

Conceptual reinforcement in concept maps using computer inference

Ernani Leite Ribeiro Filho¹, Orivaldo de Lira Tavares²,
Crediné Silva de Menezes^{2,3}, Juliana Hiroko Kowata

¹IFES, ²PPGI/UFES, ³UFRGS

eribeiro.unedserra@gmail.com, tavares@inf.ufes.br,
credine@gmail.com, juliana.kowata@gmail.com

Abstract. *The importance of a concept on a concept map is proportional to the amount of relationships in which it participates. By adding new relationships to a concept, we are highlighting it and reinforcing its role on the map, with several pedagogical applications. We use language processing techniques and rules for automatic inference from concept maps to identify and suggest new relationships on a map. This article presents a model for finding new relationships in concept maps; shows a software architecture that implements this model and reports some test cases and perspectives for future work.*

Resumo. *A importância de um conceito em um mapa conceitual é proporcional à quantidade de relacionamentos em que participa. Adicionando novos relacionamentos a um conceito, estamos destacando-o e reforçando seu papel no mapa, com várias aplicações pedagógicas. Usamos técnicas de processamento de linguagem e regras para inferência automática a partir dos mapas conceituais para identificar e sugerir novos relacionamentos num mapa conceitual. Este artigo apresenta um modelo para encontrar novos relacionamentos em mapas conceituais; mostra uma arquitetura de software que implementa esse modelo e relata alguns casos de teste e perspectivas para trabalhos futuros.*

1. Introduction

The concept map is a specific type of graphic organizer that students can use to clarify text structure and improve reading comprehension [Liu, 2014], [Kandiko et al., 2013], [Oliver, 2009] [Redford et al., 2012]. Concept maps (Cmaps) may also be defined as diagrams used to visually express the conceptual structure of a domain of knowledge, according to the way the author perceives it [Kandiko et al., 2013]. Cmaps are used to organize, describe and communicate knowledge through representation of concepts and the relationships between these concepts. Both the concept map construction process as its output, the map itself, are important elements in the evaluation of cognitive structures used by learners for organize, in a conceptual model, their perceptions about a proposed subject. The construction of the map, in its turn, reveals partial products of the cognitive processes involved in knowledge acquisition.

A concept map is an artifact that could be continuously improved in a never-ended process. In a collaborative context, even when one person considers a map as finished, it yet could be modified by understandings of others, resulting in new relationships and concepts. This is possible because each person has its own perception about a subject and, usually, those people with a high degree of mastery in a topic tend to find out more relationships when compared to those less experienced. So discovering new relationships in concept maps enriches the whole learning process and challenges

the learner to perceive concepts in different ways. These new relationships are called latent relationships.

There are many applications for this approach in computer supported learning environments, such as: (i) expanding a concept map cooperatively built by a group of students, for reference in educational activities that encourage reflection on concepts and relationships over a certain domain. In these activities, a digital resource to find out new relationship in a map can be used to explore, in a pedagogical way, the differences between the map of the student group and an expanded map with latent relationships; (ii) in scenarios where the teacher is much in demand or when there are many students, intelligent computational tools that use the latent relationships in concept maps of students are important to challenge students to explore a knowledge domain and have immediate feedback to improve the concept maps constructed; (iii) improving student knowledge on a particular subject.

On the premise that the more links a concept has, the more central it is in a map, we consider that by adding new connections to a concept, we are, in fact, highlighting it and reinforcing its role in the whole map. From a computational perspective, we are talking about using simplified language processing techniques and textual inference on the concept maps, in order to improve them, by using an engine to identify and suggest new relationships based on existing relationships in a concept maps. In this research, we named each of these new relationships as "latent relationship" in concept maps.

The present paper is organized as follows: the Section 2 presents an overview about concept maps; in the Section 3, we expose the definition and context of latent relationships and some correlated work; Section 4 details the design, implementations and evaluation of the proposed tool and, finally, Section 5 puts forward our final considerations and perspectives to future projects.

2. Concept Maps

Concept Maps are graphical tools used to organize, construct, express and communicate knowledge. Besides being represented by a diagram, concept maps are predominantly informal and textual, structured on natural language, and were originally created to be used by humans, specifically, by children in scholar age [Novak; Cañas, 2006].

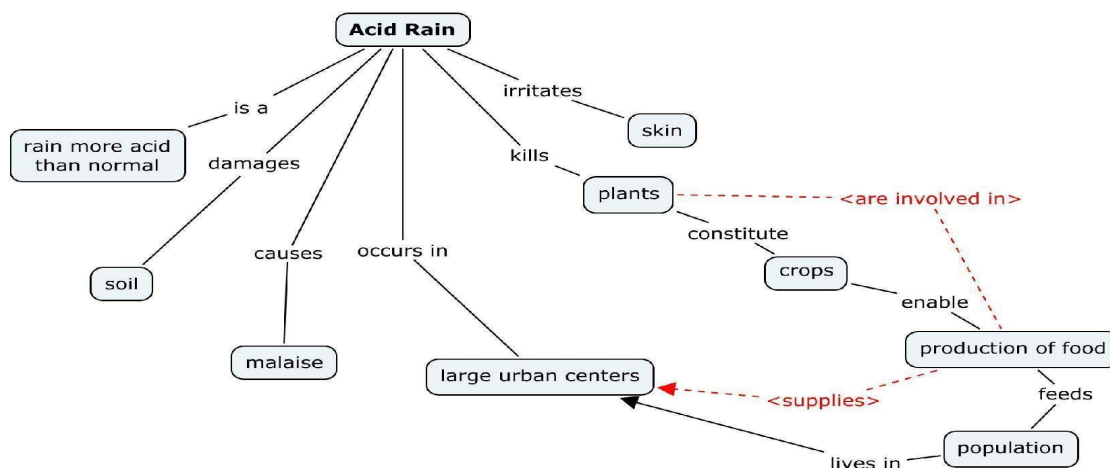


Figure 1 - A concept map about acid rain.

During the construction of a concept map, either in individual or collaborative context, a well-formed map exposes the results of a continuous and growing process of understanding in which partial maps reveal the cognitive state of an individual in a particular moment.

The semantic unit of a concept map is a proposition: a set made up of two concepts linked by a relationship that guarantees readability to a map. In Figure 1, the concepts "*Acid Rain*" and "*skin*" are linked by a relationship labeled as "*irritates*", constituting a proposition which is read as "*Acid rain irritates [the] skin*". This proposition is represented as a true fact inside the concept map.

3. The meaning of latent relationships on concept maps

Given a concept map " cm_1 ", propositions defined in it represent knowledge described by its author in an explicit way. In case another person reads the same " cm_1 ", this person might contribute by adding: (i) new concepts and relationships; (ii) new relationships based on existent concepts; (iii) new relationships based on existent relationships.

When new concepts and relationships between them are identified, a concept map meaning is enlarged with new information. On the other hand, when new relationships are defined based on existent concepts, details are inserted in a map, in order to complete it. But, when readers contribute by adding new relationships based on existent relationships, we consider that they are reasoning to find out implicit knowledge in a concept map, in a non-trivial inference process.

New relationships discovered based on existent relationships were labeled as "latent relationships" of a concept map. We consider them "latent" because these relationships are not described in an explicit way. They were implied in the interpretation of the map's reading.

Figure 1 shows some examples of latent relationships. Red dotted lines are latent relationships. To find out the proposition "*plants are involved in [the] production of food*" it is necessary to know previously that "*plants constitute crops*", that in its turn, "*crops enable [the] production of food*". The same is valid for "*production of food supplies large urban centers*", that is inferred because we already know that "*production of food feeds population*" and "*population lives in large urban centers*".

The more experienced the reader, the more latent relationships he or she can bring to the map. These occurs because what a person already knows about a knowledge domain, either conceptualizations or vocabularies, support him or her in the comprehension of a concept map, allowing him/her to find the missing links and to improve it.

Given a concept map about a subject " s ", a beginner in " s " will probably identify less latent relationships than an expert in " s ". Sometimes, latent relationships could be purposely omitted by an author, or yet, unnoticed by him or her.

In this paper, we propose a computer resource to play the role of an expert who should reason about the map produced by a beginner and find out latent relationships.

3.1. Correlated researches

Leake et al (2004) establish parameters to identify the most important concepts in a concept map, founded in a classification of concepts according with the number of

relationships that begins from or converges to them. They develop a module called "Suggest" in CmapTools [Novak, 2006] to validate this approach. The module uses a concept map to identify the most important concepts and based on them executes a search in Web or in a concept map collection to present a list with words correlated with the original concepts of a map. These suggested concepts can be used by the author to improve a concept map.

Lin (2002) researches the use of inferences in a database of common sense facts in order to generate new facts from it. A common sense is a type of non formal logic that includes premises and conclusions about entities and events related to daily issues. This researcher proposed some methods to enrich the knowledge base with new propositions. One of these methods represents propositions using a graph. By structuring a proposition in a graph format <agent, relationship, object>, a cycle with three templates are defined, each of them can be considered as an inference rule. In a cycle with three rules, in which "bats like darkness" and "you might find bats near cave interiors" are propositions, by applying some syntactic formulas as "?A relationship ?B", we are able to conclude that if "?A= bats" and "?B = darkness" and "?C = cave interiors", then "?A likes ?B", "you might find ?B near ?C", so by analogy "cave interiors have by property darkness" is a fact about "?C have by property ?B". Lin's work proves that it is possible to use graphs to find out new propositions based on a cycle of propositions.

Bessmertnyy e Kulagin (2009) use semantic nets to build knowledge bases structured in a topological form. The research applies the format Subject → Predicate → Object, to organize facts about a knowledge domain. After each fact gets fitted to the format, derived structures compile predicates in a standard format of facts f(subject, predicate, object) to be stored in a knowledge base. This knowledge base is supported by a dictionary of synonyms d, in which elements are represented as d(["word", "synonym", "synonym"]) used in order to promote disambiguation of words.

Yen-Ling Kuo and Jane Hsu (2010) merge facts from different knowledge bases to construct new propositions. The authors use analogies to search for similar concepts in different sources, in order to bridge the gaps between multiple knowledge bases and to make it possible to answer queries based on knowledge collected from multiple sources without a common ontology. New assertions are found by computing graph similarity to compute analogies across multiple knowledge bases. Considering two domains: Target Domain and Source Domain, concepts and relationships in the target domain are searched in the source domain, resulting in identification of different concepts used in similar relational structures.

3.2. A Model to Find Out Latent Relationships

Considering a concept map as a directed graph in which directions are used to read propositions in the map, Figure 2 illustrates a concept map and latent relationships (RLi) identified in it.

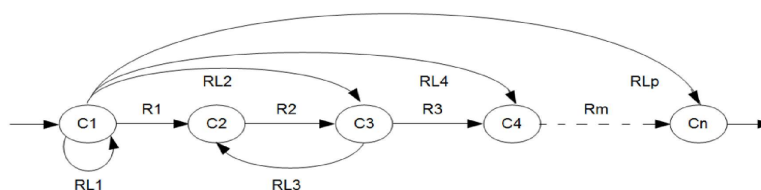


Figure 2 - Latent relationships in a concept map

In general, a latent relationship could be defined by edges that join concepts located in the beginning and ending of a set of propositions. In other words, if there is a path between two concepts that pass through other concepts, it is possible to define new relationships between the nodes that bind the path. Relationships in the path are used as premises to derive a name to the latent relationship.

Formally, a proposition is formed by a triple (C_i, R_j, C_k) where C_i and C_k are concepts and R_j is a relationship of a concept map, i and k are limited by the number of concepts and j by the total number of relationships in a concept map. If there is a triple $(Universe, is_comprise_of, Energy)$ it should be read as "Universe is comprised of Energy".

In Figure 2, RL1 is a relationship with reflexive property. RL3, in its turn, is symmetric to R2. Relationships like RL1 and RL3 are rarely represented in concept maps, because they express obvious relationships about concepts.

Transitive relationships are easily inferred from two propositions. For example, if $\{(Sun, is_greater_than, planet) AND (Planet, is_greater_than, Satellite)\}$, then it is simple to conclude that $(Sun, is_greater_than, Satellite)$. We are not interested in this kind of latent relationship because its contribution is insignificant, besides it is useless to pollute the map with obvious statements.

To define new relationships in a concept map based on explicit relationships we followed some premises: (i) new relationships cannot modify the intended meaning the author tried to express; (ii) new relationships should be inferred based on existing relationships and should not be obvious and redundant information that has been purposely omitted by the author.

Based on defined premises, in this research, we consider only latent relationships that link two or more propositions in a non transitive path because we believe that non trivial relationships are the ones that really matter in the learning process. Figure 3 puts forward an example of pattern used to find out latent relationships. Relationships "r4" and "r5" are valid relationships about a subject, from those the latent relationship "rL" could be inferred as a conclusion.

Finding a latent relationship such as "rL" in Figure 3, depends on the capability of the person who reads the map, i.e., it depends on a particular vision and previous knowledge about the theme. After including "rL" into a map, C4 and C5 were valorized. Also, a pattern for recognizing latent relationships was established: IF $\{(C_4, r_4, C_6) AND (C_6, r_5, C_5)\}$ THEN (C_5, r_L, C_4) .

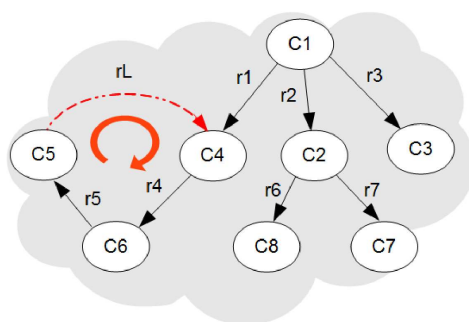


Figure 3 - A pattern to find latent relationships based on existent relationships

4. A Computer System to Find Latent Relationships

A computer system to find latent relationships, named as "Conceptual Reinforcement System" (CRS), was designed to be composed of two main modules: one responsible for constructing the facts and another for finding out latent relationships in a concept map. Figure 4 depicts the architecture of CRS.

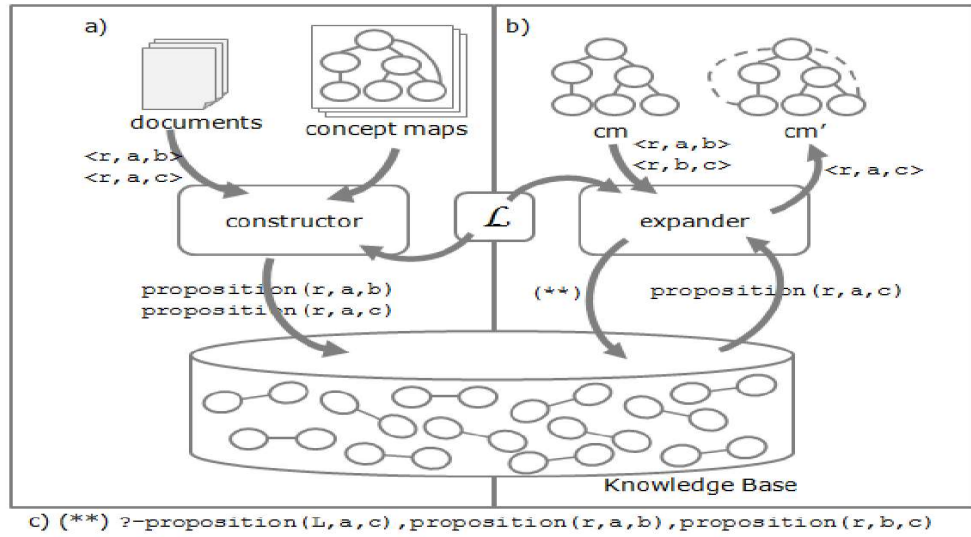


Figure 4 - Conceptual Reinforcement System (CRS)

Figure 4(a) shows the constructor module. Basically, it receives propositions represented as triples of the form $\langle \text{concept}, \text{relationship}, \text{concept} \rangle$, and transforms them into PROLOG facts that are stored in a Knowledge Base, following the pattern $\text{proposition}(\text{relationship_name}, \text{argument1}, \text{argument2})$.

Given a concept map, the expander module, illustrated in Figure 4(b), will detect all paths between two concepts that contain two linked propositions and share a same concept. For each found path, the expander does searches in the knowledge base, using propositions as parameters, as expressed in Figure 4(c). If a latent relationship is found, the system advises the user who accepts or not the system's suggestion.

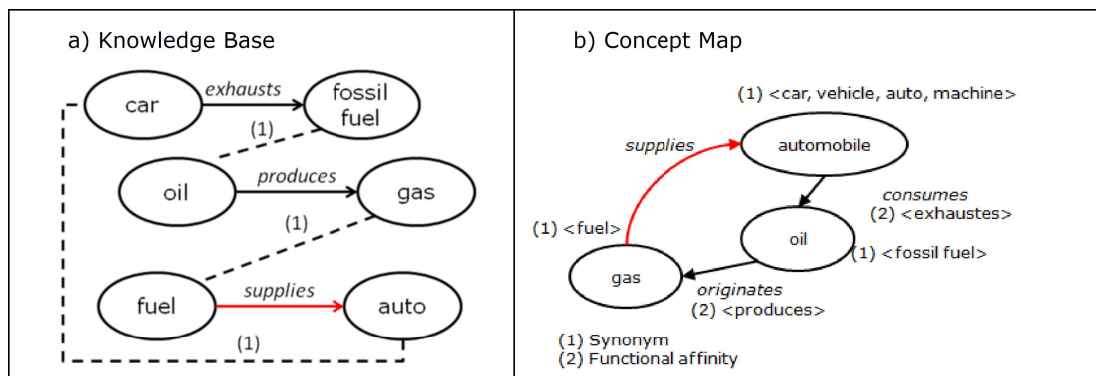


Figure 5 - Example of lexical resources to improve searches in a Knowledge Base

The expander module depends on linguistic resources in order to identify synonyms of words and verbs for concepts and relationships of a map, with the purpose of not restricting the searches only to existent concepts and relationships. The expander module could use synonyms from a Wordnet [Miller, 1995] and a Verbnets [Schuler, 2005] to improve its probability of producing results.

Figure 5 illustrates how the use of lexical resources could improve the search engine. Instead of using only the label of a concept, the search engine can use synonyms as parameters. For instance, in Figure 5, "gasoline, petrol" are synonyms of "gas", as well as "car, vehicle, auto, machine" are synonyms of "automobile".

In a WordNet each set of synonyms, called a "synset", has a unique identity. For example, the synset related to the word "house" and to all of its synonyms is 100432164. In CRS, this ID number is used to identify concepts of a concept map. Thus, all synonyms of a concept will be identified by the same ID number. Figure 6 shows how this ID number will be applied. In (a) an original concept map and in (b) the concepts were substituted by synset IDs.

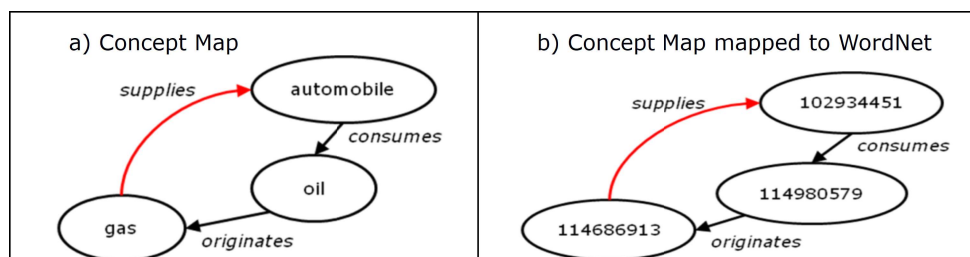


Figure 6 - WordNet used in a concept map

Classes	Classe (arquivo) allow-64
accompany-51.7.xml	<pre><?xml version="1.0" encoding="UTF-8"?> <!DOCTYPE VNCLASS SYSTEM "vn_class-3.dtd"> <VNCLASS ID="allow-64" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="vn_schema-3.xsd"> <MEMBERS> <MEMBER name="allow" wn="allow%2:32:00" grouping="allow.01"/> <MEMBER name="approve" wn="approve%2:32:00 approve%2:31:00" grouping="approve.01 approve.02"/> <MEMBER name="permit" wn="permit%2:32:00" grouping="permit.01"/> <MEMBER name="sanction" wn="sanction%2:32:01 sanction%2:41:01" grouping="sanction.01"/> <MEMBER name="tolerate" wn="tolerate%2:32:01" grouping="tolerate.01"/> </MEMBERS> <THEMROLES> <THEMROLE type="Agent"/> </THEMROLES> </VNCLASS></pre>
acquiesce-95.xml	
addict-96.xml	
adjust-26.9.xml	
admire-31.2.xml	
admit-65.xml	
adopt-93.xml	
advise-37.9.xml	
allow-64.xml	
appeal-31.4.xml	
appear-48.1.1.xml	
appoint-29.1.xml	

Figure 7 - VerbNet class allow-64

VerbNet implements the classes applied to the verbs. For example, the class allow-64 includes verbs related to the meaning of "left things/events happen", as in the verbs *allow*, *approve*, *permit*, *sanction* e *tolerate*, as showed in Figure 7.

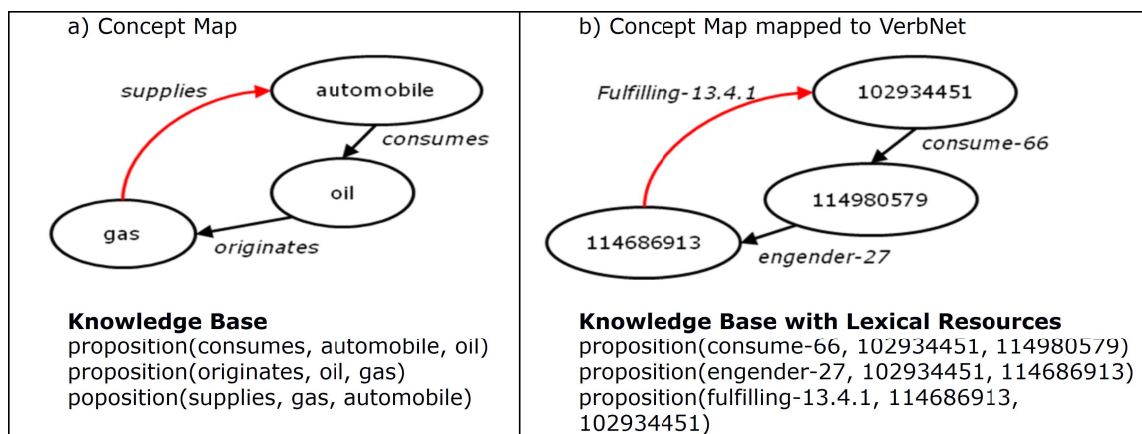


Figure 8 - Generating propositions with VerbNet classes

Figure 8 presents an example of use of VerbNet classes in a concept map. The original map in 8(a) is rewritten with VerbNet lexical units replacing the relationships,

as in 8(b). To populate a Knowledge Base with facts based on WordNet and VerbNet lexicon, the CRS needs to collect and store information suitable with these linguistic resources. Then, from a traditional concept map in Figure 8(a), CRS needs generate propositions in form of Figure 8(b).

The constructor module is responsible to populate the knowledge base and is able to read and process a concept map corpus to generate propositions with WordNet and VerbNet references. The expander module must be able to find out a latent relationship and translate the VerbNet reference into a natural language expression.

4.1 Tests

In order to validate the implemented system, an experiment was conducted using two distinct samples: S_1 and S_2 . The S_1 sample consists of 412 concept maps selected from the IHMC base, totalizing 15,143 explicit relationships that were used to populate our knowledge base. S_2 comprises 10 maps, totalizing 383 explicit relationships related to the Science domain, that were prepared to be processed in order to find out latent relationships.

In S_1 , the concept maps were processed before the experiment in order to maximize the probability of finding out latent relationships. In this pre-processing step the cmap: (i) maintained only relationships with a verb or a verb and a complement; (ii) maintained only concepts formed by a noun or a noun and an adjective; (iii) discarded concepts that don't have correspondences in the WordNet; (iv) discarded verbs that don't belong to any class in the VerbNet.

After the processing in maps of S_1 , the explicit relationships were reduced from 15,143 to 1,247, presenting a reduction of 92%. These 1,247 explicit relationships were stored in an empty Knowledge Base created by the constructor module.

In the second part of the experiment, the S_2 sample was submitted to the expander module.

From the 10 initial maps of S_2 just in 3 of them - "Fish", "Animals and Plants" and "Cells"- latent relationships were found:

(i) in the "Fish" map, from "*fish has gills*" and "*gills obtain oxygen*" was inferred: "*fish breathe oxygen*";

(ii) in "Animals and Plants", from "*animals need food*" and "*food is energy*" was inferred "*animals need energy*";

from "*plants produce food*" and "*food is energy*" was inferred "*plants produce energy*";

(iii) In "Cells", from "*chromossomes contain DNA*" and "*DNA defines proteins*" was inferred "*chromossomes define proteins*".

In this experiment, 4 latent relationships were found and added to the original 383 explicit relationships, i.e., about 1% of the total. In only 3 of the 10 maps the original map was expanded (only 30%).

Although the number of latent relationships found was small, the results prove that it is possible to find them in concept maps by using inference techniques based on a knowledge base and lexical resources. Furthermore, inferences built on existent premises can be considered relevant and coherent from a human perspective.

It is also important to notice that we worked with a reduced number of entries in the Knowledge Base. In order to address this deficiency it is necessary to populate the base with a larger number of samples. The larger is the Knowledge Base, the bigger is the probability of finding out latent relationships in a map. Another observation of our experiment was that due to concept maps being constructed in a free way, without formalisms, only a reduced number of concepts and relationships were considered apt to be matched with entries of our Knowledge Base and lexical resources.

5. Final Considerations

We consider that the exposition of latent relationships of a concept map is a product of a refined process of reasoning based on what a person understands about the map and on what a person already knows about its content.

The most expert individuals in a domain are able to propose more propositions in order to enrich the meaning of a concept map in this domain. Because of this, interaction between experts and beginners should be stimulated in the construction of concept maps. We propose that in a computer aided environment, newbie learners can be supported by a computer agent that plays the role of a more knowledgeable person.

In this research, we addressed a specific type of interaction in a computational environment of concept maps construction in which a computer agent could suggest latent relationships in concept maps based on explicit relationships in it and supported by a knowledge base and lexical resources.

To present new relationships to a concept map is not a trivial task to a machine. A human being is capable of doing this kind of task in a few minutes though. Human inference to find out latent relationships is based not only in what a person is reading on a map, but also in what he or she already knows about the content of the map: while this person is reading the map and identifying explicit concepts and relationships, its mind is, at same time, retrieving information in order to match with the elements of map.

A computational implementation of inference requires some stratagem to provide the computer with a satisfactory level of knowledge about a given theme. In our research, we created a knowledge base to circumvent this difficulty and we systematized an approach that allows the system to learn facts from concept maps. Also, lexical resources were added with the purpose of handling linguistic phenomena such as getting the stem of words, finding synonyms, getting the infinitive of verbs, etc.

This work showed that: (i) it is possible to populate a knowledge base with facts extracted from concept maps; (ii) it is important to use lexical resources in order to normalize words and handle linguistic phenomena existent in concept maps; (iii) it is possible to find out new relationships based on a concept map itself and in a knowledge base using inference.

Future works could be concerned with: (i) applying the same approach used in this research in other languages; (ii) comparing different maps and promoting their merging; (iii) implementing new classes to access lexical resources other than Verbnet and WordNet, such as FrameNet, ConceptNet e OpenCyc; (iv) develop ways to populate the Knowledge Base directly from documents in natural language; (v) use this approach in concept maps developed by students in a school setting.

6. References

- Bessmertnyy, I., A., Kulagin, V., S. (2009). Semantic network as a knowledge base in training system. Saint Petersburg State University of Information Technologies, Mechanics and Optics. Saint Petersburg, Russia.
- Cañas, A. J., Novak, J. D. (2006). Re-Examining The Foundations for Effective Use of Concept Maps. In A. J. Cañas & J. D. Novak (Eds.), *Concept Maps: Theory, Methodology, Technology*. Proceedings of the Second International Conference on Concept Mapping (Vol. 1, pp. 494-502). San Jose, Costa Rica: Universidad de Costa Rica.
- Kandiko, C., Hay, D., Weller, S. (2013). Concept mapping in the humanities to facilitate reflection: externalizing the relationship between public and personal learning. *Arts and Humanities in Higher Education*, 12(1), 70-87.
- Leake, D. B., Maguitman, A., Reichherzer, T. (2004). Understanding knowledge models: Modeling assessment of concept importance in concept maps. In R. Alterman & D. Kirsch (Eds.), *Proceedings of the twenty-sixth annual conference of the cognitive science society* (pp. 795-800). Mahwah, NJ: Lawrence Erlbaum.
- Lin, T. (2002). Analogical inference over a common sense database. In *Eighteenth national conference on Artificial intelligence*, pages 955–956, Menlo Park, CA, USA. American Association for Artificial Intelligence.
- Liu, P.L. (2014). Using eye tracking to understand learners' reading process through the concept-mapping learning strategy. *Computers & Education*, 78, 237-249.
- Miller, A.G. (1995). Wordnet: A lexical database for English. *Communications of the ACM*, 38(11), 39-41.
- Novak J. D., Cañas A. J. (2006), *The Theory Underlying Concept Maps and How to Construct Them* (Technical Report No. IHMC CmapTools 2006-01). Pensacola, FL: Institute for Human and Machine Cognition.
- Oliver, K. (2009). An investigation of concept mapping to improve the reading comprehension of science texts. *Journal of Science Education and Technology*, 18(5), 402-414.
- Redford, J. S., Thiede, K. W., Wiley, J., Griffin, T. D. (2012). Concept mapping improves metacomprehension accuracy among 7th graders. *Learning and Instruction*, 22(4), 262-270.
- Schuler, K.(2005). "VerbNet: A broad-coverage, comprehensive verb lexicon", Dissertations available from ProQuest. Paper AAI3179808. University of Pennsylvania.