

Natural Language Interface for Smart Homes

M. Fernández, J.B. Montalvá, M.F. Cabrera-Umpierrez, and M.T. Arredondo

Life Supporting Technologies

ETSIT Universidad Politécnica de Madrid

Ciudad Universitaria; 28040 MADRID, ESPAÑA

Tel.: (+34) 91 549 57 00 ext. 3407

Fax: (+34) 91 336 68 28

{mfernandez, jmontalva, chiqui, mtarredondo}@lst.tfo.upm.es

Abstract. The development of new ICT technologies, like Ambient Intelligence (AmI), and smart homes technologies has been proven to be a key factor to allow people with disabilities gain independency at their homes, vehicle or working environments. The biggest problem that users with disabilities face at the moment of using these technologies is the difficult use of their interfaces. In this paper we present the methodology and implementation used for the development of an interface with smart homes, based on natural language, which provides an easier way, especially for people with physical disabilities and the elderly, to perform the usual tasks at home without the need of a previous learning and complex processes.

Keywords: Smart Home, natural language, ambient intelligence.

1 Introduction

According to the European Union Statistics Office [1], around 17% of the European population is above 60 years old, and it is foreseen that this rate will increase up to 20-30% during the next years due to the improvements in the healthcare and the users' quality of life. On a European level, Spain is positioned as the 5th country with respect to the proportion of elderly people, and for the middle of the XXI century, it will be one of the oldest countries in the world. In this sense, a close relationship between ageing and disability exists: according to the Instituto Nacional de Estadística 31,2% of the Spanish citizens over 64 present some type of disability [2]. The same source reveals that the group of people with disability constitutes 8,5% of the national population, proportion that is maintained on a worldwide level. Such demographic changes imply an increase of the dependent population that is facing important problems for their active participation in the society as full members.

Most of the published reports about social support and dependency confirm the objective benefit, both emotional and physical for the dependent person, to continue living at home rather than moving to residences or sheltering houses. Currently in Spain, almost eight out of ten old people stay at home and in 76% of the situations, the relatives are in charge of providing support in order to achieve the daily activities

[2]. However, disabled people associations complain both about the insufficiency of assistive appropriated services and support to the family assistant.

The development of new ICT technologies, like Ambient Intelligence (AmI), can help to create adequate environments so that the claimed and necessary Independent Living can be achieved. The use of home automation technologies and assistive devices has been proven to be a key factor to allow people with disabilities gain independency at their homes, vehicle or working environments [3]. The biggest problem that users with disabilities face at the moment of using these technologies is the difficult use of their interfaces.

This paper presents a new modality of interaction with smart homes, based on natural language, which provides an easier way, especially both for people with physical disabilities and the elderly, to perform the usual tasks at home without the need of previous learning difficult processes [4].

2 Methodology

The human language is a very complex way of communication. There are many issues that must be considered when a language-based interaction is developed. For the implementation of the interface, we developed a specific methodology based on the combination of the user centered approach provided by Userfit [5] and the usability evaluation of spoken language dialogue systems [6]. The methodology consists of 4 steps that are further described below.

2.1 Step 1: Vocabulary Framework

This first step is used to analyze the context where the interface is going to be built, the elements of interaction, the actions and the users. It is based on three phases.

- Context analysis: What and Where

The context analysis is meant for answering the questions, What and Where. The What question is meant for identifying the different interaction elements. These are the blinds, the lights, and in general the home appliances that will be connected to the interface. The where question is meant for differentiating the elements that are the same or have the same function. Obviously two devices cannot be at the same place at the same time. For example: The light of the kitchen or the light of the bedroom.

- Task analysis. Which and When

The task analysis is used to identify the different actions that the user is able to perform with the interaction devices. For example programming the temperature of the room or turning off the lights.

- User profile: Who

One important question is who uses the interface. By identifying the person that has permission to use the interface, security access parameters can be applied, as well as preferences depending on the user's profile. This enables a personalization of the interface upon the preferences and needs of the user.

2.2 Step 2: Syntactic Construction

The next step determines which are the syntactic structures that should be taken into account. For one request there could be different syntactic structures, being the only difference among them the order of the words. If the user wants “to turn on the light of the bedroom”, the same user could ask the same request in different ways, and in order to be coherent with the natural language, a great amount of the most common syntactic structures have been developed. Some of these structures could be: turn on the light of the bedroom, from de bedroom turn on the light, please could you turn on the light.

2.3 Step 3: Semantic Interpretation

In the previous step there are some aspects that have not been taken into consideration. The order of the words is not the only important issue in the interface; it is more important to find a request with sense in relation to the connection between words and with the actual situation of the smart home. The system can accept a sentence like “switch on the window from the kitchen”, and the semantic interpretation of the interface understands the request. The same applies to the coherence between the request and the actual state of the home. The interface can understand an action that cannot be performed in this moment. For instance, the system has understood “switch off the TV” and the TV is already switched off.

2.4 Step 4: Answer Construction

In order to use the natural language to provide the system answers, a module that automatically generates the answers to each user’s request has been developed, taking into account the order of the words in the answers. Apart from the definition of the collection of words, the possible structures with the different combination were identified in order to make more natural the provided answers.

3 Implementation

The developed interface consists on a software application implemented with C programming language and Loquendo libraries. The voice recognition has two phases: the syntactic analysis and the semantic analysis. The action diagram is shown in Figure 1. When a request is executed, the feedback is provided to the user.

The syntactic analysis firstly defines the operation field of the request: if it is a query or an action, and its type. Secondly, it outlines the words that describe the request and provides with information to perform the demand (action, element, localization, parameter, value, etc.). The request structure recognized by the system has to match any of the defined grammars in use. Once the recognition is done, the system performs a semantic analysis to verify the correctness of the request. Then, the consistency of the result is established. In case of any inconsistency between the request’s components, the system provides an interaction module that exchanges questions (made by the system) and responses (made by the user) until it becomes a consistent request to be executed.

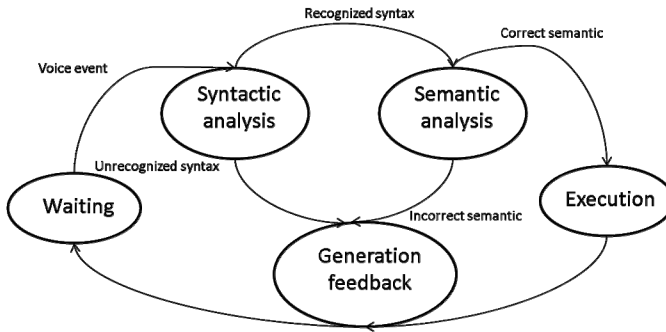


Fig. 1. Action diagram

Once the analysis is done and the request is consistent, two different behaviors are distinguished:

- If the request is to implement an action, the main function is the amendment of the relevant databases that origin an active task immediately (turn off a light, increase the volume of the television) or a delayed one (programming of the oven).
- If the request corresponds with a query (state, parameter or programming), the interface only looks for the value in the relevant database and provides the user with the answer.

3.1 Example of an Action

If we consider an action on a device that produces a change in its state, the request corresponds with the basic function that any smart home interaction system should have. For its analysis and treatment the so-called tree structures is used, consisting on a root that correspond with the system application, and successive levels of leaves that determine the components of the activity. The tree structure that defines the achievement of the request shows two different versions. In a first version the structure is centered in the element (see Figure 2), after its determination, it is analyzed the first level of leaves that correspond with the possible actions that can be carried out. The second level determines the localization of the element, to uniquely identify the device to perform the action.

In the second version, there is a tree structure centered in the action (see Figure 3), where the first level of leaves are the possible elements on which a task can take place. After that, a second level appears that determines the location to define the device uniquely.

It is in these parts of the tree in which it is not possible to determine the component of the next level when a man-machine dialogue is established to resolve any ambiguity.

Whether through the first or the second structure and after defining completely the request, a feedback is provided to the user indicating the result of the request, either indicating the action taken (current status of a device or modification of this state), or giving specific details (indicating the values of the parameters needed to program a device).

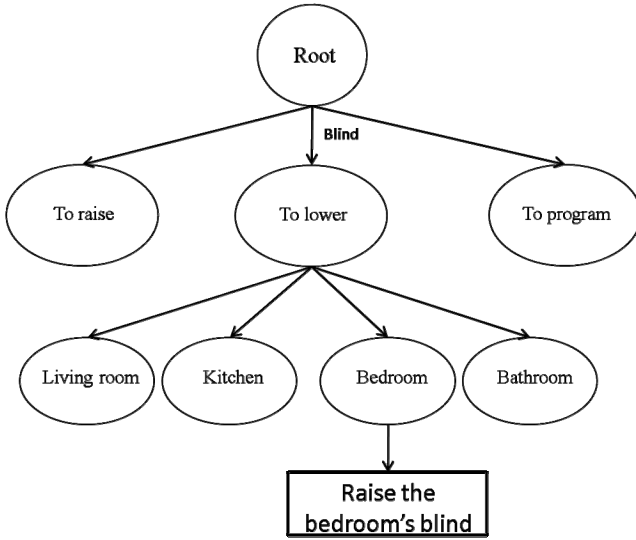


Fig. 2. Tree diagram centered in the element

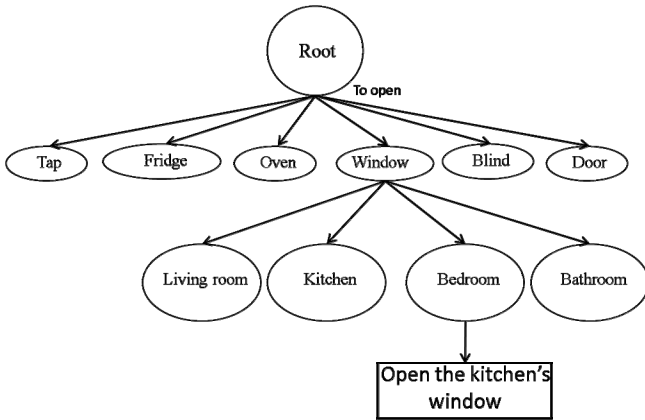


Fig. 3. Tree diagram centered in the action

The final objective is to simulate a conversation as natural as possible, and therefore giving variable feedback responses that randomly vary their structure. The three different answers before the conclusion of an action are composed of the initial part of the sentence (I have already..., I just..., I have...). Then, the general structure: Connector (I have already, I just, I have) + Action + Element + Room (in those cases where it is necessary to indicate it, because more than one device of the same type exists). For example, before a specific request like “turning on the light in the kitchen”, the possible system answers can be: I have already turned on the light of the

kitchen, I have turned on the light of the kitchen, I have just turned on the light of the kitchen.

3.2 Example of a Query

If we consider a system query about the state of a specific device, two possible queries are defined, conditioned by the type of request that the system has recognized.

A first type of query is centered in the element (see Figure 4) of a device, for example: How is the window of the kitchen? The syntactic structure that the system expected has the following components: a connector that differentiates this request from others (How is, Which is the state of, etc.), the device (indicating the element or the location in those cases of ambiguity) and the state the user wants (How is the light of the kitchen?, ¿Is the light of the kitchen turned on?). The interface’s feedback is the state of the device.

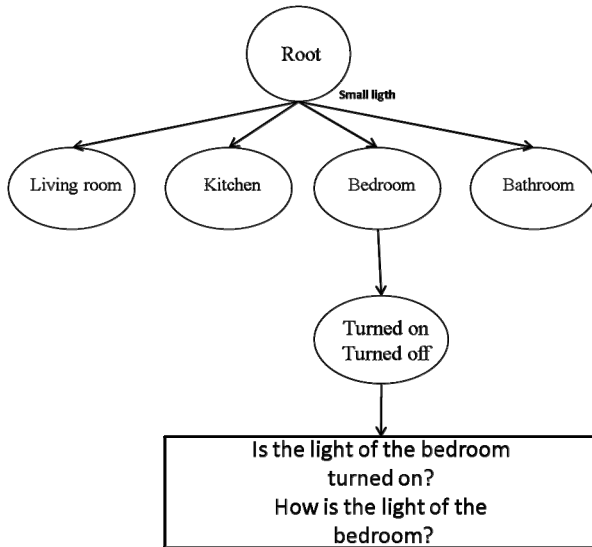


Fig. 4. Query centered in the element

A second type of query centered in the action (see Figure 5), in which the user wants to know which devices are in a specific state, for example: What device is switched on in the bedroom? In this case, the elements that define this request are similar to the above one: a connector (is it, what is ...), a state and a location if the user wants to limit his query to a specific room in the house. In this situation, the feedback offered is the list of devices that share this state. The tree structure is used to define the request.

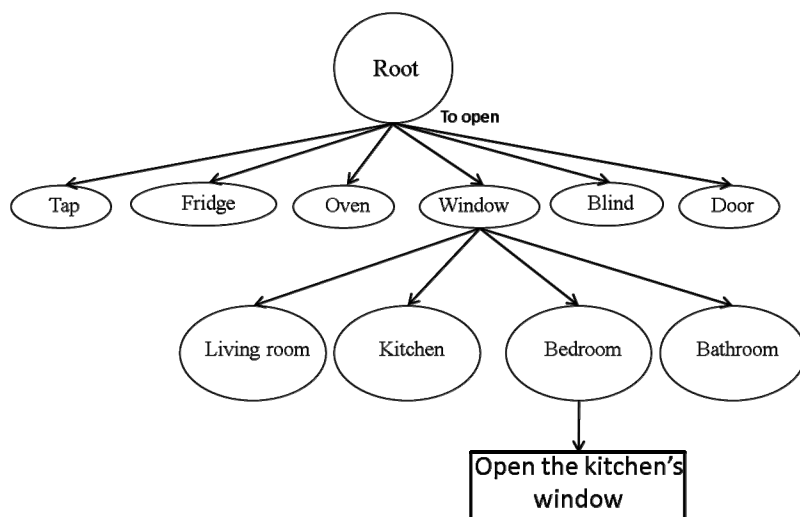


Fig. 5. Query centered in the action

4 Results

The validation of the system is currently being carried out with real users through a series of questionnaires and interviews, to get the information as possible not only related to the acceptance of the system, but also to the new recognition possibilities that have not been considered previously.

The evaluation consists of the following steps:

- The user selects the type of voice.
- The user performs different tasks according to the evaluation scenario based on the execution of different use cases.
- The user fills in a questionnaire based on qualitative and quantitative questions.
- After the interview, the experts fill in the evaluation questionnaire for each use case providing information about the user interaction with the system.

The main problem encountered during the evaluation has been the difficulty that the users experiment in differentiating a failure on a request where the system makes a correct recognition, but the quality of the recognition was not good enough, and the failure of the request when the system was not able to make a proper recognition due to a non-existent grammar.

The expected results after a massive evaluation will be related to the different aspects of its usability like: how many interactions have been required to conduct an ambiguous request, how much time the user has spent in a task, if there is any element or action that has not been taken into account, or if the user has had a problem with the natural language interaction.

6 Future Work

Further work will be needed to migrate the Spanish version of the system to other languages. The syntactic construction and semantic interpretation should be done to adapt them to the new language structures. It is important to take into account the vocabulary of each new language as well as the grammatical structures.

A module for biometric identification of the users is under development. It will enable the voice identification of the users, without the need of a specific password. This will enable the implementation of the user profile with the adaptation of the interface to security issues, users' preferences and needs.

5 Conclusions

The evolution of information technologies can facilitate the enhancement of the quality of life of dependent people both elderly and disabled. The difficulties these people may encounter with the use of these technologies rely in the way they interact with them.

The main advantage of the methodology used is the fast development of an easy and useful interface based on natural language, focused in the responses and reactions of the users after the evaluation. One of the most important points to emphasize from the design is the independence of operation with the user, giving simplicity and reducing the maximum possible learning and training, to get wider acceptance.

The different expected results imply various aspects of usability like; how many interactions have been required to conduct an ambiguous request, how long the user has used to do a task, is there any element or action that has not been taken into account, or if the user has a problem with the type of interaction based on natural language.

One of the big limitations of this system is that the final users are not familiarised with the use of new technologies. Additionally, the reluctance to change may be an obstacle for the daily life operation of such systems.

References

1. Eurostat regional yearbook 2008 (2008), <http://epp.eurostat.ec.europa.eu/>
2. Encuesta de Discapacidad, Autonomía personal y situaciones de Dependencia Instituto Nacional de Estadística (2008), <http://www.ine.es>
3. Montalvá, J.B., Rodríguez, A., Conde, E., Gaeta, E., Arredondo, M.T.: Middleware Architecture for Users Interfaces in Ambient Intelligent supporting Independent Living. In: I Workshop on Technologies for Healthcare & Healthy Lifestyle, Valencia, Spain (2006)
4. Landucci, L., Baraldi, S., Torpei, N.: Natural Interaction. A Remedy for the Technology Paradox (2008), <http://ercim-news.ercim.org/content/view/310/497/>
5. Saz, J.T.: El diseño centrado en el usuario para la creación de productos y servicios de información digital. Depto. CC. de la Documentación, Univ. de Zaragoza (2004)
6. Dybkjær, L., Bernsen, N.O., Minker, W.: Evaluation and usability of multimodal spoken language dialogue systems. *Speech Communication* 43, 33–54 (2004)