Dialog Strategy Acquisition and Its Evaluation for Efficient Learning of Word Meanings by Agents

Ryo Taguchi, Kouichi Katsurada, and Tsuneo Nitta

Graduate School of Engineering, Toyohashi University of Technology
1-1 Hibariga-oka, Tempaku-cho, Toyohashi-city, 441-8580 Japan
{taguchi, katurada, nitta}@vox.tutkie.tut.ac.jp
http://www.vox.tutkie.tut.ac.jp/

Abstract. In word meaning acquisition through interactions among humans and agents, the efficiency of the learning depends largely on the dialog strategies the agents have. This paper describes automatic acquisition of dialog strategies through interaction between two agents. In the experiments, two agents infer each other's comprehension level from its facial expressions and utterances to acquire efficient strategies. Q-learning is applied to a strategy acquisition mechanism. Firstly, experiments are carried out through the interaction between a mother agent, who knows all the word meanings, and a child agent with no initial word meaning. The experimental results showed that the mother agent acquires a teaching strategy, while the child agent acquires an asking strategy. Next, the experiments of interaction between a human and an agent are investigated to evaluate the acquired strategies. The results showed the effectiveness of both strategies of teaching and asking.

1 Introduction

As demand grows for more natural communication with human-like agents such as anthropomorphic agents, avatars, animated agents, talking heads, etc., research and development have begun on allowing people to communicate with such agents using the advanced interface of multi-modal interaction (MMI).

When receiving input modalities such as speech and gesture, the agents with MMI integrate the multiple modalities and interpret them, then they hold a conversation sometimes according to the context. However, because the current MMI technologies enable us only to interact with the agents along dialog scenarios described by system designers, the applications are quite limited. MMI without dialog scenarios requires autonomous agents who acquire the knowledge of multimedia objects in the real world and interact with humans under this commonly grounded knowledge. Namely, the agents need to acquire the meanings of words through interactions with end-users in the real world. Moreover, automatic acquisition of dialog strategies used for word meaning acquisition is another important function when making the learning effective.

Roy et al. [1] and Iwahashi et al. [2] respectively proposed mechanisms to acquire word meanings that represent relations among visual features of objects and acoustic features of human speeches using machine learning methods. With the help of these

P. Vogt et al. (Eds.): EELC 2006, LNAI 4211, pp. 45 – 56, 2006.

mechanisms, robots learn to understand word meanings in the real world. On the other hand, Levin et al. [3] and Singh et al. [4] investigated how to adapt dialog strategies to the environment by applying reinforcement learning to a human-machine interaction corpus. In the two approaches, interpretation of unknown utterances and adaptation of dialog strategies are separately investigated, however, we should apply both of them to the interaction at the same time because interpretation and dialog control depend on each other. In the case of children's language acquisition, a parent teaches his/her child through cooperative interaction and the child acquires not only word meanings but also dialog strategies.

We are developing Infant Agents (IAs) that are modeled after the word meaning acquisition and the dialog strategy acquisition process of human infants [5]. IAs automatically acquire dialog strategies through interactions among IAs. In this paper, we propose a method for dialog strategy acquisition that uses each other's comprehension level which is inferred from an agent's or a human's facial expressions and utterances. The experiments on dialog strategy acquisition are carried out through interaction between two IAs. Then, we confirm the effectiveness of the acquired strategies through the interaction between a human and an IA.

In section 2 we explain IAs and a dialog strategy acquisition method. In section 3 we conduct the experiments on automatic acquisition of dialog strategies. In section 4 we evaluate the acquired strategies. Lastly, in section 5, we describe the conclusions of this paper.

2 Infant Agent

We are developing IAs modeled after human infants [5]. IAs learn word meanings through human–IA interaction, then share the word meanings acquired by each other via a network. This knowledge-sharing is achieved through interaction in the same manner that humans learn from one another. Furthermore, IAs, through interaction between themselves, automatically acquire dialogue strategies which are needed to efficiently learn each other's knowledge.

2.1 Learning of Word Meanings

The experiments are executed in a virtual space on a computer. There are nine objects in this space. Each object has visual features such as shape and color. Such object features are categorized into six types (globe, triangle, cube, red, blue, and white). In the first step of the learning, an IA chooses an object and asks a question to a counterpart, which is a human or another IA. This question is performed by pointing at the object, not by speaking. We call this asked object a topic object. Then, the counterpart teaches one or two words representing the object features by his/her/its utterance. For simplicity, we assume that IAs have a mechanism for converting the utterance into a phoneme sequence, and IAs directly receive a phoneme sequence as the utterance (see Fig. 1). However, each phoneme contained in the sequence is incorrectly received with probability 0.1. When receiving a pair of a phoneme sequence and object features, the IA divides the phoneme sequence into words by referring to its Word Memory that stores frequencies of words and object features. If the phoneme sequence contains an unknown word, the IA newly registers the pair of the word and the object features in its Word Memory. If known words appear in the phoneme sequence, the IA increments their frequencies. Then, the IA calculates the conditional probability $P(x \mid w)$, where w is a word and x is an object feature. If the probability meets the following three conditions, the IA acquires a relation between the word w and the feature x as the meaning of w.

(1) A word w has been learned more than 3 times.
(2) $P(x \mid w) > 0.9$
(3) Only one feature x meets the condition (2).

Object Feature	Phoneme Sequence
Globe	/maru/
Triangle	/saNkaku/
Cube	/sikaku/

Object Feature	Phoneme Sequence
Red	/aka/
Blue	/ao/
White	/siro/

Fig. 1. Object features and phoneme sequences. Each phoneme sequence represents the object feature written on the left.

2.2 IA's Actions and Facial Expressions

In the above section, we explained the algorithm for the learning of word meanings that is applied to IAs. In the IA–IA interaction or the human–IA interaction, IAs carry out in turn the actions such as asking a question and teaching words. The actions are chosen according to the IA's dialog strategy. The dialog strategies should be controlled according to the comprehension level of each other. For example, when there are few words that are known by a counterpart, it is more efficient to teach one word because it is difficult for IAs to divide an utterance containing two words into correct words. And IAs have to choose according to each other's comprehension level either by teaching or asking. However, the counterpart's comprehension level must be inferred by IAs because it cannot be referred to directly. Therefore, in this paper, we assume that IAs and humans change their own facial expression as a representation of their comprehension of the counterpart's last utterance. We also give IAs a mechanism that infers the comprehension level from the counterpart's facial expressions and utterances. In the following sections, we explain the actions, facial expressions, and inference mechanism of IAs.

2.2.1 Actions

IAs can carry out the following six actions.

(1) "NO ACT": Nothing is done.
(2) "CHANGE A TOPIC": An IA changes a topic by choosing an object randomly.
(3) "ASK": An IA asks a question by pointing at a topic object.
(4) "IMITATE": An IA imitates a counterpart's utterance.
(5) "ADD A WORD": An IA randomly chooses a word from its Word Memory and adds it to its Speech Register.
(6) "SPEAK": An IA utters the contents of its Speech Register.

A dialog starts when one of the IAs carries out "ASK" after "CHANGE A TOPIC". Then, IAs perform in turn one of the six actions according to their strategies which are under learning. After an IA carries out the above (1), (3), (4) or (6), the counterpart's turn comes, or the IA takes its turn again after carrying out "CHANGE A TOPIC" or "ADD A WORD".

"CHANGE A TOPIC" can be carried out repeatedly up to 9 times. This means that an IA can strategically choose one from nine objects. For example, according to a strategy it can perform "CHANGE A TOPIC" until it finds an unknown object.

"ADD A WORD" can be carried out repeatedly up to 2 times and "SPEAK" can be carried out after it. "ADD A WORD" is an action to randomly choose a word from the IA’s Word Memory and add it to its Speech Register that stores words to be uttered. However, we assume that the added words represent features of a topic object. IAs construct a teaching utterance by iterating this action. "SPEAK" is an action to speak the content stored in the Speech Register.

"IMITATE" is an action to imitate a counterpart's utterance. By performing this action, IAs can check whether taught utterances were conveyed correctly.

2.2.2 Inference of Comprehension Level

When the counterpart's facial expression becomes comfortable or neutral after an IA teaches words, the IA considers that the counterpart already knows the correct meanings of the words, and so adds the words to its own Shared Word Memory. The Shared Word Memory stores those words that have been shared with a counterpart. And, when a counterpart's utterance does not conflict with the knowledge held by an IA, the IA adds the uttered words to its own Shared Word Memory. For their own dialog strategies, IAs use the information of the Shared Word Memory as a reflection of the counterpart's comprehension level.

2.2.3 Facial Expressions

IAs have three types of facial expression: neutral, comfortable, and uncomfortable. When new word meanings are acquired or new words are shared, an IA's facial expression becomes comfortable. If an unknown word is contained in a counterpart’s utterance, the IA's facial expression becomes uncomfortable. In other cases, the IA's facial expression remains neutral.

2.3 Dialog Strategy Acquisition

We use Q-learning as a strategy acquisition mechanism. Q-learning [6] is one of the online Reinforcement Learning algorithms and is widely used to optimize an agent's behavior. The following sections explain the states and rewards which are used in Q-learning.

2.3.1 States of Q-Learning

In order to acquire dialog strategies based on a past dialog history and the comprehension level of a counterpart, IAs recognize states by using not only the information for expressing current dialog situations (such as a counterpart's facial expressions and actions) but the contents of each IA’s Word Memory and Shared Word Memory. Specifically, a state is recognized for the following 8-dimensional information.

(1) Counterpart's action
(2) Counterpart's facial expression
(3) Own last action
(4) Own facial expression
(5) The number of words that represent features of a topic object
(6) The number of shared words in (5)
(7) The number of words in the Speech Register
(8) The number of shared words in the Speech Register

2.3.2 Rewards of Q-Learning

In order to realize cooperative learning of word meanings, rewards should be given according to not only an IA's own learning situation but also the counterpart's learning situation. Therefore, we calculate rewards r as follows.

$$r = \frac{r_1 + r_2}{2}, \tag{1}$$

where r_1 is a reward according to the IA's own learning situation and r_2 is a reward according to the counterpart's facial expression.

Table 1 shows actual values of rewards.

Table 1. Rewards of Q-learning

	r_1
New word meanings were acquired, where n is the number of the acquired words at once.	20n
New words were added into IA's Shared Words Memory.	5
"CHANGE A TOPIC" or "ADD A WOED" was paformed.	0
The other cases.	-1

	r_2
The counterpart's facial expression became comfortable.	5
The counterpart's facial expression became neutral or uncomfortable.	-1

3 Experiments of Dialog Strategy Acquisition

3.1 Experimental Setup

We investigate the acquisition of dialog strategies through the interaction between two IAs designed in the previous section. Six words are used in our experiments, where all of them are given as initial knowledge to the IA that is called IA1. The other IA is called IA2 and does not have the initial knowledge.

We define a dialog as a sequence of interaction until IA2 acquires all word meanings or until 100 turns have passed. At the beginning of each dialog, the word meanings learned by IA2 are reset, however, the strategy of each IA is preserved and is continuously learned across the dialogs. In the experiment, 100,000 dialogs are iterated and both IAs acquire dialog strategies according to their own initial knowledge. Each IA decides its action based on the ε-greedy policy in the learning step, where ε is set to 0.1. The learning rate of Q-learning is decreased from 0.1 to 0.001, according to the frequency of learning, and the discount rate of Q-learning is set to 0.9.

After the dialog strategy acquisition, we execute 1,000 dialogs between IA1 and IA2 following the acquired strategies in order to evaluate their efficiency. In the dialogs to be evaluated, the IAs do not learn dialog strategies. Each IA chooses its action on the basis of the greedy policy that is following the highest Q-value at any time. Then, we compare its efficiency with the efficiency of the interaction between two agents whose strategies have been designed by us; one of the agents is an IA that has been given the role of teacher and is called TA, and the other is an IA that has been given the role of learner and is called LA. The TA has the initial knowledge while the LA has none. Their strategies are as follows: The TA randomly chooses between "ADD A WORD" and "SPEAK". The LA randomly chooses an object and asks about it.

3.2 Experimental Results

First, we show how the actions of IA1 and IA2 change by the strategy acquisition. Figure 2 shows that the IAs before the strategy acquisition choose each action with a constant proportion regardless of dialog situations. On the other hand, Figure 3 shows that the IAs after the strategy acquisition choose actions according to the progress of the dialogs. And, after the strategy acquisition, IA1 has a tendency to choose teaching actions such as "ADD A WORD" and "SPEAK", and IA2 has a tendency to choose learning actions such as "CHANGE A TOPIC", "ASK" and "IMITATE". These results show that each IA acquired a teaching strategy or an asking strategy according to its own initial knowledge.

Next, we show an example of interaction between the IAs that follow the acquired strategies in Fig. 4. In the first turn, IA1 taught the word “maru”. Then IA2 tried to imitate IA1's utterance but heard it wrongly and said, "matu." In the second turn, IA1 made IA2 correctly learn the word by teaching "maru" again. When IA2 could correctly speak the word, IA1 considered that the word had been shared with IA2. Then IA1 began to teach two words as shown in the third turn. In each of the first, second and fifth turns, IA1 taught one word, because "maru" or "sikaku" had not been shared. IA1’s strategy can efficiently teach words according to the counterpart’s comprehension level. IA2 could learn correct words by imitating IA1’s utterance as shown in the second turn. Moreover, when IA2's facial expression was neutral or IA2 acquired new word meanings, IA2 changed the topic object, as shown in the fifth, eighth and ninth turns. IA2’s strategy is efficient because it can appropriately change a topic object according to IA2's own comprehension level.

Finally, we show the efficiency of the IA1–IA2 interaction and the TA–LA interaction in Fig. 5. Figure 5 shows that the IA1–IA2 interaction is more efficient than the TA–LA interaction.

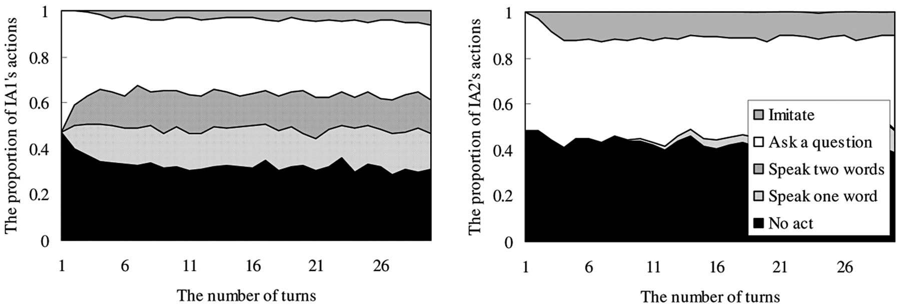

Fig. 2. Actions of IA1 (left) and IA2 (right) before strategy acquisition

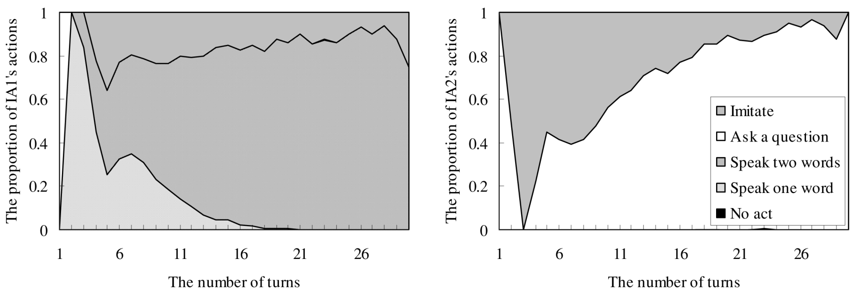

Fig. 3. Actions of IA1 (left) and IA2 (right) after strategy acquisition

0	start of dialog
1	IA2:(- -) IA2 indicated a bule globe. IA1:(- -) IA1 said, "maru."
2	IA2:(x x) IA2 imitated IA1's utterance and said, "matu." IA1:(x x) IA1 said, "maru."
3	IA2:(x x) IA2 imitated IA1's utterance and said, "maru." IA1:(^_^) IA1 considered "maru" as a shared word and said, "maruao."
4	IA2:(x x) IA2 imitated IA1's utterance and said, "maruao." IA1:(^_^) IA1 considered "ao" as a shared word and said, "aomaru."
5	IA2:(- -) IA2 indicated a red cube. IA1:(- -) IA1 said, "sikaku."
6	IA2:(x x) IA2 imitated IA1's utterance and said, "sikaku." IA1:(^_^) IA1 considered "sikaku" as a shared word and said, "sikakuaka."
7	IA2:(x x) IA2 imitated IA1's utterance and said, "sikakuaka." IA1:(^_^) IA1 considered "aka" as a shared word and said, "akasikaku."
8	IA2:(- -) IA2 indicated a red triangle. IA1:(- -) IA1 said, "sankakuaka."
9	IA2:(x x) IA2 acquired the meaning of "aka" and indicated a blue cube. IA1:(- -) IA1 said, "sikakuao."
10	IA2:(^_^) IA2 acquired the meanings of "sikaku" and "ao".

Fig. 4. Interaction example between IA1 and IA2

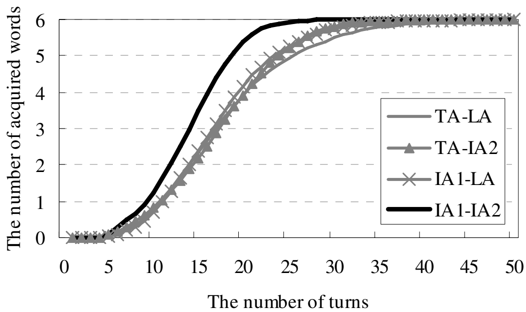

Fig. 5. Performance of acquired strategies. The horizontal axis represents the number of turns while the vertical axis represents the number of words correctly acquired by IA2 or the LA.

4 Evaluation of the Acquired Strategies Through Human–IA Interaction

In this section, through interaction between a human and an IA, we evaluate the strategies that have been acquired by IA1 and IA2 in the above section.

In order to realize human–IA interaction on the basis of the same conditions as the interaction between IAs, we developed an experimental system that enables a human to teach or learn word meanings through interaction with an IA. Figure 6 shows the execution screen of the experimental system. In the virtual space of the experiment, there are an IA and nine objects. A human, or an experimental subject, can choose the face icon that represents his/her facial expression every time an IA speaks. When it is the human’s turn, he/she can point at an object or input his/her utterance from a keyboard. The IA's utterance is displayed on a computer screen for three seconds.

We prepare the following two types of experiment.

Experiment (1): Evaluation of acquired teaching strategies: an experimental subject plays the role of a learner and learns word meanings from utterances of IAs.

Experiment (2): Evaluation of acquired learning strategies: an experimental subject plays the role of a teacher and teaches word meanings to IAs.

Fig. 6. Human–IA Interaction System

We evaluate the performance of each acquired strategy. In both experiments, the subjects are seven undergraduates. In experiment (1), they interact with IA1 and the TA given only the role of a teacher. In experiment (2), they interact with IA2 and the LA. In the experiments, subjects are requested to understand IAs' strategies (roughly) in preliminary interaction with the IAs.

4.1 Evaluation of the Acquired Teaching Strategies

In this experiment, experimental subjects learn word meanings from utterances of IAs.

4.1.1 Experimental Setup

The IAs teach the same six words as the above experiments of Section 3. Each word represents the shape or color of objects such as "cube" or "red". However, the experiments will fail if the experimental subjects already know these words. Therefore, in order to prepare words that are not known by any subject, IAs automatically generate the phoneme sequences of the words at the beginning of each dialog. Specifically, IAs randomly choose 2 to 4 syllables from 45 types of syllable, which are prepared beforehand by us, and combine them to create words such as "HEKE" and "EUREKA". The experimental subjects know that a word spoken by IAs represent either the shape or color of objects. Namely, in this task, the experimental subjects translate IAs' language into our language. The IA's utterance is displayed on a computer screen for three seconds. Every time IAs speak, the experimental subjects write down the meanings of the words on specified paper.

We compare IA1 that acquired the teaching strategy through the above experiment with the TA that decides randomly on the number of teaching words and speaks them.

4.1.2 Experimental Results and Discussion

In the TA–human interaction, the experimental subjects took about 12 turns to acquire the above six words on average. On the other hand, in the case of the IA1–human interaction, it was about 8 turns (see Fig. 7).

By using the paired t-test, we assessed the significance of the difference between TA and IA1. The result showed a significant difference between them (level of significance of the test $P = 0.02 < 0.05$) and demonstrated the effectiveness of IA1's strategy on the interaction with humans.

Compared with the above IA–IA interaction (see Fig. 5), the experimental subjects were able to acquire the words about two times as quickly as IAs. One of the reasons for the result is the fact that humans use relationships between words to acquire word meanings. For example, when a red cube is called "KOKE" and then a red triangle is called "KOKEMOMO", IAs acquire the meanings of "KOKE" as a word representing the color red. Also, the experimental subjects too are able to acquire the meanings. Furthermore, by using the knowledge "word meanings do not overlap each other", they can acquire the meaning of "MOMO" as a word representing the shape of a triangle.

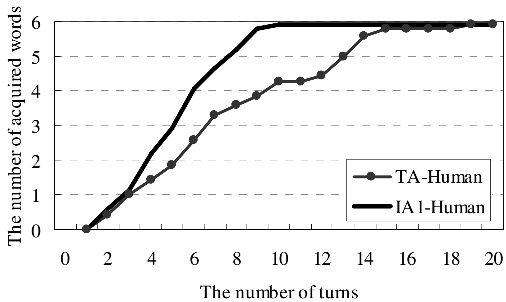

Fig. 7. The number of words correctly acquired by humans when IAs are teachers

4.2 Evaluation of the Acquired Learning Strategies

In this experiment, experimental subjects teach word meanings to IAs.

4.2.1 Experimental Setup

The experimental subjects teach the same six words as in the above experiment.

We compare IA2 that acquired the learning strategy through the experiment of Section 3 with the LA that invariably asks about a different object. In the above IA–IA interaction, IAs can teach multiple words at a time, but in this interaction, the experimental subjects invariably teach only one word in order to simplify their task. We assume that there is no recognition error.

4.2.2 Experimental Results and Discussion

In the human–LA interaction, the LA took about 30 turns to acquire the above six words on average. On the other hand, in the case of the human–IA2 interaction, it was about 24 turns (see Fig. 8).

By using the paired t-test, we assessed the significance of the difference between LA and IA2. The result showed a significant difference between them (level of significance of the test $P = 0.03 < 0.05$) and demonstrated the effectiveness of IA2's strategy on the interaction with humans.

IA2 becomes more efficient as it acquires more words, because, as discussed, IA2 asks preferentially about objects for which it has not yet acquired many words.

Next, in order to compare the teaching strategies of the experimental subjects with IA's strategy under the same conditions, we show the results of the interaction between the TA' that invariably teaches only one word and the LA or IA2 in Fig. 8. In this interaction, we assume that there is no recognition error. Figure 8 shows that the experimental subjects can teach more efficiently words to both LA and IA2 than the TA'. One of the reasons for the result is the fact that humans can memorize teaching history including IAs' response and can intentionally teach those words that have not yet been acquired by IAs. Moreover, the experimental subjects were using not only their memories but also high reasoning ability. For example, when an IA2 sequentially indicated cubes of different colors, an experimental subject inferred that "IA2 does not know the word representing cube" from the behavior of IA2. However, there were some ineffectual teachings, such as an experimental subject taught words that

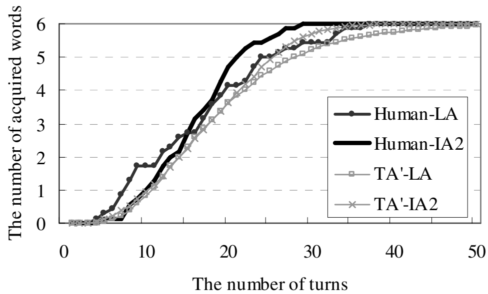

Fig. 8. The number of words correctly acquired by IA2 or the LA,when humans are teachers

had already been acquired by an IA, so the difference between human's teachings and IA's teachings is smaller than that of Section 4.1.2.

5 Conclusion

In this paper, we proposed a novel method for dialog strategy acquisition that uses the counterpart's comprehension level estimated from the agent's facial expressions and utterances. The experimental results from investigating the interaction between two IAs showed that both IAs can acquire efficient strategies according to their own initial knowledge. The actions of IAs are effectively selected by the strategies according to the estimated knowledge of the counterpart. Moreover, in the experiments of human–IA interaction, the acquired strategies of IAs are also effective for humans.

In a future work, in order to make IAs acquire both strategies of teaching and learning, we will conduct experiments in which initial knowledge is given to both IAs, as well as analyze the detailed human behavior in human–IA interaction and use the findings to study the IA's dialog strategy.

Acknowledgments

This research was supported by a 21st Century COE Program Grant for "Intelligent Human Sensing".

References

[1] D. Roy, "Integration of speech and vision using mutual information," in Proc. ICASSP, Vol. 4, pp. 2369-2372, 2000.

[2] N. Iwahashi, "Language acquisition through a human-robot interface by combining speech, visual, and behavioral information," Information Sciences, Vol. 156, pp. 109-121, 2003.

[3] S. Singh, M. Kearns, D. Litman, and M. Walker, "Empirical Evaluation of Reinforcement Learning Spoken Dialog System," in Proc. AAAI, pp. 645-651, 2000.

[4] E. Levin, R. Pieraccini, and W. Eckert, “A stochastic model of human-machine interaction for learning dialog strategies,” IEEE Transactions on Speech and Audio Processing, Vol. 8(1), pp. 11-23, 2000.
[5] R. Taguchi, K. Katsurada and Tsuneo Nitta, “Automatic Acquisition of Dialog Strategies for Concept Learning through Interaction among Agents”, Proc. of CIS'04, FM7.1, 2004.
[6] C.J.C.H. Watkins, P. Dayan, “Technical Note: Q-learning”, Machine Learning 8, pp. 279-292, 1992.