

I Remember You!: SUI Corpus for Remembering and Utilizing Users' Information in Chat-oriented Dialogue Systems

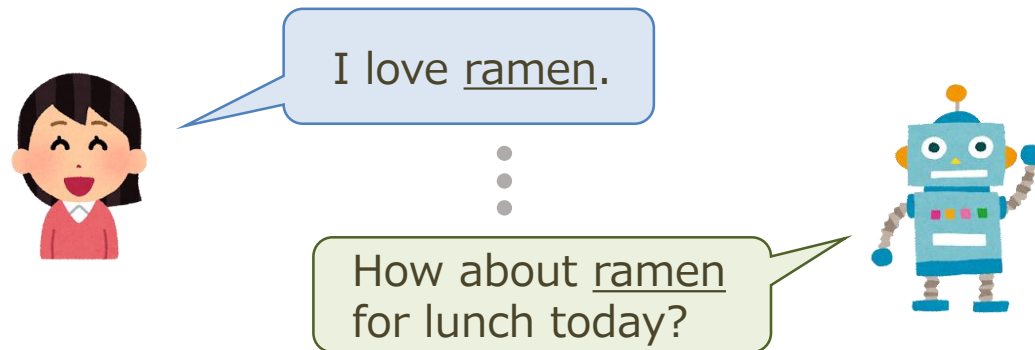
Yuiko Tsunomori and Ryuichiro Higashinaka

Graduate School of Informatics, Nagoya University, Japan



Background

- To construct a chat-oriented dialogue system that can be used for a long-term, it is important to build a good relationship with users [Richards+'14]
- In human-to-human dialogue, it is effective to use **information on the dialogue partner** for building a good relationship [Hall+'19]
- In system-to-human dialogue, several existing methods use **user information** in system utterances to build a good relationship with users



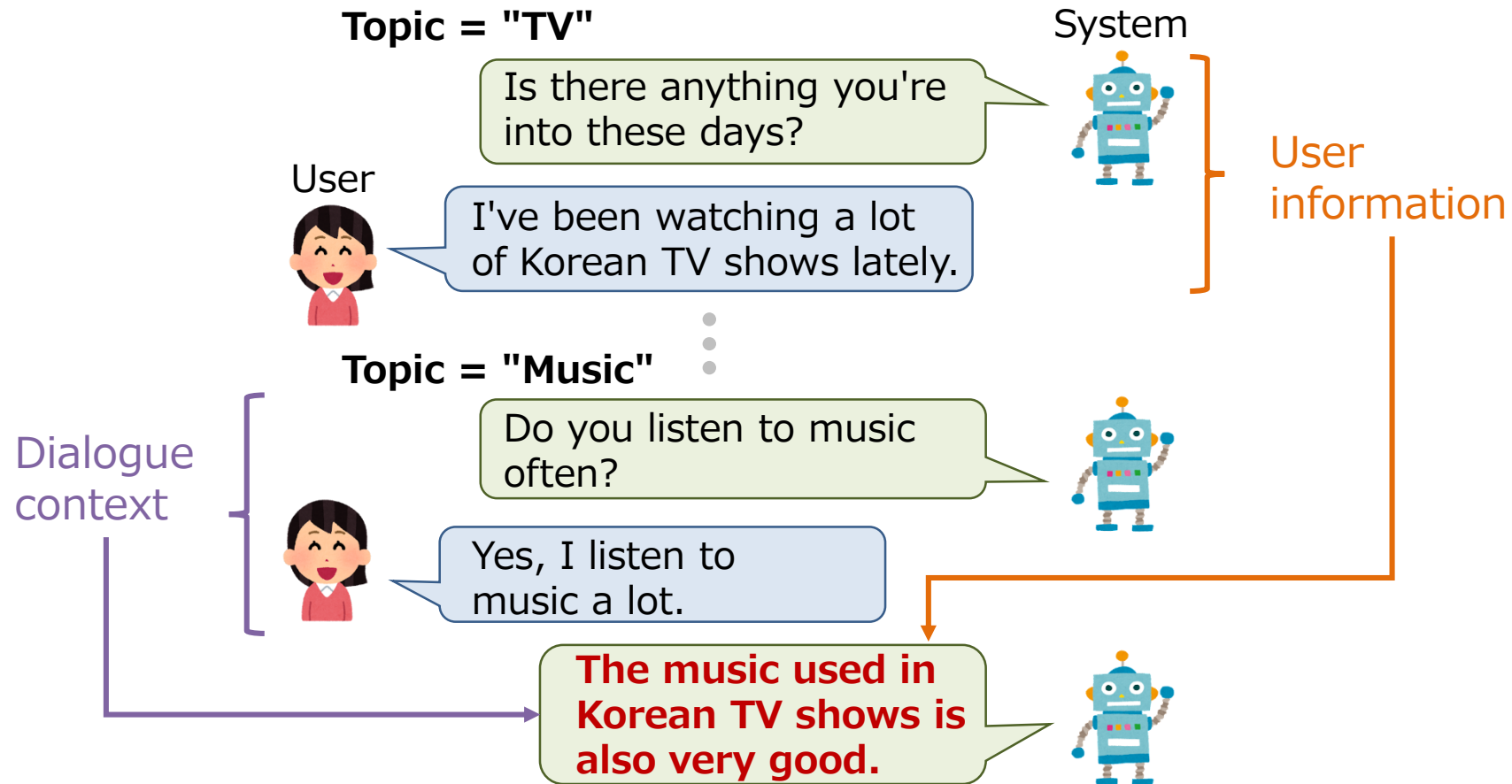
Previous studies

- Rule-based utterance generation using templates to be filled with user information [Tsunomori+'19]
 - 😊 The system using user information improved users' familiarity
 - 😞 The utterances were often unnatural with regard to the dialogue context due to the **expressions limited by utterance templates**
- Neural-based utterance generation using user information which is similar to the current dialogue context [Xu+'22]
 - 😊 The system generates natural utterances on the basis of dialogue contexts
 - 😞 **Limited user information** can be incorporated into utterances

Previous studies could not generate system utterances appropriate for dialogue context using various user information

Goal

Chat-oriented dialogue systems that can generate utterances using various user information while retaining appropriateness for dialogue contexts





Approach

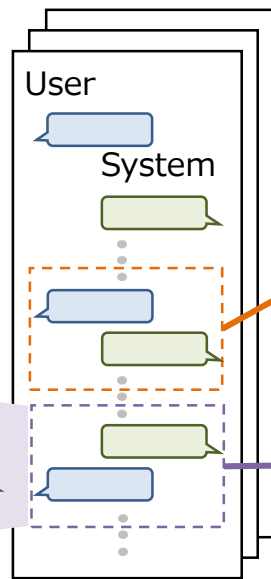
Constructing a corpus consisting of **system utterances incorporating various user information with a wide degree of relevance to the dialogue context**

- System utterance based on User Information corpus (SUI corpus)
 1. Prepare user information and dialogue context, which are on different topics
 2. Collect a system utterance based on the user information and the dialogue context (**expanded system utterance**)
- Verify the effectiveness of the model trained by the SUI corpus
 1. Fine-tune an utterance generation model by the SUI corpus and evaluate the generated utterances
 2. Incorporate the fine-tuned model into a dialogue system and evaluate the dialogues with users

Construction of the SUI corpus

- Extended the Osaka University Multimodal Dialogue Corpus (Hazumi) [Komatani+'19]
- Collected 10,801 **expanded system utterances** by crowdsourcing

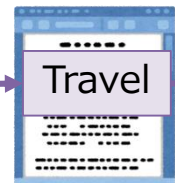
Divided each dialogue in Hazumi into topic segments



Topic segments in different topics



Dialogue-1



Dialogue-2

Crowdworkers conducted two tasks

(a) Extracted all user information from dialogue-1

- I like Harry Potter.
- I want to watch movies everyday.

(b) Created expanded system utterances based on one user information and dialogue-2

Don't you want to visit the Harry Potter filming locations in England?

Analysis of the SUI corpus

- Calculated the similarity of topics between pairs of dialogue-1 and dialogue-2
 1. Extracted embeddings of the topics using FastText [Bojanowski+'2017]
 2. Calculated the cosine similarity between them
- Results
 - The highest score was 0.48 and the lowest score was 0.09
 - ◆ The score 0.20-0.30 had the highest frequency
 - ◆ For reference, the average score of Japanese synonyms in WordNet was 0.40
 - SUI corpus consists of expanded system utterances incorporating various user information with a wide degree of relevance to the dialogue context

Example of the SUI corpus

User information	<p>S Do you like to drink alcohol? (お酒はお好きですか?)</p> <p>U Yes, I drink beer, sake, shochu, and most other alcohol. (そうですね ビールとか 日本酒 酎ハイ 大概のものは飲みます)</p>	Alcohol
Dialogue context	<p>S What kind of music do you listen to, Japanese, Western, or classical? (聞くとしたら、邦楽、洋楽、クラシックなどどんなジャンルを聞きますか?)</p> <p>U I often listen to Japanese music, but sometimes I also listen to classical music. (そうですねえ 聴く 邦楽が多いですがたまにクラシックも聴きますね)</p> <p>S What genre of music do you like? (どんなジャンルの曲が好きなんですか?)</p> <p>U Yes, I like pop music. For classical music, I like Tchaikovsky. (そうですね ポップス系が好きですね あとクラシックだと チャイコフスキーなどが好きですね)</p> <p>S You like classical music? Is there a favorite composer or music that you often listen to? (クラシックが好きなんですか。よく聞く「作曲家」やきょくってありますか?)</p> <p>U I like Tchaikovsky's Violin Concerto No. 1 the best. (チャイコフスキーのヴァイオリン協奏曲第一番が一番好きですね)</p>	
Hazumi	S What do you like about that? (そのきょくのどういうところが好きなんですか?)	
Expanded	S Do you ever enjoy your favorite classical music while drinking alcohol ? (お好きなクラシックとお酒を、一緒に楽しめることも多いんですか?)	

Naturally associated user information with dialogue context

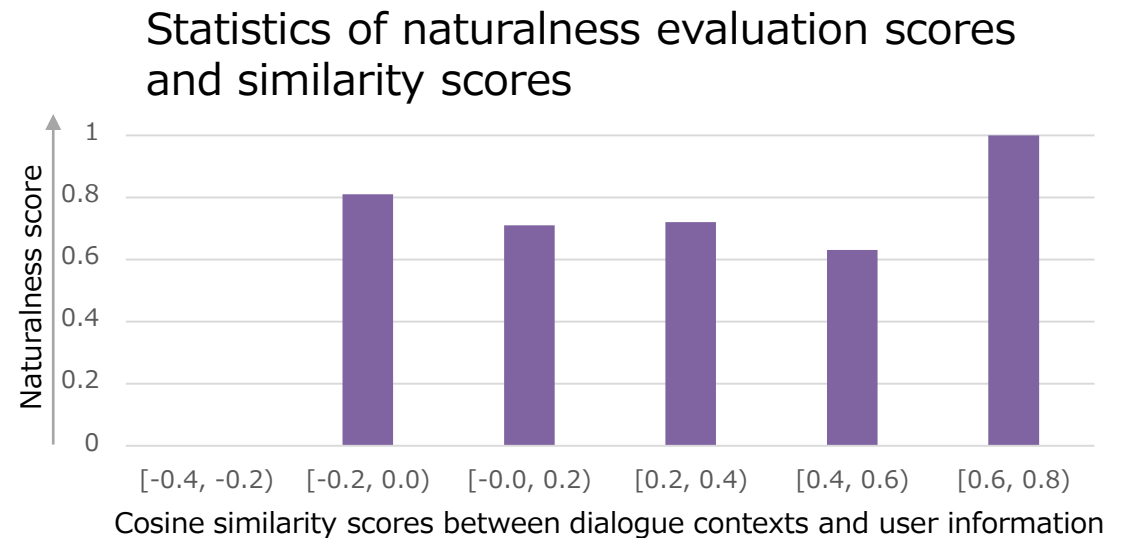
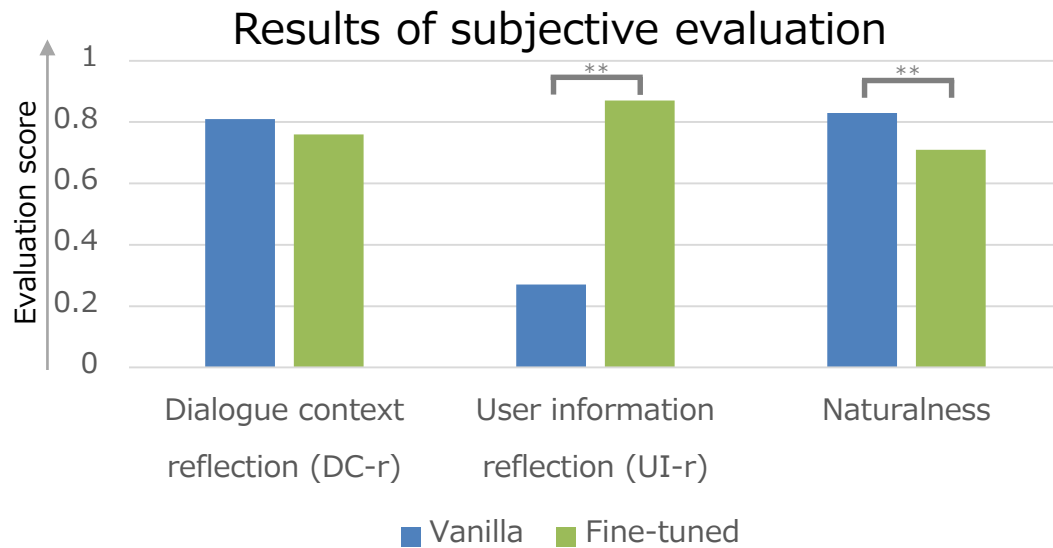
Evaluation-1

- Utterance generation models for comparison
 - Vanilla
 - ◆ Japanese Transformer encoder-decoder dialogue model trained with Twitter reply pairs [Sugiyama+'23]
 - ◆ The input format was “dialogue context”
 - Fine-tuned
 - ◆ Vanilla fine-tuned with the SUI corpus
 - ◆ The input format was “user information [SEP] dialogue context”
- Settings:

Crowdworkers evaluated a total of 200 utterances broken down into 100 system utterances generated by each of the models

Evaluation-1. Results and analysis

- Fine-tuned had higher UI-r score than Vanilla and comparable DC-r score with Vanilla
- Fine-tuned had slightly lower naturalness than Vanilla
- Utterance quality does not depend on the similarity between dialogue contexts and user information
- **Fine-tuned model could incorporate various user information into system utterances regardless of the degree of the similarity**



Evaluation-2

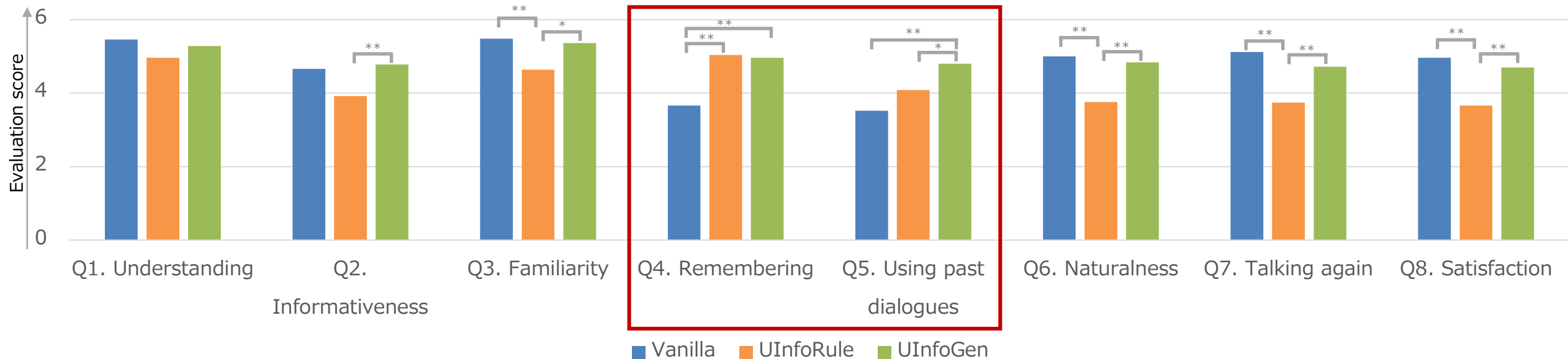
- Dialogue systems for comparison
 - Vanilla: Utterances were generated by the vanilla model [Sugiyama+'23], not using user information
 - UInfoRule [Tsunomori+'19]: Utterances were randomly generated by templates using user information with 30% and by Vanilla with 70%
 - UInfoGen [Ours]: Utterances were randomly generated by the fine-tuned model with 30% and by Vanilla with 70%
- Settings:

50 crowdworkers conducted dialogues in a text-chat interface with the systems

 1. Read a dialogue (user information) displayed in the interface as their own past dialogue with the system
 2. Conducted a dialogue with each system lasting 15 turns and evaluated the systems subjectively

Eval-2. Results and analysis

- UInfoRule had the low overall score
- UInfoGen and Vanilla had equally high scores except for Q4 and Q5
- UInfoGen had high scores for Q4 and Q5; **the fine-tuned model enabled a dialogue system to utilize various user information appropriately in dialogues with users**



Conclusion

- Summary
 - Constructed **the SUI corpus*** consisting of utterances incorporating various kinds of user information
 - Subjectively evaluated our fine-tuned model;
 - ◆ generated system utterances
 - ◆ dialogues with users
- Future work
 - Analyze the timing for effectively incorporating user information in dialogues
 - Apply the SUI corpus to large language models (LLMs) by few-shot learning methods

*<https://github.com/nu-dialogue/sui-corpus>

