Journal of Behavioral and Brain Science, 2012, 2, 195-220
http://dx.doi.org/10.4236/jbbs.2012.22024 Published Online May 2012 (http://www.SciRP.org/journal/jbbs)

Mathematical Model of Embodied Symbols: Cognition and Perceptual Symbol System

Leonid Perlovsky[1,2], Roman Ilin[2]
[1]Harvard University, Cambridge, USA
[2]Air Force Research Laboratory, Wright-Patterson Air Force Base, Greene & Montgomery, Ohio, USA
Email: leonid@seas.harvard.edu, roman.ilin@hanscom.af.mil

Received June 5, 2011; revised October 3, 2011; accepted November 10, 2011

ABSTRACT

A mathematical model of perceptual symbol system is developed. This development requires new mathematical methods of dynamic logic (DL), which have overcome limitations of classical artificial intelligence and connectionist approaches. The paper discusses these past limitations, relates them to combinatorial complexity (exponential explosion) of algorithms in the past, and relates it further to the static nature of classical logic. DL is a process-logic; its salient property is evolution of vague representations into crisp. We first consider one aspect of PSS: situation learning from object perceptions. Next DL is related to PSS mechanisms of concepts, simulators, grounding, embodiment, productivity, binding, recursion, and to the mechanisms relating embodied-grounded and amodal symbols. We discuss DL capability for modeling cognition on multiple levels of abstraction. PSS is extended toward interaction between cognition and language. Experimental predictions of the theory are discussed. They might influence experimental psychology and impact future theoretical developments in cognitive science, including knowledge representation, and mechanisms of interaction between perception, cognition, and language. All mathematical equations are also discussed conceptually, so mathematical understanding is not required. Experimental evidence for DL and PSS in brain imaging is discussed as well as future research directions.

Keywords: Perceptual Symbol System; Embodied Cognition; Perception; Knowledge Representation; Grounded Symbols; Dynamic Logic

1. Perceptual Symbol System

Barsalou [1] developed Perceptual symbol system (PSS) that embodies cognition and grounds it in perception. "Grounded cognition… rejects the standard view that amodal symbols represent knowledge in semantic memory". PSS emphasized the roles of simulation in cognition. "Simulation is the reenactment of perceptual, motor, and introspective states acquired during experience with the world, body, and mind… when knowledge is needed to represent a category (e.g., chair), multimodal representations captured during experiences… are reactivated to simulate how the brain represented perception, action, and introspection associated with it". Simulation is an essential computational mechanism in the brain. These simulation mechanisms are well known in case of mental imagery [2,3]. According to PSS cognition supports action. Although simulation is a central mechanism of PSS, yet rarely, if ever, they recreate full experiences. Based on the mechanism of simulators, which approximately correspond to concepts and types in amodal theories, PSS implements the standard symbolic functions of type-token binding, inference, productivity, recursion, and propositions. Using these mechanisms PSS retains the symbolic functionality. "Thus, PSS is a synthetic approach that integrates traditional theories with grounded theories." [1,4,5].

During the Cognitive Revolution in the middle of the last century, according to Barsalou, cognitive scientists were inspired by new forms of representation "based on developments in logic, linguistics, statistics, and computer science." Amodal representations were adopted, such as feature lists, semantic networks, and frames [6]. Virtually no empirical evidence supports amodal symbolic mechanisms [1]. The amodal symbols were adopted largely because they promised to provide "elegant and powerful formalisms for representing knowledge, because they captured important intuitions about the symbolic character of cognition, and because they could be implemented in artificial intelligence". The next section discusses that these promises were unfulfilled due to fundamental mathematical difficulties.

Past and ongoing developments of computational implementations of PSS include [7-9]. Yet, computational models for PSS require new mathematical methods dif-

ferent from traditional artificial intelligence, pattern recognition, or connectionist methods. The reason is that the traditional methods encountered combinatorial complexity (CC), an irresolvable computational difficulty, when attempting to model complex systems. Cognitive modeling requires learning combinations of perceptual features and objects or events [10-14].

This article develops a realistic and scalable mathematical model of grounded symbols and formalization of PSS based on a new computational technique of dynamic logic, DL [15,16]. The developed mathematical formalism is quite general. We first concentrate on one example of PSS mechanism: a mathematical description of models and simulators for forming and enacting representations of situations (higher level symbols) from perceptions of objects (lower level symbols), and then we discuss its general applicability. In addition to simulators, we consider concepts, grounding, binding, dynamic aspect of PSS (DIPSS), abstract concepts, the mechanism of amodal symbols within PSS, and the role of logic.

Past mathematical difficulties are considered in Section 2. They are related to classical logic, and a new computational technique of dynamic logic (DL) is introduced, which overcomes past computational limitations. Whereas classical logic is a static logic of statements (e.g., "if A then B"), DL describes a process capable of modeling the main components of PSS, including simulators. Section 3 illustrates the important properties of DL. Section 4 illustrates how DL models essential mechanisms of PSS considering an example of learning situations from objects (a difficult problem due to its inherent combinatorial complexity). Section 5 discusses DL as a general mechanism of interacting bottom-up and top-down signals, applicable to all levels of cognitive processing. Section 6 continues this discussion concentrating specifically on DL modeling amodal vs perceptual symbols. Section 7 discusses experimental evidence confirming DL predictions of the mind mechanism, and formulates further predictions that could be tested experimentally in the near future. Section 8 describes future theoretical research as well as proposed verifiable experimental predictions of DL.

2. Past Mathematical Difficulties and Recent Development

In modern neuroscience, a fundamental process in object perception is an interaction of bottom-up signals from sensory organs and top-down signals from internal mind's representations (memories) of objects. During perception, the mind matches subsets of bottom-up signals corresponding to objects with representations of object in the mind (and top-down signals). This produces object recognition; it activates brain signals leading to mental and behavioral responses [3,17-20]. This section briefly summarizes mathematical development in artificial intelligence, pattern recognition, and other computational methods used in cognitive science for modeling brain-mind processes. The fundamental difficulties preventing mathematical modeling of perception, cognition, and PSS are discussed; then overcoming these difficulties using DL is discussed.

2.1. Computational Complexity

Mathematical modeling of the above recognition process has not been easy, a number of difficulties have been encountered during the past fifty years. These difficulties were summarized under the notion of combinatorial complexity (CC) [21]. CC refers to multiple combinations of bottom-up and top-down signals, or more generally to combinations of various elements in a complex system; for example, recognition of a scene often requires concurrent recognition of its multiple elements that could be encountered in various combinations. CC is computationally prohibitive because the number of combinations is very large: for example, consider 100 elements (not too large a number); the number of combinations of 100 elements is 100^{100}, exceeding the number of all elementary particle events in life of the Universe; no computer would ever be able to compute that many combinations.

CC was first identified in pattern recognition and classification research in the 1960s and was named "the curse of dimensionality" [22]. It seemed that adaptive self-learning algorithms and neural networks could learn solutions to any problem "on their own", if provided with a sufficient number of training examples. However the following development of artificial intelligence led to a conclusion that the required number of training examples often was combinatorially large. This remains true about recent generation of algorithms and neural networks, which are much more powerful than those in the 1950s and 60s. Training had to include not only every object in its multiple variations, angles, etc., but also combinations of objects. Thus, learning approaches encountered CC of learning requirements.

In the 1970's rule systems were proposed to solve the problem of learning complexity [23,24]. Minsky suggested that learning was a premature step in artificial intelligence; Newton "learned" Newtonian laws, most of scientists read them in the books. Therefore Minsky has suggested, knowledge ought to be input in computers "ready made" for all situations and artificial intelligence would apply these known rules. Rules would capture the required knowledge and eliminate a need for learning. Chomsky's original ideas concerning mechanisms of language grammar related to deep structure [25] were also based on logical rules. Rule systems work well when all aspects of the problem can be predetermined. However in the presence of variability, the number of rules grew;

rules became contingent on other rules and combinations of rules had to be considered. The rule systems encountered CC of rules.

To overcome these difficulties model systems were proposed in the 1980s to combine advantages of learning and rules-models by using adaptive models [10,11,26-32]. Available knowledge was to be summarized in models and parameters of the models were to describe unknown aspects of concrete situations. Similar was the Chomsky idea [33] of principles and parameters. However, fitting models to data (top-down to bottom-up signals) required selecting data subsets corresponding to various models. The number of subsets is combinatorially large. A general popular algorithm for fitting models to the data, multiple hypothesis testing [34] is known to face CC of computations. Model-based approaches encountered computational CC (N and NP complete algorithms).

2.2. CC, Amodal Symbols, and Logic

Perceptual symbols and amodal symbols, as described by PSS, differ not only in their representations in the brain, but also in their properties that are mathematically modeled in this paper. This mathematically fundamental difference and its relations to CC of matching bottom-up and top-down signals are discussed in this section.

It has been demonstrated that CC is related to the use of formal logic in algorithms and neural networks [11,12, 35]. Logic serves as a foundation for many approaches to cognition and linguistics; it underlies most of computational algorithms. But its influence extends far beyond, affecting cognitive scientists, psychologists, and linguists, including those not using complex mathematical algorithms for modeling the mind. Formal logic is more than 2000 years old, it influences the entire science. Most psychologists make a more or less conscious assumption that the mechanisms of logic serve as the basis of cognition. As discussed in Section 7, our minds are unconscious about its illogical foundations. Only approximately logical part of the mind mechanisms is accessible to consciousness. Although this is a minor part of the mind operations it fundamentally influences scientific intuition. It is unconsciously affected by the bias toward logic. Even when the laboratory data drive our thinking away from logical mechanisms it is difficult to overcome the logical bias [11,12,16,18,19,36-41].

Relationships between cognition and logic have been a source of longstanding myth. The widely accepted story is that Aristotle founded logic as a fundamental mind mechanism, and only during the recent decades science overcame this influence. I would like to emphasize the opposite side of this story. Aristotle assumed a close relationship between logic and language. He emphasized that logical statements should not be formulated too strictly and language inherently contains the necessary degree of precision. According to Aristotle, logic serves not for thinking but to communicate already made decisions [42,43]. The mechanism of the mind relating language, cognition, and the world Aristotle [44] described as forms. Today we call these mechanisms mental representations, or concepts, or simulators in the mind. Aristotelian forms are similar to Plato's ideas with a marked distinction, forms are dynamic: their initial states, before learning, are different from their final states of concepts. Aristotle emphasized that initial states of forms, forms-as-potentialities, are not logical (*i.e*. vague), but their final forms, forms-as-actualities, attained in the result of learning, are logical. This fundamental idea that forms are processes evolving from illogical to logical was lost during millennia of philosophical arguments. As discussed below this Aristotelian process of dynamic forms corresponds to Barsalou idea of PSS simulators, and the mathematical model, DL, for this process described in this paper.

A contradiction between logic and language was emphasized by the founders of formal logic. In the 19th century George Boole and the great logicians following him, including Gottlob Frege, Georg Cantor, David Hilbert, and Bertrand Russell (see [45] and references therein) eliminated the uncertainty of language from mathematics, and founded formal mathematical logic, the foundation of the current classical logic. Hilbert developed an approach named formalism, which rejected intuition as a matter of scientific investigation and formally defined scientific objects in terms of axioms or rules. In 1900 he formulated famous Entscheidungs problem: to define a set of logical rules sufficient to prove all past and future mathematical theorems. This was a part of "Hilbert's program", which entailed formalization of the entire human thinking and language. Formal logic ignored the dynamic nature of Aristotelian forms and rejected the uncertainty of language. Hilbert was convinced that his logical theory described the mind mechanisms. "The fundamental idea of my proof theory is none other than to describe the activity of our understanding, to make a protocol of the rules according to which our thinking actually proceeds." (see [46]). However, Hilbert's vision of formalism explaining mysteries of the human mind came to an end in the 1930s, when Gödel [47] proved internal inconsistency of formal logic. This development called Gödel theory is considered among most fundamental mathematical results of the previous century. It demonstrated that logic, that was believed to be a sure way to derive truths, turned out to be basically flawed. This is a reason why theories of cognition and language based on formal logic did not work.

Logic is fundamentally related to CC. CC turned out to be a finite-system manifestation of the Gödel's theory [48-50]. If Gödelian theory is applied to finite systems

(such as computers and brain-mind), CC is the result, instead of the fundamental inconsistency. Algorithms matching bottom-up and top-down signals based on formal logic have to evaluate every variation in signals and their combinations as separate logical statements. Combinations of these variations cause CC.

This property of logic manifests in various algorithms in different ways. Rule systems are logical in a straightforward way, and the number of rules grows combinatorially. Pattern recognition algorithms and neural networks are related to logic in learning procedures: every training sample is treated as a logical statement ("this is a chair") resulting in CC of learning. Multivalued logic and fuzzy logic were proposed to overcome limitations related to logic [51,52]. Yet the mathematics of multivalued logic is no different in principle from formal logic [15,16]. Fuzzy logic uses logic to set a degree of fuzziness. Correspondingly, it encounters a difficulty related to the degree of fuzziness: if too much fuzziness is specified, the solution does not achieve a needed accuracy, and if too little, it becomes similar to formal logic. If logic is used to find the appropriate fuzziness for every model at every processing step, then the result is CC. The mind has to make concrete decisions, for example one either enters a room or does not; this requires a computational procedure to move from a fuzzy state to a concrete one. But fuzzy logic does not have a formal procedure for this purpose; fuzzy systems treat this decision on an ad-hoc logical basis.

Is logic still possible after Gödel's proof of its incompleteness? The contemporary state of this field was reviewed in [35]. It appears that logic after Gödel is much more complicated and much less logical than was assumed by founders of artificial intelligence. CC cannot be solved within logic. Penrose argued that Gödel's results imply incomputability of the mind processes and testify for a need for new physics "correct quantum gravitation", which would resolve difficulties in logic and physics [53]. An opposite view in [12,15,16,54] is that incomputability of logic is not related to incomputability of the mind. These publications add mathematical arguments to the Aristotelian view that logic is not the basic mechanism of the mind.

To summarize, various manifestations of CC are all related to formal logic and Gödel theory. Rule systems rely on formal logic in a most direct way. Even mathematical approaches specifically designed to counter limitations of logic, such as fuzzy logic and the second wave of neural networks (developed after the 1980s) rely on logic at some algorithmic steps. Self-learning algorithms and neural networks rely on logic in their training or learning procedures: every training example is treated as a separate logical statement. Fuzzy logic systems rely on logic for setting degrees of fuzziness. CC of mathematical approaches to the mind is related to the fundamental inconsistency of logic. Therefore logical intuitions, leading early cognitive scientists to amodal brain mechanisms, could not realize their hopes for mathematical models of the brain-mind.

The outstanding mathematicians of the 19th and early 20th c. believed that logic is the foundation of the mind, why? Even more surprising is that after Gödel. Gödelian theory was long recognized among most fundamental mathematical results of the 20th c. How is it possible that outstanding minds, including founders of artificial intelligence, and many cognitive scientists and philosophers of mind insisted that logic and amodal symbols implementing logic in the mind are adequate and sufficient? The answer is in the "conscious bias". As we have mentioned above and discuss in details in Section 7, non-logical operations represent more than 99.9% of the mind functioning, but they are not accessible to consciousness [19]. However, our consciousness functions in a way that makes us unaware of this. In subjective consciousness we usually experience perception and cognition as logical. Our intuitions are "consciously biased". This is why amodal logical symbols, which describe a tiny fraction of the mind mechanisms, have seemed to many the foundation of the mind [11,12,14-16,18,19,36-41].

We have already mentioned another aspect of logic relevant to PSS, logic lacks dynamics; it is about static statements such as "this is a chair". Classical logic is good at modeling structured statements and relations, yet it misses the dynamics of the mind and faces CC, when attempts to match bottom-up and top-down signals. The essentially dynamic nature of the mind is not represented in mathematical foundations of logic. Dynamic logic is a logic-process. It overcomes CC by automatically choosing the appropriate degree of fuzziness-vagueness for every mind's concept at every moment. DL combines advantages of logical structure and connectionist dynamics. This dynamics mathematically represents the process of Aristotelian forms, different from classical logic, and serves as a foundation for PSS concepts and simulators.

2.3. DL, a Logic-Process

DL models perception and cognition as an interaction between bottom-up and top-down signals [12,15,16,35, 38,39,41,42,55-62]. This section concentrates on the basic relationship between the brain processes and the mathematics of DL. To concentrate on this relationship, we much simplify the discussion of the brain structures. We discuss visual recognition of objects as if the retina and the visual cortex each consists of a single processing layer of neurons where recognition occurs (which is not true, detailed relationship of the DL process to brain is considered in given references). Perception consists of the association-matching of bottom-up and top-down

signals. Sources of top-down signals are mental representations, memories of objects created by previous simulators [1]; these representations model the patterns of bottom-up signals. In this way they are concepts (of objects), symbols of a higher order than bottom-up signals; we call them concepts or models. In perception processes the models are modified by learning and new models are formed; since an object is never encountered exactly the same as previously, perception is always a learning process.

The DL processes along with concept-representations are mathematical models of the PSS simulators. The bottom-up signals, in this simplified discussion, are a field of neuronal synapse activations in visual cortex. Sources of top-down signals are mental representation-concepts or, equivalently, model-simulators (for short, models; please notice this dual use of the word model, we use "models" for mental representation-simulators, which match-model patterns in bottom-up signals; and we use "models" for mathematical modeling of these mental processes). Each mental model-simulator projects a set of priming, top-down signals, representing the bottom-up signals expected from a particular object. Mathematical models of mental models-simulators characterize these mental models by parameters. Parameters describe object position, angles, lightings, etc. (In case of learning situations considered later, parameters characterize objects and relations making up a situation). To summarize this highly simplified description of a visual system, the learning-perception process "matches" top-down and bottom-up activations by selecting "best" mental models-simulators and their parameters and fitting them to the corresponding sets of bottom-up signals. This DL process mathematically models multiple simulators running in parallel, each producing a set of priming signals for various expected objects.

The "best" fit criteria between bottom-up and top-down signals were given in [12,15,16,35,62]. They are similar to probabilistic or informatic measures. In the first case they represent likelihood that the given (observed) data or bottom-up signals corresponds to representations-models (top-down signals) of particular objects. In the second case they represent information contained in representations-models about the observed data (in other words, information in top-down signals about bottom-up signals). These similarities are maximized over the model parameters. Results can be interpreted correspondingly as a maximum likelihood that models-representations fit sensory signals, or as maximum information in models-representations about the bottom-up signals. Both similarity measures account for all expected models and for all combinations of signals and models. Correspondingly, a similarity contains a large number of items, a total of M^N, where M is a number of models and N is a number of signals; this huge number is the cause for the previously discussed combinatorial complexity.

Maximization of a similarity measure mathematically models a cognitive mechanism of an unconditional drive to improve the correspondence between bottom-up and top-down signals (representations-models). In biology and psychology it was discussed as curiosity, cognitive dissonance, or a need for knowledge since the 1950s [63,64]. This process involves knowledge-related emotions evaluating satisfaction of this drive for knowledge [12,15,16,36,38,41,56,60-62,65-67]. In computational intelligence it is even more ubiquitous, every mathematical learning procedure, algorithm, or neural network maximizes some similarity measure. In the process of learning, mental concept-models are constantly modified.

The DL learning process, let us repeat, consists in estimating parameters of concept-models (mental representations) and associating subsets of bottom-up signals with top-down signals originating from these models-concepts by maximizing a similarity. Although a similarity contains combinatorially many items, DL maximizes it without combinatorial complexity [11,12,15,16,35,38, 39,48] as follows. First, vague-fuzzy association variables are defined, which give a measure of correspondence between each signal and each model. They are defined similarly to the a posteriori Bayes probabilities, they range between 0 and 1, and as a result of learning they converge to the probabilities, under certain conditions. This mathematical breakthrough led to solving many problems that could not have been solved previously [68-80].

After giving a short mathematical description of the DL process, we summarize it conceptually, so that understanding of mathematics is not essential. DL is defined by a set of differential equations given in the above references; together with models discussed later it gives a mathematical description of the PSS simulators. Here we summarize these equations simplifying the visual system, as if there is just one neural layer between the visual cortex and object recognition. Bottom-up signals $\{\mathbf{X}(n)\}$ is a field of neuronal synapse activations in visual cortex. Here and below curve brackets $\{\cdots\}$ denote multiple signals, a field. Index $n = 1,\cdots N$, enumerates neurons and $\mathbf{X}(n)$ are the activation levels. Sources of top-down signals are representations or models $\{\mathbf{M}_m(n)\}$ indexed by $m = 1,\cdots M$. Each model m $\mathbf{M}_m(n)$ projects a set of priming, top-down signals, representing the bottom-up signals $\mathbf{X}(n)$ expected from a particular object, m. Models depend on parameters $\{\mathbf{S}_m\}$, $\mathbf{M}_m(\mathbf{S}_m,n)$. Parameters characterize object position, angles, lightings, etc. (In case of learning *situations* considered in Section 3, parameters characterize objects and relations making up a situation). We use n to enumerate the visual cortex neurons, $\mathbf{X}(n)$ are the "bottom-up" activation levels of these neurons

coming from the retina, and $\mathbf{M}_m(n)$ are the "top-down" activation levels (priming) of the visual cortex neurons. The learning-perception process "matches" these top-down and bottom-up activations by selecting "best" models and their parameters and the corresponding sets of bottom-up signals. This "best" fit is given by maximizing a similarity measure between bottom-up and top-down signals; it is designed so that it treats each object-model as a potential alternative for each subset of signals [12,13,15,16,35]

$$L(\{X\},\{M\}) = \prod_{n\in N} \sum_{m\in M} r(m) l(\mathbf{X}(n)|\mathbf{M}_m(n)); \quad (1)$$

Here, $l(\mathbf{X}(n)|\mathbf{M}_m(n))$ (or simply $l(n|m)$) is called a conditional similarity between one signal $\mathbf{X}(n)$ and one model $\mathbf{M}_m(n)$. Parameters $r(m)$ are proportional to the number of objects described by the model m. Expression (1) accounts for *all combinations* of signals and models in the following way. Sum $\sum$ ensures that any of the object-models can be considered (by the mind) as a source of signal $\mathbf{X}(n)$. Product $\prod$ ensures that all signals have a chance to be considered (even if one signal is not considered, the entire product is zero, and similarity L is 0; so for good similarity all signals have to be accounted for. This does not assume exorbitant amount of attention to each minute detail: among models there is a vague simple model for "everything else"). In a simple case, when all objects are perfectly recognized and separated from each other, there is just one object-model corresponding to each signal (other $l(n|m) = 0$). In this simple case expression (1) contains just 1 item, a product of all non-zero $l(n|m)$. In the general case, before objects are recognized, L contains a large number of *combinations* of models and signals; a product over N signals is taken of the sums over M models; this results in a total of M^N items; this huge number is the cause for the combinatorial complexity discussed previously.

The DL learning process consists in estimating model parameters $\mathbf{S}_m$ and associating subsets of signals with concepts by maximizing the similarity (1). Although (1) contains combinatorially many items, DL maximizes it without combinatorial complexity [11,12,48,49]. First, fuzzy association variables $f(m|n)$ are defined,

$$f(m|n) = r(m) l(n|m) \Big/ \sum_{m'\in M} r(m') l(n|m'). \quad (2)$$

These variables give a measure of correspondence between signal $\mathbf{X}(n)$ and model $\mathbf{M}_m$ relative to all other models, m'. They are defined similarly to the a posteriori Bayes probabilities, they range between 0 and 1, and as a result of learning they converge to the probabilities under certain conditions.

DL process is defined by the following set of differential equations, as the Equation (3) below.

The principles of DL can be adequately understood from the following conceptual description and examples. As a mathematical model of perception-cognitive processes, DL is a process described by differential equations given above; in particular, fuzzy association variables f associate bottom-up signals and top-down models-representations. Among unique DL properties is an autonomous dependence of association variables on models-representations: in the processes of perception and cognition, as models improve and become similar to patterns in the bottom-up signals, the association variables become more selective, more similar to delta-functions. Whereas initial association variables are vague and associate near all bottom-up signals with virtually any top-down model-representations, in the processes of perception and cognition association variables are becoming specific, "crisp", and associate only appropriate signals. This we call a process "from vague to crisp". (The exact mathematical definition of crisp corresponds to values of f = 0 or 1; values of f in between 0 and 1 correspond to various degrees of vagueness).

DL processes mathematically model PSS simulators and not static amodal signals. Another unique aspect of DL is that it explains how logic appears in the human mind; how illogical dynamic PSS simulators give rise of classical logic, and what is the role of amodal symbols. This is discussed throughout the paper, and also in specific details in Section 6.

An essential aspect of DL, mentioned above, is that associations between models and data (top-down and bottom-up signals) are uncertain and dynamic; their uncertainty matches uncertainty of parameters of the models and both change in time during perception and cognition processes. As the model parameters improve, the associations become crisp. In this way the DL model of simulator-processes avoids combinatorial complexity because there is no need to consider separately various combinations of bottom-up and top-down signals. Instead, all combinations are accounted for in the DL simulator-processes. Let us repeat, that initially, the models do not match the data. The association variables are not the narrow logical variables 0, or 1, or nearly logical, instead they are wide functions (across top-down and bottom-up signals). In other words, they are vague, initially they take near homogeneous values across the data (across bottom-up and top-down signals); they associate all the representation-models (through simulator processes) with all the input signals [12,16,39].

$$df(m|n)/dt = f(m|n) \sum_{m'\in M} \{[\delta_{mm'} - f(m'|n)][\partial \ln l(n|m')/\partial M_{m'}](\partial M_{m'}/\partial S_{m'})\, dS_{m'}/dt$$

$$dS_m/dt = \sum_{n\in N} f(m|n)[\partial \ln l(n|m)/\partial M_m]\partial M_m/\partial S_m, \delta_{mm'} = 1 \text{ if } m = m', 0 \text{ otherwise} \quad (3)$$

Here we conceptually describe the DL process as applicable to visual perception, taking approximately 160 ms, according to the reference below. Gradually, the DL simulator-processes improve matching, models better fit data, the errors become smaller, the bell-shapes concentrate around relevant patterns in the data (objects), and the association variables tend to 1 for correctly matched signal patterns and models, and 0 for others. These 0 or 1 associations are logical decisions. In this way classical logic appears from vague states and illogical processes. Thus certain representations get associated with certain subsets of signals (objects are recognized and concepts formed logically or approximately logically) [81]. This process "from vague-to-crisp" that matches bottom-up and top-down signals has been independently conceived and demonstrated in brain imaging research to take place in human visual system [19,82]. Various neural mechanisms of dynamic logic in the brain is discussed in [62]. Thus DL models PSS simulators, describes how logic appears from illogical processes, and actually models perception mechanisms of the brain-mind.

2.4. Object Perception in Noise

Here we illustrate the DL processes, modeling multiple simulators running in parallel as described above. In this example, DL searches for patterns in noise. Finding patterns below noise can be an exceedingly complex problem. If an exact pattern shape is not known and depends on unknown parameters, these parameters should be found by fitting the pattern model to the data. However, when the locations and orientations of patterns are not known, it is not clear which subset of the data points should be selected for fitting. A standard approach for solving this kind of problem, which has already been mentioned, is multiple hypotheses testing [34]; this algorithm exhaustively searches all logical combinations of subsets and models and faces combinatorial complexity. In the current example, we are looking for "smile" and "frown" patterns in noise shown in **Figure 1(a)** without noise, and in **Figure 1(b)** with noise, as actually measured (object signals are about 2 - 3 times below noise and cannot be seen).

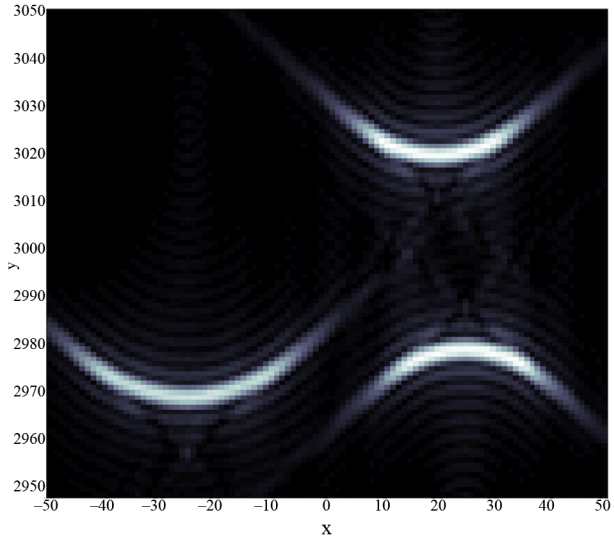
Several types of models are used in this example: parabolic models describing "smiles" and "frown" patterns (unknown size, position, curvature, signal strength, and number of models), circular-blob models describing approximate patterns (unknown size, position, signal strength, and number of models), and noise model (unknown strength). Exact mathematical description of these models is given in the reference cited above.

We consider the image size of 100 × 100 points (N = 10,000 bottom-up signals, corresponding to the number of receptors in an eye retina), and the true number of models is 4 (3 + noise), which is not known. Therefore, at least M = 5 models should be fit to the data, to decide that 4 fits best. The complexity of logical combinatorial search is $M^N = 10^{5000}$; this combinatorially large number is much larger than the size of the Universe and the problem was considered unsolvable. **Figure 1** illustrates DL operations: **(a)** true "smile" and "frown" patterns without noise; **(b)** actual image available for recognition; **(c)** through **(h)** illustrates the DL process, they show improved models at various steps of solving DL Equations (A3), total of 22 steps (noise model is not shown; **Figures (c)** through **(h)** show association variables, **(f)**, for blob and parabolic models). By comparing **(h)** to **(a)** one can see that the final states of the models match patterns in the signal. Of course, DL does not guarantee finding any pattern in noise of any strength. For example, if the amount and strength of noise would increase ten-fold, most likely the patterns would not be found (this would provide an example of "falsifiability" of DL; however more accurate mathematical description of potential failures of DL algorithms is considered later). DL reduced the required number of computations from combinatorial 10^{5000} to about 10^9. By solving the CC problem DL was able to find patterns under the strong noise. In terms of signal-to-noise ratio this example gives 10,000% improvement over the previous state-of-the-art (In this example DL actually works better than human visual system; the reason is that human brain is not optimized for recognizing these types of patterns in noise).

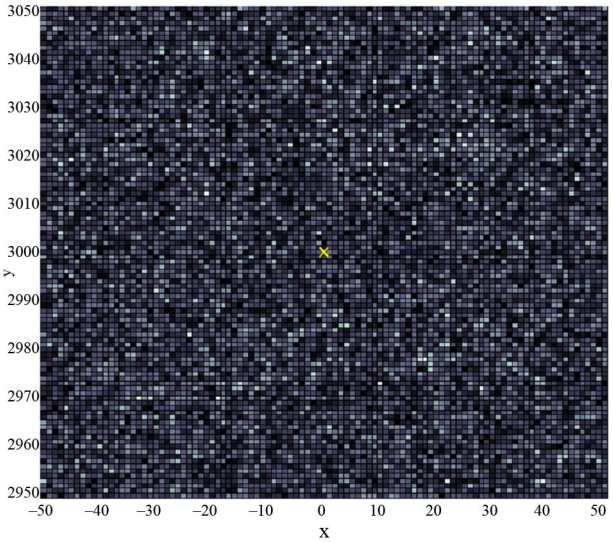
The main point of this example is that DL simulator-perception is a process "from vague-to-crisp", similar to visual system processes demonstrated in [19]. We would also like to take this moment to continue the arguments from Sections 2.1, 2.2, and to emphasize that DL is a fundamental and revolutionary improvement in mathematics [41,43,83]; it was recognized as such in mathematical and engineering communities; it is the only mathematical theory that have suggested vague initial states; it has been developed for over 20 years; yet it might not be well known in cognitive science community. A large number of mathematical and engineering applications of DL is given in references and references therein.

3. Relations between DL and PSS

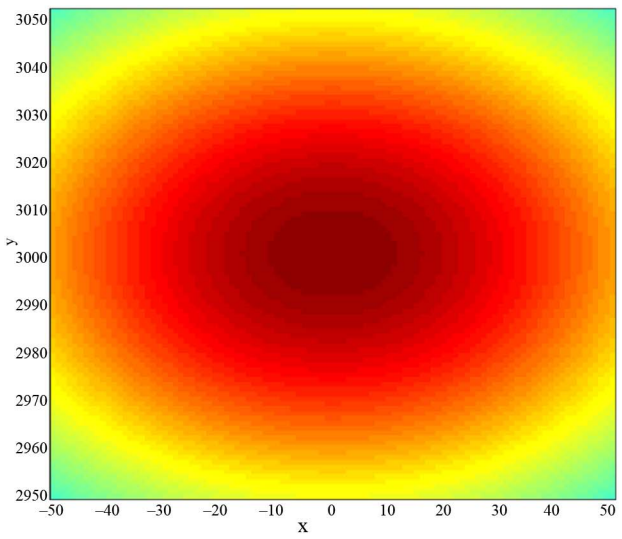
The previous section illustrated DL for recognition of simple objects in noise, a case complex and unsolvable for prior state-of-the-art algorithms, still too simple to be directly relevant for PSS. Here we consider a problem of situation learning, assuming that object recognition has been solved. In computational image recognition this is called "situational awareness" and it is a long-standing unsolved problem. The principled difficulty is that every situation includes many objects that are not essential for recognition of this specific situation; in fact there are many more "irrelevant" or "clutter" objects than relevant



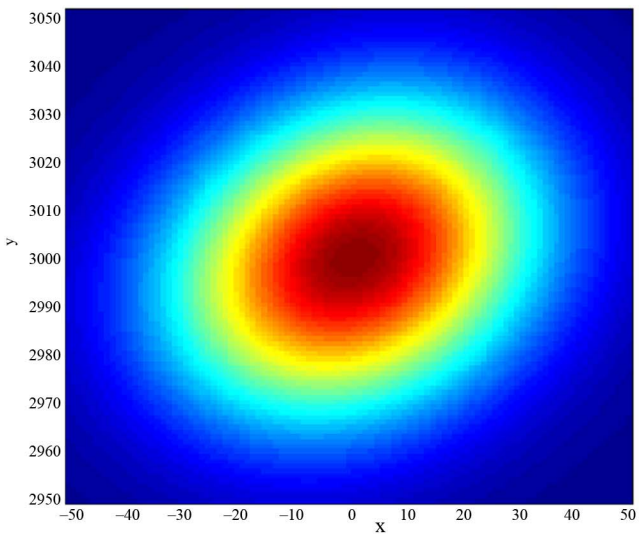
(a)



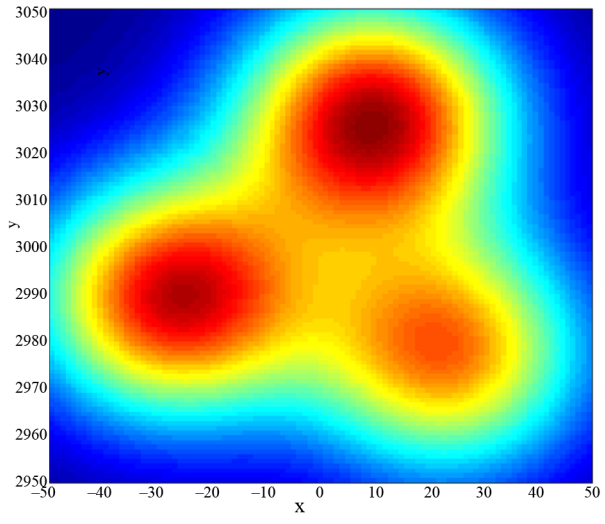
(b)



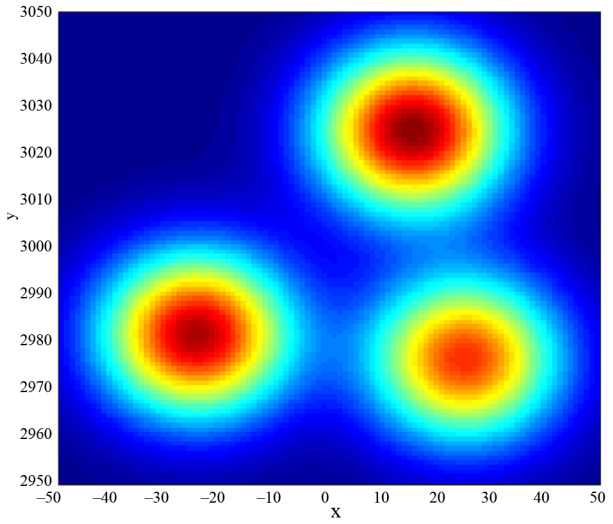
(c)



(d)



(e)



(f)

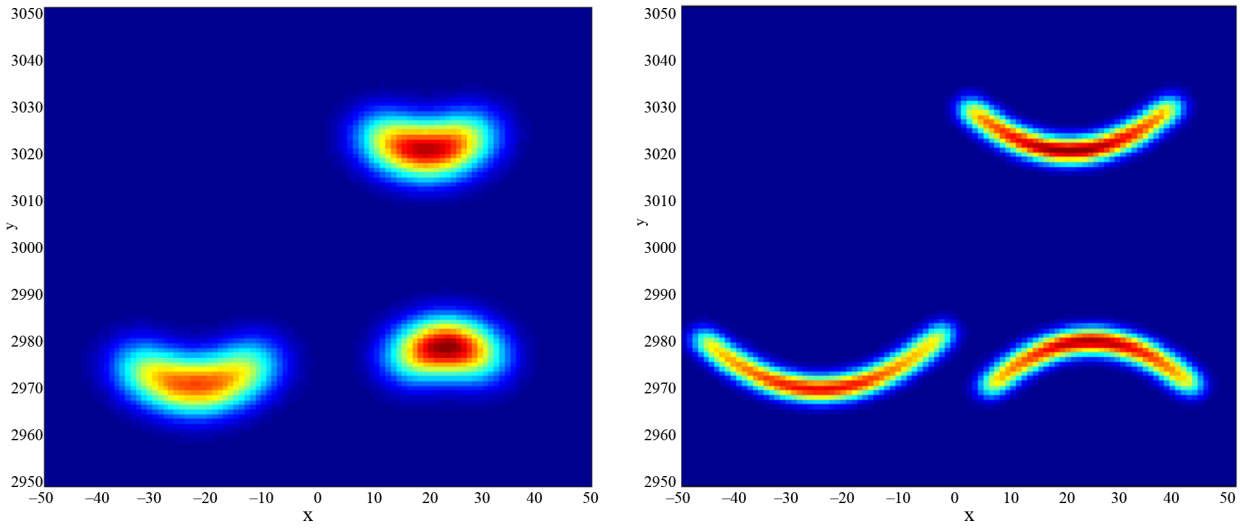


Figure 1. Finding "smile" and "frown" patterns in noise, an example of dynamic logic operation: (a) True "smile" and "frown" patterns are shown without noise; (b) Actual image available for recognition (signals are below noise, signal-to-noise ratio is between 1/2 and 1/4, 100 times lower than usually considered necessary); (c) An initial fuzzy blob-model, the vagueness corresponds to uncertainty of knowledge; (d) Through (h) show improved models at various steps of DL (Equation (3) are solved in 22 steps). Between stages (d) and (e) the algorithm tried to fit the data with more than one model and decided, that it needs three blob-models to "understand" the content of the data. There are several types of models: one uniform model describing noise (it is not shown) and a variable number of blob-models and parabolic models, which number, location, and curvature are estimated from the data. Until about stage (g) the algorithm "thought" in terms of simple blob models, at (g) and beyond, the algorithm decided that it needs more complex parabolic models to describe the data. Iterations stopped at (h), when similarity (1) stopped increasing.

ones. Let us dwell on this for a bit. Objects are spatially-limited material things perceptible by senses. A situation is a collection of contextually related objects that tend to appear together and are perceived as meaningful, e.g., an office, a dining room. The requirement for contextual relations and meanings makes the problem mathematically difficult. Learning contexts comes along with learning situations; it reminds of the problem of a chicken and egg. We subliminally perceive many objects, most of which are irrelevant, e.g. a tiny scratch on a wall, which we learn to ignore. Combinations of even a limited number of objects exceed what is possible to learn in a single lifetime as meaningful situations and contexts (e.g. books on a shelf) from random sets of irrelevant objects (e.g. a scratch on a wall, a book, and a pattern of tree branches in a window). Presence of hundreds (or even dozens) irrelevant objects makes learning by a child of mundane situations a mathematical mystery. In addition, a human constantly perceives large numbers of different objects and their combinations, which do not correspond to anything worth learning; the human mind successfully learn to ignore them.

A most difficult part of learning-cognition for mathematical modeling is to learn which sets of objects are important for which situations (contexts). The key mathematical property of DL that made this solution possible, same as in the previous section, is a process "from vague-to-crisp." Concrete crisp models-representations of situations are formed from vague models in the process of learning (or cognition-perception). We illustrate below how complex symbols, situations, are formed by situation-simulators from simpler perceptions, objects, which are simpler perceptual symbols, being formed by simulators at "lower" levels of the mind, comparative to "higher" situation-simulators. Situation-simulators operate on PSS representations of situations, which are dynamic and vague assemblages of situations from imagery (and other modalities), bits and pieces along with some relations among them perceived at lower levels. These pieces and relations may come from different past perceptions, not necessarily from a single perceptual mode, and not necessarily stored in a contiguous parts of the brain. The dynamic process of DL-PSS-simulation, which assembles these bits into situations attempting to match those before the eyes, is mostly unconscious. We will discuss in details in section 6 that these are perceptual symbols as described in [1]. DL mathematically models PSS simulators [1], processes that match bottom-up perceptions with top-down signals, assemble symbols in cognition-perception, and assemble conceptual representations by recreating patterns of activation in sensorimotor brain areas (as discussed later in the paper). An essential mechanism of DL cognition is a process of simulation of perceptual imagination-cognitions; these situation-symbols

are simulated from simpler perceptions-objects. The same mechanism can simulate plans and more complex abstract thoughts, as discussed in later sections.

3.1. DL for Learning Situations

Here we consider a problem, where an intelligent agent (a child) is learning to recognize certain situations in the environment; while it is assumed that a child has learned to recognize objects. In real life a child learns to recognize situations, to some extent, in parallel with recognizing objects. But for simplicity of the illustration examples and discussions below, we consider a case of objects being already known. For example, situation "office" is characterized by the presence of a chair, a desk, a computer, a book, a book shelf. Situation "playground" is characterized by the presence of a slide, a sandbox, etc. The principal difficulty is that many irrelevant objects are present in every situation.

We use D_o to denote the total number of objects that the child can recognize in the world (it is a large number). In every situation he or she perceives D_p objects. This is a much smaller number compared to D_o. Each situation is also characterized by the presence of D_s objects essential for this situation ($D_s < D_p$). Normally nonessential objects are present and D_s is therefore less than D_p. The sets of essential objects for different situations may overlap, with some objects being essential to more than one situation. We assume that each object is encountered in the scene only once. This is a minor and nonessential simplification, e.g. we may consider a set of similar objects as a new object.

A situation can be mathematically represented as a vector in the space of all objects, $\mathbf{X}_n = (x_{n1}, \cdots x_{ni}, \cdots x_{nDo})$ [84,85]. If the value of x_{ni} is one the object i is present in the situation n and if x_{ni} is zero, the corresponding object is not present. Since D_o is a large number, $\mathbf{X}_n$ is a large binary vector with most of its elements equal to zero. A situation model is characterized by parameters, a vector of probabilities, $\mathbf{p}_m = (p_{m1},\cdots p_{mi},\cdots p_{mDo})$. Here p_{mi} is the probability of object i being part of the situation m. Thus a situation model contains D_o unknown parameters. Estimating these parameters constitutes learning. We would like to emphasize that although notations like x_{ni}, p_{mi} might look like amodal symbols, such an interpretation would be erroneous. Correct interpretation of notations in a mathematical model depends on what actual physical entities are referred to by the notations. These notations refer to neural signals, elements from which simulator-processes assemble symbols of a higher level. As discussed, for simplicity of presentation of the results, we assume that lower-level simulator-processes responsible for object recognition have already run their course, and objects have been recognized at a lower level, therefore x_{ni} are 0 or 1. Given mathematical formulation can use dynamic signals x_{ni}, parts of object-recognition simulators. We remind that simulators of interest in this example are situations; in addition to x_{ni} these simulators involve dynamic neural signals referred by p_{mi}. These are constituent signals of the ongoing simulator processes at the considered level of situations, which learn to recognize situations, symbols at a higher level (relative to objects).

The elements of vector $\mathbf{p}_m$ are modeled as independent (this is not essential, if presence of various objects in a situation actually is correlated, this would simplify learning, e.g. perfect correlation would make it trivial). Correspondingly, conditional probability of observing vector $\mathbf{X}_n$ in a situation m is then given by the standard formula

$$l\left(\mathbf{X}(n)\middle|\mathbf{M}_m(n)\right) = \prod_{i=1}^{Do} p_{mi}^{x_{ni}}\left(1-p_{mi}\right)^{(1-x_{ni})} \qquad (4)$$

Consider N perceptions a child was exposed to (N includes real "situations" and "irrelevant" random ones); among them most perceptions were "irrelevant" corresponding to observing random sets of objects, and M-1 "real" situations, in which D_s objects were repeatedly present. All random observations we model by 1 model ("noise"); assuming that every object has an equal chance of being randomly observed in noise (which again is not essential) the probabilities for this noise model, m = 1, are $p_{1i} = 0.5$ for all i. Thus we define M possible sources for each of the N observed situations.

The total similarity for the above M models (M-1 "real" and 1 noise) is given by the same equation as similarity in the previous example [16]. And the same DL equations maximize it over the parameters, which in this case are the probabilities of objects constituting various situations. Specifics of this case of situations are given by models, which exact form is given by Equation (4). The general DL equations in Section 2.3 can be significantly simplified in this case [84] and we will describe them now in details. The DL is an iterative process, it involves a sequence of interactions between bottom-up and top-down signals; this sequence involves first, association variables connecting these signals, and second, parameter update equations, which improve parameter values in this interaction-learning. We use shorthand notations l(n|m) for conditional probabilities (1). The association variables connecting bottom-up signals n with top-down projections-simulators m are given by the general equation, as in Section 2.3,

$$f\left(m\middle|n\right) = r(m)l\left(m\middle|n\right)\Big/\sum_{m'\in M} r(m')l\left(n\middle|m'\right) \qquad (5)$$

For intuitive understanding we point out that these association variables are different from Equation (1) in that they are normalized by the denominator, the sum of l(m|n) for a given bottom-up signal over all active simulators m. Whereas l(m|n) could vary greatly in their values, f(m|n)

vary between 0 and 1. Also, Equation (5) contain parameters r(m), which are needed for the following reason: it is convenient to define conditional probabilities (1) assuming the simulator m actually is present and active; therefore r(m) are needed to define the actual probability of the simulator m in the current process. These association variables (5) are defined similarly to the a posteriori Bayesian probabilities that a bottom-up signal n comes from a situation-simulator m. Yet they are not probabilities as long as parameters p_{mi} have wrong values. In the process of learning, these parameters attain their correct values and, at the end of the DL-simulator processes, the association variables can be interpreted as probabilities.

These association variables are used to update parameter values, which is the second part of the DL process. In this case parameter update equations are simple,

$$p_{mi} = \sum_{n\in N} f(m|n)\, x_{ni} \Big/ \sum_{n'\in N} f(m|n') \qquad (6)$$

These equations have very simple interpretations: they estimate parameters p_{mi} of the m-th simulator as weighted averages of the bottom-up signals, x_{ni}. Note, the bottom-up signals "do not know" which situation they came from. The weights (f/Σf) are normalized association variables, associating data n with simulator m. These equations are easy to understand: if the object i never appears in a situation m, at the end of the DL-simulator learning process, f(m|n) = 0, and p_{mi} = 0, as it should be, even if x_{ni} are not 0 because object i appears in other situations. The role of the normalizing denominator (Σf) is easy to understand; for example, if object i is actually present in situation m, then x_{ni} = 1 for each set of bottom up signal n, whenever situation m is observed. In this case, at the end of DL-simulator process, f(m|n) = 1 for all n, and Equation (3) yields p_{mi} = 1, as it should be.

We did not mention modeling relations among objects. Spatial, temporal, or structural connections, such as "to the left," "on top," or "connected" can be easily added to the above DL formalism. Relations and corresponding markers (indicating which objects are related) are no different mathematically than objects, and can be considered as included in the above formulation. This mechanism is "flat" in the hierarchical structure of the brain, meaning that relations "reside" at the same level as entities they relate. Alternatively, some relations are realized in the brain hierarchically: relations could "reside" at a higher level, with markers being implemented similar to parametric models. Experimental data might help to find out, which relations are "flat" and which are "hierarchical". Other types of relations are principally hierarchical, e.g. objects-features, situations-objects, etc. We would also add that some relations are not "directly observable", as objects; say to differentiate between "glued to" or "stack to" might require knowledge of human actions or how the world is organized. Prerequisites to some of this knowledge might be inborn [86,87]. We suggest that directly observable relations are learned as parts of a situation, similar to objects, and this learning is modeled by the DL formalism described above. Relations that require human cultural knowledge may be learned with the help of language, as discussed later, and inborn mechanisms should be further elucidated experimentally. This discussion implies several predictions that could be experimentally tested: existence of two types of relation mechanisms, flat and hierarchical; suggestions of which types of mechanisms are likely to be used in the brain for which types of relations; and suggestions of mechanisms conditioned by culture and language.

The above formulation, let us repeat, assumes that all the objects have already been recognized, still the above formulation can be applied without any change to real, continuously working brain with multiplicity of concurrently running simulators at many levels, feeding each other. Also modality of objects (various sensorial or motor mechanisms) requires no modifications; emotions can be included as well, some emotions are reducible to representations and learned to be a part of a situation similar to objects; other involve entirely different mechanisms discussed later [60,88-92]). The bottom up signals do not have to be definitely recognized objects, these signals can be sent before objects are fully recognized, while object simulators are still running and object representations are vague; this would be represented by x_{ni} values between 0 and 1. The bottom up signals do not have to correspond to complete objects, but could recreate patterns of activetions in sensorimotor brain areas associated with perception of objects; similarly, top-down signals corresponding to situations, $\mathbf{p}_m$, correspond to patterns of activations recreating experience of this situation. The presented formalization therefore is a general mechanism of simulators. A fundamental experimentally testable prediction of the developed theory is that top-down signals originate from vague representations, and the vagueness is determined by degrees of uncertainty of association between bottom-up signals and various higher-level representations.

3.2. Learning Symbol-Situations

This example considers the total number of recognizable objects equal to 1000 (D_o = 1000). The total number of objects perceived in a situation is set to 50 (D_p = 50). The number of essential objects is set to 10 (D_s = 10). The number of situations to learn (M-1) is set to 10. Note that the true identities of the objects are not important in this simulation so we simply use object indexes varying from 1 to 1000 (this index points to neural signals corresponding to a specific object-simulators). The situation names are also not important and we use situation indexes (this index points to neural signals corresponding to a specific situation-simulators). We would emphasize that the use

of numbers for objects and situation, while may seem consistent with amodal symbols, in fact is nothing but notations. We repeat that the principled differences between PSS and amodal systems are mechanisms in the brain and their modeling, not mathematical notations. Among these mechanisms are simulators, mathematically described by DL.

Let us repeat, amodal symbols are governed by classical logic, which is static and faces CC. DL is a process and overcomes CC. DL operates on PSS representations (models $\mathbf{p}_m$), which are vague collections of objects (some of these objects could also be vague, not completely assembled yet representations). Another principled difference is interaction between perceptual-based bottom-up and top-down neural fields $\mathbf{X}_n$ and $\mathbf{M}_m$; indexes n and m are just mathematical shorthand for corresponding neural connections. In this paper we consider object perception and situation perception in different sections, but of course the real mind-brain operates continuously, "objects" in this section are neural signals sent to situation-recognition brain area (and corresponding simulators) by excited neuron fields corresponding to models of recognized-objects (partially, as described in Section 2; and as discussed, these signals are being sent before objects are fully recognized, while object simulators are still running).

This example uses simulated data generated by first randomly selecting $D_s = 10$ specific objects for each of the 10 groups of objects, allowing some overlap between the groups (in terms of specific objects). This selection is accomplished by setting the corresponding probabilities $p_{mi} = 1$. Next we add 40 more randomly selected objects to each group (corresponding to $D_p = 50$). We also generate 10 more random groups of 50 objects to model situations without specific objects (noise); this is of course equivalent to 1 group of 500 random objects. We generate N' = 800 perceptions for each situation resulting in N = 16000 perceptions (data samples, n = 1⋯ 16000) each represented by 1000-dimensional vector $\mathbf{X}_n$. These data are shown in **Figure 2** sorted by situations.

Then the samples are randomly permuted, according to randomness of real life perceptual situations, in **Figure 3**. The horizontal lines disappear; the identification of repeated objects becomes nontrivial. An attempt to learn groups-situations (the horizontal lines) by inspecting various horizontal sortings (until horizontal lines would become detectable) would require $M^N = 10^{16000}$ inspections, which is of course impossible. This CC is the reason why the problem of learning situations has been standing unsolved for decades. By overcoming CC, DL can solve this problem as illustrated below.

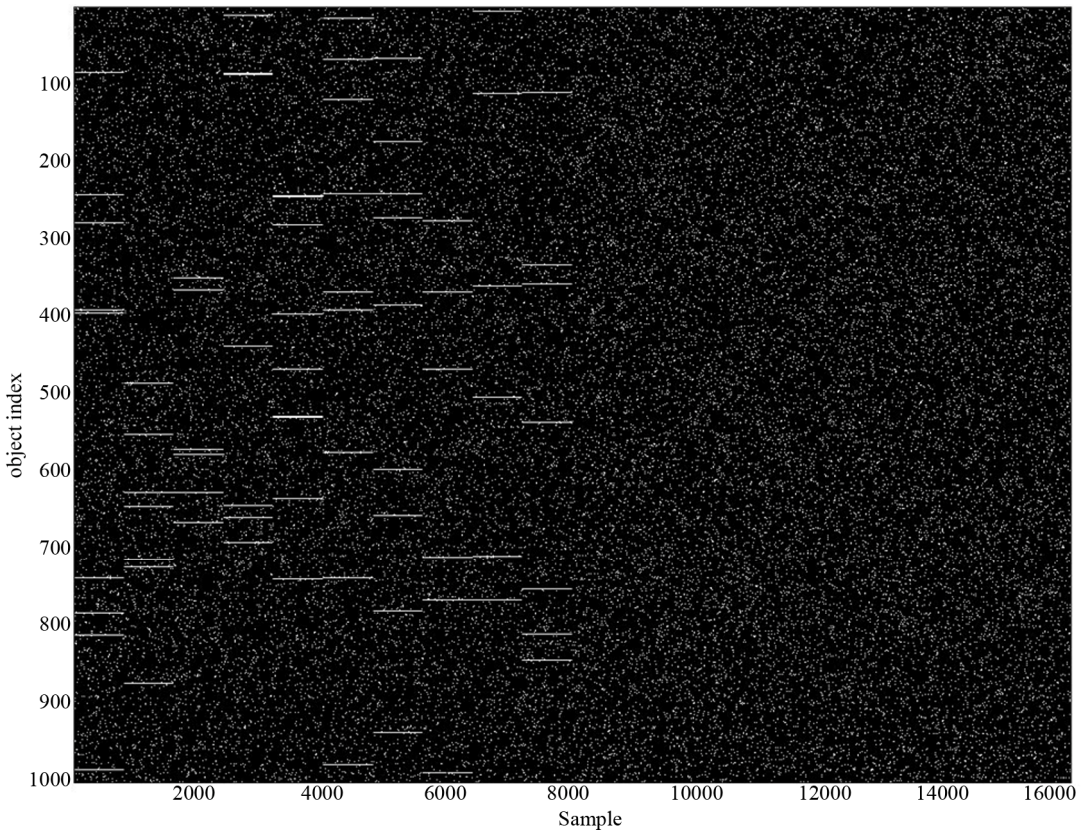


Figure 2. Generated data; object index is along vertical axes and situation index is horizontal. The perceptions (data samples) are sorted by situation index (horizontal axis); this makes visible the horizontal lines for repeated objects.

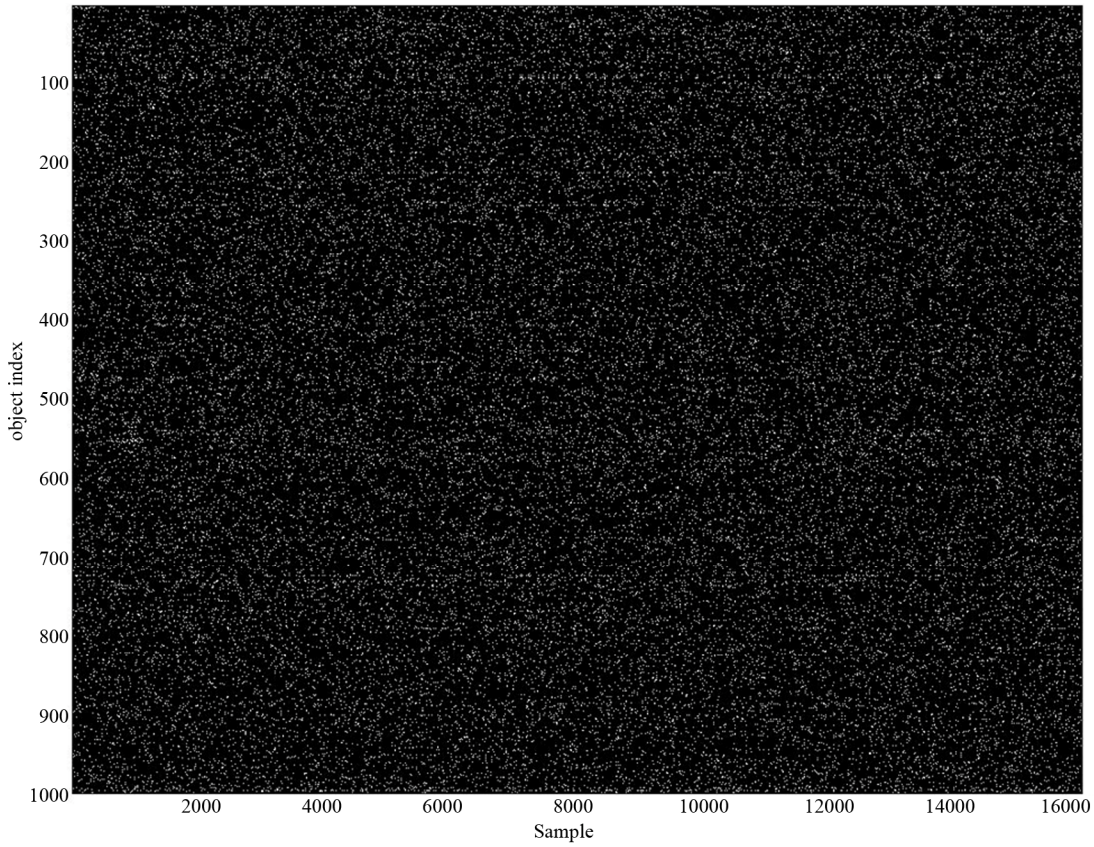


Figure 3. Data, same as Figure 2, randomly sorted by situations (horizontal axis), as available to the DL algorithm for learning.

The DL algorithm is initiated similarly to Section 2 by defining 20 situational models (an arbitrary selection, given actual 10 situations) and one random noise model to give a total of M = 21 models (in Section 2.4, **Figure 1** models were automatically added by DL as required; here we have not done this (because it would be too cumbersome to present results). The models are initialized by assigning random probability values to the elements of the models. These are the initial vague perceptual models, which assign all objects to all situations.

The initialization and the iterations of the DL algorithm (the first 3 steps of solving DL equations) are illustrated in **Figure 4**. Each subfigure displays the probability vector $\mathbf{p}_m$ for each of the 20 models. The vectors have 1000 elements corresponding to objects (vertical axes). The values of each vector element are shown in gray scale. The initial models assign nearly uniformly distributed probabilities to all objects. The horizontal axes are the model index changing from 1 to 20. The noise model is not shown. As the algorithm progresses, situation grouping improves, and only the elements corresponding to repeating objects in "real" situations keep their high values, the other elements take low values. By the third iteration the 10 situations are identified by their corresponding models. The other 10 models converge to more or less random low-probability vectors.

This fast and accurate convergence can be seen from **Figures 5** and **6**. We measure the fitness of the models to the data by computing the sum-squared error, using the following equation.

$$E = \sum_{m\in\{B\}} \sum_{i=1}^{Do} \left(p_{mi} - p_{mi}^{Ture}\right)^2$$

In this equation the first summation is over the subset {B} containing top 10 models that provide the lowest error (and correspondingly, the best fit to the 10 true models). In the real brain, of course, the best models would be added as needed, and the random samples would accumulate in the noise model automatically; as mentioned, DL can model this process and the reason we did not model it, is that it would be too cumbersome to present results. **Figure 5** shows how the sum squared error changes over the iterations of the DL algorithm. It takes only a few iterations for the DL algorithm to converge. Each of the best models contains 10 large and 990 low probabilities. Iterations stop, when average error of probabilities reached a low value of 0.05 resulting in the final error E(10) = 1000*(0.05^2)*10 = 25.

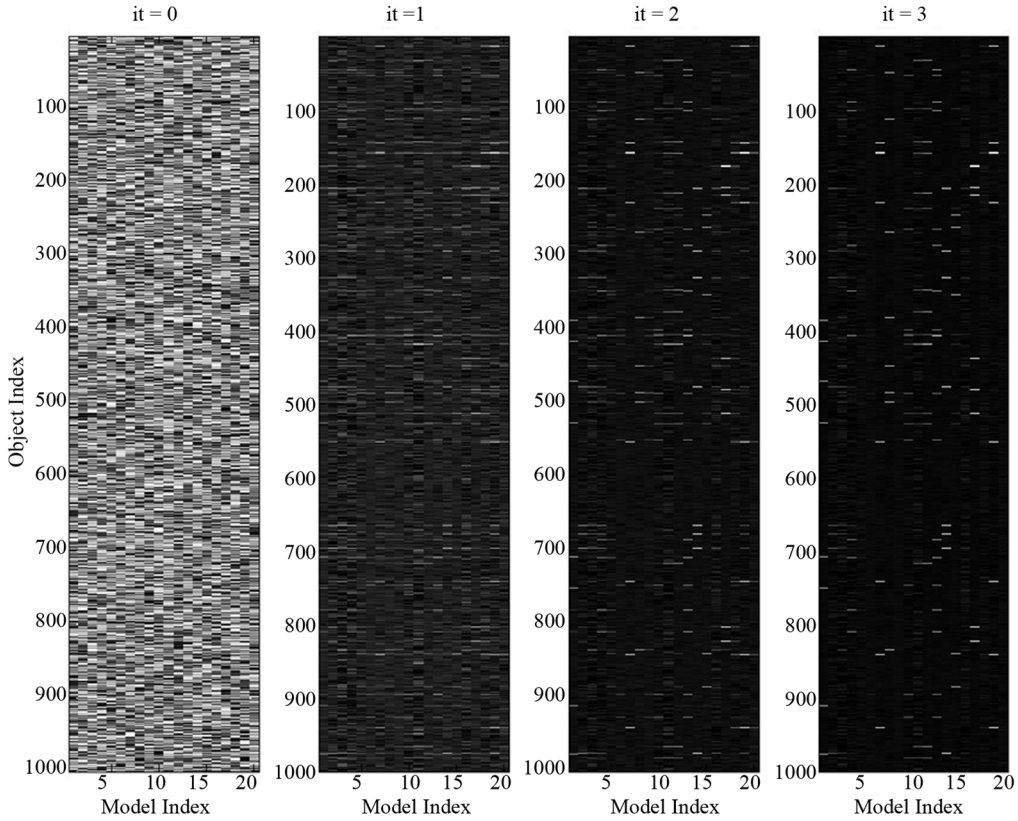


Figure 4. DL situation learning. Situation-model parameters converge close to true values in 3 steps.

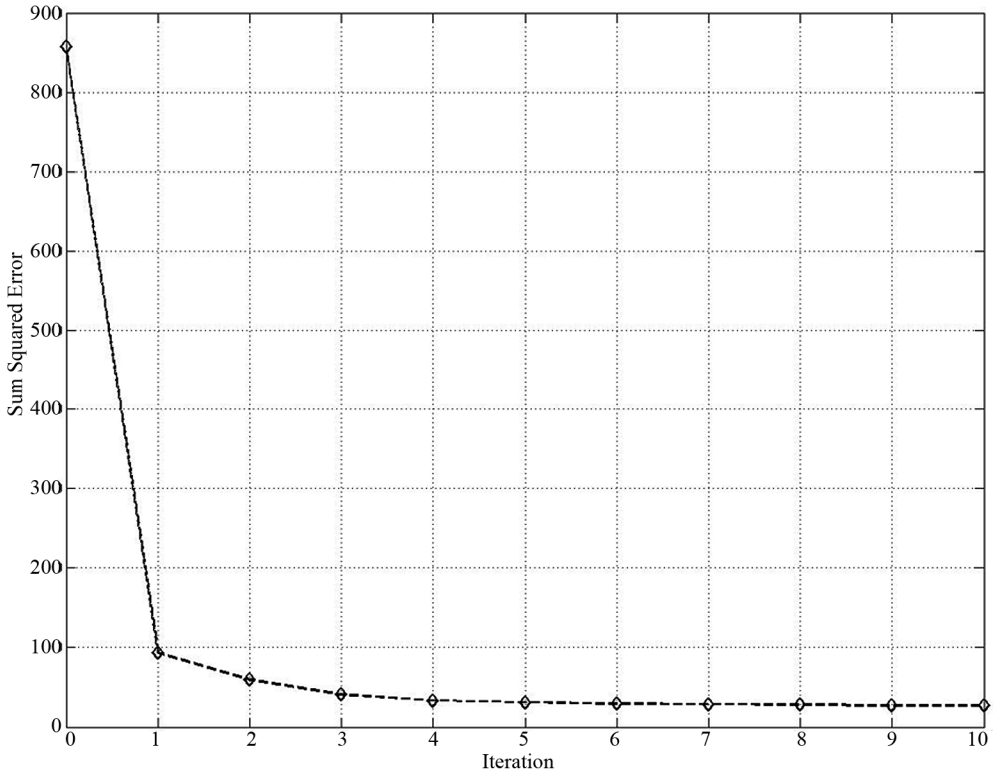


Figure 5. Errors of DL learning are quickly reduced in 3 - 4 steps, iterations continue until average error reached predetermined threshold of 0.05 (10 steps).

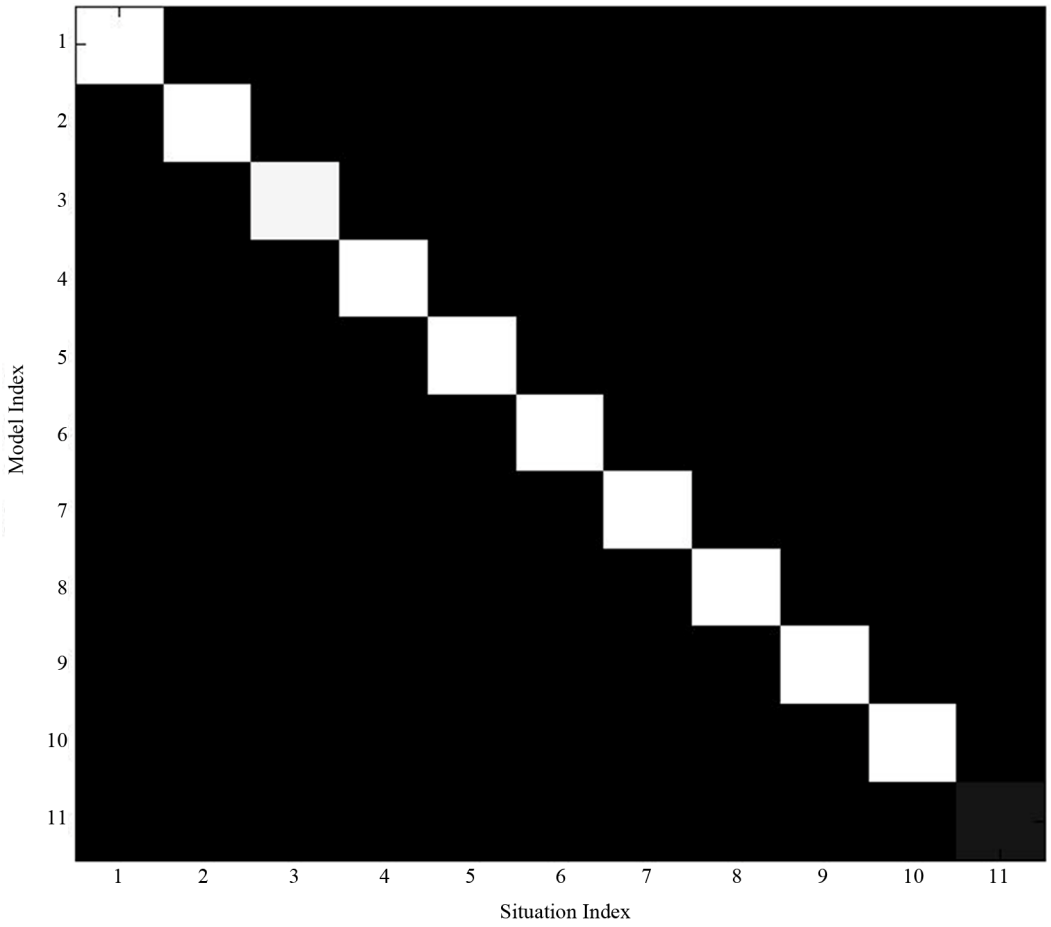


Figure 6. Correct associations are near 1 (diagonal, except noise) and incorrect associations are near 0 (off-diagonal).

Figure 6 shows average associations, A (m,m') among true (m) and computed models (m'); this is an 11 × 11 matrix according to the true number of different models (it is computed using association variables between models and data, f(m|n))

$$A(m,m') = (1/N')\sum_{n=1}^{N} f(m|n) * f(m'|n),\ m' \in \{B\},$$

$$A(m,11) = (1/10 * N') \sum_{m' \notin \{B\}} \sum_{n=1}^{N} f(m|n) * f(m'|n),\ m' \notin \{B\}$$

Here, f(m|n) for true 10 models m is either 1 (for N' data samples from this model) or 0 (for others), f(m'|n) are computed associations, in the second line all 10 computed noise models are averaged together, corresponding to one true (random) noise model. The correct associations on the main diagonal in **Figure 6** are 1 (except noise model, which is spread among 10 computed noise models, and therefore equals 0.1) and off-diagonal elements are near 0 (incorrect associations, corresponding to small errors shown in **Figure 5**).

Errors in **Figure 5** have not converged exactly to 0. The reason is numerical, p and x variables in Equation (4) cannot be allowed to take exactly 0 value, since expressions 0^0 are numerically non-defined, therefore all values have been bounded from below by 0.05 (a somewhat arbitrary limit). **Figure 6** demonstrates that nevertheless, convergence to the global maximum was achieved (the exactly known solution in terms of learning the correct situations).

Similar to Section 2, learning of perceptual situation-symbols has been accomplished due to the DL process-simulator, which simulated internal model-representtations of situations, **M**, to match patterns in bottom-up signals **X** (sets of lower-level perceptual object-symbols).

4. Concepts, Simulators, Embodiment, and Binding

PSS grounds perception, cognition, and high-level symbol operation in modal symbols, which are ultimately grounded in the corresponding brain systems. Previous section provides an initial development of formal mathematical description suitable for PSS: the DL process "from vague-to-crisp" models PSS simulators. We have considered just one subsystem of PSS, a mechanism of learning, formation, and recognition of situations from objects making up the situations. (More generally, the formalized mecha-

nism of simulators includes recognition of situations by recreating patterns of activations in sensorimotor brain areas, from objects, relations, and actions making up the situations). The mind's representations of situations are symbol-concepts of a higher level of abstractness than symbol-objects making them up. The proposed mathematical formalism can be advanced to "higher" levels of more and more abstract concepts. Such an application to more abstract ideas, however, may require an additional grounding in language [42,90,92,93,96-101] as we briefly consider in the next section.

The proposed mathematical formalism can be similarly applied at a lower level of recognizing objects as constructed from their parts; mathematical techniques of Sections 2 and 3 can be combined to implement this PSS object recognition idea as described in [1]. Specific models considered in Section 2 are likely to be based on inborn mechanisms specific to certain aspects of various sensor and motor modalities; general models of Section 3 can learn to represent and recognize objects as collections of multi-modal perceptual features and relations among them. In both cases principal mechanisms of object perception such as discussed in [102] can be modeled, either as properties of object models, or as relations between perceptual features. Since relations specific to object recognition, according to this reference are learned in infancy, the mechanism of Section 3 seems appropriate (it models learning of relations, whereas models in Section 2 do not readily contain mechanisms of learning of all their structural aspects and are more appropriate to modeling inborn mechanisms). Object representations, as described by Barsalou are not similar to photographs of specific objects, but similar to models in **Figure 4** are more or less loose and distributed (among modalities) collections of features and relations.

The developed DL theory, by modeling the simulators, also mathematically models productivity of the mind concept-simulator system. The simulated situations and other concepts are used not only in the process of matching bottom-up and top-down signals for learning and recognizing representations, but also in the motor actions, and in the processes of imagination and planning.

Presented examples are steps toward general solution of the binding problem discussed in [103]. That publication discusses the role of context similar to the DL scene modeling. DL is a general approach to solving the binding problem. Here we would emphasize two mechanisms of binding modeled in the developed theory. First, binding is accomplished hierarchically: e.g. object representations-simulators bind features into objects, similarly situation representations-simulators bind objects into situations, etc. Second, binding is accomplished by relations that are learned similarly to objects and "reside" at the same level in the hierarchy of the mind with the bound entities. These two types of binding mechanisms is another novel prediction of the DL theory that could be tested experimentally.

Now we discuss other relationships between the mathematical DL procedures of previous sections and the fundamental ideas of PSS. Section 2 concentrated on the principal mathematical difficulty experienced by all previous attempts to solve the problem of complex symbol formation from less complex symbols, the combinatorial complexity (CC). CC was resolved by using DL, a mathematical theory, in which learning begins with vague (non-specific) symbol-concepts, and in the process of learning symbol-concepts become concrete and specific. Learning could refer to a child's learning, which might take days or months or an everyday perception and cognition, taking approximately $1/6^{th}$ of a second (in the later case learning refers to the fact that every specific realization of a concept in the world is different in some respects from any previous occurrences, therefore learning-adaptation is always required; in terms of PSS, a simulator always have to re-assemble the concept). In the case of learning situations as compositions of objects, the initial vague state of each situation-symbol is a nearly random and vague collection of objects, while the final learned situation consists of a crisp collection of few objects specific to this situation. This specific of the DL process "from vague-to-crisp" is a prediction that can be experimentally tested, and we return to this later. In the learning process random irrelevant objects are "filtered out," their probability of belonging to a concept-situation is reduced to zero, while probabilities of relevant objects, making up a specific situation is increased to a value characteristic of this object being actually present in this situation. Relation of this DL process to PSS is now considered. First we address concepts and their development in the brain. According to [1], "The central innovation of PSS theory is its ability to implement concepts and their interpretative functions using image content as basic building blocks".

DL implements this aspect of PSS theory in a most straightforward way. Concept-situations in DL are collections of objects (symbol-models at lower levels, which are neurally connected to neural fields of object-images). As objects are perceptual entities-symbols in the brain, concept-situations are collections of perceptual symbols. In this way situations are perceptual symbols of a higher order complexity than object-symbols, they are grounded in perceptual object-symbols (images), and in addition, their learning is grounded in perception of images of situations. A PSS mathematical formalization of abstract concepts [104], not grounded in direct perceptions, is considered in the next section. Here we just mention that the proposed model is applicable to higher levels, "beyond" object-situations; it is applicable to modeling interactions between bottom-up and top-down signals at every level.

According to PSS, concepts are developed in the brain by forming collections of correlated features [105,106]. This is explicitly implemented in the DL process described in Section 3. The developed mathematical representation corresponds to multimodal and distributed representation in the brain. It has been suggested that a mathematical set or collection is implemented in the brain by a population of conjunctive neurons [9].

DL processes are mathematical models of PSS simulators. DL symbol-situations are not static collections of objects but dynamic processes. In the process of learning they "interpret individuals as tokens of the type" [105]. They model multi-modal distributed representations (including motor programs) as described in the reference.

The same DL mathematical procedure can apply to perception of a real situation in the world as well as an imagined situation in the mind. This is the essence of imagination. Models of situations (probabilities of various objects belonging to a situation, and objects attributes, such as their locations) can depend on time, in this way they are parts of simulators accomplishing cognition of situations evolving in time. If "situations" and "time" pertain to the mind's imaginations, the simulators implement imagination-thinking process, or planning.

Usually we perceive-understand a surrounding situation, while at the same time thinking and planning future actions and imagine consequences. This corresponds to running multiple simulators in parallel. Some simulators support perception-cognition of the surrounding situations as well as ongoing actions, they are mathematically modeled by DL processes that converged to matching internal representations (types) to specific subsets in external sensor signals (tokens). Other simulators simulate imagined situations and actions related to perceptions, cognitions, and actions, produce plans, etc.

Described here DL mathematical models corresponds to what [104] called dynamic interpretation of PSS (DIPSS). DIPSS is fundamental to modeling abstraction processes in PSS. Three central properties of these abstractions are type-token interpretation; structured representation; and dynamic realization. Traditional theories of representation based on logic model interpretation and structure well but are not sufficiently dynamical. Conversely, connectionist theories are dynamic but are inadequate at modeling structure. PSS addresses all three properties. Similarly, the DL mathematical process developed here addresses all three properties. In type-token relations "propositions are abstractions for properties, objects, events, relations and so forth. After a concept has been abstracted from experience, its summary representation supports the later interpretation of experience". Correspondingly in the developed mathematical approach, DL models a situation as a loose collection of objects and relations. Its summary representation (the initial model) is a vague and loose collection of property and relation simulators, which evolves-simulates representation of a concrete situation in the process of perception of this concrete situation according to DL. This DL process involves structure (initial vague models) and dynamics (the DL process).

5. Extension of DL beyond PSS: Abstract Concepts, Language, the Mind Hierarchy, and Emotions

DL is a general model of interacting bottom-up and top-down signals throughout the hierarchy-heterarchy of the mind-brain, including abstract concepts. The DL mathematical analysis suggests that modeling the process of learning abstract concepts has to go beyond PSS analysis in [1]. In particular, we discuss the role of language in learning abstract concepts [90,93,96,99,106-112] and connect it to the PSS mechanisms. The mind-brain is not a strict hierarchy, interactions across layers are present and this is some times addressed as heterarchy [18]. To simplify discussion we would use a term hierarchy.

In Section 3 situation representations have been assembled from object representations. This addresses interaction between top-down and bottom-up signals in two adjacent layers of the mind hierarchy. The mathematical description in Section 3 addresses top-down and bottom-up signals and representations without explicit emphasis on their referring to objects or situations. Accordingly, we would emphasize here that the mathematical formulation in Section 3 equally addresses interaction between any two adjacent layers in the entire hierarchy of the mind-brain, including high-level abstract concepts. A fundamental question of embodiment is discussed now.

Barsalou assumed that higher level abstract concepts remain grounded-embodied in PSS since they are based on lower level grounded concepts, and down the hierarchy to perceptions directly grounded in sensory-motor signals. The DL modeling suggests that this aspect of the PSS theory has to be revisited. First, each higher level is vaguer than a lower level. Several levels on top of each other would result in representations too vaguely related to sensory-motor signals to be grounded in them with any reliability. Second, the Section 3 example is impressive in its numerical complexity, which significantly exceeds anything that has been computationally demonstrated previously. We would like to emphasize again that this is due to overcoming difficulty of CC. Still statistically, learning of situations was based on these situations being present among the data with statistically sufficient information to distinguish them among each other and from noise. In real life however, human learn complex abstract concepts, such as "system", "rationality", "survival", and many other abstract concepts, without statistically sufficient information been experienced. This is

possible due to language.

Language is learned at all levels of the hierarchy of the mind-brain and cultural knowledge from surrounding language. This is possible because language models exist "ready made" in the surrounding language. Language can be learned without real life experience, therefore kids can talk about much of cultural contents by the age of five or seven. At this age kids can talk about many abstract ideas, which they cannot yet adequately use in real life. This corresponds to a significant difference in language and cognitive representations, and their different locations in the brain. Language concepts are grounded in surrounding language at all hierarchical levels. But learning corresponding cognitive concepts grounded in life experience takes entire life. Learning language, like learning cognition can be modeled by DL. Linguists consider words to be learned by memorizing them [87]. Learning meaningful phrases and syntax is similar to learning situations and relations among objects in Section 3. Morphology is not unlike object composition.

A fundamental difference of language from cognition is embodiment-grounding. Let us repeat, language is grounded in direct experience (of talking, reading) at all levels of the hierarchy, whereas cognition is grounded in direct perceptions only at the bottom of the hierarchy. Higher abstract levels of cognition are grounded in language. The detailed theory of interaction between cognition and language is considered in [16,37,43,83,93,94, 100-108,107,110,112-117]. It leads to a number of verifiable experimental predictions, which we summarize in Section 8. The main mechanism of interaction between language and cognition is the dual model, which suggests that every mental model-representation has two neurally connected parts, language model and cognitive model. Language models are learned by simulator processes, similar to PSS simulators, however, "perception" in case of language refers to perception of language facts. Through neural connections between the two parts of each model, the early acquired abstract language models guide the development of abstract cognitive models in correspondence with cultural experience stored in language. The dual model leads to the dual hierarchy illustrated in **Figure 7**.

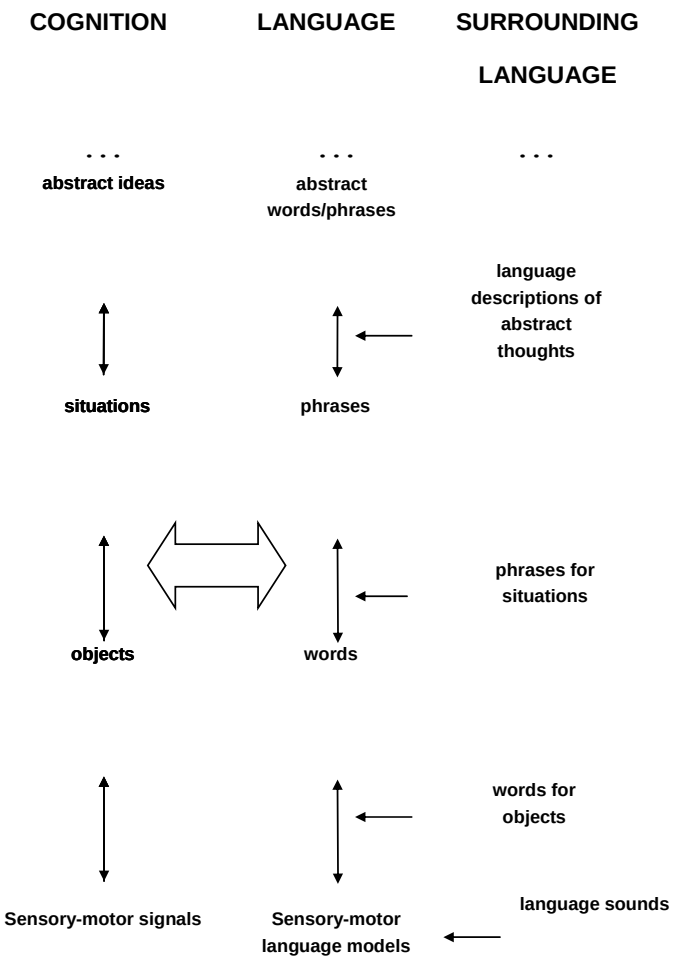


Figure 7. The dual-model architecture, modeling interaction of language and cognition. Learning of cognition is grounded in experience and guided by language. Learning of language is grounded in the surrounding language at all hierarchical levels.

The dual model mechanism explains how language affects cognition. It follows that language affects culture and its evolution [56,88,108,118,119].

6. Mathematical Models and Perceptual vs Amodal Symbols

Any mathematical notation may look like an amodal symbol, so in this section we discuss the roles of amodal vs perceptual symbols in DL. This would require clarification of the word symbol. We touch on related philosophical and semiotic discussions and relate them to mathematics of DL and to PSS. For the sake of brevity within this paper we limit discussions to the general interest, emphasizing connections between DL, perceptual, and amodal symbols; extended discussions of symbols can be found in [16,43,83,120]. We also summarize here related discussions scattered throughout the paper.

"Symbol is the most misused word in our culture" [121]. Why the word "symbol" is used in such a different way: to denote trivial objects, like traffic signs or mathematical notations, and also to denote objects affecting entire cultures over millennia, like Magen David, Swastika, Cross, or Crescent? Let us compare in this regard opinions of two founders of contemporary semiotics, Charles Peirce [122] and Ferdinand De Saussure [123]. Peirce classified signs into symbols, indexes, and icons. Icons have meanings due to resemblance to the signified (objects, situations, etc.), indexes have meanings by direct connection to the signified, and symbols have meaning due to arbitrary conventional agreements. Saussure used different terminology, he emphasized that signs receive meanings due to arbitrary conventions, whereas symbol implies motivation. It was important for him that motivation contradicted arbitrariness. Peirce concentrated on the process of sign interpretation, which he conceived as a triadic relationship of sign, object, and interpretant. Interpretant is similar to what we call today a representation of the object in the mind. However, this emphasis on interpretation was lost in the following generation of scientists. This process of "interpretation" is close to the DL processes and PSS simulators. We therefore follow Saussurean designation of symbol as a motivated process. Motivationally loaded interpretation of symbols was also proposed by Jung [124]. He considered symbols as processes bringing unconscious contents to consciousness. Similar are roles of PSS simulators and DL processes.

In scientific understanding of symbols and semiotics, the two functions, understanding the language and understanding the world, have often been perceived as identical. This tendency was strengthened by considering logical rules to be the mechanism of both, language and cognition. According to Russell [125], language is equivalent to axiomatic logic, "[a word-name] merely to indicate what we are speaking about; [it] is no part of the fact asserted… it is merely part of the symbolism by which we express our thought". Hilbert [46] was sure that his logical theory also describes mechanisms of the mind, "The fundamental idea of my proof theory is none other than to describe the activity of our understanding, to make a protocol of the rules according to which our thinking actually proceeds". Similarly, logical positivism centered on "the elimination of metaphysics through the logical analysis of language"—according to Carnap [126] logic was sufficient for the analysis of language. As discussed in Section 2.2, this belief in logic is related to functioning of human mind, which is conscious about the final states of DL processes and PSS simulators; these final states are perceived by our minds as approximately logical amodal symbols. Therefore we identify amodal symbols with these final static logical states, signs.

DL and PSS explain how the mind constructs symbols, which have psychological values and are not reducible to arbitrary logical amodal signs, yet are intimately related to them. In Section 3 we have considered objects as learned and fixed. This way of modeling objects indeed is amenable to interpreting them as amodal symbols-signs. Yet, we have to remember that these are but final states of previous simulator processes, perceptual symbols. Every perceptual symbol-simulator has a finite dynamic life, and then it becomes a static symbol-sign. It could be stored in memory, or participate in initiating new dynamical perceptual symbols-simulators. This infinite ongoing dynamics of the mind-brain ties together static signs and dynamic symbols. It grounds symbol processes in perceptual signals that originate them; in turn, when symbol-processes reach their finite static states-signs, these become perceptually grounded in symbols that created them. We become consciously aware of static sign-states, express them in language and operate with them logically. Then, outside of the mind-brain dynamics, they could be transformed into amodal logical signs, like marks on a paper. Dynamic processes-symbols-simulators are usually not available to consciousness. These PSS processes involving static and dynamic states are mathematically modeled by DL in Section 3 and further discussed in Section 4.

To summarize, in this paper we follow a tradition using a word sign for an arbitrary, amodal, static, unmotivated notation (unmotivated means unemotional, in particular). We use a word symbol for the PSS and DL processes-simulators, these are dynamic processes, connecting unconscious to conscious; they are motivationally loaded with emotions. As discussed in Section 2, DL processes are motivated toward increasing knowledge, and they are loaded with knowledge-related emotions, even in absence of any other motivation and emotion. These knowledge-related emotions are called aesthetic emotions since Kant. They are foundations of higher

cognitive abilities, including abilities for the beautiful, sublime, and they are related to musical emotions. More detailed discussions can be found in [11-13,15,16,36,38, 40-42,54, 66,83,88,89,93,95,96,100,101,126-138].

DL mathematical models (in Section 3) use mathematical notations, which could be taken for amodal symbols. Such an interpretation would be erroneous. Meanings and interpretations of mathematical notations in a model depends not on the appearance, but on what is modeled. Let us repeat, any mathematical notation taken out of the modeling context, is a notation, a static sign. In DL model-processes these signs are used to designate neuronal signals, dynamic entities evolving "from vague to crisp" and mathematically modeling processes of PSS simulators-symbols. Upon convergence of DL-PSS simulator processes, the results are approximately static entities, approximately logical, less grounded and more amodal.

Grounded dynamic symbol-processes as well as amodal static symbols governed by classical logic can both be modeled by DL. DL operates on a non-logical type of PSS representations, which are vague combinations of lower-level representations. These lower-level representations are not necessarily complete images or events in their entirety, but could include bits and pieces of various sensor-motor modalities, memory states, as well as vague dynamic states from concurrently running simulators—DL processes of the on-going perception-cognition (in Section 3, for simplicity of presentation, we assumed that the lower-level object-simulators have already run their course and reached static states; however, the same mathematical formalism can model simulators running in parallel on multiple hierarchical levels). The mind-brain is not a strict hierarchy, the same-level and higher-level representations could be involved along with lower levels. DL models processes-simulators, which operate on PSS representations. These representations are vague and incomplete, and DL processes are assembling and concretizing these representations. As described in several references by Barsalou, bits and pieces from which these representations are assembled could include mental imagery as well as other components, including multiple sensor, motor, and emotional modalities; these bits and pieces are mostly inaccessible to consciousness during the process dynamics.

DL also explains how logic and ability to operate amodal symbols originate in the mind from illogical operations of PSS: mental states approximating amodal symbols and classical logic appear as the end of the DL process-simulators. At this moment they become conscious static representations and loose that component of their emotional-motivational modality, which is associated with the need for knowledge (to qualify as amodal, these mental states should have no sources of modality, including emotional modality). The developed DL formalization of PSS, therefore suggests using a word signs for amodal mental states as well as for amodal static logical constructs outside of the mind, including mathematical notations; and to reserve symbols for perceptually grounded motivational cognitive processes in the mind-brain. Memory states, to the extent they are static entities, are signs in this terminology. Logical statements and mathematical signs are perceived and cognized due to PSS simulator symbol-processes and become signs after being understood. Perceptual symbols, through simulator processes, tie together static and dynamic states in the mind. Dynamic states are mostly outside of consciousness, while static states might be available to consciousness.

7. Experimental Evidence

Neuroimaging experiments demonstrated that DL is a valid model for visual perception [19]. Experimental procedures in this reference used functional Magnetic Resonance Imaging (fMRI) to obtain high-spatial resolution of processes in the brain, which they combined with magneto-encephalography (MEG), measurements of the magnetic field next to the head, which provided high temporal resolution of the brain activity. Combining these two techniques the experimenters were able to receive high resolution of cognitive processes in space and time. Bar *et al*. concentrated on three brain areas: early visual cortex, object recognition area (fusiform gyrus), and object-information semantic processing area (OFC). They demonstrated that OFC is activated 130 ms after the visual cortex, but 50 ms before object recognition area. Their conclusion has been that OFC represents the cortical source of top-down facilitation in visual object recognition. This top-down facilitation is unconscious. They demonstrated that the imagined image generated by top-down signals facilitated from OFC to cortex is vague (the authors in this publication refer to low spatial-frequency content images), confirming the essential mechanism of DL. Conscious perception of an object occurs when vague projections become crisp and match the crisp and clear image from the retina, and an object recognition area is activated.

From sensory data the brain predicts expected perception and cognition. This includes predictions of complex information, such as situations and social interactions [139,140]. Predictions are initiated by gist information rapidly extracted from sensory data. At the "lower"-object level this gist information is a vague image of an object. At higher levels "the representation of gist information is yet to be defined". The model presented here defines this higher-level gist information as vague collections of vague objects, with relevant objects for a specific situation having just slightly higher probabilities than irrelevant ones. The model is also consistent with the hypothesis in [139] that cognition at higher levels

relies on mental simulations. Mathematical predictions in this paper suggest specific properties of these higher-level simulators (vagueness of top-down representations is represented by their associations with bottom-up signals).

8. Future Research

The DL mathematical description of PSS should be addressed throughout the mind hierarchy; from features and objects "below situations" in the hierarchy to abstract models and simulators at higher levels "above situations". Modeling across the mind modalities will be addressed including diverse modalities, symbolic functions, conceptual combinations, predication. Modeling features and objects would have to account for suggestions that perception of sensory features are partly inborn [1]. The developed DL formalization of PSS corresponds to observations in [141] and it will be used for generating more detailed experimentally verifiable predictions. A number of predictions have been made in this paper, including influence of perception on cognition [142].

The developed theory provides solutions to classical problems of conceptual relations, binding, and recursion. Binding is a mechanism connecting events into meaningful "whole" (or larger-scale events). The DL model developed here specifies two types of binding mechanisms "flat" and "hierarchical", and suggests which mechanisms are likely to be used for various relations. Our model also suggests existence of binding mechanisms conditioned by culture and language. Recursion has been postulated to be a fundamental mechanism in cognition and language [103], however, that reference has not proposed specific mechanisms of how recursion creates representations, nor how it maps representations into the sensory-motor or conceptual-intentional interfaces. In our opinion this is an erroneous assumption, and the error is similar to mistaking logical signs for symbol-processes (recursion is an important logical operation). According to the developed theory recursion is not a fundamental mechanism, instead the hierarchy is proposed as a mechanism of recursion. Successive hierarchical levels accomplish recursive cognitive and linguistic functions. These predictions will be experimentally tested.

Predicted properties of higher-level simulators will be addressed in experimental research [19,139]. These include a prediction that early predictive stages of situation simulations are vague. Whereas vague predictions of objects resemble low-spatial frequency of object imagery [19], "the representation of gist information on higher levels of analysis is yet to be defined" [139]. According to the developed model, vague predictions of situations should contain many less-relevant (and likely vague) objects with lower probabilities. Since the mathematical model proposed here is applicable to higher levels ("above" object-situations), this hypothesis should be relevant to the nature of information of higher-level gists: initial representation of abstract concepts are vague in terms of associations with constituent bottom-up signals (these associations are not exact, but vague; probabilistically, they are not close to zeroes and ones).

Another topic discussed in [139], "how the brain integrates and holds simultaneously information from multiple points in time" also can be represented by the present model. Two different mechanisms will be explored: first, explicit incorporation of time into models (so that model parameters-probabilities depend on time), and second, categorized temporal relations, such as "before", "after" can be included as any other relations-objects into models. A joint mathematical-experimental approach will be appropriate.

Experimental and theoretical future research will address interaction between language and cognition. Language is acquired from surrounding language, rather than from embodied experience; language therefore is closer aligned with amodal symbols than with perceptual symbols. According to the developed theory, higher abstract concepts could be stronger grounded in language than in perception; using language the mind may operate with abstract concepts as with amodal symbols, and therefore have limited embodied understanding grounded in experience of how abstract concepts relate to the world. Higher-level concepts may be less grounded in perception and experience than in language. The developed theory suggests several testable hypotheses: (i) the dual model, postulating separate cognitive and language mental representations; (ii) neural connections between cognitive and language mental representations; (iii) language mental representations guiding acquisition of cognitive representations in ontological development; (iv) abstract concepts being more grounded in language than in experience; and (v) this shift from grounding in perception and experience to grounding in language progresses up the hierarchy of abstractness; while grounding in perception and experience increases with age. These make a fruitful field for future experimental research.

The suggested dual model of interaction between language and cognition bears on language evolution, and future research should address theoretical and experimental tests of this connection between evolution of languages, cognition, and cultures [42,61,88,97,112-117,134, 143-148].

Emotions in language and cognition have been addressed in [88,149]. Future research would explore roles of emotions (i) in language-cognition interaction), (ii) in symbol grounding, (iii) the role of aesthetic and musical emotions in cognition [36,54,128,131,134-136,150], and (iv) emotions of cognitive dissonances [151,152].

REFERENCES

[1] L. W. Barsalou, "Perceptual Symbol Systems," *Behavioral and Brain Sciences*, Vol. 22, No. 4, 1999, pp. 577-660.

[2] S. M. Kosslyn, "Image and Mind," Harvard University Press, Cambridge, 1980.

[3] S. M. Kosslyn, "Image and Brain," MIT Press, Cambridge, 1994.

[4] L. W. Barsalou, "Abstraction as Dynamic Interpretation in Perceptual Symbol Systems," In: L Gershkoff-Stowe and D. Rakison, Eds., *Building Object Categories*, *Carnegie Symposia Series*, Erlbaum, Mahwah, 2005, pp. 389-431.

[5] L. W. Barsalou, "Grounding Symbolic Operations in the Brain's Modal Systems," In: G. R. Semin and E. R. Smith, Eds., *Embodied Grounding*: *Social*, *Cognitive*, *Affective*, *and Neuroscientific Approaches*, Cambridge University Press, New York, 2007.

[6] L. W. Barsalou and C. R. Hale, "Components of Conceptual Representation: From Feature Lists to Recursive Frames," In: I. Van Mechelen, J. Hampton, R. Michalski and P. Theuns, Eds., *Categories and Concepts*: *Theoretical Views and Inductive Data Analysis*, Academic Press, New York, 1993.

[7] A. Cangelosi, A. Greco and S. Harnad, "From Robotic Toil to Symbolic Theft: Grounding Transfer from Entry-Level to Higher-Level Categories," *Connection Science*, Vol. 12, No. 2, 2000, pp. 143-162. doi:10.1080/09540090050129763

[8] A. Cangelosi and T. Riga, "An Embodied Model for Sensorimotor Grounding and Grounding Transfer: Experiments with Epigenetic Robots," *Cognitive Sciences*, Vol. 30, No. 4, 2006, pp. 673-689. doi:10.1207/s15516709cog0000_72

[9] W. K. Simmons and L. W. Barsalou, "The Similarity-in-Topography Principle: Reconciling Theories of Conceptual Deficits," *Cognitive Neuropsychology*, Vol. 20, 2003, pp. 451-486. doi:10.1080/02643290342000032

[10] L. I. Perlovsky, "Computational Concepts in Classification: Neural Networks, Statistical Pattern Recognition, and Model Based Vision," *Journal of Mathematical Imaging and Vision*, Vol. 4, No. 1, 1994, pp. 81-110. doi:10.1007/BF01250006

[11] L. I. Perlovsky, "Physical Concepts of Intellect," *Proceedings of Russian Academy of Sciences*, Vol. 354, No. 3, 1997, pp. 320-323.

[12] L. I. Perlovsky, "Neural Networks and Intellect: Using Model-Based Concepts," 3rd Edition, Oxford University Press, New York, 2001.

[13] L. I. Perlovsky, "Physical Theory of Information Processing in the Mind: Concepts and Emotions," *SEED on Line Journal*, Vol. 2, No. 2, 2002, pp. 36-54.

[14] L. I. Perlovsky, "Statistical Limitations on Molecular Evolution," *Journal of Biomolecular Structure and Dynamics*, Vol. 19, No. 6, 2002, pp. 1031-1043.

[15] L. I. Perlovsky, "Toward Physics of the Mind: Concepts, Emotions, Consciousness, and Symbols," *Physical Life Review*, Vol. 3, No. 1, 2006, pp. 22-55. doi:10.1016/j.plrev.2005.11.003

[16] L. I. Perlovsky, "Fuzzy Dynamic Logic," *New Mathematics and Natural Computation*, Vol. 2, No. 1, 2006, pp. 43-55. doi:10.1142/S1793005706000300

[17] L. I. Perlovsky, "Cognitive High Level Information Fusion," *Information Sciences*, Vol. 177, 2007, pp. 2099-2118. doi:10.1016/j.ins.2006.12.026

[18] S. Grossberg, "Neural Networks and Natural Intelligence," MIT Press, Cambridge, 1988.

[19] M. Bar, K. S. Kassam, A. S. Ghuman, J. Boshyan, A. M. Schmid, *et al*., "Top-Down Facilitation of Visual Recognition," *Proceedings of the National Academy of Sciences*, Vol. 103, No. 2, 2006, pp. 449-454. doi:10.1073/pnas.0507062103

[20] D. L. Schacter and D. R. Addis, "The Ghosts of Past and Future," *Nature*, Vol. 445, No. 7123, 2007, p. 27. doi:10.1038/445027a

[21] L. I. Perlovsky, "Conundrum of Combinatorial Complexity," *IEEE Transactions on PAMI*, Vol. 20, No. 6, 1998, pp. 666-670. doi:10.1109/34.683784

[22] R. E. Bellman, "Adaptive Control Processes," Princeton University Press, Princeton, 1961.

[23] M. L. Minsky, "A Framework for Representing Knowledge," In: P. H. Whinston, Ed., *The Psychology of Computer Vision*, McGraw-Hill Book, New York, 1975.

[24] P. H. Winston, "Artificial Intelligence," Addison-Wesley. Reading, 1984.

[25] N. Chomsky, "Language and Mind. Harcourt Brace Javanovich," Harcourt Brace Jovanovich, Inc., New York, 1972.

[26] R. Nevatia and T. O. Binford, "Description and Recognition of Curved Objects," *Artificial Intelligence*, Vol. 8, No. 1, 1977, pp. 77-98. doi:10.1016/0004-3702(77)90006-6

[27] P. P. Bonnisone, M. Henrion, L. N. Kanal and J. F. Lemmer, "Uncertainty in Artificial Intelligence 6," North Holland, Amsterdam, 1991.

[28] L. I. Perlovsky, "Multiple Sensor Fusion and Neural Networks," DARPA Neural Network Study, 1987.

[29] L. I. Perlovsky, "Cramer-Rao Bounds for the Estimation of Means in a Clustering Problem," *Pattern Recognition Letters*, Vol. 8, No. 1, 1988, pp. 1-3. doi:10.1016/0167-8655(88)90015-3

[30] L. I. Perlovsky, "Cramer-Rao Bounds for the Estimation of Normal Mixtures," *Pattern Recognition Letters*, Vol. 10, No. 3, 1989, pp. 141-148. doi:10.1016/0167-8655(89)90079-2

[31] L. I. Perlovsky, "A Model Based Neural Network for Transient Signal Processing," *Neural Networks*, Vol. 7, No. 3, 1994, pp. 565-572. doi:10.1016/0893-6080(94)90113-9

[32] L. I. Perlovsky, "Cramer-Rao Bound for Tracking in Clutter and Tracking Multiple Objects," *Pattern Recognition Letters*, Vol. 18, No. 3, 1997, pp. 283-288. doi:10.1016/S0167-8655(97)00009-3

[33] N. Chomsky, "Principles and Parameters in Syntactic Theory," In: N. Hornstein and D. Lightfoot, Eds., *Expla-*

nation in Linguistics. The Logical Problem of Language Acquisition, Longman, London, 1981.

[34] R. A. Singer, R. G. Sea and R. B. Housewright, "Derivation and Evaluation of Improved Tracking Filters for Use in Dense Multitarget Environments," *IEEE Transactions on Information Theory*, IT-20, 1974, pp. 423-432. doi:10.1109/TIT.1974.1055256

[35] L. I. Perlovsky and M. M. McManus, "Maximum Likelihood Neural Networks for Sensor Fusion and Adaptive Classification," *Neural Networks*, Vol. 4, No. 1, 1991, pp. 89-102. doi:10.1016/0893-6080(91)90035-4

[36] L. I. Perlovsky, "Beauty and Mathematical Intellect," *Zvezda*, Vol. 9, 2000, pp. 190-201.

[37] L. I. Perlovsky, "Integrating Language and Cognition," *IEEE Connections*, *Feature Article*, Vol. 2, No. 2, 2004, pp. 8-12.

[38] L. I. Perlovsky, "The Mind vs Logic: Aristotle and Zadeh," *Society for Mathematics of Uncertainty*, *Critical Review*, Vol. 1, No. 1, 2007, pp. 30-33.

[39] L. I. Perlovsky, "'Vague-to-Crisp' Neural Mechanism of Perception," *IEEE Transactions on Neural Networks*, Vol. 20, No. 8, 2009, pp. 1363-1367. doi:10.1109/TNN.2009.2025501

[40] L. I. Perlovsky, "Intersections of Mathematical, Cognitive, and Aesthetic Theories of Mind," *Psychology of Aesthetics*, *Creativity*, *and the Arts*, Vol. 4, No. 1, 2010, pp. 11-17. doi:10.1037/a0018147

[41] L. I. Perlovsky, "Neural Mechanisms of the Mind, Aristotle, Zadeh, and fMRI," *IEEE Transactions on Neural Networks*, Vol. 21, No. 5, 2010, pp. 718-733. doi:10.1109/TNN.2010.2041250

[42] L. I. Perlovsky, "Evolution of Languages, Consciousness, and Cultures," *IEEE Computational Intelligence Magazine*, Vol. 2, No. 3, 2007, pp. 25-39. doi:10.1109/MCI.2007.385364

[43] L. I. Perlovsky, "Symbols: Integrated Cognition and Language," In: R. Gudwin and J. Queiroz, Eds., *Semiotics and Intelligent Systems Development*, Idea Group, Hershey, 2007, pp. 121-151.

[44] Aristotle, "The Complete Works," J. Barnes, Ed., Princeton University Press, Princeton, 1995.

[45] M. Davis, "The Universal Computer: The Road from Leibniz to Turing," W. W. Norton and Company, New York, 2000.

[46] D. Hilbert, "The Foundations of Mathematics," In: J. van Heijenoort, Ed., *From Frege to Gödel*, Harvard University Press, Cambridge, 1928/1967, p. 475.

[47] K. Gödel, "Collected Works, Vol. I, 1929-1936," S. Feferman, J. W. Dawson Jr. and S. C. Kleene, Eds., Oxford University Press, New York, 1994.

[48] L. I. Perlovsky, "Gödel Theorem and Semiotics," *Proceedings of the Conference on Intelligent Systems and Semiotics*'96, Vol. 2, Gaithersburg, 1996, pp. 14-18.

[49] L. I. Perlovsky, "Logic versus Mind," Logica Universalis, 2012, in Press.

[50] E. E. Vityaev, L. I. Perlovsky, B. Y. Kovalerchuk and S. O. Speransky, "Probabilistic Dynamic Logic of the Mind and Cognition," 2012, in Press.

[51] L. A. Zadeh, "Fuzzy Sets," *Information and Control*, Vol. 8, No. 3, 1965, pp. 338-352. doi:10.1016/S0019-9958(65)90241-X

[52] V. Kecman, "Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models (Complex Adaptive Systems)," The MIT Press, Cambridge, 2001.

[53] R. Penrose, "Shadows of the Mind," Oxford University Press, Oxford, 1994.

[54] L. I. Perlovsky, "Music—The First Priciple," Musical Theatre, 2006. http://www.ceo.spb.ru/libretto/kon_lan/ogl.shtml

[55] L. I. Perlovsky, "Neural Networks, Fuzzy Models and Dynamic Logic," In: R. Köhler and A. Mehler, Eds., *Aspects of Automatic Text Analysis* (*Festschrift in Honor of Burghard Rieger*), Springer, Heidelberg, 2007, pp. 363-386.

[56] L. I. Perlovsky, "Neural Dynamic Logic of Consciousness: The Knowledge Instinct," In: L. I. Perlovsky and R. Kozma, Eds., *Neurodynamics of Higher-Level Cognition and Consciousness*, Springer Verlag, Heidelberg, 2007. doi:10.1007/978-3-540-73267-9_5

[57] L. I. Perlovsky, "Dynamic Logic," 2010. http://www.scitopics.com/Dynamic_logic.html

[58] L. I. Perlovsky, "Dynamic Logic, Computational Complexity, Engineering and Mathematical Breakthroughs," 2010. http://www.scitopics.com/Dynamic_logic_computational_complexity_engineering_and_mathematical_breakthroughs.html

[59] L. I. Perlovsky, "Dynamic Logic," *Cognitive Breakthroughs*, 2010. http://www.scitopics.com/Dynamic_Logic_Cognitive_Breakthroughs.html

[60] L. I. Perlovsky, "Mind Mechanisms: Concepts, Emotions, Instincts, Imagination, Intuition, Beautiful, Spiritually Sublime," 2010. http://www.scitopics.com/Mind_mechanisms_concepts_emotions_instincts_imagination_intuition_beautiful_spiritually_sublime.html

[61] L. I. Perlovsky and R. Mayorga, "Preface," In: R. Mayorga and L. I. Perlovsky, Eds., *Sapient Systems*, Springer, London, 2008.

[62] L. I. Perlovsky and R. Kozma, "Editorial-Neurodynamics of Cognition and Consciousness," In: L. Perlovsky and R. Kozma, Eds., *Neurodynamics of Cognition and Consciousness*, Springer Verlag, Heidelberg, 2007. doi:10.1007/978-3-540-73267-9_1

[63] H. Harlow, "Mice, Monkeys, Men, and Motives," *Psychological Review*, Vol. 60, No. 1, 1953, pp. 23-32. doi:10.1037/h0056040

[64] L. Festinger, "A Theory of Cognitive Dissonance," Row, Peterson, Evanston, 1957.

[65] S. Grossberg and D. S. Levine, "Neural Dynamics of Attentionally Modulated Pavlovian Conditioning: Blocking, Inter-Stimulus Interval, and Secondary Reinforcement," *Psychobiology*, Vol. 15, No. 3, 1987, pp. 195-240.

[66] L. I. Perlovsky, "Aesthetics and Mathematical Theories of Intellect," *Iskusstvoznanie*, Vol. 2, No. 2, 2002, pp. 558-594.

[67] L. I. Perlovsky, "Sapience, Consciousness, and the Knowledge Instinct (Prolegomena to a Physical Theory)," In: R. Mayorga and L. I. Perlovsky, Eds., *Sapient Systems*, Springer, London, 2008.

[68] R. W. Deming and L. I. Perlovsky, "Concurrent Multi-Target Localization, Data Association, and Navigation for a Swarm of Flying Sensors," *Information Fusion*, Vol. 8, No. 3, 2007, pp. 316-330. doi:10.1016/j.inffus.2005.11.001

[69] L. I. Perlovsky and R. W. Deming, "Neural Networks for Improved Tracking," *IEEE Transactions on Neural Networks*, Vol. 18, No. 6, 2007, pp. 1854-1857. doi:10.1109/TNN.2007.903143

[70] R. W. Deming, J. Schindler and L. I. Perlovsky, "Multi-target/Multisensor Tracking Using Only Range and Doppler Measurements," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 45, No. 2, 2009, pp. 593-611. doi:10.1109/TAES.2009.5089543

[71] R. Mayorga and L. I. Perlovsky, Eds., "Sapient Systems," Springer, London, 2008.

[72] L. I. Perlovsky, R. P. Coons, R. L. Streit, T. E, Luginbuhl and S. Greineder, "Application of MLANS to Signal Classification," *Journal of Underwater Acoustics*, Vol. 44, No. 2, 1994, pp. 783-809.

[73] L. I. Perlovsky, and J. V. Jaskolski, "Maximum Likelihood Adaptive Neural Controller," *Neural Networks*, Vol. 7, No. 4, 1994, pp. 671-680. doi:10.1016/0893-6080(94)90044-2

[74] L. I. Perlovsky, J. A. Chernick and W. H. Schoendorf, "Multi-Sensor ATR and Identification Friend or Foe Using MLANS," *Neural Networks*, Vol. 8, No. 7-8, 1995, pp. 1185-1200. doi:10.1016/0893-6080(95)00078-X

[75] L. I. Perlovsky, W. H. Schoendorf, D. M. Tye and W. Chang, "Concurrent Classification and Tracking Using Maximum Likelihood Adaptive Neural System," *Journal of Underwater Acoustics*, Vol. 45, No. 2, 1995, pp. 399-414.

[76] L. I. Perlovsky, C. P. Plum, P. R. Franchi, E. J. Tichovolsky, D. S. Choi and B. Weijers, "Einsteinian Neural Network for Spectrum Estimation," *Neural Networks*, Vol. 10, No. 9, 1997, pp. 1541-1546. doi:10.1016/S0893-6080(97)00081-6

[77] L. I. Perlovsky, W. H. Schoendorf, L. Garvin and C. Chang, "Development of Concurrent Classification and Tracking for Active Sonar," *Journal of Underwater Acoustics*, Vol. 47, No. 2, 1997, pp. 375-388.

[78] L. I. Perlovsky, W. H. Schoendorf, B. J. Burdick and D. M. Tye, "Model-Based Neural Network for Target Detection in SAR Images," *IEEE Transations on Image Processing*, Vol. 6, No. 1, 1997, pp. 203-216. doi:10.1109/83.552107

[79] L. I. Perlovsky, V. H. Webb, S. R. Bradley and C. A. Hansen, "Improved ROTHR Detection and Tracking Using MLANS," *AGU Radio Science*, Vol. 33, No. 4, 1998, pp. 1034-1044.

[80] R. Kozma, M. Puljic and L. I. Perlovsky, "Modeling Goal-Oriented Decision Making through Cognitive Phase Transitions," *New Mathematics and Natural Computation*, Vol. 5, No. 1, 2009, pp. 143-157. doi:10.1142/S1793005709001246

[81] B. Kovalerchuk, L. I. Perlovsky and G. Wheeler, "Modeling of Phenomena and Dynamic Logic of Phenomena," *Journal of Applied Non-Classical Logics*, Vol. 22, No. 1, 2012, in press.

[82] H. Yardley, L. I. Perlovsky and M. Bar, "Predictions and Incongruency in Object Recognition: A Cognitive Neuroscience Perspective," In: *Detection and Identification of Rare Audiovisual Cues*. *Studies in Computational Intelligence Series*, Springer Publishing, New York, 2012, in Press. doi:10.1007/978-3-642-24034-8_12

[83] L. I. Perlovsky and R. Kozma, Eds., "Neurodynamics of Higher-Level Cognition and Consciousness," Springer-Verlag, Heidelberg, 2007.

[84] L. I. Perlovsky, "Modeling Field Theory of Higher Cognitive Functions," In: A. Loula, R. Gudwin and J. Queiroz, Eds., *Artificial Cognition Systems*, Idea Group, Hershey, 2007, pp. 64-105.

[85] R. Ilin and L. I. Perlovsky, "Cognitively Inspired Neural Network for Recognition of Situations," *International Journal of Natural Computing Research*, Vol. 1, No. 1, 2009, in print.

[86] R. Ilin and L. I. Perlovsky, "Cognitively Inspired Neural Network for Recognition of Situations," *International Journal of Natural Computing Research*, Vol. 1, No. 1, 2010, pp. 36-55. doi:10.4018/jncr.2010010102

[87] S. Pinker, "The Language Instinct: How the Mind Creates Language," William Morrow, New York, 1994.

[88] S. Pinker, "The Stuff of Thought: Language as a Window into Human Nature," Penguin, New York, 2008.

[89] L. I. Perlovsky, "Language and Emotions: Emotional Sapir-Whorf Hypothesis," *Neural Networks*, Vol. 22, No. 5-6, 2009, pp. 518-526. doi:10.1016/j.neunet.2009.06.034

[90] L. I. Perlovsky, "Music and Emotions. Functions, Origins, Evolution," 2010. http://www.scitopics.com/Music_and_Emotions_Functions_Origins_Evolution.html

[91] L. I. Perlovsky, "'High' Cognitive Emotions in Language Prosody," *Physics of Life Reviews*, Vol. 8, No. 4, 2011, pp. 408-409. doi:10.1016/j.plrev.2011.10.007

[92] S. Petrov, F. Fontanari and L. I. Perlovsky, "Categories of Emotion Names in Web Retrieved Texts," *Emotions*, 2012, in Press.

[93] L. I. Perlovsky, "Computational Intelligence Applications for Defense," *IEEE Computational Intelligence Magazine*, Vol. 6, No. 1, 2011, pp. 20-28. doi:10.1109/MCI.2010.939581

[94] L. I. Perlovsky, "Language and Cognition," *Neural Networks*, Vol. 22, No. 3, 2009, pp. 247-257. doi:10.1016/j.neunet.2009.03.007

[95] L. I. Perlovsky, "Language and Cognition Interaction Mechanisms," 2010. http://www.scitopics.com/Language_and_Cognition_Inte

raction_Mechanisms.html

[96] L. I. Perlovsky, "Beauty and Art. Cognitive function and Evolution," 2010. http://www.scitopics.com/Beauty_and_art_Cognitive_function_and_evolution.html

[97] L. I. Perlovsky, "Science and Religion: New Cognitive Findings Bridge the Fundamental Gap," 2010. http://www.scitopics.com/Science_and_Religion_New_cognitive_findings_bridge_the_fundamental_gap.html

[98] L. I. Perlovsky, "Abstract Concepts in Language and Cognition," *Physics of Life Reviews*, Vol. 8, No. 4, 2011, pp. 375-376. doi:10.1016/j.plrev.2011.10.006

[99] L. I. Perlovsky, "Emotions of 'Higher' Cognition," *Brain and Behavior Sciences*, 2012, in press.

[100] L. I. Perlovsky, "Nonlinear Dynamics and Higher Cognitive Mental Functions," *Physics of Life Reviews*, Vol. 9, No. 1, 2012, pp. 74-75. doi:10.1016/j.plrev.2011.12.004

[101] L. I. Perlovsky and R. Ilin, "Neurally and Mathematically Motivated Architecture for Language and Thought, Special Issue 'Brain and Language Architectures: Where We Are Now?'" *The Open Neuroimaging Journal*, Vol. 4, 2010, pp. 70-80. http://www.bentham.org/open/tonij/openaccess2.htm doi:10.2174/1874440001004020070

[102] L. I. Perlovsky and R. Ilin, "CWW, Language, and Thinking, New Mathematics and Natural Computations," 2012, in press.

[103] E. Spelke, "Principles of Object Perception," *Cognitive Science*, Vol. 14, No. 1, 1990, pp. 29-56. doi:10.1207/s15516709cog1401_3

[104] M. D. Hauser, N. Chomsky and W. T. Fitch, "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?" *Science*, Vol. 298, No. 5988, 2002, pp. 1569-1579. doi:10.1126/science.298.5598.1569

[105] L. W. Barsalou, "Abstraction in Perceptual Symbol Systems," *Philosophical Transactions of the Royal Society of London*: *Biological Sciences*, Vol. 358, No. 1435, 2003, pp. 1177-1187. doi:10.1098/rstb.2003.1319

[106] L. W. Barsalou, "Grounded Cognition," *Annual Review Psychology*, Vol. 59, No. 1, 2008, pp. 617-645. doi:10.1146/annurev.psych.59.103006.093639

[107] W. Yeh and L. W. Barsalou, "The Situated Nature of Concepts," *American Journal of Psychology*, Vol. 119, No. 3, 2006, pp. 349-384. doi:10.2307/20445349

[108] L. I. Perlovsky, "Joint Acquisition of Language and Cognition," *Webmed Central Brain*, Vol. 1, No. 10, 2010, Article ID WMC00994. http://www.webmedcentral.com/article_view/994

[109] L. I. Perlovsky, "Language, Emotions, and Cultures: Emotional Sapir-Whorf Hypothesis," *Webmed Central Psychology*, Vol. 2, No. 2, 2011, Article ID WMC001580.

[110] L. I. Perlovsky, "Language and Cognition Interaction Neural Mechanisms," *Computational Intelligence and Neuroscience*, 2011, Article ID 454587. http://www.hindawi.com/journals/cin/contents/

[111] L. I. Perlovsky, "Brain: Conscious and Unconscious Mechanisms of Cognition, Emotions, and Language," *Brain Sciences*, Special Issue, 2012, in press.

[112] L. I. Perlovsky, "The Cognitive Function of Emotions of Spiritually Sublime," *Review of Psychology Frontier*, 2012, in Press.

[113] L. I. Perlovsky, "Emotionality of Languages Affects Evolution of Cultures," *Review of Psychology Frontier*, 2012, in Press.

[114] L. I. Perlovsky, "Languages and Cultures: Emotional Sapir-Whorf Hypothesis (ESWH)," 2010. http://www.scitopics.com/Languages_and_Cultures_Emotional_Sapir_Whorf_Hypothesis_ESWH.html

[115] J. F. Fontanari and L. I. Perlovsky, "Evolving Compositionality in Evolutionary Language Games," *IEEE Transactions on Evolutionary Computations*, Vol. 11, No. 6, 2007, pp. 758-769.

[116] L. I. Perlovsky, "Mirror Neurons, Language, and Embodied Cognition," *Neural networks*, 2012, in press.

[117] J. F. Fontanari and L. I. Perlovsky, "How Language Can Help Discrimination in the Neural Modeling Fields Framework," *Neural Networks*, Vol. 21, No. 2-3, 2008, pp. 250-256. doi:10.1016/j.neunet.2007.12.007

[118] J. F. Fontanari and L. I. Perlovsky, "A Game Theoretical Approach to the Evolution of Structured Communication Codes," *Theory in Biosciences*, Vol. 127, No. 3, 2008, pp. 205-214. doi:10.1007/s12064-008-0024-1

[119] J. F. Fontanari, V. Tikhanoff, A. Cangelosi, R. Ilin and L. I. Perlovsky, "Cross-Situational Learning of Object-Word mapping Using Neural Modeling Fields," *Neural Networks*, Vol. 22, No. 5-6, 2009, pp. 579-585. doi:10.1016/j.neunet.2009.06.010

[120] L. I. Perlovsky, "Jihadism and Grammars. Comment to "Lost in Translation," *Wall Street Journal*, 2010. http://online.wsj.com/community/leonid-perlovsky/activity

[121] L. Boroditsky and M. Ramscar, "The Roles of Body and Mind in Abstract Thought," *Psychological Science*, Vol. 13, No. 2, 2002, pp. 185-188. doi:10.1111/1467-9280.00434

[122] V. Tikhanoff, J. F. Fontanari, A. Cangelosi and L. I. Perlovsky, "Language and Cognition Integration through Modeling Field Theory: Category Formation for Symbol Grounding," *Book Series in Computer Science*, Vol. 4131, Springer, Heidelberg, 2006.

[123] T. W. Deacon, "The Symbolic Species: The Co-Evolution of Language and the Brain," Norton, New York, 1998.

[124] F. De. Saussure, "Course in General Linguistics," McGraw-Hill, New York, 1916-1965, p. 98.

[125] C. S. Peirce, "Logic as Semiotic: The Theory of Signs," In: Buchler, Ed., *The Philosophical Writing of Peirce*, Dover, 1897, 1903; In: Hartshorne and Weiss, Eds., *Collected Papers V. II*, *Elements of Logic*, Belknap, Cambridge, 1955.

[126] C. G. Jung, "Psychological Types," In: *Collected Works*, Vol. 6, Bollingen Series XX, 1971, Princeton University Press, Princeton, 1921.

[127] B. Russell, "Introduction to Mathematical Philosophy," George Allen and Unwin, London, 1919, p. 175.

[128] R. Carnap, "The Logical Syntax of Language," Littlefield,

Adams and Co, Totowa, 1959.

[129] D. S. Levine and L. I. Perlovsky, "Neuroscientific Insights on Biblical Myths: Simplifying Heuristics versus Careful Thinking: Scientific Analysis of Millennial Spiritual Issues," *Zygon*, *Journal of Science and Religion*, Vol. 43, No. 4, 2008, pp. 797-821.

[130] L. I. Perlovsky, "Musical Emotions: Functions, Origin, Evolution," *Physics of Life Reviews*, Vol. 7, No. 1, 2010, pp. 2-27. doi:10.1016/j.plrev.2009.11.001

[131] L. I. Perlovsky, "The Mind Is Not a Kludge," *Skeptic*, Vol. 15, No. 3, 2010, pp. 51-55.

[132] L. I. Perlovsky, "Science and Religion: Scientific Understanding of Emotions of Religiously Sublime," 2010, arXiv: 1010.3059v1.

[133] L. I. Perlovsky, "Music. Cognitive Function, Origin, and Evolution of Musical Emotions," *Webmed Central Psychology*, Vol. 2, No. 2, 2011, Article ID WMC001494.

[134] L. I. Perlovsky, "Consciousness and Free Will, a Scientific Possibility Due to Advances in Cognitive Science," *Webmed Central Psychology*, Vol. 2, No. 2, 2011, Article ID WMC001539

[135] L. I. Perlovsky, "Free Will and Advances in Cognitive Science," *Open Journal of Philosophy*, Vol. 2, No. 1, 2012, pp. 32-37. doi:10.4236/ojpp.2012.21005

[136] L. I. Perlovsky, "Cognitive Function, Origin, and Evolution of Musical Emotions," *Musicae Scientiae*, 2012, in Press.

[137] L. I. Perlovsky, "Cognitive Function of Music," *Interdisciplinary Science Reviews*, 2012, in Press.

[138] L. I. Perlovsky, "Cognitive Function of Musical Emotions," *Psychomusicology*, 2012, in press.

[139] N. Masataka and L. I. Perlovsky, "Music Can Reduce Cognitive Dissonance," 2012, in Press.

[140] L. I. Perlovsky, M.-C. Bonniot-Cabanac and M. Cabanac, "Curiosity and Pleasure," 2010, arXive: 1010.3009.

[141] M. Bar, "The Proactive Brain: Using Analogies and Associations to Generate Predictions," *Trends in Cognitive Sciences*, Vol. 11, No. 7, 2007, pp. 281-289. doi:10.1016/j.tics.2007.05.005

[142] M. Bar, "Predictions: A Universal Principle in the Operation of the Human Brain (Introduction)," *Philosophical Transactions of the Royal Society B*, Vol. 364, No. 1521, 2009, pp. 1181-1182. doi:10.1098/rstb.2008.0321

[143] L. Wu and L. W. Barsalou, "Perceptual Simulation in Conceptual Combination: Evidence from Property Generation," *Acta Psychologica*, Vol. 132, No. 2, 2009, pp. 173-189. doi:10.1016/j.actpsy.2009.02.002

[144] S. Van Danzig, D. Pecher, R. Zeelenberg and L. W. Barsalou, "Perceptual Processing Affects Conceptual Processing," *Cognitive Science*, Vol. 32, 2008, pp. 579-590. doi:10.1080/03640210802035365

[145] H. Brighton, K. Smith and S. Kirby, "Language as an Evolutionary System," *Physics of Life Reviews*, Vol. 2, No. 3, 2005, pp. 177-226. doi:10.1016/j.plrev.2005.06.001

[146] L. I. Perlovsky, "Evolving Agents: Communication and Cognition," In: V. Gorodetsky, J. Liu and V. A. Skormin, Eds., *Autonomous Intelligent Systems*: *Agents and Data Mining*, Springer-Verlag GmbH, Berlin, 2005, pp. 37-49. doi:10.1007/11492870_4

[147] L. I. Perlovsky, "Jihad and Arabic Language. How Cognitive Science Can Help Us Understand the War on Terror," 2010. http://www.scitopics.com/Jihad_and_Arabic_Language_How_Cognitive_Science_Can_Help_Us_Understand_the_War_on_Terror.html

[148] L. I. Perlovsky, "New 'Arrow of Time': Evolution Increases Knowledge," 2010. http://www.scitopics.com/New_Arrow_of_Time_Evolution_increases_knowledge.html

[149] L. I. Perlovsky and A. Goldwag, "The Grammatical Roots of Jihadism: How Cognitive Science Can Help Us Understand the War on Terror," *Webmed Central Psychology*, Vol. 2, No. 2, 2011, Article ID WMC001581.

[150] L. F. Barrett and M. Bar, "See It with Feeling: Affective Predictions in the Human Brain," *Royal Society Philosophical B*, Vol. 364, No. 1513, 2009, pp. 1325-1334.

[151] L. I. Perlovsky, "Music and Consciousness," *Journal of Arts, Sciences and Technology*, Vol. 41, No. 4, 2008, pp. 420-421.

[152] J. F. Fontanari, M.-C. Bonniot-Cabanac, M. Cabanac and L. I. Perlovsky, "A Structural Model of Emotions of Cognitive Dissonances," *Neural Networks*, 2012, in Press. doi:10.1016/j.neunet.2012.04.007

