Comparison of Quantum and Bayesian Inference Models

Jerome R. Busemeyer and Jennifer Trueblood[1]

[1]Cognitive Science, Indiana University
1101 E. 10th Street, Bloomington Indiana, 47405 jbusemey@indiana.edu

Abstract. The mathematical principles of quantum theory provide a general foundation for assigning probabilities to events. This paper examines the application of these principles to the probabilistic inference problem in which hypotheses are evaluated on the basis of a sequence of evidence (observations). The probabilistic inference problem is usually addressed using Bayesian updating rules. Here we derive a quantum inference rule and compare it to the Bayesian rule. The primary difference between these two inference principles arises when evidence is provided by incompatible measures. Incompatibility refers to the case where one measure interferes or disturbs another measure, and so the order of measurement affects the probability of the observations. It is argued that incompatibility often occurs when evidence is obtained from human judgments.

1 Introduction

Quantum theory was originally invented by physicists to explain findings that seemed paradoxical from the classical physical view point. Later Dirac (1930) and Von Neumann (1932) provided an axiomatic foundation for quantum theory, and by doing so, they discovered that it implied a new type of logic and probability theory. Consequently, there are now at least two general theories for assigning probabilities to events: classic theory and quantum theory [7], [10], [16].

An important application that should be addressed by any general probability theory is the problem of inference – that is, the evaluation of hypotheses on the basis of evidence. Inference is a general problem that arises in many applications. For example, a detective must infer the person who committed a crime on the basis of facts collected from the crime scene and testimony of witnesses. A physician must infer the cause of an illness based on medical symptoms and medical tests. A commander must infer the location of an enemy on the basis of sensory data and intelligence reports. According to classic probability theory, Bayes rule is used to model this kind of probabilistic inference. Quantum probability theory provides an alternative model, and the purpose of this paper is to compare models of probabilistic inference based on Bayesian versus quantum principles.

The evidence used to make an inference is based on observations and measurements. It is well known that the key point upon which Bayesian and quantum models differ is the concept of compatibility of the measures. If all the measures are compatible, then Bayesian and quantum models always agree and the two models assign exactly the same probabilities. Differences only arise when incompatible measures are involved. It is argued that incompatibility can arise when measurements are based on human judgments which interfere with each other and change depending on order of presentation [3].

There already exists a mature literature on hypothesis testing in the area of quantum information theory [8], [9], [10]. This literature is concerned with the problem of determining which of several quantum states describe the true state of a quantum system by performing measurements. The main concern of this literature is designing tests and analyzing the probability of incorrectly deciding which quantum state is the true state. This differs from the present goal which is to describe the revision of a quantum state on the basis of new evidence.

There is also a growing literature concerned with quantum networks that are comparable to Bayesian networks [13], [15], [17]. This literature is concerned with efficient algorithms for computing quantum probabilities from graphical representations of relations. However, the present paper examines more directly the effect of employing sequences of compatible and incompatible measurements on revision of quantum probabilities.

New research is beginning to appear on quantum models of human probability judgments [2], [6], [11]. This work focuses on explaining some paradoxical findings about the way humans make judgments. In contrast, the goal of this paper is to develop a general model that uses human judgments as sources of evidence for making coherent and rational inferences.

2 Probabilistic Inference Task

To begin, we limit our discussion to finite sets (although the number of elements can be very large). The ideas can be extended to infinite sets, but the latter requires more careful handling of convergence and so it is simpler to start with the finite case.

It is assumed that there is a finite set of m hypotheses labeled $\{h_1, \ldots, h_i, \ldots, h_m\}$. For example, these might be suspects for a crime, or causes of an illness, or possible locations of an enemy, or intentions of an opponent.

Evidence is obtained from a sequence of measurements that are taken across time, $t = 1,2, \ldots, T$. Different *types* of measures may be selected at each time step. The notation $X(t) = x_t$ symbolizes that the measure selected at time t produced outcome x_t. For example, a physician may first measure the patient's temperature (producing a degree on a digital thermometer), then ask the patient to judge how much pain he or she experiences (providing a rating on a one to ten scale), and finally ask the patient how long the pain lasts (evaluated in minutes). Each measure is assumed to produce one of a finite set of outcomes.

The task is to determine the probability of each hypothesis after observing a sequence of outcomes:

$$p(h_i \mid X(1) = x_1, X(2) = x_2, \ldots, X(t) = x_t) \text{ for } i = 1,\ldots,m;\ t = 1, \ldots, T. \quad (1)$$

2.1 Classic Probability

Classic probability theory [12] assigns probabilities to events defined as subsets of a sample space, S, which is the universal set. Suppose the cardinality of S equals N. Then we can define N elementary events $S = \{z_1, z_2, \ldots, z_N\}$. Two elementary events can be joined by union to form a new set. Joining elementary events this way, one can generate a family of 2^N sets (including the empty set). This forms a Boolean algebra of sets.

Classic probability postulates the existence of a probability function p that assigns a probability, $0 \leq p(z_i) \leq 1$ to each elementary event. The probability of an arbitrary A event is then defined by $p(A) = \sum_{i \in A} p(z_i)$. Classic probabilities must satisfy $p(S) = \sum_{i \in S} p(z_i) = 1$ for the universal event and $p(\emptyset) = 0$ for the null event.

For purposes of comparison, it is worthwhile to describe classic probability theory using vectors and projection operations. First, we can define an $N \times 1$ vector $|z_j\rangle$ corresponding to elementary event z_j that has all zeros except for a one in row j. Then we can define a projector for event z_j as the outer product $\mathbf{P}_j = |z_j\rangle\langle z_j|$, which is an $N \times N$ matrix full of zeros except a one on the diagonal for row j. The projectors corresponding to different elementary events are orthogonal, $\mathbf{P}_i \cdot \mathbf{P}_j = \mathbf{0}$ for $i \neq j$, and they are complete in the sense that $\sum_j \mathbf{P}_j = \mathbf{I}_N$, where $\mathbf{I}_N$ is the identity matrix.

The projector for an arbitrary event A then equals $\mathbf{P}(A) = \sum_{j \in A} \mathbf{P}_j$. We can represent the probability function by an $N \times 1$ vector of probabilities

$$\pi = \sum_{j \in S} p(z_i) \cdot |z_j\rangle \ . \quad (1)$$

This vector π can be interpreted as the state of the classic probability system. This is called a mixed state. The probability of an event A is determined by the projection of the mixed state followed by a sum of the projection:

$$p(A) = \mathbf{1} \cdot \mathbf{P}(A) \cdot \pi, \text{ with } \mathbf{1} = [\ 1\ 1\ \ldots\ 1]\ . \quad (2)$$

In particular, the probability of the event corresponding to elementary event $|z_j\rangle$ is simply

$$p(z_j) = \mathbf{1} \cdot \mathbf{P}_j \pi\ . \quad (3)$$

Also note that $p(S) = \mathbf{1} \cdot \mathbf{P}(S) \cdot \pi = \mathbf{1} \cdot \mathbf{I} \cdot \pi = 1.0$ and $p(\emptyset) = \mathbf{1} \cdot \mathbf{P}(0) \cdot \pi = \mathbf{1} \cdot \mathbf{0} \cdot \pi = 0$.

2.2 Quantum Probability

Quantum probability theory [5] assigns probabilities to events defined as subspaces of a Hilbert space, H, which is the universal space. Suppose the dimensionality of H is N. Then we can define N orthonormal basis vectors, $\{|z_1\rangle, \ldots, |z_i\rangle, \ldots, |z_N\rangle\}$ where each basis vector represents a one dimensional subspace (corresponding to an elementary event). Two basis vectors can be joined to form a subspace that spans the two vectors. Joining basis vectors this way, one can form a family of 2^N subspaces (including the zero point). As discussed in the concluding section, this forms a partial Boolean algebra of events.

In quantum theory, each basis vector $|z_j\rangle$ corresponds to a projector $\mathbf{P}_j = |z_j\rangle\langle z_j|$ that projects unit length vectors in H onto this basis vector. This forms a complete set of orthogonal projectors $\mathbf{P}_i \cdot \mathbf{P}_j = \mathbf{0}$ for $i \neq j$, and $\sum_j \mathbf{P}_j = \mathbf{I}_N$. The projector corresponding to the join of basis vector $|z_i\rangle$ with basis vector $|z_j\rangle$ equals $\mathbf{P}_i + \mathbf{P}_j$. This implies that each event A is also defined by a projector $\mathbf{P}(\mathrm{A}) = \sum_{j\in \mathrm{A}} \mathbf{P}_j$. Note that the projector for H is $\mathbf{P}(\mathrm{H}) = \sum_{j\in \mathrm{H}} \mathbf{P}_j = \mathbf{I}_N$ (the identity operator) and the projector for the null event is $\mathbf{P}(\varnothing) = \mathbf{0}$ (the zero operator).

Quantum probability postulates the existence of a state vector, denoted here as $|Z\rangle$, which is a unit length vector in the Hilbert space. This state vector can be expressed in terms of coordinates of the basis states as follows:

$$|Z\rangle = \mathbf{I}_N \cdot|Z\rangle = (\textstyle\sum_{j\in \mathrm{H}} |z_j\rangle\langle z_j|)\cdot|Z\rangle = \sum \langle z_j|Z\rangle|z_j\rangle = \sum \alpha_j\cdot|z_j\rangle \,. \quad (4)$$

The coefficient, $\alpha_j = \langle z_j|Z\rangle$ is called probability amplitude corresponding to basis state $|z_j\rangle$. The state vector $|Z\rangle$ is a superposition of the basis states.

The probability of an event A is determined by the squared length of the projection of the state vector onto the subspace that defines the event: $q(\mathrm{A}) = \|\,\mathbf{P}(\mathrm{A})|Z\rangle\|^2$. In particular, the probability of the event corresponding to basis state $|z_j\rangle$ is simply

$$q(z_j) = \|\,\mathbf{P}_j|Z\rangle\|^2 = \|(|z_j\rangle\langle z_j|\cdot|Z\rangle\|^2 = \|\alpha_j\cdot|z_j\rangle\|^2 = \|\alpha_j\|^2 \,. \quad (5)$$

The analogy between classic and quantum theory can be made even clearer if we work directly with the coordinates of the superposition state. According to quantum theory, the coordinates of the state vector $|\mathrm{Z}\rangle$ with respect to the $|z_i\rangle$ basis forms an $N \times 1$ complex vector α. Also with respect to this basis, the projector $\mathbf{P}_j = |z_j\rangle\langle z_j|$ is simply an $N \times N$ matrix full of zeros except a one on the diagonal for row j. Finally, the probability for the event A is simply

$$q(\mathrm{A}) = \|\,\mathbf{P}(\mathrm{A})|Z\rangle\|^2 = \textstyle\sum_{j\in \mathrm{A}} \|\alpha_j\|^2 \,. \quad (6)$$

2.3 Conditional Probabilities

Both classic and quantum probabilities revise (collapse, reduce) the state after observing an event. First consider state revision given by the conditional probability rule of classical probability theory. Suppose the vector π describes the probability distribution across the elementary events prior to a measurement. The probability distribution across the elementary events following an observation of event A equals the projection onto event A followed by normalization:

$$\pi|\mathrm{A} = (\mathbf{P}(\mathrm{A})\cdot\pi) \,/\, (\mathbf{1}\cdot\mathbf{P}(\mathrm{A})\cdot\pi) \,. \quad (7)$$

Normalization guarantees that the elements of the new state vector sum to unity. All subsequent conditional probabilities are then computed from projections on the new mixed state $\pi|\mathrm{A}$. Specifically, the probability of a new event B given A is

$$p(\mathrm{B}|\mathrm{A}) = \mathbf{1}\cdot\mathbf{P}(\mathrm{B})\cdot(\pi|\mathrm{A}) = \mathbf{1}\cdot(\mathbf{P}(\mathrm{B})\cdot\mathbf{P}(\mathrm{A})\cdot\pi)/(\mathbf{1}\cdot\mathbf{P}(\mathrm{A})\cdot\pi) \,. \quad (8)$$

Next consider the state revision given by Luder's rule [14] of quantum probability theory. Suppose the vector $|Z\rangle$ describes the superposition state prior to a measurement. The state following an observation of event A equals the projection onto event A followed by normalization:

$$|Z|\mathrm{A}\rangle = \mathbf{P}(\mathrm{A})|Z\rangle \,/\, ||\mathbf{P}(\mathrm{A})|Z\rangle|| \,. \tag{9}$$

Normalization guarantees that the new state vector has length equal to one. Probabilities of new events, conditioned on already observing event A, are then computed from projections on the new state vector $|Z|\mathrm{A}\rangle$. The coefficients for this conditional state vector, with respect to the $|z_j\rangle$ basis, are defined by

$$\alpha|\mathrm{A} = \mathbf{P}(\mathrm{A})\cdot\alpha \,/\, ||\mathbf{P}(\mathrm{A})\cdot\alpha|| \,. \tag{10}$$

For example the probability of a new event B given A is

$$q(\mathrm{B}|\mathrm{A}) = ||\mathbf{P}(\mathrm{B})\cdot(\alpha|\mathrm{A})||^2 = ||\mathbf{P}(\mathrm{B})\cdot\mathbf{P}(\mathrm{A})\alpha||^2/||\mathbf{P}(\mathrm{A})\cdot\alpha||^2 \,. \tag{11}$$

2.4 Inference Based on a Single Measurement

Suppose T=1 and one wishes to make an inference based on a single measure denoted X. Classic and quantum theories provide the same answers to this problem. First we present the classic Bayesian inference model (using projection operators), followed by the quantum inference model.

Suppose the measure X can take n different values, with x representing an arbitrary outcome. When we combine each possible outcome x of X with each of the m possible hypotheses, we obtain $N = m\cdot n$ unique elementary joint events such as $(h_i \wedge x)$. (Later this number will change.) The classic inference process starts with initial distribution over these N events represented by the $N \times 1$ vector π. The prior probability is given by

$$p(h_i) = \mathbf{1}\cdot\mathbf{P}(h_i)\cdot\pi \,, \tag{12}$$

where $\mathbf{P}(h_i)$ is a $N \times N$ matrix with ones on the diagonal corresponding to the rows matching hypothesis h_i.

The marginal probability equals

$$p(X(1) = x) = \mathbf{1}\cdot\mathbf{P}(X(1) = x)\cdot\pi \,, \tag{13}$$

where $\mathbf{P}(X(1)=x)$ is a $N \times N$ matrix with ones on the diagonal corresponding to the rows matching $X(1)= x$. The new state after observing $X(1) = x$ becomes

$$\pi|x = \mathbf{P}(X(1) = x)\cdot\pi \,/\, (\mathbf{1}\cdot\mathbf{P}(X(1) = x)\cdot\pi) \,. \tag{14}$$

If hypothesis h_i is known to be true then the new state is

$$\pi|h_i = \mathbf{P}(h_i)\cdot\pi \,/\, (\mathbf{1}\cdot\mathbf{P}(h_i)\cdot\pi) \,. \tag{15}$$

Finally, the likelihood is then given by

$$p(X(1) = x|\, h_i) = \mathbf{1}\cdot\mathbf{P}(X(1) = x)\cdot(\pi|h_i) \,. \tag{16}$$

Bayes inference rule follows from the definition of conditional probability:

$$p(h_i \mid X(1) = x_1) = \mathbf{1}\cdot\mathbf{P}(h_i)\cdot(\pi|x)$$
$$= \mathbf{1}\cdot\mathbf{P}(h_i)\cdot\mathbf{P}(X(1){=}x)\cdot\pi \,/\, (\mathbf{1}\cdot\mathbf{P}(X(1){=}x)\cdot\pi)$$
$$= \mathbf{1}\cdot\mathbf{P}(X(1){=}x)\cdot\mathbf{P}(h_i)\cdot\pi \,/\, (\mathbf{1}\cdot\mathbf{P}(X(1){=}x)\cdot\pi)\ . \tag{17}$$

Substituting $\pi|h_i = \mathbf{P}(h_i)\cdot\pi \,/\, (\mathbf{1}\cdot\mathbf{P}(h_i)\cdot\pi)$ into the above

$$p(h_i \mid X(1) = x_1) = (\mathbf{1}\cdot\mathbf{P}(h_i)\cdot\pi)\cdot(\mathbf{1}\cdot\mathbf{P}(X(1){=}x)\cdot(\pi|h_i)) \,/\, (\mathbf{1}\cdot\mathbf{P}(X(1){=}x)\cdot\pi)$$
$$= p(h_i)\cdot[\ p(X(1) = x \mid h_i) \,/\, p(X(1){=}x)\ ]\ . \tag{18}$$

Next we examine the quantum inference model for a single measure. We begin with an initial state defined on an N dimensional Hilbert space. This Hilbert space can be represented by $N = m\cdot n$ basis vectors such as $|h_i \wedge x\rangle$ representing the elementary event $(h_i \wedge x)$. (Later this dimension will change.) The initial state can be represented by the $N \times 1$ coordinate vector α with respect to this basis.

The prior probability of hypothesis h_i equals

$$q(h_i) = ||\mathbf{P}(h_i)\cdot\alpha||^2\ . \tag{19}$$

The marginal probability of event $X(1) = x$ is

$$q(X(1) = x) = ||\mathbf{P}(X(1) = x)\cdot\alpha||^2\ . \tag{20}$$

After observing first observation, $X(1)=x$, the initial state α changes to a new state

$$\alpha|x = \mathbf{P}(X(1){=}x)\cdot\alpha \,/\, ||\mathbf{P}(X(1) = x)\cdot\alpha||, \text{ and } ||\alpha|x|| = 1\ . \tag{21}$$

Suppose we assume that h_i is true. Then the conditional state given h_i equals

$$\alpha|h_i = \mathbf{P}(h_i)\cdot\alpha \,/ ||\mathbf{P}(h_i)\cdot\alpha\,|| \text{ and } ||\alpha|h_i|| = 1\ . \tag{22}$$

If we assume that h_i is true, then the conditional probability of observing $X(1) = x$ equals

$$q(X(1) = x \mid h_i) = ||\mathbf{P}(X(1) = x)\cdot(\alpha|h_i)||^2\ . \tag{23}$$

Finally, quantum inference follows from Luder's rule [14]

$$q(h_i \mid x) = ||\mathbf{P}(h_i)\cdot(\alpha|x)||^2$$
$$= ||\mathbf{P}(h_i)\mathbf{P}(X(1) = x)\cdot\alpha||^2 \,/\, ||\mathbf{P}(X(1) = x)\cdot\alpha||^2$$
$$= ||\mathbf{P}(X(1) = x)\mathbf{P}(h_i)\cdot\alpha||^2 \,/\, ||\mathbf{P}(X(1) = x)\cdot\alpha||^2\ . \tag{24}$$

Substituting $\mathbf{P}(h_i)\cdot\alpha = (\alpha|h_i)\cdot||\mathbf{P}(h_i)\cdot\alpha||$ yields

$$q(h_i \mid x) = ||\mathbf{P}(h_i)\cdot\alpha||^2\cdot||\mathbf{P}(X(1){=}x)\cdot(\alpha|h_i)||^2 \,/ ||\mathbf{P}(X(1){=}x)\cdot\alpha||^2$$
$$= q(h_i)\cdot[\ q(\,X(1) = x \mid h_i) \,/\, q(\,X(1) = x)\ ]\ . \tag{25}$$

This is identical to Bayes rule if the classic probability function p replaces the quantum probability function q.

3 Representation of Measurements

3.1 Change of Basis Vectors

A key issue arises from the idea of using a Hilbert space representation of events. One can choose different sets of basis vectors for spanning a Hilbert space. Two different sets of basis vectors are related by a unitary transformation, denoted **U** with $\mathbf{U}\mathbf{U}^{\dagger} = \mathbf{U}^{\dagger}\mathbf{U} = \mathbf{I}_N$:

$$\{|z_1\rangle, \ldots, |z_i\rangle, \ldots, |z_N\rangle\} = \{\mathbf{U}|z'_1\rangle, \ldots, \mathbf{U}|z'_i\rangle, \ldots, \mathbf{U}|z'_N\rangle\},$$
$$\{|z'_1\rangle, \ldots, |z'_i\rangle, \ldots, |z'_N\rangle\} = \{\mathbf{U}^{\dagger}|z_1\rangle, \ldots, \mathbf{U}^{\dagger}|z_i\rangle, \ldots, \mathbf{U}^{\dagger}|z_N\rangle\}. \quad (26)$$

A state vector can be expressed with respect to either one of these two sets of basis vectors:

$$|Z\rangle = \mathbf{I}_N \cdot|Z\rangle = (\textstyle\sum_{j\in H} |z_j\rangle\langle z_j|)\cdot|Z\rangle = \sum \langle z_j|Z\rangle|z_j\rangle = \sum \alpha_j\cdot|z_j\rangle,$$
$$= \mathbf{I}_N \cdot|Z\rangle = (\textstyle\sum_{j\in H} |z'_j\rangle\langle z'_j|)\cdot|Z\rangle = \sum \langle z'_j|Z\rangle|z'_j\rangle = \sum \beta_j\cdot|z'_j\rangle. \quad (27)$$

The coordinates of the state vector $|Z\rangle$ with respect to the $|z_i\rangle$ basis forms a $N \times 1$ complex vector α, and the probability of elementary event z_i is $||\alpha_i||^2$. The coordinates of $|Z\rangle$ with respect to the $|z'_i\rangle$ basis forms a $N \times 1$ complex vector denoted β, and the probability of elementary event z_i is $||\beta_i||^2$. These coordinates are related by a unitary matrix $U = [u_{ij}] = [\langle z'_i|z_j\rangle] = [\langle z_i|\mathbf{U}|z_j\rangle]$: $\beta = U\cdot\alpha$, and $\alpha = U^{\dagger}\cdot\beta$.

By changing basis vectors, we change the nature of the set of elementary events under consideration. In particular, the $|z_i\rangle$ basis is needed to define elementary events $\{z_1, z_2, \ldots, z_N\}$, and the projector $|z_j\rangle\langle z_j|$ defines event z_j; but $|z'_i\rangle$ is needed to define elementary events $\{z'_1, z'_2, \ldots, z'_N\}$, and the projector $|z'_j\rangle\langle z'_j|$ defines event z'_j. In other words, if we ask questions about the elementary events $\{z_1, z_2, \ldots, z_N\}$, then we need to use the α coordinates to compute probabilities, and $||\alpha_i||^2$ determines the probability of elementary event z_i; but if we ask questions about the elementary events $\{z'_1, z'_2, \ldots, z'_N\}$, then we need to use the β coordinates to compute probabilities and $||\beta_i||^2$ determines the probability of elementary event z'_i.

3.2 Compatibility of Measures

The concept of compatibility is unique to quantum theory. It is concerns the possible disturbing effect of one measure, say *X*, on another measure, say *Y*. We could take these measurements in two different orders: *X* first followed by *Y*, or *Y* first followed by *X*. If the probability of the two events produced by the two measurements does not depend on the order, then these two measures are compatible; otherwise they are incompatible [5]. Human judgments frequently exhibit order effects, hence the concern for compatibility.

One measure is labeled *X* and it yields an event such as A = (X=x) where x is one of the n possible outcomes produced by *X*. We assume that these $n(X)$ outcomes correspond to a set of $n(X)$ orthogonal projectors

$\{\mathbf{P}(X = x_1),\ldots,\mathbf{P}(X = x),\ldots,\mathbf{P}(X = x_{n(X)})\}$ operating on the N dimensional Hilbert space that forms a complete set so that

$$\mathbf{P}(X = x)\cdot\mathbf{P}(X=y) = \mathbf{0} \text{ for } x \neq y,\ \sum_x \mathbf{P}(X = x) = \mathbf{I}_N\,. \tag{28}$$

Suppose the A = (X=x) event uses the $|z_j\rangle$ set of basis vectors for its definition and corresponds to the projector

$$\mathbf{P}(X = x) = \sum_{j\in \mathrm{A}} |z_j\rangle\langle z_j|\,. \tag{29}$$

The other measure is labeled Y and it yields an event such as B = (Y = y) where y is one of the $n(Y)$ possible outcomes produced by Y. Again we assume that these $n(Y)$ outcomes correspond to a set of $n(Y)$ orthogonal projectors
$\{\mathbf{P}(Y=y_1),\ldots,\mathbf{P}(Y = y),\ldots,\mathbf{P}(Y = y_{n(Y)})\}$ that forms a complete set so that

$$\mathbf{P}(Y = x)\cdot\mathbf{P}(Y=y) = \mathbf{0} \text{ for } x \neq y,\ \sum_y \mathbf{P}(Y = y) = \mathbf{I}_N\,. \tag{30}$$

The B = (Y=y) event requires the $|z'_j\rangle$ set of basis vectors for its definition and corresponds to the projector

$$\mathbf{P}(Y = y) = \sum_{j\in \mathrm{B}} |z'_j\rangle\langle z'_j| = \sum_{j\in \mathrm{B}} \mathbf{U}^\dagger|z_j\rangle\langle z_j|\mathbf{U}\,. \tag{31}$$

The probability of the A followed by B sequence is

$$||\mathbf{P}(Y = y)\cdot\mathbf{P}(X = x)\cdot|Z\rangle||^2\,. \tag{32}$$

The probability of the opposite sequence of events is

$$||\mathbf{P}(X = x)\cdot\mathbf{P}(Y = y)\cdot|Z\rangle||^2\,. \tag{33}$$

Two measures are said to be compatible if the probability distribution over the joint outcomes from the two measures does not depend on the order of measurement; otherwise they are incompatible.

These two sequences give the same probability for any arbitrary state vector $|Z\rangle$ and any pair of outcomes x and y if and only if the commutator is always zero:

$$[\mathbf{P}(X = x)\cdot\mathbf{P}(Y = y) - \mathbf{P}(Y = y)\cdot\mathbf{P}(X = x)] = \mathbf{0}\,. \tag{34}$$

If the commutator is zero for all pairs of values that can be observed on the two measures, then the two measures are compatible; otherwise they are incompatible. The commutator is always zero when

$$(\,|z_j\rangle\langle z_j|\,)\cdot(\,|z'_k\rangle\langle z'_k|\,) - (\,|z'_k\rangle\langle z'_k|\,)\cdot(\,|z_j\rangle\langle z_j|\,) = \mathbf{0} \tag{35}$$

for all pairs, which holds when $\langle z_j|z'_k\rangle = 0$ for $j \neq k$ and $\langle z_j|z'_j\rangle = 1$. This implies that
$\mathbf{U} = \mathbf{I}_N$, or in other words, the basis set $|z_j\rangle = \mathbf{U}\cdot|z'_j\rangle$ used for X is identical to the basis set $|z'_j\rangle$ used for Y. If $\mathbf{U} \neq \mathbf{I}_N$ then the measure X will be incompatible with the measure Y.

If the measures are compatible, then we can define the joint event

$$\mathbf{P}(X=x\wedge Y=y) = \mathbf{P}(X=x)\cdot\mathbf{P}(Y=y) = \mathbf{P}(Y=y)\cdot\mathbf{P}(X=x)\,. \tag{36}$$

This forms a new complete set of $n(X)\cdot n(Y)$ orthogonal projectors
$\{\ldots,\mathbf{P}(X=x\wedge Y=y),\ldots\}$ so that $\mathbf{P}(X=x\wedge Y=y)\cdot\mathbf{P}(X=u\wedge Y=v) = \mathbf{0}$, $x \neq u$ or $y \neq v$,

$\sum_x \sum_y$ $\mathbf{P}(X{=}x \wedge Y{=}y) = \mathbf{I}_N$. These projectors are then used to define the joint probabilities for these two measures

$$q(X{=}x \wedge Y{=}y) = \|\mathbf{P}(X{=}x \wedge Y{=}y)|Z\rangle\|^2 \; . \tag{37}$$

Classic probability theory assumes that it is always possible to define these joint probabilities between measures. However, in quantum theory, this joint probability does not exist for incompatible measures.

When two measures are compatible, then the first measure does not disturb or affect the second measure, order of measurement does not matter, and both measures can be determined simultaneously. However, when two measures are incompatible, then determining the value of one measure necessarily makes the values of the other measure uncertain. To see how this uncertainty principle arises with incompatible measures, suppose the inference state $|Z\rangle$ is placed at the following point after a measurement:

$$|Z|x\rangle = \mathbf{P}(X = x)|Z\rangle \,/\, \|\mathbf{P}(X = x)|Z\rangle\| \; . \tag{38}$$

We can express this state using the coordinates defined by $|z_j\rangle$ as follows:

$$(\alpha|x) = \mathbf{P}(X = x)\cdot\alpha \,/\, \|\mathbf{P}(X = x)\cdot\alpha\|, \text{ and so } \|\alpha|x\|^2 = 1 \; . \tag{39}$$

Here $\mathbf{P}(X = x)$ is the matrix representation of the projector with respect to the $|z_j\rangle$ basis (it is simply a matrix with zeros everywhere except for ones on the diagonal in the rows corresponding to combinations that satisfy $X = x$). Given this state, the outcome x is certain to occur again with measure X:

$$\begin{aligned} q(X = x) &= \|\mathbf{P}(X = x)\cdot(\alpha|x)\|^2 \\ &= \|\mathbf{P}(X = x)^2\cdot\alpha\|^2 \,/\, \|\mathbf{P}(X = x)\cdot\alpha\|^2 = 1 \; . \end{aligned} \tag{40}$$

Now let us examine this same state using the coordinates defined by $|z'_j\rangle$:
$(\beta|x) = U\cdot(\alpha|x)$, and note that $\|(\beta|x)\|^2 = \|U\cdot(\alpha|x)\|^2 = 1$ because U is unitary. The probability of the outcome y for the measure Y is determined by

$$q(Y = y) = \|\mathbf{P}(Y = y)\cdot(\beta|x)\|^2 \tag{41}$$

where $\mathbf{P}(Y = y)$ is the matrix representation with respect to the $|z'_j\rangle$ basis. That is, it is a matrix with zeros everywhere except ones on the diagonal for rows that satisfy $Y = y$. Now we find that $q(Y = y) = \|\mathbf{P}(Y{=}y)\cdot(\beta|x)\|^2 < 1$because $\|(\beta|x)\|^2 = 1$ and $\mathbf{P}(Y{=}y)$ is a projection on a subspace of H.

In other words, if X and Y are incompatible, and if we are certain about the outcome that X will produce, then we must be uncertain about the outcome produced by Y.

3.3 Constructing a Hilbert Space Representation

We construct our Hilbert space using the principles initially described by Dirac [5]. The dimension, N, of the Hilbert space used to represent all T measures is determined by a maximum number $K \leq T$ of mutually compatible measures. Incompatible measures, being unitary transformations of a set of compatible measures, remain within the same space, and so they do not increase the dimensionality of the Hilbert space.

If all of the measures are compatible with each other, and $K = T$, then we can use the same set of basis vectors to represent events for all T of the measures. This is exactly the key assumption of classical probability theory. In fact, quantum probability assigns the same probabilities to all of the events as classical probability when all of the measures are compatible.

Hereafter we will assume that there are only $K \leq T$ compatible measures labeled $\{X_1,\ldots,X_k,\ldots,X_K\}$. As described earlier, an incompatible measure Y_k can be constructed from one of the compatible set X_k by a unitary transformation of the basis.

Any given measure from the compatible set, say X_k, has $n(k)$ possible outcomes $\{x_1,x_2,\ldots,x_k,\ldots,x_{n(k)}\}$. This produces a total of $n = [n(1)\cdot n(2)\cdot\ldots\cdot n(k)\cdot\ldots\cdot n(K)]$ combinations of possible outcomes from all K compatible measures, such as $z_j = (X_1{=}x_1)\wedge\ldots\wedge(X_k{=}x_k)\wedge\ldots\wedge(X_K{=}x_K)$.

To model inference, we also need to include the m possible hypotheses, $\{h_1,\ldots,h_m\}$. We assume the hypotheses are compatible with all of the measures. Combining each hypothesis with each combination of measurement outcomes defines an elementary event, and so this produces a set of $N = m\cdot n$ elementary events, with typical element $(h_i \wedge z_j) = h_i\wedge(X_1{=}x_1)\wedge\ldots\wedge(X_K{=}x_K)$. These events are represented as subspaces (rays) of an N dimensional Hilbert space.

The $n(k)$ outcomes produced by X_k are represented by a complete set of $n(k)$ orthogonal projectors $\{\ldots,\mathbf{P}(X_k{=}x),\ldots\}$, $\mathbf{P}(X_k{=}x)\cdot\mathbf{P}(X_k{=}y)=\mathbf{0}$, $x\neq y$, $\sum_x \mathbf{P}(X_k{=}x) = \mathbf{I}_N$. There is one such set of projectors from each measure, and all K compatible measures are defined by the same basis. Finally, each hypothesis is represented by a set of orthogonal projectors $\{\ldots,\mathbf{P}(h_i),\ldots\}$, $\mathbf{P}(h_i)\cdot\mathbf{P}(h_j) = \mathbf{0}$, $i \neq j$, $\sum_i \mathbf{P}(h_i) = \mathbf{I}_N$. The projectors for the hypotheses are defined by the same basis as the projectors for the compatible measures.

The projectors for the outcomes of each measure can be combined to form a projector for each combination of outcomes. The projector for an elementary event $(h_i \wedge z_j)$ is $\mathbf{P}_{ij} = \mathbf{P}(h_i)\cdot\mathbf{P}(X_1{=}x_1)\cdot\ldots\cdot\mathbf{P}(X_k{=}x_k)\cdot\ldots\cdot\mathbf{P}(X_K{=}x_K)$. This forms a complete set of N orthogonal projectors: $\{\ldots,\mathbf{P}_{ij}\ldots\}$, $\mathbf{P}_{ij}\cdot\mathbf{P}_{i'j'} = \mathbf{0}$, $ij \neq i'j'$, $\sum_i \sum_j \mathbf{P}_{ij} = \mathbf{I}_N$. Each projector $\mathbf{P}_{ij}$ for an elementary event has a single and unique (unit length) eigenvector

$$|h_i \wedge z_i\rangle = |h_i \wedge (X_1{=}x_1)\wedge\ldots\wedge(X_k{=}x_k)\wedge\ldots\wedge(X_K{=}x_K)\rangle \,. \tag{43}$$

These basis vectors span the $N \times 1$ dimensional Hilbert space denoted H.

It is useful to construct each basis vector $|h_i \wedge z_j\rangle$ of H from a tensor product of vectors representing the K compatible measures and the hypotheses as follows. Consider the measure X_k that has $n(k)$ possible values. For this measure, we can define a set of $n(k)$ orthonormal vectors $\{|X_k{=}x_1\rangle,\ldots,|X_k{=}x\rangle,\ldots,|X_k{=}x_{n(k)}\rangle\}$, where each possible outcome, say x, from the measure X_k corresponds to a normalized vector $|X_k{=}x\rangle$. The set of orthonormal vectors for X_k spans an $n(k) \times 1$ subspace H_k. Then the tensor product of the individual measurement vectors produces a vector representing a combination of measurement outcomes:

$$\begin{aligned}|z_j\rangle &= |X_1{=}x_1\rangle\otimes\ldots\otimes|X_k{=}x_k\rangle\otimes\ldots\otimes|X_K{=}x_K\rangle \\ &= |\,(X_1{=}x_1)\wedge\ldots\wedge(X_k{=}x_k)\wedge\ldots\wedge(X_K{=}x_K)\,\rangle \,.\end{aligned} \tag{44}$$

There are total of n such vectors which form an orthonormal set that spans an n dimensional tensor product space $H_1 \otimes H_2 \otimes \ldots \otimes H_k \otimes \ldots \otimes H_K$.

The hypotheses are represented by a set of orthonormal vectors $\{|h_1\rangle, \ldots, |h_i\rangle, \ldots, |h_m\rangle\}$. This orthonormal set forms an m-dimensional Hilbert space denoted H_m. Then each basis vector of H can be constructed from a tensor product of vectors from each measure and hypothesis:

$$\begin{aligned} |h_i \wedge z_j\rangle &= |h_i\rangle \otimes |z_j\rangle \\ &= |h_i\rangle \otimes |X_1{=}x_1\rangle \otimes \ldots \otimes |X_k{=}x_k\rangle \otimes \ldots \otimes |X_K{=}x_K\rangle \\ &= |h_i \wedge (X_1{=}x_1) \wedge \ldots \wedge (X_k{=}x_k) \wedge \ldots \wedge (X_K{=}x_K)\rangle \,. \end{aligned} \tag{45}$$

This spans an N dimensional tensor product space $H = H_m \otimes H_1 \otimes \ldots \otimes H_k \otimes \ldots \otimes H_K$. Finally the state of the quantum system is defined as

$$|\psi(t)\rangle = \sum \sum \psi_{ij}(t) \cdot |h_i\rangle \otimes |z_j\rangle \,. \tag{46}$$

The $N \times 1$ vector of coefficients, $\psi(t)$, represents the state with respect to the basis formed by $|h_i\rangle \otimes |z_j\rangle$. So if we wish to compute the joint probability that hypothesis h_i is true and that elementary event z_j occurs, then this is given by

$$||\mathbf{P}(h_i \wedge z_j) \cdot \psi(t)||^2 = ||\psi_{ij}(t)||^2 \,, \tag{42}$$

where $\mathbf{P}(h_j \wedge z_j)$ is a $N \times N$ matrix with zeros everywhere except a one on the diagonal in the row corresponding to the event $(h_j \wedge z_j)$.

4 Quantum Inference

4.1 Quantum Inference after the First Measurement

The first measure to be selected is denoted $X(1)$. If this is one of the measures in the compatible set $\{X_1, \ldots, X_k, \ldots, X_K\}$ then we proceed by using the basis with vectors $|z_j(1)\rangle = |X_1{=}x_1\rangle \otimes \ldots \otimes |X_k{=}x_k\rangle \otimes \ldots \otimes |X_K{=}x_K\rangle$.

If the first measure is incompatible with one of the measures in compatible the set $\{X_1, \ldots, X_k, \ldots, X_K\}$, then we need to change the basis by applying a unitary transformation. Suppose the first measure is Y_k which is incompatible with X_k. Then the basis vectors for these elementary events are defined by

$$\begin{aligned} |z_j(1)\rangle &= |X_1{=}x_1\rangle \otimes \ldots \otimes |Y_k{=}y_k\rangle \otimes \ldots \otimes |X_K{=}x_K\rangle \\ &= |X_1{=}x_1\rangle \otimes \ldots \otimes \mathbf{U}^{\dagger}{}_k |X_k{=}x_k\rangle \otimes \ldots \otimes |X_K{=}x_K\rangle \,. \end{aligned} \tag{48}$$

where we have replaced $(X_k{=}x_k)$ with $(Y_k{=}y_k)$. In other words, we start with a basis $|z_j(1)\rangle$ that is defined by the first measure. The initial state is represented by:

$$|\psi(0)\rangle = \sum \sum \psi_{ij}(0) \cdot |h_i\rangle \otimes |z_j(1)\rangle \,. \tag{49}$$

The $m \cdot N$ vector of coordinates $\psi(0)$ represents the state with respect to this initial basis.

In section 2.4, we calculated the quantum inference based on a single measurement. By letting $\alpha = \psi(0)$ in these calculations we have the inference after the first observation for the initial state $\psi(0)$.

4.2 Quantum Inference after a Second Measurement

Suppose we take another measurement $X(2) = x_2$ which defines our second event. If it is compatible with the first measure, $X(1)$, then we simply continue working with the same basis by setting $\psi(1|x_1) = \alpha(1|x_1)$, where $\alpha(1|x_1)$is the new state defined by Equation 22. If it is incompatible with the first measure, then we need to change coordinates.

Suppose the first measure was chosen to be $X(1) = X_k$ and the coefficients for $X(1)$ are initially expressed in terms of the $|z_j(1)\rangle$ basis with coordinates $\alpha(1|x_1)$ given by Equation 22. Now suppose the second measure is $X(2) = Y_k$, which is incompatible with X_k. So we need to change the coordinates from α which are defined by the $|z_j(1)\rangle$ basis to β which are defined with respect to the new basis:
$\beta(1|x_1) = (\mathbf{I}\otimes\ldots\otimes\mathbf{I}\otimes U_k\otimes\mathbf{I}\otimes\ldots\mathbf{I})\cdot\alpha(1|x_1)$. Finally we set $\psi(1|x_1) = \beta(1|x_1)$ and continue as before. The projector for $\mathbf{P}(X(2) = x_2)$ is simply an $N \times N$ matrix with zeros everywhere except ones on the diagonals of the rows that correspond to the event $X(2) = x_2$. $\mathbf{P}(h_i)$ is simply an $N \times N$ matrix with zeros everywhere except ones on the diagonals of the rows that correspond to hypothesis h_i.

The prior after the first measure but before the second observation equals

$$q(h_i|x_1) = \|\mathbf{P}(h_i)\cdot\psi(1|x_1)\|^2 . \qquad (50)$$

The probability of event $X(2) = x_2$ given $X(1)=x_1$ is

$$q(X(2)=x_2|X(1)=x_1) = \|\mathbf{P}(X(2) = x_2)\psi(1|x_1)\|^2 . \qquad (51)$$

After observing the second observation, $X(2)=x_2$, the state $\psi(1|x_1)$ changes to a new state

$$\alpha(2|x_1, x_2) = \mathbf{P}(X(2)=x_2)\psi(1|x_1)/\|\mathbf{P}(X(2) = x_2)\psi(1|x_1)\| \qquad (52)$$

and $\|\alpha(2| x_1, x_2)\| = 1$.

Suppose we assume that h_i is true. Then the conditional state given h_i and $X(1)=x_1$ equals

$$\psi(1|x_1, h_i) = \mathbf{P}(h_i)\psi(1|x_1)/\|\mathbf{P}(h_i)\psi(1|x_1)\| \qquad (53)$$

and $\|\psi(1|x_1, h_i)\| = 1$.
If h_i is true, and we already observed $X(1) = x_1$, then the conditional probability given of observing $X(2) = x_2$ equals

$$q(X(2)=x_2 |x_1, h_i) = \|\mathbf{P}(X(2) = x_2)\psi(1|x_1, h_i)\|^2 . \qquad (54)$$

Finally, our inference after the second observation equals

$$q(h_i \mid x_1, x_2) = \|\mathbf{P}(h_i)\alpha(2|x_1, x_2)\|^2$$

$$= ||\mathbf{P}(h_i)\mathbf{P}(X(2)=x_2)\psi(1|x_1)||^2 \,/\, ||\mathbf{P}(X(2)=x_2)\psi(1|x_1)||^2$$

$$= ||\mathbf{P}(X(2)=x_2)\mathbf{P}(h_i)\psi(1|x_1)||^2 \,/\, ||\mathbf{P}(X(2)=x_2)\psi(1|x_1)||^2$$

$$= ||\mathbf{P}(h_i)\psi(1|x_1)||^2 \times ||\mathbf{P}(X(2)=x_2)\psi(1|x_1, h_i)||^2 \div ||\mathbf{P}(X(2)=x_2)\psi(1|x_1)||^2$$

$$= q(h_i|x_1) \cdot [q(X(2) = x_2 \,|x_1, h_i) \,/\, q(X(2) = x_2 \,|\, x_1)] \,. \tag{55}$$

Once again, this corresponds with Bayes rule.

4.3 Quantum Inference after Several Observations

A new measure is denoted $X(t+1)$. The process continues with the coefficients produced by the last measurement, $\alpha(t|x_1, \ldots,x_t)$. For example, if $t+1=3$, then we would start with Equation 52. If $X(t+1)$ is compatible with $X(t)$ then we use the same basis states as used for the last measurement. That is, we continue using the same coordinates $\psi(t|x_1, \ldots,x_t) = \alpha(t|x_1, \ldots,x_t)$. If $X(t+1) = Y_k$ is incompatible with $X(t)$, then we transform the coordinates to the new basis for the new measure: $\beta(t|x_1, \ldots,x_t) = (\mathbf{I}\otimes\ldots\otimes U_j\otimes\ldots\otimes\mathbf{I})\cdot\alpha(t|x_1, \ldots,x_t)$. In this case, we express the coordinates of the current inference state as $\psi(t|x_1, \ldots,x_t) = \beta(t|x_1, \ldots,x_t)$.

The prior after the first measure but before the second observation equals

$$q(h_i|x_1,\ldots,x_t) = ||\mathbf{P}(h_i)\cdot\psi(t|x_1,\ldots,x_t)||^2 \,. \tag{56}$$

The probability of event $X(t+1) = x_{t+1}$ given the previous history is

$$q(X(t+1)=x_{t+1}|x_1,\ldots,x_t) = ||\mathbf{P}(X(2) = x_2)\psi(t|x_1,\ldots,x_t)||^2 \,. \tag{57}$$

After observing the next observation, $X(t+1)=x_{t+1}$, the state $\psi(t|x_1,\ldots,x_t)$ changes to a new state

$$\alpha(t+1|x_1, \ldots,x_{t+1}) = \mathbf{P}(X(t+1)=x_{t+1})\psi(t|x_1,\ldots,x_t) \div ||\mathbf{P}(X(t+1)=x_{t+1})\psi(t|x_1,\ldots,x_t)|| \tag{58}$$

and $||\alpha(t+1|\, x_1,\ldots, x_{t+1})|| = 1$.

Suppose we assume that h_i is true. Then the conditional state given h_i and the past history equals

$$\psi(t|x_1, \ldots,x_t, h_i) = \mathbf{P}(h_i)\psi(t|x_1,\ldots,x_t)/||\mathbf{P}(h_i)\psi(t|x_1,\ldots,x_t)|| \tag{59}$$

and $||\psi(t|x_1,\ldots,x_t,\; h_i)|| = 1$. If h_i is true, and we already observed a history of events, then the conditional probability of $X(t+1) = x_{t+1}$ given h_i and this history equals

$$q(\, X(t+1)=x_{t+1}|x_1,\ldots,x_t,h_i) = ||\mathbf{P}(X(t+1)=x_{t+1})\cdot\psi(t|x_1,\ldots,x_t, h_i)||^2 \,. \tag{60}$$

Finally, our inference after the next observation equals

$$q(h_i \,|\, x_1, \ldots,x_{t+1}) = ||\mathbf{P}(h_i)\alpha(t+1|x_1,\ldots, x_{t+1})||^2$$

$$= ||\mathbf{P}(h_i)\mathbf{P}(X(t+1)=x_{t+1})\cdot\psi(t|x_1,\ldots,x_t)||^2 \div ||\mathbf{P}(X(t+1)=x_{t+1})\psi(t|x_1,\ldots,x_t)||^2$$

$$= ||\mathbf{P}(X(t+1)=x_{t+1})\mathbf{P}(h_i)\cdot\psi(t|x_1,\ldots,x_t)||^2 \div ||\mathbf{P}(X(t+1)=x_{t+1})\cdot\psi(t|x_1,\ldots,x_t)||^2$$

$$= ||\mathbf{P}(h_i)\psi(t|x_1,\ldots,x_t)||^2 \times ||\mathbf{P}(X(t+1)=x_{t+1})\psi(t|x_1,\ldots,x_t,h_i)||^2$$
$$\div ||\mathbf{P}(X(t+1)=x_{t+1})\cdot\psi(t|x_1,\ldots,x_t)||^2$$

$$= q(h_i|x_1,\ldots,x_t) \times [q(X(t+1)=x_{t+1}|x_1,\ldots,x_t,h_i) \div q(X(t+1)=x_{t+1}|\, x_1,\ldots,x_t)] \,. \qquad (61)$$

Again, this corresponds with Bayes rule if the classic probability function *p* replaces the quantum probability function *q*.

5 Summary and Concluding Comments

This paper began with the assumption that the abstract mathematical basis of quantum theory is not tied to physics per se, but rather it can be used as a generalized probability theory with meaningful applications outside of physics. If so, it should be applicable to probabilistic inference problems.

If we do this, we find that quantum inferences are updated in manner that correspond exactly to Bayesian updating except that the coordinates of the state must be transformed by unitary matrices to coordinates of a different a basis for changes between incompatible measurements. Determining the unitary matrices that transform from one set of coordinates to another is a critical step that remains to be achieved for applications outside of physics.

Quantum inference is identical to Bayesian inference when only compatible measures are involved. But quantum inference can depart dramatically from Bayesian inference when incompatible measurements are involved. In particular, one can start out certain about a particular value of a measure, but if this is followed later by an incompatible measure, then one will become uncertain again about the value of the earlier (certain) measure. This results from the disturbance of one incompatible measure on another.

When all the measures are compatible, we have one set of elementary events and this forms a single Boolean algebra of events. When incompatible measures are involved, we need to define different incompatible sets of elementary events, which correspond to different sets of basis vectors within the same Hilbert space. These sets of events cannot be combined into a single comprehensive set of events using Boolean logic.[1] Thus we are forced to work with a partial Boolean algebra of events.

So the most crucial question is whether incompatible measurements occur outside of physics. There is clear evidence that one type of human judgment can disturb another and the order of human judgments changes the probabilities [3]. This suggests that it may be fruitful to employ quantum probabilities when human judgments are involved.

Cognitive psychologists have attempted to describe the disturbing effect of one judgment on another by building cognitive models that describe a separate probability distribution for each order. However, they have implicitly assumed a partial Boolean algebra to formulate these models, and thus these models are not really consistent with classic probability theory either. An important empirical question is whether simpler yet more generalizable probabilistic models can be found using quantum

[1] The events from incompatible measures follow quantum logic (not discussed here), which obeys all of the rules of Boolean logic except for the distributive axiom.

probabilities. The success in physics suggests this may be the case. This remains to be seen outside of physics (but see, [1], [11]).

Acknowledgments. This research was supported by NSF SES-0753164 and SES-0753168. Thanks for the IU quantum group (Amr Sabry, Larry Moss, Andrew Hanson, Gerardo Ortiz, Michael Dunn) for discussions of the ideas described in this paper.

6 References

1. Aerts, D., Broekaert, J., Gabora, L.: A Case for Applying an Abstracted Quantum Formalism to Cognition. In: R. Campbell (ed.) Mind in Interaction. John Benjamin, Amsterdam (2003)
2. Bordley, R. F.: Quantum Mechanical and Human Violations of Compound Probability Principles: Toward a Generalized Heisenberg Uncertainty Principle. Operations Research 46, 923--926 (1998)
3. Busemeyer, J. R., Wang, Z., Townsend, J. T.: Quantum Dynamics of Human Decision Making. Journal of Mathematical Psychology 50 (3), 220--241 (2006)
4. Degroot, M. H.: Optimal Statistical Decisions. McGraw – Hill, New York (1970)
5. Dirac, P.A.M.: The Principles of Quantum Mechanics. Oxford University Press, Oxford (1930, 2001)
6. Franco,R.: The Conjunction Fallacy and Interference Effects, arXiv:0708.3948v1 [physics.gen-ph]. to appear in Journal of Mathematical Psychology (2007)
7. Guder, S. P.: Stochastic Methods in Quantum Mechanics. Dover Press (1979)
8. Hayashi, M.: Quantum Information Theory: An Introduction. Springer, Berlin: (2006)
9. Helstrom, C. W.: Quantum Detection and Estimation Theory. Academic Press, New York (1976)
10. Holevo, A. S.: Statistical Structure of Quantum Mechanics. Springer, Berlin (2001)
11. Khrennikov, A.: Can Quantum Information be Processed by Macroscopic Systems? Quantum Information Theory, in press (2007)
12. Kolmogorov, A. N.: Foundations of the Theory of Probability. Chelsea Publishing Co., New York (1950)
13. Laskey, K. B.: Quantum Causal Networks. In: Bruza, P. D., Lawless, W., van Rijsbergen, C.J., Sofge, D., (eds.) Proceedings of the AAAI Spring Symposium on Quantum Interaction, March 27-29. Stanford University, 2007. AAAI Press (2007)
14. Luders, G.: Uber die Zustandsanderung durch den Messprozess. Annalen der Physik 8, 322--328 (1951)
15. La Mura, P., Swiatczak, L.: Markov Entanglement Networks. In: Bruza, P. D., Lawless, W., van Rijsbergen, C.J., Sofge, D., (eds.) Proceedings of the AAAI Spring Symposium on Quantum Interaction, March 27-29. Stanford University, 2007. AAAI Press (2007)
16. Pitowski, I.: Quantum Probability, Quantum logic. Lecture Notes in Physics 321. Springer, Heidelberg (1989)
17. Tucci, R. R.: Quantum Bayesian Nets. quantph/9706039 (1997)
18. Von Neumann, J.: Mathematical Foundations of Quantum Theory. Princeton University Press, Princeton, NJ (1932)