Quantum Information Processing Theory

Jennifer S. Trueblood and Jerome R. Busemeyer
Indiana University
Bloomington, IN
USA
jstruebl@indiana.edu and jbusemey@indiana.edu

Synonyms

Applied quantum probability theory

Definition

Quantum information processing theory is an alternative mathematical approach for generating theories of how an observer processes information. Typically, quantum information processing models are derived from the axiomatic principles of quantum probability theory. This probability theory may be viewed as a generalization of classic probability. Quantum information processing models do not make assumptions about the biological substrates. Instead this approach provides new conceptual tools for constructing social and behavioral science theories.

Theoretical Background

There are two mathematical approaches to constructing probabilistic systems: classic Kolmogorov probabilities and quantum von Neumann probabilities. The majority of information processing models in cognitive science and psychology use the classical probability system. However, classic probability and information processing theory impose a restrictive set of assumptions on the representation of complex systems such as the human cognitive system. Quantum information processing theory postulates a more general method for representing and analyzing these types of complex systems. This chapter begins by providing motivation for adopting the quantum approach and then gives a mathematical comparison of classical and quantum probability theories.

Reasons to adopt the quantum probability framework for cognitive science

1. Cognitive measures create cognitive states rather than record states. For example, suppose an oberver is trying to understand the emotional state of a person after the person is presented with an arousing stimulus. Classical information processing theory posits that the cognitive system of the individual is in a definite state before any measurements are taken. The process of imposing the measurement has no effect on this state other than to simply record it. On the other hand, quantum information processing theory postulates that the state of the cognitive system is undetermined before measurement, and it is the process of imposing measurements that determines the state. Thus,

judgment is not a simple read out from a pre-existing or recorded state, instead it is constructed from the current context and question.

2. Cognition behaves like a wave rather than a particle. For example, suppose a juror is deciding whether a defendant is guilty or not. Classical information processing theory assumes that the cognitive system is always in one of two states, guilty or innocent. So, at any given moment in time, the juror's cognitive state is clearly known to him or her. However, quantum theory suggests that beliefs are superimposed and do not jump from one state to another. Thus, the juror can feel a sense of ambiguity about all of the states simultaneously.
3. Cognitive measures disturb each other, creating uncertainty. For example, an experimenter questions a subject about his or her preferences for cars. In one scenario, the experimenter asks the subject directly the type of car he or she would like to buy, and the subject responds with a specific preference. However, in another scenario, the experimenter first asks what type of car the subject's spouse would like to buy. Then, when the experimenter asks about the subject's own preferences, they become less certain. Classical theories cannot capture the effects of measurement disturbances. However, using quantum information processing theory, questions can be represented as incompatible thus allowing for one question to disturb the answer to another.
4. Cognitive logic does not obey classic logic. Returning to the jury decision-making example, a juror might believe the defendant is guilty or innocent, and he or she might also feel that the defendant is good or bad. According to classic logic, the distributive axiom yields:

$$Guilty \wedge (Good \vee Bad) = (Guilty \wedge Good) \vee (Guilty \wedge Bad)$$

On the other hand, quantum theory allows for the existence of a superimposed state, and does not always obey the distributive axiom:

$$Guilty \wedge (Good \vee Bad) \neq (Guilty \wedge Good) \vee (Guilty \wedge Bad)$$

Thus, quantum logic is more generalized than classical logic and can model human judgments that do not obey Boolean logic.

Consider the four points above in terms of 'physical measurements' and 'physical systems' instead of 'human judgments' and 'cognitive systems', then these points are similar to the ones that faced physicists in the 1920s that forced them to develop quantum theory. In other words, physical findings that seemed paradoxical from the classical point of view led physicists to invent quantum theory. Similarly, paradoxical findings in cognitive psychology suggest that classical probability theory is too limited to fully explain various aspects of human cognition. These phenomena include violations of the sure thing axiom of decision making, violations of the conjunctive and disjunctive rules of classic probability theory, and cooperation in Prisoner Dilemma games.

A Formal Comparison of Classical and Quantum Probability Theories

Classical probability theory postulates a set of all possible outcomes Ω called the outcome space. For example, when flipping a coin the outcome space $\Omega = \{head, tail\}$. Events are defined as subsets of the Ω and correspond to things that might or might not happen. New events can be formed from other events in three different ways. Specifically, let $x, y,$ and z be events in the outcome space Ω. The negation operator, $\neg x$, denotes the complement of x. The conjunction operator, $x \wedge y$, denotes the intersection of x and y. The disjunction operator, $x \vee y$, denotes the union of x and y. Since events are mathematically represented as sets, they obey the rules of Boolean algebra:

1. Commutative: $x \vee y = y \vee x$
2. Associative: $x \vee (y \vee z) = (x \vee y) \vee z$
3. Complementation: $x \vee (\neg y \wedge y) = x$
4. Absorption: $x \vee (x \wedge y) = x$
5. Distributive: $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$

A probability distribution, $\Pr$, is a function of events that obeys the Kolmogorov axioms:

1. Non-negative: $\Pr(x) \geq 0$
2. Addition: If $x_1, \ldots, x_n$ is a partition of x, then $\Pr(x) = \Pr(x_1) + \ldots + \Pr(x_n)$
3. Total one: $\Pr(\Omega) = 1$

Quantum probability theory replaces the outcome space Ω with a Hilbert space H (i.e. a complex vector space). Quantum events are defined geometrically as a subspace (e.g. a line or place, etc.) within this Hilbert space. Similar to classical probability theory, new events can be formed in three ways. Let L_x, L_y and L_z represent three different events in H. The negation operator, $L_x^{\perp}$, denotes the maximal subspace that is orthogonal to L_x. The meet operator, $L_x \wedge L_y$, denotes the intersection of the two subspaces, L_x and L_y. The join operator, $L_x \vee L_y$, denotes the span of L_x and L_y. Quantum logic obeys all of the rules of Boolean logic except for the distributive axiom:

$$L_x \wedge (L_y \vee L_z) = (L_x \wedge L_y) \vee (L_x \wedge L_z).$$

Quantum probability postulates the existence of a unit length state vector $|z\rangle \in H$. (The use of Dirac, or Bra-ket, notation is in keeping with the standard notation used in quantum mechanics. Specifically, $|x\rangle$ denotes a column vector whereas $\langle x|$ denotes a row vector.) This state vector depends on the context of the situation being modeled. For example, the state vector might represent an individual's cognitive state during a decision-making task. Quantum probabilities are computed by projecting $|z\rangle$ onto subspaces representing events.

Specifically, for each event L_x there is a corresponding projection operator P_x that projects $|z\rangle$ onto L_x. The probability of the event L_x is equal to the squared length of this projection:

$$\Pr(L_x) = |(P_x|z\rangle|^2$$

Since events are mathematically represented as subspaces, there exists a basis for each event. For example, let $|x_1\rangle,...,|x_n\rangle$ be a set of basis vectors for the event corresponding to the subspace L_x. This is similar to a partition of an event in classical probability theory. When selecting a basis for L_x, select one that is orthonormal: inner products $\langle x_i|x_j\rangle = 0$ and lengths $|\langle x_i|x_j\rangle| = 1$. Now, use this basis to construct projectors. The projector, P_{x_i}, projects the state vector $|z\rangle$ onto the subspace L_{x_i}, and it is constructed from the outer product $P_{x_i} = |x_i\rangle\langle x_i|$. The projector, P_x, for the event L_x is constructed by adding the projects corresponding to the basis vectors:

$$P_x = P_{x_1\vee...\vee x_n} = P_{x_1} + ... + P_{x_n} = |x_1\rangle\langle x_1| + ... + |x_n\rangle\langle x_n|.$$

Now, write the probability of the L_x as a sum of separate probabilities:

$$\Pr(L_x) = |(P_x|z\rangle|^2 = |(|x_1\rangle\langle x_1| + ... + |x_n\rangle\langle x_n|)|z\rangle|^2$$

$$= ||x_1\rangle\langle x_1|z\rangle + ... + |x_n\rangle\langle x_n|z\rangle|^2 = |\langle x_1|z\rangle|^2 + ... + |\langle x_n|z\rangle|^2$$

where the final step follows from the orthogonality property. Finally for any state vector $|z\rangle$, $P_H \cdot |z\rangle = |z\rangle$ showing $|P_H \cdot |z\rangle|^2 = |\langle z|z\rangle|^2 = 1$. From this we see that $P_H = \sum_i^m P_{x_i} = \sum_i^m |x_i\rangle\langle x_i| = I$ where $|x_1\rangle,...,|x_m\rangle$ is a basis for the Hilbert space H and I is the identity operator.

From these properties quantum probabilities obey rules analogous to the Kolmogorov rules:

1. Non-negative: $\Pr(L_x) = |(P_x|z\rangle|^2 \geq 0$
2. Addition: If $x_1,...,x_n$ is a basis of L_x and $L_{x_i} \wedge L_{x_j} = 0$ for all i, j, then

   $\Pr(L_x) = \Pr(L_{x_1}) + ... + \Pr(L_{x_n})$
3. Total one: $\Pr(H) = 1$

Another important concept in quantum probability theory is the notion of superposition and quantum measurement. Any state vector $|z\rangle$ prior to measurement can be expressed in terms of the basis states as follows:

$$|z\rangle = I \cdot |z\rangle = \sum_i^m |x_i\rangle\langle x_i| \cdot |z\rangle = \sum_i^m |x_i\rangle\langle x_i|z\rangle = \sum_i^m \langle x_i|z\rangle \cdot |x_i\rangle.$$

Since the inner product called the probability amplitude $\langle x_i | z \rangle$ is a scalar, the state vector is a linear combination, or superposition, of the basis states. The act of measurement results in the normalized projection of the state vector onto the corresponding subspace. For example, performing measurement X changes the initial state $| z \rangle$ to a new state $| (z | x) \rangle$. This is a achieved by projecting $| z \rangle$ onto the subspace corresponding to X, L_x. Then the projection is normalized so the new state has unit length.

When taking more than one measurement of a state vector, quantum theory allows for the measurements to either be compatible or incompatible. Intuitively, compatibility means that measurements X and Y can be accessed simultaneously or sequentially without interfering with each other. On the other hand, if X and Y are incompatible, they cannot be accessed simultaneously. From a cognitive standpoint, this implies that the two measurements are processed serially and one measurement interferes with the other. Mathematically, incompatible measurements, or events, are represented by different bases for the same n-dimensional subspace. For example, if measurement X is represented by the subspace L_x with basis $| x_1 \rangle, \ldots, | x_n \rangle$ and measurement Y is represented by the subspace L_y with basis $| y_1 \rangle, \ldots, | y_n \rangle$, the $| x_i \rangle$ basis is a linear transformation of the $| y_i \rangle$ basis. If X and Y are compatible, then there is one basis representation for both measurements. In this case, quantum probability theory reduces to classic probability theory (Hughes, 1989; Nielsen & Chuang, 2000).

Although classical and quantum probability theory share many features such as Kolmogorov-like rules, these two theories have important differences. First, quantum theory does not necessarily obey the distributive axiom of Boolean logic. Also, quantum probability postulates that the state of a system is a superposition of all possible states before measurement, and it is the act of measurement that changes the state. Finally, quantum theory allows for both compatible and incompatible measurements whereas classical probability theory assumes all measurements are compatible.

Important Scientific Research and Open Questions

To illustrate how quantum information processing theory can be applied to empirical data, consider a categorization and decision-making study conducted by Busemeyer, Wang and Lambert-Mogiliansky (2009). In this experiment, participants were shown pictures of faces that varied on two dimensions (face shape and lip thickness) and were asked to categorize the faces as ‘good people’ or ‘bad people’ and to decide to act friendly or aggressive. In one condition, participants made an action decision without reporting a category. In a second condition, participants made an action decision after categorizing the face. In the first condition, the categorization task can be considered a hidden inference. Thus, by the law of total probability from classical probability theory,

$$\Pr(aggressive \mid face) = \Pr(good \mid face) \cdot \Pr(aggressive \mid good) + \Pr(bad \mid face) \cdot \Pr(aggressive \mid bad).$$

However, data collected from the experiment does not conform to the law of total probability. From the first condition, Busemeyer et al. found $\Pr(aggressive|\ face) = 0.69$. Then, from the second condition, they found $\Pr(good \mid face) = 0.17$, $\Pr(aggressive| good) = 0.42$, $\Pr(bad \mid face) = 0.83$, and $\Pr(aggressive| bad) = 0.63$. By applying the law of total probability to the data from the second condition, the result is

$$\Pr(aggressive \mid face) = 0.17 \cdot 0.42 + 0.83 \cdot 0.63 = 0.59.$$

Obviously, this does not match the data collected in condition one.

Now, consider a quantum probability explanation for the data. Let $\langle good \mid face \rangle$ be the amplitude for transiting from the face stimulus to a categorization of good. Thus, $|\langle good \mid face \rangle|^2$ is the probability of a face to 'good person' transition. Since the only two categorization responses are good and bad, it is the case that $|\langle good \mid face \rangle|^2 + |\langle bad \mid face \rangle|^2 = 1$. Similarly, let $\langle aggressive| good \rangle$ be the amplitude for transiting from the 'good person' categorization to the aggressive action decision. Then, the probability of this transition is $|\langle aggressive| good \rangle|^2$ and $|\langle aggressive| good \rangle|^2 + |\langle friendly| good \rangle|^2 = 1$. Thus,

$$\Pr(aggressive \mid face) = |\langle aggressive \mid face \rangle|^2$$

$$= |\langle good \mid face \rangle \langle aggressive \mid good \rangle + \langle bad \mid face \rangle \langle aggressive \mid bad \rangle|^2$$

$$= |\langle good \mid face \rangle \langle aggressive \mid good \rangle|^2 + |\langle bad \mid face \rangle \langle aggressive \mid bad \rangle|^2$$

$$+ 2 \cdot |\langle good \mid face \rangle \langle aggressive \mid good \rangle \langle bad \mid face \rangle \langle aggressive \mid bad \rangle| \cdot \cos(\theta)$$

where $\cos(\theta)$ is the interference effect term. This result shows that quantum theory does not obey the law of total probability. So, there is a value for θ such that the quantum probability model matches the experimental data.

Researchers are currently working towards building unified theories of human cognition based on the principles of quantum theory. To date, quantum information processing models have been used to explain cognitive phenomena including violations of rational decision-making principles, paradoxes of conceptual combination, human judgments, perception, and order effects on human inference (Atmanspacher et al., 2004; Busemeyer et al., in press; Pothos & Busemeyer, 2009).

Cross-References

→ Human information processing
→ Mathematical models/theories of learning
→ Modeling and simulation
→ Quantum analogical modeling

References

Atmanspacher, H., Filk, T., & Romer, H. (2004). Quantum zero features of bistable perception. *Biological Cybernetics*, 90 , 33-40.

Busemeyer, J. R., Pothos, E. M., Franco, R. & Trueblood, J. S. (in press). A Quantum Theoretical Explanation for Probability Judgment Errors. *Psychological Review*.

Busemeyer, J.R., Wang, Z., & Lambert-Mogiliansky, A. (2009). Empirical comparisons of Markov and quantum models of decision making. *Journal of Mathematical Psychology*, 53, 423-433.

Hughes, R.I.G. (1989). *The structure and interpretation of quantum mechanics*. Harvard University Press.

Nielsen, M.A. & Chuang, I.L. (2000). *Quantum computation and quantum information*. Cambridge University Press.

Pothos, E. M., & Busemeyer, J. R. (2009). A quantum probability explanation for violations of 'rational' decision theory. *Proceedings of the Royal Society B*.

Definitions

Interference effect: Empirical violations of the law of total probability are termed interference effects. These effects can be found both in physical systems and human cognitive systems. Quantum theory was initially invented to explain these effects in particle physics. Psychologists have also found interference effects in humans, motivating researchers to apply quantum theory to cognitive science.