Visualization of Bayesian Learner Models

Jim E. Greer, Juan-Diego Zapata-Rivera, Carina Ong-Scutchings, John E. Cooke
ARIES Laboratory
Department of Computer Science
University of Saskatchewan
Saskatoon, Canada

Abstract. Bayesian Belief Networks (BBNs) have been suggested as a suitable representation and inference mechanism for learner models (Reye, 1998). Having the goal to construct inspectable Bayesian learner models, we have encountered two problems. The first is the problem of visualization of BBNs and the second is the problem of using BBNs to model changes in knowledge state over time due to gradual learning and forgetting. Both of these problems are being addressed in current research at the ARIES Laboratory and both are briefly discussed in this paper.

1 Introduction and Motivation

Bayesian Belief Networks (BBNs) have become accepted and used widely to handle uncertainty situations and cause - effect relationships. BBNs have been used in many areas, such as: diagnosis of medical patients, diagnosis of malfunctioning systems, planning in uncertain domains, speech recognition and understanding of stories. The causal information encoded in BBNs facilitates the analysis of action sequences, observations, consequences, and expected utility (Pearl, 1997). BBNs have been suggested as a suitable representation and inference mechanism for learner models in various traditional ITSs (Reye, 1998) and have been used as the basis for cognitive modelling in the Self-Explanation Coach (Conati, et.al, 1997).

As a mechanism to visualize causality and probabilities, BBNs offer an intuitive approach where causes and effects are represented by circles (nodes) and arrows are used to connect each cause with its effects. Using conditional probabilities for each node based on its direct dependencies, it is possible to propagate probability values (Russell & Norvig, 1995).

We have recently developed a system for visualization of BBNs named VisNet. The VisNet system is built upon the SMILE Bayes Net Library (from the University of Pittsburgh) combined with the OpenGL graphics library and Togl (a Tcl/Tk widget for graph drawing).

VisNet provides various visualization and animation techniques to illustrate the propagation of beliefs in a BBN. If the BBN represents a learner model overlaid on a concept map (with nodes corresponding to concepts and links indicating learning sequences or aggregations or abstractions), and knowledge levels of concepts represented as probability values, VisNet can assist in visualizing the learner model.

2 Visualization Techniques in VisNet

One visualization technique in VisNet uses primary colours and their combinations to represent causes and effects. Colours of nodes are determined based on parents' colour. For example, if parents are blue and red, their children are magenta; if parents are blue and yellow, their children are green. Colour intensity is used as indicator of the marginal probability (strength of belief that the concept is known). Colour (hue) of each node can vary according to the variation of the parents' probabilities and the relative influence of the parent nodes on a descendant node. In this way, colour variation of each node can be attributed to the parent (cause) that seems to be having the stronger affect on the beliefs about concept knowledge.

Another technique uses sizes of nodes to indicate strength of belief that a concept is known. Smaller spheres indicate lower marginal probabilities. A third technique uses proximity of nodes in the network (length of the links) to indicate strength of association between nodes. Nodes linked by higher conditional probabilities are in closer proximity.

Combining these techniques of colour, node size and proximity, one can produce a relatively intuitive picture of the beliefs represented in the BBN. It is possible to see which concepts are closely related to other concepts and the concepts thought to be known and not known by the learner. Figure 1 contains a screenshot from VisNet for a fragment of a learner model, a larger sized node would be one that is more likely to be known. Nodes in close proximity are connected with stronger influence, and colour mixing indicates the relative influence of various ancestors on a node.

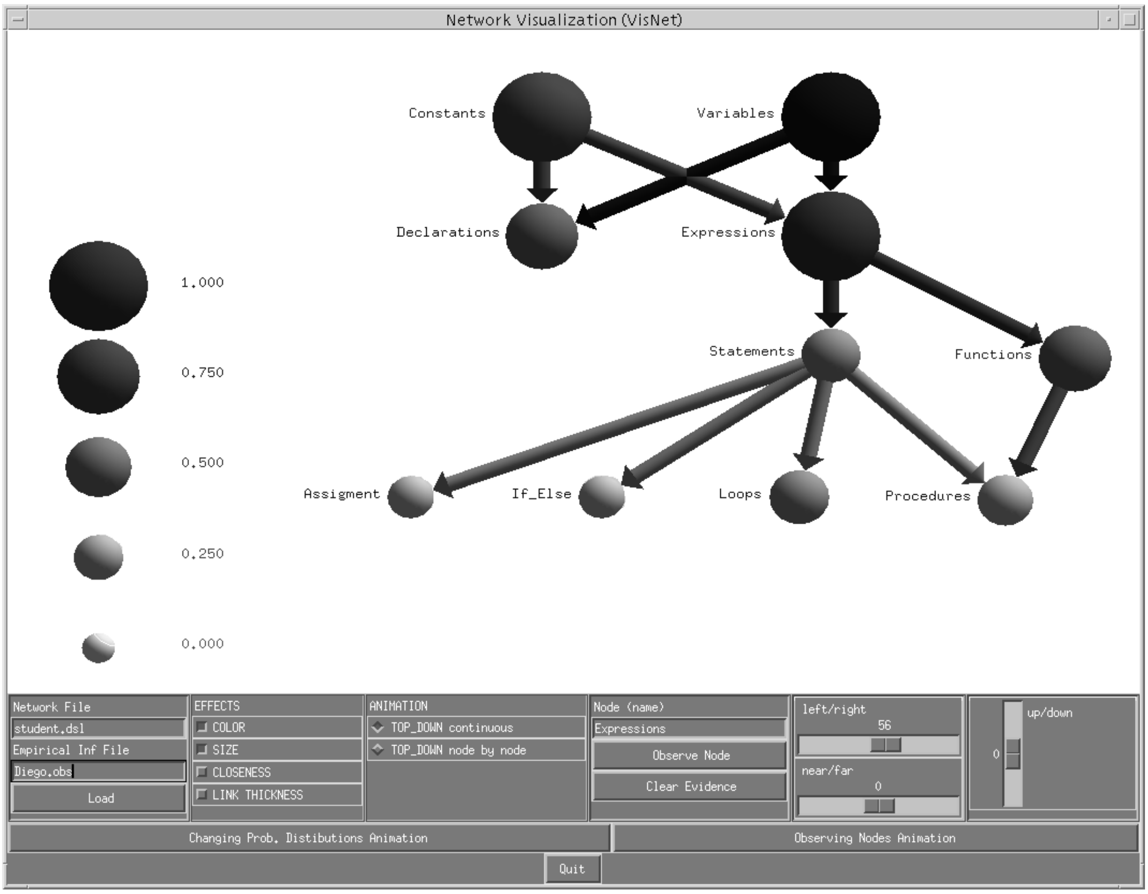

Figure 1. Learner Model Visualization.

A more interesting feature of VisNet is its ability to animate changes in marginal probabilities (belief propagation) in a BBN. For example, when a concept is observed to be known (perhaps as a result of some external evidence such as a quiz item or success on an assignment) beliefs within the BBN are propagated to reflect the new knowledge. This propagation is graphically animated in VisNet to illustrate what impact the newly gained knowledge about the learner has on the overall learner model. That is, steps in the propagation of belief across the BBN are displayed by repeatedly redrawing each node (changing slightly the size colour or proximity). This animation effect has the potential to dynamically depict learning as it occurs. We believe this will have an interesting effect on a learner who is interacting, say, with an ITS and watching the system's beliefs evolve about his or her knowledge.

It must be pointed out that VisNet has not yet been used in conjunction with an ITS to animate an open learner model in an ITS. The visualization tool has only recently been built and it will take some time to integrate into a running ITS. Nevertheless, this concept of animating the evolving learner model has been connected with a second research initiative described next.

3 Modelling the Effect of Gradual Learning and Forgetting

Learning does not happen instantaneously. Rather, learning is cumulative, with knowledge about a concept growing with time as the concept is being taught and as it is practiced. It has been suggested that people gradually forget details of an event and only remember relevant parts of an experience. Learning occurs when a relevant but novel event is experienced, modifying the general idea of the event (Schank, 1984).

Snapshots might be taken to assess a learner's knowledge at some point in time by using quizzes or examinations. On the other hand, if the concept is not being taught and not being practiced, the learner may tend to forget what had been learned. This leads to a very simple model of expected learning in the absence of any evidence. That is, while a concept is being studied or taught, learning at some rate is likely to occur. After a concept has been studied or taught, forgetting will tend to occur (especially if subsequent concepts do not reinforce that original concept). An approach called *data aging* (Webb & Kuzmycz, 1998) attempts to take into account the importance of recent observations compared to older ones by adjusting weights in a student model. A simple transient learning and forgetting cycle will tend to occur for most learners.

Incorporating even these simplistic learning and forgetting cycles into a Bayesian learner model provides an interesting challenge. Given a learner model that is an overlay on a concept map (as suggested above), and given a schedule for the teaching or study of the various concepts, one can determine which concepts are likely to be in a "pre-learned" state, which are likely to be in a "learning" state and which are likely to be in a "forgetting" state. Assuming simple learning and forgetting curves can be associated with concepts, a Bayesian learner model should be able to capture this behaviour.

Dynamic Bayes Nets (sometimes called temporal Bayes Nets) could be used for this purpose, but static BBNs can also be used to simulate this sort of behaviour. This is accomplished by providing to the BBN an input regimen of external evidence corresponding to the learning and forgetting curves expected to be in effect as time progresses.

In our first prototype of this approach, three distinct layers of BBNs are used to model the learner. The first layer is the Concept Map Layer. Each node in this layer carries the belief measure of the learner's knowledge of that concept. In a sense, each node holds an estimate of the probability that the learner knows the Concept - denoted $P(K(C_i))$. This layer is the main layer appropriate for visualization since changes here would show the student's learning and retention of knowledge.

The second layer is a Skills Layer. Here each node corresponds to a belief that a learner can successfully carry out some task or achieve mastery on some quiz (and is called a "CanDo" node). The underlying assumption is that if $P(K(C_i))$ is sufficiently high, this should affect the belief that $P(CanDo(Skill_i))$ as well. This reasoning could also work the other way. If there is the observation that the student can perform a skill associated with a concept, then $P(K(C_i))$ should be positively affected.

The third layer, the Schedule Layer, is comprised of several Learning Schedule nodes, each paired with the corresponding Concept node. Each Learning Schedule node may hold the value $Learning(C_i)$ or $Forgetting(C_i)$, which represents the belief that the student is currently learning the concept or currently forgetting the concept. The belief measure of $Knows(ConceptX_i)$ is affected by the value of its corresponding Learning Schedule node according to a specified learning or forgetting function.

These three layers can model with a BBN the effect of gradual learning and forgetting of concepts over time according to a schedule of learning events or teaching events. This provides us with a stereotypical learner model that can be continually updated through a course of study independent of evidence about the individual learner. Of course, if there is observed evidence for the individual's ability to do a skill or knowledge of a concept, this evidence can be directly asserted in the BBN.

There are specific, desirable behaviours incorporated when designing the BBN learner model. These models should capture behaviours such as: effects of prerequisite concepts or topics on subsequent concepts, reinforcement effect of the subsequent concept over the prerequisite concept, and positive effect of $P(CanDo(Skill_i))$ on the corresponding $P(K(C_i))$. The joint conditional probability values were constructed having these desired effects in mind. By running simulations of typical student behaviour on VisNet, intuitive notions of probability were verified and fine tuned.

VisNet was a useful tool in building and tuning this model of learning and forgetting. In fact, changes in prior probability and their effects in the whole network could be visualized immediately. This feature simplifies the refinement phase of the engineering of the learner model.

4 Conclusion and Future Work

Taken together, these two projects provide the basic tools to model gradual learning and forgetting in a Bayesian learner model and to provide an inspectable graphical animated representation of the learner model. With such a system, the possibility exists to record learner model data throughout a course, observe snapshots of the learner model at any time, and run animations of the accumulation of evidence as the learning (and forgetting) progresses. This evidence can be informed by observations associated with explicit assessment activities or with observations stemming from passive assessment done by an ITS.

This work is in its early stages. We have only begun to scratch the surface of the issues associated with animating the evolving beliefs represented in Bayesian learner models. In the current model, there are no dependecies among the Skill nodes. The model is further simplified by having a one to one relationship between $P(CanDo(Skill_i))$ and $P(K(C_i))$ nodes. Intuitively, this corresponds to having exactly one quiz to demonstrate knowldege for each specific concept. Further work can be done by enhancing the model to have a richer skills layer which is interconnected with true dependency links.

Beyond its potential for visualizing Bayesian learner models, VisNet is a useful tool for engineering and tuning BBNs and for engineering and maintaining BBN representations of learner models. Although the first version of VisNet was developed as a general BBN visualization tool, special interfaces are being developed to allow teachers and students to inspect and modify the learner's model. Future work might include refining VisNet into an authoring tool for creating concept maps and Bayesian learner models.

5 References

(Conati et al, 1997). Conati C., Larkin J., VanLehn K., *A Computer Framework to Support Self-Explanation* In du Boulay B.and Mizoguchi R. (eds) Artificial Intelligence in Education (AIED '97), 279-286.

(Pearl, 1997). Pearl, J., *Graphical Models for Probabilistic and Causal Reasoning* in The Computer Science and Engineering Handbook, Editor Tucker, A., CRC Press, Boca Raton, FL, 699-711.

(Reye, 1998). Reye, J., *Two-Phase Updating of Student Models Based on Dynamic Belief Networks.* In Goettl, B., Halff H., Redfield C., and Shute V. (eds). Intelligent Tutoring Systems (ITS '98), 276-283.

(Russell & Norvig, 1995). Russell, S., and Norvig, P. (1995). *Artificial Intelligence: A Modern Approach*. Prentice Hall, New Jersey, 436-470.

(Schank, 1984). Schank R., *Dynamic Memory: A Theory of Reminding and Learning in Computers and People*,, Cambridge University Press, New York, NY, 12-14.

(Webb & Kuzmycz, 1998). Webb G. and Kuzmycz M., *Evaluation of Data Aging: A Technique for Discounting Old Data During Student Modeling.* In Goettl, B., Halff H., Redfield C., and Shute V. (eds). Intelligent Tutoring Systems (ITS '98), 384-393.