Chapter 6
The Significance of a New Mathematical Model for Web Science, E-science, and E-commerce

Abstract The significance of the theory of K-representations for e-science, e-commerce, and Web science is shown. The following possibilities provided by SK-languages are analyzed: building semantic annotations of Web-sources and Web-services, constructing high-level conceptual descriptions of visual images, semantic data integration, and the elaboration of formal languages intended for representing the contents of messages sent by computer intelligent agents (CIAs). It is also shown that the theory of SK-languages opens new prospects of building formal representations of contracts and records of commercial negotiations carried out by CIAs. A theoretically possible strategy of transforming the existing Web into Semantic Web of a new generation is proposed.

6.1 The Problem of Semantic Data Integration

6.1.1 The Purpose of Semantic Data Integration in E-Science and Other E-Fields

In the modern world, the objective demands of science and technology often urge research groups in different countries to start the projects aimed at solving the same or similar tasks. A considerable part of obtained results is available via Web. Since the obtained results of the studies can be expressed in different formats, it is important to elaborate the software being able to semantically integrate the data stored in Web-documents, that is, to present the meaning of available documents in a unified format.

As a consequence, a researcher or research group starting to investigate a problem will be able to quickly get adequate information about the state of affairs in the field of interest.

On the other hand, the problem of semantic data integration emerges in connection with numerous practical tasks. Imagine, for instance, that it is required to

V.A. Fomichov, *Semantics-Oriented Natural Language Processing*, IFSR International Series on Systems Science and Engineering 27, DOI 10.1007/978-0-387-72926-8_6,

elaborate an itinerary for a ship transporting goods across an ocean and several seas. In this case, it is necessary to take into account the geophysical, economical, and political data about many areas of the world, in particular, about many areas of the ocean and several seas.

One of the fields of professional activity where the necessity of semantic data integration is most acute is *e-science*. This term (it emerged only in this decade) unifies the studies in different fields of science based on the extensive use of large volumes of obtained data stored on the Web-servers. This applies, in particular, to the studies on bioinformatics, physics, ecology, life sciences [19, 35, 116, 117, 202, 209, 212].

In e-science, a considerable part of obtained results is described in natural language texts being available via Web: in scientific articles, technical reports, encyclopedic dictionaries, etc. That is why an important demand of e-science consists in developing the formal means allowing for representing in a unified format both the meanings of NL-texts and the pieces of knowledge about the application domains.

It is one of the principal goals of *semantic e-science* – a subfield of e-science aimed at creating the semantic foundations of e-science [25, 115].

6.1.2 Ontologies in Modern Information Society

The notion of ontology is one of the most important notions for the studies on semantic data integration.

An ontology can be defined as a specification of a conceptualization [110]. The term “conceptualization” is used for indicating a way an intelligent system structures its perceptions about the world. A specification of a conceptualization gives a meaning to the vocabulary used by an intelligent system for processing knowledge and interacting with other intelligent systems.

In the last decade, one has been able to observe a permanent growth of interest in building and studying ontologies. The reason is that the researchers and systems developers have become more interested in reusing or sharing knowledge across systems. Different computer systems use different concepts and terms for describing application domains. These differences make it difficult to take knowledge out of one system and use it in another. Imagine that we are able to construct ontologies that can be used as the basis for multiple systems. In this case different systems can share a common terminology, and this will facilitate sharing and reuse of knowledge.

In a similar way, if we are able to create the tools that support merging ontologies and translating between them, then sharing knowledge is possible even between systems based on different ontologies.

The main source for automatically building ontologies is a great amount of available texts in natural language (NL). Taking this into account, we need the powerful formal means for building semantic representations (SRs) of (a) NL-definitions of concepts and (b) sentences and discourses in NL expressing knowledge of other kinds about an application domain.

The analysis of formal approaches to representing knowledge provided by the Theory of Conceptual Graphs, Episodic Logic, and Description Logic shows that these approaches give formal means with rather restricted expressive possibilities as concerns building SRs of definitions of notions and SRs of sentences and discourses representing the pieces of knowledge about the world.

That is why we need to have much more powerful and convenient formal means (in comparison with the broadly used ones) for describing structured meanings of natural language (NL) texts and, as a consequence, for building ontologies.

6.1.3 The Language UNL and the Problem of Sharing Knowledge

Since the second half of the 1990s, one has been able to observe the progress of two parallel approaches to adding to the existing Web the ability to understand the meanings of electronic documents.

On one hand, it has been the activity of the research laboratories of the World Wide Web Consortium (W3C) and a number of other research centers in the world aimed at developing a Semantic Web. Though officially the task of creating Semantic Web was announced in the beginning of 2001, the possibility to pose this task was created by a number of preliminary studies resulted, in particular, in the development of Resource Description Framework (RDF) – a language for describing the metadata about informational sources and RDF Schema Specification Language (RDFS). In the first decade of this century, RDF and RDFS became the basis for the development of DAML + OIL and its successor OWL – two languages intended for constructing ontologies [123, 136, 186, 204].

On the other hand, the following fundamental problem has emerged in the mid-1990s: how to eliminate the language barrier between the end users of the Internet in different countries. For solving this problem, H. Uchida et al. [199] proposed a new language-intermediary, using the words of English language for designating informational units and several special symbols. This language, called the Universal Networking Language (UNL), is based on the idea of representing the meanings of separate sentences by means of binary relations.

The second motive for the elaboration of UNL was an attempt to create the language means allowing for representing in one format the various pieces of knowledge accumulated by mankind and, as a consequence, to create objective preconditions for sharing these pieces of knowledge by various computer systems throughout the world [211].

Since 1996, UNO has been funding a large-scale project aimed at the design of a family of natural language processing systems (NLPSs) transforming the sentences in various natural languages into the expressions of UNL and also transforming the UNL-expressions into sentences in various natural languages. For several years the coordinator of this project was the UNO Institute for Advanced Studies by the Tokyo University. At the moment, under the framework of this project, the NLPSs for six official UNO languages are being elaborated (English, Arabic, Spanish, Chinese, Russian, and French), and also for nine other languages, including Japanese, Italian, and German. Since the beginning of the 2000s, the studies in this direction have been coordinated by the Universal Networking Digital Language Foundation.

The initially scheduled duration of the UNL project started in 1996 is 10 years. That is why it is just the time to analyze the achieved results and to take the right decisions concerning further studies in this direction. It is shown in the papers [84, 90, 93, 94] with respect to the online monographs [200, 201] that the expressive possibilities of UNL are rather restricted.

First of all, the language UNL is oriented at representing the contents of only separate sentences but not arbitrary discourses. Even the UNL specifications published in 2006 don't contain a theory of representing the meanings of discourses.

That is why in the papers [84, 90, 93, 94] it is proposed to interpret the language UNL (despite the linguistic meaning of its title) as a semantic networking language of the first generation.

With respect to the fact that the expressive power of UNL is rather restricted, it seems reasonable to look for another, more powerful formal approach to describing structured meanings of natural language texts with the aim to find (if possible) a model for constructing a universal or broadly applicable semantic networking language for adding a meaning-understanding ability to the existing Web and for contributing to semantic integration of Web data.

6.2 Building Semantic Annotations of Web Data

The analysis of a number of publications studying the problem of transforming the existing Web into Semantic Web allows for drawing the following conclusion: an ideal configuration of Semantic Web would be a collection of interrelated resources, where each of them has both an annotation in natural language (NL) and a formal annotation reflecting the meaning or generalized meaning of this resource, i.e. *a semantic annotation*. NL-annotations would be very convenient for the end users, and semantic annotations would be used by question-answering systems and search engines.

Most likely, the first idea concerning the formation of semantic annotations of Web data would be to use the formal means for building semantic representations of NL-texts provided by mathematical and computational linguistics.

However, the analysis shows that the expressive power of the main popular approaches to building SRs of NL-texts, in particular, of Discourse Representation Theory, Theory of Conceptual Graphs, and Episodic Logic is insufficient for effectively representing contents of arbitrary Web data, in particular, arbitrary biological, medical, or business documents.

First of all, the restrictions concern describing semantic structure of (a) infinitives with dependent words (e.g., representing the intended manner of using things and procedures); (b) constructions formed from the infinitives with dependent words by means of the logical connectives "and," "or," "not"; (c) the complicated designations of sets; (d) the fragments where the logical connectives "and," "or" join not the designations of assertions but the designations of objects ("the product A is distributed by the firms B1, B2, ..., BN"); (e) the explanations of the terms being unknown to an applied intelligent system; (f) the fragments containing the references to the meanings of phrases or larger fragments of a discourse ("this method," etc.); (g) the designations of the functions whose arguments and/or values may be the sets of objects ("the staff of the firm A," "the number of the suppliers of the firm A," etc.).

Taking into account this situation and the fact that the semantic annotations of Web-sources are to be compatible with the format of representing the pieces of knowledge in ontologies, a number of researchers undertook the efforts of constructing computer intelligent systems, using the languages RDF, RDFS, or OWL for building semantic annotations of Web-sources [161, 176].

However, the expressive power of RDF, RDFS, or OWL is insufficient for being an adequate formal tool of building semantic annotations of scientific papers, technical reports, etc.

Meanwhile, the formulated idea of where to get the formal means for building semantic annotations from is correct. The main purpose of this section is to set forth the principal ideas of employing the SK-languages for building semantic annotations of informational sources, in particular, Web-based sources.

Example. Let's consider a possible way of employing SK-languages for building a semantic annotation of the famous paper "The Semantic Web" by T. Berners-Lee, J. Hendler, and O. Lassila published in "Scientific American" in May 2001 [18].

Suppose that there is a Web-source associating the following NL-annotation with this paper:

> It is proposed to create such a net of Web-based computer intelligent agents (CIAs) being able to understand the content of almost every Web-page that a considerable part of this net will be composed by CIAs being able to understand natural language. The authors consider the elaboration of ontologies as a precondition of sharing knowledge by CIAs from this net and believe that it is reasonable to use the languages RDF and RDFS as primary formal tools for the development of ontologies of the kind.

A semantic annotation corresponding to this NL-annotation can be the K-string (or l-formula) of the form

$$certn\, inf.ob * (Kind1, sci\_article)(Source1,$$

$$certn\, journal1 * (Name1, \text{“}Scientific\_American\text{”}) : x1)$$

$$(Year, 2001)(Month, May)(Authors,$$

$$certn\, group1 * (Numb, 3)(Elements1,$$

$$(\langle 1, certn\, scholar * (First\_name, \text{“}Tim\text{”})(Surname,$$

$$\text{“}Berners - Lee\text{”}) : x2\rangle \wedge \langle 2, certn\, scholar*$$
$$(First\_name, \text{“}James\text{”})(Surname, \text{“}Hendler\text{”}) : x3\rangle$$
$$\wedge \langle 3, certn\, scholar * (First\_name, \text{“}Ora''\,)(Surname,$$
$$\text{“}Lassila''\,) : x4\rangle)) : S1)\,(Central\_ideas,$$
$$(\langle 1, Semrepr1\rangle \wedge \langle 2, Semrepr2\rangle \wedge$$
$$\langle 3, Semrepr3\rangle \wedge \langle 4, Semrepr4\rangle)) : v\,,$$

where the variable $S1$ designates the group consisting of all authors of this article, v is a variable being a mark of the constructed semantic annotation as an informational object, and $Semrepr1$ – $Semrepr4$ are the K-strings defined by the following relationships:

$$Semrepr1 = Proposed(S1, creation1 * (Product1\,,$$
$$certn\, family1 * (Qual - compos, intel\_comp\_agent*$$
$$(Property, web - based)(Ability, understanding1 * (Inf\_object\,,$$
$$Content(almost\_every\, web\_page)))) : S2)$$
$$(Time, certn\, time\_interval * (Part1,$$
$$Nearest\_future(decade1, \#now\#)))),$$

$$Semrepr2 = Proposed(S1, achieving\_situation*$$
$$(Description1, \exists S3(set)(Subset(S3, S2) \wedge$$
$$Estimation1(Numb(S3)/Numb(S2), considerable1) \wedge$$
$$Qual - compos(S3, intel\_comp\_agent * (Property\,,$$
$$web - based)(Ability, understanding1 * (Inf\_object\,,$$
$$almost\_every\, text * (Language1, certn\, language*$$
$$(Belong, NL\_family)))))))),$$

$$Semrepr3 = Believe(S1, Precondition(elaboration*$$
$$(Product1, certn\, family1 * (Qual - compos\,,$$
$$ontology) : S4), knowledge\_sharing*$$
$$(Group\_of\_intel\_systems, S2)),$$

$$Semrepr4 = Believe(S1, Reasonable(\#now\#\,,$$

$$using * (Object1, (RDF \wedge RDFS))$$
$$(Role1, primary\,formal\_tool)(Purpose,$$
$$elaboration1 * (Product1, S4)))).$$

To sum up, a comprehensive formal tool for building semantic annotations of Web data is elaborated. This tool is the theory of SK-languages. A very important additional expressive mechanism of SK-languages in comparison with the mechanisms illustrated in the example above is the convenience of building semantic representations of discourses with references to the meanings of phrases and larger parts of a discourse.

The analysis of expressive power of the class of SK-languages allows for conjecturing that it is both possible and convenient to construct semantic annotations of arbitrary Web data by means of SK-languages. That is why the theory of SK-languages can be interpreted as a powerful and flexible (likely, universal) formal metagrammar of semantic annotations of Web data.

6.3 Conceptual Descriptions of Visual Images

It is interesting that SK-languages, initially developed for representing structured meanings of NL-texts, open also new prospects for building high-level conceptual descriptions of visual images. This fact can be explained very easily: we are able to use NL for constructing high-level conceptual descriptions of visual images. Since the expressive possibilities of SK-languages are very rich, it is possible to use SK-languages for building conceptual descriptions of arbitrary visual images.

Example. We are able to describe Fig. 6.1 as follows:

The scene contains two groups of objects. The quadrant of the scene including the upper-left corner contains a figure being similar to a ellipse; this ellipse is formed by eight squares, the side of each square is 1 cm.

The quadrant of the scene including the bottom-right corner contains a figure being similar to a rectangle formed by 10 circles, the diameter of each circle is 1 cm. The longer sides of this rectangle are horizontally oriented, each of them consists of 4 circles.

This meaning can be expressed by the following K-string:

$$Number-of-groups(certn\,scene1 : x1, 2) \wedge$$
$$Groups(x1, (Gr1 \wedge Gr2)) \wedge Isolated(Gr1, Gr2)$$
$$\wedge Loc(Gr1, top-left-quadrant(x1)) \wedge$$
$$Similar\_shape(Gr1, certn\,ellipse : z1) \wedge$$
$$(Horiz\_diameter(z1) \equiv Multip\ (0.5, Length(x1))) \wedge$$
$$(Vertic\_diameter(z1) \equiv Multip\ (0.25, Length(x1))) \wedge$$

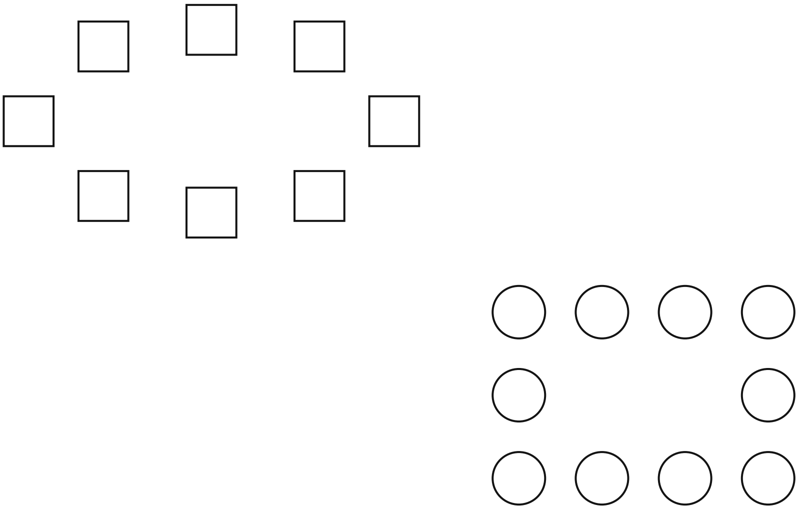

Fig. 6.1 An example of a visual scene to be associated with a high-level conceptual description being a K-string

$$Loc(Gr2, bottom - right - quadrant(x1)) \wedge$$
$$Similar\_shape(Gr2, certn\, rectangle : z2) \wedge$$
$$(Height(z2) \equiv Multip\ (0.5, Height(x1))) \wedge$$
$$(Length(z2) \equiv Multip\ (0.4, Length(x1))) \wedge$$
$$Indiv - composition(Gr1, (x1 \wedge x2 \wedge \ldots \wedge x8))$$
$$\wedge Is1((x1 \wedge x2 \wedge \ldots \wedge x8),$$
$$square1 * (Side1, Multip\ (0.06, Length(x1))) \wedge$$
$$Indiv - composition(Gr2, (y1 \wedge y2 \wedge \ldots \wedge y10))$$
$$\wedge Is1((y1 \wedge y2 \wedge \ldots \wedge y10),$$
$$circle * (Diameter, Multip\ (0.055, Length(x1))) \wedge$$
$$(Object\_number(every\, side1 * (Part1, z2)(Orient, horiz)) \equiv 4))\,.$$

Though the image presented on Fig. 6.1 is rather simple, the used method of building a high-level conceptual description of this image in the form of a K-string is rather general. This method is as follows:

1. Distinguish the principal groups of objects presented on the image and their number.
2. Calculate the positions of these groups.
3. Describe a shape (if possible) of each group.
4. Describe the number and a shape of objects being the elements of each group.
5. Describe the connections of distinguished groups of objects.

Thus, the theory of K-representations can be used in the design of multi media databases (with Web access too) for building in the same form, as the expressions of SK-languages, both semantic annotations of textual informational sources and high-level conceptual descriptions of visual images, in particular, being the components of textual informational sources.

6.4 Representation of Knowledge in Biology and Ecology

Let's consider a number of new important possibilities of building formal definitions of concepts provided by standard K-languages. If T is an expression in NL and a string E from an SK-language can be interpreted as a semantic representation (SR) of T, then E is called a K-representation (KR) of the expression T.

Example 1. Let Def1 = "A flock is a large number of birds or mammals (e.g. sheep or goats), usually gathered together for a definite purpose, such as feeding, migration, or defence." Def1 may have the K-representation $Expr1$ of the form

$$Definition1(flock, dynamic - group * (Compos1,$$

$$(bird \vee mammal * (Examples, (sheep \wedge goal)))), S1,$$

$$(Estimation1(Quantity(S1), high) \wedge$$

$$Goal - of - forming(S1,$$

$$certn\, purpose * (Examples,$$

$$(feeding \vee migration \vee defence))))).$$

Example 2. The definition Def1 is taken from a certain book published in a certain year by a certain publishing house. The SK-languages allow for building SRs of definitions in an object-oriented form reflecting their external connections. For instance, object-oriented SR of the definition Def1 can be the expression

$$certn\, inform - object * (Kind, definition)$$

$$(Content1, Expr1)(Source1, certn\, dictionary*$$

$$(Title, Longman\_Dictionary\_of\_Scientifc\_Usage)$$

$$(Publishing\_house, (Longman\_Group\_Limited/Harlow$$

$$\wedge Russky\_Yazyk\_Publishers/Moscow))$$

$$(City, Moscow)(Year, 1989)).$$

Example 3. Let T1 = "All granulocytes are polymorphonuclear; that is, they have multilobed nuclei." Then T1 may have the following K-representation:

$$(Property(arbitr\, granulocyte\ :\ x1,$$
$$polymorphonuclear)\ :\ P1 \wedge Explanation(P1,$$
$$If - then(Have1(x1,\ certn\, nucleus\ :\ x2),$$
$$Property(x2,\ multilobed)))).$$

Here $x1$ is the variable marking an arbitrary granulocyte, $x2$ is the variable marking the nucleous of the granulocyte $x1$, and $P1$ is the variable marking the meaning of the first phrase of T1.

Example 4. Consider the text D1 = "An adenine base on one DNA strand links only with a thymine base of the opposing DNA strand. Similarly, a cytosine base links only with a guanine base of the opposite DNA strand."

For constructing a KR of D1, the following remark may be helpful. A molecule of deoxyribonucleic acid (a DNA molecule) is composed of thousands of nucleotides (combinations of three basic elements: deoxyribose, phosphate, and a base). There are four kinds of bases: adenine, guanine, cytosine, and thymine. The nucleotides of a DNA molecule form a chain, and this chain is arranged in two long strands twisted around each other.

Taking into account this remark, one can associate with the first sentence of D1 a KR $Semrepr1$ of the form

$$\forall x1\,(dna - molecule)(Link(arbitr\, base1 * (Is1,\ adenine)$$
$$(Part,\ arbitr\, strand1 * (Part,\ x1) : y1) : z1,\ certn\, base1*$$
$$(Is1,\ thymine)(Part,\ certn\, strand1 * (Part,\ x1)$$
$$(Opposite,\ y1) : y2) : z2) \wedge$$
$$\neg\exists z3\,(base1)(Is1(z3,\ \neg thymine) \wedge$$
$$Part(z3,\ y2) \wedge Link(z1,\ z3) : P1.$$

In the string $Semrepr1$, the variables $y1$ and $y2$ are used to mark the descriptions of two strands of arbitrary DNA molecule $x1$; the variables $z1$, $z2$, $z3$ mark the bases.

The variable $P1$ (with it the sort "meaning of proposition" is associated) is used to mark the semantic representation of the first sentence of the discourse D1. This allows for building a compact semantic representation of the second sentence of D1, because the occurrence of the word "similarly" in the second sentence of D1 indicates the reference to the meaning of the first sentence.

In particular, the second sentence of D1 in the context of the first sentence may have a K-representation $Semrepr2$ of the form.

$$(Similarly(P1,\ P2) \wedge (P2 \equiv \forall x1\,(dna - molecule)$$
$$(Link(arbitr\, base1 * (Is1,\ cytosine)$$
$$(Part,\ arbitr\, strand1 * (Part,\ x1)\ :\ y3)\ :\ z4,$$

$$certn\, base1 * (Is1, guanine)(Part,$$
$$certn\, strand1 * (Part, x1)(Opposite, y3) \,:\, y4) \,:\, z5) \wedge$$
$$\neg \exists z6(base1)(Is(z6, \neg guanine)$$
$$\wedge Part(z6, y4) \wedge Link(z4, z6))))).$$

Then we can associate with the text D1 the K-string *Semrepr*3 of the form

$$(Semrepr1 \wedge Semrepr2),$$

where *Semrepr*1 and *Semrepr*2 are the K-strings defined above. Such string can be interpreted as a possible K-representation of the discourse D1.

The K-string *Semrepr*3 illustrates an important opportunity afforded by SK-languages: to mark by variables the fragments of K-strings being semantic representations of narrative texts, infinitive groups, or questions. This opportunity allows us to effectively describe structured meanings of discourses with references to the meanings of fragments being statements, infinitive groups, or questions.

The presence of such references in discourses is often indicated by the following words and word combinations: "this recommendation," "for instance," "e.g.," "that is," "i.e.," "the idea discussed above," "in other words," etc.

The constructed KR *Semrepr*3 of the discourse D1 illustrates several additional original features of K-strings (besides features discussed above). First, the symbol $\equiv$ connects a variable and a semantic representation of a sentence. Second, the symbol of negation $\neg$ can be connected with designations of notions. In such a way the substrings $\neg thymine$ and $\neg guanine$ are built.

Some additional useful properties of SK-languages are analyzed below.

6.5 Representation of Knowledge in Medicine

Example 1. Let T1 be the definition "The Eustachian tube is a canal leading from the middle ear to the pharynx." One can associate with T1, in particular, the following K-string interpreted as a semantic representation of T1:

$$Definition1(Eustachian - tube, canal1, x1,$$
$$\exists z\,(person)\,Lead1(x1,\; certn\, middle - ear *$$
$$(Part, z),\; certn\, pharynx * (Part, z))).$$

Example 2. If T2 = "Sphygmomanometer is instrument intended to measure blood pressure," then T2 may have a K-representation

$$(sphygmo - manometer \equiv instrument * (Purpose1, measuring1 *$$
$$(Param, blood - pressure)(Subject, any\, person))).$$

Here the semantic item $Purpose1$ is to be interpreted as the name of a binary relation. If a pair (A, B) belongs to this relation, A must be a physical object, and B must be a formal semantic analogue of an infinitive group expressing the intended manner of using this physical object.

Example 3. Let T3 be the definition "Thrombin is an enzyme helping to convert fibrinogen to fibrin during coagulation." Then the K-string

$$(thrombin \equiv enzyme * (Purpose1, helping * (Action,$$

$$converting1 * (Object1,\ certn\, fibrinogen)(Result1,\ certn\, fibrin)$$

$$(Process,\ any\, coagulation))))$$

can be interpreted as a possible KR of T3.

6.6 Representation of Semantic Content in Business

Let's demonstrate the possibility to represent semantic content of business sentences and discourses with references to the meanings of the fragments being phrases or larger parts of the discourse.

Example 1. If T1 = "Freight forward is a freight to be paid in the port of destination" then T1 may have a KR of the form

$$(freight - forward \equiv freight * (Description,$$

$$\langle x1, Payment - at(x1,\ certn\, port1 * (Destination - of, x1))\rangle)).$$

The element *certn* is interpreted as the informational item corresponding to the word combination "a certain."

Example 2. Let T2 be the following definition: "A small or medium enterprise is a company with at most 50 employees". This definition may have a K-representation

$$Definition(small\_med\_enterpr, \forall x1\, (company1)\, (Is1(x1,\ small\_med\_enterpr) \equiv$$

$$\neg\, Greater(Number(Employees(x1)), 50))))$$

or a K-representation

$$((small\_med\_enterpr \equiv company1 * (Description, P1)) \wedge$$

$$(P1 \equiv \forall x1\, (company1)(Is1(x1,\ small\_med\_enterpr) \equiv$$

$$\neg\, Greater(Number(Employees(x1)), 50)))).$$

Example 3. If T3 = "Mr. Green had asked to send him three containers with ceramics. That request was fulfilled on March 10," a possible KR of T3 may be constructed as follows:

$$((Situation(e1,\ request*(Agent1,$$

$$certn\,man*(Name,\ \text{`}Green\text{'})\ :\ x1)(Goal,$$

$$sending1*(Object1, certn\,set*(Number,\ 3)$$

$$(Qualitative-composition,\ container1*$$

$$(Contain1,\ ceramics)):x2)(Time,\ t1)))$$

$$\wedge Before(t1,\ \#now\#)):P1\wedge$$

$$Fulfilled(certn\,request*(Description,\ P1):x3,t2)\wedge$$

$$(t2\equiv\langle 10,\ 03,\ current-year\rangle)\wedge$$

$$Before(t2,\ \#now\#)\wedge Before(t1,\ t2)).$$

6.7 SK-Languages as a Tool for Building E-Contracts

6.7.1 Formal Languages for E-Contracting

During the last several years in E-commerce two interrelated fields of researches have emerged called e-negotiations and electronic contracting. The birth of these fields was formally denoted by means of the organization at the beginning of the 2000s of several international conferences and workshops.

The collection of central problems faced by the researchers in these fields includes the creation of formal languages for representing contents of the records of negotiations conducted by computer intelligent agents (CIAs) and for forming contracts concluded in the course of such negotiations. These tasks can be considered as important particular cases of the problem of constructing general-purpose formal languages for business communication [118, 148].

Hasselberg and Weigand underline in [118] that if the messages in the field of E-commerce are to be processed automatically, the meaning must be formalized. This idea coincides with the opinion of Kimbrough and Moore [148] about the necessity of developing logical-semantic foundations of constructing formal languages for business communication (FLBC).

It is suggested in [145–148] to use first-order logic insofar as possible and reasonable for expressions in any FLBC. However, the expressive possibilities of the class of first-order logic languages are very restricted as concerns describing semantic structure of arbitrary business documents.

The analysis shows that the records of commercial negotiations and contracts can be formed with the help of expressive means of natural language (NL) used for the construction of arbitrary NL-texts pertaining to medicine, technology, law, etc. In particular, the texts from such documents may include

(a) the infinitives with dependent words expressing goals, offers ("to sell 30 boxes with apples"), promises, commitments, or intended manners of using things;
(b) the constructions formed from the infinitives with dependent words by means of the logical connectives "and," "or," "not," and expressing compound designations of goals, offers, promises, commitments or destinations of things;
(c) complex designations of sets ("a consignment consisting of 50 boxes with apples");
(d) the fragments where the logical connectives "and," "or" join not the designations of assertions but the designations of objects;
(e) the fragments containing the references to the meanings of phrases or larger fragments of a discourse ("this proposal," "that order," "this promise," etc.);
(f) the designations of the functions whose arguments and/or values may be the sets of objects ("the staff of the firm A," "the suppliers of the firm A," "the number of the suppliers of the firm A");
(g) the questions with the answer "Yes" or "No";
(h) the questions with interrogative words.

Meanwhile, the first-order predicate logic provides no possibility to build the formal analogues (on the semantic level) of the texts from business documents where the NL phenomena listed in items (a)–(g) are manifested.

That is why the problem of developing formal languages allowing for representing contents of the records of commercial e-negotiations carried out by CIAs and for forming contracts concluded in the course of such negotiations is very complicated. Hence it seems to be reasonable to use for solving this problem the most broadly applicable theories (ideally, universal) of representing meanings of NL-texts provided by mathematical linguistics and mathematical computer science.

6.7.2 The Possibilities of Forming Contracts and Records of E-Negotiations by Means of SK-Languages

The analysis shows that the SK-languages possess the expressive possibilities being necessary and sufficient for representing in a formal way the contents of contracts and of the records of commercial negotiations.

For illustrating an important part of such possibilities, let's consider a multi-partner scenario of the interaction of business partners in the course of handling a car damage claim by an insurance company (called AGFIL). The names of the involved parties are Europe Assist, Lee Consulting Services (Lee C.S.), Garages, and Assessors. Europe Assist offers a 24-h emergency call answering service to the policyholders. Lee C.S. coordinates and manages the operation of the emergency service on a day-to-day level on behalf of AGFIL. Garages are responsible for car repair. Assessors conduct the physical inspections of damaged vehicles and agree upon repair figures with the garages.

The process of a car insurance case can be described as follows: The policyholder phones Europe Assist using a free phone number to notify a new claim. Europe Assist will register the information, suggest an appropriate garage, and notify AGFIL which will check whether the policy is valid and covers this claim. After AGFIL receives this claim, AGFIL sends the claim details to Lee C.S. AGFIL will send a letter to the policyholder for a completed claim form. Lee C.S. will agree upon repair costs if an assessor is not required for small damages, otherwise an assessor will be assigned. The assessor will check the damaged vehicle and agree upon repair costs with the garage.

After receiving an agreement of repairing car from Lee C.S., the garage will then commence repairs. After finishing repairs, the garage will issue an invoice to Lee C.S., which will check the invoice against the original estimate. Lee C.S. returns all invoices to AGFIL. This firm processes the payment. In the whole process, if the claim is found invalid, all contractual parties will be contacted and the process will be stopped [31, 210].

This scenario provides the possibility to illustrate some properties of SK-languages making them a convenient tool for formally describing contracts.

Property 1. The possibility to build compound designations of goals.

Example 1. Let T1 = "The policyholder phones Europe Assist to inform about a car damage." Then T1 may have the following K-representation (KR), i.e., a semantic representation being an expression of a certain SK-language:

$$Situation(e1,\ phone-communic * (Agent1,\ certn\,person*$$
$$(Hold1,\ certn\,policy1\ :\ x1)\ :\ x2)(Object2,\ certn\,firm1*$$
$$(Name1,\ \text{“}Europe Assist\text{”})\ :\ x3)(Purpose,$$
$$inform-transfer * (Theme1,$$
$$certn\,damage1 * (Object1,\ certn\,car1)\ :\ x4))).$$

Property 2. The existence of the means allowing for representing in a compact way the time and causal relations between the situations.

Property 3. The possibility to construct compact semantic representations of such fragments of sentences which are obtained by means of joining the designations of things, events, concepts, or goals with the help of logical connectives AND, OR.

Example 2. Let T2 = "After receiving a repair invoice from the firm 'Lee C.S.' and a claim from the policyholder, the company 'AGFIL' pays the car repair to the garage." Then a KR of T2 can be the expression

$$(Situation(e1,\ (receiving1 * (Agent2,\ certn\,firm1*$$
$$(Name1,\ \text{“}AFGIL\text{”})\ :\ x1)(Object1,\ certn\,invoice*$$
$$(Theme,\ certn\,repair\ :\ e2)\ :\ x2)(Sender1,$$

$$certn\,firm1 * (Name1,\ \text{“}LeeC.S.\text{”})\ :\ x3) \wedge receiving1*$$
$$(Agent2,\ x1)(Object1,\ certnclaim1\ :\ x4)(Sender1,$$
$$certn\,person * (Hold1,\ certn\,policy1\ :\ x5)\ :\ x6)))\wedge$$
$$Situation(e2,\ payment1 * (Agent2,\ x1)(Addressee1,$$
$$certn\,garage\ :\ x7)(Sum,\ Cost(e2))) \wedge Before(e1,\ e2)).$$

The analysis of additional precious properties of SK-languages (as concerns the applications of the kind) can be found in [87].

6.8 Simulation of the Expressive Mechanisms of RDF, RDFS, and OWL

This section shows how it is possible to simulate the expressive mechanisms of the language systems RDF (Resource Description Framework) [171–173], RDF Schema (RDFS) [174, 175], and OWL (Ontology Web Language) being the basic languages of the Semantic Web project [162–164].

6.8.1 Simulation of the Expressive Mechanisms of RDF and RDFs

Example 1. According to [171], the sentence T1 = “The students in the course 6.001 are Amy, Tim, John, Mary, and Sue” is translated (in some pragmatic context) into the RDF structure

$$\langle rdf\ :\ RDF\rangle\langle rdf\ :\ Description$$
$$about\ =\ \text{“}U1/courses/6.001\text{”}\rangle$$
$$\langle s\ :\ students\rangle\langle rdf\ :\ Bag\rangle$$
$$\langle rdf\ :\ Rliresource\ =\ \text{“}U1/stud/Amy\text{”}/\rangle$$
$$\langle rdf\ :\ Rliresource\ =\ \text{“}U1/stud/Tim\text{”}/\rangle$$
$$\langle rdf\ :\ Rliresource\ =\ \text{“}U1/stud/Jolm\text{”}/\rangle$$
$$\langle rdf\ :\ Rliresource\ =\ \text{“}U1/stud/Mary\text{”}/\rangle$$
$$\langle rdf\ :\ Rliresource\ =\ \text{“}U1/stud/Sue\text{”}/\rangle$$
$$\langle /rdf\ :\ RBag\rangle\langle /s\ :\ students\rangle$$
$$\langle /rdf\ :\ RDescription\rangle\langle /rdf\ :\ RRDF\rangle, where\ U1\ is\ an\ URL$$

In this expression, the item *Bag* is the indicator of a bag container object. It is possible to construct the following similar expression of a certain SK-language:

$$certn\,course1 * (W3ad, \text{“}U1/courses/6.001\text{”})$$

$$(Students,\ certn\,bag * (Compos2,$$

$$(certn\,stud * (W3ad, \text{“}U1/stud/Amy\text{”})\wedge$$

$$certn\,stud * (W3ad, \text{“}U1/stud/Tim\text{”})\wedge$$

$$certn\,stud * (W3ad, \text{“}U1/slud/John\text{”})\wedge$$

$$certn\,stud * (W3ad, \text{“}U1/stud/Mary\text{”})\wedge$$

$$certn\,stud * (W3ad, \text{“}U1/stud/Sue\text{”})))).$$

Here the symbol *certn* is interpreted as the referential quantifier, i.e., as the informational item corresponding to the word combination “a certain” in cases when it is used for building the word combinations in singular (“a certain book,” “a certain personal computer,” etc.).

Example 2. Following [171], the model for the sentence T2 = “The source code for $X11$ may be found at $U3$, $U4$, or $U5$” (where $U3$, $U4$, $U5$ are some URLs) may be written in RDF (with respect to a certain pragmatic context) as

$$\langle rdf\ :\ RDF\rangle$$

$$\langle rdf\ :\ Descriptionabout\ =\ \text{“}U2/packages/X11\text{”}\rangle$$

$$\langle s\ :\ DistributionSite\rangle\langle rdf\ :\ Alt\rangle$$

$$\langle rdf\ :\ liresource\ =\ \text{“}U3\text{”}/\rangle$$

$$\langle rdf\ :\ liresource\ =\ \text{“}U4\text{”}/\rangle$$

$$\langle rdf\ :\ liresource\ =\ \text{“}U5\text{”}/\rangle$$

$$\langle /rdf\ :\ Alt\rangle\langle /s : DistributionSite\rangle$$

$$\langle /rdf\ :\ Description\rangle\langle /rdf\ :\ RDF\rangle.$$

Here the informational item *Alt* is the indicator of an alternative container object. The theory of K-representations suggests the following similar expression:

$$certn\,resource * (W3ad, \text{“}U2/packages/X11\text{”})$$

$$(DistributionSite,\ (certn\,resource * (W3ad, \text{“}U3\text{”})\vee$$

$$certn\,resource * (W3ad, \text{“}U4\text{”}) \vee certn\,resource * (W3ad, \text{“}U5\text{”}))).$$

Example 3. Consider the sentence T3 = “Ora Lassila is the creator of the resource U6” and the corresponding RDF-structure

$$\langle rdf\ :\ RDF\rangle\langle rdf\ :\ Descriptionabout\ =\ \text{“}U6\text{”}$$

$$s\ :\ Creator\ =\ \text{“}OraLassila\text{”}/\rangle\langle /rdf\ :\ RDF\rangle.$$

Using a certain SK-language, we can build the following description of the mentioned resource:

$$certn\,resource * (W3ad, \text{“}U6\text{”})$$
$$(Creator, \text{“}OraLassila\text{”}).$$

Example 4. The theory of K-representations enables us to build reified conceptual representations of statements, i.e. the representations in the form of named objects having some external ties: with the set of the authors, the date, etc. For instance, we can associate the sentence T3 = “Ora Lassila is the creator of the resource U6” with the expression of an SK-language

$$certn\,info-piece * (RDF-type, Statement)$$
$$(Predicate, Creator)(Subject, \text{“}U6\text{”})$$
$$(Object, OraLassila)\ :\ il024,$$

where $il024$ is the name of an information piece.

This form is very close to the RDF-expression [171]

$$\{type,\ [X],\ [RDF\ :\ statement]\}$$
$$\{predicate,\ [X],\ Creator]\}$$
$$\{subject,\ [X],\ [U6]\}$$
$$\{object,\ [X],\ \text{“}OraLassila\text{”}\}.$$

Proceeding from the ideas considered in the examples above and in the previous sections of this chapter, we would be able to approximate all RDF-structures by the similar expressions of RSK-languages.

Example 5. The RDF Schema (RDFS) description of the class “Marital status” from [174]

$$\langle rdfs\ :\ Classrdf\ :\ ID = \text{“}MarStatus\text{”}/\rangle$$
$$\langle MarStatusrdf\ :\ ID = \text{“}Married\text{”}/\rangle$$
$$\langle MarStatusrdf\ :\ ID = \text{“}Divorced\text{”}/\rangle$$
$$\langle MarStatusrdf\ :\ ID = \text{“}Single\text{”}/\rangle$$
$$\langle /rdfs\ :\ Class\rangle$$

can be represented by the following K-string:

$$(any\#MarStatus = (Married \vee Divorced \vee Single)).$$

The theory of K-representations provides the possibility to approximate all RDF-structures by the similar expressions of SK-languages. The same applies to the RDF Schema Specification Language (RDFS) [174, 175]. The analysis of RDF and RDFS expressive means supports such basic ideas of the theory of K-representations as: building compound formal designations of sets; joining by logical connectives not only the designations of assertions but also the designations of things, events, and concepts; considering assertions as objects having some external ties: with a date, the set of the authors, a language, etc.

6.8.2 Simulation of OWL Expressive Mechanisms

OWL (Ontology Web Language) is the principal language for constructing ontologies under the framework of the Semantic Web project [162–164].

It is easy to show that the expressive mechanisms of SK-languages demonstrated above in this and previous chapters allow for simulating all manners of describing the classes of objects in OWL and allow for defining such algebraic characteristics of the properties as reflexivity, symmetricity, and transitivity. Consider only several examples of the kind.

Example 5 above may be interpreted as an illustration of the manner to define the classes in OWL by means of explicit enumeration of all elements of this class. Another illustration is the following K-string being the definition of the notion "a working day of the week":

$$(arbitrary\,working\_day \equiv (Monday \vee$$

$$Tuesday \vee Wednesday \vee Thursday \vee Friday))\,.$$

For defining the classes as the unions or the intersections of other classes, the theory of K-representations enables us to use the relational symbols $Union$ and $Intersection$ with the type

$$\{(\{[entity]\}, \{[entity]\})\}\,,$$

where $[entity]$ is the basic type "entity," for constructing the K-strings (or l-formulas) of the form

$$Union(Y, Z),\ Intersection(Y, Z),$$

where Y and Z are the designations of the classes.

An alternative way of defining the unions of classes illustrates the K-string

$$(air\_transp\_means \equiv (glider \vee airplane \vee$$

$$helicopter \vee deltaplane \vee dirigible))$$

interpreted as a semantic representation of the definition "Air transport means is a glider, airplane, helicopter, deltaplane, or dirigible."

The restrictions on the cardinality of the sets can be expressed with the help of the function $Numb$ (the number of elements of a set) with the type

$$tp(Numb) = \{(\{[entity]\}, nat)\}\,,$$

where *nat* is the sort "natural number."

Example 6. The knowledge piece "All people have two parents" can have the following K-representation:

$$\forall x1\,(person)\,Implies((S1 \equiv all\;person*$$

$$(Parent, x1)), (Numb(S1) \equiv 2))\,.$$

The SK-languages allow for expressing such algebraic characteristics of the properties (or binary relations) as reflexivity, symmetricity, and transitivity.

Example 7. The property of transitivity of a binary relation on the set of all space objects (the concretizations are physical objects and geometric figures) can be expressed by the following K-string:

$$\forall r1\,(property1)\,(Transitive(r1) \equiv$$

$$\forall x1\,(space.ob)\,\forall x2\,(space.ob)\,\forall x3\,(space.ob)$$

$$Implies((Associated(r1, x1, x2) \wedge Associated(r1, x2, x3)),$$

$$Associated(r1, x1, x3)))\,.$$

The construction of this formula is based on the following assumptions:

- $r1, x1, x2, x3$ are the variables from the set $V(B)$, where B is the considered conceptual basis;
- $tp(r1) = tp(x1) = tp(x2) = tp(x3) = [entity]$;
- $tp(property1) = \uparrow \{(space.ob, space.ob)\}$;
- $Implies, Associated \in X(B)$, where $X(B)$ is the primary informational universe of the conceptual basis B;
- $tp(Implies) = \{(P, P)\}$, where $P = P(B)$ is the distinguished sort "a meaning of proposition";
- $tp(Associated) = \{([entity], [entity], [entity])\}$.

The principal advantage of the theory of K-representations in comparison with the language systems RDF, RDFS, and OWL is that it indicates a small collection of operations enabling us to build semantic representations of arbitrary NL-texts and, as a consequence, to express in a formal way arbitrarily complicated goals and plans of actions, to represent the content of arbitrary protocols of negotiations and to construct formal contracts concluded in the course of e-negotiations.

6.9 A Metaphor of a Kitchen Combine for the Designers of Semantic Technologies

It seems that a metaphor can help to better grasp the significance of the theory of SK-languages for the designers of semantic informational technologies, first of all, for the designers of Natural Language Processing Systems (NLPSs). It establishes a connection between the problems of house keeping and the problems associated with the development of semantic informational technologies.

When a woman having a full-time job enters the kitchen, she has a lot of things to do in a short time. The kitchen combines are constructed in order to make the work in the kitchen easier and to diminish the time needed for the work of the kind. For this, the kitchen combines can chop, slice, stir, grate, blend, squeeze, grind, and beat.

The designers of NLPSs have a lot of things to do in a very restricted time. That is why they need effective formal tools for this work. Like a kitchen combine for housekeeping, the theory of SK-languages can help the designers of semantic informational technologies to do many things. In particular, the theory of SK-languages is convenient for

- constructing formal definitions of concepts,
- representing knowledge associated with concepts,
- building knowledge modules (in particular, definitions of concepts) as the units having both the content (e.g., a definition of a concept) and the external characteristics (e.g., the authors, the date of publishing, the application fields),
- representing the goals of intelligent systems,
- building semantic representations of various algorithms given in a natural language form,
- representing the intermediate results of semantic-syntactic processing of NL-texts (in other words, building underspecified semantic representations of the texts),
- forming final semantic representations of NL- texts,
- representing the conceptual macro-structure of an NL discourse,
- representing the speech acts,
- building high-level conceptual descriptions of the figures occurring in the scientific papers, textbooks, technical patents, etc.

No other theory in the field of formal semantics of NL can be considered as a useful tool for all enumerated tasks. In particular, it applies to Montague Grammar and its extensions and to Discourse Representation Theory, Theory of Conceptual Graphs, and Episodic Logic.

In case of technical systems, a highly precious feature is the simplicity of construction. Very often, this feature contributes to the reliability of the system and the easiness of its exploitation.

The theory of SK-languages satisfies this criterion too, because it makes the following discovery in both non mathematical and mathematical linguistics: a system of such 10 operations on structured meanings (SMs) of NL-texts is found 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. Such operations will be called quasilinguistic conceptual operations. Hence the theory of K-representations suggests a complete collection of quasilinguistic conceptual operations (it is a hypothesis supported by many weighty arguments).

The useful properties of SK-languages stated above allow for the conclusion that the theory of SK-languages can be at least not less useful for the designers of NLPSs and for a number of other semantic informational technologies as a kitchen combine is of use for making easier the work in the kitchen

6.10 The Significance of the Theory of K-Representations for Semantic Web and Web Science

6.10.1 Theory of K-Representations as a Universal Resources and Agents Framework

It appears that RDF, RDFS, and OWL are only the first steps of the World Wide Web Consortium along the way of developing semantically structured (or conceptual) formalisms, and hence the next steps will be made in the future. The emergence of the term "Web X.0" supports this conclusion. That is why let's try to imagine what may be the result of the evolution of consequent Web conceptual formalisms, for instance, one decade later.

In order to formulate a reasonable assumption, let's consider such important applications of Web as Digital Libraries and Multi media Databases. If the resources are articles, books, pictures, or films, then, obviously, important metadata of such resources are semantic representations of summaries (for textual resources and films) and high-level conceptual descriptions of pictures. As for e-commerce, the conceptual (or semantic) representations of the summaries of business documents are important metadata of resources. That is why it seems that the Web conceptual formalisms will evolve during the nearest decade to a Broadly Applicable Conceptual Metagrammar.

Hence the following fundamental problem emerges: how to construct a Universal Conceptual Metagrammar (UCM) enabling us to build semantic representations (in other words, conceptual representations) of arbitrary sentences and discourses in NL? Having a UCM, we will be able, obviously, to build high-level conceptual descriptions of visual images too.

The first answer to this question was proposed in [70]: the hypothesis is put forward that the theory of restricted K-calculuses and K-languages (the RKCL-theory) may be interpreted as a possible variant of a UCM. With respect to this hypothesis and the fact that the RKCL-theory enables us to effectively approximate the expressive means of RDF and RDFS, we may suppose that the more general theory of K-representations can be used as an effective tool and as a reference-point for developing comparable, more and more powerful and flexible conceptual formalisms for the advanced Web.

The analysis (in particular, carried out above) shows that the theory of K-representations is a convenient tool for constructing formal representations of the contents of arbitrary messages sent by intelligent agents, for describing communicative acts and metadata about the resources, and for building high-level conceptual representations of visual images. That is why the theory of K-representations together with the recommendations concerning its application in the mentioned directions may be called a Universal Resources and Agents Framework (URAF); this term was introduced for the first time in [78].

6.10.2 The Need for the Incentives for Semantic Web

During last several 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 Berners-Lee et al. [18].

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 [10, 153].

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 et al. [191]. In this chapter, 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 this 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.

6.10.3 Toward a New Language Platform for Semantic Web

In [191], 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 e-science, first of all, 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 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, law, and politics. In particular, it is necessary to store the semantic content of the articles from newspapers, of TV-presentations, etc.

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 adequate language platform for Semantic Web is to allow for reflecting the results of applying ten partial operations on conceptual structures explicated by the mathematical model constructed in Chaps. 3 and 4 of this monograph.

The reason for this conclusion is the hypothesis set forth in the final part of the previous chapter: there are weighty grounds to believe that, combining ten partial operations determined, in essence, by the rules $P[1]-P[10]$, we are able to construct (and it is convenient to do) a semantic representation of arbitrarily complex NL-text pertaining to arbitrary field of professional activity.

6.10.4 A Possible Strategy of Developing Semantic Web of a New Generation

Let's consider the principal ideas of a new, theoretically possible strategy aimed at transforming the existing Web into a Semantic Web (Fig. 6.2).

The proposed strategy is based on (a) the mathematical model constructed in Chaps. 3 and 4 and describing a system of ten partial operations on conceptual structures and (b) the analysis of the expressive mechanisms of SK-languages carried out in this and previous chapters. The new strategy 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 an NL-text T" means in this chapter 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 an 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."

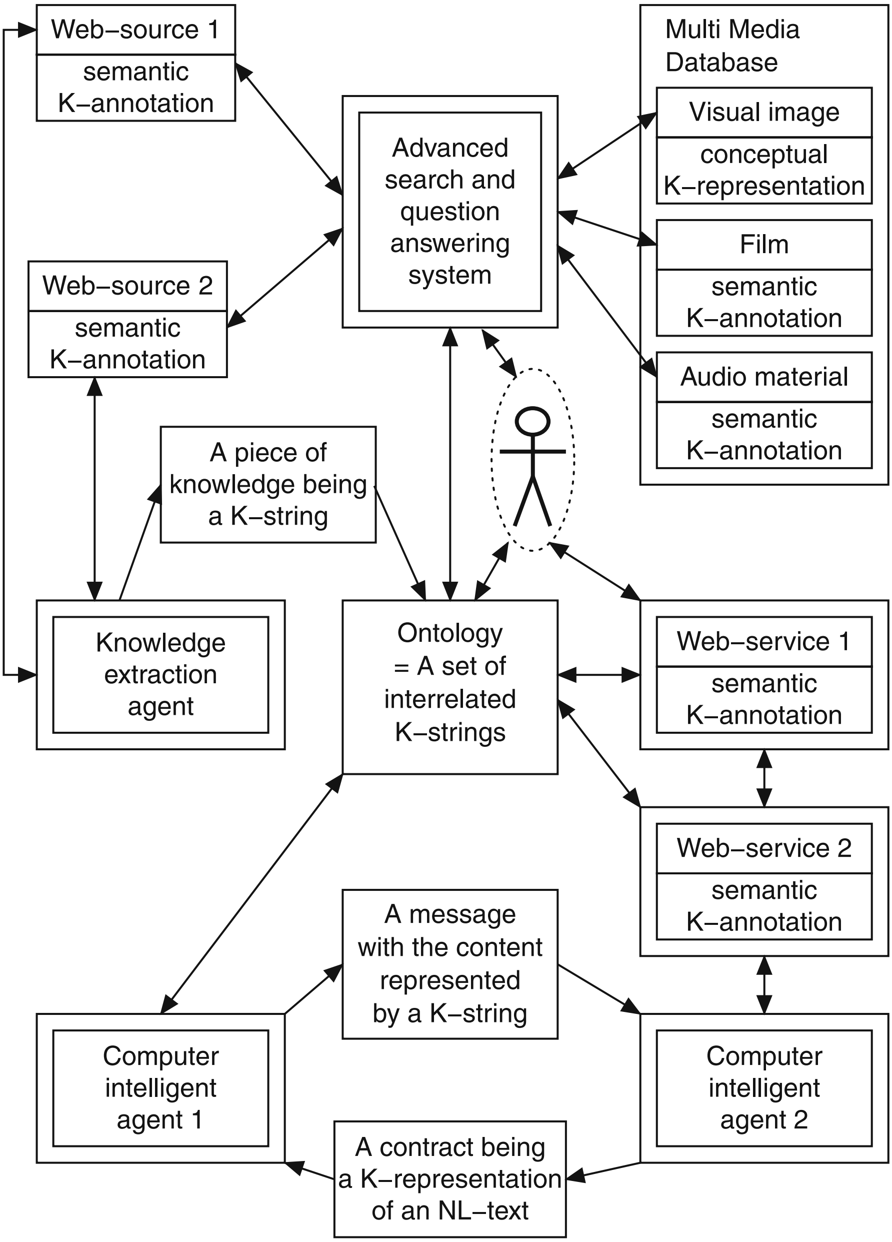

Fig. 6.2 The theoretically possible scheme of Semantic Web of a new generation

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 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 the requests are being designed.
7. The advanced Web-based search and question-answering systems are being created that are 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 an NL-text to be synthesized are being constructed. Besides, a reasonable direction of research seems to be the design of applied intelligent systems 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 an NL-expression but containing, maybe, a number of brackets, variables, and 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 strategy will depend on the results of its discussion by the international scientific community.

Problems

1 Describe the main ideas of building semantic annotations of informational sources with the help of SK-languages.
2 What new expressive mechanisms of SK-languages can be used for building high-level conceptual descriptions of visual images?
3 What new expressive mechanisms of SK-languages are useful for building compound denotations of notions?
4 Describe the proposed new interpretation of the principle of the language of least power for Semantic Web project.