

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年 6月12日現在

機関番号：82636

研究種目：若手研究（B）

研究期間：2009～2011

課題番号：21700216

研究課題名（和文） 音声対話システムにおける対話コーパスの検索・適応に基づく応答生成

研究課題名（英文） Response Generation by Retrieval and Adaptation of Dialog Corpus for Spoken Dialog Systems

研究代表者

翠 輝久（MISU TERUHISA）

独立行政法人情報通信研究機構 ユニバーサルコミュニケーション研究所

音声コミュニケーション研究室

研究者番号：00531609

研究成果の概要（和文）：

一問一答処理では対応できない高度な音声対話（たとえば、誰かに何かを相談する対話）の実現を目指して、自然な対話制御および応答生成を行う手法について研究を行った。観光案内のプランニングを対象として、音声対話コーパスをもとに、システムを利用するユーザの対話モデル、また、要求に対して適切な回答を生成する手法を提案し、被験者実験により提案手法の有効性を確認した。

研究成果の概要（英文）：

To realize a complex spoken dialog (such as a consulting dialog) by computers, we investigated methods for natural spoken dialog management and response generation. A dialog model of consulting dialog and a method for natural language generation were proposed. The methods are implemented on a dialog system with sightseeing guidance domain, and then evaluated by user experiment.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	1,300,000	390,000	1,690,000
2010年度	1,200,000	360,000	1,560,000
2011年度	800,000	240,000	1,040,000
年度			
年度			
総計	3,300,000	990,000	4,290,000

研究分野：音声対話処理

科研費の分科・細目：情報学 知覚情報処理・知能ロボティクス

キーワード：音声対話、音声言語処理

1. 研究開始当初の背景

音声インタフェースは携帯電話などを用いて場所を問わずに利用可能なユビキタスな情報アクセスを提供する。また、我々が人同士のコミュニケーションにも用いるユニバーサルなインタフェースであるため、情報格差（デジタル・ディバイド）を埋める技

術として注目されている。

従来多くの研究で行われてきた音声対話システムは、システム開発者が人手によりアドホックに作成した規則に基づいて構築されることがほとんどであったが、人同士の自然な会話を基にしてボトムアップに対話戦略を学習することで、システム構築のコスト

を削減するのみならず、人にとってより自然な対話を行うことができると考えられる。そのためには本研究においては、音声対話システムにおいて、人間の対話をもとにシステムが次に発話する内容を決めたり（対話制御）、どのような発話文を用いるか（発話文生成）を決める手法を提案することを目指す。

2. 研究の目的

本研究では、音声を用いた対話を通じた相談を行うタスクに置いて、対話の状況やコンテキストに適応的な応答を生成する手法の確立のために、以下の3点の基礎的な技術に関する研究を行う。

(1) タスクに応じたユーザモデル構築

どのような対話コンテキストにおいて、どのように応答を行うべきかを決定する手がかりとして、対話中に起こりえる状況を記述する数理モデル（ユーザモデル）を提案する。また、このモデルに基づいて、ユーザが対話システムを相手に、どのように振る舞うかを予測するモデルを構築する。

(2) コーパスベースの応答生成

従来の応答生成では、人手により記述したルールを利用した応答生成がほとんどであった。本研究ではこれに対して、人同士の会話（対話コーパス）や、Webなどから対話の類似事例（応答文）を検索し、現在の対話の状況に合うように適応したうえでユーザに提示する手法を確立する。

(3) 対話状況に応じた音声合成出力技術

生成した発話文を自然な発話で読み上げるための音声対話を志向した音声合成エンジンの構築を行う。

さらに、これらの手法を、我々が対象に研究を行っている観光案内プランニングタスクに適用する。手法を備えた音声対話システムの構築を行い、被験者実験により手法の有効性を確認する。

3. 研究の方法

(1) 2009 年度

研究の基礎となるデータベースの作成を行うとともに、コーパスの分析を行い、ユーザモデルの検討を行う。

① 音声対話コーパスの整備

情報通信研究機構で構築している音声対話コーパスに、ユーザの発話の意図（対話行為）の付与を行い、ユーザモデルの構築のための基盤を作成する。また、対話行為記述に用いるタグセットの構築を行う。

② 対話コーパスの分析

発話の対話の構成に影響を与えている箇所をコーパスから抽出する。特に、後続の対話に影響を与えている箇所を調べて、

対話への影響の大きさを調べる。

(2) 2010 年度

2009 年度作成したコーパスを利用して以下の2点を中心に取り組む。

① 音声対話システムの対話状況のモデル化（ユーザモデル構築）

対話中における現在の状態（対話状態）とコーパス中の対話の状態類似度尺度を計算するために、具体的には、観光案内コーパス中における相談を行う対話を対象に、ユーザの嗜好や知識状態を基に対話状態を定義することを目指す。

② 発話表現の適応方法の検討

検索を行うことにより、現在の状況に完全にマッチした発話が存在する場合には、それをそのまま利用することができるが、コーパスのスパースネスを考えると、常に完全にマッチした発話が検索できるとは限らない。そのため、検索された類似状況の発話を現在の状況に合ったものに修正（適応）する。

(3) 2011 年度

前年度までに研究で開発した対話状態表現手法および検索に基づく応答生成手法を利用して、以下の2点に取り組む。

① システムの構築

前年度までに提案した、観光案内コーパス中における相談を行う対話を対象とした、ユーザの嗜好や知識状態に基づいた対話表現を利用した音声対話システムを構築する。

② 被験者実験による手法の有効性の検証

構築したシステムの被験者実験を行い、本研究で構築した提案手法の有効性を検証する。

4. 研究成果

(1) アノテーション付き音声対話コーパスの構築

京都観光案内対話コーパスに対して、発話中のコンテンツや言い回しや韻律などによって表出される話し手の態度を記述するタグのスキームを作成し、作業員によるアノテーションを行った。具体的には、コーパスに対して、ドメインに依存しない発話の意図を記述するための発話行為タグと、ドメイン・タスク依存に発話が何について言及しているかを記述するための意味内容タグを作成し、コーパスに対してアノテーションを行った。さらに、アノテーションを行った付けられたコーパスに対して、観光地を決定する意思決定が行われている区間（＝エピソード）の付与を行い、エピソード区間中で対話制御を行うためにどのようなタグが重要な箇所になっ

ているかを分析した。その結果、本研究が対象としている相談型の音声対話では、ガイド(システム側)がユーザの嗜好を対話的に推定し、推定されたユーザの嗜好に応じて観光スポットを推薦し、ユーザが推薦された結果に基づいて決定を行うという流れがあることが分かった。

(2) 相談対話に置けるユーザモデル(対話状態表現)の提案

観光案内コーパス中における相談を行う対話を対象に、ユーザの嗜好および知識状態を基にした対話状態の記述表現を提案した。提案モデルは、対話中のユーザの知識と嗜好に着目したものである。これらは、システム側から直接観測できないパラメータであり、この不確実性を表現するために部分観測マルコフ決定過程(POMDP)を用いた対話のモデル化を行った。また、実際の対話データに基づいてパラメータの情報を更新する手法を提案した。さらに、実際の対話中のインタラクションから、これらのパラメータを推定できることを確認した。

(3) テキスト検索に基づく応答生成

人手により対話システムの応答を記述するコストを削減するために、対話の状況にあった応答を自動で構築する手法を提案した。現在の状況に類似したテキストをWebから取得することで、応答文としてふさわしい文を検索できることを確認した。その際に、テキストがドメインとあっているかだけでなく、文の焦点が何であるかに注目して選択することが重要であることが分かった。

(4) 音声対話用の合成エンジンの構築

機構の音声合成グループと共同で、音声対話用の合成エンジンの構築を行った。

(5) システムへの実装と評価

提案した手法を、実際の音声対話システムに実装し、被験者実験により評価を行った。合計のべ100名以上の被験者により評価を行った。手法を用いないベースライン対話システムと比較して、システムがよりよい応答生成を行えることを確認した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計2件;すべて査読あり)

- (1) 翠輝久, 水上悦雄, 志賀芳則, 川本真一, 河井恒, 中村哲.
ユーザの相槌・頷きを喚起する音声対話

システム. 電子情報通信学会論文誌, Vol. J95-A, No.1, pp.16--26, 2012.

- (2) Teruhisa Misu, Komei Sugiura, Tatsuya Kawahara, Kiyonori Ohtake, Chiori Hori, Hideki Kashioka, Hisashi Kawai and Satoshi Nakamura.

"Modeling Spoken Decision Support Dialogue and Optimization of its Dialogue Strategy" ACM Transactions on Speech and Language Processing, Vol. 7, Issue 3, pp.10:1--10:18, 2011.

[学会発表](計19件)

- (1) Teruhisa Misu, Etsuo Mizukami, Hideki Kashioka, Satoshi Nakamura and Haizhou Li. "A Bootstrapping Approach for SLU Portability to a New Language by Inducting Unannotated User Queries." In Proc. IEEE-ICASSP, 2012年3月28日, 国立京都国際会館(京都府)

- (2) Teruhisa Misu, Etsuo Mizukami, Yoshinori Shiga, Shinichi Kawamoto, Hisashi Kawai and Satoshi Nakamura.

"Analysis on Effects of Text-to-Speech and Avatar Agent in Evoking Users' Spontaneous Listener's Reactions" In Proc. IWSDS, pp.73--85, 2011年9月2日, Saray hotel in Granada (スペイン)

- (3) Teruhisa Misu, Kiyonori Ohtake, Chiori Hori, Hisashi Kawai and Satoshi Nakamura. "User Study of Spoken Decision Support System." In Proc. INTERSPEECH, pp.797--800, 2011年8月30日, Firenze Fiera venue (イタリア)

- (4) Teruhisa Misu, Etsuo Mizukami, Yoshinori Shiga, Shinichi Kawamoto, Hisashi Kawai and Satoshi Nakamura

"Toward Construction of Spoken Dialogue System that Evokes Users' Spontaneous Backchannels" In Proc. SIGDIAL, pp.259--265, 2011年7月17日, Oregon Health & Science University (アメリカ)

- (5) Teruhisa Misu, Komei Sugiura, Kiyonori Ohtake, Chiori Hori, Hideki Kashioka, Hisashi Kawai and Satoshi Nakamura.

"Dialogue Strategy Optimization to Assist User's Decision for Spoken Consulting Dialogue Systems" In Proc. IEEE-SLT, pp.342--347, 2010年12月14日, The Claremont Hotel Club & Spa (アメリカ)

- (6) Teruhisa Misu, Kiyonori Ohtake, Chiori Hori, Hideki Kashioka, Hisashi Kawai and Satoshi Nakamura.

“Web Text Classification for Response Generation in Spoken Decision Support Dialogue Systems”

In Proc. 4th International Universal Communication Symposium (IUCS), pp.130--133, 2010年10月19日, Crowne Plaza Beijing Zhongguancun (中国)

- (7) Teruhisa Misu, Kiyonori Ohtake, Chiori Hori, Hideki Kashioka, Hisashi Kawai and Satoshi Nakamura

“Construction and Experiment of a Spoken Consulting Dialogue System”

In Proc. IWSDS, pp.169--175, 2010年10月2日, Gotemba Kogen Resort (静岡県)

- (8) Teruhisa Misu, Komei Sugiura, Kiyonori Ohtake, Chiori Hori, Hideki Kashioka, Hisashi Kawai and Satoshi Nakamura.

“Modeling Spoken Decision Making Dialogue and Optimization of its Dialogue Strategy”

In Proc. SIGDIAL, pp.221--224, 2010年9月25日, 東京大学(東京都)

- (9) Teruhisa Misu, Komei Sugiura, Tatsuya Kawahara, Kiyonori Ohtake, Chiori Hori, Hideki Kashioka and Satoshi Nakamura

“Online Learning of Bayes Risk-Based Optimization of Dialogue Management for Document Retrieval Systems with Speech Interface”

In Proc. IWSDS, 2009年12月11日, Kloster Irsee (ドイツ)

- (10) Teruhisa Misu, Kiyonori Ohtake, Chiori Hori, Hideki Kashioka and Satoshi Nakamura

“Annotating Communicative Function and Semantic Content in Dialogue Act for Construction of Consulting Dialogue Systems.”

In Proc. INTERSPEECH, pp.1843--1846, 2009年9月9日, Brighton Centre (イギリス)

[図書] (計1件)

- (1) Teruhisa Misu, Komei Sugiura, Tatsuya Kawahara, Kiyonori Ohtake, Chiori Hori, Hideki Kashioka and Satoshi Nakamura, “Online learning of Bayes Risk-based Optimization of Dialogue Management for Document Retrieval Systems with Speech Interface (Chapter 2)” in “Spoken Dialogue Systems Technology and Design”, Springer, pp. 29--62, 2010年

[産業財産権]

○出願状況 (計2件)

名称: 対話装置、対話方法、およびプログラム

発明者: 翠 輝久

権利者: 情報通信研究機構

種類: 特許

番号: 特許 2010-119950

出願年月日: 平成22年 5月26日

国内外の別: 国内

名称: 学習システム、シミュレーション装置、および学習方法

発明者: 翠 輝久

権利者: 情報通信研究機構

種類: 特許

番号: 特許 2010-157623

出願年月日: 平成22年7月12日

国内外の別: 国内

[その他]

ホームページ等

<http://www2.nict.go.jp/univ-com/slc/member/xtmisu/index.html>

6. 研究組織

(1) 研究代表者

翠 輝久 (MISU TERUHISA)

(独)情報通信研究機構・ユニバーサルコミュニケーション研究所・音声コミュニケーション研究室

研究者番号: 00531609

(2) 研究分担者

研究者番号: ()

(3) 連携研究者

研究者番号: ()