Degrees of Mixed-Initiative Interaction in an Intelligent Tutoring System

Reva Freedman

Department of CSAM
Illinois Institute of Technology
10 W. 31st Street 236–SB
Chicago, IL 60616
freedman@steve.csam.iit.edu

Abstract

We are currently implementing CIRCSIM-Tutor v. 3, a conversation-based intelligent tutoring system (ITS) which tutors medical students on the baroreceptor reflex, a topic in cardiovascular physiology. In order to provide the most natural conversational experience possible, we would like to let the student take the initiative where possible. On the other hand, because of the increased complexity of the required infrastructure, the difficulty of understanding full free-text input, and the tutor's desire to accomplish the tutoring agenda, we must restrict the types of initiatives which the system will attempt to respond to. We classify initiatives according to the nature of the student's utterance and according to the type of processing required by the tutor to handle them. We describe how we encourage the student to give responses we can handle. We explain why we believe that these methods do not restrict the student's ability to communicate with the system or to learn the material. We illustrate the phenomena described with examples from human-to-human tutoring sessions.

Introduction

The SCHOLAR system, developed by J. R. Carbonell (1970), is often considered in the United States to be the first intelligent tutoring system. Carbonell called SCHOLAR a *mixed-initiative* system because it had two modes of operation: the teacher could ask the student questions or vice versa. However, SCHOLAR could not hold a continuous conversation which required cooperative dialogue behavior. In fact, SCHOLAR did not attempt to create a coherent conversation. The following excerpt is from a tutor-led section of a dialogue with SCHOLAR; the student-led sections are similar.

T: The capital of Chile is Santiago. Correct or incorrect?
S: Correct.
T: That's right. Approx. what is the area of Brazil?
S: 2,500,000 square miles.
T: Wrong. Please indicate if the following statement is correct or incorrect: The area of Paraguay is approx. 47432 square miles.

(excerpted from Carbonell 1970, fig. 1)

Our tutoring system, CIRCSIM-Tutor, is capable of participating in this type of mixed-initiative interaction, although our coherence heuristics would not generate such an abrupt topic change. But we cannot handle more sophisticated types of mixed-initiative processing such as fully cooperative conversations. We define a fully cooperative conversation as one where participation requires a multi-agent model where each speaker has goals to achieve and a plan to achieve them. The intent is to distinguish conversations requiring such a model from those which can be understood with a simpler model where one agent has a goal and a plan to achieve it, and the other agent's job is mainly to go along with the plan. It has been suggested that a fully cooperative conversation can be compared to two children building a tower of blocks together, as opposed to one child telling the other what to do. In the cooperative case, neither may be able to predict the shape of the resulting tower.

This brings up the question: Does a tutoring system like CIRCSIM-Tutor need to participate in fully cooperative conversations in order to achieve its tutoring goals? After all, most conversations between humans and computers today are led by one party or the other. For example, in a database front-end, the human user might ask questions which the program answers. Conversely, in an advice-giving system or an automatic teller machine, the program leads and the human's job is to respond. Many human-to-human tutoring sessions are also led mainly by one party or the other. In this regard, one might contrast our person-to-person tutoring experiments, where the tutor has an agenda to fulfill, with tutoring sessions such as those

This work was supported by the Cognitive Science Program, Office of Naval Research under Grant No. N00014–94–1–0338, to Illinois Institute of Technology. The content does not reflect the position or policy of the government and no official endorsement should be inferred.

described by Fox (1993), where the student chooses the agenda for the tutoring session.

In this paper we characterize the types of mixed-initiative interactions which CIRCSIM-Tutor can participate in. We classify potential student utterances and the reasoning power required by the tutor to respond to them. We also describe types of mixed-initiative interactions which the system cannot handle and how we reduce the frequency of such occurrences.

The CIRCSIM-Tutor Planner

Problem Domain and Tutoring Task

CIRCSIM-Tutor conducts conversations with students about the baroreceptor reflex, a topic in cardiovascular physiology which all beginning medical students need to master. The baroreceptor reflex is the negative feedback loop which attempts to maintain a steady blood pressure in the human body. The focus of CIRCSIM-Tutor is on tutoring students on material which they have already studied, not on teaching new material. This orientation leads us to place greater emphasis on the interactive aspects of the conversation, rather than on elaborate ways to present material. Thus generating cohesive turns which fit coherently into the evolving dialogue is more important than generating complex explanations, for example.

In the beginning physiology course, students are given problems to solve based on a simplified qualitative model of the heart. In each problem, something happens to change the processing of the heart. The student is then asked to predict the direction of change of seven core variables in each of three resulting physiological stages. Students can solve such problems in many settings: alone, in conversation with a human tutor, or while interacting with a tutoring system. When the student's predictions are complete, either the human tutor or the ITS conducts a dialogue with the student to help the student learn the correct answers and the reasoning behind them.

Planning the Tutor's Response

In order to model both pedagogical and linguistic strategies, the CIRCSIM-Tutor project has collected over 5000 turns of keyboard-to-keyboard tutoring sessions where students solve problems with the aid of live tutors. Our analysis of these transcripts shows that CIRCSIM-Tutor needs both a global, plan-oriented model and a turn-by-turn model such as that used by the Conversation Analysis school (Sinclair & Coulthard 1975).[1]

The tutor maintains global control of the conversation while responding turn by turn to the student's utterances. To be able to respond flexibly to the student, the planner refines plans only until the next move is clear. Thus it can afford to replan whenever the student gives an unexpected response.

The global plan is visible in the hierarchical structure of the dialogue, which contains the following levels:

- Physiological stage
- Core variable
- Attempts to teach each variable

Within each physiological stage, the tutorial dialogue is divided into segments, one for each core variable. Usually, only the variables which the student missed are discussed. Each segment is divided into one or more attempts to teach the value of a variable. The basis of each attempt is a correction schema chosen from a plan library. The schema contains one or more goals which must be satisfied in the coming turn or turns. Goals are satisfied recursively until primitive speech acts are obtained.

Each attempt ends with the tutor requesting the correct value of the variable being tutored. If the student gives the correct answer, the segment ends. Otherwise, the attempt fails, causing any remaining goals associated with it to be removed from the agenda. The tutor can make another attempt or give the student the answer. A more detailed description of the CIRCSIM-Tutor planner can be found in Freedman (1996).

The speech acts are assembled into turns, then realized as surface text and displayed to the student. In a typical tutorial dialogue, every turn has the following basic structure. Note that each of the sections is optional.

Response to student's previous statement
  Acknowledgment of student's statement
    (e.g. *yes, no, you're right*)
  Content-oriented reply
    (e.g. rebuttal or statement of support)
New material
  Next part(s) of current schema
  Question for student to answer

This model, based on the tenets of Conversation Analysis, is regularly observed in our human-to-human transcripts.

Although human tutors don't necessarily do so, the mechanized tutor always ends a turn by explicitly requesting information from the student, usually with a question.

Some Sample Dialogues

The following is the most common schema used to correct one category of variables, those controlled by the nervous system.[2]

[1] This work was brought to our attention by Cawsey (1992).

[2] The actual schemata are written in LISP. For simplicity, the prerequisites are not shown.

Correct-neural (V):

1. Make sure student knows that V is controlled by the nervous system.
2. Make sure that student knows that current stage is pre-neural.
3. Make sure student knows that correct value of V is no-change.

Figure 1 shows in condensed form a subset of the dialogues which can be generated from this schema. Each item in italics represents text which could be generated by one element of the schema. Where convenient, we have shown two possible realizations separated by a slash. The numbers correspond to the subgoals in the schema. The student's responses are shown in roman type.

In the leftmost path, the student gives the desired answer immediately. In the second path, the student gives a partially correct answer, in this case an answer which is true but does not use the tutor's desired language. The tutor adds a goal to correct the student's language before continuing with the schema. In the third branch, the student gives an answer which is on the path toward the correct answer. The tutor helps the student toward the correct answer in two different ways. In the rightmost branch, the tutor chooses to give the student the answer.

Due to space limitations, we have only shown one possible realization for the second and third subgoals of the schema. As multiple realizations are possible for these subgoals as well, one can see how the CIRCSIM-Tutor planner can generate a large number of coherent dialogues from a few basic elements. Although CIRCSIM-Tutor does not have the syntactic ability or semantic range of a human tutor, it can emulate all of the major dialogue patterns used by our domain experts.

Handling Student Initiatives

Classification of Student Input

The reply generated by the system depends on the type of the student's utterance. The following response types, which are the most common ones, were illustrated in Figure 1:

- Correct answer
- Wrong answer
- Physiological near-miss: a step toward the correct answer
- Linguistic near-miss: linguistically close but not exact answer

Each of these response types can refer to the tutor's most recent question or to a higher question on the agenda.

The student can also say something which does not relate to an open question posed by the tutor. Here are two examples:

- Student adds new information, e.g. an explanation.
- Student changes the topic.

For the purposes of this paper, we will consider any student statement which adds new content to the dialogue as a student initiative. This definition subsumes a narrower definition which would only include statements which change the topic of conversation. The broader definition is convenient because much of the processing is similar.

Problems with Unrestricted Student Initiatives

Student initiatives present a difficult problem. On the one hand, we would like to let students respond as freely as possible. Open-ended questions force students to think more deeply. Additionally, open-ended questions permit students to focus in on their specific problems. This ability is one of the justifications for building natural-language based ITSs.

But there are several problems with permitting unrestricted student initiatives:

- First, the student's utterance can be too difficult to understand at the purely mechanical levels of spelling, syntax and basic semantic processing.
- Second, just because we can understand a statement at a literal level does not mean that we can understand the student's model of the domain (Borchardt 1994). This can be especially difficult if the student's domain model is invalid.
- Third, even if we can understand what the student is telling us, we may not have a constructive response available.
- Fourth, even if we have a constructive response available, responding to the student initiative may not help the tutor achieve its agenda.

Reducing Unwanted Student Initiatives

Asking the students to learn a set of rules about what they can and cannot say detracts from the goal of a natural conversation. Therefore we rely on the fact that students are cooperative conversation partners (Grice 1975) to encourage the type of response we prefer.

One way to avoid difficult student initiatives is to ask short-answer questions instead of open-ended ones. Reducing the size and complexity of the expected response reduces the chance of misunderstanding a student utterance and the attendant frustration on the part of the student. In fact, this informal restriction on our part encourages students to use the system to its limits. Students are sensitive to the capabilities of the system, and they won't generate non-trivial utterances if we have proved ourselves incapable of responding to them before.

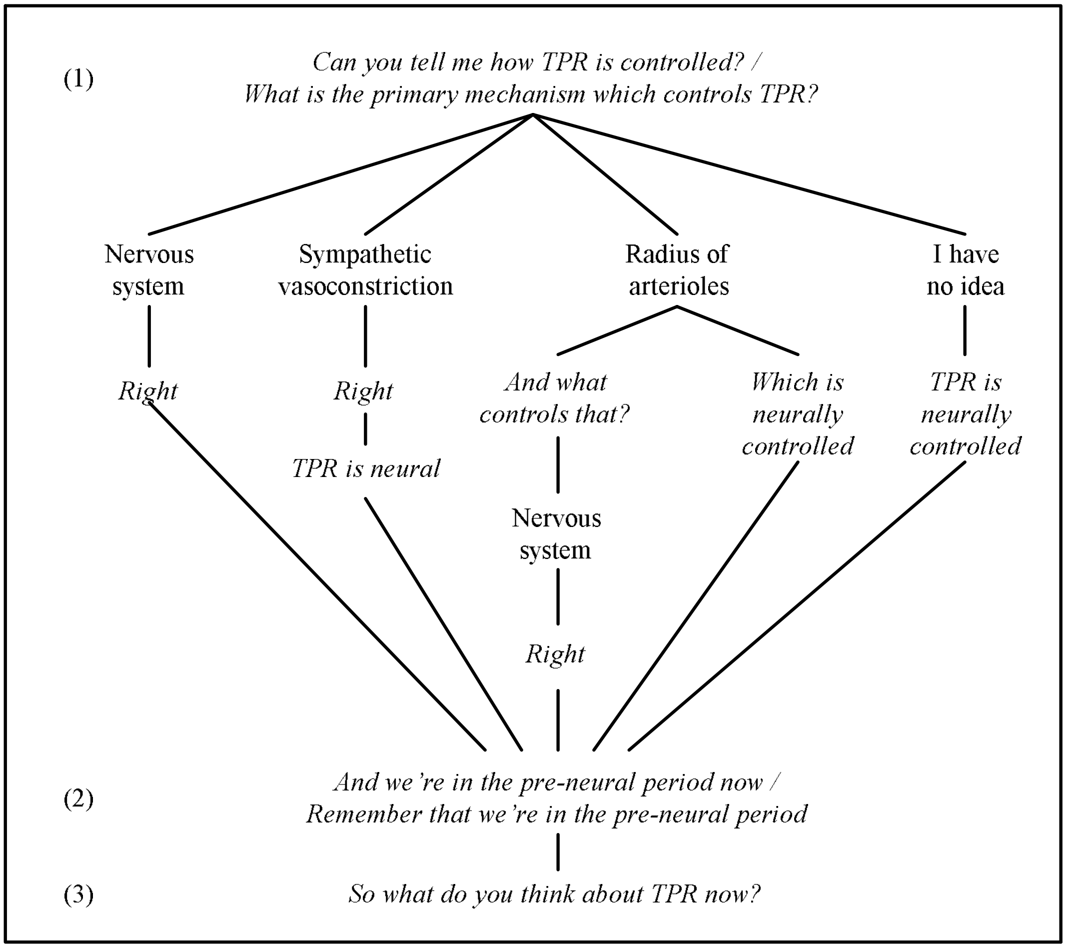

Figure 1: Dialogues which CIRCSIM-Tutor can generate

A second way to restrict unwanted student initiatives is to ensure that each turn ends with an explicit request. With an explicit request on the table, it is more likely that the student will answer the question rather than change the topic. An additional benefit of this restriction is that it provides students with an unambiguous indicator of when it is their turn to respond. Turn-taking rules work in person-to-person conversation because we are socialized to understand and use them (Sacks, Schegloff & Jefferson 1974). But people expect to have a more explicit interaction with a computer (Dahlbäck & Jönsson 1991).

Although it might seem that restricting the types of discourse we generate and hope to receive might diminish the teaching ability of the system, this is not necessarily so. According to one of our domain experts[3], human tutors have two problems with open-ended questions:

• Even with human tutors, students frequently misunderstand what the tutor is asking. Hence, even when the tutor can understand the student's response, it may not relate directly to what the tutor is looking for.

• It is often difficult for a human tutor to understand (and hence to use) what the student says in response to an open-ended question. There are two reasons for this. First, students do not always understand to what organizational level (e.g. heart, myocardial cells, or membrane receptors) the question refers to. Second, medical students do not naturally reason causally. When they are having trouble with causal reasoning, they sometimes switch to using teleological statements ("what the myocardial cells want is …"), and these can be difficult for the human tutor to interpret.

Thus human tutors have the same type of problems with open-ended questions as the computerized tutor, although not nearly to the same extent.

Desirable Types of Student Initiatives

On the other hand, we would like to encourage student initiatives whenever we can understand and respond to

[3] Dr. Joel Michael, personal communication.

them. In general, the current CIRCSIM-Tutor infrastructure can handle an initiative which satisfies the following criteria:

- Does not require a deep understanding of the student's (possibly buggy) domain model.
- Does not require reasoning about the student's plan.
- Does not require the use of stacked discourse contexts, as would be required, for example, if the student asked a hypothetical question.

Although classifying the student's response requires identifying the student's goal, most student statements can be handled without needing to reason about the student's plan. Transcript analysis shows that for the type of problem-solving involved in the baroreceptor reflex domain, expert tutors lead the conversation, although they give the student wide freedom in answering the questions. As one of the tutors' goals is to implicitly teach a model of problem-solving, this is not an unreasonable approach.

Stacked discourse contexts might be required if tutor and student needed to switch among multiple topics or to describe events in a variety of possible worlds. In general, these features are not required for discussing the values of variables and the causal relationships among them.

From another point of view, we can handle an initiative if the tutor can assimilate the student's desire into its own agenda. One category of initiatives satisfying these criteria is simple requests, e.g. asking for help, a definition or an explanation.[4] The tutor can respond to these and other simple initiatives in several ways:

- Put the student's request on the agenda above the current plan, i.e. respond to student's statement, then return to plan
- Put the student's request on the agenda instead of the current plan, i.e. switch to new plan
- Put the student's request elsewhere on the agenda
- Acknowledge student's input without responding
- Ignore student's input

In each of these cases, the student's request has been incorporated into the tutor's agenda.

Initiatives in Human Conversation

The following fragment from a human-to-human tutoring session shows some examples of cooperative phenomena which are outside the range of our model.

→ T: … How can CC go up if it's under neural control?
S: *(gives a confused explanation of a false statement)*
T: *(attempts to deal with student's confusion)* Do you see the difference?
S: No, this concept is hard for me to grasp.
T: *(explains further)* OK?
→ S: Is increased calcium the only thing that can increase contractility?
T: Yes.
→ S: OK. So would it be accurate to say that *(asks a follow-up question)*?
T: Yes, and *(gives further information)*. OK?
S: OK. (K10:43–52)

In the first line of this example, the tutor uses an open-ended question to encourage the student to give an explanation in terms of a deeper-level concept map. In the cooperative conversation which follows, both tutor and student contribute questions and ideas. In the second and third marked statements, the student takes the initiative, referring to concepts mentioned earlier in the dialogue which are now in the shared mental model of tutor and student. The more complex and/or error-ridden the student's explanation is, the less likely CIRCSIM-Tutor is to be able to understand it and reply constructively. For this reason we try to avoid such explanations on the student's part by not asking open-ended questions such as the one the tutor started with here.

On the other hand, consider the following example.

T: *(asks about the value of TPR)*
S: I thought TPR would increase due to higher flow rate through vasculature. (K10:32)

CIRCSIM-Tutor can handle a student response like this one. Although we cannot make use of the explanation given by the student, we can understand and respond to the core content ("TPR will increase"). In fact, the human tutor does not necessarily use the additional information either.[5]

In the following example, the tutor responds to a student statement for which the truth value can be determined but with difficulty. This example is intermediate in difficulty to the two previous examples. To answer the student's question correctly, the tutor needs a deeper domain model than that required just to compute the correct values of the core variables. Note that the tutor may need to expend considerable resources on such a statement before it can determine whether it can indeed understand and answer such a statement.

T: What's your first prediction?
S: TPR, HR, CC all increase simultaneously.
T: That's pretty good except for HR. Remember in this case this guy's HR is solely determined by his broken artificial pacemaker.

[4] Although a request for help can also be an indicator of a complex student plan, good results can often be obtained by taking such requests literally.

[5] Note, however, that if the student's reason is incorrect, the human tutor can correct the student in cases where the computerized tutor cannot.

→ S: Wouldn't his other myocardial cells respond to sympathetic stimulation and couldn't they override his artificial pacemaker?

T: They might and then again they might not. We're assuming in this case that they don't. So what do you say about HR? (K18:36–40)

This case is at the limit of our ability to understand natural language input. Although we strive to increase our ability to handle such input, we try to reduce the frequency of such utterances by ensuring that every turn ends with an explicit request of the student. For example, the tutor might terminate the turn previous to the marked one by asking:

T: … Now what do you think about the value of HR?

Conclusions

In a tutoring system like CIRCSIM-Tutor, the importance of cooperative conversation lies not in the conversation itself, but in interactivity, which keeps the student actively involved with the material. CIRCSIM-Tutor has several strategies for increasing interaction with the student in spite of the unavoidable limitations on natural language processing. In particular, by asking short-answer questions instead of open-ended questions and ending each turn with a question (or its equivalent), CIRCSIM-Tutor can maintain a general control over the conversation while still giving the student a great deal of leeway for discussing the desired subject matter.

CIRCSIM-Tutor can handle most things which happen in a tutor-led tutoring session, but it cannot handle a true cooperative conversation with two independently planning agents. Thus we do not attempt to handle student utterances which would require understanding the student's plan, extending our understanding of the student's domain model, or using stacked discourse contexts.

By accepting these restrictions, we were able to reduce the complexity of the architecture for CIRCSIM-Tutor v. 3. This has permitted us to reduce the response time and increase our content coverage. For a domain like the baroreceptor reflex, we believe that a tutor-led system can provide students with a large variety of coherent and helpful conversations. Since CIRCSIM-Tutor logs all of its conversations, we will be able to update our model as we acquire real-world experience.

References

Borchardt, G. C. 1994. *Thinking Between the Lines: Computers and the Comprehension of Causal Descriptions.* Cambridge, MA: MIT Press.

Carbonell, J. R. 1970. AI in CAI: Artificial Intelligence Approach to Computer Assisted Instruction. *IEEE Transactions on Man-Machine Systems* 11(4): 190–202.

Cawsey, A. 1992. *Explanation and Interaction: The Computer Generation of Explanatory Dialogues.* Cambridge, MA: MIT Press.

Dahlbäck, N. and Jönsson, A. 1989. Empirical Studies of Discourse Representations for Natural Language Interfaces. In Proceedings of the Fourth Conference of the European Chapter of the Association for Computational Linguistics, Manchester, 291–298.

Fox, B. 1993. *The Human Tutorial Dialogue Project.* Hillsdale, NJ: Lawrence Erlbaum.

Freedman, R. 1996. Interaction of Discourse Planning, Instructional Planning and Dialogue Management in an Interactive Tutoring System. Ph.D. diss., Dept. of Electrical Engineering and Computer Science, Northwestern Univ.

Grice, H. P. 1975. Logic and Conversation. In P. Cole and J. L. Morgan, eds., *Speech Acts* (*Syntax and Semantics*, v. 3), 41–58. New York: Academic Press.

Sacks, H., Schegloff, E. A. and Jefferson, G. 1974. A Simplest Systematics for the Organization of Turn-Taking in Conversation. *Language* 50(4): 696–735.

Sinclair, J. M. and Coulthard, R. M. 1975. *Towards an Analysis of Discourse: The English Used by Teachers and Pupils.* London: Oxford University Press.

Acknowledgments

Professor Martha W. Evens has shepherded the CIRCSIM-Tutor project since its inception. Professors Joel A. Michael and Allen A. Rovick of Rush Medical College were generous with their time and knowledge about both pedagogical and domain issues.