2014 International Joint Conference on Neural Networks (IJCNN)
July 6-11, 2014, Beijing, China

# Cognitive Neural Network for Cybersecurity

Leonid Perlovsky, and Olexander Shevchenko

*Abstract*—**This chapter discusses the future of cybersecurity as warfare between machine learning techniques of attackers and defenders. As attackers will learn to evolve new camouflaging methods for evading better and better defenses, defense techniques will in turn learn new attacker's tricks to defend against. The better technology will win. Here we discuss theory of machine learning based on dynamic logic that are mathematically provable to learn with the fastest possible speed. We also discuss cognitive functions of dynamic logic and related experimental proofs. This new mathematical theory, in addition to being provably fastest machine learning technique, is also an adequate model for several fundamental mechanisms of the mind.**

## I. INTRODUCTION

TODAY'S networks and their users are under attack from an ever-expanding universe of threats and malware. Malware are malicious software codes that typically damages or disables, takes control of, or steals information from a computer system. Malware broadly includes botnets, viruses, worms, Trojan horses, logic bombs, rootkits, boot kits, backdoors, spyware, adware, and other types of threats. The ever increasing danger of the future threat is its ability to evolve for avoiding system defenses. Future threats will be using machine learning to outsmart the defenses. Therefore the future of cybersecurity if a warfare of machine learning techniques. The more capable machine learning technique will win.

Correspondingly, an important direction of cybersecurity concentrates on machine learning techniques (Blowers and Williams, 2014; Dua and Du, 2011; Gesher, 2013; Mugan, 2013). Shabtai et al, 2012). In this Chapter we discuss machine learning techniques based on dynamic logic (DL), which can be mathematically proven to have the fastest possible learning ability (Perlovsky, Deming, & Ilin, 2011). Steps toward developing such cybersecurity methods are discussed below.

This chapter describes an adaptive machine learning techniques based on abstract models. The approach to detecting novel attacks is anomaly detection: we develop algorithms learning models of attack-free traffic, and then detect deviations identifying malware. Gradual learning is a fundamental aspect of this approach. We begin assuming that an adequate protection system exists, and we can learn characteristics-models of attack-free traffic. The developed algorithms learn evolution of the malware as it attempts to hide its harmful nature. For the success of this approach, learning of the defensive system must be faster than evolution of the threat.

The defensive system learns to recognize threats as combinations of basic elements, words or n-grams. In principle, this is a most general and universal approach, potentially capable of recognizing any threat. The difficulty of realizing this universal potential is computational complexity and slow learning of most existing algorithms. The reason for these difficulties is fundamental: the number of combinations is very large, even relatively few n-grams can be used to form a very large number of combinations. The number of combinations of only 100 n-grams is $100^{100}$, this number exceeds all interactions of all elementary particles in the Universe during its entire lifetime. Therefore even if the entire Universe could be made to learn combinations of n-grams, it will not be able to perform its job fast enough. Later we will relate this fundamental difficulty to Gödelian difficulties of logic.

We describe DL, a mathematical technique overcoming Gödelian limitations. DL has been used to overcome combinatorial complexity (CC), a difficulty that for 50 years has prevented classical pattern recognition and artificial intelligence to solve many complex problems, such as detection of patterns below noise and among unrelated signals (clutter). The developed DL algorithms have overcome CC and improved detection performance by orders of magnitude. After introducing DL we discuss CC specific to cybersecurity. This CC is related to learning structures of threat models. The past research in dynamic logic can be understood as developing continuous mathematical representation of associations between signals and models. The current overcoming of CC of learning cyber-security models requires continuous mathematical representation of model structures. Model structure is a combination of inherently discrete mathematical constructs; representing it continuously is equivalent to eliminating a difference between continuous and discrete mathematics in this wide field. This will overcome Gödelian limitations of classical logic in this field.

This requires the new mathematical method described here. We outline an approach to proving the fastest possible learning of DL. We illustrate the new DL technique of machine learning using an abstract simulated data set, and finally, we demonstrate DL using publicly available malware data bases.

In addition to being a provably fastest machine learning technique, DL is also an adequate model for several fundamental mechanisms of the mind. We briefly discuss theoretical foundations for cognitive-emotional functions of DL and their experimental proofs in several brain-imaging labs. The combination of mathematical and cognitive

L. Perlovsky is with Harvard University, Athinoula Martinos Imaging Center, Charlestown, USA, tel: 617-259-0197; e-mail: lperl@ rcn.com.
O. Shevchenko is with OleSoft, Denver, CO 80012, USA (e-mail: sashko.shevchenko@gmail.com).

978-1-4799-1484-5/14/$31.00 ©2014 IEEE

superiority promises reliable future cybersecurity.

## II. CC AND THE GÖDELIAN PROBLEM IN LOGIC

Developing machine learning techniques exceeding human learning abilities started in the 1950s, when computers become available. Since then, hundreds and thousands learning algorithms have been developed. Many look like they can solve a variety of problems, but practically they have been limited to "toy" problems and could not be generalized to different or more complex problems. Gradually it became accepted that all previously used paradigms faced the difficulty of CC.

CC as a general problem encountered by all paradigms attempting to model the mind and to create equally capable machine learning techniques have been discussed in (Perlovsky, 1998). This publication analyzed all major paradigms of machine learning and for each identified a reason for CC. In parallel a fundamental reason for CC has been identified (Perlovsky, 1996, 2001a). Using today's understanding (Perlovsky, Deming, & Ilin, 2011; Perlovsky, 2013) it can be formulated succinctly: CC is related to logic; it is a manifestation of difficulties of logic discovered by Gödel in the 1930s (Gödel, 2001).

As discussed in given references, the argument summarized above closely follows Gödelian arguments that lead to fundamental difficulties in logic. Gödel (2001) considered all logical statements, including potentially infinite ones. And he demonstrated that although he logically listed all logical statements, he was able to prove that there have to be logical statements not in the list. This Gödelian argument can be related to our argument above: the Gödelian list included combinations of the original elements forming logical statements. And although the number of original elements was infinite, the number of combinations turns out to be a "significantly larger" infinity. Whereas the original infinity was countable, the number of combinations turn out to be uncountable infinity. To prove a fundamental difficulty in logic, Gödelian argument has to be applied to an infinite system, such as logic. Applying the Gödelian argument to a finite system, such as computer, does not result in a "fundamentally irresolvable" difficulty. Instead it results in CC, a "practically fundamental" difficulty. Thus for the purpose of designing efficient machine learning algorithm, CC is as fundamental as Gödelian difficulty in logic.

Understanding of this fundamental reason for CC is essential for making progress in machine learning. Thousands of algorithms have been designed for machine learning since the 1950s. These attempts still continue. The argument above demonstrates that unless fundamental reliance on logic during algorithm design is avoided, CC will persist. But what does it mean to avoid logic? Isn't entire science based on logic? Should the entire science be abolished? What could be used instead? The mind solves problems that computers cannot solve. Young kids and even birds solve problems that computers cannot. Can we understand how minds do this? Can this understanding be scientifically formulated?

## III. DYNAMIC LOGIC

For thousands of years logic has been the best way to conduct arguments, including scientific arguments. The Newtonian physics, quantum physics, theory of relativity are based on logic. Only recently, when facing problems related to working of the mind we encountered insufficiency of logic. It is interesting that Aristotle, the founder of logic, did not use logic when explaining working of the mind. To explain mind, Aristotle developed theory of forms. Aristotelian forms are different from Platonian ideas and from contemporary understanding of concepts of the mind. Ideas of Plato and concepts in contemporary psychology are static entities, similar to logical statements, such as: "this is a chair." Instead, Aristotelian forms are dynamic entities, processes in which "mind meets matter." Today we describe it as an interaction between top-down and bottom-up neural signals. Before Aristotelian forms meet matter they exist as potentialities; in interacting with matter they become actualities (Aristotle, 1995).

What this process-logic means mathematically, in which way is it fundamentally different from usual classical logic? And why after Gödel has proven that logic has fundamental irresolvable difficulties in the 1930s, computer scientists still attempt to develop machine learning using logic? Did not neural networks and fuzzy logic attempt to overcome difficulties of logic?

An original mathematical theory of dynamic logic, theory that closely follows a process-logic of Aristotelian forms has been presented in (Perlovsky, 2001a, 2006a).

## IV. CONTINUOUS REPRESENTATION OF MODEL STRUCTURE AND LEARNING MALWARE CODES

Applying DL to learning models of objects and events in Internet network requires overcoming CC of a more complex nature than in the above references. In these references associations between signals and models have been transformed into continuous representations, which makes possible the DL processes avoiding combinatorial complexity. For Internet models we face a requirement to represent continuously structures of these models. Signals in networks are moved in packets, and each packet contain a large number of symbols, words, or n-grams. Contents of most of these n-grams are benign. Very few are a part of a malware message. Usually the dangerous or destructive nature of an n-gram cannot be determined from a single n-gram. Several of them have to be assembled into a message before their dangerous content can be determined. This requires sorting through a huge number of benign n-grams and messages before a specific structure can be identified. Structural constituents of a model are considered inherently discrete elements. This view based on classical logical analysis of a model leads to considering and evaluating combinations of model elements, and therefore to CC in model learning. These combinations are of an entirely different nature than combinations of signals and models in the previous example, and the previously developed mathematical approach is not applicable here.

Below we describe a further development of DL, which turns identifying of a model structure into a continuous problem (Perlovsky, Deming, Ilin, 2011). Instead of the logical consideration of a model as consisting of its elements, so that every signal or n-gram in the network either belongs to a model or does not, DL considers every n-gram as potentially belonging to a model. Starting from a vague potentiality of a model, to which every n-gram could belong, the DL learning process evolves this into a model-actuality containing definite n-grams, and not containing others.

We denote n-grams in the network as x(n,j), n=1,...N enumerates messages, j=1,...J enumerates n-grams, and m=1,...M enumerates models. Model parameters, in addition to r(m) are p(m,j), potentialities of n-gram j belonging to model m. Data x(n,j) have values 0 or 1; potentialities p(m,j) start with vague value near 0.5 and in the DL process of learning they converge to 0 or 1. Mathematically this construct can be described as

$$\ell(n|m) = \prod_{j=1}^{J} p(m,j)^{x(n,j)}(1 - p(m,j))^{(1- x(n,j))} \quad (1)$$

A model parameter p(m,j), modeling a potentiality of n-grams j being part of model m, starts the DL process with initial value near 0.5 (exact values 0.5 for all p(m,j) would be a stationary point of the DL process). Value p(m,j) near 0.5 gives potentiality values of x(n,j) with a maximum near 0.5, in other words, every n-gram has a significant chance to belong to every model. If p(m,j) converge to 0 or 1 values, these would describe which n-grams j belong to which models m. The DL process estimating parameters p(m,j) is given by:

$$f(m|n) = r(m)\,\ell(n|m) \,/ \sum_{m'\in M} r(m')\,\ell(n|m'). \quad (3)$$

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

$$*dp(m',j)/dt, \;\; \delta_{mm'} = 1 \text{ if } m=m', 0 \text{ otherwise}$$

$$dp(m,j)/dt = \sum_{n\in N} f(m|n)[\partial \ln l(n|m)/\partial p(m,j)] \,. \quad (2)$$

Parameter t here is an internal time of the DL process; in digital computer implementation it is proportional to an iteration number. Functions f(m|n) associate n-grams and models, they can be interpreted as estimated probabilities.

The data used for testing this DL algorithm and the results of the analysis are shown in Fig. 1. We simulated 16,000 messages shown on the left. They are arranged in their sequential order along the horizontal axis, n. For this simplified example we simulated 1,000 total number of possible n-grams in the network, they are shown along the vertical axis, j. Every message has or does not have a particular n-gram as shown by a white or black dot at the location (n,j). This figure looks like random noise corresponding to pseudo- random content of messages. On the right figure, messages are sorted so that messages having similar n-grams appear next to each other. These similar n-grams appear as white horizontal strikes and reveal several groups. Most of message contents are pseudo-random n-grams; about a half of messages have several similar n-grams. These messages with several specific n-gram values have specific contents, they belong to certain models, and could be malware codes.

Since the data for this example have been simulated, we know the true number of various groups, and the identity of each message as belonging to a particular groups. All messages have been assigned correctly to its group without a single error. Convergence is very fast and took two to four iterations (or steps) to solve eqs.(2).

This algorithm have been applied to a publicly available data set of malware codes, KDD (Dua and Du, 2011; Gesher, 2013; Mugan, 2013). This data set originated from 1998 DARPA Intrusion Detection Evaluation; under the sponsorship of DARPA and the Air Force Research Laboratory, MIT Lincoln Labs has collected and distributed the datasets for the evaluation of computer network intrusion detection system. This data set includes 41 features extracted from Internet packets and one class attribute enumerating 21 classes of four types of attacks. Our algorithm identified all classes of malware and all malware messages without a single false alarm. This performance in terms of accuracy and speed is better than other published algorithms. An example of the algorithm performance on this data in give in Fig. 2.

The machine learning based on DL achieves the fastest possible learning. This is a consequence of DL performing the maximum likelihood model estimation. This is known to lead to algorithms reaching the Cramer-Rao Bound, the information-theoretic bound on speed of learning (Perlovsky 2001; Perlovsky, Deming, & Ilin, 2011). Thus in the future battles of machine learning technologies between cyber-security threats and defenses, DL offers a mathematically-provable technique with the fastest adaptive capability.

## V. COGNITIVE AND EMOTIONAL FUNCTIONS OF DL IN THE MIND

In addition to being a mathematical breakthrough in several areas of Machine Learning, DL is a cognitive mathematical theory, a basis for a number of cognitive algorithms. Calling DL a cognitive mathematical theory we mean that it mathematically models several functions of the

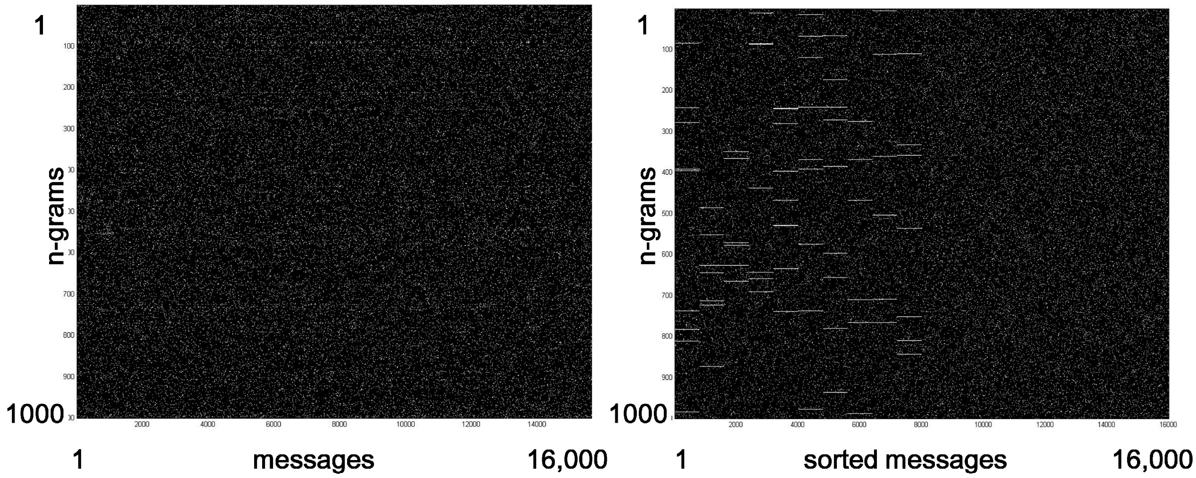


Fig.2. On the left are 16,000 messages arranged in their sequential order along the horizontal axis, n. The total number of possible n-grams in the network is 1000, they are shown along the vertical axis, j. Every message has or does not have a particular n-gram as shown by a white or black dot at the location (n,j). This figure looks like random noise corresponding to pseudo- random content of messages. On the right figure, messages are sorted so that messages having similar n-grams appear next to each other. These similar n-grams appear as white horizontal stricks. Most of message contents are pseudo-random n-grams; about a half of messages have several similar n-grams. These messages with several specific n-gram values have specific contents, they belong to certain models, and could be malware codes.


Fig.3. This figure shows sorted messages from 6 groups (1 normal and 5 malware) of the KDD data set corresponding to Fig.2 right. (We do not show unsorted messages, similar to Fig.2 left). Along the horizontal axis: 67343 normal, 2931 portsweep, 890 warezclient, 3633 satan, 41214 neptune, 3599 ipsweep. Groups are significantly different in size, and not all features important for a group are necessarily present in each vector belonging to the group, therefore the look of the figure is not as clear cut as Fig.2 right. Nevertheless all vectors are classified without errors, and without false alarms.

mind-brain. Some of these functions are well appreciated, others have seemed mysterious. The DL models made a number of experimentally testable predictions, some nontrivial and unexpected, some confirmed experimentally, none have been disconfirmed. This section briefly summarizes this cognitive aspect of DL.

The first salient and unexpected prediction of dynamic logic is the process from vague to crisp as a foundation of perception and cognition. A simplified experiment confirming this prediction can be conducted by anyone in 1/2 a minute. Concentrate on an object in front of your eyes, then close the eyes and imagine the object. Imagination is usually not as clear and crisp as perception with opened eyes. It is known that imagination is produced by top-down neural signals from representations of objects stored in memory (Kosslyn, 1994). Therefore, representations are not as crisp and clear as perceptions with open eyes, representations are vague. The more abstract are imagined ideas, the vaguer are representations. This experiment have been performed using brain imaging (Bar et al, 1996; Perlovsky, 2009c). This experiment confirmed that representations are vague, and less conscious than perceptions with opened eyes. Predictions about vaguer nature of cognitive representations have been confirmed in (Kveraga et al, 2007).

According to the theory of instincts and emotions (Grossberg & Levine, 1987), instincts are sensor-like neural mechanisms in the mind, which measure vital bodily parameters and indicate their safe ranges to decision-making parts of the brain. The neural signals connecting instinctual and decision-making parts of the brain are emotional signals indicating satisfaction of instinctual needs. DL extended this theory toward the knowledge instinct, which measures similarity between mental models-representations and patterns in sensor signals, eq.(1) (Perlovsky, 1987, 2001a; Perlovsky & McManus, 1991). Emotional signals measuring satisfaction of the knowledge instinct are aesthetic emotions, serving as a foundation for all human higher mental abilities, including abilities for the beautiful (Perlovsky, 2000, 2001a,b, 2010a). Existence of this specific aesthetic emotions related to knowledge have been first postulated by Kant (1790) and experimentally confirmed in (Perlovsky, Bonniot-Cabanac, & Cabanac, 2010).

DL has led to a theory of interaction between language and cognition (Perlovsky, 2004, 2007, 2009a,b, 2010b,c, 2013d; Fontanari & Perlovsky, 2007, 2008a; Tikhanoff et al, 2006; Perlovsky & Ilin, 2010). This theory explains why children can talk without full understanding, why language is acquired earlier than cognition, several other mysteries of language and cognition, it predicts that language and cognition are closely connected but separate brain functions, that abstract concepts are understood mostly due to language, without full cognitive understanding. Some of these predictions have been experimentally confirmed (Binder et al, 2005; Price, 2012). DL has led to a theory of language emotionality (Perlovsky 2009b, 2012c) and to a theory explaining the cognitive function, origin, and evolution of musical emotions that Darwin (1871) called "the greatest mystery" (Perlovsky, 2006c, 2008, 2010b,d, 2012a,b, 2013a,c). Predictions of this theory have been experimentally confirmed in (Cabanac et al, 2013; Masataka et al, 2012a,b, 2013; Perlovsky et al, 2013).

To summarize, the mathematics of DL has overcome combinatorial complexity and achieved the maximum likelihood model estimation and malware detection. This results in fast learning at the information-theoretic bound on the speed of learning. DL ability to model cognitive-emotional mechanisms opens new perspectives for cyber-security discussed below.

## VI. FUTURE RESEARCH

The next step of applying DL to cyber security will include learning syntactic and semantic aspects of models. Fast learning of these complicated models is necessary to counter advanced threats, including evolving malware using stealthy, mutating, self-camouflaging, and "Frankensteinian" technologies (Cisco, 2013). These types of threat are capable to evolve and mutate for avoiding existing anti-malware technology, operate stealthy, and assemble itself from parts of other codes (so that no "local" syntax-based detection is possible). We repeat: future malware codes will utilize machine learning technology, and countering these cyber-threats will be only possible by using a superior machine learning technology. Future cyber-security will be a battle of machine learning technologies. Here we described a step toward the machine learning technology that can be mathematically proven to reach information theoretic bounds on speed of learning (Perlovsky et al, 2011; Perlovsky, 2013).

Cognitive and emotional aspects of DL machine learning will be used for a different type of cyber-security than the one at the focus of this paper. It will be possible to analyze cognitive and emotional contents of network traffic, identify perpetrators and their intents, and instead of countering cyber-threats, attack the perpetrators of cyber-attacks.

## REFERENCES


[1] Aristotle, "The complete works. The revised Oxford translation," ed. J. Barnes, Princeton, NJ: Princeton Univ. Press. Original work IV BCE, 1995.

[2] Bar, M., Kassam, K.S., Ghuman, A.S., Boshyan, J., Schmid, A.M., Dale, et al., "Top-down facilitation of visual recognition," USA: Proceedings of the National Academy of Sciences, 103, 449-54, 2006.

[3] Binder, J.R., Westbury, C.F., McKiernan, K.A., Possing, E.T., & Medler, D.A., "Distinct Brain Systems for Processing Concrete and Abstract Concepts," Journal of Cognitive Neuroscience 17(6), 1–13, 2005.

[4] Blowers, M. and Williams, J. (2014). Machine Learning Applied to Cyber Operations. In Pino, R.E. (ed.) Network Science and Cybersecurity, Springer, New York, NY.

[5] Cabanac, A., Perlovsky, L.I., Bonniot-Cabanac, M-C., Cabanac, M. (2013). Music and Academic Performance. Behavioural Brain Research, 256, 257-260.

[6] Cisco. (2013). Annual Security Report; https://www.cisco.com/web/offer/gist_ty2_asset/Cisco_2013_ASR.pdf

[7] Dua, S. & Du, X. (2011). Data Mining and Machine Learning in Cybersecurity. Taylor & Francis, Boca Raton, FL.

[8] Fontanari, J.F. and Perlovsky, L.I. (2007). Evolving Compositionality in Evolutionary Language Games. IEEE Transactions on Evolutionary Computations, 11(6), pp. 758-769; doi:10.1109/TEVC.2007.892763

[9] Fontanari, J.F. & Perlovsky, L.I. (2008a). How language can help discrimination in the Neural Modeling Fields framework. Neural Networks, 21(2-3), 250–256.

[10] Gesher, A. (2013). Adaptive adversaries: building systems to fight fraud and cyber intruders. In Proceeding of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, Pages 1136-1136, ACM New York, NY, US.

[11] Gödel, K. (2001). Collected Works, Volume I, Publications 1929–1936. Feferman, S., Dawson, J.W., Jr., Kleene, S.C., Eds.; Oxford University Press: New York, NY.

[12] Grossberg, S. & Levine, D.S. (1987). Neural dynamics of attentionally modulated Pavlovian conditioning: blocking, inter-stimulus interval, and secondary reinforcement. Psychobiology, 15(3), pp.195-240.

[13] Jones, J. Bradstreet, J., Kozak, M., Hughes, T., & Blount, M. (2004). Ground moving target tracking and exploitation performance measures. Pentagon Report A269234, approved for public release.

[14] Kant, I. (1790/1914). Critique of Judgment, tr. J.H.Bernard, London: Macmillan & Co. Kant, 1798/1974.

[15] Kosslyn, S.M. (1994). Image and Brain. Cambridge, MA: MIT Press.

[16] Kovalerchuk, B., Perlovsky, L., & Wheeler, G. (2012). Modeling of Phenomena and Dynamic Logic of Phenomena. Journal of Applied Non-classical Logics, 22(1), 51-82. http://arxiv.org/abs/1012.5415

[17] Kveraga, K., Boshyan, J., & M. Bar. (2007) Magnocellular projections as the trigger of top-down facilitation in recognition. Journal of Neuroscience, 27, 13232-13240.

[18] Masataka, N. & Perlovsky, L.I. (2012a). Music can reduce cognitive dissonance. Nature Precedings: hdl:10101/npre.2012.7080.1; http://precedings.nature.com/documents/7080/version/1

[19] Masataka, N. & Perlovsky, L.I. (2012b). The efficacy of musical emotions provoked by Mozart's music for the reconciliation of cognitive dissonance. Scientific Reports 2, Article number: 694 doi:10.1038/srep00694; http://www.nature.com/srep/2013/130619/srep02028/full/srep02028.html

[20] Masataka, N. & Perlovsky, L.I. (2013). Cognitive interference can be mitigated by consonant music and facilitated by dissonant music. Scientific Reports 3, Article number: 2028 (2013) doi:10.1038/srep02028; http://www.nature.com/srep/2013/130619/srep02028/full/srep02028.html

[21] Mugan, J. (2013). A developmental approach to learning causal models for cyber security. Proc. SPIE 8751, Machine Intelligence and Bio-inspired Computation: Theory and Applications VII, 87510A (May 28, 2013); doi:10.1117/12.2014418

[22] Perlovsky, L.I. (1987). Multiple sensor fusion and neural networks. DARPA Neural Network Study, 1987.

[23] Perlovsky, L.I. & McManus, M.M. (1991). Maximum Likelihood Neural Networks for Sensor Fusion and Adaptive Classification. Neural Networks 4 (1), 89-102.

[24] Perlovsky, L.I. (1996). Gödel Theorem and Semiotics. Proceedings of the Conference on Intelligent Systems and Semiotics '96. Gaithersburg, MD, v.2, pp. 14-18.

[25] Perlovsky, L.I. (1998). Conundrum of Combinatorial Complexity. IEEE Trans. PAMI, 20(6) pp. 666-670.

[26] Perlovsky, L.I. (2000). Beauty and Mathematical Intellect. Zvezda, 2000(9), 190-201 (Russian)

[27] Perlovsky, L.I. (2001a). Neural Networks and Intellect: using model-based concepts. Oxford University Press, New York, NY (3rd printing).

[28] Perlovsky, L. I. (2001b). Mystery of sublime and mathematics of intelligence. Zvezda, 2001(8), 174-190, St. Petersburg.

[29] Perlovsky, L.I. (2004). Integrating Language and Cognition. IEEE Connections, Feature Article, 2(2), 8-12.

[30] Perlovsky, L.I. (2006a). Toward Physics of the Mind: Concepts, Emotions, Consciousness, and Symbols. Phys. Life Rev. 3(1), 22-55.

[31] Perlovsky, L.I. (2006c). Music – The First Principle. Musical Theatre, http://www.ceo.spb.ru/libretto/kon_lan/ogl.shtml

[32] Perlovsky, L.I. (2007). Evolution of Languages, Consciousness, and Cultures. IEEE Computational Intelligence Magazine, 2(3), 25-39

[33] Perlovsky, L.I. (2008). Music and Consciousness, Leonardo, Journal of Arts, Sciences and Technology, 41(4), pp.420-421.

[34] Perlovsky, L.I. (2009a). Language and Cognition. Neural Networks, 22(3), 247-257. doi:10.1016/j.neunet.2009.03.007.

[35] Perlovsky, L.I. (2009b). Language and Emotions: Emotional Sapir-Whorf Hypothesis. Neural Networks, 22(5-6); 518-526. doi:10.1016/j.neunet.2009.06.034

[36] Perlovsky, L.I. (2009c). 'Vague-to-Crisp' Neural Mechanism of Perception. IEEE Trans. Neural Networks, 20(8), 1363-1367.

[37] Perlovsky, L.I. (2010a). Intersections of Mathematical, Cognitive, and Aesthetic Theories of Mind, Psychology of Aesthetics, Creativity, and the Arts, 4(1), 11-17. doi: 10.1037/a0018147.

[38] Perlovsky L.I. (2010b). Physics of The Mind: Concepts, Emotions, Language, Cognition, Consciousness, Beauty, Music, and Symbolic Culture. WebmedCentral PSYCHOLOGY 2010;1(12):WMC001374; http://arxiv.org/abs/1012.3803

[39] Perlovsky, L.I. (2010c). Joint Acquisition of Language and Cognition; WebmedCentralBRAIN;1(10):WMC00994; http://www.webmedcentral.com/article_view/994

[40] Perlovsky, L.I. (2010d). Musical emotions: Functions, origin, evolution. Physics of Life Reviews, 7(1), 2-27. doi:10.1016/j.plrev.2009.11.001

[41] Perlovsky, L.I. (2012a). Cognitive function, origin, and evolution of musical emotions. Musicae Scientiae, 16(2), 185 – 199; doi: 10.1177/1029864912448327.

[42] Perlovsky, L.I. (2012b). Cognitive Function of Music, Part I. Interdisciplinary Science Reviews, 37(2), 129–42.

[43] Perlovsky, L.I. (2012c). Emotionality of Languages Affects Evolution of Cultures. Review of Psychology Frontier, 1(3), 1-13. http://www.academicpub.org/rpf/paperInfo.aspx?ID=31

[44] Perlovsky, L.I. (2013a). A challenge to human evolution – cognitive dissonance. Front. Psychol. 4:179. doi: 10.3389/fpsyg.2013.00179; http://www.frontiersin.org/cognitive_science/10.3389/fpsyg.2013.00179/full

[45] Perlovsky, L.I. (2013b). Learning in brain and machine - complexity, Gödel, Aristotle. Frontiers in Neurorobotics; doi: 10.3389/fnbot.2013.00023; http://www.frontiersin.org/Neurorobotics/10.3389/fnbot.2013.00023/full

[46] Perlovsky, L.I. (2013c). Cognitive Function of Music, Part II. Interdisciplinary Science Reviews, 38(2), 149-173.

[47] Perlovsky, L.I. (2013d). Language and cognition – joint acquisition, dual hierarchy, and emotional prosody. Frontiers in Behavioral Neuroscience, 7:123; doi:10.3389/fnbeh.2013.00123; http://www.frontiersin.org/Behavioral_Neuroscience/10.3389/fnbeh.2013.00123/full

[48] Perlovsky, L.I., Deming R.W., & Ilin, R. (2011). Emotional Cognitive Neural Algorithms with Engineering Applications. Dynamic Logic: from vague to crisp. Springer, Heidelberg, Germany.

[49] Perlovsky, L. I., Bonniot-Cabanac, M.-C., Cabanac, M. (2010). Curiosity and Pleasure. WebmedCentral PSYCHOLOGY 2010;1(12):WMC001275; http://www.webmedcentral.com/article_view/1275; http://arxiv.org/ftp/arxiv/papers/1010/1010.3009.pdf

[50] Perlovsky, L.I., Cabanac, A., Bonniot-Cabanac, M-C., & Cabanac, M. (2013). Mozart Effect, Cognitive Dissonance, and the Pleasure of Music. ArXiv 1209.4017; Behavioural Brain Research, 244, 9-14.

[51] Perlovsky, L.I. & Ilin, R. (2010). Neurally and Mathematically Motivated Architecture for Language and Thought. Special Issue "Brain and Language Architectures: Where We are Now?" The Open Neuroimaging Journal, 4, 70-80. http://www.bentham.org/open/tonij/openaccess2.htm

[52] Price, C.J. (2012). A review and synthesis of the first 20 years of PET and fMRI studies of heard speech, spoken language and reading. NeuroImage, 62, 816–847.

[53] Shabtai, A., Moskovitch, R., Feher, C., Dolev, S., & Elovici, Y. (2012). Detecting unknown malicious code by applying classification techniques on OpCode patterns, Security Informatics, 1:1; http://www.security-informatics.com/content/1/1/1.

[54] Singer, R.A., Sea, R.G. & Housewright, R.B. (1974). Derivation and Evaluation of Improved Tracking Filters for Use in Dense Multitarget Environments, IEEE Transactions on Information Theory, IT-20, 423-432.

[55] Tikhanoff. V., Fontanari, J. F., Cangelosi, A. & Perlovsky, L. I. (2006). Language and cognition integration through modeling field theory: category formation for symbol grounding. In Book Series in Computer Science, v. 4131, Heidelberg: Springer.

[56] Vityaev, E.E., Perlovsky, L.I., Kovalerchuk, B.Y., Speransky, S.O. (2011). Probabilistic dynamic logic of the mind and cognition, Neuroinformatics, 5(1), 1-20.

[57] Vityaev, E.E., Perlovsky, L.I., Kovalerchuk, B. Y., & Speransky, S.O. (2013). Probabilistic dynamic logic of cognition. Invited Article. Biologically Inspired Cognitive Architectures 6, 159-168.