Propensities and probabilities (revised)*

Nuel Belnap

October 31, 2010

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]

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

*This is a revision of a paper that appeared several years ago. I am indebted to Tomasz Placek for telling me about a number of errors in that paper, and for Thomas Müller for finding errors and infelicities in earlier versions of this one.

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 shan't 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'll be looking at some of the discussion as it has emerged.

My plan is this. In §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 §3 I describe the little bit of probability theory that the discussion requires, and with this as background, in §4 I lay out and discuss what has come to be known as "Humphreys's Paradox." After some preliminary remarks in §4.1, in §4.2, I rehearse Humphreys's argument, based on his photon example, that the "paradox" causes trouble for the identification of propensity and probability. In the course of presenting his *objection* to propensity theory, Humphreys introduces some new notation for putative propensities, as described in §4.3, using it to sharpen his argument. I critically discuss Humphreys's use of the new notation in §5, on the grounds that he has not given it a clear meaning. In §6 I discuss the contrasting use by Miller 1994 of the new notation to *defend* the identification of propensities as probabilities. Although I criticize some of the Miller details, I provide a reworking of the meaning of the new Humphreys-Miller notation

[1]BSTP is explored in the following papers listed as references: Belnap 1992 (BST), 2003a (BSTpp), , Belnap et al. 2001 (FF), Belnap 2002a (DTR), Belnap 2002b (FBBST), 2003 (NCCFB), 2005 (CC), Müller 2002, 2005, Müller et al. 2008 (IMC), Placek 2000a, 2000b.

that justifies a conclusion that is much the same as Miller's. (The reworking gives a formally correct meaning to the new notation that is intended to be half way between the meaning originally awarded by Miller and the meaning that finally emerges in the context of BSTP.) In order further to motivate BSTP, I return to Humphreys in §7, depending on a tinker with one of his principle examples to make my deepest criticism, by arguing that his causal analysis is faulty. 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 §8 and §9, and then in §10 I apply BSTP to yet another propensity story (better: a putative 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 §11.

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 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 until §8 we have to live with looseness.

1. Popper 1959, p. 32, explains a propensity as "the probability of the result of a single experiment, with respect to its conditions" (p. 68). 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 clear path from here to fastening a probability-of-occurrence onto the individual experiment

itself, nor does a conjectural reference to conditional probabilities seem to help in the single case.

Reason: 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 don't 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 came 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 §8.2—not other-worldly as in Lewis's account.)

2. Popper (*ibid.*, 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, but see §11(11). 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 (1959, 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, 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 objective (*ibid.*, 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 surprising point (1957) 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 along 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 had the possibility of carrying out one of several actions, you may have had 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 are 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 nonrepeatable 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's just the same-old same-old of scientific methodology.

11. *There is always some causal claim involved in an ascription of propen-*

sity. 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 outlined in CC, 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 remind ourselves of 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 nonempty and finite "sample space," F is the Boolean algebra of all subsets of Z, P is a function from F into the interval $[0, 1]$, $P(\varnothing) = 0$, $P(Z) = 1$, and for all $A, B \in F$, if $A \cap B = \varnothing$ then $P(A \cup B) = P(A) + P(B)$. [3]

The definition doesn't 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 literally 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 1986, 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 §8.3.)

An event in BSTP theory, whether possible or actual, is as concrete as a Los Angeles freeway. I also say 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":

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

where $A \cap B$ is set intersection and B is presumed nonempty. (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:

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

> $P(A|B)$ is to be read in English as "the probability that A *will occur* given that B *has occurred*." [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. [6]

For practical application, it often doesn't matter how you read conditional probability in sloppy old English; that's the nature of atheoretical practicality. The tensed English reading [5] just won't 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's 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 As among the Bs 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 calculus exhibiting such an inversion is this:

$$P(A|B) \times \frac{P(B)}{P(A)} = P(B|A). \qquad [7]$$

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 "$P(A|B)$."[3] In that case, we have Bayes's own theorem, which involves an inversion without invoking absolute probabilities:

$$P(B|(A \cap C)) = \frac{P(A|(B \cap C)) \times P(B|C)}{(P(A|(B \cap C)) \times P(B|C)) + (P(A|(\overline{B} \cap C)) \times P(\overline{B} \cap C))} \qquad [8]$$

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 (§2 and §3), we must deal with Humphreys 1985, which initiated discussions of the relations between the two.

4.1 Preliminary remarks

Humphreys 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.[4] My own view, expressed for example in FF, is indeed incompatibilist, but the argument of FF is principally to the effect that agency presupposes indeterminism—but not "chances" in the everyday sense that connotes mindlessness, or at least lack of deliberation. FF identifies a minimal sense of indeterminism that does not involve contrast with such heavy ideas as "laws" or "scientific theories."

[3] See, however, the discussion just after [10].

[4] I am reluctantly using "compatibilism/incompatibilism" for the opposed theses that Free Action is/is not compatible with Determinism. If I had my way, the opposition would be between the claims that Free Action is/is not compatible with Indeterminism; but I bow to common usage.

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.[5]

Avoiding the word, "chances," I describe Humphreys as simply arguing 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 as a provocatively crisp thesis in the very title of Humphreys 1985:

> *Propensities cannot be probabilities.* [9]

Humphreys always has in mind that propensities are conditional, as in "there is a propensity p for an electron in the metal to be emitted, conditional upon the metal being exposed to light above the threshold frequency" (p. 558), and so symbolizable by the form $pr(A|B)$. (I uniformly write this in preference to Humphrey's $Pr(A|B)$ in order to heighten the contrast with the standard probability-theory notation for conditional probabilities, $P(A|B)$.) 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 §3, standard probability theory generally gives $P(A|B)$ the English reading, [6] "the probability of A given B." $P(A|B)$ is virtually always taken as interdefinable with absolute probabilities, either via [4] when absolute probabilities are primitive, or via $P(A)=_{df}P(A|Tautology)$ if conditional probabilities are primitive. Either way, absolute probabilities are taken as respectable; BSTP, however, rejects this doctrine.

Humphreys's principal lemma for [9] is that *Bayes theorem*, [8], *fails for propensities*. I shouldn't 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). Given that there is no call to

[5] Many determinists and embarrassed indeterminists think that laws or theories or epistemic notions are required in order to make indeterminism intelligible, a thought that I do not share. See chapters 6, 7, and 8 of FF.

question the standard much-worked-over foundations of applied probability theory, it is clear that Humphreys's intent is to attack the idea that the inversion principles of probability theory apply to conditional propensities. 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.[6] 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 Humphreys's main thesis, [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 §8 and §9. It behooves us first to look at the current state of the topic.

[6] 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."

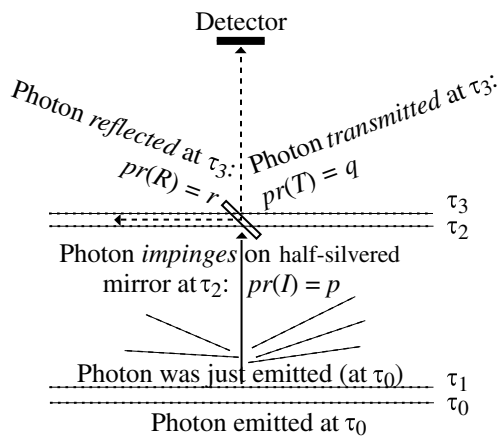

Figure 1: A nanosecond in the life of a photon. I = Impinged, R = Reflected, T = Transmitted.

4.2 Humphreys's photon argument

Many of the properties attributed to propensities, as listed informally above in §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 contributions was to argue that, important as is the suggestion, maintaining this identity thesis is difficult or impossible. I'll go over some of the argument, because, as indicated, I plan eventually to argue that it does make 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.

The story concerns a photon. (I take certain liberties with the story, in the belief that they are irrelevant to its point.) Figure 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 $r = (1-q)$ to wind up Reflected ($= R$) off to the left. Figure 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]

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).[7]

The first four items on the list below, namely, (i)–(iii) and (CI), are intended by Humphreys as local assumptions, governing just the emit-impinge-transmit set-up that is pictured spatially in Figure 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

[7] I have simplified a little in ways that I think don't 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 gone astray in replacing Humphreys's use of concatenation as in "$C_{\tau_k}B_{\tau_i}$" with an explicit sign "$\cap$" 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.

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].

4-1 Assumptions. *(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.*],

Humphreys derives an inconsistency from these assumptions by straightforward argument in probability theory (p. 562). I am going to pursue, however, some more basic themes: In §5 and §6, I express doubt that either Humphreys or Miller has given the notation, [13], a clear meaning, and in §7 I challenge the reasoning leading to Humphreys's Assumption **4-1**(iii).

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 relative to a stated domain), $A \cup B$ (union), etc. The theory evidently also invites conditional probability expressions "$P(A|C)$," since such an 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's not that [13] couldn't be given a meaning sharp enough to support a proof, but only that it hasn't 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 Assumption **4-1(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 rather 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 inklings of a language proper to the "branching time" representation of indeterminism were (1) a passing suggestion in an unpublished letter from S. Kripke to A. N. Prior dated September 3, 1958, and (2) a short and casual discussion in Prior (1967). Only in Thomason 1970 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: As I understand it, the idea of a universal time for our world is admissible only relative to an inertial frame. 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 Figure 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 Figure 1, it is dead right that Bayes-type inversion is no challenge; for example, in Figure 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 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 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 Figure 1, each of "reflect" and "transmit" can occur at τ_3 (but of course their joint occurrence at τ_3 is

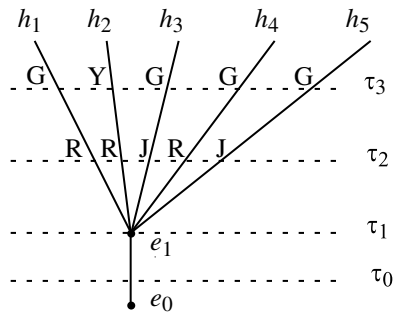

Figure 2: Bayes inversion acceptable.

not possible). Here is a homely example. It's 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's 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_{τ_3} and R_{τ_2} to be characterized only semi-generically, by a set of alternative possible events or histories, but we need e_1—the *initial event* itself—in all its 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's check this with a wholly unrealistic but simple assignment of propensities. With Laplace in mind, and only for ease of calculation, we'll assign e_1 equal propensities towards each transition into a "history," that is, a possible course of events[8] h_i $(1 \leqslant i \leqslant 5)$ to which e_1 belongs, as pictured in Figure 2. (Events belong to histories, but times do not belong to histories. Also, to anticipate with an idea and notation from BSTP introduced in §9, in BSTP propensities in the first instance attach neither to initial events nor to their outcomes, but rather to "basic transitions" from point events to their immediate outcomes, and derivatively to non-basic transitions from perhaps complex initial events to complex and more distant outcomes.) Now ask first,

> With reference to Figure 2, what is the proper value of p in the time-switched $pr_{e_1}(R_{\tau_2}|G_{\tau_3 2}) = p$, and of q in $pr_{e_1}(G_{\tau_3}|R_{\tau_2}) = q$? [18]

As long as one has equipropensity with respect to histories, the recipe for calculating the answer to [18], is (misleadingly) easy.

1. Let g be the number of Gs, in Figure 2, expressing alternatives in the future of e_1 for time τ_3, that is, let $g = 4$.

[8] From this point, instead of "course of events" we begin to use the BSTP word "history," defined in §8.3. There are two caveats tied to the overwhelming lack of realism exhibited in Figure 2: the idea of a single history being the outcome of an indeterministic transition is absurd, and the idea of equipropensity, although not absurd, is certainly so far from paradigmatic as to be downright misleading.

2. Let $r = 3$ since three histories to which e_1 belongs contain a possible occurrence of R 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, for [18], 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."[9] (File away the following observation for later consideration: Figure 2 is a "modal" diagram, without any representation of space, whereas Figure 1 is largely a "spatial" diagram, with modality represented by spatially diverging paths, whose meaning is conveyed either by extrinsic notes or by annotations.)

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 propensities to give rise to outcomes such as, in the misleading easiest case, one obtains by considering proportions. There is no call to invert the propensity claim by switching outcome and initial.

The example of Figure 2 suggests that the language of "proportions" can be helpful in the easiest cases as long as one keeps in mind the ways in which it can be misleading. So I recommend in those toy cases the following as a long-winded, whenever-you-have-the-time-and-the-energy type of reading. (The reading [19] 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 $\tau_i \neq \tau_j$.)

[9] 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* (certain privileged *sets* of histories) of which a given initial may have finitely many. See especially Belnap 2005b.

> $pr_{e_1}(A_{\tau_j}|B_{\tau_i})$ should then be read as *the proportion of cases (courses of events) through* e_1 *in which* A_{τ_j} *and* B_{τ_i} *both occur among all the cases (courses of events) in which* B_{τ_i} *occurs.* [19]

[19] is a faithful reading of $pr_{e_1}(A_{\tau_j}|B_{\tau_i})$ since that proportion, *when available*, accurately represents the propensity for e_1 to give rise to A_{τ_j} given that it gives rise to B_{τ_i}.[10] 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.[11] McCurdy 1996, when wrestling with Miller's treatment of Humphreys's Paradox comes to much the same conclusion: "I find [Miller's] treatment quite successful" (p. 106).

7 Humphreys: two mirrors

It behooves us carefully to return to §4.2 in order to consider (iii) of Assumptions **4-1** on p. 15 above: 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 Figure 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's a clear-cut feature of the one-mirror example, which is visually reenforced in Figure 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 Figure 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

[10] See note 9 above, and §10.3 below.

[11] Red flag: In this section I have avoided employing the full technical apparatus of BSTP; this has inevitably led to less than rigorous exposition. See §8, §9, and §10 for improved accuracy. Another warning: This discussion is intended to say nothing at all about *epistemic* probabilities.

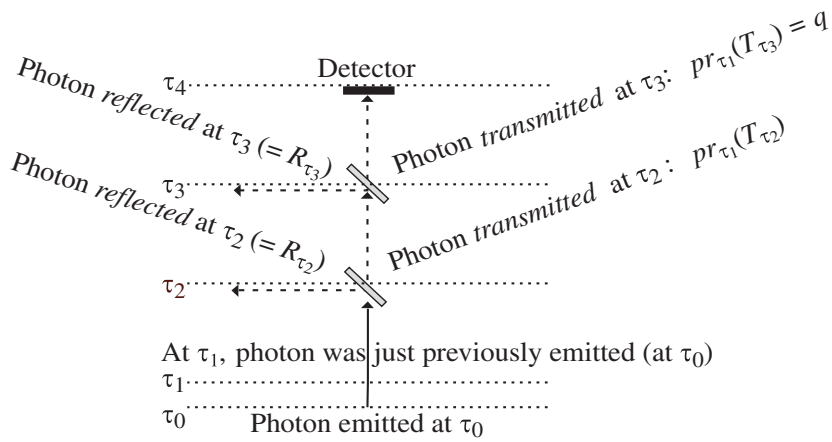

Figure 3: Two half-silvered mirrors yield two occasions of indeterministic transmission-or-reflection.

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 Figure 1 by adding a second half-silvered mirror. I trade in "impingement vs. non-impingement" at τ_2 (as in Figure 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 Figure 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."[12] I'll 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 Figure 1. As a side description of the two-mirror set-up (not drawn in the diagram), we indicate the modal fact that exactly one of transmission

[12]In the spatio-temporal Figure 3, I am reverting to taking *times* instead of *events* as the loci of indeterminism. The only reason for this change is that Figure 3 is already overcrowded.

or reflection can occur at τ_2, and, assuming transmission at τ_2, 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 what I take the difficulty to be, as seen in the two-mirror example of Figure 3. The notation allows four possible combinations:

$$T_{\tau_2}T_{\tau_3},\ T_{\tau_2}R_{\tau_3},\ R_{\tau_2}T_{\tau_3},\ R_{\tau_2}R_{\tau_3}. \qquad [20]$$

Humphreys's account that goes along with the notation, however, is not so generous in what it allows. His Assumptions **4-1**(iii) lays it down that $pr_{\tau_1}(T_{\tau_3}|\overline{I}_{\tau_2}) = 0$, that is, that the propensity at τ_1 for T to occur at τ_3 conditional upon $\overline{I}$ occurring at τ_2 is zero. Since Figure 3 sharpens the form of the example by using R_{τ_2} (reflection at τ_2) in place of $\overline{I}_{\tau_2}$ (doesn't impinge at τ_2), the sharpening of Assumptions **4-1**(iii) amounts to "the propensity for transmission at τ_3 conditional upon reflection at τ_2 is zero." By parity of reasoning, we should also have "the propensity for reflection at τ_3 conditional on reflection at τ_2 is zero." That is, 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? But then the last two of the four combinations listed in [20] are impossible. Therefore, we must have one of the first two combinations: either $T_{\tau_2}T_{\tau_3}$ or $T_{\tau_2}R_{\tau_3}$. Here is the weird result: In either case, we would have T_{τ_2}. This would prove that from the perspective of the photon at time τ_1, the transmission of the photon at τ_2 is *guaranteed*. Evidently something has gone wrong.[13]

The conclusion appears to agree with Humphreys: Standard probability theory won't do for propensities. It is notable, however, that *probabilities didn't enter my version of the argument at all.* All the trouble lay in the prior causal analysis. The two diagrams of Figure 4 are intended to show you the difficulty and how to repair the damage. To appreciate the diagrams,

[13]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.

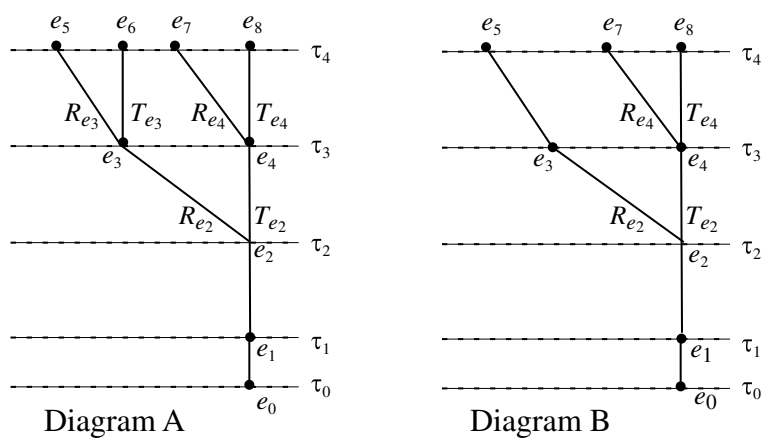

Figure 4: Two causal (or temporal-modal) diagrams for two half-silvered mirrors.

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 Figure 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 Figure 4 have no diagrammatic features indicating spatial relations, although, again, one can jam in spatial information by annotation. Diagram A of Figure 4 lays out a causal (or modal) representation of the same phenomenon that is spatially pictured in Figure 3. In both kinds of diagram, up-down keeps track of time; in Figure 3 left-right indicates a spatial relation, whereas Figure 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 e_3, which has a reflection in its immediate past, vs. e_4, which which immediately follows a transmission).

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 possible singular 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. Still with reference to Diagram A, after 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. (Permit me to repeat 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 half 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 Figure 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 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 Figure 4B, imaginatively suppose yourself to be now located at event e_5 at time τ_4. 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, whereas if situated at e_4, whose time, τ_3, is exactly the same as that of e_3, your future of possibilities would allow each of e_7 and e_8—but not e_3—as alternative possibilities. 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?" [21]

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.[14]

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 Ts and Rs, is, I think, nonsense. Humphreys's

[14] I am courting confusion by using "event" both for *point events*, symbolized by e_i, and *transition events*, symbolized by R_i and T_i. Only a desire for brevity excuses such usage. The matter is sorted out in Belnap 2005b.

reductio shows this. 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 were just, then all the weight of propensity (as seen from e_1) would have to go on the right side of Diagram A; which is to say, $pr_{e_1}(T_{e_2}) = 1$. Absurd.

The trouble is not with Diagram A itself. The trouble lies in the talk about the diagram. Equivalently it lies in the use to which Diagram A is put, namely, Diagram A is used to report events that cannot occur. 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 time 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 propensities. Diagram B is right, and Diagram A is wrong as a representation of the causal-modal propensity situation envisioned in the spatial Figure 3.

8 Branching space-times: a foundation for BSTP

BST should provide a good platform for propensities as probabilities. Here are two quick reasons: (1) BST offers an attractive theory of objective *possibilities*, something that that ought to be (but seldom is) the foundation of

any theory of objective probabilities. (2) BST supplies a cogent theory of objective event-event *causation*, a necessary condition of offering an intelligible account of propensities in terms of probabilistic causation, as called for by, for instance, Salmon 1984. In this section I review the basics of BST, and in the next I indicate how BSTP arises by the addition of probabilities to BST.

BSTP began, in Belnap 1992, as just BST, a theory of objective *possibilities* in our relativistic world. BST theory 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, in Belnap 2002b and Belnap 2003b, something simple and intelligible about correlations in the EPR family of apparently weird goings-on in quantum mechanics. (See additionally Müller et al. 2008.) From this platform, under the doubtless false but certainly entertaining proviso that there are no EPR-like phenomena, BST, still with the same primitives, offered in Belnap 2005b an explicit theory of originating causes (called *causae causantes*), taking transitions from initial-event to outcome-event as both caused and causing. The most recent work, largely represented by Weiner and 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*. Such a structure offers a viable candidate to serve as a 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 don't *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 isn't even jargon, it's a code word. Events for the objective probabalist 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."

8.1 Grammar of propensities in BSTP

The BSTP 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. Skipping, for the present, many intermediate techicalities, we may say that a *BSTP propensity*, which is measured like an objective probability, attaches to certain ordered pairs of concrete events, locatable in *Our World*. Which "kinds" of event-pairs can accept propensities will be important to BSTP propensity theory. The pair is a "transition," the left entry is an "initial" event, $\mathbf{I}$, and the right entry is an "outcome" event, $\overline{\mathbf{O}}$, each in a sense to be soon defined. (I use $\overline{\mathbf{O}}$ as a variable ranging over three kinds of outcomes, the kinds to specified later.) I write the ordered pair as $\mathbf{I} \rightarrowtail \overline{\mathbf{O}}$. Assume now (and also throughout this section) that $0 \leqslant p \leqslant 1$. Then, to make a propensity statement, I write

$$pr(\mathbf{I} \rightarrowtail \overline{\mathbf{O}}) = p \qquad [22]$$

How to read [22] in English? Such a propensity statement is saying something essentially relational. English, not being fond of relations, in using the idiom of propensity tends to call for the initial event to "have" the propensity, and to call for the propensity be "for" the outcome event, or "to [some active verb]." The following fits this pattern:

> (Possible) initial event $\mathbf{I}$ *has* a propensity of p *for* giving rise to (possible) outcome event $\overline{\mathbf{O}}$; or "the propensity *of* $\mathbf{I}$ *to* give rise to $\overline{\mathbf{O}}$ is p." [23]

The idiomatic syntactic form [23] awards the propensity to $\mathbf{I}$. Other ways of saying the same thing might award the propensity to $\overline{\mathbf{O}}$, as in

> From the standpoint or perspective of event $\mathbf{I}$, outcome event $\overline{\mathbf{O}}$ *has* a propensity of p to occur. [24]

Or one might give a personalized version of [24]: If you were ever to get to event $\mathbf{I}$, from that standpoint you could truly assert that the measure of the propensity for $\overline{\mathbf{O}}$ to occur is p.

I favor, however, the following reading, which by refusing to describe either $\mathbf{I}$ or $\overline{\mathbf{O}}$ as *having* the propensity, emphasizes the relational nature of propensities in BST theory:

> There is a propensity of p for initial event $\mathbf{I}$ to give rise to outcome event $\overline{\mathbf{O}}$. [25]

These English readings of [22], while important, speak only to the grammar of the language of propensities. The remainder of this section is devoted to explaining some concepts of branching space-times theory needed in order to be more specific. Our entire approach is by way of "event" concepts; the chief contrast is with enduring substances, but also with repeatable "experimental arrangements" as in §2(1) on p. 3. 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 benighted 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.[15] 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 study. In BSTP theory these concepts are both objective and perspectival—when tensed.

[15]More or less. Fixing e as the origin of the perspective, some events *are* future possibilities, whereas events that are space-like related to e *were* future possibilities.

8.2 Possible point events e, the causal order $<$, and *Our World* in BST

"Event" is at worst hopelessly ambiguous and at best vague. That won't 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) representing it as a nonempty set, and then (3) defining "e is a possibly-occurring point event" as equivalent to $e \in$ *Our World*.[16] (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 suggestive 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.

[16] "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? Wouldn't I rather point out that *Our World* is intended as a *representation* of our world, to be judged as more or less adequate for its purposes.

8.3 Histories, consistency, categorizing events in BST

Additional concepts and axioms require the definable idea of "history," which in turn depends on "directedness." A subset of *Our World* is *directed* iff whenever two (possible) point events e_1 and e_2 belong to it, 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. The idea of the definition is that if each of two possible events actually happened from the point of view of some event, then they must be compatible. A *history*, h, is a maximal directed subset of *Our World*, and a set of point events is defined as *consistent* iff it is a subset of some history. A history, h, is a complete course of possible point events, although perhaps "course" is the wrong word.[17] The chief warning is that you should not picture h as linear; a history permits its members to be separated either in time or space or both. The idea of history is, indeed, essential to the definition of "space-like related": $e_1 \,\text{SLR}\, e_2$ 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).

There is in BST a maximally generic sense of *event*, namely, a nonempty 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.

> An *initial event* (or just initial) **I** is a nonempty set of point events all of which are contained in some one history (**I** is consistent in the sense proper to a set of point events considered as an initial event: There must be a history in which the initial finishes, which is to say, a history in which all parts of the initial occur). We may treat a point event e as special case of an initial event. [26]

[17] Placek and Belnap 2010 establishes that there are BST models in which every history is isomorphic to special-relativistic Minkowski space-time.

> *Outcome events* $\overline{\mathbf{O}}$ fall into three classes: outcome chains O, scattered outcome events $\mathbf{O}$, and disjunctive outome events $\boldsymbol{O}$.
> An *outcome chain*, O, is a nonempty, linearly ordered, and lower bounded chain of point events. It is easy to see the following critical point: A chain event can begin (come to be) iff it is an outcome chain.
> A *scattered outcome event*, $\mathbf{O}$, is a nonempty set of outcome chains, O, all of which begin in (that is, overlap) some one history [27]
> ($\mathbf{O}$ is consistent in the sense proper to a set of outcome chains taken to represent a scattered outcome event: There must be a history in which all the chains begin).
> A *disjunctive outcome event*, $\boldsymbol{O}$, is a set of pairwise inconsistent scattered outcome events. ("Inconsistency" of a pair here means that there is no history in which both members of the pair begin.)
> I let $\overline{\mathbf{O}}$ range over all three sorts of outcome events.

In the spirit of Russell's "at-at" theory of motion, and adapting from the definition of "transition" of von Wright 1963, a *transition event* is simply an ordered pair of an initial event, $\mathbf{I}$, and an outcome event, $\overline{\mathbf{O}}$, subject to the requirement that, in the most plausible sense, $\overline{\mathbf{O}}$ begins after $\mathbf{I}$ finishes. See D5–D9 in Belnap 2005b for the unavoidably convoluted and superficially unintuitive spelling out of "most plausible sense." Briefly:

$$e < O \leftrightarrow \forall e' \in O(e < e'); e < \mathbf{O} \leftrightarrow \exists O \in \mathbf{O}(e < O); e < \boldsymbol{O} \leftrightarrow \forall \mathbf{O} \in \boldsymbol{O}(e < \mathbf{O}); \mathbf{I} < \overline{\mathbf{O}} \leftrightarrow \forall e \in \mathbf{I}(e < \overline{\mathbf{O}}) \quad [28]$$

The existential form of the definition of $e < \mathbf{O}$ may appear "implausible" until one calculates that the definition underwrites the critical fact that $e < \mathbf{O}$ implies that the occurrence of e is a necessary condition of the occurrence of the scattered outcome, $\mathbf{O}$. Von Wright takes transitions (defined differently) as events, and so in endorsement I shall speak of *transition events*. I write $\mathbf{I} < \overline{\mathbf{O}}$ for the causal-order requirement, and use the notation $\mathbf{I} \rightarrowtail \overline{\mathbf{O}}$ for the transition event.

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 has no "simple location" (Whitehead): As an entity, a choice is precisely the transition from the initial to the settled outcome; any other view leads, I believe, to severe muddle, if not to outright Zenoesque contradiction.[18]

A full and accurate study (as for example the account of "funny business" in FBBST) would have to be more subtle, 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 concepts. Of absolute import, however, is the given idea of a transition event. 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. Observe that since a transition event is an ordered pair of initial and outcome, there is no simple answer to "Where (in *Our World*) is $\mathbf{I} \rightarrowtail \overline{\mathbf{O}}$?".

Initials and outcomes also give rise to "occurrence" propositions represented as a particular set of histories, H, such that a history, h, is a member of H just in case the event can reasonably said to "occur" in h. Some metaphysicians take the expression, "event E occurs," or "event E occurs in world w," as having an intuitive meaning. That won't do, neither as a foundation for probabilities, nor for a rigorous theory of causation, in *Our World*.

> **Occurrence propositions.** Given an initial, $\mathbf{I}$, $H_{[\mathbf{I}]} = \{h : \mathbf{I} \subseteq h\}$ is the set of histories containing all of $\mathbf{I}$, so that the initial, $\mathbf{I}$, *occurs* in h just in case $\mathbf{I}$ finishes in h, in the minimal sense of being wholly contained in h. Also $H_{(e)}$ is the set of histories containing e. Given an outcome chain event, O, $H_{\langle O \rangle} = \{h : h \cap O \neq \varnothing\}$ is the set of histories containing some of O; so that the outcome, O, *occurs* in h just in case O has a beginning in h. [29]
> For $\mathbf{O}$ a scattered outcome event, $H_{\langle \mathbf{O} \rangle} = \bigcap \{H_{\langle O \rangle} : O \in \mathbf{O}\}$, so that $\mathbf{O}$ *occurs* in h iff all parts of $\mathbf{O}$ occur in h. Where $\boldsymbol{O}$ is a disjunctive outcome event, $H_{\langle \boldsymbol{O} \rangle} = \bigcup \{H_{\langle \mathbf{O} \rangle} : \mathbf{O} \in \boldsymbol{O}\}$, so that $\boldsymbol{O}$ *occurs* in h just in case some scattered outcome event that is disjunctive part of $\boldsymbol{O}$ occurs in h. Thus, I have given a meaning to all cases of $H_{\langle \overline{\mathbf{O}} \rangle}$.

[18] Moving from branching space-time to the simpler branching time, one may say that at each moment, either the choice has not yet been made or it has already been made. There can be no "moment of choice"—unless you wish, with some artificiality, to identify it as the last moment at which the choice has not yet been made.

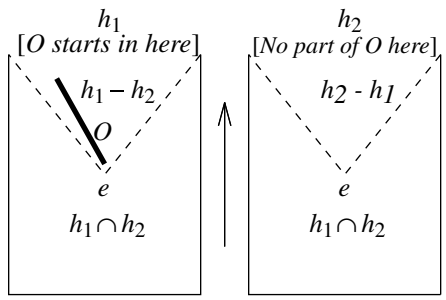

Figure 5: Prior choice principle: How histories divide.

Two notes regarding transitions. (1) For technical smoothness, we allow initials without upper bounds in h, even though such an initial cannot be the initial of any transition. (2) It is often useful to consider as a transition event a pairing $\mathbf{I} \rightarrowtail H_{\langle \overline{\mathbf{O}} \rangle}$, provided $\mathbf{I} < \overline{\mathbf{O}}$, even though the proposition, $H_{\langle \overline{\mathbf{O}} \rangle}$, gives less information than the event, $\overline{\mathbf{O}}$.

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 Figure 5. 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; I write $h_1 \perp_e h_2$ for either one of "e is a choice point for h_1 and h_2," or "h_1 and h_2 split at e." Figure 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 Figure 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 histories. Indeed, there seems to be no purely graphic method for representing in the same diagram both space-like information and splitting information. Yes, Figure 5 represents both kinds

of information at once, but does not confine itself to graphics. It is critical to Figure 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 at e (that is, $h_1 \perp_e h_2$).[19] Splitting from h is naturally generalized to sets of histories H ($h \perp_e H$) and therefore to outcomes $\overline{\mathbf{O}}$: $h \perp_e \overline{\mathbf{O}}$, defined by $h \perp_e H_{\langle\overline{\mathbf{O}}\rangle}$, signifies that up to and including e, each of h and $\overline{\mathbf{O}}$ are possible, but that after e either h or $\overline{\mathbf{O}}$ is no longer possible.[20]

Since you cannot overlook the overlap between the histories h_1 and h_2 that constitute the model of *Our World* illustrated in Figure 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 don't overlap is a central 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, §7B.2.

8.4 Concrete transitions, prior causal loci, *causae causantes*, and inus conditions in BST

The goal of this section is to explain the notion of *causa causans*, or *originating cause*, which is key to the BST account of causation, and which lies at the base of the present account of propensities. I begin by observing that every chain in every scattered outcome has, by previous definition, a lower bound, and therefore, by postulation, a greatest such. Consider a scattered outcome $\mathbf{O}$ such that all its contained chains share the same (proper) greatest lower bound, say e, in which case one writes $\mathit{inf}(\mathbf{O}) = e$. In this case there is no room between e and $\mathbf{O}$ that would permit causal lines from the past to slide into $\mathbf{O}$ without passing through e, so that one may say that $\mathbf{O}$ is an *immediate* outcome of e, and that $e \rightarrowtail \mathbf{O}$ is an *immediate* transition. In the case of immediate transitions, one loses no essential causal

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

[20]That is, $H_{(e)} \cap \{h\} \neq \varnothing$ and $H_{(e)} \cap H_{\langle\overline{\mathbf{O}}\rangle} \neq \varnothing$ but $e < e' \rightarrow H_{(e')} \cap \{h\} = \varnothing$ or $H_{(e')} \cap H_{\langle\overline{\mathbf{O}}\rangle} = \varnothing$

information in passing from the scattered-outcome form, $e \rightarrowtail \mathbf{O}$, either to the chain form, $e \rightarrowtail O$, or to the propositional form, $e \rightarrowtail H_{\langle \mathbf{O} \rangle}$, or indeed $e \rightarrowtail H_{\langle O \rangle}$, for any $O \in \mathbf{O}$; and so we may use *basic transition* for any one of these forms of immediate transition, and *basic outcome of* e for the respective immediate outcomes.[21] In some of the development here, it is necessary to use the propositional form of outcomes, as in the treatment of conditional propensities in §10.4. For this purpose, recall that $H_{(e)}$ is the set of histories to which e belongs, and let Π_e be the set of all propositional basic outcomes of e: $\Pi_e = \{H_{\langle \mathbf{O} \rangle} : e \rightarrowtail \mathbf{O}$ is a basic transition$\}$. It is a calculation that Π_e is a (possibly trivial) partition of $H_{(e)}$. When, however, we are relying on the causal ordering of BST, it is essential to invoke the event form of outcome.

A basic transition can supply no causal account of its outcome other than itself; but it can sometimes supply causal information about other transitions, $\mathbf{I} \rightarrowtail \mathbf{O}$. Thus, let $\mathbf{I} \rightarrowtail \mathbf{O}$ be an arbitrary transition to a scattered outcome event.

> e is a *prior cause-like (or causal) locus for* $\mathbf{I} \rightarrowtail \mathbf{O}$ if what can happen immediately after e might make a difference to whether $\mathbf{I} \rightarrowtail \mathbf{O}$ occurs. The technical definition requires that e lies in the past of $\mathbf{O}$, and that there be a history h in which the initial event $\mathbf{I}$ finishes (i.e., $\mathbf{I} \subseteq h$) such that e is a choice point between h and the occurrence of $\mathbf{O}$: $h \perp_e H_{\langle \mathbf{O} \rangle}$. [30]

It is consistent and even interesting to suppose that some cause-like loci for $\mathbf{I} \rightarrowtail \mathbf{O}$ fail to be in the past of $\mathbf{O}$, a situation that is so strange that we officially label it "funny business" as explored in FBBST, NCCFB, and IMC. I do not know how the existence of funny business should influence our thoughts on causation and propensities; therefore, in order to move on with the work, I simply lay down the assumption, for the remainder of this essay, that there is no funny business: All cause-like loci of $\mathbf{I} \rightarrowtail \mathbf{O}$ are assumed to lie in the past of $\mathbf{O}$.[22] Now add that when e is in the past of $\mathbf{O}$, one can calculate that there

[21] This terminology is a variant of that in earlier BST publications, but I foresee no confusion. Incidentally, remark that transitions to disjunctive outcome events are not classified as "basic."

[22] Alan Ross Anderson used to advise that if some exploration of yours is hung up by not knowing whether an essential presupposition is true or not, just assume its truth and press on.

is a unique member, $\Pi_e\langle\mathbf{O}\rangle$, of Π_e that is consistent with the occurrence of $\mathbf{O}$, and we are ready to say which basic transitions are *causae causantes*, or *originating causes*, of $\mathbf{I}\rightarrowtail\mathbf{O}$.

> By definition, a transition is a *causa causans* of $\mathbf{I}\rightarrowtail\mathbf{O}$ just in case it has the form $e \rightarrowtail \Pi_e\langle\mathbf{O}\rangle$, where e is a prior cause-like locus for $\mathbf{I}\rightarrowtail\mathbf{O}$. I write $cc(\mathbf{I}\rightarrowtail\mathbf{O})$ for the set of all *causae causantes* of $\mathbf{I}\rightarrowtail\mathbf{O}$. [31]

Such a transition is a transition from "it is open whether $\mathbf{O}$ will remain possible" to "at least for now, the occurrence of $\mathbf{O}$ remains a possibility." It is critical that the set of *causae causantes* of $\mathbf{I}\rightarrowtail\mathbf{O}$ form a set each member of which is an *inns* condition (*i*nsufficient but *n*on-redundant part of a *n*ecessary and *s*ufficient condition of $\mathbf{I}\rightarrowtail\mathbf{O}$); see CC for the proof as well as for a general theory of causation. Here I merely mention the generalization to Mackie's well-known *inus* conditions, where the "u" stands in for "unnecessary." Inus conditions are appropriate when an outcome event, $\boldsymbol{O}$, is an event that is *essentially* disjunctive in nature, consisting of a set of pairwise inconsistent outcome events, each with its own past history. The idea is that the event, $\boldsymbol{O}$, would count as *the same*, even if one considered alternative ways in which *it* could have happened.[23] The prior causal loci of $\boldsymbol{O}$ are the prior causal loci of any one of the scattered outcome events, $\mathbf{O}$, contained in $\boldsymbol{O}$; and similarly the *causae causantes* of a transition $\mathbf{I}\rightarrowtail\boldsymbol{O}$ are the *causae causantes* of any of the transitions $\mathbf{I}\rightarrowtail\mathbf{O}$, for $\mathbf{O}$ a member of $\boldsymbol{O}$. In other words,

$$cc(\mathbf{I} \rightarrowtail \boldsymbol{O}) = \bigcup\{cc(\mathbf{I} \rightarrowtail \mathbf{O}) : \mathbf{O} \in \boldsymbol{O}\}.$$

Thus, each of the *causae causantes* of $\mathbf{I}\rightarrowtail\boldsymbol{O}$ is an inus condition.[24]

[23]Disjunctive outcome events contrast with scattered outcome events in the same way that Lewis's "robust" events contrast with his "fragile" events. One should not, however, neglect that the BST concepts are defined in the context of a mathematically rigorous theory.

[24]The following may be used as an informal guide to the "inus" terminology. Picture a proposition, P, and put it into disjunctive normal form, say $(A\&B\&C)\vee(D\&E)$. Suppose, for each conjunction, that none of its conjuncts is implied by the conjunction of the rest of its conjuncts. Suppose further that no disjunct is implied by the disjunction of the rest of the disjuncts. Then B, for instance, will be an *i*nsufficient but *n*on-redundant "part" of $A\&B\&C$, which in turn will be an *u*nnecessary but *s*ufficient condition of P. Do not, however, become so bewitched by the picture that you neglect that the topic is concrete events, not mere propositions.

9 Probabilities added to BST (BSTP): the WBM analysis

Weiner and 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 a certain amount of mutual assistance.[25] One shared idea of the "WBM" constructions is that you look to the set $cc(\mathbf{I} \rightarrowtail \boldsymbol{O})$ of *causae causantes* as containing all and only the information needed for assigning a propensity to $\mathbf{I} \rightarrowtail \boldsymbol{O}$ (p. 491). There are now two ways to proceed, one more global and one more local. Each basic transition $e \rightarrowtail H_1$ among those *causae causantes* has its own *local* causal alternatives $e \rightarrowtail H_2$ based on keeping its initial e and letting its basic outcome H_2 vary over all the basic outcomes, Π_e, 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 of $\mathbf{I}$ giving rise to $\boldsymbol{O}$. By taking the family of all subsets of the sample space, Müller obtains the boolean algebra, F, that must underlay the causal-probability field of sets in the context of which giving a propensity to $\mathbf{I} \rightarrowtail \boldsymbol{O}$ makes good 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 $\{cc(\mathbf{I} \rightarrowtail \boldsymbol{O}))\}$ is bound to be an element of the algebra, F, that is enough. On the global scheme, one may now define the *causal-probability space* appropriate to $\mathbf{I} \rightarrowtail \boldsymbol{O}$ as $\langle A, F, \mu \rangle$. This approach is smoothest, deepest, most powerful, and most flexible.

A second way to proceed, exploited in Weiner and Belnap 2006, derives the wanted propensity of $\mathbf{I} \rightarrowtail \boldsymbol{O}$ (that is, the propensity for $\mathbf{I}$ to give rise to $\boldsymbol{O}$) from a calculation based on assigning propensities to its *causae causantes* in a more local fashion. Define a *local propensity space* for $\mathbf{I} \rightarrowtail \boldsymbol{O}$ as a pair $\langle T, pr \rangle$, where T is the set of all basic transitions each of whose initials is a

[25] There is, however, no doubt that B learned much more from W and M about probabilities than the other way around; furthermore, the cited article by M greatly exceeds this essay in both extent and depth. A third teacher has been the author of Placek 2000c. As I said to Alan Ross Anderson over half a century ago, and have since frequently repeated in conversations, probabilities make my head go woggly.

prior causal locus of $\mathbf{I} \rightarrowtail \boldsymbol{O}$ (that is, $T = \{e \rightarrowtail H : e \in pcl(\mathbf{I} \rightarrowtail \boldsymbol{O})\ \&\ H \in \Pi_e\}$), and where pr assigns a propensity in the interval $[0, 1]$ to each member of T, subject only to the *local* restriction that for each $e \in pcl(\mathbf{I} \rightarrowtail \boldsymbol{O})$,

$$\sum_{H \in \Pi_e} pr(e \rightarrowtail H) = 1.$$

That is, from the perspective of e, something must happen; in other words, it is guaranteed that e gives rise to Π_e considered as a disjunctive outcome event—the propensity for e to do so is 1. This second way, which begins with a separate, local, and independent measure for each set of basic transitions based on a prior causal locus, e, of $\mathbf{I} \rightarrowtail \boldsymbol{O}$, although not as widely applicable as the "global" way, has the merit of firm and direct rooting in the idea of propensity. It quickly leads to intuitive judgements in easy cases, cases in which simple multiplication and addition fit the phenomena, as follows:

> First, for each scattered outcome, $\mathbf{O}$, contained in $\boldsymbol{O}$, obtain the propensity of $\mathbf{I} \rightarrowtail \mathbf{O}$ by multiplying together all the propensities of all its *causae causantes*. Second, where $\boldsymbol{O}$ is a disjunctive outcome, obtain the propensity of the transition, $\mathbf{I} \rightarrowtail \boldsymbol{O}$ by adding the propensities of each $\mathbf{I} \rightarrowtail \mathbf{O}$, for $\mathbf{O}$ a member of $\boldsymbol{O}$. [32]

The justification of the *multiplication* is based partly on the non-redundancy of the the *causae causantes* of $\mathbf{I} \rightarrowtail \mathbf{O}$, which is a crucial part of the idea of inns conditions. Negatively put, it would make little sense to multiply the propensities attached to each *causa causans* of $\mathbf{I} \rightarrowtail \mathbf{O}$ were one or more of the *causae causantes* redundant. BSTP provides this justification by rigorous proof of non-redundancy (in Belnap 2005b) based on very general assumptions about the causal structure of *Our World*. Also at work is the absence of causal or modal dependencies between SLR-related *causae causantes*. In other words, SLR-related *causae causantes* must be causally or modally independent, so that any combination of basic outcomes, one from each, is possible. This is precisely a matter of "no (modal) funny business" in the sense of FBBST, NCCFB, and IMC. Finally, *causae causantes* of $\mathbf{I} \rightarrowtail \mathbf{O}$ must be stochastically independent, which is just to say that multiplication works for *causae causantes*; or, in other words, that there is no stochastic funny business in the phenomena:

$$pr(\mathbf{I} \rightarrowtail \mathbf{O}) = \prod_{(e \rightarrowtail H) \in cc(\mathbf{I} \rightarrowtail \mathbf{O})} pr(e \rightarrowtail H).$$

"No stochastic funny business" is by no means a universal principle rooted in nature; it simply marks the limitation to simple cases of the applicability of the present notions.

The justification of the *addition* is indeed also via *causal* considerations, but it is closer to the surface, since I included "pairwise inconsistency" of *events* (not mere propositions) in the definition of disjunctive outcomes $\boldsymbol{O}$. It is then natural to lay it down that addition works for the phenomena under consideration:

$$pr(\mathbf{I} \rightarrowtail \boldsymbol{O}) = \sum_{\mathbf{O} \in \boldsymbol{O}} pr(\mathbf{I} \rightarrowtail \mathbf{O}).$$

10 Application to corkscrew story

Here I show how the simple multiply-and-then-add account works out in connection with an "inversion" story that Salmon invoked in defending his reluctance to give a probabilistic reading to propensities. Indeed, 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 propensities 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).[26]

In retelling the story, I shall rely on an annotated BSTP diagram, namely Figure 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 in the sense of being the maximum, in the causal order, of those point events at which it is historically open whether the output of M_i will be good or defective. e_3 is the choice point for the Inspector (when he has a choice), the last point event at which he has not yet chosen which corkscrew to pick. D_i $[G_i]$ is a defective [good] corkscrew produced by machine M_i. d_i, g_i $(i = 1, 2)$ represent two possible immediate scattered outcomes of each machine choice point e_1, e_2, whereas $d_i^*(i = 1, 2)$ represent two possible immediate scattered outcomes of the picking by the Inspector from Box-β at e_3. If two features have the same label in the diagram, they represent the same event; otherwise, not. The boxes α and β are only receptacles. In Salmon's story, machines M_1 and M_2 (the grey ellipses) make different numbers of corkscrews per day, and there are also different propensities for each machine to make, on each run, a defective corkscrew. Figure 6 pictures a much simpler story. Why? To illustrate a BSTP story about many physical objects requires drawing many tracks through the space-time of each history, which is at best confusing. For this reason, with the expectation that it won't 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

[26]Salmon's language attaches propensities to things; without further comment, I translate his story into language that attaches propensities to transitions.

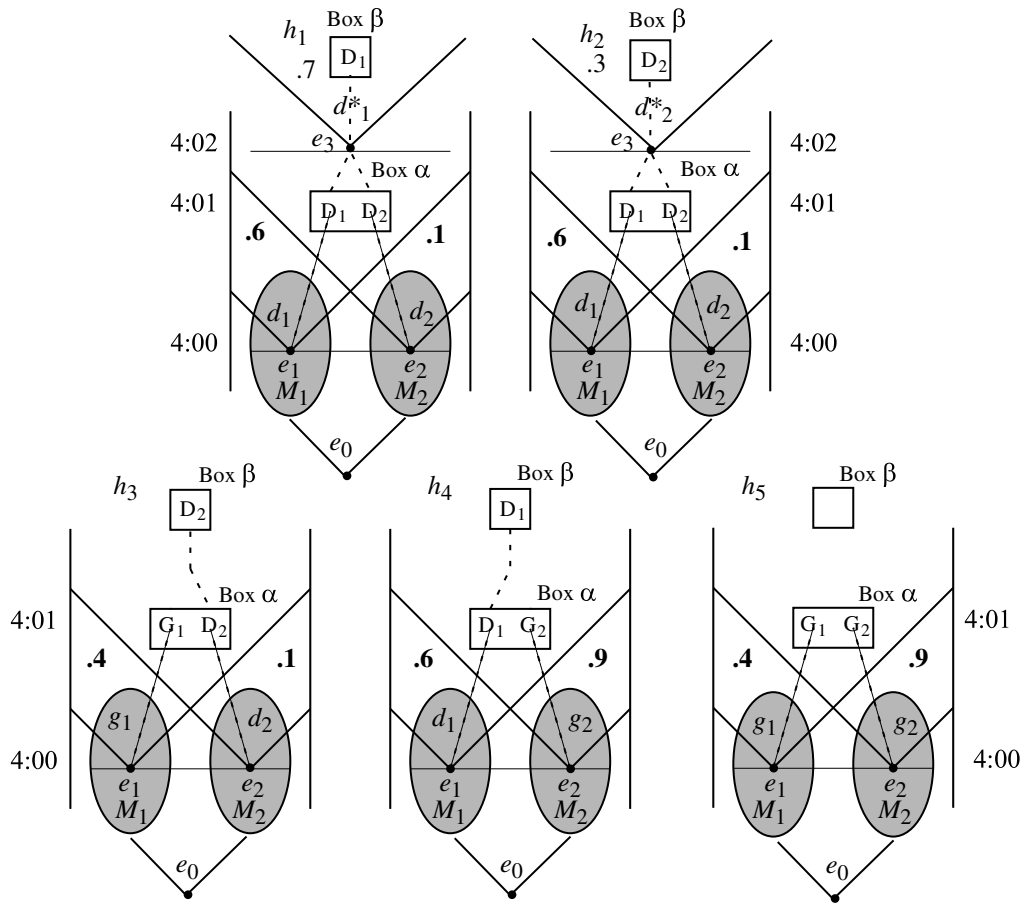

Figure 6: The Corkscrew Story.

make just one corkscrew per day. Furthermore, I shall be speaking of the possible happenings on a single day.[27] (This expository strategy also reduces the temptation to measure propensities in terms of frequencies.) The causal analysis of the story as pictured leads us to four *causae causantes*, each of which we baptize:

$$\begin{aligned} t_1 &= (e_1 \rightarrowtail g_1) \\ t_2 &= (e_1 \rightarrowtail d_1) \\ t_3 &= (e_2 \rightarrowtail g_2) \\ t_3 &= (e_2 \rightarrowtail d_2) \end{aligned} \qquad [33]$$

The propensity for each machine to turn out a defective corkscrew on the given day is given numerically in the interval $[0, 1]$; say there is a .4 propensity for e_1 to give rise to d_1, which is an immediate scattered outcome event that represents the making of a defective corkscrew by machine M1. Consequently there is a .6 propensity for e_1 to give rise to g_1, which is an immediate scattered outcome event that represents the making of a good corkscrew by M1. Let the propensities be .1 for e_2 to give rise to the making by M2 of a defective corkscrew (d_2), and .9 for e_2 to give rise to the making of a good one (g_2). In BSTP these propensities qualify, respectively, four of the relevant *causae causantes* having one of e_1 or e_2 as initial:

$$\begin{aligned} pr(t_1) &= pr(e_1 \rightarrowtail g_1) = .4 = p_1 \\ pr(t_2) &= pr(e_1 \rightarrowtail d_1) = .6 = p_2 \\ pr(t_3) &= pr(e_2 \rightarrowtail g_2) = .9 = p_3 \\ pr(t_4) &= pr(e_2 \rightarrowtail d_2) = .1 = p_4 \end{aligned} \qquad [34]$$

The rightmost column represents an additional abbreviative convention.

It is part of Salmon's story that one day

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

[27] Although it is not explicitly diagrammed, Salmon evidently supposed that propensities in the story were stable over time. That's fine. It is to be borne in mind, however, that the rock-bottom idea of propensity is for a unique concrete initial event to give rise to a unique outcome event. Enduring propensities amount to (ideally) universal generalizations.

Salmon's Humphreys-inspired Bayes-type question then arises:

> What shall we say about the chosen corkscrew's propensity to have been made by M_1? [35]

His answer: "it would hardly be reasonable to speak of the *propensity* of that corkscrew to have been produced by" M_1 (ibid.) It certainly does sound like double talk. The Inspector examines a corkscrew, only to find it defective. Would anyone want to say that the corkscrew lying quietly in his hand has a certain propensity to—what? A propensity to *have been made* by machine $M1$ rather than $M2$? But that seems nonsense. The reason it is nonsense has precious little to do with propensities, probabilities, or Bayes: It's a purely causal matter. There is further discussion of "inversions" in §10.3, where we find some sensible inversions by following Miller's lead.

Recasting the Bayes question in BSTP event terms makes it even clearer, if possible, that propensities cannot run backwards. "The propensity for d_1 to give rise to e_1" is not a piece of English for which one ought to spend time in search of something that it could mean.[28] 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 picked by the Inspector was in fact made in the settled past either by machine $M1$ or by machine $M2$, and current chancy propensity is nonexistent.

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

[28] In contrast, "the propensity for d_1 to have been given rise to by e_1" is merely a harmless (albeit confusing) passive transform of "the propensity for e_1 to give rise to d_1.

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 of a defective corkscrew to be turned out by each machine," is, though sounding rather like a standard passive transformation, more obscure 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 don't make the attribution 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 pick of a corkscrew (or can opener) a *causa causans*, and one not to be ignored in drawing philosophical pictures or morals. In §10.1 below, I specify some details governing the Inspector's contribution.

Figure 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 Figure 6 shows ten 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 Figure 6, except for those interrupted by e_3, 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 [34] that stretch forward from the initials e_1 and e_2.

10.1 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 chooses which defective corkscrew to transfer to Box β.[29] (Otherwise we might say that the inspection was fixed.) I represent the two possible picks by two additional *causae causantes*,

$$t_5 = (e_3 \rightarrowtail d_1^*) \text{ and } t_6 = (e_3 \rightarrowtail d_2^*),$$

to which I will shortly award propensities.

You might infer 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 doesn't 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, there is a propensity of .3 for the picking situation at e_3 to give rise to a pick of M_1's defective output, as represented by scattered outcome event d_1^*, and (of course) a .7 propensity for picking the defective output of M_2, as represented by scattered outcome event d_2^*. Perhaps we shouldn't; perhaps we should just assume that the chances are 50/50, in order to stay closer to the presumed intent of Salmon's

[29] Recall that Salmon described the Inspector's (or someone's) pick as "random." One standard use of "random pick" connotes "mindless pick," but I avoid the language of randomness in the firm belief that a choice among multiple options can be indeterministic without being mindless. See §4.1.

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, is given by adding to [34] the two possible transitions from e_3. (I include the p_i-names of these propensities for easier cross-reference.)

$$\begin{array}{l} pr(e_1 \rightarrowtail g_1) = .4 = p_1 \\ pr(e_1 \rightarrowtail d_1) = .6 = p_2 \\ pr(e_2 \rightarrowtail g_2) = .9 = p_3 \\ pr(e_2 \rightarrowtail d_2) = .1 = p_4 \\ pr(e_3 \rightarrowtail d_1^*) = .7 = p_5 \\ pr(e_3 \rightarrowtail d_2^*) = .3 = p_6 \end{array} \qquad [36]$$

From these tables, because they include propensities of all the relevant *causae causantes*, we can construct a table of some non-basic propensities, in particular a table of propensities of various (non-basic) transitions to certain scattered outcome events. The following five may be of some interest since the scattered outcome event of each resides in a single history, h_i.[30] First one must locate by causal analysis the various *causae causantes* of these transitions:

$$\begin{array}{l} cc(e_0 \rightarrowtail d_1^*) = \{t_2, t_4, t_5\} \\ cc(e_0 \rightarrowtail d_2^*) = \{t_2, t_4, t_6\} \\ cc(e_0 \rightarrowtail (g_1 \cup d_2)) = \{t_1, t_4\} \\ cc(e_0 \rightarrowtail (d_1 \cup g_2)) = \{t_2, t_3\} \\ cc(e_0 \rightarrowtail (g_1 \cup g_2)) = \{t_1, t_3\} \end{array} \qquad [37]$$

Having done one's causal homework, the five propensities fall out by multiplication:

$$\begin{array}{l} pr(e_0 \rightarrowtail d_1^*) = p_2 \times p_4 \times p_5 = .6 \times .1 \times .7 = .042 \\ pr(e_0 \rightarrowtail d_2^*) = p_2 \times p_4 \times p_6 = .6 \times .1 \times .3 = .018 \\ pr(e_0 \rightarrowtail (g_1 \cup d_2)) = p_1 \times p_4 = .4 \times .1 = .04 \\ pr(e_0 \rightarrowtail (d_1 \cup g_2)) = p_2 \times p_3 = .6 \times .9 = .54 \\ pr(e_0 \rightarrowtail (g_1 \cup g_2)) = p_1 \times p_3 = .4 \times .9 = .36 \end{array} \qquad [38]$$

[30] Recall that the g_i and d_i are scattered outcome events, so that a set-theoretical union of a consistent pair of them issues in another scattered outcome event. $(d_1 \cup d_2)$ is not on the list because that scattered outcome event occurs in two histories.

Since causal analysis decrees that given the occurrence of e_0, exactly one of the five scattered outcomes must occur, it is hardly a surprise that the sum of these five propensities is 1.

None of the propensities of table [38] attach to the transition from e_0 either to $\boldsymbol{O}_1$ = "the inspector picks a defective corkscrew" or to $\boldsymbol{O}_2$ ="the inspector picks a corkscrew made by machine M_1," because each of these two are transitions to disjunctive events. The outcome of the first, which can be causally situated in four ways, is $\boldsymbol{O}_1 = \{d_1^*, d_2^*, (g_1 \cup d_2), (d_1 \cup g_2)\}$, and of the second, which can be causally situated in two ways, is $\boldsymbol{O}_2 = \{d_1^*, (d_1 \cup g_2)\}$. Calculation of the propensities of transitions to these disjunctive outcomes from e_0 is not far behind; just add, using table [38]. This gives

$$\begin{aligned} pr(e_0 \rightarrowtail \boldsymbol{O}_1) &= .042 + .018 + .04 + .54 = .64, \text{ and} \\ pr(e_0 \rightarrowtail \boldsymbol{O}_2) &= .042 + .54 = .582. \end{aligned} \qquad [39]$$

10.2 Transitions are relational

Let us pause to re-emphasize the relational character of transitions in BSTP. Consider first, the attribution of a forward propensity with "the propensity for each machine to make a defective corkscrew" on the day in question. By Figure 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, $pr(e_1 \rightarrowtail d_1)$, which by [36] is .6; and also the propensity of $pr(e_2 \rightarrowtail d_2)$, which, is .1. So ".6" and ".1" are the straightforward answers to "For each machine, what is the propensity (at 4:00) for it to make a defective corkscrew?"[31] The English reading is intended to emphasize 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 they deserve consideration.

[31] Although this sounds as if the propensity is attached to a time, in fact I am using the time only to coordinate with "the machine" in order to indicate, for each machine, a particular initial *event*—namely, e_1 and e_2 respectively.

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, giving due recognition to each of initial and outcome. The point is that varying either initial or outcome varies the propensity. This is illustrated by comparing the list [38], which kept the initial the same, while varying the outcome, with the next list, which varies the initial while keeping the outcome constant. (The first two entries turn out the same because $cc(e_1 \rightarrowtail d_1^*)$ and $cc(e_2 \rightarrowtail d_1^*)$ are, identically, $\{t_1, t_4, t_5\}$.)

$$\begin{aligned} pr(e_1 \rightarrowtail d_1^*) &= p_1 \times p_4 \times p_5 = .042 \\ pr(e_2 \rightarrowtail d_1^*) &= p_1 \times p_4 \times p_5 = .042 \\ pr(e_3 \rightarrowtail d_1^*) &= p_5 = .7 \end{aligned} \qquad [40]$$

10.3 Inversion

Having the advantage of a thorough working-out of the story, let's go back now and say what can be said about Salmon's inversion question [35], and in particular 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 (unless there is funny business) the initial must *precede* the outcome in the causal ordering provided by *Our World*. Keep in mind that this has nothing to do with states, nor with 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 an initial event to one of its possible outcomes, an outcome that must inevitably commence after (possibly immediately after) the initial finishes. So given that $\mathbf{I} \rightarrowtail \overline{\mathbf{O}}$ is a proper transition,

one finds that $\overline{\mathbf{O}} \rightarrowtail \mathbf{I}$ cannot be a well-formed transition. That is so deeply ingrained in BST that it seems unimaginable that we should be able, in that theory, to make sense of inversion of initial and outcome simply by tacking on probabilities. One should not, however, use this observation to prove too much: It is entirely possible that the same set of point events be both the initial of one transition and the outcome of another.

Second, even with all this, it remains worthwhile to make all possible sense of inversion as sometimes expressed in English by finding something relevant in BSTP theory. Three things are useful, one from Müller's 2005 work on BSTP, one from Miller, and one from our discussion of Miller: (1) By Müller 2005, there is *always* a boolean algebra and a field of probabilities that are *naturally* generated by a set of BSTP transitions. Boolean algebras permit arbitrary inversions if conditional probability is defined in the standard way. (2) If the English is parsed as Miller suggests in his extended notation (§6), some inversions in English seem to make sense. (3) I suggested in §6, especially in [17], 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 (*not* time) in the extended notation of Miller. I next work out this suggestion at some length by way of treating "conditional propensities" from the point of view of the propensity-space structuring of BSTP.

10.4 Conditional propensities

I begin by calling attention to the idea of a three-termed relation corresponding to the relation $pr_{\tau_i}(A_{\tau_j}|B_{\tau_k})$ used in common by Humphreys and Miller, and read by them, respectively, in the following ways:

> $[pr_{\tau_i}(A_{\tau_j}|B_{\tau_k})]$ is the propensity at time τ_i for A to occur at time τ_j, conditional upon B occurring at time τ_k (Humphreys; see [13]).
>
> $[pr_{\tau_i}(A_{\tau_j}|B_{\tau_k})]$ is 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 (Miller; see [15]). [41]

Since neither writer has a definite *theory* of "A occurs (or comes to pass) at time τ," this may seem an unprofitable beginning. BSTP, however, does have a clear and rigorous theory of the occurrence of an event, which makes it reasonable for BSTP to adopt the idea of the Humphreys-Miller three-termed relation to serve as the carrier of conditional propensities, making such changes as are proper. For reasons given earlier, in §6, I eliminate all references to times such as occur in both versions, and also to the world, as in the Miller version. The three terms of the relation, then, will be local events, as defined in BSTP. There are three restrictions. First, the subscript on pr must name an initial event, $\mathbf{I}$. Second, the remaining two argument-places must be filled by names of outcome events, $\overline{\mathbf{O}}_1$ and $\overline{\mathbf{O}}_2$: $pr_{\mathbf{I}}(\overline{\mathbf{O}}_1|\overline{\mathbf{O}}_2)$. Third, the initial, $\mathbf{I}$, must be causally prior to each of $\overline{\mathbf{O}}_1$ and $\overline{\mathbf{O}}_2$ in the "most plausible sense" of [28]. There is no similar restriction governing the causal order of $\overline{\mathbf{O}}_1$ and $\overline{\mathbf{O}}_2$ relative to each other. As I indicated in connection with Figure 2, the logic of conditional probabilities need pay no attention to the causal- or time- order of condition and conditioned, and the same is true of conditional propensities.

Thus, I am entering

$$pr_{\mathbf{I}}(\overline{\mathbf{O}}_1|\overline{\mathbf{O}}_2) \qquad [42]$$

as new BSTP notation, with a BSTP reading that corresponds to the Humphreys and Miller readings of $pr_{\tau_i}(A_{\tau_j}|B_{\tau_k})$, namely,

> the propensity for an initial event $\mathbf{I}$ to give rise to outcome event $\overline{\mathbf{O}}_1$, given that it gives rise to outcome event $\overline{\mathbf{O}}_2$. [43]

Although certainly not mandatory, it seems good to say that an instance of $pr_{\mathbf{I}}(\overline{\mathbf{O}}_1|\overline{\mathbf{O}}_2)$ is a *conditional propensity*. In contrast with Humphreys and Miller, I do not take the notation for conditional propensities to be adequately explained by its English reading [43]. Here is a try for an exact definition—a definition, however, that turns out to be unsatisfactory:

$$pr_{\mathbf{I}}(\overline{\mathbf{O}}_1|\overline{\mathbf{O}}_2) =_{df}(?)\ pr(\mathbf{I} \rightarrowtail (\overline{\mathbf{O}}_1 \cap \overline{\mathbf{O}}_2)) \div pr(\mathbf{I} \rightarrowtail \overline{\mathbf{O}}_2). \qquad [44]$$

This trial definition of a conditional *propensity* should remind you of the standard definition of a conditional *probability*, $P(A|B) = P(A \cap B) \div P(B)$.

The key structural commonality is that for "$\cap$" to make sense here, not only must its arguments be sets, but in addition its arguments must be from (and into) a family of sets interpretable as sets of "cases"; outcome events, although sets, cannot be counted on in this way to be sets of "cases." The trouble with [44] is that the intersection of two outcome events need not be an outcome event. To take the simplest counterexample, remark that the intersection of two chain outcome events is very likely to be the empty set. The most straightforward solution to the problem is to transfer to the propositional form of outcome, namely, $H_{\langle\overline{\mathbf{O}}\rangle}$, which is the set of histories in which $\overline{\mathbf{O}}$ occurs.[32] Taking this tack, we may (and do) replace [44] with the following.

$$pr_{\mathbf{I}}(\overline{\mathbf{O}}_1|\overline{\mathbf{O}}_2) =_{df} pr(\mathbf{I} \rightarrowtail (H_{\langle\overline{\mathbf{O}}_1\rangle} \cap H_{\langle\overline{\mathbf{O}}_2\rangle})) \div pr(\mathbf{I} \rightarrowtail H_{\langle\overline{\mathbf{O}}_2\rangle}). \qquad [45]$$

I enter two examples drawn from the corkscrew factory. For an example that parrots Salmon's query [35] while still exhibiting the form of a BSTP conditional propensity, consider the following question.

> What is the propensity for the event e_0 (in Figure 6) to give rise to the Inspector's taking a corkscrew made by machine $M1$, given that e_0 gives rise to his taking a defective corkscrew? [46]

This seems an adequate replacement for [35]. 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'll persevere. The calculation is this:

The transition from e_0 to "The Inspector takes a defective corkscrew" is a transition to a disjunctive outcome that can happen in four causally distinct ways: $\boldsymbol{O}_1 = \{d_1^*, d_2^*, (g_1 \cup d_2), (d_1 \cup g_2)\}$, whence $H_{\boldsymbol{O}_1} = \{h_1, h_2, h_3, h_4\}$. The outcome of the transition from e_0 to "The Inspector takes a corkscrew produced by $M1$" can happen in two causally distinct ways, represented by $\boldsymbol{O}_2 = \{d_1^*, (d_1 \cup g_2)\}$, so that $H_{\boldsymbol{O}_2} = \{h_1, h_4\}$. Note: "Take" is not intended to imply "picked" or "chose," so the Hobson's choices count. Since $(H_{\langle\boldsymbol{O}_1\rangle} \cap H_{\langle\boldsymbol{O}_2\rangle}) = H_{\langle\boldsymbol{O}_2\rangle}$, the calculation [39] suffices for the conditional propensity,

[32] The definition of $H_{\langle\overline{\mathbf{O}}\rangle}$, in accord with [29], depends on the outcome-type of $\overline{\mathbf{O}}$. In the simplest case, when $\overline{\mathbf{O}}$ is an outcome chain, $H_{\langle\overline{\mathbf{O}}\rangle} = \{h : h \cap \overline{\mathbf{O}} \neq \varnothing\}$

$$pr_{e_0}(\boldsymbol{O}_2|\boldsymbol{O}_1) =_{df} pr(e_0 \rightarrowtail (H_{\langle \boldsymbol{O}_2 \rangle} \cap H_{\langle \boldsymbol{O}_1 \rangle})) \div pr(e_0 \rightarrowtail H_{\langle \boldsymbol{O}_1 \rangle}). \qquad [47]$$

and it turns out that $pr_{e_0}(\boldsymbol{O}_2|\boldsymbol{O}_1) = .582 \div .64 = .909$.

The inversion, although of little interest, makes perfectly good technical sense: $pr_{e_0}(\boldsymbol{O}_1|\boldsymbol{O}_2)$ is the propensity for e_0 to give rise to "The Inspector taking a defective corkscrew" given that e_0 gives rise to "The Inspector taking a corkscrew made by machine $M1$," which is a boring 1, since in my story, among the corkscrews made by $M1$, the Inspector takes only defective corkscrews. That is, since $(H_{\langle \boldsymbol{O}_2 \rangle} \cap H_{\langle \boldsymbol{O}_1 \rangle}) = H_{\langle \boldsymbol{O}_2 \rangle}$, $pr_{e_0}(\boldsymbol{O}_1|\boldsymbol{O}_2) =_{df} pr(e_0 \rightarrowtail H_{\langle \boldsymbol{O}_2 \rangle} \div pr(e_0 \rightarrowtail H_{\langle \boldsymbol{O}_2 \rangle}) = 1$.

A slightly more entertaining case is provided by the propensity for e_0 to give rise to a "real pick" by the Inspector (as opposed to a Hobson's choice) given that it gives rise to a taking from $M1$, compared with its inverse. The "real pick" is represented by $\boldsymbol{O}_3 = \{d_1^*, d_2^*\}$—or, equivalently, by $\{(d_1 \cup d_2)\}$, the latter being a disjunctive outcome with a single outcome-event alternative, but one that nevertheless manifests disjunctiveness by occurring in more than one history. In either case, $H_{\langle \boldsymbol{O}_3 \rangle} = \{h_1, h_2\}$, and thus $(H_{\langle \boldsymbol{O}_2 \rangle} \cap H_{\langle \boldsymbol{O}_3 \rangle}) = \{h_1\}$, so that $pr_{e_0}(\boldsymbol{O}_3|\boldsymbol{O}_2) = pr(e_0 \rightarrowtail (H_{\langle \boldsymbol{O}_2 \rangle} \cap H_{\langle \boldsymbol{O}_3 \rangle})) \div pr(e_0 \rightarrowtail H_{\langle \boldsymbol{O}_2 \rangle}) = pr(e_0 \rightarrowtail \{h_1\}) \div pr(e_0 \rightarrowtail \{h_1, h_4\} = .042 \div .582 = .072$. The inverse, $pr_{e_0}(\boldsymbol{O}_2|\boldsymbol{O}_3)$, calculates to $pr(e_0 \rightarrowtail \{h_1\}) \div pr(e_0 \rightarrowtail \{h_1, h_4\}) = .042 \div .06 = .7$.

So that illustrates how, starting with causal probabilities and propensities attached only to *causae causantes*, one can have a rich feast of propensities, including conditional propensities.

11 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 by itself have the resources to represent. *Objective probability, standing alone, 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 apparently 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 mere probability. It was surprising to find out, as it were, that my BSTP reconstruction of the Humphreys 1985 argument didn't involve probabilities; the argument was entirely in terms of the *causal order*, 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 conditional propensities) 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 seems misguided. One cannot know unless the part of the theory involving causal and spatiotemporal 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 et al. 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 *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 given a certain amount of effort to trying to remain faithful, but, although I am far from sure, I doubt that there is much benefit at this stage of getting straight on all the variations in the syntax of the noun "propensity." Perhaps later. At this stage, the theory doesn't 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 §4 of Humphreys's Paradox, and in §10 of Salmon's corkscrew story.

7. Placek 2009 develops an alternative to BST that, while keeping much of its spirit, eliminates the notion of "history" in favor of more finite objects. Perhaps propensity theory would find this a home more comfortable than BST. That work is yet to be done.

8. 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 $\varnothing$, 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* cries out for the distinction between conditional probabilities $P(A|B)$ and propensities $pr(\mathbf{I} \rightarrowtail \mathbf{O})$. $P(A|B)$ is dealing with probabilities only, whereas $pr(\mathbf{I} \rightarrowtail \mathbf{O})$ absolutely requires attention to objective causal structure. It is hardly a new thought that inferring objective causal structure from statistics is no matter of definition.

9. How *should* we then interpret the propensity $pr(\mathbf{I} \rightarrowtail \mathbf{O})$? Departing from the WBM "local propensity space" of §9, let's go globally 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 *causae causantes* 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, representing $\mathbf{I} \rightarrowtail \mathbf{O}$. 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 antepenultimate position, is the most powerful result that anyone has obtained in BSTP theory.

10. 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."

11. A final challenge. BSTP privileges *events* over enduring (or perduring, or whatever) *things*. It does, however, cater to initial events $\mathbf{I}$ and outcome events $\mathbf{O}$ that are spread out spatio-temporally, and to transition events consisting of an initial and an outcome. Is this enough to get straight on an appropriate family of concepts that comes to grips with the intuitively appealing attribution of propensities qua tendencies or abilities or capacities or dispositions or powers to enduring things? No, because there is no apparatus for signalling which events are occurrences in the life of which things.

References

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

Belnap, N. (2002a). **(DTR)** Double time references: Speech-act reports as modalities in an indeterminist setting. In Wolter, F., Wansing, H., de Rijke, M., and Zakharyaschev, M., editors, *Advances in modal logic*, volume 3, pages 37–58. World Scientific Co. Pte. Ltd., Singapore. A preprint of this essay may be obtained from `http://www.pitt.edu/~belnap`.

Belnap, N. (2002b). **(FBBST)** EPR-like 'funny business' in the theory of branching space-times. In Placek and Butterfield 2002, pages 293–315. A preprint of this essay may be obtained from `http://philsci-archive.pitt.edu`.

Belnap, N. (2003a). **(BSTpp)** 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). **(NCCFB)** 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 Vanderveken, D., editor, *Logic, thought and action*, pages 291–313. Springer, Dordrecht.

Belnap, N. (2005b). **(CC)** 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., and Xu, M. (2001). ***(FF)*** *Facing the future: Agents and choices in our indeterminist world*. Oxford University Press, Oxford.

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

Glymour, C., Spirtes, P., and Scheines, R. (2001). *Causation, prediction and search*. MIT Press, Cambridge, 2nd revised edition edition. 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.

Lewis, D. (1986). *On the plurality of worlds*. Basil Blackwell, Oxford and New York.

Mackie, J. L. (1974). *The cement of the universe*. Oxford University Press, Oxford.

McCurdy, C. I. S. (1996). Humphrey's paradox and the interpretation of inverse conditional propensities. *Synthese*, 108:105–125.

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

Müller, T. (2002). Branching space-time, modal logic, and the counterfactual conditional. In Placek and Butterfield 2002, pages 273–291.

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

Müller, T., Belnap, N., and Kishida, K. (2008). **(IMC)** Funny business in branching space-times: infinite modal correlations. *Synthese*, 164:141–159.

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

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. (2000c). Stochastic outcomes in branching space-times. *British Journal for the Philosophy of Science*, 51:445–475.

Placek, T. (2009). Possibilities without possible worlds/histories.

Placek, T. and Belnap, N. (2010). Indeterminism, branching spacetimes, and Earman's pruning. *forthcoming*, ??:??

Placek, T. and Butterfield, J., editors (2002). *Non-locality and modality*. Kluwer Academic Press, Dordrecht.

Popper, K. (1959). The propensity interpretation of probability. *The British journal for the philosophy of science*, 10:25–42.

Prior, A. (1967). *Past, present, and future*. Oxford University Press, Oxford.

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

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

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

Simon, H. A. (1957). *Models of man: social and rational*. John Wiley & Sons, New York.

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

Thomason, R. H. (1970). Indeterminist time and truth-value gaps. *Theoria*, 36:264–281.

Thomason, R. H. (1984). Combinations of tense and modality. In Gabbay, D. and Guenthner, G., editors, *Handbook of philosophical logic, vol. II: extensions of classical logic*, volume 165 of *Synthese Library, Studies in epistemology*, pages 135–165. D. Reidel Publishing Company, Dordrecht.

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

Weiner, M. and 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 University Press, Cambridge.

Contents

1 Introduction **1**

2 What is a propensity? **3**

3 Probability theory **7**

4 Humphreys's Paradox **10**

4.1 Preliminary remarks 10

4.2 Humphreys's photon argument 13

4.3 Extended notation showing time-dependence 14

5 What does the extended notation mean? **16**

6 Miller: extended notation and time reversal **17**

7 Humphreys: two mirrors **22**

8 Branching space-times: a foundation for BSTP **28**

8.1 Grammar of propensities in BSTP 30

8.2 Possible point events e, the causal order $<$, and *Our World* in BST 32

8.3 Histories, consistency, categorizing events in BST 33

8.4 Concrete transitions, prior causal loci, *causae causantes*, and inus conditions in BST 37

9 Probabilities added to BST (BSTP): the WBM analysis **40**

10 Application to corkscrew story **42**

10.1 The Inspector 48

10.2 Transitions are relational 50

10.3 Inversion 51

10.4 Conditional propensities 52

11 Closing remarks **55**