Using Qualitative Reasoning in the Classroom and in Electronic Teaching Systems

Evelyn Lulis (lulieve@iit.edu)
Department of Computer Science, Illinois Institute of Technology
10 West 31st Street, Chicago, IL 60616 USA

Joel Michael (jmichael@rush.edu)
Department of Molecular Biophysics and Physiology, Rush Medical College
1750 W. Harrison St., Chicago, IL 60612 USA

Martha Evens (evens@iit.edu)
Department of Computer Science, Illinois Institute of Technology
10 West 31st Street, Chicago, IL 60616 USA

Abstract

Qualitative reasoning is employed in teaching basic physiological principles to medical students. A review of electronic teaching/tutoring systems and an examination of human tutoring sessions revealed methods for implementing analogies in our electronic tutoring system.

Introduction

For the past thirty years Joel Michael and Allen Rovick have been teaching the first year Medical Physiology course at Rush Medical College in Chicago. Physiology is difficult for most medical students in part because it is heavily based in physics, a subject with which many students are not comfortable. As a result the baroreceptor reflex and other negative feedback systems are a major source of confusion for students. Michael and Rovick have written a number of Computer-Aided Instruction (CAI) programs to help their students learn this material. Medical students are very good at memorizing material, but rote memorization is not enough to enable them to understand and use this material effectively, so the focus of these CAI systems is on active learning, on helping the students learn to solve problems (Michael, 1993).

In this paper we describe how Michael and Rovick (Michael et al., 1990) came to emphasize qualitative causal reasoning and how they implemented systems that begin by asking the students to make qualitative predictions about the consequences of perturbations to the system. Next we discuss the current systems that they use with students: CIRCSIM, a conventional CAI version of this approach, and CIRCSIM-Tutor, which embodies this same approach but does so by generating a Socratic dialogue with the student. We go on to describe our current attempts to add facilities to teach more complex qualitative reasoning and to provide support for the reasoning process. We want to make our system understand and interpret student initiatives and student responses to open questions. We also want to implement analogies effectively and explain the underlying structural mapping to students who get confused.

MacMan and Heartsim

One way to initiate active learning is to get students to design experiments. When Dickinson et al. (1973) developed a mathematical model of the baroreceptor reflex at McMaster University, called MacMan, Rovick and Michael felt that this model might provide a way for students to design and carry out “experiments” in a computer laboratory. Although it was not easy to use (it could only be run via an acoustic coupler modem and a telephone line connected to the computer), they tried out the system at length and arranged for some students to try it as well. The user could specify what parameters to change and what variables to display (either by graphing them or printing a table of outputs) as the model ran. Thus, it was feasible for students to hypothesize a relationship between two variables and discover how one changed as a function of another. These first attempts to use MacMan were so promising that Michael and Rovick decided to try to convert it into a PLATO program to make it accessible to the whole class. The PLATO system (Hody and Avner, 1978) was a computer-based instructional system at the University of Illinois at Champaign-Urbana. The first step was to obtain permission to translate the original FORTRAN model into TUTOR, the authoring language for the PLATO system. Michael and Rovick wrote a PLATO interface for the program and a laboratory manual that described experiments to be carried out by varying parameters systematically and recording the results. The manual also included questions for the students to answer and discussed phenomena that the students should observe.

This PLATO version of MacMan was only used for a few years because the students did not seem to learn much from it unless there was a skilled instructor sitting next to them throughout the whole experience. First of all, the students did not know how to design an experiment by selecting a dependent variable of interest and then varying an independent variable in order to observe the relationship. Second, the students did not know enough to generate relevant hypotheses, which is where intelligent experimental design must begin. Third, they did not know enough about the cardiovascular system to know what ranges of values for variables were physiologically possible or interesting. Finally, they were so confused by the problems of collecting numerical data that they almost never managed to extract the important relationships that they were supposed to discover from these experiments. They wound up talking about numbers instead of the meaning of those numbers. The enormous variety of variables and values available in MacMan compounded these problems.

After much discussion, Michael and Rovick decided that what their students needed most was to learn the qualitative relationships between the important parameters and how to use qualitative causal reasoning to make qualitative predictions about the behavior of those parameters (whether they would increase, decrease, or remain the same). Rovick and Brenner (1983) wrote a PLATO program named Heartsim to embody this new focus on making qualitative predictions. Heartsim was made up of two components. One component was the MacMan model, but the other was a whole new teaching module. This didactic component defined a set of experiments for the students to think about. It evaluated their answers and focused on helping students reason about how changes in one parameter would affect the others.

The students entered their predictions in a table that they (Rovick and Michael, 1992) named the "Prediction Table" (see Table 1). The table has a row for each of the seven most important cardiovascular parameters, ending with the Mean Arterial Pressure or MAP, which is the regulated variable. The columns represent the three periods in the body's response to the perturbation. The Direct Response or DR phase (the first column) occurs right after the perturbation, before the baroreceptor reflex begins; the Reflex Response or RR phase in the second column records how these parameters change in response to the reflex; the third phase, the Steady State or SS phase, is recorded in the third column. The student entered an up-arrow if the variable increased, a down-arrow if it decreased, and a zero to represent no change. Once the student has been given several chances to make corrections, the system fills in the correct predictions along side those given by the student.

Table 1. The CIRCSIM Prediction Table (inherited from HEARTSIM).

Parameters	DR	RR	SS
Cardiac Contractility (CC)	↑	↓	↓
Right Atrial Pressure (RAP)	↓	↑	↓
Stroke Volume (SV)	↑	↓	↑
Heart Rate (HR)	0	↓	↓
Cardiac Output (CO)	↑	↓	↑
Arterial Resistance (RA)	↓	0	↓
Mean Arterial Pressure (MAP)	↑	↓	0

Heartsim turned out to be a success. Rovick and Michael were amazed to find that students came into the computer laboratory to use it for many hours outside of the scheduled laboratory. The fact that chronically overworked medical students went out of their way to use the system not once, but many times, suggested that they found that it really helped them learn important material.

Qualitative Reasoning and CIRCSIM

It was difficult to demonstrate Heartsim to colleagues since few medical schools used PLATO, so in the mid 1980's, Joel Michael decided to write a new program to do the job that Heartsim did and that could run on the small and relatively cheap computers that were becoming available at that point. This was an especially fortunate decision since Control Data Corporation, which supplied the computer support for the PLATO system, began to run into trouble in the late 1980's and the system rapidly became much less available. (For a brief history of Don Bitzer, the electrical engineer who built the system and of the PLATO System itself look for PLATO on the Web.)

As Joel Michael contemplated building the new system, the task of rewriting the mathematical model underlying MacMan to run on a small computer with the current restrictions on speed and memory was daunting. He and Allen Rovick discussed what the new program should do and just what parts of the model they needed, they came to a fundamental insight. The understanding of physiological systems that they expected of their students required the ability to reason about qualitative changes to the system (did a variable increase, decrease, or stay the same) and NOT quantitative ones. Their students did not need to know the numerical values of the variables that MacMan computed or the size of the change. They needed to know whether the variable changed or not and the direction of the change if a change occurred. A major goal of the physiology course is teaching physicians to reason about these problems. But much of the reasoning that physicians must carry out in a medical practice is qualitative not quantitative (Escaffre, 1985). The students still needed a set of well-defined problems and a way to record their predictions about these qualitative changes. Michael and Rovick also retained the Prediction Table because it was not only a convenient way to collect qualitative information, but it seems to help the student understand the problem.

In summary, Michael and Rovick decided that what the students needed to understand and learn to predict was these qualitative relationships between the important cardiovascular variables and how qualitative changes in one variable would affect other variables. What the student needed to remember was not the size of the change, but the direction. What they needed most from this course was to learn how to solve problems using qualitative causal reasoning.

So the new program, called CIRCSIM (Rovick and Michael, 1986) contains only a Prediction Table, the correct predictions, and a lot of small text files containing instructive remedial text about common error patterns. (An error pattern is a recognizable constellation of errors representing a frequent source of confusion.) CIRCSIM was written in Basic, since all computers built at the time came with a Basic interpreter or compiler.

Again students started using this program outside of regular assigned laboratory periods. It also became popular among second-year students reviewing for national boards. Rovick and Michael finally carried out a formal evaluation experiment (1992). The results demonstrated that the use of CIRCSIM produced significant learning gains. Furthermore, it appeared that students did significantly better using the program in pairs rather than all by themselves and better still when their instructors circulated through the laboratory asking questions here and there and doing some informal tutoring.

CIRCSIM-Tutor

As Rovick and Michael looked for new ways to improve their system, they focused on the fact that students did better when they used CIRCSIM in pairs and their own conviction that students really understood the material when they could give a verbal explanation of what was going on and decided to build an intelligent tutoring system that could carry on a natural language dialogue with the student. Chi's studies on the importance of self explanation in tutoring (1994) reinforced their belief that natural language interaction was fundamental to good tutoring. They joined up with Martha Evens at Illinois Institute of Technology to build the CIRCSIM-Tutor system (Evens et al., 1993, 2001; Michael et al., 2003). This system uses the Prediction Table from Heartsim and CIRCSIM but it also builds a student model and uses a domain knowledge base with a collection of frames to support its dialogues with the student. Most of the language generated by the system consists of questions and hints, based on studies of human tutors (Fox, 1993; Glass et al., 1999; Graesser and Person, 1994; Kim et al., 2000). This output is synthesized from tutoring strategies represented as schemas or simple plans for tutoring a particular concept. Then a text generator based on the Lexical Functional Grammar approach outputs actual sentences. Student input tends to be composed of short phrases, which are handled well by Glass's (2001) parser made up of cascaded finite state machines. The parser uses case frames (Lee et al., 2002) to translate the input into simple logic forms based on the work of Kieras (1985). The knowledge in the frames is represented using the same simple first-order calculus.

Although the system can sustain a dialogue successfully and the students like it and learn from it and demand copies to take home and run on their own PC's, its repertoire of tutoring strategies is still limited. What is more it does not really fulfill the original dream of getting students to make long explanations of their own. With the hope of stimulating longer responses from students, we set up the system to ask a number of open questions in November, 2002. We did not try to parse the answers, but we hoped that the students would take some time to realize that since we parse all the rest of the student input – or that they would answer anyway because they found the questions of interest. Many of the students blew off these questions, but most people answered some of them before they realized that the system was not analyzing this part of the input; a few answered all of these questions. To our delight, the students produced a number of complete sentences in response and even when they did not type in a complete sentence, they produced much longer and more informative than usual. For example, some of the answers received in response to the question "Why did you predict that Inotropic State would remain unchanged?" were:

M80: sym. and parasym. didn't change it yet
M85: THERE HAS NOT BEEN A BARORECEPTOR REFLEX YET
T51: Remembering the concept map, IS is affected by the baroreceptor reflex.
T61: IS is a reflex response
M55: because IS is separate from s
M58: because it's a direct response and Changing resistance wouldn't affect contractility of the heart just yet
T48: Because only the baroreceptor firing rate directly affects IS.
T49: Because it is a reflex response
T62: Because that is part of the baroreceptor reflex.
T76: because it is controlled by the baroreceptor
T86: because baroreceptor reflex has not been activated yert

The form of the question still affects the range of answers. This question and other "why?" questions received a lot of "because …" answers. The question "What does the baroreceptor reflex do?" brought us fewer complete sentences but a whole lot of verb phrases, as shown below:

T60: It tries to control MAP.
T76: it acts to return MAP to the original value
M48: Try and maintain MAP
M60: respond to changes in mean arterial pressure.
M64: regulates MAP
M82: tries to change the direction of the direct response, tries to get variables back to normal
T87: tries to keep MAP near a constant value
M58: Negative feedback system, which tries to correct the problems occurring in the cardiovascular system

We were delighted to see that students could and would type in such informative and insightful responses. Now we have to figure out how to cope with a large range of mistakes in grammar and spelling and develop strategies for parsing and responding to these inputs.

The Why-2 Atlas system developed by Kurt VanLehn and his team at the University of Pittsburgh (2002) asks students to enter essay answers to questions in qualitative physics. Rosé's parser (Rosé et al., 2002) analyzes the input and then Tacitus-Lite from Stanford is used to represent the student's argument and analyze it further. Finally Jordan's discourse generator synthesizes a response (VanLehn et al., 2002). We obviously need to choose a more powerful representation system ourselves in order to handle the responses to open questions. We also want to develop our capacity to respond to student initiatives (Shah et al., 2002). Most of all we want to add analogies to our repertoire of tutoring strategies. The question is what reasoning capabilities we need to employ.

Bredeweg & Forbus (2003) provide an insightful introduction to qualitative reasoning in ITS. Both Bredeweg (Bouwer & Bredeweg, 2001) and Forbus (Forbus et al., 2001) have constructed ITS with qualitative models that are on some ways more sophisticated than ours. How much power do we need to support analogies?

Supporting Analogies Using Qualitative Reasoning

A corpus of eighty-one hour long expert human tutoring sessions tutoring first year medical students at Rush Medical College on the baroreceptor reflex were analyzed and marked for analogies by hand, using SGML an annotation language that was used to mark up the sessions. The letters F and K are identifiers indicating whether the session was carried out face-to-face or keyboard-to-keyboard; the session number follows; st (student) and tu (tutor) indicates who is typing; following is the turn number; last, is the sentence number within the turn. Transcripts will be provided on request. The examples listed below demonstrate the expert tutor's use of analogy to facilitate understanding while using qualitative causal reasoning to make qualitative predictions about the consequences of perturbations to the system. The analogies found were of two types: those involving reflection on students' earlier work—another neural variable—and those that employed familiar objects outside the physiological domain—balloons and Ohm's Law. Below are examples of each type.

Example 1 Keyboard-to-keyboard session number one (K1) demonstrates the use of analogy to tutor the qualitative changes in neural variables in the DR period. After discussing how a neurally controlled variable (TPR) behaves in the DR period, the tutor requests that the student come up with another neurally controlled variable. He then invites the student to infer from the analogy that the other variable will behave—qualitatively change—in the same manner during this time period.

K1-tu-16-2: Now, what variable do you think will be affected first and in what direct ion?
K1-st-17-1: Co i
K1-tu-18-1: Excellent.
K1-tu-18-2: What variable is affected next and how?
K1-st-19-1: Map i
K1-tu-20-1: Super.
K1-tu-20-2: Couldn't do better myself.
K1-tu-20-3: What's next?
K1-st-21-1: Rap i
K1-tu-22-1: What causes the RAP to rise?
K1-st-23-1: Increased amount of blood coming into the ra
K1-tu-0-1: Finished?
K1-st-23-2: Remember to finish each entry with an
K1-tu-24-1: Are you finished?
K1-st-25-1: Y
K1-tu-26-1: Well, if one of the earliest things to happen is that the CO goesup, the heart will be removing increased quantities of blood fromthe central venous compartment.
K1-tu-26-2: Why would there be an increasedamount of blood coming into it?
K1-st-27-1: I guess there would not be increased RAP, but would the TPRincrease because of the increased CO?
K1-tu-28-1: Were dealing now with the DR period.
K1-tu-28-2: That's before there are anyneural, reflex changes.
K1-tu-28-3: The TPR is neurally controlled.
K1-tu-28-4: So whatdo you think would happen to it?
K1-st-29-1: During the DR period it would remain unchanged.
K1-tu-30-1: Correct.
K1-tu-30-2: What other variables are neurally controlled?
K1-st-31-1: CC, HR
K1-tu-32-1: Again correct.
K1-tu-32-2: Now we know that in this guy HR is under thecontrol of the artificial pacemaker.
K1-tu-32-3: But what would happen toCC?
K1-st-33-1: CC 0
K1-tu-34-1: Right on!

Example 2 Face-to-face session number one (F1) demonstrates the tutoring of the qualitative changes in stroke volume when right atrial pressure goes up. The human tutor compares the right atrium to a balloon, which behaves similarly to the right atrium when filled (as discussed in Lulis & Evens, 2003; Lulis, Evens, & Michael, 2003)

Structure for the balloon

- fill a balloon with water
- it will distend
- the pressure in the balloon increases as it distends

Structure for the heart

- fill the right atrium
- the right atrium will distend
- pressure will increase as it distends

F1-tu-49-6: Our situation is, we have instantly changed the cardiac output by changing the heart rate.
F1-tu-49-7: OK?
F1-st-50-1: Right.
F1-tu-51-1: But we haven't changed the venous return.
F1-tu-51-2: So the rate at which we are removing blood from the venous system is now slower then the rate at which blood is returning to the venous system.
F1-st-52-1: Oh, OK.
F1-tu-53-1: What will the consequence be?
F1-st-54-1: So the consequence there should be that the stroke volume would go down.
F1-tu-55-1: Why?
F1-st-56-1: Because the venous return would, ah, would...
F1-tu-57-1: Could we get to something that's more direct?
F1-tu-57-2: Let's just consider the venous system which is being emptied now more slowly, but being filled at the same rate.
F1-tu-57-3: What will happen to the volume in the venous system?
F1-st-58-1: Emptied more slowly but filling at the same rate, so it is going to be increasing in something like right atrial pressure?
F1-tu-59-1: So, can we try that again?
F1-tu-59-2: (both laughing)
F1-st-60-1: Let's see, Prof?
F1-tu-61-1: Yes?
F1-st-62-1: If I make an analogy of you try to fill a sink with water and you...
F1-tu-63-1: Try to fill a balloon with water, since that's what we're dealing with, a distensible object.
F1-st-64-1: OK.
F1-st-64-2: We're filling a balloon with water, and the rate at which the water empties out is slower than before, but you're filling it at the same rate, so it's going to be easier to distend the balloon.
F1-tu-65-1: So the balloon will distend.
F1-st-66-1: So the balloon will distend and we will have -- oh, OK, according to Starling's...
F1-tu-67-1: Well, let's stay with the balloon first.
F1-tu-67-2: What's going to happen when the balloon distends?
F1-st-68-1: In terms of the volume or the pressure?
F1-tu-69-1: In terms of the pressure.
F1-st-70-1: Pressure in the balloon will increase.
F1-tu-71-1: OK, good.
F1-tu-71-2: What is this balloon we're talking about?
F1-st-72-1: This balloon is the right atrium.
F1-tu-73-1: So can we say something about what happens to right atrial pressure when cardiac pressure goes down?
F1-st-74-1: Right atrial pressure goes up when cardiac output goes down.
F1-tu-75-1: They are inversely related, right?
F1-st-76-1: Right, OK.
F1-tu-77-1: When the right atrial pressure goes up...?
F1-st-78-1: When the right atrial pressure goes up --
F1-tu-79-1: Then?
F1-st-80-1: In terms of stroke volume?
F1-tu-81-1: Yeah.
F1-st-82-1: When the right atrial pressure goes up, the right ventricle will be filled more rapidly, to be more,would increase your stroke volume.
F1-tu-83-1: Absolutely.

Example 3: In face-to-face session number five (F5), Ohm's Law (voltage = current * electrical resistance) is used as a base to facilitated understanding of MAP = CO * TRP (the pressure = flow * resistance).

F5-st-343-5: Like MAP = CO x TPR.
F5-tu-344-1: Of course!
F5-st-345-1: I'll think about that...
F5-st-345-2: I just did five problems on this.
F5-st-345-3: So...
F5-st-345-4: I don't know why this is true though.
F5-tu-346-1: You don't know why that is true.
F5-tu-346-2: Do you recognize that that equation is identical in principle to Ohm's Law which says that electrical potential is equal to the current times resistance.
F5-st-347-1: V = IR.
F5-tu-348-1: Or V = IR.
F5-tu-348-2: Right.
F5-tu-348-3: The push equals the flow times the resistance.
F5-st-349-1: Well, OK, Yeah, this one makes ense to me and I can see that this is the same thing.

Results of the Use of Analogies to Facilitate Understanding

An examination of the corpus yielded fifty-one analogies proposed by the tutors in the eighty-one hour long sessions they conducted. Out of the fifty-one, nine were used by the tutor solely to enhance the student's understanding of the material and no predictions were requested. No inference was requested by the tutor in five of the cases, but, out of these five, four cases resulted in correct qualitative predictions resulting from correct mappings between the base and the target. Inferences were requested by the tutor thirty-seven times resulting in fifteen successful mappings (correct inferences) and twenty-two failed mappings (incorrect inferences). Out of the twenty-two failed mappings, the tutor successfully repaired/explained the analogy resulting in correct inferences by the student fifteen times. The tutor abandoned the analogy in favor of a different teaching plan only seven times.

Implementation of Analogies

After examining the examples of human of analogy found in our corpus, Michael selected the following bases to implement: another neural variable, another procedure, Ohm's law, balloons or compliant structures, and the accelerator/brake analogy. We are exploring the possible use of bases within the domain—another neural variable and another procedure—to test Gentner's mutual alignment theory of analogies (Kurtz et al., 2001) in our system. Bases outside of the domain—Ohm's Law, balloons, and the brake and accelerator—are much more interesting, but also leave room for greater misunderstanding and possible future misapplication of the analogy.

A Model for Implementing Analogies Outside the Domain

For analogies whose bases fall within the domain, the rule-based implementation that CIRCSIM-Tutor currently employs will be used. After surveying different models, we selected the Structure Mapping Engine (SME) (Gentner, 1998; Gentner & Markman, 1997; Forbus et al., 1995) for implementing analogies whose bases lie outside the domain. Our expert tutors' behavior seems to resemble SME closely, making it an excellent model.

Conclusion

In this paper we have tried to describe how Michael and Rovick came to believe that experience in qualitative reasoning rather than in quantitative manipulation of a simulation model would serve their students' needs more effectively and the systems that they have written to provide this experience to their students; we have also described our current efforts to support analogies in our ITS using qualitative reasoning.

Michael and Rovick (Michael et al., 1990) emphasize qualitative causal reasoning in the teaching of physiological concepts to medical students who need to make predictions about the consequences of perturbations to the system. They have employed qualitative reasoning in CIRCSIM, a conventional CAI version of this approach, and CIRCSIM-Tutor, which embodies this same approach but does so by generating a Socratic dialogue with the student.

The use of analogies will facilitate the understanding of more complex qualitative reasoning and provide support for the reasoning process. We want to make our system understand and interpret student initiatives and student responses to open questions. We also want to implement analogies effectively and explain the underlying structural mapping to students who get confused.

Acknowledgments

This work was partially supported by the Cognitive Science Program, Office of Naval Research under Grant 00014-00-1-0660 to Stanford University as well as Grants No. N00014-94-1-0338 and N00014-02-1-0442 to Illinois Institute of Technology. The content does not reflect the position or policy of the government and no official endorsement should be inferred.

References

Bouwer, A., & Bredeweg, B. (2001). "VisiGarp: Graphical representation of qualitative simulation models." In J.D. Moore, C.L. Redfield, & W.L. Johnson (Eds.), *Proceeding of Artificial Intelligence in Education.* Amsterdam: IOS Press. 294-305.

Bredeweg, B. & Forbus, K. (2003). "Qualitative modeling in education." *AI Magazine*, 24 (4): 35-46.

Chi, M., de Leeuw, N., Chiu, M.H. & LaVancher, C. (1994). "Eliciting self-explanations improves understanding." Cognitive Science, 18, 439-477.

Dickinson, C. J., Goldsmith, C. H. & Sackett, D. L. (1973). "MACMAN: A digital computer model for teaching some basic principles of hemodynamics." *Journal of Clinical Computing,* 2: 42-50.

Escaffre, D. (1985). "Qualitative reasoning on physiological systems: The example of blood pressure regulation." In I. DeLotto and M. Stefanelli (eds.) *Artificial Intelligence in Medicine,* Elsevier Science Publishers. 51-62.

Evens, M., Spitkovsky, J., Boyle, P., Michael, J.A., & Rovick, AA. (1993). "Synthesizing tutorial dialogues." *Proceedings of CogSci '93.* Boulder, Colorado, June. 137-142.

Evens, M., Brandle, S., Chang, R.C., Freedman, R., Glass, M., Lee, Y.H., Shim, L.S., Woo, C.W., Zhang, Y., Zhou, Y., Michael, J.A., & Rovick, A.A. (2001). "CIRCSIM-Tutor: An intelligent tutoring system using natural language dialogue." In F. Moore (ed.*), Proceedings of the Midwest Artificial Intelligence and Cognitive Science Conference.* 16-23.

Forbus, K. D., Gentner, D., & Law, K. (1995). "MAC/FAC: A model of similarity-based retrieval." *Cognitive Science, 19*(2), 141-205.

Forbus, K.D., Carney, K. Harris, R., and Sherin, B.L. (2001). "A qualitative modeling environment for middle-school students." Paper presented at the Fifteenth International Workshop on Qualitative Reasoning., May 15-17, San Antonio.

Fox, B. (1993). *The human tutorial dialogue project.* Hillsdale, NJ: Erlbaum.

Gentner, D. (1998). "Analogy." In W. Bechtel and G. Graham (Eds.), *A companion to cognitive science,* Oxford: Blackwell. 107-113.

Gentner, D., & Markman, A. B. 1997. "Structure mapping in analogy and similarity." *American Psychologist, 52*(1): 45-56.

Glass, M.S. (2001). "Processing language input for an intelligent tutoring system." In J.D. Moore, C.L. Redfield,

& W.L. Johnson (Eds.), *Proceedings of Artificial Intelligence in Education.* Amsterdam: IOS Press. 210-221.

Glass, M., Kim, J.H., Evens, M., Michael, J.A., & Rovick, A.A. (1999). "Novice vs. expert tutors: A comparison of style." *Proceedings of MAICS 99,* Bloomington, IN, April 24. 43-49.

Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. *American Educational Research Journal*, **31**, 104-137.

Hody, G. L., & Avner, R. A. (1978). "The PLATO system: an evaluative description." In E. C. DeLand (editor), *Information Technology in Health Science Education*. New York: Plenum Press. pp. 143-177.

Kieras, D. E. (1985). "Thematic processes in the comprehension of technical prose." In Britton, B.K. & Black, J.B. (eds.). *Understanding expository text.* Hillsdale, NJ: Erlbaum, 89-108.

Kim, J.H., Glass, M., Freedman, R., & Evens, M.(2000). "Learning the use of discourse markers in tutorial dialogue for an intelligent tutoring system." In L.R. Gleitman & A.K.Joshi (Eds.) P*roc. Cognitive Science 2000,* Mahwah, NJ. Lawrence Erlbaum. 262-267.

Kurtz, K., Miao, C., & Gentner, D. (2001). "Learning by analogical bootstrapping." *Journal of the Learning Sciences,* 10(4):417-446

Lee, C.H., Seu, J.H., & Evens, M.W. (2002). "Automating the construction of case frames for CIRCSIM-Tutor." *ICAST 2002.* Chicago, IL. 59-65.

Lulis, E., & Evens, M. (2003). "The use of analogies in human tutoring dialogues." In R. Freedman and C. Callaway (Eds.) *Natural language generation in spoken and written dialogue. AAAI Symposium,* Stanford, March, 2003. 178-183.

Lulis, E., Evens, M., & Michael, J. (2003). "Representation of analogies found in human tutoring sessions." *Proceedings of the Second IASTED International Conference on Information and Knowledge Sharing,* 88-93. Anaheim, CA:ACTA Press.

Michael, J. A. (1993). "Teaching problem-solving in small groups." In H.I. Modell and J.A. Michael (eds.) *Promoting active learning in the life science classroom.* Annals of the New York Academy of Sciences. 701, 37-48.

Michael, J.A., Rovick, A.A., Evens, M.W., & Kim, N.H. 1990."A Smart Tutor Based on a Qualitative Causal Model." Proceedings of the AAAI Spring Symposium on Knowledge-Based Environments for Learning and Teaching, Stanford, March 27-29, 1990, 112-117.

Michael, J.A., Rovick, A.A., Glass, M.S., Zhou, Y., & Evens, M. (2003). "Learning from a computer tutor with natural language capa-bilities." *Interactive Learning Environments, 11(3),* 233-262. November 2003.

Rosé, C.P., Bhembe, D., Roque, A., Siler, S., Shri-vastava, R., & VanLehn, K. (2002). "A hybrid natural language understanding approach for robust selection of tutoring goals." In S.A. Cerri, G. Gouardères, & F. Paraguaçu (eds.) *ITS 2002,* LNCS 2363. Berlin: Springer-Verlag. 552-561.

Rovick, A. A., & Brenner, L. (1983). HEARTSIM: A cardiovascular simulation with didactic feedback. *The Physiologist* **26**: 236-239.

Rovick, Allen A. & Joel A. Michael. (1986). "CIRCSIM: An IBM PC computer teaching exercise on blood pressure regulation." Paper presented at the XXX IUPS Congress, Vancouver, Canada.

Rovick, A. A., & Michael, J. A. (1992). "The prediction table: A tool for assessing students' knowledge." *American Journal of Physiology,* 263(6), Pt. 3, S33-S36. Also available as *Advances in Physiology Education,* 8(1), S33-S36.

Shah, F., Evens, M., Michael, J.A., & Rovick A.A. (2002). "Classifying student initiatives and tutor responses in human keyboard-to-keyboard tutoring sessions." *Discourse Processes,* 33(1) 23-52.

VanLehn, K., Jordan, P.W., Rosé, C.P., Bhembe, D., Bottner, M., Gaydos, A., Makatchev, M., Pappuswamy, U., Ringenberg, M., Roque, A., Siler, S., & Shrivastava, R. (2002). "The architecture of Why2-Atlas: A coach for qualitative physics essay writing." In S.A. Cerri, G. Gouardères, & F. Paraguaçu (Eds.) *ITS 2002,* LNCS 2363. Berlin: Springer-Verlag. 158-167.