

音声言語による人間 - 機械対話システムの研究

Research on Man-Machine Dialogue System through Spoken Language

(研究プロジェクト番号 : JSPS-RFTF 96R15201)

プロジェクトリーダー

藤崎 博也 東京理科大学基礎工学部・教授

コアメンバー

古井 貞熙 東京工業大学大学院情報理工学研究科・教授

中川 聖一 豊橋技術科学大学情報工学系・教授

堂下 修司 龍谷大学理工学部・教授

広瀬 啓吉 東京大学大学院新領域創成科学研究科・教授

白井 克彦 早稲田大学理工学部・教授

板橋 秀一 筑波大学電子・情報工学系・教授



1. 研究目的

社会の情報化が急速に進展し、マルチメディアによる情報の提示の多様化や、インターネットによる情報の授受の広汎化が進む現在でも、人間にとって最も容易で、しかも常に使用し得る媒体としての音声言語の優位性はゆるがない。機械による知能がさらに高度化し、人間の知的なパートナーとして重要な役割を果たす『人間-機械共生』の社会では、人間と機械が音声言語を用いて自由に対話し得ることが不可欠である。

本研究プロジェクトは、これを可能とする要素技術を開発し、それらを統合した人間-機械対話システムを実現することを目的とし、その具体例として**音声言語を主要な媒体とする学術情報検索システム**を世界ではじめてとりあげた。これは、従来の対話システムよりも格段に高度の知的処理を必要とするため、多くの独創的技術の開発をうながし、かつ、その成果が学術上・社会生活上の進歩に貢献するところが大きいのである。

2. 研究成果概要

2.1 ユーザの心のモデルに基づく対話管理手法

図1は本研究で構築した人間-機械音声対話に基づく学術情報検索システムの構成を示す。従来の対話システムにおける対話管理手法は、対話を現象的にモデル化するものが多かったが、本研究では、ユーザとシステムとを、そ

れぞれ有限状態の確率的オートマトンによって近似的に表現し、それらを用いて対話管理を行なう新しい手法を考案し、採用した。なお、これは通常のオートマトンとは異なり、現状態・外部入力・次状態・状態遷移に伴う出力、の4つ組により規定されるもので、ユーザの対話の相手はシステムのユーザインターフェース部、ユーザインターフェース部の対話の相手はユーザ及びシステムの情報検索部である。対話をこのように定式化することによって、システム側が**ユーザの現在及び将来の心の状態と次の発話を推定**することが可能となった。また、ユーザの各状態ごとに発話の内容が異なるため、状態ごとの言語モデルを設けることにより、ユーザの発話に対する音声認識率を飛躍的に向上させることが可能となった。

2.2 キー概念に基づく情報検索

従来の情報検索方式は、分類表方式またはキーワード方式のいずれかであった。このうち、後者は自動化に適しているが、キーワードの表記だけに頼るため、検索の精度が低い欠点がある。本研究では、この欠点を解消するため、表面的にはキーワード方式ながら、キーワードの持つ概念(キー概念)のレベルにまでさかのぼって検索を行なう方式(キー概念方式)を考案し、それを具体化するため、次の3つの処理を導入した。なおキーワードとキー概念の関係を図2に示す。

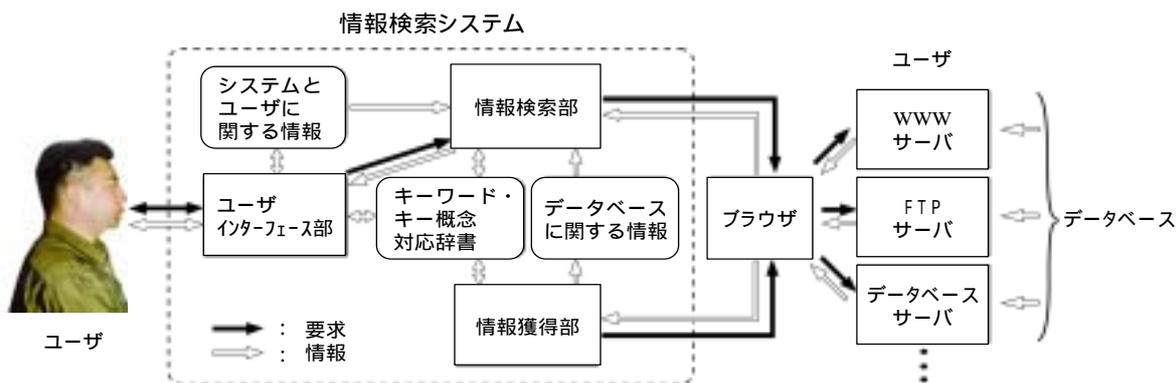


図1. 人間-機械音声対話に基づく学術情報検索システムの構成

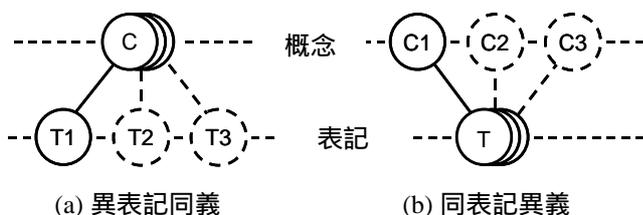


図2 キーワードとキー概念の関係

(a) キーワードの異表記同義への対応

異表記同義は検索もれ(データベース中の妥当な文書の総数に対するその割合を M とする)の原因となる。これを避けるため、まずすべてのキー概念に対して可能なキーワードを網羅した「**概念 表記辞書**」を作成し、これを用いて、ユーザーが提示するキーワードと同概念の語はすべてキーワードに加える方法を開発し、その有効性を実証した。

(b) キーワードの同表記異義への対応

同表記異義は誤検索(検索した文書の総数に対する割合を F とする)の原因となる。これを避けるため、まずすべてのキーワードに対して可能なキー概念を網羅した「**表記 概念辞書**」を作成し、キーワードを含む文書を共起する語に従って自動的にクラスタリングし、検索の場合には共起語を利用してキー概念以外のものを自動的に排除する方法を開発し、その有効性を実証した。

(c) 未知語の処理

ここで未知語とは、システムの辞書に登録されていない語をさす。学術の進歩に従って常に新語が作られ、また日本語では必要に応じて臨時に複合語を作ることが許容されるため、それらはすべて未知語となる。学术论文における未知語の調査と分類を行なった結果、約 80% が既知の形態素から構成される複合語であることを見出した。この種の未知語を対象として、その内部構造を統語及び意味の見地から自動的に分析し、**既知の構成要素の意味から未知の複合語の意味を自動的に推定**する方法を開発し、その有効性を実証した。

2.3 検索結果の妥当度の評価と検索の最適化

現在の情報検索システムの多くは、キーワード方式であり、 M も F も共に高い。これは検索結果の妥当度を自動的に評価する適切な方法を欠いているためである。本研究では、まず予備実験として、ユーザーのあげたすべてのキーワードから導かれるキー概念の論理和によって検索を行い、個々の検索結果が検索要求に対してどの程度妥当であるかを人間により 5 段階に評価させた。一方、検索式との一致度、キーワードの出現位置、出現頻度などから、**個々の検索結果の妥当度を定量的に推定する式**を導入し、妥当度の推定値と人間による評価値との相関係数が最大となるように推定式の係数を決定した。相関係数は 0.81 である。以下、この方式で推定した妥当度を単に妥当度と呼び、最大値で正規化した値を D と略記し、 D が閾値 θ を超えるもののみを最終的な検索結果とみなす。一般に M は θ の単調増大関数、 F は

θ の単調減少関数である。これらの関数は多数回の検索実験を行なうことにより、実測可能である。

検索結果の総合的な評価量としては検索者の立場により種々のものがありうる。一例として、**総合的な損失 L を M と F の荷重和として定義**する。(ϕ は荷重係数)

$$L(\phi, \theta) = \phi M(\theta) + (1 - \phi) F(\theta)$$

図3は $\phi = 0.5$ として、 \circ は本研究で採用した妥当度 D を用いた場合、 \square は情報検索で通常用いられる $tf \cdot idf$ 法により妥当度を推定した場合の L を示す。 θ のすべての値に対し、妥当度 D を用いた場合の方が $tf \cdot idf$ 法による推定値を用いた場合よりも L が小さく、特に前者が極小となる $\theta = 0.25$ の近傍ではその差が顕著である。この傾向は荷重係数 ϕ の広い範囲に対して認められ、本研究の妥当度推定法の有効性を実証している。

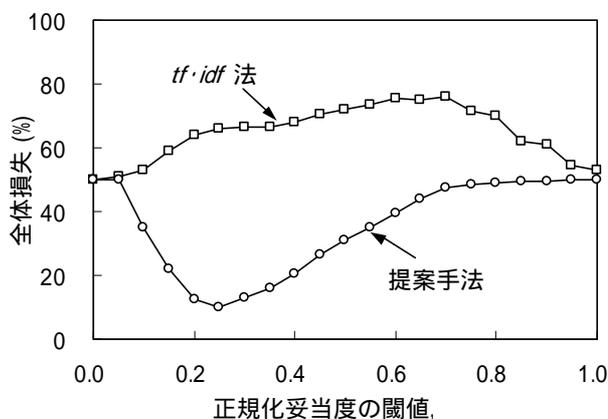


図3. 提案手法と $tf \cdot idf$ 法の総合的損失の比較

2.4 システムの構築と総合的評価

EDR 電子辞書をもとに概念→表記辞書及び表記→概念辞書を作成し、新たに開発した対話管理・概念検索・妥当度推定に関する基本的手法、および対話音声認識・合成手法を統合した音声対話による情報検索システムを構築し、学術情報センターによるテストコーパス NTCIR-1 の一部を用いて総合的評価を行なった結果、その有効性を実証した。

3. 結論

音声言語による高度の人間-機械対話システムの実現のために、対話による情報検索をとりあげ、ユーザーの心のモデルに基づく対話管理、キー概念検索、検索結果の妥当度の自動推定と総合評価等、多くの独創的かつ革新的な手法を導入することにより、**世界で最初の音声対話による学術情報検索システム**を実現し、これらの手法の有効性を実証して、学術上・社会生活上の進歩に貢献した。

主な発表論文

- (1) H. Fujisaki: "Toward an intelligent system for academic information retrieval through human-machine spoken dialogue," Proceedings of SNLP- O-COCOSDA2002 (2002) 1-11.