

( 続紙 1 )

京都大学	博士 (情報学)	氏名	神田 直之
論文題目	Open-ended Spoken Language Technology: Studies on Spoken Dialogue Systems and Spoken Document Retrieval Systems (拡張可能な音声言語技術 : 音声対話システムと音声文書検索システムにおける研究)		
(論文内容の要旨)			
<p><i>This thesis presents an open-ended spoken language technology for enabling broader applications of spoken language processing: Such applications should handle a wide range of utterances and provide a means to extend their knowledge without lots of efforts and time for improving human computer interactions. The thesis consists of seven chapters.</i></p> <p><i>Chapter 1 introduces the thesis. It begins with the background, motivation, and research objective for open-ended spoken dialogue technologies. From the viewpoint of enabling technology, the thesis focuses on two target systems, spoken dialogue systems as a man-machine interface and spoken document retrieval systems as an information source, and specifies two research problems; (1) spoken language analysis with little or no domain knowledge, and (2) integration of knowledge on arbitrary number of domains.</i></p>			

Chapter 2 reviews the literature related to spoken dialogue systems and spoken document retrieval systems. Conventional systems still depend on limited domain knowledge and can understand what users say to a severely limited extent. Thus, there remains a wide gap between what the current system can understand and what users expect it to do. In other words, the above two research problems are critical and unsolved issues in the literature.

Chapter 3 presents a solution to Problem (1) in spoken dialogue systems. A robust language understanding method is proposed for dialogue systems on a database search task. General context models for the database search task are constructed and the context information obtained from the models is incorporated into the language understanding method. Experiments with an actual spoken dialogue system demonstrate that the resulting language understanding method performs with high accuracy when one switches a database.

Chapter 4 describes a solution to Problem (2) in spoken dialogue systems. It addresses a spoken dialogue system that consists of multiple

subsystems on different domains, and can be easily extended for new domain knowledge. The main issue of the system is a domain selection method that selects an appropriate domain for each user utterance. The key idea of the thesis is to design a meta-level method instead of designing individual domain selection methods; a new domain selection meta-level method is to check whether the previously selected domain should be kept or not. Experimental results indicate that the proposed domain selection method achieved better results than conventional methods with providing the means to extend new domain knowledge.

Chapter 5 describes a solution to Problem (1) in spoken document retrieval systems. It addresses a spoken document retrieval system with new indexing and searching methods that can detect the positions of arbitrary keywords in speech database. To realize fast and accurate term detection, three types of the open vocabulary indexing method are tandemly combined. Experimental results indicate that the proposed method clearly outperforms conventional spoken term detection methods, especially when searching for out-of-vocabulary keywords.

Chapter 6 describes a solution to Problem (2) in spoken document retrieval systems. It addresses an index combination method for spoken term detection systems. Simply combining multiple indexes increases the index size, which poses problems of a high storage cost and slow search speed. To suppress the increase of the index size, a selective index combination method is developed with the out-of-vocabulary region classifier that returns the score of unknown term. The method is confirmed to be capable of suppressing the increase of the index size with gaining improvements of the search accuracy.

Chapter 7 concludes the thesis with a discussion about the contributions of our study and future directions.

注) 論文内容の要旨と論文審査の結果の要旨は1頁を38字×36行で作成し、合わせて、3,000字を標準とすること。

論文内容の要旨を英語で記入する場合は、400～1,100 wordsで作成し

審査結果の要旨は日本語500～2,000字程度で作成すること。

(論文審査の結果の要旨)

本論文は、人にとって最も自然なコミュニケーション手段の1つである音声を活用した応用システムの開発を促進する実現技術の開発を目的とし、音声言語認識技術の性能を支配するドメイン（知識領域）に注目し、複数ドメインの切り替えや新規ドメインの自由な追加が可能な音声言語処理技術に関する研究をまとめたものである。得られた主な成果は次の通りである。

1. 拡張可能な音声言語処理技術の課題として、(1) 新規ドメインだけでなく未知ドメインに対して頑健に動作する音声言語の解析技術、(2) 新規ドメインの自由な追加が可能なシステム統合技術、という2つの課題を指摘し、想定される応用として、音声対話システムと音声文書検索の2つを取り上げ、4通りの組み合わせに対して、各々の問題点を解決する手法を開発し、具体的にシステムを構築し、提案手法の有効性を実証している。これを通じて、拡張可能な音声言語処理技術の可能性を示したこと。

2. 新規ドメインの自由な追加のために、既存ドメインと新規ドメインとの判定手続きを作成するのではなく、処理中の現在のドメインに留まるかどうかを判定するメタレベルの機構を開発し、新規ドメインに発話が属するかを判定する手続きの追加だけで済むようにし、拡張性の高い実現法を開発したこと。

3. 音声文書検索で、複数システムを統合すると検索性能が向上する一方でインデックスサイズの増大と検索速度の低下を招くというトレードオフに対して、未知語領域検出技術を開発し、得られた未知語スコアをもとに複数のインデックスを選択的に統合する手法を考案し、トレードオフが解決できる可能性を評価実験により実証していること。

以上本論文は、拡張可能な音声言語処理技術に不可欠な手法を考案し、具体的なシステムを構築し、その有効性を実証したことに意義がある。本研究の成果は、音声言語処理の要素技術にとどまらず、音声インターフェースや音声文書処理への貢献が期待されており、学術上・實際上寄与するところが少なくない。よって、本論文は博士（情報学）の学位論文として価値あるものと認める。

また、平成26年2月24日実施した論文内容とそれに関連した口頭試問を行った結果合格と認めた。

注) 論文審査の結果の要旨の結句には、学位論文の審査についての認定を明記すること。  
更に、試問の結果の要旨（例えば「平成 年 月 日論文内容とそれに関連した口頭試問を行った結果合格と認めた。」）を付け加えること。

Webでの即日公開を希望しない場合は、以下に公開可能とする日付を記入すること。  
要旨公開可能日： 年 月 日以降