Language Generation for Spoken Dialogue Systems

Johanna D. Moore
School of Informatics, University of Edinburgh
Edinburgh, United Kingdom
j.moore@ed.ac.uk

The goal of spoken dialogue systems (SDS) is to offer efficient and natural access to applications and services. A common task for SDS is to help users select a suitable option (e.g., flight, hotel, restaurant) from the set of options available. When the number of options is small, they can simply be presented sequentially. However, as the number of options increases, the system must have strategies for summarizing the options to enable the user to browse the option space. In this talk, we evaluate two recent approaches to information presentation in SDS: (1) the Refiner approach (Polifroni et al., 2003) which generates summaries by clustering the options to maximize coverage of the domain, and (2) the user-model based summarize and refine (UMSR) approach (Demberg and Moore, 2006) which clusters options to maximize utility with respect to a user model, and uses linguistic devices (e.g., discourse cues, adverbials) to highlight the trade-offs among the presented items.

To evaluate these strategies, we go beyond the typical "overhearer" evaluation methodology, in which participants read or listen to pre-prepared dialogues, which limits the evaluation criteria to users' perceptions (e.g., informativeness, ease of comprehension). Using a Wizard-of-Oz methodology to evaluate the approaches in an interactive setting, we show that in addition to being preferred by users, the UMSR approach is superior to the Refiner approach in terms of both task success and dialogue efficiency, even when the user is performing a demanding secondary task. Finally, we hypothesize that UMSR is more effective because it uses linguistic devices to highlight relations (e.g., trade-offs) between options and attributes. We report the results of two studies which show that the discourse cues in UMSR summaries help users compare different options and choose between options, even though they do not improve verbatim recall.

References

V. Demberg and J.D. Moore. 2006. Information Presentation in Spoken Dialogue Systems. In *Proceedings of the 11th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*.

J. Polifroni, G. Chung, and S. Seneff. 2003. Towards the Automatic Generation of Mixed-Initiative Dialogue Systems from Web Content. In *Proceedings of Eurospeech*, pages 2721–2724.

Proceedings of the 13th European Workshop on Natural Language Generation (ENLG), page 132,
Nancy, France, September 2011.