# SPOKEN DIALOGUE SYSTEM FOR HOME HEALTH CARE

*S. Takahashi, T. Morimoto, S. Maeda, and N. Tsuruta*

Fukuoka University
Department of Electronics Engineering and Computer Science
8-19-1 Nanakuma, Jonan-ku, Fukuoka 814-180, Japan

## ABSTRACT

This paper reports a Japanese spoken dialogue system aiming at support for home health care services. The targets of this system are aged people who live alone with comparatively good health condition. In order to watch their daily health condition, the system executes medical checks through some examinations and conversations. By generating appropriate discourse plan in collaboration with a medical expert program, the system performs a efficient discourse reducing redundant or needless interactions. This paper describes the architecture and the main features of our system, and presents the results for a dialogue experiment .

## 1. INTRODUCTION

The aging of the population in Japan is proceeding more rapidly than in other countries. According to the Ministry of Health and Welfare, senior citizens, who are more than 65 years old, are expected to reach 20% of Japan's population by the year 2010. Especially the number of the elderly people in need of health care is growing. Under the circumstances described above, an advent of home health care supporting system is desired strongly.

This paper proposes a spoken dialogue system aiming at support for the home health care services. The targets of this system are aged people who live alone with comparatively good health condition. The system watches user's daily physical conditions through medical examinations (e.g. checking the user's temperature and blood pressure) and some conversations (e.g. asking some questions about the user's condition).

In order to accomplish this task, the system has to satisfy the following requirements:

1. to perform appropriate medical discourses for the user's conditions.
2. to generate the efficient discourse for medical check by reducing redundant or needless interactions.
3. to deal with the user's interruptions causing topic shifting.

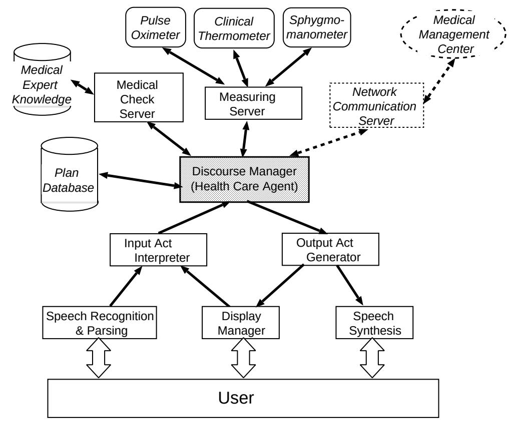


**Fig. 1**. System organization.

Taking into consideration of these requirements, we propose a plan-based dialogue system which generates appropriate discourse plans in collaboration with a medical expert program. The system uses discourse plans to describe tree structures in a health check domain. The system determines whether the current topic should be changed or not, according to the user's utterance, and if necessary, the system changes the topic or reconstructs the discourse plan. In addition, the efficient dialogue avoids the redundant or needless questions by generating a medical check plan according to a medical knowledge, and the mis-understanding of the user's utterances is reduced by limiting the input accepted according to the proceeding plan and by navigating the input with displaying menu items.

Several spoken dialogue systems targeting elderly people have been developed, which are the part of personal service robots, but they are still early prototypes and don't deal with conversations for health care service yet [1],[2]. On the other hand, over the past few decades, many expert systems have been proposed for diagnosing diseases, but they have mainly attempted to follow a reasoning or diagnostic plan, and have not been investigated in the context of a dialogue system or as a human-computer interaction problem[3].

In this paper, we first describe the architecture and the functions of our system. Secondly, we explain main features of our system, topic-shifting caused by user's interruption and dynamic generation of medical check plan. Finally, we present our prototype system and the experimental results. From the experimental results, we show that our system can execute efficient discourse and derive the likely disease names from the user's answers.

## 2. SYSTEM ORGANIZATION

The basic organization of our system is shown in Fig.1. The system communicates between each component which works parallel and asynchronously. Our system consists of the discourse manager and the individual special components. In the following, we explain the role of main components.

### 2.1. Discourse Manager

The Discourse Manager, called the Health Care Agent, is the core of our system. This component is responsible for managing the ongoing conversation, accomplishing the appropriate discourse plan according to the current topic, sending requests to or receiving responses from the other components, and generating appropriate responses to the user.

### 2.2. Medical Check Server

The Medical Check Server provides a medical check plan. Upon the requests from the Discourse Manager, this component prepares questionnaires and examinations to be performed, and sends them back to the Discourse Manager. The examination and questionnaire plans to be executed, that is, which data should be measured or what questions should be asked, are determined by referring to the medical knowledge which is prepared based on the advice of human medical doctors in advance.

### 2.3. Measuring Server

The role of this component is to measure values through some medical instruments installed with our system, for example, a clinical thermometer, a sphygmomanometer, a pulseoximeter for a $SpO_2$, which indicates the oxygen saturation, and so on.

## 3. EFFICIENT DISCOURSE MANAGEMENT

### 3.1. Dynamic Discourse Planning

Fig.2 shows a basic plan for a daily medical check. The system executes a dialogue using a discourse plan, in the forms of a tree structure, because the tree structure plan is able to deal with topic-shift in our limited domain. For example,

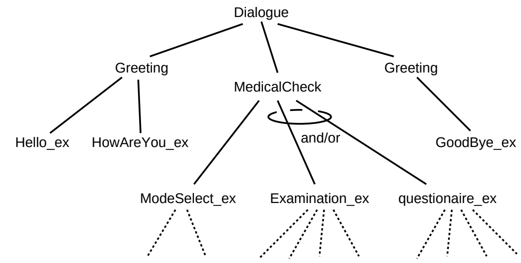


**Fig. 2**. A plan tree for daily medical checks.

when the user's physical condition is good, the system executes discourses according to the given basic plan [1]. When the user complains about his/her conditions to the system, an interruption from the user is sent to the Discourse Manager and the discourse plan asking about the user's symptom is constructed.

In the parsing process of our system, a user's utterance are segmented by morphological analyzer, key phrases are derived from it. The key phrases are, for example, noun phrases, verb phrases, and negative expressions, and so on. Matching this key phrase with the executable plan in the current state, the system instantiates the plan tree and executes the matched plan. If no plan is matched with the key phrases, the user's utterance is rejected.

The system also displays the menus to allow the user to input his/her answer. The menu items are made from expected answers in the current plan. The user can answer the system by means of not only speech but also by selecting the menu items. Selected answers from the menu items are also parsed the same way as the speech inputs.

### 3.2. Dynamic Generation of Medical Check Plan

When the system receives a message of the user's complaint about bad physical condition, the Discourse Manager sends a request to the Medical Check Server in order to execute the appropriate plan corresponding to the user's condition. The Medical Check Server generates the plans by referring to a medical knowledge called a "pre-diagnosis rule" [2]. In addition, the Medical Check Server identifies suspicious diseases with the examination's results and the questionnaire's answers, and then determines whether the user's condition needs urgent treatments or not.

The pre-diagnosis rule mainly consists of the comparatively high urgent disease which aged people often suffer with. Table 1 and 2 show the examples of the examination / questionnaire items included in our system and the name of (suspicious) diseases diagnosed by our system, respec-

[1] While the basic plans are given in this case, the details of the plans are constructed dynamically as same as the other case.

[2] This term "pre-" means that a fi nal diagnosis is determined by human medical doctors, not our system.

**Table 1**. The examples of pre-diagnosis items.

| category | examination/questionnaire items |
|---|---|
| examination items | temperature<br>pulse<br>blood pressure<br>$SpO_2$ |
| questionnaire items | head ache<br>stomach ache<br>nausea<br>the onset of a cold<br>$\cdots$ |

**Table 2**. The examples of the name of diagnosed diseases.

| assortment | disease name |
|---|---|
| serious / urgent / warning | cardiac infarction, cardiac failure, acute bronchitis, $\cdots$ |
| otherwise | anemia, cold and fever, intercostal neuralgia, hypotension, $\cdots$ |

tively. Table 3 shows the pre-diagnosis rule expressed as a data table. A mark "-" means a *don't-care* state in which the answer can be "yes" or "no". As shown in Table 3, each disease is defined with the several symptoms represented in the top row. If the user's answer doesn't match the condition of a disease, the disease is removed from the candidates of the plausible diseases. Since the questionnaire items relative to the removed disease are also removed, the system executes the discourse efficiently avoiding the redundant or needless questions.

## 4. EVALUATION OF PROTOTYPE SYSTEM

### 4.1. Prototype System

Fig. 3 shows the prototype system we have developed. This prototype system consists of the components drawn with the solid lines in Fig.1 (the parts with the dashed lines will be implemented soon).

For speech recognition, the system uses the HTK's Viterbi recognizer[3], and a Japanese acoustic model included Julius[4]. The language model is prepared as an FSA network grammar for our domain. The number of the sentence patterns accepted by this FSA is about 50, and the number of the words in the vocabulary is about 200. The system used the health check plan such as shown in Fig.2. The number of the executable plans, which correspond to the leaves in the

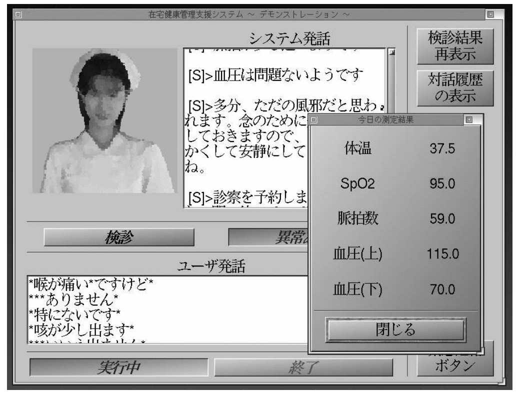


**Fig. 3**. Prototype system.

plan tree, is about 50. In the prototype system, the table of the pre-diagnosis rule consists of 36 disease names and 32 items of examination / questionnaire.

### 4.2. A Dialogue Experiment

In order to investigate the dialogue performance of our prototype system, we conducted a dialogue experiment. This experiment used six subjects from graduate students in our laboratory. None of the subjects had ever used the system before. After the brief explanation of the usage of our system, the experiment proceeded under the assumption that each subject has caught a cold or suffered from a disease like a cold. Even though the menu was displayed, subjects were requested to use speech input only.

The results of the experiment are shown in Table 4. The first six rows in Table 4 present the result for six subjects, respectively. The bottom row presents the average of the result for the subjects. The left four columns show the correct recognition rate of the sentences the user uttered, the successful and the unsuccessful rate of plan matching, and the number of rejected utterances, respectively. The fifth column presents the number of the user's utterances and the total of the user's and the system's utterances in each dialogue session. The rightmost column shows the results of the pre-diagnosis ( a mark "$*$" means that the disease is urgent).

From this result, although the recognition rates of the sentences were low, the high plan matching rates were obtained. The reason for this is that the key phrases to be matched with the plan were recognized correctly, although the all words in the sentence were not recognized.

The results also show that the number of interactions is reduced as compared to the case that the system would ask all the questions in the pre-diagnosis rule [5]. From this result, we confirmed the system executes discourse plans

[3] HTK (Hidden Markov Model Tool Kit) web page is "http://htk.eng.cam.ac.uk".

[4] Julius is Japanese Dictation Tool Kit. Julius's web page is "http://winnie.kuis.kyoto-u.ac.jp/pub/julius/".

[5] The number of the interactions in the utmost case is about 70.

**Table 3**. The example of diagnosis rules.

| temperature | pulse | · · · | headache | sputum | · · · | nausea | disease name |
|---|---|---|---|---|---|---|---|
| $\geq 37$ | - | | - | - | | no | common cold |
| $\geq 37$ | - | | no | yes | | - | suspicious of pneumonia |
| - | - | | yes | - | | yes | cold or suspicious of meningitis |

**Table 4**. The result of the dialogue experiment.

| subject | correct recognition | successful matching plan | unsuccessful matching plan | number of rejections | user's / total utterances | result of pre-diagnosis |
|---|---|---|---|---|---|---|
| A | 9 | 12 | 0 | 0 | 12 / 37 | cold and fever |
| B | 2 | 11 | 2 | 3 | 16 / 45 | cold or suspicious of meningitis $^*$ |
| C | 6 | 13 | 0 | 0 | 13 / 44 | head ache with cold |
| D | 7 | 13 | 0 | 0 | 13 / 39 | head ache with cold |
| E | 7 | 9 | 1 | 1 | 11 / 40 | common cold |
| F | 10 | 17 | 2 | 0 | 11 / 52 | suspicious of pneumonia $^*$ |
| average | 52.4% | 89.3% | 5.9% | 4.8% | 13 / 42 | |

efficiently, avoiding needless or redundant interactions.

For the results of the pre-diagnosis, the likely diseases were obtained. The explanation for the different results is that the system determined more safely the suspicious disease for each user's different answers. For example, meningitis and the common cold have the similar symptom, but the system determined the disease that needs more urgent treatments, when the user complained about serious symptoms. We intend to test the system in the field to refine the pre-diagnosis rules.

### 5. CONCLUSION AND FUTURE WORK

In this paper, we described a Japanese spoken dialogue system aimed at support for home health care services. In order to achieve the task for watching user's daily health, we proposed the plan-based dialogue system which generates appropriate discourse plans in collaboration with the Medical Check Server. From the experimental results, we confirmed that the system can execute the efficient discourse and derive the likely disease names from the user's answers.

In the future, we intend to address the following three issues. First, we intend to make our system handle more natural interactions. To do this, the system needs to perform incremental understanding and generation with more flexible turn-taking or mixed-initiative interactions [4],[5]. Second, we intend to cope with the issues of speech recognition for the elderly[6]. Finally, we intend to integrate image processing, such as the recognition of the user's image and the generation of graphical interfaces with a visualized agent, in order to achieve the more natural and user friendly conversation.

## Acknowledgment

We would like to thank Dr. Yamada, Dr. Ogawa, and Dr. Kaetsu at School of Medicine, Fukuoka University, for providing the medical knowledge.

## References

[1] N. Roy *et al.*, "Spoken dialogue management using probabilistic reasoning," in *Proc. of the 38th ACL*, 2000.

[2] R.D. Schraft *et al.*, "The concept of a system for assisting elderly or disabled persons in home environments," in *Proc of the 24th IEEE IECON*, 1998, vol. 4, pp. 183–190.

[3] A. Gertner *et al.*, "Recognizing and evaluating plans with diagnostic actions," in *Working notes of the IJCAI-95 Workshop on Plan Recognition*, 1995, pp. 48–53.

[4] M. Nakano *et al.*, "Understanding unsegment user utterances in real-time spoken dialogue systems," in *Proc. of 37th ACL*, 1999, pp. 200–207.

[5] D. Byron, "Improving discourse management in TRIPS-98," in *Proc of the 6th Eurospeech*, 1998, vol. 3, pp. 1379–1382.

[6] A. Baba *et al.*, "Elderly acoustic models for large vocabulary continuous speech recognition (in Japanese)," *Trans. of the IEICE D-II*, vol. J85-D-II, no. 3, pp. 390–397, 2002.