Quantum-like Contextual Model for Processing of Information in Brain

Andrei Khrennikov
International Center for Mathematical Modeling
in Physics and Cognitive Sciences
University of Växjö, S-35195, Sweden

Abstract

We present a quantum-like model in that contexts (complexes of physical, mental, or social conditions) are represented by complex probability amplitudes. This approach gives the possibility to apply the quantum formalism for probabilities induced in any domain of science. In this paper we propose a model of brain functioning based on the quantum-like (QL) representation of mental contexts. In this model brain can be considered as a QL (but not conventional quantum) computer. We discuss a possible mechanism of the QL-representation of information in the brain. It is based on processing of information at two different time scales: pre-cognitive (fine) and cognitive (coarse).

Introduction

We present a contextualist statistical realistic model for quantum-like representations in cognitive science and psychology (Khrennikov 2004), (Khrennikov 2005). We apply this model to describe cognitive experiments to check quantum-like structures of mental processes. The crucial role is played by the *interference of probabilities for mental observables.* Recently one such experiment based on recognition of images was performed, see (Khrennikov 2004). This experiment confirmed our prediction on the quantum-like behavior of mind. In our approach "quantumness of mind" has no direct relation to the fact that the brain (as any physical body) is composed of quantum particles. We invented a new terminology "quantum-like (QL) mind." Cognitive QL-behavior is characterized by a nonzero *coefficient of interference* λ ("coefficient of supplementarity"). This coefficient can be found on the basis of statistical data. There are predicted not only $\cos\theta$-interference of probabilities, but also hyperbolic $\cosh\theta$-interference. The latter interference was never observed for physical systems, but we could not exclude this possibility for cognitive systems. We propose a model of brain functioning as a QL-computer. We shall discuss the difference between quantum and QL-computers.

From the very beginning we emphasize that our approach has nothing to do with *quantum reductionism.* Of course, we do not claim that our approach implies that quantum physical reduction of mind is totally impossible. But our approach could explain the main QL-feature of mind – *interference of minds* – without reduction of mental processes to quantum physical processes. Regarding the quantum logic approach we can say that our contextual statistical model is close mathematically to some models of quantum logic (Mackey 1963), but interpretations of mathematical formalisms are quite different. The crucial point is that in our probabilistic model it is possible to combine *realism* with the main distinguishing features of quantum probabilistic formalism such as *interference of probabilities, Born's rule, complex probabilistic amplitudes, Hilbert state space, and representation of (realistic) observables by operators.*

Copyright © 2007, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

Observational Contextual Statistical Model

A general statistical realistic model for observables based on the contextual viewpoint to probability will be presented. It will be shown that classical as well as quantum probabilistic models can be obtained as particular cases of our general contextual model, the *Växjö model*.

This model is not reduced to the conventional, classical and quantum models. In particular, it contains a new statistical model: a model with hyperbolic *cosh*-interference that induces "hyperbolic quantum mechanics" (Khrennikov 2004).

A physical, biological, social, mental, genetic, economic, or financial *context* C is a complex of corresponding conditions. Contexts are fundamental elements of any contextual statistical model. Thus construction of any model M should be started with fixing the collection of contexts of this model. Denote the collection of contexts by the symbol $\mathcal{C}$ (so the family of contexts $\mathcal{C}$ is determined by the model M under consideration). In the mathematical formalism $\mathcal{C}$ is an abstract set (of "labels" of contexts).

We remark that in some models it is possible to construct a set-theoretic representation of contexts – as some family of subsets of a set Ω. For example, Ω can be the set of all possible parameters (e.g., physical, or mental, or economic) of the model. However, in general we *do not assume the possibility to construct a set-theoretic representation of contexts.*

Another fundamental element of any contextual statistical model M is a set of observables $\mathcal{O}$: each observable $a \in \mathcal{O}$ can be measured under each complex of conditions $C \in \mathcal{C}$. For an observable $a \in \mathcal{O}$, we denote the set of its possible values ("spectrum") by the symbol X_a.

We do not assume that all these observables can be measured simultaneously. To simplify considerations, we shall consider only discrete observables and, moreover, all concrete investigations will be performed for *dichotomous observables.*

Axiom 1: *For any observable $a \in \mathcal{O}$ and its value $x \in X_a$, there are defined contexts, say C_x, corresponding to x-selections: if we perform a measurement of the observable a under the complex of physical conditions C_x, then we obtain the value $a = x$ with probability 1. We assume that the set of contexts $\mathcal{C}$ contains C_x-selection contexts for all observables $a \in \mathcal{O}$ and $x \in X_a$.*

For example, let a be the observable corresponding to some question: $a = +$ (the answer "yes") and $a = -$ (the answer "no"). Then the C_+-selection context is the selection of those participants of the experiment who answering "yes" to this question; in the same way we define the C_--selection context. By Axiom 1 these contexts are well defined. We point out that in principle a participant of this experiment might not want to reply at all to this question. By Axiom 1 such a possibility is excluded. By the same axiom both C_+ and C_--contexts belong to the system of contexts under consideration.

Axiom 2: *There are defined contextual (conditional) probabilities $\mathbf{P}(a = x|C)$ for any context $C \in \mathcal{C}$ and any observable $a \in O$.*

Thus, for any context $C \in \mathcal{C}$ and any observable $a \in O$, there is defined the probability to observe the fixed value $a = x$ under the complex of conditions C.

Especially important role will be played by probabilities:

$$p^{a|b}(x|y) \equiv \mathbf{P}(a = x|C_y), a, b \in \mathcal{O}, x \in X_a, y \in X_b,$$

where C_y is the $[b = y]$-selection context. By axiom 2 for any context $C \in \mathcal{C}$, there is defined the set of probabilities:

$$\{\mathbf{P}(a = x|C) : a \in \mathcal{O}\}.$$

We complete this probabilistic data for the context C by contextual probabilities with respect to the contexts C_y corresponding to the selections $[b = y]$ for all observables $b \in \mathcal{O}$. The corresponding collection of data $D(\mathcal{O}, C)$ consists of contextual probabilities:

$$\mathbf{P}(a = x|C), \mathbf{P}(b = y|C), \mathbf{P}(a = x|C_y), \mathbf{P}(b = y|C_x), ...,$$

where $a, b, ... \in \mathcal{O}$. Finally, we denote the family of probabilistic data $D(\mathcal{O}, C)$ for all contexts $C \in \mathcal{C}$ by the symbol $\mathcal{D}(\mathcal{O}, \mathcal{C}) (\equiv \cup_{C \in \mathcal{C}} D(\mathcal{O}, C))$.

Definition 1. (Växjö Model) *An observational contextual statistical model of reality is a triple*

$$M = (\mathcal{C}, \mathcal{O}, \mathcal{D}(\mathcal{O}, \mathcal{C})) \tag{1}$$

where $\mathcal{C}$ is a set of contexts and $\mathcal{O}$ is a set of observables which satisfy to axioms 1,2, and $\mathcal{D}(\mathcal{O}, \mathcal{C})$ is probabilistic data about contexts $\mathcal{C}$ obtained with the aid of observables belonging $\mathcal{O}$.

We call observables belonging the set $\mathcal{O} \equiv \mathcal{O}(M)$ *reference of observables.* Inside of a model M observables belonging to the set $\mathcal{O}$ give the only possible references about a context $C \in \mathcal{C}$.

Contextual Model and Ignorance of Information

Probabilities $\mathbf{P}(b = y|C)$ are interpreted as *contextual (conditional) probabilities.* We emphasize that we consider conditioning not with respect to *events* as it is typically done in classical probability (Kolmogoroff 1933), but conditioning with respect to contexts – complexes of (e.g., physical, biological, social, mental, genetic, economic, or financial) conditions. This is the crucial point.

On the set of all events one can always introduce the structure of the *Boolean algebra* (or more general σ-algebra). In particular, for any two events A and B their set-theoretic intersection $A \cap B$ is well defined and it determines a new event: the simultaneous occurrence of the events A and B.

In contract to such an event-conditioning picture, if one have two contexts, e.g., complexes of physical conditions C_1 and C_2 and if even it is possible to create the set-theoretic representation of contexts (as some collections of physical parameters), then, nevertheless, their set-theoretic intersection $C_1 \cap C_2$ (although it is well defined mathematically) need not correspond to any physically meaningful context. Physical contexts were taken just as examples. The same is valid for social, mental, economic, genetic and any other type of contexts.

Therefore even if for some model M we can describe contexts in the set-theoretic framework, there are no reasons to assume that the collection of all contexts $\mathcal{C}$ should form a σ-algebra (Boolean algebra). This is the main difference from the classical (noncontextual) probability theory (Kolmogoroff 1933).

One can consider the same problem from another perspective. Suppose that we have some set of parameters Ω (e.g., physical, or social, or mental). We also assume that contexts are represented by some subsets of Ω. We consider two levels of description. At the first level a lot of information is available. There is a large set of contexts, we can even assume that they form a σ-algebra of subsets $\mathcal{F}$. We call them the first level contexts. There is a large number of observables at the first level, say the set of all possible random variables $\xi : \Omega \to \mathbf{R}$ (here $\mathbf{R}$ is the real line). By introducing on $\mathcal{F}$ a probability measure $\mathbf{P}$ we obtain the classical Kolmogorov probability model $(\Omega, \mathcal{F}, \mathbf{P})$, see (Kolmogoroff 1933). This is the end of the classical story about the probabilistic description of reality. Such a model is used e.g. in classical statistical physics.

We point our that any Kolmogorov probability model induces a Växjö model in such a way: a) contexts are given by all sets $C \in \mathcal{F}$ such that $\mathbf{P}(C) \neq 0$; b) the set of observables coincides with the set of all possible random variables; c) contextual probabilities are defined as Kolmogorovian conditional probabilities, i.e., by the Bayes formula: $\mathbf{P}(a = x|C) = \mathbf{P}(\omega \in C : a(\omega) = x)/\mathbf{P}(C)$. This is the Växjö model for the first level of description.

Consider now the second level of description. Here we can obtain a *non-Kolmogorovian Växjö model.* At this level only a part of information about the first level Kolmogorovian model $(\Omega, \mathcal{F}, \mathbf{P})$ can be obtained through a special family of observables $\mathcal{O}$ which correspond to a special subset of the set of all random variables of the Kolmogorov model $(\Omega, \mathcal{F}, \mathbf{P})$ at the first level of description. Roughly speaking not all contexts of the first level, $\mathcal{F}$ can be "visible" at the second level. There is no sufficiently many observables "to see" all contexts of the first level – elements of the Kolmogorov σ-algebra $\mathcal{F}$. Thus we should cut off this σ-algebra $\mathcal{F}$ and obtain a smaller family, say $\mathcal{C}$, of visible contexts. Thus some Växjö models (those permitting a set-theoretic representation) can appear starting with the purely classical Kolmogorov probabilistic framework, as a consequence of ignorance of information. If not all information is available, so we cannot use the first level (classical) description, then we, nevertheless, can proceed with the second level contextual description.

We shall see that starting with some Växjö models we can obtain the quantum-like calculus of probabilities in the complex Hilbert space. Thus in the opposition to a rather common opinion, we can derive a *quantum-like description for ordinary macroscopic systems* as the results of using of an incomplete representation. This opens great possibilities in application of quantum-like models outside the microworld. In particular, in cognitive science we need not consider composing of the brain from quantum particles to come to the quantum-like model.

Example 1. (Firefly in the box) Let us consider a box which is divided into four sub-boxes. These small boxes which are denoted by $\omega_1, \omega_2, \omega_3, \omega_4$ provides the the first level of description. We consider a Kolmogorov probability space: $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, the algebra of all finite subsets $\mathcal{F}$ of Ω and a probability measure determined by probabilities $\mathbf{P}(\omega_j) = p_j$, where $0 < p_j < 1, p_1 + ... + p_4 = 1$. We remark that in our interpretation it is more natural to consider elements of Ω as *elementary parameters,* and not as *elementary events* (as it was done by Kolmogorov).

We consider two different disjoint partitions of the set Ω :

$$A_1 = \{\omega_1, \omega_2\}, A_2 = \{\omega_3, \omega_4\},$$
$$B_1 = \{\omega_1, \omega_4\}, B_1 = \{\omega_2, \omega_3\}.$$

We can obtain such partitions by dividing the box: a) into two equal parts by the vertical line: the left-hand part gives A_1 and the right-hand part A_2; b) into two equal parts by the horizontal line: the top part gives B_1 and the bottom part B_2.

We introduce two random variables corresponding to these partitions: $\xi_a(\omega) = x_i$, if $\omega \in A_i$ and $\xi_b(\omega) = y_i \in$ if $\omega \in B_i$. Suppose now that we are able to measure only these two variables, denote the corresponding observables by the symbols a and b. We project the Kolmogorov model under consideration to a non-Kolmogorovian Växjö model by using the observables a and b – the second level of description. At this level the set of observables $\mathcal{O} = \{a, b\}$ and the natural set of contexts $\mathcal{C} : \Omega, A_1, A_2, B_1, B_2, C_1 = \{\omega_1, \omega_3\}, C_1 = \{\omega_2, \omega_4\}$ and all unions of these sets. Here "natural" has the meaning permitting a quantum-like representation (see further considerations). Roughly speaking contexts of the second level of description should be large enough to "be visible" with the aid of observables a and b.

Intersections of these sets need not belong to the system of contexts (nor complements of these sets). Thus the Boolean structure of the original first level description disappeared, but, nevertheless, it is present in the latent form. Point-sets $\{\omega_j\}$ are not "visible" at this level of description. For example, the random variable

$$\eta(\omega_j) = \gamma_j, j = 1, ..., 4, \ \gamma_i \neq \gamma_j, i \neq j,$$

is not an observable at the second level.

Such a model was discussed from positions of quantum logic, see, e.g., (Svozil 1998). There can be provided a nice interpretation of these two levels of description. Let us consider a firefly in the box. It can fly everywhere in this box. Its locations are described by the uniform probability distribution $\mathbf{P}$ (on the σ-algebra of Borel subsets of the box). This is the first level of description. Such a description can be realized if the box were done from glass or if at every point of the box there were a light detector. All Kolmogorov random variables can be considered as observables.

Now we consider the situation when there are only two possibilities to observe the firefly in the box: 1) to open a small window at a point a which is located in such a way that it is possible to determine only either the firefly is in the section A_1 or in the section A_2 of the box; 2) to open a small window at a point b which is located in such a way that it is possible to determine only either the firefly is in the section B_1 or in the section B_2 of the box. In the first case I can determine in which part, A_1 or A_2, the firefly is located. In the second case I also can only determine in which part, B_1 or B_2, the firefly is located. But I am not able to look into both windows simultaneously. In such a situation the observables a and b are the only source of information about the firefly (reference observables). The Kolmogorov description is meaningless (although it is incorporated in the model in the latent form). Can one apply a quantum-like description, namely, represent contexts by complex probability amplitudes? The answer is to be positive. The set of contexts that permit the quantum-like representation consists of all subsets C such that $\mathbf{P}(A_i|C) > 0$ and $\mathbf{P}(B_i|C) > 0, i = 1, 2$ (i.e., for sufficiently large contexts). We have seen that the Boolean structure disappeared as a consequence of ignorance of information.

Finally, we emphasize again that the Växjö model is essentially more general. The set-theoretic representation need not exist at all.

Boolean and quantum logic

Typically the absence of the Boolean structure on the set of quantum propositions is considered as the violation of laws of classical logic, e.g., in quantum mechanics (Birkhoff & von Neumann 1936). In our approach classical logic is not violated, it is present in the latent form. However, we are not able to use it, because we do not have complete information. Thus quantum-like logic is a kind of projection of classical logic. The impossibility of operation with complete information about a system is not always a disadvantages. Processing of incomplete set of information has the evident advantage comparing with "classical Boolean" complete information processing – the great saving of computing resources and increasing of the speed of computation. However, the Boolean structure cannot be violated in an arbitrary way, because in such a case we shall get a chaotic computational process. There should be developed some calculus of consistent ignorance by information. Quantum formalism provides one of such calculi.

Of course, there are no reasons to assume that processing of information through ignoring of its essential part should be rigidly coupled to a special class of physical systems, so called quantum systems. Therefore we prefer to speak about *quantum-like processing of information* that may be performed by various kinds of physical and biological systems. In our approach quantum computer has advantages not because it is based on a special class of physical systems (e.g., electrons or ions), but because there is realized the consistent processing of incomplete information. We prefer to use the terminology QL-computer by reserving the "quantum computer" for a special class of QL-computers which are based on quantum physical systems.

One may speculate that some biological systems could develop in the process of evolution the possibility to operate in a consistent way with incomplete information. Such a QL-processing of information implies evident advantages. Hence, it might play an important role in the process of the natural selection. It might be that consciousness is a form of the QL-presentation of information. In such a way we really came back to Whitehead's analogy between quantum and conscious systems (Whitehead 1929).

Supplementary ("Incompatible") Observables in the Växjö Model

Nowadays the notion of incompatible (complementary) observables is rigidly coupled to *noncommutativity.* In the conventional quantum formalism observables are incompatible iff they are represented by noncommuting self-adjoint operators $\hat{a}$ and $\hat{b}$: $[\hat{a}, \hat{b}] \neq 0$. As we see, the Växjö model is not from the very beginning coupled to a representation of information in a Hilbert space. Our aim is to generate an analogue (may be not direct) of the notion of incompatible (complementary) observables starting not from the mathematical formalism of quantum mechanics, but on the basis of the Växjö model, i.e., directly from statistical data.

Why do I dislike the conventional identification of *incompatibility with noncommutativity?* The main reason is that typically the mathematical formalism of quantum mechanics is identified with it as a physical theory. Therefore the quantum incompatibility represented through noncommutativity is rigidly coupled to the micro-world. (The only possibility to transfer quantum behavior to the macro-world is to consider physical states of the Bose-Einstein condensate type.) We shall see that some Växjö models can be represented as the conventional quantum model in the complex Hilbert space. However, the Växjö model is essentially more general than the quantum model. In particular, some Växjö models can be represented not in the complex, but in hyperbolic Hilbert space (the Hilbert module over the two dimensional Clifford algebra with the generator $j : j^2 = +1$).

Another point is that the terminology – incompatibility – is misleading in our approach. The quantum mechanical meaning of compatibility is the possibility to measure two observables, a and b *simultaneously.* In such a case they are represented by commuting operators. Consequently incompatibility implies the impossibility of simultaneous measurement of a and b. In the Växjö model there is no such a thing as fundamental impossibility of simultaneous measurement. We present the viewpoint that quantum incompatibility is just a consequence of information *supplementarity* of observables a and b. The information which is obtained via a measurement of, e.g., b can be non trivially updated by additional information which is contained in the result of a measurement of a. Roughly speaking if one knows a value of b, say $b = y$, this does not imply knowing the fixed value of a and vice versa, see (Khrennikov 2005) for details.

We remark that it might be better to use the notion "complementary," instead of "supplementary." However, the first one was already reserved by Nils Bohr for the notion which very close to "incompatibility." In any event Bohr's complementarity implies *mutual exclusivity* that was not the point of our considerations.

Supplementary processes take place not only in physical micro-systems. For example, in the brain there are present supplementary mental processes. Therefore the brain is a (macroscopic) QL-system. Similar supplementary processes take place in economy and in particular at financial market. There one could also use quantum-like descriptions (Choustova 2006). But the essence of the quantum-like descriptions is not the representation of information in the complex Hilbert space, but incomplete (projection-type) representations of information. It seems that the Växjö model provides a rather general description of such representations.

We introduce a notion of supplementary which will produce in some cases the quantum-like representation of observables by noncommuting operators, but which is not identical to incompatibility (in the sense of impossibility of simultaneous observations) nor complementarity (in the sense of mutual exclusivity).

Definition 2. *Let $a, b \in \mathcal{O}$. The observable a is said to be supplementary to the observable b if*

$$p^{a|b}(x|y) \neq 0, \tag{2}$$

for all $x \in X_a, y \in X_b$.

Let $a = x_1, x_2$ and $b = y_1, y_2$ be two dichotomous observables. In this case (2) is equivalent to the condition:

$$p^{a|b}(x|y) \neq 1, \tag{3}$$

for all $x \in X_a, y \in X_b$. Thus by knowing the result $b = y$ of the b-observation we are not able to make the definite prediction about the result of the a-observation.

Suppose now that (3) is violated (i.e., a is not supplementary to b), for example:

$$p^{a|b}(x_1|y_1) = 1, \tag{4}$$

and, hence, $p^{a|b}(x_2|y_1) = 0$. Here the result $b = y_1$ determines the result $a = x_1$.

In future we shall consider a special class of Växjö models in that the matrix of transition probabilities $\mathbf{P}^{a|b} = (p^{a|b}(x_i|y_j))_{i,j=1}^2$ is *double stochastic:* $p^{a|b}(x_1|y_1) + p^{a|b}(x_1|y_2) = 1; p^{a|b}(x_2|y_1) + p^{a|b}(x_2|y_2) = 1$. In such a case the condition (4) implies that

$$p^{a|b}(x_2|y_2) = 1, \tag{5}$$

and, hence, $p^{a|b}(x_1|y_2) = 0$. Thus also the result $b = y_2$ determines the result $a = x_2$.

We point out that for models with double stochastic matrix $\mathbf{P}^{a|b} = (p^{a|b}(x_i|y_j))_{i,j=1}^2$ the relation of supplementary is symmetric! In general it is not the case. It can happen that a is supplementary to b : each a-measurement gives us additional information updating information obtained in a preceding measurement of b (for any result $b = y$). But b can be non-supplementary to a.

Let us now come back to Example 1. The observables a and b are supplementary in our meaning. Consider now the classical Kolmogorov model and suppose that we are able to measure not only the random variables ξ_a and ξ_b – observables a and b, but also the random variable η. We denote the corresponding observable by d. The pairs of observables (d, a) and (d, b) are non-supplementary:

$$p^{a|d}(x_1|\gamma_i) = 0,\ i = 3, 4;\ p^{a|d}(x_2|\gamma_i) = 0,\ i = 1, 2,$$

and, hence,

$$p^{a|d}(x_1|\gamma_i) = 1,\ i = 1, 2;\ p^{a|d}(x_2|\gamma_i) = 1,\ i = 3, 4.$$

Thus if one knows , e.g., that $d = \gamma_1$ then it is definitely that $a = x_1$ and so on.

Test of Quantum-like Structure

We consider examples of cognitive contexts:

1). C can be some selection procedure that is used to select a special group S_C of people or animals. Such a context is represented by this group S_C (so this is an ensemble of cognitive systems). For example, we select a group $S_{\mathrm{prof.math.}}$ of professors of mathematics (and then ask questions a or (and) b or give corresponding tasks). We can select a group of people of some age. We can select a group of people having a "special mental state": for example, people in love or hungry people (and then ask questions or give tasks).

2). C can be a learning procedure that is used to create some special group of people or animals. For example, rats can be trained to react to special stimulus.

We can also consider *social contexts.* For example, social classes: proletariat-context, bourgeois-context; or war-context, revolution-context, context of economic depression, poverty-context, and so on. Thus our model can be used in social and political sciences (and even in history). We can try to find quantum-like statistical data in these sciences.

We describe a mental interference experiment.

Let $a = x_1, x_2$ and $b = y_1, y_2$ be two dichotomous mental observables: x_1=yes, x_2=no, y_1=yes, y_2=no. We set $X \equiv X_a = \{x_1, x_2\}, Y \equiv X_b = \{y_1, y_2\}$ ("spectra" of observables a and b). Observables can be two different questions or two different types of cognitive tasks. We use these two fixed reference observables for probabilistic representation of cognitive contextual reality given by C.

We perform observations of a under the complex of cognitive conditions C :

$$p^a(x) = \frac{\text{the number of results } a = x}{\text{the total number of observations}}.$$

So $p^a(x)$ is the probability to get the result x for observation of the a under the complex of cognitive conditions C. In the same way we find probabilities $p^b(y)$ for the b-observation under the same cognitive context C.

As was supposed in axiom 1, cognitive contexts C_y can be created corresponding to selections with respect to fixed values of the b-observable. The context C_y (for fixed $y \in Y$) can be characterized in the following way. By measuring the b-observable under the cognitive context C_y we shall obtain the answer $b = y$ with probability one. We perform now the a-measurements under cognitive contexts C_y for $y = y_1, y_2$, and find the probabilities:

$$p^{a|b}(x|y) = \frac{\text{number of the result } a = x \text{ for context } C_y}{\text{number of all observations for context } C_y}$$

where $x \in X, y \in Y$. For example, by using the ensemble approach to probability we have that the probability $p^{a|b}(x_1|y_2)$ is obtained as the frequency of the answer $a = x_1 = yes$ in the ensemble of cognitive system that have already answered $b = y_2 = no$. Thus we first select a sub-ensemble of cognitive systems who replies no to the b-question, $C_{b=no}$. Then we ask systems belonging to $C_{b=no}$ the a-question.

It is assumed (and this is a very natural assumption) that a cognitive system is "responsible for her (his) answers." Suppose that a system τ has answered $b = y_2 = no$. If we ask τ again the same question b we shall get the same answer $b = y_2 = no$. This is nothing else than *the mental form of the von Neumann projection postulate:* the second measurement of the same observable, performed immediately after the first one, will yield the same value of the observable).

Classical probability theory tells us that all these probabilities have to be connected by the so called *formula of total probability:*

$$p^a(x) = p^b(y_1)p^{a|b}(x|y_1) + p^b(y_2)p^{a|b}(x|y_2),\quad x \in X.$$

However, if the theory is quantum-like, then we should obtain (Khrennikov 2004) the formula of total probability with an interference term:

$$p^a(x) = p^b(y_1)p^{a|b}(x|y_1) + p^b(y_2)p^{a|b}(x|y_2) \tag{6}$$

$$+2\lambda(a = x|b, C)\sqrt{p^b(y_1)p^{a|b}(x|y_1)p^b(y_2)p^{a|b}(x|y_2)},$$

where the coefficient of supplementarity (the coefficient of interference) is given by $\lambda(a = x|b, C) =$

$$\frac{p^a(x) - p^b(y_1)p^{a|b}(x|y_1) - p^b(y_2)p^{a|b}(x|y_2)}{2\sqrt{p^b(y_1)p^{a|b}(x|y_1)p^b(y_2)p^{a|b}(x|y_2)}} \tag{7}$$

This formula holds true for *supplementary observables.* To prove its validity, it is sufficient to put the expression for

$\lambda(a=x|b,C)$, see (7), into (6). In the quantum-like statistical test for a cognitive context C we calculate
$\lambda(a=x|b,C)=$

$$\frac{p^a(x)-p^b(y_1)p^{a|b}(x|y_1)-p^b(y_2)p^{a|b}(x|y_2)}{2\sqrt{p^b(y_1)p^{a|b}(x|y_1)p^b(y_2)p^{a|b}(x|y_2)}}.$$

An empirical situation with $\lambda(a=x|b,C)\neq 0$ would yield evidence for quantum-like behaviour of cognitive systems. In this case, starting with (experimentally calculated) coefficient of interference $\lambda(a=x|b,C)$ we can proceed either to the conventional Hilbert space formalism (if this coefficient is bounded by 1) or to so called hyperbolic Hilbert space formalism (if this coefficient is larger than 1). In the first case the coefficient of interference can be represented in the trigonometric form $\lambda(a=x|b,C)=\cos\theta(x)$, Here $\theta(x)\equiv\theta(a=x|b,C)$ is the phase of the a-interference between cognitive contexts C and $C_y, y\in Y$. In this case we have the conventional formula of total probability with the interference term:

$$p^a(x)=p^b(y_1)p^{a|b}(x|y_1)+p^b(y_2)p^{a|b}(x|y_2) \tag{8}$$
$$+2\cos\theta(x)\sqrt{p^b(y_1)p^{a|b}(x|y_1)p^b(y_2)p^{a|b}(x|y_2)}.$$

In principle, it could be derived in the conventional Hilbert space formalism. But we chosen the inverse way. Starting with (8) we could introduce a "mental wave function" $\psi\equiv\psi_C$ (or pure quantum-like mental state) belonging to this Hilbert space. We recall that in our approach a mental wave function ψ is just a representation of a cognitive context C by a complex probability amplitude. The latter provides a Hilbert representation of statistical data about context which can be obtained with the help of two fixed observables (reference observables).

Wave Function Representation of Cognitive Contexts

Let C be a cognitive context. We consider only cognitive contexts with trigonometric interference for *supplementary mental observables* a and b. The interference formula of total probability (6) can be written in the following form: $p_C^a(x)=$

$$\sum_{y\in Y}p_C^b(y)p^{a|b}(x|y)+2\cos\theta_C(x)\sqrt{\Pi_{y\in Y}p_C^b(y)p^{a|b}(x|y)} \tag{9}$$

By using the elementary formula: $D=A+B+2\sqrt{AB}\cos\theta=|\sqrt{A}+e^{i\theta}\sqrt{B}|^2, A,B>0$, we can represent the probability $p_C^b(x)$ as the square of the complex amplitude:

$$p_C^a(x)=|\psi_C(x)|^2 \tag{10}$$

where

$$\psi(x)\equiv\psi_C(x)=\sum_{y\in Y}\sqrt{p_C^b(y)p^{a|b}(x|y)}e^{i\xi_C(x|y)}. \tag{11}$$

Here phases $\xi_C(x|y)$ are such that $\xi_C(x|y_1)-\xi_C(x|y_2)=\theta_C(x)$. We denote the space of functions: $\psi:X\to\mathbf{C}$ by the symbol $E=\Phi(X,\mathbf{C})$. Since $X=\{x_1,x_2\}$, the E is the two dimensional complex linear space. Dirac's $\delta-$functions $\{\delta(x_1-x),\delta(x_2-x)\}$ form the canonical basis in this space. For each $\psi\in E$ we have $\psi(x)=\psi(x_1)\delta(x_1-x)+\psi(x_2)\delta(x_2-x)$.

Denote by the symbol $\mathcal{C}^{\rm tr}$ the set of all cognitive contexts having the trigonometric statistical behaviour (i.e., $|\lambda|\leq 1$) with respect to mental observables a and b. By using the representation (11) we construct the map $\tilde{J}^{a|b}:\mathcal{C}^{\rm tr}\to\tilde{\Phi}(X,\mathbf{C})$, where $\tilde{\Phi}(X,\mathbf{C})$ is the space of equivalent classes of functions under the equivalence relation: φ equivalent ψ iff $\varphi=t\psi, t\in\mathbf{C},|t|=1$. We point out that if the matrix of transition probabilities for the reference observables is double stochastic, then $a|b$-representation is equivalent to the $b|a$-representation. In general it is not the case.

Quantum-like Processing of Information in Brain

The brain is a huge information system that contains millions of elementary mental states . It could not "recognize" (or "feel") all those states at each instant of time t. Our fundamental hypothesis is that the brain is able to create the QL-representations of mind. At each instant of time t the brain creates the QL-representation of its mental context C based on two supplementary mental (self-)observables a and b. Here $a=(a_1,...,a_n)$ and $b=(b_1,...,b_n)$ can be very long vectors of compatible (non-supplementary) dichotomous observables. The reference observables a and b can be chosen (by the brain) in different ways at different instances of time. Such a change of the reference observables is known in cognitive sciences as a *change of representation.*

A mental context C in the $a|b-$ representation is described by the mental wave function ψ_C. We can speculate that the brain has the ability to feel this mental field as a distribution on the space X. This distribution is given by the norm-squared of the mental wave function: $|\psi_C(x)|^2$.

In such a model it might be supposed that the state of our consciousness is represented by the mental wave function ψ_C. The crucial point is that in this model consciousness is created through neglecting an essential volume of information contained in subconsciousness. Of course, this is not just a random loss of information. Information is selected through the algorithm of the probabilistic representation, see (11): a mental context C is projected onto the complex probability amplitude ψ_C.

The (classical) mental state of sub-consciousness evolves with time $C\to C(t)$. This dynamics induces dynamics of the mental wave function $\psi(t)=\psi_{C(t)}$ in the complex Hilbert space.

Further development of our approach (which we are not able to present here) induces the following model of brain's functioning (Khrennikov 2006a):

The brain is able to create the QL-representation of mental contexts, $C\to\psi_C$ (by using the algorithm based on the formula of total probability with interference).

Brain as Quantum-like Computer

The ability of the brain to create the QL-representation of mental contexts induces functioning of the brain as a quantum-like computer.

The brain performs computation-thinking by using algorithms of quantum computing in the complex Hilbert space of mental QL-states.

We emphasize that in our approach the brain is not quantum computer, but a QL-computer. On one hand, a QL-computer works totally in accordance with the mathematical theory of quantum computations (so by using quantum algorithms). On the other hand, it is not based on superposition of individual mental states. The complex amplitude ψ_C representing a mental context C is a special probabilistic representation of information states of the huge neuronal ensemble. In particular, the brain is a *macroscopic* QL-computer. Thus the QL-parallelism (in the opposite to conventional quantum parallelism) has a natural realistic base. This is real parallelism in the working of millions of neurons. The crucial point is the way in which this classical parallelism is projected onto dynamics of QL-states. The QL-brain is able to solve NP-problems. But there is nothing mysterious in this ability: an exponentially increasing number of operations is performed through involving of an exponentially increasing number of neurons.

We point out that by coupling QL-parallelism to working of neurons we started to present a particular ontic model for QL-computations. We shall discuss it in more detail. Observables a and b are self-observations of the brain. They can be represented as functions of the internal state of brain ω. Here ω is a parameter of huge dimension describing states of all neurons in the brain: $\omega = (\omega_1, \omega_2, ..., \omega_N)$:

$$a = a(\omega), b = b(\omega).$$

The brain is not interested in concrete values of the reference observables at fixed instances of time. The brain finds the contextual probability distributions $p_C^a(x)$ and $p_C^b(y)$ and creates the mental QL-state $\psi_C(x)$, see the QL-representation algorithm (11). Then it works with the mental wave function $\psi_C(x)$ by using algorithms of quantum computing.

Two Time Scales as the Basis of the QL-representation of Information

The crucial problem is to find a mechanism for producing contextual probabilities. We think that they are frequency probabilities that are created in the brain in the following way.There are two scales of time: a) internal scale, τ-time; b) QL-scale, t-time. The internal scale is *finer than the QL-scale.* Each instant of QL-time t corresponds to an interval Δ of internal time τ. We might identify the QL-time with mental (psychological) time and the internal time with physical time. We shall also use the terminology: pre-cognitive time-scale - τ and cognitive time-scale - t.

During the interval Δ of internal time the brain collects statistical data for self-observations of a and b. The internal state ω of the brain evolves as

$$\omega = \omega(\tau, \omega_0).$$

This is a classical dynamics (which can be described by a stochastic differential equation).

At each instance of internal time τ there are performed nondisturbative self-measurements of a and b. These are realistic measurements: the brain gets values $a(\omega(\tau, \omega_0))$, $b(\omega(\tau, \omega_0))$. By finding frequencies of realization of fixed values for $a(\omega(\tau, \omega_0))$ and $b(\omega(\tau, \omega_0))$ during the interval Δ of internal time, the brain obtains the frequency probabilities $p_C^a(x)$ and $p_C^b(y)$. These probabilities are related to the instant of QL-time time t corresponding to the interval of internal time $\Delta : p_C^a(t, x)$ and $p_C^b(t, y)$. We remark that in these probabilities the brain encodes huge amount of information – millions of mental "micro-events" which happen during the interval Δ. But the brain is not interested in all those individual events. (It would be too disturbing and too irrational to take into account all those fluctuations of mind.) It takes into account only the integral result of such a *pre-cognitive activity* (which was performed at the pre-cognitive time scale).

For example, the mental observables a and b can be measurements over different domains of brain. It is supposed that the brain can "feel" probabilities (frequencies) $p_C^a(x)$ and $p_C^b(y)$, but not able to "feel" the simultaneous probability distribution $p_C(x, y) = P(a = x, b = y|C)$.

This is not the problem of mathematical existence of such a distribution. This is the problem of integration of statistics of observations from different domains of the brain. By using the QL-representation based only on probabilities $p_C^a(x)$ and $p_C^b(y)$ the brain could be able to escape integration of information about *individual self-observations* of variables a and b related to spatially separated domains of brain. The brain need not couple these domains at each instant of internal (pre-cognitive time) time τ. It couples them only once in the interval Δ through the contextual probabilities $p_C^a(x)$ and $p_C^b(y)$. This induces the huge saving of time and increasing of speed of processing of mental information.

One of fundamental consequences for cognitive science is that our mental images have the probabilistic structure. They are products of transition from an extremely fine pre-cognitive time scale to a rather rough cognitive time scale.

Finally, we remark that a similar time scaling approach was developed in (Khrennikov 2006b) for ordinary quantum mechanics. In (Khrennikov 2006b) quantum expectations appear as results of averaging with respect to a prequantum time scale. There was presented an extended discussion of possible choices of quantum and prequantum time scales.

We can discuss the same problem in the cognitive framework. We may try to estimate the time scale parameter Δ of the neural QL-coding. There are strong experimental evidences, see, e.g., (Mori 2002), that a moment in psychological time correlates with ≈ 100 ms of physical time for neural activity. In such a model the basic assumption is that the physical time required for the transmission of information over synapses is somehow neglected in the psychological time. The time (≈ 100 ms) required for the transmission of information from retina to the inferiotemporal cortex (IT) through the primary visual cortex (V1) is mapped to a moment of psychological time. It might be that by using

$$\Delta \approx 100\mathrm{ms}$$

we shall get the right scale of the QL-coding.

However, it seems that the situation is essentially more complicated. There are experimental evidences that the temporial structure of neural functioning is not homogeneous. The time required for completion of color information in V4 ($\approx$ 60 ms) is shorter that the time for the completion of shape analysis in IT ($\approx$ 100 ms). In particular it is predicted that there will be under certain conditions a rivalry between color and form perception. This rivalry in time is one of manifestations of complex level temporial structure of brain. There may exist various pairs of scales inducing the QL-representations of information.

References

Birkhoff, G., and von Neumann, J. 1936. The logic of quantum mechanics. *Ann. Math.* 37:823–643.

Choustova, O. 2006. Quantum bohmian model for financial market. *Physica A* 374:304–314.

Khrennikov, A. Y. 2004. *Information Dynamics in Cognitive, Psychological and Anomalous Phenomena*. Dordreht: Kluwer Academic.

Khrennikov, A. Y. 2005. The principle of supplementarity: A contextual probabilistic viewpoint to complementarity, the interference of probabilities, and the incompatibility of variables in quantum mechanics. *Foundations of Physics* 35(10):1655 – 1693.

Khrennikov, A. Y. 2006a. Quantum-like brain: Interference of minds. *BioSystems* 84:225–241.

Khrennikov, A. Y. 2006b. To quantum mechanics through random fluctuations at the planck time scale. *http://www.arxiv.org/abs/hep-th/0604011*.

Kolmogoroff, A. N. 1933. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Berlin: Springer Verlag.

Mackey, G. W. 1963. *Mathematical Foundations of Quantum Mechanics*. New-York: W. A. Benjamin Inc.

Mori, K. 2002. On the relation between physical and psychological time. In *Toward a Science of Consciousness*. Tucson, Arizona: Tucson University Press. 102.

Svozil, K. 1998. *Quantum Logic*. Berlin: Springer.

Whitehead, A. N. 1929. *Process and Reality: An Essay in Cosmology*. New York: Macmillan Publishing Company.