Empathic Tutoring Software Agents Using Real-time Eye Tracking

Hua Wang
University of Toronto/Tokyo
Tel: 1-416-978-3045
hwang@mie.utoronto.ca

Mark Chignell
University of Toronto
Tel: 1-416-978-3045
chignel@mie.utoronto.ca

Mitsuru Ishizuka
University of Tokyo
Tel: 81-3-5841-6755
ishizuka@miv.t.y-tokyo.ac.jp

Abstract
This paper describes an empathic software agent (ESA) interface using eye movement information to facilitate empathy-relevant reasoning and behavior. Eye movement tracking is used to monitor user's attention and interests, and to personalize the agent behaviors. The system reacts to user eye information in real-time, recording eye gaze and pupil dilation data during the learning process. Based on these measures, the ESA infers the focus of attention and motivational status of the learner and responds accordingly with affective (display of emotion) and instructional behaviors. In addition to describing the design and implementation of empathic software agents, this paper will report on some preliminary usability test results concerning how users respond to the empathic functions that are provided.

Keywords
Eye tracking, e-learning, character agent, tutoring, tracing, educational interface, eye movements, eye-aware interfaces

1 Introduction

Learners can lose motivation and concentration easily, especially in a virtual education environment that is not tailored to their needs, and where they may be little contact with live human teachers. As Palloff and Pratt [1] noted "the key to success in our online classes rests not with the content that is being presented but with the method by which the course is being delivered" (p. 152). In traditional educational settings, good teachers recognized learning needs and learning styles and adjusted the selection and presentation of content accordingly.

In online learning there is a need to create more effective interaction between e-learning content and learners. In particular, increasing motivation by stimulating learner's interest is important. A related concern is how to achieve a more natural and friendly environment for learning. We will address this concern by detecting the attention information from the real-time eye tracking data from each learner and modify instructional strategies based on the different learning patterns for each learner.

Eye movements provide an indication of learner interest and focus of attention. They provide useful feedback to character agents attempting to personalize learning interactions. Character agents represent a means of bringing back some of the human functionality of a teacher. With appropriately designed and implemented animated agents, learners may be more motivated, and may find learning more fun. However, amusing animations in themselves may not lead to significant improvement in terms of comprehension or recall. Animated software agents need to have intelligence and knowledge about the learner, in order to personalize and focus the instructional strategy.

Figure 1 shows an ESA as a human-like figure embedded within the content on a Web page. In this paper, we use real time eye gaze interaction data as well as recorded study performance to provide appropriate feedback to character agents, in order to make learning more personalized and efficient. This paper will address the issues of when and how such agents with emotional interactions should be used for the interaction between learners and system.

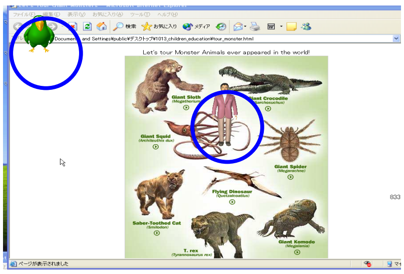

Figure 1 Interface Appearance

2 Related Work

Educational research, as well as anecdotal observation, indicates that the human qualities and psychological insights of teachers are important components in learning, motivating learning performance, and adapting content presentation in response to user needs. Cooper et al. [2] argue that showing emotions, empathy, and understanding through facial expressions and body language is essential in tutor-learner and learner-learner interaction. For instance, eye gaze (i.e., teacher looking at pupil) has been shown to significantly enhance information recall performance (Ottenson [3]; Sherwood [4]). Regarding teacher reactions to success and failure affecting children's expectations of success, Graham [5] showed that children who received sympathy from the experimenter after failing at a task, tended to attribute that failure to their lack of ability, while children who received mild anger tended to attribute their failure to lack of effort.

Several attempts have been made to build learner models that consider emotion and affective state. Klein et al. [6] described an interactive system that responds to the learner's self-reported frustration while playing a computer game. MIT's affective learning companion [7] developed the system for using affective sensing and appropriate relational-agent interactions to support learning and meta-cognitive strategies for perseverance through failure. Ou et al. [8] described a system that took account of the attention of the learner, and the expected time required to perform task, to choose when to interact with learners. In some educational contexts, character agents represent a person. For instance, Fabri et al. [9] described a system for supporting meetings between people in educational virtual environments using quasi face-to-face communication via their character agents. In other cases, the agent is a stand-alone software agent, rather than a persona or image of an actual human. Stone et al. in their COSMO system used a life-like character that teaches how to treat plants [10].

Eye tracking is an important tool for detecting users' attention information and focus on certain content. Applications using eye tracking can be diagnostic or interactive. In diagnostic use, eye movement data provides evidence of the learner's focus of attention over time and can be used to evaluate the usability of interfaces [11] or to guide the decision making of a character agent. For instance, Johnson [12] used eye-tracking to assist character agents during foreign language/culture training. In interactive use of eye tracking for input, a system responds to the observed eye movements, which can thus serve as an input modality [13]. Our approach is to use real-time eye tracking for improving interaction between learners and software character agents.

3 Education Interface by Real-time Eye Tracking

The functions of an ESA can be divided into those involving explicit or implicit outputs from the user, and those involving inputs to the user. In case of outputs from the user, empathy involves functions such as monitoring emotions and interest. In terms of output to the user, empathy involves showing appropriate emotions and providing appropriate feedback concerning the agent's understanding of the users' interests and emotions. Real time feedback from eye movement is detected by eye tracking, and the character agents use this information to interact with learners, exhibiting emotional and social behaviors, as well as providing instructions and guidance to learning content. Information about the learner's past behavior and interests based from their eye tracking data is also available to the agent and supplements the types of feedback and input.

Analyzing eye movements provides information on how learners look at images, principally in terms of where they look and how long they spend looking at different objects. Changes in eye position can also be used to infer head movements. In our system, we use eye gaze information for both real time feedback and non real-time collection of eye-tracking statistics. Figure 2 shows an example of the learner's eye position by red dots in the screen.

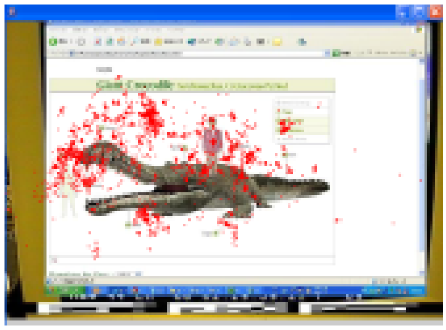

Figure 2. Eye Movements in Eye-aware System

3.1 Real-Time Eye Gaze Interaction

Figure 3 shows how the character agent reacts to feedback about the learner's status based on eye-tracking information. In this example, the eye tracker collects eye gaze information and the system then infers what the learner is currently attending to. This information is then combined with the learner's activity records, and an appropriate pre-set strategy is selected. The character agent then provides feedback to the learner, tailoring the instructions and emotions (e.g., facial expressions) to the situation.

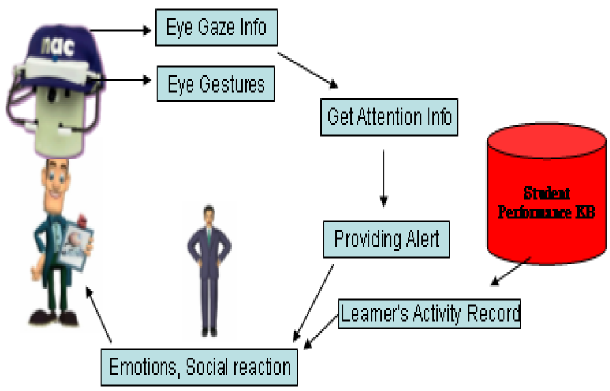

Figure 3 Real-Time Use of Eye Information in ESA

We also show some examples of the functions using real time eye tracking in the system.

3.1.1 Topic Choosing

Information about where learners are looking can be used to infer interest and the current focus of attention. For instance, in Figure 4 there are three pictures or panels and the red circle indicates that the learner is looking at the middle one. If the learner "gazes" at that panel, which indicated by the eye dwell time being longer than a chosen threshold, then the system will treat it as a selection input and will zoom in and give more details on the associated topic.

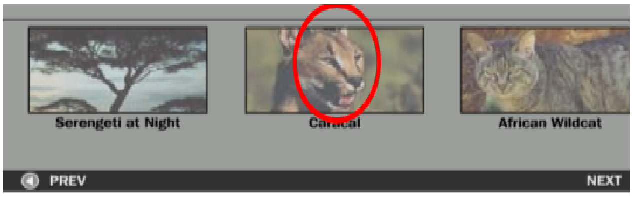

Figure 4 Topic Choosing Interface

3.1.2 Awareness of learners

Human teachers are able to read body language and understand the psychology of the learner to some extent. E-Learning agents may also benefit from an awareness of user states and the ability to customize and adapt instruction accordingly. For instance, an agent may be "eye-aware", using information about eye movements, pupil dilation, and changes in overall eye position (head movements) to make inferences about the state of the learner. After determining the learner's eye position information and current area of interest or concentration, the agents can move around to highlight the current learning topic, in order to attract or focus the learner's attention. For instance, with eye gaze data, agents react to Eye information in real time through actions such as moving to the place being looked at or by showing the detailed information content for where learners are looking at, etc.

3.1.3 Motivation

The character agent can provide motivation as well as feedback or instruction. For instance, it can remind learners to concentrate on a topic if they keep looking away from the screen.

When learners show interest in the current content the character agents provide positive reactions. However, if a learner shows less interest which is indicated by a smaller amount of pupil dilation in the eyes or less activity in the eye movement, then the interface will inquire if the learner is tired or bored. Boredom or fatigue may also be inferred based on lower overall activity in terms of mouse clicks and key selections. Providing rest time or changing topics are other strategies that character agents can use to deal with situations where the learner appears to be bored or fatigued.

3.2 Eye Input

Eye information is also used as an input method in ESA.

3.2.1 Eye gestures

When learners want to give "yes" or "no" responses, or ask the questions, they can use gestures of eyes/head. If the position of the eyes (head) moves up and down this is taken to indicate a "yes" response, while moving left to right indicates a "no" response. When the head moves in a clockwise rotation, a questioning or uncertain response is inferred.

3.2.2 Functions with Eye Gesture

When learners want to select or group the selections, they can use gestures of selection and grouping. By looking at the same place for more than a certain time, the point being gazed at is selected. The group function is used for selecting items of a category. First, the learner selects the Grouping mode and the targeted group, then s/he uses pointing gestures to select the items to put in the target group. This feature of ESA can be used to group learning contents of interest.

3.3 Eye-aware Character Agent

In our system, one or more character agents interact with learners using synthetic speech and visual gestures. The character agents can adjust their behavior in response to learner requests and, in some cases, inferred learner needs. The character agents perform several functions/behaviors including the display of different types of emotion. The agent's emotional response depends on the learner's performance. For instance, an agent shows a happy/satisfied emotion if the learner concentrates on the current study topic from their eye gaze information. In contrast, if the learner seems to lose concentration, the agent will show mild anger or alert the learner. The agent also shows empathy when the learner is stuck or gives a wrong answer. In general, the character agent interacts between the educational content and the learner. Other tasks of a character agent include explaining the study material and provide hints when necessary, moving around the screen to get or direct user attention, and to highlight information.

The character agents are "eye-aware" because they use eye movements, pupil dilation, and changes in overall eye position to make inferences about the state of the learner and to guide his behavior. After getting learner's eye position information and current area of interest or concentration, the agents can move around to highlight the current learning topic, to attract or focus the learner's attention. For instance, with eye gaze data, agents react to the eye information in real time through actions such as moving to the place being looked at, or by showing the detailed information content for where learners are looking at, etc. ESA can also accommodate multimodal input from the user, including text input, voice input and eye information input, e.g., choosing a hypertext link by gazing at a corresponding point of the screen for longer than a threshold amount of time.

ESA is currently targeted to Biology Online Education (Figure 5) and English Learning contents. The educational content within the system uses multimedia such as figures, flash animations, and video clips.

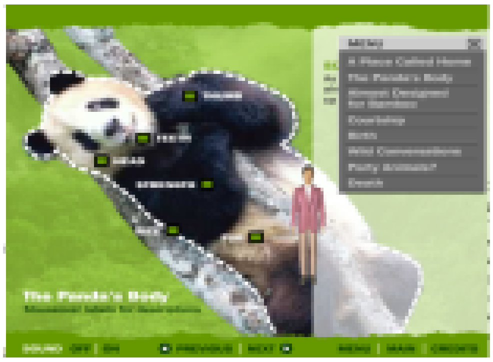

Figure 5 Biology Content Interface

In our approach, different agents have different roles in the interface. When multiple characters are used they have different roles in the interface and interact with each other. There are primary agent (interaction agent) and secondary agent (student monitoring agent). The primary agent talks with students and explains educational contents to users (Figure 6). The secondary agent provides system-related information with the eye tracking information from learners. It manages the interaction based on if the learner is concentrating on the current topic and looking at the right content, or if the learner appears to be interested or bored by the current topic, etc.

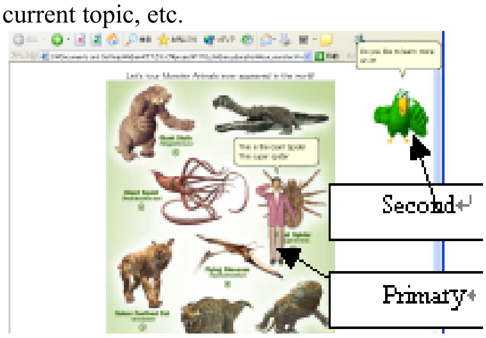

Figure 6 Topics Interface

4 Implementation

For the character agents in ESA, we used the Microsoft Agent Engine to create the characters and also the emotions, movements and control functions, using the MPML agent control language [14]. Other approaches might also have been used. For instance, Huang, et al. [15] developed a character agent that interacted with people while they were browsing Web pages. Their system was constructed mainly from Java, JavaScript and XML. The eye gaze data is stored and transferred using an XML file. Here the serial Data is transferred to an XML file using the Com port and the information is then sent to the interface using JavaScript. During this process, the eye data is mapped to the screen position.

ESA uses a two-dimensional graphical window to display character agents and education content. The graphical window interface shows the education content, flash animations, movie clips, and agent behaviors. The Eye Marker eye tracking system was used to detect the eye information and the basic data was collected using 2 cameras facing towards the eye (Figure 7). The integration of eye tracking with other sensors in the system is shown in Figure 8. We analyzed the recorded eye tracking data with character agent reacting to Eye Information versus an implementation of the interface that did not provide eye tracking feedback.

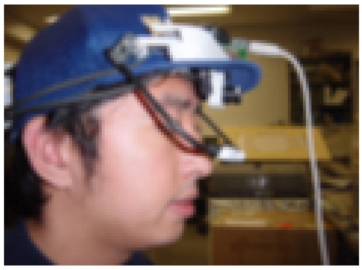

Figure 7 Eye Tracker in the System

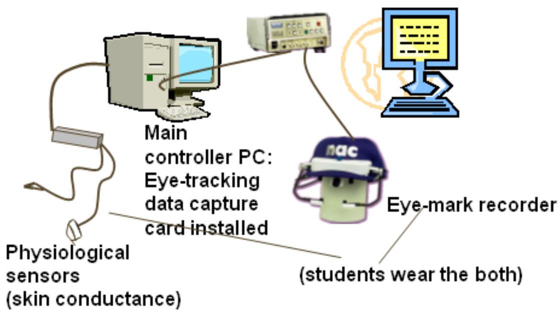

Figure 8. Experiment Setup

5 Preliminary Usability Testing

10 college students participated in a usability study of ESA

with each participant using two versions of the ESA (the character agent versions with and without real-time eye-tracking). After the calibration process required for eye tracking, the participants started using the interfaces. Each learning session lasted about 30 minutes. After the session, the subjects answered questionnaires and commented on the system.

We examined the pattern of eye movements with and without the eye tracking feedback. We measured the numbers and locations of the eye fixations made by the learner and compared the resulting data between the two interfaces. In the agent with eye tracking interface, the eye positions tended to be more focused on current topics. While the small sample in this study precluded tests of statistical significance since our focus was on prototyping and initial user testing of the system, we plan to evaluate these possible effects in a subsequent larger study. Subjective feedback from the participants in the current study also indicated that they had stronger motivation when using character agent with eye tracking interaction.

We also investigated the areas where participants looked at, along with their eye traces (Figure 9). The figure shows the areas that people paid attention to. We calculated the number of eye gazes in each area and how the eye trace moved. Of particular interest were eye movement patterns where the learner's eye traces went to the character agent and then back to the learning content. This type of interaction showed that the learners could receive feedback from the character agent and could then return to the current topic without losing track of where they were. Analysis of the data showed that the eye position moves with the focus of the current topic fairly well showing that the learners were concentrating on the content being taught.

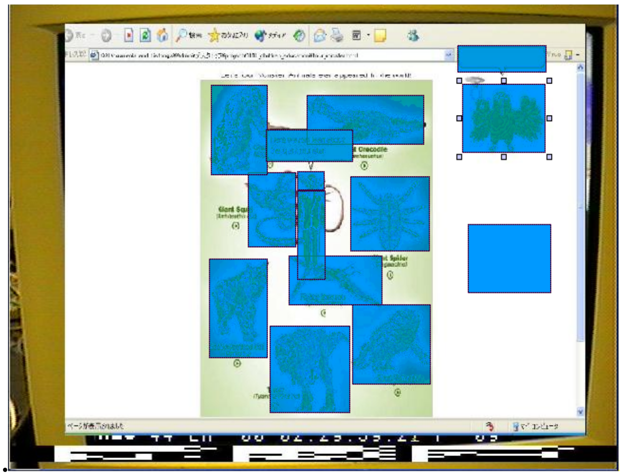

Figure 9 Eye Traces on Different Areas

We analyzed the questionnaires and comments from the subjects, Participants felt that the feedback from the character agent made them more involved in the learning process. They indicated that the information about their eye positions, made them better able to concentrate on the current topic. They also indicated that they found the character agent helpful in learning how to use the interface.

Two participants indicated that they did not like the warning feedback provided when they appeared to be showing low levels of interest. They also felt that the character agent sometimes provided too much information at once. They suggested that multiple agents may be helpful to solve this problem where each agent deals with different types of information and only one agent is speaking/active at any time. Participants in this initial study said that they found the character agents useful and that they listened to the explanation of contents from the agents more carefully than if they had been reading the contents without the supervision and assistance of the character agent.

6 Discussions and Future Work

By getting information about learner response (such as eye movement data and bio-signals data), character interfaces such as ESA can interact with the learner more efficiently and provide appropriate feedback. From preliminary assessment of usability, ESA had a beneficial effect on learner motivation and concentration during learning. This result suggests that there may be a larger role for empathic tutor agents to play in acting as guides to learning content. Such agents can provide important aspects of social interaction when the student is independently working with e-learning content. This type of agent-based interaction can then supplement the beneficial social interactions that occur with human teachers, tutors, and fellow students within a learning community.

During the preliminary trials, we found that learners tended to pay attention to the explanation from character agents and found feedback about their own eye movements useful. They felt that they looked at the contents more attentively than without the eye tracking feedback and this feedback made them more engaged with learning contents.

The current ESA system has some limitations. It is not convenient to carry eye trackers and some noise in the eye tracking data still exists. However, we believe that as eye tracking technology improves, it will become more convenient to use and will provide increasingly precise feedback. Aside from an explicitly educational context, real-time eye gaze interaction can be used in Web navigation. By getting what part of users are more interested in, the system can provide real time feedback to users and help them to get target information more smoothly.

Promising areas for future study include improvements to the collection of eye information and bio-signals, and analysis of video data recorded during online learning sessions to search for correlations between the video and eye tracking data. The use of multiple character agents within empathic tutoring systems represents another interesting direction for this type of research.

Acknowledgment

We would like to thank Dr. Alan Kay for his advice concerning this research.

Reference

[1] Palloff, R. M. and Pratt, K. Lessons from the cyberspace classroom: The realities of online teaching. San Francisco: Jossey-Bass, 2001.

[2] Cooper B, Brna P, Martins A. Effective Affective in Intelligent Systems – Building on Evidence of Empathy in Teaching and Learning. In: Affective Interactions: Towards a New Generation of Computer Interfaces. Paiva A ed. London: Springer Verlag. 2000

[3] Otteson, J. P., and Otteson, C. R. Effect of teacher's gaze on children's story recall. Perceptual and Motor Skills, 50, 35-42, 1979

[4] Sherwood, J. V. Facilitative effects of gaze upon learning. Perceptual and Motor Skills, 64, 1275-1278, 1987.

[5] Graham, S. Communicating sympathy and anger to black and white children: The cognitive (attributional) consequences of affective cues. Journal of Personalitv and Social Psvchology. 47(1), 40-54, 1984.

[6] Klein, J., Moon, Y., and Picard, R. This computer responds to learner frustration: Theory, design, and results. Interacting with Computers, pp 119-140, 2002.

[7] Ahn, H. and Picard, R. Affective-Cognitive Learning and Decision Making: A Motivational Reward Framework For Affective Agent, The 1st International Conference on Affective Computing and Intelligent Interaction. October 22-24, 2005, Beijing, China.

[8] Qu, L., Wang, N., John, W.L., "Choosing when to interact with learners", Proceedings of the 9th international conference on Intelligent user interface table of contents,Funchal, Madeira, Portugal , Short Papers table of contents, pp 307-309, 2004.

[9] Fabri, M., D. Moore, and D. Hobbs, "Mediating the Expression of Emotion in Educational Collaborative Virtual Environments: An Experimental Study". Virtual Reality Journal, 2004

[10] Stone, B., and Lester, J., Dynamically Sequencing an Animated Pedagogical Agent, Proceedings of the 13th National Conference on Artificial Intelligence, pp. 424-431, Portland, OR, August, 1996.

[11] Duchowski, T.,. Eye Tracking Methodology: Theory and Practice. Springer, London, UK, 2003.

[12] Johnson, W.L., Marsella, S., Mote, H., Vilhjalmsson, S., Narayanan , S. and Choi, S. , Language Training System: Supporting the Rapid Acquisition of Foreign Language and Cultural Skills.

[13] Faraday, P., and Sutclie, A., An empirical study of attending and comprehending multimedia presentations. In Proceedings of ACM Multimedia, pp. 265–275, Boston MA, 1996.

[14] Prendinger, H., Descamps. S., and Ishizuka, M.,. MPML: A markup language for controlling the behavior of life-like characters. Journal of Visual Languages.

[15] Huang, Z., Elis, A., and Visser, C., Facial Expressions for Embodied Agents in STEP, Proceedings of AAMAS 2004 Workshop on Embodied Conversational Agents: Balanced Perception and Action, 2004.