WIP: From Multimedia to Intellimedia

Elisabeth Andre, Wolfgang Finkler, Winfried Graf, Karin Harbusch,
Jochen Heinsohn, Anne Kilger, Bernhard Nebel, Hans-Jiirgen Profitlich,
Thomas Rist, Wolfgang Wahlster (Authors of the WIP System),
Andreas Butz (Video Producer), Anthony Jameson (Speaker)
German Research Center for Artificial Intelligence (DFKI)
Stuhlsatzenhausweg 3, D-66123 Saarbriicken
Email: {name}@dfki.uni-sb.de

This video provides an overview of the WIP project. WIP aimed at the development of a presentation system that is able to generate a variety of multimedia documents considering generation parameters, such as target group, presentation objective, resource limitations, and target language (see Fig. 1).

The major components of the WIP system are: a *presentation planner* that is responsible for determining the contents and selecting an appropriate medium combination, *medium-specific generators* and a *layout manager* that arranges the generated output in a document. The video gives a survey of the WIP architecture and demonstrates the performance of the single modules by several system runs.

A basic assumption behind the WIP model is that not only the generation of text and dialog contributions, but also the design and presentation of graphics and multimodal documents are planning tasks. 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 the *presentation planner* which tries to find a presentation strategy which matches this goal and generates a refinement-style plan in the form of a directed acyclic graph (DAG). This DAG reflects the propositional contents of the potential document parts, the intentional goals behind the parts as well as the rhetorical relationships between them. While the top of the presentation plan is a more or less complex presentation goal (e.g., instructing the user in switching on a device), the lowest level is formed by specifications of elementary presentation tasks (e.g., formulating a request or depicting an object). These elementary tasks are directly forwarded to the medium-specific generators, currently for text and graphics.

The design of the *text generator* was strongly influenced by the quest for incremental processing. The first component that is activated during natural language generation is the *text design* component. As soon as the presentation planner decides that a particular element should be presented as part of a text, the element is handed over as input to this component. The main task of the text design component is the organization of input elements into clauses. This comprises the determination of the order in which the given input elements can be realized in the text and lexical choice. 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. The text realization component is based on the formalism of Lexicalized LD/LP Tree Adjoining Grammars.

As the text generator, the *graphics generator* includes components for *design* and *realization.* The task of the graphics design component is to transform presentation tasks received from the presentation planner into a sequence of operators to be executed by the graphics realization component. Basic knowledge about how to accomplish this transformation is represented by so-called design strategies. The realization component can be considered as an extension of an object-oriented graphics editor that handles both two-dimensional concepts and three-dimensional models of objects and object configurations. Thus, the component has to support operators that manipulate models, mappings and pictures. Beside these operators, the functionality of the realization component also encompasses evaluation operators (e.g., to check whether an object as part of an object configuration is visible from a given viewing specification, or to check whether a picture part can be discriminated from other picture components). These evaluators are necessary in order to recognize whether the effect of an achievement operator has been destroyed by the application of subsequent achievement operators.

The main task of the *layout manager* is to convey certain semantic and pragmatic relations specified by the planner by the arrangement of graphic and text fragments received from the medium-specific generators, i.e., to determine the size and the exact coordinates of the entities for positioning them on the document page. We use a grid-based approach as an ordering system for efficiently designing functional (i.e., uniform, coherent and consistent) layouts. The automatic placement of layout objects in a design space can be viewed as a combinatorical problem. Therefore, we treat layout as a constraint satisfaction problem in a finite discrete search space. We encode graphical design knowledge via prioritized constraints expressing semantic/pragmatic and geometrical/topological relations.

WIP has been designed for interfacing with heterogeneous back-end systems such as expert systems, tutoring systems, intelligent control panels, on-line documentation and help systems, which supply the presentation

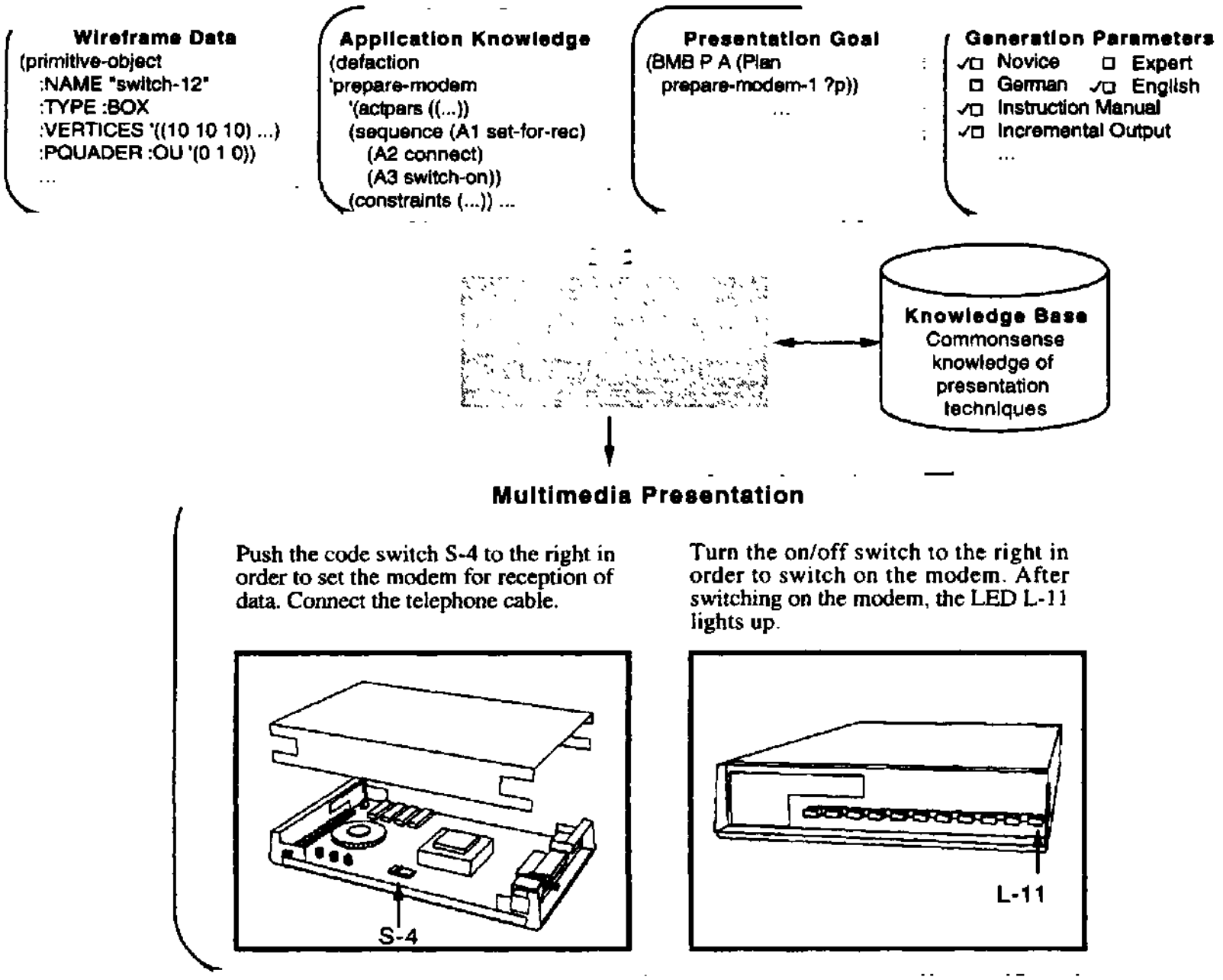

Figure 1: Functional View of the System

system with the necessary input. While for each domain the application knowledge and the wireframe model are fixed, the presentation goal and the generation parameters can be varied. The intelligence of the WIP system lies in its ability to present the same information in a variety of ways depending on the generation parameters. Thus WIP allows for tailoring presentations for individual users in particular communicative situations. To demonstrate this, the video contains a variety of presentations generated by the system starting from the same presentation goal, but different settings of the generation parameters.

WIP can also be used in a stand-alone fashion, where an author specifies the necessary domain information. This leads to the long-term vision of an intelligent authoring system, that forces one to specify information only once in a formal way and then allows the generation of a possibly infinite variety of presentations of this information tailored to various audiences and media. In contrast to the current situation in technical writing and document preparation, this approach - similar to the view concept in database design - could ensure consistency across all derived presentations, since the underlying content is stored in only one place.

The WIP project was supported by the German Ministry of Research and Technology under grant ITW 8901

References

[Andre *et al.*, 1993] E. Andre, W. Finkler, W. Graf, T. Rist, A. Schauder, and W. Wahlster. WIP: The Automatic Synthesis of Multimodal Presentations. In M. Maybury, editor, *Intelligent Multimedia Interfaces,* pages 75-93. AAAI Press, 1993.

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