Initializing Student Models in Web-based ITSs: a Generic Approach

Victoria Tsiriga, Maria Virvou
Department of Informatics,
University of Piraeus,
80 Karaoli & Dimitriou St.
Piraeus 18534, Greece
vtsir@unipi.gr, mvirvou@unipi.gr

Abstract

The issue of initializing the model of a new student is of great importance for educational applications that aim at offering individualized support to students. In this paper we introduce a general framework for the initialization of the student model in Web-based educational applications. According to the framework, the student modeler makes initial estimations of the new student's knowledge level and error proneness based on the models of other similar students. This is done by applying a machine learning technique. The similarity between students is calculated taking into account different student characteristics for different teaching domains. We have implemented the proposed methodology in two different tutoring domains, namely language learning and mathematics. An evaluation study conducted in the case of the Web-based language learning system, showed that the use of the framework can initialize student models in a sufficiently accurate way, given the little information about individuals available.

1. Introduction

The initialization of a student model is an important function of the student modeler of Intelligent Tutoring Systems (ITS). Indeed, it seems unreasonable to assume that every student starts up with the same knowledge and misconceptions about the domain being taught. However, the process of the initialization of the student model has often been neglected or it has been dealt using trivial techniques.

Aïmeur et al. in [1], distinguish between three approaches for initializing the student model:

1. The ITS may assume that a new student knows nothing (e.g. [2, 3]) or has some standard prior knowledge about the domain (e.g. [4]). Although this approach is the easiest way to address the problem of initializing the student model, it has poor performance for students who have different knowledge from the one initially assumed by the system.

2. The student's prior knowledge may be evaluated by using a pre-test. An accurate solution for the problem of assessing the prior knowledge of students would be to use exhaustive pre-tests that contain questions related to every topic in the domain being taught. However, in case of a broad domain, using this method would expect the student to answer questions for a long period of time before s/he could actually start working with the system. An alternative that would reduce the number of questions needed would be the use of adaptive pre-testing (e.g. [5]). Adaptive pre-tests are constructed dynamically for each student based on her/his answers to already posed questions. Although this approach may lead to good results, if the inferences drawn from the system do not stand for a particular student, or her/his answer to a certain question of the adaptive test is a simple guess or slip, then the accuracy of the student model is reduced.

3. The system may use patterns among students in order to group similar students to categories. The most noted example is the stereotype approach, which has been introduced in [6] and has been used in many systems (e.g. [7, 8, 9]). Although stereotypes are very powerful in providing considerable information based on few observations, they do not permit the formation of an accurate student model. Furthermore, a problem of the stereotype approach to user model acquisition is that it is quite inflexible due to the fact that stereotypes are constructed in a hand-crafted way before real users have interacted with the system and they are not updated until a human does so explicitly.

In this paper we describe a generic framework for the initialization of the student model in Web-based educational applications. The framework is called Initializing Student Models (ISM) framework. The basic idea of the ISM framework is to make initial estimations concerning the knowledge level and the error proneness of a new student in each domain concept. To accomplish this task, we propose an innovative combination of stereotypes [6] and the distance weighted k-nearest neighbor algorithm [10, 11, 12]. In particular, a student is first assigned to a stereotype category concerning her/his knowledge level of the domain being taught. Then, the model of the new student is initialized by applying the distance weighted k-nearest neighbor algorithm among the students that belong to the same stereotype category with the new student.

We have used the above generic approach in two different tutoring domains, namely language learning and mathematics. The application of our framework on these two domains, which are so different from one another, was considered a good test for its generality. Furthermore, according to an evaluation study conducted in the case of the Web-based language learning system, the proposed method was successful at providing sufficiently accurate initial student models, given the fact that very little is known about new students.

2. The ISM Framework

In this section we describe the novel methodology that we have created for initializing the model of a new student. The main architecture of the ISM framework is presented in Figure 1.

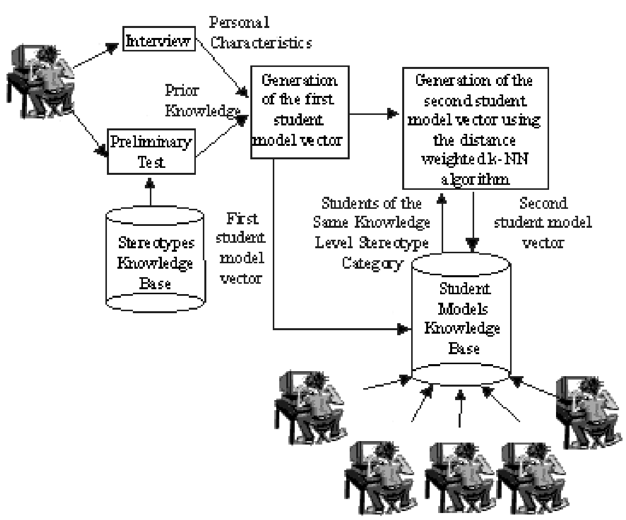

Figure1. Architecture of the ISM framework.

According to the ISM framework, initial information about a new student is acquired by the Web-based educational system, from an interview and a preliminary test. The interview is presented the first time a student interacts with the system. It contains questions related to personal student data (e.g. the student's name, age, etc.) as well as several indirectly domain dependent characteristics that may influence the students' learning process. The information acquired from the interview form the personal student characteristics.

Furthermore, according to the ISM framework, a preliminary test should be used in order to assess the prior knowledge level of the student concerning the domain being taught and/or certain important prerequisite topics. According to the student's performance on the preliminary test, the Web-based educational system assigns the student to a stereotype category concerning her/his knowledge level. The information acquired from both the interview and the preliminary test is represented as a feature vector, which is of the form:

<Student_Code, Name, Stereotype, Characteristic$_1$, Characteristic$_2$, …, Characteristic$_n$>

The n Characteristics of the vector record the student attributes that may have an effect in the student learning of the domain being taught by the ITS.

When the above student model vector has been formed, the student modeler produces a second vector that represents the system's estimations of the new student's degree of knowledge and error proneness for each concept in the domain knowledge. The second student model vector is of the form:

<Student_Code, Knowledge_Level(Concept$_1$), Errors(Concept$_1$), Knowledge_Level(Concept$_2$), Errors(Concept$_2$), …, Knowledge_Level(Concept$_n$), Errors(Concept$_n$)>

This is based on the distance weighted k-Nearest Neighbor algorithm. In particular, the degree of knowledge or error proneness of a new student concerning a concept is calculated as the weighted mean value of the degree of knowledge or error proneness of the students that belong to the same stereotype with the student in question. However, for these students, the Web-based ITS has constructed an individual student model based on their actual behavior which has been observed while they interacted with the system. It should be noted that the stereotype category of a student should be updated as the student progresses in learning the domain through the use of the Web-based ITS.

The weight of the contribution of each student depends on her/his similarity with the new student. Furthermore, the similarity between students is calculated taking into account the student characteristics that may play a significant role in the students' learning of the domain. These are the student attributes that form the first student model vector. Thus, the first step for the application of the ISM framework is to specify which students' characteristics may affect the way they learn. The specification of these characteristics result from existing expertise or empirical studies about how students learn. For example, in an ITS about language learning, an important student characteristic that influences the way a student learns a foreign language is what her/his mother tongue is. If there are no students belonging to the same stereotype with the new student, then her/his student model is initialized using the default assumptions of the active stereotype.

3. Case Study I

The first system that was implemented using the ISM framework for initializing the model of a new student was Web-Passive Voice Tutor (Web-PVT) [13]. Web-PVT is an adaptive and intelligent Web-based tutoring system that aims at teaching non-native speakers the domain of the passive voice of the English language [14]. The system constructs an individual model for each student and uses it to support her/him while studying theory and solving exercises.

3.1. Constructing Initial Student Models

The construction of initial models in Web-PVT is based on the ISM framework. In the case of Web-PVT the ISM framework is instantiated by assuming that students of similar knowledge level of English, who have the same mother tongue and know the same foreign languages have similar strengths and weaknesses when they learn English as a foreign language.

In particular, Web-PVT acquires initial information about a new student when s/he interacts with the Web-based ICALL for the first time. Personal characteristics are directly provided by the student using a set of questions posed to the student in the initial phase of the system's usage. These characteristics include the full name of the student, her/his mother tongue, prior knowledge of other languages, and a self-estimation of how careful s/he is while solving exercises. In particular, the student can select among three different categories, namely careless, averagely careful, and careful. Furthermore, Web-PVT also uses a preliminary test in order to assess the initial knowledge level of the student concerning the passive voice of the English language. The preliminary test is given to students before they have ever interacted with Web-PVT. Then, based on the student's performance on the preliminary test, s/he is classified to one of the four stereotypes, namely novice, beginner, intermediate and advanced. When all these pieces of information are acquired, the system represents them in a feature vector (first student model vector).

Then based on the first vector constructed for the new student and the vectors of all the students that belong to the same knowledge level stereotype, Web-PVT proceeds to the generation of the second domain related student model vector. In particular, for each concept in the domain knowledge there are two feature-value pairs related to it in the second student model vector. The first pair represents an estimation of the student's degree of knowledge concerning the particular concept that the pair is associated with, whereas the second represents an estimation of the student's proneness to make mistakes while using this concept. The values of the estimations are within the range [0..1]. When estimating the degree of a student's knowledge of a particular concept, 0 depicts the system's belief that the student does not know the concept at all, while 1 represents the system's belief that the student knows this concept very well. Furthermore, when referring to error proneness concerning a concept, 0 represents the system's belief that the student never makes mistakes when using a concept, and 1 represents the system's belief that the student always makes mistakes related to the concept.

Web-PVT produces the second vector for a new student taking into account the characteristics of the first student model vector of the new student as well as of the other students who belong to the same stereotype concerning their knowledge level. For example, the degree of knowledge concerning a particular concept of a new student who belongs to the intermediate stereotype would be calculated as the weighted mean value of the degree of knowledge of this concept of all the students that also belong to the intermediate stereotype. The weight of the contribution of each student is determined by her/his similarity with the new student. In Web-PVT, the similarity between students is calculated taking into account the degree of carefulness of students, their mother tongue and the other languages that they already know. One factor that guided our decision toward the choice of the student characteristics that would be used for measuring the similarity between students was the issue of language transfer. Language transfer is the interference resulting from the similarities and differences between the target language and any other language that has been previously (and perhaps imperfectly) acquired [15].

3.2. Evaluation of the Initialization Module

In order to assess the effectiveness of the proposed approach to initialize the student models in Web-PVT, we conducted an evaluation study. Three teachers of English and their students (118 students) participated in the experiment. Among the 118 students, 21 were found to belong to the novice stereotype, 38 to the beginner stereotype, 36 to the intermediate stereotype and 23 to the advanced stereotype.

The teachers were asked to evaluate five randomly chosen initial student models from each one of the four supported stereotypes (novice, beginner, intermediate and advanced) at two phases. At first the teachers evaluated five models of each stereotype, before any student of this particular stereotype had been registered to the system. At this point, the evaluation could be considered as an evaluation of the stereotypes that were hand-coded to Web-PVT. At the second phase, teachers evaluated the initial models of five randomly chosen students of each stereotype after Web-PVT had constructed the models of 15 students of each stereotype, based on direct observations of the students' interaction behavior.

The experimental hypothesis was that an initial student model that was built after the system had constructed models of other students of the same knowledge level stereotype would be superior to the initial student model constructed by using the default assumptions of the stereotype that becomes active for a specific student. In order to evaluate the hypothesis, we used a one-tailed paired t-test with the alpha level set to 0.05. The results of this study showed that the student modeler in all the cases performed better at initializing the model of a new student when it took into account other students of the same knowledge level stereotype. In most of the cases, the initial student model that was generated after 15 students of each stereotype had interacted with Web-PVT, achieved significantly higher acceptance from teachers than those produced when no similar students were found in the student model knowledge base.

4. Case Study II

In order to test the generality of the ISM framework, we applied this method for initializing student models for a Web-based algebra tutor, which is called Web-EasyMath. In the case of Web-EasyMath, the ISM framework is instantiated by assuming that students of similar knowledge level, who attend the same class (and instructors) and have similar skills in simple arithmetic operations have similar strengths and weaknesses when learning the new topic of algebraic powers.

As indicated by the ISM framework, Web-EasyMath represents student models as a pair of feature vectors. Similarly with Web-PVT, the first vector is constructed based on the data acquired during the student's first interaction with Web-EasyMath. However, due to the different nature of the tutoring domain, the student attributes that are recorded are different from Web-PVT. In particular, in the case of Web-EasyMath, the characteristics of the student that are obtained via the interview include the name of the student, the specific class that s/he belongs to, and a self-estimation of how careful the student is while solving exercises (s/he can select between three distinct categories, namely careless, averagely careful, and very careful). Furthermore, there is also a difference in the inferences made by the student's performance on the preliminary test. In particular, apart from categorizing the student to a stereotype concerning her/his knowledge level (novice, beginner, intermediate, and advanced), the system also infers the student's ability in using simple arithmetic operations. This is done based on a set of questions of the preliminary test that are related to simple arithmetic operations.

When the first vector for a new student has been formulated, Web-EasyMath uses the models of other students that belong to the same knowledge level stereotype with the student in question in order to estimate the degree of knowledge and error proneness of the new student in each concept of the domain. In particular, the prediction of the knowledge level or the error proneness of a student on a specific concept is made based on the weighted mean value of this attribute as it is acquired by the models of other students. Similarly with Web-PVT, the contribution of each of the students that belong to the same stereotype concerning their knowledge level is weighted based on her/his distance from the new student. However, in Web-EasyMath the distance between two students is estimated taking into account different student characteristics. In particular, the student attributes that are used are the specific class that students belong to, the students' performance in the calculation of basic arithmetic operations between numbers (addition, subtraction, multiplication and division) as well as their degree of carefulness while solving exercises. A more detailed description of the process of initializing student models in Web-EasyMath is presented in [16].

5. Conclusions

In this paper we have described a framework that addresses the initialization of the student model in Web-based educational applications. In particular, the framework proposes the use of a novel combination of stereotypes and the distance weighted k-Nearest Neighbor algorithm in order to set initial values to the model of a new student. Stereotypes are used to make initial hypotheses about the knowledge level of the student, whereas the distance weighted k-Nearest Neighbor algorithm is utilized to refine the estimations of the student's knowledge level of each concept and her/his proneness to make mistakes concerning this concept, based on the student's similarity with other students of the same stereotype category. The similarity between students is estimated based on the student characteristics that may play a role in the student's performance while learning the domain being taught by the application.

We have tested the generality of our framework by applying it to two totally different tutoring domains, namely language learning (Web-PVT) and mathematics (Web-EasyMath). According to the evaluation of the student modeler of Web-PVT, the system indeed increased the accuracy of the initial models of students as compared to using the stereotype approach alone.

6. References

[1] E. Aïmeur, G. Brassard, H. Dufort and S. Gamps, “CLARISSE: a machine learning tool to initialize student models”, In S. A. Cerri, G. Gouardéres and F. Paraguaçu (eds.) *Proceedings of the Sixth International Conference on Intelligent Tutoring Systems, Lecture Notes in Computer Science*, Vol. 2363, Springer-Verlag, Berlin Heidelberg, 2002, pp. 718-728.

[2] P. da Silva, R. Van Durm, K. Hendrikx, E. Duval and H. Olivie, “A simple model for adaptive courseware navigation”, In Lobodzinski, S. and Tomek, I. (eds.) *Proceedings of WebNet ’97, World Conference of the WWW, Internet and Intranet*, Charlotsville VA, AACE, 1997, pp 959-960.

[3] J. Tchètagni and R. Nkambou, “Hierarchical representation and evaluation of the student in an intelligent tutoring system”, In S. A. Cerri, G. Gouardéres and F. Paraguaçu (eds.) *Proceedings of the Sixth International Conference on Intelligent Tutoring Systems, Lecture Notes in Computer Science*, Vol. 2363, Berlin Heidelberg, Springer-Verlag, 2002, pp. 708-717.

[4] T. Heift and D. Nicholson, “Web Delivery of Adaptive and Interactive Language Tutoring”, *International Journal of Artificial Intelligence in Education*, 12, 2001, pp. 310-324.

[5] J. Kurhila, M. Miettinen, M. Niemivirta, P. Nokelainen, T. Silander and H. Tirri, “Bayesian modeling in an adaptive on-line questionnaire for education and educational research”, In H. Ruokamo, O. Nykänen, S. Pohjolainen and P. Hietala (eds.) *PEG 2001: Intelligent Computer and Communications Technology - Learning in On-line Communities, Proceedings of the Tenth International PEG Conference*, 2001, pp. 194-201.

[6] E. Rich, “User modelling via stereotypes”, *Cognitive Science*, 3, 1979, pp. 329-354.

[7] M. Virvou and M. Moundridou, “Student and instructor models: two kinds of user model and their interaction in an ITS authoring tool”, In M. Bauer, P. Gmytrasiewicz and J. Vassileva (eds.) *Proceedings of the Eighth International Conference on User Modeling, Lecture Notes in Artificial Intelligence*, Vol. 2109, Berlin Heidelberg, Springer-Verlag, 2001, pp. 158-167.

[8] K. Bontcheva, “Adaptivity, adaptability, and reading behaviour: some results from the evaluation of a dynamic hypertext system”, In P. De Bra, P. Brusilovsky, and R. Conejo (eds.) *Proceedings of Second International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems, Lecture Notes in Computer Science*, Vol. 2347, Berlin Heidelberg, Springer-Verlag, 2002, pp. 69-78.

[9] J. Garofalakis, E. Sakkopoulos, S. Sirmakessis and A. Tsakalidis, “Integrating adaptive techniques into virtual university learning environment”, In V. Petrushin, P. Kommers, Kinshuk and I. Galeev (eds.) *Proceedings of the 2002 IEEE International Conference on Advanced Learning Technologies: Media and the Culture of Learning*, Palmerston North New Zealand, IEEE Computer Society, 2002, pp. 28-33.

[10] S. Dudani, “The distance-weighted k-nearest-neighbor rule”, *IEEE Transactions on Systems, Man and Cybernetics*, 6, 1976, pp. 325-327.

[11] J. MacLeod, A. Luk, and D. Titterington, “A re-examination of the distance-weighted k-nearest-neighbor classification rule”, *IEEE Transactions on Systems, Man and Cybernetics*, 17, 1987, pp. 689-696.

[12] W. Emde, and D. Wettshereck, “Relational instance-based learning”, In L. Saitta (Ed.) *Proceedings of the Thirteenth International Conference on Machine Learning*, Morgan Kaufmann, San Francisco, CA, 1996, pp. 122-130.

[13] V. Tsiriga and M. Virvou, “Dynamically initializing the student model in a web-based language tutor”, In *Proceedings of the 2002 First International IEEE Symposium “Intelligent Systems”*, Vol. I, 2002, pp. 138-143.

[14] M. Virvou and V. Tsiriga, “Web Passive Voice Tutor: an Intelligent Computer Assisted Language Learning System over the WWW”, In Okamoto, T., Hartley, R., Kinshuk, and Klus, J. (eds.) *Proceedings of the IEEE International Conference on Advanced Learning Technologies: Issues, Achievements and Challenges*, IEEE Society Press, Los Alamitos, 2001, pp. 131-134.

[15] H. Ogata, Y. Liu, Y. Ochi and Y. Yano, “Neclle: Network-based Communicative Language Learning Environment Focusing on Communicative Gaps”, *Computers and Education*, 37, 2001, pp. 225-240.

[16] V. Tsiriga and M. Virvou, “Initializing the student model using stereotypes and machine learning”, In *Proceedings of the 2002 IEEE International Conference on System, Man and Cybernetics*, 2002.