Multi-domain Spoken Dialog System for Information Access in Mobile Environments

Junhwi Choi , Kyungduk Kim, Seokhwan Kim, Donghyeon Lee, Injae Lee, Gary Geunbae Lee

Department of Computer Sceience and Engneering, POSTECH
San 31.Hyoja-dong.Nam-gu.Pohang.Gyungbuk.Korea

{chasunee, getta, megaup, semko, lij1984, gblee}@postech.ac.kr

***Abstract*— In this demonstration, we present a multi-domain spoken dialog system for information access in mobile environments. The main technical contributions of our system are summarized as follows: (1) the system can cover user requests across multiple domains; (2) the system can handle misunderstanding errors by dynamic threshold confirmation modeling; and (3) the system is developed using an unsupervised spoken language understanding technique. We developed the system in Korean for four domains: smart TV program guide, weather, restaurant finder, and song finder.**

I. Introduction

Recently, there are many spoken dialog systems developed for single domain, and most of the conventional system is composed of five major components: automatic speech recognition (ASR), spoken language understanding (SLU), dialog management (DM), natural language generation (NLG), and text-to-speech synthesis (TTS). However, there are not many systems that can handle multiple-domain dialogs simultaneously.

In this demo, we present a dialog system to cover multiple domains by employing domain spotting technique in two-level SLU. Our system also utilizes an unsupervised technique for domain dependent SLU and employs a confirmation modeling strategy for mis-understanding handling and robustness of the system.

II. System Overview

Our system is developed by integrating ASR, SLU, DM, NLG, and TTS. Fig. 1 shows the overall architecture of our system.

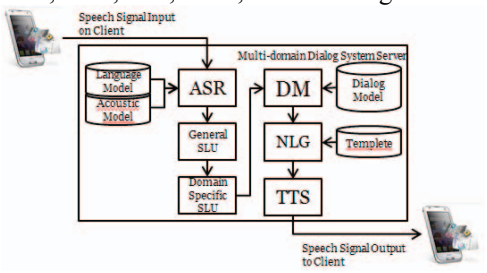

Fig. 1. Overall architecture of multi-domain dialog system on mobile environments

ASR is developed using a conventional hidden Markov model (HMM)-based continuous speech recognition engine. For SLU, we use a two-level approach, where, at first level, the result of ASR is tagged and classified by general SLU for domain spotting by employing a maximum entropy-based classifier using lexical word, dialog acts and previous domain as classification features. After domain spotting process, a domain specific SLU tagging is proceeded. To reduce human efforts, we developed an unsupervised approach based on non-parametric Bayesian HMM to domain-specific SLU training. In DM, we use an example-based dialog management (EBDM) method [1]. To handle mis-understanding errors, we model confirmation dialogs based on dynamic threshold and belief update [2]. A template-based language generation technique is used for NLG, and we use a concatenation TTS synthesizer. Constructing a multi-domain dialog model for our system, we collect dialog corpus by Wizard-of-Oz simulation technique.

Our system is developed for Korean dialogs simultaneously covering four different domains: smart TV program guide domain, weather domain, restaurant finder domain, and song finder domain.

III. System Implementation

Our system was fully implemented using a Java for Android client and C++ for Microsoft Windows server. The snapshot of our system running on a mobile phone is shown in Fig. 2.

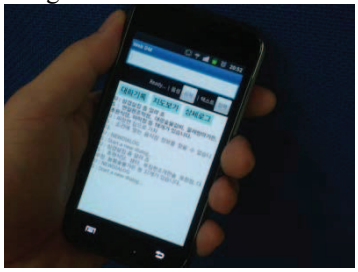

Fig. 2. Multi-domain spoken dialog system client on a mobile phone

Acknowledgment

This demo work was supported by the Ministry of Knowledge Economy, Korea, under QoLT development program, 10036458, and also under ITRC support program supervised by the National IT Industry Promotion Agency (NIPA-2011-C1090-1311-0009).

References

1. Cheongjae Lee, Sangkeun Jung, Seokhwan Kim, and Gary Geunbae Lee, "Example-based dialog modeling for practical multi-domain dialog system," in *Speech Communication*, May 2009, Volume 51, Issue 5, pp. 466–484.
2. Kyungduk Kim, Cheongjae Lee, Donghyun Lee, Junhwi Choi, Sangkeun Jung and Gary Geunbae Lee, "Modeling Confirmations for Example-based Dialog Management," in *Proc 2010 IEEE Workshop on Spoken Language Technology (SLT 2010)*, Berkeley, December 2010.