User Study of Spoken Decision Support System

Teruhisa Misu, Kiyonori Ohtake, Chiori Hori, Hisashi Kawai and Satoshi Nakamura*

National Institute of Information and Communications Technology (NICT), Kyoto, Japan

`teruhisa.misu@nict.go.jp`

Abstract

This paper presents the results of the user evaluation of spoken decision support dialogue systems, which help users select from a set of alternatives. Thus far, we have modeled this decision support dialogue as a partially observable Markov decision process (POMDP), and optimized its dialogue strategy to maximize the value of the user's decision. In this paper, we present a comparative evaluation of the optimized dialogue strategy with several baseline strategies, and demonstrate that the optimized dialogue strategy that was effective in user simulation experiments works well in an evaluation by real users.

Index Terms: Spoken dialogue systems, user modeling, POMDP, decision support systems

1. Introduction

In many situations where spoken dialogue interfaces are installed, information access by the user is not itself a goal, but a means for decision-making [1]. For example, in using a restaurant retrieval system, the user's goal may not be obtaining price information but instead making a decision based on the retrieved information about the restaurants. In these situations, users try to extract information from the system that will aid their decision-making. Yet users, often unaware not only of what kind of information the system can provide but also their own preferences or factors in which they place value, can ultimately retrieve insufficient information. Systems themselves have little knowledge of the users, or where their interests lie; thus a system must bridge these gaps by sensing (potential) user preferences and recommending information in which they would be interested, considering the trade-off with the length of the dialogue.

To handle such use cases, we have proposed a user model and dialogue state representation, which consider user preferences as well as their knowledge about the domain changing through a decision support dialogue by modeling as a partially observable Markov decision process (POMDP). A dialogue strategy for information recommendation was optimized, and its effectiveness was confirmed by user simulation [2]. Unfortunately, as mentioned in several studies, an improvement in the simulation environment does not necessarily mean an improvement in a real user experiment [3, 4]. The main point of this paper is thus to demonstrate that the optimized dialogue strategy is also effective with real users.

The paper is organized as follows. Section 2 gives an overview of spoken decision support dialogue. Section 3 covers an overview of the system and task domain. Section 4 reports on the proposed user model, dialogue state expression and their evaluation by user simulation. Section 5 gives user study of the system and Section 6 concludes the paper.

---

* currently with Nara Institute of Science and Technology (NAIST)
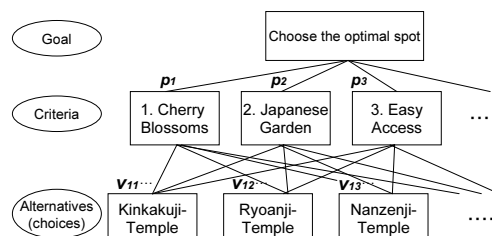
Figure 1: Hierarchy structure for sightseeing guidance dialogue

2. Spoken Decision Support Dialogue

We assume a situation in which a user selects from a given set of alternatives. In the real world this is highly likely, such as when a user selects a restaurant from a list of candidates presented by a car navigation system. This work considers a sightseeing planning task where the user determines the sightseeing destination, with little prior knowledge of the target domain. Consulting dialogues such as these are also regarded as a type of decision-making problem. That is, the user selects from a given set of alternatives based on certain criteria. Several studies in the operations research field have addressed decision support systems, with the Analytic Hierarchy Process [5] (AHP) being the typically employed method. In the AHP, the problem is modeled as a hierarchy consisting of the decision goal, the alternatives for achieving it, and the criteria for evaluating these alternatives. For the sightseeing planning task that we focus on in this paper, the goal is to decide on an optimal spot that aligns with the user's preference. The alternatives are comprised of all sightseeing spots that can be proposed and explained by the system. We adopt the decision criteria defined in our tagging scheme [6]. These include various deciding factors used in planning sightseeing activities, such as the presence of "cherry blossoms" or a "Japanese garden." Figure 1 shows an example hierarchy using these criteria.

In this model, the user's problem of making an optimal decision can be solved by fixing a weight vector $\mathbf{P}_{user} = (p_1, p_2, \ldots, p_M)$ for criteria and a local weight matrix $\mathbf{V}_{user} = (v_{11}, v_{12}, \ldots, v_{1M}, \ldots, v_{NM})$ for alternatives in terms of the criteria. The optimal alternative is then identified by selecting the spot k that maximizes the priority of $\sum_{m=1}^{M} p_m v_{km}$. In typical AHP methods, the procedure of fixing these weights is often conducted through pairwise comparisons for all possible combinations of criteria and spots in terms of the criteria, followed by weight tuning based on the results of the comparisons [5]. However, the methodology cannot be directly applied to spoken dialogue systems. To users, the information about a spot in terms of the criteria is not known, and is obtained only through the system's information navigation; thus, it is difficult to evaluate and compare the spots without this navigation. Spoken dialogue systems also usually handle several candidates and criteria, which makes pairwise comparison expensive. We therefore consider a spoken dialogue framework that estimates the weights for the user's preference (poten-

28 – 31 August 2011, Florence, Italy

tial preferences) as well as the user's knowledge of the domain through interactions of information retrieval and navigation.

3. Decision support system with spoken dialogue interface

Our dialogue system has two functions: answering users' information requests and making recommendations. When asked to explain certain spots or their properties (decision criteria), the system provides an explanation in terms of the requested property. After providing the requested information, it then provides information to aid in making a decision (e.g., instructing what the system can explain or recommending details on the current topic that the user might be interested in). Note that the latter is optimized via reinforcement learning.

Our back-end database consists of 15 sightseeing spots as alternatives and 10 decision criteria described for each spot [1]. We select decision criteria that frequently appear in our corpus [6]. Our candidate spots are evaluated and annotated in terms of these criteria if they apply to them. The value of the evaluation e_{nm} is "1" when the spot n applies to the criterion m and "0" when it does not.

The content of the recommendation is determined by one of six possible methods. The dialogue act (or action) of system recommendation a_{sys} consists of a communicative act ca (or recommendation method) and semantic content sc. The semantic content includes spots and/or criteria, which are determined by the heuristic rules defined for each method.

1. **Recommendation of criteria based on the currently focused on spot (Method 1)**
   This method is structured on the basis of the user's current focus on a particular spot. Specifically, the system selects several criteria related to the current spot whose evaluation is "1" and presents them to the user.
2. **Recommendation of spots based on the currently focused on criterion (Method 2)**
   This method functions on the basis of focus on a certain criterion.
3. **Open prompt (Method 3)**
   The system does not make a recommendation, and presents an open prompt.
4. **Listing of criteria 1 (Method 4)**
   This method lists several decision criteria to inform the user of the criteria that the system can handle, regardless of the current focus spot and criterion. The system lists the criteria in ascending order starting from what it estimates as low-level.
5. **Listing of criteria 2 (Method 5)**
   In this method, the system lists in the order of what it estimates as user's higher preference.
6. **Recommendation of user's possibly preferred spot (Method 6)**
   The system recommends spots that the users would be interested in based on the estimated preferences. It selects several spots that matches to the user's preferences based on the estimated user preferences. This collaborative filtering like method will be helpful to users if the system successfully estimates the user's preference; but it irrelevant if it does not.

---

[1] The number of alternatives is small compared to systems dealing with information retrieval, but note that this work focuses on the process of comparing and evaluating candidates that meet "essential condition" (e.g., Famous temple easily accessible on foot from Kyoto station).

4. Optimization of dialogue strategy

4.1. User modeling

We introduce a user model that consists of a tuple of knowledge vector $\mathbf{K}_{user}$, preference vector $\mathbf{P}_{user}$, and local weight matrix $\mathbf{V}_{user}$. In this paper, for simplicity, a user's preference vector or weight for criteria $\mathbf{P}_{user} = (p_1, p_2, \ldots, p_M)$ is assumed to consist of binary parameters. That is, if the user is interested in (or potentially interested in) the criterion m and places value on it when making a decision, the preference p_m is set to "1". Otherwise, it is set to "0". In order to represent a state in which the user has potential preferences, we introduce a knowledge parameter $\mathbf{K}_{user} = (k_1, k_2, \ldots, k_M)$. This parameter represents whether the user perceives that the system can accommodate his/her preferences. It is also interpreted as a parameter to indicate if the user is aware that he/she is interested in the criteria. k_m is set to "1" if the user knows that the system can handle criterion m and "0" when he/she does not. For example, the state that the criterion m is the potential preference of a user (but he/she is unaware of that) is represented by $(k_m = 0, p_m = 1)$. k_m is updated to "1" when the system informs the user of the recommendation through Methods 1, 2, 4, and 5. A user's local weight v_{nm} for spot n in terms of criterion m works as the user's knowledge on whether the preference p_m is satisfied by visiting the spot n. v_{nm} is set to "1", when the system lets the user know that the evaluation of spots are "1" through recommendation Methods 1 and 2.

4.2. Dialogue state expression

In above section we presented the user state representation. However the problem is that for the system, the state $(\mathbf{P}_{user}, \mathbf{K}_{user}, \mathbf{V}_{user})$ is not observable, but is only estimated from the interactions with the user. Therefore this model can be seen as a partially observable Markov decision process (POMDP) [7]. In order to estimate unobservable properties of the POMDP, we introduce the system's inferential user knowledge vector $\mathbf{K}_{sys}$ or probability distribution (estimate value) $\mathbf{K}_{sys} = (Pr(k_1 = 1), Pr(k_2 = 1), \ldots, Pr(k_M = 1))$ and that of preference $\mathbf{P}_{sys} = (Pr(p_1 = 1), Pr(p_2 = 1), \ldots, Pr(p_M = 1))$. In this work, we do not estimate the weight, because v_{nm} is assumed to be set to "1" only when the system lets the user know that the evaluation of the criterion m of the spot n is "1" through recommendations, thus $\mathbf{V}_{sys} = \mathbf{V}_{user}$. This consequently means that $\mathbf{V}_{user}$ is observable.

The dialogue state DS^{t+1} or estimated user's dialogue state of the step $t+1$ is assumed to be dependent only on the previous state DS^t, as well as the interactions $I^t = (a_{sys}^t, a_{user}^t)$. This approximation is often adopted in spoken dialogue management systems using a dynamic Bayesian network (DBN) representation [4]. The relation of the parameters used in our model is illustrated as DBN in Figure 2. The estimated user's preference $\mathbf{P}_{sys}$ is updated when the system observes the interaction I^t. The update is conducted using Bayes' Theorem, with the previous state DS^t as a prior. The posterior of the estimated user's knowledge of criteria k_m is updated to "1" when the system tells or the user requests the criteria. An example of this update is illustrated in Figure 3.

4.3. Reward function

The reward function that we use is based on the number of attributes agreed upon between the user preference and the decided spot. The reward $\mathbf{R}$ is then calculated based on the improvement in the number of agreed attributes between the user's actual (potential) preferences and the decided spot k over the
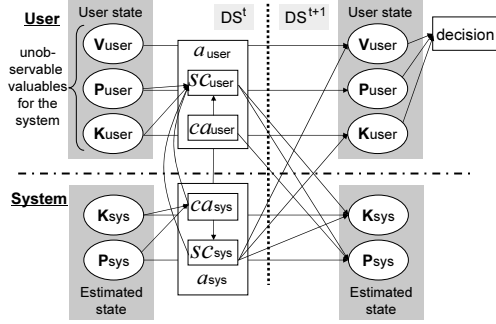
Figure 2: Dynamic Bayesian network of the model

Priors of the estimated state:
- Knowledge: $\mathbf{K}_{sys} = (0.22, 0.31, 0.02, 0.18, \ldots)$
- Preference: $\mathbf{P}_{sys} = (0.37, 0.19, 0.48, 0.38, \ldots)$

Interactions (observation):
- System recommendation:
$a_{sys} = Method1\{(Spot_5), (Det_1, Det_3, Det4)\}$
"Japanese garden (Det_1), World heritage (Det_3) and fall foliage (Det_4) are some of the areas about which information is available on Ninnaji ($Spot_5$)."
- User query: $a_{user} = Accept\{(Spot_5), (Det_3)\}$
"Tell me about world heritage sites. (Det_3)"

Posterior of the estimated state:
- Knowledge: $\mathbf{K}_{sys} = (1.00, 0.31, 1.00, 1.00, \ldots)$
- Preference: $\mathbf{P}_{sys} = (0.26, 0.19, 0.65, 0.22, \ldots)$

Figure 3: Example of state update

expected agreement by random spot selection.

$$R = \sum_{m=1}^{M} p_m \cdot e_{km} - \frac{1}{N} \sum_{n=1}^{N} \sum_{m=1}^{M} p_m \cdot e_{n,m}$$

For example, if the decided spot satisfies three preferences and the average by random selection is 1.3, then the reward is 1.7.

4.4. Experiment by simulated users

The dialogue strategy of the system was optimized using the above dialogue state expression, reward function, and user simulator using the statics of our trial system via reinforcement learning so that the manager selects the optimal recommendation method matched to the dialogue state [2]. In the typical dialogue strategy of the optimized policy, the system first bridges the knowledge gap with the user and estimates the user's preferences (Method 4 and 5), then recommends specific information that would be useful to the user (Method 6). This flow is similar to the strategy of a human guide collected in our dialogue corpus [6].

Prior to the user study, we examined the performance by user simulation. Simulated users are assumed to continue a dialogue for 5, 10, and 15 turns, and episodes are sampled using the optimized policy. They are also assumed to have four preferences[2], and to determine the spot based on their preference $\mathbf{P}_{user}$ under their knowledge $\mathbf{K}_{user}$ at that time, and select the spot k with a maximum priority of $\sum_m k_k \cdot p_k \cdot v_{km}$. The system parameters, $\mathbf{P}_{sys}$ and $\mathbf{K}_{sys}$ are initialized using the statistics obtained by the trial system [2]. The system is rewarded by the reward function of section 4.3. We simulated possible automatic speech recognition (ASR) and spoken language un-

---

[2]As a result, four parameters in $\mathbf{P}_{user}$ are "1" and the others are "0".

Table 1: Evaluation by user simulation

	Reward ($\pm std$)		
Strategy	T = 5	T = 10	T = 15
optimized	0.43 (0.59)	0.68 (0.64)	0.84 (0.62)
B1	0.22 (0.52)	0.46 (0.63)	0.65 (0.63)
B2	0.02 (0.41)	0.13 (0.55)	0.26 (0.58)

derstanding (SLU) errors [8], assuming that the semantic content is deleted at 16.7%, which is the probability that the system was not able to handle in-domain queries in the experiment using the trial system. The average reward that the system would have received when the simulated users made the decision after 5, 10, and 15 turns of interaction is listed in Table 1.

The following baseline strategies were compared with the trained policy:

1. **Random recommendation (B1)**: The system randomly chooses a recommendation from six methods.
2. **No recommendation (B2)**: The system only provides the requested information and does not generate any recommendations.

The comparison of the average reward between the baseline strategies is also listed in Table 1. The reward by the optimized strategy was significantly better than that of baseline strategies[3].

5. User Evaluation

The three dialogue strategies discussed above were evaluated by 40 subjects who had not previously used spoken dialogue systems. Subjects were requested to use the system to select one sightseeing spot from among 15 alternatives. No instructions or scenarios were given. They were also requested to use the phrase "I've decided to go to (spot_name)," signifying their commitment once they had reached a decision.

We asked 20 users to use the optimized system first and to carry out a dialogue session of selecting one spot. After the dialogue session using the optimized system, 13 out of the 20 users were asked to use the baseline system 1 and seven users the baseline system 2, and another dialogue session was conducted. The other 20 users are asked to use the systems in reverse order. That is, 13 users used the baseline system 1, then the optimized system, and seven users used the baseline system 2, then the optimized system. In total, 785 user utterances were collected. Note that only the first dialogue session for each subject was a truly valid dialogue episode for our experiment, since the first such dialogue would very likely alter the state of user knowledge. After the dialogue sessions, users were asked to select their preferences (four out of 10 criteria) they place value on when selecting sightseeing spots through questionnaires. Since subjects were asked to select from among all criteria, their selections are considered to be their preferences within the full knowledge of the system. We assumed that the values of p for selected criteria are "1".

Before comparing the dialogue strategies, we examined if the user decisions are based on the priority of $\sum_m k_k \cdot p_k \cdot v_{km}$ as assumed in the user simulator. We regarded the k_m and v_{nm} as "1", when the system lets the user know the information through recommendations, and examined the order of the selected spot in the priority. The average order was 1.4, and more than half of users selected the spot with the first priority. These results suggest that user decisions are related to the priority, although users may be unconscious of the priority.

We then focused on the users' first sessions and examined the obtained reward, length of the dialogue (number of interactions before user's commitment), out-of-domain (OOD) query

---

[3]The maximum possible reward, which is achieved when the user selects the spot that best satisfies their preferences, was 1.47.

Table 2: Evaluation by human subjects (first session)

Strategy	Reward	Length	% OOD	% Error	% Accept
optimized	**0.85**	10.1	18.2	1.0	52.5
B1	0.07	9.3	21.4	1.8	27.0
B2	0.09	11.8	21.1	1.4	-

rate, ASR/SLU error rate, and user's acceptance rate of system recommendation[4]. The results are listed in Table 2. As with the result of the simulation, the system with the optimized dialogue strategy obtained a much higher reward. The average reward of the optimized strategy 0.85 was significantly higher than that of 0.07 and 0.09 by baseline strategies ($p < .05$). The user's acceptance rate of the system recommendation generated by the optimized strategy was 52.5%, and the percentage of the OOD utterance rate was lower than that of the baseline strategies. These figures suggest that the optimized dialogue strategy can recommend appropriate information matched to the user's dialogue state, resulting in better decisions by users.

For reference, we then evaluated the result of the second sessions and compared the results with that from the first sessions. The rewards of the sessions are listed in Table 3. Interestingly, the average reward of the optimized strategy of users in the second session (= users who used the system with the optimized dialogue strategy after using the baseline system) was much smaller than that of the users in the first session. The primary reason for this would be that users selected their second-best spots, since all users selected a different spot with that which they selected in the first session. Another reason for this would stem from the fact that the user's acceptance rate for system recommendations was lower, resulting in a failure in estimating user preferences. Actually, the acceptance rate of users of the system with an optimized strategy in the second session was 30.0%, which was much lower than that of 52.5% by users in their first session. The relationship between the average acceptance rate and reward is plotted in Figure 4. Although there was no strong correlation, the rewards of the system by the users with a low acceptance rate were very likely to be low. While about half of the users who used the optimized system in the first session accepted system recommendations, at more than 50%, more than half of users who used the optimized system after using the baseline system did not accept the system recommendation at all. The reason for the low acceptance rate in the second session would be that the optimized strategy was intended for users (and trained by simulated users) with little knowledge about the system; consequently, several recommended items of information were known information for the users in the second session. In the trained strategy, the system usually estimates user preference via methods 4 and 5 at an early stage. And since many users in the second trial did not accept the recommendation, the system seemed to fail to estimate user's preferences[5] and recommending appropriate spots matched to the user. These results indicate the importance of the prior information for the system parameters in making appropriate recommendations.

Nonetheless, we believe that the optimized strategy and the dialogue state representation are effective because every user will experience the first use, and it is important to enable users to succeed at the first use, as this will likely influence the user's

---

[4]Note, that Method 1, 4 and 5 are regarded as accepted when the user requests information on either of the recommended criteria. Method 2 and 6 are regarded as accepted when the user requests information on either of the recommended spots.

[5]The system could detect only 1.1 out of four preferences of the users who used the baseline system in advance. The figure was much worse than that of 1.8 for the user who used the optimized system first.

Table 3: Comparison of first and second sessions

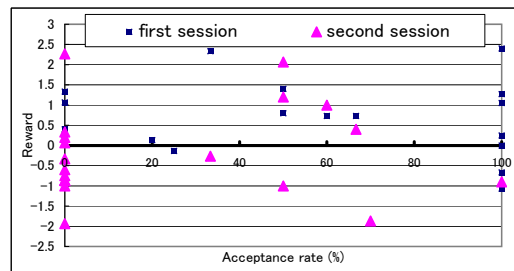| Strategy	Reward (1st session)	Reward (2nd session)
optimized	**0.85**	**-0.17**
B1	0.07	-0.06
B2	0.09	0.28

Figure 4: Relationship between recommendation acceptance rate and reward

beliefs and their impressions of the system. In addition, the problem is considered to be solved by maintaining the system use history of the user (e.g. by using phone number [9]) and coordinating prior information.

6. Conclusion

In this paper, we addressed the results of user evaluations of several dialogue strategies in spoken decision support dialogue systems. We compared the optimized dialogue strategy based on POMDP-based dialogue state expression with several baseline strategies. We confirmed that the optimized strategy that was effective in user simulation experiments worked well in evaluations by real users. However, the results also demonstrated the importance of prior information on users. Future work thus involves estimating user profiles, for instance through the system use history, as well as training in dialogue strategies using multiple profiles of the users.

7. References

[1] J. Polifroni and M. Walker, "Intensional Summaries as Cooperative Responses in Dialogue: Automation and Evaluation," in *Proc. ACL/HLT*, 2008, pp. 479–487.

[2] T. Misu, K. Sugiura, K. Ohtake, C. Hori, H. Kashioka, H. Kawai, and S. Nakamura, "Dialogue Strategy Optimization to Assist User's Decision for Spoken Consulting Dialogue Systems," in *Proc. IEEE-SLT*, 2010, pp. 342–347.

[3] J. Schatzmann, M. Stuttle, K., and S. Young, "Effects of the user model on simulation-based learning of dialogue strategies," in *Proc. Automatic Speech Recognition and Understanding Workshop (ASRU)*, pp. 220–225.

[4] B. Thomson, K. Yu, M. Gasic, S. Keizer, F. Mairesse, J. Schatzmann, and S. Young, "Evaluating semantic-level confidence scores with multiple hypotheses," in *Proc. Interspeech*, 2008, pp. 1153–1156.

[5] T. Saaty, *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. Mcgraw-Hill, 1980.

[6] K. Ohtake, T. Misu, C. Hori, H. Kashioka, and S. Nakamura, "Annotating Dialogue Acts to Construct Dialogue Systems for Consulting," in *Proc. The 7th Workshop on Asian Language Resources*, 2009, pp. 32–39.

[7] R. Sutton and A. Barto, *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[8] O. Pietquin, S. Rossignol, and M. Ianotto, "Training Bayesian networks for realistic man-machine spoken dialogue simulation," in *Proc. First International Workshop on Spoken Dialogue Systems Technology (IWSDS)*, 2009.

[9] K. Komatani and H. Okuno, "Online Error Detection of Barge-In Utterances by Using Individual Users' Utterance Histories in Spoken Dialogue System," in *Proc. SIGDIAL*, 2010, pp. 289–296.