

An Ontology-Based Model for Student Representation in Intelligent Tutoring Systems for Distance Learning

Ioannis Panagiotopoulos, Aikaterini Kalou,
Christos Pierrakeas, and Achilles Kameas

Educational Content, Methodology and Technology Laboratory (e-CoMeT Lab.)
Hellenic Open University, Patras, Greece
{gpanagiotopoulos, kalou, pierrakeas, kameas}@ecomet.eap.gr

Abstract. An Intelligent Tutoring System (ITS) offers personalized education to each student in accordance with his/her learning preferences and his/her background. One of the most fundamental components of an ITS is the student model, that contains all the information about a student such as demographic information, learning style and academic performance. This information enables the system to be fully adapted to the student. Our research work intends to propose a student model and enhance it with semantics by developing (or via) an ontology in order to be exploitable effectively within an ITS, for example as a domain-independent vocabulary for the communication between intelligent agents. The ontology schema consists of two main taxonomies: (a) student's academic information and (b) student's personal information. The characteristics of the student that have been included in the student model ontology were derived from an empirical study on a sample of students.

Keywords: Ontology, intelligent tutoring systems, stereotypes, personalized learning, student model.

1 Introduction

Intelligent Tutoring Systems (ITSs) are complex systems that can be adapted easily to each student's cognitive features, characteristics and learning progress [1]. These systems use a large amount of educational knowledge and many of them also employ pedagogical methodologies.

Traditional ITSs consist of the following four modules: (a) the domain module, which contains all the knowledge (educational content), (b) the student model, (c) the pedagogical module, which contains all the information relevant to the various pedagogical decisions and (d) the user interface which enables communication between the user and the system [2]. Especially, in multi-agent architectures the communication between these modules is achieved through the communication of intelligent agents assigned to each module. So, for example a learner model agent is responsible for answering queries from other agents about learner's information, which information is included in the student model.

The characteristics and progress of the students are captured in the student model. This is achieved by using AI techniques to represent pedagogical decisions, domain knowledge, and personal information about the student [3]. Since there are many candidate characteristics of a user that can be included in the student model, the selection of the appropriate characteristics is a very challenging and significant procedure. Consequently, we have to obtain a tradeoff of the completeness of the model so that the systems can be adapted successfully and their performance is not affected. Some of the basic student characteristics maintained in the model are: (a) demographic information, (b) knowledge of the teaching domain, (c) background and interests, (d) learning styles and interaction preferences and (e) learning goals and specificities that can affect the learning procedure. From the above characteristics (a) is a basic feature when describing a user, while (b) and (c) are essential parameters in every educational process. Characteristics (d) and (e) are considered as crucial for a user-centered intelligent tutoring system in order to deliver the appropriate educational material, according to the individual's needs.

In this paper, we propose to use an ontology to represent all the above-mentioned information in a student model. In addition, the model includes the learner characteristics, as they were identified after an empirical study conducted on a sample of students of the Hellenic Open University.

The rest of the paper is structured in the following way. In section 2, we discuss the student modeling approaches based on ontologies. In Section 3, we elaborate on the proposed student model, in terms of the learning styles, the modeling approaches and the basic characteristics of the student. In Section 4, we outline the ontology that represents the proposed student model. Finally, in Section 5, we discuss future work and summarize our conclusions.

2 Related Work

Ontologies have been widely used for student modeling mainly for two reasons: (a) ontologies support the formal representation of abstract concepts and properties in a way that they can be reused by many tasks or extended if needed and (b) they enable the extraction of new knowledge by applying inference mechanisms (e.g. reasoner) on the information presented in the ontology.

Therefore, a plethora of ontology-based approaches for student modeling have been proposed in the field of ITSs. Paneva [4] proposes an ontology-based student model for eLearning systems that adopts technologies and standards from the Semantic Web. Chen et al. [5] describe a domain-independent student model for a multi-agent intelligent educational system (IES). In [6], they propose a student model ontology for an e-Learning system. The ontology is based on the representation of prior knowledge of the student and his/her learning style. Jeremic et al. [7] describe a student model for the Design Pattern intelligent tutoring system. Similarly, many other approaches have been presented in literature (i.e. [8], [9]). Additionally, in [10] the authors propose a student modeling mechanism for Intelligent Virtual Environments for Training (IVETs). They divide the student information in three major categories: (a) student profile (personal data), (b) state of student's progress and (c) trace of student's activity.

Furthermore, many attempts have been made in order to model the learner data in a more formal way and have been resulted in a number of standards, such as PAPI (Public and Private Information) [11], IMS LIP (Learner Information Package) [12], eduPerson [13], Dolog LP [14,15], FOAF (Friend of a Friend) [16] etc. Even if these models share a set of common learner characteristics, they vary on their main purpose and the way in which a system may use their embedded information. It is a usual practice to produce a learner profile for a learner system combining different learner standards and profiting from their unique benefits.

It is a common belief that PAPI and LIP are the most significant and important among the known standards due to their extended use and the benefits that they provide when used jointly. In [17], the presentation of the main characteristics of the aforementioned standards and the comparison of them denote the importance and the completeness of PAPI and LIP.

Nevertheless, both standards have some shortcomings. For example, the IMS LIP standard is based on the notion of a classic CV, while the PAPI standard considers student's performance as the most important information. However, in the context of our work, we took them into consideration and incorporated some of their basic notions to our proposed student model so that it conforms to these international standards.

Regarding the aforementioned approaches, our idea adopts the basic principles of the student model described in [10] suitably adapted to the needs and characteristics of an adult learner (e.g. time for study, previous experience, educational level and learning goals) in a distance learning educational framework. For example, in [4] the author does not include student's learning style in their model, while in [8] they take into consideration only student's performance and his/her interaction with the system. Furthermore, in [6] the authors do not include student's preferences, learning goals and motivation state in their model.

3 Description of the Student Model

In this section, we shall give a brief description of the basic components of the proposed approach and some of the basic characteristics of the model.

3.1 Learning Style

One of the most important components of the student model in an ITS is the personal learning style of the learner. The term "*learning style*" is used to describe the individual differences in the learning process. It is based on the assumption that each person has a unique and distinctive way to learn, i.e. to collect, process and organize information [18].

Among the models and theories presented in the literature, we have adopted the Felder-Silverman theory for student modeling. Most existing learning-style based

theories classify students into few coarse grained groups, whereas Felder and Silverman describe the learning styles of a student in more detail, distinguishing between preferences on four dimensions [19].

According to the Felder-Silverman model, the learning types are categorized in the following four dimensions: (a) *active/reflective*, (b) *sensing/intuitive*, (c) *visual/verbal*, and (d) *sequential/global*.

3.2 Modeling Approach

The most common representation of a student model is the *overlay model*. The overlay model represents a learner's knowledge as a subset of the domain knowledge (expert's knowledge). Therefore, the system provides the learner with educational material until learner's knowledge coincides with the expert's knowledge [20]. Another approach which is widely used is the *buggy model*. Systems that use such models record and represent the most common/frequent mistakes made by learners based on statistics. Finally, one widely adopted approach for student modeling is the use of *stereotypes* [21]. New learners are classified into distinct categories and the system adjusts its performance based on the category assigned to the learner.

In the context of our work, we adopt a combination of the stereotype and overlay techniques. A fully stereotype-based model was excluded as a choice because (1) the initialization of the system derived from students descriptions or questionnaires may not be accurate for every knowledge domain and (2) the system would adapt to the learner's needs very slowly. So we developed a model where some attributes of the student profile (e.g. previous knowledge, experience in a specific knowledge domain) are initialized based on a stereotype. In addition, dynamic attributes related to the learning process are represented with an overlay model. After the initialization phase, the profile is dynamically modified, as the overlay model is updated with the information gathered by the interaction between learner and system.

3.3 Basic Characteristics of the Students

The users' classification in categories, called *stereotypes* constitutes a technique that has been widely used in user modeling systems. Stereotypes can be specified according to the following criteria: age, gender, educational level, working experience etc.

An empirical study was conducted by the Educational Content, Methodology and Technology Laboratory¹ among students of the Hellenic Open University² (HOU) in order to extract the basic characteristics and formulate the corresponding stereotypes of the student. The HOU was founded in 1992 and provides open distance learning at both undergraduate and postgraduate level.

¹<http://eeyem.eap.gr>

²http://www.eap.gr/index_en.php

Table 1. Students' characteristics and their corresponding stereotypes

| Characteristic | Stereotype |
|-----------------------|---|
| Learning style | active/reflective – sensing/intuitive – visual/verbal – sequential/global |
| Use of technology | adaptable - adaptive |
| Computer literacy | novice – beginner – advanced |
| Previous experience | novice – beginner – advanced |
| Time for study | no time – little – much |
| Reasons for education | career development –career change – self improvement |
| Academic literacy | poor – good – excellent |
| Socialization style | lonely – collaborative |

The students who participated in the study were chosen based on their different characteristics such as different gender, age, educational background and current course. The study included (a) personal interviews with the students and (b) observation of the face to face meetings, by a social scientist. In particular, from the 13 students who were interviewed, 5 were male and 8 female, 8 of them pursued undergraduate studies and 5 postgraduate, 10 of them are working in the public sector, 2 are unemployed and 1 is working as a freelancer. Table 1 summarizes the characteristics and the stereotypes that came up from the empirical study.

Besides the modeling approach that defines the specialization of the model, a few more model characteristics have been taken into account: (a) it is a *dynamic model* that can change over time as the system collects information about the individual, (b) it is a *long-term model* that keeps generalized information regarding the user-system interactions and (c) it is a combination of “*active*” and “*passive*” user model, i.e. in the beginning the user provides directly information about him/her and then the system indirectly collects more information.

4 Student Model Ontology

In this section, we thoroughly describe the *Student Model* ontology that has been developed in order to capture the main concepts presented in Section 3. The focus of our attempt is not restricted on modeling the static profile of the user, but encompasses both permanent and dynamic characteristics. Moreover, the developed ontology complies partly with well-known standards for student modeling, i.e. IEEE PAPI Learner [11] and IMS Learner Information Package (LIP) [12].

In order to build the ontology, we followed a widely-adopted methodology, proposed in [22]. As far as its formal representation is concerned, we adopted the Web Ontology Language (OWL), which is a W3C standard. More specifically, our ontology falls into the OWL DL sublanguage, which provides the maximum expressiveness, while maintains computational completeness (all the conclusions are measurable and all calculations are terminated in finite time). The development process of the ontology was accomplished with the aid of Protégé³ tool.

³<http://protege.stanford.edu/>

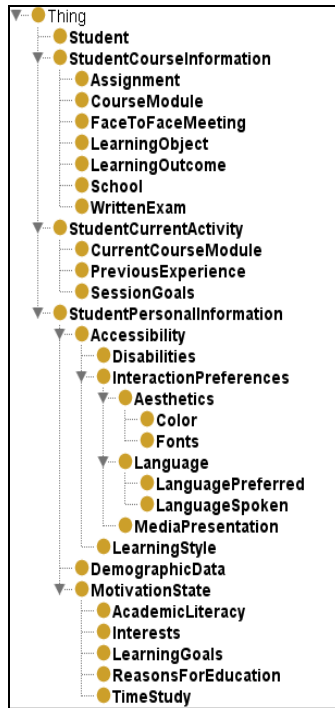


Fig. 1. The Student Model ontology as displayed in Protégé

In the proposed ontology *Student Model*, we define a set of four upper level classes, namely *Student*, *StudentCourseInformation*, *StudentCurrentActivity* and *StudentPersonalInformation*. The class hierarchy of the ontology, as displayed in Protégé, is depicted in Figure 1. The class *Student* represents any student. The *StudentCourseInformation* class comprises information relevant to the student’s performance during the overall educational process and has a number of subclasses that are listed below, together with a brief explanation:

1. *Assignment* - the written assignments that the student has to submit during a course module
2. *CourseModule* - the course modules of the course program
3. *FaceToFaceMeeting* - face to face meetings during a course module
4. *LearningObject* - the learning objects that the student has been taught
5. *LearningOutcome* - the learning outcomes succeeded by the student as indicated by the learning objects
6. *School* - the school for which the student is registered
7. *WrittenExams* – the written exams that the student has to participate during a course module

In order to capture any detail in terms of student’s activity for the current academic year, we define the class *StudentCurrentActivity*. Student’s activity for the current academic year can be specified by the following three axes: (i) current chosen course

modules (class *CurrentCourseModule*), (ii) the experience on a specific course module that the student has previously gained (class *PreviousExperience*) and (iii) the student's goals on a specific course module (class *SessionGoals*).

The more compact class in the proposed ontology, *StudentPersonalInformation*, is defined so as to represent mostly static and permanent student information, describing not only simple data, like demographic data, but more complex characteristics that concern student's interaction with the e-learning system. Table 2 lists the subclasses that exist under the upper level class *StudentPersonalInformation*. The table gives also a brief description of the entities that are represented by these classes.

Table 2. Description of the class *StudentPersonalInformation*

| Class Name | Class Description |
|------------------------|---|
| Accessibility | The overall set of features that characterizes the student's behavior during his interaction with the e-learning system |
| Disabilities | The set of student's disabilities that could affects the educational process |
| DemographicData | Student's demographic data |
| InteractionPreferences | Student's preferences regarding interaction with the e-learning system |
| MediaPresentation | Student's preferences regarding the presentation of learning objects |
| Language | Student's preferences regarding the language of the learning objects |
| LanguageSpoken | Student's native languages |
| LanguagePreferred | Language that the student prefers for the presentation of learning objects |
| Aesthetics | Aesthetic factors such as the use of highly interactive sensory and visual communication |
| Color | Student's preferences regarding the coloring scheme of learning system's environment |
| Fonts | Student's preferences regarding the fonts used by the learning system's environment |
| LearningStyle | Student's learning style - This class will be further divided to the sub classes according to the Felder-Silverman theory |
| MotivationState | Student's motivation during the educational process |
| LearningGoals | Overall goals set by the student |
| ReasonsForEducation | The reasons why the student desires to engage in the educational process |
| AcademicLiteracy | Student's previous formal educational experiences |
| Interests | Student's interests |
| TimeStudy | The average time per day that the student can use for studying |

On the other hand, relationships between instances (members of classes) are modeled as object properties. In this context we define a set of object properties (mostly of the *hasA* kind). This kind of object property is used for expressing the association of the aforementioned characteristics with students. It links an instance of the class *Student* to instances of classes that reflects student characteristics such as *StudentCourseInformation*, *StudentCurrentActivity*, *PreviousExperience*, *Demographic*

Data, *StudentPersonalInformation*, *Learningstyle*, *InteractionPreferences*, *Disabilities*, *Interests*, *TimeStudy*, *ReasonsForEducation* and *AcademicLiteracy*. Furthermore, datatype properties, that link individuals to data values, have been set in order to define more effectively the classes.

In addition, as foresaid, we have adopted a scheme inspired by the Felder-Silverman Learning Style Model [23], in order to infer the student’s learning style. The eight proposed learning styles are captured as individuals of the class *LearningStyle*, in the *Student Model* ontology (see Section 3.1 for a learning styles description).

The proposed ontology has been enriched with a set of rules in order to enable inference mechanisms (i.e. reasoner) to automatically classify the students into different stereotypic profiles (Table 1). As foresaid, these rules and the stereotypic profiles have resulted from an empirical research on a sample of adult learners. All the rules are expressed in Semantic Web Rule Language⁴ (SWRL) and a subset are given in Table 3. For example, the first rule in the table indicates that “*if the student is female, over 50 years old and doesn’t have a bachelor degree, then has little familiarity with computers*”.

Table 3. SWRL Rules

| # | Rule Body | Rule Head |
|---|--|---------------------------------------|
| 1 | IF AND x is-a <i>DemographicData</i> AND y is-a <i>Student</i> AND y hasDemographics x AND x age z AND x educational_level "secondary" AND x gender "female" AND z greaterThan "50"^^integer | x computer_literacy "beginner" |
| 2 | IF AND y is-a <i>DemographicData</i> AND x is-a <i>Student</i> AND z is-a <i>TimeStudy</i> AND x hasDemographics y AND y age w AND y educational_level "secondary" AND y gender "female" AND z time_for_study "much" AND w greaterThan "50"^^integer | x hasTime z |
| 3 | IF AND x is-a <i>DemographicData</i> AND y is-a <i>Student</i> AND z is-a <i>ReasonsForEducation</i> AND y hasDemographics x AND x age w AND x educational_level "secondary" AND x gender "female" AND z reasons_for_education "general knowledge" AND w greaterThan "50"^^integer | y hasReasons z |

⁴ <http://www.w3.org/Submission/SWRL/>

5 Conclusions and Future Work

We proposed in this paper an ontology-based approach to model student profiles especially for distance learning students. The student profile ontology we developed can be used as an integral ITS module, while it can be easily accessed from a web-based application. The proposed approach collects the characteristics of an adult student which are considered important for an ITS in order to be fully adapted to the needs of the learner. This model is a combination of international standards in user modeling and the results of an empirical study on a group of HOU students. One of the main advantages of the proposed model is the integration of semantic rules. These rules combined with inference mechanisms classify learners into stereotypic profiles which are already incorporated in the ontology and thus produce additional knowledge. The most challenging part of our research has been the selection of the characteristics to be included in the ontology. We further plan to add more rules in the student model ontology in order to cover as many stereotypic profiles as possible. Moreover, as a future work is the implementation of a web application which will communicate directly with the *Student Model* ontology and allow users to create their own profile.

Acknowledgment. This research described in this paper was partly funded by the National Strategic Reference Framework programme 2007-2013, project MIS 296121“Hellenic Open University”.

References

1. Beck, J., Stern, M., Haugsjaa, E.: Applications of AI in Education. ACM, Crossroads (1996)
2. Shimic, G., Gasevic, D., Devedzic, V.: Classroom for the Semantic Web. In: Intelligent Learning Infrastructure for Knowledge Intensive Organizations: A Semantic Web Perspective, pp. 251—283 (2005)
3. Polson, M.C., Richardson, J.J. (eds.): Foundations of intelligent tutoring systems. Routledge (1988)
4. Paneva, D.: Use of Ontology-based Student model in Semantic-oriented Access to the Knowledge in Digital Libraries. In: Proceedings of the Open Workshop Semantic Web and Knowledge Technologies Applications, Bulgaria, pp. 31–41 (2006)
5. Chen, W., Mizoguchi, R.: Communication Content Ontology for Learner Model Agent in Multi-agent Architecture. In: Workshop on Ontologies for Intelligent Educational Systems (AI-ED 1999), France (1999)
6. Pramitasari, L., Hidayanto, N.A., Aminah, S., Krisnadhi, A.A., Ramadhanie, A.M.: Development of Student Model Ontology for Personalization in an E-Learning System based on Semantic Web. In: International Conference on Advanced Computer Science and Information Systems (ICACSIS 2009), Indonesia, December 7-8, pp. 434–439 (2009)
7. Jeremic, J., Devedzic, V.: Design Pattern ITS: Student Model Implementation. In: Proceedings of the IEEE International Conference on Advanced Learning Technologies (ICALT 2004), pp. 864–865 (2004)
8. Henze, N., Dolog, P., Nejdil, W.: Reasoning and Ontologies for Personalized E-Learning in the Semantic Web. Educational Technology and Society 7, 82–97 (2004)

9. Muñoz, L.S., Palazzo, J., Oliveira, M.: Applying Semantic Web Technologies to Achieve Personalization and Reuse. In: Proceedings of the SWEL Workshop at Adaptive Hypermedia, pp. 348–353 (2004)
10. Clemente, J., Ramirez, J., de Antonio, A.: A proposal for student modeling based on ontologies and diagnosis rules. *Expert Systems with Applications Journal*, 8066–8078 (2011)
11. LTSC Learner Model Working Group of the IEEE (2000) IEEE p1484.2/d7, 2000-11-28 Draft Standard for Learning Technology - Public and Private Information (PAPI) for Learners (PAPI Learner), Technical report (2000)
12. Smythe, C., Tansey, F., Robson, R.: IMS Learner Information Package Information Model Specification. Technical report (2001)
13. EduPerson Specification. Document: Internet2-mace-dir-eduPerson-200312. Copyright © 2003 by Internet2 and/or the respective authors. Internet2 Middleware Architecture Committee for Education, Directory Working Group (December 2003)
14. Dolog, P., Nejdli, W.: Challenges and Benefits of the Semantic Web for User Modelling (2003)
15. Dolog, P.: Identifying Relevant Fragments of Learner Profile on the Semantic Web. In: Proceedings of the SWEL 2004 Workshop at ISWC 2004 Conference, Hiroshima, Japan (2004)
16. Brickley, D., Miller, L.: FOAF Vocabulary Specification -Namespace Document - ('Pages about Things' Edition) (July 27, 2005)
17. Ounnas, A., Davis, H.C., Millard, D.E.: Towards semantic group formation. In: The 7th IEEE International Conference on Advanced Learning Technologies (ICALT 2007), Niigata, Japan (2007)
18. Jones, D.C., Mungai, D.: Technology-enabled teaching for maximum learning. *International Journal of Learning* 10, 3491–3501 (2003)
19. Felder, R.M., Silverman, L.: Learning and Teaching Styles in Engineering Education. *Engineering Education Journal* 78, 674–681 (1988)
20. Conlan, O., O'Keeffe, I., Tallon, S.: Combining Adaptive Hypermedia Techniques and Ontology Reasoning to Produce Dynamic Personalized News Services. In: Wade, V.P., Ashman, H., Smyth, B. (eds.) AH 2006. LNCS, vol. 4018, pp. 81–90. Springer, Heidelberg (2006)
21. Kay, J.: Stereotypes, Student Models and Scrutability. In: Gauthier, G., VanLehn, K., Frasson, C. (eds.) ITS 2000. LNCS, vol. 1839, pp. 19–30. Springer, Heidelberg (2000)
22. Noy, N., McGuiness, D.: Ontology Development 101: A Guide to Creating Your First Ontology, Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMi-2001-0880 (2001)
23. Graf, S., Viola, S.R., Kinshuk, Leo, T.: Representative Characteristics of Felder-Silverman Learning Styles: An Empirical Model. In: Proceedings of the International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2006), Spain, pp. 235–242 (2006)