

ソフトウェアサポートタスクにおける効率的な確認を行うための対話戦略

Dialogue Strategies for Generating Efficient Confirmation in a Software Support Task

駒谷 和範

Kazunori Komatani

翠 輝久

Teruhisa Misu

河原 達也

Tatsuya Kawahara

奥乃 博

Hiroshi G. Okuno

1 はじめに

音声による自然言語テキストに対する検索システムの研究が行われつつある [1, 2]。音声対話システムでは、発話の音声認識結果から、ユーザの意図を解釈する手続きが不可欠である。従来の音声対話システムでは、ユーザ発話からキーワードを抽出することで意図を解釈し、それが同定できなければ確認するという方法論を採用することができた。しかし、マニュアル [3] や Web ページなど、テキストで記述された大規模な知識ベースを検索する際には、キーワードの集合を明確に定義することが不可能であるため、発話を自然言語として解釈する必要がある [4]。

このようなシステムを実現するには、単純に音声認識結果をそのまま検索すればよいわけではない。問題点として、(1) 音声認識誤りへの対処 (2) 音声言語表現に含まれる冗長性が挙げられる。音声認識誤りに対しては、キーワードが明確に定義されている場合には、それらを一つずつ確認することができるが、一般的な自然言語入力の場合は難しい。また、音声は入力容易であるため、ドメイン外発話や多様な文末表現など、タスク遂行には直接関係がない冗長な部分も入力されやすい。したがってタスク遂行に有用な部分を自動的に判別する枠組みが必要である。

本研究では、検索対象のみから学習した統計的言語モデルによる尤度と、音声認識結果の N-best 候補に対する検索結果の両方を用いて、音声認識結果の各文節が検索に必要かどうかを判定する。これにより、検索に有用な部分とそうでない部分を切り分け、効率よく確認を行う手法について述べる。

2 検索整合度と検索重要度を用いた対話戦略

2.1 検索整合度の定義と検索への重み付け

ユーザ発話の中から、検索に有用な部分とそうでない部分を切り分ける基準として、各文節に対する単語パープレキシティを用いる。単語パープレキシティの計算に用いる言語モデルは、検索対象である知識ベースから学習したもので、音声認識時に用いるものとは異なる。したがって単語パープレキシティ PP は、検索対象との整合性を示す尺度である。

検索対象である知識ベースから学習した言語モデルによるパープレキシティが低いということは、知識ベースにおけるその単語列の出現頻度が高いことを意味する。逆に、音声認識結果中の認識誤りである箇所は文脈的に不自然である場合が多く、またドメイン外発話の単語列も知識ベースの中では出現確率が低いため、パープレキシティは高く

なる。これにより、認識誤り部分や、認識したが検索には重要でない部分を同時に検出できる。

このパープレキシティを、以下の関数により 0~1 の間の値に変換したものを検索整合度 (relevance score; RS) とする。これは、重要語句の確認や、知識ベースとのマッチング時の各文節に対する重みとして用いる。

$$RS = \frac{1}{1 + \exp(\alpha * (\log PP - \beta))}$$

2.2 検索結果を用いた検索重要度の計算

音声認識結果の N-best 候補に対する検索結果を用いて検索重要度を定義する。音声認識結果の N-best 候補には音声認識の過程で曖昧であった部分が現れるが、この部分に対して確認すべきかどうかを規定している。

まず音声認識結果の N-best 候補間の相違箇所を同定する。この N-best 候補それぞれについて実際に検索を行い、結果が異なる場合にその相違の大きさを検索重要度として定義する。ここでは検索重要度は、検索の結果得られるスコアの高い 5 候補の中で、異なる候補の割合とする。例えば、検索結果の上位 5 候補中、4 候補が異なる場合には、音声認識結果中の相違部分の検索重要度は 0.8 となる。

2.3 検索整合度と検索重要度を用いた対話管理

認識誤り箇所が常に検索に悪影響を与えるとは限らないため、認識誤りの可能性が高い部分を全て確認するのは非効率である。そこで、文節ごとの認識誤りによる損失を考慮して、確認の方法を切り替える。認識誤りが検索に決定的な影響を与えることが予測される語句は、検索を実行する前にユーザに確認する¹。この事前確認を行うか否かの判断に検索整合度を利用し、あるしきい値以下のものには確認を行わない。しきい値は収集したデータを用いて実験的に決定した ($PP = 50$)。ユーザからの回答により認識誤りであるとされた文節は認識結果から取り除かれ、残りの部分が次のマッチングモジュールに渡される。

次に、常に決定的な影響を与えるわけではないが、結果として検索結果に影響を与える箇所を検索重要度を用いて同定し、検索後に確認する。検索重要度は、音声認識の N-best 候補の相違部分に対して計算されるが、ある部分に対する検索重要度がしきい値を超えている場合には、その相違部分をユーザに対して表示し確認する。しきい値は 0.5 とした。ユーザがその中から適切な候補を選択すると、対応する検索結果が表示される。検索重要度がしきい値以

¹ソフトウェアサポートタスクでは、プロダクト名がこれにあたる。現時点は人手で事前に用意しているが、現在これを自動的に得る方法を検討中である。

⁰京都大学 大学院 情報学研究所 知能情報学専攻

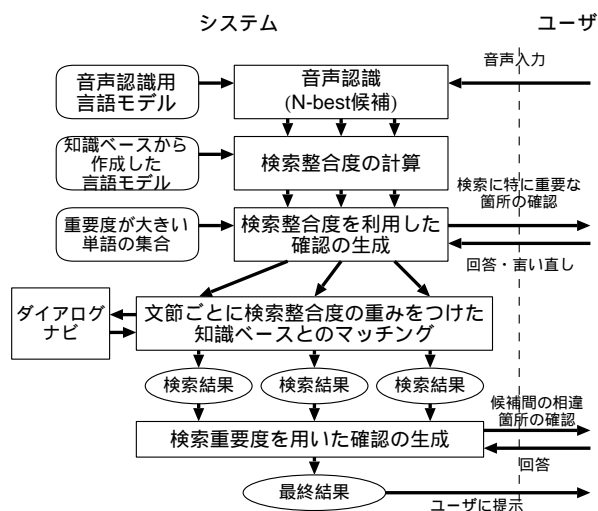


図 1: 確認対話戦略を導入した検索システムの概要

下の場合には確認を行わず、第 1 候補による検索結果をそのまま表示する。候補が全て適当でなかった場合には、現在の候補を棄却し再発話することになる。

これらの確認戦略を含めた全体の流れを図 1 に示す。

3 評価実験

評価用データは本システムを利用したことのない 10 名 (男 7, 女 3) の被験者により収集した。検索を行うバックエンドにはダイアログナビ [5] を使用した。設定した想定場面に基いて各 11 課題、これとは別に、自由に 3 課題に対して検索を行ってもらった。ただし、検索結果として、ふさわしい候補が提示されない場合には各課題につき 3 度まで言い直しを許した。その結果、合計 139 課題、235 発話を得た。10 人の発話の音声認識率は平均で 65.6% である。収集した音声データに対して、以下の 3 つの条件で検索実験を行った。検索の成功率を表 1 に示す。なお、ユーザに最終的に提示した候補の中に最初の質問の答えとなる候補が含まれていた場合に検索成功としている。

1. ユーザ発話の正確な書き起こし (人手で作成) を用いて検索した場合
2. 音声認識結果の第 1 候補をそのまま検索した場合
3. 検索整合度と検索重要度の両方を用いて検索を行い、生成する確認に対してユーザが適切に回答した場合 [提案手法]

全 235 発話の成功率では、提案手法により検索を行った場合は、音声認識結果の第 1 候補をそのまま用いて検索を行った場合よりも大幅に検索の成功率が上昇している。また、各課題のユーザの最終発話のみを比較した場合には、提案手法での成功率は、発話の正確な書き起こしを利用して検索を行った場合に近づいている。

次に、システムが生成した確認に関して検証を行った。生成された確認の回数は 126 回である。これは、1 課題あたりおおよそ確認が 1 回行われていることになる。このうち、検索整合度を用いた事前確認の回数は 51 回あり、検索重要度を用いた事後確認が 75 回であった。検索整合度を用いた確認により、確認を行わない場合と比べて 9 発話で検索が成功した。同様に、検索重要度を用いた確認によ

表 1: 検索成功率

評価対象	発話数	条件 1	条件 2	条件 3
全発話	235	184 (78.3%)	104 (44.2%)	134 (57.0%)
各課題の最終発話	139	125 (89.9%)	93 (66.9%)	118 (85.0%)

表 2: 音声認識の信頼度を用いた確認戦略との比較

	提案手法	信頼度 ($\theta_1 = 0.4$)	信頼度 ($\theta_1 = 0.6$)	信頼度 ($\theta_1 = 0.8$)
確認回数	126	76	173	284
検索成功率	134 (57.0%)	113 (48.1%)	125 (53.2%)	129 (54.9%)

り 6 発話で検索成功回数が増加した。

提案手法の確認回数を評価するために、音声認識結果の N-best 候補から計算される信頼度 [6] を用いた確認を行う場合との確認回数、検索成功率を比較した。確認を行うための信頼度の閾値 θ_1 として、0.4, 0.6, 0.8 の 3 通りを用いた。信頼度が閾値以下の自立語を確認するものとし、それが誤認識されたものであった場合には、その単語を含む文節を棄却して検索した。

この結果を表 2 にまとめる。提案手法は、音声認識の信頼度の閾値 θ_1 が 0.8 の場合に比べて確認回数を半分以下に抑えながら、高い検索成功率を得ている。以上より、提案手法の確認が効率的であることが確かめられた。

4 おわりに

大規模な知識ベースを対象とした音声による検索タスクにおいて、ユーザの多様な発話に対して頑健に検索を行うための対話戦略を提案した。音声認識結果に対して検索整合度と検索重要度の 2 つの尺度を導入し、音声認識誤りや余分な入力を含む部分に対して効率よく確認を行う戦略を考案・実装し、評価実験によりその有効性を確認した。

謝辞

本研究は、東京大学の黒橋禎夫助教授、清田陽司氏、マイクロソフト株式会社の木戸冬子氏との共同研究である。各氏の貴重な貢献に感謝します。

参考文献

- [1] S. Harabagiu, D. Moldovan, and J. Picone. Open-domain voice-activated question answering. In *Proc. COLING*, pp. 502-508, 2002.
- [2] 藤井敦. 音声による言語バリアフリーな他言語情報アクセス. 情報処理学会研究報告, SLP-44-33, 2002.
- [3] 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.
- [4] 駒谷和範, 河原達也, 清田陽司, 黒橋禎夫, Pascale Fung. 柔軟な言語モデルとマッチングを用いた音声によるレストラン検索システム. 情報処理学会研究報告, 2001-SLP-39-30, 2001.
- [5] Y. Kiyota, S. Kurohashi, and F. Kido. "Dialog Navigator": A question answering system based on large text knowledge base. In *Proc. COLING*, pp. 460-466, 2002.
- [6] 駒谷和範, 河原達也. 音声認識結果の信頼度を用いた効率的な確認・誘導を行う対話管理. 情報処理学会論文誌, Vol. 43, No. 10, pp. 3078-3086, 2002.