SEMANTIC CATEGORIZATION OF DATA BY DETERMINING THE SIMILARITIES OF ASSOCIATIONS OF THE SEMANTIC NETWORK

SEMANTIČKA KATEGORIZACIJA PODATAKA UTVRĐIVANJEM SLIČNOSTI ASOCIJACIJA SEMANTIČKE MREŽE

Miloš Stojković[1], Miodrag Manić[1], Milan Trifunović[1], Dragan Mišić[1]

[1] University of Niš, Faculty of Mechanical Engineering in Niš

Abstract

In the paper is presented concept of semantic network analysis which performs semantic categorization of data in the network. The concept is based on determining the similarity between semantic network associations which is done on two levels. On the first level, the determination of associations similarity is based on similarity of attributes values of associations. On the second level, the determination of association's similarity is performed according to the similarity of graph topology, or association's subgraphs. The concept allows highly efficient semantic categorization of new concepts, which does not depend on pre-planned inputs and predefined rules of deduction. Also, the concept allows different semantic interpretations of the same concept in different semantic contexts.

1. INTRODUCTION

Semantic categorization (classification of data according to meaning) and interpretation of data (SCD, SID), represent one of the biggest challenges faced by modern information technologies. In fact, this problem is closely related to the ability of computer applications to attach certain meaning to data which is being processed. The motive for the solution of this problem lies in the ever increasing need to enable software applications to provide meaningful answers when it is not possible to predict the input, and consequently the code by which a meaningful response is programmed. This need becomes more intense with increasing amounts of information available. Computer networking in the Internet has enabled instant access to indefinite amounts of data and information. However, it is often difficult rather than easy to find quality information - one for which the user searches. Quality of information access does not equally follow the increase of available information [1], [2]. Research in cognitive psychology suggest that the process of searching information contents should include the search of meaning (or according to the meaning) of what is required in order to conduct search significantly more precise and even faster [3], [4]. Following these recommendations, the current research trend in the field of information technologies includes the two main subdirections which complement each other: a) research and the creation of semantic data models and b) research and modeling of cognitive processes. The studies of semantic data models aim to examine the way we form information about the outside world and how information is stored in our memory in order to create data models that will be able to "accept" and "incorporate" the meaning or semantics of data. Of course, the ultimate goal is not the integration of meaning alone, but its automated and intelligent interpretation and use [5]. For that reason, studies of cognitive processes are trying to determine how this information is used to solve different types of more or less complex problems [4], and to study the mechanisms of cognition and reasoning in humans, in order to model these processes and perform them automatically by artificial-intelligence systems. These directions in science are recognized as an research area of supreme importance for the near future of the European Union [6].

2. CURRENT SOLUTIONS

In solving the problem of semantic categorization and interpretation, so-called functional modeling – FM is the most widely used approach. Basically, it uses Generic Semantic Model (GSM) whose most prominent derivatives are entity-relationship model (ER), a functional data model (FDM) and semantic data model (SDM) [7]). Essentially, this is a way of presenting the knowledge coming from the concept of semantic networks [8], and the first graphical representations of the so-called semantic memory structure. The sudden development of web technologies in the last two decades also has accelerated the evolution of functional modeling and its application [9], [10], [11]. The main motive for influence of functional modeling on web technologies and/or vice versa, is the absolute priority of improving the quality of information retrieval on the Internet [5]. In order to improve the quality of search it was necessary to use not only phenomenal but also the semantic (meaningful) combinations of keywords (some sort of semantic context). In fact, in order to quickly find the right information in a huge amount of data on the Internet it's imperative to have "information on this data" - metadata, which will allow software applications for browsing the internet to find the requested information. Metadata, therefore, should be used to describe the meaning of the requested content on the Internet. Methods and models of describing semantic features are taken from the functional modeling [9], while syntax for metadata structuring is based on extensible markup language (XML) and its derivatives (RDF / RDFS, DAML + OIL and OWL) [5]. Ontologies are often referred to as vocabularies of so-called web objects made in OWL notation [5]. Like other vocabularies, their purpose is to present the meaning of terms contained in them, and in this case, the meaning of these web objects. Web objects are actually concepts used for representation or conceptualization of data and information content on the web. Vocabularies are organized according to the semantic networks model, where the nodes are concepts or web objects and functional relationships between web objects are represented by semantic network relations. It is precisely this organization that supports and enables the semantic categorization and interpretation of web objects.

Limitations

However, ontologies also have significant limitations on in terms of their ability to enable SCD and SID. These limitations originate from imperfections in the structure of the vocabulary, especially concerning semantic relations, and the lack of appropriate methods of cognitive data processing and semantic interpretation of data (web-objects, concepts). With ontologies SID is conducted on two levels.

On the first level, functional or semantic relations in the segment of semantic network around the currently observed (new) concept are analyzed and new functional relationships from or to a new concept are created according to pre-prepared set of rules. Which functional relation will be created depends upon analyzed functional relations graph. In essence, the rule set contains an encoded first-order logic which is used by ontologies inference engine to create the adequate functional relation between currently considered concept and other concepts in the semantic network for pre-planned cases of functional relations subgraphs (Fig. 1). Creation of new functional relations between data that is currently considered and data already described in the vocabulary results in semantic categorization of a new concept.

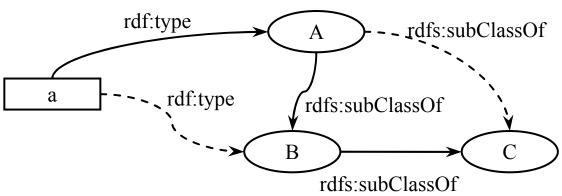

Figure 1. Propagation type inferencing schema

Although this set of rules brings partial analyticity of semantic interpretation data, autonomy and flexibility of the semantic interpretation of data is completely abolished. Each new case of functional relations subgraphs, which is not provided by the set of rules, can not be used for semantic interpretation. Another important aspect which further compromise the intelligence of ontologies "inference engine" is the fact that embedded set of rules covers only the simplest relations subgraphs. "Inference engine" is not designed to analyze large subgraphs. Thereby ontologies "inference engine" is not able to perceive the wider context of functional relations which further reduces the analyticity SID. Also, it disables ontologies "inference engine" to connect SID with semantic contexts, or to analyze the analogies between semantically distant subgraphs.

At another, deeper level, by using of SWRL (Semantic Web Rule Language) syntax an additional domain knowledge in the form of production rules can be embedded in the sole content of vocabulary. Of course, this kind of knowledge can also be used for semantic categorization of data, but only for the cases with provided input (an example of this kind of rule is presented in expressions 1 and 2):

IF

$$P_j^i \in comp(A_i), j = 1,2,\dots,m \wedge FF_{jk} \in comp(P_j^i) \quad (1)$$

where P_j^i is a part, A_i is an assembly, FF_{jk} is a form feature (surface, edge, etc.)

THEN

$$B_{jk}^{(i)}: FF_{jk}: \rightarrow P_j^i, k = 1,2,\dots,n, \quad (2)$$

where „→“ designates property: *belongsTo*

In addition to inference engine limitations regarding cognitive data processing, there is a limitation regarding vocabulary structure. In fact, the method of structuring or incorporating the meaning of data can be considered very limited. First of all, functional relations are treated as concept features, i.e. structurally relations are encapsulated into the concept definitions in the form of object-oriented programming - for appropriate so-called *domain* and *range*. This approach to building relations indicate that the meaning of the concept and knowledge of it is seen as valuable content of the box which can be reached only by opening the box (through the concept). The flexibility and analyticity of semantic categorization of data is significantly limited by this ontology feature because it is not easy to take advantage of the functional relations from one domain to another, especially if domains are semantically distant or, even worse, if domains are semantically similar, but syntactically different. In such cases, ontologies inference engine remains to rely on a limited set of rules from the first-order logic. One problem that directly arises from these limitations is the lack of so-called ontologies interoperability. Two although, semantically very close vocabularies usually only for syntactic differences can not be recognized as such therefore can not be linked, complemented or modified mutually.

Another drawback which results from modeling approach to the meaning of the concept by encapsulation into the concept structure, is the leanness of relations attributes. Given that the focus of creating the meaning is not on relations, but on the structure of concepts, there is no obvious need for more detailed semantic differentiation of relations. This, in turn, directly limits the mechanism for the analysis of relations graphs and subgraphs, and consequently all the other deficiencies of cognitive processing.

Finally, of no lesser importance is following limitation, not being flexible ontologies inference engine is unable to self-learn and reprogram itself on basis of it’s own work and feedback from users/teachers.

3. ACTIVE SEMANTIC MODEL (STRUCTURE)

An alternative approach to modeling of semantic features of data in a semantic network is based on the that the semantic designation of certain concept (network node) is contained in the semantic relations between that and other concepts. Based on this hypothesis a special model of semantic network which is called the *Active Semantic* Data *Model* - ASM has been designed and developed [13], [14]. In ASM semantic relations are modeled as separate data structures that are (structurally) independent of the concepts (network nodes). Moreover, the names of concepts that they connect are their attributes (expression 3). In addition to the concepts connected by them, semantic relations are modeled (defined) with eight more parameters.

$$\mathbf{A}(\boldsymbol{cpt_i}\leftrightarrow\boldsymbol{cpt_j})_{ij}\text{: } \{ctx_i,\ \boldsymbol{cpt_i},\ r_i,\ t,\ h,\ s,\ c,\ d,\ r_j,\ \boldsymbol{cpt_j}\} \qquad (3)$$

One of these eight parameters (ctx_i) is used to identify a semantic context to which given association belongs - the other seven parameters are considered correct for given context. (Fig. 2). Two parameters: *accuracy* (***h***) and *significance* (***s***) of the association represent so-called weight parameters used to define assumed accuracy and importance of semantic features stated by the relation for given context (Fig 2).

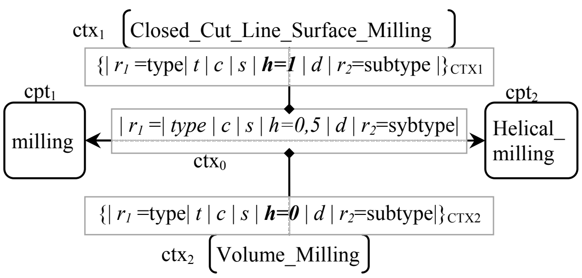

Figure 2. The value of associations' parameters depends on context

The remaining five parameters are named *topological* because in addition of stating a functional relation between two concepts (*type of semantic relation* ***t*** and their *roles* in it, $\boldsymbol{r_1}$, $\boldsymbol{r_2}$) and method of association, they define the topological features of associations graph (*character* - ***c*** and *direction* - ***d*** of associating).

4. SCD & SID BASED ON ANALOGIES

The procedure of SCD and SID in ASM is built on the claim that the use of analogy is the cognitive process core and can be considered as the primary process of cognition and communication. [15], [16], [17]. Examples of the most common heuristics as well as general strategies used in solving the problem are based on the use of analogies. [18]. In fact, well-known procedure that has proven successful in solving previous similar problems is used to solve the new problem. Condition for the success of this strategy is the *recognition* (*Einsicht* - a sudden perception of significant *relations* that leads to solutions to problems [19]) analogy between the two problems and the recall of previously implemented solutions. One of the reasons why sometimes it is difficult to recognize the analogy between the two problems is because their elements are different, although the relations are the same. Insight on the common elements of the solution is actually hiding in the similarity and/or identity of the relation between these elements [4]. This thesis was used to develop algorithms for SCD and SID in ASM.

SCD starts with *introducing* the concept whose semantic categorization is needed (the new concept) to ASM through the establishment of initial semantic relations between new concept and concepts that already exist in the network [13]. Initial introduction of a new concept with ASM is performed through so-called *communication context* - a kind of matrix that helps to define the initial associations plexus (subgraphs) between new concept and other concepts in the network.

The next step is *refinement* by allocating a subset of *significant* and *accurate* associations and for the the so-called general and the appropriate particular context (expressions 4, 5, and 6). Weighting parameters ***s*** and ***h***, and the attribute of belonging to a particular context ***ctx*** are used for precise control of the segmentation of semantic network where to look for an analog network patterns (associations plexus) on which SCD and SID will be executed.

$$\bar{S} = AVG\left(s_{A(CPT_X\leftrightarrow CPT_i)}, s_{A(CPT_N\leftrightarrow CPT_i)}\right) \qquad (4)$$

$$\text{where } \mathrm{A}_{(CPT_{X,N}\leftrightarrow CPT_i)} \in \{CTX_{\mathrm{i}}\}$$

$$\bar{H} = AVG\left(h_{A(CPT_X\leftrightarrow CPT_i)}, h_{A(CPT_N\leftrightarrow CPT_i)}\right) \qquad (5)$$

$$\bar{S}\cdot\bar{H} \geq T, \textit{ where T is a defined threshold} \qquad (6)$$

All those who are found to have sufficient accuracy and significance for given context (which are subject to the condition 6) shall be taken into consideration in further proceedings in SCD and SID.

After *refinement* SCD (and SID) implements the so-called procedure of *determining the parameters of association* - *DPA* [13]. This procedure results in the creation of new associations between the new and known concepts of the network, thus actually conducting semantic categorization of a new concept. As before launching SCD new concept has some initially created associations to known concepts in the network, SCD increases in the number of semantic relations that describe it ("deepening" of knowledge about the concept). On that occasion, with new associations, the knowledge of some known concepts is also being "deepened".

DPA procedure begins with the analysis of refined set of associations which consists of associations between new concept and existing concepts in the network (created in the communication context), and associations that continue to link these existing concepts with other concepts in the network. Of course, the procedure of DPA takes place simultaneously and separately for every identified context (Fig. 3). Determining the value of each parameter of the new association within the DPA procedure is conducted by a special subalgorithm (expression 7).

$$p\left(A_{CPT_X\leftrightarrow CPT_N}\right) = F_p\left\{\begin{matrix}\boldsymbol{P}\left(A_{CPT_X\leftrightarrow CPT_1}\right), \boldsymbol{P}\left(A_{CPT_1\leftrightarrow CPT_N}\right)\\ \cdots \\ \boldsymbol{P}\left(A_{CPT_X\leftrightarrow CPT_i}\right), \boldsymbol{P}\left(A_{CPT_i\leftrightarrow CPT_N}\right)\\ \cdots \\ \boldsymbol{P}\left(A_{CPT_X\leftrightarrow CPT_k}\right), \boldsymbol{P}\left(A_{CPT_k\leftrightarrow CPT_N}\right)\end{matrix}\right\} \qquad (7)$$

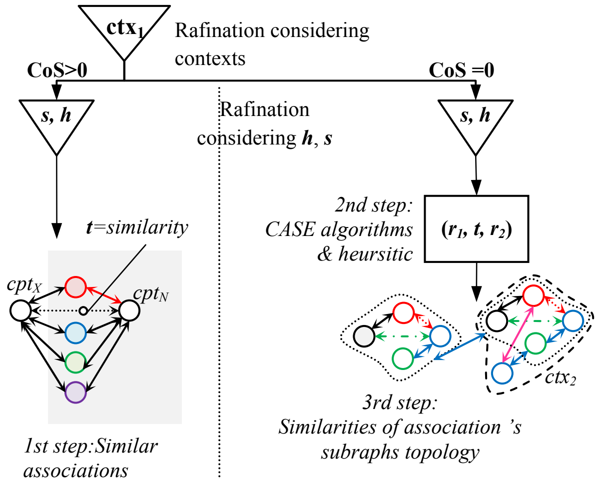

Figure 3. Branching off the *type* determination algorithm

The first in a series is a subalgorithm for determining type of association. This algorithm determines not only the type of new associations but also the role of each of the concepts that will be connected by the new association. It is implemented in three steps. However, before executing any of these three steps, additional refinement is performed. The set of association pairs whose types are the same (so-called first and upper classes of similarity CoS> 0) and a set of association pairs whose types are not the same (so-called zero class of similarity CoS = 0) are extracted.

In the first step is executing subalgorithms for determining other parameters of associations that will emerge between the new concept and those concepts whose association pairs are at least of the same type (Fig. 3). Resulting asociations from this branch of the DPA procedure are of different degree and class of *similarity*, or *differentness* (e.g., a new concept CPT_X and well known concept CPT_N are synonyms, antonyms, similar concepts with significant difference, incompletely opposite, etc.). What degree and class of semantic similarity/differentness of concepts are induced by these association depends on the obtained values of other parameters.

In the second and third step, DPA procedure processes associations pairs that are not of the same type (nor the role of concepts coincide). ASM decides the type of association that will connect the concepts CPT_X and CPT_N depending on the topological class of associations *plexus* i.e. kind of topology of association's subgraphs.

The second step is characterized by the execution of so-called CASE subalgorithms designed for pre-planned kind of association plexus (e.g. association plexus that describe an action, assembly, subject etc.). CASE subalgorithm is built based on the explicitly defined causal mechanism (rule) for each provided case of associations plexus. This part of the DPA is in many ways reminiscent of the coded first-order logic which ontologies inference engine uses for pre-planned cases of functional relations subgraphs. ASM offers the possibility of constant expansion of the CASE subalgorithms base, based on monitoring of the efficiency and factors of decisions usefulness (*knowledge crystallization algorithm* can not be presented in this paper due to its volume).

Still, true flexibility, autonomy and analyticity of the DPA procedure (i.e. SCD and SID) is shown in the third step of the subalgorithm for determining of association type. This part of the procedure is intended for upfront unforeseen topology types of association subgraphs. The algorithm is based on recognizing the similarity between topologies of association plexuses (subgraphs). ASM scans the network and extracts set of association plexuses $\{CTX_N\}$ which are topologically analogous with input association plexus CTX_x in which the new concept appears CPT^x_1 ($CTX_x \asymp F(CTX_N)$) (Fig 4.)

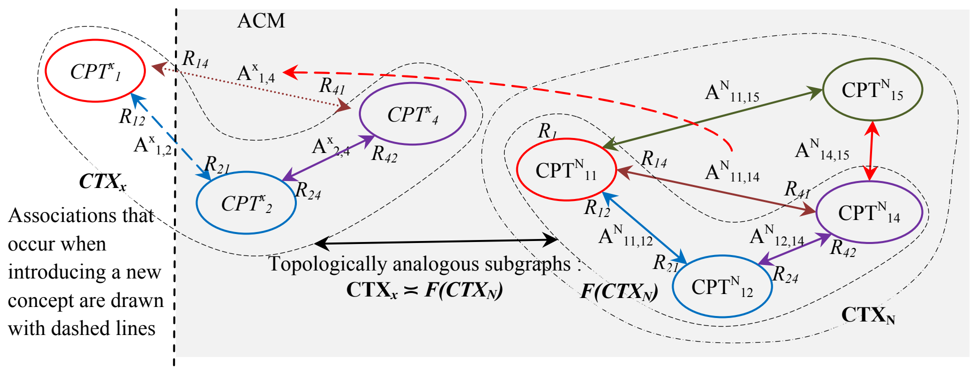

Figure 4. Type determination schema is based on similarity between the topologies of subgraphs of associations

ASM singles out this set of association plexuses by recognition of so-called *topologically corresponding associations* and *concepts*. Two associations that belong to different contexts and CTX_i и CTX_j and thereby have similar values of weight parameters and the same values of topological parameters are called *topologically corresponding associations* of two different contexts – TCAs. Concepts that belong to TCAs of two different contexts or association plexuses (CTX_i and CTX_j), and thereby have the same role in the TCAs, are called *topologically corresponding concepts* of two different contexts or TCCs. The algorithm for recognition of TCA and TCC is located in the core of ASM's SCD and SID. Basically it is a kind of algorithm for recognition of subgraph isomorphism, which was originally designed in Laboratory of Intelligent Production Systems of University of Nis [13].

Once it recognize the topological analogy between a simple input subgraph CTX_x and previously stored subgraphs $\{CTX_N\}$, a procedure for upgrading the input subgraph is preformed in accordance to the exemplar of the subgraphs that are topologically analogous to CTX_x.

ASM, in general, may allocate a number of different elementary subgraphs (basic association subgraph is one that contains two associations and three concepts) and their sets, which "offer different exemplars" for input subgraph upgrade. The suitability of selected exemplar associations is evaluated primarily by the degree of topological similarity of basic subgraphs (or sets of basic subgraphs) they belong to, and input subgraph (or a input set of basic subgraphs). The degree of topological similarity of exemplary sets of basic subgraphs and input set is measured by the number of exemplary sets of basic subgraphs that are topologically analogous. The exemplary association (and its type), which ASM first recommends is TCA that comes from the largest set of exemplary elementary subgraphs. This emphasizes the importance of the degree of topological similarity of the subgraphs of association for the type determination.

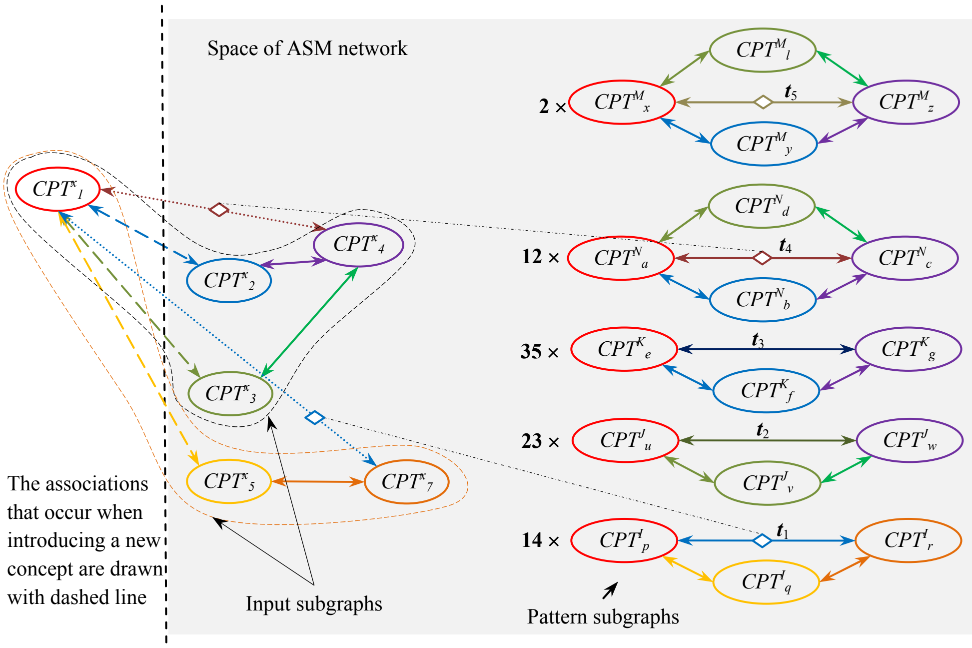

Figure 5. Type determination in the case of several candidates: basic subgraphs of associations of different type

$$\left\{\begin{matrix} CPT_1^x \leftrightarrow CPT_2^x \leftrightarrow CPT_4^x \\ CPT_1^x \leftrightarrow CPT_3^x \leftrightarrow CPT_4^x \\ CPT_1^x \left(\overset{t_4}{\leftrightarrow}\right) CPT_4^x \end{matrix}\right\} \asymp \left\{\begin{matrix} CPT_a^N \leftrightarrow CPT_b^N \leftrightarrow CPT_c^N \\ CPT_a^N \leftrightarrow CPT_d^N \leftrightarrow CPT_c^N \\ CPT_a^N \overset{t_4}{\leftrightarrow} CPT_c^N \end{matrix}\right\} \qquad (5)$$

$$\left\{\begin{matrix} CPT_1^x \leftrightarrow CPT_2^x \leftrightarrow CPT_4^x \\ CPT_1^x \leftrightarrow CPT_3^x \leftrightarrow CPT_4^x \\ CPT_1^x \left(\overset{t_3}{\leftrightarrow}\right) CPT_4^x \end{matrix}\right\} \asymp \left\{\begin{matrix} CPT_e^K \leftrightarrow CPT_f^K \leftrightarrow CPT_g^K \\ \\ CPT_e^K \overset{t_3}{\leftrightarrow} CPT_g^K \end{matrix}\right\} \qquad (6)$$

In the case that there is more topologically different sets of exemplary basic subgraphs that are equally topologically similar with the input set of basic subgraphs, ASM recommends exemplary association from that set of basic subgraphs which are more in the network. This emphasizes the importance of frequency, that is representativeness of certain topological class of subgraphs for determination of the type of association (expressions 7 and 8).

$$12 \times \left\{\begin{matrix} CPT_1^x \leftrightarrow CPT_2^x \leftrightarrow CPT_4^x \\ CPT_1^x \leftrightarrow CPT_3^x \leftrightarrow CPT_4^x \\ CPT_1^x \left(\overset{t_4}{\leftrightarrow}\right) CPT_4^x \end{matrix}\right\} \asymp \left\{\begin{matrix} CPT_a^N \leftrightarrow CPT_b^N \leftrightarrow CPT_c^N \\ CPT_a^N \leftrightarrow CPT_d^N \leftrightarrow CPT_c^N \\ CPT_a^N \overset{t_4}{\leftrightarrow} CPT_c^N \end{matrix}\right\} \qquad (7)$$

$$2 \times \left\{\begin{matrix} CPT_1^x \leftrightarrow CPT_2^x \leftrightarrow CPT_4^x \\ CPT_1^x \leftrightarrow CPT_3^x \leftrightarrow CPT_4^x \\ CPT_1^x \left(\overset{t_5}{\leftrightarrow}\right) CPT_4^x \end{matrix}\right\} \asymp \left\{\begin{matrix} CPT_x^M \leftrightarrow CPT_y^M \leftrightarrow CPT_z^M \\ CPT_x^N \leftrightarrow CPT_l^N \leftrightarrow CPT_z^N \\ CPT_x^N \overset{t_5}{\leftrightarrow} CPT_z^N \end{matrix}\right\} \qquad (8)$$

Belonging of each association to a particular context enables DPA algorithms to focus the analyses according to context. This approach in modeling the semantic content and the SCD procedure allows the same concept to be semantically interpreted differently in different contexts, which is very important feature of ASM.

5. RESULTS

The concept of semantic networks ASM and its analysis for SCD have been developed primarily in order to be used as add-in application in CAD / CAM systems, which should help the engineer in making certain decisions in situations that are characterized by significant uncertainty. Particularly interesting application of ASM could be in modules for reverse modeling and free-form design (bioforms and art design). The uniqueness of these forms intrudes the need for a system that is capable to interpret the geometry features of the model semantically and thus to help the designer to decide what features to use for further modeling and remodeling.

Preliminary tests of the application of the concept were carried out for workflow management system [20]. ASM and its algorithms for SCD and SID have shown the ability to recognize and categorize semantically unexpected exceptions that cause disruption to workflow. (e.g. the workflow of designing the bone implant). The following verification of the solution will include thorough tests that have to prove ability of ASM to automate and optimize selection and composition of the design features to create a geometric model of bio- or art-forms.

6. CONCLUSION

The concept of SCD and SID by determining the similarities of associations of the semantic network was created as a kind of modification or extension of known concepts of analogy based reasoning. SCD and SID that ASM's algorithms provide is based on recognition and determining the similarity of concepts, and this, in turn, on recognition and determining the similarity of associations between the concepts and similarity of subgraphs of associations. Determining the similarity of associations and subgraphs of associations is performed by the recognition of topological analogy between subgraphs of associations.

Compared to inference engine of OWL, ASM has autonomous, flexible and significantly more analytical mechanism of SCD. Besides the expandable base of CASE algorithms, which are used for efficient inferring in cases of known and predicted inputs and situations, ASM makes the meaningful judgments, conclusions and decisions in new and unexpected situations using the algorithms for determination of similarity of associations and topological analogy of association's subgraphs.

7. ACKNOWLEDGEMENT

The paper presents a concept that is a result of the multidisciplinary research from the domain of IT application in bioengineering. The research project is sponsored by the Ministry of Science and Technology of the Republic of Serbia - project id 41017 .

8. LITERATURA

[1] Kearney, A. T., "*Study Shows $750 Billion Waste of Time,*http://www.knowledgeboard.com/cgi-bin/item.cgi?id=44235, 2002.

[2] Krill P., "Overcoming Information Overload" InfoWorld, 7-Jan-2000, 2000.

[3] Peterson, L. R., Peterson, M. J., Miller, G. A. "*Short-term retention and meaningfulness*" Canadian Journal of Psychology, Vol 15, pp 143-147,1961.

[4] Kostić, A. *Kognitivna psihologija*, Zavod za udžbenike i nastavna sredstva, Beograd, 2006.

[5] Daconta, M. C., Obrst, L. J., Smith, K. T., *The Semantic Web.* Wiley Publishing, Indianapolis, 2003.

[6] ICT Summit. ICT-WP-Call3_Challenge2. Retrieved from http://www.ictsummit.eu/template/fs/documents/ict-wp-Call3_Challenge2_en.pdf, 2007.

[7] Hull, R., King, R. "*Semantic Database Modeling: Survey, Applications, and Research Issues*" ACM Computing Surveys , Vol 19 No. 3, pp 201-260, 1987.

[8] Quillian, R. M. "*Semantic Memory*", in M. Minsky (Ed.), *Semantic information processing*. MIT Press, Cambridge, 1968.

[9] Chandrasekaran, B. "*Representing function: Relating functional representation and functional modeling research streams*", Artificial Intelligence for Engineering Design, Analysis and Manufacturing, Vol. 19, pp 65–74, 2005.

[10] Fensel, D., *A Silver Bullet for Knowledge Management and Electronic Commerce*, Springer-Verlag, Heidelberg (Berlin), 2001.

[11] Mizoguchi, R., "*Tutorial on ontological: Engineering part 3 advanced course of ontological engineering*", New Generation Computing , Vol. 22 No. 2, 2004

[12] Kyoung-Yun, K., David, M. G., Hyungjeong, Y., "*Ontology-based assembly design and information sharing for collaborative product development*", Computer Aided Design, Vol. 38, pp 1233–1250, 2006

[13] Stojković, M. "*Analysis of the Manufacturability Parameters Based on Semantic Structures of the Digital Product Model*", PhD thesis, 2011.

[14] Stojkovich M., Manich M., Trajanovich M., Korunovich N., "*Semantic Structures In The Product Data Model*", Proceedings: Product Lifecycle Management - Assessing the industrial relevance PLM-SP3, pp 227-234, Milano, 2007.

[15] Gentner, D., Holyoak, K. J., & Kokinov, B. N., *The Analogical Mind: Perspectives from cognitive science*, MIT Press, Cambridge, 2001.

[16] Foxwell, H. J., & Menasce, D. A., "*MARVIN: A Web-Based System for Representing, Retrieving, and Visualizing Analogies*", World Wide Web: Internet and Web Information Systems, Vol. 7, pp 385-419, 2004.

[17] Bar, M., "*The Proactive brain: using analogies and associations to generate predictions*" Trends in Cognitive Science , Vol. 11 No. 7, pp 280-289, 2007.

[18] Mednick, S., "*The associative basis of the creative thinking*", Psychological Review, Vol. 69, pp 562-569, 1962.

[19] Köhler, W, *The Mentality of Apes*, Harcourt, Brace, New York, 1927.

[20] Mišić, D., Stojkovic, M., Domazet, D., Trajanovic, M., Manic, M., & Trifunovic, M., "*Exception detection in business process management systems*", Journal of Scientific & Industrial Research, Vol. 69, pp 188-193, 2010