A Causal Agent Quantum Ontology

Kathryn Blackmond Laskey

Systems Engineering and Operations Research Department
George Mason University
Fairfax, VA 22030-4444
klaskey@gmu.edu

Abstract

Quantum theory is widely regarded as the most successful scientific theory to date. As a foundational physical theory, its range extends from the sub-microscopic to the interstellar, excluding only phenomena for which gravitational spacetime curvature cannot be neglected. Even its most surprising empirical predictions have been borne out to extraordinarily high accuracy. Yet the ontological status of quantum theory remains a source of controversy. As science extends ever more deeply into the biological and social domains, there is a pressing need for a foundational scientific theory with an explicit place for cognition and agency. This paper describes a *causal agent* ontology for quantum theory. In the causal agent ontology, subsystems called *agents* make choices, unconstrained by the laws of present-day physics, that have a causal impact on the future evolution of the physical world. The causal agent ontology is in precise accordance with von Neumann's formulation of quantum theory. In von Neumann quantum theory, physical systems evolve by two distinct kinds of process: deterministic, continuous mechanical evolution, and discrete stochastic transitions. Given a specification of which kinds of transitions occur at what times, quantum theory makes highly accurate probabilistic predictions for the outcomes of the transitions. However, there is no physical theory governing the choice of which kinds of transitions occur at what times. Several authors have suggested that this choice be identified with free choices of agents. Although controversial, this identification coheres with how quantum theory is applied in practice, and has led to predictions that agree with experimental results. This paper presents the causal agent ontology, discusses some of its implications, and points to evidence from perception experiments that show good agreement with predictions based on a causal agent theory.

1. Introduction

The relationship between the external world and the minds that study the world has puzzled humankind since the dawn of science. The course of Western thought was profoundly affected by Rene Descartes' articulation of the distinction between mind and matter. Since then, the prevailing metaphysical stance in the West has been that the material universe evolves autonomously, follows a lawful dynamics independent of our thoughts, and can be described by empirically testable mathematical theories. The scientific revolution has been profoundly successful at describing those aspects of the world we label material. As a result, our ability to manipulate the material world through technology has exploded. The hypothetico-deductive approach and commitment to empirical evaluation that mark the scientific attitude have moved beyond the purely material to the biological and social sciences. The computer revolution has raised the possibility that intelligence itself can be understood scientifically, formalized, and engineered into physical devices.

Despite enormous practical success, science remains unclear about how the mind that formulates and understands scientific theories, designs and conducts experiments to test its theories, and acts on the basis of those theories, is related to the material world it studies. The past century has brought broad appreciation of the statistical regularities underlying the seeming complexity of biological and social phenomena (e.g., Gigerenzer et al., 1990). The computational metaphor has generated important insights into how cognition functions, enabling the automation of many "knowledge tasks" once thought to be the exclusive province of human cognition. Advances in neuroscience have identified the neurological correlates of many cognitive functions, enabling concomitant advances in medicine and education. However, current theories of intelligent behavior are founded on an outmoded model of physical systems as deterministic automatons. As the reach of science expands into the cognitive domain, this worldview fuels a perception of our minds as passive spectators watching our robotic bodies execute their built-in programs. Public policies that shape the context in which private decisions are made are increasingly based on a scientific approach to evaluating their likely effects. But the analysts and decision makers who practice science-based public policy operate within a scientific paradigm that has no role for efficacious conscious choice. As human activity begins to threaten our very survival, the question of the appropriate role in scientific theories for free will and deliberate choice takes on increased urgency.

There is a pressing need for a unified science that encompasses both the mental and material aspects of universe. In this unified scientific worldview, the universe evolves in a lawlike manner, yet its evolution is contingent on the choices made by *agents*. Agents are physical systems that can act on the world by choosing from among a set of physically allowable options. Although physical law constrains the options available to agents, within these constraints their choices are not restricted by physical theory. These choices have consequences. That is, interventions made by agents have tangible effects on the evolution of the physical world. In particular, agents' choices may influence their chance of physical survival. Collectively, therefore, agents influence the future evolution of the physical and biological world, which in turn creates the fitness landscape to which they adapt.

According to the ontology presented in this paper, physical evolution produces a discrete sequence of *actual occurrences*, a term taken from Whitehead (1978). These actual occurrences evolve stochastically in a way that depends on the choices made by agents. For each sequence of allowable choices, quantum theory predicts probabilities for which potential occurrences will be actualized. To formalize this intuitive idea, the traditional theory of stochastic processes is insufficient. We need mathematical theory that formalizes the notion of effective interventions performed by agents. For this purpose, we apply ideas from sequential decision theory and Pearl's (2000) interventionist theory of causality. We show that quantum theory is naturally expressed in these terms.

The next two sections present the formal mathematical basis of the causal agent ontology. The causal agent ontology is at its core an ontology of temporal unfolding of systems of agents. It must therefore be reconciled with the twentieth century discovery of the non-absolute nature of space and time. Section 4 describes a relativistic formulation of the causal agent ontology in terms of the relativistic quantum field theory of Tomonaga and Schwinger. Section 5 sketches the outlines of a unified theory of the evolution of active learning agents in a quantum universe. No such unified theory yet exists, but the section speculates on how threads from different research communities might come together to form such a theory. The final section discusses the implications of the causal agent ontology through the eyes of Alice, who is an ardent proponent of its tenets.

2. Causal Markov Processes

It has been asserted (e.g., Leifer, 2006) that quantum theory demands a new, non-commutative variety of probability theory. The causal realist ontology requires only ordinary probability theory. Rather than inventing a new probability theory, we draw on sequential decision theory and Pearl's (2000) theory of causality. We show how to formalize von Neumann quantum theory as a theory of causal Markov processes. Causal Markov processes generalize stochastic processes to allow event probabilities to depend not only on past events, but also on the choices made by agents. The influence of agents' choices is causal in the sense of Pearl (2000).

***Definition* 1:** A *stochastic process* consists of a triple (S, $\mathfrak{A}$, Pr), where S is a *state space*, $\mathfrak{A}$ is an *event algebra*, and Pr is a *probability distribution* on sequences of states, such that the following conditions are satisfied:

i. The event algebra $\mathfrak{A}$ is a subset of the power set of $S^\infty = \{(s_1, s_2, \ldots) : s_i \in S\}$, the set of infinite-length sequences of states.

ii. $\mathfrak{A}$ contains the empty set and is closed under complementation and countable unions. That is, $\emptyset \in \mathfrak{A}$; If $A \in \mathfrak{A}$ then $A^c \in \mathfrak{A}$; and if $A_1, \ldots \in \mathfrak{A}$ then $\cup_i A_i \in \mathfrak{A}$.

iii. Pr is a countably additive probability measure on $\mathfrak{A}$. That is, Pr maps each $A \in \mathfrak{A}$ to a real number $\Pr(A)$, such that $\Pr(A) \geq 0$; $\Pr(S^\infty) = 1$; and $\Pr(\cup_i A_i) = \Sigma_i \Pr(A_i)$.

A probability measure maps sets $A \in \mathfrak{A}$ in the event algebra to probabilities $\Pr(A)$, in such a way that the axioms of probability are satisfied.

Because quantum events have a discrete set of outcomes, we restrict our attention to discrete probability measures. This simplifies the mathematics and the notation. The generalization to continuous measures is straightforward. For brevity, we write $\Pr(s)$ for $\Pr(\{s\})$ and $\Pr(s_1, \ldots, s_n)$ for $\Pr(\{s_1, \ldots, s_n\})$. The restriction to discrete measures allows us to treat finite-length sequences as effectively impossible if they have probability zero. That is, sequences $(s_1, \ldots, s_n)$ for which $\Pr(s_1, \ldots, s_n) = 0$ are assumed not to occur.

A causal Markov process is a family of stochastic processes on a given state space, indexed by a set of allowable actions. Any deterministic or stochastic rule for choosing actions gives rise to one of the processes in the family.

***Definition* 2:** *A* (time-independent first-order) *causal Markov process* is a family of stochastic processes specified by the 3-tuple (S, A, π), where S is a *state space*, A is an *action space*, and π is a *transition distribution*, such that the following conditions are satisfied:

i. For each $s \in S$ and $a \in A$, the function $\pi(\cdot \mid s'; a)$ is a probability measure on S.

ii. If actions are chosen from a probability measure $\theta(a_n \mid h_n)$ that depends only on the past history $h_n = (a_1, \ldots, a_{n-1}, s_0, s_1, \ldots, s_{n-1})$ of actions and states, then the probability measure for actions and states satisfies:

$$\Pr(s_1, \ldots, s_n, a_1, \ldots, a_n \mid s_0) \qquad (1)$$

$$= \prod_{k=1}^{n} \theta(a_k \mid h_k)\pi(s_k \mid s_{k-1}, a_k) .$$

The history-dependent action distribution $\theta(a_n \mid h_n)$ is called a *policy*. Condition (1) is called the *causal Markov* condition. Conditional on each allowable policy, states evolve in a way that depends on the past history only through the most recent past state and the current action.

Interventions satisfy a *locality* condition: if we change a single action probability from $\theta(a_n \mid h_n)$ to $\theta'(a_n \mid h_n)$, the distribution for the first n-1 states remains unchanged, and the conditional distribution for future states given the chosen action a_n also remains unchanged. That is, an intervention that changes $\theta(a_n \mid h_n)$ to $\theta'(a_n \mid h_n)$ is a *local surgery* operator (Pearl, 2000). It changes the likelihood of a_n, but leaves all other causal mechanisms unchanged, so that the evolution of the system is affected only through the change in likelihood of the n^{th} action.

Definition 1 can be generalized in a natural way to systems with time-dependent transition distributions and time-dependent policies. A time-dependent process can always be formally transformed into a time-independent process by including time as a state variable, although this may not be the most natural formulation for many applications. Similarly, an n^{th}-order process, in which the next state depends on states up to n periods in the past, can be re-formulated as a first-order process on a larger state space. This paper focuses on time-independent first-order processes because this class of processes is adequate to represent von Neumann quantum theory.

A causal Markov process is characterized by the state space, action space, and transition distribution. The choice of policy is treated as an exogenous intervention, specified externally to the process. Once states, actions, and transitions and probabilistic predictions have been specified, Equation (1) provides probabilities for *any* policy.

The most common application of causal Markov processes is to sequential optimization problems. For such problems, a causal Markov process is used to represent the distribution of states given the allowable policies. Mathematically, the agent's goal is represented by a real-valued function called a *utility function* that represents the desirability to the agent of a sequence of states. The optimal policy is the one that maximizes expected utility.

***Definition* 3:** *A* (time-independent first-order) *Markov decision process* is a causal Markov process (S, A, π), together with a *utility function* that obeys the backward recursion

$$u_n = f(s_{n+1}, s_n, a_n, u_{n+1}). \qquad (2)$$

To ensure that (2) specifies a well-defined utility function, one typically either imposes a finite horizon and terminal cost, or specifies suitable restrictions on the recursion relation f. A common choice of utility function is the discounted cumulative reward $u_n = \sum_{k\geq n} \xi^k r(s_k)$, where $r(s_k)$ is an immediate reward that depends on the current state, and $0 < \xi < 1$ is a discount factor that gives future rewards less weight than current rewards. Solving a Markov decision process means finding a policy $\theta_{\text{opt}}(a_n \mid h_n)$ to maximize the agent's utility. Optimal policies for Markov decision processes may be obtained via dynamic programming, or when exact solution is intractable, through approximation methods such as value or policy iteration (cf., Bertsekas, 2001). Finding near-optimal solutions to various classes of Markov decision processes is an active area of research and application.

3. Quantum Causal Markov Processes

We have seen that a causal Markov process represents the temporal evolution of a stochastic sequence of states whose likelihoods depend on an exogenously specified policy for choosing actions. This section describes how von Neumann quantum theory can be represented as a causal Markov process. To arrive at this representation, we need to specify the state space, the action space, and the transition distribution. Each of these is considered in turn.

State space. The state of a quantum system is represented by a mathematical structure called a *density operator*. A density operator is a self-adjoint, positive semidefinite operator with unit trace on a Hilbert space $\mathcal{H}$. Each type of quantum system has a characteristic Hilbert space. States correspond to density operators acting on this characteristic Hilbert space. A density operator can be represented as a complex-valued matrix (possibly infinite-dimensional) that is equal to its conjugate transpose, with non-negative diagonal elements that sum to unity. Each density operator has many such representations, related to each other by unitary transformations. A unitary transformation can be represented as a complex-valued matrix whose inverse is equal to its conjugate transpose.

Action space. In von Neumann quantum theory, the allowable actions are called *reductions*, *projective measurements*, or *collapses*. Each reduction is characterized by a *time interval* $d > 0$ and an *operator set* $\{P_i\}$. The P_i are mutually orthogonal projection operators on $\mathcal{H}$ that sum to the identity, i.e.:

i. $P_i^2 = P_i$;
ii. $P_iP_j = 0$ for $i \neq j$; and
iii. $\sum_i P_i = I$.

That is, choosing of a policy means selecting, in a way that may depend on the past sequence of actions and states, a time interval d and a set $\{P_i\}$ of operators satisfying *i-iii*.

Transition distribution. According to Definition 2, the transition distribution is a family of probability measures on states, one for each combination of previous state and current action. For a quantum system, this means specifying a probability distribution on density operators conditional on a density operator ρ, a time interval d, and an operator set $\{P_i\}$. The density operator ρ represents the state just after the previous reduction, the time interval d represents the time interval until the next reduction, and the operator set $\{P_i\}$ represents the possible outcomes of the reduction.

First, we consider actions in which the operator set is the singleton set $\{I_{\mathcal{H}}\}$ consisting of the identity operator on $\mathcal{H}$. In this case, the initial state ρ transforms deterministically into the state $\mathcal{A}_d(\rho)$. The transition from ρ to $\mathcal{A}_d(\rho)$ represents undisturbed mechanical evolution of the system for the time period d. The evolution operator $\mathcal{A}_d$ is a completely positive trace-preserving (CPTP) map. That is, $\text{Tr}(\mathcal{A}_d(\rho)) = \text{Tr}(\rho)$; $\mathcal{A}_d(\rho)$ is a positive operator; and if τ is a density operator on the tensor product space $\mathcal{H}\otimes\mathcal{G}$, then $(\mathcal{A}_d\otimes I_{\mathcal{G}})(\tau)$ is a positive operator, where $I_{\mathcal{G}}$ is the identity operator on the auxiliary Hilbert space $\mathcal{G}$. An important special case is the unitary transformation $\mathcal{A}_d(\rho)$ =

$\exp\{-iH_k d/\hbar\}\ \rho\ \exp\{iH_k d/\hbar\}$, where H is a self-adjoint operator known as the *Hamiltonian*, and $\hbar$ is Planck's constant divided by 2π. Unitary transformations apply to systems evolving in isolation from their environments. Arbitrary quantum operations can be represented as unitary operations on a larger system that includes a system coupled to its environment. Specifically, when a supersystem undergoes a unitary transformation that involves an interaction between a subsystem and its environment, the subsystem considered alone transforms according to a quantum operaton $\mathcal{A}_d(\rho)$ (cf., Nielsen and Chuang, 2000). It can be shown that $\mathcal{A}_d(\rho)$ is continuous in d and $\mathcal{A}_0(\rho)$ is the identity operator $I_{\mathcal{H}}$.

Next, we consider the result of an arbitrary action, consisting of a time interval d and a projection set $\{P_i\}$. In this case, the set $\{\rho_i\}$ of possible outcomes of the action is in one-to-one correspondence with the operators P_i. The probability of outcome ρ_i is given by the *Born rule* $p_i = \mathrm{Tr}(P_i\mathcal{A}_d(\rho)P_i)$. The density matrix ρ_i of the i^{th} outcome state is given by projecting $\mathcal{A}_d(\rho)$ onto the subspace associated with P_i and normalizing to unit trace: $\rho_i = (P_i\ \mathcal{A}_d(\rho)\ P_i)/p_i$. This implies that the outcomes ρ_i are mutually orthogonal. As noted above, although the actions range over an uncountable set (time intervals and sets of projection operators), the set of possible outcomes of any action is finite or countably infinite. Thus, all the outcome distributions are discrete.

Because $\mathcal{A}_d(\rho)$ is continuous, if the action is repeated after a very short interval of time (i.e., d is very near zero), the same state will re-occur with a probability nearly equal to 1. In fact, for small d, the decrease in probability is quadratic in d, giving rise to the quantum Zeno effect, in which repeatedly applying the same operator set in rapid succession tends to hold the quantum system in the same state for a long sequence of events.

Summary. Quantum evolution can be represented as a causal Markov process (S, A, π), where the state space S consists of density operators on a Hilbert space $\mathcal{H}$, the action space A consists of positive time intervals and orthogonal sets of projectors summing to the identity, and the transition distribution consists of Born probabilities applied to the state $\mathcal{A}_d(\rho)$ obtained by mechanically evolving the initial state ρ for the time interval d. This representation is mathematically equivalent to von Neumann quantum theory. Evolution between reductions follows the deterministic quantum evolution rule, and reductions are stochastic events that occur with the Born probabilities.

4. Relativistic Formulation

A quantum causal Markov process as defined above represents the evolution of a sequence of spatially extended states, each occurring at a given instant of time. Each spatially extended state determines, for each possible action, the likelihood of actualizing the next spatially extended state if that action is taken. Figure 1, taken from Stapp (2006), is a schematic depiction in space-time of the sequence of states. Each state is represented in the figure as a horizontal line, constant in time, extending across space. The past is fixed; the future is open. The future unfolds stochastically in a way that depends on the time of the next reduction and the operator set that is applied.

The theory of special relativity tells us that this picture cannot tell the whole story. Special relativity demands that we replace 3-dimensional space and 1-dimensional time with 4-dimensional spacetime. But quantum theory treats space and time in fundamentally different ways. A reduction creates an instantaneous and discontinuous state change across a spatially extended region. But changing reference frames can transform an instantaneous discontinuity to one that spans a time interval. For this reason, the development of a relativistically invariant formulation of quantum theory was a non-trivial achievement. Tomonoga (1946) and Schwinger (1951) accomplished this by indexing quantum states not with time instants, but with spacelike surfaces that advance in a timelike separated sequence.

Figure 2, also taken from Stapp (2007), illustrates how states evolve in the Tomonaga-Schwinger picture. In this picture, each "now" represents a continuous wavy line running across the space-time plane. Time expands by pushing "now" forward to a new wavy line. The new "now" may partly coincide with the old "now," but in the places where it does not coincide, it must lie in the forward direction from the old "now". The numbers 1 through 9 in Figure 2 depict a sequence of "now" surfaces advancing through spacetime. Each "now" represents the occurrence of a reduction. The reduction is applied locally to the part of space-time that is strictly in the future of the previous "now." That is, the operators in the operator set act as the identity on any part of the new "now" that coincides with the old "now." The direct effect of the operator occurs only on the part of the "now" surface that has moved forward since the last reduction. Because of a phenomenon known as quantum entanglement, in which outcomes of reductions applied at one spatial location can be correlated with outcomes of reductions applied at a different spatial location, there may be indirect effects of a reduction on other parts of the "now" surface. Nevertheless, these spacelike correlations are not causal. That is, no signal (controlled message) can propagate faster than the speed of light. Thus, Tomonaga-Schwinger relativistic quantum field theory is consistent with the requirements of special relativity.

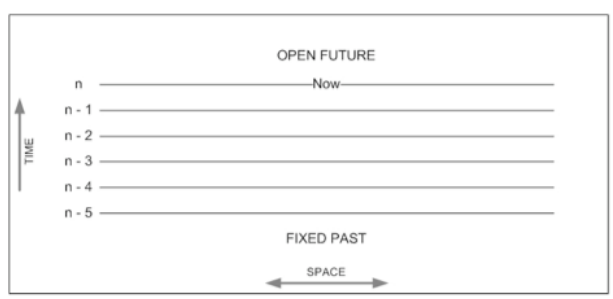

Figure 1: Representation of Space-Time Structure in Non-Relativistic Quantum Theory
(from Stapp, 2007)

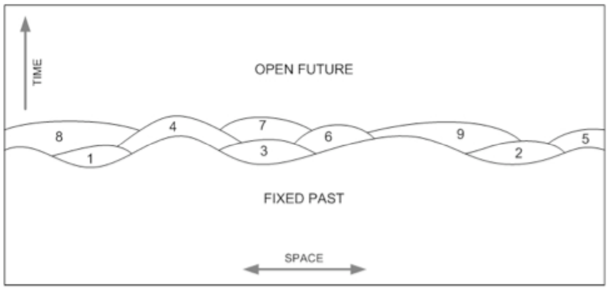

Figure 2: Representation of Space-Time Structure in Relativistic Quantum Theory
(from Stapp, 2007)

We can restate the causal Markov condition for relativistic quantum field theory as follows. Forward predictions for the result of an action taken on a local region of the advancing spacelike surface depend on the past light-cone of the local region only through (*i*) the state of the local region just prior to the action and (*ii*) the set of possible outcomes of the action. That is, suppose we are going to apply a reduction that acts non-trivially only on the spacelike surface represented by the heavy line in Figure 3 (i.e., it acts as the identity elsewhere). To make probabilistic forecasts for the outcome, it suffices to know the quantum state of the local spacelike surface represented by the heavy line and the set of possible outcomes of the reduction. No other aspect of the past light cone (represented by the shaded region spreading out to the past of the heavy line) is relevant to the prediction.

Unitary evolution is time-reversible. That is, if a unitary transformation is applied to advance a local region forward in time, the advance can be "rolled back" to the earlier state by applying a local transformation to the region affected by the transformation. However, the causal Markov property as described in the previous paragraph is not. We can predict the immediate future of a local region from its immediate past and the set of possible states for its immediate future. We cannot predict the immediate past of a local region from its immediate future and the set of possible states for its immediate past.

To illustrate this time-asymmetry, consider the following thought experiment. Alice and Bob have built a device that prepares pure states of a two-state system in a way that does not involve human intervention. The device records whether the YES or NO state has been prepared and, again without human intervention, locks the record in a box with a timing device that will open the lock in five years.

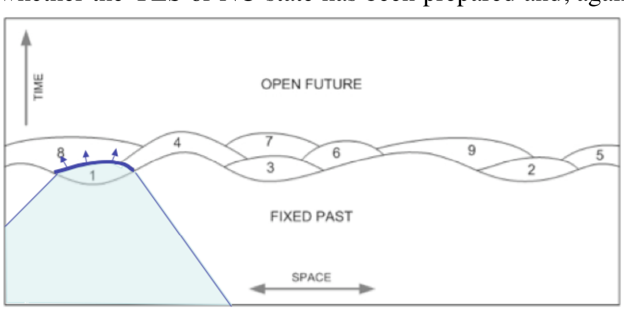

Figure 3: Causal Markov Condition and Forward Prediction

In Scenario 1, Alice prepares a state using this device, measures the state and observes its value. A short time later, she applies an operator set with outcomes Q_1, Q_2, where each of the outcomes has non-zero probability. The only information she needs to make her prediction is whether her initial observation was YES or NO, and that the possible states a very short time after the measurement are $Q_1$and Q_2. The details of the preparation procedure are irrelevant. Indeed, everything else in the past light cone of the experimental situation is irrelevant. All that matters to the future predictions are the state at the last reduction, the time until the next reduction, and the possible states resulting from the next reduction.

When we try to time-reverse this situation, the story is very different. In Scenario 2, Bob applies the same procedure as Alice, except that instead of observing the state as it emerges from the device and trying to predict the result of a measurement a short time later, he observes the result of a reduction a short time after the system emerged from the device and tries to use this result to infer the state that emerged from the device. In drawing his inference, he is allowed to use only information in the future light cone of the heavy line, represented by the shaded area in Figure 4. Again, the possible states after the reduction are $Q_1$and Q_2. Bob makes his best inference of the probabilities that YES or NO would have been observed had he done what Alice did, and "peeked" at the state prior to the reduction that yielded Q_1 or Q_2.

Bob wants to know whether he can perform this inference using only local information. Suppose he makes the best inference he can, using only the state he observed (Q_1 or Q_2) and the fact that the state just prior to his observation was either YES or NO. He asks whether there is information in the future light-cone of the experimental situation that bears on the question. Indeed, there is. Today, the best he can do is to state probabilities for whether the system was in the YES or NO state when it emerged from the device. Five years from now, when the box is unlocked, the actual state will be revealed, and he will know the truth with certainty.

In summary, quantum theory contains an essential time asymmetry. When making forward-in-time predictions of the outcome of the next reduction, the pre-reduction state and the set of possible outcomes of the reduction are sufficient. No other information from the past can improve predictions. When performing backward-in-time retrodiction of the outcome of the previous reduction, the post-measurement state and the set of possible pre-measurement states are not sufficient. Other information in the future is relevant to the retrodiction problem. Figure 4 illustrates this with a dotted line representing future information relevant to the retrodiction problem. That is, the local quantum state "screens off" the past, rendering it irrelevant to local inference about future states. The local quantum state does not "screen off" the future. The future remains relevant to local inference about past states.

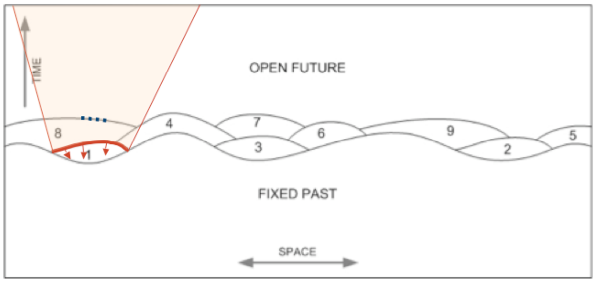

Figure 4: Causal Markov Condition Does Not Apply to Retrodiction

A natural response to the foregoing thought experiment is that if we had time-reversed the entire experimental situation, and not just the outcomes Q_1 and Q_2, we would have been able to make a maximally accurate retrodiction. If the result of the first message (YES, NO) was recorded, it must have been represented physically in some part of the experimental apparatus immediately after the experiment was performed. If this part of the apparatus had been included as part of the heavy line of Figure 4, then the information in the dotted line would have been redundant. That is, a *complete* local state just after the experiment would have been sufficient for maximally accurate retrodiction. But the point remains that the *only* relevant information for forward prediction is the pre-reduction quantum state of the system to which the operator set is being applied. For retrodiction, we need in addition the state of a system that was coupled to it prior to the reduction. This example illustrates the principle that there is in general greater conditional independence for forward inference than for backward inference (Pearl, 2000).

5. Learning and Action in a Quantum World

A universe capable of evolving adapting and learning systems must be both sufficiently structured that its regularities can be learned, and sufficiently complex that learning systems can evolve. A foundational scientific theory adequate to the physical, biological and social sciences must account for adaptation, cognition and learning. It must provide sufficient flexibility to allow for systems that can form representations of their surroundings, use sensory inputs to learn better representations over time, and use the representations to choose survival-enhancing policies. At the same time, there must be enough regularity that learnable survival-enhancing policies exist. If quantum theory is to serve as this foundational theory, then it must have enough flexibility that quantum learners can exist, but it must have enough structure that regularities can emerge for the quantum systems to learn.

A density operator ρ of dimension n can be specified from n^2–1 free parameters, and a trace-preserving completely positive map $\mathcal{A}_d(\rho)$ can be specified from $n^4 - n^2$ free parameters. This is the maximum number of free parameters that would need to be learned to predict the behavior of the system under all possible interventions. For most problems a learning agent would encounter, there are physical symmetries that further reduce the number of free parameters. Learning systems exist in environments characterized by decoherence. These environments can be approximated closely by the equations of classical Newtonian or statistical mechanics. Such environments can be predicted to close approximation by many fewer than the theoretical maximum number of parameters.

In summary, a quantum system evolving in a quasi-stable environment follows a causal Markov process that can, at least in principle, be learned by observing how it behaves and how it responds to interventions. According to the causal agent ontology, there are quantum systems called agents that can initiate reductions, and can control the timing and the operator sets. The rules of causal agent quantum theory permit agents and environments to co-evolve to quasi-stable regimes in which there exist learnable regularities the agents can exploit to enhance their survival chances, and in which agents have sufficient complexity to learn and exploit these regularities.

Much work remains to be done to develop scientifically testable theories of how agency operates. Stapp (2007) hypothesizes that agents exert control by means of the quantum Zeno effect. Stapp hypothesizes that in highly evolved agents such as humans, the brain's unconscious processing nominates templates for action, or patterns of highly organized brain activity that endure for tens or hundreds of milliseconds. Conscious attention acts to hold these templates in place for longer periods of time than they would remain in place if left to the brain's automatic processing.

This idea of a template for action is consistent with the rapidly expanding body of theory and applied work on computer and robotic systems that learn from environmental inputs and choose goal-directed sequences of actions. It is also consistent with research in neuroscience and perception. Most of this literature views brains and computers through the lens of classical physics. Stapp emphasizes that, to a large extent, the functioning of cognitive systems can be understood in classical terms. In the decoherent regime inhabited by cognitive systems, classical stochastic processes provide a good approximation to the phenomenology of perception and decision. Stapp hypothesizes that the process by which the brain processes sensory inputs and arrives at a small set of approximately optimal choices is largely automatic. Nevertheless, attention plays a fundamental role in selecting, activating, and holding in place one choice from among the alternatives nominated by the brain's unconscious processes. Unconscious processing can be approximated well by a classical stochastic process. Conscious attention functions as an executive to select and monitor execution of the agent's policy. Stapp hypothesizes the quantum Zeno effect as the mechanism by which conscious attention exerts control.

A recent paper by Manousakis (2007) contains empirical results that support Stapp's theory. Manousakis examined data from studies of binocular rivalry, a phenomenon in which perception alternates between different images presented to each eye. Using parameters derived from known neural oscillation frequencies and firing rates, Manousakis used Stapp's theory to derive predictions of dominance times for the competing images. Predictions of the theory are in good accord with experimental data gathered under various conditions, including periodic interruption of the stimulus and drug-induced alternate states of consciousness. Stapp (2008) develops a mathematical model of how QZE might operate in a warm, wet brain.

6. Ontology According to Alice and Bob

This section presents the causal agent ontology through the eyes of Alice, an unabashed proponent of its tenets. We contrast Alice's worldview with that of Bob, who subscribes to a variant of one of the most common alternative ontologies, the many worlds or many minds ontology.

Alice believes physical systems evolve as a sequence of actual occasions, and that their evolution is not fully determined by the laws of physics as currently understood. She thinks she makes genuine free choices, unconstrained by the known laws of physics, and that these choices have causal influence on how the world unfolds. For example, when she sets the knobs on the device she and Bob built, her choice of setting is not fully determined by the laws governing the mechanical evolution of the device. She believes her experienced perception, beliefs, desires and values have neural correlates in her physical brain state. She thinks there are automatic processes in her brain that operate below her conscious awareness, and these processes do most of the work of making choices that reflect her values. Yet, she believes there is a remaining element of decision-making that requires effort on her part. She has a good understanding of von Neumann quantum theory and the recent literature on its application to the neuroscience of attention and perception. She thinks her experience of having free will is a reflection of an objectively real part of the world that science is just beginning to understand. Alice thinks that when she makes a free choice, Nature responds to her choice by actualizing one of the possible outcomes. She believes Nature's choice of response is stochastic, and the probabilities of the different responses are objective propensities. She believes these propensities lie behind many of the learnable regularities she sees around her. She has noticed that when she encounters a new type of phenomenon, her predictions are initially not very accurate, but they tend to become more accurate with experience. She has read the literature comparing human reasoning under uncertainty to resource-constrained approximate Bayesian reasoning. Alice believes her brain executes a process akin to resource-constrained approximate Bayesian inference and optimization, in which her unconscious neural processes deliver candidate high-value policies to her consciousness. She thinks she must exert the mental effort necessary to actualize the policy she believes is best. There are times when this is very difficult for her, especially when her natural drives and desires conflict with what she knows is best. She is not always fully successful, but she tries to focus on, instantiate, and perpetuate the actions she believes are best.

Bob disagrees with Alice. Bob agrees with Caves, Fuchs and Schack (e.g., 2006) that quantum probabilities are subjective states of information, and do not correspond to objective properties of Nature. He believes in an Everettian multiverse, in which the myriad possibilities unfold simultaneously, each being experienced by one of his many minds. In this particular multiverse thread, Bob observes himself making ethical and responsible choices that reflect his values. Bob considers it possible that in other co-existing threads, a quantum system nearly identical to himself experiences that version of himself committing despicable acts. Nevertheless, Bob reiterates (in this thread) his commitment to ethical behavior, and continues (in this thread) to act in accordance with his values and ethical principles. He subscribes to Deutch's (1999) theory of rational decision theory in an Everettian multiverse, and believes that Deutsch's theory justifies his choosing actions that reflect his values.

Alice and Bob agree on the mathematics of quantum theory and Bayesian decision theory. They agree on the rationality of using Bayesian inference to update their beliefs and utility maximization to make decisions. When presented with a sequence of outcomes from a series of quantum experiments, over time they come to close agreement in their predictions for future outcomes. However, their interpretations of their predictions differ radically.

Alice believes her predictions are getting better because she is learning the true values of unknown propensities that are an objective property of the physical world. Being a Bayesian, she assigns a prior distribution to these unknown propensities. When an outcome occurs, she updates her prior distribution according to Bayes rule. She considers her prior distribution to be a reflection of her subjective degrees of belief about the objective propensities. She believes both the propensities and the observed outcome sequence are objectively real. Before the outcome occurs, it is a potentiality of the open future, but on its occurrence, it becomes part of the fixed past. She believes that the quantum state of her brain represents her subjective beliefs in a manner analogous to (but *not* identical with) the way a computer represents the probabilities in a Bayesian network. She believes this physical representation is the neural correlate of her experienced beliefs about the device. She believes her learning process is bringing her representation (both its physical manifestation and her experienced beliefs) into closer correspondence with the objective propensities intrinsic to the device.

Bob disagrees. He believes *all* probabilities are subjective. He agrees that his brain is a quantum system, and that its quantum state represents his beliefs, but he does not think the quantum state of his brain is objectively real. Bob thinks the multiverse contains innumerable versions of himself, each experiencing itself as having different beliefs and experiences. The life histories of all these versions exist simultaneously in the multiverse. Bob argues that quantum probabilities must be subjective, because they depend on the observer's state of information – for example, when an observer learns the outcome of a quantum experiment, the probability of the outcome that was observed becomes equal to one for that observer. In each thread of the multiverse, a version of Bob makes predictions on the basis of that version's past experiences. In the vast majority of these threads, Bob's and Alice's predictions come into increasingly close correspondence as they learn about the device. But Bob does not think this is because they are learning about an objective propensity.

Alice has adopted the causal agent ontology because it provides a principled, scientifically justifiable basis for asserting that the universe contains subsystems that can choose and act efficaciously. Alice finds the causal agent ontology attractive because it provides a plausible explanation of the process by which increasingly complex systems have evolved to have ever more sophisticated powers of cognition and agency. The causal agent ontology gives credence to our experience of making conscious choices that cause changes to the world around us. Alice thinks it is important to humanity's survival that we believe we are agents responsible for our own destiny. Despite its correspondence with intuitive experience, Alice understands that the causal agent ontology is at present a minority worldview among scientists and philosophers. She knows she is unlikely to convince someone like Bob to switch positions. As a scientist, she respects Bob's skepticism. Alice acknowledges that scientists are only just beginning to develop the theories and gather the evidence that would clearly establish a scientific basis for choosing between her ontology and other ontologies such as Bob's. Nevertheless, she considers it unfortunate, and dangerous to our future, that many educated people think belief in the reality of efficacious conscious choice is incompatible with science. She thinks people should be aware that the existence of efficacious, value-based free choice is fully compatible with the known laws of physics.

6. Conclusion

The von Neumann formulation of quantum theory can be restated in terms of causal Markov processes. A quantum system can be represented as a causal Markov process on density operators. Actions are represented as projection operators, which punctuate mechanical evolution with stochastic events in which the state is projected onto one of an orthogonal set of subspaces. The causal agent ontology postulates that the universe contains subsystems, called *agents*, that can cause reductions to the degrees of freedom associated to their own states. It is hypothesized that human beings are agents, as are many other kinds of living system. Much work remains to flesh out the details of the causal agent ontology, to provide a testable theory of what kinds of systems can be agents, and to specify precisely how free choice operates in agents.

References

Bertsekas, D. P. 2001. *Dynamic Programming and Optimal Control*. Nashua, NH, Athena Scientific.

Bohm, D., 1951. *Quantum Theory*. New York: Prentice-Hall.

Caves, C. M., Fuchs, C. A., and Schack, R. 2006. *Subjective probability and quantum certainty*. quant-ph/0608190.

Deutch, D. 1999. *Quantum Theory of Probability and Decisions*. http://arxiv.org/pdf/quant-ph/9906015

Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J. and Kruger, L., 1990. *The Empire of Chance : How Probability Changed Science and Everyday Life*. Cambridge, Cambridge University Press.

Leifer, M. S. 2006. Quantum Dynamics as an Analog of Conditional Probability. *Physical Review A 74*, http://arxiv.org /abs/quant-ph/0606022.

Leifer, M. S. 2007. Conditional Density Operators and the Subjectivity of Quantum Operations. In G. Adenier, C. A. Fuchs and A. Yu. Khrennikov (eds.), *Foundations of Probability and Physics - 4, AIP Conference Proceedings vol. 889, (AIP 2007),* pp. 172-186, http://arxiv.org /abs/quant-ph/0611233.

Manousakis, E., 2007. *Quantum theory, consciousness and temporal perception: Binocular rivalry*. http://arxiv.org/abs/0709.4516.

Nielsen, M.A. and Chuang, I.L., 2000. *Quantum Computation and Quantum Information*. Cambridge, UK: Cambridge University Press.

Pearl, J., 2000. *Causality*: *Models, Reasoning and Inference*. Cambridge, UK: Cambridge University Press.

Stapp, H., 2001. Quantum Mechanics and the Role of Mind in Nature. *Foundations of Physics 31*, 1465-99.

Stapp, H., 2007. *The Mindful Universe: Quantum Mechanics and the Participating Observer*. Springer.

Stapp, H., 2008. *The Quantum-Classical and Mind-Brain Linkages: The Quantum Zeno Effect in Binocular Rivalry*. http://www-physics.lbl.gov/~stapp/stappfiles.html.

Schwinger, J. 1951. Theory of quantized fields I. *Physical Review 82*, 914-27.

Tomonaga, S. 1946. On a relativistically invariant formulation of the quantum theory of wave fields. *Progress of Theoretical Physics 1*, 27-42.

Whitehead A.N. 1978. *Process and Reality*, corrected edition by D.R.Griffin and D.W.Sherburne. Free Press, New York. Originally published in 1929.

von Neumann, J., 1955. *Mathematical Foundations of Quantum Mechanics*. Princeton, NJ: Princeton University Press.