Neuroinformatics Presumptions of Semantic Representations Evolution

Alexander Kharlamov
Neuroontogenesis lab. Institute of Higher Nervous Activity and Neurophysiology of RAS
Moscow State Linguistic University
Higher School of Economics
Moscow
kharlamov@analyst.ru

***Abstract*—To date, traditional information technology and artificial intelligence technology have evolved independently of each other. Now is the time to fundamentally rethink the experience of using and evolution of traditional information technology and its integration with artificial intelligence technology. Currently, the key problem in the development of information technology in general and artificial intelligence technology in particular is the problem of ensuring the information compatibility of computer systems, including intelligent systems. One of the advantages of combining knowledge bases — the core of modern information systems — is the use of a more or less standard way (in the form of semantic networks) to represent knowledge. However, the sole use of semantic networks to represent knowledge does not solve the problem of standardization. For such a representation to become generally accepted, it must be based on the fundamental principles of describing semantics. The most fundamental concept in this sense is the representation of knowledge in human consciousness. Semantic representations in the modern sense – that is, thesauri, semantic networks, ontologies – are a very rough description of the highest level of reality where we live, which can only conditionally be called a model of the world. In order to understand the architecture of the semantic representations that adequately describe the human world (as the modeled object), it is necessary to reproduce the architecture of the human world model as it is formed in the human consciousness. To understand this architecture, one need to look at the human brain (a natural neural network that forms semantic representations), at its (brain) organs that form this architecture, at the cognitive networks that form in it in response to particular situations at its input and output in the process of solving specific problems, at the informatics of individual sensory and effector modalities, the levels of hierarchical representations of which contain images of events of the external and interoceptive world of human of varying degrees of complexity, at the combinations of these hierarchies in multimodal representations, including those in the description of entire situations, as well as the sequence of situations, how they are used in the process of the unconscious and purposeful behaviour – its planning and implementation control. Obviously, the brain is both very large and very heterogeneous neural network that is complex in architecture. Such representations are not only well supported by a comparison with the architecture and informatics of the brain, but are also effectively modelled in applications. TextAnalyst, a software technology for automatic semantic analysis of unstructured texts (which is based on an artificial neural network based on neurons with time summation of signals), effectively implements the functions of forming a homogeneous semantic network, automatic abstracting of texts, comparing texts by their meaning, as well as classifying and clustering texts. It can be assumed that this technology will also effectively analyze code sequences obtained in the analysis of video sequences, if this analysis is sufficiently bionic. A single approach to the processing of textual and visual information will enable constructing effective multimodal systems for processing and presenting information, which is the only accurate approach to the modelling of human intellectual functions.**

***Keywords*—integration of information technology, semantic networks as the basis for standardizing the representation of knowledge, semantic networks in the human consciousness, artificial neural networks, neurons with temporary summation of signals, semantic analysis of texts, integration of modalities**

I. Introduction

To date, traditional information technology and artificial intelligence technology have evolved independently of each other. Now is the time to fundamentally rethink the experience of using and evolution of traditional information technology and its integration with artificial intelligence technology. This is necessary to eliminate a number of shortcomings of modern information technology [1].

Currently, the key problem in the development of information technology in general and artificial intelligence technology in particular is the problem of ensuring the information compatibility of computer systems, including intelligent systems. One of the advantages of combining knowledge bases — the core of modern information systems — is the use of a more or less standard way (in the form of semantic networks) to represent knowledge. However, the sole use of semantic networks to represent knowledge does not solve the problem of standardization. For such a representation to become generally accepted, it must be based on the fundamental principles of describing semantics. The most fundamental concept in this sense is the representation of knowledge in human consciousness, which is the objective of this paper.

Semantic representations in the modern sense – that is, thesauri, semantic networks, ontologies – are a very rough description of the highest level of reality where we live, which can only conditionally be called a model of the world. In order to understand the architecture of the semantic representations that adequately describe the human world (as the modeled object), it is necessary to reproduce the architecture of the human world model as it is formed in the human consciousness. To understand this architecture, one need to look at the human brain (a natural neural network that forms semantic representations), at its (brain) organs that form this architecture, at the cognitive networks that form in it in response to particular situations at its input and output in the process of solving specific problems, at the informatics of individual sensory and effector modalities, the levels of hierarchical representations of which contain images of events of the external and interoceptive world of human of varying degrees of complexity, at the combinations of these hierarchies in multimodal representations, including those in the description of entire situations, as well as the sequence of situations, how they are used in the process of the unconscious and purposeful behaviour – its planning and implementation control. Obviously, the brain is both very large and very heterogeneous neural network that is complex in architecture.

Understanding the architecture of the brain enables correct modelling of its individual elements, their combinations and the architecture as a whole, which is the only tool to implement truly intelligent systems (including the highest level of intelligence, an integrated robot) that will solve intellectual problems no worse than humans, and maybe better if one can implement a model of this architecture compactly enough.

Understanding how the architecture of the human brain is structured will make it possible to approach the solution to the tasks of intelligence cloning, as well as the tasks of the natural informational combination of separate natural and artificial intelligence into some hybrid system (bypassing the sensory channels of exchange, a bottleneck in the exchange of information, but necessary due to significant differences in internal representations of models of the world of various people), which will enable the transition from a virtual repository of social knowledge (currently available) to some real unified storage of knowledge of mankind.

To understand all this, one has to turn to our knowledge of the human brain. The more accurate our understanding of its structure and functions is, the better the artificial intelligent systems that we are trying to build will be. Therefore, let us look at the human brain and its informatics in relation to the semantic representations that model the world of human.

II. Brain informatics

The transition from the signal level of information processing to the symbolic one occurs in the periphery of sensory organs, which differs for different modalities, but brings the streams of input sensory information of various modalities to a single form that corresponds to its (sensory information) representation in the cortical parts of the analyzers. The human brain includes three main organs that handle specific information at the symbolic level. This is the cortex of the hemispheres, cerebrum, hippocampus and thalamus. The cerebral cortex forms a model of the human world as a hierarchy of images of events of varying complexity of various modalities. The hippocampus forms representations of the images of situations and brings these representations into line with real situations encountered by human in the process of his/her activity. Finally, the thalamus is the energy commutator of the brain, which implements the attention mechanism, either focusing it (attention) on one process, or defocusing it to provide completely parallel solving of many tasks simultaneously. Let us consider the first two organs in more detail, since it is they that perform the basic processes of forming a model of the human world, manipulating specific sensory and effector information. One important point to bear in mind: this paper deals exclusively with information processing; intentional moments are left out of the scope of the paper.

A. Cerebral cortex

In the columns of the cerebral cortex, which are mainly composed of groups of pyramidal neurons of the third layer, from the flows of sensory information previously pre-processed in the periphery of sensory organs, hierarchies of images of human world events of various modalities are formed. Each modality at the periphery is processed in its own way, but at the input to the cortex, all modalities are represented by one-type sequences (either lineups or matrixes of sequences, depending on the type of information being analyzed; visual information is more complex than auditory information) of event codes.

These sequences in the columns of the cortex form sequences of pyramidal neurons that respond to them (each pyramidal neuron is a filter for its own fragment of the input sequence, including (fragment) n symbols of this sequence), which is a sequence of passed vertices of the n-dimensional signal space (trajectory), where the coordinates of the vertices correspond to the addresses of the filtering neurons in the corresponding column modelling a fragment of the n-dimensional space R^n. Such a mapping of a symbolic sequence into a multidimensional space leads to the structural processing of input information, that is, to the formation of a hierarchy of dictionaries of event images of various complexity for a particular modality [2], as well as to the formation of syntactic sequences including relations of words from the dictionary of the previous level in the input sequence received at the input of the next processing level.

Hierarchy of representations for a separate modality

The frequency of occurrence of various events of varying complexity at the input of sensory organs is different. So, if we consider textual representations, such events will be letters (or speech sounds), morphemes, words, syntactic groups, steadily repeating pairs of notions. In a natural language, this is called level-forming elements, and the corresponding language levels are called, respectively: graphematic (acoustic-phonetic, for spoken speech), morphemic, lexical, syntactic levels, and semantic level of a separate sentence. Such a representation is characteristic not only for the textual modality, but also for other modalities as well, since the sensory sequences of these modalities are encoded quasi-texts observed by human – the sequences that are meaningful and structured at several levels of complexity of their constituent elements (for example, a video sequence).

The hierarchy of such representations in the form of dictionaries of event images is formed in the columns of the cortical parts of the analyzers; each modality (in each analyzer) has its own one. A feature of the formation of such a hierarchy is the presence of a previous formed representation level as a filter needed for the formation of the next level: when forming a dictionary of the next level, the input information sequence interacting with the formed dictionary of the previous level, generates a so-called syntactic sequence (a sequence of word relations from the dictionary of the previous level in the input sequence, which is used at the input of the next level to form the dictionary of the following level). Thus, a hierarchy of dictionaries is formed in which the words of lower-level dictionaries are inserted by association into the words of dictionaries of the following levels (the words of the higher-level dictionaries are grammars for the words of the lower-level dictionaries), and this entire hierarchy can be considered a world model of the human (in terms of a particular modality), as well as a set of interconnected elements – images of events at various levels – a static model of the world in terms of this particular modality.

Combining hierarchies – a multimodal model of the world

The hierarchies in the dictionary of event images of various modalities formed in this way are virtually combined into a single representation, being combined by associative level-by-level relations between elements describing the same phenomena in terms of different modalities. The same hierarchies are combined informationally (at the higher levels of representations): syntactic sequences lose some amount of information (it remains in dictionaries) as they rise from level to level, therefore, these informational sequences that become increasingly sparse are combined, complementing each other in places of lacunae turning into a single multimodal information sequence. This combination takes place in the parietal cortex that collects information of all modalities into a single representation, a multimodal static model of the world.

Anterior and posterior cortex

Everything written above concerned the sensory (posterior) cortex. There is also the anterior cortex, which is essentially a motor cortex. It is also a hierarchy of representations, the lower level of which – the motor cortex proper – controls the human motor activity. What are the upper levels of the anterior cortex?

In processing information by the posterior and anterior cortex, a hippocampus is involved at their border. First, let us talk about what happens in the process of processing sensory information in the hippocampus, and how it participates in the formation of representations in the anterior cortex.

B. Hippocampus

The main function of the hippocampus is the formation of sensory representations of the situation, integral in space and time, that include images of events represented in the cortex in a way they are included in particular situations of the external and interoceptive world of human. In individual hippocampal lamellae (CA_3 field represents the Hopfield associative memory [3]), images of situations are formed as spatio-temporal groups of event images represented in the corresponding columns of the cortex.

Another function of the hippocampus is to filter the input flow of situations through the images of situations presented in the lamellae, and to identify the degree of their similarity to the images of situations (averaged in some sense) stored in them. After identifying the degree of similarity, the representation of the old and new images of situations is averaged.

Hippocampus in the processing of sensory information

The situations represented in the lamellae of the hippocampus are used by the higher levels of the motor (front) cortex in the process of formation and control of purposeful behaviour.

This representation can be illustrated by the extended predicate structure of the sentence (which actually describes the situation), which is a graph that includes the subject, predicate, main and secondary objects and attributes. The graph describing the situation and represented in the lamella of the hippocampus is similar to the graph of the extended predicate structure, and includes multimodal representations of the same images: subject, main and secondary objects, attributes. This representation includes both linguistic and multimodal representations, both sensory and motor ones.

Representations of situations in the hippocampus due to the influence of representations of current implementations of situations at the input to the cortex are averaged in free time of the human. The sensory image of the situation projected onto the columns of the cortex by association selects those columns where there is a representation of the images of events included in the input situation. This image is projected (also associatively) onto the hippocampus, where those lamellae respond, the image of the situation in which most closely matches the input situation. During the iterative procedure, the found lamellae affect the images of the initial events of the cortical columns (the so-called long-term potentiation [4]) providing further training for the images of events stored in the columns to an average state (average between the state stored in the lamella and the state received from the input). In the process of this procedure, further training of images of situations stored in the hippocampus also takes place.

Hippocampus in the processing of motor information

The situations presented in the lamellae of the hippocampus are used by the higher levels of the motor (anterior) cortex in the process of formation and control of purposeful behaviour.

In the images of situations represented in the lamellae of the hippocampus, motor information is also recorded in addition to sensory information. Thus, along with sensory images, their effector analogues are stored there. It is they that form at the lower level of the anterior cortex (the motor cortex proper) the representations that control the effector organs of the human.

C. *Anterior (motor) cortex*

The world represented in the effector model of the world (the anterior cortex) is, in some sense, mirrored by its representation in the sensory cortex. A hierarchy of representations is also formed in the effector part of the world model, but these representations work in the opposite direction: while in the sensory cortex information flows come from the bottom up (analysis), they are directed from top to bottom (synthesis) in the motor cortex. But before this control flow from top to bottom occurs, such images of effector (motor) images of events of various levels should be formed.

Inner speech as an example of forming a hierarchy of representations of the anterior cortex

Let us show how motor representations are formed by the example of the formation of internal speech, as A.R. Luria [5] described it (managing internal speech also belongs to motor skills). Inner speech is initially formed with the participation of the teacher. Mother tells her son: "Do this, do that." And he does. In the son's hippocampus lamella, a situation is remembered when he hears, repeats what he hears and does what he is told. In the first level of the anterior cortex (the motor cortex proper) following the last level of the posterior cortex, the least variable information is recorded – repeating the phrase ("Do this!").

Over time, a fair amount of such events accumulate in the motor cortex (lower-level dictionaries are formed). However, life moves forward: in addition to simple actions, the son, under the control of his mother, performs more complex (increasingly complex) actions, in which those simple events are components. They form representations of the following levels. These representations of the anterior cortex are related to the hippocampus by association, as well as the representations of the sensory cortex: the corresponding hippocampal lamella by association responds to images of the anterior cortex that are involved in the situation description presented in this lamella ("Do this first, then that!"). But the same lamella also includes links to images of the sensory cortex, which describe the mentioned situation.

At higher and higher levels of the anterior cortex, increasingly complex effector events are presented – sequences of lower-level events of the motor cortex with links to the lamellae of the hippocampus, that is, to entire spatio-temporal images of situations. Thus, the higher levels of the anterior cortex manipulate sequences of images of situations.

D. *Purposeful behaviour*

Actually, the formation of such sequences of situations (from the current situation to the target one), and control over the implementation of these sequences, that is, correcting the discrepancies in the situation representations in the hippocampus (plus in the columns of the cortex, on which the representation in the hippocampus is based) and in reality is a purposeful behaviour. Inner speech is also purposeful behaviour.

III. Integration of modalities as a perspective for the development of semantic representations

The representation described above is, to some degree, implemented in modern intelligent applications that model the processing of information in the human brain. However, currently such models are somewhat one-sided. If speech recognition is modeled, the task of processing speech information only is to be solved. However, sometimes processing of visual information about articulatory movements is also added (the so-called lip-reading). If the processing of visual information is modelled, the task of processing visual information only is to be solved.

The human brain is not so one-sided. To process information of any modality, any suitable information is involved. If speech is recognized, then any necessary accessible context is involved, including information not only related to the language model of the world, but also extralinguistic information of modalities other than textual. The same thing happens when processing information of other modalities. It is the involvement of the widest (necessary) context that allows achieving such accuracy in speech recognition, even in noisy conditions. But one has to pay for it with the resources needed to support the provision of the appropriate context (including other modalities).

A. World models in the dominant and subdominant hemispheres

The human model of the world, due to the features of the information representation in the dominant and subdominant hemispheres, is divided into three independent parts: two of them are in the dominant hemisphere and the third in the subdominant one [6]. Since the cortical fields responsible for the perception and articulation of speech are represented in the dominant hemisphere of the human brain, it is here where the linguistic model of the world is formed, including the language model, and the world model described in terms of language. This model is formed in parallel with another model of the dominant hemisphere (multimodal/ extralinguistic), which is formed under the influence of society through the language model. The multimodal model of the dominant hemisphere is multi-level (knowledge of society is deep and wide), but schematic, since socialized knowledge is usually represented by few examples. In the subdominate hemisphere, the situation is fundamentally different. There is no direct influence of society (there is no speech perception and synthesis), therefore the model of the world is multimodal and very individual: it contains only the information that the person handled in the process of his/her development. But this model has only two levels of representation (part-whole) due to the very great variability of the events presented there.

These three parts of the model of the world are connected level-by-level by associative relations, and represent a single whole in the process of manipulation.

B. Integral robotics

The only modern type of intelligent systems where the use of a wide context is justified from this point of view, and is even necessary regarding the task assignment, includes integrated robots, which by their nature include sensors and effectors of many modalities in their architecture. But this is an advantage of tomorrow, and today resources are being saved there, and the sensory (and motor) skills only work as much as it is strictly necessary for solving the task. If this is a speech control, then a command recognition system with speaker-dependent tuning and restrictions in the dictionary (and other restrictions) is used.

IV. Mechanisms for the formation of semantic representations

At the moment there are few approaches to the representation of semantics. Traditionally, these are: (1) logical languages; (2) production rules systems; (3) frame representations; and (4) semantic networks. Recently, they were completed by (5) ontologies [7]. However, ontologies are just another (non-graphical) form of either frame representations or semantic networks. Therefore, we will not consider them separately.

Of the above, production rules systems (in expert systems), and semantic networks are still successfully used. The effectiveness of using these representations depends solely on the availability of mechanisms for automatic generation and manipulation of the semantic representation itself. In this sense, one can only talk about the presence of mechanisms for the automatic formation of semantic networks. All other types of semantic representations are formed exclusively manually. And even the deep learning mechanisms do not leave any hope for the automation of these processes [8].

It is easiest to automatically form homogeneous (so-called associative) semantic networks. They work almost exclusively on statistics of the processed text (for now, only mechanisms for analyzing natural language texts have been implemented). Using linguistic information, it is possible to automate the formation of heterogeneous semantic networks as well.

A. Homogeneous semantic networks

Homogeneous semantic networks are graphs, where vertices are connected by arcs that represent only one type of relations, that is, “be together”. Therefore, they are called associative. They are easily generated automatically, since identification of a dictionary of words (recall that so far only analysis of texts of natural language modality has been implemented) in a text and identification of a dictionary of pairwise occurrence of words (these dictionaries are necessary for constructing an associative network) is not difficult. And the calculation of the weights of the network vertices is a completely manageable procedure [9]. It is also easy to use them in applications [10].

B. Heterogeneous semantic networks

Inhomogeneous semantic networks require for their construction information on the types of relations between the notions of the network that characterize arcs. There are two ways to automatically construct heterogeneous semantic networks. One can first construct a homogeneous semantic network, and then mark up the relations between the vertices by examining the sentences of the text (corpus of texts), on the basis of which a homogeneous semantic network is constructed. One can also first examine the sentences of the text, on the basis of which a heterogeneous semantic network is constructed, and then calculate the weights of the network vertices in the same way as in calculating the weights of the vertices of a homogeneous semantic network.

In the first case, we obtain a somewhat "roughened" network, but it is resistant to distortions generated by a linguistic interpretation of link analysis. Since the procedure is statistical in nature, it is almost independent of the language. In the second case, the construction of the network is complicated by the fact that no more than 85% of the sentences of the Russian-language text are analyzed correctly when identifying the types of relations between their constituent words [11]. In addition, the mark-up of sentences according to the types of relations depends on the language, and the resulting network with very blurry vertices (each type of relation for the same pair of notions generates additional branches in the network) is poorly ranked by vertex weights.

V. Neural network interpretation of the world model architecture

We still have to say a few words about the possibility of a neural network implementation of the procedure for automatic formation of the world model. As mentioned above, only two types of semantic representations are implemented automatically. They will be discussed below.

A homogeneous semantic network can be interpreted as a model of the world (linguistic - textual), for example. The same mechanisms can be used to form an extralinguistic model of the world (so far this has not been implemented). To do this, it is necessary to replace natural language text sequences at the input of the semantic network formation process with quasi-text sequences (for example, a video sequence encoded accordingly).

A. Neural network based on neurons with temporal summation of signals – cortical column model

Constructing a homogeneous semantic network is an ordinary procedure, if artificial neural networks based on neural-like elements with time summation of signals are used for this [2]. Such artificial neural networks are naturally designed to identify dictionaries of words and word pairs in natural language texts, as described above. They model the processing of information in columns of the cortex.

B. Hopfield associative memory - hippocampal lamella model

Applying a procedure similar to the Hopfield associative memory mechanism [12], it is just as easy to calculate the weights of the network vertices [10].

C. Formation of a heterogeneous semantic network

Probably, one can look for evidence that the mark-up of relations between network vertices can also be implemented using artificial neural networks [13]. Well, then one can rely on a completely neural network mechanism for constructing a heterogeneous semantic network.

D. Analysis of texts and quasi-texts

All that was said about the analysis of texts can be said about the analysis of quasi-texts, which are understood as meaningful code sequences relating to the interpretation of the human world formed by sensors other than textual. Since quasi-texts are similar to natural-language texts, that is, they contain level-forming units of varying degrees of complexity, then, at least, theoretically, nothing prohibits them from being processed in ways similar to those described above, when forming both a homogeneous and a heterogeneous semantic network.

VI. Possible applications

The processes described above are quite easily algorithmized at the information level. And software solutions are different from similar ones relying on simpler and less natural bases, and work for the benefit of users. So, on the basis of a simplified language model of the world formed in the columns of the cortex, which includes only some levels of language representation supplemented by an iterative procedure for knowledge reordering, as it is implemented in the Hopfield associative memory, which models the CA_3 fields of hippocampal lamellae, a program system was created for semantic analysis of texts that uses a network representation of semantics and has proved quite effective [14].

A. Semantic analysis of texts

TextAnalyst, a software system for the semantic analysis of texts, was developed on the principles presented above. It includes a limited language model where there is no syntactic level (it is replaced by a dictionary of deleted words), and where morphological analysis is implemented by the simplest means, but an artificial neural network based on neural-like elements with time summation of signals is used to represent the lexical level and semantics of a separate sentence. A simplified representation is also taken from the informatics of the hippocampal lamella, including iterative reordering of the weights for the vertices of a homogeneous semantic network obtained at the semantic processing level [2].

Such a network modeling of text semantics makes it possible to take into account the word relations in the analyzed text by n steps in the semantic network. It is convenient for analyzing the text semantics, since it does not require a huge training sample of texts as in n-gram language modelling. The ranking of the text notions obtained as a result of the iterative procedure allows the extraction of a topic tree of texts, as well as abstracting and comparison of texts by meaning and classification of texts.

B. Analysis of a video sequence as a quasi-text

An effective solution to the problem of semantic analysis of natural language texts leaves hope for the effective use of the presented approach for the analysis of video sequences (and even individual two-dimensional images), as opposed to quite successful convolutional networks.

However, it requires implementation of a very specific image preprocessing, which is very different from the usual sweep from left to right and from top to bottom. If we model image processing by the eye of a person who sees images as large contoured surfaces filled with some texture (for example, color) and scans these images from the point of greatest informativity [15] to the point of greatest informativity (as a rule, these are contour turning points) according to a given rule (rules differ depending on the task assignment), a code sequence is obtained that resembles text. It repeats elements of varying complexity (they differ from those in natural language texts), that is, it becomes possible to structurally analyze them, as well as ordinary text.

But one can save the whole information processing mechanism, similar to that of the TextAnalsyt technology. But scanning itself differs in one aspect from traditional methods of image processing: we can scan a poorly represented part of the image in the model for as long as it takes to present it in detail.

C. Combining modalities

And finally, there is another very important advantage of this way of representing semantics. A single way of presenting and processing information of textual and visual modalities makes it possible to combine these representations into a single model of the world [16], in which fragments of the visual modality are named with the corresponding names of the textual modality. Such a two-modal representation is much more resistant to recognition problems than a single-modal visual one. In a two-modal representation, one of them complements the gaps (lacunes) of the other. But, most importantly, this representation makes it possible to interpret visual images and situations in terms of a natural language, greatly simplifying the communication of the system with the user.

VII. Conclusion

The tasks to be solved when creating artificial intelligent systems are effectively achievable provided that architectures similar to the human brain architecture used to solve a similar problem are used when solving them. Including the use of representations about the model of the world implemented on the basis of any approaches based on semantic representations. A prerequisite for the formation of a model of the task being solved in this case is the automatism of the formation of the mentioned semantic representations. At the moment, only semantic networks can be efficiently generated automatically, including involvement of artificial neural networks.

Such representations are not only well supported by a comparison with the architecture and informatics of the brain, but are also effectively modelled in applications. TextAnalyst, a software technology for automatic semantic analysis of unstructured texts (which is based on an artificial neural network based on neurons with time summation of signals), effectively implements the functions of forming a homogeneous semantic network, automatic abstracting of texts, comparing texts by their meaning, as well as classifying and clustering texts. It can be assumed that this technology will also effectively analyze code sequences obtained in the analysis of video sequences, if this analysis is sufficiently bionic. A single approach to the processing of textual and visual information will enable constructing effective multimodal systems for processing and presenting information, which is the only accurate approach to the modelling of human intellectual functions.

The standardization of knowledge representation on the basis of semantic networks, as they are presented in the human consciousness, is an appropriate basis for standardizing the ways of representing knowledge in information systems.

References

[1] V. Golenkov, N. Guliakina, and I.Davydenko, “Methods and tools for ensuring compatibility of computer systems,” [Open semantic technologies for intelligent systems, 2019, iss. 3, Pp. 25-53, 2019].

[2] A.A. Kharlamov, Assotsiativnaya pamyat’ – sreda dlya formirovaniya prostranstva znanij. Ot biologii k prilozheniyam. [Associative memory as an environment for the formation of a space of knowledge. From biology to applications]. Germany: Palmarium Academic Publishing, 2017. (in Russian)

[3] E.T. Rolls, “Theoretical and Neurophysiological Analysis of the Functions of the Primate Hippocampus in Memory,” [Cold Spring Harbor Symposia on Quantitative Biology, Vol. LV, Cold Spring Harbor Laboratory Press. Pp. 995 – 1006, 1990].

[4] O.S. Vinogradova, Gippokamp i pamyat’. [Hippocampus and memory]. Moscow, USSR: Nauka, 1975.(in Russian)

[5] A.R. Luria, Yazyk i soznanie. [Language and consciousness]. Rostov-on-Don: Feniks, 1998.(in Russian)

[6] V.L. Bianki, Morfofunktsionalnaia organizatsiia mezhpolusharnykh sviazei. [Morphofunctional organization of interhemispheric connections]. Leningrad, USSR: Nauka, 1989.(in Russian)

[7] I.V. Efimenko, and V. F. Khoroshevskii, Ontologicheskoe modelirovanie: podkhody, modeli, metody, sredstva, resheniia [Ontological modelling: approaches, models, methods, tools, solutions]. Moscow: HSE Publ. House, 2011. (in Russian)

[8] N.V. Lukashevich, and B.V. Dobrov, “Combining ontological and statistical methods in the tasks of automatic processing of large text data [Kombinirovanie ontologicheskikh i statisticheskikh metodov v zadachakh avtomaticheskoi obrabotki bolshikh tekstovykh dannykh],” [Coll. of papers of the Conference “Artificial intelligence: issues and solutions”. Pp. 32-37, 2018].

[9] A.A. Kharlamov, “Formirovanie n-grammnoj tematicheskoj modeli teksta. [Formation of an n-gram topic model of the text],” J Speech technology 1-2: 15–23, 2016. (in Russian)

[10] A.A. Kharlamov, Nejrosetevaya tekhnologiya predstavleniya i obrabotki informatsii (estestvennoe predstavlenie znanij). [Neural network technology of information representation and processing (natural representation of knowledge]. Moscow, Russia: Radiotekhnika, 2006. (in Russian)
[11] G.S. Osipov, I.A. Tikhomirov, and I.V. Smirnov, "Exactus – an intelligent meta-search system on the Internet [Exactus – sistema intellektualnogo metapoiska v seti Internet]," [the 10th National Conference on Artificial Intelligence with KII-2006 participation. Moscow: Fizmatlit. Vol. 3, pp. 859-866, 2006]. (in Russian)
[12] J.J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," Proc. Natl. Acad. Sci. 79, Pp. 2554 – 2558, 1982.
[13] L. He, K. Lee, O. Levy, L. Zettlemoyer, and P.G. Allen, "Jointly Predicting Predicates and Arguments in Neural Semantic Role Labeling," [Annual Meeting of the Association for Computational Linguistics (Short Papers) , Melbourne, Australia, July 15 - 20, Pp. 364–369, 2018].
[14] D. Sullivan, Document Warehousing and Textmining. NY: Wiley publishing house, 2001.
[15] N.V. Zavalishin, and I.V. Muchnik, Model zritelnogo vospriyatiia i algoritmy analiza izobrazhenij [Visual perception model and image analysis algorithms]. Moscow: Nauka, 1974. (in Russian)
[16] A.A. Kharlamov, R.M. Zharkoy, and V.V. Arzumanov, "Kollektivnaya model mira mobilnogo robota. Yazykovoy analiz stseny i upravlenie peshekhodnym perekhodom [Collective model of the world of a mobile robot. Language scene analysis and pedestrian crossing management]," J Speech technology 1. Pp. 99-108, 2019. (in Russian)

Перспективы развития семантических представлений, основывающиеся на тенденциях нейроинформатики

Харламов А.А.

До настоящего времени традиционные информационные технологии и технологии искусственного интеллекта развивались независимо друг от друга. Сейчас настало время фундаментального переосмысления опыта использования и эволюции традиционных информационных технологий и их интеграции с технологиями искусственного интеллекта. Ключевой на текущий момент проблемой развития информационных технологий в целом и технологий искусственного интеллекта в частности является проблема обеспечения информационной совместимости компьютерных систем и в том числе интеллектуальных систем. Одной из возможностей информационного совмещения баз знаний – ядра современных информационных систем – является использование для представления знаний более или менее стандартного способа в виде семантических сетей. Однако само использование семантических сетей для представления знаний не решает проблемы стандартизации. Чтобы такое представление стало общепринятым, оно должно базироваться на фундаментальных принципах описания семантики. Наиболее фундаментальным в этом смысле представлением является представление знаний в сознании человека. Семантические представления в современном понимании – тезаурусы, семантические сети, онтологии – это очень грубое описание самого верхнего уровня действительности, в которой мы живем, которое лишь условно можно назвать моделью мира. Для понимания релевантной моделируемому предмету – реальному миру человека – архитектуры семантических представлений, адекватно описывающих этот мир, необходимо воспроизвести архитектуру модели мира человека, как она формируется в его сознании. Для понимания этой архитектуры необходимо посмотреть на мозг человека (естественную нейронную сеть, формирующую семантические представления), на его (мозга) органы, которые формируют эту архитектуру, на когнитивные сети, возникающие в нем в ответ на появление конкретных ситуаций на его входе и выходе в процессе решения конкретных задач, на информатику отдельных сенсорных и эффекторных модальностей, уровни иерархических представлений которых содержат образы событий внешнего и интероцептивного мира человека разной степени сложности, объединение этих иерархий в многомодальных представлениях, в том числе, в рамках описания целых ситуаций, а также последовательностей ситуаций, как они используются в процессе неосознанного, а также целенаправленного поведения – его планирования и контроля исполнения. Очевидно, что мозг – это и очень большая, и очень неоднородная, а потому, сложная по архитектуре нейронная сеть. Такие представления не только хорошо подтверждаются сравнением с архитектурой и информатикой мозга, но и эффективно моделируются в приложениях. Программная технология для автоматического смыслового анализа неструктурированных текстов TextAnalyst (в основе которой лежит искусственная нейронная сеть на основе нейронов с временной суммацией сигналов) эффективно реализует функции формирования однородной семантической сети, автоматического реферирования текстов, сравнения текстов по смыслу, классификации и кластеризации текстов. Можно предполагать, что также эффективно подобная технология будет анализировать кодовые последовательности, полученные при анализе видеорядов, если этот анализ будет достаточно бионичен. Единый подход к обработке текстовой и зрительной информации позволит говорить о построении эффективных многомодальных систем обработки и представления информации, что является единственно верным подходом в развитии моделирования интеллектуальных функций человека.

Received 09.01.2020