A Visual Anthropomorphic Agent with Learning Capability of Cooperative Answering Strategy through Speech Dialog

Yasufumi Takama Hiroshi Dohi
Mitsuru Ishizuka
Dept. of Information & Communication Engineering, School of Engineering, University of Tokyo
7-3-1, Hongo, Bunkyo-ku, Tokyo, 113, Japan
+81-3-3812-2111 ext. 6755
{takama, dohi, ishizuka}@miv.t.u-tokyo.ac.jp

Figure 1. The Visual Software Agent(VSA).

1. Introduction

As the opportunity of using computer systems spreads into everyday life, the importance of friendly human interfaces is increasing. As a form of next-generation human interfaces, an anthropomorphic interface agent which mimics a face-to-face communication holds a promise [8, 9], and some early developments have been started[1, 2, 3, 11, 6]. Its multimodality including facial expression and speech dialog fits to human perception and can enhance the friendliness of the interface.

Figure 1 shows our anthropomorphic interface agent called VSA (Visual Software Agent), which has a moving realistic facial image and a speech dialog function[1, 2]. (Unlike other anthropomorphic interface systems, our VSA system recently has a connection to a WWW browser (Netscape Navigator), so that it can serve as a new interface to a vast WWW information space and effectively use multimedia data written in a standardized HTML format[2, 3].) As immediate applications of the VSA, it is suitable for guidance systems which are used by various people with little knowledge of computers at, for example, department stores, company's reception desks, university campuses, etc.[7, 10]. These guidance systems may be installed in various places, and cooperative dialog strategies for guidance are different to a large extent depending on places while a basic knowledge base may be common. This makes it difficult to design and implement a guidance system fitted to a certain place. If a guide is a human, he/she can acquire heuristic knowledge from experiences to generate cooperative responses fitted to a place and situations. Therefore, we tries to implement this types of learning capability into our anthropomorphic agent. We use reinforcement learning, in particular, profit sharing method[5]; but our research is unique that the learning mechanism is implemented to acquire knowledge from speech dialogs. We show here our implemented learning mechanism with reference to a task of campus guidance.

This paper is organized as follows. We first introduce our VSA system briefly. Then we consider what is a cooperative response in the dialog of guidance tasks. In section 4, we describe a dialog management system(DMS) with a learning module. Experimental results shown in section 5 prove the capability and characteristics of our learning mechanism.

2. Outline of VSA

Our VSA shown in Figure 1 is a multimodal anthropomorphic agent system which has a moving realistic facial image and a speech dialog function. A user can interact with the VSA through a spoken language as

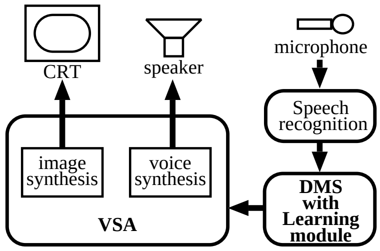

Figure 2. System configuration.

well as the moving facial image.

The configuration of our system is depicted in Figure 2. The VSA part in this figure has already developed, which produces facial and speech outputs. The VSA's face is formed with approximately 500 surface patches (triangle polygons), onto which a facial texture image of a real human is transcribed. The real-time deformation of the triangle polygons allows to make the facial expression change. The VSA occasionally winks her eyes, which gives a vividness of the agent. When speaking, the VSA moves its mouth synchronized with a speech synthesizer.

For speech communication, we are using commercial products for speech recognition and speech synthesis; a speech synthesizer is "Shaberin-bo" (NTT data), and a speech recognizer is "Phonetic Engine 200" (Speech Systems inc.) which is a continuous speech recognizer for unspecified speakers. A word dictionary and a set of grammatical rules are implemented in the speech recognizer for a conceivable situation. The speech recognition performance of our system is not necessarily superior to other systems; a question is sometimes issued when the recognition result is not certain. Although this is bothering to a user, the co-existence of the friendly face has the effect of mitigating this uneasiness of the user.

3. Cooperative answering in guidance tasks

Although there have been several researches on cooperative answering which focused on the contents of an agent's answer, many of them were for goal-oriented style dialogs [4, 12] rather than for question-answering style dialogs. The reason of this may be that in the question-answering style dialog, it has been considered that user's questions can be handled one by one to generate appropriate answers. However, people using guidance systems for example at a department store etc. differ from people using database systems in the following ways:

1. Many of users are not familiar with a computer nor a database system.
2. They don't know the data structure of a database system nor how to query it.
3. It is often the case that their desired information is too vague to submit a query to a database system.

For example, suppose a person who comes to a grocery floor on a department store to buy something for dinner. In this case, he/she may not be sure what to buy in advance, and also may want some recommendation. It will not be cooperative for such a person that a guidance system forces him/her to specify search conditions in detail, or displays a vast number of data returned from a database limitlessly.

As another example, a facility guidance system may be frequently asked by users about the location of his/her destination or the way to there. Although the location where the system is installed can be always used as a reference point regardless of user's knowledge, a good landmark must satisfies the following conditions: (a)it is located near the user's destination, and (b)the user knows it. The latter condition is dependent on users' knowledge about the facility, so it is not easy for the system installed in various places to select an appropriate landmark for the explanation.

Of course, which type of answers is more cooperative than others is dependent on an individual user. Nevertheless, it may be true that the characteristics of users (their knowledge, purposes, etc.) are dependent to a certain extent on where the system is located. Thus guidance experience at a certain place would help to improve the cooperativeness of the dialog, for example, to select an appropriate landmark for explanation. Accordingly, we attempt to generate cooperative answers based on the following two aspects:

user dependency: generating a cooperative answer dependent on the individual user's knowledge about the domain and the context of the dialog.
situation dependency: generating a cooperative answer dependent on where the system is located and the common characteristics of the users at the place.

It is very time consuming for us to take account of this situation dependency and to tune up a knowledge base even for one place. Our approach here is that we implement a certain set of knowledge for cooperative answering in a knowledge base and let the system learn a more cooperative answering strategy depending on a working place through speech dialog interactions with users. We use reinforcement learning [5] as a learning mechanism of this situation dependency.

In addition to the situation dependency, we have to consider the user dependency. For this purpose, we also incorporate a user-model approach.

4. DMS with learning module

In this section, we discuss some details of the DMS with a learning module, which is used in combination with VSA described in section 2.

Figure 2 shows the whole configuration of our speech-dialog system using VSA. The DMS receives the output of the speech recognizer, generates an appropriate answering sentence as a reply to the user's speech, and sends it to the VSA part. The answering sentence is a sequence of words with information about accent and pause. Based on this information, the VSA part synthesizes a voice output and also move the mouth of the VSA face synchronized with the voice output. This facial action increases the reality of the agent, and users can feel a stronger affinity than systems only speaking.

Figure 3 shows the details of the DMS configuration. In each module, main data (knowledge) are represented in italic. A learning module is invoked after the completion of a series of dialog with each user and modifies the strength of used rules and correlation functions. As a domain database for our campus guidance application, we use a simple relational-like database about the campus of the University of Tokyo.

4.1. Recognizer

The recognizer first extracts keywords necessary to identify the intention of an utterance from a keyword sequence sent from the speech recognizer, and then converts the resulting keywords into a semantic representation by matching these keywords with "key–sem. conversion rule".

The key–sem. conversion rules have the following items in their condition parts:

Variables: for indicating keywords that exist in the domain database.

REL_WORD: for indicating keywords not existed in the domain database.

Question Type: for indicating the type of question.

Topic: for indicating topics related to the direct intention of the user's utterances.

4.2. Interpreter

The interpreter selects answering rules (called "candidates") based on both the semantic representation of

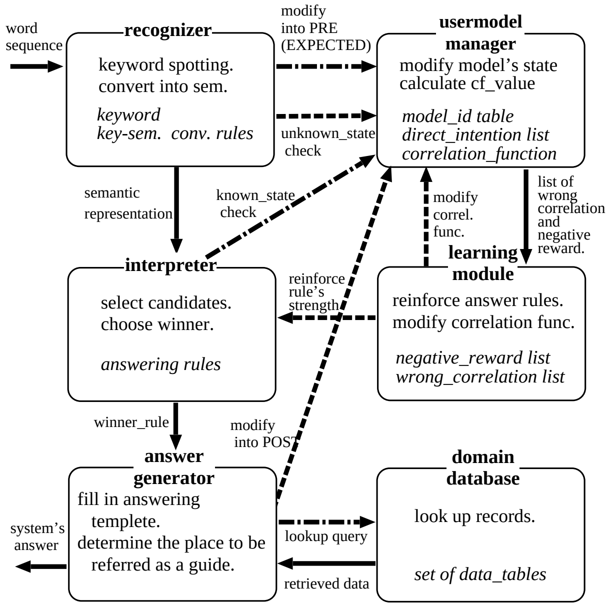

Figure 3. Configuration of DMS.

the utterance and the state of the user-model. In the case that competing candidates exist, one rule (called "winner") is selected based on both the strength of each rule and the number of times the rule was used in the dialog with the current user. That is, the possibility of selecting a certain rule is decreased as the rule is used in the current dialog so that the dialog may avoid falling into a loop.

As illustrated in Figure 4, each answering rule has a Question item and a User-model item as its condition parts, and an Answer item and a Modify item as its action parts. The Question item indicates a semantic representation to match, and the User-model item indicates the states of the user-model to match. The Answer item indicates a templete for generating an answer sentence, and the Modify item indicates instructions to modify the user-model. Also, each answering rule has a Strength part which value is used to resolve the rule competition. This value is modified by a learning module, which will be described later.

4.3. User-model manager

In this system, we adopt an overlay model as a user-model. That is, the state of the current user's knowledge about a task domain is stored in a set of tables which denote a small subset of the domain database. Each item in the user-model has one of the following status; "PRE" indicates that the user has already known the corresponding knowledge, "POST"

Question:

> **Question Type:** CONTENT_TYPE
> **Topic:** QRESTAURANT
> **Variables:**

User-model:

Answer: There are several restaurants;
AREST_ALL

Modify: TBL_FACILITY:REST:IT_NAME;

Strength: 100

Figure 4. An example of answering rule.

indicates that the system has told the user the corresponding knowledge, "KNOWN" indicates that the user seems to have already known, and "UNKNOWN" indicates otherwise.

A user-model manager modifies the state of a user-model according to the instruction from the recognizer or the answer generator. Its concrete process is as follows:

- If a keyword is appeared in the current user's utterance for the first time and the corresponding item has the status: UNKNOWN, it is changed to PRE.
- After above modifications, if the status of highly related items are UNKNOWN, they are changed to EXPECTED. The relation between items is calculated using correlation functions to be described later.
- If the item which corresponds to the direct intention of the utterance has the status POST or EXPECTED, it is changed to UNKNOWN. At this time, the correlation functions which judge the item status to be EXPECTED or the answering rule which directs to change the item status into POST are considered to be wrong, and punished at the learning phase.
- The status of the item which the system has told the user is changed to POST. This is done based on descriptions of a Modify item of the winner.

The correlation function $rel(a, b)$ returns a confident value $[0, 1]$ which represents the possibility for "when the status of item 'a' is PRE, the status of item 'b' is PRE, too". If there are more than two items $x_i (1 \leq i \leq n)$ with the status PRE, the confident value of item y, cf(y), is calculated as follows:

$$cf(y) = 1 - \prod_{i=1}^{n} (1 - rel(x_i, y)) \tag{1}$$

4.4. Learning module

A learning module is invoked when a series of dialog with one user completes. Then it updates the strength of the used answering rules based on reinforcement learning and also modifies the correlation functions.

As a reinforcement learning method, we use profit sharing method [5]. The strength of a rule after completing the tth episode, S(t), is calculated according to the following equation:

$$S(t) = (1 - \alpha)S(t - 1) + \alpha R \tag{2}$$

where R is a reward and α is a learning rate. Notice that we consider a series of dialog with one user as a single episode.

We consider that a dialog with less redundancy is more cooperative, because in such a dialog the system would be able to provide the user with his/her desired information without useless interactions. Based on this consideration, the reward R is calculated as follows:

$$R = R_{base} - R_{penalty} \left\{ 1 - \left(\frac{e - e_M}{e_m - e_M} \right)^5 \right\} \tag{3}$$

$$e = \frac{\#of \ \ utterance\text{–}pairs}{\#of \ \ tasks} \tag{4}$$

where e of Eq. (4) implys the average number of utterance-pairs per task, e_m is the minimum value of e that the system has ever experienced, and e_M is the max value. The number of tasks is calculated using heuristics based on the topics of utterances and extracted keywords. Eq. (3) is designed so as to give much larger rewards R to rules which were used in episodes with a smaller value of e. Furthermore, the R of the rules that were judged to be wrong is set as the penalty to their initial strength (in our experiment, 100).

The learning module also modifies the correlation functions. If a certain item in the user-model has a status EXPECTED at the end of one dialog, correlation functions used in this presumption are considered to be correct. When the current value of a correlation function is q/p, it is changed to $(q + 1)(p + 1)$ in the correct case and to $q/(p + 1)$ in the wrong case (see section 4.3), respectively.

5. Experiment with simulated dialogs

5.1. Overview of experiment

We have implemented an experimental campus guidance system for the University of Tokyo using our VSA

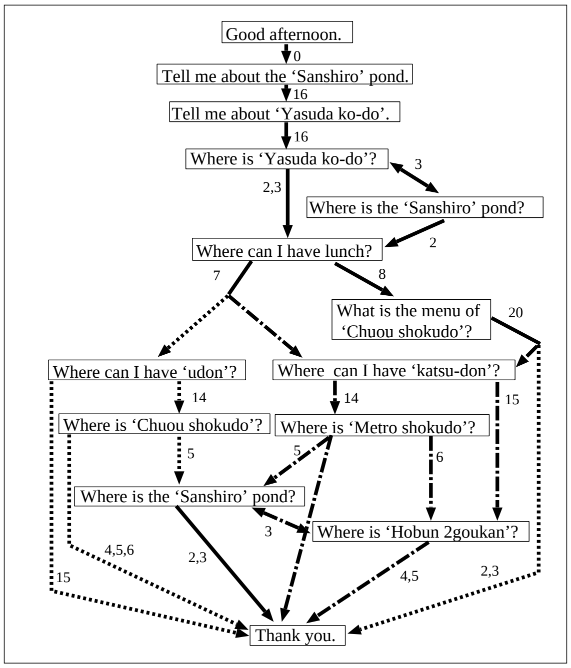

Figure 5. Transition graph of simulated dialogs.

with the learning DMS. This section reports its experimental results carried out using simulated speech dialogs.

In this experiment, we assume that there are four types of users. That is, as to knowledge about the task domain, users who already know the location of the 'Sanshiro' pond (a famous spot on the campus) are distinguished from those who don't know it. As to the user's purpose, users who want to eat 'udon' (a Japanese-style noodle) at a restaurant on the campus are distinguished from those who want to eat t to eat 'katsu-don' (rice with pork cutlet). In combination with these two types of distinction, the users are classified into four groups. In our experiment, there were four users for each group. After total 16 speech dialog interactions with the system, we evaluated the strength of each answering rule. The order that the users in the four groups interacted with the system was random. Of cource, the group category to which the current user belonged was not told to the system.

Also, it was assumed that each user tended to ask to the system in the following order:

1. information about the 'Sanshiro' pond
2. information and location of Yasuda ko-do (auditorium)
3. information about restaurants

As answering rules, we prepared a few competitive rules for telling restaurant information and for showing the way to a certain restaurant. Figure 5 shows the state transition graph that represents all the possible progress of one dialog under above assumptions. Note that these assumptions have been imposed to examine the learning characteristics of our system, and are not required in an actual guidance system. In Figure 5, an arrow indicates an answering rule. The numbers attached to the arrows correspond to those of the following answering rules:

0: Can I help you?
1: You're welcome.
2: {show the way from system's location}
3: {show the way using landmark near to the destination and known to user}
4: {show the way from system's location and internal location}
5: {show the way using landmark near to the destination and known to user, and internal location}
6: {show the way by building's name and its internal location}
7: {list all the names of restaurants}
8: {suggest 'Chuou shokudo' (the central restaurant) which is the nearest to Yasuda ko-do}
14: {tell the restaurant by its name}
15: {tell the restaurant by its name and location}
16: {guidance information}
20: {tell the menu of the restaurant}

In this case, the following sets of answering rules are competitive; rule 2 and rule 3 (regarding how to tell the way to a location); rule 4, rule 5 and rule 6 (regarding how to tell the way to a facility); rule 7 and rule 8 (regarding how to tell restaurant information); rule 14 and rule15 (regarding how to introduce a certain restaurant). Under above settings of the experiment, preferable rules to be selected in the competitive sets are rule 2, rule 4, rule 7 and rule 15, respectively.

5.2. Experimental results

Table 1 shows the setting of parameters in the learning module. Experiments were repeated 3 times. Table 2 compares the strengths of the competitive rules at the final state, i.e., after 16 speech dialog interactions with the system. Here, all rules were given the same strength of 100 at the initial state. Values within parentheses are relative strengths among the competitive rule sets.

Table 2 shows that a stable result can be obtained even after 16 interactions in the random order of the users; that is, in each trial, the strongest rule among

Table 1. Parameter setting in learning module

learning rate:α	0.2
positive base reward:R_{base}	1000
negative base reward:$R_{penalty}$	800
initial max of evaluation function:e_M	3.0
initial min of evaluation function:e_m	2.5
max difference between e_m&e_M: L_{Max}	1.0

Table 2. Experimental results

trial	1	2	3
rule2	560(**2.92**)	792(**2.01**)	517(**3.23**)
rule3	192(1.00)	394(1.00)	160(1.00)
rule4	183(**1.83**)	539(**1.93**)	141(1.04)
rule5	136(1.36)	333(1.19)	285(**2.10**)
rule6	100(1.00)	280(1.00)	136(1.00)
rule7	405(1.00)	793(**2.09**)	429(1.00)
rule8	522(1.29)	380(1.00)	664(1.55)
rule14	197(1.00)	712(**2.10**)	257(1.00)
rule15	399(**2.03**)	339(1.00)	378(1.47)

each competitive set can get about twice strength than the weakest. It can also be observed that almost every rule is used at least once in each trial. This fact indicates that our learning module makes a sufficient search.

We think that the following two factors mainly contribute to these favorable results. As the first factor, by adopting a high-order (the fifth) polynomial function of Eq. (3) rather than linear functions, the reward R can vary sharply as e exceeds e_m. This leads to the correction of wrong learning results obtained so far. The second factor is the modification of the correlation functions. For example, in the second trial in Table 2, the modification of wrong correlation functions led to resolve competition between rule 2 and rule 3; that is, only rule 2 become to be selected after the halfway of the experiment.

6. Conclusion

We have implemented a multimodal anthropomorphic agent with a learning module that can learn cooperative answering strategy from speech dialogs. Although the experiment reported here is still a small scale, obtained results show the capability of our system. Giving such an autonomy to intelligent agents like VSA will be effective for reducing the design cost in various situations. As shown here, multimodality based on multimedia technologies and intelligence based on AI technologies may be two key factors for realizing advanced user-friendly interfaces in a new information environment.

References

[1] H. Dohi and M. Ishizuka. Realtime Synthesis of a Realistic Anthropomorphous Agent toward Advanced Human–Computer Interaction. *Proc. 5th Int'l Conf. on Human–Computer Interaction(HCI'93)*, pages 152–157, 1993.

[2] H. Dohi and M. Ishizuka. A Visual Software Agent connected with the WWW/Mosaic. *Proc. Multimedia Japan'96*, pages 392–397, 1996.

[3] H. Dohi and M. Ishizuka. Visual Software Agent Interface with Realistic Face and Voice-controlled Netscape. *Proc. Int'l Conf. on Computational Intelligence and Multimedia Applications (ICCIMA-97)*, 1997.

[4] T. Gaasterland et al. An Overview of Cooperative Answering. *J. Intell. Inf. Syst.*, Vol. 1:123–157, 1992.

[5] J. J. Grefenstette. Credit Assignment in Rule Discovery Systems Based on Genetic Algorithms. *Machine Learning*, Vol. 3:225–245, 1988.

[6] O. Hasegawa et al. Active Agent Oriented Multimodal Interface System. *Proc. IJCAI–95*, pages 82–87, 1995.

[7] Y. Hiramoto, H. Dohi, and M. Ishizuka. A Speech Dialogue Management System for Human Interface employing Visual Anthropomorphous Agent. *Proc. 3rd IEEE Int'l Workshop on Robot and Human Communication(RO-MAN'94)*, 1994.

[8] A. Kay. Computer Software. *Scientific American*, Vol. 251(No. 3):52–59, 1984.

[9] B. Laurel. Interface Agents: Metaphors with Character. In *The Art of Human–Computer Interface Design (B. Laurel ed.)*, pages 355–365. Addison–Weslay, 1990.

[10] M. Mori, H. Dohi, and M. Ishizuka. A Multi–Purpose Dialogue Management System employing Visual Anthoropomorphous Agent. *Proc. 4th IEEE Int'l Workshop on Robot and Human Communication(RO-MAN'95)*, pages 187–912, 1995.

[11] A. Takeuchi and K. Nagao. Communicative Facial Display as a New Conversational Modality. *Proc. 5th Int'l Conf. on Human–Computer Interaction (HCI '93)*, pages 187–193, 1993.

[12] K. Yamada, R. Mizoguchi, and N. Harada. User's Utterance Model and Cooperative Answering for Question–Answering Systems (in Japanese). *Trans. Information Processing Society of Japan*, Vol.35(No.11):2265–2275, 1994.