

Una aproximación para la adquisición de un corpus de diálogo*

An approach for a dialog corpus acquisition

L.F. Hurtado, E. Segarra, E. Sanchis, F. García

Dpt. de Sistemes Informàtics i Computació
Universitat Politècnica de València
Camí de Vera sn, 46022 València
{lhurtado, esegarra, esanchis, fgarcia}@dsic.upv.es

D. Griol

Departamento de Informática
Universidad Carlos III de Madrid
Av. Universidad 30, 28911 Leganés
dgriol@inf.uc3m.es

Resumen: Presentamos la aproximación que hemos desarrollado para simplificar la tarea de adquisición de corpus de diálogos. Esta aproximación esta basada en el uso de dos Magos de Oz.

Palabras clave: Sistemas de Diálogo, Adquisición de Corpus, Etiquetado Semántico

Abstract: We present the approach we have developed to simplify the dialog corpus acquisition task. This approach is based on the use of two Wizards of Oz.

Keywords: Dialog Systems, Corpus Acquisition, Semantic Tagging

1. *Introducción*

El desarrollo de sistemas de diálogo hablado es un proceso complejo que comporta el diseño, implementación y evaluación de varios módulos en los que intervienen diversas fuentes de conocimiento. Actualmente, una de las aproximaciones adoptadas con más éxito es la basada en modelos estadísticos, que representan de forma probabilística los diferentes procesos implicados en los módulos y en la que los correspondientes modelos se estiman a partir de corpus de diálogos persona-máquina (Williams y Young, 2007; Griol et al., 2008). El éxito de las aproximaciones estadísticas depende principalmente de la calidad de los modelos y, por tanto, de los corpus a partir de los cuales son entrenados.

Con el objetivo de facilitar la adquisición de un corpus de diálogos para la tarea EDECAN-SPORT de reservas de pistas deportivas en el marco del proyecto EDECÁN (Lleida et al., 2006) hemos seguido el proceso que describimos a continuación.

En primer lugar, hemos analizado los diálogos persona-persona facilitados por el servicio de reservas de pistas deportivas de nuestra universidad, que constituye el dominio de trabajo de la tarea EDECAN-SPORT. A partir de estos diálogos se ha

definido la semántica de la tarea en términos de actos de diálogo, tanto para los turnos de usuario como para los de sistema, y se han etiquetado estos diálogos. Con ello disponemos de un corpus inicial muy reducido para la tarea EDECAN-SPORT. A partir de este pequeño corpus etiquetado se ha estimado una versión preliminar del gestor de diálogo (Griol et al., 2008). Este gestor de diálogo se utiliza en el proceso de adquisición de un corpus más amplio de esta tarea mediante la técnica del Mago de Oz.

En segundo lugar, como el corpus inicial no tiene un tamaño suficiente para entrenar modelos adecuados para el módulo de comprensión del habla, no disponemos de una versión preliminar de este módulo para el proceso de adquisición con el Mago de Oz. Nuestra propuesta consiste en el uso de un Mago de Oz específico para el módulo de comprensión diferente del Mago de Oz que supervisa el gestor de diálogo. El hecho de contar con dos Magos permite que como resultado del proceso de adquisición, además del corpus de diálogos, se obtienen los actos de diálogo correspondientes a los turnos de usuario y de sistema (evitando el proceso posterior de etiquetado semántico manual del corpus).

2. *Esquema de la adquisición*

Siguiendo las principales contribuciones de la literatura en el área de los sistemas

* Trabajo parcialmente subvencionado por el gobierno español con el proyecto TIN2008-06856-C05-02

de diálogo hablado, realizamos la adquisición del corpus para la tarea EDECAN-SPORT usando la técnica del Mago de Oz. La diferencia de nuestra propuesta (García et al., 2007) para esta adquisición estriba en trabajar con dos Magos de Oz: un simulador del proceso de comprensión del habla y un supervisor del gestor de diálogo. El primero escucha al locutor y simula el comportamiento de los módulos de reconocimiento y comprensión del habla y proporciona la representación semántica del turno de usuario. A partir de dicha representación, el segundo Mago supervisa el comportamiento del gestor de diálogo real.

2.1. El simulador de comprensión

El Mago de Oz de comprensión traduce la intervención del locutor en su correspondiente representación semántica. Para que el resultado sea lo más parecido posible al de un sistema automático real, la representación semántica generada por el Mago de comprensión sufre un proceso de simulación de errores. Esta simulación (García et al., 2007) está basada en el análisis de los errores de reconocimiento y comprensión generados cuando nuestros modelos son aplicados a un corpus de similares características.

2.2. El gestor de diálogo

Hemos desarrollado una aproximación para la gestión del diálogo haciendo uso de un modelo estadístico que se estima a partir de corpus (Griol et al., 2008). Este modelo de gestor se estima automáticamente a partir de un corpus de diálogos etiquetado en términos de actos de diálogo. Esta aproximación, que se ha desarrollado para una tarea de consultas de horarios y precios de trenes, se ha adaptado para su uso en el marco de la tarea de reservas de instalaciones deportivas EDECAN-SPORT. Esta adaptación tiene en cuenta los nuevos requerimientos introducidos por esta tarea, lo que incluye el uso de un manejador de aplicación que interactúa con la base de datos y que verifica si el requerimiento del usuario cumple con las normas definidas para el uso del servicio de reservas. Las acciones del manejador de aplicación pueden afectar a la decisión que tome el gestor de diálogo, aspecto que no se contemplaba en la tarea anterior.

A partir del corpus persona-persona se ha estimado una versión del gestor que está im-

plementada en el sistema de adquisición, y la misión del Mago de Oz de diálogo consiste en supervisar su funcionamiento. Esta supervisión se lleva a cabo a través de dos aplicaciones, en una se supervisa (se corrige cuando el sistema automático da una respuesta que el Mago considera poco adecuada) el funcionamiento del gestor de diálogo y en la otra se supervisa el funcionamiento del manejador de la aplicación.

3. La adquisición

Basándonos en la aproximación descrita en este artículo se ha realizado la adquisición de 240 diálogos en la que han intervenido 18 locutores de diferentes procedencias geográficas (las sedes de los equipos investigadores del consorcio EDECÁN). Las lenguas implicadas en la adquisición han sido el castellano, el valenciano y el euskera. Se ha definido un conjunto de 15 tipos de escenarios con el objetivo de abarcar toda la casuística de la tarea.

De cada diálogo adquirido se dispone de: cuatro canales de audio, la transcripción de los turnos de usuario (con una media de 5,1 turnos de usuario por diálogo y 6,7 palabras por turno) y el etiquetado semántico de los turnos de usuario y de sistema.

Bibliografía

- García, F., L.F. Hurtado, D. Griol, M. Castro, E. Segarra, y E. Sanchis. 2007. Recognition and Understanding Simulation for a Spoken Dialog Corpus Acquisition. En *TSD 2007*, volumen 4629 de *LNAI*. Springer, páginas 574–581.
- Griol, David, L.F. Hurtado, E. Segarra, y E. Sanchis. 2008. A Statistical approach to spoken dialog systems design and evaluation. *Speech Communication*, 50:666–682.
- Lleida, Eduardo, Encarna Segarra, M. Inés Torres, y J. Macías-Guarasa. 2006. EDECÁN: sistEma de Diálogo multidominio con adaptación al contExto aCústico y de Aplicación. En *IV Jornadas en Tecnología del Habla*, páginas 291–296, Zaragoza, Spain.
- Williams, J. y S. Young. 2007. Partially Observable Markov Decision Processes for Spoken Dialog Systems. En *Computer Speech and Language* 21(2), páginas 393–422.