Design and evaluation of an adaptive autonomous help agent

JUDITH MASTHOFF
University of Brighton
Watts building, Moulsecoomb
Brighton BN2 4GJ, UK
Judith.Masthoff@brighton.ac.uk

BACKGROUND

Between 1992 and 1997, an *artificial teacher* was developed based on a multi-agent architecture (Masthoff, 1997). One of the agents was the Help Agent, which adapted the content and timing of help to the individual student. This agent was based on layered explanations, inspired by the work of Wood, Wood and Middleton (1978) on layered interventions, and the work of Taylor (1988) on layered protocols in communication. It made exercises gradually easier by providing hints. The agent took the initiative to provide help when the student seemed unable to continue (in the prototype, this was decided on the basis of the time during which the student had not taken an action).

The need for help is not restricted to the traditional educational domain. Users need more and more help on how to operate systems (like phones), as the amount of functionality increases rapidly, and the user group becomes increasingly diverse. It is not possible to make a clear distinction between novice users and expert users, as the level of expertise of users can vary per functionality. So, help needs to be *adapted* to the expertise level of the individual user on their current task. In 1999, a prototype system was built for the Radiology domain, which explained its own operation, adapting its explanation to the individual user.

THE PROTOTYPE SYSTEM

The prototype system has a touch screen and voice control, but the latter was disabled in the experiment. The following interactions were relevant for the evaluation:

- Brightness and Contrast of radiological images can be set using sliders. Touching a slider puts the system in brightness-contrast mode. In this mode, moving the pen on the image changes brightness and contrast simultaneously: a horizontal movement changes the brightness, a vertical movement the contrast. We call this “dual control”.
- Zooming occurs by pressing the Zoom button, and panning by dragging the image. After zooming in, the image is zoomed out by pressing the Zoom button again.
- Individual shutters can be dragged to their desired location. Adjacent shutters can be dragged simultaneously by dragging the corner where they overlap. We call this the “corner method”.

Explanations are given at the moment when the user touched a certain function or completed an action related to that function. For instance, when the user touches the brightness slider, or releases the pen after having touched or dragged the brightness slider, an explanation related to brightness could be given. Explanations are only given if they have not been given before and if the user has not already performed the action to which the explanation is related. Explanations can be interrupted at any time, by performing any action.

The system remembers what parts of the explanation have not yet finished, and repeats those parts when a new occasion occurs. Explanations are given in audio, with a visual presentation of the Help agent and text appearing at the bottom of the screen. The text is synchronised with the audio. For some explanations, examples are given with animations of a moving finger, for instance, moving over the image to change brightness and contrast simultaneously.

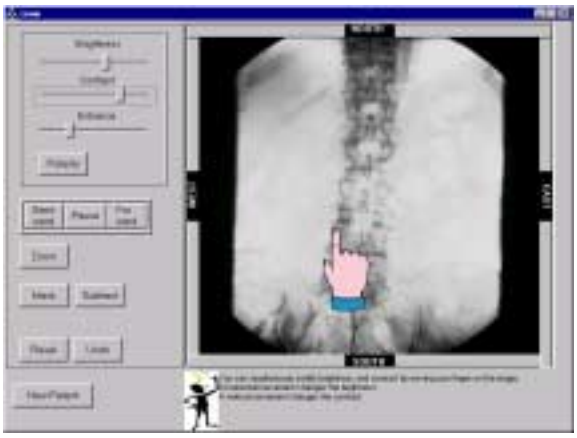

Figure 1. An explanation (bottom of screen) and example (finger on image).

The following explanations are relevant to the experimental tasks:

- *Brightness-Contrast*. Explanation parts: (1) How to use a slider, (2) How to drag a slider indicator, (3) How to use dual control, (4) Example of dual control. The dual control explanations are only given after the user has used both brightness and contrast sliders.
- *Zoom-Pan*. Explanation parts: (1) How to pan, (2) How to zoom out.
- *Shutters*. Explanation parts: (1) How to set an individual shutter, (2) How to drag a shutter, (3) Example of setting an individual shutter, (4) How to use the corner method. The latter is only given after the user has set two adjacent shutters individually.

EVALUATION

We have performed a first evaluation of the effectiveness of the explanations. A between-subjects design was used: in group I the system provided explanations and examples, in control group II the system did not. Sixteen subjects participated in the experiment. All subjects used computers in their jobs, but had no experience with touch screens. Group I was told that explanations could be given by the system, but not when and in what form.

Both groups were given the same tasks. In all tasks, the subject was asked to make an image on the system look like a target image on paper. Three types of tasks were used:

- *Brightness-Contrast* tasks (B), in which the subject was asked to change the brightness and contrast to match a target image,
- *Zoom-Pan* tasks (Z), in which the subject was asked to zoom the image in, and pan it to match a target image, and
- *Shutter* tasks (S), in which the subject was asked to position the shutters so as to match a target image.

Of each type of task, three different instances were used. The tasks were given in random order: B Z S Z B S B S Z. We measured how long it took subjects to complete the tasks, how many mistakes they made, how long it took them to discover system functionality, and their satisfaction.

- *Brightness-contrast.* Sliders were used immediately by all subjects. Therefore, no explanations were given on their use. Five subjects in group I used the dual controls compared to nobody in group II. The dual-controls did not prove to be a more efficient interaction, though.
- *Zoom-Pan.* In the first task instance, all subjects in group I performed the first step of panning (touching the image) within three actions after having zoomed in, compared to only one subject in group II. The others took at least nine actions, with two subjects never touching the image at all. Two subjects in group I misinterpreted the explanation on the second panning step and did not finish the task, compared to three in group II. The average time in group I was more than 20 seconds faster (1:06 versus 1:27). This difference was not statistically significant, but this is not surprising given the small number of subjects who completed the task. Even in the last task instance, two subjects in group II still assumed shutters played a role in panning, and were interleaving shutter presses with panning operations. They provide a nice example of how users can unknowingly cope with tasks in an unintended and inefficient way.
- *Shutters*. Unfortunately, all subjects in group II had already used shutters while trying to pan. Also, nobody in group I was explained how to set individual shutters, because tiny movements when touching the shutters were interpreted by the system as setting the shutters. The system should be modified to use a threshold. Half of group I used the corner method, compared to nobody in group II. Two subjects learned from the corner method explanation how to set individual shutters.

CONCLUSIONS

There is evidence that explanations help users to discover and use interactions. In particular, it helped the subjects to discover panning, dual control, and the corner method. In general, the subjects felt in control of the system. User control should be further improved by enabling the user to explicitly ask for explanations, and to repeat them.

The timing of the explanations is good, but can still be improved. Some time delay may be needed between two parts of an explanation that explain, respectively, a follow-up action (like panning) and an undo action (like zooming out). A repetition or elaboration of the explanation should be given when the user (1) performs the start of an interaction without completing it (like touching the image but not dragging it), or (2) performs a resultless action (like moving a shutter outwards when at its maximum position). To avoid unnecessary repetition, explanations should be broken down into short parts (the dual-control part was too long). It should be investigated whether repetition is still needed when an explanation part is interrupted *near* the end (as often happens), and how "near" should be defined in that case.

The use of a combination of audio and text seems good, as audio is more attention grabbing, while text is less transient. An indication (perhaps an audio signal) should be given when an explanation is about to start. An indication might also be given of the amount of explanation still to follow. As subtle reminders, text-only explanations could be provided when the explanation has already been given in audio.

REFERENCES

Masthoff, J. (1997). *An agent-based interactive instruction system*. Eindhoven University of Technology, Eindhoven.

Taylor, M. (1988). Layered protocols for computer-human dialogue I: Principles. *International journal of man-machine studies, 28*, 175-218.

Wood, D., Wood, H., & Middleton, D. (1978). An experimental evaluation of four face to face teaching strategies. *International journal of behavioral development, 1*, 131-147.