

# Using Learner Focus of Attention to Detect Learner Motivation Factors

Lei Qu, Ning Wang, and W. Lewis Johnson

Center for Advanced Research in Technology for Education (CARTE), USC / ISI  
4676 Admiralty Way, Suite 1001, Marina del Rey, CA, 90292  
{leiqu, ning, johnson}@isi.edu

**Abstract.** This paper presents a model for pedagogical agents to use the learner's attention to detect motivation factors of the learner in interactive learning environments. This model is based on observations from human tutors coaching students in on-line learning tasks. It takes into account the learner's focus of attention, current task, and expected time required to perform the task. A Bayesian model is used to combine evidence from the learner's eye gaze and interface actions to infer the learner's focus of attention. Then the focus of attention is combined with information about the learner's activities, inferred by a plan recognizer, to detect the learner's degree of confidence, confusion and effort. Finally, we discuss the results of an empirical study that we performed to evaluate our model.

## 1 Introduction

Animated pedagogical agent technology seeks to improve the effectiveness of intelligent tutoring systems, by enabling them to provide various interactive actions in more natural and engaging ways. However, work to date has focused mainly on improving the output side of the interface, through the inclusion of expressive, lifelike behaviors [2]. The focus of the work described in this paper is on the input side, to enable the agent to track the learner's activities and focus of attention, so that it can detect the learner's motivational factors such as confidence and confusion.

Our approach involves monitoring of the learner's activities - both interface actions and focus of eye gaze. It infers the learner's focus of attention using a Bayesian model [5], which allows reasoning under uncertainty with various sources of information. And it combines the method for tracking learner focus of attention with a plan recognition capability for interpreting the learner's actions and forming expectations of future actions. Our particular motivation for conducting this work is to create pedagogical agents that are able to interact with learners in more socially appropriate ways, sensitive to rules of politeness and etiquette and able to influence the learner motivational as well as cognitive state [3, 4].

## 2 Background Tutoring Studies

In an earlier study, we investigated how human tutors coach learners while interacting with Virtual Factory Teaching Systems (VFTS) [1, 4], an on-line factory system for teaching industrial engineering concepts and skills. We found that tutors used the following types of information, observed and inferred the learner motivation:

- The task that the learner was expected to perform next.
- The learner's focus of attention.
- The learner's self-confidence, inferred from the questions the learner asked.
- The learner's effort expended, as evidenced by the amount of time that the learner spent reading the tutorial and carrying out the tasks described there.

We therefore designed the user interface of our new system to enable an agent to have access to sufficient information about the learner, her/his activities, cognitive and motivational state. The new interface includes three major components:

- The VFTS interface, which reports each keyboard entry and mouse click that the learner performs on it.
- WebTutor, which is an on-line tutorial used to teach learner instruction and concepts of industrial engineering.
- Agent window, in which the left part of this window is a text window used to communicate with the agent (or a human tutor in Wizard-of-Oz mode) and the right part is an animated character that is able to generate speech and gestures.

The input devices consist of keyboard, mouse, and a small camera focused on the learner's face. This interface thus provides information that is similar to the information that human tutors use in tracking learner activities.

## 3 Description of Our Model

Our model includes four components for pedagogical agent to access the learner's focus of attention and track the learner's activities to detect the learner's motivation.

- WebTutor provides information about what task the learner is working on, as well the actions the learners perform as they read through the tutorial.
- The plan recognizer in VFTS monitors the learner's actions and tracks learner progress through the task.
- The focus of attention model takes input from the WebTutor interface, the VFTS interface and Agent interface as well as eye gaze information, in order to infer learner focus of attention.
- The detection model calculates the learner motivational factors based on outputs from focus of attention model and plan recognizer.

These four components can provide agents with capabilities to gather information about the learners' states and their expected tasks. Therefore agents are able to track learner attention and detect the learners' motivation.

### 3.1 Tracking Learner Focus of Attention Under Uncertainty

Information about eye gaze is extremely useful for detecting user focus of attention. In our system we want an eye tracking approach that is unobtrusive, that requires no special hardware and no calibration. We use a program developed by Larry Kite in the Laboratory for Computational and Biological Vision at USC to track eye gaze. It estimates the coordinates on a video display that correspond to the focus of gaze. The agent uses two types of information to infer the learner's focus: (1) information with certainty, i.e., mouse click, type and scroll window events in VFTS, WebTutor and Agent Window, and (2) information with uncertainty, namely data from eye track program and inferences about current state based upon past events. Then agents use Dynamic Bayesian Networks (DBNs) to infer the learner's focus of attention based on various sources of information.

### 3.2 Plan Recognition System

To help pedagogical agents track learner activities, we need to be able to track the learner's actions as well as track the learner's focus. We developed a plan recognition system to track learner actions and progress. The plan recognition system has four main components: .NET server, Student Interaction Database (SID), Plan Library and Action Pattern file in VFTS. The .NET server has two services: data service and agent service. All student interaction data in VFTS are captured and encoded into SOAP messages, sent to the data service and then saves it to SID. Interaction data represent learner action on the current object in the VFTS. An object in VFTS could be a textfield, a tab panel or a combobox. Interaction data include mouse clicks, mouse movements and keyboard input. A plan in the plan library consists of a list of tasks the user needs to achieve given a tutorial or problem. The VFTS action pattern XML file defines the actions a user could perform in VFTS.

### 3.3 Utilizing Focus of Attention

The above analyses make it possible for the agent to track learner's attention and actions. There are many factors that influence learner's motivation. We focus here on the learner's confidence, confusion and effort, factors that were shown to be important in the background tutor studies.

- Confidence represents the confidence of learners in solving problems in the learning environment. The learner's confidence is modeled as one of three levels: High, Normal and Low.
- Confusion reflects the learner's failing to understand the tutorial or decide how to proceed in the VFTS. A learner with high confusion is most likely to be stuck or frustrated.
- Effort is the duration of time that the learner spends on performing tasks. It is an important indicator of intrinsic motivation in learners, and expert human tutors often praise learners for expending effort even when they are not successful.

## 4 Evaluation and Conclusion

To evaluate our method, we designed and conducted an experimental study. With new interfaces and models, we ran 24 subjects at the University of California at Santa Barbara. The 24 participants were all undergraduate students. Most of them had computer skills but little or no knowledge of industrial engineering. With the human tutor's observation as the baseline, the recognition accuracies of our model are 82% for confidence, 76.8% for confusion, and 76.3% for effort. With the learner's self-reports as the baseline, the recognition accuracies dropped to 70.7%, 75.6% and 73.2% for the learner's motivation. In conclusion, we can say that the results of our evaluation suggest that such model can provide agents accurate information about learner's motivation. It is possible for pedagogical agents to detect learner's motivation with confidence and provide learner with proactive help in order to motivate the learner's learning.

Furthermore we wish to extend the user monitoring capability to handle a wider range of ambiguous contexts. Based upon these results, pedagogical agent can then interact with learners through a conversational system in more socially appropriate ways.

## Acknowledgements

This work was supported in part by the National Science Foundation under Grant No. 0121330, and in part by a grant from Microsoft Research. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

## References

1. Dessouky, M.M., Verma, S., Bailey, D., Richel, J.: A methodology for developing a Web-based factory simulator for manufacturing education. *IEEE Transactions*, 33 (2001) 167-180
2. Johnson, W.L., Rickel, J.W., Lester, J.C.: Animated pedagogical agents: Face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education*, 11 (2000) 47-78
3. Johnson, W.L.: Using Agent Technology to Improve the Quality of Web-Based Education. In N. Zhong and J. Liu (Eds.), *Web Intelligence*, Springer-Verlag, Berlin Heidelberg New York (2002)
4. Johnson, W.L.: Interaction Tactics for Socially Intelligent Pedagogical Agents. In *Proceedings of the Intelligent User Interfaces* (2003) 251-253
5. Pearl, J.: *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan-Kaufmann, San Mateo, CA (1988)