Adaptability through Artificial Intelligence Techniques

Maikel León[1], Zenaida García[1], Koen Vanhoof[2] and Rafael Bello[1]

[1]Central University of Las Villas, Cuba
[2]Hasselt University, Belgium

[1]{mle, zgarcia, rbellop}@uclv.edu.cu
[2]koen.vanhoof@uhasselt.be

Abstract. The use of pedagogical methods combined with Information and Communication Technologies produces a new quality that favors the task of generating, transmitting and sharing knowledge. In that case we have the pedagogical effect that produces the use of Concept Maps, which are considered a learning technique as a way to increase meaningful learning in the sciences. Also used for the knowledge management as an aid to personalize the Teaching-Learning process, to exchange knowledge, and to learn how to learn. Concept Maps provides a framework for making the internal knowledge explicit in a visual form that can easily be examined and shared. In this paper the authors present different approaches to elaborate Intelligent Systems, applied in some areas, in each approach Concept Maps and Artificial Intelligence are combined, using in the first one the Case-Based Reasoning and in the other Bayesian Nets as a knowledge representation forms and inference mechanisms for the decision making, supporting the Student Model. The authors also show the use of other techniques like Petri Nets, Neural Nets, and Fuzzy Cognitive Maps with the goal of the Interface Adaptability. The proposed models have been implemented in computational systems that have been successfully used by laymen in the Computer Science field to generate them owns adaptive systems.

1 Introduction

To construct and to share knowledge, to learn significantly, and to learn how to learn, are ideas on which researchers have been pondering for a long time as well as the use of tools that would allow taking these aspirations into practice. To achieve this, different techniques and strategies have been used. A Concept Map (CM) provides a schematic summary of what is learned and they order it in a hierarchic range. The knowledge is organized and represented in all the levels of abstraction, the general knowledge goes to the upper part and the most specific one goes to the lower [1].

CMs have increasingly got a great popularity and its integration with the technologies of the information and the communications have become a very important element in the plans for the improvement of the teaching-learning process. They have also extended its use to other areas of the human activity in which the management and the use of knowledge play an important role. The main application of the CMs has had effect in the teaching-learning process, which is its basic intention, it is important to point out that the CMs lead the attention of both students and professors on the restricted number of important ideas in which they must concentrate. The CMs are based on an instrument that combines the scientific rigidity with simplicity and flexibility, producing a general approval in the audience of students and professionals; this represents an important nexus between pedagogy and technology. Also, it constitutes an important aid for those who generate, transmit, store, and divulge information and knowledge and it comprises an important tool to obtain a highest practical value in the systems of the teaching-learning process.

The field of the Intelligent Teaching-Learning Systems (ITLS) is characterized by the application of Artificial Intelligence techniques, to the development of the teaching-learning process assisted by computers. If the key feature of an ITLS is the aptitude to adapt itself to the student, the key component of the system is the Student Model, where the information associated to the student is stored. This information must be inferred by the system depending on the information available: previous data, response to questions, etc. This process of inference is identified as diagnosis, and is undoubtedly the most complicated process inside an ITLS, who uses the Artificial Intelligence techniques to represent knowledge, to shape the human reasoning, to emphasize the learning by means of the action, to combine experiences of resolution and discovery, to be able to solve problems by their own, to formulate diagnoses and to provide explanations. So, they count on a bank of instruction strategies which helps to decide what and how to inform to the student to get an effective direction.

2 General motivations for using Concept Maps

Managing the knowledge assets of an organization requires capturing and retaining useful knowledge and making it available in a usable form when it is needed in the future. This process is complicated by difficulties in acquiring and representing knowledge, in accessing relevant knowledge, and in reapplying prior lessons to new situations. These issues are particularly acute in capturing and utilizing “internal” knowledge assets embodied in the experiences of task experts. Different technologies offer different benefits for addressing these problems. Concept Maps provide a framework for capturing experts' internal knowledge and making it explicit in a visual, graphical form that can be easily examined and shared. The Concept Maps constitute a tool of great profit for teachers, investigators of educational topics, psychologists, sociologists and students in general, as well as for other areas especially when it is necessary to mange with large volumes of information. They have become a very important element in the plans for the improvement of the Intelligent Teaching-Learning Systems and they have also extended its use to other areas of the human activity in which both management and the use of knowledge take up a preponderant place. If to define a Concept Map is relatively simple, it is simpler to understand the meaning of the definition.

The Concept Maps are descendant of the synoptic chart that the teacher of senior education would write on the black board for the students to get a better understanding of the lesson. It was defined by Novak, his creator, as a skill that represents a strategy of learning, a method to get the gist of a topic and a schematic resource to represent a set of conceptual meanings included in a set of propositions [2]. It is necessary to point out that there is not only one model of Concept Maps, several may exist. The important point is the relations that are established between the concepts through the linking words to form propositions that configure a real value on the studied object. For such a reason, in a concept there may appear diversity of real values. In fact, it turns very difficult to find two exactly equal Concept Maps, due to the individual character of knowledge.

The Concept Maps can be described under diverse perspectives: abstract, visualization, and conversation. Since significant learning is reached more easily when the new concepts or conceptual meanings are included under wider concepts, the most used Concept Maps are the hierarchic ones, the most general and inclusive concepts are placed in the upper part of the map, and the progressively more specific and less inclusive concepts, in the lower part. The subordinated relations among concepts may change in different fragments of learning, so in a Concept Map, any concept may rise up to the top position, and keep a significant propositional relation with other concepts of the map. The use of Concepts Maps designed by the professor increases both learning and retention of scientific information. The students produce maps as learning tools. Considering that the Concept Maps constitute an explicit and clear representation of the concepts and propositions, they allow both teachers and students to exchange points of view on the validity of a specific propositional link and to recognize the missing connections in the concepts that suggest the need of a new learning. For this reason, this skill has complemented so favorably with the practice of distance learning which presupposes that students and teachers are not physically in the same place at the same time. Concept Maps have particular characteristics that make them amenable to smart tools. These include:

1. Concept Maps have structure: By definition, more general concepts are presented at the top with more specific concepts at the bottom. Other structural information, e.g. the number of ingoing and outgoing links of a concept, may provide additional information regarding a concept's role in the map.
2. Concept Maps are based on propositions: every two concepts with their linking phrase forms a "unit of meaning". This propositional structure distinguishes Concept Maps from other tools such as Mind Mapping and The Brain, and provides semantics to the relationships between concepts.
3. Concept Maps have a context: A Concept Maps is a representation of a person' understanding of a particular domain of knowledge. As such, all concepts and linking phrases are to be interpreted within that context.
4. Concepts and linking phrases are as short as possible, possibly single words.
5. Every two concepts joined by a linking phrase form a standalone proposition. The proposition can be read independently of the map and still "make sense".
6. The structure is hierarchical and the root node of the map is a good representative of the topic of the map.
7. The students who analyze Concept Maps will have a wider basic knowledge; therefore, they will be more prepared to solve problems in comparison to those students who learn by memorizing.

The Concept Maps have turned into a useful instrument for teacher training and for the student's understanding of diverse subjects. From our view point, the Concept Maps constitute an instrument that merges the scientific rigidity with the simplicity and flexibility. It represents an aid for those who generate, transmit, store and spread information and knowledge. They also constitute an important tool to achieve a practical value, especially in the Artificial Intelligence systems.

3 Case-Based Systems and Bayesian Nets

In the traditional Case-Based Reasoning the solution of a problem is made taking the examples of cases stored in the case memory through the implementation of a function of distance or similarity, depending on the domain. The cases are composed by a set of attributes that describe the problem, among them the predictive features with the solution given to the problem described. The representation of the case is defined taking into account the nature of the problem, that is to say, the important attributes, the problems to be processed, the proposed solution, etc. Also, it is necessary to define the mechanisms for the retrieval of cases. The most similar case will be the one closest to the current problem, depending on its attributes; it will be the case which evaluation of the target function or function of similarity takes the best value understanding as the closest value, the value of the target function that denotes minor difference between the current case and the evaluated case.

The functioning of the Case-Based Reasoning comprises related processes, so problems and their solutions can be used to derive solutions to new problems, these processes are: the retrieval of similar cases, the suitability to the proposed solution and the storage or incorporation of the new solution to the new case. The retrieval process consists of determining the most similar cases found in the case base in order to give solution to new cases: the practice to determine the value of similarity between two cases include a wide range that involves methods such as counting the number of similar predictive features and others that consider the importance of the predictive features within the function of similarity [3]. After obtaining the value of similarity between the cases stored in the case base and the new problem, those cases to be considered in the construction phase of the new solution are selected. The suitability is the process of modification of solutions that have already been retrieved in order to give solution to the new problems. The Case-Based Systems have a group of features for creation of Intelligent Teaching-Learning Systems such as the acquisition of knowledge, the flexibility in the representation of knowledge, the preservation of knowledge and the reuse of previous solutions.

On the other hand, BNs are powerful tools both for graphically representing the relationships among a set of variables and for dealing with uncertainties in expert systems see [4] for an introduction to this field. Although the BNs are powerful in the diagnosis problems, their application in the student model is not very frequent in comparison with the great number of developed systems. We consider that the cause is that the application of the BNs demands much more effort than the application of other models of approximate reasoning (certainty factors, fuzzy logic) and the development of heuristic for define and update the student model. This additional effort is mainly caused by the specification of the net (structure and parameters) [5]:

- Specification of a BN demands a careful study of the variables that take part in the system and the relations of causal influence among them. In addition, once the net has been defined we have to estimate the conditional probabilities that are normally a quite great number of parameters difficult to consider.
- Difficulty to implement the algorithms of propagation of probabilities, which besides being more or less complex are very expensive computationally.

In our case the first factor is not a problem, because the structure of the net is determine by the structure of the CM, taken in consideration its structural similarity. This is one of the facilities that we obtain when we combine the BN and the CM. To make an inference we use the Union Tree probabilities propagation algorithms that use the conglomerates in order to reduce the computational complexity. Using the combination of the BNs and CMs we facilitate the specification of the BN. First we define a CM (its structural part) and later with the experience of the professor we obtain the tables of conditional probabilities (its parametric part), diminishing the additional effort to implement a BN. The professor who develops the CM, must formulate an initial questionnaire, that has to be able to catch the student's cognitive state, turning it in a customized CM, where the student "navigate" in an oriented way according to his knowledge and not in a free way like in the traditional CM. And precisely this characteristic allows the CM to adapt the interaction to its user's specific needs. The variables that take part in the calculation of the prior and conditional probabilities are:

1. The evaluation of the questions of the questionnaires.
2. The results obtained from the Psicopedagogical Assistant.

The number of variables conforms a table of n combinations, where n is a natural number that can be considerably great; which constitutes a difficulty whose solution could reduce the space of initial representation, so that if there are superfluous variables, it is analyzed if they stay or not, according to its importance from the methodological point of view. An alternative solution to the selection of problem variables is the calculate the reducts by the algorithm QuickReduct, that combines elements of Genetic Algorithm with Estimation of Distribution Algorithms [6].

4 Architectures of both tools

In this section the authors describe how they use Artificial Intelligence techniques combined with CM to generate ITLS through HESEI and MacBay [7], describing the functionalities of the principal modules of each tool.

The ITLS that are created by HESEI correspond with a CM, with the particular feature that in some of its nodes there appears a questionnaire, capable to get the cognitive and affective state of the student and able to guide his navigation, creating this way an "Intelligent" CM. Figure 1 shows the HESEI architecture that includes Case-Based Reasoning, and an algorithm of Patterns Recognition to obtain the implementation of the Student Model with a previous selection of characteristics.

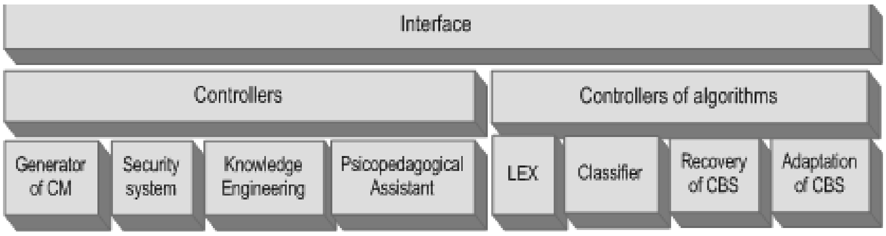

Fig. 1. HESEI architecture

Brief description of some components of HESEI:

Interface: Is provided with an editor that allows the teacher to introduce all the necessary information to prepare the ITLS. Through the Interface the system catches the cognitive and affective state of the student (Student Model). Also, through this component the students may be able to interact with the ITLS generated by HESEI using part of the information provided by the professor which best satisfies the students' need. Figure 2 shows a window of the HESEI interface with an ITLS generated for the subject of Binary Heaps in the course of Data Structures.

Knowledge Engineering and LEX: These components helps the teacher in the complex and intense work of knowledge engineering, using a Pattern Recognition, algorithm to reduce the space of initial representation (features that shape the Student Model), calculating the typical testors (discriminates features) which analyzes each questionnaire in order to facilitate the teacher the information of those questions that are worthless, they are eliminated by the teacher if they do not have a methodological value either. It was implemented the LEX algorithm for this task. A questionnaire is composed for n questions, and the teacher from the questionnaire conceives m categories (Student Model), in fact the experience has demonstrated that m is always very much minor than n.

Retrieval and CBS adapter: In this component a Case-Based System was implemented, it is composed by case-base, the retrieval algorithm and case-adaptation algorithm.

On the other hand we have MacBay, a tool to build ITLS; it is composing as we show in the Figure 3, it is composed by two fundamental modules, which allow the specifications of the functions of both users that will interact with the ITLS constructed the Professor Module and the Student Module. In the Professor Module is the module Generator of CM that offer to the professor, whom it does not have to be necessarily expert in Computer Science, a graphical interface to construct its ITLS helping itself of a CM. MacBay offers the possibility to use an assistant to make this task in order to facilitate the work of its users. When the CM is constructed the professor can assign a set of didactic materials to one specific node of the CM that is associated to the specified subject. For each node the professor can assign a questionnaire in order to evaluate the performance of the student in the subject through the Question Module.

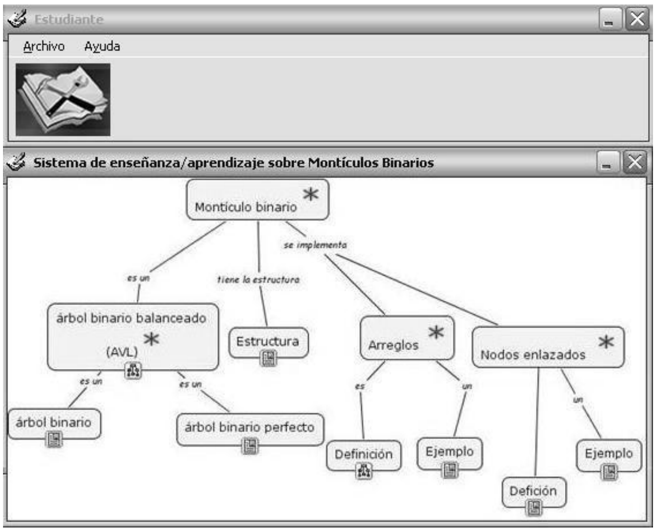

Fig. 2. HESEI Interface

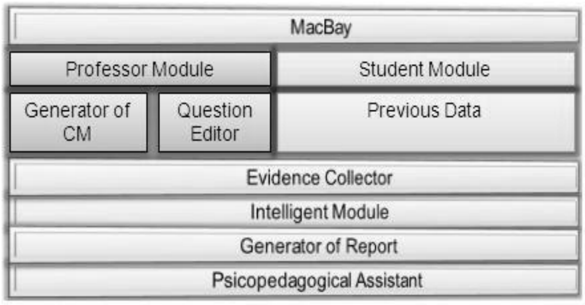

Fig. 3. MacBay Architecture

On the other hand the student can interact with the ITLS constructed having in consideration its previous performance in the subject by means of the facilities that offers the module Previous Data, which have to archive the data of all the students registered in the course to construct the distribution probabilities function associated to the BN. All the actions conducted by the professor or the student are compiled and it analyzes if these constitute an evidence by the module Evidence Collector. If evidence was found it information is passed to the Intelligent Module in order to make the inference process using the algorithm of Probabilities Propagation in Union Tree. The result of the inference is accessible for the student and for the professor. It's to be seen in report form, this is made by the Generator Module Report. This report is used by the Psicopedagogical Assistant with objective to help the student and the professor to determinate the strategy that they can choose during the teaching-learning process. Figure 4 shows an example of a CM for the Mathematical Logic course.

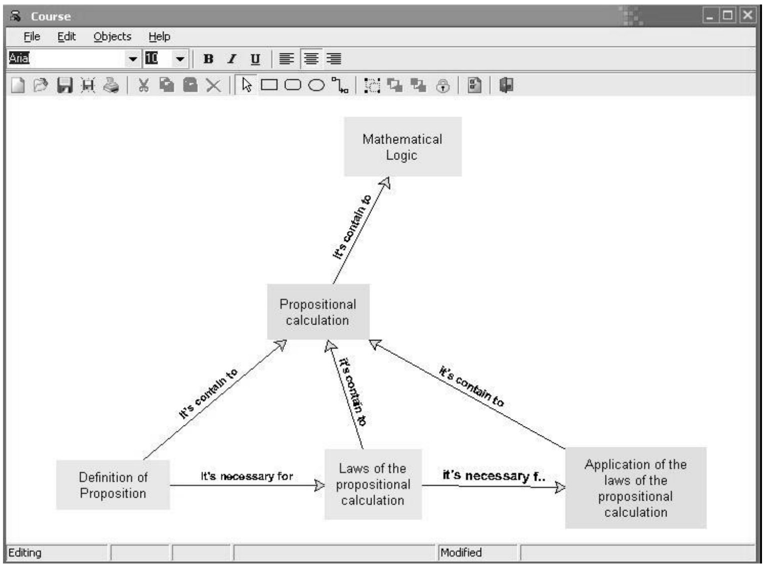

Fig. 4. Example of a CM to learn Propositional Calculation in MacBay Tool

Both architectures have a Psicopedagogical Assistant; the affective features are obtained through this assistant in order to achieve a connection between affectivity and cognition. There has traditionally existed an absolute separation between the cognitive and affective aspects when studying their influence in the learning process. So, investigators would centre their studies in the cognitive aspects rather than in the affective ones or vice versa. At present, there exists an increasing interest in studying the two components simultaneously.

The teaching-learning process is both cognitive and affective; in the development of the academic output it is necessary to take into account both the cognitive aspects and the affective ones. Learning is only possible if the chance to acquire the knowledge is given and the process is related to the capacities, knowledge, strategies, and the skills (cognitive component) of the student, but it is also necessary to have intention and the motivation to face this process (affective component). For this component the authors developed a system based on rules which team up with the learning process. It allows to be acquainted with the mood of the student and check their motivation as well as their understanding on a specific subject.

5 Using other techniques

In this section the authors describe some ideas of how to use other Artificial Intelligence techniques combined with CM to generate ITLS.

Petri Nets (PN) are powerful and versatile tools for modeling, simulating, analyzing and designing of complex workflow systems. This topic mainly discusses a hybrid approach using neural net and PN in the formal model of an ITS. Neural nets are used in different fields; classification is one of problems where they are commonly used [8]. PN are alternative tools for the study of non-deterministic, concurrent, parallel, asynchronous, distributed or stochastic systems. They can model systems in an easy and natural way. PN approach can be easily combined with other techniques and theories such as fuzzy theory, neural nets, etc. Since PN offer advantages to model systems and can interact with other techniques easily, it would be advantageous to model neural nets starting from PN models, which allow not only the design adjustment but also the initialization of the neural net weights [9]. Following the algorithm in [10], we can model a neural net starting from a PN with the application of weighty production rules in the algorithm [11].

On the other hand, FCM are an interconnected net of concepts, which represent variables, states, inputs and outputs relevant to a modeled domain of the system. The connection edges are directed to indicate the direction of causal relationships and each edge includes information on type of the relationship. The relationships can be positive (a promoting effect) or negative (an inhibitory effect). Each connection is represented by a weight which has been inferred through a method based on fuzzy rules that describe the influence of one concept to another [12].

The FCM development method is based on fuzzy rules that can be either proposed by human experts and/or derived by knowledge extraction methods. FCM consist of nodes (concepts) that illustrate the different aspects of the system's behavior [16]. These nodes interact with each other showing the dynamics of the model. FCM are developed by human experts who know the system and its behavior under different circumstances in such a way that the accumulated experience and knowledge are integrated in a causal relationship between components of the system.

In fact, FCM could be regarded as a combination of fuzzy logic and neural nets. In a graphical illustration FCM seems to be a signed directed graph with feedback, consisting of nodes and weighted arcs. Nodes of the graph stand for the concepts that are used to describe the behavior of the system and they are connected by signed and weighted arcs representing the causal relationships that exist between the concepts, all the values in the graph are fuzzy, so concepts take values in the range between [0,1] and the weights of the arcs are in the interval [-1,1]. Observing this graphical representation, it becomes clear which concept influences other concepts showing the interconnections between concepts and it allows updating in the construction of the graph, such as the adding or deleting of an interconnection or a concept. FCM are an improvement on cognitive maps with the addition of four significant characteristics:

- Edges between nodes can take on numeric values representing degrees of causality.
- FCM can model complex real-life scenarios, which are dynamic systems that evolve with time through feedback. The effect of a change in a concept node affects other nodes, which in turn can affect the node initiating the change and so on. In this respect, FCM may be viewed as recurrent artificial neural nets.
- The knowledge-base stored in an FCM can be augmented by combining a number of FCM using rules of fuzzy union.

Like artificial neural nets, FCM can be adaptive; the causal link strengths (initially specified by human experts) can be fine-tuned through learning. Between concepts, there are three possible types of causal relationships, expressing the type of influence from one concept to the others. The weights of the arcs between concept C_i and concept C_j could be positive ($W_{ij}>0$) which means that an increase in the value of concept C_i leads to the increase of the value of concept C_j, and a decrease in the value of concept C_i leads to the decrease of the value of concept C_j. Or there is negative causality ($W_{ij}<0$) which means that an increase in the value of concept C_i leads to the decrease of the value of concept C_j and vice versa [18].

Beyond the graphical representation of the FCM there is its mathematical model. It consists of an $1\times n$ state vector A which includes the values of the n concepts and an $n\times n$ weight matrix W which gathers the weights W_{ij} of the interconnections between the n concepts of the FCM. The matrix W has n rows and n columns where n equals the total number of distinct concepts of the FCM and the matrix diagonal is zero since it is assumed that no concept causes itself. The value of each one concept is influenced by the values of the connected concepts with the appropriate weights and by its previous value. So the value A_i for each concept C_i is calculated by the following rule:

$$A_i = f(\sum_{\substack{j=1 \\ j\neq i}}^{n} A_j W_{ji}) + A_i^{old} \quad (1)$$

where A_i is the activation level of concept C_i at time t+1, A_j is the activation level of concept C_j at time t, A_i^{old} is the activation level of concept C_i at time t, and W_{ji} is the weight of the interconnection between C_j and C_i, and f is a threshold function.

$$A_{new} = f(A_{old}W) + A^{old} \quad (2)$$

So the new state vector Anew is computed by multiplying the previous state vector Aold by the weight matrix W. The new vector shows the effect of the change in the value of one concept in the whole FCM. But, equation (2) includes also, the old value of each concept, and so the FCM possesses memory capabilities and there is a smooth change after each new cycling of the FCM [13].

6 Conclusions

Concept mapping is useful for knowledge management as a vehicle for externalizing "internal" expert knowledge, to allow that knowledge to be examined, refined, and reused. In this paper the authors propose two approaches combining the facilities of the CMs for the organization of the knowledge, one that uses the potentiality of the Case-Based Reasoning like inference tool and the other the Bayesian Nets as a nice way of representing knowledge and a powerful inference tool that allows systems to have the capacity to adapt the interaction to its user’s specific needs. The proposed ideas were implemented in the computational systems HESEI and MacBay, whose have been applied successful in the elaboration of ITLS by users that there are not necessarily expert in the Computer Science field. Other techniques also were used in the construction of the Student Model, combining the facilities of the CM for the organization of the knowledge and the potentiality of the PN that allows to the student to sail in an oriented way according to their knowledge and not in a free way as the CM are conceived. The suggested hybrid system combines PN and neural nets using the advantages of PN in order to overcome the neural net deficiencies concerning their original design and definition of their initial weighs.

Finally Fuzzy Cognitive Maps to support and model Intelligent Teaching-Learning Systems is proposed and analyzed. Fuzzy Cognitive Maps can very well be used as a knowledge mapping instrument and are widely used to simulate the behavior of complex systems. An indispensable characteristic of reactive environments is the need to continuously perceive and interpret its changes. This is because learning is evaluated based on the theory of qualitative reasoning in which the cause-effect actions are the base mechanisms of the reactive environments. The Fuzzy Cognitive Maps allow a faster control of the different states of the environment. With this work we propose some approach that could be considered in the construction of Intelligent Teaching-Learning Systems, where the adaptability of the interface for the user is the main goal, to offer the best option that produces a high performance for his behavior.

References

1. Cañas, A. (2004) Concept Maps and AI: an Unlikely Marriage? Simpósio Brasileiro de Informática na Educação.
2. Cañas, A. (2003) A Summary of Literature Pertaining to the Use of Concept Mapping Techniques and Technologies for Education and Performance Support. The Institute for Human and Machine Cognition.
3. Gutiérrez, I. (2003) Modelo para la Toma de Decisiones usando Razonamiento Basado en Casos en condiciones de Incertidumbre. Trabajo de Tesis en opción al grado científico de Doctor en Ciencias Técnicas. Cuba.
4. Castillo, E. (1997) Expert Systems and Probabilistic Network Models. Springer-Verlag.
5. Pearl, J. (1988) Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann.
6. Caballero, Y. (2006) Two new feature selection algorithms with Rough Sets Theory. IFIP AI 2006 pp. 209-216.
7. Medina, D. (2007) Putting Artificial Intelligence Techniques into a Concept Map to Build Educational Tools. Springer-Verlag, Berlin Heidelberg, 2007. IWINAC 2007, Part II, LNCS 4528 p. 617-627.
8. Hilera, J. (1995) Redes neuronales artificiales: Fundamentos, modelos y aplicaciones. Ed. Addison-Wesley Iberoamericana.
9. Yu, X. (1995) Dynamic Learning Rate Optimization of the Backpropagation Algorithm. IEEE Transaction on Neural Networks, Vol. 6, No. 3. pp. 669–677.
10. Liu, W. (2000) Re-investigating Dempster's Idea on Evidence Combination. Knowledge and Information Systems: An International Journal, Springer-Verlag, ISSN 0219-1377. Vol. 2, No.2. pp. 223–241.
11. Li, X. (2000) Dynamic Knowledge Inference and Learning under Adaptive Fuzzy Petri Net Framework. IEEE Transactions on Systems, Man, and Cybernetics – Part C: Applications and reviews, Vol. 30, No. 4.
12. Kosko, B. (1992) Neural Networks and Fuzzy Systems. Prentice-Hall.
13. Pelaez, C. (1995) Applying Fuzzy Cognitive Maps Knowledge-Representation to Failure Modes Effects Analysis. Annual Reliability and Maintainability Symposium. pp. 450–455.