Intelligent Tutoring System for Clinical Reasoning Skill Acquisition in Dental Students

Siriwan Suebnukarn, D.D.S., Ph.D.

Abstract: Learning clinical reasoning is an important core activity of the modern dental curriculum. This article describes an intelligent tutoring system (ITS) for clinical reasoning skill acquisition. The system is designed to provide an experience that emulates that of live human-tutored problem-based learning (PBL) sessions as much as possible, while at the same time permitting the students to participate collaboratively from disparate locations. The system uses Bayesian networks to model individual student knowledge and activity, as well as that of the group. Tutoring algorithms use the models to generate tutoring hints. The system incorporates a multimodal 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. Comparison of learning outcomes shows that student clinical reasoning gains from the ITS are similar to those obtained from human-tutored sessions.

Keywords: intelligent tutoring systems, clinical reasoning skill, problem-based learning, Bayesian networks, empirical evaluation

Submitted for publication 4/19/09; accepted 6/8/09

Educators agree that clinical reasoning is a central component of clinician competence, and objectives related to mastery of clinical reasoning skills appear in the documentation of most medical and dental schools, licensing bodies, and specialty societies. Learning clinical reasoning is an important core activity of the modern dental curriculum. Clinical reasoning is the cognitive process by which the information contained in a clinical case is synthesized, integrated with the clinician’s knowledge and experience, and used to diagnose or manage the patient’s problem. Problem-based learning (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 involves student-centered, small-group, collaborative problem-solving activities. PBL is typically carried out in three phases: problem analysis, self-directed study, and synthesis and application of newly acquired information.

While PBL has many strengths, effective PBL requires the tutor to provide a high degree of personal attention to the students in order to recognize when and where they most need help and to guide them to find their own answers. Whereas an instructor can deliver a lecture to a large lecture hall of students, a typical PBL tutorial session includes only six to eight 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.

During the last two decades, the impact of information and computer technologies in general and in education in particular has been considerable. Computers play a role in many areas in dental education. In the PBL area, we have developed an intelligent tutoring system (ITS) called COMET (Collaborative MEDical Tutor). COMET uses Bayesian networks (BNs) 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 an early laboratory study with COMET, an evaluation found a high degree of agreement between the hints generated by COMET...
and those of experienced human tutors. The results of the evaluations of the accuracy of the alternative student models in determining individual student actions provide encouraging support for the framework of COMET’s clinical reasoning model.\textsuperscript{9,10} This article presents an overview of the system and reports on the results of a formal evaluation of COMET’s effectiveness in imparting problem-solving skills to dental students.

**Methods**

The current version of COMET supports the initial problem analysis phase of PBL, in which there is intensive interaction between students and the system as well as among students in the group. COMET contains four primary components similar to any typical ITS\textsuperscript{11}: domain model, student model, pedagogic module, and user interface. Figure 1 shows the basic components of the system. Our prototype incorporates preclinical domain knowledge about head injury diagnosis, stroke, and heart attack. The system implementation is modular and the pedagogical model is generic, so that adding a new scenario requires only adding the appropriate model representing how to solve a particular case, e.g., pneumonia, peptic ulcer, dental caries, endodontic lesions (domain model).

COMET emulates the PBL environment by incorporating a multimodal 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 (Figure 2). The hypothesis board (Figure 2, lower pane) provides the central shared group workspace. Students can create hypothesis nodes and create causal links between nodes. Hypothesis node labels are created by retrieving them from the medical concept repository. The system has access to all information entered on the board. To support collaborative interaction, the discussion pane (Figure 2, upper left pane) is the place for displaying tutoring hints and student chat dialogues. Students may communicate with others in the group by typing into the text box below the discussion pane. COMET has no language-processing capabilities, so the text in the chat pane is not taken as input to the system. The interface includes an image pane (Figure 2, upper right pane) in which COMET displays images that are relevant to the current focus of the group discussion. All students see the same image and see anything that other students sketch or point to on the image. The system includes a medical concept repository to help students better understand

![Figure 1. COMET system overview](image-url)
the relationships between domain concepts, as well as to facilitate system input. All valid and invalid hypotheses that are relevant to all problem scenarios are stored in the repository, relieving the system of the need to process typed text input.

**Domain Model**

Generating appropriate tutorial actions requires a model of the students' understanding of the problem domain and of the problem solution. However, as in human-tutored PBL sessions, COMET must provide an unrestricted interaction style, which gives the students the freedom to solve the patient case without requiring them to explain the rationale behind their actions. This complicates the modeling task since we have only a limited number of observations from which to infer each student’s level of understanding. The resulting need to deal with uncertainty has lead us to use the Bayesian network (BN) as our modeling technique. To construct the model, we used information from various sources. From research papers and textbooks on problem-solving and PBL, we obtained the generic structure of the network. For each problem scenario, we consulted textbooks and expert PBL tutors. The BN structure contains two types of information: 1) the hypotheses and the causal links of the problem solution (Figure 3, right half) and 2) how students derive the hypotheses (Figure 3, left half). We represent the hypothesis structure following the illness script of Feltovich and Barrows, which defines three categories of illness features: enabling conditions, faults, and consequences.

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 causal 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.

---

**Figure 2. COMET student interface**
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 (Figure 3, left half) is represented in terms of three kinds of nodes: goals, general medical concepts, 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. Part of the modeling process is to come up with a discrete medical concept that underlies a unique 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. The student is able to apply only if the corresponding goal and concept are known.

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 Dental School. A total of fifteen groups of third-year dental 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 were mentioned. We used the EM learning algorithm provided by the HUGIN Researcher software to learn the conditional probabilities of each node.

**Figure 3. Part of the Bayesian network clinical reasoning model of the heart attack scenario**

*Note: The complete network contains 194 nodes. The model contains five types of nodes: goal, concept, apply, hypothesis, and link.*

**Student Model**

The domain model is instantiated for each student by entering that student’s background 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 as-
sumption 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. The differences in background knowledge, represented in the individual student models, result in differences in the likelihoods of yet to be created hypotheses.

Since the students work in a group, it is necessary to identify a causal path linking enabling conditions, faults, and consequences that can be used to focus group discussion, particularly when the discussion seems to be diverging in different directions. Intuitively, 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. The algorithm used by COMET identifies what we call the group path by determining the probabilities of the various candidate paths, given the hypotheses created by the students thus far.9

Pedagogic Model

Our automated tutor needs to take on the role of guiding the tutorial group to construct possible hypotheses for the case by the use of specific open-ended questions. The tutor should give hints when the group appears to be getting stuck, off track, collaborating poorly, or producing erroneous hypotheses. To do this, the tutor requires knowledge of both the problem domain and the problem-solving process. 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 downstream consequence, leaving out intermediate hypotheses, COMET asks the student for the more direct consequences.

5. **Address incomplete information:** Once 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 or her.

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.

We developed tutoring algorithms to generate each of these types of hints, using as input the interaction log and both the structure and the probabilities of the BN models. All strategies except numbers 3 and 7 use both the structure and the probabilities of the BN models. Strategy 3 uses only the structure of the model, while strategy 7 uses only a count of the number of inputs from each student. Strategies 1, 2, and 5 make use of the group path discussed in Section 3.3. Strategies 1–5 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 is given, the more specific parent medical concept node is used. If the students still cannot 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.8

Evaluation

In this study, COMET’s effectiveness in imparting problem-solving skills to students is demonstrated by comparing student clinical reasoning gains obtained using COMET versus those obtained from experienced human-tutored PBL sessions. We expected to see no differences in clinical reasoning gains obtained from both groups.

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 problem-solving gains to those obtained from a session with an experienced human PBL tutor. We recruited thirty-six second-year students

---

9. Since the students work in a group, it is necessary to identify a causal path linking enabling conditions, faults, and consequences that can be used to focus group discussion, particularly when the discussion seems to be diverging in different directions. Intuitively, 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. The algorithm used by COMET identifies what we call the group path by determining the probabilities of the various candidate paths, given the hypotheses created by the students thus far.
from Thammasat University Dental School. At that stage, they had not yet had PBL experience in stroke and heart attack. Stratified random sampling was applied to divide the students into six groups based on their background knowledge. We compared three groups of students tutored by COMET with three other groups of students tutored by experienced human tutors (Figure 4). Initial training in the use of COMET required about fifteen minutes. Three PBL tutors participated in the study. Each tutor had at least five years’ experience in conducting the stroke and heart attack course at Thammasat. The study had a pre-/post-test control group design. After the pre-test session, students participated in a two-hour problem-solving session. Each group was asked to enumerate a network of hypotheses and causal links collaboratively explaining the stroke and heart attack scenarios. In the human-tutored groups, the network was created on a whiteboard in the classroom. All of the valid hypotheses and the links created in each group were counted. The validity of hypotheses and links were verified by the expert in the area of stroke and heart attack. All students were assessed on their clinical reasoning before and after the PBL tutorial session on heart attack and stroke to determine the clinical reasoning gains for each individual student. The student clinical reasoning gains were then compared between the two groups.

We used the clinical reasoning problem (CRP) approach15 for clinical reasoning skill assessment. Each CRP consisted of a clinical scenario that was vetted for clinical accuracy and realism by a specialist physician. Two cases in the area of heart attack and two in the area of stroke in the pre-test set measured each student’s initial ability to solve the problems. Two other post-test cases in each area measured their ability to generalize the clinical reasoning acquired from the PBL tutorial session to the new related cases. Participants were asked to name the two diagnoses they considered most likely, to list the features of the case that they regarded as important in formulating their diagnoses, to indicate whether these features were positively or negatively predictive, and to give a weighting to each. To establish reference scores, ten volunteer general practitioners (GPs) were asked to complete both sets of CRPs. GPs were chosen because they have experience with a broad range of undifferentiated clinical presentations that encompass all areas of clinical practice. In this respect, they provide the most appropriate standard against which to compare students who graduate with a sound background in general medicine but without any specialist knowledge.

The mean values were compared between pre- and post-test scores obtained from the GPs with dependent t-tests. Reliability of a test was measured using Cronbach’s alpha. To examine the differences between pre- and post-test scores on the same group of students, we used the Wilcoxon test for nonparametric data and matched pairs. The Mann-Whitney test for unmatched data was used to detect any differences between COMET- and human-tutored groups. All statistical analyses were performed using SPSS 12.0 (SPSS, Chicago, IL, USA).

Figure 4. Students tutored by COMET (a) and students tutored by experienced human tutors (b)
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 1). The GPs’ scores varied from 88.20 to 91.50, indicating that the questions were not trivial. Cronbach’s alphas 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 shows that the pre-test mean scores of the COMET- and human-tutored students were almost identical, and a Mann-Whitney test confirmed there was no significant difference. The post-test mean scores were significantly higher than the pre-test mean scores in both COMET- and human-tutored students (Wilcoxon, $p=0.000$), indicating that significant learning occurred. The average post-test score for the COMET students (61.45) was not significantly different from that obtained for the human-tutored students (59.46) (Mann-Whitney, $p=0.058$), indicating that students were learning similarly in the COMET sessions and the human-tutored sessions. However, comparing the number of hypotheses and links created by each group shows that, on average, there were a greater number of hypotheses and links created by the COMET groups (stroke: hypotheses=45.33, links=48.33; heart attack: hypotheses=40.67, links=44.67) compared to the human-tutored groups (stroke: hypotheses=35, links=40; heart attack: hypotheses=33.33, links=38).

Discussion

All can agree that clinical reasoning, or one of its many synonyms (problem-solving, decision making, judgment), should be taught and tested in dental education. Problem-based learning promotes clinical reasoning skill—an iterative process of hypothesis generation, testing, refinement, and synthesis that is used by both clinicians and students when confronted with an unfamiliar or atypical case. Following identification of critical or key features contained in the case, the clinician generates several possible diagnostic hypotheses and then selects the most probable amongst them on the basis of discriminatory or predictive features. As clinical experience increases, however, pattern recognition takes over as the most frequently used technique of clinical reasoning.$^{16,17}$ Pattern recognition is extremely efficient and is associated with the almost instantaneous identification of certain highly significant features, alone or in combination, that define the case and lead rapidly to diagnosis. A consequence of introducing formal tuition in clinical reasoning is a requirement for its specific evaluation. Methods such as patient management problems,$^{18}$ key feature problems,$^{19}$ modified essay questions,$^{20}$ and clinical reasoning problems (CRPs)$^{14}$ have been introduced. CRPs$^{15}$ are designed to evaluate clinical reasoning at different levels of expertise and to monitor the evolution of clinical reasoning skill in students as they progress through their medical training. The strength of the methodol-

### Table 1. Mean score for all clinical reasoning problems (CRPs)

<table>
<thead>
<tr>
<th>CRPs</th>
<th>Students’ Score (SD)</th>
<th>GPs’ Score (SD)</th>
<th>COMET</th>
<th>Human tutor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$n=10$</td>
<td>$n=18$</td>
<td>$n=18$</td>
</tr>
<tr>
<td>Pre-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1</td>
<td>88.70 (0.78)</td>
<td>33.56 (0.48)</td>
<td>34.56 (0.80)</td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>91.50 (0.78)</td>
<td>34.00 (0.77)</td>
<td>34.78 (0.64)</td>
<td></td>
</tr>
<tr>
<td>1.3</td>
<td>88.20 (0.53)</td>
<td>38.72 (0.52)</td>
<td>38.61 (0.80)</td>
<td></td>
</tr>
<tr>
<td>1.4</td>
<td>89.80 (1.10)</td>
<td>39.17 (0.51)</td>
<td>38.22 (0.83)</td>
<td></td>
</tr>
<tr>
<td>Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1</td>
<td>89.50 (1.07)</td>
<td>62.44 (0.59)</td>
<td>58.11 (0.46)</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>87.70 (1.40)</td>
<td>62.94 (0.46)</td>
<td>58.67 (0.56)</td>
<td></td>
</tr>
<tr>
<td>2.3</td>
<td>90.60 (0.83)</td>
<td>67.56 (0.47)</td>
<td>60.00 (0.65)</td>
<td></td>
</tr>
<tr>
<td>2.4</td>
<td>89.50 (1.04)</td>
<td>67.56 (0.46)</td>
<td>61.05 (0.56)</td>
<td></td>
</tr>
</tbody>
</table>

*Note: CRPs 1.1, 1.2, 2.1, and 2.2 are chest pain cases. CRPs 1.3, 1.4, 2.3, and 2.4 are stroke cases.*
ogy using an intelligent tutoring system for clinical reasoning skill acquisition is that the recording is not limited to the outcome data (e.g., number of hypotheses and links created by students), as demonstrated in this study. Rather, it can automatically record associated process data on how students solve the problems, e.g., reasoning path, search strategy (depth or breadth), or application of concepts to goals, which are not available in the conventional training environments. Those variables are important for the development of objective scoring criteria that lead to the future development of an evaluative tool to assess clinical reasoning.

Most of the intelligent tutoring systems\(^{10,21}\) produce hints when the student requests help and bug messages when the student errs. Every entry the student makes in the problem-solving interface receives immediate feedback on whether it is correct or incorrect, e.g., green and red in ANDES, a well-known intelligent tutoring system in physics.\(^{21}\) Following PBL tutorial principles, however, students generally do not ask the tutor when they get stuck, and the tutor does not say whether the students’ ideas are right or wrong. COMET thus has no button for the student to ask for help and does not indicate whether the student’s entry is correct or incorrect. To recognize when and where the group needs help and give hints to help them continue their discussion, COMET’s strategies “focus group discussion” and “create open environment for discussion” are used when the group does not err but needs help to continue the discussion on a productive track. Strategies “avoid jumping the critical steps” and “deflect uneducated guessing” are forms of bug messages. Most strategies 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; if there is no student response or an incorrect response is given, the more specific parent medical concept node is used. Generating general and specific hints from BN nodes is consistent with hints generated by ANDES. In ANDES, if the student does not know what to do after receiving the first general hint, he or she can select a follow-up question by clicking on the three buttons: “explain further,” which gives slightly more specific information about the proposition represented by the node; “how do I do that?,” which finds the lowest probability child node assuming that is the node the student is most likely to be stuck on; and “why?,” which displays a canned description of the rule that was used to derive the node. COMET gives the next specific hint by using the medical concept of the highest probability node indicating the hypothesis the student is likely to know. If the students can still not come up with the hypothesis of interest, the strategy “refer to expert in the group” finds the student in the group who is most likely to know the answer. If this doesn’t work, COMET identifies this as a learning objective for study outside the session.

These results show that clinical reasoning gains for COMET-tutored students are similar to those for human-tutored students. Additionally, the number of hypothesis links created differed between groups. This is particularly true in light of our earlier study showing that on average 74 percent of human tutors used the same hint strategy and content as COMET.\(^9\) We believe the explanation lies primarily in the 26 percent 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.

One limitation of this work is that students’ interactions in the chat tool are beyond the capability of the student modeling module to interpret. It would be useful to add some language-processing capabilities so that the system can track and comment on the discussion. The model for each problem scenario required about one person month to build. Similar to training PBL tutors to have the skills that are necessary, creating the domain model is not a trivial task and requires significant expert knowledge.

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Pre-test</th>
<th>Post-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMET (1)</td>
<td>36.38 (0.70)</td>
<td>62.12 (0.69)</td>
</tr>
<tr>
<td>COMET (2)</td>
<td>36.58 (0.74)</td>
<td>61.33 (0.57)</td>
</tr>
<tr>
<td>COMET (3)</td>
<td>36.12 (0.76)</td>
<td>60.92 (0.69)</td>
</tr>
<tr>
<td>COMET (all)</td>
<td>36.36 (0.42)</td>
<td>61.45 (0.38)</td>
</tr>
<tr>
<td>Human tutor (1)</td>
<td>36.83 (0.64)</td>
<td>60.96 (0.51)</td>
</tr>
<tr>
<td>Human tutor (2)</td>
<td>37.42 (0.89)</td>
<td>58.63 (0.41)</td>
</tr>
<tr>
<td>Human tutor (3)</td>
<td>35.38 (0.70)</td>
<td>58.79 (0.55)</td>
</tr>
<tr>
<td>Human tutor (all)</td>
<td>36.54 (0.44)</td>
<td>59.46 (0.31)</td>
</tr>
</tbody>
</table>

October 2009 ■ Journal of Dental Education 1185
An authoring system needs to be developed by employing medical resources like the Unified Medical Language System (UMLS) Semantic Network to assist in creation of new cases. Finally, PBL typically occurs over a period of several days, with students carrying out individual learning tasks and bringing their learned knowledge back to the group. The current version of COMET has the additional capability to support single-session group PBL. The support for this aspect of PBL including investigating the effectiveness or utility of short-span PBL sessions and PBL integration into dental curriculum as a whole should be addressed.

**Conclusion**

This article has described research aimed at providing a general domain-independent framework for intelligent tutoring that facilitates clinical reasoning skill in PBL. COMET is the first system that combines aspects of ITS and CSCL to student and pedagogic modeling in collaborative problem-based learning environments. COMET utilizes probabilistic modeling of students and groups combined with pedagogic strategies elicited from real PBL sessions to offer feedback to students. Rather than giving answers or explanations, COMET actively and opportunistically provides guidance that is more than a confirmatory or negative feedback to students. The evaluation of the system shows that student clinical reasoning gains from COMET are similar to those obtained from human-tutored sessions.

**Acknowledgments**

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

**REFERENCES**