Clinical-Reasoning Skill Acquisition through Intelligent Group Tutoring

Siriwan Suebnukarn and Peter Haddawy
Asian Institute of Technology
Computer Science and Information Management Program
Pathumthani, Thailand, 12120
{Siriwan.Suebnukarn, haddawy}@ait.ac.th

Abstract

This paper describes COMET, a collaborative intelligent tutoring system for medical problem-based learning. COMET uses Bayesian networks to model individual student knowledge and activity, as well as that of the group. Generic domain-independent tutoring algorithms use the models to generate tutoring hints. We present an overview of the system and then the results of two evaluation studies. The validity of the modeling approach is evaluated in the areas of head injury, stroke and heart attack. Receiver operating characteristic (ROC) curve analysis indicates that, the models are accurate in predicting individual student actions. Comparison of learning outcomes shows that student clinical reasoning gains from our system are significantly higher than those obtained from human tutored sessions (Mann-Whitney, p = 0.011).

1 Introduction

The transformation from medical student to physician is a gradual one, requiring the assimilation of vast amount of knowledge as well as the development of clinical-reasoning skills. Clinical reasoning is the cognitive process by which the information contained in a clinical case is synthesized, integrated with the physician's knowledge and experience, and used to diagnose or manage the patient's problem [Newble *et al*., 1994]. Problem-based learning (PBL) has been introduced as an alternative to traditional didactic medical education to teach clinical-reasoning skills at the early stages of medical education. PBL is designed to challenge learners to build up their knowledge and develop effective clinical-reasoning skills around practical patient problems. PBL instructional models vary but the general approach is student-centered, small group, collaborative problem solving activities [Barrows, 1986]. The main arguments for using collaborative problem solving in medical PBL include the wider range of ideas generated and the higher quality of discussion that ensues. In addition, students obtain training in the skill of consultation and group clinical problem solving, which are important for the successful practice of clinical medicine. But effectively implementing PBL in the clinical curriculum is difficult due to the lack of standards for PBL tutoring [Das *et al*., 2002] and a lack of properly trained tutors. In addition, 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.

There has been increasing interest in application of intelligent technologies to medical training to provide rich environments for maximizing learning while minimizing risks to patients, until sufficient competency is established. The majority of the work in intelligent medical training system has focused on particular domains, such as Radiology [Sharples *et al*., 2000] or Pathology [Crowley and Medvedeva, 2003] for training students in feature perception and disease classification. Little or no work has addressed providing a general domain-independent framework for intelligent medical tutoring and no work has addressed intelligent medical tutoring in group settings.

We have developed a collaborative intelligent tutoring system for medical problem-based learning called COMET. COMET uses Bayesian networks to model individual student knowledge and activity, as well as that of the group. It uses generic tutoring algorithms applied to the models to generate tutorial hints to guide problem solving activity. In previous work [Suebnukarn and Haddawy, 2004] we presented a basic Bayes net student model and details of the tutoring algorithms. We also presented results of a study showing that the hints generated by COMET agree with those of a majority of human tutors. In this paper we present a new, more expressive Bayesian network student model, along with an ROC analysis evaluating of the model's accuracy. We also evaluate the overall effectiveness of COMET in imparting clinical reasoning skills to medical students by comparing clinical reasoning exam scores of COMET tutored students to those of human tutored students.

2 Medical Problem-Based Learning

Problem-based learning (PBL) can be described as “the learning that results from the process of working toward the understanding or resolution of a problem” [Barrows, 1986]. PBL is typically carried out in three phases: (1) *Problem analysis*: In group discussion the students evaluate the patient problem presented to them exactly as they would a real patient, attempting to determine the possible underlying anatomical, physiological, or biochemical dysfunctions and to enumerate all possible causal paths (hypotheses and their causal relations) that would explain the progression of the patient’s problems. (2) *Self-directed study*: In this phase, students work outside the tutorial session, using any relevant learning resources, e.g. literature, laboratories, specialists, to address any open issues identified in the first phase. (3) *Synthesis and application of newly acquired information*: The students analyze data and wrap up the problem collaboratively after they return from their self-study period.

One of the main issues in PBL is the role of the tutor. Like a good coach, a tutor needs enough command of whatever the learners are working on to recognize when and where they most need help [Das *et al*., 2002]. The ideal tutor should be an expert in both learning content and learning process, which is rare to find among human tutors. The tutor intervenes to as small an extent as possible, giving hints only when the group appears to be getting stuck or off track. In this way, the tutor avoids making the students dependent on him for their learning.

3 A Collaborative Medical Tutor

3.1 Conceptual Framework

COMET 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 collaboratively from disparate locations. COMET incorporates a multi-modal interface that integrates text and graphics so as to provide a rich communication channel between the students and the system, as well as among students in the group (Fig. 1). Students collaboratively create the problem solution on the hypothesis board, shown at the bottom of Fig. 1. Typically each student works from a separate computer. COMET can currently support PBL problem analysis in the domains of Head injury, Stroke and Heart attack. Note that these three domains are quite different since the knowledge used to reason about head injury is primarily anatomical, while that used to reason about stroke and heart attack is primarily physiological. Furthermore, the patho-physiology of the latter two diseases is more dynamic.

Generating appropriate tutorial actions in COMET requires a model of the students’ clinical reasoning for 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.

The system is implemented as a Java client/server combination, which can be used over the Internet or local area networks and supports any number of simultaneous PBL groups.

The system implementation is modular and the tutoring algorithms are generic so that adding a new scenario requires only adding the appropriate model representing how to solve a particular case (*domain clinical reasoning model*). The *student clinical reasoning model*, which is a probabilistic overlay of the domain clinical reasoning model, is then constructed during runtime by instantiating the nodes that represent the knowledge and activity of an individual student. The architecture of COMET differs from that of most ITS’s in that the domain model and student model are embodied in one representation. The domain model is contained in the part of the structure of the network that represents the hypotheses and the cause-effect relations among them. The student model is contained in the part of the network that represents how the hypotheses are derived and in the network’s probabilities. The probabilities do not represent the likelihood of occurrence of the hypotheses but rather the likelihood that a student will be able to create the hypotheses.

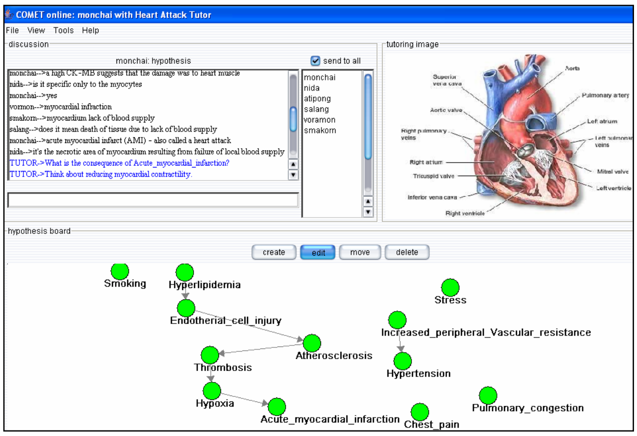

Figure 1. COMET student interface.

3.2 Domain and Student Clinical Reasoning Model

We built the domain clinical reasoning model based on the process of hypothesis generation in problem-based learning. Consider, for example, the heart attack scenario taken from a PBL session at Thammasat University Medical School.

“Mr. C, a 56-year-old who was diagnosed as having essential hypertension four years ago, is complaining of chest pain which feels like indigestion. You have noticed that he is mildly obese, pale, clammy and sweating profusely…”

Here students must enumerate possible hypotheses to explain why the patient is experiencing chest pain. Figure 2 is a photograph of the white board of the group PBL session, showing a directed acyclic graph representing cause-effect relationships among hypotheses. This graph represents the problem solution developed by the students. Since we would like to reason about the state of knowledge of each student concerning the solution, this graph is our starting point for the student model. The hypothesis graph can be conveniently represented as a Bayesian network since

Bayesian networks are also directed acyclic graphs. In addition, Bayesian networks can represent our uncertainty about the state of knowledge of the students.

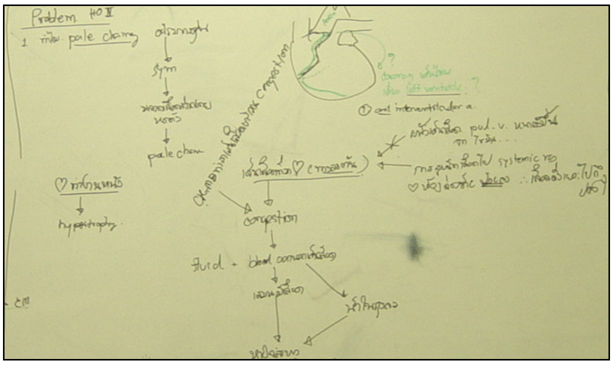

Figure 2. A photograph of the white board after a PBL session at Thammasat University Medical School. The graph shows hypotheses with arrows indicating cause-effect relations among them. (Note: Some hypotheses are written in Thai.)

The BN structure contains two types of information: (1) the hypotheses and the causal links of the problem solution (Fig. 3, right half) and (2) how students derive the hypotheses (Fig. 3, left half). 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 (e.g., compromised host factors, unusual travel, or hereditary factors). Faults are the major real malfunctions in illness (e.g., direct trauma, invasion of tissue by pathogenic organisms, or inadequate blood supply). Consequences are the secondary consequences of faults within the organism, and generally comprise different types of signs and symptoms, e.g., chest pain, breathlessness, or tachycardia. In Figure 3 (right half), we have five possible faults associated with the single consequence chest pain: *Myocardial infarction*, *Angina*, *Musculoskeletal injury*, *Gastrointestinal disorder*, and *Stress*. *Atherosclerosis* is the enabling condition of *Myocardial infarction* and *Angina*. The remaining hypothesis nodes are consequences of *Myocardial infarction*. Each hypothesis node has parent nodes, which have a direct casual impact on it. For example, *Right heart failure* has parents *Pulmonary congestion* and *Myocardial infarction*. All hypothesis nodes have two states, indicating whether or not the student knows that the hypothesis is a valid hypothesis for the case.

In the PBL sessions, the students create the hypotheses as well as the causal links between them (Fig. 2). We would like to be able to reason about the probability that students know the correct causal links. But in a Bayes net, random variables are represented with nodes. So we use *link nodes* to represent the causal links between hypotheses. For every hypothesis that is a direct cause of another hypothesis (e.g. *Atherosclerosis* and *Myocardial infarction)*, we have a node (e.g *Link_14*) representing the causal link between them. The two hypothesis nodes (*Atherosclerosis*, *Myocardial infarction*) are the parents of the link node. The intuition is that the link cannot be created unless both hypotheses are created first. Each link node has two states, indicating whether or not the student creates a causal link between two hypotheses.

The derivation of hypotheses (Fig. 3, left half) is represented in terms of three kinds of nodes: goals, general medical knowledge, and apply actions. Every hypothesis node has a unique *Apply* node as one of its parents. The *Apply* node represents the application of a medical concept to a goal in order to derive the hypothesis. For example the *Apply_13* node indicates that the student is able to use knowledge of the *Vessel Lumina occlusion* medical concept to infer that *Myocardial infarction* is a consequence of *Atherosclerosis*. 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.

Our BN student model is similar to the student model used by Conati, et al [2002]. Their model includes five types of nodes: Context-Rule, Rule-Application, Fact, Goal, and Strategy. The correspondence between their node types and ours is: Context-Rule = Concept, Rule-Application = Apply, Fact = Hypothesis, and Goal = Goal. Strategy nodes, which represent different correct solutions to a problem, are implicitly encoded in our model by the fact that students can enumerate the causal hypothesis structure in any order. Our model contains causal links among hypotheses, which are not present in their model. The reason for this is that in our medical domains a problem solution is represented by the hypotheses and causal links among them, while in their physics domains a problem solution is represented by a sequence of rule applications and the derived facts.

For each problem scenario, we consulted medical textbooks and expert PBL tutors to obtain the hypotheses, the causal relations among them, the goals, and the medical concepts used to derive the hypotheses. The conditional probability tables for each resulting network were obtained by learning from data obtained from transcripts of PBL sessions. The data for this study consisted of tape recordings and photographs of tutorial sessions for the head injury, stroke and heart attack scenarios at Thammasat University Medical School. A total of 15 groups of third year medical students were involved in this study. Each group, consisting of eight students with different backgrounds, was presented with the head injury, stroke and heart attack cases and asked to construct possible hypotheses for the case, under the guidance of a tutor. After the sessions the tape and the results on the whiteboard were analyzed to determine whether or not each goal, concept, hypothesis, and link was mentioned. We used the EM learning algorithm provided by the HUGIN Researcher software to learn the conditional probabilities of each node [Lauritzen 1995].

3.3 Individual and Collaborative Student Clinical Reasoning Modeling

The domain clinical reasoning model is instantiated for each student by entering that student's medical background

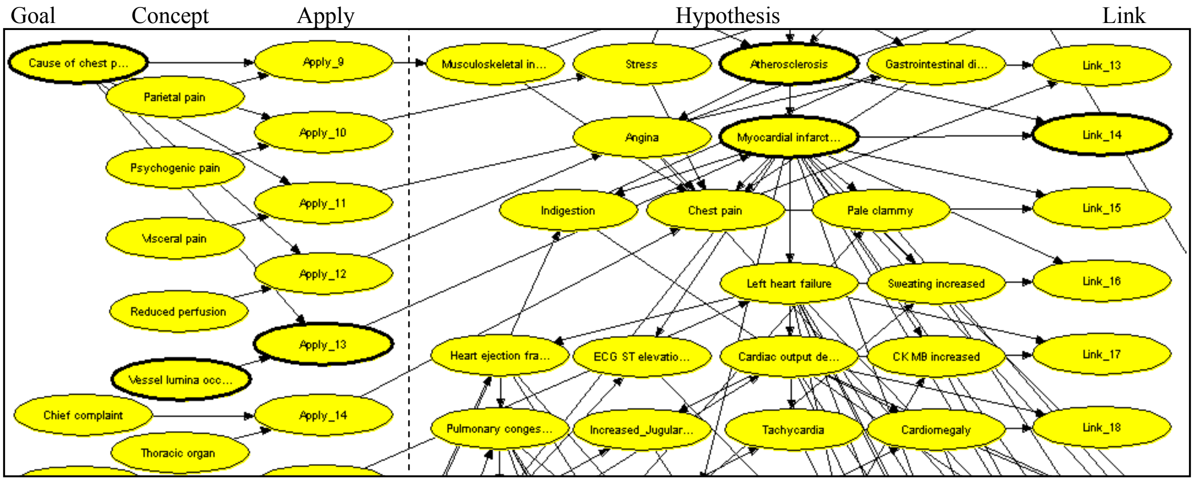

Figure 3. Part of the Bayesian Network clinical reasoning model of the heart attack scenario. The complete network contains 194 nodes. The model contains five types of nodes: goal, concept, apply, hypothesis, and link.

knowledge as evidence. For example, if a student has a background in Thoracic anatomy, we would instantiate the *Thoracic organ* node. Since all students have *basic* knowledge in Anatomy, Physiology and Pathology before they encounter the PBL tutorial sessions, we make the assumption that once a hypothesis in the domain model is created by one student in the group, every student knows that hypothesis. So as hypotheses are created, 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 in that student's model. This is computed as the joint probability of the nodes along the path. The most likely current reasoning path for each student is path that gives the maximum joint probability. Since the students work in a group, 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. We would like to identify a 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. We identify a set of candidate paths by taking the most likely path for each student. This guarantees that each candidate path has at least one student currently focused on it. We then compute the sum of the probabilities of each candidate path over all students and select the path with the highest sum. This gives us the candidate path with the highest average attention over all students.

3.4 Generating Tutorial Hints

Our automated tutor takes on the role of guiding the tutorial group to construct possible hypotheses for the case by the use of open-ended questions. From our study of the tutoring session transcripts, we identified and implemented seven hint strategies commonly used by experienced human tutors: (1) *Focus group discussion:* Members of the group may suggest various valid hypotheses without focusing on any given causal path. When such lack of focus becomes apparent, COMET intervenes by directing the students to focus on one of the hypotheses in the group path. (2) *Promote open discussion*: If a student proposes a hypothesis that is not on the current group reasoning path, COMET provides positive feedback by encouraging the student to relate the hypothesis to the current focus of discussion. (3) *Deflect uneducated guessing:* If a student creates an incorrect causal link, COMET points this out and encourages the student to correct the error. (4) *Avoid jumping critical steps*: If a student creates a link that jumps directly from one hypothesis to a down-stream consequence, leaving out intermediate hypotheses, COMET asks the student for the more direct consequences. (5) *Address incomplete information:* Once the students have completed elaborating all hypotheses on the group path, COMET identifies another path for them to work on. (6) *Refer to experts in the group*: If after COMET provides a general and then a more specific hint, the students still do not respond correctly, COMET determines the student most likely to know the answer and refers directly to him. (7) *Promote collaborative discussion*: If one student dominates the discussion, COMET asks for input from the other students. If a student does not contribute after a certain number of hypotheses have been mentioned, COMET solicits input from that student.

All strategies except strategies 6 and 7 have general and specific versions. COMET first gives a general hint using the parent goal node of the hypothesis that it has determined the students should focus on, and if there is no student response or an incorrect response, the more specific parent medical concept node is used. If the students can still not come up with the hypothesis of interest, COMET refers directly to the student in the group most likely to know the answer. If this doesn't work, COMET identifies this as a learning objective for study outside the session.

We developed algorithms to generate each of these types of hints, using as input the interaction log and the Bayesian network student models. All strategies except strategy 7 use both the structure and the probabilities of the Bayes net models. Strategy 7 uses only a count of the number of inputs from each student. Strategies 1, 2, 5 make use of the group reasoning path discussed in the previous section. The following transcript shows the interaction with the system after the students read the heart attack problem scenario, described in Section 3.2. The system selects hint strategies and content based on the student input on the hypothesis board.

Students: Gastrointestinal disorder, Smoking, Hypertension, Angina, Myocardial infarction, Chest pain *(Students in the group gradually create six hypotheses on the board, while analyzing the problem.)*

Tutor: What is the consequence of Myocardial infarction? *(Strategy 1: COMET focuses group discussion by identifying which causal path the group should focus on, finding the highest probability non-mentioned node along the path (Left heart failure), and providing a hint using its parent goal node (Consequence of Myocardial infarction).)*

Student: Cardiac output decreased *(Cardiac output decreased is a node along the group reasoning path but not the node that COMET wants the group to focus on.)*

Tutor: Think about decrease in myocardial contractility. *(Strategy 1: COMET gives the next hint using the medical concept parent node of Left Heart Failure.)*

Student: Right heart failure *(Right heart failure is a node along the group reasoning path but not the node that COMET wants the group to focus on.)*

Tutor: It seems we have a problem. Nida, can you help the group? *(Strategy 7: Nida is called on since she has the highest probability of knowing the Left heart failure node among the students.)*

4 Evaluation

4.1 Evaluation of the Student Model

In order to determine the accuracy of the model, we compared the probabilities of hypotheses and causal links from the student model with actual student actions considered as a "gold standard". We recruited 15 second-year medical students from Thammasat University Medical School. That is, they had not yet had PBL experience in Head injury, Stroke, or Heart attack. Stratified random sampling was applied to divide the students into 3 groups based on their background knowledge. Students were asked to answer pre-test questions to determine their background knowledge. This information was used to instantiate the general student model for each individual student. Students participated individually in the problem solving sessions on head injury, stroke, and heart attack with COMET. Each student was asked to enumerate hypotheses and the causal links without any help from the COMET tutor. The student actions of creating hypotheses and their links served as a gold standard for comparison with the predicted probabilities from the Bayesian network student model.

Results

To determine whether our student models are accurate in predicting student actions, we evaluated them by means of receiver operating characteristic (ROC) curve analysis. The area under the curve (AUC) represents an overall measurement of performance of the student model, with 1.0 a perfect test and 0.5 representing a model with no discriminating capacity.

Table 1 shows the ROC curve analysis of the student models for the Head injury, Stroke, and Heart attack scenarios. There were more false positive cases in the Stroke scenario than in the others, since the student model for the Stroke scenario was built from students who had already studied Cerebrovascular knowledge from the Head injury scenario, while in this study, we recruited students who had not yet studied any of the three scenarios. Averaging results over all scenarios shows high accuracy in predicting both hypotheses and causal links.

Table 1. ROC analysis showing AUC for three scenarios

Scenarios	Prediction	
	Hypotheses	Causal links
Head injury	0.909	0.898
Stroke	0.765	0.838
Heart attack	0.868	0.905
All	**0.832**	**0.900**

4.2 Evaluation of Student Clinical Reasoning Gains

To evaluate the overall impact of the system on student learning, we designed a study to test the hypothesis that a COMET tutorial will result in similar student clinical reasoning gains to those obtained from a session with an experienced human PBL tutor.

We compared three groups of students tutored by COMET with three other groups of students tutored by experienced human tutors. The study had a pre/post test control group design. All students were assessed on their clinical reasoning before and after the PBL tutorial session on heart attack and stroke to determine the reasoning gains for each individual student.

We used the Clinical Reasoning Problem (CRP) approach for clinical reasoning assessment [Groves *et al*., 2002]. Each CRP consisted of a clinical scenario that was vetted for clinical accuracy and realism by a specialist physician. Four cases in the pre-test set measured each student's initial ability to solve the problems. Four other post-test cases measured their ability to generalize the clinical reasoning acquired from tutorial session to the new related cases. Participants were asked to nominate the two diagnoses they considered most likely, to list the features of the case that they regarded as important in formulating their diagnoses, and to indicate whether these features were positively or negatively predictive. To establish reference scores, ten volunteer general practitioners (GPs) were asked to complete both sets of CRPs.

Results

There were no statistically significant differences between pre- and post-test scores obtained from the GPs, indicating that the pre- and post-tests were of approximately equal difficulty (Table 2). The GPs' scores varied from 88.20 to 91.50 indicating that the questions were not trivial. Reliability, the measure of the reproducibility of a test, was measured using Cronbach's alpha. Cronbach's alpha for pre-and post-test student scores were 0.901 and 0.921 respectively. A reliability coefficient of 0.80 or higher is commonly considered as acceptable.

Table 2. Mean score for all CRPs (CRPs 1.1, 1.2, 2.1, 2.2 are chest pain cases. CRPs 1.3, 1.4, 2.3, 2.4 are stroke cases.)

	CRPs	GP's score (SD)	Student's score (SD)	
			COMET	Human tutor
Pre-test	1.1	88.70 (2.45)	34.67 (4.51)	34.00 (2.70)
	1.2	91.50 (2.46)	34.00 (3.27)	34.78 (2.73)
	1.3	88.20 (1.69)	37.72 (2.21)	38.61 (3.39)
	1.4	89.80 (3.49)	39.17 (2.18)	38.22 (3.54)
Post-test	2.1	89.50 (3.37)	62.28 (2.11)	58.11 (1.94)
	2.2	87.70 (4.42)	63.94 (1.95)	58.67 (2.40)
	2.3	90.60 (2.63)	64.06 (1.94)	65.00 (2.74)
	2.4	89.50 (3.27)	65.56 (1.98)	64.05 (2.39)

Table 3 shows that there were no statistically significant differences between pre-test mean scores of the COMET and human tutored groups. The post-test mean scores were significantly higher than the pre-test mean scores in both COMET and human tutored groups (Wilcoxon, $p = 0.000$), indicating that significant learning occurred. But the average post-test score for the COMET groups (64.96) was significant higher than that obtained for the human tutored groups (60.46) (Mann-Whitney, $p = 0.011$), indicating that students were learning more in the COMET sessions than in the human tutored sessions.

Table 3. Mean CRP score for each cohort

Cohort	Mean score (SD)	
	Pre-test	Post-test
COMET (1)	36.38 (3.45)	66.12 (3.38)
COMET (2)	37.00 (4.11)	64.33 (2.78)
COMET (3)	35.54 (4.24)	65.42 (3.10)
COMET (all)	**36.31 (3.90)**	**64.96 (3.08)**
Human tutor (1)	36.42 (2.95)	60.96 (2.49)
Human tutor (2)	37.42 (2.37)	62.63 (1.99)
Human tutor (3)	35.38 (3.42)	58.79 (2.68)
Human tutor (all)	**36.40 (3.68)**	**60.46 (2.40)**

5 Discussion

The results showing that clinical reasoning gains for COMET tutored students are higher than those for human tutored students were unexpected. This is particularly true in light of our earlier study showing that on average 74% of human tutors used the same hint strategy and content as COMET [Suebnukarn and Haddawy, 2004]. We believe the explanation lies primarily in the 26% disagreement. Human tutors often give up after providing a general hint, jumping right to identifying the hypothesis as a learning objective. In contrast, COMET is more relentless in pushing the students, always following the sequence of general hint, specific hint, referring to expert, and finally identifying as a learning objective. It is generally agreed that students should generate as many hypotheses as possible in a PBL session, leaving only the truly difficult issues as learning objectives.

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, and to Dr. Kesorn Jongjarern for her helpful suggestions in designing the student clinical reasoning model.

References

[Barrows, 1986] H. Barrows. A taxonomy of problem-based learning methods. *Medical Education*, 20: 481-486, 1986.

[Crowley and Medvedeva, 2003] R. Crowley. and O. Medvedeva. A General Architecture for Intelligent Tutoring of Diagnostic Classification Problem Solving. In *Proceedings of AMIA Symposium*, pages 185-189, Washington DC, November 2003.

[Conati *et al*., 2002] C. Conati, A. Gertner, and K. VanLehn.: Using Bayesian Networks to Manage Uncertainty in Student Modeling, *Journal of User Modeling and User-Adapted Interaction*. 12:371-417, 2002.

[Das *et al*., 2002] M. Das, D. Mpofu, M. Hasan, and T. Stewart. Student perceptions of tutor skills in problem-based learning tutorials. *Medical Education*, 36:272-278, 2002.

[Feltovich and Barrows, 1984] P. Feltovich and H. Barrows. Issues of generality in medical problem solving. *Tutorials in problem-based learning: A new direction in teaching the health professions*. Van Gorcum, The Netherlands, 1984.

[Groves *et al*., 2002] M. Groves, I. Schott, and H. Alexander. Assessing Clinical Reasoning: A Method to Monitor its Development in PBL Curriculum. *Medical Teacher*. 5: 507-515, 2002.

[Lauritzen 1995] S. Lauritzen. The EM-algorithm for graphical association models with missing data. *Computational Statistics and Data Analysis*, 1: 191-201, 1995.

[Newble *et al*., 2000] D. Newble, G. Norman, and C. Vleuten. Assessing clinical reasoning, in: J. H. M. Jones (Ed.) *Clinical Reasoning in the Health Professions*, 2nd ed, pp. 156–168, Oxford, Butterworth-Heinemann, 2000.

[Sharples *et al*., 2000] M. Sharples, N. Jeffery, B. du Boulay, B. Teather, D. Teather, and G. du Boulay. Structured computer-based training in the interpretation of neuroradiological images. *International Journal of Medical Informatics*, 60: 263-280, 2000.

[Suebnukarn and Haddawy, 2004] S. Suebnukarn and P. Haddawy. A collaborative intelligent tutoring system for medical problem-based learning. In *Proceedings of the 9th International Conference on Intelligent User Interfaces*, pages 14-21, Madeira, Portugal, January 2004.