The PPP persona: a multipurpose animated presentation agent

Elisabeth André, Jochen Müller, Thomas Rist

The PPP Persona: A Multipurpose Animated Presentation Agent

Elisabeth André, Jochen Müller, Thomas Rist

German Research Center for Artificial Intelligence (DFKI)

Stuhlsatzenhausweg 3, D-66123 Saarbrücken, Germany

Email: {andre,jmueller,rist}@dfki.uni-sb.de

Abstract

Animated agents – either based on real video, cartoon-style drawings or even model-based 3D graphics are likely to become integral parts of future user interfaces. We present the *PPP Persona*, a tool which can be used for showing, explaining, and verbally commenting textual and graphical output on a window-based interface. The realization of the module follows the client/server paradigm, i.e, some client applications can send requests for executing presentation tasks to the server. However, to achieve a lively and appealing behaviour of the animated agent, the server autonomously performs some actions, eg. to span pauses or to react immediately to user interactions.

Introduction

Through the last decade, the design of user interfaces has undergone dramatic changes due to the introduction of multi-windowing and the exploitation of multimedia technology. In fact, most of today's application systems come with window-based interfaces and many of them already communicate computation results and retrieved data using a mixture of different media. The next major step in the evolution of interfaces is very likely to focus on highly personalized interfaces. On the one hand, personalization refers to the ability of a system to design multimedia presentations on the fly in order to meet the information needs of individual users in particular situations. This means that decisions on content selection, media choice, and medium-specific encoding of information are made with regard to models of the user, task, and situation. While this kind of personalization is widely addressed in the area of user modelling and in research on intelligent presentation systems, less attention has surprisingly enough been paid so far to the *peripheral* aspect of personalizing user interfaces. By peripheral personalization we mean that the system personifies itself audio-visually, e.g. as in our case, by means of an animated character called *PPP Persona*.

There are several motivations for using such an animated presentation agent in the interface. Most importantly in our working context - which is the development of automated presentation systems (WIP cf. [Wahlster et al. 93] and PPP cf. [André et al. 93]) - an animated character adds expressive power to a system's presentation skills.

For example, in one of our application domains, PPP persona effectively advertises accommodation offers found on the WWW.
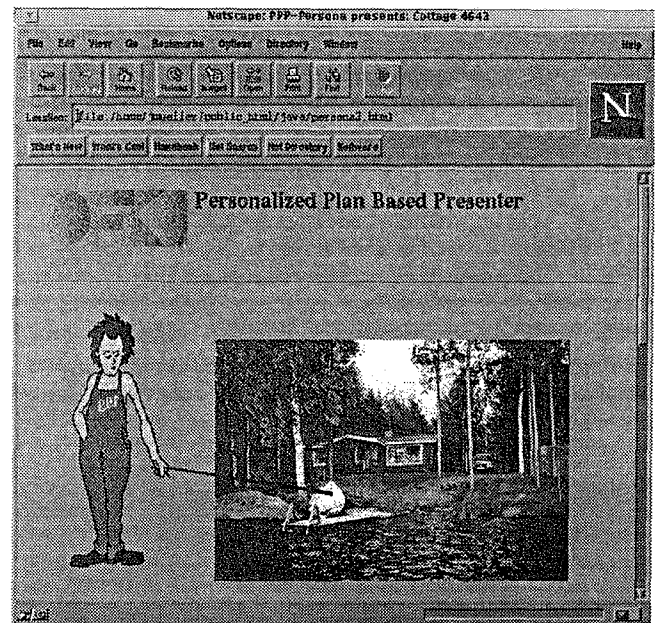

Figure 1: PPP Persona presents retrieval results from the web using the Netscape Navigator^{TM} and Java^{TM}

Suppose the user is planning to spend holidays in Finland and is therefore looking for a lake-side cottage. To comply with the user's request, the system retrieves a matching offer from the web and creates a presentation script for the PPP persona which is then sent to the presentation viewer (e.g. Netscape Navigator^{TM} including a Java^{TM} interpreter). When viewing the presentation PPP Persona highlights the fact that the cottage has private access to a lake by means of a *verbal annotation of a picture*; i.e., Persona points to the picture

during a verbal utterance (cf. Fig. 1). A further example is the establishment of cross-references between different media (possibly occurring in different windows) through a two-handed pointing gesture.
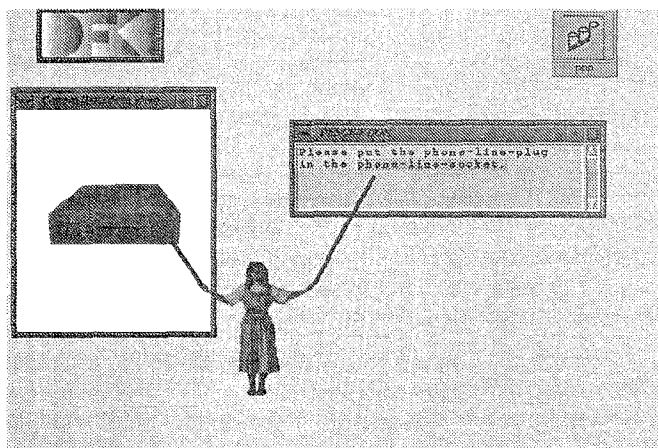

Figure 2: Establishing cross-media/window references with a "real" Persona

The screen shot shown in Fig. 2 is taken from another application domain which deals with the automated generation of instructions for technical devices such as a modem. To explain how to connect a certain modem to the phone-line Persona uses two pointing sticks to refer to the graphical depiction of the phoneline socket in the left window and its textual presentation in the right window. The screen shot also shows that the appearence of the Persona is not restricted to cartoon characters only. This time, the presentation system personifies itself as a "real" person (namely the paper's first author) composed of grabbed video material.
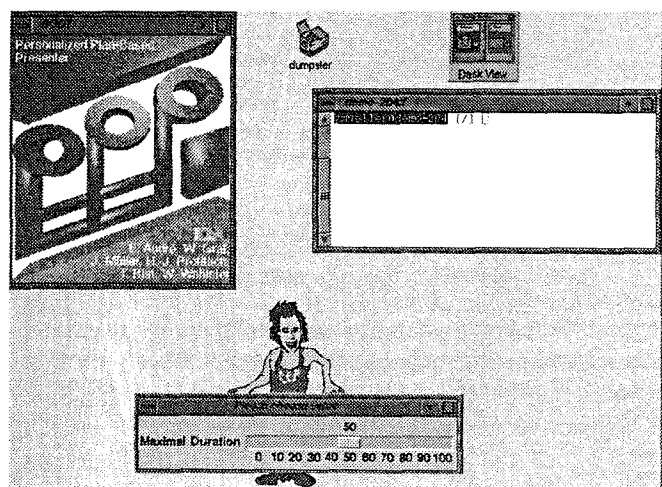

Figure 3: Requesting user input

While in the previous examples, the Persona's activities are devoted to the presentation of multimedia material, it can also take up the role of a full-fledged interface agent with which the user can communicate. The PPP system makes use of this possibility too, although in a rather simplified way. The user can directly click on the Persona to obtain the system's main menu. In this way, the user can enter a new goal to be achieved by the system, or to interrupt and "criticize" an ongoing presentation by changing generation parameters such as the degree of detail, or the preference for a certain medium. However, the Persona offers also the possibility to guide the user through input menus. Instead of presenting different input devices (i.e. text boxes, sliders, buttons, etc.) in a formsheet-style menu, the Persona can request the single values one after another giving verbal explanations on each option. The screen shot in Fig. 3 illustrates a situation where the Persona asks the user to input a numeric value.

Conception of the PPP Persona Server

The realization of the PPP Persona Server follows the client/ server paradigm, i.e, some client applications can send requests for executing presentation tasks to the server. However, there are several requirements that an animated presentation agent has to meet. According to its functional roles in a presentation (e.g. as in the examples above), the Persona must be conversant with a broad variety of presentation gestures and rhetorical body postures. Furthermore, it should adopt a reasonable and lively behaviour without being distracting. From the technical point of view, the piece of software that realizes the Persona should be highly independent of application and platform. While the next section gives a classification of the actions the Persona can perform, we refer to the extended version of this paper (cf. [André et al. 96]) for a description of the technical realization of the PPP Persona server.

Classification of Persona Actions

Since we aim at a generic presentation agent, we are most interested in domain independent actions the Persona should perform. The current repository of such general actions comprises the following types:

- *High-level presentation acts* This group of actions includes pointing gestures. A pointing gesture is described as a tuple of the form: (⟨pointing-device⟩, ⟨target-of-pointing⟩, ⟨pointing-shape⟩). The Persona's hands/arms and sticks are typical pointing devices. The target refers to the particular entity (text element, graphical object, window etc.) to which the Persona points. Punctual pointing, underscoring, and encircling are possible pointing shapes. Further high-level presentation acts are speaking and the expression of emotions, e.g. getting angry or tired. We assume that all high-level presentation acts are delivered by superordinate components such as a presentation planner in the PPP system.

- *Idle-time acts* To achieve a lively and "natural" behaviour of the Persona, it even "stays alive" in an idle phase. Typical acts to span pauses are breathing, thumb twiddling and so on. In contrast to high-level presentation acts, idle-time acts are part of a self-monitored, basic behaviour of the Persona.

- *Reactive Behaviours* In case of an interactive system, the Persona should be able to react to some user interactions immediately. For example, if the user moves the window to which the Persona is currently pointing, the consistency of the pointing gesture has to be restored as soon as possible, e.g. by a prolongation of the pointing stick or by moving the Persona to a new position.

- *Low-level navigation acts* This class of actions has been introduced for both, ontological and economical reasons. So it is quite natural to separate a pointing gesture from a preceding navigation act (e.g. moving to a certain position specified relative to a window) which enables the execution of the pointing gesture. The economical motivation for such decompositions lies in the fact that common subactions can be factored out.

- *Basic postures/acts* Any action of one of the types mentioned above is characterized as being basic if it cannot be decomposed into less complex subactions. Technically speaking, a basic posture corresponds either to a single frame of the Persona or to an uninterruptable sequence of several frames.

As in many other ontologies of actions, the concept of *action decomposition* is complemented by the concept of *specialization* (see Fig. 4).
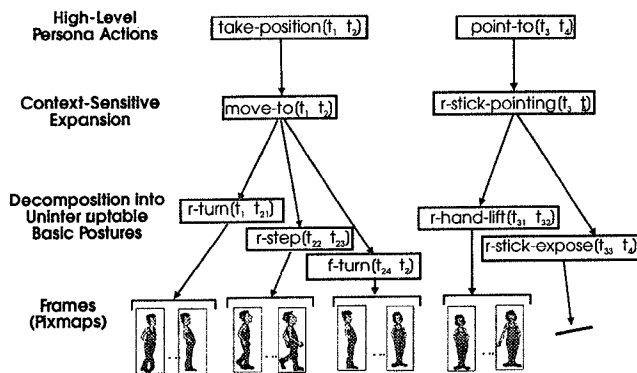

Figure 4: Action types the PPP-Persona can perform

For example, the acts "moving-to-the-right-of-window," and "moving-to-the-left-of-window" are two different specializations of the more general act "take-position-for-pointing" which in turn is part of the decomposition of a superordinate "point-to" act. In case an action has several specializations, the current context is taken into account when making a choice. Note that we associate with each action a time interval within which the action takes place. For example, the act *take-position* has to be executed during $(t_1\ t_2)$.

Conclusion

We sketched the PPP Persona server, a component that monitors the behaviour of an animated presentation agent. By means of examples, we illustrated how to take advantage of the expressive power which is added to the user interface by such an agent. While animated user interface agents have been proposed by several other authors (cf. [Laurel 90; Bos et al. 94; Gibbs et al. 93]), a distinguishing new feature of our approach is that the agent not only performs tasks requested by some application clients, but also implements a basic behaviour independent of the applications it serves. This basic behaviour comprises *idle-time actions*, and *immediate reactions* to events occurring at the user interface. Both action types have to be supported, in order to obtain a lively and appealing presentation agent. Due to a built-in mechanism for action specialization and decomposition application clients can request presentation tasks at a high-level of abstraction. For example, the server itself can find reasonable positions from which pointing gestures are executed.

While the automated hypermedia presentation system PPP currently serves as a testbed for the Persona server, the component is a highly portable tool and could be used as well for various other applications including teleshopping, home banking, and tele-communication. For example, the Persona can take the role of a synthesized shop assistant or investment consultant.

References

[André et al. 93] E. **André**, W. **Graf**, J. **Heinsohn**, B. **Nebel**, H.-J. **Profitlich**, T. **Rist**, and W. **Wahlster**. *PPP - Personalized Plan-Based Presenter*. Document D-93-5, DFKI, Saarbrücken, 1993.

[André et al. 96] E. **André**, J. **Müller**, and T. **Rist**. *The PPP Persona: A Multipurpose Animated Presentation Agent*. Report, DFKI, Saarbrücken, 1996. Extended Version.

[Bos et al. 94] E. **Bos**, C. **Huls**, and W. **Claassen**. *EDWARD: Full Integration of Language and Action in a Multimodal User Interface*. International Journal of Human-Computer Studies, 40:473–495, 1994.

[Gibbs et al. 93] S. **Gibbs**, C. **Breiteneder**, V. **de Mey**, and M. **Papathomas**. *Video Widgets and Video Actors*. In: UIST, pp. 179–184, 1993.

[Laurel 90] B. **Laurel**. *The Art of Human-Computer Interface Design*. Addison-Wesley, 1990.

[Wahlster et al. 93] W. **Wahlster**, E. **André**, W. **Finkler**, H.-J. **Profitlich**, and T. **Rist**. *Plan-Based Integration of Natural Language and Graphics Generation*. AI Journal, 63:387–427, 1993.