THE ANNALS OF ”DUNAREA DE JOS” UNIVERSITY OF GALATI
FASCICLE III, 2005 ISSN 1221-454X
ELECTROTEHNICS, ELECTRONICS, AUTOMATIC CONTROL, INFORMATICS

UNCERTAINTY MANAGEMENT USING BAYESIAN NETWORKS IN STUDENT KNOWLEDGE DIAGNOSIS

Adina COCU, Diana STEFANESCU

University “Dunarea de Jos” of Galati, Department of Computer Science,
str. Domneasca 47, Galati, Romania

Abstract: In intelligent tutoring systems, student or user modeling implies dealing with imperfect and uncertain knowledge. One of the artificial intelligence techniques used for uncertainty management is that of Bayesian networks. This paradigm is recommended in the situation when exist dependencies between data and qualitative information about these data. In this work we present a student knowledge diagnosis model based on representation with Bayesian networks. The educational system incorporate a multimedia interface for accomplishes the testing tools. The results of testing sessions are represented and interpreted with probability theory in order to ensure an adapted support for the student. The aims of the computer assisted application that contains this diagnose module are to support the student in personalized learning process and errors explanation.

Keywords: numerical uncertainty management, intelligent computer based learning system, Bayesian network, probability theory, student diagnosis.

1. INTRODUCTION

The intelligent computer based learning systems try to adapt its modules to the student needs. The educational systems must be able to manage with uncertainty in inferential process about student knowledge.

An educational system is composed through a set of modules: learning domain module, student/ user module, evaluation/ diagnosis module, pedagogical module, and user interface. These modules aren’t always implemented like independents modules they can be combined. One of these modules is possible to contain uncertain information, especially student diagnosis module. In psycho-pedagogical research fields (that are related to the educational systems research) have been investigated some aspects about imperfection in every day life, the type of the errors make by the humans. In conclusion it is a challenging work when dealing with imperfection and uncertainty associated with large amounts of data for process observations about persons.

One of the most important features of intelligent computer based learning system is adaptation to the student. To make this task possible, the systems must know the student knowledge state very well. The way to do this is to diagnose the student with help of tests. The learning system makes inferential processes in order to obtain student knowledge state, based on tests results.

In this paper we present a model of a student diagnosis module based on representation of dependencies between learning domain subjects and on probabilities associated with student knowledge assessment task. This model can be applied to declarative learning domains structured in granular structures and it uses evidence propagation in Bayesian network for inference.

2. GRAPHICAL REPRESENTATION OF ASSESSING KNOWLEDGE

The type of learning domain is important from the representational point of view. The model of learning domain knowledge must be chosen in accordance with the type of learning content. In order to provide individualized instruction, the learning systems must adapt learning content to the student, in accordance with what he/she wants (applications, theory). For an

This paper was recommended for publication by Severin BUMBARU

efficiently approach the learning domain is structured to the different thoroughly levels modules and a study curriculum is made by modules combination. *Structural* representation of learning knowledge and educational process are closely link. One consideration is that it is impossible to understand a concept or an item without to have the possibility to identify and classify them and without establish the connection with others items (Jonassen H., Beissner K., Yacci M., 1993). For our case study we consider a structural learning domain. This domain can be split into more granular learning resources (sub-modules) and represented with a graphical model. We choose a schematic representation of learning domain because the schemes play an important role in the human cognitive process (Rumelhart, Ortony, 1977).

In the learning domain literature the domain modules are designed with different knowledge representation formalism likes: semantic nets, conceptual maps, causal diagrams, deduction graphs, rules and propositional logics.

We choose to represent domain content with a directed acyclic graph. That can be usually easy implemented in practice with a XML structure that can be shared with other learning system tasks or other systems. The learning items are placed into the node of graphical model. Items are linked between them with arc. Arcs are oriented so they suggest that a parent node contain more child nodes that are sub-module of them.

In our case we consider that learning *Course* is split in more *Sections*. Each section must have both, theoretical and applicative parts. These parts can contain more or less sub-modules, which, on the other hand, can be respectively divided in a more granular manner. For example, we consider a section that contains theoretical knowledge and applicative knowledge. The student can be asked to resolve two problem notated Kai, for applicative knowledge part. Theoretical aspects are divided into three parts (KTi) that can be verified through associated questions, notate with KTij in Figure 1.

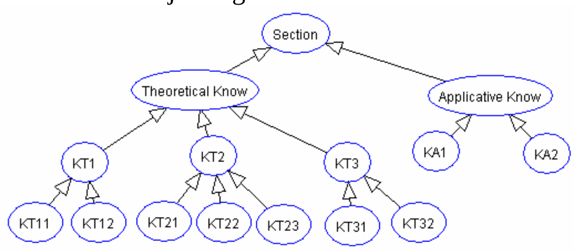

Fig.1 A domain model prepared for diagnosis module

The advantage of graphical model is it can be easily adapted for any learning domain structure. Here, we make an observation: if a parent node has more than three children then it is needed a large amount of data (not always available) and the computational time increase exponentially.

Each student learns in an individual manner, in concordance with his/her personality, learning domain type, difficulty level. After student assessment, in order to be user adaptive, the educational system must provide student error feedback and suggest the weak points in educational par course. In presented case this can be done with support of granular knowledge representation.

The evaluation module observes learner behavior, purchase assessments tasks corresponding with each knowledge level and evaluate student at an established moment. The courseware granular modularization and evaluation methods are interconnected. The knowledge representation method influence the diagnose methods of student knowledge state, also the structured information communication methods are involved in educational process. Closely linked with these methods are also learning strategies applied in learning process.

The evaluation process must create assessments information that is used in inference about current student state. The knowledge assessment task of a learner can be defined in three steps: obtaining of student knowledge with different techniques, structural representing of obtained information and comparing between these representation and domain knowledge expert. The training system must be capable to implement different assessment techniques like quizzes, multiple choice questions with either a "true or false" or "agree or disagree" responses, multiple choice question with one or more items, exercise resolving with implies numerical notation of the results.

It is a known the fact that evaluation methods can be considered learning methods, thus they have a formative implication. So, after an evaluation session, the student will know what items of educational par-course he/ she know less then normal or he/she have misunderstandings. Also the user can recognize his progress in informational gain. Thus the student is stimulated to learn exactly what he/she needs for proposed aims of his/her studies. This is the reason why we consider capability to provide explanation about errors and misconceptions an important feature of adaptive learning systems. In our application we can exactly specify what parts of course must be remake, because of dependencies graphical representation.

The learning errors can be combined between them and a wrong answer may have multiple diagnoses or there is not a unique answer for an item. This is the reason why we choose an uncertainty representational model. We implement the uncertainty management with probabilistic representation. In next section we make a short presentation of theoretical probability aspects used in Bayesian networks representation.

3. USE OF BAYESIAN NETWORK IN STUDENT DIAGNOSIS

So we saw in previous section the learning domain items are represented into a causal directed acyclic graph. If we associate to each node some numerical qualitative measures it is obtain a Bayesian network which is able to make inference. The Bayesian network is powerful representational formalism, based on probability theory.

Formally, a Bayesian network is defined by a set of variables and a directed acyclic graph defining the model of conditional dependencies among the variables. From computational point of view we consider discrete variables with a finite state number. A conditional dependency arc links a child variable to a set of parent variables and is defined by the conditional probability distribution of a child given the combination of all parent values.

In our example we consider that KTij are the parents of KTi, KTi are the parents of Theoretical Knowledge node, KAi are the parent of Applicative Knowledge, Theoretical and Applicative Knowledge are the parent of Section node.

KTij are variables that represent “true or false” or “agree or disagree” responses or multiple choice question with one or more choices. All the other variables can have discrete values corresponding to our numerical school notational interval like “unknown”, “very low” (<4), “low” (5-6), “medium” (7-8), “high” (9-10).

For the KTij and KAi variables we specify prior probability in the next tables:

	P(KTij=T)	P(KTij=F)
Prior probb.	0.5	0.5

	Prior probability
P(KAi=unknown)	0.2
P(KAi=vlow)	0.2
P(KAi=low)	0.2
P(KAi=medium)	0.2
P(KAi=high)	0.2

Then we specify the conditional probability distribution at each child node. The values of these probabilities (marginal or conditional) can be obtain from the learning domain expert or tutor based on their experience. As we will see below, the more number of values has the parents the much more conditional probabilities has the child. In our case for a child with five possible values and two parents with two values we must complete $5*2^2$ values. For applicative knowledge node that can have five probably values and that have two parents, each of them with five values, then the conditional probability table must have $5*2^5$ cells.

We observe that we need a large amount of numerical data. This fact is difficult in practice and it is not always possible to have the exactly match values. In the case when we have a large, bun not complete set of data, we can apply on these data a decision tree, which is able to compute probability distribution. Decision trees are a simple yet successful technique for supervised classification learning based on experience training data set.

In our example, suppose we have the tutor experience express through next training data:

KTi1	KTi2	KTi
true	false	medium
true	false	low
false	false	vlow
false	true	low
false	false	unknown
true	true	high
false	true	medium

Constructing on these training data a decision tree using splitting function with information gain conclude to the next calculated probabilities:

KT11	T	T	F	F
KT12	T	F	T	F
P(KT1=unknown)	1.0	0.0	0.0	0.5
P(KT1=vlow)	0.0	0.0	0.0	0.5
P(KT1=low)	0.0	0.5	0.5	0.0
P(KT1=medium)	0.0	0.5	0.5	0.0
P(KT1=high)	0.0	0.0	0.0	0.0

This conditional probability table is introduced like prior information in Bayesian network.

The same algorithm may be applied for all the nodes for we have to specify conditional distributions.

The aim of Bayesian network inference is to propagate evidence. Thus we introduce possible values for the theoretical questions and application results and the system reason the probability distribution for the Section node. In inference task the Bayesian network specifies a complete joint probability distribution over all the variables. With this joint probability it is possible to answer to all possible inference queries by marginalization.

In our case the joint probability
P(KT11,KT12,KT21,KT22,KT23,KT31,KT32,
KT1,KT2,KT3,KT,KA1,KA2,KA,S)

can be decomposed into a set of independent parent-child contributions as:

P(...)=P(KT11)*P(KT12)*P(KT1|KT11,KT12) *P(KT21) *P(KT22)*P(KT23)*P(KT2|KT21,KT22,K23)*P(KT31) *P(KT32)*P(KT3|KT31,KT32)*P(KT|KT1,KT2,KT3) *P(KA1)*P(KA2)*P(KA|KA1,KA2)*P(S|KT,KA)

The Bayesian network inference may be of two kinds:

- predictive – from causes to effects;
- diagnostic – based on the effects search the causes;

In our example we use Bayesian network to make predictions about values of a variable in a given situation specify through evidence. We make inference by computing conditional probability distribution of the variable given the values of a set of other variables in the network. For instance we are interested in the values of Section variable when the variable Ktij and Kti are observed to be in a set of states. We apply the total probability theorem to compute the marginal probability P(S):

$$P(S) = \sum_{ij} P(S \mid KT_i, KA_j) * \sum_{klm} P(KT \; | KT_{1k}, KT_{2l}, KT_{3m})$$
$$* \sum_{no} P(KA \; | KA_{1n}, KA_{2o}) * \sum_{pr} P(KT_1 \; | KT_{11p}, KT_{12r})$$
$$* \sum_{stv} P(KT_2 \; | KT_{21s}, KT_{22t}, KT_{23v}) * \sum_{xy} P(KT_3 \; | KT_{31x}, KT_{32y})$$

For example, if we have a set of evidences:
KT11=T, KT12=F, KT21=T, KT22=F, KT23=T, KT31=T, KT32=T, KA1=high, KA2=low
the results for Section node probability are presented in the graphic below:

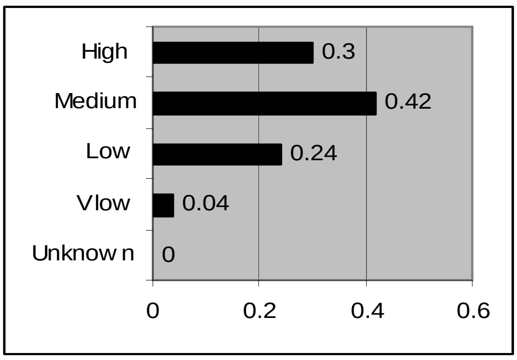

Fig. 2 Probability Section chart

We interpret these results in this way: it is more probable that the student knowledge level is medium than low or high, because the probability for medium is the greatest from all the others numeric values.

4. CONCLUSIONS

We perform more experiments and we use the gathered information for evaluate the Bayesian inference against classical evaluation method using the same assessment questions. The output of the diagnostic module concerning student knowledge level compared with the assessment of a tutor give in the most cases similar results.

We can conclude that the Bayesian network method for uncertainty management is near by the human treatment of uncertainty, but it has the advantage to give additional information about the errors or mistakes causes and a subtle distribution about student knowledge level.

5. REFERENCES

- Grigoriadou M., Kornilakis H., Papanikolaou K., Magoulas G., "Fuzzy Inference for Student Diagnosis in Adaptive Educational Hypermedia", Springer-Verlag, 2002.
- Guzman E., Conejo R., "A Model for Student Knowledge Diagnosis Through Adaptive Testing", Springer-Verlag, 2004
- Jameson A., "Numerical Uncertainty Management in User and Student Modeling: an overview of systems and issues", Pracma, 1995Ramoni M., Sebastiani P., "Bayesian Methods in Intelligent Data Analysis", Springer-Verlag, 2004.
- Jonassen H., Beissner K., Yacci M., „Structural Knowledge – Techniques for Representing, Acquiring Structural Knowledge", Lawrence Erlbaun Associates, 1993
- Murphy E., "A Brief Introduction to Graphical Models and Bayesian Networks", Berkeley, 2000.