

## 文書検索タスクにおける 検索要求明確化のための対話戦略

翠 輝久 河原 達也

京都大学 情報学研究科 知能情報学専攻  
〒 606-8501 京都市左京区吉田二本松町  
e-mail: misu@ar.media.kyoto-u.ac.jp

あらまし 一般的な文書を検索するタスクにおいて、ユーザの検索要求を明確化するための対話戦略を提案する。テキストで記述された知識ベースを自然な音声により検索する際には、断片的な発話や漠然とした検索要求が入力されやすく、結果として多くの候補が該当する場合が多い。そこで、係り受け情報を利用したボトムアップの知識と知識ベースから得られるトップダウンの知識を利用して、ユーザの意図を詳細化するための質問を生成する。質問を選択する際の基準として、候補の絞り込みの効果の大きさを情報ゲインとして計算する。これにより、効率的な検索結果の絞り込みを可能にする。ソフトウェアサポートを行うシステム「音声版ダイアログナビ」を拡張し、以上の対話戦略を実装した。評価の結果、音声認識結果のままでは正しい検索結果が得られないサンプルの過半数で、正しい検索結果を得ることができた。

## Dialogue Strategies to Clarify User's Query for Document Retrieval Task

Teruhisa Misu Tatsuya Kawahara

School of Informatics, Kyoto University, Kyoto 606-8501, Japan  
e-mail: misu@ar.media.kyoto-u.ac.jp

**Abstract** The paper addresses a dialogue strategy to clarify user's queries for document retrieval systems. In spoken dialogue systems, vague queries and fragmental utterance are often input. As a result, so many items are usually obtained. Thus, we introduce an efficient dialogue strategies to narrow down the items based on bottom-up knowledge derived from dependency structure analysis together with the top-down knowledge acquired from a knowledge base. We also introduce information gain as a criterion of efficiency of the possible questions. As a result of the experimental evaluation, correct retrieval results were obtained for more than half of the samples that had not been correctly handled by the baseline system.

## 1 はじめに

大語彙連続音声認識と情報検索技術の高度化に伴い、音声対話システムの研究対象は関係データベースの検索から、Web のテキストや新聞記事などの一般的な文書の検索へと広がりつつある [1][2][3]。このようなシステムでは、ユーザ発話の音声認識結果と文書のマッチングが行われ、そのゆがみが高いものが提示される。我々も、大規模なソフトウェアサポートマニュアルを検索するシステムとして「音声版ダイアログナビ」を構築し、試験運用を行っている [4]。

このような大規模な知識ベースを対象としたシステムにおいて、ユーザが意図する文書を正しく同定するためには、発話文中に検索に必要な情報が十分に含まれている必要がある。しかしながら、話し言葉音声は発話が断片的になったり、検索要求が漠然なまま入力されることが多い。また、音声認識の際に生じる認識誤りにより、検索に必要な情報が欠損する可能性がある。このような場合は、検索の手がかりとなる情報が少なくなるため、結果として非常に多くの候補がマッチング結果として得られることになる。そのため、ユーザに聞き返しを行い、不足している情報や音声認識誤りにより失われた情報を補完することで、候補を絞り込む必要がある。

従来の音声対話システムで主に扱われてきたホテル検索 [5] などの関係データベース検索型のタスクでは、データベースのフィールド（「地域名」や「料金」など）が明確に与えられるため、未確定のフィールドの内容を確認することにより、候補を絞り込むことができた。また、家電製品のマニュアルのように、目次などにより階層構造化されている文書を検索する場合には、その構造を利用して効率的に絞り込みを行う手法が提案されている [6]。しかし、階層構造を持たない一般的な文書に対して、このような手法を適用することはできない。

本研究では、一般的な大規模知識ベースを対象として、検索結果の絞り込みを行うために対話的に検索要求を明確化する手法を提案する。まず、知識ベースから抽出されるボトムアップの知識と、人の知見などを利用したトップダウンの知識を併用することにより、可能な質問集合を用意する。その際に、検索結果に対する情報ゲインが大きくなるように動的に質問を選択することで、効率のよい対話を実現する。さらに、生成した質問に対するユーザの回答を

表 1: ソフトウェアサポート用知識ベース

知識ベースの種類	件数	文字数
用語集	4,707	約 70 万
ヘルプ集	11,306	約 600 万
サポート技術情報	23,323	約 2200 万
合計	39,336	約 4000 万

### Windows XP で音声認識を使用する方法

この資料は以下の製品について記述したものです。

- Microsoft Windows XP Professional
- Microsoft Windows XP Home Edition

**概要** この資料では、Windows XP で音声認識を使用する方法について説明しています。Microsoft Office XP の音声認識をインストールしているか、または、Office XP がインストールされたコンピュータを新たに購入した場合は、すべての Office アプリケーションや、音声認識が利用可能なその他のアプリケーションで音声認識を使用できます。

**詳細** 音声認識は、音声をテキストに変換するオペレーティング システムの機能です。音声認識エンジンと呼ばれる内部ドライバによって、単語が認識され、テキストに変換されます。音声認識エンジンは、..

図 1: ソフトウェアサポート用知識ベースの例

用いて、検索文を更新・整形することも考える。

## 2 音声による大規模文書検索システム

本研究で検索対象とする文書は、マイクロソフト社のソフトウェアサポート用知識ベースであり、この概要を表 1 に示す。これらは自然言語によって記述されている。サポート技術情報の例を図 1 に示す。

これまで我々は、音声による検索要求によりこの知識ベースを検索するシステムとして「音声版ダイアログナビ」を開発してきた [4]。音声版ダイアログナビでは、音声認識結果に対して文節ごとに知識ベースとの整合性の尺度となる検索整合度、音声認識結果の N-best 候補により計算される検索重要度によ

り、検索の前後で確認を行い音声認識誤りや、音声言語の冗長性に対処してきた。しかし、音声版ダイアログナビは、基本的に一問一答型のシステムであり、漠然とした検索要求や、音声認識誤りにより一部の文節が削除された場合には、多数の候補がマッチングし、そのまま提示されていた。今回、この音声版ダイアログナビを拡張し、検索要求を明確化するための対話戦略を実装する。

### 3 漠然とした検索要求を明確化するための対話戦略

音声版ダイアログナビを試験運用したところ「印刷できません」といった漠然とした発話が多く見受けられた。また、テキスト入力に比べて単語のみによる入力も多い。このような検索要求に対しては、いくつもの状況・原因が考えられるため、一意に回答となる候補を同定することは不可能である。また、十分に具体的な内容を発話した場合でも、音声認識誤りや、発話中の表現と知識ベースの表現との不一致が原因で、発話の全てを知識ベースとのマッチングに利用できないことがある。このような場合に膨大な数の候補が該当する可能性が高い。

しかし、大規模なテキスト文書を検索するタスクでは、候補を絞り込むための手順をあらかじめ用意することは困難である。そこで本研究では可能な質問集合を用意しておいて、検索結果に基づいて動的に選択する枠組を考える。前提条件を満たす質問の中から最も効果的なものを選択するために、その質問により得られる情報ゲインを定義する。しかしながら、ある質問に対して複数のカテゴリに属するテキストが存在したり、逆にどのカテゴリにも属さないものが存在するため、この計算に単純なエントロピーを利用することはできない。そこで、以下の式により質問  $S$  による情報ゲイン  $H(S)$  を定義する。

$$H(S) = - \sum_{i=0}^n P(i) \cdot \log P(i)$$

$$P(i) = \frac{|C_i|}{\sum_{i=0}^n |C_i|}$$

$$|C_i| = \sum_{D_k \in i} CM(D_k)$$

ここで、 $D_k$  は最初の検索要求で検索された  $k$  番目の候補であり、 $CM(D)$  は文書  $D$  のマッチングのゆわ度である。 $C_i$  は質問  $S$  で分類されるカテゴリ  $i$  に

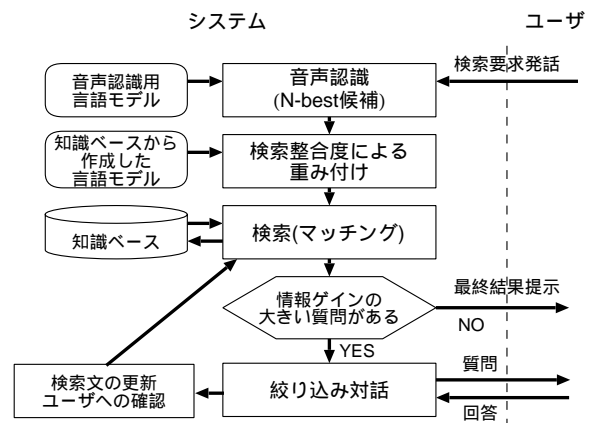


図 2: 絞り込み戦略の概要

該当する文書の数に相当するが、マッチングのゆわ度で重み付けされている。また、どのカテゴリにもマッチしないテキストはカテゴリ 0 に属するものとする。

以上の戦略を組み込んだシステムの処理の流れは以下の通りである。

1. 認識結果に対して検索整合度を計算し、文節に対して検索重みを付与する。(これにより、認識誤りや検索に不要な箇所を除去する。)
2. ダイアログナビを用いて知識ベースを検索する。
3. 全ての質問候補に対して、情報ゲインを計算する。
4. 情報ゲインが大きく、候補の絞り込みが可能である質問が存在する場合には、その中から最も情報ゲインが大きい質問を生成する。それ以外の場合には、ユーザに最終結果を提示する。
5. 質問に対するユーザの回答を用いて、検索文を更新する。
6. 3に戻る。

これら全体の流れを図 2 に表す。また、次章でその詳細について述べる。

## 4 ボトムアップの知識とトップダウンの知識を利用した質問生成

本研究では、マッチングしたテキストの文から得られるボトムアップの知識と人間の知見や知識ベー

スから抽出できるトップダウンの知識の両方を利用して、以下の3つの手法により質問候補を用意する。このうち、ボトムアップな知識を利用した質問として「係り受け情報を利用した質問」、トップダウンの知識を利用した質問として「人間の知見を利用した質問」、「知識ベースのメタデータを用いる質問」を考える。

#### 4.1 係り受け情報を利用した質問 (手法1)

検索要求に含まれる単語を係り受け対象とする語句が明示的に示されていない場合に、それを明確化するための質問を行う。例えば、「削除する」という単語を含む候補には「アプリケーションを削除する」や「アドレス帳を削除する」のように数多くの候補が存在する。そこで、これらの単語の係り元を同定することで、検索要求を明確化することができる。(テキスト版の)ダイアログナビにおいても、ユーザの検索要求にマッチした知識ベース中の文の間で、検索文にマッチしなかった文節を抽出して提示することで、ユーザの候補選択を支援する手法が実装されている[7]。本研究では、マッチした知識ベース中に含まれる特定の単語に着目し、その係り元を明示的に確認することで、検索要求の明確化を目指す。ただし、その際に全ての単語について確認すればよいわけではない。例えば「コンピューターシャットダウン」という係り受けペアのように係り元の語句がほぼ一意に同定できる語句も存在し、そのような語句に対して質問するのは冗長である。そのため、どのような単語を対象として質問を生成するかは、以下の手順により決定する。

1. 知識ベース、及び「マイクロソフト話し言葉検索<sup>1</sup>」に寄せられた検索文から全ての係り受けペアを抽出する。
2. 単語  $m$  に係る語句が単語  $i$  である頻度割合  $P(i)$  を求め、それによりエントロピー  $H(m)$  を計算する。

$$H(m) = - \sum_i^N P(i) * \log P(i)$$

今回は、このエントロピー  $H(m)$  が大きい上位50単語を確認することとした。この手法により生成する具体的な質問の例を表2に示す。なお、表中の「適

<sup>1</sup> マイクロソフト話し言葉検索は、自然言語文によりソフトウェアサポート知識ベースを検索するシステムである。

用可能な知識ベースの割合」は、それぞれの質問を適用可能な知識ベースが全知識ベースに占める割合であり、「情報ゲイン」は、適用可能な知識ベースすべてを対象として情報ゲインを計算した値である。

#### 4.2 人間の知見を利用した質問 (手法2)

本研究で対象とするソフトウェアサポートは、コールセンターのオペレータによっても行われている。そこにおいて蓄積されている人間の知見に基づいて質問を用意する。今回、この手法による質問を3種類用意した。例えば「その症状はいつ発生しましたか?」という質問では、「~した際」「~した後」という文節の直前に現れる名詞を同定するものとする。この手法による質問の例を表3に示す。この手法による質問は、ユーザの発話に対する検索結果に規則が適用可能なものが多く含まれる場合に、大きな情報ゲインが得られる。しかし、知識ベースの表現が人手により記述した規則と一致している必要があるため、実際に規則を適用できる場合は多くない。

#### 4.3 知識ベースのメタデータを用いた質問 (手法3)

検索対象の知識ベースに付与されているタグを利用して質問を生成する。一般的に、新聞記事や図書のような大規模な資料や文書を扱う際には、管理・運用を効率的に行うために、各文書にカテゴリ情報などのメタデータが付与されることが多い。本研究の検索対象のソフトウェアサポート用知識ベースにおいても、図1の3,4行目のように、その文書がどのプロダクトについて記述したものであるかを表すメタデータが付与されているものがある。このように付与されているタグ情報を利用して、ユーザの検索要求がどれに該当するかをたずねる質問を生成する。このようなタグは、多くの知識ベースに付与されているが、1つのテキストに対して複数のタグがふられているケースも多く、絞り込みの効果はタグの構成に大きく依存する。本研究で使用する質問の数は16であり、その例を表4に示す。

以上の質問を利用して、実際の検索結果に対して情報ゲインを計算し、システムが質問を生成する対話例を図3に示す。

表 2: 係り受け情報を利用した質問の例

質問の内容	適用可能な知識ベースの割合	情報ゲイン
何を削除したか	2.15 (%)	7.44
何をインストールしたか	3.17 (%)	6.61
何を挿入したか	1.12 (%)	7.12
何を保存したか	1.81 (%)	6.89
何のファイルか	0.94 (%)	6.00
何が文字化けしたか	0.19 (%)	5.62

表 3: 文書から抽出される情報を使った質問の例

質問の内容	適用可能な知識ベースの割合	情報ゲイン
症状はいつ発生したか	15.40 (%)	8.08
発生したエラーの内容	2.63 (%)	8.61
行うことの詳細な内容	6.98 (%)	8.04

表 4: 知識ベースの情報を用いる質問の例

質問の内容	適用可能な知識ベースの割合	情報ゲイン
Windows のバージョン	30.03 (%)	2.63
Application の種類	30.28 (%)	2.31
Excel のバージョン	4.13 (%)	2.44
Word のバージョン	3.76 (%)	2.71
Office のバージョン	3.63 (%)	2.11
Outlook のバージョン	3.11 (%)	2.18

## 5 システムの質問に対するユーザの回答を利用した検索文の更新

このように生成した質問に対するユーザの回答を利用して単純に候補を絞り込むこともできるが、最初のユーザ検索発話を更新・整形し、再度マッチングを行う方が、検索文中の語句の係り受け関係などの情報を利用して正確なマッチングができると期待される。また、最終的な検索要求をユーザに確認する際に、システムの質問への回答を羅列して確認するよりも更新した検索文を提示するほうがわかりやすいと考えられる。そこで、質問に対するユーザの回答を用いて最初の検索発話を更新する。本研究では、質問手法ごとにあらかじめ検索文の更新規則を用意し、それに従ってユーザの検索文を更新する。それぞれの質問手法ごとの更新規則は以下の通りである。

### 1. 係り受け情報を利用した質問

ユーザの回答を係り先となる単語の直前に追加する。たとえば「何を削除するのですか?」と

S1:	ご質問をどうぞ。
U1:	文字化けしました。
(検索結果):	1. [閉じる] ボタンや [最大化] ボタンが文字化けした場合の対処方法 2. [OL 2000] Outlook Today 画面の文字化け 3. テキストの挿入処理で、日本語入力を行なうと文字化けする 4. [ XL2002 ]ピボットテーブルを作成するとドロップダウンリストのアイテム名が文字化けする 5. V B アプリが文字化けする 6. 通知機能のメールが文字化けする ...
(情報ゲイン計算):	質問候補 1: 「何が文字化けしましたか?」 (情報ゲイン 5.27) 質問候補 2: 「その症状はいつ発生しますか?」 (情報ゲイン 2.47) 質問候補 3: 「Windows のバージョンを教えてください」 (情報ゲイン 1.43) ...
S2:	(最も情報ゲインが大きい質問を選択) 何が文字化けしますか?
U2:	ウインドウのボタンです。
S3:	(検索文の更新) 「ウインドウのボタンが文字化けする」 で検索します。

図 3: システムの質問生成の例

という質問に対する回答は、検索文中の「削除」という単語の直前に挿入する。

### 2. 人間の知見を利用した質問

生成する質問ごとに、質問へのユーザの回答を挿入する場所をあらかじめ指定しておく。たとえば「エラーはいつ出ますか?」という質問に対する回答は検索文の「エラー」という単語の直前である。

### 3. 知識ベースのメタデータを利用した質問

ユーザの検索文の文頭に「(タグ名)で」の文節を付与する。また、プロダクトのバージョンを確認する質問の場合は、検索文の「(プロダクト名)」の部分「(プロダクト名)(バージョン情報)」に置き換える。

これらの規則を用いて、システムが実際に検索文を更新しながら対話を行う例を図 4 に示す。

## 6 対話戦略の評価

提案する対話戦略の有効性を確かめるために、評価実験を行った。テストセットとして、音声版ダイ

S1:	ご質問をどうぞ。
U1:	ファイルが開けません。
S2:	何のファイルですか？
U2:	Excel のファイルです。
(更新結果):	「Excel のファイルが開けません」
S3:	Excel のバージョンを教えてください。
U3:	Excel 2002 を使っています。
(更新結果):	「Excel2002 のファイルが開けません」
S4:	その症状はいつ発生しますか？
U4:	エクスプローラから開こうとしたときです。
(更新結果):	「エクスプローラから開こうとした時に Excel2002 のファイルが開けません」
S5:	「エクスプローラから開こうとした時に Excel2002 のファイルが開けません」 で検索します。

図 4: ユーザの回答を用いた検索文の更新の例

アログナビの評価用に収集した発話データ [4] の中から、音声認識誤りあるいは検索要求の漠然性が原因で検索が失敗した発話 102 サンプルを用いた。これらのサンプルに対して、提案手法により、ユーザへの質問及び検索文の更新を行うことで、ユーザの要求に沿った回答を提示できるかを検証した。なお、今回の実験ではシステムの生成する質問に対して、ユーザは適切に回答するものとし、その際に音声認識誤りは起きないものとして評価を行った。ユーザへの質問とその回答を用いた検索文の更新は、情報ゲインが大きい質問がなくなるまで繰り返すものとし、そのような質問がなくなった時点でユーザに結果を提示するものとした。

評価実験の結果、認識結果をそのまま用いても検索が成功しなかった 102 発話のうち、全体の 55.9%にあたる 57 発話で検索が成功した。システムの生成した質問の回数を調べたところ、1 検索要求あたりの質問生成回数は、2.3 回であった。なお、2 章で述べたこれまでの音声版ダイアログナビの対話戦略 [4] では、検索成功率は 25.4%であった。この結果から、提案する対話戦略により、漠然とした検索要求や、音声認識誤りに対して、ユーザとの対話により適切に情報を補完できたといえる。

さらに、システムが最終的に提示した候補の中での正解となる候補の順位を調べたところ、検索が成功したサンプル全体の約 7 割にあたる 37 発話で正解の知識ベースが最もマッチングのゆう度が高い候補として提示されていた。これは、提案する戦略により単純に候補を絞り込むだけでなく、検索文を更新・整形することで、より正確にユーザの意図を反映した検索が行えていることを示している。

## 7 むすび

本研究では、ソフトウェアサポート用知識ベースを対象とした検索タスクにおいて、ユーザの漠然とした検索要求に対して対話的に意図を明確化するための対話戦略を提案した。そのために、係り受け関係を利用した質問、人間の知見を利用した質問、知識ベースのメタデータを利用した質問を用意し、情報ゲインの基準により選択する枠組みを採用した。さらに、その回答により検索文を更新することも行った。音声認識結果をそのまま用いても検索が成功しなかった 102 発話により評価を行った結果、これらの過半数で回答となる知識ベースを提示することができた。

## 謝辞

本研究に対し、多大な協力を頂いた東京大学の黒橋禎夫助教授、清田陽司助手、及びマイクロソフト株式会社の木戸冬子氏に深く感謝します。

## 参考文献

- [1] A. Fujii and K. Itou. Building a test collection for speech-driven web retrieval. In *Proc. EUROSPEECH*, Vol. 2, pp. 1153–1156, 2003.
- [2] C. Hori, T. Hori, H. Isozaki, E. Maeda, S. Katagiri, and S. Furui. Deriving disambiguous queries in a spoken interactive ODQA system. In *Proc. ICASSP*, Vol. 1, pp. 624–627, 2003.
- [3] 西崎博光, 中川聖一. 音声認識誤りと未知語に頑健な音声文書検索手法. 電子情報通信学会論文誌, Vol. J86-DII, No. 10, pp. 1369–1381, 2003.
- [4] 翠輝久, 駒谷和範, 清田陽司, 河原達也, 木戸冬子. 音声対話による大規模知識ベース検索システム -音声版ダイアログナビ-. 情報処理学会研究報告, 2004-SLP-52-4, 2004.
- [5] 駒谷和範, 鹿島博晶, 田中克明, 河原達也. 複合的言語制約に基づくキーフレーズ検出を用いた汎用的なデータベース検索音声対話プラットフォーム. 情報処理学会論文誌, Vol. 44, No. 5, pp. 1333–1342, 2003.
- [6] 伊藤亮介, 駒谷和範, 河原達也. 機器操作マニュアルの知識と構造を利用した音声対話ヘルプシステム. 情報処理学会論文誌, Vol. 43, No. 7, pp. 2147–2154, 2002.
- [7] 清田陽司, 黒橋禎夫, 木戸冬子. 大規模テキスト知識ベースに基づく自動質問応答 -ダイアログナビ-. 自然言語処理, Vol. 10, No. 4, pp. 145–175, 2003.