

Integración de IRIS en ATOS (Sistema de Operador Telefónico Automático).

Grupos Participantes en el Proyecto:

- Grupo de Investigación en Lingüística Computacional JULIETTA. Universidad de Sevilla
- Telefónica Investigación y Desarrollo. División de Tecnología del Habla.

Persona de contacto: Teresa López Soto. Universidad de Sevilla. Email: teresa@fing.us.es

Dirección de contacto: <http://www.fing.us.es/>

Resumen

1. Introducción

ATOS (Automatic Telephone Operator Service) [Alvarez et al. 1996] es un sistema interactivo dirigido por medio de la voz que permite al usuario activar y configurar todas las opciones que permiten las PABX actuales [Lopez-Soto et al. 1997]. Básicamente consta de 3 módulos: reconocedor de voz, módulo de PLN, y gestor de diálogo (GD). El módulo de RV es capaz de reconocer un vocabulario de unas 20.000 palabras y emplea un modelo estadístico del lenguaje. El sistema es independiente del usuario. La tasa media de error de palabra medida en condiciones de laboratorio es de alrededor de un 95%. Los errores más frecuentes son los debidos a falta de concordancia así como inserciones y omisiones de palabras. El reconocedor emplea unidades dependientes del contexto que se modelan mediante modelos ocultos de Markov semicontínuos.

Las unidades empleadas son trifenemas [Amores et al. 1994]. El módulo de PLN se denomina IRIS. IRIS recibe como cadena de entrada la secuencia generada por un reconocedor de voz y pasa la salida al GD. A continuación nos centraremos en los mecanismos de parsing que han servido para analizar las muestras de lenguaje hablado que genera el reconocedor de voz.

Asimismo presentamos el protocolo CTAC, que asegura una interfaz fluida entre el sistema de PLN y el GD.

2. IRIS: Mecanismos de parsing

Los mecanismos de parsing en IRIS se han diseñado para hacer frente a las características propias del lenguaje hablado así como a los

errores producidos en la fase de reconocimiento.

IRIS constituye una versión modificada de Episteme [Amores & Quesada 1997] para esta aplicación. Las incorporaciones más novedosas que se han introducido en la fase de análisis son las siguientes:

2.1 Categoría VOID: Palabras sin información semántica relevante.

Sólo los términos léxicos que aportan información del dominio pasan al parser. El resto caen bajo la categoría VOID y son ignorados. El sistema puede analizar o no estas entradas según cambie el dominio mediante la activación o desactivación del comando correspondiente (VoidWordsIgnore).

2.2 Palabras no incluidas en el léxico.

Mediante el comando NotFoundWordIgnore se previene la entrada al parser de aquellas palabras que no pertenecen al léxico. El comando NotFoundWordDefault permite al sistema aumentar progresivamente el vocabulario asignando una categoría por defecto a la palabra desconocida y comprobando su validez en el análisis.

2.3 Análisis de subcadenas

El módulo de parsing permite el análisis de fragmentos del total de la cadena de entrada con el comando ConfParserSolution-PartialStrings. Cuando una unidad de habla no ha sido correctamente reconocida, o cuando sólo es posible analizar parte de la información contenida en la secuencia de entrada, IRIS puede extraer las subcadenas gramaticales de la cadena completa. Así se consigue que el proceso de análisis no quede interrumpido, y llegue información parcial al GD.

2.4 Resolución de la ambigüedad en el análisis

Se ha implementado un mecanismo de control de prioridad entre reglas gramaticales asociadas a un mismo símbolo en la parte izquierda de la producción. Los niveles de prioridad se establecen en función de la plausibilidad semántica de los nodos analizados cuando se emparejan a una producción gramatical. El sistema selecciona el análisis más completo desde el punto de vista de la información que aporta en relación al dominio. En segundo lugar, se han definido unos algoritmos para seleccionar el análisis más completo según los elementos consumidos en el análisis. Hemos considerado los siguientes criterios: posición, longitud, y el denominado "criterio G" o criterio de globalidad.

2.5 CTAC: Protocolo de comunicación entre IRIS y el gestor de diálogo.

CTAC es el protocolo que permite la comunicación fluida entre IRIS y el GD. Se define como una especificación tipificada de núcleos semánticos. Este protocolo permite recuperar información aunque las oraciones no se hayan analizado completamente. Esto se consigue gracias a la estructura de rasgos (CTAC) asociada a cada elemento de análisis, desde los nodos no terminales a los raíces, y que se mantiene durante todo el proceso de análisis.

3. Resultados actuales

Con los mecanismos descritos arriba se ha conseguido un análisis correcto en el 98.2% de las oraciones del corpus (4.028). El promedio de velocidad de análisis es de 70.240 palabras por segundo.

References

- [Alvarez et al. 1996] J. Álvarez, C. Caminero, C. Crespo & D. Tapias. 1996. The Natural Language Processing Module for a Voice Assisted Operator at Telefónica I+D. ICSLP-96. Philadelphia
- [Amores & Quesada 1997] J. G. Amores & J. F. Quesada. 1997. Episteme En Procesamiento del Lenguaje natural, 21:1-16.
- [Amores et al. 1994] J. G. Amores, J. F. Quesada & D. Tapias. 1994. Traducción

automática basada en el formalismo LFG con entrada y salida por voz. En Comunicaciones de Telefónica I+D, 21:1-16

- [Lopez-Soto et al. 1997] M. T. Lopez-Soto, J. F. Quesada & J. Álvarez-Cercadillo. 1997. Aplicación de LEKTA al entorno ATOS en Procesamiento del Lenguaje Natural. 21:49-68.

- [Quesada 1997] J. F. Quesada 1997. El algoritmo SCP de análisis sintáctico mediante propagación de restricciones Tesis doctoral. Universidad de Sevilla. Abril de 1997.