

The Influence of User Tailoring and Processing Load on Performance in SDS Applications

Andi Winterboer & Vera Demberg

School of Informatics
University of Edinburgh
A.Winterboer@ed.ac.uk
V.Demberg@sms.ed.ac.uk



1 Introduction

Spoken dialogue systems (SDS) are systems developed to provide natural and efficient access to applications and services using speech as the main interaction mode (e.g., calendars, public transport information, flight booking). This poster presents the results of three experiments comparing two recently proposed information presentation strategies, and assesses the influence of cognitive processing load on the user through the SDS.

2 Information presentation strategies

Previously proposed: The **summarize and refine approach (SR)** [1,3] groups a large number of options into a small number of clusters that share attributes. The system summarizes the clusters and suggests additional constraints to the user.

User: I'd like to book a flight from Edinburgh to Brussels for Thursday, 27th July, arriving at about noon.
System: There are 28 flights from Edinburgh to Brussels. Most of them require a connection and cost from 100 to 230 £.
User: Are there any direct flights with availability in business class that arrive around 3 pm?
System: There are no direct business class flights that arrive around 3 pm. However, there is one direct business class flight that arrives at 5:30 pm. It costs 111 pounds.
User: No, that's too late... what about connecting business class flights that arrive around 3pm?
System: ...

Our new approach: In the **user-model based summarize and refine approach (UMSR)** [2], information from a user model is utilized to reduce dialogue duration by considering all options that are relevant to the user and by introducing a content structuring algorithm that supports stepwise refinement based on the ranking of attributes in the user model.

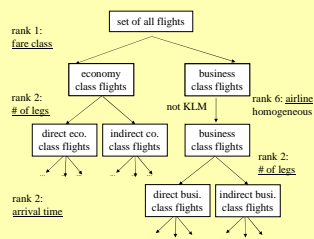


Figure 1: Option tree for business user

UMSR dialogue sample for business class user:

User: I'd like to book a flight from Edinburgh to Brussels for Thursday, 27th July, arriving in early afternoon.
System: There are four flights with availability in business class. None of them are on KLM. And the only direct flight arrives at 5:30 pm. To arrive earlier, you'll have to make a connection. If you're willing to travel economy, there is a direct flight on BMI, arriving at 1:10pm. There's also a KLM flight arriving at 1:50pm, but it requires a connection in Amsterdam. All other economy flights take longer or arrive later than 4 pm.

A new variant of the UMSR approach: Modified to produce more **concise** system responses

User: I'd like to book a flight from Edinburgh to Brussels for Thursday, 27th July, arriving in the early afternoon.
System: There are four flights with availability in business class, but none are on KLM. The only direct flight arrives at 5:30pm on Lufthansa. If you're willing to travel economy, there is a direct flight on BMI, arriving at 1:10 pm. There's also a flight on KLM arriving at 1:50pm, but it requires a connection in Amsterdam.

3 Evaluation + Results

Experiment 1: SR vs. UMSR

- participants overhear dialogues between a user and both the SR and the UMSR system
- questionnaire assessing user satisfaction on a 7 point Likert scale

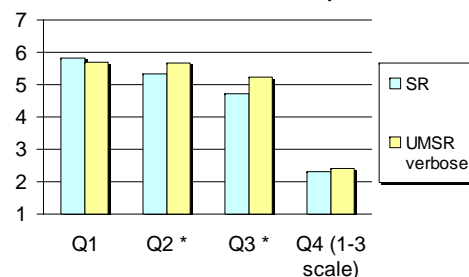
Results

- participants prefer the UMSR system over the SR system in a forced choice question
- UMSR system provides better overview and quicker access

Evaluation based on four questions:

- Q1: Did the system give the information in a way that was easy to understand?
Q2: Did the system give you a good overview of the available options?
Q3: Do you think there may be flights that are better options for the user that the system did not tell him/her about?
Q4: How quickly did the system allow the user to find the optimal flight?

Answers to the 4 evaluation questions

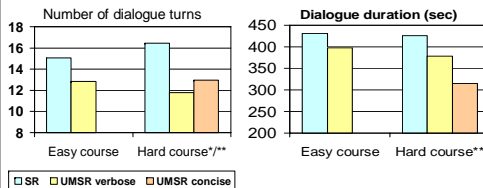


Experiment 2: Dual task SR vs. UMSR verbose

- drivers in easy vs. difficult driving conditions
- secondary task to place additional cognitive burden on participants

Results

- booking flights with system adopting the UMSR approach required significantly fewer dialogue turns and shorter dialogue duration
- no significant difference in terms of stop sign and traffic light violations, or number of speeding tickets.
- UMSR approach was more fun than SR in easy driving condition, reverse was true for difficult driving condition
- however, participants had significantly more minor driving errors (centerline crossings, road edge excursions, ...) with UMSR
- => long UMSR turns caused high cognitive load



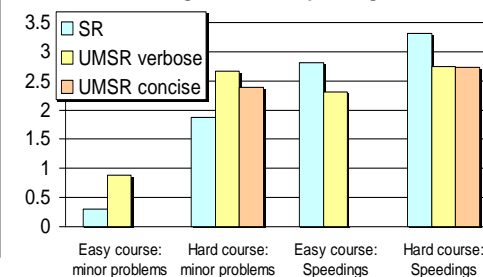
Experiment 3: Dual task with UMSR concise

- controlling for turn length and information density
- only on difficult driving course
- attempting to reduce cognitive load

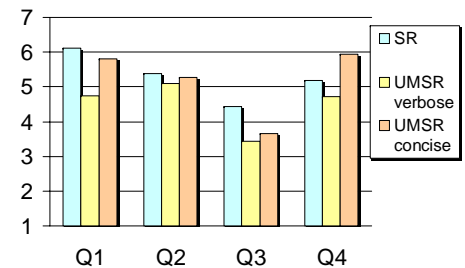
Results

- even shorter dialogue duration for UMSR
- fewer dialogue turns required with UMSR
- no significant difference in number of driving errors between systems

Driving errors and speedings



Answers to the 4 evaluation questions



4 Conclusion

Experiment 1: SR vs. UMSR

- listeners prefer UMSR over SR style of presentation in the single task experiment

Experiment 2: Dual task SR vs. UMSR verbose

- more efficient information retrieval with UMSR approach

- however, in difficult driving condition SR was preferred and caused fewer driving errors

Experiment 3: Dual task with UMSR concise

- due to concise presentations UMSR requires less cognitive load than in 1st experiment
- number of driving errors as well as dialogue duration reduced
- users found the system more pleasant to use and selected the "best" flight more often
- UMSR based messages enable more efficient information retrieval and do not negatively affect driving safety

The authors would like to thank Johanna Moore for helpful discussions and supervision.

References:

- Chung, G. (2004). Developing a flexible spoken dialog system using simulation. In Proceedings of ACL '04, Barcelona, Spain, pages 63-70.
- Demberg, V., and Moore, J.D. (2006) Information presentation in spoken dialogue systems. In Proceedings of EACL '06, Trento, Italy, pages 65-72.
- Polifroni, J., Chung, G., and Seneff, S. (2003). Towards automatic generation of mixed-initiative dialogue systems from web content. In Proceedings of Eurospeech'03, pages 193-196, Geneva, Switzerland. AAI Press.