

Sistema de diálogo para el Proyecto DIHANA

L. F. Hurtado, F. Blat, F. García, S. Grau, D. Griol, E. Sanchis, E. Segarra, F. Torres

Departament de Sistemes Informàtics i Computació

Universitat Politècnica de València. E-46022 València, Spain

{lhurtado,fblat,fgarcia,sgrau,dgriol,esanchis,esegarra,ftgoterr}@dsic.upv.es

Resumen: En este trabajo se describe un sistema de diálogo desarrollado para el Proyecto DIHANA. El sistema está compuesto por siete módulos: un reconocedor automático del habla, un módulo de comprensión del habla, un gestor del diálogo, un módulo de consulta a la base de datos, un generador de respuestas en lenguaje natural, un sintetizador de texto a voz y finalmente por un gestor central de comunicaciones. Para la implementación del sistema se ha optado por una arquitectura siguiendo el paradigma cliente-servidor, donde el gestor central actúa como cliente, gestionando las comunicaciones, y el resto de módulos actúan como servidores.

Palabras clave: Sistemas de diálogo, reconocimiento automático del habla, comprensión del habla, procesamiento del lenguaje natural, gestión del diálogo.

Abstract: In this work we describe a dialog system developed into the DIHANA project. This system consists of seven modules: an automatic speech recognizer, a language understanding module, a dialog manager, a module that manages the queries to the database, a natural language answer generator, a text-to-speech converter and, finally, a central communication manager. For the implementation of the system, we built an architecture based on the client-server paradigm, where the central communication manager works as the client and manages the communications, and the other modules work as servers.

Keywords: Dialog systems, automatic speech recognition, speech understanding, natural language processing, dialog managing.

1. Introducción.

Un sistema de diálogo es un sistema automático capaz de emular a un ser humano en un diálogo con otra persona, con el objetivo de que el sistema cumpla con una cierta tarea (dar una cierta información, o proporcionar ciertos servicios).

El objetivo del proyecto DIHANA es el estudio y desarrollo de un sistema robusto de diálogo modular y distribuido para el acceso a sistemas de información. En concreto, se pretende profundizar en aspectos metodológicos fundamentales en los campos del tratamiento del habla espontánea, modelado del lenguaje, comprensión y diálogo. La tarea del proyecto es proporcionar información en lenguaje natural sobre servicios, horarios y precios de trenes de largo recorrido en español.

En esta contribución se resumen las características básicas de cada uno de los módulos que conforman el sistema de diálogo desarrollado para el proyecto DIHANA, así como el formato de comunicación adoptado para la comunicación entre los diferentes componentes del sistema.

2. Descripción de los módulos.

El esquema utilizado para el desarrollo de sistemas de diálogo suele englobar una serie de módulos genéricos para tratar las múltiples fuentes de conocimiento, que deben coordinarse para responder a los requerimientos del usuario. La Figura 1 resume la arquitectura del sistema, mostrándose a continuación las características principales de los módulos más relevantes.

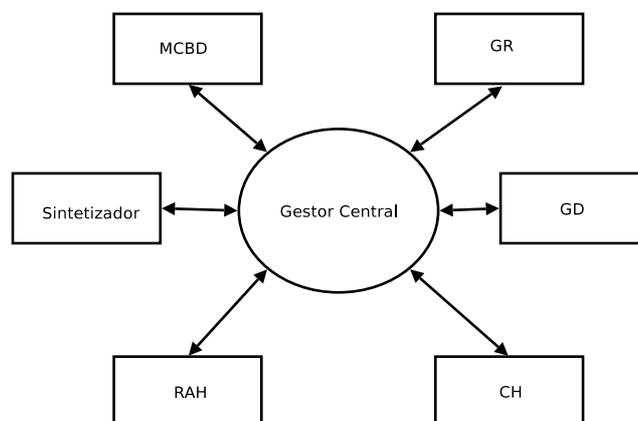


Figura 1: Descripción de la arquitectura del sistema.

1. Módulo de Reconocimiento Automático del Habla (RAH): Facilita la secuencia de palabras que se ajusta a la señal vocal de entrada con una mayor probabilidad. Se ha desarrollado un reconocedor automático del habla expresamente para el proyecto. Asimismo, se ha realizado la integración del reconocedor del habla Sphinx-II y el entrenamiento de los modelos acústicos con SphinxTrain, ambos desarrollados por la Carnegie Mellon University.
2. Módulo de comprensión (CH): Extrae el significado de las palabras reconocidas en el módulo anterior, expresándolo en términos de un lenguaje semántico especificado para la tarea. Para la definición de la semántica de la tarea se utiliza el concepto de frame: cada intervención del usuario genera uno o más frames que representan el significado de la intervención.
3. Gestor de diálogo (GD): Decide qué paso debe dar el sistema tras cada intervención del usuario. Para ello se basa en la interpretación semántica generada por comprensión, el historial del diálogo, el estado actual del sistema, información obtenida de la base de datos, la estrategia definida, etc. En la actualidad hay disponibles cuatro gestores de diálogo adaptados a la filosofía de comunicación del sistema. Dos de ellos están basados en reglas, uno de ellos es estocástico y el último se basa en un modelo mixto.
4. Módulo de consulta a la base de datos de la aplicación (MCBD): Recibe peticiones de consulta a la base de datos por parte del gestor de diálogo, las procesa y devuelve el resultado al gestor. El sistema utiliza una base de datos en Post-GreSQL.
5. Módulo de generación de respuestas (GR): Tiene como función la generación de una frase, gramaticalmente correcta y en un lenguaje lo más cercano posible al lenguaje natural, que transmita el mensaje generado por el gestor de diálogo. La estrategia que sigue el sistema consiste en la utilización de plantillas de respuestas adaptadas a los tipos de consulta.
6. Módulo de síntesis texto-voz: Transforma la frase de respuesta al usuario en

señal de audio. Se ha realizado la integración de un sintetizador en castellano desarrollado expresamente para el proyecto.

7. Módulo central gestor de comunicaciones: Este módulo es el responsable de establecer las comunicaciones con el resto de módulos del sistema, encargándose de recibir la totalidad de mensajes enviados (excepto los de audio) y de encaminarlos al servidor destino.

3. *Formato de comunicación entre los módulos.*

La comunicación entre los diversos módulos se lleva a cabo mediante el envío de paquetes en formato XML a través de sockets dedicados. La Figura 2 muestra un ejemplo de paquete XML, dirigido del módulo de comprensión al gestor de diálogo, en el formato de comunicación desarrollado.

```
<?xml version="1.0" encoding="ISO-8859-15"?>
<paquete destino="GD" origen="CH" version="1.0">
  <comprension>
    <segmentacion_semantica>
      <segmento etiqueta="consulta" valor="quiero">
        <confianza valor="0.0219018941131"/>
        <confianza valor="0.126746506986"/>
      </segmento>
      <segmento etiqueta="hora_s" valor="horarios">
        <confianza valor="0.14695464419"/>
        <confianza valor="0.0763993948563"/>
      </segmento>
    </segmentacion_semantica>
    <frames>
      <frame tipo="(HORA-SALIDA)">
        <confianza_frame valor="0.14695464419"/>
        <confianza_frame valor="0.0763993948563"/>
      </frame>
    </frames>
  </comprension>
</paquete>
```

Figura 2: Ejemplo de paquete definido para la comunicación en el sistema.

4. *Objetivos de la demostración.*

La demostración se centrará en una descripción de la arquitectura desarrollada para la implementación del sistema de diálogo, así como de las características más relevantes de los módulos descritos en esta comunicación. Seguidamente, se llevarán a cabo diferentes diálogos con entrada textual y vocal, a partir de los cuales pueda realizarse una visión del sistema y una comparativa entre los diferentes gestores de diálogo disponibles en la actualidad.