International Conference on Control, Automation and Systems 2008
Oct. 14-17, 2008 in COEX, Seoul, Korea

Implementation of Dialogue System for Intelligent Service Robots

Changyoon Lee[1] , You-Sung Cha[2] and Tae-Yong Kuc[3]

[1] Department of Electrical and Computer Engineering, Sungkyunkwan University, Suwon, Korea
(Tel : +82-31-290-7202; E-mail: leech84@ece.skku.ac.kr)
[2] Center for Robot Industry Promotion, Korea Institute of Industrial Technology, Ansan, Korea
(Tel : +81-31-400-3848; E-mail: webchaly@gmail.com)
[2] School of Information and Communication Engineering, Sungkyunkwan University, Suwon, Korea
(Tel : +82-31-290-7137; E-mail: tykuc@yurim.skku.ac.kr)

Abstract: This paper presents a dialogue system for intelligent service robot application. The dialogue system consists of an audio speech recognition module, a text-to-speech conversion module, and a dialogue manager for supervision of timely conversion of input command to output robotic service. The dialogue manager is implemented in the framework of task supervisor for HRI(Human-Robot-Interface) module and makes use of finite state machine module for each robotic service. Since the number of services generated grows larger as the complexity of robotic task increases, the set of finite state machine modules is designed in compact modular form in the dialogue manager enhancing the capability of intelligent service robot handling various missions. The effectiveness of proposed dialogue system is demonstrated by applying it to information delivery service of weather and news, scheduling service, and entertainment game service for robotic user.

Keywords: dialogue system, intelligent service robot, dialogue manager

1. INTRODUCTION

As robot intelligence has been developed, various research topics on the intellectual service robot which can render service to man have been explored in recent years. For example, there are many researches on navigation of robot to freely move in the environment of office or hospital, image processing technique to recognize the surrounding environment using camera, manipulation technique which controls the robot arms to hold a desired object, human-robot interaction for robot to communicate with man, and integration and service implementation of robot to ensure new service generation by applying such technologies.

Intellectual service robots must be able to freely communicate with man because they directly render the target service to man. In such a case, the voice recognition technique is essential which helps the robot to recognize the man's voice and invokes the voice synthesis algorithm to deliver the result of certain event to man by language. In addition, the dialogue system which integrates the two individual technologies to ensure the communication must be implemented for target robot services. There are lots of researches on the implementation of dialogue system in the field of telephony. Among the various implementations, a common feature is that most of the dialogue systems include dialogue manager which plays a key role in generating various services as well as communicating with human. In recent years, similar researches on dialogue system for robotic application have been started to communicate with man and deliver robotic services. This important research subject of dialogue system in service robotics area has a few results worldwide from university research groups [1-4].

This paper describes design and implementation of dialogue system for intelligent robotic services which use a general mobile robot platform with SBC(Single Board Computer) and MSRS(Microsoft Robotics Studio) for software integration. Only functions required for provision of the necessary robotic services are implemented along with the designed dialogue system.

The outline of this presentation is as follows. Chapter 2 describes the overall software structure of robot implemented and Chapter 3 explains the dialogue system designed with detailed structure. Chapter 4 demonstrates experiment results through real execution of robotic service scenarios. The final chapter briefly summarizes and concludes the paper.

2. ROBOT'S SOFTWARE STRUCTURE

As the robot's software platform, MSRS has been used in this application in which each software module for robotic service is also implemented based on MSRS.

Robot recognizes the man's voice through the voice interface service module and delivers it to the dialogue manager in the task manager. The dialogue manager controls the robot to enable the robot to properly practice the received command by the present state of FSM(Finite State Machine). Alternately, it can be controlled through a remote access.

2.1 Task Manager (Dialogue Manager)

Dialogue manager is the service module which determines the next state of the robot by FSM, by using the robot's present state information and the voice command value gained from the voice interface module, and corresponds to the main FSM. Task manager controls the robot by operating the service module which is the state determined by dialogue manager.

Also, if the service requires certain process, it can execute the process through the task manager.

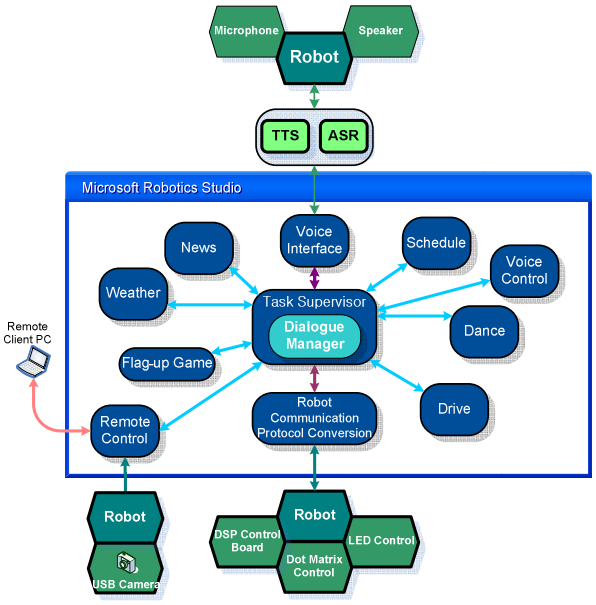

Fig. 1 Software structure based on MSRS

2.2 Robot Communication Protocol Conversion Module

To control the robot's driving part, both arms, ears and both arm ends' LED and eyes' Dot Matrix, the command must be given through protocol. Robot communication protocol conversion module receives the command from another service module and converts it into the corresponding protocol. Also, it receives the robot motor's encoder information and sensor information.

2.3 Voice Interface Module

To configure the dialogue system, ASR(Audio Speech Recognition) and TTS(Text To Speech) products of HCILAB were used. But, these products were supplied in the form of extensible DLL, so that they couldn't be directly implemented as the service module of MSRS. It was because all services of MSRS were implemented based on .Net Framework. Accordingly, there were two methods for using the above products; the method of making and using the COM Object, and the method of sending the message recognized from ASR to MSRS service module and of implementing TCP/IP-based communication program which sends the voice synthetic message from MSRS service module. When we are prepared for the development of the robot system, we were pressed for time and we didn't know about COM object. So, we selected the latter. Voice Interface Module is the MSRS service module which helps the task manager and the communication program to exchange a message through TCP/IP with the communication program.

2.4 Weather, News and Schedule Module

These services provide information of robot's text file by voice through using TTS. Currently, the system which searches the weather or news information online and the schedule input system by remote access haven't been implemented yet. Service is provided by classifying 'today (tomorrow)'s weather', 'yesterday (today)'s news', and 'today (tomorrow)'s schedule according to the voice command. When the voice synthesis is completed, the service is completed.

2.5 Flag-up Game and Voice Control Module

Flag-up game module and voice control module is the module which controls the robot by receiving the voice command based on it with the another FSM, that is sub-FSM. Flag-up game module controls both arms according to voice command such as 'Blue flag up' and 'White flag down', and changes the arm's led and eyes' dot matrix. In the flag-up game, after arms are up and down, arms are not returned to a basic pose again (forward stretching pose).

Voice control module controls both arms and the driving part's two wheels and head according to the voice commands such as 'move forward', 'move backward', 'left hand up', and 'look up'. Also, it finishes the service by voice command.

2.6 Dancing and Navigation Module

Dancing module is the module which sequences the robot's actions in the order by music's rhythm. Because of music sound, it is difficult to recognize the voice command. Accordingly, the robot's front bumper must be pressed for finishing the service. Navigation module is the service of autonomous navigation using the robot's ultrasonic sensor and bumper sensor. The service can be finished through voice command.

2.7 Remote Control Module

Remote control module helps the outside computer to control the robot or to check the robot's state. Because the robot doesn't have the display such as LCD monitor, this module is absolutely necessary. With the camera on the robot, the robot's surrounding images can be obtained through the remote access. It can be executed regardless of voice command.

3. DIALOGUE SYSTEM

3.1 DIALOGUE SYSTEM'S STRUCTURE

Dialogue system applied to the robot is described in Fig. 2. The voice interface part is not significant because it doesn't exist in general dialogue systems [5].

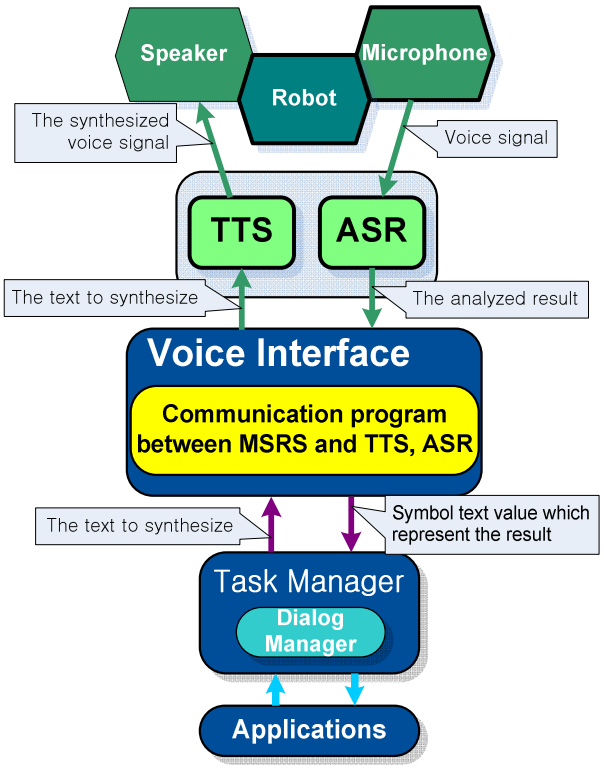

Fig. 2 Dialogue system structure.

The progressive flow of dialogue system by time is described as follows. The man's voice is received by mike, and the received voice signal is analyzed by ASR. Symbol text value which represents the results of analyzing is delivered to the voice interface. (Because there are expressions with different speeches but same meaning (synonym), different speeches with the same meaning are grouped in one symbol value.) Dialogue manager operates the necessary application program (MSRS Service Module) according to symbol value and stores it as the present state. Also, when the application program requires the voice synthesis, it delivers the text of voice synthesis contents to the task manager. Task manager delivers that text to the voice interface and the voice synthesized by speaker is reproduced through TTS. The example by scenario execution is described in chapter 4.

3.2 FSM(Finite State Machine) Design

Robot basically waits for the voice input in the initial state of main FSM shown in Fig. 3. When the voice recognition value comes through the dialogue system, the robot determine one of 'Flag up game', 'voice control', news', 'weather', 'schedule', 'dancing', and 'navigation state' as the next state. When the next state is determined, the service for each state is executed. When the service is finished, the robot's state is restored to the original state again.

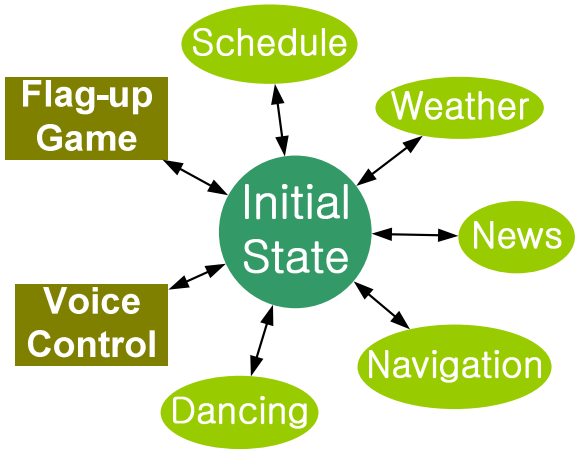

Fig. 3 Main FSM

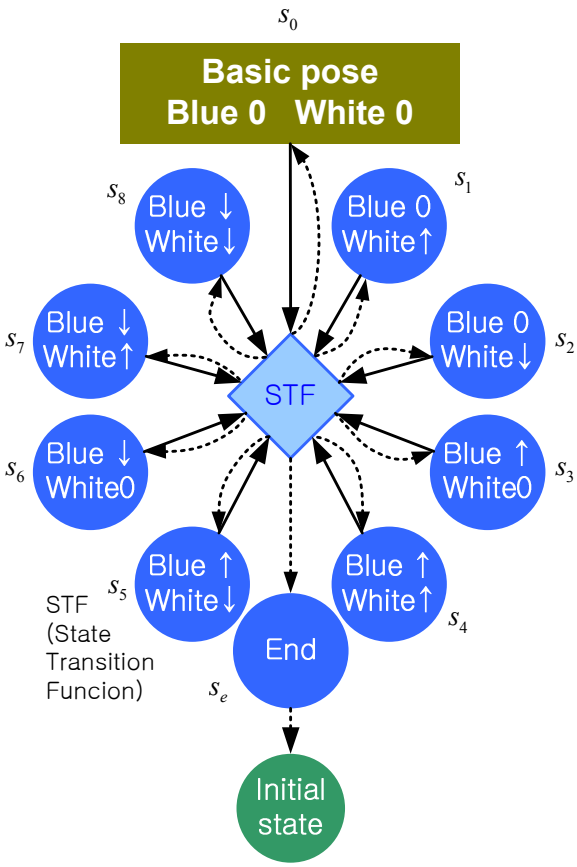

Fig. 4 Sub FSM. The solid line represents the moving direction of the state and symbol. The dotted line represents the transition direction of the state.

Flag-up game service or voice control service has the sub FSM, because many states exist during the execution of each service. Fig. 4 shows FSM of Flag-up game service, as sub FSM. STF(State Transition Function) is the function which decides the next state in the flag-up game through inputting the present state and the recognized voice, as the function which decides the

next state.

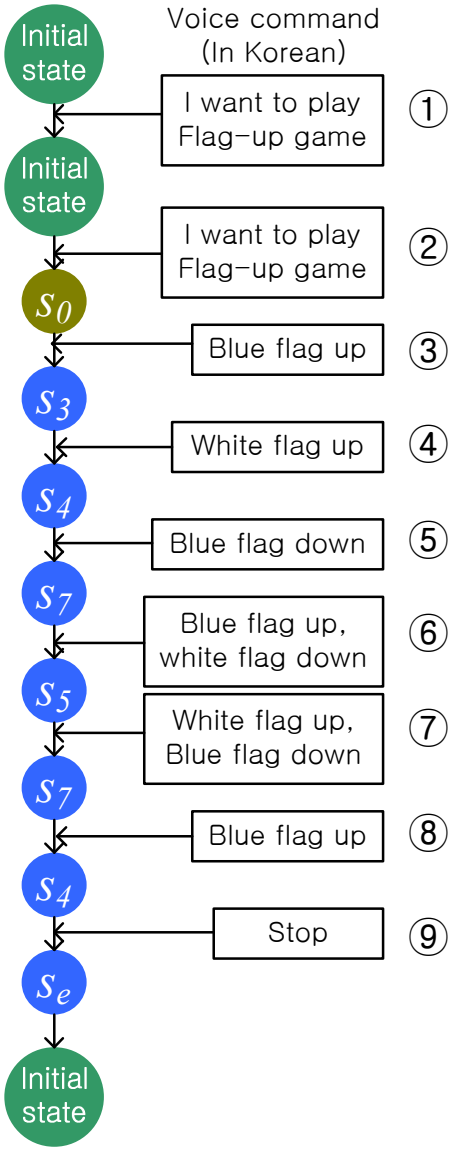

Fig. 5 Flag-up game scenario to validate the dialogue system.

4. EXPERIMENT

By executing the flag-up game service according to the scenario of Fig. 5, whether it was possible to execute the process properly through designed dialogue system was verified. ① The voice command is given to start the flag-up game service in the initial state. When the voice command is not properly recognized by robot, the state is not changed, and the robot outputs the voice message asking the voice command again to the user. ② The user gives the command to start the flag-up game service again, and the flag-up game service starts as the robot properly recognizes the command. The robot's state is s_0 which is the basic pose state of flag-up game. Before the flag-up game is over, the robot's state is changed in sub FSM. ③~⑧ Robot's state is moved into the next state according to the robot's present state and inputted voice command, and the robot executes the state-related action. ⑨ Robot receives the game over command and the state becomes s_e. Robot outputs the game over message and puts down both arms. The flag-up game service is finished. Robot's state is restored to the initial state again.

Fig. 6 The robot appearance of state s_7

5. CONCLUSIONS

The proposed dialogue system can be easily implemented because it is easy to understand and straightforward to design for target robotic services. By using the voice recognition module which can recognize only a short sentence or words, FSM-based dialogue system was configured. The whole FSM was composed of main FSM and many sub FSMs. Each sub FSM takes charge of one service. The state in FSM is changed by user's voice command input and robot's present state and the robot executes the next action which corresponds to changed state (all actions including the voice synthesis). In the dialogue manager, since the number of services generated grows larger as the complexity of robotic task increases, the finite state machine modules are designed in compact modular form enhancing the capability of intelligent service robot handling various missions.

Even if the dialogue system has been designed by considering only user's voice input, the dialogue can be possible with other kind of inputs in place of the voice. For example, it is possible to use gesture recognition by image processing, character recognition or tactile sensor information, ultrasonic sensor information, odor sensor information input, etc. in designing the dialogue system. Accordingly, these are some design factors which configure the dialogue system, and a dialogue system which integrates them needs to be further researched.

REFERENCES

[1] I. Toptsis, A. Haasch, S. Hwel, J. Fritsch and G.

Fink, "Modality Integration and Dialog Management for a Robotic Assistant", in Proc. *European Conf. on Speech Communication and Technology*, Lisboa, Portugal, 2005.

[2] R. Gockley, R. Simmons, J. Wang, D. Busquets, C. DiSalvo, K. Caffrey, S. Rosenthal, J. Mink, S. Thomas, W. Adams, T. Lauducci, M. Bugajska, D. Perzanowski and A. Schultz, "Grace and George: Social Robots at AAAI", Technical Report WS-04-11, *AAAI Mobile Robot Competition Workshop*. pp. 15-20, 2004.

[3] L. Aryananda and J. Weber, "MERTZ: A Quest for a Robust and Scalable Active Vision Humanoid Head Robot", in Proc. *IEEE-RAS International Conference on Humanoid Robotics*, Santa Monica, Los Angeles, CA, USA., 2004. IEEE Press.

[4] H. Jung, Y. Seo, M.S. Ryoo and H.S. Yang, "Affective communication system with multimodality for the humanoid robot, AMI", in Proc. *IEEE-RAS HUMANOIDS*, vol.2, pp.690–-706, 2004.

[5] S. Young, “Speech Understanding and Spoken Dialogue Systems”, In Proceedings of the *IEE Colloquium on Speech and Language Engineering – State of the Art (Ref. No. 1998/499)*. London, UK, 1998.