Pedagogical Agents for Learning by Teaching: Teachable Agents

Kristen Blair & Daniel L. Schwartz
Stanford University
Gautam Biswas & Krittaya Leelawong[1]
Vanderbilt University

To appear in *Educational Technology*

We describe pedagogical agents called Teachable Agents (TAs) where students learn by teaching the computer. We describe an example of a TA, and discuss the features that allow students to capitalize on learning-by-teaching interactions. These include (i) explicit well-structured shared visual representations, (ii) independent performance of the agent, (iii) the agent's ability to model productive learner behavior, and (iv) embedding the agent in environments that support teaching. Finally, we describe new directions for TAs including new models of homework practice, assessment, and video game environments.

Almost everyone has had the experience of learning when teaching. Many graduate students observe they never really understood a topic until they had to teach it. To cultivate the benefits of learning by teaching, we have created a special kind of pedagogical agent that we call a Teachable Agent (TA). Students teach their TA and then assess its knowledge by asking it questions or by getting it to solve problems. The TA uses artificial intelligence techniques to generate answers based on what it was taught. Depending on the TA's answer, students can revise their agents' knowledge (and their own).

TAs do not replace real students. But, they do provide unique opportunities to optimize learning-by-teaching interactions. TAs, for example, always make their thinking visible, something that not all students can do. This raises the question: *What aspects of learning-by-teaching can we maximize with TAs?* We start with a concrete description of a TA. We then describe four core learning-by-teaching design principles that we believe can maximize learning-by-teaching. We conclude with instances of how these principles enable us to introduce exciting new technologies that leverage learning-by-teaching.

[1] Currently at Mahidol University International College, Thailand

BETTY: AN EXAMPLE OF A TEACHABLE AGENT

For this article, we spotlight our agent named Betty. Students teach Betty by creating a network of entities and their relations, much like a concept map. Figure 1 provides an example. Students use a point-and-click editor to create nodes (e.g., LDL, arterial plaque) and connect pairs of nodes with links. The links are labeled (e.g., LDL builds up arterial plaque) and categorized using a pull-down menu (e.g., builds up implies and increase) (Biswas et. al, 2005).

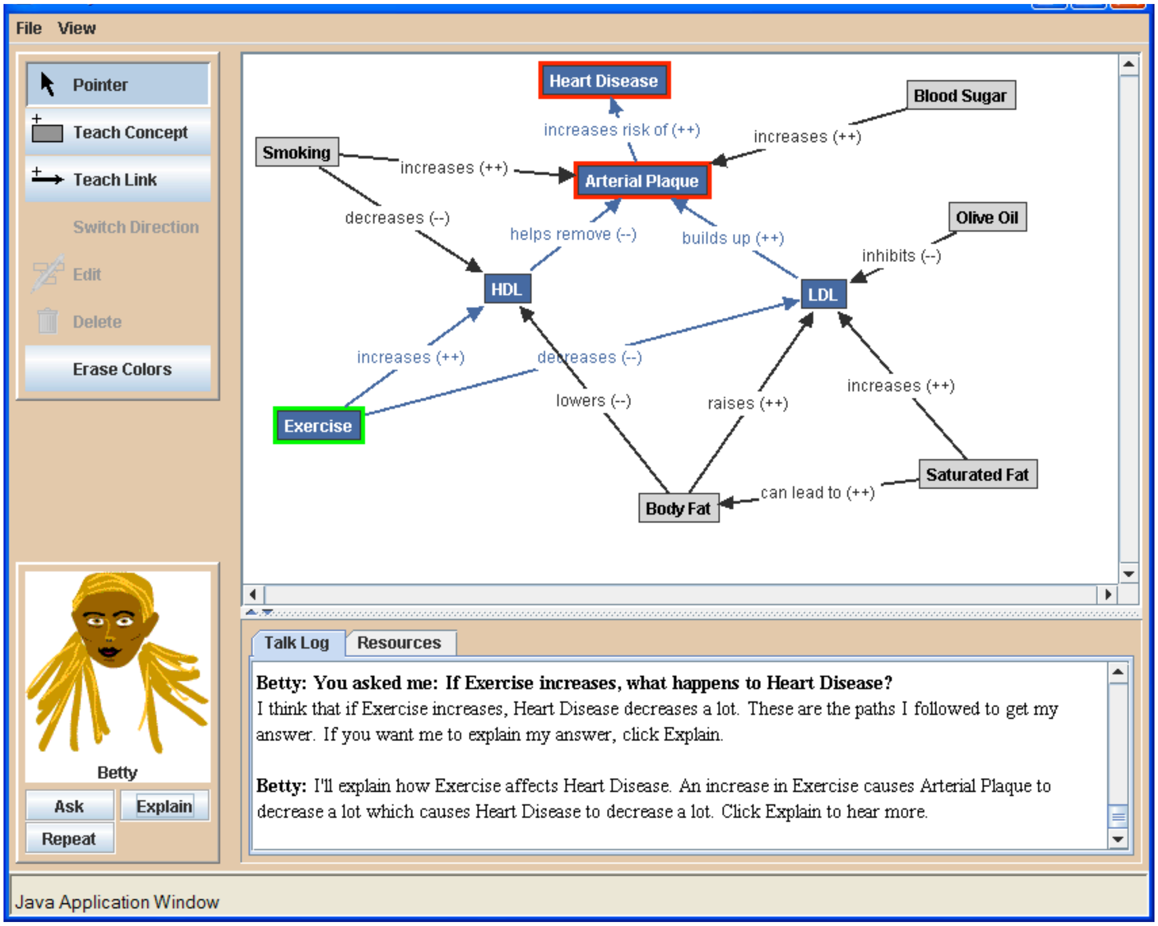

Figure 1. The Teachable Agent Betty. Students teach Betty by making a concept map. Once Betty has been taught, she can answer questions by tracing links through the concept map.

At any point, students can ask Betty a question to see how well she is learning. Figure 1 shows how Betty animates her reasoning for the question, *"What happens to heart disease if exercise increases?"* To make it easier for the student to follow her reasoning process, Betty breaks down the explanation into parts. For example, Betty reasons that exercise increases HDL cholesterol and that exercise decreases LDL. Increasing HDL and decreasing LDL together result in decreased arterial plaque, which in turn decreases the risk of heart disease. To complement her graphical thinking, Betty unfolds her reasoning in text (lower panel). Students can observe Betty's conclusions and decide whether they need to revise what they have taught Betty. Betty can also take a quiz composed by a classroom instructor but automatically scored by the computer. So, instead of students taking the quiz, they can watch their agent perform and receive projective feedback on their own knowledge.

FOUR CORE PRINCIPLES OF TEACHABLE AGENTS

Betty is one instance of a Teachable Agent (for other instances, see Schwartz et. al.; in press; Blair and Schwartz, 2004). Like all our TAs, Betty includes features that we think enhance the experience of learning-by-teaching. Some features are implicit in the TA metaphor. For example, unlike a human learner, a computer agent will not be hurt if the student is a bad teacher or has missing or incomplete knowledge. Other features derive from learning-by-teaching design principles that we have tested with hundreds of students of all ages (for empirical reviews see, Biswas et al., 2001; Biswas et. al., 2005; Schwartz et al., in press). We describe the four most important principles.

Use Explicit and Well-Structured Visual Representations

A critical component of learning-by-teaching is knowing what one's tutee has in mind. Unlike human learners, we design the computer-based TAs to make their thinking visible using well-structured visual representations and intuitive reasoning mechanisms. For example, Betty's representation of causal relations provides a well-structured graph that is both intuitive and common among experts reasoning about causal propagation. Betty's visual representation helps students formalize and organize their own thinking. For example, in one study, we asked college students to read a passage on metabolism. Half of the students taught Betty and half wrote a summary of the passage. Students who taught Betty adopted her knowledge representation, for example, by exhibiting more complex chains of causal reasoning on a posttest.

Enable the Agent to Take Independent Actions

The second design principle is that TAs take independent actions based on what they have been taught. It is not until students see Betty reason independently that they can get useful feedback and models of reasoning. This opportunity to observe how one's student performs is much more informative than just teaching and not seeing the effects of that teaching. In one study, for example, we found that students learned more when they saw their agent perform compared to when the students taught their agent and then solved the problems themselves instead of the agent. Moreover, it is quite motivating to see how one's tutee performs. For example, middle-school students had to help Betty pass a test so she could join a science club. Students were eager to help Betty improve and retake the test several times, something they are not always willing to do when they need to retake a test themselves.

Model Productive Learner Behaviors

Another important component of learning-by-teaching is the need to monitor the tutee's performance, and to pay attention to new ideas the tutee might introduce. Young children do not spontaneously know how to monitor thinking. Therefore, to help young students learn these important skills, TAs model good learners (Biswas, Schwartz, Leelawong, Vye, & TAG-V, 2005). Under specific conditions, Betty can spontaneously offer a learning strategy or a concern. For example, as students build Betty's map, she occasionally starts to check her understanding. She then remarks (right or wrong) that the answer she is deriving does not seem to make sense. These spontaneous prompts help students reflect on what they are teaching, and hopefully like a good teacher, check on their tutee's learning progress.

Include Environments that Support Teaching Interactions

Teaching occurs in a larger context that includes goals such as preparing a student for a test, a performance, a job, or a fun but complex game. The larger context of teaching helps to motivate and focus a teacher. As we mentioned above, preparing Betty to take a test is highly motivating and sets high standards. At the same time, a larger context typically also provides resources that facilitate achieving these goals. There can be books on the shelf, practice tests, model answers, a homework hotline, a nearby friend, and so on. In the computerized context of TAs, we build environments that provide students with a number of learning resources. Betty can appear in an environment that includes searchable hypertext explanations of key topics. Betty can also appear in an environment with a Mentor Agent who provides a combination of domain knowledge and feedback, plus offers tips on how to monitor Betty's knowledge. Importantly, the learning-by-teaching environments should not undermine students' need to actively decide how to use the resources and feedback available in the environment. In one study, we compared 5^{th}-grade students who had to make decisions about how to use resources and feedback versus students who were told exactly what to do. Students who had to make decisions about resource use and teaching were able to learn more effectively a month later on their own, even when they were no longer learning in the context of teachable agents (Biswas, Schwartz, Leelawong, Vye, & TAG-V, 2005).

ADDITIONAL LEARNING-BY-TEACHING POSSIBILITIES

The four core principles are common to all the TAs. At the same time, we can extend each TA in different directions to further support learning. TAs have a modular software architecture that allows us to repurpose and combine them into different applications. People learn by teaching in many ways, and in the following section, we provide some examples of how we repurpose Betty to optimize different aspects of learning-by-teaching

The independent performance of TAs offers excellent assessment opportunities. In the preceding examples, we described how students could assess their agents (and themselves) one-on-one. We have extended this capability into a front-of-the-class assessment system so students can see the performance of their own agent, as well as that of their classmates. The agents created by each student can be displayed side-by-side. The classroom teacher can ask a question of all the agents simultaneously. A hidden expert map determines the correct answer. The results are tabulated and indicated by color coding. The classroom teacher can zoom in on a map to see why an agent gave the answer it did, and then compare it to another map as shown in Figure 2,. In a formal study with college classes, we have found that the front-of-the-class system helps students learn the structure of a domain better than just seeing the performance of their own agent.

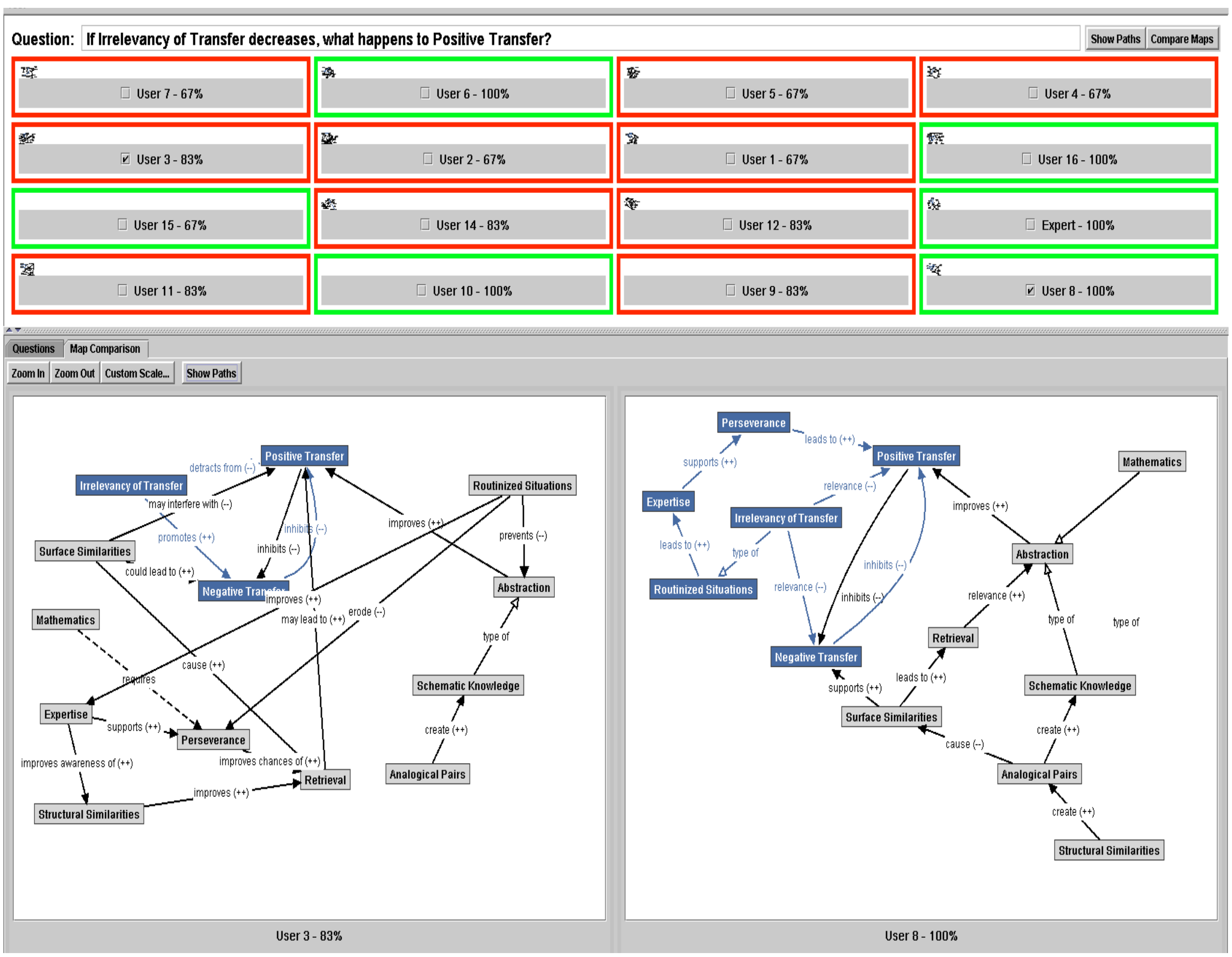

Figure 2. Front of class quiz system for showing agents perform. The classroom teacher can ask all the agents the same question simultaneously. The top of the panel indicates whether each agent answered the question correctly with red and green highlighting. The bottom of the panel shows how the classroom teacher sets up a class discussion by highlighting how two different agents reasoned about the same problem.

Rethinking Homework

With TAs, we can reach beyond the classroom. The Triple-A Game Show is an internet application designed to change homework practices (Figure 3). The student teaches his or her agent and customizes the agent's look. The student and agent then participate in an on-line game show with other students and their agents. Students can log on from home or school. The game host asks agents to answer questions and show their thinking. Students can see their own agent reason, and also observe the reasoning of other students' agents. The application also includes a chat environment so students can discuss and cheer (or jeer) an agent's performance. Our hope is that students will find this socially rich environment both engaging and educational, and it will prepare them for their lessons in school the next day or week.

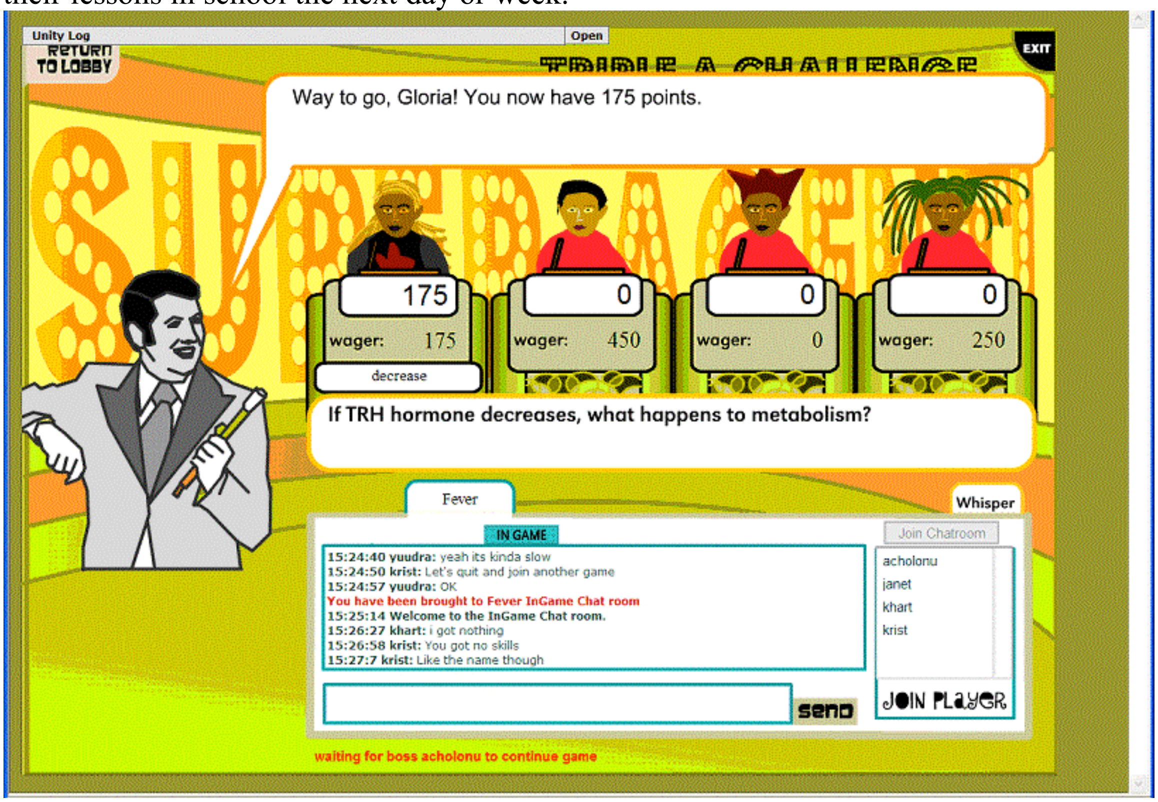

Figure 3. A customized agent performs in an on-line game show with other agents.

Videogames that Capture the Many Ways that People Learn

People learn in many ways. In our final example, we combine many forms of learning and teaching interactions into a guided-discovery video game called *Pumpkin World* (Blair & Schwartz, 2005). Figure 4 provides a sample screen shot. Betty takes the form of an embodied agent, and students teach her how to grow giant pumpkins so she can win contests at a pumpkin festival. In Pumpkin World, Betty can take actions (instead of just answer questions). Betty, for example, reasons about what nutrients to add to the soil to increase pumpkin growth. As she takes actions in the game world, the student can see the functional consequences as the pumpkin grows or wilts. The environment and narrative of the game support many other types of learning interactions and resources as well. For example, students learn about the role of nitrogen through experimentation; they learn about phosphorous by observing another agent; they learn about "energy" by being told by another agent; and so on. Connecting TAs with video games is not only motivating, it also provides a way to bring together a variety of resources and interactions into one coherent problem solving environment.

Figure 4. Betty in a guided-discovery video game. The window in the corner shows Betty reasoning about a challenge in the game world.

CONCLUSIONS

Using the common wisdom that people learn by teaching we have developed pedagogical agents that students must explicitly teach. Unlike most instructional technology where the computer teaches and tests the students, with TAs students teach and assess computer. Students readily adopt the fiction of teaching an agent. They find it motivating, and it helps them to organize otherwise complex human-computer interactions and learning tasks. Our research has also shown valuable learning benefits of teachable agents. Perhaps, equally important, we have found that the TA metaphor to be a generative source of novel ideas for creating new forms of instruction.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grants REC-0231946, REC-0231771, and Stanford's Media-X. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation or Media-X.

REFERENCES

Biswas, G., Schwartz, D. L., Bransford, J. D., & TAG-V. (2001). Technology support for complex problem solving: From SAD Environments to AI. In K. Forbus & P. Feltovich (Eds.), Smart machines in education (pp. 71-98). Menlo Park, CA: AAAI/MIT Press.

Biswas, G., Schwartz, D. L., Leelawong, K., Vye, N., & TAG-V. (2005). Learning by teaching: A new paradigm for educational software. Applied Artificial Intelligence, 19(3).

Blair, K. P., and Schwartz, D. L. (2005) Guided Discovery Games with Teachable Agents. Games, Learning and Society Conference, Madison, WI.

Blair, K. P., and Schwartz, D. L. (2004) Milo and J-Mole: Computers as constructivist teachable agents. In Y. Kafai et al. (Eds.) Embracing Diversity in the Learning Sciences: The proceedings of the Sixth International Conference of the Learning Sciences (p. 588). Mahwah, NJ: Erlbaum.

Schwartz, D. L., Blair, K., Biswas, G., Leelawong, K., & Davis, J. (in press). Animations of thought: Interactivity in the teachable agent paradigm. To appear in R. Lowe & W. Schnotz (Eds). Learning with animation: Research and implications for design. UK: Cambridge University Press.