Model Learning for Dialog Management

Finale P Doshi

Abstract: Spoken dialog managers can allow for natural human-robot interaction, but noisy voice recognition and linguistic ambiguities make it difficult to decipher the user's intent. Partially Observable Markov Decision Processes (POMDPs) have succeeded in dialog management applications because they are robust to uncertainties in the dialog [1]. However, POMDPs are generally specified using large probabilistic models with many parameters that are difficult to specify from domain knowledge. Gathering enough data to accurately determine the parameters apriori is also expensive.

Drawing on Bayesian techniques to learning in MDPs [2], we present continuing work on approaches to learn the dialog POMDP parameters online. The parameter statistics provide a compact representation from which we can derive and refine policies. Although sub-optimal, these intermediate policies can reduce the number of poor decisions made during the learning process--a factor that is crucial in domains such as our robotic wheelchair.

We first use explicit rewards to learn the problem parameters over time and demonstrate a heuristic that allows the dialog manager to intelligently replan its policy given data from recent interactions. The drawback to this approach is that the agent must still experience the consequences of an action to discover the cost of a poor decision. Thus, we explore the use of meta-actions--queries about actions the dialog manager should have taken--as an implicit means of learning the user model.

References

[1] J. Williams and S. Young (2005). "Scaling up POMDPs for Dialogue Management: the Summary POMDP Method." IEEE workshop on Automatic Speech Recognition and Understanding (ASRU2005), Cancun, Mexico.

[2] R. Dearden, N. Friedman, and D. Andre (1999). "Model based Bayesian Exploration." Proceedings of Fifteenth Conference on Uncertainty in Artificial Intelligence. San Francisco: Morgan Kaufmann.