Interactive Multimedia Explanation for Equipment Maintenance and Repair

Kathleen McKeown and Steven Feiner

Department of Computer Science
450 Computer Science Building
Columbia University
New York, N.Y. 10027

PROJECT GOALS

We are developing COMET, an interactive system that generates multimedia explanations of how to operate, maintain, and repair equipment. Our research stresses the dynamic generation of the content and form of all material presented, addressing issues in the generation of text and graphics, and in coordinating text and graphics in an integrated presentation.

COMET contains a static knowledge base describing objects and plans for maintenance and repair, and a dynamic knowledge source for diagnosing failures. Explanations are produced using a content planner that determines what information should be communicated, a media coordinator that determines which information should be realized in graphics and which in text, and separate text and graphics generators. The graphics and text for a single explanation are laid out on the screen by a media layout component. A menu interface allows users to request explanations of specific procedures or to specify failure symptoms that will invoke a diagnostic component. The diagnostic component can ask the user to carry out procedures that COMET will explain if requested. In contrast to hypermedia systems that present previously authored material, COMET has underlying models of the user and context that allow each aspect of the explanation generated to be based on the current situation.

Our focus in the text generation component has been on the development of the Functional Unification Formalism (FUF) for non-syntactic tasks, of a large syntactic grammar in FUF, of lexical choice in FUF using constraints from underlying knowledge sources and from past discourse, and of models of constraints on several classes of word choice. Important results in knowledge-based graphics generation include the automated design of 3D technical illustrations that contain nested insets, algorithms for and rule-based application of illustrative techniques such as cutaway views, a design-grid–based methodology for display layout, and development of a testbed for knowledge-based animation.

Finally, we have had significant results in the development of our media coordinator which, unlike other systems, features a common description language that allows a fine-grained division of information between text and graphics. The media coordinator maps information to media specific resources, and allows information expressed in one media to influence realization in the other. This allows for tight integration and coordination between different media.

RECENT RESULTS

- Incorporated user model constraints on word selection in order to use words appropriate to user's vocabulary level. This includes both word substitution and replanning of sentence content when there is no word that can be substituted for unknown word (e.g., "Check the polarity." is replaced by "Make sure the plus lines up with the plus.")

- Completed sentence-picture coordination, allowing longer sentences to be broken into shorter ones that can separately accompany each generated picture when necessary.

- Added all m&r procedures for the radio from the manual to the knowledge base and augmented the lexicon to include new words for the procedures.

- Continued implementation of cross-references between text and graphics, including query facilities for the graphics representation that allow the text generator to determine where and how an object is displayed, use of these facilities along with the underlying knowledge base to construct cross-references (e.g., "The battery is shown in the cutaway view of the radio."), and development of a lexicon for such cross-references.

- Extended the graphics generator to support the maintenance of visibility constraints through a set of illustrative techniques modeled after those used by technical illustrators. These involve detecting objects that obscure those that must remain visible and rendering the obscuring objects using transparency, cutaway views, and "ghosting" effects. The effects are invoked automatically as the graphics generator designs its illustrations.

- Developed facilities for *dynamic illustrations* that are incrementally redesigned to allow users to explore the generated pictures by choosing viewpoints different from those selected by the system.

PLANS FOR THE COMING YEAR

We plan to finish implementation of cross references between text and graphics, to increase the ways in which the user model can influence lexical choice, and to incorporate all extensions as part of our demo system. Following that, we will move to a new contract, where we will begin work on identifying usage constraints on a variety of lexical classes through automatic and manual examination of large text corpora.