The Many Worlds Interpretation of QM: A Hylomorphic Critique and Alternative

Robert C. Koons

October 19, 2016

1 Introduction

2 Wallace's Oxford-Style Everett Interpretation

2.1 The Problems of Probability

	P	Q	Q only	P and Q
A	$\langle 1,1\rangle$	$\langle 1,1\rangle$	2	4
B	$\langle 1,1\rangle$	$\langle 0.6,1.8\rangle$	$2.4 - \frac{(0.6+0.6)}{2} = 1.8$	$4.4 - \frac{(0.1+0.1+0.5+0.7)}{4} = 4.05$

2.2 Recovering the Manifest Image through Ramseyfication

In this paper, I will use a model-theoretic version of Ramseyfication, in which, instead of introducing second-order variable and quantifiers for the predicates, we simply extend the interpretation function of a given model in order to turn a model of the original, base language into a model of the emergent theory in an appropriately expanded language. Since we are interested in providing a model for the functional or dispositional features of the emergent theory, we will need to have something like subjunctive or counterfactual conditions (i.e., 'this would happen if that happened') already in the base language. This will require use of possible-worlds semantics, which requires a model *frame* (consisting of a set of worlds and a set of sets used to interpret subjunctive conditionals) plus an interpretation, including both a domain of individuals and an interpretation function that assigns individuals to constants and sets of n-tuples of individuals to n-ary predicates in the language.

1. A model frame F consists of a set of a set of worlds W, a designated actual world $w^* \in W$ and, as in David Lewis's semantics, a system of spheres S, consisting of nested subsets of W, centered on $w*$. I will assume that the sphere-system S is dense (between any two concentric spheres there is always a third), and that the number of worlds in every sphere-membership equivalence class is equal to the cardinality of the powerset of the set of atomic formulas of the language.

2. A model M consists of a model frame F plus a domain of worldbound individuals D and an interpretation function I.
3. The set of individuals D is partitioned into disjoint cells, one for each world in W. We can think of D as a function from W into a set of disjoint sets, with $D(w)$ designating the worldbound individuals of world w.
4. For any n-ary predicate F, $I(|F|)$ is a function whose domain is W, and for each world $w \in W$, $I(|F|)(w)$ is a subset of the n-ary cross-product of $D(w)$.
5. For any constant c, $I(|c|)$ is a function whose domain is W, and for each world $w \in W$, $I(|c|)(w)$ is a member of $D(w)$.
6. For any atomic sentence $F(c_1, c_2, ..., c_n)$, $I(|F(c_1, c_2, ..., c_n)|)$ is a set of worlds in W, where each world w belongs to $I(|F(c_1, c_2, ..., c_n)|)$ if and only if the n-tuple $\langle I(c_1)(w), I(c_2)(w), ..., I(c_n)(w)\rangle$ belongs to $I(|F|)(w)$.
7. $I(|(\phi \,\&\, \psi)|) = I(|\phi|) \cap I(|\psi|)$, and similarly for the other sentential connectives.
8. $I(|\exists x \phi(x)|)$ = the infinite union of the sets $I(|\phi(c)|)$, for each constant c in the language L. (We'll assume that the language L has been enriched with enough constants to provide a witness for every existential generalization true in M.)
9. $I(|(\phi \,\Box\!\!\rightarrow \psi)|) = W$ if tthere is an $I(|\phi|)$-permitting sphere s in S such that every world in $I(|\phi|) \cap s$ is also in $I(|\psi|)$). Otherwise $I(|(\phi \,\Box\!\!\rightarrow \psi)|) = \emptyset$.

A model $M = \langle F, D, I\rangle$ is a model of a theory T just in case, relative to I and D, the actual world w^* belongs to $I(|T|)$, where $I(T)$ is the intersection of the sets $I(|\phi|)$, for each formula ϕ in T. As usual, a theory is defined as a set of formulas closed under logical implication.

Let's suppose that we start with a model $M_{base} = \langle F, D, I\rangle$, defined for our base language L_{base}, which represents the fundamental level of reality. Now suppose that we extend the language L_{base} to a language $L_{base+emergent}$, by adding constants and predicates that signify an emergent level. A theory $T_{emergent}$ of this emergent world is realized in our base model M just in case the interpretation function I can be extended to a new function $I_{realizer}$, defined for $L_{base+emergent}$, such that the model $M_{extended} = \langle F, D, I_{realizer}\rangle$ is a model of $T_{emergent}$. In such a case, we can say that the function $I_{realizer}$ is a *realization* of the emergent theory $T_{emergent}$ in the original base model M_{base}. This model-theoretic version is a generalization of Ramsey's original idea (1927).

2.3 Classical phenomenalism and Russell's structuralism

We c select the model $M_{true-phen}$ that incorporates all the actual truths about actual and counterfactual experiences. The set of theories true in $M_{true-phen}$ is the set $TRUE_{phen}$, the set of all truths expressible in the vocabulary of L_{phen}. We now enrich the language by adding terms referring to physical objects, which will now be assigned locations and trajectories in a single three-dimensional (public) space, indexed by universal time.

- Call the resulting language $L_{phen+phys}$.
- Consider each theory expressible in $L_{phen+phys}$ that is consistent with the set of phenomenal truths, $TRUE_{phen}$.
- Let T_0 be one such a theory.

- Since T_0 is consistent with the set of phenomenal truths, we can extend the interpretation function I_{phen} to a function $I_{phen+phys}$ in such a way that theory T_0 is true in the model $M_{true-phen}$ relative to $I_{phen+phys}$.

In general, there will be many realizations of any theory T_0 in the model $M_{true-phen}$, and there will be many other theories in the enriched language besides T_0 that are consistent with the set of all phenomenal truths (and which therefore have realizations in $M_{true-phen}$). In order to cut down the number of theories and realizations, we need some further constraints both on our theory T_0 and on the permissible realizations of that theory. We can accomplish both of these at once simply by restricting the interpretation function. We can then hope to pick out the one true theory of physics that has a unique permissible realization in the model $M_{true-phen}$. In the case of both the phenomenalist and Russellian-structuralist program, these constraints consist in the *laws of perspective* that link geometrical properties described in terms of public four-dimensional spacetime with properties described in terms of egocentric phenomenal space and time.

2.4 Analytic Functionalism about the Mind

The true model of the world $M_{true-phys}$ yields a set of physicalistically acceptable truths, $TRUE_{phys}$ in a language of purely physical (and chemical, biological, and neurological) vocabulary L_{phys}. We want to extend this language to a language $L_{phys+psy}$ that includes the vocabulary of psychology, with predicates that assign beliefs, desires, and sensory experiences to the class of human beings. Lewis assumes that we are already given, not only the vocabulary of $L_{phys+psy}$, but also a fairly rich theory of *folk psychology* T_{folk}, that specifies a large number of connections between psychological and physical states. This will include facts about the sensory experiences resulting from sensory-organ stimulations, coordinated in such a way that experiences are veridical under normal conditions. It will also include connections between belief-desire pairs and overt behavior, and certain kinds of overt behavior that results directly from certain experiences or desires, like wincing from pain.

- Let's assume that T_{folk} is consistent with the set of physical truths, $TRUE_{phys}$.
- If so, we can find an interpretation function $I_{phys+psy}$, relative to which T_{folk} is true in the true model of the physical world, $M_{true-phys}$.
- If there is such a function, it will be a *realization* (in Ramsey's sense) of the folk theory of psychology.
- If there is a unique such function, then we can use it to define the set of *all* psychological and psychophysical truths bysimply identifying it with the set of sentences $TRUE_{phys+psy}$ in the language $L_{phys+psy}$ that are verified by the model $M_{true-phys}$ as extended by the interpretation function $I_{phys+psy}$.

Lewis is entitled to help himself to the psychophysical language $L_{phys+psy}$ and the folk theory T_{folk} in that theory, since the facts about what language humans speak and what sentences in that language they assert can be recovered with a high degree of determinacy from the physicalistically acceptable set of facts, simply by consulting users' overt verbal behavior (including their counterfactual behavior under all possible circumstances). This is the sort of task that Donald Davidson described as *radical interpretation*. In addition to or as an alternative to reliance on the folk theory T_{folk}, we could rely, as Donald Davidson recommended, on a Principle of Charity, which could serve as a constraint on acceptable interpretation functions. We could require that the interpretation of sentences that attribute the belief with content ϕ be assigned intensions in which ϕ is also verified, at least to as great an extent as possible.

2.5 Wallacian Functionalism

We can take the language of pure quantum mechanics (with its description of the cosmic wavefunction and its deterministic Schrödinger evolution) and supplement it with a counterfactual or subjunctive conditional. This will require a model that contains a domain of worlds, each of which consists of a single quantum wavefunction evolved through time, one world designated as actual (which picks out the world's actual wavefunction), and a set-selection function $COND$ for the evaluation of subjunctive conditionals. The system of spheres could be based, as in David Lewis's semantics, on a relation of comparative similarity between quantum worlds.

So, let's turn now to the emergent domain. Our first problem is a very basic one: What language do we use, and what theory in that language? It seems that we must use *every* possible language and *every* possible theory. There are no languages or theories and no language users or believers explicit at the level of quantum reality. Any constraints we place on these theories (besides their sheer interpretability in the model of quantum mechanics) are going to be constraints of internal coherency. We might reasonably demand of any theory $T_{emergent}$ of the emergent world that, according to $T_{emergent}$ itself, the human beings speak the language of $T_{emergent}$ and have beliefs and sensory and mnemonic experiences that mostly accord with $T_{emergent}$. We can also require that $T_{emergent}$ have the theoretical virtues valued by most people (as depicted in $T_{emergent}$), and that $T_{emergent}$ be well-confirmed, according to itself. Call the theories that meet these constraints the *internally ideal* or *coherent* theories.

2.6 Putnam's Permutation Argument for Semantic Indeterminacy

I will argue, in a way inspired by Putnam's argument for anti-realism , that there is a radical indeterminacy of meaning for all the names and predicates of the languages of our emergent theories. This isn't surprising, since in Wallace's picture, all that matters is that we find an interpretation of those theories in the true model of quantum mechanics that makes all of the formulas of that theory come out true (or at least approximately true) under that interpretation. The constraint on the meaning of the emergent theories is utterly holistic in character.

Suppose that $T_{emergent}$ is a theory of a world that is emergent relative to the model $M_{true-QM} = \langle F, D, I \rangle$. That means that there is an interpretation function, call it $I_{intended}$ that extends I to the language of $T_{emergent}$, resulting in a new model $M_{QM+emergent} = \langle F, D, I_{intended} \rangle$, with the theory $T_{emergent}$ true in $M_{QM+emergent}$. It is obvious that there are an infinite number of alternative extensions of I that will also produce an extension of $M_{true-QM}$ relative to which $T_{emergent}$ is true. Take any permutation $\pi(w)$ for any world $w \in W$ of the objects in $D(w)$. Now apply the permutation $\pi(w)$ to the interpretation $I_{intended}$ with respect to the interpretation of all constants and predicate symbols at w. The resulting interpretation $I_{intended-\pi(w)}$ will also be a realization of $T_{emergent}$. Apply similar permutations to every world in W, resulting in the thoroughly scrambled interpretation $I_{bizarro}$. The extension of $M_{true-QM}$ by $I_{bizarro}$ will also be a model of $T_{emergent}$, and so $I_{bizarro}$ will be a realization of $T_{emergent}$ in $M_{true-QM}$.

As Alexander Pruss has pointed out, all the predicates that apply truthfully to the emergent world as it exists today (including mental-property predicates) could be interpreted in such a way that they apply truthfully only to the cosmos as it was 12 billion years ago.

Any two worlds that are isomorphic under an isomorphism of the quantum structure (i.e., of the Hilbert spaces and the operator algebras) have the same functional properties. Now consider two worlds w_1 and w_2. Both are short-lived worlds: the temporal sequence of each is only a billion years old. Each world is an exact duplicate of a temporal portion of our world. Thus, w_1 is an exact duplicate of the temporal portion of our world from 13 billion years ago to 12 billion years ago, while w_2 is an exact duplicate of the temporal portion of our world from a billion years ago to the present. Then w_2 has the same kind of mental properties that obtained in our world over the last billion years. And w_1 has the same kind of mental properties that obtained in our world from 13 to 12 billion years ago. But there is a quantum-structure preserving isomorphism from w_1 to w_2. This isomorphism is simply given by the time-evolution operator U_{12} (where we measure time in billions of years). This operator is an isomorphism of the quantum structure. Hence w_1 and w_2 are exactly alike with respect to mental properties. Hence our world had exactly the same mental properties in the early 13-to-12 billion-years-ago period as in the last billion years. That's absurd. (For one, it makes us question how we could possibly know that the world is as old as we think it is.)[**?**]

2.7 Model-Theoretic Indeterminacy Guarantees Truth of our Emergent Theories

Let $T_{emergent}$ be one of our target theories of the world: folk psychology or a scientific theory of "emergent" phenomena. We can suppose that $T_{emergent}$ is internally ideal and that it has a realization in the model of quantum mechanics, $M_{true-QM}$. Let $I_{intended}$ be the "intended" interpretation of the theory $T_{emergent}$ in the model $M_{true-QM}$, with a domain consisting of the spacetime regions and quantum subsystems of the quantum world and with the predicates of the language $L_{emergent}$ assigned appropriate intensions in the corresponding model $M_{true-QM}$.

Now consider a theory $T_{bizarro}$, whose intended model includes the same interpretation function $I_{intended}$ but includes a different, counterfactual model of the quantum world, $M_{counterfactual-QM}$. Both $M_{true-QM}$ and $M_{counterfactual-QM}$ have infinite domains, and so both T_{emerge} and $T_{bizarro}$ are consistent with the hypothesis of a domain of infinite cardinality. By the Skolem-Löwenheim theorems, there is an interpretation $I_{bizarro}$ of $T_{bizarro}$ in the actual model of the quantum world, $M_{true-QM}$. Thus, the bizarro emergent world represented by $T_{bizarro}$ is realized in the actual quantum world in just the same way as $T_{emergent}$ is.

In fact, *all possible theories of emergent domains are actually true*: if they are logically consistent (in the logic of quantified counterfactual conditionals), and they contain no quantum-mechanical vocabulary and make no claims about the finite size of reality, then (by the Skolem-Löwenheim theorems), they have a model that extends $M_{true-QM}$.

In fact, the situation is even worse than this, since Wallace doesn't require perfect realization in $M_{true-QM}$—just a reasonable degree of approximation to such perfect realization. So, even *inconsistent theories* or theories that entail the existence of a finite domain or that entail falsehoods about the structure of spacetime will nonetheless have quantum realizations and so will be *actually true* theories of a world that emerges from the quantum world.

The upshot is this: we are free to believe and say *whatever we want* about the emergent world of macroscopic objects, and we are guaranteed to believe and speak the truth (so long as our stories are internally coherent and not massively inconsistent).

2.8 Epistemological and Pragmatic Consequences

If classical physics is understood as true by stipulation, this undermines any rational confidence we might have in quantum mechanics, since a large part of our evidence for QM consists in its agreement with classical mechanics when interference terms are small. In fact, Wallace's functionalism leads quickly to an epistemological catastrophe: if we cannot interpret our theories of the emergent world realistically, then no belief in such a theory can count as objective knowledge. And yet all of our knowledge of the truth of quantum mechanics depends on our having objective knowledge of experimental data that belongs to an emergent domain.

In fact, we couldn't even interpret our emergent scientific theories as *instrumentally* valuable in an objective way, since *any* theory of our future experiences would be equally true (just one more realizable emergent theory). In addition, what counts as *the same* qualitative properties of experience is itself up for grabs via the interpretation function. By choosing a suitable function, we can make any set of predictions about future experience come out true. Thus, pragmatism itself is inconsistent with radical indeterminacy of meaning, as Plato recognized in the *Theaetetus*:

> SOCRATES But, Protagoras (we'll say), what about the things which are going to be, in the future? Does he [the individual human being] have in himself the authority for deciding about them, too? If someone thinks there's going to be a thing of some kind, does that thing actually come into being for the person who thought so? Take heat, for example. Suppose a layman thinks he's going to catch a fever and there's going to be that degree of heat, whereas someone else, a doctor, thinks not. Which one's judgment shall we say the future will turn out to accord with? Or should we say that it will be in accordance with the judgments of both: for the doctor he'll come to be neither hot nor feverish, whereas for himself he'll come to be both?
>
> THEODORUS. No, that would be absurd. (*Theaetetus* 178c1-10)

Every claim about the future, practical consequences about believing and acting on an emergent theory of the world will itself be part of some emergent theory of the world. I have shown that every such theory, so long as it is not massively inconsistent and doesn't entail the finitude of the universe, will be realizable in $M_{true-QM}$ and so will be true. Thus, we cannot appeal to pragmatic considerations (like avoiding being eaten by a tiger) as grounds for preferring some theories over others.

2.9 The Argument's Upshot in a Nutshell

- Radical indeterminacy of content, via Putnam's paradox: there an infinite number of alternative interpretation functions mapping our actual theory of the world into M_{QM}. In particular, there is no fact of the matter as to the quantum probability associated with any given emergent world.
- Every consistent and internally coherent story (more precisely, every story consistent with an infinite domain) represents an emergent reality in M_{QM}, on a par with our current best theories about the macroscopic world.
- So, we can't go wrong in proposing theories about the emergent world we inhabit, so long as our theories are consistent with an infinite domain, and so long as they are internally coherent from a semantic and epistemological point of view.

- These facts undermine any claim to know that quantum mechanics is true, on the basis of experiments and observations that depend in any way on the emergent. There are equally real emergent worlds in which the available data flatly contradicts quantum theory. Therefore, Wallacian functionalism is epistemologically self-defeating.

3 Putnam's Paradox: The Problem of the Missing External Constraints

3.1 Can the concept of *emergence* fix the interpretation?

3.2 Can causal constraints fix the interpretation?

3.3 Can we use "our" actual language and folk theory?

3.4 Why not simplicity?

There are two independent parameters of simplicity to consider: (1) the simplicity of the *theory* of the emergent world (is it, for example, finitely or recursively axiomatizable, or at least approximately so?), and (2) simplicity of the *interpretation function* that interprets the non-quantum vocabulary in the quantum model.

Maximizing simplicity

We can't maximize simplicity of either theory or interpretation, since the simplest possible interpretation function is just the original interpretation function of the quantum model (with no addition), and the simplest possible theory is the just theory of the original quantum model (the totality of purely quantum truths, including the theorems of logic).

Setting a minimal degree of simplicity

- Any requirement of relative simplicity will have to be quite loose and permissive, since we know that the entities and properties of the emergent manifest image are, under the most optimistic assumptions, far from natural. Neurological and other biological properties will be highly disjunctive and gerrymandered from the viewpoint of fundamental quantum mechanics, and phenomenological, intentional, and semantic properties even more so. If the requirement of simplicity is too strong, it would give us *no* emergent worlds at all; if too weak, it would give us far *too many.*
- What could be the truthmaker or metaphysical ground of the correct standard of minimum simplicity? How are we supposed to explain the connection between complexity and unreality?
- There has to be some counterweight to simplicity, or we should embrace an eliminativist theory (a no-emergence theory), in which our theory of the "manifest" world just is fundamental quantum mechanics. We need some reason not to set the standard at the maximum level of simplicity. So, what is the counterweight? Usefulness? Apparent truth? But these criteria only make sense given a manifest theory. We need people, organisms, perceptions, beliefs, purposes, etc. in order to make these judgments.

Degrees of reality

3.5 Can we appeal to natural or eligible properties?

3.6 Can we use spacetime to restrict the interpretation function?

3.7 What about decoherence?

How is that relevant to our theories of the emergent world? It's relevance seems to depend on two assumptions:

1. The dynamics of the emergent world must be (approximately) that of classical (Newton-Maxwell) mechanics.
2. The dynamics of the emergent world should closely mimic those of the underlying quantum reality.

Of course, if we keep loading up conditions on a *proper* emergent world, we will eventually isolate the theory we want. The following five conditions might work:

1. Extend the model $M_{true-QM}$ of quantum mechanics to a model $M_{QM+branches}$ with a branch parameter, each branch being assigned a probability weight, a period of time, and a set of particles in each world.
2. Privilege the basis consisting of position and momentum by having the extended model $M_{QM+branches}$ assign definite but branch-relative position and momentum to each particle that belongs to the branch and at each time assigned to the branch.
3. Require that the sum of branch-probabilities corresponding to a set of particle positions (or momenta) be a good approximation to corresponding sum of probability amplitudes in the original model $M_{true-QM}$.
4. Require that the dynamics assigned to particles by branches approximate a dynamic theory that is both simple and relies on highly localized, separable quantities (i.e., very like classical mechanics) .
5. Require that the branch structure include as many particles and as much time as possible, given the other constraints.

We can then stipulate that the only real emergent worlds correspond to the set of truths verified by some such extended model $M_{QM+branches}$. In addition, we could count as an emergent theory a theory that is expressed in a reduced language of $L_{QM+absolute}$, reduced by the replacement of each branch-relativized predicate with an absolute version of the same predicate (including location and momentum predicates). A theory $T_{absolute}$ in the reduced language could count as *realized* by $M_{QM+branches}$ just in case there is a consistent assignment of branch-parameters to the formulas of $T_{absolute}$ results in a theory that is verified by $M_{QM+branches}$.

However, such a move has several disadvantages:

- The account is no longer tied to and no longer provides a *general theory* of emergence. Consequently, we would have to deny the emergent reality of other special sciences, like chemistry, thermodynamics, biology, psychology, and the social sciences.
- We would be offering no account of why these conditions are of great *metaphysical* significance? Why must an emergent world satisfy just these conditions to count as real?

- We would give up the claim that decoherence generates the privileged basis by itself. Instead, we would simply be stipulating what we shall count as the correct basis. If we try to get around this by deleting conditions (2) and (3), we will be unable to dissolve the Putnam paradox, since permutations of the intended model of the emergent world will meet the other three conditions. This would leave Pruss's time shift argument, with the superfluity of minds in obviously mindless regions, untouched.
- We would be making the many-worlds or many-branches structure of emergent reality true by stipulation.

4 The Solution: Real Essences and Extra-Conceptual Grounding

4.1 Two Forms of Grounding

We enrich our base model $M_{ture-QM}$, representing fundamental reality, by supplementing it with a set of natural-kind essences K and a fundamental composition relation $COMP$. The new model, M_{QM+HM} (HM for 'hylomorphism') would be defined over a language that contains constants for each of the natural substance-kinds in K along with a four-place *part-of* predicate P, where $P(k, p_1, p_2, t)$ represents the fact that both particles p_1 and p_2 are at time t proper parts of a substance of kind k. The truth-conditions for the P predicate will be given by the fixed $COMP$ relation in the model, with the stipulation that, for fixed time t, each particle can be part of at most one substance. .

The natural kinds will make a real difference by virtue of constraining acceptable models to connect substances of each kind with appropriate branches in the branching-extension of M_{QM} defined by the five conditions in section 3.7 above.

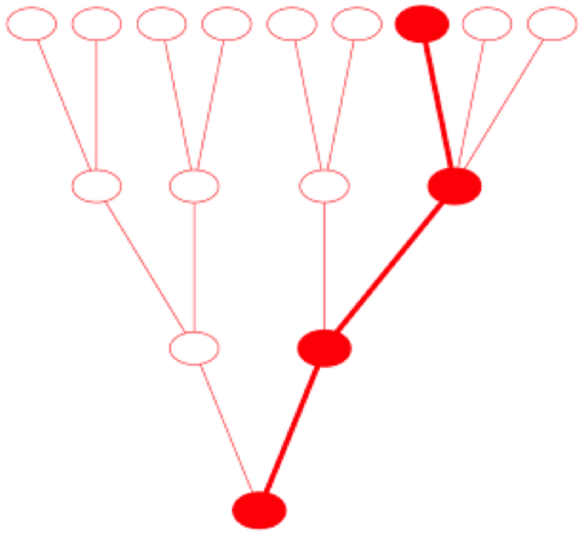
> M_{QM+HM} is an **acceptable model of the emergent world** if and only if there is a branch extension of the base model $M_{QM+branches}$ meeting the five conditions in 3.7 such that, for any kind k in K, for any world w in W, for any particles p_1 and p_2 and time t in $D(w)$, if $M_{QM+HM} \models P(k, p_1, p_2, t)$, then there is a branch b in $M_{M+branches}$ such that p_1 and p_2 belong to b in w at t, and the tuple $\langle p_1, P_2, t, b\rangle$ satisfies all of the metaphysical conditions associated with natural kind k.

The appeal to natural kinds of substantial forms enable us to overcome the four problems we identified at the end of section 3.7:

- The account is a general theory of emergence. Every emergent domain depends on an appropriate set of natural kinds (macrophysical, thermodynamic, chemical, biological, etc.).
- It is the fundamental existence of the emergent natural kinds that lends metaphysical significance to the constraints.
- The essences of the natural kinds select the privileged basis of operators, by being requiring a range of values for the corresponding parameters.
- The essences of the natural kinds provide the ground for the truth of the multiple-branch structure within the wavefunction, since each essence requires the existence of a branch of an appropriate kind for its actualization.

5 Extra Bonus: Restoring the Real World's Unity

> On the physical side, this interpretation is just like the many-worlds interpretation. It is a dualist interpretation like the many-minds one: we each have a non-physical mind. But there is only one mind per person, as per common sense, and minds never split. Moreover our minds are all stuck together: they always travel together. When we come to a branching point, the physical world splits just as on the many-worlds interpretation. But the minds now collectively travel together on one of the outgoing branches, with the probability of the minds taking a branch being given by the indeterministic theory. (Pruss 2014)



Traveling minds

5.1 Traveling Forms and Ontic Vagueness

6 Second Bonus: Solving the Problem of Anti-Darwinian Branches

7 Third Bonus: Grounding Scientific Inference in Localized Causal Powers

8 Conclusion