

AN APPROACH TO COMPUTER-AIDED LEARNING ASSESSMENT

*Francisco Edson Lopes da Rocha, Júlio Valente da Costa Júnior, Eloi Luiz Favero
Federal University of Pará, Brazil*

Abstract. Advances in Computer Science have supported considerably better results in the development of educational environments. Such technologies are the foundation of tools and environments aimed at facilitating teaching, learning, and assessment. The goal of these initiatives is, ultimately, to favor learning and to respect students' idiosyncrasies, their unique ways of learning. In this paper we present an approach to computer-aided learning assessment: its supporting educational concept is *meaningful learning* and the computational tools used to implement it are *domain ontologies* and *genetic algorithms* (GAs); its main focus is *learning assessment* via *concept maps* (CMs). In the proposed approach, domain ontologies drawn by teachers for topics of disciplines are searched by a GA, which builds best-match populations of CMs semantically comparable to students' CMs. This way, we can assess individual learning, taking into account idiosyncratic forms of learning.

1 Introduction

In the area of school education, many researchers support the idea that the educational process can be greatly enhanced with the use of technologies resulting from advances in Computer Science (Jonassen, 1996). These technologies allow for the construction of computational environments that aim at facilitating teaching, learning, and sometimes learning assessment. With the dissemination of distance learning, however, learning assessment has become a constant concern. In large-scale virtual learning environments, teachers have to cope with the assessment of virtual groups of, for example, hundreds of students. Concerning the construction of educational environments, two fundamental problems arise: the choice of a learning theory to serve as the basis for the environment and, even more important, the identification of implementable aspects of this theory (Giraffa, 1999). In Educational Psychology, Cognitivism has played a major role in the last decades. It emphasizes the internal processes that lead to the construction of knowledge. Under Cognitivism, self-knowledge and awareness of idiosyncratic ways of learning are important issues, as people learn the same thing in different ways.

In the sixties, David Paul Ausubel, a cognitivist author, developed the *Assimilation Theory* (Ausubel, 1968; Ausubel, 2000). According to Ausubel, human beings learn meaningfully via acquisition and retention of concepts and propositions, which are stored in their cognitive structures in a particular, idiosyncratic way. This particular way of storing concepts and propositions is what forms the meanings human beings assign to experiences. A new meaningful learning process starts with the definition of an anchorage point in the cognitive structure, called *subsumer*, to which new concepts are connected. As a result, new learning essentially depends on the quantity and quality of the subsumers, as well as their stability and differentiability in the apprentice's cognitive structure (Ausubel, 1968). Applying adequate mental processes, called *progressive differentiation* and *integrative reconciliation*, human beings construct the knowledge stored in their cognitive structures. An immediate consequence of this theory are that humans can construct the same knowledge differently – using different subsumers, or different connections between new concepts and subsumers. Another consequence is that one of the major aspects of teaching is finding sufficiently mature subsumers able to serve as stable anchorage points to the new concepts.

Although apparently simple, Ausubel's ideas are not easily put into classroom practice without proper understanding of the processes through which people learn meaningfully, i.e., there is a fundamental necessity of learning the concepts of progressive differentiation and integrative reconciliation before applying them to usual school topics. Aware of this difficulty, Joseph D. Novak developed a pedagogical tool called *concept map* (CM) (Novak, 1998; Novak & Gowin, 1984). According to Novak, a CM represents part of a person's cognitive structure, revealing his or her particular understanding of a specific knowledge area. It contains concepts and propositions in graphical form, and it is constructed by the continued application of progressive differentiation and integrative reconciliation. This way, a sequence of CMs constructed by a person can illustrate the evolution of this person's understanding of the topic (Rocha & Favero, 2004). The step-by-step construction of a CM can also highlight personal preferences, as some people prefer to specialize new concepts from more general ones (in a top-down approach to learning), while others prefer to generalize new concepts from specialized instances (bottom-up approach).

Consequently, CMs are a viable, computable, and theoretically-sound solution to the problem of expressing and assessing students' learning. They can be used, among many other things, as alternatives to usual essays, decreasing the amount of work demanded from the teacher during assessment. Nevertheless, the assessment of

hundreds of CMs is still a considerable source of effort. Educational environments based on CMs focus on automating parts of this process.

In this article, we present an approach to computer-aided learning assessment via Concept Maps. It is based on Artificial Intelligence techniques (Rocha, Costa Jr., & Favero, 2004), such as ontologies and genetic algorithms. This approach was developed to facilitate the construction of CM-based environments aimed at teaching, learning and, most of all, assessment. It is an alternative to mere CM comparisons, which hinder personal ways of constructing knowledge. Some environments, based on CMs and aimed at assessment, have already been described in the literature (e.g. Araújo, Menezes, & Cury, 2003; Cabral & Giraffa, 2002). The general tendency of these environments is to compare the CM developed by student to a reference CM constructed by the teacher or by a specialist. This approach forces the comparison between potentially different – and potentially correct – understandings of the same reality. The result of this comparison can be used to assign a degree to a student, but it can hardly be considered learning assessment, from a cognitivist perspective.

As an alternative approach, we propose the use of two complementary Artificial Intelligence techniques: (i) domain ontologies, which serve as repositories of knowledge related to concepts of a topic of a discipline; (ii) a genetic algorithm (GA) capable of emulating the cognitive processes described in the Assimilation Theory, and capable of generating various CMs based on the domain ontologies, with the help of an inference engine. We suggest that domain ontologies, which concentrate the knowledge of an area or topic, be explicitly built by teachers or automatically extracted from a set of previously validated CMs. Furthermore, students are expected to express their conceptual learning, mapping the concepts used in the ontology. In order to do this, students can use a CM editor, for example. Eventually, students submit their CMs for automatic assessment. Detailed explanations on the roles of teachers and students in the approach, as well as how assessment is accomplished in an idiosyncratically, cognitivist-aware manner, are provided in the next sections of this article.

This article encompasses seven sections. Section 2 details the relationship between Ausubel’s Assimilation Theory and Novak’s Concept Maps. Section 3 describes the proposed approach from the teachers’ perspective. Section 4 describes the role of the students in the approach (working individually or cooperatively). Section 5 describes the steps necessary to accomplish assessment: the generation, by the GA, of populations of CMs comparable to the student’s CM, and the results generated by an assessor component. Finally, Section 6 describes an educational environment developed according to the approach proposed in this article (called CMTTool) and lists future research. Section 7 presents our final considerations.

2 The Assimilation Theory and Concept Maps

The Assimilation Theory (Ausubel, 1968; Ausubel, 2000) was developed by David Paul Ausubel in the 1960s and explains learning as the immersion of new concepts in the individual’s cognitive structure – a mental structure in which knowledge organization and integration are processed. The main concept of this theory is meaningful learning: a process in which new information is linked to some specific relevant aspect of the individual’s cognitive structure (subsumer). A CM about meaningful learning can be seen in Figure 1. Other basic principles described in the Assimilation Theory are progressive differentiation (in which learners increase the degree of elaboration of a concept as they increase their knowledge about it) and integrative reconciliation (when the learner identifies dimensions of relationships between components not previously connected), which are cognitive processes that explain the subsumption of new concepts in the cognitive structure. Integrative reconciliation can be of two distinct types: superordinate or combinatorial (Rocha, Costa Jr. & Favero, 2007).

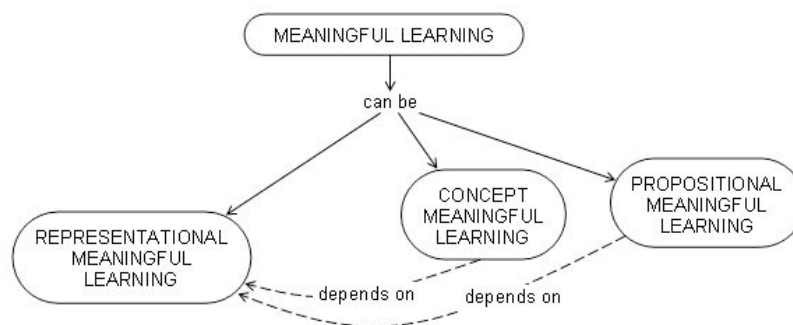


Figure 1. Concept map about meaningful learning (adapted from (Rocha, 2007)).

The systematic use of progressive differentiation, superordinate or combinatorial integrative reconciliation, during CM construction, can shed light on individual learning preferences. Some students prefer a top-down approach to learning, favoring progressive differentiations when constructing their CMs, while others tend to favor a bottom-up approach, making extensive use of integrative reconciliations. Awareness of personal learning styles allows for individualization of teaching (from the teacher's perspective) and more efficient learning, via self-knowledge (from the student's perspective).

CMs are semi-formal knowledge representation tools that use natural language to represent concepts and propositions. As such, they profit from the ease of creation and use: CMs have been used to teach a variety of different disciplines, to many different ages and teaching levels, including kindergarten (Mancinelli et al., 2004; Afamasaga-Fuata'I, 2004). They have also been used as a tool to organize and present information, for course or curriculum development, for navigation support, and for learning assessment. Nevertheless, this ease of use causes an undesirable side effect: ambiguity, which makes it difficult to assess the knowledge expressed in CMs (Costa Jr., Rocha, & Favero, 2004). Much has been done, however, in the field of CM disambiguation (Cañas et al., 2008).

As mentioned in the previous section, the assessment accomplished through mere comparison of a student's CM and a reference CM is not in accordance with cognitivist principles, as it forces students to construct their knowledge in a way that mimics the knowledge constructed by the teacher or expert who built the reference CM. This approach does not address the fact that humans construct knowledge in a number of different ways. An alternative is to compare students' CMs to populations of CMs generated by a mechanism responsible for building correct CMs based on an ontology.

Our approach recommends the use of ontologies to generate search spaces of possible correct CMs. These search spaces are then searched by a genetic algorithm (GA), responsible for finding the *best-match CMs*, i. e., the CMs in the search space that can be compared to the student's CM. Assessment is accomplished by analyzing the semantic difference between the student's CM and the CMs found by the GA. This is a general approach to learning assessment capable of coping with situations not addressed by the simple comparison of the student's CM to a reference CM. For example, a student who claims that "plants have leaves" will be assessed similarly to another who states that "leaves are part of plants" (Synonymy). Another student who states that "plants generate oxygen" will be assessed correctly, even if the ontology contains only the propositions "leaves generate oxygen" and "plants have leaves" (Inference).

3 Ontology Creation

The approach proposed in this paper recommends that domain ontologies, which concentrate the knowledge of an area or topic, be either explicitly built by teachers or automatically extracted from a set of previously validated CMs. Ontology mining has been a proficuous field in Knowledge Engineering in the last few years (Cheng et al., 2004). If automatic ontology consolidation is not provided, educational environments compliant with this approach should allow for the explicit creation of ontologies by teachers. In accordance with cognitivist principles, the role of teachers in the approach proposed in this article is to help students in their journey towards the construction of their own knowledge. Later on, teachers are expected to assess students' learnings, respecting idiosyncrasies. To help teachers, the approach recommends the use of a user-friendly tool to create domain ontologies related to the topics of their disciplines. The goal of this step is to allow for the creation of a repository of knowledge related to the topic being taught, in order to liberate teachers from the work overload deriving from the analysis of a multitude of CMs.

In the approach, domain ontologies are used to create common vocabularies for the different topics of a discipline. They are also used to create classification and semantic relationship rules between these concepts, so as to make it possible to infer new knowledge from the knowledge expressed in the ontology and, as a consequence, help in students' automatic learning assessment. In order to infer new knowledge from ontologies, environments developed according to the approach must make use of an inference engine, capable of making inferences from axioms described in the ontology. Figure 2 illustrates the graphical representation of a domain ontology about data communication. It was created in an ontology editor (On_Tool), which is part of CMTool (Rocha & Favero, 2004), an environment developed according to the approach described in this article.

On the right side of Figure 2, it is possible to see the taxonomy of linking phrases prescribed by the approach. It contains the possible semantic dimensions of relationships between concepts, and the linking

phrases that instantiate these dimensions. For example, the semantic dimension *process* can be instantiated by the linking phrases *is used by* or *is supported by*.

When teachers build the ontologies, they are not required to inform linking phrases, but only the semantic dimensions of relationships between concepts. This is one of the major differences between CM construction and ontology construction. However, in some cases, teachers can limit the linking phrases that can be used in the construction of propositions, in order to improve the accuracy of the CMs generated by the genetic algorithm. As detailed in Section 5, if a student's CM contains the propositions (i) <DIRECT COMMUNICATION has MANY PHYSICAL CONNECTIONS>, (ii) <DIRECT COMMUNICATION is characterized by MANY PHYSICAL CONNECTIONS>, or (iii) <DIRECT COMMUNICATION is not CHEAP>, all of them will be considered valid by the assessment mechanism, because propositions (i) and (ii) denote explicit knowledge (the *characterization* dimension can be validly instantiated by <is characterized by> and <has>), and proposition (iii) denotes knowledge validly inferred from the ontology.

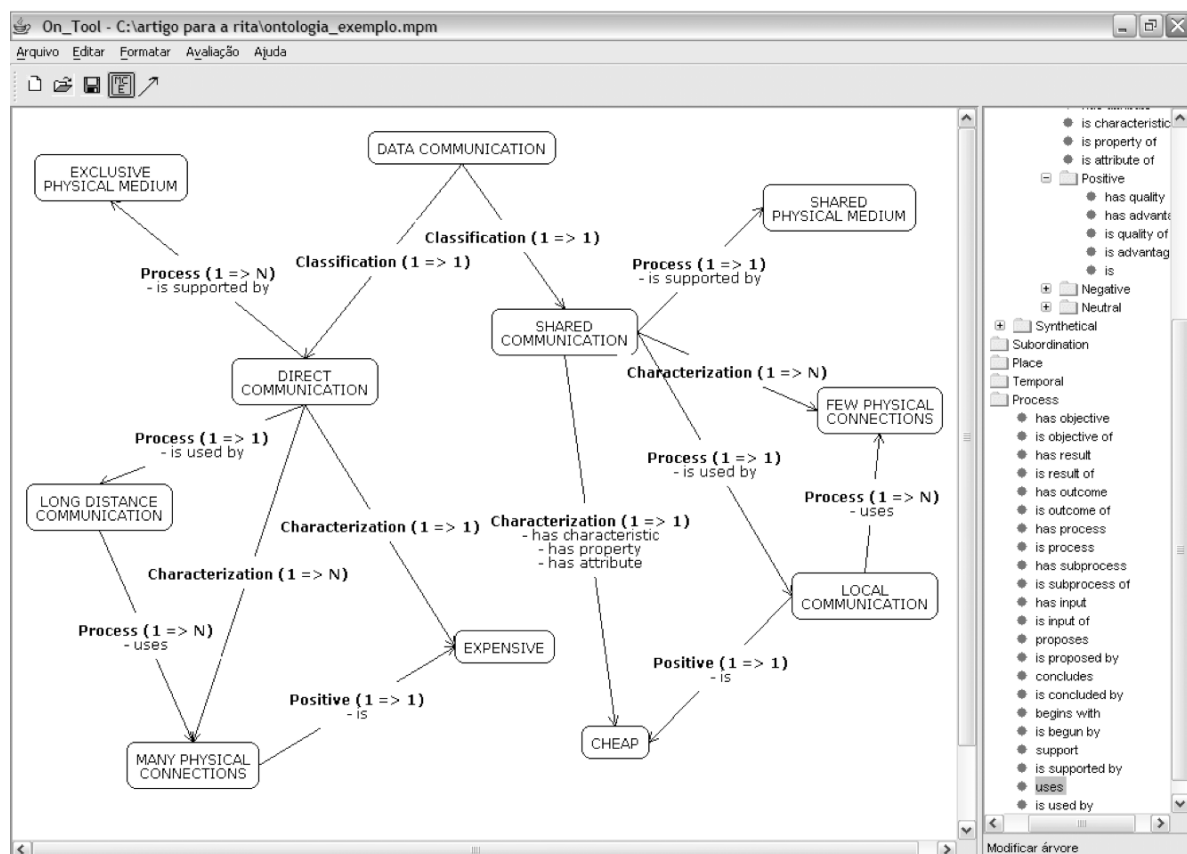


Figure 2. An ontology about data communication constructed in On_Tool.

If students construct propositions that cannot be validated by the assessment mechanisms based on the ontology, it is desirable that teachers be notified by the environment, as this event can denote the occurrence of valid propositions not expressed in – and not inferable from – the ontology. Consequently, teachers can begin a negotiation process with students, in order to reach an agreement about the validity of the suggested proposition. If the validity is confirmed, the environment should insert the proposition in the ontology, with the teachers' acknowledgement.

4 Concept Mapping

In accordance with the cognitivist view of meaningful learning, students are supposed to construct their knowledge by establishing semantic connections between the concepts related to the study of a specific topic of a discipline (described in the domain ontology built by the teacher). They are expected to construct their concept maps by the continued application of progressive differentiation and integrative reconciliation. The knowledge represented in the CMs can then be submitted for learning assessment.

In order to help students, environments compliant with the approach should provide a CM editor for student construction of concept maps. The CMs should use the concepts of an ontology stored in the environment. After construction, they should be submitted for assessment. Figure 3 presents a CM built from the concepts of the ontology illustrated in Figure 2. This CM represents a student's understanding of the topic, and can be submitted to the learning assessment mechanism prescribed by the approach.

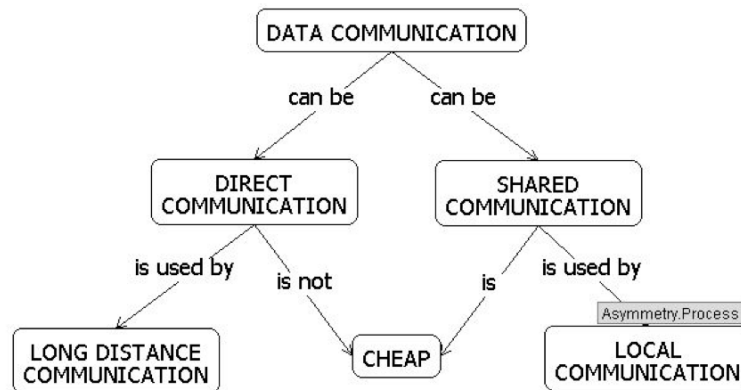


Figure 3. A concept map about Data Communication.

When a CM is being built, environments should help students during the choice of linking phrases, because this is the moment in which mappers explicitly define meanings. This can be done by showing pre-categorized semantic dimensions – and their respective linking phrases – to students. When a student decides to connect two concepts, he/she must be aware of the semantic dimension under which these concepts will be connected. This step is crucial, because it is the input for other definitions, like the choice of the most inclusive concept in a proposition, i.e., the concept that will be the subject of the assertion corresponding to the proposition in the CM.

In Figure 3, the student chose to connect the concepts SHARED COMMUNICATION and LOCAL COMMUNICATION with the *process* dimension, and the linking phrase that instantiated this dimension was <is used by>. Under these circumstances, the most inclusive concept is SHARED COMMUNICATION. The student could have chosen another semantic dimension for the relationship between these two concepts as, for example, the *classification* dimension. With this choice, the student would be able to construct the proposition <SHARED COMMUNICATION can be LOCAL COMMUNICATION>. However, this proposition is expected to be considered inaccurate by the assessment mechanism, because it is not supported by the underlying ontology. Figure 3 shows the complete taxonomy dimension of the linking phrase chosen by the student. This dimension should be stored internally, and the final CM should present only the linking phrases, for readability purposes.

Variations of the steps described in this section can be used during the learning process. Among other possibilities, it is possible to use interdisciplinary ontologies, contextualize and assess CMs under more than one ontology, or assess CMs produced collaboratively by a group of students, as a result of meaning negotiation among them.

5 Assessment

In this section, we present an example that shows the functioning of the GA. For this learning task, the teacher constructed the ontology illustrated in Figure 2. Internally, the ontology should be stored as axioms, in order to allow for inferences (and further exploration of the search space by the GA).

Based on the axioms, the GA can generate propositions (using concepts from the ontology and linking phrases from the taxonomy) and ask the ontology if they are valid. Valid propositions should be stored for posterior creation of individuals (CMs) in the populations generated by the GA. These individuals have to be evaluated, and a fitness value must be assigned to each one, based on its semantic difference from the student's CM (the GA privileges CMs that use the same concepts and phrases found in the student's CM). The final objective is to find a set of *best-match* CMs: those that are similar to the student's CM and valid according to the ontology. For more information on GAs, please refer to (Goldberg, 1989), and for a detailed description of the inner workings of the GA defined by the approach, refer to (Rocha, Vieira, Costa Jr., & Favero, 2004).

The first step taken by the GA, after a CM submission for assessment, is to generate several propositions (sets of two concepts connected by a linking phrase). Each proposition has to be evaluated by an inference engine, based on the ontology. Invalid propositions should be discarded. Valid propositions should be kept for creating populations of CMs. CMs are, thus, formed by grouping propositions in a number equal to the propositions present in the CM submitted for assessment.

Individuals in the population (CMs created in the previous step) have to be evaluated according to a fitness function that measures their semantic difference to the student's map (maps similar to the student's are scored highly). Afterwards, the GA selects the best individuals to be the parents of the next population. The next generation is created with the propositions (genes) of the parents. The best individuals of the previous generation are kept in the current generation. Additionally, as in nature, mutation should be allowed with a certain probability. When a mutation occurs, the GA should use a new proposition (formed from the ontology and from the taxonomy, and considered valid by the inference engine). Mutations are an important part of GAs, as they allow for further exploration of the search space, and inhibit premature convergence. This process should be repeated until a best-match CM is found. The best-match CM is then presented to the student as an alternative to the initial CM. The GA ensures this map is similar to the student's, so that it is simple for the student to analyze possible misconceptions in the original CM submitted for assessment.

Table 1 presents excerpts of a CM assessment. The results are organized in four parts. Part (a) reports if the concepts used in the CM submitted for assessment are related to the learning task underway, and if the learner's propositions are valid in the context under analysis. In order to do this, relationship dimensions should be validated. Part (b) presents the semantic comparison of the assessed CM to the best-match CMs generated by the GA. The objective is to present to the learner valid forms of mapping the knowledge represented in the ontology of the learning task underway. The semantic distance between the assessed CM and each one of the best-match CMs should be calculated. If any of the calculated values is different from zero, detailed information containing the possible alternatives to the identified misconception should be presented to the learner. Part (c) details the actions taken by the GA to construct the best-match CMs presented in part (b). The objective is to show to learners how to construct forms of knowledge representation alternative to their own (presented in part (a)). Finally, part (d) presents the list of concepts that, although present in the domain ontology, were not mapped by the learner. The list may indicate the need for reinforcement of specific topics of the discipline.

Assessment Results
<p>a) Hierarchical Structure and Learning Types 1. Assessed CM: {<DT COMM,can be,DR COMM>,<SH COMM,has type,DT COMM>} 1.1. Concepts: { DT COMM, DR COMM, SH COMM} 1.2. Propositions: p1=< DT COMM, can be, DR COMM >, p2=<SH COMM, has type, DT COMM > 1.3. Valid Hierarchies: {<DT COMM,Asymmetric.Definition.Synthetical.Classification,DR COMM >} 1.4. Invalid Hierarchies: {<SH COMM,Asymmetric.Definition.Synthetical.Classification,DT COMM >} 1.5. Valid Propositions: p1=<DT COMM,can be,DR COMM> 1.6. Invalid Propositions: p2=<SH COMM,has type,DT COMM> ...</p>
<p>b) Semantic Analysis 1. Best-Match CMs Generated by the GA: CM1={<DT COMM,can be,DR COMM>,<DT COMM,has type,SH COMM>} 1.1. Concepts: CM1 Ì { DT COMM, DR COMM, SH COMM} 1.2. Propositions: CM1 Ì {p1=<DT COMM,can be,DR COMM>, p2=<DT COMM, has type, SH COMM>} ...</p>
<p>c) Actions for the Reconstruction of the Best-Match CMs CM1: Create propositions p1, p2 Combine propositions p1, p2 (differentiate <DT COMM> progressively) ...</p>
<p>d) Ontology Concepts Absent in Assessed CM: {EXCLUSIVE PHYSICAL MEDIUM, LONG DISTANCE COMMUNICATION, MANY PHYSICAL CONNECTIONS, EXPENSIVE, CHEAP, SHARED PHYSICAL MEDIUM, FEW PHYSICAL CONNECTIONS, LOCAL COMMUNICATION}</p>

Table 1: Excerpts of a CM assessment

the automation of cognitivist practices. We are aware that enhancements can be made and our next goal is to further develop the approach.

References

- Afamasaga-Fuata'i, K. An Undergraduate Student's Understanding of Differential Equations through Concept Maps and Vee Diagrams. In: Proceedings of the First International Conference on Concept Mapping, Pamplona, Spain. 2004.
- Araújo, A. M. T., Menezes, C. S. & Cury, D. Apoio Automatizado à Avaliação da Aprendizagem Utilizando Mapas Conceituais." In: Proceedings of XIV Simpósio Brasileiro de Informática na Educação. Rio de Janeiro, RJ: Sociedade Brasileira de Computação. 2003.
- Ausubel, D. P. Educational Psychology: A Cognitive View. New York: Holt, Rinehart and Winston. 1968.
- Ausubel, D. P. The Acquisition and Retention of Knowledge. New York: Kluwer Academic Publishers. 2000.
- Cabral, A. R., Giraffa, L. M. Avaliação de Cursos WBT Utilizando Mapas Conceituais. In: Proceedings of XIII Simpósio Brasileiro de Informática na Educação. Porto Alegre, RS: Sociedade Brasileira de Computação. 2002.
- Cañas, A. J., Valerio, A., Lalinde-Pulido, J, Carvalho, M & Arguedas, M. Using WordNet for Word Sense Disambiguation to Support Concept Map Construction. Available: <<http://cmap.ihmc.us/Publications/ResearchPapers/SPIRE-2003%20-%20Using%20Wordnet%20for%20Word%20Sense%20Disambiguation%20to%20Support%20Concept%20Map%20Construction.pdf>> . Access: mar. 2008.
- Cañas, A. J., Hill, G., Carff, R., Suri, N., Lott, J., Gómez, G., Eskridge, T. C., Arroyo, M., Carvajal, R. CmapTools: A Knowledge Modeling and Sharing Environment. In: Proceedings of the First International Conference on Concept Mapping, Pamplona, Spain. 2004.
- Cheng, J., Sun, S., Tracy, A., Hubbell, E., Morris, J., Valmееkam, V., Kimbrough, A., Cline, M., Liu, G., Shigeta, R., Kulp, D., Siani-Rose, M. NetAffx Gene Ontology Mining Tool: a Visual Approach for Microarray Data Analysis. In: Bioinformatics, 12, 20(9). 2004.
- Costa Jr, J. V., Rocha, F. E. L. & Favero, E. L. Linking Phrases in Concept Maps: A Study on the Nature of Inclusivity. In: Proceedings of the First International Conference on Concept Mapping, Pamplona, Spain. 2004.
- Giraffa, L. Uma Arquitetura de Tutor Utilizando Estados Mentais. PhD Thesis in Computer Science. Instituto de Informática, Federal University of Rio Grande do Sul, Brazil. 1999.
- Goldberg, D. E. Genetic Algorithms in Search, Optimization & Machine Learning. Addison-Wesley. 1989.
- Jonassen, D. H. Computers in the Classroom: Mindtools for Critical Thinking. New Jersey: Prentice-Hall. 1996.
- Mancinelli, C., Gentili, M., Priori, G. & Valitutti, G. Concept Maps in Kindergarten. In: Proceedings of the First International Conference on Concept Mapping, Pamplona, Spain. 2004.
- Novak, J. D. Learning, Creating, and Using Knowledge: CMs as Facilitative Tools in Schools and Corporations. New Jersey: Lawrence Erlbaum Associates. 1998.
- Novak, J. D. & Gowin, D. B. Learning How to Learn. New York: Cornell University Press. 1984.
- Rocha, F. E. L. Avaliação da Aprendizagem: Uma Abordagem Qualitativa Baseada em Mapas Conceituais, Ontologias e Algoritmos Genéticos. PhD Thesis. Centro Tecnológico, Federal University of Pará, Brazil. 2007.
- Rocha, F. E. L. & Favero, E. L. CMTool: A Supporting Tool for Conceptual Map Analysis. In: Proceedings of World Congress on Engineering and Technology Education, Santos, Brazil. 2004.
- Rocha, F. E. L., Costa Jr, J. V. & Favero, E. L. Qualitative Learning Assessment via Concept Maps. In: Revista de Informática Teórica e Aplicada, 14(1). 2007.
- Rocha, F. E. L., Costa Jr, J. V. & Favero, E. L. A New Approach to Meaningful Learning Assessment Using Concept Maps: Ontologies and Genetic Algorithms. In: Proceedings of the First International Conference on Concept Mapping, Pamplona, Spain. 2004.
- Rocha, F. E. L., Vieira, R. V. L, Costa Jr., J. V. & E. Favero. Especificação de um Algoritmo Genético para Auxiliar na Avaliação da Aprendizagem Significativa com Mapas Conceituais. In: Proceedings of XV Simpósio Brasileiro de Informática na Educação. Manaus, AM: Sociedade Brasileira de Computação. 2004.