Automatic Generation of Students' Conceptual Models underpinned by Free-Text Adaptive Computer Assisted Assessment*

Diana Pérez-Marín[1], Enrique Alfonseca[1,2], Manuel Freire[1], Pilar Rodríguez[1], José María Guirao[3], Antonio Moreno-Sandoval[4]
[1]*Computer Science Department, Universidad Autonoma de Madrid*
[2]*Precision and Intelligence Laboratory, Tokyo Institute of Technology*
[3]*Language and Computer Systems Department, Universidad de Granada*
[4]*Computational Linguistics Department, Universidad Autonoma de Madrid*
{Diana.Perez,Enrique.Alfonseca,Pilar.Rodriguez,Manuel.Freire}@uam.es; sandoval@maria.lllf.uam.es ; jmguirao@ugr.es

Abstract

In this paper, we present an automatic procedure to generate students' knowledge conceptual models from their answers to an automatic free-text scoring system. The conceptual model is defined as a simplified representation of the concepts and relationships among them that each student keeps in his or her mind about an area of knowledge. It is considered that each area of knowledge comprises several topics and each topic several concepts. Each concept can be identified by a term that the students should use. A concept can belong to one topic or to several topics. The conceptual model is graphically displayed to the teachers as a conceptual map so that they can instantly see which concepts have already been assimilated and which ones should still be reviewed as they have been misunderstood.

1. Introduction

In 1978, Ausubel presented a new theory of learning whose main premise is that meaningful learning occurs when learners link new information with relevant, preexisting concepts or propositions in their cognitive structure [1].

According to this theory, concepts are objects, events, situations or properties that possess common criterial attributes and are designated by some sign or symbol. Two methods of concept learning can be distinguished: concept formation and concept assimilation. Concept formation is for young children that acquire the criterial attributes of the concepts through direct experience. While concept assimilation (the predominant method for concept learning) is for school children and adults that actively incorporate new knowledge by linking it with previous one (anchorage of new information to existing ideas).

The theory of Meaningful Learning is the fundamental pillar of the conceptual maps that are useful and powerful tools to visually represent as a graph the conceptual structure that someone has about an area of knowledge [2,3]. The basic elements of a conceptual map are three: the concepts (represented in the graph as the nodes), the labels (represented in the links that join the nodes to indicate the type of relationship between these nodes) and the propositions (the semantic unit created from the combination of the concepts and labels).

Conceptual maps have been extensively used in many different applications and by very different users. In this paper, we are going to focus on their use as knowledge evaluators. In particular, as many teachers complain about lack of feedback from their students to know if they are understanding the concepts exposed in the lessons, the conceptual maps can be used as a powerful tool to visually represent what the students know, that is, their knowledge conceptual models. Moreover, as an alternative of the traditional approach, we present a system that, for the first time, allows the students not to be directly asked to create the conceptual map, but to generate it from their answers to the automatic free-text scoring system [4] Willow [5].

* This work has been sponsored by Spanish Ministry of Science and Technology,project number TIN2004-0314.

Willow is the first free-text Adaptive Computer Assisted Assessment system that combines the techniques of Natural Language Processing to assess the students' answers and Adaptive Hypermedia (AH) to take into account static and dynamic information about the student and to adapt the assessment to each student's particular profile and performance during the assessment section. This allows the system to keep track of how each student deals with the terms (concepts) of the topic or topics under assessment and group the topics by areas of knowledge. Each student's conceptual model can be shown to the teachers as a conceptual map graphically displayed by the CLOVER system [6]. In this way, teachers can have instant feedback about how well each student is able to use the concepts taught.

The paper is organized as follows: Section 2 defines what we consider as the conceptual model of a student; Section 3 focuses on the goals and utility of these models; Section 4 outlines Willow and explains the procedure to generate them automatically; Section 5 describes how to visualize the conceptual models as conceptual maps. Finally, Section 6 ends with the conclusions and lines of future work.

2. Definition

First of all, in order to define what a conceptual model is, it is necessary to have a clear understanding of what a model and a concept is.

Many different definitions can be found for model in the literature. Nevertheless, all of them share a main idea: a model is a simplified representation of the real world. In fact, we define a model as an imitation of a particular phenomenon of the real world in a smaller scale, but preserving all the details, necessary to make easier to study it. Models have been used for many different applications such as testing a theory or predicting economical investment evolutions.

Regarding a concept, there are also many different ways to define it. However, we prefer the definition given in the introduction. If we examine this statement we can identify two important features of a concept: its abstract nature and its necessity of being denominated. Firstly, it does not specify a concrete object or fact, but a regularity in a set of them. For instance, a concept is not a particular book or books but what we, human beings, understand as a book: a set of pages fastened together and fixed inside a cover of stronger paper or cardboard talking about some topic or group of topics, with a beginning, an argument and an ending. Secondly, every concept needs to be assigned a label, something that identifies it since without this label the concept is inaccessible. In this paper, this label will be called a term. When we talk or write, our language involves the use of many different words. Not all of them can be considered terms, since they only serve to express actions and to link them. In fact, it can be seen that terms are usually expressed by nouns.

Once model and concept have been defined, we can say that a conceptual model is a particular kind of model that captures the cognitive structure of a human being about an area of knowledge. That is, a conceptual model is a simplified representation of the concepts and relationships among them that somebody has in his or her mind (the portion of real world that we want to model) about an area of knowledge in a certain instant.

It is very important to highlight the fleeting nature of a human being conceptual model. As we are all the time subject to new information and experiences that modify our previous assumptions and create new relationships between already existing concepts and new ones, the conceptual model is continuously changing itself. Thus, we can only model snapshots of its configuration at a given moment.

Each concept has a confidence value that reflects how well it is understood at the time that the model is shown. A lower value means that the student does not know the concept as he or she does not use it, while a higher value means that the student confidently uses that concept. This confidence value is automatically updated as the student keeps answering questions to a free-text scoring system.

According to Ausubel's theory, the hierarchy of concepts of a student is an evidence of his or her understanding of the topic. Therefore, we have not given the same relevancy to all the concepts but we have created a hierarchical model in which three different kinds of concepts can be distinguished: basic-concepts (BCs) such as *"book"*, *"semaphore"*, *"process"* that only refer to what Novak calls a regularity; topic-concepts (TCs) that group several basic-concepts as these concepts represent a category of them such as *"concurrency"* that comprises basic-concepts such as *"semaphore"* or *"process"* but not *"book";* and, the area of knowledge-concepts (ACs) that represent a group of topic-concepts, a group of categories such as *"operating system"* that comprises topic-concepts such as *"concurrency"*. Thus, if the AC of a student's conceptual model has a high confidence value, just by looking at this high-level concept we can see that this student seems to have assimilated it.

Regarding the relationships among the concepts, three kinds of links can be identified: type 1 between ACs and TCs; type 2 between TC and BC; and type 3 between two BCs. A concept can belong to one topic or to several topics but, in any case, it can only appear once in the conceptual model.

Type 3 links are very important as they reflect how the concepts are related in the cognitive structure of the student. Type 2 are also important since they give us information about how the basic-concepts are grouped in topic-concepts and how well each topic-concept is understood. From a BC, several type 2 links to TCs with a certain strength can be created. The higher the strength, the more relevancy that the BC has for this topic. Besides, for each BC that belongs to different TCs, the student's ability to deal with the BC in the different contexts provided by the different topics can be studied. TCs are not linked among them, as the relationships between the topics are already captured by the links of type 3. Finally, from a TC, only one type 1 link can be created to an area-of-knowledge, as all the TCs are included in the more representative AC.

3. Goals

In [3] Novak and Gowin stated that conceptual maps (a graphical visualization of someone's conceptual model) can be used to represent the cognitive structures. This is very helpful in several fields such as student guidance in e-learning courses, automatic diagnosis of diseases in expert systems, and in education, for all ages and phases, from the design of the lessons to their evaluation. In particular, we are going to focus on its use as knowledge evaluators.

Traditionally, students are asked to draw their conceptual model in a conceptual map. In this way, the conceptual organization of the learner about an area of knowledge can be seen [7]. We believe that it is interesting to directly ask the students to draw the conceptual map in order to know how they are structuring the information in their minds. However, we also believe that this approach can be complemented and improved if the conceptual model is automatically generated. Our motivation is not to put an automatic score of the student or to reduce the tasks assigned to the students, but to make easier for the teachers to have instant access to the fleeting conceptual model of each student and what is more important, to have the possibility of automatically generating the conceptual model of the class. This is very hard to do by hand from hundreds of particular student models.

To sum up, two main goals of the automatic generation of each student's knowledge conceptual model both in e-learning and presential courses can be identified: to have instant feedback to know how well the concepts taught have been understood and to identify misunderstandings.

4. Procedure

In order to automatically generate the student knowledge conceptual model, we use the answers that the students give to an automatic free-text scoring system called Willow.

Willow is an on-line application (available at http://orestes.ii.uam.es:8080/ateneaAdaptativa/jsp/loginAtenea.jsp) that interactively asks open-ended questions to the students and, immediately evaluates the students' answers (in Spanish or in English) to give them instant feedback. Its main aim is to review the concepts exposed in the lessons. It works by comparing the student's answer against a set of correct answers (the references). The first year a course is inserted in Willow, teachers are asked to write a set of references per question. The following years, a genetic algorithm automatically chooses the best answers of the students (the ones that achieved the highest scores) and it stores them as new references. In this way, the source of correct answers is not only the teachers but also the good students. Moreover, the references dynamically grow as more students introduce good (high-scored) answers into the system. The feedback provided to the students can be a numerical score, the processed answer with the best points of the answer marked in green and the references with the terms identified underlined [5].

The terms that label the concepts per topic and area of knowledge have been extracted semi-automatically from the Willow's references. Firstly, a program analyzes the text and extracts candidates. Afterwards, a human expert checks the list and selects the relevant terms for the evaluation.

The automatic procedure looks for single words or sequences of words (multi-words) that meet two conditions. Firstly, that the term neither starts nor ends with a grammatical word (i.e. articles, pronouns, prepositions, conjunctions, etc.). However, these non-lexical words can appear inside the term (e.g. *tiempo de retorno*). Secondly, that the term is repeated more than three times.

The candidates (single words or multi-words) are strings that appear several times in the text. From this list, the "actual" terms are selected.

Once the basic-concepts have been identified, they are stored in Willow's database with its global frequency (the number of times the term appears in all

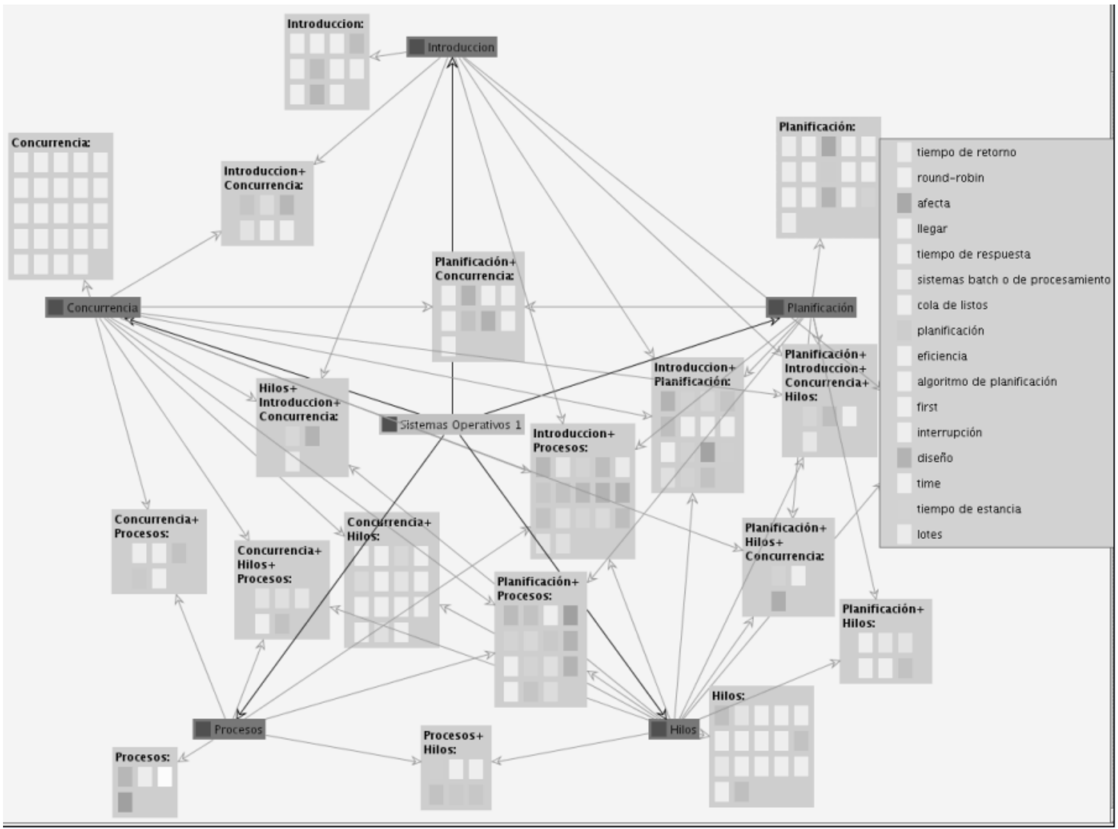

Figure 1. Representation of an example student's generated conceptual model

the references) and its local frequency per question (the number of times the term appears in each particular question). Thus, whenever a student logs into Willow and answers a question, the system looks for the concepts stored in the database. For each concept found (Z), it updates the confidence-level that the student knows that concept and how to use it correctly, by recalculating the mean between two intermediate values that can be called c1 and c2:

$$c1 = \frac{\sum_{Y} scStX \times freqRfs}{freqGlob}$$

where the numerator is the result of adding per each question Y the sum of the product of the score that each student X has achieved in the question Y and the frequency of the term in the references of Y, and the denominator is the frequency of Z in the references of all questions.

$$c2 = \frac{\sum_{Y} \frac{freqAns}{freqRfs}}{numQuest}$$

where the numerator is the result of adding per each question Y the quotient of dividing the number of times Z appears in the student's answer by the number of times Z appears in the references of the question Y and the denominator is the total number of questions.

It can be seen that the first heuristic (c1) is more related to Willow's score, that is, the better the score Willow gives a student, the higher the confidence-level that the terms appear in the student's answer. While the second heuristic (c2) is more related to the frequency of the term in the student's answer compared with the frequency of that term in the teachers'.

5. Representation

We have developed a viewer of the student model that shows it as a conceptual map, that is, as a graph in which the nodes are the concepts (BCs, TCs or ACs) and the links are the relationships between the concepts. The links can be of type 3 (between BCs), 2 (between BC and TC) and 1 (between TC and AC).

For the nodes, a color schema has been established in two levels to indicate the type of concept and the confidence level that Willow has about whether this student knows that concept. The type of concept is indicated by the background color: white for BCs, blue for TCs and grey for ACs. The confidence level is indicated by the foreground color: red for unknown concepts, yellow for unsure concepts and green for known concepts. Besides, the lighter the color is, the lower is the confidence level. Thus, a lighter green means that although the concept is known, the student sometimes uses it inaccurately.

For all links, the labels are “talks about”. For instance, if we have the TC “concurrency” and the BC “semaphore”, it is possible to find in some students' conceptual map the proposition “concurrency talks about semaphores” indicating that they know the concepts “semaphore” and “concurrency” and that they have meaningfully learnt it as they have linked it with concurrency. The conceptual map is shown in a spider-like form in Figure 1 as generated by a customized version of CLOVER [6]. It shows a real Spanish course of Operating System comprising five topics: Introduction, Processes, Threads, Concurrency and Scheduling .

CLOVER is a graph visualization environment that uses clustering to aggregate related nodes together and reduce visual clutter. To represent concept maps, we have clustered BCs according to the TCs they have been assigned to. For instance, all BCs assigned to TCs 'A' and 'B' will be grouped in a single cluster, 'A+B'.

Only one concept map is represented at a given time, with the AC node fixed in the center of the graph and TCs arranged radially at a fixed distance around it. The initial layout is performed automatically, however the user is free to move nodes around and otherwise interact with the graph once created. When hovering the mouse pointer over a node, a tooltip with the node's contents is displayed (e.g., rightmost box of Figure 1). Similar context-aware tooltips are available for edges. If the user clicks on a group of terms, the group will be exploded and the individual terms will be automatically layed out on the graph by themselves. When individually displayed, the layout algorithm seeks to locate each term nearer to a given topic so that the greater the strength of a type-2 relation between a term and a topic, the closer that term is placed to that topic. When a different node group is selected, concepts composing the previously expanded one will be collapsed again, and those corresponding to the new one will be expanded.

6. Conclusions and future work

In this work, a procedure to automatically generate students' conceptual models from their answers to the open-ended questions formulated by an automatic free-text scoring system and to show them to the teachers as conceptual map has been presented for the first time.

Due to the fleeting nature of the conceptual models, the fact that they are automatically generated is very helpful, since it allows teachers to instantly see anytime they want the particular knowledge of the concepts and their relationships for a student and, as many times as they choose. Moreover, it could be done not only for one student, but for a whole class. This procedure can be used not only for teachers of presential courses but for e-learning tutors. In particular, it is interesting for e-learning tutors that are given the possibility of keeping track of the students' progress.

Some lines of future work are to improve the terms identification module by using generic corpora and performing more complex statistical analyses; not to use all the concepts identified by the terms-identification module, but to automatically filter them so that there is no more than *N* terms (e.g., 50) per area of knowledge, and the teachers do not feel overloaded with too much information; to give a weight to each concept to help the teachers to identify which BCs are more important than others; and, to give them the possibility of adding the concepts that they consider more relevant for the evaluation of their students.

7. References

[1] D.P. Ausubel, J.D. Novak, and H. Hanesian, *Educacional Psychology: a cognitive view*, 2nd. ed., Holt, Reinhart and Winston, New York., 1978.

[2] J.D. Novak, *A Theory of Education*, Cornell University Press, Ithaca, New York, 1977.

[3] J.D. Novak, and D.B. Gowin, *Learning How to Learn,* Cambridge University Press, U.K., 1984.

[4] S. Valenti, F. Neri and A. Cucchiarelli, “An Overview of Current Research on Automated Essay Grading", *Journal of Information Technology Education,*2, published by the Informing Science Institute, Santa Rosa, California, USA, 2003, pp. 319-330.

[5] D. Perez-Marin, E. Alfonseca and P. Rodriguez, “On the dynamic adaptation of Computer Assisted Assessment of free-text answers”, *Proceedings of the Adaptive Hypermedia Conference*, Dublin, Ireland, 2006.

[6] M. Freire M, and P. Rodríguez, “A graph-based interface to complex hypermedia structure visualization”, *Proceedings of the Working Conference on Advanced Visual Interfaces* (AVI), ACM Press, Gallipoli, Italy, 2004, pp. 163-166

[7] M.A. Moreira, “Mapas conceptuales y aprendizaje significativo”, *O ENSINO Revista Galáico Portuguesa de Sócio Pedagogia y Sócio-Lingüística*, Pontevedra/Galicia/España and Braga/Portugal, no. 23-28:87-95, 1988