In Proceedings of the Sixth International Conference on Cognitive Modeling, 416-417.
Mahwah, NJ: Lawrence Earlbaum.

Student Modeling for Collaborative Medical Problem-Based Leaning

Siriwan Suebnukarn (ssiriwan@cs.ait.ac.th)
Computer Science and Information Management Program, Asian Institute of Technology
PO Box 4, Pathumthani, 12120 Thailand

Background and Objectives

Today a great many medical schools have turned to a problem-based learning (PBL) approach to teaching. PBL instructional models vary but the general approach is student-centered, small group, collaborative problem-based learning activities. While PBL has many strengths, effective PBL requires the tutor to provide a high degree of personal attention to the students. In the current academic environment where resources are becoming increasingly scarce and costs must be reduced, providing such attention becomes increasingly difficult. This is exacerbated by the fact that medical school faculty, in particular, often have limited time to devote to teaching. As a consequence, medical students often do not get as much facilitated PBL training as they might need or want.

Our proposed work combines concepts from Intelligent Tutoring System with those from Computer Supported Collaborative Learning to develop an intelligent group-based medical PBL system. Our proposed work departs from previous efforts to incorporate user modeling into computer supported collaborative environments by focusing on modeling individual and group problem solving behavior. Since problem solving in group PBL is a collaborative process, modeling individuals and the group is necessary if we wish to develop tutoring algorithms that can do things like focus the group discussion, promote collaboration, and suggest peer helpers.

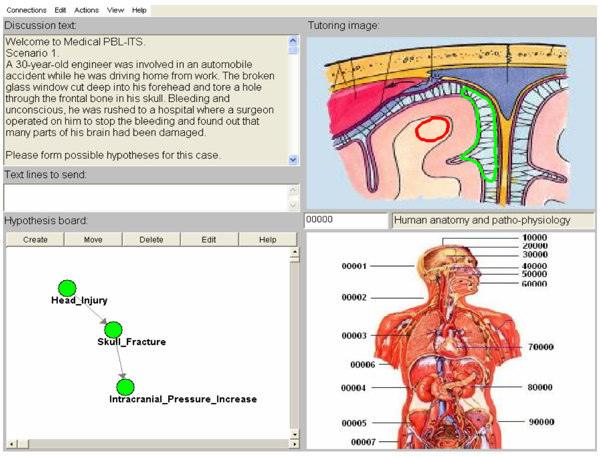

Figure 1: COMET Student interface

We have developed an initial prototype system called COMET (Suebnukarn & Haddawy, 2004) that is designed to provide an experience that emulates that of live human-tutored medical PBL sessions as much as possible while at the same time permitting the students to participate from disparate locations. The system uses Bayesian networks to model individual student knowledge and activity, as well as that of the group. It uses the models to generate tutorial hints to guide group problem solving activity (Fig. 1).

Student Clinical Reasoning Model

Generating appropriate tutorial actions requires a model of the students' understanding of the problem domain. This modeling task is necessarily wrought with uncertainty since we have only a limited number of observations from which to infer each student's level of understanding. Thus we have chosen to use Bayesian networks as our modeling technique. We model each student with an instance of our general Bayesian network student model. The group is reasoned about by combining information from the models of the individual students, as described below. Figure 2 shows a portion of the student model corresponding to the head injury scenario in Figure 1. The model contains two types of information; (1) the hypothesis structure based on the differential diagnosis of the scenario (the right group of nodes in Fig. 2); and (2) the application of medical concepts in terms of anatomy and patho-physiology (the left group of nodes in Fig. 2) to derive the hypotheses.

We represent the hypothesis structure following the model of Feltovich and Barrows (1984), which defines three categories of illness features: enabling conditions, faults, and consequences. Enabling conditions are illness features associated with the acquisition of illness, which can make an individual more susceptible than usual to illness in general or to particular illnesses. Faults are the major real malfunctions in illness. Consequences are the secondary consequences of faults within the organism, and generally comprise different types of signs and symptoms. Each hypothesis node has parent nodes, which have a direct casual impact on it. All hypothesis nodes have two states, indicating whether or not the student knows that the hypothesis is a valid hypothesis for the case.

The application of medical concepts is represented in terms of three kinds of nodes: goals, general medical knowledge, and apply actions. Every hypothesis node (except the root, which represents the scenario itself) has a unique *Apply* node as parent. The *Apply* node represents the application of a medical concept to a goal in order to derive the hypothesis. For example the *Apply3* node indicates that the student is able to use knowledge of the *Blood_Flow_Decrease* medical concept to infer that *Brain_ Damage* is a consequence of *Brain_Infection*. Each hypothesis node thus has a conditional probability table specifying the probability of the hypothesis being known conditioned on whether the parent hypotheses are known and whether the student is able to apply the appropriate piece of knowledge to determine the cause-effect relationship. The conditional probability tables for the *Apply* nodes are simple AND gates.

To construct the model which took about one person-month, we started from an initial Bayesian network built based on information extracted from medical textbooks and from interviews with a neurosurgeon and a physiologist from Thammasat University Medical School. The initial

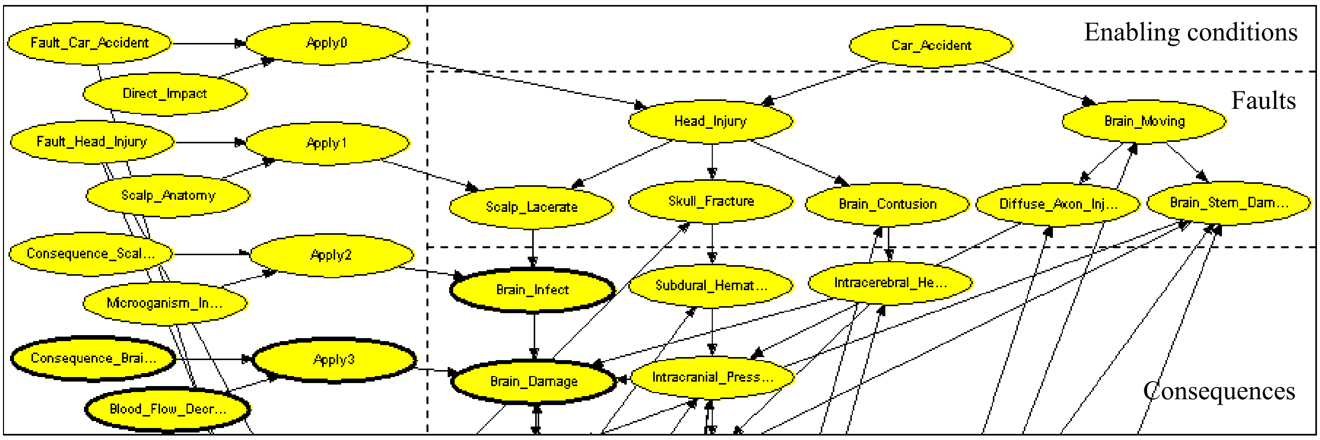

Figure 2: Part of the Bayesian network student model. The complete network contains 66 nodes.

model was refined by learning structure and parameters using data collected from a medical problem-based learning tutorial.

The student model is used to reason about the state of knowledge of each student, as well as about problem solving behavior of each student and of the group. The general student model is instantiated for each student by entering that student's medical background knowledge as evidence. We make the assumption that once a hypothesis in the domain model is mentioned by one student in the group, every student knows that hypothesis. So as hypotheses are mentioned, they are instantiated in each student model. Following commonly accepted practice in medical PBL, we assume that students should and generally do enumerate the possible hypotheses by focusing sequentially on the various causal paths in the domain, linking enabling conditions with faults and consequences. So for each student, we must determine what causal path he is reasoning along, which we do by identifying the path of highest probability. This is computed as the joint probability of the nodes along the path. It is also necessary to identify a causal path that can be used to focus group discussion, particularly when the discussion seems to be diverging in different directions. This is done as follows. We take the most likely path for each student and compute the sum of the probabilities of that path over all the students. We then take the path with the highest sum. The idea is that we want to identify the path that has much of the attention of much of the group and has at least one member whose attention is focused on that path.

Our automated tutor takes on the role of guiding the tutorial group to construct possible hypotheses for the case by the use of specific open-ended questions. From our study of the tutoring session transcripts, we identified eight hint strategies commonly used by experienced human tutors: 1) focus group discussion using general hint, 2) focus group discussion using specific hint, 3) promote open discussion, 4) deflect uneducated guessing, 5) avoid jumping critical steps, 6) address incomplete information, 7) refer to experts in the group, and 8) promote collaborative discussion. We developed algorithms to generate each of these types of hints, using as input the interaction log and the Bayesian network student models.

Evaluation

In order to evaluate the appropriateness and quality of the hints generated by our system, for each of the eight hinting strategies, we created three scenarios under which COMET would generate that hint. Ten experienced medical tutors were asked to analyze each situation and indicate whether they would provide a hint and what hint they would provide. Our results show that COMET's hints agree with the hints of the majority of the human tutors with a high degree of statistical agreement (McNemar test, p = 0.652, Kappa = 0.773).

Conclusions and Future Work

We have described a Bayesian network clinical reasoning model that integrates the hypothesis structure based on the differential diagnosis of the case and the application of the corresponding medical concepts in the problem solving process. Results from our initial evaluation show a high degree of agreement of tutoring dialogues between our system and human tutors.

We are currently working on evaluating the accuracy of a number of alternative Bayesian network student models that capture student's actions of creating hypotheses as well as the causal links between them. The models are being evaluated in terms of accuracy in predicting individual student actions and in terms of accuracy of identifying group reasoning paths. The ultimate test of the effectiveness of our work is how it impacts student learning. So we intend to compare the effectiveness of student learning with COMET versus student learning with human tutors.

Acknowledgments

We thank Hugin Expert for providing us the use of the Hugin Researcher software. Thanks to Thammasat University Medical School for their participation in the data collection and system evaluation.

References

Suebnukarn, S., & Haddawy, P. (2004). A collaborative intelligent tutoring system for medical problem-based learning. *Proceedings of the 9th International Conference on Intelligent User Interface* (pp. 14-21). Madeira, Portugal.

Feltovich, P. J., & Barrows, H. S. (1984) Issues of generality in medical problem solving. *Tutorials in problem-based learning: A new direction in teaching the health professions*. The Netherlands: Van Gorcum.