USING THE LEARNING-BY-TEACHING PARADIGM TO DESIGN

INTELLIGENT LEARNING ENVIRONMENTS

By

Krittaya Leelawong

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

In partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Computer Science

August, 2005

Nashville, Tennessee

Approved: Date:

Copyright © 2005 by Krittaya Leelawong
All Rights Reserved

For my grandparents,
Duangduean and Piroj Leelawong

In memory of my great uncle,
Dr. Laan Leelawong, M.D.
2004

ACKNOWLEDGEMENTS

I am deeply grateful for my family's support, especially my mother who encourages higher education. I am in dept to my grandparents, who wish to have a "doctor" among their grandchildren; it has been the source of patient and strength to finish my degree. Thanks a lot, Sis, for those long-length phone calls about all the problems in the world I have had. You always listen to me.

I would like to thank Dr. Gautam Biswas, Dr. Robert Bodenheimer, Dr. John Bransford, Dr. Douglas Fisher, Dr. David Noelle, and Dr. Nancy Vye for their guidance, helps and patience through the years. I would also like to thank all the current and previous members of the Teachable Agents Group at Vanderbilt and Stanford Universities, all members of the former Learning Technology Center, and the faculty members, staff and students in the department of Electrical Engineering and Computer Science. Everyone played a part in making Vanderbilt my second home. In addition, I would like to thank Cynthia Mayfield, Amanda Lynn Peltz, Matthew Lee, Dr. David Dilts, and Dean John Veillette for their help. This work would not be possible without the financial support from the National Science Foundation (Grant Numbers NSF KDI 9873520 and NFS ROLE 0231771).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

HPL	**H**ow **P**eople **L**earn: The framework for effective learning environments
ITS	An experimental group using a directed-learning environment described in Chapter 6
LBT	An experimental group using a basic learning-by-teaching environment described in Chapter 4
MSLQ	**M**otivated **S**trategies for **L**earning **Q**uestionnaire
SRL	An experimental group using a learning-by-teaching environment with self-regulated-learning feedback described in Chapter 5
TA	**T**eachable **A**gent
TAG-V	The **T**eachable **A**gents **G**roup at **V**anderbilt University

CHAPTER I

INTRODUCTION

A research review of science education research from 1955 to 1994 by Ponder and Kelly (1997) outlined the prevalence of a crisis in science education in U. S. school systems. Billions of dollars have been invested in research to resolve this problem but the low levels of science literacy among students remain. Science curricula need to focus on increasing literacy, link concepts learned in classrooms to real-life issues, encourage conceptual understanding, motivate students, and develop concrete problem solving skills (Bransford, Brown, and Cocking 2000; Ponder and Kelly 1997).

Problem solving requires the abilities to apply *higher-order cognitive skills* (Johnstone 1993) to *new* situations (Tsaparlis 2001). These skills include goal setting, planning, making choices about solution methods, and imposing structure on the problem-solving process by systematic application of knowledge and techniques. In contrast, typical classroom work is overly constrained, and, therefore, students seem to develop only *lower-order cognitive skills* (Stoney and Oliver 1999), such as memorizing facts and procedures and applying them without deep understanding to solving problems in homework, quizzes, and tests. Clearly, effective learning must develop in students the ability to solve complex problems, especially if the knowledge and skills learned in classrooms are to be applied in the workplace and in everyday life.

Stoney and Oliver (1999) have stated that higher-order-thinking skills are developed as students *reflect* on their learning experiences, and through this process students can incorporate new knowledge with what they already know. Traditional secondary classroom instruction does not allow much time for reflection and individual response (Ponder and Kelly 1997). Collins and Brown (1988) have conjectured that computers can be a great tool that allows students to work at their own pace and, at the same time, self-assess their knowledge, monitor their learning progress, and reflect on what they have learned. Computer-based systems for instruction and learning can be used to monitor students' progress individually while providing tailored feedback that fits their learning abilities and styles (Wenger 1987). The goal of this thesis is to design and develop computer-based learning environments that help students develop effective learning methods and problem solving skills that apply to life-long learning.

Lifelong learning is an occurring process throughout one's lifetime that supports the construction and transformation of experience into knowledge and skills (Jarvis 2002; Sloman 2002). Lifelong learning involves four types of development—**personal** (inter- and intrapersonal skills), **course-based** (schooling and training), **accidental** (unplanned learning), and **experience-based** (learning-by-doing) (Holmes 2002). This dissertation focuses on supporting the experience-based aspect of lifelong learning through the design and implementation of effective computer-based learning environments.

Effective learning includes two important issues—**understanding** and **transfer** (Bransford, Brown, and Cocking 2000). Understanding is the ability to connect new information to existing knowledge as well as the ability to realize when, where, and how to use the knowledge appropriately. Transfer is the ability to generalize one's understanding in order to tackle new problems in a variety of contexts that may include a number of different problem-solving situations.

The National Research Council has proposed a general framework for effective learning environments (the **H**ow **P**eople **L**earn framework) (Bransford, Brown, and Cocking 2000). The HPL framework defines four characteristics:

(i) **Learner-centered**: focuses on relating school subjects to students' prior experiences and understanding;
(ii) **Knowledge-centered**: focuses on imparting knowledge and skills necessary for problem solving;
(iii) **Assessment-centered**: emphasizes that learners need to receive feedback both during (formative assessment) and after (summative assessment) the teaching process to help them stay on track in terms of meeting their learning goals;
(iv) **Community-centered**: recognizes that learning can happen outside of classroom environments and encourages learning by collaboration.

This dissertation looks at adapting a learning-by-teaching model to design an effective learning environment. The potential for learning by teaching is supported by a number of studies in reciprocal teaching (Palincsar and Brown 1984), peer-assisted tutoring (Cohen, Kulik, and Kulik 1982), programming (Papert 1993), small-group interaction (Tubbs 1984), and self-explanation (Chi et al. 1994). More directly, Bargh and Schul (1980) found that people who prepared to teach others to take a quiz on a passage learned more than those who prepared to take the quiz themselves. The literature on tutoring has shown that tutors benefit as much from tutoring as their tutees (Chi et al. 2001; Graesser, Person, and Magliano 1995). In addition, the effectiveness of the learning-by-teaching paradigm finds further support in the HPL framework described above.

Teaching is a problem-solving activity that involves multiple constructive and active-learning activities. Before teaching others, teachers must assess if they have sufficient knowledge in the domain. Next, teachers must *plan* what to teach and *organize* the material to be taught into a sequence of topics. To do a good job in teaching the material, teachers must pay careful attention to both the knowledge-centered and learner-centered aspects of the HPL framework during the preparation stage. During the actual teaching process, the teacher interacts with students, while attempting to convey the knowledge in the context of what they already know. This may involve dialog processes that include probing questions to determine the students' background knowledge. During classes, teachers ask questions to gauge students' understanding of the material. These activities directly relate to the learner-centered and assessment-centered aspects of the HPL framework. Students might also ask questions, and the teacher may encourage discussions in the classroom. Afterward, teachers *reflect* on the teaching experience, using this reflection to reformulate the content of their knowledge and organization so that they may improve their teaching in future sessions. This reflection is related to the assessment-centered aspect of the HPL framework. Outside the classroom, teachers may interact with peers to exchange to domain knowledge and teaching information. This interaction can be linked to the community-centered aspects of the HPL framework.

Current efforts in building computer-based intelligent learning environments can be grouped into three primary categories: (i) Intelligent Tutoring Systems (ITS) (Wenger 1987), (ii) Cognitive Tools (Kozma 1987), and (iii) Computer Support for Collaborative Learning (CSCL) environments (Koschman 1994) that include learning companions (Ramírez Uresti 1998). None of these provide a complete framework in which to design and implement computer-based learning-by-teaching environments.

Intelligent tutoring systems have shown success in improving learning performance by focusing on problem-solving tasks and tailoring feedback to address the student's immediate needs. The strength of intelligent tutoring systems comes from their focus on knowledge-centered issues by modeling the domain and students' knowledge and using these models to aid students in their problem-solving tasks to provide tailored explanations. However, the learner-centered and assessment-centered aspects are often weak because accurate student-modeling is difficult. Some intelligent tutoring systems, such as the early cognitive tutors (Anderson et al. 1995), have focused on immediate feedback to improve problem-solving efficiency. However, immediate feedback may impede deep understanding and transfer.

Cognitive tools, on the other hand, have focused more on open-ended exploratory activities; their goal is to facilitate cognitive and high-order thinking processes through exploration with open-ended problem-solving tasks and constructions of knowledge structures to facilitate understanding. (Kozma 1987; Brown 1985). In other words, cognitive tools let learners generate and test self-constructed hypotheses (Lajoie and Derry 1993). Because these environments support active and constructivist learning, they are strongly aligned with the learner-centered aspect of the HPL framework. Nonetheless, they often suffer from the lack of knowledge-focused feedback (Cognition and Technology Group at Vanderbilt 1997). This may create situations where students cannot proceed beyond the sub-optimal levels of performance. Therefore, the cognitive-tool approaches may be weak on the knowledge-centered aspect of the HPL framework.

Researchers, such as Lajoie (Lajoie and Derry 1993) and others (Crews 1995), have proposed Interactive Learning Environments (Lajoie and Derry 1993) that combine the strengths of intelligent tutoring systems and cognitive tools. For example, a system called AdventurePlayer (Crews et al. 1997) combines exploratory learning with coaching feedback (Burton 1982) to help students progress when they reach plateaus of performance.

Computer Support for Collaborative Learning (CSCL) environments let participants drive the dynamics of their learning processes by sharing and discussing information with other participants. Therefore, these environments are solidly learner- and community-centered. However, this can also present problems because group dynamics and differences in participants' abilities affect the amount of individual learning that occurs in these environments. A proposed solution to this problem is the **learning companion** (Ramírez Uresti 1998). Learning-companion systems combine the strengths of intelligent tutoring systems and CSCL environments. The learning companion is a software agent that acts as the user's learning companion even though they may not truly learn the domain along with the user. These systems combine the internal mechanisms of intelligent tutoring systems that address the knowledge-centered aspect of the HPL framework and social constructs that address the community-centered aspect of the framework. In some learning-companion systems, the companion takes on the role of the tutee, as in learning-by-teaching systems; in **reciprocal teaching systems**, the agent and the user take turns at being the tutor and the tutee (Reif and Scott 1999).

The above discussion implies that intelligent tutoring systems, designed as problem-solving environments with tutorial components, such as lesson plans, domain resources, and knowledge-focused feedback (external assessment), mainly focus on direct knowledge acquisition. On the other hand, exploratory learning environments place more emphasis on the process of learning and transfer by letting students explore the features of the environments. This allows students to actively generate and verify (internally assess) hypotheses as parts of the learning and problem-solving tasks. CSCSL systems facilitate peer learning through different types of interactions that supply motivation and support collaborative learning and problem solving. Figure 1.1 illustrates the mapping of these main characteristics of each class of systems into the HPL framework. The lack of the check mark does not imply the absence of the feature but the inconclusiveness of its strength in the particular paradigm.

Systems	*Learner-Centered*	*Knowledge -Centered*		*Assessment-Centered*	*Community-Centered*
		Understanding	*Transfer*		
Intelligent Tutoring Systems		✓		✓	
Exploratory Learning Environments	✓		✓		
CSCL systems	✓				✓

Figure 1.1 Characteristics of the Three Paradigms of Intelligent Learning Environments

The primary focus of this dissertation, a learning-by-teaching environment, attempts to bring together the benefits of the computer systems described above. In this learn-by-teaching environment, learners teach a software tutee, called the **Teachable Agent**, using a shared knowledge representation and a shared responsibility scheme. Even though the representation is not syntax-free, it allows students to express their own knowledge. These shared components should motivate participants to become more knowledgeable in the domain while helping them develop strategies to acquire knowledge as needed to accomplish problem-solving tasks (Schwartz 1995). Students may review the domain resources provided in the environment (knowledge-centered) and then teach what they have learned to the teachable agent. Students can verify how much the agent has learned by composing queries and evaluating the agents' answers and by checking the agent's progress on a quiz designed by a subject matter expert (formative assessment). Thus, in this environment students' primary task is to learn the domain knowledge through the activity of teaching. Students have the freedom to conduct activities of their choices (researching resources, teaching, asking queries, sending the teachable agent to take a quiz) using the representations that they construct (learner-centered).

Goals of Thesis Research

The purpose of this study is to investigate the effectiveness of a new design for computer-based learning-by-teaching systems that promote learning with understanding and the ability to transfer the learning to other domains. The research has three primary components:

1. The systematic design and implementation of computer-based, learning by teaching systems with self-regulated learning feedback.

2. An agent-based implementation that allows for the design of modular and re-usable agent components and provides the flexibility of adding multiple pedagogical agents that can be embedded within different learning and problem solving contexts. The focus of this thesis has been on the implementation of two primary agents:

   A *Teachable Agent*, named Betty Bashinal, who learns explicitly from the human student using a shared representation scheme. Once taught, Betty can reason with her knowledge to answer questions and solve problems. She also demonstrates characteristics of a good student, who remembers everything she has been taught and applies self-regulated-learning strategies during the learning process to provide students with examples of how to apply self-regulation strategies to improve their learning.

   A *Mentor Agent* who combines the characteristics of a domain knowledge expert and a good teacher. This agent helps students improve their understanding of the domain and provides them with pointers for self-assessment on their own learning as well as strategies that help them become better learners and teachers. The agent also provides feedback to help students devise and implement strategies to improve Betty's problem solving performance in the form of quiz scores.
3. An experimental study conducted in a middle school science classroom to compare the effectiveness of three computer-based learning environments:

   Baseline learning by teaching (LBT): This environment included components that allowed students to prepare to teach, teach, and receive outcome and structural feedback.

   Guided learning by teaching (SRL): In addition to learning-by-teaching activities, this environment focused on self-regulated-learning strategies to help students develop certain strategies to achieve learning goals based on their insight to their own knowledge and skill states.

   Directed learning (ITS): In this environment, students were directed by the mentor agent through learning and problem-solving tasks. The mentor agent provided appropriate feedback to correct users' misconceptions and errors. The feedback received was the same as in the baseline learning-by-teaching system.

Research Questions

This thesis investigates two related research questions.

Research Question 1: Will students who learn by teaching a computer agent exhibit significantly greater understanding in a science domain that involves complicated reasoning processes than students who have been taught?

Research Question 2: Will participants who are exposed to self-regulated learning exhibit significantly greater learning understanding in a science domain that involves complex reasoning processes than students are not exposed to self-regulated learning?

Hypothesis: Computer-based learning-by-teaching environments (the SRL & LBT groups) are more effective in helping students to gain a deeper understanding and ability to transfer in science domains than the system where the learning is directed by a tutor agent (the ITS group). In addition, the computer-based learning-by-teaching environment with feedback based on self-regulated learning strategies (the SRL group) is more effective in helping students gain better understanding and develop greater ability to transfer in science domains (the LBT group).

In this dissertation, both research questions are analyzed in term of students' abilities to learn domain knowledge, develop and apply learning strategies, and demonstrate their abilities in another related domain.

Dissertation Layout

The second chapter discusses effective learning as defined by theories and models in education, cognitive psychology, and the learning sciences. This chapter also describes the process of teaching and how it can be built into learning environments to create effective learning-by-teaching environments.

The third chapter examines existing learning-by-teaching systems. This chapter classifies the systems according to the representations students use to teach the agents and the learning mechanisms that some of these agents employ. In addition, this chapter describes lessons learned from previous work and how they have been employed to design and implement a new generation of learning-by-teaching environments.

Chapter 4 presents implementation details of one of our *Teachable Agent* systems, called **Betty's Brain**. This chapter also discusses a previous experiment that has studied the contribution of each of the components of the environment to the learning-by-teaching process. The outcome of this study has led to the design and implementation of the system that is the subject of this dissertation, a more advanced and refined system with a more interactive teachable agent that uses self-regulated learning strategies.

Chapter 5 describes the design of the Betty's Brain environment that is modified with self-regulated learning (Zimmerman 1989) features. Self-regulated learning features make the teachable agent more interactive, more responsive during the teaching phase, and more aware of her own learning progress. The chapter also presents a systematic multi-agent approach to the design of teachable agent systems, and applies this design approach to introduce a second software agent, the **Mentor Agent,** who gives feedback on cognitive processes of learning and teaching.

Chapter 6 describes the study designed to test the contribution of the self-regulated learning approach toward learning with understanding and transfer. The study investigated the contributions of three learning processes in enhancing learning outcomes for fifth-grade students—directed learning (ITS), self-guided learning (LBT), and self-guided learning with feedback supporting self-regulated learning (SRL)—in the domain of river ecosystems.

The seventh chapter reports the outcome of this study and implications for the design of intelligent learning environments. This study measures knowledge gains, learning behaviors, and the ability to transfer. Results from this study illustrate the importance of different tasks and the type of feedback given when designing intelligent learning environments for children learning complex, scientific domains.

The last chapter presents the conclusions of the research and outlines future work that extends the Teachable Agents framework.

CHAPTER II

AN OVERVIEW OF LEARNING AND TEACHING

The cognitive science and educational literature agrees that learning in a particular domain is effective when one gains a conceptual *understanding* of the domain knowledge and is able to apply this knowledge to solve nontrivial problems in the domain. True learning occurs when the learner can *transfer* their understanding to different situations that he has not encountered before (Bransford, Brown, and Cocking 2000). Effective learning occurs in *active* and *constructive* environments, where learners receive support to develop cognitive and meta-cognitive strategies that can be generalized and applied to a variety of problem solving situations. This research focuses on a certain aspect of meta-cognition, namely **self-regulation**, which is considered to be an important factor in the context of learning new materials (Moses and Baird 1999). The *framework for effective learning environments* (Bransford, Brown, and Cocking 2000) serves as a guide to integrating these learning theories and practices to create effective learning environments.

Teaching is a problem-solving task (Artzt and Armour-Thomas 1999) with multiple phases, preparation to teach, actual teaching, and assessment. Before teaching others about a domain, one has to first understand the domain. In addition to a deep understanding of the subject, successful teachers must also have monitoring skills and the ability to reflect on the feedback they receive when they teach and also when they grade their students' homework and tests.

This chapter provides an overview of relevant learning theories, and discusses a framework for integrating these theories and models into an effective learning environment.

Learning Theories

As stated in Chapter 1, this dissertation focuses on designing an environment that supports effective learning. This section characterizes effective learning and learning theories that support effective learning, namely active learning, constructivism, and meta-cognition.

Effective Learning

Experts in the learning sciences characterize effective learning as the process that leads to deep understanding, long-term retention, and the ability to apply what is learned to different problem-solving situations (Bransford, Brown, and Cocking 2000). Deep **understanding** supports generalization and transfer to other problem-solving situations and domains, enabling one to tackle a wide variety of problems. In other words, effective learning goes beyond remembering facts to realizing the essence of knowledge and how to apply it in various situations. Effective learning is supported by prolonged, authentic interaction with domain knowledge such that learners have ample opportunity to discover when and where to apply their learning to obtain solutions to complex problems or create new approaches to solving problems.

Transfer describes the ability to apply knowledge appropriately in a variety of contexts (Detterman 1993; Haskell 2001) that may include a number of different problem-solving situations. Transfer also covers the ability to assimilate and learn new knowledge and the ability to relate it with existing knowledge (Cormier and Hagman 1987). The literature describes various forms of transfer as **far** and **near** (Detterman 1993), also called **lateral** and **vertical** transfer (Cormier and Hagman 1987), respectively. When the new situation has a partial relationship to the previous experience, that transfer process is *near*. Otherwise, it is classified as a *far* transfer. Another interesting classification is **specific,** transfer of the domain contents, and **non-specific**, transfer of general skills and principles (Detterman 1993). Transfer has very strong implications in education since it means applying knowledge or skills learned in classrooms and training courses to real-life situations. If transfer is low, years in school and money put in training may be wasted.

The definitions of understanding and transfer described above are the key characteristics that differentiate effective learning from such passive learning and learning by rote, which have been known to lead to shallow understanding, the inability to remember information for significant periods of time, and difficulty applying learnt knowledge to problem solving situations (Cognition and Technology Group at Vanderbilt 1997). Understanding and transfer are closely linked. Involvement in various types of known and unknown problem solving exercises often helps students develop an understanding of how to define a problem space and identify problem-solving options. However, problems solving exercises have limited applicability; being able to solve problems in a familiar domain does not directly imply the ability to solve non-trivial problems in other domains. The transfer can only happen if one develops a deep understanding of fundamental principles in the domain. Understanding and transfer are also linked to the development of meta-cognition (Bransford, Brown, and Cocking 2000), an important characteristic of effective learners.

Active Learning

Active learning suggests that students should control their own learning processes (Dewey 1933; Bonwell and Eison 1991). Learning is more effective when students are actively involved in learning activities (Bonwell and Eison 1991), such as discussing the ideas, writing essays about what they have learned, relating their past experiences to new information, and applying knowledge to real-life situations (Chickering and Gamson 1987). Listening to a lecture can be an active learning activity if the instructor allows time for students to recall and reflect on the notes taken during the lecture (Ruhl, Hughes, and Schloss Winter 1987). Other examples of activities that can support active learning are reading, discussion, problem solving, and higher-order thinking tasks, such as analysis, synthesis, and evaluation (Chickering and Gamson 1987). A research review has shown that active learning not only facilitates the mastering of content knowledge but it also supports the development of thinking skills (Bonwell and Eison 1991).

Several barriers hinder active learning in traditional classroom environments (Bonwell and Eison 1991). First, teachers need to acquire a practice that successfully promotes active learning with the limited time and resources that are available in a classroom.

Active learning activities are very likely to increase class preparation time for teachers. Applying active learning strategies in large classrooms may present great challenges since students may come to school with widely different backgrounds and competencies.

There are also barriers for students in active learning classrooms (Bonwell and Eison 1991). To take ownership of their learning process, students must build and refine strategies to assess their understanding of information and mechanisms for evaluating its value. If students do not understand a piece of information, they need to know how to find related information that might help them. Then, students must decide if the newly-found information conforms or contradicts their existing knowledge. In case of conformity, students need to acquire the skill to create theories of phenomena by connecting the information to their existing knowledge structures. In case of contradictions, students need to locate the contradicting facts and to determine a way to resolve them. Finally, to ensure that these new theories are accurate, students need opportunities to test them.

These challenges to active learning in the classroom motivate the need to develop learning environments that support students' active-learning activities and assist teachers in managing and supplementing classroom activities to achieve effective learning. The National Research Council has proposed characteristics for such environments (*How People Learn* (Bransford, Brown, and Cocking 2000)). This framework is discussed later in this section.

Constructivism

Constructivism is a theory of knowing (Piaget 1978; Vygotsky 1978). Constructivism focuses on learning as a cumulative process where one constructs new knowledge from active engagement in tasks, and the interpretation and assimilation of new information as a function of previous knowledge. A constructivist classroom respects students' prior experiences and beliefs, and encourages students to play an active role in classroom activities under the guidance of the instructor. The following paragraph provides an illustrative definition of *constructivist learning*:

> *"Constructivist learning is based on students' active participation in problem-solving and critical thinking regarding a learning activity that they find relevant and engaging. They are "constructing" their own knowledge by testing ideas and approaches based on their prior knowledge and experience, applying these to a new situation, and integrating the new knowledge gained with pre-existing intellectual constructs* (Briner 1999)*."*

Constructivist learning environments often challenge learners to construct multiple representations of real-world situations and make students realize the level of complexity of the real world while discouraging them from generating oversimplified solutions to problems. By being active in learning activities, learners are encouraged to reflect on their experiences and motivated to fill shortcomings in their knowledge.

Two primary constructivist paradigms discussed in the educational literature are **cognitive** (Piaget 1954) and **social constructivism** (Vygotsky 1978). Cognitive constructivism is based on Jean Piaget's theory that transferable knowledge cannot be imparted in an inert manner, but has to be constructed through experiences. According to Piaget, learning is not drill-and-practice on isolated skills; learning must involve complex-task environments to expose students to a wide range of opportunities to develop skills in meaningful and realistic contexts. Social constructivism entails cognitive constructivism but emphasizes the social context of learning. Proponents of social constructivism assert that learning is a social activity, and it should not be separated from the real world. Students construct knowledge through social interactions where new knowledge is constructed through dialog and discussion with others, such as peers with different experiences, teachers, and parents. Furthermore, students learn via collaboration and negotiation of cognitive conflicts rather than just competition for recognition, such as making better grades for the sake of the instructor or their parents (Jonassen 1994).

Evidence shows that learning is more effective when teachers pay attention to students' experiences and beliefs (Bransford, Brown, and Cocking 2000). Teachers can employ a number of methods to implement constructivism, such as problem solving and anchored instruction (Cognition and Technology Group at Vanderbilt 1997). However, many of the same barriers that make active learning difficult to employ in the traditional classroom also hold true for constructivist approaches to teaching.

Metacognition

Metacognition is the process of thinking about one's own cognitive processes. As defined by Flavell,

> *"Metacognition refers to one's knowledge concerning one's own cognitive processes or anything related to them, e.g., the learning-relevant properties of information or data. For example, I am engaging in metacognition if I notice that I am having more trouble learning A than B; if it strikes me that I should double check C before accepting it as fact* (Flavell 1976)."

Flavell's statement specifies that metacognition is important to learning in general. Without evaluating one's current cognitive state, learning is goalless and, therefore, not meaningful.

There are three basic elements in metacognition: (i) action planning, (ii) progress monitoring and (iii) plan evaluation (Pressley, Borkowski, and Schneider 1987). Before engaging in a learning activity, one should plan on how to achieve his learning goals. This planning process includes realizing the role of existing knowledge in fulfilling learning goals and deciding on activities that need to be performed if the existing knowledge is insufficient. During plan execution, one needs to monitor the effects of actions and determine if the plan needs to be modified to achieve the desired goal. After the plan is finished, one should evaluate how well the actions have led to the accomplishment of the goal. In addition, one should evaluate if the plan is optimal, if there are alternatives, and if some tasks should be completed again to improve the learning experience. Many research studies have shown that metacognitive training can improve both learning ability and attitude toward learning (Alexander, Carr, and Schwanenflugel 1995; Franks et al. 1982; Jaušovec 1994).

Active learning and constructivism enhance metacognitive-skill development since they involve two critical learning skills—sense making and self-assessment (Bransford, Brown, and Cocking 2000)—that are Flavell's examples in the quotation above. In addition to active learning and constructivism, self-regulation of learning is another method that can aid the development of metacognitive ability (Moses and Baird 1999).

Self-Regulated Learning

"The construct of self-regulation refers to the degree that individuals are metacognitively, motivationally, and behaviorally active participants in their own learning process. (Zimmerman 1986)*"*

Students can self-regulate their learning process by committing to desired academic goals, initiating self-regulated learning strategies to achieve these goals, and monitoring **self-efficacy**, the self-perception of performance, while learning. **Self-regulated learning strategies** involve actions and processes to acquire knowledge and skills. Self-regulation is a critical set of skills in developing problem solving ability (Pintrich and DeGroot 1990). Zimmerman (1989) described the fifteen categories of self-regulated learning skills:

1. *Self-evaluating*: Self-initiated evaluation of work quality and progress. An example is ensuring the correctness of one's work.
2. *Organizing and transforming*: Self-initiated arrangement of instructional materials to improve learning. This could involve making an outline of topics to be studied, and then organizing them to facilitate understanding of the relationships between the topics.
3. *Goal-setting and planning*: Setting of academic goals and sub-goals, and planning activities to achieve the goals.
4. *Seeking information*: Self-initiated efforts to obtain new non-social information that could be useful to a specific task.
5. *Keeping records and monitoring*: Self-initiated efforts to record events and results important to monitoring learning progress. An example is taking notes of past errors and analyzing them so that they are not repeated.
6. *Environmental structuring*: Self-initiated actions to rearrange the physical workspace to improve learning. An example is turning the radio or the television off while studying
7. *Self-consequating*: Self-reinforcement efforts in a physical or imaginary form. Examples are going to a movie only after finishing homework, and commitment to work harder when a grade is not as good as expected.
8. *Rehearsing and memorizing*: Self-initiated efforts to memorize procedures and facts important to solving the problem.

9–11. Self-initiated efforts to seek assistance from (9) peers, (10) teachers and (11) Adults.

12–14. Self-initiated efforts to review records, such as (12) *notes*, (13) *tests* and (14) *textbooks*.

Other: Efforts to execute learning behaviors suggested by others, such as doing what the teacher recommends.

Since learning-by-teaching is a self-directed and constructive activity, teachers can use self-regulated learning strategies to improve the quality of their teaching. For example, while preparing materials to teach, a teacher *sets the goals* about what to teach. Then, he *evaluates his state of knowledge* of the topic. If he is not satisfied with the assessment of his knowledge, he needs to *decide* what the missing pieces are. From this information, he *plans* what to learn and how to learn it. He also *evaluates his state of knowledge* while studying. During and after learning, he has *made an outline* of what to teach his students. If he is not satisfied with his existing knowledge, he *seeks further information*. While teaching, the teacher interacts with his students in many ways, such as asking them questions and giving them quizzes, to evaluate if they understand what he has taught. Often, he learns what his students misunderstand from these interactions. He also *records and analyzes feedback* to improve his knowledge and quality of teaching.

Framework of Effective Learning Environments

Learning theories emphasize that learning environments should be tailored to suit students' needs. According to a National Research Council research review (Bransford, Brown, and Cocking 2000), effective learning environments should combine knowledge-centered, learner-centered, assessment-centered, and community-centered learning paradigms and activities. Note that the term "environment" in this section refers to generic learning environments, not specifically a computer-based learning environment that is the focus of this research work.

By being **knowledge-centered**, the environment supports students in gaining expertise in the knowledge and skills necessary for problem solving. Being "knowledgeable" means the development of understanding of the material and the ability to apply knowledge and skills gained to problem solving tasks, which can be encouraged by active learning. This is in contrast to rote-learning situations where students memorize information only to score well in standardized tests. The curriculum must be well planned to lead students to develop well-organized bodies of knowledge. This is critical for the development of metacognitive skills so that students can understand and make sense of new information and apply what they have learned to problem solving situations.

By being **learner-centered**, the environment relates the subject matter to students' prior experiences and understanding. This supports students in developing knowledge by building on their own reasoning skills, beliefs, and experience. In addition, a learner-centered environment respects discourse, both between the teacher and students and among students, as articulation can bring deeper understanding and transfer. Sharing opinions on cognitive conflicts can clarify misconceptions and encourage students to learn to think critically about their own and others' ideas. In short, the environment should allow students to be active and constructive and encourage self- assessment.

By being **assessment-centered**, the environment provides learners with opportunities to receive feedback and then revise their learning processes. This aspect is very important for learning (Bransford, Brown, and Cocking 2000) as it helps develop metacognitive skills. Assessments reflect one's level of understanding and make one become aware of the gap between the current state of knowledge and skills and the desired goal (Ramaprasad 1983; Sadler 1989). School systems employ two major types of assessment, **formative** and **summative**. Formative assessments evaluate progress during the learning process to guide students to achieve their goals. This can range from discussions with the teacher and peers to feedback on homework and quizzes. Formative feedback is especially helpful when it specifically comments on errors, makes suggestions for improvement, and offers encouragement, instead of simply focusing on getting correct answers (Bangert-Drowns, Kulick, and Morgan 1991; Elawar and Corno 1985). Summative assessments are evaluations that determine the final outcomes upon the completion of a set of activities, such as tests at the end of the semester (Angelo and Cross 1993). Summative assessments help students, and teachers, reflect on the effectiveness of the learning environment, and the strategies of the learner.

Learning environments need to properly combine these two kinds of assessments to help students attain their learning goals. An environment that employs only formative assessment may not provide the students the right perception of their aggregate cognitive-levels. On the other hand, an environment that offers student only summative assessments may provide the help too late to aid students in the learning process.

Not all assessments are useful, especially those that focus on delivering the correct answers rather than the process of thoughts and strategies (Black and Wiliam 1998). Thus the emphasis of evaluations and feedback should not be on memorization and reproduction of information but on conceptual understanding of that information and its transfer. Good assessments should make students' thinking visible to both the teacher and the students. This visibility can aid students in realizing gaps in their understanding, in revising their work, and in developing the skills of self-assessment, another important factor for the development of meta-cognitive skills, which includes self-regulation.

By being **community-centered**, the environment encourages learning by collaborating in a learning environment, which could be a classroom, a school, a neighborhood, or society at large. Expanding learning beyond the individual level can motivate one to learn both intrinsically and extrinsically. Being asked to present their work to outside audiences can motivate students to focus more on work quality (Bransford, Brown, and Cocking 2000). Interactions within a small community via structured learning activities, such as tutoring, can improve learning significantly because of the articulation about thinking and learning (Goodman et al. 1998; Chi et al. 2001). Other than learning opportunities, a community also provides chances for learning that one can apply and expand one's knowledge in real-life situations. Hence, the community-centered aspect suggests that interactions with others in a community can lead to effective learning anywhere and anytime, especially if students realize the opportunities to do so.

To summarize, the HPL framework is a base for implementing active and constructivist learning. The framework also extends to include self-regulated learning that helps students develop independent learning strategies. These ideas are very important in the development of learning-by-teaching environments. The next section begins with a discussion of the teaching process. It is crucial to emphasize that the teachable-agent research focuses on enhancing the learning of the students who participate in the study, and not on developing them into real-life teachers. Therefore, we have paid attention to the detailed activities of teaching to decide which activities benefit middle school students as learners.

Teaching Process and Techniques

Teaching Process

Teaching is a reflective and iterative process that consists of three main phases, *decision-making*, *performing-actions*, and *monitoring* (Colton and Sparks-Langer 1993). In the decision-making phase, teachers evaluate their current knowledge against goals and plan actions to acquire new knowledge, if necessary. Teachers can gain knowledge on the subjects to be taught from a variety of sources and methods, such as reading recent professional journals, books and reports, conducting research to gain more information, and consulting personally with experts in the domain. This results in the creation and organization of knowledge structures that facilitate understanding and problem solving skills. In the performing-action phase, teachers decide on the content they will teach and create lessons plans to communicate this knowledge in a structured form to students. In the monitoring phase, teachers analyze and reflect on their interactions and students' performance in class, on homework assignments, and on examinations to draw conclusions on their students' knowledge. The outcome of what students have or have not learned affects further planning and activities teachers will perform in the next iterative process of teaching. Figure 2.1 illustrates the flow of this process.

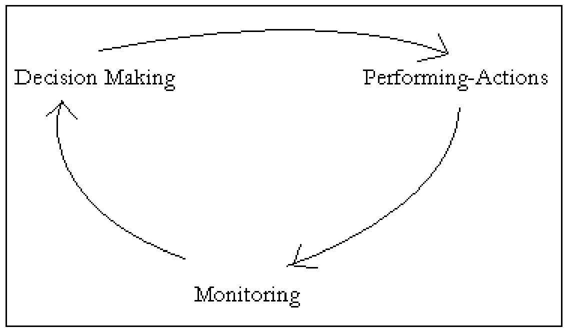


Figure 2.1 A Teaching Process

Teaching Techniques

Teaching can be executed using various techniques. The most common technique in schools is lecturing where teachers have the full control of classes and often are more active than their students. This sub-section presents a brief overview of other selected techniques described in (Bransford, Brown, and Cocking 2000) that enable students to be more active in their learning.

Inquiry-based teaching is based on John Dewey's theory (Dewey1933) that curiosity motivates learning. Novak (1964) stated that "*inquiry is the [set] of behaviors involved in the struggle of human beings for reasonable explanations of phenomena about which they are curious.*" Inquiry-based teaching encourages students to be curious and ask questions, reflect on what they know, and search for information so that they may learn to answer those questions to satisfy their own curiosities. Therefore, inquiry-based learning can promote critical-thinking and self-assessment skills. There are many practices in inquiry-based teaching, such as problem-based teaching (Neufeld and Barrows 1974) and discovery learning (Bruner 1987).

Individual or group teaching. Individual teaching involves an instructor and a student and typically describes one-on-one tutoring. The one-on-one format allows customization of knowledge to fit the student's cognitive state, and the student receives the full attention of the instructor. In contrast to learning directly from the instructor, students in a group setting learn through collaboration among peers and the instructor. For example, group projects monitored by the instructor are an example of group teaching. However, a group of students in a lecture may not benefit as much from the group teaching if there are few interactions during the lectures.

The last category, **technology-enhanced teaching,** is not really a teaching technique in it own right, but rather a way to enhance other real-world teaching techniques, such as classroom teaching. Teaching techniques in this category range from audio and visual media to computer-assisted techniques, such as computer simulation.

Different teaching techniques can be combined , such as a lecture-based classroom where the instructor pauses every 15 to 20 minutes to allow students to reflect on and discuss what they have learned (Ruhl, Hughes, and Schloss Winter 1987). This implementation enhances lecture-based instruction with a degree of inquiry. In the same manner, a combination of teaching techniques can support active and constructivist learning in a learning environment. The next section outlines a design of learning-by-teaching systems that results from the materials presented previously.

Creating an Effective Learning-by-Teaching Environment

Previous sections discussed the many obstacles that hinder creating effective learning environments in traditional classrooms. Introducing changes in the school system will naturally take extensive time and effort (Joolingen and Jong 1996), but changing the way individuals learn is a more feasible task. Computer technology has advanced at a rapid pace in terms of computational power, memory, interfaces, and networking. Computers can be a great tool for learning since we can tailor software to suit students' individual characteristics and needs (Wenger 1987). In addition, the ability to store information in an organized way makes computers a great tool for reflecting on the learning process (Collins and Brown 1988; Stoney and Oliver 1999). Therefore, it is possible to engineer software that supports the HPL framework, learning-by-teaching, and even active, constructivist, self-regulated, inquiry learning.

Using the HPL framework, computer-based learning environments, especially intelligent learning environments, can be built to be learner-, knowledge-, assessment-, and community-centered. An intelligent learning environment is a computer application that allows students to explore a complex domain with guidance that is tailored to individual needs of each student (Lajoie and Lesgold 1989). An important characteristic of intelligent learning environments is that users are allowed to progress at their own pace (Lajoie and Derry 1993). Some systems, such as CSCL systems, support social constructivism by allowing users to retain the locus of control. In systems where students collaborate, they can become active participants in the learning process (Koschman 1994). Moreover, environments with cognitive tools are designed to let students construct new knowledge or cognitive structures based on their previous experience (Kozma 1987).

Intelligent learning environments can also be knowledge-centered. They often provide resources, tutorials, and lessons-these can be tailored to users' needs. However, in some cases, these systems provide general knowledge on a set of topics rather than emphasizing problem solving and deep understanding (Bransford, Brown, and Cocking 2000; Schmidt 1997). Such systems may not enable knowledge connectivity (Ponder and Kelly 1997). It is important to develop mechanisms in the systems that allow students to connect isolated pieces of knowledge in the context of understanding complex phenomena while they are involved in problem solving tasks.

Intelligent learning environments can also be assessment-centered. The role of assessment, both formative and summative, is the key to active learning, understanding, and the ability to transfer. Formative assessment helps students reflect on the state of their knowledge and evaluate how far they are from their learning goals. This sets up a framework where students can be advised on how to achieve learning goals on topics where they have not performed well. For instance, intelligent tutoring systems give formative assessment in the context of problem-based learning (Wenger 1987). Students try to solve problems and receive feedback on how to improve performance. This is where self-regulated learning comes in. For feedback to be useful, students must know what to do after receiving feedback. They may have to adjust their plans, carry them out, and monitor their own progress as they execute their plans.

Finally, intelligent learning environments can be community-centered. The World Wide Web provides universal connectivity, and this has been exploited in games and CSCL environments. There are examples where some of the participants are software agents, as seen in learning-companion systems, such as (Ramírez Uresti 2000). Not all activities in collaborative environments are conducive to learning. Collaborative environments that support learning encourage different participants to initiate and participate in discussions, employ active-learning skills, and assess their own and the group's performances (Soller 2001).

An effective environment should satisfy all four aspects. Lacking any aspect may have a negative effect on the effectiveness of the learning environment. For example, students using a system that does not have assessment mechanisms will have difficulty realizing the actual state of their knowledge, and those in another environment without knowledge components will not have a source of new information that could lead to detection and correction of misconceptions.

The next chapter examines existing learning-by-teaching systems. The lessons learned from these systems are used in conjunction with the principles presented in this chapter to design effective learning-by-teaching environments.

CHAPTER III

LEARNING-BY-TEACHING SYSTEMS

Pedagogical agents are defined as "*animated characters designed to operate in a human setting for supporting and facilitating learning* (Shaw, Johnson, and Ganeshan 1999)". They typically combine a tutoring approach (Wenger 1987) with more human-like behavior and animations, with the understanding that the more natural interactions provide the greater motivation to students in their learning tasks (Clark and Mayer 2003). In general, pedagogical agents have been designed to take on different roles, such as a tutor, a coach, and a peer. In our work, we extend the notion of pedagogical agents to include those that are designed to incorporate the learning-by-teaching paradigm. In other words, these agents play the explicit role of a tutee, such as some one who is taught by a teacher. This chapter examines only pedagogical agents that assume the role of the tutee in learning-by-teaching systems. These agents must have means that allow their users to teach them explicitly. Example systems include those where students teach their agent by example (Michie, Paterson, and Hayes-Michie 1989), and by creating knowledge structures (Nichols 1994; Palthepu, Greer, and McCalla 1991).

We categorize pedagogical agents that adopt the learning-by-teaching paradigm into three groups.

1. **Learning-by-teaching agents that learn directly** from the users of the system via representations, such as causal maps, that are shared by both the LBT agent and the user. This category can be further split into:
   - 1.1. Agents that use **explicit shared-representations**: Everything users have taught the agents is visible in the shared representation between students and the agents.
   - 1.2. Agents that use **implicit representations**: The information taught by students affects the agents' knowledge, but there is no shared visual representation where students can see what they have taught the agents.
2. **Learning-by-teaching agents that learn indirectly** store the information students have taught using internal representations that are different from those used by the students to teach these agents.
3. **Pseudo learning-by-teaching agents** pretend to learn from what students have taught them but actually have full knowledge of the domains of study.

This chapter discusses these categories and systems that fall under them in details. Lessons learned from these systems have led to a new design for learning-by-teaching environments presented in the last section that can satisfy the specifications of effective intelligent learning environments. Because these discussions involve both human and software-based agents, for clarity human students are referred to as *students*, and software-based agents as *agents*, *LBT agents* and *systems*.

Learning-by-Teaching Agents that Learn Directly

The agents in these systems develop their knowledge for solving problems and answering questions from what their users teach them. Systems in this category may have *a priori* knowledge of the domain or may start *tabula rasa*. This section presents three learning-by-teaching systems using different representation schemes, namely a qualitative model, examples, and semantic networks.

Explicit, Shared Knowledge-Representation Systems

Systems in this sub-category learn directly from students via viewable, shared representations. In other words, students teach their agents using graphical-user interfaces, and the changes students have made to the visual representations directly affect the contents of the representation structures that the agents use for answering questions and solving problems. A literature survey indicates that there is no existing learning-by-teaching system that fall into this category.

Implicit Knowledge-Representation Systems

Systems in this category learn directly from their users using shared representation schemes that are not made visible to the students. Therefore, students can modify the knowledge structure, but receive no visual feedback on how those changes affect their agents' learning. Instead, students have to query the agents to examine the consequences of the modifications and additions they have made to the agents' knowledge structure. This sub-section presents two of such systems in this category.

DENISE

DENISE (Development Environment for an Intelligent Student in Economics) was a learning-by-teaching system where the LBT agent assumed the role of a tutee in a peer tutoring situation (Nichols 1994, 1994). DENISE was built as an exploration of the feasibility of implementing learning-by-teaching systems that the student agent had no domain knowledge (*tabula rasa*) to start with and students directly taught the agent. In this case, the student was asked to construct a causal qualitative model via a dialog, like the one shown in Figure 3.2, to teach DENISE. Unlike intelligent tutoring systems, this system did not incorporate student modeling. DENISE's interactions were driven by the framework of a Socratic teaching strategy based on what the students taught. For example, in Figure 3.2, DENISE asked the student to complete a relationship starting with the concept "Investment." The relation types were limited to the six shown in the figure. Interestingly the Socratic strategy employed here was the reverse of the tradition one that the teacher probed the student by asking questions. In other words, in DENISE the agent who was playing the role of the learner probed the student who was playing the role of the teacher for more information.

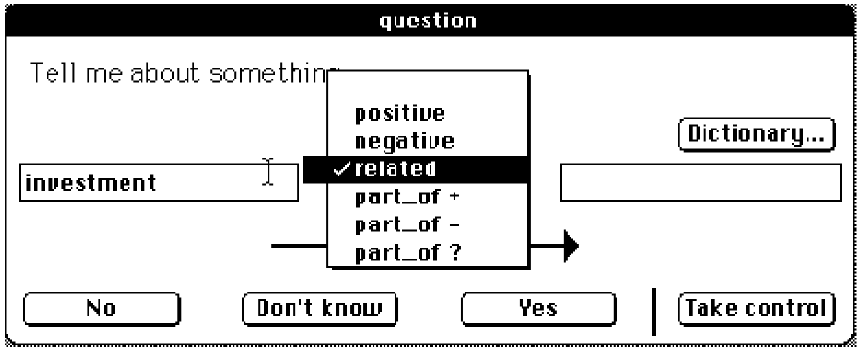


Figure 3.2 Syntax of DENISE (Nichols 1994)

Students could take control from DENISE at any time by clicking the "Take control" button at the right-bottom corner. This gave them the freedom to specify any relationship of their choice instead of being guided by DENISE. For example, the idea that the demand was positively related to price was specified in DENISE as the link "Demand *positive* Price" where "Demand" and "Price" were concepts and "positive" was the type of link selected. Because the causal model was not visible to students, there were often redundant concepts in the knowledge structure, such as "money demand" and "demand for money". Therefore, the dictionary feature displayed in Figure 3.3, called by clicking on the "Dictionary…" button on the right-top corner of the dialog in Figure 3.2, showed all concept labels that the student had taught the agent thus far.

Students could also compose queries to DENISE. The agent applied qualitative-reasoning techniques to its causal model to answer these queries. This feature motivated self-explanations (Chi et al. 1989). However, the details of how DENISE answers questions was not specified in (Nichols 1994a, 1994b).

Figure 3.3 Dictionary of Used Concept Names (Nichols 1994)

An experiment was conducted with five participants to test DENISE. The participants perceived the idea of teaching to be motivating. Nonetheless, the lack of a visual representation of the qualitative model frustrated the participants after they had a number of concepts and links in DENISE. This was primarily because the participants reported difficulty in recalling concepts and relationships they had taught DENISE earlier. The query feature was used sparingly. We speculate that this was also a consequence of the implicit representation because, without a visual representation, it could be difficult for students to understand why DENISE answered a question in a particular way.

Learning-by-Teaching Agents that Learn Indirectly

Learning-by-teaching systems in this category create internal knowledge representations from the information that their users provide during the teaching processes. Example systems of this category are Math Concept Learning System (Michie, Paterson, and Hayes-Michie 1989) and a learning-by-teaching system presented in (Palthepu, Greer, and McCalla 1991).

Math Concept Learning System

The Math Concept Learning System (MCLS) (Michie, Paterson, and Hayes-Michie 1989) learns a general set of rules for solving sets of linear-equations. The system has built-in knowledge of the structure of linear equations but has no knowledge of strategies that may be used to solve these problems. To teach the system, the student first creates a linear equation or asks the system to create one. Then, the student generates a solution for this linear equation by constructing a sequence of actions from the six actions provided by the system as defined below.

1. Divide both sides of the equation by a common factor.
2. Collect like terms on the same side of the equation.
3. Multiply out bracketed terms.
4. Cancel out a term that appears on both sides.
5. Divide by the coefficient of the unknown.
6. Stop because the equation is solved.

In response, the system creates a set of general rules for solving linear equations using an inductive machine-learning algorithm, Iterative ID3 (Shapiro 1983). Figure 3.4 displays two examples of such rules; (a) is considered to be close to an ideal rule for solving linear equations, and (b) contains many errors. If two solutions generate two different sets of rules, the system asks the user to decide whether he would like to correct one of the solutions.

IF one side has more than one like term
THEN combine like terms
OTHERWISE
IF there are like terms on opposite sides
THEN collect like terms on the same side
OTHERWISE
IF the unknown term has a coefficient of one
THEN stop because the equation is solved
OTHERWISE divide by the coefficient of the unknown

(a) A Good Rule

IF one side has more than one like term
THEN EITHER collect like terms on the same side
        OR combine like terms
OTHERWISE EITHER collect like terms on the same side
                OR divide by the coefficient of the unknown

(b) A Bad Rule

Figure 3.4 Examples of Rules Produced by the MCLS

At any time, students can assess the rules the system has created using three functions:

1. *Look*: Request the system to display the current rule.
2. *Test*: Request the system to solving a linear equation using the current rule.
3. *Ask*: Debug the current rule by requesting the system to step through its current rule.

A preliminary study was conducted to compare MCLS to three other systems, MCLS/NLV, *Success with Math*, and the control condition (no treatment). The MCLS/NLV system was a version of the MCLS system without the learning part. Therefore, it was similar to a drill-and-practice program. Students still performed steps 1 and 2, but each action the student selected was compared against the action that the system selected based on the master rule pre-defined in the system. If they were different, the system suggested that the student changed that action. The Success with Math system had a slightly different set of actions and feedback. Students were told of the correct action every time they made a mistake.

The study was a pre- and posttest design. Thirty students were divided into four groups ($n_{MCLS} = 8$, $n_{\text{MCLS/NLV}} = 7$, $n_{SWM} = 8$, and $n_{control} = 7$), and each group used one program throughout the experiment with the exception of the control group being untreated. All students except those in the control group spent fifteen to thirty minutes each day working with the assigned programs for seven days. Even though the total time spent was not reported, the author implied that all the students using the systems spent approximately the same amount of time on the systems.

The sample sizes were so small that the significance levels were not reported. However, the MCLS condition had the highest gain of pretest to posttest scores (+2.5). The other groups reported gains of SWM = +1.7, MCLS/NLV =0.3, and the control group = -0.7.

A Learning-by-Teaching Systems with the Natural-Language Communication

Palthepu, Greer and McCalla (1991) proposed the design of a learning-by-teaching system that could be taught via a natural language. The system was designed to be domain independent and initially had no domain knowledge. It gradually learned concepts and relations using a semantic net structure (Quillian 1968) by communicating with students through a natural language interface. Like DENISE, this system adopted a reverse form of the Socratic teaching style (Seeskin 1987). The system asked questions to get students to explain what they had communicated (see Figure 3.5), and parsed the responses into a semantic-net structure. The expectation was that the system's queries would lead students to perform self-explanations (Chi et al. 1989), which would help them gain deeper understanding on the topic of conversation. Unfortunately, the details of the implementation and experiments conducted on the system are not mentioned in the paper and there has been no follow-up publication.

Student: I will tell you something about the animal kingdom.
System: OK. Please proceed.

Student: There are two types of living things. Plants and animals.
System: So every living thing is either a plant or an animal.

Student: Yes. And all mammals are animals. And so all reptiles.
System: That means there are two kinds of animals?

Figure 3.5 Sample Dialogs of Teaching via English in Palthepu et al's Proposed System

Pseudo-Learning LBT Agents

Systems in this category give the appearance of being taught by students, but actually the agents have full *a priori* knowledge of the domain of study. In interacting with the students, these agents give the impression that they do not know much about the domain but gradually learn as they are taught by the students.

A Virtual Classroom

Obayashi, Shimoda and Yoshikawa (2000) reported an implementation of a learning-by-teaching system that is based on a traditional learning approach. First, students attended a set of lectures, and then solved a set of problems on the system. The system diagnosed students' weaknesses in domain knowledge from the solutions they generated. Next, each student entered the virtual discussion room, where a virtual student asked the student questions that focused on his weaknesses. Sample questions asked were "do acids act on metal?", "how do they act?", and "why do they act?" The virtual student used the students' answers in the last phase of the system in a virtual classroom that was monitored by a virtual tutor. In this phase, students could observe all virtual students' performance when answering the virtual tutor's questions. The tutor also provided correct solutions to the questions that it asked.

A pilot experiment was conducted as a modified switching-replication design (Trochim 2001) to test the system on forty college students, split into two groups of twenty. The experiment had no pretest. In the first phase, one group used the system and the other group used a comparable text-page-styled version without the virtual student. All students took a posttest at the end of two hours. Then, the students switched groups to work on the other system for an unreported period of time. At the end of this second phase, students answered a questionnaire that asked them to compare these two systems.

The study reported that the group that used this learning-by-teaching CAI system scored significantly higher than the group that used its text-page version without the virtual student at the end of the first phase. In addition, in response to the questionnaire the participants reported higher levels of motivation and effectiveness of learning using the virtual-agent version of the system than using the text-page-styled version, but the levels of significance of the differences were not specified.

Diagnosis-Hint Tree

Chan and Chou (1997) conducted a preliminary study to compare different approaches of intelligent learning environments that included peers. The study was constructed around two types of agents, the tutor and the peer. The peer could be real, software-based, or could be absent. The tutor could only be software-based (an intelligent tutor) except in two cases, (i) the Distributed-Responsibility Sharing condition where a real student acted as the tutor and (ii) the Self-Diagnosing and Working Alone conditions where both the peer and tutor were absent. The tutees, real students, were the users of these systems. Some possible combinations, shown in Figure 3.6, are used as the conditions in this experiment. Some conditions had the same agent combinations but differed in tools that were available for learning.

1. Agents			*2. Condition*	*3. Tools*	
4. Tutee	*5. Peer*	*6. Tutor*		*7. PLS*	*8. DHT*
9. Real	10. Real	11. Real + Virtual	12. Distributed-Responsibility Sharing	13. ✓	14. ✓
		15. Virtual	16. Distributed-Reciprocal Tutoring	17. ✓	18. ✓
	19. Virtual	20. Virtual	21. Centralized Reciprocal Tutoring	22. ✓	23. ✓

			24. Learning by Tutoring	25.	26. ✓
	27. Absent	28. Virtual	29. Intelligent Tutoring System	30. ✓	31.
		32. Absent	33. Self-Diagnosing	34. ✓	35. ✓
			36. Working Alone	37. ✓	38.

Figure 3.6 Combinations of Three Types of Agents in Chan's and Chou's Study

The tools were discussed first because it helps in defining the tasks performed by the students in each condition. Two tools were used in this study, (i) the **Diagnosis-Hint-Tree** (DHT), a diagnosis tool for LISP programs and (ii) the **Petal-Like-System** (PLS), a LISP programming helper. The DHT tool was a scaffolded, LISP debugging-tool. The interface of the DHT tool, shown in Figure 3.7, displayed a tree of possible (correct and incorrect) solutions to a buggy program that the system selected for the student. Starting from the root node of the tree, the user clicked on it to see the possible sub-solutions in the form of child nodes of the root node. The user started to diagnose the problem by selecting one of these child nodes. This action revealed more child nodes further down in the tree. Each node corresponded to a step in the solution. The user kept selecting nodes until he reached the leaf level of the tree. The system provided immediate feedback during this debugging process. It expanded a node only when the node was on the solution path. Once the user reached a leaf node, he had to choose a hint to give to the system from a list of hints that was provided with that node.

PLS is a tool that assists students in writing LISP programs. The tool scaffolds the programming task by completing parts that students have not yet mastered in the beginning and gradually reduces the support along the learning process (Chan and Chou 1995). Also, the tool reduces syntax errors by providing blocks of open and close parentheses when students create their programs.

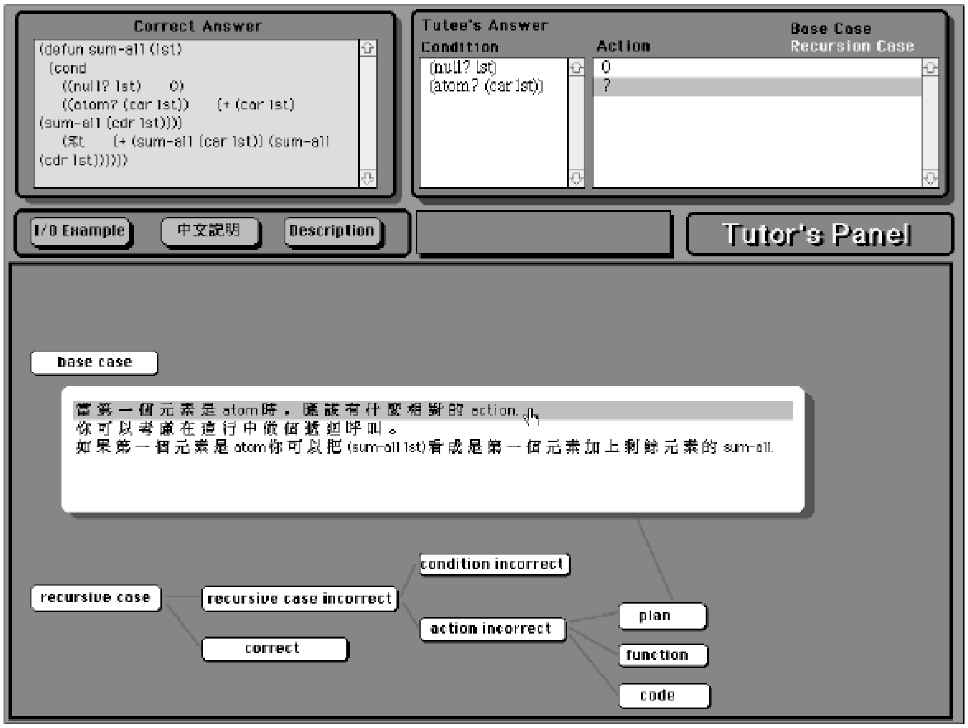


Figure 3.7 Interface of the DHT (Chan and Chou 1997)

Below were the descriptions of the tasks conducted by the students in the different conditions shown in Figure 3.6. The use of the PLS and the DHT tools within each condition was also specified.

Distributed-Responsibility Sharing: Three human students took turns being the designer, the tutor, and the translator. The designer generated solutions as LISP code using the PLS tool. Then, the solutions were passed to the human tutor for approval. Finally, the translator recoded the solution programs into LISP programs via another version of the PLS tool that had the help of the software-based tutor to reduce the syntactic errors.

Distributed-Reciprocal Tutoring: Two human students working on remotely-connected computers took turns being the tutor and the tutee. The tutee used the PLS tool to construct LISP programs, and the tutee used the DHT tool to help the tutee debug them. The human tutee received help from the virtual tutor in debugging and correcting the tutee's error.

Centralized Reciprocal Tutoring: The human student and the virtual peer took turns playing the tutor and the tutee roles with the help of the virtual tutor. Similar to the Distributed-Reciprocal Tutoring condition, the student used the PLS tool when taking the role of the tutee and the DHT tool when taking the role of the tutor.

Learning by Tutoring: The human student helped the virtual peer debug the peer's LISP program using the DHT tool with the help of the virtual tutor

Intelligent Tutoring System: The human students used the PLS tool to write LISP programs as solutions to the problems posted by the system. The virtual tutor diagnosed the programs and gave hints to correct the students' solutions.

Self-Diagnosing: The human student used the DHT tool to diagnose the programs that he built with the PLS tool to solve problems given by the system.

Working Alone: The human student used the PLS tool on his own.

Because the underlying architecture of systems for all conditions except for the *Working Alone* group was an intelligent tutoring system, these systems use the ITS components in selecting problems for students to work on. In addition, the systems that have the DHT tool also selected hints to be included at the leaf node to best correct the students' weaknesses in knowledge.

The experiment conducted on these systems was an archived proxy-pretest design (Trochim 2001). Thirty students from a LISP class were divided into six groups of five or six, and they were evenly distributed by their mid-term scores in class. The participants were asked to use the assigned systems for two hours to practice LISP recursion problems, and took the posttest with a LISP language interpreter. The grading procedure was not mentioned in the paper.

The grades for the posttest are shown in Table 3.1 ranking from the highest to the lowest. The best three scores were obtained for the systems involved sharing responsibility in developing solutions. The next three systems were those that asked students to learn for themselves. This indicated the possibly positive effect of sharing responsibility toward learning gains. However, another shared-responsibility group that was of interest to us, learning-by-tutoring, did not fit this pattern. This condition had the lowest posttest-performance. An explanation might be that the PLS, reported to have several benefits toward learning (Chan and Chou 1995), was present in every condition except the learning-by-tutoring one. This fact raised the question how much the PLS tool influenced the success of the other systems as compared to the learning-by-teaching system. Another observation is that the DHT tool may not be a good teaching tool because diagnosing somebody else's code requires a different set of skills than programming one's own code, and may not convey the same responsibility as teaching somebody to program.

Table 3.1 The Comparison of Seven Systems

Condition	*Posttest Score*	*Standard Deviation*
Distributed-Reciprocal Tutoring	87	14.69
Distributed-Responsibility Sharing	80	5.47
Centralized Reciprocal Tutoring	78	4.78
Intelligent Tutoring System	77	13.86
Self-Diagnosing	81	14.62
Working Alone	70	15.16
Learning by Tutoring	64	9.69

Discussions

Our survey of the literature indicates that only a small numbers of computer-based, learning-by-teaching systems have been developed and the studies conducted with them have been preliminary. It is clear from the analysis that three important factors uniquely characterize these systems, (i) explicit teaching, (ii) shared representations and (iii) shared responsibility.

First of all, all learning-by-teaching systems in this survey have explicit teaching processes even though some systems only give the appearance of learning. Students need to teach the LBT agents in order to make progress in the systems. In addition, these LBT agents have the fixed role of being the tutee. This is in contrast with pedagogical agents of other types, such as learning companions (Ramírez Uresti 1998), which normally take on the role of a peer but can switch to the role of a tutor.

Second, all experiments conducted on these learning-by-teaching systems demonstrate the positive influence of shared responsibility. (Obayashi, Shimoda, and Yoshikawa 2000) reported significantly higher performance in the group that taught virtual students than the group that studied for themselves. Also, students who taught the MCLS scored higher than those who learned for themselves (significant levels were not reported) (Michie, Paterson, and Hayes-Michie 1989). Even in (Chan and Chou 1997) that reported negative results on the learning-by-teaching system, the results indicated that lack of shared responsibility seemed to have a negative impact on students' performances.

An interesting observation is that, in many LBT systems, the pedagogical agents share responsibility in learning with students via explanation mechanisms, such as querying mechanisms in MCLS and DENISE, the question-answer mechanism in the Virtual Classroom, and the reverse Socratic interactions in DENISE and the system proposed by Palthepu, et al. In other words, these studies were conducted with systems that encouraged self-explanation.

In the virtual classroom, the virtual students repeated the user's answers. In the MCLS, students asked the system to show the rule it used for solving linear equations it had learned earlier from the user's examples. In DENISE, students could query the system about the links they had created. Also, in Palthepu et al's system, the system asked the user to explain further, and if a connection it made between two pieces of information is correct. These explanation activities were actually a form of self-explanation because the agent basically explained the knowledge that students had created. This indicates the importance that researchers in learning-by-teaching systems have given to self-explanation.

Finally, there is no pattern of selecting knowledge representations for learning-by-teaching systems. DENISE used a shared representation, but it had very limited view-ability through the dictionary and the query mechanism. MCLS had a viewable knowledge representation but it had a different structure from what the user taught the agent. The system presented in (Palthepu, Greer, and McCalla 1991) did not share its knowledge representation with students. In other words, its representation was neither viewable nor query-able. Even though the LBT agents in (Obayashi, Shimoda, and Yoshikawa 2000) did not have any knowledge representation, they appeared to share knowledge with students by repeating their previous answers.

The knowledge representations in DENISE and MCLS are the most interesting. Both systems defines more realistic LBT agents that are not only learn from what they been taught, but can also reason with that knowledge. Students taught DENISE single links and it used a qualitative reasoning algorithm to answer questions about the relation between two concepts that were not directly linked. MCLS derived a more complete rule for solving linear equations from examples provided by its user. This alerted the user of connections that can be made between pieces of information. As discussed in Chapter 2, this realization can lead to better learning and transfer than learning by rote.

Interestingly, only the system presented in (Obayashi, Shimoda, and Yoshikawa 2000) reported significant, positive results toward LBT agents even though these agents simply repeated what students typed in. This might be because the answers were shared and viewable. Students had direct impact on the agents because the agents gave exactly the same answers that were taught. The knowledge of the agents was also viewable to the students because it is exactly the text that they had explicitly typed in. Therefore, the viewable, shared representation can be another key in the success of learning by teaching. It can give the realistic feeling of the effects of teaching. In addition, such representations can also elicit self-explanation because they make reasoning explicit. Students can be more motivated when they have the direct control and can see the immediate effect of their activities on their tasks (Brasell 1987a, 1987b).

The need for viewable, shared visual representations is conformed by its relationship to self-explanation. In the virtual classroom, the users shared with their virtual agents their verbatim answers. MCLS shared the rule it had learned from the user via a visual representation, a tree, and students could examine the rule by asking the system to solve problems and explain the use of the rule in deriving the solution. DENISE has a similar mechanism that let students compose queries, but the study reported that students did not use this query feature. A possible answer is that students could not view what they have taught DENISE, and after teaching it a number of concepts and links it became difficult for students to recall the global structure. After some teaching, when users asked DENISE a question, there was no mechanism to help them understand how DENISE derived the answers. Therefore, the benefit of the query mechanism, a self-explanation mechanism or a debugging mechanism, was not apparent to the users.

In summary, in addition to supporting explicit teaching, learning-by-teaching systems should have components that support shared responsibility, shared representation, and self-explanation. The shared representation should also heighten the sense of shared responsibility of the student because the teachable agents apply this knowledge to solve problems, therefore, helping the student to reflect his own knowledge level.

These implications add more dimensions to the process of teaching, presented in Chapter 2, that is defined based on teachers who are experts in the domain and are trained in teaching. The process is also adjusted for one-on-one tutoring situations that are characteristics of intelligent learning systems. Another issue of importance is that the student teaching the agent lacks both domain knowledge and teaching experience. This implies the need for added support in the learning-by-teaching environment.

- *Decision-Making* Phase: In addition to having resources to study, the learner can be more motivated to share the responsibility of learning with his teachable agent from the start of the process.
- *Performing-Actions* Phase: This is where the learner performs the actual teaching. Our additional requirement is that the knowledge representation the learner uses to teach with is both explicit and visual, and the teachable agent also shares this representation. The main reason for this is discussed next.
- *Monitoring* Phase: Because the teachable agent shares a visual knowledge representation with the student, the student can see how his teaching affects the teachable agent's performance. Therefore, we can build internal and external assessment mechanisms around the knowledge-representation structure.

Recall the framework of effective learning environment also presented in Chapter 2. By applying that framework to implement this modified teaching-process, our teachable-agent environments now includes four categories of components:

- Learning methods that are both knowledge- and learner-centered by letting students develop their own learning plans based on a number of resources
- Teaching methods that are learner-centered by supporting explicit teaching via viewable, shared representation that lets students actively express their self-formed knowledge that becomes the teachable agent's knowledge
- Formative assessment features that support both internal and external assessments and also cognitive feedback in addition to outcome feedback
- Teaching activities as methods for collaboration that let the user share the responsibility in learning with the teachable agent

As mentioned previously, there are two types of formative assessment, outcome (correct or incorrect) and cognitive (related to the status of knowledge and the process of learning) feedback. Cognitive feedback is more effective in supporting learning and problem solving than outcome feedback (Butler and Winne 1995). Therefore, the formative feedback mechanism in an effective learning-by-teaching system must provide cognitive feedback and encourage students to access this type of mechanisms more than outcome-feedback Our design of the Betty's Brain environment, presented in the next chapter, adopts these four principles.

CHAPTER IV

A TEACHABLE AGENT ENVIRONMENT: BETTY'S BRAIN

This chapter describes the first version of the **Betty's Brain** environment and the study the Teachable Agents Group (TAG) at Vanderbilt University conducted with the system in spring 2002. (See (Leelawong et al. 2001) for earlier prototypes.) The environment was jointly designed and implemented by the TAG members, and I was the primary designer and implementer within this group. Components of the environment were evaluated in terms of how they supported the learning-by-teaching process. The effectiveness of the shared representation as a means for communicating knowledge between the student and the teachable agent was also studied. These evaluations led to the design and implementation of an improved version of the Betty's Brain system, which is the main focus of this dissertation.

It is important to differentiate our **Teachable Agent** (TA), Betty, in the Betty's Brain environment from software student agents used in previous learning-by-teaching systems (Michie, Paterson, and Hayes-Michie 1989; Palthepu, Greer, and McCalla 1991; Nichols 1994; Chan and Chou 1997; Obayashi, Shimoda, and Yoshikawa 2000). The primary difference is the *viewable*, *shared representation* that human students use to directly teach the agents. This shared representation can take on a variety of standard knowledge representation formats, such as concept maps, graphs, logical expressions, and statistical distributions. The teachable agents behave like perfect students, and this simplifies the teaching task for their novice teachers, i.e., the students. The agents do not forget, change, or induce new knowledge from what is explicitly taught to them. In other words, unlike a number of other agent-based systems, teachable agents are not endowed with machine-learning algorithms that learn from examples, explanations, or by induction. By using systematic and well-defined representation schemes, teachable agents can reason with their taught knowledge to answer questions and explain how they derive their answers. Finally, because teachable agents do not possess prior knowledge, they cannot evaluate their knowledge for coverage of domain content.

The teachable agent in Betty's Brain environment, **Betty Bashinal**, learns and reasons using a version of the concept map that focuses on how entities in the domain relate to one another by cause-effect relationships (Tolman 1948; Axelrod 1976). Betty is introduced to middle-school students as an enthusiastic seventh-grader who would like to join a nearby high-school science club. She wants to learn about river ecosystems and use that knowledge to save the river that runs by her house. She suspects that some form of pollution is killing off the fish in the river. The high-school science teacher is reluctant to accept her. He can see that she is enthusiastic, but he is not sure she has the maturity or knowledge to participate in the science club projects. So, he makes a deal. He will let her join the club if she can pass a test on river ecosystems. At this point, the students are told to teach Betty about river ecosystems to help her pass the science test.

The Betty's Brain environment provides a number of tools that assist students in their teaching tasks. Students have resources they can read as they prepare to teach. They can teach Betty using a graphical interface. In addition, they can monitor Betty's and their own progress by asking her queries and sending her to take quizzes.

The first part of this chapter, *Design Principles*, explains the design foundations for teachable-agent systems. These foundations form the six primary components of the Betty's Brain environment that are described in the *Betty's Brain* section. Then, the *Study* section, describes an experimental study conducted in a fifth-grader classroom in the spring of 2002 to test the effectiveness of the different components of the environment and their effects on the quality of students' concept maps. The *Results and Discussions* section reports and analyzes the findings and uses the analysis to suggest improvements to the Betty's Brain environment. This improved system is described in Chapter 5.

Design Principles

The design of the Betty's Brain environment was influenced by three important considerations—(i) the ability to model all phases of the *teaching process* (Colton and Sparks-Langer 1993), (ii) the HPL framework (Bransford, Brown, and Cocking 2000), and (iii) an explicit, visual, and shared domain-knowledge representation and corresponding reasoning mechanisms. This last consideration ensured that students could teach Betty with low programming overhead and could work together to answer questions related to domain concepts. Recall that Chapter 2 discusses the teaching process and the HPL framework in detail. This section discusses how the principles from these two frameworks are translated into the design and implementation of the system while satisfying the requirements of low programming overhead, shared representation, and shared responsibility.

Implementing the Teaching Process in Betty's Brain

One of the design principles of the Betty's Brain system was to ensure that all three distinct phases of the teaching process (Colton and Sparks-Langer 1993), (i) decision making (preparing to teach), (ii) performing actions (teaching and interacting with the student), and (iii) monitoring (assessing and reflecting on) what the agent has learned, were explicit and visible components of the environment. Since our target participants are middle school students who lack teaching experience as well as domain knowledge in the particular field of study, the system had to provide additional support and scaffolds to help students through each of the stages of their teaching and learning process.

To help students learn as they teach, resources are made available and cover all the relevant knowledge content in the domain of study. In the teaching phase, students can teach and query the teachable agent. Each student demonstrates his understanding of the information by creating a concept map to teach Betty. During this phase, students can also evaluate how much Betty has learned by asking her specific questions, requesting her to explain her answers, and reflecting on the answer and the explanations provided. Finally, in the monitoring phase, students can ask the teachable agent to take quizzes that are provided by the mentor agent. They can use the quiz performance to reflect on how much their agent, and they themselves have learned. In addition, they can use the mentor agent's feedback on the quiz results to assess what they should do to improve the teachable agent's performance.

The graphical user interface allows students to move seamlessly from one phase to another as they perform their tasks of learning, teaching the agent, and monitoring and reflecting on the agent's and their own performance (see in Figure 4.1). In addition, it plays the role of a visual programming language that alleviates the students' programming burdens in creating the concept map structure to teach their agent. The graphical user interface has been successful in keeping our students' primary focus on gaining knowledge, teaching the agent, and reflecting on their own learning rather than worrying them about programming details.

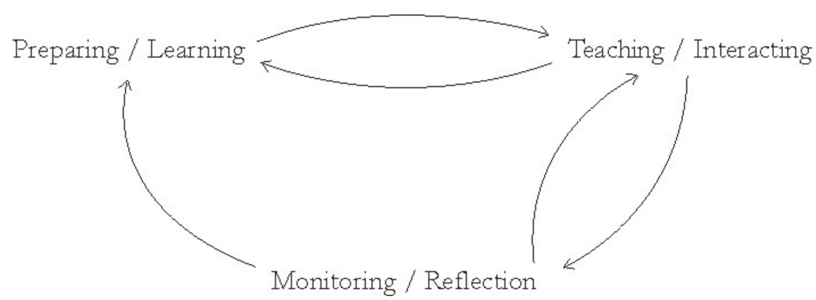


Figure 4.1 Flow within the Teaching Process in the Betty's Brain Environment

Applying the HPL framework

The second goal of our design is to take into consideration the HPL framework (Bransford, Brown, and Cocking 2000) that combines the learner-centered, knowledge-centered, assessment-centered, and community-centered aspects to define effective learning environments. The Betty's Brain environment incorporates the learner-centered aspect of the HPL framework by giving students the flexibility to teach their agents at their own pace on the topics they choose as well as to apply their individual styles of learning to acquire information from the text resources.

The environment is knowledge-centered because it focuses students on gaining domain knowledge and expressing it in an organized form to teach the agent in a consistent and explicit manner. This is achieved by using a causal-map structure as a shared representation that allows the student and the agent to communicate effectively. In addition, the environment provides a number of on-demand resources in the particular domain of study.

Assessment-centered issues, especially formative assessment, are addressed by letting students query their agents and evaluate their responses. Furthermore, students can also send their agents to take quizzes and receive feedback on the quiz performance.

Finally, the environment supports a community-centered aspect to some extent because one-on-one teaching (tutoring) is a social activity outside the traditional classroom environment between two agents, in this case, the student and Betty. True community-centered learning environments would include discussions about the system and river ecosystems that students may have with peers, parents, and others. However, that is beyond the scope of this thesis work.

Domain-Knowledge Representation

In a real-world teaching environment, teachers need to communicate lessons to their students. Instead of verbal communication, in Betty's Brain students teach the teachable agent using a visual adaptation of a concept map (Novak 1964) where concepts represent objects (such as "fish") or ideas (such as "overgrown plants") and links represent causal-effect relationships between concepts. Concept mapping is suitable to the Betty's Brain environment because it enables (Biswas et al. 2004):

1. Visible and organized knowledge structures, which aid constructivist learning by allowing explicit integration of existing and new knowledge, enhancing assessment of knowledge and diagnosis of misconceptions (Novak 1996)
2. Low-cost domain-knowledge programming to teach the agent
3. Teaching interactions that benefit the tutor, such as interactions which involve asking and answering questions (Chi et al. 2001)

Several researchers have discussed the effectiveness of traditional concept maps in promoting learning in scientific domains (Kinchin and Hay 2000; Stoyanov and Kommers 1999; Novak 1996). The concept map provides a mechanism for structuring and organizing knowledge into hierarchies and allows for analysis of phenomena as cause-effect relations. Hence, the concept map is a powerful tool that lets students represent their understanding of domain knowledge in a well-organized format (Kinchin and Hay 2000). By making thoughts visible, concept map structures can provide a framework for reflection and revision of one's knowledge with the goal of improving problem-solving performance. Moreover, an intelligent software-agent that uses this concept-map structure with reasoning and explanation mechanisms provides executable mechanisms (as opposed to static structures) that help students analyze the effect of a change in one concept on others, such as "if fish increases, what happens to dissolved oxygen?"

It is important to make clear that, even though this adaptation of concept mapping is similar to causal mapping (Axelrod 1976), the Betty's Brain environment has its own version of causal reasoning. Its reasoning procedure is adapted and simplified from the actual qualitative reasoning process to encourage learning and reflection in young students rather than to accurately model qualitative effects.

The Components of Betty's Brain

The Betty's Brain environment has been designed to include three categories of components that correspond to the three phases of teaching as shown in Figure 4.2.

The **Resources** are on-line text-based materials organized as book chapters that contain information about river ecosystems. Students are expected to read the resources and extract the important concepts and their relations from the text to teach them to Betty.

The **Concept Map Editor** is a graphical tool with a point-and-click interface and a number of function buttons that students use to construct and modify their concept maps.

The **Query Mechanism**, a self-assessment feature, is a template-based interface with pull-down menus that let students compose questions for the teachable agent. The query mechanism assists both the teaching and assessment phases since it allows teaching interactions that assess Betty's current performance based on what she has learned. This helps students reflect on their own knowledge and understanding.

The **Reasoning Mechanisms** implement qualitative-reasoning algorithms that the teachable agent employs to answer (i) questions that students compose using the query mechanism and (ii) quiz questions that students select via the quiz mechanism.

The **Explanation Mechanisms** are algorithms that Betty uses to explain her process of deriving an answer to a query in a step-by-step fashion.

The **Quiz Mechanism**, a formative-assessment feature, is a set of predefined questions that students use to evaluate Betty's and their own knowledge. They then receive (i) outcome feedback and (ii) specific feedback on how to improve the concept map.

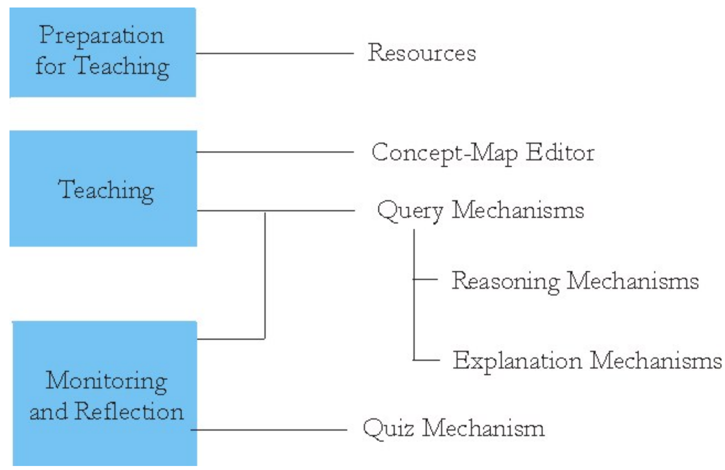


Figure 4.2 Components for Teaching Phases

Figure 4.3 illustrates the Betty's Brain environment. The figure labels the main, graphical components of the system.

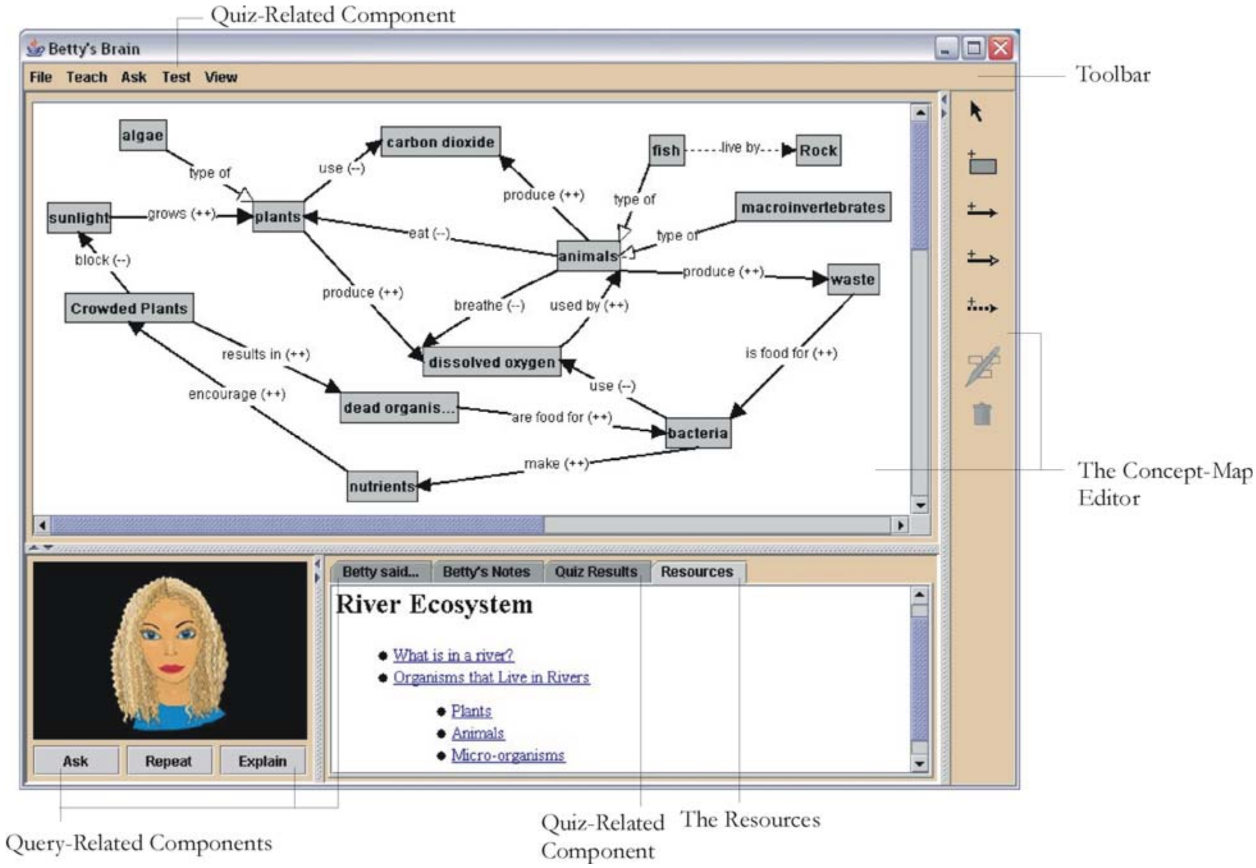


Figure 4.3 Betty's Brain

The Betty's Brain Environment

We have implemented Betty's Brain environment in Java (Java 2 SDK v1.3.8) with a Microsoft's text-to-speech engine (Microsoft Speech SDK v5.1) in conjunction with a Java-Speech Application-Program-Interface for Windows platforms (Cloud Garden v1.3).

The Concept Map Editor

Novak (1996) defines a **concept map**, a collection of concepts and links between these concepts, as a mechanism for representing domain knowledge. In the Betty's Brain environment, **concepts** are entities that are of interest in the domain of study. For example, common concepts in a river ecosystem are fish, plants, bacteria, dissolved oxygen, carbon dioxide, algae, and waste. Links are unidirectional, binary relationships between two concepts. They help express interactions among concepts. For example, information that fish breathe dissolved oxygen can be translated into a link from the concept *fish* to the concept *dissolved-oxygen*.

In the Betty's Brain environment, the student can teach Betty three kinds of links in a concept map, (i) **causal** (cause-and-effect), (ii) **type-of** (hierarchical), and (iii) **descriptive** links. The causal link (Axelrod 1976) specifies an active relationship on how a change in the originating concept affects the destination concept. Two examples of this type of link are "Animals inhale Dissolved Oxygen" and "Plants produce Dissolved oxygen." The causal links are further qualified by **increase** ('+') and **decrease** ('–') labels. For example, "inhale" implies a decrease relation and "produce" an increase. Therefore, an introduction of more animals into the river ecosystem causes a decrease in dissolved oxygen, but an increase in plants causes an increase in dissolved oxygen. The causal relation can be refined further by specifying the degree of change (increase or decrease) of a causal link as **small** ('–', '+'), **normal** ('– –', '++'), or **large** ('– – –', '+++'). To reduce the complexity for middle-school students, this feature has been disabled in the construction of the concept map. In other words, we have simplified the system so that students have to deal with only one degree of change for links in the concept maps, namely "normal increase" and "normal decrease". The refined degrees of changes are used in the causal reasoning mechanism, which is described in a later section.

When there is a two-way relationship between a pair of concepts, the student specifies two causal links, one forward and another backward. For example, the student can capture the bi-directional relationship between animals and dissolved oxygen as two links, (i) "Animals *inhale (–)* Dissolved Oxygen" and its reverse effect, (ii) "Dissolved Oxygen *helps (+)* Animals." The rationale for the forward link is that more animals in the ecosystem decrease the amount of dissolved oxygen and fewer animals imply more dissolved oxygen. The backward link captures the relation that more dissolved oxygen facilitates an increase in the amount of animals, whereas a decrease in dissolved oxygen implies less essential life support; therefore, the number of fish is likely to reduce.

Hierarchical links let students establish class structures to organize the domain knowledge. Consider an example where students deal with a variety of fish, such as trout, bass, blue gill, and catfish. All of these fish types breathe dissolved oxygen and nibble on plants. To simplify the knowledge construction process, the student first creates the generic concept "Fish" and expresses the links "Fish eat Plants" and "Fish breathe Dissolved oxygen." Then the student can create additional concepts, such as "trout" and "bass," and link them to the "Fish" concept using "type of" links. This way, the individual types of fish inherit all relations associated with the "Fish" concept unless they are over-ridden by more specific links (Russell and Norvig 1995).

The "descriptive*"* link is used for information that a student considers to be relevant but not essential to reasoning about changes in the domain. For example, a river ecosystem normally contains animals, such as fish and macro-invertebrates. The descriptive link, "Animals *live in* Rivers," can represent this information. In addition, descriptive links are used for information that does not fit the other two link types described above. An example is "Fish live by Rocks." Fish do not have any effect on the amount of rocks and vice versa.

To keep the semantics of the representation structure simple, all relationships in our environment are binary. Other more complex forms of relations, such as conditional causal relationships, have not been implemented in this version of Betty's Brain. The alternative in the Betty's Brain environment is to model such a phenomenon with two links instead of one. For example, the relationship "Plants need Sunlight to produce Dissolved Oxygen" is represented by the links "Sunlight grows (+) Plants" and "Plants produce (+) Dissolved Oxygen."

The two panels at the top-right of the environment shown in Figure 4.3 form the Concept Map Editor. The drawing canvas illustrates a concept map in the domain of river ecosystems. The arrow indicates the direction of the link. Each class of links has a different visual representation as shown in Figure 4.4. The labeled boxes correspond to entities (the labels are entity names) and the labeled links relations. Initially, this is an empty white panel indicating that Betty has no initial knowledge of the domain. As students create concepts and links the drawing canvas depicts the current concept-map structure.

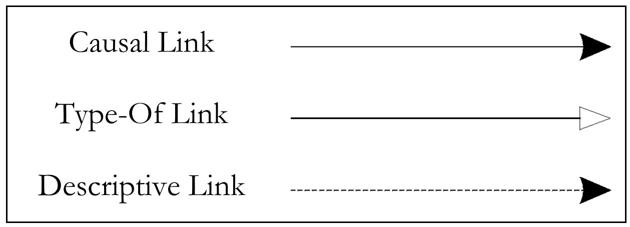


Figure 4.4 Visual Representation of Three Types of Links in Betty's Brain

Students can add, modify, and delete concepts and links using a visual, drag-and-drop editor. Figure 4.5 shows the set of buttons used to create, modify, and delete concepts and links. The icons and their descriptions are shown in the figure. The concept-map-editor control can also be accessed from the toolbar on the top of the Betty's Brain environment. When the student clicks on the *Teach* menu item, the teaching-activity drop-down menu appears as shown in Figure 4.6. This menu contains the same set of icons as the control panel shown in Figure 4.5 except the pointer.

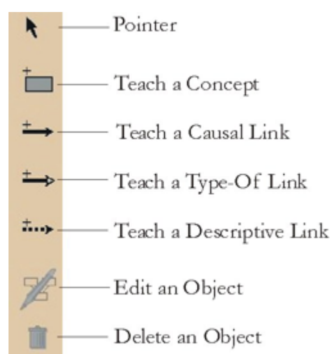


Figure 4.5 The Control Panel of the Concept Map Editor

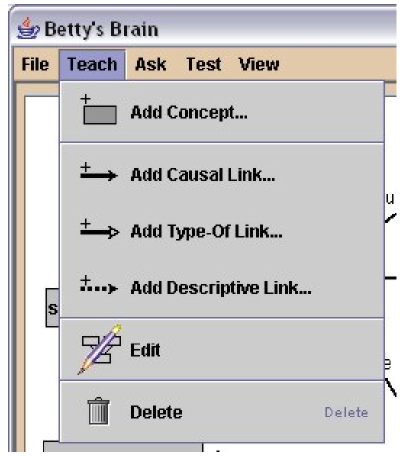


Figure 4.6 The Teaching-Activity Menu

To add a concept, the student can either click on the "Add Concept" icon () on the control panel and then in the map, or double-click where the concept should be added. An editable box then appears, as shown in Figure 4.7, and the student can enter the name of the concept in this box.

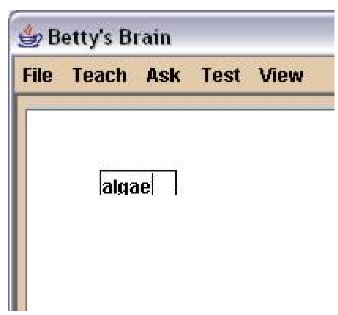


Figure 4.7 Adding a Concept

To add a link between a pair of concepts, the student first chooses the type of link ("causal" , "type-of" , and "descriptive"), and then clicks on the concept box where the link originates. When the link appears on the screen, the student can stretch it by dragging the mouse pointer to the destination concept-box. The green-woven line indicating this new link appears simultaneously with an appropriate dialog box where the student enters information about the link. An example dialog-box for a casual link is displayed in Figure 4.8.

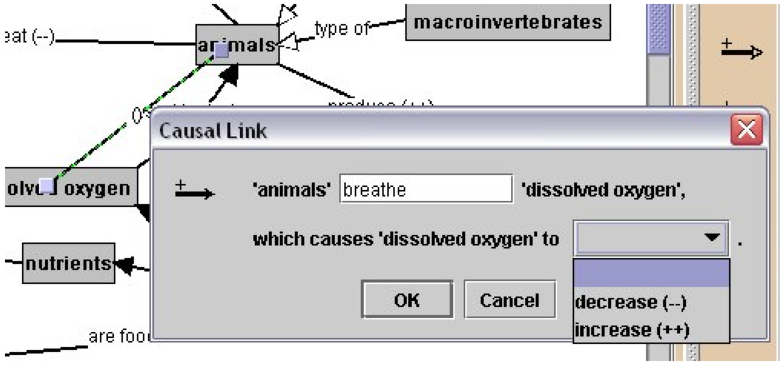


Figure 4.8 Adding a Causal Link

The addition of type-of links requires no dialog box. The addition of descriptive links is similar to that of casual links except for the fact that it does not require trend information as shown in Figure 4.9.

Figure 4.9 Add-a-Descriptive-link Dialog

Once the student has taught Betty a concept map that has one or more links, he or she can ask Betty questions using the query mechanism described in the next section.

Query Mechanism

In a teaching setting, a tutor typically asks the tutee questions during the teaching process to assess the tutee's understanding of the materials. Therefore, the Betty's Brain environment supports this interaction through the query mechanism. This interaction is intended to give students a chance to revisit and reflect on the knowledge structures they have created and to improve them when they seem appropriate. The teachable agent, Betty, uses reasoning mechanisms that allow her to analyze the concept map in order to answer domain-relevant questions posted by her tutor. Students can observe Betty's answers and query Betty further to get more detailed explanations of how she generated an answer.

The Betty's Brain environment provides students with two types of question templates (displayed in Figure 4.10):

1. Reasoning about causes and effects: "*If a CONCEPT1 CHANGES, what happens to another CONCEPT2?*"
2. Descriptive Explanation: "*Tell me about CONCEPT1.*"

CONCEPT1 and *CONCEPT2* are concepts that the student has taught Betty, and *CHANGES* can be the value "increase" or "decrease." The first question template enables students to learn how Betty reasons through chains of causal effects in the concept map. This helps students investigate how a change in one concept is propagated to and affects other concepts. In this process, students can reflect on whether they have taught her correctly. The second type of question lists the link between a concept and adjacent concepts. The rest of this dissertation refers to these two types of questions as **causal** (for the category of reasoning about causes and effects) and **tell-me** (for the general explanation category) for briefness.

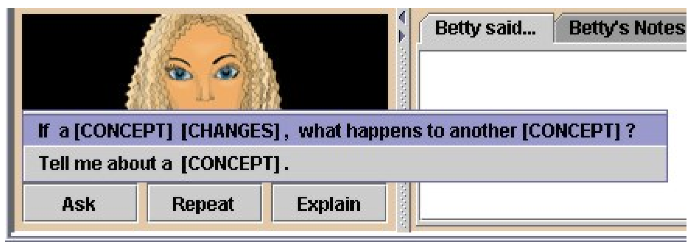


Figure 4.10 Query Dialog

To ask a tell-me question, the student clicks on the item "*Tell me about CONCEPT*" from the pop-up menu shown in Figure 4.10. The dialog box, depicted in Figure 4.11, appears and the student can select one of the concept names from his concept map from an alphabetical list that appears when clicking on the down arrow at the end of the dialog line. The use of the pull-down menu reduces the student's typing effort and chance for typing errors. Clicking on the "*OK*" button at the bottom of the dialog box sends this question to Betty.

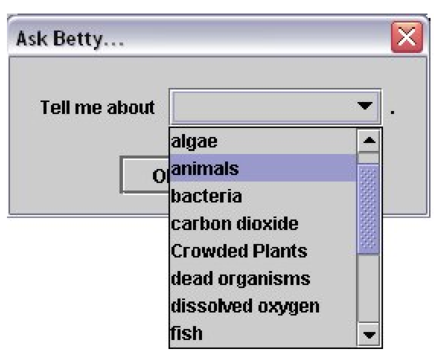


Figure 4.11 Asking a Tell-Me Question

The student asks causal questions by selecting the template "*If a CONCEPT1 CHANGES, what happens to CONCEPT2?*" from the pop-up menu shown in Figure 4.10. The dialog box, depicted in Figure 4.12, appears, and the student can select the source concept (*CONCEPT1*), the triggering effect (*CHANGES;* increase or decrease) of the source concept, and the target concept (*CONCEPT2*). Similar to the control used in the tell-me dialog, the student can click on the down arrows to select the choice for each field. Students click on the "OK" button to ask Betty the query they have just formulated. If one or more fields are not filled in, an error message appears on the screen.

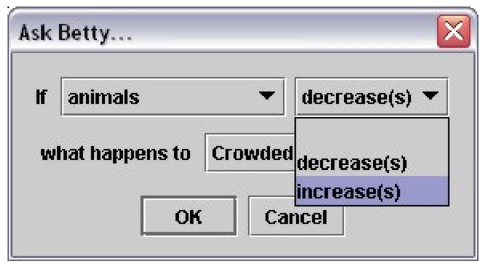


Figure 4.12 The Causal Query Dialog Box

Once Betty is asked a question, she reasons with her concept map to find an answer. The next section describes the reasoning mechanisms that Betty uses to answer causal and tell-me questions.

Reasoning Mechanisms

Betty uses two types of reasoning mechanisms to generate answers to questions that students formulate—one for answering casual questions and the second for tell-me questions. To answer a "tell-me" question, Betty searches for all adjacent concepts that are linked by both incoming and outgoing links to the chosen concept. For example, given the concept map in Figure 4.3 for the query "Tell me about *animals,*" the part of the concept map that Betty uses to answer this question is shown in Figure 4.13. How Betty communicates this answer to the student is explained in the next section, *Explanation Mechanisms*.

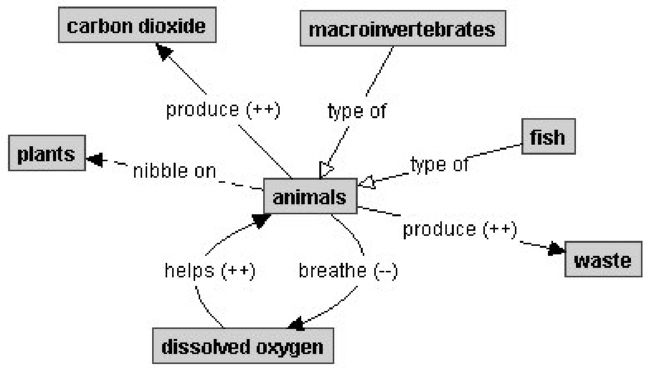


Figure 4.13 A Reasoning Graph for a Tell-Me Question

The reasoning mechanism for causal questions uses a qualitative-reasoning procedure to propagate a change in an originating concept though a sequence of connected causal links to determine the effect on the target concept. Because the current version of the concept map does not include temporal information, the reasoning mechanism does not accommodate causal links that form loops. To derive the effect of a change (either an *increase* or a *decrease*) in Concept *A* on Concept *B*, the reasoning engine performs the following steps:

1. Generate an acyclic, directed sub-graph with causal links that connect the change in Concept *A* to Concept *B* using a modified depth-first-search algorithm (Russell and Norvig 1995) that is modified to continue until each link in the forward direction (such as from concept A to concept B) is visited only once. Therefore, paths that include loop structures are deliberately excluded.
2. The flow of reasoning follows the direction of the arrows of the causal links. Propagate the change from Concept A along all outgoing links using the pair wise propagation table shown in Figure 4.14. A "+" represents an increase effect and a "–" a decrease effect. For finer-grained calculations, to avoid inherent ambiguity in qualitative reasoning, increases and decreases are further qualified by "S" and "L" subscripts to represent small and large changes, respectively. From Figure 4.14, one can infer that a large decrease combined with a small increase results in a normal decrease in Concept B (row 1, column 3).
3. Repeat Step 2 using the target concepts in the last step as the source concepts for this step, and propagate forward as long as there is only one incoming link to each source concept or the goal node is not reached. For the case that a source concept has more than one incoming link, wait until all the effects have accumulated at that node before going to Step 4. Note that the reasoning mechanism cannot get into an infinite loop because cycles are not considered in forming all of the forward paths. The reasoning process explicitly ignores paths that form loop structures. In other words, the reasoning progresses only along forward paths from the source concept (*A*) to the target concept (*B*).
4. For nodes that have to aggregate information from exactly two incoming links, the pair-wise combination table shown in Figure 4.15 is employed to generate the aggregated effect on the node. For more than two incoming links, we first create pair of the same level of opposite effects (such as a small increase and a

small decrease), and if they are equal in number they cancel each other, implying that there is no change. The rest of the effects are them combined in a pair-wise manner.

Change in Relation

Change in Entity	$+_L$	$+$	$+_S$	$-_S$	$-$	$-_L$
$+_L$	$+_L$	$+_L$	$+$	$-$	$-_L$	$-_L$
$+$	$+_L$	$+$	$+_S$	$-_S$	$-$	$-_L$
$+_S$	$+$	$+_S$	$+_S$	$-_S$	$-_S$	$-$
$-_S$	$-$	$-_S$	$-_S$	$+_S$	$+_S$	$+$
$-$	$-_L$	$-$	$-_S$	$+_S$	$+$	$+_L$
$-_L$	$-_L$	$-_L$	$-$	$+$	$+_L$	$+_L$

Figure 4.14 The Pair wise Effects

	$+_L$	$+$	$+_S$	$-_S$	$-$	$-_L$
$+_L$	$+_L$	$+_L$	$+_L$	$+$	$+_S$	0
$+$	$+_L$	$+_L$	$+$	$+_S$	0	$-_S$
$+_S$	$+_L$	$+$	$+$	0	$-_S$	$-$
$-_S$	$+$	$+_S$	0	$-$	$-$	$-_L$
$-$	$+_S$	0	$-_S$	$-$	$-_L$	$-_L$
$-_L$	0	$-_S$	$-$	$-_L$	$-_L$	$-_L$

Figure 4.15 Integrating Results from Two Paths

In general, opposite effects in a qualitative reasoning scheme (Kuipers 1994) implies ambiguous effects. In this work, to simplify the scheme for middle-school students, we avoid ambiguous effects. We have carefully checked the expert concept map for the task to ensure that this simplistic scheme does not produce incorrect results.

To further explain the reasoning mechanism, the question "If *bacteria* increase, what happens to *animals*?" was applied to the concept map shown in Figure 4.3.

1. The reasoning sub-graph generated by step 1 of the algorithm is displayed in Figure 4.16. The dotted links represent an example of a path that is removed because it creates a loop back to *bacteria*.
2. Propagate the change in *bacteria* (increase) forward through the network. Follow the link from *bacteria* to *dissolved oxygen*, as shown in Figure 4.17. The propagation cannot continue at this time because there are two incoming links to dissolved oxygen.
3. Another path is followed from the source concept, as shown in Figure 4.18.
4. The result from combining the two incoming paths to *dissolved oxygen* is a large decrease, as shown in Figure 4.19. Now we can propagate the effect forward to the target concept. The answer is a large decrease in *animals*.

Betty communicates her answer to the query and her explanation for how she derives her answer using text, speech, and animation. This is described in the next section, *Explanation Mechanisms*. The explanation mechanisms are developed to help students understand how Betty applies her reasoning processes to answer questions.

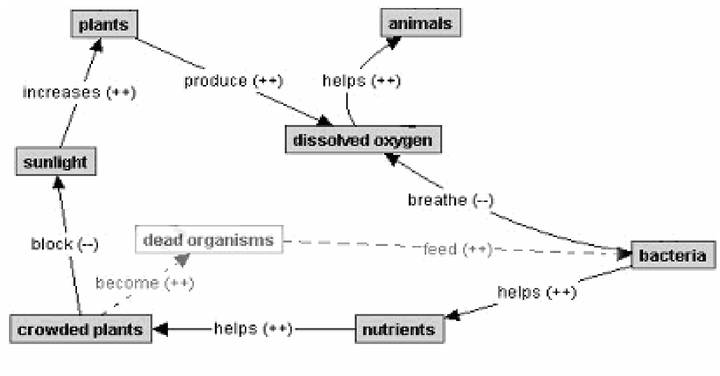


Figure 4.16 The Resulting Graph with the Removed Path indicated with Dotted Lines

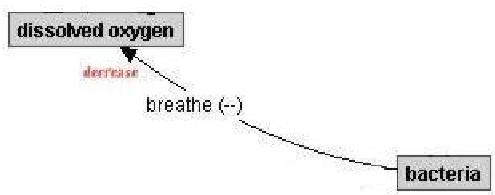


Figure 4.17 Propagation of Causal Effects: The Direct Path from Bacteria to Dissolved Oxygen

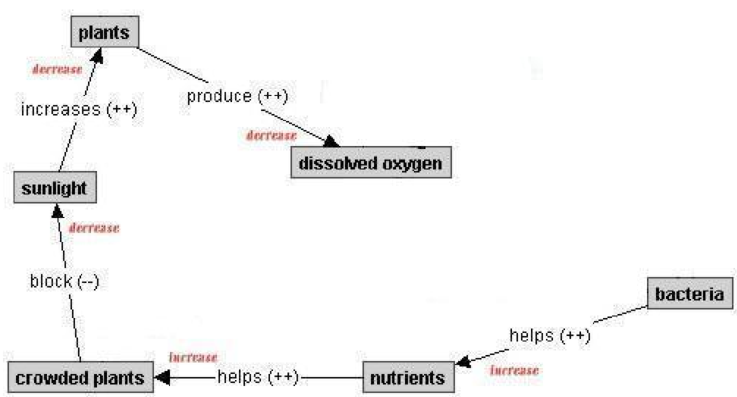


Figure 4.18 Propagation of Causal Effects: The Indirect Path from Bacteria to Dissolved Oxygen

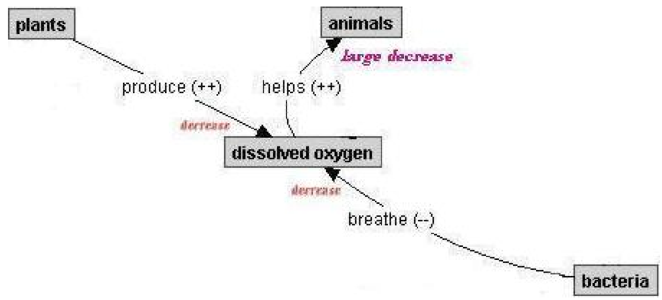


Figure 4.19 Propagation of Causal Effects: Combining the Results

Explanation Mechanisms

Betty uses a combination of visual representations and simple dialog-structures to explain her reasoning processes. As she proceeds through the explanations, she uses animation to display the process step-by-step on the concept map. The explanation is also saved in the "*Betty said…*" tab on the lower-right panel of the interface. This allows students to refer back to explanations that Betty has previously generated. Students can also click on the "Repeat" button in the lower-left panel to hear the explanations Betty has just spoken again.

We first present the algorithm for generating explanations for tell-me questions. The tell-me explanation is based on the information of links (the source concept's label, the link label, and the target concept's label) that are associated with the concept regardless of any causal information. Figure 4.20 illustrates the concept map that is used to answer the query "Tell me about *animals*." Betty reads and highlights her answer link by link, such as "Fish are a type of animals" and "Animals produce waste." If Betty cannot generate any answer to a query, she reports to the student that the query concept is not linked to any other concepts though any link.

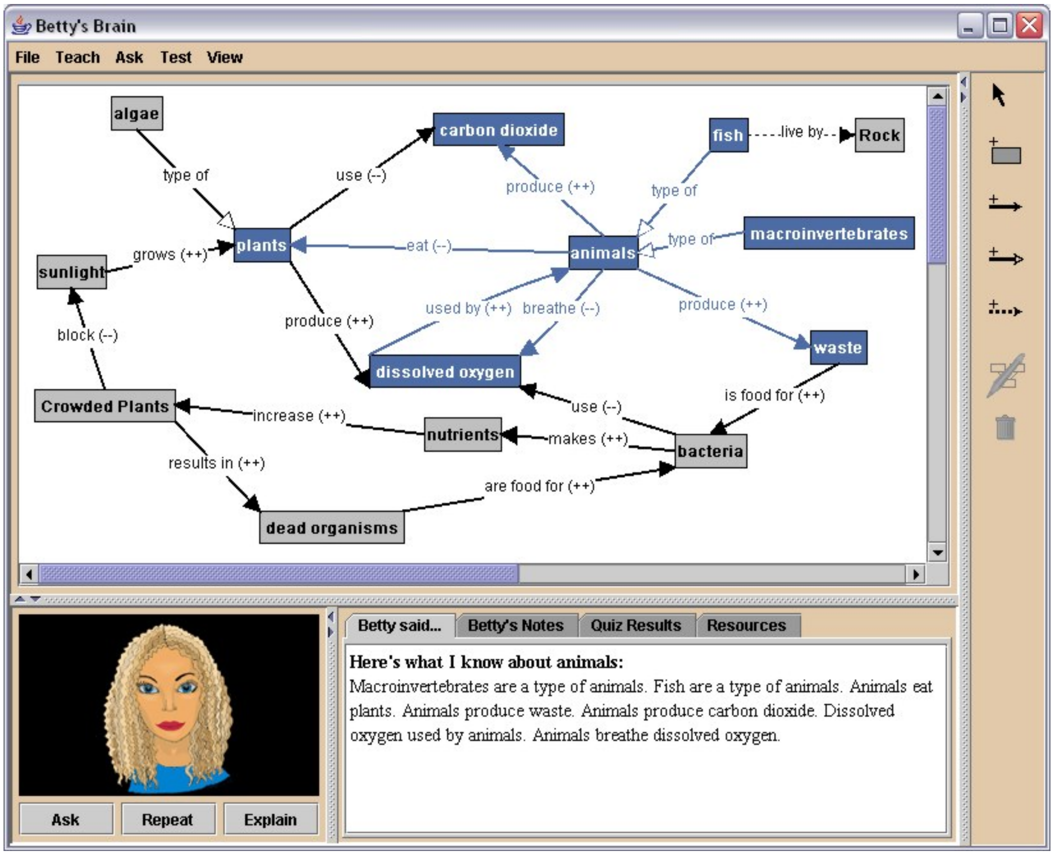


Figure 4.20 Answer to a Tell-Me Question

To scaffold the causal explanation, Betty splits long explanations into chains of events that start from a general overview of the result to a detailed link-by-link explanation. The algorithm for generating explanations for answers to causal questions is presented below. The answer graph, such as the one in Figure 4.16, is organized into three different data structures to build the explanation: path, block, and abstracted path. These structures are illustrated in Figure 4.21. A **path** is a series of connected links with no branching or merging points. A **block** is a set of multiple structures that connects one concept to another. A block must start with a concept with two or more outgoing answer-links and end with another concept with multiple incoming answer-links. An **abstract path** is a path that has blocks as its nodes, and it must either start or end with a normal path. Every structure has three components—the beginning and the end nodes and an internal data structure.

1. Start from the source concept.
    a. Initiate a path if the concept has only one outgoing link in the answer structure. Put the link in the path.
    b. Otherwise, initiate a block with as many paths as the number of the outgoing answer links.
    c. Make the source concept as the beginning node and the target concept the end node of the structure because the explanation is structured to explain how the source concept affects the target concept.

Note that the current structure may change when we traverse deeper in the graph (such as converting a path to an abstract path when encountering a block).

2. For each path, use the target concept of the last link in the path as the current node.
   a. If this node has more than one incoming link in the answer graph, check if the number of structures in the block is equal to the number of links. (In this case, the current path was a part of a block.):
      i. If equal, set the end node of this block to this concept and proceed to Step 2.b.
      ii. If the first number is less, set the end node of the path to be this concept and traverse back to the start node of the block, and repeat this step. Like the reasoning algorithm, we wait until the explanation is constructed for the whole block before proceeding toward the destination concept.
   b. If this node has one outgoing link in the answer graph:
      i. If this node is not the end of a block, append that link to the end of the path.
      ii. Otherwise, initiate another path and put this link in. If the block is already part of an abstract path, simply append this new path to it. Otherwise, convert this block into an abstract path and append the new path to its end.
   c. Otherwise, this node has two or more outgoing links in the answer graph:
      i. Initiate a block that starts with this node.
      ii. Initiate as many paths as the number of the links, and put each link in one path.
      iii. Convert the current path into an abstract path by having the current path as the start structure following by the block. Repeat Step 2 for each path.

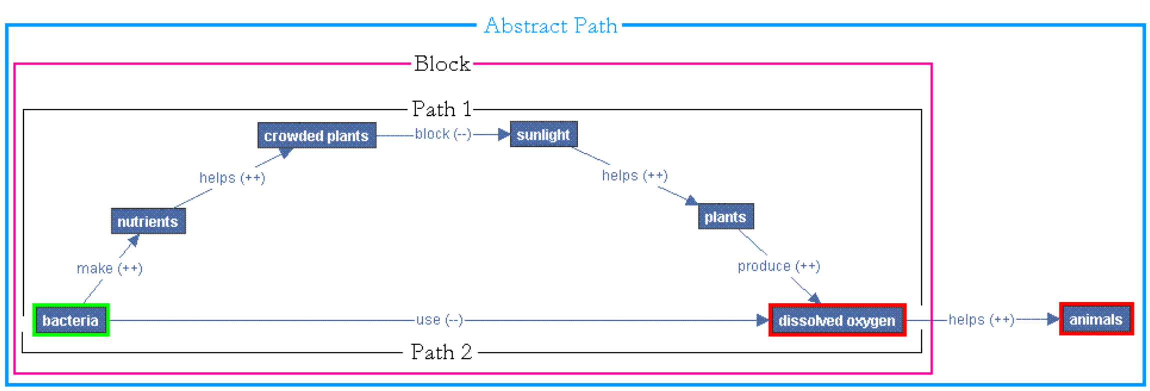


Figure 4.21 An Answer Labeled with Explanation Structures

By organizing the reasoning graph into sub-structures with blocks, the explanation can be divided into a set of aggregate paths. The explanation structure provides a top-down view with the source to target effect being presented first, and then details within each aggregated block are presented in a step-by-step fashion. We illustrate the working of the causal explanation-mechanism with an example. Assume that the question is, "If *bacteria* increase, what happens to *animals*?" The sub-graph extracted from the concept map to answer this question is displayed in Figure 4.16. Betty first gives a verbal answer to the question, and the text of her response appears in the "Betty said…" tab at the bottom of the environment. Betty also animates parts of the concept map corresponding to the particular part of the explanation. In this case, it is the source and the target concepts.

After Betty answers the question, the students can ask Betty to explain her answer by clicking on the "Explain" button. Using the algorithm described previously, Betty structures her explanation from the answer graph, illustrated in Figure 4.16, by propagating along two paths, then combining information from the paths, and showing how the aggregate conclusion is derived. The list below contains the steps of explanation. The thicker path in Figure 4.22 - Figure 4.26 shows the part of the concept map animated when Betty communicates that step of the explanation, and the text she speaks in the step is appended to the "Betty said…" tab at the bottom of each image.

Aggregate (Top) Step (Figure 4.22): Final result: An increase in *bacteria* results in a large decrease in *animals*. Betty's verbal explanation in this step is, "*I think that if bacteria increase, animals decrease a lot. These are the paths I followed to get my answer. If you want me to explain my answer, click Explain. To see my notes, click Notes.*"

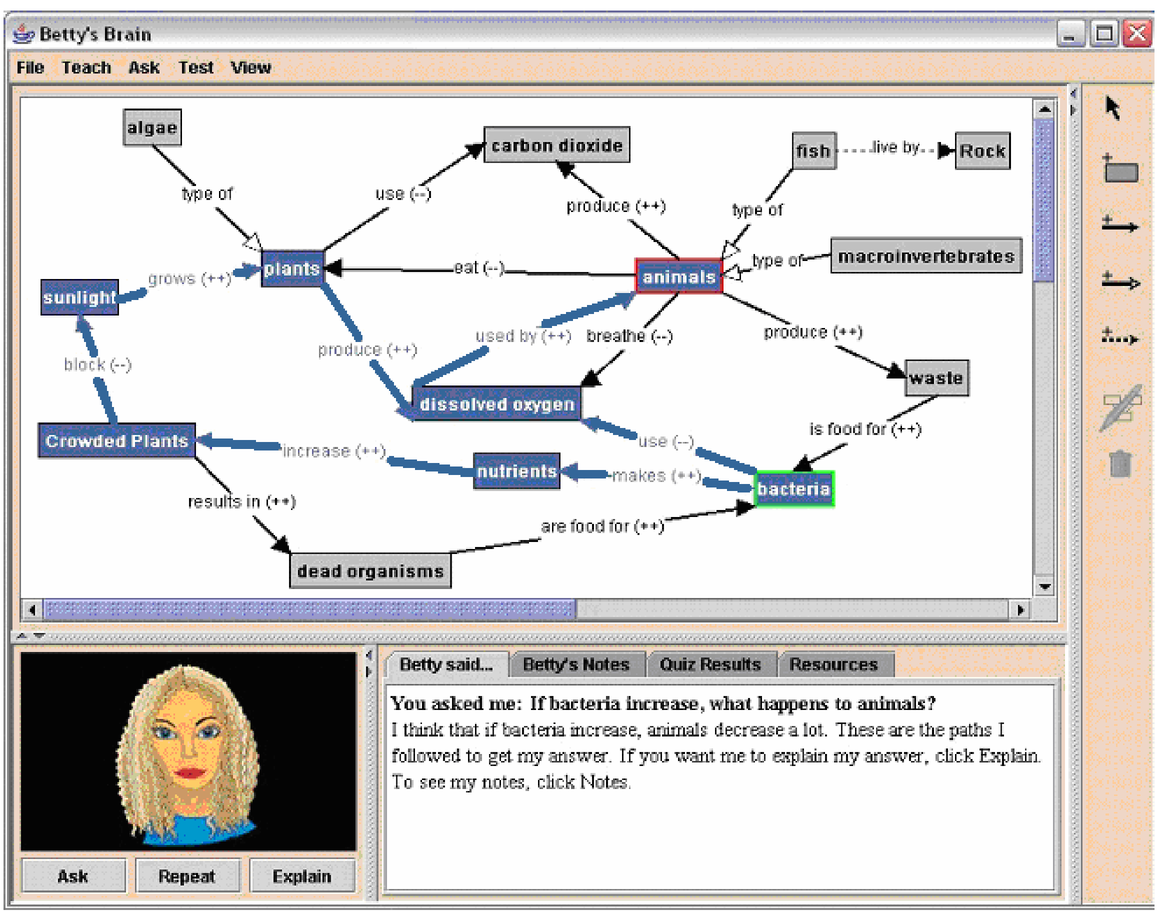


Figure 4.22 Explanations to the Question "If *bacteria* increase, what happens to *animals*?": Overview of the Result

<u>Step 1</u> (Figure 4.31): To deliver the final result, we need to consider the block that consists of two paths both starting with *bacteria* and ending with *dissolved oxygen*. Betty's verbal explanation for this step is, "*To find out what happens to dissolved oxygen when bacteria increase, I must first know what happens to bacteria and plants. Both directly affect dissolved oxygen.*"

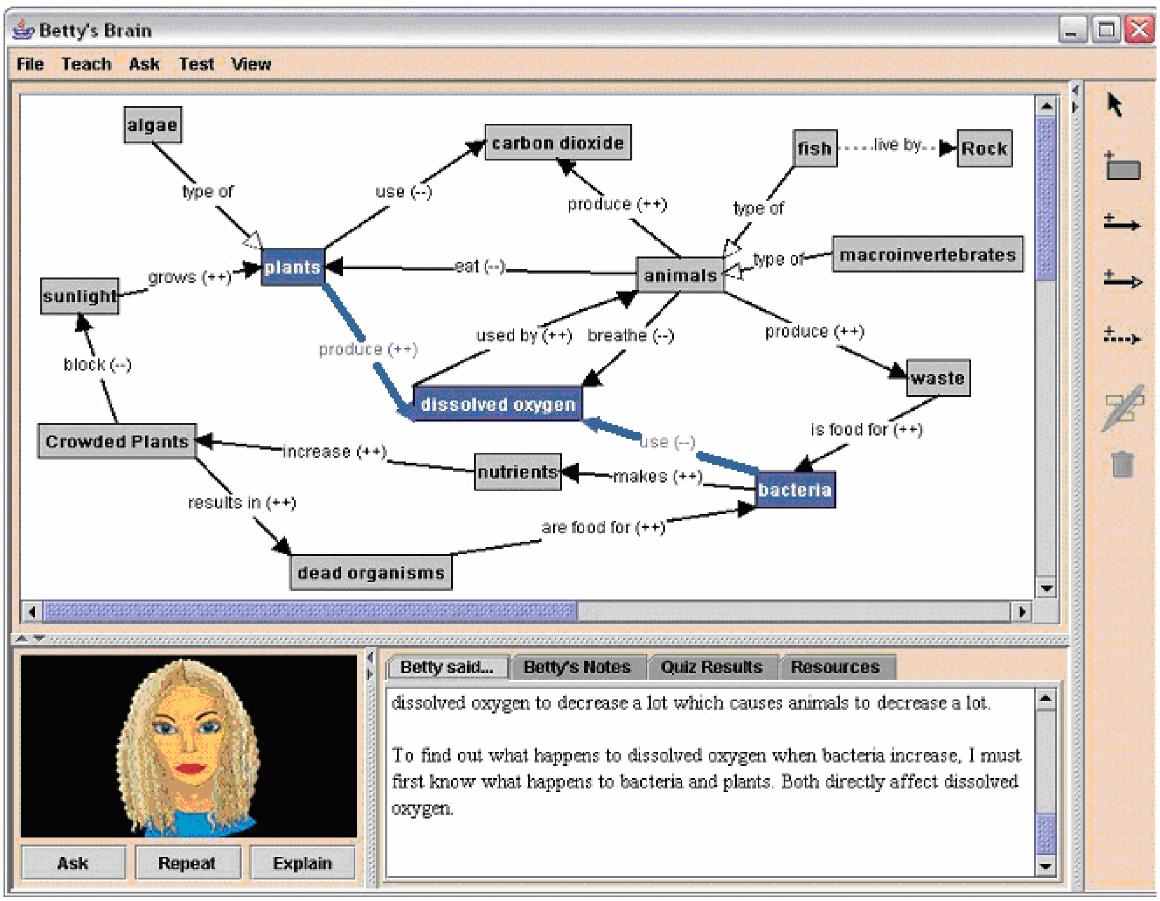


Figure 4.31 Explanations to the Question "If bacteria increase, what happens to animals?": The Merging Point

<u>Step 2</u> (Figure 4.32): The first path starts from *bacteria* through *nutrients*, *crowded plants*, *sunlight*, and *plants*, and ends at *dissolved oxygen*. Betty's verbal explanation for this step is, "*I'll explain how bacteria affect dissolved oxygen through plants. An increase in bacteria causes nutrients to increase, which causes crowded plants to increase, which causes sunlight to decrease, which causes plants to decrease, which causes dissolved oxygen to decrease.*"

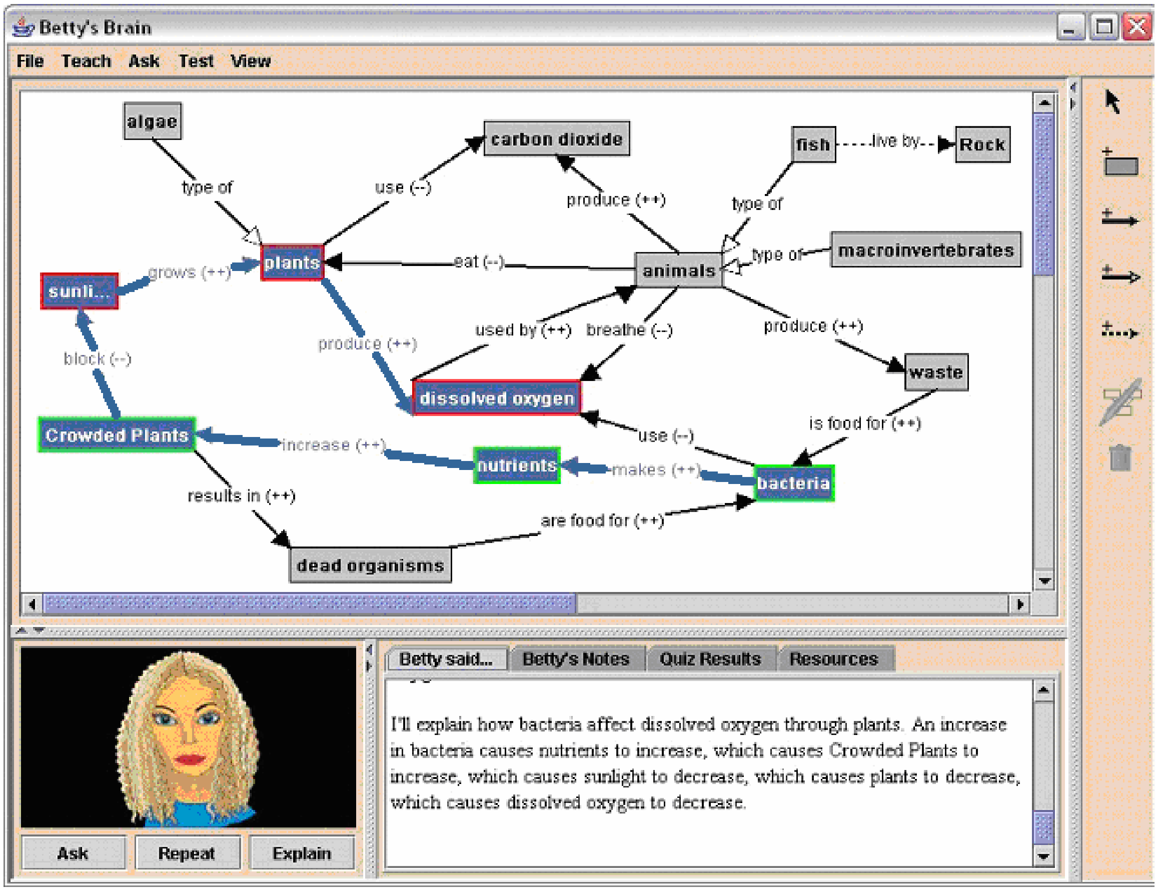


Figure 4.32 Explanations to the Question "If bacteria increase, what happens to animals?": The First Path

Step 3 (Figure 4.33): The second path is a direct link from *bacteria* to *dissolved oxygen*. Betty's verbal explanation for this step is, "*An increase in bacteria causes dissolved oxygen to decrease.*"

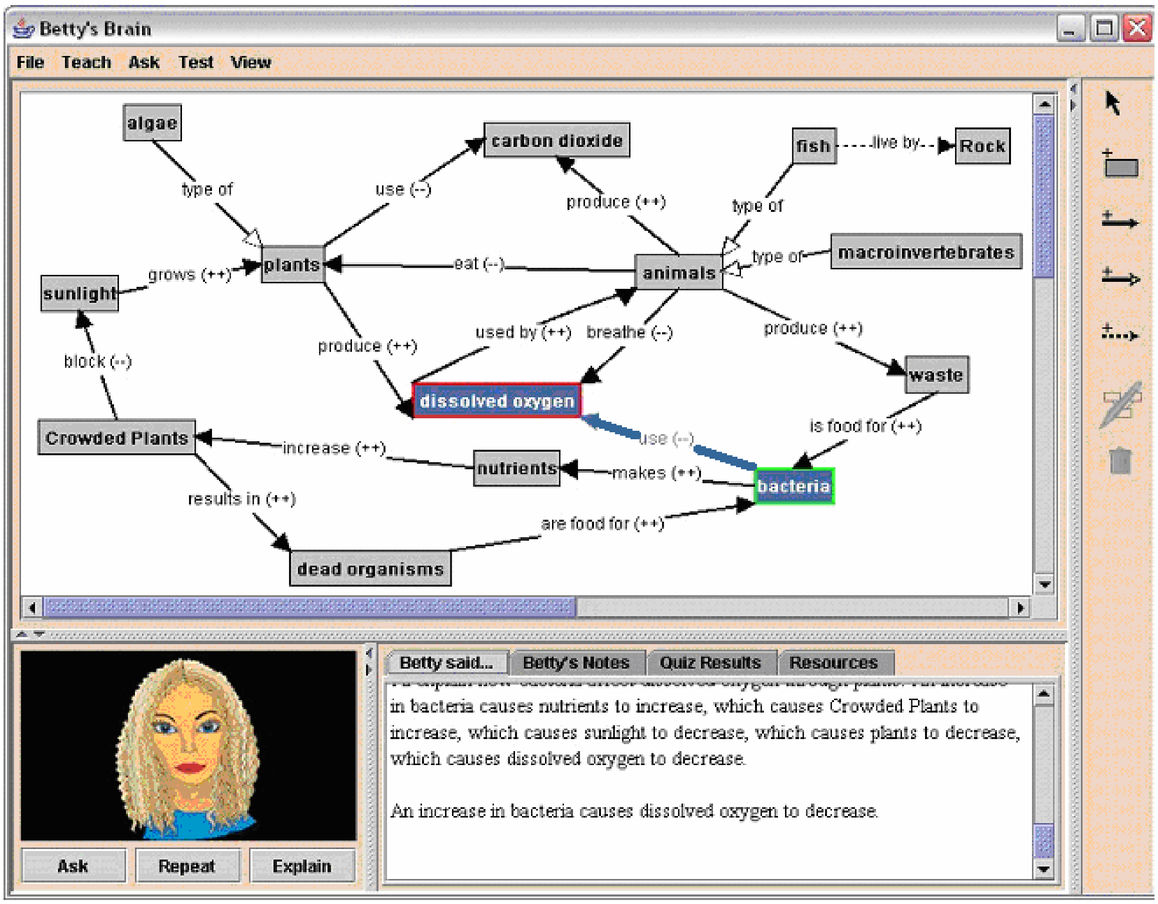


Figure 4.33 Explanations to the Question "If bacteria increase, what happens to animals?": The Second Path

Step 4 (Figure 4.26): The effects are aggregated at *dissolved oxygen* and this combined effect is propagated forward to the target concept, *animals*. Betty's verbal explanation for this step is, "So, a decrease in *dissolved oxygen* from bacteria and a decrease in *dissolved oxygen* from plants cause a large decrease in *dissolved oxygen*." Even though this step does not include the target concept, it is the end of this set of explanations because every branch of the reasoning structure has been explained at this point. Therefore, the conclusion (the answer) that is already given in the top step is not given again here.

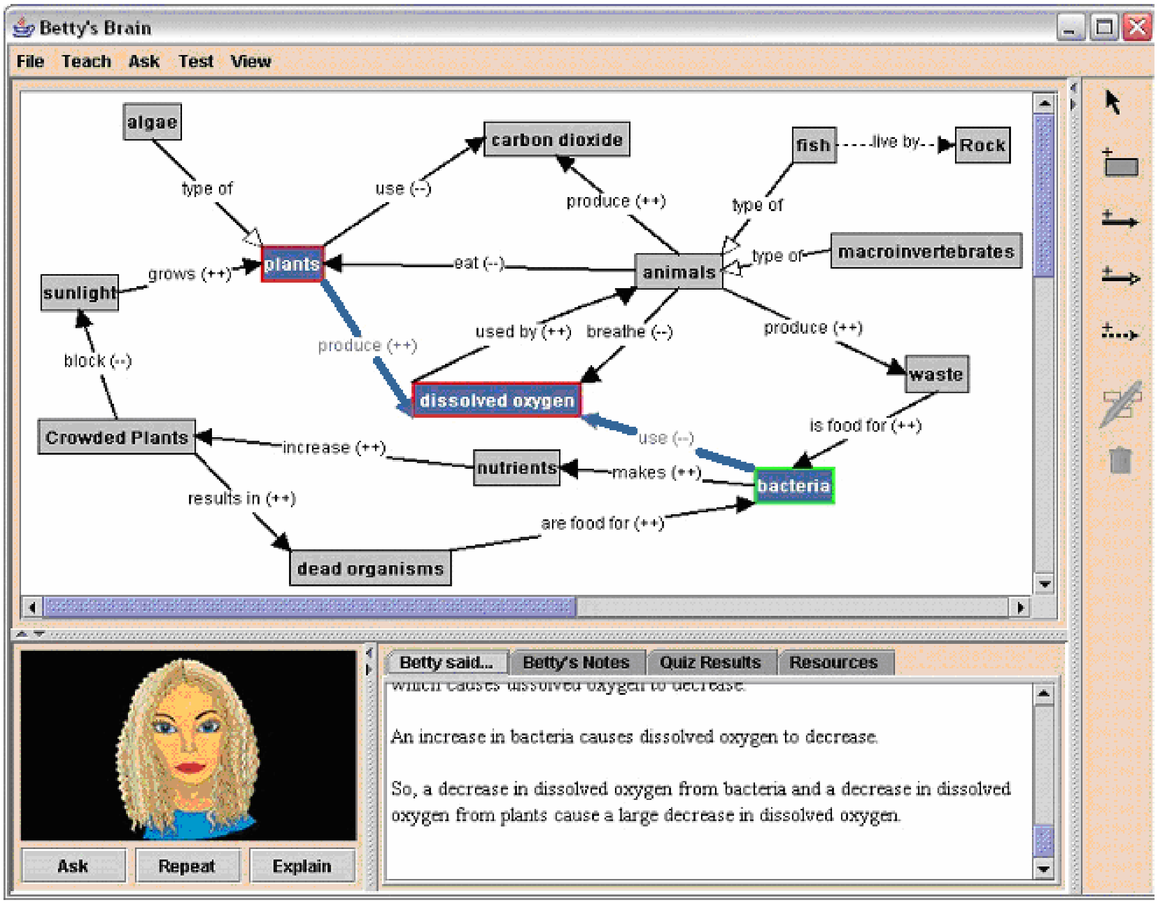


Figure 4.26 The Explanations to the Question "If *bacteria* increase, what happens to *animals*?": The Causal-Effect Conclusion

Quiz Mechanism

An effective learning environment must provide a means for formative assessment (Bransford, Brown, and Cocking 2000). Formative assessment in the Betty's Brain environment is provided via a set of pre-defined quiz questions that are typically generated by the domain expert or the classroom teacher. Students can ask Betty to take a quiz at any point during the teaching process. The mentor agent produces two kinds of feedback after Betty takes a quiz, outcome and corrective feedback. First, Betty and the student are told of the number of correct answers in the quiz. Second, the mentor agent gives corrective suggestions that can be made to the concept map structure to improve Betty's performance. The quiz mechanism also provides a scaffold because it indirectly alerts students to concepts and relations they need to pay attention to in the domain of study.

This version of Betty's Brain includes three quizzes of four to six questions per quiz. Each quiz covers a specific part of the expert concept-map in the domain of study. The number of quizzes and the number of questions in each quiz can be varied. Furthermore, they are easily tailored to any domain of study because they are based on the causal-query template and can be authored via a text-based file (see this system's quiz configuration file in Appendix A). For example, the quizzes can be customized to different difficulty levels or to cover different parts of the expert's concept map. In addition, the expert map can be changed to any concept map constructed in the concept-map editor of Betty's Brain.

To get Betty to take a quiz, the student starts by clicking the "Test" item on the menu bar and select the only item, "Quiz." When the quiz dialog, illustrated in Figure 4.35, appears, the student selects a quiz by clicking on the quiz number from the drop-down menu, as shown at the top of the dialog box. The student can further select questions that Betty should take in a quiz by checking the boxes in front of the question. If there is no question selected, Betty answers all questions in that quiz.

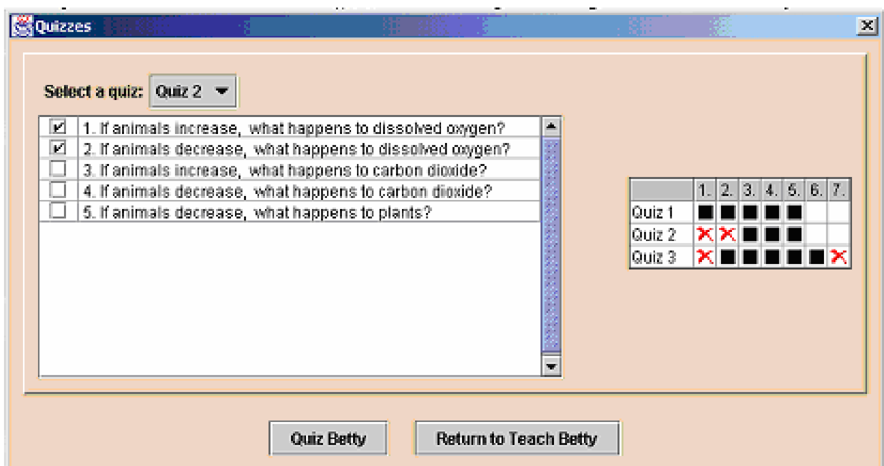
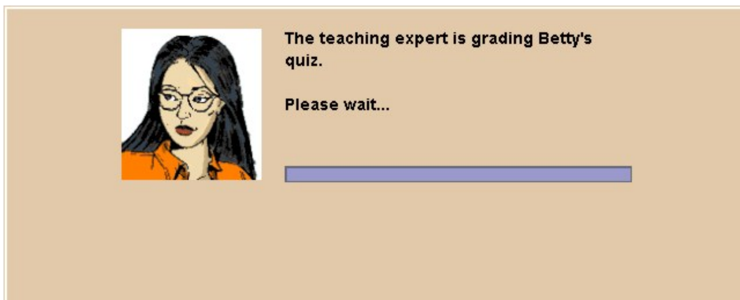


Figure 4.35 Quiz Dialog

Once the student clicks the "*Quiz Betty*" button, a screen appears showing that she is taking a quiz, as illustrated in Figure 4.36 (a). A screen shown in Figure 4.36 (b) appears for a short period of time on the display. The short delay is intentional because it gives a realistic impression that Betty needs some time to answer the quiz questions and the mentor agent also needs some time to grade Betty's answers.

Internally, Betty generates answers to the quiz question-by-question using the causal reasoning mechanism described in the previous section. The mentor agent grades each of Betty's answers by comparing her answers to those generated from the expert concept-map. The right-hand side of quiz-feedback panel in Figure 4.35 displays Betty's quiz grades. The questions that Betty answers correctly are marked with a check icon (✓) and the ones that she answers incorrectly are marked with a cross icon (✕). If Betty has not yet attempted to answer a particular quiz question (because the student has not asked her to), that position is marked with a black rectangle (■).

(a) Betty is taking a Quiz



(b) The Mentor Agent is grading a Quiz

Figure 4.36 The Interface of Taking a Quiz

The mentor agent generates feedback for Betty and the student by overlaying the part of the expert map used to answer the quiz questions on Betty's concept map. The mentor then looks for the following in the order stated, (i) missing expert concepts, (ii) missing expert links, and (iii) incorrect expert links, to generate the feedback content. The following algorithm is applied to each Quiz question to find a concept or a link to hint on.

1. If neither the source nor the target concept exists in Betty's map, the expert suggests that Betty needs to study about the concept.
2. If both the source and target concepts are present, the mentor overlays her answer graph over Betty's map. By starting the traversal of the answer graph at the source concept, the mentor picks the first missing concept or link or the first incorrect link. The mentor then generates the following response:
    a. If the target concept of this link is missing from Betty's answer, the expert suggests that Betty needs to study it.
    b. If the target concept exists but the link is missing, the expert suggests that Betty should study the relationship between the source and target concepts.
    c. If the link exists in Betty's answer but its information, including its type, trend, and direction, is different from that link of the expert, the expert suggests that relationship between the source concept and the target concept is not correct.

The mentor agent ignores extraneous concepts and links in a student's concept map.

The Betty's Brain environment uses three levels of hints that range from general to specific as listed below. One word within a pair of square brackets will be chosen to display at a time depending on the type of error.

Level	Concept	Missing Link	Incorrect Link
1	A concept is missing. Check the resources.	Read about *Concept A* and *Concept B* in the resources.	Read about *Concept A* and *Concept B* in the resources.
2	Read the resource about A.	*Concept A* and *Concept B* should be linked.	The link between *Concept A* and *Concept B* is wrong.
3	You should add A to your concept map.	There should be a link from *Concept A* to *Concept B*.	You should change the [type / trend / direction] of the link between *Concept A* and *Concept B*.

To allow students to reflect on Betty's performance and revise their concept maps, the environment provides on-line resources, discussed in the next sub-section, that students can access during the interaction with of the environment.

Resources

The resources for Betty's Brain were written to cover the content that students need to study in order to model a river ecosystem in the level for a fifth-grade science curriculum. The resources were domain-related and separated in sections as follows:

- *What is in a river?*
- *Organisms that Live in Rivers*
    - *Plants*
    - *Animals*
    - *Micro-organisms*

The Study

The purpose of this study was to examine the benefits of different activities associated with "learning by teaching" in the Betty's Brain environment. Details of the study and its results are reported in (Leelawong et al. 2002; Davis et al. 2003). The (Leelawong et al. 2002) paper is included in Appendix B. The system had three main features, teaching, querying, and quizzing. Crossing these variables created four versions of the teachable agent environment with features as shown in Figure 4.37.

	40.	41. Features			
	43.	44. Teach	45. Query	46. Quiz	47. Resources
Grou	49. TEACH only	50. ×	51.	52.	53. ×
ps	54. QUERY	55. ×	56. ×	57.	58. ×
	59. QUIZ	60. ×	61.	62. ×	63. ×
.	64. FULL	65. ×	66. ×	67. ×	68. ×

Figure 4.37 Four Experimental Groups and their Features

All students had the opportunity to teach their agent by creating a concept map using the concept map editing facilities. In the other three conditions, we manipulated whether students could query Betty and/or observe her quiz performance as they taught Betty. Every group had access to the same resources. In the baseline condition (TEACH only), students just created a concept map to represent their knowledge of river ecosystems. They did not have other interactions with the system other than teaching (resources included). In the QUERY condition, students could ask Betty questions as they taught her, but the quiz feature was not provided. Similarly, in the QUIZ condition, students could ask Betty to take a quiz as they taught her but did not have access to the query feature. In the FULL condition, students had access to both the query and quiz features as they taught Betty.

We hypothesized that having opportunities to query and/or quiz Betty would positively, but differentially, impact students' learning. The query feature should help students debug their own thinking and reasoning in the domain of study. If Betty answered questions in unexpected ways, students would realize that they needed to add to or modify their concept maps. In addition, and perhaps more important, when Betty explained her answers, she made explicit the process of reasoning across links in a concept map (i.e., infer the effect of one concept on another through one or more chains of relations). Therefore, we expected that students who used the QUERY version of the software would create maps containing more inter-linked concepts. With respect to the quiz condition, we expected that students would become better at identifying important concepts and links to include in their maps because they could map backward from the quiz questions. We also expected that overall they would produce more accurate concept maps because they had access to feedback on Betty's quiz performance.

Procedures

Fifty high-achieving fifth grade students from a science class in an urban public school located in a southeastern city participated in the study. (The exact information is restricted due to the agreement with the Institutional Review Board at Vanderbilt University.) Students were randomly assigned to one of four versions of the software: TEACH only, QUERY, QUIZ or the FULL version.

The software was used in 3 sessions of one hour each. At the beginning of the first session, students were introduced to the features of the system they would work with. Every group was asked to teach Betty about river ecosystems—to represent the primary entities in rivers and how they interact with each other to maintain the balance of the natural system. In between sessions with Betty, students engaged in independent study to prepare themselves to teach Betty. Reference materials were also available for students to access as needed when preparing to teach and when teaching Betty.

Results and Discussions

The detailed results appear in (Leelawong et al. 2002) that is included in Appendix B. In summary, results from this study indicated that both the query and quiz features had beneficial effects on students' learning about ecosystems. The *query* feature resulted in more inter-linked concept maps. By being able to query Betty about concepts relations in their concept maps, students realized the importance of causal interdependences among concepts in describing and understanding the river ecosystems domain. The results also indicated that students who had access to *quizzes* generated by the teacher or a domain expert naturally focused more on those concepts and their interrelationships. This decreased the number of irrelevant concepts, increased the number of valid causal links and the number of expert causal links in the students' maps. Overall, the quiz feature was effective in helping students determine relevant domain concepts and relations to teach Betty so she could improve her performance. Students reasonably inferred that if a concept or relationship was in the quiz, it was important for Betty to know.

Surprisingly, students in the QUERY condition produced as many valid relevant causal links as the students in the QUIZ group even though they did not have the benefit of quiz feedback. This demonstrated the value of explicitly illustrating the reasoning process (by having Betty explain her answers) so that students understood causal structures.

The FULL group did not generate significantly higher-quality maps than the QUIZ and the QUERY groups. An investigation of the activity logs of the FULL-group students revealed a pattern where the students' primary focus was on producing the correct answers rather than being concerned whether Betty (and they themselves) understood the interdependence relations involved and their effects on the overall system (Davis et al. 2003). After getting Betty to take the quiz, they used the teaching expert's hints to make corrections to their maps and used the query feature only to check if Betty would now answer the questions correctly. They then quickly returned to the quiz mode to see how well Betty performed. In other words, the query mechanism was not used to reflect on the reasoning mechanisms and to gain a deeper understanding of the causal structures they had created, before corrections were attempted on the concept map. Thus the feedback had inadvertently focused students on making local changes to their maps instead of reasoning more globally in their maps.

In summary, the results indicated that providing students with opportunities to quiz their agent decreased the amount of irrelevant information and increased the proportion of causal information in students' maps. Having opportunities to query their agent helped students develop an understanding of the interrelationships of entities, living and non-living, in river ecosystems. However, one of the primary shortcomings of this system was that it did not promote the use of strategies for effective learning. An explanation was that the teaching expert's directed feedback resulted in students making local changes to their concept maps to get the correct answers to quiz questions. Students relied heavily on this feedback from the mentor agent (Davis et al. 2003) and, therefore, did not seem to learn as much about river ecosystems. The results pointed to the importance of various forms of feedback when designing teachable agent environments that promoted learning.

Summary

Our analyses of the accuracy of the final concept maps and student activities indicated that the participants became proficient in using the concept map structure and the features available in the Betty's Brain environment:

- The query mechanism helped students develop an understanding of the interrelationships of entities in an ecosystem.
- The quiz feature helped students determine relevant domain concepts and relations to teach Betty so she could improve her performance.

However, it was not clear as to how much their domain knowledge had improved in term of understanding the concepts of interdependence and balance among the entities in the river ecosystem.

A previous study (Leelawong et al. 2001) indicated that Betty's Brain had helped college students become more aware of the interdependence between sets of concepts. Students working with pencil and paper tended to create single causal links, whereas students working with Betty's Brain created longer chains of causal structures. However, the study reported in (Leelawong et al. 2002) demonstrates that younger students had difficulty comprehending the global consequences of interdependence among entities and how that affected balance in the river ecosystem.

In summary, the findings of this study suggested the need to modify Betty's Brian, especially the quiz and feedback features, to focus students on the interrelationships between concepts and the consequences of these relationships on the ecosystem. The quiz feature needed to promote more reflective learning by students. Furthermore, in exit interviews a number of students indicated that while they found the overall environment very interesting and easy to work with, they would like Betty to be more active and participatory in the learning process. Betty was passive and only responded when asked questions. We believed that, to create a true learning by teaching environment, Betty needed to be more interactive and demonstrate more human student-like qualities.

The approach to make Betty more interactive and participatory was to adopt **self-regulated learning** (Zimmerman 1989) mechanisms into the new design of the environment. The idea was that, if Betty demonstrated these characteristics, students would assimilate them and become more independent learners with better meta-cognitive skills. Self-regulated learning, discussed in detail in the next chapter, was emphasized to help students develop general learning strategies. Previous research had shown that self-regulated skills could enhance the development of meta-cognitive skills (Moses and Baird 1999). However, a question was whether it would also enhance students' learning abilities. Therefore, the next step was to investigate if introducing self-regulated learning i in the Betty's Brain environment would result in better understanding and transfer of knowledge.

From the knowledge-centered standpoint, we needed to modify the environment's features to make students more cognizant of interdependence issues among entities and their consequences in the occurrence of chains of events rather than local interdependence, particularly with regard to balance. For example, a river ecosystem maintains balance because of the interactions between processes, such as the carbon dioxide and oxygen cycle, food chain, and decomposition of waste. Unless students understood how interdependence produces balance, it would be difficult for them to understand the effects of external phenomena that might affect the system. In addition, students should have a greater understanding of the idea that local changes affected global behavior.

Consistent with this need, we also decided to redesign the student resources to emphasize the processes and cycles that described domain phenomena, as opposed to the individual entities that make up the domain. The on-line resources were reformatted in hypertext format to enable keyword access, and to explicitly reflect the phenomena of interdependence and balance through interactions among the primary cycles. The roles of the different entities in the river ecosystem were also developed in this framework. As a result, entities such as plants and bacteria appear in different roles in different cycles and processes. This realization should make the student realize the global implications of interdependence in the concept map.

In addition, other aspects of effective learning environments had to be incorporated in conjunction with the knowledge-centered aspects. Analyses of the protocols showed that the students struggled to learn while teaching (Davis et al. 2003). Students attempted to use all features, the query mechanisms, the quiz mechanism, the mentor agent's feedback and paper and the on-line resources. However, our environment was not sufficiently well-designed to cope with the learner-centered aspect, especially for situations where the student lacked both domain knowledge and teaching experience.

To improve the learner-centered aspect, our environment had to more actively assist students in learning and mastering domain knowledge. The mentor agent could guide and help students to adopt constructive learning techniques, and once they did, encourage them to continue on this path. Betty could also encourage students by asking if they would like to discuss the results of the quiz with her, allowing her to demonstrate learning strategies through her dialog and actions.

We also believed that the Time for Telling strategy (Schwartz and Bransford 1998) should be integrated into Betty's Brain. Betty could construct questions that involve long chains and multiple-path events in the current concept map and ask the student to analyze the answers she generates. Discussion of these answers would lead to better understanding of the concept map structure and Betty's reasoning mechanisms. The Mentor agent could suggest reading materials related to the discussion.

We would also like to associate Betty with behaviors that were linked to good learning practice. Betty could be endowed with self-regulation strategies that drive her interactions with the student. For example, to counter the student's sole focus on performance in the quiz, Betty might refuse to take the quiz if the student repeatedly ignores the teacher agent's feedback. Along the same lines, Betty could express skepticism if the student did not look up resources before attempting to make corrections in the concept map.

The assessment-centered aspect of our design framework could be improved if the interface allowed a different quiz structure that emphasized global system behavior instead of local changes. For example, the feedback could better emphasize the global nature of changes at the expense of a local change made to get one particular quiz question correct. The quiz feedback could also promote reflective learning to a higher degree if Betty incorporated reflective activities, potentially encouraged by the teacher agent. Our interactive Betty could also enhance the community-centered aspect. By collaborating on the quiz results and discussing the local and global causal effects in the concept map, Betty shared a common space in learning with the student.

The next chapter presents the concept of self-regulated learning and discusses how to integrate this concept and the modifications mentioned above into the Betty's Brain environment.

CHAPTER V

BETTY'S BRAIN ENHANCED WITH SELF-REGULATED-LEARNING STRATEGIES

This chapter presents two primary contributions of the thesis in the design of the Betty's Brain environment: (i) feedback based on self-regulated learning and (ii) an agent architecture implementation for the system. The ultimate goal of the Betty's Brain environment is to prepare students for life-long learning. The experiment with the previous version of the Betty's Brain environment, described in Chapter 4, demonstrated its effectiveness in assisting students to learn new domains. However, the results did not indicate that it helped students to prepare for future learning tasks. The previous version of Betty's Brain had only one type of feedback, which was directed at correcting errors in students' concept maps. Its study discussed previously demonstrated that this kind of feedback is not sufficient to promote deep learning and transfer. As discussed at the end of the previous chapter, based on educational theories a potential solution is to guide the students learning processes by encouraging them to follow self-regulated-learning strategies. The details are discussed in the first part of this chapter.

This decision led to a number of changes to the Betty's Brain environment. For clarity, in this chapter we call the version in Chapter 4 the **baseline** Betty's Brain environment, and the new version presented in this chapter, the **SRL** (self-regulated learning) Betty's Brain environment. Due to the large scale of changes and the effort to keep the environment flexible and customizable in case that learning and instruction methods were to be added and removed, we decided that the implementation of the SRL version should use a multi-agent architecture. The multi-agent architecture design and implementation are discussed in the latter part of this chapter.

Framework for Learning by Teaching and Self-Regulated Learning

To promote deep learning and transfer students need to learn self-regulation and meta-cognition strategies to guide their learning and problem-solving tasks. To help students do so, this dissertation has proposed improvements in two areas, (i) making Betty practice self-regulated learning ,and (ii) changing the Mentor agent's feedback from being focusing on changes to make corrections in the students' concept maps, to providing more generic feedback that discusses issues in learning, teaching, and the domain content. The mentor agent's new feedback is also based on self-regulated learning (defined in Chapter 2), and the domain-content feedback focuses on chains of events rather than local corrections to the concept map.

In addition, these revisions address previous concerns of the students' use of the quiz feature for gaming with the system, and focusing mainly on getting quiz answers correct. By redesigning features to encourage independent learning, we can also reduce the degree of dependence on the quiz feedback as the primary mechanism for learning about the domain.

The following list describes a set of self-regulated learning and pedagogical skills we have implemented to support self-regulated learning in this version of Betty's Brain. This list is selected from Zimmerman's social cognitive view of self-regulated academic learning (Zimmerman 1989) presented in Chapter 2:

1. *Monitoring knowledge*: The Betty's Brain environment allows students to (i) assess Betty's learning progress and (ii) revise Betty's knowledge after receiving feedback from a quiz. Students can assess the consequences of their own learning processes by asking Betty questions. This may be considered equivalent to a self-explanation (Chi, Bassok et al. 1989) process but has the added advantage of being incorporated into a social interaction. The quiz feature provides formative assessment, where students receive feedback during the learning process, provided in the Betty's Brain system. This feedback should enable students to reflect on their domain knowledge, which is explicitly depicted using in a visual concept map structure. Therefore, students can systematically monitor their progress in learning through the shared representation (concept maps) and shared activities (such as quiz taking).
2. *Self-assessment*: Betty can take quizzes with the Mentor agent and discuss the results with students by commenting on her performance on a quiz. An example response from Betty is:
   "*Hi, I'm back. I'm feeling bad because I could not answer some questions in the quiz. They seem to deal with concepts that I don't know about. Ms. Davis said that I should study more before I take the quiz again.*"
   Note that the Mentor agent's initial comments are very general, but they may become more specific if students have repeated the errors or ask for more help.
3. *Goal setting*: The mentor agent helps students to set goals for studying domain knowledge by pointing out processes in the system and sometimes, concepts and relations that they should pay attention to.
4. *Keeping records and monitoring*: Betty keeps records of questions she has been asked and quizzes that she has taken. Students can access these records at anytime to reflect on their own learning experience.
5. *Seeking information*: The on-line resources are always available to the student.
6. *Self-consequating*: This goal is not explicit to the students, but we hope they are motivated to help Betty learn so that she can join the high-school science club.

7-8. Seeking social assistance from:

7. *Peers*: Our new Betty behaves more like a peer rather than a passive tutee since she uses her knowledge about learning strategies to drive her interactions with the student. She has expertise in reasoning with knowledge created as causal concept maps, but she has no *a priori* knowledge of river ecosystems. She only knows what she has been taught by the student. In addition, Betty has perfect memory of what she has been taught.
8. *Mentor*: The mentor agent's help is on demand. Mr. Davis offers three categories of help, learning-related, teaching-related, and the domain-related (the river ecosystem)

9-11. Reviewing records, such as rereading:

9. *Notes*: The student can access the records of interactions that Betty and the mentor agent keep at anytime during the use of the environment.

10. *Tests*: Betty keeps a record of quiz questions she has answered and the feedback she has received from the Mentor agent. Betty displays this information on demand.
11. *Textbooks*: We have developed structured resources that still contain the same information as in ordinary textbooks but organize the contents for easy access and search.
12. *Trying strategies suggested by the others*: The student can decide to try strategies suggested by the mentor agent and Betty.

Together, Betty and the mentor agent should help students realize their progress in learning domain materials and regulate their learning, as well as their learning process through conversation and feedback. Students have chance to see Betty practice her learning strategies, received guidance from the mentor, and practice what they have seen or been told.

To ensure the implementation of self-regulated learning strategies, we have modified a number of the components in the Betty's Brain environment, three main flows of activities in the environment, the concept-map editor, Betty's responses to quiz requests, and the feedback that students receive from the mentor agent. The rest of this section discusses these changes. Other changes in the environment that are secondary, namely the search feature and the descriptive editor-menu, will be described later. To accommodate these modifications, the environment's graphical user-interface has been modified as shown in Figure 5.38.

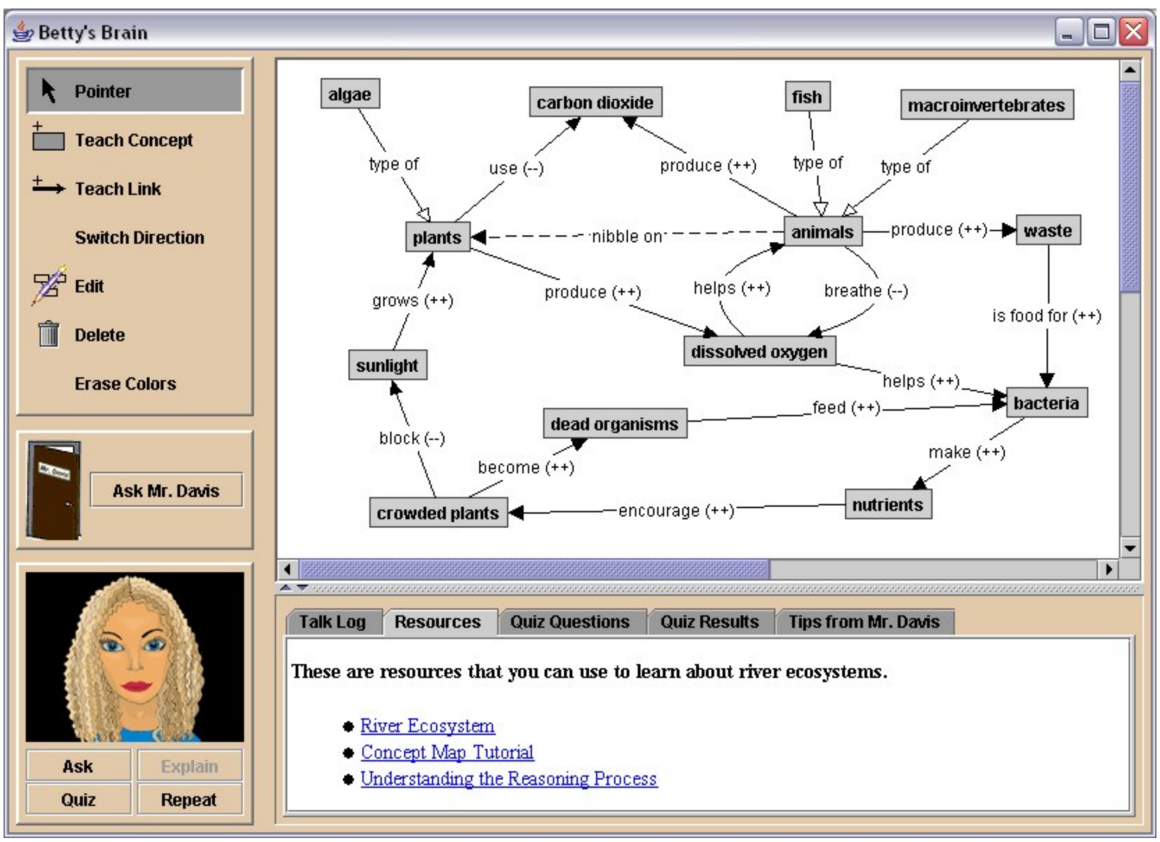


Figure 5.38 The SRL-enhanced Betty's Brain Environment

Modifications to the Concept-Map Editor

Two issues governed the modifications made to the concept-map editor, (i) the interface for link-type selection and (ii) the reasoning procedure. In the first version of the system, some students did not seem to realize that causal links were the key to expressing the concept of interdependence and balance in a river ecosystem. In addition, some students expressed confusion about the type of link to use for a particular situation. Hence, in this version of Betty's Brain we modified the interface that students use to teach Betty about links. In addition, Betty responds to students as they create and modify their concept maps.

Recall that in the baseline version, there was a separate interface component for adding each kind of link. In the SRL version, we use one pop-up window in which students have to select which one of the three types of links they will use to teach Betty. After clicking on the "Teach Link" button, students click inside the source concept box and drag the link's arrow head that appears on the screen to the target concept box. The add-link dialog box, displayed in Figure 5.39, appears, and the student has to select a link type from the dialog box.

Once the student clicks on a radio button in front of a link type, a customized dialog for that type of link appears at the bottom of the previous window. If the student selects a different kind of link, the lower pane will change according to the type, as displayed in Figure 5.40. Students are required to make a choice when adding or modifying a link, and this gives them the opportunity to compare the differences among the three types of links. This should add some cognitive value to this interaction, which can ultimately lead to a better understanding of the link types and their roles in representing and reasoning with information. However, a formal study is still needed to confirm this conjecture.

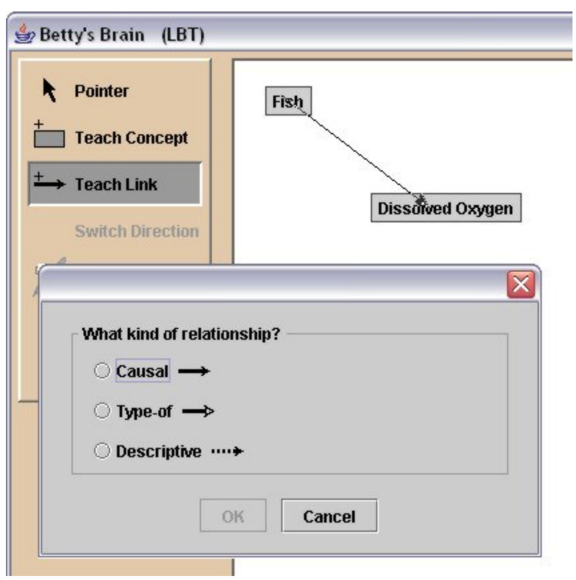


Figure 5.39 Link Addition

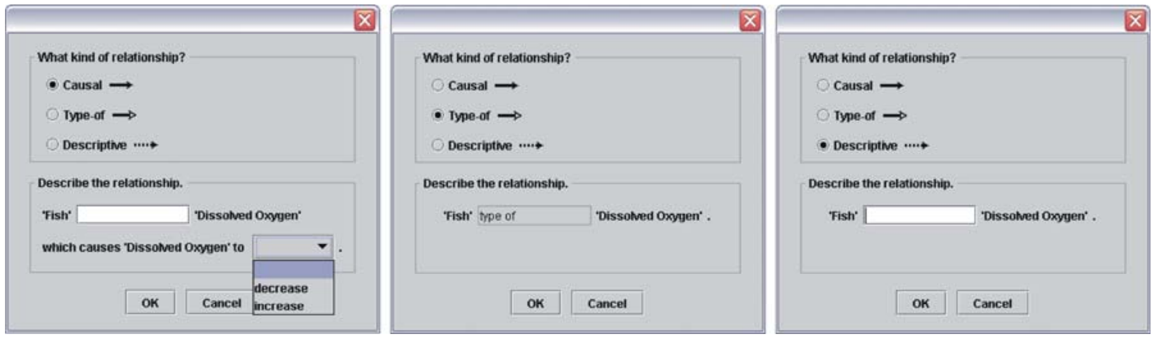


Figure 5.40 Link Dialogs for adding Causal, Type-of and Descriptive Links, respectively

Another modification enables Betty to interact with students on her own initiative. When the student teaches Betty a new causal link the first time in a session, Betty automatically reasons with that link. For example, when the student adds the causal link, "Animals *breathe* (decrease) Dissolved Oxygen," Betty responds as shown in Figure 5.41. Furthermore, the first time in each session that the student adds a causal link so that Betty can make a longer chain of reasoning (three links or more) that includes this link, Betty elaborates the chain of events on this path. Betty performs each of these actions only once every session so that students are not disturbed by a repetitive behavior, which may not have much value after a while. These mechanisms serve two purposes. First, this makes Betty more responsive, and students feel that she is more like a "real" student. Second, by spontaneously illustrating how to reason with causal links, Betty demonstrates the causal reasoning procedure. We believe that this helps students gain a better understanding of how to model interdependencies among entities and study their effects.

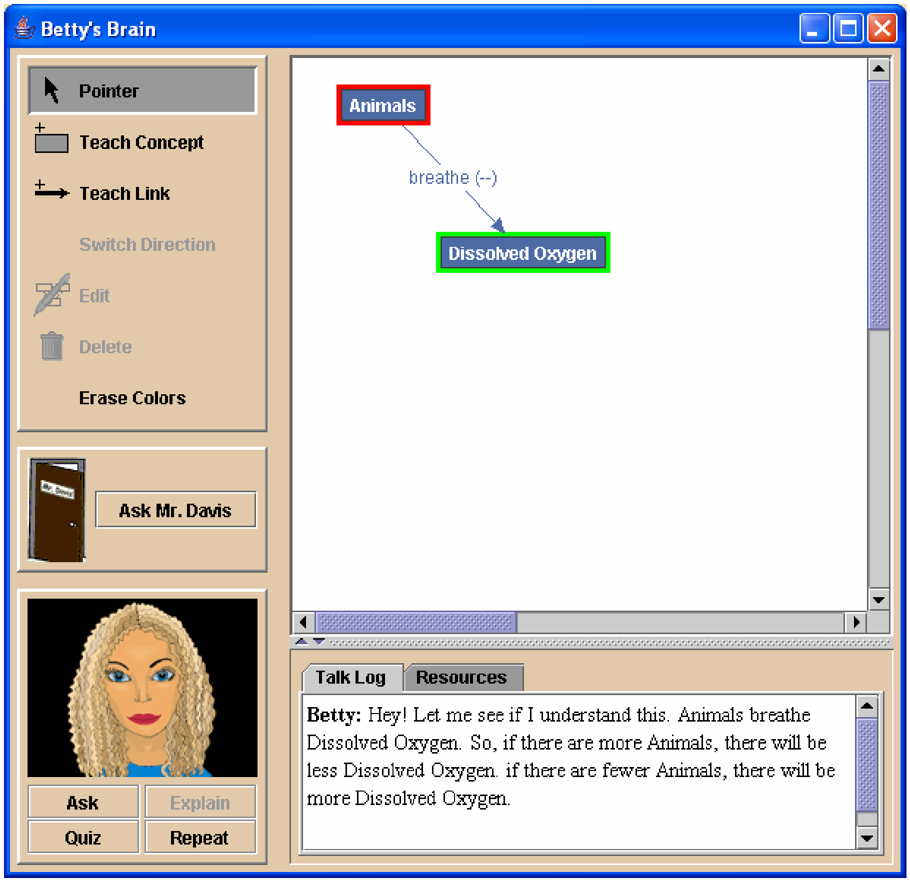


Figure 5.41 Betty's Automatic Reasoning when the first Causal Link is added to her Concept Map in each session

Modifications to the Quiz Feature

As we expected, the quiz feature as a formative-assessment tool proved to be very useful in directing students to think about relevant entities and their relationships in the domain. Furthermore, students frequently used the quiz to debug their concept maps. However, an unfortunate side effect observed from our previous study was the students' overwhelming dependence on the quiz results and local feedback received from the mentor agent to revise their concept maps (Davis, Leelawong et al. 2003). The students' focus seemed to shift from trying to learn the domain so that they could teach Betty better to trying to get the quiz questions correct one by one. As discussed in the paper, we observed that a number of students adopted the strategy of sending Betty to take a quiz, editing the concept maps only when directed, and then resending Betty to take the same quiz. In other words, those students relied heavily on the mentor's local feedback to modify their concept maps and rarely made attempts to understand the concept-map structure via the self-assessment feature, queries, or read the resources before making changes to their maps.

In summary, four modifications were made to the quiz feature to reduce this dependency. The details of these modifications are discussed next.

1. Betty can decide whether she is ready to take a quiz.
2. The quiz interface is redesigned to lead students to conduct formative assessment and self-reflection instead of being the feature that students receive localized, structural feedback.
3. The quiz questions are made accessible to students at all times once Betty has taken an initial quiz.
4. The quizzes are ordered in a way that they are increasingly difficult. This scaffolds the students' efforts as they make attempts to learn about the domain.

First, the SRL Betty decides whether she is ready to take a quiz when students ask her using the new interface shown in Figure 5.42. Her decision to take a quiz is based on two rules:

1. There must be three causal links or more in the student's concept map.
2. The student has asked at least one causal query since the last time Betty took a quiz.

The first condition is to ensure that students have paid enough attention to causal relations that are important to natural, dynamic systems, such as river ecosystems. As discussed, the second condition ensures that students self-reflect on Betty's quiz performance and spend some time debugging and correcting their maps using the query feature before they request Betty to take a quiz again. Betty responds to take a quiz only if both conditions are true. When Betty refuses to take a quiz, she explains to the student why she is not ready.

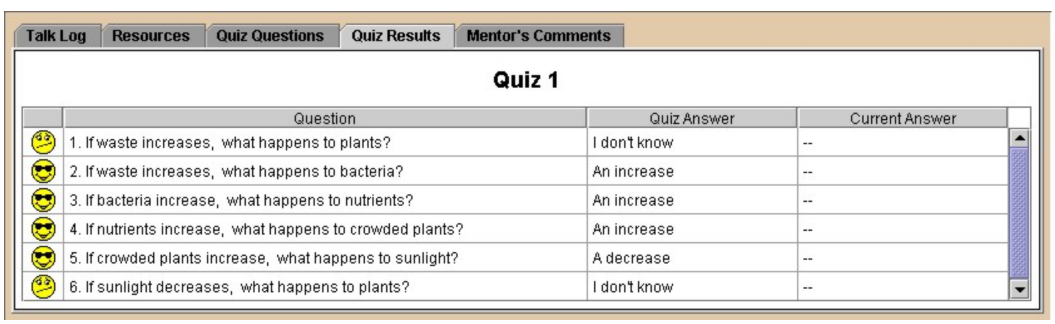
Figure 5.42 Requesting Betty to Take a Quiz

Second, we have redesigned the quiz feature in a way that it provides better support for formative assessment and self-reflection. Two major modifications help to achieve this: (i) the removal of the immediate, localized quiz feedback, and (ii) setting up a connection between the quiz and query features. To remove localized feedback, the mentor does not give direct help on missing concepts and links in the student's concept map as a part of the immediate feedback for quizzes. Only outcome feedback is given at this point. In the SRL version, students need to seek the domain-specific help from the mentor as discussed in the *Modifications to the Mentor's Feedback* section in this chapter.

We would like to see students use the quiz feature as a formative assessment tool to aid their learning as opposed to using it solely as a performance measurement. To achieve this, we provide the connection between the query and the quiz features to help students use the quiz results as an assessment of their concept maps. In the SRL version, students can transfer a a quiz question to the query window. This helps the students to, (i) better debug their concept maps if there are missing or incorrect concepts and links and (ii) reflect on the quality of their concept maps if they are good enough to answer the quiz. This feature is explained with a walkthrough example presented next.

We sill keep the feature that students need to wait for a brief period of time when Betty is taking a quiz, as shown in Figure 5.43. Betty reappears after the mentor agent has graded the quiz. Figure 5.44 shows the new interface of the quiz-grade panel that consists of four columns. The first column provides outcome feedback (correct or incorrect) for the corresponding quiz question is in the second column. The correct answer has a happy-face icon (), and the incorrect one has an unhappy-face icon (). The third column lists Betty's answers to the quiz questions. The last column is empty at this point.

Figure 5.43 Betty is taking a Quiz



	Question	Quiz Answer	Current Answer
	1. If waste increases, what happens to plants?	I don't know	--
	2. If waste increases, what happens to bacteria?	An increase	--
	3. If bacteria increase, what happens to nutrients?	An increase	--
	4. If nutrients increase, what happens to crowded plants?	An increase	--
	5. If crowded plants increase, what happens to sunlight?	A decrease	--
	6. If sunlight decreases, what happens to plants?	I don't know	--

Figure 5.44 The Quiz-Result Panel

Now the student can click on the quiz question that he wishes to debug and then click on the "Ask" button as if to ask a question. In this case, he selects the sixth question. Now there is an additional menu item, "What is your answer to this quiz question?" at the bottom of the pop-up menu, as shown in Figure 5.45. After Betty gives the answer to this query, which is the same as the quiz answer, the student can ask Betty to explain her answer as he would do when querying Betty with his own question. In this case, Betty cannot answer this question (the third column shows "I don't know" as her quiz answer).

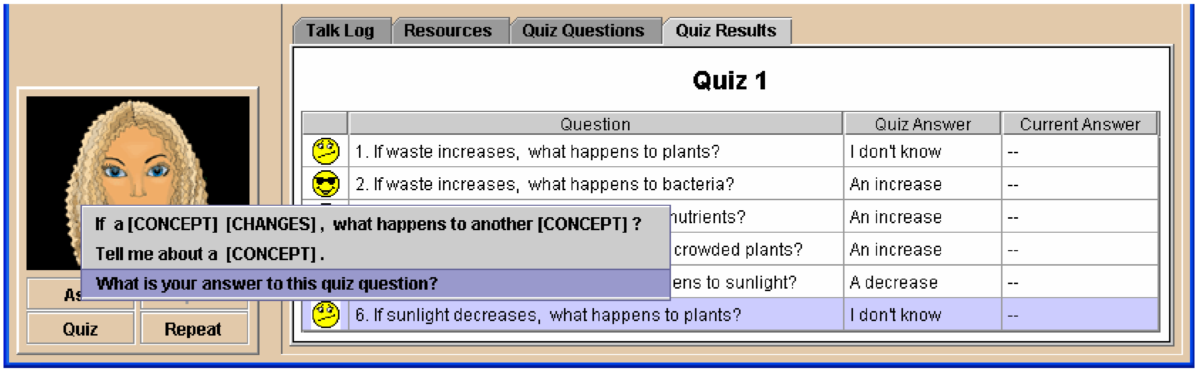


Figure 5.45 Asking a Quiz Question

The student figures out that there is no causal path from the source to the target concept. To fix this, he may study the resources and then add a few causal links to his concept map. To test his actions, the student asks Betty the sixth quiz question again. Betty responds to this request by putting the query result, a decrease, in the forth column, "Query Answer," as illustrated in Figure 5.46. This permits the student to analyze the quiz questions without having to get Betty to take the quiz repeatedly. We believe that this feature should reduce the gaming behaviors that students exhibited before, and get them to focus on higher level tasks as understanding the domain and debugging their concept maps. This fourth column is reset to be empty every time Betty takes a quiz to be ready for the next debugging cycle.

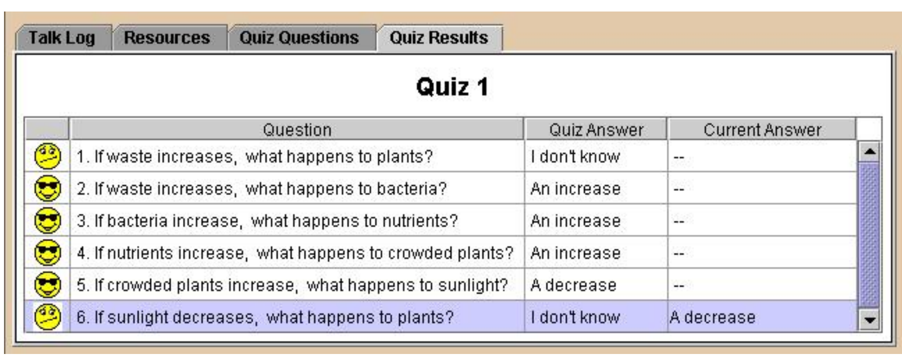


Figure 5.46 Result of Querying the sixth Quiz Question in Quiz 1

In addition to the feature explained above, students can view the complete list of quiz questions in Figure 5.47 at any time after Betty has taken her first quiz. In the previous version, students saw quiz questions only when they took a quiz. Now that all quiz questions are visible, we hope that students would utilize them as a scaffold to develop their own thinking process. See Appendix C for the complete list of the quiz questions.

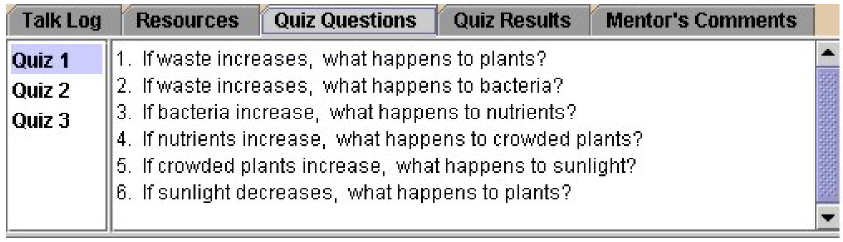


Figure 5.47 Quiz Question List

Fourth, each quiz is increasingly difficult. The first quiz covers only a specific part of the mentor's concept map and most questions require only single-link answers. The last quiz covers the entire map, and each question requires multiple links. Furthermore, each quiz is scaffolded to have one major question, and after students answer all other questions in the quiz, only a few links additional are required to answer the major question correctly.

It is important to clarify that the main purpose of the features described in this sub-section is to encourage students to use the quiz feature strategically as a learning scaffold rather than to look at them purely as a performance measure. As a scaffold of the domain knowledge we would like students to learn, the quizzes contain all concepts we consider important in a river ecosystem. Therefore, it is an acceptable strategy if a student builds his concept map with these concepts and uses other features of the system to discover the relationships between them.

Modifications to the Query Feature

There are two modifications to the query feature. The first modification, the ability to transfer a specific quiz question to the query mechanism without retyping it, was presented in the previous section. The second modification is related to how students ask Betty questions. In this version of the system, students can specify the effect on the target concept when composing a causal question for Betty, as shown in the lower part of the dialog box in Figure 5.48.

Figure 5.48The idea behind this feature is that students can create their own quiz questions for Betty as a real tutor would do and derive their own answers before Betty generates hers. Even though students are not forced to enter their own response to queries, we hope that students can gain more insight especially if Betty's answers happen to be different from their own.

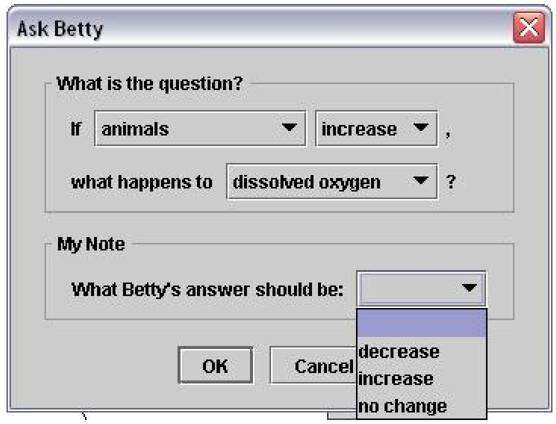


Figure 5.48 The Cause-and-Effect Query Dialog

Modifications to the Mentor Agent's Feedback

As stated before, the previous study indicated the participants' dependency of localized, directed feedback from the baseline mentor agent. This implies that the students have not developed well-established learning strategies of their own, and are not making sufficient effort to understand this material before teaching someone else, in this case Betty. This is because their experience and focus in traditional classroom work is to make sure they score well on standardized and multiple-choice tests, and not to develop the abilities that enable them to apply learnt knowledge to complex problem solving tasks. To help students become independent learners, the mentor agent's feedback in the SRL version covers issues in learning, teaching, and understanding of river ecosystems content. This new feedback suggests a number of strategies that point to the features available in the Betty's Brain environment to help students independently regulate their own learning progresses.

The on-demand feedback is organized into three categories, (i) learning, (ii) teaching, and (iii) domain knowledge of river ecosystems. All of the hint mechanisms and contents of the hints are listed in Appendix D. In summary, the learning feedback helps students establish their own strategies to learn with some guidance from the mentor agent but no explicit feedback on how to correct particular errors in their concept maps. The mentor agent encourages them to conduct research using the on-line resources and monitor their own knowledge by studying Betty's answers to queries and her performance on quizzes.

Fifth-grade students usually do not have experience in teaching. Therefore, the mentor provides feedback on being a better teacher. It is also impressed on the student that learning and teaching in the Betty's Brain environment is a shared responsibility between the student and the teachable agent, Betty. Students need to query Betty to check if she is learning what they have meant to teach, and Betty takes responsibility for answering queries using her reasoning mechanisms. Also, after receiving feedback based on Betty's performance from the formative assessment, the quizzes, students are responsible to reflect the quiz feedback on Betty's concept maps. In short, Betty knows nothing about the domain until the student teaches her about river ecosystems, and typically the student does not know much about how to reason causally with concept maps till Betty shows them how to answer questions.

The SRL mentor still gives domain-content feedback under the "River Ecosystems" category. Unlike the previous version where the mentor agent's feedback focused on missing concepts and links, the feedback in this version of the system is more generic. Students are reminded that interdependence and balance among entities in the river ecosystem are governed by chains of events. Like the other help categories, the quiz feedback is progressive, and the mentor will start giving domain-content help when the two conditions below are met:

1. All concepts that appear in the quiz questions are present in the student's concept map, and
2. Betty has answers to all quiz questions (right or wrong).

If this is not the case, the mentor keeps reminding the student that he needs to read the resources to find relevant information to teach Betty further. Once these two conditions are satisfied, the mentor provides progressively more and more feedback on part to the whole chain of events that are related to those quiz questions every time the user makes some progress and gets Betty to take a quiz. Once the complete chain of events feedback is provided by Mr. Davis but the student is still struggling with the contents of the quiz, the mentor starts to give localized feedback, just as in the previous version. This is done so as not to overly frustrate the student and to keep them moving along in their learning tasks.

Students can access Mr. Davis's on-demand help at any time by clicking on the "Ask Mr. Davis" button, shown in Figure 5.49, in the middle of the left panel of Figure 5.38. On clicking this button, the help topics appear on the top of the dialog box depicted in Figure 5.50. The first help category is about teaching, where students are introduced to various issues in teaching including how to use the concept map to teach Betty. Once the student clicks on the "Teaching" menu item, more detailed choices appear as listed in Figure 5.51. Some menu items lead to more selections, but the one selected in Figure 5.51 provides a hint, shown at the bottom of the screen. Depending on the student's past activities, the help information may vary.

Figure 5.49 Asking the Mentor Agent for Help

Help on other categories are structured in the same fashion. Appendix C provides details of help screens and how the mentor agent responds to different help requests. To go back to the previous screen, click on the "Back" button. To go to the main dialog, click on the "Home" button. Finally, to go back to the primary environment, the student can click on the "Quit" button. Every hint the mentor agent has given is stored, and can be accessed by clicking on the "Tips from Mr. Davis" tab, shown in Figure 5.52.

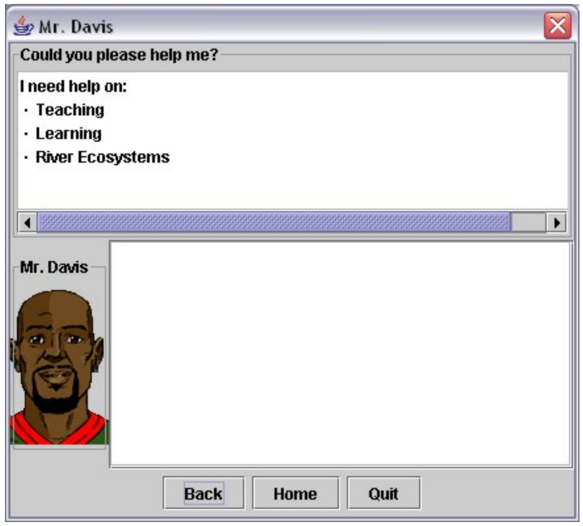


Figure 5.50 On-Demand-Help Dialog

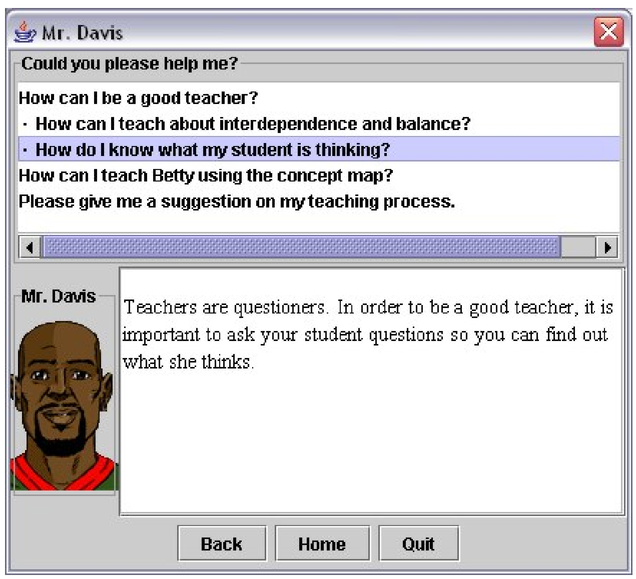


Figure 5.51 Teaching Help

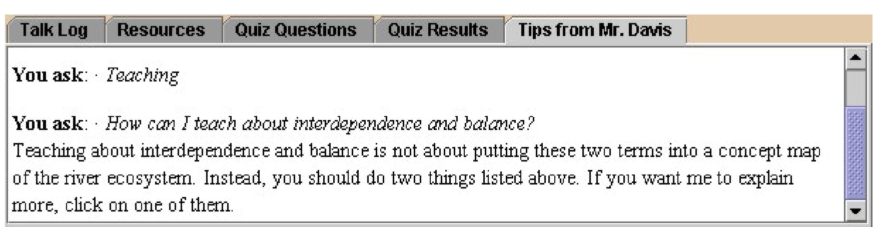


Figure 5.52 The Mentor's Panel

Some sub-categories require dynamic feedback; they change according to the user's activities and progress the student has made in the environment. The mentor agent works with another agent, the Interface agent, to assess the user's activities. The details of these two agents are described in the last section of this chapter.

Other Changes in the Betty's Brain Environment

Two of the changes in the environment were introduced to reduce the cognitive load on students but are not the focus of this dissertation work. The first change was in the control panel of the concept-map editor where each icon and its function are displayed side by side, as illustrated in Figure 5.53. The new control, "Switch Direction," of the link was introduced to accommodate results of a user study (Viswanath, Balachandran et al. 2003) that showed that students prefer to have a feature that allow them to change the direction of a link by one click.

Another change is in the resources that have been expanded to cover the concept-map and the reasoning-process tutorials. In addition, students can search the river-ecosystem resources as illustrated in Figure 5.54. Since links in the Betty's Brain environment are binary, students can enter up to two phrases to conduct a keyword search on the resource document.

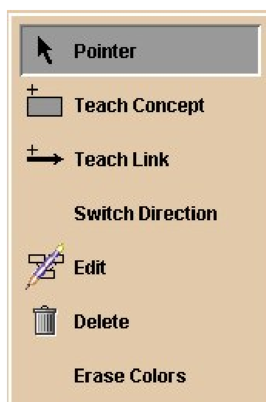


Figure 5.53 Control Panel of the Concept-Map Editor

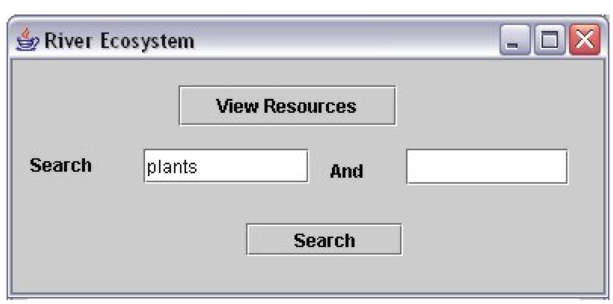


Figure 5.54 Searching Resources

We have already described the main ideas of the SRL-enhanced Betty's Brain environment. However, the changes go beyond the look and feel of the environment. The next section describes the changes within the implementation to facilitate easier interactions between the different components of the environment.

The Betty's Brain Environment and Its Agent Architecture

Following all of the design changes described earlier in this chapter, substantial changes had to be made in the implementation of the baseline Betty's Brain environment. However, these modifications were not easily done because of the coupling between the different components in the baseline system. Another issue was that these changes were experimental. After we conducted an experiment with this SRL version, we might decide to extend certain features or adapt other approaches. Therefore, we decided to re-implement the baseline version using a multi-agent architecture and then extended this implementation to cover the SRL design.

The new version of the system includes four agents: (i) the *Teachable* agent (Betty), (ii) the *Mentor* agent (Mr. Davis), (iii) the *Interfacer* agent, and (iv) the *Pattern Tracker* agent, and one shared module, the graphical user interfaces. The last two agents are new to the environment. We present brief overviews of these agents in the new version of the Betty's Brain environment. These agents are explained in greater details later on in this section.

The **Teachable Agent** ("Betty") models a student who possesses self-regulated learning strategies in addition to the ability to reason with the concept map. This knowledge drives the interactions between Betty and the student.

The **Mentor Agent** ("Mr. Davis") provides the learning and pedagogical feedback, in addition to the help on domain knowledge (implemented in the previous version), to move the user away from dealing with the local changes and focus more on a global understanding of the domain. The mentor's goal is to emphasize pedagogical and learning issues in addition to providing help on domain knowledge and concepts. The agent's feedback has been redesigned to guide students in their learning tasks as opposed to directing them to correct errors in their concept maps. In other words, the feedback helps students gain a better understanding of the global structure of the domain and how to regulate their learning in the environment.

The **Interfacer Agent** acts as a middleman between the graphical user interface and the other agents to simplify the individual agent- design. In this manner, the graphic user interface components and the agent mechanisms can be developed independently.

The **Pattern Tracker Agent** plays the role of a "listener" in the environment. It looks at the student's and other agents' activities and combines them into predefined patterns. Betty and the Mentor agent use these patterns to interact with and provide feedback to the student.

Figure 5.55 displays the interactions of these agents. The user and the software agents interact through the environment. The interfacer agent acts as the representative of the environment on the software-agent side. The interfacer agent communicates the user's actions in the environment to the other three agents. If the pattern tracker finds any set of activities that matches its pre-defined patterns, it notifies the teachable and the mentor agents. Using this information and the single events they have been sensed, both agents respond to the user by signaling the interfacer to relay it to the environment. This relay of communication via the interfacer agent keeps the agent implementation independent from the graphical user interface.

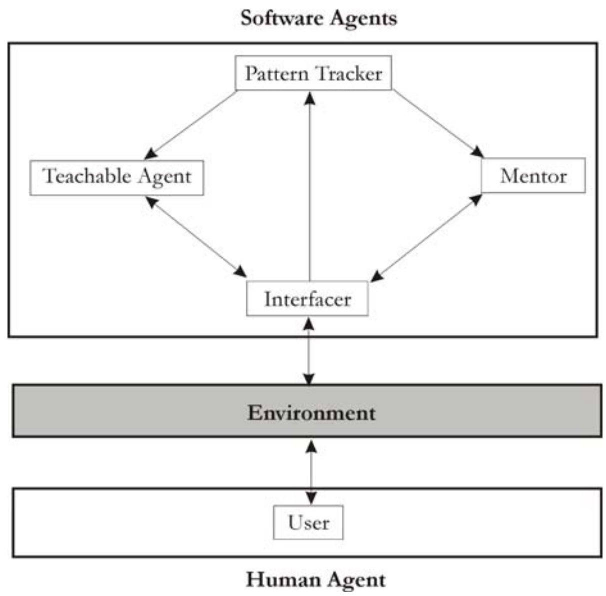


Figure 5.55 The Multi-Agent Structure of Betty's Brain

Figure 5.56 represents the implementation structure for software agents in the environment simplified from (Arkin and Balch 1997). An agent has four components:

1. Monitor: Senses actions performed by other agents in the environment.
2. Decision Maker: Processes the information received from the monitor to decide actions that may be performed by that particular agent in the environment or to other agents.
3. Executive: Communicates decisions made by the decision maker to the environment and to other agents.
4. Memory: Keeps information used by the Decision Maker, namely a history of past actions and states and stores information necessary to the agent.

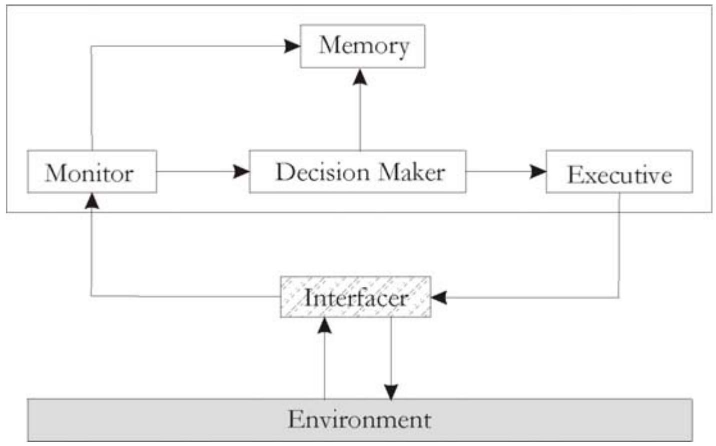


Figure 5.56 General Agent Architecture in the Betty's Brain Environment

The new computational architecture for Betty is illustrated in Figure 5.57. Betty has access to the actions performed by the user and the mentor agent as well as high-level patterns linked to learning and teaching strategies generated by the Pattern Tracker agent. Betty's memory module is modified to reflect this change. The Decision Maker module contains two components:

1. Methods for reasoning with the concept map to answer queries and generating explanations
2. Methods for generating meta-cognitive prompts based on patterns recognized by the Pattern Tracker

Betty is interested in self-assessment-related patterns. The students need to edit the concept map and ask at least one causal question before she agrees to take a quiz. Betty also encourages the use of causal links by responding to the first few additions of causal links in each session.

The new item in the decision maker module is decision making related to the meta-cognitive strategies. Betty interacts with students when she thinks they are using a trial-and-error approach to solving problems, such as send Betty to take a quiz, then switching the trend of a causal link, and then trying to send her to take the quiz again. In this case, she suggests a possibility of using a cognitive pattern, such as asking her the quiz questions and reflecting on her answer to see make sure if she can answer the questions better.

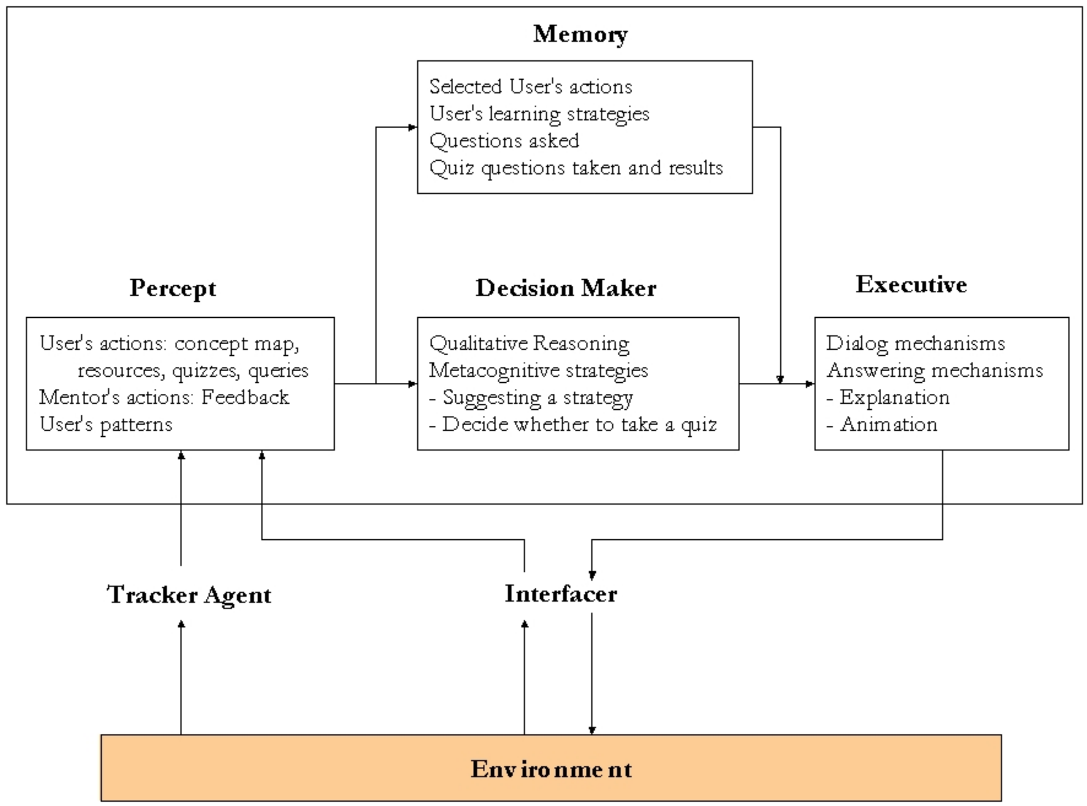


Figure 5.57 Betty's Architecture

The mentor agent's architecture, shown in Figure 5.58, is separated from the main environment instead of being embedded in it as in the previous version. The Mentor agent monitors students' interactions with Betty and also keeps track of Betty's performance in the quizzes. When students ask the Mentor agent for help, it uses pre-defined patterns to decide if it should give help in the requested category or any other category and, if so, at which level of detail. For example, a student conducts activities that show a trial-and-error approach to solving problems, such as sending Betty to take a quiz, switching a trend of a causal link, and then sending Betty to take a quiz again. In this case, the Mentor will suggest a possibility of using a cognitive pattern, such as asking Betty the quiz questions and reflect on her answer to see if there is any error.

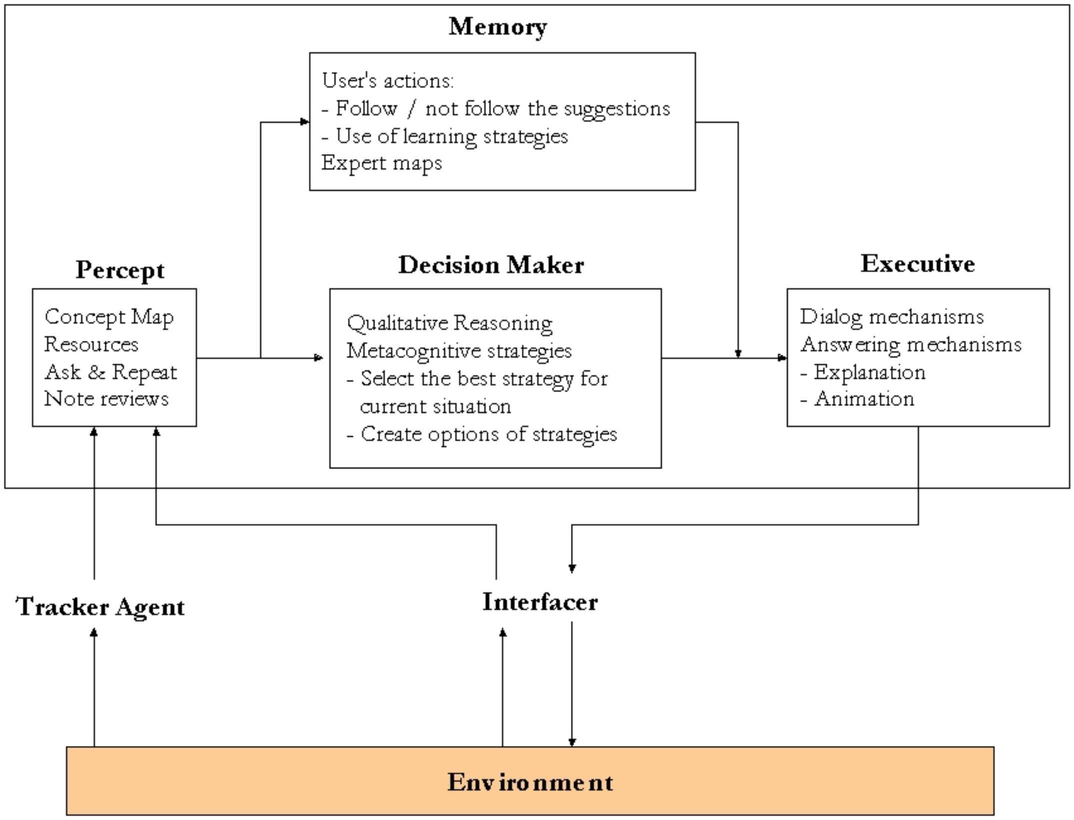


Figure 5.58 The Mentor Agent's Architecture

The interfacer agent's architecture, displayed in Figure 5.59, is a simple reflective one; it responds as soon as it senses an action in the environment.

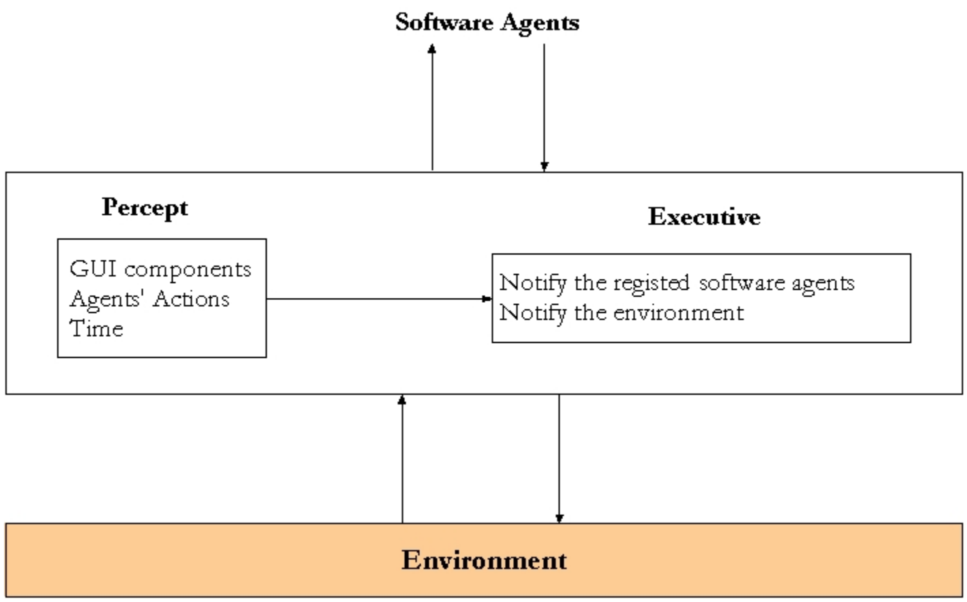


Figure 5.59 The Interfacer Agent's Architecture

The pattern tracker also listens to the actions in the environment, but its behavior, shown in Figure 5.60, is more complex than that of the interfacer agent. Because the task of Pattern Tracker agent is to detect possible learning strategies the user has employed as defined in Appendix E, it needs to store potential patterns in its memory. When new actions arrive, it tries to fit them with the existing partial pattern structures. Once a pattern is confirmed, the agent notifies this information to the other agents in the system.

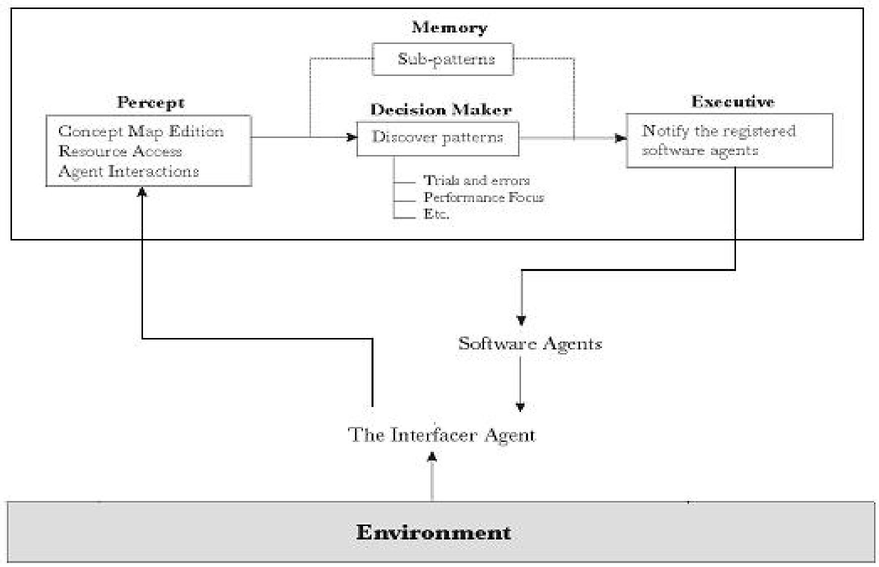


Figure 5.60 The Pattern Tracker Agent's Architecture

The result of the redesign and reimplementation was a new version of Betty's Brain that implements self-regulated learning. The next chapter discusses experiments conducted with this new self-regulated-learning version of the Betty's Brain environment to demonstrate the effectiveness of the learning-by-teaching environments toward self-regulated learning.

CHAPTER VI

METHODS

This chapter discusses the experimental studies that were conducted to demonstrate the effectiveness of learning-by-teaching environments along two different dimensions. One goal of this research was to discover if students gain a deeper understanding of domain concepts and their interrelationships in complex science domains, such as river ecosystems when they are put in a situation where they have to teach a computer agent. Another goal of these studies was to discover if, by exposing students to self-regulated learning strategies, students would learn more and be better able to transfer knowledge and strategies they learned while using the Betty system to other domains.

While the previous chapter discussed in detail the different components of the Betty's Brain environment, this chapter focuses on the design of the experiments and the data analysis methods. The *Experimental Design*, section presents the framework for our comparative study of learning gains for three different conditions. The measures of learning gains and transfer are also defined. The *Analysis Procedures* section describes the methods that are used for collecting and grading data collected from students' activities on the system. Finally, the *Data Analysis* section explains the statistical tests conducted to verify the significance of the claims made in this thesis.

Experimental Design

The goal of the study was to measure the effectiveness of a learning-by-teaching system that explicitly incorporated self-regulated learning feedback as students worked with the system. This system (presented in Chapter 5), called the **SRL** system, was compared against two other systems.

LBT: Basic Learning-by-Teaching environment (presented in Chapter 4)

ITS: A directed-learning environment that emulated some aspects of intelligent tutoring systems

Students' learning was directed by Mr. Davis, the mentor agent, in the ITS environment. Students did not teach Betty. Instead Mr. Davis asked them to create a concept map that would answer a set of quiz questions. The mentor agent also provided focused feedback to help students correct their concept maps and get the quiz questions correct.

In the LBT and SRL systems, students explicitly taught Betty by creating a concept map. They were told that their goal was to teach Betty so that she could pass a test and join the high school science club (details discussed in Chapter 4). The system for the LBT condition had all of the basic learning-by-teaching features. Students taught Betty by creating concept maps. They could query Betty to see how well she was learning what they had taught her, and get her to take quizzes with Mr. Davis to determine how much she had learned about the domain. Mr. Davis also provided focused feedback after Betty took a quiz. This feedback was the same as that provided for the ITS system.

The system for the SRL participants was based on a guided learning approach. As discussed in Chapter 5, the teachable agent was more responsive during the teaching process. The mentor agent provided pedagogical and learning on-demand feedback that was tailored to students' current activities on the system. The domain feedback that SRL participants received was also different from that provided to the ITS and LBT groups in that it was based on chains of events rather than focusing on single concepts or links.

Research Questions and Hypotheses

Given these three systems, we formulated the following research questions and hypotheses.

Research Question 1: Will students who learn by teaching a computer agent exhibit significantly greater understanding and ability to transfer in science domains that involve complicated reasoning processes than students who have been taught by a pedagogical agent?

Comparing the learning-by-teaching treatments to the directed-learning treatment, we hypothesized that learning by teaching environments would be more effective than directed learning environments. The learning-by-teaching environments were grounded in an active, constructivist practice that allowed students to develop and apply learning strategies in a domain specific context. The focused-feedback environments (ITS and LBT) were not quite as learner-centered as the SRL environment because the mentor agent in those systems, tended to provide input on how to correct specific errors students had made in their concept maps without taking into account how the student had learned this knowledge. Thus, the locus of control moved away from the students, especially in the ITS system. The mentor agent provided immediate feedback, and though this often results in short-term improvements in learning outcomes in traditional classroom settings, it is possible that the directed learning environment does not encourage or guide participants to explore and develop their own ideas about the domain content.

Research Question 2: Will students who use a computer-based environment where the agents promote social interaction such as expressing, encouraging, and demonstrating the use of self-regulated learning strategies exhibit significantly greater understanding and ability to transfer in a science domain that involves complex reasoning processes than students who do not see self-regulated learning behaviors modeled?

These studies also compare the effect of feedback provided by the LBT and SRL systems. We conjectured that feedback that focuses on self-regulated learning skills helped students learn more effectively in scientific domains. Data collected in the last study of our learn-by-teaching systems showed that students had some difficulties evaluating the importance of information and reflecting on feedback received from the mentor agent (Leelawong, Davis et al. 2002; Davis, Leelawong et al. 2003). This kind of feedback, used in the previous version of the Betty's Brain environment, is grounded in problem-based learning, where users learn by receiving feedback while solving a set of problems. This feedback is typically provided on a problem-by-problem basis. However, domains with complex knowledge structures cannot be easily learned piecemeal, one problem at a time, since understanding complex systems requires learners to develop a global view of interactions between entities in the system. The underlying phenomena that govern the answers to a set of questions about complex, dynamic systems require the analysis of multiple paths that determine the effect of one entity on another. Eventually learners must come to understand that answers to questions about interdependent and changing relationships between entities can affect answers to questions about other entities in the system.

An example of such a complex, dynamic system is a river ecosystem. The processes within this system, such as the oxygen-carbon dioxide cycle, involve direct relationships between dissolved oxygen, plants, and the animals that live in the river. However, when considering other processes in the system, such as waste decomposition and the food chain, a number of interactions involve waste, bacteria, and nutrients- and these in turn affect dissolved oxygen, plants, animals, and other entities that participate in the oxygen cycle. Therefore, the explanation structure is intricate, and in some cases, the overall effect of one entity on another may change when one considers all the interactions.

Self-regulated learning should be an effective framework for providing feedback in a learning-by-teaching situation of a complex scientific domain because of several reasons. Self-regulated learning strategies are among the highest form of higher-order cognitive skills (Corno and Mandinach 1983), and these skills are critical to the development of problem solving ability (Pintrich and DeGroot 1990). In addition, cognitive feedback is more effective than outcome feedback for decision-making tasks (Butler and Winne 1995). Cognitive feedback informs students about ways to monitor their learning and identify their needs (achievement relative to goals) while guiding them to achieve their learning objectives (cognitive engagement by applying tactics and strategies).

A common way to learn is by observation (Bransford, Brown et al. 2000). By modeling self-regulated learning strategies through the software agents, we expect that students will take notice and begin to develop similar strategies when using the computer-based learning environment. However, the main issue seems to be how to design these agents so that the students are motivated to learn these strategies. If the social interaction between the agent and the student is not designed properly, learning may suffer because of a variety of reasons—the students may experience frustration in understanding the agents, or the system may distract students in a way that gets them off-track in terms of learning about the domain. Because there are few studies in the literature about software agents used to teach self-regulated learning in learning-by-teaching situations, this research serves as an initial experiment for evaluating the value of such agents in computer-based learning-by-teaching systems.

Adding more interactions to the agent does not imply that students would learn more. On the contrary, if the design of the environment facilitates the learning process by modeling for students the strategies that aid in the acquisition of new knowledge, this should also assist them in developing good pedagogical strategies. Although there have been other studies with computer-based learning environments that support self-regulated learning while teaching (Armstrong 1998), the use of software agents that model and explain self-regulated learning strategies in a learning-by-teaching situation is new. Therefore, this study can provide preliminary data about the effects of including such agents in computer-based learning-by-teaching environments.

Based on these two research questions, we formulate the following hypothesis:

Hypothesis: Computer-based learning-by-teaching environments (the SRL & LBT groups) are more effective in helping students to gain a deeper understanding and ability to transfer in science domains than the system where the learning is directed by a tutor agent (the ITS group). In addition, the computer-based learning-by-teaching environment with feedback based on self-regulated learning strategies (the SRL group) is more effective in helping students gain better understanding and develop greater ability to transfer in science domains (the LBT group).

In other words, this dissertation hypothesizes that the SRL environment is the most effective in helping students achieve effective learning. The LBT environment is hypothesized to be the second most effective. Finally, the ITS environment is hypothesized to be the least effective.

As stated in Chapter 2, effective learning can be measured by understanding and ability transfer. In addition, deep understanding normally results in better retention of knowledge. The ability to transfer includes the capabilities to apply the knowledge and strategies learned in other problem-solving situations and domains. Therefore, the learning of strategies is as important as the learning of domain knowledge. We divided this hypothesis into three sub-hypotheses.

Hypothesis 1: (Understanding of the Domain) The SRL group will gain better understanding of the domain content than the LBT and SRL groups, and the LBT group will gain better understanding than the ITS group. The understanding will be measured in three areas:

<u>Hypothesis 1.1</u>: The SRL group will show greater improvement from the pretest to the posttest than the LBT and SRL groups, and the LBT group will show greater improvement than the ITS group.

<u>Hypothesis 1.2</u>: The SRL group will develop knowledge structures that are richer than the LBT and SRL groups, and the LBT group will develop richer knowledge structures than the ITS group.

<u>Hypothesis 1.3</u>: The SRL group will show better retention of the domain knowledge than the LBT and SRL groups, and the LBT group will show better retention of the domain knowledge than the ITS group.

Hypothesis 2: (Learning Behaviors) The SRL group will demonstrate learning behaviors that are more independent than the LBT and SRL groups, and the LBT group will demonstrate learning behaviors that are more independent than the ITS group.

Hypothesis 3: (Ability to Transfer) The SRL group will exhibit greater ability to transfer their learning strategies to another domain than the LBT and ITS groups, and the LBT group will exhibit greater ability to transfer their learning strategies to another domain than the ITS group.

<u>Hypothesis 3.1</u>: The SRL group will develop knowledge structures that are richer than the LBT and SRL groups, and the LBT group will develop richer knowledge structures than the ITS group.

<u>Hypothesis 3.2</u>: The SRL group will exhibit more strategic learning-behaviors than the LBT and SRL groups, and the LBT group will exhibit more strategic learning-behaviors than the ITS group.

Sample

Fifty four fifth-grade science students from two classrooms in a public school in Southeastern United States participated in the study. Compared to the district average, these students are classified from medium to high-achieving students. All students were required to submit a signed parent-consent form before they began the study.

Procedures

	R	O_1	X_1	O_1	O_2	O_3
C	R	O_1	X_2	O_1	O_2	O_3
	R	O_1	X_3	O_1	O_2	O_3
	R	O_1	X_1	O_1	O_2	O_3
C	R	O_1	X_2	O_1	O_2	O_3
	R	O_1	X_3	O_1	O_2	O_3
	R	O_1	X_1	O_1	O_2	O_3
C	R	O_1	X_2	O_1	O_2	O_3
	R	O_1	X_3	O_1	O_2	O_3

Figure 6.1 Experiment Design

This study was designed as a randomized regression-discontinuity experiment (Trochim 2001) using a stratified random sampling method. The participants' previous academic performances were used to create three heterogeneous groups. Figure 6.1 illustrates the design of this study in a notational form where:

C = Group Assignment using cutoff scores of the means of participants' standardized test scores in mathematics and reading (see details below)

R = Random Group Assignment after controlling for achievement. There were three groups containing an equal number of participants in this study (see details below)

O_1 = Paper test on the river ecosystem knowledge (Appendix F) to test the participants' knowledge about river ecosystems and the Motivated Strategies for Learning Questionnaire (Appendix G) to assess the participants' motivation and use of learning strategies

O_2 = Memory test: Participants recreated the concept maps they had at the end of the main study in limited-feature versions of the systems assigned to them during the main study

O_3 = Transfer test: Participants used limited-feature versions of the systems assigned to them during the main study to create a concept map about the land-based nitrogen cycle

X_1 = **SRL**: The learning-by-teaching environment with self-regulated-learning feedback

X_2 = **LBT**: The baseline learning-by-teaching environment

X_3 = **ITS**: The directed-learning environment without the learning-by-teaching and self-regulated learning features

We divided participants into three groups using cutoff scores from standardized tests taken from the previous year. These scores reflected their academic achievement in reading and mathematics. Because students were required to read and understand a significant amount of domain material during the experiment, their reading ability was taken into account. In this way, we avoided differences in performance level that could be attributed to groups. The procedure for creating groups was as follows. First, we calculated the mean scores for all students in both subjects. Based on these scores, students were separated into three stratified groups. The first group had both mathematics and reading scores above the class means, the second group had both scores under the class means, and the last group had the mathematics score above the subject mean and the reading score under the subject mean, or vice versa. From each group, we them assigned students equally to the three experimental groups by random assignment.

At the beginning of the experiment, all students took the domain-knowledge pretest and the Motivated Strategies for Learning Questionnaire (MSLQ). The LBT and SRL groups were introduced to Betty and the motivating story of her reason for wanting to learn about river ecosystems. All three groups were also given a preliminary tutorial on concept maps and the reasoning processes. Then the participants worked with the assigned systems for five sessions of fifty minutes each. Figure 6.2 shows the features of the environment available to each group in the main study. At the end of the main study, the participants took the same posttest and questionnaire.

The details of these components were described in Chapters 4 and 5. In particular, in the main study the LBT group used the basic learning-by-teaching system presented in Chapter 4. Students in this group had the limited versions of the teachable agent (no self-regulated-learning strategies) and the mentor agent (providing directed as opposed to guided feedback).

Features	Concept Map Editor	Query	Quiz		Resources	Betty		Mentor			
			Self-paced	Controlled		Visualization	Learning Help	Visualization	Pedagogical	Learning	Domain
SRL	×	×	×		×	×	×	×	×	×	×
LBT	×	×	×		×	×					×
ITS	×	×		×	×			×			×

Figure 6.2 Features available in the Main Study

The system used by the ITS group had the same functionality as that for the LBT group, except the teachable agent was not present in the system. Students in this group were directed by the Mentor agent, who asked them to construct concept maps to answer quiz questions. Therefore, social interactions between Betty and the participants, and the notions of shared representation and responsibility, did not exist in this condition. The students in the ITS group composed causal queries to the mentor agent instead of Betty, and the "tell-me" query template was excluded. Moreover, students in the ITS group took a quiz to see how good their concept maps were, but the LBT and SRL groups sent Betty to take a quiz to see how well they had taught her.

The SRL group received the full version of the system, which included the self-regulated teachable agent and the on-demand Mentor agent. The SRL Betty could initiate short conversations with students about her current learning experience and her feelings on her performance in the quizzes. The original mentor agent used for the ITS and LBT groups gave domain-related feedback with quiz results based on the Expert concept map. In addition to this feature, the SRL Mentor agent provided pedagogical and learning feedback that was tailored to students' current activities in the system, and the help was provided on-demand instead.

The memory test took place six weeks after the main study ended, and lasted for one fifty-minute session. In this test, students started with an empty concept map and attempted to recreate from memory the map that they had created during the main study. The participants used the same systems they had used in the main study except that only the concept-map editor was enabled (see Figure 6.3). However, we maintained the interface of the environments because suddenly introducing new interfaces might cause confusion. Therefore, the agents' presentations were present but they did not participate in the process.

Features	Concept Map Editor	Query	Quiz		Resources	Betty		Mentor			
			Self-paced	Controlled		Visualization	Learning Help	Visualization	Pedagogical	Learning	Domain
SRL	×					×					
LBT	×					×					
ITS	×							×			

Figure 6.3 Features available in the Memory Test

After completing the memory test, the students worked on the transfer test for two fifty-minute sessions. All students were asked to create a concept map that included the entities involved in the land-based nitrogen cycle and their interrelationships. Again, all three groups used the same versions of the system for the main study with no help from the agents other than the outcome feedback from the quiz feature. In addition, there was only one quiz with three unscaffolded questions in this phase. The features enabled in the versions used in the transfer test are summarized in Figure 6.4.

Features	Concept	Query	Quiz	Resources	Betty	Mentor

	Map Editor		Self-paced	Controlled		Visualization	Learning Help	Visualization	Pedagogical	Learning	Domain
SRL	×	×	×		×	×					
LBT	×	×	×		×	×					
ITS	×	×	×		×			×			

Figure 6.4 Features available in the Transfer Test

Measures

There were four categories of measures for the three hypotheses. These measures were described in the sub-sections below.

1. Domain knowledge
2. Learning behaviors
3. Retention of Knowledge
4. Ability to Transfer

<u>Domain Knowledge</u>

Two measures were used to study students' domain-knowledge gains as a result of the interventions:

1. Improvements in students' scores from the pre-test to the post-test.
2. Improvements in the quality of the students' concept maps as they had progressed in the study

The paper-based pre-test and post-test had an identical set of questions that included open-ended, multiple-choice, and event-ordering questions (see Appendix F). A brief description of each question appears below.

Question 1 (Open-ended): Definition of Interdependence

Question 2 (Open-ended): Definition of Balance

Question 3 (Open-ended): Definition of Chain of events

Question 4 (Multiple-choice): Role of bacteria

Questions 5 – 12 are in the context of a self-sustaining ecosphere, a small, sealed glass container with a number of brine shrimp, algae, bacteria, and air. Question 5 had 16 sub-questions.

Question 5 (True-or-False): How a change in one entity might affect others in the ecosphere

Question 6 (Ordering): The chain of events that would take place if too much algae was added to the ecosphere

Question 7 (Ordering): The chain of events that might occur if no algae was introduced into the ecosphere

Question 8 (Ordering): The chain of events when too much bacteria in the ecosphere were added into the ecosphere

Question 9 (Ordering): The chain of events when there were too many macro-invertebrates in an ecosphere

Question 10 (Ordering): The chain of events when no macro-invertebrate were added into the ecosphere

Question 11 (Ordering): The chain of events that would occur if the ecosphere was always in a lighted environment

Question 12 (Ordering): The chain of events that would occur if the ecosphere was always in a dark (no sunlight) environment

In addition to the pre- and post-tests, we used a systematic procedure to grade the students' concept maps at the end of each session to determine the numbers of concepts and links that:

1. Corresponded to those in the expert concept map
2. Existed in the system's resources on the river-ecosystem domain but did not appear in the Expert concept map

The expert concept-map is shown in Appendix H. Note that students never saw the expert map.

Learning Behaviors

There were two types of measures in this category:

1. The Motivated Strategies for Learning Questionnaires (MSLQ) that was given along with the pre- and post-tests
2. The frequency with which students employed learning strategies in each session using the Betty's Brain environment: Log files were checked to find out the frequency with which participants employed the following learning strategies in each session.
   2.1. *Seek information*: Accessing on-line resources and asking the Mentor agent for help
   2.2. *Monitor*: Asking the teachable agent queries and sending the agent to take the quiz

Retention of Knowledge (Memory Test)

Participants completed a measure of how much of the knowledge about river ecosystems gained in the main study they could recall. We measured the following variables:

1. *Recollection*: How many concepts and links participants could recall compared to how many they had in their concept maps in the main study
2. *Accuracy*: The quality of recalled concepts and links in terms of how many correct and incorrect concepts and links were recalled (see the details of score calculation in the grading procedures)

3. *Maturation*: How many valid and invalid objects students forgot from the main-study concept-maps and how many new valid and invalid concepts and links students had in their memory-test concept maps

Ability to Transfer

For the transfer study, students used a system similar to the one used in the main study, but the domain of study was changed from river ecosystems to the land-based nitrogen cycle. The measures described for the main study were also applied to the transfer test.

Grading Procedures

Because most of the participants' data are qualitative, we applied coding schemes to translate the data to comparable quantitative scales. This section details the grading procedures for each of the measures described in the last section. To ease the grading effort and reduce human errors, computer programs, discussed in detail in Appendix I, were developed to assist in the grading process.

All of qualitative grading involved two human graders. Each grader separately graded about twenty percent of each data set first. Then their grading results were checked for consistency. If the consistency level was better than 80%, the graders assigned grades to the rest of that data set independently (without more checking for consistency). Otherwise, the graders needed to reconcile their differences, and repeat this process until the consistency level was reached.

Grading the Domain-Knowledge Paper Test

The first three questions were open-ended and asked for definitions and examples of three key features usually found in natural dynamic systems. The list below includes correct answers.

1. *Interdependence*: Every entity in the river ecosystem depends on other entities in the system to live, grow, and reproduce. For example, fish depend on plants to produce the dissolved oxygen they need to survive.
2. *Balance*: A system is in balance when there is not too much or too little of something that affects interdependence in the ecosystem. For example, if there is too little dissolved oxygen, fish will die.
3. *Chain of events*: A change in an entity triggers changes in other entities in the system, and this effect is then propagated to another set of entities. For example, when there are too many plants, there will be less sunlight that reaches the plants below the surface of the water in the river. When sunlight is blocked, some of these plants die. This increases the amount of dead organisms in the river. As a consequence, the amount of bacteria also increases (because bacteria feed on dead organisms). A chain is defined as containing at least two consecutive cause-effect events, such as $A \rightarrow B \rightarrow C$.

Open-ended answers were first categorized as one of the following types:

- General definition: Describing the given term without specifically referring to any entities in the domain
- Example: Giving a scenario that by example illustrates the phenomena in the given domain
- Others: Any other answer

After this initial categorizing, answers were further graded as correct, partially correct, or incorrect. Only the correct answers received points as shown in the diagram in Figure 6.5.

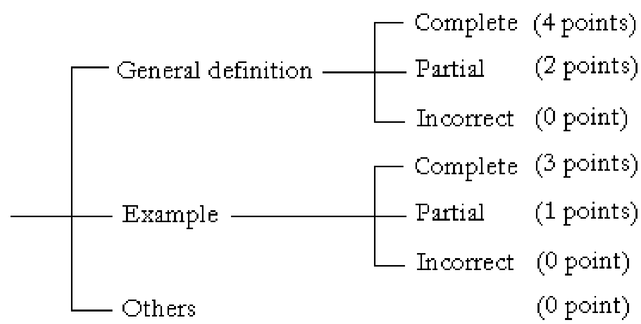


Figure 6.5 Grading Answers to Open-ended Questions

The grading procedure for multiple-choice and true-or-false questions was simple—one point if students marked the correct choice and zero points, otherwise. Students' answers were entered into two comma-delimited spreadsheets, each by a different grader. Then, the spreadsheets were checked for consistency of the entered data to ensure the accuracy of manual input.

The score for the ordering questions was calculated by the Longest Common Subsequence algorithm (Hirschberg 1977) (the `LCSorderingGrader` class). The score given to an order was the number of positions that have the same sequence as those in the correct order. For example, an order "3, 2, 1, 4, 5" received the score of 3 given the correct order "1, 3, 2, 5, 4" because the longest common sequence in the answer was "3, 2, 4". If the student explicitly specified that he did not know the answer to a particular question, such as by writing "I don't know" for the answer, he was awarded zero points for that answer.

Grading Concept Maps

Many of the steps for grading concept maps were implemented as Java programs that were compiled into the `mapGrader` package (see Appendix I). These programs were designed to be independent from concept-map versioning; they use text files as the input and output formats. There were seven steps in grading concepts and links:

1. Run the `ConceptMapConverter` program to convert concept maps to comma-delimited text files, one for concepts and one for links.

2. Run the concept- and the link-database programs, `ConceptDatabase` and `Link-Database`, respectively, to grade concepts and links. These programs load the existing grades from text files (also in the same format as the input and output files) and use them to assign grades to the concepts and links in the given files according to the algorithms discussed later in this sub-section. Each object in the output file is labeled as either "Done", "Graded", or nothing. See the assignments algorithms for concepts and links below.
3. The human graders can ignore the objects marked with "Done", but should check the "Graded" codes. Coders must grade the objects with no status marks. The details of assigning status to the objects are discussed below.
4. Run the `GradeCombiner` program to check if there are any differences in grades between coders.
5. If differences are found, the human coders resolve their differences in the grade assignments.
6. At the end of this process there is only one copy of the grades for each set. This file is added to update the appropriate database. The database reports if any new grades cause conflicts with the existing grades.
7. Count concept and link grades by category for each participant using the `ConceptGradeCounterNoSyns` and `LinkGradeCounterNoSyns` programs, respectively.

To assign a status for a concept:

1. If its label is the same as one of the objects in the concept database, assign the grade "Done".
2. If it is a synonym of one of the objects exist the concept database, assign the grade "Graded".
3. Otherwise, do not assign the grading status to the concept.

Assigning a status for a link is more complicate than that for a concept. The link database stores three sub-databases, one for each type of links (causal, type-of, and descriptive). The causal-link database indexes its link by its source and target concepts' labels and its trend. The type-of-link database indexes its link only by its source and target concepts' labels because its link label is fixed ("type-of"). The descriptive-link database is indexed by its source and target concepts' labels and its link label.

1. If a link of the same type exists in the link database with the same source and target concepts' labels as this link:
    - 1.1. If they are causal links:
        - 1.1.1. If they have the same trend, assign the grade "Done".
        - 1.1.2. Otherwise, assign the grade "Graded".
    - 1.2. If they are type-of links, assign the grade "Done".
    - 1.3. If they are descriptive links,
        - 1.3.1. If they have the same label, assign the grade "Done".
        - 1.3.2. Otherwise, assign the grade "Graded".
2. If a link of the same type exists in the link database but the source and target concepts were switched compared to this link, assign the grade "Graded".
3. If a link of a different type exists in the link database with the same source and target concepts' labels as this link, assign the grade "Graded".
4. Otherwise, do not assign the grading status to the concept.

Concepts were graded before links because the concept grades were used to inform the grades for the links. Each concept was classified into four categories.

1. *Expert* (E): The concept is in the expert map.
2. *Relevant* (R): The concept is not a part of the expert map but is found in river ecosystems or in the on-line resources.
3. *Irrelevant* (I): The concept is not found in river ecosystems nor is found in the on-line resources.
4. *Uncodable* (U): The concept is incomprehensible or ambiguous.

If two concepts in the same concept map are synonyms to each other, the less general concept is graded as redundant and receives the tag "S" (synonym) in addition to grades. For example, assume that "macro-invertebrates" and "bugs" appear in the same student's map. Both concepts can be expert concepts, but "bugs" is a synonym of "macro-invertebrates". Therefore, the concept "bugs" received the "S" tag.

Link grading also had two steps, (i) judging the relevance to the domain and (ii) deciding the correctness of the relationship. The relevance of a link was determined from the relevance of its concepts according to Figure 6.6. For example, if the relevance of the source concept was "expert (E)" and that of the target concept was "irrelevant (I)", the relevance of the link was also "irrelevant (I)".

		Target Concept			
	Relevance	**E**	**R**	**I**	**U**
	E	E	R	I	U
Source Concept	**R**		R	I	U
	I			I	U
	U				U

Figure 6.6 Determining Link Relevance

Each link coded "expert" or "relevant" was further graded by the seven subcategories listed below.

1. *Valid* (V): the relationship is true in this world.
2. *Reversed direction* (XD): the link direction is switched, otherwise is would be valid.
3. *Wrong typed* (XL): the link is valid, but it was assigned the wrong type of link by the participant.
4. *Missing Mechanism* (M): (for casual links only) the relationship corresponding to the link is valid, but the relationship is indirect (there should be more concepts and links in between the two concepts).
5. *Reversed trend* (XT): (for casual links only) the relationship might happen in a river ecosystem if the trend (increase / decrease) is switched to the opposite one.
6. *Alternative concept* (AC): all other links that could not fit the categories above, including concepts and links that express an idea that is not correct according to current scientific models of river ecosystems.

Redundant links received the tag "S" (synonym) in addition to grades. Two links are redundant if they possess the same source and target concepts or their synonyms. They can be of the different types. Similar links with worse grades are classified as being redundant.

For the statistical (quantitative) tests described in Chapter 7, the grades assigned to concepts and links were summarized into three categories as shown in Figure 6.7. The six numbers in the table were captured for each of the students' concept maps and used for judging the validity of the map. The finer grained labels presented above were only invoked in situations where we decided to take a more detailed look at the data.

Report Category	*Concept Grading Category*	*Link Grading Category*
Expert	n_E	$n_{E\text{-}V}$
Valid	$n_E + n_R$	$n_{E\text{-}V} + n_{R\text{-}V}$
Others	$\sum_{concept} - (n_E + n_R)$	$\sum_{link} - (n_{E\text{-}V} + n_{R\text{-}V})$

Figure 6.7 Report Categories for Concepts and Links

For a student concept-map,
n_E = Number of the *expert* concepts in the map that are not synonyms of other concepts in the map
n_R = Number of the *relevant* concepts in the map that are not synonyms of other concepts in the map
$\sum_{concept}$ = Total number of concepts in the map
$n_{E\text{-}V}$ = Number of the valid, *expert* links in the map that are not synonyms of other links in the map
$n_{R\text{-}V}$ = Number of the valid, *relevant* links in the map that are not synonyms of other links in the map
$\sum_{link}$ = Total number of link in the map
Therefore, the "other" group of grades includes invalid, uncodable, and synonym concepts and links.

Learning Behaviors

In this study, we also collected data on students' activities in the Betty's Brain environment, and analyzed them to infer students' learning behaviors. Students' activities were recorded in log files, which were later analyzed to derive the following information for each session of a student:

1. Number of times the student composed queries for Betty
2. Number of times the student asked Betty to take a quiz (or take the quiz themselves in the ITS condition)
3. Number of times the student accessed the resources (including accessing the mentor's on-demand feedback for the SRL condition)
4. Amount of time the student spent in reading resources

Retention of Knowledge

The grading procedure for the dimension of retention of knowledge had three steps, (i) comparisons of the main-study and memory-test concept maps, (i) correctness grading, and (iii) accuracy grading (linking of the comparisons and the correctness grading). The comparison measures compute the number of concepts and links that students recalled in the memory test map from the concept maps in the main study. This comparison was done primarily by the `ConceptMapComparator` program, and its output was checked by two coders. These two coders then compared their results and resolved any conflicts.

The grading for correctness was the same as that of the concept-map grading described previously. Then, the correctness grades were used to refine the comparison grades to include validity because of the need to exclude misconceptions. This part was also done by the programs `ConceptGradeFinder` and `LinkGradeFinder` for concepts and links, respectively. All of the numbers generated for each student are summarized in Figure 6.8 and Figure 6.9.

Concept Map	*Classification*	*Grade*	
		Valid	*Others*
Main Study		v_1	o_1
Memory Test	Recalled	v_2	o_2
	Added	v_3	o_3

Figure 6.8 Memory-Test Grades for Concepts

Concept Map	*Classification*	*Grade*				
		Causal Links			*Other Links*	
		Valid	*Partially Valid*	*Others*	*Valid*	*Others*
Main Study		vcv_1	vcp_1	vco_1	vlv_1	vlo_1
Memory Test	Recalled	vcv_2	vcp_2	$vcop_2$	vlv_2	vlo_2
	Added	vcv_3	vcp_3	vco_3	vlv_3	vlo_3

Figure 6.9 Memory-Test Grades for Links

Using the computed variables in the two tables, three measures were derived as is shown below:

1. Recollection

	Score	*Ratio*
Concept	$v_2 + o_2$	$\frac{v_2 + o_2}{v_1 + o_1}$
Link	$vcv_2 + vcp_2 + vlv_2 + vlo_2$	$\frac{vcv_2 + vcp_2 + vlv_2 + vlo_2}{vcv_1 + vcp_1 + vlv_1 + vlo_1}$

2. Accuracy: Because causal links were important to modeling dynamic systems, positive causal-link pairs received more points.

	Score	*Ratio*
Concept	$v_2 - o_2$	$\frac{v_2 - o_2}{v_1 - o_1}$
Link	$3 \times vcv_2 + 2 \times vcp_2 + vlv_2 - vlo_2$	$\frac{3 \times vcv_2 + 2 \times vcp_2 + vlv_2 - vlo_2}{3 \times vcv_1 + 2 \times vcp_1 + vlv_1 - vlo_1}$

3. Maturation: This variable was a composite of several grade variables because the purpose of this variable was to check if participants in each group matured in a similar manner. The separated numbers could be analyzed further if this score was significant.
   score = (*added_valid* + *forgotten_invalid*) – (*added_invalid* + *forgotten_valid*)
   where:
   added_valid = $v_3 + vcv_3 + vlv_3$
   forgotten_invalid = $(o_1 - o_2) + ((vcp_1 + vco_1 + vlo_1) - (vcp_2 + vco_2 + vlo_2))$
   added_invalid = $o_3 + vcp_3 + vco_3 + vlo_3$
   forgotten_valid = $(v_1 - v_2) + ((vcv_1 + vlv_1) - (vcv_2 + vlv_2))$

Transfer Test

The grading procedure of this transfer test is exactly the same as that in the domain knowledge except that the domain of study changes to the nitrogen cycle.

Statistical Analyses

This section described the analysis of data for every sub-hypothesis. We conducted statistical analyses using the SPSS software for Windows version 11.5. The analyses of all the quantitative data was conducted assuming a general-linear-model (GLM) ANOVA test when the data for each independent variable satisfied the assumptions listed below (Gardner 2001).

- The data was interval-valued.
- The variances of the three groups (ITS, LBT, and ITS) were determined equal according to the Levene's test.
- The data for each group was normally distributed.

When these three assumptions were not met, the Mann-Whitney U test was used instead. Note that the MSLQ data was analyzed using ANOVA also. Even though the score for each question was based on a Likert scale, the data analysis was conducted on the summations of scores of several questions.

The GLM MANOVA tests were used for repeated-measures data, and we used Wilks' Lambda tests to examine group differences. The Wilks' Lambda test is most commonly used when there are more than two independent variables, and in our case, we had three conditions. Tukey's Honestly Significant Difference test was used as the post hoc test for the MANOVA and ANOVA tests.

If a participant did not participate in any phase of the study, his or her data was excluded from the data set for analysis. Some participants finished the main study toward the end of the fourth session. In this case, their concept-map data in the fifth session was the same as those of the fourth session, and for other measures the average value of the corresponding data from sessions 1 to 4 was used.

CHAPTER VII

EXPERIMENTAL RESULTS AND DISCUSSIONS

Forty four students from the class of original fifty four students completed every phase of the studies. Data gathered during the study has been analyzed using the procedures discussed in Chapter 6. As discussed in Chapter 6, the three primary hypotheses analyzed are:

Hypothesis 1: Students' understanding of the domain of river ecosystems.

Hypothesis 2: Learning behaviors during the learning and teaching process.

Hypothesis 3: Ability to transfer from one domain to another, especially when scaffolds are removed.

The first three sections of this chapter present the experimental findings by hypothesis. Then the last section aggregates all of the results and their implications, and draws conclusions about how our approach contributes to the design of effective intelligent learning environments. The grading and scoring procedures were discussed in Chapter 6.

Hypothesis 1: Understanding of the Domain

The hypothesis was broken down into three sub-hypotheses:

Hypothesis1.1: Performance gain from the pre-test to the posttest

Hypothesis1.2: Concept-map quality in term of the expert and valid concepts and links from the main study

Hypothesis1.3: Ability to retain knowledge after the main study was completed. We allowed for a delay of six weeks after the main study

The next three sub-sections reports the experimental results for these three hypotheses.

Hypothesis 1.1: Pre- & Post-test Results

The pre- and posttest administered to the students contained 12 questions (listed in Appendix F), but only six of the twelve questions was analyzed for this study. The details and reasons for choosing these six and not the others appear in Appendix I. Questions 1 to 3 were open-ended; they required students to explain the notion of interdependence, balance, and chain of events, respectively. Question 5 had five sub-questions that required True-False answers. The rest of the questions required students to put a given set of events in the right sequence. The scores obtained by the students for the six questions chosen for analysis are shown by group in Table 7.2. Difference in the pretest and posttest scores between groups was not significant for any of the six questions as shown in Table 7.3. (The reasons for using the Mann-Whitney U tests for this analysis appear in Appendix I).

Table 7.2 Pre-test and Post-test Results by group

Question	*ITS*		*LBT*		*SRL*	
	Pretest Mean (SD)	*Posttest Mean (SD)*	*Pretest Mean (SD)*	*Posttest Mean (SD)*	*Pretest Mean (SD)*	*Posttest Mean (SD)*
1. Definition of Interdependence (full score = 4)	1.07 (0.34)	1.80 (0.50)	0.94 (0.35)	2.2.5 (0.43)	1.23 (0.43)	2.51 (0.54)
2. Definition of Balance (full score = 4)	1.33 (0.43)	1.60 (0.41)	0.69 (0.35)	0.63 (0.32)	0.92 (0.40)	1.31 (0.40)
3. Definition of Chain of Events (full score = 4)	2.53 (0.34)	2.20 (0.39)	1.63 (0.42)	1.56 (0.39)	2.23 (0.46)	2.46 (0.40)
5: Roles of Bacteria & Macro-invertebrates (full score = 5)	3.60 (0.34)	4.20 (0.14)	3.00 (0.35)	4.13 (0.18)	3.31 (0.33)	3.69 (0.31)
6. Chain of Events: Extra Algae (full score = 8)	4.53 (0.27)	4.73 (0.37)	4.56 (0.29)	5.00 (0.43)	4.77 (0.23)	5.00 (0.48)
11. Chain of Events: 24-Hour Light (full score = 10)	6.40 (0.47)	6.13 (0.51)	6.44 (0.43)	6.13 (0.50)	6.00 (0.48)	6.08 (0.49)

Table 7.3 Significance Levels for Differences between the Pretest and Posttest Scores (Mann-Whitney U Tests)

Question	*ITS & LBT*		*ITS & SRL*		*LBT & SRL*	
	U	*Significance*	*U*	*Significance*	*U*	*Significance*
1. Definition of Interdependence	96.0	$p = .36$	92.0	$p = .82$	94.5	$p = .68$
2. Definition of Balance	108.0	$p = .66$	85.5	$p = .59$	74.5	$p = .20$
3. Definition of Chain of Events	102.5	$p = .50$	70.5	$p = .22$	87.0	$p = .48$
5: Roles of Bacteria & Macro-invertebrates	91.5	$p = .26$	89.5	$p = .72$	64.0	$p = .08$
6. Chain of Events: Extra Algae	108.0	$p = .65$	92.5	$p = .82$	91.5	$p = .59$
11. Chain of Events: 24-Hour Light	98.0	$p = .40$	94.0	$p = .89$	97.0	$p = .78$

However, when one looks at post-test gains for all the students, there is an overall improvement in the scores of four out of six questions, and the improvements are statistically significant for Questions 1 and 5 ($p < 0.05$). Figure 7.10 displays the overall average scores of the chosen six questions with the error bars displaying the 95% confidence intervals. Table 7.4 summarizes the results of the Mann-Whitney tests.

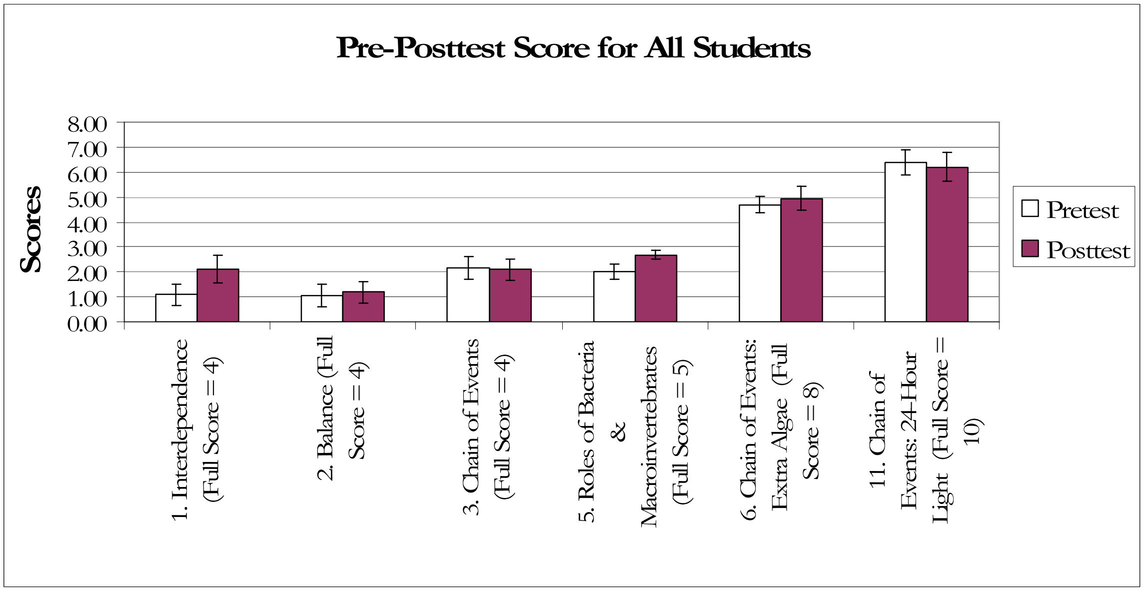


Figure 7.10 Average Pretest and Posttest Scores for all Students (Error bars represent the 95% confidence intervals of the differences between means.)

Table 7.4 Significance Levels for Pretest and Posttest Scores

Question	*U*	*Significance*
1. Definition of Interdependence	681.0	$p < .01$
2. Definition of Balance	891.0	$p = .47$
3. Definition of Chain of Events	942.5	$p = .83$
5: Roles of Bacteria & Macro-invertebrates	644.5	$p < .005$
6. Chain of Events: Extra Algae	899.0	$p = .55$
11. Chain of Events: 24-Hour Light	909.0	$p = .62$

Hypothesis 1.2: Quality of Concept Maps

The quality of the students' concept maps were evaluated at the end of each session of the main study to determine whether students' understanding of river ecosystem concepts improved as they worked with the pedagogical agents. Because no significant difference between groups was observed in the pretest (see Appendix I), the subsequent data analysis did not need a covariate.

Information Evaluation

Recall the concept map grading scheme from Chapter 6. Each concept in a student's concept map was graded as one of four categories—*Expert*, *Relevant*, *Irrelevant, and Uncodable.* Figure 7.11 shows the number of expert concepts, the number of relevant concepts, and the number of irrelevant and erroneous concepts in the students' concept maps at the end of the main study by group.

For the data analysis, the concept grades were grouped into three categories: (i) **expert,** (ii) **valid** (expert + relevant), and (iii) **other** (irrelevant + uncodable + synonymous). The ANCOVA results implied that the LBT group had significantly more expert concepts than the ITS group ($p < 0.05$) (see Table 7.5). The Mann-Whitney U tests were conducted to study the differences in the number of valid concepts between groups. The LBT-ITS and the SRL-ITS differences were significant, but the LBT-SRL difference was not (Table 7.5). The reasons for using different statistical tests for the two analyses are presented in Appendix J. The number of invalid concepts was not significantly different between groups.

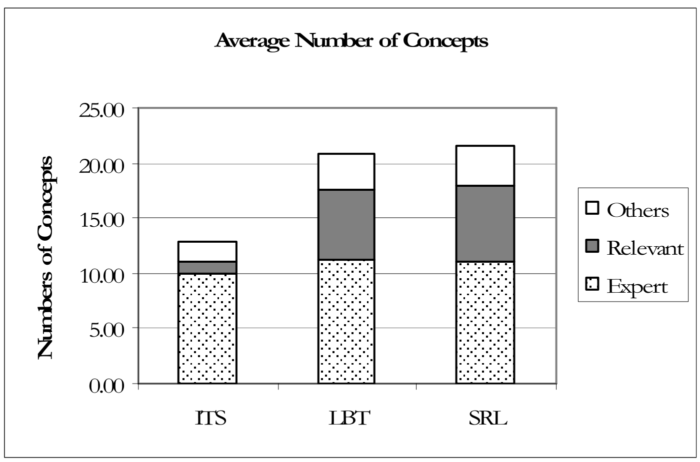


Figure 7.11 Concept Grades for Students' Final Ecosystem Concept-Maps

Table 7.5 Significance Levels for Concept Grades

	Expert Concepts	*Valid Concepts*	*Others*
Test	GLM ANOVA $F_{(2, 41)} = 5.72$, p = .006	*Mann-Whitney U*	*Mann-Whitney U*
ITS & LBT	Tukey's: p ≤ .05	U = 19.5, p < .0005	U = 81.0, p = .13
ITS & SRL	Tukey's: p > .05	U = 26.0, p < .001	U = 74.0, p = .29
SRL & LBT	Tukey's: p > .05	U = 99.0, p = .85	U = 97.0, p = .78

As discussed in Chapter 6, the link grades were also grouped into three categories:

- **Expert**: The number of correct *expert* links according to the grading scheme,

- **Valid**: The number of correct *expert* and *relevant* links according to the grading scheme, and
- **Other**: The number of links that did not belong to the other two types above. This includes *irrelevant*, *uncodable*, and synonymous links.

Figure 7.12 illustrates the link counts by these categories in the students' final concept-maps of the main study. See Table 7.6 for the details of the Mann-Whitney U tests. In summary, there was no significant difference in the number of expert links between groups in the final concept-maps. However, both learning-by-teaching groups created significantly more valid links than the ITS group ($p < 0.05$). There was no statistically significant difference between the LBT and SRL groups for the number of expert and the number of valid links. In addition, there was no significant difference in the number of *other* links between groups.

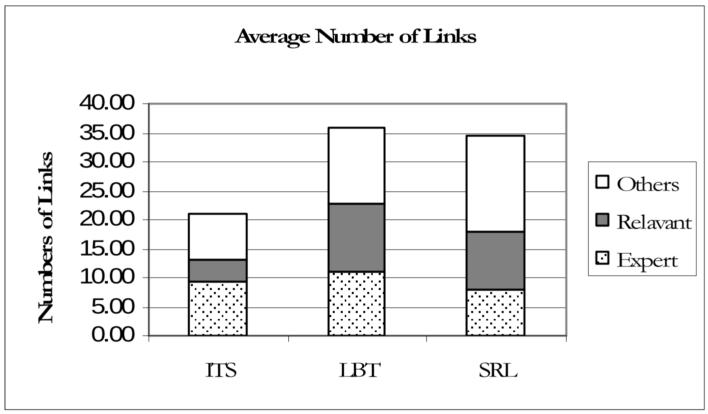


Figure 7.12 Link Grades for the final Ecosystem Concept-Maps

Table 7.6 Significance Levels for Link Grades (Mann-Whitney U Tests)

Groups	*Expert Links*	*Valid Links*	*Others*
ITS & LBT	U = 79.5, p = .10	U = 55.0, p < .01	U = 78.0, p = .10
ITS & SRL	U = 69.0, p = .36	U = 46.0, p < .05	U = 72.0, p = .25
SRL & LBT	U = 64.5, p = .09	U = 92.5, p > .62	U = 98.5, p = .81

All groups showed better performance in the quality of their concept map as they progressed from session 1 to session 5 in the main study. The average-per-group data for the expert and relevant concepts and links, illustrated in Figures 7.4 – 7.7 were analyzed by multiple analyses of variance, and the statistical results are shown in Table 7.7.

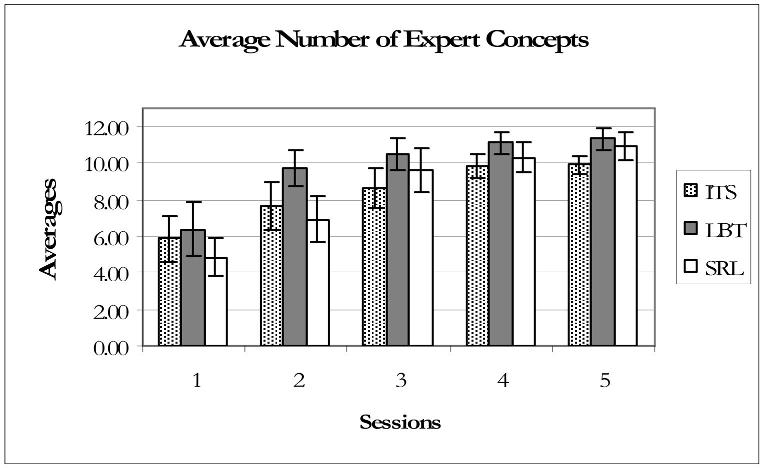


Figure 7.13 Average Number of Expert Concepts in Student Maps at the end of each session of the Main Study (Maximum Number = 13; Error bars represent the 95% confidence intervals of the differences between means.)

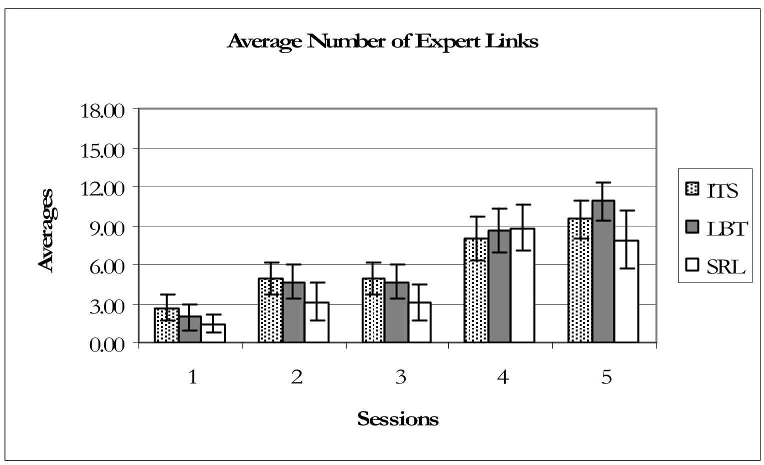


Figure 7.14 Average Number of Expert Links in Student Maps at the end of each session of the Main Study (Maximum Number = 18; error bars represent the 95% confidence intervals of the differences between means.)

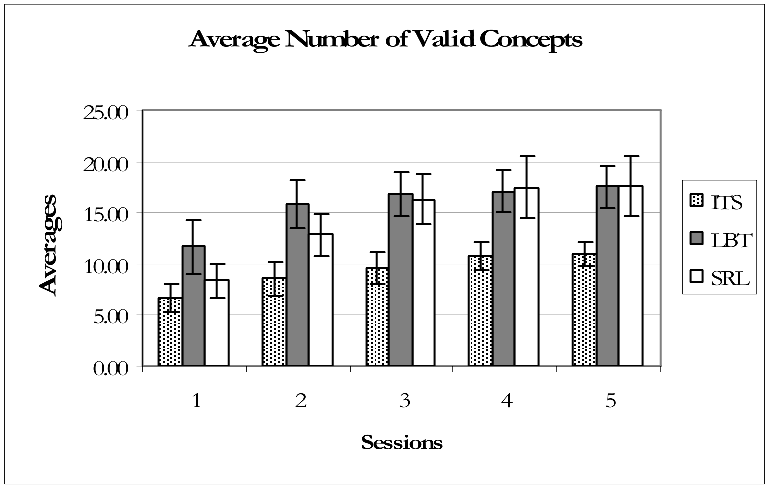


Figure 7.15 Average Number of Valid Concepts in Student Maps at the end of each session of the Main Study (Error bars represent the 95% confidence intervals of the differences between means.)

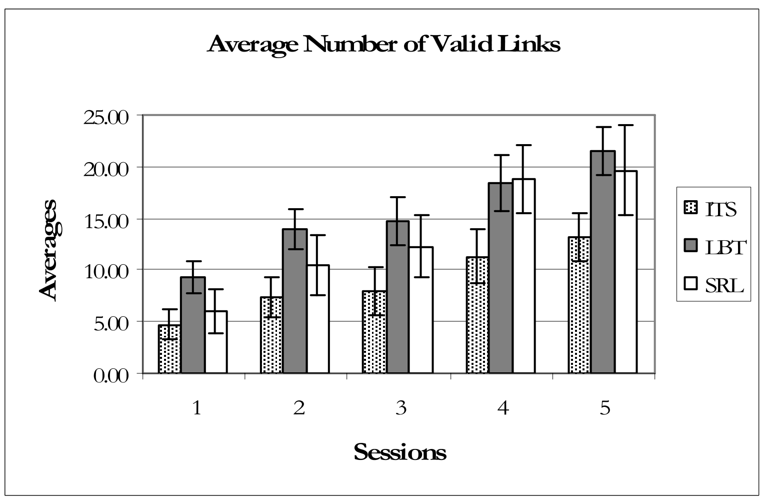


Figure 7.16 Average Number of Valid Links in the Student Maps at the end of each session of the Main Study (Error bars represent the 95% confidence intervals of the differences between means.)

All effects of time of measurement were significant as indicated in the *Time* row in Table 7.7. Also, there was a significant between-subjects interaction effect between time of measurement and group for the numbers of expert and valid concepts. The Tukey HSD post hoc analysis indicated that the LBT group identified significantly more expert concepts than the ITS group. In addition, both the LBT and the SRL groups had significantly more valid concepts and links than the ITS group. However, there was no significant effect between the three groups in the number of expert links.

Table 7.7 Significance Levels of Concept-Map Grades (GLM MANOVA tests)

	Expert Concepts	*Expert Links*	*Valid Concepts*	*Valid Links*
Time	$F_{(4, 38)} = 48.8$ $p < .0005$	$F_{(4, 38)} = 59.5$ $p < .0005$	$F_{(4, 38)} = 40.3$ $p < .0005$	$F_{(4, 38)} = 98.4$ $p < .0005$
*Time * Group*	$F_{(8, 76)} = 2.7$ $p < .05$	$F_{(8, 76)} = 1.1$ $p = .40$	$F_{(8, 76)} = 30.7$ $p < .001$	$F_{(8, 76)} = 2.8$ $p < .01$
ITS & LBT	Tukey: $p \leq .05$	Tukey: $p > .05$	Tukey: $p \leq .05$	Tukey: $p \leq .05$
ITS & SRL	Tukey: $p > .05$	Tukey: $p > .05$	Tukey: $p \leq .05$	Tukey: $p \leq .05$
SRL & LBT	Tukey: $p > .05$	Tukey: $p > .05$	Tukey: $p > .05$	Tukey: $p > .05$

An interesting observation is that the SRL group started with lesser number of expert concepts than the other two groups, but, by the end of Session 5, they had caught up with the other two groups (see Figure 7.13). The overall map quality of the SRL group was not worse than that of the LBT group, and the SRL group had significantly more valid concepts and links in their maps than the ITS group. Our conjecture is that the self-regulated learning components, the strategies that guided Betty's interactions and the mentor's feedback, may have slowed down the SRL group in the first two sessions. For example, the SRL group had the least number of expert concepts during the first two sessions but then had generated more expert concepts than the ITS group in the last three sessions.

At the end of the main study, the overall quality of concept maps of the SRL group was not worse than the LBT and ITS groups. This conclusion was further supported by the numbers of correct quiz answers, shown in Figure 7.17. The average quiz scores of the three groups significantly increased over time (GLM MANOVA, $F_{(4, 38)} = 32.8$, $p < .0005$) but the difference in the performance at the end of five sessions was not statistically different (GLM MANOVA, $F_{(8, 76)} = 1.1$, $p = .37$).

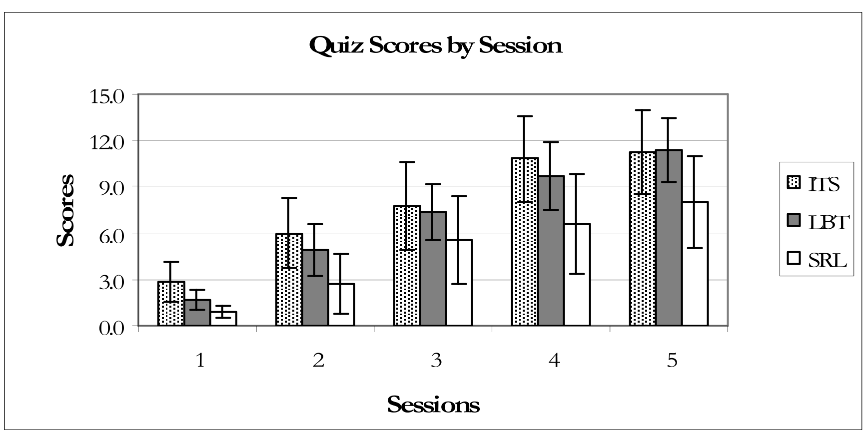


Figure 7.17 Average Number of Correct Quiz Answers by session (Maximum score = 15; error bars represent the 95% confidence intervals of the differences between means.)

More analyses of the components in students' concept maps confirmed that the ITS group relied mainly on the local feedback provided by the mentor agent. This was confirmed by Figures 7.9 – 7.10 that show the ratios of the expert concepts and links to the total number of valid concepts and links, respectively in the students' concept maps. There were significant differences in these numbers between the ITS group and the LBT and SRL groups. The post hoc test results appear in Table 7.7. Most of the valid concepts and links in the ITS students' concept maps were expert ones (the average for ITS students was 90% and 67%, respectively). As opposed to the trend seen in the ITS group, the LBT and SRL groups had as many or more expert concepts and links in their maps, but the percentage of expert concepts and links in comparison to the valid concepts and links was less than 65% and 40%, respectively. In other words, the LBT and SRL groups seemed to be learning on their own, and did not rely solely on the feedback they received from the mentor agent.

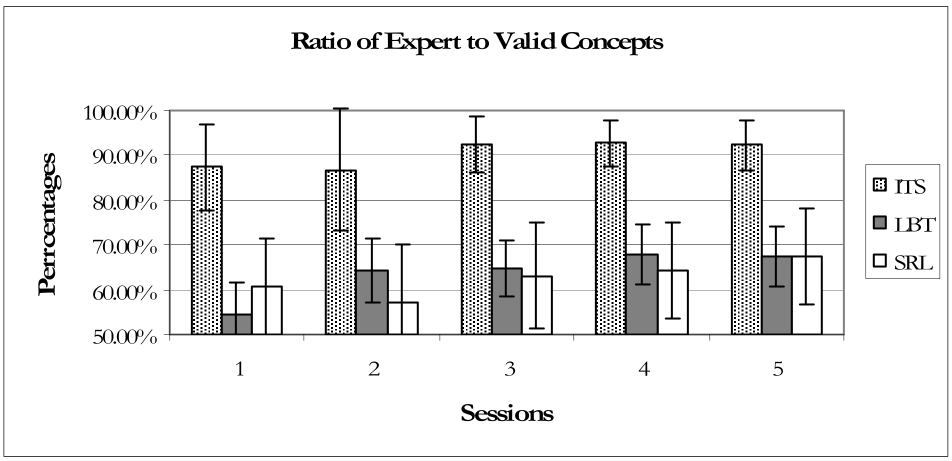


Figure 7.18 Ratio of the Number of Expert Concepts to the Number of Valid Concepts in Students' Concept Maps at the end of each session of the Main Study (Error bars represent the 95% confidence intervals of the differences between means.)

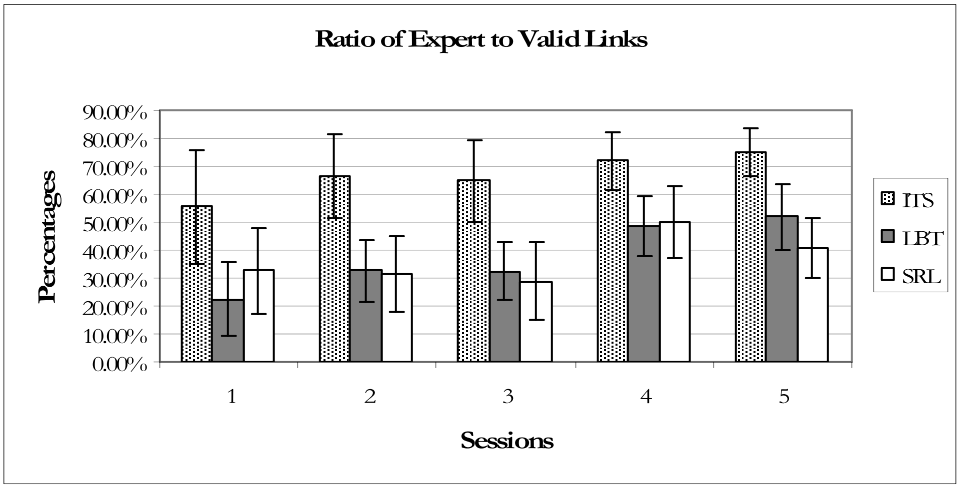


Figure 7.19 Ratio of the Number of Expert Links to the Number of Valid Links in Students' Concept Maps at the end of each session of the Main Study
(Error bars represent the 95% confidence intervals
of the differences between means)

Table 7.8 GLM MANOVA on the Ratio of the Expert Concepts and Links in Students' Concept Maps to the Numbers of Valid Concepts and Links in the Maps, respectively

	Ratio of Expert Concepts	*Ratio of Valid, Expert Links*
Time	$F_{(4, 38)} = 3.7, p < .05$	$F_{(4, 38)} = 6.3, p < .001$
*Time * Group*	$F_{(8, 76)} = 2.3, p < .05$	$F_{(8, 76)} = 0.7, p = .73$
ITS & LBT	Tukey: $p \leq .05$	Tukey: $p \leq .05$
ITS & SRL	Tukey: $p \leq .05$	Tukey: $p \leq .05$
SRL & LBT	Tukey: $p > .05$	Tukey: $p > .05$

<u>Summary</u>

Over time students in all three groups of the main study improved their knowledge about river ecosystems. This was partially demonstrated by the overall improvement in the posttest scores. In addition, the main study analysis of the students' concept maps showed that all three groups improved their concept maps and Betty's quiz scores over time. In terms of final quiz and scores, the ITS and LBT groups were ahead of the SRL group (see Figure 7.8) but the difference in the scores were not statistically significant.

Furthermore, the two learning by teaching conditions (SRL & LBT) seemed to demonstrate more initiative in learning on their own. These groups had a significantly larger number of valid concepts and valid links than the ITS group (see Table 7.6). Also, if one looks at the ratio of expert to valid concepts and the ratios of expert to valid links (Figures 7.9 and 7.10) the ITS group percentages were much higher (>90% for expert concepts and >70% for expert links) than the LBT and SRL groups, and the results are statistically significant (see Table 7.7). When one looks only at the expert concepts and links, the LBT group had more expert concepts than the ITS groups, but none of the other measures applied to the students' final concept maps of the main study were statistically significant (Table 7.6). However, there was no notable difference between the LBT and SRL groups. The MANOVA across the five sessions of the main study confirmed the same findings. Our conclusion from the data is that the LBT and SRL groups created richer concept maps than the ITS group which demonstrate that they probably had a better understanding of the domain, and the learning by teaching environments also prepared them to be better learners. The ITS groups seemed to focus primarily on the quiz questions and mentor feedback, and, therefore, created maps that mainly included expert concepts and links. It is not clear whether they developed a good understanding of the river ecosystem domain. On the other hand, the learning by teaching groups seemed to have spent a lot of effort in trying to develop a better understanding of the domain so they could teach Betty better. They looked beyond the quiz questions and mentor feedback to learn about relevant concepts and relations in river ecosystems, and then teach these to Betty.

Hypothesis 1.3: Memory Test

This section discusses the results of the memory test. Students were asked to recreate their concepts maps after a six-week delay using the same interface for map creation as in the original Betty's Brain system. None of the other features of the system, namely the query and quiz mechanisms, the resources, and the Mentor Agent were available to the students for this part of the study. The maps created were evaluated along three dimensions: recollection, accuracy, and maturation. To calculate data for these measures, the raw data were into the form shown in Tables 7.8 and 7.9. These tables have the same format as the ones presented previously in Chapter 6 in the Retention of Knowledge sub-section.

Table 7.9 Concept-Comparison Results

(a) Results for the ITS Group

Concept Map	*Classification*	*Grade*	
		Valid	*Invalid*
Main Study		11.4	1.5
Memory Test	Recalled	6.1	0.1

	Added	2.1	0.8

(b) Results for the LBT Group

Concept Map	*Classification*	*Grade*	
		Valid	*Invalid*
Main Study		18.1	2.7
Memory Test	Recalled	8.3	0.4
	Added	1.9	0.8

(c) Results for the SRL Group

Concept Map	*Classification*	*Grade*	
		Valid	*Invalid*
Main Study		19.0	2.5
Memory Test	Recalled	7.9	0.2
	Added	2.7	1.4

Table 7.10 Link-Comparison Results

(a) Results for the ITS Group

Concept Map	*Classification*	*Grade*			
		Valid		*Partially Valid (Causal only)*	*Invalid*
		Causal	*Others*		
Main Study		13.5	1.7	0.4	5.3
Memory Test	Recalled	1.7	0.5	0.1	0.7
	Added		3.9		3.3

(b) Results for the LBT Group

Concept Map	*Classification*	*Grade*			
		Valid		*Partially Valid (Causal only)*	*Invalid*
		Causal	*Others*		
Main Study		15.1	8.4	0.9	9.8
Memory Test	Recalled	1.7	1.1	1.4	1.6
	Added	5.9			4.2

(c) Results for the SRL Group

Concept Map	*Classification*	*Grade*			
		Valid		*Partially Valid (Causal only)*	*Invalid*
		Causal	*Others*		
Main Study		14.2	7.5	0.8	13.5
Memory Test	Recalled	2.6	0.5	0.5	1.9
	Added	4.8			4.5

Recollection

The recollection measure counts the number of concepts and links from the students' concept maps at the end of the main study that appear in maps that they created for the memory test. Figure 7.20 shows the average numbers of recalled concepts and links by group, and Figure 7.21 displayed the average ratios of recalled concepts and links to the total number of concepts and links, respectively, in the corresponding main-study concept maps. The average numbers of concepts and links recalled by the LBT and SRL groups were higher than those recalled by the ITS group. As shown in Table 7.11, these results were statistically significant. However, when one compared the ratio of the recalled concepts (GLM ANOVA, $F_{(2, 41)} = 0.99$, $p = .38$) and the ratio of recalled links (GLM ANOVA, $F_{(2, 41)} = 0.51$, $p = .60$), the results were not statistically significant. The general conclusion, therefore, is that students did not show any significant differences in their ability to recall the concept maps they had created in the main study. An interesting observation is that they were much better at remembering concepts than they were at remembering links.

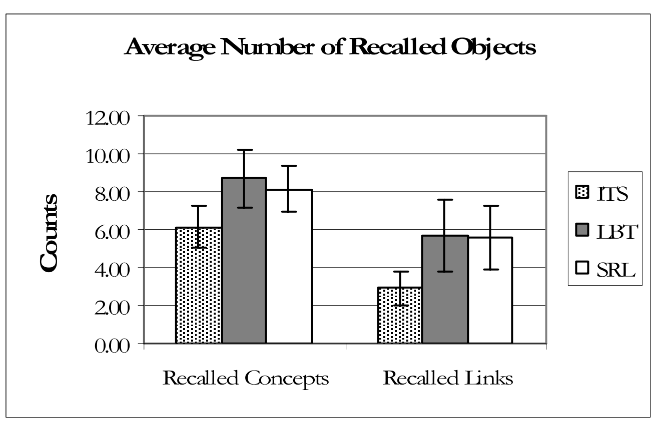


Figure 7.20 Average Number of Concepts and Links recalled in the Memory-Test Maps (Error bars represent the 95% confidence intervals of the differences between means.)

Table 7.11 Significance Levels for recalling the Main-Study Concept Maps

	Recalled Concepts	*Recalled Links*
Test	GLM ANOVA, $F_{(2, 41)} = 4.30$, $p = .02$	Mann-Whitney U
ITS & LBT	Tukey: $p \leq .05$	$U = 48.5$, $p < .05$
ITS & SRL	Tukey: $p > .05$	$U = 62.0$, $p < .05$
SRL & LBT	Tukey: $p > .05$	$U = 101.0$, $p = .91$

The fact that the learning-by-teaching groups could recall concepts and links in the same proportion as the ITS group even though they had more to remember can be construed to be a positive result, and may indirectly be attributed to a better understanding of the domain structure. Also, recollecting the main study results, many of the concept and links created by the LBT and SRL groups were valid and not expert concepts and links. In other words, the LBT and SRL groups had concepts and links in their maps that went beyond the concepts and links the mentor told them about. They had a better understanding of their map and this may have facilitated in their remembering more about the map.

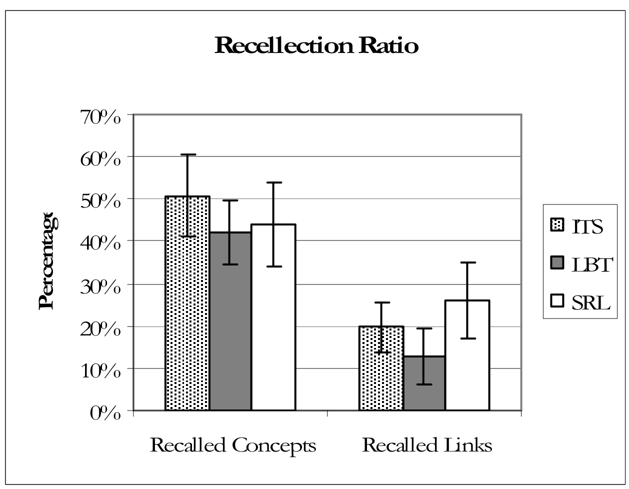


Figure 7.21 Average Ratio of Concepts and Links recalled in the Memory-Test Concept Maps to the Total Number of Concept and Links in the Main-Study Concept Maps, respectively (Error bars represent the 95% confidence intervals of the differences between means.)

Accuracy

This section presents the accuracy of recalling correct concepts and links from the main-study concept maps. The accuracy was calculated by deducting the misconceptions, which were incorrect concepts and links that were recalled, from the total scores. These numbers used the data from Tables 7.8 and 7.9 to compute the accuracy measures (the procedure is discussed in Chapter 6). Figure 7.22 shows the accuracy of the memory recollections for concepts and links, and Figure 7.23 shows the same results as a ratio of accuracy scores of memory-test concept maps to those of main-study concept maps. There was no significant difference in the accuracy measure as shown in Table 7.12.

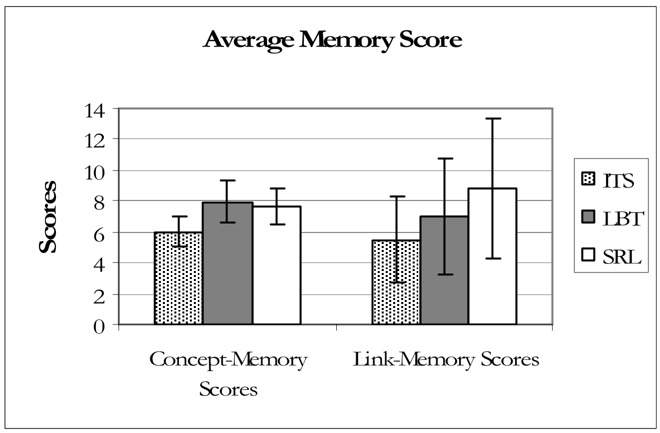


Figure 7.22 Accuracy of Memorization (Error bars represent the 95% confidence intervals of the differences between means.)

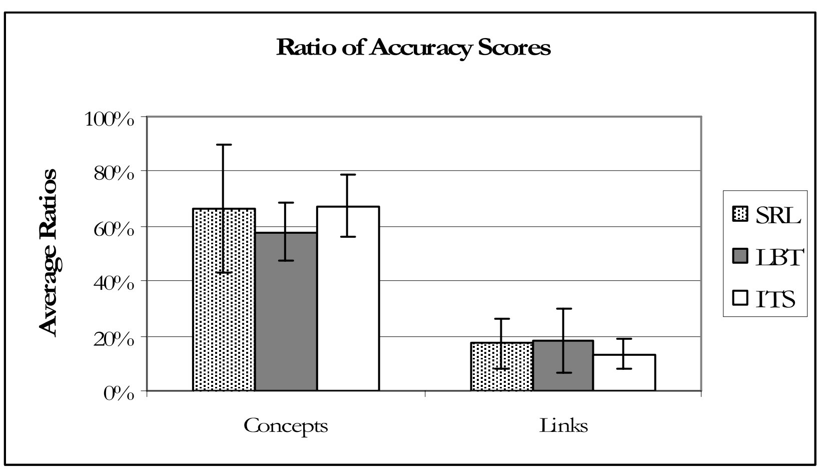


Figure 7.23 Ratio of Accuracy Scores of Memory-Test Concept Maps to those of Main-Study Concept Maps (Error bars represent the 95% confidence intervals of the differences between means)

Table 7.12 Significance Levels for the Accuracy of Memory (Mann-Whitney U Tests)

	Concept Accuracy	*Link Accuracy*	*Concept-Accuracy Ratio*	*Link-Accuracy Ratio*
Test	GLM ANOVA, $F_{(2, 41)} = 3.07$, $p = .06$	Mann-Whitney U		
ITS & LBT		U = 97 $p = .78$	U = 89 $p = .23$	U = 118 $p = .95$
ITS & SRL		U = 84.5 $p = .56$	U = 76.5 $p = .34$	U = 82.5 $p = .50$
SRL & LBT		U = 109.5 $p = .68$	U = 102.5 $p = .95$	U = 93 $p = .65$

<u>Maturation</u>

This section reports the changes in the students' memory-test concept maps from the main-study in terms of omitted and added concepts and links, as shown in Figure 7.24 and Figure 7.25. Because there were six weeks in between the main study and the memory test, the students might have received information in the interim that changed their knowledge. Therefore, we calculated the score of these possible changes by subtracting the number of the negative changes (invalid, added concepts and links and valid, omitted concepts and links) from the number of the positive changes (valid, added concept and links and invalid, omitted concepts and links) in each student's memory-test map. As expected, there was no significant difference in the trend scores (the number of positive objects minus the number of negative objects) between groups (Mann Whitney U Tests; U = 90.0, $p < .25$; U = 77.0, $p < .36$; U = 104.0, $p < 1.00$).

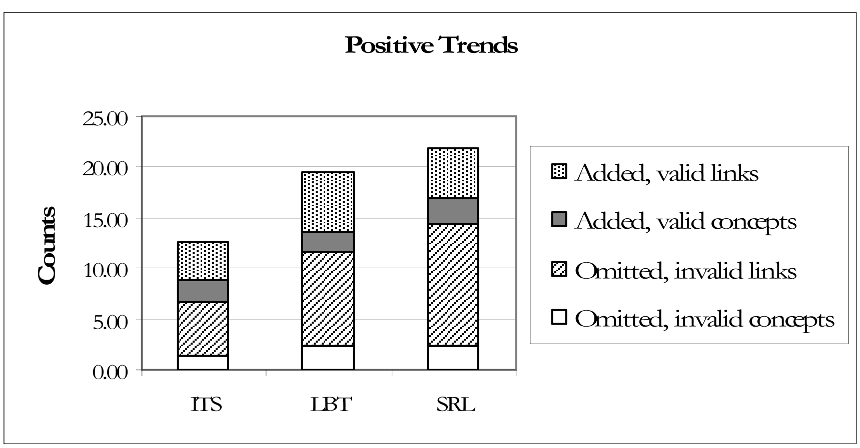


Figure 7.24 Grade Distribution of Positive Trends: Valid Objects that were added to the Memory-Test Maps and Invalid Objects that were omitted from the Main-Study Maps

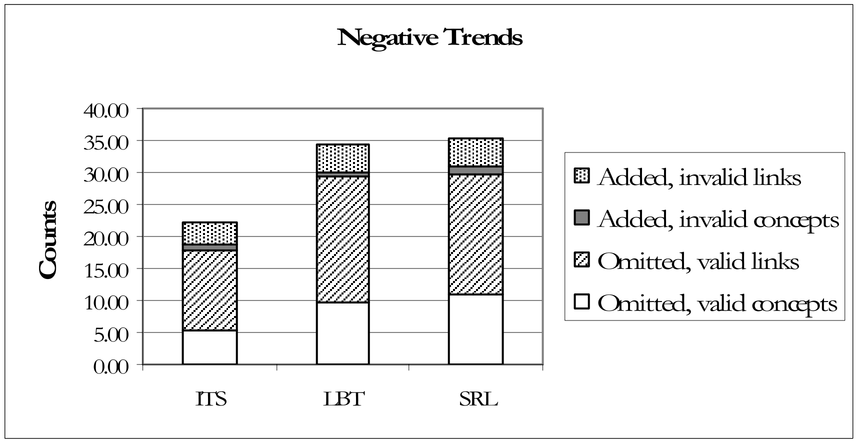


Figure 7.25 Grade Distribution of Negative Trends: Valid Objects that were omitted from the Main-Study Maps and Valid Objects that were added to the Memory-Test Maps

Summary

There was no conclusive difference between the three groups in recalling the concepts and the links that each student created in the main study. Although the learning-by-teaching groups (SRL and LBT) recalled more concepts and links, the recalling ratio and the accuracy were not statistically different between the three groups. According to the maturation measure, these results were not affected by the maturation of the students during the pause of six weeks between the main study and the memory test.

Hypothesis 2: Learning Strategies

This section evaluates how students in the three groups utilized the features in the Betty's Brain environment when constructing their concept maps of the river ecosystem. Among other things, one of our goals is to prepare students for future learning. The SRL version of Betty's Brain, tries to make students aware of self recognition and metacognitive strategies through interactions with Betty and the mentor. We believe that learning and adopting the self regulated strategies will make students better learners, and, therefore, better prepared for future learning in new domains. The rest of this section looks for behaviors that will help establish the students' use of good learning strategies.

The number of times students accessed the online resources, including the domain resources, the concept-map and the reasoning-process tutorials, and the mentor's on-demand help (the SRL group only), illustrated in Figure 7.26, is used as an indication of their effort to seek help when needed, and to gain knowledge about the river ecosystems on their own (as opposed to being told). There were no significant effect of time of measurement and no significant interaction effect of time of measurement and group (see Table 7.12).

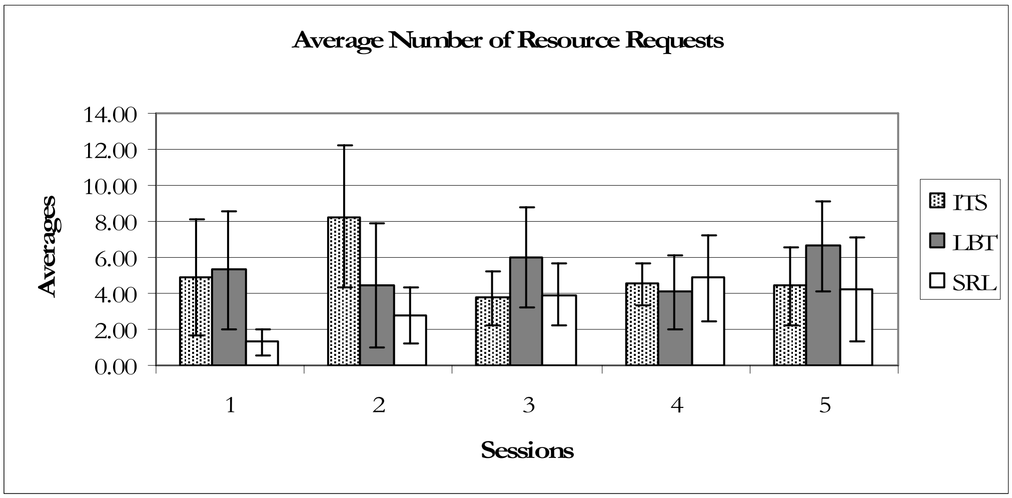


Figure 7.26 Average Number of Resource Accesses in the Main Study by group and by session (Error bars represent the 95% confidence intervals of the differences between means.)

Table 7.13 GLM MANOVA on the Average Number of Resource Requests and Time spent reading Resources in each session of the Main Study

	Numbers of Resource Requests	*Time spent reading Resources*
Time	$F_{(4, 38)} = 0.6$, $p = .70$	$F_{(4, 38)} = 2.3$, $p = .07$
*Time * Group*	$F_{(8, 76)} = 2.1$, $p < .05$	$F_{(8, 76)} = 1.9$, $p = 0.07$
ITS & SRL	Tukey: $p > .05$	Tukey: $p \leq .05$
SRL & LBT	Tukey: $p > .05$	Tukey: $p \leq .05$

However, the ITS and LBT groups spent more time reading the resources (statistically significant, p<0.005) than the SRL group. This would seem like a contradiction, but if one looks at other activities, such as asking causal queries and studying Betty's explanations, overall student behavior becomes clearer.

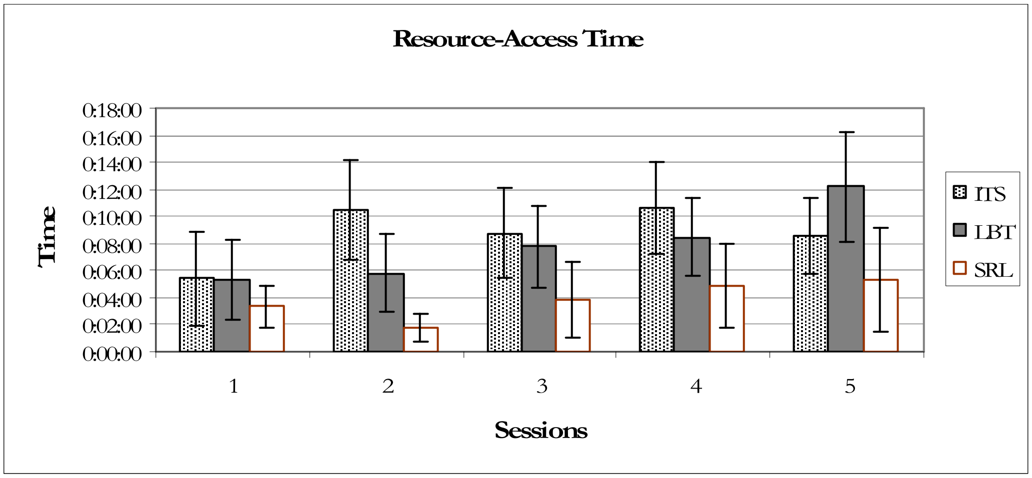


Figure 7.27 Average Amount of Time spent reading Resources (Error bars represent the 95% confidence intervals of the differences between means.)

The numbers of causal queries that students asked Betty, an indication of the student's effort to monitor and debug the knowledge they taught Betty, is illustrated in Figure 7.28, session by session for the three groups. The data was analyzed by a repeated measures analysis of variance with between-subjects factors, and the statistical results are shown in Table 7.14. The effect of time of measurement and the interaction effect of time and group of measurement were both significant. The post hoc results showed that the SRL group asked more casual questions than the other two groups (Tukey HSD, $p \leq 0.05$).

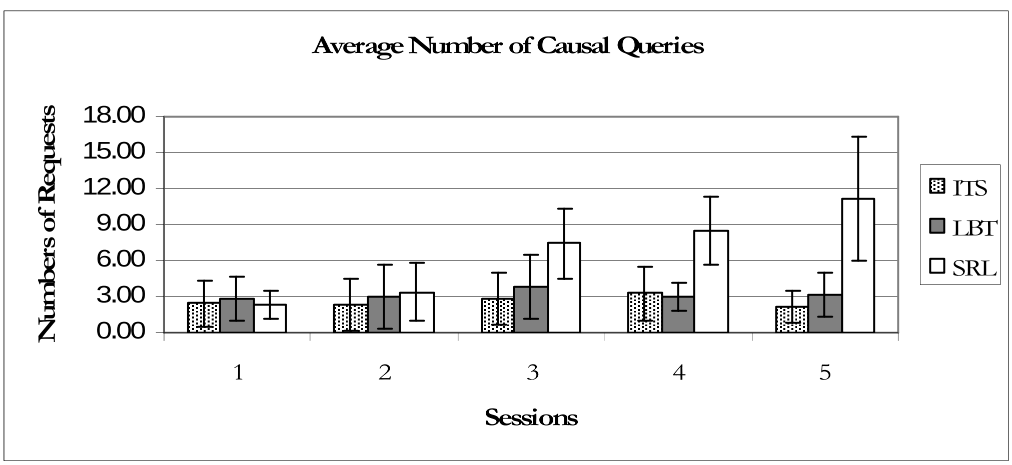


Figure 7.28 Average Number of Causal Queries asked in the Main Study by group and by session (Error bars represent the 95% confidence intervals of the differences between means.)

Table 7.14 Repeated-Measures Analyses of Variance on Averages of Causal Queries in the Main Study

Time	$F_{(4, 38)} = 5.4, p < .001$
*Time * Group*	$F_{(8, 76)} = 3.6, p < .001$
ITS & LBT	Tukey: $p > .05$
ITS & SRL	Tukey: $p \leq .05$
SRL & LBT	Tukey: $p \leq .05$

The numbers of explanation requests, illustrated in Figure 7.29, as discussed earlier, indicates the students' efforts to debug their concept maps, and understand the reasoning mechanism. Explanations helped students understand the reasoning mechanisms Betty employed to answer causal queries. We consider students' analysis and reflection of explanations to be the key in their understanding of how to organize domain knowledge, understand chain of events, and get a global understanding of interdependence among entities in a river ecosystem.

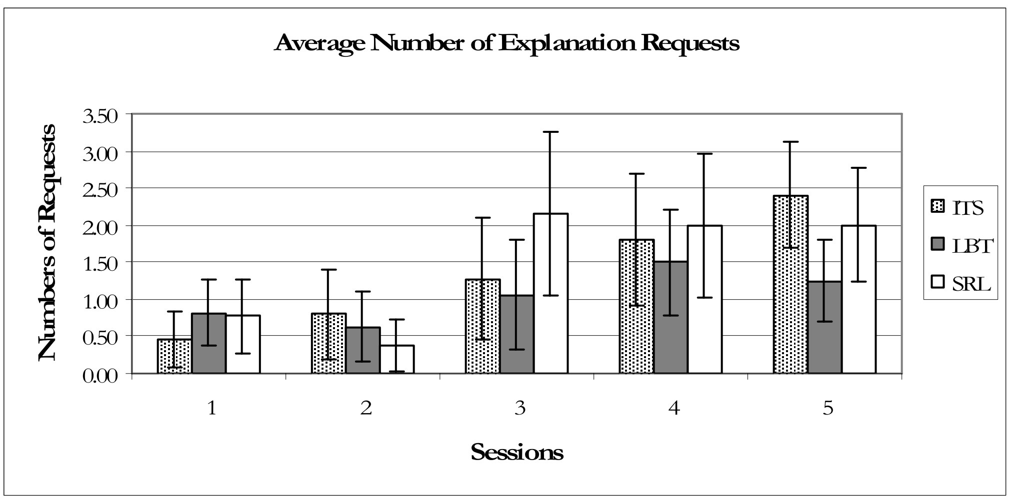


Figure 7.29 Average Number of Explanation Requests in the Main Study by group and by session (Error bars represent the 95% confidence intervals of the differences between means.)

The data were analyzed by GLM MANOVA. The effect of time of measurement was significant ($F_{(4, 38)} = 9.3, p < .0005$) and the interaction effect of time and group of measurement was not significant ($F_{(8, 76)} = 1.6, p = .13$). However, there was no significant result in the post hoc test (Tukey HSD). Therefore, all three groups illustrated an increase in using this feature over time.

The numbers of quiz requests, illustrated in Figure 7.30, indicates the students' efforts to assess their knowledge (formative assessment). The effect of time of measurement was significant (GLM MANOVA, $F_{(4, 38)} = 21.6$, $p < .0005$) and the interaction effect of time and group of measurement was not significant (GLM MANOVA, $F_{(8, 76)} = 2.6$, $p < .05$). However, there was no significant result in the post hoc test (Tukey HSD). Therefore, all three groups illustrated an increase in using this feature over time.

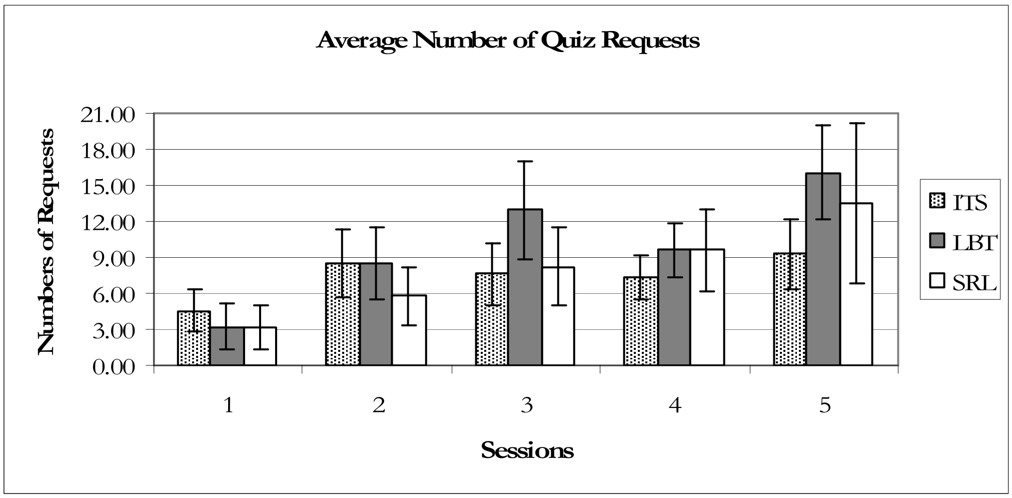


Figure 7.30 Average Number of Quiz Requests in the Main Study by group and by session (Error bars represent the 95% confidence intervals of the differences between means.)

In summary, all students had learned to utilize features provided in the Betty's Brain environment to improve the quality of their concept maps and were successful at that. The SRL group had asked more causal queries but this behavior did not help this group outperform the other two groups in the main study. However, the results of the transfer test in the next section provide greater insights into students' preparation for future learning.

Hypothesis 3: Ability to Transfer

The domain of study was changed from river ecosystems to the land-based nitrogen cycle in this test to measure how well students learned a new domain in an environment that demanded them to learn independently. The measures of quality of concept maps and learning behaviors for the main study were also applied to the transfer test.

Hypothesis 3.1: Quality of Concept Maps

Figure 7.31 displays the numbers of expert concepts and links in the students' concept maps for the land-based nitrogen cycle, and Figure 7.32 displayed the numbers of valid concepts and links. There was no significant difference between groups in the number of expert concepts in the maps they created for the land-based Nitrogen cycle as shown in Table 7.15.

Table 7.15 also shows that the LBT group had more valid concepts in their maps than the ITS group (Mann Whitney U tests, $p < 0.05$), but there was no significant differences between the LBT and SRL groups and the SRL and ITS groups. We could not perform data analysis on the numbers of expert and valid links because the data were skewed (see Appendix K). Overall, it seems that two sessions were not adequate for students to understand the new domain (for the transfer study). Therefore, applying statistical tests may not be meaningful.

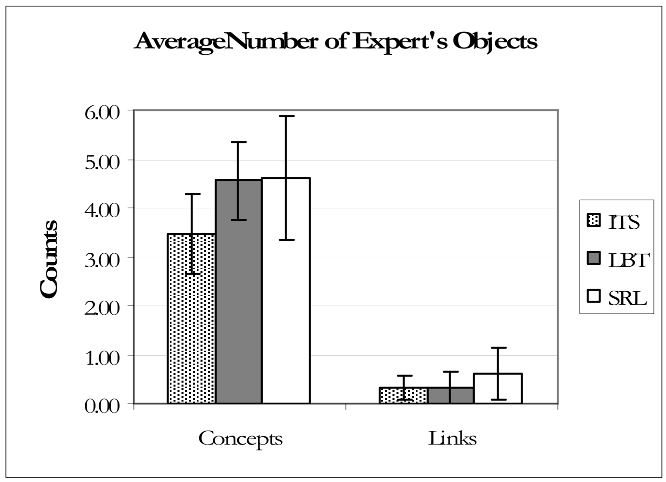


Figure 7.31 Average Number of Expert Concepts and Links in the Students' final Concept Maps for the Nitrogen Cycle (Error bars represent the 95% confidence intervals of the differences between means.)

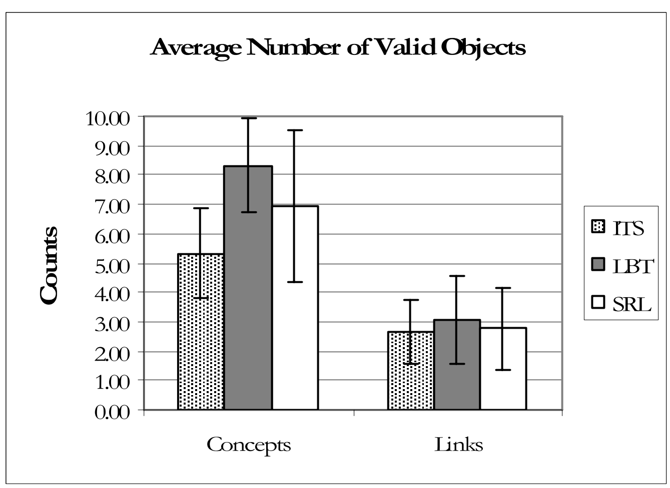


Figure 7.32 Average Number of Valid Concepts and Links in the Students' final Concept Maps for the Nitrogen Cycle (Error bars represent the 95% confidence intervals of the differences between means.)

Table 7.15 Mann-Whitney U Tests on the Correctness Grades of the Students' Concept Maps at the end of the Transfer Test

Variable	*ITS & LBT*	*ITS & SRL*	*SRL & LBT*
Number of expert Concepts	U = 74.5, p = .20	U = 76.5, p = .09	U = 75, p = .32
Number of expert Links	N/A		
Number of Valid Concepts	U = 61.5, p < .05	U = 91, p = .79	U = 74.5, p = .20
Number of Valid Links	N/A		

Hypothesis 3.2: Learning Behaviors

Even though the differences between groups could not clearly established through concept map quantity, the activities conducted by the group showed significant differences (see Figure 7.33 for data on session 2 of the transfer test). Unfortunately, a number of log files of the first session were lost, and therefore, student activities in the first session were not counted.

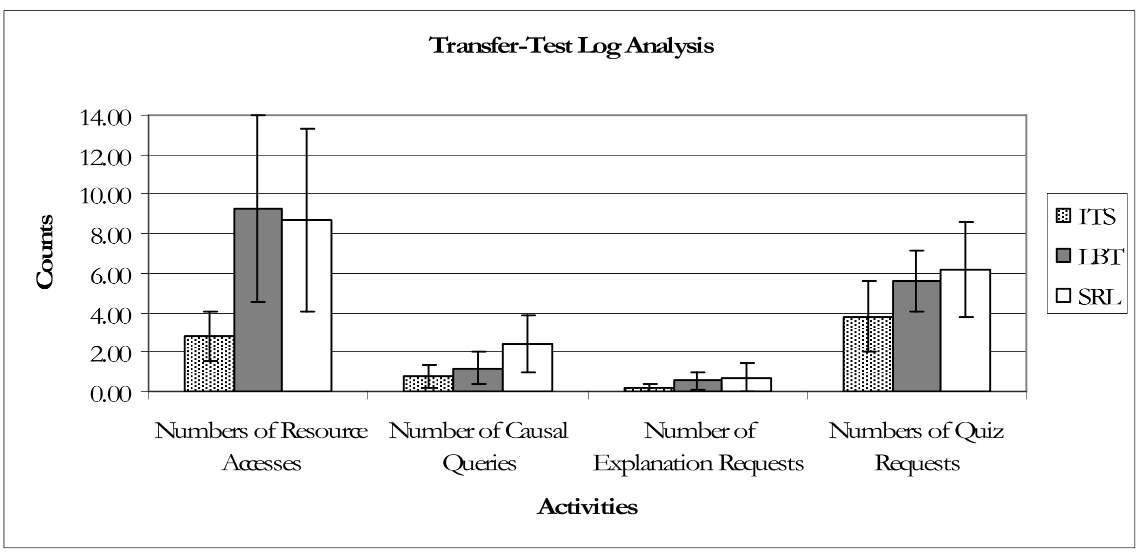


Figure 7.33 Average Number of Resource, Query, and Explanation Request and Quiz Activities during the second session of the Transfer Test (Error bars represent the 95% confidence intervals of the differences between means.)

The results of Mann-Whitney U tests, shown in Table 7.16, indicated that the ITS group accessed the resources significantly less frequently than the LBT and the SRL groups (Mann-Whitney U test, level of significance, $p < 0.01$). Even though the LBT and SRL groups did not differ in the numbers of resource assesses, the SRL group tended to spent more time reading resources ($p = 0.052$), as shown in Figure 7.34. The SRL group spent more time reading the resources than the ITS group (Mann Whitney U Test, level of significance, $p < 0.05$). There were no significant differences in the number of quiz requests for all three groups.

Table 7.16 Mann-Whitney U Tests on the Frequencies of Activities during the second session of the Transfer Test

	ITS & LBT	*ITS & SRL*	*SRL & LBT*
Numbers of Resource Accesses	U = 49, $p < .005$	U = 41.5, $p < .01$	U = 104, $p = 1.00$
Time spent reading Resources	U = 110.5, $p = .71$	U = 55, $p = .052$	U = 59, $p < .05$
Numbers of Quiz Requests	U = 71, p = .054	U = 70.5, p = .22	U = 99.5, p = .85

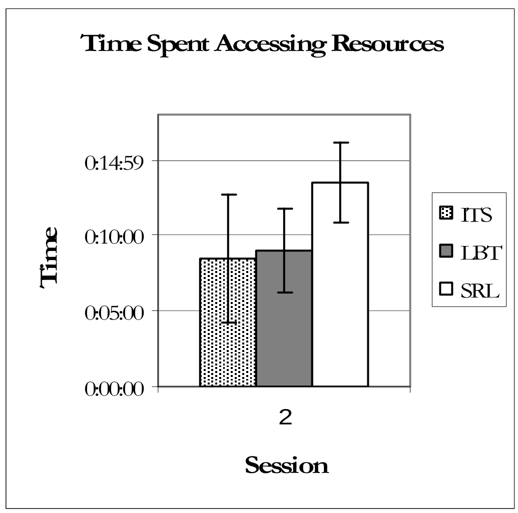


Figure 7.34 Average Amount of Time spent reading Resources in the Transfer Test by group (Error bars represent the 95% confidence intervals of the differences between means.)

We could not perform statistical tests on the number of causal queries and the numbers of explanation requests (see Appendix K). This data was skewed by a large number of zeros because there were a number of students who asked no queries and did not ask for explanations. As an ad hoc analysis, we have counted the number of non-zero value in each category. These counts are displayed in Table 7.17.

Table 7.17 Number and Percentage of Students who asked Casual Queries and made Explanation Requests

Variable	*ITS*		*LBT*		*SRL*	
	Count	*Ratio*	*Count*	*Ratio*	*Count*	*Ratio*
Number of Causal Queries	6	40%	9	56%	9	69%
Number of Explanation Requests	3	20%	6	38%	4	31%

Even though we could not perform statistical tests on this data, the results were interesting. The SRL group had the largest ratio of non-zero data points to the total number of its data points for the number of causal queries (69%). The LBT group had the second largest ratio (56%). For the number of explanation requests, the LBT and SRL groups had about the same ratio of non-zero data points (38% and 31%, respectively). The ITS group had the lowest ratios for both variables, 40% for the number of causal queries and 20% for the number of explanation requests. This implies the LBT and SRL groups made more attempts to understand and debug the concept maps while creating concept maps for the nitrogen cycle.

None of the students' concepts maps were advanced to the point where they could answer any of the quiz questions correctly. Recall that there were no intermediate scaffolding questions in the transfer test to help students construct their concept map in parts. All three quiz questions in the nitrogen cycle are non-trivial because each answer involves multiple links that form a long chain of events to get the answers right. Therefore, the lack of correct answers in the quiz is not surprising.

Even though there was no statistically significant finding in the quality of concept maps constructed in the transfer test, the SRL group demonstrated significantly better learning strategies than the LBT and ITS groups. The fact that the quiz feature in the transfer test provided little scaffolding as compared to the main study did not make the SRL students give up. They adapted, and made a lot of effort to learn the new and difficult domain by spending proportionally more time reading resources than they had done during the main study. This is supported by the significance levels of the number of resource accesses and the total amount of time the SRL group used the resource feature.

In addition, both learning-by-teaching groups seemed to be more motivated while working on the transfer test. More of them asked Betty queries and requested her to explain the query answers than those in the learning-by-being-taught group.

Summary

Differences between Directed Learning and Learning-by-Teaching

This research presented a more comprehensive method for comparing learning-by-teaching and learning-by-being-taught environments. We have examined students' performance in learning and understanding domain material in the main study and the transfer test, and also looked at learning behaviors in the main and transfer studies. The findings indicated that students in the learning-by-teaching groups made an effort in gaining a better understanding of the domain, and this was above and beyond what was required to get the quiz questions right. In other words, students who taught Betty made greater efforts to learn more and teach Betty what they had learnt. This was demonstrated by the greater number of valid concepts and links in the students concept maps in comparison to the maps of the ITS group.

Even though the SRL group (or even the SRL and LBT together) did not demonstrate superior performance in the transfer test, the students, especially the ones in the SRL group, seemed to demonstrate the proper learning behavior of one being faced with a complex domain that he or she knows little about. Compared to the other two groups, the SRL students spend much more time reading resources, had more expert concepts and links and made more attempts to debug their maps by asking causal queries and asking Betty to explain her answers. Overall, our observations during the study showed that the students working with the teachable agent were more focused on the task than those who were directed by the mentor agent in the main study. In the exit interviews, students who worked with Betty expressed the desire to work more with the agent while the students directed by the mentor did not.

Effects of Self-Regulated Learning on Learning by Teaching

This thesis is the first longitudinal study examining the effect of self-regulated learning toward the ability to transfer in terms of both performance and learning behaviors. The results indicated that receiving general, self-regulated feedback was a burden to the learners initially as we could see from the slow start in performance during the first two sessions of the main study. (If one looks at Figures 7.17 to 7.21, the SRL group was universally behind the other two groups for all performance and behavior measures in the first two sessions.) However, by the end of the main study, these students produced concept maps that were comparable to the ones created by the LBT group, and better than the one created by the ITS group.

Even in a short time to learn and understand a new, complex domain, the SRL group demonstrated better performance in their transfer task compared to the other two groups. The learning strategies that the students in the SRL group employed were clearly the ones that lead to better learning and understanding. As discussed, this was demonstrated by the time they spend in reading resources and their attempt to debug their maps by asking the teachable agent casual queries. Overall, this demonstrates that SRL strategies may better prepare students for future learning.

CHAPTER VIII

CONCLUSIONS

This dissertation research has made four primary contributions in the area of intelligent learning environments. First, this dissertation has introduced a new, systematic multi-agent framework for the design of learning-by-teaching systems that is geared toward helping novice learners in complex science domains. Second, this design brings together the cognitive and social aspects of the learning process in modular form into the multi-agent framework. The Betty's Brain environment developed for this dissertation incorporates self-regulated learning strategies into the teachable-agent framework to aid students in developing metacognitive abilities that can be transferred to learning in other domains. Third, the extensive studies conducted in a fifth-grade science classroom demonstrate that learning-by-teaching environments produce a greater degree of effective learning than environments where students are taught by computer-based pedagogical agents. Last, this research demonstrates the implications of feedback based on self-regulated learning strategies in promoting learning that goes beyond the computer environment that students have used in the main study.

A Learning-by-Teaching Design with Self-Regulated Learning Strategies

This dissertation develops a learning-by-teaching environment that differs from the previous systems that incorporate learning by teaching in five aspects.

1. The agent only knows what it has been explicitly taught by the student. To keep the overhead low, the student teaches Betty using a visual representation structure. Unlike some other learning-by-teaching systems, the teachable agent does not apply inductive learning techniques to generalize what it has learnt from the student.
2. The knowledge model that students build is visible to them and is in an intuitive and well-structured format that is easy to understand and manipulate. Moreover, the student and the teachable agent share the same visual knowledge representation. The representation is active in that the agent can reason with this knowledge and demonstrate its reasoning process using speech, animation, and text.
3. Students receive both cognitive and outcome feedback. The cognitive feedback focuses on self-regulated learning strategies. The outcome feedback informs students how well the teachable agent is performing on the formative assessment provided by the mentor agent. This helps students reflect on their own knowledge and understanding of domain materials.
4. The environment employs a multi-agent architecture that will allow for greater scalability and applicability to other domains of study.

5. This environment strongly supports the framework of effective learning environments as defined by the HPL framework. This is apparent because the Betty's Brain system:
   - provides both domain knowledge and learning strategies (is knowledge-centered),
   - lets students have the control of the pace and style of learning, the activities that that they wish to engage in, and the materials students want to learn and teach the teachable agent (is learner-centered),
   - includes mechanisms for formative assessment, both internal and external, and external feedback (is assessment-centered), and
   - provides an environment for social interactions along with learning (is community-centered); students take on the responsibility for teaching Betty to help her succeed in her goal of joining the high school science club. Students also have opportunities to learn by seeking help from the mentor agent.

Agent Architecture

To customize the learning strategies and guidance feedback, the design of the Betty's Brain system follows a modular and reusable approach that allows flexibility in designing multiple agents that can be attributed with different learning and meta–cognitive strategies and behaviors[1]. The computational architecture also facilitates the design and implementation of multiple pedagogical agents and the reusability of components of agents in different learning situations. The agent architecture incorporates three primary components: (i) the agent persona and appearance, (ii) the knowledge representation and reasoning mechanisms that the agent uses for answering questions and for problem solving, and (iii) multi-modal interfaces (including text, speech, and animation) that it uses for communicating with other agents and the student. Also, the overall design is modular in that one can change domains by changing the resources, the expert concept map, and the representation and the reasoning mechanisms associated with the representation along with some of the graphical displays related to the components.

[1] We have implemented Teachable Agents with different behavior characteristics, but studying the effects of different behaviors was not a topic of this dissertation.

There are two types of agents in the Betty's brain environment: pedagogical agents and service agents. The pedagogical agents, the teachable agent (Betty) and the Mentor agent (Mr. Davis), are visible to students and form the core of the students' learning process in the Betty's Brain system. The teachable agent can reason about the visual, shared knowledge structure created by the students to answer queries from students and take quizzes provided by the mentor agent. In addition, the teachable agent employs self-regulated learning strategies to interact with the students and demonstrate good learning strategies by example. The mentor agent provides the formative assessment scheme in the form of quiz questions and on-demand feedback on learning, teaching, and the domain of study. Agents in the second category, the service agents, act as computational and communication aids in the environment and are invisible to the students. The two agents in this category are the Interfacer agent and the Pattern Tracker agent. The interfacer agent provides a common communication link between all agents and the other modules so that they are least independent to each other. The pattern tracker looks for learning patterns, such as trial-and-error patterns, in the students' activities so that the pedagogical agents have information about the students' behaviors as an aggregate of a set of actions.

The knowledge representation in the Betty's Brain system is based on the concept map. This module includes the internal representation of the concept map and the reasoning mechanisms that the pedagogical agents can use. Grouping the reasoning mechanisms with the representation increases the flexibility of the agent architecture when agents work with multiple representations and the reusability of the representation with other agents in the future.

The resources are in a separate module and easy to customize because each of the topics is a separate hypertext document. When the teachable agent environment is adapted to another domain or representation with new resources, we only need to replace the hypertext links from the resource index panel, and possibly the text associated with the links.

The last module of the Betty's Brain environment is the graphical user interfaces (GUIs) that visually display all components of the system to the users of the Betty's Brain system. All of the actions that students perform using these components are aggregated into a package and send to the corresponding agents via the interfacer agent. When the agents need to communicate to the users, they also send packages to the GUI units to display their messages visually and vocally. This mechanism allow multiple teachable and mentor agents to work on the same set of GUIs.

Comparative Studies

The extensive studies have been conducted to compare three learning methods:

1. Directed learning (the ITS group)
2. Learning by teaching with directed feedback (the LBT group)
3. Learning-by-teaching with guided feedback based on self-regulated-learning strategies (the SRL group)

All students improved their knowledge on river ecosystems using the Betty's Brain environment in the main study on river ecosystems. All of the groups used the visual editor successfully to create concept maps on river ecosystems. As far as expert concepts and links, all three groups had maps of about equal quality, but the SRL and LBT groups had many more valid concepts and links (the differences were statistically significant). In terms of overall qui performance, there were no significant differences between the groups. Their learning behaviors did not show significant differences when looked at how they accessed resources, asked Betty to explain her reasoning mechanisms, and the number of times they asked Betty to take quizzes. However, the SRL group asked Betty more causal questions than the other two groups even though it spent less time reading resources by average (these differences were statistically significant). Our conjecture was that the students in the SRL group spent more time thinking about questions to ask Betty so that they could debug her concept map. In summary, at the end of the main study, both learning-by-teaching groups produced richer concept maps than the directed-learning group. The directed learning group produced concept maps that contained mostly expert concepts and links, and they seemed to have relied heavily on the qui questions and the directed feedback provided by the mentor agent. The larger number of valid concepts and links in the LBT and SRL group maps indicates that these students picked up domain knowledge on their own, and made efforts to teach Betty this knowledge by including it in their concept maps. In other words, it was not clear that the directed-learning group developed strategies on how to learn on their own but the learning by teaching groups did. This was further illustrated in the transfer test.

Both learning-by-teaching groups recalled more concepts and links from the main study after six weeks had passed. However, there was no obvious different in the recalling rates and quality. Therefore, it was conclusive which group could longer retain the knowledge structure.

In the transfer test, the SRL group showed the most independent learning behaviors even though they did not result in significant results in the quality of the students' concept maps. This SRL group on the average spent more time than the other two groups in reading resources. In addition, the SRL group produced the highest average number of expert links in the nitrogen cycle, average number of causal queries, and average number of modifications per quiz even though these numbers were not significant.

Interestingly, the SRL students' behaviors in the transfer test differed from those in the main study. Asking queries is the activity that made the SRL group differed from the other groups in the main study. However, in the transfer test, it was the time spent reading resources that differentiated the group. They monitored their learning and adjusted their learning strategies so as to be most effective to the situation and their knowledge state.

This highly adaptive strategy showed that the SRL group was more adaptive to their learning context and environment than the other two groups. In addition, The LBT group also continued the trend of producing rich concept maps by having significantly more valid concepts and number of resource accesses than the ITS group.

In summary, we believe that the advantages of learning by teaching, directed learning, and guided learning can be combined to enable students to become more independent learners and to be better prepared for future learning tasks. The directed feedback in the ITS and LBT groups enhanced their performance in the early sessions of the main study because the students were given hints that allowed them to construct initial concept maps. In the memory test, the learning-by-teaching groups remembered more concepts and links. In the transfer test, the benefits of the guided learning group (SRL) were visible but, unfortunately, the study did not give them sufficient time to effectively demonstrate their learning strategies.

These results indicate that different learning methods can positively affect different phases of the learning process. Directed feedback can be used at the beginning of the learning process to motivate students and increase their initial performance. This gives students an entry into a new domain and also increases their levels of comfort in working in the learning environment. Later on when students start to adjust to the new learning environment, the learning process can be made more exploratory, which, in our case, would correspond to being moved from a directed-learning situation to a learning-by-teaching situation. To refine the students' learning process, self-regulated strategies can become the major framework to guide and help the students in their learning, teaching, and reflection tasks. In the meanwhile, the scaffolding framework by directed learning can be gradually removed.

Overall, the studies demonstrated a number of interesting trends that support learning by teaching and mentoring using self-regulated learning principles. However, in a number of situations, statistically significant results could not be obtained for a number of reasons.

1. The period of the main study was not long enough.

2. The time given for the transfer test, two forty-five-minute sessions, was too short.
3. For pragmatic reasons (such as the lack of resources), the study was only run in two sections of the fifth-grade classroom. Therefore, the three groups had between thirteen to sixteen students each. As a result of smaller size per group, in many cases, the data did not meet the strong requirement of ANOVA, and the Mann-Whitney test was used instead.

Implications of Self-Regulated-Learning Feedback in a Learning-by-Teaching Environment

The participants that received self-regulated-learning feedback showed a greater tendency to apply these principles in their own learning tasks when they faced a new challenge. However, this behavior did not result in statistically significant results in the concept map quality in the transfer test. This, very likely, can be attributed to the fact that students did not have enough time to apply their learned skills in the new and complex domain. The reasons for this were explained in the previous section.

Nonetheless, there are a number of lessons we have learned from these studies. First, instruction styles and types of feedback can affect students' behaviors even when they are introduced by software agents who take on the role of pseudo humans. Introducing characteristics that directly imply learning and understanding is important. Having the same tools, students who were told to help the teachable agent were more willing to explore an unfamiliar domain and tools to gain knowledge to teach, but students who were told to learn for themselves focused more on getting good scores. The observation during the transfer test suggested that the group that learned by being taught tended to give up on the transfer test, which was difficult, where they received no direct hints.

On the other hand, providing directed feedback early in the learning process as a scaffold to help students get started may be a more effective way to aid students in their learning and teaching tasks. From the concept-map analysis, students with self-regulated feedback spent three first sessions trying to understand the learning strategies introduced by Betty and Mr. Davis. Once they had figured this out, we started to see the improvement in their performance.

Finally, the development of self-regulating strategies in students is hard to measure because the same pattern can mean different things depending on students' goals and the current situation the students are in. For example, the composite pattern of behavior that includes taking a quiz, editing a link in the concept map, and then taking the same quiz could mean two things. This may reflect a trial-and-error behavior that the student is exhibiting if it is observed when the student is still learning the materials of the domain. On the other hand, late in the process, the student may be in a situation where he is making an attempt to correct one small error in his concept map.

Future Research

There are a number of lessons that we have learned from this dissertation that translate into a number of design revisions we would like to incorporate in our future learning-by-teaching environments. Some of these issues are being examined by other members of the Teachable Agents Group at Vanderbilt University.

Design of Learning-by-Teaching Systems

There are two main areas in which the design could be improved. The first issue deals with a major challenge, how to improve the self-regulated learning features. The teachable agents group's current solution is to expand, strengthen, and synchronize the self-regulated-learning behaviors in the teachable agent and the self-regulated-learning feedback from the mentor agent. An example of this expansion is goal setting. The story told to the participants at the beginning of the study should set the goal for them to help Betty or to learn for themselves. However, the system provided no direct support for breaking down that big goal into sub-goals to set students off in the right direction and help them track their achievements. Therefore, the next version of the Betty's Brain environment should expand the role of the mentor agent so that he can help students establish their learning sub-goals.

Strengthening the existing features can be done by combining directed and guided feedback as discussed previously. Help types and levels can be varied as students progress, and as a function of the situation that students and Betty are in. Nonetheless, detecting when students start to develop their own learning strategies and identify which strategies they are developing are non-trivial tasks. To detect patterns more accurately, the pattern trackers need to look at longer sequences of activities over a continuous period of time and knows the user's learning goal at the moment. These difficulties appear in both online and after-the-fact analyses.

Synchronization of features is also important. It will reduce redundancy yet can remind students what they should be doing. For example, at one point the mentor tells the student that he should ask Betty questions before sending her to take a quiz. Later on, Betty corroborates this point by telling the user that she is not ready to take a quiz until he helps her review her knowledge by asking her some causal questions. Such synchronization and dual reinforcement should clarify to students about what good learning practices are.

Agent Architecture

The agent architecture still needs refinement to streamline its functionality. In a newer version of Betty's Brain, the communication tasks of the Interface agent are taken over by a new agent, the Environment agent. This enables all of the communication issues to be handled by a common framework, as opposed to direct pair-wise communication for every pair of agents in the system. This will address the scalability issue for such systems.

Another change is the Pattern Tracker agent. The idea of having only a single pattern tracker is that all pedagogical agents are interested in similar information. Therefore, the Pattern Tracker can share it the aggregated pattern it derives with them. However, the disadvantage of this design is that the individual agent design is not well encapsulated. Another way to design the pattern tracking unit is letting each agent have its own pattern tracker and share the information it finds with other agents.

Grading Programs

The grading programs reduce the time spent by and errors made by human graders in every phase of the grading process, namely the pre- and posttest grading, the concept map grading in both domains of the river ecosystem and the nitrogen cycle, and the activity-log analysis. Even without the database, the manual grades can be checked with each other and their previous grading to determine whether the grading is consistent. For example, if an early concept record, "elodea", is graded as "relevant" but later on it is graded again as irrelevant, the grading program will report this conflict. Moreover, the grading programs are independent from the concept map structure can be used if the concept-map structure is updated from one study to the next. When one switches to a new domain, the list of synonyms and relevant concepts has to be created, but otherwise, the grading process remains the same.

These grading programs have components that can improve the Betty's Brain environment. For example, the `WordComparator` class checks for misspellings. This has not been done in Betty's Brain; students have to spell names correctly for the quiz grading procedure to recognize their Expert's concepts. Another example is the `ConceptComparator` class, which returns a list of possible concepts in the student's concept map that can match a particular concept in the Expert's concept map. The current concept finding algorithm returns the first-found, exact match or synonym. Having a list of the synonyms return will reduce redundancy in the student's concept map. A real example of such concept maps is a map that contains concepts "plants", "green plants", and "live plants." The "plants" concept seems to be the best match judging from only the labels. However, it depends on which concept is connected to the other related Expert's concepts such as "dissolved oxygen" and "sunlight."

APPENDIX A

QUIZ CONFIGURATION FILE

```
#NUM_QUIZZES=3
#QUIZ=Quiz 1
#QUESTION=waste;2;plants
#QUESTION=waste;2;bacteria
#QUESTION=bacteria;2;nutrients
#QUESTION=nutrients;2;crowded plants
#QUESTION=crowded plants;2;sunlight
#QUESTION=sunlight;-2;plants
#QUIZ=Quiz 2
#QUESTION=dead organisms;2;animals
#QUESTION=dead organisms;2;bacteria
#QUESTION=bacteria;2;dissolved oxygen
#QUESTION=dissolved oxygen;-2;animals
#QUIZ=Quiz 3
#QUESTION=sunlight;-2;animals
#QUESTION=sunlight;-2;plants
#QUESTION=plants;-2;dissolved oxygen
#QUESTION=dissolved oxygen;-2;animals
#QUESTION=animals;-2;carbon dioxide
#END=DONE
```

APPENDIX B

THE EFFECTS OF FEEDBACK IN SUPPORTING LEARNING BY TEACHING IN A TEACHABLE AGENT ENVIRONMENT

Leelawong, K., Davis, J., Vye, N., Biswas, G., Schwartz, D., Belynne, T., Katzlberger, T., & Bransford, J. (2002). The effects of feedback in supporting learning by teaching in a teachable agent environment. In P. Bell, R. Stevens, & T. Satwicz (Eds.), *Keeping Learning Complex: The Proceedings of the Fifth International Conference of the Learning Sciences (ICLS)* (pp. 245-252). Mahwah, NJ: Erlbaum.

Krittaya Leelawong, Joan Davis, Nancy Vye, Gautam Biswas
Box 1679-Station B, Vanderbilt University, Nashville, TN 37235
Tel: 615-322-7763, Fax: 615-343-5459
kiktaya@vuse.vanderbilt.edu, joan.davis@vanderbilt.edu, nancy.vye@vanderbilt.edu,
biswas@vuse.vanderbilt.edu

Daniel Schwartz
Stanford University
daniel.schwartz@stanford.edu

Kadira Belynne, Thomas Katzlberger & John Bransford
Vanderbilt University
kadira.belynne@vanderbilt.edu, katzlbt@vuse.vanderbilt.edu,
john.bransford@vanderbilt.edu

Abstract: The idea that teaching others is a powerful way to learn is both intuitively compelling, and one that has garnered support in the research literature. The present study investigates aspects of the "learning by teaching" process that contribute to enhanced learning outcomes for students. We developed a computer-based teachable "agent" that students explicitly teach using concept maps. Results indicate that providing students with opportunities to quiz their agent decreases the amount of irrelevant information and increases the proportion of causal information in students' maps, whereas having opportunities to query their agent increases the interconnectedness of concepts in students' maps. The results point to the importance of including various forms of feedback in designing teachable agent environments that promote learning.

Introduction

The idea that teaching others is a powerful way to learn is both intuitively compelling, and one that has garnered support in the research literature. For example, Bargh and Schul (1980) found that people who prepared to teach others to take a quiz on a passage learned the passage better that those who prepared to take the quiz themselves. The literature on tutoring suggests a similar conclusion in that tutors have been shown to benefit as much from tutoring as their tutees (Graesser, Person, & Magliano, 1995; Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001). Biswas and colleagues (Biswas, Schwartz, Bransford, & TAG-V, 2001) report that students preparing to teach made statements about how the responsibility to teach forced them to gain deeper understanding of the materials; other students focused on the importance of having a clear conceptual organization of the materials.

Reflection on these studies and others lead us to conjecture that the creation of a computer-based system, where students can assume the role of "teacher," may provide an effective and motivating environment for learning. We designed an environment that lets students explicitly teach a computer agent. Once taught, the agent reasons with its knowledge and answers questions. Students observe the effects of their teaching by analyzing these responses and by getting additional feedback from a teaching expert. Although the research literature (e.g., Bargh & Schul, 1980) suggests that learning benefits accrue from preparing to teach, it is not clear if and how other aspects of the learning by teaching process contribute to enhanced outcomes. In addition to preparatory activities, teachers provide explanations and demonstrations during teaching and receive questions and feedback from students. These activities also seem significant from the standpoint of their cognitive consequences. For example, we might expect that teachers' knowledge structures would become better organized and differentiated through the process of communicating key ideas and relationships to students and reflecting on students' questions and feedback (Chi, et al., 2001).

The purpose of the present study was to examine the learning benefits of different activities associated with "learning by teaching" in our teachable agent environment. This was done by constructing computer-based agents that students could teach domain knowledge. In particular, we created an agent environment called Betty's Brain which can operate in three modes: (i) the TEACH mode, where students impart knowledge to the agent Betty by means of a dynamic concept map interface, and access content materials as needed to learn information for teaching, (ii) the QUERY mode, where students ask Betty questions (using question templates) which she answers by reasoning with information that the student has taught her, and (iii) the QUIZ mode, where students evaluate how well they have taught Betty by observing her performance on a quiz. At times, an expert teacher agent intervenes to make suggestions that may help Betty (and the student) correct her answers.

In the present study we examined the effects of the interactive features of the teachable agent environment that emulate the feedback that instructors receive from students during teaching. All students had the opportunity to TEACH their agent, and we manipulated whether students could QUERY Betty and observe her QUIZ performance following their teaching efforts. Crossing these variables created four versions of the teachable agent environment: 1. TEACH Only version (No QUERY or QUIZ), 2. QUERY version, 3. QUIZ version and 4. FULL version (QUERY & QUIZ).

We hypothesized that having opportunities to query and/or quiz Betty would positively, but differentially, impact students' learning. The query feature helps students debug their own thinking and reasoning in the problem domain. If Betty answers questions in unexpected ways, students know that they need to add to or modify their concept maps. In addition, and perhaps more important, when Betty explains her answers, she makes explicit the process of reasoning across links in a concept map (i.e., infer the effect of one concept on another through a chain of relations). Therefore, we might expect that students who use the QUERY versions of the software would create maps containing more inter-linked concepts. With respect to the quiz condition, we expected that students would become better at identifying important concepts and links to include in their maps because they could map backward from the quiz questions. We also expected that overall they would produce more accurate concept maps because they had access to feedback on Betty's quiz performance.

Methods

Description of Betty's Brain

Figure 1 illustrates the Betty's Brain interface. The system possesses multimedia capabilities. Students use a graphical drag and drop interface to create and modify their concept maps. When queried, Betty can provide explanations for how she derives her answers, and simultaneously depict the derivation process on the concept map by animation. In the sections below, we describe the software's 3 modes: TEACH, QUERY and QUIZ.

TEACH Betty

Students teach Betty by means of a concept map interface. A concept map is a collection of concepts and relations between these concepts (Novak, 1996). A relation is a unidirectional link connecting two entities. Concept maps help to categorize groups of objects and express interactions among them. They also provide a mechanism for representing knowledge hierarchies and cause-effect relations (Stoyanov and Kommers 1999). This makes the concept-mapping technique very amenable to applications in scientific domains, in particular, for modeling dynamic systems. In this study, we asked students to teach Betty about rivers--the living and non-living things in a river and how living things survive. In this way, the focus is on creating a model of a river ecosystem and on ideas of balance and interdependence.

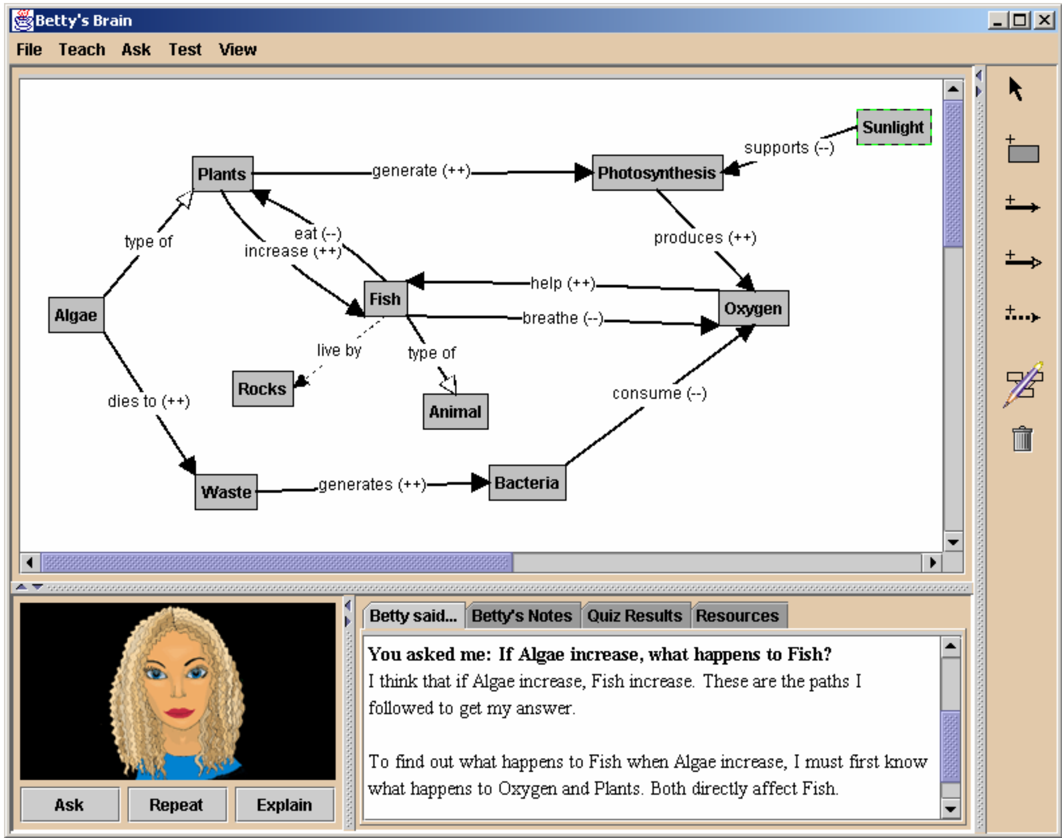


Figure B.35 Betty's Brain Interface

Figure B.35 displays an example of a concept map that student created in Betty's Brain—the map represents what the student has taught Betty. Note that this map is not totally accurate or complete. The labeled boxes correspond to concepts (the labels are concept names), and the labeled links correspond to relations. Students can use three kinds of links, (i) causal, (ii) hierarchical, and (iii) descriptive. Students use descriptive links to embed notes or interesting characteristics of an object in their concept map (e.g., "Fish live by Rocks"). Hierarchical links let students establish class structures to organize domain knowledge (e.g., "Fish is a type of Animal").

A causal link specifies an active relationship on how a change in the originating concept affects the destination concept. Two examples of this type of relation are "Fish eat Plants" and "Photosynthesis produces Oxygen". The causal relations are further qualified by increase ('++') and decrease ('--') labels. For example, "eat" implies a decrease relation, and "produce" an increase. Therefore, an introduction of more fish into the ecosystem causes a decrease in the number of plants, but an increase in the number of plants causes an increase in oxygen.

QUERY Betty

Students are able to query Betty about what they have taught her. The query mode consists of two mechanisms: (i) a reasoning mechanism and (ii) an explanation mechanism. The reasoning mechanism enables Betty to analyze the knowledge that the student has taught her so that she can answer questions. The explanation mechanism enables Betty to produce a detailed explanation of how she generated her answer. Currently, Betty's Brain has templates for four question types:

Type 1: What will happen to Concept A when we increase/decrease Concept B?

Type 2: What will happen when we increase/decrease Concept A?

Type 3: What can cause Concept A to increase/decrease?

Type 4: Tell me what you know about Concept A.

In this paper, we discuss the reasoning and the explanation mechanisms for the first type of question. The mechanisms of the second and the third types are similar (see Leelawong, Wang, Biswas, Vye, Bransford, & Schwartz 2001). The last type of question is descriptive; no reasoning mechanism is involved. The explanation mechanism uses a simple process to convert concept-link-concept pairs to a specific text form.

The reasoning mechanism is based on a simple chaining procedure to deduce the relationship between a set of connected concepts. If the teachable agent is asked to answer the question about the effects of an addition of Algae on Fish, she searches the concept map and deduces that Fish increase when more Algae are added to a river.

As mentioned earlier, Betty employs animation and speech to explain her thinking to the students. A written explanation is also available. The structure of Betty's explanations is closely tied to the reasoning algorithm. To avoid information overload, the explanation is broken down into segments. If users ask for more explanation, Betty works backward, and links the concept back to the closest nodes. When asked "What happens to fish when algae increase?", Betty's initial response is: "I think that when Algae increase, Fish increase." Students can then ask Betty for a more detailed explanation. The following is the sequence of explanation that Betty generates:

1. "You asked me: If Algae increase, what happens to Fish? I think that if Algae increase, Fish increase."
2. "To find out what happens to Fish when Algae increase, I must first know what happens to Oxygen and Plants. Both directly affect Fish."
3. "An increase in Algae causes Plants to increase, which causes Fish to increase."
4. "Now let me explain how Algae affect Fish through Oxygen. An increase in Algae causes Oxygen to remain the same which causes Fish to remain the same."
5. "So, no change in Fish from Oxygen and an increase in Fish from Plants cause an increase in Fish."

After each step in the explanation, students have the option to: (i) ask Betty to repeat the explanation or (ii) look at the explanation in text form.

QUIZ Betty

During the quiz phase, the student observes Betty's responses to a set of pre-scripted questions. The teaching expert informs Betty (and the student) if Betty's answers are right or wrong. The teaching expert also gives hints to help the student debug the concept map.

The teaching expert employs a simple mechanism for generating feedback. The system is provided with an expert concept map (built by a teacher or other expert) in the domain of study. The student's concept map structure is superimposed on the expert's, and the teaching expert searches for a missing concept (first) or relation that is considered essential for the right answer, and uses this to generate a hint for the student. A hint is given, if necessary, for each quiz question. Currently, the system implements three levels of hinting. First, the expert suggests that the student read a section of the resource materials that relates to the concept or link. The second hint for the same question explicitly mentions the name of the missing concept or relation, and asks the student to look for more information on that topic. The third hint tells how to correct the concept or relation in the map.

Procedures

Study participants were 50 high-achieving fifth grade students from a science class in an urban public school located in a southeastern city. Students were randomly assigned to one of four versions of the software: TEACH only, QUERY, QUIZ, or FULL.

The software was used in 3 sessions of one hour each. At the beginning of session 1, students were introduced to features of the software. They were asked to teach Betty about river ecosystems—to model things contained in rivers and how living things in a river meet their survival needs. In between sessions with Betty, students engaged in independent study to prepare themselves to teach Betty. Reference materials were also available for students to access as needed when preparing to teach and when teaching Betty.

Results

Analysis of the scope of students' maps and the types and accuracy of links contained therein suggest several conclusions. Figure B.36 shows the mean number of concepts contained in students' maps at the end of each session with Betty's Brain. A 2X2 ANOVA with repeated measures was used to analyze the data. The factors were: Quiz (Quiz or No Quiz), Query (Query or No Query) & Session (Sessions 1, 2 and 3). The Quiz ($F(1,46)=8.47$, $p<.01$), Session ($F(2,92)=77.3$, $p<.001$) and Quiz by Session ($F(2,92)=20.13$, $p<.001$) effects were significant in the ANOVA. Comparisons of the means (Tukey HSD) indicate that at the end of Session 1 groups did not differ in the number of concepts in their maps, but by the end Session 3, the TEACH only group had significantly more concepts than the FULL and QUIZ groups. The QUERY group had an intermediate number of concepts and did not differ from the other groups. Examination of the maps suggests that maps in the TEACH only group contain many extraneous/irrelevant concepts, whereas the QUIZ group maps contain fewer, but as predicted, more relevant concepts (i.e., more of the concepts contained in the quiz questions). The quiz served to limit the amount of information in students' maps to those concepts that are most important for modeling ecosystem interdependence (e.g., plants, fish, macro-invertebrates, dissolved oxygen, carbon dioxide, waste, bacteria, and sunlight). The intermediate position of the QUERY group may imply that querying, which made students more aware of the causal reasoning process, also helped reduce irrelevant concepts in the concept map.

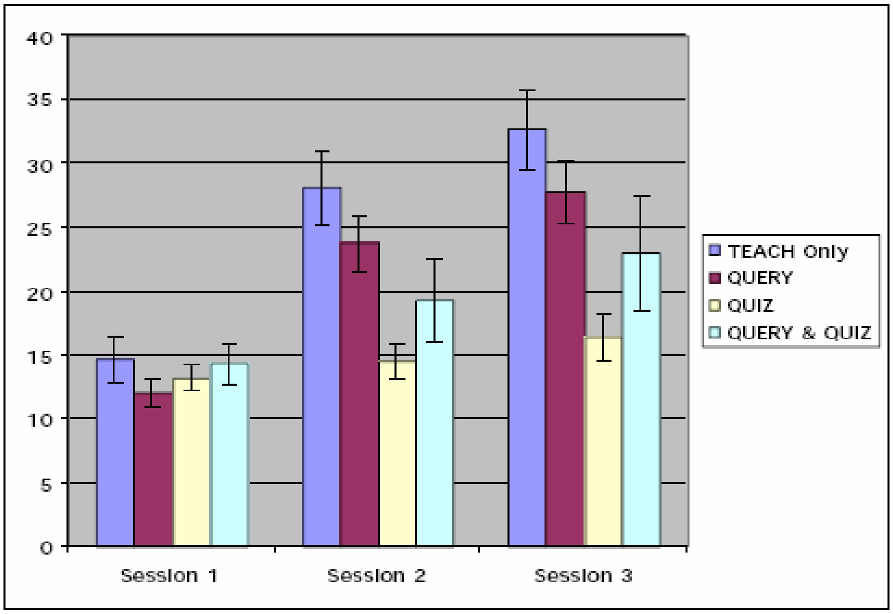


Figure B.36 Number of Concepts in Students' Concept Maps

Figure B.37 displays the data on the proportion of causal links in students' maps. The main effect of Quiz and the Quiz by Session interaction were significant ($F(1,46)=4.82$, $p<.05$ and $F(2,92)=18.52$, $p<.001$, respectively). None of the groups differed at the end of Session 1, but by Session 3 students in the QUIZ and FULL version had a significantly higher proportion of causal links in their maps than students in the TEACH only version. In Session 3 the QUERY group did not differ from and was intermediate to the other groups on this measure. These results suggest that quizzing, along with the teacher feedback, focuses students on modeling causal relationships among concepts. It is precisely these relationships that enable students to understand the more global concepts of balance and interdependence related to ecosystems.

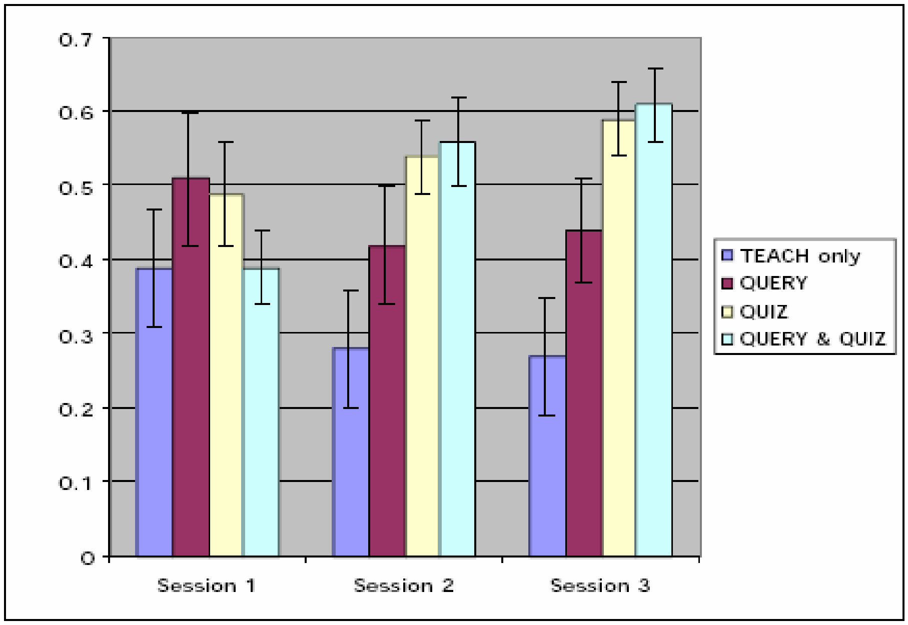


Figure B.37 Proportion of Causal Links in Students' Concept Maps

Figure B.38 shows the ratio of links to concepts in students' maps, and is a measure of the interconnectedness or density of their maps. The main effects of Query was significant ($F(1,46)=4.87$, $p<.05$). Overall, QUERY and FULL students had significantly denser maps in than other students. Evidently, having the opportunity to query Betty, which made the reasoning process more explicit, helped students interrelate concepts in their maps. The effect of Quiz ($F(2,92)=20.10$, $p<.001$) and the Quiz by Session interaction ($F(2,92)=3.36$, $p<.05$) were also significant. Quiz students' maps became increasingly dense over sessions.

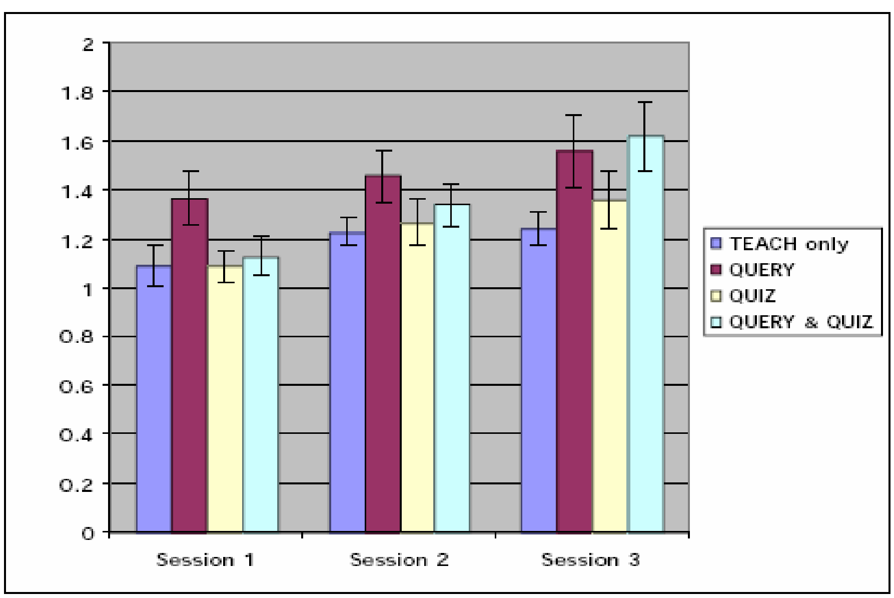


Figure B.38 Ratio of Links to Concepts in Students' Concepts Maps

Figure 5 shows the number of valid causal links contained in students' maps. The main effects of Session ($F(2, 90)=45.39$, $p<.001$) and Query ($F(1,45)=4.92$, $p<.05$) were significant. The Query by Quiz $F(1,46)=4.06$, $p<.05$), Query by Session ($F(2,90)=3.2$, $p<.05$), and Quiz by Session ($F(2,90)=4.69$, $p<.05$) interactions were also significant. Comparisons of the means indicate that by Session 3, QUERY students had significantly more valid links in their maps than students in the TEACH only group. QUIZ and FULL students were intermediate and did not differ from the QUERY and TEACH only groups.

When coding the validity of the links in students' maps credit was given for correct links comprising the quiz questions (i.e., links comprising the teaching expert's map), as well as for other relevant links related to water ecosystems (e.g., "Fungi makes CO_2"). Although the QUERY group had more valid links (expert and relevant combined) than the TEACH only group, the QUIZ and FULL groups had more links from the teaching expert's map than students in the TEACH only group (the QUERY group was intermediate). The main effects of Quiz and Session, and the interaction of Quiz by Session were significant ($F(1,45) = 15.78$, $p<.001$), $F(2,90) = 40.8$, $p<.001$, $F(2,90) = 16.99$, $p<.001$, respectively) in the analysis of the number of expert links comprising students' maps. This shows that students in the quiz conditions were guided by the quiz in determining concepts and relations to teach Betty. However, it was not clear how much global understanding the QUIZ only group had of their concept maps.

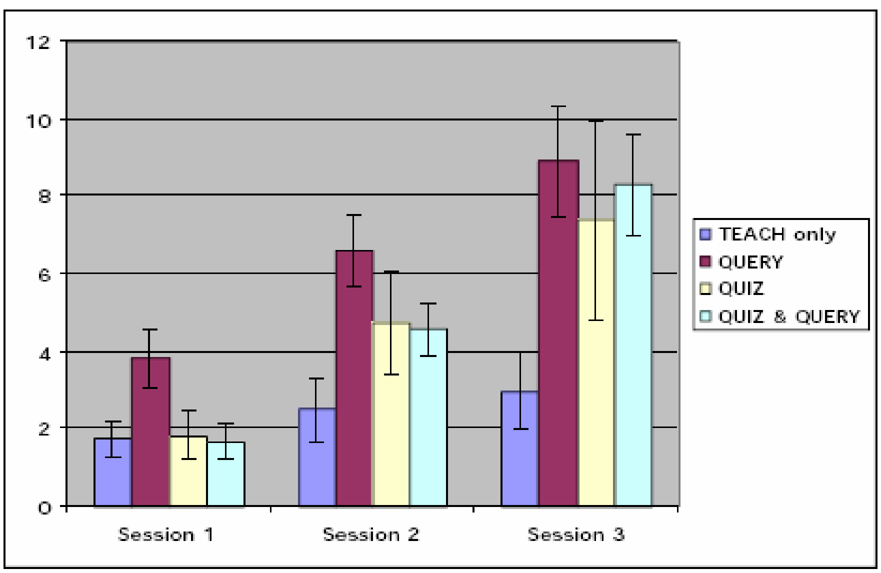


Figure B.39 Number of Valid Causal Links in Students' Concepts Maps

Discussion

Results from the study indicate that both the query and quiz features had beneficial effects on students' learning about ecosystems. In the query feature, Betty models cause-effect reasoning by animating her map and explaining links in her causal reasoning chain. Students who had access to this feature had a significantly higher ratio of links to concepts, indicating that concepts in their maps were more inter-linked. In this way, the query appears to be effective in helping students develop an understanding of the interrelationships of things—living and non-living—in an ecosystem. (It is interesting to speculate as to whether these students would produce longer causal reasoning chains if queried about the consequences of specific changes to an ecosystem.)

Results indicated that providing students with opportunities to quiz their agent decreased the number of irrelevant concepts, increased the proportion of causal information, and increased the number of expert causal links in students' maps. In these ways, the quiz feature was effective in helping students decide the important domain concepts and types of relationships to teach Betty. Students inferred—and reasonably so--that if a concept or relationship was in the quiz, it was important for Betty to know. This notwithstanding, our observations of students during the study suggest that quiz students may have been overly-focused on "getting the quiz questions correct" rather than "making sure that Betty (and themselves) understood the information." We believe that some of this could be attributed to the nature of the suggestions provided by the teacher agent, which led students to focus on making local changes to their maps, and not paying attention to consequences at the level of the (eco)system.

Surprisingly, students in the QUERY condition produced as many valid relevant causal links as the conditions with the quiz feature, and without the benefit of quiz feedback. This demonstrates the value of explicitly illustrating the reasoning process (by having Betty explain her answers) so that students understand causal structures.

The FULL group did not generate significantly higher-quality maps than the QUIZ and the QUERY groups. An investigation of the activity logs revealed a pattern where students' primary focus was to get the quiz questions correct. After getting Betty to take the quiz, they used the teacher agent's hints to make corrections to their maps, and used the query feature only to check if Betty now would answer the questions correctly. They then quickly returned to the quiz mode to see how well Betty performed. In other words, the query mechanism was not used to reflect on the reasoning mechanisms and to gain a deeper understanding of the causal structures they had created, before corrections were attempted on the concept map. As noted above, the feedback we designed for the teaching agent may have inadvertently focused students on making local changes to their maps instead of reasoning more globally in their maps.

These findings suggest that we need to modify the quiz feature so as to focus students on the interrelationships between concepts and the consequences of these relationships on ecosystem. The quiz feature needs to promote more reflective learning by students. Furthermore, in exit interviews a number of students indicated that while they found the overall environment to be quite interesting and easy to work with, they would like Betty to be more active and participatory in the learning process. Betty is passive and only responds when asked questions. We believe that to create a true learning by teaching environment, Betty needs to be more interactive and demonstrate more human-like qualities.

In summary, our goal is to develop Betty's Brain as a generic teachable agent that can be applied to a variety of scientific domains, where reasoning with cause-effect structures helps in learning about the domain. Results indicate that providing students with opportunities to quiz their agent decreases the amount of irrelevant information and increases the proportion of causal information in students' maps; whereas having opportunities to query their agent helps students develop an understanding of the interrelationships of things—living and nonliving—in an ecosystem. The results point to the importance of various forms of feedback when designing teachable agent environments that promote learning.

References

Bargh, J. A. & Schul, Y. (1980). On the cognitive benefits of teaching. *Journal of Educational Psychology, 72(5)*, 593-604.

Biswas, G., Schwartz, D., Bransford, J. & TAG-V. (2001). Technology support for complex problem solving: From SAD environments to AI. In Forbus and Feltovich, (Eds.), *Smart Machines in Education* (pp. 71-98). Menlo Park, CA: AAAI Press.

Chi, M. T.H., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, *25(4*), 471-533.

Graesser, A. C., Person, N., & Magliano, J. (1995). Collaborative dialog patterns in naturalistic one-on-one tutoring. *Applied Cognitive Psychologist*, *9*, 359-387.

Leelawong K., Wang, Y., Biswas, G., Vye, N., Bransford, J., & Schwartz, D. (2001). Qualitative reasoning techniques to support learning by teaching: The Teachable Agents project. In *Proceedings of the Fifteenth International Workshop on Qualitative Reasoning* (pp. 109–116). San Antonio, Texas: AAAI Press.

Novak, J. D. (1996). Concept mapping as a tool for improving science teaching and learning. In: D. F. Treagust; R. Duit; and B. J. Fraser (Eds.), *Improving Teaching and Learning in Science and Mathematics* (pp. 32–43). London: Teachers College Press.

Stoyanov, S., & Kommers, P. (1999). Agent-support for problem solving through concept-mapping. *Journal of Interactive Learning Research*, *10(3/4*), 401-42.

Acknowledgements

This research was supported by National Science Foundation Grant NSF-9873520. The opinions expressed here are those of the authors and should not be attributed to the funding agency. We would like to thank Babs Rector, Suzanne Cassel and Cynthia Mayfield for their invaluable help on this project.

APPENDIX C

QUIZ QUESTIONS IN THE BETTY'S BRAIN ENVIRONMENT

Quiz 1: If wastes increase, what happens to plants?
If wastes increase, what happens to bacteria?
If bacteria increase, what happens to nutrients?
If nutrients increase, what happens to crowded plants?
If crowded plants increase, what happens to sunlight?
If sunlight decreases, what happens to plants?

Quiz 2: If dead organisms increase, what happens to animals?
If dead organisms increase, what happens to bacteria?
If bacteria increase, what happens to dissolved oxygen?
If dissolved oxygen decreases, what happens to animals?

Quiz 3: If sunlight decreases, what happens to animals?
If sunlight decreases, what happens to plants?
If plants decrease, what happens to dissolved oxygen?
If dissolved oxygen decreases, what happens to animals?
If animals decrease, what happens to carbon dioxide?

APPENDIX D

THE MENTOR AGENT'S ON-DEMAND HELP

The *Hint Tree* section describes the structure of the mentor agent's on demand help. Some nodes are marked with text in parentheses. This means they are redundant to previous screens indicated by the parenthesized text. All nodes are static except when they are marked with text in square brackets. Such screens are dynamic and may depend on the patterns that the pattern-tracker agent has detected. See the *Screens and Decisions* section for their details.

Hint Tree

I need help on:

1. Teaching
    1.1. How can I be a good teacher?
        1.1.1. How can I prepare to teach Betty?

Hint Level 1: To be a good teacher, it is important to set goals for your teaching and for Betty's learning. This system has resources to help you learn about rivers. Make sure that you know how to use the online resources about rivers.

- ♦ You may not have time to read all of the resources in one session. The introduction contains the most important information and is a good place to begin reading.
- ♦ Look over the table of contents so you will know how to find important information when teaching Betty.

Hint Level 2: It can be difficult to teach about a complicated subject like rivers. There is so much to think about that it is not easy to remember everything when you are teaching Betty.

- ♦ Think about the relationships described in the resources.
- ♦ Choose one relationship in a river to teach Betty.
- ♦ Teach her about each thing in the relationship using the concept button, then use the link you think is best to teach her the relationship.
- ♦ If you want Betty to reason about how one thing increases or decreases another thing, be sure to use the causal link.

Hint Level 3: Some important concepts in the resources are summed up with words that scientists use to describe very big ideas about rivers (and all ecosystems). Simply teaching Betty these words will not help her understand what they mean.

- ♦ For example, you may want to teach Betty about some ways that things in the river depend on each other, instead of teaching her the word interdependence.

- If you just teach Betty the word interdependence, she will use it, but she will not be able to explain what it means. If she can't explain what it means, she doesn't understand the concept.

Hint Level 4: When reading the resources, you may be tempted to think that all of the information is equally important, but this cannot be true.

- Some information is very important, if you want Betty to understand how a healthy river works.
- Think about the information in the resources and decide what is very important, what is sort of important, and what is not important at all. This will help you decide how to best spend your time teaching Betty.

1.1.2. How do I know what my student is thinking?

Hint Level 1: Teachers are questioners. To be a good teacher, it is important to ask Betty questions so you can find out what she thinks.

Hint Level 2: If you ask Betty questions, her answers will show you what she thinks, based on what you have taught her so far.

Hint Level 3: Good teachers will take the time to find out what their students really think. Asking questions can help, but be sure to find out how Betty reasons about her answer by asking her to explain once she has answered a question.

- Ask Betty questions.
- Have Betty explain her answers by clicking the Explain button.
- Keep clicking the button until her explanation is complete.
- If Betty's explanation is not correct, think about what you need to change in the concept map to improve her explanation.
- Decide if you need to teach Betty more, or if you need to change something you have already taught.
- Make sure you understand what you are teaching.
- If you need more information about rivers, check the river resources.
- If you need more information about teaching with concept maps, check the concept map tutorial.

1.1.3. How can I teach about interdependence and balance?

Hint: Teaching about interdependence and balance is not about putting these two terms into a concept map of the river ecosystem. Read the two sub-topics to see what it is about.

1.1.3.1. Learn about the river ecosystems from the resources.

Hint Level 1: The online resources will help you learn about interdependence and balance in river ecosystems. Become familiar with the resources so you will know how to look up information when you need help teaching Betty.

- Click on the resource button- the table of contents will appear.
- If you click on a topic in the table of contents, it will appear in the resource window at the start of that section.
- You can use the search feature to look up specific information about things in a river.
- Begin teaching Betty by adding concepts to the map.

- ♦ Link concepts together to teach Betty about their relationship.
- ♦ Remember that Betty will reason about how one thing increases or decreases another only if you teach her using the causal link.

<u>Hint Level 2</u>: To help Betty learn about interdependence and balance in river ecosystems you must first understand these ideas so you can teach Betty.

- ♦ Teaching Betty about the relationships among living and non-living things in a river is the best way to teach her about interdependence and balance.
- ♦ To learn more about the important relationships in a river, read the introduction to the resources.
- ♦ To learn more about each relationship, read the sections on **Interdependence** and **Balance**.
- ♦ Ask yourself questions as you read- this will help you find out if you understand the ideas.
- ♦ To get started, think about the relationships in the river and decide which one Betty should learn about first.

1.2. How can I teach Betty using the concept map? (Nuts & Bolts and conceptual)

1.2.1. How do I name the concepts I want to teach?

<u>Hint</u>: To teach Betty about interdependence and balance in a river, make a concept map in the big screen. Betty will use what you put in the map to reason about what happens in a river.

- ♦ The name for a concept should be something in the river that you think is important.
- ♦ The name should be only one word or phrase.
- ♦ Sometimes, you may have to use extra words but the concept should be only one idea.

1.2.2. What kind of links can I use when I teach about relationships?

<u>Hint</u>:

- ♦ If a concept changes the amount of another concept, use a causal link. For example, when we eat vegetables, vegetables decrease. We can represent this idea with a link, "**Human** *eat* (decrease) **Carrots**".
- ♦ If a concept belongs to a group, or is an example of something that you have already taught your student, use a type-of link. For example, "**Carrots** *type of* **Vegetables**".
- ♦ If the concept you would like to teach about does not fit any of the categories above, use a descriptive link. This is especially important if what you want to teach describes a concept, but does not change the amount of the concept you describe. For example, carrots are sold at the market. The sale of carrots does not change the amount or size of the market, and carrots are not a type of market, so the descriptive is the best one to use to teach about this concept. We can represent this idea with a link, "**Carrots** *sold at* **Market**".

1.3. Please give me a suggestion on my teaching process.

<u>Hint</u>: [Teaching Decision Screen]

2. Learning

2.1. How can I be a better learner?

2.1.1. What is the most useful information about river eco-systems?

Hint Level 1: In order for Betty to understand river eco-systems, she must understand the ideas of interdependence and balance. Simply teaching Betty these two words is not enough for her to understand the ideas. Read the introduction to the on-line resources to find out information about interdependence and balance in river eco-systems. If you have already read this information, ask for this type of help again for a different hint.

Hint Level 2: Interdependence and balance depend on the relationships between things that are in the rivers. Some of these things are alive and some are not. Find out what is in a river and teach Betty those things using the concept button. Once Betty know what is in a river, use the link button to teach her the relationship between those things. If one thing helps another thing, then it will probably cause it to increase. If something hinders another thing it might cause it to decrease.

- ♦ Use the causal link to describe the relationships between things where one thing increases or decreases because of the relationship.
- ♦ Betty uses the causal links in the concept map to reason about changes in the river. If you teach her about relationships using the causal link, she can tell you what will happen when something increases or decreases that affect other things in the river. This is what understanding balance in the river is all about how everything affects everything else in a river.
- ♦ If you teach Betty about increase/decrease relationships by using another link, she will not be able to tell you what happens when one thing increases or decreases.

Hint Level 3: Some living things in a river get what they need through an indirect relationship. This means that you must teach Betty about these relationships by linking up several things using causal links. If one thing helps something that helps something else, you must make all of these relationships clear to Betty if you want her to understand balance in a river ecosystem.

- ♦ Use only one word or phrase per concept. Concepts are usually the names of things, and are most often nouns. Make sure that however many words you use, the concept names only one thing.
- ♦ Use a separate link for each relationship. Link labels should contain only one word or idea and should describe the relationship between the two things in the concept boxes- links labels are usually a verb.
- ♦ Once you have named and linked several relationships you will have taught Betty about the balance between those things in a river. Betty will be able to reason about changes in everything that is linked together when one thing increases or decreases.

2.1.2. How can I learn about interdependence and balance?

Hint Level 1: If you want to understand interdependence and balance in a river ecosystem, think about how everything in a river depends on other things that are also in the river. Answering the questions below might help you decide which concepts to teach Betty.

- ♦ What are the living things in a river?

- What are non-living things in a river?
- How do these non-living things get in the river?

Hint Level 2: The living things in a river depend on other living things, and some non-living things, to stay alive. The relationships between the living things and the non-living things are what keep the river in balance. Answering the following questions may help you decide how to link concepts when teaching Betty.

- How do the living things depend on other living things in a river?
- How do the living things depend on the non-living things in a river?
- How do living things affect the other things in a river?
- Does the amount of something change because a living thing depends on it?
- If living things use up something they need, how does it get back in the river so that everything stays balanced?

Hint Level 3: Some things in a river directly affect other things. These things should be connected with only one link between them.

- For example, if one thing eats another thing, these two things directly affect each other.
- Some things in a river indirectly affect other things. Some processes in a river involve many things that all affect each other indirectly. These things should be linked together with one or more link in between.
- Photosynthesis is a process that makes food for some animals in the river. Sunlight is important for photosynthesis to occur. Although the animals do not depend on the sun directly for food, they do depend on their being enough sunlight for photosynthesis to happen. Therefore, some animals indirectly depend on sunlight for their food.

2.1.3. How can I understand what I am reading in the resources better?

Hint Level 1: To be a good learner, it is important to realize when you don't understand what you are reading.

- First, decide if you understand all of the words in the text.
- If you don't understand a word, look it up in the glossary, or in a dictionary.
- If it is highlighted in blue, you can click on the word and get a definition from the resources.

Hint Level 2: If you understand all of the words in the resource section, but you still do not understand what you are reading, you may have missed an important word or phrase because you read too fast.

- Read the section again, but more slowly and carefully.

Hint Level 3: If you understand all of the words and you have read the section carefully, you still might not understand what you are reading because you are missing some important information.

- Ask yourself: What is it that I don't understand?
- Once you have an idea about what you need to understand, you can find more information by reading another section of the resources.

2.1.4. I heard that teachers often learn from their students. How can I do that?

Hint: Sometimes, teachers will find out that they don't understand something when a student asks a question or wants a more complete explanation. Most teachers do not understand every single thing about what they teach unless they have been teaching the subject for a long time.

- ♦ If you do not understand something that is important for Betty to learn, read about it in the resources and ask yourself: what does not make sense to me, what else do I need to understand?
- ♦ Perhaps you think about things differently than what you have read. Think about how your ideas fit in with what you are reading.
- ♦ Once you have thought about your ideas and experience, read the resource section again and ask yourself: do the resources make sense based on what I already know?
- ♦ If the ideas in resources do not make sense to you, you may need more information. Come up with some questions. Your questions will help you decide where to look for more information.

2.2. Please give me a suggestion on my learning process.
Hint: [Learning-Decision Screen]

3. River Ecosystems

3.1. Please help me on the current quiz.
Hint: [Quiz-Help Screen]

3.2. (See Screen 2.1.1.)

3.3. (See Screen 2.1.2.)

3.4. Please evaluate my knowledge about river ecosystems.
Hint: [Map-Evaluation Decision]

3.5. Should I teach my student anything else about the cycles below?

3.5.1. The food chain
Hint: Type **food chain** in the search box of the on-line resources.

3.5.2. The carbon-Dioxide and Oxygen cycles
Hint Level 1: Read the on-line resources about plants, dissolved oxygen and carbon dioxide.
Hint Level 2: Ask Betty questions about plants, dissolved oxygen and carbon dioxide.
Hint Level 3: It is important for Betty to learn enough from you to answer the following questions:

- ♦ How do animals affect dissolved oxygen?
- ♦ How do bacteria affect dissolved oxygen?
- ♦ How do plants affect dissolved oxygen?

3.5.3. Decomposition
Hint Level 1: Read the on-line resources about animals, bacteria, dissolved oxygen, carbon dioxide, waste and dead organisms.
Hint Level 2: Ask [TA Name] questions about animals, bacteria, dissolved oxygen, carbon dioxide, waste and dead organisms.
Hint Level 3: It is important for [TA Name] to learn enough from you to answer the following questions:

- ♦ How does animals' waste affect dissolved oxygen?
- ♦ How do bacteria affect dissolved oxygen?

- ♦ How do dead organisms affect dissolved oxygen?

Also ask her to explain her answer and look at the explanation in details.

Screens and Decisions

Teaching-Decision Screen

If an *AT suggestion* is applicable, fire it.
Else if a *CT suggestion* is applicable, fire it.
Else if an *AQ suggestion* is applicable, fire it.
Else if an *RL suggestion* is applicable, fire it.
Else if an *RS suggestion* is applicable, fire it.

Learning-Decision Screen

If an *AL suggestion* is applicable, fire it.
Else if a *CL suggestion* is applicable, fire it.
Else if an *AQ suggestion* is applicable, fire it.
Else if an *RL suggestion* is applicable, fire it.
Else if an *RS suggestion* is applicable, fire it.

Quiz-Help Screen

If the student has not sent the teachable agent to take any quiz yet, brief the student on the concepts of interdependence and balance in river ecosystems and how to compose a concept map to teach her.

Else if the student is done with all quizzes (receiving full scores for all quiz questions), congratulate the student on successfully helping Betty to reach her goal.

Else if the student is done with this quiz but not with the rest, congratulate the student on passing the quiz and suggest that he or she moves on to another quiz.

Else if the teachable agent cannot answer some questions in the quiz, list these questions and suggest the student to construct his or her concept map so that the teachable agent can use it to answer them.

Else if the score of the current quiz is lower than forty or the increase from the previous score is lower than ten points, ask the student to apply previous hints to improve the concept map before coming back for more help.

Else give the corresponding part of the chain of events for this quiz.

Chain-of-Events Hints on Quiz 1

Level 1: In order for Betty to know what happens to plants when there is too much waste in the river, she must understand the chain of events that happens when the river becomes out of balance because of too much waste. I will give you a small part of this chain at a time. Make sure that the relationships in the chain are in Betty's map.

Here is the first part. When animal waste enters the river, bacteria increase, because waste is food for bacteria. Bacteria decompose waste by breaking it down into smaller parts, from which they take in what they need to live, grow and reproduce.

Level 2: Here is the second part of the waste-plants chain of events. When bacteria decompose waste, they break down the waste and change it into simple molecules called nutrients. These nutrients are chemicals that plants need to live, grow and reproduce. Therefore, if bacteria increase because of waste, nutrients will increase too.

Level 3: This is the third part of the waste-plants chain of events. If there is too much waste, bacteria will multiply a lot, and put too many nutrients in the river. Then the plants will increase because of all the nutrients until there are too many plants. When there are too many plants in a river, they become crowded.

Level 4: This is the final part of the waste-plants chain of events. When crowded plants increase, most will die because they don't get enough sunlight. Plants need sunlight to provide energy for photosynthesis. When crowded plants increase, sunlight to the plants decreases, and if plants don't get enough sunlight, they die. This is the chain of events that can cause a river to get out of balance because of too much waste.

Chain-of-Events Hints on Quiz 2

Level 1: In order for Betty to know what happens to animals when there are too many dead organisms in the river, she must understand the chain of events that happens when the river becomes out of balance because of too many dead organisms. I will give you a small part of this chain at a time. Make sure that the relationships in the chain are in your map.

When plants and animals in a river die, bacteria decompose them and change their remains into nutrients. If there are many dead organisms, there will be a lot of extra food for bacteria, and the bacteria will increase.

Level 2: Here is the second part of the dead-organisms-animals chain of events. When decomposing dead organisms, bacteria use dissolved oxygen, because bacteria use up oxygen in decomposition. Most bacteria also need oxygen to live, grow, and reproduce. Therefore, when bacteria increase, dissolved oxygen decreases.

Level 3: Here is the third part of the dead-organisms-animals chain of events. In order for Betty to know what happens to animals when there are too many dead organisms in the river, she must understand the chain of events that happens when the river becomes out of balance because of too many dead organisms.

Level 4: This is the forth part of the dead-organisms-animals chain of events. When plants and animals in a river die, bacteria decompose them and change their remains into nutrients. If there are many dead organisms, there will be a lot of extra food for bacteria, and the bacteria will increase.

Level 5: This is the final part of the dead-organisms-animals chain of events. If there is a lot of bacteria in the river water, they will use up so much dissolved oxygen that there will not be enough left for the fish and macro-invertebrates. When animals in the river do not get enough oxygen, they suffocate and die. Therefore, when dissolved oxygen decreases, animals decrease too. This puts even more dead organisms in the rivers, and more bacteria. This is the chain of events that cause the river to get out of balance because of too many dead organisms.

Chain-of-Events Hints on Quiz 3

Level 1: In order for Betty to know what happens to animals when sunlight on the river decreases, she must understand the chain of events that happens when plants do not get enough sunlight. I will give you a small part of this chain at a time. Make sure that the relationships in the chain are in Betty's map.

When there are too many plants in a river, they become crowded. When crowded plants increase, most will die because they don't get enough sunlight. Therefore if sunlight decreases, plants will decrease too.

Level 2: Here is the second part of the sunlight-animals chain of events. Plants produce oxygen when they make food through the process of photosynthesis. This oxygen becomes dissolved in the river water. Therefore, when plants decrease, dissolved oxygen decreases too.

Level 3: Here is the third part of the sunlight-animals chain of events. Animals that live in the river need dissolved oxygen to live, grow and reproduce. When animals in the river do not get enough oxygen, they suffocate and die. Therefore, when dissolved oxygen decreases, animals decrease too.

Level 4: Here is the final part of the sunlight-animals chain of events. Animals breathe in dissolved oxygen, and breathe out carbon dioxide. Plants use the carbon dioxide produced by animals in photosynthesis. When animals decrease, carbon dioxide will decrease too. This is the chain of events that cause a river to get out of balance when there is too little sunlight for the plants.

Map-Evaluation Decision

If this quiz is the same as the previous quiz
  - If the score decreases, tell the student that the map looks worse.
  - Else if the score is three points better or worse than the previous score, tell the student that what he or she has modified to the concept map since the last quiz did not improve the concept map.

Else if the score is greater than 90 points,
  - If the previous score is less, praise the student about the improved quality of the concept map.
  - Else tell the student that the part of the concept map that answers the particular quiz is as good as that of an expert in river ecosystems.

Else if the score is greater than 75 points,

If the previous score is less, praise the student about the improved quality of the concept map.
Else tell the student that the part of the concept map that answers the particular quiz has three quarters of concepts and links in that of an expert in river ecosystems.

Else if the score is greater than 50 points,

If the previous score is less, praise the student about the improved quality of the concept map.
Else tell the student that the part of the concept map that answers the particular quiz has half of concepts and links in that of an expert in river ecosystems.

Else if the score is greater than 10 points,

If the previous score is less, praise the student about the improved quality of the concept map.
Else tell the student that the part of the concept map that answers the particular quiz has half of concepts and links in that of an expert in river ecosystems.

Else tell the students to read resources and make sure they understand them before teaching.

Suggestions

Every time a suggestion is accessed, it gives the lowest level of hints that has not been given before. When all hints in a suggestion have been given, it returns null when accessed.

AL Suggestion (Ask a question to learn better)

<u>Level 1</u>: Ask [TA Name] some questions. I haven't seen you doing that today. You can learn a lot by observing how " + name + " answers questions.
<u>Level 2</u>: I haven rarely seen you asking [TA Name] questions today. Do you know that a good learner often ask questions to make sure he understands things correctly?
<u>Level 3</u>: If you don't know what question to ask [TA Name], ask your self these questions:

- What is one important thing I need to understand in order to teach someone else about rivers?
- Am I understanding what I am reading?
- Does what I read match what I already thought about river ecosystems?
- Do I need to change the way I think because of new information, or am I misunderstanding what I am reading in the resources, and my thinking is correct?

<u>Level 4</u>: Another thing you may want to ask yourself: Have I missed an important connection or relationship when I reason about river ecosystems? Do I know about all the important relationships, or should I read further to get more information?

AQ Suggestion (Ask quiz-questions)

Level 1: Ask [TA Name] some quiz questions. I haven't seen you doing that today. Asking [TA Name] those questions are important because you can see how well she can handle questions from experts in the field of river ecosystems.
Level 2: You should ask [TA Name] quiz questions that she has answered wrong in the previous quiz. See if her answer has changed.
Level 3: It's good that you have many questions to ask [TA Name], but asking questions posted by somebody else can give you a new insight to [TA Name]'s learning progress.
Level 4: Asking [TA Name] quiz questions that she has answered wrong can lead you to understand why she got them wrong. Don't look only at the answer but also at how she uses the concept map that you have taught her to get that answer.

AR Suggestion (Ask questions about what you have read in the online resources)

Level 1: You have done a good job reading the resources and use this information to teach [TA Name]. But you also need to make sure that [TA Name] is learning what you are teaching her correctly. Ask her questions about things you have learned from the on-line resources.
Level 2: I saw that you have read the on-line resources on [section]. Pick two concepts from that section to ask [TA Name] a question.

AT Suggestion (Ask a question to teach better)

Level 1: Ask [TA Name] some questions. I haven't seen you doing that today. You need to make sure your student is learning the right things by observing how [TA Name] answers questions.
Level 2: A good teacher asks students questions to make sure they understand things correctly. You can ask [TA Name] by clicking on the **Ask** button underneath her window.
Level 3: Asking [TA Name] questions can help you determine if you are teaching her correctly.

CL Suggestion (Ask more causal questions in order to learn)

Level 1: Ask [TA Name] some causal questions (what happen to **A** when **B** increase/decrease?) I haven't seen you doing that today. You can learn a lot by observing how [TA Name] answers those questions.
Level 2: Ask [TA Name] a causal question (what happen to **A** when **B** increase/decrease?), and check how she reasons through the concept map that you have taught her.
Level 3: I know that you are teaching [TA Name] good things, but you should know that you can learn the qualitative reasoning process from [TA Name]. She is very good at that. Ask her some causal question (what happen to **A** when **B** increase/decrease?)

Level 4: Knowing how to reason about the concept map can give you an advantage in understanding what [TA Name] has learned from you. Read the resources on the reasoning process, ask [TA Name] causal questions, and try to understand the process using her explanation as examples.

CT Suggestion (Ask more causal questions in order to teach)

Level 1: Ask [TA Name] some causal questions (what happen to **A** when **B** increase/decrease?) I haven't seen you doing that today. You can improve the quality of your teaching by observing how [TA Name] answers questions and figure out if something is wrong or if there is something to be added or removed.
Level 2: If you ask [TA Name] a casual question, you can see how she reasons through the concept map that you have taught her. This can help you decide you have taught [TA Name] correctly.
Level 3: Knowing how to reason about the concept map can give you an advantage in understanding what [TA Name] has learned from you. When [TA Name] answers a question in a funny way, it would be easier for you to see why by tracing her reasoning process for that particular answer.

RL Suggestion (Read the on-line resources to learn)

Level 1: I haven't seen you checking the on-line resources today. Use some new words you've seen in the quiz questions in the search boxes of the on-line resources. I am sure you will find something useful.
Level 2: I haven't seen you checking the on-line resources today. The search feature is useful when you want to check the relationship between two concepts. Search the resources using the names of concepts that you are interested in the search boxes.
Level 3: The search feature is useful when you want to check the relationship between two concepts. If some section is highlighted, these two concepts are related in the opinion of a river-ecosystem expert. If no section is highlighted, these two concepts may be related, but their relationship is not critical to the interdependence of the river ecosystem.

RS Suggestion (Reflection strategy)

Main Algorithm

1. If the previous-hint list does not contain a hint from the Quiz-Query category, given hint from this category. Add the hint to the previous-hint list.
2. Else if the previous-hint list does not contain a hint from the Quiz-Resources category, given hint from this category. Add the hint to the previous-hint list.
3. Else if the previous-hint list does not contain a hint from the Query-Edit category, given hint from this category. Add the hint to the previous-hint list.

4. Else if the previous-hint list does not contain a hint from the Query-Resources category, given hint from this category. Add the hint to the previous-hint list.
5. Else if the previous-hint list does not contain a hint from the Resources-Edit category, given hint from this category. Add the hint to the previous-hint list.
6. Else reset the previous-hint list to be empty, and go to step 1.

Query-Quiz Category

1. If the pattern tracker has the Specific-Query-Quiz pattern since the last quiz has been taken:
   Hint: I notice that you have asked [TA Name] some quiz question in the quiz she has taken, and that has improved the concept map you use to teach [TA Name]. Keep up the good work!
2. Else if the pattern tracker has the Nonspecific-Query-Quiz pattern since the last quiz has been taken:
   Hint: I notice that you have asked [TA Name] some quiz question in the quiz she has taken. Keep up the good practice!
3. Else:
   Hint: Asking [TA Name] questions she has answered wrong in the last quiz will help you in figuring out what is needed to help her pass the quiz.

Quiz-Resources Category

1. If the pattern tracker has the Specific-Quiz-Resources pattern since the last quiz has been taken:
   Hint: I notice that you have asked [TA Name] questions to understand what she knows before correcting her, and that has improved the concept map you use to teach her. Keep up the marvelous work!
2. Else if the pattern tracker has the Nonspecific-Quiz-Resources pattern since the last quiz has been taken:
   Hint: I notice that you have looked up the resources to help [TA Name] after she has taken the quiz. Keep up the good practice!
3. Else:
   Hint: Looking up the on-line resources to find more information when the mentor gives you some tips on the quiz will help you in figuring out what is needed to help her pass the quiz.

Query-Edit Category

1. If the pattern tracker has the Specific-Query-Edit pattern since the last quiz has been taken:

Hint: I notice that you have asked [TA Name] questions to understand what she knows before correcting her, and that has improved the concept map you use to teach her. Keep up the marvelous work!

2. Else if the pattern tracker has the Nonspecific-Query-Edit pattern since the last quiz has been taken:
   Hint: I notice that you have asked [TA Name] questions to understand what she knows before correcting her. Keep up the good practice!
3. Else:
   Hint: Asking [TA Name] questions will help you determine what are important and what are not for [TA Name] to know.

Query-Resources Category

1. If the pattern tracker has the Specific-Query-Resources pattern since the last quiz has been taken:
   Hint: I notice that you have asked [TA Name] questions to understand what she knows and double-checked with the on-line resources before correcting her, and that has improved the concept map you use to teach her. Keep up the great work!
2. Else if the pattern tracker has the Nonspecific-Query-Resources pattern since the last quiz has been taken:
   Hint: I notice that you have asked [TA Name] questions to understand what she knows and double-checked with the on-line resources before correcting her. Keep up the good practice!
3. Else:
   Hint: Asking [TA Name] questions will help you determine which parts of resources to look for information about what are important and what are not for [TA Name] to know.

Resources-Edit Category

1. If the pattern tracker has the Specific-Resources-Edit pattern since the last quiz has been taken:
   Hint: I notice that you have checked with the on-line resources before correcting [TA Name], and that has improved the concept map. Keep up the great work!
2. Else if the pattern tracker has the Nonspecific-Resources-Edit pattern since the last quiz has been taken:
   Hint: I notice that you have checked with the on-line resources before correcting her. Keep up the good practice!
3. Else:
   Hint: Looking up the resources will help you determine on what are important and what are not for [TA Name] to know.

APPENDIX E

PATTERNS

There are two types of patterns for the mentor agent, specific and nonspecific. A pattern of events is specific when its events are related, such as taking a quiz and then asking a question from that quiz. Otherwise, the same pattern is classified as a non-specific one. Thus, the patterns are shown here without the classifications (specific / nonspecific).

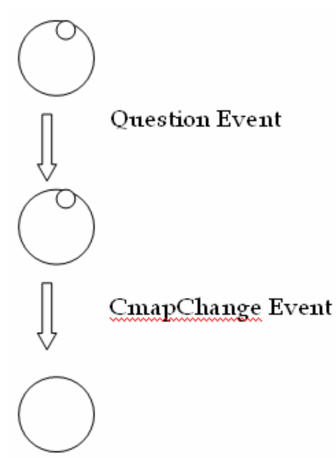


Figure E.40 Query-Edit Pattern

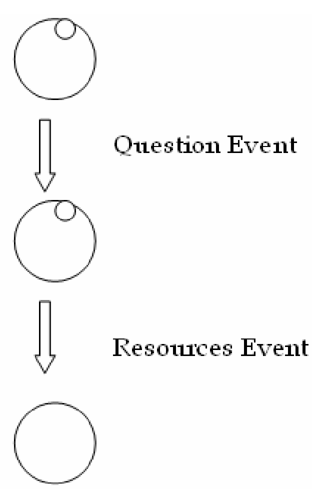


Figure E.41 Question-Resources Pattern

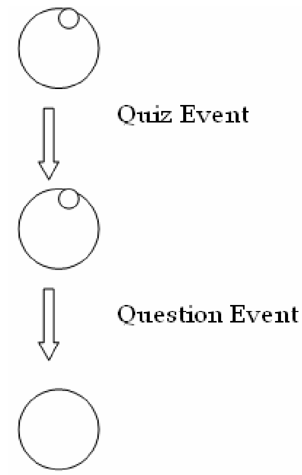


Figure E.42 Quiz-Question Pattern

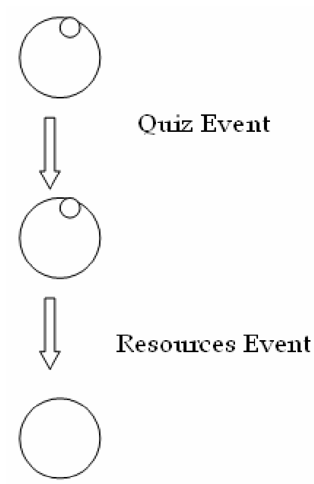


Figure E.43 Quiz-Resources Pattern

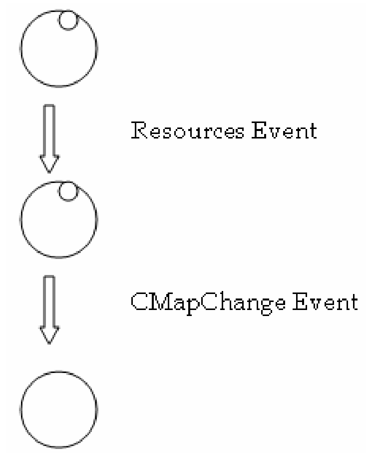


Figure E.44 Resources-Edit Pattern

APPENDIX F

DOMAIN KNOWLEDGE TEST ON THE RIVER ECOSYSTEM

1. We recently purchased an eco-sphere from the Living Systems company. An eco-sphere is a small glass container that holds a small eco-system, and is sealed so that nothing can get in or out. The company told us it was sealed up with water, a few small rocks, a plastic branch, algae, macro-invertebrates, and some bacteria inside. We were told to put it in a window where it will get some sunlight and then leave it alone.
   The company that makes the eco-sphere guarantees that everything that lives in the eco-sphere will stay alive for many years because of interdependence. We are not sure what they mean by interdependence in the eco-sphere.
   Please explain what you think the company means by interdependence in the space below.
2. Some of us thought that it would be a good idea to open up the eco-sphere so we could feed the macro-invertebrates and maybe put in some fish. When we wrote the company with our idea they wrote back and said if we want to do that we should think about the balance in the eco-sphere. We are not sure what the company means by balance in the eco-sphere.
   In the space below, please explain what you think the company means by balance.
3. We decided to make our own eco-spheres and wrote to the company for advice. They wrote back and told us that we must be careful what we put in our eco-spheres. They said if we are not careful, adding certain things could cause a chain of events to happen in the eco-sphere that would not be good for the living things. We are not sure what the company means by chain of events.
   In the space below, please explain what you think the company means by chain of events. It would help us understand if you would also give us an example.
4. Some of us have been wondering if the company put bacteria in the water, or if it is there and they cannot find a way to get it out because it is so small. Some of us think the bacteria help to keep the eco-sphere alive, but some of us think the bacteria could cause a problem, but something in the eco-sphere cleans them up. We have three different ideas about the bacteria, and they are listed below. Please circle the idea that you think is right about bacteria in the eco-sphere.
   a. The bacteria in the eco-sphere are there because bacteria live in water and inside macro-invertebrates, not because the company put bacteria inside. The bacteria in the eco-sphere could make the macro-invertebrates sick, but algae eat bacteria, so it isn't a problem.
   b. The bacteria in the eco-sphere are there because bacteria live in water and inside macro-invertebrates. This is a good thing because bacteria break down waste products and dead organisms and turn these into nutrients that algae use during photosynthesis to make food and oxygen.
   c. The company puts bacteria in the eco-sphere because bacteria are food for the macro-invertebrates. When macroinvertebrates eat bacteria, the bacteria release carbon dioxide that algae use during photosynthesis to make food and oxygen.

5. When we thought about making our own eco-spheres, we decided to do some experiments to help us find out how to take care of our eco-sphere, and how much of each thing we should put in the container. Before we did the experiments, we thought it would be a good idea to make some predictions about what would happen if we did certain things. Please circle the **True** for every prediction you think will turn out to be true in our experiments, and circle **False** for every prediction you think will turn out to be false in our experiments.
    a. If we put the eco-sphere in a dark place, it will be too dark for the fish to find food and they will die. **True** **False**
    b. If we put the eco-sphere in a window where it gets some light in the daytime, the algae will make food and oxygen for the other living things in the eco-sphere, through the process of photosynthesis. **True** **False**
    c. If we put bacteria in the eco-sphere, it will attack the fish and macro-invertebrates and kill them. **True** **False**
    d. If we put bacteria in the eco-sphere, it will make the fish and macro-invertebrates very sick, but they may not die. **True** **False**
    e. If we put bacteria in the eco-sphere it will decompose dead organisms and waste and put nutrients in the water for the algae to use. **True** **False**
    f. If we put bacteria in the eco-sphere, it will eat the fish, algae, and macro-invertebrates. **True** **False**
    g. If we don't put any bacteria in the eco-sphere, everything will be healthy and nothing bad will happen. **True** **False**
    h. If we put algae in the eco-sphere, it will make food and oxygen for the macroinvertebrates through the process of photosynthesis, and this will allow the macro-invertebrates to live for a long time. **True** **False**
    i. If we put a whole lot of algae in the eco-sphere, it will provide the macro-invertebrates with extra food and they will be able to live for a very, very long time. **True** **False**
    j. If we do not put any algae in the eco-sphere, the macroinvertebrates will be better off because they eat bacteria, and without the algae they will have more space to move around. **True** **False**
    k. If we put a lot of extra algae in the eco-sphere, the algae will reproduce a lot. The algae will eventually take over the eco-sphere and make things too crowded for the macro-invertebrates. **True** **False**
    l. If we put a lot of extra algae in the eco-sphere, the algae will reproduce a lot. The algae will eventually become so crowded that most of the algae will not get sunlight and will die. **True** **False**
    m. If we put macro-invertebrates in the eco-sphere, they will eat the bacteria and get very sick. **True** **False**
    n. If we put macro-invertebrates in the eco-sphere, they will eat the algae and get very sick. **True** **False**
    o. If we put macro-invertebrates in the ecosphere, they will produce carbon dioxide that the algae need for photosynthesis. **True** **False**
    p. If we put macro-invertebrates in the ecosphere, they will produce carbon dioxide that the bacteria need to stay alive. **True** **False**

We finally got around to doing our eco-sphere experiments and checking our predictions. In each experiment we did something and several events happened because of what we did. We kept notes on each experiment, but one of us mixed up the pages in our binders and we can't remember in what order the events happened. We used a different binder for each experiment (thank goodness!), but we need help getting our notes back in order.

Below are the events that happened in each experiment. Please put them in order according to what happened first, then second, then third, and so on, to make a complete chain of events. Please put the number 1 by the thing you think happened first, the number 2 by the thing you think happened second, the number 3 by the third event, and so on.

6. In experiment 1, we made an eco-sphere and put in a lot of extra algae.

____________ Macro-invertebrates die.

____________ Bacteria multiply a lot.

____________ Algae begin to die.

____________ Sunlight is blocked to most of the algae.

____________ Algae reproduce a lot and become very crowded.

____________ Bacteria feed on dead algae.

____________ Bacteria use up most of the dissolved oxygen.

____________ Macro-invertebrates seem fine for 2 weeks.

7. In experiment 2, we made an eco-sphere and didn't put any algae in it.

____________ Bacteria decompose dead macro-invertebrates.

____________ Macro-invertebrates run out of dissolved oxygen.

____________ Macro-invertebrates die.

____________ Macro-invertebrates run out of food.

____________ Macro-invertebrates seem fine for 2 days.

8. In experiment 3, we made an ecosphere and put lots of extra bacteria inside.

____________ Macro-invertebrates run out of dissolved oxygen.

____________ Bacteria multiply a lot.

____________ Bacteria use up most of the dissolved oxygen.

____________ Macro-invertebrates seem fine for 2 days.

____________ Macro-invertebrates die.

____________ Bacteria decompose dead macro-invertebrates.

9. In experiment 4, we made an eco-sphere and put in extra macro-invertebrates.

____________ Bacteria decompose dead macro-invertebrates.

____________ Macro-invertebrates run out of dissolved oxygen.

____________ Macro-invertebrates seem fine for a week.

____________ Macro-invertebrates die.

____________ Macro-invertebrates wastes begin to pile up in the eco-sphere.

____________ Algae becomes scarce.

____________ Macro-invertebrates run out of food.

____________ Macro-invertebrates eat most of the algae.

10. In experiment 5, we made an eco-sphere and we didn't put any macro-invertebrates inside.

____________ Algae run out of carbon dioxide.

____________ Algae cannot make food through photosynthesis.

____________ Algae cannot reproduce.

____________ Algae seem fine for a month.

____________ All algae eventually die.

____________ Bacteria decompose dead algae.

11. In experiment 6, we made an eco-sphere and put it under a grow light 24 hours each day.

____________ Macro-invertebrates die.

____________ Bacteria multiply a lot.

____________ Algae become very crowded.

____________ Bacteria feed on dead algae.

____________ Bacteria use up most of the dissolved oxygen.

____________ Algae begin to die.

____________ Light is blocked to most of the algae.

____________ Algae reproduce a lot.

____________ Grow light causes photosynthesis in algae 24 hours a day.

____________ Bacteria decompose dead algae and macro-invertebrates.

12. In experiment 7, we made an eco-sphere and put it in a dark place where it would never get any light.

____________ Macro-invertebrates die

____________ Bacteria multiply a lot

____________ Bacteria feed on dead algae

____________ Bacteria use up most of the dissolved oxygen

____________ Algae begin to die

____________ Algae do not get any sunlight

____________ Algae cannot create food or oxygen through photosynthesis

____________ Bacteria decompose dead algae and macro-invertebrates

APPENDIX G

MOTIVATED STRATEGIES FOR LEARNING QUESTIONNAIRES

The following scales and their questions are selected from the Motivated Strategies for Learning Questionnaire (MSLQ) developed by Pintrich and DeGroot (1990). The number in front of each item indicates its actual position on the questionnaire. Items marked with "(* R)" were reflected before scale construction. There were 56 questions on the full questionnaire, but only forty-four items we may use are presented here. Note that these items are original and their wording may need to be modified to fit our study.

Questionnaires

Motivational Beliefs

A. Self-Efficacy

2. Compared with other students in this class I expect to do well.
7. I'm certain I can understand the ideas taught in this course.
10. I expect to do very well in this class.
11. Compared with others in this class, I think I'm a good student.
13. I am sure I can do an excellent job on the problems and tasks assigned for this class.
15. I think I will receive a good grade in this class.
20. My study skills are excellent compared with others in this class.
22. Compared with other students in this class I think I know a great deal about the subject.
23. I know that I will be able to learn the material for this class.

B. Intrinsic Value

1. I prefer class work that is challenging so I can learn new things.
5. It is important for me to learn what is being taught in this class.
6. I like what I am learning in this class.
9. I think I will be able to use what I learn in this class in other classes.
12. I often choose paper topics. I will learn something from them even if they require more work.
17. Even when I do poorly on a test, I try to learn from my mistakes.
18. I think that what I am learning in this class is useful for me to know.

21. I think that what we are learning in this class is interesting.
25. Understanding this subject is important to me.

C. Test Anxiety

3. I am so nervous during a test that I cannot remember facts I have learned.
14. I have an uneasy, upset feeling when I take a test.
24. I worry a great deal about tests.
27. When I take a test, I think about how poorly I am doing.

Self-Regulated Learning Strategies

D. Cognitive Strategy Use

30. When I study for a test, I try to put together the information from class and from the book.
31. When I do homework, I try to remember what the teacher said in class so I can answer the questions correctly.
33. It is hard for me to decide what the main ideas are in what I read. (* R)
35. When I study, I put important ideas into my own words.
36. I always try to understand what the teacher is saying even if it doesn't make sense.
38. When I study for a test, I try to remember as many facts as I can.
39. When studying, I copy my notes over to help me remember material.
42. When I study for a test I practice saying the important facts over and over to myself.
44. I use what I have learned from old homework assignments and the textbook to do new assignments.
47. When I am studying a topic, I try to make everything fit together.
53. When I read material for this class, I say the words over and over to myself to help me remember.
54. I outline the chapters in my book to help me study.
56. When reading I try to connect the things I am reading about with what I already know.

E. Self-Regulation

32. I ask myself questions to make sure I know the material I have been studying.
34. When work is hard, I either give up or study only the easy parts. (* R)
40. I work on practice exercises and answer end of chapter questions even when I don't have to.
41. Even when study materials are dull and uninteresting, I keep working until I finish.
43. Before I begin studying I think about the things I will need to do to learn.
45. I often find that I have been reading for class but don't know what it is all about. (* R)

46. I find that when the teacher is talking I think of other things and don't really listen to what is being said. (* R)
52. When I'm reading I stop once in a while and go over what I have read.
55. I work hard to get a good grade even when I don't like a class.

Results

Figure G.1 displays the pre- and posttest MSLQ scores grouped accordingly to its manual (Pintrich 1991). There was no difference in scores between groups but over time as shown in Table G.1.

Even though the decreases in the scores after the main study were unexpected, they were not surprising. The participants in this study, being convenient samples, were average to high-achieving students and had been successful in completing their school work to a degree. In this study, they had to face a challenge that asked them to change the ways they think and act. The challenge asked the participants to teach the domain they did not have the expertise, the river ecosystem, with the features that were novel to them. Thus, the drops in the self-evaluated test could be a positive effect because the difficulty of the challenge helped the participants to realize that there were more to learn outside classroom textbooks. Using the Betty's Brain environment was comparable to receiving an external-assessment tool on how students approached learning, which was important to independent learning.

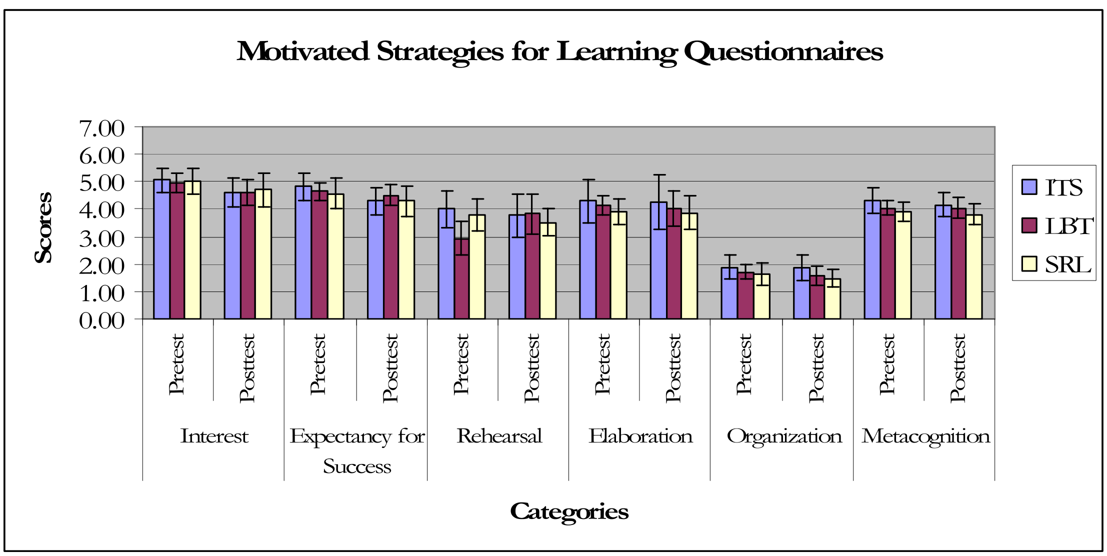


Figure G.1 MSLQ Results by category

Table G.1 Repeated-Measures Analyses of Variance on the MSLQ Scores

Categories	Time	Time * Group

Interest	$F_{(1,41)} = 11.41, p < .005$	$F_{(2,41)} = .17, p = .85$
Expectancy of Success	$F_{(1,41)} = .35, p < .005$	$F_{(2,41)} = 1.5, p = .24$
Rehearsal	$F_{(1,41)} = 11.41, p = .56$	$F_{(2,41)} = 2.9, p = .07$
Elaboration	$F_{(1,41)} = .19, p =.66$	$F_{(2,41)} = .01, p =.99$
Organization	$F_{(1,41)} = .71, p < .41$	$F_{(2,41)} = 1.3, p < .90$
Meta-cognition	$F_{(1,41)} = .59, p = .45$	$F_{(2,41)} = .23, p = .80$

APPENDIX H

THE EXPERT CONCEPT-MAP

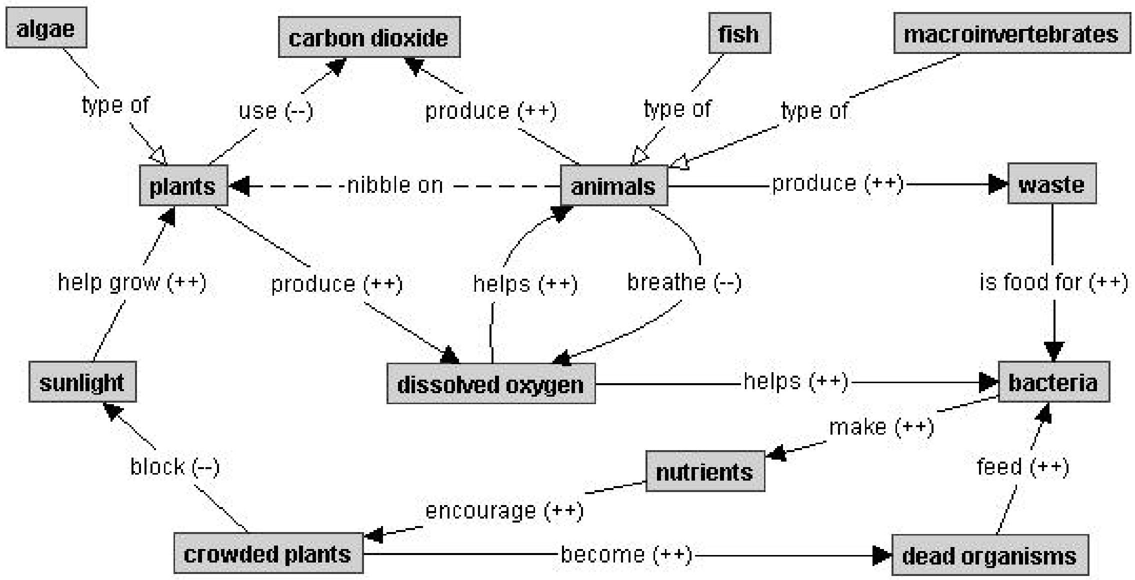


Figure H.2 The expert concept-map

APPENDIX I

GRADING PROGRAMS

The grading programs were developed by Krittaya Leelawong in Java (Java 2 SDK version 1.4.02) with help from John Kilby, Joan Davis, and Karun Viswanath. Figure I.3 illustrates the hierarchy of the `BettyGrader` project that is structured as follows:

- `gradeCounter`: Counts the frequencies of grade codes from various grading tasks and summarizes them by students IDs
- `log`: Discovers the behaviors and patterns of the users from their information saved in log files during the use of the Betty's Brain environment
    - `logcounter`: Counts the frequencies of using specific features of the Betty's Brain environment
    - `patterncounter`: Searches for patterns from series of activities
- `maps`: Grades concepts and links in students' by student and by session
    - `interconnectivity`: Counts the number of links in each answer to a given set of causal questions
    - `mapcomparator`: Identifies which concepts and links two maps share and which each of them exclusively contains
    - `mapgrader`: Grades the correctness of concepts and links in students' by student and by session and also manages the databases of concepts' and links' correctness grades
    - `quizgrader`: Grades students' maps according to a given set of causal questions
- `preposttest`: Grades items in the pre- and post-test
    - `ordering`: Grades students' answers to the ordering questions
    - `scoring`: Grades students' answers to the multiple-choice and true-or-false questions
- `toolbox`: maintains shared code among the above packages

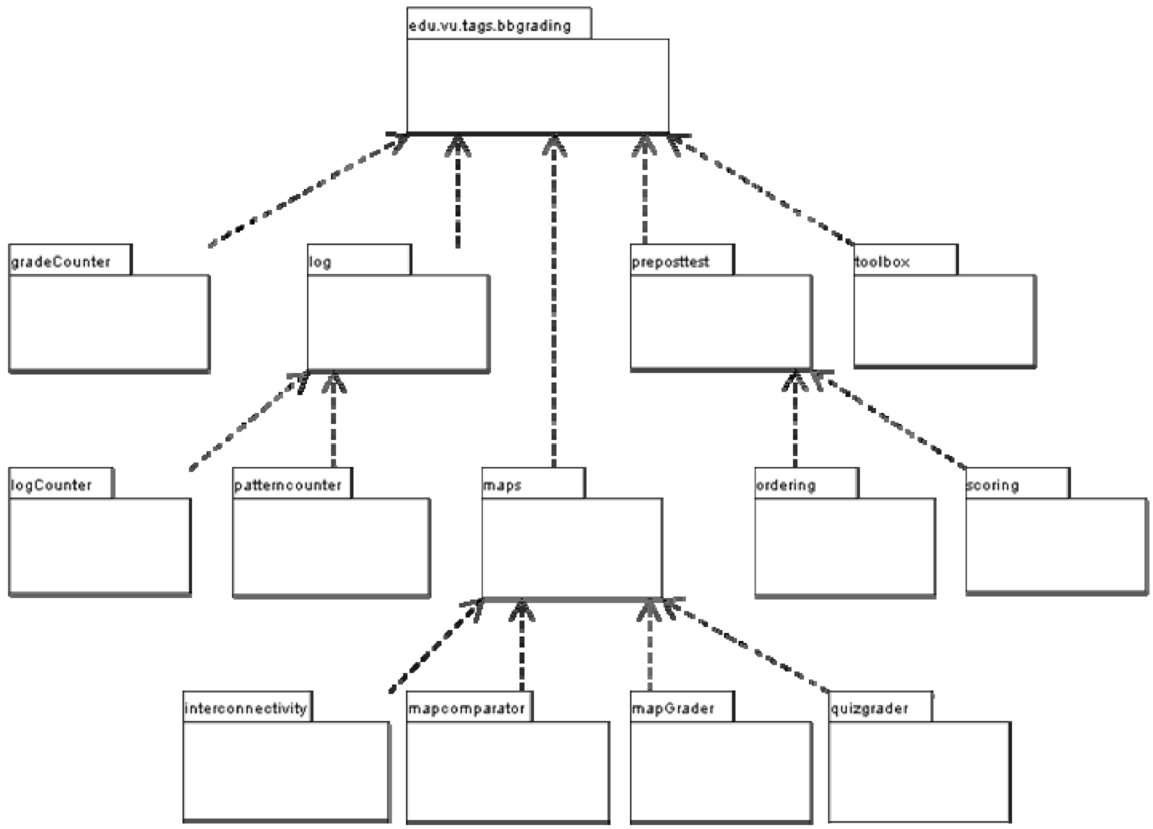


Figure I.3 Overview of Packages in the `BettyGrader` Project

These programs were tested by a black-box method. Because outputs of these programs are by student ID, for each set of outputs six IDs were randomly selected and their outputs were verified by a human grader to have the desired outcomes. If there were any mistakes in the outputs of any program, it and its related classes will be examined and corrected, and three more cases were be verified. This process was repeated until the output was 100% correct. The testers were Mathew Lee and Amanda L. Peltz.

Because there are too many classes in this project to be mentioned in details, this appendix introduces only the classes having user interfaces in each package. In the case that several runnable classes extend the same class, only the base class is presented.

Figure I.4 displays the classes in the `gradeCounter` package and its relationship to other packages. The `GradeCounter` class, shown in Figure I.5, is the base class for all grade-counting classes in this project.

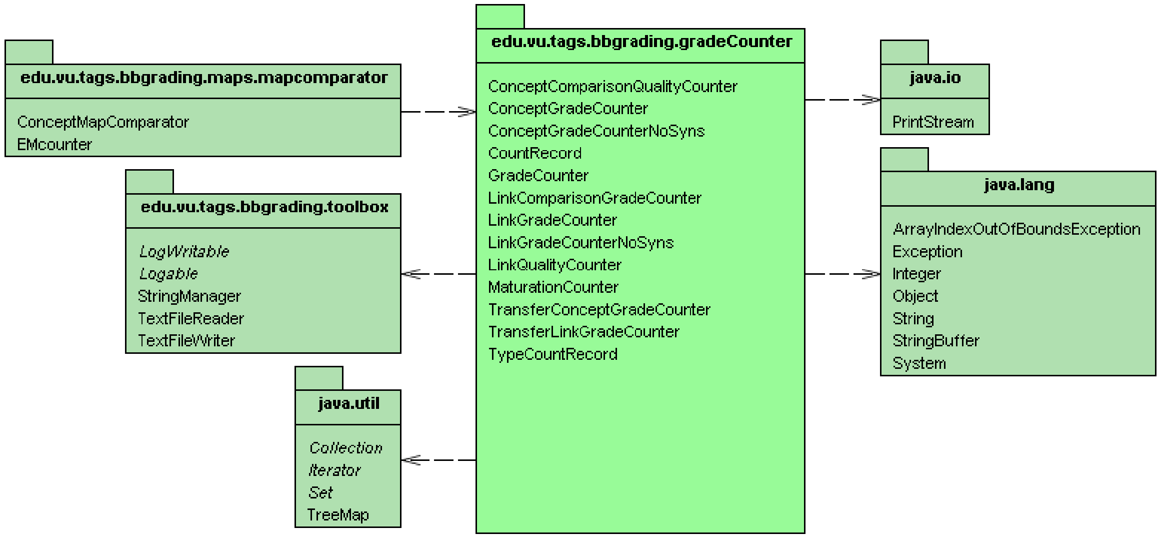


Figure I.4 The `gradeCounter` Package and Its Dependency

There are two methods that other programs can call this class and its sub-classes:

1. `countAndWrite()`'s: Reads grades from a text file, count them, and write the results to another text file
2. `count()`'s: Reads grades from a text file, count them, and return the results in a `TreeMap` object

A sample of comma-delimited input files is shown in Figure I.6. We can specify the indices of the input file to let the program know which column is the student ID and which columns have the grades to count. This means the grade can be a composite of more than one column of data. In this case, the student ID is in the first column and the grade is in the forth column. The counts are shown in Figure I.7.

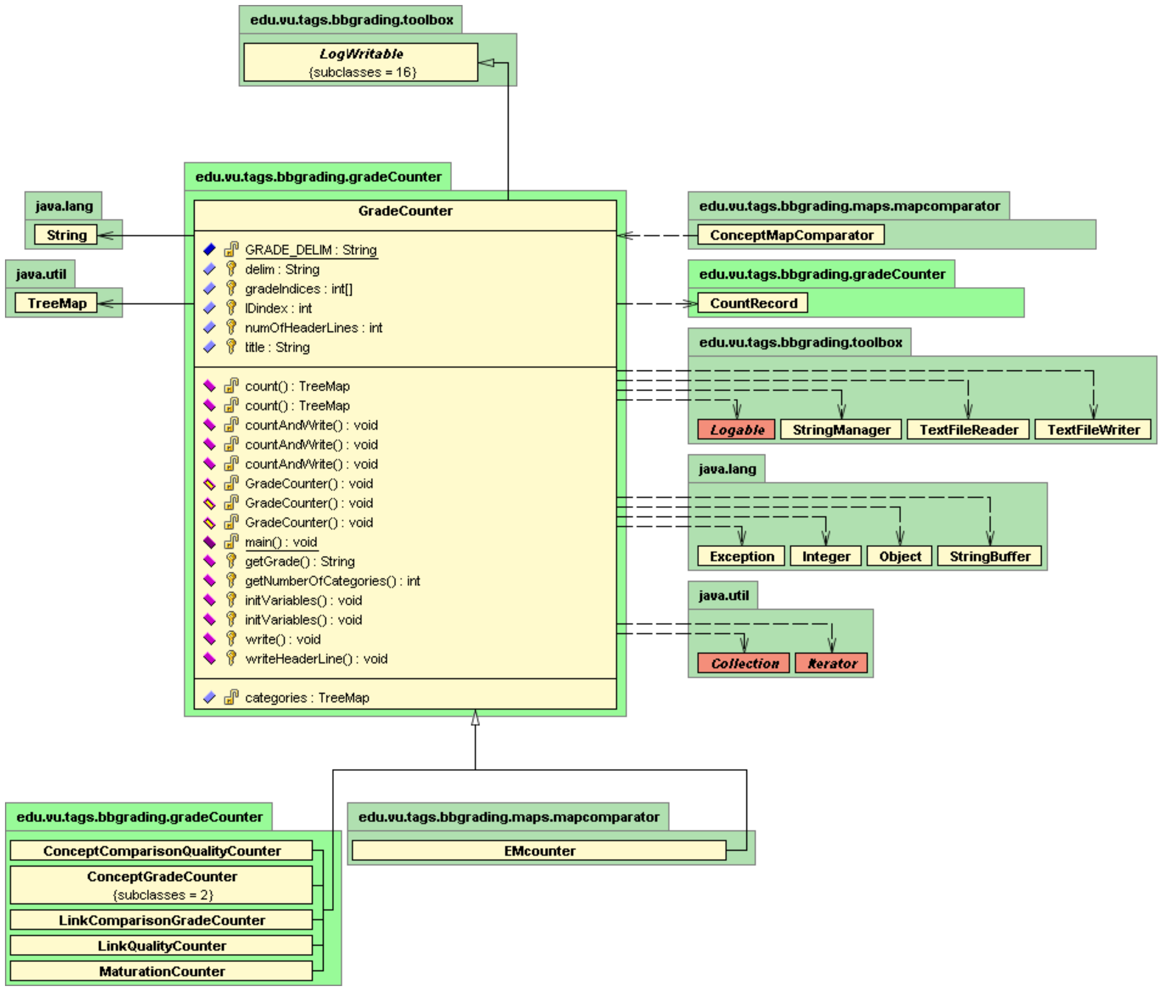


Figure I.5 The `GradeCounter` Class

ID	Label	Status	Relevance
111	dead organisms		E
111	carbon dioixde		E
111	nutrients		E
111	wastes		E
112	animals	Done	E
112	bacteria	Done	E
112	carbon dioxide	Done	E
112	crowded plants	Done	E
112	dead organisms	Done	E
112	decreased animals	Graded	U
112	decreased sunlight	Graded	U
112	dissolved oxygen	Done	E
112	increased dead organisms	Graded	U
112	nutrients	Done	E
112	plants	Done	E
112	reduced sunlight	Graded	U
112	sunlight	Done	E
112	wastes	Done	E
113	animals	Done	E
113	bacteria	Done	E

Figure I.6 A Sample Input File for a `GradeCounter` Class

Student ID	E	R	I	U
111	10	0	0	0
112	10	0	0	4
113	10	0	0	0
114	11	4	1	2
115	10	0	0	0
116	12	5	1	0
117	10	1	0	0
118	9	2	1	1
121	13	8	0	1
122	8	0	0	1

Figure I.7 A Sample Output File for a `GradeCounter` Class

The details of the `logcounter` package and its dependency to other packages are shown in Figure I.8. There are two runnable classes:

- `ActivityTimer` (Figure I.9): Sums the time spending doing specific activities
- `LogCounter` (Figure I.10): Counts the frequencies of specific activities conducted in the Betty's Brain system

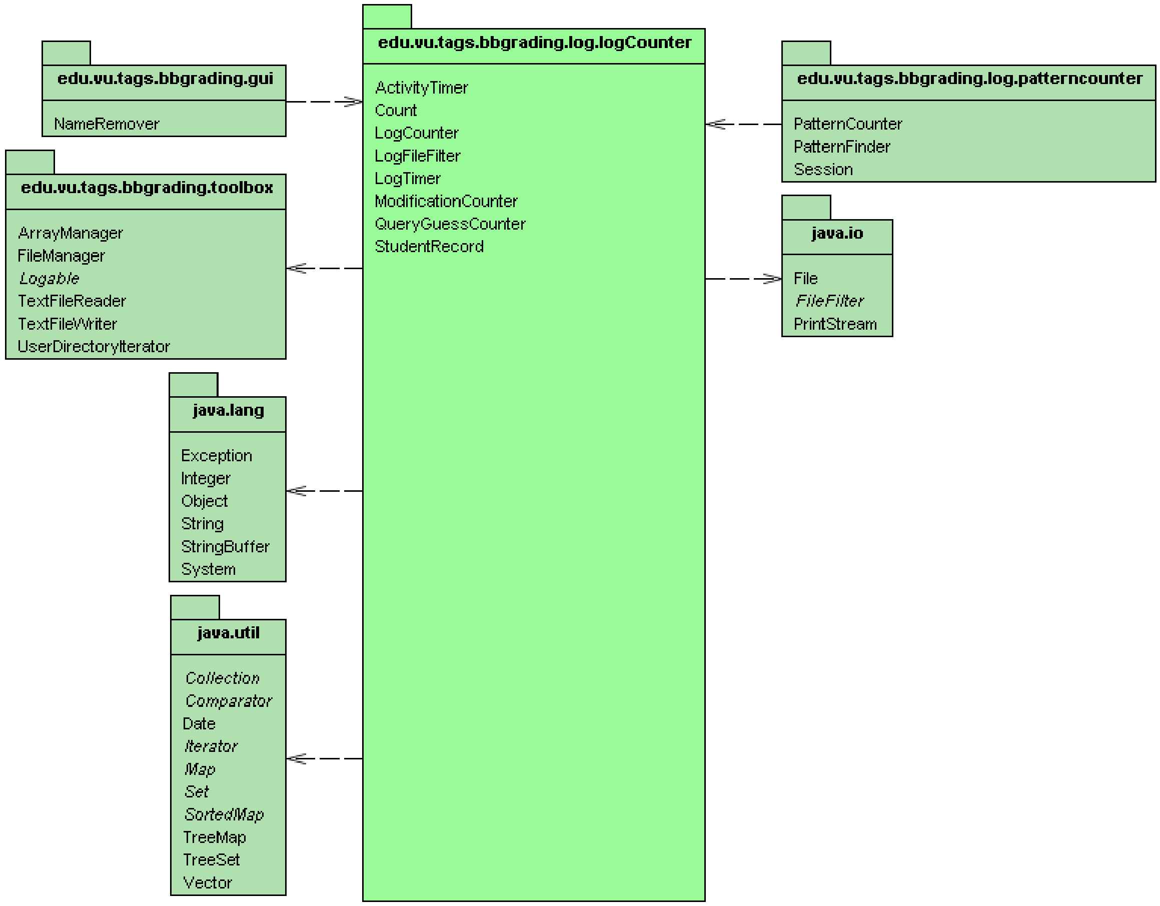


Figure I.8 The `logCounter` Package and Its Dependency

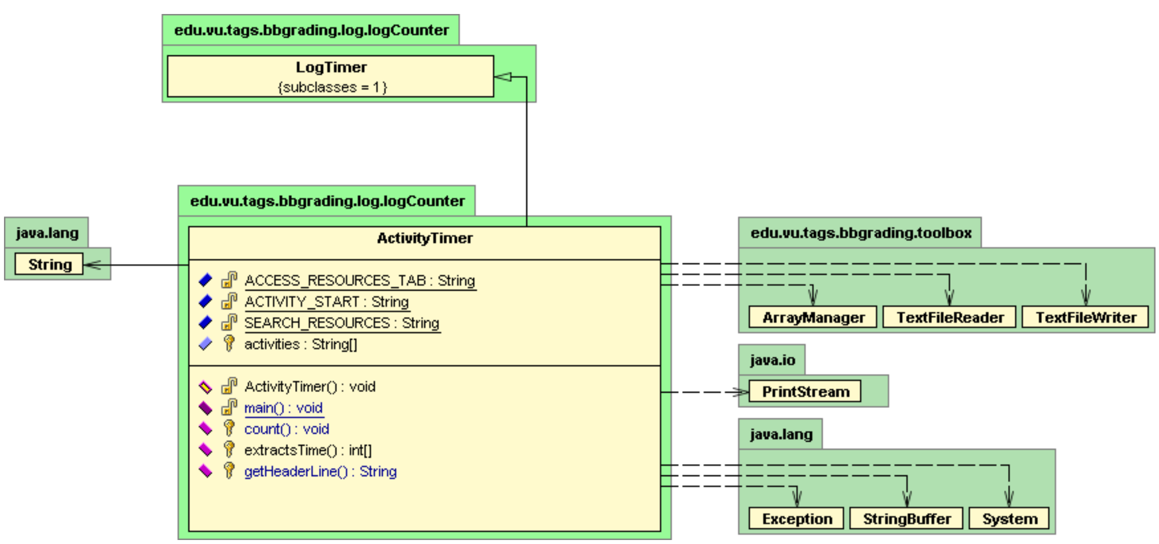


Figure I.9 The `ActivityTimer` Class

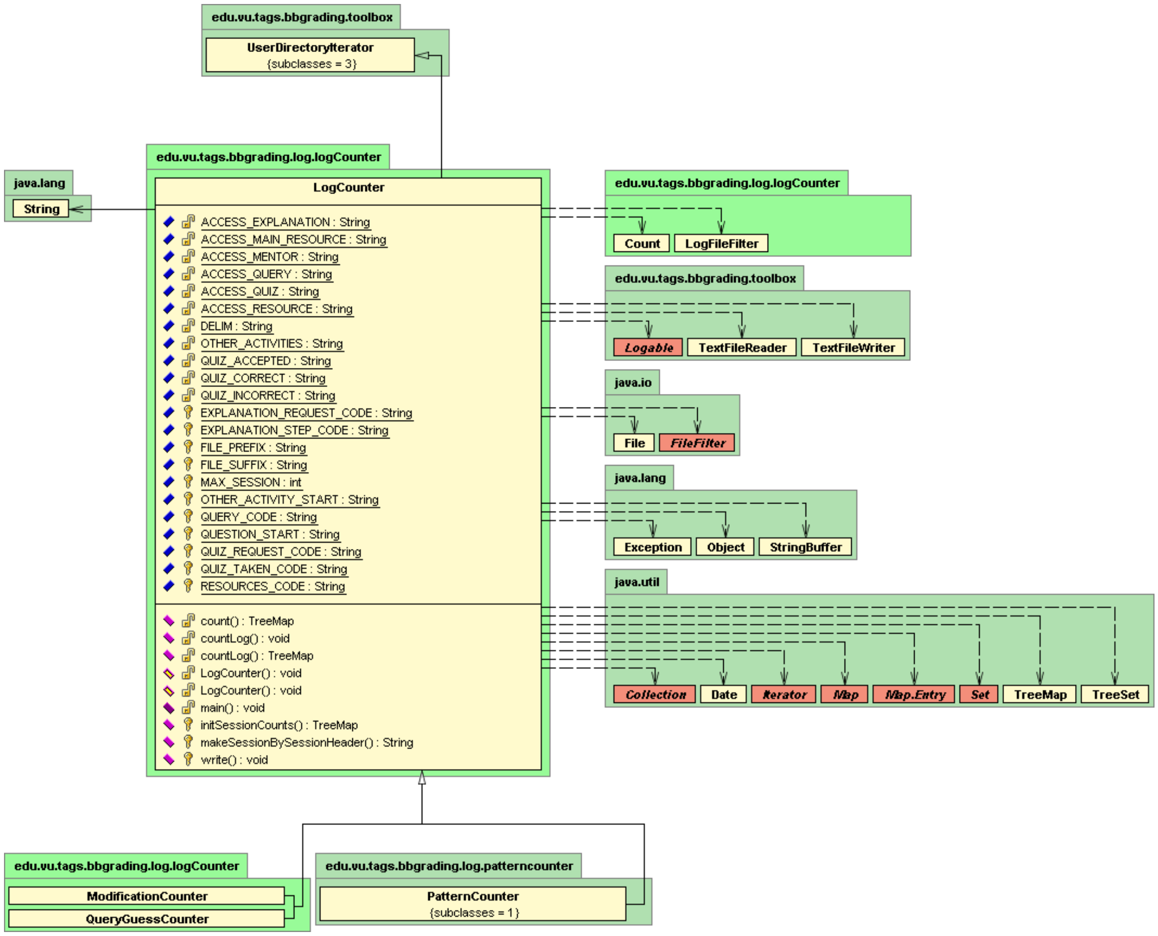


Figure I.10 The `LogCounter` Class

The details of the `patterncounter` package and its dependency to other packages are shown in Figure I.11. This package has not been fully developed yet because of several difficulties in accurately detecting patterns. There are two runnable classes:

- `PatternCounter` (Figure I.12): Is a simplified version of the `Pattern-Finder` class focusing on only three patterns
- `PatternFinder` (Figure I.13): Counts the frequencies of pre-defined patterns in log files

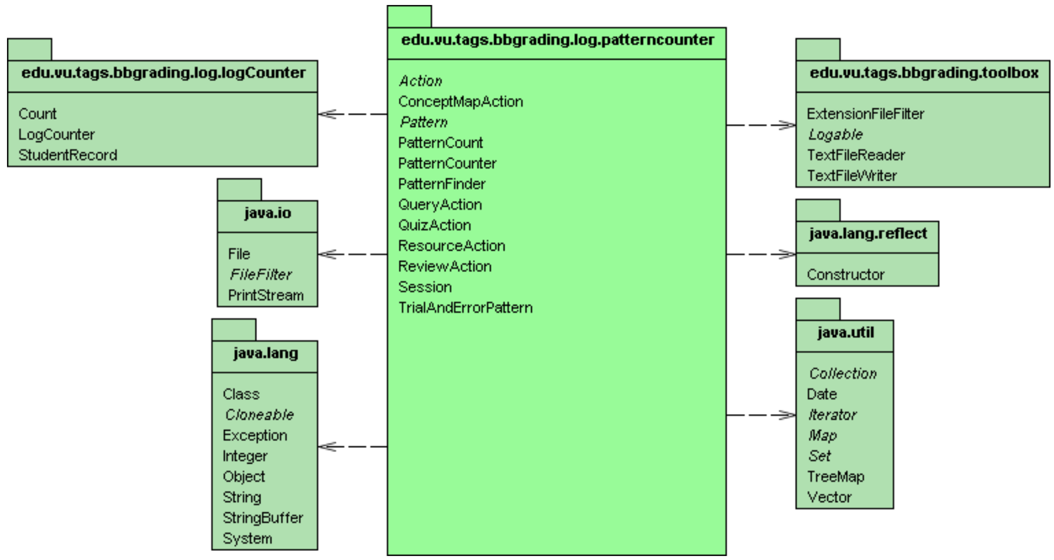


Figure I.11 The `patterncounter` Package

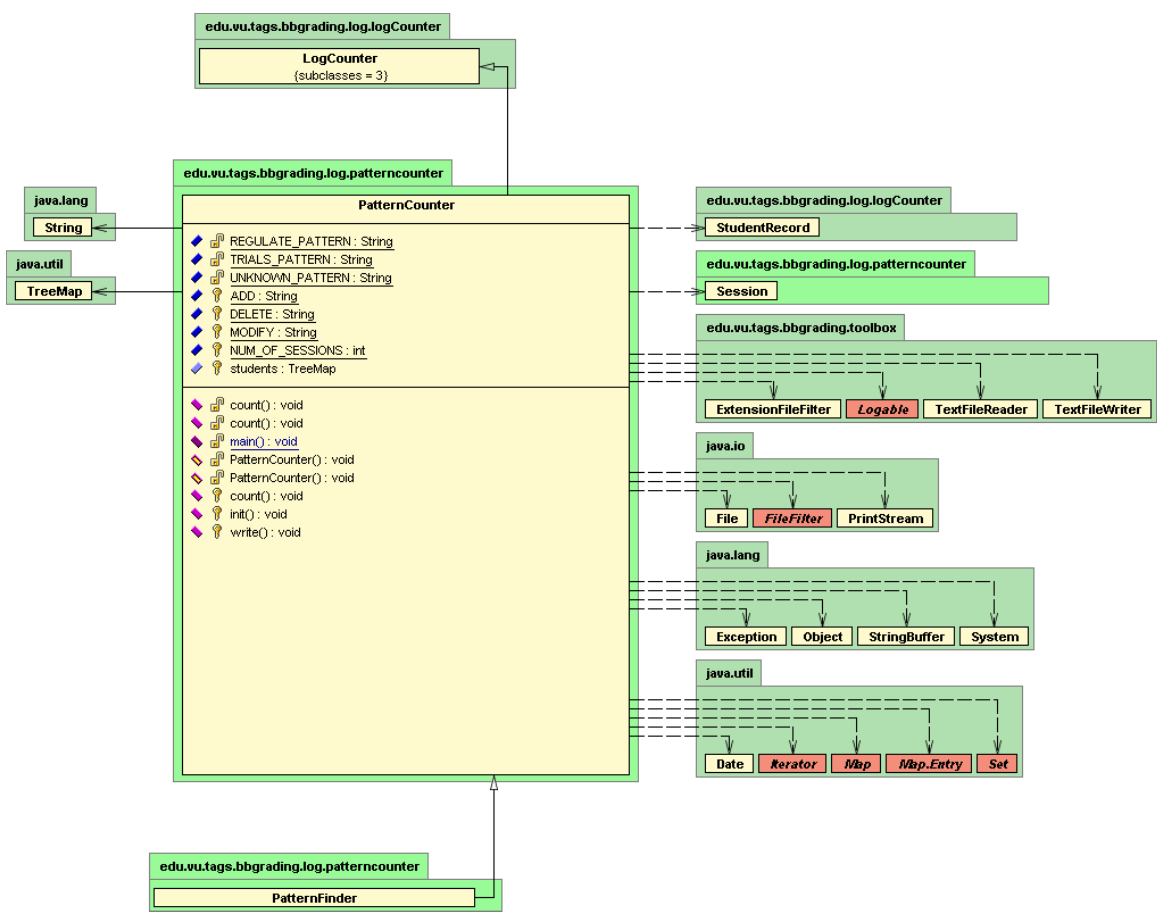


Figure I.12 The `PatternCounter` Class

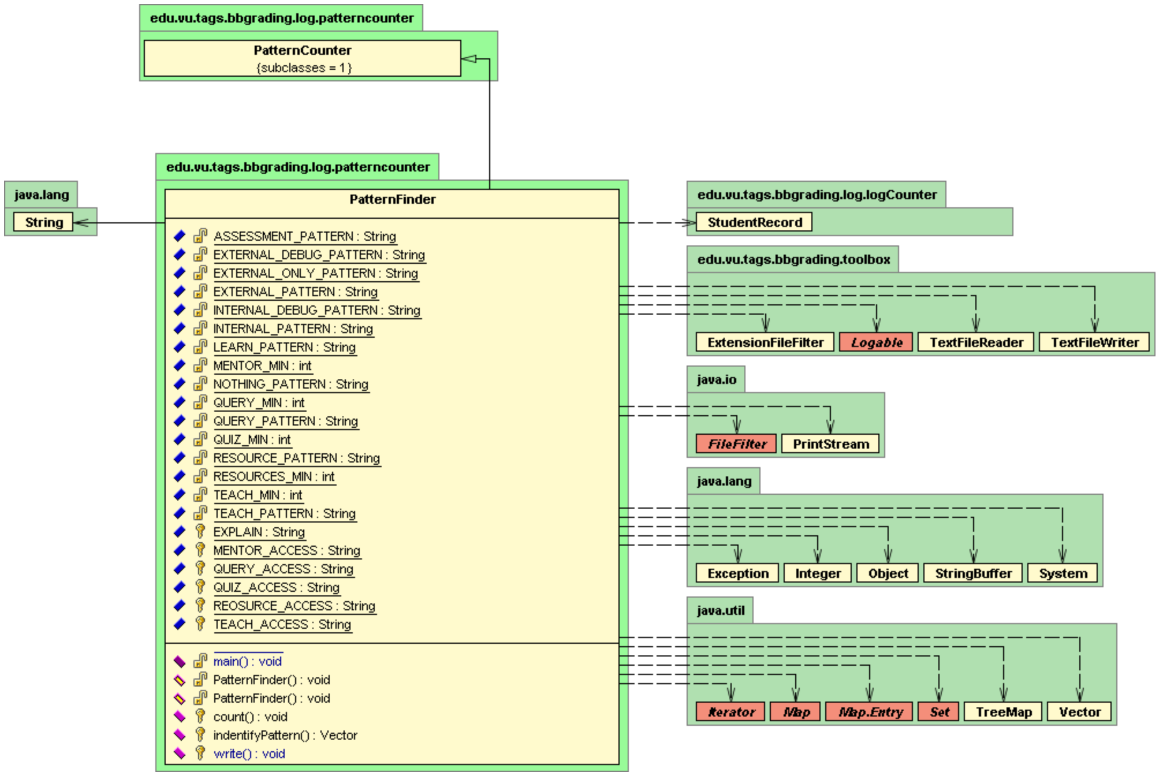


Figure I.13 The `PatternFinder` Class

The details of the `interconnectivity` package and its dependency to other packages are shown in Figure I.14. There are two runnable classes:

- `InterconnectivityGrader` (Figure I.15): Given a set of causal questions, reasons with a given concept map to find the number of links in this map answering each question and the average number of links of all answers
- `QuestionInitiator` (Figure I.16): Generate a causal question from each pair of concepts including the reversed-direction question ($a \longrightarrow b$ and $b \longrightarrow a$)

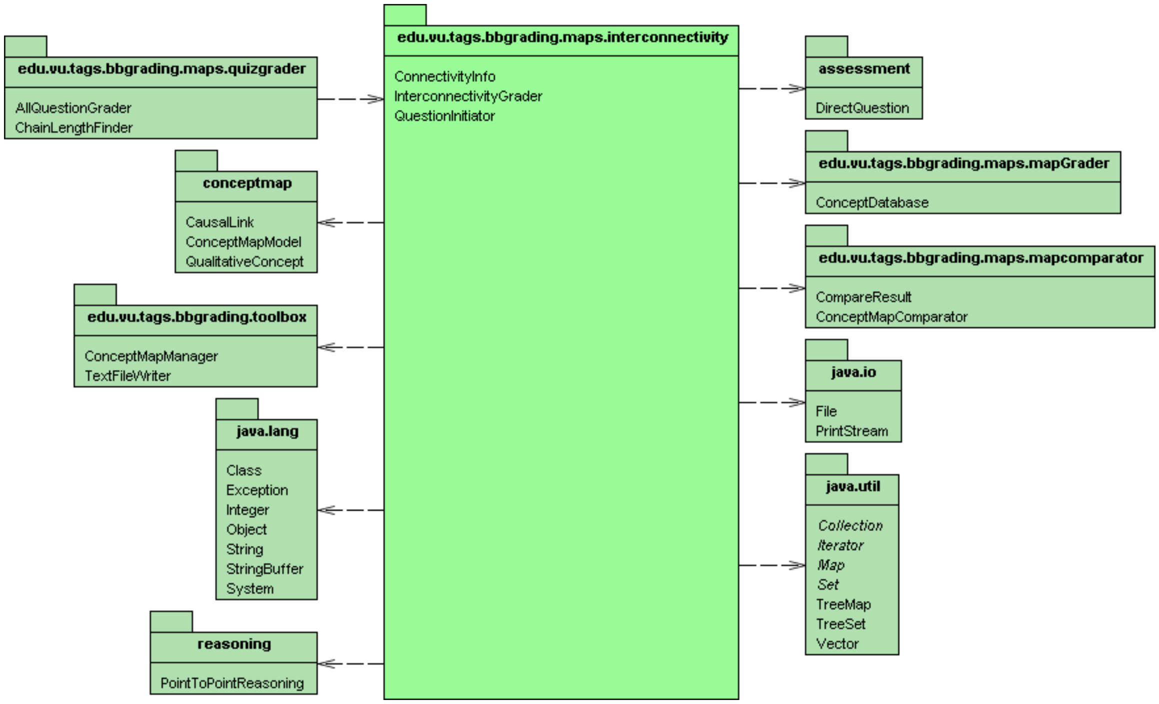


Figure I.14 The `interconnectivity` Package and Its Dependency

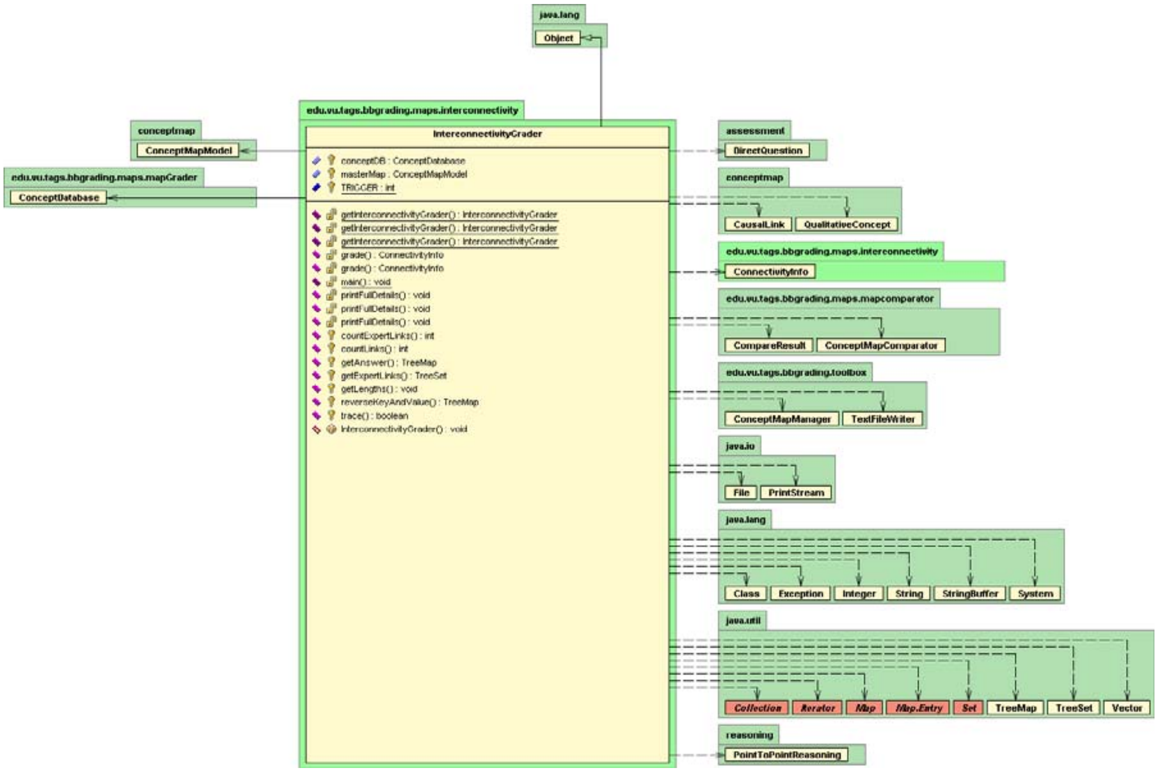


Figure I.15 The `InterconnectivityGrader` Class

The details of the `mapcomparator` package and its dependency to other packages are shown in Figure I.17. There are three main runnable classes, `ConceptMapComparator` (Figure I.18), `GradeChecker` (Figure I.19), and `MemoryQualityGradeFinder` (Figure I.20).

`ConceptMapComparator` compares two concept maps to find what concepts and links they shared by having exact labels and being synonyms and what each map owns exclusively.

`GradeChecker` checks if the comparison grades are consistent with the existence of concepts and links in two concept maps to assure the accuracy of human graders. For example, the grade should be a match code if a concept is present in both map or be a missing code if it appears in only one map and not in another.

`MemoryQualityGradeFinder` is the base class for those that also consider the correctness of the recalled, forgotten, and added concepts and links (see the definitions in Chapter 6) in calculating the score of the memory-test concept-maps.

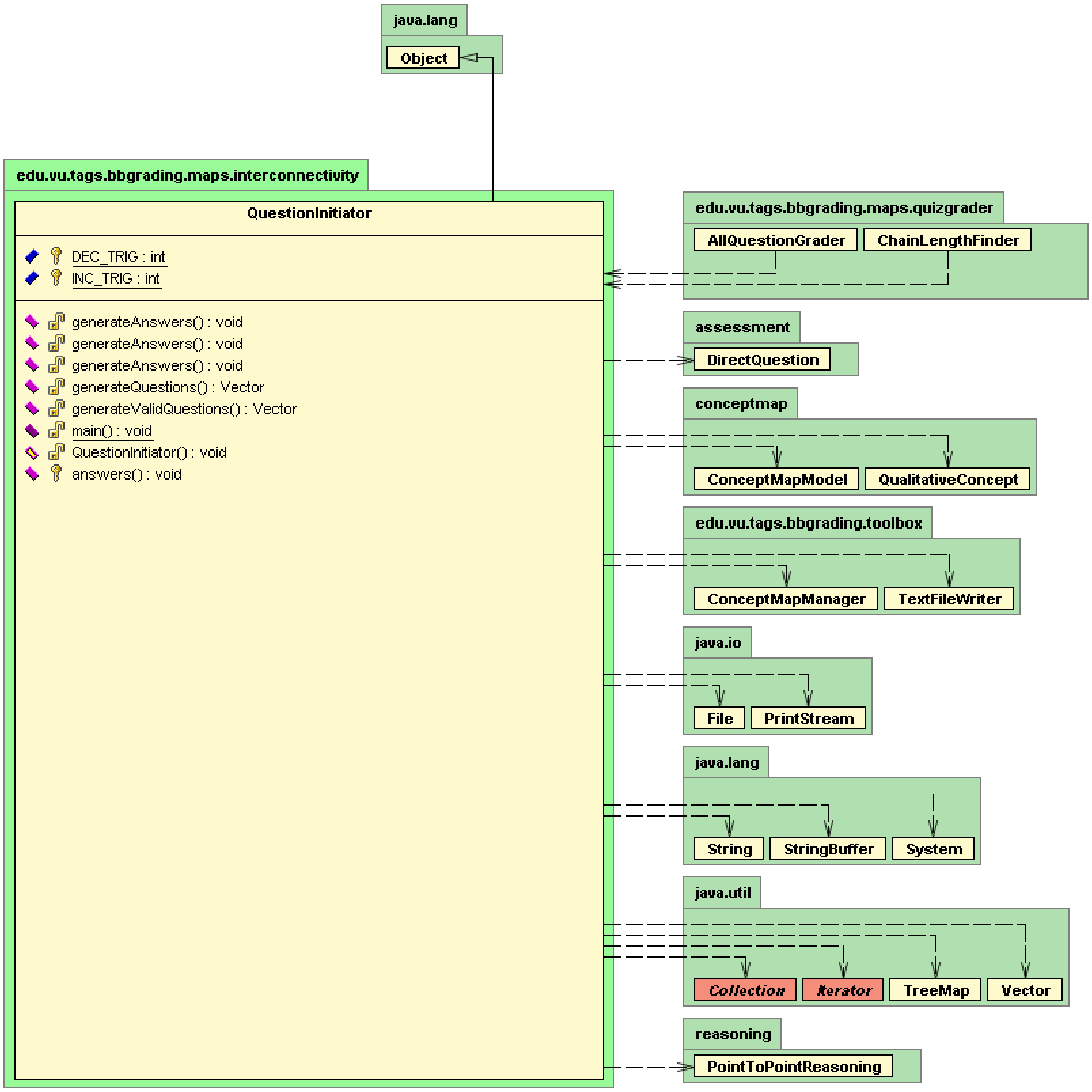


Figure I.16 The `QuestionInitiator` Class

Figure I.17 The `mapcomparator` Package and Its Dependency

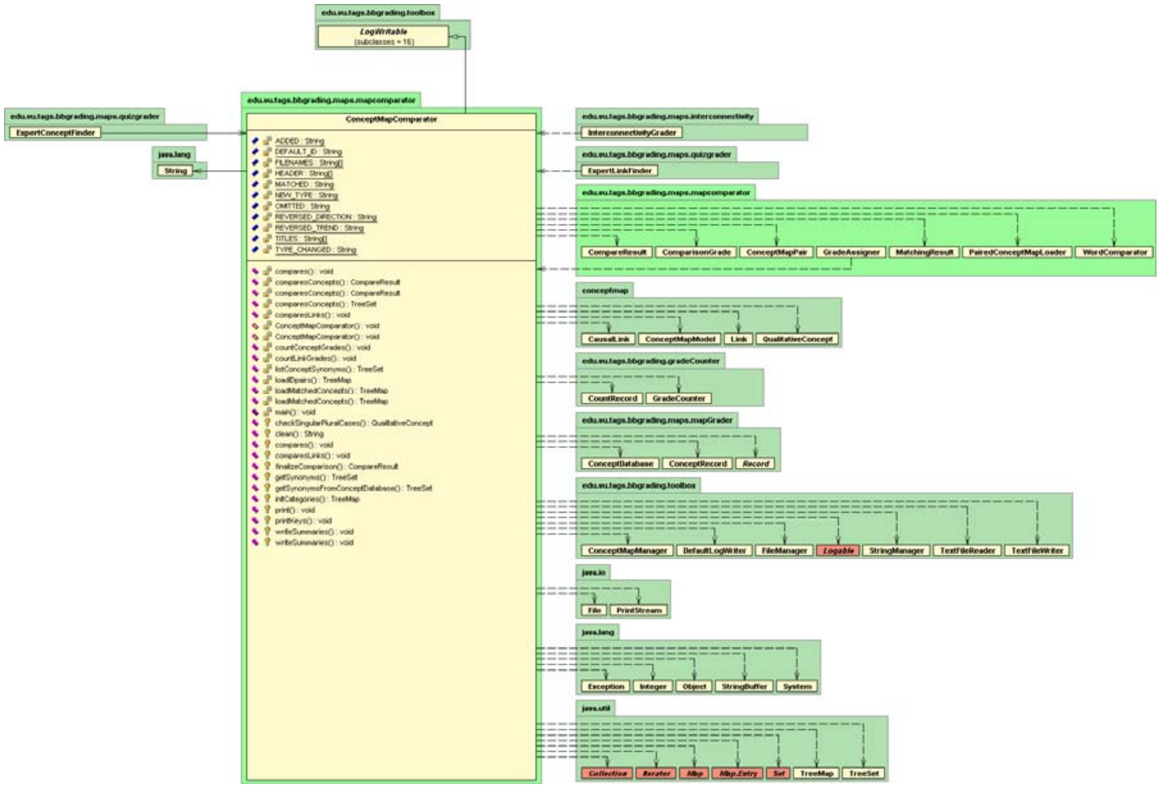


Figure I.18 The `ConceptMapComparator` Class

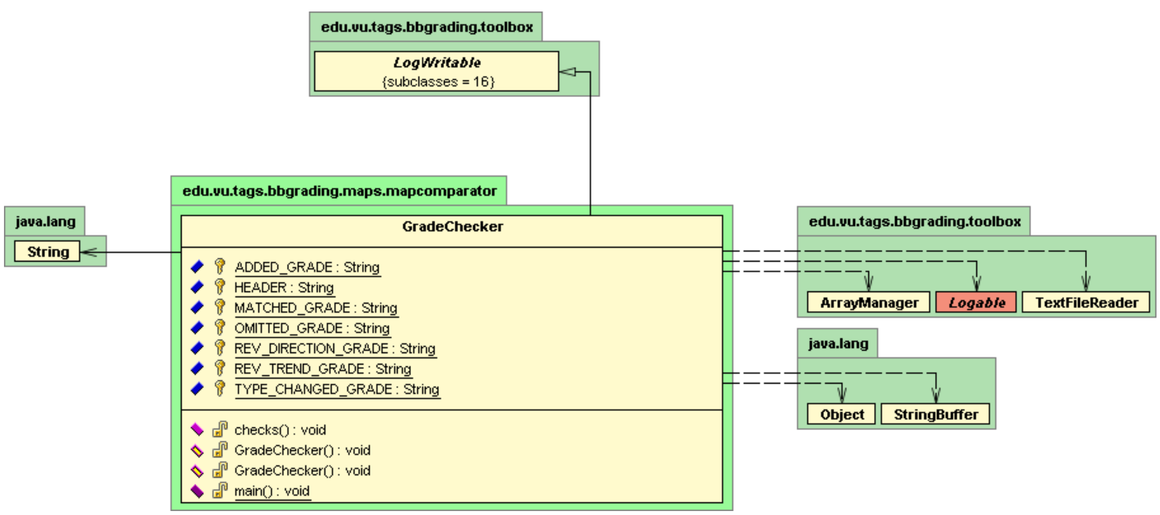


Figure I.19 The `GradeChecker` Class

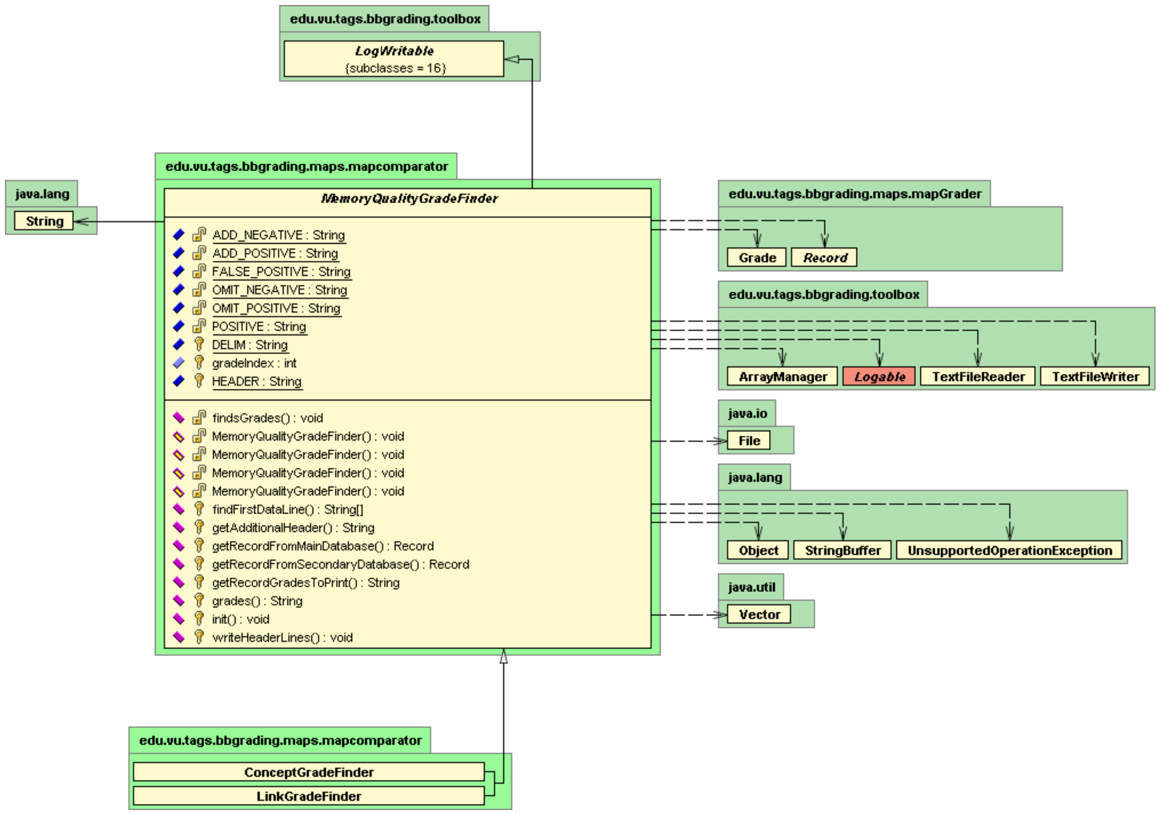


Figure I.20 The `MemoryQualityGradeFinder` Class

The details of the `mapgrader` package and its dependency to other packages are shown in Figure I.21. There are three main, runnable classes, `ConceptMapConverter` (Figure I.22), `Database` (Figure I.23), and `GradeCombiner` (Figure I.24).

`ConceptMapConverter` converts concept maps into text files, one for concepts (shown in Figure I.6) and one for links, for human graders.

`Database` is the base classes for all concept and link graders (see the list at the bottom of Figure I.23). These graders grade concepts and links according to the correctness information stored in their databases. The input-file and database-record format is the same as that produced by the `ConceptMapConverter` class.

`GradeCombiner` compares grades from more than one human grader. This program lists all different grades for each record.

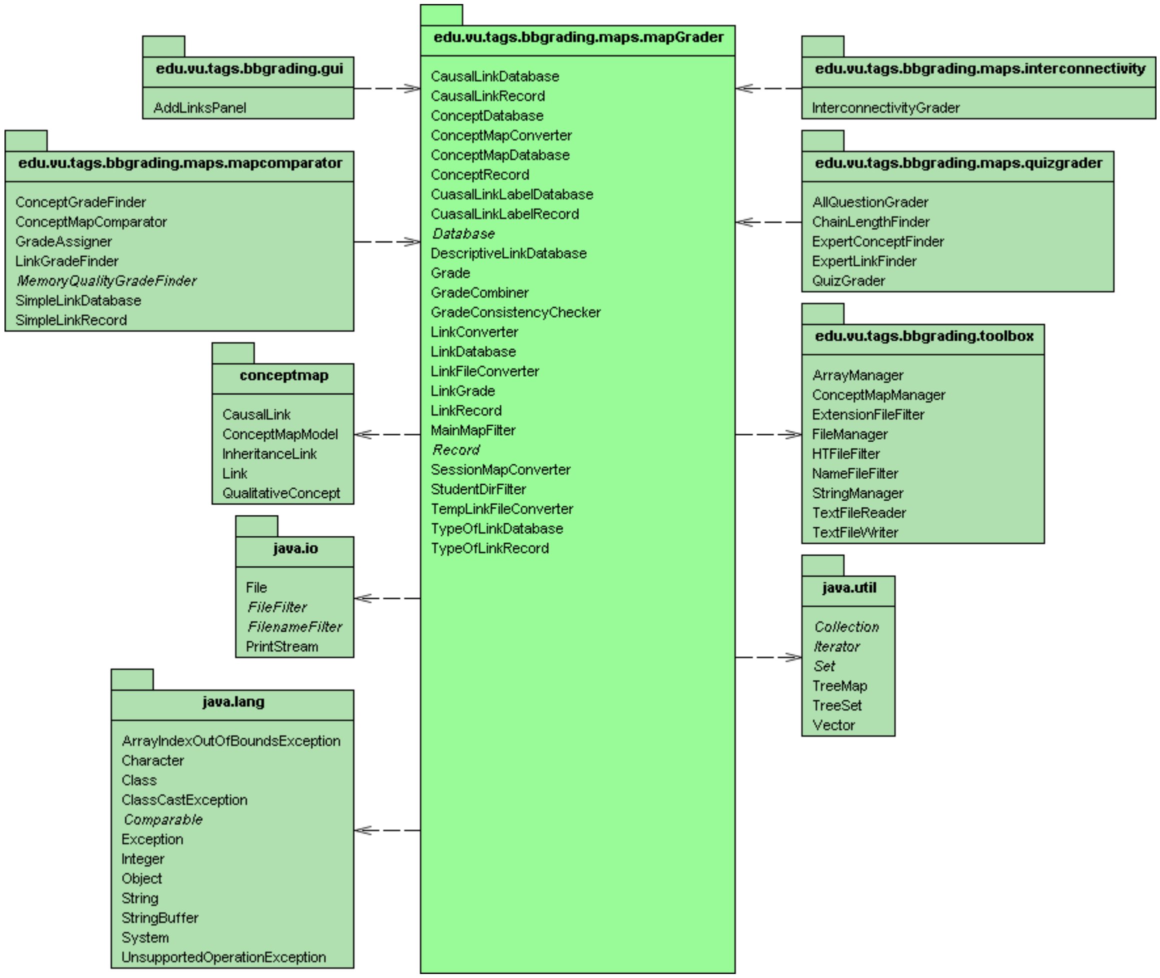


Figure I.21 The `mapgrader` Package and Its Dependency

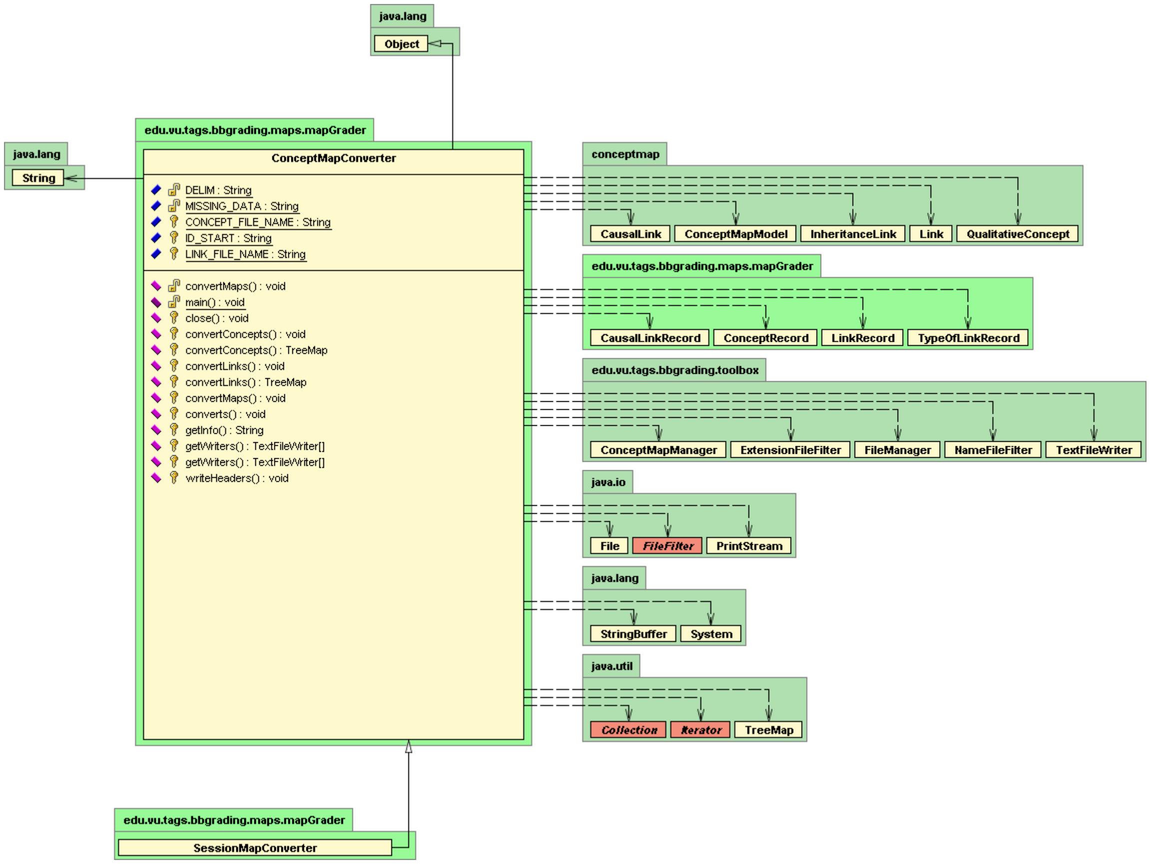


Figure I.22 The `ConceptMapConverter` Class

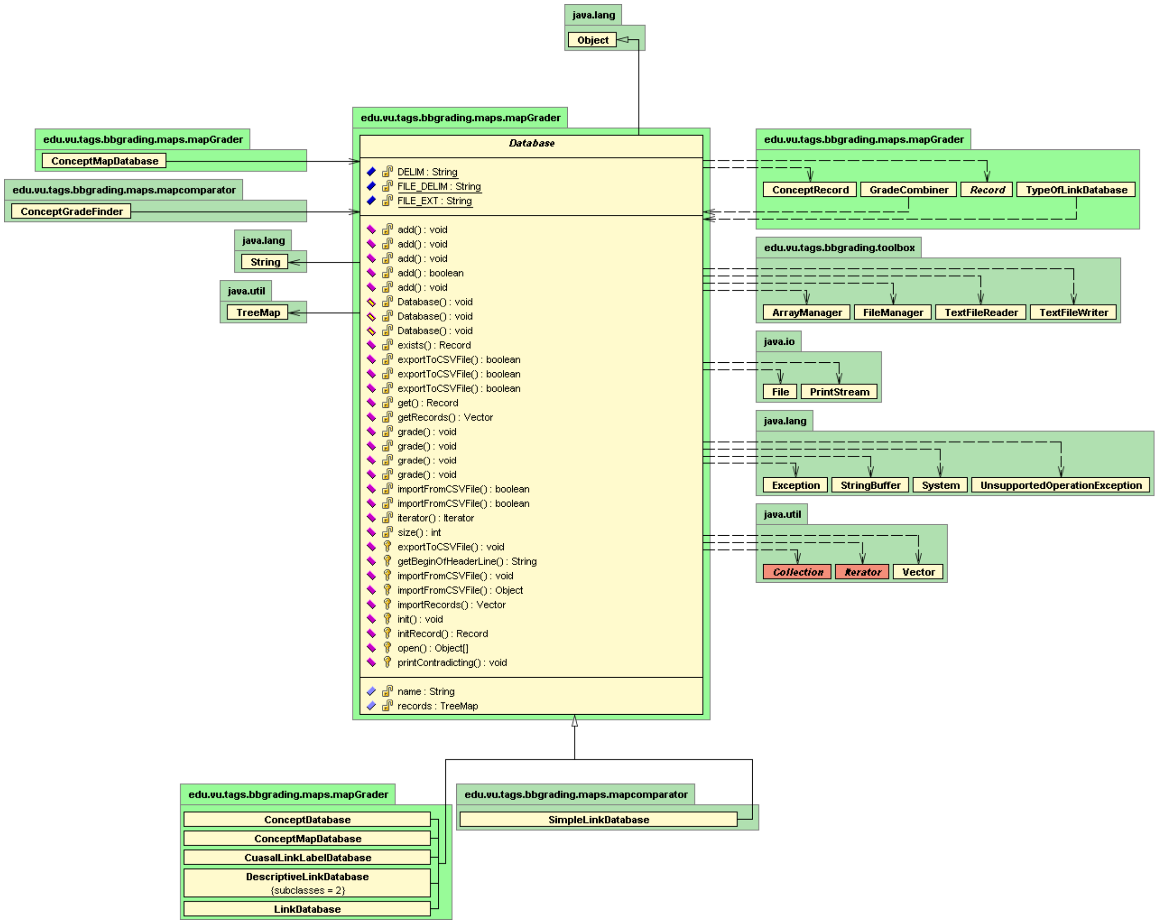


Figure I.23 The `Database` Class

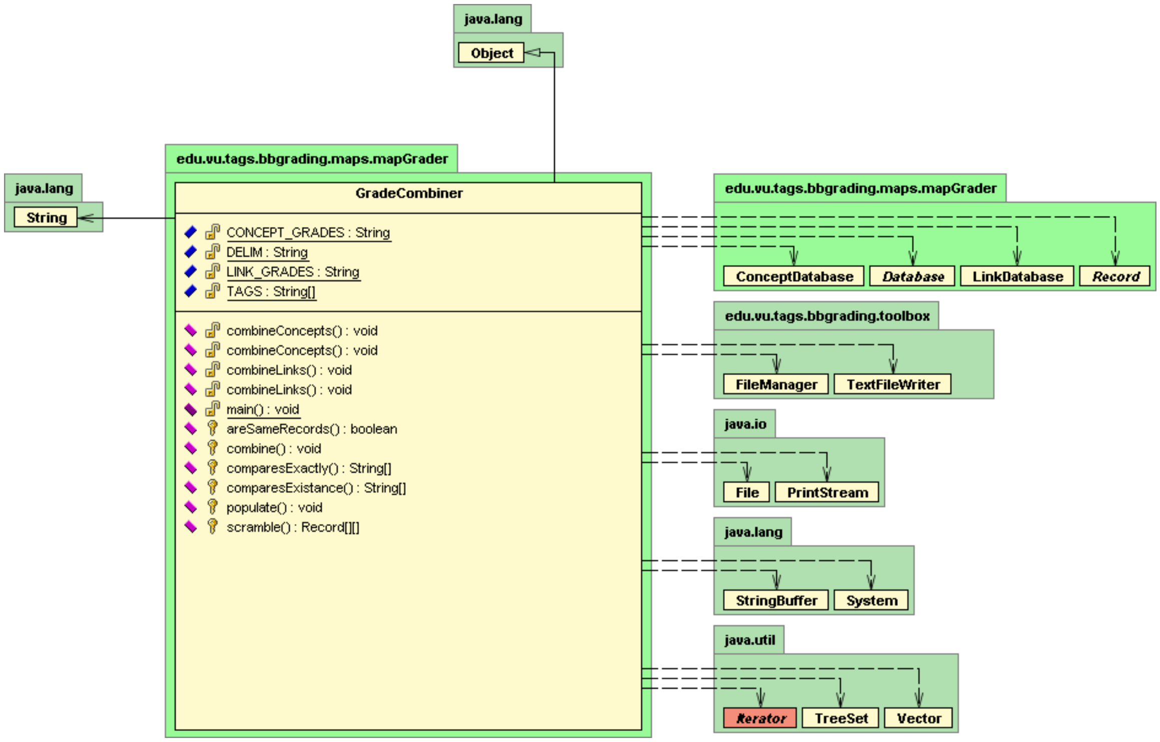


Figure I.24 The `GradeCombiner` Class

The details of the `quizgrader` package and its dependency to other packages are shown in Figure I.25. There is only one runnable class in this package, `QuizGrader` (Figure I.26). This class grades a concept map according to a given set of quizzes of causal questions. The output can be the score of answering each question, the number of correct answers in each quiz, and the number of links in each answer.

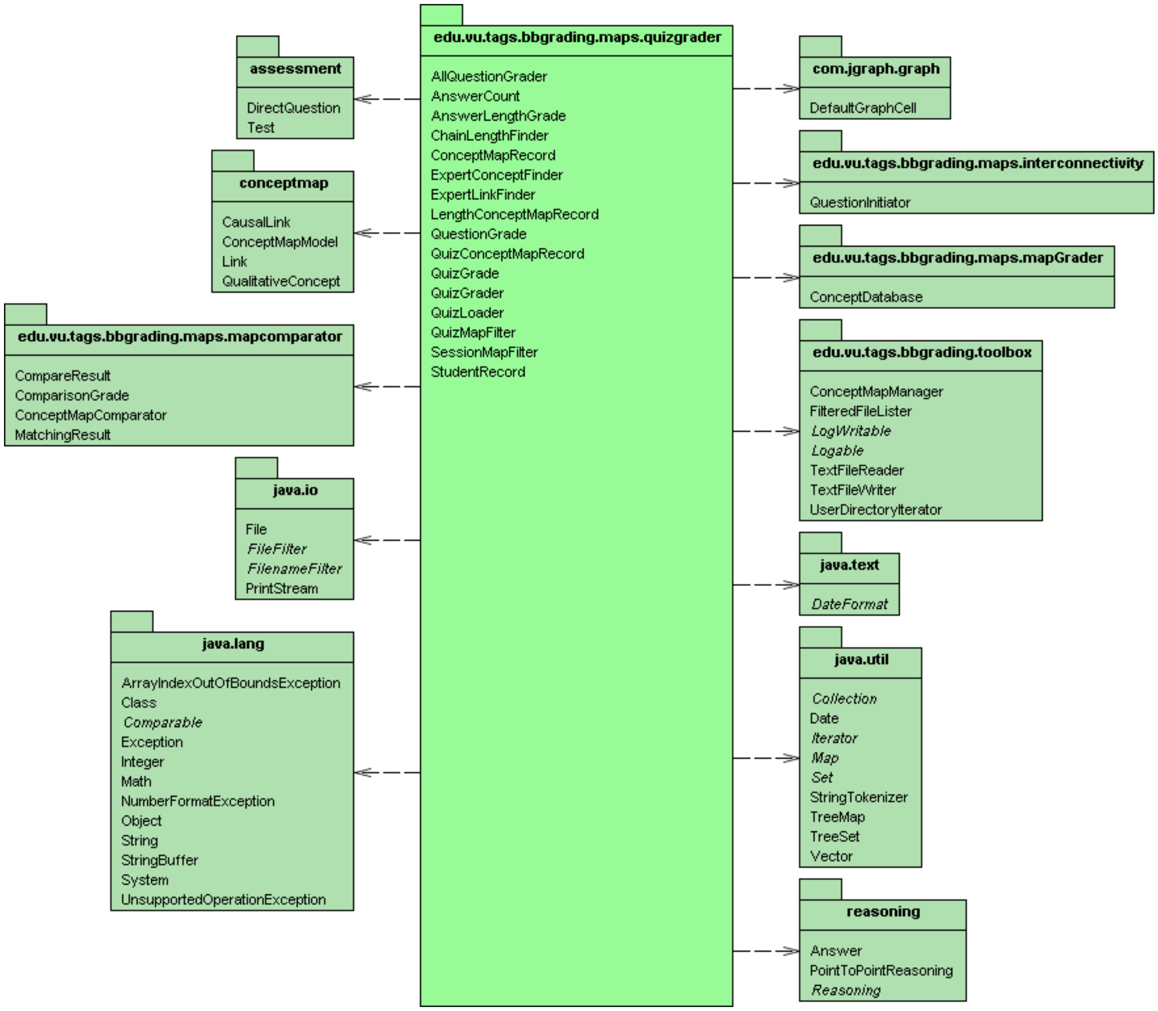


Figure I.25 The `quizgrader` Package and Its Dependency

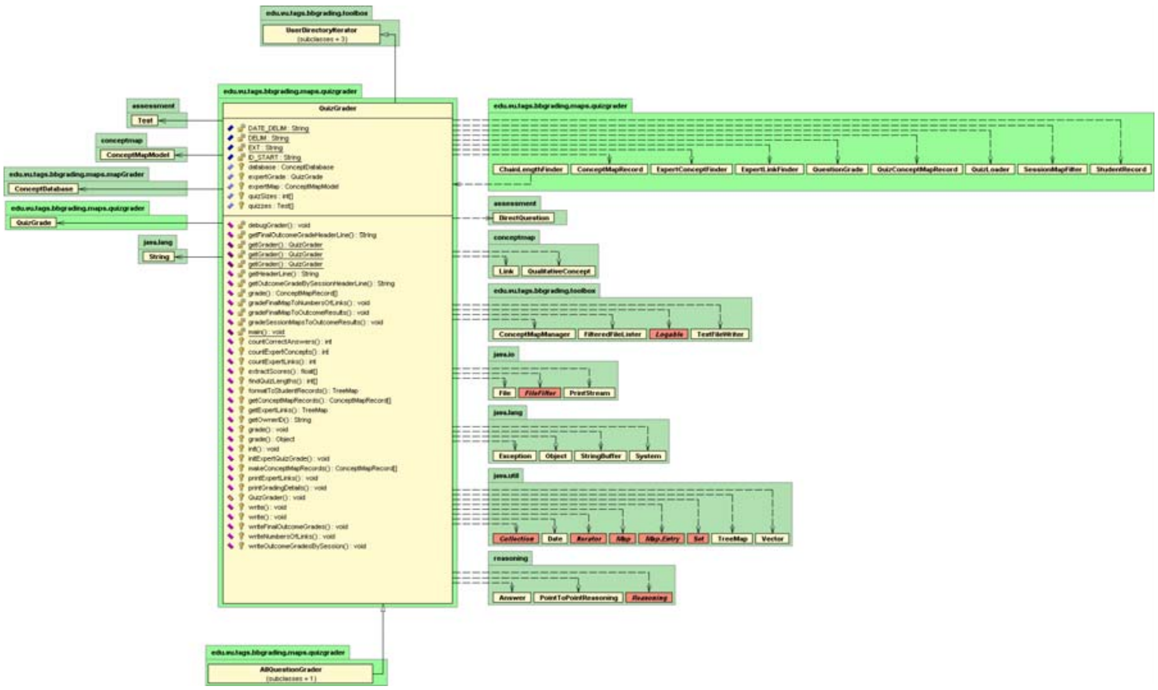


Figure I.26 The `QuizGrader` Class

The details of the `ordering` package and its dependency to other packages are shown in Figure I.27. There are three runnable class in this package, `LongestCommonSub-sequence` (Figure I.28), `OrderingQuestionGrader` (Figure I.29), and `OrderingGrader` (Figure I.30).

To grade an answer to an ordering question, `LongestCommonSubsequence` uses the longest-common subsequence algorithm (Hirschberg 1977) to calculate the similarity of a student's order to the master order.

The `OrderingQuestionGrader` class is a finer-grained grading method for ordering questions. While the `LongestCommonSubsequence` class considers only the positions of each object in the sequence, this method is sensitive to how two consecutive objects in the master order are arranged in the student's order:

1. Event i is immediately followed by Event i + 1 (…, e_i, e_{i+1}, …) = 4 points
2. Event i is followed by Event i + 1 (…, e_i, …, e_{i+1}, …) = 3 points
3. Event i immediately follows Event i + 1 (…, e_{i+1}, e_i, …) = 2 points
4. Event i follows Event i + 1 (…, e_{i+1}, …, e_i, …) = 1 points

"e" represents an event and the subscription "i" represents the correct position of the event in the chain of events, which could be different than its position in the test question. This algorithm considers each two consecutive pairs in the master order and sums the score for the whole sequence. If the student does not specify a position of an event in the pair, it receives zero point. This class can also print the details report for each student.

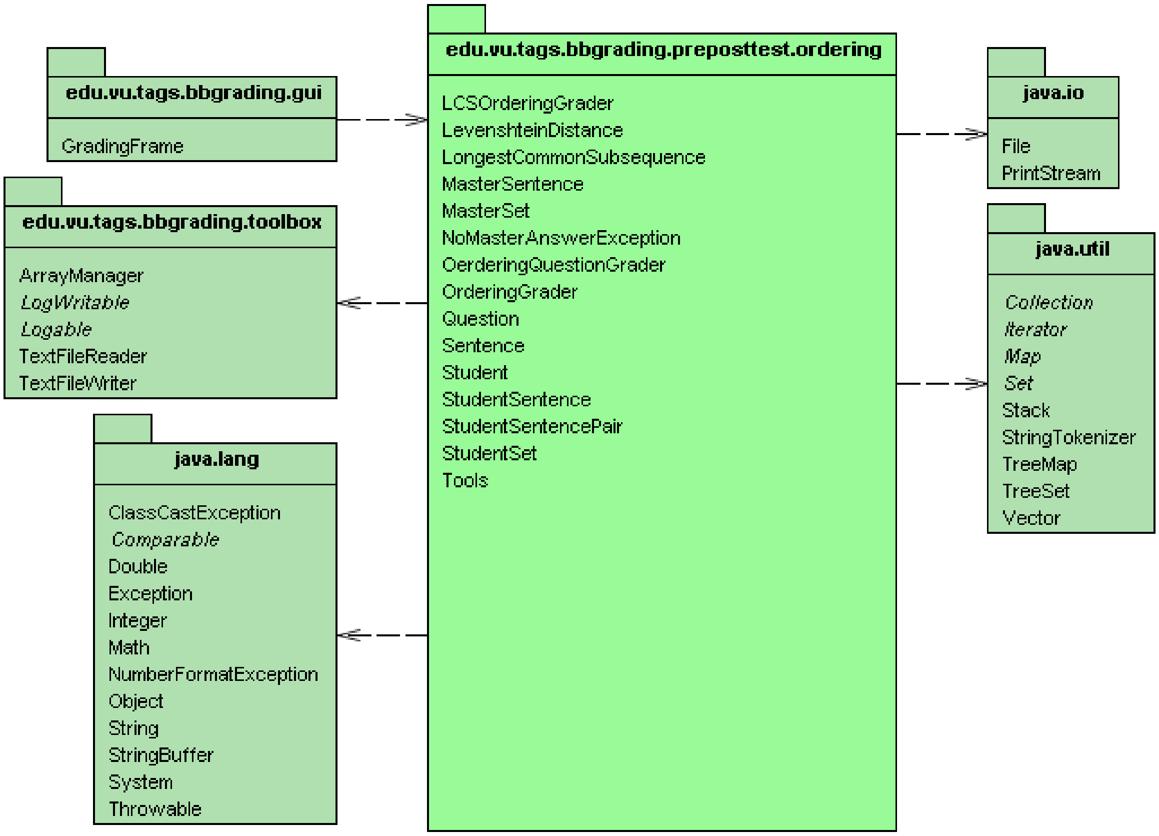


Figure I.27 The `ordering` Package

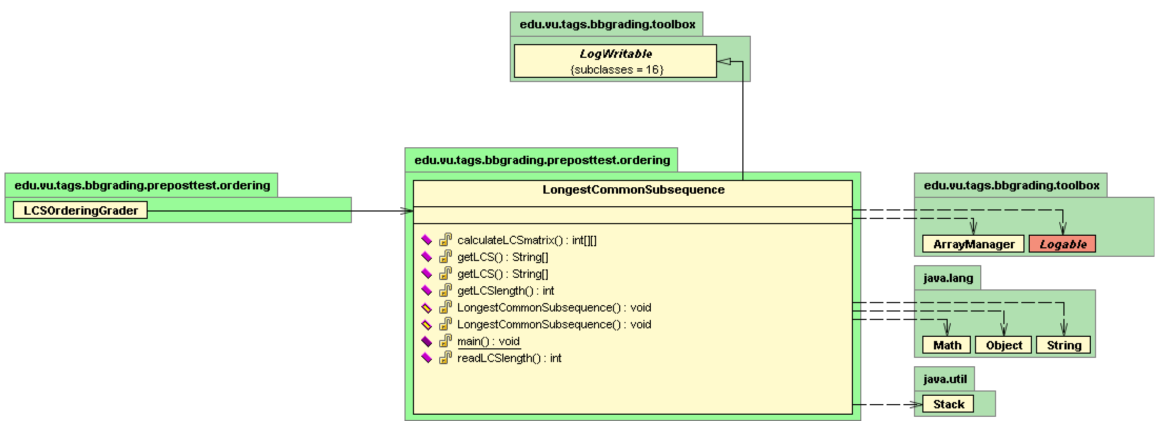


Figure I.28 The `LongestCommonSubsequence` Class

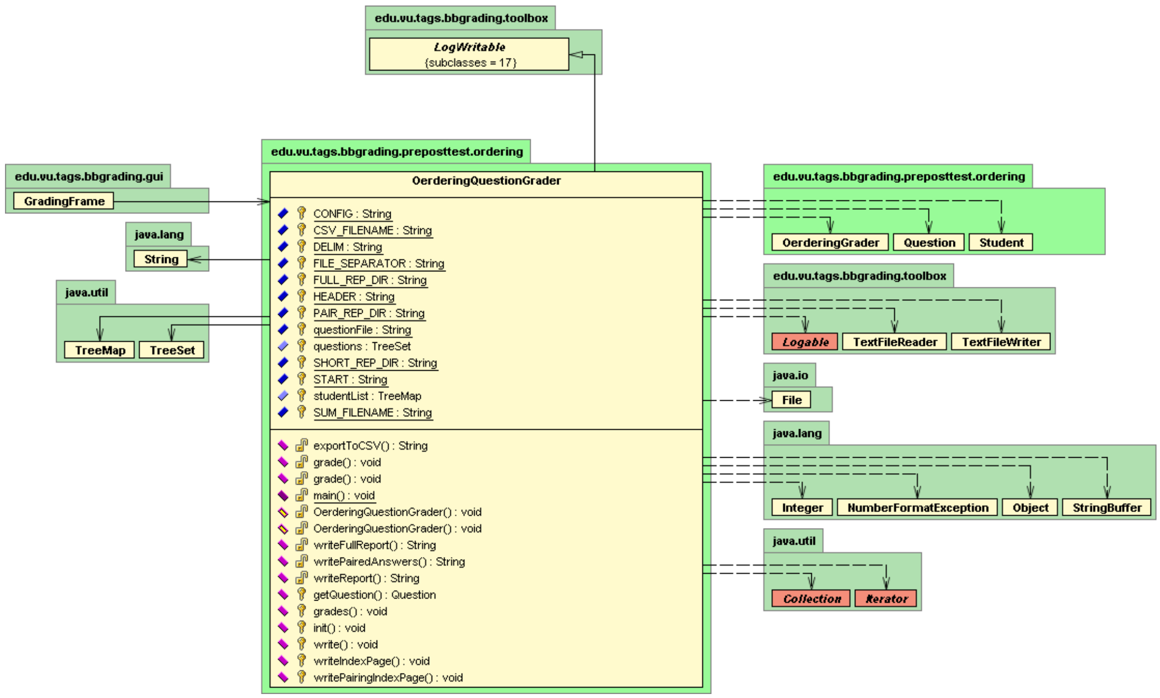


Figure I.29 The `OrderingQuestionGrader` Class

Similar to the `OrderingQuestionGrader` class, the `OrderingGrader` class grades ordering answers in the same fashion but it cannot generate detailed reports. This class is a rewritten version that simplifies the code for ease in calling and debugging it.

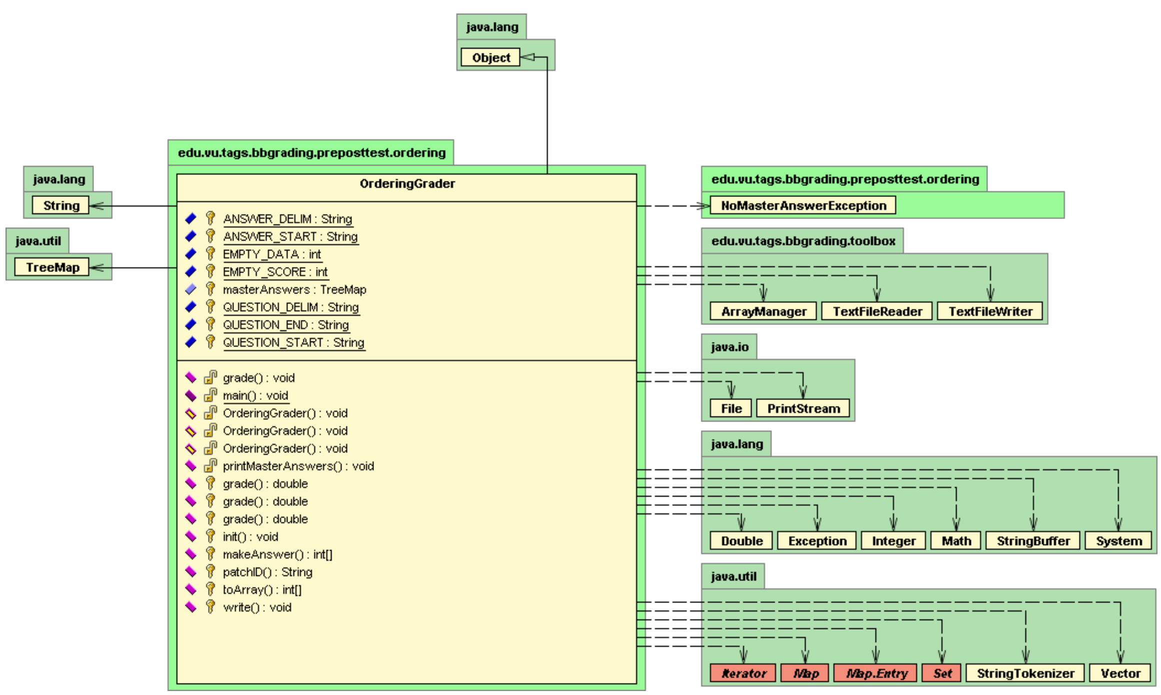


Figure I.30 The `OrderingGrader` Class

The details of the `scoring` package and its dependency to other packages are shown in Figure I.31. There only two runnable classes in this package:

- `MultipleChoiceGrader` (Figure I.32): Gives one point to the same answer in the master list and zero otherwise to the answers to multiple-choice and true-or-false questions
- `OpenEndedQuestionScorer` (Figure I.33): Assign scores to the graded code as discussed in Chapter 6.

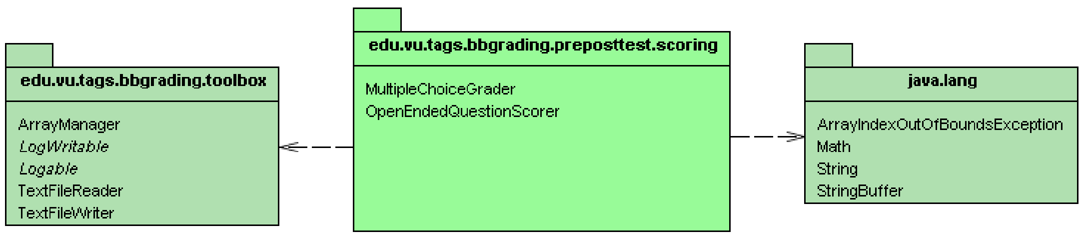


Figure I.31 The `scoring` Package and Its Dependency

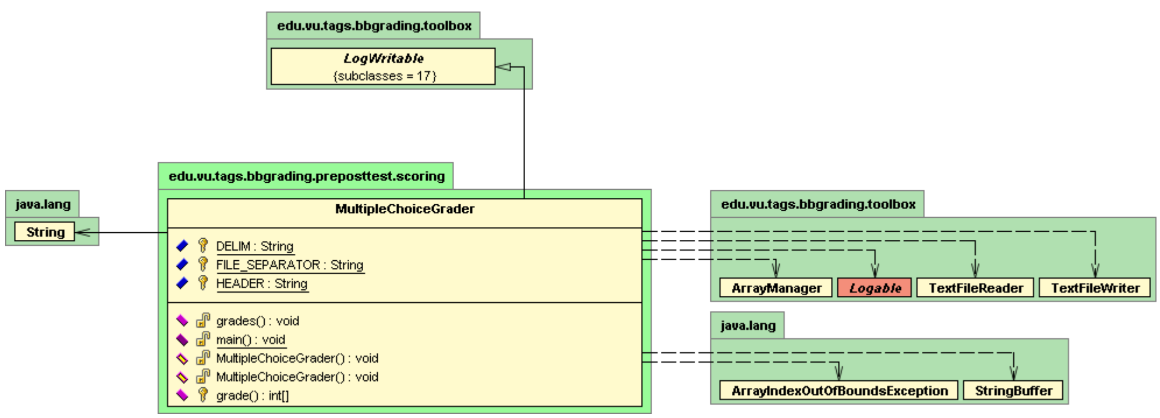


Figure I.32 The `MultipleChoiceGrader` Class

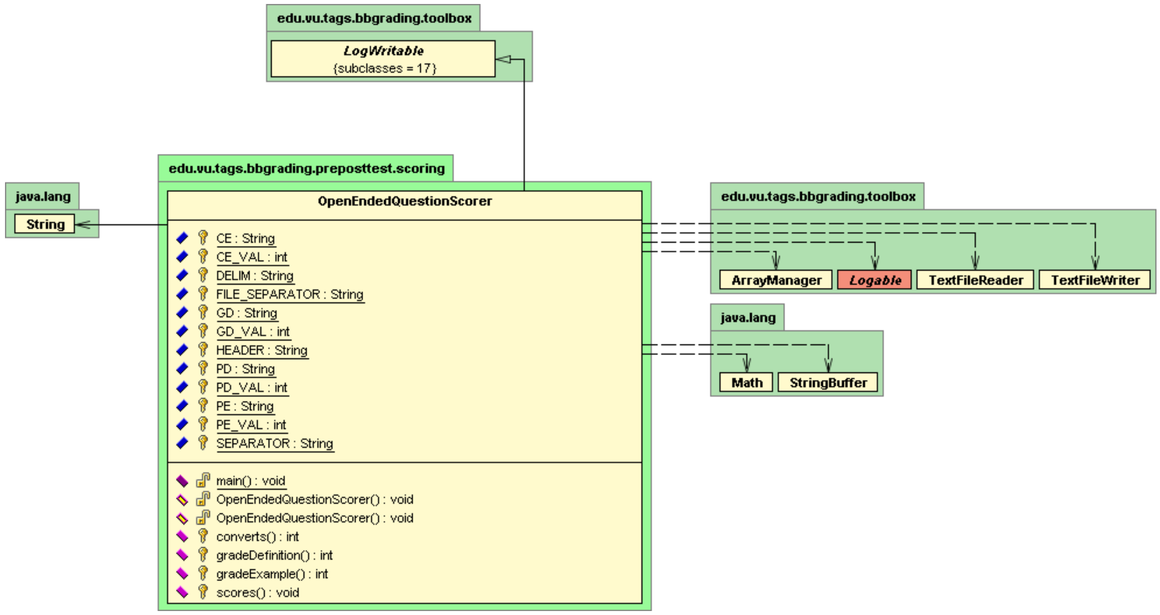


Figure I.33 The `OpenEndedQuestionScorer` Class

The `toolbox` package, shown in Figure I.34, are a collection of code that is shared among packages to reduced the redundancy of code and maintain the modularity of the programs calling these files to focus on their specific tasks. For example, opening a text file to read or write is not a trivial task in Java. Three classes must be initiated and the `IOException` object must be set up to be caught every time a class tries to read from or write to the file stream. In addition, the Betty's Brain project often deals with reading and writing text files. Therefore, the `TextFileReader` and `TextFileWriter` classes manage the initiating the required objects, catch the required exception, and close the stream when the job is done.

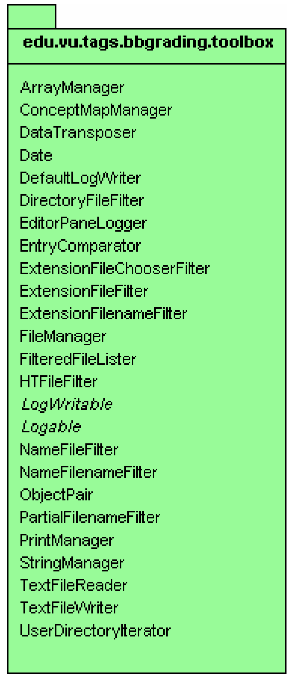


Figure I.34 The `toolbox` Package

APPENDIX J

PRE- AND POSTTEST DETAILED RESULTS AND ANALYSIS

The raw scores of pre- and posttest results are shown in Table J.1. The items highlighted in faded red in Figure J.35 are excluded from the analysis in Chapter 7 because they have either ceiling or floor effects or regressions toward the means. An item has:

- Ceiling effect: Its score is equal to or more than 85% of the full score for this item
- Floor effect: Its score is equal to or less than 15% of the full score for this item
- Regression toward the mean: The posttest score of the group is lower than the pretest score, or the posttest score of the group is the same as the pretest while the posttest scores of the other groups increase from the corresponding pretest score

Table J.2 Pre- and Posttest Scores

Question	SRL				LBT				ITS			
	Full Score	15% 85%	Pre-test Score	Post-test Score	Full Score	15% 85%	Pre-test Score	Post-test Score	Full Score	15% 85%	Pre-test Score	Post-test Score
1	4	0.6 3.4	1.23	2.15	4	0.6 3.4	0.94	2.25	4	0.6 3.4	1.07	1.80
2			0.92	1.31			0.69	0.63			1.33	1.60
3			2.23	2.49			1.63	1.56			2.53	2.20
4	13	1.95	6	10	16	2.4	9	13	15	2.25	12	12

5	a		11.05	12	9		13.6	14	15		12.75	14	12
	b			14	13			16	16			14	13
	c			14	13			16	15			15	15
	d			9	12			7	12			9	13
	e			13	13			13	15			14	14
	f			13	14			16	16			15	15
	g			12	13			12	16			13	15
	h			12	11			16	13			12	14
	i			12	11			14	11			14	13
	j			12	12			15	13			15	14
	k			1	2			1	3			2	4
	l			11	13			11	14			14	11
	m			11	13			10	16			12	15
	n			10	12			15	16			13	13
	o			9	11			12	16			11	13
	p			7	4			7	6			9	7
6		8	1.2 6.8	4.77	5.00	8	1.2 6.8	4.56	5.00	8	1.2 6.8	4.53	4.73
7		5	0.75 4.25	4.85	4.46	5	0.75 4.25	4.56	4.81	5	0.75 4.25	4.33	4.07
8		6	0.9 5.1	4.85	4.38	6	0.9 5.1	5.31	4.06	6	0.9 5.1	4.80	4.33
9		8	1.2 6.8	6.85	6.31	8	1.2 6.8	5.81	6.13	8	1.2 6.8	5.80	5.93
10		6	0.9 5.1	4.62	4.62	6	0.9 5.1	5.00	4.75	6	0.9 5.1	4.87	4.73
11		10	1.5 8.5	6.00	6.08	10	1.5 8.5	6.44	6.13	10	1.5 8.5	6.40	6.13
12		8	1.2 6.8	5.31	5.92	8	1.2 6.8	6.00	4.94	8	1.2 6.8	5.53	5.67

Question		Pretest		Posttest		Regression toward the Mean
		Floor Effect	Ceiling Effect	Floor Effect	Ceiling Effect	
1						
2						
3						
4						ITS
5	a		SRL, LBT, ITS		LBT	
	b		SRL, LBT, ITS		SRL, LBT, ITS	
	c		SRL, LBT, ITS		SRL, LBT, ITS	

	d				SRL, LBT, ITS	
	e		SRL, LBT, ITS		SRL, LBT, ITS	
	f		SRL, LBT, ITS		SRL, LBT, ITS	
	g		SRL, LBT, ITS		SRL, LBT, ITS	
	h		SRL, LBT			
	i		SRL, ITS		ITS	
	j		SRL, LBT, ITS		SRL, ITS	
	k	SRL, LBT, ITS		LBT		
	l		ITS		SRL, LBT	
	m				SRL, LBT, ITS	
	n		LBT, ITS		SRL, LBT, ITS	
	o				LBT, ITS	
	p					
6						
7			SRL, LBT, ITS		SRL, LBT, ITS	
8			SRL, LBT		ITS	
9			SRL			
10			LBT, ITS			
11						
12						LBT

Figure J.35 Ceiling and Floor Effects in Pre- and Posttest Scores

Table J.3 Normality and Variance Equality Tests for the Pretest and Posttest Data

Question		Normality Test: Shapiro-Wilk			Variance Equality: Levene's Test
		ITS	LBT	SRL	
Pretest	1	0.00	0.00	0.00	.57
	2	0.00	0.00	0.00	.32
	3	0.03	0.00	0.02	.20
	5	0.01	0.26	0.39	.78
	6	0.01	0.00	0.00	.68
	11	0.05	0.04	0.15	.98
Post-test	1	0.00	0.00	0.00	.32
	2	0.01	0.00	0.02	.26

3	0.01	0.01	0.06	.90
5	0.00	0.00	0.02	.04
6	0.08	0.23	0.12	.79
11	0.62	0.26	0.26	.79

Table J.4 Mann-Whitney Tests on the Pretest Results

Question	ITS & LBT		ITS & SRL		LBT & SRL	
	U	Significance	U	Significance	U	Significance
1	117.5	p = .92	91.5	p = .79	94.5	p = . 68
2	93.5	p = .30	85.5	p = .59	92.0	p = .62
3	82.0	p = .14	90.5	p = .75	83.5	p = .37
5	88.5	p = .22	81.0	p = .47	92.5	p = .62
6	118.5	p = .95	83.0	p = .53	83.5	p = .37
11	118.5	p = .95	90.0	p = .75	95.0	p = .71

APPENDIX K

MISCELLENEOUS MAIN-STUDY RESULTS AND ANALYSIS

Table K.5 indicates that we can use ANOVA only with the numbers of expert concepts because the data is normal and the variance are not significantly different (the first row, all *p*'s > 0.05). The other variables cannot be tested with ANOVA because either data for some groups were not normal ($p \leq .05$). As stated in Chapter 6, ANOVA is a stronger test than the Mann-Whitney U test. Therefore, we have tried to use ANOVA wherever it is applicable.

Table K.5 Normality and Variance Equality Tests for the Concept-Map Data from Session 5

Variable	Normality Test: Shapiro-Wilk			Variance Equality: Levene's Test
	SRL	LBT	ITS	
Expert concepts	.07	.16	.09	.52
Valid concepts	.52	.01	.17	.04
Invalid concepts	.01	.07	.003	.11
Expert links	.87	.56	.002	.09
Valid links	.94	.64	.91	.04
Invalid links	.002	.000	.29	.02

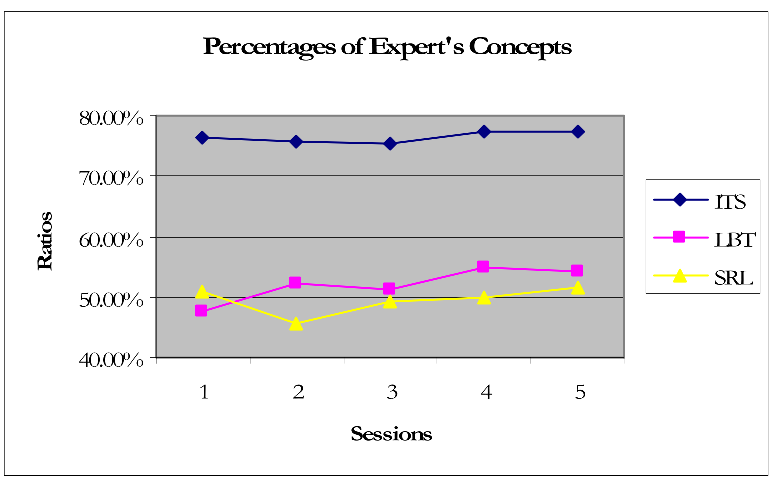


Figure K.1 Ratio of the Number of expert Concepts to the Total Number of Concepts in Students' Concept Maps at the End of Each Session of the Main Study

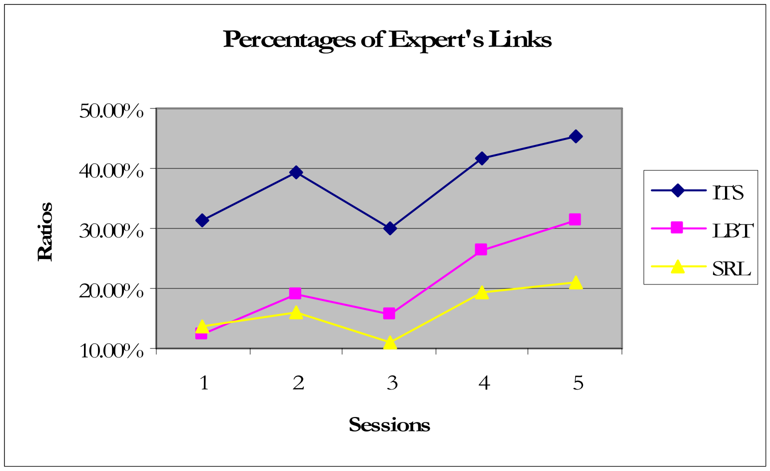


Figure K.2 Ratio of the Number of expert Links to the Total Number of Links in Students' Concept Maps at the End of Each Session of the Main Study

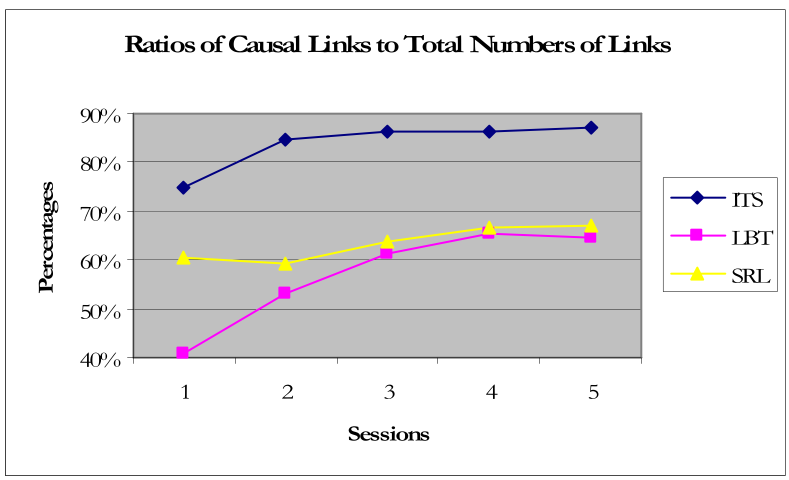


Figure K.3 Ratio of the Number of Causal Links to the Total Number of Links in the Participants' Concept Maps at the End of Each Session of the Main Study

Table K.6 Repeated-Measures Analysis of Variance on the Ratio of Causal Links in the Participants' Concept Maps at the End of Each Session of the Main Study

	Ratios of Causal Links
Time	$F_{(2, 80)} = 9.6, p < .001$
*Time * Group*	$F_{(3, 80)} = 1.6, p > .05$

ITS & LBT	Tukey: $p \leq .05$
ITS & SRL	Tukey: $p > .05$
SRL & LBT	Tukey: $p > .05$

Awareness of Interdependence

Causal links are important to students' concept maps because they represent the dynamic nature of natural domains, specifically the river ecosystem. Figure K.4 depicts the average ratios of numbers of valid, causal links to the total number of links in the students' concept maps at the end of each session of the main study. A repeated-measures analysis of variance was conducted on the data, and the results appeared in Table K.7. The main effect of time of measurement was significant, and also the interaction effect of time and group. The whole class of participants had had more valid, causal links as they progressed, but the ITS group had higher ratio.

We conjectured that the feedback in terms of global chains of events that only the SRL group received would increase the numbers of causal links in their concept maps. The other two groups, ITS and LBT, receiving concept-by-concept and link-by-link feedback may not be able to connect these pieces to create chains of actions. However, this was no significant difference between the SRL group and the others. A possible explanation was that the new quiz questions were scaffolded to focus first on a part of expert concept map and expand the coverage area to cover the whole map in the last quiz. Therefore, the students in the ITS and LBT groups who were heavily depended on the quiz feature added causal links according to the mentor's hints. Even though the feedback was still in the piece-by-piece fashion, these pieces now were also given in a way that created chains of events. This positive effect of the quiz structure as a scaffold made this measure obsolete.

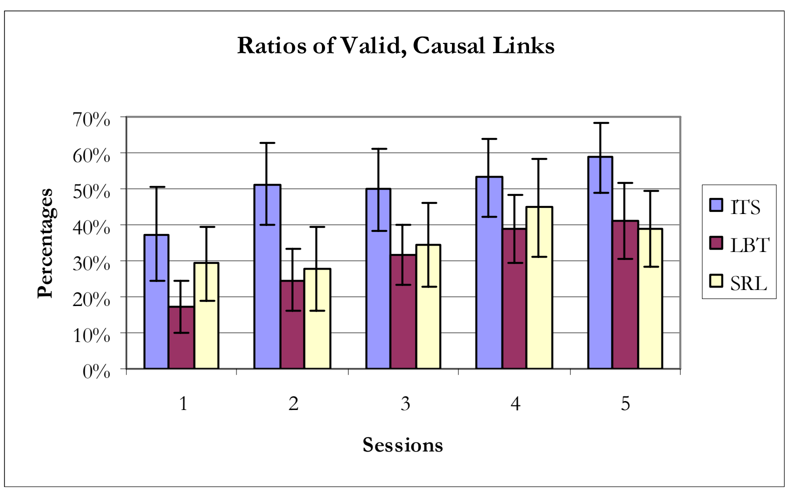


Figure K.4 Ratio of Valid, Causal Links to the Total Number of Links in the Participants' Concept Maps at the End of Each Session of the Main Study (Error Bars Represent the 95% Confidence Intervals of the Differences between Means)

Table K.7 GLM MANOVA on the Ratio of Valid, Causal Links in the Participants' Concept Maps at the End of Each Session of the Main Study

	Ratios of Valid Causal Links
Time	$F_{(4, 38)} = 7.7, p < .0005$
*Time * Group*	$F_{(8, 76)} = 1.9, p = .08$
ITS & LBT	Tukey: $p \leq .05$
ITS & SRL	Tukey: $p \leq .05$
SRL & LBT	Tukey: $p > .05$

Awareness of the Dynamic Nature of the River Ecosystems

The numbers of causal links per answer indicated the degree of chaining in the concept maps. The question set for the results shown in Figure K.5 was generated by a permutation of all concepts in expert concept map, and the set for the results shown in Figure K.6 a permutation of all concepts in the student's own concept map at the end of each session of the main study. Both set of questions were applied to the concept map at the end of the main study. There was significant different in time and there was no significant result between groups as shown in Table K.8. The significant effect of time was as expected. The participants had had more connected concept maps when they spent more time using the Betty's Brain environment. The slight drop in the fifth session of the SRL group did not disturb the direct dependency of the degree of interconnectivity and the number of sessions.

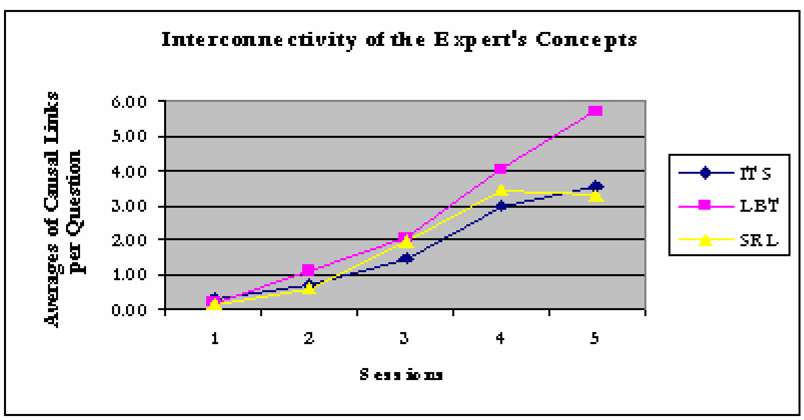


Figure K.5 Average Number of Causal Links per Question Permutated from expert Concepts

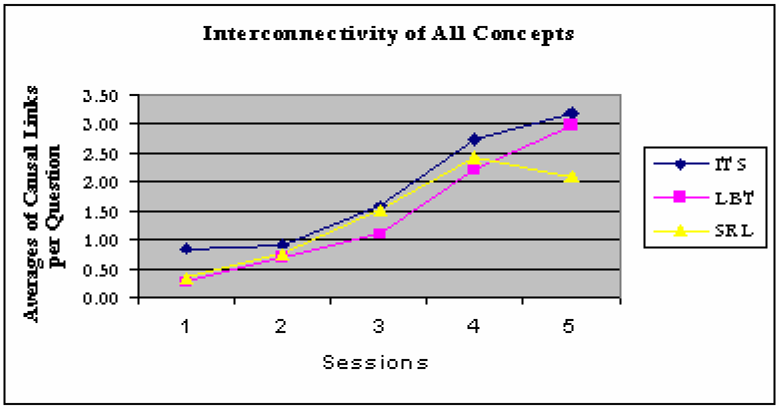


Figure K.6 Average Number of Causal Links per Question Permutated from Concepts in Students' Concepts Maps

Table K.8 GLM Analysis of Variance on the Ratio of Expert Objects in Students' Concept Maps to the Total Number of the Objects in the Maps

	Numbers of Links per expert Question	*Numbers of Links per the student's Question*
Time	$F_{(2, 68)} = 27.5, p < .001$	$F_{(2, 68)} = 18.4, p < .001$
*Time * Group*	$F_{(4, 68)} = 1.0, p > .05$	$F_{(4, 68)} = 0.4, p > .05$
ITS & LBT	Tukey: $p > .05$	Tukey: $p > .05$
ITS & SRL	Tukey: $p > .05$	Tukey: $p > .05$
SRL & LBT	Tukey: $p > .05$	Tukey: $p > .05$

We conjectured that the feedback in terms of global chains of events that only the SRL group received would increase the degree of interconnectivity. However, the non-significant results were not surprising because of the same reason as that of the awareness-of-independence measure

APPENDIX L

MISCELLENEOUS MEMORY-TEST AND TRANSFER-TEST RESULTS AND ANALYSIS

Correctness of the Memory-Test Concept Maps

Figure L.1 displays the average performance by group when their maps are compared against the expert map. The number of expert concepts and links, valid concepts and links, and incorrect concepts and links are plotted in the figure. The ITS group had less valid concepts than both the LBT and the SRL groups but other results were not significant between groups as indicated by results from Mann-Whitney U tests shown in Table L.1.

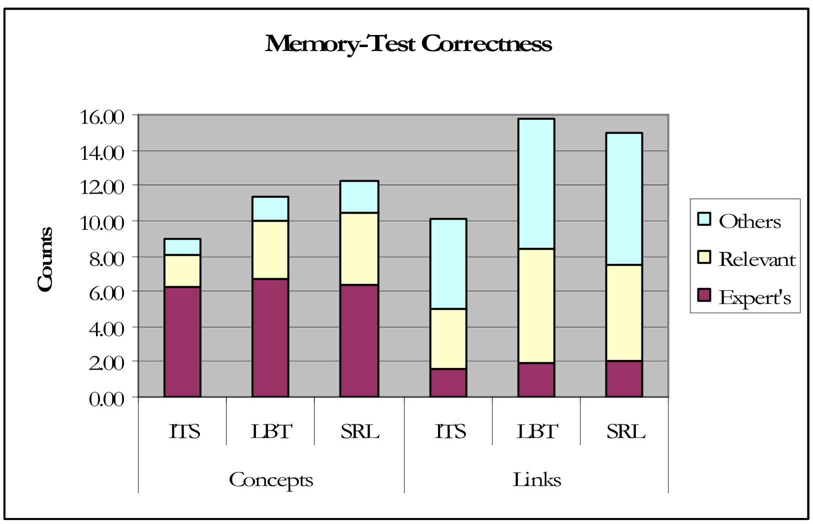


Figure L.1 Grade Distributions of the Memory-Test Concept-Maps

Table L.1 Significance Levels of the Correctness of Memory-Concept-Map from Mann-Whitney U Tests

	Expert Concepts	*Expert Links*	*Valid Concepts*	*Valid Links*
ITS & LBT	U = 100 $p = .45$	U = 97 $p = .38$	U = 63.5 $p < .05$	U = 71 $p = .054$
ITS & SRL	U = 97 $p = 1.00$	U = 81.5 $p = .47$	U = 53 $p < .05$	U = 67 $p = .17$

SRL & LBT	U = 91 p = .59	U = 103.5 p = .98	U = 96.5 p = .75	U = 102 p = .95

Normality and Variance Equality Tests on Memory-Test Results

The following variables can be tested by ANOVA because their group data are normal and the variances are not statistically different among three groups, Numbers of Recalled Concepts, Ratios of Recalled Concepts, Ratios of Recalled Links, and Concept-Accuracy Scores. The rest of the variables is tested with the Mann-Whitney U Tests.

Table L.2 Normality and Variance Equality Tests for the Memory-Test Data

Variable		*Normality Test: Shapiro-Wilk*			*Variance Equality: Levene's Test*
		ITS	*LBT*	*SRL*	
Recollection	Numbers of Recalled Concepts	.60	.25	.46	.23
	Numbers of Recalled Links	.25	.11	.39	.04
	Ratios of Recalled Concepts	.70	.51	.34	.42
	Ratios of Recalled Links	.13	.95	.20	.47
Accuracy	Concept-Accuracy Scores	.27	.50	.71	.53
	Link-Accuracy Scores	.01	.07	.25	.04
	Concept-Accuracy Ratios	.69	.82	.000	.51
	Link-Accuracy Ratios	.45	.000	.75	.23
Maturation		.20	1.00	.60	.04

Normality and Variance Equality Tests on Transfer-Test Results

As shown in Table L.3, the standard deviations of the numbers of expert links for all three groups and the number of valid links are greater than their means. This means that the data are extremely skewed. The examination of data has revealed that there are many zero data points in these sets (between 67% - 81%). Therefore, we have not performed data analysis on these two variables. For the rest of the variables, some data sets are not normal, as shown in Table L.4. Therefore, we have conducted Mann-Whitney U tests on them.

Table L.3 Descriptive Statistic Results of the Transfer-Test Data

Variables	*ITS*		*LBT*		*SRL*	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
Number of expert concepts	3.47	1.60	4.56	1.63	4.62	2.43
Number of expert links	0.33	0.49	0.31	0.70	0.62	1.04
Number of valid concepts	5.67	3.37	8.88	3.36	7.77	5.05
Number of valid links	3.67	2.85	6.44	4.34	4.00	4.47

Number of resource requests	2.80	2.43	9.31	9.65	8.69	8.51
Time spent accessing resources (minutes)	8:31	8:22	8:59	5:39	12:52	4:51
Number of causal queries	0.80	1.15	1.19	1.72	2.38	2.66
Number of explanation requests	0.20	0.41	0.56	0.89	0.69	1.32
Number of quiz requests	3.80	3.53	5.63	3.12	6.15	19.97

Table L.4 Normality and Variance Equality Tests for the Transfer-Test Data

Variable	*Normality Test: Shapiro-Wilk*			*Variance Equality: Levene's Test*
	ITS	*LBT*	*SRL*	
Number of expert concepts	0.01	0.48	0.40	0.29
Number of expert links	N/A	N/A	N/A	N/A
Number of valid concepts	0.67	0.61	0.004	0.02
Number of valid links	N/A	N/A	N/A	N/A
Number of resource requests	0.06	0.001	0.01	0.04
Time spent accessing resources	.020	0.06	0.06	0.08
Number of causal queries	N/A	N/A	N/A	N/A
Number of explanation requests	N/A	N/A	N/A	N/A
Number of quiz requests	0.000	0.48	0.02	0.06

BIBLIOGRAPHY

Alexander, J. M., M. Carr, et al. (1995). "Development of metacognition in gifted children: Directions for future research." Developmental Review **15**: 1-37.

Anderson, J. R., A. T. Corbett, et al. (1995). "Cognitive tutors: Lessons learned." The Journal of Learning Sciences **4**: 167-207.

Angelo, T. A. and K. P. Cross (1993). Classroom Assessment Techniques: A Handbook for College Teachers. San Francisco, Jossey-Bass.

Arkin, R. C. and T. Balch (1997). "AuRA: principles and practice in review." Journal of Experimental & Theoretical Artificial Intelligence **9**(2): 175-189.

Armstrong, A.-M. (1998). The Effects of Self-Regulation on Efficacy and Problem Solving. Instructional Design & Development Program, College of Education. Mobile, University of South Alabama**:** 129.

Artzt, A. F. and E. Armour-Thomas (1999). "Cognitive Model for Examining Teachers' Instructional Practice in Mathematics: A Guide for Facilitating Teacher Reflection." Educational Studies in Mathematics **40**(3): 211-235.

Axelrod, R. M. (1976). The Analysis of Cognitive Maps. Structure of decision : the cognitive maps of political elites. R. M. Axelrod. Princeton, N.J., Princeton, N.J.

Bangert-Drowns, R. L., J. A. Kulick, et al. (1991). "The instructional effect of feedback in test-like events." Review of Educational Research **61**(2): 213-238.

Bargh, J. A. and Y. Schul (1980). "On the cognitive benefits of teaching." Journal of Educational Psychology **72**(5): 593-604.

Biswas, G., K. Leelawong, et al. (2004). Incorporating Self Regulated Learning Techniques into Learning by Teaching Environments. The twenty sixth Annual Meeting of the Cognitive Science Society, Chicago, Illinois, Lawrence Erlbaum.

Black, P. and D. Wiliam (1998). "Assessment and classroom learning." Assessment in Education **5**(1): 111 - 122.

Bonwell, C. C. and J. A. Eison (1991). Active Learning: Creating excitement in the Classroom. Washington, DC, School of Education and Human Development, George Washington University.

Bransford, J. D., A. L. Brown, et al., Eds. (2000). How People Learn. Washington, D.C., National Academy Press.

Brasell, H. (1987). "The effect of real-time laboratory graphing on learning graphic representations of distance and velocity." Journal of Research in Science Teaching **24**: 385-395.

Brasell, H. (1987). The role of microcomputer-based laboratories in learning to make graphs of distance and velocity. the Annual Meeting of the American Educational Research Association, Washington, D.C.

Briner, M. (1999). Constructivism. `http://www.dean.usma.edu/math/activities/cape/Constructivism/501const.htm`. **2002**.

Brown, J. S. (1985). "Process versus product:a perspective on tools for communal and informal eletronic learning." Journal of Educational Computing Research **1**: 179-201.

Bruner, J. S. (1987). Making sense : the child's construction of the world. New York, Methuen.

Burton, R. R. (1982). Diagnosing bugs in a simple procedural skill. Intelligent Tutoring Systems. D. H. Sleeman and J. S. Brown. Orlando, Academic**:** 157–183.

Butler, D. L. and P. H. Winne (1995). "Feedback and Self-Regulated Learning: A Theoretical Synthesis." Review of Educational Research **65**(3): 245-281.

Chan, C. K. and C.-Y. Chou (1995). Simulating a Learning Companion in Reciprocal Tutoring System.

Chan, T.-W. and C.-Y. Chou (1997). "Exploring the Design of Computer Supports for Reciprocal Tutoring." International Journal of Artificial Intelligence in Education **8**: 1-29.

Chi, M. T. H., M. Bassok, et al. (1989). "Self-explanations: How students study and use examples in learning to solve problems." Cognitive Science **13**(2): 145-82.

Chi, M. T. H., N. De Leeuw, et al. (1994). "Eliciting self-explanations." Cognitive Science **18**: 439-477.

Chi, M. T. H., S. A. Siler, et al. (2001). "Learning from Human Tutoring." Cognitive Science **25**(4): 471-533.

Chickering, A. W. and Z. Gamson (1987). "Seven principles for good practice in undergraduate education." AAHE Bulletin **39**(7): 3-7.

Clark, R. and R. E. Mayer (2003). e-Learning and the Science of Instruction. San Francisco, CA, Pfeiffer.

Cognition and Technology Group at Vanderbilt (1997). The Jasper Project: Lessons in Curriculum, Instruction, Assessment, and Professional Development. Mahwah, New Jersey, Lawrence Erlbaum Associates.

Cohen, P. A., J. A. Kulik, et al. (1982). "Educational Outcome of Tutoring: A Meta-analysis of Findings." American Educational Research Journal **19**(2): 237-248.

Collins, A. and J. S. Brown (1988). The Computer as a Tool for Learning Through Reflection. Learning Issues for Intelligent Tutoring Systems. H. Mandl and A. Lesgold. New York, NY, Springer-Verlag New York, Inc.**:** 1 - 18.

Colton, A. B. and G. M. Sparks-Langer (1993). "A Conceptual Framework to Guide the Development of Teacher Reflection and Decision Making." Journal of Teacher Education **44**(1): 45-54.

Cormier, S. M. and J. D. Hagman (1987). Introduction. Transfer of Learning: Contemporary Research and Applications. S. M. Cormier and J. D. Hagman. San Diego, Academic Press**:** 1-46.

Corno, L. and E. B. Mandinach (1983). "The role of cognitive engagement in classroom learning and motivation." Educational Psychology **19**(2): 88-108.

Crews, T., G. Biswas, et al. (1997). "Anchored Interactive Learning Environments." International Journal of Artificial Intelligence in Education **8**: 142-78.

Crews, T. R. (1995). Adventureplayer : a microworld anchored in a macrocontext. Computer Science. Nashville, TN, Vanderbilt University**:** 139.

Davis, J. M., K. Leelawong, et al. (2003). Intelligent User Interface Design for Teachable Agent Systems. International Conference on Intelligent User Interfaces, Miami, Florida, The Association for Computing Machinery.

Detterman, D. K. (1993). The Case for the Prosecution: Transfer as an Epiphenomenon. Transfer on Trial: Intelligence, Cognition, and Instruction. D. K. Detterman and R. J. Sternberg. Norwood, NJ, Ablex Publishing**:** 1-24.

Dewey, J. (1933). How We Think. Boston, MA, Houghton Mifflin.

Elawar, M. C. and L. Corno (1985). "A factorial experiment in teachers' written feedback on student homework: Changing teacher behaviour a little rather than a lot." Journal of Educational Psychology **77**(2): 162-173.

Flavell, J. H. (1976). Metacognitive aspects of problem solving. The Nature of Intelligence. L. B. Resnick. Hilldale, N.J., Lawrence Erlbaum**:** 231-235.

Franks, J. J., N. J. Vye, et al. (1982). "Learning from explicit versus implicit texts." Journal of Experimental Psychology, General **111**: 414-422.

Gardner, R. C. (2001). Psychological Statistics Using SPSS for Windows. Upper Saddle River, NJ, Prentice-Hall.

Goodman, B., A. Soller, et al. (1998). "Encouraging Student Reflection and Articulation using a Learning Companion." International Journal of Artificial Intelligence in Education **9**: 237-55.

Graesser, A. C., N. Person, et al. (1995). "Collaborative dialog patterns in naturalistic one-on-one tutoring." Applied Cognitive Psychologist **9**: 359-387.

Haskell, R. E. (2001). Transfer of Learning: Cognition, Instruction, and Reasoning. San Diego, CA, Academic Press.

Hirschberg, D. S. (1977). "Algorithms for the Longest Subsequence Problem." Journal of the Association for Computing Machinery **24**(4): 664-675.

Holmes, A. (2002). Lifelong Learning. Oxford, United Kingdom, Capstone Publishing.

Jarvis, P. (2002). Certificating Teachers of Adults: an epistemological discussion. Fokus på voksenlæreren: Om kvalificering af lærere inden for den almene voksenundervisning. `http://pub.uvm.dk/2002/fokus/12.htm`. Copenhagen, Denmark, Danish Ministry of Education. **2004**.

Jaušovec, N. (1994). "Edge giftedness Be taught?" Roeper Review **16**: 210-214.

Johnstone, A. H. (1993). Introduction. Creative problem solving in Chemistry. C. Wood and R. Sleet. London, The Royal Society of Chemistry**:** iv-vi.

Jonassen, D. H. (1994). "Thinking technology: Toward a constructivist design model." Educational Technology **34**(4): 34-37.

Joolingen, W. R. v. and T. d. Jong (1996). "Design and implementation of simulation-based discovery environments: the SMISLE solution." International Journal of Artificial Intelligence in Education **7**: 253-76.

Kinchin, I. M. and D. B. Hay (2000). "How a qualitative approach to concept map analysis can be used to aid learning by illustrating patterns of conceptual development." Educational Research **42**(1): 43-57.

Koschman, T. (1994). "Toward a Theory of of computer support for collaborative learning." Journal of the Learning Sciences **3**: 219-225.

Kozma, R. (1987). "The Implications of Cognitive Psychology for computer-based learning Tools." Educational Technology **40**(11): 20-25.

Kuipers, B. (1994). Qualitative Reasoning: Modeling and Simulation with Incomplete Knowledge. Cambridge, MA, MIT Press.

Lajoie, S. P. and S. J. Derry (1993). Computer Environments as Cognitive Tools for Enhancing Learning. Computer as Cognitive Tools. S. P. Lajoie and S. J. Derry. Hillsdale, NJ, Lawrence Erlbaum Associates.

Lajoie, S. P. and S. J. Derry (1993). A Middle Camp for (Un)intelligent Instructional Computing: An Introduction. Computer as Cognitive Tools. S. P. Lajoie and S. J. Derry. Hillsdale, NJ, Lawrence Erlbaum Associates.

Lajoie, S. P. and A. Lesgold (1989). "Apprenticeship training in the workplace: Computer coached practice environment as a new form of apprenticeship." Machine Mediated Learning **3**(1): 7-28.

Leelawong, K., J. Davis, et al. (2002). The Effects of Feedback in Supporting Learning by Teaching in a Teachable Agent Environment. The Fifth International Conference of the Learning Sciences, Seattle, Washington.

Leelawong, K., Y. Wang, et al. (2001). Qualitative reasoning techniques to support learning by teaching: The Teachable Agents project. Fifteenth International Workshop on Qualitative Reasoning, San Antonio, Texas, AAAI Press.

Michie, D., A. Paterson, et al. (1989). Learning by Teaching. Second Scandinavian Conference on Artificial Intelligence (SCAI) 1989, Tampere, Finland, IOS Press.

Moses, L. J. and J. Baird (1999). Metacognition. The MIT Encyclopedia of the Cognitive Sciences. R. A. Wilson and F. C. Keil. Cambridge, Mass., MIT Press**:** 964.

Neufeld, V. R. and H. S. Barrows (1974). "The McMaster philosophy: An approach to medical education." Journal of medical education **49**: 1040-1050.

Nichols, D. M. (1994). Intelligent Student Systems: an application of viewpoints to intelligent learning environments. Computing. Lancaster, UK, Lancaster University.

Nichols, D. M. (1994). Issues in Designing Learning by Teaching Systems. Proceedings of the East-West International Conference on Computer Technologies in Education (EW-ED'94), Crimea, Ukraine.

Novak, A. (1964). "Scientific inquiry." Bioscience **14**: 25-8.

Novak, J. D. (1996). Concept Mapping as a tool for improving science teaching and learning. Improving Teaching and Learning in Science and Mathematics. D. F. Treagust, R. Duit and B. J. Fraser. London, Teachers College Press**:** 32-43.

Obayashi, F., H. Shimoda, et al. (2000). Construction and Evaluation of CAI System Based on Learning by Teaching to Virtual Student'. World Multiconference on Systemics, Cybernetics and Informatics, Orlando, Florida.

Palincsar, A. S. and A. L. Brown (1984). "Reciprocal teaching of comprehension-fostering and comprehension -monitoring activities." Cognition and instruction **1**: 117-175.

Palthepu, S., J. E. Greer, et al. (1991). Learning by teaching. The International Conference on Learning Sciences, Illinois, USA.

Papert, S. (1993). The Children's Machine: Rethinking school in the age of the computer. New York, NY, Basic Books.

Piaget, J. (1954). The construction of reality in the child. New York, Basic Books.

Piaget, J. (1978). Success and Understanding. Cambridge, MA, Harvard University Press.

Pintrich, P. R. (1991). A Manual for the use of the motivated strategies for learning questionnaire (MSLQ). Ann Arbor, Mich., University of Michigan ; [Washington, DC] : U.S. Dept. of Education, Office of Educational Research and Improvement, Educational Resources Information Center.

Pintrich, P. R. and E. V. DeGroot (1990). "Motivational and self-regulated learning components of classroom academic performance." Journal of Educational Psychology **82**(1): 33-40.

Ponder, G. and J. Kelly (1997). "Evolution, chaos, or perpetual motion? A prospective trend analysis of secondary science curriculum advocacy, 1955-1994." Journal of Curriculum and Supervision **12**(3): 238-245.

Pressley, M., J. G. Borkowski, et al. (1987). Cognitive strategies: good strategy users coordinate metacognition and knowledge. Annals of Child Development. R. Vasta. Greenwich, CT, JAI Press. **5:** 89-129.

Quillian, M. R. (1968). Semantic Memory. Semantic Information Processing. M. Minsky, MIT Press**:** 216-270.

Ramaprasad, A. (1983). "On the definition of feedback." Behavioral Science **28**(1): 4-13.

Ramírez Uresti, J. A. (1998). Teaching a learning companion. The tenth White House papers. Graduate research in the cognitive and computing sciences at Sussex. J. Halloran and F. Retkowsky. Brighton, UK, University of Sussex.

Ramírez Uresti, J. A. (2000). Should I Teach My Computer Peer? Some Issues in Teaching a Learning Companion. Proceedings of the 5th International Conference on Intelligent Tutoring Systems, Montréal, Canada, Springer-Verlag.

Reif, F. and L. A. Scott (1999). "Teaching scientific thinking skills: Students and computers coaching each other." American Journal of Physics **67**: 819-831.

Ruhl, K. L., C. A. Hughes, et al. (Winter 1987). "Using the Pause Procedure to Enhance Lecture Recall." Teacher Education and Special Education **10**: 14-8.

Russell, S. and P. Norvig (1995). Artificial Intelligence: A Modern Approach. Upper Saddle River, NJ, Prentice Hall.

Sadler, D. R. (1989). "Formative assessment and the design of instructional systems." Instructional Science **18**(2): 119-144.

Schmidt, D. R. (1997). Technology in the Classroom. the 14th annual Excellence in Education Conference, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Schwartz, D. L. (1995). "The Emergence of Abstract Representations in Dyad Problem Solving." Journal of the Learning Sciences **4**(3): 321-354.

Schwartz, D. L. and J. D. Bransford (1998). "A Time for Telling." Cognition and Instruction **16**(4): 475-522.

Seeskin, K. (1987). Dialogue and discovery : a study in Socratic method. Albany, State University of New York Press.

Shapiro, A. (1983). The Role of Structured Induction in Expert Systems. Machine Intelligence Unit. Edinburgh, UK, University of Edinburgh.

Shaw, E., W. L. Johnson, et al. (1999). Pedagogical Agents on the Web. the Third Int'l Conf. on Autonomous Agents, Seattle, Washington.

Sloman, M. (2002). The E-Learning Revolution : How Technology Is Driving a New Training Paradigm. New York, NY, AMACOM.

Soller, A. L. (2001). "Supporting Social Interaction in an Intelligent Collaborative Learning System." International Journal of Artificial Intelligence in Education **12**(1): 40-62.

Stoney, S. and R. Oliver (1999). "Can higher order thinking & cognitive engagement be enhanced with multimedia?" Interactive Multimedia Electronic Journal of Computer-Enhanced Learning **1**(2).

Stoyanov, S. and P. Kommers (1999). "Agent-Support for Problem Solving Through Concept-Mapping." Journal of Interactive Learning Research **10**(3/4): 401-42.

Tolman, E. C. (1948). "Cognitive Maps in Rats and Men." The Psychological Review **55**(4): 189-208.

Trochim, W. M. K. (2001). The Research Methods Knowledge Base. Cincinnati, OH, Atomicdogpublishing.com.

Tsaparlis, G. (2001). Problem Solving in Chemistry and Science Education. Research on Mathematics and Science Education: From Beliefs to Cognition, from Problem SOlving to Understanding. M. Ahtee, O. Bjorkqvist, E. Pehkonen and V. Vatanen. Jyväskylä, Finland, Institute for Educational Research, University of Jyväskylä**:** 136.

Tubbs, S. L. (1984). A Systems Approach to Small Group Interaction. New York, Random House.

Viswanath, K., R. Balachandran, et al. (2003). Effective User Interface Design for Teachable Agent Systems. Proceedings of the ACM Southeast Conference (ACMSE03), Savannah, GA.

Vygotsky, L. S. (1978). Mind in Society: The Development of the Higher Psychological Processes. Cambridge, MA, The Harvard University Press.

Wenger, E. (1987). Artificial Intelligence and Tutoring Systems. Los Altos, California, Morgan Kaufmann Publishers.

Zimmerman, B. J. (1986). "Development of self-regulated learning: Which are the key sub processes?" Contemporary Educational Psychology **16**: 307-313.

Zimmerman, B. J. (1989). "A Social Cognitive View of Self-Regulated Academic Learning." Journal of Educational Psychology **81**(3): 329-339.