

Sistema *Mayordomo*: Uso de un Entorno de Inteligencia Ambiental a Través de un Sistema de Diálogo Multimodal

Mayordomo system: Using an Ambient Intelligence Environment Through a Multimodal Dialogue System

G. Espejo*, N. Ábalos, R. López-Cózar, Z. Callejas

Dpto. de LSI, CITIC-UGR, Universidad de Granada

18071 - Granada

{gonzaep,nayade}@correo.ugr.es, {rlopezc,zoraida}@ugr.es

D. Griol

Universidad Carlos III Madrid

28911- Madrid

dgriol@inf.uc3m.es

Resumen: *Mayordomo* es un sistema de diálogo multimodal que permite la interacción con un entorno de Inteligencia Ambiental para una vivienda. Este sistema permite la manipulación de los electrodomésticos instalados en la vivienda de manera oral, escrita y a través de una interfaz gráfica. También incluye una serie de funcionalidades adicionales, incluyendo, acceso identificado de usuarios, control parental e historial de acciones realizadas por los usuarios.

Palabras clave: Sistema de Diálogo, Inteligencia Ambiental, Reconocimiento del Habla

Abstract: *Mayordomo* is a multimodal dialogue system designed to interact with an Ambient Intelligence environment implemented in a home. The system can control appliances installed in a house through speech, text and a GUI interface. It also includes a number of additional features, including system access for identified users only, parental control and record of actions carried out by users.

Keywords: Dialogue system, Ambient Intelligence, Speech Recognition

1 Introducción

La Inteligencia Ambiental (AmI) es un campo de desarrollo reciente que se centra en la incursión de tecnología en espacios habitables (viviendas, oficinas, museos, etc.) de manera que facilite y mejora las prestaciones y comodidad de dichos espacios para los usuarios. Esta incursión tecnológica ha de ser lo más transparente posible, haciendo de esta manera invisible la tecnología al usuario (IST Advisory Group, 2001)

Por otro lado, un sistema de diálogo es una aplicación informática en la que la interacción persona-ordenador se hace de manera oral y en lenguaje natural. Un sistema multimodal añade la peculiaridad de que la interacción humano-ordenador puede hacerse mediante varias vías de comunicación. Esta interacción multimodal no tiene por qué ser expresamente verbal. Podría interactuarse con el sistema informático a través de gestos, movimientos corporales o expresiones faciales.

2 Sistema de Diálogo *Mayordomo*

Mayordomo es un sistema de diálogo multimodal que permite la interacción con los distintos electrodomésticos de una vivienda, así como la consulta sobre el estado de los mismos. Esta interacción puede hacerse de manera oral o escrita mediante lenguaje natural, o a través de una interfaz gráfica. En este trabajo se presentan las principales características del sistema en cuanto a reconocimiento automático de habla, comprensión de habla, gestión del diálogo, generación de frases, síntesis de voz, e interacción mediante interfaz gráfica.

2.1 Reconocimiento Automático del Habla

El sistema usa el software *Windows Vista Speech Recognition*, que incluye tanto el motor de reconocimiento automático de habla (RAH) como el sintetizador de voz que permite realizar

* Este trabajo ha sido financiado por el proyecto HADA TIN2007-64718

la conversión texto-habla (TTS, Text-To-Speech). Windows Vista incluye dos herramientas para programadores: SAPI 5.3 (Speech API) y System.Speech (espacio de nombres de .NET Framework 3.0).

2.2 Comprensión del Habla

La comprensión de habla se realiza usando *frames* (Allen, 1995). Cada frame consta de una serie de huecos que incluyen información necesaria para el procesamiento de órdenes o consultas de los usuarios.

El campo *Acción*, es el único que se debe rellenar obligatoriamente. Consta de cuatro partes: habitación, electrodoméstico, atributo y valor. Así, al analizar la frase “Switch on the light in the kitchen” (Enciende la luz en la cocina), *Mayordomo* reconocería como habitación *cocina*, como electrodoméstico *luz*, como atributo el *estado* del dispositivo, y como valor *encendido*. En caso de que uno de los elementos de *Acción* no se haya reconocido, *Mayordomo* preguntará por aquél sobre el que dude.

2.3 Gestión del Diálogo

Una vez se han rellenados todos los campos de la estructura *Acción*, *Mayordomo* procesa la información en la misma. Como resultado, puede ejecutar una acción sobre un electrodoméstico, o bien proporcionar al usuario la información que éste solicita (ver Figura 1).

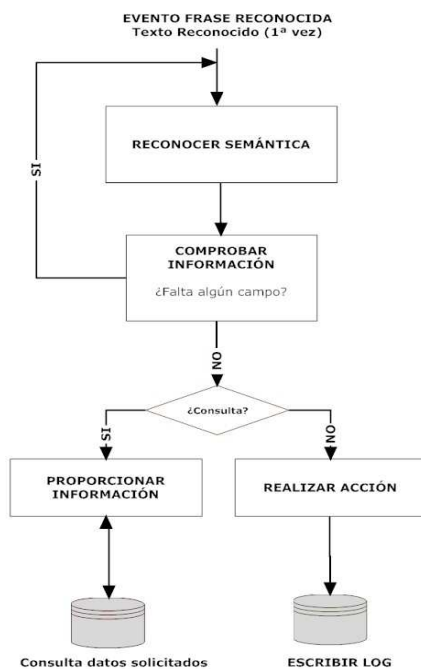


Figura1. Proceso seguido por el gestor del diálogo

2.4 Generación de Frases

Para generar las respuestas, *Mayordomo* utiliza un conjunto de patrones que se instancian con unos valores u otros según el electrodoméstico, habitación, atributo y valor implicados. Por ejemplo, “You have changed to (valor) the (atributo) of the (electrodoméstico) in the (habitación)”.

2.5 Síntesis de Voz

El sistema utiliza síntesis de habla mediante conversión texto-habla (Text-To-Speech, TTS) para comunicarse de forma oral con el usuario. Para ello usa el software *System.Speech.Synthesis*.

2.6 Interfaz Gráfica

Mayordomo posee una interfaz gráfica para la interacción con los electrodomésticos (ver Figura 2). Esta interfaz supone otro modo de interacción con el sistema.

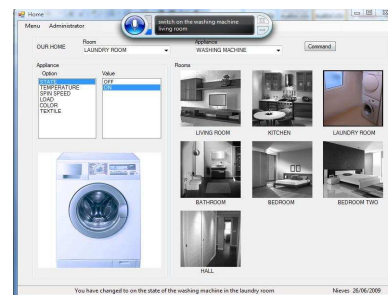


Figura 2. Interfaz gráfica de *Mayordomo*

Esta interfaz proporciona una serie de menús, a través de los cuales se gestionan funcionalidades de la aplicación (administración de usuarios, control parental, etc.), y pantallas, que permite la navegación por las distintas habitaciones de la vivienda. Dentro de cada habitación permite, a su vez, la gestión de las funcionalidades (atributos) de los electrodomésticos, facilitando el cambio de valor de cada uno de ellos.

Una barra de estado en la parte inferior muestra por escrito cada una de las frases emitidas de manera oral por el sistema.

Bibliografía

Allen J. 1995. Natural language understanding. The Benjamin/Cummings Publishing Company Inc.

IST Advisory Group. 2001. The european union report, scenarios for ambient intelligence in 2010.