Propensities and probabilities

Nuel Belnap

1028-A Cathedral of Learning, University of Pittsburgh, Pittsburgh, PA 15260, USA

Abstract

Popper's introduction of "propensity" was intended to provide a solid conceptual foundation for objective single-case probabilities. By considering the partly opposed contributions of Humphreys and Miller and Salmon, it is argued that when properly understood, propensities can in fact be understood as objective single-case causal probabilities of transitions between concrete events. The chief claim is that propensities are well-explicated by describing how they fit into the existing formal theory of branching space-times, which is simultaneously indeterministic and causal. Several problematic examples, some commonsense and some quantum-mechanical, are used to make clear the advantages of invoking branching space-times theory in coming to understand propensities.
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Propensities; Probabilities; Space-times; Originating causes; Indeterminism; Branching histories

1. Introduction

You are flipping a fair coin fairly. You ascribe a probability to a *single case* by asserting

> The probability that heads will occur on this very next flip is
> about 50%. (1)

The rough idea of a single-case probability seems clear enough when one is told that the contrast is with either *generalizations* or *frequencies* attributed to populations asserted while you are flipping a fair coin fairly, such as

> In the long run, the probability of heads occurring among flips is about 50%.

(2)

E-mail address: belnap@pitt.edu

Once one begins to study probability theory, however, one rapidly loses ones grip on what it could mean to fasten a probability on an isolated case without thinking of it as part of a group of similar cases—perhaps stretching over time and space. The reason is that the *frequency interpretation* of probability, which is not only as venerable as, but also as attractive as, a California grandmother fresh from the spa, gives a kind of theoretical support for the doctrine of the meaningless of single-case ascriptions of probability, since it speaks always of the relative size of sets, and never of the single case. Popper (1959) articulates and defends the intelligibility of single-case ascriptions. After criticizing the frequency interpretation in well-known ways that I shall not describe, Popper thought to use the idea of a *propensity* to make sense out of applying probabilities to the single case, whereas starting with Humphreys (1985), a number of philosophers have argued in their turn that talk of propensities has not been justified, and in particular that conditional probabilities cannot be construed as propensities. As kind of a third party, I am going to participate in the dispute over propensity theory: I will be defending the idea of identifying propensities in the single case as causal probabilities in the sharply defined sense of branching space-times with probabilities (BSTP).[1] The way in which I am going to mount this defense is by showing, eventually, how a concept of propensities derivable from the literature fits *exactly* into the theory of BSTP. Before I get to that, however, I will be looking at some of the discussion as it has emerged.

My plan is this. In Section 2, I try to make clear the topic by offering as clear a description as I can manage of earlier discussions of propensities as derived from Popper, emphasizing the variety of sometimes incompatible features that have been attributed to them. Then in Section 3, I describe the little bit of probability theory that the discussion requires, and with this as background, in Section 4, I discuss what has come to be known as "Humphreys's Paradox" by means of Humphreys's own examples. In the course of presenting his *objection* to propensity theory, Humphreys introduces some notation for putative propensities; I discuss this new notation in Section 5. Miller (1994, 2002) uses the new notation in order to *defend* the identification of propensities as probabilities, a development some of the details of which I critique in Section 6. With that out of the way, I appear to come to the same conclusion as Miller, my argument of Section 7 depending on a tinker with one of Humphreys's principle examples. Up to this point there has been a great deal of informal argument by all parties. At last I review some key ideas of BSTP, an axiomatic theory, in Section 8, and then in Section 9, I apply BSTP to yet another propensity story (better: an anti-(propensity = probability) story), a story due to Salmon. I suggest that the idea of a propensity is illuminated by identifying propensities with certain causal probabilities—the probabilities of "causae causantes"—in BSTP theory, a theory governed by rigorous definitions and axiomatic principles (no informality). Finally I summarize and review conclusions in Section 10.

2. What is a propensity?

This is the place in this essay where I offer a preliminary idea of propensities. Since for "propensity" no crisp understanding (much less a definition) is available, I assume the

---

[1]BSTP is explored in the following papers listed as references: Belnap (1992, 2003a) (BST), Belnap, Perloff, & Xu (2001) (FF), Belnap (2002a) (DTR), Belnap (2002b)(FBBST), Belnap (2003a; 2003b)(NCCFB), Belnap (2005a; 2005b) (CC), Müller (2002, 2005), Placek (2000a, 2000b).

privilege of directing your attention by listing some loosely characterized features that, in various combinations, have been attributed to propensities. In the end we should have a theoretically sharp account of propensities so we can relate them to single-case causal probabilities, but for the present we have to live with looseness.

1. Popper (1959, p. 68) explains a propensity as "the probability of the result of a single experiment, with respect to its conditions." The rider, "with respect to its conditions," seems essential, but at best the meaning of that phrase is murky. Propensities are taken as characterizing "the whole physical situation," a phrase that also clouds the mind. On the preceding page Popper says that "Every experimental arrangement is *liable to produce*, if we repeat the experiment very often, a sequence with frequencies which depend upon this particular experimental arrangement." It is evident that there is no suggestion here of fastening a probability-of-occurrence onto the experiment itself; hence, in harmony with Humphreys, classical conditional probability is the wrong tool.
   If we take "single experiment" seriously, uniqueness is implied. Each experiment has the conditions that it has. Therefore, as long as we identify the experiment by its spatio-temporal location as well as in terms of all the possible histories with which it is consistent, "its conditions" are uniquely determined by the experiment itself. ("Histories" is a technical term in BSTP, with a precise definition.) The phrase "with respect to its conditions" might be thought redundant, and it might be thought that the phrase "the whole physical situation" calls for the same treatment: BSTP sees a difference here. (1) If "conditions" are in the past, then, as long as we do not waffle, conditions of a single experiment are unique, whether known or not. But (2) the phrase, "the whole physical situation," leaves room for alternate possibilities. BSTP says that if heads come up on a coin-toss, part of the *actual* total situation is that it might have come up tails. (Possibilities are intraworldly in BSTP theory—see Section 8.1—not other-worldly as in Lewis's account.)
2. Popper (1959, p. 37) also characterizes propensities as *dispositional properties of singular events*. Generally dispositions are taken to characterize enduring objects, but that is not appropriate here (the idea is too complicated for an elementary discussion). At this point we want to think about something simpler: the disposition of a singular event to be followed by events of a certain character. Propensities and dispositions, in the required sense, are valenced toward the future (p. 36).
3. That propensities *explain* probabilities is an appealing notion that lies behind a great deal of thought about them. This species of explanation in this case, however, should vividly remind you of the "dormative powers of opium" as a putative explanation of why opium puts one to sleep.
4. Propensities for Popper are *unashamedly indeterminist*, since they "influence future situations without determining them." There being a propensity for such and such may make it likely, but is never a guarantee. That idea suits BSTP.
5. Propensities are objectives (p. 32). Subjective probability theory need not apply.
6. Propensities always, or for the most part, create a time-asymmetric situation. One year Lance Armstrong had a propensity to win the *Tour de France* the next year; but after winning that second year, he does not have a propensity for having won the Tour the first year. Pure probability theory is incapable of making the distinction. You have to have become clear on indeterminist tense logic to appreciate Prior's (1957) surprising

point that this temporal asymmetry nevertheless effortlessly allows that in early 2005, Lance Armstrong had a propensity to *have been born* a seven-time winner of the *Tour de France*. See FF, p. 243, for the short version, and DTR for the long.

7. "Having a propensity…" can be used in perfectly idiomatic English to describe persons, animals, or non-living things (enduring objects), or even characteristics. Examples: "Angry folks have a propensity to act irrationally." Put abstractly, "Anger has a propensity to lead to violence." Also "Dogs have a propensity to bite the hand that feeds them," and "The jar on the left has a propensity to leak." Also "This experimental coin-flipping set-up has more of a propensity to produce heads than it does to produce tails."

   "Having a propensity…" can also describe actions or events, especially in generalizations such as "Taking a pepper spray when setting out to walk the streets of New York has a propensity to lengthen your life," or "Measuring coffee grounds accurately has a propensity to result in a cheerful breakfast."

   These uses of "propensity" in generalizations flow easily. It is not so idiomatic to characterize a particular singular event as having a propensity unless one surrounds the use with considerable context. "I understand that you are going walking on Wall Street first thing in the morning. I suggest that you see to it that you take a pepper spray. That action, should you carry it out, would have a propensity to lengthen your life." Atheoretical philosophers often or always take the use of "propensity" or "disposition" in generalizations as primary, or even exclusively intelligible. I believe that one cannot have a satisfactory theory of generalizations if one has no understanding of individual cases; for this reason, I am ignoring the frequent awkwardness of attributing a propensity in a single case. In this study I shall assume that when we have our theories straight, we can always find some language that is good enough to warrant application of propensities to singular events.

8. *Propensities come in degrees of more or less*, *high or low*, *etc.* In the common sense situations in which we speak of them, we hardly ever assign them numerical degrees. In this respect, as well as many others, *propensities resemble probabilities. That observation encourages us to look for a way to bring probability theory to bear on propensities*.

9. If you have the possibility of carrying out one of several actions, you may have a certain propensity for more than one, or even for each. For instance, "Last night, before bed, I was trying to decide whether to have apple juice or orange juice. Because, however, my propensity for apple juice was the same as my propensity for orange juice, I gave up the internal debate in favor of making myself a chocolate milkshake." In such cases, in spite of linguistic temptation, we *should not identify* "having a propensity for" with "having some propensity for," just as we refrain from confusing "being probable that" with "There is some probability that." In this respect, too, propensities are much like probabilities.

10. Gillies (2000) recommends saving "propensity theory" for "any theory which tries to develop an objective, but non-frequency, interpretation of probability" (p. 808). Gillies eventually suggests that there are two kinds of propensity theories: long-run propensity theories and single-case propensity theories. Gillies plumps for the former. His chief argument against the latter is that they are "metaphysical" in a pejorative sense since speculations concerning them in their non-repeatable glory "cannot be tested against data" (p. 824). It is against this line that Popper is protesting in (1) above. What is

non-repeatable is the event to which the propensity is attached, but this does not stand in the way of taking the singular event to be of a certain repeatable kind that can well be involved in universal (law-like) generalizations. For a homely example, to say that a certain coin is fair can be taken to say that each and every flipping of the coin is fair. One cannot, of course, use the finite amount of evidence at one's disposal to fashion a deductive argument to the conclusion that a particular toss (or sequence of tosses) was fair. But that is just the same-old of scientific methodology.

11. *There is always some causal claim involved in an ascription of propensity.* Salmon (1989) writes as follows:

> Propensities, I suggest, are best understood as some sort of probabilistic causes.

Presence or absence of causality stands as the enormous gap between a mere probability and a propensity. That is a repeated theme of most of the literature, and of this essay. In the theory of causation of Belnap (2005b), causation is analyzed in such a way as to require the underlying idea of concrete transitions (see von Wright, 1963), that is, transitions from event to event; which is another theme that I shall eventually exploit.

So much for some not-necessarily coherent abstract characterizations of *propensities*. There will be more about some of them, but first let us think about standard *probabilities*.

3. Probability theory

The "explanatory question" is, *Do propensities explain probabilities*? As I indicated in (3) above, I fail to be enamored of the suggestion that they do. The "identity question" is this: *Are propensities probabilities*? Having skimmed various characterizations of propensities, let us turn to probabilities in order to sharpen the question. For definiteness, I sketch here rudiments of Kolmogorov probability theory, emphasizing those few parts that are of use in understanding the debate about propensities, just to be sure that we are all talking the same language.

> A *probability space* is a triple $\langle Z, F, P \rangle$, where Z is a non-empty
> and finite "sample space," F is the boolean algebra of all subsets
> of Z, P is a function from F into the interval $[0, 1]$, $P(\emptyset) = 0$,
> $P(Z) = 1$, and for all $A, B \in F$, if $A \cap B = \emptyset$ then $P(A \cup B)$
> $= P(A) + P(B)$. (3)

The definition does not allow for infinite probability spaces. As far as the discussion goes in this essay, finitude suffices to illustrate the crucial points.

Users of this calculus are generally happy to call each $A \in F$ (so that according to (3) A is formally a set), an "event," or perhaps a "proposition," or to speak of the members of A as events, and in either case to help themselves informally to the language of occurrence, often tensed. Thinking of A as a property, it is common to speak using (1) "A occurred five times," or (2) "five instances of A occurred," or (3) various nominalizations such as "some of the occurrences of A will be occurrences of B", or (4) "the chances that A will occur." Readers accustomed to such usages should explicitly keep in mind that I shall reserve "event" for certain concrete occupants of our world, which may actually occur, but also

may only possibly occur.[2] The dominating association is with Lewis, who persuasively describes the difference between the actual and the possible as perspectival rather than intrinsic. (The fundamental difference from Lewis is that BSTP postulates just one world; see Section 8.2.)

An event in BSTP theory, whether possible or actual, is as concrete as a Los Angeles freeway. I also say (in adjectival form) of an event e that it may "possibly occur"; that is a short form for the more properly tensed form "e is occurring or did occur or will occur, or might be occurring or might have occurred or might occur"; and even that litany leaves out events that are space-like related to me, which have the feature that it might be that they have occurred. Sometimes I use "possible event" with exactly the same meaning. It is critical that every possible event is in a perhaps complicated space-time-like causal relation to me-here-now. I have in mind the ancestral of "*either* e_1 is in the causal past of e_2, *or* vice versa." So as not to lengthen this essay unduly, I ask that you consult the BST papers if you wish a rigorous and I trust persuasive explanation.

For our purposes, the most significant defined concept of probability theory is that of "conditional probability":

$$P(A|B) =_{df} \frac{P(A \cap B)}{P(B)}, \tag{4}$$

where $A \cap B$ is set intersection and B is presumed non-empty. (Here and everywhere we ignore the awkwardness of keeping track of division by zero.) There are many proposed English readings of $P(A|B)$. Here is one found on the net:

> $P(A|B)$ is to be read in English as "the probability that A *will*
>
> *occur* given that B *has occurred*." $\tag{5}$

Those tangled tenses are at best difficult to penetrate. To avoid the complications of tense (by no means always a commendable aim), the following generic reading would appeal to many:

> the probability of A given B. $\tag{6}$

For practical application, it often does not matter how you read conditional probability in sloppy old English; that is the nature of atheoretical practicality. The tensed English reading (5) just will not matter in the thick of scientific discovery or testing.

It is often good informal practice, because safest, to think of $P(A|B)$ as the probability $P(A)$ of A with respect to the probability space obtained from the entire sample space, Z, of (3) by reducing Z to just its subset B. That is accurate, but hard to read. No wonder the frequency interpretation is so attractive! On that interpretation, in the simplest cases $P(A|B)$ is simply the proportion of A's among the B's as given by (4).

Argle-bargle aside, the formal feature of the probability calculus on which Humphreys (1985) rests his argument for the thesis (9) below that "propensities cannot be probabilities" is that in its theory of conditional probability, the probability calculus permits a Bayes-like inversion of condition and conditioned. "Inversion" does not, of course, mean that $P(A|B) = P(B|A)$; "Inversion" means only that $P(A|B)$ and $P(B|A)$ are equally grammatical and understandable. Perhaps the simplest theorem of the probability

---

[2]This is spelled out in detail in publications listed as references that are authored by any of the following: Belnap, Perloff, Xu, Weinstein, Müller, and Placek.

calculus exhibiting such an inversion is this:

$$P(A|B) \times \frac{P(B)}{P(A)} = P(B|A). \tag{7}$$

This is a consequence of taking (4) as the definition of conditional probability in terms of absolute probabilities. It is more persuasive (?), however, to follow Popper in restricting the language of the abstract theory by omitting absolute probability terms such as "$P(A)$," taking as well-formed only notation for conditional probabilities "$P(A|B)$." In that case, we have Bayes's own theorem, which involves an inversion without invoking absolute probabilities:

$$P(B|(A \cap C)) = \frac{P(A|(B \cap C)) \times P(B|C)}{(P(A|(B \cap C) \times P(B|C)) + (P(A|(\overline{B} \cap C)) \times P(\overline{B} \cap C))}. \tag{8}$$

I am using the overline here as the sign of set complementation relative to the sample space Z of (3). The thing to note is that on the left of (8) there is $(B|A)$ (with "$\cap C$" as extra context), while in the first upper term on the right there is the inversion $(A|B)$ (also with "$\cap C$" as extra context). And that is enough or more than enough probability theory.

4. Humphreys's Paradox

Now that we have our propensities and probabilities (Sections 2 and 3), we must deal with Humphreys (1985), which initiated discussions of the relations between the two. Humphrey uses "chance" for an objective phenomenon to be contrasted with degrees of rational belief or confirmation. For a perhaps contentious reason, I avoid this use of "chance": The usage so often serves as a bludgeon with which to flatten an incompatibilist without responsible argument. My own view, expressed for example in *FF*, is indeed compatibilist, but moves in the polar opposite direction: I take it that agency presupposes indeterminism, but not "chances" in the everyday sense that connotes mindlessness, or at least lack of deliberation. The New Compatibilism, which, as a co-author of FF, I espouse, identifies a minimal sense of indeterminism that does not involve contrast with such heavy ideas as "laws" or "scientific theories." Minimal indeterminism is the doctrine that there are occasions in this world of ours on which there is more than one alternative for the future, and that an occasion of deliberation is a paradigm example.

Putting aside the word, "chances," Humphreys argues that probability theory in its present form cannot serve as a true theory of propensities. His essay was clearly the first to recognize problems in understanding propensities in terms of probabilities, putting his conclusion with a provocatively crisp thesis:

Propensities cannot be probabilities. (9)

Humphreys always has in mind that propensities are conditional, so that propensities are usefully symbolized by $pr(A|B)$, in contrast with standard conditional probabilities, written in the standard notation of probability theory as $P(A|B)$. Conditional propensities are tied to English by Humphreys with the equivalence,

$pr(A|B)$ is "the propensity for A to occur, conditional on the

occurrence of B" (p. 560). (10)

As noted in Section 3, conditional probabilities, in standard probability theory, are generally given the English reading, (6) "the probability of *A*, given *B*," and virtually always interdefinable with absolute probabilities, either via (4) when absolute probabilities are primitive, or via $P(A) =_{df} P(A|\text{Tautology})$ if conditional probability expressions are primitive. Do note that $P(B)$ is, in either case, an absolute (unconditional) probability, which must certainly be a horse of an altogether different color.

Humphrey's principal lemma for (9) is that *Bayes theorem fails for propensities*. I should not even need to remind you of Bayes theorem here, for the following reason: Arguments against it are "so clearly directed against inversion principles that any considerations involving other parts of the calculus seemed to be quite separate" (p. 567). This is a big claim since it calls into question either the much-worked-over foundations of applied probability theory or the intelligibility of the idea of a propensity. Humphreys's (2004) brief article on Salmon reports that even after many, many years of conversations between the pair, Humphreys himself believes that he and Salmon were divided on propensities. Humphreys thinks of Salmon as more attracted to frequencies than to propensities.[3] Humphrey's own belief is that

> it is causal relevance that is important for single-case propensities,
>
> not statistical relevance.   (11)

Humphreys concludes that

> the theory of probability [is] an inappropriate constraint on any
>
> theory of single-case propensities (p. 945),   (12)

where single-case propensities are objective, causal probabilities. You should expect the following.

- I will endorse (9), provided "probabilities" are understood as standard conditional probabilities with no causal element in their constitution.
- I will argue *against* (12).
- I will heartily support the thesis, *it is causal relevance that is important for propensities*.
- But not causality alone; I will also suggest that standard Kolmogorov probabilities are of critical importance for propensities. I will eventually explain how in BSTP causality and probability cooperate.

But all of that comes later, in Section 8. It behooves us first to look at the current state of the topic.

4.1. Conditional propensities

Many of the properties attributed to propensities, as listed informally above in Section 2, make them sound rather like probabilities with a causal flavor, as noted by those who think about propensities, which is doubtless why Popper and others have taken it to be a good idea to *identify* propensities as probabilities. One of Humphreys's most significant

---

[3]See p. 942. Note that Humphreys, in this paper, uses "chances" as synonymous with "single-case propensities." For consistency, when quoting Humphreys (2004), I shall feel free to replace his preference, "chances," with my preference, "single-case propensities."

contributions was to argue that, important as is the suggestion, maintaining this identity thesis is difficult or impossible. I will go over some of the argument, because, as indicated, I plan eventually to argue that it makes theoretical sense to identify propensities with probabilities of causal transitions in branching space-times.

As exhibited in (10), Humphreys introduces the use of $pr(A|B)$ for conditional propensities such as (7) and (8). Humphreys tells a story in which a Bayesian calculation gives the wrong answer.

4.2. Humphreys's photon argument

The story concerns a photon. (I take certain liberties with the story, in the belief that they are irrelevant to its point.) Fig. 1 tells of a photon that has just been emitted in a laboratory. The photon either impinges ($= I$) on a half-silvered mirror with propensity p, or it drifts off somewhere. If it impinges on the mirror, then it acquires a fixed propensity q to be transmitted ($= T$) straight ahead through the mirror and onto a detector, and also of course a companion propensity of about $1 - q$ to wind up reflected ($= R$) off to the left. Fig. 1 spells this out in a two-dimensional spatial diagram, looking at the apparatus from above. Although I expect that the figure is sufficiently intelligible, I will go over it again after bringing in a bit of Humphreys's notation.

4.3. Extended notation showing time-dependence

Humphreys introduces an extended notation in place of $pr(A|B)$ of (10) in order to reflect the time-dependence of propensities:

$pr_{\tau_i}(A_{\tau_j}|B_{\tau_k})$ is to be read in English as *the propensity at time τ_i*

for A to occur at time τ_j, conditional upon B occurring at time
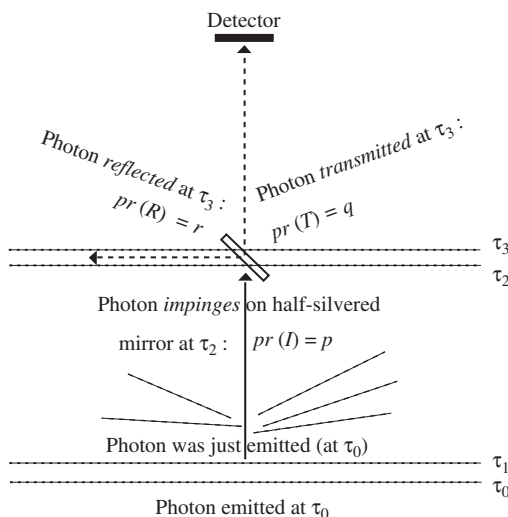
τ_k (Humphreys, 1985, p. 561). (13)

Fig. 1. A nanosecond in the life of a photon. I = impinged, R = reflected, T = transmitted.

Later I will mention some respects in which the notation is inadequately explained, but for now that will not hold us up. Humphreys codifies in this notation the story and the assumptions it presupposes (pp. 559–563).[4]

The first four items on the list below, namely, (i)–(iii) and (CI), are intended as local assumptions, governing just the emit–impinge–transmit set-up that is pictured spatially in Fig. 1. The rest, (TP) and (MP), express general principles of probability theory, including conditional-probability theory. The hypothesis we are testing is whether a theory of propensities can serve as an interpretation of conditional-probability theory. On that hypothesis, those six assumptions should hold good of conditional propensities. Humphreys (1985), however, derives a contradiction, and enters this as an argument that classical probability theory does not give a correct account of conditional propensities; that is, it can be that $pr(A|B) \neq P(A|B)$.

Letters and times are as follows. (The times are to be in the order of their names: $\tau_0 < \tau_1 < \tau_2 < \tau_3 < \tau_4$.)

τ_0 is the last moment before emission. (Think Dedekind cut.)

τ_1 is a time just after emission. By τ_1 it is settled whether or not the photon has been emitted.

τ_2 is the time of impingement/no impingement. I_{τ_2} is the (possible) event of the photon impinging upon the mirror at time τ_2, and $\overline{I}_{\tau_2}$ (which is used in (iii) below) is the (possible) event of the photon failing to impinge on the mirror at τ_2.

τ_3 is the time of transmission/no transmission (noting that no transmission = reflection, R_{τ_3}). T_{τ_3} is the (possible) event of the photon being transmitted through the mirror at time τ_3. $\overline{T}_{\tau_3}$ is the (possible) event of the photon failing to be transmitted through the mirror at τ_3.

Humphreys characterizes the situation in the notation of (13).

Assumptions 4-1 (*Humphreys's assumptions for reductio*). (i) $pr_{\tau_1}(T_{\tau_3}|I_{\tau_2}) = p$, $p > 0$. [The propensity at τ_1 for T to occur at τ_3 conditional upon I occurring at τ_2 is p.]

(ii) $pr_{\tau_1}(I_{\tau_2}) = q$, $q > 0$. [The propensity at τ_1 for I to occur at τ_2 is q.]

(iii) $pr_{\tau_1}(T_{\tau_3}|\overline{I}_{\tau_2}) = 0$. [The propensity at τ_1 for T to occur at τ_3 conditional upon $\overline{I}$ occurring at τ_2 is 0.]

(CI) $pr_{\tau_1}(I_{\tau_2}|T_{\tau_3}) = pr_{\tau_1}(I_{\tau_2}|\overline{T}_{\tau_3}) = pr_{\tau_1}(I_{\tau_2})$. ["The propensity for a particle to impinge upon the mirror is unaffected by whether the particle is transmitted or not."]

(TP) $pr_{\tau_i}(A_{\tau_j}) = pr_{\tau_i}(A_{\tau_j}|B_{\tau_k}) \times pr_{\tau_i}(B_{\tau_k}) + pr_{\tau_i}(A_{\tau_j}|\overline{B}_{\tau_k}) \times pr_{\tau_i}(\overline{B}_{\tau_k})$. [A version of the principle of total probability is assumed for propensities.]

(MP) $pr_{\tau_i}(A_{\tau_j} \cap B_{\tau_k}) = pr_{\tau_i}(B_{\tau_k} \cap A_{\tau_j}) = pr_{\tau_i}(A_{\tau_j}|B_{\tau_k}) \times pr_{\tau_i}(B_{\tau_k})$. [The "multiplication principle" is assumed for propensities.]

---

[4]I have simplified a little in ways that I think do not matter here, but it is hard to be sure. The most significant departure is this: Humphreys (1985) uniformly includes "a complete set of background conditions B_{τ_1}," which I uniformly omit in order to improve readability at the cost of giving a false appearance of countenancing absolute propensities. If you transform each of my "$pr_{\tau_i}(A_{\tau_j}|C_{\tau_k})$" into Humphreys's "$pr_{\tau_i}(A_{\tau_j}|C_{\tau_k} \cap B_{\tau_i})$," you will return to the land of verbatim. In addition, I have tinkered a little with notation, for uniformity. In so doing, I may have blundered in replacing Humphreys's use of concatenation as in "$C_{\tau_k}B_{\tau_i}$" with an explicit sign "∩" of intersection; since e.g. B_{τ_i} seems to be sentential: When read in English, it may be that concatenation here signifies conjunction rather than intersection. The quote about background conditions and assumptions occurs on p. 561.

Humphreys derives an inconsistency from these assumptions by straightforward argument in probability theory (p. 562). I am going to discuss, however, something more basic.

5. What does the extended notation mean?

In this section I suggest that the extended notation $pr_{\tau_i}(A_{\tau_j}|B_{\tau_k})$ is, given what we have been told to date, less than luminous. The English reading of $pr_{\tau_i}(A_{\tau_j}|B_{\tau_k})$ that Humphreys gives in (13) helps, but in certain respects falls short of a completely rigorous explanation. Rigor is of course just one ideal; it has to compete with other ideals, and indeed must usually give way. It nevertheless seems certain that the reading (13) is not given in language that bears its structural meaning on its face. In contrast, probability theory writes "$P(A) = q$," where A is a set from a boolean algebra of sets F presupposed as background, and where P is an appropriate function defined on all of F and taking values in the interval [0, 1]. "A" is *said* to name an "event," but common as it is, that is the loosest of talk, as will be granted by any probability theorist when in a good mood. What is clear is that since the domain of P is a boolean algebra of sets, the place of A can be taken by the result of any ordinary set-theoretical construction, such as $A \cap B$ (intersection), $\overline{A}$ (set-theoretical complement), $A \cup B$ (union), etc. The theory evidently also invites conditional probability expressions "$P(A|C)$," since the expression is fully defined by $P(A \cap C) \div P(C)$.

Propensity theory writes, as in (13), "$pr_{\tau_i}(A_{\tau_j}|C_{\tau_k})$." My suggestion at this point is a weak one: It is not that (13) could not be given a meaning sharp enough to support a proof, but only that it has not been given one yet. It is not enough to give a holistic meaning to (13); the *parts* of (13) must themselves be given independent meaning since in (MP), for example, the meaning of the sentence depends on being sure that the various sign-designs bear the same meaning in every occurrence. So we need to know the categorematic meaning of e.g. A_{τ_j}. It seems inescapable that it is sentential or propositional, some syntactic variation of "A occurs at τ_j." But then can we any longer be speaking of singular events? On p. 561, however, we are offered the example,

> "the event of a photon impinging upon the mirror at time τ_j," (14)

which looks as if A_{τ_j} is a named singular event rather than anything sentential or propositional. It may snap to your mind that this is all cavil; but there is more to it than that. There are the notations "AB" and "$A_{\tau_1} \cap B_{\tau_2}$" and "$I_{\tau_2}T_{\tau_3}$" (p. 562). These notations need to be rigorously explained if we are to be able to prove something using them. And if one writes "$I_{\tau_2}T_{\tau_3}$," one should be prepared to say whether propositional conjunction or set intersection is intended, and if the result is another entity of type A_{τ_1}, and if so, which one.

I take it that the suspicious tone of the foregoing remarks is justified, at least to an extent, by the detailed work in delivering a rich account of events and propositions in FF and the BSTP papers listed in "References," work that offers an alternative. It also needs to be said, however, that all parts of "tense logic" become sensitive to subtleties in the presence of indeterminism and need to be treated with delicacy. Consider that the first inkling of a language proper to "branching time" was the insightful (but casual) representation of indeterminism in Prior (1957), and that only in Thomason (1970, 1984) was tense logic accurately and rigorously adapted to indeterminism. Thomason himself let

his theories lie fallow until 1984, and then yet another batch of years elapsed before much further work was done on the philosophical side of branching-time theory. See FF.

6. Miller: extended notation and time reversal

Miller (1994), who adheres to the view that propensities *are* probabilities, suggests the following reading of $pr_{\tau_i}(A_{\tau_j}|B_{\tau_k})$ as a replacement for (13).

> the propensity of the world at time τ_i to develop into a world
>
> in which A comes to pass at time τ_j, given that it (the world at
>
> time τ_i) develops into a world in which B comes to pass at the
>
> time τ_k (p. 113; lettering has been systematically altered). (15)

The most comforting aspect of the Miller reading is that it seems that something is provided to *have* the propensity at time τ_i, namely, "the world." The least comforting is that what is provided is enormous: the entire world. If we took this seriously, we should be giving up all hope of a theory of *local* propensities, which, while far from a *reductio*, comes too quickly to be satisfying. Note especially that the Miller reading of $pr_{\tau_i}(A_{\tau_j}|B_{\tau_k})$ is clearly and definitely inconsistent with the fundamental ideas of special relativity. The idea of a universal time for our world will not do. I shall come back to this later; until then I shall ignore special relativity. At this point I wish to stress that on Miller's causal analysis of propensities in (15), all the indeterminism is stuffed into the single time τ_i, way down at the bottom. In contrast, Humphreys's reading (13) requires *two* indeterministic initials; you can see in Fig. 1 that there is one indeterministic initial at τ_0 and another at τ_2. On the Miller hypothesis, which, because it restricts itself to only the single indeterministic episode, cannot apply to Fig. 1, it is dead right that Bayes-type inversion is no challenge; for example, in Fig. 2 we will see later that whenever $pr_{\tau_1}(R_{\tau_2}|G_{\tau_3}) = p$ makes sense, so does its Bayes-inspired inversion $pr_{\tau_1}(G_{\tau_3}|R_{\tau_2}) = q$.

First, however, it is essential to put in place a "metaphysical" change suggested by considerations in the theory of branching space-times: *The locus of the indeterminism cannot be a time.* It must be a time-filling (so to speak) concrete event, what Thomson (1977) terms a "super-event", that has alternatives. That is because there must be alternative ways of "filling" the same time, so that what is absolutely required for propensity theory is the idea of an event. We do not need to ask if the (possible) event "has
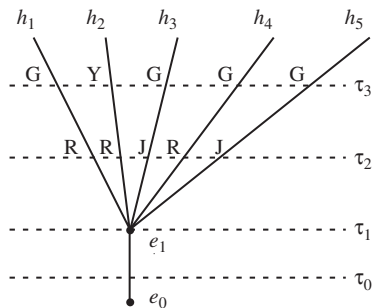

Fig. 2. Bayes inversion acceptable.

actually occurred or will occur'' (Humphreys, 1985, p. 560). The conclusion is that we must be able to say that a family of events can consist in pairwise inconsistent alternatives for occurring at the same time. This is not mysterious, just hard to say: In Fig. 1, each of "reflect" and "transmit" can occur at τ_3 (but of course their joint occurrence at τ_3 is not possible). Here is a homely example. It is 4:00 p.m. A ball is bouncing down the mountain. It might be in any of many places at 4:05. It is certain, however, that it cannot be in two places at the same time. If by 4:05 it has bounced northward, it will be faced with certain possibilities, whereas if at that time it has bounced in a southerly direction, the possibilities open to it will be entirely different. In the first case, the ball may have a strong propensity to continue, at 4:05, to the north on a green path, whereas in the second case it may have a strong propensity to travel, at 4:05, westward on a red path.

That is already a lot of words, but still not enough. There is here an essential if subtle point. We have just seen that the ball may be endowed with alternative, radically distinct, and inconsistent strong propensities at 4:05. Since it is impossible that the ball have both strong propensities at 4:05 (although from an earlier point of view each strong propensity was possible),

> *It makes no sense to index propensities with times, in the fashion*
> *of Humphreys and Miller. Propensities must be indexed to*
> *possible events, not times.* (16)

This is spelled out in detail in BST, and further elaborated in later papers. In any case, however, (16) strongly suggests that we need to change the subscripts on *pr* from names of times τ_1 to names of possible events e_1. According to BST theory, however, we do not need to change the language of outcomes: Let R_{τ_2} be read as "Red occurs at time τ_2," and G_{τ_3} as "Green occurs at time τ_3." Our argument will therefore be that whenever $pr_{e_1}(R_{\tau_2}|G_{\tau_3}) = p$ is a sensible alternative for occurring (or not) at time τ_1, so is its inversion $pr_{e_1}(G_{\tau_3}|R_{\tau_2}) = q$. The critical point recorded in this choice of syntax is that, for a sane expression of propensity, we need the *outcome events* involving *G* and *R* to be characterized only generically, by properties, times, histories, etc., but we need e_1, the *initial event* itself, in all its glorious concrete individuality. In short:

> *Time-reversal itself causes no problems for Bayes, as long as what*
> *you reverse are the times of   only outcomes; when you make that*
> *reversal, you must not touch the initial event where the propensity begins.* (17)

Let us check this with an unrealistic but simple assignment of propensities. With Laplace in mind, and only for ease of calculation, we will assign e_1 equal propensities toward each transition to a "history," that is, a possible course of events[5] h_i ($1 \leqslant i \leqslant 5$) to which e_1 belongs, as pictured in Fig. 2. (Events belong to histories, but times do not belong to histories. Also, to anticipate with an idea and notation from BSTP introduced in Section 8.4, in BSTP propensities in the first instance attach neither to point events nor to their outcomes, but rather to "basic transitions" such as $pr(e_1 \rightarrowtail \pi_{e1}\langle h_i \rangle)$ ($1 \leqslant i \leqslant 5$) in Fig. 6, and derivatively to non-basic transitions such as $pr(e_0 \rightarrowtail \pi_{e_0}\langle h_i \rangle)$.) Now ask first: What is the

---

[5]From this point we begin to use the BSTP word "history," defined in Section 8.2.

proper value of p in the time-switched $pr_1(R_{\tau_2}|G_{\tau_3}) = p$, and of q in $pr_1(G_{\tau_3}|R_{\tau_2}) = q$? As long as one has equipropensity with respect to histories, the recipe for calculation is easy. Equipropensity with respect to each history is absurd, but the pretense sometimes, as here, has a point, as long as you maintain awareness of the pretense.

1. Let g be the number of G's expressing alternatives for time τ_3, that is, let $g = 4$. More accurately, you should be counting sets of histories rather than G's, but since all our sets here are unit sets, that baggage is not needed on this trip.
2. Let $r = 3$ since three histories to which e_1 belongs make R true at time τ_2.
3. Let s be the number of histories that feature both G at time τ_3 and R at time τ_2, namely, $s = 2$.
4. Finally, calculate that $p = (s \div g) = (2 \div 4) = \frac{1}{2}$, and $q = (s \div r) = \frac{2}{3}$.

Therefore, what? Recall that in the example, all the occurrences of red $= R$ came at an earlier time that the occurrences of $G =$ green. Nevertheless, the Humphreys–Miller formulas make good "scientific sense" regardless of whether the expression is $pr_{e_1}(R_{\tau_2}|G_{\tau_3 2}) = p$ or $pr_{e_1}(G_{\tau_3}|R_{\tau_2}) = q$.

Once everything is set up with causal care, that calculation just follows the probability-theory definition of "conditional probability."[6] (You may have noticed that Fig. 2 is a "modal" diagram, without any representation of space, whereas Fig. 1 is a "spatial" diagram without any representation of modality.)

We find support in noticing that when Gillies (2000) comes to Humphreys's Paradox, the general character of his "solution" lies in showing by examples that the reversal at issue has to do only with non-propensity probabilities such as, in the easiest case, one obtains by considering proportions. There is no call to invert the propensity claim.

The example of Fig. 2 suggests that the language of "proportions" is more suitable than talk of "conditions," and since we are (I recommend) giving up awarding a causal gloss to conditional probabilities, a good deal of confusion can likely be avoided in this way. So I recommend at least in the easiest cases the following as a long-winded, whenever-you-have-the-time-and-the-energy type of reading. (The reading (18) presupposes, in addition to finiteness, that event B_{τ_i} is later than e_1 on at least one history to which point event e_1 belongs, that $e_1 \leqslant \tau_j$, and that $\tau_i \neq \tau_j$.)

$pr_{e_1}(A_{\tau_j}|B_{\tau_i})$ should then be read as *the proportion of cases (histories)*
through e_1 in which A_{τ_j} and B_{τ_i} both occur among all the
cases (histories, courses of events) in which B_{τ_i} occurs. (18)

(18) is a faithful reading of $pr_{e_1}(A_{\tau_j}|B_{\tau_i})$ since that proportion, when available, accurately represents the probability, from the perspective of e_1, of A_{τ_j} given B_{τ_i}.[7] It has, I think, in total agreement with Miller, nothing to do with any direct propensity relation between A and B, in spite of looking so much like it should.

---

[6]This example would have misled if you had inferred that "counting histories" is typical. There are, of course, uncountably many histories. Instead, you count *outcomes* (sets of histories) of which a given initial may have finitely many.
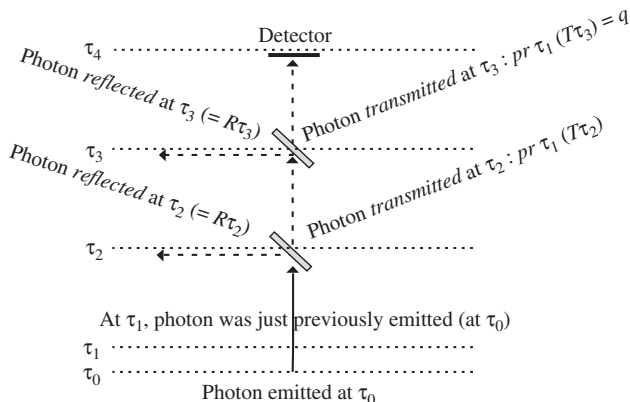
[7]See footnote 6.

Fig. 3. Two half-silvered mirrors yield two occasions of indeterministic transmission or reflection.

7. Two mirrors

It behooves us carefully to return to Section 4.2 in order to consider (iii) of Assumptions 4-1: Even without time reversal, something seems to call for amendment. We can more easily see what might be "off" if we change the example a little. In the original "one mirror" situation as pictured in Fig. 1 and described in (i)–(iii) of Assumptions 4-1, T_{τ_3} is sharply characterized, as is also the alternative to T_{τ_3}. That is, the transmission at τ_3 is equally as sharp as reflection at τ_3. The side-description tells us that exactly one of the possible events transmission at τ_3 or reflection at τ_3 can happen. That is a clear-cut feature of the one-mirror example, which is visually reenforced in Fig. 1 by the labeling. In contrast, the alternative to I_{τ_2} in the one-mirror case, that is, the alternative to impingement at τ_2, is unlabeled in Fig. 1 and described in a soft, indefinite way by using the overline in Assumptions 4-1(iii), presumably as a kind of negation signifying "non-impingement occurs at τ_2" or perhaps "impingement does not occur at τ_2." These candidates express operations on sentences or propositions, which are a very different kettle of ontology from events. Humphreys does not explain the overline as applied to events. In the meantime our skepticism is called for, since it is proper to doubt that good theory can be made from the idea of forming the negation or complement of a singular event.

I am going to change the one-mirror example of Fig. 1 by adding a second half-silvered mirror. I trade in "impingement vs. non-impingement" at τ_2 (as in Fig. 1) by adding a second mirror that allows for an alternative that is *exactly* as sharp as that given for τ_3, namely, "transmission vs. reflection" at τ_2. The new experiment, or at least my spatial picture of it in Fig. 3, makes the alternative to T_{τ_2}, that is, the alternative to transmission at τ_2, the positively and narrowly described event "reflection at τ_2."[8] I will call this the "two-mirror" situation or example. Using now "R" for "reflection," we may use "R_{τ_2}" and "R_{τ_3}" as labels in the two-mirror example in perfect correspondence to Humphreys's notation for transmission in the one-mirror example of Fig. 1. As a side description of the

---

[8]In the spatio-temporal Fig. 3, I am reverting to taking *times* instead of *events* as the loci of indeterminism. The only reason for this change is that Fig. 3 is already overcrowded.
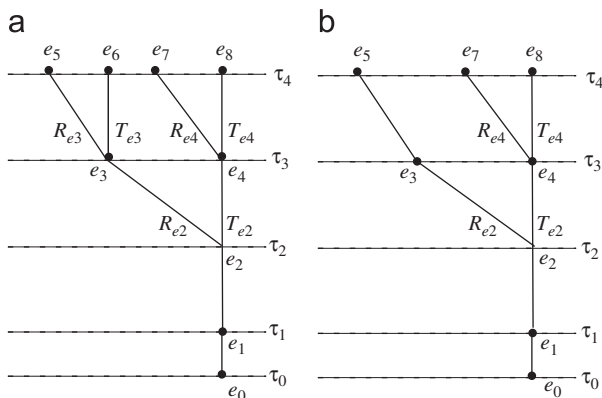
Fig. 4. Two causal (or temporal-modal) diagrams for two half-silvered mirrors.

two-mirror set-up (not drawn in the diagram), we indicate that exactly one of transmission or reflection can occur at τ_2, and exactly one of those alternatives at τ_3. The causal situation is such that at each of τ_2 and τ_3 exactly one of transmission and reflection can occur. But, and this is a big "but", that is not happily expressed by an unexplained symbol suggesting any of logical complement or set-theoretical complement, or negation applied to properties. In the changed example, we do not have to deal with the problem of giving a simple meaning to "does not impinge."

Here is the difficulty seen in the two-mirror example of Fig. 3. The notation allows four possible combinations: $T_{\tau_2}T_{\tau_3}$, $T_{\tau_2}R_{\tau_3}$, $R_{\tau_2}T_{\tau_3}$, $R_{\tau_2}R_{\tau_3}$. Humphreys's account that goes along with the notation, however, is not so generous in what it allows. In accord with Assumptions 4-1(iii), and with very little more than common sense, the last two combinations are forbidden: If reflection at τ_2, then neither transmission nor reflection at τ_3. How could there be, given that the photon has left the playing field? Therefore, we must have one of the first two combinations: either $T_{\tau_2}T_{\tau_3}$ or $T_{\tau_2}R_{\tau_3}$. In either case, we would have T_{τ_2}. This proves that the photon is *guaranteed* (from the perspective of the photon at time τ_1) to be transmitted at τ_2. This is absurd.[9]

The conclusion appears to agree with Humphreys: Standard probability theory will not do for propensities. It is notable, however, that *probabilities did not enter my version of the argument at all.* All the trouble lay in the prior causal analysis. The two diagrams of Fig. 4 are intended to show you the difficulty and how to repair the damage. To appreciate the diagrams, you need to distinguish in your mind (1) spatial or spatio-temporal diagrams from (2) modal or temporal-modal diagrams. As drawn here, spatio-temporal diagrams such as Fig. 3 do not themselves give out modal information, although of course one can convey just about anything one likes by adding words to a diagram. Contrariwise, modal or temporal-modal diagrams such as Fig. 4 have no diagrammatic features indicating spatial relations, although, again, one can jam in spatial information by annotation. Diagram A of Fig. 4 lays out a causal (or modal) representation of the same phenomenon that is spatially pictured in Fig. 3. In both kinds of diagram, up–down keeps track of time;

---

[9]This argument follows the structure of the horse-racing example of Müller (2005), which is used to draw a BSTP conclusion more positive than the one I draw here.

in Fig. 3 left–right indicates a spatial relation, whereas Fig. 4 uses left–right to help indicate inconsistency (for example, left–right separated event-representations that are on the same vertical level must be inconsistent, like "was reflected a moment ago" vs. "was transmitted a moment ago").

Keep in mind that Humphreys is exploring indeterminism, so that it is reasonable to let the various upward-traveling lines represent various possible courses of events. The e_i are concrete events in our world, idealized to point events in order to avoid irrelevant confusion. We are asking for propensities at point event e_1 in Diagram A, which occurs at the time τ_1 of (deterministic) emission, before the action gets started. At e_2, which occurs at time τ_2, indeterminism sets in: From the perspective of our origin at e_1, there are two possibilities for the immediate future (Dedekind style) of e_2, when the photon strikes the first half-silvered mirror: (1) T_{e_2} = transmitted, or (2) R_{e_2} = reflected. Following, in Diagram A, the possible transmission T_{e_2}, we arrive at time τ_3, at which time the photon strikes the second mirror (a possible event that we may call e_4), and is either reflected, R_{e_4}, or transmitted, T_{e_4}. At τ_4 the detector passively registers what has happened. According to Diagram A, exactly one of four point events will occur at the detector at time τ_4; that is, e_5–e_8 are alternative possibilities for the detector at τ_4. Keep in mind that there is no representation of space in these modal diagrams: The horizontal dimension represents alternative possibilities, not spatial separation. Turning now to the critical left portion of Diagram A, at e_2 (which occurs at time τ_2), there is reflection R_{e_2}, and then at event e_3 (which occurs at time τ_2) there are two possibilities for the future, reflection R_{e_3} and transmission T_{e_3}, possibilities that would be detected at τ_4 as either e_5 or e_6, respectively.

Before we go further, we can use Fig. 4 to repeat with emphasis that changing from times τ to possible events e as individualizing subscripts (so to speak) is, once you take indeterminism seriously, not a matter of whimsy. The difference is that an *event*, e, has an absolutely unique place among the exciting events of *Our World*, whereas a time, τ, has a unique place only in the temporal order. Given e, "the past" (of e) is a unique course of events or happenings leading up to e. For example, in Fig. 4B, imaginatively suppose you yourself to be now located at event e_3. Then a particular reflection event, namely, R_{e_2}, would be in your actual past. Furthermore, when situated at e_3 (at time τ_3 in Diagram B), your future at τ_4 would be determined as e_5, if situated at e_4, whose time is exactly the same as that of e_3, both your actual past and your future of possibilities would be decidedly different. In contrast, if your imagination supposes only that you are located at the *time τ_3*, *it makes no sense* (without specifying an event) to ask,

"What actually happened immediately after time τ_2?" (19)

The status of R_{e_2} and T_{e_2} is exactly the same: Each is a possible event that *might have* occurred after time τ_2; neither is singled out as what *did* occur in the past of τ_3. In the same way, given you are at e_3 (and so at time τ_3), your future is determined as e_5 at time τ_4; but if you are located at e_4 (and so also at time τ_3), e_5 is not a real possibility in your future at τ_4.

I now return to a consideration of Humphreys's Assumption 4-1(iii). First a minor (even optional) point. An event, in the sense needed here, is not subject to replication. Each event marks a unique piece of our world. This is the reason that I drop the "background conditions" for which Humphreys calls; an event, e, already uniquely determines its past, so that adding a description of it, such as Humphreys's B_{τ_i} (p. 561) is redundant. (Proviso:

You are not tempted to play around with mind- or language-involving phrases such as "under a description.")

The very neatness with which Diagram A was drawn and described has the power to mislead. The left part of Diagram A, including the off-stage accounts of the meanings of the T's and R's, is, I think, nonsense. This time, for variety, I will argue this by going through the easy steps with explicit attention to propensities. Since R_{e_3} cannot follow R_{e_2} (a reflected photon is not (in this set-up) ripe for further reflection), it must be that $pr_{e_1}(R_{e_3}|R_{e_2}) = 0$; and by a parallel argument $pr_{e_1}(T_{e_3}|R_{e_2}) = 0$. So since (to restate) both combinatorial possibilities at e_1 for what can happen next when R_{e_2} is "given" have a propensity of zero, it better be that $pr_{e_1}(R_{e_2}) = 0$. Otherwise, it would be possible to arrive at e_3 via R_{e_2} (recall that $R_{e_2} = $ reflection at e_2), and then have an absolutely empty future, whatever that could mean. But if this reasoning is just, then all the weight of propensity (as seen from e_1) must go on the right side of Diagram A; which is to say, $pr_{e_1}(T_{e_2}) = 1$. Absurd.

You replace Diagram A with Diagram B if you forbid representation of events that on combinatorial grounds cannot occur. (We can put that more rigorously, but for now let us just go with the pictures.) In Diagram B you immediately see that there are only three possible courses of events, not four as in Diagram A. Either you have transmission–transmission (at e_2, e_4), or transmission–reflection (at e_2, e_4), or simply reflection (at e_2, followed perhaps by a "go directly to e_3 and thence directly to e_5; do not pass e_4"). Red herring to be avoided: There are certainly four fine-grained *propositions* relevant to τ_3 in Diagram B, including R_{e_3} (the photon is reflected at e_3) and T_{e_3} (the photon is transmitted at e_3). Whether fine-grained or course-grained, however, these are impossibilities, and *should not be taken into consideration* when your topic is objective possibilities and probabilities.

8. BSTP: Branching space-times with probabilities

BSTP should provide a good home for propensities as probabilities. BSTP began as just BST, a theory of objective *possibilities* in our relativistic world. It took *initial-event to outcome-event transitions* as fundamental, its only primitives being the idea of a point event and a representation of *Our World* as the set of all possibly occurring point events. The theory stepped gingerly into the realm of causality, but without an explicit theory of causation, in order to say something simple and intelligible about correlations in the EPR family of apparently weird goings-on in quantum mechanics. From this platform, under the presumably false but certainly entertaining proviso that there are no EPR-like phenomena, BST, still with the same primitives, offered an explicit theory of causation, taking transitions from initial-event to outcome-event as both caused and causing. The most recent work, largely represented by Weiner & Belnap (2006) and Müller (2005), has shown how probabilities may be taken to fit into our relativistic and indeterministic world, thus yielding an *explicit axiomatic theory of causal probabilities* in the guise of probability distributions on *causae causantes*. How can such a structure fail to provide a candidate home for propensities?

I begin by describing the branching space-times theory, BST, of *Our World* and its events. There are a number of detailed branching space-times publications available; for that reason, I can be brief here. You may well wonder why I do not *altogether* skip a discussion of events. After all, the probability theorist *always* starts with a field of

"events." Well, yes, but "event" in probability theory is not even jargon, it is a code word. Events for the objective probabilist are plain old set-theory sets, so that the operations the theorist performs on these sets neither needs nor wants additional structure. In that sense, probability theory contains no theory of events, and indeed asks its users to interpret "event" in just about any way they can while still remaining sensible in how they attach probability-numbers to "events." The theory of propensities, on the other hand, is intended to refer to events in the concrete sense, characterized by an antecedent theory of causation and of causal probabilities. To spring the theory on you first instead of last, a *BSTP propensity*, which is measured like a probability, attaches to certain ordered pairs of concrete events, locatable in *Our World*. Which "kinds" of event-pairs can accept propensities is important to BSTP propensity theory, a topic to which we shall return. In the end, when you attach a propensity-symbol, to a pair of names of concrete, unrepeatable possible events A and B, you are saying (I use personal terms for an impersonal idea),

> *If you were ever to get to event A, from that standpoint you could*
> *truly assert that the measure of the propensity for B to occur is*
> p $(0 \leqslant p \leqslant 1)$.
> Or put impersonally,
> *From the standpoint or perspective of event A, event B has a*
> *propensity to occur of p $(0 \leqslant p \leqslant 1)$.* (20)

That is intended to convey only a rough, preliminary idea of "propensity." For one thing, both recipes suppress the requirement that propensities-to-occur apply to future possibilities, not to past facts.

 The remainder of this section is devoted to explaining some useful jargon concerning branching space-times. Our entire approach is by way of "event" concepts; the chief contrast is with enduring substances. There are many "concepts" of events, often involving things or persons. I need a much more abstract idea. I am thinking of a representation of our world as made up out of small, local events, and at the lower limit, point events. An event, however, is not a place—time nor a collection of such; it is a happening. It has a time and a place, which partly describes its locus in our world, but that is not enough to confer uniqueness. An event has a causal past history, and a causal future of possibilities. Nor is a past history a mere array of times and places; a past history consists in concrete events. A point of view that can accurately label these gone-by events as "past" is entitled also to label them as "actual", a category that BSTP takes as just as perspectival as "past." Furthermore, unless you are a beknighted metaphysical determinist, since there are many possibilities for the future, you are going to have to face up to the fact that the concept of "event" has a modal character. Its "could be" is part of its nature. BSTP goes farther: Space-like related events generally fall into the same basket as future events. That is, events of both sorts are perspectivally "possible" rather than "actual." Naturally, as speakers of English, we all tend to use "possible," "actual," etc. as *non*-perspectival adjectives. Rightly or wrongly, however, I have no patience with those who wish to make metaphysics out of this observation. Similarly, "actual" and "possible" are used epistemically, but not in this paper. In BSTP theory these concepts are both objective and perspectival—when tensed.

8.1. Possible point events e, the causal order <, and Our World

"Event" is at worst hopelessly ambiguous and at best vague. That will not do for careful theorizing. There are big events and little events, with the smallest being possible point events. They are supposed to be *ideally* small, in strict analogy to Euclidean points or Newtonian instants of time or the physicists' mass points. A possible point event, e, is as concrete as your actual eye-blinking, has a definite relationship to this very eye-blinking, and is equally related to all other point events that from your point of view are either actual (such as those in your past) or really possible (such as those in your future of possibilities—BST does not deal with imaginary possibilities). Point events are small, and at the other end of the scale we come to *Our World*, which contains each and every possibly occurring point event as a member. It is a matter of taste which one takes as primitive; I am in the habit of (1) taking *Our World* as primitive, (2) characterizing it as a non-empty set, and then (3) defining "e is a possibly occurring point event" as equivalent to $e \in Our\ World$.[10] (I will use—and have used—"point event" as a stylistic variant of "possible point event," which in turn is to be synonymous with the more accurate phrase, "possibly occurring point event.")

Our World may be a set, but it is also a connected whole. What, one might ask, is the connector between (possible) point events? The connector is an adaptation to indeterminism of what relativity theorists and cosmologists, for example, call "the causal ordering," a binary relation symbolized by "$<$," which is the second (and last, until we come to probabilities) primitive of BSTP theory. In application, one may read "$e_1 < e_2$" as either "e_1 is in the causal past of e_2," or "e_2 is in the future of possibilities of e_1." Axioms tell us that $<$ is a strict partial order, so that its companion "$\leqslant$" is a partial order on *Our World*. There are just a few more axioms, most (but not all) of which are easy, for example: density, no "last" point events, existence of infima of lower bounded chains.

8.2. Histories, consistency, categorizing events

Additional concepts and axioms require the critical idea of "history," which in turn depends on "directedness." A subset E of *Our World* is *directed* iff whenever two (possible) point events e_1 and e_2 belong to E, so does some point event e_3 that is in the future of possibilities of both: $e_1 \leqslant e_3$ and $e_2 \leqslant e_3$; or, in equivalent prose, e_1 and e_2 both lie in the causal past of e_3. A *history*, h, is a maximal directed subset of *Our World*, and a set of point events E is *consistent* iff E is a subset of some history h. A history, h, is a complete course of possible point events, although perhaps "course" is the wrong word. The chief warning is that you should not picture h as linear; a history permits its members to be separated either time-like or space-like, or both.

There is a generic sense of *event*, namely, a non-empty and consistent set of point events. Most useful, however, are various species and subspecies of events. Each of the following defined concepts is general enough to be widely useful, and particular enough to earn its keep.

A *chain* is a non-empty and linearly ordered event; an *outcome chain*, O, is a properly lower bounded chain; and an *initial chain*, I, is a properly upper bounded chain (counting a point event e as a special case).

---

[10] "You mean to tell me that *Our World* is a *set*—an abstract entity—that has, for instance, a determinate number of members?" Would I tell you something like that?

It is easy to see the following critical point: a chain event can begin (come to be) iff it is an outcome chain, and a chain event can finish (pass away) iff it is an initial chain.

These ideas may be extended to more complicated initials and outcomes. An *outcome event* (or just "outcome"), **O**, is a non-empty set of outcome chains all of which begin in some one history. It turns out that we can bring the idea of "initial" down a level, identifying it as a generic event: An *initial event* (or just initial) **I** is a non-empty subset of some one history. A full and accurate study (as for example the account of "funny business" in FBBST) would have to keep track of outcomes and initials some of which are chains and some of which are not, but for what we are doing with propensities, and for the level of precision at which we aim, we need to keep only rough track of these distinctions. Of absolute import, however, is the idea of a "transition." Science and its philosophers generally speak of a transition between "states," but the BSTP set of ideas is more concrete—and less subject to waffle—than that. In the spirit of Russell's "at-at" theory of motion, and adapting from the definition of "transition" of von Wright (1963), a *transition* is simply an ordered pair of a (concrete) initial event and a (concrete) outcome event such that the outcome event begins after (possibly immediately after) the initial event finishes (see Belnap, 1999). von Wright takes transitions (defined somewhat differently) as events, and so in endorsement I shall speak of *transition events*. A transition event is an ordered pair of initial and outcome, and therefore disallows a simple answer to "Where (in *Our World*) is it?". As notation, $e \rightarrowtail O$ is the transition from point event e to chain outcome O. Although it bears a fancy name and involves fancy notation, $e \rightarrowtail O$ is nothing but an ordered pair. (Theoretical considerations also support a type of transition, $e \rightarrowtail H$, where H is a set of histories doing duty as a proposition.)

Example. A choice is a paradigm of a transition event, and thus an event without a simple location. A choice must have an initial that finishes while it is still open to incompatible resolutions, and it must have an outcome whose very occurrence testifies that the matter at issue has already been settled. The choice itself, as an entity, is the transition from the open initial to the settled outcome; any other view leads, I believe, to severe muddle, if not outright Zenoesque contradiction.

The idea of history is essential to the definition of "space-like related": e_1 SLR e_2 if neither $e_1 \leqslant e_2$ nor $e_2 \leqslant e_1$, but nevertheless e_1 and e_2 are consistent (there is a history, h, to which both belong). Histories branch in their own way, as governed by the following axiom, called the *prior choice principle*:

> Let an outcome chain O be a subset of history h_1 and excluded
> from history h_2; then there is a point event e that is in the past
> of O and maximal in $h_1 \cap h_2$. See Fig. 5.   (21)

The point e is called a *choice point*: e and all point events in its causal past belong to both h_1 and h_2, whereas every pair of point events $\{e_1, e_2\}$ with $e < e_1 \in h_1$ and $e < e_2 \in h_2$ is inconsistent. Fig. 5 illustrates a simple case. Time-like goes up and down, space-like goes left and right, and the dotted lines are tracks of the light-speed limitation on velocities. Obviously when left–right represents space-like separation, you cannot in the same figure let it represent alternative histories, and vice versa. In Fig. 5, the picture of each history already uses left–right to indicate space-like separation. For this reason, you should not, on pain of confusion, use left–right also to indicate the modal idea of the splitting of
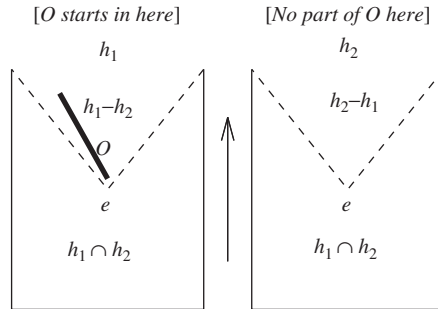
Fig. 5. Prior choice principle: how histories divide.

histories. Indeed, there seems to be no purely graphic method for representing in the same diagram both space-like information and splitting information. Yes, Fig. 5 represents both kinds of information at once, but does not confine itself to graphics. It is critical to Fig. 5 that there are *annotations*. Two examples: (1) The labeling of the two history-pictures as "h_1" and "h_2" ensures that we have no need of their left–right arrangement. (2) The occurrence of "e" in both h_1 and h_2 helps to represent how those two histories split.[11]

Since you cannot overlook the overlap between the histories h_1 and h_2 that constitute the model of *Our World* illustrated in Fig. 5, perhaps this is a good place to make it explicit that according to BSTP, there is just one "world" in Lewis's sense (everything is connected to everything by an external relation). Histories have something in common with Lewis worlds, but that histories do not overlap is a key Lewis doctrine, whereas there is no doubt that the overlapping of histories is an *essential* part of BSTP. The matter is thoroughly discussed in FF.

8.3. Concrete transitions, prior causal loci, causae causantes, and inus conditions

Every chain in every outcome has, by previous definition, a proper lower bound, and therefore, by postulation, a greatest such. An outcome **O** is *immediate* if all such contained chains share the same proper lower bound, say e, in which case one writes $inf(\mathbf{O}) = e$. In this case there is no room between e and **O** that would permit causal lines from the past to slide into **O** without passing through e, so that one may say that the transition $e \rightarrowtail \mathbf{O}$ is *immediate*. An immediate transition can supply no causal account of its outcome other than itself; but it can sometimes supply causal information about other transitions. Thus, let $t_1 = e_1 \rightarrowtail \mathbf{O}_1$ be an arbitrary transition. e_0 is a *prior cause-like (or causal) locus for* t_1 if what can happen at e_0 might make a difference to whether t_1 occurs. And $t_0 = e_0 \rightarrowtail \mathbf{O}_0$ is a *causa causans* of t_1 just in case one immediate outcome of e_0 is consistent with $\mathbf{O}_1$, whereas the rest are not. We can then show that t_0 is an *inus* condition (*i*nsufficient but *n*on-redundant part of a *n*ecessary and *s*ufficient condition) of t_1 (see CC for the proof as well as for a general theory of causation).

---

[11]For an example of space-like information without modal information, see Fig. 1; and for modal information without space-like information, see Fig. 2.

8.4. Probabilities in branching space-time: the WBM analysis

Weiner & Belnap (2006) and Müller (2005), in spite of being independent and offering distinct results and perspectives, benefited to a certain extent from pre-publication conversations; it is proper to think of their work in terms of mutual assistance. One fundamental idea of the "WBM" construction is that you look to the set $cc(\mathbf{I} \rightarrowtail \mathbf{O})$ of *causae causantes* (certain basic transitions; see CC) as containing all and only the information needed for assigning a propensity to $\mathbf{I} \rightarrowtail \mathbf{O}$ (p. 491). There are now two (equivalent) ways to proceed. Each basic transition $e \rightarrowtail H$ among those *causae causantes* has its own *local* causal alternatives $e \rightarrowtail H_1$ based on keeping its initial e and letting its basic outcome H_1 vary over all the basic outcomes of e. Müller develops more "global" causal transitions by piecing together those basic-transition causal alternatives to the maximum extent that they can consistently be fit together. These maximal consistent sets of basic causal alternatives are the *sample space*, A, underlying one way of calculating the propensity belonging to $\mathbf{I} \rightarrowtail \mathbf{O}$. By taking the family of all subsets of the sample space, you inevitably get the boolean algebra, F, that must underlay the causal-probability field of sets in the context of which giving a propensity to $\mathbf{I} \rightarrowtail \mathbf{O}$ makes inescapable sense (recall that we are thinking finite). From this point of view, one has only to add a standard Kolmogorov measure, μ, defined on the algebra, F, and yielding values in the interval [0, 1]. Since the unit set $\{cc(\mathbf{I} \rightarrowtail \mathbf{O})\}$ is bound to be an element of the algebra, F, that is enough. One may now define the *causal-probability space* appropriate to $\mathbf{I} \rightarrowtail \mathbf{O}$ as $\langle A, F, \mu \rangle$.

A second way to proceed, exploited in Weiner & Belnap (2006), derives the wanted propensities $pr(\mathbf{I} \rightarrowtail \mathbf{O})$ from a distributed measure $\mu = \{v_e | e$ is a prior causal locus of $\mathbf{I} \rightarrowtail \mathbf{O}$ and v_e is a probability measure on the boolean algebra, π_e, of all immediate outcomes of e, and thus a measure (the very "same" measure) defined on each of the basic transitions $e \rightarrowtail O$. Define a *propensity space* for $\mathbf{I} \rightarrowtail \mathbf{O}$ as a pair $\langle T, v \rangle$, where T is the set of all *causae causantes* of $\mathbf{I} \rightarrowtail \mathbf{O}$ together with, for each prior causal locus e, the set of all basic transitions $e \rightarrowtail H$; and where v maps each member of T into the interval (0, 1), subject only to the restriction that for each prior causal locus, e, that figures in T,

$$\sum_{h \in H_{(e)}} v(e \rightarrowtail \pi_e \langle h \rangle) = 1.$$

That is, from the perspective of e, something must happen.

The second way, which begins with a separate, local, and independent measure for each set $e \rightarrowtail \pi_e$, has the virtue of firm and direct rooting in the idea of propensity. It quickly leads to correct intuitive judgments in easy cases, cases in which we obtain the propensity of $\mathbf{I} \rightarrowtail \mathbf{O}$ by multiplying together all the propensities of all its *causae causantes*.

On the other hand, the relation of the second way to global Kolmogorov probabilities is less obvious; it is surely not made clear in Weiner & Belnap (2006). The second way nevertheless turns out to be conceptually equivalent to the first way, as one would hope regarding a pair of ways one of which is top-down and the other bottom-up.

9. Application to corkscrew story

Salmon told (at least) two "inversion stories," one, the can opener story, in Salmon (1984), and one, the corkscrew story, in Salmon (1989). With these stories he subscribed to

the position that propensities "make sense as direct probabilities ..., but not as inverse probabilities (because the causal direction is wrong)" (p. 88). Salmon furthermore expressed the view that "probabilistic causality is fraught with difficulties" that such as Reichenbach (1956), and Suppes (1970) do not overcome. I agree with both verdicts, while still endorsing the BSTP theory of causal probability as a sensible account of some of the deepest properties of causal probabilities. BSTP is sufficiently different from any worked-out theory in the literature that it needs to be evaluated by itself, untainted by association with other theories.

My plan is to retell the corkscrew story with the help of BSTP, leaving the reader to judge whether BSTP helps. Here is how Salmon (1989) put the matter.

> Imagine a factory that produces corkscrews. It has two machines, one old and one new, each of which makes a certain number per day. The output of each machine contains a certain percentage of defective corkscrews. Without undue strain, we can speak of the relative propensities of the two machines to produce corkscrews (each produces a certain proportion of the entire output of the factory), and of their propensities to produce defective corkscrews. If numerical values are given, we can calculate the propensity of this factory to produce defective corkscrews. So far, so good. Now, suppose an inspector picks one corkscrew from the day's output and finds it to be defective. Using Bayes's theorem we can calculate the *probability* that the defective corkscrew was produced by the new machine, but it would hardly be reasonable to speak of the *propensity* of that corkscrew to have been produced by the new machine (p. 88).

In retelling the story, I shall rely on an annotated BSTP diagram, namely Fig. 6, which you should put before you. Each h_j $(j = 1, \ldots, 5)$ in that figure is a history. e_0 is a reference point event before all the action. The ellipses are machines M_i $(i = 1, 2)$, with e_i the (idealized) point event of production by M_i. e_3 is the choice point for the Inspector (when he has a choice). D_i $[G_i]$ is a defective [good] corkscrew produced by machine M_i. $\{d_i, g_i$ $(i = 1, 2)\}$ represent the two possible immediate outcomes of each machine choice point e_1, e_2, whereas $d_i^*(i = 1, 2)$ represent the two possible pickings by the Inspector from Box-β at e_3. If two areas have the same label, they are identical; otherwise, not. The boxes α and β are only receptacles. In the story that Fig. 6 illustrates, machines M_1 and M_2 (the gray ellipses) make different numbers of corkscrews per day, and also have different propensitiesfor making, on each run, a defective corkscrew.

To illustrate a BSTP story about many physical objects requires drawing many tracks through the space-time of each history, which is a pain. For this reason, with the expectation that it will not make any difference, and because this is in part a test of making sense of single-case propensities, I am going to have each machine make just one corkscrew per day. (This expository strategy also reduces the temptation to measure propensities in terms of frequencies.) The propensity of each machine to turn out a defective corkscrew on a given day is given numerically in the interval [0, 1]; say M_1 has a .4 propensity of making a defective corkscrew, and consequently a .6 propensity for making a good one. Let the propensities for M_2 be .1 defective and .9 good. In BSTP these propensities qualify, respectively, four of the relevant basic transitions (each expressed in two ways) having one
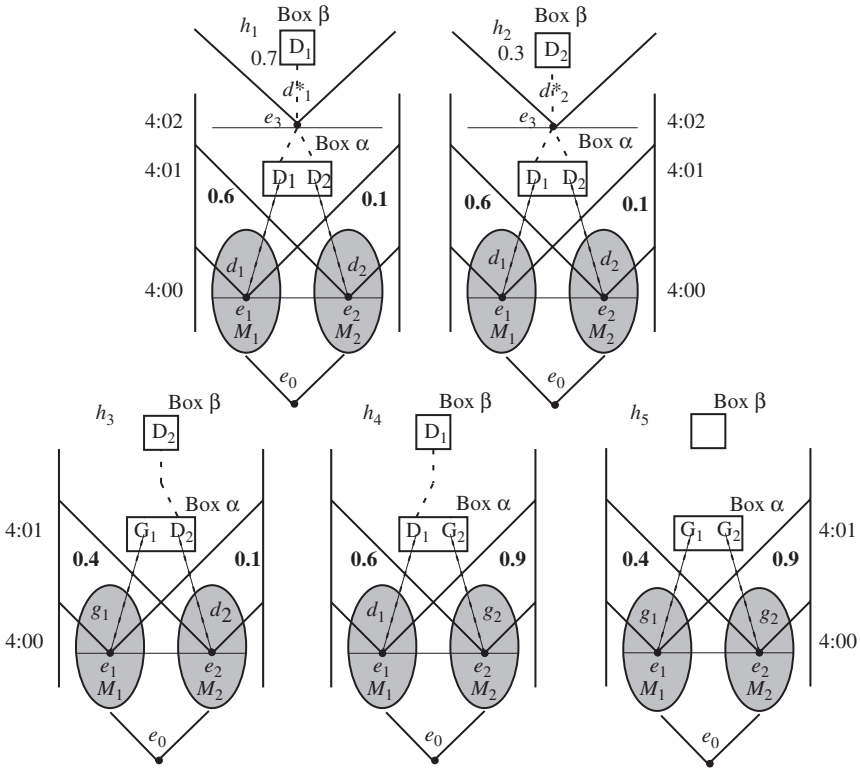
Fig. 6. The corkscrew story.

of e_1 or e_2 as initial:

$$pr(e_1 \rightarrowtail \pi_{e_1}\langle g_1 \rangle) = .4 = pr(e_1 \rightarrowtail \{h_3, h_5\}),$$
$$pr(e_1 \rightarrowtail \pi_{e_1}\langle d_1 \rangle) = .6 = pr(e_1 \rightarrowtail \{h_1, h_2, h_4\}),$$
$$pr(e_2 \rightarrowtail \pi_{e_2}\langle g_2 \rangle) = .9 = pr(e_2 \rightarrowtail \{h_4, h_5\}),$$
$$pr(e_2 \rightarrowtail \pi_{e_2}\langle d_2 \rangle) = .1 = pr(e_2 \rightarrowtail \{h_1, h_2, h_3\}). \quad (22)$$

It is part of Salmon's story that one day

> the Inspector picks one corkscrew from the day's output and finds it defective (Salmon, 1989).

Salmon's Humphreys-inspired Bayes-type question then arises: What shall we say about the corkscrew's propensity to have been made by M_1? His answer: "it would hardly be reasonable to speak of the *propensity* of that corkscrew to have been produced by" M_1. It certainly does sound like double talk. The Inspector examines a corkscrew, only to find it defective. Does anyone want to say that the corkscrew lying quietly in his hand has a certain propensity to—what? Propensity to *have been made* by machine $M1$ rather than $M2$? But that seems nonsense. Calculate a subjective probability if you wish; there is nothing wrong with "Given what I know, it is reasonable to believe that there is a good

chance that the corkscrew was made by machine $M1$," so that you can likely find a guide for your inferring, should you wish to so engage yourself. Or work out whatever objective correlations you like, survive the Problem of the Reference Group, and put your money down. Correlations may be objective, but they have no "direction," whereas propensities are all about influencing the future. The corkscrew was in fact made in the settled past either by machine $M1$ or machine $M2$, and current chancy propensity is non-existent.

We nevertheless need to go further into the "deep structure" of the story in two ways. First, Salmon uses language that attributes propensities to *objects*: "we can speak of the relative propensities of the two machines ... to produce defective corkscrews." (The same language is used in the can opener story of Salmon, 1984, p. 205.) But in attributing propensities to objects, we threaten to abandon the "single case." *Basic single-case propensities are more nearly tied to* transition events, *not to* objects, and not only because of wanting to adhere to the single case. In addition, the English passive can be used to ill effect: Salmon's active-voice "The propensity of each machine to turn out a defective corkscrew on a given day" could be turned into "the propensity for a defective corkscrew to be turned out by each machine," is, though equivalent by the standard passive transformation, less clear as to the bearer of the propensity. It is, however, perfectly rational to relate propensities for objects to transition propensities by way of complex events, for example by attributing a propensity to a whole series (or cloud or continuum) of transitions rooted in point events in a space-time-like event-locale. BSTP has the resources for this generalization (see Müller, 2005). Just do not make it without noticing what you are doing, or without careful conceptualization. If, when theorizing, we consistently attribute propensities to transition events, we shall be less likely to fall into the soup.

Second, the rhetoric suggests that all the indeterminism is fastened on the respective occasions e_1 and e_2 of production of the corkscrews, respectively, by M_1 and M_2. Not so: Salmon is explicit that the Inspector *picks* one corkscrew from the day's output. He is even more explicit in telling the can opener story: "Someone randomly picks a can opener out of the box" (Salmon, 1984, p. 205). That makes the Inspector's random pick of a corkscrew (or can opener) a *causa causans*, and one not to be ignored in drawing philosophical pictures or morals.

Fig. 6 features Box α and Box β as part of my retelling of the corkscrew story. Boxes, like corkscrews and inspectors, are substances that can be re-identified even when coincident with radically distinct point events in BSTP. (Were I to be distracted by problems surrounding this pronouncement, I would take my argle-bargle from Wiggins (2001), extending the arguments from linear time to branching space-time. In the meantime, I beg you not to suggest that Fig. 6 shows 10 boxes!) For the corkscrew story we need two boxes, Box α and Box β. The minute the corkscrews come out of the two machines, they are mechanically dumped into Box α. (All the dotted lines in Fig. 6 represent paths in a single history; motion along them is deterministic.) From the perspective of the point event, e_0, in the causal past of the whole of a day's running of the machines, Box α can be filled in one of four ways: (1) A good corkscrew from M_1 and a defective corkscrew from M_2, which we label, respectively, G_1 ("G" for "good") and D_2 ("D" for "defective"); (2) the reverse combination labeled D_1 and G_2; (3) two good corkscrews, labeled G_1 and G_2; or, finally, (4) two defective corkscrews labeled D_1 and D_2. We have up to this point emphasized only the four transitions (22) that stretch forward from the initials e_1 and e_2.

9.1. The Inspector

Then along comes the Inspector. He or she ignores any good output in Box α, paying attention only to defective corkscrews. In a case in which Box α has a single defective corkscrew, as happens in histories h_3 and h_4, the Inspector is offered only a Hobson's Choice, and that uniquely defective corkscrew is deterministically moved to Box β. In the case in which Box α contains no defectives, as in h_5, the Inspector deterministically leaves all as it is, making no transfer to Box β. Naturally in this case the Inspector cannot succeed in gathering up a defective corkscrew. There remains the case in which two defective corkscrews were put into Box α, as illustrated in histories h_1 and h_2. Since the Inspector is to pick, in this case the story tells us that the Inspector indeterministically chose which defective corkscrew to transfer to Box β. (Otherwise we might say that the inspection was fixed.)

You can tell from Salmon's calculations that the Inspector's choice was Laplacian, with an equal chance for each corkscrew in Box α. That is, however, a thesis of no bearing on the point of the story, so with good reason Salmon simply does not say what the chances are for this propensive picking of a defective corkscrew from Box α. Let us (you and I) say that in the case when there are two defective corkscrews in Box α, the Inspector has a propensity of .3 for M_1's defective output and (of course) a .7 propensity for picking the defective output of M_2. Perhaps we should not; perhaps we should just assume that the chances are 50/50, in order to stay closer to the presumed intent of Salmon's story. But let us go with 30/70 just to have calculations that seem easier to keep straight. The choice point for the Inspector's pick, that is, e_3, must come after all the defective corkscrews are already in Box α (or else you have a different story). The propensity distribution for all of the *causae causantes* based on each of the choice points e_1, e_2, and e_3, can now be described as follows:

$$pr(e_1 \rightarrowtail \pi_{e_1}\langle g_1 \rangle) = .4 = pr(e_1 \rightarrowtail \{h_3, h_5\}) = .04 + .36,$$
$$pr(e_1 \rightarrowtail \pi_{e_1}\langle d_1 \rangle) = .6 = pr(e_1 \rightarrowtail \{h_1, h_2, h_4\}) = .042 + .018 + .54,$$
$$pr(e_2 \rightarrowtail \pi_{e_2}\langle g_2 \rangle) = .9 = pr(e_2 \rightarrowtail \{h_4, h_5\}) = .54 + .36,$$
$$pr(e_2 \rightarrowtail \pi_{e_2}\langle d_2 \rangle) = .1 = pr(e_2 \rightarrowtail \{h_1, h_2, h_3\}) = .042 + .018 + .04, \tag{23}$$
$$pr(e_3 \rightarrowtail \pi_{e_3}\langle d_1^* \rangle) = .7 = pr(e_3 \rightarrowtail \{h_1\}) = .7,$$
$$pr(e_3 \rightarrowtail \pi_{e_3}\langle d_2^* \rangle) = .3 = pr(e_3 \rightarrowtail \{h_2\}) = .3.$$

From this table, because it includes all the relevant *causae causantes*, we can construct a table of the non-basic propensities, in particular the table of propensities of the various transitions $e_0 \rightarrowtail \pi_{e_0}\langle h_i \rangle$ for $1 \leqslant i \leqslant 5$:

$$pr(e_0 \rightarrowtail \pi_{e_0}\langle h_1 \rangle) = .042 = .6 \times .1 \times .7,$$
$$pr(e_0 \rightarrowtail \pi_{e_0}\langle h_2 \rangle) = .018 = .6 \times .1 \times .3,$$
$$pr(e_0 \rightarrowtail \pi_{e_0}\langle h_3 \rangle) = .04 = .4 \times .1,$$
$$pr(e_0 \rightarrowtail \pi_{e_0}\langle h_4 \rangle) = .54 = .6 \times .9,$$
$$pr(e_0 \rightarrowtail \pi_{e_0}\langle h_5 \rangle) = .36 = .4 \times .9. \tag{24}$$

You can also start with this table, (24), whose entries show propensities from a suitably early initial to a maximally precise outcome, and finish with the previous table, (23).

9.2. Inversion

Having the advantage of a thorough working-out of the story, let us go back now and say what needs to be said about the inversion in question. First, the attribution of a forward propensity with "the propensity of each machine to make a defective corkscrew." By Fig. 6 we are asking for the propensity of the basic transition from the production-point e_i of each machine M_i to the immediate outcome d_i that guarantees a defective corkscrew from that machine; namely, $pr(e_1 \rightarrowtail \pi_{e_1}\langle d_1 \rangle) = pr(e_1 \rightarrowtail \{h_1, h_2, h_4\})$, which by (23) is .6; and also the propensity of $pr(e_2 \rightarrowtail \pi_{e_2}\langle d_2 \rangle) = pr(e_2 \rightarrowtail \{h_1, h_2, h_3\})$, which, is .1. So ".6" and ".1" are the straightforward answers to "For each machine, what propensity does it have (at 4:00) to make a defective corkscrew?"

It should be clear that the probability of a transition is not the same as a standard conditional probability, whether defined in terms of absolute probabilities or not. In fact, I wonder if there is anything conditional here whatsoever. In an earlier publication I suggested that a transition such as $pr(e_1 \rightarrowtail \pi_{e_1}\langle d_1 \rangle)$ is an "event-conditional," a suggestion that arose from the observation that the "antecedent" must be an event rather than a proposition, since it must supply a locus in *Our World*. (Thinking of events as sets of point-events and of propositions as sets of histories, the former do have such a locus, whereas propositions do not. This point arises repeatedly in the BST papers.) It was nevertheless, I now think, not a profitable suggestion. What is wrong is the "conditional" part: The antecedent simply does not supply any kind of condition. I then suggested that, inasmuch as we are looking at a transition from e_1 to d_1, the antecedent event supplies a "perspective." One might try reading $pr(e_1 \rightarrowtail \pi_{e_1}\langle d_1 \rangle) = .4$ as "from the perspective of e_1, the propensity for d_1 coming to pass is .4." I think that may help some, but it rubs somewhat against the grain of BSTP: To suggest that the "action" lies with the outcome, d_1, with e_1 as only a passive onlooker, certainly seems to underplay the relational aspect of a transition. Attributing a propensity-strength to a transition $\mathbf{I} \rightarrowtail \mathbf{O}$ is, in this respect as well, a major difference from attributing a conditional probability, for the reading "Given B, the probability of $A = p$," which rhetorically fastens the probability onto A, is eminently satisfactory. The reading moves B into the background, where it belongs, so that the role of A can be emphasized. These are of course rhetorical points, but I do think that they deserve consideration.

It is perhaps helpful to pursue the analogy between a BSTP transition $e \rightarrowtail O$ that represents a propensity and a position vector as represented by an arrow. (1) The outcome, O, gives a "direction" to the propensity in analogy to the direction of the arrow, (2) the initial e provides a "starting position" in *Our World* in analogy to the position of the tail of the arrow, and (3) the strength p of the propensity is analogous to the length of the arrow. That analogy urges that propensities are essentially relational. It is of course all right, even necessary, in English to speak of the propensity as being "of" the initial and "for" the outcome, but for accuracy it all has to be said in one rush of air. In the following lists, note how the degree of the propensities change as one varies *either* the initial (what follows the "of" in "the propensity of . . .") *or* the outcome (what follows the "for," or some other preposition, in "propensity for . . .").

This batch varies the initial while keeping the outcome constant:

$$pr(e_1 \rightarrowtail \pi_{e_1}\langle d_1^* \rangle) = .042,$$
$$pr(e_2 \rightarrowtail \pi_{e_2}\langle d_1^* \rangle) = .07,$$
$$pr(e_3 \rightarrowtail \pi_{e_3}\langle d_1^* \rangle) = .7. \tag{25}$$

This batch keeps the initial constant while varying the outcome:

$pr(e_1 \rightarrowtail \pi_{e_1} \langle d_1 \rangle) = .6,$

$pr(e_1 \rightarrowtail \pi_{e_1} \langle H_{(e_3)} \rangle) = .06,$

$pr(e_1 \rightarrowtail \pi_{e_1} \langle d_1^* \rangle) = .7.$

What, however, about the inversions challenged by Humphreys, Salmon, and so on? There is more than one thing to say. First and foremost, through the lens of BSTP one sees that inversion of a BSTP transition makes no sense. Forgetting about probabilities, and going way back to pre-probabilistic BST, initials and outcomes could never be exchanged, and that for two reasons. First, and I suppose foremost, it is part of the BST concept of a transition that the initial *precedes* the outcome in the causal ordering provided by *Our World*. Keep in mind that this has nothing to do with states, events, initials or outcomes that occur at different times. We sorted through that in discussing Miller. Instead we are thinking about a single transition from a point event to one of its immediate outcomes. This preliminary characterization, in turn, generated the need for separate characterizations of initials and outcomes: Initials must be able to finish, and outcomes must be able to begin. These two requirements arose out of inescapable theoretical necessities, although certainly supported by intuitive considerations. So given that $I \rightarrowtail O$ is proper, one finds that $O \rightarrowtail I$ is generally not well-formed. That is so deeply ingrained in BST that it seems unimaginable that we should be able to make sense of inversion in that theory simply by adding probabilities.[12]

Second, even with all this, it remains tempting to try to make what sense one can of inversion as sometimes expressed in English by finding something intelligible in BSTP theory that is relevant to at least the example. Three things are required, one from Miller, one from our discussion of Miller, and one from Müller's 2005 work on BSTP: (1) If the English is parsed as Miller suggests, some inversions in English seem to make sense. (2) Our suggestion is that the way to make simple BSTP sense out of Miller's examples is to treat everything that one wishes to invert, both conditioning and conditioned (I mean both A and B in the conditional probability $P(A|B)$) as outcomes of some one initial event, or even better for commonsensical understanding, as *stable* events whose mode of occurrence is such that each is fully contained in a history h (that is, comes to completion in h) just in case it overlaps h (that is, begins in h). Observe, however, that one cannot do with stable events alone, unless one plans to drop the entire causal analysis featuring *causae causantes*, which is not a good idea. (3) Remind ourselves that by Müller (2005), there is *always* a boolean algebra and a field of probabilities that are *naturally* generated by a set of BSTP transitions.

Finally, what about probability? The essential idea is, not surprisingly, that *causal probabilities belong essentially to basic transitions*, and (derivatively provided there is no funny business, but holistically if there is funny business) to sets of basic transitions. Consult Placek (2004) for some ideas on how to think about these EPR-like phenomena. As Müller has established in the no-funny-business case (I put it roughly), a properly chosen or constructed set of these always forms a standard probability space $\langle A, F, \mu \rangle$, with A a sample space containing certain well-chosen sets of basic transitions, F a boolean algebra of all subsets of A, and μ a standard probability measure on F. Since we have probabilities of sets of basic transitions, we have probabilities of each non-basic transition as well by taking the probability of the set of its *causae causantes*. Also we have lots and

---

[12]This discussion assumes no funny business in the sense of FBBST and NCCFB. If there is funny business, everything must be re-thought.

lots of standard conditional probabilities, which will in fact have a causal meaning of sorts since they are based on probabilities of basic transitions; that is, since they are based on propensities.

For a single example of a standard conditional probability that parrots Salmon's query,

> From the perspective of e_0, what is the probability that the Inspector takes a corkscrew made by machine $M1$, given that he takes a defective corkscrew?

My calculation is disappointingly simple because of the extreme simplicity of my version of the corkscrew example, but since it is only principle that is at stake, I will persevere. The calculation is this:

> The Inspector takes a corkscrew produced by $M1$ in histories h_1 and h_4. The Inspector takes a defective corkscrew in histories $h_1, h_2, h_3,$ and h_4. (Just look at what is in box β. Note: "Take" is not intended to imply "picked" or "chose," so the Hobson's choices count.)

$$pr(e_0 \rightarrowtail \{h_1, h_2, h_3, h_4\}) = .64,$$
$$pr(e_0 \rightarrowtail \{h_1, h_4\}) = .582,$$
$$P_{e_0}(\{h_1, h_4\} | \{h_1, h_2, h_3, h_4\}) = \frac{.582}{.64},$$
$$P_{e_0}(\{h_1, h_2, h_3, h_4\} | \{h_1, h_4\}) = 1. \tag{26}$$

The subscript e_0, which here represents a perspective, is intended to fill out an echo of the three-termed relation of Miller and of Humphreys that we earlier canvassed. The last line is the conditional probability of the Inspector taking a defective corkscrew given that he takes a product of $M1$. The answer in this case is an uninteresting "1" because the set-up implies that the Inspector takes only defective corkscrews.

So that illustrates how, starting with causal probabilities and propensities attached only to *causae causantes*, one can have a rich feast of standard probabilities with standard conditional probabilities and all that jazz. Which of course is far from saying that propensities are standard probabilities, or that standard probabilities are propensities. The upshot is rather that BSTP lets one investigate both, together with their interconnections, without sacrificing either.

10. Closing remarks

So what should we conclude? Is BSTP causal probability theory right for objective propensities or not? It seems best to close with some short answers.

1. A partial answer to our covering question is that propensities have a causal aspect, which Kolmogorov probability theory does not see or need to see. *Objective probability, by itself, has nothing to say about causation.* The work of Humphreys, Miller, etc., show that much to a fare-thee-well, although they refrain from drawing the strong conclusion. Nor is this conclusion impugned by decades of work issuing in suggestions of how to get something objectively causal out of objective probabilities, for in every case of which I know, the "how to" involves (sometimes implicitly, sometimes explicitly) something doxastic or epistemic.

2. BSTP theory provides a useful background for sorting out some of the arguments concerning the relations between probability and propensity; BSTP helps us see that the new language shared by Humphreys and Miller, while intuitively all right, is in need of a full working out that will meet the standards of Frege. To that extent, you cannot fully trust any conclusions of the debate that are stated in those terms.

3. Are propensities correctly described by probability theory? Humphreys took one side, coming to a negative conclusion, which I share: You cannot get a propensity out of a probability. It was surprising to find out, as it were, that the essence of the Humphreys (1985) argument—at least according to the BSTP analysis—did not involve probabilities; the argument was entirely *causal*, in the sense that BST borrows from special relativity.

4. Miller took the other side, coming to the positive conclusion that the bug-a-boo of Bayesian inversion can be dismissed as a non-problem. BSTP agrees with Miller that time-reversal is not a problem: I relied on ideas from BSTP to argue that as long as one is dealing only with BSTP *outcomes*, time-reversal (or Bayesian reversal of $(A|B)$) is fine. So there is a place for Bayes reversal in propensity theory, as long as outcomes ride on both sides of the vertical. On the other hand, BSTP says that you cannot reverse initial and outcome of a transition. They must sit in the causal order. Exactly what this means and why it is true is part of the theory of *causae causantes*, Belnap (2005b).

5. Sometimes it is tempting to try to get causality or "the direction of time" out of probability theory in the sense of the Kolmogorov axioms, but this entire enterprise may be misguided. One cannot know unless the part of the theory involving causal and spatio-temporal concepts is laid out with a rigor equal to that of the probability part. Certainly the classic examples, and there are many such, do not live up to this ideal. My knowledge of this research area is not deep, so I do not plead for a grant of authority. I am thinking of pioneers such as Simon (1957), Lazarsfeld (1958), Suppes (1970), and more recently Glymour, Spirtes, & Scheines (2001), all of whom are strictly rigorous with respect to probability theory, but not comparably rigorous with respect to space-time and the like. It seems to me just another case of "you only get out what you put in."

6. BSTP does not encourage treating $pr(\mathbf{I} \rightarrowtail \mathbf{O})$ as a conditional probability that should ever be defined by

$$pr(\mathbf{I} \cap \mathbf{O}) \div pr(\mathbf{O}).$$

What stands in the way is not only that an initial $\mathbf{I}$ when intersected with one of its outcomes $\mathbf{O}$ typically yields Ø, the empty set. In addition, the idea of an "absolute propensity" such as might be nominated by $pr(\mathbf{I})$ or $pr(\mathbf{O})$ seems to be unintelligible. Indeed, BSTP *always* distinguishes conditional probabilities $P(A|B)$ from propensities $pr(\mathbf{I} \rightarrowtail \mathbf{O})$. $P(A|B)$ is dealing with proportions, $pr(\mathbf{I} \rightarrowtail \mathbf{O})$ with causal structure. It is hardly a new thought that inferring objective causal structure from statistics is no matter of definition.

7. How *should* we then interpret the propensity $pr(\mathbf{I} \rightarrowtail \mathbf{O})$? Let us go wholesale along with Müller (2005). The interpretation is definitely and explicitly and rigorously going to be a probability—though not *merely* a probability: It wears its causal structure on its face. You first find the set $cc(\mathbf{I} \rightarrowtail \mathbf{O})$ of basic transitions that stand as *causae causantes* of $\mathbf{I} \rightarrowtail \mathbf{O}$. An arbitrary (finite) set of basic transitions gives rise via *BST structure* to an eminently satisfying notion of "maximal consistent subset of $cc(\mathbf{I} \rightarrowtail \mathbf{O})$."

The terminology guides us to see how natural it is to form the boolean algebra, B, with those maximal consistent subsets, A, as *atoms*. From there one needs to reach out but a hand's breadth in order to identify a suitable representative of $\mathbf{I} \rightarrowtail \mathbf{O}$ as a particular set, S, in the algebra B. Backing up just a little, one can see how the atoms serve as "middle terms": a propensity for each basic transition yields, by multiplication, a "propensive probability" for each atom, and hence, by addition, a propensive probability, p, for the set of atoms, S. Finally, the special relationship between S and $\mathbf{I} \rightarrowtail \mathbf{O}$ makes sense out of transferring p from S to $\mathbf{I} \rightarrowtail \mathbf{O}$: $pr(\mathbf{I} \rightarrowtail \mathbf{O}) = p$. I think that this analysis and theorem of Müller (2005), though tucked away here in penultimate position, is the most powerful result that anyone has obtained in BSTP theory.

8. I add two remarks. (1) It seems to me that BSTP helps out propensities very much more than vice versa, chiefly because BSTP is an exact theory and propensity-theory is not. Indeed, there would be little in the BSTP portions of this essay that would need rewording (for sense) if "propensity" was dropped in favor of some jaw-breaker like "objective single-case causal transition probability as idealized in BSTP." (2) BSTP *theory* attaches propensities to basic transitions (and to transitions more generally), but it is difficult or impossible to stay faithful to that attachment in informal English. I have not tried very hard, since, although I am far from sure, I do not think that there is much benefit at this stage of getting straight on all the variations in the syntax of the noun "propensity." Maybe next year. At this stage, the theory does not do well at teaching one how to talk. It is, however, pretty good at saying how one should think about propensities, as I take it is illustrated by the rigorous treatment in Section 4 of Humphrey's Paradox, and in Section 9 of Salmon's corkscrew story.

References

Belnap, N. (1992). Branching space-time. *Synthese*, *92*, 385–434 See Belnap, N. (2003a).

Belnap, N. (1999). Concrete transitions. In: G. Meggle (Ed.), *Actions, norms, values: Discussions with Georg Henrik von Wright* (pp. 227–236). Walter de Gruyter, Berlin. A "postprint" (2002) may be obtained from ⟨http://philsci-archive.pitt.edu⟩.

Belnap, N. (2002a). Double time references: Speech-act reports as modalities in an indeterminist setting. In F. Wolter, H. Wansing, M. de Rijke, & M. Zakharyaschev (Eds.), *Advances in modal logic*, Vol. 3 (pp. 37–58). Singapore: World Scientific. A preprint of this essay may be obtained from ⟨http://www.pitt.edu/~belnap⟩.

Belnap, N. (2002b). EPR-like 'funny business' in the theory of branching space-times. In T. Placek, & J. Butterfield (Eds.), *Non-locality and modality* (pp. 293–315). Dordrecht: Kluwer Academic Publishers A preprint of this essay may be obtained from ⟨http://philsci-archive.pitt.edu⟩.

Belnap, N. (2003a). Branching space-time, postprint, January, 2003. This is a postprint of Belnap (1992) that includes a number of additional explanations and a little re-structuring. It may be obtained from ⟨http://philsci-archive.pitt.edu⟩.

Belnap, N. (2003b). No-common-cause EPR-like funny business in branching space-times. *Philosophical Studies*, *114*, 199–221 A non-quotable preprint may be obtained from ⟨http://philsci-archive.pitt.edu⟩.

Belnap, N. (2005a). Agents and agency in branching space-times. In D. Vanderveken (Ed.), *Logic, thought and action* (pp. 291–313). Dordrecht: Springer.

Belnap, N. (2005b). A theory of causation: *Causae causantes* (originating causes) as inus conditions in branching space-times. *British Journal for the Philosophy of Science*, *56*, 221–253 A preprint of this essay may be obtained from ⟨http://philsci-archive.pitt.edu⟩.

Belnap, N., Perloff, M., & Xu, M. (2001). *Facing the future: Agents and choices in our indeterminist world.* Oxford: Oxford University Press.

Gillies, D. (2000). Varieties of propensity. *British Journal for the Philosophy of Science*, *51*, 807–835.

Glymour, C., Spirtes, P., & Scheines, R. (2001). *Causation, prediction and search* (2nd revised ed.). Cambridge: MIT Press First edition, 1993, by C. Glymour, Springer.

Humphreys, P. (1985). Why propensities cannot be probabilities. *The Philosophical Review*, *94*, 557–570.

Humphreys, P. (2004). Some thoughts on Wesley Salmon's contributions to the philosophy of probability. *Philosophy of Science*, *71*, 942–949.

Lazarsfeld, P. F. (1958). Evidence and Inference in Social Research. *Daedalus*, *87*, 99–130.

Miller, D. (1994). *Critical rationalism. A restatement and defense*. Chicago and La Salle: Open Court Publishing Company.

Miller, D. (2002). Propensities may satisfy Bayes's theorem. In R. Swinburne (Ed.), *Bayes's theorem, Proceedings of the British Academy*, Vol. 113 (pp. 111–116). Oxford: Oxford University Press.

Müller, T. (2002). Branching space-time, modal logic and the counterfactual conditional. In T. Placek, & J. Butterfield (Eds.), *Non-locality and modality* (pp. 273–291). Dordrecht: Kluwer Academic Publishers.

Müller, T. (2005). Probability theory and causation: A branching space-times analysis. *British Journal for the Philosophy of Science*, *56*, 487–520.

Placek, T. (2000a). *Is nature deterministic?* Kraków: Jagiellonian University Press.

Placek, T. (2000b). Stochastic outcomes in branching space-time: Analysis of Bell's theorem. *British Journal for the Philosophy of Science*, *51*, 445–475.

Placek, T. (2004). *Quantum state holism*: A case for holistic causation. Jagiellonian University.

Popper, K. (1959). The propensity interpretation of probability. *The British Journal for the Philosophy of Science*, *10*, 25–42.

Prior, A. (1957). *Time and modality*. Oxford: Oxford University Press.

Reichenbach, H. (1956). *The direction of time*. Berkeley: University of California Press.

Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press.

Salmon, W. C. (1989). *Four decades of scientific explanation*. Minneapolis: University of Minnesota Press.

Simon, H. A. (1957). *Models of man: Social and rational*. New York: Wiley.

Suppes, P. (1970). *A probabilistic theory of causality, Acta philosophica fennica*, Vol. 24. Amsterdam: North-Holland.

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, & G. Guenthner (Eds.), *Handbook of philosophical logic, Vol II: Extensions of classical logic, Synthese Library, Studies in epistemology*, Vol. 165 (pp. 135–165). Dordrecht: Reidel.

Thomson, J. J. (1977). *Acts and other events*. Ithaca: Cornell University Press.

von Wright, G. H. (1963). *Norm and action. A logical inquiry*. London: Routledge & Kegan Paul.

Weiner, M., & Belnap, N. (2006). How causal probabilities might fit into our objectively indeterministic world. *Synthese*, *149*, 1–36.

Wiggins, D. (2001). *Sameness and substance renewed*. Cambridge: Cambridge University Press.