NAIST-IS-MT1551205

Master's Thesis

Considering deception information in management of negotiation dialog

Nguyen The Tung

14th September, 2017

Graduate school of Information Science
Nara Institute of Science and Technology

A Master's Thesis
submitted to Graduate School of Information Science,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
MASTER of ENGINEERING

Nguyen The Tung

Thesis Committee:

Professor Satoshi Nakamura	(Supervisor)
Professor Yuji Matsumoto	(Co-supervisor)
Assistant Professor Koichiro Yoshino	(Co-supervisor)
Assistant Professor Sakriani Sakti	(Co-supervisor)

NARA INSTITUTE OF SCIENCE AND TECHNOLOGY

Abstract

Graduate school of Information Science

Master of Engineering

Considering deception information in management of negotiation dialog[1]

by Nguyen The Tung

In the past few years, there have been an increasing number of works in negotiation dialog topic. Almost all of existing negotiation systems assume that its interlocutor (the user) is telling the truth. However, in this negotiation, participants can tell lies to earn a profit, and the partner need to change the negotiation plan for making the negotiation advantageous. As a result, we cannot use existing systems to act as a negotiator and handle this conversation with user. In this research, I proposed a negotiation dialog system that detects user's lies and designed a dialog behavior on how should system react when user is lying. As a typical case of deception negotiation, we built a dialog model of doctor-patient conversation. I showed that we can use partially observable Markov decision process to model this conversation and use reinforcement learning to train the system's action according to the designed behavior. Experimental evaluations proved that the proposed system outperforms the baseline negotiation system that does not use user's deception information for dialog management.

Keywords:

dialog system, negotiaion, POMDP, reinforcement learning, deception detection

[1]Master's Thesis, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT1551205, 14th September, 2017.

Declaration of Authorship

I, Nguyen The Tung, declare that this thesis titled, 'Considering deception information in negotiation dialog' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: ______________________________

Date: ______________________________

"What is negotiation but the accumulation of small lies leading to advantage?"

Felix Dennis

Acknowledgements

First and foremost, I would like to express my deepest thanks toward Prof. Satoshi Nakamura for his great supervision during two years of my master course. His kind advices and guidance, which helped me realize my weaknesses and improved myself to become not only a better student but also a better researcher. I have been and will always be looking up to him as a great teacher.

Professor Yuji Matsumoto; who have been taking part as members of my master thesis's committee, I wish to thank thim for his careful reviews as well as many insightful advices for my research.

I would like to thank Asistant Prof. Koichiro Yoshino for his tireless support that helped me overcame all the obstacles during my research. His knowledge and professionalism really impresed me as well as other members of my lab. I will always be looking at him as a model for me to follow.

I wish to send my gratitude to Asistant Prof. Sakriani Sakti; her advices for my research were extremely valuable and helped me a lot to complete it with success.

I would also like to thank Assoc. Professor Graham Neubig, whose guidance and suggestions have helped me a lot to shaping up the idea of this research. Along with Assistant Professor Koichiro Yoshino, his tutorship at the beginning of my master course has equipped me with lots of useful skills that helped me to complete the research.

Especially, I want to thank Ms. Manami Matsuda, my lab's assistant, who has always given great helps, both in the lab and in my daily life. I also want to send my thanks to all the lab's members, from whom I received great support.

Finally, I want to send my deepest gratitude to my family. During my master course, they have been continuously giving me motivation and inspiration to keep going forward. Without them, this work could not have been completed.

Contents

List of Figures

List of Tables

Abbreviations

Acronym	**W**hat (it) **S**tands **F**or
MDP	**M**arkov **D**ecision **P**rocess
POMDP	**P**artially **O**bservable **M**arkov **D**ecision **P**rocess
MLP	**M**ulti **L**ayer **P**erceptron
OSCE	**O**bjective **S**tructured **C**linical **E**xamination
WoZ	**W**izard **of** **O**z
DA	**D**ialog **A**ct
AU	(facial) **A**ction **U**nit

For/Dedicated to/To my...

Chapter 1

Introduction

1.1 Background

A negotiation is defined as a process between many parties that are trying to find a common ground and reach an agreement among them because each party has their own views, needs and aims that may conflict with the others. A negotiation dialog system is a system that tries to persuade the interlocutors (which can be either a human user or other systems) to agree on the system's preferences by using most appropriate types of arguments [1]. Negotiation has always been a popular topic in the community of spoken dialog researchers. Other than improving the efficiency of dialog management, the focus of these studies have switched from passive role systems (eg. restaurant, tourist information provider) to more active systems that can influence the user's decision using persuasive techniques [2–5].

Since the goals of parties in a negotiation can be conflict with each other, the use of deception in negotiation is expected [6**?**]. Negotiators can use lies such as giving wrong information so that the outcome of the negotiation will be more favorable to them. For existing negotiation systems, when the users use lies as arguments, they gain the advantages and are more likely to win the negotiation because the system believes in the lies Efstathiou and Lemon [7]. Therefore, when system detects that user is lying, the most logical choice is to continue defending its goal by giving more argumentation instead of changing the outcome according to what the user wants.

A typical example of negotiation is the doctor-patient conversation. In this type of dialog, the patient has their own perspective and opinions about their health condition. In other words, the doctor needs to take these opinions into consideration when making a treatment plan and negotiate with users to reach a plan that satisfy both user's demand

and requirements of the treatment. The application of using dialog system to play the role of the doctor and negotiate with the patient in this situation is very promising. One problem happens on this conversation is that the user does not always tell the truth to lead the conversation to their demands. However, almost all of existing negotiation systems assume that the user is speaking the truth. Especially with situation that patients talk with doctors about their health condition, there are about 23% to 38% of patients admit to have told lies to their doctors (Wall Street Journal (2009), WebMD (2004), Zocdoc (2016)). Thus, the assumption used in existing negotiation dialog system is not correct, especially in the doctor-patient conversation. As a result, if we use them to handle this kind of situation, negative consequences will happen as discussed further in the next part of this thesis. Therefore, in this research, I propose a dialog system for negotiation that detects the patient's lies by facial and acoustic clues then use this information for dialog management to decide more appropriate system's action for the user.

1.1.1 Current works in negotiation dialog system

Studies about negotiation dialog started as early as 1990s, when all of the systems were built using rule-based approach [8–10]. It wasn't until recently that statistical approach like Markov Decision Process (MDP) and Partially Observable Markov Decision Process (POMDP) are applied to model the negotiation process [1, 11–13]. All these works assume that user's statements are honest, or in other words, both system and user are behaving in a cooperative manner.

This assumption does not work well since deceptive behaviors are clear indication of non-cooperation in the negotiation process. As pointed out by [14, 15], non-cooperative negotiation dialog "easily leads to miscommunication and an unnecessarily long, complicated, and perhaps failed dialog because of the systems limited abilities to detect, handle, and recover from a non-cooperative dialog flow." As a result, this certain type of negotiation dialog is more difficult than cooperative situation.

In spite of having this disadvantage, there are numbers of practical application that are very promising and worth investigation. Such applications include: educational simulation, commercial bargaining that the customer and the seller cannot reach a mutual objective; negotiation, persuasion and debate [16], zero-sum games scenario [7, 17, 18] and especially the doctor-patient conversation, which is being addressed in this thesis.

1.1.2 Using deception in negotiation process

In fact, deceptive behavior is generally considered to be a part of negotiation. Craver [19] mentioned that he "rarely participated in legal negotiations in which both participants did not use some misstatements to further client interests" and according to Lewicki et al. [6] "lying and deceit are an integral part of effective negotiation".

The benefits that deception brings to negotiator include increasing of power [20] or profit [21–23]. On the other hand, the cost of lying is usually uncertain and can be delayed while the penalty when being caught can be reduced by providing adequate explanations [24]. A survey by Lewicki and Robinson [20] with MBA students in US who are attending negotiation course show that some form of deceptive tactics are considered to be ethically appropriate. Another survey using similar method by Al-Khatib et al. [25] had the same conclusion. These studies shows that deception happens not only in different types of negotiation but also among different countries and cultures.

The doctor-patient conversation is not the only type of negotiation that has the existence of deception. In the doctor-patient conversation, the reasons behind lying are quite different. A study by Iihara et al. [26] showed that the main factor associating with patients' lies is their lack of information and knowledge which leads to distrust in the doctor's treatment plan and poor understanding about the necessity and benefits the treatment plan can bring them. Therefore, to tackle the problem of deception in doctor-patient, we need to carefully designed a system behavior that suits this situation.

1.1.3 Current works about deceptive-countering strategies in negotiation dialog

With such a low number of research about deception in negotiation dialog, as a result, studies about countering deceptive strategy are almost non-existent. Using the same setting as in a previous work by Efstathiou and Lemon [7], Vourliotakis et al. [18] put a learning agent (system) that can tell lies against an adversary that can detect lies with contradiction. Results show that the winning rate of the learning agent in this situation is much less than going against normal adversary that does not detect lies. This work shows that using deception information can help a system to be more success at persuading its interlocutor in the case when interlocutor is using lies. However, application of this work to the doctor-patient conversation is not suitable since the behavior of the adversary when it detects lies is against the principal of the type of negotiation being considered in this research. Another problem is that the penalty of being exposed of using deception in this trading game scenario is not realistic according to survey conducted by Shapiro [24]. Thus, using the work in [18] to situations other than trading game is very difficult.

1.2 Common type of deception

There are several ways to define different types of deceptions. In an early work about deceptive behavior, Bok [27] defined misleading statements as lying and misleading omission as deception. Within the scope of this research, these two terms can be used interchangeable. This is because their purposes are similar in this situation, to make the doctor give treatment plan that is preferred by the patient. Studies by Lewicki and Stark [28], Lewicki and Robinson [20], [29] evaluated common ethically controversial negotiation tactics and grouped them into five factors. DePaulo et al. [30] developed a taxonomy of lies and divide type of lies according to their motivation, content, magnitude and referent (whether the subject of the lies are about liars themselves, other people, objects or events.) Using the analysis from these studies, for the doctor-patient conversation, lies can be divided into two types: information misrepresentation and false promises. While previous work only dealt with the first type of deception [18], in this work, we also consider the use of "false promises" - patient pretends to agree with doctor's treatment plan.

1.2.1 Information mispresentation

This type of deception refers to the situation where user gives false information about himself or hide important facts when giving his arguments. Such situations are illustrated in the example below.

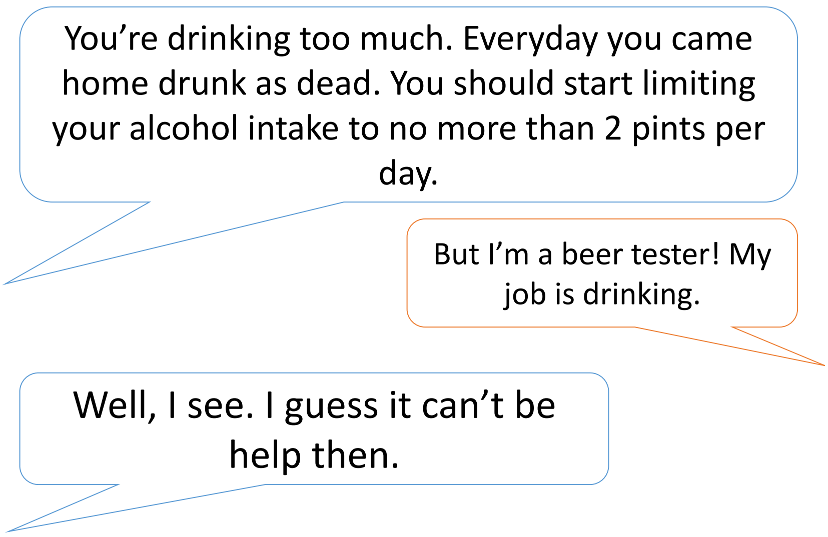

FIGURE 1.1: *Lying by misrepresenting information.*

In this example, the user is lying about his job to convince system that his drinking habit cannot be changed. In this situation, if the system knows that user is lying, a better response will be "I see, but I still recommend you to drink less than that amount. Drinking too much will damage your liver and causes many health problems."

1.2.2 False promises

"False promises" deception means the cases when user pretends to have an agreement with the system. In real life, this situation is commonly found in conversations where a person is getting advices about his personal lifestyles or habits (eg. doctor-patient conversation). Figure 1.2 shows an example of such situation.

In comparison with "information misrepresentation", this kind of deception has a direct impact to the outcome of negotiation. A special characteristic of the doctor-patient conversation is that the treatment plan will be given by the doctor (system) but the patient is the one who decide to follow it or not. As illustrated in Figure 1.2, user pretends that he will stop smoking but he does not really intend to do so. Normal negotiation system believes in the user and end the conversation. As a result, the system fails to persuade user to give up smoking.

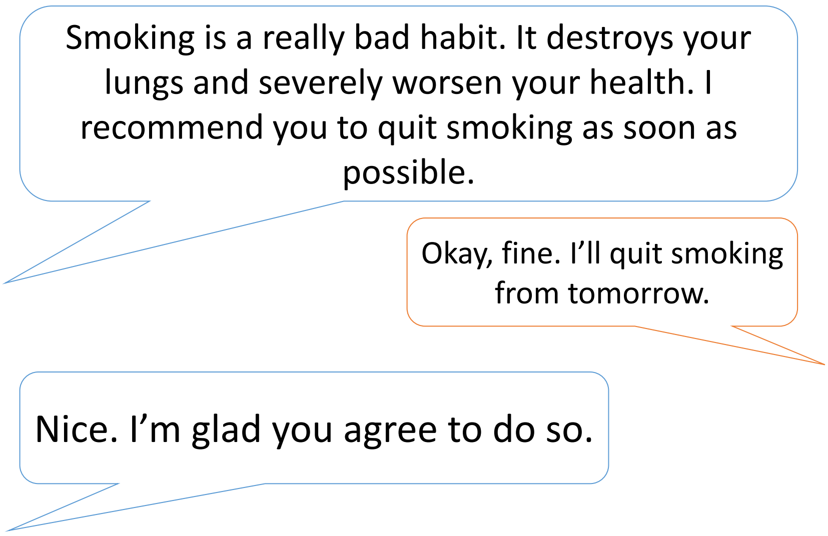

FIGURE 1.2: *Lying by using false promises.*

1.3 Approaches

In this thesis, I proposed a dialog system that can handle the problem of deceptive responses from the user in the doctor-patient conversation. Information about user's deception will be tracked using multi-modal approach. The system uses facial expressions and acoustic features from the user for deception detection. While trying to reach its own goal, the system also takes this deception information into consideration and gives the most suitable respond.

1.3.1 Multimodal deception detection

Deception detection is a topic that is gaining lots of attention from linguistic, speech and multi-modal processing community recently. There are various clues that can help us to detect deceit in other people's utterances, which include lexical modal (the way someone use words), acoustic modal (the way they speaks), their body gestures and facial expressions. As a result, multi-modal approach that can combine those modalities was proved to be very efficient in detecting deception [31, 32]. Polygraph is also an useful clue that is being widely used in lie detector tests. However, the set up is complicated and the cost of using polygraph is very high (estimated cost of a 2-hour polygraph test is about 550$) while the results are still unreliable [33, 34]. Another drawback of this method is that the person who is questioned may feel anxious by the procedure and his utterances will be mistakenly detected as lies.

A research from Hirschberg et al. [35] shows that acoustic features can be used for deception detection. Works by Pérez-Rosas et al. [31, 32] shows that visual clues such as body movement, facial expressions are also useful to decide if someone is lying or not. Therefore, in this research, acoustic and facial features were used to classify the user's statement into lie or truth. The details about deception detection will be discussed further in the next chapter of this thesis.

1.3.2 Negotiation dialog policy that handle user's deception

To build the proposed system, first, I designed a dialog behavior (dialog policy) to respond to user's lies. As pointed out above, there are two situations where the user uses deceptions in a negotiation process: telling lies to gain advantages (to reach an agreement that is more preferred by user) or to pretend to agree with system's recommendation.

As shown in Shapiro [24], liars can mitigate the penalties to their deceptive behaviors by giving excuses. Furthermore, in the doctor-patient conversation, exposing patient's lies without concrete evidence will lead to more distrust from patient. Thus, the system should not focus on trying to expose and confront user's deception. With the first situation, conventional negotiation system will give up convincing the user and switch to a new offer that is more preferred by the user. For the proposed system, since it knows that the user is using fake reasons in his argument, the system should prioritize in providing more arguments to support the current recommendation. For the second situation, let's take a look again at the fast foods example shown in Figure 1.2. When facing this "false promises" from user, a better negotiation system should respond with Framing - dialog action that gives user more information about the recommendation and show them the necessary and benefits from system's recommendation. An example of Framing action in this case is "I'm glad that you agreed with my recommendation; this is because fast foods usually contains lots of cholesterol and made from low-quality ingredients, which can severely worsen your health." By providing the argument that points out bad effects of fast foods, the system expects to draw the attention from the user and persuade them to truly accept the recommendation.

1.3.3 Policy management using statistical method

In the past, dialog management which decides the next most suitable action for the system has switched from manually crafted rules-based to statistical approaches, especially reinforcement learning were also investigated and shown to be very effective. Statistical approaches also make it easier to scale up the system with new data or to increase complexity of the system's behavior. In particular, MDP (Markov Decision Process) and POMDP (Partially Observable Markov Decision Process) are the most successful and now widely used for modeling and training dialog manager. POMDP is an extension of MDP that deals with the situations where the states of decision process is not deterministically observable by the system. In the problem that we focus in this thesis, the result of deception detection cannot be recognized with 100% accuracy. Thus, POMDP was chosen to model the dialog decision process of the system and reinforcement learning was used to train the dialog manager.

1.4 Outline of the thesis

Chapter 2 presents overview of the proposed system; where the general structure of the system, the dialog scenario that system is working on, the dialog behavior (policy) of a conventional negotiation system when working with this dialog scenario and the proposed dialog behavior are described. In comparison with the conventional one, with the proposed dialog behavior, the system reacts to the user based on both user's action and user's deception, thus it can handle user's lies correctly.

In chapter 3, the construction of deception detection module and dialog manager are represented. Deception detection module is a multi layer perceptron classifier that takes facial expressions and acoustic features of the user as input, then combines them and classify user's statements into either lie or truth. The proposed system's dialog decision process is modeled by partially observable Markov decision process (POMDP), which is widely used in control of dialog system. System's dialogue manager receives the user's dialog acts and user's deception and selects an appropriate response by a policy, which is optimized by reinforcement learning. The reward function is defined according to the designed behavior in Chapter 2. In addition, the system will also receive high reward if it can successfully persuade user to accept it's recommendation.

Chapter 4 shows evaluation of the proposed dialog system. Experiment results with simulated user shows that proposed system performs better than conventional one in terms of negotiation efficiency. Experimental evaluations with real human-to-human dialogs demonstrated that the proposed system outperforms the conventional negotiation system that does not track the user's deception in term of action selection.

Chapter 5 concludes this thesis and discusses the contribution of this research as well as future directions.

Chapter 2

Overview of the system

In general, the negotiation covers lots of different topics and many of them require knowledge about specific domain that is too difficult to work with, especially for the medical field. Thus, this research focuses on negotiation between a doctor and a patient about the patient's living habits. In this scenario, the system (doctors) gives recommendation that users (patients) should change their living habit. This chapter describes the general structure of the proposed dialog system as well as the "living habit" scenario that the system is working on.

2.1 System architecture

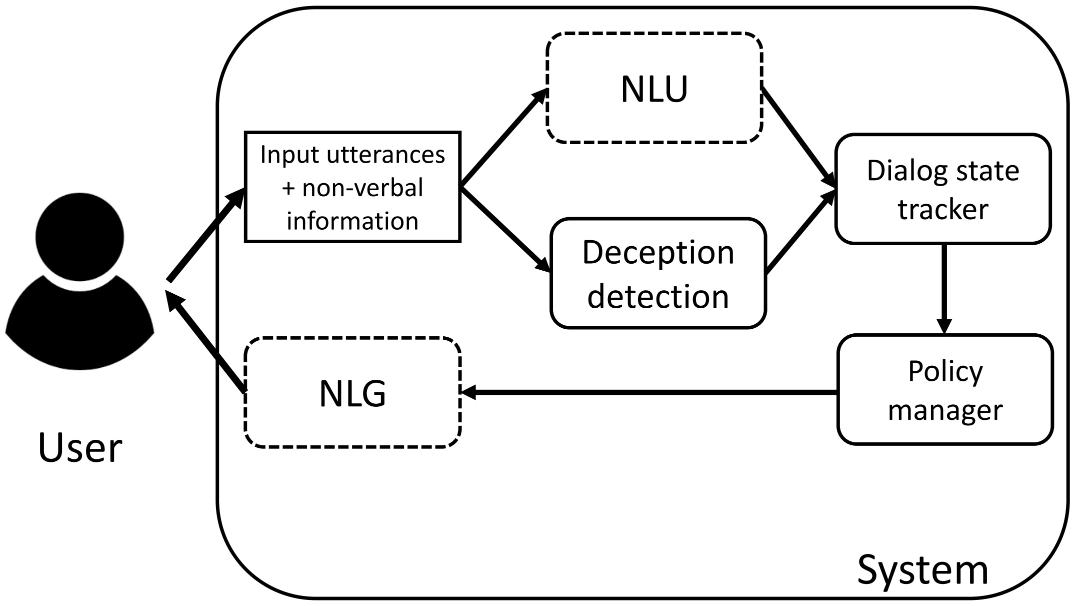

FIGURE 2.1: *General structure of the proposed system.*

Figure 2.1 illustrates the structure of a dialog system that uses the proposed behavior. User's utterance will be converted into text by using the automatic speech recognition (ASR) component. This information is then put into natural language understanding (NLU) unit to recognize the action type of the user's dialog act. In addition to verbal information (text sentence), other non-verbal clues such as voice and facial expressions of the user are also captured. These information will then be fed to a deception detection module, which will determine whether the user is telling the truth or not. The output of deception detection, NLU and dialog histories will be used to track dialog state by considering the long-term user behavior with belief update. The output of the dialog state tracker will be stochastic variable of current state over the space of all possible dialog states. The information of the current state is processed by the policy manager module, which decides the best action to do next. This module is trained by using Q-learning, a popular reinforcement learning method. The decided action will be used by the sentence generation (or natural language generation - NLG) module to output a response to the user.

The paragraph above describes general structure of the proposed dialog system. However, in reality, the system has not yet used any speech recognition or language generation technologies since they are not the target of this research. The main focus of this research is the building of deception detection and dialog manager modules. Therefore, all interactions between the user and the system will be done via an operator.

2.2 Changing living habit scenario

In this scenario, the dialog is a one-to-one conversation between a system (doctor) and a user (patient). They will be discussing about user's living habit. This habit can be one of the following topics: sleeping, food, working/studying, exercise, social media usage and leisure activities. The system's goal is to convince the user that they need to change to a more healthy living habit. The system persuades the user by giving them information about the new habit (system's recommendation), health benefits of the new habit and negative effects of user's current habit. This action is denoted as Framing in this research. In general, Framing is an argument that supports the goal of the system or opposes the interlocutor's goal. On the other hand, the user wants to continue the current habit and gives reasons to show that it is too difficult for him to change. The system behaves in a cooperative manner, which means when user's reason is honest, system will give a new recommendation (another habit) that is easier for the user. For example: the system starts the negotiation with a recommendation that the user should do running 3 times per week. The user, who is currently having no physical activities, rejects this recommendation by saying that they are too busy and do not have enough time to do running so often. Thus, the system makes a new recommendation that the user should do running twice week. This new recommendation requires less time and thus, easier for the user to adapt to.

To make the conversation simpler and easier to be understood, only the system can propose recommendations, the user cannot make suggestion about what habit they should change to. On the other hand, users are allowed to use dishonest reasons to make the system offer an easier recommendation. If the recommendations of the system is not attractive enough for the users, they can choose to pretend to agree with system's recommendation while they actually do not intend to change their current habit.

2.3 Dialog behavior of conventional negotiation system

A conventional negotiation dialog policy, which treats all user's response as truth and reacts to the user only based on user's action is illustrated in Figure 2.2.

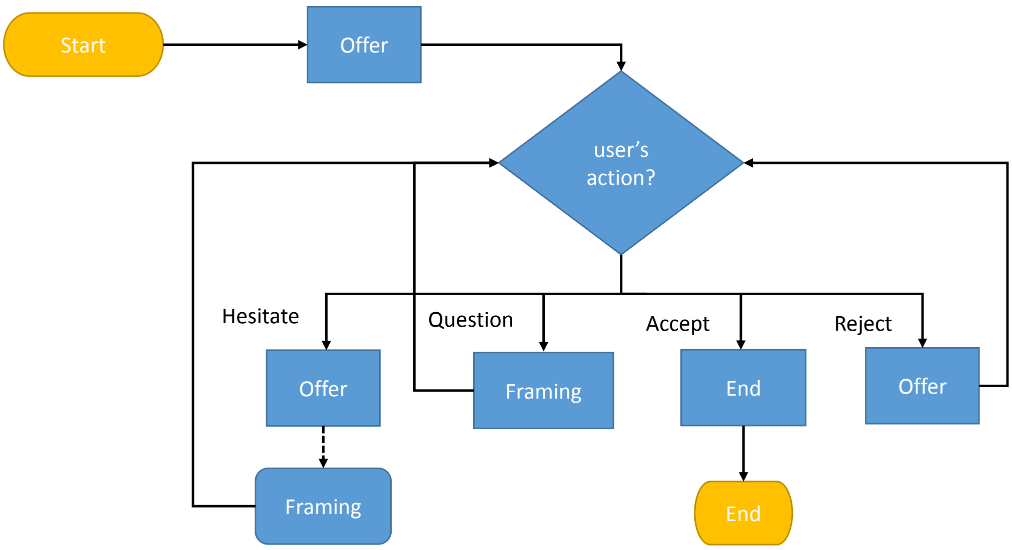

FIGURE 2.2: *Conventional negotiation dialog policy.*

In this flowchart, rectangles indicate system actions. As we can see, the system reacts to user based on user's action only. The set of dialog acts for the system includes:

Offer: the system offers a recommendation to the user, which means the system suggests the user should change to a new habit.

Framing: the system provides information about the new living habit, the necessary and health benefits of the habit and consequences if the user keep the current habit.

End: the system ends the conversation.

For persuasive dialog, framing technique was shown to be a very effective way to direct users to accept the system's goal (Hiraoka et al. [3].) By definition, framing means using emotionally charged words to explain particular alternatives (Levin et al. [36].) In this thesis, Framing actions are any statements that provide supporting arguments to the system's current recommendation, which means either show to benefits of changing the habit or to point out negative effects of the user's current habit or give detailed information about the recommended new habit.

On the other hand, the user can react with different actions as described below:

Accept: the user accepts the recommendation (offer) that was given by the system. User can pretend that they accept the system's offer.

Reject: the user rejects the recommendation given by the system and gives reason why they reject it. In order to make the system to give an easier habit, the user can reject with a deceptive reason.

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

Question: the user asks the system about the benefits of doing the recommended task or asks the system to give a clear explanation about the task.

As shown in Figure 2.2, with conventional negotiation policy, the system starts the conversation by giving the user a recommendation (Offer). When the user replies the system with Hesitate, either offering new recommendation (Offer) or giving argument (Framing) is acceptable choice. Meanwhile, if the user gives a Question, the only suitable response will be Framing, which gives information about the new habit. If the user rejects the system's recommendation by giving a reason why changing habit is too difficult for him, system will change to offer a new habit that is easier for user to change to (since the system believes in user's reason). Finally, if the user accepts the recommendation, no matter if the user is lying or not, the system will take it as truth and end the negotiation.

The dialog behavior described above is used for the baseline system in the evaluation. Similar to the previous work in [1], the dialog decision process of this baseline system is modeled by POMDP and the system learns the behavior using reinforcement learning.

2.4 Dialog behavior to handle user's deception

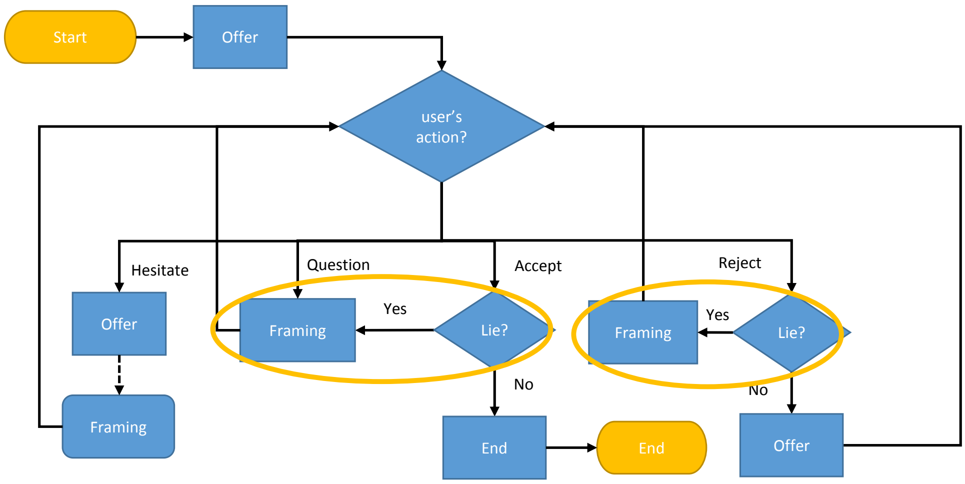

Figure 2.3: *Flow chart of the conversation.*

Figure 2.3 describes the proposed dialog behavior, which considers deceptions of the user. When the user is telling the truth the behavior of the proposed system is the same as conventional one. In this "living habit" scenario, there are two situations in which the user can give a deceptive answer: Reject and Accept, which correspond to "information misrepresentation" and "false promises" respectively. According to the studies by Horne [37] and Kjellgren et al. [38], when the patient is lying, the most suitable reaction from doctor is to explain to the patient about the treatment plan, showing the necessary and benefits of it and consequences if patient refuses to follow the plan. Applying to the "living habit" scenario, the most logical reaction of the system when the user is telling lie is Framing.

2.5 System's goal and functionalities

The ultimate goal of the proposed system is to successfully persuade the user to accept its recommendation and change to the system's recommended habit. When the user gives honest reason why he cannot accept the system's current recommendation (ie. it is too difficult for the user to change to the recommended habit) then the system should offer a new recommendation (another new habit) that is easier for the user.

The proposed system is also expected to be capable of detecting lies in user's statements and provide suitable respond according to the dialog behavior described previously.

In addition to these points, the system also has a limit to the number of times it can give persuasive arguments (Framing). This is because Framing action is related to information about the new living habit and this information is limited. Furthermore, there is also a limit to the number of times the system can do Offer. Such a limit is also natural because after giving Offer for a certain number of times, the system's recommendation will be exactly the same as user's current habit. For example, system's original recommendation is that the user should start running 3 times per week and user wants to continue the current habit of doing no physical activities (run 0 times per week.) With each new Offer, system's recommendation is reduced by 1 time per week. That means with the second time, the recommendation will be running twice per week. Thus, at the fourth time doing Offer, system will recommend the user to run 0 times per week, which is pointless. Therefore, in this situation, the system should stop giving new Offer after recommending user to do running once per week. In this example, the maximum number of time system can do Offer is 3.

Three hypotheses will be investigated in the evaluation section of this thesis.

- Q-learning for POMDP can be utilized to successfully learned the proposed dialog behavior.
- The proposed system has better negotiation performance than the baseline conventional system (evaluation with simulated user).
- The proposed system performs better than conventional systems in term of dialog act selection (evaluation with real human-to-human dialogs.)

Chapter 3

System modeling

This chapter describes the construction of main components of the proposed system. We developed two modules to manage a dialog system with deception detection: multimodal deception detection and POMDP based dialog manager which is extended to use deceptive information. Deception detection module is built using multi-modal approach, which uses non-verbal cues such as facial expressions and acoustic characteristics of users to decide whether they are telling the truth or not. Dialog management is based on partially observable Markov decision process (POMDP), which decides the next best action to do in the negotiation process given observations of the user states that include recognition noises. The dialog management module uses the information of user dialog acts and user's deception to decide what to do next.

The focus of this thesis is to investigate the effect of deceptive information in negotiations, thus, we did not build dialog act classifier and the module is conducted by a human operator. The detail is described in the section of evaluation.

3.1 Deception detection based on multi-modal approach

We built a multi-modal deception detection by using facial and acoustic features. To extract facial features, we used the OpenFace toolkit developed by Baltrušaitis et al. [39]. For face tracking purpose, this toolkit provides tracking of humans' face action units (AUs), which represent facial expressions. By using OpenFace, we were able to extract 14 face AU regression values, 6 AU classification values as well as head position, head direction parameters. All these values are then normalized and discretized into 4 different levels of intensity. For each normalized value, the maximum of it was also taken as a feature for classification. The normalization ensures that the features will represent changes of user's expressions when giving his arguments instead of their intensities. The lists of face action units being used in this thesis are shown in Table 3.1. This AU coding system was developed by [40] to record the movement of muscle groups in human face. The AUs shown in the table are the major muscle groups and they cover the whole face area of a human. Therefore, these AUs were chosen as facial features in this research.

AU list
AU regression values: $AU01$, $AU02$, $AU04$, $AU05$, $AU06$, $AU09$, $AU10$,$AU12$,$AU14$, $AU15$, $AU17$, $AU20$, $AU25$, $AU26$
AU classification values: $AU04$, $AU12$, $AU15$, $AU23$, $AU28$, $AU45$

TABLE 3.1: List of facial action units.

Acoustic features are extracted from audio files using OpenSMILE tool [41]. The acoustic features are selected from the work by Hirschberg et al. [35] as shown in Table 3.2.

Features description
Pitch: $F0\_max$, $F0\_min$, $F0\_mean$, $F0\_std$, $F0\_max - F0\_mean$, $F0\_min - F0\_mean$, $F0\_max/F0\_mean$,$F0\_min/F0\_mean$, $(F0\_max - F0\_mean)/F0\_std$, $(F0\_min - F0\_mean)/F0\_std$
Loudness: Same as above
Duration: %voiced_region, %falling_pitch, %raising_pitch

TABLE 3.2: Acoustic features description

From the pitch and loudness values measured from the audio, we calculated maximum (*max*), minimum (*min*), mean (*mean*) and standard deviation (*std*). We also include duration-related features: percentage of frame with voice, percentage of frame with lower pitch than previous frame (falling pitch), percentage of frame with higher pitch than previous frame (raising pitch).

The classification model for deception detection module was built using Multi Layer Perceptron with hierarchical structure to combine acoustic and facial features. This method was proposed by Tian et al. [42] in a work for emotion detection. It is shown that this combination method outperforms the traditional ways of features combining such as feature-level and decision-level combination. The structure of MLP for each of the combination method is shown in Figure 3.1.

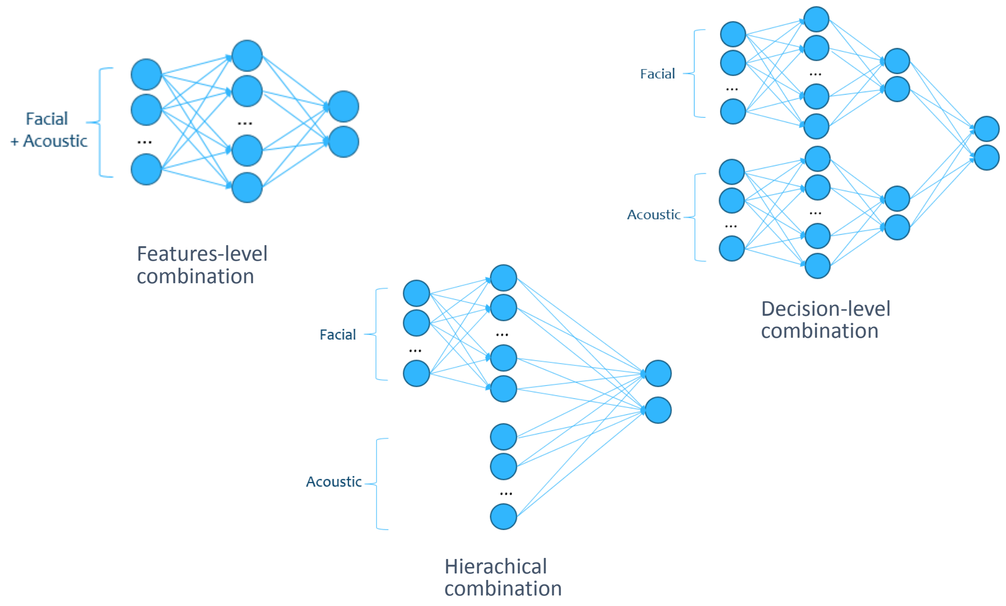

FIGURE 3.1: *Structures for different combination methods.*

To test the effectiveness of the deception detection module, we used data from recorded conversations between 2 participants. The scenario of these conversations is set to be an interview between a doctor and a patient. However, the role of the doctor is played by a student who does not have medical-field knowledge. Thus, the contents of this conversation was prepared using available examples of objective structured clinical examination (OSCE). In these conversations, the "patient" will try to get a prescription from the "doctor" by telling lies about his health condition. The recordings include both video and audio data of the"patient", which were used for extracting facial and acoustic features respectively. Labels of deception are given by the participant by pressing a button when they make any deceptive answer. We separated the corpus into two types of data:

Short: Answers from the user, which are very short (mainly just contains yes or no.) For this data, we have 3 participants, one of them plays the role of "doctor" and the others plays the role of "patient". This data contains 4 dialogs with a total of 87 utterances, 67 of them are honest and 20 are deceptive answers.

Long: Answers from the user, which are longer and usually contains 1 or 2 sentences. For this data, we have 3 participants, one of them plays the role of "doctor" and the others plays the role of "patient". This data contains 2 dialogs with a total of 56 utterances, 42 of them are honest and 14 are deceptive answers.

Results of the experiment are shown in Table 3.3. For each type of data, we trained 3 different classifiers using different combination methods shown in Figure 3.1. For the

Short data, 20 utterances (10 honest, 10 deceptive) are taken out as testing data. For the **Long** data, 14 utterances (7 honest, 7 deceptive) are taken out as testing data We observed the same result as in [42], hierarchical combination method outperformed the others.

Data	Features combine	Decision combine	Hierarchical
Short	55.00%	55.00%	65.00%
Long	57.2 %	64.3%	71.4%

Table 3.3: Deception detection accuracy.

Evaluation results showed that the highest accuracy that deception detection achieved was about 70%. This performance is in line with [31, 32, 35]. However, if the dialog system uses this module to detect user's deception, we cannot make a concrete conclusion whether user is lying or not. How the proposed system overcame this problem will be described in the next section.

3.2 Dialog management using POMDP

As mentioned in the previous section, using result from deception detection module, we cannot reach an exact conclusion about if the user is lying or not. On the other hand, the designed dialog behavior requires the system reacts (make a decision) to the user based on both the user's action and deception (unobservable state.) Therefore, we introduced POMDP as the framework of the proposed dialog manager. POMDP consists of two functions: belief update function and policy function. We extended the architecture of POMDP by introducing the deception state to fit the framework for our system.

POMDP works with belief, a probability distribution over all possible states. Therefore, all of the states are taken into consideration instead of the state with highest possibility as in MDP. The belief is updated at every dialog turn using sequence of observation and transition probabilities. The calculation of these transition probabilities and observation requires data from dialog corpus, which have dialog acts and user's deception information manually annotated beforehand.

The policy function outputs the actual system action as the next behavior of the system given a belief of current turn. In learning of the policy function, the system learn the optimal dialog policy via interactions with an user simulator. The user simulator is an artificial agent that generates user's actions and deception based on n-gram model created from a dialog corpus. The process of learning is described as follow:

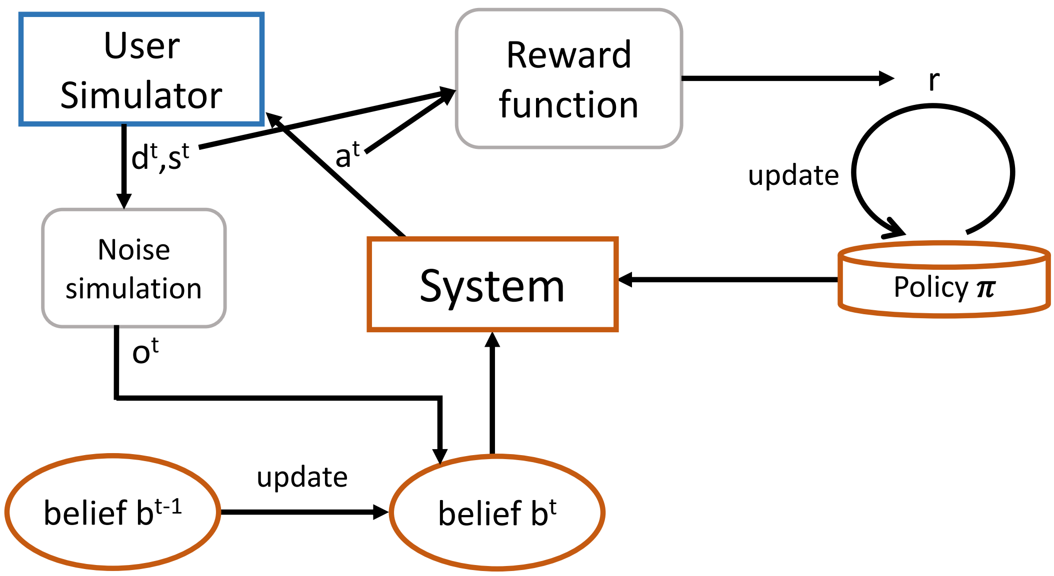

FIGURE 3.2: *Process of system's learning mode.*

1. The user simulator generates an user dialog act and a deception state (either 0 - honest or 1 - deceptive). This dialog act and deception state are denoted as current user action s^t and current user's deception state d^t.

2. The user state (which includes s^t and d^t) is then run through the noise simulation process, which simulates the noise caused by errors from deception detection or user dialog acts classification. The result of this process is observation o^t, which is the system's observation of current user state.

3. Using this observation o^t, along with transition probability and the previous belief state (or belief) b^{t-1}, we calculate the current belief b^t.

4. The system will decide the next best action to do for this dialog turn by using the current policy π, which is denoted as a^t.

5. From the system's action a^t and the current actual user state (s^t,d^t), a reward function gives reward to the system. The reward function is a mapping between a pair of a user state and a system's dialog act $((s^t, d^t), a^t)$ to a real number value. This function reflects the goal of the system, which is to successfully persuade the user. System also receives higher reward if it acts according to the designed dialog policy.

6. Using the reward received, system performs update on the current policy. After this step, we move to next turn of the dialog and repeat the steps above.

The dialog ends when system perform End action. The system will learn the optimal policy by interact with user simulator in many dialogs. Usually, the learning process stops after a certain number of dialogs have passed (ie. 50,000 or 100,000).

When performing discourse, the system can either interact with real human or with another user simulator. The process of discourse is similar to learning but without the use of the reward function since we have already got the optimal policy π^* from the learning phase. Another difference is that when performing discourse with real human, the noise simulation process is also not necessary because the noises are now directly generated by deception detection component. Discourse mode is used for the purpose of system evaluation.

Optimizing POMDP means producing a policy function that maps from the beliefs to system dialog actions with maximum future reward. First, let us assume that we are at turn number t of the dialog; denote the current user state as $s^{t+1} \in S$, $a^t \in A$ as the current system action and the observation of the current state is o^t. Then, the current belief of state s^{t+1} will be denote as b^t, which is calculated by:

$$b^t = P(s^{t+1}|o^{1:t+1}, a^{1:t}) \tag{3.1}$$

In Equation 3.1, $o^{1:t}$ and $a^{1:t-1}$ being the sequence of observations and system actions from the beginning until this turn respectively. Denote the user state of the next turn to be s^{t+1}. Under a certain policy function π, we rewrite the current system action as $\hat{a}^t$, then the equation can be rewritten as:

$$b^{t+1} \propto P(o^{t+1}|s^{t+1}) \sum_{s_i} P(s^{t+1}|s^t, \hat{a}^t) b^t \tag{3.2}$$

Apart from user intention, the proposed dialog system also uses deception information for dialog management. To solve this problem, we used method in a similar work for user focus by Yoshino and Kawahara [43]. By extending Equation (3.2) with deception information of the current turn d^t and next turn d^{t+1}, we have the belief update of the proposed system:

$$\begin{aligned} &b_{s,d}^{t+1} \propto \\ &P(o_s^{t+1}, o_d^{t+1}|s^{t+1}, d^{t+1}) \sum_s \sum_d P(s^{t+1}, d^{t+1}|s^t, d^t, \hat{a}^t) b_{s,d}^t \end{aligned} \tag{3.3}$$

With the observation result come from SLU and Deception Detection modules being denoted as o_s and o_d respectively. The second component of the right-hand side of

Equation 3.3 is developed as:

$$P(s^{t+1},d^{t+1}|s^t,d^t,\hat{a}^t) = \underbrace{P(s^{t+1}|d^{t+1},s^t,d^t,\hat{a}^t)}_{\textbf{intention model}}\underbrace{P(d^{t+1}|s^t,d^t,\hat{a}^t)}_{\textbf{deception model}} \tag{3.4}$$

Thus, the next user's dialog act and deception state will be determined by **intention model** and **deception model**, respectively. The graphical illustration of these transition probabilities (models) are shown in Figure 3.3. These two models are used for the construction of user simulator from the dialog corpus.

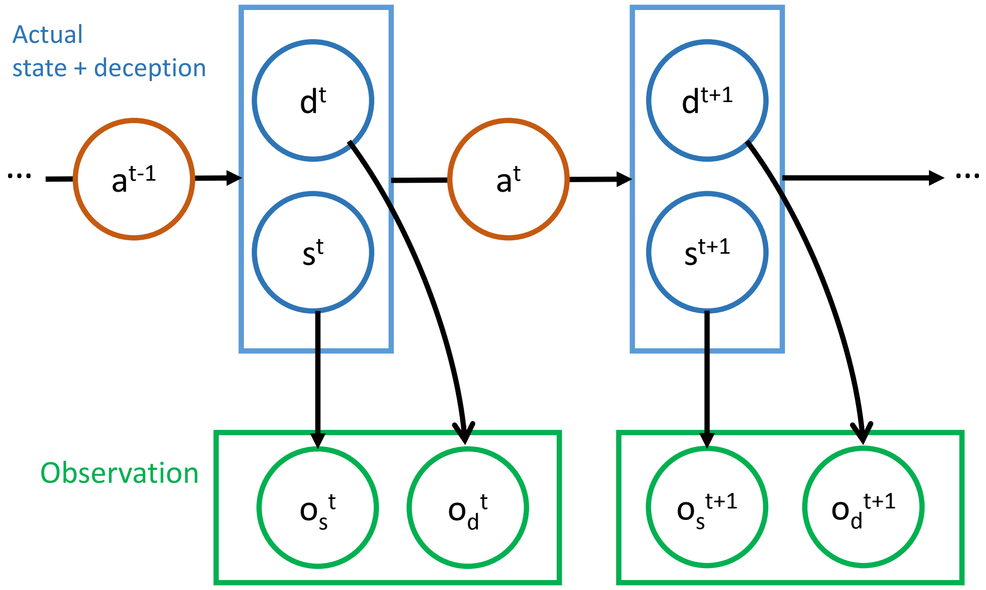

Figure 3.3: *Observation model and state transition model.*

In this research, we use Q-learning, a popular method to train the optimal policy π^* [44, 45]. In Q-learning method, we estimate value of Q-function, which gives estimation of 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:

$$\begin{aligned} Q(b^t, a^t) \leftarrow (1-\epsilon)Q(b^t, a^t) + \epsilon[R(s^t, a^t) \\ +\gamma \max_{a^{t1}} Q(b^{t+1}, a^{t+1})] \end{aligned} \tag{3.5}$$

This algorithm is called $\epsilon - greedy$ [45, 46], where ϵ is the learning rate, γ is discounted factor and $R(s^t, a^t)$ being the reward system received when perform action a^t given user state or user dialog act s^t. For the training of the proposed system, the update function is modified as below:

$$Q(b^t, a^t) \leftarrow (1-\epsilon)Q(b^t, a^t) + \epsilon[R(s^t, d^t, a^t) \\ +\gamma \max_{a^{t_1}} Q(b^{t+1}, a^{t+1})] \tag{3.6}$$

Since the space of belief can be extremely large, it is impossible to calculate all possible value of Q-function. Thus, in this research we utilize Grid-based Value Iteration method proposed by Bonet [47] . The belief is discretized by:

$$b_{s_i} = \begin{cases} \mu & \text{if } s = o \\ \frac{1-\mu}{|S|-1} & \text{otherwise} \end{cases} \tag{3.7}$$

μ represents the rounded probability for every 0.1 that the observation comes from SLU and Deception Detection equal to actual user state and deception information.

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

Table 3.4: Rewards in each turn.

In order to train the system to act according to the proposed dialog behavior, the reward function plays a critical role. Table 3.4 shows the reward the system receives for each turn, which we defined. In this table, d represents user's deception, 0 denotes no deception (truth) and 1 denotes lie. When the user truly accepts the recommendation, the system receives a very high reward (+100). On the other hand, when the user pretends to accept, if the system chooses to do Framing, it will receive positive reward (+10), all other options will lead to negative reward. When the user rejects with an honest reason, the system should do Offer (new recommendation of changing living habit) and the system receives reward of +10. In contrast, if the user rejects by telling a lie, then the system should continue the persuasion (doing Framing) than making new the recommendation (giving Offer.) Thus, in this case, doing Framing gives system higher reward (+10) than doing Offer (-10). In the situation where the user gives Question to the system, the only suitable reaction the system should make in this case is Framing, which will give the answer about the benefits or to explain about the system's recommendation. Hesitate is a situation where user's intention to reject or to accept is not clear. In this case, either Framing or Offer is acceptable. Thus, the reward system receives in both cases is similar (+10).

To make the system stop giving a Framing or Offer action when their limit is reached, a rule-based method is used. To do so, first we keep track of the number of times the system has done Framing or Offer since the beginning of the dialog. At the beginning of each dialog in the training phase, the limit of Framing and Offer is randomly generated. At each dialog turn, before the system looks at the Q-values to choose the best action, the system will check if it has passed the limit or not. If the system reaches the limit of Framing, then Framing action will be removed from the list of available dialog actions that the system can choose. The same rule is applied for Offer limit. Therefore, the final policy that the system learns is a hybrid method resulted from combination of POMDP and hand-crafted rules.

Chapter 4

System evaluation

In this chapter, we assess the following hypotheses.

- Q-learning for POMDP can be utilized to successfully learned the proposed dialog behavior.
- The proposed system has better negotiation performance than the baseline conventional system (evaluation with simulated user).
- The proposed system performs better than conventional systems in term of dialog act selection (evaluation with real human-to-human dialogs.)

In order to perform these evaluations, we set up 2 experiments to test the effectiveness of the system:

- The system is put against another user simulator created from test data.
- The system is tested by using actual dialogs in the test data.

4.1 Dialog corpus

Due to the difficulties of recruiting actual doctors and patients to collect the data, the participants who took part in data collection are students who have good level of English fluency. All participants are working at the same academic environment at the time of data collection.

We setup 2 notebook computers, one is used to record the audio data and the other one is used to record video data using integrated web-camera. The videos are recorded in 1280x720 resolution, 29.97 frames per second and then stored in h.264 compression standard with yuv420p pixel format. Audio data was captured using 2 Crown CM-311A head-worn microphones, which are connected to a Roland QUAD-CAPTURE UA-55 USB audio interface. The audio files were recorded at 44.1 KHz, 16-bit PCM quality. We capture the sound with 2 mono audio signals, one for each of the participants in the recording session. After finishing the sessions, audio and video data was manually synchronized.

The data used for the experiments are recorded using the "living habits" dialog scenario as explained in Chapter 2. Each recording session is carried out by 2 participants who play the role of the system and the user respectively. Each recording session consists of 6 dialogs for each of the living habit topic. The participants who play the role of the user (patient) are given payment as reward for the outcome of the conversation. If they pretend to agree with the system's offer, they will receive lower payment. On the other hand, if they choose to truly agree with the system's offer they will get higher payment with a condition that they will need to change to the new habit for one week. The payment is to create the situation where the user has to choose between an easy activity (continue current habit) with low reward and an activity that is difficult but has higher reward in return (change to new habit). In addition, since the payment is the same for every level of difficulty of recommendation, participants will engage more in the negotiation process.

For training, the recordings are done using wizard-of-oz (WoZ) setup. With this setup, the two participants sit in front of the notebook machine in two separate rooms. The utterances spoken by participant in one room is transferred to the other room and output by a speaker. In this research, the procedure of WoZ setting is carried out as below:

- A set of habits and the corresponding Offer, Framing sentences are prepared for the participant who is playing role of the doctor before the recording starts. In addition, the "doctor" were briefly explained about the designed dialog behavior, which they need to follow. During the session, "doctor" will input these sentences into a web-based TTS application, which converts them into spoken utterances. This procedure is to avoid the "patients" to realize that they are talking with a human.

- The participants playing the role of the patient will be informed that they will be interacting with a computer. They also receive information about the payment reward before the sessions starts. In addition, "patients" are also informed that "doctor" are trying to detect their lies as well.

- For the participant who is playing the role of the doctor, only audio data are recorded. On the other hand, for the participant who is playing the role of the patient; both video and audio data are recorded. The "patients" will label their deceptive utterances by using a stopwatch application. When the "patient" tells a lie, he/she will push a button to keep track of the time stamp. After finishing, the record of time stamps will be matched with the video to determine the deceptive utterances.

By using this setting, I was able to collect dialog transcript (text), audio data (acoustic features), video data (facial feature) and deception labels of the participants. The WoZ setup also helped to make the flow of conversations more controlled and the dialog turns more easily separated. This is important since disruptive speeches are quite common for negotiation dialog. There are total of 7 participants who take part in the recordings. 4 of them played the role of the system and 6 played the role of the user (participants who played the role of the user can switch their role to doctor in later recordings.) In the end, the recorded training data are about 3 hours 20 minutes long with a total of 883 utterances. Labels of DA are annotated by one expert and labels of deception are provided by the participant who made the deception.

With test corpus, recordings are done as direct conversations between participants. The recording setup is similar to WoZ scenario but without the help of TTS since the participants are now talking directly with each other. This type of evaluation data contains 936 utterances and about 2 hours 35 minutes in length.

Data	System DA			User DA				% user's utterances lie
	End	Framing	Offer	Hesitate	Question	Accept	Reject	
WoZ	14.43%	43.30%	42.27%	21.69%	3.61%	51.81%	22.89%	18.07%
Direct Conversation	17.54%	36.26%	46.20%	17.64%	9.86%	19.72%	52.82%	14.08%

Table 4.1: Deception and dialog acts statistic.

Table 4.1 shows statistics of the collected data. With both type of data, we can see that behavior of the system is quite similar. In WoZ data, we can see that the users tend to use Accept more often than Reject while an opposite trend is observed in Direct Conversation data. This shows that people are more eager to negotiate with human than with a computer, which is expectable. Regarding to user's deception, the table shows that people are more likely to tell a lie when interacting with a computer than with a human (*p-value* <0.05). This behavior can be explained that when people are chatting with a computer, they do not have the same feeling of guilty as when they are talking with another human. As a result, with WoZ data, users are more willing to use lies so that the outcome of negotiation is more beneficial to them.

4.2 Evaluation results

4.2.1 Learning curve during the training phase

First, we show the reward received when training a policy using 100,000 dialogs with simulated users created from training corpus. The user simulator is constructed using an n-gram model. In this evaluation, there is no limit to the number of times system can do Framing or Offer, thus the policy is purely trained from the proposed strategy in Figure 2.3 using POMDP.

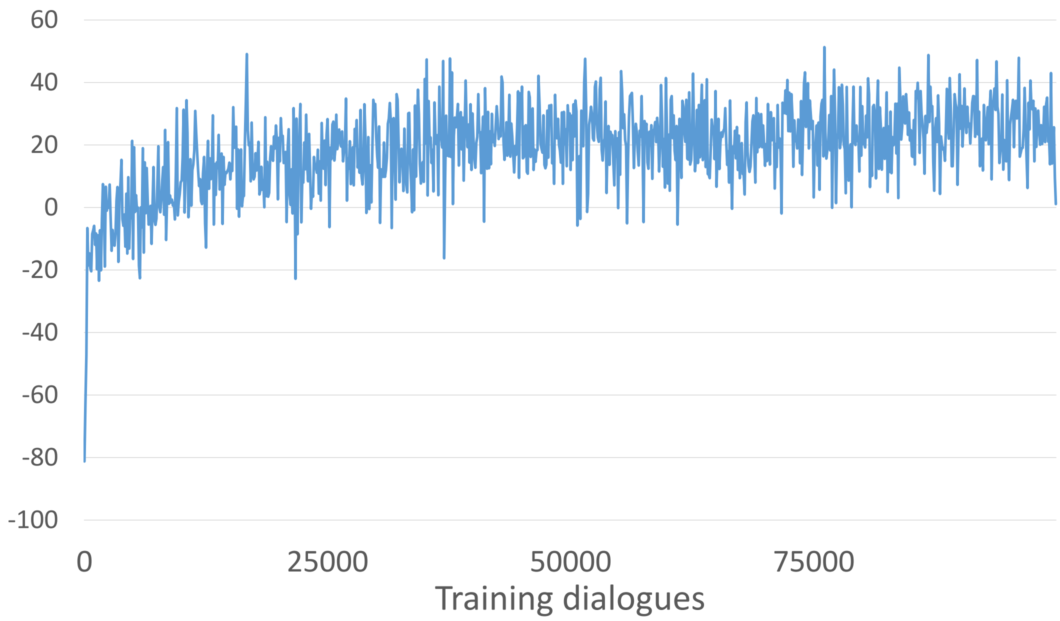

FIGURE 4.1: *Average reward of every 100 dialogs during training process .*

As shown in Figure 4.1 there is a sharp increase in reward received in the first 25,000 dialogs then the gain becomes much slower but the steadily for the remaining of the training process. This proves that the proposed POMDP model can be applied successfully to learn the dialog strategy.

4.2.2 Evaluation with simulated user

In this section, we examine the effectiveness of the learned policy by comparing it with a baseline negotiation policy that does not concern about user's deception. The baseline system uses the dialog behavior described in Figure 2.2. Similar to the proposed system, dialog management of the baseline system are done by using POMDP and it is trained by hybrid of reinforcement learning and rule-based method. This evaluation is conducted with theoretical setup, which means we let the system run in discourse mode with a simulated user created from **Normal** evaluation data for 100,000 dialogs. In this experiment, both system are constrained by Framing and Offer limit of actions. For each dialog, Framing limit and Offer limit are numbers randomly chosen in the range [2,5] .

In this test, performance is evaluated with success rate and average offer per succeeded dialog. Success rate is the percentage of dialogs in which user truly accepts the system's offer and agrees to change their habit. For the second evaluation metric, average offer; as explained in Chapter 2, the dialog system has cooperative manner and if the system makes a **Offer** action, the new habit will be easier but gives less health benefit, thus it is less favorable for the system. Therefore, a system that uses less offers to successfully persuading the user is better at negotiation. From the results shown in Table 4.2, it is clear that our proposed dialog behavior outperforms the baseline in both criteria.

Dialogue policy	Success rate	Avg. offer
Conventional	36.18%	1.866
Proposed policy	**42.60%**	**1.713**

Table 4.2: Success rate and average offer per succeeded dialogs.

4.2.3 Evaluation with real human-to-human dialogs

In this evaluation, we test the proposed system's dialog acts choice accuracy by using manually annotated dialogs in the evaluation data. First, an expert (the author) chooses the best system reaction in each dialog turn by looking at the transcript that contains gold-label of user's dialog action and user's deception. For each dialog turn, we let the system choose the best reaction based on current user's action and user's deception. Accuracy is measured by comparing system's dialog act choice against the choice of expert (annotator).

Dialog system	DA accuracy	Deception handling
Conventional	74.51%	35.00%
Proposed policy + gold-label deception	**77.45%**	**75.00%**
Proposed policy + predicted deception	**77.94%**	**65.00%**

Table 4.3: Dialog acts accuracy.

The accuracy is measured with 2 metrics. DA choice accuracy refers to the accuracy of the system's chosen dialog acts against human's action in general. Deception handling indicates the accuracy of dialog act selection in the turns where the user is lying. From the results of Table 4.3, we can see that in general, the dialog act choice accuracy of the proposed dialog system is higher than the conventional system that does not use deception for dialog management. In term of the deception handling accuracy, the proposed dialog system is also better in both cases: using gold-label deception result and using predicted deception result. Deception prediction was performed on video and audio recording of the corresponding dialog.

Chapter 5

Conclusions

5.1 Contribution of this work

This thesis investigates the lying phenomenon in negotiation dialog by using the conversation between a doctor and a patient. The scenario that is considered in this thesis is set to be a conversation about user's living habits. A dialog system is built to act as the doctor to handle the negotiation. The system tries to persuade the users to change from their current (bad) living habits to a more healthy one. In this situation, the user uses lies so that the outcome of negotiation will be more beneficial to him (ie. the habit is easy for the patient to adapt to). Such lies are either using false information as arguments (information misrepresentation) or pretending that they accept the system's recommendation (false promises). Following analysis of medical experts, a dialog behavior (dialog policy) to react to user's deception is proposed and a dialog system is built based on this behavior. The system uses a deception detection module, which can detect lies in user's responses. However, the goal of the system is not to expose or confront user about their lies. Instead, the deception results from this detection module is used as additional information for the system's dialog manager, which decides the best reaction to the user's statements. In conclusion, the system reacts to user by using both user's action and user's deception information instead of just user's action as in previous works.

In terms of deception detection, the proposed system utilizes multi-modal approach. In this case, the modals being used are facial expressions and acoustic features. The performance of the deception detection module is shown to be comparable to previous works. This performance is achieved by using a small data set for training. If we use bigger deception corpus to train the classifier used for deception detection, increase in performance is expected.

The system's dialog manager was successfully trained using POMDP framework and Q-learning. POMDP is used to model the system's dialog decision process. Deception information is put directly into dialog's belief state and thus system's reaction is based on both user's dialog acts and user's deception. Due to other characteristics of the dialog being considered in this thesis, the final policy is trained by hybrid of reinforcement learning and rule-based methods. Evaluation shows that the system successfully learns the proposed dialog behavior using this method.

From experiments, it is clear that the proposed system outperforms conventional ones both theoretically (testing with simulated user) and practically (turn-by-turn evaluation with real human dialogs.)

5.2 Future directions

The major disadvantage of the system is the lack of speech recognition and language generation usage. Therefore, in future, these technologies will be incorporated to make the system fully automatic and ready to be used in real application. After that, evaluation with real humans will be carried out to measure the performance of this complete system.

In addition, I also would like to improve the accuracy of the deception detection module, which plays a crucial role in the system proposed in this thesis. Such improvement can be achieved by using better classification methods like recurrent neural network (RNN.)

A full doctor-patient communication contains two processes: interview and negotiation. This thesis addressed the problem of patient lying in during the negotiation. In reality, the user can also tell lies during the interview phase as well. This problem is also worth studying in the future.

Appendix A

An Appendix

From what time do you usually start sleeping?

Usually between 1 and 2 a.m. Sometimes even later, depends on the day.

At what time do you usually wake up in the morning?

May be 8 or 9 a.m

How often do you stay up later than 2 a.m?

Quite often I think. Usually 3 or 4 days in a week I stay up that late.

I see. From your answers just now, I recommend you to start sleeping earlier, at 11 p.m.

Well, how does that work out for me? I feel that my current habit is okay.

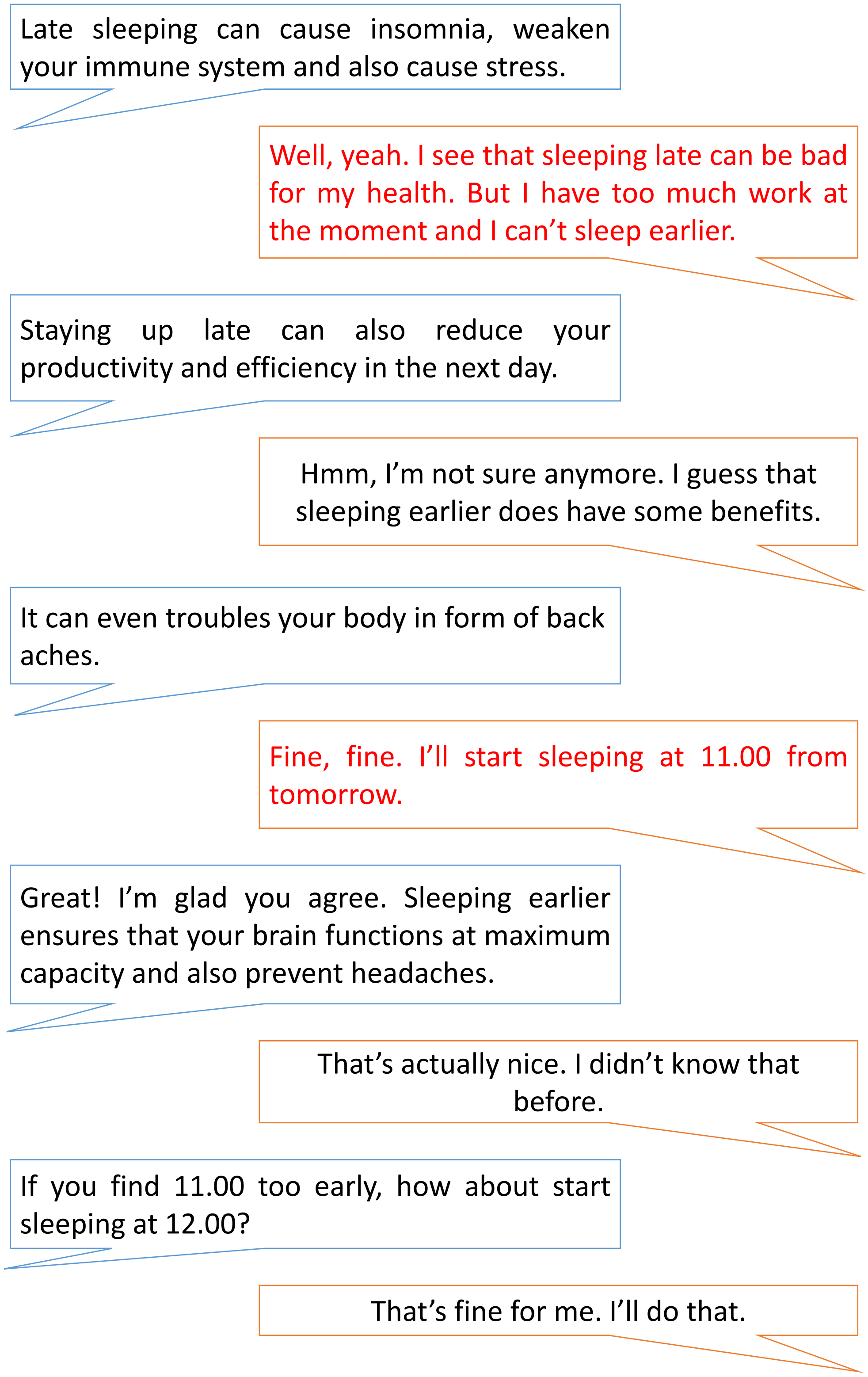

Figure A.1: *Example of a conversation in health consultation .*

Figure A.1 shows an example of health consultation conversation that is investigated in this thesis. The blue call-outs indicate statements from doctor (system) and the orange ones is patient's (user) responses. In this situation, the user has a habit of sleeping very late at night to play computer games or watching movies. The statements from user with red text means it is a lie.

Bibliography

[1] Kallirroi Georgila and David Traum. Reinforcement learning of argumentation dialogue policies in negotiation. In *Twelfth Annual Conference of the International Speech Communication Association*, 2011.

[2] Takuya Hiraoka, Yuki Yamauchi, Graham Neubig, Sakriani Sakti, Tomoki Toda, and Satoshi Nakamura. Dialogue management for leading the conversation in persuasive dialogue systems. In *Automatic Speech Recognition and Understanding (ASRU), 2013 IEEE Workshop on*, pages 114–119. IEEE, 2013.

[3] Takuya Hiraoka, Graham Neubig, Sakriani Sakti, Tomoki Toda, and Satoshi Nakamura. Reinforcement learning of cooperative persuasive dialogue policies using framing. In *COLING*, pages 1706–1717, 2014.

[4] Kristian Torning and Harri Oinas-Kukkonen. Persuasive system design: state of the art and future directions. In *Proceedings of the 4th international conference on persuasive technology*, page 30. ACM, 2009.

[5] Harri Oinas-Kukkonen and Marja Harjumaa. Persuasive systems design: Key issues, process model, and system features. *Communications of the Association for Information Systems*, 24(1):28, 2009.

[6] RJ Lewicki, J Litterer, JW Minton, and DM Saunders. Negotiation, richard d. *Irwin, Homewood, II*, 1985.

[7] Ioannis Efstathiou and Oliver Lemon. Learning non-cooperative dialogue behaviours. In *Proceedings of the 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL)*, pages 60–68, Philadelphia, PA, U.S.A., June 2014. Association for Computational Linguistics. URL `http://www.aclweb.org/anthology/W14-4308`.

[8] Jan Alexandersson, Elisabeth Maier, and Norbert Reithinger. A robust and efficient three-layered dialogue component for a speech-to-speech translation system. In *Proceedings of the seventh conference on European chapter of the Association*

for Computational Linguistics, pages 188–193. Morgan Kaufmann Publishers Inc., 1995.

[9] Leila Amgoud, Simon Parsons, and Nicolas Maudet. Arguments, dialogue, and negotiation. In *Proceedings of the 14th European Conference on Artificial Intelligence*, pages 338–342. IOS Press, 2000.

[10] Leila Amgoud, Nicolas Maudet, and Simon Parsons. Modelling dialogues using argumentation. In *MultiAgent Systems, 2000. Proceedings. Fourth International Conference on*, pages 31–38. IEEE, 2000.

[11] Peter A Heeman. Representing the reinforcement learning state in a negotiation dialogue. In *Automatic Speech Recognition & Understanding, 2009. ASRU 2009. IEEE Workshop on*, pages 450–455. IEEE, 2009.

[12] Kallirroi Georgila. Reinforcement learning of two-issue negotiation dialogue policies. In *SIGDIAL Conference*, pages 112–116, 2013.

[13] Praveen Paruchuri, Nilanjan Chakraborty, Roie Zivan, Katia Sycara, Miroslav Dudik, and Geoff Gordon. Pomdp based negotiation modeling. In *Proc. of the IJCAI Workshop on Modeling Intercultural Collaboration and Negotiation (MICON)*, 2009.

[14] Marion Klein, N Ole Bernsen, Sarah Davies, Laila Dybkjær, Juanma Garrido, Henrik Kasch, Andreas Mengel, Vito Pirrelli, Massimo Poesio, Silvia Quazza, et al. Supported coding schemes. *MATE Deliverable D*, 1:1, 1998.

[15] Melita Hajdinjak and France Mihelič. Information-providing dialogue management. In *International Conference on Text, Speech and Dialogue*, pages 595–602. Springer, 2004.

[16] David R Traum. Computational models of non-cooperative dialogue. 2008.

[17] Ioannis Efstathiou and Oliver Lemon. Learning to manage risks in non-cooperative dialogues. In *Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue (SemDIAL)*, pages 173–175, Edinburgh, Scotland, U.K., September 2014.

[18] Aimilios Vourliotakis, Ioannis Efstathiou, and Verena Rieser. Detecting deception in non-cooperative dialogue: A smarter adversary cannot be fooled that easily. In *Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue (SemDIAL)*, pages 252–254, Edinburgh, Scotland, U.K., September 2014.

[19] Charles B Craver. Negotiation ethics: How to be deceptive without being dishonest/how to be assertive without being offensive. *S. Tex. L. Rev.*, 38:713, 1997.

[20] Roy J. Lewicki and Robert J. Robinson. Ethical and unethical bargaining tactics: An empirical study. *Journal of Business Ethics*, 17(6):665–682, 1998. ISSN 1573-0697. doi: 10.1023/A:1005719122519. URL `http://dx.doi.org/10.1023/A:1005719122519`.

[21] Jerome M Chertkoff and Suzanne L Baird. Applicability of the big lie technique and the last clear chance doctrine to bargaining. *Journal of Personality and Social Psychology*, 20(3):298, 1971.

[22] Kathleen M O'Connor and Peter J Carnevale. A nasty but effective negotiation strategy: Misrepresentation of a common-value issue. *Personality and Social Psychology Bulletin*, 23(5):504–515, 1997.

[23] Maurice E Schweitzer and Rachel Croson. Curtailing deception: The impact of direct questions on lies and omissions. *International Journal of Conflict Management*, 10(3):225–248, 1999.

[24] Debra L Shapiro. The effects of explanations on negative reactions to deceit. *Administrative science quarterly*, pages 614–630, 1991.

[25] Jamal A Al-Khatib, Avinash Malshe, and Mazen AbdulKader. Perception of unethical negotiation tactics: A comparative study of us and saudi managers. *International Business Review*, 17(1):78–102, 2008.

[26] N Iihara, T Tsukamoto, S Morita, C Miyoshi, K Takabatake, and Yuji Kurosaki. Beliefs of chronically ill japanese patients that lead to intentional non-adherence to medication. *Journal of clinical pharmacy and therapeutics*, 29(5):417–424, 2004.

[27] Sisela Bok. Lies to the sick and dying. *Lying: Moral choice in public and private life*, pages 232–255, 1978.

[28] Roy J Lewicki and Neil Stark. What is ethically appropriate in negotiations: An empirical examination of bargaining tactics. *Social Justice Research*, 9(1):69–95, 1996.

[29] Robert J Robinson, Roy J Lewicki, and Eileen M Donahue. Extending and testing a five factor model of ethical and unethical bargaining tactics: Introducing the sins scale. *Journal of Organizational Behavior*, pages 649–664, 2000.

[30] Bella M DePaulo, Deborah A Kashy, Susan E Kirkendol, Melissa M Wyer, and Jennifer A Epstein. Lying in everyday life. *Journal of personality and social psychology*, 70(5):979, 1996.

[31] Verónica Pérez-Rosas, Mohamed Abouelenien, Rada Mihalcea, and Mihai Burzo. Deception detection using real-life trial data. In *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, pages 59–66. ACM, 2015.

[32] Verónica Pérez-Rosas, Mohamed Abouelenien, Rada Mihalcea, Yao Xiao, CJ Linton, and Mihai Burzo. Verbal and nonverbal clues for real-life deception detection. In *EMNLP*, pages 2336–2346, 2015.

[33] Leonard Saxe and Gershon Ben-Shakhar. Admissibility of polygraph tests: The application of scientific standards post-daubert. *Psychology, Public Policy, and Law*, 5(1):203, 1999.

[34] David Thoreson Lykken. *A tremor in the blood: Uses and abuses of the lie detector.* Plenum Press, 1998.

[35] Julia Hirschberg, Stefan Benus, Jason M Brenier, Frank Enos, Sarah Friedman, Sarah Gilman, Cynthia Girand, Martin Graciarena, Andreas Kathol, Laura Michaelis, et al. Distinguishing deceptive from non-deceptive speech. In *Interspeech*, pages 1833–1836, 2005.

[36] Irwin P Levin, Sandra L Schneider, and Gary J Gaeth. All frames are not created equal: A typology and critical analysis of framing effects. *Organizational behavior and human decision processes*, 76(2):149–188, 1998.

[37] Rob Horne. Compliance, adherence, and concordance: implications for asthma treatment. *Chest Journal*, 130(1_suppl):65S–72S, 2006.

[38] Karin I Kjellgren, Johan Ahlner, and Roger Säljö. Taking antihypertensive medicationcontrolling or co-operating with patients? *International journal of cardiology*, 47(3):257–268, 1995.

[39] Tadas Baltrušaitis, Peter Robinson, and Louis-Philippe Morency. Openface: an open source facial behavior analysis toolkit. In *Applications of Computer Vision (WACV), 2016 IEEE Winter Conference on*, pages 1–10. IEEE, 2016.

[40] Paul Ekman, Wallace V Friesen, and J Hager. The facial action coding system (facs): A technique for the measurement of facial action. palo alto. *CA: Consulting Psychologists Press, Inc. Ekman, P. Levenson. RW, & Friesen WV (1983). Autonomic nervous system activity distinguishes among emotions. Science*, 221:1208–12, 1978.

[41] Florian Eyben, Martin Wöllmer, and Björn Schuller. Opensmile: the munich versatile and fast open-source audio feature extractor. In *Proceedings of the 18th ACM international conference on Multimedia*, pages 1459–1462. ACM, 2010.

[42] Leimin Tian, Johanna Moore, and Catherine Lai. Recognizing emotions in spoken dialogue with hierarchically fused acoustic and lexical features. In *Spoken Language Technology Workshop (SLT), 2016 IEEE*, pages 565–572. IEEE, 2016.

[43] Koichiro Yoshino and Tatsuya Kawahara. Conversational system for information navigation based on pomdp with user focus tracking. *Computer Speech & Language*, 34(1):275–291, 2015.

[44] Christopher John Cornish Hellaby Watkins. *Learning from delayed rewards.* PhD thesis, King's College, Cambridge, 1989.

[45] Peter. Dayan Watkins, Christopher JCH. Q-learning. *Technical note, Machine Learning*, 8:279–292, 1992.

[46] Richard S Sutton and Andrew G Barto. *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge, 1998.

[47] Blai Bonet. 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.