Measurement Outcomes and Probability in Everettian Quantum Mechanics

David J. Baker

Department of Philosophy,1879 Hall

Princeton University

Princeton, NJ 08544-1006

`djbaker@princeton.edu`

April 11, 2006

Abstract

The decision-theoretic account of probability in the Everett or many-worlds interpretation, advanced by David Deutsch and David Wallace, is shown to be circular. Talk of probability in Everett presumes the existence of a preferred basis to identify measurement outcomes for the probabilities to range over. But the existence of a preferred basis can only be established by the process of decoherence, which is itself probabilistic.

Keywords: Quantum mechanics, Everett interpretation, Many-worlds interpretation, Decoherence, Probability.

1   Introduction

The Everett or many-worlds interpretation of quantum mechanics has long suffered from a problem about probability. It is all very well to say that the wavefunction never collapses and the linear dynamical laws always apply, leading to a branching universe in which every possible outcome of each measurement is actual. But how are we then to explain the fact that we are justified in treating some measurement outcomes as more likely than others? Quantum

mechanics is an empirically successful theory only because it combines the linear dynamics with a rule for determining the probabilities of possible outcomes from the amplitude of the wavefunction. This is the Born rule. If they are to explain this success, Everett's adherents must explain how this explicitly probabilistic rule enters into a theory that they interpret as deterministic.

David Deutsch (1999) has offered a supposed derivation of the Born rule in Everettian quantum mechanics from the linear dynamics combined with axioms of decision theory. More recently, David Wallace (2003a) has clarified and expounded on Deutsch's proof. If correct, the Deutsch-Wallace approach justifies the use of the Born rule as a guide to rational action: rational agents are constrained to apportion their preferences about measurement outcomes according to this probability measure.

But all is not well with this approach to quantum probability – or so I will argue here. Before Everettians can coherently speak of probability, they must provide a definition of what constitutes a measurement outcome under their interpretation. Defenders of Everett such as Wallace (2003a) and Hilary Greaves (2004) have identified measurement outcomes with *people* – the post-branching successors of the person making the measurement. This cannot be right, I claim, for on this construal amplitude is not a measure over outcomes of measurements, and so probabilities (which must apply to outcomes) cannot be derived from amplitude.

This misunderstanding is not difficult to correct, if one takes measurement outcomes to be classes of "branches," or elements of a superposition of states. The problem of identifying measurement outcomes thus reduces to the problem of selecting a preferred basis: which terms are we to count as branches? The Everettian may now claim to stand on solid ground – after all, decoherence leads to the elimination of interference between branches and the

(vague or approximate) selection of a preferred basis.

But this decoherence-based approach leads to a further problem. Proofs of decoherence depend implicitly upon the truth of the Born rule. Without first justifying the probability rule, Everettians cannot establish the existence of a preferred basis or the division of the wave-function into branches. But without a preferred basis or a specification of branches, there can be no assignment of probabilities to measurement outcomes. This has two unfortunate consequences for the Everett interpretation: it reveals a circularity in the Deutsch-Wallace approach, and it calls into question the coherence of the primitivist approach to probability advocated by Greaves.[1]

2   The Quantum Representation Theorem

The objective for Deutsch and Wallace is to prove the Born rule, which expresses the connection between probability and the wavefunction's amplitude. Suppose that a quantum system is in a state $\Psi = c_1\phi_1 + c_2\phi_2 + ...c_n\phi_n$, where the ϕ_i are eigenstates of the observable to be measured (with corresponding eigenvalues λ_i) and the c_i are (complex) constants. Then the Born rule states that the likelihood of a measurement giving result λ_i is $|c_i|^2$, the normed squared amplitude or, to use Wallace's term, the *weight* of the i'th component of the superposition.

To show the applicability of this rule to the Everett interpretation, Deutsch and Wallace employ the axioms of causal decision theory. Decision theory governs rational action under statistically weighted uncertainty, and so the decision-theoretic description of any process includes a measure of probability or likelihood. The role of this measure is to dictate rational

---

[1]As should by now be clear to the reader, my target is what might be called the "Oxford reading of Everett," which I take to involve a decoherence-based solution to the preferred basis problem and a decision-theoretic account of probability. In particular, I am not concerned with "many-minds" interpretations.

action through the calculation of *expected utility.* For example, suppose that I would prefer to have more money, and a friend offers me the opportunity to bet that a coin toss will come up heads. If I believe that the probability of heads is P, it is rational for me to take any bet for which $P \times$(the agreed-upon payoff) is greater than or equal to $(1 - P) \times$(the stakes). In fact, a rational agent's probability measure can be defined purely in terms of that agent's ordering of preferences.

Deutsch and Wallace seek to define quantum probability in exactly this way. They consider a quantum game: an act of measurement in which the observer is given rewards which vary according to which eigenvalue λ_i is the result of his measurement. It is further assumed that the agent accepts the Everett interpretation, and the following constraints on his preferences:

1. Some axioms of decision theory, whose legitimacy I will not contest here.

2. *Physicality*: If two quantum games describe the same physical process, a rational agent is indifferent between them.

3. *Measurement Neutrality:* Once the system to be measured, the state vector of that system and the observable to be measured have been specified, further details will be irrelevant for decision-making purposes.

Then it can be shown (Wallace 2002, 2003a) that the observer is rationally obliged to apportion his probabilities in accordance with the Born rule. This is the Quantum Representation Theorem (QRT).

3   Measurement outcomes in Everett

In order to understand what the Born rule is saying, and thereby understand what the QRT purports to prove, we must form a conception of what it means for a measurement to give the result λ_i on the Everett picture. This is made more difficult by the deterministic nature of the interpretation. In one sense, any measurement (or any physical process, for that matter) has only one possible outcome according to Everett: the observer sees every result in one branch or another. This poses a seeming problem for the invocation of decision theory, which is normally applied only to decision-making in the face of uncertainty.

The branching structure which arises from decoherence provides the first hint at how to proceed. The Everett interpretation tells us that any "measurement" leads to a superposition of states of macroscopic objects, including the observer, his instruments, and so on. This state can be decomposed into sets of what we shall call "worlds": projections of the wavefunction onto the elements of an orthonormal basis of the Hilbert space of states. Since there are an infinite number of distinct bases, there are infinitely many such decompositions. But a vaguely defined *preferred basis* soon emerges. By the process of decoherence the components of the superposition (expressed as projections along the elements of this basis) become causally isolated from one another and no longer interfere. Each is now a separate "branch" which behaves in rough accordance with the laws of classical mechanics.[2] There are now multiple people who are equally similar to, and causally connected with, the original observer, much as in cases of human fission discussed by Parfit (1984). One possible strategy for the Everettian is thus to consider the applicability of decision theory to cases in which a rational agent undergoes fission.

---

[2]Throughout this paper, I will follow Butterfield (2002) in using this notation: a *world* is a component of any decomposition of the state; a *branch* is a component of the state relative to an element of the decoherence basis.

3.1 Subjective certainty and uncertainty

There are two ways of understanding the application of decision theory to human fission. The first, which Wallace (2002, 2003a) and Simon Saunders (1998) have defended, is the position of subjective uncertainty (SU). The SU picture comes about in the following way. Consider the attitudes that the pre-fission experimenter might form about his successors. First, he might expect to experience nothing after fission. On a reductionist view of personal identity (which is the view assumed by all parties to the debate), this attitude makes little sense, for each of his post-fission successors individually meets all the criteria of mental continuity and connectedness normally required for identity with the original person. Second, he might expect to experience the lives of all of his successors. But this is incoherent, since no one can experience multiple lives at once. The only option remaining, the SU advocate argues, is *uncertainty* as to which successor's experiences will be his.

Therefore, the subject of fission should assume an attitude of uncertainty as to which successor he will become. Thus quantum probabilities are an expression of how likely it is that the initial observer will become any one of his successors, and so come to experience the corresponding branch. In the simplified example of a fifty-fifty quantum coin toss, the observer should have a 50 percent degree of belief that he will become one of the successors who sees heads.

The alternative, more radical position is subjective certainty (SC, also sometimes called objective determinism or the fission program), advocated by Greaves (2004). The believer in SC holds that anyone who is certain about all the physical facts must also be certain about his expectations, and so to view the quantum probabilities as uncertainty about which successor I will become is incoherent. Rather, the probabilities should be viewed as the degree to which I should care about my successors – in effect, the degree to which they *are* my successors.

As Greaves explains,

> I agree that the R-relation – mental continuity and connectedness – captures much of what matters in survival ... I will argue that given the Everett interpretation and Deutsch's axioms, it is *not* all that matters. I therefore introduce the term 'R^*-relation', initially as a place-holder, to denote *whatever it is that matters in survival and in caring for the future.*
>
> ... The Deutsch-Wallace proof is to be understood as claiming to establish that my Everettian successors bear the R^*-relation to me-now in proportion to their weight. With this conceptual base, decision theory can be applied without uncertainty. (2004, p. 444)

So in the case of the equal-weight coin toss, I should care about my heads successors exactly as much as I care about my tails successors, and act accordingly, just as I might equally weigh the interests of the two people resulting from a fission because I know that both will have equally valid claims to be me. But if the weights were unequal – for example, if the tails successors possessed higher total weight than the heads successors – then I would be rationally bound to favor the tails successors. They would, in a sense, be my successors to a greater degree because of the higher weight of the branches they inhabit.

One noteworthy point of agreement between these two perspectives on probability is how measurement outcomes are implicitly identified. Both SU and SC assign probabilities to *future people*. On SU probability provides a measure of the degree to which the experimenter should expect to become certain future people. On SC it measures the experimenter's degree of care for future people. In either case it is a measure over the observer's successors, and so successors fulfill the role of measurement outcomes. This makes intuitive sense, for it is the fission of the observer into multiple successors which creates indeterminacy about

which result the observer will see. This makes it possible to view the successors' experiences as distinct outcomes of the measurement, rather than viewing the experiment as merely a deterministic process with a single outcome.

I do not mean to argue that either SU or SC identifies *individual* future people as measurement outcomes. Rather, it seems accurate to say that both views treat classes of future people as measurement outcomes, in the following way. There may be many branches in which a measurement of z-spin comes out spin up. But the decision-theoretic approach assigns probabilities to the spin eigenvalue, even though it is the observer's fission into multiple successors which permits the application of decision theory. It therefore seems appropriate to identify as an outcome the class of successors who observe spin up.

3.2 Branching and identity

Unfortunately, the identification of measurement outcomes with future people leads to the unacceptable result that probability is not a measure over outcomes of experiments, if one assigns probabilities according to the Born rule. To see why, we must consider a thought experiment which illustrates how personal survival can vary independent of quantum branching. This is the quantum suicide experiment, which has been much discussed in previous literature. It is seen by some (Lewis 2002) as a *reductio ad absurdum* of Everett, and by others (Tegmark 1998) as a weird and fun consequence of the interpretation. I want to show that it precludes us from treating future people as measurement outcomes on the Everett picture.

The quantum suicide experiment is much like the Schrödinger cat experiment, except it is the experimenter's life which is put on the line. The experimenter prepares a machine which measures the z-spin state of a particle which is in the superposition $1/\sqrt{2}(|\uparrow\rangle + |\downarrow\rangle)$,

and wires it to a gun or some other lethal device so that if the particle is measured as spin-up, the experimenter will be killed. After the experiment is carried out, under the Everett interpretation the experimenter will be in the state $1/\sqrt{2}(|alive\rangle + |dead\rangle)$.[3] And after decoherence, we are left with two branches (or classes of branches): one in which the particle is spin-down and the observer is alive, another in which the particle is spin-up and the observer is dead.

The Born rule assigns a probability of 0.5 to each of these two outcomes. A problem now arises for both the SU and SC positions. In one of the branches, there is no person who can claim to be the original observer's successor.[4] So if measurement outcomes in Everett are nothing more than future people who are causally connected with the initial observer, the total probability of all measurement outcomes is 0.5, and the Born rule assigns a probability of 0.5 to something which is not a possible measurement outcome. Adopting the SU perspective, the observer cannot think it 50 percent likely that he will become the successor who sees spin-up – there is no such person. Likewise, on the SC perspective, he cannot care equally for the spin-up and spin-down successors, since there is no spin-up successor to care for.

In order to be useful in predicting experimental results, quantum probability must be a quantity which applies to measurement outcomes. It must be a measure over outcomes, such that the sum total probability of all outcomes is one.[5] On both the SU and SC viewpoints,

---

[3]Or, more accurately, the universe's wavefunction will be in an equal superposition of a state including a living experimenter and a state including a dead one.

[4]To answer the worry that the observer in the "death branch" does survive for some period of time, however brief, a more elaborate experiment can be constructed. For example, the observer might be killed *before* the experiment, but in such a way that he can be revived by life-support equipment. In the spin-down branch he is revived, while in the spin-up branch his corpse is disposed of. And to those who consider this example too arcane to motivate my argument, I would point out that *every* realistic experiment resembles quantum suicide, in that there will be some (low-weight) branches in which, for example, the observer disintegrates as his atoms spontaneously move apart.

[5]If this criterion is not met, we will be vulnerable to a lose-lose "Dutch book" wager.

measurement outcomes are people – classes of post-branching successors of the initial observer. But according to the Born rule, probability is normed squared amplitude, and normed squared amplitude is not a measure over people. It is a measure over eigenvalues of observables. There may be eigenvalues with non-zero weight which do not coincide with any of the initial observer's successors – branches of the universal wavefunction in which the observer has no successor. If the Born rule is correct, which it must be if we accept the QRT, these branches must also correspond to measurement outcomes. Thus it cannot be that quantum probability is a measure of our uncertainty about which of our successors we will become, or of our degree of care for our successors. A new understanding of probability will be needed if the Everett interpretation is to succeed.[6]

3.3 Caring about branches

Consider the "death branch" of the quantum suicide experiment, in which the particle is measured spin-up and the observer is dead. Prior to measurement, the observer cannot adopt the SC or SU viewpoint when considering this branch, since it contains no successor for him to care for. But if we abandon the considerations of personal identity that motivate SU and SC, it may still be possible to give a decision-theoretic account of his subjective probabilities. Although he cannot coherently weigh the interests of the successor who sees

---

[6]Another account of Everettian uncertainty besides SU has also been proposed. On this view, uncertainty enters after the measurement has been made but before the results have been collected – at this time the observer will presumably be uncertain as to which branch he is in. Clearly this view also identifies measurement outcomes with people. This post-measurement uncertainty view is incapable of coping with the quantum suicide example, in which there is no room for uncertainty about the outcome at any point after the measurement. If the post-measurement observer can entertain such uncertainty, he is presumably alive and can therefore infer with certainty that he is not in the death branch. I can see no way of modifying this approach to identify outcomes with branches (see Section 3.3), so I conclude that it fails. (Greaves (2004, 442-443) has suggested that this approach can provide a foundation for probability via a decision-theoretic reflection principle, but this is impossible in the quantum suicide case. Reflection only applies when no new information has been gained, but in quantum suicide the post-measurement observer gains information about the result just by noticing that he's still alive.)

spin-up, he can certainly have preferences about what happens in the death branch. Standard decision theory is able to account for rational agents who possess preferences about events which occur after their deaths. It is exactly this sort of decision-making which enters into such acts as the drafting of a last will and testament. On this desire-satisfaction picture of utility, an agent may rationally assign value to any state of the world, whether it affects his life directly or not.

For any particular Everettian quantum game, there is only one possible state of the world at any given time: the state of the universal wavefunction. But that state is a linear sum of terms – eigenstates of the observable to be measured – which will each correspond to a causally isolated branch of the wavefunction after measurement. Each of these branches represents a macroscopic series of events to which the observer may assign some utility, regardless of whether one of his successors inhabits the branch. It is thus open to the agent to assign an effective expected utility to the entire quantum game by a weighted sum of the utility he ascribes to each branch. This weight might as well be normed squared amplitude – in fact, we may take the QRT as establishing that it *must* be – and it occupies the same role that probability normally fills in decision theory on the desire-satisfaction picture.[7]

On this third view, probability is a measure of the agent's care about what occurs in particular branches, rather than his care for post-fission successors.[8] Measurement outcomes are thus to be identified with non-interfering branches – or, more precisely, with classes of branches. Prior to observing a particle's spin, I consider spin-up to be one possible outcome.

---

[7]Note, however, that the desire-satisfaction picture of utility is not universally accepted. Some hold that the only viable approach is hedonism, on which only subjective experiences can be of value to an agent. It seems that hedonists of this sort will be unable to accept Everettian QM! Though I won't press this point any further, this extraneous commitment to a desire-satisfaction picture of utility strikes me as an inelegant feature for a scientific theory – especially given that no other interpretation of QM involves any such commitment.

[8]This view of measurement outcomes was first suggested to me by Wayne Myrvold.

On a proper understanding of Everett, this means I am identifying as a single outcome the collection of all branches in which the particle is measured spin-up. A class of branches identified in this way, by its correspondence with an eigenvalue of the observable to be measured, avoids the pitfalls of SU and SC. Unlike post-fission people (or classes of people), the classes of branches determined in this way will possess a sum total weight of one. For each of the observable's eigenvalues, there will be exactly one such class of branches, and the summed weight of all eigenvalues is always normalized to one.

Note that this third view does not invoke uncertainty about the outcome of branching. In this way it rejects the SU approach while making only a minor departure from SC. Branches, rather than people, are treated as measurement outcomes; all else remains the same. The current view should thus be seen as a refinement of the SC viewpoint rather than something entirely new.[9]

3.4 Branches and uncertainty

Is there a fourth view available, one which adopts an SU-like approach but accepts that branches, not people, are measurement outcomes? I am skeptical. The SU-like position would have to be something like the following: the Everettian observer should adopt an attitude of uncertainty about which future branch the branch he's currently in will become. This is sounding less and less like *subjective* uncertainty, and more like uncertainty about some objective matter of fact about the future. But all physical facts about the future are fixed by the state of the world right now – so how is this uncertainty possible?

One can know all of the physical matters of fact and still be uncertain about self-locating propositions. For example, if know that I am to be awakened, given amnesia and then

---

[9]In fact, Greaves sometimes seems to indicate that this is one way of understanding SC: "Under an Everettian picture... I am more concerned to maximize my utility in high-weight worlds." (2004, p. 450)

awakened again, I may be uncertain on waking whether this is the first or the second time – even if I know the entire physical history of the world. It is this sort of uncertainty that Wallace (2005) suggests as a new basis for the SU viewpoint. His approach assumes the picture of identity over time (for persons and, in the more general case, branches) proposed by David Lewis (1976).[10] According to Lewis, a persisting object such as a person (or perhaps a branch) is a four-dimensional entity that can be partitioned into temporal parts called person-stages (or branch-stages). In cases of human fission it is possible for two continuant persons to share a person-stage, and so Wallace suggests that two or more branches can also share a branch-stage. So when an Everettian measurement is made and a branch "splits," really there were multiple branches all along.

Subjective uncertainty is then supposed to be uncertainty about which continuant branch I am in. This cannot, of course, be uncertainty about which continuant branch my current branch-stage is a part of. It is part of all the continuant branches that will split off from it. It must be uncertainty about which continuant branch I (the continuant person) timelessly inhabit. Since "the correct referent of utterances and mental states is a person [or branch] at a time (rather than a person [or branch]-stage)" (Wallace 2005, 12), it makes sense to assign some credence or other to the proposition that I am (for example) the person in the spin-up branch, as opposed to the one in the spin-down branch.

This view seems deeply mistaken to me. Let me begin by noting that, on this view, it must make sense to say things like "I'm the man who will go on to see spin-up after this measurement I'm about to do; I also share this person-stage with another guy who'll go on to see spin-down." This is exactly the sort of proposition about persons or branches that the

---

[10]Again we see that the Everett interpretation requires extraneous philosophical commitments, this time in metaphysics instead of decision theory. A Parfitian about personal identity will, as Wallace himself notes (2005, 11-12), be unable to adopt the SU viewpoint.

SU advocate claims we can believe, be uncertain about, and assign some degree of credence. Peter Lewis (2006) argues persuasively that this kind of uncertainty is impossible. Wallace is committed to the notion that when an observer asserts something like "I will see spin-up," there are really two acts of assertion, one attributable to each of the two people who share his current person-stage. This seems odd in light of the fact that the assertion (unlike a person) is an instantaneous event, and so is not subject to problems about identity over time. It makes far more sense, Lewis claims, to say that there is a single act of assertion carried out by the person-stage, and not attributable uniquely to either continuant person. This means that the pronoun 'I' cannot refer uniquely to either person, and so the sort of claims that the SU advocate wants to be uncertain about cannot really be asserted.

Lewis is correct, but I think his point can be put more clearly. The SU advocate claims that we should be uncertain, prior to measurement, about certain self-locating claims like "I am in the spin-up branch." But self-locating propositions are odd in the following way: it is not always possible to believe or express them. Consider the case I mentioned above, in which I am awakened, given amnesia and then awakened again. After I wake up (either the first or the second time) I can coherently say "This is the first time I've woken up." It makes sense to be uncertain about whether or not this is true. But that uncertainty cannot be an appropriate (or even entertainable) attitude prior to my first awakening. The occasion has not yet arisen for me to locate myself in the world, so I cannot yet consider whether the self-locating statement might be true – even though I know that I will later be uncertain. The same is true, I suggest, of the Everettian branching case. Prior to branching, I cannot yet claim to be in the spin-up branch, because there are not yet any physical facts that could make this claim true or false. (Note that self-locating propositions do not typically have any non-trivial *objective chance* of being false, and Wallace has given us no reason to

suppose that the Everettian case should be different in this regard.) It is not possible for me to self-locate until after the branching occurs, and so pre-measurement uncertainty cannot be uncertainty about which branch I am in. The SU position has not advanced a coherent notion of what uncertainty might be in the Everett interpretation.[11]

The reader may agree or disagree with my opinion of SU. Perhaps some version of SU is viable; the central point (and the only point that will be important to the conclusions of this paper) is that any Everettian must accept branches, and not people, as measurement outcomes. Probability will then be some measure over classes of branches: either my degree of care about what happens in certain branches (on SC) or my degree of belief about which branch I am in (on SU), whatever that is supposed to mean.

We now possess a means of determining which aspects of the quantum wavefunction correspond to measurement outcomes on the Everett interpretation. The Born rule assigns to these outcomes the proper quantum probabilities, and the QRT aims to justify the Born rule from Everettian assumptions. Responding to such critics as Barnum *et al.* (2000), Wallace has claimed that the QRT's assumptions are non-probabilistic. This is a crucial point, for any assumption of probability prior to the QRT's establishment could constitute a circularity in the theorem. I will now reveal how one of the premises of the Everett interpretation itself – the process of decoherence – involves a hidden probabilistic assumption which undermines the QRT in exactly this way.

---

[11]Wallace (2005,14-16) also offers an argument from "interpretive charity" for a view like SU, but Lewis (2006, 12-17) argues persuasively that this either fails by breaking the conceptual connection between uncertainty and probability or collapses into the SC view.

4 The premise of decoherence

Among the physical assumptions he accepts as premises of the QRT, Wallace lists the following:

> In the Hilbert space of the theory, a *de facto* preferred basis is defined by decoherence theory. This basis is approximately diagonalized in both position and momentum, and allows the Universe to be interpreted, to a very good approximation, as a branching system of approximately classical regions. (2002, p. 17)

I mean to show that this statement rests on an assumption about how probable certain events are, and so presumes the very principle (the Born rule) that the QRT seeks to prove.

4.1 Decoherence, probability – and circularity

Standard accounts of decoherence, as reviewed in Joos et al. (1996) and Zurek (2003), identify a preferred set of states by environment-induced superselection, or einselection. Einselection takes place roughly as follows (for the simplified case of a two-state system).[12] We begin with a system to be measured, a single particle in a superposition

$$|\Psi_S\rangle = A|s_1\rangle + B|s_2\rangle. \tag{1}$$

Supposing the external environment to be composed of a large number N of spin-$\frac{1}{2}$ particles, the system S and the environment E form a composite system SE characterized by an initial state

$$|\Psi_{SE}\rangle = (A|s_1\rangle + B|s_2\rangle) \bigotimes_{k=1}^{N} (a_k|\uparrow\rangle_k + b_k|\downarrow\rangle_k) \tag{2}$$

---

[12]My presentation here parallels that of Zurek (2003, 15-16).

and an interaction Hamiltonian

$$H_{SE} = (|s_1\rangle\langle s_1| - |s_2\rangle\langle s_2|) \sum_{k=1}^{N} g_k(| \uparrow\rangle\langle\uparrow | - | \downarrow\rangle\langle\downarrow |)_k. \qquad (3)$$

Thus (2) will time-evolve under the influence of this Hamiltonian into

$$|\Psi_{SE}(t)\rangle = A|s_1\rangle|e_1(t)\rangle + B|s_2\rangle|e_2(t)\rangle \qquad (4)$$

where

$|e_1(t)\rangle = \bigotimes_k (a_k \exp(ig_k t)| \uparrow\rangle_k + b_k \exp(-ig_k t)| \downarrow\rangle_k)$ and

$|e_2(t)\rangle = \bigotimes_k (a_k \exp(-ig_k t)| \uparrow\rangle_k + b_k \exp(ig_k t)| \downarrow\rangle_k).$

Thus the entire environment of N particles can be considered as a two-state system with eigenstates $|e_1(t)\rangle$ and $|e_2(t)\rangle$, which has become entangled with the system being measured. The entire system is characterized by a density operator

$$\rho_{SE}(t) = |\Psi_{SE}\rangle\langle\Psi_{SE}| \qquad (5)$$

while the measured sub-system S is characterized by the reduced density operator

$$\rho_r(t) = Tr_e\rho_{SE}(t) = |A|^2|s_1\rangle\langle s_1| + AB^*r(t)|s_1\rangle\langle s_2| + A^*Br^*(t)|s_2\rangle\langle s_1| + |B|^2|s_2\rangle\langle s_2|. \qquad (6)$$

where $r(t) = \langle e_1(t)|e_2(t)\rangle$. Clearly the size of the off-diagonal terms, which determine the extent to which the reduced state behaves non-classically (i.e. exhibiting interference effects) is dictated by $r(t)$. To derive the desired result, that $|s_1\rangle, |s_2\rangle$ form a preferred basis of the measured system and the elements of the decomposition (4) behave like branches, we must

show that the off-diagonal terms can be ignored.

As time passes,[13] the evolution of the system under H_{SE} leads $|e_1(t)\rangle$ and $|e_2(t)\rangle$ to approach orthogonality. For a fixed time t, let $|f_1\rangle$ and $f_2\rangle$ denote two vectors which are exactly orthogonal and are very close in Hilbert space to $|e_1(t)\rangle$ and $|e_2(t)\rangle$, respectively, and whose span includes $|e_1(t)\rangle$ and $|e_2(t)\rangle$.[14] If the state at t were given by

$$A|s_1\rangle|f_1\rangle + B|s_2\rangle|f_2\rangle, \tag{7}$$

instead of by (4), it would be uncontroversial that the two terms in the expansion denote two branches in the Everettian's sense, since there would then be no interference terms and the off-diagonal elements of the reduced density operator (6) would be exactly zero. But there is no time t for which the state is exactly (7), so the Everettian must establish that (4) is "close enough" to (7). It must be shown that the differences between (4) and (7) are somehow minor enough to be ignored.

We can expand (4) to

$$A\langle e_1(t)|f_1\rangle|s_1\rangle|f_1\rangle + A\langle e_1(t)|f_2\rangle|s_1\rangle|f_2\rangle + B\langle e_2(t)|f_1\rangle|s_2\rangle|f_1\rangle + B\langle e_2(t)|f_2\rangle|s_2\rangle|f_2\rangle. \tag{8}$$

The differences between (7) and (4) can then be ignored iff we can justify ignoring the middle two terms in this expansion (8). We know that the weights of these terms (determined by the inner products $\langle e_1(t)|f_2\rangle$ and $\langle e_2(t)|f_1\rangle$) are very low thanks to decoherence. But how can we infer from this that (8) can be treated like (7) from the perspective of a rational

---

[13]At least until the characteristic time of the system has elapsed. This interval is dictated by g_k and $N!$ and in general is very long.

[14]There are many such f's we might choose; this is one sense in which decoherence provides us with a vaguely-defined set of branches.

agent in the Everett interpretation?

It must be that the near-zero weights of these terms permit us to ignore them in decision-making. But we must carefully consider what this claim might mean on the Everett approach. Under a collapse interpretation of quantum mechanics, we ignore low-weight events because they are very unlikely to occur, where likelihood is defined in terms of objective chance. But the Everettian cannot appeal to such concepts.

As we've seen, the Everettian does appeal to a different notion of likelihood: the decision-theoretic approach wherein probability is a measure of our degree of care or self-locating belief about outcomes (classes of branches). So the Everettian might claim that inter-branch interference is improbable in this sense. I can think of three ways of formulating this claim:

B_1 After decoherence we have four branches, given by the terms in the expansion (8), and the two branches involving interference are low-weight and therefore improbable.

B_2 After decoherence we have two branches, given by the terms in (4). Each branch is constituted by an interference term and a non-interfering term, as in (8). Since the interference terms are low-weight, they are improbable.

B_3 After decoherence we have two branches as in B_2. Since these branches are close in Hilbert space to the non-interfering branches in (7), they will appear to be non-interfering branches.

I will now examine these three positions in turn. We shall see that none of them can be given a non-circular justification on the Everett picture.

B_1 is a strange claim for the Everettian to assert, since it entails the existence of more than one branch for each element of the preferred basis. But consider nonetheless what the Everettian must presume in order to assert B_1. The Everett approach, as we've seen, defines

19

'probability' as our degree of care or belief about branches. Until we've identified which worlds are the branches, and indeed whether there are any components of the state which meet this criterion, this term would seem to lack a determinate meaning. Surely we must at least assume that there *is* a branching structure to the universal wavefunction in order to make assertions about (our degree of care or belief about) certain branches, as B_1 requires. But in order to justify the assumption that there are classes of branches for probability to measure over, B_1 must be asserted – otherwise decoherence does not imply the emergence of branching structure, since we have no other grounds for treating (4) as indiscernible from (7). Thus B_1 cannot be justifiably asserted unless we know which worlds (if any) are the branches – but we need to know B_1 in order to know which worlds are the branches. So B_1 can only be justified by circular reasoning.

What about B_2? On the face of it, B_2 is incoherent from the perspective of Everettian decision theory. It makes a claim about the probability of some structure within a branch, but we have seen that probability in Everett is a measure over classes of branches. So to assert B_2 the Everettian must make a category error, attributing some probability to a structure that is not in fact a measurement outcome. I take this to rule out B_2 as a conceptual possibility. But even if it does not, we will see in Section 4.2 that, like B_1, B_2 can be given only circular justification by appeal to the QRT.

B_3 may be the best option for the Everettian. But it rests on the assumption that two state vectors which are close in Hilbert space describe similar worlds. We must then consider whether this assumption can be justified without the further assumption that low-weight events are improbable. If it cannot, then the Everettian cannot assert B_3 without first committing himself to B_1 or B_2.

This further assumption is clearly necessary. Consider the expansion (8). If we suppose

that high (rather than low) probability can be inferred from low weight, I can see no way of justifying the claim that (8) behaves much like a universe of non-interfering branches. To the contrary, it would seem that the low-weight interference terms would, in such a case, dictate the behavior of the branches. A branch in Everett is, after all, supposed to be approximately stable, and naturally-occurring "measurements" are going on all the time. This means that in supposing that the branch given by $A\langle e_1(t)|f_1\rangle|s_1\rangle|f_1\rangle + A\langle e_1(t)|f_2\rangle|s_1\rangle|f_2\rangle$ behaves (or appears from within) like a branch given by $A|s_1\rangle|f_1\rangle$, we are supposing that if the physical quantity associated with the f_i were measured in the former branch, it would reliably give the result f_1. This is so only if high probability can be inferred from high weight.

This is a special case of a more general fact: we only appeal to closeness in Hilbert space as a mark of similarity for quantum states because nearby states give similar statistics. If we drop this assumption, closeness in Hilbert space becomes a mere mathematical relation with no physical relevance. But this assumption, that nearby states assign similar probabilities to events, must for the Everettian be a claim about degrees of care or belief in branches – and so we are back to assuming B_1 or B_2. The latter is incoherent and the former admits only of circular justification.

4.2   Circularity in the QRT

Let's examine this accusation of circularity in the context of the QRT. I claim that, since the employment of decoherence to identify branches depends upon the unlikeliness of low-weight events, the framework of quantum games in which the theorem is formulated presumes its conclusion. Unless the Born rule (or some similar rule permitting inferences from low weight to low probability) holds, there are no quantum games in Everett. Since the QRT implicitly assumes that there are quantum games, it is circular.

Formally, a quantum game is defined as a bet on a measurement procedure. Payoffs are assigned to readout states, which are (and here comes the circularity) defined as elements of the final state of the measured system and measuring device, *as decomposed in the basis selected by decoherence* (Wallace 2003a, 5; 2002, 23). In other words, payoffs are allotted to outcomes as I have defined them – classes of branches indexed by eigenvalues of the measured observable. So for there to be any quantum games, decoherence must result in a state that can be treated as a set of approximately classical branches.

We've seen that this can only be established by asserting one of the three claims – B_1, B_2 and B_3 – explained in the last subsection. But we have already seen that none of these claims can be justified except by the assumption that low-weight events are improbable. (B_1 and B_2 just *are* claims about the probability of low-weight events, and B_3 has been shown to depend on some such claim.) The Everettian will of course agree that low-weight events are improbable, but only (I presume) because the Born rule entails this. If the QRT's conclusions were different, and some other probability rule correlating low weight with (for example) high probability, the Everettian would be obliged to deny B_1, B_2 and B_3. Quantum games assume that decoherence results in a branching universe. This is only true in Everett if one of the three B's is true. But the three B's are only true in Everett if they are implied by the conclusions of the QRT – that is, if the Born rule (or something very similar) applies to rational decision-making. So in assuming that there are quantum games, the QRT's framework assumes the truth of the theorem's conclusions.

In plainer language, Wallace's description of the quantum game requires that payoffs be allotted to specific measurement outcomes. But as we have seen, a measurement outcome is a class of branches. Thus one aspect of the framework in which the QRT is formulated (the description of quantum games) requires that we know which worlds (if any) are the branches;

but in order to know which worlds are the branches we must appeal to the probability rule which is the QRT's conclusion.

Perhaps this circle can be broken by identifying the branches without appeal to the Born rule – that is, by establishing one of the three B's without appeal to a specific probability rule. It might be possible to establish decoherence by resorting to some weaker statement about the relationship between amplitude and probability. But any such assumption would have to be accepted almost *a priori*, from some first principle. I know of only one promising possibility, which was suggested to me by Wayne Myrvold.

Even prior to establishing the QRT, it seems uncontroversial for the Everettian to assume that zero-weight events have a probability of zero. These outcomes are, after all, not components of the wavefunction, and so there are no worlds in which they occur. We can then justify ignoring low-weight events by supposing that we can ignore small variations in *any* physical quantity when determining the functional relationship between that quantity and an event's likelihood.

This crucial assumption can perhaps be justified by appeal to the methodology of theory construction. In practice, it is impossible for us to establish exactly the quantum state of the system we are studying prior to measurement. There is a practical limit to how precisely physical processes, including our instruments, can prepare a system. In order for quantum mechanics to be useful in predicting experiments, we must be justified in assuming that states which are nearly equal will predict similar patterns of outcomes.

If this amounts to just the assumption that probabilities are related to weights by a continuous function, it is too weak to imply that branching behavior occurs in any finite time. If that is the best the Everettian can do, my argument is in good shape. But this way of formalizing the assumption is not very fair to Myrvold's point, which is that acceptable

theories must entail a relationship between their states and experimental predictions which is continuous *at a scale that is measurable and manipulable by us.* The result would seem to be that no theory which predicts large variations in probability between states that we cannot reliably differentiate prior to measurement (for instance, by preparing them through emission from different radioactive isotopes or particle accelerators) could be empirically satisfactory.

This sort of "anthropocentric" continuity assumption may well be strong enough to establish the intended conclusion, that after decoherence the leftover interference effects are improbable enough to be ignored. But this is no help to the Everettian without some further argument that the continuity assumption is actually true of Everettian QM. And the line of argument above which establishes that $\text{weight}(A) = 0$ implies $\Pr(A) = 0$ does not support any further claims about the probability of outcomes with nonzero weight, however small. If the wavefunction of the universe assigns zero weight to a particular eigenstate, then after decoherence there are no branches in which the corresponding measurement result obtains. It is therefore natural to assign a probability of zero to that event – the outcome does not happen. The same cannot be said of any event with nonzero weight. And it will not do to say that any such event approximately does not happen. It either happens (i.e., there exist branches in which it obtains) or it does not. The same reasoning that allows us to derive zero probability from zero weight is entirely inconclusive when applied to near-zero weights.

One might still complain that any (anthropocentrically) discontinuous relationship between weight and probability would make quantum mechanics empirically useless, but this is hardly a telling criticism of an argument which attempts to show that the Everett interpretation does not provide probabilistic predictions in the first place. To suppose that the continuity assumption holds of all possible scientific theories is clearly mistaken – we can

easily imagine theories that fail to satisfy this condition. And to suppose that it holds of all empirically satisfactory theories is beside the point. The continuity assumption can then be assumed to apply only if Everettian QM is a satisfactory theory, and to assume Everettian QM is empirically satisfactory is to beg the question against my argument.

4.3   A holistic approach?

If I am right, Everettian decision theory depends for its justification on circular reasoning. There is a natural response open to the Everettian: embrace the circle. Plausibly, not all circular arguments are vicious, in the sense of undermining our justification for their conclusions. Couldn't the interpretation make do with a holistic approach, employing the decision-theoretic probabilities, once derived from the QRT, as antecedent justification for their talk of decoherence?

The idea here is that a probability rule, once established, can identify which worlds are the branches even if its proof depends on a specification of the branches. To see why, take a universe governed by Everettian quantum mechanics in which we begin by assuming the Born rule. From such a standpoint we can establish via decoherence which worlds are the branches, and then prove via the QRT that our probability rule is correct. A probability rule in Everett is nothing more than a rationality principle (as opposed to some matter of physical or ontological fact). Thus it may be that anyone who starts out employing the rule, and can prove to his satisfaction that it is justified, is rational.

First, an uncharitable reply: I find it hard to be convinced by an argument that includes its conclusion among its premises! I do not mean to endorse an epistemic foundationalism that rejects any possibility of circular justification. Rather, I suggest that we are rarely (if ever) justified in adopting a circular set of beliefs, none of which we currently hold. Yet this is

exactly what the holistic Everettian is urging, at least when evangelizing the non-Everettian.

Let me try to engage more thoroughly with this holistic approach by bringing to light one further reason for concluding that the circularity I've identified is a vicious one. Both the QRT and related primitivist approaches to Everettian decision theory aspire to solve two distinct problems: the *quantitative problem* (what is the quantitative rule for assigning probabilities to outcomes?) and the *coherence problem* (what *are* probabilities on the Everett interpretation?). This distinction makes it possible to formulate in a new way the circularity in Everettian decision theory. Specifically, any decoherence-based Everettian solution to the coherence problem must depend for its justification upon a solution to the quantitative problem, and vice versa.

Any solution to the qualitative problem should amount to an analysis of the concept of probability. A probability measure is, in general, a measure over a certain *sample space*, that is, a set of possible outcomes. In statistical mechanics, for example, the sample space is given by the set of possible initial conditions for a system; likewise for Bohmian mechanics. In these cases, probability is analyzed as a measure of our degrees of belief about the initial conditions. In a stochastic theory like a collapse interpretation of quantum mechanics, sample spaces are sets of possible measurement outcomes, where these are complete physical states of affairs that might result from a measurement. Probability is then a measure of the objective chance of each possible outcome. Note that in each of these cases, a specification of necessary and sufficient conditions for membership in the sample space is central to the analysis of probability for a theory.

What about in the case of Everett? Certainly the account proposed in Section 3 is, in some sense, an analysis of probability within the framework of the interpretation: probability is a measure of our care or self-locating belief about branches. But it is not clear that

this analysis provides necessary and sufficient conditions for membership in a sample space in the absence of a probability rule. As I argued in Section 4.2, a complete description of an Everettian quantum state does not specify the measurement outcomes (that is, the branches) unless it is supplemented by a quantitative probability rule. This requires an appeal to decoherence *and* to a quantitative rule which entails that low-weight outcomes are improbable. But the sample space for an experiment just is the set of outcomes. So, without a prior solution to the quantitative problem, the Everettian analysis of probability cannot fully describe the conditions under which an element of the state is a member of a sample space.

I think it plausible that a complete solution to the qualitative problem must include a full specification of the conditions under which a world should be included in the set of outcomes. Everettians may contest this point, but suppose that it is correct (it is, at least, satisfied by other major interpretations of QM). Then an Everettian solution to the coherence problem can only be justified by a solution to the quantitative problem. Solving the quantitative problem obviously requires appeal to a coherent concept of probability, so any such solution must likewise be justified by a solution of the coherence problem.

It appears that the circularity previously identified in the QRT actually encompasses any Everettian attempt to offer a coherent notion of probability along with an explanation of the Born rule. Primitivist approaches to probability like that endorsed by Greaves (2004) appeal to the same decision-theoretic solution to the coherence problem invoked by the QRT. Thus the criticism of circularity undermines these approaches as well as the Deutsch-Wallace proof.

Further, this dependence of the Everettian concept of probability on the quantitative rule governing its apportionment calls into question the methodological soundness of the

interpretation. This is one sense in which the circularity can be seen as vicious. Everett does away with the notion – to which common sense, if nothing else, commits us – that we can advance a correct concept of probability (or at least a full set of criteria for membership in the sample space for an experiment) without appeal to any facts about which events are probable or improbable. Other interpretations of quantum mechanics, for example, appeal to notions of objective chance or epistemic credence which can be defined and understood independently of the Born rule. To the extent that we believe probability should be so definable, so that knowing which is the correct concept of probability presupposes no knowledge of which particular events are probable, we should favor these interpretations over Everett.

The Everettian may legitimately claim at this stage that we have no strong pre-theoretic commitment to the independence of the quantitative and qualitative problems. On this view, it has been shown only that the correct concept of probability and the correct rule governing its distribution must be inter-defined in the quantum case. But neither the bare fact of this circularity nor the interdependence of the qualitative and quantitative solutions is an indictment of the Everett approach if we simply accept these as among the unique features of the interpretation. Once this step is taken, there is little left to argue. I have shown that these are consequences the Everettian must accept; I claim they are unacceptable.

5   Conclusion

Defenders of the Everett interpretation have so far failed to ground their decision-theoretic approach to probability on a viable notion of measurement outcomes. I have shown that measurement outcomes in Everett must be classes of branches. Once this is understood, the circularity inherent in the decision-theoretic approach is not difficult to see. For nothing

about the bare theory of quantum mechanics guarantees that there *are* any branches – considerations from decoherence only establish that branching behavior is extremely likely. An appeal to some probability rule is therefore required to determine which components of the state are measurement outcomes, but a specification of the measurement outcomes is itself needed to tell us which components of the state can be bearers of probability. Unless this circularity is accepted, the decision-theoretic program cannot get off the ground. Talk of an Everettian representation "theorem" would thus appear to be deceptive. Everettians are not asking us to accept the conclusions of a theorem, but rather to buy into a circular picture of the quantum world that can only be justified from within. Even granting their claim that our world is fully described by the quantum state, a further step – that of entering into the circle – is required before we can accept their decision-theoretic notion of probability.

So, despite the impressive progress that has been made in clarifying the Everettian program during this decade and the last, it is still not clear that the interpretation can legitimately employ either the concept of probability or any specific quantitative rule for assigning probabilities. Nor is it clear that a solution to the problems I've posed will vindicate the decision-theoretic approach to probability. The most significant goal of such an approach is, as noted earlier, 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,[15] could ever be confirmed or disconfirmed by experimental results. This requires a leap from 'ought' to 'is' – from the preferences we should possess to the outcomes we will experience – that needs to be justified by further argument.[16]

---

[15]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 SC.

[16]In other words, it remains to be seen whether the decision-theoretic account of probability, even if successful, makes the Everett interpretation *empirically coherent*. See Barrett (1996; 1999, 116-117) for

Acknowledgements

References

[1] Barnum, H., C.M. Caves, J. Finkelstein, C.A. Fuchs and R. Schack (2000). Quantum Probability from Decision Theory? *Proceedings of the Royal Society of London* A456: 1175-1182.

[2] Barrett, J. (1996). Empirical Adequacy and the Availability of Reliable Records in Quantum Mechanics. *Philosophy of Science* 63: 49-64.

[3] Barrett, J. (1999). *The Quantum Mechanics of Minds and Worlds.* Oxford: Oxford UP.

[4] Butterfield, J. (2002). Some Worlds of Quantum Theory. In R.Russell et al (ed.), *Quantum Physics and Divine Action,* 111-140. Vatican Observatory Publications.

[5] Deutsch, D. (1999). Quantum theory of probability and decisions. *Proceedings of the Royal Society of London* A455: 3129-3137.

[6] Greaves, H. (2004). Understanding Deutsch's Probability in a Deterministic Multiverse. *Studies in History and Philosophy of Modern Physics* 35: 423-456.

[7] Joos, E., C. Kiefer, J. Kupsch, I. Stamatescu, H. Zeh and D. Giulini (1996). *Decoherence and the Appearance of a Classical World in Quantum Theory.* Springer.

[8] Lewis, D. (1976). "Survival and Identity," reprinted in *Collected Papers Vol. I.*

[9] Lewis, P. (2002). What is it like to be Schrödinger's cat? *Analysis* 60: 22-29.

[10] Lewis, P. (2006). Uncertainty and probability for branching selves. Forthcoming in *Studies in History and Philosophy of Modern Physics.* Available online at http://philsci-archive.pitt.edu/archive/00002636/.

---

more on empirical incoherence.

[11] Parfit, D. (1984). *Reasons and Persons.* Oxford: Oxford UP.

[12] Saunders, S. (1998). Time, Quantum Mechanics and Decoherence. *Synthese* 102: 235-266.

[13] Tegmark, M. (1998). The Interpretation of Quantum Mechanics: Many Worlds or Many Words? *Fortschritte Physik* 46: 855-862.

[14] Wallace, D. (2002). Quantum Probability and Decision Theory, Revisited. Available online at http://www.arxiv.org/abs/quant-ph/0211104/.

[15] Wallace, D. (2003a). 'Everettian Rationality. *Studies in History and Philosophy of Modern Physics* 34: 87-105. This paper is a shorter version of the above.

[16] Wallace, D. (2003b). Quantum Probability from Subjective Likelihood: improving on Deutsch's proof of the probability rule. Available online at http://philsci-archive.pitt.edu/archive/00001523/.

[17] Wallace, D. (2005). "Epistemology Quantized." Forthcoming in *British Journal for the Philosophy of Science.* Available online at http://users.ox.ac.uk/ mert0130/papers/epist.pdf.

[18] Zurek, W.H. (2003). Decoherence, einselection and the quantum origins of the classical. *Reviews of Modern Physics* 75: 715-775.