World Transactions on Engineering and Technology Education
Vol.2, No.3, 2003

The application of adaptive students' classification to the determination of a learning strategy in an e-learning environment

Czesław Daniłowicz & Elżbieta Kukla

Wrocław University of Technology
Wrocław, Poland

ABSTRACT: The article presents the application of adaptive students' classification to the determination of a learning strategy in an e-learning environment. A learning strategy is defined within the context of the knowledge domain, which is explicitly expressed by one of the many possible sequences of hypermedia pages that represent subsequent knowledge units. The presentation techniques that are utilised for the construction of Web pages correspond to those learning and teaching methods that are applied to the class of students and are characterised by a certain definite set of features. Learner classifications aim at determining similar students' categories and the set of learners' features that are considered important from an effective didactic process perspective. An analysis of learning strategies is initially assigned to a given course and class of learners, and the final strategies that have brought the student success are applied to both classification and student profile modifications. A new classification that has been adapted to real learning effects is created on the basis of the procedures presented in the article.

INTRODUCTION

Learning strategies refer to those methods that students use to learn. They are strongly influenced by individual learner preferences with respect to his/her own manner of knowledge acquisition [1]. Learning results depend on both the student's mental predisposition and cognitive process performance. Mental predispositions apply to learner individual abilities that describe the peak performance in a certain domain. In turn, cognitive process performance within a learning environment refers to an individual student's learning style that is determined typical for a given person's way of receiving and processing information, ie thinking, remembering and problem solving. Although people are unique, it is possible to distinguish a final number of learning styles that they represent. Moreover, the information typical for each style's method of thinking and processing information could be used to arrange and shape the knowledge that is transferred to students during a didactic process in order to increase its effectiveness.

This idea is the basis of a solution proposed for the personalisation of e-learning courses. Firstly, it is assumed that students are similar, regarding different features (including individual learning styles), and learn in a similar manner. As such, students taking a course are classified according to these values. Each class of learners is assigned a learning strategy that matches best the needs and abilities of that category's class members. As the learning process advances, the systems gather information about students and the learning effects. This information is used to modify learner characteristics and verify the set of features supposed to influence the learning process. It is also the basis for adapting the classification to an actual system of knowledge of learners. The characteristics of newly created class members is then used to determine the learning strategy for new students that may be classified to this category and who seek to undertake the same course [2].

LEARNING STRATEGY

In order to explain the conception of learning strategies and the procedure of its determination in an e-learning context, it is necessary to assume a general knowledge structure (see Figure 1) and a learning process [3].

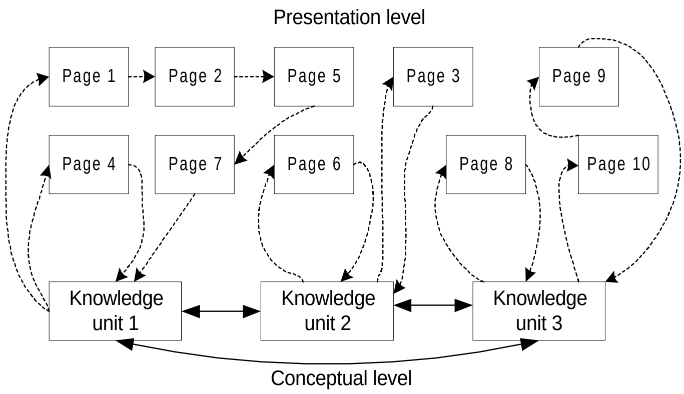

Figure 1: Outline of a general knowledge structure.

Let K be the finite set of knowledge units resulting from the whole knowledge domain partition on elementary, indivisible and meaningful pieces that are connected to each other by different relations (nodes and arcs, respectively, at the conceptual level in Figure 1). It is assumed that there is a defined partial order α in K, which represents the obligatory order in learning these knowledge elements. R_k denotes the set of all presentations of a knowledge unit $k \in K$. Presentations are the sequences of hypermedia pages that are related to every node in the conceptual level (the presentation level). Each sequence corresponds to a teaching method suitable to the knowledge unit presentation. E_k identifies the set of tests serving to verify learners' competences referring to k. A presentation from R_k and a test from E_k are called *corresponding* to k.

Let $R = \bigcup_{k\in K} R_k$ and $E = \bigcup_{k\in K} E_k$. A course L can be defined as a subset of knowledge units, that is $L \subseteq K$. Let $\alpha_L = \{(k,k'): k,k' \in L \text{ and } (k,k') \in \alpha\}$.

Definition 1: Given a learning strategy s for course L, a sequence belonging to Cartesian product $(R \cup E)^{2n}$ for n is called the cardinality of set L, which fulfils the following conditions:

- Each knowledge unit $k \in L$ has exactly one corresponding presentation and one corresponding test in s, and s does not contain any other presentations and tests that correspond to knowledge unit k.
- The order of any two presentations in s should correspond to the order (if one exists) of their knowledge units in relation to α_L.
- The order of any two tests in s should match the order (if one exists) of their knowledge units in relation to α_L.
- Any test should have a presentation as its predecessor and any presentation should be followed by a test.

The following example should explain the definition. Let $L = <k_1,k_2,k_3>$, $R_{k1} = \{r_{k11},r_{k12}\}$, $R_{k2} = \{r_{k21},r_{k22}\}$, $R_{k3} = \{r_{k31},r_{k32}\}$, $E_{k1} = \{e_{k11},e_{k12}\}$, $E_{k2} = \{e_{k21}\}$, $E_{k3} = \{e_{k31},e_{k32}, e_{k33}\}$ and $\alpha_L = \{(a,b)\}$. Potential learning strategies for course L are as follows:

- $s_1 = <r_{k21}, e_{k21}, r_{k32}, e_{k33}, r_{k11}, e_{k12}>$;
- $s_2 = < r_{k12}, e_{k12}, r_{k22}, e_{k21}, r_{k31}, e_{k32}>$;
- $s_3 = < r_{k12}, e_{k11}, r_{k21}, e_{k21}, r_{k31}, e_{k33}>$.

General concepts of the learning process assume that similar students should learn in the same way, namely by the same learning strategy. So, when a learner takes up a course, the system creates an initial profile of the student's characteristics. In the following step, the profile is used to classify the student with a set of similar learners. This isolated set of learners serves to determine an opening learning strategy. At the start of a system functioning, an opening learning strategy is assigned to a given class by experts or pedagogical agents. When there is a class of students who are similar to a new learner and who have finished the chosen course successfully, the opening strategy is determined from an analysis of their final learning strategies.

When the learning procedure begins, the student is offered a first sequence of hypermedia pages indicated by the strategy for the first element of knowledge, and the student starts to learn. Next, the student has to pass a test suitable for the knowledge just learnt. According to the test results a pedagogical agent decides if the student should continue learning in keeping with the earlier learning strategy. If the test results are insufficient, the strategy may be changed to another style for a given knowledge unit: one that has not yet been used and would better suit the learner's characteristics. The agent may also alter the order of the presented knowledge units. The actions of the pedagogical agent always target the identification of a more effective method of learning for a given student.

All decisions that have been made during the course are recalled and used to modify each student's profile, and to generate a final learning strategy. Consequently, a preliminary learners' classification could be no more correct. Therefore, at the end of the learning process, a student may be reclassified. The final strategy that brings the student success, together with the final strategies of other members of the class, are used to determine an opening tutoring strategy for each new student who is classified into the same category.

STUDENT CHARACTERISTICS

It is proposed that students' characteristics be represented in an e-learning environment utilising their individual learner profiles. This should include all of the information concerning the student and takes into account a system adaptability point of view. This data can then be divided into two main categories: personal features and user-system interaction history.

Initial information would be the data that gives the student's identification and those connected with the learner's education level, ie learner's identifier, first and surname, birthday, age, completed education (primary, secondary, graduate and postgraduate). The following values have been assigned for their age: young, middle and advanced, corresponding to the age ranges (18-35〉, (35-65〉 and (65 and over) respectively.

Personal features comprise the individual student's Intelligence Quotient (IQ) with the following values: low (below average), middle (average) and high (above average). The student's learning style will be determined using the model developed by Felder and Silverman [4]. This model considers the learner's behaviour in four bipolar dimensions: perception (sensitive-intuitive), receiving (visual-verbal), processing (active-reflective) and understanding (sequential-global). The student's preference along each particular dimension will be assessed utilising the Index of Learning Styles Questionnaire developed by Felder and Soloman [5]. The results obtained are presented as a pair, where the first element refers to the learner's preferred direction for every dimension, and the second is a score on a scale from 1 to 11, indicating the intensity of the student's behaviour in a given direction. This information will determine alternatives for a given student should a particular process be unsuccessful.

Definition 2: Individual learning style (ILS) is derived as a Cartesian product ILS = PER×REC×PRO×UND, where: PER={SEN, INT}, REC={VIS, VER}, PRO={ACT, REF}, UND={SEQ, GLO}, and SEN = (sensitive, i), INT = (intuitive, j), VIS = (visual, i), VER= (verbal, j), ACT = (active, i), REF = (reflective, j), SEQ = (sequential, i), GLO = (global, j), where i, j = 1, 2, …, 11 and if i>0, then j = 0. Otherwise, if j>0, then i = 0 for: SEN and INT, VIS and VER, ACT and REF, SEQ and GLO.

The second category of the learner's profile data covers dynamic information gathered by the system during the learner-system interaction. This comprises the information concerning the completed courses, ie course identifier, opening and final learning strategies, as well as the currently realised course (ie course identifier and opening strategy). All learning strategies identified in the student's profile have a structure that is consistent with Definition 1. For completed and current courses, all of the tests scores in the opening strategies are equal to 0. The final strategies will have the real scores obtained by the student in successfully passed tests following the related knowledge units presentations.

It should also be noted that, when looking for similarities between students, not all of the attributes are of the same level of importance. So, it was decided that only those attributes that are considered to significantly influence the effects of learning would be taken into account. These form the basis of the classification and are hence called *basic attributes*. The set of *basic attributes B* includes the following (using the first letters of attributes' names and values):

- Age (AG) = {YO, MI, AD};
- Gender (Sex) (SX) = {MA, FE};
- Intelligence Quotient (IQ) = {LQ, MQ, HQ};
- Education (ED) = {EE, SE, HE, PE};
- Individual style of learning: PER, REC, PRO and UND.

The remaining student's data contains the following:

- Learner identifier: PI;
- Name and first name: NF;
- Date of birth: BD;
- Finished courses (opening and final learning strategy): LF = {(I1,Os1,Fs1),...,(Ik,Osk,Fsk)};
- Course being currently realised course (opening learning strategy): LA = (Ia,Osa);
- Score on the scale 1-11 for each attribute: SEN, INT, VIS, VER, ACT, REF, SEQ and GLO.

These are included in the set of *supplementary attributes A* and do not participate in a student's classification.

Initial Students' Classification

The basic attributes define the characteristics of a class and, consequently, they determine the learners' profile distribution into classes. The remaining (supplementary) attributes, which reflect the data about students, could be used for the actualisation of classifications or for future investigations.

Definition 3: Let $B = \{B_1,...,B_N\}$. Profile classification C is called Cartesian product $C = B_1 \times ... \times B_N$. Each element from $C = (b_1,...,b_N)$ of product C is called a class of classification. In other terms, $C \in$ C only if $b_1 \in B_1,...,b_N \in B_N$.

Definition 4: Value b of the attribute B represents a class C = (b1,...,bN) only if b belongs to set b1,...,bN.

Definition 5: The learning strategy s belongs to a class $C = (b_1,...,b_N)$ only if there exists at least one profile p that belongs to C, for which s is the final strategy.

The following provides an example to demonstrate this; let C = AG×SX×IQ×ED×PER×REC×PRO×UND, where *AG*, *SX*, IQ, ED, PER, REC, PRO, UND are *basic attributes* with values defined in the previous section. The category C_j = (YO, MA, MQ, SE, INT, VIS, ACT, SEQ) includes the profiles of students who have the following characteristics:

- Are young (YO);
- Are male (MA);
- Have a middle intelligence quotient (MQ);
- Have completed the secondary education level;
- Like innovations and dislike routine calculations (INT);
- Prefer knowledge presented in graphical form (VIS);
- Like active experimentation (ACT);
- Prefer material presented in a steady progression of complexity and difficulty (SEQ).

In the classification mentioned above as *basic attributes*, it is assumed that these attributes and their values influence the effects of learning. However, it cannot be assumed that the selection of attributes and their values is entirely consistent with the postulate of dependency between learning strategies and basic attributes. Verification of the strategy and classification takes place during the classification process.

LEARNING STRATEGY DETERMINATION

Learning strategies should comply with different students' interests, preferences and skills, which may change periodically. The entire timeframe of system functions can be divided into the intervals $[t_0, t_1)$, $[t_1, t_2)$, $[t_2, t_3)$.... During each one, the information about the learning process will gathered and at the discrete moments t_1, t_2, t_3, ..., the data will be analysed. Obviously, the aggregated information about a particular student's learning process is contained in the final strategy of a course that the student has just completed. So the set of final strategies should form the basis for the analysis of changes in the student's interests, preferences and skills that have occurred in the preceding time interval. The analysis aims at determining new opening strategies, which will be assigned to those students beginning to learn in the subsequent period. Therefore, it is necessary to establish a method of opening strategy determination, based on the foundations of the known final strategies.

Let S_L be a set of all strategies for course L; $s_1,\ldots,s_q$ will be denoted a subset of S_L and by $\beta_1,\ldots,\beta_q$, which is the number of learners who have used strategies $s_1,\ldots,s_q$, respectively.

Definition 6: By the representation of strategies $s_1,\ldots,s_q$, a strategy is called $s_\gamma \in S_L$, which satisfies the following condition:

$$\sum_{i=1}^{q} \beta_i d(s_r, s_i) = \min_{s \in S_L} \sum_{i=1}^{q} \beta_i d(s, s_i) \qquad (1)$$

ADAPTIVE MODIFICATION OF CLASSIFICATIONS

When the student completes the course without changing the opening strategy, he/she stays in a previously determined class. If the strategy is changed, then the pedagogical agent transfers the student to another class. After some period of time, the data gathered by the system, especially in the opening and final learning strategies of the completed courses, is analysed in detail so as to modify the classification and identify a new opening strategy for each class and every course of actualised classification. Some new concepts should be introduced in order to describe the principles of classification modification.

Definition 7: The function of the measurement of the cohesion of learning strategies S is defined as:

$$c(S) = 1 - \frac{\mathrm{diam}(S)}{\mathrm{diam}(S_L)} \qquad (2)$$

where $\mathrm{diam}(S) = \max_{x,y \in S} d(x,y)$ denotes the diameter of the set S, while $\mathrm{diam}(S_L) = \max_{x,y \in S_L} d(x,y)$ is the diameter of the set S_L, $S \subseteq S_L$.

It should be noted that the greater the value of $c(S)$ is, the greater is the cohesion of learning strategies. It is then possible, without committing a significant mistake, to use any learning strategy from the set *S* or to replace one strategy with another.

Definition 8: Let S_F ($S_F \subseteq S_L$) be a set of final learning strategies of a course *L*, belonging to all classes of classification *C* that are represented by values of an attribute B belonging to the subset B_U ($B_U \subseteq B$). The subset B_U is undistinguishable with respect to the course L only if $c(S_F)>\tau_s$, where τ_s is the threshold of the distinguishable ability.

Definition 9: The values of an attribute B that belongs to the subset B_U are undistinguishable only if they are undistinguishable in respect of all the courses.

The undistinguishable subset B_U can be reduced to just one value of the attribute B. This is illustrated in the following example. Attribute $AG = \{YO, MI, AD\}$ belongs to the set of basic attributes but, after some time of system performance, the learning strategies assigned to those classes, wherein the values of the attribute AG belong to set $\{YO, MI\}$, do not differ significantly from each other. Given this, the undistinguishable values YO and MI of the attribute AG could then be replaced by another value. If denoted by YM, the reduction would then be $AG = \{YM, AD\}$.

Definition 10: Attribute B is undistinguishable in respect of course L if, and only if, a set of all values of the attribute B is undistinguishable in respect of course L.

Definition 11: Attribute B is undistinguishable if, and only if, it is undistinguishable in respect of the learning strategies of every course.

An undistinguishable attribute may be removed from, or transferred to, the set of supplementary attributes. The following provides an example.

Attribute SX = {MA, FE} belongs to the set of basic attributes. However, after some time, it appears that the tutoring strategies assigned to those classes, where the values of the attribute SX belongs to the set {MA, FE} slightly differ from each other, ie $c(SX)>\tau_s$. So, the attribute SX has no significant influence on the strategy selection and can, therefore, be removed from the set of basic attributes. The deletion of this attribute from the set of basic attributes does not mean that the learner data, represented by this attribute, should no longer be gathered. This would be a mistake in this case because gender is very important information that characterises a learner and should, therefore, be transferred into the supplementary attributes set.

Attributes can be removed from the basic attributes set only if it is possible to reduce a subset of their values. The procedure for attribute values reduction or attribute transfer to supplementary attributes set is presented below.

Let $L = \{L_1,\ldots,L_Z\}$ and be a set of all courses, $B = \{b_1,\ldots,b_M\}$ and $V=\{v_0,v_1,\ldots,v_W\}$, where $v_0 = B$, then:

$$
\begin{array}{llll}
v_1=B\backslash\{b_1\} & v_{M+1}=v_{\binom{M}{1}+1}= & \cdots & v_{\binom{M}{M-3}+1}= \\
 & =B\backslash\{b_1,b_2\} & & =B\backslash\{b_1,b_2,\ldots,b_{M-3},b_{M-2}\} \\
v_2=B\backslash\{b_2\} & v_{M+2}=v_{\binom{M}{1}+2}= & \cdots & v_{\binom{M}{M-3}+2}= \\
 & =B\backslash\{b_1,b_3\} & & =B\backslash\{b_1,b_2,\ldots,b_{M-3},b_{M-1}\} \\
\vdots & \vdots & \cdots & \vdots \\
v_M=B\backslash\{b_M\} & v_{\binom{M}{2}}= & \cdots & v_W=v_{\binom{M}{M-2}}= \\
 & =B\backslash\{b_{M-1},b_M\} & & =B\backslash\{b_3,b_4,\ldots,b_{M-1},b_M\}
\end{array}
$$

Let S_{ij} be a set of final learning strategies of a course L_j that belongs to the classes of classification C, which then are represented by values belonging to a subset v_i of values of the attribute B ($v_i \subseteq B$).

The procedure for the reduction of the attribute $B=\{b_1,\ldots,b_M\}$ values set is as follows:

BEGIN
1. I: = 0.
2. j: = 1.
3. Calculate $c(S_{ij})$.
4. If $c(S_{ij})>\tau_s$, then go to 5, else go to 9.
5. If $j=Z$, then go to 7, else go to 6.
6. j:= j+1, go to 3.
7. If i = 0 then $C:=C \backslash\{B\}$, $A:=A \cup\{B\}$ (the attribute B has been transferred from *basic attributes set* to *supplementary attributes set*), else go to 8.
8. If $i<W$, then $B:= B\backslash v_i\cup\{v\}$ (values set v_i of the attribute B has been replaced by an attribute v).
9. if $i=W$ then there are no such values of the attribute B that could be reduced, else go to 10.
10. i:= i+1, go to 2.

END

Definition 12: Let S_b be a set of tutoring strategies in those classes represented by a value b of the attribute B. Values b_1 and b_2 replace the value b of the attribute B only if $c(S_b)<\tau_s$ and there are two tutoring strategies sets S_{b1} and S_{b2} where $S_{b1}\cap S_{b2} = \emptyset$, $S_{b1}\cup S_{b2} = S_b$, $c(S_{b1})>\tau_s$, and $c(S_{b2})>\tau_s$.

Attribute replacement could be generalised by introducing the possibility of replacing n values with m values. For example, the values, 20-35, 36-60, and over 60 of age attributes could be replaced by the values, 20-30, 31-40, 41-50, 51-60, and over 61.

CONCLUSIONS

A procedure for the adaptive modification of learners' classification and its application to learning strategy determination in an e-learning environment is presented and discussed in this article. Future works should focus on defining the concrete structures of learner profiles, conceiving algorithms for the profile classifications, as well as concrete methods for the optimal learning strategy determination of a given student and a course.

REFERENCES

1. Daniłowicz, C., Modeling of user preferences and needs in Boolean retrieval system. *Processing and Management*, 30, **7**, 363-378 (1994).
2. Kukla, E., Nguyen, N.T. and Sobecki, J., The consensus-based tutoring strategy selection in CAL systems. *World Trans. on Engng. and Technology Educ.*, 1, **1**, 95-98 (2002).
3. Jenkins, C. and Inman, D., Adaptive automatic classification on the Web. *Proc. Inter. Workshop on Database and Expert Systems Applications*, Los Alamitos, USA, 504-511 (2000).
4. Felder, R. and Silverman, L., Learning and teaching styles in engineering education. *ASEE J. of Engng. Educ.*, 78, **7**, 674-681 (1988).
5. Felder, R. and Soloman, B., Index of Learning Styles (1998), http://www2.ncsu.edu/unity/lockers/users/f/felder/public/ILSdir/ILS-a.htm.