LEARNING STYLE ESTIMATION
USING BAYESIAN NETWORKS

S. Botsios, D.A. Georgiou, N.F. Safouris

Department of Electrical & Computer Engineering, School of Engineering,
Democritu University of Thrace, GR 67100 Xanthi, Greece
smpotsio@ee.duth.gr , dgeorg@ee.duth.gr, nsafouri@ee.duth.gr

Abstract: In order to improve the efficiency of Learning Style estimation, we propose an easily, applicable, Web based, expert system founded on Bayesian networks. The proposed system takes under consideration learners' answers to a certain questionnaire, as well as classification of learners who have been examined before. As a result, factors such as cultural environment will add value to the learning style estimation. Moreover, the influence of wrong answers, caused by various reasons, is expected to be reduced.

1   INTRODUCTION

The development of artificial intelligence methodology has been recognized as an important requirement in complex asynchronous e-learning situations. Cognitive Style (CS) estimation is a particularly good example, because of the complexity of the learner behaviour and style as well as of our limited and vague knowledge of how these interact to each other. This estimation is also influenced by the teacher's expertise. Such difficulties mean that a degree of uncertainty is involved in Learning Style (LS) estimation. Moreover, acquisition of Learning Objects (LO) in Adaptive Educational Hypermedia Systems (AEHS) requires analysis of the learner's CS. The link between LS estimation and LO retrieval thus produces large numbers of cause-effect relations at many interacting levels of both description and function. The relations are necessarily poor approximations of complex dynamic systems, and some allowance must be made for uncertainty at this level of description.

There exists a great variety of models and theories in the literature regarding LS and CS. Although some authors do not distinguish between LS and CS (Kaltz, Rezaei, 2004), there are others who clearly do (Smith, 2001). In any case, both of them are considered relevant for the adaptation process in the user model, and have been used as a

basis for adaptation in AEHS (Georgiou, Makry, 2004). Related models have been proposed by Kolb (Kolb, 1984), Honey and Mumford, Dunn R. and Dunn K. (Dunn, Dunn 1985 & Dunn, Dunn 1992), Felder and Silverman (Felder, Silverman, 1988), Murray (Murray, 1999) and others.   Most of the authors categorize LS and/or CS into groups and propose certain inventories and methodologies capable of classifying learners accordingly.

Such procedures can be influenced by a wide variety of errors which may be caused by reasons such as diverse as misconception, false use of the space that has been alloted for the answer or bad formulation of the questionnaire. Learners may also respond to the questions in a wrong way, as slippery answers or lucky guesses due to misconceptions appear (VanLehn, Martin, 1995 & Reye, 2004). Another significant source of poor LS estimation can be deficiencies in the formulation of the questionnaire itself. Barros et al (Baros, Verdejo, Read, Mizoguchi, 2002) address the issue of cultural environment influence on learners' behavior. A review of the vast literature shows that such factors lead to controversial comments on the model's applicability and efficiency (Murray 1999). Despite the bottleneck caused by such reasons, it is worthwhile developing LS estimation techniques.

In order to improve the efficiency of LS estimation, we propose an expert system based on Bayesian Networks (BN). The proposed system takes under consideration learners' answers to a

certain questionnaire, as well as classification of learners who have been examined before. BNs, and their close cousins, influence diagrams, have been proved to be both a natural representation of probabilistic information and the basis for inference mechanisms that are suitably efficient in practice. A BN is a direct, acyclic graph that consists of nodes and arcs (Pearl 1988). Nodes represent random variables and arcs qualitatively denote direct dependence relationships between the connected nodes (Milan, de la Cruz, Suarez, 2000). A BN indirectly specifies the joint probability distribution of the random variables, so 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 causal networks. In fact, degrees of relation are conditional probabilities adapted as weights to the Bayesian network's edges.

In this paper we introduce a BN capable of classifying learners in a predefined set of classes. It is expected that our method, which takes advantages of previously accumulated knowledge, will be more accurate than LS direct estimation, i.e. an estimation based only on single user responses. Since such knowledge is based on the responses to the given questionnaire made by antecedent users, their classification in LS classes provides information that contributes to the random variables' degree of relation. It is noted that the use of the proposed BN restricts the LS grey areas, i.e. the areas where the estimation does not provide a clear output.

In order to implement the BN we propose, we made use of the Kolb's Learning Style Inventory (LSI) (Kolb, 1999).

2 RELATED WORK

Work has been published that accords with LS recognition via BN. Bund et al. (Bunt, Conati, 2003), address this problem by building a BN capable of detecting when the learner is having difficulty exploring, and of providing the types of assessments that the environment needs to guide and improve the learner's exploration of, the available material. In Garcia et al. (Garcia, Amandi, Sciaffino, Campo, 2005), 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. They used the Felder – Silverman classification method. Zapata-Rivera et al. (Zapata-Rivera, Greer, 2004), present SModel, a BN student-modeling 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 as 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.

3 THE MODEL

Let $LS=\{C_1,C_2,...,C_v\}$ be the set of LSs. A learner is recognized being as of class C_i, $(i=1,2,...,v)$ according to his/her responses to a given set of m questions. Each question can be answered by *yes* or *not*. Let $M=\{Q_1^{(k)},\ Q_2^{(k)},...,Q_m^{(k)}\}$ be the set of answers where k is a Boolean operator taking the values *TRUE* or *FALSE* whenever $Q_l^{(k)}$ 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 $j \in \{1,2,...,2^m\}$. A learner's responses to the set of questions formulates an element

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

where $r_j \in M$. Obviously, $r_i \neq r_j$ for any pair $r_i, r_j \in M$, with $i \neq j$. Let n be the number of learners who made use of the system, and n_{ri} be the number of them who responded to the questionnaire with an r_i. The a priori probability that the $(n+1)^{th}$ user responded to the questionnaire with an element r_i is

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

In this case, the BN in use is a weighted and oriented K^v_{2m} graph, i.e. a weighted and oriented
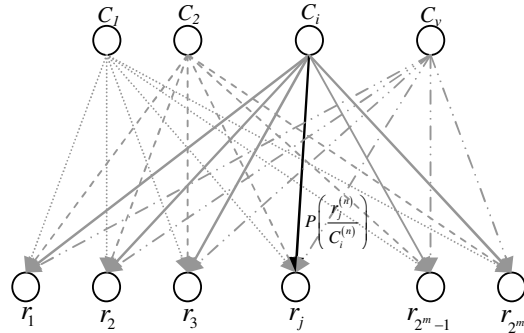

Figure 1: The proposed BN

complete bipartite graph on n and 2^m nodes. Figure 1 represents the proposed BN.

4 THE IMPLEMENTATION

Kolb's learning theory sets out four distinct learning

Table 1: Kolb's Learning Cycle

D	As	C	Ac
Diverging (Feel and Watch)	Assimilating (Think & Watch)	Converging (Think & Do)	Accommodating (Feel & Do)
C_1	C_2	C_3	C_4	C_1
Concrete Experience (CE - Feeling)	Reflective Observation (RO – Observing)	Generalization and Abstract Conceptualization (AC – Thinking)	Active Experimentation (AE - Doing)	Concrete Experience (CE - Feeling)

At each edge of the network's graph we adjust the conditional probability $P(r_i^{(n)}/C_j^{(n)})$. This probability expresses the ratio of users who responded to the questionnaire with the element r_i and were finally classified to 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

$$P\left(C_j^{(n+1)}\right) = \sum_{i=1}^{2^m} P\left(C_j^{(n)}/r_i^{(n)}\right) P\left(r_j^{(n+1)}\right) \qquad (3)$$

where

$$P\left(C_j^{(n)}/r_i^{(n)}\right) = \frac{P\left(r_i^{(n)}/C_j^{(n)}\right) P\left(C_j^{(n)}\right)}{\sum_{k=1}^{n} P\left(r_i^{(n)}/C_k^{(n)}\right) P\left(C_k^{(n)}\right)}, \qquad (4)$$

$$\forall (i,j) \in \{1,2,...,n\} \times \{1,2,...,2^m\}$$

Finally, the learner's dominant LS is given by

$$P\left(C_0^{(n+1)}\right) = \max_{1 \le j \le v} P\left(C_j^{(n+1)}\right) \qquad (5)$$

Where $P(C_i^{(n+1)}) = P(C_j^{(n+1)})$ for $i \neq j$, the learner can be classified either in class C_i or in class C_j. Since this conflicts with the procedure, the system, in order to avoid such a situation, redirects the programme flow to a subsystem where the whole procedure is repeated on a BN which has only the dominated classes C_i and C_j.

In what follows, the proposed model is applied using the Kolb's Adaptive Style Inventory (Kolb, 1999).

styles (or preferences), which are based on a four-stage learning cycle (figure 2), which might also be interpreted as a 'training cycle'.

Based on Kolb's Learning Cycle, the set LS has four elements which represent the four LSs as they appear in table 1.

Let us consider $LS=\{CE,RO,AC,AE\}$ the set of four classes. The set M has $card(M)=2^{48}$ which indicates all the possible elements, i.e. the arrays of answers to Kolb's inventory. It follows that the BN is a weighted and oriented $K^4_{2^{48}}$ graph having as weights at its edges the conditional probabilities $P(r_i^{(n)}/C_j^{(n)})$.

To start with, we define the initial conditional probabilities. The BN is therefore trained by a direct classification via Kolb's inventory. Special attention has been paid to avoiding an initial uniform joint distribution that results in the system's inability to detect the user's LS. To this end, further direct classification via Kolb's inventory is made, skipping the use of BN. The data produced enrich the system's database and modulates the conditional probabilities. Practically, such implications are not expected to occur after the initial system's training.
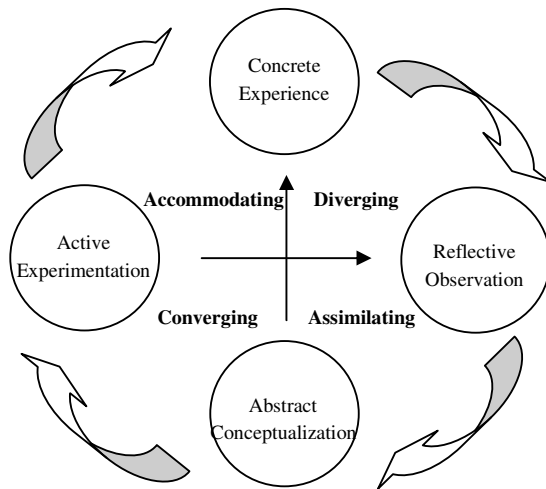
Figure 2: Kolb's Learning Cycle

In the proposed algorithm one recognizes the following steps:

1. The system's training. This is necessary at the beginning, as there are no stored data. So, as the system recognizes a certain response $r_i^{(n)}$, having no other identical match stored in the database, it skips the BN part of the algorithm and simply stores the response $r_i^{(n)}$ and the ratios, as they appear in the Kolb's calculations, in the data base.

2. The BN application. This part of the algorithm makes use of the stored data to calculate conditional probabilities $P(r_i^{(n)}/C_j^{(n)})$. In this step, formulas (3) and (4), the program calculates probabilities of the elements in *LS*. The system therefore, returns an LSs hierarchy. According to Kolb's learning cycle, the two leading LSs characterize the learner as D, As, C, or Ac. As soon as a response $r_i^{(n)}$ (different from the stored responses) appears, step 1 is activated.

4 CONCLUSION AND FUTURE WORK

Using collected data from various test groups, we shall compare LS direct diagnoses to the diagnoses that are outcomes of the proposed algorithm. We expect to have explicit diagnoses even in cases where the direct application of Kolb's inventory leads to equal LS scores.

ACKNOWLEDGEMENTS

REFERENCES

Barros, B., Verdejo, M.F., Read, T., Mizoguchi, R., 2002. *Lecture Notes in Computer Science; vol 2313. Proceedings of the Second Mexican International Conference on Artificial Intelligence: Advances in Artificial Intelligence.* Springer-Verlag. London.

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

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

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

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

Garcia, P., Amandi, A., Schiaffino, S., Campo, M., 2005. Evaluating Bayesian network's precision for detecting students' learning styles. In *Computers and Education (in press)*

Georgiou, D., Makry, D., 2004. A Learner's Style and Profile Recognition via Fuzzy Cognitive Map. In. *Proceedings of the IEEE International Conference on Advanced Learning Technologies (ICALT04)*. IEEE.

Kaltz, L., Rezaei, R., 2004. Evaluation of the reliability and validity of the cognitive style analysis. In *Personality and Individual Differences, vol 36.*

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

Kolb, D., 1999. *Learning Style Inventory – version 3: Technical Specifications*, TRG Hay/McBer, Training Resources Group.

Millán, E., Pèrez-de-la-Cruz J., Suárez, E., 2000. Adaptive Bayesian Networks for Multilevel Student Modelling, Intelligent Tutoring Systems. In *Proceedings of the Fifth International Conference*. Montreal, Canada.

Murray, W., 1999. *An Easily Implemented Linear – time Algorithm for Bayesian Student Modelling in Multi-level Trees, Artificial Intelligence in Education: Open Learning Environments: New Conceptual Technologies to Support Learning, Exploration and Collaboration, Frontiers in Artificial Intelligence and Applications*, vol. 50, pp. 413 – 420.

Pearl, J., 1988. *Probabilistic Inference in Intelligent Systems*, Morgan Kaufmann, San Mateo, California.

Reye, J., 1996. *Lecture Notes In Computer Science; vol 1086. Proceeding of the Third International Conference on Intelligent Tutoring Systems*. Springer-Verlag. London.

Reye, J., 2004. Student Modelling based on Belief Networks. In *International Journal of Artificial Intelligence in Education, vol 14.*

Smith, SE., 2001. The relationship between learning style and cognitive style. In *Personality and Individual Differences, vol 30.*

Van Lehn, K., Martin, J., 1995. Student assessment using Bayesian Nets. In *International Journal of Human Computer Studies, vol 42.*

Zapata-Rivera, JD., Greer, J., 2004. Inspect able Bayesian student modeling servers in multi-agent tutoring systems. In *International Journal of Human-Computer Studies, vol 61.*