Mining Student Learning Behavior and Self-assessment for Adaptive Learning Management System

Konstantina Moutafi[1], Paraskevi Vergeti[1], Christos Alexakos[1], Christos Dimitrakopoulos[1], Konstantinos Giotopoulos[1], Hera Antonopoulou[2], and Spiros Likothanassis[1]

[1] Pattern Recognition Lab., Dept. of Computer Engineering and Informatics, University of Patras, Greece
{moutafi,vergetip,alexakos,dimitrakop, kgiotop,likothan}@ceid.upatras.gr
[2] Technological Educational Institute of Patras, Greece
santonop@teipat.gr

Abstract. The specific contribution aims to provide a web-based adaptive Learning Management System (LMS), named EVMATHEIA, which integrates specific innovative fundamental aspects of Student Learning Style and Intelligent Self-Assessment Mechanisms. More specifically the proposed adaptive system encapsulates an integrated student model that facilitates the decision about the learning style of the student monitoring his/her behavior. Furthermore, the platform utilizes semantic modeling techniques for the representation of the knowledge, semantically annotated educational material and an intelligent mechanism for the self-assessment and recommendation process.

Keywords: learning management systems, adaptive e-learning, user modeling, personalized learning, learning styles, assessment.

1 Introduction

It is commonly known that e-learning environments are widely spread in all levels of human education. The specific aspect has imposed scientific research to enhance the efforts in the field of adaptive and intelligent learning platforms aiming to contribute significantly in the provision of high quality services towards the end users of e- learning systems. Over the last years scientific research aim to provide integrated systems that are intelligent and adaptive, and special attention has been given to specific key features of the learning style of the student and the self-assessment mechanisms.

The proposed EVMATHEIA platform aims to the provision of personalized learning adapted to student's receptivity. Its main target is to deliver knowledge, through a web platform, to individuals based on their interactions with the system, reducing the interference of the tutor and the collaboration with other students. System's key characteristic is the identification of student's learning style providing educational content aiming to a faster study and easier learning. Furthermore, it provides a mechanism detecting student knowledge's weakness through intelligent evaluation of self-assessment questionnaires and stimulates him/her to study specific additional material.

L. Iliadis, H. Papadopoulos, and C. Jayne (Eds.): EANN 2013, Part II, CCIS 384, pp. 70–79, 2013.

In order to realize these aspects, EVMATHEIA encapsulates a *student model* storing information about student's preference, knowledge and learning style. The last denotes the way a student better perceives the provided knowledge. The student model is accompanied by a monitoring mechanism of the student behavior and decides his/her learning style. System's *knowledge* (subject of education) is represented on a semantic model setting as basic knowledge unit named the *concept*. The result is an ontological network of concepts that depicts the various relations between them (i.e. if a concept prerequisites another one). Another key element is the semantically annotated *educational material* which contains the appropriate information for the personalized and recommendation mechanisms. Finally, system affords an *intelligent mechanism* which interacts with the student during the *self-assessment process*, deciding the level of the accumulated knowledge for each concept. The results are processed by a recommendation mechanism that takes into account the learning style and the education material providing suggestions to the student for further reading.

Section 2 presents a brief reference to the work done the past years on the relative fields of learning style and learner assessment in personalized e-learning systems. An overview of the proposed EVMATHEIA architecture is depicted in Section 3. Sections 4, 5 and 6 present the realization of student modeling, educational material personalization and student assessment in the system. Finally, some conclusion is given in Section 7. Work presented in this paper has been partially developed in the framework of the project LOC PRO II-Support and Promotion of Local Products and SMEs through ICT, Operational Programme Greece – Italy 2007-2013, s.c.: I1.12.01.

2 Learning Style and Student Assessment in LMS

The integration of learning styles in the adaptation process of Learning Management Systems (LMS) has become a major field of study the last years. Many models for learning styles are proposed and many techniques have been used to infer the learning style from the behavior of the student. One of the most widely used is the Felder-Silverman Learning Style Model (FSLSM) [1] which is proposed for engineering students. According to this model (which Felder revised in 2002) a student is classified according to his/her preference for one of the categories in each of the four learning style dimensions [2]: a) *Sensing* (concrete, practical, oriented toward facts and procedures) or *Intuitive* (conceptual, innovative, and oriented towards theories and meanings) depending on the type of information he/she prefers to perceive. b) *Visual* (prefer visual representations of presented material: pictures, diagrams, flow charts) *or Verbal* (prefer written and spoken explanations) depending on the way he/she prefers to receive information. c) *Active* (learn by trying things out, working with others) or *Reflective* (learn by thinking things through, working alone) depending on the way he/she prefers to process information. d) *Sequential* (linear, orderly, learn in small incremental steps) or *Global* (holistic, systems thinkers, learn in large leaps) learners (learn by thinking things through, working alone) depending on the way he/she prefers to process information. For the assessment of the preferences on the four dimensions of FSLSM, Felder and Soloman developed the Index of Learning

Styles (ILS), a questionnaire consisting of 44 questions, 11 for each learning style dimension [3]. According to the answers, student receives a score for each dimension: a) *balance* b) *moderate* or c) *strong* according to its preference for one of the two categories.

A crucial issue in automatic student modeling is to determine which student's behavior is indicative about his/her learning style. Graf et al [4] and García et al [5] are utilizing the FSLSM as their basic learning style model. Popescu at [6] proposes a combination of learning styles models. These approaches describe a great number of navigational, temporal and performance indicators identifying the learning style preferences according to FSLSM and propose thresholds that are necessary to classify the behavior of students. Furthermore, experimental studies were conducted investigating the behavior of students with different learning styles in online courses [7, 8].Useful conclusions derived from these studies contributed to the selection of patterns from the literature for online learning.

A key aspect of adaptive LMS is related with the intelligent self-assessment mechanism. Several efforts have been made towards this direction utilizing Computational Intelligence Techniques to support self-assessment in LMS [9]. The main idea behind this approach is to use Bayesian Networks and Genetic Algorithms simplifying the assessment by predicting student's answers [10, 11]. Several issues from the aforementioned efforts resulted in the present contribution. It is important also to note the efforts in personalized e-learning system with self-regulated learning assisted mechanisms to help learners promote their self-regulated learning abilities [12].

3 System Overview

Figure 1 depicts the main architectural components used for the realization of the main concepts in the EVMATHEIA approach. The core platform is a full-functional web-based application providing the main e-learning services to the users (students and tutors) such as registration, course structure, presentation and management.

Student model is a database segment that stores all the information needed from the system regarding student's characteristics and preferences. *Assessment questions* and *Educational Material* contain the questions related to each knowledge concept and the digital real educational material respectively. Two ontologies are used in order to define the knowledge and the educational material. *Knowledge ontology* contains the provided concepts, defining the student's knowledge and the relationship between educational material and the knowledge's concepts. *Educational Material Annotation Ontology* provides the annotation layer to the stored material.

Three supported modules execute the additional functionalities of the proposed system. *Student modeling module* collects information about student's behavior and updates student model. *Student assessment module* assists the student to the self-assessment process by providing the appropriate questions and evaluates the results. It also evaluates the answers and updates the student's accumulated knowledge. In addition, it forwards the results to Education Material Selection Module which by utilizing the information of *Educational Material Annotation Ontology* recommends additional material to the student. The third module is also responsible to provide student with the appropriate material according to his/her learning style.

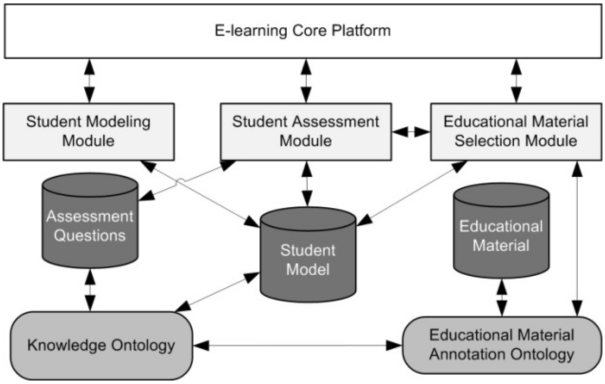

Fig. 1. Proposed system architecture

4 Student Model

Student model contains the information about the student that is needed by EVMATHEIA modules in order to deliver the desired level of personalization. The basic groups of information stored in the student model are: a) Student's *Personal Information* saves data such as name, surname, age, communication details, etc. b) *Language Knowledge* depicts student's level of knowledge (average, good, very good) of various languages. *c) Student's knowledge* is an overlay of system's knowledge which depicts the level of student's understanding on each concept. *d) Accessibility* contains special information regarding student's capability to access and study digital educational material. This group stores information such as visual, hearing, physical and cognitive disabilities. *e) Learning Style* includes student's score in every dimension of FSLSM.

4.1 Student Learning Style Modeling

The identification of student's learning style plays a key role in the EVMATHEIA system and as a result a modeling mechanism is developed combining the basic educational and psychological concepts around this issue. The modeling procedure is conducted in two phases: a) an initial approximation from student's answers in the classic ILS questionnaire and b) the continuous monitoring of student's behavior in the system and the re-calculation of student's learning style.

The indicative behavior patterns of the student's learning style preferences are based on the literature regarding the FSLSM and the features of our system. An online course in the system consists of sections and each section presents learning objects

(LOs) for a set of concepts. For each section educational material is provided that contains content in different types (text, video, sound, image, etc.), exercises and examples. At the end of each section the knowledge can be assessed through self-evaluation tests. Students navigate in the course through the course outline, the navigation menu and Next-Previous buttons. Thus the set of monitored behavior patterns consists of navigational, temporal and performance indicators correlated to the above features.

The system uses thirty (30) behavior patterns that have been selected after analyzing the approaches of [4], [5] and [6]. The main selection criterion was the monitoring feasibility of a pattern in a system without collaborating functionalities. Some examples of the behavior patterns are: percentage of time spent on examples, relative time spent by the student on content type text versus the relative study time for content type text and relative number of visits of content type video versus the total relative number of content type video available in the course. For each pattern, two thresholds define three ranges of values that conclude pattern's score to low, medium and high. Furthermore each pattern is associated with weights that represent how indicative is the respective behavior in the online course on identifying student's learning preference. These weights mainly derived from [7] and [8].

A set of N relevant patterns (P_{ij}) has been assigned in each dimension (D_j) of FSLSM. Since the two categories related to each dimension of FSLSM are opposed, if a high value of a pattern is associated with one category of a specific dimension, the low value of the same pattern is associated with the other category of the same dimension. Therefore calculations can be done only for one category in each dimension. Equation 1 defines the calculated score S_j that corresponds to the dimension D_j of FSLSM.

$$S_j = \frac{\sum_{1 \le i \le N} w_{ij} * p_{ij}}{\sum_{1 \le i \le N} w_{ij}} \tag{1}$$

The p_{ij} is the numerical value of i-th pattern P_{ij} of a particular category of the dimension D_j. The numerical values are (-) 1 for low, (-) 2 for medium and (-) 3 for high values. The positive values are if the pattern corresponds to the particular category and negative values for the opposite one. Pattern's weights w_{ij} are enumerated with 0.2, 0.5 and 1, indicating low, medium and high importance respectively. The calculated S_j is a number ranging from -3 to 3. According to absolute value of S_j, the student is classified for the pointed category as balanced ($0 \le S_j \le 1$), moderate ($1 < S_j \le 2$) or strong ($2 < S_j \le 3$) preference.

5 Knowledge and Educational Material

The proposed system utilizes semantic annotation defining knowledge and educational material, permitting their combination with the student model in order to infer the personalized presentation of the educational material to the student.

5.1 Knowledge Representation

An ontological approach is used for the representation of the knowledge structure, which is simply based on concepts with given relations between them. The ontology used derives from the education domain ontology proposed in [13], in which three kinds of relations are given: a) HasPart (an inclusion relation), b) IsRequiredBy (an order relation) and c) SuggestedOrder (a 'weak' order relation). These relations form a graph, where the nodes are the concepts and the edges are the relations. Figure 2 depicts an example of the knowledge semantic representation graph.

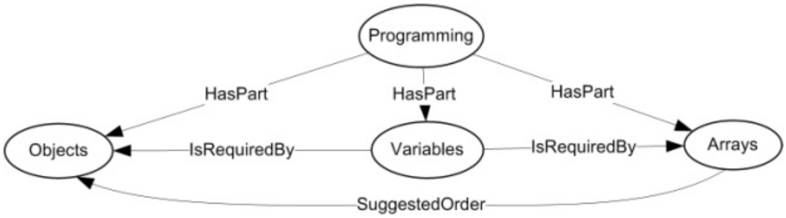

Fig. 2. Example portion of knowledge representation

5.2 Educational Material Annotation

EVMATHEIA educational material is defined/annotated using the OWL LOM ontology proposed by Hartonas C. [14] utilizing the IEEE Learning Object Metadata (IEEE LOM) [15] standard. LOM is the most common standard used for the description of learning material. The structure of LOM consistes of nine categories of the educational material: general, lifecycle, meta-metadata, technical, educational, rights, relation, annotation and classification. The schema used in the system is a subset of LOM, consisting of metadata mainly from the educational category. The metadata entities that support the adaptation mechanism are the following: language, format, typical age range, semantic density, interactivity type, typical learning time and associated concept. *Language* is a general characteristic referring to the language or languages that are used in the educational material. *Format* typically is the technical data type of the learning material, but here we are borrowing the set of values defined in [16]. These values are text, image, streaming media and application. *Typical age range* of the intended users is an educational characteristic to match the age of the learner. *Semantic density* of the educational material is the degree of conciseness and its value space is: very low, low, medium, high and very high. *Interactivity type* according to LOM takes values: active, expositive and mixed. The *typical learning time* of the material denotes the average time a student needs to study it. This characteristic has link to the active/reflective characteristic of the learning style of the student. Therefore, for a student with active learning style will be more accurate to propose an 'active' material, e.g. a questionnaire. *Associated concept* defines the relationship of the particular material to the relative concept in the knowledge ontology graph. Each material describes at least one concept.

5.3 Student Personalized Educational Material

The annotation of educational material in conjunction with student model permits the system to decide the best material for each student. The decision mechanism rates the available material based on a set of rules from the definition of FSLSM dimensions and the student's preferences. First the system chooses the relative, to a particular concept, educational material taking into account student's previous knowledge and his/her language. Then, for each retrieved educational material k the Sc_k score is calculated by the equation 2.

$$Sc_k = \frac{\sum_{1\le i\le \max\_\text{associations}} LS_i * ws_i}{\sum_{\substack{1\le i\le \max\_\text{associations} \\ LS_i \neq 0}} ws_i} + Ln_s + A_s \quad (2)$$

The LS_i values define the association of educational material characteristics to student's FSLSM dimensions score. LS_1 depicts the Active/Reflective dimension, the LS_2 the Sensing/Intuitive and the LS_3 the Visual/Verbal. The respective weights ws_i denote student's trend for a particular axis, i.e. strong verbal. Weights vales are 0, 0.5 and 1 for "well balanced", "moderate" and "very strong" score respectively.

For LS_1, value equal to 1 is assigned in case student's score to Active/Reflective dimension is moderate or above in the corresponding axis of material's *Interactivity type* value (active or expositive). In other case 0 is assigned. In case of the educational material's definition as the value 0.5 is assigned. For LS_2, value equal to 1 is assigned to this parameter when following combinations are true: a) student tends to sensing, material's interactivity type is active and its semantic density is low or very low and b) student tends to intuitive, material's interactivity type is expositive and it's semantic density is high or very high. Otherwise value equals to 0. The conditions are derived from the hypothesis that intuitive students prefer non active and of high semantic density educational material, whereas sensing learners prefer active and of low semantic density material. Finally, LS_3 is graded with 1 when the format of the educational material is corresponding to the student's learning style in the Visual/Verbal dimension. When a student is verbal then prefers text format. In the opposite case, when a student is visual prefers image, streaming media or application formatted material.

Ln_s parameter takes values of 0.5, 1 and 2 regarding student's level of knowledge (average, good, very good) in the material's language. The value of A_s is set to 0.5 in case student's age is contained to material's *typical age range* and to 0 otherwise. Finally, the educational material with the higher Sc_k is proposed to the student.

6 Student Assessment and Recommendation

The self-assessment mechanism aims to precisely identify student's acquired knowledge and to find the concepts that he/she has weakly learned. The proposed assessment procedure consists of a set of questions, related with one of the section's knowledge concepts, at the end of each section. The assessment questionnaire is created on the fly from a pool of questions for each concept. The questions are selected randomly, trying to avoid the repetition when a student conducts the questionnaire several times. The algorithm identifies the level of student's understanding in each concept. The idea is to provide him/her questions and predict the answer; if the answers are similar with the prediction then the algorithm concludes regarding how well the student understands a concept. In this concept the algorithm tries to be indifferent to student's random answers that are correct and to give a precise result, taking into account the correlations between concepts and the fact that the more "sincere" the student is, the faster he/she will complete the self-assessment.

6.1 Question Answer Prediction - Simple Majority Voting

For a particular student k, the algorithm sets the state of a question to +1 if the question has been answered correctly and to -1 if the question has been answered incorrectly. Next, the algorithm selects an unanswered question i by the student k repeatedly, and sets its state s_i to -1 or +1 according to the following rule:

$$S_i = sgn\left(\sum_{1\leq j\leq n_i} S_j - \theta\right) \quad (3)$$

Where n_i is the number of questions answered from the student k either correctly or incorrectly and belong to the same concept that the unanswered question belongs, s_j is the state of brother j and θ is an activation threshold. The questions that belong to the same concept with question i and have been answered from the student k previously are considered as "brothers" of question i. The right hand side of this equation computes the sum of the states of the brothers of question *i* and sets its state to +1 if the sum is > θ, and to −1 otherwise. In this implementation the activation threshold was chosen for each student and concept independently and was set at an integer value that yielded the best F-measurement score in cross-validation tests. Finally, the algorithm predicts the answer to a question to be "Correct" by the student k if the state computed was set to +1.

Furthermore, the algorithm incorporates the information of linked concepts in order to improve the prediction performance of the algorithm. For each unanswered question q_i, the state of a particular question q_j is set to +1, if the question is answered correctly by the student k and belongs to the same concept that q_i belongs **or** if the question is answered correctly by the student k and belongs to a concept that is linked with the concept that q_i belongs. Otherwise the state of the unanswered question q_i is set to −1 (in case the question is answered incorrectly and belongs to either the concept of q_i or to a concept that is linked to that of q_i). Next, the unanswered question q_i is assigned a state of +1 or −1 using the same rule as before. Now the brothers (n_i) of the unanswered question q_i for student k are chosen as the questions that belong to the same concept with q_i or to a concept linked to that of q_i, and have been answered from student k previously either correctly or incorrectly. Finally, the unanswered question q_i is predicted to be answered correctly by the student k if its state was set to +1. The activation threshold was optimized for every student/concept separately by using cross-validation. The activation rule in the case of the linked concepts is modified as follows:

$$S_i = sgn\left(w_n\left(\sum_{1\le j\le n_i} S_j - \theta_n\right) + w_m\left(\sum_{1\le j\le m_i} S_j - \theta_k\right)\right) \quad (4)$$

where n_i are the questions that belong to the same concept with q_i and m_i the questions that are assigned to a concept linked to that of q_i. We suppose that the concept of the question is more significant from the concepts linked to the question. As a result we define two weights w_n=0.8 and w_m=0.2 that represent the significance of the question's concept and of the linked concepts.

6.2 Self-assessment Result and Recommendation

The aforementioned algorithm is used in the EVMATHEIA system in order to ensure a valid result of student's assessment procedure. If the student has answered a considerable amount of questions then his/her answers are used to predict his/her answers to the remaining unanswered questions. If the remaining questions have been predicted to be answered correctly by the student k in a high degree then the assessment for the particular concept stops and the concept is considered as adapted by the student in a high (80%) or fundamental (100%) extend. If the remaining questions' prediction cannot lead to a certain result (same or near to same proportion of correctly and incorrectly answers predicted) or the remaining questions have been predicted to be answered incorrectly in a high degree, the system continues to assess the student with questions until a valid concept adaptation prediction is detected or the available questions are finished. If the questions finish without assimilation halt, then the student is assigned a below 80% knowledge extend.

Following this approach, the system predicts the level of knowledge for each assessed concept. After this, the system updates the student model using the procedure depicted in section 4.1 and the student's score for each knowledge concept. For the concepts ranked below 80% score, the system recommends material using the rules defined in section 5.3 and rejecting educational material that he/she has studied.

7 Conclusion

The proposed EVMATHEIA platform provides a new perspective in the provision of personalized learning adapted to student's receptivity. The main focus of the current work is to monitor the interaction of each individual student and reduce the interference of the tutor and the collaboration of the students. A self-managed learning process has been presented aiming to deliver knowledge in a personalized approach.

The key features that were exploited are the identification of the learning style of the student aiming to minimize the time needed for learning process. Furthermore, new and emerging concepts were presented in the specific field aiming to enhance the research in the specific filed and to open the path for future related work.

Last, but not least, it is important to emphasize in the intelligent assessment and recommendation mechanism that utilizes a new algorithm for the identification the level of student's understanding in each concept.

References

1. Felder, R.M., Silverman, L.K.: Learning and Teaching Styles in Engineering Education. Engineering Education 78(7), 674–681 (1988)
2. Felder, R.M., Spurlin, J.E.: Applications, Reliability, and Validity of the Index of Learning Styles. Intl. Journal of Engineering Education 21(1), 103–112 (2005)
3. Felder, R.M., Soloman, B.A.: Index of Learning Styles Questionnaire, North Carolina State University (2001), `http://www4.ncsu.edu/unity/lockers/users/f/felder/public/ILSpage.html` (accessed at April 2013)
4. Graf, S., Kinshuk, L.T.-C.: Supporting Teachers in Identifying Students' Learning Styles in Learning Management Systems: An Automatic Student Modelling Approach. Educational Technology & Society 12(4), 3–14 (2009)
5. García, P., Amandi, A., Schiaffino, S., Campo, M.: Evaluating Bayesian networks' precision for detecting students' learning styles. Computers & Education 49(3), 794–808 (2007)
6. Popescu, E.: Diagnosing Students' Learning Style in an Educational Hypermedia System. In: Mourlas, C., Tsianos, N., Germanakos, P. (eds.) Cognitive and Emotional Processes in Web-based Education: Integrating Human Factors and Personalization. Advances in Web-Based Learning Book Series, pp. 187–208. IGI Global (2009)
7. Graf, S., Kinshuk: Analysing the Behaviour of Students in Learning Management Systems. In: Wallace, M., Angelides, M.C., Mylonas, P. (eds.) Advances in Semantic Media Adaptation and Personalization. SCI, vol. 93, pp. 53–73. Springer, Heidelberg (2008)
8. Popescu, E.: Learning Styles and Behavioral Differences in Web-based Learning Settings. In: 9th IEEE International Conference on Advanced Learning Technologies, pp. 446–450. IEEE Press, New york (2009)
9. Giotopoulos, K., Alexakos, C., Beligiannis, G., Stefani, A.: Bringing AI to E-learning: the case of a modular, highly adaptive system. International Journal of Information and Communication Technology Education (IJICTE) 6(2), 24–35 (2010)
10. Mangalwede, S.R., Rao, D.H.: Application of Bayesian Networks for Learner Assessment. E-Learning Systems International Journal of Computer Applications 4(4), 23–28 (2010)
11. Dascălu, M.-I.: Application of Particle Swarm Optimization to Formative E-Assessment in Project Management. Informatica Economica 15(1/2011), 48–61 (2011)
12. Chen, C.-M., Huang, T.-C., Li, T.-H., Huang, C.M.: Personalized E-Learning System with Self-Regulated Learning Assisted Mechanisms for Promoting Learning Performance. In: 7th IEEE International Conference on Advanced Learning Technologies, pp. 637–638. IEEE Press, New York (2007)
13. Gaeta, M., Orciuoli, F., Ritrovato, P.: Advanced ontology management system for personalized e-Learning. Knowledge-Based Systems 22, 292–301 (2009)
14. Hartonas, C.: A Learning Object Metadata Ontology (2013), `http://wdok.cs.uni-magdeburg.de/Members/miotto/diplomarbeit/ontologien-und-beispiel/lom.owl/view` (accessed at April 2013)
15. Learning Technology Standards Comittee of the IEEE, Draft Standard for Learning Objects Metadata, IEEE Draft P1484.12.1/D6.412 (2002)
16. Nikolopoulos, G., Kalou, A., Pierrakeas, C., Kameas, A.: Creating a LO Metadata Profile for Distance Learning: An Ontological Approach. In: Dodero, J.M., Palomo-Duarte, M., Karampiperis, P. (eds.) MTSR 2012. CCIS, vol. 343, pp. 37–48. Springer, Heidelberg (2012)