DIALOGUE MANAGEMENT FOR LEADING THE CONVERSATION IN PERSUASIVE DIALOGUE SYSTEMS

Takuya Hiraoka, Yuki Yamauchi, Graham Neubig, Sakriani Sakti, Tomoki Toda, Satoshi Nakamura

Graduate School of Information Science, Nara Institute of Science and Technology, Japan
{takuya-h, yuki-ya, neubig, ssakti, tomoki, s-nakamura}@is.naist.jp

ABSTRACT

In this research, we propose a probabilistic dialogue modeling method for persuasive dialogue systems that interact with the user based on a specific goal, and lead the user to take actions that the system intends from candidate actions satisfying the user's needs. As a baseline system, we develop a dialogue model assuming the user makes decisions based on preference. Then we improve the model by introducing methods to guide the user from topic to topic. We evaluate the system knowledge and dialogue manager in a task that tests the system's persuasive power, and find that the proposed method is effective in this respect.

Index Terms— Spoken dialogue system, Persuasive dialogue, Dialogue management, Leading conversations, Estimating user's preference

1. INTRODUCTION

Traditionally, dialogue systems have been developed to help users perform a specific task [1, 2, 3], help users with uncertain needs discover the information they are interested in [4, 5, 6], or entertain users through chat [7, 8, 9]. These systems are all similar in that the dialogue system is a tool that is used to achieve the *user goal* of finishing a task, learning new information, or simply being entertained (left ellipse in Figure 1).

On the other hand, there has also been a focus in recent years on persuasive technology and computational deception, where the computer is not simply a passive actor, but actively tries to change the thoughts or habits of the users [10, 11] (right ellipse in Figure 1). In other words, in persuasive technology there is a *system goal*, and the system attempts to persuade the user in such a way that this goal may be achieved. Persuasive technology has been used to identify factors of user decisions [12], or for selling items, interactive advertisement, and helping to improve bad habits [10]. There is also some related research in dialogue on optimizing policies of dialogue systems for argumentation [13] and analyzing important factors for persuasion, such as message style or number of participants [14].

Acknowledgment: Part of this work was supported by JSPS KAKENHI Grant Number 24240032, by the collaborative research with I2R and by the Commissioned Research of National Institute of Information and Communications Technology (NICT), Japan.

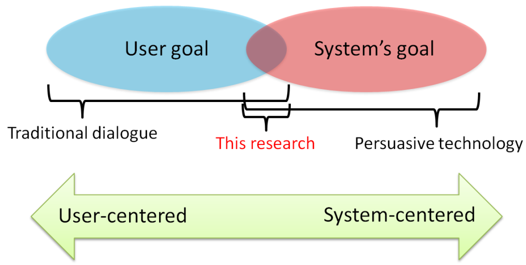

Fig. 1. Categorization of dialogue systems by goal to be achieved.

Our interest lies in an application at the crossroads of persuasive technology and traditional dialogue systems: persuasive dialogue systems that attempt to satisfy both the *user goal* and *system goal* simultaneously. While this sort of general framework can be used for any number of tasks, in this paper we will use as an example a system that helps incoming graduate students decide which laboratory they would like to join. When the system considers only the user's goal, this is an example of decision support dialogue [6] in which the user's preference is not clear to the system and the system needs to estimate it through the dialogue. However, we can also think of a situation where the system may have an auxiliary goal, such as maintaining a relatively even balance of the number of students between the laboratories in the department. In this case, the system would like to lead the user to a decision that both satisfies the user's goal (of finding an appropriate laboratory), and the system goal (of leading the user to a laboratory that is under capacity).

In this research, we propose the first system for persuasive dialogue that leads the user to take actions that the system intends while still considering the user's goal and satisfaction (the overlapping zone in the two ellipses in Figure 1). In particular, we propose methods for knowledge-base construction and a dialogue management that help us achieve this sort of persuasive dialogue. We examine three methods for constructing knowledge bases, with a focus on acquiring knowledge that enables the system to guide users towards topics related to the system goal. The proposed persuasive dialogue manager is based on the Bayesian network (BN) framework

Table 1. An example of persuasive dialogue. S is a system utterance, U is a user utterance, and the system's target is laboratory A.

Utterance
S1: *"What research fields are you interested in?"*
U1: *"Dialogue and communication support."*
S2: *"Laboratory A is working on dialogue."*
U2: *"But not communication support?"*
S3-1: *"You are interested in communication support."*
S3-2: *"Speech translation helps communication support."*
S3-3: *"Are you interested in speech translation?"*
U3: *"Yes."*
S4-1: *"Laboratory A works on speech translation."*
S4-2: *"How about joining laboratory A?"*
U4: *"Sure."*

[15, 16], and further implements new functionality including actions to guide the user to the target topic.

We experimentally evaluate performance of the proposed persuasive dialogue manager on two tasks, one controlled task with a small number of topics and where the user has almost no knowledge about the domain, and a larger scale task where the system handles a large number of topics and the user has some prior knowledge. Experimental results indicate that guiding the user to the target topic contributes to the achievement of the system goal. If the user has little knowledge about the domain, guiding the user to the target topic also contributes to an improvement in user satisfaction.

2. AN OUTLINE OF PERSUASIVE DIALOGUE

As mentioned in the introduction, as an example of a persuasive dialog, we suppose a situation where a system is trying to recommend a laboratory for incoming students. The dialogue system has an intended target laboratory, and attempts to persuade the user (i.e., the student) to join the target laboratory. Table 1 shows an example of a persuasive dialogue where the target laboratory is set to laboratory A. In this research, we assume that 1) The goal of the system is pre-determined and invariant throughout the dialogue, 2) the user has interest in topics, at least one of which is covered by the target laboratory.

In order to formalize a system that can perform a dialogue similar to that in Table 1, we describe a persuasive dialogue system that uses four types of dialogue act:

Inform the user of information about a topic or the relationship between topics.
Question the user about his/her preferred topic.
Confirm the content of recognized user utterance.
Solicit the user to choose the target alternative.

In Table 1, utterances S2, S3-2 and S4-1 are system actions based on *Inform* and utterance S4-2 is a system action based on *Solicit*. And, for promoting smooth dialogue, the system

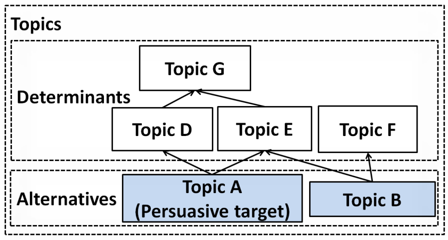

Fig. 2. Example of knowledge of the system

also handles the *Question* (S1, S3-3) and *Confirm* (S3-1) actions.

In order to persuade the user to select a candidate that the system intends from the user's decision candidates, it is natural that the system should know about the topics that may be covered in the conversation. Furthermore, to allow the system to lead the user from topic to topic it is also desirable for the system to have knowledge about the relationships between topics. To achieve high user satisfaction, it is also necessary that the system understands whether the user is interested in each topic, and manages the dialogue based on this user interest. We will introduce techniques for creating the knowledge-base and dialogue manager in the following two sections.

3. BUILDING A KNOWLEDGE-BASE FOR PERSUASIVE DIALOGUE

Using a knowledge base, we provide the system with knowledge about topics and the relationships between them as shown in Figure 2. Here, *topics* can be further divided into *alternatives* representing topics among which the user is trying to decide (i.e. laboratories), and *determinants* representing topics that can affect the user's selection of alternatives (i.e. research topics). In the context of persuasive dialogue, we must guide the user from topic to topic, and thus we also make a particular effort towards defining the relationship between topics.

3.1. Manually created knowledge base

Traditional knowledge bases such as WordNet [17] define hypernym/hyponym relations between topics. As shown in Fig 3, we define 3 types of relationship over hypernym/hyponym graphs: 1) parent-child, 2) grandparent-grandchild, and 3) common ascendant. We further define a response template for each relation. For example, the response template for 1) is "y is one of the research fields of x."

3.2. Building a knowledge base through web search

While the method described in the previous section has the advantage of being able to use existing knowledge bases, it is also limited by the coverage of manually created resources, and limited to the response templates that can be defined using hypernym-hyponym relations. In order to solve this problem,

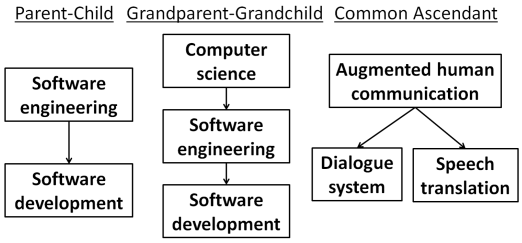

Fig. 3. Example of the knowledge based on hypernym/hyponym relations.

we propose a method to expand the knowledge base using Web search in the following manner:

1. Create a response template rt by hand. Response templates are chosen that have high potential for guiding the user from the current topic x to the next topic y. (i.e. "x is used in y")
2. Create a search query based on rt. (i.e. "x used in y")
3. For each x/y pair use a search engine to calculate statistics including the total number of pages covered N and the number of hits $C(x, y, rt)$, $C(x)$, and $C(y)$.
4. Score each x/y pair according to the number of hits $C(x, y, rt)$ or the mutual information

$$I(x, y, rt) = \frac{P(x, y, rt)}{P(x) * P(y)} = \frac{N * C(x, y, rt)}{C(x) * C(y)}. \quad (1)$$

The response template for a particular x/y is selected according to $\underset{rt}{\operatorname{argmax}}\, I(x, y, rt)$ or $\underset{rt}{\operatorname{argmax}}\, C(x, y, rt)$.

This knowledge-base expansion method has the advantage of being automated, and directly using the response template to express relationships other than hypernym and hyponym.

4. DIALOGUE MANAGEMENT FOR PERSUASION

Now that we have defined our set of dialogue acts and knowledge base, we require a dialogue manager that can choose appropriate actions. We first describe a traditional baseline that attempts to achieve the user goal, and follow with the proposed persuasive dialogue manager.

4.1. Baseline model

As a baseline that focuses on the user goal, we use the dialogue manager introduced in the decision support system of [6]. We assume that 1) the user makes a decision to choose an alternative satisfying the user preference from the available candidates and 2) the user's latent preference for each topic is invariant through the dialogue. As mentioned in Section 2, at least one determinant matches the user preference.

In order to find alternatives that achieve the user goal, the baseline model estimates the user's preference over time using a dynamic Bayesian network (DBN). While the dialogue system uses information about the user's preferences to manage the conversation, these preferences are initially unknown to the system, and must be estimated through dialogue. To manage this process, the dialogue state of the system is modeled in a probabilistic manner by the following equation:

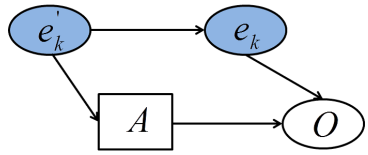

Fig. 4. Description of the dialogue state by Bayesian network in the baseline model. Each node represents a variable, and each edge represents a probabilistic dependency.

$$P(e_k|A, O) = \frac{P(O|A, e_k) \sum_{e_k^{'}} P(e_k|, e_k^{'}) P(e_k^{'})}{P(O|A)} \quad (2)$$

where e_k and $e_k^{'}$ represent the user's estimated preference in a topic k at the previous and current turns respectively, A represents the previous action of the system, and O represents the observed user action. Each action is further divided into dialogue act da such as *inform* and *question*, for topics $\boldsymbol{k}$, so that $A = \langle da_A, \boldsymbol{k_A} \rangle$ and $O = \langle da_O, \boldsymbol{k_O} \rangle$.

Fig 4 shows a Bayesian network for the dialogue state according to this model. The dialogue history from the beginning of the dialogue to the present time is represented as a dialogue state (called a belief state). Specifically, the belief state considers the estimated user preference, dialogue history including turn and recognized user action, and the alternative the system would like user to choose (the persuasive target). Because of the previous assumption that the user's latent preference is invariant, we set $P(e_k|e_k^{'}) = 1$ if $e_k = e_k^{'}$, and $P(e_k|e_k^{'}) = 0$ otherwise. Thus allowing us to simplify (2) to

$$P(e_k = 1|A, O) = \frac{P(O|A, e_k = 1) P(e_k^{'} = 1)}{P(O|A)} \quad (3)$$

and similarly for $e_k = 0$. $P(O|A, e_k)$ is determined on the basis of questionnaire and dialogue corpus of conversation between students and a system whose $P(O|A, e_k)$ is manually set. Procedure of data collection is same to those of experiment in Section 5.2. We determine the topics in which true user preference exists by using a questionnaire asking "which research field is important when you decide the laboratory to enter?" Next, subjects use the system, and we collect and annotate dialogue acts from those conversations. We use user preference and actions of the user and system to calculate probabilities

$$P(O|A, e_k) = \frac{N(O, A, e_k)}{N(A, e_k)} \quad (4)$$

where $N(O, A, e_k)$ represent the count of the tuples of system action O, user action A and the user answer about preference

e_k. If a particular topic $e_{k'}$ was not selected as a user's preferred topic even once in the dialogue corpus, we use only dialogue act da to calculate a smoothed version of Equation (4). The prior probability on the preference $P(e_k)$ can be set to a uniform distribution over all laboratories.

The next system action is determined on the basis of the present belief state. The policy of the baseline model according to the decision support system can be summarized below.

1. If the recognized user utterance is a question, the system answers the question.
2. If the estimated user preference in the persuasive target is higher than any of the other alternatives, or if a certain number turns has elapsed, the system *solicits* the user to join the laboratory.
3. If neither of the above two conditions are fulfilled, the system *informs* the user of the relationship between the current topic (or the topic of highest preference) and a topic randomly selected from the topics which have a direct relation to the current topic.

4.2. Proposed persuasive dialogue model

In this section, we describe our proposed expansions to the baseline that allow the system to also persuade the user to choose the alternatives satisfying the system goal. We first define the reward function for the system based on the system and user goals. The system goal reward is defined as

$$R_{system}(k_1, ..., k_n) = e_{k_{n-1}} \cdot w_{system}(k_n) \quad (5)$$

where $k_1, .., k_n$ represents the series of topics mentioned by the system and e_i represents the users preference for the topic. $w_{system}(k_n)$ represents how important topic k_n is as a persuasive target, and $e_{k_{n-1}}$ represents the user preference for k_{n-1}. As the system's final action is to inform the user that k_{n-1} has a relationship to the persuasive target k_n, we also consider the user's preference $e_{k_{n-1}}$. As a first step for this research, we assume that $w_{system}(k_n)$ is set to non-zero only if $k_n = k_{target}$.

$$R_{user}(k_1, ..., k_n) = \prod_{i=1}^{n} e_{k_i} \quad (6)$$

represents the user goal reward, which is higher if the dialogue covers many of the user's preferred topics[1].

In order to perform a dialogue maximizing these two rewards, it is important to talk about topics related to the target alternative, both to present more information about the target, and to efficiently estimate the user's preference for the determinants covering the target alternative. Therefore, functionality to use the previously described knowledge base to guide the user from the present topic to another topic that the system desires is central to the success of our system.

We propose a method that provides the system with an action that can be used to lead the conversation from topic to topic. The present topic is determined by the user's action at the present turn. As the next topic, the system selects a topic more closely related to the target alternative based on the estimated user preference as follows:

1. The system builds a weighted graph with nodes representing topics, edges representing relations between topics from the system knowledge base, and edge weights representing log probabilities $\log(P(e_k|A, O))$ based on the system knowledge and the estimated user preference $P(e_k|A, O)$ described in Section 4.1.
2. The starting point e_{k1} is set to the current topic and the ending point e_{kn} is set to the persuasive target. Then, the system finds the sequence $e_{k1}...e_{kn}$ that maximizes the sum of the edge weights.
3. The system selects e_{k2} as the next topic and generates an action to guide the user to the chosen topic.

Finally, we incorporate this functionality of guiding the user from topic to topic by using the strategy of the baseline model, but replacing step 3 with the process described above.

5. EXPERIMENTAL EVALUATION

5.1. Evaluation of the knowledge base

First, in order to experimentally evaluate the effectiveness of the knowledge-base described in Section 3, we evaluate quality of the responses generated by each method.

We first select 148 topics from the home page of the "Nara Institute of Science and Technology, Graduate School of Information Science." For the web-based knowledge-base, we used 4 types of queries, (a) "x is used in y," (b) "x technology y," (c) "research field of x y," (d) "x y"[2]. As the system utterance, we use the pair with the highest number of hits or mutual information for any of the 4 queries. However, when the number of hits is used as a measure, we use (d) only when search using (a), (b), and (c) all fail to produce any results.

We empirically compare 3 methods using the manually created knowledge base, 6 methods using Web search, and 1 method that does not use the relationship between terms:

- Manually created knowledge base [Section 3.1]: Three types of relations in the knowledge-base: parent-child (Dic1), grandparent-grandchild (Dic2), and common ancestor (Dic3).
- Web search [Section 3.2]: The web hit $C(x, y, rt)$ and mutual information $I(x, y, rt)$ critera are used to select topic pairs in the top 0-10% (Hit1, MI1), top 10-50% (Hit2, MI2), and top 50-100% (Hit3, MI3).
- Baseline: All links between the 148 topics (ALL).

[1]Typically, user goal is evaluated by user satisfaction, but it is not trivial to judge automatically. We assume that user satisfaction has a high relation with how much the dialogue touches upon preferred topics.

[2]We used "CiNii", http://ci.nii.ac.jp as a search engine.

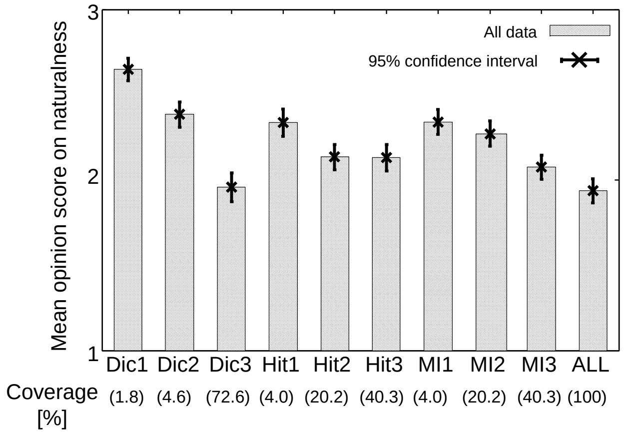

Fig. 5. Naturalness and coverage of topic pairs chosen by each method.

We randomly sampled 20 utterances from the utterances generated by each of the 10 methods, and evaluate a total of 200 statements at each subject. As evaluation criteria for each of the methods for generating topic pairs we use the coverage of term pairs and naturalness of the generated utterance (1: unnatural, 2: neutral, 3: natural). The coverage is the ratio of utterances that each method can generate out of all combinations of topic pairs[3]. There were a total of 17 evaluators.

We show the result of the evaluation in Fig 5. The experimental results show that the knowledge-base Dic1 achieves the highest naturalness, but at the cost of low coverage. In contrast, ALL achieves the highest coverage, but low naturalness. HIT1, MI1, MI2 achieve essentially the same naturalness as using grandparent-grandchild relations in the knowledge base Dic2. Thus, if we add the knowledge acquired using mutual information to the manually created knowledge, total coverage can be increased, while maintaining relatively high naturalness.

Based on these results, in the experiments in Section 5.3, we use system knowledge consisting of the knowledge Dic1, Dic2, Dic3 and the knowledge acquired using MI1, MI2. The priority of each method is chosen based on naturalness and success rate of guiding in the present dialogue. If there are some methods that can be used to guide between a topic pair, only the one what has the highest score is selected as the method for the pair.

5.2. Evaluation of dialogue management in a prototype system

Next, to evaluate the effectiveness of the proposed dialogue management technique, we first compare systems in a small task similar to that described in Section 2. We conducted a subjective evaluation of the baseline model (Baseline) and a system expanding the baseline model with functionality to guide users to target topics (+Guiding). In this experiment, as alternatives, two laboratories (A and B) exist, and the system attempts to encourage the user to join one of the two laboratories. We use 14 research fields as determinants.

The experimental procedure is as follows:

1. We instruct subjects to select two research fields satisfying their preference from a list of research fields.
2. We determined the research fields that each laboratory is working on. In order to cover tasks of varying difficulty, we assign one of the following conditions for each subject.
   Easy The persuasive target covers two of the chosen research areas and the other laboratory covers one of the research fields.
   Medium Both laboratories cover both two of the research fields.
   Hard The persuasive target covers one of the chosen research areas and the other laboratory covers two of the research fields.
3. Subjects start to use the systems without knowledge of the research fields each laboratory is working on.
4. If subjects request that they want to close the dialogue, the dialogue is finished. Upon completion of the dialogue, we took a questionnaire for each user about satisfaction and naturalness of the dialogue.

In order to reduce the influence of errors in speech recognition, language understanding, and speech synthesis, we substitute these modules with a human wizard-of-Oz, with subjects being aware that they were talking to a human. The number of subjects is 8, with each subject using each of the 2 systems 2 times each. The evaluation criteria are:

Success rate The percentage of time that the user chose the laboratory that the system intended.
Satisfaction 5 level score of user satisfaction
(1: Not satisfied, 3: Neutral, 5: Satisfied)

The results of the evaluation as measured by success rate and satisfaction are shown in Figure 6. We see increases in both evaluation measures for +Guiding over Baseline. Based on an analysis of the experiment data, Baseline could correctly estimate the preferred determinant in early stages, but does not make a connection between the research fields that match user preference to the laboratories, and as a result the user was not able to select a research field in many cases, resulting in a low success rate and satisfaction. On the other hand +Guiding could make the connection guiding from preferred determinant to target laboratory, both increasing the success rate and satisfaction.

5.3. Evaluation of the systems at large scale domain

We also conducted an evaluation of +Guiding on a larger task. The system policy was incrementally modified by repeated manual evaluation and selecting the better system from the

[3]Note that the methods based on Web search might not generate all of topic pairs due to lack of coverage of the search engine.

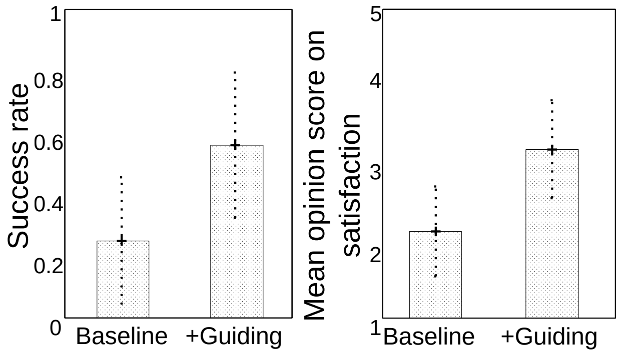

Fig. 6. Evaluation result. Success rate of persuasion (left), mean opinion score on satisfaction (right). The error bar represents 95% confidence interval.

revised system and previous system. $P(O|A, e_k)$ was retrained using the questionnaire results and dialogue corpus, and $P(e_k)$ and $P(e_k|e_k^{'})$ are the same as the previous section.

As system knowledge, there are 128 determinants and 22 alternatives. The experimental procedure and evaluation criteria are almost same as the previous section. The first difference is that the user was chosen to have some knowledge about the relationship of the research areas and each laboratory. Another difference is persuasive goal selection. In this experiment, we randomly select the systems persuasive goal from the laboratories which has direct relationship to the research fields selected by user in first step of experimental procedure.

As a result of the evaluation, the system achieved a success rate of 55%, and average satisfaction of 2.75. The average number of alternatives that satisfy at least one determinant that was selected by the user before starting the dialogue was 3.4. Therefore, if the user selected randomly from these alternatives, the success rate would be 29%. Thus, the system has been shown to be able to persuade effectively at a larger-scale task. In contrast, user satisfaction is low, mainly because the system persuasive target is invariant throughout the dialogue, and the system continues to guide to the persuasive target even if the user is not interested in the target alternative (i.e. persuasion implicitly failed). For further improving user satisfaction, we must consider methods to determine when the system and user goals are mutually exclusive, so the system can give up on persuasion and proceed solely on user interest.

6. CONCLUSION

In this paper, we constructed a dialogue manager for persuasive dialogue systems that encourage the user to make a decision that fulfill the system goals. We proposed a method to guide the user to target topics within a probabilistic framework for modeling the dialogue state. Experimental results indicate that proposed methods are effective for improving persuasive power, as well as the user's satisfaction.

In the future, we plan to optimize the system using actual data from persuasive dialogues. The current system also has the restriction that the persuasive target is invariant. For a more general persuasive dialogue model, we need to consider changing or abandoning the persuasive target.

7. REFERENCES

[1] Wayne Ward, "The CMU air travel information service: Understanding spontaneous speech," *Proceedings of the Workshop on Speech and Natural Language*, 1990.

[2] Antoine Raux, Brian Langner, Dan Bohus, Alan W Black, and Maxine Eskenazi, "Let ' s go public! taking a spoken dialog system to the real world," *Proceeding of the International Speech Communication Association*, 2005.

[3] Lori Lamel, Sophie Rosset, Jean-Luc Gauvain, and Samir Bennacef, "The LIMSI ARISE system for train travel information," *Proceeding of International Conference on Acoustics, Speech, and Signal Processing*, 1999.

[4] Pontus Warnest, "User evaluation of a conversational recommender system," *Proceedings of Workshop on Knowledge and Reasoning in Practical Dialogue Systems*, 2005.

[5] Paolo Viappiani and Craig Boutilier, "Regret-based optimal recommendation sets in conversational recommender systems," *Proceedings of the Workshop on Knowledge and Reasoning in Practical Dialogue Systems*, 2005.

[6] Teruhisa Misu, Komei Sugiura, Kiyonori Ohtake, Chiori Hori, Hideki Kashioka, Hisashi Kawai, and Satoshi Nakamura, "Modeling spoken decision support dialogue and optimization of its dialogue strategy," *ACM Transactions on Speech and Language Processing*, 2011.

[7] Joseph Weizenbaum, "ELIZA – a computer program for the study of natural langage communication between man and machine," *Communications of the Association for Computing Machinery*, 1966.

[8] Bayan Abu Shawar and Eric Atwell, "Using dialogue corpora to train a chatbot," *Proceedings of the Corpus Linguistics Conference*, 2003.

[9] Shinsuke Higuchi, Rafal Rzepka, and Kenji Araki, "A casual conversation system using modality and word associations retrieved from the web," *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 2013.

[10] B.J.Fogg, *Persuasive Technology*, Morgan Kaufman, 2003.

[11] Anton Nijholt, Ronald C Arkin, Sebastien Brault, Richard Kulpa, Frank Multon, Benet Bideau, David Traum, Hayley Hung, Eugene Santos Jr, Dequing Li, Fei Yu, Lina Zhou, and Donsong Zhang, "Computational deception and noncooperation," *IEEE Intelligent Systems*, 2012.

[12] Marco Guerini, Oliviero Stock, and Massimo Zancanaro, "Persuasion model for intelligent interfaces," *Proceedings of the IJCAI Workshop on Computational Models of Natural Argument*, 2003.

[13] Kallirroi Georgila and David Traum, "Reinforcement learning of argumentation dialogue policies in negotiation," *Proceedings of INTERSPEEECH*, 2011.

[14] Hien Nguyen, Judith Masthoff, and Pete Edwards, "Persuasive effects of embodied conversational agent teams," *Proceedings ofHuman-Computer Interaction*, 2007.

[15] Stephen G Pulman, "Conversational games, belief revision and baysian networks," *Proceedings of the Association for Computational Linguistics*, 1996.

[16] Eric Horvitz and Tim Paek, "A computational architecture for conversation," *Proceedings of the Seventh International Conference on User Modeling*, 1999.

[17] George Armitage Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Kathering Miller, "Introduction to WordNet: An on-line lexical database," *International Journal of Lexicography*, 1990.