TECHDOC: Multilingual generation of online and offline instructional text

Dietmar Rösner
FAW Ulm
P.O. Box 2060, 89010 Ulm, Germany
roesner@faw.uni-ulm.de

Manfred Stede
(University of Toronto and) FAW Ulm
P.O. Box 2060, 89010 Ulm, Germany
stede@faw.uni-ulm.de

Project idea

Supplying technical documentation accompanying a product in multiple languages is a growing problem, particularly in Europe with its legislation on the common market. A huge amount of translation work needs to be done when creating and updating technical documentation. In response to these needs, the TECHDOC project aims at supporting the creation and maintenance of technical documentation by *knowledge-based, multilingual generation*. The idea is to build up a knowledge base that includes a model of the product in question, and to produce documentation in multiple languages automatically. At present, the system produces maintenance instructions in English, German, and French.

The application domain of TECHDOC is *technical manuals*, where one has to do with "real-world" texts: the domain is *practical* instead of a "toy world". At the same time, the language that is used in such manuals tends to be relatively simple; one mostly finds straightforward instructions that have been written with the intention to produce text that can be readily understood by a person who is executing some maintenance activity. Moreover, the *structure* of manual sections is largely uniform and amenable to formalization.

System architecture

The system is based on a KB encoding knowledge about the technical domain and about schematic text structure in LOOM, a KL-ONE dialect (LOOM 1991). A typical manual section first describes the location of the object to be repaired/maintained and then lists possible replacement parts/substances; next, the activities are described, which fall into the three general categories of checking some attribute (e.g., a fluid level), adding a substance and replacing a part/substance. These actions are represented as plans in the traditional AI sense, i.e. with pre- and postconditions, and with recursive structure (steps can be elaborated through complete refinement plans). We call the schematic ordering of a manual section its *macrostructure*.

For a language-independent *text* representation, a level was sought that captures the commonalities of the correponding sections of the German, English and French texts, i.e. that is not tailored towards one of the specific languages. Rhetorical Structure Theory (RST) (Mann and Thompson 1987) turned out to be a useful formalism: for almost every section we investigated in our extensive corpus studies of multilingual manuals, the RST analyses for the different language versions were identical. The system implements this finding by mapping the schematic plan structure to a tree that captures the *microstructure* of documents by means of rhetorical relations holding among elementary propositions or sub-trees, and a number of specific annotations. See (Rösner and Stede 1992) for details.

This document representation is successively transformed into a sequence of sentence plans (together with formatting instructions in a selectable target format; SGML, LaTeX, Zmacs and – for screen output – formatted ASCII are currently supported), which are handed over to sentence generators. For English, we use Penman (Penman 1991) and it s sentence planning language (SPL) as input terms. To produce German text, we have implemented a (partial) German version of Penman's grammar (NIGEL), which is enhanced by a morphology module, and a fragment of a French grammar in the same style.

Interactive, multimodal instructions

The interactive variant of the system, TECHDOC-I (Peter and Rösner 1994), produces instructions online, by engaging in a dialog with the user. On the basis of initial questions, users are assigned to a stereotypical class (ranging from 'novice' to 'expert'), which subsequently influences the degree of detail in which instructions are produced.

The generated texts on the screen are mouse-sensitive, so that the user can click on words or phrases (in any of the languages displayed) and request additional information. Depending on the unit selected, a menu offers a range of applicable items for further inquiries. For example, the location of an object referred to in the text can be clarified by

an image in which that part is highlighted. Or, if an action is selected, the system offers to show a short video sequence on the screen, which illustrates how that action is to be performed.

In essence, these facilities have paved the way to move from static, inactive strings as output to an active and dynamic interface for the associated knowledge sources and their various presentation modalities. The key is that all kinds of information (lexemes in various languages, images and object's location therein, and video sequences) are associated with the underlying KB objects.

Document authoring support

Apart from supporting the end user with an interactive system, the system is now being enhanced on the "opposite end", where the input to the document generation process is created. A first version of an *authoring tool* has been designed and implemented and tested with a number of users. This tool allows to interactively build up knowledge base instances of maintenance plans, including the actions and objects involved, and to convert them immediately into multilingual documents. The process is menu-driven, and the available options are determined dynamically in accordance with the current state of the knowledge base — in other words, the knowledge already encoded in the KB constrains the range of possible instances that are to be built with the authoring tool.

Portable KB: the "middle model"

Our KB design focuses on the targets that the representations can be used for more than language generation (tutoring, diagnosis systems, etc.) and that they should be portable across different technical domains. We have collected recurring knowledge about various technical devices like types of connections, about switches, about objects that turn up in a variety of designs and functions, for example fluid tanks, etc., in a *middle model* that can be carried, together with the basic ontology (the so-called 'upper model' (Bateman 1990)), from one technical domain to the next. As the first practical domain extension, we moved from automobile manuals to several sections of aircraft maintenance instructions, where the model of the specific objects and devices was subsumed under the middle model, in the place where the automobile model previously resided.

The middle model is a set of quite abstract concepts; separating lexical from conceptual information allows us to account for lexical differences between the languages; in our corpus studies of automobile manuals, we found, for instance, that English often uses the general *remove* where German verbs characterize the physical action more specifically (*abziehen, abnehmen, herausschrauben, herausziehen,* etc.). Our lexicalization procedure captures these language-specific differences by exploring

subsumption (see (Stede 1993)); the important point to note here is that such lexical matters are independent of the "conceptual kernel", which is represented in a language-independent way and allows for sharing as much information as possible.

Perspective: From research prototype to practical usage

TECHDOC is being developed into two complementary directions: toward online, interactive instruction dialog on the one hand, and toward production of fully formatted paper documents on the other. Each is suitable for its particular purposes; both are important. Online systems, for example, will make sense in settings where the system can be given sensory data from the device (temperature, pressure values, etc.) and thus exactly those instructions can be generated that are relevant in the current situation. Paper documents, on the other hand, will still be used for getting familar with a product prior to using it, as well as for reference purposes — they won't vanish altogether.

References

John A. Bateman. Upper modeling: A level of semantics for natural language processing. In *Proc. of the 5th International WS on Natural Language Generation*, Pittsburgh, PA, 1990.

The LOOM Knowledge Representation System. Documentation package, USC/ISI, Marina Del Rey, CA, 1991.

William C. Mann and Sandra A. Thompson. Rhetorical structure theory: A theory of text organization. In L.Polanyi, editor, *The Structure of Discourse*. Ablex, Norwood, NJ, 1987.

The Penman documentation package. USC/ISI, Marina Del Rey, CA, 1991.

Gerhard Peter and Dietmar Rösner. User-model-driven generation of instructions. To appear in *User Modeling and User-Adapted Interaction*, 1994.

Dietmar Rösner and Manfred Stede. Customizing RST for the automatic production of technical manuals. In R. Dale, E. Hovy, D. Rösner, and O. Stock, editors, *Aspects of Automated Natural Language Generation - Proc. of the 6th International WS on Natural Language Generation*. Springer, Berlin/Heidelberg, 1992.

Manfred Stede. Lexical options in multilingual generation from a knowledge base. In: Working notes of the 4th European Workshop on Natural Language Generation.