

## 音声対話システムにおける発話文の 自動クラスタリングに基づく応答選択

垣 鍔 亮 太<sup>†</sup> 南 角 吉 彦<sup>†</sup>  
李 晃 伸<sup>†</sup> 徳 田 恵 一<sup>†</sup>

公共の場における音声情報案内システムでは、幅広いユーザの自然な発話を対象とするため、さまざまな内容や認識誤りを含む発話に対して適切な応答を選択する必要がある。本稿では、発話文の自動クラスタリングに基づく統計的応答選択手法を提案する。システムに対するユーザ発話の書き起こし文集合に対して、文間の類似度からクラスタリングを行い、代表発話文（テンプレート）の集合を得る。入力時には、発話の認識結果の文に対して最も類似度の高いテンプレートを求め、それに対応した応答を出力する。発話文の自動クラスタリングには Y-クラスタリングを用い、さらに精度良く文のクラスタリングを行うため付属語ペナルティおよび固有名詞のタグ化を行った。応答選択システムを構築し、公共音声情報案内システム「たけまるくん」のデータを用いて実験を行った結果、従来のキーワードベースの応答選択手法による応答正解率が 74.4% であるのに対して 75.2% の性能が得られた。また、様々なパラメータ設定による性能の比較や、K-means 法との比較も行った。

### Question Answering based on Automatic Utterance Clustering in Spoken Dialogue System

RYOTA KAKITSUBA,<sup>†</sup> YOSHIHIKO NANKAKU,<sup>†</sup> AKINOBU LEE<sup>†</sup>  
and KEIICHI TOKUDA<sup>†</sup>

A public spoken dialog system operated on real situation has to deal with vast variety of utterances with disfluencies and ill-formedness from various speakers, and their recognition errors are also inevitable. In this paper, a robust method of statistical response selection on question answering system based on automatic utterance clustering is proposed. A set of transcribed user utterances are first clustered by their similarity between sentences to generate a set of representative sentences (templates). On question answering, the similarities between a recognized sentence and the templates are calculated, and a response that corresponds to the template of the highest similarity will be chosen as an output. A clustering method called "Y-clustering" is used for the automatic utterance clustering, and two modifications are made for more accurate clustering: ancillary words penalty and lexical tagging on proper noun. Experimental results on the data of public speech-oriented information guidance system "Takemaru-kun" showed that the proposed system achieved a response correct rate of 75.2%, which outperforms the conventional method based on keyword matching whose correct rate was 74.4%. The clustering performance on various configurations are also investigated, and a comparison with K-means method are also reported.

#### 1. はじめに

音声対話システムにおいて、ユーザの望む応答を返すことは重要である。ユーザの広範囲かつ認識誤りを含むような発話一つ一つに対して的確な応答を行うことが求められる。

実際の音声対話システムにおいては、ユーザの自発発話の範囲やゆれが大きく、その全てをあらかじめ想

定することは困難である。また話し言葉による発話のゆれが大きく、これらに起因する認識誤りの問題がある。この問題に対する既存のアプローチとしては、ルールに基づいて文の意味を定義し、ユーザ意図を取り出す手法<sup>1)2)</sup> や、認識結果からキーワードを抽出し対応する応答を返すアプローチ<sup>3)4)5)</sup>、音声認識結果の信頼度を用いる手法<sup>6)</sup> 等がある。例えば公共情報音声案内システム「たけまるくん」<sup>7)</sup> では、想定されるユーザの質問とそれに対するシステムの応答の組で形成される質問応答データベース (DB) を用いたキーワードによる応答選択を行っている。しかし、応答選択のためのスコアリングは明確な基準の設定が難しい。

<sup>†</sup> 名古屋工業大学大学院 工学研究科  
Department of Computer Science and Engineering,  
Nagoya Institute of Technology

本稿では、システムに対する発話文の集合に対して、文間の類似度に基づいて自動クラスタリングを行い、表層的に似た発話を1つのセントロイドに結びつけることにより、発話のくずれや認識誤りを含んだ入力文に対してより頑健な応答生成を実現することを目指す。文のクラスタリング手法として、H. Yeらが提案したY-クラスタリング<sup>8)</sup>を用いる。彼らは入力発話文に対して用意されたスロットに対応する単語の位置を同定する目的で用い、特に書き起こし文が少ない場合によい性能が得られることを示している。本稿ではこのクラスタリング手法の類似文のクラスタリング能力に着眼し、応答選択システムに適用するとともにいくつかの拡張を行う。

以下、第2章で発話文の自動クラスタリング、第3章で発話文のクラスタに基づく応答選択手法について述べる。第4章で実験条件および結果を示し、最後に第5章でまとめと今後の課題について述べる。

## 2. 発話文の自動クラスタリング

ユーザの発話間の類似度に基づいて書き起こし文をあらかじめクラスタリングしてセントロイドとなる文の集合を求め、入力文を最も近いクラスタのセントロイドの文に置き換えて応答処理を行う。これにより発話のくずれや音声認識誤りに対し頑健な応答選択が期待できる。

発話文の自動クラスタリングにはY-クラスタリング<sup>8)</sup>を用いる。 $N$ 個の学習文 $\{S_1, \dots, S_N\}$ が与えられたとき、Y-クラスタリングは各クラスから代表文(セントロイド)を同定する。以降、代表文をテンプレートと呼ぶ。クラスタリングにおける距離は文間の類似度(Similarity)を用い、その類似度計算には単語特徴度(Word Saliency)を用いる。Y-クラスタリングは、テンプレートと割り当てられた各学習文の類似度が、あらかじめ設定する閾値以上であることが保証されるという特徴を持つ。この手法を用いることで、従来人手で作成していた想定質問文を収録発話文から自動構築ができる上、入力文とテンプレート間の類似度を定義できることから、統計的な情報に基づいた応答選択スコアリングを行えることが可能と期待できる。

以下、2.1節でY-クラスタリングにおける単語特徴度と類似度の定義を述べ、2.2節で具体的なアルゴリズムを述べる。

### 2.1 単語特徴度と類似度の定義

発話文間の類似度計算のため、Y-クラスタリングでは文中の各単語の特徴度(Word Saliency)という尺度を定義する。対象とする学習文の集合に合計 $M$ 個のクラスがあり、そのうちのあるクラス $C$ を考える。クラス $C$ が $H$ 個の学習文を持ち、それらのうち $\alpha$ 個の学習文が単語 $w$ を含み、また $M$ クラス中 $\beta$ 個のクラスが単語 $w$ を含む学習文を持つとする。このとき

クラス $C$ における単語 $w$ の単語特徴度は式(1)で定義される。

$$T(w) = \sqrt{\frac{\alpha}{H} \times \left(1 - \frac{\beta}{M}\right)} \quad (1)$$

式(1)における $\alpha/H$ はクラス内における単語 $w$ を含む文の出現確率を表し、 $\beta/M$ は全クラスにおける単語を含むクラスの割合である。値は0から1の値となる。クラス内において頻繁に出現し、そのクラス外ではほとんど出現しない単語に対して高い単語特徴度が与えられる。

次に、式(1)で与えられた単語特徴度の定義を用いて、学習文 $S$ とテンプレート $T$ 間の類似度を定義する。 $L$ を学習文 $S$ を構成する単語数、 $K$ をテンプレート $T$ を構成する単語数とする。また、学習文 $S$ とテンプレート $T$ 間でDPマッチングを用いて単語アライメントを取り、マッチした単語数を $J$ とする。このとき、学習文 $S$ とテンプレート $T$ 間の類似度は式(2)で定義される。

$$d(S, T) = \sqrt{\frac{J}{L} \times \frac{\sum_{j=1}^J T(w_j)}{\sum_{i=1}^K T(w_i)}} \quad (2)$$

### 2.2 Y-クラスタリングのアルゴリズム

Y-クラスタリングを用いた発話文の自動クラスタリングのアルゴリズムを以下に示す。

- (1) 事前に類似度の閾値 $p(0 \sim 1)$ および1回の試行で生成されるクラスの最大数 $M$ を決める。
- (2) 全ての学習文にOOC(Out Of Class)タグをつけ、コーパスから $M$ 個の同一でない文を選んで初期テンプレート集合とする。このとき全ての単語特徴度は1とする。
- (3) 各学習文 $S$ ごとに、全てのテンプレート $T$ に対する類似度 $d(S, T)$ を式(2)を用いて順に計算する。その類似度が閾値 $p$ 以上であれば文 $S$ をテンプレート $T$ をセントロイドとするクラスに割り当て、OOCタグを外し次の文に進む。一方類似度が全てのテンプレートに対して閾値 $p$ より小さい場合、一番高い類似度のクラスに割り当てるが、このときその文からOOCタグは外さない。
- (4) 全ての学習文の割り当てが終わったら、各テンプレートが属するクラスに割り当てられた学習文数の降順でテンプレートをソートし、空のクラスを消去する。
- (5) OOCタグのついた学習文の数が前回の試行から減少しなくなった場合、(6)に進む。それ以外の場合には、単語特徴度を式(1)を用いて更新し、テンプレートを以下に示す方法に従い再生成する。その後(3)に戻る。

(テンプレート再生成): 各クラスにおいて、クラス内類似度 $D(S)$ が最大となる文 $S^T$ を式(3)により求め、 $S^T$ を各クラスのテンプレ

トとする.

$$S^T = \arg \max_{S_i} \{D(S_i)\} (i = 1, \dots, H) \quad (3)$$

なおクラス内類似度  $D(S)$  は式 (4) で与えられ、これはクラス中の全学習文と学習文  $S$  間の類似度の合計である.

$$D(S) = \sum_{h=1}^H d(S, S_h) \quad (4)$$

$H$  はそのクラスに割り当てられた学習文の数である.

- (6) 現在のテンプレートの集合を保存する. OOC タグが残っている全ての学習文に対して, (2) からのステップを再帰的に行う. 全ての学習文から OOC タグが外れるまでこの処理を繰り返す.
- (7) 生成された全てのテンプレートを集め, 各学習文を最も類似度の高いクラスに割り当てる. その後全クラスを割り当てられた学習文の数で降順にソートする.

図 1 に文を 2 次元空間における点と考えたときのクラスタリングの処理のイメージを示す. 図中の各点は文, 円はクラスタ (半径は閾値  $p$ ), 三角はテンプレートをそれぞれ示す. 1 回の (2)~(5) の処理で,  $K$ -means と同一の方法で  $M$  個のテンプレートの選択を行い, 同時に各テンプレートに対して類似度  $p$  以上の文を取り出す. その後, 残った文に対して上記の処理を繰り返す.

このアルゴリズムは各クラスにおける文間の類似度が閾値以上であることを保証する. また, 生成されるクラスタの最終的な数は閾値  $p$  および学習文に依存する.

### 3. 発話文のクラスタに基づく応答選択

前節の発話文クラスタリングに基づく応答選択手法を以下に述べる. 図 2 に提案システムの概要を示す. まず, システムに対するユーザ発話の書き起こしデータに対して  $Y$ -クラスタリングを行い, 発話文のクラスタと各クラスタのセントロイド (質問文テンプレート) を構築する. 各テンプレートはクラスタリング時の単語特徴度を保持している. つぎに, 得られた各テンプレート毎に応答文の対応付けを行い, 応答選択システムを構築する. このシステムに発話文が入力されたとき, 入力文と各テンプレート間で類似度を計算し, 最も高い類似度を出したテンプレートと対となる応答文をシステム応答として出力する.

本稿では, より精度良くクラスタリングを行うために,  $Y$ -クラスタリングにいくつかの拡張を行った. 3.1, 3.2 節でその拡張を説明し, 3.3 節で本稿で提案する応答選択アルゴリズムを述べる.

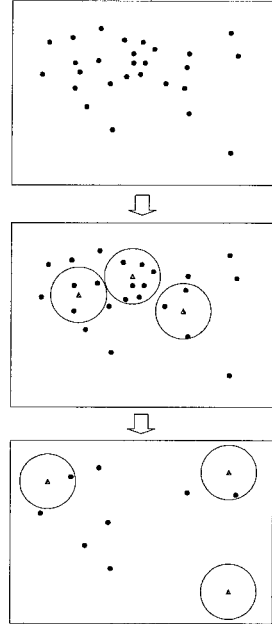


図 1 クラスタリングの処理のイメージ

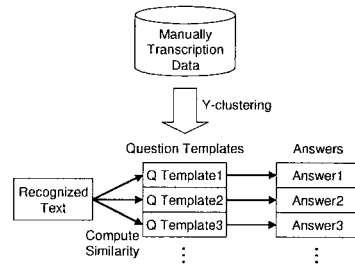


図 2 システムの概要図

#### 3.1 固有名詞のタグ化

一般に特定のタスクに特化した音声対話システムでは, 固有名詞を含む発話が多い. このような固有名詞のみ異なる似た発話が異なる発話文クラスタとされるのを避けるため, クラスタリングの前処理として, 学習文の固有名詞をタグ名に置き換える. これにより一部の固有名詞のみ異なる似たクラスタが大量に生成されるのを防ぐことが期待できる.

例を挙げると,

はばたきホールってどこですか  
図書館ってどこですか

という文は

[施設]ってどこですか

```

Input :
頑張っ+ガンバツ+47/17/6 て+テ+64 下+シタ+2 -+17+0 さい+サイ+2
Similarity Template
0.515948 頑張っ+ガンバツ+47/17/6 て+テ+64 ね+ネ+69
0.338215 泣い+ナイ+47/8/5 て+テ+64
0.319273 こんにち+コンニチ+ハ+75 -+17+0
0.314779 さようなら+サヨウナラ+75 -+17+0
0.313637 教え+オシエ+47/6/4 て+テ+64
0.268997 躡っ+オドツ+47/17/6 て+テ+64 下さい+クダサイ+48/19/9
0.264638 こんち+コンチ+17 は+ハ+65 -+17+0
0.262270 は+ハ+2 -+17+0 か+カ+70
0.260476 教え+オシエ+47/6/4 て+テ+64 下さい+クダサイ+48/19/9
0.251435 もしも+モシモ+56 -+17+0 し+シ+47/3/7
0.228893 お+オ+42 名前+ナマエ+2 は+ハ+65 -+17+0
0.224318 たけ+タケ+2 まるく+マルク+51/50/6 -+17+0 ん+ン+21
0.221389 す+ス+42 ご+ゴ+42 -+17+0 い+イ+47/6/4
0.220560 歌っ+ウタツ+47/21/5 て+テ+64 み+ミ+48/6/4 て+テ+64

```

図3 類似度計算例

```

Input : 郵便局はどこ?
Tagged : [施設]+シセツ+4 は+ハ+65 どこ+ドコ+14
Template : [施設]+シセツ+4 は+ハ+65 どこ+ドコ+14 です+デス
+74/56/1 か+カ+70
Similarity : 0.932539527923347
Answer : 最寄りの郵便局は、タカヤマ郵便局です。正面玄関前の道を北上してください。

```

図4 応答選択の例

のように置き換えられる。

### 3.2 付属語ペナルティの導入

文間の類似度計算時に、自立語は意味的により重要な単語と考え、自立語の強調を行う。これは、自立語と付属語とでスコアに差をつける付属語ペナルティを導入することで実現する。付属語ペナルティは式(5)、付属語ペナルティを導入した類似度計算の式は式(6)で与えられる。

$$P(w) = \begin{cases} 1 & w \text{ が自立語} \\ \gamma & w \text{ が付属語} \quad (0 \leq \gamma \leq 1) \end{cases} \quad (5)$$

$$d(S, T) = \sqrt{\frac{\sum_{j=1}^J P(w_j)}{\sum_{i=1}^I P(w_i)} \times \frac{\sum_{j=1}^J T(w_j) P(w_j)}{\sum_{i=1}^K T(w_i) P(w_i)}} \quad (6)$$

### 3.3 応答選択アルゴリズム

本応答選択システムに認識結果が入力された際の応答選択アルゴリズムを示す。

- (1) 入力文中の固有名詞をタグ化する。このときタグの値を保存しておく。
- (2) (1)の手順を施した入力文と収録発話文から得られた各テンプレートの類似度を式(6)を用いて計算する。
- (3) 最も類似度の高いテンプレートとペアになっている応答文をシステム応答とする。

入力文とテンプレート間の類似度計算の例を図3に、図4に応答選択の例を示す。Inputが入力文、Tagged

```

入力文      : わたちはユカコですお名前は?
正解応答   : 私は、竹からうまれた、タケマルと申します。
システム応答 : わたしは、たけまる、3才です。
入力文      : 地図を
正解応答   : マピオンにアクセスします。
システム応答 : 生駒市の地図を表示します。

```

図5 正解外であるが許容すべき応答の例

が固有名詞のタグ化を行った入力文、Templateが最も高い類似度を出したテンプレート、Similarityが上記テンプレートに対する類似度、Answerが選ばれた応答である。

なお、音声認識結果のN-bestを用いる場合、各認識候補それぞれに対し上記手順(1)~(3)を行い、(3)において最も類似度が高いテンプレートとペアになっている応答文に類似度を投票し、得票数が最大となる応答文をシステム応答とする。

## 4. 実験的評価

### 4.1 実験条件

提案法の評価のため、一問一答形式の音声対話システムである公共音声情報案内システム「たけまるくん」<sup>7)</sup>における応答選択<sup>9)</sup>と同じデータを用いて比較実験を行った。本システムの発話は、非常に自由かつ短い発話が多いのが特徴である。

実験データは「たけまるくん」によって収録したユーザ発話のうち成人のみを用いた。表1に諸元を示す。2002年11月から2004年10月の期間において収録された発話データのうち、2003年8月を除いた期間の有効発話中、書き起こし文の出現頻度が2回以上の発話データを学習データ、2003年8月において収録された有効発話すべてをテストセットとしている。

入力として、書き起こし文および音声認識結果の双方を評価する。評価尺度としては、入力に対する応答がテストセットに対してあらかじめ付加されている応答正解ラベルと一致するかどうかで応答正解率を算出した。

今回、テストセットは文献<sup>9)</sup>にあるデータから、応答正解文の決定に曖昧性のある発話を除いたものを用いることとした。テストセットの発話文に対し正解応答文が定められているが、それ以外にもシステム上許容されるべき応答文が存在する。例を図5に示す。この例では「地図を」という入力文に対して「マピオンにアクセスします。」という応答が正解応答文と定義されているが、「生駒市の地図を表示します。」という応答もシステム上正解と許容されるべきである。テストセット中約1割のデータについて、上記のようなあいまいな応答があったため、これを除外して評価を行った。

提案する応答選択システムを、表1の学習データを

表 1 データの諸元

	学習データ	テストセット
実験データの期間	2002.11~2004.10 (2003.08 は除く)	2003.08
抽出条件	出現回数 2 回以上のみ	1 回も含む
抽出発話数	13487	999
音声認識の単語正解精度	91.1%	80.3%

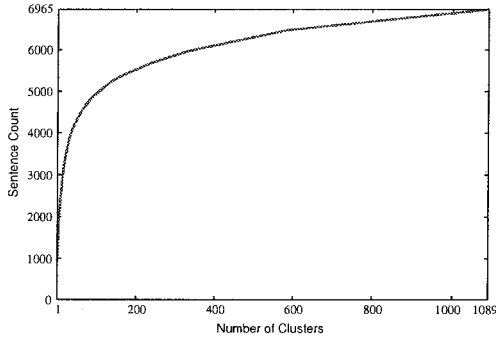


図 6 クラスタのメンバ数の累積分布

用いて 3.3 節で述べた手順で構築した。付属語パネルティは全て  $\gamma = 0.2$  で固定した。なお、クラスタリング時はプログラムの制限から表 1 中の学習データにおいて各文の出現回数を半分と仮定した 6965 文を用いた。システム応答は従来法<sup>9)</sup>と同じ応答文 242 文を各テンプレートに対応づけた。なお、この対応づけはテストセットを参照せずに行った。また、音声認識結果を入力する場合は音声認識結果の上位 10 候補それぞれに対し最も類似度が高いテンプレートを求めて、そのテンプレートとペアになっている応答文に類似度を投票し、得票数が最大となる応答文をシステム応答とした。

#### 4.2 クラスタリングの結果

クラスタリング条件を閾値  $p = 0.9$ 、1 ステップ当たりの生成クラス数  $M = 400$  としてクラスタリングを行った。結果として総数 1089 のテンプレートが得られた。これらのうち、メンバ数が多い上位 5 つを表 2 に示す。なお、メンバ数が 1 (セントロイドのみ) のクラスタが約 500 であった。また、メンバ数が 4 以下のクラスタが全クラスタの約 9 割弱を占めた。これは、出現頻度が極端に低く、他の発話との類似度も低い外れ発話によるものと考えられる。

図 6 にクラスタのメンバ数の累積分布を示す。図より上位約 50 位のクラスタが特にメンバ数が多く、それ以降のクラスタのメンバ数が少ないことが分かる。

#### 4.3 応答選択実験結果

##### 4.3.1 各種パラメータの関係

1 ステップ当たりの生成クラス数  $M$  を 400 に固定し閾値  $p$  を変化させた場合の結果を表 3 に、 $p$  を固定

表 2 上位クラスタのメンバ数

順位	テンプレート	メンバ数
1	こんにちは+コンニチハ+75	910
2	[施設]+シセツ+4 は+ハ+65 どこ+ドコ+14 です+デス+74/56/1 か+カ+70	423
3	今+イマ+16 何時+イツ+16 です+デス+74/56/1 か+カ+70	267
4	お+オ+42 名前+ナマエ+2 は+ハ+65	248
5	何+ナン+19 歳+サイ+35 です+デス+74/56/1 か+カ+70	146

表 3 閾値  $p$  を変化させたときの応答正解率 ( $M = 400$ )

	総テンプレート数	応答正解率	
		書き起こし文	音声認識結果
$p = 0.7$	637	73.2%	69.2%
$p = 0.8$	879	76.3%	72.9%
$p = 0.9$	1089	79.0%	75.2%

表 4 1 ステップ当たりの生成クラス数  $M$  を変化させたときの応答正解率 ( $p = 0.9$ )

	総テンプレート数	応答正解率	
		書き起こし文	音声認識結果
$M = 10$	1084	76.1%	72.5%
$M = 100$	1093	77.5%	74.0%
$M = 400$	1089	79.0%	75.2%
$M = 1000$	1093	78.7%	74.4%

し  $M$  を変化させた場合の結果を表 4 にそれぞれ示す。各表は生成された総テンプレート数および応答正解率をそれぞれ示す。

表 3 より、閾値  $p$  を大きくすることで総テンプレート数は多くなるが正解率が高くなることがわかる。閾値が小さいときは文間の類似度が低くても同一クラスに分類されることにより、本来同一にされるべきでない文がクラスに統合されるため、総テンプレート数は少なくなるが応答正解率が低下する。閾値を大きく設定した場合は逆に文間の類似度が高いものだけがクラス化されるため、テンプレート数は多くなるが応答正解率は良くなる傾向にある。

表 4 より、1 ステップ当たりの生成クラス数  $M$  は応答正解率および生成される総テンプレートに対し閾値  $p$  ほど大きな影響を与えないことがわかる。

##### 4.3.2 従来法との比較

次に、構築したシステムを従来法<sup>9)</sup>と比較した。従来法は、人手で作成した想定質問応答 DB を使い、自立語のみのキーワードマッチングスコアリングに基づく応答選択を行う。従来法は想定質問文 3726 文を用いる。提案法は前出のうち最も高い正解率を出した  $p = 0.9$ 、 $M = 400$  のものとする。

表 5 に応答選択実験結果を示す。表の各値は従来法・提案法それぞれにおける応答正解率 (%) である。こ



表5 従来法との比較

	書き起こし文	音声認識結果
従来法 (キーワード)	79.7%	74.4%
提案法	79.0%	75.2%

表6 クラスタリング手法ごとの応答正解率

クラス数	手法	応答正解率	
		書き起こし文	音声認識結果
400	K-means	67.5%	67.9%
	Y-クラスタリング	75.4%	72.1%
1000	K-means	66.3%	68.2%
	Y-クラスタリング	79.0%	75.2%

の表より、従来法と比較して、書き起こし文の入力に対してより少ない質問文で同等の性能が得られた。さらに、音声認識結果の場合は、従来法が74.4%なのに対し75.2%と改善された。この理由として、提案法は従来法と比較し、音声認識誤りに起因するゆれに対して頑健性が高いことが示唆される。

なお、書き起こし入力へのテストセット全体で選択されたテンプレートの種類数は316であった。総テンプレート数が1089であることから、実際には選択されないテンプレートが多いことが分かった。今後の課題として発話文クラスタリングアルゴリズムの改善が挙げられる。

#### 4.3.3 K-means アルゴリズムとの比較

発話文の自動クラスタリング手法として用いたY-クラスタリングをK-meansと比較した。K-meansアルゴリズムに基づくクラスタリングを用いて発話文から発話文クラスタを得て、それに基づく応答選択実験を行った。

表6に実験結果を示す。K-meansアルゴリズムは分割数Kを400と1000で行った。比較のため、Y-クラスタリングは前節の条件(閾値 $p = 0.9$ ,  $M = 400$ )に加え、テンプレート数をそろえるため、2.2節のステップ(7)において得られたテンプレート集合のうち、上位Y個以降を消去したものをを用いた。

表6のとおり、どちらの場合でもY-クラスタリングを用いた方がより高い正解率を出したことがわかる。これは、従来のK-meansアルゴリズムで文のクラスタリングを行うと各学習文が最も類似度の高いテンプレートに割り当てられるため、テンプレートと各学習文の類似度が保証されないが、Y-クラスタリングにおいては各クラスにおける文間の類似度が閾値以上であることが保証されるためと考えられる。

## 5. むすび

本稿では、一問一答形式の音声対話システムにおいて、発話文の自動クラスタリングに基づいた応答選択

手法を提案し評価した。実験的評価によって、従来のQAデータベースを用いた応答選択手法と比較して、約3割少ない質問文テンプレートで、音声認識結果の入力では従来を若干上回る性能を示した。本手法が対話システムにおいて発話のゆれや認識誤りのゆれに対処する統計的な枠組みとして有効であることが示された。今後の課題としては、不要なテンプレート生成を防ぐためのクラスタリングアルゴリズムの改善、および本学に設置予定の音声情報案内システムに実装し、実際の運用からさまざまな探索を行うことがあげられる。

## 6. 謝 辞

本稿の実験を行うにあたって、音声情報案内システム「たけまるくん」のデータを貸与していただいた、鹿野清宏教授をはじめとする奈良先端科学技術大学院大学の皆様に感謝します。

## 参 考 文 献

- 1) Roberto Pieraccini, Evelyne Tzoukermann, Zakhari Gorelov, Esther Levin, Chin-Hui Lee, Jean-Luc Gauvain, "Progress Report on the Chronus System: ATIS Benchmark Results", in *Proc. of the DARPA Speech and Natural Language Workshop*, pp.67-71, (1992).
- 2) Yulan He, Steve Young, "Semantic processing using the Hidden Vector State model", in *Computer Speech and Language* 19, pp.85-106, (2005).
- 3) 竹林 洋一. 音声自由対話システム TOSBURGII-ユーザ中心のマルチモーダルインタフェースの実現に向けて. 電子情報通信学会論文誌, vol.J77-D-2, no.8, pp.1417-1428, (1994).
- 4) 西村 竜一, 内田 賢志, 李 晃伸, 猿渡 洋, 鹿野 清宏. Juliusを用いた学内案内ロボット用音声対話システムの作成. 情報処理学会研究報告, 2001-SLP-39-16, pp.93-98, (2001).
- 5) 駒谷 和範, 上野 晋一, 河原 達也, 奥乃 博. ユーザモデルを導入したバス運行情報案内システムの実験的評価. 情報処理学会研究報告, 2003-SLP-47-12, pp.59-64, (2003).
- 6) 駒谷 和範, 河原 達也. 音声認識結果の信頼度を用いた頑健な混合主導対話の実現法. 情報処理学会研究報告, 2000-SLP-30-9, pp.39-44, (2000).
- 7) 西村 竜一, 西原 洋平, 鶴見 玲典, 李 晃伸, 猿渡 洋, 鹿野 清宏. 実環境研究プラットホームとしての音声情報案内システムの運用. 電子情報通信学会論文誌, Vol.J87-D-2, No.3, pp.789-798, (2004).
- 8) Hui Ye, Steve Young, "A Clustering Approach to Semantic Decoding", in *Proc. of ICSLP*, pp.5-8, (2006).
- 9) 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. 音響学会講演論文集, 3-2-7, pp.85-86, (Sep. 2006).