

音声案内システムにおける質問応答データベースの最適化

— 音声認識結果を用いた質問応答データベースによる応答生成と最適化 —

竹内 翔大[†] ツインツアレク トビアス[†] 川波 弘道[†] 猿渡 洋[†] 鹿野 清宏[†]

[†] 奈良先端科学技術大学院大学 情報科学研究科 〒 630-0192 奈良県生駒市高山町 8916-5

E-mail: †{shota-t,cincar-t,kawanami,sawatari,shikano}@is.naist.jp

あらまし 実環境で頑健に動作する音声対話システムには多彩な発話表現を柔軟に処理できる対話制御技術が必要となる。多彩な発話表現に対して頑健な応答生成手法として、用例ベースの応答生成が研究されている。この手法では質問応答データベース (QADB) を用いて入力音声の認識結果に最も近い質問例を選択し、入力に対して最も適切な応答を生成する。これまで、QADB は発話データの書き起こし文を用いて作成されていた。本稿では、音声認識結果を用いた QADB (認識結果 QADB) による応答生成とその最適化を提案する。この手法により、入力に含まれる音声認識誤りに頑健な用例選択が行える。システム応答の適切さを応答正解率で評価した結果、6 候補程度の認識結果を用いて QADB を作成し、最適化を行うことで、従来と同等以上の応答正解率が得られた。

キーワード 音声対話システム, 用例ベースの応答生成, 質問応答データベースの最適化, Leave-one-out 交差検定法, 認識結果 QADB

Optimization of Question and Answer Database for Speech-oriented Guidance System

— Response Generation Based on Question and Answer Database Using Automatic Speech Recognition Results and Its Optimization —

Shota TAKEUCHI[†], Tobias CINCAREK[†], Hiromichi KAWANAMI[†], Hiroshi SARUWATARI[†],
and Kiyohiro SHIKANO[†]

[†] Graduate School of Information Science, Nara Institute of Science and Technology, 8916-5 Takayama-cho, Ikoma-shi, Nara-ken, 630-0192, Japan

E-mail: †{shota-t,cincar-t,kawanami,sawatari,shikano}@is.naist.jp

Abstract A speech-oriented dialog system employed in real-environment requires dialog control techniques which enable flexible processing of various utterance expressions. As a robust response generation method for various utterance expressions, an example-based response generation method has been studied. This method employs a question and answer database (QADB) to generate the most appropriate response by selecting an example question which is nearest to an input. Conventionally, a QADB is constructed from transcriptions of utterance data. In this paper, we propose response generation based on a question and answer database using automated speech recognition results (ASR-QADB), which enables to perform robust selection of examples against recognition errors. Evaluating the performance of system responses by response accuracy, the same response accuracy as with the conventional method is achieved by constructing and optimizing ASR-QADB using 6-best recognition results.

Key words speech-oriented dialog system, example-based response generation, QADB optimization, leave-one-out cross-validation, QADB using automated speech recognition results

1. はじめに

音声対話システムとは、音声を用いて機械が人とコミュニケーションを図るためのシステムである。たとえば音声入力を受け付ける案内システムでは、ユーザはシステムに話しかけるだけで所望の案内情報を得ることができる [1] [2] [3]。このように、音声対話システムでは発話というごく簡単な操作で機械に指示や要求を伝えることができる。音声対話システムを実現するためには音声認識技術だけではなく、認識された音声に対して適切な応答を提示するための対話処理技術が欠かせない。

実環境下で音声対話システムを用いると、実験室環境下では生じにくい様々な問題が生じる。入力音声への環境雑音や背景会話の混入だけではなく、基本的に不特定ユーザによる利用が前提であるため、未知話者の入力による音声認識精度の低下が常に生じる。前者に対しては音声信号処理による雑音除去、後者に対しては不特定話者モデルの使用といった技術が知られている。これらの技術は音声認識の精度を高めるためのものだとはいえる。

さらに入力音声に関する問題だけが実用上の音声対話システムで生じる問題ではなく、システムに対して多種多様の発話表現がなされることも重要な問題である。たとえば天気予報を聞く表現として、「明日の天気は？」、「天気予報を教えてください」、「晴れますか？」のようにいくつもの表現が考えられる。システムの初心者話しかけ方が分からないという事に加え、ベテランユーザであっても常に同一の表現を行うわけではないためである。

このように、実環境下での音声対話システムには多彩な発話表現を的確に処理する対話処理技術が求められる。実環境下で頑健に動作する対話処理技術の開発のため、本研究では対話の基本単位となる一問一答の性能を高めることを目標としている。

一問一答での応答生成方法として、用例ベースの応答生成技術が知られている [3] [4]。これは実際の質問例のデータベース（質問応答データベース、QADB）を用いて応答生成を行う手法で、意味解析による応答生成よりも簡便な処理で応答を行うことができる。用例ベースの応答生成ではデータベースを拡張するほど応答の適切さが向上するが、一方では誤ったデータの混入に対して比較的脆弱である。そのため、誤ったデータを除去することで応答の適切さを向上させる、質問応答データベースの最適化手法が提案されている。[5] [6]

従来の用例ベースの応答生成では、質問例に書き起こし文を用いて質問応答データベースを作成していた。しかし実用上の音声対話システムでは、入力音声に対して誤りのない音声認識結果が常に得られるとは限らない。

システム応答の適切さを向上させる方法として音声認識の性能を向上させることが考えられるが、本稿では音声認識結果が誤りを持つ状況での対話システムの性能向上を検討する。誤りを含む入力に対して頑健な応答生成を行うため、入力と同じ傾向の誤りを含む質問例を用いた応答生成、具体的には質問例として学習発話データの音声認識結果を用いた質問応答データベース（認識結果 QADB）による応答生成を検討した。

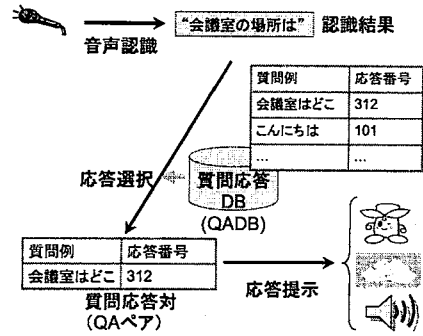


図 1 用例ベースの応答生成の模式図。

Fig. 1 An illustration of the example-based response generation.

2. 音声案内システム「たけまるくん」

本研究室では音声対話システム「たけまるくん」[1]を実環境下で運用し、得られた自然発話データや知見を基に、実環境で頑健に動作する音声対話システムの研究を行っている。このシステムは一問一答の対話戦略に沿って、入力音声に対して 300 以上の固定応答のうち最も適切なものを一つ応答することで音声案内を行う。応答はテキスト合成音声のほか、地図などの画像情報やキャラクター「たけまるくん」のアニメーションを提示する。

主なタスクは設置されている市民センターの案内だが、他に簡単なあいさつや「たけまるくん」のプロフィールを答えることができる。キャラクター性を付与することで対話システムはユーザにとって親しみのあるものとなる。一方でこのシステムは子供の利用が多いことが報告されており [1]、また子供音声に対する応答の精度は大人に対するものよりも低い [5] ため、子供音声への応答の精度を高めることが音声対話システムの運用上では重要な課題といえる。

「たけまるくん」では 2002 年 11 月の運用開始時から現在まで、研究目的でユーザの自由発話を収録している。そのうち 2 年分の全収録音声について、人手による発話内容の書き起こし文や適切な応答（正解応答）などの発話情報をもった音声データベースが構築されている [7]。本研究ではこのデータベースのうち有効音声（雑音ではない、システムへの入力を意図した音声）から質問応答データベースを作成している。なお、正解応答は固定応答のうちただ一つに限定している。

3. 用例ベースの応答生成

3.1 原理

用例ベースの応答生成では質問応答データベースを用い、入力に最も近い質問例から応答を生成する。具体的には、あらかじめ質問例と正解応答の対を QA ペアとしてデータベース化しておき、入力が生じたときにはデータベースから類似した質問例を選択し、対応する正解応答を提示する (図 1)。

3.2 質問例の選択

質問例の選択方法としてはk-近傍法がよく知られている。k-近傍法では、入力に類似した質問例を複数集めてシステム応答を多数決などで決定する。本研究では簡便な最近傍法を用いている。つまり、入力に対して最も類似した質問例一例のみを参照し、対応する応答一つをシステムの応答として決定する。

3.3 形態素マッチングによる類似度

質問例と入力との類似度の計算方法は用例ベースの応答生成にとって重要な要素である。本研究では質問例・入力音声認識結果間で形態素単位（単語単位）のマッチングを取ることで類似度を算出している。

一般的なベクトル空間モデルでは単語ベクトルとしてモデル化されたデータサンプル間の特徴の類似度を測るため、ベクトル間の「角度」の尺度であるコサイン距離を用い、またベクトルの特徴に寄与しない不要語の除去を行う。対して用例ベースの応答生成では実際のデータである質問例を用いるため、入力と質問例との「距離」を直接測ることで、入力に類似した質問例を精度良く選択できる。この方針に基づき、本研究や先行研究[8]では入力と質問例の2点間の距離に関する尺度として形態素マッチングによる類似度（5.2節で説明）を用い、また質問例から不要語を除去せずに用例選択を行っている。

3.4 応答正解率

手法の評価尺度として、応答正解率（response accuracy）という尺度を用いている。これはテストセットを音声対話システムに入力したとき、システムが正解することのできた発話の割合である。質問応答データベース Q 、応答生成手法 G が与えられたときのテストセット T の応答正解率 $r(T|Q, G)$ を次の式で与える。

$$r(T|Q, G) = \frac{T_C}{T_A} \quad (1)$$

ここで、 T_C は正解した発話の件数、 T_A はテストセットの総数を表す。応答正解率を自動的に求めるときは正解応答の付与された発話セットを用いる。

応答正解率はテストセット、質問応答データベース、応答生成手法（用例選択、類似度計算）に依存する。したがって評価の際には評価項目以外の要因を統一する必要がある。本稿では質問応答データベースの変化に関して評価するので、後述の実験ではテストセットと応答生成手法を固定して手法評価を行っている。

4. 質問応答データベースの設計

4.1 概要

質問応答データベース（QADB）は質問例と正解応答の対からなるデータベースである。開発者が考案した質問例を基にQADBを作成することもできるが、大規模なシステムにおいては発話事例を収集してQADBを作成することでシステム応答の適切さが向上することが知られている[4]。本研究でもこの知見に基づき、先述の音声データベースを基にQADBを作成している。

先行研究ではデータベース内で複数回出現する書き起こし文

について応答が付与されていたが、本研究の開始にあたって全有効発話について正解応答の付与が完了したため、本稿での実験でも全有効発話を基に質問応答データベースを作成している。

4.2 質問応答データベースの最適化

用例ベースの応答生成において、誤った応答が付与された質問応答対（QAペア）や互いに表現の良く似ているが応答の異なるQAペアが混って選択され、対話システムが不適切な応答を行うことがある。このようなQAペアをQADBから除去することで、対話システムはより適切な応答が可能となる。不要なQAペアを自動検出してQADBから除去する技術を本研究では質問応答データベースの最適化[5]と呼んでいる。

質問応答データベースの最適化では、学習データに対する応答正解率を自動的に計算し、除去によって応答正解率が向上するQAペアをQADBから除去している。実験で用いた手法ではLeave-one-out交差検定法での応答正解率を計算し、除去するQAペアを決定している[6]。なお、Leave-one-out交差検定法での応答正解率の計算では、学習データについて発話単位で交差検定法を行い、正解した学習データの件数から応答正解率を算出している。

5. 書き起こしQADBによる応答生成

5.1 QADBの作成

書き起こし文を用いた従来のQADBを本稿では「書き起こしQADB」と呼ぶ。書き起こしQADBの作成では、「たけまるくん」音声データベースから有効発話データに付与された書き起こし文と正解応答を一对のQAペアとして抽出し、QAペアの重複を除いている。また検索順序に依存するシステム応答の変動を防ぐため、QADB内のQAペアは内部的にソートしている。

5.2 形態素マッチングによる類似度

ある2文が同じ単語を同じ数だけ持つ時、各文は等しいといえる。2文の等しさの度合いを、2文間で共通する単語の割合で表したのが単語単位の類似度である。英語における単語は日本語の形態素に相当するため、日本語の音声入力を扱う本研究では形態素単位の類似度を考える。

1候補の入力 I と質問例 E に関して、形態素マッチングによる類似度 $s(I, E)$ を次の式で与える。

$$s(I, E) = \frac{|I \cap E|}{\max(|I|, |E|)} \quad (2)$$

ここで、 $|\bullet|$ は単語集合（入力・質問例）の形態素数を表す。この類似度は入力と質問例とで一一致する形態素の個数を形態素あたりの割合へと正規化したものとなる。正規化係数として入力文と質問例の形態素数の最大値を用いることで、精度よく用例を選択できる。[8]

なお、入力が複数の候補を持つときは、一致形態素数と入力形態素数として候補あたりの平均値を実数値で求め、類似度を計算している。また、単語の連呼のように文章内で形態素が重複するときは、それぞれ別の形態素とみなし、形態素数とマッチングを求めている。

6. 認識結果 QADB による応答生成

6.1 動機

実用上の音声対話システムでは入力音声に対して誤りのない音声認識結果が常に得られるとは限らない。本研究での音声認識率（単語正解精度）はクローズドテストでの性能は大人データに関して 90%に達しているが、子供データについては 70%程度にとどまっている（7.1 節の表 2）。

用例ベースの応答生成では一般的に質問例の分布（データベース内の質問例の傾向）と入力の分布（ユーザ発話表現の傾向）に近いほど応答の精度は高い。そのため、入力のもつ誤りに一定の傾向が存在すると仮定すると、その傾向を質問例にも含めることで音声対話システムの応答の適切さは向上すると考えられる。このとき、質問例の基となった発話の認識結果と正解応答の対から QADB を構築し、応答生成に用いることになる。本稿では質問例として音声認識結果を用いた QADB を「認識結果 QADB」と呼ぶ。

認識結果 QADB では入力に用いる音声認識器・認識モデルと同一の状況で質問例側の認識結果を生成させる必要がある。また質問例側の認識候補が多いほど入力誤りの傾向をより忠実に表現できると考えられる。なお、この方法の別の効果として、QADB の作成に発話データの書き起こし文を必要としないため、新規システムを開発する際に QADB を作成するコストを軽減できる点が挙げられる。

6.2 認識結果 QADB の作成

認識結果 QADB は書き起こし QADB と同様の方法で作成できる。つまり、まず正解応答が付与された QADB 作成用の発話データを音声認識し、次に発話の正解応答と認識結果を一对の QA ペアとして抽出し、重複を除くことで認識結果 QADB を作成する。本稿では認識結果として平文の形式の候補を用いている。後述の実験では候補数を変化させて認識結果 QADB を作成し、音声対話システムの性能を比較している。なお認識結果の候補数が増えるにしたがって、QA ペアの重複は見られなくなる。

6.3 形態素マッチングの拡張

応答生成に認識結果 QADB を用いるとき、質問例もまた複数の候補をもつこととなる。従来の形態素マッチングでは質問例が複数候補を持つ場合を想定していなかったため、類似度を正しく求めることができない。そのため、複数文の認識候補同士の類似度を計算できるように、形態素マッチングにファジィ集合の概念を導入し、平均形態素数でのマッチングによって形態素の一致度を求める。（図 2）

これまでの形態素マッチングは、入力文や質問例を各要素が 1 以上の重みを持って存在する単語の集合として扱い、集合間の類似度を求めていた、と解釈できる。また、従来の形態素マッチングでは入力が複数の候補を持つときは一致形態素数と入力形態素数の平均値を実数で求めていたが、この従来法は入力に対してファジィ集合の概念を暗黙的に利用していたと言える。

ファジィ集合論では集合の各要素が $[0, 1]$ の重みで属していると考えられる。一方、入力と質問例は一候補の中で複数の形態素

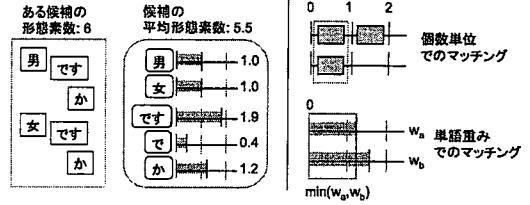


図 2 拡張を試みる形態素マッチングの概念。左：形態素重みの概念による平均形態素数の解釈。右：形態素重みでのマッチング。
Fig. 2 An illustration of the extended morpheme matching.

を持ちうるので、形態素の重みとして 1 以上の値をとることを許すのが自然な拡張だといえる。そこで拡張する形態素マッチングは各形態素が正の実数の重みをもつと考えてマッチングに相当する計算を行う。

認識候補同士の形態素マッチングによる類似度 $s_f(I, E)$ は次式で表される。入力や質問例の形態素の重み $w_I(i), w_E(e)$ は本稿では各候補の尤度の差を考えず、単純に候補あたりの平均値とする。正の実数の一致形態素数 $w(I, E)$ は入力と質問例の両者に存在する形態素の重みの最小値、正規化係数は各集合のサイズ n_I, n_E （形態素重みの総和）の最大値とする。

$$s_f(I, E) = \frac{w(I, E)}{\max(n_I, n_E)}, \quad (3)$$

$$w(I, E) = \sum_{\{i_k, e_l | i_k \in I, e_l \in E, i_k \equiv e_l\}} \min(w_I(i_k), w_E(e_l)), \quad (4)$$

$$n_I = \sum_{i \in I} w_I(i), \quad (5)$$

$$n_E = \sum_{e \in E} w_E(e), \quad (6)$$

$$w_I(i) = \frac{c_I(i)}{C_I}, \quad (7)$$

$$w_E(e) = \frac{c_E(e)}{C_E}. \quad (8)$$

ここで、 $c_I(m), c_E(m)$ は入力・質問例認識結果 I, E 中の形態素 m の個数、 C_I, C_E ：入力 I （質問例 E ）における認識結果の候補数を表す。

7. 実験結果

7.1 実験

認識結果 QADB を用いて応答生成を行うときのシステム応答の適切さを、QADB 最適化の反復による応答正解率の変化によって評価した。また実験環境下で運用される音声対話システムを用いる時は応答速度も重要であるため、QADB のサイズを表す QA ペアの異なり数も比較した。

実験条件は表 1 のとおり。大人データと子供データでは別の音声認識モデルを用いるため、音声データベースを分割して個別に実験を行った。音声認識エンジンには Julius [9] を使い、最適化と評価の入力には上位 10 候補の認識結果を用いた。認

表 1 実験に用いた学習データと評価データの件数.

Table 1 Number of the training set and the test set on the experiments.

評価データ	収集時期	2003年8月	
	データ件数	大人	1053件
		子供	6543件
学習データ	収集時期	2002年12月 - 2004年10月 (評価データは含まない)	
	データ件数	大人	計 19383件
		子供	計 79346件
音声認識エンジン	Julius Ver. 3.5.3		
入力認識結果	上位 10 Best		
質問例認識結果	上位 1, 3, 6, 10 Best		
最適化の反復処理	除去ステップのみ反復し、最後に復帰処理		
評価方法	応答正解率、一意の正解 ID と機械的に比較		

表 2 クローズドテストでの音声認識器の認識性能.

Table 2 Recognition accuracy of automated speech recognizer on the closed test.

大人	単語正解率	単語正解精度
学習データ	90.26%	85.66%
評価データ	90.33%	85.10%
子供	単語正解率	単語正解精度
学習データ	74.44%	66.81%
評価データ	74.76%	67.18%

識結果 QADB は音声認識結果の上位の候補数を変えて 4 種の QADB を作成した。候補数はそれぞれ 1,3,6,10 候補とした。なお学習データと評価データの両者について、クローズドな音声認識モデル (音響モデル・言語モデル) を使用して音声認識を行っている (表 2)。

QADB 最適化の反復処理では除去ステップのみ反復し、反復後に除去 QA ペアの候補を一部取り消すアルゴリズム [6] を用いた。

また認識結果 QADB による応答生成手法のベースラインとして、書き起こし QADB を用いた応答生成を評価した。

7.2 考 察

QADB 最適化の指標となる Leave-one-out 交差検定法による応答正解率を図 3 に示す。どの認識結果 QADB についても最適化による応答正解率の向上が見られた。また 3 Best の認識結果 QADB において、書き起こし QADB を上回る応答正解率となった。このことから、3 Best 以上の認識結果を用いた QADB による応答生成により、書き起こし QADB よりも精度よく応答生成を行えることが期待される。1 Best での応答正解率が全体的に低い結果は、候補数が少ないほど質問例の形態素重みの誤差が大きくなり、入力の傾向と一致しなくなるためだと考えられる。

次に未知の発話に対する応答の適切さの指標として、オープンテストでの応答正解率を図 4 に示す。

大人データの応答正解率に関して、最適化前 (0 回目) では書き起こし QADB での 74.2% に対して 3 Best 以上の認識

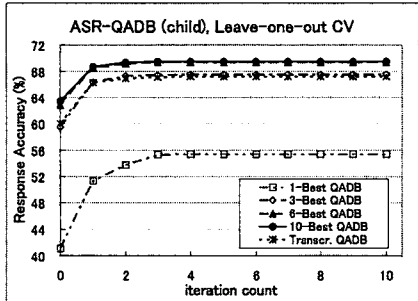
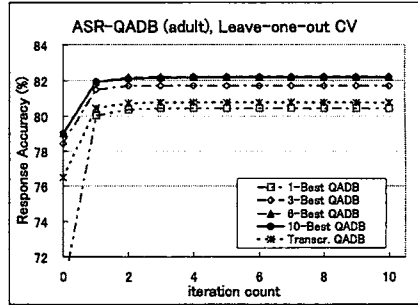


図 3 QADB 最適化による Leave-one-out 応答正解率の変化。上: 大人データでの実験結果, 下: 子供データでの実験結果。横軸は最適化の反復回数。0 回目は最適化をしない時の結果。10-Best QADB: 認識結果 QADB, 認識候補に上位 10 候補を使用。1-6 Best についても同様。Transcr. QADB: 書き起こし QADB
Fig.3 Response accuracy improvement on the leave-one-out cross-validation by QADB optimization.

結果 QADB の結果が優位だった。最大値は 6 Best QADB で 75.0% だった。このことから認識結果 QADB の使用により入力音声認識結果の誤りを QADB によって吸収することができた。

しかし最適化後の結果では、書き起こし QADB で最大 75.5%, 3 Best の QADB で最大 75.2% と、書き起こし QADB が認識結果 QADB よりわずかに優位な結果となった。この結果は、学習データ内の入力に対して認識結果 QADB の最適化を行ったことによる過学習が原因だと考えられる。

子供データに関しては、6 Best の認識結果 QADB での応答正解率が最大で 59.6% と書き起こし QADB での結果と同等の正解率を示し、10 Best の認識結果 QADB では最大 60.0% とわずかに上回った。

このことから、音声認識誤りの多い入力を扱う音声対話システムでは認識結果 QADB を用いることで応答の精度を向上させることが示唆される。

QADB のサイズに関しては、最適化によってサイズが 3~10% 程度軽くなるものの、認識結果の候補が多いほど QA ペアの異なり数が増加した (図 5)。特に 10 Best の認識結果 QADB での QA ペア異なり数は学習データ数とほぼ同数になるため、より大規模なデータベースを用いて候補数の多い認識結果 QADB を構築する際には注意が必要となる。

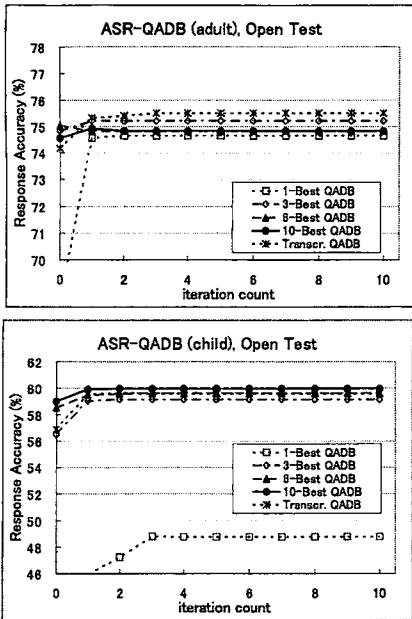


図4 オープンテストでの応答正解率の変化。上：大人データでの実験結果，下：子供データでの実験結果。横軸と凡例の意味は図3と同じ。子供データ，1-Best QADBの0回目の応答正解率は40.6%だった。

Fig. 4 Response accuracy improvement on the open test by QADB optimization.

8. まとめ

一問一答形式の音声対話システムに用いられる用例ベースの応答生成に関して，音声認識結果を用いた QADB (認識結果 QADB) による応答生成を提案した。この方法により，6 候補程度の認識結果 QADB を用いることで入力に含まれる誤りに対して頑健な応答生成が行える。

用例ベースの応答生成では用いる用例選択手法によって性能が左右される。特に本稿で提案した形態素マッチングの手法にはまだ検討の余地がある。今後は試みた定式化を基に，より精度の高い用例選択手法を検討する。

謝辞 本研究は文部科学省のリーディングプロジェクト「e-Society 基盤ソフトウェアの総合開発」によって行われたものである。

文 献

- [1] Ryuichi Nisimura, Akinobu Lee, Hiroshi Saruwatari, Kiyohiro Shikano, "Public Speech-Oriented Guidance System with Adult and Child Discrimination Capability," Proc. of ICASSP2004, vol.1, pp.433-436, 2004.
- [2] Teruhisa Misu, Tatsuya Kawahara, "Speech-based interactive information guidance system using question-answering technique," Proc. ICASSP2007, pp.145-148, 2007.
- [3] A. L. Golin *et al.*, "How May I Help You?," Speech Communication, vol.2 3, pp.113-127, 1997.
- [4] Hiroya Muraio, Nobuo Kawaguchi, Shigeki Matsubara, Yukiko Yamaguchi, Kazuya Takeda, Yasuyoshi Inagaki,

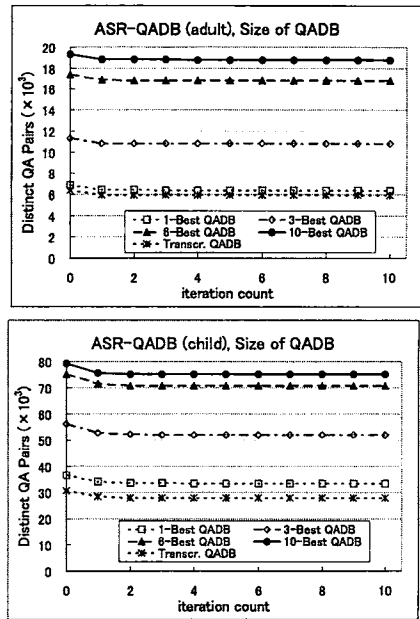


図5 QADBに含まれるQAペアの異なり数の変化。上：大人データでの実験結果，下：子供データでの実験結果。横軸と凡例の意味は図3と同じ。

Fig. 5 Decrease of distinct QA pairs by QADB optimization.

"Example-based Spoken Dialog System with Online Example Augmentation," Proc. of ICSP2004, Spec4402p-7, pp.3073-3076, 2004.

- [5] 木田学, 川波弘道, 猿渡洋, 鹿野清宏, "音声情報案内システムにおける質問応答データベース設計手法の検討," 日本音響学会講演論文集, 3-2-7, pp.85-86, September 2006.
- [6] 竹内翔大, ツインツアレク トピアス, 川波弘道, 猿渡洋, 鹿野清宏, "質問応答データベースに対する Leave-one-out 最適化の反復処理," 日本音響学会講演論文集, 3-Q-14, pp.241-242, September 2007.
- [7] Hiromichi Kawanami, Takahiro Kitamura, Kiyohiro Shikano, "Long-term Analysis of Prosodic Features of Spoken Guidance System User Speech," the 5th international conference on Language Resources and Evaluation (LREC), pp. 2586-2589, 2006.
- [8] 早川直樹, 木田学, ツインツアレク・トピアス, 川波弘道, 猿渡洋, 鹿野清宏, "音声情報案内システムの応答文選択におけるスコアリング手法の改善," 日本音響学会講演論文集, 3-2-8, pp.87-88, September 2006.
- [9] 大語彙連続音声認識エンジン Julius
<http://julius.sourceforge.jp/>.