

‘Vague-to-Crisp’ Neural Mechanism of Perception

Leonid I. Perlovsky

Abstract—The paper describes Neural Modeling Fields (NMF) for object perception, a bio-inspired paradigm. I discuss previous difficulties in object perception algorithms encountered since the 1950s, and describe how NMF overcomes these difficulties. NMF mechanisms are compared to recent experimental neuro-imaging observations, which have demonstrated that initial top-down signals are vague and during perception they evolve into crisp representations matching the bottom-up signals from observed objects. Neural and mathematical mechanisms are described and future research directions outlined.

I. NEURAL NETWORKS AND MECHANISMS OF THE MIND

The field of neural networks has achieved significant success in engineering applications [1,2,3,4] and modeling mechanisms of the brain-mind [5,6,7,8,9,10]. Still, most neural paradigms have not addressed a recently discovered property of the visual perception mechanisms, a vague-fuzzy nature of internal representations [11], which gives rise to top-down priming signals. In this paper we argue that this property is essential for understanding the workings of the mind at lower levels such as object perception, as well as at higher levels associated with abstract concepts, higher emotions including the beautiful and sublime, and their roles in cognition, imagination, and intuition. At lower levels a process “from vague-to-crisp” is essential for fast operation of perception. Mathematically, it reduces the complexity of computation from (often) combinatorial to linear [1,12,13]. At higher levels, in addition to reducing complexity, it is essential for understanding mechanisms that were not previously understood and seemed mysterious [13,14, 15].

The paper also touches on a role of logic in the mind mechanisms. For the first time we describe how logical states emerge from vague-fuzzy states in the continuous process from vague-to-crisp. Whereas fuzzy logic [16] is usually perceived in opposition to Aristotelian logic [17], we note that Aristotle did not consider logic a fundamental mechanism of the mind. His views on the mind operations incorporated vague-fuzzy states of the mind and were more close to the process “from vague-to-crisp” considered here [18].

The next three sections summarize neural modeling fields (NMF) and dynamic logic (DL) forming the mathematical foundation for the paper. Section II summarizes difficulties of the past algorithms, which NMF-DL overcomes; these difficulties are related to complexity and logic. Section III describes the neural architecture of NMF-DL and its operations. Section IV presents an example

illustrating application of NMF-DL to object detection in clutter, which significantly exceeds the performance of previously known algorithms and neural networks. Section V relates NMF-DL operations to mechanisms of the mind. Sections VI and VII discuss relations among DL, fuzzy logic, classical logic, AI, and views of some of the founders of these fields, including Aristotle, Godel, and Zadeh. Experimental validation of DL in neuroimaging experiments is discussed in Section VIII. Section IX discusses further directions.

II. PAST DIFFICULTIES, COMPLEXITY AND LOGIC

Biological object perception involves signals from sensory organs and the internal mind’s representations (memories) of objects. During perception, the mind associates subsets of signals corresponding to objects with representations of object [5,6,11,19], the so-called matching of bottom-up and top-down signals. This matching produces object recognition.

Mathematical descriptions of the very first *recognition* step in this seemingly simple association-recognition-understanding process have met a number of difficulties during the past fifty years. These difficulties were summarized under the notion of combinatorial complexity (CC) [12]. CC refers to multiple 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 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.

Algorithmic complexity was first identified in pattern recognition and classification research in the 1960s and was named “the curse of dimensionality” [20]. It seemed that adaptive self-learning algorithms and neural networks [see 21] could learn solutions to any problem ‘on their own’, if provided with a sufficient number of training examples. The following thirty years of developing adaptive statistical pattern recognition [22] and neural network algorithms led to the conclusion that the required number of training examples often was combinatorially large. Not only individual objects had to be presented for training, but also combinations of objects. Thus, self-learning approaches encountered *CC of learning requirements* [1,12,13].

Rule-based systems were proposed in the 1970’s to solve the problem of learning complexity [23, 24]. An initial idea was that rules would capture the required knowledge and eliminate the need for learning. However in the presence of variability, the number of rules grew; rules became contingent on other rules; combinations of rules had to be considered; thus rule systems encountered *CC of rules*.

Model-based systems were proposed in the 1980s [25,26,27,28]. They used models, which depended on adaptive parameters. The idea was to combine the advantages of rules with learning-adaptivity by using

Harvard University, Cambridge, MA; and the US Air Force Research Laboratory, Sensors Directorate, Hanscom, MA; leonid@seas.harvard.edu

adaptive models. The knowledge was encapsulated in models, whereas unknown aspects of particular situations were to be learned by fitting model parameters. Fitting models to data required selecting data subsets corresponding to various models. The number of subsets, however, is combinatorially large. A popular algorithm for fitting models to the data, multiple hypotheses testing [28], is known to face CC of computations. Model-based approaches encountered *computational CC* (N and NP complete algorithms).

Later research related CC to the type of logic, underlying various algorithms and neural networks. CC of algorithms based on logic is related to Gödel theory: it is a manifestation of Gödel theory in finite systems [1,29]. Various manifestations of CC are all related to formal logic and Gödel theory. Rule systems rely on formal logic in a most direct way. 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 therefore turned out to be related to the fundamental inconsistency of logic [30].

Neural mechanisms of visual perception were described in detail by several authors (see [5,6,19,31,32,33,34] and references therein). A fundamental aspect of perception mechanisms, as mentioned, is matching bottom-up and top-down signals. Several processing layers have been identified, each responsible for specific mechanisms (such as detection of points, edges, elimination of lighting effects, binocular vision, etc.) However, one fundamental aspect of matching bottom-up and top-down signals was missing in these publications. According to Neural Modeling Fields (NMF) [1,27], initial top-down projections onto visual cortex are vague; during perception they become crisp, matching the bottom-up signals from observed objects. This aspect of the matching mechanism is fundamental, it eliminates CC: matching vague representations is fast, it does not require combinations, and it explains several properties of the mind that were not previously understood. This mechanism was confirmed in recent neuro-imaging experiments [11,19].

MF-DL corresponds to these findings and overcomes Gödelian limitations by using a new type of logic.

III. MATHEMATICAL FORMULATION OF NMF-DL

NMF gives a functional description of biological perception. Rather than describing individual neurons and their interconnections, NMF describes functional mechanisms of neural aggregates in visual systems. NMF is a multi-level, hetero-hierarchical system [1]. It mathematically implements several mechanisms of the mind-brain described in this publication. At every level in the hierarchy, top-down signals are produced by representations-memories-models (of objects or concepts) at this level. They interact with bottom-up signals produced at lower levels by models, which were recognized-excited by close match with bottom-up signals at the corresponding (lower) levels. At the lowest level (say retina), signals are coming from the surrounding world. The hierarchy is not perfect, interactions

at each level are affected by several higher and lower levels, hence it is called heterarchy [1]. For shortness we will usually call it the hierarchy. This paper considers in details interaction between two adjacent hierarchical levels of bottom-up and top-down signals (fields of neural activation).

The interaction of bottom-up and top-down signals involves neurons indexed by $n = 1, \dots, N$. Bottom-up signals are denoted as $\mathbf{X}(n)$. $\mathbf{X}(n)$ is a field of bottom-up neuronal synapse activations, coming from neurons at a lower level. Each neuron has a number of synapses; for generality, we describe each neuron activation as a set of numbers, $\mathbf{X}(n) = \{X_d(n), d = 1, \dots, D\}$. Top-down, or priming signals to these neurons are sent by concept-models (also called representations), $\mathbf{M}_m(\mathbf{S}_m, n)$; we enumerate models by index $m = 1, \dots, M$. Each model is characterized by its parameters, \mathbf{S}_m ; in the neuron structure of the brain they are encoded by strength of synaptic connections, mathematically, we describe them as a set of numbers, $\mathbf{S}_m = \{S_{a,m}, a = 1 \dots A\}$. Models represent signals in the following way. Say, signal $\mathbf{X}(n)$, is coming from a sensory neuron n activated by an object m , characterized by parameters \mathbf{S}_m . These parameters may include position, orientation, or size of an object m . Model $\mathbf{M}_m(\mathbf{S}_m, n)$ predicts a value $\mathbf{X}(n)$ of a signal at neuron n . For example, during visual perception, a neuron n in the visual cortex receives a signal $\mathbf{X}(n)$ from retina and a priming signal $\mathbf{M}_m(\mathbf{S}_m, n)$ from an object-concept-model m . A neuron n is activated if both bottom-up signal from lower-level-input and top-down priming signal are strong. As mentioned, the above described interaction of top-down and bottom-up signals has been pioneered by ART [35]; however using parametric models along with matching top-down and bottom-up signals in a neural architecture has been a novel mechanism introduced in NMF-DL [1,27,36,37] along with functional dependence of neural weights on the models. Various models compete for evidence in the bottom-up signals, while adapting their parameters for better match as described below. This is a simplified description of perception. The most benign everyday visual perception uses many levels from retina to object perception. The NMF premise is that the same laws describe the basic interaction dynamics at each level. Perception of minute features, or everyday objects, or cognition of complex abstract concepts is due to the same mechanism described below. Perception and cognition involve models and learning. In perception, models correspond to objects; in cognition models correspond to relationships and situations, similar to [38].

In NMF, bottom-up signals are unstructured data $\{\mathbf{X}(n)\}$ and output signals are recognized or formed concepts $\{m\}$. Top-down, “priming” signals are models, $\mathbf{M}_m(\mathbf{S}_m, n)$, which upon recognition become bottom-up signals for the next, higher level. This architecture is illustrated in Fig. 1.

Learning is essential for perception and cognition. NMF learns driven by the knowledge instinct, an internal “desire” to improve correspondence between top-down and bottom-up signals. Mathematically it increases a similarity measure between the sets of models and signals, $L(\{\mathbf{X}\}, \{\mathbf{M}\})$ [1],

$$L(\{\mathbf{X}\}, \{\mathbf{M}\}) = \prod_{n \in N} l(\mathbf{X}(n)). \quad (1)$$

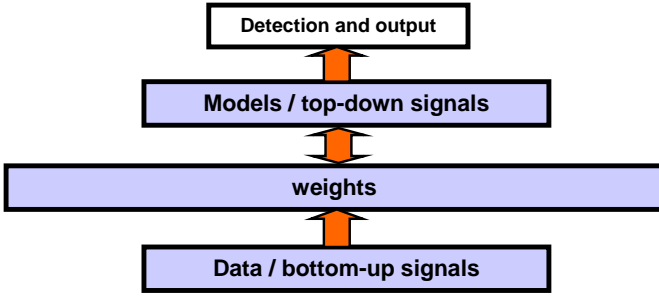


Fig. 1. A single-layer of the NMF architecture.

We now consider two ways to formulate $l(\mathbf{X}(n))$ mathematically, one inspired by Bayes probabilistic approach and another by Shannon information [1,39]. We start with Bayesian formulation as more intuitive. Expression (1) contains a product of partial similarities, $l(\mathbf{X}(n))$. Before perception occurs, the mind does not know which retinal neuron corresponds to which object. Therefore a partial similarity measure is constructed so that it treats each model as an alternative (a sum over models) for each input neuron signal. Its constituent elements are conditional partial similarities between signal $\mathbf{X}(n)$ and model \mathbf{M}_m , $l(\mathbf{X}(n)|m)$, or $l(m|n)$ for short. This measure is “conditional” on object m being present, therefore, when combining these quantities into the overall similarity measure, L , they are multiplied by $r(m)$, which represents the measure of object m actually being present. Combining these elements, a similarity measure is constructed as follows [1]:

$$L(\{\mathbf{X}\},\{\mathbf{M}\}) = \prod_{n \in N} \sum_{m \in M} r(m) l(\mathbf{X}(n) | m). \quad (2)$$

It is convenient to consider its logarithm,

$$LL(\{\mathbf{X}\},\{\mathbf{M}\}) = \sum_{n \in N} \ln \left[\sum_{m \in M} r(m) l(\mathbf{X}(n) | m) \right]. \quad (3)$$

Partial similarities $l(m|n)$ can be taken approximating probability density functions (pdf) or likelihoods of signal $\mathbf{X}(n)$ originating from object m . If models exactly correspond to objects for some values of parameters, the knowledge instinct (at a single NMF level) is equivalent to the maximum likelihood estimation of the model parameters. However, if models are approximate at best, this interpretation does not stand. Therefore it is advantageous to consider similarity based on mutual information. For this purpose we replace (3) by

$$LL = \sum_{n \in N} \text{abs}(\mathbf{X}(n)) \cdot \ln \left[\sum_{m \in M} r(m) l(\mathbf{X}(n)|m) \right]. \quad (4)$$

Eq.(4) gives an approximation of mutual information in models about signals, even if models are not perfect [1].

The learning process consists in estimating model parameters \mathbf{S} and associating signals with concepts by maximizing the similarity (1). Note that all possible combinations of signals and models are accounted for in

expressions (2), (3) or (4). This can be seen by expanding a sum in (2), and multiplying all the terms; it would result in M^N items, a huge number. This is the source of combinatorial complexity (CC) of many algorithms used in the past. For example, multiple hypothesis testing algorithms [28] attempts to maximize similarity L over model parameters and associations between signals and models, in two steps. First it takes one of the M^N items, which is one particular association between signals and models; and maximizes it over model parameters; this is performed over all items. Second, the largest item is selected (that is the best association for the best set of parameters). Such a program inevitably faces a wall of CC, the number of computations on the order of M^N .

NMF overcomes this fundamental difficulty of many learning algorithms and, as mentioned, solves this problem without CC by using dynamic logic, which evolves ‘from vague-to-crisp’ [1,36]; relations of dynamic logic to other types of logic were considered in [40]. An important aspect of dynamic logic is matching vagueness or fuzziness of similarity measures to the uncertainty of models. Initially, parameter values are not known, and uncertainty of models is high; so is the fuzziness of the similarity measures. In the process of learning, models become more accurate and the similarity measure more crisp, the value of the similarity increases. This is the mechanism of dynamic logic.

Mathematically this process is described as follows. First, assign any values to unknown parameters, $\{\mathbf{S}_m\}$. Then, compute association variables $f(m|n)$,

$$f(m|n) = r(m) l(\mathbf{X}(n) | m) / \sum_{m' \in M} r(m') l(\mathbf{X}(n)|m'). \quad (5)^{(2)}$$

Next, iteratively improve the association variables and parameters, according to:

$$df(m|n)/dt = f(m|n) \sum_{m' \in M} \{[\delta_{mm'} - f(m'|n)] \cdot \quad (2)$$

$$[\partial \ln l(n|m') / \partial \mathbf{M}_{m'}] (\partial \mathbf{M}_{m'} / \partial \mathbf{S}_{m'}) \cdot d\mathbf{S}_{m'} / dt,$$

$$d\mathbf{S}_m / dt = \sum_{n \in N} \text{abs}(\mathbf{X}(n)) f(m|n) [\partial \ln l(n|m) / \partial \mathbf{M}_m] \partial \mathbf{M}_m / \partial \mathbf{S}_m,$$

$$\delta_{mm'} = 1 \text{ if } m=m', 0 \text{ otherwise.} \quad (6)$$

Parameter t is the time of the internal dynamics of the NMF system (like a number of internal iterations). Here $\partial(\cdot)/\partial(\cdot)$ stands for partial derivatives and $d(\cdot)/dt$ are temporal derivatives defining the dynamic logic process. Eqs.(6) are the first order differential equations in time (t) and can be solved by a standard solver. Initially, parameter values are wrong and similarity values are vague, every data point has an appreciable similarity to each model. After few iterations parameter values improve, similarity (1) increases, association variables $f(m|n)$ tend to zeroes or ones that is to crisp associations. This dynamic logic process “from vague-to-crisp” sets NMF-DL apart from all other algorithms and neural paradigms: it results in orders of improvements over

previous algorithms as illustrated in the next section example and documented for many dozens of applications in given references and hundreds of references therein. Biologically, increase in similarity satisfies the knowledge instinct, which is felt subjectively as a positive aesthetic emotion [13], NMF-DL enjoys learning.

Convergence of these equations was proven in [1]. Many applications of this theory have been described in [1,7,9,13,14,27] and in other references in this paper; an application to multiple target tracking in strong clutter was presented in [41]. This and other applications resulted in significant savings in complexity, in about two orders of magnitude improvement in signal-to-clutter ratio, and often in the best possible performance. Here we consider an application to image recognition (perception) using similarity measure (3).

IV. EXAMPLE OF NMF OBJECT PERCEPTION

Finding patterns buried in clutter (a large number of signals coming from extraneous sources of no interest) is 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 discussed, is multiple hypotheses testing [28]; since all combinations of subsets and models are exhaustively searched, it faces the problem of combinatorial complexity. In the current example, we are looking for ‘smile’ and ‘frown’ patterns in noise shown in Fig.1a without noise, and in Fig.1b with noise, as actually measured.

The NMF internal knowledge in this case is given by three types of parametric models: a uniform model for noise, Gaussian blobs for vague, poorly resolved patterns, and parabolic models for ‘smiles’ and ‘frowns.’ Each parabolic model is made of a superposition of K Gaussian components.

Noise:

$$r(1) = r_1; l(n|1) = 1;$$

Gaussian blobs:

$$r(m) = r_m; l(n|m, \text{Gaussian}) = G(\mathbf{n}|\mathbf{n}_m, \sigma);$$

Parabolic:

$$r(m)=1; l(n|m, \text{parabolic}) = \sum_{k \in K} r_{km} G(\mathbf{n}|\mathbf{n}_{mk}, \sigma); \quad (13)$$

$$\mathbf{n} = (n_x, n_y); \mathbf{n}_m = (n_{mx}, n_{my});$$

$$\mathbf{n}_{mk} = (n_{mkx}, n_{mky}) = (k-k_{mx}, a_m(k-k_{my})^2).$$

Here, \mathbf{n} is a two-dimensional point in (n_x, n_y) -image plane; \mathbf{n}_m is a two-dimensional center of the m -th Gaussian blob in this image plane; \mathbf{n}_{mk} is a two-dimensional center of the k -th Gaussian component of the m -th parabolic model and a_m determines curvature of the m -th parabolic model. Parabolic models approximate pattern shapes (as in Fig. 2a below) as somes of Gaussians. Parameters of the models are $r_m, r_{km}, \mathbf{n}_m, \mathbf{k}_{mk}, a_m, \sigma$; and also the number of models of each type, except K . Standard deviations σ were kept equal for all models for simplicity, also K was set to 10. (Possibly K depending on iterations, and σ, K depending on m, k would result in a more efficient algorithm; it would be definitely so for more complex problems. Here, the number of parameters per model is 7.)

The (n_x, n_y) image plane in this example is a square of 100x100 points, and the true number of parabolic patterns is 3, which is not known to the algorithm. At the first iteration (the 1st step in solving the differentiation eqs.(6)) two models were initiated, one noise and one Gaussian. On the following iterations an additional model is initiated if its energy $r(m)$ exceeds a predetermined threshold (set at 0.001 of the total signal energy). There is always one Gaussian and one parabolic model available for initiation in addition to already initiated models; this is the procedure for estimating the number of objects present.

Using a multiple hypothesis testing (brute-force) approach will take about $M^N = 10^{6,000}$ operations. Alternatively, fitting $7 \times 4 = 28$ parameters to 100x100 grid by testing all parameter values would take about 10^{47} to 10^{48} operations, a prohibitive computational complexity in both cases (which grows exponentially with the problem complexity). The DL “from vague-to-crisp” process solved

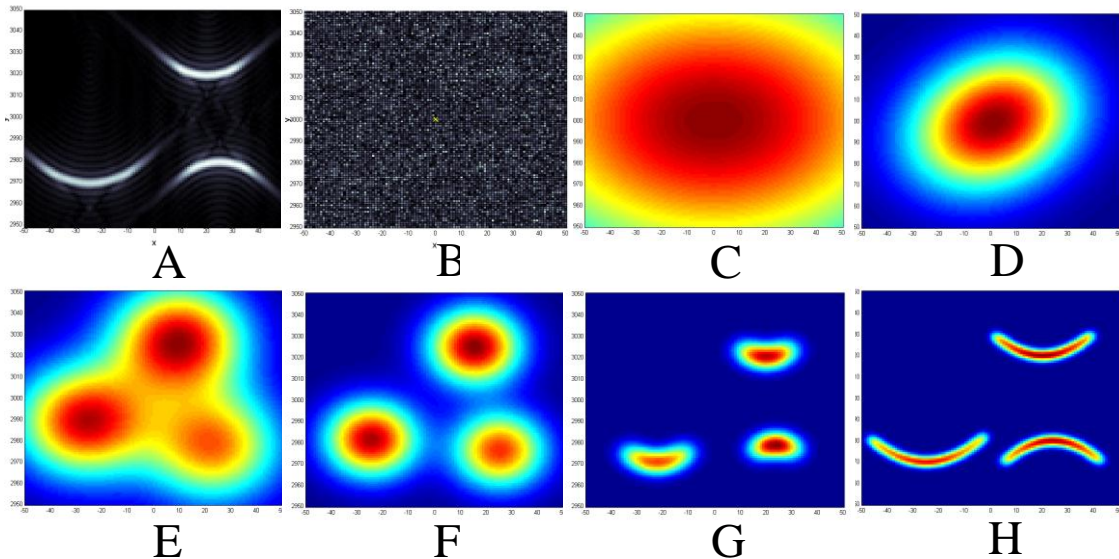


Fig.1. Finding ‘smile’ and ‘frown’ patterns in noise, an example of dynamic logic “from vague-to-crisp” process: (a) true ‘smile’ and ‘frown’ patterns are shown without clutter; (b) actual image available for recognition (signal is below clutter, signal-to-clutter ratio is about 1/3); (c) an initial fuzzy blob-model, the vagueness corresponds to uncertainty of knowledge; (d) through (h) show improved models at various iteration stages (total of 22 iterations). 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. Until 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. Initial models contain low-spatial frequencies compared to the final one. Iterations stopped at (h), when similarity (4) stopped increasing.

this problem with 10^9 operations; and resulted in two orders of magnitude improvement in signal-to-clutter ratio over the best algorithms available previously [42]. We would emphasize that detection of signals below clutter has not been previously possible.

During an adaptation process, initial fuzzy and uncertain models are associated with structures in the input signals, and fuzzy models become more definite and crisp with successive iterations. The type, shape, and number, of models are selected so that the internal representation within the system is similar to input signals: the NMF concept-models represent structure-objects in the signals. There are several types of models: one uniform model describing clutter (it is not shown) and a variable number of blob models and parabolic models, which number, location, and curvatures are estimated from the data.

VI. BIOLOGICAL VALIDATION OF DYNAMIC LOGIC

Biological experimental validation of dynamic logic can be obtained by everyone in just a few seconds. Just close your eyes and imagine a familiar object in front of you just a second ago. Your imagination is vague, not as crisp as perception of the object with open eyes. Interaction of top-down and bottom-up signals involved in this experiment were described in [5,6,43,44]: imagination is produced in the visual cortex by top-down signals from models in our memory. This proves that in the initial stages of perception memories-models are vague, as in dynamic logic.

Recently, detailed neurological and fMRI neuroimaging studies [11,19,45] confirmed that conscious perceptions are preceded by activation of cortex areas, where top-down signals originate from memories-models. Also, initial top-down signals (models) are driven by *vague* contents of bottom-up signals. Thus, dynamic logic processes ‘from vague to crisp’ were confirmed in neuroimaging experiments. This experimentally confirmed aspect of the mathematical modeling in section III is the novel content of this paper. It supports more complicated models of cognition based on NMF-DL briefly discussed below.

VII. FUTURE DIRECTIONS

The next step is to extend NMF and dynamic logic to perception of complex objects. For engineering applications, initial vague models can be obtained from exact models by convolving the model with a low resolution kernel. The kernel uncertainty, like in the demonstrated example, should be wide enough to account for uncertainty in knowledge of model parameters. Parameterization of the model should allow for a reasonably good match with sensor data. If constructing an analytic expression for the image shape is

unfeasible, an alternative could be to construct a weighted combination of possible views, with weights included into the model parameter list. For understanding workings of the brain, neural model “parameterization,” “parameter fitting,” and model acquisition are topics of future research.,

NMF relies on parametric models

Longer-term research should concentrate first on implementing multi-layer NMF, for modeling higher level cognition, including complex situations. Second, it should be combined with language-learning NMF, for higher-level cognition is impossible without language [14,15,46]; these references also discuss how NMF-DL can learn complicated models beyond parametric shapes, similar to “perception symbol systems” [38]. Third, this integrated cognitive-language hierarchical NMF system should be used for acquiring, in interaction with human users, language along with high-level cognition.

In parallel, multi-agent systems, where each agent has an NMF mind, can be developed for several applications.

Further experimental brain-imaging studies should extend results of [11] to high cognitive functions, to cognition of abstract concepts, and to mechanisms of language and cognition interaction.

ACKNOWLEDGMENTS

I am thankful to M. Bar, R. Brockett, R. Deming, D. Levine, R. Kozma, and B. Weijers, for discussions, help and advice, and to AFOSR for supporting part of this research under the Lab. Task 05SN02COR, PM Dr. Jon Sjogren.

REFERENCES

- [1] L. I. Perlovsky. *Neural Networks and Intellect: using model based concepts*. New York: Oxford University Press, 2001.
- [2] C. M. Bishop. *Neural Networks for Pattern Recognition*. New York: Oxford University Press, 1996.
- [3] S. Haykin. *Neural Networks: A Comprehensive Foundation*. New York: Prentice Hall, 1998.
- [4] C. M. Bishop. *Pattern Recognition and Machine Learning*. New York: Springer, 2007.
- [5] S. Grossberg. *Studies of Mind and Brain*. Dordrecht, Holland: D. Reidel Publishing Co., 1982.
- [6] S. M. Kosslyn. *Image and Brain*. Cambridge, MA: MIT Press, 1994
- [7] L. I. Perlovsky and R. Kozma, Eds. *Neurodynamics of Higher-Level Cognition and Consciousness*. Heidelberg, Germany: Springer-Verlag, 2007.
- [8] R. Hecht-Nielsen. *Confabulation Theory: The Mechanism of Thought*. Heidelberg, Germany: Springer-Verlag, 2007.

-
- [9] R. Mayorga and L. I. Perlovsky, Eds. *Sapient Systems*. London, UK: Springer, 2007.
- [10] W. Gnadt and S. Grossberg. "SOVEREIGN: An autonomous neural system for incrementally learning planned action sequences to navigate towards a rewarded goal". *Neural Networks*, 21, pp. 699-758, 2008
- [11] M. Bar, K. S. Kassam, A. S. Ghuman, J. Boshyan, A. M. Schmid, A. M. Dale, M. S. Hamalainen, K. Marinkovic, D. L. Schacter, B. R. Rosen, E. Halgren. "Top-down facilitation of visual recognition." *PNAS*, vol. 103, pp. 449-454, 2006.
- [12] L. I. Perlovsky. "Conundrum of Combinatorial Complexity." *IEEE Trans. PAMI*, (20)6, pp. 666-670, 1998.
- [13] L. I. Perlovsky. "Toward Physics of the Mind: Concepts, Emotions, Consciousness, and Symbols." *Phys. Life Rev.* 3(1), pp. 22-55, 2006.
- [14] L. I. Perlovsky. "Integrating Language and Cognition." *IEEE Connections*, 2(2), pp. 8-12, 2004.
- [15] L. I. Perlovsky, "Modeling Field Theory of Higher Cognitive Functions." In *Artificial Cognition Systems*, Eds. A. Loula, R. Gudwin and J. Queiroz. Hershey, PA: Idea Group, 2006, pp. 64-105.
- [16] L. A. Zadeh. "Fuzzy sets." *Information and Control*, 8, pp. 338-353, 1965.
- [17] Aristotle. *The Complete Works of Aristotle*, Ed. J. Barnes, Princeton, NJ, IV BCE / 1995.
- [18] Aristotle. *Metaphysics*. In *The Complete Works of Aristotle*, Ed. J. Barnes, Princeton, NJ, IV BCE / 1995.
- [19] D. L. Schacter, I. G. Dobbins, D. M. Schnyer. "Specificity of priming: A cognitive neuroscience perspective." *Nature Reviews Neuroscience*, vol. 5, pp. 853-862, 2004.
- [20] R. E. Bellman. *Adaptive Control Processes*. Princeton, NJ: Princeton University Press, 1961.
- [21] Wikipedia, "Backpropagation Neural Networks." <http://en.wikipedia.org/wiki/Backpropagation>
- [22] N. J. Nilsson. *Learning Machines*. New York: McGraw-Hill, 1965.
- [23] M. L. Minsky. "A Framework for Representing Knowledge." In *The Psychology of Computer Vision*, ed. P. H. Winston, New York: McGraw-Hill Book, 1975.
- [24] P. H. Winston. *Artificial Intelligence*. Reading, MA: Addison-Wesley, 1984.
- [25] R. Nevatia and T. N. Binford. "Description and recognition of curved objects." *Artificial Intelligence* 8(1), pp. 77-98, 1977.
- [26] R. A. Brooks. "Model-based interpretation of three-dimensional images." *IEEE Trans. Pattern Analysis and Machine Intelligence*, 5(2), pp. 140-150, 1983
- [27] L. I. Perlovsky and M. M. McManus. "Maximum Likelihood Neural Networks for Sensor Fusion and Adaptive Classification." *Neural Networks*, 4(1), pp. 89-102, 1991.
- [28] 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*, vol. IT-20, pp. 423-432, 1974.
- [29] L. I. Perlovsky. "Gödel Theorem and Semiotics." In *Proc. Conference on Intelligent Systems and Semiotics '96*. Gaithersburg, MD, vol.2, pp. 14-18, 1996.
- [30] Gödel, K. (1986). *Kurt Gödel collected works, I*. (Ed. S. Feferman at al). Oxford University Press.
- [31] H. Neumann, A. Yazdanbakhsh and E. Mingolla. "Seeing surfaces: The brain's vision of the world," *Physics of Life Reviews*, vol. 4, no. 3, pp. 189-222, 2007
- [32] R. P. N. Rao and D. H. Ballard, "Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects," *Nature Neuroscience* 2, 1, 79-87, 1999.
- [33] T. S. Lee and D. Mumford. "Hierarchical Bayesian Inference in the Visual Cortex". *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, 20, pp. 1434-1448, 2003.
- [34] S. Grossberg and M. Versace. "Spikes, synchrony, and attentive learning by laminar thalamocortical circuits." *Brain Research*, 1218, pp. 278-312, 2008.
- [35] G.A. Carpenter and S. Grossberg. "A massively parallel architecture for a self-organizing neural pattern recognition machine." *Computer Vision, Graphics, and Image Processing*, 37, pp. 54-115, 1987.
- [36] L. I. Perlovsky and M. M. McManus. "Maximum Likelihood Neural Network for Adaptive Classification." *International Joint Conference on Neural Networks*, Washington, DC, 1989.
- [37] L. I. Perlovsky. *Multiple Sensor Fusion and Neural Networks*. DARPA Neural Network Study, MIT/Lincoln Laboratory, Lexington, MA, 1987.
- [38] L. W. Barsalou. "Perceptual Symbol Systems," *Behavioral and Brain Sciences*, 22, pp. 577-660, 1999.
- [39] C. E. Shannon and W. Weaver. "A mathematical theory of communications." *Bell System Technical Journal*, vol. 27, pp. 370-423, 1948.
- [40] B. Kovalerchuk and L. I. Perlovsky. "Dynamic Logic of Phenomena and Cognition." In *proc. The World Congress on Computational Intelligence, WCCI'08, Hong Kong, China*, 2008.
- [41] L.I. Perlovsky and R. W. Deming. "Neural Networks for Improved Tracking." *IEEE Trans. Neural Networks*, vol. 18, no. 6, pp. 1854-1857, 2007.
- [42] R. Linnehan, C. Mutz, L.I. Perlovsky, B. Weijers, J. Schindler, R. Brockett. "Detection of Patterns Below Clutter in Images." *Int. Conf. Integration of Knowledge Intensive Multi-Agent Systems*, Cambridge, MA, 2003.
- [43] S. Grossberg. "How does a brain build a cognitive code?" *Psychol. Rev.* vol. 87, pp. 1-51, 1980.
- [44] D. L. Schacter. "Implicit memory: History and current status." *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 13, pp. 501-518, 1987
- [45] D. L. Schacter and D. R. Addis. "The ghosts of past and future." *Nature*, vol. 445, p. 27, 2007.
- [46] L. I. Perlovsky. "Symbols: Integrated Cognition and Language." In *Semiotics and Intelligent Systems Development*. Eds. R. Gudwin, J. Queiroz. Hershey, PA: Idea Group, pp.121-151. 2006.