The Seventeenth International Symposium on Artificial Life and Robotics 2012 (AROB 17th '12),
B-Con Plaza, Beppu, Oita, Japan, January 19-21, 2012

A method of teaching a new action to a communication robot through spoken commands

Eiichi Inohira[1] and Yudai Okuda[2]

[1] Kyushu Institute of Technology, Kitakyushu, Fukuoka 808-0196, Japan
(Tel: 81-93-695-6050, Fax: 81-93-695-6050)
(inohira@life.kyutech.ac.jp)
[2] Hoshiden Kyushu Corporation, Kurate, Fukuoka 807-1312, Japan

Abstract: We study on an approach to teach a new action to a communication in spoken commands. It is supposed that these robots are used for household use. Usually, a spoken command corresponds to a predefined action in this type of system. However, it is difficult to prepare many predefined actions to meet user's expectations in any situations. In this paper, we proposed an approach to teaching a new action which is composed of basic actions through spoken commands. And also, we demonstrated our proposed approach through a communication robot PaPeRo, which has a speech recognition system in Japanese. In our experiment, it took approximately 50 seconds to instruct the five combinations of a basic action and an instruction modifier through 10 spoken commands. After teaching the new action, it took approximately 20 seconds to complete a series of stored actions through a single spoken word.

Keywords: Communication robot, speech recognition, teaching

1 INTRODUCTION

There are several researches on service robots communicating with people through spoken dialogue [1, 2, 3]. It is supposed that these robots are used for household use. In general, a spoken command corresponds to a predefined action in this type of system. It means that the robots can perform only a predefined action through spoken commands. Therefore the robots can serve in limited situations. However, a designer is very difficult to prepare many predefined actions in any situations. A user cannot adjust predefined actions without programming according to its expectations. Our solution to this problem is that a user can teach a new action to the robot according to its need. In this paper, we proposed an approach to teaching a new action which is composed of basic actions through spoken commands [5]. And also, we demonstrated our proposed approach through a communication robot PaPeRo [4], which is manufactured by NEC Corporation and has a speech recognition system and speech synthesis in Japanese.

2 OUR PROPOSED TEACHING SYSTEM THROUGH SPOKEN COMMANDS

Our basic idea is that a new complex action is composed of basic actions. It means that a user can teach a new action to a robot via spoken words which describe combination of basic actions. A simplest case is that a robot memorizes a sequence of basic actions. A more complex case is that a robot memorizes a sequence of basic actions and responses to some situations, where a robot is interrupted during acting. A much more complex case is that a robot communicates with a user and modifies apart of a memorized sequence. In stage of practical use, a robot needs to communicate with people via multimodal interfaces such as touch and vision in order to response to contacts with perople and to memorize people and places.

Our proposed teaching system is based on a speech recognition system and a learning system which can store a spoken word corresponding to a sequence of basic actions. Design of our system depends on implementation of the used robot, that is, PaPeRo. Therefore, our implemented teaching system is described in Section 3.

3 EXPERIMENTS

3.1 Overview

In order to perform preliminary experiments to evaluate our proposed approach, we defined the four basic actions (go forward, go backward, turn right, turn left) and the three instruction modifiers (small, middling, big), which specifies magnitude of travel distance or rotation angle in a single action. We prepared a program to store the five combinations of a basic action and an instruction modifier in the robot's memory and to play the new action through a single spoken command. After the robot accepted the spoken command for 'teaching a new action', it waits for a spoken command for basic actions. Next, the robot waits for a spoken command for instruction modifiers. For instance, when a user says, 'go forward', 'big', it means the robot travel long distance forward. Then, a user can give five basic actions to the robot. Here, limitation on the number of stored actions

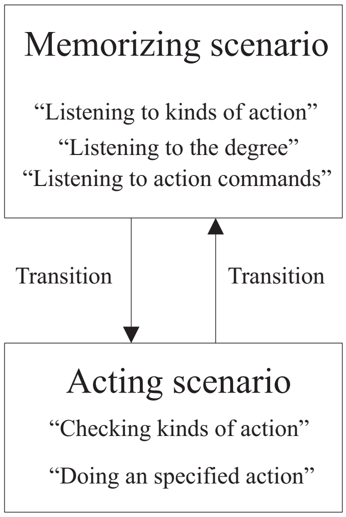


Fig. 1. The teaching scenario

is provisional for preliminary experiment and can be easily changed by modifying the program. After a user taught a new action, the robot can play the new action through a spoken command.

3.2 Experimental setup

PaPeRo has functions of speech recognition and speech synthesis in Japanese. In addition, PaPeRo has two active wheels for moving on a floor and several touch sensors and two cameras, which we do not use in this paper. PaPeRo acts according to a specific scenario, which describes its internal states, external events, robot actions, and state transitions, that is, a finite state machine. We described scenarios to memorize a sequence of basic actions and to invoke a new action. Fig. 1 shows a diagram of the teaching scenario consisting of the memorizing scenario and the acting scenario. In the memorizing scenario, a robot listens to kinds of basic actions and the degree, that is, the instruction modifiers. After memorizing a sequence of basic actions, a robot listens to action commands such as invoking a new action. The acting scenario is prepared in order to travel or turn the robot. The aim of separating scenarios is to remove complex commands to move the robot from the memorizing scenario.

Fig. 2 shows a flow chart of the memorizing scenario. In the memorizing scenario, a robot listens to a pair of kinds of basic actions and the degree until five times. After a robot listens to a pair, it performs an instructed. After a robot memorizes a sequence of the pairs, a robot listens to the action command "all" and performs a new action according to the memorized sequence.

Fig. 3 shows a flow chart of the acting scenario. In the acting scenario, a robot checks if a specified action is travelling or turning. The reason is that the travelling command and the turning command are not the same programmatically.

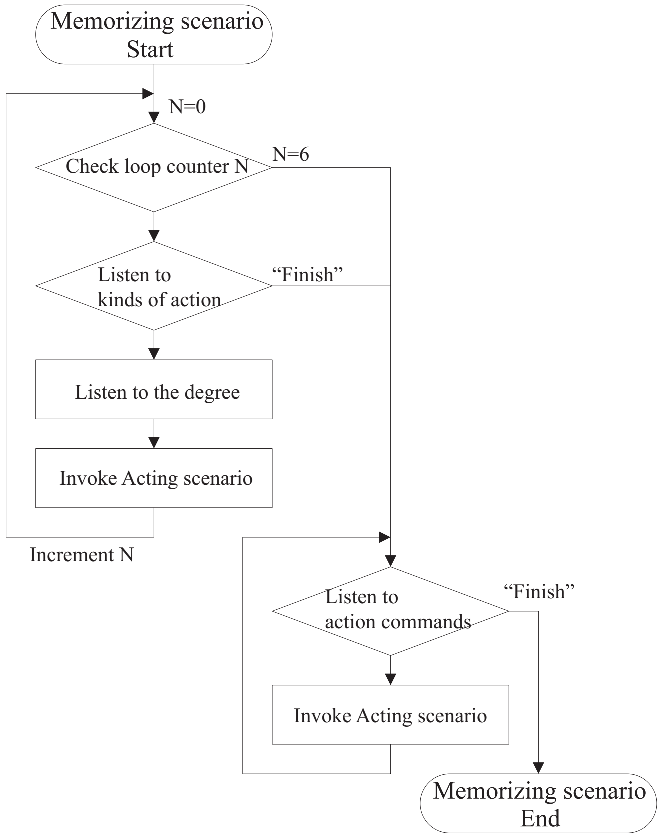


Fig. 2. The memorizing scenario

We defined pairs of a spoken command and a basic actions as shown in Table 1. Table 2 shows definition of instruction modifiers.

3.3 Experimental results

We demonstrated teaching of the new action to travel the robot along a route as shown in Fig. 4. Fig. 5 shows a time chart of our demonstration. We give five pairs of basic actions and instruction modifiers to the robot. The robot memorized the 10 spoken commands as follows:

1. "Go forward" and "ordinarily"

2. "Turn left" and "big"

3. "Go forward" and "big"

4. "Turn right" and "big"

5. "Go backward" and "ordinarily"

Note that a user actually spoke Japanese words in the experiment. After teaching, we give a spoken command "all" to invoke a new action according to memorized sequence of basic commands. A robot performed the new action successfully as shown in Fig. 6.

Table 1. Definition of basic actions

Spoken command	Definition
Go forward	A robot goes forward with constat speed for a specified period
Go backward	A robot goes backward with constat speed for a specified period
Turn right	A robot turns right in the center of its current position by specified rotation angle
Turn left	A robot turns left in the center of its current position by specified rotation angle

Table 2. Definition of instruction modifiers

	Duration [s]		Rotation angle [deg]	
	Go forward	Go backward	Turn right	Turn left
small	1	1	-30	30
ordinarily	3	3	-60	60
big	6	6	-90	90

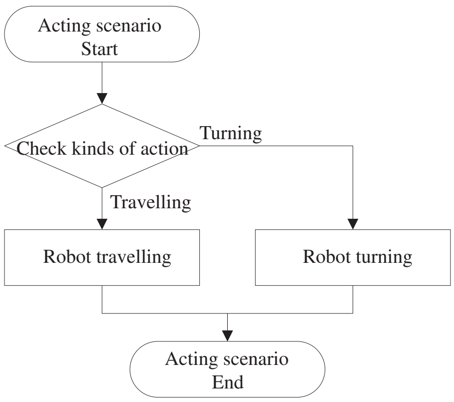


Fig. 3. The acting scenario

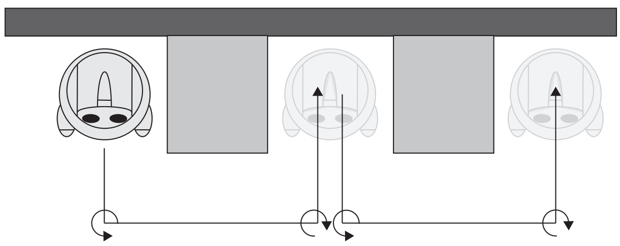


Fig. 4. The travelling route of the robot

In our experiment, it took approximately 50 seconds to instruct the five combinations of a basic action and an instruction modifier through 10 spoken commands. On the other hand, it took approximately 20 seconds to complete a series of stored actions through a single spoken command. We showed that our proposed teaching system realizes to teach a new action to the robot through spoken commands without developing a new program and that teaching a new action via spoken commands is useful for saving time and labor.

4 DISCUSSION

In our demonstration as mentioned above, a robot memorized a sequence of basic actions and instruction modifiers. It means that we realized a simplest teaching system mentioned in Section 2. In the demonstration, it is difficult for us to memorize 10 spoken commands. For instance, when the tenth command is spoken, it is hard for us to recall the first command because we cannot memorize too many commands at a time. Therefore, we need to confirm given basic actions until a certain point before completion of teaching. And also, when a user gives mistaken command to a robot, modification of a taught action is needed. In order to deal with these cases, we need to improve the teaching system as a user smoothly teaches a new action to a robot.

5 CONCLUSION

In this paper, we proposed an approach to teaching a new action which is composed of basic actions through spoken commands. A spoken command corresponds to a predefined action in conventional communication robots. We demonstrated our proposed teaching system through a communication robot PaPeRo, which has a speech recognition system in Japanese. In order to perform preliminary experiments, we defined the four basic actions (go forward, go backward, turn right, turn left) and the three instruction modifiers (small, middling, big). Our teaching system can memorize a squence of basic actions and instruction modifiers via spoken commands and invoke a new action after teaching.

In the future work, we will improve the teaching system in order to deal with modification of a taught action via spoken commands.

REFERENCES

[1] Tikhanoff V, Cangelosi A, Metta G (2011), Integration of speech and action in humanod robots: iCub simu-

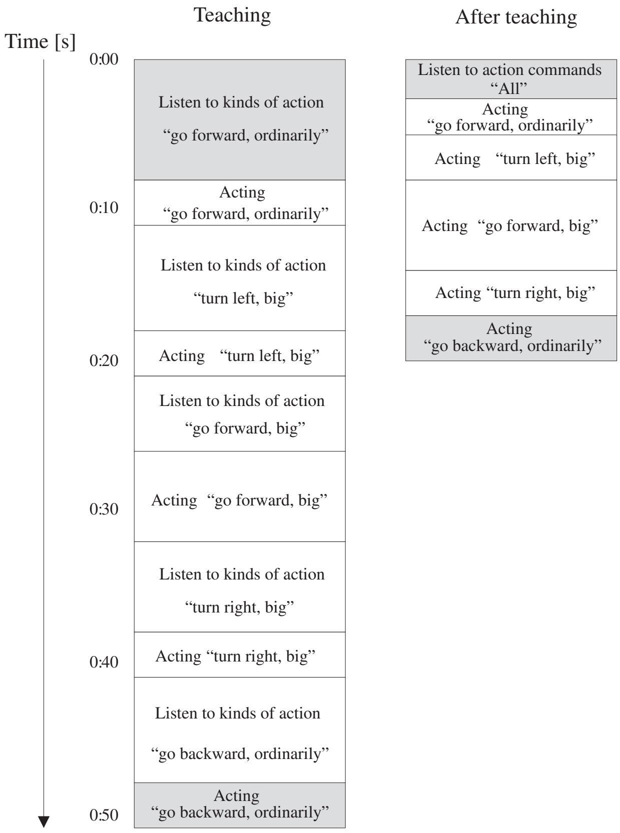


Fig. 5. A time chart of teaching the robot

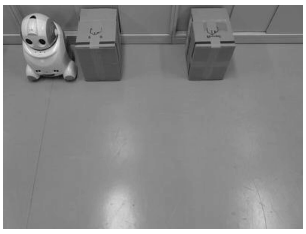
lation experiments. IEEE trans on Autonomous Mental Development 3(1):17-29

[2] El-Emary IMM, Fezari M, Attoui H (2011), Hidden Markov model/Gaussian mixture models (HMM/GMM) based voice command system: A way to improve the control of remotely operated robot arm TR45. Scientific Research and Essays 6(2):341-350

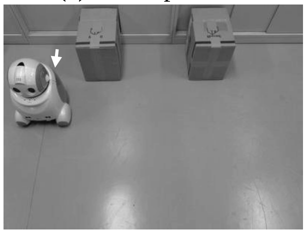
[3] Budiharto W, Jazidie A, Purwanto D (2010), Indoor Navigation using Adaptive Neuro Fuzzy Controller for Servant Robot. 2010 Second Intl Conf on Computer Engineering and Applications (ICCEA 2010) 582-586

[4] PaPeRo, http://www.nec.co.jp/products/robot/en/index.html

[5] Okuda Y, Inohira E (2010), Acquisition of action by using speech recognition in a communication robot. Proc of the 12th SOFT Kyushu Chapter Annual Conf 67-70

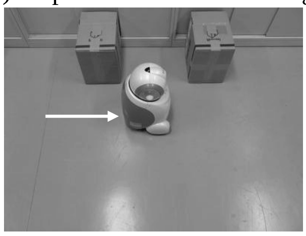


(a) Initial position

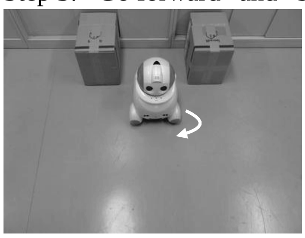


(b) Step 1: "Go forward" and "ordinarily"

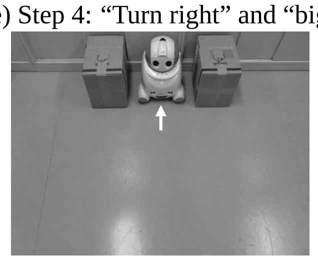
(c) Step 2: "Turn left" and "big"



(d) Step 3: "Go forward" and "big"



(e) Step 4: "Turn right" and "big"



(f) Step 5: "Go forward" and "ordinarily"

Fig. 6. Snapshots of the experiment of teaching the robot