Lexical and Semantic methods in design of the Problem-oriented Linguistic resources

Olga Kozhunova[1]
[1] Institute for Informatics Problems, The Russian Academy of Sciences,
Moscow, Russia

***Abstract* -** In the paper up-to-date methods and approaches to lexical and semantic modelling are analyzed, namely, in the works of such prominent specialists in the field as N. N. Leontjeva, E.V. Paducheva, J.D. Apresyan, R. Jackendoff, J. Sowa, and others. Some of the linguistic resources (integrated into information systems) with semantic- oriented features and other language-specific attributes are considered. Apart from the analyzing and reviewing of the existing approaches and resources, some suggestions and additions are given by the author, including the approaches of the artificial intelligence techniques and JSM-method formalisms.

Keywords: Lexical and Semantic modelling, Linguistic resources of the information systems, JSM-method.

1 Introduction

Last decades research preferences emphasize the significance of computer linguistics methods and models applied to various interdisciplinary tasks. Modelling of natural language mechanisms for problem-oriented systems touches up the overall quality of information systems and their performance, especially when it comes to the semantic aspect of text processing.

The above mentioned methods and models are used primarily for knowledge representation, formalization of their structures and entities, and adjustment to facilitate its automatic processing. Today the majority of information systems includes semantic component integrated into their linguistic software. Among them are well-known WordNet, EuroWordNet, Ontolingua, Russian dictionary Ruslan, etc.

In the paper lexical and semantic methods are viewed from the perspective of design and development of the problem-oriented linguistic resources of the domain-specific information systems. By the methods we mean a complex of theoretical terms designed to analyze and acquire basic semantic units from natural language texts at lexical and phraseological level. An effort to generalize the existing methods of this kind and to suggest a new enhancing approach is made.

Though the methods mentioned above are characterized as theoretical ones, nowadays they imply plenty of essential applications, namely, building of coordinated systems of domain-specific terms, analysis and formation of links between various lexemes, modelling of semantic relations at lexical and phraseological levels of any language, classification and typology of the domain-specific concepts, private reference tasks, coordination of various concepts understanding among experts of a field or domain, semantic component analysis, etc. This implies plenty of possible research directions and benefits when dealing with semantic and lexical methods integrated into linguistic resources.

2 Lexical and Semantic methods

Lexical and Semantic methods are applied in diverse fields and tasks. For instance, according to N.N. Leontjeva, almost every automatic text interpretation involves need in dictionaries that refer text units as symbol objects to their semantic equivalents [1, 2]. Among the dictionaries capable of such function is automatic semantic dictionary RUSLAN (Russian Dictionary for Analysis) designed and developed at The Moscow State University, Russia [1].

In RUSLAN Semantic information is encoded by specific descriptors indicating semantic category of a word considered. In total, 120 descriptors divided into two subsets are used. The first one is meant for Semantic Attributes (SA) description including such of the type ARTEFACT, SUBSTANCE, SITUATION, INFORMATION, etc. For example (1):

$$SA(\text{"electricity"}) = PHENOMENON \,\&\, ENERGY \qquad (1)$$

The second subset contains descriptions of binary semantic relations (SR) of the type PART(A,B), FORM(A,B), BELONG(A,B), etc. Descriptors may construct logic forms – conjuctions and disjunctions. As a rule, conjunctions combine generic and specific in the lexical meaning. They help for modelling of component semantic word structure [2, 3]. The Dictionary functions as a data base, new concepts/relations are entered there in specific patterns predetermined in the system [3].

Apart form the above mentioned, there is an approach to Lexical and Semantic modelling of the natural language processes which has been implemented in the system Lexicograph [4]. Lexicograph is a semantic dictionary also designed as a data base. Traditional dictionaries indicate a list of separate meanings for each word they contain. In order to reestablish the semantic unity of a word, Lexicograph formulates two interdependent tasks [5]:

- Present each given word meaning in the form allowing for getting a clear explanation of the language behavior of the word within its specific meaning;
- Visualize links between words, i.e. build a hierarchy of words meanings, or even a paradigm of meanings applicable to a certain category of words.

The system gives formats for presenting basic meanings types, and its user may edit the base entering new words and meanings. The current version is devoted to the category of words "Russian verbs" [5].
E.V. Paducheva is the principal ideologist of the system [4, 5]. She claims reestablishment of a word unity as one of the urgent tasks of semantics as a discipline. Up-to-date research techniques require a given word to be parsed into separate lexemes and postulate polysemy which speakers seldom perceive. In order to solve the polysemy problem Paducheva and her colleagues are guided by Kurilovich approach [6]. It concerns the concept of derivation and description of derivative meanings with the help of the original one.J.D. Apresyan, the leading expert in the field of lexical semantics and the leader of the Moscow Semantic School [7], states that today semantics is a specific component of the complete language description regarded as a formal tool which is modelling language behavior of humans. He also assumes that Lexical and Semantic modelling is merely modelling of the language skills rather than of knowing reality. These two assumptions are usually taken into account when implementing some applications of Computational linguistics.

Apresyan claimed one of the basic semantic rules regulating correct understanding of a text by a listener as one more postulate of lexical semantics. He formulated it, as follows: one chooses such understanding of a sentence given which is characterized by maximal repetitiveness of the semantic elements. In other words, this postulate (The Rule of Semantic Coordination) states that the meaning of a polysemic word is getting obvious from the context [8,9].

Theoretic research made by Apresyan was applied in the model "Meaning<->Text" (by I.Melchuk and A.Zholkovski) when he designed Explanatory and Combinatorial Dictionary. It became a dictionary of a new type due to its capability of demonstrating nontrivial combinations of lexemes. Он становится словарём нового типа, поскольку отражает, прежде всего, нетривиальную сочетаемость лексем. Words semantics in the dictionary is described as expanded formalized explanations involving only limited set of units. More complex elements (in semantic sense) are explained through common ones until it comes to "semantic primitives". This approach of complex semantic description had much in common with the principal concept of the Polish Semantic School headed by A. Wezbicka and A.Boguslavsky [10]. Apart from the dictionary mentioned Apresyan has designed plenty of other applications of his theories, e.g. New Explanatory Dictionary of Synonyms, Linguistic software of the ETAP system (machine translation), etc. [11, 12].

Among other prominent experts in the field of lexical semantics, cognitive linguistics, and generative linguistics are Ray Jackendoff and John Sowa. Notably, the latter being not linguist but a specialist in the field of Informatics and IT-technologies, made a significant contribution to the domain of Knowledge representation and Knowledge acquisition from natural language structures.

As regards Ray Jackendoff, he constructed the theory of semantic form expression and a comprehensive theory on the foundations of language [13-15].
One of the main goals of the theory suggested is characterization of the conceptual elements (e.g., a word and a sentence), thus constructing an explanatory semantic representation to conduct semantic analysis of the language units [13]. Jackendoff suggested to assume as the basic postulate that meaning in natural language is an information structure encoded by people at the mental level. This postulate enables us both to extend. boundaries of language representations and meaning acquisition in any language, and to apply theory of conceptual semantics in the fields studying problems of conceptualization of higher orders (logic, psychology, behavior science, etc.). In respect to linguistic tasks and formalization of its representations and processes this theory allowed of systemizing of language expressions in terms of semantic, or conceptual structures [13, 14]. According to Jackendoff, there is no need in interpretation of a conceptual structure by a structure of other level, since expressions of the conceptual structure are interpretations themselves. That is, the question is in the choice of the structures formalization and representation tools. For example, if an entity C from the real world has no representation in the mind of a person X (no mappings at the mental level), it involves that C either doesn't exist for X, or C is unavailable for X. It means that without mental representation of C person X wouldn't be able to address it as an expression, a language structure. Therefore one of the disambiguation methods is constructing mental relations with outer structures that already possess features necessary for formalizing of the entity C representation [15].

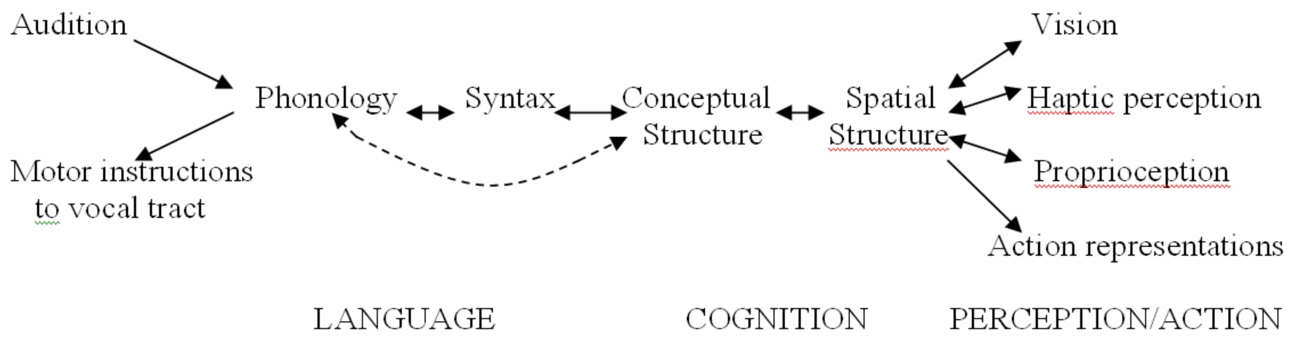

Figure 1. Architecture of the mind (according to R. Jackendoff)

Thus, Jackendoff considers that conceptual semantics, as linguistics in general, can't exist autonomously but closely relates to all the spheres of human perception which allows him/her of outworld comprehension and participation in various types of activities (Figure 1). This idea is very close to what Apresyan's and Wezbicka's ideas mentioned above. Moreover, such approach to Semantic modelling leads us to abstractions of higher level when studying lexical and semantic relations and concepts in natural language. To our mind, it happens by reason that context and semantic primitives are not only words, phrases and even large fragments of text encircling an expression/sentence under analysis, but also its situation background, genre, intonations, amount of speakers/listeners, various extralinguistic aspects of their perception, and many other factors. This approach when formalized and applied in the tasks of Lexical and Semantic modelling, Text analysis, Machine translation, and other applications will give us an opportunity to refine all the above mentioned operations with natural language structures.

Formalization of different language processes and units is downside of Lexical and Semantic modelling problems. Such thorough language analysis requires adequate formalisms to fix all relations between language concepts in a complete and accurate way. Recognized expert in the field of language knowledge representation and acquisition, semantic relations modelling, and also of Semantic Web and language ontologies design is John Sowa [16].

He developed new methods of logic and ontology application in the systems of language understanding and relevant logic reasoning [17, 18, 19]. Language of conceptual graphs description which he designed as well was accepted as one of the principal dialects of the standard ISO/IEC [20].

Conceptual graphs (CGs) are a system of based on the existential graphs of Charles Sanders Peirce and the semantic networks of artificial intelligence [21].With a direct mapping to language, conceptual graphs serve as an intermediate language for translating computer-oriented formalisms to and from natural languages. CGs have been implemented in a variety of projects for information retrieval, database design, expert systems, and natural language processing (www.jfsowa.com/cg/index.htm). Below are some examples of natural language structures formalization using John Sowa's CGs (Figure 2, 3):

Figure 2. Conceptual graphs built for the expression «Every cat is on a mat».

In KIF language (Knowledge Interchange Format), whose author is also John Sowa, the expression visualized at Figure 2 could be written as follows:

(forall ((?x Cat)) (exists ((?y Mat)) (On ?x ?y))) (2)

Such representation form is convenient for natural language structures representation.

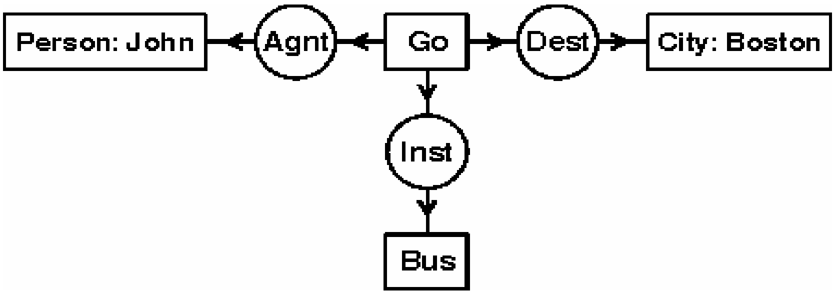

Figure 3. Conceptual graphs built for the expression «John goes to Boston by bus».

3 Additions to existing approaches of Lexical and Semantic modelling

In the previous paragraph some of principal methods and approaches to Lexical and Semantic modelling and its applications have been overviewed.

In general, while studying existing techniques developed to produce thorough analysis and parsing of natural language lexical structures, the following basic assumptions of Lexical semantics and its applications have been revealed:

- Usage of the concept "Semantic category" appeared to be rather fruitful when it came to describing lexical groups and their interrelations (particularly, in special problem-oriented dictionaries). Lexemes belonging to a certain category are marked by special descriptors [1];
- Lexeme descriptors, their attributes and relations proved their effectiveness when divided into two subsets. The first one is meant for describing semantic characteristics like ARTEFACT, SUBSTANCE, SITUATION, INFORMATION, etc. The second one – for describing binary semantic relations as PART(A,B), FORM(A,B), BELONG(A,B), etc. [1];
- Generic and specific in lexical meaning combined by logical operators allow of modelling the component semantic word structure [1, 2];
- Electronic semantic dictionaries are easy-to-use when they have data base architecture. This gives an opportunity to fix and formalize basic concepts, their attributes and relations[1, 7];
- Application of system of patterns minimizes mistakes of manual input of new concepts and associated structures. It also contributes to accurate Lexical and Semantic modelling of the natural language structures [2];
- Every word meaning considered should be represented in the form promoting a clear and complete explanation of the "language behavior" of a word in this particular meaning [4, 5];
- It is necessary to demonstrate in what way word meanings are associated with each other, i.e. to build a word meanings hierarchy or even a paradigm of its meanings general for a certain category of words [4, 5];
- The concept "Semantic role" is vital to words meanings descriptions, their links to language reality and reality in general. On one side, this furthers exposing those invariant components in semantics of a word that remain when it's deriving new forms. On the other, it allows of detecting the communicative shifts characteristic of various transformations not revealed before [4];
- The Rule of Semantic Coordination should be taken into account when design and development of new Lexical and Semantic methods are taking place, namely: one chooses such understanding of a sentence given which is characterized by maximal repetitiveness of the semantic elements, i.e. the meaning of a polysemic word is getting obvious from the context [7, 8];
- Meanings of language units correlate not immediately with the reality itself but with a language speaker notion about it [7, 10];
- Principle of the Integrative language description: vocabulary and grammar form a compact unity and are "adjusted" to each other [7];
- Meaning in any natural language is represented as an information structure encoded by people at the mental level [13];
- There is no need in interpretation of a conceptual structure by a structure of other level, since expressions of the conceptual structure are interpretations themselves (so, the question is in the choice of the structures formalization and representation tools) [14];
- Context and semantic primitives are not only words, phrases and even large fragments of text encircling an expression/sentence under analysis, but also its situation background, genre, intonations, amount of speakers/listeners, various extralinguistic aspects of their perception, and many other factors [10, 14].

Apparently, many methods and approaches from this list of basic assumptions of the Lexical and Semantic modelling have been further developed and applied at various stages of Semantics evolution. However, this doesn't mean that all existing techniques and approaches have been already developed, and problems of Lexical Semantics have been solved as well. Problems of lexical and structural polysemy, optimal representation of lexical and semantic structures in text parsing and analysis tasks, integration into information systems and domain-specific adjustment, and many other problems are still urgent. Hence, we would like to make some additions to the existing approaches.

As it was emphasized in the paper, it is a complicated task to develop Lexical and Semantic methods according to some abstract situation, or provide for all possible cases of language structures representation and usage. As a rule, everything depends on the particular applications, or some category of words and their specific features (say, English verbs), or semantic categories (e.g., semantic binary relations), etc. This leads us to a conclusion that in order to optimize approaches of Lexical Semantics at large, one needs to optimize some part of its range of problems.

Recently more and more efforts aim at maximal adjustment of linguistic resources to precise and complete modelling of natural language entities and processes. This implies qualitative change in design of problem-oriented information systems.

Therefore, as an approach adding new features to the existing ones, we introduce a formalism that has many fruitful applications in various domains. This is JSM-method of the automatic hypothesis generation originally elaborated by V.K. Finn [22, 23]. This method was named after John Stuart Mill, the English eminent philosopher, sociologist, economist, etc. He also was the author of some formalisms (Mill induction, rules of induction reasoning, inductive methods, and so on) [24]. These formalisms were applied in the JSM-method and were founded on the assumptions that it is necessary to generate hypotheses about objects properties and effects they produce using partly defined data bases. The result of the method work would be confirmation either disproof of the basic hypotheses. For instance, method of resemblance of Mill's induction is formulated as follows: resemblance of facts implies presence/absence of an effect and its repetitiveness [24].

The main idea of the method is related to formalization of cognitive procedures interaction, namely, of induction, analogy, and abduction. This interaction coordinates Mill's ideas about induction with Peirce's abduction [25], Popper's claim to falsify generated hypotheses [26] and plausible reasoning for knowledge discovery according to Polya [27].

JSM-based systems involve instruments that may be applied in various fields of science in which knowledge is weakly formalized, data is structured, and data base contains both positive and negative examples of some objects. But JSM-method proved its particular effectiveness in machine learning. It was corroborated in the task of semantic dictionary completion [28]. However, as semantic dictionary and associated intellectual system used for testing JSM-method approaches were small-scale prototypes, and data array was rather limited, one can't judge about the method's effectiveness for Computational linguistics in general.

So the given below propositions implying JSM-method are considered only from the perspective of the task of Lexical and Semantic modelling of the Information Monitoring system.

In general case, the Information Monitoring system is designed to monitor, analyze, and evaluate some domain/system/program, etc. using some data integrated into the system or external sources as a base for necessary calculations and assessment.

Since plenty of different specialists are involved in the monitoring process (e.g., subject domain experts, evaluators, IT-specialists, decision-makers, etc.), coordination of indicators[1)] meanings and their understanding by specialists becomes problematic. Moreover, indicators are characterized by novelty and weak associations between indicators names and concepts they are designating. Taking toponyms as analogous examples which names are rarely associated with geographic position they are attached to, one can hardly grasp an indicator meaning from its name. For instance, users of the Information Monitoring system consider "Performance indicator" and "Effectiveness indicator" equivalent. But to understand meanings of these indicators when working with the system, one needs to be aware of the algorithms of their calculation first [29].

Sometimes substantial information about some indicators might be found in regulations and standards, or scientific papers but it's unpredictable and rare.

All these aspects affect the final result of the Information monitoring system performance and decision-making. That is, one of the dramatic factors of its successful performance is unambiguous coordinated indicators names and their explicit definitions.

This involved the necessity in terms dictionary design which would be integrated into the system and take into account all mentioned specific features of the subject domain, thus facilitating monitoring and evaluation procedures. The dictionary was called Semantic dictionary, as the system required coordinated meanings explanations of its terms and representation of their interrelations [30].

Thus, methods and approaches of Lexical and Semantic modelling gained significance in the field of Monitoring and associated information systems.

As it was noted at the beginning of the section, to accomplish semantic analysis one feels necessity in applying the Rule of Semantic Coordination which in its brief form states that the meaning of a polysemic word (one may be interested in) is getting obvious from the context [7, 9].

In JSM-formalism a notion of context is one of the basic ones as well. For instance, if an inductive conclusion depends on some condition X expressed by a formula not equivalent of such Y and Z that correspond to cause and effect of the conclusion, such conclusion is called *context-dependant* [24]. Therefore, when analyzing some regulations or papers containing terms of monitoring and their definitions, one may apply such inductive conclusion which would take into account basic information about these terms and their context – to add new terms/concepts, their relations and other characteristics into the Semantic dictionary.

As to terms, definitions, examples, and relations already included into the dictionary, according to JSM-method they might be related to Base of facts. It is a base containing initial information for a logical conclusion. In the Base of facts this data would correspond to (+)-facts and (–)-facts that allow of detecting cause-effect relations in texts and construct hypotheses about their belonging/not belonging (i.e., (+)-hypothesis and (–)-hypothesis) to some natural language structure from Monitoring and indicators system described in the dictionary. It should be noted that descriptors of binary semantic relations might be used as initial facts (see the beginning of the section, basic assumptions of Lexical and Semantic modelling):

BE_INDICATOR (X, У={False, True}), PART_OF_DEFINITION (Z, Def) (3)

As (+)-hypotheses and (–)-hypotheses one may use components of new terms/concepts and their derivatives context (i.e., phrasal and supra-phrasal structures encircling the term sought-for). Consider an example of JSM-method approaches application and usage of the basic assumptions of Lexical and Semantic modelling from the perspective of the Semantic dictionary completion in the Information Monitoring system (Example 1).

[1)] Indicator – an index of quantitative estimate calculated on the base of information resources of a Monitoring system.

Example 1

(+)-facts

Performance indicator
Index of the RAS program effectiveness
Peer review of RAS Institutes activities

(–)-facts

Effectiveness of a medicine
Intelligence quotient
Indicator of oxidation-reduction processes

Fragments from the Information Monitoring system taken for analysis:

<…>On the base of the Decree 201 lies the law on system of result indicators of Federal programs of the RAS basic research. In order to define effectiveness of the program one should take into account the following indicators <…>

<…>Apparently, it is not enough to use results of logic-maths tests and assotiation tests as a general intelligence quotient of the senior age groups <…>

<…>accomplishing peer review of research departments of the RAS led to <…>

(+)-hypotheses

On the base of the decree lies the law
To define effectiveness
Accomplishing peer review led to

(–)-hypotheses

Apparently, it is not enough to use… as a quotient

When using such data in the context-dependant inductive conclusion of the JSM-method of the automatic hypotheses generation the Semantic dictionary is completed by terms/concepts «*Result indicators of Federal programs of the RAS basic research*» and «*Peer review of research departments of the RAS*».

Obviously, even such a small example demonstrates that there is necessity in conclusion results validation (new terms/concepts in the dictionary) and more detailed processing of the conclusion elements. But Example 1 makes it evident that formalism of the JSM-method can be adapted to the task of contextual retrieval and semantic analysis with further completion of the Semantic dictionary of the Information Monitoring system. But this is determined by initial data, limitations and requirements to the result of the linguistic resources performance.

4 References

[1] N.N. Leontjeva. Categorization of the units in the Russian all-semantic dictionary. (Леонтьева Н.Н. Категоризация единиц в русском общесемантическом словаре (РОСС) //Труды Международного семинара Диалог'98 по компьютерной лингвистике и ее приложениям. Т.2.- С.519-532). – Published in Russian.

[2] N.N. Leontjeva, S.J. Semenova. On reflecting polysemy in the automatic semantic dictionary. (Леонтьева Н.Н., Семенова С.Ю. Об отражении полисемии в прикладном семантическом словаре //Компьютерная лингвистика и интеллектуальные технологии. Труды международного семинара Диалог 2002. Протвино, 6-11 июня 2002 года. – М., "Наука", 2002. – Т.2. – С. 489-496.) – Published in Russian.

[3] N.N. Leontjeva, S.J. Semenova. The Semantic dictionary RUSLAN as an instrument of the computer comprehension. (Леонтьева Н.Н., Семенова С.Ю. Семантический словарь РУСЛАН как инструмент компьютерного понимания // Понимание в коммуникации. Материалы научно-практической конференции. М., 2003.) – Published in Russian.

[4] G.I.Kustova, J.V. Paducheva. Dictionary a lexical data base. (Кустова Г.И., Падучева Е.В. Словарь как лексическая база данных. - Вопросы языкознания, 1994, N3.) – Published in Russian.

[5] Lexicograph project site: http://lexicograf.ru/ideology/

[6] J. Kurilovich. Notions about word meaning (Курилович Е. Заметки о значении слова // в его кн.: Очерки по лингвистике, пер. с польск., англ., франц., нем. – М., 1962.) – Published in Russian.

[7] J.D. Apresyan. Selected works. (Ю.Д.Апресян. Избранные труды. Том 1. Лексическая семантика. Синонимические средства языка. М., 1995.) – Published in Russian.

[8] V.K.Gak. Semantic structure of a word as a component of the semantic structure of an utterance. (В.Г.Гак. Семантическая структура слова как компонент семантической структуры высказывания // Семантическая структура слова. Психолингвистические исследования. –М., 1971.) – Published in Russian.

[9] V.K.Gak. About semantic syntagmatics problem. (В.Г.Гак. К проблеме семантической синтагматики // Проблемы структурной лингвистики. – М ., 1972.) – Published in Russian.

[10] Wierzbicka, Anna. Semantic Primitives. – Frankfurt/M., Athenaum-Verl., 1972. – 235 p.

[11] J.D. Apresyan, I.M. Boguslavsky, L.L.Iomdin et al. Linguistic software for the system ETAP-2. (Ю. Д. Апресян, И. М. Богуславский, Л. Л. Иомдин и др. Лингвистическое обеспечение системы ЭТАП-2. – М.: Наука, 1989.) – Published in Russian.

[12] J.D. Apresyan, I.M. Boguslavsky, L.L.Iomdin et al. Linguistic processor for complex information systems. (Ю. Д. Апресян, И. М. Богуславский, Л. Л. Иомдин и др. Лингвистический процессор для сложных информационных систем. – М.: Наука, 1992.) .) – Published in Russian.

[13] Jackendoff, Ray. Semantic Interpretation in Generative Grammar. Cambridge, MA: MIT Press, 1972. – 400 pp.

[14] Jackendoff, Ray. Semantics and Cognition. Cambridge, MA: MIT Press, 1983. – 283 pp.

[15] Jackendoff, Ray. Foundations of Language: Brain, Meaning, Grammar, Evolution. Oxford: Oxford University Press, 2002. – 477 pp.

[16] Sowa, John F. Knowledge Representation: Logical, Philosophical, and Computational Foundations. Brooks/Cole Publishing Co., Pacific Grove, CA, 2000.

[17] Sowa, John F. Language Games: A Foundation for Semantics and Ontology. In Game Theory and Linguistic Meaning, edited by Ahti-Veikko Pietarinen, Elsevier, 2007. – pp. 17-37.

[18] Sowa, John F. The role of logic and ontology in language and reasoning. Chapter 11 of Theory and Applications of Ontology: Philosophical Perspectives, edited by R. Poli & J. Seibt, Berlin: Springer, 2010. – pp. 231-263.

[19] Sowa, John F., Majumdar, Arun and Stewart, John. Pursuing the goal of language understanding. P. Eklund & O. Haemmerlé, eds, Proceedings of the 16th ICCS, LNAI 5113, Berlin: Springer, 2008. – pp. 21-42.

[20] Sowa, John F. Conceptual Graphs. In F. van Harmelen, V. Lifschitz, and B. Porter, eds., Handbook of Knowledge Representation, Elsevier, 2008. – pp. 213-237.

[21] Max H. Fisch. Peirce Semiotic and Pragmatism. Indiana University Press, 1986.

[22] V.K. Finn. On the role of machine learning in the intellectual systems. (Финн В.К. О роли машинного обучения в интеллектуальных системах // Научно-техническая информация, Сер. 2, № 12. — М.: ВИНИТИ, 1999. — С.1-3.) .) – Published in Russian.

[23] Artificial intelligence. Reference book (Искусственный интеллект. Справочник в 3 кн. . Под ред. Попова Э.В //М.: Радио и связь, 1990.) .) – Published in Russian.

[24] J.S. Mill. System of the syllogistic and inductive logic. (Милль Д.С. Система логики силлогистической и индуктивной: Изложение принципов доказательства в связи с методами научного исследования. Пер. с англ./ Предисловие и приложения В.К. Финна. – М.: ЛЕНАНД, 2011. – 832 с.) – Published in Russian.

[25] Peirce C.S. Abduction and induction. Philosophical Writings of Peirce. Ed. Buchler. Dover Publications NY, pp. 150-156.

[26] Popper, K. R. Objective Knowledge, corrected edition, Oxford, Oxford University Press (1973).

[27] G. Polya. Mathematica and plausible reasoning. (Пойа Д. Математика и правдоподобные рассуждения. – М.: Книжный дом «Либроком»/URSS, 2010.) – Published in Russian.

[28] V.K.Finn, D.V.Vinogradov, O.S. Kozhunova. Intellectual system of semantic dictionaries completion. (Финн В.К., Виноградов Д.В, Кожунова О.С. Интеллектуальная система пополнения семантических словарей // Программные продукты и системы, № 2, 2006. – с.27-30.) – Published in Russian.

[29] I.M. Zatsman, O.S. Kozhunova. Semantic dictionary of the Information Monitoring system in the field of science: tasks and functions. (Зацман И.М., Кожунова О.С. Семантический словарь системы информационного мониторинга в сфере науки: задачи и функции. // Системы и средства информатики. Вып. 17.- М.: Наука, 2007.- С. 124-141.) .) – Published in Russian.

[30] Olga S. Kozhunova. Semantic Dictionary: heading toward ontology of the Information Monitoring System // Proceedings of WorldComp’10, Las Vegas, 2010. Pp. 934-941.

Despite their importance, lexical semantic relations are severely underrep-resented in current linguistic ontologies. As Morris and Hirst [2] point out, cur-rent linguistic ontologies only capture what they call classical relations â€" basi-cally, WordNet relations such as hyponymy, hypernymy, troponymy, meronymy, antonymy, and synonymy.Â In this research, we tackle the problem of automated learning of lexical se-mantic relations from text. We present an iterative algorithm in Sect. 2 that expands a small set of sample relation instances to a much larger set by mak-ing use of a dictionary of lexical analogies. In Sect.Â We refer to the semantic relation between a pair of words as the underlying relation of the word-pair. A lexical analogy A = W1 : W2 :: W3 : W4 is two. Some of the linguistic resources (integrated into information systems) with semantics oriented features and other language-specific attributes are considered. Besides analyzing and reviewing of the existing approaches and resources, some suggestions and additions are given by the author, including the approaches of the artificial intelligence techniques and JSM (John Stewart Mill) method formalisms. Keywords: lexical and semantic modeling; linguistic resources of the information systems; JSM-method. Full text: PDF file (399 kB) References: PDF file HTML file. Citation: O. S. Kozhunova, â€œApproa including linguistics. In other words, we deal with the inductive method of inquiry. The key role of the opening phase of linguistic analysis is that the statements of fact must be based on observation, not on unsupported authority, logical conclusions or personal preferences. â€¢ Classification is the second phase that comes after observation.Â Statistical inquiries have considerable importance because of their relevance to certain problems of the selection of vocabulary items for the purposes of language usage and language teaching. For instance, very few people know more than 10 % of the words in their mother tongue. During the day we usually pronounce about 48 000 words.