Pre-Conference Proceedings of the Focus Symposium on Knowledge Management Systems (August 4, 2009, Focus Symposia Chair: Jens Pohl) in conjunction with InterSymp-2009, 21st International Conference on Systems Research, Informatics and Cybernetics, August 3 – 7, 2009, Baden-Baden, Germany). Sponsored by The International Institute for Advanced Studies in Systems Research and Cybernetics and Society for Applied Systems Research. Published by Collaborative Agent Design Research Center, California Polytechnic State University, San Luis Obispo, CA, USA, 2009. ISBN 978-1-897546-00-0. P. 39-50.

Scheme and Formal Tools for Transforming the Existing Web into Semantic Web of a New Generation

V. A. Fomichov

Department of Innovations and Business
in the Sphere of Informational Technologies
Faculty of Business Informatics
State University – Higher School of Economics
Kirpichnaya str. 33, 105679 Moscow, Russia
E-mail: vdrfom@aha.ru and vfomichov@hse.ru
Fax: +7-495-771-3238

Abstract

This paper indicates the basic model of the theory of K-representations (knowledge representations) elaborated by the author as a source of ideas for constructing an advanced language platform for Semantic Web. It is to allow for modeling a system of operations on conceptual structures enabling us to build semantic representations of practically arbitrary texts in Natural Language (NL) pertaining to arbitrary field of professional activity. The next subject of the paper is a new scheme of transforming the existing Web into a Semantic Web of a new generation with the well developed mechanisms of understanding natural language texts. For the realization of this scheme, the theory of K-representations provides a number of broadly applicable formal tools. The third subject of this paper is a description of the correspondence between the input texts of two elaborated algorithms of semantic-syntactic analysis and the semantic representations of these texts being the expressions of SK-languages (standard knowledge languages). The first algorithm has a program implementation in PHP. It is a computer program transforming the definitions of a part of the notions into the OWL-expressions The forth subject of this paper is a short description of the transformations fulfilled by this program.

Keywords

Semantics-oriented natural language processing; semantic representation; theory of K-representations; formal model of a linguistic database; SK-languages; algorithm of semantic-syntactic analysis

Introduction

The immense social significance of the Web has caused the awareness of the necessity to create the Web Science. One of the central problems of Web Science is to elaborate the theory of transforming the existing Web into a Semantic Web (Berners-Lee et al. 2006; Hendler et al. 2008).

It has been possible to observe the permanent expansion in the scientific literature of the following opinion: a Semantic Web satisfying the initial goal of this project will be created in an evolutionary way as a result of the efforts of many research groups in various fields. This point of view is expressed, in particular, in (Shadboldt et al. 2006). In this paper, the e-science international community is indicated as a community playing now one of the most important roles in quick generation of semantic content in a number of fields. The activity of this community seems to give a sign of future success of Semantic Web project.

In (Shadbolt et al. 2006), the authors ground the use of RDF as the basic language of the Semantic Web project with the help of the principle of least power: "the less expressive the language, the more reusable the data". However, it seems that the stormy progress of, first of all, e-science urges us to find a new interpretation of this principle in the context of the challenges faced nowadays by the Semantic Web project. E-science (in particular, bio informatics) needs to store on the Web the semantic content of the definitions of numerous notions, the content of scientific articles, technical reports, etc. The similar requirements are associated with semantics-oriented computer processing of the documents pertaining to economy, technology, medicine, law, politics. In particular, it is necessary to store the semantic content of the articles from newspapers, of TV-presentations, etc.

The substantial discussions of the role of semantics-oriented natural language processing mechanisms for constructing a Semantic Web satisfying the demands of numerous end users can be found in the papers (Fomichov 2000, 2008) and in the monographs (Wilks and Brewster 2006; Fomichov 2010).

That is why it can be conjectured that, in the context of the Semantic Web project, the following new interpretation of the principle of least power is reasonable: an advanced language platform for Semantic Web is to allow for modeling a system of operations on conceptual structures enabling us to build semantic representations (SRs) of practically arbitrary texts in Natural Language (NL) pertaining to arbitrary field of professional activity.

Following (Fomichov 2009), this paper indicates the basic model of the theory of K-representations (knowledge representations) elaborated by the author as a source of ideas for constructing an advanced language platform for Semantic Web. *The next subject of the paper* is a new scheme of transforming the existing Web into a Semantic Web of a new generation with the well developed mechanisms of understanding natural language texts.

For the realization of this scheme, the theory of K-representations provides a number of broadly applicable formal tools. *The third subject of this paper* is the description of the correspondence between the input texts of two elaborated algorithms of semantic-syntactic analysis and the semantic representations of these texts being the expressions of SK-languages (standard knowledge languages). The input texts of the second algorithms may belong to the sublanguages of English, German, and Russian languages. The first algorithm has a program implementation in PHP. It is a computer program transforming the definitions of the notions (these definitions form a subset of all definitions in the Russian language) into the OWL-expressions (as it is known, OWL is the principal language for developing ontologies under the framework of the Semantic Web project). *The forth subject of this paper* is a short description of the transformations fulfilled by this program.

Conceptual Operations Introduced by the Theory of K-representations

The question immediately emerges what a system of operations on conceptual structures satisfying the mentioned requirement might look like. A possible answer to this question is given by the theory of K-representations (knowledge representations) stated in numerous publications of the author in English and Russian, in particular, in (Fomichov 1996 - 2010). The basic mathematical model of this theory describes a system consisting of 10 partial operations on conceptual structures (Fomichov 2002b, 2002c, 2005a, 2007, 2010). The model determines a new class of formal languages for building SRs of sentences and complex discourses in NL – the class of SK-languages (standard knowledge languages). An early version of this model set forth in (Fomichov 1996 - 1998) determines the class of RSK-languages (restricted standard knowledge languages).

Let's consider the central ideas of determining the class of SK-languages At the first step (consisting of a rather long sequence of auxiliary steps), a class of formal objects called *conceptual bases (c.b)* is defined. Each c.b. *B* is equivalent to a system of the form $(c_1, ..., c_{15})$ with the components $c_1,..., c_{15}$ being mainly finite or countable sets of symbols and distinguished elements of such sets. In particular, $c_1 = St$ is a finite set of symbols called sorts and designating the most general considered notions (concepts); $c_2 = P$ is a distinguished sort "meaning of proposition"; $c_4 = X$ is a countable set of strings used as elementary blocks for building knowledge modules and semantic representations (SRs) of texts; *X* is called a primary informational universe; $c_5 = V$ is a countable set of variables; $c_7 = F$ is a subset of *X* whose elements are called functional symbols.
Each c.b. *B* determines three classes of formulas, the first class *Ls(B)* being considered as the principal one and being called *the SK-language (standard knowledge language) in the basis B*. Its strings (they are called K-strings) are convenient for building SRs of NL-texts. We'll consider below only the formulas from the first class *Ls(B).*

In order to determine for arbitrary c.b. *B* three classes of formulas, a collection of inference rules P[0], P[1],..., P[10] is defined. The rule P[0] provides an initial stock of formulas from the first class. E.g., there is such c.b. B_1 that, according to P[0], $Ls(B_1)$ includes the elements

house1, blue, city, set, France, 17, all, any,
Weight, Distance, Staff, Suppliers, Quantity, x1, x7, P1, P3.

For arbitrary c.b. *B*, let *Degr(B)* be the union of all Cartesian m-degrees of *Ls(B),* where *m 1* .
Then the meaning of the rules of constructing well-formed formulas P[0], P[1],..., P[10] can be explained as follows: for each k from 1 to 10, the rule P[k] determines a partial unary operation *Op[k]* on the set *Degr(B)* with the value being an element of *Ls(B).*

For instance, there is such conceptual basis B that the value of the partial operation *Op[7]* (it governs the use of logical connectives AND and OR) on the four-tuple
< *Ú , Austria, France, Germany* > is the K-string *(Austria Ú France Ú Germany).*

Thus, the essence of the basic model of the theory of SK-languages is as follows: this model determines a partial algebra of the form

(*Degr(B), Operations(B)*) ,

where *Degr(B)* is the carrier of the partial algebra, *Operations(B)* is the set consisting of the partial unary operations *Op[1], …, Op[10]* on *Degr(B).*

The volume of complete descriptions in (Fomichov 2005a, 2007, 2010) of the mathematical model introducing, in essence, the operations *Op[1], …, Op[10]* on *Degr(B)* and, as a consequence, determining the class of SK-languages considerably exceeds the volume of this paper. That is why, due to objective reasons, this model can’t be included in this paper. The short characteristics of thes partial operations on conceptual structures can be found, in particular, in (Fomichov 2002a, 2005b, 2008).

The Context for the Proposed Scheme of Developing Semantic Web of a New Generation

During several last years, it has been possible to observe that the achieved state of Semantic Web and a state to be relatively soon achieved are considerably different from the state of affairs outlined as the goal in the starting publication on Semantic Web by T. Berners-Lee, J. Hendler, and O. Lassila (Berners-Lee et al. 2001).

The principal reason for this conclusion is the lack of large-scale applications implemented under the framework of Semantic Web project. This situation is implied by the lack of a sufficiently big amount (of "a critical mass") of formally represented content conveyed by numerous informational sources in many fields. This means the lack of a sufficiently big amount of Web-sources and Web-services with semantic annotations, of the visual images stored in multimedia databases and linked with the high-level conceptual descriptions, rich ontologies, etc.

This situation is characterized in the Call for Papers of the First International Symposium on Incentives for Semantic Web (Germany, Karlsruhe, October 2008) as *the lack of a critical mass of semantic content.*

That is why it has been possible to observe the permanent expansion in the scientific literature of the following opinion: a Semantic Web satisfying the initial goal of this project will be created in an evolutionary way as a result of the efforts of many research groups in various fields. In particular, this opinion is expressed in (Angelova 2003).

It is important to underline that this point of view is also expressed in the article "Semantic Web Revisited" written by the pioneers of Web: N. Shadboldt, W. Hall, T. Berners-Lee (Shadboldt et al. 2006). In this paper, the e-science international community is indicated as a community playing now one of the most important roles in quick generation of semantic content in a number of fields. The activity of this community seems to give a sign of future success of Semantic Web project.

One of the brightest manifestations of the need for new, strong impulses to developing Semantic Web is the organization of the First International Symposium on Incentives for Semantic Web under the framework of the Semantic Web International Conference - 2008.

The content of the next section is to be considered in the context of the broadly recognized need for the incentives for Semantic Web, in particular, for the incentives on the models stimulating the development of Semantic Web.

A Possible Scheme of Developing Semantic Web of a New Generation

The analysis of the expressive power of the class of SK-languages allowed for putting forward the following hypothesis in (Fomichov 2002c, 2005a, 2007): SK-languages are a convenient tool of building semantic representations of arbitrary complex natural language texts (sentences and discourses) pertaining to arbitrary field of professional activity.

Let's consider the principal ideas of a new, theoretically possible scheme aimed at transforming the existing Web into a Semantic Web. The proposed scheme is based on (a) the mathematical model constructed in (Fomichov 2005a, 2007, 2010) and describing a system of 10 partial operations on conceptual structures and (b) the analysis of the expressive mechanisms of SK-languages carried out in the mentioned works. The new scheme can be very shortly formulated as follows:

1. An XML-based format for representing the expressions of SK-languages (standard knowledge languages) will be elaborated. Let's agree that the term "a K-representation of a NL-text T" means below a semantic representation of T built in this format and that the term "a semantic K-annotation" will be interpreted below as a K-representation of a NL-annotation of an informational source. The similar interpretations will have the terms "a K-representation of a knowledge piece" and "a high-level conceptual K-description of a visual image".
2. The NL-interfaces for different sublanguages of NL (English, Russian, German, Chinese, Japan, etc.) helping the end users to build semantic K-annotations of Web-sources and Web-services are being designed.
3. The advanced ontologies being compatible with OWL and using K-representations of knowledge pieces are being elaborated.
4. The new content languages using K-representations of the content of messages sent by computer intelligent agents (CIAs) in multi-agent systems are being worked up. In particular, this class of languages is to include a subclass being convenient for building the contracts concluded by the CIAs as a result of successful commercial negotiations.
5. The visual images of the data stored in multimedia databases are being linked with high-level conceptual K-descriptions of these images.
6. The NL-interfaces transforming the NL-requests of the end users of Web into the K-representations of are being designed.
7. The advanced Web-based search and question-answering systems are being created being able (a) to transform (depending on the input request) the fragments of a discourse into the K-representations, (b) to analyze these K-representations of the discourse fragments, and (c) to analyze semantic K- annotations of Web-sources and Web-services.
8. The NL processing systems being able to automatically extract knowledge from NL-texts, to build the K-representations of knowledge pieces, and to inscribe these K-representations into the existing ontologies are being elaborated.
9. The generators of NL-texts (the recommendations for the users of expert systems or of recommender systems, the summaries of Web-documents, etc.) using the SK-languages for

representing the meaning of a NL-text to be synthesized are being constructed. Besides, a reasonable direction of research seems to be the design of applied intelligent systems being able to present the semantic content of a message for the end user as an expression of a non-standard K-language being similar to a NL-expression but containing, may be, a number of brackets, variables, markers.

Fulfilling these steps, the international scientific community will create in a reasonable time a digital conceptual space unified by a general-purpose language platform. The realization of this scheme will depend on the results of its discussion by the international scientific community.

The Formal Tools Provided by the Theory of K-representations

The monographs (Fomichov 2005a, 2010) stating two versions of the theory of K-representations propose one universal (most likely) and several broadly applicable formal tools for the realization of this scheme.

The theory of K-representations is an expansion of the theory of K-calculuses and K-languages (the KCL-theory). The basic ideas and results of the KCL-theory are reflected in numerous publications both in Russian and English, in particular, in (Fomichov 1996 – 2005a).

The *first basic constituent* of the theory of K-representations is the theory of SK-languages (standard knowledge languages), stated, in particular, in (Fomichov 1996 – 2005b). The kernel of the theory of SK-languages is a mathematical model describing a system of such 10 partial operations on structured meanings (SMs) of natural language texts (NL-texts) that, using primitive conceptual items as "blocks", we are able to build SMs of arbitrary NL-texts (including articles, textbooks, etc.) and arbitrary pieces of knowledge about the world. The outlines of this model can be found in two papers published by Springer in the series “Lecture Notes in Computer Science” (Fomichov 2002a, 2005b).

The analysis of the scientific literature on artificial intelligence theory, mathematical and computational linguistics shows that today the class of SK-languages opens the broadest prospects for building semantic representations (SRs) of NL-texts (i.e., for representing meanings of NL-texts in a formal way).

The expressions of SK-languages will be called below the K-strings. If T is an expression in natural language (NL) and a K-string *E* can be interpreted as a SR of T, then *E* will be called a K-representation (KR) of the expression T.

The *second basic constituent* of the theory of K-representations is a widely applicable mathematical model of a linguistic database (LDB). The model describes the frames expressing the necessary conditions of the existence of semantic relations, in particular, in the word combinations of the following kinds: “Verbal form (verb, participle, gerund) + Preposition + Noun”, “Verbal form+ Noun”, “Noun1 + Preposition + Noun2”, “Noun1+ Noun2”, “Number designation + Noun”, “Attribute+Noun”, “Interrogative word + Verb”. The expressive power of SK-languages enables us to associate the lexical units with the appropriate simple or compound semantic units. The model describes the logical structure of linguistic databases being the components of natural-language interfaces to intelligent databases as well as to other applied computer systems.

The *third basic constituent* of the theory of K-representations is two complicated, strongly structured algorithms carrying out semantic-syntactic analysis of texts from some practically interesting sublanguages of NL. These algorithms, called *SemSyn* and *SemSynt1* respectively, are based on the elaborated formal model of a linguistic database. The algorithm *SemSyn* transforms a NL-text in its semantic representation being a K-representation, the algorithm *SemSyn* is described in two final chapters of the monograph (Fomichov 2005a), and the algorithm *SemSynt1* is set forth in Chapters 9 and 10 of the monograph (Fomichov 2010).

An important feature of these algorithms is that they don’t construct any syntactic representation of the inputted NL-text but directly finds semantic relations between text units. Since numerous lexical units have several meanings, the algorithm uses the information from a linguistic database and linguistic *context* for choosing one meaning of a lexical unit among several possible meanings.

The other distinguished feature is that these complicated algorithms are completely described with the help of formal tools, that is why they are problem independent and don’t depend on a programming system. The algorithm *SemSyn* is implemented in the Web programming language PHP.

The Principles of Selecting the Form of a Text's Semantic Representation

Let's consider the examples illustrating the correspondence between the natural language sentences in English and their semantic representations (SR) being the expressions of a certain SK-language, that is, being the K-representations of the input texts. In these examples, the SR of the input text T will be the value of the string variable *Semrepr (Semantic representation).*

Example 1. Let T1 = "The writer Igor Somov lives in Saratov". Then
Semrepr = Situation (e1, living * (Time, #now#)
(Agent1, certn person * (Qualification, writer)
(Name, 'Igor') (Surname, 'Somov') : x2)
(Place1, certn city * (Name1, 'Saratov') : x3)).

Example 2. Let T2 = "Deliver a box with details to the warehouse ^{1} 3". Then
Semrepr = (Command (#Operator#, #Executor#, #now#, e1)
$\wedge$ Target (e1, delivery1 * (Object1, certn box1 *
(Content1, certn set * (Qual-compos, detail)) : x1)
(Place2, certn warehouse * (Number1, 3) : x2).)

Example 3. Let T3 = "Did the international scientific conference "COLING" take place in Asia?". Then
Semrepr = Question (x1, (x1
Truth-value (Situation (e1, taking_place *
(Time, certn moment * (Earlier, #now#) : t1)

(Event1, certn conference * (Type1, international)
(Type2, scientific) (Name1, 'COLING') : x2)
(Place, certn continent * (Name, 'Asia') : x3))))).

Example 4. Let T4 = "What publishing house has released the novel "Winds of Africa"?". Then
Semrepr = Question (x1, Situation (e1, releasing1 *
(Time, certn moment * (Earlier, #now#) : t1)
(Agent2, certn publ-house : x1)
(Product1, certn novel1 * (Name1, 'Winds of Africa') : x2))).

Example 5. Let T5 = "What foreign publishing houses the writer Igor Somov is collaborating with?". Then
Semrepr = Question (S1, (Qual-compos (S1, publish-house *
(Type-geographic, foreign)) ∧
Description(arbitrary publish-house * (Element, S1) : y1,
Situation (e1, collaboration * (Time, #now#)
(Agent1, certn person * (Occupation, writer)
(Name, 'Igor') (Surname, 'Somov') : x1)
(Organization1, y1))))).

Example 6. Let T6 = "Who produces the medicine "Zinnat"?". Then
Semrepr = Question (x1, Situation (e1, production1 *
(Time, #now#) (Agent2, x1)
(Product2, certn medicine1 * (Name1, 'Zinnat') : x2))).

Example 7. Let T7 = "For whom and where the three-ton aluminum container has been delivered from?".
Semrepr = Question ((x1 ∧ x2),
Situation (e1, delivery2 *
(Time, certn moment * (Earlier, #now#) : t1)
(Recipient, x1) (Place1, x2)
(Object1, certn container1 * (Weight, 3/ton)
(Material, aluminum) : x3))).

Example 8. Let T8 = "How many people did participate in the creation of the textbook on statistics?". Then
Semrepr = Question (x1, ((x1 Numb(S1))
∧ Qual-compos (S1, person) ∧
Description (certn person * (Element, S1) : y1,
Situation (Õ1, participation1 *
(Time, certn moment * (Earlier, #now#) : t1)
(Agent1, y1) (Type-of-activity, creation1 *
(Product1, certn textbook * (Area1, statistics) : x2))))).

Example 9. Let T9 = "How many times Mr. Stepan Semenov flew to Mexico?". Then
Semrepr = Question (x1, ((x1 Numb (S1))

∧ Qual-compos (S1, sit) ∧
Description (arbitrary sit * (Element, S1) : e1,
Situation (e1, flight * (Time, certn moment *
(Earlier, #now#) : t1) (Agent1, certn person *
(Name, 'Stepan')(Surname, 'Semenov') : x2)
(Place2, certn country * (Name, 'Mexico') : x3))))).

A Transformer of Natural Language Knowledge Descriptions into OWL-expressions

In the context of transforming step by step the existing Web into Semantic Web (see Fomichov 2008, 2009), the need for large Web-based and interrelated collections of formally represented pieces of knowledge covering many fields of professional activity is a weighty ground for increasing the interest of the researchers to the problem of automated formation of ontologies.

It seems that the most obvious and broadly applicable way is to construct a family of natural language processing systems being able to transform the descriptions of knowledge pieces in NL (in English, Russian, German, French, Chinese, Japanese, etc.) into the OWL-expressions and later, possibly, into the expressions of an advanced formalism for developing ontologies.

This idea underlay the design of the computer system NL-OWL1, it is a Russian-language interface implementing a modification of the algorithm of semantic-syntactic analysis *SemSyn* stated in (Fomichov 2005a). The main directions of expanding the input language of the algorithm *SemSyn* are as follows:
- the definitions of notions in restricted Russian language can be the input texts of the system;
- a mechanism of processing the homogenious members of the sentence is added to the algorithm of semantic-syntactic analysis;
- a part of input sentences (the descriptions of the events and the definitions of notions) is transformed not only into the K-representations (i.e. into the expressions of a certain SK-language) but also, at the second stage, into the OWL-expressions.

Let's consider the examples illustrating the principles of processing NL-texts by the experimental Russian-language interface NL-OWL1, implemented in the Web programming system PHP.

Example 1. Definition: "Carburettor is a device for preparing a gas mixture of petrol and air"

K-representation:

ModuleOfKnowledge (definition; carburettor; x1;
(Definition1 (certn carburettor : x1, certn device : x2) ∧
Purpose (x2, certn preparation1 : x_e1) ∧
Description (preparation1, Object1 (certn mixture *
(Type, gas) : x3)) ∧ Product1 (x3, certn petrol : x4)
∧ Product1 (x3, air))))

OWL-expression:

```
< owl   Class  rdf   ID = "ModuleOfKnowledge"/>
< hasFormModule  rdf   resource = "#definition"/>
< hasConcept  rdf   resource = "#carburettor_x1"/>
< owl   Class  rdf   ID = "Action"/>
< Action  rdf   ID = "Concept"/>
< owl   Class  rdf   ID = "Situation"/>
<Situation  rdf   ID = "x1">
< has Action  rdf   resource = "#Concept"/>
< hasDetermination  rdf   resource = "#Device_x2"/>
</Situation>
< owl   Class  rdf   ID = "Device"/>
< Device  rdf   ID = "Device_x2"/>
< Destination  rdf   resource = "#preparation_x_e1"/>
</Device>
< owl   Class  rdf   ID = "Destination"/>
< Destination  rdf   ID = "Destination_x_e1"/ >
< Object1  rdf   resource = "#Mixture_x3"/>
</Destination>
< owl   Class  rdf   ID = "Mixture"/>
< Mixture  rdf   ID = "Mixture_x3"/>
< Form  rdf   resource = "#gas"/>
< Product1  rdf   resource = "#Petrol_x4"/>
< Product1  rdf   resource = "#Air_x5"/>
```

</Mixture>

Due to the broad expressive possibilities of SK-languages, the intelligent power of the transformer NL-OWL1 can be considerably enhanced. That is why the formal methods underlain the design of the system NL-OWL1 enrich the theoretical foundations of the Semantic Web project.

References

1. Angelova, G. (2005); Language technology meet ontology acquisition. In: Dau, F., Mugnier, M.-L., Stumme, G. (eds.) International Conference on Conceptual Structures 2005. LNCS 3596, Springer, Heidelberg (pp. 367-380).
2. Berners-Lee, T., Hendler, J. and O. Lassila (2001); The Semantic Web; Scientific American (pp. 34-43)
3. Berners-Lee, T., Hall, W., Hendler, J.A., O'Hara, K., Shadbolt, N. and D.J. Weitzner (2006); A Framework for Web Science; Foundations and Trends in Web Science, Vol. 1, No. 1, now Publishers Inc. (pp. 134)
4. Fomichov, V.A. (1996); A mathematical model for describing structured items of conceptual level; Informatica. An Intern. J. of Computing and Informatics (Slovenia), 1996, Vol. 20, No, 1. P. 5-32.
5. Fomichov, V.A. (1998); Theory of Restricted K-calculuses as a Comprehensive Framework for Constructing Agent Communication Languages; Special Issue on NLP and Multi-Agent Systems. In: Fomichov V.A., Zeleznikar A.P. (eds.). Informatica. An International Journal of Computing and Informatics (Slovenia). Vol. 22, No 4 (pp. 451-63)
6. Fomichov, V.A. (2000); An ontological mathematical framework for electronic commerce and semantically-structured Web. In: Zhang Y., Fomichov V.A., Zeleznikar A.P. (eds.), Special Issue on Database, Web, and Cooperative Systems. Informatica. An International Journal of Computing and Informatics (Slovenia, Europe), Vol. 24, No. 1, 39--49 (2000)
7. Fomichov, V.A. (2002a); Theory of K-calculuses as a Powerful and Flexible Mathematical Framework for Building Ontologies and Designing Natural Language Processing Systems (ed. Troels Andreasen, Amihai Motro, Henning Christiansen, Henrik Legind Larsen), Flexible Query Answering Systems. 5th International Conference, FQAS 2002, Copenhagen, Denmark, October 27 - 29, 2002. Proceedings; LNAI 2522 (Lecture Notes in Artificial Intelligence, Vol. 2522), Springer Verlag (pp. 183-196)
8. Fomichov, V.A. (2002b); Mathematical foundations of representing texts' meanings for the development of linguistic informational technologies. Part 1. A model of the system of primary units of conceptual level. Informational Technologies (Moscow) No. 2002/10 (pp. 16-25)
9. Fomichov, V.A. (2002c); Mathematical foundations of representing texts' meanings for the development of linguistic informational technologies. Part 2. A system of the rules for constructing semantic representations of phrases and complex discourses. Informational Technologies (Moscow), No. 2002/11 (pp. 34-45)
10. Fomichov, V.A. (2005a); The Formalization of Designing Natural Language Processing Systems. Moscow, MAX Press; in Russian (pp. 368)

11. Fomichov, V.A. (2005b); Standard K-Languages as a Powerful and Flexible Tool for Building Contracts and Representing Contents of Arbitrary E-Negotiations. In: Bauknecht K., Proell B., Werthner H. (eds.), The 6th Intern. Conf. on Electronic Commerce and Web Technologies "EC-Web 2005", Copenhagen, Denmark, Aug. 23 - 26, Proceedings. Lecture Notes in Computer Science, Vol. 3590, Springer Verlag (pp. 138-47)
12. Fomichov, V.A. (2007); Mathematical Foundations of Representing the Content of Messages Sent by Computer Intelligent Agents; Moscow, State University – Higher School of Economics, The Publishing House "TEIS", 2007 (in Russian).
13. Fomichov, V.A. (2008); A comprehensive mathematical framework for bridging a gap between two approaches to creating a meaning-understanding Web; International Journal of Intelligent Computing and Cybernetics (Emerald Group Publishing Limited, UK). 2008, Vol. 1, No. 1 (pp. 143-63)
14. Fomichov, V.A. (2009); Theory of K-representations as a Source of an Advanced Language Platform for Semantic Web of a New Generation; On-line Proceedings of the First International Conference on Web Science, Athens, Greece, March 18-20, 2009; available at http://journal.webscience.org/221/1/websci09_submission_128.pdf
15. Fomichov, V.A. (2010); Semantics-Oriented Natural Language Processing: Mathematical Models and Algorithms; New York, Berlin, Heidelberg, Springer US (pp. 348; in press)
16. Hendler, J., Shadboldt, N., Hall, W., Berners-Lee, T. and D. Weitzner, D. (2008); Web science: an interdisciplinary approach to understanding the web. Communications of the ACM, 2008, Vol. 51, No. 7 (pp. 60-69)
17. Shadboldt, N., Hall, W., Berners-Lee, T. (2006); Semantic Web Revisited. IEEE Intelligent Systems, Vol. 21, No. 3.
18. Wilks, Y. and C. Brewster (2006); Natural Language Processing as a Foundation of the Semantic Web; Foundations and Trends in Web Science, Vol. 1, No. 3, now Publishers Inc.