



Philosophy of Mind

Computational Complexity and the Origin of Universals

Leonid I. Perlovsky
Nichols Research Corporation
lperl@bellatlantic.net

ABSTRACT: This paper establishes close relationships between fundamental problems in the philosophical and mathematical theories of mind. It reviews the mathematical concepts of intelligence, including pattern recognition algorithms, neural networks and rule systems. Mathematical difficulties manifest as combinatorial complexity of algorithms are related to the roles of a priori knowledge and adaptive learning, the same issues that have shaped the two-thousand year old debate on the origins of the universal concepts of mind. Combining philosophical and mathematical analyses enables tracing current mathematical difficulties to the contradiction between Aristotelian logic and Aristotelian theory of mind (Forms). Aristotelian logic is shown to be the culprit for the current mathematical difficulties. I will also discuss connections to Gödel's theorems. The conclusion is that fuzzy logic is a fundamental requirement for combining adaptivity and apriority. Relating the mathematical and philosophical helps clarifying both and helps analyzing future research directions of the mathematics of intelligence.

I. Introduction: Mathematics and Philosophy

The two-thousand year old debate on the origins of universal concepts of mind was about the roles of adaptivity or learning from experience vs. the a priori knowledge (the inborn or God-given). It is closely related to the epistemological problem of the origins of knowledge. The problem of combining adaptivity and a-priority is fundamental to computational intelligence as well as to understanding human intelligence. There is an interrelationship among concepts of mind in mathematics, psychology, and philosophy, which is much closer than currently thought among scientists and philosophers of today. From the contemporary point of view, the questions about mind posed by ancient philosophers are astonishingly scientific. A central question to the work of Plato, Aristotle, Avicenna, Maimonides, Aquinas, Occam, and Kant was the question of the origins of universal concepts. Are we born with a priori knowledge of concepts or do we acquire this knowledge adaptively by learning from experience? This question was central to the work of ancient philosophers, medieval theologians, and it was equally important to theories of Freud, Jung, and Skinner. The different answers they gave to this question are very similar to the answers given by McCulloch, Minsky, Chomsky and Grossberg.

When 2300 years ago Plato faced a need to explain our ability to conceptualize, he concluded that concepts are of a priori origin. The philosophy based on the transcendental, a priori reality of concepts was named *realism*. During the following 2000 years the

concept of a-priority was tremendously strengthened by the development of monotheistic religion in Europe, to the extent that it interfered with empirical studies. At the end of scholastic era, human spirit felt strong enough to question a priori truths on the empirical ground. Occam rejected the concept of a-priority; he held nominalistic views that are opposite to realism. Following Antisthenes, *nominalism* considers ideas to be just names for classes of similar empirical facts. Occam prepared the way for empiricism of Lock and Hume, that is among foundations of the scientific method.

Time has obscured the influence of Occam on the development of the scientific method, and his name is hidden behind the figures of great philosophers and scientists that came after him. However, despite the realism of Descartes, Leibnitz, and Newton, nominalism of the forerunner of contemporary scientific thinking continues to pervade scientific attitudes of today. One of the reasons for the influence of nominalism is the unbreakable tie between the scientific method and objectivization of the subject of inquiry. In physics, theoretical tradition of the Newton's realism counterbalanced the influence of nominalism, but in the area of empirical sciences, such as psychology in the last century, the reality of facts seemed more significant than the reality of ideas that have not been clad in a mathematical form.

Near the end of 19th century, the success of the mathematical method in physics had advanced a requirement of objectivization and, in the empirical sciences, where the only criterion of objectivity was seen in the reproducible experiments, questioned a possibility of a theoretical consideration of a priori concepts. A priori concepts started loosing ground, became lowered to the level of (at best) unproved hypothesis, and I would risk to say that in some areas of science a temptation of objectivity eliminated a possibility for deep theoretical scientific thinking. Concepts dressed not in the strict language of mathematical computations, seemed compromised. In this atmosphere, to resolve the dilemma between the objectivity and depth of investigation, there was born behaviorism, a new scientific direction redefining psychology as a science of human behavior (Watson, 1913) and an accompanying intellectual and philosophical movement (Skinner, 1974).

A concept of behaviorism that attempted to explain the entire human psychology as a sequence of stimuli and reflexes and denied a need for consciousness in understanding of the intellect, dominated American psychology from about 1920 to 1960 (Jaynes, 1976). One of the reasons for the past popularity of behaviorism was a striving toward scientific strictness in the absence of mathematical methods adequate for the complicated problem of the analysis of mind. Seeing the only criteria of scientific objectivity in reproducible experimental results, behaviorism had to forgo considerations of deep mental processes (Grossberg, 1988). Behaviorism as a scientific school, as a temporary idealization of a complicated problem, created a scientific methodology of experimental psychology, established an importance of the environment as a determining factor in human behavior, showed that the role of mental factors is often incorrectly exaggerated in everyday life, and successfully described multiple aspects of behavior in terms of external factors alone. However, behaviorism as a philosophy maintaining that the concepts of *consciousness*, *free will*, *idea*, are not needed in psychology and should be discarded (Skinner, 1974), exerted an inhibiting influence on the development of concepts of mind. As an attempt to reduce psychology exclusively to external factors, — behaviorism is a continuation of an ancient philosophical tradition of nominalism expressed in psychological terms of the twentieth century.

Emergence of cybernetics proceeded under the influence of the dominating psychological concept of behaviorism, which can be seen from the cybernetics' program paper (Rosenblueth, Wiener, & Bigelow, 1943). The mutual influence of behaviorism, nominalistic philosophy, and cybernetics was enhanced by the fact that available cybernetic

models were relatively simple linear Wiener filters, suitable for utilization of only simple a priori knowledge. It was truly revolutionary that despite of these prevailing nominalistic orientation, McCulloch came to a conclusion that under the influence of nominalistic concepts since Occam, the realistic logic (based on the a-priority of ideas) decayed, which caused problems for scientific understanding of mind (McCulloch, 1961; 1965). The basis of the search for the material structures of intellect McCulloch founded on a realistic philosophy, created by the school of Plato and Aristotle. However, early neural network research in 1950s and 1960s did not follow this direction and pursued nominalistic concept of learning from examples, without using complicated a priori knowledge, until the demise of behaviorism in 1960s.

The early research in neural networks from 1940s to 1960s has generated tremendous interest as it promised to resolve the mystery of mind. Why did the Goliath-to-be fell down in 1960s? How did it happen that a relatively mild criticism by Minsky and Papert (1969) had a devastating effect on the interest in artificial neural systems? The question of why did this happen was widely discussed in a scientific community. However, the often offered explanations pointing to personal opinions can not be accepted, as unscientific and relatively useless. A personal opinion can produce a large scale effect in a society only if it captures, embodies, and serves as a conduit for a new philosophical trend. The crisis in the field of early neural networks coincided with the contemporaneous downfall of behavioristic psychology and philosophy, which was but a milestone in the age old debate between realism and nominalism. Emergence of cybernetics proceeded under the heavy influence of behaviorism (Rosenblueth, Wiener & Bigelow, 1943). Similarly, behaviorism influenced early neural network research in 1950s and 1960s. It pursued nominalistic concept of learning from examples and did not follow the realistic philosophical direction outlined by McCulloch in 1940s. However, behaviorism, as a philosophy, impoverished study of mind and was rejected in 1960s. The downfall of early neural network research is related to its association with the behaviorism and nominalism, a philosophy untenable any longer as a philosophy of mind.

Notwithstanding, today nominalism still forms the basis for many algorithms and neural networks, which do not utilize complicated a priori information in the process of learning and adaptation. Jung has explained the schism between philosophies of realism and nominalism due to the two types of deep seated psychological attitudes. Nominalism and empiricism are related to an extroverted psychological attitude, which is at a premium in our pluralistic society. Thus, it is not a coincidence or chance, that nominalism continue to exert significant influence on scientific concepts in this century despite of the realistic philosophies of the founders of science. However, a concerted research effort toward combining a priori knowledge and learning is emerging. And today, tracing the relationships between philosophical and mathematical theories of the intellect and outlining future research directions, mathematicians move away from Occam, who stands near the roots of scientific objectivization toward the idealistic realism of Plato and Aristotle.

II. Apriority, Adaptivity and Conundrum of Combinatorial Complexity

Mathematical methods of recognition of complex patterns have met with difficulties that are often expressed in terms of the complexity of a recognition process. Various recognition paradigms have their own sets of difficulties, but it seems that there always is a step in the recognition process that is exponentially or combinatorially complex. A well known term used in this regard is "the curse of dimensionality" (Bellman, 1961). This designates a phenomenon of exponential (or combinatorial) increase in the required number of training samples with the increase of the dimensionality of a pattern recognition problem. The curse of dimensionality is characteristic of *adaptive* algorithms and neural networks.

Another set of difficulties is encountered by those approaches to the problem of recognition that utilize systems of *a priori* rules. In the case of rule systems, the difficulty is in a fast (combinatorial) growth of the number of rules with the complexity of the problem (Winston, 1984). Model-based approaches that utilize *a priori* object models in the recognition process encounter difficulties manifested as combinatorial complexity of required computations (Nevatia & Binford, 1977; Brooks, 1983; Grimson & Lozano-Perez, 1984). The difficulties of various mathematical paradigms of intelligence have been summarized in recent reviews as follows. "... Much of our current models and methodologies do not seem to scale out of limited 'toy' domains" (Negahdaripour & Jain, 1991). The key issue is the "combinatorial explosion inherent in the problem" (Grimson & Huttenlocher, 1991).

The seemingly inexorable combinatorial explosion that reincarnates in every paradigm of mathematical intelligence is related in this paper to a fundamental issue of the roles of a priori knowledge vs. adaptive learning. This relationship has been discussed recently for geometric patterns in (Perlovsky, 1994) and for function approximation in (Girosi, Jones & Poggio, 1995). The issue of the roles of a priori knowledge vs. adaptive learning has been of an overriding concern in the research of mathematics of intelligence since its inception. This controversy is here traced throughout the entire history of the concepts of mind throughout the Middle Ages to Aristotle and Plato. The philosophical thoughts of the past turn out to be directly relevant to the development of mathematical concepts of intellect today.

A contemporary direction in the theory of intellect based on modeling neural structures of the brain was founded by McCulloch and his co-workers (McCulloch and Pitts, 1943). In search of a mathematical theory unifying neural and cognitive processes they combined an empirical analysis of biological neurons with the theory of information and mathematically formulated the main properties of neurons. McCulloch believed that the material basis of the mind is in complicated neural structures of a priori origin. Specialized, genetically inherited a priori structures have to provide for specific types of learning and adaptation abilities. An example of such a structure investigated by McCulloch was a group-averaging structure providing for scale-independent recognition of objects, which McCulloch believed serves as a material basis for *concepts* or *ideas* of object independent of their apparent size (Pitts & McCulloch, 1947).

However, this investigation into the a priori aspect of the intellect was not continued during the neural network research in 1950s and 60s and neural networks developed at that time utilized simple structures. These neural networks were based on the concept of general, non-specific adaptive learning using empirical data. By underlining the adaptive aspect of intellect and neglecting its a priori aspect, this approach deviated from the program outlined by McCulloch. Simple structures of early neural networks and learning based entirely on the empirical data were in agreement with behaviorist psychology dominant at the time. When the fundamental, mathematical character of limited capabilities of perceptrons was analyzed by Minsky and Papert (1969), interest in the field of neural networks fell sharply.

Concurrent with early neural networks, adaptive algorithms for pattern recognition have been developed based on statistical techniques and the concept of classification space (Nilsson, 1965; Fukunaga, 1972; Duda and Hart, 1973; Watanabe, 1985). In order to recognize objects (patterns) using these methods, the objects are characterized by a set of classification features that are designed based on a preliminary analysis of a problem and thus contains a priori information needed for a solution of this type of problems. However, general mathematical methods of the design of classification features utilizing a priori information have not been developed. Design of classification features based on a priori

knowledge of specific problems remains an art requiring human participation. When a problem complexity is not reduced to a few classification features by a human analyst, these approaches lead to difficulties related to exorbitant training requirements.

Exorbitant training requirements of statistical pattern recognition algorithms can be understood due to geometry of high-dimensional classification spaces (Perlovsky, 1994). Due to the fact that volume of a classification space grows exponentially with the dimensionality (number of features), training requirements for non-constrained paradigms are exponential in terms of the problem complexity. This is essentially same problem that was encountered earlier in the field of adaptive control and was named "the curse of dimensionality" (Bellman, 1961). The father of cybernetics, Wiener, has also seen this problem, he underlined that using higher order predictive models, or combining many simple models is inadequate for the description of complex non-stationary systems, because of insufficient data for learning (Wiener, 1948).

Facing exorbitant training requirements of statistical pattern recognition algorithms and being dissatisfied with limited capabilities of mathematical methods of modeling neural networks, which existed at the time, Minsky suggested a different concept of artificial intelligence based on the principle of a-priority. He argued that intelligence could only be understood on the basis of extensive systems of a priori rules (Minsky, 1968). This was the next attempt (after McCulloch) to develop the mathematics of intellect from the principle of a-priority. The main advantage of this method is that it requires no training, because it explicitly incorporates detailed, high level a priori knowledge into the decision making. This knowledge is represented in a "symbolic" form similar to high level cognitive concepts utilized by a human in conscious decision making processes, thus the name of this approach: "Symbolic Artificial Intelligence".

The main drawback of this method is the difficulty of combining rule systems with adaptive learning; while modeling the a priori aspect of the intellect, rule systems were lacking in adaptivity. Although, Minsky emphasized that his method does not solve the problem of learning (Minsky, 1975), notwithstanding, attempts to add learning to rule-based artificial intelligence continued in various fields of modeling the mind, including linguistics and pattern recognition (Winston, 1984; Koster & May, 1981; Botha, 1991; Bonnisone et al, 1991; Keshavan et al, 1993). In linguistics, Chomsky has proposed to build a self-learning system that could learn a language similarly to a human, using a symbolic mathematics of rule systems (Chomsky, 1972). In Chomsky's approach, the learning of a language is based on a language faculty, which is a genetically inherited component of the mind, containing an a priori knowledge of language. This direction in linguistics, named the Chomskyan Revolution, was about recognizing the two questions about the intellect: first, how is it possible? and second, how is learning possible? as the center of a linguistic inquiry and of a mathematical theory of mind (Botha, 1991). However, combining adaptive learning with a priori knowledge proved difficult: variabilities and uncertainties in data required more and more detailed rules leading to combinatorial complexity of logical inference (Winston, 1984).

Model-based approaches in machine vision have been used to extend the rule-based concept to 2-D and 3-D sensory data. Use of physically based models permits utilization of detailed a priori information on objects' properties and shape in algorithms of image recognition and understanding (Nevatia & Binford, 1977; Brooks, 1983; Winston, 1984; Grimson & Lozano-Perez, 1984; Chen & Dyer, 1986; Michalski et al, 1986; Lamdan & Wolfson, 1988; Negahdaripour & Jain, 1991; Bonnisone et al, 1991; Segre, 1992; Keshavan et al, 1993; Califano & Mohan, 1994). Models used in machine vision typically are complicated geometrical 3-D models that require no adaptation. These models are useful in applications where variabilities are limited and types of objects and other

parameters of the recognition problem are constrained. When unforeseen variabilities are a constant factor in the recognition problem, utilization of such models faces difficulties that are common to rule-based systems. More and more detailed models are required, potentially leading to a combinatorial explosion.

Parametric model-based approaches have been proposed to overcome the difficulties of previously used methods and to combine the adaptivity of parameters with a-priority of models. In these approaches adaptive parameters are used to adapt models to variabilities and uncertainties in data. Parametric adaptive methods date back to Widrow's Adaline (1959) and linear classifiers. These early parametric methods can be efficiently trained using few samples, however, they are limited to simple decision regions and are not suitable for many complicated problems. Complicated problems, such as routinely solved by human perception mechanisms, require utilization of multiple flexible models. In the process of recognition, an algorithm has to decide which subset of data corresponds to which model. This step is called segmentation, or association, and it requires a consideration of *multiple combinations* of subsets of the data. Because of this, complicated adaptive models often lead to combinatorial explosion of the complexity of the recognition process.

Fifty years of experience with classical mathematical concepts of intelligence led to three important conclusions. First, intelligent algorithms have to combine learning and adaptivity with complicated a priori structures, second, they should utilize complicated internal models learned on the basis of a priori structures, and third, all classical approaches to this problem led to combinatorial complexity. A mathematical analysis leads to the conclusion that the specific types of combinatorial complexity are closely related to the roles of apriority and adaptivity (Perlovsky; 1994; 1996b,f; 1997a,c). While methods based on adaptivity face combinatorial explosion of *the training process*, those based on a-priority face combinatorial explosion of *the complexity of rule systems*, and attempts to combine the two face combinatorial explosion of *the computational complexity*. Existing approaches to this problem has not resolved the conundrum of combinatorial complexity. To repeat again, "... Much of our current models and methodologies do not seem to scale out of limited 'toy' domains" (Negahdaripour & Jain, 1991); "The key issues (is)... "combinatorial explosion inherent in the problem" (Grimson & Huttenlocher, 1991).

III. Aristotelian Contradiction, Gödel, and Zadeh

Tracing metaphysical origins of our mathematical concepts of intellect is helpful for understanding the dynamics of changing scientific paradigms. In particular, two concepts due to Aristotle were examined (Perlovsky, 1996a,c,d). One is the Aristotelian logic conceived to describe eternal truths. Another is the Aristotelian theory of mind describing adaptive, changeable Forms. The mathematical difficulties we are facing today were traced to a contradiction in the Aristotelian treatment of these concepts. This contradiction is related to the Aristotelian disagreement with Plato, and to the Aristotelian rejection of Plato's Ideas for the new concept of Form. "Symbolic AI" utilized internal structures based on Aristotelian logic similar to the Plato's Ideas. Similarity between logical rule systems and the Plato's conception of mind has been discussed by Chomsky (1972). He has directly related the principle of a-priority in algorithm design to the philosophy of Plato. He has also hoped that the problem of learning can be solved using rule-based approach to intelligence. As discussed in Section 1, this approach faced combinatorial computational complexity. The combinatorial explosion has been related to Gödelian theorems, which revealed the combinatorial nature of Aristotelian logic (Perlovsky, 1996g;1997c;1998).

The most striking fact is that the first one who pointed out that learning can not be achieved in Plato's theory of mind was Aristotle. Aristotle recognized that in Plato's formulation there could be no learning, since Ideas (or concepts) are given a priori in their final forms

of eternal unchangeable truths. Thus, learning is not needed and is impossible, and the world of ideas is completely separated from the world of experience. Searching to unite the two worlds and to understand learning, Aristotle developed a concept of Form having, on the one hand, a universal and higher a priori reality like Plato's Ideas, but on the other, being a formative principle in an individual experience (Metaphysics). Forms can exist as potentialities and as actualities. In Aristotelian theory of Form, the adaptivity of the mind was due to a meeting between the a priori Form and matter, forming an individual experience. The major point of Aristotelian theory departure from Plato's Ideas was that before a Form meets matter it exists as a potentiality, thus, it has to be not in its final form of a concept; it becomes a concept in the process of experience. In the process of learning a Form-as-potentiality evolves into a Form-as-actuality that is the crisp concept of logic. This theory was further developed by Avicenna (XI AD), Maimonides (1190), Aquinas (XIII), and Kant (1781) among many other philosophers during the last 2300 years.

But, Aristotelian logic is unsuitable for describing Forms, because Aristotelian logic deals explicitly with the eternal truths in their final crisp forms of concepts. For example, consider a "law of excluded third", which is a central law of Aristotelian logic. According to the law of excluded third, every statement is either true or not true, and there is no third alternative. It might be applicable to eternally valid truths, but it is not applicable to our everyday intelligence, nor to fluid and adaptable Aristotelian Forms describing the process of learning. Since Aristotelian logic is a foundation of most of our algorithms including the logic of propositions and "Symbolic AI", the difficulties and contradictions of "Symbolic AI" are traced to Aristotle. Fuzzy logic is needed for Aristotelian theory of Form — theory of mind. Thus, the 2300 year old contradiction between theory of mind and logic is being resolved with fuzzy logic (Perlovsky, 1996e;1997a).

For 2000 years philosophers-realists, followers of Plato and Aristotle, analyzed ontological differences between Plato's Ideas and Aristotelian Forms, but the principled epistemological difference was not noticed. Ontology refers to existence: while Plato assumed that Ideas exist in a separate world, Aristotle considered Forms as existing in our mind. Epistemology refers to the ways in which knowledge is acquired: in Plato's theory, Ideas are unchangeable eternal truths, while Aristotelian Forms are dynamic entities. Only, when scientists have applied the Aristotelian logic to mathematical modeling of mind, the contradiction between Aristotelian logic and theory of mind led to difficulties, contradictions, and impasse, which is being resolved today in shifting scientific paradigms. Analyzing the original contradiction will help us to understand the future directions in the research of mind, both, mathematical and metaphysical.

Plato-Aristotelian conception of mind based on a priori structures was further developed by Kant. As well known, Kant identified a priori inner models as a separate faculty of mind that he called Understanding. Mind's operations with a priori concepts Kant calls the domain of Pure Reason (1781). The main question that the analysis of Pure Reason shall answer, according to Kant, is "How are *synthetic judgments a priori* possible?" In the mathematical theory of mind, this specific faculty is represented in hierarchical models: next levels in a hierarchy contain synthesis of the lower level concepts. Thus, development of a priori hierarchical models is a key to mathematical modeling of the Understanding or Pure Reason. Making this hierarchy fuzzy, adaptable and situationally dependent to enable learning is the future challenge.

References

Albus, J. (1990). *Theory of Intelligent Systems*. 5th IEEE Symp. on Intelligent Control. Philadelphia, PA.

- Aristotle. (IV BC). *Metaphysics*. Trans. H.G.Apostle, 1966, Indiana University Press, Bloomington, IN.
- Avicenna. (XI AD). *Kitab al-Shifa*. Tr. in Avicenna, S. Afnan, G. Allen & Unwin, LTD, 1958, London, GB.
- Aquinas, T. (XIII) *Summa contra Gentiles*. Translated by the English Dominican Fathers, 1924, London.
- Bellman, R.E. (1961). *Adaptive Control Processes*. Princeton University Press, Princeton, NJ
- Bonnisone, P.P., Henrion, M., Kanal, L.N., & Lemmer, J.F. (1991). *Uncertainty in Artificial Intelligence 6*. Nort Holland, Amsterdam, The Netherlands.
- Botha, R. P. (1991). *Challenging Chomsky. The Generative Garden Game*. Basil Blackwell, Oxford, UK.
- Brooks, R.A. (1983). Model-based three-dimensional interpretation of two-dimensional images. *IEEE Trans. Pattern Anal. Machine Intell.*, 5 (2), 140-150.
- Califano A. & Mohan, R. (1994). Multidimensional indexing for recognizing visual shapes. *IEEE Trans. Pattern Anal. Machine Intell.*, 16 (4), 373-392.
- Chen, R.T. & Dyer, C.R. (1986). Model-based recognition in robotic vision. *ACM Computing Surveys*, 18, pp.67-108.
- Chomsky, N. (1972). *Language and Mind*. Harcourt Brace Javanovich, New York, NY.
- Chomsky, N. (1981). *Principles and Parameters in Syntactic Theory*. In N.Hornstein and D.Lightfoot (eds), *Explanation in Linguistics. The Logical Problem of Language Acquisition*, Longman, London.
- Duda, R.O. & Hart, P.E. (1973). *Pattern Classification and Scene Analysis*. Wiley & Sons, New York, NY.
- Fukunaga, K. (1972). *Introduction to Statistical Pattern Recognition*. Academic Press. New York, NY.
- Girosi, F., Jones, M. & Poggio, T. (1995). Regularization theory and neural networks architectures. *Neural Computation*, 7 (2), pp.219-269.
- Grimson, W.E.L. & Lozano-Perez, T. (1984). Model-based recognition and localization from sparse range or tactile data. *Int. J. Robotics Research*, 3 (3), pp.3-35
- Grimson, W.E.L. & Huttenlocher, D.P. (1991). Introduction to the special issue on interpretation of 3-D scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13 (10), pp.969-970; 14 (2).
- Grossberg, S. (1980). *How Does a Brain Build a Cognitive Code?* Psychological Review 87, pp.1-51.
- Grossberg, S. (1988). *Nonlinear Neural Networks*. Neural Networks, 1 (1), pp.17-61.
- Jaynes (1976). *The Origin of Consciousness in the Breakdown of the Bicameral Mind*. Houghton Mifflin Co., Boston, MA.

- Jung, C.G. (1951). *Aion, Researches into the Phenomenology of the Self*. In the Collected Works, v.9 part II, Bollingen Series XX, 1969, Princeton University Press, Princeton, NJ.
- Kant, I. (1781). *Critique of Pure Reason*. Tr. J.M.D. Meiklejohn, 1943. Willey Book, New York, NY.
- Keshavan, H.R., Barnett, J., Geiger, D. & Verma, T. (1993). *Introduction to the Special Section on Probabilistic Reasoning*. IEEE Trans. PAMI, 15(3), pp. 193-195.
- Koster, J. & May, R. (1981). *Levels of Syntactic Representation*. Foris Publications, Dordrecht.
- Lamdan, Y. & Wolfson, H.J. (1988). Geometric hashing: a general and efficient recognition scheme. *Proc. 2nd Int. Conf. Computer Vision*.
- Maimonides, M. (1190). *The Guide for the Perplexed*. Transl. M. Friedlander, 2nd ed., 1956, Dover, New York, NY.
- McCulloch, W. and Pitts, W. (1943). A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics*, 7, pp.115-133.
- McCulloch, W.S. (1961). *What Is a Number that a Man May Know It, and a Man, that He May Know a Number?* The 9th Alfred Korzybski Memorial Lecture. *General Semantics Bulletin*, 26, 27, pp.17,18.
- McCulloch, W.S. (1965). *Embodiments of Mind*. MIT Press, 2nd edition, Cambridge, MA, 1988.
- Meystel, A., (1996). Intelligent Systems: A Semiotic Perspective. *Int. Journ. Intelligent Control and Systems*, 1 (1), pp.31-58.
- Michalski, R.S., Carbonell, J.G., Mitchell, T.M. (1986). *Machine Learning*. Kaufmann, Los Altos, CA.
- Minsky, M.L. (1968). *Semantic Information Processing*. The MIT Press, Cambridge, MA.
- Minsky, M.L. (1975). *A Framework for Representing Knowledge*. In *The Psychology of Computer Vision*, ed. P. H. Winston, McGraw-Hill Book, New York.
- Morris, C.W. (1938). *Foundation of the Theory of Signs*. In *Morris Writings on the general theory of signs*, ed. T.A. Sebeok, 1971, Mouton, Hague.
- Negahdaripour, S. & Jain, A.K. (1991). Final Report of the NSF Workshop on the Challenges in Computer Vision Research; *Future Directions of Research*. National Science Foundation.
- Nevatia R. & Binford, T.O. (1977). Description and recognition of objects. *Artificial Intelligence*, 8 (1), pp.77-98
- Nilsson, N.J. (1965). *Learning Machines*. McGraw-Hill, New York, NY.
- Occam, W. (XIV). *Summa logicae*. Transl. M.J.Loux, Occam's Theory of Terms, 1974, and A.J.Freddoso and H.Schuurman, Occam's Theory of Propositions, 1980, University of Notre Dame Press, Notre Dame, IN.

- Perlovsky, L.I. (1994). Computational concepts in classification. *J.Math.Imaging and Vision*, 4 (1).
- Perlovsky, L.I. (1996a). *Fuzzy Logic of Aristotelian Forms*. Proc. Conference on Intelligent Systems and Semiotics '96. Gaithersburg, MD, v.1, pp. 43-48.
- Perlovsky, L.I. (1996b). *Mathematical Concepts of Intellect*. Proceedings of World Congress on Neural Networks, San Diego, CA; Lawrence Erlbaum Associates, NJ, pp. 1013-1016.
- Perlovsky, L.I. (1996c). *Aristotle, Complexity and Fuzzy Logic*. Proceedings of World Congress on Neural Networks, San Diego, CA; Lawrence Erlbaum Associates, NJ, p. 1152.
- Perlovsky, L.I. (1996d). *Complexity of Recognition: Aristotle, Göedel, Zadeh*. IFAC Triennial World Congress. San Francisco, CA.
- Perlovsky, L.I. (1996e). *Fuzzy Logic of Aristotelian Forms*. Proceedings of the Conference on Intelligent Systems and Semiotics '96. Gaithersburg, MD, v.1, pp. 43-48.
- Perlovsky, L.I. (1996f). *Intelligence of Recognition. Introduction*. Intelligent Systems and Semiotics '96. Gaithersburg, MD.
- Perlovsky, L.I. (1996g). *Gödel Theorem and Semiotics*. Proceedings of the Conference on Intelligent Systems and Semiotics '96. Gaithersburg, MD, v.2, pp. 14-18.
- Perlovsky, L.I. (1997a). *Mathematical Aspects of Cyberaesthetics*. Proceedings of the Conference on Intelligent Systems and Semiotics '97. Gaithersburg, MD, pp. 319-324.
- Perlovsky, L.I. (1997b). *Towards Quantum Field Theory of Symbol*. Proceedings of the Conference on Intelligent Systems and Semiotics '97. Gaithersburg, MD, pp. 295-300.
- Perlovsky, L.I. (1997c). *Modeling Fields, Evolutionary Computations, and Hierarchies*. Conference on Intelligent Systems and Semiotics '97. Gaithersburg, MD.
- Perlovsky, L.I. (1998). *Physics of Mind*. Oxford University Press.
- Pitts, W. & McCulloch, W.S. (1947). How we know universals: the perception of auditory and visual forms. *Bulletin of Mathematical Biophysics*, 9, pp.127-147.
- Plato. (IV BC). *Phaedrus*. Translated in Plato, L. Cooper, Oxford University Press, New York, NY.
- Pribram, K. (1971). *Languages of the Brain*. Prentice Hall. New York, NY.
- Rosenblueth, A., Wiener, N., & Bigelow, J. (1943). *Philosophy of Science*, 10 (1), pp. 18-24.
- Segre, A.M. (1992). *Applications of Machine Learning*. IEEE Expert, 7 (3), pp. 31-34.
- Skinner, B.F. (1974). *About Behaviorism*. Alfred A. Knopf. New York, NY.
- Watanabe, S. (1985). *Pattern Recognition: Human and Mechanical*. John Wiley & Sons, New York, NY.
- Watson, J.B. (1913). *Psychology as the Behaviorist Views It*. Psychological Review, 20, pp.158-177

Widrow, B. (1959). 1959 WESCON Convention Record, Part 4, 74-85.

Wiener, N. (1948). *Cybernetics*. Wiley, New York, NY.

Winston, P.H. (1984). *Artificial Intelligence*. 2nd edition. Addison-Wesley. Reading, MA.

Zadeh, L.A. (1965). Fuzzy Sets. *Information and Control*, 8, pp.338-352.