Computers & Education 79 (2014) 126–137

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Computational organization of didactic contents for personalized virtual learning environments

Francisco Ramos de Melo [a,*], Edna Lucia Flôres [b], Sirlon Diniz de Carvalho [b,c], Ricardo Antonio Gonçalves de Teixeira [b,d], Luis Fernando Batista Loja [b,c], Renato de Sousa Gomide [b,c]

[a] Universidade Estadual de Goiás, Unidade de Ciências Exatas e Tecnológicas, Anápolis, GO, Brazil
[b] Universidade Federal de Uberlâdia, Faculdade de Engenharia Elétrica, Uberlândia, MG, Brazil
[c] Instituto Federal de Goiás, Goiânia/Luziania, GO, Brazil
[d] Universidade Federal de Goiás, Faculdade de Engenharia Elétrica e Computação/Faculdade de Educação, Goiânia, GO, Brazil

A R T I C L E I N F O

Article history:
Received 12 November 2013
Received in revised form
24 July 2014
Accepted 24 July 2014
Available online 7 August 2014

Keywords:
Personalized learning content
Proximal learning patterns
Intelligent tutoring systems
Artificial neural networks
Personalized teaching

A B S T R A C T

This paper presents an organization model for personalized didactic contents used in individual study environments. For many students the availability of contents in a general form might not be effective. A multilevel structure of concepts is proposed to provide different presentation combinations of the same content. Our work shows that it is possible to personalize the didactic content in order to encourage students, by using proximal learning patterns. These patterns are obtained from the analysis of the actions of students with positive results in the individual content organization. The system uses artificial intelligence techniques to reactively organize and personalize content. Personalization is made possible by means of an artificial neural network that classifies the student's profile and assigns it a proximal learning pattern. Expert rules are used to mediate and adjust the contents reactively. Experimental results indicate that the approach is efficient and provides the student a better use of the content with adaptive and reactive personalized presentation.

1. Introduction

Since the dawn of mankind, acquiring and transmitting knowledge has been the element that distinguishes and instigates human development. Knowledge has been the instrument used to promote and ensure the survival of humanity, personal and social development and national sovereignty The oldest and most widely used method to transmit knowledge is still based on face-to-face interaction. In this mode, both teacher and student are in the same space and time. Along the presentation of contents, the teacher interacts with the student, mediating knowledge according to the student's development. The process is driven and controlled by the teacher in a dynamic and immediate manner. However, economic constraints of space and time have restricted the scope and availability of education for many students (Horton, 2000; Jonassen, 2001).

A recent teaching mode, called Distance Learning (DL), gave rise to a greater availability and scope of education, both socially and demographically. In this mode the interaction between teacher and student is asynchronous, not occurring in the same space at the same time. The teacher provides content that can be studied by students in a different time and place. The student takes charge of his or her own learning. When difficulties arise, the process may be blocked until the teacher's intervention takes place. The student's profile, discipline and persistence in the presence of difficulties will be crucial to sustain the learning process (Horton, 2000; Phelan, Mendoza-Diaz, & Mathews, 2002).

* Corresponding author. Rua Bernardo Sayão, 175, Vila Nossa Senhora da Abadia, Anápolis, Goiás, Brazil. Tel.: +55 (62) 9963 7437.
E-mail addresses: chicorm@yahoo.com.br, francisco.melo@ueg.br (F. Ramos de Melo), edna@ufu.br (E.L. Flôres), sirlondiniz@yahoo.com.br (S. Diniz de Carvalho), professorricardoteixeira@gmail.com (R.A. Gonçalves de Teixeira), luizloja@gmail.com (L.F. Batista Loja), renato.s.gomide@gmail.com (R. de Sousa Gomide).

http://dx.doi.org/10.1016/j.compedu.2014.07.012
0360-1315/

The organization and presentation of content have great relevance in the process of imparting knowledge. Its relevance increases in self-study learning environments. Several technologies have been developed to support Distance Learning. The introduction of electronic computers in the process revealed important solutions to aid the student's learning. Much of the developed computer applications consist of a sophisticated electronic version of books. The development of computer technologies that employ artificial intelligence (AI) provided important solutions for distance learning environments. Intelligent tutoring systems (ITS) are among the solutions that employ AI for educational applications. ITS was developed with the purpose of establishing links between the learning object and student's knowledge. A major effort in developing an ITS is the search for a mechanism that can compensate for the teacher's absence while the content is delivered (Horton, 2000; Viccari, 2005).

ITS research seeks to develop techniques for personalizing the presentation of content that is reactive to the student. The development of ITS usually utilizes symbolic AI techniques. In these techniques, knowledge processes are abstracted and organized into rules or equivalent structures that allow the system to automate the operation (Dastbaz, Mustafa, & Stoneham, 2006; Duque & Jiménez, 2006; González & Ruggiero, 2009; Vicari, 2005).

An alternative research line sought to simplify the knowledge representation using connectionist AI techniques (Carvalho, 2004; Martins, Melo, Meireles, Nalini, 2004; Melo, Flores, & Carvalho, 2011). In connectionist AI, knowledge is abstracted, organized and manipulated in the form of patterns.

This paper presents a mechanism for reactive personalization of didactic contents. The proposal is a way to restructure and formalize a given content in different presentation levels of the same subject. Thus, it is possible to combine these levels to produce a different content for each student. By using a mathematical formulation with partial derivatives, it is possible to represent the teacher/student interaction process, i.e., the didactic path of the contents presentation. The proposed method identifies the student's profile and establishes a relationship with a proximal learning pattern. From these patterns, expert rules complement the personalization, providing reactivity for the different moments of contents presentation. The model discusses and formally organizes the elements involved in the system.

This paper is structured as follows: Section 2 present the rationale for the process of contents personalization. The proposed method, the description of the experiment and the results are shown in Sections 3 through 5, respectively. Finally, conclusions are discussed in Section 6.

2. Concepts, environments and patterns

2.1. Transmission of knowledge

The process of transmitting knowledge can be represented by the combined actions of teacher and student, with the purpose of assimilating knowledge. Three elements take part in this representation: the teacher, the student and the delivered content. The student is the interacting part to which all the effort of the process is directed, in order to develop his or her skills. The teacher is the key agent of the process and is responsible for the direction and organization of the means that enable the student to assimilate knowledge from the content. The content is related to the knowledge that is conveyed (Fontenla, Caeiro, & Llamas, 2010; Patten, Chao, & Reigeluth, 1986; Phelan et al., 2002).

The success of a knowledge transmission process should take into account the organization and presentation of content (Patten et al., 1986). The content is an organization of ideas involved in the inception of the knowledge to be transmitted. Ideas are knowledge units called concepts. A didactic content can be understood as an element structured in such a way that the ideas constitute a whole.

Typically, for designing, selecting and implementing programs and procedures for teaching numerous classes, teachers shape the content in a single and general format for all students. This format assumes a student of average ability and skill. The content selected by the teacher is presented in the same way for all students. The learning process happens with the interaction of teacher, student and content. This interaction involves the mediation by the teacher in face of the student's reactivity to the content that is being presented. The reactivity is the reaction of the student to the content that is presented (good understanding, doubts, etc.). The perception of the student's reactivity enables the teacher to interact in the process, taking the necessary action in the study sequence. This action performed by the teacher while interacting with students and content is called mediation, and favors the development of knowledge (Fontenla et al., 2010; Jonassen, 2001).

In a face-to-face teaching environment, the teacher interacts with students during the presentation of the content, identifies their difficulties and offers options to conduct the learning process. This interaction between teacher and student occurs throughout the contents development process. In this kind of educational environment, the single format can deliver effective results, because the teacher interacts with students when difficulties arise (Horton, 2000).

In a distance learning environment, interactivity is a complicating factor. Considering the lack of mediation by an educator, single-format content may not meet individual learning situations. During the presentation of content, the teacher's pedagogic intervention towards the student's difficulty may take longer to be administered. General pedagogic interventions are made impossible by the unknown plurality of students and all the typical situations of a distance learning environment. Due to its asynchronous aspect, DL does not enable immediate interaction between instructor and student when problems arise. Thus, the student's learning may be jeopardized by the delay of the teacher's intervention (Fontenla et al., 2010; Jonassen, 2001).

2.2. Computers in education

The need to provide cost-effective education to increasingly larger and diverse populations brought forth the development of new instructional alternatives. The interaction between the different elements involved in a non-face-to-face education process has established disciplinary ties between the Exact Sciences and the Humanities (Horton, 2000). The introduction of electronic computers enabled the development of many technologies to assist in educational processes.

Skinner (1968) points out the importance and rationale for organizing the teaching sequence in order to transmit knowledge, which requires patterns for the systematic process. Skinner's work comprises definitions for format and presentation of the contents, which enable the mediation of knowledge using electronic devices.

The use of electronic computers in education began in the '50s, with the introduction of tutoring systems. The first applications of these systems are considered as mere “electronic page turners” due to its typically static presentation of contents. Several educational applications that have come to use tutoring system technologies were classified as “Computer-Assisted Instruction” (CAI) (Rosenberg, 2001).

Artificial intelligence (AI) is a computational technology that offers interaction solutions for distance learning contexts. Several AI applications have been developed as a means of establishing links between the learning object and the student's knowledge. A tutoring system that uses AI is considered an ITS or “Intelligent Computer-Assisted Instruction” (ICAI). A major effort in developing an ITS is the search for a mechanism that can compensate for the absence of a teacher that conducts the didactic contents. For developing personalized tutorials, AI techniques have been introduced to provide a system that is reactive and adaptive to the student during the tutoring process (Dastbaz et al., 2006; Duque & Jiménez, 2006; Fontenla et al., 2010; González & Ruggiero, 2009).

In order to contextualize the proposal of tutoring system that uses AI based on artificial neural networks, it is important to present the main structures used at the moment. Usually, an introduction marks the beginning of the lesson and, as a final step, a summary is presented for revision of the concepts, following by test or other activity to measure the acquired knowledge. In classical tutorial, users access the content in basic, intermediary and advanced levels progressively. In the tutorial focused in activities, another activity with some information or additional motivations precedes the accomplishment of the goal activity. In the tutorial customized by the apprentice, between the introduction and the summary, there are cycles of pages of options (navigation) and content pages. The page of options presents a list of alternatives for the apprentice or a test in the sense of defining the next step. In the progress by knowledge tutorial, the apprentice can omit contents dominated already, being submitted to tests of progressive difficulty to determine the entrance point in the sequence of contents. In exploratory tutorial, the initial page of exploration has access links to documents, databases or other information sources. In lesson generating tutorial, the result of the test defines the personalized sequence of topics to be exposed the apprentice (Horton, 2000).

The development of a computer system to assist in the organization of didactic presentation of the contents offered by the teacher, respecting the student's individuality, is justified by the need to create mechanisms that allow for the democratization of knowledge. This democratization is necessary precisely to promote the inclusion of persons who, for adverse reasons, would otherwise not have access to a human tutor that could mediate the knowledge internalization process. For such, it must be taken in consideration that each student has a particular profile of intelligence. It is important that Distance Learning should be carried out so that the transmitted knowledge has significance to the student.

2.3. Artificial neural networks

The development of intelligent systems can be accomplished with the use of AI techniques, either classic or connectionist. In classic AI, knowledge is abstracted, organized and represented by experts. Experts are individuals with knowledge of a problem and of the actions and decisions that must be resolved within the problem domain. The expert establishes facts, rules and procedures to be considered in order to model knowledge in classic AI systems (Norvig & Russel, 2004). In connectionist AI, knowledge is established and manipulated by the structure and relationship of the elements that define the problem. The expert defines the features that should be considered in the problem domain. Knowledge may be established by techniques that arrange the relationship of the elements observed in the problem (Haykin, 2000).

The process of knowledge acquisition is the part of the development of intelligent systems that defines the knowledge manipulation structures. In classic AI this process of knowledge acquisition is usually performed by manual techniques such as interviews. The process of acquiring knowledge is organized and conducted by a knowledge engineer, where the expert's experience and knowledge are very important in defining and driving the process. For an efficient modeling, it is also important that the expert clearly and correctly stipulates the knowledge. The difficulty of such manual techniques lies in the quality of the communication between the engineer and the expert. In classic AI systems, the modeling of knowledge depends on the quality of the communication. In connectionist AI, the modeling of the knowledge acquisition and representation are simplified by automatic or semi-automatic techniques. The expert determines which features to consider from classes of knowledge. The organization and knowledge representation can be performed automatically using some machine learning technique. The quality of that technique at organizing and structuring will define the knowledge modeling quality in connectionist AI systems (Bergman, 1991).

Artificial neural networks (ANN) are processing structures that mimic the processing activity of the human brain. The main feature of an ANN is the ability to determine the relationship of the input data, and then perform pattern classification in the network's output (Haykin, 2000).

The Multi-Layer Perceptron (MLP) neural network is a model that performs pattern classification with high capacity and flexibility. The pattern classification undertaken by MLP networks results from a function complex which considers the relevance and combination of the inputs (Haykin, 2000).

The MLP network uses the supervised learning paradigm. In this paradigm, a set of examples of the pattern to be classified is presented. Each set consists of a group of features (used as inputs) and the corresponding pattern, which represents the expected output. After the presentation of these features to the network input, the output that the network calculates is compared to the expected output. If there is a difference, the weights of the network connections are readjusted. The adjustment of the weights is performed iteratively until the network is considered trained. The backpropagation algorithm is the most widely used method for training MLP networks (Haykin, 2000).

After the training stage, the ANN has the ability to identify new patterns, different from those used for training, thus establishing a new classification intermediate to the known patterns (interpolation) (Haykin, 2000). Hence, the model is especially useful in situations where other pattern classification methods do not perform well either because of the problem complexity or due to the difficulty at establishing the relationship of the parameters of the selected method (Duda, Hart, & Stork, 2000).

2.4. Behavior patterns

In analyzing student profiles, it can be easily observed that there is not a unique profile model. In fact, there are a wide variety of personality profiles with various behavior patterns.

Psychology has developed several tools to classify the patterns of the human mind. Studies in this area establish, using psychological tests, the relationship between behavior patterns and the various dimensions that characterize the pattern (Duda et al., 2000; Gregory, 2007).

According to Jung (1971), the psychological type is an explanation of the human personality. He noted that human behavior is not random, i.e. human actions are not the result of chance. Instead, Jung noted that the behavior follows patterns developed from the structure of the human mind. Thus, Jung developed a theory of psychological types based on four functions and two attitudes. These functions are Feeling, Reasoning, Intuition and Sensitivity. And the attitudes are Extraversion and Introversion.

3. System structure

A personalized content can be considered a different sequence of concepts. This difference lies in the fact that the sequence is organized in such a way that matches the student's profile as much as possible. Contents personalization allows the student to benefit more from the presented educational topic.

Since the conventional contents design is directed to a particular student profile, it is usually presented in the same way for any student, regardless of any differences. Consequently, the possibility of adjusting the content for other student profiles is almost nil. It is therefore necessary to create a method in which the same content may be treated in different ways.

The personalization process should consider the identification of the student profile. In order to benefit the most from the presented educational topic, an association may be established between the student profile and a pattern of content organization.

3.1. Multilevel didactic content

A given conventional content ***(C)*** can be mathematically defined by a set consisting of a sequence of concepts. The concepts are organized in a logical sequence conductive to knowledge. The organization of this sequence can be achieved in different ways, ranging from the classic organization to personalized organizations. The presentation of this sequence is usually preceded by an introduction followed by an initial test and, at the end, an overview of the discussed concepts is presented and then a final test is performed.

The most common organization is the classic, presented in Fig. 1, where the contents sequence is developed and presented in the same way for any student. Personalized organization is a way of customizing the content because it allows for the presentation of parts of the content as needed by the student.

The sequence of the conventional didactic content is inappropriate for the personalization process. However, if each concept is rewritten with different levels of presentation, it is possible to compose the same content using different combinations of these levels. The restructuring of the concept in different presentation sequences, presented in Fig. 2, is called multilevel concept. The structure of this concept is similar to a teacher's mediation of knowledge, i.e. after the presentation of the concept the teacher seeks a way to define the organization of knowledge according to the student's reaction.

The different combinations of multilevel concept levels produce the personalized content presentation for each student profile. In addition, it allows mediating interventions for different moments of study. Thus, it is possible that for the same content, each multilevel concept may be presented differently. The presentation of this concept can be performed using one or more levels as needed to establish the direction of knowledge.

Conventional concepts and multilevel concepts differ in that the latter is more complex and requires a greater effort to be developed. In multilevel concept, the number of presentation levels must be initially defined, as well as at which level the presentation starts. The initial level is the reference to develop other levels. In all concepts, identical levels must have the same presentation pattern.

3.2. Student profile

The goal of personalization is to allow the student to benefit the most from the presented topic. After defining the multilevel content, in order to proceed with the personalization process, it is necessary to define a way to identify the student profile and how to determine the most appropriate content organization for that profile.

In order to identify the student profile, the proposed method provides classification patterns. These patterns are: personal, ability and pre-knowledge of content.

The personal pattern is defined by characteristics obtained from a psychological test. An initial test containing questions about the content being studied defines the pattern for pre-knowledge of content. The ability pattern is defined from a questionnaire on sociocultural conditions of the student and his or her familiarity with technology.

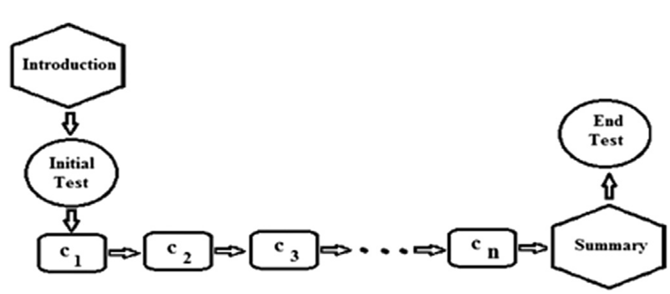

Fig. 1. Organization of the classic tutorial.

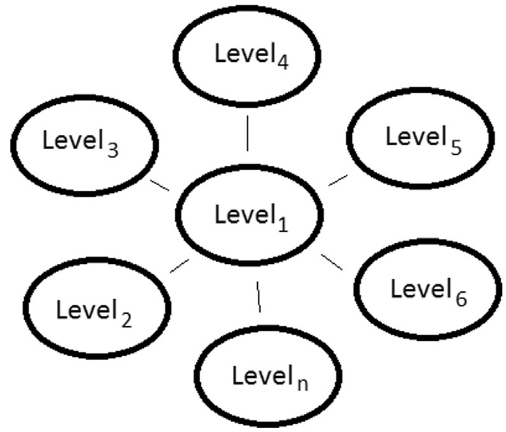

Fig. 2. Organization of the multilevel content.

Psychological tests are part of the psychology field that attempt to associate a person's everyday items to dimensions that express a behavioral characteristic. Typically, these dimensions are represented by dichotomies. Dichotomies are two attributes of a certain dimension positioned at opposite ends. The response to a psychological test identifies the trend in each considered dimension. Fig. 3 shows an example of the dimension "Judging". In this example, the dichotomy is formed by the characteristics of "Feeling" represented by the value 1 and "Reasoning" represented by the value −1. As seen in this figure, dimension "Judging" tends to the value 1, i.e., to the attribute "Feeling".

The scope of a dimension is defined by a set of items in a psychological test. Hence, a given dimension can be represented by any sum of the answers of each item to the regarded dimension. The set of dimensions considered in the psychological test determine the personal pattern *(Ψ)*.

The ability pattern *(A)* is a sociocultural evaluation the student's condition. This evaluation may include items such as age, gender, handling technology, etc. This pattern complements the personal pattern *(Ψ)* in order to establish elements closer to the student's reality and his or her interaction with the content. The ability pattern *(A)* is a region defined by the sum of each item of the ability test.

The pattern of pre-knowledge *(K)* of the content is a pre-test performed before presenting the content. The purpose of this pre-test is to verify the student's level of knowledge about the regarded content. For each concept of the content a question is formulated in the pre-test. Pattern *(K)* is a score obtained by summing up the answers to each question in the pre-test.

The student profile *(Ω)* is a multidimensional region resulting from the interaction of regions defined by the following patterns: psychological *(Ψ)*, ability *(A)* and pre-knowledge *(K)*.

Considering the psychological pattern *(Ψ)*, each student has a unique profile. However, if we consider only this parameter, the wide variety of psychological patterns *(Ψ)* results in a great number of student profiles. Nonetheless, according to the purpose, a psychological pattern *(Ψ)* can be compared to another pattern *(Ψ)* similar to the first one. Using other patterns as a complement to the psychological pattern *(Ψ)* can improve the scope in the student profile's classification region. As this region expands, the probability of common areas increases. In this sense, the ability pattern *(A)* and pre-knowledge pattern *(K)* offer a greater scope in the students profile's classification region. As that scope increases it becomes possible to classify by approximation a greater number of student profiles in the same region.

3.3. *Proximal learning pattern*

In multilevel content, a personalization path is the definition of how much the student prefers each of the levels of this content. The learning pattern *(L)* can be defined as the group of preference for each level on multilevel content.

Once the student profile *(Ω)* and learning pattern *(L)* are defined, it is necessary to establish an association between them. In this paper, the proposed organization of personalized didactic content consists in classifying the student profile *(Ω)* and associate it with the learning pattern *(L)*. From the student profile *(Ω)* it is possible to find by approximation the learning pattern *(L)* which in turn becomes the proximal learning pattern (L_Ω).

The learning pattern *(L)* is determined directly from the observation of a student. The proximal learning pattern (L_Ω) is obtained in the same way as L, but L_Ω is estimated by approximation considering the student profile *(Ω)*.

If learning pattern *(L)* led to a greater effectiveness for one student with profile Ω, consequently the proximal learning pattern (L_Ω) can result in greater effectiveness for another student with similar profile Ω.

The proposed method uses a Multi-Layer Perceptron (MLP) neural network to define the student profile *(Ω)* and estimate the proximal learning pattern (L_Ω), as illustrated in Fig. 4.The ANN is trained with data selected from the observation of students that had positive results

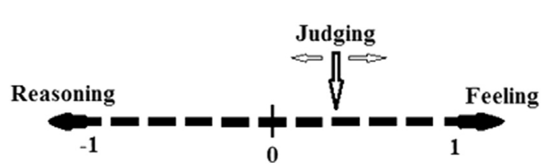

Fig. 3. Example of a dichotomy dimension.

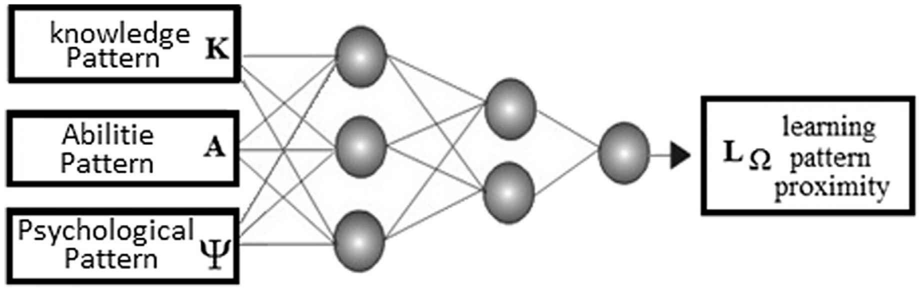

Fig. 4. Structure of the ANN to define the $L_Ω$ proximal learning pattern.

with their ability to organize individual study. At the end of the training stage, the ANN will be able to classify the student into a pattern $Ω$ and indicate the proximal learning pattern ($L_Ω$).

Fig. 5 shows some examples of proximal learning patterns ($L_Ω$) for different students. This figure exemplifies the patterns ($L_Ω$) for a multilevel content *(MC)* with 5 levels. The probabilities *P(a)*, *P(b)*, *P(c)*, *P(d)* and *P(e)* for each level define the vector for pattern ($L_Ω$).

3.4. Personalizing multilevel didactic content

The proximal learning pattern ($L_Ω$) is not sufficient to define the whole sequence of levels for a personalized content presentation. This pattern is general and does not support reactivity to contents presentation. Moreover, it is not possible to represent with a single function the whole sequence of personalized content due to the student's reactivity at different moments of the presentation of each level in the sequence.

One of the reasons why the proximal learning pattern ($L_Ω$) does not suffice for the personalization is about how this pattern is treated. For example, when it is decided, within the presentation, to show only the student's most likely preferred level, the content will be the same for all students with ($L_Ω$) maximized at that level. This happens for the student regardless of the percentage probability of the maximized level. This situation is similar to the presentation of classic content, but at the maximized level. This format cannot be considered a personalization of contents presentation.

One possibility to change the multilevel content *(CM)* into a personalized one can be a random drawing from the probabilities of each level. This can result in a different content for each student, but can present inconsistencies in the presentation. An example of inconsistency is when the system indicates the study of a concept at a higher difficulty level when the student failed the test on the lower difficulty level.

The personalized content presentation must be reactive, and must also be related to a proximal learning pattern ($L_Ω$). Reactivity should occur at various moments along the presentation in order to correct potential problems.

The reaction to the content presentation varies according to the individual profile of each student. According to this reaction, the teacher can modify the way of presenting the subject.

Some points should be established to represent the decision process for the correction of the didactic path in the personalized presentation using multilevel contents *(MC)*.

While navigating through conventional contents *(C)*, the student tends to follow the didactic path of the originally defined sequence. This sequence shows only one level of each concept. This is because this concept is the only level available. The conventional content sequence is shown in Fig. 6, which shows that, for all the concepts, the path passes through a single point which is the level $N_{0.}$

In multilevel content *(MC)*, the didactic path tends to be similar to the conventional content path. However, according to the student's reaction after the presentation of any given level, the teacher can apply a variation along the path. This variation takes the proximal learning pattern ($L_Ω$) in consideration and may lead the student to another concept level, as seen in Fig. 7.

In the proposed system, the multilevel concept was divided in five levels: three main levels and two auxiliary levels. The main levels are medium, easy and advanced. Auxiliary levels are Examples and Frequently Asked Questions (FAQs). The medium level is the reference for

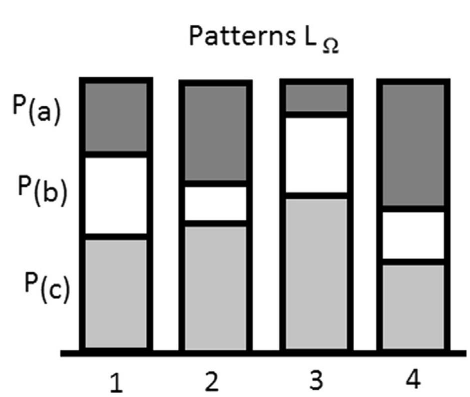

Fig. 5. Examples of proximal learning patterns ($L_Ω$).

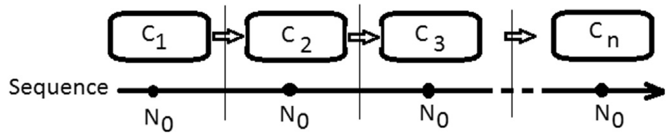

Fig. 6. Sequence of the didactic path for conventional content.

other levels and, for each concept, it is the first level to be presented to the student. The proposed system considers that the main and auxiliary levels are positioned in the *xy* plane as shown in Fig. 8.

The student's didactic path tends to follow a center position. However, according to the student's reaction upon presentation of a certain level, the teacher can change the path toward another level. The next level to be presented to the student is a region located between the student's pattern (L_Ω) and a region defined by the teacher for that situation.

In order to correct local situations of the student's reactivity, a study was conducted with experts (teachers) about their intervention upon each possible situation at the end of a content presentation. From that study, local intervention rules have been created, called expert rules.

In the system proposed in this work, personalization is implemented by combining the proximal learning pattern (L_Ω) with the expert rules. The selection of the rule that is appropriate to the student's reactivity is defined by presenting the student with a multiple-choice exercise at the end of the presentation of each level. This exercise is the retention test of the level and it is the system's sensor used to identify the student's reaction.

The insertion of the expert rules together with the ANN defines a hybrid system. The expert rules in the proposed system have the same structure as the proximal pattern (probabilities for each level). However, each rule is specific to each situation in the retention test performed after the presentation of content. Each one of this situations performs the local pattern. Fig. 9 shows an example of a set of expert rules for any given level. In this figure, each column represents the set of probabilities of level presentation according to the evaluation by the retention test.

In order to personalize the content, both the global pattern and the local pattern are necessary. The global pattern (L_Ω) is established only once before the presentation starts. The local pattern (expert rules) is redefined after the presentation of each concept level. The next step in the organization of the content is to define the next concept level to be presented.

The organization adjustment of the content sequence can be carried out by combining the global pattern (student) with the local pattern (teacher). In the proposed system, this combination is obtained by the product of the probabilities of local and global patterns, as illustrated in Fig. 10. The result of this product is a new pattern, called the probabilistic decision pattern. In this pattern, probabilities can be increased or reduced in proportion to their relevance in each of the patterns. In this method it is important to note that the attenuations can suppress inconsistencies in local situations.

Each level in the probabilistic decision pattern is represented by probabilities according to the student profile adjusted by expert rules in relation to the student's local performance in the retention test. With this pattern, the next content level to be presented can be obtained by using the Monte Carlo technique. This technique chooses a level by means of a probabilistic drawing. Greater probabilities are more likely to be chosen, but small probabilities can also be selected.

3.5. System structure for reactive personalization

In order to personalize didactic content, it is possible to structure a system considering multilevel content *(MC)*, the student profile *(Ω)*, the proximal learning pattern (L_Ω) and multilevel content personalization.

The proposed system comprises introduction, initial tests, presentation sequence and final test.

The introduction displays general information about the system and the topic being presented. After the introduction, tests are performed in order to assess profile, ability and pre-knowledge. The test results become input data to the ANN, which will classify the student profile and will yield, as output data, the proximal learning pattern (L_Ω). This pattern is stored in the system memory as a global reference for content presentation.

The system starts the sequence of contents presentation from the first concept all the way through the final concept, as defined by the teacher. The presentation of each concept always starts at the medium level. After all content has been presented, a final test is performed. Fig. 11 shows the structure of the proposed system.

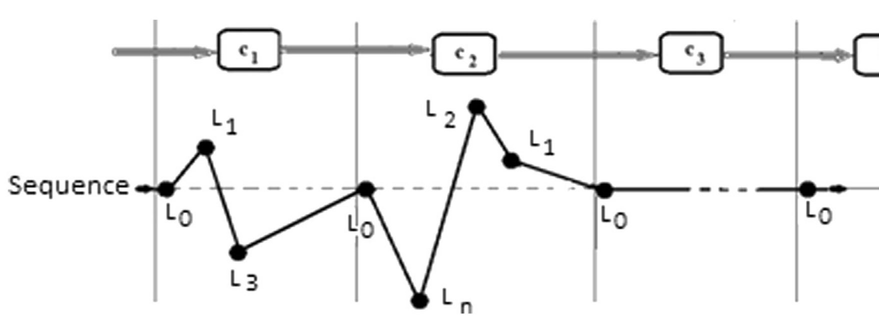

Fig. 7. Sequence of didactic path in Multilevel Content.

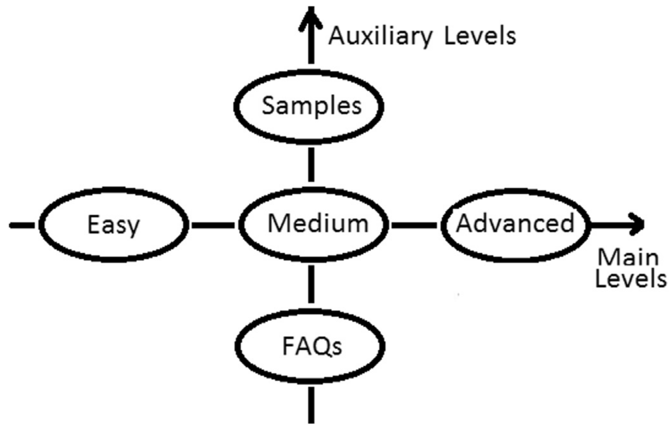

Fig. 8. Layout of a multilevel concept.

Content presentation is organized in a combination of two sequences: general and localized.

The general sequence is usually organized by the teacher, who defines the order in which the concepts are presented. This sequence presents the concepts in a logical order, structuring the content from the first to the last concept. The general sequence has a structure similar to the classic tutorial organization.

The localized sequence is organized by the proposed system, which defines which levels will be presented and their presentation order in accordance with the general sequence. After content presentation in any given level, the student may react in different ways. According to the reactivity, a level is selected to be presented. The localized sequence corresponds to the teacher's didactic interventions according to the student's reactivity.

In the proposed system, the localized sequence organization begins always at the medium level. After the presentation of each concept level, the student takes a retention test at that level. After evaluating the response of the student, the system selects, from the set of expert rules, the local rule that is appropriate to the situation. Combining the local rule and the global pattern (L_{Ω}), the system determines the next step of content presentation by means of a probabilistic drawing. This process uses the aforementioned Monte Carlo technique. The definition of the next step is repeated until the conclusion of the content. Consequently, the system organizes content in a way that it is personalized and reactive to the student.

4. Experiment

To conduct the experiment, three systems were developed with similar software interfaces: a free navigation system, a random navigation system and an intelligent navigation system. The difference between these systems is the decision mechanism for choosing in which level the organization of content is to be presented. The intelligent navigation system utilizes the modeling techniques proposed in this paper. Fig. 12 illustrates the organization of experiment used in this work.

In free navigation, the student decides the next level in the content presentation sequence. The goal of using this navigation in the experiment is to collect, analyze and select data for the training stage of the neural network that the intelligent system uses; the collected data is also saved for comparison with data from the random and intelligent navigation systems. After the data is collected, the navigation performance of each student is analyzed. The best navigations are selected according to criteria established by teachers participating in the experiment. Some of the criteria considered are: score, gain (improvement), consistency, etc.

An example of inconsistency in the organization of content occurs when the student, after finding difficulty at the easy level, moves to a more advanced level.

The gain measures the percentage of student improvement within the range of his or her score and the maximum possible score of the test. Equation (1) shows the calculation of gain.

$$Gain = \left(\frac{final\_score - initial\_score}{max\ score - initial\_score}\right) * 100 \tag{1}$$

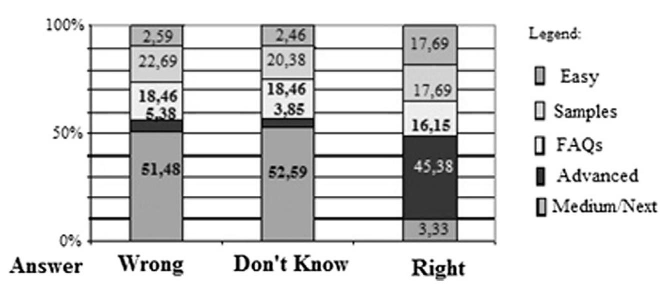

Fig. 9. Example of expert rules.

Fig. 10. Operation of the probabilistic decision pattern.

In this experiment, only valid samples have been collected for analysis. A valid sample is one in which the student has achieved all navigation stages of each system.

The first system to be sampled was the free navigation system. The students involved were either in high school or in early college years. This group of students was selected in attempt to increase the probability of a sample with a better profile of individual study organization. The total number of valid samples was 148 students. From these 148 samples, only 60 met the minimum criteria. The following criteria have been established for this experiment: final score greater than or equal to 6, minimum of 10% improvement and no inconsistencies.

In the random navigation system, the next level of concept to be presented is chosen randomly by the software. The general concept sequence follows the presentation order defined by the teacher. However, due to the random nature of how the next level is chosen, any concept level may be presented in any order. Besides the randomness, the content level presentation sequence is organized by chance, without any rules or logic. The purpose of sampling this type of navigation is to draw a comparison with the intelligent navigation system. With this comparison, it is possible to demonstrate that the effective actions of the intelligent system (alternative hypothesis − $H_{0)}$ are not due to chance (null hypothesis − $H_{n.}$). A total of 31 random samples were collected with the random navigation system.

In intelligent navigation, the system decides the next content to be presented using a series of techniques described in the system proposed in this paper. The purpose is to validate the applicability of these techniques in the personalized and reactive organization of the contents. In this navigation 31 valid samples were collected.

The expert rules have been defined based on data collected from teaching experts. The data were collected from 27 teachers with experience in several areas of higher education. The data were obtained from a form of local situations. In this form, a local situation is considered to be each of the possible answers after the retention test for the content level presented. For each local situation, the teacher defines an action to present another content level. This action is a set of probabilities for directing another level according to the student's reactivity. Table 1 contains a portion of the teacher's actions form according to the student's performance after the retention test at the FAQs contents level.

After data from the teaching experts were collected, the expert rules of the proposed system were defined by averaging the teachers' directions for each local situation. Consequently, a specific rule was defined for each of these situations. After identifying the pattern of the local situation, the specific rule is triggered by the system indicating the probability of presentation for each contents level, i.e., the possibilities of didactic intervention.

The data collected from the free navigation system were used for training the neural network. The MLP model was chosen for its ability to classify new patterns, different from those used in training. The chosen network structure was the one that yielded the smallest error between the training set and the validation set. After training, the ANN became the system's component that determines proximal learning pattern (L_{Ω}) for the intelligent navigation system.

In order to collect sample data for the experiment described in this work, a multilevel content with the theme "Introduction to Computers" was created to be presented in textual form. The general sequence of this content was organized into 15 concepts. Because a

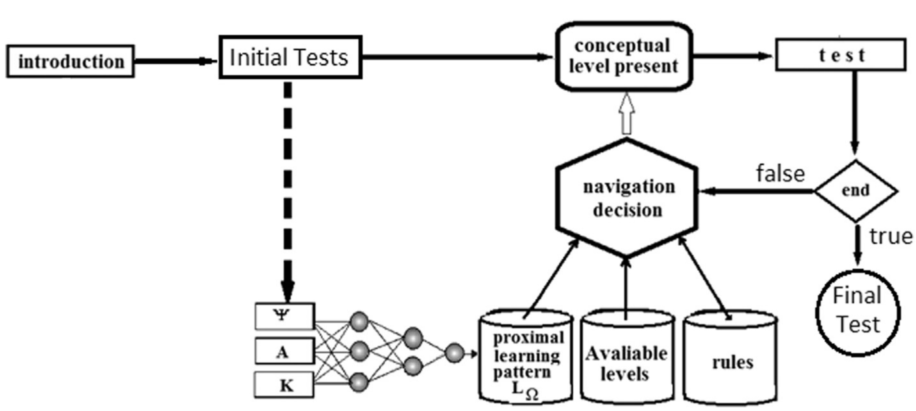

Fig. 11. Structure of the proposed system for personalization of didactic content.

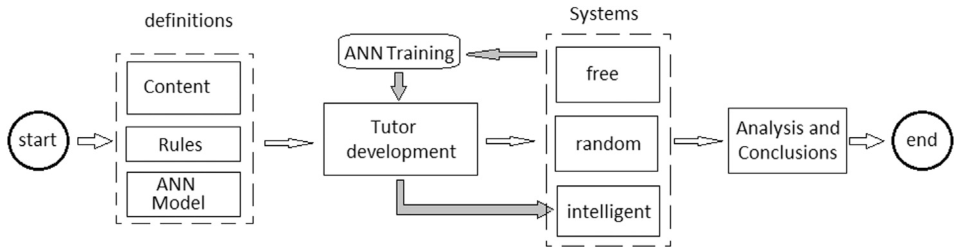

Fig. 12. Experiment.

five-level structure was used in this experiment, 75 texts were developed, i.e., 15 concepts in five levels. Fig. 13 illustrates the multilevel structure used in this work.

5. Results

In order to analyze the experiment, three distinct data samples have been collected, namely from the free navigation system, the random navigation system and the intelligent navigation system. The collected data were analyzed in a comparative study of the samples from each navigation system. This study compared characteristics in order to validate the use of the techniques proposed in this paper.

The first aspect of the comparative study was the analysis of the behavior and performance of the initial and final scores of students participating in the experiment. The range of the scores is zero to ten; zero being the lowest score and ten the highest. Table 2 shows the averages of these scores for each system.

As seen in Table 2, the mean initial scores are similar for the three types of navigation. The average of the initial scores was 4.16. This value shows that the participants' knowledge about the subject was merely sufficient. Among these three navigation systems, the mean initial score was highest for the free navigation system, and that was due to the pre-selection of the sample. The purpose of this pre-selection was to obtain the largest possible number of examples of more efficient content organization.

Apart from the initial scores, it can be seen in Table 2 that the average of the final scores was 6.6. This value indicates that there was an improvement in the final performance of participating students. From the three samples, the highest mean final score was observed for the intelligent navigation system, followed by free navigation and random navigation systems.

At this point, it should be emphasized that the mean final score for the intelligent navigation system was higher than the random navigation score. This indicates that the systematization of the techniques used in the proposed system is responsible for organizing the content in a personalized way so that the student benefits the most from the studied topic. This organization differs from random navigation, where the presentation sequence of each level has its order defined completely by chance.

In order to compare the performance of the three systems considered in the experiment, the results of some navigation attributes were observed. These attributes are: total time needed to complete the study of the content, number of levels used in the sequence organization, number of inconsistencies, final score and gain. Table 3 shows the comparison of average performance for the three navigation systems.

It can be seen in Table 3 that the intelligent navigation, in comparison to the other navigation systems, produced better results in terms of time, number of visited levels, final score, and gain.

The global pattern (L_Ω) defined by the ANN, and the local pattern defined by the expert rules are the key elements in the content personalization. The global pattern is defined only once and remains unaltered until the end of the presentation. The local pattern is variable and is redefined after each control point.

6. Conclusions

This paper presented a formal basis to represent the teacher/student interaction process, in other words, the didactic path for content presentation. In virtual learning environments, the organization and presentation of course content is an important factor in knowledge transfer process. The custom organization may be more attractive to students, encouraging their interest in the content presented.

This work demonstrated that restructuring the content into a multilevel arrangement is a favorable factor to the process of contents personalization. By restructuring the contents, the diversification of contents presentation becomes possible due to different ways of presenting the same concept to the student.

Since the customization process continuously involves an interaction with the student, the presentation of contents also varies. The formalization of the elements involved in organizing the content was very relevant to the techniques employed for personalizing the content, making it reactive, identifying the student profile and associating the profile with a proximal learning pattern.

Table 1
Example of teacher's actions probability rule after the presentation of level FAQ.

Source level	Test answer	Next level (probability of direction)				
		Easy	Advanced	FAQ	Example	Next
FAQ	Correct	10	20	0	30	40
	Incorrect	50	5	0	40	5
	Does not know	30	5	0	50	15

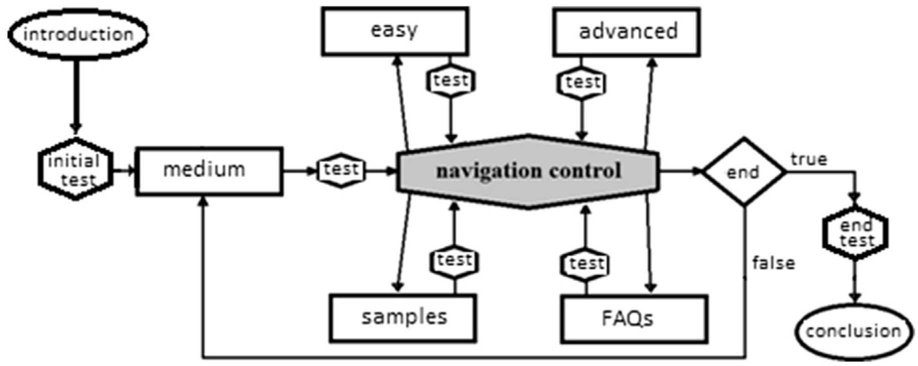

Fig. 13. Multilevel content structure of the proposed system.

Table 2
Descriptive analysis of the performance for the navigation systems.

Type of navigation	Number of cases	Mean initial score	Mean final score
Free	148	4.56	6.87
Random	31	3.99	5.93
Intelligent	31	3.92	7.21

Table 3
Comparison of the average performance for the navigation systems.

Navigation	Time in minutes	Number of visited levels	Inconsistencies	Final score	Gain
Free	37.88	35.34	0.63	6.87	39.59
Random	35.97	45	1.06	5.93	32.60
Intelligent	26.80	26.71	0	7.21	58.02

In order to ensure the process of personalizing the content, it was necessary to establish mechanisms to identify the characteristics of the student and associate them with a learning pattern. This identification was achieved using a test set of the student's characteristics. These characteristics were used as input to a neural network, which enabled the identification of the student profile and its association with a proximal learning pattern.

Regarding the organizational capacity of the personalized and reactive content sequence, the results showed that the proposed system using intelligent navigation has achieved such organization with a smaller amount of texts and a shorter time for completing the studying activity, which resulted in greater benefit for the student.

The use of proximal learning patterns derived from students with good ability to organize individual study proved to be a promising strategy to help students who have difficulties in organizing a study exercise by themselves. Thus, the use of the proximal pattern proved to be efficient at creating a personalized organization of contents, and at aiding the student in a distance learning environment.

The organization and presentation of content is a significant factor in the transmission of knowledge process. While reading text, the student may have difficulty that can be solved by teaching action. In virtual learning environments teaching action is not immediate or this action may doesn't exist. These difficulties may discourage the student and promote the study disinterest. The system may simulate the teacher's action submitting a dynamically adjustable text when the difficulties shows. This action may help the student to continue reading and learning. Therefore, the ability to reactively organize personalized content, as shown in this work, may be a favorable factor for in promoting the study support in virtual learning environments, respecting students in their different individualities and difficulty factor.

Acknowledgments

Thanks to the support promoted by Programa de Bolsa de Incentivo ao Pesquisador (BIP) of Universidade Estadual de Goiás/Pró Reitoria de Pesquisa – Goiás – Brazil.

References

Bergman, R. (1991). Knowledge acquisition by generating skeletal plans from real world cases. In *Lectures Notes in Artificial Intelligence: Contemporary Knowledge Engineering and Cognition* (Vol. 622)Springer-Verlag.
Carvalho, S. D. (2004). An intelligent tutoring system based on self-organizing maps design, implementation and evaluation. In *International Conference on Intelligent Tutoring Systems, 2004, Maceió-Al. Proceedings of International Conference on Intelligent Tutoring Systems* (pp. 573–579).
Dastbaz, M., Mustafa, A., & Stoneham, R. (2006). Issues in design and development of Personalised e-learning systems. In *Proceedings of the ED-MEDIA 2006. World Conference on Educational Multimedia, Hypermedia & Telecomunications. Orlando. USA*.
Duda, R. O., Hart, P. E., & Stork, D. G. (2000). *Pattern classification* (2nd ed.). Wiley.
Duque, N. O., & Jiménez, D. J. (2006). Artificial intelligence for automatic generation of customized courses. In *Proceedings of the ED-MEDIA 2006-World Conference on Educational Multimedia, Hypermedia & Telecommunications, USA*.
Fontenla, J., Caeiro, M., & Llamas, M. (2010). A SOA architecture to improve the tailorability and extensibility of e-learning Systems. *IEEE Latin-America Transactions, 8*(2).

González, L. A. G., & Ruggiero, W. V. (2009). Collaborative e-learning and learning objects. *IEEE Latin-Transactions, 7*(5).
Gregory, R. J. (2007). *Psychological testing: History, principles, and applications* (5ª Edition). Pearson.
Haykin, S. S. (2000). *Redes neurais artificiais: princípio e prática*. São Paulo, SP: Bookman.
Horton, W. K. (2000). *Designing Web-based training*. USA: Wiley and Sons.
Jonassen, D. H. (2001). *The handbook of research for educational communications and technology*. Bloomington, USA: The Association for Educational Communications and Technology.
Jung, C. G. (1971). Psychological types. In *Bollingen series XX* (Vol. 6)Princeton University Press.
Martins, W., Melo, F. R., Meireles, V., & Nalini, L. E. G. (2004). A novel hybrid intelligent tutoring system and its use of psychological profiles and learning style. In *ITS 2004 – 7th International Conference on Intelligent Tutoring Systems, 2004, Maceió/AL, Brazil, August – September, 2004. Proceedings* (Vol. 3220, pp. 830–832).
Melo, F. R., Flores, E. L., & Carvalho, S. D. (2011). Multilevel Content's structure for personalization in conexionists intelligent tutor systems. In *8º CONTECSI – 8th International Conference on Information Systems and Technology Management. São Paulo/SP, Brazil, June*.
Norvig, P., & Russel, S. (2004). *Artificial intelligence: A modern approach*. New Jersey, NJ: Prentice-Hall.
Patten, J. V., Chao, C. I., & Reigeluth, C. M. (1986). A review of strategies for sequencing and synthesizing instruction. *Review of Educational Research, 56*(4), 437–471.
Phelan, K. C., Mendoza-Diaz, N. V., & Mathews, S. (2002). An example of converting a traditional course into distance learning: an open discussion. In *9th Annual International Distance Education Conference, Austin, Texas, USA*.
Rosenberg, M. J. (2001). *E-Learning strategies for delivery knowledge in the digital age*. New York : NY, USA: McGraw-Hill.
Skinner, B. F. (1968). *The technology of teaching*. Appleton-Century-Crofts.
Viccari, R. (2005). Agentes Inteligentes en Educación Virtual. In *Proceedings of Seminario Internacional. Inteligencia Artificial y Seguridad. Pereira, Colombia*.