Implementing Tutoring Feedback into a Clinical Reasoning Learning Simulator

Annabelle CHARNEAU⁠¹, Froduald KABANZA⁠¹, Guy BISSON²,

Sylvain CLAVETTE⁠¹ and Marc FRAPPIER⁠¹

University of Sherbrooke
Sherbrooke, Québec J1K2R1, Canada
¹ Department of Computer Science
annabelle.charneau@usherbrooke.ca, kabanza@usherbrooke.ca,
sylvain.clavette@usherbrooke.ca, marc.frappier@usherbrooke.ca
² Faculty of Medicine
Guy.Bisson@Usherbrooke.ca

Abstract. TeachMed is a patient simulator for clinical reasoning learning (CRL). It allows students to evaluate a simulated patient, in a one-on-one problem-based learning mode, by iteratively collecting evidences and formulating hypotheses until converging to one or two final hypotheses, that is, the most probable diagnosis. Evidences are collected by asking questions to the patient, performing a physical exam and ordering labs. In this paper we discuss our ongoing work for modeling tutoring feedback into TeachMed by using Bayesian networks. The student’s actions (e.g., evidence collections and hypothesis formulations) are monitored to provide tutoring feedback based on the quality of these actions with respect to a Bayesian network modeling expert clinical knowledge for the patient case.

Keywords: Intelligent Tutoring System, Clinical Reasoning Learning, Bayesian Network, Influence Diagram.

Introduction

Medical diagnostic problems that are the focus of clinical reasoning are examples of complex decision making processes for which solutions depend not only upon clinical knowledge and medical experience, but also on well drilled strategies for collecting evidences about the problem presented by the patient, analyzing them, evaluating them based on collected evidences and formulate appropriate hypotheses. Clinical reasoning learning (CRL) in medicine deals with teaching medical students to acquire the basics of such cognitive strategies and to apply them efficiently [5].

In many medical schools, CRL is taught using a problem-based learning approach (PBL). Students, in small groups (typically six to eight students) under the direct supervision of a tutor who is a domain expert, are presented a clinical case and asked to diagnose it, propose a treatment and based on the pedagogical objectives a management plan. In our institution, CRL sessions are done under the supervision of a professor where a student play the role of a patient based on a scenario prepared according to pedagogical objectives. Another student
play the role of the attending physician and the others observe the encounter and will give feedback to the ‘attending physician’ when appropriate. The ‘attending physician’ will begin to collect evidences by asking questions to the patient (e.g., the motive of the consultation, in which parts of the body he is experiencing problems, how long and frequently the symptoms occur, his lifestyle, and so on), then they will proceed with a directed physical exam (e.g., sensing reflexes, auscultation,…) and after that order laboratory/imaging tests (e.g., urine tests, blood tests, X-rays or MRI). The evidences are analyzed and then the hypotheses are reevaluated in an ongoing process to discard those not fitting with his analysis, to keep those still uncertain, and to collect again new evidences in order to discard or confirm remaining hypotheses or generate new ones. This process of evidence collection, evidence analysis, hypothesis evaluation and problem reformulation is iterated until the students are able to narrow their list of probable hypotheses to the one or two most probable diagnosis.

Students conduct their investigation based upon their declarative knowledge they have learned in PBL sessions and also on their organized clinical knowledge database which is the result of similar cases they have already seen in clinical settings or in previous CRL sessions. The clinical knowledge linking symptoms to diseases is at least in part a probabilistic exercise whereas the actions used to collect evidences (i.e., questions, physical exams and lab tests) involve tradeoffs about costs, harm to the patient, and the value of the information obtained from the tests with regard to the current hypotheses. The professor uses his own experience and knowledge to provide hints to the students during their investigation and sometimes to comment on variant of similar cases.

Because of the cognitive complexity of CRL, it is crucial that students have a varied, very well organized and structured knowledge base, and be exposed to as many cases as possible in their curriculum, since this is mostly an experience-based learning process. To foster the acquisition of their clinical reasoning skills, it is mandatory that the CRL involves small group sizes. One of the drawbacks is that this learning activity requires a lot of resources in terms of professor’s availability for supervising the CRL. These professors are for the most part also physicians and/or researchers, and since most countries are faced with a severe shortage of physicians, the burden is even higher on them.

Many medical schools are therefore turning towards the use of software simulations to support at least in part the CRL activities. One possible strategy is to keep an affordable number of CRL courses directly supervised by professors and to have students practice on computer-based cases that are variants of those learned in classes, in groups or individually, with the assumption that computer-based cases will require little professor supervision or no supervision at all if the computer has enough intelligence capabilities to automatically provide the required feedback to students.

Providing feedback to students in computer-based CRL requires a system that not only is able to assess the student’s strengths and weaknesses of his clinical reasoning process to solve a patient case, but is also able to apply metacognitive strategies to identify and correct flaws in his reasoning strategies, for instance, with regard to how he interleaves evidence collection and the hypothesis formulation processes. However, computer-based CRL systems proposed to date focus mostly on providing feedback with regard to the application of declarative clinical knowledge, with very little emphasis on feedback related to cognitive strategies about the sequencing of clinical reasoning sub-processes [2,3,8,9].

For instance, DxR is one of the most used commercial CRL software with a database of about a hundred clinical cases, covering both patient diagnosis and management and illustrates this kind of approach [2]. However, DxR has no built-in student’s model other than a simple decision tree checking whether the student has reached the correct diagnosis within a given deadline and has considered all relevant hypotheses. The decision tree is explicitly specified by a professor who authors the case. Consequently, students using DxR autono-
mously learn mostly by reinforcement as the decision tree indicates whether they have the right hypotheses or not, but with little ongoing feedback about their reasoning processes. Other commercial CRL such as VIPS have the same limitations [9].

The Adele [3] and COMET [9] systems provide more advanced automated tutoring feedback by using an influence diagram (i.e., a Bayesian network with action nodes and action utility functions) to model the clinical knowledge about the patient. In fact, the influence diagram represents expert knowledge about how symptoms probabilistically relate to diseases and how evidence-collect actions or clinical-knowledge-apply actions relate to evidences. It is essentially the same kind of Bayesian network used in medical decision support systems to recommend diagnosis, to collect new evidences, or treatment to clinicians in real-life cases [6,7]. The generation of tutoring feedbacks is done by comparing the recommendations obtained from the network to actions, evidences or hypotheses found by the students and providing hints accordingly.

More specifically, at any step of the interaction between the student and the CRL interface that simulates a patient, we can recognize the evidence collected by the student and we can also determine the hypothesis he is currently working on either by asking them to the student directly, or by inferring them. Thus, given the current evidences as collected by the students, we can use Bayesian inference algorithms to determine the most likely hypotheses from the expert point of view and compare them with the hypotheses currently inferred by the student. Based on the comparison, we can for example give hints to the students about whether he is on the right track or not, or give him contextualized clinical knowledge that is most appropriate to help him. Conversely, given the current hypotheses collected by the student, we could use again Bayesian inference algorithms to determine the most valuable evidences to collect next, and exploit this information to give hints about which evidences to collect next and how.

In this paper we discuss how such an approach is being implemented into TeachMed, a patient simulator that we have developed, and present some use-case scenarios that illustrate the approach. We begin in the next section by first presenting the TeachMed. Then we discuss the implementation of tutoring feedbacks, followed with some illustrative scenarios. We conclude with a discussion on future developments.

2 Patient Simulation System

TeachMed has a graphic user interface (GUI) that provides general functionalities similar to DxR [2] and VIPS [9]. The student GUI allows the student to select a patient case, access the patient’s medical record, interview the patient, proceed with a physical exam on a 3D model of the patient (Fig. 1), interact with a laboratory module to order tests (Fig. 2), and finally treat and manage the patient. Professors use another GUI to author patient cases.
3 Automated Tutor

A key original component of TeachMed is an automated tutor that monitors the student’s actions to provide rapid feedback. Tutoring feedback is implemented based on a clinical influence diagram, the current system state, feedback rules triggered by the current system state, and a clinical reasoning automaton that sequences clinical reasoning sub-processes and corresponding feedback rules.

3.1 Clinical Influence Diagram and Bayesian Inference Processes

A clinical influence diagram is a Bayesian network encoding the clinical diagnostic knowledge from an expert point of view. It specifies causal links between evidences, symp-
toms and pathologies with corresponding probability distributions, as well as actions that investigate these symptoms with corresponding utilities. Fig. 3 shows examples of action nodes (rectangular, begin with “A_”), evidence nodes (“E_” nodes) and hypothesis nodes (“H_” white nodes) extracted from our influence diagram for a case of pelvic pain. Probabilities and utility specifications are omitted for clarity. The fragment illustrates only some nodes with few evidence-collection actions (questions to the patient and lab tests). The entire model for the case contains 70 nodes, including physical exam actions and curriculum-knowledge application actions to collect evidences.

The influence diagram is implemented using the Smile Bayesian network library [1]. Evidences gathered by the student are recorded into a student evidence table. Given this table, we use Smile library to compute the posterior probability for a set of hypotheses from the expert point of view and rank them according to their likelihood, that is, the expert hypothesis table. Intuitively, assuming the influence diagram correctly models the expert clinical knowledge, the expert would infer these hypotheses from the given evidences. On the other hand, based on his evidence table, the student will infer his own hypothesis table (i.e., the student hypothesis table) which may or may not match the expert hypothesis table.

Fig 3. Extract of the Influence Diagram for the pelvic pain case

In principle, the next best evidence-gathering process should be one yielding a value of information that exceeds the cost of gathering information. However, in general it is very difficult to express the utilities of evidence-gathering actions with a level of accuracy that takes into account all possible evidence observations [6,7]. Therefore, we do not rely entirely on the utilities specified for evidence-gathering actions to provide feedback about the next evidence-gathering process. Rather we use these utilities as heuristic input to feedback rules that provide the actual hints about the next evidence-gathering steps. More specifically, based on the current student’s evidence and hypothesis tables, we use Smile library to compute evidence-collection actions and store them into a next student evidence-action table, on top of which we can express facts that are preconditions of feedback rules. For instance, we can specify a feedback rule stating that “if not Ectopic pregnancy and contra-
ceptive method question’s value of information is less than sexual partner question no more by x threshold, then perform contraceptive method question”.

In order to provide feedback TeachMed needs to determine the current hypothesis the student is trying to confirm or reject (i.e., the student’s hypothesis goal) or the current evidence he is trying to collect (i.e., the student’s evidence goal). We could obtain this information by simply asking it to the student, but this would result in too much intervention, disturbing the tutoring strategy, particularly for advanced students who just need occasional help. Instead, based on the observed student’s actions (i.e., questions he asks, physical exams he makes, or lab tests he orders), we use Smile library to infer the evidence he is trying to collect, similarly for the student’s hypothesis goal. If the output of this goal recognition inference is below a validity threshold, we prompt a question to the student to get directly from him the goal he is working on. If the recognition output is a list of at most three elements with valid probability above the threshold, the question looks like “are you trying to establish x, y or z?” Otherwise, the question is a direct one such as “what hypothesis are you trying to establish?”

In practice, experienced practitioners formulate hypotheses unconsciously, particularly in the early stages of their cognitive investigation [5]. They start to think about them explicitly only in later stages as their reasoning sub-processes become more involved, requiring them to weigh the benefits and disadvantages of the final decisions they are about to make (e.g., ordering a costly MRI versus waiting for the symptoms to develop further in order to enable clearer evidences). Nevertheless, when teaching, it is a good strategy to require novice students to express explicitly their current hypotheses they have, even during the early stages of their investigation; that way, the professor can evaluate for example whether their evidence gathering is indeed guided by their hypotheses and their validity. For this purpose, TeachMed’s menu provides the possibility for the student to update directly his evidence table and his hypothesis table, if this is a requirement set by the professor when he authored the case. Nevertheless, for tutoring feedback we cannot entirely rely on these tables because students may forget to record the hypothesis in trying to collect evidences based on hypotheses in their minds but not in the table. In fact, TeachMed still invokes goal-recognition to infer the hypothesis the student may be working on and if different from those in his edited table, an appropriate feedback will be triggered.

3.2 Student Model and Model Tracer

TeachMed maintains a student model, reflecting the student’s progress information inferred from the influence diagram as explained above (i.e., student evidence table, expert evidence table, student’s hypothesis table, the student’s hypothesis goal, and the student’s evidence goal), and student’s input for feedback request (e.g., the student can ask TeachMed to help him decide between two interpretations of a lab test). A process, called the student’s model tracer, monitors the student’s actions and periodically invokes the update processes described above to maintain the student’s model.

3.3 Feedback Rules and Feedback Generator

The feedback generator is a process that periodically checks the current student’s model to trigger tutoring feedback to the student, using production rules that are preconditioned on
data in the student’s model and having consequents that are tutoring hints. These are “teaching” expert rules and can be as efficient as the available teaching expertise allows.

To illustrate, a feedback rule of the form “if student-asks-help (interpret, test, results) then hint-test-interpretation (test, results)” will trigger a hint helping the student about the interpretation of a lab test. Following Andes’ approach [4], hints are implemented using template-based heuristic dialogues, with slots replaced by data in the student’s model. With the previous example, TeachMed will engage a tree-structured dialog with the student by asking him his impressions about the interpretation and issuing hints until the dialog concludes with the correct interpretation. Feedback rules can also be triggered by TeachMed initiative by having rules with antecedents based upon the evaluation of the student’s progress as reflected by the student’s model.

4 Example Scenarios

4.1 Case No. 1: Feedback while changing hypotheses

At given point, the student may be asking a series of questions to the patient focusing one particular hypothesis among his current list, to confirm or reject it. We say that this is the working hypothesis. Usually it’s not a good idea to switch the focus of the working hypothesis back and forth without exhausting all relevant questions on the current working hypothesis, until more evidence is introduce in to make re-activation of the hypothesis relevant. Therefore, whenever the student asks a question that is irrelevant to the working hypothesis, TeachMed determines: (1) whether the student has changed his working hypothesis without explicitly stating it (hence the question may actually be relevant to the new working hypothesis); (2) or whether the student is still assuming the same working hypothesis (hence the current question suggests a lack of sufficient clinical knowledge by the student).

In case 1, TeachMed checks whether the student has exhausted all questions related to the working hypothesis at this point. We have two possibilities: (a) if the student has indeed exhausted all relevant questions, Teachmed displays a message, as feedback to the student, asking him to formulate his new working hypothesis; (b) otherwise, Teachmed initiates a dialogue with the student aimed at making her realize that there are still evidences to collect using questions, for the current working hypothesis; the dialogue may involve strategic references to clinical knowledge related to the case. In case 2, TeachMed initiates a dialogue aimed at making the student that his current question is irrelevant to the current working hypothesis, again referring whenever appropriate to clinical knowledge about the case.

The following dialogue illustrates Case 1.a:

<table>
<thead>
<tr>
<th>Collected Evidence:</th>
<th>Collected Evidence: Acute lower abdominal pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student:</td>
<td>Working Hypothesis: Urinary infection</td>
</tr>
<tr>
<td>Student:</td>
<td>Question: “Any fever?”</td>
</tr>
<tr>
<td>TeachMed Patient:</td>
<td>“I don’t know.”</td>
</tr>
<tr>
<td>Student:</td>
<td>Question: “Could you describe your pain? Is it like a cramp, a burning sensation?”</td>
</tr>
<tr>
<td>TeachMed Patient:</td>
<td>“It’s not a cramp but it’s very painful.”</td>
</tr>
<tr>
<td>Student:</td>
<td>Question: “Is it worse on one side?”</td>
</tr>
</tbody>
</table>
TeachMed Patient: “No.”
Student: Question: “How many times did you urinate since the beginning of your pain?”
TeachMed Patient: “I don’t know. More often I think”
Student: Question: “Do you have a burning sensation on urination?”
TeachMed Patient: “No.”
Student: Question: “Do you have a sexual partner?”
TeachMed Tutor: Is this question relevant to the hypothesis you are working with? (Yes/No, I’m examining another hypothesis.)
Student: No, I’m examining another hypothesis.
( Gathered enough evidence to dismiss “urinary infection” as a hypothesis )
TeachMed Tutor: Formulate your new working hypothesis.
Student: Working Hypothesis: Sexually transmitted disease (STD)

The following dialogue illustrates Case 1.b:

Collected Evidence: Acute lower abdominal pain
Student: Working Hypothesis: Urinary infection
Student: Question: “Could you describe your pain? Is it like a cramp, a burning sensation?”
TeachMed Patient: “It’s not a cramp but it’s very painful.”
Student: Question: “How many times did you urinate since the beginning of your pain?”
TeachMed Patient: “I don’t know. More often I think”
Student: Question: “Do you have a burning sensation on urination?”
TeachMed Patient: “No.”
Student: Question: “Do you have a sexual partner?”
TeachMed Tutor: Is this question relevant to the hypothesis you are working with? (Yes/No, I’m examining another hypothesis.)
Student: No, I’m examining another hypothesis.
( Didn’t gather enough evidence to dismiss “urinary infection” as an hypothesis )
TeachMed Tutor: Did you finish working with “urinary infection” as your working hypothesis? (Yes/No)
Student: Yes.
TeachMed Tutor: Are you sure? (Yes/No)
Student: Yes.
TeachMed Tutor: What do you think about the “body temperature”? (It isn’t related to “urinary infection”./It is related to a “urinary infection”./I don’t know.)
Student: It is related to a “urinary infection”.
TeachMed Tutor: Ask the relevant question.
Student: Question: “Any fever?”
TeachMed Patient: “I don’t know”.
Student: Working Hypothesis: STD
TeachMed Tutor: Have you finished with “urinary infection”? (Yes/No)
Student: Yes.
TeachMed Tutor: Are you sure? (Yes/No)
Student: Yes.
TeachMed Tutor: What do you think about “lateralization of the pain”? (It isn’t related to “urinary infection”./It is related to a “urinary infection”./I don’t know.)
Student: I don’t know.

TeachMed Tutor: “lateralization of the pain” influences negatively “urinary infection” as your working hypothesis? (You can verify it in the library).

TeachMed Tutor: Ask the necessary question.

Student: Question: “Is it worse on one side?”

TeachMed Patient: “No.”

Student: Working Hypothesis: STD

3.2 Case No. 2: Skipping steps in the reasoning activity

Students tend to conduct unstructured hypothesis formulation. They often consider a specialized symptom deep in the clinical knowledge network before examining its parent, more general symptoms. For example, the student may consider Salpingitis hypothesis before examining Gynecology or Gynecologic Infection (Fig. 4). In such case, TeachMed asks the student to give a more general hypothesis first. Then: (1) if the student is able to find one, TeachMed let him continue working with it; (2) if not, TeachMed begin a dialogue with the student, by point him to evidences that would make him realize which hypothesis are parents for the one he is currently considering (3) if this dialogue fails without the student being able to realize the relevant hypotheses, TeachMed will simply give the answer.

The following dialogue illustrates this case:

Collected Evidences: Young woman, acute abdominal pain not lateralized, temperature (38.5C)

Student: Added hypothesis: Salpingitis

TeachMed Tutor: Could you specify your hypothesis hierarchy before proposing that hypothesis? (Yes/No)

Student: Yes.

TeachMed Tutor: Formulate your hypothesis.

Student: Added hypothesis: Gynecology

Student: Added positive link: Gynecology – young woman

Student: Added positive link: Gynecology – abdominal pain

Student: Added hypothesis: Salpingitis

TeachMed Tutor: Could you specify your hypothesis hierarchy before proposing that hypothesis? (Yes/No)

Student: No

TeachMed Tutor: How about pain location? fever? Can you give me an hypothesis related to these evidences? (Yes/No)

Student: Yes

TeachMed Tutor: Formulate your hypothesis.

Student: Added hypothesis: Gynecologic Infection

Student: Added positive link: Gynecology – Young woman

Student: Added positive link: Gynecology – Abdominal pain not lateralized

Student: Added positive link: Gynecology – Temperature (38.5C)
5 Conclusion

Future work on this project includes the implementation of an interface between TeachMed and a real-world electronic health record to feed information directly to TeachMed based on pedagogical objectives. As noted by Gertner et al. in their physics tutoring system [4], a Bayesian network implicitly models solutions, but not the solution process itself. Yet, being able to model the solution process can improve the quality of feedback given to the student. A similar observation holds in CRL, with the difference that the solving process is harder to formalize other than giving general strategies about the interleaving of the different clinical reasoning sub-processes. Nevertheless, we feel that the possibility of having one working expert-solution, albeit not encompassing all experts’ solutions, can be exploited to provide enriched feedback. In the more trivial case, we could illustrate the solution to the student as one example of expert’s approach. Given that an electronic health-record contains clinical reasoning traces in the audit trail, the challenge will be to find techniques for inferring expert solving processes from those actual expert solutions in real life cases.

References

2. DxR Clinician: http://www.dxrgroup.com/