# On the Everettian epistemic problem

Hilary Greaves

*Department of Philosophy, Rutgers University, Davison Hall, 26 Nichol Ave, New Brunswick, NJ 08901-2882, USA.*
*Email: hgreaves@rci.rutgers.edu.*

## Abstract

Recent work in the Everett interpretation has suggested that the problem of probability can be solved by understanding probability in terms of rationality. However, there are *two* problems relating to probability in Everett — one practical, the other epistemic — and the rationality-based program *directly* addresses only the practical problem. One might therefore worry that the problem of probability is only 'half solved' by this approach. This paper aims to dispel that worry: a solution to the epistemic problem follows from the rationality-based solution to the practical problem.

**Keywords:** Everett interpretation, probability, decision theory, confirmation

## 1  Introduction

Stochastic quantum mechanics (QMS) is an *indeterministic* theory: that is, it is a theory according to which an initial segment of a history of the universe has more than one dynamically *possible* future continuation. Further, it is a *chance* theory: that is, the theory assigns *chances* to the various possible futures. In the case of stochastic quantum mechanics, the chances are given by the Born rule.[1]

Everettian quantum mechanics (QME) is a *branching-universe theory*. A branching-universe theory is one according to which the history of reality has a branching structure: as one goes forward in time, branches split in to multiple copies. An initial branch-segment has, in general, more than one

---

[1] An (approximate) example: the spontaneous collapse theory of Ghirardi, Rimini and Weber (1986).

*real* future continuation. Further, QME is a *weighted* branching-universe theory: the theory assigns *weights* to the various branches. In Everettian quantum mechanics, the branch weights are given by a particular function that we'll refer to as the 'Everettian weight function'; this function is mathematically equal to the Born rule.[2]

The weight function in a weighted branching-universe theory is, *mathematically* speaking, a probability function on the space of branches: that is, it satisfies the Kolmogorov axioms. But more is presumably required: not every function that satisfies the Kolmogorov axioms will do the work that we need 'physical probability' to do. (For example: the standard area measure on the surface of my desk satisfies the Kolmogorov axioms, but I don't have much use for the phrase 'the probability of the aging inkblot on my desktop'. (Not the probability *that it exists*: it certainly exists. The probability of the inkblot itself.)) The question then arises: can branch weights do all the necessary work that chances do? If not, the Everett interpretation will be in trouble.

## 1.1  Two problems of probability

I said that not every function that satisfies the Kolmogorov axioms will 'do the work that we need physical probability to do'. But what exactly *is* that work? To specify exactly why Everettian probability might be problematic, we need a precise answer to this question.

The easiest way to identify the required roles of physical probability is to imagine stripping Everettian quantum mechanics of its branch weights, and to ask ourselves what would be deficient about the resulting theory. In a previous paper (Greaves, 2004), I suggested that such a theory would be deficient in exactly two respects: it could not answer either of the following.

---

[2]Deterministic theories, where the agent is ignorant of the future because ignorant of the initial conditions, present a third type of case. Examples of this third type include classical statistical mechanics, and Bohmian mechanics. For simplicity, I will omit such cases from the discussion of this paper. I take it that the special probability-related problems that are supposed to plague the Everett interpretation will have been solved if we can show that QME can provide an account of probability as satisfactory as that available to QMS.

The term 'weighted branching universe', and the distinction betweeen a 'weighted branching universe' and a 'minimal branching universe', was introduced by David Wallace (unpublished work). I am grateful to Wayne Myrvold for emphasizing the importance, for the purpose of seeing the analogies between Everettian quantum mechanics and a chance theory, of this distinction between weighted branching universes in general and Everettian quantum mechanics in particular. My presentation of the problem (with which he may or may not agree) draws heavily on Myrvold's suggestions.

**The Practical Problem:** How are we rationally to act, if we interpret quantum mechanics along Everettian lines? Suppose we are faced with a choice between, say, disaster on the spin-up branch and disaster on the spin-down branch. Given only that, whichever choice we make, there will be a disaster branch and a non-disaster branch, how could we ever have grounds for choosing?

**The Epistemic Problem:** How can we justify believing the theory on the basis of our empirical evidence, if we interpret quantum mechanics along Everettian lines? Given only that the theory predicted that the evidence that we have in fact observed would occur *on some branch* (and that the same is true of every other 'possible' string of evidence), how can we reasonably take our evidence to confirm the theory?

It is crucial to note that the same two problems would crop up in a stochastic theory that was stripped of its measure over possibilities (that is, in an indeterministic theory that was not a chance theory). It would be difficult to choose between disaster on spin-up and disaster on spin-down if one was told only that both spin-up and spin-down were *possible*; it would be difficult to take one's evidence to confirm one's theory if the theory predicted only that that evidence (along with every other string of evidence) was *possible*. But we normally think that admitting a *chance* measure over *alternative possibilities* can solve these problems: Choose disaster on spin-up iff spin-up has a *lower chance of occurring* than spin-down. Take your evidence to confirm your theory iff your theory predicted that evidence *with high chance*.[3] The worry for Everettian quantum mechanics, then, is that the analogous moves may make no sense (or make less sense, or be less justifiable) if we replace 'chance' with 'branch weight', and replace 'alternative possibilities' with 'coexisting branches'. That is why Everettian quantum mechanics has a *prima facie* problem with probability.

## 1.2   Solving the Practical Problem

The rationality-based program brings decision theory[4] to bear on the Practical Problem. Consider the set of all 'quantum bets': assignments of 'rewards' (things the agent cares about, e.g. monetary reward, ice cream, nuclear disaster) to Everettian branches. Expanding on an original suggestion by David

---

[3]I simplify. A more refined account of empirical confirmation will be used in section 3.

[4]The present section contains only a rather loose sketch of the decision-theoretic claims required to deal with chance theories and weighted branching theories. A more precise account will be given in section 2.

Deutsch (1999), David Wallace has argued (in several papers, e.g. (2003), (2005c)) for the following claim:

**(MEUB) Maximization of expected utility; branching case.** The rational agent's preferences over branching bets can be represented by a utility function $U$ on rewards-*on-branches* and a probability measure $p$ over branches, such that, for any two bets $a$, $b$, the agent prefers $a$ to $b$ iff $EU(a) > EU(b)$, where the expected utility $EU(a)$ of a bet $a$ is given by

$$EU(a) := \sum_{s \in S} p(s) \cdot U(a(s)). \tag{1}$$

Call $p$ the *decision-theoretic branch measure.* (We follow Savage (1972) in representing an act as a function: in this case, a function from branches to rewards.)

**(MEUB)** secures the existence of *some* probability measure $p$ over branches, but it does not say anything about what that measure should be. Deutsch and Wallace further argue for the following quantitative claim:

**(PE) Probability constraint: Everettian case.** Conditional on an assumption that Everettian quantum mechanics is true, the decision-theoretic branch measure $p$ for any rational agent is numerically equal to the Everettian branch weight measure.

This claim, *if* true, goes a long way towards putting Everettian QM on the same footing as a chance theory *as far as the Practical Problem is concerned.*[5] This is because the best way we have to precisify the intuitive advice 'choose disaster on spin-up iff spin-up has a *lower chance of occurring* than spin-down' is to make the analogous decision-theoretic claims for the case of a chance theory:

**(MEUI) Maximization of expected utility; indeterministic case.** The rational agent's preferences over indeterministic bets can be represented by a utility function $U$ on rewards-on-possible-outcomes and a probability measure $p$ over possible outcomes, such that, for any two bets $a$, $b$, the agent prefers $a$ to $b$ iff $EU(a) > EU(b)$, where the expected utility $EU(a)$ of a bet $a$ is given by

$$EU(a) := \sum_{s \in S} p(s) \cdot U(a(s)). \tag{2}$$

---

[5] Why not '*all* the way'? Because the claims **(MEUB)** and **(PE)** need to be generalized slightly before they can actually be applied to realistic cases; again, see section 2.

**(PC) Probability constraint; chance case.** Conditional on an assumption that a given chance theory is true, the decision-theoretic probability measure $p$ for any rational agent is numerically equal to the chance measure given by the theory in question.

The conditional claim — the claim that *if* the Deutsch-Wallace claims **(MEUB)** and **(PE)** are as well-justified as the analogous claims **(MEUI)** and **(PC)** *then* the Everettian is no worse off than (say) the GRW theorist vis-a-vis the Practical Problem — is relatively uncontentious. The contentious claim is that **(MEUB)** and **(PE)** *are* as well-justified as **(MEUI)** and **(PC)**. This claim has been defended extensively by Wallace in the papers already cited. I will have a little more to say about it below; for the most part, however, my approach will be to take this claim for granted — at least for the sake of argument — since my current interest is in examining the question of whether, even if true, it is *enough*.[6]

The (main) reason that **(MEUB)** and **(PE)** may not be enough is obvious, in the light of our above discussion of what we require from 'physical probability': we have so far said nothing at all about the Epistemic Problem. There is thus the danger that, while we would know how rationally to act if we had (somehow!) come to believe that QME was true, we can have no empirical justification for believing QME in the first place.

I will say more about this problem below. First, we need to distinguish two ways one might think about branching universes.

## 1.3   Thinking about branching: 'subjective uncertainty' and 'objective determinism'

Consider an agent who knows that she is in a branching universe, and that her branch is about to split. She therefore knows that, very soon, there will be two (or more) copies of her (along with the rest of her world). It will be possible for different things to happen on the two branches. For example, suppose that a pointer swings to point at the words 'spin-up' on one branch, while the other branch contains a pointer pointing at 'spin-down', and that our (pre-splitting) agent knows this too.

How should this agent, *prior* to the splitting, describe the future of her branch? There seem to be two *prima facie* possibilities:

---

[6] Wallace actually makes a stronger claim: that **(PE)** is *more* defensible than **(PC)** (e.g. (2005a), section 3.6). A similar claim has been made by Simon Saunders (2005). I am skeptical of this stronger claim, but I won't discuss it in this paper. I am exclusively interested, here, in the question of whether or not Everettian probability is in *at least as good* a shape as, say, probability in GRW.

**'Subjective uncertainty' (SU) semantics.** 'X might happen' is true iff X happens on some branch. So, in the example under consideration, spin-up might happen, and spin-down might happen. (It is not the case that spin-up-and-spin-down might happen, since that happens on no branch: there is a distinction between '$X \wedge Y$ might happen' on the one hand, and '$X$ might happen $\wedge Y$ might happen' on the other.)

**'Objective-deterministic' (OD) semantics.** 'X will happen' is true iff X happens on some branch. So, in the example under consideration, spin-up will happen, and spin-down will happen. (It is not the case that spin-up-and-spin-down will happen, since that happens on no branch: there is a distinction between '$X \wedge Y$ will happen' on the one hand, and '$X$ will happen $\wedge Y$ will happen' on the other.)

Note that on the SU semantics *but not on the OD semantics*, it makes sense for the agent to say that she is *uncertain* about whether or not spin-up will happen. According to the SU semantics, spin-up might happen and spin-down might happen, but it is not the case that both will happen; since the agent knows all this, she must be uncertain about which outcome will happen. According to the OD semantics, on the other hand, the agent knows that spin-up will (certainly) happen, and that spin-down will (certainly) happen, so there is nothing for her to be uncertain about. This will be crucial in our discussion of the Epistemic Problem.

Which is the *right* semantics? Well: all are agreed that the 'OD' semantics is a viable option. The disagreement is over whether or not the 'SU' semantics is *also* a viable option. This is an important question, but this, too, is a debate that the present paper will not enter.[7]

The relevance of the SU/OD debate to the present paper is the following. While there is little consensus on the question of whether or not the subjective-uncertainty semantics is viable, there *is* a fairly widespread consensus, shared by both 'Everettians' and 'anti-Everettians' alike, that the Everett interpretation is defensible *if and only if* SU is viable. Thus, 'Everettians' tend to argue by defending the SU semantics and showing how, *given* that semantics, one can defend the Everett interpretation. 'Anti-Everettians' tend not to dispute the claim that *if* the SU semantics *were*

---

[7]A brief guide to the literature on this debate: The first argument in defense of SU was given by Simon Saunders (1998). In (Greaves, 2004), I argued against SU and against Saunders' argument. David Wallace (2005b) gives a different argument in favor of SU, based on considerations of interpretive charity. Peter Lewis (2006) argues against Wallace. For the record, I am convinced by (a close cousin of) Wallace's argument, so I am no longer opposed to the use of the SU semantics.

available, *then* the Everett interpretation could be defended; they argue that the SU semantics is incoherent and that, in the absence of that semantics, one cannot defend the Everett interpretation.[8]

I think that this consensus is mistaken: I think that the Everett interpretation can be defended *whether or not* the SU semantics is viable. In a previous paper (Greaves, 2004), I argued for this position in the case of the Practical Problem: I argued, that is, that the decision-theoretic account of Everettian branch weights developed by Deutsch and Wallace is just as defensible on the OD semantics as on the SU semantics. The purpose of the present paper is to extend the claim to the Epistemic Problem: to argue, that is, that if the Practical Problem can indeed be solved in this way, then the Epistemic Problem can be solved on the OD semantics, as well as on the SU semantics.

Since agreement that the theory can be defended *if* the SU semantics is viable is relatively widespread, the more contentious claim is the claim that it can be defended without relying on the SU semantics. That is the claim I will try to establish in this paper. If I am right, then we need not get bogged down in the 'semantic dispute' if our only concern is the question of whether or not Everettian quantum mechanics is an acceptable physical theory. The two issues — physics and semantics — can be disentangled. We can first agree that the theory is acceptable, and afterwards discuss how we should talk about it. Or so I claim.

The rhetorical aims of the present paper are therefore as follows. I aim to convince 'Everettians' (more precisely: those Everettians who agree with the above consensus) that they have got the logic of the defense wrong: SU may happen to be true, but the theory would be perfectly defensible even if SU were false. It is therefore misleading to argue that SU is true and that *therefore* the theory's problems with probability can be solved. And I aim to convince 'non-Everettians' that, insofar as their reason for rejecting an Everett interpretation relied on an assumption that the *epistemic* prob-

---

[8]Thus David Baker, for example, writes in a recent paper:

> The most significant goal of [the decision-theoretic approach to Everettian probability] is ...to show that Everettian quantum mechanics is empirically verifiable. But it is not obvious how the "predictions" of the decision-theoretic approach, which take the form of normative prescriptions dictating how much we should care about particular outcomes*, could ever be confirmed or disconfirmed by experimental results.
>
> *[Baker's footnote:] This is not true of the SU version of Everett, of course. But as I have shown, if SU has a leg to stand on it is at least an extremely weak position compared with [OD]. (Baker, 2006)

lem cannot be solved *because* SU is incoherent, they should reconsider: the epistemic problem can be solved *even if* SU is incoherent.

**Structure of the paper.** The structure of the remainder of the paper is as follows. In section 2, I note that we need to generalize our decision-theoretic framework before we have a fully general solution even to the Practical Problem. I carry out the required generalization, and I suggest (briefly!) how the defense of the more general framework might follow from the defenses of the special framework that are already in the literature. It is this more general framework that we will need to use, in order to discuss the Epistemic Problem. In section 3, I argue that a Bayesian model of the Epistemic Problem is adequate, and I review the standard Bayesian account of empirical confirmation outside of the context of branching-universe theories. I sketch the way in which this account would need to generalize to the 'general case' in which both branching-universe and indeterministic theories are under consideration, in order for empirical confirmation to work in the way the Everettian requires. Section 4 discusses and rejects an initially plausible alternative updating policy for the branching-universe case that, if correct, would create havoc. In section 5, I argue that that generalization required to place Everettian quantum mechanics on the same epistemic footing as stochastic quantum mechanics is indeed the correct one; my strategy here will be to argue that the usual justifications for updating beliefs by conditionalization support the belief-updating policy that the Everettian requires in the branching-universe case. In section 6, I consider and respond to a possible objection to my account. Section 7 is the conclusion.

## 2  An integrated decision theory

In presentations of the rationality-based approach to date, and in the sketch given in section 1 above, the decision theory presented has assumed that the agent is certain she lives in a universe governed by Everettian quantum mechanics, and just has to decide how she should go about living her life, under that assumption.

This is a useful idealization for the purposes of getting started on the Practical Problem, but it is not general enough to provide a full solution even to that problem. No-one (or: no-one rational!) is ever actually certain of any given physical theory. The human predicament is one of making decisions without knowing *for sure* what the chances, or the branch weights, are. What we really need to know, therefore, is how a rational agent *whose*

8

*credences are distributed thus-and-so between such-and-such theories* should act.

The inadequacy of the idealization to a single physical theory becomes even more obvious when our concern is the Epistemic Problem. To investigate the Epistemic Problem at all, we *must* model an agent who has nonzero credence in *several* theories, of which QME is just one. There is no question of how the agent should manage her credences in theories, on the assumption that she already has credence 1 in some particular theory: she will be stuck with having credence 1 in that theory until the day she dies. What we need to know is how credences distributed non-trivially between *more* than one theory evolve in time, in response to evidence. The first task of this paper is therefore to sketch that more general decision-theoretic framework.

## 2.1 Histories, branches, centered worlds, worlds

Consider a 'general case': the case of agent who has nonzero credence that her world is a non-branching one, *and* non-zero credence that her world is a branching multiverse. Then her set of epistemic possibilities might have the structure illustrated in Figure 1; in the example illustrated, she is uncertain which of the possible worlds $w_1, \ldots, w_5$ is actual.
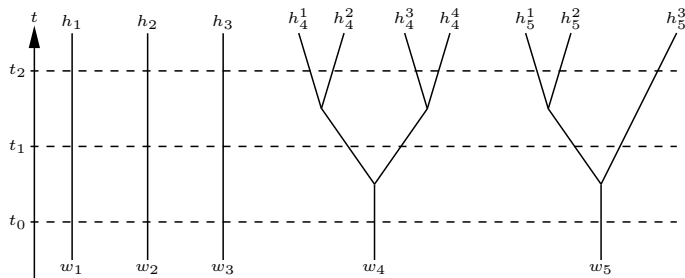


Figure 1: A 'general case'. Each tree structure (including the trivial 'trees' on the left) represents a single possible world. The agent's epistemic possibilities include both non-branching possible worlds $(w_1, w_2, w_3)$ and branching possible worlds $(w_4, w_5)$.

But there is more that our agent is uncertain about. Suppose the time now is $t_1$. Conditional on the assumption that she is in a particular branching multiverse, $w_4$ or $w_5$, our agent will still, in general, be uncertain as

to which branch at $t_1$ she is in, since each of these multiverses contains two branches at time $t_1$. (This is self-locating uncertainty, but self-locating uncertainty is uncertainty nonetheless.)

What of the future? Consider some time $t' > t$. Let $t'$ be far enough in the future that, if our agent is in a branching world, her branch will undergo splitting between $t$ and $t'$. Then, there is more than one branch-at-$t'$ that bears the 'is a future of' relation to our agent's branch-at-$t$; what is Alice's attitude to these futures? Well: *if we were working with the SU semantics*, we would say that Alice's uncertainty, *at $t$*, extended to which of the branches $b_1, b_2$ she will be in at $t'$. But we are eschewing the SU semantics. We aim to model the Practical and Epistemic problems on the assumption that this sort of uncertainty is incoherent. So we will not say that, even conditional on the assumption that Alice knows which branch she is now in, she is uncertain about which branch she will be in in the future. In our language, rather (that is, on the OD semantics), if Alice knew which branch she is on now, she would have no uncertainty about the future either.

We will need to regiment our talk of possible worlds, centered worlds and credence functions; so, let us formalize all this.

- We start from a *set of histories*,

$$\mathcal{H},$$

  taken as primitive. (Heuristically: a history $h \in \mathcal{H}$ is a complete future-directed path through a branching tree. We can represent this by putting history labels $h_1, h_2$ etc. at the 'future end' of each branch; see figure 1. In the example depicted,

$$\mathcal{H} = \{h_1, h_2, h_3, h_4^1, h_4^2, h_4^3, h_4^4, h_5^1, h_5^2, h_5^3\}.)$$

- Admit the following (symmetric, reflexive, transitive) binary relation on $\mathcal{H}$ (also taken as primitive): the 'same world as' relation,

$$SW \subseteq \mathcal{H} \times \mathcal{H}.$$

  (In our example, this relation is generated by taking

$$SW(h_4^1, h_4^2), SW(h_4^1, h_4^3), SW(h_4^1, h_4^4);$$

$$SW(h_5^1, h_5^2), SW(h_5^1, h_5^3),$$

  and closing under symmetry, reflexivity and transitivity.)

- Similarly, for any time $t$, let there be a 'same branch at $t$' equivalence relation on $\mathcal{H}$ (also an equivalence relation).

  (In our example, for time $t_1$, $SW_{t_1}$ is generated by:

  $$SB_{t_1}(h_4^1, h_4^2), SB_{t_1}(h_4^3, h_4^4), SB_{t_1}(h_5^1, h_5^2).)$$

- Place the following constraint on each 'same branch at $t$' relation:

  $$\forall h_1, h_2 \in \mathcal{H}, \forall t, SB_t(h_1, h_2) \text{ only if } SW(h_1, h_2). \tag{3}$$

  (Two histories cannot be in the same branch at any time $t$ unless they are in the same world.)

- Define the set $\mathcal{W}$ of *possible worlds* and, for each time $t$, the set $\mathcal{W}_t^C$ of *centered worlds at $t$*, as follows:

  $$\mathcal{W} := \mathcal{H}/SW; \tag{4}$$
  $$\mathcal{W}_t^C := \mathcal{H}/SB_t. \tag{5}$$

  (The set of possible worlds, $\mathcal{W}$, is formed by taking the set of histories, and quotienting by the 'same world as' relation. The set of centered worlds at $t$, $\mathcal{W}_t^C$, is formed by taking the set of histories, and quotienting by the 'same branch at $t$' relation.)

  In our example:

  $$\mathcal{W} = \left\{ \{h_1\}, \{h_2\}, \{h_3\}, \{h_4^1, h_4^2, h_4^3, h_4^4\}, \{h_5^1, h_5^2, h_5^3\}, \right\};$$
  $$\mathcal{W}_{t_1}^C = \left\{ \{h_1\}, \{h_2\}, \{h_3\}, \{h_4^1, h_4^2\}, \{h_4^3, h_4^4\}, \{h_5^1, h_5^2\}, \{h_5^3\} \right\}.$$

**'Is in the future of' relation.** For times $t, t'$ with $t' > t$, say that a centered world $w' \in \mathcal{W}_{t'}^C$ is in the *future of* a centered world $w \in \mathcal{W}_t^C$ iff, regarded as a set of histories, $w'$ is a subset of $w$.[9]

**Credence function.** The agent's epistemic state at time $t$ is represented by a probability function $Cr_t$ on $\mathcal{W}_t^C$. (The agent is uncertain about which possible world is actual, and, conditional on the assumption that she is in some particular branching world, she is uncertain about which branch within that world she is in.)

---

[9]One may want to amend this definition if one is considering branching universes in which branches may recombine, as well as split. I will not consider such universes, since they do not (in practice) fit the Everettian case, and they introduce complications.

## 2.2 Maximization of expected utility

Next, let us sketch our integrated decision theory. We proceed in two stages: first (in the present section), we make a maximization-of-expected-utility claim; second (in section 2.3), we place quantitative constraints on the probability measure with respect to which the expectation value (for the purposes of maximizing *expected* utility) is to be taken.

As in the decision theory familiar from Savage, we have the following notions of a set of *consequences* and a *utility function*:

**Consequences.** There is a set $\mathcal{C}$ of consequences, representing the locus of value to the agent.

**Utility function.** The agent has a utility function $U : \mathcal{C} \to \mathcal{R}$ on the set of consequences, representing the degree to which the agent values each consequence.

At this point, we must make a crucial decision regarding how the Savage framework is to be applied, when some of the epistemically possible worlds are branching multiverses. The key question is what, in our framework of centered worlds and so forth, we should identify with Savage's 'States'. There seem to be two options:

1. For a decision taken at time $t$, take the set of States to be the set $\mathcal{W}_t^C$ of centered worlds at time $t$.

2. Take the set of States to be the set $\mathcal{H}$ of histories.

When there are no branching multiverses on the table, these two options coincide; so, our existing usage of decision theory fails to commit us to one or the other. Option (1) is the generalization that follows if we insist on hanging on to Savage's idea that the set of States must be the locus of *uncertainty*. Option (2) is the generalization that follows if we insist on hanging on to the idea that the set of consequences must be straightforward things like 'I get an ice-cream' or 'There is nuclear war', rather than things like 'A branching multiverse with branches with such-and-such weights is realized to my future; on branch $i$ I get an ice-cream; on branch $j$ there is a nuclear war; ...'. What to do?

It is absolutely crucial to the Everettian program that option (2), not option (1), is the right extension. (2) is the option that leads to the strictly stronger decision theory, according to which branch weights can function,

for all practical purposes, just like chances. If only option (1) is defensible, the Everett interpretation is sunk.

Why think that option (2) *is* the right extension? Indeed, why think that it has anything going for it at all? Surely (a skeptic might rhetorically ask), the idea that the probability measure should be quantifying *uncertainty* is far more fundamental than the idea that the idea that the Consequences should have short descriptions in language familiar to the man on the street?

This challenge certainly has some force, and a full defense of option (2) lies well outside the scope of this paper. Here I will only make the following brief remarks towards its defense. Suppose one has already accepted that, in a case of uncertainty in which no branching worlds are under consideration, Savage was right (i.e. one should maximize expected utility with respect to some $p$, $U$, where $p$ is a measure over the set of possible [centered] worlds). This much is relatively uncontentious. Suppose further that one has also accepted that, in a hypothetical case in which the agent was certain that she lives in an Everettian multiverse, decision theory should be applied in the way Deutsch and Wallace suggest (i.e. $p$ should be a measure over possible *future branches*, not over possible worlds). That this is true given the SU semantics has been persuasively argued by Wallace; for defense of the claim that it remains true even on the OD semantics, see (Greaves, 2004) and (Wallace, 2005c).[10] If those two suppositions are granted, then it seems rather implausible that, in the general case, one should revert to insisting that the probability measure be over present centered worlds, rather than over histories or over future centered worlds. It seems far more plausible that the same sorts of considerations that defend Savage in the non-branching case, and Deutsch and Wallace in the Everettian case, will combine to defend option (2) in the general case. In this sense, the framework I have proposed above strikes me as an innocuous extension of claims that have been defended elsewhere in the literature.

This is (to say the least) hardly a knock-down argument in defense of the decision theory I plan to use. Perhaps the reader is unconvinced, either by my claim that the extension of option (2) from the pure branching case to the general case is innocuous, or by the claim that Deutsch and Wallace are justified in using (2) in the pure branching case. If so, let us put that

---

[10]The basic idea behind the latter claim is: conditional on the assumption that she will have multiple successors in parallel future branches, a decision-maker will care about *all* her futures; but sometimes the interests of her various future selves will conflict; the decision-maker must therefore weigh up the interests of her various successors against one another; the way to achieve this is to maximize expected utility with respect to a normalized measure over parallel coexisting branches.

dispute aside for the time being, since it is a dispute about whether or not the *Practical* Problem can be solved. In this paper, our project is to examine the *conditional* claim: *if* the Practical Problem can be solved along the lines proposed by the rationality-based program, *then* a solution to the Epistemic Problem follows. We will therefore take for granted the claim that option (2) is the right way to build a decision theory for the general case — at least for the sake of argument, at least for the time being — and investigate what follows from it.

On with building our general-case decision theory, then:

**States.**  The set of states is the set of histories, $\mathcal{H}$.

**Acts.**  An *act* is represented by a function $a : \mathcal{H} \to \mathcal{C}$. Since realistic acts will involve receiving a reward at some particular time $t$, one can usually represent an act (instead) as a function from the set of *centered worlds at some particular (future) time $t$* to $\mathcal{C}$: $a : \mathcal{W}_t^C \to \mathcal{C}$.

**Expected utility.**  The *expected utility* of an act $a : \mathcal{H} \to \mathcal{C}$ with respect to a probability measure $p$ over $\mathcal{H}$ is given by:

$$EU_p(a) = \sum_{h \in \mathcal{H}} p(h) \cdot U(a(h)). \tag{6}$$

With these definitions in hand, we are ready to state our core normative claim:

**Maximization of expected utility; general case (MEUG).**  For any rational agent and any time $t$, there exists a probability measure $\overline{Cr}_t$ on $\mathcal{H}$ and a utility function $U : \mathcal{C} \to R$ such that, for any two actions $a_1, a_2$, the agent weakly prefers (at $t$) $a_1$ to $a_2$ iff the expected utility of $a_1$, calculated with respect to $\overline{Cr}_t$, is at least as high as that of $a_2$:

$$a_1 \succ_t a_2 \text{ iff } EU_{\overline{Cr}_t}(a_1) \geq EU_{\overline{Cr}_t}(a_2). \tag{7}$$

Say that $\overline{Cr}_t$ is the agent's *quasi-credence function at time $t$*. (The relationship of $\overline{Cr}_t$ (a normalized measure on $\mathcal{H}$) to the agent's genuine *credence* function at $t$, $Cr_t$ (a normalized measure on $\mathcal{W}_t^C$), will be discussed below.)

14

## 2.3 Quantitative constraints on the quasi-credence function

(**MEUG**) states only that there must be *some* probability measure ('quasi-credence function') over $\mathcal{H}$ with respect to which the agent maximizes expected utility. We wish to impose some stronger constraints: the generalizations of our 'Principal Principles' (**PC**) and (**PE**) to the general case.

**(PC): Principal Principle (chance).** If $E$ is any admissible[11] proposition that entails that there is a chance $x$ that $A$ occurs with branch weight 1, and if the agent has no inadmissible information at $t$, then

$$\overline{Cr}_t(A|E) = x. \tag{8}$$

**(PW): Principal Principle (weights).** If $E$ is any proposition that entails that $A$ occurs with branch weight $x$, then

$$\overline{Cr}_t(A|E) = x. \tag{9}$$

From these two principles, one can derive the following two (heuristically useful) special cases.

- Let $T_{nb} \wedge S$ be the proposition that a particular non-branching, indeterministic, chance theory $T_{nb}$ is true and that the present state of the world is $S$. Let $Ch_{T_{nb},S}(A)$ be the chance that theory $T_{nb}$ assigns to event $A$, given initial conditions $S$. Then, for any proposition $A$,

$$\overline{Cr}_t(A|T_{nb} \wedge S) = Ch_{T_{nb},S}(A). \tag{10}$$

- Let $T_b \wedge S$ be the proposition that a particular weighted branching theory $T_b$ is true and that the present state of the agent's branch is $S$. Let $W_{T_b,S}(A)$ be the branch weight that theory $T_b$ assigns to event $A$, given initial conditions $S$. Then, for any proposition $A$,

$$\overline{Cr}_t(A|T_b \wedge S) = W_{T_b,S}(A). \tag{11}$$

Next, let us define 'conditional credence function' and 'caring measure' as follows:

---

[11]I ignore a number of subtleties concerning the notion of admissibility, since they are orthogonal to the concerns of the present paper. The reference to admissibility can be done away with at the expense of (arguably) making the Principal Principle less user-friendly: see, e.g., (Hall, 2004).

- The agent's *credence function at t conditional on $T_{nb} \wedge S$* is defined by

$$Cr_t(\cdot|T_{nb} \wedge S) \equiv \overline{Cr}_t(\cdot|T_{nb} \wedge S).$$

- The agent's *caring measure at t given $T_b \wedge S$* is defined by

$$Car_{t,T_b,S}(\cdot) := \overline{Cr}_t(\cdot|T_b \wedge S).$$

Then the above normative constraints have the following implications: a rational agent's credence function conditional on $T_{nb} \wedge S$, $Cr_t(\cdot|T_{nb} \wedge S) \equiv \overline{Cr}_t(\cdot|T_{nb} \wedge S)$, must be equal to the chance function given by the theory $T_{nb}$ for possible histories consistent with present state $S$. A rational agent's caring measure given $T_b \wedge S$, $Car_{t,T_b,S}$, is equal to the weight function given by the theory $T_b$ for branches in the future of a branch in state $S$. (In subsequent sections, I will drop explicit mention of the state $S$, for convenience.)

By way of illustration: note that these constraints ensure, in particular, that in an idealized case in which the agent is certain that the actual world is quantum-mechanical and that the initial state of a system about to be measured is $|\psi\rangle$, she would accept and reject bets as if she had credence $|\langle\phi_i|\psi\rangle|^2$ that outcome $|\phi_i\rangle$ would occur, *regardless* of her views as to whether it is a stochastic or a branching version of QM that is true. Thus, we have captured, in our more general framework, Deutsch's claim that 'the rational agent [who thinks that QME is true] makes decisions as if [the Born rule] were true' (Deutsch, 1999).

## 3   A Bayesian account of empirical confirmation

Suppose that the decision theory sketched in section 2 is defensible; then the Practical Problem is solved. What now of the Epistemic Problem?

The Epistemic Problem (recall) is the threat that, on the Everett interpretation, the sorts of empirical data that we ordinarily take to be evidence for quantum mechanics can no longer rationally be regarded as evidence in favor of the theory.

Let us be more precise about this threat. A large part of our empirical evidence for quantum mechanics takes the following form:

**Matching:** In experiments that have been repeated many times, observed relative frequencies have roughly matched the single-case chances predicted by quantum mechanics for experiments of that type.

I take the core Epistemic Problem to be the task of explaining why Matching still counts as evidence for quantum mechanics when quantum mechanics is understood as a branching-universe theory, rather than a stochastic theory.[12] The *prima facie* problem is that whereas, in the stochastic case, we take Matching to confirm quantum mechanics because stochastic quantum mechanics (QMS) predicted that Matching would occur *with high chance*, Everettian quantum mechanics (QME) ('merely'?) says that Matching will occur with high branch weight. Meanwhile, if Everettian QM is true, there (certainly!) also exist branches on which Matching does *not* occur. To be sure, those other branches have low weight — but, the worry continues, what of that? They are still *real*. Our problem is that it is unclear whether, or why, observing high-*weight* outcomes should count as confirming a theory.

**Adequacy of a Bayesian approach.** Before attempting to solve this problem, it is worth reminding ourselves of how empirical data is supposed to support belief in the theory in a non-branching case.

Bayesian epistemology offers a systematic account of belief revision. This account is subject to a variety of criticisms, mainly that all the interesting epistemological questions are shifted to the question of which priors are rational. Be this as it may, there is at least one question that *is* well handled by the Bayesian framework: the question of the relevance of relative frequency data to theory confirmation. I therefore take it that, if the Bayesian account yields the result that Matching confirms Everettian quantum mechanics just as much as it confirms stochastic quantum mechanics, then the Epistemic Problem will be solved. My project is therefore to show that the Bayesian account, when properly applied to the Everettian case, does indeed yield that result.

## 3.1   The Bayesian framework

Our first task is to review the standard Bayesian framework, and its explanation of the epistemic relevance of Matching in the non-branching case.

The following four features are key to the Bayesian account of this case:

- *Probabilism* holds. That is, an ideally rational epistemic agent has, at any given time, a well-defined probability function on the space of

---

[12]Note: the Epistemic Problem has nothing to do with the issue of gaining empirical evidence that can preferentially confirm QME *relative to other versions of quantum mechanics.*

possible worlds, representing the agent's degrees of belief (so, we call the probability function a 'credence function').

- *Empirical evidence* comes in the form of subsets of the set of all possible worlds (the interpretation being that the agent has just learnt that the actual world is a member of the subset in question).

- On receiving a given piece of empirical evidence, the agent updates her beliefs by *conditionalizing* on that evidence.

- *Empirical confirmation* of a theory consists in the increase of degree of belief in that theory as a result of conditionalizing on evidence.

To explain why Matching constitutes empirical evidence for quantum mechanics according to the Bayesian account, it suffices to consider the following simple scenario. We begin with an agent, Alice, at time $t_0$. Alice is entertaining two possible **theories** of the world. One is our non-branching, stochastic version of quantum mechanics (QMS), according to which measurement outcomes are uncontroversially assigned genuine non-trivial *chances*. The other is an alternative, non-quantum theory that I will simply label $T$.[13] So, we have:

**Alice's candidate theories.** The set of theories that Alice is considering is given by $\mathcal{T} \equiv \{T, \text{QMS}\}$.

Alice is about to perform an experiment that has $n$ possible outcomes, $\{E_i\} : i = 1, \ldots, n$:

**Possible outcomes of the experiment.** The set of possible outcomes is given by $\mathcal{O} \equiv \{E_1, \ldots, E_n\}$.

Let there be one outcome, $E_1$ say, that has the following feature: the chance assigned to $E_1$ by QMS is significantly higher than Alice's credence, conditional on $T$, that $E_1$ will be the outcome that actually occurs.

Note that, if the experiment in question happens to be a relative-frequency experiment, then $E_1$ could be the proposition that the observed relative frequency approximately matches the single-case chances predicted by QMS

---

[13]$T$ is whatever Alice thinks is true, conditional on the assumption that QMS is false. It is a 'theory' only in some rather loose sense of the term; it presumably resembles some long disjunction of non-quantum theories more closely than it resembles a single physical theory such as quantum mechanics. $T$ is best thought of as a heterogeneous collection of possible worlds, over which the agent has some unspecified but (in the limit of idealization) well-defined conditional credence function.

(that is, the proposition that Alice's experiment yields an instance of Matching), and $E_2$ just the negation of $E_1$. But this will not be essential to our argument.

Next, Alice will perform her experiment. There are now $2n$ possibilities that our model must treat as distinct if it is to handle the issue of theory confirmation. Either QMS is true, or $T$ is true. And, within each of these two cases, there are $n$ further possibilities: the outcome might be any one of the possibilities $\{E_1, \ldots, E_n\}$. To make the decision theory more manageable, let us coarse-grain our history set so that each of these possibilities corresponds to just one 'history'[14], and, with slight abuse of terminology, let us call our coarse-grained histories 'possible worlds':

**Possible worlds.** The set $\mathcal{W}$ of possible worlds is given by

$$\mathcal{W} \equiv \{T \wedge E_1, \ldots, T \wedge E_n, \mathrm{QMS} \wedge \mathrm{E_1}, \ldots, \mathrm{QMS} \wedge \mathrm{E_n}\}.$$

**Alice's credence function at time $t_0$.** Since probabilism holds, Alice has, at time $t_0$, some well-defined credence function $Cr_{t_0}$ over the set $\mathcal{W}$ of possible worlds. Since Alice at time $t_0$ respects the Principal Principle with respect to the theory QMS (and we assume that she has no 'inadmissible information'), her conditional degrees of belief $Cr_{t_0}(\cdot|\mathrm{QMS})$ equal the chances given by the theory[15]:

$$Cr_{t_0}(E_i|\mathrm{QMS}) = \mathrm{Ch_{QMS}(E_i)}. \tag{12}$$

**Person-stages.** It will be useful to talk in terms of person-stages. We will usually have no need to distinguish between the person-stages in the various possible worlds at $t_0$. So, I will often write $Alice_0$ for the person-stage that represents Alice at time $t_0$, regardless of which world is supposed real. When discussing what goes on at $t_1$, we *will* sometimes need to distinguish between the various person-stages that may exist at that time. So, I will write $\mathrm{Alice_1^{QMS,E_i}}$ for the person-stage in possible world $\mathrm{QMS} \wedge \mathrm{E_i}$ at time $t_1$, and similarly for $t_2$. (See figure 2.)

---

[14] This is a standard move. One typically works with 'small-world' decision problems, not with the 'grand-world' decision problem in which the set of states is given by the fully fine-grained set of histories, or 'possible worlds' in the usual philosophers' sense.

[15] This is required in order to ensure that observing high-*chance* outcomes confirms the theory for our agent, rather than observing outcomes that just happened to be given high conditional credence $Cr_{t_0}(\cdot|\mathrm{QMS})$ by our agent's initial credence function.
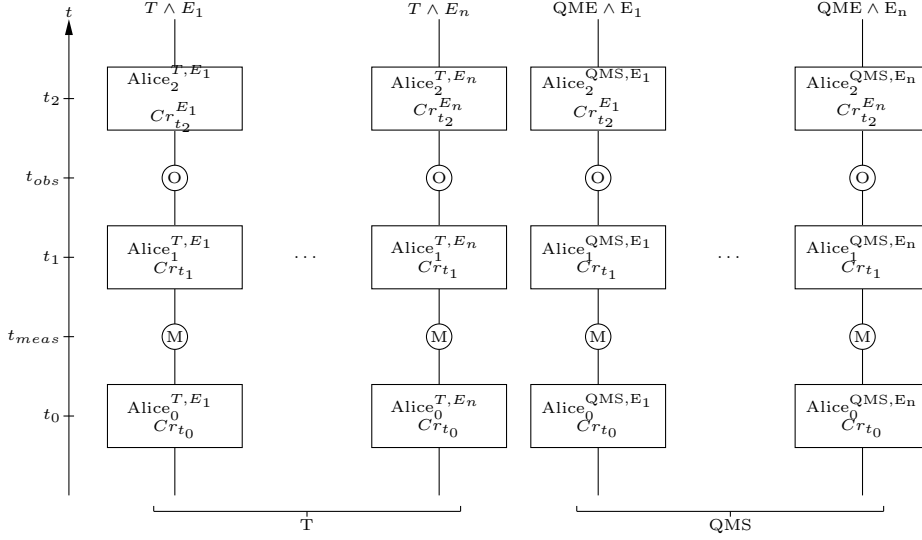
Figure 2: A diagrammatic representation of the Bayesian model of confirmation, when Alice is entertaining theories QMS (stochastic quantum mechanics) and $T$ (her 'default theory') only. Each of the $2n$ possible worlds in our model is represented by a timeline running up the page (although only four have been drawn here, for simplicity); the three person-stages that exist in each possible world at the three times $t_0, t_1, t_2$ we are considering, plus the measurement event and the observation event (Alice's looking at the apparatus), are represented by nodes on the corresponding timeline. The nodes representing each person-stage contain two pieces of information: the name we are giving to that person-stage (e.g. $\text{Alice}_0^{T,E_1}$ for the stage of Alice in possible world $T \wedge E_1$ that exists at time $t_0$), and the epistemic state of that person-stage (represented by a probability function on $\mathcal{W}_t^C$, e.g. $Cr_{t_0}$). At time $t_0$, Alice is in the epistemic state $Cr_{t_0}$, regardless of which theory is true, and of which outcome is going to occur. Similarly, at time $t_1$, Alice is in epistemic state $Cr_{t_1}$ in all possible worlds. At time $t_2$, Alice is either one of the epistemic states $Cr_{t_2}^{E_i}$ (depending on which outcome $E_i$ has occurred in her possible world).

20

**Credence function $Cr_{t_1}$: after the measurement has taken place, but before the result(s) has/have been observed.** At time $t_{meas}$, the measurement occurs; by $t_1$ it is complete. Since the result has not yet been observed, Alice has received no information between $t_0$ and $t_1$; we *must* therefore insist that all (possible) $t_1$-person-stages are in the same epistemic state *as one another*. So, the epistemic situation at time $t_1$ can be fully described by a single credence function $Cr_{t_1}$. Further, and again because Alice has received no information between $t_0$ and $t_1$, we *should* insist that her epistemic state has not changed during this time interval: $Cr_{t_1} = Cr_{t_0}$. That is, we should insist that Alice satisfies a reflection condition. (Since $Cr_{t_0}$ and $Cr_{t_1}$ are thus assumed equal, the distinction between the two plays no interesting role in the present account. It is mentioned here only to facilitate later comparison to the general case.)

**Credence functions $\{Cr_{t_2}^i\}$: after the result(s) has/have been observed.** A little while later (at time $t_{obs}$), Alice gets curious, and decides actually to look at the measuring apparatus. In each possible world, Alice sees some particular outcome $E_i$. Depending on which outcome she sees, she may change her epistemic state. Since it is possible for her to move to different epistemic states depending on which outcome she sees, we will in general need $n$ credence functions to describe the epistemic situation at time $t_2$: we call these $\{Cr_{t_2}^{E_i}\} : i = 1, \ldots, n$. Again, these credence functions are probability functions on the set $\mathcal{W}$ of possible worlds.

**Rational belief-updating.** Rationality places constraints on how the posterior credence functions $\{Cr_{t_2}^{E_i}\}$ may be related to $Cr_{t_1}$. According to the Bayesian account, rational belief-updating is by conditionalization:

$$Cr_{t_2}^{E_i}(\cdot) = Cr_{t_1}(\cdot|E_i), \tag{13}$$

for each possible outcome $E_i$.

The term 'conditionalization' will be in danger of ambiguity later in our discussion. We will therefore refer to the updating policy (13) *for cases in which no branching-universe theories are under consideration* as **Minimal Conditionalization**. (So: by stipulation, Minimal Conditionalization just does not apply to cases in which one or more branching-universe theories *is* under consideration, and we have as yet said nothing about how rational belief-updating proceeds in such a 'general case'.)

In particular, if Alice updates by Minimal Conditionalization, then her possible credences at time $t_2$ in her two candidate theories are given by

$$
\begin{aligned}
Cr_{t_2}^{E_i}(\text{QMS}) &= Cr_{t_1}(\text{QMS}|\text{E}_i); \\
Cr_{t_2}^{E_i}(T) &= Cr_{t_1}(T|E_i).
\end{aligned}
$$

It is easy to see that, given this updating rule, observing $E_1$ (recall: by stipulation, $E_1$ is an outcome with the feature that the chance assigned to $E_1$ by QMS is greater than Alice's prior credence, conditional on $T$, that $E_1$ will occur) will lead Alice to increase her credence in QMS at the expense of her credence in T. Meanwhile, observing $E_2$ will lead Alice to *decrease* her credence in QMS, increasing her credence in $T$ to compensate:

$$
\begin{aligned}
Cr_{t_2}^{E_1}(\text{QMS}) &> Cr_{t_1}(\text{QMS}); \\
Cr_{t_2}^{E_1}(T) &< Cr_{t_1}(T); \\
Cr_{t_2}^{E_2}(\text{QMS}) &< Cr_{t_1}(\text{QMS}); \\
Cr_{t_2}^{E_2}(T) &> Cr_{t_1}(T).
\end{aligned}
$$

That is, according to the Bayesian account, observing $E_1$ empirically confirms QMS (and disconfirms $T$), while observing $E_2$ empirically disconfirms QMS (and confirms $T$).

Recall again that, as a special case, $E_1$ might be the proposition that the relative frequency approximately matches the single-case QMS chances, and $E_2$ the proposition that it does not. The above account therefore gives a Bayesian explanation of why, in the *non*-branching case, Matching counts as empirical evidence for quantum mechanics. It is time now to turn to the general case.

## 3.2 Generalizing the Bayesian framework

The Bayesian framework sketched above must be generalized slightly in order to handle the question of rational belief-updating when one of the theories our agent is considering is a branching-universe theory, such as *Everettian* QM. This is because we now need to be able to handle the co-existence of more than one copy of Alice at a give time in a given possible world. Our next task is to carry out that generalization.

Again, to set up an illustrative toy example, we begin with an agent, Alice, at time $t_0$. Alice is now entertaining *four* possible theories of the world:

- QMS: a stochastic version of quantum mechanics;

- QME: Everettian quantum mechanics;

- $T_{nb}$: Alice's 'default non-branching theory', i.e. whatever Alice thinks is true if QMS is false but the actual world is non-branching;

- $T_b$: Alice's 'default branching theory', i.e. whatever Alice thinks is true if QME is false but the actual world is a branching multiverse.

That is, we now have:

**Alice's candidate theories.** The set $\mathcal{T}$ of theories that Alice is considering is given by
$$\mathcal{T} \equiv \{T_{nb}, \mathrm{QMS}, \mathrm{T_b}, \mathrm{QME}\}.$$

Alice performs the same experiment as before, with the same set of possible outcomes,
$$\mathcal{O} = \{E_1, \ldots, E_n\}.$$
Her candidate theories say the following things about these outcomes:

- QMS (as before) provides a normalized measure ('chance function') $Ch_{\mathrm{QMS}}$ over $\mathcal{O}$; we have, in particular,

$$Ch_{\mathrm{QMS}}(E_1) > Cr_{t_0}(E_1|\mathrm{QMS} \vee \mathrm{T_{nb}});$$

$$Ch_{\mathrm{QMS}}(E_2) < Cr_{t_0}(E_2|\mathrm{QMS} \vee \mathrm{T_{nb}}).$$

- According to each of Alice's branching-universe theories ($T_b$, QME), there will be $n$ branches at times later than $t_{meas}$: for each $E_i$, there will be some branch on which $E_i$ occurs. The theory provides a normalized measure ('weight function') $W_{\mathrm{QME}}$ over $\mathcal{O}$; we stipulate that

$$W_{\mathrm{QME}}(E_1) > Car_{t_0,T_b}(E_1);$$

$$W_{\mathrm{QME}}(E_2) > Car_{t_0,T_b}(E_2).$$

Again, we simplify the model by coarse-graining our history space, so that we have $4n$ histories: $T_{nb} \wedge E_i, \mathrm{QMS} \wedge \mathrm{E_i}, \mathrm{T_b} \wedge \mathrm{E_i}, \mathrm{T_b} \wedge \mathrm{E_i}$ for each $i \in \{1, \ldots, n\}$.[16] Given that all (coarse-grained) histories in which QME is

---

[16]There is a further simplification: we have not bothered to represent parallel branches that are in the same world as $Alice_0$, but that she is certain she is not on. Such branches are irrelevant to her decision-making and belief-updating. This move is structurally analogous to our 'failure' to discuss non-branching worlds that are allowed by (say) the non-branching theory QMS, but that the agent is certain are not actual.

true are then in the same (coarse-grained) possible world, and similarly for $T_b$, there are $2n + 2$ (coarse-grained) possible worlds to consider:

**Possible worlds.** The set $\mathcal{W}$ of possible worlds is given by

$$\mathcal{W} \equiv \{T_{nb} \wedge E_i, \ldots, T_{nb} \wedge E_n, \mathrm{QMS} \wedge \mathrm{E_1}, \ldots, \mathrm{QMS} \wedge \mathrm{E_n}, \mathrm{T_b}, \mathrm{QME}\}.$$

**Person-stages.** Again, it will be useful to talk in terms of person-stages. The only new feature in the general case is that there are $n$ person-stages in each of the branching possible worlds QME, $\mathrm{T_b}$ at each time later than $t_{meas}$: for instance, at $t_1$ we have $\mathrm{Alice}_1^{\mathrm{QME,E_1}}, \ldots, \mathrm{Alice}_1^{\mathrm{QME,E_n}}$, and $\mathrm{Alice}_1^{T_b,E_1}, \ldots,$ $\mathrm{Alice}_1^{T_b,E_n}$. (See figure 3.)

**Alice's credence function at time $t_0$.** As before, Alice at $t_0$ has a credence function $Cr_{t_0}$ on the set $\mathcal{W}$ of possible worlds. Since Alice respects the Principal Principle, $Cr_{t_0}$ respects the constraints (12) as before.

No assumption is made about the value of $Cr_{t_0}(\mathrm{QME})$, except that it lies strictly between 0 and 1. (That is, we are interested in the epistemic dynamics of agents who are just barely considering QME, as well as those who already have fairly high credence in QME. This is important: the present discussion does *not* treat only of 'Everettians', if by 'Everettian' one means 'agent who has high credence in QME'.[17])

**Caring measures and the Everettian weight function.** In one sense, our description of Alice's epistemic state at $t_0$ is now complete: that epistemic state is simply $Cr_{t_0}$. However, there are two more objects related to

---

[17]This observation is made in reply to Wallace's objection that the Bayesian approach suggested in (Greaves, 2004) involves some sort of illicit presupposition that Everettian quantum mechanics is true:

> Greaves (2004) ... argues that *on the assumption that we live in a branching, Everettian-style universe*, we can construct an analogue of the Bayesian update rule and prove its validity. But it is unclear at best how this strategy can help us where we are concerned with evidence for the Everett interpretation *itself.* (Wallace, 2005a, p.20; first *emphasis* added)

Wallace's suggestion here is that, because of the need for this assumption, the Epistemic Problem cannot be solved without appealing to the SU semantics. My point is that no such assumption is being made. We are considering a rational agent who *has nonzero credence in the proposition that* the world is Everettian one, and asking how such an agent should manage her beliefs. No assumption that the agent *is in fact* in an Everettian multiverse (or that she is not) plays any role in the argument. (This is not to say that Wallace's reading of (Greaves, 2004) is unreasonable; I am grateful for the insistence on clarification.)
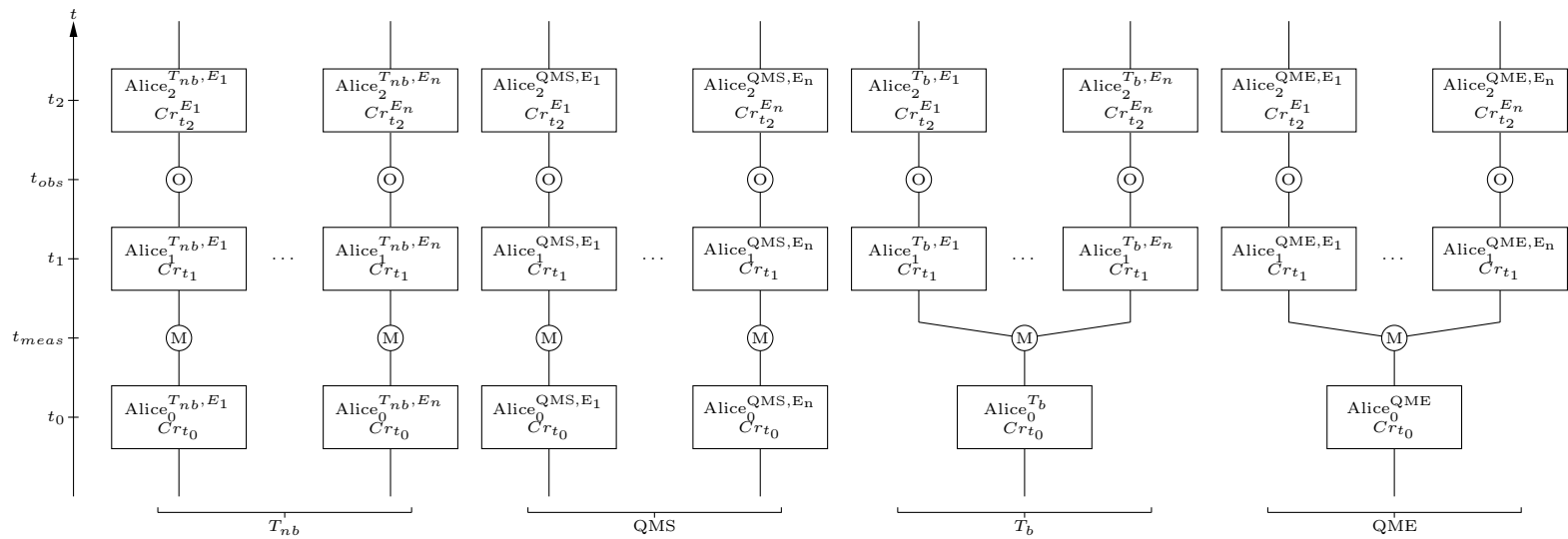
Figure 3: The possible worlds and person-stages relevant to the Bayesian account of updating in a 'general-case' scenario, i.e. a case in which the agent has non-zero credence in branching multiverses as well as in non-branching worlds.

Alice's psychology at $t_0$ that will turn out to be crucial to the *dynamics* of Alice's credences. These are her 'caring measures' for the two branching multiverses under consideration, $Car_{t_0,\text{QME}}$ and $Car_{t_0,T_b}$. Recall that each caring measure is (mathematically) a probability function over the set of histories corresponding to the possible world in question.

Mathematics aside, what do these caring measures *mean*? Well: they can be given operational significance in terms of rational action choice *via* decision theory. In evaluating the decision-theoretic value of a branching future in a world governed by a particular weighted-branching-universe theory, Alice$_0$ will take the average of utility-on-a-branch over all branches that are in her future, averaged according to her caring measure for the theory in question. We are granting (by **(PE)**) that $Car_{t_0,\text{QME}}$ is dictated by the Everettian weight function: $\forall E_i \in \mathcal{O}, Car_{t_0,\text{QME}}(E_i) = W_{\text{QME}}(E_i)$.

The relevance of $Car_{t_0,\text{QME}}$ and $Car_{t_0,T_b}$ to Alice's epistemic dynamics will be drawn out below (section 5).

**The Epistemic Problem.** Our ultimate interest, as far as the Epistemic Problem goes, is in the dymanics of credences *in theories*. That is, we are interested in the relationships between the quadruples that describe the agent's epistemic states at various times and following observation of various outcomes:

$$\text{At } t_0: \quad \langle Cr_{t_0}(T_{nb}), Cr_{t_0}(\text{QMS}), \text{Cr}_{t_0}(\text{T}_b), \text{Cr}_{t_0}(\text{QME})\rangle;$$
$$\text{At } t_2, \text{having observed } E_i: \quad \langle Cr_{t_2}^{E_i}(T_{nb}), Cr_{t_2}^{E_i}(\text{QMS}), \text{Cr}_{t_2}^{E_i}(\text{T}_b), \text{Cr}_{t_2}^{E_i}(\text{QME})\rangle.$$

The Epistemic Problem is the threat that, under rational updating of beliefs in the light of observations $E_i$, the quantities $Cr_{t_2}^{E_i}(\text{QME})$ will take different values than those they would take under Minimal Conditionalization if QME were a stochastic theory and the branch weights were chances.

**Sufficient condition for solving the Epistemic Problem.** The Epistemic Problem will be solved if we can show that, under rational belief-updating, the quantities $Cr_{t_2}^{E_i}(\text{QME})$ *do* evolve from $Cr_{t_1}$ in the same way that the quantities $Cr_{t_2}^{E_i}(\text{QMS})$ do (with branch weights playing the role that chances play for QMS). In that case, we will have explained why observing outcomes that Everettian quantum mechanics predicted would occur *with high weight* (for example, but not necessarily: relative frequencies that match single-case branch weights) counts as evidence in favor of Everettian quantum mechanics, just as observing outcomes that a stochastic quantum

26

mechanics predicted would occur *with high chance* (for example, but not necessarily: relative frequencies that match single-case chances) counts as evidence in favor of stochastic quantum mechanics.

The task of section 5 will be to argue that this condition is indeed met. First, however, it is worth investigating in more detail why one might have a *prima facie* suspicion that the condition might *not* be met. The thought here is that there is a way of extending Minimal Conditionalization to the general case that both (a) is intuitive, and (b) has that disastrous consequence. I will call that intuitive extension Naive Conditionalization. We will see, however, that Naive Conditionalization is implausible for independent reasons.

# 4    Interlude: Naive Conditionalization in the general case

The Epistemic Problem arises because it is not (yet) clear what rationality constraints can be placed on the posterior credence functions $Cr_{t_2}^{E_i}$. The natural move is to find an updating policy that is a natural generalization of Minimal Conditionalization for the general case, defend it as the unique rational updating policy, and note how Alice's credences in the various theories at $t_2$ relate to her credences at $t_0$ according to this updating policy. (Recall that, by stipulation, Minimal Conditionalization itself applies only to cases in which no branching-universe theories are under consideration.) However, there is more than one way to do this. The most naively obvious such policy — 'Naive Conditionalization' — yields thoroughly unsatisfactory results. Let us investigate.

**Naive Conditionalization.**    We want to find some sort of conditionalization that will apply to the present case. The obvious move seems to be:

$$Cr_{t_2}^{E_i}(\cdot) \quad = \quad Cr_{t_0}(\cdot | E_i) \tag{14}$$

$$\equiv \quad \frac{Cr_{t_0}(\cdot \wedge E_i)}{Cr_{t_0}(E_i)}. \tag{15}$$

The problem now is that we have not yet defined the quantities on the right hand side of this equation for the case in which branching theories are among the possibilities: recall that $Cr_{t_0}$ is defined only on $\mathcal{W}_{t_0}^C$, not on $\mathcal{W}_{t_2}^C$. One wants to say, for instance, that $Cr_{t_0}(E_i) = Cr_{t_0}(T_{nb} \wedge E_i) + Cr_{t_0}(\text{QMS} \wedge \text{E}_i) + \text{Cr}_{t_0}(\text{T}_b \wedge \text{E}_i) + \text{Cr}_{t_0}(\text{QME} \wedge \text{E}_i)$, but we have not defined $Cr_{t_0}(\text{QME} \wedge \text{E}_i)$ or $Cr_{t_0}(T_b \wedge E_i)$.

Very well: let us define those quantities. How? Well, the most obvious values to assign to $Cr_{t_0}(\text{QME} \wedge \text{E}_i)$ and $Cr_{t_0}(T_b \wedge E_i)$, if we are to assign values at all, are 'surely' (on the OD semantics!) $Cr_{t_0}(\text{QME} \wedge \text{E}_i) := \text{Cr}_{t_0}(\text{QME})$ and $Cr_{t_0}(T_b \wedge E_i) := Cr_{t_0}(T_b)$, since (on the OD semantics), if either of QME and $T_b$ is true, each $E_i$ will certainly occur (on some branch). That is, according to this proposal, we are to identify $E_i$ (for each $i$) with the set $\{T_{nb} \wedge E_i, \text{QMS} \wedge \text{E}_i, \text{T}_b, \text{QME}\} \subset \mathcal{W}$. This has the consequence that the $\{E_i\}$ are no longer mutually exclusive propositions, but that does not prevent us from conditionalizing on any of them.

As is easily verified, however, applying this operation of Naive Conditionalization to the present case has the unwanted result that a straightforward piece of empirical evidence such as $E_i$ confirms QMS and QME to different degrees. This is not the result that the Everettian wants.

Note well, though: Naive Conditionalization would *not* have the result that the sorts of evidence that we ordinarily take to be confirmatory of quantum mechanics (for example, Matching) fail to count as confirmatory of Everettian quantum mechanics. Rather, as a simple calculation can show, Naive Conditionalization actually has the result that *every* event that is quantum-mechanically possible, but not necessary, confirms QME at the expense of theories that assign probability less than unity to that event (namely, in our example, QMS and $T_{nb}$). (This result is not really surprising: what we have done, in constructing Naive Conditionalization, is to give formal expression to the intuition that disconfirmation of a theory occurs only when something that, according to the theory in question, is *improbable*, occurs; and (at the same time) that, according to QME or $T_b$, nothing is improbable since everything is certain to occur.) This observation is (presumably!) a *reductio* of the suggestion that Naive Conditionalization is the rational updating policy when branching-universe theories are among those under consideration.

The upshot of all this is as follows. Our Epistemic Problem, at the present point in the discussion, takes a somewhat different form from that originally suggested. The problem is *not* that there is some obviously correct updating policy that has the consequence that Matching fails to count as empirical confirmation of QME relative to $T_{nb}$ and $T_b$. Rather, the problem is that we have not yet found *any* adequate account of how empirical confirmation — including confirmation of a non-branching theory, such as QMS — should proceed, once any branching-universe theories are afforded any (nonzero) credence. This is a problem for everyone who wants a coherent account of empirical confirmation for the general case, not only for those who want the Everett interpretation to turn out to be acceptable.

I conclude that, whatever *prima facie* intuitive support Naive Conditionalization has, the claim that it captures epistemic rationality for the general case is untenable. We must look elsewhere. In section 5, we will find independent arguments that not only demonstrate (again) that Naive Conditionalization is irrational, but that also show which updating policy is rational.

# 5 Solving the Epistemic Problem without SU

## 5.1 The confirmation question is a cognitive decision problem

To find the right updating policy for the general case, and thence to address the Epistemic Problem, let us first take a step back.

In a transition from non-Everettian to Everettian quantum mechanics, many things change. Among the things that change, perhaps, are the status of sentences like 'the outcome was spin-up'; whereas such sentences can, on a non-branching interpretation, be taken to pick out propositions in the sense of sets of possible worlds, on the Everett interpretation they must, at least if we eschew the SU semantics, be taken to pick out self-locating propositions (sets of possible centered worlds). This observation is closely related to the thought that confirmation becomes more problematic in the Everettian case.

However this may be, not *everything* changes in the move to a branching-universe interpretation. The key observation here is a close cousin of Deutsch's insight, in the case of the Practical Problem, that the rational agent still has to make decisions, and that therefore we have to say something about the boundary between rational and irrational practical decision in a branching multiverse. In particular, the fact remains that we have person-stages who

1. hold well-defined credence functions over $\mathcal{W}_{t_0}^C$ or $\mathcal{W}_{t_2}^C$;

2. observe experimental outcomes $E_i$, and

3. have to update their credences (trivially or otherwise) in reaction to their observations.

We must therefore say something about the location of the boundary between rational and irrational updating policies in such a case.

**Alice's epistemic decision problem.**  Alice's problem is an epistemic decision problem: she has to decide (and she might as well decide at time $t_0$), for each experimental outcome, which credence function she will move to if/when/where she observes that outcome. That is, she has to choose $n$ preferred probability functions, $\{Cr_{t_2}^{E_i}\}$, defined on $\mathcal{W}_{t_2}^C$. That is, she has to choose a preferred map from $\mathcal{O}$ (the set of possible outcomes) to the set $\mathcal{P}(\mathcal{W}_{t_2}^C)$ of probability functions on $\mathcal{W}_{t_2}^C$, representing her updating policy. Let $f_{pref}$ denote the map that Alice chooses. (This 'choice' may not be a voluntary one, but what of that?)

It is crucial to realize that *this is all that Alice can do.* She cannot, for instance, decide to adopt a different credence distribution on observing a given outcome $E_i$ if QMS is true than on Everettian branches where $E_i$ is observed if QME is true. The possible world (which may or may not be actual) in which QMS is true and $E_i$ is observed on the one hand, and the Everettian branch (which may or may not exist) on which Alice observes $E_i$ on the other hand, are epistemically indistinguishable environments for Alice; it is not in her power to decide to react differently to those two situations. (Would that science were that easy!)

**How to solve the Epistemic Problem.**  The Epistemic Problem will be solved if we can justify the claim that one particular function — the one according to which belief-updating treats Everettian branch weights in the same way that Minimal Conditionalization treats the chances of a stochastic theory — is the uniquely rational updating policy. Our next task is to construct this privileged function, using only resources that are available to the Everettian according to the OD semantics. (Only to *construct* it; we will *justify* the claim that this function represents the uniquely rational choice later.) This construction proceeds in three steps.

1. First, we need to identify the set $\mathcal{W}_{t_2}^C$ of *possible centered worlds.* In our toy example, this is given by

$$
\begin{aligned}
\mathcal{W}_{t_2}^C \;=\; \{ & T_{nb} \wedge E_1, \ldots, T_{nb} \wedge E_n, \\
& \mathrm{QMS} \wedge \mathrm{E_1}, \ldots, \mathrm{QMS} \wedge \mathrm{E_n}, \\
& T_b \wedge E_1, \ldots, T_b \wedge E_n, \\
& \mathrm{QME} \wedge \mathrm{E_1}, \ldots, \mathrm{QME} \wedge \mathrm{E_n} \}.
\end{aligned}
$$

2. Next, let the probability measure $\overline{Cr}_{t_0} : \mathcal{W}_{t_2}^C \longrightarrow [0,1]$ be defined as follows.

$$
\overline{Cr}_{t_0}(T_{nb} \wedge E_i) \;=\; Cr_{t_0}(T_{nb}) \times Cr_{t_0}(E_i | T_{nb});
$$

30

$$\begin{aligned}
\overline{Cr}_{t_0}(\mathrm{QMS} \wedge \mathrm{E_i}) &= Cr_{t_0}(\mathrm{QMS}) \times \mathrm{Cr}_{t_0}(\mathrm{E_i}|\mathrm{QMS}) \\
&= Cr_{t_0}(\mathrm{QMS}) \times \mathrm{Ch}_{\mathrm{QMS}}(\mathrm{E_i}); \\
\overline{Cr}_{t_0}(T_b \wedge E_i) &= Cr_{t_0}(T_b) \times Car_{t_0,T_b}(E_i); \\
\overline{Cr}_{t_0}(\mathrm{QME} \wedge \mathrm{E_i}) &= Cr_{t_0}(\mathrm{QME}) \times \mathrm{Car}_{t_0,\mathrm{QME}}(\mathrm{E_i}) \\
&= Cr_{t_0}(\mathrm{QME}) \times \mathrm{W}_{\mathrm{QME}}(\mathrm{E_i}).
\end{aligned}$$

(The notation $\overline{Cr}_{t_0}$ is chosen because this probability measure is the quasi-credence function of Alice at time $t_0$, in the terminology of section 2.)

3. We now have the resources to define the map $f_{EC} : \mathcal{O} \to \mathcal{P}(\mathcal{W}_{t_2}^C)$ as follows.

$$f_{EC}(E_i) = \overline{Cr}_{t_0}(\cdot|E_i). \tag{16}$$

$f_{EC}$ is an updating rule in the required sense. In honor of its formal similarity to Minimal Conditionalization, I will refer to $f_{EC}$ as Extended Conditionalization.

So far, we have made no philosophically contentious claims; we have just performed a mathematical construction. (It is a straightforward mathematical fact that $\overline{Cr}_{t_0}$ and the $\{f_{EC}(E_i) : E_i \in \mathcal{O}\}$ satisfy the probability axioms.) The claim that *is* philosophically contentious, but (I claim) correct nonetheless, is the claim that $f_{EC}$, i.e. Extended Conditionalization, is the uniquely rational updating rule: $f_{pref} = f_{EC}$. This is the claim that, if true, solves the Epistemic Problem: if Alice updates her beliefs according to Extended Conditionalization, then the 'quasi-credence function' $\overline{Cr}_{t_0}$ will be behaving like a credence function with respect to belief-updating, as well as with respect to practical decision. This, of course, will have the consequence that a given outcome $E_i$ will count as evidence for (or against) Everettian quantum mechanics in just the same way that $E_i$ counts as evidence for (or against) stochastic quantum mechanics. The next task is to justify this contentious claim.

## 5.2 Justifying conditionalization

We noted that Extended Conditionalization is the formal analog of Minimal Conditionalization. Now, Minimal Conditionalization, in the absence of branching-universe theories, strikes most people as the *obviously* rational way to update credences. Minimal Conditionalization simply amounts to

setting one's credence to zero on the proposition whose falsity one has just learnt, and renormalizing. What else is one supposed to do?

If this appeal to intuitive obviousness were the only justification of conditionalization, then we would have a bit of a problem. Whatever the virtues of Everettian QM, immediate widespread agreement about what is intuitively obvious when branching universes are under consideration is unfortunately not among its privileges.

Fortunately, however, we are not stuck here. The Bayesian literature also contains several *arguments* for Minimal Conditionalization. This gives us a systematic way to tackle the justification question: we can remind ourselves of how those arguments for Minimal Conditionalization are supposed to work in the non-branching case, and then examine the extent to which those arguments also justify Extended Conditionalization in the general case. My claim is that straightforward adaptations of those arguments equally well justify Extended Conditionalization in the general case. (Note that there is no conflict here because, in the non-branching case, Minimal Conditionalization, Naive Conditionalization and Extended Conditionalization all coincide.)

### 5.2.1 The Principle of Minimum Information

Our first classical (i.e. non-branching) argument in favor of Minimal Conditionalization is that of Paul Williams (1980). This argument will yield inconclusive results for the general case, but the reasons for its semi-failure will be illuminating.

Williams proposed the following principle of probabilist belief dynamics:

> **Principle of minimum information.** Given the prior distribution $P^0$, the probability distribution P appropriate to a new state of information is one that minimises $I(P, P^0)$ subject to whatever constraints the new information imposes.

Here, $I(P, P^0)$ is the *information in P relative to $P^0$*. In the case in which the sample space has finite size, $I(P, P^0)$ is given by

$$I(P, P^0) = \sum_{j=1}^{n} p_j log \frac{p_j}{p_j^0},$$

where the summation ranges over elements of the sample space.

It is easy to show that, in the special case in which the 'constraint' has the form $P(X) = 0$ for some subset $X$ of the sample space, the Principle

of Minimum Information entails that the 'appropriate' posterior is the one given by conditionalizing the prior $P^0$ on the set $\neg X$. That is, in the non-branching case, the Principal of Minimum Information recommends Minimal Conditionalization.

Not much is offered by way of justification for the normative status of this principle, so it would in any case be difficult to assess the question of whether or not the principle carries the same normative force in the general case. However, we have a prior problem with trying to carry this justification across to the general case. The problem is that, in the cases Williams considers, the *sample space* over which the probability distributions are defined is the same both prior and posterior to updating. In contrast, in the general case, whereas Alice's *credence* function at $t_0$ is defined over the set $\mathcal{W}_{t_0}^C$ of centered worlds at time $t_0$, her credence function at $t_2$ is defined over the set $\mathcal{W}_{t_2}^C$ of centered worlds at time $t_2$. When any branching is going on, these two sets fail to be in 1:1 correspondence. Unless we can reconstrue the branching problem as one in which pre- and post-updating sample spaces are identical, the Principle of Minimum Information is not applicable.[18]

There is, in fact, a way of reconstruing the general problem so that it meets this demand. The key is to notice that, while we have a problem if we compare $t_0$ and $t_2$, Alice's credence functions at $t_1$ and $t_2$ are defined on the same sample space: $\mathcal{W}_{t_2}^C$ in both cases. (This is because $\mathcal{W}_{t_1}^C = \mathcal{W}_{t_2}^C$.) Then we could try to argue as follows:

**Generalization of the Principle of Minimum Information argument to the general case.**

1. By appeal to a decision-theoretic Principle of Reflection (or a particularly simple Dutch Book argument — see below for the general idea of Dutch Book arguments), argue that, since Alice receives no relevant information between $t_0$ and $t_1$, her betting quotients at those two times had better be equal to one another.

2. Since (we are assuming) Alice's betting quotients at $t_0$ are given by the quasi-credence function $\overline{Cr}_{t_0}$ defined in section 3.2, and those at $t_1$ are given by her credence function $Cr_{t_1}$ at that time, we obtain the result that $Cr_{t_1} = \overline{Cr}_{t_0}$.

---

[18]Of course, Alice's *quasi*-credence function at $t_0$ is defined over the same set as her credence function at $t_2$, but it would be question-begging to claim that that was relevant.

3. Nothing in the Principle of Minimum Information leads us to think that that principle should be restricted to cases of purely propositional (i.e. non-self-locating) belief. So, insofar as the Principle justifies Minimal Conditionalization in the non-branching case, it also justifies Minimal Conditionalization for Alice between $t_1$ and $t_2$: $Cr_{t_2}^{E_i}(\cdot) = Cr_{t_1}(\cdot|E_i) \equiv \overline{Cr}_{t_0}(\cdot|E_i)$. This has the result that $f_{EC}$ is the unique preferred updating policy.

I think this argument is, in fact, sound. However, the strategy of *relying* on this argument alone in order to defend Everettian QM would have two rather significant drawbacks:

- The adapted Minimum Information argument *essentially relies on the fact that Alice at $t_1$ is genuinely uncertain about which outcome has occurred (on her branch).* This reliance is bad for two reasons, one of which is far more worrying than the other.

  - The first (and very weak) reason is that, since the argument proceeds *via* considering Alice's post-measurement, pre-observation self-locating uncertainty, the argument cannot *directly* be applied in cases in which such uncertainty does not occur. This is a common situation: it occurs in any measurement in which the agent discovers the result of the experiment so soon after branching that there is no time period during which she *does* have the self-locating uncertainty of Alice$_1$.

    The reason that this is a weak objection is as follows. A straightforward continuity argument can assure us that if updating works in the normal way when the post-measurement, pre-observation period of uncertainty has any finite temporal duration, then updating should continue to work in that normal way in the limit in which the duration of the period of uncertainty shrinks to zero.

    However, perhaps not everyone will be convinced by this reply. An argument that does not appeal to credences at $t_1$ at all would allow us to bypass this issue.

  - Second, and more seriously, exclusive reliance on the adapted Minimum Information argument would promote the idea that genuine *uncertainty* about which outcome (if not pre-measurement, then post-measurement) occurs is an *essential* part of the Everettian's epistemological story, even in cases in which all details of the branching structure are known. One of the central aims

34

of this paper is to dispel this notion. To fulfil this aim, we must supply an argument that does not depend on this sort of 'branch uncertainty' at all.

- The argument cannot get off the ground without the initial step linking the 'quasi-credence function' $\overline{Cr}_{t_0}$ at $t_0$ with the credence function $Cr_{t_1}$ at $t_1$. This step is highly nontrivial. We must not rely on a question-begging choice of the terminology 'quasi-credence function' in order to smuggle the argument past intuition; it must be *argued* that $\overline{Cr}_{t_0}$ deserves the name we have given it, with respect to epistemic as well as practical matters. In fact, the sorts of considerations that should lead us to accept this initial step of the adapted Minimum Information argument should also lead us on a more direct route to accepting the unique rationality of Extended Conditionalization. In that case, the appeal to Minimum Information itself becomes something of a free-spinning wheel. (This point will become clearer following sections 5.2.2 and 5.2.3.)

In summary, while the Principle of Minimum Information argument is (I claim) sound, to rely on this argument alone to justify Extended Conditionalization for the general case would be both misleading and question-begging. I therefore turn now to two further justifications for conditionalization that will not have these flaws: the Dutch Book argument, and the appeal to expected epistemic utility.

### 5.2.2 The Dutch Book argument

**The original Dutch Book argument.** The second justification of Minimal Conditionalization that we shall consider is the Dutch Book argument (first proposed in print by Teller (1976)). The key idea in this argument is the following. If the agent (in a non-branching case) chooses any updating policy other than conditionalization, and if in addition the agent is willing to accept all bets that are "fair" according to her own degrees of belief at the time at which the bet is offered, then a bookmaker who had no more knowledge than the agent but who was aware of the agent's updating policy could construct a "Dutch Book" against that agent.[19] A 'Dutch Book' is a set of bets, each of which the agent considers fair at the time it is offered, such

---

[19]In (Greaves, 2004), I wrote that the Dutch Book could not be made against an agent who updated her beliefs trivially: $Cr_{t_2}^{E_1} = Cr_{t_2}^{E_n} = Cr_{t_0}$. As the proof given in appendix A makes clear, this was just incorrect.

that, if the agent accepts *all* the bets offered, then the agent is guaranteed a net loss.

The existence of a Dutch Book for the scenario represented in figure 2, in the case in which Alice's post-observation credence functions $Cr_{t_2}^{E_1}, \ldots, Cr_{t_2}^{E_n}$ are related to her pre-measurement credence function $Cr_{t_0}$ other than by conditionalization, is proved in appendix A.

Proponents of the Dutch Book argument take susceptibility to a Dutch Book to be a sufficient condition for epistemic irrationality; thence they conclude that conditionalization is the unique epistemically rational updating policy.

**The Dutch Book argument and branching universes.** Our next task is to consider the following questions. How would a would-be Dutch Bookie proceed if one of the theories that he and Alice were contemplating was a branching-universe theory, i.e. if the set of possibilities was as represented in figure 2? Would he meet with the same degree of success as in the non-branching case?

It is fairly straightforward to see that the Dutch Book argument supports the (unique) epistemic rationality of Extended Conditionalization (cf. section 3.2) in the general case, in much the same way that it supports Minimal Conditionalization in the purely non-branching case. The crucial observation is the following. We have already accepted (as part of our solution to the Practical Problem, in section 2) that the probability function $\overline{Cr}_{t_0}$ operates as a credence function as far as rational action is concerned; the only point at which the genuine credence function $Cr_{t_0}$ enters the Dutch Book justification of Minimal Conditionalization for the non-branching case is via its role in acceptance and rejection of bets; choices of which bets to accept and reject are choices of action. It follows that, in the general case, the bookie can make a Dutch Book against Alice if Alice updates other than by $f_{EC}$, i.e. other than by 'conditionalization' from her effective credence function $\overline{Cr}_{t_0}$.

**Criticism of Dutch Book arguments.** I have argued that the Dutch Book argument justifies Extended Conditionalization for the general case *just as well as* it justifies Minimal Conditionalization for the purely non-branching case. However, we must also consider how good (or bad) this justification is. There are a number of reasons to be suspicious.

In the first instance, the bets that the bookie must use in constructing the Dutch Book are, in part, bets on the truth of a scientific theory such

36

as QMS or QME, rather than simply bets on experimental outcomes. It is obviously going to be difficult actually to carry out such bets. The Dutch Book could be implemented in practice only if Alice and the bookie agree on some procedure for conclusively determining whether or not the bet has been won; unless there happens to be a trusted god around, this will be difficult to arrange.

Whether or not one considers this to be a serious objection depends on what one takes to be the exact reasoning from "Alice is susceptible to a Dutch Book" to "Alice is epistemically irrational". If the point was supposed to be that susceptibility to a Dutch Book is a real practical liability that self-interested rational agents will take steps to avoid, then the objection is clearly a serious one — an agent need not worry, in practice, about losses she might incur as a result of bets that are in-practice-impossible to implement.

However, as has been noted in the Dutch Book literature, understanding the logic of a Dutch Book argument in this way is anyway implausible. The point of a Dutch Book argument surely cannot be to highlight a real practical liability that the agent had better take steps to avoid. There are plenty of ways to update beliefs in wild and wonderful (non-conditionalizing) ways while thwarting would-be Dutch Bookies: for instance, one could refrain from advertising one's intended updating policy, and/or one could simply refuse to take any bets. If such arguments have a point at all, it must rather be that, "somehow", vulnerability to a Dutch Book points out an underlying incoherence in one's beliefs. I conclude that the fact that the bets in question cannot actually be implemented does not *further* detract from the force of the Dutch Book argument.

The broader issue raised by this objection, however, is an important and controversial one. The challenge is to expand the "somehow" in the previous paragraph into a coherent account of just why susceptibility to a Dutch Book *should* be taken as an indicator of epistemic irrationality. Several authors believe that it should not. A common objection is that the Dutch Book shows only that updating other than by conditionalization would, under certain exotic circumstances, be *prudentially* irrational. There is a distinction (these authors insist) between prudential and epistemic rationality; to pursue riches is not necessarily to respect one's epistemic duties. While there have been attempts (e.g. (Christensen, 1996)) to reply to this objection and defend the relevance of Dutch Book arguments to *epistemic* rationality, the issue remains controversial.

It is therefore worth looking for an argument that can bypass this controversy, by avoiding reliance on the relevance of *prudent* betting behaviour to epistemic rationality. This was the motivation behind a third attempt to

justify conditionalization — the argument from expected epistemic utility. We turn now to this argument, and its generalization to the general case.

### 5.2.3 The Expected Epistemic Utility argument and Extended Conditionalization

The source of the objection to the Dutch Book argument is not that that argument uses considerations of expected utility. Rather, it is that the Dutch Book argument uses *the wrong kind of utility function*. To elaborate: the reason that the Dutch Book argument seems to play only to considerations of *prudential* rationality is that it assumes that utility is linear in dollars. The worry here is not that the *linearity* assumption is implausible; rather, it is that there is no reason to think that dollars are relevant *in any way whatsoever* to considerations of epistemic rationality.

What sort of utility function would do better justice to the spirit of epistemic rationality, then? The central idea behind our third argument for conditionalization (Greaves & Wallace, 2006) is that if we had a utility function that captured *the epistemic value of being in a given belief state*, then epistemic rationality would consist in maximizing the (subjective) expectation value of *that* utility function. It turns out that, on such an approach, conditionalization can be justified on the basis that, of all possible updating policies, conditionalization is the one that yields the highest value of expected epistemic utility. (An heuristic gloss: 'Conditionalization is epistemically rational because it is likely to lead you to epistemically good belief states.'[20])

The expected-epistemic-utility argument carries over just as well as the Dutch Book argument does to the Everettian case. The crucial point, once again, is that the Everettian branch weight measure $W_{\mathrm{QME}}$ plugs into expected utility calculations in exactly the same way in which a mathematically identical chance function would plug in if QME were a chance theory. It follows that, if Alice is rational, she will select Extended Conditionalization as her updating policy. ('Extended Conditionalization is rational because it is likely to lead you to epistemically good belief states on the branches you care the most about.')

To summarize: We have considered three arguments for the epistemic rationality of conditionalization: the argument from the Principle of Minimum Information, the diachronic Dutch Book argument, and the argument

---

[20]Since the details of this argument, and the framework required to state the claim precisely, are somewhat involved, I will not repeat them here; the reader who wishes to assess this argument in the non-Everettian case is referred to the cited paper.

from expected epistemic utility. When applied to a case in which one of the agent's candidate theories is a branching-universe theory, each of the arguments uniquely sanctions the Extended Conditionalization that places Everettian quantum mechanics on an epistemic footing equal to that of any other interpretation of quantum mechanics; none sanctions the pathological Naive Conditionalization that led us into trouble in section 4.

This concludes the core of our argument for the claim that the Epistemic Problem can be solved within the rationality-based approach to probability in the Everett interpretation.

# 6   Remarks on intertemporal consistency

There is, however, something curious about the argument that I have presented so far. The task of this section is to highlight what that is, and to present responses. The discussion will throw some light on the place of the above arguments in an overall defense of the Everett interpretation.

Under relatively uncontentious auxiliary assumptions, a necessary and sufficient condition for solving the Epistemic Problem is as follows: establish that the agent who believes that QME is true and that branching has occurred, but who has gained no post-split information as to which branch she is on, should set her self-locating credences equal to the assumed branch weights. But the arguments of section 3 and 5 do *not*, *directly*, establish this. What they (more directly) show is that a *pre*-measurement Everettian agent, if she was able to do something that would constrain her future selves to adopt particular sets of degrees of belief on every future branch, should choose to constrain her future selves to set their (self-locating) degrees of belief according to those branch weights. In the absence of such a constraining mechanism, whether or not her future selves choose to obey is (until we have invoked some principle of intertemporal consistency) an open question. This observation invites the following objection:

**'Irrelevance of hypothetical predecessors' objection.** In order to solve the Epistemic Problem, we require an explanation of why we should, *now*, conditional on the assumption that the Everett interpretation is true, have a high degree of belief that high-weight outcomes have occurred on the branch that we are now on. According to the Everettians' rationality-based program (without subjective-uncertainty), this explanation is: well, *yesterday*, a predecessor of mine would have cared a lot about branches on which high-weight outcomes occur, and intertemporal consistency requires that my present degrees of belief

39

match the 'caring measure' that she had (or would have had) yesterday. But how much someone yesterday would have cared about some branch or another cannot possibly be relevant to what I, now, should believe (even if that 'someone' is me). This point becomes even more obvious when we note that we are often concerned with evidence resulting from quantum measurements that took place before we were even born. In those cases, the predecessors in question are not even real. This shows that there is something fishy about the argument that has been offered so far.[21]

This is an important and illuminating objection, but it can be answered. There are two (mutually complementary) responses.

**Response 1: the post-measurement credence function can, if desired, be obtained without appeal to predecessors.** The key to the first response is to remember that the notion of probability in the Everett interpretation seemed problematic, in the first place, only from the *pre-*measurement perspective. From the post-measurement perspective, we can (even on the OD semantics) make sense of genuine uncertainty: the agent is uncertain about which outcome occurred on the branch she is on.

From the post-measurement point of view, therefore, it is only the *quantitative* problem (why should our credences respect the branch weights?), not the *incoherence* problem (why is the notion of credence appropriate at all?), that requires a solution. To answer the quantitative problem, two remarks should be made. First remark: in the context of a non-branching theory, one typically accepts a primitive rationality postulate (the Lewisian Principal Principle) without argument. To demand more in the case of Everettian quantum mechanics is therefore to invoke double standards. Second remark: one could apply the sort of symmetry argument that was supposed to justify the Everettian Principal Principle from the pre-measurement perspective (i.e. the sort of symmetry argument involved in proofs of the representation theorems given by Deutsch (1999) and Wallace (2003), (2005c)) directly to the post-measurement perspective. The idea that the post-measurement agent should 'have equal credence in all branches', for instance, falls foul of exactly the same objections as the idea that the pre-measurement agent should 'care about all branches equally' (viz. that this suggestion presupposes more structure than there in fact is; cf. (Greaves, 2004, section 5.3) and/or (Wallace, 2005c, section 9)).[22]

---

[21]I am grateful to Tim Maudlin for raising this issue.

[22]David Wallace has suggested (2005a, p.17) that the Everettian representation theo-

**Response 2: the arguments for Extended Conditionalization serve as a consistency check.** Response 1, however, leaves it somewhat mysterious whether the arguments adduced in this paper have any role to play in a defense of the Everett interpretation: if we can argue from the pre-measurement perspective that Everettian branch weights play the same role as chances in practical action, and we can argue from the post-measurement perspective that Everettian branch weights play the same role as chances in theory confirmation, then we have solved both the Practical Problem and the Epistemic Problem; who cares, then, about the link between the two?

We must care also about the link, for the following reason. If, having independently advocated a particular strategy for rational action and a particular epistemic strategy, we were to find that the combination of the two led to failures of intertemporal consistency — if, for example, it turned out that the pre-measurement agent should desire to constrain her post-measurement selves to adopt credence functions that do *not* count as reasonable from the post-measurement perspective — then our argument as a whole would have landed us in paradox. This would have been the case, for instance, if Naive Conditionalization (section 4) had turned out to be the unique otherwise-rational updating policy in the general case. The defenses of Extended Conditionalization presented in this paper show that such intertemporal inconsistency does not in fact occur. The strategies for rational action and for rational belief in the presence of a branching-universe theory, even on the OD semantics, complement one another in just the same way that they complement one another in the absence of a branching-universe theory.

# 7    Conclusion

There was an intuitive worry that the usual sorts of empirical evidence for quantum mechanics would not count as evidence also for Everettian quan-

rems (ERTs) succeed where non-Everettian symmetry arguments, similarly aimed at deriving rationality constraints on credence functions, fail, and that the reason is that there is (from the *pre-measurement Everettian* perspective, but not from the post-measurement Everettian perspective or from the non-Everettian perspective) no fact of the matter as to which outcome will occur. *If* Wallace is correct, then the ERT arguments do indeed have more force from the pre-measurement perspective than from the post-measurement perspective; the 'Irrelevance of hypothetical predecessors' objection then returns as an objection to Wallace's claim that the Everettian can do *better* than the non-Everettian. As mentioned in footnote 6, I am skeptical of Wallace's suggestion, but this is not the place to investigate further.

tum mechanics, unless we accept the controversial 'subjective-uncertainty' semantics. But working through the details shows this worry to be unfounded: the arguments behind our best account of empirical confirmation, when applied to the Everettian case, show that such evidence confirms stochastic quantum mechanics and Everettian quantum mechanics in exactly the same way. This leads to my first conclusion: that the rationality-based approach can solve the Epistemic Problem of probability in the Everett interpretation.

We should also draw a second conclusion from this discussion. The rationality-based approach to probability in the Everett interpretation is usually taken, by advocates and critics alike, to require the 'subjective-uncertainty' (SU) thesis defended by Saunders and Wallace. But this consensus is mistaken. In a previous paper, I argued that the rationality-based solution to the Practical Problem can be justified just as well without SU as with SU. The present paper has presented a solution to the Epistemic Problem which, again, does not *require* SU. The argument could easily be rephrased in terms more congenial to the SU viewpoint (we could, for instance, simply refer to $\overline{Cr}_{t_0}$ as a credence function, rather than introducing the terminology 'quasi-credence function', and then run the argument as before). If SU is correct, that rephrasing should be preferred. But such a rephrasing would neither add to nor detract from the force of the argument in defense of Extended Conditionalization. This leads to my second conclusion: that the issue of whether or not SU is correct is orthogonal to the issue of whether or not Everettian quantum mechanics is an acceptable physical theory.

## Acknowledgements

## A   The Dutch Book argument

This appendix contains the details of the Dutch Book argument appealed to in section 5.2.2.

Consider the following pair of bets:

| | $T \wedge E_i$ | $\neg T \wedge E_i$ | $\neg E_i$ |
|---|---|---|---|
| Bet 1 | $R \cdot (1 - \overline{Cr}_{t_0}(T|E_i))$ | $-R \cdot \overline{Cr}_{t_0}(T|E_i)$ | 0 |
| Bet 2 | $S \cdot (1 - Cr_{t_2}^{E_i}(T))$ | $-S \cdot Cr_{t_2}^{E_i}(T)$ | 0 |

Each row of this table represents one bet; each column represents one proposition. The table entries give the amount (which may be positive or negative) that the parties agree that Bob will give Alice if Alice accepts the bet in question, in the event that the corresponding possible world turns out to be actual.

Bet 1 is offered at time $t_0$, i.e. before the measurement has occurred. Bet 2 is offered, after the measurement, iff the outcome is $E_i$.

$\overline{Cr}_{t_0}$ and $Cr_{t_2}^{E_i}$ are (as usual) Alice's quasi-credence function at time $t_0$ and her credence function at time $t_2$ having observed $E_i$, respectively. (Recall that, when no branching-universe theories are under consideration, the 'quasi-credence function' just is the credence function.) $R$ and $S$ are monetary amounts whose values are yet to be fixed; they may be positive or negative.

First, we note that Alice's own quasi-credence function commits her to regarding each of these bets as fair at the time it is offered:

**Definition:** Say that a bet is *fair for Alice at time $t$* iff the expected value of the bet, evaluated according to Alice's credence function at time $t$, is zero.

**Claim 1:** Bet 1 is fair for Alice at time $t_0$. Bet 2 is fair for Alice at time $t_2$ if $E_i$ has occurred.

**Proof of claim 1:** The expected value of bet 1, evaluated with respect to $\overline{Cr}_{t_0}$, is given by

$$
\begin{aligned}
EU_{\overline{Cr}_{t_0}}(\text{Bet1}) &= \overline{Cr}_{t_0}(T \wedge E_i) \cdot R \cdot (1 - \overline{Cr}_{t_0}(T|E_i)) \\
&\quad + \overline{Cr}_{t_0}(\neg T \wedge E_i) \cdot R \cdot (-Cr_{t_0}(T|E_i)) \\
&= \frac{1}{\overline{Cr}_{t_0}(E_i)} \cdot R \cdot (\overline{Cr}_{t_0}(T|E_i) - (\overline{Cr}_{t_0}(T|E_i))^2 \\
&\quad - \overline{Cr}_{t_0}(T|E_i) + (\overline{Cr}_{t_0}(T|E_i))^2) \\
&= 0.
\end{aligned}
$$

The second component of the claim follow from a similarly trivial calculation.

Next, we claim that, if Bob knows Alice's quasi-credence function and updating policy, and if in addition Alice updates her credence in $T$ other than by conditionalization, then Bob can choose the values of $R$ and $S$ so that Alice is guaranteed a net loss if outcome $E_i$ occurs, whether $T$ is true or not:

**Claim 2:** If $Cr_{t_2}^{E_i}(T) \neq \overline{Cr}_{t_0}(T|E_i)$, then there exist values of $R$ and $S$ such that Alice makes a net loss in each of the possible worlds in which $E_i$ occurs.

**Proof of Claim 2:** By symmetry, it suffices to consider the case $Cr_{t_2}^{E_i}(T) > \overline{Cr}_{t_0}(T|E_i)$. In this case, take $R < 0$, and let $S = -R$. Then Alice's net gain, in the case in which $T \wedge E_i$ holds, is given by:

$$
\begin{aligned}
& R \cdot (1 - \overline{Cr}_{t_0}(T|E_i)) + S \cdot (1 - Cr_{t_2}^{E_i}(T)) \\
= \; & R \cdot ((1 - \overline{Cr}_{t_0}(T|E_i)) - (1 - Cr_{t_2}^{E_i}(T))) \\
= \; & R \cdot (-\overline{Cr}_{t_0}(T|E_i) + Cr_{t_2}^{E_i}(T)) \\
< \; & 0,
\end{aligned}
$$

since $R < 0$ and $-\overline{Cr}_{t_0}(T|E_i) + Cr_{t_2}^{E_i}(T) > 0$.

Similarly, Alice's net gain in case $\neg T \wedge E_i$ holds is given by:

$$
\begin{aligned}
& -R \cdot \overline{Cr}_{t_0}(T|E_i) - S \cdot Cr_{t_2}^{E_i}(T) \\
= \; & R \cdot (-\overline{Cr}_{t_0}(T|E_i) + Cr_{t_2}^{E_i}(T)) \\
< \; & 0.
\end{aligned}
$$

Repeating the argument for each candidate theory and each outcome, we obtain the full Dutch Book result: if Alice updates her credences (in *any* theory, given *any* outcome) other than by conditionalization on $\overline{Cr}_{t_0}$, then there exists a set of bets, each of which Alice regards as fair at the time it is offered, such that, if Alice accepts all bets in the set, she makes a net loss in every possible world.

# References

Baker, D. (2006). *Measurement outcomes and probability in Everettian quantum mechanics.* (Forthcoming in *Studies in History and Philosophy of Modern Physics*)

Christensen, D. (1996, September). Dutch books depragmatized. *Journal of Philosophy*, *93*(9), 450-479.

Deutsch, D. (1999). Quantum theory of probability and decisions. *Proceedings of the Royal Society of London*, *A455*, 3129-37. (Preprint available online from www.arxiv.org.)

G. Ghirardi, A. R., & Weber, T. (1986). Unified dynamics for micro and macro systems. *Physical Review D*, *34*, 470.

Greaves, H. (2004, September). Understanding Deutsch's probability in a deterministic multiverse. *Studies in History and Philosophy of Modern Physics*, *35*(3), 423-56.

Greaves, H., & Wallace, D. (2006, July). Justifying conditionalization: Conditionalization maximizes expected epistemic utility. *Mind*, *115*, 607-632. (Preprint available online from http://philsci-archive.pitt.edu)

Hall, N. (2004, March). Two mistakes about credence and chance. *Australasian Journal of Philosophy*, *82*(1), 93-111.

Lewis, P. (2006). *Uncertainty and probability for branching selves.* (Forthcoming in *Studies in History and Philosophy of Modern Physics.*)

Saunders, S. (1998). Time, quantum mechanics, and probability. *Synthese*, *114*, 373-404.

Saunders, S. (2005). What is probability? In S. D. A. Elitzur & N. Kolenda (Eds.), *Quo vadis quantum mechanics.* New York: Springer.

Savage, L. J. (1972). *The foundations of statistics* (2nd ed.). New York: Dover.

Teller, P. (1976). Conditionalization, observation, and change of preference. In W. Harper & C. A. Hooker (Eds.), *Foundations of probability theory, statistical inference, and statistical theories of science.* Dordrecht: D. Reidel.

Wallace, D. (2003). Everettian rationality: defending Deutsch's approach to probability in the Everett interpretation. *Studies in History and Philosophy of Modern Physics*, *34*, 415-439. (Available online at http://www.arxiv.org/abs/quant-ph/0303050.)

Wallace, D. (2005a). *Epistemology quantized: Circumstances in which we should come to believe the Everett interpretation.* (Forthcoming in *British Journal for the Philosophy of Science.* Available online at http://philsci-archive.pitt.edu/archive/00002368.)

Wallace, D. (2005b). *Language use in a branching universe.* (Available online at http://philsci-archive.pitt.edu/archive/00002554/02/branch_dec05.pdf and from http://philsi-archive.pitt.edu.)

Wallace, D. (2005c). *Quantum probability from subjective likelihood: im-*

*proving on Deutsch's proof of the probability rule.* (Forthcoming in *Studies in History and Philosophy of Modern Physics.* Available online at http://arxiv.org/abs/quant-ph/0312157 and from http://philsci-archive.pitt.edu.)

Williams, P. (1980, June). Bayesian conditionalisation and the principle of minimum information. *The British Journal for the Philosophy of Science, 31*(2), 131-144.