Modeling Students' Views on the Advantages of Web-Based Learning with Bayesian Networks

Petri Nokelainen
Tomi Silander
Henry Tirri
firstname.lastname@cs.helsinki.fi
University of Helsinki
Department of Computer Science
CoSCo (Complex Systems Computation Group)
P.O.Box 26 FIN-00014 Helsinki, Finland
Tel: +358 9 708 44173, Fax: +358 9 708 44441

Anne Nevgi
Kirsi Tirri
firstname.lastname@helsinki.fi
University of Helsinki
Department of Education
P.O.Box 39 FIN-00014 Helsinki, Finland
Tel: +358 9 191 28018, Fax: +358 9 191 28073

Abstract. This study discusses student's experiences on Web-based learning in a Helsinki Virtual Open University HEVI [01] environment with the help of a new modeling technique, Bayesian networks. The advantage of Bayesian network models relies both on improved modeling capabilities and in the possibility to use such models to implement new integrative modules that enable interactive activities (e.g. tests, quiz's, questionnaires) in computer supported learning environments. The results of our experimental study showed that students' views on learning in Virtual University could be used in the process of building more learner-centered approaches to Web-based teaching. The results of this study give information about the preferences in learning by different learners.

Keywords: adaptive and distributed learning environments, meaningful learning, Bayesian network

1 Introduction

The objective of our research is to study how students experience Web-based learning in Helsinki Virtual Open University HEVI [01] environment. In addition to this objective the research reported also aims at gaining experience with a new modeling technique, Bayesian networks, which to our knowledge has not been used in this application domain prior to our study. The advantage of Bayesian network models relies not only in the improved modeling capabilities (see e.g., Ruohotie et al. 1999, 85-120, Ruohotie & Nokelainen 2000, 207-238), but also in the possibility to use such models in the further development of new adaptive learning environments (Niemivirta et al. 2000). The paper discusses both the results of our experimental study, and conducts an extensive comparison of the Bayesian network analysis and traditional factor analysis results.

2 The Study

2.1 Educational theory

The educational theoretical framework for the study is built on constructivism, which is seen as a philosophy of learning based on the idea that knowledge is constructed by learners (Kirschner 1999). Consistent with this philosophy, learning must be situated in a rich context and it needs to be reflective of real world contexts. The growing number of students who wish to participate in higher education challenges universities to develop distributed learning environments (Dede 1996, 1997). A distributed learning environment is intended to be learner-centered, enabling both synchronous and asynchronous interaction through the integration of appropriate technologies. The approach gives instructors the flexibility to customise learning environments to meet the needs of diverse student populations, while aiming to provide both high quality and cost-effective learning opportunities. Learning environments promote the use of the Internet (usually Web-based material) to help students find, evaluate and process information, solve problems, communicate ideas, work collaboratively, and learn how to learn (Kirschner 1999). The basis for the questionnaire was Jonassen's criteria list of meaningful learning (Jonassen 1995). The combination of these qualities is centered around seven abstract criteria which derive from student-centered and constructivist learning: "Constructive", "Active", "Collaborative", "Intentional", "Contextualised", "Conversational" and "Reflective". This list of meaningful learning was operationalised into 49 propositions in the questionnaire.

2.2 Bayesian theory

Uncertainty is something we need to model when we create models of the domain. Uncertainty can be modeled in many different ways, probability being one of such methods, and so-called "fuzzy set" (Manton et al. 1994) another one. Probability is a mathematical construct that behaves in accordance with certain rules (Bernardo and Smith 2000, Berry 1996) and can be used to represent uncertainty. In order to be able to perform inferences using the model, probability needs to be interpreted somehow. Depending on this interpretation, we end up in different inference frameworks; the classical statistical inference is based on a long-run frequency interpretation of probability, and the Bayesian inference is based on the "degree of belief" interpretation.

¨Paper presented in PEG2001, Tampere, Finland, June 23-26, 2001.

In frequency interpretation the probability of an (observable) event is the long run proportion of the time it happens compared with the total number of observations—where long-run means in the limit as the total number of observations tends to infinity. Alternatively probability can be defined as a subjective assessment concerning whether the event in question will occur (or has occurred). At first sight it might be quite astonishing that anything useful comes from a definition that involves subjective assessment. Now degree of belief depends on the person who has the belief, as well as on the event on question. In Bayesian inference, this person could be any experimenter or observer. One should always be aware that there is not such a thing as the probability P(A) of an event A, as the probability will always depend on the state of knowledge of the one who believes. Obviously, some opinions are based on more information than others.

Subjective degree of belief interpretation applies any time the subject in question has an opinion, and if one counts ignorance as an opinion, this includes every setting. More importantly, subjective information can change when new information arrives. It should be observed that subjectivity in this context does not mean arbitrariness, i.e., that since all probabilities are subjective, everybody has different probabilities. The degree of belief definition of probability says that with different information one may get different probability. However, all subjects sharing the same information will always assign the same probability to the event. Thus the state of knowledge determines the value of the probability. Bayesian inference is based on this degree of belief interpretation of probability. All Bayesian probabilities depend on the available information, and they actually are mathematical concepts known as conditional probabilities.

Bayesian inference is based on this degree of belief interpretation of probability. Since all Bayesian probabilities depend on the available information, they actually are mathematical concepts known as conditional probabilities, and are denoted P(A | I), where I represents the information affecting the probability assignment. The right hand side of the bar "|" is called the conditioning part. In many cases, however, to shorten the mathematical notation the conditioning part is dropped if the conditioning is clear from the context.

Since the Bayesian inference uses conditional probabilities to represent uncertainty, we are interested in P(M | D,I)—the probability of unknown things (M) given the data (D) and background information (I). In EDUFORM we are typically interested in the answers that the student will give to rest of the questions given some initial responses. The initial uncertainty about M is also represented as a conditional probability P(M | I). For example we could have some initial belief that some answers are more likely than others. Now the essence of Bayesian inference is in the rule that tells us how to update our initial probabilities P(M | I) if we see data D, in order to find out P(M | D,I). If we return to our example case this means that we could "update" our beliefs in the various alternative answers based on the answers the student has already given. This update rule is known as Bayes' theorem and can be formally expressed in Equation 1.1.

$$P(M \mid D, I) = \frac{P(D \mid M, I)P(M, I)}{\int P(D \mid M, I)P(M, I)dM} \tag{1.1}$$

Consequently Bayesian inference briefly comprises the following principal steps:

- Obtain the initial probabilities P(M | I) for the unknown things. These probabilities are called the prior (distribution).
- Calculate the probabilities of the data D given different values for the unknown things, i.e., P(D | M,I). This function of the unknowns is called the likelihood.
- Finally the probability distribution of interest, P(M | D,I), is calculated using the Bayes' theorem given above. This so called posterior (distribution) will then express what is known about M after observing the data.

This is all there is to Bayesian inference at this conceptual level. The Bayes' theorem can be used sequentially, i.e., if we first receive some data D, and calculate the posterior P(M | D,I), and at some later point in time receive more data D', the calculated posterior can be used in the role of prior to calculate a new posterior P(M | D,D',I) and so on. The posterior P(M | D,I) expresses all the necessary information to perform predictions. The more data we get, the more certain we will become of the unknowns, until all but one value combination for the unknowns have probabilities so close to zero that they can be neglected.

2.3 Benefits of using Bayesian over traditional techniques

Bayesian methods have two major benefits in this study over traditional statistic techniques. The first is the ability to analyse almost any kind of data. There is no 'invalid' or 'incomplete' data for Bayesian techniques when at least two cases exists. The second, discussed later in "Results", is that Bayesian probability theory allows us to produce naïve causal relationships between variables.

Comparing pretensions issued on data in Bayesian and traditional techniques is easy due to fact that former has very few, but latter has more than enough. When testing if data is applicable for traditional techniques we must be very careful; for example many traditional multivariate techniques are dependent of Pearson product moment correlation which requires linear relationship between variables.

Next we state some pretensions issued on data to be analysed with traditional techniques: *1) The measurement scale of variables should be at least interval or preferably continuous*. Likert –scale questionnaires produce in most cases nominal and ordinal data. The solution for educational researcher has for years been simple – he or she has specifically decided to treat the data as interval! *2) Data should be approximately normally distributed*. This is a difficult rule to obey due to fact that educational phenomenon very seldom follows uniform sampling fraction. Reader should notice that some traditional modeling techniques are less sensitive to non-normality (i.e. factor analysis) than others (i.e. discriminant analysis), but the general rule prefers normality. *3) Data should also be screened graphically with boxplots of the within-group distributions of each variable.* Researcher should also take care of variances, transformations and relations among variables and use scatterplots to study relations among pairs of variables. Printing covariance matrix also helps to compare between different variables across the groups. During the variable selection process a researcher applying traditional statistical techniques must reject all the variables that can not keep up with the requests stated in the previous paragraph.

2.4 The data

The participants in the study (N=412) represent students who have studied in Helsinki Virtual Open University (HEVI) [01] during the period 1995-1999. HEVI is a Web-based learning environment, where students can study, get advice, and receive help from tutors. We later refer to this course as "Web-course". HEVI is a project, which is implemented in co-operation with different departments of the University of Helsinki and its Open University. The Ministry of Education supports it and it will be completed at the end of the year 2000. The students were asked to evaluate the advantages of Web-based learning with the help of a questionnaire, which measured basic components of learner-centered and constructivist ideas in learning (Bonk & Cunningham 1998, 25-50, Jonassen 1995). The majority of students (83%) lived in the capital area or Southern parts of Finland. Half of them (58%) used their own computer at home for studying. Majority of the respondents were female (73%) and born in the 1950s or 1960s (69%). Half of them (51%) had a university degree.

2.5 Method

The data gathering method of the study was a survey that was mailed to all the students of the Virtual University during the years 1995-1999. The total size of this total population was 875. The questionnaire was sent to the students by mail in September 1999. After two second-mailings the total number of returned questionnaires was 412, i.e., approx. 50% answering rate.

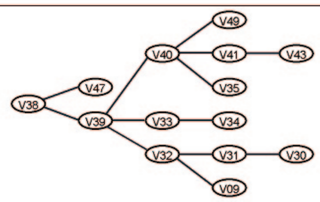
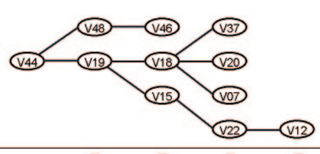
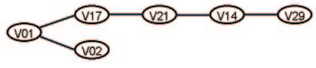
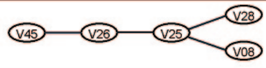
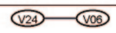
The questionnaire used a 5-point Likert-scale. The students were asked to assess the advantages of learning in virtual university. The answers were analysed concurrently with both traditional linear and Bayesian statistical methods in two phases. The first phase was a traditional variable selection with standard statistic indicators (means and standard deviations). The second phase examined more abstract dimensions underlying students' ratings with the help of factor analysis and Bayesian Network Modeling. The chosen factor solution was based on the "variance greater than 1.0" rule (Kaiser 1970, 401-415) and the "scree" of Cattell (1978, 76-91). Bayesian networks were compared to the original theoretical model by competing them against the reference network and predicting the probability of each applicant network.

3 Results

In our analysis the six-factor solution was found to be most appropriate. The factors represented the following abstract advantages of learning in a Virtual University (Jonassen's names in parenthesis): Factor 1 "Transferability of learning into situation of life (Transfer)", Factor 2 "Communication and interaction with other students (Co-operation)", Factor 3 "Independence in planning the studies (Intentionality)", Factor 4 "Learning and Individual Support", Factor 5 "Constructivity" and Factor 6 "Opportunity to build on previous knowledge".

Extracted Bayesian networks were compared with rotated factorial solution. Comparison of the results is presented in Table 1. Results show that the actor analysis (PCA) and Bayesian network solution (B-Course [02]) are almost identical. Additional value of applying Bayesian networks is their ability to indicate naïve causal relationships between variables. B-Course also allows researcher to study interactively causal relationships. (Table 1.)

Table 1. Comparison of factor analysis and Bayesian Network solution.

FACTOR	DESCRIPTION	SELECTED VARIABLES	
		PCA (Variables ordered by factorial loadings)	B-Course
Factor 1	Transferability of Learning into Situation of Life	V39, V38, V40, V32, V31, V47, V41, V43, V49, V10, V33, V35, V09, V30, V42, V11, V45, V23, V36	
Factor 2	Communication and Interaction with Other Students	V18, V19, V15, V07, V20, V37, V48, V12, V44, V22, V46	
Factor 3	Independence in Planning the Studies	V21, V14, V29, V17, V02, V16	
Factor 4	Learning and Individual Support	V25, V08, V28, V26	
Factor 5	Constructivity	V05, V03, V34, V04	
Factor 6	Learner-centered Approach	V06, V24	

One interesting observation was that the findings of the study indicated that only one variable was omitted by the PCA factor analysis, but up to nine individual unlinked variables were found in Bayesian analysis. Table 2 represents comparison of the factors by principal axis factoring with Varimax rotation and groups by Bayesian analysis.

Table 2. Comparison of the factors and groups by Bayesian analysis.

FACTOR	GROUP	DESCRIPTION
Factor 1 Transferability of Learning into Situation of Life (Extraction Method: Principal Axis Factoring- Rotation: Varimax – 7 factors solution)	7th group – Transfer and contextuality	Intrinsic motivation is stimulated by tasks of optimal novelty and difficulty, relevant to personal interests, and providing for personal choice and control. (Bonk and Cunningham 1998, 29.)
Factor 2 Communication and Interaction with Other Students - Collaborative Learning	4th group – Collaborative learning	In factor analysis both conversational and collaborative variables are combined in the same factor, but in the Bayesian analysis, the two different factors of meaningful learning; conversational and collaborative are differentiated.
Factor 3 Independence in Planning the Studies	1st group – Self-directiveness 6th group – Independence in learning	Active, learners are engaged by learning process in mindful processing of information, where they are responsible for result. Intentional learners are actively and willfully trying to achieve a cognitive object. (Jonassen 1995, 60 – 63.)
Factor 4 Learning and Individual Support	5th group – Teacher support and assistance for learning	Web-based learning environment offers new possibilities for the teacher-student interaction and teachers' role varies from an expert to a facilitator and a mentor of learning. On the Web-based courses as an instructional approach, a cognitive apprenticeship challenges teachers to promote students thinking and teamwork skills. Web offers teachers new possibilities to support collaborative learning processes of students. (Bonk & Reynolds 1997, 172-174.)
Factor 5 Opportunity to Build on Previous Knowledge	2nd group – Constructivity	Constructive learners accommodate new ideas into prior knowledge. (Jonassen 1995, 60 – 63.)
Factor 6 Learner-centered Approach	3rd group – Individuality	Developmental and social factors of learning. Learning is most effective, when differential development across physical, intellectual, emotional, and social domains is taken account. (Bonk & Cunningham 1998, 29.)

Following variables had no links to node variables in the Bayesian analysis: (V04) “I had a chance to influence on given learning tasks”, (V11) “Links in this Web-course helped me to learn new things”, (V13) “I had a chance to study at any place I wanted to”, (V16) “To study in this Web-course activated me to get information independently”, (V23) “Some substances in this Web-course helped me to design my personal curriculum”, (V27) “I was responsible for my own studies”, (V30) “To study in this course developed my criticism”, (V36) “In this Web-course authentic situations presented with the means of multimedia facilitated my learning.” Based on the research evidence it seems that application of constructivist approaches to Web-based learning have both advantages and disadvantages as evaluated by students of the virtual university.

Figure 1 presents all variables as a Bayesian model derived from the data. The final model is the most probable model B-Course [02] could find given the time used for searching. However, there may be other models that are almost as probable as our final model. Natural candidates for other probable models are the ones that can be obtained by slightly changing the final model by removing naïve causal relationships (arcs). The model is named ‘naïve’ because there are no latent (unmeasured) variables in the domain that could cause the dependencies between variables.

We learn from Figure 1 that only four variables are left ‘outside’ the model: (V03) “New subjects covered in learning material were linked to my previous knowledge structures”, (V04) “I had a chance to influence on given learning tasks”, (V05) “I had a chance to utilise my previous knowledge on the subject.” Figure 1 supports our previous model of six dimensions.

Next we examine one factor, “Independence in planning studies”, with B-Course to interpret relationships between variables. Figure 2 presents variable cluster containing following variables: (V01) “I had chance to study by my own personal manners”, (V02) “I was able to get information needed by myself in the network learning environment”, (V14) “I advanced in my studies in my own pace”, (V17) “I directed studies by myself”, (V21) “I planned my own schedules”, (V29) “I progressed in the network course by my own goals from task to another.”

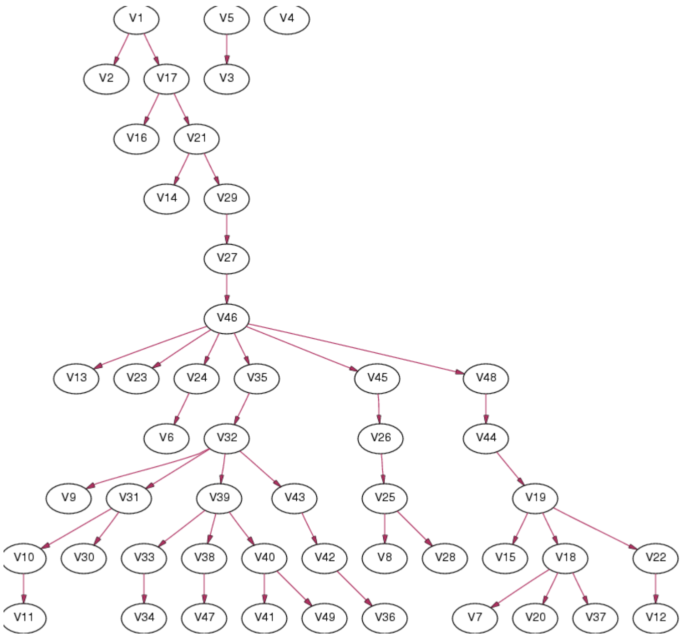


Figure 1. The final model of Bayesian network.

B-Course allows researcher to play a predictive game in which he or she will see how the dependencies cause the knowledge on some matter affect the probability of others.

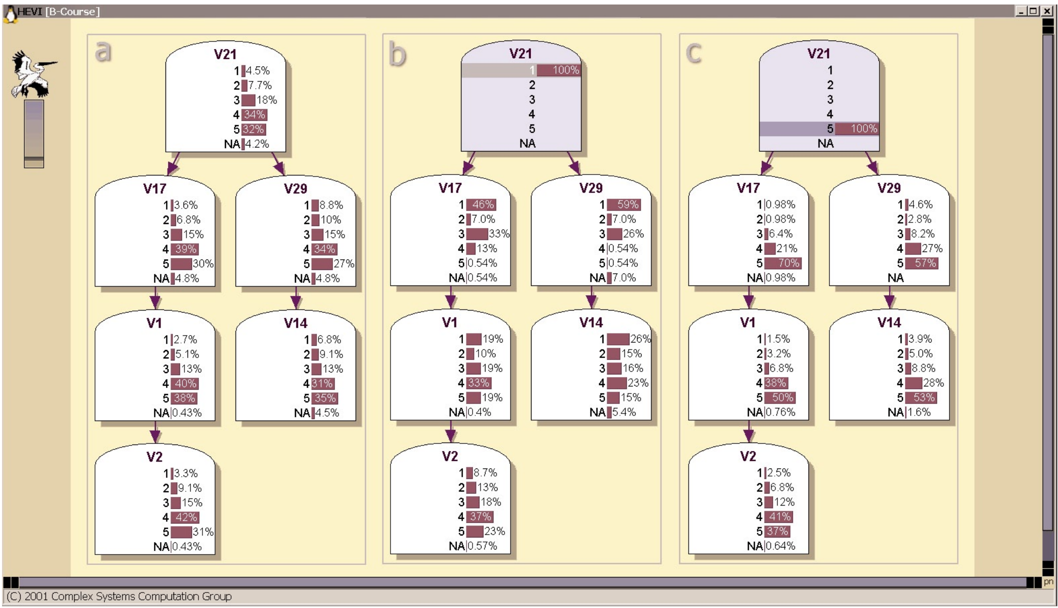


Figure 2. Factor 3 in initialised mode (a) and one variable fixed to negative (b) and positive (c) value.

Figure 2 represents variable cluster "Independence in Planning the Studies" in initialised mode (a) and variable (V21) fixed to values "I was not able to plan my own schedules" (b) and value "I planned my own schedules" (c). We see that the change in variable (V21) has the most dramatic effects on variables (V17) "I controlled my own studies" and (V29) "I progressed in the network course by my own goals from task to another" which could be interpreted as follows: *"When student feel that he or she has no resorts to influence on his or her study schedule to get the information or knowledge needed to complete given task, feeling of controlling ones learning process is getting weaker and the locus of control is moving from human to computer.*"

4 Conclusion

The results of the study have both theoretical and practical value. The empirical results reflect the ideas of constructivist approaches in practice. The results of factor analysis can guide the theory building in this domain. Bayesian networks provide new possibilities for educational researcher to confirm theoretical models and - in addition to traditional factor analysis – allow her to generate and test theoretical scenarios with different data sets. Students' views on learning in virtual university can be used in the process of building more learner-centered approaches to Web-based teaching. The results of this study give information about the preferences in learning by different learners. The respondents represent different types of learners with their own interests. The empirical results show some of these preferences that can be applied in developing the virtual university for the future.

References

Bernardo, J.M & Smith, A.F. 2000. *Bayesian Theory*. John Wiley & Sons: New York. 2nd edition.

Berry, D.A. 1996. *Statistics - A Bayesian perspective*. Duxbury Press.

Bonk, C. & Cunningham, D. 1998. Searching for learner-centered, constructivist, and sociocultural components of collaborative educational learning tools. In Bonk, C. & King, K. (Eds.) 1998. *Electronic collaborators. Learner-centered technologies for literacy, apprenticeship, and discourse*. New Jersey: Lawrence Erlbaum Associates.

Bonk, C. & Reynolds, T. 1997. Learner-Centered Web Instruction for Higher-Order Thinking, Teamwork, and Apprenticeship. In Badrul H. Khan (ed.) *Web-Based Instruction*. Educational Technology publications. 167-178.

Cattell, R. 1978. *The scientific use of factor analysis in behavioral and life sciences*. New York: Plenum Press.

Dede, C. 1996. Distance learning-distributed learning: making the transformation. *Learning and Leading with Technology* 23 (7), 25-30.

Dede, C. 1997. Distributed learning: how new technologies promise a richer educational experience. *Connection- New England's Journal of Higher Education and Economic Development* 12 (2), 12-16.

Jonassen, D. 1995. Supporting communities of learners with technology: a vision for integrating technology with learning in schools*.* *Educational Technology* 35 (4), 60-63.

Kaiser, H. 1970. A second generation Little Jiffy. *Psychometrika* 35, 401-405.

Kirschner, P. 1999. *Using integrated electronic environments for collaborative teaching/learning*. EARLI Keynote speech presented at the 8th Annual Conference of the Association for Research on Learning and Instruction (EARLI 99) held in Gothenburg, Sweden.

Manton, K.G., Woodbury & M.A., Tolley, H.D. 1994. *Statistical Applications Using Fuzzy Sets*. John Wiley & Sons: New York.

Niemivirta, M., Nokelainen, P., Kurhila, J., Miettinen, M., Silander, T. & Tirri, H. 2000. Studying the Role of Individual Differencies in CSCL – The Adaptive Assessment of Motivational Profiles. Submitted to *Euro-CSCL 2001 Conference*.

Ruohotie, P., Tirri, H. & Nokelainen, P. 1999. Professional Growth Determinants – Comparing Bayesian and Linear Approaches to Classification. In P. Ruohotie, H. Tirri, P. Nokelainen & T. Silander (Eds.), *Modern Modeling of Professional Growth*, pp.85-120. Research Centre for Vocational Education, University of Tampere. Saarijärvi : Saarijärven Offset.

Ruohotie, P. & Nokelainen, P. 2000. Modern Modeling of Student Motivation and Self-regulated Learning. In P. R. Pintrich & P. Ruohotie (Eds.), *Conative Constructs and Self-regulated Learning*. Research Centre for Vocational Education, University of Tampere. Saarijärvi : Saarijärven Offset.

Internet References

[01] Helsinki Virtual Open University (HEVI)   http://www.avoin.helsinki.fi/english.htm

[02] B-Course   http://b-course.cs.helsinki.fi