Neural Networks, Fuzzy Models and Dynamic Logic

Leonid I. Perlovsky

Air Force Research Lab
Leonid.Perlovsky@hanscom.af.mil

Summary. The paper discusses possible relationships between computational intelligence, known mechanisms of the mind, semiotics, and computational linguistics. Mathematical mechanisms of concepts, emotions, and goals are described as a part of information processing in the mind and are related to language and thought processes in which an event (signals from surrounding world, text corpus, or inside the mind) is understood as a concept. Previous attempts in artificial intelligence at describing thought processes are briefly reviewed and their fundamental (mathematical) limitations are analyzed. The role of emotional signals in overcoming these past limitations is emphasized. The paper describes mathematical mechanisms of concepts applicable to sensory signals and linguistics; they are based on measures of similarities between models and signals. Linguistic similarities are discussed that can utilize various structures and rules proposed in computational linguistic literature. A hierarchical structure of the proposed method is capable of learning and recognizing concepts from textual data, from the level of words and up to sentences, groups of sentences, and towards large bodies of text. I briefly discuss a role of concepts as a mechanism unifying thinking and language and their possible role in language acquisition. A thought process is related to semiotic notions of signs and symbols. It is further related to understanding, imagination, intuition, and other processes in the mind. The paper briefly discusses relationships between the mind and brain and applications to understanding-based search engines.

1 Language and the Mind

Language and thinking are distinctly human abilities. Even if one prefers to consider the difference between human and animal minds in terms of degrees, the difference is formidable. Close relationships between language and thinking encouraged equating these abilities in the past. Rule-based systems, using the mathematics of logic, implied significant similarities between the two. The situation has changed, in part due to the fact that logic-rule systems have not been sufficiently powerful to explain thinking, nor language abilities, and in part due to improved scientific understanding (psychological, cognitive, neural, linguistic) of the mechanisms involved. Among contemporary linguists there is a growing appreciation of a possibility that language and thinking could be distinct and different abilities of mind (see [43] for further references). Comparing apes with parrots, the first having significant intellectual capabilities and the second having significant linguistic capabilities, one may conclude that language and thinking might have evolved along separate evolutionary paths; and some researchers believe there are reasons for this conclusion.

Human language mechanisms include abilities to acquire a large vocabulary, rules of grammar, and to use the finite set of words and rules to generate virtually infinite number of phrases and sentences [16, 44]. Human thinking includes abilities to understand the surrounding world in terms of objects, their relationships (scenes and situations), relationships among relationships, and so on [40]. Researchers in computational linguistics, mathematics of intelligence and neural networks, cognitive science, neuro-physiology and psychology during the last twenty years significantly advanced understanding of the mechanisms of the mind involved in learning and using language, mechanisms of perception and cognition – for the discussions and further references see [10, 16, 25, 28, 40, 44, 46] Much less advance was achieved toward deciphering mechanisms relating linguistic competence to understanding and thinking. Although it seems clear that language and thinking are closely related abilities, intertwined in evolution, ontogenesis, and everyday use, still the currently understood mechanisms of language are mainly limited to relations of words to other words and phrases, but not to the objects in the surrounding world, not to cognition and thinking. Possible mathematical approaches toward integrating language and thinking, words and objects, phrases and situations are discussed in this paper.

The paper starts with a mathematical description of thinking, which still is an issue of much controversy. Among researchers in mathematical intelligence it has become appreciated, especially during the last decades that thinking is not just a chain of logical inferences [10, 28, 40]. Yet, mathematical methods describing thinking as processes involving concepts, instincts, emotions, memory, imagination are not well known, although significant progress in this direction was achieved [10, 28, 40]. A brief historical overview of this area including difficulties and controversies is given in the next two sections from mathematical, psychological and neural standpoints; it is followed by a mathematical description of thinking processes. Then the paper discusses the ways in which the mathematical description of thinking can be combined with language, taking advantage of recent progress in computational linguistics. It touches upon novel ideas of computational semiotics relating language and thinking through signs and symbols. In conclusion, I briefly discuss relationships between mathematical, psychological, and neural descriptions of thinking processes and language as parts of the mind.

Words like *mind, thought, imagination, emotion, concept* are often used colloquially in many ways, but their use in science and especially in mathematics of intelligence has not been uniquely defined and is a subject of active research and ongoing debates [10, 28]. According to a dictionary [42], mind includes conscious and unconscious processes, especially thought, perception, emotion, will, memory, and imagination, and it originates in brain. These constituent notions will be discussed throughout the paper.

A broad range of opinions exists on the mathematical methods suitable for the description of the mind. Founders of artificial intelligence thought that formal logic was sufficient [31] and no specific mathematical techniques would be needed to describe the mind [29]. An opposite point of view is that there are few specific mathematical constructs, "the first principles" of the mind organization. Among researchers taking this view is Grossberg, who suggested that the first principles include a resonant matching between lower-level signals [10] and higher-level representations and emotional evaluation of conceptual contents [12]; several researchers suggested specific principles of the mind organization [18, 27, 40, 51]. Hameroff, Penrose, and the author (among others) considered quantum computational processes that might take place in the brain [14, 33, 36]. Although, it was suggested that new unknown yet physical phenomena will have to be accounted for explaining the working of the mind [33]. This paper describes mechanisms of the mind that can be "implemented" by classical physics mechanisms of the brain neural networks and, alternatively, by using existing computers.

2 Theories of the Mind and Combinatorial Complexity

Understanding signals coming from sensory organs involves associating subsets of signals corresponding to particular objects with internal representations of these objects. This leads to recognition of the objects and activates internal brain signals leading to mental and behavioral responses, which constitute the understanding of the meaning (of the objects).

Developing mathematical descriptions of the very first *recognition* step of this seemingly simple association-recognition-understanding process has not been easy, a number of difficulties have been encountered during the past fifty years. These difficulties have been summarized under the notion of combinatorial complexity (CC) [37]. The problem was first identified in pattern recognition and classification problems in the 1960s and was named "the curse of dimensionality" [2]. The following thirty years of developing adaptive statistical pattern recognition and neural network algorithms designed for self-learning led to a conclusion that these approaches often encountered *CC of learning requirements*: recognition of any object, it seemed, could be learned if "enough" training examples were used for an algorithm self-learning. The required examples had to account for all possible variations of "an object", in all possible geometric positions and in *combinations* with other objects, sources of light, etc., leading to astronomical (and worse) numbers of required examples.

By the end of the 1960s a different paradigm became popular: logic-rule-based systems (or expert systems) were proposed to solve the problem of learning complexity. An initial idea was that rules would capture the required knowledge and eliminate a need for learning. The first Chomskian ideas concerning mechanisms of language grammar related to deep structure [4] were also based on a similar idea of logical rules. Rule systems work well when all aspects of the problem can be predetermined. However, rule systems and expert systems in the presence of unexpected variability, encountered *CC of rules*: more and more detailed sub-rules and sub-sub-rules, one contingent on another, had to be specified.

In the 1980s model-based systems became popular, which were proposed to combine advantages of adaptivity and rules by utilizing adaptive models. Existing knowledge was to be encapsulated in models and unknown aspects of concrete situations were to be described by adaptive parameters. Along similar lines were *rules and parameters* ideas of Chomsky [5]. Model-based systems encountered *computational CC* (N and NP complete algorithms). The reason was that considered algorithms had to evaluate multiple combinations of elements of data and rules (models). CC is prohibitive because the number of combinations is very large: for example, consider 100 elements (not too large a number) whose combinations had to be evaluated; the number of combinations of 100 elements is 100^{100}, a number comparable to the number of elementary particles in a Universe; no computer would ever be able to compute that many combinations. The CC became a ubiquitous feature of intelligent algorithms and seemingly, a fundamental mathematical limitation.

Combinatorial complexity has been related to the type of logic, underlying various algorithms and neural networks [37]. Formal logic is based on the "law of excluded third", according to which every statement is either true or false and nothing in between. Therefore, algorithms based on formal logic have to evaluate every little variation in data or internal representations as a separate logical statement; a large number of combinations of these variations cause combinatorial complexity. In fact, combinatorial complexity of algorithms based on logic has been related to the Gödel theory: it is a finite system manifestation of the incompleteness of logic [34]. Multivalued logic and fuzzy logic were proposed to overcome limitations related to the law of excluded third [17]. Yet the mathematics of multivalued logic is no different in principle from formal logic. Fuzzy logic encountered a difficulty related to the degree of fuzziness: if too much fuzziness is specified, the solution does not achieve a needed accuracy, if too little, it becomes similar to formal logic.

3 Mind: Concepts and Emotions

Seemingly fundamental nature of mathematical difficulties discussed above led many to believe that classical physics cannot explain the working of the mind. Yet, I would like to emphasize another aspect of the problem: often mathematical theories of the mind where proposed before the necessary physical intuition of how the mind works was developed. Newton, as often mentioned, did not consider himself as evaluating various hypotheses about the working of the material world, he felt that he possesses what we call today a physical intuition about the world [50]. An intuition about the mind points to mechanisms of concepts, emotions, instincts, imagination, behavior generation, consciousness and unconscious. An essential role of emotions in the working of the mind was analyzed from the psychological and neural perspective by Grossberg [13], from the neuro-physiological perspective by Damasio [6], and from the learning and control perspective by the author [8, 38, 39]. One reason for engineering community being slow in adopting these results is the cultural bias against emotions as a part of thinking processes. Plato and Aristotle thought that emotions are "bad" for intelligence, this is a part of our cultural heritage ("one has to be cool to be smart"), and the founders of Artificial Intelligence repeated this truism about emotions [31]. Yet, as discussed in the next section, combining conceptual understanding with emotional evaluations is crucial for overcoming the combinatorial complexity as well as the related difficulties of logic.

Let me summarize briefly and in a much simplified way several aspects of the working of the mind, which seem essential to the development of the mathematical descriptions of the mind mechanisms: instincts, concepts, emotions, behavior generation. The mind has evolved for the purpose of survival and therefore it serves for a better satisfaction of the basic instincts, which have emerged as survival mechanisms even before the mind. Instincts operate like internal sensors: for example, when a sugar level in our blood goes below a certain level an instinct "tells us" to eat. The most accessible to our consciousness mechanism of the mind is concepts: the mind operates with concepts. Concepts are like internal models of the objects and situations; this analogy is quite literal, e.g., during visual perception of an object, an internal concept model projects an image onto the visual cortex, which is matched there to an image projected from retina (this simplified description will be refined later).

An ability for concepts evolved for instinct satisfaction, and the mechanism linking concepts and instincts involves emotions. Emotions are neural signals connecting instinctual and conceptual brain regions. Whereas in colloquial usage, emotions are often understood as facial expressions, higher voice pitch, exaggerated gesticulation, these are the outward signs of emotions, serving for communication. A more fundamental role of emotions within the mind system is that emotional signals evaluate concepts for the purpose of instinct satisfaction. This evaluation is not according to rules or concepts (like in rule systems of artificial intelligence), but according to a different instinctual-emotional mechanism described in the next section. This emotional mechanism is crucial for breaking out of the "vicious circle" of combinatorial complexity.

The results of conceptual-emotional understanding of the world are actions (or behavior) in the outside world or within the mind. In this paper we touch on only one type of behavior, the behavior of improving understanding and knowledge of the language and world (including self). In the next section we describe a mathematical theory of a "simple" conceptual-emotional recognition and understanding. As we will discuss, in addition to concepts and emotions, it involves with necessity mechanisms of intuition, imagination, conscious, unconscious, and aesthetic emotion. And this process is intimately connected to an ability of mind to form symbols and interpret signs.

The mind involves a hierarchy of multiple levels of concept-models, from simple perceptual elements (like edges, or moving dots), to concept-models of objects, to complex scenes, and up the hierarchy ... toward the concept-models of the meaning of life and purpose of our existence. Hence the tremendous complexity of the mind, yet relatively few basic principles of the mind organization go a long way explaining this system.

4 Modeling Field Theory (MFT)

Modeling field theory [40], summarized below, associates lower-level signals with higher-level concept-models (or internal representations), resulting in understanding of signals, while overcoming the difficulties of CC described in Section 2. It is achieved by using measures of similarity between the concept models and the input signals combined with a new type of logic, i.e. the fuzzy dynamic logic. Modeling field theory is a multi-level, hetero-hierarchical system. This section describes a basic mechanism of interaction between two adjacent hierarchical levels of signals (fields of neural activation); sometimes, it will be more convenient to talk about these two signal levels as an input to and output from a (single) processing level.

At each level, the output signals are concepts recognized (or formed) in input signals. Input signals $\mathbf{X}$ are associated with (or recognized, or grouped into) concepts according to the representation models and similarity measures at this level. In the process of association-recognition, models are adapted for better representation of the input signals; and similarity measures are adapted so that their fuzziness is matched to the model uncertainty. The initial uncertainty of models is high and so is the fuzziness of the similarity measure; in the process of learning models become more accurate and the similarity more crisp, the value of the similarity measure increases. I call this mechanism *fuzzy dynamic logic*.

4.1 Internal Models, Learning, and Similarity

During the learning process, new associations of input signals are formed resulting in evolution of new concepts. Input signal $\{\mathbf{X}(n)\}$, $n = 1, \ldots, N$, a field of input neuronal synapse activation levels, enumerates the input neurons. $\mathbf{X}(n)$ are the activation levels. A set of concept-models $h = 1, \ldots, H$ is characterized by the models (or representations) $\{\mathbf{M}_h(n)\}$ of the signals $\mathbf{X}(n)$.

Each model depends on its parameters $\{\mathbf{S}_h\}, \mathbf{M}_h(\mathbf{S}_h, n)$. In a highly simplified description of a visual cortex, n enumerates the visual cortex neurons, $\mathbf{X}(n)$ are the "bottom-up" activation levels of these neurons coming from the retina through visual nerve, and $\mathbf{M}_h(n)$ are the "top-down" activation levels (or priming) of the visual cortex neurons from previously learned object-models[1]. Learning process attempts to "match" these top-down and bottom-up activations by selecting "best" models and their parameters. Mathematically, learning increases a similarity measure between the sets of models and signals, $L(\{\mathbf{X}(n)\}, \{\mathbf{M}_h(n)\})$. The similarity measure is a function of model parameters and associations between the input synapses and concept-models. It is constructed in such a way that any of a large number of objects can be recognized, no matter if they appear on the left or on the right. Correspondingly, a similarity measure is designed so that it treats each concept model as an alternative for each subset of signals

$$L(\{\mathbf{X}\}, \{\mathbf{M}\}) = \prod_{n \in N} \sum_{h \in H} r(h) l(\mathbf{X}(n)|\mathbf{M}_h(n)); \tag{1}$$

$l(\mathbf{X}(n)|\mathbf{M}_h(n))$ (or simply $l(n|h)$) is a conditional partial similarity between one signal $\mathbf{X}(n)$ and one model $\mathbf{M}_h(n)$ – all possible combinations of signals and models are accounted for in this expression. Parameters $r(h)$ are proportional to the number of signals $\{n\}$ associated with the model h.

In the process of learning, concept-models are constantly modified. From time to time a system forms a new concept, while retaining an old one as well; alternatively, old concepts are sometimes merged. (Formation of new concepts and merging of old ones require a modification of the similarity measure (1); the reason is that more models always result in a better fit between the models and data. This is a well-known problem, it can be addressed by reducing the r.h.s. of equation (1) using a "penalty function", $p(N, M)$ that grows with the number of models M, and this growth is steeper for a smaller amount of data N. For example, an asymptotically unbiased maximum likelihood estimation leads to multiplicative $p(N, M) = \exp(-N_{par}/2)$, where N_{par} is the total number of adaptive parameters in all models (this penalty function is known as *Akaike Information Criterion*, see [40] for further discussion and references).

4.2 Fuzzy Dynamic Logic and MFT

The learning process consists in estimating model parameters $\mathbf{S}_h$ and associating subsets of signals with concepts by maximizing the similarity (1). Note, that equation (1) contains a large number of combinations of models and signals, a total of H^N items; this was a reason for the combinatorial complexity of the past algorithms discussed in section 2. Modeling field theory (MFT)

[1] In fact, there are many levels between the retina, visual cortex, and object models.

solves this problem by fuzzy dynamic logic [35, 40]. MFT introduces fuzzy association variables $f(h|n)$:

$$f(h|n) = \frac{r(h)l(n|h)}{\sum_{h' \in H} r(h')l(n|h')} \tag{2}$$

These variables give a measure of correspondence between signal $\mathbf{X}(n)$ and model $\mathbf{M}_h$ relative to all other models, h'. A mechanism of concept formation and learning, an internal dynamics of the modeling fields (MF) is defined as follows,

$$\mathbf{S}_h = \mathbf{S}_h + \alpha \sum_n f(h|n) \left(\frac{\partial \ln l(n|h)}{\partial \mathbf{M}_h} \right) \frac{\partial \mathbf{M}_h}{\partial \mathbf{S}_h} \tag{3}$$

$$r(h) = \frac{N_h}{N}; N_h = \sum_n f(h|n); \tag{4}$$

Parameter α determines the iteration step and speed of convergence of the MF system; N_h can be interpreted as a number of signals $\mathbf{X}(n)$ associated with or coming from a concept object n. As already mentioned, in the MF internal dynamics, similarity measures are adapted so that their fuzziness is matched to the model uncertainty. Mathematically, this can be accomplished in several ways, depending on the specific parameterization of the conditional partial similarity measures, $l(n|h)$; for example, they can be defined as Gaussian functions,

$$\begin{aligned} l(n|h) = &(2\pi)^{-\frac{d}{2}} (det \mathbf{C}_h)^{-\frac{1}{2}} \exp\{-0.5(\mathbf{X}(n) - \mathbf{M}_h(n))^T \\ &\mathbf{C}_h^{-1}(\mathbf{X}(n) - \mathbf{M}_h(n))\} \end{aligned} \tag{5}$$

In this formula, d is the dimensionality of the vectors $\mathbf{X}$ and $\mathbf{M}$, and $\mathbf{C}_h$ is a covariance. The dynamics of fuzziness of the MF similarity measures is defined as

$$\mathbf{C}_h = \sum_n f(h|n)(\mathbf{X}(n) - \mathbf{M}_h(n))(\mathbf{X}(n) - \mathbf{M}_h(n))^T / N_h \tag{6}$$

Initially, models do not match data, covariances are large, and association variables, $f(h|n)$, take homogeneous values across the data, associating all concept-models h with all input signals n. As matching improves, covariances become smaller, and the association variables, $f(h|n)$, tend to high values 1 for some subsets of signals and models and zero for others; thus certain concepts get associated with certain subsets of signals (objects are recognized and concepts formed). The following theorem was proven in [40]:

Theorem 1. *Equations (2) through (6) define a convergent dynamic system MF with stationary states given by* $\max_{\{\mathbf{S}_h\}} L$.

In plain language this means that the above equations indeed result in concept-models in the "mind" of the MFT system, which are most similar – in terms of similarity measure (1) – to the sensory data. Despite a combinatorially large number of items in equation (1), a computational complexity of the MF method is relatively low, it is linear in N and could be implemented by a physical system (like a computer or a brain). These equations describe a closed loop system, which is illustrated in figure 1. A reference to the *closed* loop emphasizes that the loop sustains its operations on its own, the loop is not closed in that there are input signals into the loop and output concepts from the loop.

Comment. Equation (5) of conditional partial similarities using Gaussian functions can be considered a basis for the following probabilistic interpretation: A model $\mathbf{M}_h(\mathbf{S}_h, n)$ is a conditional statistical expectation of signals from object h described by parameters $\mathbf{S}_h$. A similarity measure (1) is a total likelihood. Let me emphasize that such an interpretation could be valid if for some values of the parameters, the models are accurate (that is, models actually *are* conditional statistical expectation). If models are approximate in a non-statistical sense, other similarity measures could be more preferable mathematically, like mutual information in the models about the data [40]. I would also like to emphasize that unlike usual "Gaussian assumption" this model is quite general, it *does not assume that the signal distribution is Gaussian*, but only the deviations between the models and signals are, this model can represent any statistical distribution [40].

Summary of the MF convergence: during an adaptation process initial fuzzy and uncertain models (internal structures of the MF system) are associated with structures in the input signals, fuzzy models are getting more definite and crisp. The type, shape and number of models are selected so that the internal representation within the system is similar to input signals: The MF concept-models represent structure-objects in the input signals. Mathematical equations which describe this process are called *fuzzy dynamic logic* [40] which in terms of mind-internal processes describes an elementary thinking process involving instincts, imagination, emotions and concepts. But before discussing this cognitive-psychological interpretations, lets us briefly look into integrating this process with language.

4.3 Integrating Language and Thinking

During visual perception, when internal representation-models are matched in the visual cortex to retinal signals, cortex representations maintain their spatial topology and continuity. A number of MFT models have been developed for visual perception, for other sensor modalities, and for cognition of simple situations [40]. By using concept-models with multiple sensor modalities, a MFT system can integrate signals from multiple sensors, while adapting and improving internal concept-models. Similarly, MFT can be used to integrate

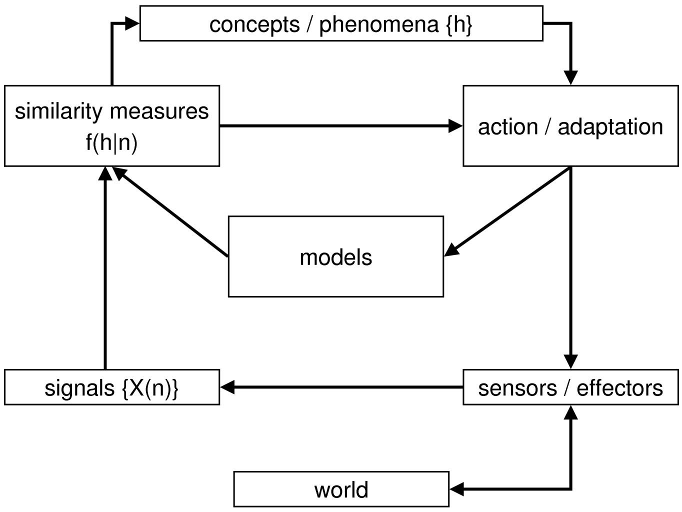

Fig. 1. For a single level of MFT, input signals are unstructured data $\{\mathbf{X}(n)\}$ and output signals are recognized or formed concepts $\{h\}$ with high values of similarity measures. The MFT equations (2) through (6) describe a continuous loop operation involving input signals, similarity measures, models, and actions of the model adaptation (the inner loop in this figure). Psychologically, a similarity measure corresponds to the knowledge instinct and its changes to aesthetic emotions.

language and thinking. This requires the development of linguistic MFT models. Here, I briefly outline an approach to the development of MFT linguistic models. Like MFT, language is a hierarchical system. Among other things, it involves sounds, phonemes, words, phrases, sentences, and texts, where each level operates with its own models. Like other models of the mind, these models are a result of evolution; for computational intelligent systems we have to develop them, and this development at each level is a research project, which is added by a number of already described linguistic models [16, 26, 44, 46].

In order to give an illustration, I discuss an approach to the development of models of phrases from words in the context of text understanding which, for example, could be used for an understanding-based search engine. The input data, $\mathbf{X}(n)$, in this "phrase-level" MF system are word strings of a fixed length S. Thus: $\mathbf{X}(n) = \{w_{n+1}, w_{n+2}, \ldots, w_{n+S}\}$. w_n are words of a given dictionary $W = \{w_1, w_2, \ldots, w_K\}$ of size K, and n is the word position in a body of texts. A simple phrase model is "a bag of words", that is, a model is a subset of words from a dictionary, without any order or rules of grammar,

$$\mathbf{M}_h(\mathbf{S}_h, n) = \{w_{h,1}, w_{h,2}, \ldots, w_{h,S}\} \; . \tag{7}$$

The parameters of this model are its words, $\mathbf{M}_h(\mathbf{S}_h, n) = \mathbf{S}_h = \{w_{h,1}, w_{h,2}, \ldots, w_{h,S}\}$. The language acquisition project in this simplified context consists in defining models-concepts-phrases best characterizing the given body of texts in terms of a similarity measure.

Conditional partial similarities between a string of text, $\mathbf{X}(n)$, and a model $\mathbf{M}_h$ could be defined by a proportion of the matches between the two sets, $\mathbf{X}(n)$ and $\mathbf{M}_h, l(n|h) = |\mathbf{X}(n) \cap \mathbf{M}_h|/S$. Thus, similarity in terms of equation (1) would be defined and could be maximized over the unknown parameters of the system, $\{\mathbf{S}_h\}$, that is, over the word contents of phrases. This would result in learning models-concepts-phrases, accomplishing the goal of the language acquisition project. The difficulty of the above approach is that the dynamics of MFT cannot be used for the similarity maximization, in particular, equation (3) requires evaluating derivatives, which requires a smooth dependence of models on their parameters. Without the fuzzy dynamic logic of MFT, the computational complexity of this language acquisition project becomes combinatorial, i.e. $\sim K^{(H^*N^*S)}$, which is a prohibitively large number.

The combinatorial complexity of the above solution is related to a "logic-type" similarity measure, which treats every potential phrase model (every combination of words) as a separate logical statement. The problem can be solved by using dynamic fuzzy phrase contents as follows. First, define fuzzy conditional partial similarity measures:

$$l(n|h) = (2\pi\sigma_h^2)^{-\frac{S}{2}} \exp\{-0.5 \sum_s e(n,h,s)^2/\sigma_h^2\}, \tag{8}$$

where $e(n,h,s)$ is a distance (measured in the numbers of words) between the middle of the word sequence $\mathbf{X}(n)$, that is $n+S/2$, and the closest occurrence of the word $w_{h,s}$; the sum here is over words belonging to the phrase model h. In practical implementations, the search for the nearest word can be limited by $\pm 3\sigma_h$ words, and $e(n,h,s)$ falling outside this range can be substituted by a $(3\sigma_h + 1)$. The dynamics of fuzziness of this similarity measure is given by a modification of equation (6),

$$\sigma_h^2 = \sum_n f(h|n) \sum_s e(n,h,s)^2/N_h \ . \tag{9}$$

Second, define fuzzy phrase contents, that is a degree of the word $w_{h,s}$ "belonging" to a model-phrase h, $\phi(s|h)$; this is a function of the average distance of the word $w_{h,s}$ from the phrase model $\varepsilon(s,h)$:

$$\varepsilon(h,s) = \sum_n f(h|n) e(n,h,s)^2/N_h; \tag{10}$$

$$\begin{aligned}\phi(s|h) =& p(h|s)/\sum_{s' \in h} p(h|s');\\ p(h|s) =& (2\pi\sigma_h^2)^{-1/2} \exp\{-0.5 \sum_s \varepsilon(h,s)/\sigma_h^2\},\end{aligned} \tag{11}$$

The dynamics of the word contents of the phrase models is given by modifying S (the number of words in phrases) in the iteration process, say, by defining $S_h \sim S\sigma_h$, or by requiring $\phi(s|h)$ to be above a threshold value, and keeping in each phrase model the words satisfying this criteria. The dynamics defined in this way results in learning phrase models (concepts) and accomplishes the goal of the language acquisition project without combinatorial complexity, the computational complexity is moderate, $\sim H^*K^*S^2$.

The "bag-of-words" phrase models considered above are much simpler than tree-like dependencies or known structures of natural languages [16, 25, 26, 44, 46, 48]. These more complicated "real" linguistic models can be used in place of a simple distance measure $e(n, h, s)$ in equation (8). In this way the models of noun and verb phrases and tree structures can be incorporated into the above formalism of MFT.

Integration of language and cognition in MFT is attained by characterizing objects and situations in the world with two types of models, linguistic models considered above and cognitive models considered in section 4.2 and in [40]. Such integrated MFT system learns – similarly to human – in parallel in three realms: (1) linguistic models can be learned to some extent independently from cognition, when linguistic data are encountered for the first time with limited or no association with perception and cognition (like in a newborn baby); (2) similarly, cognitive models can be learned to some extent independently from language, when perception signal data are encountered for the first time in limited or no association with linguistic data; and (3) linguistic and cognitive models are learned jointly, when linguistic data are present in some association with perception signals, like during mother talking to a baby: "this is a car" (perception models and word models), and like during more complicated conversations: "Look at Peter and Ann, they are in love" (cognitive models and phrase models).

A Constructed Example. A real-life example of this approach would be too voluminous and boring to follow. Here is a simplified constructed example to illustrate some of the main points of learning phrase models. It starts with a large text data base (hundreds of millions of words) and partitions it into 10-word chunks. Four of these chunks are shown here containing a word *chair*:

fifth chair foundation not-for-profit organization devoted fostering online bridge
education hickory chair furniture catalog register wish list store locator contact
fork picnic table set choice chairs benches sets fork table site give information
university course provide software engineering chair involved

After several iterations the algorithm learned phrase models limited to 6-word length; one hundred thousand of the mostly often used phrase models were retained, among them the following four were connected to the previous chunks with appreciable probabilities:

chair foundation nonprofit organization online education
online furniture catalog store brand discount

ascii table hexadecimal octal set character
university chair professor appointment invitation name

After several more iterations ten thousands of most useful 4 word phrase models were retained, among them the following five were connected to the previous chunks with appreciable probabilities:

organization nonprofit community service
online furniture catalog website
brand name furniture discount
ascii table hexadecimal octal
university chair professor appointment

Higher levels of generalization (fewer word sentences with broader meanings) require moving to a higher level of a multi-level hierarchical system.

4.4 MFT Hierarchical Organization

The previous subsections described a single processing layer in a hierarchical MFT system. Inputs to each layer are signals $\mathbf{X}(n)$, or in neural terminology, an input field of neuronal activations. Outputs are the activated models $\mathbf{M}_h(\mathbf{S}_h, n)$; it is a set of models or concepts recognized in the input signals. Equations (2-6) and (8-11), as shown in figure 1, can be interpreted as a loop process: at each iteration the equations contain association variables $f(h|n)$ and model parameters computed at the previous iteration. In other words, the output models "act" upon the input to produce a "refined" output model (at the next iteration). This process is directed at increasing the similarity between the models and signals. It can be described as an internal behavior of model adaptation.

The output models initiate other actions as well. First, activated models (neuronal axons) serve as input signals to the next processing layer, where more general concept-models are recognized or created. Second, concept-models along with the corresponding instinctual signals and emotions may activate behavioral models and generate behavior directed into the outside world (a process not contained within the above equations). In general, a higher level in a hierarchical system provides a feedback input into a lower level. For example, sensitivities of retinal ganglion cells depend on the objects and situations recognized higher up in the visual cortex; or, a gaze is directed based on which objects are recognized in the field of view. These interactions within this hierarchical organization are illustrated in figure 2.

Concept objects identified at the output of the lower level of MFT system in figure 2 become input signals to the next MFT level which identifies more general concepts of relationships among objects and situations; at the same time more general concepts of understanding identified at a higher level activate behavioral concept-models that affect processes at a lower level. The

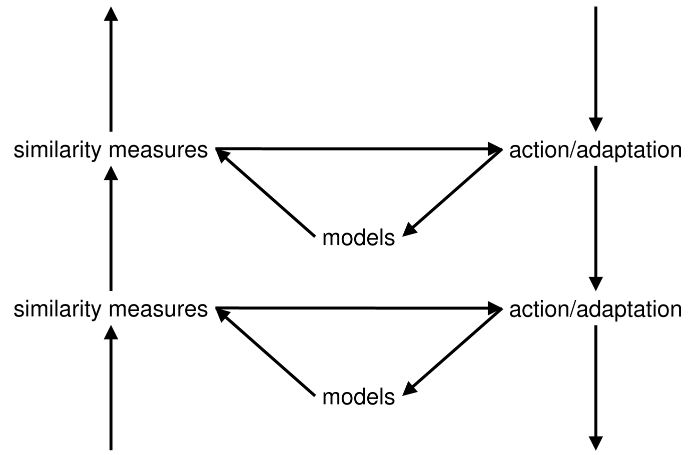

Fig. 2. Hierarchical organization of the MFT system. High levels of similarity measures correspond to concepts recognized at a given level in the hierarchy; these are the input signals to the next, higher level. Also concepts affect behavior (actions). Models at a higher level are more general than models at a lower level.

agent processes, or the loop processes of model concept adaptation, understanding and behavior generation continue up and down the hierarchy of the MFT levels.

The loop of operations of MFT can also be described as multiple loops each involving a single model $h, h = 1, \ldots, H$. To some extent these multiple loops are independent, yet some models interact when they are associated with the same input signals. Each model along with its adaptation mechanism is an intelligent agent, which possesses a degree of autonomy and is interacting with other agents. Thus MFT is an intelligent system composed of multiple adaptive intelligent agents. Each agent, including its concept model along with the similarity measure and behavioral response, is a continuous loop of operations, interacting with other agents from time to time; an agent is "dormant" until activated by a high similarity value. When activated, it is adapted to the signals and other agents, so that the similarity increases. A subset of data in input signals may activate several concept agents, in this way data provide evidence for the presence of various objects (or concepts). Agents compete with each other for evidence (matching to signals), while adapting to the new signals.

A multi-level hierarchical linguistic MFT system can be developed by adding more levels similar to a word phrase level described in section 4.3. A relatively simple system can use similar "bag" models for each layer, like "bag of phrases" model for the next level of concepts (say, sentence), and so on. Alternatively, more realistic linguistic models of sentences, paragraphs and large bodies of texts can be utilized (cf. Rieger [47] and Mehler [24]). Among many possible commercial applications of such systems could be understanding-based search engines; everybody familiar with the frustration of the web searches would appreciate a search engine that even remotely understands user queries and contents of the web pages.

5 MFT Theory of Mind

5.1 MFT Dynamics

Equations (2-6) and (8-11) describe elementary processes of perception or cognition, in which a number of model concepts compete for incoming signals, model concepts are modified and new ones are formed, and eventually, more or less definite connections (high values of $f(h|n)$, close to 1) are established among signal subsets on the one hand, and some model concepts on the other, accomplishing perception and cognition.

A salient mathematical property of this processes ensuring a smooth convergence is a correspondence between uncertainty in models (that is, in the knowledge of model parameters) and uncertainty in associations $f(h|n)$. In perception, as long as model parameters do not correspond to actual objects, there is no match between models and signals; many models poorly match many objects, and associations remain fuzzy (between 0 and 1). Eventually, one model (h') wins a competition for a subset $\{n'\}$ of input signals $\mathbf{X}(n)$, when parameter values match object properties, and $f(h'|n)$ values become close to 1 for $n \in \{n'\}$ and 0 for $n \notin \{n'\}$. This means that this subset of data is recognized as a specific object (concept). Upon the convergence, the entire set of input signals $\{n\}$ is divided into subsets, each associated with one model object, uncertainties become small, and fuzzy a priori concepts become crisp concepts. Cognition is different from perception in that models are more general, more abstracts, and input signals are the activation signals from concepts identified (cognized) at a lower hierarchical level; the general mathematical laws of cognition and perception are similar and constitute a basic principle of the mind organization. Kant was the first one to propose that the mind functioning involves three basic abilities: Pure Reason (concept-models), Judgment (emotional measure of correspondence between models and input signals), and Practical Reason (behavior; we only considered here the behavior of adaptation and learning) [21, 20, 22]. Let us discuss relationships between the MFT theory and concepts of mind originated in psychology, philosophy, linguistics, aesthetics, neuro-physiology, neural networks, artificial intelligence, pattern recognition, and intelligent systems.

5.2 Elementary Thought Process, Conscious, and Unconscious

A thought process or thinking involves a number of sub-processes and attributes, including internal representations and their manipulation, attention, memory, concept formation, knowledge, generalization, recognition, understanding, meaning, prediction, imagination, intuition, emotion, decisions, reasoning, goals, behavior, conscious and unconscious [25, 27, 40]. Here and in the following subsections we discuss how these processes are described by MFT.

A "minimal" subset of these processes, *an elementary thought process*, has to involve mechanisms for afferent and efferent signals [10], in other words, bottom-up and top-down signals coming from outside (external sensor signals) and from inside (internal representation signals). According to Carpenter and Grossberg [3] every recognition and concept formation process involves a "resonance" between these two types of signals. In MFT, at every level in a hierarchy the afferent signals are represented by the input signal field $\mathbf{X}$, and the efferent signals are represented by the modeling fields $\mathbf{M}_h$; resonances correspond to high similarity values $l(n|h)$ for some subsets of $\{n\}$ that are "recognized" as concepts (or objects). The mechanism leading to the resonances between incoming signals and internal representations is given by equations in sections 4.2 and 4.3. The elementary thought process also involves elements of conscious and unconscious processes, imagination, memory, concepts, instincts, emotions, understanding and behavior as described later.

A description of working of the mind as given by the MFT dynamics was first provided by Aristotle [1], describing thinking as a learning process in which an a priori form-as-potentiality (fuzzy model) meets matter (sensory signals) and becomes a form-as-actuality (a concept). Jung suggested that conscious concepts are developed by the mind based on genetically inherited structures, archetypes, which are inaccessible to consciousness [19], and Grossberg [10] suggested that only signals and models attaining a resonant state (that is signals matching models) reach consciousness. Fuzzy uncertain models are less accessible to consciousness, whereas more crisp and certain models are better accessible to consciousness.

5.3 Understanding

In the elementary thought process, subsets in the incoming signals are associated with recognized model objects, creating *phenomena* (in the MFT-mind) which are *understood* as objects, in other words *signal subsets* acquire *meaning* (e.g., a subset of retinal signals acquires a meaning of a chair). There are several aspects to understanding and meaning. First, object-models are connected (by emotional signals [8, 12, 38, 39, 40]) to instincts that they might satisfy, and also to behavioral models that can make use of them for instinct satisfaction. Only two instincts and types of behavior are described within equations of section 4: (1) the knowledge instinct and behavior of learning of perception and cognition models (that is improving and adapting these models for better correspondence to the world), and (2) the language instinct and behavior of learning linguistic models (that is improving and adapting these models for better correspondence to the language data, like words and grammar). A formulation is proposed where these two instincts are closely related and can be considered as two aspects of the same instinct.

The second aspects of understanding and meaning is that an object, situation, or phrase is understood in the context of a more general situation in the next layer, consisting of more general concept-models, which accepts as input signals the results of object recognition. That is, each recognized object model (phenomenon) sends (in neural terminology, activates) an output signal; and a set of these signals comprises input signals for the next layer models, which 'cognize' more general concept-models. And this process continues up and up the hierarchy of the models and mind toward the most general models a system could come up with, such as models of universe (scientific theories), models of self (psychological concepts), models of meaning of existence (philosophical concepts), models of a priori transcendent intelligent subject (theological concepts).

5.4 Imagination

Visual imagination involves excitation of a neural pattern in a visual cortex in absence of an actual sensory stimulation (say, with closed eyes) [10]. Imagination was often considered to be a part of thinking processes; Kant [20] emphasized the role of imagination in the thought process, he called thinking "a play of cognitive functions of imagination and understanding". Whereas pattern recognition and artificial intelligence algorithms of recent past would not know how to relate to this [29, 31], Carpenter and Grossberg's resonance model [3] and the MFT dynamics both describe imagination as an inseparable part of thinking: imagined patterns are top-down signals that *prime* the perception cortex areas (*priming* is a neural terminology for making neural cells to be more readily excited). In MFT, models $\mathbf{M}_h$ give the imagined neural patterns. MFT (in agreement with neural data) just adds details to Kantian description: thinking is a play of *higher-hierarchical level* imagination and *lower-level* understanding. Kant identified this "play" (described by equations (2-6) or (8-11)) as a source of aesthetic emotions discussed later.

5.5 Mind vs. Brain

Historically, the mind is described in psychological and philosophical terms, whereas the brain is described in terms of neurobiology and medicine. Within scientific exploration the mind and brain are different description levels of the same system. Establishing relationships between these description is of great scientific interest. Today we approach solutions to this challenge [11], which eluded Newton in his attempt to establish physics of "spiritual substance" [50]. General neural mechanisms of the elementary thought process (which are similar in MFT and ART [3]) have been confirmed by neural and psychological experiments, this includes neural mechanisms for bottom-up (sensory) signals, top-down "imagination" model signals, and the resonant matching between the two [9, 10, 52]. Adaptive modeling abilities are well studied and adaptive parameters identified with synaptic connections [15, 23]; instinctual learning mechanisms have been studied in psychology and linguistics [4, 7, 41, 43]; identifying neural structures responsible for knowledge and language instincts is a next challenge for the neural sciences.

5.6 Instincts and Emotions

Functioning of the mind and brain cannot be understood in isolation from the system's "bodily needs". For example, a biological system (and any autonomous system) needs to replenish its energy resources (*eat*); this and other fundamental unconditional needs are indicated to the system by instincts, which could be described as internal sensors. Emotional signals, generated by this instinct are perceived by consciousness as "hunger", and they activate behavioral models related to food searching and eating. In this paper we are concerned primarily with the behavior of recognition: instinctual influence on recognition modify the object perception process (3) - (6) in such a way that desired objects "get" enhanced recognition. It can be accomplished by modifying priors, $r(h)$, according to the degree to which an object of type h can satisfy a particular instinct. Details of these mechanisms are not considered here, except for the two instincts considered in this paper.

5.7 Aesthetic Emotions and Instinct for Knowledge

Recognizing objects in the environment and understanding their meaning is so important for human evolutionary success that there has evolved an instinct for learning and improving concept models. This instinct (for knowledge and learning) is described in MFT by maximization of similarity between the models and the world according to equation (1). Emotions related to satisfaction/dissatisfaction of this instinct we perceive as harmony/disharmony (between our understanding of how things ought to be and how they actually are in the surrounding world). According to Kant [20] these are aesthetic emotions (emotions that are not related directly to satisfaction or dissatisfaction of bodily needs). Aesthetic emotions in MFT correspond to changes in the knowledge instinct (1). The aesthetic emotion is negative, when new input signals do not correspond well to existing models. The mathematical basis for the theorem stated after equation (6) can be interpreted psychologically: During iterations defined by the equations (2-6) the aesthetic emotion is always positive.

In sections 4.2 we considered perception and cognition concept models and similarity measures; using them in equation (1) yields an instinct driving the MFT system to improve the knowledge about the world. Similarly, using linguistic models in equation (1) and the similarity measures considered in section 4.3, yields the MFT system improving the knowledge of language, or language instinct. Combining cognitive and linguistic models results in a system with combined linguistic and thinking abilities: language and sensory information together help adapting both, linguistic and cognitive models. A specific mathematical mechanism combining language and cognition described in section 4 associates both types of models with every object and situation. We do not know if the mind works this way. Finding out actual neural mechanisms combining language and cognition is a future challenge.

5.8 Beauty and Intuition

Harmony is an elementary aesthetic emotion related to improvement of object-models. Higher aesthetic emotions are related to the development of more complex "higher" models: we perceive an object or situation as aesthetically pleasing if it satisfies our learning instinct, that is the need for improving the models and increasing similarity (1). The highest forms of aesthetic emotion are related to the most general and most important models. According to Kantian analysis [20], among the highest models are models of the meaning of our existence, of our purposiveness or intentionality, and beauty is related to improving these models: we perceive an object or a situation as beautiful, when it stimulates improvement of these highest models of meaning. Beautiful is what "reminds" us of our purposiveness.

Intuition includes an intuitive perception (imagination) of object models and their relationships with objects in the world, higher-level models of relationships among simpler models, and behavioral models. Intuition involves fuzzy unconscious concept models, which are in a state of being learned and being adapted toward crisp and conscious models (a "thought" or a theory); such models may satisfy or dissatisfy the knowledge instinct in varying degrees before they are accessible to consciousness, hence the complex emotional feel of an intuition. The beauty of a physical theory discussed often by physicists is related to satisfying our feeling of purpose in the world that is satisfying our need to improve the models of the meaning through understanding of the universe.

5.9 Theory Testing and Future Directions

The general neural mechanisms of the elementary thought process, which include neural mechanisms for bottom-up (sensory) signals, top-down "imagination" model signals, and the resonant matching between the two have been confirmed by neural and psychological experiments (these mechanisms are similar in MFT and ART [9, 10, 40, 52]). Adaptive modeling abilities are well studied and adaptive parameters have been identified with synaptic connections [15, 23]; instinctual learning mechanisms have been studied in psychology and linguistics [4, 7, 41, 43]. Ongoing and future research will confirm, disprove, or suggest modifications to specific mechanisms of parameter adaptation (equation 2-5), reduction of fuzziness during learning (equation 6), similarity measure (equation 1) as a foundation of aesthetic instinct for knowledge, relationships between psychological and neural mechanisms of learning on the one hand and, on the other, aesthetic feelings of harmony and emotion of beauty. Differentiated forms of (1) need to be developed for various forms of the knowledge instinct (differentiation between cognition and language instincts in this paper is a step in this direction). Future experimental research needs to study in details the nature of hierarchical interactions: to what extent the hierarchy is "hardwired" vs. adaptively emerging in ontogenesis and throughout life; theory of emerging hierarchical models will have to be developed. For a combined theory of language and cognition, future experimental research ought to identify neural mechanisms combining linguistic and cognitive concepts, prove or disprove the mechanisms proposed in this paper, and also study the ontogenesis of these mechanisms in child development processes.

5.10 Thinking Process and Semiotics

Semiotics studies symbol content of culture [49]. For example, consider a written word "chair". It can be interpreted by a mind to refer to something else: an entity in the world, a specific chair, or the concept "chair" in the mind. In this process, the mind, or an intelligent system is called *an interpreter*, the written word is called *a sign*, the real-world chair is called *a designatum*, and the concept in the interpreter's mind, the internal representation of the results of interpretation is called *an interpretant* of the sign. The essence of a sign is that it can be interpreted by an interpreter to refer to something else, a designatum. This process of sign interpretation is an element of a more general process called *semiosis*, which consists of multiple processes of sign interpretation at multiple levels of the mind hierarchy.

In mathematics and in "Symbolic AI" there is no difference between signs and symbols. Both are considered as notations, arbitrary non-adaptive entities with axiomatically fixed meaning. This non-differentiation is a "hangover" from an old superstition that logic describes mind, a direction in mathematics and logical philosophy that can be traced through the works of Frege, Hilbert, Russell, to its bitter end in Gödel theory, and its revival during the 1960s and 1970s in artificial intelligence. In general culture, symbols are understood also as psychological processes of sign interpretation. Jung emphasized that symbol processes connect conscious and unconscious [19], Pribram wrote of symbols as adaptive, context-sensitive signals in the brain, whereas signs he identified with less adaptive and relatively context-insensitive neural signals [45].

In classical semiotics [30, 32] words *sign* and *symbol* were not used consistently; in the context of the mathematical description in this paper, a sign means something that can be interpreted to mean something else (like a mathematical notation, or a word), and the process of interpretation is called a symbol process, or symbol. Interpretation, or understanding of a sign by the mind according to MFT is due to the fact that a sign (e.g., a word) is a part of an object model (or a situation model at higher levels of the mind hierarchy). The mechanism of a sign interpretation therefore involves first an activation of an object model, which is connected to instincts that the object might satisfy, and also to behavioral models that can make use of this object for instinct satisfaction. Second, a sign is understood in the context of a more general situation in the next layer consisting of more general concept-models, which accepts as input signals the results of lower-level sign recognition. That is, recognized signs comprise input signals for the next layer models, which 'cognize' more general concept-models.

A symbol process of a sign interpretation coincides with an elementary thought process. Each sign interpretation or elementary thought process, a symbol, involves conscious and unconscious, emotions, concepts, and behavior; this definition connecting symbols to archetypes (fuzzy unconscious model concepts) corresponds to a usage in general culture and psychology. As described previously, this process continues up and up the hierarchy of models and mind toward the most general models. In semiotics this process is called *semiosis*, a continuous process of creating and interpreting the world outside (and inside our mind) as an infinite hierarchical stream of signs and symbol processes.

References

[1] Aristotle. Metaphysics. In J. Barnes, editor, *Complete Works of Aristotle*. Princeton University Press, Princeton, NJ, 1995.

[2] R. E. Bellman. *Adaptive Control Processes*. Princeton University Press, Princeton, NJ, 1961.

[3] G. A. Carpenter and S. Grossberg. A Massively Parallel Architecture for a Self-organizing Neural Pattern Recognition Machine. *Computer Vision, Graphics and Image Processing*, 37:54–115, 1987.

[4] N. Chomsky. *Language and Mind*. Harcourt Brace Javanovich, New York, 1972.

[5] N. Chomsky. Principles and Parameters in Syntactic Theory. In N. Hornstein and D. Lightfoot, editors, *Explanation in Linguistics. The Logical Problem of Language Acquisition*. Longman, London, 1981.

[6] A. R. Damasio. *Descartes' Error: Emotion, Reason, and the Human Brain*. Avon, New York, 1995.

[7] T. W. Deacon. *The Symbolic Species: The Co-Evolution of Language and the Brain*. W. W. Norton & Company, 1998.

[8] V. A. Dmitriev and L. I. Perlovsky. Art Form as an Object of Cognitive Modeling (Towards Development of Vygotsky's Semiotics Model). In *Proceedings of the 1996 Conference on Intelligent Systems and Semiotics*, volume 2, pages 385–389, Gaithersburg, 1996.

[9] W. J. Freeman. *Mass Action in the Nervous System*. Academic Press, New York, 1975.

[10] S. Grossberg. *Neural Networks and Natural Intelligence.* MIT Press, Cambridge, MA, 1988.
[11] S. Grossberg. Linking Mind to Brain: The Mathematics of Biological Intelligence. *Notices of the American Mathematical Society*, 47:1361–1372, 2000.
[12] S. Grossberg and D. S. Levine. Neural Dynamics of Attentionally Modulated Pavlovian Conditioning: Blocking, Inter-stimulus Interval, and Secondary Reinforcement. *Psychobiology*, 15(3):195–240, 1987.
[13] S. Grossberg and N. A. Schmajuk. Neural Dynamics of Attentionally Modulated Pavlovian Conditioning: Conditioned Reinforcement, Inhibition, and Opponent Processing. *Psychobiology*, 15(3):195–240, 1987.
[14] S. R. Hameroff. *Toward a Scientific Basis for Consciousness.* MIT Press, Cambridge, MA, 1994.
[15] D. Hebb. *Organization of Behavior.* J. Wiley & Sons, New York, 1949.
[16] R. Jackendoff. *Foundations of Language: Brain, Meaning, Grammar, Evolution.* Oxford University Press, New York, 2002.
[17] J.-S. R. Jang, C.-T. Sun, and E. Mizutani. *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence.* Prentice Hall, Upper Saddle River, 1996.
[18] B. D. Josephson. An Integrated Theory of Nervous System Functioning Embracing Nativism and Constructivism. In *International Complex Systems Conference*, Nashua, NH, September 21-26 1997.
[19] C. G. Jung. Archetypes of the Collective Unconscious. In *The Collected Works*, volume 9.II of *Bollingen Series XX, 1969.* Princeton University Press, Princeton, NJ, 1934.
[20] I. Kant. *Critique of Judgment.* Macmillan & Co., London, 2nd edition, 1914.
[21] I. Kant. *Critique of Pure Reason.* Wiley Book, New York, 1943.
[22] I. Kant. *Critique of Practical Reason.* Hafner, 1986.
[23] C. Koch and I. Segev, editors. *Methods in Neuronal Modeling: From Ions to Networks.* MIT Press, Cambridge, MA, 1998.
[24] A. Mehler. A Multiresolutional Approach to Fuzzy Text Meaning. A First Attempt. In A. Meystel, editor, *Proceedings of the 1996 International Multidisciplinary Conference on Intelligent Systems: A Semiotic Perspective*, volume I, pages 261–273, Gaithersburg, 1996. National Institute of Standards and Technology.
[25] A. Mehler. Components of a Model of Context-Sensitive Hypertexts. *Journal of Universal Computer Science*, 8(10):924–943, 2002.
[26] A. Mehler. Hierarchical Orderings of Textual Units. In *Proceedings of the 19th International Conference on Computational Linguistics, COLING'02, Taipei*, pages 646–652, San Francisco, 2002. Morgan Kaufmann.
[27] A. Meystel. *Semiotic Modeling and Situational Analysis.* AdRem, Bala Cynwyd, PA, 1995.
[28] A. M. Meystel and J. S. Albus. *Intelligent Systems: Architecture, Design, and Control.* Wiley, New York, 2001.

[29] M. Minsky. *The Society of Mind.* MIT Press, Cambridge, MA, 1988.
[30] C. Morris. *Writings on the General Theory of Signs.* Mouton, The Hague, 1971.
[31] A. Newell. Intellectual Issues in the History of Artificial Intelligence. In F. Machlup and U. Mansfield, editors, *The Study of Information.* J. Wiley, New York, 1983.
[32] C. S. Peirce. *Collected Papers of Charles Sanders Peirce.* Harvard University Press, Cambridge, MA, 1935-66.
[33] R. Penrose. *Shadows of the Mind.* Oxford University Press, Oxford, 1994.
[34] L. I. Perlovsky. Gödel Theorem and Semiotics. In *Proceedings of the 1996 Conference on Intelligent Systems and Semiotics*, volume 2, pages 14–18, Gaithersburg, 1996.
[35] L. I. Perlovsky. Mathematical Concepts of Intellect. In *Proceedings of the World Congress on Neural Networks*, pages 1013–1016, San Diego, 1996. Lawrence Erlbaum Associates.
[36] L. I. Perlovsky. Towards Quantum Field Theory of Symbol. In *Proceedings of the 1997 Conference on Intelligent Systems and Semiotics*, pages 295–300, Gaithersburg, 1997.
[37] L. I. Perlovsky. Conundrum of Combinatorial Complexity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(6):666–670, 1998.
[38] L. I. Perlovsky. Cyberaesthetics: Aesthetics, learning, and control. In *STIS'98*, Gaithersburg, 1998.
[39] L. I. Perlovsky. Emotions, Learning, and Control. In *Proceedings of the International Symposium on Intelligent Control, Intelligent Systems & Semiotics*, pages 131–137, Cambridge, MA, 1999.
[40] L. I. Perlovsky. *Neural Networks and Intellect: Using Model-based Concepts.* Oxford University Press, New York, 2001.
[41] J. Piaget. *The Psychology of the Child.* Basic Books, 2000.
[42] J. P. e. Pickett, editor. *The American Heritage College Dictionary.* Houghton Mifflin, Boston, MA, 3rd edition, 2000.
[43] S. Pinker. *The Language Instinct: How the Mind Creates Language.* Harper Perennial, 2000.
[44] S. Pinker. *Words and Rules: The Ingredients of Language.* Harper Perennial, 2000.
[45] K. Pribram. *Languages of the Brain.* Prentice Hall, 1971.
[46] B. B. Rieger. Empirical Semantics II. A Collection of New Approaches in the Field. In *Quantitative Linguistics*, volume 13. Brockmeyer, Bochum, 1981.
[47] B. B. Rieger. Situation Semantics and Computational Linguistics: Towards Informational Ecology. In K. Kornwachs and K. Jacoby, editors, *Information. New Questions to a Multidisciplinary Concept*, pages 285–315. Akademie-Verlag, Berlin, 1995.
[48] B. B. Rieger. Tree-like Dispositional Dependency Structures for Non-propositional Semantic Inferencing: A SCIP Approach to Natural Lan-

guage Understanding by Machine. In B. Bouchon-Meunier and R. Yager, editors, *Proceedings of the 7th International Conference on Information Processing and Management of Uncertainty in Knowledge-based Systems (IPMU-198)*, pages 351–358, Paris, 1998.

[49] T. A. Sebeok. *Sign: An Introduction to Semiotics.* University of Toronto Press, Toronto, 1995.

[50] R. S. Westfall. *Never at Rest: A Biography of Isaac Newton.* Cambridge University Press, Cambridge, 1983.

[51] L. A. Zadeh. Information Granulation and its Centrality in Human and Machine Intelligence. In *Proceedings of the 1997 Conference on Intelligent Systems and Semiotics*, pages 26–30, Gaithersburg, 1997.

[52] S. Zeki. *A Vision of the Brain.* Blackwell, Oxford, 1993.