



25th DAAAM International Symposium on Intelligent Manufacturing and Automation, DAAAM
2014

Homonyms and Synonyms in NOK Method

Marina Rauker Koch^{a*}, Mile Pavlič^b, Martina Ašenbrener Katić^b

^a *Railway Technical School Moravice, Školska 2a, 51325 Moravice, Croatia*

^b *Department of Informatics, University of Rijeka, Radmile Matejčić 2, 51000 Rijeka, Croatia*

Abstract

Knowledge representation is one of the areas covered by artificial intelligence. Method Nodes of Knowledge (NOK) is one of the methods which allows a graphical representation and organization of knowledge in the form of a DNOK diagram (Diagram of Nodes Of Knowledge) as a network of nodes. With this diagram the meaning of sentences is preserved. The problem occurs when words that have synonyms or homonyms, and whose meaning depends on the context, appear in the sentence which is being modeled. Every term used in modeling sentences is written in a dictionary, including homonyms and synonyms. This paper presents a metamodel mode of writing homonyms and synonyms into the dictionary from which words are used for modeling. This demonstrates that the NOK method can be applied to modeling dictionaries of natural languages.

© 2015 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Peer-review under responsibility of DAAAM International Vienna

Keywords: Knowledge representation; Nodes Of Knowledge (NOK); homonyms and synonyms

1. Introduction

Knowledge representation [1] is an unavoidable field of artificial intelligence. For the conversion of knowledge into a form that can be encoded on a computer many methods, applicable to the various types of knowledge [2, 3, 4], have been developed.

This paper describes the problem of representation of words and their meanings through different connections in dictionaries of natural human languages. Knowledge representation in dictionaries will be demonstrated by using NOK method [5]. Particular problems are writing the meanings of words with multiple meanings (homonyms) and

* Corresponding author. Tel.: +385-51-877-120; fax: +385-51-877-523.

E-mail address: marina.rauker@skole.hr

writing the meanings of several different words that have the same meaning (synonyms). As a possible solution to these problems, metamodels of writing synonyms and homonyms from the dictionary using the NOK method are presented. The proposed models preserve all the meanings and connections of specific terms from the dictionary. Also, their application on the modeling of a specific example will be shown.

2. Review of Relevant Publication

In the field of artificial intelligence different approaches to knowledge representation formalisms are defined [2, 4]. Conceptual graphs [1, 6, 7, 8], semantic networks [9], frames [10] and predicate logic [11] are traditional formalisms that eventually produced numerous other formalisms and methods for knowledge representation. Ontologies have had a significant impact in the field of knowledge representation in the past twenty years [12, 13, 14]. They emphasize the semantic components of knowledge and language [15]. Furthermore, the field of computer analysis of natural language (NLP) is also involved in the formalization of knowledge expressed in language and text [16, 17, 18, 19, 20, 21, 22].

For a graphical representation of knowledge different methods have been developed. They use nodes and the links between them in their representations. Some of those methods are:

- Basic Conceptual Graphs (BG) - consisting of nodes (conceptual and relational) which are connected with links [24,25]
- Multi-layered extended semantic networks (MULTINET) - a very complex method in which nodes are classified using the method of conceptual ontology [25,26]
- Hierarchical Semantic Form (HSF) - consisting of group and link. The concept of the group is used in the labeling of a particular character, group of characters, words, semantic categories and complex patterns [23, 25, 27]
- Resource Description Framework (RDF) - a method that shows the relation between web resources (data, documents, images, etc.) by applying the appointed characteristic and its value [25, 28]
- Nodes of Knowledge (NOK) [5, 25, 29]

Comparison of these methods with the NOK method is analyzed in [25].

3. The Node of Knowledge Method (NOK)

The Nodes of Knowledge Method (NOK) belongs to a group of semantic networks in which knowledge is displayed as a graph [4]. Its aim is to show knowledge in textual form as a knowledge network [25].

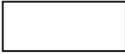
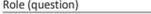
NOK is a method for modeling nodes of knowledge by which knowledge is displayed graphically in the form of a DNOK diagram (Diagram of Nodes Of Knowledge) [29, 30]. Sentences written in human language are converted into a graphical model from which it is possible to interpret the meaning because it is preserved in this model. It consists of nodes and links between them, and together they form a network of knowledge.

A node is the smallest element of knowledge that cannot be divided, and each new term or new meaning is a new node and is indivisible [25]. Accordingly, nodes can contain named persons, things, events, actions, ideas. Graphical representation of Nodes and concepts in the method NOK is given in Table I. The name of the node is an attribute of the node which gives it semantic meaning. The name may also consist of several words that have certain meaning.

Textual knowledge in a natural language is, by NOK method, organized into a network of knowledge (knowledge network, KN), which consists of three different levels of knowledge [29] (Fig. 1). These are: language knowledge (LK), which is at the highest level, then general knowledge (GK) at the next level and working knowledge (WK) which is a compound of the upper levels and represents concrete knowledge.

For each word it is necessary to create a model of knowledge that will be located in the dictionary, which is part of the LK.

Table 1. Symbols of NOK method.

Symbol	Concept
	NODE - part of knowledge in the reality or minds
	PROCESS NODE - action, event, occurrence
	LINK (one-way) and QUESTION (role identifier) – connecting nodes

3. Problem of Synonyms and Homonyms

Problem of dictionaries at the level of language knowledge (LK) in knowledge representation appears because of differences between dictionaries in terms of synonyms and homonyms, word semantics described through the meaning of other words and possible connections between words that are not part of the "analog" dictionaries. Dictionaries are important for the development of a system for understanding the meaning of sentences. On the basis of the semantics of words contained in dictionaries begins the process of interpreting the meaning of a sentence.

Knowledge structure in NOK can be shown schematically, as in Fig.1 [29]. Each word from a dictionary or in a sentence is a node by NOK method. The nodes are connected by other linking concepts and they are grouped into several levels of knowledge. Knowledge network, as can be seen in the presented model, consists of a series of levels of knowledge. At the lowest level of WK (Working knowledge) are sentences. At the next level is general knowledge (GK), such as days of the week, family relationships (husband - wife - child - father - son ...), etc.

It is possible to interconnect the nodes from any level where the node from higher level called a context node, is connected with lower levels' node called described node. The link is then called contextual link. Contextual link is a special link between the context node (node at a higher level of abstraction, general, superior, class, generic, super-term) and described node (a specialized, specific, node at a lower level of abstraction, the appearance, the associated), and for each described node it answers the following questions: what is a node, what kind of a node is it, which type or class is it, which group does it belong to, etc.

If two nodes from different levels are connected, the contextual link is displayed as an arrow line. The arrow points to the node at the lower level. The question is how to introduce synonyms and homonyms with the NOK method.

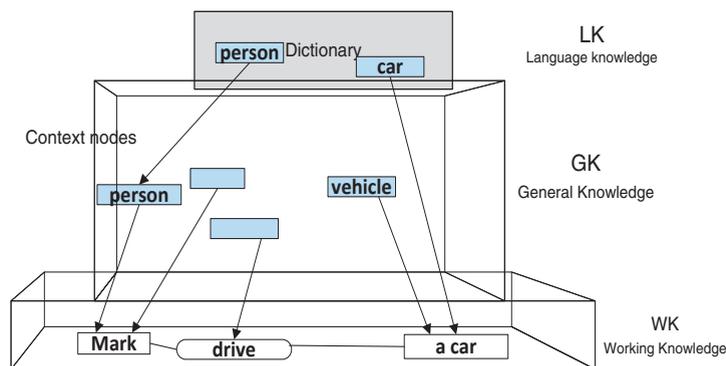


Fig. 1. Relationship LK, GK and WK in the network of knowledge.

3.1. Model of Embedding Synonyms

Linguistically, synonym [31, 32] is a word that differs in spelling from another word with which it shares:

1. The same meaning, the absolute synonym.
2. Similar meaning, relative synonym.

The model of knowledge should include knowledge from different dictionaries, so some general term "A" can be represented as A1 in the dictionary "r1" and as "A2" in the dictionary, "r2". In different dictionaries one term can have same or different synonyms. For example, the term A1 in one dictionary has synonyms B, C, D, and in the other the same term A2 has synonyms in this order: B, E, D, C.

Different synonyms of one term are connected via process node "Synonym", abbreviated as "S". Explanations of the term in various dictionaries are connected with the contextual link between them, which is named the same as the dictionary. This means that the general term A is listed in the dictionary r1 as A1. The link of process node synonym is marked with ordinal numbers of explanations in the dictionary.

Every synonym (e.g. C) is at the same time also a term in the dictionary (e.g. C1) that may have its own synonyms defined. In this way a network of connected terms is developed. Metamodel of synonyms is shown in Fig. 2.

3.2. Model of Embedding Homonyms

Linguistically, homonyms [31, 33] are words with the same spelling and pronunciation but different meanings.

The question is how to display the concept of homonyms with the NOK method; will it be a process node, used for synonyms, or an ordinary node. The idea is that the concept of homonyms in dictionaries is modeled by the concept of contextual link in the method NOK. This means that a term is interpreted with a super-term, that is, it belongs to the group of super-terms.

The assumption on which this concept for modeling knowledge lies is that some of the knowledge can be displayed and organized in nodes between which the relationship "abstract term - specific term" exists.

Metamodel of homonyms is presented with the method NOK in Fig. 3. A general term A can be shown as A' in a dictionary. The names of homonyms are not listed in dictionaries, but only the number of homonyms, such as 1., 2a., 2.b, etc. Suppose that term A' in the selected dictionary has two meanings, two homonyms, 1A' and 2A'. In DNOK, homonym is presented via concept of contextual link "H". Homonym can further be described by its synonyms (shown with processing node "S"). Each meaning of the term is explained with new terms (B, C, D) that have specific meanings and are synonymous with the term.

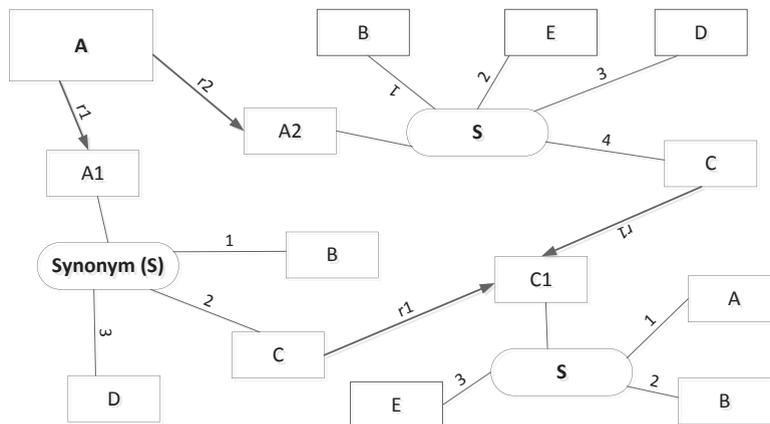


Fig. 2. Metamodel of synonyms displayed with NOK method.

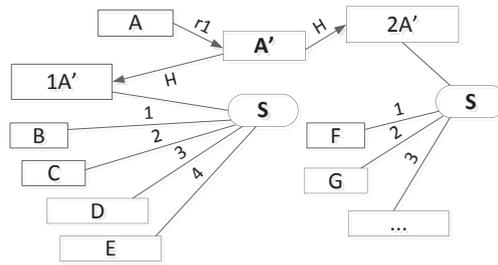


Fig 3. Metamodel of homonyms.

3.3. Application of Metamodel

We will show the suggested ideas of modeling homonyms and synonyms using a specific example. Take, for example, the word "zemlja" (earth, country, land, ground, dirt) from the Vladimir Anić Dictionary of the Croatian language [34] (Table 2).

Table 2. The word "zemlja" (Land) from the Vladimir Anić Dictionary of the Croatian language [34].

Croatian	English
<p>“zèmlja ž N mn zemlje, G zemljā</p> <p>1.a. (Zemlja) planet na kojem živimo, treći unutarnji planet Sunčeva sustava s jednim prirodnim satelitom (Mjesec) b. mjesto, prostor života i ljudske djelatnosti; svijet</p> <p>2. površina Zemlje, kopno, suho, opr. voda</p> <p>3. površina tla, gornji sloj Zemljine kore [nad zemljom i pod zemljom]</p> <p>4.a. zemljište, tlo kao izvor dobara i hrane b. (i u mn) parcela, prostor koji se obrađuje, iskorištava; zemljište kao imovina, vlasništvo [privatna zemlja; zadružna zemlja; državna zemlja]</p> <p>5. tip tla koje se obrađuje, na kojem se gradi itd. [pjeskovita zemlja; raskvašena zemlja]</p> <p>6. pov. državno-upravna jedinica u smislu državne, administrativne i političke podjele Habsburške Monarhije i Austro-Ugarske Monarhije”</p>	<p>“land n. pl lands</p> <p>1.a. (Earth) planet where we live, the third inner planet of the Solar System with one natural satellite (Moon) b. place, habitation of humans and place of human activity, world</p> <p>2. the solid part of the earth’s surface, ground, mainland (opp. water, air, sea)</p> <p>3. the surface of the ground, the top layer of the Earth’s crust [above and under the earth’s surface]</p> <p>4.a. a terrain, field, soil as a source of goods and food</p> <p>b. (also in pl.) plot, area that can be cultivated, exploited; land as assets, property, estates [private land, collective land, state land]</p> <p>5. type of soil which is cultivated, built on, etc. [sandy land, sodden land]</p> <p>6. his. a country, a state, a nation with its own government, occupying a particular territory</p>

Metamodels of synonyms and homonyms were applied to create models of the word "zemlja" and their synonyms and homonyms (Fig. 4) The word "zemlja" as defined above, has six different meanings, of which the first and the fourth have two submeanings that are similar to each other, and are its homonyms. In the model, the word "zemlja" is linked by contextual relationships marked with the letter H and a number with associated explanation from the dictionary (H1a to H6) which is then connected via a process node with its synonyms (according to explanation H1a "zemlja" is associated via processing node S with its synonym “planet”).

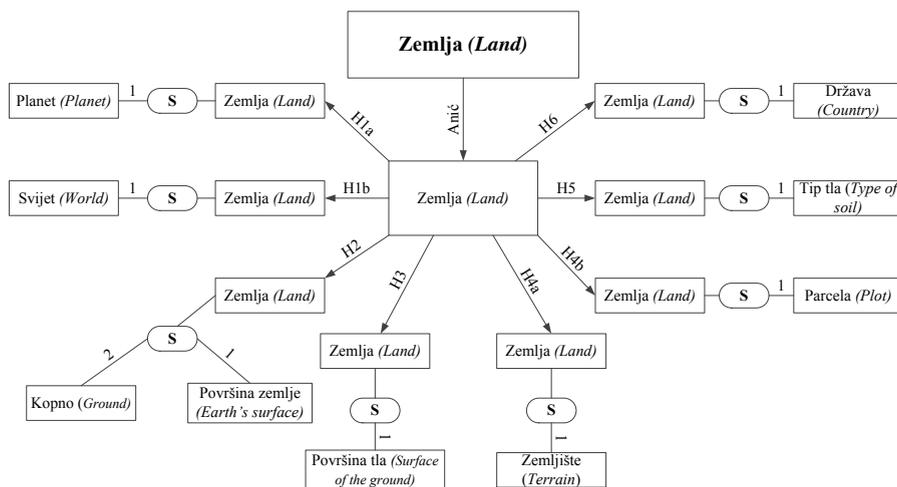


Fig. 4. Model of the word “zemlja” (Land) from the Vladimir Anić Dictionary of the Croatian language [34].

Conclusion

The NOK method is applicable to modelling a text written in a natural human language. The method NOK, besides keeping terms, also preserves connections between them, i.e., it preserves the syntax and the meaning of sentences. Writing concepts must be unambiguous, and the problem occurs if a sentence which is being modeled contains homonyms, words that have multiple meanings. Synonyms, words with same/similar meaning, should also be considered. This problem can be solved by creating a network of knowledge of words, or dictionaries, which would include all the terms and their relationships, either as homonyms or synonyms.

In this paper the metamodels of homonyms and synonyms formed with NOK method are shown. These models confirm the hypothesis that the NOK method can be applied to modeling dictionaries of natural languages. Synonyms are modeled by concept of process node that connects all synonyms for the chosen term. A synonym, also a term in the selected dictionary, has a synonymous process node that displays all its synonyms as well. It should be investigated whether one synonymous process node gathering all synonyms is sufficient, or should there be as many nodes as there are synonyms. For each synonymous word, it is possible that all its synonyms are not listed, or the relationship between them is not regarded as bijective (if a term has a synonym then itself is also a synonym of its synonyms). The situation in specific dictionaries should also be investigated; if there is bijection, if it is mandatory, whether there are errors in dictionaries.

Homonyms are modeled by the context link that connects the term and its explanations. The term can have several homonyms and all of them should find homonymous connections. Homonymous link, unlike synonymous link, cannot be an inversion. Dictionary is actually a list of meanings of homonyms of individual words (terms) and the list of synonyms of each homonyms.

This paper describes the application of NOK method on a specific example. Using the proposed method, we can make an interconnected network of concepts from one dictionary, where the interpretation of some terms is given through other terms via synonymous or homonymous connection. Every single dictionary represents a separate DNOK network. It is possible to connect different dictionaries by bringing together different DNOK diagrams. The connection is simply achieved through general dictionary that contains all the words of all dictionaries without definitions. The terms from the general dictionary, and concepts from a variety of dictionaries represent the highest level of language knowledge, LK.

In future research it is necessary to explore other problems with the recording of language knowledge, such as compound words and derivative words, which are integral parts of dictionaries.

Acknowledgements

The research has been conducted under the project "Extending the information system development methodology with artificial intelligence methods" (reference number 13.13.1.2.01.) supported by University of Rijeka (Croatia).

References

- [1] J. F. Sowa, Knowledge representation: logical, philosophical and computational foundations. Pacific Grove: Brooks/Cole Publishing Co., 2000.
- [2] F. Harmelen, V. Lifschitz, B. Porter, Handbook of Knowledge Representation, Oxford: Elsevier, 2008.
- [3] S. Russell, P. Norvig, Artificial Intelligence: A Modern Approach, Third edition, Prentice Hall, 2010.
- [4] R. J. Brachman, H. Levesque, Knowledge representation and reasoning, The Morgan Kaufmann Series in Artificial Intelligence, Morgan Kaufmann, 2004.
- [5] M. Pavlič, A. Jakupović, A. Meštrović: Nodes of knowledge method for knowledge representation, Informatologia. 46, 3, 206-214, 2013.
- [6] J. F. Sowa, Conceptual graphs for a data base interface. IBM Journal of Research and Development. 20(4). 336 - 357., 1976.
- [7] J. F. Sowa, Semantics of conceptual graphs. Proceedings of the 17th annual meeting on Association for Computational Linguistics. Stroudsburg: Association for Computational Linguistics. 39 - 44. 1979.
- [8] A. Jakupovic, M. Pavlic, Z. Dovedan Han, Formalisation method for the text expressed knowledge, Expert systems with applications, pp. 5308-5322, 41, 2014.
- [9] R. Quillian, Semantic memory. Semantic Information Processing. MIT Press. 227 - 270., 1968.
- [10] M. Minsky, Framework for Representing Knowledge. Technical Report. Cambridge. 1974.
- [11] J. McCarthy, Programs with common sense. Teddington Conference on the Mechanization of Thought Processes. Teddington: National Physical Laboratory. 75 - 91. 1958.
- [12] A. Gomez-Perez, M. Fernandez-Lopez, O. Corcho, Ontological Engineering. London: Springer., 2004.
- [13] S. Grimm, P. Hitzler, A. Abecker, Knowledge Representation and Ontologies, Logic, Ontologies and Semantic Web Languages, 2007.
- [14] T. R. Gruber, A translation approach to portable ontology specifications. Knowledge Acquisition. 5(2). 199 - 220., 1993.
- [15] A. Garcia-Crespo, B. Ruiz-Mezcua, J. L. Lopez-Cuadrado, I. Gonzalez-Carrasco, Semantic model for knowledge representation in e-business. Knowledge-Based Systems. 24 (2). 282 – 296, 2011.
- [16] J. Allen, Natural Language, Knowledge Representation and Logical Form. In Bates, M., & Weischedel, R. (1993). Challenges in Natural Language Processing. Cambridge University Press. 1993.
- [17] J. Allen, M. Swift, W. Beaumont, Deep Semantic Analysis of Text. Proceedings of the 2008 Conference on Semantics in Text Processing. Stroudsburg: Association for Computational Linguistics. 343 - 354., 2008.
- [18] N. Chomsky, Studies on Semantics in Generative Grammar. Janua Linguarum Series Minor. Mouton De Gruyter, 1980.
- [19] D. Jurafsky, J. H. Martin, Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. New York: Prentice Hall PTR. 2000.
- [20] T. McQueen, A. A. Hopgood, T. J. Allen, J. A. Tepper, Extracting finite structure from infinite language. Knowledge-Based Systems. 18(4-5). 135-141, 2005.
- [21] S. C. Shapiro, J. W. Rapaport, Knowledge representation for natural language processing, Natural Language Planning Workshop. New York: Northeast Artificial Intelligence Consortium. 56 - 77. 1987.
- [22] S. Wermter, E. Riloff, G. Scheler. Connectionist, statistical and symbolic approaches to learning for natural language processing. New York: Springer-Verlag., 1996.
- [23] M. Ašenbrener Katić, Predstavljanje znanja: pregled područja (Knowledge representation: overview of the field), Department of informatics, University of Rijeka, kvalifikacijski rad, 2013.
- [24] M. L. Mugnier, M. Chein. "Graph-based Knowledge Representation." Advanced Information and Knowledge Processing. Springer, London 2009.
- [25] A. Jakupović, M. Pavlič, A. Meštrović, V. Jovanović, Comparison of the Nodes of Knowledge method with other graphical methods for knowledge representation u Proceedings of the 36th international convention /CIS/, MIPRO 2013, Rijeka, 2013.
- [26] H. Helbig, Knowledge Representation and the Semantics of Natural Language, Springer, 2006.
- [27] M. Stanojević, S. Vraneš, Knowledge representation with SOUL, Expert Systems with Applications, 33, pp. 122-134, 2007.
- [28] K. Tolle, Introduction to RDF 2000., <http://www.dbis.informatik.uni-frankfurt.de/~tolle/RDF/DBISResources/RDFIntro.html>, 8.3.2014.
- [29] M. Pavlič, Development of a method for knowledge modeling 2013., Department of informatics, University of Rijeka, Rijeka, 2013. (unpublished)
- [30] M. Pavlič, Oblikovanje baza podataka (Database design), Department of informatics, University of Rijeka, Rijeka, 2011.
- [31] <http://hjp.novi-liber.hr/index.php?show=main> (10. 6. 2014.).
- [32] <http://www.oxforddictionaries.com/definition/english/synonym> (10.6.2014.)
- [33] <http://www.oxforddictionaries.com/definition/english/homonym> (10.6.2014.)
- [34] V. Anić, Rječnik hrvatskoga jezika (Dictionary of the Croatian language). Novi Liber, Zagreb, 1991.