}^{T1}$  in a  $T \times W$ -frame is extended to an assignment  $V^{T1}$  for  $\mathcal{L}_{\mathcal{A}}^{T1}$  as follows: For all  $\varphi$  in  $\mathcal{L}_{\mathcal{A}}^{T1}$ ,

$$\begin{aligned} \text{If } \varphi \in \mathcal{L}_{\mathcal{A}} : V^{T1}(\varphi)(t)(w) &= V^1(\varphi)(t)(w) \\ V^{T1}(\varphi^t)(t)(w) &= V^{T1}(\varphi)(t')(w) \end{aligned}$$

“Chance” is the objective probability that a proposition is true (Lewis, 1980; cf. also van Fraassen, 1981). It is encoded as a prior probability distribution over the worlds (now world-lines) in the model, similarly to the time-less model from Definition 1.

DEFINITION 11 (Chance model). A *chance model* is a structure  $\langle W, T, <, \approx, \text{Pr} \rangle$  where  $\langle W, T, <, \approx \rangle$  is a  $T \times W$ -frame and  $\text{Pr}$  is a probability distribution over  $W$ .

Chance does not remain constant over time. The probability distribution changes over time by conditionalization on what is settled. What is settled in a world  $w$  at a time  $t$  is represented by the proposition  $[w]_t^{\approx}$ .

DEFINITION 12 (Chance in history). The *chance* of a proposition  $X$  at a time  $t$  in a world  $w$  is  $Pr(X|[w]_t^{\approx})$ .

Thus “interesting” chances – ones which differ from the corresponding truth values – arise only with respect to the future. The chance that the match was struck is either 1 (if it was struck) or 0 (if it wasn’t). The chance that it will be struck may fall between 1 and 0.

Notice that chances are fully determined by the world and time of evaluation, since these two parameters determine the set of historical alternatives in Definition 12. In this sense, the values of conditionals, though dependent on the probability distribution, are nevertheless objective. Thus a potential objection to the Jeffrey conditional is avoided which was already voiced by Lewis (1976), who insisted that the conditional must have an interpretation that is independent of *subjective* probabilities: “Else how are disagreements about a conditional possible, or changes of mind?” Assuming an objective interpretation of the probability, Lewis’ question concerns the relationship between it and subjective probabilities, which is beyond the scope of this paper.<sup>20</sup>

#### 4.2. Interpretation

The probabilities of sentences, too, depend on the time of evaluation. In addition, they depend on the world, since different histories give rise to different probability distributions over the open futures. Thus:<sup>21</sup>

$$\begin{aligned}
 (24) \quad P_{w,t}(\varphi) &= E[V(\varphi)(t)|[w]_t^{\approx}] \\
 &= \sum_{x \in [0,1]} x \cdot \Pr(V(\varphi)(t) = x|[w]_t^{\approx}) \\
 &= \Pr(V(\varphi)(t) = 1|[w]_t^{\approx}) \text{ if } \varphi \text{ is truth-functional}
 \end{aligned}$$

Conditionals are evaluated according to Jeffrey’s Formula (22), with the consequence that their probabilities equal the conditional probabilities at the time of evaluation:

$$(25) \quad V(\varphi \rightarrow \psi)(t)(w) = \begin{cases} V(\psi)(t)(w) & \text{if } V(\varphi)(t)(w) = 1 \\ E[V(\psi)(t)|V(\varphi)(t) = 1, [w]_t^{\approx}] & \\ \text{if } V(\varphi)(t)(w) = 0 & \end{cases}$$

<sup>20</sup> In defining chance as a prior probability distribution, I sidestepped some deep metaphysical questions, such as how chance is related to the facts of the world, whether it could in principle be known (Lewis, 1980, 1994) and, indeed, whether it exists. I take it to be self-evident that speakers talk *as if* there was such a thing, but for my purposes it may just as well be a special kind of subjective belief – special in that we take it to be the belief that anyone would hold who knew everything about the past and could correctly project expectations from those facts.

<sup>21</sup> Recall that the right-hand side of the first line can be further rewritten as  $\sum_{x \in [0,1]} x \cdot \Pr(\{w' | w \approx_t w', V(\varphi)(t)(w') = x\}) / \Pr([w]_t^{\approx})$ .



I will drop the subscripts from  $P_{w,t}$  below. The connection between the values and the probability of a conditional is again straightforward (the argument is similar to footnote 17).

$$(26) \quad P_{w,t}(\varphi \rightarrow \psi) = E[V(\varphi \rightarrow \psi)(t)|[w]_t^{\approx}] \\ = \Pr(V(\psi)(t) = 1|V(\varphi)(t) = 1, [w]_t^{\approx})$$

It is not hard to find a conceptually satisfactory explanation for the value assignment in (25) in the context of time and chance. According to the Ockhamist view, the current truth or falsehood of a non-conditional prediction is determined by future facts. The uncertainty arises because it is impossible to tell which historical alternative is the actual one, not because the actual world is such that the prediction does not (yet) have a truth value. As a guide in *discovering* the truth value, this notion implies that once the relevant facts are settled, one should be able to tell with hindsight whether the prediction *was* true or false at the time it was made. Such a finding would not imply that it was *likely* to be true at the time – it may turn out true by pure coincidence.

The same is only partly the case for conditional predictions like (27). As it happens, both antecedent and consequent are true, so the sentence was true throughout the decades leading up to the events in question (although it was in no way foreseeable).<sup>22</sup>

- (27) If Gorbachev visits East Berlin in October, 1989, the Wall will fall in November, 1989.

But what will the future reveal about the truth of a conditional whose antecedent is false? I agree with Ramsey (1929), who maintains that the truth or falsehood of such “unfulfilled” conditionals is not determined by the facts. In a world in which the antecedent of (27) is false, there seems to be no way to determine with certainty, at any subsequent time, whether the conditional was true or false.

For a less “historical” and perhaps even clearer example, consider an upcoming toss of a fair die. There surely is a number  $x$  such that the outcome will be  $x$ . It is also clear that if the outcome is an even number, then there will be a number  $y$  such that the outcome is  $y$ . But unless the outcome is in fact an even number, there is no number  $z$  such that if the outcome is an even number, it will be  $z$ .

<sup>22</sup> Mackie (1973) discussed conditionals that turn out true or false by accident (see also Edgington, 1995). Pendlebury (1989) and Read (1995) hold the opposing view that the conjunction should not entail the conditional.

But even if the antecedent turns out false, the world did “support” the conditional to some degree, given the facts up to the time in question. This may be taken as an intuitive rationale for assigning *some* value in this case, rather than leaving it undefined.

## 5. RIGHT-NESTED CONDITIONALS

The previous sections showed how the random-variable interpretation of conditionals circumvents the triviality results, and how it can be integrated in an overall probabilistic semantics. One of the merits of this approach is that it yields predictions about the probabilities of compounded and embedded conditionals, such as (28a), which I consider equivalent to (28b).

- (28) a. If the match is wet, then if you strike it, it will light.  
 b. If the match is wet, then it will light if you strike it.  
 c.  $W \rightarrow (S \rightarrow L)$

The standard probabilistic calculus does not provide probabilities of expressions of the form (28c): ‘ $P((L|S)|W)$ ’ is not defined. But it is possible to take the conditional expectation of its consequent  $S \rightarrow L$ , given that its antecedent  $W$  is true, and assign that expectation as the value of (28) at those worlds at which  $W$  is false. No modification of Jeffrey’s truth assignment is needed to accomplish this.<sup>23</sup>

Unfortunately, however, the predictions that Jeffrey’s account makes for such conditionals turn out to be counterintuitive (Edgington, 1991; Lance, 1991). To see this, consider a concrete example.

- (29) The probability that . . .  
 a. it gets wet is “low” (0.1)  
 b. you strike it is 0.5  
 c. it lights given that you strike it and it is dry is “high” (0.9)  
 d. it lights given that you strike it and it is wet is “low” (0.1)  
 and the probabilities of the striking and the wetness are independent of each other.

<sup>23</sup> The interpretation of left-nested and conjoined conditionals involves some complications; see Stalnaker and Jeffrey (1994) for these cases, which I do not discuss in this paper.

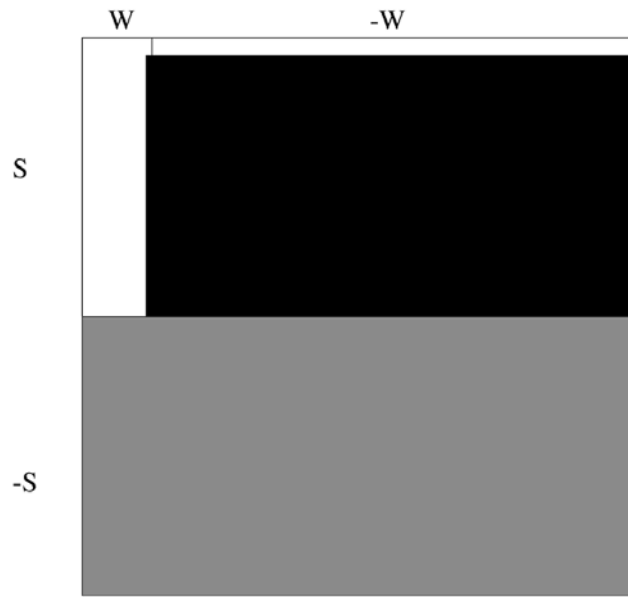


Figure 2. Distribution of values of (30).

Let us see how in this scenario the values of the consequent (30) are distributed over the set of worlds.

(30) If you strike the match, it will light.

In Figure 2 and below, the values are represented as shades of gray: black for the value 1 at worlds at which the match is struck and lights, white for the value 0 at worlds where it is struck and does not light, and gray for the intermediate value 0.82 assigned at those worlds at which it is not struck.<sup>24</sup> The value assignment is as in (31) (for the sake of readability, I omit reference to times here and later on in the section).

$$(31) \quad V(S \rightarrow L)(w) = \begin{cases} V(L)(w) & \text{if } V(S)(w) = 1 \\ E[V(L)|V(S) = 1] & \text{if } V(S)(w) = 0 \end{cases}$$

<sup>24</sup> The value 0.82 is the conditional probability that it lights, given that it is struck and can be calculated from the numbers given above as follows (recall that  $W$  and  $S$  are independent, so  $P(W|S) = P(W)$  and  $P(\bar{W}|S) = P(\bar{W})$ ):

$$\begin{aligned} P(L|S) &= P(LS)/P(S) = [P(LSW) + P(LS\bar{W})]/P(S) \\ &= [P(L|SW)P(W|S)P(S) + P(L|S\bar{W})P(\bar{W}|S)P(S)]/P(S) \\ &= P(L|SW)P(W) + P(L|S\bar{W})P(\bar{W}) = 0.1 \cdot 0.1 + 0.9 \cdot 0.9 = 0.82 \end{aligned}$$

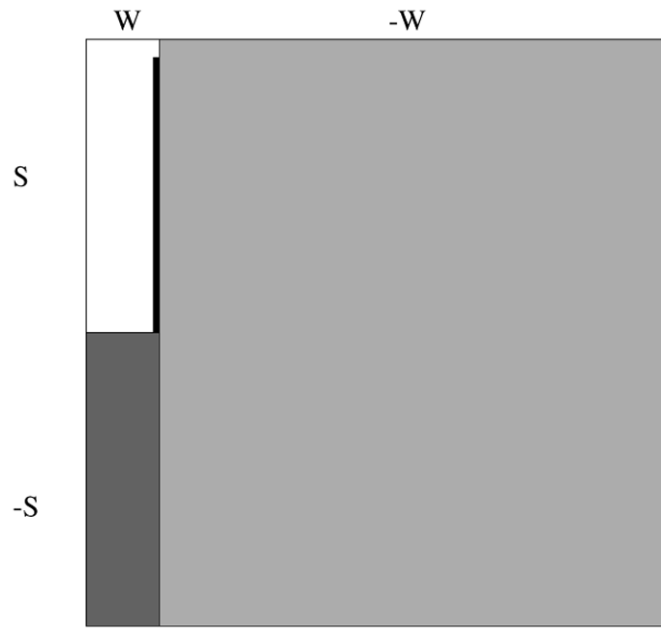


Figure 3. Distribution of values of (28).

$$= \begin{cases} 0 & \text{if } V(S)(w) = 1, V(L)(w) = 0 \\ 1 & \text{if } V(S)(w) = 1, V(L)(w) = 1 \\ 0.82 & \text{if } V(S)(w) = 0 \end{cases}$$

The conditional in (28) is evaluated as follows: The values of its consequent (30) are assigned at those worlds at which the antecedent is true, i.e., at which the match is wet. The expectation of these values is then distributed over the set of worlds at which the antecedent is false, i.e., the match is dry. This expectation is as follows.

$$\begin{aligned}
 (32) \quad & E[V(S \rightarrow L) | V(W) = 1] \\
 &= \sum_{x \in [0,1]} x \cdot \Pr(V(S \rightarrow L) = x | V(W) = 1) \\
 &= 0 \cdot \Pr(V(S) = 1, V(L) = 0 | V(W) = 1) \\
 &\quad + 1 \cdot \Pr(V(S) = 1, V(L) = 1 | V(W) = 1) \\
 &\quad + 0.82 \cdot \Pr(V(S) = 0 | V(W) = 1) \\
 &= 0 \cdot 0.45 + 1 \cdot 0.05 + 0.82 \cdot 0.5 = 0.46
 \end{aligned}$$

The values assigned to (28) according to (32) are given in (33). The shaded lines will later turn out to be the source of a problem with this approach.

$$\begin{aligned}
 (33) \quad & V(W \rightarrow (S \rightarrow L))(w) \\
 &= \begin{cases} V(S \rightarrow L)(w) & \text{if } V(W)(w) = 1 \\ E[V(S \rightarrow L)|V(W) = 1] & \text{if } V(W)(w) = 0 \end{cases} \\
 &= \begin{cases} V(L)(w) & \text{if } V(W)(w) = 1, V(S)(w) = 1 \\ E[V(L)|V(S) = 1] & \text{if } V(W)(w) = 1, V(S)(w) = 0 \\ E[V(S \rightarrow L)|V(W) = 1] & \text{if } V(W)(w) = 0 \end{cases} \\
 &= \begin{cases} 0 & \text{if } V(W)(w) = 1, V(S)(w) = 1, V(L)(w) = 0 \\ 1 & \text{if } V(W)(w) = 1, V(S)(w) = 1, V(L)(w) = 1 \\ 0.82 & \text{if } V(W)(w) = 1, V(S)(w) = 0 \\ 0.46 & \text{if } V(W)(w) = 0 \end{cases}
 \end{aligned}$$

The distribution of these values is illustrated in Figure 3. Their expectation is again 0.46:

$$\begin{aligned}
 (34) \quad & P(W \rightarrow (S \rightarrow L)) = E[V(W \rightarrow (S \rightarrow L))] \\
 &= \sum_{x \in [0,1]} x \cdot \Pr(V(W \rightarrow (S \rightarrow L)) = x) \\
 &= 0 \cdot 0.045 + 1 \cdot 0.005 + 0.82 \cdot 0.05 + 0.46 \cdot 0.9 = 0.46
 \end{aligned}$$

But this result is intuitively wrong. Given the scenario, the probability of the conditional should be small, rather than close to 0.5.

Examples like this have been used by Lance (1991) and Edgington (1991) to point out a number of similarly counterintuitive predictions. In the next subsection I will discuss what the problem is and how it might be addressed.

### 5.1. *The Values at Non-Antecedent Worlds*

Intuitively, the probability of (28) should not be 0.46 as predicted by the Jeffrey-style value assignment. What, then, should it be? It seems like 0.1, the probability of (29d), would be a much better estimate. I take this probability to be that of the simple conditional in (35):

$$(35) \quad \text{If you strike the match and it is wet, it will light.}$$

As before, the conditional has clearcut truth values only at worlds at which its antecedent is true, i.e., at which the match is wet and struck. The

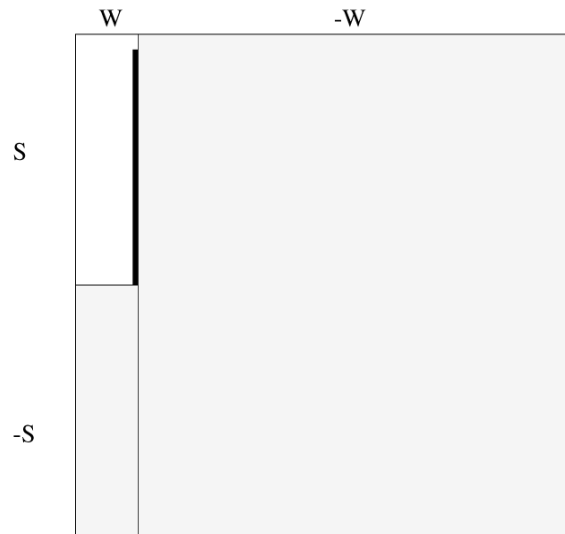


Figure 4. Distribution of values of (35).

expectation of these values, given as 0.1 in (29d), is uniformly distributed over the worlds at which the antecedent is false. The values are calculated in (36); their distribution is shown in Figure 4.

$$\begin{aligned}
 (36) \quad V((SW) \rightarrow L)(w) &= \begin{cases} V(L)(w) & \text{if } V(SW)(w) = 1 \\ E[V(L)|V(SW) = 1] & \text{if } V(SW) = 0 \end{cases} \\
 &= \begin{cases} 0 & \text{if } V(SW)(w) = 1, V(L)(w) = 0 \\ 1 & \text{if } V(SW)(w) = 1, V(L)(w) = 1 \\ 0.1 & \text{if } V(SW)(w) = 0 \end{cases}
 \end{aligned}$$

Comparing Figures 3 and 4, it seems that the expectation of the values in the latter is closer to intuitions. And indeed, that expectation, given in (37), is what one would expect.

$$\begin{aligned}
 (37) \quad P((SW) \rightarrow L) &= E[V((SW) \rightarrow L)] \\
 &= 0 \cdot \Pr(V(SW) = 1, V(L) = 0) \\
 &\quad + 1 \cdot \Pr(V(SW) = 1, V(L) = 1) \\
 &\quad + 0.1 \cdot \Pr(V(SW) = 0) \\
 &= 0 \cdot 0.045 + 1 \cdot 0.005 + 0.1 \cdot 0.95 = 0.1
 \end{aligned}$$

That the probability of (35) is a good estimate of that of (28) is quite clear. Some authors have generalized this observation and based on it proposals for dealing with all right-nested conditionals. Thus McGee (1989)

stipulates the equality in (38a). This definition applies recursively, unpacking and flattening arbitrarily deep right-embeddings of conditionals.<sup>25</sup> The same can be done in the present framework with the definition in (38b).

- (38) a.  $P(\varphi \rightarrow (\psi \rightarrow \chi)) =_{df} P((\varphi\psi) \rightarrow \chi) = P(\chi|\varphi\psi)$   
 b.  $V(\varphi \rightarrow (\psi \rightarrow \chi)) =_{df} V((\varphi\psi) \rightarrow \chi)$

This “Import-Export Principle” is logically valid for the material conditional. The diagrams in Figures 3 and 4 show why this move would improve the predictions about the conditional in (28). The truth values at worlds in which the match is both wet and struck are the same in both cases. The difference lies in the worlds at which the antecedents are false. In Figure 4, these are all the worlds at which the match is either dry or not struck, or both. In contrast, in Figure 3 the non-antecedent worlds are only those at which the match is dry. Those at which it is wet but not struck verify the antecedent of (28); therefore the conditional there receives the value of its consequent. (Recall that the values of the consequent were given in Figure 2.)

It would not do to stipulate somehow that the values of the consequent should be uniformly low in this example. For in the case of (39a), their expectation should arguably be high. This is again correctly predicted if the sentence is treated as equivalent to (39b) in accordance with the Import-Export Principle (38b). The resulting values would then be as shown in Figure 5 (I skip the details of the calculation).

- (39) a. If the match is dry, then it will light if you strike it.  
 b. If you strike the match and it is dry, it will light.

To rectify this situation, it would be necessary to assign to (40a), the consequent of (28), different values at non-antecedent worlds depending on whether the match is wet or not: At worlds at which it is wet, the value is that of (40b), and at worlds at which it is dry, it is that of (40c).

- (40) a. If you strike it, it will light.  
 b. If you strike it and it is wet, it will light.  
 c. If you strike it and it is dry, it will light.

---

<sup>25</sup> Notice that McGee does not also require (38a) to equal  $P(\psi \rightarrow \chi|\varphi)$ . He does not conditionalize conditional consequents on their antecedents, thus avoiding to treat them as proposition-denoting. This is what enables him to “wiggle out of the trivialization argument” (Stalnaker, 1991).

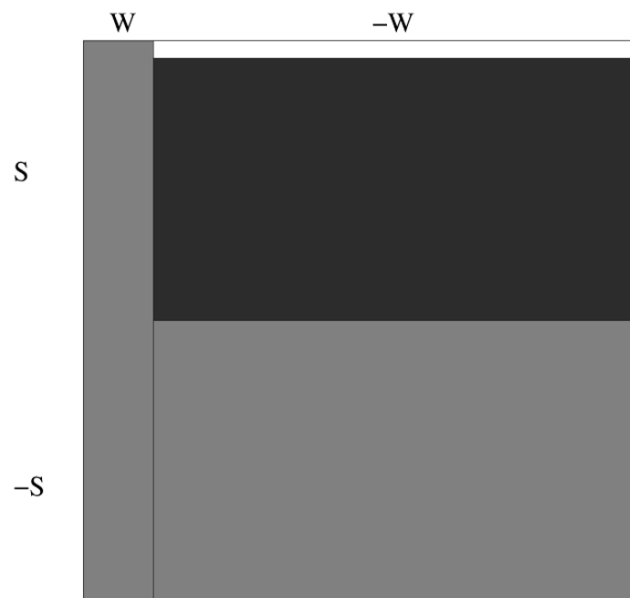


Figure 5. Distribution of the values of (39b).

If values were assigned in this way, they would be distributed as shown in Figure 6. The values of the right-nested conditionals (28) would then come out as in Figures 4 and 5, respectively. This move looks promising, but it is at odds with what Jeffrey (1991) himself assumed about the values at non-antecedent worlds:

I take it that if  $A \rightarrow C$  is to be an *indicative* conditional it must have the same value at all worlds  $w$  where  $A$  is false ...

Jeffrey does not elaborate on this claim, and the example shows that it is to blame for some problems. The shaded lines in (33) above illustrate this vividly: There the value of (30) (i.e., the consequent of (28)) at worlds at which the match is not struck is independent of whether the match is wet or not.

It is not clear yet how an amended assignment along the lines of (40b,c) could come about in a principled way, nor why we should not simply go with the Import-Export Principle. In the next subsection I will discuss two arguments pertaining to these questions, one against the Import-Export Principle and one for an assignment following (40b,c).

### 5.2. *The Import-Export Principle*

Adams (1975) argues that the equivalence of  $\varphi \rightarrow (\psi \rightarrow \chi)$  and  $(\varphi\psi) \rightarrow \chi$  should not hold as a general principle. Consider the special case of  $A \rightarrow$



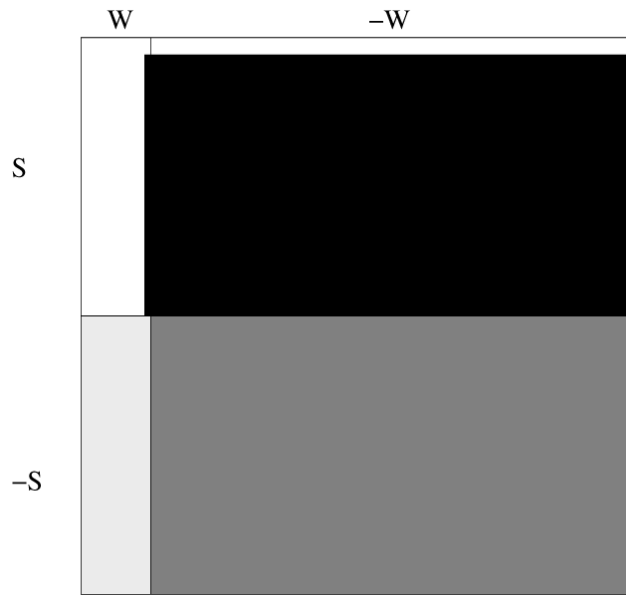


Figure 6. Distribution of the values of (40a) according to (40b) in  $W$  and (40c) in  $\bar{W}$ .

$(B \rightarrow A)$  and  $(AB) \rightarrow A$ . The latter is a logical truth, and so, given to the equivalence, would be the former. But then  $B \rightarrow A$  could be inferred from  $A$  and a logical truth by Modus Ponens.

Gibbard (1981, fn. 17), who defends the Import-Export Principle, is not convinced that this prediction is all that unreasonable. However, it does seem to go against intuitions in examples like Adams'. The prediction is that (41a) and (41b) are equivalent and therefore both are logical truths.

- (41) a. If the match lights, it will light if you strike it.
- b. If you strike the match and it lights, it will light.

While it is undeniable that (41b) is necessarily true, this is far less clear for (41a). Suppose again that the probability that the match lights if struck is 0.82 as above, and consider a situation in which it is very likely that the match is not struck but instead tossed into the camp fire, where it will light without being struck. In this case, the probability of (41a) should clearly be less than 1. The fact that intuitions come apart so tangibly between (41a) and (41b) suggests that the substitution is not applicable across the board as a general logical principle.

One way in which the preceding argument might be challenged is by pointing out that (28) and (41a) differ with respect to the temporal order of the constituents: In (28), the wetness of the match precedes, or is at

least simultaneous with both the striking and the lighting, thus it might be suggested that the value of the conditional at worlds at which it is wet or dry, respectively, should be the one it has *at that time*. This, it could be argued, preserves the effect of the Import-Export Principle in this particular case but not in the case of (41a).

However, while such a patch would fix this particular case, it would not solve the problem: There are other examples where the effect of the Import-Export Principle *should* obtain, but where the temporal order of the constituents is reversed like in (41a). The conditional in (42a) is such an example.

- (42) a. If the coin comes up heads, I will lose if I bet on tails.  
 b. If I bet on tails and the coin comes up heads, I will lose.

Suppose the betting precedes the coin toss and (42a) is asserted prior to the betting. The temporal relations are as in (41b): The outcome of the toss and my winning or losing are determined simultaneously, and the betting precedes both. But unlike (41a), the sentence in (42a) is certainly true and indeed equivalent to (42b).

Thus the order of the reference times is not responsible for the failure of the Import-Export Principle in (41a,b), and so an appeal to temporal relations (alone) is not likely to decide when the Principle should apply.

## 6. COUNTERFACTUALS

In the previous section I argued that the values assigned at worlds at which the antecedents are false are responsible for the intuitively wrong probabilities of conditionals under Jeffrey's definition. In asking what those values *should* be, it is natural to consult intuitions about counterfactuals. In this section I will discuss the latter to the extent that they inform the solution to the above problems. Consider (43b), the counterfactual counterpart of (43a).

- (43) a. If you strike the match, it will light.  
 b. If you had struck the match, it would have lit.

Clearly in the event that the match was not struck, the truth – after the fact – of the counterfactual in (43b) depends on whether it was wet or not. Similarly, in case it was not struck, but lit because it was tossed into the fire, the fact that it lit does not make (43b) any more of a certainty.

Thus intuitions about counterfactuals seem to accord well with what we found the values of the indicative conditional should be at non-antecedent worlds. Harper (1981, fn. 18) presumably also hinted at this connection when he noted about embedded conditionals like (41a) that “the reading of the [embedded] conditional is implicitly subjunctive”.

### 6.1. *Past Predominance*

In similarity-based accounts of counterfactuals in time, the similarity between worlds is to some extent determined by the temporal order in which facts come about. Thomason and Gupta (1981) considered (and ultimately dismissed) a radical version of this principle of *Past Predominance*:

In determining how close [ $w$  at time  $t$ ] is to [ $w'$  at time  $t$ ], past closeness predominates over future closeness; that is, the portions of  $w$  and  $w'$  not after  $t$  predominate over the rest of  $w$  and  $w'$ . (p. 301)

The simplest implementation of this idea is to re-run history from the perspective of an earlier time at which it was still possible for the antecedent to turn out true, then examine, from the perspective of that time, those continuations in which it *is* true. Thus at a world in which the match is not struck, the value of (43a) before the fact should be that of (43b) after the fact. This intuition is often appealed to (Lewis 1979; Tedeschi 1981; Dahl 1997), but it does not always yield the right predictions. It is prone to “undoing” too much of the history of the world of evaluation. While some posterior facts – at times after that of the antecedent – should not bear on similarity, others should.<sup>26</sup>

#### 6.1.1. *Irrelevant Posterior Facts*

A well-known argument for Past Predominance is provided by Fine’s (1975) example (44), presented as a counterexample to Lewis’ “overall similarity” account.

- (44) If Nixon had pressed the button, there would have been a nuclear holocaust.

The sentence is true, even granting that there are worlds at which the button was pressed but no nuclear holocaust ensued, and which are there-

<sup>26</sup> In probabilistic settings, the corresponding assumption that the prior probability of (43a) must equal the posterior probability of (43b) is implemented in *epistemic past* (Adams 1975) and *prior propensity* (Adams 1976; Skyrms 1981) accounts. This view, too, faces similar problems (Barker 1998). Kaufmann (2001c) argues that the right way to think of the relationship is by distinguishing between values and their expectations, just like in the present paper. Then pairs like (43a,b), while not generally *equiprobable*, are *equivalent*.

fore more similar to the actual world than those at which the button was pressed and a nuclear holocaust did ensue. But the former are more similar only at times after that of the antecedent; the intuition that (44) is true suggests that such similarities should be disregarded.<sup>27</sup>

### 6.1.2. *Relevant Posterior Facts*

In contrast, (45), which already appeared in the right-nested conditional (42) above, shows that in some cases posterior facts *do* affect the interpretation of the counterfactual. Slote (1978) attributes (45) to Sidney Morgenbesser; I cite it after Bennett (1984):

At  $t_1$  I bet that when the coin is tossed at  $t_2$  it will come up heads; and in the upshot it does just that; but this is a purely chance event, with no causally sufficient prior conditions. Now consider the conditional

(45) If I had bet on tails at  $t_1$  I would have lost.

I agree with Bennett's judgment and that of his informants that (45) should come out true.<sup>28</sup> However, this is not predicted by Past Predominance as it stands: In worlds which are part of the same history as the actual one at  $t_1$ , the coin toss has not yet occurred. The prior chance of tails is not zero, not even small, hence the counterfactual is wrongly predicted to be false.

Notice that these intuitions are not accounted for by the probabilities of the scenario: Imagine that instead of one fair coin there are two, used depending on my bet. Then a different coin would have been used if I had

<sup>27</sup> Lewis (1979) responded by developing his famous hierarchy of "miracles". Many authors have complained that this notion is too vague to test the predictions of the theory (cf. Nute 1980, 1984).

<sup>28</sup> One reviewer takes issue with this intuition of Bennett's, his informants, and mine. According to the reviewer, the following sentence (a) is true in the situation described, simply due to the fact that the outcome was "genuinely chancy", i.e., that both sentences in (b) were true at  $t_1$ .

- (a) If I had bet on tails I might have won.
- (b) It might land heads. / It might land tails.
- (c) I might have bet on tails and won.

I think it is an open empirical question how easy the reviewer's reading for (a) is to come by. Clearly (c) is true in the situation by virtue of the fact that the outcome was chancy, but things are much less clear for (a). The reviewer's suggestion deserves further investigation; meanwhile, however, I suspect that most speakers would agree that (a) strongly suggests that the outcome somehow causally depended on the bet.

bet on tails. Since it, too, would have been fair, the probabilities are the same throughout; but (45) is no longer intuitively true.<sup>29</sup>

### 6.1.3. *Tichý's Puzzle*

The following story, due to Tichý (1976) and frequently discussed in this connection (Veltman 1985, and others) serves to summarize the points made so far.

- (T1) Consider a man – call him Jones – who is possessed of the following disposition as regards wearing his hat. Each morning he flips a coin before he opens the curtains to see what the weather is like. ‘Heads’ means he is going to wear his hat if the weather is fine, ‘tails’ means he is not going to wear his hat in that case. Bad weather, however, invariably induces him to wear a hat. This morning, ‘heads’ came up when he flipped the coin; furthermore, the weather is bad, so Jones is wearing his hat; and (46) If the weather were fine, Jones would be wearing his hat.

Had the weather been fine, the coin would nevertheless have come up heads. Thus (46) is correctly predicted true under past predominance, since Jones tossed the coin before he knew what the weather was like.

Consider now the following minor variant of the story:

- (T2) This morning, Jones was too impatient to toss the coin first. He opened the curtains and saw that the weather was bad. But he tossed the coin all the same, since doing so is a time-honored habit of his. The coin came up ‘heads’. Later in the day, he is wearing his hat.

In this case, (46) should still be true, but this is no longer predicted by Past Predominance. When Jones found out about the weather, he had not yet tossed the coin. A re-run of history from that point on would involve another toss of the coin, whose outcome might be different. Thus Past Predominance predicts, contrary to intuitions, that if the weather were fine, Jones might not be wearing his hat.

This prediction would be correct, however, if Jones’ decision whether or not to toss the coin depended on the weather, as in the following variant:

- (T3) Jones always opens the curtains first. If the weather is bad, he will wear his hat. If it is good, he tosses the coin. This morning the weather was bad, so he is wearing his hat.

<sup>29</sup> Adams (1975, p. 129) already discussed such cases and pointed out the importance of distinguishing stochastic from causal dependence.

If the weather had been fine, he would have tossed the coin to decide whether to wear his hat, so he might not wear it and (46) is false.

What if he tosses the coin no matter what, but the consequences of the toss (i.e., its effect on John's behavior) depend on the weather? Consider the following variation on Tichý's theme:

- (T4) Jones always opens the curtains and then tosses the coin. In bad weather, the coin decides which of his two bad-weather hats he will wear – 'heads' for blue, 'tails' for black. In good weather, the toss decides whether he will wear his straw hat ('heads') or none at all ('tails'). Today the weather is bad and the coin came up 'heads', so he is wearing his blue hat.

Clearly (46) is once again true in this scenario. If the weather were fine, the outcome would still have been 'heads' and it would, jointly with the weather, have induced Jones to wear his straw hat.

In none of these cases did I assume that the probabilities of heads and tails, if a toss is made, depend in any way on the weather. Notice also that in (T1), (T2) and (T4) what makes the counterfactual true is not the fact that *a* toss was made, but rather the particular "toss token", along with its outcome. For suppose we are told in addition that in bad weather Jones tosses the coin in the living room, whereas in good weather he tosses it on the balcony. Now it is no longer the case that in good weather it would have been the *very same* toss which would have decided, so the possibility that it might have come up 'tails' is again open, and (46) is intuitively false.

## 6.2. *Causal Independence*

Examples like the ones discussed above suggest that if the unraveling and re-running of history is to yield the right predictions about counterfactuals, simple temporal precedence must be augmented with further restrictions on which alternative courses of events are to be considered. Some facts that are settled only after the reference time of the antecedent must be held constant.

What are the facts that should be held constant? Obviously those that would not be different if the truth value of the antecedent had been different. But this is itself a counterfactual. This circularity has plagued theories of counterfactuals ever since it was first made explicit by Goodman (1947). The notion of relative similarity employed by Stalnaker and Lewis is a way of breaking it, or rather sidestepping it, by using "a model theory with room for wired-in answers to various counterfactual questions" (Jeffrey 1991, p. 162).

Another way of avoiding the circularity takes causal relations as basic: Some processes are not causally affected by the antecedent, and their outcomes are carried over in choosing alternative worlds. The relation of being (or not being) “causally affected”, simply an assumption speakers seem to be making, remains unanalyzed and is given as part of the model.<sup>30</sup> Ultimately, therefore, the avoidance of circularity is only apparent, since no analysis of causality seems to be forthcoming that would not in turn refer to counterfactuals. But taking causality as basic seems appropriate if the goal is a semantic analysis of conditionals, rather than a metaphysical analysis of causality.

### 6.3. *What Causal Relations Relate*

To integrate causal relations into the model, some clarifications are in order. First, the causal relations I have in mind hold between event *tokens* rather than *types*.<sup>31</sup> The former is typically described by episodic sentences like (47a), the latter by generics like (47b).

- (47) a. Her lack of exercise caused her heart attack.
- b. Lack of exercises causes heart attack.

Although pairs such as (47a,b) superficially appear to make the same claim, the relations referred to have different properties. It is natural, for instance, to assume that causal relations between event tokens are transitive, asymmetric, and irreflexive; however, the same does not hold for event types (Davis 1988).

Related to this distinction is the second question, how to encode the relation formally. In most probabilistic theories, the formal analog of the claim that *C causes E* is the statement that *C increases the probability of E (under certain circumstances)*, accompanied by the condition that the high probability *E* does not come about in other ways or cannot be explained otherwise. Suppes (1970) applied a definition of this kind indiscriminately to both token and type causation, but subsequent authors came to agree that it is more appropriate for the latter than the former (see Davis 1988 for an overview).

<sup>30</sup> See Mårtensson (1999) and Bennett (2003) for a related and more detailed analysis in terms of causal “trails” and “ramps”, “early departure” and other notions.

<sup>31</sup> This talk of “events” as the carriers of the relation is somewhat loose and sidesteps a number of complications. Hausman (1998) settles on “tropes”, parts or aspects of events, of which he writes: “A trope is a located value of a variable or an instantiation of a property at a place and time” (p. 26). I will continue to speak of events for now, and switch to indicator variables in the formal implementation.

With regard to event tokens, I take causation to be a deterministic relation between events and probabilities: The statement that  $C$  causes  $E$  asserts that the *probability* of  $E$  depends on whether or not  $C$  occurs. To adapt Hausman's (1998) example, "rather than regarding the match's lighting (if it does) as probabilistically caused, one should regard the probability that it lights as deterministically caused" (p. 201). Hausman lines up an array of good arguments for this view. I adopt it without further discussion.

#### 6.4. Implementation

Events, the bearers of causal relations, are not directly represented in a model defined as above, in which worlds have no structure other than a temporal succession of states. However, for each event there is a characteristic function which identifies the set of those worlds in which it occurs. The language may contain a sentence that is true if and only if the event occurs, whose denotation then will coincide with this characteristic function. There is no *a priori* reason to restrict the possible causes and effects to those that can be named, however.

An effect is not an event, but the probability of an event. The event itself, like its cause, is represented by the characteristic function of the proposition that it occurs. The expectation of this function is the chance that the event has of occurring. It can be read off the model: Each world at each time  $t$  is a member of a set  $[w]_t^{\approx}$  of historical alternatives. The chance of the event is  $\Pr(\text{the event occurs} | [w]_t^{\approx})$ .

The relation of causal dependence is defined solely in terms of the variables involved, and separately from the probability distribution (see also Woodward, 2001): A statement of the form  $X$  causally affects  $Y$ , where  $X$  and  $Y$  are random variables, asserts that the *value* of  $X$  causally affects the *expectation* of  $Y$ .

In Definition 13, a subset  $\Phi$  of "causally relevant" variables is singled out from the set of all random variables in the model.

**DEFINITION 13** (Causal chance model). A *causal chance model* is a structure  $\langle W, T, <, \approx, \Pr, \Phi, \prec \rangle$ , where  $\langle W, T, <, \approx, \Pr \rangle$  is a chance model (cf. Definition 11) and  $\langle \Phi, \prec \rangle$  is a set of random variables on  $W$ , where  $\prec$  is a strict partial order.<sup>32</sup>

---

<sup>32</sup> Transitivity is commonly assumed for token causation (Davis, 1988, p. 146), but it only holds for the relation itself, not necessarily for the probabilistic "impulse" transmitted along a causal chain; the latter may be canceled out by the values of intermediary variables; cf. also Hausman (1998, p. 204), Pearl (2000, p. 237).



The idea is that  $\varphi \prec \varphi'$  means  $\varphi$  *causally affects*  $\varphi'$  in the above sense. In the interpretation of conditionals, in fact, it will be more useful to identify those variables which are *causally unaffected* by a given variable. These are all its non-descendants in the causal order:

DEFINITION 14 (Causal independence). Given a causal structure  $\langle \Phi, \prec \rangle$ , for all  $\varphi, \varphi' \in \Phi$ :  $\varphi'$  is *causally independent* of  $\varphi$  if and only if  $\varphi \not\prec \varphi'$ .

As in the case of probabilities, this causal structure can be given a subjective or an objective interpretation. One may speculate that objective chance comes with an “all-embrasive Bayesian net,” a frame “containing all variables needed for a complete description of empirical reality” (Spohn, 2001, p. 167). If, on the other hand, the evaluation in the model is to be a realistic reflection of the interpretations carried out by actual speakers, the set of variables that are considered relevant in any given instance would probably be rather sparse.<sup>33</sup>

### 7. VALUE ASSIGNMENTS

With an explicit encoding of causal dependencies as in Definition 13, we can now take their effect on the interpretation of conditionals into account by making the value assignment sensitive to them. Given an ordered set  $\langle \Phi, \prec \rangle$  of relevant variables as defined above, the Jeffrey-style formula (22) is changed to (48).

$$(48) \quad V(\varphi \rightarrow \psi)(t)(w) = \begin{cases} V(\psi)(t)(w) & \text{if } V(\varphi)(t)(w) = 1 \\ E[V(\psi)(t)|V(\varphi)(t) = 1, [w]_t^\approx, X = X(w)] & \text{for all } X \in \Phi \text{ s.t. } V(\varphi)(t) \not\prec X \\ V(\varphi)(t)(w) & \text{if } V(\varphi)(t)(w) = 0 \end{cases}$$

The condition in (48) reads ‘ $V(\varphi)(t) \not\prec X$ ’ rather than ‘ $V(\varphi)(t) \not\prec X$ ’ because it would otherwise result in a contradiction if  $V(\varphi)(t) \in \Phi$ , since  $V(\varphi)(t) \not\prec V(\varphi)(t)$ .

Recall from the discussion in Section 5 on nested conditionals that in order for the predictions about embedded conditionals to come out right,

---

<sup>33</sup> As we will see, which causal structure is adopted can greatly affect the beliefs about particular conditionals. Therefore the discovery of the “true” (or a better) causal structure may constitute progress even without any new information about relative frequencies.

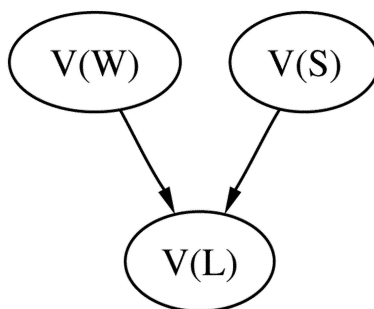


Figure 7. Causal dependencies in (49).

the value of (49a) should be that of (49b) at worlds at which the match is wet, and that of (49c) at worlds at which it is dry.

- (49) a. If you strike the match, it will light.  
 b. If you strike the match and it is wet, it will light.  
 c. If you strike the match and it is dry, it will light.

According to (48), these are indeed the values of the conditional, provided that  $\langle \Phi, < \rangle$  holds the information that the wetness of the match does not causally depend on the striking.<sup>34</sup> Let us see how in this case the distribution of the values of (49) comes out as shown in Figure 6, which I argued above is correct.

The structural dependencies among the relevant variables are simple: Whether the match lights depends both on its wetness and on whether it is struck or not, but the wetness and the striking are causally independent from each other. Adopting the graphical conventions of Bayesian networks, these facts may be represented schematically as in Figure 7. By the above definition, it follows that four different values are assigned to the conditional in (49a), since its value at worlds at which the match is not struck depends on whether it is wet or not. The assignment is as shown in (50). See Figure 6 for illustration.

<sup>34</sup> Notice again that this is orthogonal to any *stochastic* dependence between two: Perhaps you won't strike the match unless it is dry, so your not striking it is evidence that it is wet. But that has no bearing on the question of what would have happened if you had struck it, given that it is dry.

$$(50) \quad V(S \rightarrow L)(w) = \begin{cases} V(L)(w) & \text{if } V(S)(w) = 1 \\ E[V(L)|V(S) = 1, V(W) = 0] & \text{if } V(S)(w) = 0, V(W)(w) = 0 \\ E[V(L)|V(S) = 1, V(W) = 1] & \text{if } V(S)(w) = 0, V(W)(w) = 1 \end{cases}$$

$$= \begin{cases} 0 & \text{if } V(S)(w) = 1, V(L)(w) = 0 \\ 1 & \text{if } V(S)(w) = 1, V(L)(w) = 1 \\ 0.9 & \text{if } V(S)(w) = 0, V(W)(w) = 0 \\ 0.1 & \text{if } V(S)(w) = 0, V(W)(w) = 1 \end{cases}$$

The expectation of these values is the probability of (49a), predicted to be high:

$$(51) \quad P(S \rightarrow L) = 0 \cdot \Pr(V(S) = 1, V(L) = 0) \\ + 1 \cdot \Pr(V(S) = 1, V(L) = 1) \\ + .9 \cdot \Pr(V(S) = 0, V(W) = 0) \\ + .1 \cdot \Pr(V(S) = 0, V(W) = 1) \\ = 0 \cdot 0.09 + 1 \cdot 0.41 + 0.9 \cdot 0.45 \\ + 0.1 \cdot 0.05 = 0.82$$

Turning to the right-nested conditional (52a), its values are now different because the values of its consequent are distributed differently.

- (52) a. If the match is wet, then if you strike it it will light.
- b. If the match is dry, then if you strike it it will light.

Notice in particular the gray area in (53), which highlights the contrast with the problematic (33). As a result of this change, the value at those worlds at which the match is not wet, calculated as the conditional expectation of the values of the consequent over those worlds at which it is wet, is low as well.

$$(53) \quad V(W \rightarrow (S \rightarrow L))(w) \\ = \begin{cases} V(S \rightarrow L)(w) & \text{if } V(W)(w) = 1 \\ E[V(S \rightarrow L)|V(W) = 1] & \text{if } V(W)(w) = 0 \end{cases}$$

$$= \begin{cases} V(L)(w) & \text{if } V(W)(w) = 1, V(S)(w) = 1 \\ E[V(L)|V(S) = 1, V(W) = 1] & \text{if } V(W)(w) = 1, V(S)(w) = 0 \\ E[V(S \rightarrow L)|V(W) = 1] & \text{if } V(W)(w) = 0 \end{cases}$$

$$= \begin{cases} 0 & \text{if } V(W)(w) = 1, V(S)(w) = 1, V(L)(w) = 0 \\ 1 & \text{if } V(W)(w) = 1, V(S)(w) = 1, V(L)(w) = 1 \\ 0.1 & \text{if } V(W)(w) = 1, V(S)(w) = 0 \\ 0.1 & \text{if } V(W)(w) = 0 \end{cases}$$

A diagram of the resulting values appeared in Figure 4. It is obvious from that picture that the expectation of the values will be small. This is indeed the case: The probability of (52a) under the new value assignment is much more in line with intuitions, as (54) shows.

$$\begin{aligned} (54) \quad & P(W \rightarrow (S \rightarrow L)) \\ &= 0 \cdot \Pr(V(W) = 1, V(S) = 1, V(L) = 0) \\ &\quad + 1 \cdot \Pr(V(W) = 1, V(S) = 1, V(L) = 1) \\ &\quad + 0.1 \cdot \Pr(V(W) = 1, V(S) = 0) \\ &\quad + 0.1 \cdot \Pr(V(W) = 0) \\ &= 0 \cdot 0.045 + 1 \cdot 0.005 + 0.1 \cdot 0.05 + 0.1 \cdot 0.9 = 0.1 \end{aligned}$$

On the other hand, the values and their expectations are high for the conditional in (52b). The value assignment is as follows:

$$\begin{aligned} (55) \quad & V(\bar{W} \rightarrow (S \rightarrow L))(w) \\ &= \begin{cases} V(S \rightarrow L)(w) & \text{if } V(W)(w) = 0 \\ E[V(S \rightarrow L)|V(W) = 0] & \text{if } V(W)(w) = 1 \end{cases} \\ &= \begin{cases} V(L)(w) & \text{if } V(W)(w) = 0, V(S)(w) = 1 \\ E[V(L)|V(S) = 1, V(W) = 0] & \\ \quad \text{if } V(W)(w) = 0, V(S)(w) = 0 \\ E[V(S \rightarrow L)|V(W) = 0] & \text{if } V(W)(w) = 1 \end{cases} \\ &= \begin{cases} 0 & \text{if } V(W)(w) = 0, V(S)(w) = 1, V(L)(w) = 0 \\ 1 & \text{if } V(W)(w) = 0, V(S)(w) = 1, V(L)(w) = 1 \\ 0.9 & \text{if } V(W)(w) = 0, V(S)(w) = 0 \\ 0.9 & \text{if } V(W)(w) = 1 \end{cases} \end{aligned}$$

Again, the shaded area marks the effect of the new definition, according to which the values of the conditional  $S \rightarrow L$  depend on whether the match is wet or dry at the world of evaluation. The distribution of these values is as shown in Figure 5.

The probability of (52b), the expectation of the values defined in (55), is accordingly high, as desired:

$$\begin{aligned}
(56) \quad & P(\bar{W} \rightarrow (S \rightarrow L)) \\
&= 0 \cdot \Pr(V(W) = 0, V(S) = 1, V(L) = 0) \\
&\quad + 1 \cdot \Pr(V(W) = 0, V(S) = 1, V(L) = 1) \\
&\quad + 0.9 \cdot \Pr(V(W) = 0, V(S) = 0) \\
&\quad + 0.9 \cdot \Pr(V(W) = 1) \\
&= 0 \cdot 0.045 + 1 \cdot 0.405 + 0.9 \cdot 0.45 + 0.9 \cdot 0.1 = 0.9
\end{aligned}$$

## 8. FURTHER ISSUES

In this paper I dealt with simple and right-nested predictive conditionals and argued that a causal account, which is independently needed for counterfactuals, corrects the problematic predictions of the random-variable approach. I believe that this improved account brings us closer to a unified probabilistic semantics for the various classes of conditionals (and ultimately not just conditionals), but some issues that I did not touch on in this paper need to be addressed before such a theory will take shape. I conclude by summarizing what is left for future work.

As Stalnaker and Jeffrey (1994) showed, the random variable approach can be extended to a language containing conjunctions and arbitrary embeddings of conditionals, including ones with conditional antecedents. This general case requires a more complex model theory, which Stalnaker and Jeffrey adopt from van Fraassen (1976). But the predictions about the probabilities of these more complex sentences are again counterintuitive in some cases (Lance 1991; Edgington 1991). These problems may be overcome by modifying the account along the lines of the proposal in this paper, but considerable complexities are involved in working this out.

The restriction to predictive conditionals and their counterfactual counterparts was, as I mentioned, for the sake of simplicity. Both the problems discussed here and their solution in fact apply as well in the case of epistemic conditionals and subjective probability. But while the solution works for each of these interpretations, a theory in which both are combined requires extra care in setting up the relationship between them.

Finally, I already pointed out that the value assignment I proposed in Section 7 relies on an assumption that cannot be sustained in general: that for each non-antecedent world, a set of antecedent worlds with positive probability can be accessed while keeping constant the values of all variables that do not causally depend on the antecedent. The assignment in

(48) does not define a value if the values of those variables at the world of evaluation are such that the antecedent is not only false but has probability 0. More generally, this paper only imposed *restrictions* on the accessibility of antecedent worlds. That restrictions are needed is clear from the examples; however, assuming that patches of undefined values in the set of historical alternatives are undesirable, the restrictions on accessibility must be counterbalanced in order to ensure that the values of conditionals be defined everywhere if they are defined anywhere.<sup>34</sup>

I leave each of the above issues for another occasion. From a metaphysical perspective, whether the appeal to causal independence truly solves the problem or merely begs the question may be debatable. However, it is interesting to know that once this move is made, the relationship between counterfactual and indicative conditionals becomes very straightforward and predictions about embedded conditionals become sensible. This should be seen in the context of the rising recognition of the role of causality in other areas, notably Artificial Intelligence.<sup>35</sup>

#### ACKNOWLEDGEMENTS

This paper grew out of Kaufmann (2001a). Fragments of it appeared in Kaufman (2001b, 2001c). I thank Stanley Peters, David Beaver, Cleo Condoravdi, Peter Sells, Robert van Rooy, the audiences at SALT XI, the 13th Amsterdam Colloquium, colloquia at Northwestern, Kyoto and Tokyo Universities, the 2001 CSLI LLC Workshop and the NTT Communication Science Lab, and above all the editor and two anonymous reviewers for comments and suggestions. Remaining errors, omissions and misrepresentations are my own. I also gratefully acknowledge the support from the Japan Society for the Promotion of Science (JSPS, Grant #P01063) and

<sup>34</sup> This problem was addressed by Skyrms (1984, 1988, 1994), cited above, in his set selection account. There, too, the values of conditionals are calculated locally within the cells of partitions induced by causal background factors. Cells in which the antecedent does not have positive probability are merged with larger ones in which it does. How exactly this merger proceeds is not very clear; Skyrms does not assume that it is guided by an order on the background factors, as would be natural in the present proposal. Here, the corresponding solution would consist in conditioning on fewer variables in (48).

In Section 2.4, I mentioned that Skyrms' account is in some respects similar to mine. This similarity breaks down in other areas; for instance, for counterfactuals he assumes a prior propensity account, which is vulnerable to Barker's (1998) objections.

<sup>35</sup> Readers familiar with probabilistic methods in knowledge representation will recognize a connection between the approach advocated here and the theory of causal Bayesian networks (Cowell et al. 1999; Spirtes et al. 2000; Pearl 2000, and elsewhere), on which I did not elaborate in this paper.

Satoshi Sato's hospitality at Kyoto University, where much of this work was carried out.

## REFERENCES

- Adams, E.: 1965, 'The Logic of Conditionals', *Injury* **8**, 166–197.
- Adams, E.: 1970, 'Subjunctive and Indicative Conditionals', *Foundations of Language* **6**, 89–94.
- Adams, E.: 1975, *The Logic of Conditionals*, Reidel, Dordrecht.
- Adams, E.: 1976, 'Prior Probabilities and Counterfactual Conditionals', in W. L. Harper and C. A. Hooker (1976), pp. 1–21.
- Adams, E.: 1998, *A Primer of Probability Logic*, CSLI Publications, Stanford.
- Anderson, A. R.: 1951, 'A Note on Subjunctive and Counterfactual Conditionals', *Analysis* **12**, 35–38.
- Aristotelian Society: 1991, 'The Symposia Read at the Joint Session of the Aristotelian Society and the Mind Association at the University of Durham', *The Aristotelian Society*, Supplementary Volume 65.
- Aristotelian Society: 1995, 'Proceedings of the Aristotelian Society', *The Aristotelian Society*, Supplementary Volume 69.
- Athanasiadou, A. and R. Dirven (eds.): 1997, *On Conditionals Again*, Vol. 143 of *Amsterdam Studies in the Theory and History of Linguistic Science*, John Benjamins, Amsterdam/Philadelphia.
- Barker, S. J.: 1998, 'Predetermination and Tense Probabilism', *Analysis* **58**, 290–296.
- Barwise, J.: 1986, 'Conditionals and Conditional Information', in E. C. Traugott et al. (1986), pp. 21–54.
- Bennett, J.: 1984, 'Counterfactuals and Temporal Direction', *The Philosophical Review* **93**(1), 57–91.
- Bennett, J.: 1988, 'Farewell to the Phlogiston Theory of Conditionals', *Mind* **97**(388), 509–527.
- Bennett, J.: 2003, *A Philosophical Guide to Conditionals*, Oxford University Press, Oxford.
- Braithwaite, R. B. (ed.): 1931, *The Foundations of Mathematics and other Logical Essays*, Routledge and Kegan Paul, London.
- Burgess, J. P.: 1979, 'Logic and Time', *The Journal of Symbolic Logic* **44**(4), 566–582.
- Cohen, A.: 2003, 'Probabilistic Approaches to Semantics', in R. Bod, J. Hay, and S. Jannedy (eds.), *Probabilistic Linguistics*, Chap. 9, MIT Press, Cambridge, MA, pp. 343–379.
- Comrie, B.: 1985, *Tense*, Cambridge University Press, Cambridge.
- Comrie, B.: 1986, 'Conditionals: A Typology', in E. C. Traugott et al. (1986), pp. 77–99.
- Cowell, R. G., A. P. Dawid, S. L. Lauritzen, and D. J. Spiegelhalter: 1999, *Probabilistic Networks and Expert Systems*, Springer, New York.
- Dahl, O.: 1997, 'The Relation between Past Time Reference and Counterfactuality: A New Look', in A. Athanasiadou and R. Dirven (1997), pp. 97–112.
- Dancygier, B.: 1998, *Conditionals and Prediction*, Cambridge University Press, Cambridge.
- Davis, W. A. 1988, 'Probabilistic Theories of Causation', in J. H. Fetzer (1988), pp. 133–160.
- Downing, P. B.: 1959, 'Subjunctive Conditionals, Time Order, and Causation', *Proceedings of the Aristotelian Society* **59**, 125–140.

- Dudman, V. H.: 1984, 'Conditional Interpretations of *If*-sentences', *Australian Journal of Linguistics* **4**, 143–204.
- Dudman, V. H.: 1988, 'Indicative and Subjunctive', *Analysis* **48**(3), 113–122.
- Dudman, V. H.: 1994, 'On Conditionals', *Journal of Philosophy* **91**, 113–128.
- Edgington, D.: 1986, 'Do Conditionals Have Truth Conditions?', *Critica* **18**(52), 3–20.
- Edgington, D.: 1991, 'The Mystery of the Missing Matter of Fact', in Aristotelian Society (1991), Supplementary Volume 65, pp. 185–209.
- Edgington, D.: 1995, 'On Conditionals', *Mind* **104**(414), 235–329.
- Eells, E. and B. Skyrms (eds.): 1994, *Probabilities and Conditionals: Belief Revision and Rational Decision*, Cambridge University Press, Cambridge.
- Ellis, B.: 1978, 'A Unified Theory of Conditionals', *Journal of Philosophical Logic* **7**, 107–124.
- Fetzer, J. H. (ed.): 1988, *Probability and Causality*, Vol. 192 of *Studies in Epistemology, Logic, Methodology, and Philosophy of Science*, Reidel, Dordrecht.
- Fine, K. 1975, 'Review of Lewis (1973)', *Mind* **84**, 451–458.
- von Fintel, K.: 1998, 'The Presupposition of Subjunctive Conditionals', in U. Sauerland and O. Percus (eds.), *The Interpretive Tract*, Vol. 25 of *MIT Working Papers in Linguistics*, MIT Press, Cambridge.
- Fleischman, S.: 1989, 'Temporal Distance: A Basic Linguistic Metaphor', *Studies in Language* **13**(1), 1–50.
- van Fraassen, B. C.: 1976, 'Probabilities of Conditionals', in W. L. Harper and C. A. Hooker (1976), pp. 261–308.
- van Fraassen, B. C.: 1981, 'A Temporal Framework for Conditionals and Chance', in W. L. Harper et al. (1981), pp. 365–414.
- Funk, W.-P.: 1985, 'On a Semantic Typology of Conditional Sentences', *Folia Linguistica* **19**(3/4), 365–414.
- Gabbay, D. and F. Guenther (eds.): 1984, *Extensions of Classical Logic*, Vol. 2 of *Handbook of Philosophical Logic*, Reidel, Dordrecht.
- Galavotti, M. C., P. Suppes, and D. Costantini (eds.): 2001, *Stochastic Causality*, CSLI Publications, Stanford.
- Gärdenfors, P.: 1982, 'Imaging and Conditionalization', *Journal of Philosophy* **79**, 747–760.
- Gärdenfors, P. 1988, *Knowledge in Flux: Modeling the Dynamics of Epistemic States*, MIT Press, Cambridge.
- Gibbard, A.: 1981, 'Two Recent Theories of Conditionals', in W. L. Harper et al. (1981), pp. 211–247.
- Goodman, N.: 1947, 'The Problem of Counterfactual Conditionals', *The Journal of Philosophy* **44**, 113–128.
- Grandy, R. E. and R. Warner (eds.): 1986, *Philosophical Grounds of Rationality*, Clarendon Press, Oxford.
- Hájek, A.: 1994, 'Triviality on the Cheap?', in E. Eells and B. Skyrms (1994), pp. 113–140.
- Hájek, A. and N. Hall: 1994, 'The Hypothesis of the Conditional Construal of Conditional Probability', in E. Eells and B. Skyrms (1994), pp. 75–110.
- Hall, N.: 1994, 'Back in the CCCP', in E. Eells and B. Skyrms (1994), pp. 141–160.
- Harper, W. L.: 1976, 'Rational Belief Change, Popper Functions, and Counterfactuals', in W. L. Harper and C. A. Hooker (1976), pp. 73–115.
- Harper, W. L.: 1981, 'A Sketch of Some Recent Developments in the Theory of Conditionals', in W. L. Harper et al. (1981), pp. 3–38.



- Harper, W. L. and C. A. Hooker (eds.): 1976, *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, Vol. 1 of *The University of Western Ontario Series in Philosophy of Science*, Reidel, Dordrecht.
- Harper, W. L., R. Stalnaker, and G. Pearce (eds.): 1981, *Ifs: Conditionals, Belief, Decision, Chance, and Time*, Reidel, Dordrecht.
- Hausman, D. M.: 1998, *Causal Asymmetries*, Cambridge University Press, Cambridge.
- Hughes, G. and M. Cresswell: 1996, *A New Introduction to Modal Logic*, Routledge, Oxford.
- Iatridou, S.: 2000, 'The Grammatical Ingredients of Counterfactuality', *Linguistic Inquiry* **31**(2), 231–270.
- Jackson, F.: 1979, 'On Assertion and Indicative Conditionals', *Philosophical Review* **88**, 565–589.
- Jackson, F.: 1987, *Conditionals*, Basil Blackwell, Oxford.
- James, D.: 1982, 'Past Tense and the Hypothetical: A Cross-Linguistic Study', *Studies in Language* **6**(3), 375–403.
- Jeffrey, R. C.: 1964, 'If', *Journal of Philosophy* **61**, 702–703.
- Jeffrey, R. C. (ed.): 1980, *Studies in Inductive Logic and Probability*, Vol. 2, University of California Press.
- Jeffrey, R. C.: 1991, 'Matter-of-fact Conditionals', in Aristotelian Society (1991), Supplementary Volume 65, pp. 161–183.
- Karttunen, L. and S. Peters: 1979, 'Conventional Implicature', in C.-K. Oh and D. Dinneen (eds.), *Presupposition*, Vol. 11 of *Syntax and Semantics*, Academic Press, New York, pp. 1–56.
- Kaufmann, S.: 2001a, 'Aspects of the Meaning and Use of Conditionals', Ph.D. thesis, Stanford University, Stanford.
- Kaufmann, S.: 2001b, 'Probabilities of Conditionals', in R. Hastings, B. Jackson, and Z. Zvolenszky (eds.), *Proceedings of SALT XI, held at New York University*, CLC Publications, Ithaca, NY, pp. 248–267.
- Kaufmann, S.: 2001c, 'Tense Probabilism Properly Conceived', in R. van Rooy and M. Stokhof (eds.), *Proceedings of the Thirteenth Amsterdam Colloquium*, pp. 132–137.
- Kaufmann, S.: 2004, 'Conditional Truth and Future Reference', semanticsarchive.net.
- Kratzer, A.: 1979, 'Conditional Necessity and Possibility', in U. Egli, R. Bäuerle, and A. von Stechow (eds.): *Semantics from Different Points of View*, Springer, Berlin, pp. 117–147.
- Kratzer, A.: 1991, 'Conditionals', in A. von Stechow and D. Wunderlich (eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung. [=Semantics]*, de Gruyter, Berlin/New York, pp. 651–656.
- Lance, M.: 1991, 'Probabilistic Dependence among Conditionals', *Philosophical Review* **100**, 269–276.
- Lewis, D.: 1973, *Counterfactuals*, Harvard University Press, Cambridge, MA.
- Lewis, D.: 1976, 'Probabilities of Conditionals and Conditional Probabilities', *Philosophical Review* **85**, 298–315.
- Lewis, D.: 1979, 'Counterfactual Dependence and Time's Arrow', *Noûs* **13**, 455–476.
- Lewis, D.: 1980, 'A Subjectivist's Guide to Objective Chance', in R. C. Jeffrey (1980), pp. 263–293.
- Lewis, D.: 1981, 'Ordering Semantics and Premise Semantics for Counterfactuals', *Journal of Philosophical Logic* **10**(2), 217–234.
- Lewis, D.: 1986, 'Probabilities of Conditionals and Conditional Probabilities II', *Philosophical Review* **95**, 581–589.

- Lewis, D.: 1994, 'Humean Supervenience Debugged', *Mind* **103**, 473–490.
- Mackie, P.: 1973, *Truth, Probability and Paradox*, Clarendon Press, Oxford.
- Mårtensson, J.: 1999, *Subjunctive Conditionals and Time: A Defense of the Classical Approach*, No. 10 in *Acta Philosophica Gothoburgensia*, Department of Philosophy, University of Göteborg, Göteborg.
- McGee, V.: 1989, 'Conditional Probabilities and Compounds of Conditionals', *The Philosophical Review* **98**(4), 485–541.
- Mellor, D. H. (ed.): 1978, *Foundations: Essays in Philosophy, Logic, Mathematics and Economics*, Routledge and Kegan Paul, London.
- Mellor, D. H. (ed.): 1990, *Philosophical Papers: F. P. Ramsey*, Cambridge University Press, Cambridge.
- Mellor, D. H.: 1993, 'How to Believe a Conditional', *The Journal of Philosophy* **90**, 233–248.
- Milne, P.: 1997, 'Bruno de Finetti and the Logic of Conditional Events', *British Journal for the Philosophy of Science* **48**(2), 195–232.
- Nute, D.: 1980, *Topics in Conditional Logic*, Reidel, Dordrecht.
- Nute, D.: 1984, 'Conditional Logic', in D. Gabbay and F. Guentner (1984), pp. 387–439.
- Palmer, F. R.: 1983, 'Future Time Reference in the Conditional Protasis: A Comment on Comrie', *Australian Journal of Linguistics* **3**, 241–243.
- Pearl, J.: 2000, *Causality: Models, Reasoning, and Inference*, Cambridge University Press, Cambridge.
- Pendlebury, M.: 1989, 'The Projection Strategy and the Truth Conditions of Conditional Statements', *Mind* **98**(390), 179–205.
- Prior, A.: 1967, *Past, Present and Future*, Oxford University Press, Oxford.
- Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik: 1985, *A Comprehensive Grammar of the English Language*, Longman, London.
- Ramsey, F. P.: 1929, 'General Propositions and Causality', printed in R. B. Braithwaite (1931), pp. 237–255, Meller (1978), pp. 133–151; Meller (1990), pp. 145–163.
- Read, S.: 1995, 'Conditionals and the Ramsey Test', in *Aristotelian Society* (1995), Supplementary Volume 69.
- Reichenbach, H.: 1956, *The Direction of Time*, University of Berkeley Press, Berkeley/Los Angeles, CA.
- Roberts, C.: 1989, 'Modal Subordination and Pronominal Anaphora in Discourse', *Linguistics and Philosophy* **12**, 683–721.
- van Rooy, R.: 1997, 'Attitudes and Changing Contexts', Ph.D. thesis, Institut für maschinelle Sprachverarbeitung (Stuttgart).
- Skyrms, B.: 1981, 'The Prior Propensity Account', in W. L. Harper et al. (1981), pp. 259–265.
- Skyrms, B.: 1984, *Pragmatics and Empiricism*, Yale University Press, New Haven/London.
- Skyrms, B.: 1988, 'Conditional Chance', in J. H. Fetzer (1988), pp. 161–178.
- Skyrms, B.: 1994, 'Adams Conditionals', in E. Eells and B. Skyrms (1994), pp. 13–26.
- Slote, M. A.: 1978, 'Time in Counterfactuals', *The Philosophical Review* **87**(1), 3–27.
- Spirtes, P., C. Glymour, and R. Scheines: 2000, *Causation, Prediction, and Search*, MIT Press, Cambridge, MA.
- Spohn, W.: 2001, 'Bayesian Nets are All there is to Causal Dependence', in M. C. Galavotti et al. (2001), pp. 157–172.
- Stalnaker, R.: 1968, 'A Theory of Conditionals', in *Studies in Logical Theory, American Philosophical Quarterly, Monograph: 2*, Blackwell, Oxford, pp. 98–112.
- Stalnaker, R.: 1970, 'Probability and Conditionals', *Philosophy of Science* **37**, 64–80.

- Stalnaker, R.: 1975, 'Indicative Conditionals', *Philosophia* **5**, 269–286.
- Stalnaker, R.: 1976, 'Letter to van Fraassen', in W. L. Harper and C. A. Hooker (1976), pp. 302–306.
- Stalnaker, R.: 1991, 'Letter to Richard Jeffrey', included in Jeffrey (1991), pp. 178–183.
- Stalnaker, R. and R. Jeffrey: 1994, 'Conditionals as Random Variables', in E. Eells and B. Skyrms (1994), pp. 31–46.
- Strawson, P. F.: 1986, 'If' and '⊃', in R. E. Grandy and R. Warner (1986), pp. 229–242.
- Suppes, P.: 1970, *A Probabilistic Theory of Causality*, North Holland, Amsterdam.
- Tedeschi, P.: 1981, 'Some Evidence for a Branching-Futures Semantic Model', in P. Tedeschi and A. Zaenen (eds.), *Tense and Aspect*, Vol. 14 of *Syntax and Semantics*, Academic Press, New York, pp. 239–270.
- Thomason, R. H.: 1970, 'Indeterminist Time and Truth Value Gaps', *Theoria* **36**, 264–281.
- Thomason, R. H.: 1984, 'Combinations of Tense and Modality', in D. Gabbay and F. Guenther (1984), pp. 135–165.
- Thomason, R. H. and A. Gupta: 1981, 'A Theory of Conditionals in the Context of Branching Time', in W. L. Harper et al. (1981), pp. 299–322.
- Tichý, P.: 1976, 'A Counterexample to the Lewis–Stalnaker Analysis of Counterfactuals', *Philosophical Studies* **29**, 271–273.
- Traugott, E. C., A. ter Meulen, J. Snitzer Reilly, and C. A. Ferguson (eds.): 1986, *On Conditionals*, Cambridge University Press, Cambridge.
- Veltman, F.: 1976, 'Prejudices, Presuppositions and the Theory of Conditionals', in J. Groenendijk and M. Stokhof (eds.), *Proceedings of the First Amsterdam Colloquium [=Amsterdam Papers in Formal Grammar, Vol. 1]*, pp. 248–281.
- Veltman, F.: 1985, 'Logics for Conditionals', Ph.D. thesis, University of Amsterdam, Amsterdam.
- Woods, M.: 1997, *Conditionals*, Clarendon Press, Oxford.
- Woodward, J.: 2001, 'Probabilistic Causality, Direct Causes and Counterfactual Dependence', in M. C. Galavotti et al. (2001), pp. 39–63.

*Linguistics Department*  
*Northwestern University*  
*2016 Sheridan Road*  
*Evanston, IL 60208-4090*  
*USA*