Abstract

Spoken dialog systems have been widely used across many domains. For example, voice applications are popular these days in environments such as smart phones or cars. Such speech systems are built using the developers' understanding of the application domain and of the potential users in the field. This understanding may be driven by observations collected from a sampled population at a given time. However, the deployed models may not perfectly fit the real-life usage or may no longer be valid with the dynamics of the domain/users over time. Therefore, an agent which automatically adapts to the domain and users after deployment is intuitively desired. In this thesis, we focus on realistic problems in human-machine communication via natural language where adaptation can contribute to the quality of interaction. We mainly focus on speech understanding in a spoken dialog system, in order for the agent to figure out meaning from speech input. To do so, the system needs to recognize sentences (sequences of words) and interpret, among other semantics, user's intention. We discuss these two aspects in details. In short, our system can 1) adjust its vocabulary to improve speech recognition and understanding performance; 2) infer a user's high-level intention to better assist the user in the interaction. The former enables the system to accommodate user's (and domain's) language. The latter provides smooth and personalized interaction across multiple existing domains and enables the system to communicate at the task level, in addition to individual domain's level.