Talkin' bout a revolution (statistically speaking)

Oliver Lemon
Heriot-Watt University
Edinburgh, United Kingdom
`o.lemon@hw.ac.uk`

This talk will describe new methods for generating Natural Language in interactive systems – methods which are similar to planning approaches, but which use statistical machine learning to develop adaptive NLG components. Employing statistical models of users, generation contexts, and of Natural Languages themselves, has several potentially beneficial features: the ability to train models on real data, the availability of precise mathematical methods for optimisation, and the capacity to adapt robustly to previously unseen situations. Rather than emulating human behaviour in generation (which can be sub-optimal) these methods can even find strategies for NLG which improve upon human performance.

Recently, some encouraging results have been obtained with real users of 3 different systems developed using these methods, for the tasks of Information Presentation in an automated tourist guide, Referring Expression Generation in a technical support system, and generation of Temporal Referring Expressions in an appointment scheduling system. The results show that optimised NLG significantly outperforms related prior approaches, and can also improve the global performance of dialogue systems.

As well as explaining the core Reinforcement Learning and user modelling methods and concepts behind this work, I will also cover some recent work from other researchers which fits with this general perspective on NLG. Finally, I discuss some future directions for this research area, for example the issues of incremental generation and generation under uncertainty.

References

Srini Janarthanam and Oliver Lemon. 2010a. Learning to adapt to unknown users: Referring expression generation in spoken dialogue systems. In *Proceedings of ACL*.

Srinivasan Janarthanam and Oliver Lemon. 2010b. Adaptive referring expression generation in spoken dialogue systems: evaluation with real users. In *Proceedings of SIGDIAL*.

Srinivasan Janarthanam, Helen Hastie, Oliver Lemon, and Xingkun Liu. 2011. 'the day after the day after tomorrow?' a machine learning approach to adaptive temporal expression generation: training and evaluation with real users. In *Proceedings of SIGDIAL*.

Oliver Lemon, Srini Janarthanam, and Verena Rieser. 2010. Generation under uncertainty. In *Proceedings of the Generation Challenges Session at INLG*.

Oliver Lemon. 2011. Learning what to say and how to say it: joint optimization of spoken dialogue management and natural language generation. *Computer Speech and Language*, 25(2).

Verena Rieser and Oliver Lemon. 2009. Natural language generation as planning under uncertainty for spoken dialogue systems. In *Proceedings of EACL*.

Verena Rieser, Oliver Lemon, and Xingkun Liu. 2010. Optimising information presentation for spoken dialogue systems. In *Proceedings of ACL*.