WIP/PPP: Automatic Generation of Personalized Multimedia Presentations

Elisabeth André, Jochen Müller, Thomas Rist
DFKI GmbH, Stuhlsatzenhausweg 3,
D-66123 Saarbrücken, andre@dfki.uni-sb.de

MOTIVATION

For an increasing number of applications the rnarrual creation of multimediahypermedia presentations is no longer feasible. In order to meet the specific needs of the individual presentation consumer, information has to be communicated fast and flexibly. Although powerful authoring tools are available, they are not likely to meet these requirements since they only facilitate the presentation editing. Over the past years, we have developed methods for fully automated presentation design. In the system demonstration described here, we illustrate our approach by means of two prototype systems, The first system, WIP [1], generates multimedia instructions for the maintenance of technical devices. The benefit of WIP lies in its ability to produce presentations depending on a given set of presentation parameters. The more advanced system PPP [2, 3] additionally addresses the temporal coordination of different media as it comes into play when using a life-like character to present multimedia material.

APPROACH

In order to build up multimedia presentations automatically, we have developed principles for describing the structure of coherent text-picture combinations. Essentially, these principles are based on (a) the generalization of speech act theory to the broader context of communication with multiple media, and (b) the extension of RST (Rhetorical Structure Theory) to capture relations that occur not only between presentation parts realized within a particular medium but also those between parts conveyed by different media.

Since we consider the generation of multimedia presentation as a goal-directed activity, it seemed most appropriate to implement a goal-driven, top-down planning approach. This presentation planning process is illustrated in Fig. 1, The planning component receives as input a communicative goal (for instance, the user should be able to localize the internal parts of a modem) and a set of generation parameters, such as target group, presentation objective, resource limitations, and target language. The task of the component is to select parts of a knowledge base (see left-hand side of Fig. 1) and to transform them into a multimedia presentation structure (see right-hand side of Fig. 1). Whereas the root node of such a presentation structure corresponds to a more or less complex communicative goal, such as describing a technical device, the leaf nodes are elementary retrieval or generation acts, currently for text, graphics, animations and gestures.

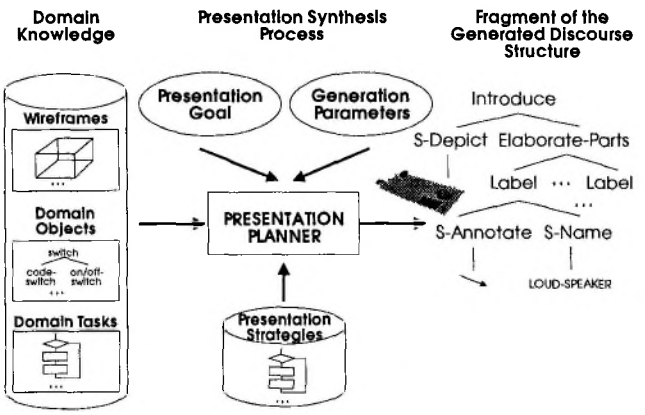

Figure 1: From a Domain Plan to a Multimedia Discourse Structure

Design knowledge is represented by so-called presentation strategies which encode knowledge on: (a) how to select relevant content, (b) how to structure selected content, and finally (c) which medium to use for conveying a content.

To enhance the effectivity of computer-based communication, we embarked on a new project, called PPP, in which a life-like character, the so-called PPP Persona, acts as a presenter, showing, explaining, and verbally commenting textual and graphical output on a window-based interface. That is PPP has to design multimedia material, and to plan presentation acts and their temporal coordination.

In order to cope with the dynamic nature of presentations made by an animated agent, several extensions became necessary. These include the distinction between *production* and *presentation acts*, and the development of a mechanism for propagating temporal constraints and building up schedules

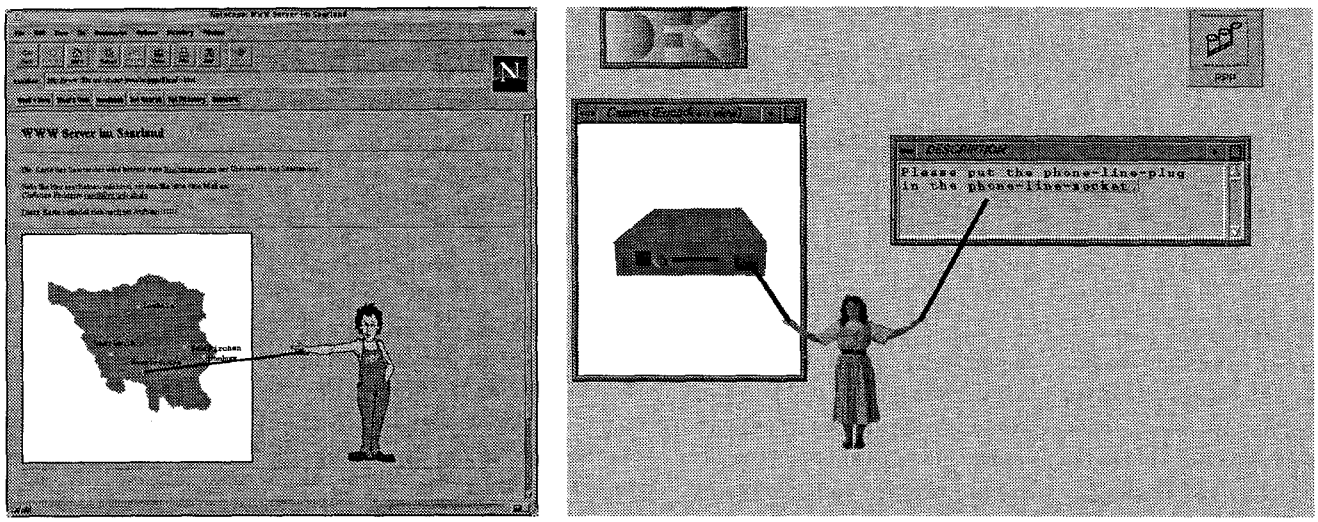

Figure 2: PPP Examples

to be executed by the PPP persona.

Whereas production acts refer to the creation/retrieval of material, presentation acts are display acts, such as *Display-Text*, or acts which are carried out by the PPP Persona, e.g. *Point*. To run the presentation in an intelligible way, the presentation acts need to be temporally coordinated. Thus, we had to allow for the specification of the temporal behavior of a presentation by means of qualitative and metric constraints. Qualitative constraints are represented in an "Allen-style" fashion which allows for the specification of thirteen temporal relationships between two named intervals, e.g. *(Speak1 (During) Point2)*. Quantitative constraints appear as metric (in)equalities, e.g. *(5 $\leq$ Duration Point2)*. When designing a presentation, PPP builds up a temporal constraint network which is checked for consistency by computing numeric ranges on each interval endpoint, the difference between endpoints and the possible Allen relationships between each pair of intervals. Finally, a partial schedule is built up by resolving all disjoint temporal relationships between intervals and computing a total temporal order. Since the temporal behavior of presentation acts may be unpredictable at design time, the schedule will be refined at runtime by adding new metric constraints to the constraint network.

The realization of the Persona Server follows the client/server paradigm; i.e, client applications can send requests for the execution of presentation tasks to the server. However, the Persona Server not only enables the execution of complex presentation acts, but also implements a basic behavior independent of the applications it serves. This basic behavior comprises *idle-time actions*, *immediate reactions* to events occurring in the user interface and *low-level navigation acts*.

GENERATION EXAMPLES

Fig. 2 shows two sample applications. In the screen shot on the left-hand side, the Persona presents predesigned material retrieved from the WWW. It points to the mouse-sensitive regions of a map and describes important sites verbally after following the links associated with the regions and filtering the retrieved data according to the user's interest profile. The screen shot on the right was taken from another application domain, namely the automated generation of instructions for technical devices, in this case a modem. Here, the Persona uses two pointing sticks to refer to a graphical object depiction in the left window and its textual presentation in the right window. The screen shot also shows that the appearance of the Persona is not restricted to cartoon characters only. Similar to changing a text font, the system may personify itself as a "real" person composed of grabbed video material.

TECHNICAL REALIZATION

X11-Implementations of our systems are available for Sun Sparc and Silicon Graphics workstations. The WWW version requires a browser which runs Java applets.

REFERENCES

1. E. André and T. Rist. The Design of Illustrated Documents as a Planning Task. In M. Maybury, editor, *Intelligent Multimedia Interfaces*, pp. 94–116. AAAI Press, 1993.

2. E. André and T. Rist. Coping with temporal constraints in multimedia presentation planning. In *Proc. of AAAI-96*, Vol. 1, pp. 142–147, Portland, Oregon, 1996.

3. T. Rist, E. André, and J. Müller. Adding Animated Presentation Agents to the Interface. In *Proc. of the 1997 International Conference on Intelligent User Interfaces*, Orlando, Florida, 1997. to appear.