

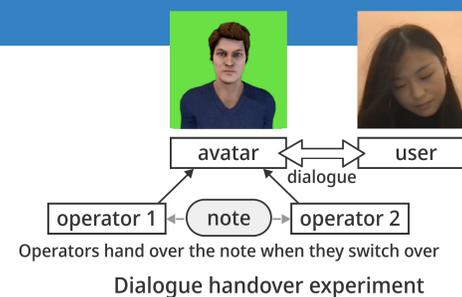
Data Collection for Empirically Determining the Necessary Information for Smooth Handover in Dialogue

Sanae Yamashita, Ryuichiro Higashinaka (Graduate School of Informatics, Nagoya University, Japan)



Overview

- Despite recent advances, dialogue systems still struggle to achieve fully autonomous transactions.
- When a system encounters a problem, human operators need to take over the dialogue to complete the transaction.
 - e.g., call center (Gorin et al., 1997; Walker et al., 2000), framework of parallel conversations with semi-autonomous avatars (Kawahara et al., 2021)
- However, it is unclear **what information should be presented to the operator** when the handover takes place.
- We investigated the information needed for handing over a dialogue.
 - We conducted a dialogue handover experiment in which one of two operators talked to a user and switched with the other operator periodically while exchanging notes.
- We found that **adjacency pairs**, which can express basic exchanges and include the comments and follow-up utterances of speakers were found to be useful. In task-oriented dialogues, **key-value pairs** were also found to be useful to grasp at a glance the information that has already been mentioned in tasks.



Dialogue Handover Experiment

We experimentally investigated the information needed for handing over a dialogue by dialogue handover experiment.

Dialogue

- Two operators act as one operator, swapping at regular intervals, and interact with one user on Zoom.
- Three dialogue tasks were conducted to identify the general elements for a handover of various types of dialogue.
 - Chat task: Operator listens to user's favorite things by asking questions and responding.
 - Consultation task: Operator consults with user on user's travel plan.
 - Sales task: Operator recommends one of three vacuum cleaners to user.

Operator switching

- Two operators **swap at every 2 minutes**.
- The users were not informed of the switching. To prevent the user from detecting the switching, the operators changed their voices and appeared as an avatar.

Note settings

- Each operator **takes notes and passes them on to the next operator every time they switched**.
- Notes are handwritten on A4 paper.

Operator's discussion

- In the discussion, the operators discussed ways of improving the note-taking process.
- In the interview, the operators described the handover situations in which they felt they had failed.

Questionnaire

- The operators and the users filled out questionnaires regarding naturalness and content understanding.
- The evaluations from the users were typically around four out of seven points, indicating that they were able to interact with the operator reasonably well even when the operators periodically switched.

Operator's interview

Dialogue settings

operators	4 (two pairs)
users	40
dialogues	60
dialogue duration	10 min.
operator switches	240

Statistics of dialogues

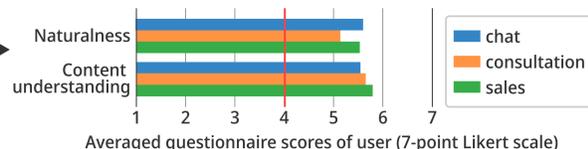
	chat		consultation		sales	
	operator	user	operator	user	operator	user
duration	3m57s	6m26s	6m19s	4m21s	6m13s	4m39s
utterances	86.1	71.3	82.9	85.5	86.1	87.6
utterance length	15.2	31.7	25.0	17.2	23.9	18.6

Statistics of notes

	chat		consultation		sales	
	operator	user	operator	user	operator	user
characters		208.4		115.3		120.5
words		105.4		57.8		63.6
arrows		151		116		115

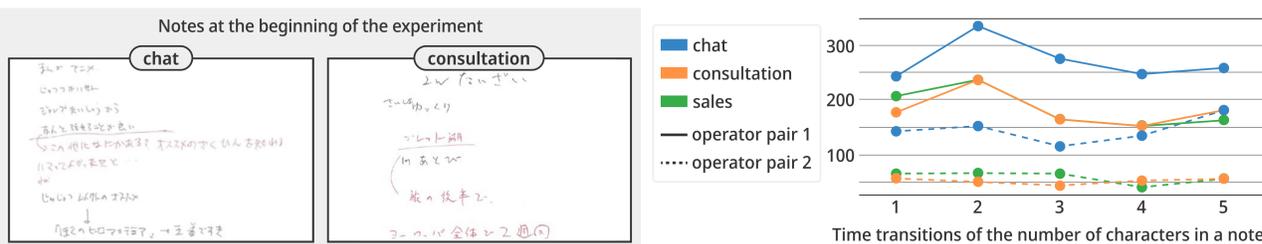
Example of chat dialogue

user	The name of its maker is Gibson.
operator 1	Oh, it's a really expensive one, yes.
user	I got excited and bought it.
operator 1	Ah, yes, it's a bit pricey, isn't it?
user	That's right.
operator 1	I see. Gibson's, what model is it?
-	<switching>
user	Well, I don't remember the model number, but it's a black.
operator 2	Black!
user	That's right.
operator 2	Gibson. Heh.



Analysis of the Notes

The number of characters and the **format in the notes roughly converged** for both operator pairs and dialogue tasks at the end of the experiment.



Notes at the beginning of the experiment

Notes at the end of the experiment

Adjacency Pairs

O: How much is the guitar?
U: One million yen.

O: The view is beautiful.
U: I'm looking forward to it.

O: It is expensive, but high-spec.
U: For example?

Adjacency Pairs

- Operator and user utterances were separated into left and right and connected by arrows.
- This format represents a **basic exchange**. (e.g., question-answer)
- This format can make it possible to track basic exchanges (who said what to whom) and avoid inconsistency of their own utterances.
- The notes included extended adjacency pairs:
 - pre/post-expansions
 - comment and follow-up
 - third pair parts (e.g., initiation-response-follow-up)

Three types of adjacency pairs

conventional adjacency pairs	50
extended adjacency pairs (containing comment and follow-up)	59
extended adjacency pairs (containing three or more parts)	49
total of adjacency pairs in all notes	158

Key-value Pairs

- This format represents **slot-value pairs in a particular domain**.
- This format can make it possible to grasp at a glance the information that has already been mentioned.

Analysis of Discussions and Interviews

- One operator stated that, after switching, she referred to only the most recent information because she did not immediately have time to read the whole note carefully. From the next turn onward, she went back through the notes to grasp the history of the note.
 - It is necessary for the operators to have two types of information: **most recent information** and **dialogue history**.
- One operator stated that it was difficult for two operators to behave as one unless they knew the personality of the other.
 - It is important for the operators to share the **same profile information** in order to keep the consistency of the dialogue.

Example of chat dialogue

operator 1	I think it's delicious, too. (about chocolate)
-	<switching>
operator 2	I'm not a fun of sweets.
	(conflicting self-disclosures against operator 1)
user	Oh, really? I was just thinking that you like sweet things, since you said earlier that chocolate is delicious.

Future Work

- We aim to implement these findings in an interface suitable for dialogue handover.
- Several underlying techniques will be necessary, such as extracting adjacency pairs and key-value pairs from the dialogue history.
- It may also be necessary to format the extracted adjacency pairs in a more concise way for understanding the content quickly.