FALA 2010
VI Jornadas en Tecnología del Habla and II Iberian SLTech Workshop

A System for User Location by Means of RFID Devices for a Dialogue System Designed to Interact in an Ambient Intelligence Environment

Gonzalo Espejo[1], *Nieves Ábalos*[1], *Ramón López-Cózar*[1], *Zoraida Callejas*[1], *David Griol*[2]

[1] Dep. of Languages and Computer Systems. CITIC-UGR. University of Granada
[2] Computer Science Department. Carlos III University of Madrid
{gonzaep,nayade}@correo.ugr.es, {rlopezc,zoraida}@ugr.es, dgriol@inf.uc3m.es

Abstract

This paper presents our current work in locating users through RFID devices for a multimodal dialogue system called *Mayordomo*. The paper explains how RFID works and describes the system architecture to connect hardware devices (RFID) and software applications (*Mayordomo)*. Finally several examples are shown to explain how this implementation improves the dialogue between the system and users.

Index Terms: speech recognition, multimodal dialogue system, ambient intelligence, RFID.

1. Introduction

Ambient Intelligence (AmI) is a research area that has attracted a lot of efforts by the scientific community in the last years [1, 2, 3, 4, 5, 6]. The aim of AmI is to create environments in which users are able to interact in a natural and transparent way with systems that help them carrying out their daily leisure and work activities.

AmI has been developed for some different environment, such as medical fields [7], home [8], public spaces [9] and learning spaces, as museums [10].

One of the most important characteristics in Ambient Intelligence systems is proactivity. This means that the system should predict the behaviour and preferences of the users to anticipate them.

In this paper we propose the automatic location of users for a multimodal dialogue system in a home environment through RFID devices. This way, the system is able to adapt its behaviour in several situations once it knows where the user is. For example, in spoken interfaces like ours, the system can vary the dialogue with the user so that he does not have to explicitly provide information about his location.

RFIDs are considered one of the main ways for physical browsing, which is an interaction paradigm that associates digital information with physical objects [1]. RFID devices have been used in several research projects concerned with user localization. For example, [11] used RFID devices to locate students in an educational environment, whereas [2] employed these devices in a home environment. Also, in [12] RFID technology is used in a medical environment.

The paper is organized as follows. In section 2, we describe some features of our *Mayordomo* multimodal dialogue system, focusing on the speech-based interaction module of automatic speech recognition. Section 3 explains some features about RFID devices and how they work. Next, we describe the overall system architecture in order to explain how hardware and software are connected and work together. Finally, conclusions and future work are presented.

2. *Mayordomo* Multimodal Dialogue System

2.1. General Overview

Mayordomo is a multimodal dialogue system developed in our laboratory, to be integrated in an Ambient Intelligence environment. The goal of *Mayordomo* is to centralize the control of the appliances located in a home. The interaction with the appliances can be carried out through spontaneous speech or with the GUI interface shown in Figure 1.

In addition to handling appliances, the system supports different types of users, depending on the administrator level and the experience with *Mayordomo*. The system administrator has privileges to perform special actions, for example, installing and uninstalling appliances Also, restrictions are allowed to some users. For instance, parents can forbid that their children watch TV after 10 p.m. Moreover, the system creates a log of all actions carried out within the environment by any user which is only accessible to the administrator user.

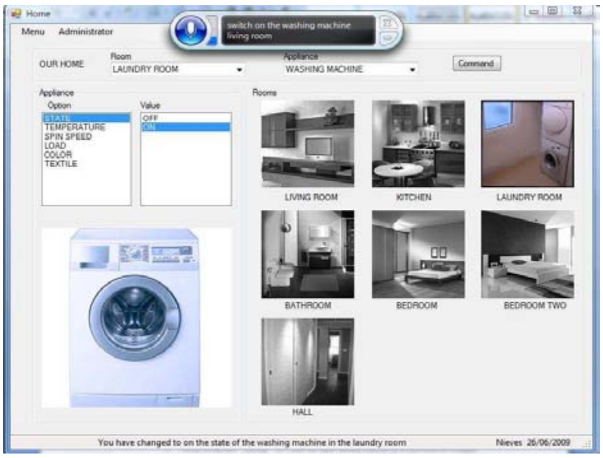

Figure 1: *GUI interface of* Mayordomo.

Scalability is one of the main characteristics of *Mayordomo*. Different appliances can be installed dynamically, and a reboot of the system is not necessary. Scalability is possible due to configuration and grammar files.

2.2. Speech-based Interaction

We use Windows Vista Speech Recognition (WVSR) to implement the speech-based interaction. This package includes both the engine for automatic speech recognition (ASR) and the engine for text-to-speech synthesis (TTS).

Figure 2, shows the detailed architecture of the speech-based interface, focusing on the dialogue manager we have implemented.

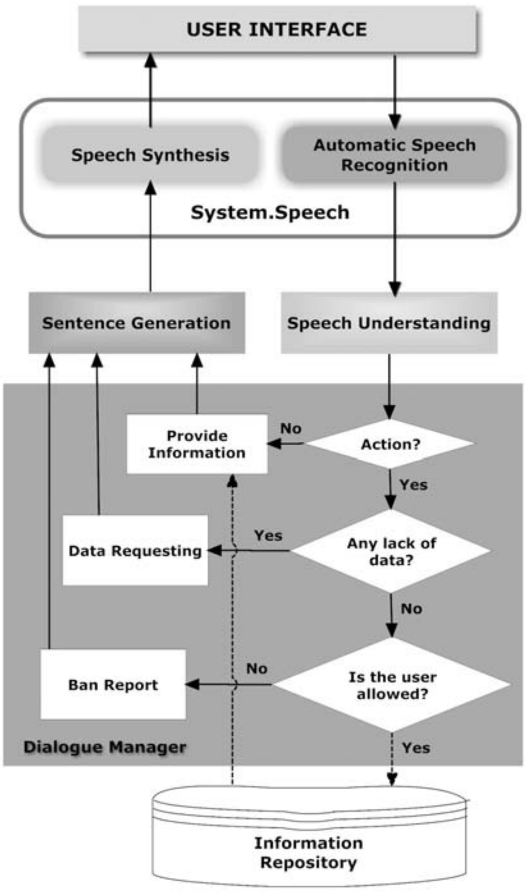

Figure 2. *Speech-based interaction in Mayordomo*.

As mentioned above, each appliance has an associated configuration file that allows the user to control it orally. This file contains a specific grammar for interacting with the appliance that is used for ASR. This grammar is specified in SRGS (Speech Recognition Grammar Specification) format[1].

When specifying grammars for the different home appliances of a home, we can consider three strategies: (1) allowing keyword recognition using the specific subrule *keywords*, (2) allowing keyword recognition without using this subrule, (3) not allowing keyword recognition.

Using the first strategy (which is the one used by the system) and the second one, grammars allow the recognition of keywords. The main difference between these two strategies lies in the way that the recognition is carried out. Using the first one (see Figure 3), the initial rule of the grammar contains four subrules: *order*, *sentence*, *request* and *keywords*. The subrule for *keywords* includes all the words related to the control a particular appliance. For example, keywords related to the T.V. may be concerned with the place where this appliance is (e.g. *living room* or *kitchen*), the attribute or characteristic the user wants to change (e.g. *volume* or *channel*) and the action to be performed with the appliance (e.g. *switch on* or *turn off*).

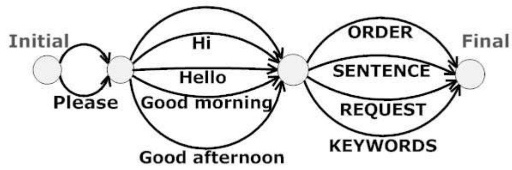

Figure 3. *Initial rule of ASR grammar including a subrule for recognition of keywords*.

This strategy is the most suitable for users who do not provide all the data required to perform an action with an appliance. If any data is missing to perform the action, the dialogue manager of the system prompts the user for the missing information. In this case, the user can provide just the missing data, that is, he does not need to utter the entire order. The advantage of this strategy is that the interaction is more comfortable for the user, particularly if the orders are complex and long. For example, if when processing the order:

"*Set the temperature of the washing machine of the laundry room to thirty degrees*"

the system does not understand the room where the appliance is located, *Mayordomo* may prompt "*Where?*" and the user may answer "*In the laundry room"*.

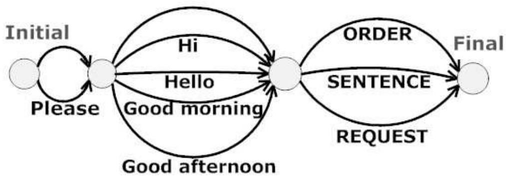

Figure 4. *Initial rule of ASR grammar without subrule for keywords.*

Using the second strategy for ASR (see Figure 4), grammars also allow the recognition of keywords. However, in this case, they do not use the specific subrule for *keywords* and all the elements of the subrules *order*, *sentence* and *request* are optional. An optional element in a rule means that it can be provided or not by the user, thus, providing it is not necessary to trigger the rule. For example, the subrule *order* has three elements: verb prep (subrule with a list of elements such as *"switch on"*), object (i.e. *"the light"*) and where (subrule with the rooms, for instance, *"of the kitchen"*). In this rule, if the element *where* is optional, the phrase *"Switch on the light"* will trigger the rule and will be recognized by the ASR. Therefore, this strategy permits all kinds of combination of words, resulting in a greater number of ASR errors as it allows incorrect combinations of words, for example, *"Hello washing off hall"*.

Using the third strategy for ASR, the initial rule of the grammar is the same as in the second strategy. However, it does not allow optional elements in the subrules. Thus, this strategy is not advisable when the sentences are complex and long.

Grammars are loaded into the system's memory at the beginning of interaction, and remain there throughout the dialogue so that the recognition engine can use them at any time during the interaction.

[1] http://www.w3.org/TR/speech-grammar/

3. RFID Devices

User location is known by the multimodal dialogue system through the information obtained from RFID devices, installed throughout the environment.

The interaction takes place using mobile terminals, such as mobile phones or cards that have a tag associated with a reader device. Each tag contains a universal resource identifier (URI, or Web address). The readers are distributed all over the environment. Each RFID reader has a serial or identification number. When one of the tagged objects is scanned by the reader, the tag and the identification number of the reader are sent to the dialogue system. The distance between the physical object and the RFID reader can range from a centimetre to four meters. Usually, in indoor and small environments this distance is not very large.

To be useful for our purposes, the dialogue system must know which user carries each tag. Taking into account this information, the system can find out which was the last RFID reader used by the user, and thus deduce his location.

Figure 5 shows how the different elements interact with each other. The smaller arrows represent the connection established between the RFID device and the server which manages the information. The big arrows show the way that the information flows from the user to the server. Firstly the user's card is scanned by a RFID device, and then the server receives the information related to the user from the device.

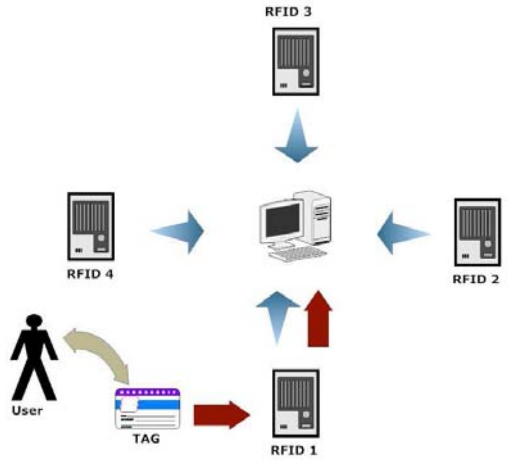

Figure 5. *Elements used in location though RFID*

These RFIDs devices used in *Mayordomo* are Phidgets[2]. We have selected them because the easy way to use and the possibility of extend the architecture to other devices in a simple way.

4. System Architecture

In this section we present the global system architecture. As it shown in Figure 6 the system in ready to use several hardware devices (RFIDs), and these devices are managed by some software application.

A middleware layer has been developed in order to connect hardware devices and software applications. This middleware layer consists in a repository in which information about devices, rooms and users is stored. Moreover, this information can be modified by hardware devices that are triggered when an event happens, for instance, when a device RFID detects a user's tag. In the same way, software applications are allowed to access the information stored with the aim of obtaining information about users or events occurring in the environment.

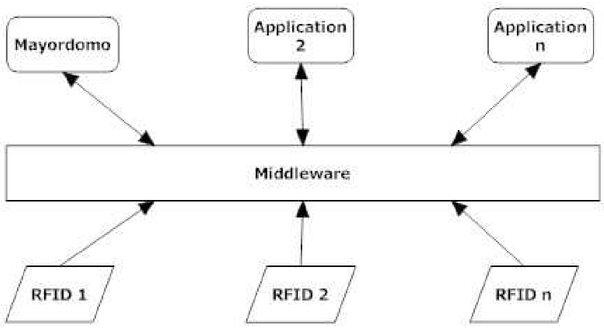

Figure 6. *System Architecture*

The information needed and stored in the middleware layer implemented in *Mayordomo* are a list of pairs users-tag, a list of pairs RFID devices-rooms and a list of RFID devices-tag triggered.

Using this information, the dialogue manager is able to detect who is the user and in which room is interacting with the system.

5. Automatic Location

The main goal of using RFID devices with a dialogue system is to reduce the information that the user must provide in order to provide shorter dialogues in an easy and natural way.

Knowing the room where the user is the dialogue becomes simpler. Comparing dialogue 1 and dialogue 2 with dialogue 3 it is clear how in the second one it is not necessary to explicitly provide to the system the room in which the user is located.

Starting with dialogue 1, the user asks the system to switch on the light, in a natural way, but the system does not know his location yet, and thus some additional turns are necessary for the user to specify his exact location.

User: Please, switch on the light.
System: Where?
User: In the living room.
System: You have changed to on the state of the light in the living room.
(The user goes to the hall)
User: Please, switch on the light.
System: Where?
User: In the hall.
System: You have changed to on the state of the light in the hall.

Dialogue 1.

In dialogue 2 the user includes in his order the room in which he wants to switch on the light, which makes the sentence longer and less natural than in dialogue 1.

[2] http://www.phidgets.com

User: Please, switch on the light in the living room.
System: You have changed to on the state of the light in the living room.
(The user goes to the hall)
User: Please, switch on the light in the hall.
System: You have changed to on the state of the light in the hall.

Dialogue 2.

Finally, in dialogue 3 the system includes RFIDs and these devices have been used by the user in order to identify where he is. This way of interaction is the best for the users: the system requires less information during the dialogue making it more friendly and natural.

(The user goes to the living room and uses the tag with the RFID device)
User: Please, switch on the light.
System: You have changed to on the state of the light in the living room.
(Now, the user goes to the hall and uses the tag with the RFID device)
User: Please, switch on the light.
System: You have changed to on the state of the light in the hall.

Dialogue 3.

6. Conclusions and future work

In this paper we have presented the current status of the design and implementation of location through RFID in our multimodal dialogue system.

We have described how RFID works, the global architecture and the main advantages using RFID devices. The architecture described has been designed to be used, not only with RFID, but with any other type of devices such as touch sensors or temperature sensors.

As it has been shown, using RFID devices less information is required in dialogues between users and our system. This feature makes the dialogues more natural as the user does not have to provide information which is implicit in the context of the interaction.

As future work we plan to include more devices in *Mayordomo*. In addition, we plan to include users' profiles describing their preferences and usual activities in order to make the system proactive. In this way it would be possible, for instance, to detect which is the favourite channel of any user in our environment.

7. Acknowledgements

This research has been funded by the Spanish Ministry of Science and Technology, under project TIN2007-64718 Adaptive Hypermedia for Attention to Different User Types in Ambient Intelligence Environments.

8. References

[1] Aghajan, H., López-Cózar, R, Augusto, J. C. "Human-centric Interfaces for Ambient Intelligence". Academic Press. 2010

[2] Haya, P. A., Montoro, G., Alamán, X. "A Prototype of a Context-Based Architecture for Intelligent Home Environments". CoopIS/DOA/ODBASE (1): 477-491. 2004

[3] Augusto, J. C. y McCullagh, P. "Ambient Intelligence: Concepts and Applications" Int'l J. Computer Science and Information Systems, vol. 4, no. 1, pp. 1–28.5

[4] Augusto, J. C. "Ambient Intelligence: Basic Concepts and Applications." Series: Communications in Computer and Information Science, Vol. 10. Springer Verlag, pp 14-24. 2008

[5] Ramos, C., Augusto, J. C., Shapiro, D. "Ambient Intelligence – The next step for Artificial Intelligence". IEEE Intelligent Systems
23(2) (March/April 2008) pp 15 – 18. 2008

[6] Remagnino, P. y Foresti, G. L. "Ambient Intelligence: A New Multidisciplinary Paradigm". IEEE Transactions on Systems, Man and Cybernetics, Vol. 35(1), pp. 1-6, Jan. 2005.

[7] Corchado, J. M.; Bajo, J.; de Paz, Y.; Tapia, D. I.: Intelligent environment for monitoring Alzheimer patients, agent technology for health care. Decision Support Systems, 44, pp. 382–396, 2008

[8] Pérez-Castrejón, E.; Andrés-Gutiérrez, J. J.: AAL and the Mainstream of Digital Home. Lecture Notes in Computer Science 5517, pp. 1070–1082, 2009

[9] Corchado, J. M., Pavón, J., Corchado, E., Castillo, L. F.: Development of CBR-BDI Agents: A Tourist Guide Application. Lecture Notes in Artificial Intelligence 3155, pp. 547-559, 2004

[10] Wakkary, R.; D. Evernden, Museum As Ecology: A Case Study Analysis Of An Ambient Intelligent Museum Guide, in J. Trant and D. Bearman (eds.). Museums and the Web 2005: Proceedings, Toronto: Archives & Museum Informatics, published March 31, 2005

[11] López-Cózar, R., Callejas, Z., Montoro, G., Haya, P. " DS-UCAT: Sistema de Diálogo Multimodal y Multilingüe Para un Entorno Educativo". Proc. IV Jornadas en Tecnología del Habla, pp. 135-140. . 2006

[12] Collins G. R. "Usable Mobile Ambient Intelligent Solutions for Hospitality Customers" Journal of Information Technology Impact, Vol 10, No 1. pp. 45-54. 2010