The Multimodal Natural Language Generator – MNLG

Paul Piwek – ITRI

What is the MNLG?
The MNLG is a platform for multimodal generation. The input is a series of communicative acts. For each act, the input specification includes the dialogue act type, semantics (in terms of DRT), speaker (including information about personality), addressee and common ground at the outset of the act. The output of the MNLG is the input representation enriched with specifications for the realization of the communicative acts by means of language and gestures. Both input and output are represented in the XML-compliant RRL language ([3]). Generation consists of tree construction: the goal is to construct a tree with only terminal leaves that has the input specification as its mother node. Selection of trees for expansion from a tree repository relies on unification and supports seamless integration of canned text, templates (cf. [5]) and generation rules (see [2] for details) and personalized referring expressions generation (see [4] for details).

MNLG Tree Repository Examples

Canned Text for “Yes, I agree.”

Template NP is N years old.

Rule for S → NP VP

Applications

The MNLG is one of the components of the NECA eShowroom application. This application generates dialogues between a virtual car seller and buyer. Users can control the personality, role and interests of the interlocutors. NECA was an EU project involving ITRI (Brighton), DFKI, OFA, Syss AG and IPUS (Boaard).

EPOCH iGuide

The MNLG was used in the EPOCH EU Network of Excellence for a demonstrator in the cultural heritage domain. It is the generation component for a virtual tour guide for the world heritage site of Wolfsburg. The demonstrator subproject involved ITRI (Brighton) and the Universities of Braunschweig and East Anglia.

Implementation

The MNLG has been implemented in Sicstus Prolog, using Profil [1] for representing feature structures. The MNLG is fully documented and makes extensive use of the Sicstus module system to facilitate maintenance and reusability.

Performance

For the eShowroom application, with a total of 138 generation trees, the generation time for individual dialogue acts ranges from 1/100 to 1/10 of a second, depending on the generation strategy (select-first-match versus generate-first-N-and-select).

Acknowledgements

Thanks are due to my colleagues in the NECA project, especially Kees van Deemter who coordinated the NECA project at ITRI. Integration of and resources development for the MNLG for the EPOCH iGuide application was carried out by Ondrej Pacovský, Michel Gérénoux and Karna Rodriguez.

References


Contact Paul.Piwek@itri.bton.ac.uk