Influence of Data-derived Individualities on Persuasive Recommendation

Masashi Inoue[1][0000-0002-9364-3114] and Hiroshi Ueno[2]

[1] Tohoku Institute of Technology, Yagiyama Kasumicho 35-1, Sendai, Japan
m.inoue@acm.org
http://www.ice.tohtech.ac.jp/ inoue/index.html
[2] Yamagata University, Jyonan 4-3-16, Yonezawa, Japan

Abstract. In this study, two machine learning based approaches have been compared that can add personal communication traits to a conversational recommender system. The first approach involves the creation of generative models for reactive tokens such as backchannels. The second approach involves a method for rewriting the conversational text by applying machine translation. Both approaches can impart personal communication traits to systems that incorporate a dialogue corpus. Two methods were implemented for a persuasive recommender system and their positive or negative effects based on an individual's personality were experimentally analyzed through a restaurant ranking task. The results suggest that addition of personal communication traits decrease objective persuasiveness while increasing the individual's impression on recommender systems.

Keywords: Recommendation · Conversational · Persuasive.

1 Introduction

In this study, we consider a recommender system that achieves persuasiveness by adopting conversational aspects. Persuasive technologies have been studied in various fields including item recommendation[3]. The task involves changing a user's conceptualization of a product or service. One method to change users' minds is to present an explanation for a recommendation instead of simply a list of offering choices. The assumption is that if the explanation is reasonable, then users can apply logic regarding an offer, and hence, the user may accept the offer. However, according to our observation, users are often reluctant to change their mind even though the offerings are logical. That is, in addition to the message being delivered, how they are delivered should be taken into account in designing recommender systems. When the recommender system is conversational, one approach to modify the message delivery style is to add characters to the system. In this paper, we examined if there are either positive or negative effects of added personal communication traits to conversational recommender systems.

2 Related Works

Addition of characteristics to dialogue systems have been considered mostly by changing parameters reflecting predetermined generic personality dimensions such as degree of extraversion [7]. The change of behavior is considered to be a factor influencing personalities. For example, the effects of self-disclosure has been studies in a spoken dialogue system [9]. In this study, we considered text-based conversational systems and extract the specific profile of individuals.

3 Method

3.1 Reactive Token Generation (TokenGen)

To add personal communication traits, we employed two approaches. They are either non-lexical or lexical. For a non-lexical approach, we used the reactive token-based method [5]. This method extracts usage patterns of reactive tokens for a particular speaker in a corpus. Then, a probabilistic model for reactive token generation is created, which adds reactive tokens to the conversational system outputs. For example, tokens "I see" or "Uh-ha" are added. We call this method **TokenGen**. This method changes the manner in which information is delivered but does not change what is expressed.

3.2 Lexical Modification (LexMod)

For lexical modification, we adopted a machine translation-based method [8]. This method estimates probabilities of translation between a default word to the word that is peculiar to the speaker based on corpus statistics. Then, the utterances from the conversational system are modified based on the model. For example, "a nasty kid" may be changed into "a rude boy" and "Do you know?" may be changed into "You know, don't you?". Such replacement were conducted for function words in Japanese. We call this method **LexMod**. This method changes the manner in which information is expressed in a message. There were some unnatural utterances after automatic rewriting but we did not modified them manually.

3.3 Corpus

We need a corpus of different speakers as the basis for preparing personal communication traits. For the purpose, we used the Nagoya University Conversation Corpus (NUCC)[4] that is a transcribed corpus of Japanese natural conversation. The corpus contained the speakers' attributes, including their gender, age, and hometown. For the purpose of the pilot study, we extracted personal traits from three typical speakers: F1, F2, and M1. We selected both female (F) and male (M) speakers to compare difference in gender. Since there are more female speakers than male speakers (118 females and 20 males), we selected young and old female speakers (F1 and F2) for the purpose of examining age factors.

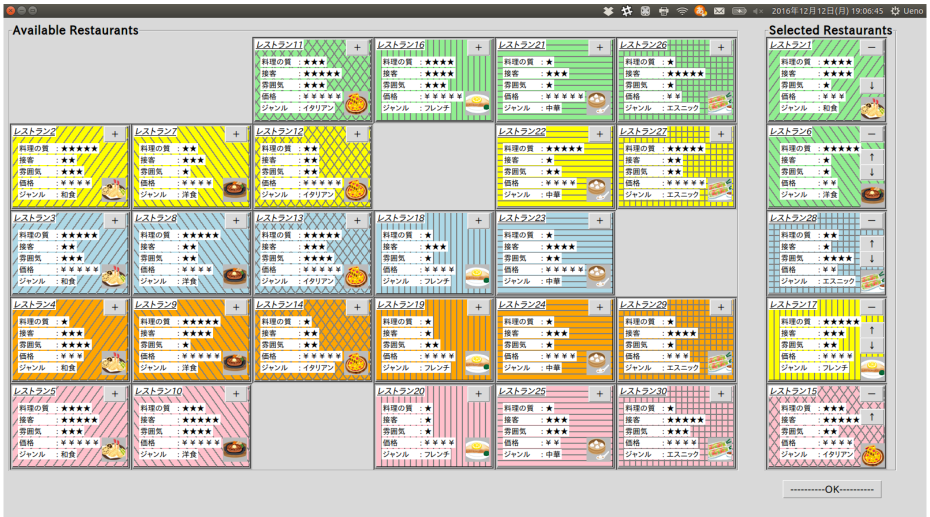

Fig. 1. Restaurant information provided to the task participants.

4 Experiment

4.1 Task

The task that the user is asked to carry out is restaurant ranking. First, task participants are asked to input preference information when selecting restaurants. The information consists of preferred food genre and preferred order of the restaurants' characteristics for selecting the relevant restaurant. For example, a user may describe the preference information as {(Genre), (Preference)} ={(Chinese), (food quality > service level > interior quality > price)} by using natural language. Second, the participants select five of thirty restaurants based on the information shown in the system interface. restaurants' information panel (Figure 1). Each restaurant is described based on its food quality, service level, interior quality, price, and genre. Then, the participants discuss their selection with the conversational recommender system that suggests re-ordering. This restaurant selection task is based on Andrews's work [2]. The recommendation can be conducted deterministically, not probabilistically. When there is logical inconsistency in user preferences based on the initial list of selected restaurants in terms of rankings, the system suggests changes to modify the ranking so that it conforms to the user-supplied preference data. For example, if a user stated that food quality was more important that the price range but the initial list was created based on the price range, the system asks to change the order of restaurants according to the food quality scores.

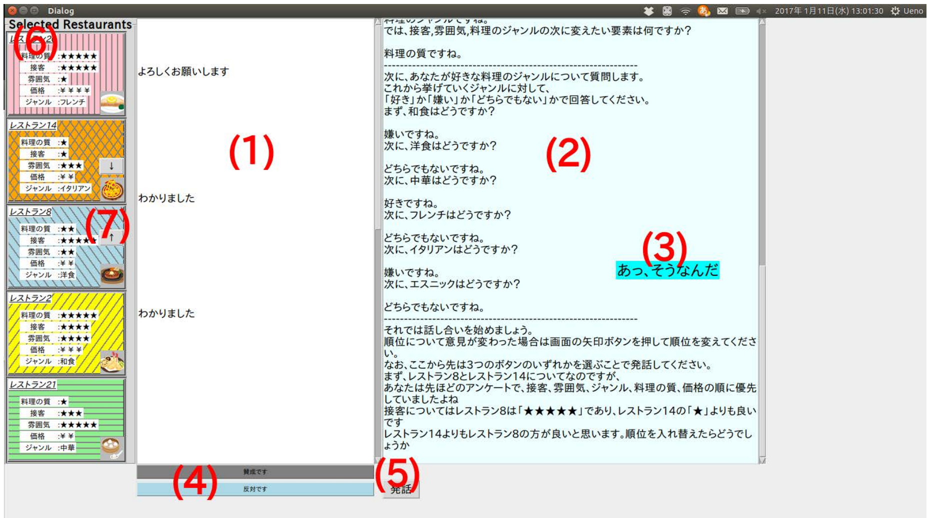

Fig. 2. Interface of the restaurant ranking system.

4.2 System Interface

The interface shown to the experiment participants is shown in Figure 2. The interface is a text-based one; users type in the text to send a message (1) and receive feedback from the system through textual modality. Simultaneously, there is a visual display of the current restaurant ranking (6). The candidate restaurants and any selected restaurants are shown as tile displays in separate windows. The goal of the persuasive recommender system is to change the ranking so that it does not conform to the initially described participants preferences. To the system suggestions, user can react by clicking either the acceptance button or the rejection button (4). There is a confirmation button for their decisions (5). When users accept the suggestions by the system, they can change the ranking by clicking the upward arrow button or the downward arrow button (7). While users and the system converse, there are back-channeling utterances appear in the system utterance window (3).

4.3 Experimental Procedure

The experiment was conducted as follows.

1. The task is explained using the dialogue system interface (on screen).
2. The participant select 5 favorite restaurants and rank them.
3. The participant answers the questionnaire on the importance of restaurant selection criteria.
4. The participant answers the questionnaire on the preferences for the cuisines.

Table 1. Estimation accuracies of individuals who are the source of added communication traits.

	F1	F2	M1
TokenGen Method	0.89	0.22	0.56
LexMod Method	0.67	0.56	0.78
Both Methods	0.63	0.00	0.63

5. The participant conducts interactions with the system. During the interaction session, the participant can change the restaurant ranking.
6. The participants answers the questionnaire on the system impression.

Each participant interact with one of three systems (TokenGen Method, LexMod Method, and Both Methods). Each participant evaluate both baseline default system (no individuality) and three different personal traits (F1, F2, and M1) extracted from the corpus. In the experiment, we have 26 participants. Among them, nine interacted with the system of TokenGen method, nine used the system of LexMod method, and eight experienced the system of both methods.

4.4 Evaluation Measure

Two measures were used for assessing the influence of added personal communication traits. The first is the objective persuasiveness, the second is subjective persuasiveness, and the third is the degree of satisfaction on the interactive session with the system. The objective degree of persuasiveness is measured by the ratio given by the occurrences of swapping the ranking by the users with those suggested by the system in the session. Further, the subjective persuasiveness was measured using a five-point scale questionnaire on the feeling of being persuaded. The degree of satisfaction was measure using questionnaire.

5 Results

First, we examined if our methods can add personal communication traits sufficiently so that the participants can feel the individualities from the systems. To assess the degree of personal trait representations when speakers are not well-known public figures and whose identities are not known in advance, utterance consistency had been used [6]. However, our recommendation scenario, the diversity of utterances are limited and the consistency is not considered to be a meaningful measure. Therefore, we used the person identification test [5]. In our system, each individuality has its source speaker in the corpus. We asked participants if they can identify the dialogue logs from the corpus that belong to the source individuals after interacting with the system. The results of this experiment are shown in Table 1. Although the accuracy scores for F2 was not satisfactory, we assumed that they are exceed the change rates and added modification gives information on the source speakers. The results of the experiments

Table 2. Influences of added individualities in terms of re-ranking occurrences (objective) and system impression scores evaluated by the participants (subjective).

	Objective Persuasiveness (%)	**Subjective Persuasiveness**	**Degree of Satisfaction**
Baseline	37	3.50	3.73
TokenGen Method	36	3.56	4.07
LexMod Method	35	3.26	3.63
Both Methods	33	2.96	3.88

are shown in Table 2. We compared four settings. The first setting, baseline, use the default system output as it is. The second setting employed the TokenGen method, the addition of reactive tokens to the default system outputs. The third setting utilized the LexMod method, rewriting default output text by using machine translation. The fourth is the combination of both methods. Since the first method is non-lexical and the second method is lexical, they can be used simultaneously. As shown in the left column of Table 2, the baseline method achieved higher objective persuasiveness. Subjective persuasiveness and the degree of satisfaction are higher in the TokenGen method than the baseline and LexMod systems based on the five-point scale, as shown in the middle and right columns. It may be natural that adding the personalities that are not necessarily persuasive did not improve the degree of persuasiveness of the recommender systems. An interesting result is that even though the objective persuasiveness for TokenGen system is lower than the default system, the subjective persuasiveness is higher for the TokenGen system. We should examine the cause of this discrepancy further.

6 Conclusion

In this study, we compared two methods to introduce individuality in a restaurant recommender system in order to evaluate their persuasiveness. Individuality is realized by extracting personal communication traits from the face-to-face dialogue corpus of diverse speakers. The first method uses reactive token selection and the second method involves text rewriting. Experimental results on system users suggest that both methods do not improve objective persuasiveness but they do differ in subjective persuasiveness and degree of satisfaction.

There are several topics to be considered. The lexical method we employed was based on the utterance selections. There are attempts to modify utterance style after generating utterances [1]. It would be interesting to examine the influences of base dialogue systems. Also, for open domain dialogue system, various evaluation measures were considered [10]. Utilities of these metrics in our restricted recommendation scenario can be considered.

References

1. Akama, R., Inada, K., Inoue, N., Kobayashi, S., Inui, K.: Generating stylistically consistent dialog responses with transfer learning. In: Proceedings of the Eighth International Joint Conference on Natural Language Processing. pp. 408–412. Taipei, Taiwan (2017)
2. Andrews, P.Y.: System personality and persuasion in human-computer dialogue. ACM Trans. Interact. Intell. Syst. **2**(2), 1–27 (2012)
3. Fogg, B.J.: Persuasive Technology: Using Computers to Change What We Think and Do. Science & Technology Books (2002)
4. Fujimura, I., Chiba, S., Ohso, M.: Lexical and grammatical features of spoken and written Japanese in contrast: Exploring a lexical profiling approach to comparing spoken and written corpora. In: Proceedings of the VIIth GSCP International Conference. Speech and Corpora. pp. 393–398 (2012)
5. Inoue, M., Ueno, H.: Dialogue system characterization by back-channelling patterns extracted from dialogue corpus. In: LREC. pp. 2736–2740 (2016)
6. Li, J., Galley, M., Brockett, C., Spithourakis, G., Gao, J., Dolan, B.: A persona-based neural conversation model. In: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics. vol. 1, pp. 994–1003 (2016)
7. Mairesse, F., Walker, M.: Personage: Personality generation for dialogue. In: Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics. pp. 496–503 (2007)
8. Mizukami, M., Neubig, G., Sakti, S., Toda, T., Nakamura, S.: Linguistic individuality transformation for spoken language. 6th International Workshop on Spoken Dialog Systems (IWSDS) (2015)
9. Ogawa, Y., Miyazawa, K., Kikuchi, H.: Assigning a personality to a spoken dialogue agent through self-disclosure of behavior. In: Proceedings of the Second International Conference on Human-agent Interaction. pp. 331–337. HAI '14 (2014)
10. Venkatesh, A., Khatri, C., Ram, A., Guo, F., Gabriel, R., Nagar, A., Prasad, R., Cheng, M.: On evaluating and comparing conversational agents. In: NIPS 2017 Conversational AI workshop. Long Beach (2018)