A Probabilistic Relational Student Model for Virtual Laboratories

Julieta NOGUEZ[1], Enrique SUCAR[2], Fernando RAMOS[2]
Instituto Tecnológico de Estudios Superiores de Monterrey
[1]Campus Ciudad de México. Calle del Puente 222,
Col Ex Ejidos de Huipulco, México D.F. CP14380 México D.F.
jnoguez@itesm.mx
[2]Campus Cuernavaca. Paseo de la Reforma 182-A Col. Lomas de Cuernavaca
C.P. 62589, Temixco Morelos, México.
{esucar, framos}@itesm.mx

Abstract

The main purpose of this work is to develop an intelligent tutor system coupled to a robotics virtual laboratory, in order to offer a tutored virtual learning environment. The student model is the main component of this intelligent tutoring systems (ITS). We propose a probabilistic relational model to characterize the students, so the domain can be represented in terms of entities, their properties, and the relations between them. In this way we obtain a compact and natural representation, which can be learned from data, by combining the results from the students interaction with the virtual laboratory and the models for similar students.

Introduction

In the current state of the art in virtual laboratories, a couple of opposing postures were detected, simulation based labs and tutor based tools. The former is oriented to develop simulation software tools with a high degree of complexity able to represent situations as real as possible. No interruptions or intromissions in the operation of the system should occur [11]. This approach has permitted very complex modeling systems especially helpful in training. On the other hand, important efforts to develop virtual laboratories with the support of intelligent tutors have been identified. Nevertheless, even though a layout towards the use of intelligent tutor systems has been reflected, the quality and complexity of the simulation tools is located below the expected level, and the student usage and profit is restricted to particular system conditions [2].

Until now, the principal concern in the development of virtual laboratories has been oriented toward the technical construction of the simulation tools or remote manipulation. So the necessity of having elements from intelligent tutoring systems, capable of giving the students benefits of a virtual learning environment, has been detected.

This proposal is part of a project which considers the integration of an intelligent tutor within a virtual laboratory. An important aspect to consider is a balance between the use of a simulator or remote manipulator versus the tutoring labor based on decisions such as when to interrupt, the student performance follow up, task planning, etc. [8]. The key component of such a system is the student model, which is the focus of this work.

The student model is the main component of intelligent tutoring systems (ITS). Student models adapt the learning experience to suit the learner's perceived needs [9]. However, the practical difficulty in building reliable students models is defining the role for student models in virtual learning environments. The debate is really over what kind of information the student model should provide. At one extreme, student models provide only information about the latest student input to which the system has to react. At the other extreme, student models provide a detailed description of student's knowledge and psychological make-up. Neither extreme is feasible or desirable.

In the student model the cognitive state of a learner is inferred from two parts: the previous data about the student and student's behavior during the interaction with the system [6]. Both involve uncertainty, so a model that considers it is required. Thus, the main objective of this work is to provide a student model representation for an intelligent tutoring system of a virtual laboratory, using probabilistic relational models.

1. Student Model.

Student models based on Bayesian networks (BN) have been used to allow important simplifications in diagnosis, the task to infer the cognitive state of the student from observable data [1,3,7,10]. However, the effort required to define the network structure, the difficulty to obtain the parameters and the computational complexity of the inference algorithms, have to be considered to implement this type of models.

Other difficulties for using Bayesian nets for student modeling have been detected. The human expert should define the network structure, the conditional and unconditional dependencies, and also the initial parameters to feed the model. This task is hard because he/she needs to find a general model for several students, but each student has different knowledge, abilities, preferences and academic antecedents. Some authors decided to simplify the problem using nodes representing student's abilities and nodes which represent *binary knowledge questions*. They grouped different level questions according to five categories [12]. But this kind of system can only diagnose the abilities one by one with binary questions using a tree structured BN.

Other systems, like OLAE [7], generate the Bayesian net parameters for each resolved problem using student behavior during the solution process. It considers four node types, including rules, rules application, facts and actions. Off line, the system obtains it and propagates the evidence across the net. This makes it difficult to apply it in virtual laboratories and other web based systems, which work on line. Additionally, the solution process does not give enough information to obtain the student model.

We propose the use of *Probabilistic Relational Models* (PRM) [4] to characterize the student model, allowing the domain to be represented in terms of entities, their properties, and the relations between them. Koller states in [4] "…The basic entities in a probabilistic relational models are objects or domain entities. Objects in the domain are partitioned into a set of disjoint classes $X_1,...,X_n$. Each class is associated with a set of attributes $A(X_i)$. Each attribute $A_j \in A(X_i)$ takes on values in some fixed domain of values $V(A_j)$".

The dependency model is defined at the class level, allowing it to be used for any object in the class. Also, PRM's explicitly use the relational structure of the model, an attribute of an object to depend also on attributes of related objects. A PRM specifies the probability distribution using the same underlying principles used in specifying Bayesian networks. The assumption is that each of the random variables in a PRM, in this case the attributes $x.a$ of the individual objects x, is directly influenced by only a few others. A PRM therefore defines for each attribute $x.a$, a set of parents, which are the directed influences on it, and a local probabilistic model that specifies probabilistic parameters.

Probabilistic relational models allow to represent each attending student in the same model. Each attribute, $x.a$, of the individual student x, is directly influenced by only a few other parameters related with this student. Each class represents the set of parameters of several students, like in databases, but the model also includes the probabilistic dependences between classes for each student.

The ability to capture more data that just one learner response is crucial to provide valuable information about the Bayesian net structure and parameters. So, PRM's allow a compact and natural representation of student models for virtual laboratories. The model can be obtained from data of previous students interaction using statistical techniques. The basic idea is to obtain an initial model of the student from other student models with similar characteristics; and then to improve the model with data from her/his interaction with the virtual laboratory.

2. Case Study: tutor for a virtual robotic laboratory.

As a case study we propose an intelligent tutoring system for a mobile robots virtual laboratory. The students use this laboratory in a basic robotics course at the B. Sc. Level. We initially consider an experiment in line following, which requires some basic knowledge in control theory and programming from the students. The experiment assumes that the mechanical and sensor configuration of the robot is previously defined.

Using the UML methodology [5], some use cases were defined. These are depicted in figure 1. The experiment considers as main actor the student, and it includes two pedagogical agents, an observer which follows the interaction and result of each experiment; and a pedagogical agent which has the control and record for each student. At the end of an experiment, the observer agent reports the results to the pedagogical agent, performs the diagnosis and updates the student model.

The initial model considers five classes: student, pedagogical agent, observer agent, robot, and design. A human expert defined the structure of this model, and the dependencies between attributes of the different classes. This initial model is shown in figure 2. When the classes are defined, the human expert looks for relations between classes. For example, *Student-Design*, *Design-Robot*, *Observer Agent-Robot*, etcetera.

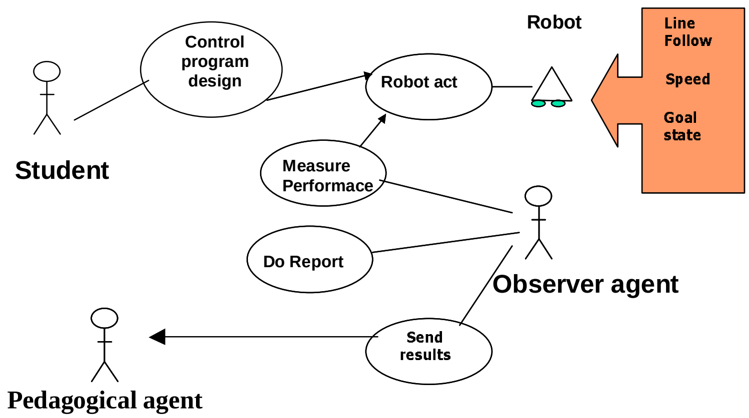

Figure 1. Student model use cases.

Each class has several attributes representing significant information for the student model. A more detailed analysis must be done to find conditional dependencies between attributes of the same class and attributes of different classes. For example, figure 2 shows conditional dependencies between the attributes of each class and between different classes.

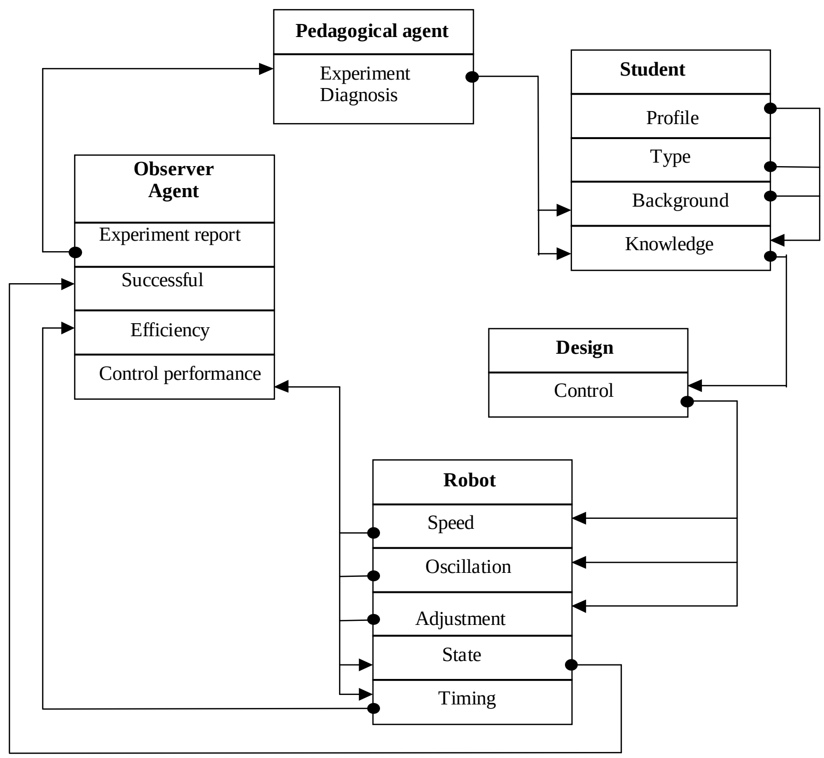

Figure 2. Initial PRM defining conditional dependences between attributes and classes.

From the PRM we can extract a Bayesian network that represents each particular student. A partial view of this network for one student is shown in figure 3. That is, for each student and theme knowledge, the PRM has different probabilities as we show in Table 1. These values were obtained from evidence in the Student.Profile, Student.Type and Student.Background by probability propagation. Each ID corresponds to one student of table 2. The Student.Background is obtained from experiment performance, propagating evidence from each experiment performed by each student as shown in table 3.

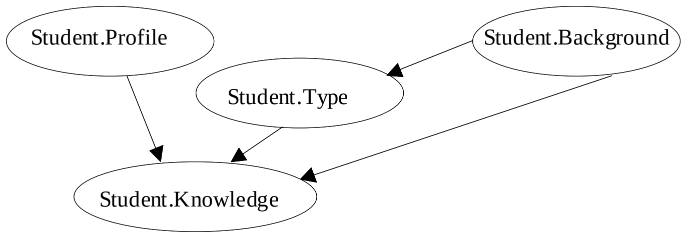

Figure 3. Conditional dependencies between attributes of the class student in the PRM.

Table 1 shows probabilistic values of Student.Knowledge (un_known, doubt, well_known) for several course themes (theme_number, theme_name) for two different students. Profile and Type Student attributes characterization is shown in table 2. The results of two experiments that correspond to some attributes (successful, efficiency and control performance) of the observer agent are shown in table 3.

Table 1. Probabilistic values of Student.Knowledge

Knowledge : Tabla

ID	theme_numbe	theme_name	un_known	doubt	well_known
954760	1	Opened lasso	0.9	0.05	0.05
954760	2	Closed lasso	0.8	0.1	0.1
954760	3	ON-OFF Control	0.4	0.4	0.2
954760	4	Proportional Control	0.7	0.15	0.15
954760	5	Derivative Proporcional Control	0.6	0.2	0.2
966058	1	Opened lasso	0.05	0.15	0.8
966058	2	Closed lasso	0.05	0.15	0.8
966058	3	ON-OFF Control	0.05	0.05	0.9
966058	4	Proportional Control	0.4	0.2	0.4
966058	5	Derivative Proporcional Control	0.2	0.5	0.3

Table 2. Profile and Type Student attributes

Profile : Tabla

ID	Name	Profile	Type
954760	Rodrigo Pérez	ISC	novice
966058	Sonia Hernández	IEC	medium
994588	Alejandra Tenorio	BI	expert

Table 3. Successful, Efficiency and Control Performance attributes of an observer agent.

Background : Tabla

ID	Num_exp	date	hour	Succesful	low_efficie	medium_effic	hight_efficie	bad_performa	regular_perfo	good_perform
954760	1	10/10/02	10:00:00 a.m.	☐	0.7	0.2	0.1	0.9	0.05	0.05
966058	1	12/11/01	06:00:00 p.m.	☑	0.1	0.1	0.8	0.1	0.2	0.7

The initial structure of the model is defined by an expert, and the parameters can be learned from data, by combining the results from the students interaction with the virtual laboratory and the models for similar students.

3. Conclusions and Future Work

This paper has introduced a PRM for student modeling in virtual laboratories. Probabilistic relational models provide a new approach to student modeling, that integrates the expressive power of Bayesian networks and of relational models. This model has several advantages: *flexibility*, it allows to consider different models for each student in a common framework, *adaptability*, by obtaining an initial model of a new learner from similar student models, and *modularity*, it can be easily extended to include more students, and more classes and attributes.

As future work, we plan to integrate this probabilistic relational student model in an intelligent tutor for a mobile robots virtual laboratory. We will implement, test and validate this model for student tutoring in a basic robotics course based on the virtual lab. Another interesting direction for future work, is to consider an extension of the PRM that incorporates dynamics, that is a dynamic probabilistic relational model, and to use it for student modeling.

References

[1] A. Gertner, K. VanLehn, Andes: A Coached Problem Solving Environments for Physics, In G. Gauthier, C. Frasson, K. VanLehn (Eds.), Proceedings of 5th International Conference of Intelligent Tutoring Systems, Springer Verlag (2000), 133-142,

[2] B. Wagner, From Computer Based Teaching to Virtual Laboratories in Automatic Control, 29th ASEE/IEEE Frontiers in Education Conference, Session 13d6-6. San Juan Puerto Rico (1999).

[3] C. Conati, A. Gertner, K. VanLehn, On-Line Student Modeling for Coached Problem Solving Using Bayesian Networks, In Kay J. (Ed), Proceedings of the 7th International Conference on User Modeling, Springer Verlag (1997), 231-242.

[4] D. Koller, Probabilistic Relational Models, 9th International Workshop, Inductive Logic Programming (1999), 3-13.

[5] G. Booch, J. Rumbaugh, I. Jacobson, The Unified Modeling Language User Guide, Addison Wesley (1999).

[6] J. Collins, Adaptative assessment using granularity hierarchies and Bayersian nets, Proceedings of the Third International Conference ITS'96, Springer-Verlag (1996). 569-577.

[7] J. Martin, K. VanLehn, Student assessment using Bayesian nets. International Journal of Human-Computer Studies. Vol 42 (1995), 575-591.

[8] J. Noguez, L. E. Sucar, M. A. López, Un sistema tutor inteligente para un laboratorio virtual de robótica, Mexican Internacional Conferencia on Artificial Intelligence, Workshop Intelligence Virtual Environments, Special Issue (2002), 27-34.

[9] J. Self, The Role of Student Models in Learning Environments, Lancaster University, Technical Report No. 194 (1994).

[10] K. VanLehn, Bayesian student modeling, user interfaces and feedback: a sensitivy analysis, International Journal of Artificial Intelligence in Education (2001), 154-184.

[11] L. Richard, G. Gourderes, An agent-Operated Simulation-Base Training System –Presentation of the CMOS Project, A. I., In S.P. Lajoie and M. Vivet. Artificial Intelligence in Education, IO Press (1999), 343-351.

[12] W. Murray, A Practical Approach to Bayesian Student Modeling. In B. Goettle, H. Halff, C. Redfield, and V. Shute (Eds.), Proceedings of 4th International Conference of Intelligent Tutoring Systems, Springer Verlag (1998), 424-433.