Using Natural Language Generation to Support Interactive Concept Mapping

Harry Kornilakis, Kyparisia A.Papanikolaou, Evangelia Gouli, Maria Grigoriadou

Department of Informatics & Telecommunications, University of Athens,

Panepistimiopolis, GR-15784 Athens, Greece

{harryk, spap, lilag, gregor}@di.uoa.gr

Abstract. Learners can use concept mapping as a knowledge elicitation technique as they try to articulate and synthesize their actual states of knowledge during the learning process. In this paper we propose a framework that interactively supports learners to construct concept maps. In particular, learners construct a concept map using a list of concepts and relationships and when they ask for support, the system identifies errors on their map and accordingly annotates the map. Along with this visual information the system provides the learner with verbal information interpreting his/her wrong propositions in natural language. To this end, natural language generation techniques are used to formulate questions that reflect also the specific type of error by combining expert knowledge with learners' concept maps. The aim is to provide learners with multiple representations of their errors in order to stimulate them to reflect on their concept maps and correct them.

1 Introduction

Concept maps provide a means to capture, elicit and represent qualitative aspects of the learners' knowledge on specific topics [1], [2]. In such a map learners represent meaningful relationships between concepts in the form of propositions. In its simplest form, a concept map is composed of just two concepts connected by a linking word to form a single proposition. For example, "*Distance Education and Learning* <u>is a tool for</u> *Open Education*" represents a simple concept map forming a valid proposition about the concepts of "Distance Education and Learning" and "Open Education". Concept maps differ from other representational forms of meanings such as flow charts, organizational charts, semantic networks, in that none of these forms of maps are based on the theory of learning and theory of knowledge that underlie concept mapping strategies and their application to education [2].

Along this line, concept mapping is the process of organizing concepts in a hierarchical manner and forming meaningful relationships between the concepts. It promotes and assists meaningful learning by encouraging students to identify concept meanings, establish relationships between concepts, re-arrange the existing relationships, relate new concepts to prior concepts, organize the concepts in a hierarchical and integrated manner and refine the completed map resulting in generalized schemata for certain concepts [1], [2], [3], [4]. Thus, it requires from learners to reflect carefully on their understanding of important concepts and their relationships. Reflective thinking is controlled doing, involving a pushing and pulling of concepts, putting them together and separating them again [2].

There are several computer applications for editing and presenting concept maps providing learners with the appropriate tools like Inspiration (URL: http://www.inspiration.com/). Moreover, research has been conducted in assisting learners construct a concept map [5], [7], [8]. For example, in [5], the system gives appropriate hints to the learner according to the comparison between learner's and expert's concept maps. The hints are provided in a partial proposition type made of concepts and relationships represented on the learners' concept map. In [8] the system verifies learners' concept maps and provides learners with hints about specific errors, such as missing propositions and missing concepts. These hints are predefined and entirely controlled by the teacher.

In this paper we propose a framework for supporting learners to construct concept maps by stimulating them to reflect on their errors/misconceptions and correct them. In this context, concept maps become an important learning experience for learners as well as a unique evaluation experience. Learners construct a concept map using a list of concepts and relationships and when they ask for support the system identifies errors on their map and accordingly annotates the map. In addition to this visual information the system provides the learner with verbal information, transforming his/her wrong proposition in natural language and associating a question with it. To this end, techniques from the area of natural language processing have been employed to formulate questions and explanations by combining expert knowledge with learners' concept maps. In particular we employ the Functional Unification Grammar [11] framework for natural language generation in order to transform propositions from the learners' concept maps into English sentences.

In order to illustrate the proposed approach we use the results of a research study investigating learners' usual misconceptions about the subject of "Distance Education and Learning". This study was conducted at the Department of Informatics and Telecommunications of the University of Athens during the spring-semester of the academic year 2001-2002 in the context of a relevant postgraduate course [6]. In this research, after the experimenters identified the key concepts of the topic, they asked a sample of learners to construct concept maps using all or some of these concepts and encouraged them to add other relevant concepts. From these maps a number of valid propositions and many misconceptions, or invalid propositions as well were identified.

The paper is organized as follows. In Section 2 the categorization of learners' errors on which the generation of questions is partly based is presented. Moreover, the system-learner interaction procedure during the process of concept mapping is described. Section 3 presents the framework for supporting learners in the concept mapping process, analyzing its main components such as the knowledge base, the error detection algorithm, and the natural language generation mechanism. The paper ends, in Section 4, with concluded remarks.

2 Interactive Construction of Concept Maps

The construction of a concept map requires considerable creativity in organizing the structure of the map, selecting important, relevant concepts to add to the map and searching out salient cross-links, indicating relationships between concepts in different sections of the map [2]. Aiming to formulate a framework for evaluating learners' knowledge based on their concept maps, we analyzed a sample of learners' concept maps where we investigated repeated patterns of valid and invalid propositions. In particular we identified several common errors, which *instruction* should face, that led us to draw conclusions about learners' knowledge [6]:

- learners omit specific concepts (which are considered fundamental concepts of the subject matter) from their maps. The usual omission of specific concepts led us to the conclusion that these are *unknown concepts* to the learners,
- (a) learners use specific relationships between two or more concepts, which are not false but they do not correctly/fully address the relation of these concepts in the context of the subject matter, and/or (b) they do not relate two or more concepts denoting their relationship. These cases were both considered as evidence of *incomplete understanding*, and
- learners' *false beliefs* are signalled as:
  - learners relate two or more concepts (a) that should not be related, and/or (b) with incorrect relationships that lead to clearly false propositions;
  - learners use incorrect concepts in propositions;
  - learners construct propositions which are not false, but they are characterized as false due to the omission of other relevant propositions.

Aiming to develop an *evaluation* scheme, which supports and facilitates the identification and categorization of faulty propositions in a concept map, we classified learners' errors to the categories that are presented in Table 1. Moreover, this categorization is used as the basis for the construction of questions that reflect the different types of errors. In particular, a specific form of question is associated with each category of errors (for more details see Section 3.3). These questions aim to probe each learner's cognitive structure to ascertain whether or not misconceptions exist and if so, how they are related to other ideas held in the learner's mind [1].

Table 1. Categorization and interpretation of learners' errors based on the common errors that were identified by analyzing their concept maps.

Interpretation of learners' errors	**Categories of learners' errors**
Unknown Concepts	*Missing concept and its relationships* *when* specific concepts defined by the tutor as fundamental concepts [6] are omitted from the learner's concept map *E.g.* The concept of "Open Education" and its relationship with the concepts of "Distance Education and Learning" and "UK Open University" are missing from the learner's concept map.

Incomplete Understanding	*(A) Incomplete relationship* *when* the relationships between two concepts that appear on the learner's concept map are incomplete, i.e. several relationships have been omitted. *E.g.* the proposition "*Tutors* <u>teach</u> *Learners*" showed that the learners were not able to specify all the relationships between these two concepts, i.e. "Support", "Teach", "Advice", "Assess", as the role of the tutor is quite different in the context of Distance Education.
	(B) Missing relationship *when* the relationship between two concepts is missing on the learner's concept map. *E.g.* the concepts "Tutors" and the "Communication Means" are not related although they should be linked with the relationship "use".
False beliefs	*(A) Superfluous relationship* *when* two concepts are related even though they should not. *E.g.* the proposition "*Tutors* <u>determine</u> *Place of study*" is incorrect, as the concepts "Tutors" and the "Place of study" are not related. So, the relationship "determine" should be omitted.
	(B) Incorrect relationship *when* two concepts are related with an incorrect relationship which should be substituted. *E.g.* the relationship "is better than" between the concepts of "Distance Education and Learning" and the "Traditional Education" is incorrect and should be replaced by the relationship "operates supplementary with".
	(C) Incorrect concept *when* a concept is related to an incorrect concept which should be replaced with another concept. *E.g.* In the proposition "*Traditional education* <u>is offered by</u> *UK Open University*", the concept "University of Athens" should replace the concept "UK Open University".
	(D) Incomplete propositions *when* the relationship of a concept to other concepts is incomplete due to the omission of one or more propositions. *E.g.* The propositions "*Learners* <u>determine</u> *Place of study*", "*Learners* <u>determine</u> *Time of study* " should also include the proposition "*Learners* <u>determine</u> *Pace of study*".

The learner - system interaction procedure during the concept mapping process is as follows. Learners are asked to construct a concept map as a response to an open-ended question posed by the system. This beginning question is an open-ended one that cannot be answered by ‘yes’ or ‘no’ or a simple statement of fact following Novak’s suggestion [2]. A list of concepts and a number of relationships are given to the learner. We adopted this approach as we believe that providing a specific list of concepts and their relationships prevent the learners from floundering and constrain their thinking to “productive” directions.

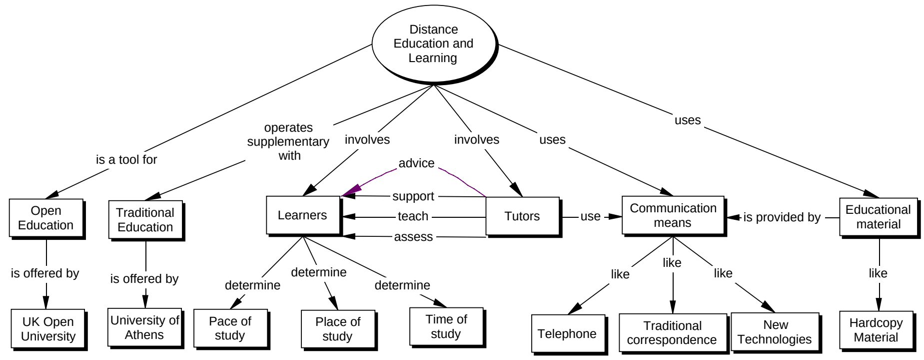


Fig. 1. The target concept map: Expert’s concept map

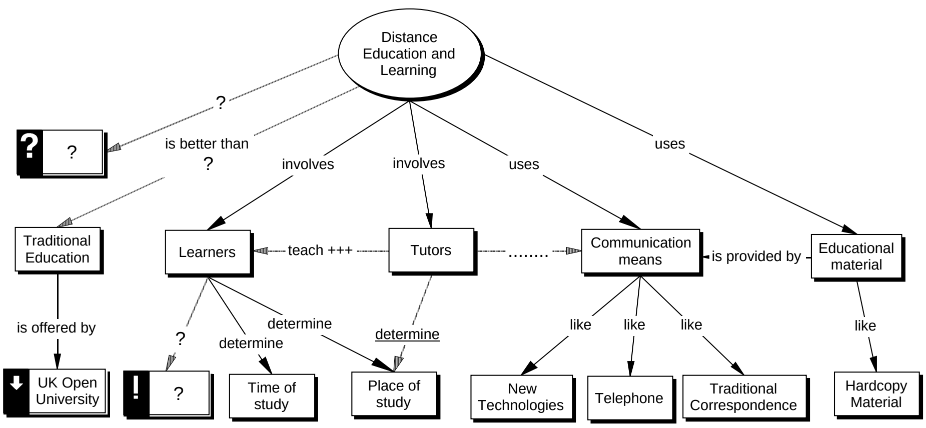


Fig. 2. An annotated concept map constructed by a learner which includes the categories of errors presented in Table 1: (i) *missing concept and its relationships* is denoted through the symbol “?” on the labels of the concept and the relationship, (ii) *incomplete relationship* is denoted through the symbol “+++” next to the label of the relationship, (iii) *missing relationship* is denoted through the symbol “…” on the label of the relationship, (iv) *superfluous relationship* is denoted by underlining the relationship, (v) *incorrect relationship* is denoted through the symbol “?” next to the label of the relationship, (vi*) incorrect concept* is denoted through the symbol ‘🡻’ on the label of the concept, and (vii) *incomplete propositions* are denoted by using the symbol ‘!’ on the label of the concept.

Moreover, the central concept of the concept map is provided – in our example the concept of “Distance Education and Learning” (see Fig.1 the root concept). Then ,the learner should choose from the concept list the appropriate concepts and arrange them on the map by selecting the appropriate relationships to reflect hierarchy.

After the learner has completed the map, or in case s/he asks for support, the system checks the map and graphically annotates the wrong propositions, if any, one by one. Learners’ errors, which can be identified and accordingly annotated (see Fig. 2) belong to the aforementioned error categories (see Table 1). At the same time, a question appears in natural language which reflects the error made by the learner, i.e. the system uses verbal representation of the already provided graphical annotation. In response to the question posed, the learner constructs a new proposition by adding new concepts/relationships and/or correcting the existing ones.

3 A Framework for Supporting Learners in Concept Map Construction

Once the learner has completed constructing the concept map or decides to ask for support, the system checks the map for any of the errors presented in Table 1. This is performed by comparing the learner's concept map to the expert-teacher’s concept map, which are stored in the knowledge base of the system. After detecting an error the system presents a question in English to the learner. A schematic representation of this interaction between the learner and the system and of the data flow within the system appears in Fig. 3.

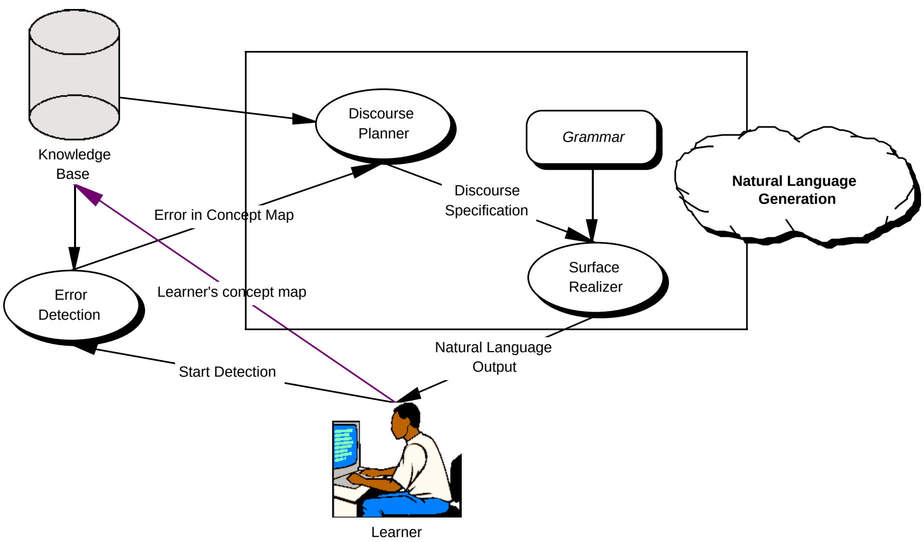


Fig. 3. Interaction between the learner and the system and the data flow within the system.

```
Insert root node in queue
While the queue is not empty repeat
    n <- the first node in the queue
    n' <- the node corresponding to n in the expert graph
    E <- the set of all edges coming out of n
    E' <- the set of all edges coming out of n'
    For every edge e' in E'
        d' <- the destination node of e'
        If e' not in E
            If there exists an edge in E ending on d'
                 If there exists an edge in the expert graph from n' to
                 d' other than e'
                     Return incomplete relationship error
                 Else
                     Return incorrect relationship error
            Else
                 Search the learner graph for d'
                 If found
                     Return missing relationship error
                 Else
                     Return missing concept and its relationship error
        Else
            e <- the edge corresponding to e' in the learner graph
            If the destination node of e is not d'
                If there exists a node other that d' in the expert graph
                connected to n' with an edge e'
                    Return incorrect concept error
                Else
                    Return incomplete concept error
    For every edge e in E
        If e not in E'
            Return superfluous relationship error
    Add all child nodes of n in the queue
```

Fig. 4. The error detection algorithm

3.1 Knowledge Base

In the knowledge base, the concept maps of both the expert and the learner are stored. A concept map is internally represented as a directed acyclic graph, where each node in the graph corresponds to a concept and each edge corresponds to a relationship between two concepts. Only one node without incoming nodes is allowed to exist in the graph. In particular, this is the central concept of the map, which was provided by the system at the beginning of the concept mapping process ("Distance Education and Learning" in Fig. 1). Additionally it is important that we make sure that no cycles appear on the concept map. This is achieved by running a cycle detection algorithm [10] on the graph every time the learner adds a new relationship to the concept map. Naturally, the same constraints also hold true for the expert's concept map.

3.2 Error Detection

The algorithm for detecting errors in a learner’s concept map is based on a breadth-first search through the corresponding graph stored in the knowledge base, beginning at the root node (central concept). A queue is used to contain the nodes that have not yet been searched. The algorithm for detecting errors in the learner graph appears in Fig. 4.

It is important to note here that when an error is found, the algorithm will not continue to search for more errors until it has been corrected by the learner. Combining this with the breadth-first nature of the algorithm we can be certain that it will always be possible to find the node corresponding to **n** in the expert graph. Furthermore since the input graph contains no cycles and is connected, the algorithm is guaranteed to end either after finding an error or after confirming that the learner's concept map matches the concept map of the expert.

3.3 Natural Language Generation Mechanism

After detecting an error on the learner's concept map, the system annotates that error on the map and along with this visual information it provides the learner with verbal information, presenting a question about the wrong proposition in natural language. In order to achieve this we use Natural Language Generation (NLG) techniques. Natural language generation may be defined as the process of constructing natural language output from non-linguistic input [12]. In our case the non-linguistic input is the concept map of the learner and in particular the type of error and the false proposition of the learner.

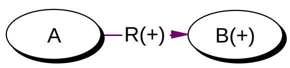
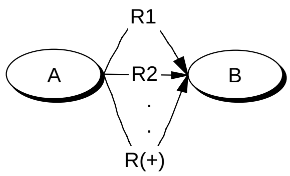
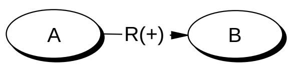
The natural language output is generated by two main components, which operate sequentially, the Discourse Planner and the Surface Realizer. The Discourse Planner initially produces a specification of the output sentence and then the Surface Realizer transforms this specification into a grammatically and syntactically correct English sentence.

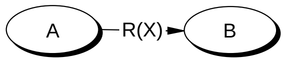
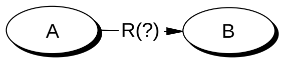
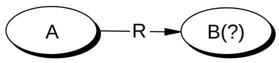
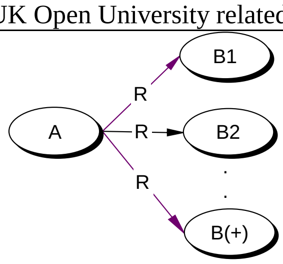
Discourse Planner. The operation of the Discourse Planner is to take the concept map as input and generate the specification of the question that will be given to the learner in order to help him\her correct the error on the concept map. This specification is expressed in terms of features [14] which follow a specific feature structure, which in our case is a subset of the feature structure for the English language defined as part of the SURGE grammar [11] (see also description of Surface Realizer). These features assume values that depend on the type of error that the learner has made and on the false proposition. In fact, it is very often the case that the values of features are either concepts or relationships taken from the proposition. The definition of these features and their values is done by using a subset of the Sentence Planning Language (SPL) [9], which offers an interface for specifying lexical information that defines English sentences. For example the proposition *"Distance Education and Learning uses Communication means"* is specified as:

```
(:PROCESS
    ("Use"
    :AGENT
          ("Distance Education and Learning"
          :NUMBER Singular
          :PERSON 3)
    :AFFECTED
          ("Communication Means"
          :NUMBER Plural
          :PERSON 3)
    :MOOD Declarative)
```

The actual specification depends to a large extend on the type of error made by the learner since the Discourse Planner has to generate the specification based on the type of error and on the false proposition. Depending on the error different features are used and depending on the proposition different values are assigned to those features. A sample of the used features and their assigned values for the various types of errors and propositions of the concept map of Fig. 2 appears in Table 2.

Table 2. Assignment of values to sentence features for the various types of errors and propositions accompanied by specific examples, where symbol "?" next to a concept / relationship means that it is wrong; symbol "X" that a concept or a relationship is superfluous on the learner's concept map; symbol "+" that it is missing from the learner's concept map.

Error categories	Visual representation of error	Sentence features and their values
Missing concept and its relationships		Process: "add" Agent: "you" Affected: B On-Location:"map" Mood: Interrogative (Wh) Modality: Possible
E.g. How can you add on the map the concept of Open Education?		
Incomplete relationship		Process: R1, R2 Agent: A Affected: B Determiner {Process}: "only" Mood: Interrogative (Yes-No)
E.g. Do the Tutors only teach the Learners?		
Missing relationship		Relation: "related to" Participant-1: A Participant-2: B Mood: Interrogative (Wh)
E.g. How are the Tutors related to the Communication Means?		

<table>
<tr><td>Superfluous relationship</td><td></td><td>Relation: "related to"
Participant-1: A
Participant-2: B
Determiner {Process}: "indeed"
Mood: Interrogative (Yes-No)</td></tr>
<tr><td colspan="3">E.g. Are Tutors indeed related to the Place of study?</td></tr>
<tr><td>Incorrect relationship</td><td></td><td>Process: R
Agent: A
Affected: B
Determiner {Process}: "really"
Mood: Interrogative (Yes-No)</td></tr>
<tr><td colspan="3">E.g. Is really Distance Education and Learning better than Traditional Education?</td></tr>
<tr><td>Incorrect concept</td><td></td><td>Relation: “related to”
Participant-1: A
Participant-2: B
Determiner {Process}: "really"
Mood: Interrogative (Yes-No)</td></tr>
<tr><td colspan="3">E.g. Is really UK Open University related to Traditional Education ?</td></tr>
<tr><td>Incomplete propositions</td><td></td><td>Process: R
Agent: A
Affected: B1, B2, ...
Determiner {Affected}: "Only"
Mood: Interrogative (Yes-No)</td></tr>
<tr><td colspan="3">E.g. Do Learners determine only the Place of study and the Time of study?</td></tr>
</table>

Surface Realizer. The Surface Realizer receives the sentence specification generated by the discourse planner and generates the individual sentences taking into account the language specific lexical and grammatical constraints. This is achieved by using the Functional Unification Grammar (FUG) [11] framework. The idea behind FUG is to build a Unification Grammar [13] based on a feature structure and then to unify this structure with the input specification, which is built using the same sort of feature structure. The unification process then takes the features specified in the input and reconciles them with those in the grammar, producing a feature structure which can be linearized to form a sentence in natural language.

As a grammar for our implementation we have used a subset of the SURGE grammar which is a very extensive and widely used Unification Grammar of the English language developed by Elhadad as part of his ADVISOR II system [11]. The sentence specification in SPL provided by the Discourse Planner is unified with the grammar producing the output sentence, which in our case is the question shown to the learner. In Table 2 we give an example of such a question for each type of error, based on the learner’s map that is illustrated in Fig 2.

4 Conclusions

Concept mapping enables learners to externalize their understanding of a domain. This process is inherently reflective as it requires from learners to reflect carefully on their knowledge of important concepts and their interrelations. In this paper we propose a framework for supporting learners to identify their errors and misconceptions on their concept maps and accordingly refine them. To this end, we propose the use of both visual information in the form of annotation of the concept map and verbal information in the form of questions posed in natural language which reflect learners' wrong statements. Building element of this framework is a comprehensive categorization of learners' common errors. This categorization provides the basis for annotating learners' errors on their maps and formulating questions that stimulate learners to further elaborate on their concept maps.

References

1. Novak, J.: Learning, creating and using knowledge: Concept maps as facilitative tools in schools and corporations. Lawrence Erlbaum Associates, Mahwah, NJ (1998)
2. Novak, J., Gowin, D.: Learning how to learn. Cambridge University Press, Cambridge, UK (1984)
3. Novak, J.: Concept mapping: A useful tool for science education. Journal of Research in Science Teaching *27* (10) (1990) 937-949
4. Fisher, K., Faletti, J., Patterson, H., Thornton, R., Lipson, J., & Spring, C.: Computer-based concept mapping. Journal of College Science Teaching *19* (6) (1990) 347-352
5. Chang, K., Sung, T., & Chen, S-F.: Learning through computer-based concept mapping with scaffolding aid. Journal of Computer Assisted Learning *17* (1) (2001) 21-33
6. Gouli, E., Gogoulou, A., Grigoriadou, M.: A Coherent and Integrated Framework Using Concept Maps for Various Educational Assessment Functions. Journal of Information Technology Education, Vol. 2 (2003) 215-240
7. Reichherzer, T., Canas, A., Ford, K., & Hayes, P.: The Giant: An Agent-based Approach to Knowledge Construction & Sharing. The Eleventh International Florida Artificial Intelligence Research Symposium Conference (FLAIRS), Sanibel Island, (1998) 136-140
8. Cimolino, L., Kay, J., & Miller, A.: Incremental student modeling and reflection by verified concept-mapping. Proceedings of the AIED2003 Learner Modeling for Reflection Workshop (2003)
9. Kasper, R. T.: *A Flexible Interface for Linking Applications to Penman's Sentence Generator*. In Proceedings of the DARPA Speech and Natural Language Workshop, Philadelphia, February 1989.
10. Johnsonbaugh R.: Discrete Mathematics (5th Edition). Prentice Hall (2000).
11. Elhadad M.: Using Argumentation to Control Lexical Choice: A Functional Unification Implementation. PhD Thesis. Columbia University (1992).
12. Vander Linden K.: Natural Language Generation. In Jurafsky D. & Martin J. H.: Speech and Language Processing. Prentice Hall (2000).
13. Kay M.: Functional Grammar. In BLS-79, Berkeley, CA (1979) 142-158.
14. Chomsky N.: Aspects of the Theory of Syntax. MIT Press, Cambridge, MA (1965).