Botsios, S., Georgiou, D., & Safouris, N. (2008). Contributions to Adaptive Educational Hypermedia Systems via on-line Learning Style Estimation. *Educational Technology & Society*, *11* (2), 322-339.

Contributions to Adaptive Educational Hypermedia Systems via on-line Learning Style Estimation

Sotiris Botsios, Demetrius Georgiou and Nikolaos Safouris
Department of Electrical & Computer Engineering, School of Engineering, Democritus University of Thrace, Greece // Tel: +30 254 107 9969 // Fax: +30 254 107 9968 // smpotsio@ee.duth.gr, dgeorg@ee.duth.gr, nsafouri@ee.duth.gr

ABSTRACT

In order to establish an online diagnostic system for Learning Style Estimation that contributes to the adaptation of learning objects, we propose an easily applicable expert system founded on Bayesian Networks. The proposed system makes use of Learning Style theories and associated diagnostic techniques, simultaneously avoiding certain error sources. It takes under consideration learners' answers to a specific questionnaire, as well as a classification of learners who have been previously examined. System application based on Kolb's Learning Style Inventory is also presented, in order to study the changes to the final estimation made by the system. The questionnaire in use is the David Kolb's Learning Style Inventory revised, an on-line form which avoids logical fault implications. With this aim, a case study has been carried out involving 885 university students, from which conclusions are drawn. As a result of the expert system use, the influence of factors, such as reflection of cross-cultural environment differences and/or of slippery or lucky answers' influence on learning style estimation is reduced.

Keywords

Learning style estimation, Adaptive educational hypermedia systems, Bayesian networks, Probabilistic expert systems, User modelling

Introduction

Education via internet represents great and exciting opportunities for both educators and learners. Wagschal (1998) suggests that the internet and World Wide Web have made the computer a dynamic force in distance education, providing a new and interactive means of overcoming time and distance to reach learners. As there is an increasing demand for e-learning procedures tailored to the needs of learners, Learning Management Systems (LMSs) designers take under consideration adaptivity issues. An approach to adaptive learning is considered to be Adaptive Educational Hypermedia Systems (AEHSs). According to Brusilovsky (2001) Adaptive Hypermedia is an alternative to the traditional "one-size-fits-all" approach in the development of Hypermedia Systems. Adaptive Hypermedia Systems build a model of the goals, preferences and knowledge of each individual user, and use this model throughout the interaction with the user, in order to adapt to the needs of that user. For example, a student in an AEHS will be given a presentation that is adapted specifically to his or her knowledge of subject and a suggested set of most relevant links to proceed further. In Papanikolaou, Mabbott et al (2006) work they consider Learning Styles (LSs) among adaptivity parameters and play a crucial role in AEHSs (Papanikolaou, Mabbott et al, 2006).

Learning Theories diverge with respect to the fact that students learn and acquire knowledge in many different ways, which have been classified as LSs. Felder and Silverman (1988) claim that students learn by observing and hearing; reflecting and acting or by reasoning logically and intuitively. Students also learn by memorizing and visualizing; drawing analogies and building mathematical models. Learning behavior has been extensively examined in cognitive psychology. There exists a great variety of models and theories in the literature regarding learning behavior and cognitive characteristics i.e. LSs or Cognitive Styles (CSs) i.e. Sternberg and Zhang (2001). Although some authors do not distinguish between LSs and CSs (Kaltz, Rezaei, 2004), there are others who clearly do (Papanikolaou, Mabbott et al, 2006, Smith, 2001). According to Riding and Rayner (1998) CS refers to an individual's method of processing information. The building up of a repertoire of learning strategies that combine with cognitive style, contribute to an individual's LS. In particular, as Jonassen and Grabowski (1993) reported, LSs are applied CSs, removed one more level from pure processing ability usually referring to learners' preferences on how they process information and not to actual ability, skill or processing tendency. Our research in not intended in focusing on such differences in terminology, so the terms LS and CS will be used interchangably. LSs classifications have been proposed by Kolb (1984) and others (Honey and Mumford, 1992; Dunn, Dunn, 1985; Dunn, Dunn, 1992; Felder, Silverman, 1988; Murray, 1999). Most of the authors categorize LSs and/or CSs into groups and propose certain inventories and methodologies capable of classifying learners accordingly. Despite their efforts for classifications, most of the above authors notice that the LS variable is a continuous. That means that an effort for dichotomous classification of learners could be proved pointless. Most of the proposed LS diagnostic methodologies are addressed to educators.

Educators are humans and are therefore flexible systems. They can adapt to unfamiliar situations in class, and they are able to gather information in an efficient manner, disregarding irrelevant details. The information, which is gathered, could be general, qualitative and vague because humans can reason, infer, and deduce new information. Subsequently, educators exploit all the information concerning their students in order to teach in the best possible way. Educators have common sense, as they can make decisions on teaching strategy, and provide logical explanations for those decisions, as referred by De Köning et al (2000). They can learn, perceive, and improve their skills through experience. Furthermore, they can be creative, inventive, and innovative. However, teachers have weaknesses, too. They can be slow, inaccurate, forgetful, and emotional. In the presence of such characteristics their teaching behaviour becomes unstable. According to Pearl (1988), these exact characteristics reduce the teacher's ability to adapt his/her teaching to the learners' LS.

Educational Technology faces a very challenging task in seeking to develop and possess even a few of the simple human abilities, combined with the computer qualities: speed, accuracy, capacity, reliability. In particular, the challenge faced by researchers in the field of AEHSs mainly considers learner preferences, interests, and browsing behaviours in providing personalized services. Stathacopoulou, Magoulas, et al (2005) propose that the idea is to introduce human knowledge and experience in the computer in order to make it behave like the best possible tutor, who can adapt his/her teaching to the learners' characteristics and abilities. It has been pointed out that LS effects e-learning more than traditional instructor based learning. Moreover, Manocher (2006) among others suggests that e-learning becomes more effective for learners with a particular LS. LSs are considered relevant for the adaptation process in the user model, and have been used as a basis for adaptation in AEHS (Brown et al, 2005; Karampiperis, Sampson, 2005; Georgiou and Makry, 2004). In specific, the property of adaptation should lead to the design of sophisticated Learning Objects (LOs) acquisition methods. Rehak and Mason (2003) consider LO as a digitized entity which can be used, reused or referenced during technology supported learning. Practically, LOs acquisition is achieved by querying LOs repositories distributed over the internet, using LOs metadata standards. The database queries must have solid structure with strictly defined parameters. To this end, researchers as Reye (2004) face the question of establishing suitable techniques for handling the abstraction and uncertenty of the classification proposed by the cognitive theorists. According to Milosevic, Brkovic et al (2007) combining reusable LOs with adaptation scenario brings faster creating of e-learning courses and greater possiblility to tailor the teaching process to individual student needs. Mustaro and Silveira (2006) propose a solution for adaptive LOs retriaval through LSs. In order to accurately provide an adequate fulfillment ot individual learning to individual learning requirements, LMSs must keep track of more than historical aspects of one's evolution in his/her learning process during lifetime: different LS, naturally developed by each students, are to be considered, properly modeled and effectively used for learning purposes.

Dunn et al (1990) studied the influence of cultural behavior to students' responses to LS inventory questions. Cultural behavior is considered to be the common extragenetic behavior exhibited by the members of a community. The non uniform LS distribution of 885 university students from a specific region of Greece test group indicated the affect of extragenetic parameters to the classification. This raises the need for a methodology that takes under consideration cultural behavior, in order to improve user's LS estimation.

A third question concerns the reliability of user's responses to questionnaires. Uncertainties in student responses to such tests come from several sources. Those most commonly cited are slips and lucky guesses (Martin, VanLehn, 1995; Mislevy, Gitomer, 1996; Millán et al., 2003; Reye, 2004). In this study the Kolb's Learning Style Inventory (1984) (KLSI) was applied. David Kolb (1999) introduced his LS inventory that includes 8 items. Each item in KLSI consists of 4 statements that appear in every possible combination of pairs, i.e. six pairs of statements. Therefore, there are six choices and the student is asked to choose one of the two statements for every pair. Depending on the selection of each statement it is possible, due to implication reasons, to determine some future selections. It has been noted that a significant number of the test group members chose statements randomly, ignoring these logical implications.

To sum up we focus on the following questions:

- How LS estimation can contribute to AEHS?

- How cultural behavior influences LSs and in which way such relation can be reflected in LS estimation proceedure?
- In which way the users' lucky and slippery answers effect on LS estimation can be reduced?

In this paper we face the above questions.

- We propose an algorithm to estimate user's strongest preference for a specific LS. Moreover, it provides a LSs' descending taxonomy according to the user's questionnaire responses. This supports AEHS need for LOs retrival through their metadata to the purpose of constructing a learning interface tailored to the learner's needs. It is the LMS designers' decision wether they take under consideration just the user's dominant LS or the complete taxonomy. The algorithm is modulated on the basis of widely known and valid LS theory.
- We also implement the Learning Style Estimation via Bayesian Network (LSEvBN) proposed Botsios, Georgiou and Safouris (2007), a probabilistic expert system based on Bayesian Networks (BNs). Such networks allow learners' common characteristics to add value to their LS estimation. Our system analyzes information from responses to questionnaire supplied by the system's antecedent users (users that completed the questionnaire before the present user) and the system's present user as well. The proposed corresponding algorithm utilizes what appears to be reasoning capabilities so as to reach LS estimations.
- Finally, we monitor the influence of lucky and slippery answers on LS estimation. This is achieved by introducing the Fault Implication Avoidance Algorithm (FIAA). We expect that this algorithm contributes to the reduction of such influence.

The scope of this research is not to provide new psychometrical evidence on the correctness of a certain Learning Style Theory. It also does not propose a new questionnaire. The results presented do not stand up for or against the effectiveness and accuracy of those that already exist. Our effort is in the direction of diminishing the influence of factors that hinder an accurate estimation and providing service to an AEHS. Also, in cases were methodologies as those of David Kolb's concludes with two LSs of equal weight, our method provides a dominant LS as far as it makes use of the system knowledge, i.e. the LS estimation of previous users. Practically, this system can be applied, with minor modifications, to inventories of any kind, making them capable of taking under consideration both the examined user's responses and past users' classifications.

This article is structured as follows: the next section discuss the related work cited in literature. Section 3 describes the FIAA. Brief introductory description of the BNs and how they are applied to expert systems is given in section 4. In section 5 we establish a general model with its mathematical formulations and in the following section we present the model's application specifically in the KLSI. In section 7 the results and the evaluation of our research are presented and the conclusions and future work are discussed in the last section.

Related Work

The issue of estimating a learner's LS in the scope of providing tailored education has been addressed in the literature several times. ACE (Adaptive Courseware Environment) is a WWW-based tutoring framework, developed by Specht (2000), which combines methods of knowledge representation, instructional planning and adaptive media generation to deliver individualized courseware over the WWW. Experimental studies within ACE showed that the successful application of incremental linking of hypertext is dependent on students' LS and their prior knowledge. Kinshuk and Graf (2007) show in their research how cognitive traits and learning styles can be incorporated in web-based learning systems by providing adaptive courses. Such courses fit to the individual characteristics of learners and therefore make learning easier and better accessible for those who have difficulties with the "one-size-fits-all" courses. The adaptation process includes two steps. Firstly, the individual needs of learners have to be detected and secondly, the courses have to be adapted according to the identified needs. The LS estimation in their work is made by a 44-item questionnaire based on Felder-Silverman LS model. In another work, Papanikolaou, Mabbott et al, (2006), experimental results from two empirical studies performed on two educational systems (Flexi-OLM and INSPIRE) to investigate learners' learning and cognitive style, and preferences during interaction, are described. The Index of Learning Styles questionnaire was used to asses the style of each participant according to the four dimensions of the Felder-Silverman LS model. It was found that learners do have a preference regarding their interaction, but no obvious link between style and approaches offered, was detected. Part of the Carver, Howard et al (1999) work was to develop an adaptive hypermedia interface that provided dynamic tailoring of the presentation of course material based on the individual student's learning style. By tailoring the presentation of material to the student's LS, the authors believe students learned more efficiently and more effectively. Students determine their LS by answering a series of 28 questions. These forms were based on an assessment tool developed at North Carolina State University (B.S. Solomon's Inventory of Learning Styles). The Milosevic, Brkovic et al (2007) approach tend to pursue adaptation according to generated user profile and its features which are relevant to the adaptation, e.g. the user's prefernces, knowledge, goals, nabigation history and prossibly other relevant aspects that are used to provide personilized adaptations. They discuss about designing lesson content tailored to individual users, taking into consideration LS (Kolb LS) and subject matter learning motivation and how could learning objects metadata be used for LO retrieval according to the specific needs of the individual learner. Analyzing coordination between student's LS and his motivation for specific teaching material we give guidelines for preparing learning materials according to different learner's characteristics. Those guidelines are based on pedagogical strategy and motivation factor with a strong psychological background.

Work has been published concerning a methodogical approach based on BNs for modelling the behaviour of students on certain courses. Also some efforts have been made for LS recognition via BNs, e.g. Bunt and Conati (2003). To the best of the authors' knowledge, such published works are not based on the use of proposed LS recognition questionnaires or related inventories. Instead they address this problem by building a BN capable of monitoring users' interactions with a LMS, and of providing the type of assessment that the environment needs to guide and improve the learner's exploration of the available material.

In Garcia et al (2007), a BN that detects the student's LS is evaluated. The BN's input is the student's interactions with the Web-based educational system, and the Felder – Silverman classification method is used. The ILS questionnaire proposed by Felder was applied only for results comparisons reasons. Graf and Kinshuk (2006) propose an architecture for LS estimation independent from a specific LMS. Their work is based on the Felder-Silverman LS model. While Garcia et al work is focused on the use of BN, Graf and Kinshuk's (2006) approach sums up indications on preferences based on behaviour patterns, equally to the approach of LS questionnaires. Zapata-Rivera and Greer (2004) present the SModel, a BN student-modelling server used in a distributed multi-agent environment. They implemented their Bayesian student models on a modified version of the belief net backbone structure for student models proposed by Reye (1996). The above-mentioned work applies BN, while the learning process is in progress. It bases LS estimation on the learner's behaviour, avoiding the use of inventories proposed by cognitive science specialists. In addition, Liu (2005) formulates applications with BN and elaborates on the applications of mutual information to adaptive item selection. Moreover Liu creates a BN-based simulation environment that is employed for generating students' data. The simulated data will be used in evaluating different approaches for item selection and student classification.

Furthermore, in the work presented by Danine et al (2006), BN provide an accurate analysis of the errors made by a student while solving a subtraction problem. Finally, the system makes a deeper and comprehensive diagnosis of errors. In Conati et al (2002), BN are used as a comprehensive, sound formalism to handle certain uncertainties. The models that are presented have been implemented in Andes, a tutoring system for Newtonian physics whose philosophy is to maximize student initiative and freedom during the pedagogical interaction. Andes helps students study examples effectively, a novel task that adds new sources of uncertainty to the student modelling problem.

Fang and Blank (2006), present a student model to diagnose students' knowledge level in CIMEL-ITS, an intelligent tutoring system that helps beginners learn object oriented design. This student model provides a refined representation of prerequisite relationships, adds prerequisites to estimate the current students' knowledge level, and guarantees real-time responsiveness using an atomic BN. Moreover, Millán and Pérez-de-la-Cruz (2002) propose a solution which is founded on the definition of a new integrated student model based on BNs, and on the application of Computer Adaptive Tests theory to improve the efficiency and accuracy of the diagnostic process. This new Bayesian student model allows measurement of a student's knowledge at different levels of granularity (that is, the subject domain is curriculum-structured), as well as substantial simplifications when defining the BN. Mislevy, Gitomer (1996) present the Hydrive model to assess the student's competence at troubleshooting an aircraft hydraulics system. The student model evaluates the student's actions and characterizes student understanding in terms of more general constructs (dimensional Variables) that express the student's knowledge of the system, strategies and procedures. A BN is used to express and update these student-model dimensional variables.

In Collins et al (1996), BNs are applied to granularity hierarchies in McCalla and Greer (1994). Assessment is the main purpose rather than diagnostic student modeling. They decompose the domain into learning objectives. The evidence used to evaluate whether the student masters a particular learning objective consists of his/her answers to related test items. To choose the most informative test item they use a utility measure that maximizes the change in expected probability for the learning objective being evaluated. Bayesian updating is used to propagate values from the evidence to higher-level nodes.

In the above mentioned approaches, the authors avoid using questionnaires, and monitored user's interactions with the system for diagnosing the user's LS. In our work, we apply the well established questionnaires of the cognitive psychologists to make the diagnosis, supplementary to the expert BN system, which is regarded this work's main contribution.

Fault Implication Avoidance Algorithm

Let us consider three selection pairs consisting of the statements (a), (b), and (c). Logical implication determines that once the statement (a) is chosen between (a) and (b) in the first selection pair, and (b) is chosen between (b) and (c) in the second selection pair, the choice of (a) instead of (c) is obligatory (Table 1). As the first two selections lead to (a)>(b)>(c) order of preference. Alternatively, reverse choices in pairs 1 and 2 ((b) and (c) instead of (a) and (b) correspondingly) leads to the order (c)>(b)>(a). In every other combination of choices in pairs 1 and 2, no logical implication appears and pair 3 remains open to chose from its statements. At this point a question arises: What if a selection in pair 3 can better represent the user's preference than pair 1 or 2, do not allow a choice to be made in pair 3 and moreover those choices lead to wrong order of (a) and (c). The answer is that pair 3 can only be "locked", ranking statements (a) and (c) in a wrong way, in the very rare case the user's choices in pairs 1 and 2 are both against his/her preferences. In case were only one choice from pairs 1 or 2 is against the user's real preferences, pair 3 remains "unlocked" waiting the user's selection. Obviously, the probability of two sequential "wrong" choices is considerably smaller than making one "wrong" choice even in cases of statistical dependence.

Table 1. Example of fault implication avoidance

pair	statement	input method
1	*a*	user selection
	b	
2	*b*	user selection
	c	
3	*a*	automatic selection
	c	

Analogously, for more than three selections, the final ranking can be reached by responding to a subset of the set of selections pairs. Figure 1 presents a binary tree which is the logical diagram for a set of 4 statements as appears in each item of KLSI. The paths (figure 1) end in every possible combination of responses a user can give in an item. Nodes of the tree represent the "logical ifs" i.e. the user's choices in every pair of statements. For example the leave [a b c d]' denotes the end of a sequence of choices at nodes ("logical ifs") which are (Table 2).

Table 2. Sequence of choices

a instead of b
c instead of d
a instead of d
a instead of c
b instead of d
b instead of c

Apart from "logical ifs", the parallelograms represent the statements that are "locked" because of FIAA. The "locked" statements are disabled (they are faded in the form), making them unable to be selected (figure 2). For example the leave [c d a b]' denotes the end of a sequence of unabled and disabled choices as appears in Table 3.

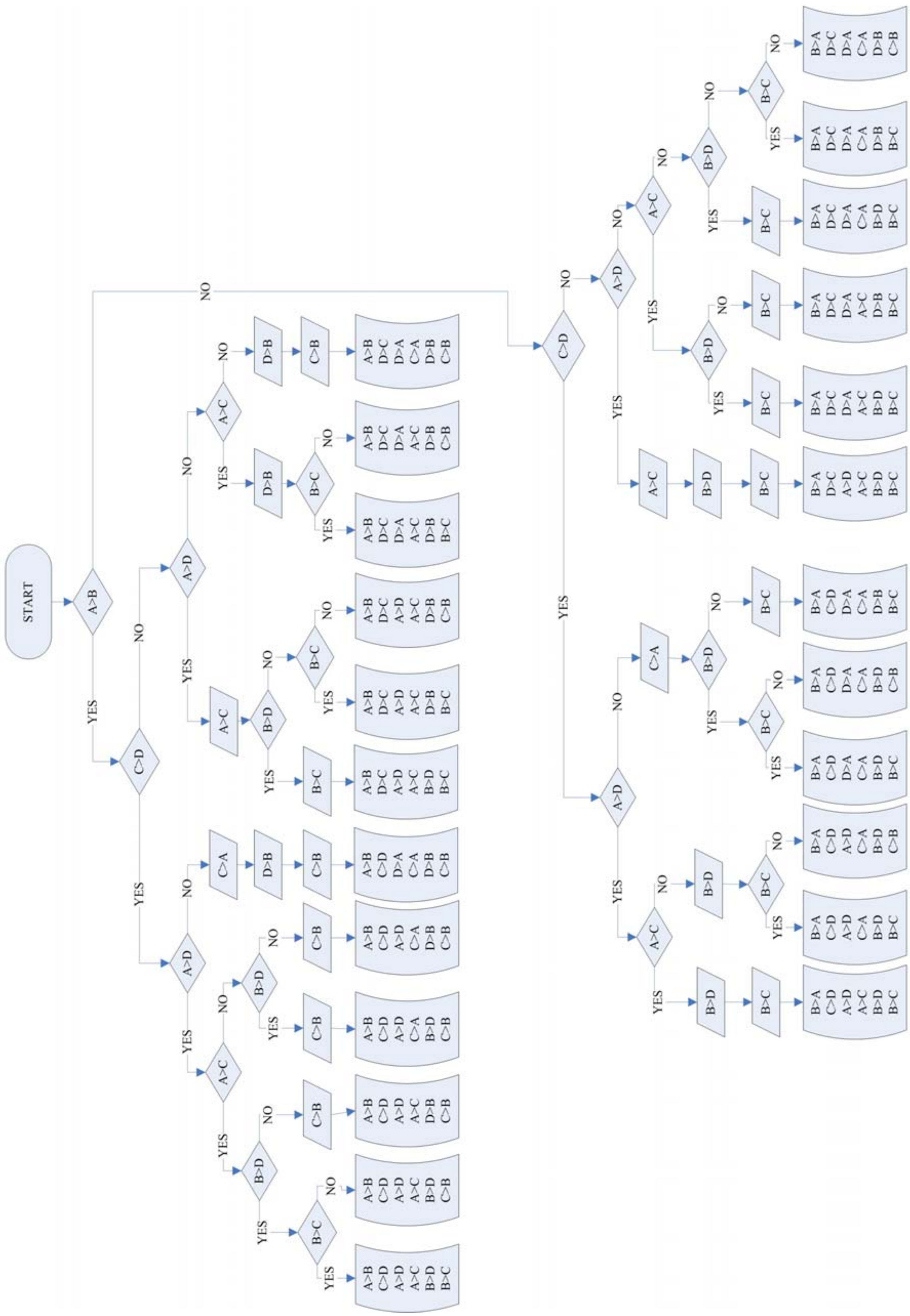


Figure 1. Logical diagram of FIAA

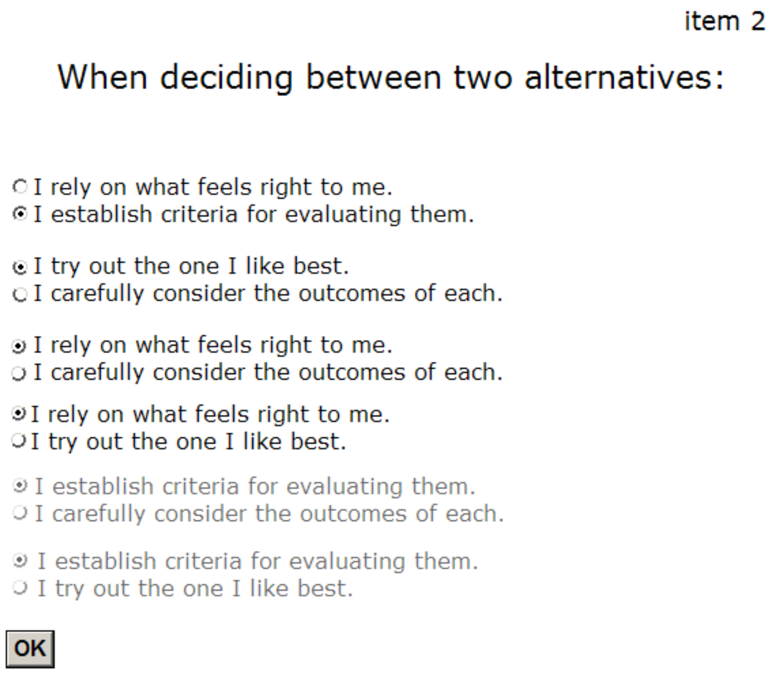


Figure 2. Item example of revised on-line inventory. Pairs five and six are locked and automatically completed due to implication limitations. (no radio buttons are marked by default, the user must make the selection)

Table 3. Sequence of choices

a instead of b
c instead of d
a instead of d
c instead of a (disabled)
d instead of b (disabled)
c instead of b (disabled)

In the printed KLSI there are no such possibilities, as the student has to deal with every single selection pair in the item. It has been noticed that some students who succeeded an early final ranking, they conflict it by their late responses. The original printed KLSI reduces fault logical implication influence on the final estimation by repeating the ranking procedure 8 times (8 items). Taking advantage of the computer capabilities the proposed FIAA makes a step further to face possible fault logical implications.

In our work, the application of FIAA in KLSI provided the revised form of the inventory. In every inventory's item, users respond to limited number of pairs which varies from three to six (figure 2). The remaining pairs take the right values automatically.

Bayesian Networks

The development of artificial intelligence methodology has been recognized as an important requirement in complex asynchronous e-learning situations. LS estimation is a particularly good example, because of the complexity of the learner behaviour and style and of our limited and vague knowledge of how these interact with each other. This estimation is also influenced by the teacher's expertise. Therefore, in order to develop AEHSs capable of estimating LS, researchers face a wide range of relations which arise in complex dynamic systems. The existing relations (which might be unrecognizable) are necessarily poor approximations of complex dynamic systems. As a result some allowance must be made for uncertainty at this level of description. From a probabilistic point of view, such hidden relations mean that there is a degree of uncertainty involved in the LS estimation.

BNs and their close cousins, influence diagrams, have proved to be both a natural representation of probabilistic information and the basis for inference mechanisms that are suitably efficient in practice. Over the last decade, the BN has become a popular representation for encoding uncertain expert knowledge in expert systems, as presented by Heckerman et al (1995).

A BN is a direct, acyclic graph that consists of nodes and arcs. Nodes represent random variables and arcs qualitatively denote direct dependence relationships between the connected nodes. A BN indirectly specifies the joint probability distribution of the random variables, so that we can compute any conditional probabilities that involve variables in the network. Edges in the graph represent causal relationships between random variables, and thus such networks are sometimes called belief networks. In fact, degrees of relation are conditional probabilities adapted as weights to the BN's edges.

By a strictly mathematical definition, a BN is a triplet $(\boldsymbol{V}, \boldsymbol{A}, \boldsymbol{P})$, where
1. $\boldsymbol{V}$ is a set of variables. Each element of $\boldsymbol{V}$ is attached to a single node of the graph.
2. $\boldsymbol{A}$ is a set of arcs, which together with $\boldsymbol{V}$ constitutes a directed acyclic graph $\boldsymbol{G}=(\boldsymbol{V}, \boldsymbol{A})$.
3. $\boldsymbol{P}=\{P(\boldsymbol{M}/\boldsymbol{\pi_\upsilon}): \boldsymbol{\upsilon} \in \boldsymbol{V}\}$, where π_υ stands for the set of parents of $\boldsymbol{M}$ ($\boldsymbol{M} \subset \boldsymbol{V}$).

In words, $\boldsymbol{P}$ is the set of conditional probabilities of all variables given their respective parents (Note that when $\boldsymbol{M}$ is a root, $\boldsymbol{\pi_\upsilon}$ is empty). In such a case, the expression $P(M/\pi_\upsilon)$ simply stands for the prior probability of $\boldsymbol{M}$ (Xenos, 2003; Millán et al., 2000; Nadkarni and Shenoy, 2001; Zhang, Poole, 1994).

The model of Learning Style Estimation via Bayesian Networks

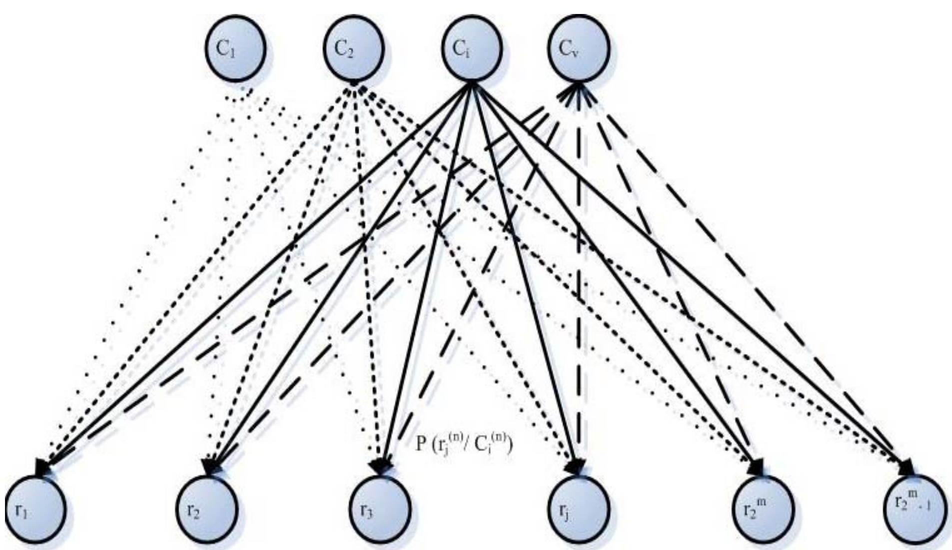


Figure 3. The model of LSEvBN

Let us consider the $BN=(\boldsymbol{V},\boldsymbol{A},\boldsymbol{P})$ where $\boldsymbol{V}=\boldsymbol{\pi_\upsilon} \cup \boldsymbol{M}$ and $\boldsymbol{\pi_\upsilon}=\boldsymbol{LS}=\{C_1, C_2, \ldots, C_v\}$ is the set of LSs. A learner is recognized as being of class C_i, ($i=1,2,\ldots,v$) according to his/her responses to a given set of m questions. Each question can be answered by yes or not. Let $\boldsymbol{Q}=\{Q_1^{(k)},Q_2^{(k)},\ldots,Q_m^{(k)}\}$ be the set of answers where k is a Boolean operator taking the values TRUE or FALSE whenever $Q_1^{(k)}$ represents the answer YES or NOT respectively. There are 2^m different sets of such responses to the questionnaire. Let us consider the index j, where $j \in \{1,2,\ldots,2^m\}$. A learner's responses to the set of questions formulates an element

$$r_j = \bigcup_{l=1}^{m} Q_l^{(k)} \tag{1}$$

where $r_j \subset \boldsymbol{M}$ the set of BN leaves. Obviously, $r_i \neq r_j$ for any pair $(r_i, r_j) \subset \boldsymbol{M}$, with $i \neq j$.

Inference, or model evaluation, is the process of updating probabilities of outcomes based upon the relationships in the model and the evidence known about the situation at hand. When our Bayesian model is actually used, the end user provides evidence about his or her LS through the responses to a given questionnaire. This information is applied to the model by "instantiating" or "clamping" a variable to a state that is consistent with the responses. Thus, we proceed with a probabilistic expert system based on the previously described BN. The probabilistic expert system emphasizes the basic computational principles that make probabilistic reasoning feasible in expert system. The key to computation in these systems is the modularity of the probabilistic model (Shafer, 1987). Nieto et al (2001) suggest an extension of the Classical Probability Theory that allows combine new evidence about an hypothesis along with prior knowledge (or assumptions) in order to arrive to an estimate of the likelihood the hypothesis being true.

The necessary mathematical mechanics are performed to update the probabilities of all the other LS and the array of responses variables that are connected to the variable representing the new evidence. Let n be the number of learners who made use of the system, and n_{r_i} be the number of those who responded to the questionnaire with r_i. The a priori probability that the $(n+1)^{th}$ user responded to the questionnaire with an element r_i is

$$P\left(r_i^{(n+1)}\right) = \frac{\left(n+1\right)_{r_i}}{n+1} \tag{2}$$

In this case, the LSEvBN in use is a weighted and oriented $K_{2^m}^{v}$ graph, i.e. a weighted and oriented complete bipartite graph on n and 2^m nodes. Figure 3 represents the proposed LSEvBN. At each edge of the network's graph we adjust the conditional probability $P(r_i^{(n)}/C_j^{(n)})$, i.e. a probability which dynamically changes as a new user enters the system. This probability expresses the ratio of users who responded to the questionnaire with the element r_i and were finally classified in C_j, in terms of the total number of r_i responses. Thus, the measure $P(C_j^{(n+1)})$ is the probability that the LS of the $(n+1)^{th}$ learner belongs to C_j. This probability is given by the relation

$$P\left(C_j^{(n+1)}\right) = \sum_{i=1}^{2^m} P\left(C_j^{(n)} \big/ r_i^{(n)}\right) P\left(r_i^{(n+1)}\right), \quad j = 1, 2, \ldots, v \tag{3}$$

where

$$P\left(C_j^{(n)} \big/ r_i^{(n)}\right) = \frac{P\left(r_i^{(n)} \big/ C_j^{(n)}\right) P\left(C_j^{(n)}\right)}{\sum_{k=1}^{v} P\left(r_i^{(n)} \big/ C_k^{(n)}\right) P\left(C_k^{(n)}\right)}, \quad \forall (i,j) \in \{1,2,\ldots,n\} \times \{1,2,\ldots,2^m\} \tag{4}$$

Let $score_j^{(n+1)}$, $j=1,2,\ldots,v$ be the score for the j^{th} LS, that the $(n+1)^{th}$ student gets by responding to the revised inventory. Then, by the contribution of LSEvBN, the learner's j^{th} LS final score is given by

$$ls_j = P\left(C_j^{(n+1)}\right)\left(score_j^{(n+1)}\right) \tag{5}$$

Then the dominant LS is the maximum value of ls_j, $j=1,2,\ldots,v$.

In case where $ls_i = ls_j$ $(i \neq j)$ holds, the learner can be classified either in class C_i or in class C_j. In order to avoid such conflict (or even with more than two ls participating in the equality), the system redirects the programme flow to a subsystem where the algorithm is repeated for the dominant classes C_i and C_j (or more). To this purpose, for every user the system updates different probabilistic expert system subgraphs, where $\pi_\upsilon \subset \mathcal{P}(LS) \cap \{\{C_1\},\{C_2\},\ldots,\{C_v\},\varnothing\}^C$. In words, this is the set of all subsets of *LS* with more than one element. In this way, all possible sample spaces of the *LS*×M are constructed and systematically updated, ready for use at first demand. Although this is a rather rare case, we consider it as necessary supplement of the algorithm. In our test group no such cases have been observed.

Application

Kolb's learning theory sets out four distinct learning styles (or preferences), which are based on a four-stage learning cycle (figure 4), which might also be interpreted as a ‘training cycle’.

When implementing KLSI, the set LS has four elements which represent the four LSs as they appear in figure 3.

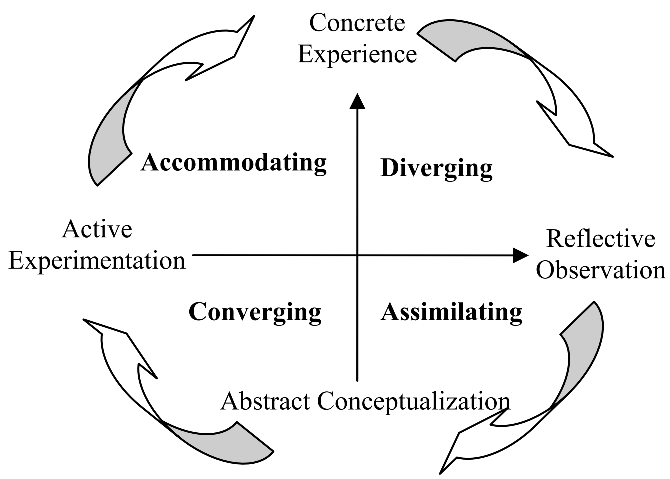


Figure 4. Kolb’s Learning Cycle

According to David Kolb (1999), diagnosis of LS can be based on the learner’s response to the inventory proposed by him. Based on a description of the way one learns as well as the way one deals with ideas and day-to-day situations in his/her life this inventory has proven to be a useful diagnostic tool. The learner who responds to the inventory, marks the scores he or she has received on a separate scoring sheet. These scores are finally represented on a two-dimensional Cartesian plane giving a dominated vector located on a quadrant. In some cases the resulting vector lays on, or in the vicinity of the bisectors.

In the developed application we made use of the KLSI revised version (FIAA algorithm). Following the approval of our research request by the Hay Resources Direct, Hay Group Inc. we translated the inventories questions-statements and created an on-line questionnaire.

Follow to the welcome / log-in screen (figure 5) and some brief instructions, the items and their corresponding statements appear, according to D. Kolb’s order of appearance. As noted before, the FIAA disables some statements according to the statements already selected (figure 2).

The user responses are recorded in a database and the LS estimation is instantly made. The dominant LS and a brief text about the user’s most suitable way of learning are presented to the user.

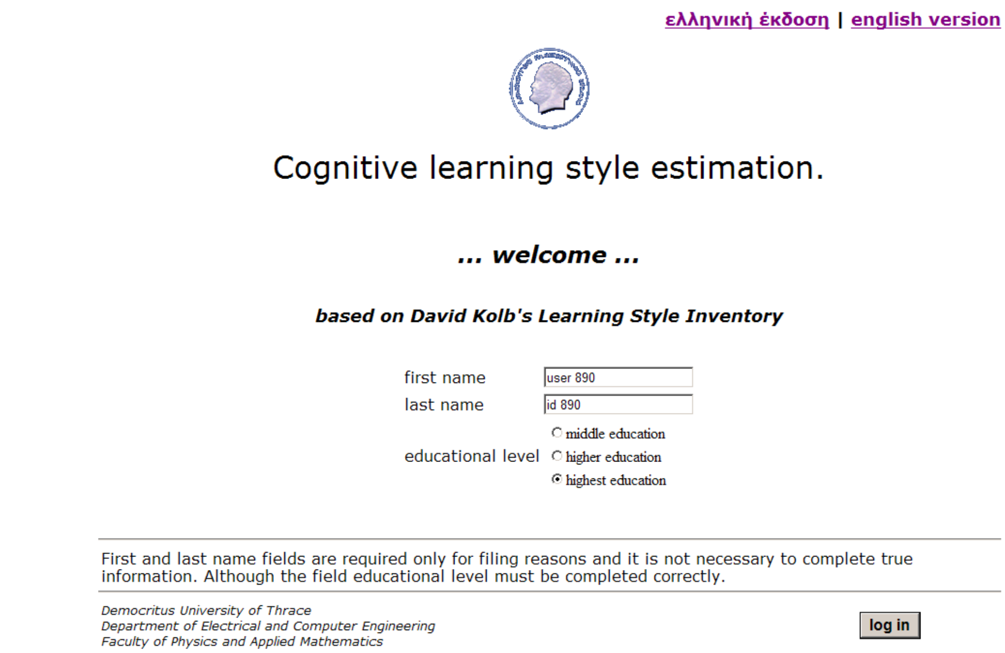


Figure 5. Welcome and user login screen of revised on-line inventory questionnaire (English version)

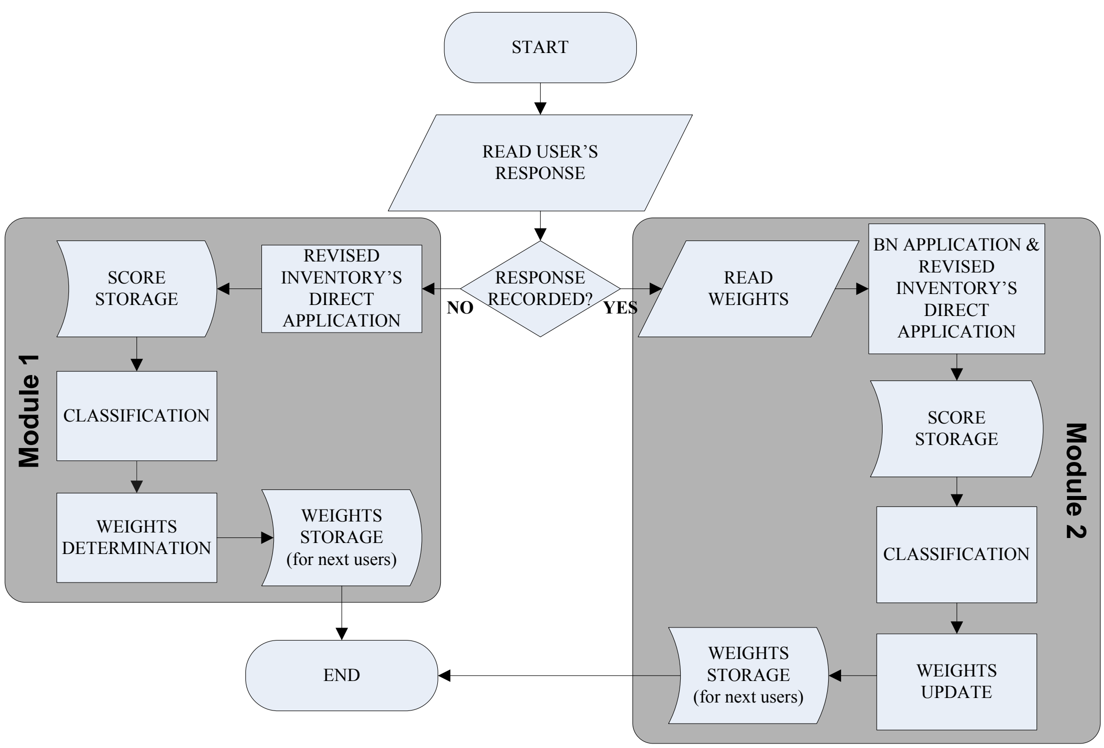


Figure 6. Block diagram of the LSEvBN

According to the model discussed in the previous section, let us now consider that the set $\pi_{\upsilon}=LS=\{CE,RO,AC,AE\}$ has elements belonging to Kolb's distinct Learning Styles. The set M consists of all possible responses to the proposed inventory. Normally, each response is a 96-digits binary word. That is because there are 8 items with 6 pairs of statements each. When referring to any pair of statements, the user is forced to choose a single one. Therefore, $card(M) = 2^{48}$ indicates all the possible elements, i.e. the total number of different responses to Kolb's originally proposed inventory. The revised online inventory we use, accepts a considerably smaller number of responses ($card(M) \leq 2^{32}$) due to the FIAA. It follows that the LSEvBN is a weighted and oriented $K^{4}_{2^{32}}$ graph having weighted edges the conditional probabilities $P(r_i^{(n)}/C_j^{(n)})$.

For the purpose of applying the proposed LSEvBN it is necessary to mark *a priori* conditional probabilities of the expert system. At the starting point no data had been stored and so there was direct use (i.e. skipping the probabilistic expert system's application) of the revised online inventory. Both the users' responses and the LS classes appointed to them, were stored. Also the initial conditional probabilities were calculated "waiting" for the descendant users. In cases where the system recognized a response which was yet unknown, the classification is made from the direct application (i.e. skipping the probabilistic expert system's application) of the revised inventory. If there was a previously recorded response of the same kind, the LSEvBN used current *a priori* probability values, in order to classify the system's user. At the same time, some of the LSEvBN's conditional probabilities were updated according to the evidence they gathered.

In the proposed algorithm one recognizes the following modules (figure 6):

1. Every time an unrecorded response $r_i^{(n)}$ appears the system recognizes that it has no other identical one stored in the database. In this case the system skips the probabilistic expert system application and simply stores the response $r_i^{(n)}$ and the $score_i^{(n)}$. Otherwise the system's control goes to module 2.
2. The probabilistic expert system application. One of the strong points of the model is that the probabilities are updated dynamically every time a new user makes use of it. This part of the algorithm makes use of the stored data to retrieve conditional probabilities $P(r_i^{(n)}/C_j^{(n)})$. In this step, using formulas (3) and (4), the program calculates probabilities of the elements in *LS*. The system therefore returns to a LSs hierarchy. As soon as a response $r_i^{(n+1)}$ (that differs from the stored responses) appears, module 1 is activated. This part of the algorithm continuously improve the quality of LS estimations based on experience gained from previous applications of the model.

Results and evaluation

We employ the previously described algorithms to evaluate this research. The test group consisted of 885 (66.1% male and 33.9% female) undergraduate university students (ages between 19-22 years old from the Engineering Faculty) from Northern Greece. Their responses to the revised online inventory were stored to a database and statistical results were drawn and presented.

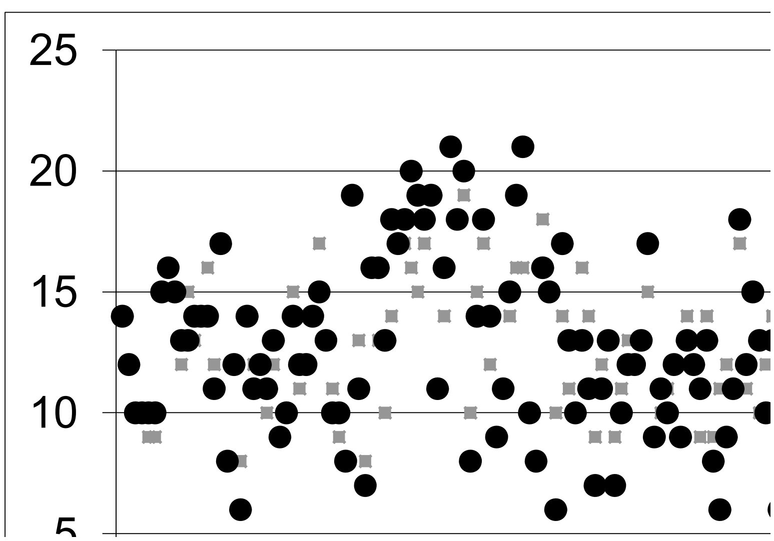


Figure 7. The distribution of users' questionnaire scores

Evaluation of FIAA

In order to evaluate the FIAA efficiency, a randomly selected sample of the test group completed the KLSI. The sample size was 124 students. In figure 7 the distribution of users' questionnaire scores (KLSI and revised inventory) are presented.

The correlation coefficient has been calculated $\rho_{XY}=0.925$. Therefore, the relation between KLSI scores and revised questionnaire scores is strong. Moreover, in this sample we found that 16.13% of students did not fall into implication errors and 83.87% made at least one implication error. For the users who did not fall into fault logical implications the LS taxonomy as estimated from the printed KLSI and the revised form is identical. A detailed table concerning the distribution of fault implication errors is given (Table 4):

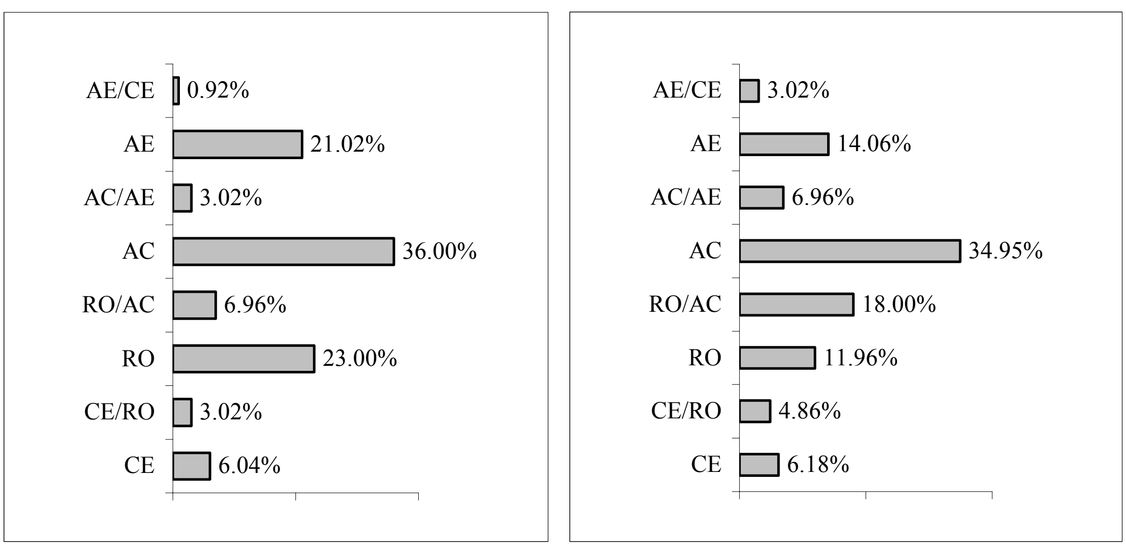
Table 4. Distribution of fault implication errors

Number of errors	Relative frequency (%)	Number of errors	Relative frequency (%)
1	16.13	5	9.68
2	22.58	6	12.90
3	9.68	7	3.23
4	3.23	9	3.23
		11	3.23

Furthermore, it has been noticed that 23.08% of the students who felt into implication errors using the KLSI, scored equally in two LS classes. 50% of them were classified in a single class with the application of FIAA. The algorithm assigned other LS classes than those that have been diagnosed by the KLSI, only to 9.67% of the sample. A percentage of 3.23 of the sample did not classify in a certain LS class, neither with the original KLSI, nor with the FIAA. Even for these students the LSEvBN algorithm assigns a certain LS class.

Evaluation of LSEvBN

We employed the LSEvBN algorithm to the group of 761 out of 885 students to test its capability to overcome the cases where the inventory's direct application results to more than one LSs of equal score.



Figures 8 and 9. Results from the direct use of the inventory

Using the LS inventory directly, 14% of the examined cases resulted with equal scores and 33% with scores of one unit difference. We took student answers with a LS score of one unit's difference under consideration as we believe that one unit is not sufficient to clarify a student's LS due to the uncertainty introduced by numerous reasons already mentioned above. Detailed picture of those results appears in figures 8 & 9 (that two LSs tie, or differ by one unit). The CE/RO, RO/AC, AC/AE, AE/CE bars present the percentage of the population who scored equally in two consecutive LSs in Kolb's Learning Cycle (figure 8) and at most one-unit difference (figure 9).

Using the LSEvBN all of that cases had been driven into one clear estimation, as it appears in figure 10.

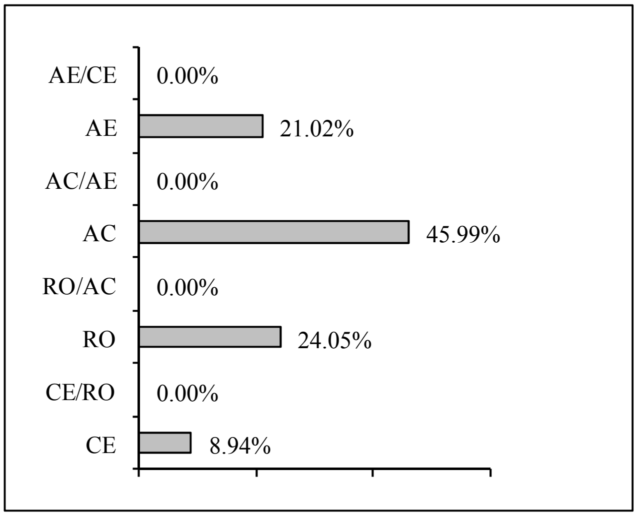


Figure 10. Results from the use of LSEvBN

Figure (10) presents the LS distribution, as resulting from the LSEvBN application. The percentages of undefined LSs are zero, which means that the proposed algorithm always proposes an estimation, which is considered this work's main contribution, based both on questionnaire methodology and previous knowledge. It has to be stressed out that the users that scored equally in two LSs with the direct use of the questionnaire (i.e. skipping the probabilistic expert system's application), they are now classified in one of these LSs that participated in the equality. AC still gathers the most students. Its high percentage functions as a source of attraction for the LS estimation of those test group members who scored equally or with a one-unit difference. For those who scored with a difference of two or more units, no significant influence has been observed.

Some interesting cases of scores per LS are presented in Table 5. The first column contains the user's identification label (Id). The next 4 columns show the LS score obtained by each student through the inventory direct use (KI). The last 4 columns contain corresponding scores gained through the LSEvBN (BN).

Table 5. Scores per LS of selected learners

Id	CE/KI	RO/KI	AC/KI	AE/KI	CE/BN	RO/BN	AC/BN	AE/BN
805	16	11	7	14	3.916	2.989	2.020	3.516
806	16	11	7	14	3.935	2.983	2.014	3.507
215	14	10	14	10	3.413	2.482	3.486	2.491
721	1	17	17	13	0.191	4.636	4.919	3.208
841	11	14	13	10	2.702	3.501	3.412	2.419
284	16	15	5	12	3.832	4.009	1.214	2.914

In the first two rows, the response and the duplicate of a certain learner's response have been recorded. The first row (user id 805) is a case where CE is clearly the dominant LS. Both methodologies, direct questionnaire use and Bayesian network, conclude that the highest score is CE. It is interesting to note that the recorded duplicate user's id 806, regardless of whether it gets the same score after responding to inventory, he/she is classified by the LSEvBN in CE again, but the score is slightly affected by the history (antecedent users). As a result CE/BN is higher while RO/BN, AC/BN and AE/BN are slightly lower.

In the third and fourth row, two cases are recorded where there is a tie through direct questionnaire use. User 215 is considered as being both CE and AC and user id 721 as being RO and AC. The LSEvBN classifies those two users as AC, because, as mentioned earlier, the highest percentage of antecedent users have been classified as such. Furthermore, user id 841 case, where RO/KI=14 and AC/KI=13, is classified by LSEvBN as RO. This is the case situation where the dominant LS scores differ by one unit. Again, LSEvBN provides a straight classification.

The last row user id 284 is quite confusing. The direct questionnaire's estimation is CE=16 and RO=15 (one unit's difference). The LSEvBN estimation is CE<RO. When user 284 answered, there was a high concentration of responses that classified this user as RO instead of CE, and this resulted in $P(CE) << P(RO)$. Similar situation appeared only ten times, especially when our test group was still small.

Conclusions and future work

As it appears in cases presented in Table V, our method avoids the undesirable unclear LS classifications (classifications with more than one leading LSs of equal scores). As a matter of cases where the questionnaire clearly classifies a user into a certain class, our system concluded in the same manner. In this way our research contributes to the economy of AEHS functionality, by reducing the range of LOs metadata characterisations concerning LSs. We restrict such LOs metadata characterisations to n classes instead of $n!$. Nevertheless, it is on the LMS designer choice to use more than one of the leading classes of the ranking in the LO database queries. For example using the LSEvBN algorithm the LMS designer may either decide to retrieve LOs according to the leading LS of the resulted taxonomy, or to retrieve LOs partially from the first and the second top LSs. The late case, allows educators, who will use the LMS, to tune up the system according to their preferred learning methodologies.

It should be mentioned that it was in the authors' intentions to develop an easily applicable on-line LS estimation method based on prevalent learning theories. As appeared in the system's evaluation, we examined the cases where the dominant LS classes scored at most by one unit difference. Nevertheless we did so, one can easily re-evaluate the system using higher differences. In such cases the conclusion of the evaluation does not change.

Referring to the FIAA, we provided evidence that the effect of slippery answers and lucky guesses can not be neglected. This evidence supports the application of the FIAA in the revised form of the inventory. As it has been shown in Table IV a considerably large number of students felt into more than three fault implications. The randomly selected sample of students has been asked to respond the KLSI. Later their responses copied to the revised inventory. The comparison showed that a significant number of students were driven into fault implications. It has been proved for data in figure 7 that for users who did not fall into fault implications the LS taxonomy as estimated from the printed KLSI and the revised form is identical. Also, the correlation factor calculated denotes a strong relation between KLSI and revised form. Slight appearing differences are due to users' fault logical implications and their influence on LS estimation can not be neglected.

As Figure 1 shows, FIAA results in 24 taxonomies of different kind. One recognizes 4 paths (16.7%) as a result of six consecutive choices, 14 paths (58.3%), 4 paths (16.7%) and 2 paths (8.3%) as result of five, four, and three consecutive choices correspondingly. By avoiding the FIAA application the fields of the tree would be as many as 2^6=64. The drastic cut down of the selection pairs a user has to complete contributes to the shrink of the boredom effect as it may appears during the questionnaire completion. Thus, the smaller number of choices is expected to result on the accuracy's increment.

Let us recall, that the effectiveness of the model depends on the cognitive characteristics of the test group under examination. In this study, we noticed that there is strong evidence that test groups taken from populations of various origins formulate LSs distributions of different types, and we intend to further investigate this issue.

Following this, efforts will be made to use fuzzy methods for the purpose of LS estimation. It is considered valuable to use both probabilistic and fuzzy methods in order to detect LSs, as a comparison of such LS estimations will add value to online LS estimation methods.

According to Brusilovsky (2001), while adaptive hypermedia researchers have begun exploring the use of individual traits for adaptation in several areas, it cannot be described as a success story at present. Many researchers agree on the importance of modelling and using individual traits, but there is little agreement on which features can and should be used, or how to use them. We believe that our work provides an alternative prospective and contributes in the controversial and “blur” use of LS estimation questionnaires in the field of AEHS.

Acknowledgements

This work is supported in the frame of Operational Programme “COMPETITIVENESS”, 3rd Community Support Program, co financed (75% by the public sector of the European Union – European Social Fund and 25% by the Greek Ministry of Development – General Secretariat of Research and Technology).

References

Botsios, S., Georgiou, D.A., & Safouris, N.F. (2007). Learning Style Estimation Using Bayesian Networks. *Paper presented at the 3rd International Conference on Web Information Systems and Technologies*, March 3-6, 2007, Barcelona, Spain.

Brown, E., Cristea, A., Stewart, C., & Brailsford, T. (2005). Patterns in Authoring of Adaptive Educational Hypermedia: A Taxonomy of Learning Styles. *Educational Technology & Society, 8* (3), 77-90.

Brusilovsky P. (2001). Adaptive Hypermedia. *User Modeling and User-Adapted Interaction, 11*, 87-110.

Bunt, A., & Conati C. (2003). Probabilistic Student Modeling to Improve Exploratory Behavior. *User Modeling and User-Adapted Interaction, 13* (3), 269-309.

Carver, C.A., Howard, R.A. and Lane, W.D. (1999). Addressing different learning styles through course hypermedia. *IEEE Transactions on Education, 42* (1):33-38.

Collins, J.A., Greer, J.E. & Huang, S.X. (1996). Adaptive Assessment Using Granularity Hierarchies and Bayesian Nets. *Paper presented at the Third International Conference on Intelligent Tutoring Systems*, June 12-14, 1996, Montréal, Canada.

Conati, C., Gertner, S. A., & VanLehn, K. (2002). Using Bayesian Networks to Manage Uncertainty in Student Modeling. *User Modeling and User-Adapted Interaction, 12* (4), 371-417.

Danine, A., Lefebvre, B., & Mayers, A. (2006) TIDES - Using Bayesian Networks for Student Modeling. *Paper presented at the 6th IEEE International Conference on Advanced Learning Technologies*, July 5-7, 2006, Kerkrade, The Netherlands.

De Köning, K., Breweweg, B., Breuker, J. & Wielinga, B. (2000) Model – based reasoning about learner behavior. *Artificial Intelligence, 117* (2), 173-229.

Dunn, R., & Dunn, K. (1992), *Teaching elementary students through their individual learning styles: Practical approaches for grades 3-6*, Boston, MA: Allyn & Bacon.

Dunn, R., Dunn, K., & Price, G. (1985), Learning Style Inventory Research Manual. *Price Systems.*

Dunn, R., Gemake, J., Jalali, F., & Zenhausern, R. (1990). Cross-cultural differences in learning styles of elementary-age students from four ethnic backgrounds. *Journal of multicultural counselling and development, 18* (2), 68-93.

Fang, W., & Blank, G. D. (2006). Student Modeling with Atomic Bayesian Networks. *Paper presented at the 8th International Conference on Intelligent Tutoring Systems*, June 26, 30, 2006, Jhongli, Taiwan.

Felder, R., & Silverman, L. (1988). Learning and Teaching Styles in Engineering Education. *Engineering Education, 78* (7), 674-681.

Garcia, P., Amandi, A., Schiaffino, S., & Campo, M. (2007). Evaluating Bayesian network’s precision for detecting students’ learning styles. *Computers & Education, 49* (3), 794-808.

Georgiou, D., & Makry, D. (2004). A Learner's Style and Profile Recognition via Fuzzy Cognitive Map. *Paper presented at the 4th IEEE International Conference on Advanced Learning Technologies*, August 30 – September 1, 2004, Joensuu, Finland.

Graf, S., & Kinshuk (2006). An Approach for Detecting Learning Styles in Learning Management Systems. *Paper presented at the 6th IEEE International Conference on Advanced Learning Technologies*, July 5-7, 2006, Kerkrade, The Netherlands.

Heckerman, D., Geiger, D., & Chickering, D. (1995). Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning, 20* (3), 197-243.

Honey, P., & Mumford, A. (1992).*The learning styles helper's guide*, Maidenhead: Peter Honey.

Jonassen, D.H., & Grabowski, B.L. (1993). *Handbook of individual differences: Learning and instruction*, Hove: Lawrence Erlbaum.

Kaltz, L., & Rezaei, R. (2004). Evaluation of the reliability and validity of the cognitive style analysis. *Personality and Individual Differences, 36* (6), 1317-1327.

Karampiperis, P., & Sampson, D. (2005). Adaptive Learning Resources Sequencing in Educational Hypermedia Systems. *Educational Technology & Society, 8* (4), 128-147.

Kinshuk, & Graf, S. (2007). Considering Cognitive Traits and Learning Styles to Open Web-Based Learning to a Larger Student Community. *Paper presented at the first International Conference on ICT & Accessibility*, April 12-14, 2007, Hammamet, Tunisia.

Kolb, D. (1984). *Experiential Learning: Experience as the Source of Learning and Development*, Englewood Cliffs: Prentice Hall.

Kolb, D. (1999). *Learning Style Inventory – version 3: Technical Specifications*, Boston, MA: Hay Group.

Liu, C. L. (2005). Using Mutual Information for Adaptive Item Comparison and Student Assessment. *Educational Technology & Society, 8* (4), 100-119.

Manochehr, N. N. (2006). The influence of learning styles on learners in e-learning environments: An empirical study. *Computers in Higher Education Economic Review, 18*, 10-14.

Martin, J., & VanLehn, K. (1995). Student assessment using Bayesian nets. *International Journal of Human Computer Studies, 42* (6), 575-591.

McCalla, G.I., & Greer, J.E. (1994). Granularity-based reasoning and belief revision in student models. In J. E. Greer, J.E. & McCalla, G. I (Eds.), *Student Modeling: The Key to Individualized Knowledge-Based Instruction, NATO ASI Series, Series F, 125*, 39-62.

Millán, E., & Pérez-de-la-Cruz, J.L. (2002). A Bayesian Diagnostic Algorithm for Student Modeling and its Evaluation. *User Modeling and User-Adapted Interaction, 12* (2-3), 281-330.

Millán, E., Pérez-de-la-Cruz, J. L., & García, F. (2003). Dynamic versus Static Student Models Based on Bayesian Networks: An Empirical Study. *Paper presented at the Seventh International Conference on Knowledge-Based Intelligent Information and Engineering Systems*, September 3-5, 2003, Oxford, UK.

Millán, E., Pèrez-de-la-Cruz, J.L., & Suárez, E. (2000). Adaptive Bayesian Networks for Multilevel Student Modelling. *Paper presented at the Fifth International Conference on intelligent tutoring systems*, *June 19-23, 2000,* Montreal, Canada.

Milosevic, D., Brkovic, M., Debevc, M., & Krneta, R. (2007). Adaptive Learning by Using SCOs Metadata. *Interdisciplinary Journal of Knowledge and Learning Objects, 3*, 163-174.

Mislevy, R., & Gitomer, D. H. (1996). The role of probability-based inference in an intelligent tutoring system. *User Modeling and User-Adapted Interaction, 5*, 253-282.

Murray, W. (1999). An Easily Implemented Linear – time Algorithm for Bayesian Student Modelling in Multi-level Trees. *Paper presented at the Ninth International Conference of Artificial Intelligence in Education*, July 19-23, 1999, Le Mans, France.

Mustaro, P.N., & Silveira, I.F. (2006). Learning Objects: Adaptive Retrieval through Learning Styles. *Interdisciplinary Journal of Knowledge and Learning Objects, 2*, 35-46.

Nadkarni, S., & Shenoy P. P. (2001). A Bayesian network approach to making inferences in causal maps. *European Journal of Operational Research, 128*, 479-498.

Nieto, K., Yanez, E., Silva, C., & Barbieri, M.A. (2001). Probable fishing grounds for anchovy in northern Chile using an expert system. *Paper presented at the Geoscience and Remote Sensing Symposium*, July 9-13, 2001, Sydney, Australia.

Papanikolaou, W., Mabbott, A., Bull, S., & Grigoriadou, M. (2006). Designing learner-controlled educational interactions based on learning/cognitive style and learner behaviour. *Interacting with Computers, 18*, 356-384.

Pearl, J. (1988). *Probabilistic Inference in Intelligent Systems*, California: Morgan Kaufmann.

Riding R., & Rayner, S. (1998). *Cognitive Styles and Learning Strategies*, London: David Fulton Publishers.

Rehak, D. R., & Mason, R. (2003). Keeping the learning in learning objects. In Littlejohn, A. (Ed.), *Reusing online resources: a sustainable approach to e-Learning*, London: Kogan Page, 22-30.

Reye, J. (1996). A Belief Net Backbone for Student Modelling. *Paper presented at the Third International Conference on Intelligent Tutoring Systems*, June 12-14, 1996, Montréal, Canada.

Reye, J. (2004). Student Modelling based on Belief Networks. *International Journal of Artificial Intelligence in Education, 14*, 63-96.

Shafer, G. (1987). *Probabilistic Expert Systems*. *Paper presented at the CBMS-NSF Regional Conference Series in Applied Mathematics*, Cambridge University Press, UK.

Shih, S. C., Kuo, B. C., Khosla, R., Howlett, R.J., & Jain, L.C. (2005). Using bayesian networks for modeling students' learning bugs and sub-skills. *Paper presented at the Knowledge-based intelligent information and engineering systems*, September 14-16, 2005, Melbourne, Australia.

Smith, S.E. (2001). The relationship between learning style and cognitive style. *Personality and Individual Differences, 30* (8), 609-616.

Specht, M. (2000). ACE-Adaptive Courseware Environment. *Lecture Notes in Computer Science, 1892*, 380-383.

Stathacopoulou, R., Magoulas, G.D., Grigoriadou, M., & Samarakou, M. (2005). Neuro-fuzzy knowledge processing in intelligent learning environments for improved student diagnosis. *Information Sciences, 170*, 273-307

Sternberg, R.J., & Zhang, L. F. (2001). *Perspectives on Thinking, Learning, and Cognitive Style,* Netherlands: Springer.

Wagschal, P.H. (1998). Distance Education Comes to the Academy: But are we asking the right questions? *The internet and higher education, 1*, 125-129.

Xenos, M. (2003). Prediction and assessment of student behaviour in open and distance education in computers using Bayesian networks. *Computers & Education, 43* (4), 345-359.

Zapata-Rivera, J.D., & Greer, J. (2004). Inspectable Bayesian student modeling servers in multi-agent tutoring systems. *International Journal of Human-Computer Studies, 61* (4), 535-563.

Zhang N.L., & Poole, D. (1994). A simple approach to Bayesian network computations. *Paper presented at the Tenth Biennial Canadian Artificial Intelligence Conference,* May 16-20, 1994, Banff, Canada.