

Estos informes se pueden pedir a Toni Badia (tbadia@upf.es).

Bibliografía

Markantonatou, S., Sadler, L. eds., (1994): "Grammatical Formalisms: Issues in Migration", Studies in Machine Translation and Natural Language Processing, Volum 4, Luxemburgo.

Badia, T. (1993): "Inicios de una gramática para el español en ALEP, un formalismo de unificación", Boletín de la SEPLN, Nº14, Santiago de Compostela.

Pollard, C., Sag, I. (1992) Head-Driven Phrase Structure Grammar, Final Draft, June 15, 1992.

TRADE (MLAP93/003)

Nuria Bel Rafecas

Maite Melero Nogués

GILCUB (FBG/Universitat de Barcelona)

Introducción

La Traducción Automática se considera una de las aplicaciones más importantes de la Ingeniería Lingüística, desde el punto de vista comercial.

A pesar de que el problema de la TA está lejos de haber sido resuelto, se pone de manifiesto la necesidad de disponer de productos operativos que cubran, al menos parcialmente, la demanda del mercado en este sentido, proporcionando ayudas y herramientas para la traducción.

Así, desde instancias europeas, se favorece el apoyo a los proyectos aplicados de TA que sean capaces de reutilizar los recursos y las tecnologías existentes, con objeto de crear productos industriales aptos para ser comercializados.

TRADE está basado en un sistema prototipo de traducción automática, conocido como E-STAR, desarrollado bajo los auspicios del programa EUROTRA, y que supone una extensión y mejora del propio prototipo de EUROTRA, capaz de servir como base para un sistema industrial. El proyecto TRADE pone énfasis en los requerimientos principales del usuario, como son, rendimiento, extendibilidad, robustez, asequibilidad y tratamiento de "textos

reales". El sistema ha sido especialmente diseñado para superar los problemas propios de los sistemas estratificacionales (según el modelo EUROTRA, basado en la transferencia entre niveles generadores), pero conservando, sin embargo, sus características más deseables. Así, el formalismo lingüístico, basado en la unificación, siendo totalmente compatible con el de EUROTRA, está dotado de mayor poder expresivo e incorpora mecanismos de "seguridad" en la traducción, lingüísticamente motivados, que permiten alcanzar la deseada robustez (relajación de condiciones, reglas de preferencia, ...).

Descripción del proyecto

El proyecto TRADE, se centra, en la actualidad, en tres lenguas europeas y en la traducción de cualquiera de ellas a las otras dos. Estas son: inglés, español e italiano. Se prevee que el sistema sea capaz de traducir textos pertenecientes a ámbitos limitados, definidos por los propios usuarios finales. Las áreas sobre las que actualmente se está trabajando son: la Seguridad Social y la informática de usuario.

Participan en el proyecto los siguientes centros de investigación:

Centro de Cálculo de Sabadell, Barcelona (España)

Gruppo DIMA, Turín (Italia)

Universidad de Manchester, Manchester (Reino Unido)

GILCUB (FBG/Universitat de Barcelona), Barcelona (España)

El proyecto se inició en enero de 1994 y su desarrollo está organizado de la siguiente manera:

Fase 1: Definición de especificaciones sobre la arquitectura general del sistema y sobre la cobertura lingüística, selección de corpus y elaboración de baterías de prueba.

Fase 2: Desarrollo del software básico y de las herramientas de procesamiento de texto. Migración de los recursos lingüísticos de las gramáticas de Eurotra al nuevo formalismo. Optimización de las implementaciones en base a las características innovadoras del E-STAR. Cobertura extensiva de los fenómenos según el estudio de corpus de la fase previa. Desarrollo de la interficie de usuario. Integración de todos los módulos.

Fase 3: Verificación, validación y documentación del sistema.

Hasta este momento se ha cubierto la primera fase y, prácticamente, la segunda, habiéndose alcanzado, en líneas generales, los objetivos propuestos. El proyecto fue sometido recientemente a la evaluación externa por parte de expertos en el ámbito de la TA, con resultados altamente satisfactorios.

Informes publicados

Espejo J.M., Pérez F., Somá E. "System architecture and design specifications".

Mazzini G., Steinberger R., Melero M. "Corpus study and coverage definition".

Allegranza V. "Criteria for testing and evaluation of linguistic components".

Espejo J.M, Pérez F., Chambers C. "Text Handler".

Somà E., Tesio R. "Core Software Enhancement Specification".