

堀 智織 鈴木 潤 堀 貴明 磯崎 秀樹 前田 英作  
 (日本電信電話株式会社 NTT コミュニケーション科学基礎研究所)

## 1. はじめに

これまで、新聞記事などの大規模コーパスを対象とした質問応答(QA)システム(SAIQA)の研究を進めてきた[1]。現在、より簡便で高精度なQAシステムを目指し、音声インタラクティブQAの研究を行っている。テキスト以外のインタフェースを利用したQAシステムとして、音声入力を適用したQAシステムが検討されている[2]など、音声対話を組み込んだインタラクティブQAシステムが注目されつつある。本研究では、ユーザとシステムが音声対話を行い、逐次的に知識を共有しつつ、最終的な答えを出力することを目指している。本稿では、音声入力された質問文を解析し、その曖昧性を解消するための応答対話文自動生成法を提案する。

## 2. 曖昧性解消のための応答対話文生成法

質問文「ワールドカップの優勝国は南米のどこ？」に対する構文解析例を図1に示す。この例では、破線

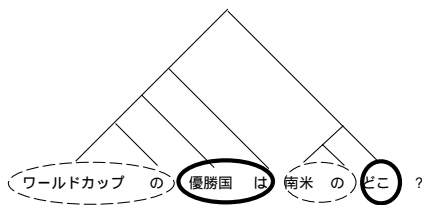


図1. 構文解析例

の丸で囲まれた「ワールドカップ」「南米」という単語は、他の単語から係られていないことから「ワールドカップの」という文節に係られている「優勝国」に比べ、意味が限定されておらず、曖昧性が高い。一般的に、質問文において、他の単語から係られる回数(係られ度)が低い実質語は、曖昧性が高いと考えられる。また、検索対象となるコーパス中に高頻度で出現している単語は、質問文を特徴づけておらず、曖昧性が高いと考えられる。ただし、「南米」などの固有名詞や数値表現の曖昧性は低い。

この例では、「ワールドカップ」の持つ曖昧性により、質問文の答えは一意に限定されない。それ故、質問の答えを絞り込むには、「いつのワールドカップ」、「何のワールドカップ」という情報を追加し、「ワールドカップ」の曖昧性を解消する必要がある。

そこで、質問文中で係られ度の低い情報で、なおかつコーパス中に高頻度で出現する情報を対象とし、曖昧性を解消するための応答対話文を生成する。ただし、曖昧性解消の対象となる情報は、文節、および、名詞節とする。具体的には、全ての文節、およ

び名詞節に、「いつの」、「どこの」などの疑問詞を先頭に付与する。さらに、文末に「ですか?」を接続させることで応答対話文を生成する。以下に応答文生成の例を示す。

疑問詞 + 抽出された単語列 + 疑問符  
 「いつの」+「ワールドカップ」+「ですか?」  
 「どこの」+「ワールドカップ」+「ですか?」

次に、自動生成された応答対話文の候補の中から、最適な応答対話文を決定する。質問文中の  $n$  番目の文節を、 $P_n = c_{n1}, c_{n2}, \dots, c_{nt_c}, f_{n1}, f_{n2}, \dots, f_{nt_f}$  と表す。ただし、 $c$  は実質語、 $f$  は機能語を示す。文節は、1個以上の実質語+0個以上の機能語で構成されるものとする。文節  $P_n$ 、および同文節中の名詞節  $C_n = c_{n1}, c_{n2}, \dots, c_{nt_c}$  を対象として、曖昧性を解消する。疑問詞  $Q$  と  $P_n$  または  $C_n$  の組み合わせで生成される応答対話文を  $S$  とする。

曖昧性を解消する応答対話文の適正を示す尺度を、次式のように定義する。

$$A(S) = \lambda_I \sum_{x=1}^{t_c} I(c_{nx}) + \lambda_D \sum_{y=1}^{n-1} D(P_y, P_n) + \max\{L(P_n), L(C_n)\}$$

ただし、単語の汎用性を示す単語汎用度スコア  $I$ 、質問文としての適正を示す言語スコア  $L$ 、他の文節から係られている度合いを示す係られ度スコア  $D$  の重みつき総和で定義する。 $\lambda_I$ 、 $\lambda_D$  は、各スコアのバランスをとるための重み係数である。

$L$  は、単語  $n$ -gram 確率に基づく  $S$  の生成確率の対数値である。 $I$  は、単語 unigram 確率の対数値である。日本語は前方から後方への係り受けなので、当該文節  $P_n$  と、それに先行する各文節との間の係り受け確率  $D$ [3] の総和を係られ度スコアと定義する。

## 3. 評価実験

### 3.1 実験条件

質問文 69 文に対して、3 人の男性被験者が読み上げ、音声認識を行った。ただし、読み上げた質問文は、文法的に正しく、書き言葉で書かれた文である。平均単語正解精度は 81% であった。提案手法により曖昧性を解消するための応答対話文を自動生成した。誤認識を含む認識結果から、失った情報の補完を促す応答対話文が自動生成できるかを評価した。上位 10 位の応答対話文に対し、2 名の被験者による主観評価を行った。

\* Deriving Disambiguous Queries in a Spoken Interactive QA System

### 3.2 音声認識部

話し言葉コーパスから構築された、各状態 16 混合ガウス分布の不特定話者音素文脈依存 HMM(3000 状態)の音響モデル [4] を適用した。毎日新聞 1994 年～2001 年までの 8 年分を JTAG により形態素解析し、語彙数 20k で作成した bigram, trigram を言語モデルとして適用した。単語グラフを中間表現とする 2 パスデコーダを用いた。

### 3.3 応答対話文生成部

単語の汎用度スコア  $I$  は、毎日新聞 8 年分を適用した。質問文作成に用いた疑問詞は、「誰の」「何の」等、これまで SAIQA に入力された疑問詞 82 個を適用した。質問文としての適正を示す言語モデル  $L$  には、毎日新聞 3 年分と質問文 1000 を併せて学習した trigram を適用した。係られ度スコア  $D$  については、京大コーパスに基づき学習した SCFG[3] を適用した。

### 3.4 自動応答対話文の評価方法

自動生成される応答対話文は、効率的に答えを絞り込む役割と、認識誤りによる質問の曖昧性を解消する役割とがある。前者の役割を評価するには、対話の状態、補完すべき情報の候補、および対話戦略を考慮する必要がある。本稿では、評価を単純化するため、質問に対する答えをコーパスから一意に同定できる質問文を対象とした。従って、補完されるべき情報は質問文中にある。つまり、100%の精度で音声認識が行われれば、曖昧性の解消を必要とせず、完全に情報が補完されたことになる。一方、認識誤りがある場合、欠落した原文の情報の補完を促す応答対話文が生成できれば、ユーザの答えにより最初の質問の答えを絞り込める。

本実験では、自動生成された応答対話文を以下の 3 段階で評価した。ただし、単語正解精度 100%の質問文は評価の対象外とした。

- 適切 : 認識誤りによる情報の欠落を補完
- 補助 : 間接的な情報の補足
- 不適切 : 欠落した情報とは無関係

### 3.5 評価結果

自動生成された応答対話文の例を表 1 に示し、評価結果を表 2 に示す。表 2 に示すように、誤認識の含まれる質問に対し、約 6 割の質問で 10-best 中に曖昧性を解消する応答対話文が生成できたことが分かる。生成された応答対話文を分析したところ、以下の問題点が明らかになった。

まず、応答対話文の生成法として、疑問詞の前方接続のみを適用したが、「 $L$  したのは誰？」のような後方接続も必要である。また、全ての文節(名詞節)に対し、全ての疑問詞を組み合わせたが、意味カテゴリに基づく疑問詞の絞り込みが有効と考えられる。さらに、固有名詞の同定を行わなかったが、曖昧性の少ない固有名詞には特殊な処理も必要である。

表 1. 自動応答対話文の例

読み上げ質問	キューバからの医療援助に対し、地元の医師会が撤収を要請しているのはどこの国ですか？
音声認識結果	キューバからの医療援助に対し、地元の医師会が停止を要請しているのはどこの国ですか？
自動応答対話文	何を要請しているのですか？

表 2. 自動応答対話文の原文の情報に対する補完の評価

被験者	全認識率	テストセット認識率	適切	補助	不適切
A	82%	76%	22	8	22
B	81%	77%	38	3	15
C	79%	73%	31	8	14
平均	81%	75%	30	6	17

表中の数値は文数を示す。

一方、以下に示す誤認識への対処により、応答対話文が改善し得る。まず、「何」「誰」といった疑問詞を誤認識した場合「誰ですか？」というユーザの質問に対し、「誰ですか？」という応答対話文が重複して生成されることがある。そこで、質問の認識結果を、Support Vector Machine (SVM) により質問タイプ分類 [5] をし、ユーザが質問していると推測される質問タイプを考慮することで、改善できるものと考えられる。

また、音声認識結果を対象とすることから、誤認識された単語による誤応答文が生成されている。誤認識された質問文から適正な応答対話文を自動生成するため、誤りによる不要な情報を排除するスクリーニング処理が必要となる。この問題には、音声認識結果から文意を保持した部分単語列を抽出する音声自動要約の枠組を適用できる [3]。

## 4. まとめ

音声インタラクティブ QA システムにおける入力質問文の曖昧性解消を目的とした自動応答対話文生成法を提案した。誤認識を含んだ質問文から応答対話文を自動生成し、その結果生成された文章の評価を行った。約 6 割の質問文で、10-best 中に曖昧性の解消を促す応答文を生成できた。今後、コーパス中の単語間の係り受け関係や意味カテゴリを考慮し、さらに曖昧性解消の性能を改善していきたい。

### 参考文献

- [1] 佐々木, 磯崎, 平, 平尾, 賀沢, 鈴木, 国領, 前田, “SAIQA: 大量文書に基づく質問応答システム”, 情報処理学会 情報学基礎研究会 F1-64, pp.77-82(2001).
- [2] 秋葉, 伊藤, 藤井, 石川, “音声入力による質問応答システムのための音声認識用言語モデルの検討”, 言語処理学会第 8 回年次大会発表論文集, pp.244-247(2002).
- [3] C. Hori and S. Furui, “A New Approach to Automatic Speech Summarization,” To appear in the IEEE Transactions on Multimedia (2002).
- [4] 篠崎, 細川, 古井, “話し言葉コーパスを用いた音声認識の検討”, 音講論, Vol.1, 1-3-14, pp.31-32(2001).
- [5] 鈴木, 佐々木, 前田, “統計的機械学習を用いた質問タイプ同定法”, 情報科学技術フォーラム (FIT) 掲載予定, (2002).