

# Choice of Knowledge Representation Model for Development of Knowledge Base: Possible Solutions

Sabina Katalnikova, Leonids Novickis  
Faculty of Computer Science and Information Technology  
Riga Technical University  
Riga, Latvia

**Abstract**—In current society knowledge, information and intelligent computer systems based on knowledge base play a great role. The ability of an intelligent system to efficiently implement its functions depends on the efficiency of organizing knowledge base, and on the fact whether the applied knowledge representation models comply with the set requirements. The article is devoted to the research of the problem of choosing the knowledge representation models. Based on the requirement analysis for knowledge representation models, one of the solutions for the researched problem shown is application of extended semantic networks. Analysis of extended semantic networks' properties is carried out, as well as relevant examples of representing knowledge of extended semantic networks' application for various spheres offered.

**Keywords**—Extended semantic networks; knowledge base; knowledge representation model; semantic networks

## I. INTRODUCTION

The end of 20th and the beginning of 21st century can be characterized by transfer from industrial to the so-called information society the peculiarity of which is a significant increase of the role of knowledge and information. Transfer from economy which was dominated by such traditional factors as land, labour and capital to knowledge based economy marked a new approach to the concept of economic efficiency. In the circumstances of the new reality society's interest in information and knowledge grew both in theoretical and practical aspect as apparently the lack of innovation, innovative products and services significantly decreases economic efficiency.

Computer systems which use data bases to solve several typical formalized tasks are based on developed rules, models and algorithms. On the other hand, the arisen complicated tasks are not always solvable with formalized rules and algorithms. In order to solve new problems, knowledge rooted intelligent systems are applied, which are based on knowledge bases and where the main focus is processing knowledge.

One of the central and most complex problems solved in development of knowledge-based systems is the problem of knowledge representation and processing: the efficiency of system being developed and the correctness of solutions obtained by its means depend on success of this problem's solution [1].

Many articles in the world are dedicated to the problem of knowledge representation [2]-[4]. Success of adjusting

knowledge representation models directly depends on the fact whether the applied models comply with the set requirements. Problem of choosing knowledge representation model and processing methods can be defined in the following way: how to represent knowledge structures from such sources as special literature and knowledge of highly qualified professionals (namely, to choose knowledge representation model), so that their automatized processing could allow efficient solutions of domain tasks and gain positive results.

The paper is organized as follows. Section II describes requirements for knowledge representation model. Section III outlines one of the knowledge representation models - semantic networks and their advantages. Section IV presents a general description of extended semantic networks, while Section V discusses in detail their possibilities. Section VI concludes the paper by briefly discussing the future direction of research.

## II. REQUIREMENTS FOR KNOWLEDGE REPRESENTATION MODEL

The basis for intelligent system is knowledge base which comprises all the information that the system uses in a systemized way. Thus, all the used and workable information within intelligent systems is presented in a semantically structured unified knowledge base which represents a notion of a wholesome world where this system "lives".

The aim of knowledge representation is the organization of necessary information in such a form that the intelligent system would be able to apply it for decision making, planning, analysis, judgment output and other function implementation process. In order for the knowledge representation model to be applicable in development of an intelligent system, it has to ensure representation of all the knowledge types necessary for the operation of the intelligent system.

Firstly, the knowledge base of the intelligent system must contain knowledge on the domain where this system will function. That is knowledge on basic units (concepts and objects) within the domain, as well as knowledge on how these units are related. Such knowledge includes relations that directly connect concepts as well as more sophisticated types of knowledge that represent several types of dependency among the domain concepts (logical and functional).

The majority of intelligent system knowledge bases also contain particular knowledge on the domain (subject

knowledge) that is represented as concept specimens (in a form of particular objects) and the relation between them – in a form of relations' or restrictions' specimens.

Another important type of knowledge necessary to be represented in an intelligent system is knowledge on problems and solutions within the modelled domain (methods and algorithms). This knowledge characterizes problem environment of the intelligent system. Such knowledge can be declarative and procedural. Declarative knowledge describes division of tasks in subtasks and their link to the solution methods. Such information is knowledge received from the user. This information is included in the system, it constantly changes and determines the system solutions. Procedural knowledge is task solving methods and particular algorithms. Such knowledge is developed once in such a way that the setup of the system is done based only on declarative knowledge.

In addition to knowledge included in the knowledge base, it is necessary that the knowledge describing a fragment of reality (situation) that defines the context and entry data for tasks solved by the intelligent system are represented in the intelligent system. Such knowledge, similarly to subject knowledge, usually is given as concept and relation and/or restriction specimen type.

There are many requirements submitted to a knowledge representation model. By analyzing these requirements, it is possible to define a requirement cluster for knowledge representation model in intelligent systems, namely, [5]-[8]:

- representation of knowledge meaning; acquiring a unified character of knowledge to be represented with an intention to comply with all the substantial objects from the viewpoint of the solvable task, their characteristics and relations, and ignore the irrelevant ones;
- representation of knowledge within concepts of natural language of the domain to be studied; clearness of development and representation of logical links and semantic relations of domain to be studied;
- preservation of initial information and acquisition of new information;
- representation of hierarchical structure of knowledge;
- possibilities of representing fuzzy knowledge,
- representation of both declarative and procedural knowledge;
- representation possibilities of logical operations and quantifiers;
- representation possibilities of intentional and extensional;
- possibilities of recognizing contradiction in knowledge to be represented;
- model uniformity;
- provision of integrity of knowledge to be represented;

- possibility of merging the knowledge structures.

### III. SEMANTIC NETWORKS

One of the knowledge representation models is net-type model where the domain is examined as a body of objects and their binding relations (for instance, semantic networks, conceptual graphs). Knowledge representation in network models is the closest to knowledge in natural language texts. Initially semantic network was made as representation model of long-term memory structure in psychology, but later it became one of the basic types of knowledge representation. The task of semantic networks is representation of concept clusters, namely, establishment of basic organisation of domain notions. The necessities for development of a semantic network [9]:

- analysis of structural interworking of content to be researched;
- exhaustive description of concepts and their relations;
- thorough processing of knowledge;
- the link between the new concepts and the existing concepts and notions.

Knowledge representation concept in the semantic network type is based on the idea that all the knowledge can be represented as a cluster of objects (concepts) and links (relation) between them. Semantic network possesses such characteristics from the viewpoint of the requirements mentioned in the previous section: knowledge representation in natural language notions, declarative knowledge representation, domain semantic link representation, clearness of knowledge description, integrity of knowledge structure representation. Thus, it is possible to conclude that a necessary model can be continuation of semantic network model adding to its logic and computing property [1].

In the example of the semantic network it is possible to establish the difference of data base (working memory) and knowledge base. Domain is a cluster of possible conditions of its entities. This cluster which is represented through common terminology, concepts, relations and laws creates knowledge base as an intentional semantic network. But in every particular situation characteristics of this domain entities have particular values. This particular data is represented in extensional semantic network (data base or working memory). Working memory is used to store temporary data. Information on aims, current tasks, finished tasks, incoming and outgoing messages and short-term connections are located here.

Advantages of semantic models' information processing [10], [11]:

- similarity of semantic network structure to semantic structure of natural language phrases;
- clearness of knowledge representation model;
- allness that is achieved as a result of choosing the appropriate relation cluster;
- knowledge representation with semantic networks allows significant simplification of knowledge

integration process that is implemented as identification and pasting of synonymic elements of integrated semantic networks;

- properly developed intelligent system's knowledge base as a semantic network completely eliminates doubling of information within such knowledge base;
- knowledge representation as a semantic network simplifies the associative access for various types of knowledge based fragments;
- knowledge processing semantic models are well suited for parallel asynchronic processing of information.

#### IV. EXTENDED SEMANTIC NETWORKS

As it was said, knowledge base is a depository of various types of knowledge that constantly, without any restrictions has to store, change, and adjust. It is possible if knowledge representation model allows rather arbitrary modification of its constructions. Thus, it is preferable for the inner language to include simple, uniform structures, which could be removed and added. On basis of such considerations conception of uniformity developed. Semantic networks comply with requirements of uniformity, but are restricted, for instance, from the viewpoint of generalized information, relations between situations or relation representation. In relation to this, extended semantic networks were developed in which node-concept necessary clusters and special complex elements – link nodes were introduced [12]-[15]. Extended semantic networks can be used for representation of sophisticated objects, logical, generalized information, different requirements and so on.

In extended semantic networks the nodes correspond not only to objects or concepts, but also to relations, logical components of information (truth or untruth facts), complex objects and so on. All that could be regarded as an independent unit must correspond to a separate node. In such networks instead of semantic network edges the so-called link nodes are used. This node does not correspond to any object or relation, it is used only for indicating the link and ensures unified significance for nodes that correspond to separate components or information units. As a result a fragment that corresponds to the elementary situation forms. From such fragments the networks are composed.

There are also special constructions that are called semantic graphs. With their help it is additionally listed which component should be searched first, which – afterwards and so on. Every graph gives its own operations that are carried out on the network and leads to finding or distributing to network nodes. Network can be regarded as a special case of graph, namely, network is a graph where the processing direction is not set. Networks and graphs are composed of uniform fragments, every fragment can be removed or added to the network without damaging the correctness of syntax or semantics.

A special complex element – link node – connects with the help of marked edges to node-relation and nodes-objects, as a result a fragment is made that corresponds to an elementary situation, that is, objects related to a certain relation. This fragment is called an elementary fragment in the following way:  $D_0 (D_1, D_2, \dots, D_k / D_{k+1})$ , where  $D_0$  is the relation word,  $D_1, D_2, \dots, D_k$  – objects participation in relation,  $D_{k+1}$  – link node that describes the whole cluster participating in relation,  $D_0, D_1, D_2, \dots, D_{k+1} \in D$ ,  $D$  – node cluster,  $k > 0$ . Extended semantic networks are regarded as a final cluster of elementary fragments. With the help of semantic networks relation cluster, various situations, scenarios are represented. Every elementary fragment has its own link node that describes its fragment [14].

Formal description of extended semantic network [14]:

- 1) if  $\{D_0, D_1, D_2, \dots, D_k, D_{k+1}\} \subseteq D$ ,  $k > 0$ , then  $D_0 (D_1, D_2, \dots, D_k / D_{k+1}) = T_0$ ;
- 2) every  $T_k$  is extended semantic network;
- 3) if  $T_1$  and  $T_2$  – extended semantic networks, then compositions  $T_1 T_2$  and  $T_1 T_2$  are also extended semantic networks, moreover,  $T_1 T_2 \equiv T_2 T_1$ .

In an extended semantic network a rather free placement of nodes in various positions is allowed. Any node that stands in a position of any elementary fragment (and represents a particular object or object cluster, or relation word) can be placed in a different position of a different fragment. As a result, it is possible to represent a situation where relation words have the role of objects and close their relations. Link node of an elementary fragment can also be included in different elementary fragments but in a different role. With the help of such elementary fragments representation of such cases when some cluster is reviewed as a complex object that, for its turn, comes into relation is ensured. It defines the model's uniformity.

For processing of extended networks comparison by example principle is applied using a method of two network overlay. This principle is based on identification rules that allow linking the nodes and compare the networks on the basis of logic laws [10].

#### V. REVIEW OF EXTENDED SEMANTIC NETWORKS

We shall review the compliance of extended semantic networks to a requirement cluster that is proposed for knowledge representation model in intelligent systems. Extended semantic network model characteristics with informative examples are summarized in Table I.

By analyzing the examples and explanations offered in the table, it is possible to conclude that the extended semantic networks comply with the requirements of knowledge representation model in intelligent systems and can be used for development of knowledge base.

TABLE I. EXTENDED SEMANTIC NETWORK MODEL POSSIBILITIES/OPTIONS

Model Properties	Explanation
Knowledge meaning representation, acquisition of uniform character of representable knowledge	Semantic network by definition is a knowledge system with a definite meaning in a network, the nodes of which correspond to concepts and objects, and edges – to concepts and object relations, in an integrated character type [16]. Implementation of knowledge meaning representation requirement is achieved by including in the network structure relations that exist among the object of domain to be studied.
Knowledge representation in natural language notions, knowledge description clearness	For example, natural language expression “Cranberry – red, sour berry, which grows in a bog” is depicted in the following way [13]: SUB(‘berry’, ‘cranberry’) COLOUR(‘cranberry’, ‘red’) TASTE(‘cranberry’, ‘sour’) GROWS_IN(‘cranberry’, ‘bog’) SUB(‘cranberry’, ‘cranberry1’) SUB(‘cranberry’, ‘cranberry2’).. where relation SUB(‘cranberry’, ‘cranberry1’) SUB(‘cranberry’, ‘cranberry2’).., renders particular objects (for instance, those could be cranberries bought in a market or picked somewhere).
Knowledge hierarchical structure representation	Hierarchical structures are the knowledge basis on which catalogues, explanatory dictionaries, etc. are constructed. An example of representing such structures [17]: SUB(‘human’, ‘man’) SUB(‘human’, ‘woman’) SUB(‘man’, ‘Janis1’) SUB(‘man’, ‘Karlis1’) SUB(‘woman’, ‘Mara1’) SUB(‘woman’, ‘Una1’) For class relations node SUB is used. Here it is depicted that a human – it is men and women. In the lower level there are nodes representing particular humans. Each such node can have its own connection that represents characteristics and/or relations. Characteristics (relations) can also be for nodes-classes, and they are referable to all the class elements. SUB branches must form a tree (it cannot be a cycle). It is necessary from the viewpoint of characteristic’s succession – every high level node characteristics are inherent to all the lower level nodes. This principle allows significant decrease of the knowledge amount, using object class characteristics (relations) and automatically distributing them to particular objects.
Fuzzy knowledge representation	Expression “A1 is not very little” looks like this [12]: LENGTH(a1, x11) ‘NOT’(x21,x11) EVALUATION(x21, ‘little’) ‘VERY’(x21)
Declarative knowledge representation	An example of declarative knowledge representation [14]: network SUB(human, man) SUB(man,a1) SUB(man,a2) NAME(Jānis,a1) NAME(Jānis,a2) FATHER(a1,a2) represents that there are two men, who are human, who are named Jānis, one of them is the father of the other man.
Procedural knowledge representation	An example of procedural knowledge representation [13]: It is necessary to calculate the value of variable B3 by the following formula: B3:=(B1+25)*B2 The network looks like this: VALUE(‘B1’, x1 <sup>1</sup> ) VALUE(‘B2’, x2 <sup>1</sup> ) +(x1 <sup>1</sup> , ‘25’, x3 <sup>1</sup> ) *(x2 <sup>1</sup> , x3 <sup>1</sup> , x4 <sup>1</sup> ) ?(x4 <sup>1</sup> ) VALUE(‘B3’, x4 <sup>1</sup> ), where the relation “VALUE” means that the node x1 <sup>1</sup> corresponds to the value of variable B1; upper index 1 means that x <sup>1</sup> corresponds to some one single variable.
Logical operation representation	For example, the expression $\neg(P1 \wedge P2) \vee (P2 \wedge 1) = 0$ is depicted in the following way [14]: $\wedge(x1, x2, x3) \neg(x3, x4) \wedge(x2, t, x5) \vee(x4, x5, f)$ , where the variables P1 and P2 correspond to nodes x1 and x2, constants 1 and 0 correspond to nodes t and f.
Quantifier representation options	Predicate expression $(\forall X1 \in M1) R1(X1)$ is depicted in the following way [14]: $\in(x1, m1) \forall(x1, x2) r1(x2)$ , where node x1 corresponds to a cluster, node x2 corresponds to cluster representatives which are all inherent to common (additional) characteristics or relations r1. In the case of quantifier $\forall$ node x2 corresponds to the whole cluster, quantifier $\exists$ - sub-cluster, quantifier $\exists_1$ – one cluster element.
Preservation of initial information and acquisition of new information	Outer world is dynamic, although something is continuously changing. As it was mentioned before, special type of extended semantic networks are used – productions which have an important role in representing different types of time and cause-consequence dependence, definitions, explanations, etc. For example, we shall review a production which depicts the following changes – at first the object X1 possessed the characteristic R2, afterwards – R3. We are talking about an object that has relation R1 with A1 [13]: $R1(x1^1, a1) R2(x1^1/\gamma1) \text{ IN BEGINNIN-THEN}(\gamma1, \gamma2) R1(x1^1, a1) R3(x1^1/\gamma2)$ , where $\gamma1$ and $\gamma2$ mean nodes, in the cluster of the corresponding situation, $\gamma1$ – the initial situation, $\gamma2$ – the end situation.
Intensional and extensional representation options	A possibility to represent extensional and intensional means a possibility to represent in a model a cluster of objects that is described with a particular word (word extensional) and cluster or object characteristics that is described with a particular word (word intensional) [7]. We shall review the following example [18]. Study object is any material that can be used for teaching. When designing a study object the following main characteristics must be provided: <ul style="list-style-type: none"> <li>• they are little subject quants that last for 2 to 15 minutes;</li> <li>• they are closed, that is, they can be used separately;</li> <li>• repeatedly usable;</li> <li>• can be aggregated in one group;</li> <li>• marked with metadata.</li> </ul> Study objects can have different types: theoretical information; explanation; example; question; task; commentary etc. Such definition can be depicted in the following way: SUB(‘study material’, ‘study object’) MIN_DURATION(‘study object’, ‘2 min’) MAX_DURATION(‘study object’, ‘15 min’) CHARACTERISTIC1(‘study object’, ‘closed’) CHARACTERISTIC2(‘study object’, ‘repeatedly usable’) CHARACTERISTIC3(‘repeatedly usable’, ‘can be aggregated with others’) CHARACTERISTIC4(‘study object’, ‘marked with metadata’) SUB(‘study object’, ‘theoretical information’) SUB(‘study object’, ‘explanation’) SUB(‘study object’, ‘example’) SUB(‘study object’, ‘question’) SUB(‘study object’, ‘commentary’) SUB(‘study object’, ‘task’). Here six subtypes of a study object are established (theoretical information, explanation etc.) that are node words and make extensional of the word ‘study subject’. The words ‘completed’, ‘repeatedly applicable’ etc. make intensional of the word ‘study object’.

TABLE I EXTENDED SEMANTIC NETWORK MODEL POSSIBILITIES/OPTIONS (CONTINUATION)

Model Properties	Explanation
Contradiction recognition	When a new fragment of knowledge is received, it must be established whether it contradicts a fragment known before. For example, the system knows that <i>Janis</i> (who is our focus) <i>is the only child in the family</i> . Then the expression: <i>his brother's friend is definitely married</i> should form actions that are related both to outer actions (repeated question; statement that it cannot be etc.) and inner actions (change of notions) [13].
Knowledge structure representation integrity	Every complicated object or situation can be regarded as a unified whole. Then it is allocated its own node (link node).
Model uniformity	Network is constructed from uniform fragments, every fragment can be removed or added to the network without damaging the correctness of syntax or semantic.
Option to merge knowledge structures	<p>Possibility to merge knowledge structures means the possibility of the knowledge representation model to merge different networks structure which can contain, for example, nodes that represent words with the same meaning. Expediency of network structure merging is connected with the necessity to merge knowledge structures that describe domains with an intention to acquire a possibility to make decision within the framework of integrated domain. Before merging of knowledge structures it is necessary to check implementation of several provisions [13]:</p> <ul style="list-style-type: none"> <li>• there should be no alternatives; if we talk about separate objects, they are to be found, if several objects exist, they must be specified;</li> <li>• all the search possibilities must be used, including those that are connected with equivalent transformations, logical conclusion elements; search can be ineffective, for example, because there are relations in the expression that are depicted differently within knowledge; then reassessment of notions is necessary;</li> <li>• it is necessary to find out whether merging causes a contradiction to the information available before.</li> </ul>

In the end we shall consider an informative example of application of extended semantic networks.

The most important attributes of the study course are its name, laboriousness, sections (course content), the competences formed in the process of studying the course,

keywords. The study course is studied at a certain stage, is part of a study program, is completed with certain report, requirements of employers were taken into account, etc. Fragment of study course description with means of an extended semantic network is shown in Fig. 1.

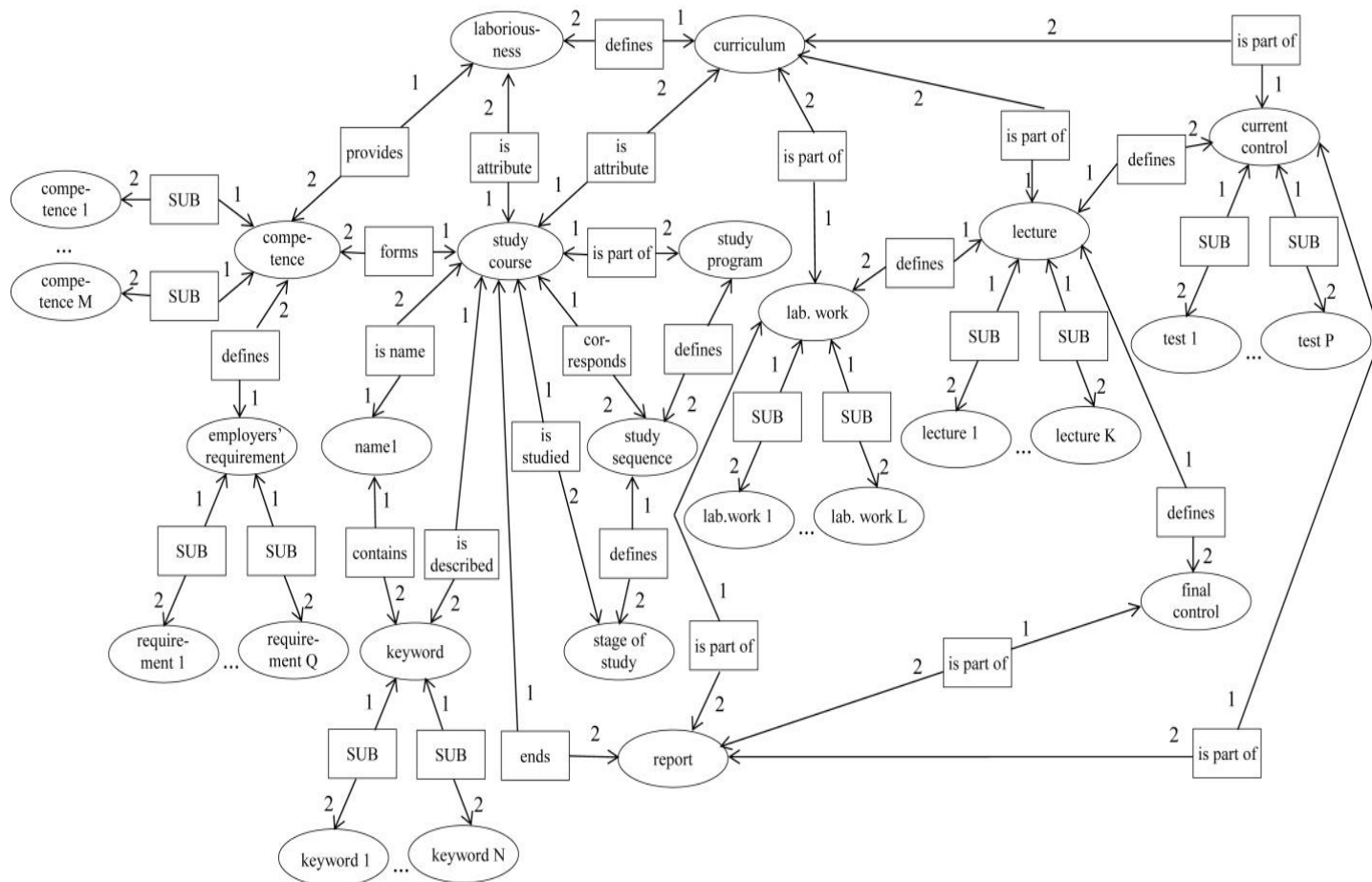


Fig. 1. Fragment of study course description.

This network is described as follows:  
IS PART OF('study course', 'study program')  
CORRESPONDS('study course', 'study sequence')  
IS NAME('name' 'study course')  
IS DESCRIBED('study course', 'keyword')  
FORMS('study course', 'competence')  
..... etc.

## VI. CONCLUSION

Topicality of knowledge representation problem is evident. Choice of knowledge representation model is one of the main problems in intelligent system's development, and its essence is to choose such a model that would satisfy the set requirements.

Software developers frequently try to describe complicated domains, where informative sophisticated tasks are solved, by using monotone regular structures that are too primitive for representation of the whole variety of domain's meaning nuances, even though they are convenient for further processing [18].

The article examines body of requirements for a knowledge representation model in intelligent systems, it is offered to use extended semantic networks for knowledge representation and with examples it is shown that this model complies with the aforementioned body of requirement. The offered material could be useful for developers of intelligent systems and other researchers to continue the work in knowledge representation and processing problem solution.

The authors believe that use of extended semantic networks in their work is expedient. Future work will focus on the evolution of extended semantic networks model for its use in development of collaborative intelligent educational systems.

## REFERENCES

[1] A. L. Yalovets (2011). "Representation and Processing of Knowledge from the Point of View of Mathematical Modeling". Naukova Dumka, Kiev [In Russian].  
[2] J. Bentahar, B. Moulin, M. Bélanger (2010). "A taxonomy of argumentation models used for knowledge representation" // Artificial

Intelligence Review, vol. 33 (3), pp. 211-259. doi: 10.1007/s10462-010-9154-1  
[3] M.M Abdul Jalil, C.P. Ling, N.M. Mohamad Noor, F. Mohd (2017). "Knowledge Representation Model for Crime Analysis" // Procedia Computer Science, vol. 116, pp. 484-491  
[4] I.N. Gluhiih, R.K. Akhmadulin (2017). "Problem-Oriented Corporate Knowledge Base Models on the Case-Based Reasoning Approach Basis" // IOP Conference Series: Materials Science and Engineering, vol. 221, issue 1, article number 012025  
[5] P. Clark (2007). "Requirements for a Knowledge Representation System Working", Note 10, pp. 1-10.  
[6] R. Davis, H. Shrobe, P. Szolovits (1993). "What Is a Knowledge Representation?" In: AI Magazin, Vol. 14(1), pp. 17-33.  
[7] H. Ueno, M. Ishizuka (1989). "Knowledge Representation and Usage". Moscow, Mir [In Russian].  
[8] D. H. Jonassen (1995). "Computers as Cognitive Tools: Learning with Technology, not from Technology, in: Journal of Computing in Higher Education", Spring, Vol. 6(2), pp. 40-73.  
[9] A. Bashmakov (2005). "Intelligent Information Technology". MSIT, Moscow [In Russian].  
[10] V. Golenkov, N. Guliakina (2014). "Project of Open Semantic Technology of the Componential Design of Intelligent System"s. Part 1: The Principles of Creation, in "Ontology of Designing" Scientific Journal, 1(11), pp. 42-64 [In Russian].  
[11] E. V. Zolotov, I. P. Kuznetso (1982). "Expanding Systems of Active Dialogue". Nauka, Moscow [In Russian].  
[12] I. P. Kuznetsov (1986). "Semantic Representation". Nauka, Moscow [In Russian].  
[13] I. P. Kuznetsov (1989). "The System of Knowledge Processing on Extended Semantic Network"s (in the form of a scientific report), Moscow [In Russian].  
[14] E. B. Kozerenko, I.P. Kuznetsov (2010). "Evolution of Linguistic Semantic Presentations in the Intelligent Systems Based on the Extended Semantic Networks", in: Computer Linguistics and Intelligent Technologies, Vol. 9, pp. 205-211  
[15] S. P. Habarov (2017). "Knowledge Representation in Information Systems", <http://www.habarov.spb.ru/bz/index.htm> (accessed: October 04, 2017 )  
[16] I. P. Kuznetsov, A.G.Matskevich (2006). "Semantic-Oriented Systems Based on Knowledge Bases", MTUSI, Moscow [In Russian].  
[17] J. Bule (2011). "A Set of Models for Managing Adaptive Computer-Based Learning", Doctoral Thesis, Riga, RTU [In Latvian].  
[18] A. Novikov (2016). "Problem Definition of Representation of Knowledges in Computer Systems" // Applied Informatics, vol. 11, Nr. 2(62), pp. 107-143, ISSN: 1993-8314 [In Russian]