

Dialog Management of Healthcare Consulting System by Utilizing Deceptive Information

The Tung Nguyen

Nara Institute of Science and Technology, Japan
nguyen.tung.np5@is.naist.jp

Koichiro Yoshino

Nara Institute of Science and Technology, Japan
PRESTO, Japan Science and Technology Agency, Japan
koichiro@is.naist.jp

Sakriani Sakti

Nara Institute of Science and Technology, Japan
RIKEN, Center for Advanced Intelligence Project AIP, Japan
s-sakti@is.naist.jp

Satoshi Nakamura

(affiliation as previous author)
s-nakamura@is.naist.jp

keywords: dialog system, reinforcement learning, Q-learning, POMDP, multimodal, deception.

Summary

In the past few years, there has been an increasing number of works on negotiation dialog. These studies mainly focus on situations where interlocutors work cooperatively to agree on a mutual objective that can fulfill each of their own requirements. However, in real-life negotiation, such situations do not happen all the time, and participants can tell lies to gain an advantage. In this research, we propose a negotiation dialog management system that detects when a user is lying and a dialog behavior for how the system should react when faced with a lie. We design our system for a living habits consultation scenario, where the system tries to persuade users to adopt healthy living habits. We show that we can use the partially observable Markov decision process (POMDP) to model this conversation and use reinforcement learning to train the system's policy. Our experimental results demonstrate that the dialog manager considering deceptive states outperformed a dialog manager without this consideration in terms of the accuracy of action selection, and improved the true success rate of the negotiation in the healthcare consultation domain.

1. Introduction

Over the past few decades, various types of dialog systems have been developed. Due to the growth of dialog-related technologies, the recent focus of dialog system studies covers more extensive areas of dialog: not just simple navigation tasks such as restaurant or tourist information navigation [Devillers 98] but also more complex ones such as persuasion or negotiation [Hiraoka 14, Torning 09]. In persuasion or negotiation, the system has its own goal and will attempt to convince the user to accept an agreement that is similar or close to this goal.

In cooperative persuasion or negotiation, the system and the user work together to reach a balanced point between their goals. Most of the current works mainly deal with this situation (cooperative case). However, there are some cases where users have goals that conflict with the system's goal (non-cooperative case), especially if the user is

required to pay a cost to compromise with the system's goal. In these cases, the mutual goal cannot be reached if both of the system and the user refuse to change their goals [Efstathiou 14, Traum 08].

In the non-cooperative case, deception is a common tactic [Chertkoff 71, Craver 97, Lewicki 85, O'Connor 97, Pérez-Rosas 15]. Using misstatements often provides an immediate reward rather than the delayed reward that is obtained if the negotiation is truly agreed upon. For example, in the case of healthcare consulting, a consultee can use misstatements and receive as an immediate reward less obligation to change their living habits, as compared to when the consultee truly agrees with the consultant to change. The cost of lying is usually uncertain and can be delayed, and the penalty for being caught can be mitigated by providing adequate explanations [Shapiro 91]. The system needs to decide its action for finding a new concession point in these cases, in order to come to a true agreement

with the user.

In this paper, we focus on healthcare consulting as a typical negotiation case, where the user and the system discuss how to find the best treatment for the user. The system tries to convince the users that they should change their living habits to improve their health, and the users hesitate to change their living habits, because disrupting their comfortable lifestyle comes at a cost, even though these habits will have a bad influence on their health in the future. The system needs to consider the user's opinions and preferences when coming up with a treatment plan (recommendation of new habits). In this situation, the user does not always tell the truth in an effort to keep their current habit. According to numerous studies and surveys, this reaction from the consultee is common in health consultation [Blackwell 73, Davis 66].

In response to the above situation, we propose a dialog management system for system action decisions that selects the best actions to make the user honestly agree with the system. We model the persuasive dialog process on partially observable Markov decision process and introduce the deception detection result of a user utterance as a state of the process in addition to the dialog state of the user, as it is difficult to detect deceptions perfectly. The dialog manager tries to find the best policy to increase the rate of true agreement in the persuasion, by utilizing the policy trained by reinforcement learning. The novel contribution of our work is the incorporation of user's deceptive information into the system's dialog management process. As shown further in Section 4.2, automatic deception detection accuracy is not high; however, we can combat the uncertainty of deception information with POMDP framework, which can treat states as stochastic variables to model the system's dialog management.

2. Related Works

A myriad of literature has documented deceptive behavior in human negotiations [Chertkoff 71, Craver 97, Lewicki 85, O'Connor 97]. Subsequently, interest in negotiation agents/systems that can detect lies is increasing. There are a number of works on deception in non-cooperative negotiation using the trading game scenario. [Efsthathiou 14] developed an agent that can tell lies in negotiation with a rule-based adversary. It was shown that the chance of winning for the agent is improved by using lying, thus indicating that deceptions can bring advantages to the negotiators. Furthermore, misstatements tend to result in the interlocutor finding the system's offer to be fairer and feeling more satisfied with the negotiation

outcome [Gratch 16].

The number of studies on countermeasures against deception in negotiation is still very limited. [Gratch 16] gave some examples of how to make counter-arguments when faced with deception in a negotiation, but no further studies on this problem were conducted. With a similar setup to [Efsthathiou 14], [Vourliotakis 14] found that if the adversary is aware that the learning agent is lying, the agent's chance of winning the game will be lower. This suggests that deception information is important if we are negotiating with someone who is lying.

3. System Overview

This section describes the general structure of the proposed dialog system and the focus of this paper.

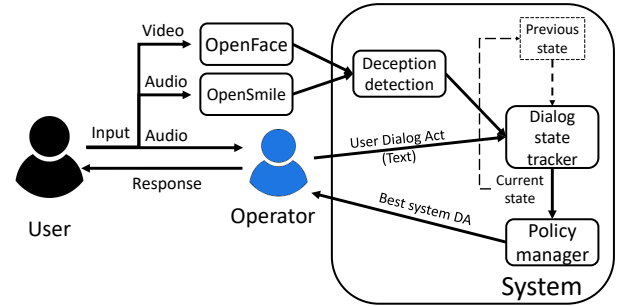


Fig. 1: General structure of the proposed system.

The structure of the proposed system is shown in Figure 1. Verbal and non-verbal information is used for natural language understanding (NLU) and deception detection modules. The NLU module gives the dialog act (DA) of the user utterance. This part is played by the system operator in our experiments; thus, the dialog state tracker always receives the correct dialog act as the output of the NLU module. We have two external modules, OpenFace and OpenSmile, that respectively extracts visual and acoustic (nonverbal) information. Section 4.2 explains in details the feature extraction process. The deception detection module uses these nonverbal information to output the deception label of the user utterance. The deception label is sent to the dialog manager and consists of a dialog state tracker and policy manager. To investigate the impact of deception in our system, we conducted experiments using gold deception labels and estimated deception labels. Dialog histories (previous belief state), the output of the deception detection, and the results of NLU are used to track dialog state by considering the long-term user behavior with a belief update. The output of the dialog state tracker is stochastic variables of the current states over the

space of all possible dialog states, this variable is called belief state. The belief state can be seen as the system's image about dialog situation in the current turn, and the purpose of belief update step is to transit from situation of previous turn to the current situation. Detail about this update step is explained in Section 4.3. The policy manager module uses this information to decide the best action to take next and is trained using a reinforcement learning method. The decided action is sent to the counselor (consultant) who is interacting with the user to decide how to convey the chosen contents (however, the utterance generation part is not the focus of this work).

4. Scenario and Modeling

4.1 Dialog scenario

This work considers a dialog scenario between a system (consultant) and a user (consultee). They discuss the user's living habits, which include: Sleeping, food, working/studying, exercise, social media usage, and leisure activities. The system tries to convince the user to change to more healthy living habits and give up current bad habits. It does by giving the user information about a new habit (system's recommendation), the health benefits of the new habit, and the negative effects of the current habit. This action is denoted as "Framing", defined in existing research [Hiraoka 14]. On the other hand, the user wants to continue the current habit and tries to keep it by using some actions: misstatement (deception,) and reasoning. The objective of the system is to persuade the user to agree with its suggestion truly.

In order to make the conversation more straightforward, only the system can propose recommendations, i.e., the user cannot suggest what habit they should change. However, the user is allowed to give dishonest reasons to manipulate the system into offering an easier recommendation. The user can also pretend to accept the system's offer when they do not actually intend to change the current habit.

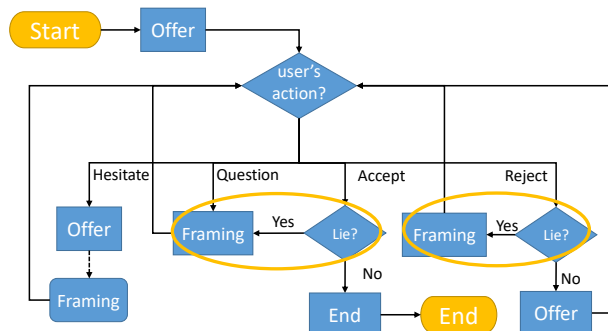


Fig. 2: Flow chart of the conversation.

Figure 2 shows the proposed dialog behavior, which considers user deceptions. Rectangles in this flowchart, indicate system actions. The dialog acts for the system include:

Offer: The system suggests the user should change to a new habit.

Framing: The system provides arguments to persuade the user.

End: The system ends the conversation.

Similar to the work in [Hiraoka 14], we use Framing as one of system's dialog acts. The user can react with different actions:

Accept: The user agrees to change the habit.

Reject: The user gives reasons why they cannot change their habit.

Hesitate: The user says he/she is unsure about whether to accept the offer or not.

Question: The user asks the system about the new habit.

We show an example dialog in this scenario.

System-1: Hello, thank you for coming today. Let's talk about your working habits first. I think that you are spending too much time sitting continuously during working and you should take breaks more often. I suggest that you take a break every hour and during breaks, you should stand up, move away from your desk, and do some light stretches outside of your office. **Offer**

User-1: Why should I move away from my desk during breaks? I feel more comfortable and relaxed sitting in my chair, reading news than going outside. **Question Truth**

System-2: It may be more comfortable for you doing so, but if you still work with computer like that during breaks, your eyes cannot rest. Continuously straining your eyes like that is not good at all and worsen your eye health. **Framing**

User-2: I see that taking breaks is good for my eyes, however, if I go out to take a break, I feel disrupted, and my productivity will be reduced when return. I prefer to take breaks at my desk. **Reject Deception**

System-3: Maybe your productivity will be reduced a little bit; however, in return, your health will be better in the long term. **Framing**

User-3: Okay, I'll try to do as you suggested. **Accept Deception**

System-4: That's great! In case you also want to know, sitting continuously for a long time also puts a lot of stress on your spine and neck, causing many

severe health problems. Doing some light stretches help you regain posture and release stress from your body. **Framing**

User-4: Oh, I see. Sometimes, I also feel my neck hurt a bit while doing work, too. Maybe your suggestion can help. **Hesitate Truth**

System-5: If you find it too distracted for your work to go out every break, how about going out and do stretches just once or twice a day? For example, you can take 3-4 breaks per day, and during one of those, you do some stretches; for the remaining breaks, you stay at your desk. **Offer**

User-5: That sounds good to me. Okay then, I will give it a try. **Accept Truth**

System-6: That's right. If you follow my recommendation, soon you will see the problems related to your neck gone in no time! **End**

In the scenario, the system gives an offer of consulting (System-1), and the user asks a question (User-1). After the framing of the system (System-2), the user makes some deceptions to end the consultation (User-3 and 4). According to the study by [Kjellgren 95], when a consultee is lying, the consultant should tell the consultee about the necessity and benefits of the treatment plan and emphasize the consequences. Applying this theory to the “living habit” scenario, the most logical reaction when the user is telling a lie (uses fake reasons or pretends to agree) is Framing (System-3 and 4). This is in contrast to a conventional negotiation system that does not consider user deception. When the user is lying, such system would probably offer a new recommendation if the user rejects and end the conversation in case the user agrees regardless of whether or not they are telling the truth.

4.2 Deception detection using multimodal approach

There are various clues that can help us to detect lies, including lexical information, acoustic features, gestures, and facial expressions. A multimodal approach that combines these modalities has proven very efficient in detecting deception [Pérez-Rosas 15]. The results from [Pérez-Rosas 15] showed that using linguistic features had the lowest performance out of three modalities. Furthermore, in this research, we perform deception detection at the utterance level, which makes the frequency-based linguistic features used in previous works less effective. Therefore, we utilize the multimodal approach with acoustic and facial features to build the deception detection module.

The classification model we used for deception detection is Multi-Layer Perceptron with one single hidden layer.

The output layer consists of two neurons represent the labels of deception: Lie and Truth. All the layers are fully connected in a feed-forward manner. Our deception detection model used in the proposed system uses hierarchical fusion to combine acoustic and facial features [Tian 16]. As shown in Section 5.1, the observed detection accuracy was 64.71%, which indicates that it is not possible to be sure whether the user is lying or not with this deception detection model.

4.3 Policy management using POMDP

In order to find the best strategy for the dialog system against the user deception, it is necessary to consider errors of the deception detection for user utterances, as the model does not have very high accuracy. The partially observable Markov decision process (POMDP) is widely used to learn the best strategy of dialog systems for such error-containing cases [Yoshino 15]. The user's deception information is included in the dialog state. We separated dialog act states with deception states to model the effect of deceptions explicitly. The detail will be described below (Equation (3)).

POMDP is premised on belief - that there is a probability distribution over all possible states. This belief is updated at every dialog turn using sequence of observation. Optimizing POMDP means producing a policy function that maps the beliefs to the system dialog actions with maximizing future rewards. First, assume that we are at turn number t of the dialog; denote the current user's action (intention) as u^t , $s^t \in S$ as the current state, $a^t \in A$ as the current system's action and o^t as the observation of the current state. Note that, general dialog system only concerns about user intention, thus s^t can be treated as u^t in Equation (1) and (2). The current belief of turn t will be denoted as b^t , which is calculated by:

$$b^t = P(s^t | o^{1:t}, a^{1:t-1}), \quad (1)$$

with $o^{1:t}$ and $a^{1:t-1}$ being the sequence of observations and system actions from the beginning until this turn or the previous turn respectively. We have the user intention of the next turn as s^{t+1} . Under a certain policy function π , we rewrite the current system action as \hat{a}^t , and the equation can be rewritten as

$$b^{t+1} \propto P(o^{t+1} | s^{t+1}) \sum_{s^{t+1} \in S} P(s^{t+1} | s^t, \hat{a}^t) b^t, \quad (2)$$

We can understand (2) as a product of the observation probability $P(o^{t+1} | s^{t+1})$ and summing of state transition probabilities from all possible states $\sum_{s^{t+1} \in S} P(s^{t+1} | s^t, \hat{a}^t) b^t$.

The state transition probabilities revise errors on the observation by using possible transitions from the previous belief b^t .

In addition to user intention, the proposed system also uses deception information for dialog management. For this problem, we use a method in a similar work on user focus by [Yoshino 15] and the dialog state will be $s^t = (u^t, d^t)$. By extending Equation (2) with deception information of the current turn d^t and next turn d^{t+1} , we have the belief update of the proposed method:

$$b_{u,d}^{t+1} \propto P(o_u^{t+1}, o_d^{t+1} | u^{t+1}, d^{t+1}) \sum_u \sum_d P(u^{t+1}, d^{t+1} | u^t, d^t, \hat{a}^t) b_{u,d}^t, \quad (3)$$

with the observation result from SLU and from the deception detection modules being denoted as o_u and o_d respectively. We can see that the belief state is a multivariate categorical distribution of two variables u and d . In our study, transition probability is estimated by from the training data by using maximum likelihood estimation. The probability is also used in the simulator. For training and evaluation with simulation, observation probabilities (o_u^{t+1}, o_d^{t+1}) are randomized from $[0, 1]$. For evaluation with actual human user, the observation probability o_d^{t+1} is taken from confidence score of deception detection and o_u^{t+1} is set to 1 since user's utterance classification into DA is performed by the human operator.

In this research, we use Q-learning [Watkins 92], a popular method to train the optimal policy π^* . Specifically, we estimate the value of the Q-function, which gives an estimation of the discounted future reward of a system action a given a certain belief b . For the proposed system, the training of Q-learning is done by iteratively performing the following update:

$$Q(b^t, a^t) \leftarrow (1 - \epsilon)Q(b^t, a^t) + \epsilon[R(s^t, a^t) + \gamma \max_{a^{t+1}} Q(b^{t+1}, a^{t+1})], \quad (4)$$

where ϵ is the learning rate, γ is the discounted factor and $R(s^t, a^t)$ is the reward the system receives when it performs an action a^t given a user state or a user dialog act s^t . We can apply parametric models such as Deep Q-Network (DQN) for estimating the Q-function. However, since the dimension size of dialog state in our task is small (a state is represented as a tuple of (u, d)), using DQN is not efficient and make the training more difficult. Thus, we did not applied such parametric models in this study. Since the space of belief can be extremely large, it is impossible to calculate all possible values of the Q-function. In this research, we utilize the grid-based value iteration method

proposed by [Bonet 02]. The belief is discretized by the following function:

$$b_s = \begin{cases} \mu & \text{if } s = o. \\ \frac{1-\mu}{|S|-1} & \text{otherwise} \end{cases} \quad (5)$$

where μ represents the rounded probability for every 0.1. In this formula, S is the state space and $|S|$ is the number of elements of S , and b_s is discretization of the belief b of state s .

Using this formula, any belief b can be mapped into a certain fixed point b' in the belief state, as shown in the example below:

$$b = \{(A:0.147, H:0.386, Q:0.235, R:0.232); (L:0.735, T:0.265)\} \\ \rightarrow b' = \{(A:0.2, H:0.4, Q:0.2, R:0.2); (L:0.7, T:0.3)\}$$

The tuple (L, T) is the variable of deception information (*Lie, Truth*) and (A, H, Q, R) is the variable of the user's dialog act (*Accept, Hesitate, Question, Reject*). These probabilities come from the NLU and deception detection, respectively. Note that the task of the NLU module is conducted by a human operator in our experiment; thus, the posterior probability coming from the NLU module is always 1.0 for the decided dialog act of the user.

Table 1: Rewards in each turn.

Dialog state		Rewards		
User DA (s)	d	Offer	Framing	End
Accept	0	-10	-10	+100
	1	-10	+10	-100
Reject	0	+10	+10	-100
	1	-10	+10	-100
Question	0	-10	+10	-100
Hesitate	0	+10	+10	-100

In training the system to act in accordance with the proposed dialog behavior, the reward function plays a critical role. Table 1 shows the reward we defined for the system to receive for each turn, where d represents user deception, 0 denotes no deception (truth), and 1 denotes deception. When the user truly accepts the recommendation, the system receives a very high reward (+100). In contrast, when the user pretends to accept, if the system chooses to perform Framing, it will receive a positive reward (+10), and all other options lead to negative reward. When the user rejects with an honest reason, the system should perform Offer (new recommendation of changing living habit), and it receives a reward of +10. When the user rejects by telling a lie, the system should continue the persuasion (by doing Framing) than making a new recommendation (Offer). Thus, in this case, doing Framing gives

the system a higher reward (+10) than doing Offer (-10). In the situation where the user gives Question to the system, the only suitable reaction that the system should give is Framing, which provide an answer about the benefits or explain about the system's recommendation. Hesitate is used in a situation where the user's intention to reject or accept is not clear. In this case, either Framing or Offer is acceptable. Thus, the reward the system receives in both cases is the same (+10). In general, with this reward function, we encourage the system to follow the behavior shown in Figure 2 with minor reward (+10) and penalty (-10). On the other hand, the system also tries to learn a policy that persuades user to truly accept its offers (reward +100). We observed that when increasing the final reward (from +100 to +200 or +500), the system tends to user Offer more often to increase the successful rate. Translating to an actual conversation, this behavior is equivalent to a consultant who offers easier recommendation to make the patient agree. A drawback of this strategy is that the new recommendation usually provides less health benefit since every time system makes an offer, the new recommendation deviates further from the original recommended habit and goes closer to the user's current habit. The balance of these two kinds of rewards should be determined by considering these trade-off. The reward function shown in Table 1 allows us to train a balance policy model that does not only have reasonable persuasion success chance but also ensures that its recommendations bring enough health benefits.

As explained in the previous section, in our dialog scenario, the system decides when to end the conversation. Thus, there is a limit to the number of times the system can use Framing or Offer; otherwise, the dialog could continue endlessly. In order to impose these constraints on the system, we use three different methods.

Rule-based: The first method uses a rule-based approach. First, we keep track of the number of times the system has performed Framing or Offer since the beginning of the dialog. At each dialog turn, before the system looks at the Q-values to choose the best action, it checks if it has passed the limit or not. If the system has reached the limit of Framing, the Framing action is removed from the list of available dialog actions that the system can choose.

RL limit: The second method uses reinforcement learning to make the system follow the Framing and Offer limits. We use two variables ls_f, ls_o (for Framing and for Offer, respectively) to track the "limit state" of the current dialog turn. These variables show whether the number of times the system uses a certain action has

reached the limit or not. We incorporate these variables into the belief state and train the system using Q-learning. A negative reward (-100) is given if the system violates the constraints.

RL count: The third method is similar to the second one but instead of using "limit state", we incorporate the action count directly into the belief state. The system receives a negative reward if it chooses to perform a certain action (Framing or Offer) when the number of times it has already performed that action is equal to or greater than the limit.

4.4 User simulator

We built a user simulator for both training and testing, following the learning strategy used in [Georgila 11, Yoshino 15]. The simulator generates the next user state, the dialog act, and the deceptive state according to parameter calculated from the corpus by using maximum likelihood estimation. The simulator used in Q-learning was trained from the training data, and the simulator used in the evaluation was trained from the test data. For each dialog, Framing and Offer limits were randomly chosen in the range from 2 to 5. The simulator was created using the same method described in [Yoshino 15]. Specifically, user's dialog acts and deceptions were generated using an **intention model** and a **deception model**, as below:

$$\begin{aligned} \text{intention model} : P(u^{t+1} | d^{t+1}, u^t, d^t, \hat{a}^t) \\ \text{deception model} : P(d^{t+1} | u^t, d^t, \hat{a}^t) \end{aligned} \quad (6)$$

For the evaluation, these probabilities were calculated from test data using maximum likelihood estimation. Our system assumes a small number of user dialog act (DA) (4) and deception (2) and system DA (3). When we consider the belief space of POMDP and the average length of the collected dialogs (about 5 turns per dialog), the size of our data set (203 utterances for training and 178 utterances for evaluation) is small but minimum amount of data necessary to model the intention and deception model. We applied Q-learning that uses simulator with this reason.

5. Experiments

5.1 Experiment of deception detection

To extract facial features, we used the OpenFace toolkit developed by [Baltrušaitis 16]. First, 3D facial landmark points are detected using Conditional Local Fields (CLNF). After that, head position and direction can be directly estimated from these landmark points. Next, facial action intensity and class are detected by Support Vector Machines (SVM) and Support Vector Regression (SVR), based on

features from the landmark points. From the OpenFace toolkit, we were able to extract 14 face Action Unit (AU) regression and 6 AU classification values as well as the head position and, head direction parameters. These values were then normalized and discretized into five different levels of intensity to be used as features for deception detection.

Acoustic features were extracted from audio files using the OpenSMILE tool [Eyben 10]. The acoustic feature template was taken from [Hirschberg 05]. From the pitch and loudness values extracted by OpenSmile, we calculated maximum (*max*), minimum (*min*), mean (*mean*), and standard deviation (*std*). The duration-related features include the percentage of frames with voice, percentage of frames with a lower pitch than the previous frame (falling pitch) and percentage of frames with a higher pitch than the previous frame (rising pitch).

In the experiment for deception detection, we used data from recorded conversations between two participants, one of whom plays the role of a doctor and the other of a patient. In this conversation, the patient tries to get a prescription from the doctor by telling lies about his health condition. The deception label for each utterance by the patient was manually annotated by himself. The total number of utterances was 146. We used 34 of them (17 honest, 17 deceptive) as test data.

Table 2: Deception detection accuracy.

Models	Single facial	Single acoustic	Early fusion	Late fusion	Hierarchical fusion
Accuracy	55.88%	52.95%	61.76%	58.82%	64.71%
Recall	47.06%	35.29%	47.06%	41.18%	52.94%
F1-score	51.61%	42.86%	55.17%	56.25%	60.00%

We trained five different models, single facial (only uses facial features), single acoustic (only uses acoustic features), early fusion, late fusion, and hierarchical fusion [Tian 16]. In early fusion, acoustic and facial features are concatenated and fed into the neural network. For late fusion, features from each modality are fed to two different networks that act as two separate deception classifiers. The results of these networks are then combined by concatenation and fully connect to the output layer. Finally, with hierarchical fusion, acoustic features vector is concatenated with a hidden layer that fully connects to the visual features input vector. All the models were trained using stochastic gradient descent for 100 epochs. The size of the hidden layer was 100, and the learning rate was 0.1. Results of the detection accuracy is shown in Table 2. For our consulting dialog system, a higher recall is preferable since mis-detecting a feint agreement leads to a fail dialog.

On the other hand, taking the user’s honest reason or acceptance as deception makes the user feel irritated but the system can still reach an agreement with the user. In Table 2, we can see that our deception detection performance is not enough to reliably tell whether the user is lying or not. Therefore, our policy manager, which uses deception information for management, needs to decide best system action based on uncertain information. This is one of the main reason why we decided to use POMDP for modeling the dialog management process.

5.2 Experiments of dialog management system

§ 1 Data

We collected dialog data in our scenario for the system training from Wizard-of-Oz (WoZ) setting. This data was also used to model a user simulator for both training and evaluation. We evaluated the proposed dialog management system using dialogs with both the simulated and real users.

The corpus we collected consists of dialogs between two students fluent in English. Both participants role-played as consultant and consultee, and the consultant tried to persuade the consultee to change some actual living habits. All participants were working at the same academic environment at the time of data collection. A total of seven participants took part in the recordings. Four of them played the role of consultant (system), and six played the role of consultee (user).

For training, the recordings were done using the “living habits” dialog scenario in the Wizard-of-Oz (WoZ) setup. Each recording session was carried out by two participants, one playing the role of the system and the other of the user. Each session consisted of six dialogs for each of the living habit topics. The participants who played the role of consultee were given payment as a reward for the outcome of the conversation. If they pretended to agree with the system’s offer, they would receive a lower payment, if they chose to agree truly with the system’s offer they would get a higher payment with the condition that they would need to adopt the new habit for at least one week. The payment was intended to create a situation where the user has to choose between an easy activity (continuing a current habit) with low reward and a difficult activity with a higher reward (changing to a new habit) in order to observe more lies. The recorded training data was about three hours and 20 minutes long and contained 35 dialogs, with an average of 5.8 turns per dialog. DA labels were annotated by one expert, and deception labels were provided by the participant who made the deceptions. In this setup, the two participants sit in front of a laptop in

two separate rooms and cannot see each other. The utterances spoken by a participant in one room were transferred to the other room and output by a speaker. In this experiment, the procedure of the WoZ setting was carried out as below:

- A set of habits and their corresponding Offer and Framing sentences were prepared for the participants who were playing the consultant before the recording sessions start. Besides, they were briefly filled in on the designed dialog behavior that they needed to follow. During the session, the consultants input these sentences into a Web-based TTS application, which converted them into spoken utterances. This procedure is to avoid having the consultee realize that they are talking with a human.
- The participants playing the role of consultee were informed that they would be interacting with a computer. They also received information about the payment reward before the sessions start. In addition, we told the consultees that the consultant would be trying to detect if they were lying.
- For the consultants, only audio data were recorded, while for the consultees, both video and audio data were recorded. The consultees labeled their deceptive utterances by using a stopwatch application. Specifically, when the consultees told a lie, they pushed a button to record the timestamps. After finishing, we matched the timestamps with the video to confirm the deceptive utterances.

With the test corpus, the recordings are direct conversations between participants. The recording setup was similar to the WoZ scenario but without the help of TTS, since the participants were now talking directly with each other. The dialogs collected using this setup are more complex and thus more helpful for assessing the system performance. The test data set was about two hours and 35 minutes in length and contained 36 dialogs, with an average of 4.94 turns per dialog. Table 3 lists the statistics of the training and testing data.

Table 3: Statistics of deception and dialog acts.

Data		Train	Test
Consultant DA	End	14.43%	17.54%
	Framing	43.30%	36.26%
	Offer	42.27%	46.20%
Consultee DA	Hesitate	21.69%	17.64%
	Question	3.61%	9.86%
	Accept	51.81%	19.72%
Reject		22.89%	52.82%
% lie in user utterances		18.07%	19.72%

§ 2 Evaluation with simulation

First, we evaluated the behavior of the trained dialog policy in the virtual conversation with the simulator trained from the test set. The performance was measured by success rate and average offers per successful dialog. The success rate is the percentage of dialogs in which the simulated user truly accepts the system's offer. The goal of our proposed system is to "finds the best treatment and persuades the user to accept it", therefore, the first and foremost important is success rate since it shows us whether the system can convince the user to truly agree with it or not. The second metric, average offer, reflects the "best treatment" point in the system's goal. Every time the system makes an Offer action, the new habit will be more comfortable (closer to the user's current habit) but gives less health benefit, so it is less favorable for the system. In addition, the task of the system is persuading the user truly in less dialog turns, because the exhaustive search will be stressful to the user. Therefore, using fewer offers to persuade the user successfully is better. In this experiment, we used a baseline negotiation system that does not consider user deception. The dialog strategy of the baseline system is similar to Figure 2 but in the branch check for deception, the answer is always "Yes". Both the proposed and baseline systems were trained using Q-learning and grid-based value iteration with a method to limit the number of times they use Offer or Framing by utilizing a rule ("Rule-based" in section 4.3). For the baseline system, we used a similar reward function to Table 1 but the deception information d is set to 1 in all cases. We let the systems interact with the simulated user for 100,000 dialogs. From the results shown in Table 4, it is clear that our proposed system outperformed the baseline in terms of success rate. The bold score indicates that the score has a significant improvement compared with the score of the compared method ($p < 0.05$). On the other hand, average offer score of baseline and proposed system are similar to each other. This is expectable, since our reward function focus on true user agreement and there is no additional reward for using less Offer action to persuade the user. We will explain later why this additional reward is not necessary in the next experiment.

We also examined the performance of the proposed system when using different models to train the dialog policy. These models were distinguished in terms of the methods they use to limit the number of times it can perform a certain action (Framing or Offer), as described in Chapter 3. In detail, **Hybrid** used **Rule-based** for limiting the actions while **POMDP limit** and **POMDP count** respectively used **RL limit** and **RL count** methods.

Table 4: Effect of deception information on dialog management.

Limit	System	Success rate	Avg. offer
2	Baseline	16.76%	1.926
	Proposed	21.08%	1.892
3	Baseline	19.71%	2.558
	Proposed	26.87%	2.562
4	Baseline	21.24%	3.014
	Proposed	29.85%	3.013
5	Baseline	21.79%	3.345
	Proposed	32.66%	3.347

Table 5: Performance of policies using different methods to limit actions usage.

Limit	Model	Success rate	Avg. offer
2	Hybrid	21.08%	1.892
	POMDP limit	26.15%	1.907
	POMDP count	26.34%	1.783
3	Hybrid	26.87%	2.562
	POMDP limit	29.85%	2.521
	POMDP count	30.79%	2.245
4	Hybrid	29.85%	3.058
	POMDP limit	32.30%	3.013
	POMDP count	33.23%	2.546
5	Hybrid	32.66%	3.447
	POMDP limit	33.85%	3.347
	POMDP count	34.60%	2.827

In this experiment, we tested the performance of each model with different limit values ranging from 2 to 5 (Framing and Offer limits are set to be equal). Bold numbers indicate the significant improvement ($p < 0.05$) of **POMDP limit** and **POMDP count** over the **Hybrid** model. As we can see, **POMDP limit** has higher success rate than the **Hybrid** model but shows no significant gain in term of average offer. On the other hand, in the case where Framing and Offer limits were set to 3 ~ 5, the **POMDP count** model outperforms **Hybrid** in both metrics and also beats the **POMDP limit** model. Let us recall that a RL agent does not only learn to solve a task but also learn how to solve it in the lowest number of actions due to discounted reward used to train the Q-value function. Therefore, the **POMDP count**, which takes number of Offer actions used in the past into account, can use the lowest number of Offer to reach true agreement with the user. In conclusion, the POMDP limit makes it possible to learn a good dialog strategy in our scenario.

§ 3 Evaluation of action decision by trained policies

We evaluated the performance of the system’s dialog acts decision with trained policies. The following two metrics were used for the evaluation. DA accuracy refers to the accuracy of the system’s chosen dialog acts against reference actions that were chosen by participants. The appropriate dialog acts in the context are annotated by the participants and another expert; the agreement was 0.78 (Cohen’s kappa). The annotations made by the participants were chosen as the gold standard for this evaluation. Deception handling indicates the accuracy of the dialog acts decision when the user is lying. In this experiment, the proposed system was tested using three types of deception label. The first was “gold label”, which have deception labels that were manually annotated by the participants themselves. The “predicted label” refers to the deception results from the deception detection module described in section Section 4.2. The final one is “chance rate deception label”, we randomly assigned “Lie” labels to 20% of the utterances. This probability was decided on the basis of the statistics shown in Table 3. From the results in Table 6, we can see that our proposed system outperformed the baseline in both metrics when the deception label was either gold label or predicted ($p < 0.05$; bold scores have significant improvements over the baseline). With the chance rate deception labels; the proposed system beat the baseline in term of DA selection accuracy, but the deception handling score is similar to the baseline. More importantly, these results demonstrate that deception information is essential to negotiation dialog management and that the more accurately we predict deception, the better the system will perform.

Table 6: Accuracy of dialog acts selection.

Dialog system	DA accuracy	Deception handling
Baseline	68.54%	35.00%
Proposed system + chance rate deception	69.66%	35.00%
Proposed system + gold-label deception	79.77%	80.00%
Proposed system + predicted deception	75.28%	65.00%

§ 4 Evaluation in interaction with humans

We also tested the system’s performance when interacting with human users. In order to make the system able to interact with humans, natural language understanding and generation parts were performed by a human operator (Figure 1). This person classified the user’s utterance into a user dialog act that was used by the system to choose the best response action. In addition, the operator responded to the user on the basis of the output system dialog act.

A total of nine participants took part in this experiment, one of whom played the role of an operator with the remaining eight as users. The participant that was chosen as the operator was not a medical expert but did have knowledge about healthy living. Before the experiment started, each participant answered a questionnaire about their living habits to decide which topics would be discussed. For each participant, a list of recommendations (new habits) was created on the basis of their answers. Each recommendation in this experiment corresponded to a conversation between the operator and the user. The dialog proceeds as follows:

- (1) The operator recommends (makes an Offer) that the user adopt a new living habit.
- (2) The operator listens to the user's response and translates it into a user dialog act. When the user is replying, his/her face and voice are also recorded for deception detection. The recording procedure is controlled by the operator.
- (3) The system extracts facial and acoustic features from the recorded files and performs deception detection. After that, utilizing the user dialog act given by the operator and the deception results, the system's policy manager module outputs the best response in the form of a system dialog act.
- (4) The operator translates the output of the system dialog act into a full sentence and speaks it to the user. If the output action is **Offer**, the operator recommends another habit (still on the same topic) that is easier to be adopted compared to the previous recommendation. If the system's output is **Framing**, the operator chooses one from a list of arguments that were prepared before the experiment started and spoke it to the user.

After each conversation, the user was required to answer an evaluation questionnaire with the following criteria.

Difficulty of recommendation: How difficult was the recommendation that user agreed to. 1—very easy, 3—not too difficult, 5—impossible. If the user did not agree with any recommendation, this question was skipped.

In this experiment, we evaluated the performance of the proposed system with the **POMDP limit** model against a baseline that also uses the **POMDP limit** model but does not consider user deception. Thus, the dialog for each recommendation was carried out twice. For a fair comparison, the second dialog (using the baseline system) was conducted several days after the first one.

Similar to the previous experiments, evaluation criteria includes negotiation success rate. However, with this experiment, we used the recommendation's difficulty score given by the operator and the user instead of average offers

per successful dialog since this metric more accurately reflects the persuasive performance in real life. Similar to the "average offer" metric, recommendation's difficulty score also reflects the "best treatment" point in the system's goal. For the evaluation with simulation, we assume that all the change in level of difficulty is the same, and thus a recommendation in the second offer will be more difficult (closer to the system's goal and further from the user's goal) than a recommendation given in the third offer. Such assumption is not correct in an actual negotiation (health consultation in our case). On the other hand, the average difficulty score tell us directly how hard it is for the user to change from the current habit to the recommended one. A system that can persuade user to make a bigger change in their habit (which also bring more health benefits) is better. This point is demonstrated by the average difficulty metric. The total number of dialogs in this experiment was 33, and the results are shown in Table 7. Numbers in brackets indicate a 95% confidence interval.

Table 7: Success rate and average difficulty score of successful dialogs.

Dialog system	Baseline	Proposed
Success rate	33%	48%
Average difficulty (user)	2.27 (± 0.36)	2.63 (± 0.57)
Average difficulty (operator)	2.91 (± 0.47)	3.00 (± 0.55)

As we can see from Table 7, our proposed system again has a higher rate of successfully persuading the user. We also observed no significant change in terms of difficulty of recommendation; however, the difficulties of the proposed system were higher than those of the baseline. In general, a negotiation system can increase the chance of success by lowering its bid (in our case, recommend a more comfortable habit to the user). This phenomenon did not occur in our experiment, thus indicating that the proposed system has a higher negotiation performance than the baseline that does not consider deceptions of users.

6. Discussion and Conclusion

In this paper, we tackled the problem of lying in negotiation by utilizing consulting dialogs in the living habit domain. We proposed a dialog manger that can detect user lies and uses this information to choose the best behavior for the system in the interaction. Experimental results showed that the proposed system significantly outperformed a conventional negotiation system that does not consider user deception, beating it by more than 8% for rate of successful persuasion and by more than 11% for system dialog act selection.

In general, the proposed system's performance outperformed baseline systems in our defined metrics; however, there are still several points that are open for improvement. The reward function we used to train the policy management only focuses on user true agreement; thus, the proposed model's main target is to successfully persuade the user. By including the action count into dialog state (*POMDP count* model), we can train a policy manager that learns to persuade using the lowest amount of Offer actions as possible. The experiment results also show that *POMDP count* model was the best in term of average offer per succeeded dialog. However, to measure the performance of the system in term of "find the best treatment for the users (balanced point)", the difficulty of recommended habit needs to be taken into account. In the future, we will focus on incorporating this difficulty level into the system's objective function (the rewards) and train a new system that is able to cooperate with the user.

Acknowledgments

Part of this work was supported by JSPS KAKENHI Grant Number JP17H06101.

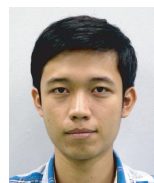
◆ References ◆

- [Baltrušaitis 16] Baltrušaitis, T., Robinson, P., and Morency, L.-P.: Openface: An open source facial behavior analysis toolkit, in *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1–10 (IEEE (2016))
- [Blackwell 73] Blackwell, B.: Patient compliance, *New England Journal of Medicine*, Vol. 289, No. 5, pp. 249–252 (1973)
- [Bonet 02] Bonet, B.: An e-optimal grid-based algorithm for partially observable Markov decision processes, in *Proc. of the 19th Int. Conf. on Machine Learning (ICML-02)* (2002)
- [Chertkoff 71] Chertkoff, J. M. and Baird, S. L.: Applicability of the big lie technique and the last clear chance doctrine to bargaining, *Journal of Personality and Social Psychology*, Vol. 20, No. 3, p. 298 (1971)
- [Craver 97] Craver, C. B.: Negotiation ethics: How to be deceptive without being dishonest/How to be assertive without being offensive, *S. Tex. L. Rev.*, Vol. 38, p. 713 (1997)
- [Davis 66] Davis, M. S.: Variations in patients' compliance with doctors' orders: Analysis of congruence between survey responses and results of empirical investigations, *Academic Medicine*, Vol. 41, No. 11, pp. 1037–48 (1966)
- [Devillers 98] Devillers, L. and Bonneau-Maynard, H.: Evaluation of dialog strategies for a tourist information retrieval system, in *5th International Conference on Spoken Language Processing* (1998)
- [Efsthathiou 14] Efsthathiou, I. and Lemon, O.: Learning non-cooperative dialogue behaviours, in *Proc. of the 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*, pp. 60–68, Philadelphia, PA, U.S.A. (2014)
- [Eyben 10] Eyben, F., Wöllmer, M., and Schuller, B.: Opensmile: The munich versatile and fast open-source audio feature extractor, in *Proc. of the 18th ACM International Conference on Multimedia*, pp. 1459–1462 (ACM (2010))
- [Georgila 11] Georgila, K. and Traum, D.: Reinforcement learning of argumentation dialogue policies in negotiation, in *12th Annual Conference of the International Speech Communication Association* (2011)
- [Gratch 16] Gratch, J., Nazari, Z., and Johnson, E.: The misrepresentation game: How to win at negotiation while seeming like a nice guy, in *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pp. 728–737 (2016)
- [Hiraoka 14] Hiraoka, T., Neubig, G., Sakti, S., Toda, T., and Nakamura, S.: Reinforcement learning of cooperative persuasive dialogue policies using framing, in *COLING*, pp. 1706–1717 (2014)
- [Hirschberg 05] Hirschberg, J., Benus, S., Brenier, J. M., Enos, F., Friedman, S., Gilman, S., Girand, C., Graciarena, M., Kathol, A., Michaelis, L., et al.: Distinguishing deceptive from non-deceptive speech, in *Interspeech*, pp. 1833–1836 (2005)
- [Kjellgren 95] Kjellgren, K. I., Ahlner, J., and Säljö, R.: Taking antihypertensive medication — controlling or co-operating with patients?, *International Journal of Cardiology*, Vol. 47, No. 3, pp. 257–268 (1995)
- [Lewicki 85] Lewicki, R. and Litterer, J. A.: *Negotiation*, Richard D. Irwin, Inc., Homewood, Illinois (1985)
- [O'Connor 97] O'Connor, K. M. and Carnevale, P. J.: A nasty but effective negotiation strategy: Misrepresentation of a common-value issue, *Personality and Social Psychology Bulletin*, Vol. 23, No. 5, pp. 504–515 (1997)
- [Pérez-Rosas 15] Pérez-Rosas, V., Abouelenien, M., Mihalcea, R., and Burzo, M.: Deception detection using real-life trial data, in *Proc. of the 2015 ACM on International Conference on Multimodal Interaction*, pp. 59–66 (2015)
- [Shapiro 91] Shapiro, D. L.: The effects of explanations on negative reactions to deceit, *Administrative Science Quarterly*, pp. 614–630 (1991)
- [Tian 16] Tian, L., Moore, J., and Lai, C.: Recognizing emotions in spoken dialogue with hierarchically fused acoustic and lexical features, in *2016 IEEE Spoken Language Technology Workshop (SLT)*, pp. 565–572 (2016)
- [Torning 09] Torning, K. and Oinas-Kukkonen, H.: Persuasive system design: State of the art and future directions, in *Proc. of the 4th International Conference on Persuasive Technology*, p. 30 (ACM (2009))
- [Traum 08] Traum, D. R.: *Computational models of non-cooperative dialogue*, Citeseer (2008)
- [Vourliotakis 14] Vourliotakis, A., Efsthathiou, I., and Rieser, V.: Detecting deception in non-cooperative dialogue: A smarter adversary cannot be fooled that easily, in *Proc. of the 18th Workshop on the Semantics and Pragmatics of Dialogue (SemDIAL)*, pp. 252–254, Edinburgh, Scotland, U.K. (2014)
- [Watkins 92] Watkins, C. J. and Dayan, P.: Q-learning, *Machine learning*, Vol. 8, No. 3–4, pp. 279–292 (1992)
- [Yoshino 15] Yoshino, K. and Kawahara, T.: Conversational system for information navigation based on POMDP with user focus tracking, *Computer Speech & Language*, Vol. 34, No. 1, pp. 275–291 (2015)

[担当委員：翠 輝久]

Received May 18, 2019.

Author's Profile



The Tung Nguyen

received his Bachelor degree in Computer Science in 2014 from University of Engineering and Technology, Vietnam National University and the M.Eng degree in 2017 from Nara Institute of Science and Technology (NAIST). He is currently a doctoral candidate at Augmented Human Communication Laboratory, NAIST. He is a recipient of the scholarship from the Japanese Ministry of Education, Culture, Sport, Science, and Technology (MEXT). His research area includes multimodal processing, reinforcement learning, and spoken dialog system.



Koichiro Yoshino

received his B.A. degree in 2009 from Keio University, M.S. and Ph.D. degrees in informatics in 2011 and 2014 from Kyoto University, respectively. From 2014 to 2015, he was a research fellow (PD) of Japan Society for Promotion of Science. From 2015, he is an assistant professor of the Graduate School of Information Science, Nara Institute of Science and Technology (NAIST). He is also a researcher of PRESTO, JST, concurrently. He is working on areas of spoken and natural language processing, especially on spoken dialogue systems. Dr. Koichiro Yoshino received the JSAI SIG-research award in 2013, the ANLP outstanding paper award in 2018, and the ACL-NLP4ConvAI workshop best paper award in 2019. He is a member of IEEE, ACL, ISCA, IPSJ, ANLP and RSJ.

received his B.A. degree in 2009 from Keio University, M.S. and Ph.D. degrees in informatics in 2011 and 2014 from Kyoto University, respectively. From 2014 to 2015, he was a research fellow (PD) of Japan Society for Promotion of Science. From 2015, he is an assistant professor of the Graduate School of Information Science, Nara Institute of Science and Technology (NAIST). He is also a researcher of PRESTO, JST, concurrently. He is working on areas of spoken and natural language processing, especially on spoken dialogue systems. Dr. Koichiro Yoshino received the JSAI SIG-research award in 2013, the ANLP outstanding paper award in 2018, and the ACL-NLP4ConvAI workshop best paper award in 2019. He is a member of IEEE, ACL, ISCA, IPSJ, ANLP and RSJ.



Sakriani Sakti

received her B.E. degree in Informatics (cum laude) from Bandung Institute of Technology, Indonesia, in 1999. In 2000, she received DAAD-Siemens Program Asia 21st Century Award to study in Communication Technology, University of Ulm, Germany, and received her MSc degree in 2002. During her thesis work, she worked with the Speech Understanding Department, DaimlerChrysler Research Center, Ulm, Germany. Between 2003-2009, she worked as a researcher at ATR SLC Labs, Japan, and during 2006-2011, she worked as an expert researcher at NICT SLC Groups, Japan. While working with ATR-NICT, Japan, she continued her study (2005-2008) with Dialog Systems Group University of Ulm, Germany, and received her Ph.D. degree in 2008. She actively involved in collaboration activities such as Asian Pacific Telecommunity Project (2003-2007), A-STAR, and U-STAR (2006-2011). In 2009-2011, she served as a visiting professor of the Computer Science Department, University of Indonesia (UI), Indonesia. In 2011-2017, she was an assistant professor at the Augmented Human Communication Laboratory, NAIST, Japan. She also served as a visiting scientific researcher of INRIA Paris-Rocquencourt, France, in 2015-2016, under "JSPS Strategic Young Researcher Overseas Visits Program for Accelerating Brain Circulation". Currently, she is a research associate professor at NAIST, as well as a research scientist at RIKEN, the Center of for Advanced Intelligent Project AIP, Japan. She is a member of JNS, SFN, ASJ, ISCA, IEICE, and IEEE. She is also the officer of ELRA/ISCA Special Interest Group on Under-resourced Languages (SIGUL) and a board Member of Spoken Language Technologies for Under-Resourced Languages (SLTU). Her research interests include statistical pattern recognition, graphical modeling framework, deep learning, multilingual speech recognition & synthesis, spoken language translation, affective dialog system, and cognitive communication.

received her B.E. degree in Informatics (cum laude) from Bandung Institute of Technology, Indonesia, in 1999. In 2000, she received DAAD-Siemens Program Asia 21st Century Award to study in Communication Technology, University of Ulm, Germany, and received her MSc degree in 2002. During her thesis work, she worked with the Speech Understanding Department, DaimlerChrysler Research Center, Ulm, Germany. Between 2003-2009, she worked as a researcher at ATR SLC Labs, Japan, and during 2006-2011, she worked as an expert researcher at NICT SLC Groups, Japan. While working with ATR-NICT, Japan, she continued her study (2005-2008) with Dialog Systems Group University of Ulm, Germany, and received her Ph.D. degree in 2008. She actively involved in collaboration activities such as Asian Pacific Telecommunity Project (2003-2007), A-STAR, and U-STAR (2006-2011). In 2009-2011, she served as a visiting professor of the Computer Science Department, University of Indonesia (UI), Indonesia. In 2011-2017, she was an assistant professor at the Augmented Human Communication Laboratory, NAIST, Japan. She also served as a visiting scientific researcher of INRIA Paris-Rocquencourt, France, in 2015-2016, under "JSPS Strategic Young Researcher Overseas Visits Program for Accelerating Brain Circulation". Currently, she is a research associate professor at NAIST, as well as a research scientist at RIKEN, the Center of for Advanced Intelligent Project AIP, Japan. She is a member of JNS, SFN, ASJ, ISCA, IEICE, and IEEE. She is also the officer of ELRA/ISCA Special Interest Group on Under-resourced Languages (SIGUL) and a board Member of Spoken Language Technologies for Under-Resourced Languages (SLTU). Her research interests include statistical pattern recognition, graphical modeling framework, deep learning, multilingual speech recognition & synthesis, spoken language translation, affective dialog system, and cognitive communication.



Satoshi Nakamura (Member)

is Professor of Graduate School of Science and Technology, Nara Institute of Science and Technology, Japan, Project Leader of Tourism Information Analytics Team of RIKEN, Center for Advanced Intelligence Project AIP, Honorarprofessor of Karlsruhe Institute of Technology, Germany, and ATR Fellow. He received his B.S. from Kyoto Institute of Technology in 1981 and Ph.D. from Kyoto University in 1992. He was Associate Professor of Graduate School of Information Science at Nara Institute of Science and Technology in 1994-2000. He was Director of ATR Spoken Language Communication Research Laboratories in 2000-2008 and Vice president of ATR in 2007-2008. He was Director General of Keihanna Research Laboratories and the Executive Director of Knowledge Creating Communication Research Center, National Institute of Information and Communications Technology, Japan in 2009-2010. He is currently Director of Augmented Human Communication laboratory and a full professor of Graduate School of Information Science at Nara Institute of Science and Technology. He is interested in modeling and systems of speech-to-speech translation and speech recognition. He is one of the leaders of speech-to-speech translation research and has been serving for various speech-to-speech translation research projects in the world including C-STAR, IWSLT, and A-STAR. He received Yamashita Research Award, Kiyasu Award from the Information Processing Society of Japan, Telecom System Award, AAMT Nagao Award, Docomo Mobile Science Award in 2007, ASJ Award for Distinguished Achievements in Acoustics. He received the Commendation for Science and Technology by the Minister of Education, Science and Technology, and the Commendation for Science and Technology by the Minister of Internal Affairs and Communications. He also received the LREC Antonio Zampolli Award 2012. He has been Elected Board Member of International Speech Communication Association, ISCA, since June 2011, IEEE Signal Processing Magazine Editorial Board Member since April 2012, IEEE SPS Speech and Language Technical Committee Member since 2013, and IEEE Fellow since 2016.

is Professor of Graduate School of Science and Technology, Nara Institute of Science and Technology, Japan, Project Leader of Tourism Information Analytics Team of RIKEN, Center for Advanced Intelligence Project AIP, Honorarprofessor of Karlsruhe Institute of Technology, Germany, and ATR Fellow. He received his B.S. from Kyoto Institute of Technology in 1981 and Ph.D. from Kyoto University in 1992. He was Associate Professor of Graduate School of Information Science at Nara Institute of Science and Technology in 1994-2000. He was Director of ATR Spoken Language Communication Research Laboratories in 2000-2008 and Vice president of ATR in 2007-2008. He was Director General of Keihanna Research Laboratories and the Executive Director of Knowledge Creating Communication Research Center, National Institute of Information and Communications Technology, Japan in 2009-2010. He is currently Director of Augmented Human Communication laboratory and a full professor of Graduate School of Information Science at Nara Institute of Science and Technology. He is interested in modeling and systems of speech-to-speech translation and speech recognition. He is one of the leaders of speech-to-speech translation research and has been serving for various speech-to-speech translation research projects in the world including C-STAR, IWSLT, and A-STAR. He received Yamashita Research Award, Kiyasu Award from the Information Processing Society of Japan, Telecom System Award, AAMT Nagao Award, Docomo Mobile Science Award in 2007, ASJ Award for Distinguished Achievements in Acoustics. He received the Commendation for Science and Technology by the Minister of Education, Science and Technology, and the Commendation for Science and Technology by the Minister of Internal Affairs and Communications. He also received the LREC Antonio Zampolli Award 2012. He has been Elected Board Member of International Speech Communication Association, ISCA, since June 2011, IEEE Signal Processing Magazine Editorial Board Member since April 2012, IEEE SPS Speech and Language Technical Committee Member since 2013, and IEEE Fellow since 2016.