A Dynamic Spoken Dialogue Interface for Ambient Intelligence Interaction

Germán Montoro, Universidad Autónoma de Madrid, Spain
Pablo A. Haya, Universidad Autónoma de Madrid, Spain
Xavier Alamán, Universidad Autónoma de Madrid, Spain

ABSTRACT
In this article, we present the interpretation and generation processes of a spoken dialogue interface for ambient intelligence. The interface is automatically created for each specific environment and the interpretation and generation vary depending on the environment and its context. These processes rely on a dialogue tree structure. Several modules process the tree structure and the context information to produce specific dialogues for the current environment state. The interface has been implemented and evaluated in an ambient intelligence environment. Satisfactory objective and subjective evaluation results are shown at the end of the article.

Keywords: Ambient Intelligence, Dialogue Management, Dynamic Spoken Interface

1. INTRODUCTION
In recent years a new research area has appeared within the ubiquitous computing (Weiser, 1991) field under the name of active environments, intelligent environments or ambient intelligence. The aim of ambient intelligence is to provide a more natural interaction between the environment and its inhabitants. The environment must help people in their everyday life, offering more human-like ways of communication. Therefore classrooms, offices, laboratories and homes should be capable of assisting their occupants in their tasks. This interaction must be adapted to the task, the environment, its occupant and the available devices. This implies the offering for selection of a specific communication modality.

Nevertheless, not all the projects related with ambient intelligence consider the necessity of deploying interaction interfaces. This is the case of some projects of the Housing department (Munguia Tapia et al., 2004), which employ low level sensors to recognize the activity of its occupants. A similar idea is shown by the MavHome project (Das & Cook, 2004). Other projects, like The Adaptive House (Mozer, 2005), do not consider appropriate the environment has any kind of new interface.
On the other hand, some projects try to obtain more natural forms of communication to integrate them with the environment. That is the case of Aire (Adler and Davis, 2004), which has studied the possibilities for combining sketching with speech for multimodal design. Another project that has explored the use of speech for interacting with the environment is Homey (Milward & Beveridge, 2004). This project aims to carry out research on an intelligent dialogue interface designed to develop a dialogue between a tele-medicine interface and a patient. Considering the environment characteristics, this dialogue interface requires dynamic adaptation. Furthermore, the interaction can be multimodal. One of the main contributions to this field was the project Smartkom (Wahlster, 2006). This interface recognized speech or gestures and generated text, graphics or speech. Users could employ any of these modalities in three different scenarios: at home or in the office, at a communications booth and on the move with mobile devices.

In this article we present a Spanish spoken dialogue interface for ambient intelligence environments. A dialogue control structure is automatically created according to the specific environment and it allows to interact with the environment and control its devices by means of spoken language interaction. Adaptation occurs at the interface creation and interaction processes. In both cases the interface and its behaviour automatically vary depending on the environment and its state. Contextual information obtained from the environment is employed to assist dialogue processes such as simple pronominal anaphora resolution, sentence interpretation, or recognition error recovering.

The article is organized as follows. In Section 2 we introduce the concept of spoken dialogue interfaces in ambient intelligence environments. Section 3 presents the implemented ambient intelligence environment. We provide an overview of the environment representation in Section 4 and of the dialogue representation in Section 5. In Section 6 we give a more concise description of the interpretation and generation algorithms. Section 7 provides real examples of interaction. The interface evaluation is explained in Section 8. Finally we give some conclusion in Section 9.

2. DIALOGUE INTERFACES IN AMBIENT INTELLIGENCE ENVIRONMENTS

Although the presence of sound is not an essential characteristic for an ambient intelligence environment, we consider that a spoken dialogue interface is an important aspect in the development of these environments. Speech is a common, spontaneous and simple mean of communication (Clark & Brennan, 1991). This way, although it cannot always be the best input mechanism, it is a powerful method for the development of person-computer communication environments (Karat et al., 1999). This kind of interaction provides ambient intelligence environments with a more natural and intuitive way of communication. A continuous interaction in a daily occupied highly interactive environment without the possibility of using the voice could be a considerable effort and decrease significantly the capabilities of its occupants.

Moreover, a field research study carried out with real subjects to know their expectations about ambient intelligence environments shows that people prefer to employ their voice to control the home devices and, when they can choose between different modalities, they mainly choose oral communication (Brumitt & Cadiz, 2001).

Most of the spoken dialogue interfaces developed so far have focussed on the desktop classic environment or telephone-based agents for bank assistance, route planning or ticket reservation. These approaches have to be modified in the context of an ambient intelligence environment, where the interaction is addressed to a heterogeneous set of physical devices. Another differential factor for these interfaces is established by its idiosyncrasy. They are highly dynamic spaces whose configurations