Title	
Author(s)	,
Citation	
Issue Date	2004-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/947
Rights	
Description	Supervisor: , ,

Japan Advanced Institute of Science and Technology

A Study on an Architecture for Spoken Dialogue Systems

Kohji Dohsaka

School of Information Science,
Japan Advanced Institute of Science and Technology

March 24, 2004

Abstract

In this thesis, we discuss an architecture for spoken dialogue systems that achieves smooth conversation between a system and its users through appropriate cooperation among system components.

A spoken dialogue system is composed of utterance understanding, utterance generation, and dialogue control components. We discuss a system architecture with emphasis on how the components should cooperate with each other to achieve a natural and efficient dialogue. First, we present an utterance understanding method based on pragmatic constraints for identifying the referents of zero-pronouns. Through cooperation between the utterance understanding and generation components sharing contextual information, the proposed method not only understands user utterances but also serves for generating relevant system utterances under the context, and it enables a system to achieve a natural dialogue.

Second, we present an incremental utterance production method that allows a system to generate utterances while handling user responses. In spoken dialogue, dialogue participants use utterance units that are smaller than sentences in written language. Therefore, conventional methods utilizing sentences as utterance units cannot rapidly handle user responses during system utterances. The proposed method generates utterances incrementally in small utterance units specific to spoken dialogues. The utterance generation component cooperates with the dialogue control and utterance understanding components based on collaborative principles derived from an analysis of human-human dialogues. Through this cooperation, a system can generate natural utterances while handling user responses.

Third, we present dialogue control methods called dual-cost and trial dual-cost methods for a system to convey the desired information to the user in as short a dialogue as possible. The dialogue length is estimated as the sum of the confirmation cost depending on the speech recognition rate and the information transfer cost depending on the system database content. By choosing a system action at each point of a dialogue so as to minimize the total dialogue length through cooperation among the dialogue control, utterance understanding, and utterance generation components, the proposed methods achieve a more efficient dialogue than conventional ones. The results of experiments by dialogue simulation and with human users prove the effectiveness of the proposed methods.

Key Words: **spoken dialogue, utterance generation, dialogue control, dialogue management, utterance understanding**