

論文 / 著書情報
Article / Book Information

論題(和文)	信頼度尺度を用いた音声対話システムの検討
Title(English)	
著者(和文)	本田征嗣, 斉藤 朗, 古井貞熙
Authors(English)	SADAOKI FURUI
出典(和文)	日本音響学会 2000年春季講演論文集, Vol. , No. 3-8-11, pp. 87-88
Citation(English)	, Vol. , No. 3-8-11, pp. 87-88
発行日 / Pub. date	2000, 3

◎本田征嗣 斎藤朗 古井貞熙(東工大)

1. はじめに

音声対話による情報検索システムでは、ユーザの発話からタスクに必要な語(キーワード)のみを抽出し、尤度の高い候補を結果として出力する。不要語をキーワードとして誤認識しないためには、抽出されたキーワードに対し認識結果として適切であるかどうかを検証し、不要語を棄却することが必要である。ここではキーワード検証の手段として、2種類の簡便な方法による対立語を用いた検証モデルを作成し、認識実験を行った。

2. 音声対話システム

これまで、我々は店舗情報検索をタスクとした音声対話システムに関する研究を行ってきた[1]。今回は、同様のタスクを用いて抽出されたキーワードの検証実験を行った。今回のタスクでは、キーワードを駅名、店名の2種類とした。それ以外の文中に挿入される可能性のある語彙をガーベージモデルとして表した。店名の種類を増やしたため店名に続く語彙が複雑化したので、新たに店名を後に続く語彙の可能性によって4種類に分類した。図1において eki は駅名を、mise1~mise4 は店名を表すキーワードモデル。g1、g5 はそれぞれ文頭、文末に挿入される語彙、g3 は店名につく接頭語、g2 は駅名に、g4.1~g4.4 は店名につく接尾語である。

ユーザには、キーワードの上位N位までの候補を認識結果としてディスプレイに提示し、目的のキーワードに到達するまで対話的に絞り込みを行う。

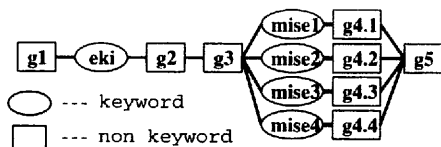


図1. 基本ネットワーク文法

本研究では、NTT-AT社の音声認識開発キット Rex Ver1.1.1を使用し、上記のモデルを記述した。

3. キーワードの検証

音声対話システムでは、正解キーワードのみをユーザに提示することは難しく、不要語または文法外の語彙をキーワードとして認識する誤りが発生する。キーワードの脱落は候補数を増やすことで解決できるが、絞り込みの段階でのユーザの負担が大きくなってしまふ[2]。そこで、抽出されたキーワードが認識結果として適切であるかを検証するシステムを考え

る。この検証で受理されたキーワードのみをユーザに提示、それ以外は不適当な候補とみなし棄却する。これにより不要な候補を削減し、認識結果をより洗練されたものとするのが可能である。

3.1 キーワードの検証方法

以下で用いられているスコアは、Rexにより正規化された単語ごとの音響スコアを示す。

スコアによる検証

認識されたキーワードのスコアを直接用いて検証を行う。対数尤度スコアが閾値以上のキーワードを受理、閾値に満たないものは棄却する。

対立語による検証

各キーワードに対し、キーワードと音響的に類似した単語をモデルに書き加える。この単語をキーワードの検証用モデルとして使用する。以下、この単語を元のキーワードの『対立語』、各キーワードに対立語を書き加えたモデルを『検証モデル』と呼ぶ(図2)。抽出されたキーワードのスコアと対立語のスコアの比が閾値以上の場合、このキーワードを受理し、満たない場合は棄却する。

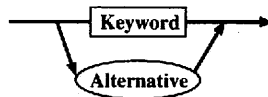


図2. 検証モデル

3.2 検証モデル

これまで発話検証用として、キーワードと競合させるモデルに関する種々の研究が行われてきた(例[3])。これは、キーワードモデルと尤度を比較し、キーワードの信頼性を測る尺度として用いられるモデルである。したがって、競合するモデルはキーワードと一致せず、音響的に類似することが必要である。

本研究では以下の2通りの方法で検証モデルを作成した。

母音列を用いた検証モデル

キーワードから母音のみを抜き出した単語を対立語とする。

(例) しぶや (sibuya) → いうあ (iua)

よこはま (yokohama) → おおああ (ooaa)

* A study on a spoken dialog system using verification measures.

By Seiji Honda, Akira Saito and Sadaoki Furui (Tokyo Institute of Technology)

認識実験結果に基づく検証モデル

実際の認識結果に基づき、各音素が他のどの音素に誤認識されやすいかを調査し、キーワードの各文字を最も誤認識されやすい文字に変換した単語を対立語とする。

3.3 提案手法

以下の3つの手法について、検証を行う。

手法1 通常のモデルを用いて認識し、スコアによる検証のみを行う。

手法2 母音列による検証モデルを用いて認識し、スコアによる検証と対立語による検証の両方を行う。

手法3 認識実験結果に基づく検証モデルを用いて認識し、スコアによる検証と対立語による検証の両方を行う。

4. 実験と評価

今回のシステムでは計361単語(駅名142、店名219)のキーワードと、217単語のノンキーワードから成る単語辞書を用いてネットワークを構成した。また、検証モデルにおいてはこれにキーワードと同数の対立語がネットワークに加わる。ネットワークにおいて、キーワード単語は駅名、店名の順で出現するものとし、駅名、店名のいずれかの脱落は認めるものとする。また、言い直しに対処するため各キーワードは2回までの出現を許すものとする。

4.1 実験条件および評価データ

音声認識実験はRexを用いて行った。評価データは男女計8名の話者による計320文章で、すべての発声は研究室のマイクを用いて収録された。評価データは「駅名、店名ともに含む発声」「駅名のみを含む発声」「店名のみを含む発声」「駅名、店名ともに含まない発声」各80文から構成されている。また、手法3で用いる検証モデルは同様の環境で男女計5名に50音、濁音、半濁音を各5回づつ発声してもらい、その結果を元に作成した。

4.2 評価方法

同一の音声320文を、通常のモデルと検証モデルを用いて認識を行う。抽出された候補単語に対し、3.3で述べた手法1~3を用いて受理または棄却を行う。それぞれの場合で、不要候補の棄却率とキーワード認識率をプロットし、同程度の認識率における棄却率の高さを比較検討する。

4.3 実験結果

表1に各検証手法での認識率ごとの不要語棄却率を示した。ここで認識率とは、発声されたキーワードが認識結果の上位n位までの候補に含まれ、キーワード検証によって受理される確率とし、不要語棄却率とは、上位n位までに入る不要語がキーワード検証によって棄却される割合とする。

この結果、手法1に比べ手法2、3の方が高い棄却率を得ることが出来た。これにより、キーワード検

表1. 認識率と不要語棄却率

キーワード認識率		不要語棄却率		
n=1		手法1	手法2	手法3
駅名 (92.9%)	80%	66.2%	89.2%	70.3%
	85%	58.1%	86.5%	66.2%
	90%	35.1%	81.1%	39.2%
店名 (82.1%)	70%	66.7%	66.7%	74.8%
	75%	64.6%	61.2%	70.8%
	80%	47.6%	-	40.1%
平均 (87.5%)	70%	71.5%	83.2%	80.1%
	75%	65.1%	77.4%	74.7%
	80%	61.5%	70.1%	70.6%
n=5		手法1	手法2	手法3
駅名 (97.1%)	85%	85.0%	93.7%	89.6%
	90%	81.6%	91.2%	85.8%
	95%	65.3%	78.0%	70.5%
店名 (90.0%)	75%	64.6%	86.9%	84.6%
	80%	76.2%	82.7%	81.5%
	85%	58.9%	68.2%	55.3%
平均 (93.6%)	80%	81.2%	90.1%	86.1%
	85%	77.9%	86.9%	83.3%
	90%	59.6%	70.2%	59.3%

()内は不要語の棄却を行わない場合の認識率

太字の数字は対立語を用いて手法1よりも

5%以上棄却率が上昇したものの

証において検証モデルを用いることは有効な手段であると言える。また、手法2と手法3を比較した場合、n=1の店名を除いて手法2の棄却率が上回った。

5. まとめ

本研究では、音声対話システムにおいて抽出されたキーワードの検証を行い不要語を棄却する手法を提案した。2種類の対立語について認識実験を行い、認識率と不要語棄却率を比較し評価した。この結果、提案した検証モデルが不要語棄却に有効であることがわかった。母音列による検証モデルを用いた手法で最も大きな効果が得られ、n(候補数の上限)=5、認識率80%のシステムでは不要な候補の90%を棄却することが出来た。

謝辞

様々なアドバイスを頂いたNTTサイバースペース研究所の大附克年氏、ならびに本研究に協力頂いた東工大古井研究室の皆様へ深く感謝致します。

参考文献

- [1] 山口晃一郎, 古井貞照 "音声対話を用いた情報検索システムの検討", 日本音響学会春季講演論文集 3-6-19, 1998.
- [2] 山口晃一郎, "情報検索のための音声認識を用いた対話システムの研究", 東京工業大学修士論文, 1998.
- [3] 河原達也, 石原健太郎, 堂下修司, Chin-Hui Lee "会話スタイル依存の言語モデルを用いたキーフレーズの検出・検出", 電子情報通信学会信学技報SP97-78, 1997.