grammatically augmented ontology, language processing, machine translation, grammatical analysis

Maksym DAVYDOV[*], *Olga LOZYNSKA*[**]

SPOKEN AND SIGN LANGUAGE PROCESSING USING GRAMMATICALLY AUGMENTED ONTOLOGY

Abstract

The mathematical model of grammatically augmented ontology was introduced to address this issue. This model was used for grammatical analysis of Ukrainian sentences. Domain specific language named GAODL for description of grammatically augmented ontology was developed. The grammar of the language was defined by means of Xtext extension for Eclipse. The developed language was used as an auxiliary part of the information technology for bidirectional Ukrainian sign language translation.

1. INTRODUCTION

The problem of developing a machine translation system for sign language has been studied by scientists for a long time [1]. The solution to this problem can provide new communication opportunities for people with hearing impairments. The challenge of translation from Ukrainian sign language (UKL) to Ukrainian Spoken language (UKR) refers to tasks of machine translation.

One of the problems that arises when translating from one language to another is the problem of word sense disambiguation (WSD), i.e. the selection of one of possible meanings of a word listed in linguistic resources. The fuzziness of the problem suggests that there is no exact solution, but there are numerous heuristic methods developed to tackle it: methods of learning with and without a teacher [2], knowledge based methods [3], etc. Despite this, the problem of word sense disambiguation is not completely solved yet, because the solution of this problem requires structuring of human knowledge in different subject areas for each of the target languages.

[*] Information Systems and Networks Department, Lviv Polytechnic National University,
S. Bandery Str., 12, Lviv, 79013, UKRAINE, +38 (032) 258 26 38, maks.davydov@gmail.com
[**] Information Systems and Networks Department, Lviv Polytechnic National University,
S. Bandery Str., 12, Lviv, 79013, UKRAINE, +38 (032) 258 26 38, kanmirg@gmail.com

For a long time a source of knowledge for translation systems were glossaries and bilingual dictionaries [4]. Later ontological dictionaries replaced glossaries, because glossaries were not complete and lacked special structure for expressing semantic relationships between words and concepts [2]. Modern ontological dictionaries contain lexical and semantical information: antonyms, synonyms, relative words, nominalization, hypernymy and meronymy relations, etc. This information is used to decrease word ambiguity in different ways. However, the ontological vocabularies do not describe language constructions (predicates, templates) for expressing semantic relations between words. Thus, more detailed models and dictionaries are required for better language processing and translation.

The article describes a new model of grammatically augmented ontology (GAO), as well as examples of its usage for syntactic and semantic analysis of Ukrainian sentences. The work is a part of a larger project conducted by authors to tackle bidirectional Ukrainian sign language translation problem.

The article consists of problem statement, analysis of related work, description of GAO mathematical model and GAODL language, sentence parsing results and discussion.

2. PROBLEM STATEMENT

Extending ontologies with grammatical information is a task that requires analysis of possible language constructions or "templates" that are used to develop a discourse. The knowledge of these "templates" can be used to generate better translation and to perform semantic and syntactic parsing.

During the development of UKL translation system the following information was proved to be important and thus was included into GAO:

1) synonymous grammatical constructions and their relationship to concepts or synsets;
2) grammatical attributes of words: which one can be freely changed across templates, which one should be consistent, and which one can't be changed at all.

The inclusion of grammatically augmented ontology into the machine translation system requires the development of tools for description of ontologies. The main problem of the research was to extend ontological dictionaries with a grammatical information in a way that is expressive, extendable, consistent, and friendly to users and applications. In order to solve this problem a mathematical model of GAO was developed. The developed model was used to support storage format for ontology dictionaries.

An approach based on domain specific language was utilized to describe the storage format. GAO description language (GAODL) was developed for this purpose. Xtext extension for Eclipse was used to develop GAODL rules and to create user-friendly editing environment based on Eclipse.

3. RELATED WORK

Ontologies are widely used today for structuring knowledge from different subject areas. They are useful for formal specification of concepts and relations between them. Ontologies can be used to describe a particular area of knowledge (domain ontology) or to develop common sense relationships (upper ontologies). The main advantage of ontologies is their formal structure that facilitates computer processing [5].

Well-known example of domain ontologies are Gene Ontology for the annotation of biomedical data [6], the Unified Medical Language System (UMLS) [7], ontology in the field of goods and services such as UNSPSC [8], etc.

Upper ontologies are widely used for research in the field of computer linguistics. Such ontologies include WordNet for English [9], plWordNet for Polish language [10], CWN (Chinese Wordnet) for Chinese [11], WOLF (WordNet Libre du Français) for the French language [12], MultiWordNet project for Italian language [13], BalkaNet project for six European languages (Bulgarian, Czech, Greek, Romanian, Turkish and Serbian) [14], GermaNet for the German language [15], IndoWordNet for Indian language [16], the project RussNet [17] for the Russian language.

WordNet it is one of the largest electronic database, organized in a semantic network, which consists of various relations between words that includes synonymy, antonymy and generalization.

In Ukraine an ontology similar to WordNet is being developed by scientists Kulchitsky I. M. and others [18]. For now scientists have developed a small fragment of Ukrainian common ontology, which consists of 194 synsets, related to each other by links of hypo-/hypernymy, antonymy, and in addition by connections of meronymy/holonymy.

The scientists of Taras Shevchenko National University of Kyiv developed Ukrainian ontological lexical-semantic knowledge base UkrWordNet (UWN) [19]. The project lasts for several years and it was successful in creation and filling of UWN database. Special tools were used to fill the database. At the moment, the ontology contains about 80,000 concepts.

Ontologies are widely used for word sense disambiguation. In article [3] scientists A. Romaniuk and others have investigated the problem of word sense disambiguation and have analyzed the main methods of its solution. It was revealed that the WordNet network can be successfully used for automatic disambiguation of word meaning, however no percentage of correctly disambiguated words was given.

Ontologies are often used in conjunction to statistical models for word sense disambiguation that take into account the frequency of word mutual occurrence in a particular context [2]. This approach can provide high-quality translation only when large training corpuses are used. Besides that this approach has an obvious limitation because of the combinatorial explosion.

Thus, an alternative approach is required in order to use ontologies for translation between languages that have no large bilingual corpuses. It was shown that the use of grammatical rules itself is not sufficient for WSD [20]. A better approach can be developed when using ontologies in conjunction to language grammar and common expression templates to disambiguate word meaning.

Modern software tools for building ontologies such as DOE (Differential Ontology Editor) [21], Ontolingua [22], OntoEdit [23], WebOnto [24], Protege [25] have no means for linking concepts to possible grammatical constructions were they can be used. As to the authors' best knowledge, there are no papers that describe models of ontologies supplemented by grammatical relations and expression templates.

4. MATHEMATICAL MODEL OF GRAMMATICALLY AUGMENTED ONTOLOGY

There are several alternative mathematical models of classic ontologies [26]. The mathematical model of ontology from article [27] was chosen as the basis of the developed grammatically augmented ontology. It is defined as a tuple $O = \langle L, C, F, R_c \rangle$, where $L = \{w_i\}$ is a vocabulary of a subject area, $C = \{c_i\}$ is a set of the subject area concepts, $F \subset L \times C$ – a relation between appropriate terms and concepts, R_c is a set of relations on concepts (hyponymy, hyperonymy, meronymy, holonymy, etc).

The introduced grammatically augmented ontology was defined as a tuple:

$$O_G = \langle O, P, E, T, R_p \rangle \tag{1}$$

where: O is an ontology,

$P = \{p_i\}$ is a set of predicates,

$E = \{e_i\}$ is a set of expressions, where each expression $e_i = ((w_1, g_1), (w_2, g_2), \ldots, (w_n, g_n))$ is a tuple of grammatically augmented ontology terms (w_i, g_i),

$T = \{t_j\}$ is a set of parametrized expressions, where $t_j = (e_j, f_j, p_j)$ is a triple of expression e_j, argument positioning function $f_j : \{1,2,...,Len(e_j)\} \to \{0,1,...,N(p_j)\}$, and a related predicate p_j. $Len(e_j)$ denotes the length of tuple e_j, $N(p_j)$ is the number of places of predicate p_j,

R_p is a relation that matches predicates to verb concepts.

For some predicate p_j and some expression e_j argument positioning function $f_j(k)$ was defined to be 0 for the term in position k of the expression e_j that can't be changed without breaking the expression relation to predicate p_j. The value $f(k) > 0$ means that appropriate term in position k represents an argument of the predicate with ordinal number $f(k)$, and it can be replaced with another term from the set of hyponyms of term w_k. If the related predicate has n places and for each $i \in \{1,2,...,n\}$ exists $k \in \{1,2,...,Len(e_j)\}$ such that $f(k) = i$ then expression e_j completely defines predicate p_j. Otherwise, some arguments of the predicate are considered to be undefined in the sentence. They can be either completely unknown or can be devised from the context of speech or from a situation.

The definition of grammatically augmented ontology provided the possibility to express links between concepts, predicates and means of their expression in the form of language constructions.

For example, the predicate $GIVE(a,b,c)$, where a is someone who gives, b is someone who obtains and c is something that is passed, can be expressed using expressions $e_1 =$ "(somebody) (give) (somebody) (something)" or $e_2 =$ "(somebody) (give) (something) (to somebody)" (or in Ukrainian $e_3 =$ "(хто) (давати) (кому) (що)"). Both statements completely define the predicate and their argument position functions are $f_1(1) = 1$, $f_1(2) = 0$, $f_1(3) = 2$, $f_1(4) = 3$, and $f_2(1) = 1$, $f_2(2) = 0$, $f_2(3) = 3$, $f_2(4) = 2$ (Ukrainian expression e_3 has argument positioning function $f_3(1) = 1$, $f_3(2) = 0$, $f_3(3) = 2$, $f_3(4) = 3$).

In spoken languages the grammatical forms of subject, object, predicate, and complement comply to certain grammatical rules. These rules in the grammatically augmented ontology are defined by grammatical attributes of the expression terms.

These grammatical attributes were divided into 3 groups:

1) attributes that can't be modified (for example, preposition and casus of a complement);
2) attributes that can be freely modified (usually, number and gender of an object);
3) attributes that should be matched (like person and number of a subject or predicate).

In the example above expression e_1 has two terms “somebody”. The first of them has a grammatical attribute of subject, and the second has an attribute of object. It means that person and number of the first “somebody” should be matched to person and number of the predicate, and number of the second “somebody” can be independently modified, that is the default behaviour for subject and objects. In the Ukrainian expression e_3 the first term should be in nominativus and should be matched in number, gender and person to the predicate. The third term should be in dativus and its number and gender can be freely modified.

5. LANGUAGE FOR DESCRIPTION OF GRAMMATICALLY AUGMENTED ONTOLOGY

The creation and use of the GAO requires the development of special means for its representation. The approach based on domain specific language (DSL) was chosen due to its good extensibility, portability and verificability [28]. The term DSL indicates a language, that is used to solve specific problems using terminology that is as close to the subject area as possible.

The DSL named GAODL was created to facilitate uniform editing and processing of grammatically augmented ontologies. These ontologies could be created for specific subject areas and lately merged to obtain upper ontologies. The GAODL language contains means for definition of new grammatical attributes, synsets, relations on synsets, predicates and expressions.

The notion ***grammatical_attribute <name of category> [=description]*** was used to introduce new grammatical attribute. For example basic attributes for Ukrainian spoken and sign language were introduced using the following code:

grammatical_attribute noun_nom = noun in the nominative case
grammatical_attribute noun_gen = noun in the genitive case
grammatical_attribute noun_dat = noun in the dative case
grammatical_attribute noun_acc = noun in the accusative case
grammatical_attribute noun_loc = noun in the ablative case
grammatical_attribute noun_voc = noun in the vocative case
grammatical_attribute v_inf = verb (infinitive)
grammatical_attribute v_sign = verb (sign)
grammatical_attribute noun_sign = noun (sign)

Special attributes were added to distinguish expressions in Sign Language (***v_sign, noun_sign***). This approach was helpful to keep spoken and sign expressions in the same file.

Noun synsets were introduced using expression ***synset <name of synset> [=description]***. Every synset presents certain language concept. Synsets can be considered as a set of words or expressions that express the same concept. All nouns, verbs, adjectives and adverbs of Ukrainian spoken and Ukrainian sign language create sets of synonyms (synsets), each one is presenting one semantic concepts. For example:

synset knowledge_container
synset contains_knowledge
synset read_understand
synset read_recite
synset listen_understand
synset obtain_knowledge
synset entity
synset human_person
synset can_learn = any entity that can learn something
synset can_teach = any entity that can teach someone
synset object

Verb synsets with associated expressions were defined by expression ***synset <name of synset> [=description]*** (<newline> expression)* (<newline> description of domains)*.

synset teach = pass knowledge
teach 1.noun 2.v_inf
teach 1.noun 2.noun
explain 1.noun to 2.noun
0: can_teach
1: can_learn
2: knowledge_domain

Numbers were used to link expression terms to predicate variables. Number 0 was reserved for the subject. In the example provided above synset "teach" corresponds to some predicate $teach(a,b,c)$ were a is someone who teaches, b is an entity that obtains knowledge and c is some piece of knowledge.

The main relation that was used for sentence parsing was a hypernymy relation on synsets. This relation was defined using expression ***hypernym(<synset1>) = <synset2>***. For example:

hypernym(book) = contains_knowledge

There was no special keyword for hyponymy relation because it is the opposite relation to hypernymy. Besides hypernym/hyponym relation we found useful to use relation of association and meronymy(whole/part) for WSD tasks. These relations are defined by the following expressions:

parts(<synset1>)=<synset2> [,<synsetN>]*
optional_parts(<synset1>)=<synset2> [,<synsetN>]*
associations(<synset1>)=<synset2> [,<synsetN>]*
optional_associations(<synset1>)=<synset2> [,<synsetN>]*

These relations were used to increase the probability of word meanings that were associated or were in the meronymy relationship with neighbor words.

parts(book) = page
optional_parts(page) = page_number
associations(student) = university
optional_associations(teacher) = school

GAODL language was implemented using Xtext framework for Eclipse IDE. The Xtext framework provides a set of tools for development, editing and verification of domain specific languages. It includes the language code analyzer, code formatting tool, compiler, code editor etc. A parser and an editor for the GAODL language were automatically generated when its grammar was declared in Xtext. The screenshot of text editor for GAODL is depicted in Fig. 1.

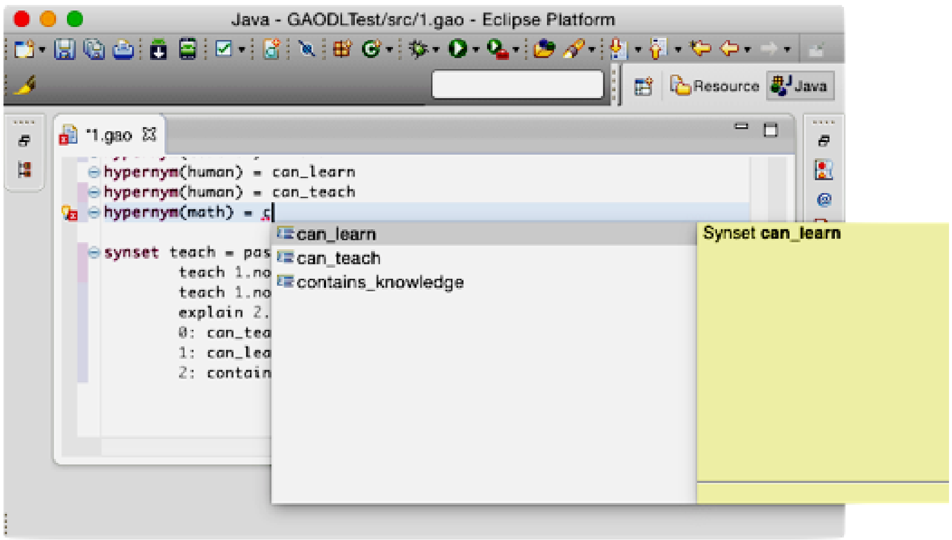

Fig. 1. The editor of GAODL files based on Eclipse IDE [source: own study]

6. THE RESULTS OF USING GAO FOR PARSING SENTENCES

The developed GAO was used for parsing sentences in Ukrainian Spoken Language and Ukrainian Sign Language. UKL sentences were represented as glosses.

Grammatically augmented ontology for "Education", "Nature", "Journey", "State", "Family", "Production", "Profession", "Army", "Theatre", "Culture", and "Hospital" subject areas were built. For this purpose, 1200 words were collected from these subject areas and the meaning of each word was verified using the Ukrainian glossary [29]. The meaning of UKL signs was clarified with teachers of Lviv Maria Pokrova Secondary Residential School for Deaf Children because there are no glossaries for UKL yet. GAO description was built using the collected words as synsets. Expressions in UKR and UKL were added for all verb synsets.

Affix probabilistic context free grammar (APCFG) parser UkrParser [30] was used for parsing sentences. All experiments were conducted for Ukrainian language and examples below are English equivalents of them.

The algorithm for parsing a sentence comprises the following steps:

1. Look up all possible meanings of every word from the sentence.
2. Add base forms for every word and detect its grammatical attributes.
3. Add hypernyms for every meaning of the words.
4. Add all expressions for every verb in the sentence.
5. Parse the sentence using UkrParser.

Consider parsing sentences “Professor teaches math to students” and “Professor teaches math to the car”. The parsing starts by adding all possible meanings of all words from the sentences, their base forms and all possible hypernyms (steps 1–3 of the algorithm). This process is outlined in table 1. GAO relation “hypernym” is not limited to be a simple tree structure. It can be used to define different groups of words that share some common property. In the example provided property “can_learn” is common for all individuals (“person_individual”). It is shown by relation person_individual→can_learn.

Tab. 1. Possible meaning of words from the example and their hypernyms [source: own study]

Word	Grammatical attributes	Meaning	Hypernyms
Professor	Noun, person3, singular	Someone who is a member of the faculty at a college or university	Professor → faculty_member → educator→ professional → adult → person_individual → organism_being → living_thing → whole_unit → physical_object → physical entity → entity
Teaches (teach)	Verb, present-simple, person3, singular	Impart skills or knowledge to	Teach → inform → intercommunicate → interact → act
Math	Noun, person3, uncountable	A science (or group of related sciences) dealing with the logic of quantity and shape and arrangement	Math → science → discipline_subject → knowledge_domain → content → noesis→ psychological_feature → abstraction
To	Preposition		–
	Part of infinitive		–
Student	Noun, person3, singular	A learner who is enrolled in an educational institution	Student → enrollee → person_individual → organism_being → living_thing → whole_unit → physical_object → physical_entity → entity. person_individual → can_learn
The	Definite article	Definite article	–
	Adverb	Used to modify an adjective or adverb in the comparative degree	–
Car	Noun, person3, singular	A motor vehicle with four wheels (automobile)	Car → motor_vehicle → self-propelled_vehicle → wheeled_vehicle → vehicle → transport → instrumentality → artifact → whole_uni t→ physical_object → physical_entity → entity
		A wheeled vehicle adapted to the rails of railroad (railcar)	Car → wheeled_vehicle → vehicle → transport → instrumentality → artifact→ whole_unit → physical_object → physical_entity → entity
		Elevator car where passengers ride up and down	Car → compartment → room → area → structure_construction → artifact → whole_unit → physical_object → physical_entity → entity

The next step is to add expressions for these words. Only verb “teach” contains an associated expressions, so it is added to the set of APCFG rules:

VP → teach <knowledge_domain>[NP] to <can_learn>[NP] (1.1)

VP → teach <can_learn>[NP] <skill>[VP, GERUND] (1.1)

where VP means verb phrase, NP means noun phrase and the numbers in braces mean multiplicative weight of the rules. In the conducted experiment all grammatical rules were weighted 1.0 and the weight of all expression rules was set to 1.1. This helped the parser to prefer expressions over the grammatical rules where it was possible.

The result of parsing the sample sentences is depicted in Fig. 2. An expression “teach” was used when the first sentence was parsed, thus the weight of the result is 1.1. In the second sentence expression “teach” could not be used because “the car” does not belong to the group of entities who “can_learn”. Thus, the second sentence was parsed using only grammatical rules.

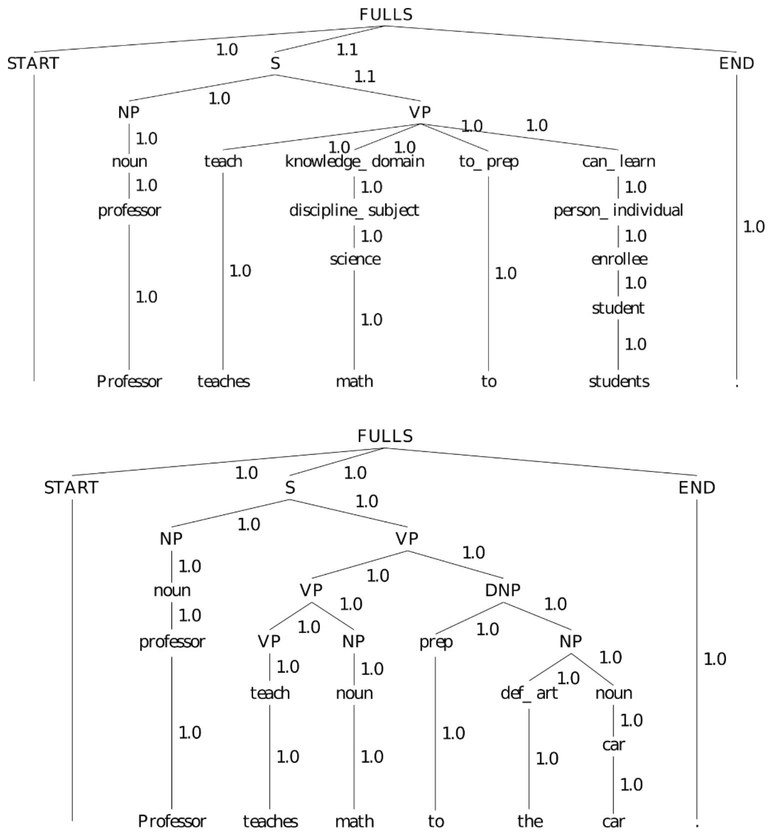

Fig. 2. The result of parsing sentences “Professor teaches math to students” and “Professor teaches math to the car”. FULLS stands for “full sentence”, S – a part with major clause, VP – verb phrase, NP – noun phrase, DNP – object or complement [source: own study]

The results of the experiment with parsing 200 test sentences in UKL and UKR language are given in table 2. The percentage of correctly parsed sentences was low when only the grammatical rules were used. This percentage is small especially for spoken language. It was due to the fact that Ukrainian spoken language grammar has flexible word order and word order in sign language is fixed in most expressions.

Tab. 2. Percentage of correctly parsed UKL and UKR sentences

Rule set	Ukrainian Sign Language	Ukrainian Spoken Language
Grammatical rules only	72%	65%
Grammatical rules + rules generated from GAO	91%	90%

7. CONCLUSIONS

The mathematical model and GAODL language for description of grammatically augmented ontology was developed. The model provides the possibility of integrating expressions into ontologies and supports the means for description of grammatical attributes.

The use of the developed grammatically augmented ontology for parsing sentences in Ukrainian Spoken and Ukrainian Sign Languages improved the performance of APCFG parser. The major increase in percentage of correctly parsed sentences was achieved for Ukrainian Sign Language.

GAODL language and the environment for editing grammatically augmented ontologies was developed using Xtext framework for Eclipse. This approach lets to use Eclipse environment with intelligent code completion for editing ontology files.

The language can be easily extended to incorporate more relations between concepts. Such relations can be relations like “role”, “instrument”, “locations” and others.

However, we faced challenges of verification ontology files from different sources, automatization of the process of building GAO ontologies from other known ontologies and large text corpuses. Besides that optimal weights for rules generated from GAO expressions and grammatical rules should be determined to achieve better performance of APCFG parser. These challenges would be a subject of further research.

REFERENCES

[1] BUNGEROTH J., NEY H.: *Statistical Sign Language Translation*. Proc. LREC 2004, Workshop proceedings: Representation and Processing of Sign Languages, 2004, pp. 105–108.

[2] NAVIGLI R.: *Word Sense Disambiguation: A Survey*. ACM Computing Surveys, ACM Press, no. 41(2), 2009, pp. 1–69.

[3] ROMANYUK A. B., SUNDUTOVA I. A., ROMANISHIN M. M. *Metodi virishennya leksichnoyi bagatoznachnosti. Vikoristannya WORDNET dlya virishennya problem bagatoznachnosti*, no. 711, NULP, pp. 147–157.

[4] KOBRITSOV B. P.: *Metodyi snyatiya semanticheskoy mnogoznachnosti*. Nauchno-tehnicheskaya informatsiya, no. 2, 2004,

[5] VORONOY A. S.: *Vikoristannya ontologIy dlya pidvischennya yakosti poshuku informatsii dlya popovnennya baz znan intelektualnih system*. Kompyuterni nauki I Inzheneriya, 2009, pp. 364–366.

[6] BROCHHAUSEN M.: *Establishing and Harmonising Ontologies in an Interdisciplinary*. Health Care and Clinical Research Environment, 2007. (http://ontology.buffalo.edu/medo/eHealth.pdf)

[7] KUMAR A., SMITH B.: *The Unified Medical Language System and the Gene Ontology: Some Critical Reflections*. In: KI 2003: Advances in Artificial Intelligence, 2003, pp. 135–148.

[8] UNSPSC. www.unspsc.org, 2014.

[9] AL-HALIMI R., BERWICK R.C., BURG J. F. M.: *WordNet: An Electronic Lexical Database*. MA: MIT Press, Cambridge, 1998, pp. 422.

[10] PlWordNet. http://plwordnet.pwr.wroc.pl/wordnet, 2006.

[11] CHU-REN H., SHU-KAI H., JIA-FEI H., YUN-ZHU C., I-LI S., YONG-XIANG C., SHENG-WEI H.: *Chinese WordNet: Design and Implementation of a Cross-Lingual Knowledge Processing Infrastructure*. Journal of Chinese Information Processing, no. 24(2), 2010, pp. 14–23.

[12] BENOÎT S., DARJA F.: *Building a free french wordnet from multilingual resources*. In Proc. of Ontolex. Marrakech, Maroc, 2008.

[13] PIANTA E., BENTIVOGLI L., GIRARDI C.: *MultiWordNet: Developing an aligned multilingual database.* In Proc. of the 1st International Conference on Global WordNet, India, 2002, pp. 21–25.

[14] TUFIS D., CRISTEA D., STAMOU S.: *Balkanet: Aims, methods, results and perspectives. A general overview.* Romanian J. Sci. Tech. Inform, no. 7(1–2), 2004, pp. 9–43.

[15] HENRICH V., HINRICHS E., VODOLAZOVA T.: *Semi-Automatic Extension of GermaNet with Sense Definitions from Wiktionary.* In: Proceedings of 5th Language & Technology Conference, Poznań, Poland, 2011, pp. 126–130.

[16] BHATTACHARYYA P.: *IndoWordNet*. Lexical Resources Engineering Conference, 2010.

[17] BALKOVA V., SUHONOGOV A., YABLONSKY S. A.: *Russian WordNet*. In: Proceedings of the Second International WordNet Conference, Brno, Czech Republic, 2004, pp. 31–38.

[18] KULCHITSKIY I.M., ROMANYUK A.B., HARIV H.B.: *Rozroblennya wordnet – podibnogo slovnika ukrainskoi movi.* InformatsIynI sistemi ta merezhi NULP, no. 783, 2011, pp. 306–319.

[19] ANISIMOV A.V., MARCHENKO O. O., NIKONENKO A.O.: *UWN: UnIversalna ontologIchna baza znan ukranskoi movi.* Problemi programuvannya, no. 2–3, 2012, pp. 348–355.

[20] LOZINSKA O.V., M.V. DAVIDOV, V.V. PASICHNIK.: *Mashinniy pereklad na osnovi pravil dlya perekladu na ukrainsku zhestovu movu.* Informatsiyni tehnologii ta kompyuterna inzheneriya, Vinnitsya, no. 29, 2014, pp. 11–17.

[21] DOE. http://opales.ina.fr/public/index.htmlFarquhar, 2006.

[22] FICKAS A., RICE J.: *The Ontolingua Server: a tool for collaborative ontology construction*. Proceedings of the 10yh KAW, 1995.

[23] SURE Y., ERDMANN M., STUDER R.: *OntoEdit: Collaborative Engineering of Ontologies*. Proceedings of the First ISWC, Sardinia, Italia, 2002, pp. 221–235.

[24] DOMINGUE J., MOTTA E., CORCHO GARCIA O.: *Knowledge modelling in WebOnto and OCML: A user guide*. 1999.

[25] NOY N. F., SINTEK M., DECKER S., CRUBEZ M., FERGERSON R. W., MUSEN M. A.: *Creating Semantic Web Contents with Protege-2000*. IEEE Intelligent Systems, no. 16, 2011, pp. 60–71.

[26] ANTONOV I. V.: *Model ontologii predmetnoy oblasti dlya sistem semanticheski-orientirovannogo dostupa*. Trudyi Pskovskogo politehnicheskogo instituta. Elektrotehnika. Mashinostroenie, no. 14.3, 2011, pp. 339–343.

[27] ZAHAROVA I. V., TIMCHENKO M. S.: *Sposobyi avtomaticheskogo postroeniya ontologii dlya zadach analiza tekstov*. Znaniya – Ontologii – Teorii: trudyi Vserosiysk. konf. ZONT-09, Novosibirs, no. 2, 2009, pp. 164–167.

[28] FREUDENTHAL M.: *Domain Specific Languages in a Customs Information System*. Software, IEEE, no. 99, pp. 1.

[29] BILODID I. K.: *Slovnik ukrayinskoi movi*. Naukova dumka, Kyiv, 1970–1980, vol. 1–11.

[30] UkrParser. https://github.com/mdavydov/UkrParser, 2013.