Dialog-based User Interface for a Smart Home System

Alan Liu[1]

[1]Department of Electrical Engineering
National Cheng Cheng University
Chiayi, 621, Taiwan

aliu@ee.ccu.edu.tw

Abstract. *A collection of patterns involving the use of a dialog-based user interface is presented in this study to aid the design of a spoken language interface for a smart home system. A smart home system is regarded as a system which assists its user for the needs within a space where the user resides. Speaking to the system is considered a natural way of communication, but there are many challenges involved. It is inefficient to develop the components like speech recognition, natural language understanding, dialog management, natural language generator, and speech synthesis from scratch. Patterns of designing voice command systems and dialog systems have been observed in a speech-based user interface system. We present an example of designing a dialog-based user interface through patterns to enhance the interaction between the user and the system.*

Categories and Subject Descriptors

•Software and its engineering → Software design techniques

•Software and its engineering → Design patterns

General Terms

Smart home systems, user interface, user interaction, patterns

Keywords

user interface, user interaction, smart home, dialog system

1. Introduction

A smart home system is regarded as a system which assists its user for the needs within an environment where the user resides. The autonomy of providing services is expected, but the user still is responsible for issuing requests in many situations because the detection of user needs by relying solely on observation is difficult. The input from the user is usually required, and there are many different approaches proposed, such as panel control, gesture, brain-signal control, or voice-command. If all services are initiated based on user requests, then the system becomes simply a passive system waiting for commands. It remains just a digital home system with appliances controlled by a computer, awaiting for user commands. Thus, a smart home needs to be proactive, and capturing both explicit and implicit user requests becomes a key research challenge. A smart interface in the form of intelligent assistants and companions has been considered by many researchers [FEDI][MTDM]. One of the most important goals is to provide a natural way for interacting with the intelligent system. Speaking to the system is considered natural, but audio signal processing itself is a challenge. In addition, natural language processing (NLP) afterwards is also a great issue while providing a meaningful feedback is critical to success.

In our study, we consider a dialog-based approach for providing a proactive interface for the user. By knowing the timing of when to talk to the user, the system could speak to the user on reminders or suggestions and also can prompt the user for questions. The goal is to let the user talk to the smart home with simple sentences to receive services, and the system can also collect feedback from the user at the same time. In designing such systems, we have observed that a pattern language can be used in explaining the transformation of a passive system to a more proactive system with the introduction of a dialog system. This paper introduces the Dialog-Based Interface pattern along with other possible patterns to be used in designing a proactive user interface for a smart home system.

2. Dialog-Based Interface

A voice-based user interface is usually proprietary. Having a pattern will reduce the complexity in leveraging such system in designing a smart home system. By understanding the problem of the needs in more proactive system interaction, we consider a dialog-based user interface as the solution for users to communicate with a smart home system.

2.1 Problem

Having only a voice-based interface comes little short in a smart home system.

A smart home system is supposed to be more autonomous compared to a digital home system which is more passive. Both systems rely on appliances which are controlled by a computer. Controlling the appliances autonomously in order to provide services to the user at a right moment is difficult because of uncertainty in detecting user needs in a context. Interaction with the user becomes necessary to ensure user needs, but without proper system initiatives, the system sits there and awaits for user commands.

The voice commands can be implemented as voice labels to corresponding functions of a system if the voice commands are limited to simple words. As illustrated in Figure 1, a command-based interface acts in a single direction, which is to send user commands to the smart home system. Such commands may be something like turning an appliance on and off or getting some information from the system. A command may be considered as a finger to turn on a switch. Thus, the user expects something to happen like lights being turned on, but does not care what the system is returning to the user.

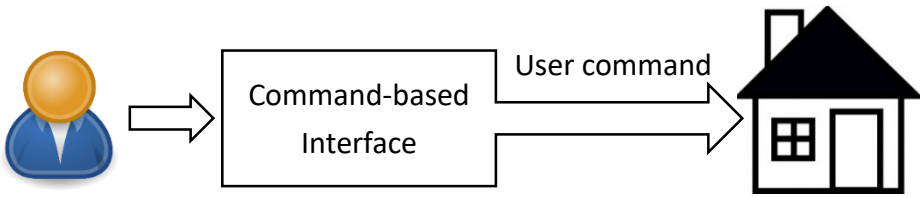

Figure 1. Command-based interface

If the system only accepts a preset command which is matched with voice label, then it becomes too restricted and predictable. In terms of interface, there is no distinction between a digital home and a smart home. A proactive interface should be considered to accompany more advanced functions that a smart home system provides.

2.2 Misfit and Forces

As pointed out in the problem, a simple command-based interface is unsuitable for a smart home system because it makes a smart home more restricted and passive. Other types of interface like gesture or brain control also work in a one-way direction shown in Figure 1. If we consider a smart home as an assistant or a companion, then the interface needs to be bidirectional. Speaking to each other is a common practice in human interaction. Thus, if a smart home system works like a human assistant, then one reasonable solution is to make it speak. One important question then is "when and what" to speak. In a simple command-response mode, the system may accept a word and match that word with a voice label and triggers a predefined responding message while performing the associated function. Another common response is that the system may double-check with the user for confirmation. However, these two scenarios do not reflect common interaction between two people and are not capable of handling more complicated user requests.

A smart home is expected to provide a more advanced service tailored to the user preferences, and these advanced services are the composition of basic services. Thus, using a simple command-function matchmaking will rely on the user to produce a sequence of orders or to predefine different sets of composite services.

If the system allows the user to speak not just commands but a sentence, then the voice input may contain a sequence of words which may need NLP to analyze it. Considering a service-oriented system, the system allows the user to make a request. Figure 2 shows the sequence of processing a user request in a sentence form, in which the sentence processing part usually involves NLP. After the user request is made known, service discovery and service composition take place before the system delivers the service. Services then can be considered as the responses from the system. Applying these concept in a smart home will result in a system like Figure 3 that shows this interaction.

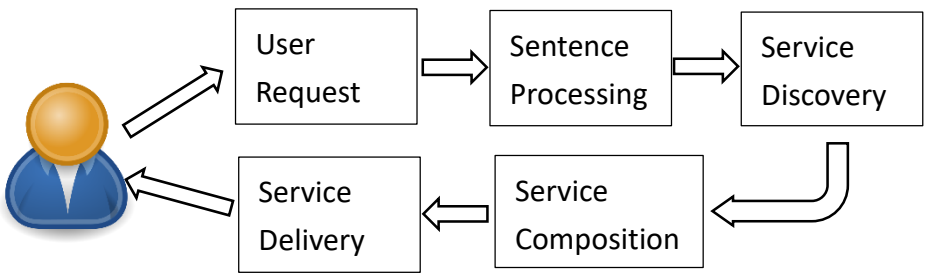

Figure 2. Response of a smart home system to a user request

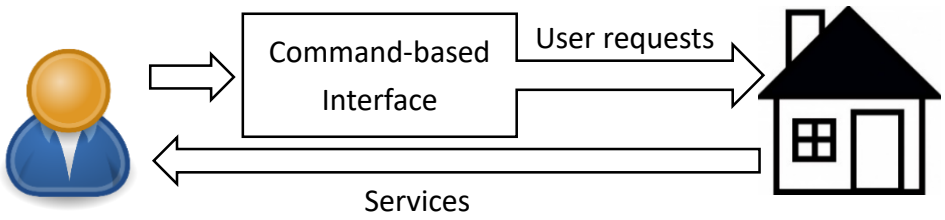

Figure 3. Command-based interface with a complex service as a response

To summarize the above observation, we can see that one-directional command-based interface (Figure 1) does not reflect a more natural interaction between two people. Speaking to a system needs a system to talk back to make the interaction a dialog. Another point is that the one-to-one command-function pairs may not be able to handle more abstract or personalized services. Thus, system initiation is important, but what to say remains a problem. Here, we list the forces in the following points:

- Speaking to the system is similar to asking a person for help.
- Hearing a system giving you suggestion or asking you questions may generate more productive interaction with the system.
- Using a dialog-based interface seems to be a natural way to engage the user in the interaction with the system.
- Handling a more advanced request should not depend on the user to construct all detailed requests.
- Confusion might occur if the user is not used to technologies.

2.3 Solution

A dialog-based interface is then proposed for a smart home system over a command-based interface, so that the system can talk to the user. The purpose is to let the system initiate interaction, so that the system is not just a receiving side of the commands. Figure 4 shows different types of messages that a system can talk to the user in order to receive a response. As an assistant, the system could ask the user for service requests and the system could also give suggestions or recommendation to the user as a steward. In addition, the system could take a role of companion to have a conversation with the user like giving greetings or engaging in conversation.

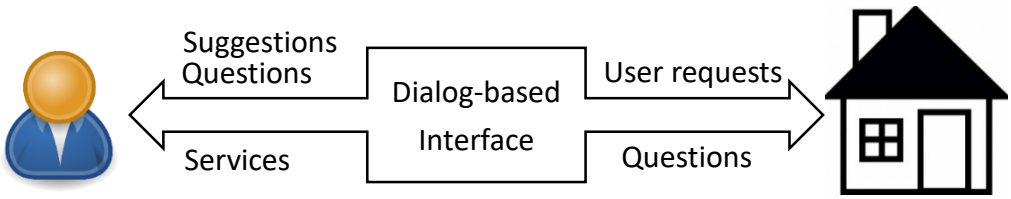

Figure 4. Dialog-based interface

If the system becomes more proactive, then it may initiate interaction like issuing a statement of suggestion to the user, asking for answers, or soliciting questions as shown in Figure 4. As the sentences become more complex or the conversation becomes more involved, then knowledge is certainly needed to support the dialog feature of the system. For our ongoing research in smart home systems, we have developed a dialog system which initiates interaction with the user [HULZ].

In the process of studying dialog-based interface systems, we have relied on a common practice called automated speech recognition (ASR) in designing a voice command system [VLSA] as shown in Figure 5. For processing voice input, usually microphone is used in capturing the

input signal that is sent to the ASR system. By analyzing the requirements of ASR systems [VCPM], we observe patterns that reside in design consisting of a sequence of components shown in Figure 6. The signal then is used for speech-to-text processing in order to produce a sequence of words, and NLP is performed for treating such sequences of words to produce meaningful information.

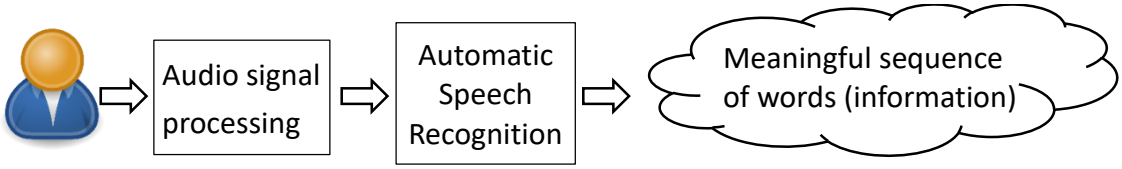

Figure 5. Processing spoken words to information

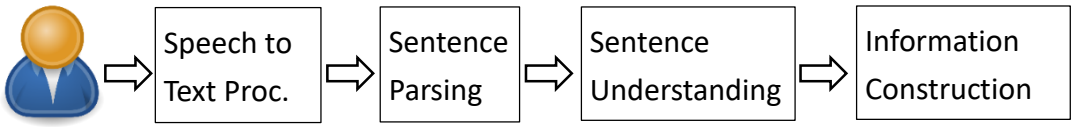

Figure 6. Automatic Speech Recognition

Besides the above, a smart home system also needs to consider context information and user preferences in order to provide a personalized service to the user in a timely manner. That will involve more implicit information which the system needs to discover. The system can also collect such information from the user. However, asking too many questions to the user makes the interaction awkward. Thus, a dialog-based interface system depends on knowledge about language, content, service, and context. Such knowledge is processed by an ASR system, a context-aware system, a service matchmaker, and a dialog system.

According to [JOMC], there are five important components in a dialog system: speech recognition, natural language understanding, dialog management, natural language generator, and speech synthesis. In our research, we integrate third party speech recognition, speech generator, and natural language understanding systems as a language interpreter as in Figure 7 and implement other components to support a dialog-based interface system for a smart home. Other components like home information manager, device controller, service planner, and service broker are needed for service provision in our system [HULZ].

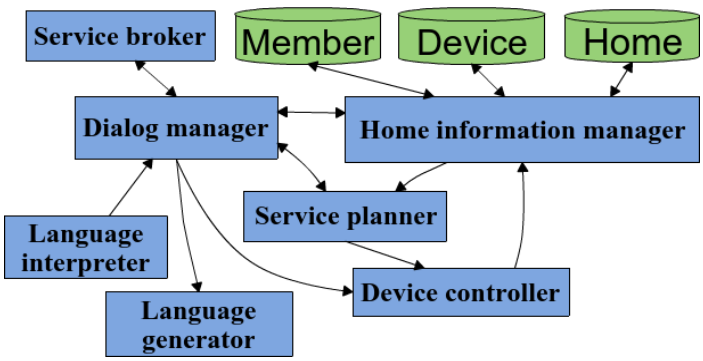

Figure 7. System architecture

If a user issues a service request, the system needs to know which available appliances are capable of delivering such service. If the user orders an appliance to do something, the permission and availability of the corresponding appliance is checked. The system needs to check the location or the identity of the appliance if more than one similar appliances exist. Thus, information concerning appliances and users along with context information is considered.

Figure 8 shows a class diagram which represents the main components considered in our system. The dialog manager is at the center for interacting with other components. The language interpreter and the language generator act as the user interface, and the rest of the components are responsible for service provision through controlling the appliances.

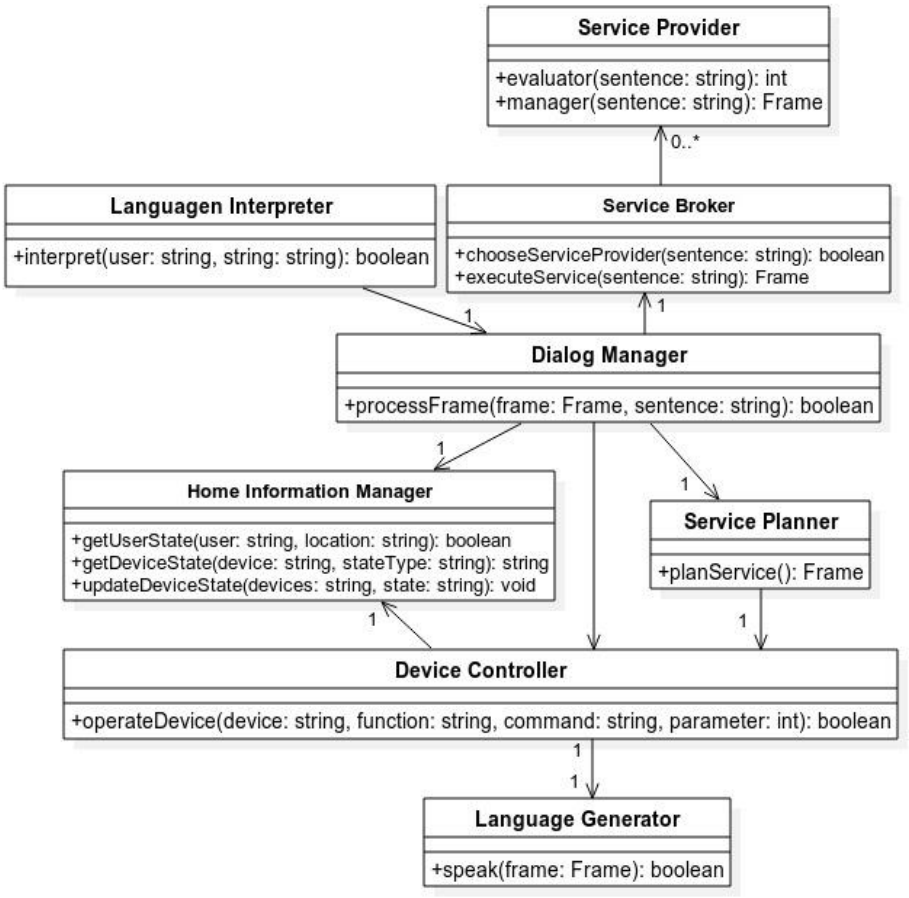

Figure 8. Relationship between key components

2.4 Related Patterns

The development of a smart home involves technologies from different fields, and the system needs to be adaptive and extendible. The need of a software process for developing a smart home has led to the piecemeal service engineering (PSE) approach [OVKU]. The goal of PSE is to reuse design practices in development of applications in a smart space. In addition, a design pattern suitable for developing an Ambient Assisted Living (AAL) system is considered in [NIHC]. The taxonomy of activities (or activity classification) is used for organizing a design pattern. The purpose is to include the special needs of elderly people, so that the integration of infrastructure can be adjusted to the constraints of the users. For enhancing the user interaction, the involvement of NLP and speech processing fields in the assistive technology has been growing for the past few years [PCRA]. The semantic web field has also been active in AAL for considering the challenges of web accessibility by elderly people, in which multilingual applications are provided [ANAS].

The known use of patterns are observed in the Sweet Home Project, which uses microphones in rooms to collect voice commands from user [VALP]. The main challenges are distant speech recognition and sentence spotting, in which language and acoustic models are combined with ASR systems. By using individual microphones in each room, the system needs to select best channel dynamically to adapt to different users [VLSA]. This project has claimed to be the first

in having people use their system in daily life. A few elderly people and a couple of visual impairment users have tried the interface, and many more challenges are identified from the feedback including unexpected ways of conversation with the system [VCPM]. A great effort of detecting and processing user input has taken place in the Sweet Home Project. Many other subsystems are developed in processing services. The integration of different components and subsystems requires flexibility in meeting different needs.

As for supporting patterns, we can focus on the following patterns which are the initial members in the pattern language for a dialog system for smart home systems. The detail definition of these patterns will be the future study of the pattern language.

Speech2Text and Text2Speech. Input and output through dialog can be processed by a standalone system which provides a speech to text service and a text to speech service.

Commands for simple services. There are services like turning on and off a device or opening or closing a door that can be provided without interaction. These services can be processed with simple commands

Dialog as a two-way street. Whether answering a question asked by a user or giving a comment to the user need be an interaction rather than a one directional input or output of a system.

Talk to the user. The system initiative is an important trait of a system as a companion. A proactive service of bringing up conversation is needed.

Know the background knowledge. In order to provide a better service to the user or to have a meaningful conversation, a system needs to have certain background knowledge of the user in particular context.

3. Conclusion

In this paper, we have introduced a pattern of integrating a dialog system and a context aware system in enhancing the interaction between the user and a smart home system. The main goal is to improve the user-machine interaction from one direction to both directions through dialog. By leveraging the features of a dialog system, we could make the system take the initiative to talk to the user instead of waiting for the user to initiate interaction.

Many researchers have studied speech-based user interaction for a smart home because of its naturalness of issuing commands to the system. When user requests involve more than just issuing commands but more of asking questions or requesting services, a simple interface is not enough since it only allows ordering the system to do something. As the system response is considered, the dialog-based user interface can support these requirements by both providing responses in speech and services to the user. With personalized knowledge and context information in a smart home, more thorough services and information can be given to the user. The structure of dialog-based information can be used in delivering such services.

Acknowledgements

Thanks go to my shepherd, Yu-Ching Cheng, for his valuable comments and feedback during the AsianPLoP 2016 shepherding process and also to the members in 2016 AsianPLoP Writers Workshop Group, Alice, Ayaka, Erik, Hiro, Kei, Takashi, Tetsuro, Tomoki, Yuma, for their valuable comments.

This research is supported in part by grant from Ministry of Science and Technology, Taiwan, under contract number 104-2221-E-194 -059.

References

[ANAS] Anastasiou, Dimitra. 2015. “Multi-Lingual and -Modal Applications in the Semantic Web: The Example of Ambient Assisted Living.” *Advances in Computer Science : An International Journal* 4 (4): 27–36.

[FEDI] Ferreira, J. F., and J. Dias. 2014. “Attentional Mechanisms for Socially Interactive Robots--A Survey.” *IEEE Transactions on Autonomous Mental Development* 6 (2): 110–25. doi:10.1109/TAMD.2014.2303072.

[HULZ] Huang, C. C., A. Liu, and P. C. Zhou. 2015. “Using Ontology Reasoning in Building a Simple and Effective Dialog System for a Smart Home System.” In *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*, 1508–13. doi:10.1109/SMC.2015.267.

[JOMC] Jokinen, Kristiina, and Michael F. McTear. 2009. *Spoken Dialogue Systems*. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers. http://dx.doi.org/10.2200/S00204ED1V01Y200910HLT005.

[MTDM] Mayer, Simon, Andreas Tschofen, Anind K. Dey, and Friedemann Mattern. 2014. “User Interfaces for Smart Things -- A Generative Approach with Semantic Interaction Descriptions.” *ACM Trans. Comput.-Hum. Interact.* 21 (2): 1–25.

[NIHC] Ni, Qin, Ana García Hernando, and Iván de la Cruz. 2015. “The Elderly’s Independent Living in Smart Homes: A Characterization of Activities and Sensing Infrastructure Survey to Facilitate Services Development.” *Sensors* 15 (5): 11312–62. doi:10.3390/s150511312.

[OVKU] Ovaska, E., and J. Kuusijarvi. 2014. “Piecemeal Development of Intelligent Applications for Smart Spaces.” *Access, IEEE* 2: 199–214. doi:10.1109/ACCESS.2014.2309396.

[PCRA] Portet, François, Heidi Christensen, Frank Rudzicz, and Jan Alexandersson, eds. 2015. “Perspectives on Speech and Language Interaction for Daily Assistive Technology: Overall Introduction to the Special Issue Part 3.” *ACM Trans. Access. Comput.* 7 (2): 1–8.

[VCPM] Vacher, Michel, Sybille Caffiau, François Portet, Brigitte Meillon, Camille Roux, Elena Elias, Benjamin Lecouteux, and Pedro Chahuara. 2015. “Evaluation of a Context-Aware Voice Interface for Ambient Assisted Living: Qualitative User Study vs. Quantitative System Evaluation.” *ACM Trans. Access. Comput.* 7 (2): 1–36.

[VALP] Vacher, Michel, Benjamin Lecouteux, and François Portet. 2015. “On Distant Speech Recognition for Home Automation.” In *Smart Health - Open Problems and Future Challenges*, 161–88. http://dx.doi.org/10.1007/978-3-319-16226-3_7.

[VLSA] Vacher, Michel, Benjamin Lecouteux, Javier Serrano-Romero, Moez Ajili, François Portet, and Solange Rossato. 2015. “Speech and Speaker Recognition for Home Automation: Preliminary Results.” In *8th International Conference Speech Technology and Human-Computer Dialogue ”SpeD 2015”*, 181–90. Proceedings of the 8th International Conference Speech Technology and Human-Computer Dialogue. Bucarest, Romania: IEEE. https://hal.archives-ouvertes.fr/hal-01207692.