

SYSTEM DEMONSTRATION

INTERACTIVE GENERATION AND KNOWLEDGE ADMINISTRATION IN MULTIMETEO

José Coch, *ERLI*
261, rue de Paris
93556 Montreuil Cedex, FRANCE
jose.coch@erli.fr

1. Several important problems to solve for applying Generation to the real world

We believe that it is important from the practical point of view to use natural language generation (NLG) in real world applications. The key benefits are outlined as follows:

- Higher quality of generated texts
- Cost effective maintenance and adaptability
- Usability (including acceptability by the users).

Generally, it takes "exotic" (linguists, knowledge engineers) manpower to maintain and adapt a NLG system. To avoid this problem, we have developed a **Knowledge Administration** station, which is usable by the target population (in our project, weather forecasters).

On the other hand, the system is designed to help the forecasters and not replace them. It is able to adapt to each forecaster's style and manage enhancements they wish to bring to their texts. With that in mind, the **Interactive Generation** environment was designed to allow forecasters to modify generated texts in their native language, and then generate weather forecasts in several foreign languages based on those modifications.

Interactive Generation is a viable alternative to Automatic Translation.

2. MultiMeteo

2.1. General motivation

The European weather forecast market is no longer limited to the national level. Travelling is growing rapidly, with an increasing number of people going abroad, for both personal and professional reasons. Markets for meteorological forecasts tend to overlap borders and thus need multilingual tools.

The volumes involved are high, and time constraints for producing fresh usable information are stringent: several hundreds of bulletins for each Weather Office, produced several times a day.

End users have also become much more demanding about quality and tailoring, which here are realised by different styles of weather forecast.

1.2. Goal and characteristics of MultiMeteo

The goal of the MultiMeteo project is to build an automatic multilingual short-report generation system, and to use it in the Meteorology field for producing weather forecasts from structured data¹.

¹ Other weather-forecast generation systems are FoG [Goldberg et al. 94], Weathra [Sigurt et al. 92], MET90 [Sigurt et al. 96], and METEOVIS [Kerpedjiev & Noncheva 90].

This system will allow European forecasters to automatically produce texts in English, French, German, Spanish, and Dutch.

MultiMeteo is a 3-year project funded partially by the Language Engineering programme of the European Commission, and partially by European weather offices. It involves around 40 people in France, Spain, Austria and Belgium. There are around 15 pilot sites in Europe, which are representative of different meteorological characteristics (south, north, plain, mountain, sea, etc.). In each site, 3 or 4 styles will be developed (local and regional general public, seaside general public, fun-board sport oriented, etc.).

3. Interactive Generation and Knowledge Administration

The Interactive Generation environment allows the user (the forecaster) to:

- Read the generated text
- Modify it in terms of concepts, modifiers, links, etc.
- Read the modified version

The following example shows a generated text, and the menu proposed by the system when the user clicks on a clause (here "le ciel couvert devient peu nuageux") - "the grey sky becomes less cloudy"

The screenshot shows the MultiMeteo application window. The title bar reads "MULTIMETEO". The menu bar includes "Analyse", "Modifier", "Sauvegarder", "Navigation", and "EXIT". The main text area displays a weather forecast for "METEO-FRANCE, centre départemental du Mans" for December 16th. The forecast is divided into sections: "Ce jour :", "Pour la nuit :", and "Pour demain :". The "Pour demain :" section contains the clause "Le ciel couvert devient peu nuageux, avec de la bruine, le vent de Nord-Est souffle de 30 à 45 km/h". A context menu is open over this clause, listing several concepts and modifiers:

- Concept N1 (P1) Devenant N2**
- Concept Dvpt D Eclaircies
- Ciel Se Dégageant
- Concept Nuages Devenant Moins Nombreux
- Concept Ciel Devenant Moins Nuageux
- Concept Dvpt De Timides Eclaircies
- Concept Eclaircies Devenant Un Peu Plus Franches

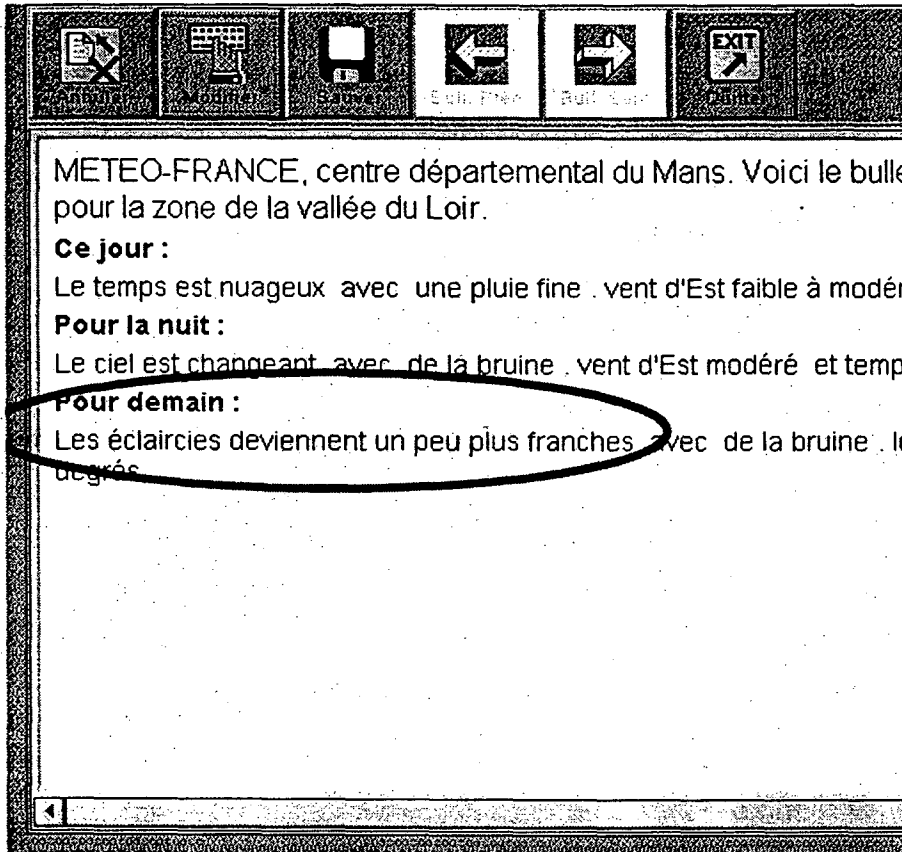
- devrai(en)t
- pourra(ont)
- pourrai(en)t

- commencera(ont)
- continuera(ont)

- encore
- toujours

In the above menu, alternative concepts are proposed, and several types of modifiers (modality, phase, etc.) can be added.

For example, if the concept "Eclaircies devenant un peu plus franches" (Sunny spells becoming more extensive) is selected, the system generates the following text:



Once the forecaster saves the result, MultiMeteo generates the weather forecast in all the selected languages (which are English, Spanish and German). Of course his/her modifications are taken into account in all of the languages.

In addition, MultiMeteo offers a Knowledge Administration station, which allows the administrator to modify the style of the texts to be generated, in terms of default structure, titles, order between paragraphs, available models of paragraphs, type of sentences used (telegraphic / non telegraphic), terms to be used, etc.

In fact, there are two levels of administration: central and local. Certain functions are specific to the central administration, such as the creation of a new style from scratch, or the management of the multilingual terminology.

4. Technical Characteristics

MultiMeteo runs under PC/Windows 95 (or more), Windows NT and Unix station/Solaris2. The input is a numerical matrix, where columns are the meteorological parameters and rows are time steps. The output is an HTML or RTF file that contains one or more weather forecasts.

The MultiMeteo software has 4 main modules:

- The Front-End Module (FEM)
- The Administration Server
- The Generation Server
- The Help facilities.

The FEM, the Administration, and the Help are written in JAVA.

The Generation Server is based on ERLI's AlethGen toolbox, which has already been used for generating texts, in particular for producing correspondence for a leading French mail-order company (see [Coch 1996a] and [Coch 1996b]). AlethGen is divided in an engine (written in C++) and Knowledge Bases. The approach is inspired by the Meaning-Text Theory

Acknowledgements

The MultiMeteo project is partially funded by the European Commission's Language Engineering programme (LE 1612), and we are very grateful to the members of the European Commission's DGXIII for their support of this work.

References

- [Coch 1996a] Coch José, 1996. "Overview of AlethGen". Proceedings of the International Workshop on Natural Language Generation (INLG-96). Herstmonceux, England.
- [Coch 1996b] Coch José, 1996. "Evaluating and comparing three text-production techniques", in *Proceedings of the 16th Conference of Computational Linguistics, Coling 96*, Copenhagen, Denmark.
- [Coch 1997a] Coch José, 1997. "Applications industrielles de la génération : pourquoi et comment", in *Proceedings of GAT'97*. Grenoble, France.
- [Coch 1997b] Coch José, 1997. "Quand l'ordinateur prend la plume : la génération de textes", in *Document Numérique Vol.1, N°3*, Hermès, France.
- [Goldberg et al 1994] Goldberg Eli, Driedger Norbert, Kittredge Richard, 1994. "Using natural-language processing to produce weather forecasts". in *IEEE Expert*, 9(2):45-53.
- [Kerpedjiev & Noncheva 1990] Kerpedjiev Stephan, Noncheva Veska, 1990. "Intelligent handling of weather forecasts" in *Proceedings of the 13th Conference of Computational Linguistics, Coling 90*, Helsinki.
- [Sigurd et al 1992] Sigurd Bengt, Willners Caroline, Eeg-Olofsson Mats, Johansson Christer, 1992. "Deep comprehension, generation and translation of weather forecasts (Weathra)" in *Proceedings of the 14th Conference of Computational Linguistics, Coling 92*, Nantes.
- [Sigurd et al 1996] Sigurd Bengt, Lastow Birgitta, Våvargård Tomas, 1996. "Computer generation of weather overviews" in *Proceedings of the ECAI 96 Workshop Gaps and Bridges: New directions in Planning and Natural Language Generation*.