

User-adaptive Coordination of Agent Communicative Behavior in Spoken Dialogue

Kohji Dohsaka

NTT Communication Science Laboratories
NTT Corporation
2-4, Hikaridai, Seika-cho,
Kyoto 619-0237, Japan

Atsushi Kanemoto

Graduate School of
Information Science and Technology
Osaka University, 1-1, Yamadaoka,
Suita, Osaka 565-0871, Japan

Ryuichiro Higashinaka

NTT Cyber Space Laboratories
NTT Corporation
1-1, Hikarinooka, Yokosuka,
Kanagawa 239-0847, Japan

Yasuhiro Minami and Eisaku Maeda

NTT Communication Science Laboratories
NTT Corporation
2-4, Hikaridai, Seika-cho,
Kyoto 619-0237, Japan

{dohsaka, minami, maeda}@cslab.kecl.ntt.co.jp
higashinaka.ryuichiro@lab.ntt.co.jp

Abstract

In this paper, which addresses smooth spoken interaction between human users and conversational agents, we present an experimental study that evaluates a method for user-adaptive coordination of agent communicative behavior. Our method adapts the pause duration preceding agent utterances and the agent gaze duration to reduce the discomfort perceived by individual users during interaction. The experimental results showed a statistically significant tendency: the duration of the agent pause and the gaze converged during interaction with the method. The method also significantly improved the perceived relevance of the agent communicative behavior.

1 Introduction

Conversational agents have been studied as an effective human-computer interface for such purposes as training decision-making in team activities (Traum and Rickel, 2002), learning support (Johnson et al., 2002), museum guides (Kopp et al., 2005), and community facilitators (Zheng et al., 2005; Fujie et al., 2009). They will play a crucial role in establishing a society where humans and robots collaborate through natural interaction. However, agents cannot produce their intended effects when the smooth flow of interaction is disturbed. To fully exploit the promise of agents, we need to achieve smooth interaction between human users and agents.

Although various types of modalities have been used in human-computer interfaces, speech has drawn a great deal of interest because it is one of the most pervasive communication methods in our daily lives and we usually perform it without any special effort (Nass and Brave, 2005). In this paper, we are interested in smooth spoken dialogues between users and agents.

A spoken dialogue is a joint activity among participants (Clark, 1996). For such a joint activity to be smooth and successful, participants need to coordinate their communicative behaviors in various ways. In human dialogues, participants agree on lexical choices to refer to objects (Brennan and Clark, 1996) and coordinate eye gaze (Richardson and Dale, 2005) and whose turn it is to speak (Sacks et al., 1974). They become more similar to their partners as the dialogue proceeds in many aspects such as pitch, speech rate, and pause structure (Burgoon et al., 1995; Hayashi et al., 2009). Such coordination serves to make conversation flow easily and intelligibly (Garrod and Pickering, 2004).

The coordination of communicative behaviors also plays a crucial role in smooth human-agent interaction. Previous work addressed human behavior adaptation to agents (Oviatt et al., 2004), agent behavior adaptation to human partners (Mitsunaga et al., 2005; Tapus and Matarić, 2007), and the mutual adaptation of human and agent behavior (Breazeal, 2003).

In this paper, which addresses smooth spoken interaction between human users and agents, we focus on the adaptation of agent communicative behavior to individual users in spoken dialogues

with flexible turn-taking. We present a method for user-adaptive coordination of agent communicative behavior to reduce the discomfort perceived by individual users during the interaction and show experimental results that evaluate how the method influences agent communicative behavior and improves its relevance as perceived by users. For evaluation purposes, we used a quiz-style multi-party spoken dialogue system (Minami et al., 2007; Dohsaka et al., 2009). A quiz-style dialogue is a kind of thought-evoking dialogue that can stir user thinking and activate communication (Higashinaka et al., 2007a; Dohsaka et al., 2009). This characteristic is expected to be advantageous for evaluation experiments since it encourages involvement in the dialogue.

Our method adapts agent communicative behavior based on policy gradient reinforcement learning (Sutton et al., 2000; Kohl and Stone, 2004). The policy gradient method has been used for robot communicative behavior adaptation (Mitsunaga et al., 2005; Tapus and Matarić, 2007). However, both studies dealt with scenario-based interaction in which a user and a robot acted with predetermined timing. In contrast, we focus on spoken dialogues in which users and agents can speak with more flexible timing. In addition, we allow for two- and three-party interactions among a user and two agents. It remains unclear whether the policy gradient method can successfully adapt agent communicative behavior to a user in two- or three-party spoken dialogues with flexible turn-taking. Although this paper focuses on agent behavior adaptation to human users, we believe that our investigation of the agent behavior adaptation mechanism in flexible spoken interaction will contribute to conversational interfaces where human users and agents can mutually adapt their communicative behaviors.

As agent communicative behavior to be adapted, this paper focuses on the pause duration preceding agent utterances and the agent gaze duration. In conversation, the participant pause duration is influenced by partners, and the coordination of pause structure leads to smooth communication (Burgoon et al., 1995; Hayashi et al., 2009). Without pause structure coordination, undesired speech overlaps or utterance collisions are likely to occur between users and agents, which may disturb smooth communication. Funakoshi *et al.* proposed a method to prevent undesired speech overlaps in human-robot speech interactions by using

a robot's subtle expressions produced by a blinking LED attached to its chest (Funakoshi et al., 2008). In their method, a blinking light notifies users about such internal states of the robot as processing or busy and helps users identify the robot pause structures; however we are concerned with the adaptation of robot pause structures to users.

Gaze coordination is causally related to the success of communication (Richardson and Dale, 2005), and the amount of gaze influences conversational turn-taking (Vertegaal and Ding, 2002). The relevant control of agent gaze duration is thus essential to the smooth flow of conversation. Moreover, since the amount of gaze is related to specific interpersonal attitudes among participants and is also subject to such individual differences as personalities (Argyle and Cook, 1976), agent gaze duration must be adapted to individual users.

In the following, Section 2 describes our quiz-style multi-party spoken dialogue system. Section 3 shows our method for the user-adaptive coordination of agent communicative behavior. Section 4 explains the experiment, and Section 5 describes its results. Section 6 concludes our paper.

2 Quiz-Style Spoken Dialogue System

To evaluate a method for agent communicative behavior adaptation, we used a quiz-style multi-party spoken dialogue system based on a quiz-style two-party spoken dialogue system (Minami et al., 2007) and extended it to perform multi-party interaction (Dohsaka et al., 2009).

In this system, a human user and one or two agents interact. The two agents include a quizmaster and a peer. The quizmaster agent creates a “Who is this?” quiz about a famous person and presents hints one by one to the user and the peer agent, who participates in the interaction and guesses the correct answer in the same way that the user does.

The hints are automatically created from the biographical facts of people in Wikipedia¹ and ranked based on the difficulty of solving the quizzes experienced by users (Higashinaka et al., 2007b). Since users must consider the hints to offer reasonable answers, the system can stimulate their thinking and encourage them to engage in the interaction (Higashinaka et al., 2007a). In addition, the peer agent's presence and the agent's empathic expressions improve user satisfaction and

¹<http://ja.wikipedia.org/>



Figure 1: User interacting with two agents using the quiz-style spoken dialogue system

increase user utterances (Dohsaka et al., 2009).

Figure 1 shows a human user interacting with the two agents, both of whom are physically embodied robots. The system utilizes an extremely large vocabulary with continuous speech recognition (Hori et al., 2007). Agent utterances are produced by speech synthesis. The agents can gaze at other participants by directing their faces to them. At each point of the dialogue, the system chooses the next speaker and its utterance based on the dialogue state that the system maintains, the preconditions of the individual utterances, and a few turn-taking rules (Dohsaka et al., 2009). The agent pause and gaze durations are controlled based on the adaptation method described in Section 3.

A sample dialogue among a user and two agents is depicted in Figure 2. *Master* is the quizmaster agent, and *Peer* is the peer agent. The agent utterances are classified as either spontaneous or responsive. Spontaneous utterances are those made after an agent takes his turn in an unforced manner, and responsive utterances are responses to the other's utterances. In the sample dialogue, *spon* identifies spontaneous and *res* identifies responsive utterances.

Quizmaster agent *Master* makes spontaneous utterances such as presenting hints (lines 1 and 5), indicating the quiz difficulty, and addressing listeners. It also makes such responsive utterances as evaluating the other's answers (lines 3, 9, and 11). Peer agent *Peer* makes such spontaneous utterances as showing its own difficulty (line 4), giving an answer (line 8), giving feedback when its own or the other's answer is right (line 12), and addressing listeners. It also makes such responsive utterances as showing empathy to the user (line 7).

3 Method for Agent Communicative Behavior Adaptation

We apply policy gradient reinforcement learning (Sutton et al., 2000; Kohl and Stone, 2004)

- 1 *Master* Who is this? First hint. He graduated from the University of Tokyo. (*hint/spon*)
- 2 *User* Yoshida Shigeru? (*answer/spon*)
- 3 *Master* No, not even close! He's not a politician. (*evaluation/res*)
- 4 *Peer* I don't know. Very difficult. (*show difficulty/spon*)
- 5 *Master* It's time for the second hint: He's a novelist and a scholar of British literature. (*hint/spon*)
- 6 *User* Oh, I think I know it but I can't remember his name. That's so frustrating. (*show difficulty/spon*)
- 7 *Peer* Difficult for me, too. (*show empathy/res*)
- 8 *Peer* Haruki Murakami? (*answer/spon*)
- 9 *Master* Close! You are half right, because he is a novelist. (*evaluation/res*)
- 10 *User* Natsume Soseki? (*answer/spon*)
- 11 *Master* That's right. Wonderful. (*evaluation/res*)
- 12 *Peer* Good job. (*feedback/spon*)

Figure 2: Sample dialogue between user and two agents: quizmaster *Master* and peer *Peer*. *Spon* identifies spontaneous and *res* identifies responsive utterances.

to the user-adaptive coordination of agent communicative behavior. A policy gradient method is a reinforcement learning (RL) approach that directly optimizes a parameterized policy by gradient ascent based on the gradient of the expected reward with respect to the policy parameters. Although RL methods have recently been applied to optimizing dialogue management in spoken dialogue systems (Williams and Young, 2007; Minami et al., 2009), these previous studies utilized RL methods based on the value-function estimation. The policy gradient method is an alternative approach to RL that has the following merits. It can handle continuous and large action spaces (Kimura and Kobayashi, 1998) and is usually assured to converge to a locally optimal policy in such action spaces (Sutton et al., 2000). Moreover, it does not need to explicitly estimate the value function, and it is incremental, requiring only a constant amount of computation per learning step (Kimura and Kobayashi, 1998).

Due to these advantages, the policy gradient method is suitable for adapting agent communicative behavior to a user during interaction, because

- (1) $\Theta = [\theta_j] \leftarrow$ initial policy (policy parameter vector of size n)
- (2) $\epsilon = [\epsilon_j] \leftarrow$ step size vector of size n
- (3) $\eta \leftarrow$ overall scalar step size
- (4) $maxA \leftarrow 0$ (greatest absolute value of reward ever observed in adaptation process)
- (5) while dialogue continues
- (6) for $i = 1$ to T
- (7) for $j = 1$ to n
- (8) $r_j \leftarrow$ random choice from $\{-1, 0, 1\}$
- (9) $R_j^i \leftarrow \theta_j + \epsilon_j * r_j$
(R^i is T random perturbations of Θ)
- (10) for $i = 1$ to T
- (11) Perform a hint dialogue based on policy R^i , and evaluate reward
- (12) for $j = 1$ to n
- (13) $Avg_{+\epsilon,j} \leftarrow$ average reward for all R^i with positive perturbation in parameter ϵ_j
- (14) $Avg_{0,j} \leftarrow$ average reward for all R^i with zero perturbation in parameter ϵ_j
- (15) $Avg_{-\epsilon,j} \leftarrow$ average reward for all R^i with negative perturbation in parameter ϵ_j
- (16) if ($Avg_{0,j} > Avg_{+\epsilon,j}$ and $Avg_{0,j} > Avg_{-\epsilon,j}$)
- (17) $a_j \leftarrow 0$
- (18) else
- (19) $a_j \leftarrow Avg_{+\epsilon,j} - Avg_{-\epsilon,j}$
- (20) $\forall j (a_j \leftarrow \frac{a_j}{|A|} * \epsilon_j * \eta)$
- (21) $maxC \leftarrow$ maximum of absolute value of reward in current adaptation cycle
- (22) if ($maxC > maxA$)
- (23) $maxA \leftarrow maxC$ (update $maxA$)
- (24) else
- (25) $A \leftarrow A * \frac{maxC}{maxA}$
- (26) $\Theta \leftarrow \Theta + A$

Figure 3: Pseudocode for user-adaptive coordination of agent communicative behavior

it can naturally incorporate such continuous parameters as pause and gaze duration and incrementally adapt agent behavior. In fact, the policy gradient method has been successfully used for robot behavior adaptation (Mitsunaga et al., 2005; Tapus and Matarić, 2007). In this paper, we apply this method to agent communicative behavior adaptation in spoken dialogues with flexible turn-taking.

Figure 3 shows our method for the user-adaptive coordination of agent communicative behavior. This method is a modification of an algorithm presented by Kohl and Stone (2004) in that the gradient is adjusted based on the maximum absolute

value of the reward obtained during each adaptation cycle.

The agent communicative behaviors are determined based on a policy that is represented as vector $\Theta (= [\theta_j])$ of n policy parameters. In the quiz-style dialogues, the behavior of both the quizmaster and peer agents is controlled based on the same policy parameters. The method adapts the behavior of both agents to individual users by adapting the policy parameters. In this experiment, we used the following four parameters ($n = 4$):

- pre-spontaneous-utterance pause duration σ_{spon} : duration of pauses preceding agent spontaneous utterances
- pre-responsive-utterance pause duration σ_{res} : duration of pauses preceding agent responsive utterances
- gaze duration σ_{gaze} : duration of agent’s directing its face to the other while it is speaking or listening
- hint interval σ_{hint} : interval of presenting quiz hints

As shown above, we used two types of pause duration since the relevant pause duration can be dependent on dialogue acts (Itoh et al., 2009). Although our main concern is the pause and gaze duration, we examined the hint interval as a parameter particular to quiz-style dialogues.

To adapt the policy parameters to individual users, we first generate T random perturbations $[R^1, \dots, R^T]$ of current policy Θ by randomly adding $\epsilon_j, 0, -\epsilon_j$ to each parameter θ_j of Θ in lines 6 to 9, where ϵ_j is a step size set for each parameter. In the experiment, we set T to 10. The step sizes of the parameters used in the experiment will be shown later in Table 1.

Dialogue per hint (a hint dialogue) is then performed based on each perturbation policy R^i , and the reward for each hint dialogue is obtained in lines 10 to 11. All agent behaviors in a hint dialogue are determined based on the same perturbation policy. As we will explain in Section 4, in the experiment, we regarded the magnitude of discomfort perceived by users during a hint dialogue as a negative reward. Users signified discomfort by pressing buttons on the controller held in their hands. After performing hint dialogues for all T perturbations R^i , gradient $A (= [a_j])$ is computed in lines 12 to 19. The gradient is normalized by

Parameters	σ_{spont} (sec.)	σ_{res} (sec.)	σ_{gaze} (sec.)	σ_{hint} (sec.)
Initial value	4.96	0.53	3.04	27.7
Step size	0.50	0.20	0.30	2.5

Table 1: Initial values and step sizes of policy parameters: σ_{spont} (pre-spontaneous-utterance pause duration), σ_{res} (pre-responsive-utterance pause duration), σ_{gaze} (gaze duration), and σ_{hint} (hint interval)

overall scalar step size η and individual step size ϵ_j for each parameter in line 20. Overall scalar step size η is used to adjust the adaptation speed, which we set to 1.0.

Next we get the maximum $maxC$ of the absolute value of the reward in the current adaptation cycle. As in lines 21 to 25, the gradient is adjusted based on the ratio of $maxC$ to the greatest absolute value $maxA$ of reward ever observed in the overall adaptation process. Finally, the current policy parameters are updated using the gradient in line 26.

This is an adaptation cycle. By iterating it, the agent communicative behavior is adapted to reduce the discomfort perceived by each user.

4 Experiment

We recruited and paid 32 Japanese adults (16 males and 16 females) for their participation. The mean ages of the male and female groups were 33.2 and 36.8, respectively. They were divided into two groups: two-party dialogues (user and quizmaster) and three-party dialogues (user, quizmaster, and peer). In each group, the numbers of males and females were identical.

For this experiment, we used a quiz-style spoken dialogue system. We chose the quiz subjects in advance and divided them into sets of five so that the difficulty level was approximately the same in all sets. For this purpose, we made several sets of five people of approximately identical PageRankTM scores based on Wikipedia’s hyperlink structure.

The users first rehearsed the dialogues for a set of five quizzes to familiarize themselves with the system. After practicing, they performed the dialogues to evaluate the adaptation method and took a break per five-quiz set. The presentation order of the quiz sets was permuted to prevent order effect. For each user, the dialogues continued until the user received 150 hints. The adaptation

method was applied during the interaction, and the policy parameters were updated per 10 hint dialogues. As a result, the parameters were updated 15 times through the dialogues. It took about two hours for each user to complete all dialogues.

The policy parameters were updated based on the magnitude of discomfort perceived by users. In this experiment, users were told to concentrate on the discomfort caused by agent pause and gaze duration and signified it by pressing buttons on the controller held in their hands at three levels of magnitude: ‘3’, ‘2’, and ‘1’. The sum of discomfort obtained during a hint dialogue was normalized with respect to the hint dialogue length, and the normalized values were regarded as negative rewards. Ideally we should estimate user discomfort from such user behaviors as pause structure and eye gaze. However, as the first step toward that goal, in this experiment we adopted this setting in which users directly signified their discomfort by pressing buttons.

Table 1 shows the initial values and the step sizes of the policy parameters used in the experiment. To obtain the relevant initial values, we conducted a preparatory experiment in which ten other participants performed quiz-style dialogues under the same conditions as this experiment. The initial values in this experiment were set to the averaged final values of the policy parameters in the preparatory experiment. The step sizes were determined as approximately one-tenth of the initial values except for the pre-responsive-utterance pause, for which the step size was set to 200 msec based on the limits of human perception.

Before and after the adaptation, the users filled out the following questionnaire items (7-point Likert scale) to evaluate the relevance of agent pause and gaze duration:

- Did you feel that the pause duration preceding the agent utterances was relevant?
- Did you feel that the agent gaze duration was relevant while the agents were speaking or listening to you?

5 Results

5.1 Convergence of policy parameters

The policy parameters were updated based on the adaptation method during the user-agent interaction. Figure 4 exemplifies how the policy parameter values changed during the adaptation cycles with a user engaged in the two-party dialogue.

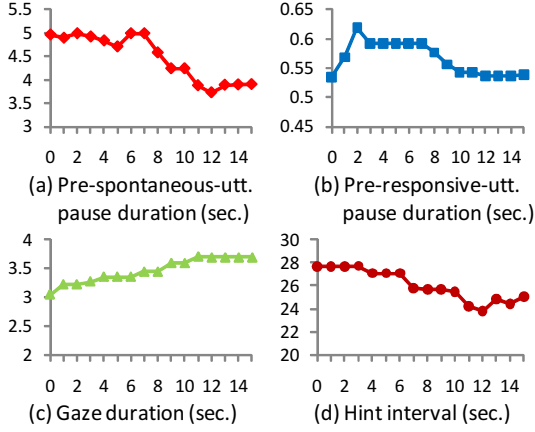


Figure 4: Change of policy parameter values during adaptation cycles with a user engaged in two-party dialogue. Horizontal axis shows adaptation cycles and vertical axis shows parameter values.

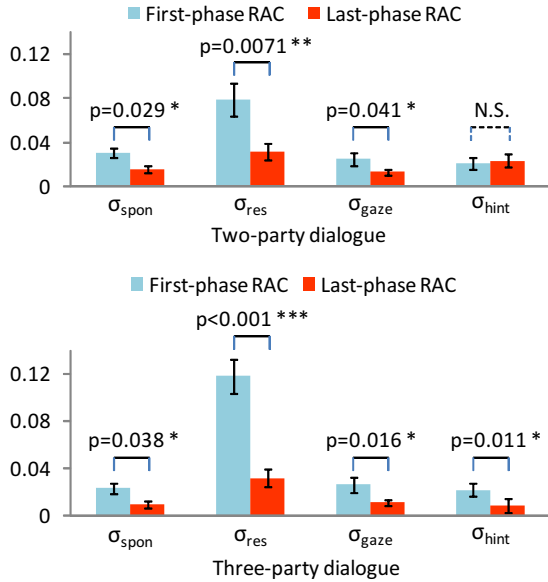


Figure 5: For each policy parameter, average and standard error of first- and last-phase RACs (relative amount of change in parameter values)

Table 2 shows the statistics of the final values of the policy parameters at the end of the adaptation process. Since the initial values were appropriately determined based on the preparatory experiment, the final value averages were not greatly different from the initial values. However, judging from the maximum, minimum, and standard deviations, the final values reflected individual users.

If the adaptation method works successfully, the policy parameter values should converge during the user-agent interaction. From this viewpoint, we examined the relative amount of change in the policy parameters (RAC). Given parameter value p_{k-1} at $(k-1)$ -th adaptation cycle and parameter value p_k at k -th cycle, RAC is defined as

Two-party dialogues				
Parameters	σ_{spon} (sec.)	σ_{res} (sec.)	σ_{gaze} (sec.)	σ_{hint} (sec.)
Average	5.04	0.62	3.10	25.8
Min	3.90	0.39	2.40	19.5
Max	6.17	1.18	3.69	31.2
Sd.	0.72	0.21	0.36	2.7
Three-party dialogues				
Parameters	σ_{spon} (sec.)	σ_{res} (sec.)	σ_{gaze} (sec.)	σ_{hint} (sec.)
Average	4.86	0.62	3.15	27.4
Min	4.07	0.35	2.52	22.0
Max	5.54	0.90	3.58	32.7
Sd.	0.44	0.18	0.27	2.5

Table 2: Statistics of final values of policy parameters: σ_{spon} (pre-spontaneous-utterance pause duration), σ_{res} (pre-responsive-utterance pause duration), σ_{gaze} (gaze duration), and σ_{hint} (hint interval)

$$\frac{|p_k - p_{k-1}|}{p_{k-1}}$$

For each policy parameter, we compared the RAC averages in the first and in the last three adaptation cycles: the first-phase RAC and the last-phase RAC. As shown in Figure 5, the last-phase RAC tends to be smaller than the first-phase RAC. The Kolmogorov-Smirnov test showed that the assumption of normality ($p > 0.2$) was met for each group. By applying the paired Welch’s t-test, as shown in Figure 5, we found that the last-phase RAC is significantly smaller than the first-phase RAC except for the hint interval in the two-party dialogues. This shows that the agent pause and gaze duration converged during the interaction in both the two- and three-party dialogues.

The hint interval is unlikely to converge, probably because it is a longer period than the pause and gaze duration and is subject to various factors. Moreover, it greatly depends on user interest.

5.2 User evaluations

Figure 6 shows the subjective user evaluations of the relevance of agent pause and gaze duration before and after the adaptation. Each user evaluation was measured by a Likert question. The rating of a single Likert question is an ordinal measure, and we generally cannot apply a parametric statistical test to an ordinal measure. Therefore we used a nonparametric test, the Wilcoxon signed-rank test, to compare user evaluations before and after the

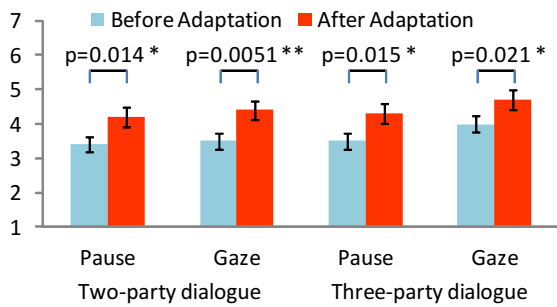


Figure 6: Average and standard error of user evaluations of relevance of agent pause and gaze duration before and after adaptation

adaptation. The F-test for the homogeneity of variances ($p > 0.1$) showed that the data satisfied the statistical test assumption.

We found that in both the two- and three-party dialogues, the relevance of the agent pause and gaze duration significantly improved during the two-hour adaptation process ($p < 0.01$ for gaze duration in the two-party dialogues, $p < 0.05$ for other cases). The p-values are shown in Figure 6. No significant differences between gender were found.

These results on the convergence of policy parameters and user evaluations show that the policy-gradient-based method can adapt agent communicative behavior to individual users in spoken dialogues with flexible turn-taking.

6 Conclusion

In this paper, addressing smooth spoken interaction between human users and conversational agents, we presented a method for user-adaptive coordination of agent communicative behavior and experimentally evaluated how it can adapt agent behavior to individual users in spoken dialogues with flexible turn-taking. The method coordinates agent pause and gaze duration based on policy gradient reinforcement learning to reduce the discomfort perceived by individual users during interaction. We experimentally evaluated the method in a setting where the users performed two- and three-party quiz-style dialogues and signified their discomfort by pressing buttons held in their hands. Our experimental results showed a statistically significant tendency: the agent pause and gaze duration converged during interaction with the method in both two- or three-party dialogues. The method also significantly improved the perceived relevance of the agent communicative behavior in both two- and three-party di-

alogues. These results indicate that in spoken dialogues with flexible turn-taking, the policy-gradient-based method can adapt agent communicative behavior to individual users.

Many directions for future work remain. First, we will analyze how users adapt their communicative behaviors with our method. Second, we need to automatically estimate user discomfort or satisfaction based on such user behaviors as pause structure, prosody, eye gaze, and body posture. Third, we will extend the adaptation method to regulate agent behavior based on dialogue states, since one limitation of the current method is its inability to recognize them. Fourth, we are interested in the adaptation of additional higher-level actions like the relevant choice of dialogue topics based on the level of user interest.

Acknowledgments

This work was partially supported by a Grant-in-Aid for Scientific Research on Innovative Areas, "Founding a creative society via collaboration between humans and robots" (21118004), from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), Japan.

References

- Michael Argyle and Mark Cook. 1976. *Gaze and Mutual Gaze*. Cambridge University Press.
- Cynthia Breazeal. 2003. Regulation and entrainment for human-robot interaction. *International Journal of Experimental Robotics*, 21(10-11):883–902.
- Susan E. Brennan and Herbert H. Clark. 1996. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22:1482–1493.
- Judee K. Burgoon, Lesa A. Stern, and Leesa Dillman. 1995. *Interpersonal Adaptation: Dyadic Interaction Patterns*. Cambridge University Press.
- Herbert H. Clark. 1996. *Using Language*. Cambridge University Press.
- Kohji Dohsaka, Ryota Asai, Ryuichiro Higashinaka, Yasuhiro Minami, and Eisaku Maeda. 2009. Effects of conversational agents on human communication in thought-evoking multi-party dialogues. In *Proc. of SIGDIAL 2009*, pages 217–224.
- Shinya Fujie, Yoichi Matsuyama, Hikaru Taniyama, and Tetsunori Kobayashi. 2009. Conversation robot participating in and activating a group communication. In *Proc. of Interspeech 2009*, pages 264–267.

- Kotaro Funakoshi, Kazuki Kobayashi, Mikio Nakano, Seiji Yamada, Yasuhiko Kitamura, and Hiroshi Tsujino. 2008. Smoothing human-robot speech interactions by using a blinking-light as subtle expression. In *Proc. of ICMI 2008*, pages 293–296.
- Simon Garrod and Martin J. Pickering. 2004. Why is conversation so easy? *Trends in Cognitive Sciences*, 8:8–11.
- Takanori Hayashi, Shohei Kato, and Hidenori Itoh. 2009. A synchronous model of mental rhythm using paralinguistic for communication robots. In *Lecture Notes in Computer Science (PRIMA 2009)*, volume 5925, pages 376–388.
- Ryuichiro Higashinaka, Kohji Dohsaka, Shigeaki Amano, and Hideki Isozaki. 2007a. Effects of quiz-style information presentation on user understanding. In *Proc. of Interspeech 2007*, pages 2725–2728.
- Ryuichiro Higashinaka, Kohji Dohsaka, and Hideki Isozaki. 2007b. Learning to rank definitions to generate quizzes for interactive information presentation. In *Proc. of ACL 2007 (Poster Presentation)*, pages 117–120.
- Takaaki Hori, Chiori Hori, Yasuhiro Minami, and Atsushi Nakamura. 2007. Efficient WFST-based one-pass decoding with on-the-fly hypothesis rescoring in extremely large vocabulary continuous speech recognition. *IEEE Transactions on Audio, Speech and Language Processing*, 15:1352–1365.
- Toshihiko Itoh, Norihide Kitaoka, and Ryota Nishimura. 2009. Subjective experiments on influence of response timing in spoken dialogues. In *Proc. of Interspeech 2009*, pages 1835–1838.
- W. Lewis Johnson, Jeff W. Rickel, and James C. Lester. 2002. Animated pedagogical agents: face-to-face interaction in interactive learning environments. *International Journal of Artificial Intelligence in Education*, 11:47–78.
- Hajime Kimura and Shigenobu Kobayashi. 1998. Reinforcement learning for continuous action using stochastic gradient ascent. In *Proc. of the 5th International Conference on Intelligent Autonomous Systems*, pages 288–295.
- Nate Kohl and Peter Stone. 2004. Policy gradient reinforcement learning for fast quadrupedal locomotion. In *Proc. of ICRA 2004*, volume 3, pages 2619–2624.
- Stefan Kopp, Lars Gesellensetter, Nicole C. Krämer, and Ipke Wachsmuth. 2005. A conversational agent as museum guide: design and evaluation of a real-world application. In *Lecture Notes in Computer Science (IVA 2009)*, volume 3661, pages 329–343.
- Yasuhiro Minami, Minako Sawaki, Kohji Dohsaka, Ryuichiro Higashinaka, Kentaro Ishizuka, Hideki Isozaki, Tatsushi Matsubayashi, Masato Miyoshi, Atsushi Nakamura, Takanobu Oba, Hiroshi Sawada, Takeshi Yamada, and Eisaku Maeda. 2007. The World of Mushrooms: human-computer interaction prototype systems for ambient intelligence. In *Proc. of ICMI 2007*, pages 366–373.
- Yasuhiro Minami, Akira Mori, Ryuichiro Higashinaka, Kohji Dohsaka, and Eisaku Maeda. 2009. Dialogue control algorithm for ambient intelligence based on partially observable Markov decision processes. In *Proc. of IWSDS 2009*.
- Noriaki Mitsunaga, Christian Smith, Takayuki Kanda, Hiroshi Isiguro, and Norihiro Hagita. 2005. Human-robot interaction based on policy gradient reinforcement learning. In *Proc. of IROS 2005*, pages 1594–1601.
- Clifford Nass and Scott Brave. 2005. *Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship*. The MIT Press.
- Sharon Oviatt, Courtney Darves, and Rachel Coulston. 2004. Toward adaptive conversational interfaces: modeling speech convergence with animated personas. *ACM Transactions on Computer-Human Interaction*, 11(3):300–328.
- Daniel C. Richardson and Rick Dale. 2005. Looking to understand: the coupling between speakers’ and listeners’ eye movements and its relationship to discourse comprehension. *Cognitive Science*, 29:1045–1060.
- Harvey Sacks, Emanuel A. Schegloff, and Gail Jefferson. 1974. A simplest systematics for the organization of turn-taking in conversation. *Language*, 50:696–735.
- Richard S. Sutton, David McAllester, Satinder Singh, and Yishay Mansour. 2000. Policy gradient methods for reinforcement learning with function approximation. In *Advances in Neural Information Processing Systems*, volume 12, pages 1057–1063.
- Adriana Tapus and Maja J. Matarić. 2007. Hands-off therapist robot behavior adaptation to user personality for post-stroke rehabilitation therapy. In *Proc. of 2007 IEEE International Conference on Robotics and Automation*, pages 1547–1553.
- David Traum and Jeff Rickel. 2002. Embodied agents for multi-party dialogue in immersive virtual worlds. In *Proc. of AAMAS 2002*, pages 766–773.
- Roel Vertegaal and Yaping Ding. 2002. Explaining effects of eye gaze on mediated group conversations: amount or synchronization. In *Proc. of CSCW 2002*, pages 41–48.
- Jason D. Williams and Steve Young. 2007. Partially observable Markov decision processes for spoken dialog systems. *Computer & Speech Language*, 21(2):393–422.
- Jun Zheng, Xiang Yuan, and Yam San Chee. 2005. Designing multiparty interaction support in Elva, an embodied tour guide. In *Proc. of AAMAS 2005*, pages 929–936.