Intellimedia: Making Multimedia Usable by Exploiting AI Methods

ELISABETH ANDRÉ

German Research Center for Artificial Intelligence (DFKI), Stuhlsatzenhausweg 3, D-66123 Saarbrücken ⟨andre@dfki.uni-sb.de⟩

Rapid progress in technology for information processing, storing, distribution, and exchange is paving the way for the information society of the next century. It is one thing to have a great potential for producing and accessing vast amounts of information but quite another to make information available to human users in a profitable way; therefore methods for effectively presenting information are becoming more and more crucial in an expanding field of applications. *Intellimedia presentation systems* attempt to incorporate multimedia technology into intelligent presentation systems. Such systems go far beyond current hypermedia/multimedia environments because they incorporate the following:

—effectiveness through the coordinated use of several media and presentation styles;
—adaptivity by automatically generating presentations on the fly;
—interactivity based on a good understanding of the user's goals and plans; and
—personalized communication by animated interface agents.

This paper sketches how we integrated multiple AI methods, such as planning, constraint processing, temporal reasoning, natural processes processing, and knowledge representation, into the presentations systems WIP and PPP.

WIP: AUTOMATED GENERATION OF MULTIMEDIA DOCUMENTS

WIP [Wahlster et al. 1993] is a knowledge-based presentation system that generates a variety of multimedia documents from an input consisting of a formal description of the communicative intent of a planned presentation. The current WIP prototype generates illustrated instructions for assembling, using, maintaining, or repairing physical devices. Figure 1(a) shows an annotated 3D graphic of a circuit board generated by WIP.

The basic assumption behind the WIP project is that the design of multimedia presentations is a planning task [André and Rist 1995a]. As input, WIP receives a presentation goal (e.g., the user should know how to prepare a modem for reception of data). This goal is forwarded to a presentation-planning component that looks for a presentation strategy matching this goal and generates a refinement-style plan in the form of a directed acyclic graph. Although the top of the presentation plan is a more or less complex presentation goal, the lowest level is formed by specifications of elementary presentation tasks (e.g., formulating a request or depicting an object) that are directly forwarded to medium-specific design components, currently a text and a graphics generator.

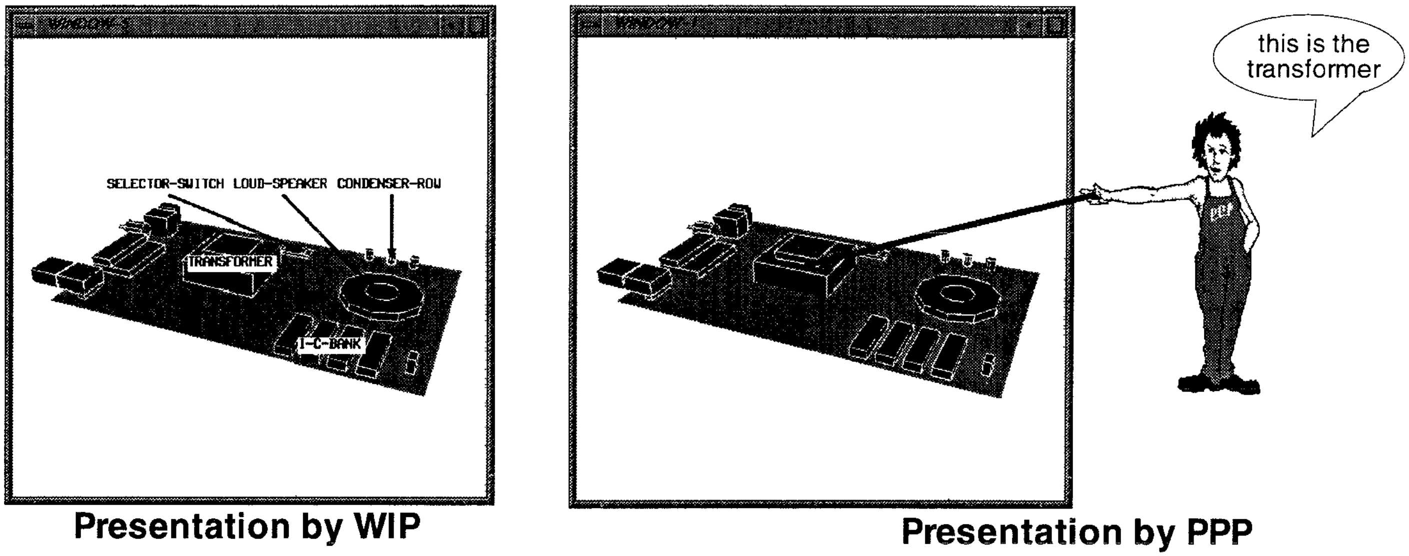

Figure 1. (a) Presentation by WIP; (b) presentation by PPP.

As coherent presentations cannot be generated by simply merging verbalization and visualization results into multimedia output, the processes for content determination, medium selection, and content realization in different media have to be carefully coordinated. A distinguishing feature of our work is that we use a uniform planning mechanism for these subtasks that facilitates their simultaneous coordination.

In contrast to document-retrieval systems, WIP does not use any prestored document fragments, such as predesigned icons or canned text, but generates all parts of a presentation from scratch. As soon as the presentation-planning component has decided which generator should encode a certain piece of information, this piece is passed on to the respective generator. Each generator consists of a design and a realization component.

The main task of the text design component is the organization of input elements into clauses. The results of the text designer are preverbal messages. These preverbal messages are forwarded in a piecemeal fashion to the text realization component where grammatical encoding, linearization, and inflection take place.

To perform elementary pictorial presentation acts, the graphics designer builds up a so-called design plan consisting of sequences of graphical operators. These operators fall into the following three classes: operators for creating and manipulating wireframe models of 3D objects, operators that constrain projection parameters and map wireframe models onto images, and operators that are defined on the picture level. The design plan is then passed onto the graphics-realization module for execution.

Because the results of the different generators should be tailored to each other, each generator has to know how information has been conveyed by other generators. Therefore, each generator provides an explicit representation of its encodings. Through its clear separation of content and form, WIP goes well beyond conventional hypermedia systems.

Presentation fragments provided by the generators have to be arranged in a multimedia output. A purely geometrical treatment of the layout task would, however, lead to unsatisfactory results. Rather, layout must be considered as an important carrier of meaning. To determine effective screen layouts for the generated multimedia material, a constraint-based layout component is used. In our approach, we map relations between presentation parts onto spatial constraints and rely on a finite-domain constraint solver to determine an arrangement of the presentation parts in a way consistent with the structure of the underlying information.

WIP's generation process is controlled by a set of generation parameters such as target group, presentation objective, resource limitations, and target language. The benefit of WIP lies in its ability to present the same information in a variety of ways, depending on the value combination of these generation parameters. This has become possible because WIP postpones all design decisions until run time.

PPP: GENERATION OF PERSONALIZED MULTIMEDIA PRESENTATIONS

The PPP project continues work done in WIP by adding fundamental extensions. First, PPP allows user interaction; that is, PPP responds to follow-up questions concerning the domain as well as to meta-comments on the act of presentation. To make this possible, we have extended automated presentation design for hypermedia and developed a uniform framework for representing the structure of hypermedia discourse. Among other things, this makes possible the coherent continuation of a presentation both for expected and unexpected user interactions [André and Rist 1995b].

Second, we have introduced an animated interface agent that plays the role of a presenter, showing, explaining, and verbally commenting textual and graphical output on a window-based interface (see Figure 1(b)). That is, PPP has to plan presentations as well as presentation acts and their temporal coordination. To do this, we distinguish between the creation of multimedia material and its display and explicitly represent the temporal course of a presentation by qualitative (Show-Picture before Explain-Object) and quantitative ($10 \leq$ Duration Show-Picture ≤ 40) temporal constraints [André et al. 1995].

REFERENCES

André, E. and Rist, T. 1995a. Generating coherent presentations employing textual and visual material. *Artif. Intell. Rev.* (Special volume on the Integration of Natural Language and Vision Processing) *9* (2–3), 147–165.

André, E. and Rist, T. 1995b. Towards a new generation of hypermedia systems: Extending automated presentation design for hypermedia. Document, DFKI, Saarbrücken. (Submitted for publication.)

André, E., Müller, J., and Rist, T. 1995. The PPP persona: A multipurpose animated presentation agent. Document, DFKI, Saarbrücken. (Submitted for publication.)

Wahlster, W., André, E., Finkler, W., Profitlich, H.-J., and Rist, T. 1993. Plan-based integration of natural language and graphics generation. *AI J. 63*, 387–427. Also DFKI Res. Rep. RR-93-02.