Research in Multimedia Systems at DFKI

ELISABETH ANDRÉ and THOMAS RIST

*German Research Center for Artificial Intelligence (DFKI), Stuhlsatzenhausweg 3,
D-66123 Saarbrücken; Email: {andre,rist}@dfki.uni-sb.de*

Abstract. The German Research Center for Artificial Intelligence (Deutsches
Forschungszentrum für Künstliche Intelligenz, DFKI) is a non-profit organization
that conducts application-oriented basic research in the field of Artificial Intelligence
and other related subfields of computer science. One of its current research areas is
devoted to Intelligent User Interfaces. The group, headed by Prof. Dr. Wolfgang
Wahlster, is concerned with issues of multimedia/multimodal communication. In this
paper, an overview of two projects carried out by the group is given.

1. WIP (KNOWLEDGE-BASED PRESENTATION OF INFORMATION)

In 1989, our group started work with the project WIP. WIP aimed at the devel-
opment of a presentation system that is able to generate a variety of multimedia
documents from input consisting of a formal description of the communicative
intent of a planned presentation. In the implemented system, the generation
process is controlled by a set of generation parameters such as target group,
presentation objective, resource limitations, and target language (see Figure 1).

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 system with
the necessary input. Since WIP's internal processing schemes are independent
of any particular back-end system it requires only a limited effort to adapt the
system to a new application. In order to validate WIP's transportability the system
was tested in three different application domains (generating illustrated expla-
nations and instructions on using an espresso machine, assembling a lawn-mower,
and maintaining a modem). Starting from the original espresso-machine domain
we did not have to change a single line of code in going to the two new domains.
Only the declarative knowledge sources coded in RAT (Representation of Actions
in Terminological Logics), the lexicon and the geometric information are dif-
ferent. While for each domain the application knowledge and the wireframe model
are fixed, the presentation goal and the generation parameters can be varied.
The benefit of the WIP system lies in its ability to present the same informa-
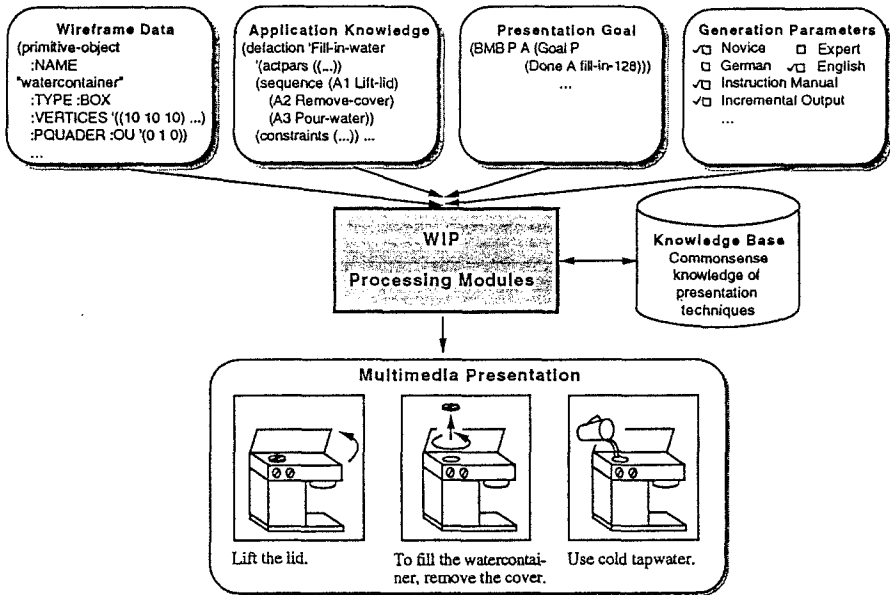tion in a variety of ways depending on the generation parameters. Thus WIP

Fig. 1. Functional view of the system.

allows for tailoring presentations for individual users in particular communicative situations.

One of the important insights we gained from building the WIP system is that it is actually possible to extend and adapt many of the fundamental concepts developed to date in AI and computational linguistics for the generation of natural language in such a way that they become useful for the generation of graphics and text-picture combinations as well. This means that an interesting methodological transfer from the area of natural language processing to a much broader computational model of multimedia communication seems possible. In particular, semantic and pragmatic concepts like coherence, speech acts, anaphora, and rhetorical relations take on an extended meaning in the context of text-picture combinations. For a more detailed description of this approach we refer to the paper by André and Rist in this volume.

The WIP project was supported by the German Ministry of Research and Technology under grant ITW 8901 8. The project team, headed by Prof. Dr. Wolfgang Wahlster, was divided into three subgroups for presentation planning (PP), language generation (TAG-GEN) and knowledge representation (KR). The presentation planning group focused on problems of context-directed selection of contents, automated graphics design, coordination of text and graphics (Elisabeth André and Thomas Rist), and constraint-based layout (Winfried Graf). The language generation group worked on the incremental and parallel generation of text using lexicalized tree-adjoining grammars with feature unification (Dr. Karin Harbusch, Anne Kilger and Wolfgang Finkler). The knowledge representation group extended the expressiveness of the terminological logic used

in WIP with regard to the representation of temporal relations, action structures, default values and exceptions (Prof. Dr. Bernhard Nebel, Prof. Dr. Jochen Heinsohn and Hans Jürgen Profitlich).

2. PPP (PERSONALIZED PLAN-BASED PRESENTER)

In February 1994, the members of the PP and the KR group started with the project PPP (under grant ITW 9400). The PPP project continues work done in WIP by adding three fundamental extensions:

- *Planning Multimedia Presentation Acts*
  In contrast to WIP, the PPP system not only synthesizes multimedia documents, but also plans how this material is to be presented to various users. One particular objective of the PPP project is to emulate more natural and efficient presentations by the use of an animated character as a presenter who will show and explain the generated material.

- *Interactive Multimedia Presentations*
  Since it is impossible to anticipate the needs and requirements of each potential user, a presentation system should allow for interaction. The PPP system responds to follow-up questions concerning the domain as well as to meta comments on the act of presentation. Furthermore, PPP allows for user interaction, i.e., PPP responds to follow-up questions concerning the domain as well as to meta comments on the act of presentation.

- *Monitoring the Effectiveness of a Presentation*
  In order to find out whether a user really understood an instruction, the presentation system must keep track of the user's behavior in the world. Given such an additional source of feedback, the system is able to monitor the effectiveness of its presentations and can continuously adapt the instruction strategies to the current situation.

Furthermore, in PPP we aim at a firm representational foundation to allow for easy adaptations of new domains. This includes the development of representational techniques which are flexible and powerful enough to support a wide range of applications.

SELECTED PAPERS FROM THE WIP AND PPP PROJECTS

André, E. & Rist, T. (1990). Synthesizing Illustrated Documents: A Plan-Based Approach. In *Proceedings of InfoJapan '90* 2: 163–170. Tokyo. Also as DFKI Research Report RR-91-06.
André, E. & Rist, T. (1990b). Towards a Plan-Based Synthesis of Illustrated Documents. In *Proceedings of The Ninth ECAI*, 25–30. Stockholm. Also as DFKI Research Report RR-90-11.
André, E. & Rist, T. (1993). The Design of Illustrated Documents as a Planning Task. In Maybury, M. (ed.) *Intelligent Multimedia Interfaces*, 94–116. AAAI Press. Also as DFKI Research Report RR-92-45.
André, E. & Rist, T. (1994). Generating Coherent Presentations Employing Textual and Visual Material. *Artificial Intelligence Review* (this volume).

André, E. & Rist, T. (1994). Multimedia Presentations: The Support of Passive and Active Viewing. In *Working Notes of the AAAI Spring Symposium on Intelligent Multi-Media Multi-Modal Systems*, 22–29. Stanford University.

André, E. & Rist, T. (1994). Referring to World Objects with Text and Pictures. In Proceedings of *The Fifteenth COLING* 1: 530–534. Kyoto, Japan.

André, E., Finkler, W., Graf, W., Rist, T., Schauder, A. & Wahlster, W. (1993). WIP: The Automatic Synthesis of Multimodal Presentations. In Maybury, M. (ed.) *Intelligent Multimedia Interfaces*, 75–93. AAAI Press. Also as DFKI Research Report RR-92-46.

André, E., Graf, W., Heinsohn, J., Nebel, B., Profitlich, H.-J., Rist, T. & Wahlster, W. (1993). *PPP – Personalized Plan-Based Presenter*. Document D-93-5, DFKI, Saarbrücken.

André, E., Herzog, G. & Rist, T. (1994). Multimedia Presentation of Interpreted Visual Data. In *AAAI-94 Workshop on the Integration of Natural Language and Vision Processing*, 74–82. Seattle, USA.

Butz, A. (1994). Betty: Planning and Generating Animations for the Visualization of Movements and Spatial Relations. In *Advanced Visual Interfaces* (Proc. of AVI '94, Bari, Italy) (to appear).

Graf, W. & Neurohr, S. (1994). *Using Graphical Style and Visibility Constraints for a Meaningful Layout in Visual Programming Interfaces*. Report, DFKI, Saarbrücken (to appear).

Graf, W. (1992). Constraint-Based Graphical Layout of Multimodal Presentations. In *Advanced Visual Interfaces* (Proc. of AVI '92, Rome, Italy), 365–385. Singapore: World Scientific Press. Also as DFKI Research Report RR-92-15.

Harbusch, K., Finkler, W. & Schauder, A. (1991). Incremental Syntax Generation with Tree Adjoining Grammars. In Brauer, W. & Hernandez, D. (eds.) *Verteilte Künstliche Intelligenz und kooperatives Arbeiten: 4. Internationaler GI-Kongreß Wissensbasierte Systeme, Proc.*, 363–374. Springer: Berlin, Heidelberg.

Heinsohn, J., Nebel, B., Profitlich, H.-J., Rist, T. & Wahlster, W. (1992). *RAT: Representation of Actions Using Terminological Logics*. Document, DFKI, Saarbrücken.

Kilger, A. (1994). Using UTAGs for Incremental and Parallel Generation. *Computational Intelligence* (to appear).

Rist, T. & André, E. (1992a). From Presentation Tasks to Pictures: Towards a Computational Approach to Graphics Design. In Proceedings of *The Tenth ECAI*, 764–768. Wien. Also as DFKI Research Report RR-92-44.

Rist, T. & André, E. (1992b). Incorporating Graphics Design and Realization into the Multimodal Presentation System WIP. In Costabile, M. F., Catarci, T. & Levialdi, S. (eds.) *Advanced Visual Interfaces* (Proc. of AVI '92, Rome, Italy), 193–207. World Scientific Press: Singapore.

Rist, T. & André, E. (1992b). Designing Coherent Multimedia Presentations. In Salvendy, G. & Smith, M. J. (Hrsg.) *Human-Computer Interaction: Software and Hardware Interfaces* (Proc. of HCI–93), Volume 19B, 434–439. Elsevier: Amsterdam, London.

Rist, T., Krüger, T., Schneider, G. & Zimmermann, D. (1994). AWI – A Workbench for Semi-Automated Illustration Design. In *Advanced Visual Interfaces* (Proc. of AVI '94, Bari, Italy) (to appear).

Wahlster, W., André, E., Graf, W. & Rist, T. (1991). Designing Illustrated Texts: How Language Production is Influenced by Graphics Generation. In Proceedings of *The Fifth EACI*, 8–14. Berlin, Germany.

Wahlster, W., André, E., Bandyopadhyay, S., Graf, W. & Rist, T. (1991). WIP: The Coordinated Generation of Multimodal Presentations from a Common Representation. In Ortony, A., Slack, J. & Stock O. (eds.) *Communication from an Artificial Intelligence Perspective: Theoretical and Applied Issues*, 121–144. Berlin, Heidelberg: Springer. Also as DFKI Research Report RR-91-08.

Wahlster, W., André, E., Finkler, W., Profitlich, H.-J. & Rist, T. (1991). Plan-Based Integration of Natural Language and Graphics Generation. *AI Journal* 63: 387–427. Also as DFKI Research Report RR-93-02.

Wazinski, P. (1992). Generating Spatial Descriptions for Cross-Modal References. In Proceedings of *The Third Conference on Applied Conference on Applied Natural Language Processing*, 56–63. Trento, Italy.