STUDENT MODELING FOR OPEN LEARNING ADAPTIVE HYPERMEDIA

Cristina Conati, Andrea Bunt,
University of British Columbia
2366 Main Mall, Vancouver, BC, V6T1Z4
Canada

Abstract
In this paper we discuss the importance of incorporating into adaptive hypermedia a student model that can assess the effectiveness of a learner's exploratory behavior. We describe the details of the student model in the Adaptive Coach for Exploration (ACE), an open learning hypermedia for mathematical functions. Based on our experience of having evaluated the model with human subjects, we describe the factors such a model should incorporate to effectively provide adaptive guidance to learning through exploration for those students who cannot explore adequately on their own. We discuss how, in addition to domain dependent elements (e.g., coverage of relevant domain concepts and prior knowledge), these factors include domain independent meta-cognitive skills such as self-explanation and self-monitoring.

Key Words
Open learning hypermedia, student modeling, meta-cognitive skills.

1. Introduction

In this work, we focus on issues related to student modeling for a specific type of adaptive hypermedia – open learning hypermedia. Open learning environments place less emphasis on supporting learning through structured, explicit instruction and more on providing the learner with the opportunity to freely explore the available instructional material, acquiring knowledge of relevant concepts and skills in the process [1][1]. In theory, this type of active learning should enable students to acquire a deeper, more structured understanding of concepts in the domain [1]. Also, owing to the unguided nature of the interaction, the hope is that, in addition to skills in the target instructional domain, the learner can practice and acquire meta-cognitive skills associated with effective exploration [2]. The vast amount of exploratory material provided by the Web makes the techniques employed by these types of educational environments particularly relevant.

Empirical evaluations, however, have shown that a student's ability to benefit from interacting with open learning environments depends on a number of student-specific features, including activity level, whether or not the student already possesses the meta-cognitive skills necessary to learn from exploration and general academic achievement (e.g., [1], [2], [3]). Students who are inactive or lack the necessary cognitive skills often fail to initiate enough meaningful exploratory actions and they can have difficulty interpreting and generalizing the results of the actions that they do initiate. Furthermore, the larger the exploration space, the higher the danger that these students could experience information overload and get lost in the exploration process.

Given the above findings, we argue that any highly exploratory adaptive hypermedia should be instrumented with a student model that can assess the effectiveness of a learner's exploratory behaviour and detect when the learner needs supports for the exploration process. The student model should be the basis for pedagogical strategies designed to provide this support in a tailored and timely manner, without taking away the sense of freedom and control that is a key asset of open learning environments. In this paper, we discuss issues related to the specification of this model and related pedagogical strategies. The discussion is based on our experience in building ACE (Adaptive Coach for Exploration), an intelligent exploratory learning environment for mathematical functions [4][5]. We first introduce ACE and its student model for exploratory behavior. We then illustrate some lessons that we have learned from empirical evaluations of ACE, as well as their implications for student modeling in highly exploratory hypermedia. We conclude with a discussion of related and future work.

2. The ACE Open Learning Environment

2.1 ACE Interface

ACE is an adaptive open learning hypermedia for the domain of mathematical functions. ACE's activities are divided into units and exercises. Units are collections of exercises whose material is presented with a common

[1] Another interpretation of "open" is that there is unconstrained access to instructional material and activities outside the environment, but this is not the intended meaning in the paper.

theme and mode of interaction. Exercises within the units differ in function type and equation.

Figure 1 and 2 show the main interaction window for two of ACE's units: the *Machine Unit* and the *Plot Unit*. ACE also has a third unit, the *Arrow Unit*, as well as help pages and a feedback panel, not displayed for lack of space.

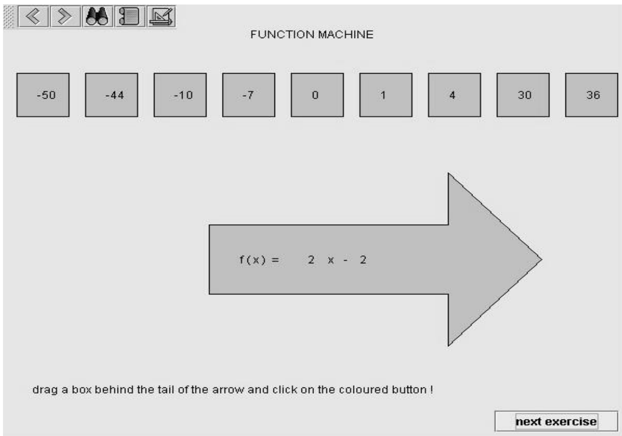

Figure 1: ACE Machine Unit

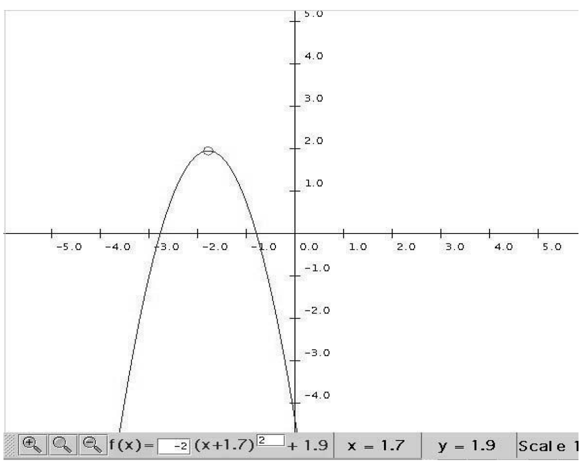

Figure 2: ACE Plot Unit

The Machine Unit and the Arrow Unit allow the student to explore the relationship between an input and the output of a given function. In the Machine Unit, the exploration consists of dragging any number of inputs displayed at the top of the screen to the tail of the function ``machine" (the large arrow shown in Figure 1), which then computes the corresponding output. The Arrow Unit allows the student to match a number of inputs with the correct outputs and is the only activity within ACE that has a clear definition of correct and incorrect behaviour. In the Plot Unit (Figure 2), the student can explore the relationship between the graph of a function and its equation, by manipulating either the graph or the equation, and then observing the corresponding changes in the other entity.

Although the above units and corresponding exercises can be traversed by the student in a predefined sequence, we want students to also be able to freely explore the curriculum. Therefore, ACE includes a Lesson Browser which shows all units and exercises, and allows the student to go to any exercise by clicking on it. ACE also provides the Exploration Summary (see Figure 3), a tool that helps students organize their exploration process by summarizing the exploration actions they have performed so far within a given exercise, in terms of relevant exploration categories (described in the next section).

In addition to these tools to facilitate student exploration, ACE also include a coaching component that provides tailored hints when students have difficulties exploring effectively. The coach's interventions are based on a model of student exploration that monitors the student's interaction with ACE and assesses the effectiveness of the resulting exploration. We first describe this model and then discuss how the ACE's coaching component uses it.

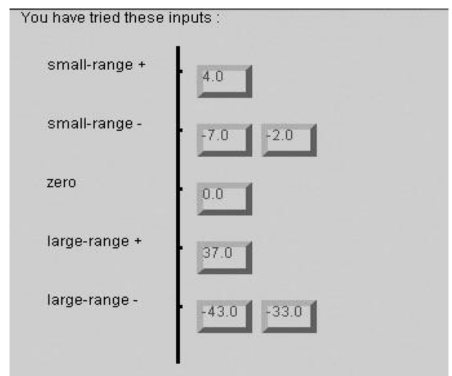

Figure 3: ACE Exploration Summary Tool

2.2 ACE Student Model

Modelling students' exploration presents unique challenges, for two main reasons. First, in more structured educational activities, such as problem solving and question answering, there is usually a definition of correct behaviour, which allows this behaviour to be represented and recognized in a formal model. In contrast, in open learning environments there is no clear understanding of what constitutes successful exploration in general. Second, it is hard to obtain reliable information on the student's exploratory behaviour. The amount and quality of information available to a user model to perform its assessment is referred to as the bandwidth issue [6]. The less explicit information on the user's relevant traits or behaviours the model is able to obtain, the higher the uncertainty in the modelling process. The bandwidth problem is especially difficult for student modelling in open learning environments. Both exploratory behaviour and related meta-cognitive skills necessary for effective learning are not easily observable unless the environment's interface forces students to make them explicit. However, forcing students to articulate their exploration steps clashes with the unrestricted nature of open learning environments. Thus, a model for exploratory behaviour is bound to deal with low-bandwidth information, which introduces a high level of uncertainty into the modelling task.

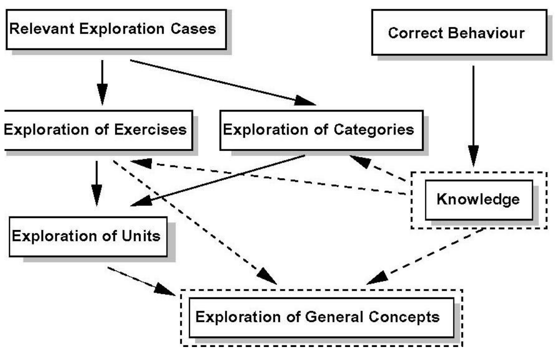

Figure 4: High-Level Description of ACE Student Model

ACE's student model uses Bayesian networks to manage this uncertainty. One of the main challenges in using Bayesian Networks is to define the network's structure to accurately represent the relevant probabilistic dependencies. In our model, this problem is exacerbated by the difficulty of defining correct exploratory behaviour. Our approach was to use an iterative design process. We first built a version of the Student Model using our intuition of what constitutes effective exploration in ACE and evaluated the model in a formal study. We then used the study results to redesign the model and conducted an evaluation of the changes. By building and evaluating two versions of the model, we were able to gain valuable insight into (i) what factors contribute to effective exploration of the ACE environment and (ii) how to formalize these factors in the Student Model (details on this iterative process can be found in [5]).

Figure 4 shows a high-level description of the different types of nodes in the current version of the student model. As the figure shows, the model includes several types of exploration nodes, to assess exploratory behaviour at different levels of granularity.

- *Relevant Exploration Cases*: the exploration of individual exploration cases in an exercise (e.g., dragging the number 3, a small positive input, to the back of the function machine in the Machine Unit).
- *Exploration of Exercises*: the exploration of individual.
- *Exploration of Units*: the exploration of groups of related exercises (e.g., all of the exercises in the Plot Unit).
- *Exploration of Categories*: the exploration of groups of relevant exploration cases that appear across multiple exercises (e.g., all of the exploration cases involving small positive inputs or large negative inputs in the Machine and Arrow unit; all the cases involving a linear function with a positive slope or a negative intercept in the Plot unit).
- *Exploration of General Concepts*: the exploration of general domain concepts, such as the input/output relation for different types of functions.

In ACE's student model, the links among these types of exploration nodes represent how they interact to define effective exploration. Individual operations that a student performs within an exercise (*Relevant Exploration Cases*) define effective exploration of each individual exercise (*Exploration of Exercises*), which in turns influences effective *Exploration of Units* and exercises involving similar cases. How well the student explores the available exercises and units then influences how well she has explored general domain concepts (*Exploration of General Concepts*). Exploration of individual cases also defines how well the student is exploring groups of relevant cases that appear across exercises (*Exploration of Cases*). Each exploration node can take on the value of either True of False. A True value means that the learner has sufficiently explored the item associated with the node (i.e., the relevant exploration case, exercise, unit, category, or general concept), where sufficient means that the learner has performed enough interface actions to indicate thorough exploration.

Initial studies with ACE showed that how much exploration a student needs to understand a concept depends on how much knowledge of that concept she already has. Thus, ACE's student model also includes binary nodes representing the probability that the learner understands the relevant pieces of knowledge. (summarized by the node *Knowledge* in Figure 4) and links indicating the aforementioned relation between knowledge and exploration. Knowledge nodes are updated only through actions for which there is a clear definition of correctness (e.g. linking inputs and outputs in the Arrow Unit) and directly contribute to the assessment of how much further understanding the learner can gain from exploratory activities.

2.3 ACE adaptive support to effective exploration

Assessment of exploration at different levels of granularity allows the coaching component to provide a wide range of tailored feedback. In order to remain consistent with the philosophy of exploratory learning environments, it is crucial that ACE's support for student exploration be as unobtrusive as possible. Thus, this feedback is designed to provide different levels of guidance, according to the needs of the individual learner.

The first level of guidance consists of a generic suggestion to continue exploring when a student tries to leave an exercise before the student model assesses that it has been adequately explored. Currently, ACE does not interrupt a student's exploration of an exercise. Once the students signal that they wish to move to a new exercise, ACE queries the Student Model for two pieces of information: the probability that the student has

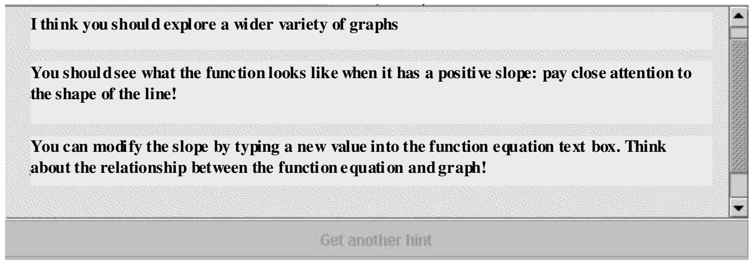

Figure 6: Sample of ACE incremental hints

adequately explored the current exercise and the probabilities for the relevant exploration categories. ACE remains silent if either the current exercise exploration is satisfactory (i.e. the related probability is above a predefined threshold currently based on our subjective estimate and equal for every exercise), or if the probabilities of the associated exploration categories are satisfactory. If neither of these criteria is met, then a message is shown, suggesting that the student explore more and ask for hints if she needs help. This first hint omits any specifics on *how* to explore, to force the students to be as self-directed as possible in the exploration process and to take initiative in obtaining hints. Since we want to maintain a high level of learner control, the student may always choose to disregard ACE's suggestion and leave the exercise at any point. If a student does decide to stay, ACE may suggest to open the Exploration Summary to recapitulate her recent exploratory actions. Figure 3 shows the tool open for an exercise in the Machine unit; it has organized the various inputs that the student has explored so far into relevant categories represented in the Student Model (such as Small-Positive-Range inputs, Zero inputs, etc).

As students explore, they can ask for a hint at any time. ACE generates hints on which element of an exercise to explore next by traversing the relevant exploration categories that are stored in the Student Model and by returning the first category that has a low probability of having been explored. Hints are provided at an incremental level of detail, to stimulate the students to do as much work as possible on their own. Figure 5 shows an example hint sequence generate for a student who is working with a linear function in the plot unit, has explored both positive and negative intercepts extensively and then has requested a hint.

Initial studies on ACE have generated encouraging evidence that the system, in its current form, can indeed help students learn from exploratory learning environments (see [4] and [5]). In particular, the studies have shown a high and statistically significant correlation between student learning and exploration related interface events triggered by ACE's tailored hints. The lack of correlation between different event types provides further evidence that learning was triggered by ACE's hints rather than alternative factors such as student general academic ability and conscientiousness. However, these studies have also uncovered a limitation that reduces ACE's potential. This limitation is that ACE's student model does not currently assess any meta-cognitive skills relevant to effective exploration. Thus, ACE sometimes misjudges' student behaviour and when it correctly interprets this behavior as suboptimal exploration, it cannot tell what caused it, as we will discuss in the next session.

4. Role of Meta-Cognitive Skills in Effective Exploration

In one of the ACE's studies [5], we analyzed the log files of the learners' interactions with ACE to see, among other things, how often ACE allowed a learner to leave an exercise without warning, even though the student had not learned the associated concepts (as indicated by the study post-test). We refer to this event count as the number of *premature passes*.

The data analysis showed that 9% of all student transitions between exercises could be classified as premature passes, indicating that ACE's model sometimes overestimates the learners' exploratory behaviour: when learners perform a large number of exploratory interface actions, the model assesses this to be good exploratory behaviour, even though some of the learners do not learn from these actions. We conjecture that a likely cause for this problem is that ACE only considers as evidence the interface actions that the student performs, without considering whether the student is *self-explaining* the outcome of these actions.

Self-explanation is defined as the meta-cognitive skill of spontaneously explaining to oneself available instructional material, in terms of the underlying domain knowledge. It has been shown to greatly influence learning [7]. To understand how self-explanation plays a key role in effective exploration, consider a learner who, in the ACE Plot Unit, moves a function graph around the screen, and never looks at how the movements change the function equation. Although this learner is performing many exploratory actions in that exercise, he can hardly learn from them because he is not reflecting on (self-explaining) their outcomes. We observed this exact behaviour in one of our subjects, who in fact showed in the post test that he had not learned the concepts underlying the actions he performed. Thus, we argue that an adaptive open learning environment must include ways to track student self-explanation. Factors that can be used to assess self-explanation behaviour include the amount of time spent on each exploration action [8][9], whether or not the learner is observing the results of exploration actions, whether the learner is actually generating explanations of these results, and the learner's known tendency to spontaneously self-explain. Consider, for instance, the case of the learner who is continually altering a graph in the Plot Unit. ACE's student model could be more confident that this is a sign of good exploration if the learner looks at both the graph and the equation as they change, and leaves the graph at the key positions for long enough to self-explain the correspondence with the current equation. Knowing a priori that the learner has a tendency to self-explain (many learners do not, as shown by the studies summarized in [7]) could further increase the model's confidence that the learner's behaviours are conducive to learning. Having ways to actually track what explanations the student is generating, if any, would provide even more solid evidence. However, detailed monitoring of the student focus of attention and explanations can interfere with the student exploration process. In [9], for instance, the authors describe an interface that allows monitoring student attention by forcing the student to explicitly uncover parts of the relevant instructional material, and tracks students' explanations by providing menu-based tools for explanation composition. These mechanisms were judged to be non-intrusive by the majority of the subjects in a study to evaluate the proposed interface, but the system described in [9] was not trying to promote learning through exploration. Subjects' tolerance of additional interface actions may be lower in an open learning environment. On the other hand, it is possible that switching to a more constrained interaction may help those students who do not have the meta-cognitive skills to explore effectively on their own. In particular, we conjecture that students who do not spontaneously self-explain, or who cannot effectively self-monitor their learning progresses may benefit from the stronger guidance provided by interface mechanisms that allow them to explicitly reflect on their interaction with the learning environment. To test this hypothesis, and to see how more detailed information on student self-explanation can improve ACE's effectiveness, we are working on giving ACE the capability to guide the generation of relevant self-explanations. We have started to design hints to stimulate relevant self-explanations in the different ACE's units. For instance, Figure 6 shows a dialogue box designed to help the student generalize the input-output behaviour of a given function, after the student has tried a few inputs in the Machine or in the Arrow unit.

The next step in providing ACE with the capability to track and stimulate student self-explanations will be to add to the student model the relevant self-explanation nodes. The high level relations between self-explanation information and exploratory behavior is shown if Figure 8. Adequate self-explanation of each exploration case can be assessed by leveraging any existing information on the time the student spent on this case, the self-explanation actions the student has performed using the available tools, and the student's pre-existing tendency to self-explain. This tendency could be tested before the student starts interacting with ACE by observing how she studies a piece of instructional text, and further evidence could be derived directly from the student's usage of the self-explanation tools available in the interface.

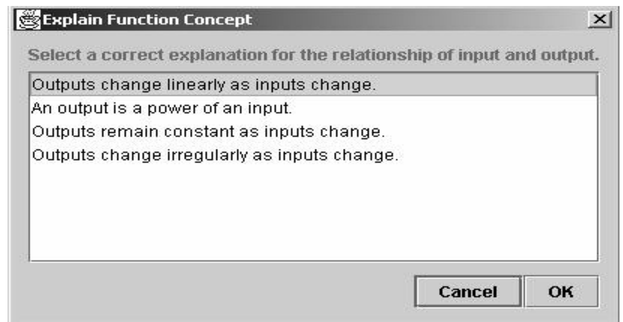

Figure 7: Self-explanation box for the Machine Unit.

The conditional probabilities for these relations in Figure 8 must encode the different reliability of each source of information, as well as how each source affects the reliability of the other two. For instance, a long time spent on a given exploration case for a student who is known to have a low tendency to self-explain will correspond to a lower probability of effective exploration than if the student is known to have a high tendency to self-explain.

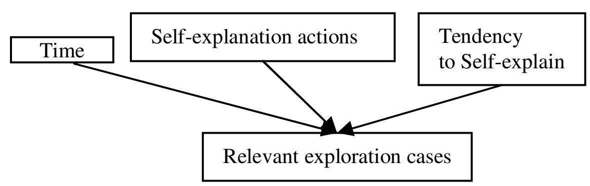

Figure 8: Addition of self-explanation variables to model exploration of relevant cases

5 Discussion and More Future Work

Because of the difficulty in monitoring student behavior in an open learning environments, there have been very few attempts so far to model student exploration. The open learning environments described in [1] and [10] monitor and support exploration subskills related to correct hypothesis selection, but rely on the assumption that students actively generate hypotheses and test them. Thus, they are unable to support students who are not active experimenters. The student model described in the ALE environment [11] does assess whether students touch all the relevant domain concepts during their experimentation, but it generates its assessment by explicitly testing the students. Thus, ALI departs from the open-ended and learner-directed style of interaction typical of pure open learning environments, which ACE's Student Model aims to preserve. We are not aware of any adaptive open learning environment that tries to tie effective exploration to self-explanation.

Although our efforts are currently concentrated on adding to ACE the capability to model and stimulate self-explanation, other meta-cognitive skills and students' traits could be added to the ACE's student model to improve its accuracy and to provide effective and unobtrusive support to exploration.

For instance, the capability of self-monitoring one's learning progresses seems to play an important role in effective exploration [1], since it allows the learner to judge if more exploration is needed, and of what concepts. The lack of this meta-cognitive skill could be one of the causes of students moving prematurely to a new exercise, or of overexploring. We did see several cases of student over-exploration with ACE. This was especially noticeable when students accepted ACE mistaken suggestions (based an early version of the student model) to stay in the current exercise when the study pretest showed that the student had already mastered all the necessary concepts. By being able to identify lack of self-monitoring capability as the cause of the above suboptimal behaviors, ACE could provide more appropriate support to improve student exploration. It could, for instance, make a more principled use of the Exploration Summary tool to help students monitor their progresses in the interaction.

Other factors that can influence the effectiveness of student exploration include motivation, personality traits, and learner's knowledge of effective exploration strategies. Incorporating these factors, along with the aforementioned factors affecting self-explanation and self-monitoring, is expected to improve both the model's assessment and its diagnostic capabilities, thus allowing the implementation of coaching strategies that can more precisely address the causes of poor exploration.

These coaching strategies must always take into account the trade-off between the added pedagogical effectiveness that an explicit hint may buy vs. the danger of interfering with the student engagement in the exploratory activities. Currently ACE does not take into account the student's attitude toward receiving pedagogical hints during the exploration process. Although ACE always allows a student to ignore its hints, the very fact that a hint was provided might be detrimental for some students with a specific personality or learning style. Thus, we are planning to add to the student model the capability to assess the impact of ACE's hints on student engagement, as well as decision theoretic mechanisms to compute at any given time what is ACE's intervention (or lack thereof) with the highest expected value, considering the preferences of the current student and the current state of the interaction.

References

[1] V.J Shute & R Glaser, A large-scale evaluation of an intelligent discovery world: Smithtown, *Interactive Learning Environments, 1*, 1990, 55-77.

[2] M. Njoo & T. de Jong, Exploratory learning with a computer simulation for control theory: learning processes and instructional support, *Journal of Research in Science Teaching, 30(8),* 1993, 821-844.

[3] W.R. van Joolingen & T. de Jong, Supporting hypothesis generation by learners exploring an interactive computer simulation, *Instructional Science, 20*, 1991, 389-404.

[4] A. Bunt, C. Conati, M. Huggett & K. Muldner, On improving the effectiveness of open learning environments through tailored support for exploration, *Proc. AIED 2001, 10th International Conf. on Artificial Intelligence in Education*, San Antonio, Texas, 2001, 365-376.

[5] A. Bunt & C. Conati, Probabilistic assessment of exploratory behaviour, *User Modeling and User-Adapted Interaction, 13(3)*, 2003, 269-309.

[6] K. VanLehn, Student modeling, *Foundations of Intelligent Tutoring System,* (Hillsdale, NJ: Erlbaum, 1988) 55-78.

[7] M.T.H. Chi, Self-explaining expository texts: the dual processes of generating inferences and repairing mental models, *Advances in Instructional Psychology*, 2000, 161-238.

[8] C. Conati & K. VanLehn, Further results from the evaluation of an intelligent tutor to coach self-explanation, *Proc. 5th International Conf. On Intelligent Tutoring Systems*, Montreal Canada, 2000, 304-313.

[9] C. Conati & K. VanLehn, Providing adaptive support to the understanding of instructional material, *Proc. of the International Conf. on Intelligent User Interfaces*, Santa Fe, NM, 2001, 41-47.

[10] Veermans, K. and W. van Joolingen, Using induction to generate feedback in simulation based discovery learning environments, *Proc. of Intelligent Tutoring Systems, 4th International Conference* (ITS '98), 1998. San Antonio, TX, 196-205.

[11] D'Souza, A., J. Rickel, B. Herreros, and W. Johnson, An automated lab instructor for simulated science experiments, *Proc. AIED 2001, 10th International Conf. on Artificial Intelligence in Education*, San Antonio, Texas, 2001, 65-75.