Visualization of Bayesian Belief Networks

Juan-Diego Zapata-Rivera, Eric Neufeld, Jim E. Greer
Department of Computer Science. University of Saskatchewan.
Saskatoon, Canada

Abstract

Concepts like marginal probability, changes in probability, probability propagation and cause-effect relationships are important when reasoning about causality and uncertainty.

To help people to understand these concepts, a Bayes net visualization tool (VisNet) has been developed. VisNet shows how temporal order, colour, size, proximity (closeness), and animation techniques can be used to visualize each of these probabilistic concepts. In addition, a usability study using simple BBNs was conducted to determine how people react to each of the visualization techniques.

This paper describes the visualization techniques in VisNet and reports the main results found in the usability study.

Keywords

Bayesian Belief Networks, visualization applications, visualization techniques, temporal order, colour, size, proximity (closeness), and animation.

1. INTRODUCTION

Bayesian Belief Networks (BBNs) have become accepted and used widely to model uncertain reasoning situations and cause - effect relationships. BBNs have been used in such areas as: diagnosis of medical problems, diagnosis of malfunctioning systems, planning in uncertain domains, speech recognition, user modelling and story understanding. The causal information encoded in BBNs facilitates the analysis of action sequences, observations, consequences, and expected utility [1].

As a mechanism to visualize causality and probabilities, BBNs offer a relatively intuitive approach where causes and effects are represented by circles (nodes) and arrows are directed from each cause to its effects. Using conditional probabilities attached to each node based on its direct dependencies, it is possible to propagate changes probability values on receipt of evidence[2].

Although BBNs originated in the AI/CS community as an effective computational tool for manipulating joint distributions of many variables, BBNS are beginning to be seen by some philosophers and social scientists (e.g. Cartwright) as providing a complete treatment of path analysis (Sewell Wright), which has significant application in the social sciences.

This paper focuses on BBNs visualization. Although simple static directed graphs can convey a lot of information, an ordinary user of BBNs can easily be overwhelmed. This paper reports on the utility of temporal order, colour, size, proximity (closeness), and animation techniques for helping people understand concepts inherent in BBNs such as marginal probability, changes in probability, probability propagation and cause-effect relationships.

2. VisNet (NETWORK VISUALIZATION)

VisNet was implemented using C++, SMILE (a Bayes net library) [3], OpenGL (Mesa libraries) [4], and Togl (a tk widget) [5]. VisNet can use different BBN formats, including: dsl, dsc, net, dne, and erg. With VisNet, it is possible to observe different nodes (that is, receive evidence) and visualize how probability values subsequently change through the net. The present version runs on Sun Solaris, since it can be easily ported to any other platform (e.g., PC). VisNet visualizes probability changes with such visualization techniques as: colour, size, closeness, link thickness, and animation.

3. VISUALIZATION IN VisNet

The different visualization techniques in VisNet can be applied singly or in combinations.

3.1. Temporal Order Technique

Temporal order is a way of laying out a Bayes net for a user to most naturally understand cause-effect relationships. In this technique, causes are placed in the first level (top level), and their immediate effects directly underneath, and so on, until leaf nodes are placed in the last level (bottom level). This arrangement makes the graph easy to read (Top-down) from causes to effects (See Figure 1).

Temporal order can be seen as an example of layered drawing of acyclic digraphs. Hence, algorithms to reduce edge crossings, assign layers, and assign horizontal coordinates can be used to make the graph more readable [6].

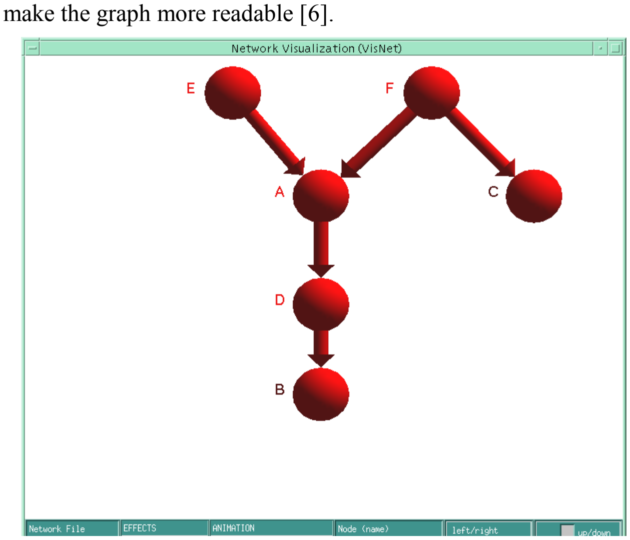

Figure 1: *Temporal order.*

3.2. Colour Technique

This colour technique uses primary colours and their combinations to represent causes and effects. Colours are assigned based on parents' (i.e., causes') hues; saturation is based on the probability of the node. For example, if parents are blue and red then their children are magenta; if parents are blue and yellow then their children are green, etc.

Colour intensity indicates the marginal probability based on parents' probabilities. For example, Figure 2 shows that "Tampering" can have probability values from 0.0 (light red) to 1.0 (dark red).

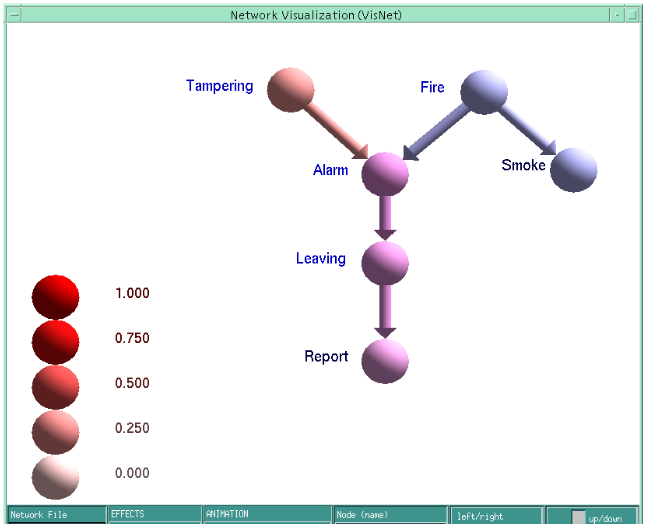

Figure 2: *Colour technique.*

Thus, a node with a single parent inherents the parent's hue, but its intensity is a function of its marginal probability.

Colour (hue) of a node varies with variations of the parents' probabilities and the relative influences of parent nodes on a descendant node. In this way, colour variation of a node can be attributed to the parent (cause) that appears to have the stronger affect according to its conditional probability distribution. Figure 3 shows how "Alarm" can take on different shades of magenta (changing hue from red to blue) according to the parents' probabilities and its conditional probability distribution.

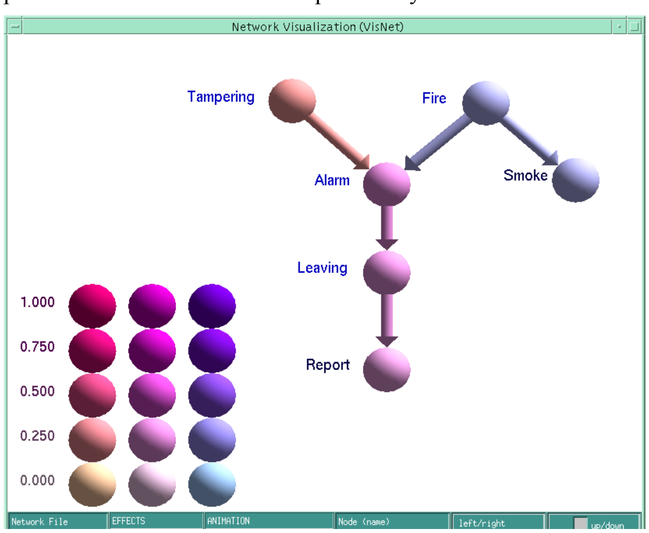

Figure 3: *Alarm using colour (two parents).*

3.3. Size Technique

In this technique, node size represents marginal probability. Small spheres indicate probability values near to 0.0, and big spheres indicate values close to 1.0. Figure 4 shows a Bayes net with node probabilities between 0.0 and 1.0.

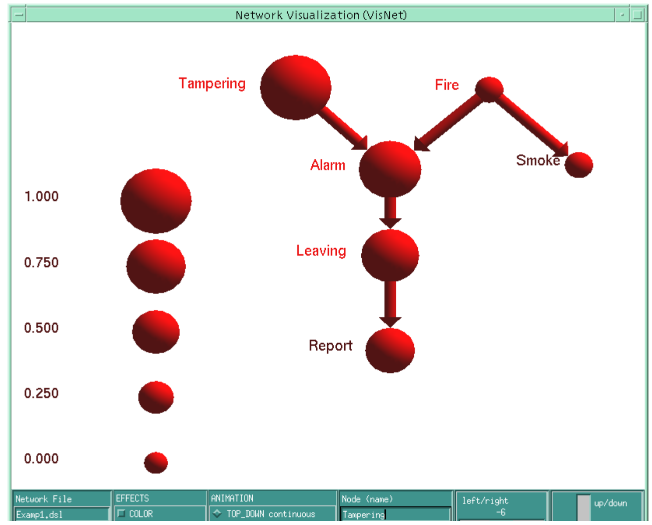

Figure 4: *Size and probability.*

3.4. Proximity Technique (Closeness)

This technique uses the distance between nodes to represent the strength of their cause-effect relationship. Hence, if the probability of a node given its parent is high, the two nodes appear near each other; otherwise, the two nodes appear far apart.

In this technique, after a node has been observed (new evidence is available), nodes relocate according to the new probability values propagated through the net. Figure 5 illustrates such relocation.

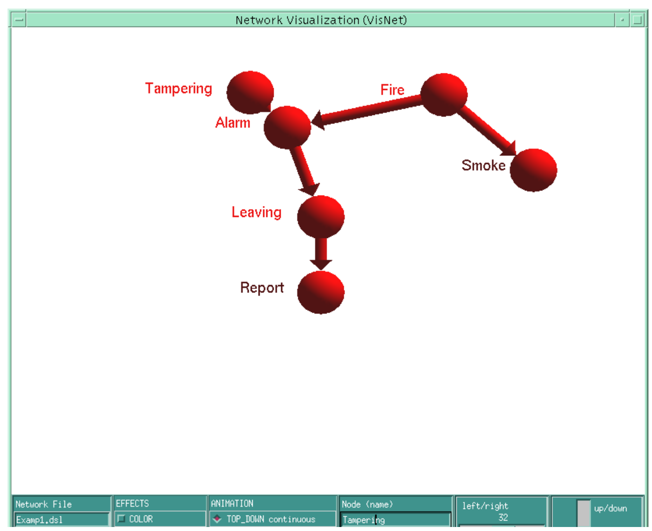

Figure 5: *Closeness before "Tampering" was observed.*

3.5. Mixing Techniques.

Different authors have pointed out potential risks when using colour, such as: "some colours can cause visual discomfort" [7], "Eight percent of males and 0.5 percent of females have colour deficient vision", and "cultural differences can affect the number and categories individuals recognise" [8]. In consideration of this, it is desirable to provide redundant information (reinforcement) though combination of different techniques. Some combinations of techniques like colour and size, size and proximity, colour and proximity, and colour, size and proximity were tested during a usability experiment. (See Figure 6).

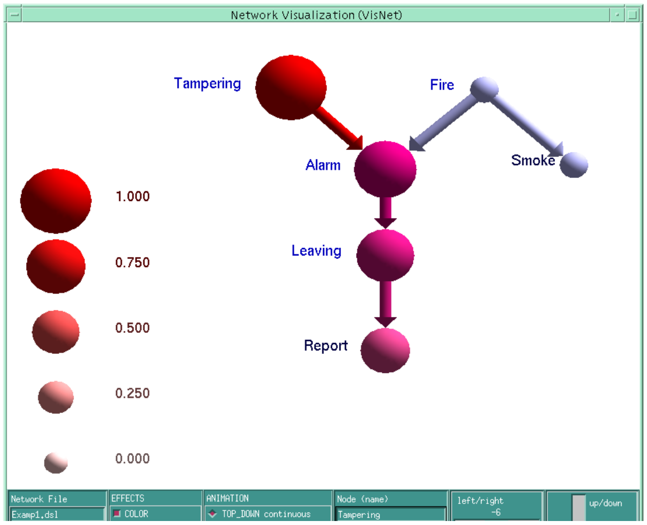

Figure 6: *Colour and size after "Tampering" was observed.*

3.6. Animation

Animation techniques can dynamically illustrate changes in probability when new evidence is presented. Two kinds of animations were explored: continuous TOP-DOWN, and node by node TOP-DOWN.

Continuous TOP-DOWN animation shows how the whole net changes using a small probability increment, starting from an initial state (no new evidence), updating probabilities, and redrawing until a final state (new evidence) is reached. Node by node TOP_DOWN animation uses the same technique, but it shows the probability propagation node by node trough the complete net. Both, individual effects (colour, size or closeness) and their combination are available in either of the animation styles. Figures 7 and 8 show a continuous animation of the coma or cancer network[1] when a brain tumour is being observed. (An included AVI animation file is further illustrates these animation effects.)

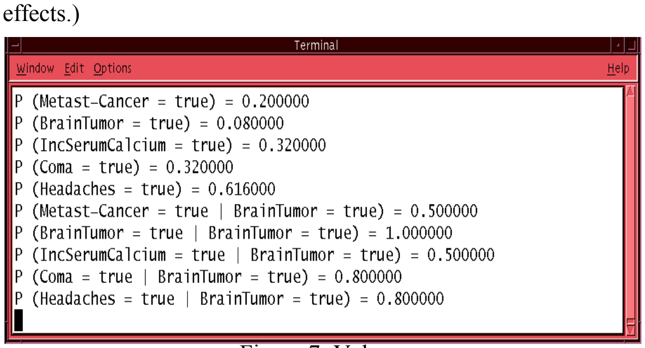

Figure 7: Values.

[1] The coma or cancer network appeared first in Greg Cooper's doctoral dissertation: Cooper, Gregory F. (1984). NESTOR: A computer-based medical diagnostic aid that integrates causal and probabilistic knowledge. PhD thesis, Medical Information Sciences, Stanford University, Stanford, CA, 1984.

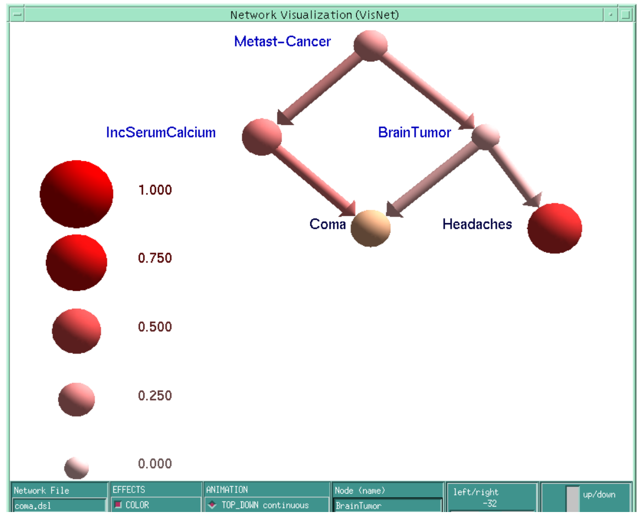

Figure 8a.

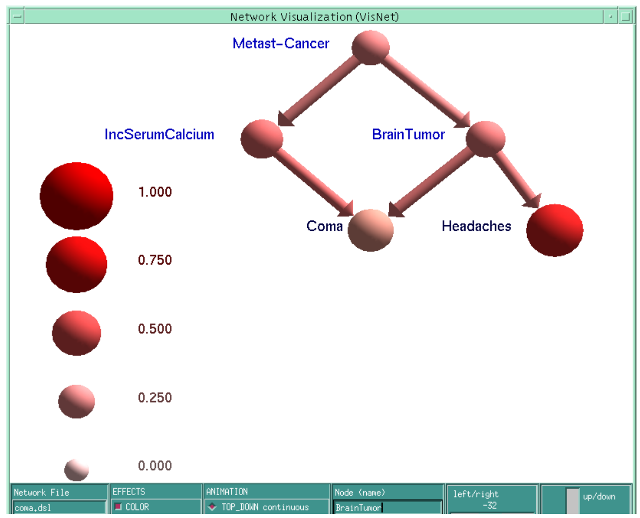

Figure 8b.

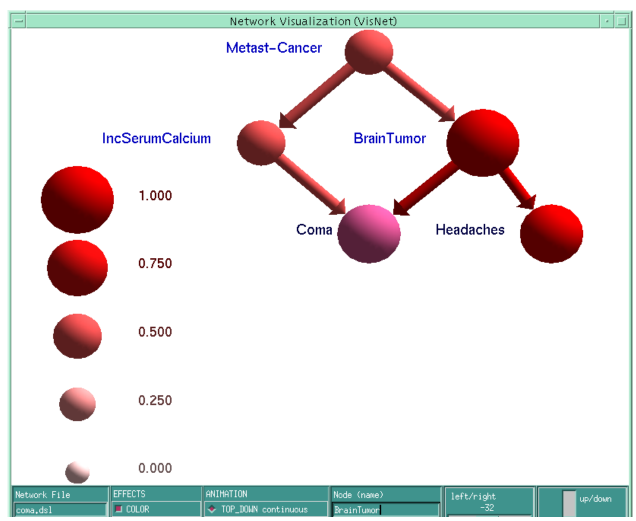

Figure 8c.

Figures 8a, 8b, and 8c continuous animation on observing brain tumour.

4. USABILITY STUDY SETUP

A simple usability study was conducted to test VisNet. Participants were ten Computer Science graduate students. They were introduced to BBNs using a simple example and a short explanation about cause-effect relationships and directed acyclic graphs (DAGs).

5. TASKS

To determine how useful each technique was for representing marginal probability, changes in probability, probability propagation and cause-effect relationships, several tasks were developed. In addition, participants' comments and suggestions were carefully studied to determine useful extensions to this work.

Participants were asked general and specific questions about what was intended to be shown in each visualization technique. For example, in the case of temporal ordering, participants were asked to select the correct causal sequence from three different layouts of the DAG in Figure 1. They did not know that the two figures were representations of the same Bayes net. They were also asked for main causes (origin of causality on each net) and general comments about the usefulness of temporal order applied to BBNs.

In the case of colour technique and size technique, participants were asked to select the three more probable nodes (marginal probability), and to give general comments about how colour or size were useful to them. The closeness technique was tested using animation when a specific node ("Tampering") was observed.

During the test, animation was used to show how the net was affected by new evidence. Animation technique was used to show changes in probability with a single visualization technique and combination of techniques. Participants answered specific questions such as "what was (were) the event(s) observed", and "how strong was the effect (big, small, or no effect) on the rest of the nodes", and general questions like "which kind of animation (continuos or node by node) was better to represent probability propagation".

6. RESULTS

- *Temporal order* was chosen as an appropriate way to show *cause-effect relationships* in BBNs. Indeed, using temporal order (Figure 1) the participants found the correct sequence easier than using non-temporal order. Nine of the ten participants selected two nodes (E and F) as main causes (origin in the causal sequence) using temporal order, while only two participants found them using non-temporal order. An interesting observation is that most of the participants (eight) found only one cause using non-temporal order. It can be attributed to the issue that people are not willing to check all possible nodes and directions using an unorganised net. Some of the reasons the participants selected temporal order are: hierarchical order is a natural way to show DAGs, it is simple and consistent, and it is more readable (top-down)
- Nine of the ten participants preferred *size* over *colour* to represent *marginal probability*. The main reason is that changes in size are easier to perceive than changes in colour intensity or hue.
- Given the issue that *big graphs* are very sensitive to changes in size and position of the nodes, *colour* is perhaps a better alternative for large BBNs. In addition, users' mental maps are not changed dramatically by using colour.
- *Closeness* technique proved to be an interesting and powerful way to show *probability propagation*, *changes in probability*, and to determine which cause seems to have a *strong influence* on a specific node.
- Although *colour* and *closeness* can be used to represent which parent has the strongest influence on a child, *participants* (eight) *preferred closeness* because changes in position are easier to perceive than changes in hue.
- *Combinations of techniques* appear to be clearer than a single technique to most of the participants (nine). In fact, Participants chose *size and colour* as a good alternative to represent *marginal probability* (size) and *strength of a relationship* (colour). These selections can be explained by the issue that by having two cues, it is possible to visualize different dimensions (complementary effect) or confirm what is shown using only one technique (reinforcement effect).
- *Animation* was very useful for representing *probability propagation*. Nine participants preferred node by node TOP-DOWN animation because it shows a sequence of events and probability propagation was more clear to them.

7. CONCLUSIONS

VisNet is a useful tool to represent, analyse and inspect Bayesian Belief Networks. Future work might include refining VisNet into an authoring tool for creating and tuning and maintaining BBNs. Special interfaces are being developed to allow teachers and students to inspect and modify BBNs in an educational application.

References

[1] Pearl, J. Graphical Models for Probabilistic and Causal Reasoning in *The Computer Science and Engineering Handbook*, Editor Tucker, A., CRCPress, Boca Raton, FL, 1997, 699-711.

[2] Russell, S., and Norving, P. *Artificial Intelligence: A modern Approach*. Prentice Hall, New Jersey, 1995.

[3] SMILE. Decision System Laboratory. University of Pittsburgh.

[4] Paul, B. Mesa 3.0. Copyright (C) 1995-1998

[5] Paul, B. and Bederson, B. Togl - a Tk OpenGL widget Version 1.5. Copyright (C) 1996-1998.

[6] Di Battista, G., Eades, P., Tamassia, R., Tollis, L., *Graph Drawing Algorithms for the Visualization of Graphs*. Prantice Hall, New Jersey, 1999, 265-301.

[7] Marcus, A. Principles of Effective Visual Communication for Graphical User Interface Design in *Reading in Human-Computer Interaction: Toward the year 2000*, Editors Baecker, R., Gruding, J.,Buxton, W. and Greenberg, S., Morgan Kaufmann Publishers. San Francisco, California, 1995, 425-441.

[8] Salomon, G., New Uses for Color in *The Art of Human-Computer Interface Design*, Editor Laurel, B., Addison-Wesley Publishing Company, Reading, MA, 1990, 269-278.

[9] Sherman, W.R., Craig, A.B., Baker, M.P., and Bushell, C. Scientific Visualization *in The Computer Science and Engineering Handbook*, Editor Tucker, A., CRCPress, Boca Raton, FL, 1997, 821-845.