The DIPLOMAT Rapid-Deployment Speech MT System

Robert E. Frederking, Ralf D. Brown, and Christopher Hogan

Center for Machine Translation

Language Technologies Institute

Carnegie Mellon University

5000 Forbes Avenue

Pittsburgh, PA 15213-3890 USA

{ref,ralf,chogan}+@cs.cmu.edu

1. Overview of DIPLOMAT project

The DIPLOMAT project is designed to test the feasibility of rapid-deployment, wearable speech translation systems. By "rapid-deployment", we mean being able to develop an MT system that performs initial translations at a useful level of quality between a new language and English within a matter of days or weeks, with continual, graceful improvement to a good level of quality over a period of months. While beginning our investigations into new semi-automatic techniques for knowledge-base development, we have already produced an initial bidirectional system for Serbo-Croatian/English speech translation in less than a month, and are currently developing Haitian-Creole/English and Korean/English systems.

The full DIPLOMAT system also involves research in speech understanding and synthesis, as well as wearable computer systems. The speech understanding system used is the well-known SPHINX II HMM-based continuous speech recognition system [5,8]; the speech synthesis system is a newly-developed subword concatenative system [6]. The user interface uses the UNICODE multilingual character encoding standard to facilitate rapid addition of new languages.

2. MT Technology used in DIPLOMAT

The MT component of DIPLOMAT is the Pangloss-Lite (PanLite) system [3]. PanLite is a standalone C++ re-implementation of several major components from the Pangloss machine translation system [7]. Pangloss was a joint project between three sites: the Computing Research Laboratory of New Mexico State University, the Information Sciences Institute of the University of Southern California, and the Center for Machine Translation of Carnegie Mellon University. It was funded by the U.S. Department of Defense. PanLite incorporates the Pangloss Example-Based MT (EBMT) [1] and Transfer-Based MT engines, and its statistical language modeller [2], as well as the newly-implemented iCelos morphological analyzer, within the Multi-Engine MT architecture [4] developed during the course of the Pangloss project (described further below).

Due to improved design and the C++ implementation, PanLite runs very quickly. For example, the EBMT engine formerly required several minutes to translate a typical newswire sentence; it now requires under 10 seconds (and this with a much larger corpus).

To allow its use in the widest variety of applications, PanLite has been designed to translate strings provided either on the standard input or via network sockets, and to produce as output either the best composite string or the full chart of scored alternative translations of phrases. The latter is necessary, for example, when the output will be supplied to an external graphical user interface (GUI) for post-editing. We envision PanLite being employed in any number of applications involving translation of unrestricted text; one application under consideration is for browsing of pages in foreign languages on the World Wide Web.

2.1. The Multi-Engine MT Architecture

In the Multi-Engine MT (MEMT) architecture, several MT engines, each employing a different MT technology, are applied in parallel to each input. Each engine attempts to translate the entire input text, segmenting each sentence in whatever manner is most appropriate for its technology, and putting the resulting output segments into a shared chart data structure after giving each segment a score indicating the engine's internal assessment of the quality of the output segment. The output segments are indexed in the chart based on the positions of the corresponding input segments. Since the scores produced by the engines are of variable reliability, we use statistical language modelling techniques adapted from speech recognition research to select the best overall set of outputs [2]. (In our demonstration, we will describe the technology in somewhat greater detail.)

3. Demonstration

We expect to demonstrate the PanLite system at the MT Summit running on a wearable "Language Translator Smart Module" currently under development by the Wearable Computer group here at CMU, plus two Toshiba laptops, which take the place of future wearable user interface modules. The demonstration will include speech recognition and synthesis, and user interaction to correct both MT and speech recognition via Graphical User Interfaces (GUIs). We currently have Serbo-Croatian/English and Spanish/English bidirectional MT systems, and are developing Haitian-Creole/English and Korean/English versions.

4. References

[1] Ralf Brown, 1996. "Example-Based Machine Translation in the Pangloss System." In *Proceedings of the 16th International Conference on Computational Linguistics* (COLING-96).

[2] Ralf Brown and Robert Frederking, 1995. "Applying Statistical English Language Modeling to Symbolic Machine Translation." In *Proceedings of the Sixth International Conference on Theoretical and Methodological Issues in Machine Translation* (TMI-95), pages 221-239.

[3] Robert Frederking and Ralf Brown, 1996. "The Pangloss-Lite Machine Translation System." In *Proceedings of the Conference of the Association for Machine Translation in the Americas* (AMTA).

[4] Robert Frederking and Sergei Nirenburg, 1994. "Three Heads are Better than One." In *Proceedings of the fourth Conference on Applied Natural Language Processing* (ANLP-94), Stuttgart, Germany.

[5] Xuedong Huang, Fileno Alleva, Hsiao-Wuen Hon, Mei-Yuh Hwang, Ronald Rosenfeld, 1992. "The SPHINX-II Speech Recognition System: An Overview." Carnegie Mellon University Computer Science Technical Report CMU-CS-92-112.

[6] Kevin Lenzo, 1997. Personal Communication.

[7] Sergei Nirenburg (ed.), 1995. "The Pangloss Mark III Machine Translation System." Joint Technical Report, Computing Research Laboratory (New Mexico State University), Center for Machine Translation (Carnegie Mellon University), Information Sciences Institute (University of Southern California). Issued as CMU technical report CMU-CMT-95-145.

[8] Mosur Ravishankar, 1996. "Efficient Algorithms for Speech Recognition." Ph.D. Thesis. Carnegie Mellon University.