

Concepts and Bounded Rationality: An Application of Niestegge's Approach to Conditional Quantum Probabilities

Reinhard Blutner

Universiteit van Amsterdam, ILLC, Nieuwe Doelenstraat 15, 1012 CP Amsterdam, The Netherlands

Abstract. Recently, Gerd Niestegge developed a new approach to quantum mechanics via conditional probabilities developing the well-known proposal to consider the Lüders-von Neumann measurement as a non-classical extension of probability conditionalization. I will apply his powerful and rigorous approach to the treatment of concepts using a geometrical model of meaning. In this model, instances are treated as vectors of a Hilbert space \mathcal{H} . In the present approach there are at least two possibilities to form categories. The first possibility sees categories as a mixture of its instances (described by a density matrix). In the simplest case we get the classical probability theory including the Bayesian formula. The second possibility sees categories formed by a distinctive prototype which is the superposition of the (weighted) instances. The construction of prototypes can be seen as transferring a mixed quantum state into a pure quantum state freezing the probabilistic characteristics of the superposed instances into the structure of the formed prototype. Closely related to the idea of forming concepts by prototypes is the existence of interference effects. Such interference effects are typically found in macroscopic quantum systems and I will discuss them in connection with several puzzles of bounded rationality. The present approach nicely generalizes earlier proposals made by authors such as Diederik Aerts, Andrei Khrennikov, Ricardo Franco, and Jerome Busemeyer. Concluding, I will suggest that an active dialogue between cognitive approaches to logic and semantics and the modern approach of quantum information science is mandatory.

Keywords: Bayesian view, bounded rationality, cue validity, interference effects in cognition, orthoalgebra, prototype theory.

PACS:, 02.50.Cw, 07.05.Mh.

1. INTRODUCTION

For many years, the idea of rationality was predominantly seen in the context of logic. It has been claimed that the idea of being rational is simply the idea of following a system of logical rules. The view that rational thought is governed by logic can be called *the logicist conception of the mind* [following 1]. This view contrasts with *the Bayesian view* of rationality which is defined by certain standards to reason about uncertainty. Interestingly, the Bayesian approach has applied in a variety of areas of artificial intelligence and cognitive modelling including vision [2, 3], natural language parsing [4], concept-learning [5], categorization [6], cognition of musical structures [7], logical reasoning [8]. The great success of the Bayesian modelling approach has made it a dominant research area in the cognitive sciences within the last few years.

However, also the Bayesian approach suffers from puzzling aspects and apparent irrationalities which were convincingly demonstrated by a series of celebrated experiments [for overviews, see 9, 10].

In this paper, I will propose a general approach to modelling concepts that saves the advantage of the Bayesian approach but overcomes its shortcomings. The present approach considers (propositional) concepts as constituted by projection operators in a Hilbert-space framework. This approach allows to contrast and to compare two traditional views of concept formation within a uniform framework: the classical set-theoretic conception (relating to the Bayesian view) and the prototype view (dominantly taken by cognitive psychologists). I will argue that a kind of “freezing” mechanism relates between these two views, a mechanism that transfers a mixed quantum state into a

pure quantum state (in some sense, this mechanism can be seen as the opposite of the *measuring process* in orthodox quantum theory).

The present work is based on Gerd Niestegge's [11] conception of conditional probabilities in an ortho-algebraic framework. This framework is especially useful for integrating different views (such as the classical Bayesian view and the prototype view) within a single theory. Niestegge's framework is outlined in section 2. Section 3 discusses the application of his framework to two ways of modelling concepts. In section 4 I shortly introduce the notion of bounded rationality. Section 5 demonstrates how to overcome the puzzles described as "irrationalities" in the Bayesian framework in a completely rational way. Section 6 makes some remarks about the composition of concepts based on prototypes and on modelling typicality. Finally, section 7 concludes with an outline of further perspectives.

2. G. NIESTEGGE'S APPROACH TO CONDITIONAL PROBABILITIES

Gerd Niestegge's new approach to quantum mechanics via conditional probabilities develops the well-known proposal to consider the Lüders-von Neumann measurement as a non-classical extension of probability conditionalization.

Due to Gleason's theorem and its generalizations (for references see [12]) we have the following fact considering a Hilbert space \mathcal{H} with dimension > 2 : Every countably additive probability measure on a projection lattice of \mathcal{H} has a unique linear extension defined for all (bounded) operators on \mathcal{H} . Using this fact it is a simple exercise to prove that each additive probability measure μ on a projection lattice of \mathcal{H} (dimension > 2) can be represented by a density operator. Niestegge applied the very same fact to derive a new understanding of conditional probabilities in an orthoalgebraic framework.

For characterizing conditional probabilities we stipulate the following condition for all projections \mathbf{a}, \mathbf{b} in \mathcal{H} :

$$\mu(\mathbf{b}|\mathbf{a}) = \mu(\mathbf{b})/\mu(\mathbf{a}) \text{ if } \mathbf{b} \leq \mathbf{a} \text{ (i.e. } \mathbf{ba} = \mathbf{b}), \mu(\mathbf{a}) \neq 0 \quad (1)$$

This condition has a close pendant in the classical formulation of conditional probabilities [e.g. 12]. Whereas in the classical case the standard formulation of conditional probability can be derived – as probability of the conjunction of \mathbf{a} and \mathbf{b} normalized by $\mu(\mathbf{a})$ –, we get a slidely different formulation in the quantum case:

Fact 1: Let \mathbf{a} and \mathbf{b} be projection operators, then

$$\mu(\mathbf{b}|\mathbf{a}) = \mu(\mathbf{aba}|\mathbf{a}) = \mu(\mathbf{aba})/\mu(\mathbf{a}) \quad (2)$$

Proof: Using the decomposition $\mathbf{b} = \mathbf{aba} + \mathbf{a'ba} + \mathbf{aba}' + \mathbf{a'ba}'$ (where \mathbf{x}' denotes the orthogonal projection operator for \mathbf{x}) we get $\mu(\mathbf{b}|\mathbf{a}) = \mu(\mathbf{aba}|\mathbf{a}) + \mu(\mathbf{a'ba}|\mathbf{a}) + \mu(\mathbf{aba}'|\mathbf{a}) + \mu(\mathbf{a'ba}'|\mathbf{a})$. It is simple to see that $\mu(\mathbf{a'ba}|\mathbf{a}) = \mu(\mathbf{aba}'|\mathbf{a}) = \mu(\mathbf{a'ba}'|\mathbf{a}) = 0$ (e.g. $\mu(\mathbf{a'ba}|\mathbf{a}) \leq \mu(\mathbf{a}'|\mathbf{a}) = 0$, etc.). Thus, $\mu(\mathbf{b}|\mathbf{a}) = \mu(\mathbf{aba}|\mathbf{a}) = \mu(\mathbf{aba})/\mu(\mathbf{a})$.

I will use the notion $(\mathbf{a}; \mathbf{b})$ for a sequence of two projection operators (also called *asymmetric conjunction*). The probability for sequences of projections is defined as follows:

$$\mu(\mathbf{a}; \mathbf{b}) \equiv \mu(\mathbf{a}) \mu(\mathbf{b}|\mathbf{a}) = \mu(\mathbf{aba}) \quad (3)$$

If \mathbf{a} and \mathbf{b} commute we get $\mu(\mathbf{a}; \mathbf{b}) = \mu(\mathbf{b}; \mathbf{a})$, i.e. $\mu(\mathbf{a}) \mu(\mathbf{b}|\mathbf{a}) = \mu(\mathbf{b}) \mu(\mathbf{a}|\mathbf{b})$ (Bayesian formula). If \mathbf{a} and \mathbf{b} don't commute the Bayesian formula can be violated.

Next, we consider probabilistic conditioning in the quantum case, i.e. we calculate the probability for \mathbf{b} conditioning on \mathbf{a} and its orthocomplement \mathbf{a}' . This can be done by using the following fact:

Fact 2: Let \mathbf{a} and \mathbf{b} be projection operators, then

$$\mu(\mathbf{b}) = \mu(\mathbf{a}; \mathbf{b}) + \mu(\mathbf{a}'; \mathbf{b}) + \mu(\mathbf{aba}' + \mathbf{a'ba}) \quad (4)$$

Proof: $\mathbf{a} + \mathbf{a}' = \mathbf{b} + \mathbf{b}' = \mathbf{1}$; $\mathbf{aa}' = \mathbf{bb}' = \mathbf{0}$. Consequently, we can write $\mathbf{b} = \mathbf{aba} + \mathbf{a'ba} + \mathbf{aba}' + \mathbf{a'ba}'$. Using Gleason's theorem yields the result.

In equation (4), the term $\mu(\mathbf{a}\mathbf{b}\mathbf{a}' + \mathbf{a}'\mathbf{b}\mathbf{a})$ describes an *interference effect* which vanishes in the “classical case” of commuting operators. Let’s consider the interference term in case \mathbf{a} and \mathbf{a}' represent *pure states*:

Fact 3: Let \mathbf{a} and \mathbf{b} be projection operators of pure states, i.e., $\mathbf{a} = |a\rangle\langle a|$, $\mathbf{a}' = |a'\rangle\langle a'|$, $\mathbf{a}\mathbf{a}' = 0$, $\mathbf{a} + \mathbf{a}' = \mathbf{1}$. Then it holds:

$$\mu(\mathbf{a}\mathbf{b}\mathbf{a}' + \mathbf{a}'\mathbf{b}\mathbf{a}) = 2 \mu^{1/2}(\mathbf{a}; \mathbf{b}) \mu^{1/2}(\mathbf{a}'; \mathbf{b}) \cos(\delta) \quad (5)$$

Proof: $\mu(\mathbf{a}\mathbf{b}\mathbf{a}' + \mathbf{a}'\mathbf{b}\mathbf{a}) = \langle a|b\rangle\langle a'|b\rangle^* |a\rangle\langle a'| + \langle a|b\rangle^*\langle a'|b\rangle |a'\rangle\langle a|$. Using the factorization $\langle a|b\rangle\langle a'|b\rangle^* = |\langle a|b\rangle||\langle a'|b\rangle| \exp(i\delta)$ with phase angle δ we get the following result: $\mu(\mathbf{a}\mathbf{b}\mathbf{a}' + \mathbf{a}'\mathbf{b}\mathbf{a}) = 2 |\langle a|b\rangle| |\langle a'|b\rangle| \mu(|a\rangle\langle a'| \cdot \exp(i\delta) + |a'\rangle\langle a| \cdot \exp(-i\delta))$. This conforms to the final result $\mu(\mathbf{a}\mathbf{b}\mathbf{a}' + \mathbf{a}'\mathbf{b}\mathbf{a}) = 2 \mu^{1/2}(\mathbf{a}; \mathbf{b}) \mu^{1/2}(\mathbf{a}'; \mathbf{b}) \cos(\delta)$.

3. TWO WAYS OF REPRESENTING CONCEPTS

In this section we contrast the usual set theoretic conception of concepts with the prototype conceptions. According to the set theoretic conceptions a category is considered as a set of instances. In this paper we will consider instances as orthonormal basic vectors of a Hilbert space \mathcal{H} . This conforms to a geometrical (vector model) model of meanings [e.g. 13, 14]. In the following it is assumed that the basic instances which form a concept are independent from each other. Further, we assume that we can index these base vectors. I will use index sets such as A, B, C . Then we can form projection operators in the following way, e.g. for an index set A :

$$\mathbf{a} = \sum_{i \in A} |i\rangle\langle i| \quad (6)$$

In formal theories of semantics the conception of a possible world is essential [e.g. 15, 16]. Building propositional concepts can now be done by summing up the possible worlds (seen as vectors) that are considered to validate the propositional concept.

If we calculate the corresponding projection operators as in (6), then the probabilities of the ortho-lattice conform to the classical probabilities – assuming that the probabilities $\mu(|i\rangle\langle i|)$ can be identified with an initial probability distribution of the index set: $\mu(|i\rangle\langle i|) = \mu(i)$. Hence, we get:

$$\mu(\mathbf{a}) = \sum_{i \in A} \mu(|i\rangle\langle i|) = \sum_{i \in A} \mu(i) = \mu(A) \quad (7)$$

$$\mu(\mathbf{b}\mathbf{a}) = \mu(B|A)$$

Using the formalism of density matrixes we could apply $\rho = \sum_j \mu(j) |j\rangle\langle j|$ to obtain the same results. Hence, the underlying state is a mixed state (mixing the different instances of the concept).

The standard prototype idea sees categories formed by a single prototype. In the present vector model, this prototype is calculated by superposing all the instances corresponding to the relevant index set, say A . These instances have to be weighted with a corresponding weight factor $w(i)$ reflecting their frequency and other factors of use. The formal representation of the prototype concept $\underline{\mathbf{a}}$ results by projecting the corresponding sum vector:

$$\underline{\mathbf{a}} = |a\rangle\langle a| \text{ with } |a\rangle = \sum_{i \in A} w(i) |i\rangle \quad (8)$$

By stipulating $|w(i)|^2 = \mu(i)$, we get

$$\mu(\underline{\mathbf{a}}) = |\langle i|a\rangle|^2 = \mu(i|A) \quad (9)$$

$$\mu(\underline{\mathbf{b}}\underline{\mathbf{a}}) = |\langle a|b\rangle|^2 = \mu^{1/2}(B|A) \cdot \mu^{1/2}(A|B)$$

Hence, from the concept defined by a single prototype, we are able to reconstruct the probability $\mu(i|A)$ by calculating the absolute square of the scalar product between the instance vector i and the prototype state a . Since the conditional probability $\mu(\underline{\mathbf{b}}\underline{\mathbf{a}})$ is symmetric we cannot reconstruct the classical (asymmetric) conditional probabilities. What we get for $\mu(\underline{\mathbf{b}}\underline{\mathbf{a}})$ is the square root of the product $\mu(B|A) \cdot \mu(A|B)$.

In contrast to the mixed case discussed before, the value of the probabilities in the present case stems from the *algebraic relation* between the two events and not from any (initial) probability measure. Gerd Niestegge [12] calls this the *objective* conditional probability. The importance of the objective conditional probability lies in the fact that it only depends on the two measurements results for, say, \underline{a} and \underline{b} , but not on any initial state of the physical system. It is therefore not necessary to assume that a physical system is in a certain, perhaps unknown, state before the measurement testing \underline{a} starts. Moreover, $\mu(\underline{b}|\underline{a})$ is independent of any measuring apparatus or method. The objective conditional probabilities are transition probabilities between the results of successive measurements and depend on nothing else but the system properties that are tested with the measurements. Without the existence of objective conditional probabilities, the knowledge of the outcome of the first measurement \underline{a} is not sufficient for any prediction concerning the results of the future measurement \underline{b} . (cf. [12], p. 245).

The phenomenon that frequencies are important to construct structural objects is familiar from cognitive science and linguistics. The standard case is the construction of prototypes as a kind of mental representation reflecting the stochastic properties of our natural environment. For constructing these objects stochastic learning mechanisms are highly relevant. I will consider the general process as a *fossilization* phenomenon since the stochastic nature of the environment is frozen into an internal representation. Equation (9) illustrates how structural properties of some relevant vectors reflect the statistical properties of the environment.

There is another aspect that makes a difference for the two cases of concept formation, i.e. mixing instances versus superposing instances. This aspect has to do with the operator algebra. Whereas in the first case the two projection operators (when constructed from index sets A and B) always commute, this is not generally valid in the second case: \underline{a} and \underline{b} commute only if the prototype states a and b are orthogonal. This makes an important difference when it comes to discuss the combination of concepts and the appearance of interference effects.

Summarizing, we have two simple possibilities to form categories (or propositions if you consider possible worlds instead of objects). The first possibility sees categories as a set of instances: $\mathbf{a} = \sum_{i \in A} |i\rangle\langle i|$. In this case we get the classical probability theory including the Bayesian formula. For the base projections $\mathbf{i} = |i\rangle\langle i|$ probabilities are required which constitute a density matrix. Conditional probabilities are dependent on this initial distribution. The second possibility sees categories formed by a single prototype: $|a\rangle = \sum_{i \in A} w(i) |i\rangle$. Also in this case the correct probabilities of the instances (given A) can be reconstructed. However, these probabilities are objective ones and result from the frozen probabilistic characteristics of the superposed instances. Further, conditional probabilities are symmetric in the prototype case. Hence, the Bayesian formula cannot be satisfied. The most important idea we find in the prototype case is the existence of interference effects.

4. BOUNDED RATIONALITY

More than a half century again, Herbert Simon [e.g. 17] claimed that we have to replace the “global rationality of the economic man” with a kind of behaviour that is compatible with the boundedness of the Human decision maker and the particular kinds of environments in which such organisms exist. In the words of Simon (1955): “Boundedly rational agents experience limits in formulating and solving complex problems and in processing (receiving, storing, retrieving, transmitting) information”.

There is a number of dimensions along which “classical” models of rationality can be made more realistic without giving up rigorous formalization. These include (i) limiting what sorts of utility functions there might be; (ii) recognizing the costs of gathering and processing information, and (iii) the possibility of having a “multi-valued” utility function. Looking at the huge literature of recent approaches to bounded rationality [e.g. 9] suggest an adaptive toolbox of different approaches and models to handle the list of phenomena headed under the same title. From an explanatory point of view such an approach is not encouraging, and a more systematic approach would be highly welcome. It is the claim of the present paper that something fundamentally is wrong with the Boolean conception of probability that silently underlies this approach. There are some hard problems to the bounded rationality issues, such as the disjunction puzzle, the conjunction puzzle, the Ellsberg paradox, and the Allais paradox, which recently were reconsidered by a number of people demonstrating that the interference terms found in quantum probability theory can give a deeper grounding of these effects [e.g. 18, 19-25].

The approach taken in the present paper generalizes some of these earlier approaches. It is based on the new formulation to conditional probabilities developed by Gerd Niestegge (see section 2).

5. SOLVING SOME PUZZLES OF BOUNDED RATIONALITY

Let's start with the *disjunction puzzle* which relates to violations of Savage's sure-thing principle. This principle was described as follows: "A businessman contemplates buying a certain piece of property. He considers the outcome of the next presidential election relevant to the attractiveness of the purchase. So, to clarify the matter for himself, he asks whether he would buy if he knew that the Republican candidate were going to win, and decides that he would do so. Similarly, he considers whether he would buy if he knew that the Democratic candidate were going to win, and again finds that he would do so. Seeing that he would buy in either event, he decides that he should buy, even though he does not know which event obtains." [26, p.21]

In practical reasoning, however, people seldom follow the sure thing rationality principle. For example, Tversky and Shafir [27] show that significantly more students report they would purchase a non-refundable Hawaiian vacation if they were to know that they have passed or failed an important exam than report they would purchase if they were not to know the outcome of the exam. Conte et al. [25] gave a quantitative judgment of the deviations from the sure thing principle (= *disjunction effect*) in a series of experiments. In one of these experiments subjects had to decide whether the two objects in Test A (Test B) are of the same size or not. Test A and Test B are realized separately and in turn (first A then B after 2 seconds).

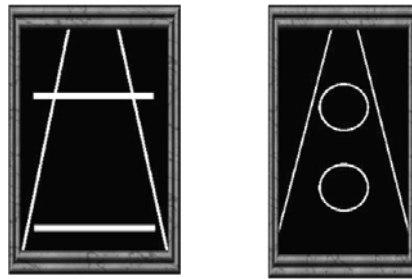


FIGURE 1. Test A (left) and Test B (right) of an experiment measuring the disjunction effect [from 25]

This experiment allows a measurement of the disjunction effect, which is defined as follows:

$$\nabla(A, B) = \mu(B) - [\mu(B|A) \mu(A) + \mu(B|\neg A) \mu(\neg A)] \quad (10)$$

In the present framework, when using equation (4), the disjunction effects relates to the interference term $\mu(\mathbf{a}b\mathbf{a}' + \mathbf{a}'b\mathbf{a})$, which is calculated in (5) for the pure case. Using the empirical results $\mu(B) = 0.45$, $\mu(A) = 0.7$, $\mu(B|A) = 0.43$, and $\mu(B|\neg A) = 1$ we observe a significant disjunction effect $\nabla(A, B) = 0.15$. From this result we obtain $\cos(\delta) = 0.25$.

Next, consider the conjunction effect [28]. In one of their experiments subjects are presented with a story such as the following one: "Linda is 31 years old, single, outspoken and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-nuclear demonstrations." After the presentation of the story the subjects are asked to assess the probabilities of several propositions on a number scale. I represent the critical propositions only (together with the averaged judgements of the probabilities):

(A)	Linda is active in the feminist movement	(0.61)
(B)	Linda is a bank teller.	(0.38)
(A&B)	Linda is a bank teller and is active in the feminist movement	(0.51)

Obviously, the results contradict the Kolmogorov axioms of probability theory: The conjunction of two propositions can never get a higher probability than each of the two conjuncts. Closely related results were obtained when asking for ranking the probabilities of the different events or when asking for frequency judgments.

I will demonstrate now how asymmetric conjunction resolves this conjunction puzzle in the probabilistic case. We can define the conjunction effect as follows:

$$\Delta(A, B) = \mu(A; B) - \mu(B) \quad (11)$$

According to classical probability theory the value of the conjunction effect is always negative or zero. However, when using equation (4) and inserting (5) for the interference term we get the following expression for the conjunction effect: $-\mu(\mathbf{a}';\mathbf{b})-2\cdot\mu^{1/2}(\mathbf{a};\mathbf{b})\cdot\mu^{1/2}(\mathbf{a}';\mathbf{b})\cdot\cos(\delta)$, and this value can be positive if $\cos(\delta)$ is negative. For example, assume $\delta=\pi$. Then the conjunction effect is always positive (or zero in some special cases): $\mu^{1/2}(\mathbf{a}';\mathbf{b})\cdot[\mu^{1/2}(\mathbf{a}';\mathbf{b})+2\cdot\mu^{1/2}(\mathbf{a};\mathbf{b})]$. In the example case with $\mu(B) = 0.38$ (Linda is a bank teller), $\mu(A) = 0.61$ (Linda is a feminist), $\mu(A; B) = 0.51$ (Linda is a feminist bank teller), we get a conjunction effect $\Delta(A, B) = 0.13$. This corresponds to a parameter $\cos(\delta) = -0.7$.

The present approach followed Khrennikov [e.g. 21-23, 25] who pioneered the investigation of interference effects in cognitive macro-systems. Franco [29] suggested to apply basically the same framework for explaining the Ellsberg paradox. Hence, the existence of concepts based on a single prototype state and their adequate treatment within an ortho-algebraic framework provides a rational foundation of some hard puzzles of bounded rationality.

6. TYPICALITY AND CONCEPTUAL COMBINATION

It should be noticed that there are related experiments where representativeness (or prototypicality) is judged [e.g. 30, 31, 32]. In these experiments the subjects are asked for the degree to which an instance is *representative* of a concept (or it is asked for the degree of correspondence between an instance and a category). As a first approximation I suggest to identify representativeness/typicality – written as $c_i(A)$ – with cue validity, i.e. to identify it with the conditional probability $\mu(i|A)$:

$$c_i(A) = \mu(i|A) \quad (12)$$

Taking *classical probabilities* this model already makes some interesting predictions. Let's start with the conjunction effect of typicality:

$$c_i(A) / c_i(A \cap B) \leq 1 \quad (13)$$

This prediction is verified in some cases. For instance, take $i = guppy$, which is a poorish example of a *fish*, and a poorish example of a *pet*, but it's quite a good example of a *pet fish*: $c_i(A \cap B) > c_i(A)$. However, there are also many counterexamples. For instance in cases where i is a typical instance of A but not of $A \cap B$, this prediction is clearly wrong.

Another prediction of the model is as follows. Let be $i \in A$ and $j \in A$, then the following equivalence should be valid:

$$c_i(A) \geq c_j(A) \text{ iff } \mu(i) \geq \mu(j) \quad (14)$$

Hence, in comparing the typicality of different instances of a concept only the (absolute) probability of the instances should matter, and the considered concept should not influence the comparison. It is simple to find counterexamples against this claim (14). For example, assume i as an instance of *focal red* and j as an instance of the colour of *red beans*. Plausibly, we can assume that focal colours are more salient (probably) than non-focal ones: $\mu(i) \geq \mu(j)$. In case A the colour of a *red flag* then (14) seems to be satisfied but (14) is clearly wrong if A is the colour of *red beans*.

I will show now that we get a more adequate theory of prototypicality when we replace the classical probabilistic conception of cue-validity by its ortho-algebraic pendant. It should be mentioned that the present proposal improves earlier proposals made by Aerts & Gabora [33, 34] and Blutner [35].

As in the previous section we consider instances of concepts as vectors of a Hilbert space \mathcal{H} , and we take the prototype view of concepts and calculate a as a superposition of instances, $|a\rangle = \sum_{i \in A} w(i) |i\rangle$. Using the prototype concept \underline{a} as given in (8) we get:

$$c_i(A) = \mu(\mathbf{i}|\underline{a}) = |\langle i|a\rangle|^2 = |w(i)|^2 \quad (15)$$

The crucial point corresponds to the construction of the quantum pendant to conjoined concepts $A \cap B$. The important assumption is to construct it as the tensor product $|a\rangle \otimes |b\rangle$ corresponding to the two prototypes (for A and

B). This idea corresponds to the idea of using tensor products for constructing complex activation vectors in connectionist models [cf. 36, 37, 38].

Next, we need a diagonalization operation in order to build an entangled state $\Delta(|a\rangle \otimes |b\rangle)$ representing the conjoined concept. The unitary operation of diagonalization is defined as follows:

$$\Delta\left(\sum_{u,v} x_{ij} |i\rangle \otimes |j\rangle\right) = \sum_{u,v} \delta(i, j) x_{ij} |i\rangle \otimes |j\rangle \quad (16)$$

The typicality of an instance i for the conjoined concept is the quantum probability that the entangled state $\Delta(|a\rangle \otimes |b\rangle)$ collapses into the product state $|i\rangle \otimes |i\rangle$.

Figure 2 illustrates that we can explain the conjunction effect of prototypicality in that way, but avoiding the problems in connection with the *classical* cue validity.

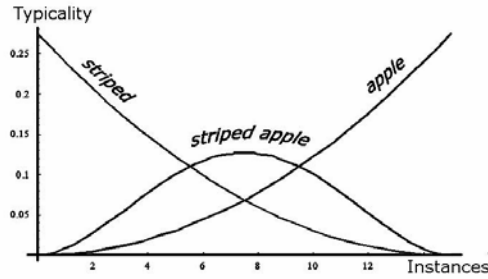


FIGURE 2. Conjunction effect of prototypicality calculated in a one-dimensional space of instances

Figure 3 illustrates the conjunction effect of prototypicality in a more realistic two-dimensional case.

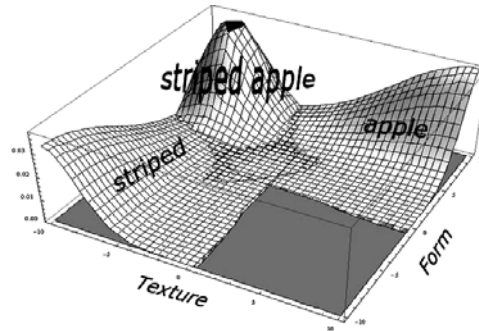


FIGURE 3. Conjunction Effect of Typicality calculated in a two-dimensional space of instances (using the dimensions *form* and *texture*)

Applications of the proposed model concern the calculation of the most typical colour value for combinations such as *red beans*, *white beans*, *black beans*, *red nose*, *red flag* etc. There are only a few (rather informal) proposals in the literature how to handle this problem of composing prototypes, one is due to suggestions made by Peter Gärdenfors within the framework of his geometrical model of meaning [cf. 39, 40]. It is a future task to compare the main characteristics of the present tensor product model with these alternatives.

7. CONCLUSIONS

Recently, a number of researchers have argued that the basic framework of quantum theory can find useful applications in the cognitive domain [e.g. 18, 19-21, 41, 42] and in the domain of personality diagnostics [43]. Hence, an active dialogue between cognitive approaches to concept formation, logic and semantics and the modern approach of quantum information science is mandatory.

The main argument for applying the formal apparatus of quantum theory to the domain of cognition has to do with the flexibility, instability, and context-dependency of natural concepts that manifest themselves as fleeting contents of conscious experience. For example, in the domain of language, words are floating freely in a polyvalent state representing a variety of different uses. As the properties of small particles are not absolute and determined until observing them, the properties of word tokens are not determined until conscious apprehension. Similarly, impressions, ideas and opinions are conceptual entities with analogous properties and likewise invite an analysis in terms of quantum theory.

In the present paper I make the distinction between concepts based on a mixture of instances (as described in the traditional Bayesian view) and concepts based on a single prototypes (described by *pure* quantum states), and I added the idea that there is a transfer between these two types of concepts. The relevant transfer from a mixed state into a pure state relates to the idea of “freezing” which is a very popular concept in natural language research (related terms are *grammaticalization*, *fossilization*, *conventionalization*) [cf. 44, 45, 46]. Using Niesteegge’s [12] conception of conditional quantum probabilities, this transfer relates to the evolution of states where the *algebraic relation* between the relevant events determine the conditional probabilities (and not any initial probability measure).

I have argued that Niesteegge’s [11] theory of conditional probabilities is an especially useful framework for integrating the classical Bayesian framework with the idea of prototype construction. This possibly adds a useful theoretical element to the toolbox of bounded rationality (as it was illustrated by using the operation of asymmetric conjunction to overcome some puzzles of bounded rationality in the case of probability judgements). Further, Niesteegge’s theory was useful for modelling typicality in terms of non-classical cue validity. The problem of conceptual combination was solved by using tensor products and a special diagonalization operation (leading to entangled quantum states).

The present investigation is limited in several ways. For example, we considered prototype concepts as consisting of *one* prototype state only. On the other hand, we considered only a special case of mixed states – one that corresponds to classical (Boolean) concepts. There are several reasons why concepts should be formed by mixing several prototype states. One reason relates to the peculiarities of propositional reasoning where it was clearly demonstrated that neither the Bayesian approach nor an approach based on the strict prototype view – with its symmetry $\mu(\mathbf{b}|\mathbf{a}) = \mu(\mathbf{a}|\mathbf{b})$ – is sufficient to fit the existing data [8]. Another reason relates to the general idea of fossilization, which relates to a transfer from mixed conceptual states to pure (prototype) states. Obviously, fossilization relates to a graded, stepwise process. Such a process can best be characterized as a *decrease* of quantum entropy [47, 48]. Of course, this process is not a single big jump from the excess statistical entropy of mixtures to the minimum uncertainty entropy of pure states.

REFERENCES

1. Oaksford M., Chater N., Against logicist cognitive science, *Mind and Language* 6 (1991) 1-38.
2. Kersten D., High-level vision as statistical inference, in: Gazzaniga M.S. (Ed.), *The new cognitive neuroscience*, MIT Press, Cambridge, Mass., 1999, pp. 353-363.
3. Kersten D., Yuille A.L., Vision as Bayesian Inference: Analysis by Synthesis?, *Trends in Cognitive Sciences* 10 (2006) 301-308.
4. Manning C., Schütze H., *Foundations of statistical natural language processing*, MIT Press, Cambridge, Mass., 2000.
5. Tenenbaum J.B., Bayesian modeling of human concept learning, in: Kearns M.S., Solla S.A. and Cohn D.A. (Eds.), *Advances in neuronal information processing systems*, Vol. 11, MIT Press Cambridge, Mass., 1999, pp. 59-65.
6. Anderson J.R., The adaptive nature of human categorization, *Psychological Review* 98 (1991) 409-429.
7. Temperley D., *Music and probability*, MIT Press, Cambridge, Mass., 2007.
8. Oaksford M., Chater N., *Bayesian Rationality: The Probabilistic Approach to Human Reasoning*, Oxford University Press, Oxford, 2007.
9. Gigerenzer G., Selten R., *Bounded Rationality: The Adaptive Toolbox*, MIT Press, Cambridge, MA 2001.
10. Tversky A., *Preference, Belief, and Similarity: Selected Writings* edited by Eldar Shafir, MIT Press, Cambridge, MA, 2004.
11. Niesteegge G., An Approach to Quantum Mechanics via Conditional Probabilities, *Foundations of Physics* 38 (2008) 241-256.
12. Halpern J.Y., *Reasoning about Uncertainty*, MIT Press, Cambridge, MA, 2003.
13. Widdows D., *Geometry and Meaning*, CSLI, Stanford University, 2004.
14. Widdows D., Peters S., *Word Vectors and Quantum Logic: Experiments with negation and disjunction*, Eighth Mathematics of Language Conference, Indiana University, Bloomington, Indiana, 2003, pp. 141-154.
15. Lewis D., *Philosophical Papers Volume I*, Oxford University Press, Oxford, 1983.
16. Montague R., Universal grammar, *Theoria* 36 (1970) 373-398.
17. Simon H.A., A behavioral model of rational choice, *The Quarterly Journal of Economics* 69 (1955) 99-118.

18. Aerts D., Czachor M., D'Hooghe B., Do We Think and Communicate in Quantum Ways? On the Presence of Quantum Structures in Language, in: Gontier N., Bendegem J.P.V. and Aerts D. (Eds.), *Evolutionary Epistemology, Language and Culture*, John Benjamins Publishing Company, Amsterdam, 2005.
19. Busemeyer J.R., Wang Z., Townsend J.T., Quantum dynamics of human decision-making, *Journal of Mathematical Psychology* 50 (2006) 220-241.
20. Franco R., Quantum mechanics and rational ignorance, Arxiv preprint physics/0702163 (2007).
21. Khrennikov A., Quantum-like formalism for cognitive measurements, *BioSystems* 70 (2003) 211-233.
22. Khrennikov A., Reconstruction of quantum theory on the basis of the formula of total probability, Arxiv preprint quant-ph/0302194, 2003.
23. Khrennikov A., Quantum-like brain: "Interference of minds", *BioSystems* 84 (2006) 225-241.
24. Franco R., Risk, ambiguity and quantum decision theory, Arxiv preprint arXiv:0711.0886, 2007.
25. Conte E., Khrennikov A.Y., Todarello O., Robertis R.D., Federici A., Zbilut J.P., A Preliminary Experimental Verification on the Possibility of Bell Inequality Violation in Mental States, *NeuroQuantology* 6 (2008) 214-221.
26. Savage L.J., *The Foundations of Statistics*, Wiley, New York, NY, 1954.
27. Tversky A., Shafir E., The disjunction effect in choice under uncertainty, *Psychological Science* 3 (1992) 305-309.
28. Tversky A., Kahneman D., Extension versus intuitive reasoning: The conjunction fallacy in probability judgment, *Psychological Review* 90 (1983) 293-315.
29. Franco R., Risk, ambiguity and quantum decision theory, Arxiv preprint arXiv:0711.0886 (2007).
30. Osherson D.N., Smith E.E., On the adequacy of prototype theory as a theory of concepts, *Cognition* 9 (1981) 35-58.
31. Smith R., Osherson D., Rips L., Keane M., Combining prototypes: A selective modification model, *Cognitive Science* 12 (1988).
32. Smith E.E., Osherson D.N., Conceptual combination with prototype concepts, *Cognitive Science* 8 (1984) 337-361.
33. Aerts D., Gabora L., A state-context-property model of concepts and their combinations I: The structure of the sets of contexts and properties, *Kybernetes* 34 (2005) 151-175.
34. Aerts D., Gabora L., A state-context-property model of concepts and their combinations II: A Hilbert space representation, *Kybernetes* 34 (2005) 176-205.
35. Blutner R., Combining Prototypes: Quantal Macrostates & Entanglement, Quantum Mind Symposium, Salzburg. Available from: http://amor.rz.hu-berlin.de/~h0998dgh/Contributions%202007/salzburg-talk_handout.pdf, 2007.
36. Smolensky P., Tensor product variable binding and the representation of symbolic structures in connectionist systems. *Artificial Intelligence* 46 (1990) 159-216.
37. Smolensky P., Constituent structure and explanation in an integrated connectionist/symbolic cognitive architecture, in: Macdonald C. and Macdonald G. (Eds.), *Connectionism: Debates on Psychological Explanation*, Blackwell, Oxford, 1995, pp. 221-290.
38. Smolensky P., Legendre G., *The Harmonic Mind: From neural computation to optimality-theoretic grammar*, MIT Press, Cambridge, Mass., 2006.
39. Gärdenfors P., Some tenets of cognitive semantics, in: Allwood J. and Gärdenfors P. (Eds.), *Cognitive semantics. Meaning and cognition*, John Benjamins, Amsterdam, 1999, pp. 19-36.
40. Gärdenfors P., *Conceptual spaces: The geometry of thought*, The MIT Press, Cambridge, Mass., 2000.
41. Atmanspacher H., Römer H., Walach H., *Weak Quantum Theory: Complementarity and Entanglement in Physics and Beyond*, *Foundations of Physics* 32 (2002) 379-406.
42. Aerts D., Broekaert J., Gabora L., A case for applying an abstracted quantum formalism to cognition, *New Ideas in Psychology*. Archive reference and link: <http://uk.arxiv.org/abs/quant-ph/0404068>, 2008.
43. Blutner R., Hochnadel E., Two Qubits for C.G. Jung's Theory of Personality, Universiteit van Amsterdam, ILLC. Available from www.quantum-cognition.de, Amsterdam, 2008.
44. Blutner R., Zeevat H., *Optimality-Theoretic Pragmatics*, in: Maienborn C., Heusinger K.v. and Portner P. (Eds.), *Semantics: An International Handbook of Natural Language Meaning*, Mouton de Gruyter, Berlin, to appear.
45. Levinson S., *Presumptive meaning: The theory of generalized conversational implicature*, MIT Press, Cambridge, Mass., 2000.
46. Traugott E., Dasher R.B., *Regularity in Semantic Change*, Cambridge University Press, Cambridge, 2005.
47. von Neumann J., *Mathematical foundations of quantum mechanics*, Princeton University Press Princeton, NJ, 1955.
48. Stotland A., Pomeransky A.A., Bachmat E., Cohen D., The information entropy of quantummechanical states, *Europhysics Letters* 67 (2004) 700-706.