

# 情報検索を目的とした対話の分析

藤崎 博也 大野 澄雄 飯島 岐勇 阿部 賢司 黒川 一滋 津守 裕

東京理科大学

## 1. はじめに

情報検索においては、多くの場合、ユーザは検索すべき対象を当初から明確に意識しているわけではなく、また、十分な知識を持ち合わせているとは限らない。人間が介入する検索においては、検索の専門家(サーチャー)がユーザとの音声による対話を通じてユーザの検索要求を明確化させ、迅速かつ確かな検索を行なっている。したがって、機械による情報検索システムにおいても、ユーザとの音声対話によりその検索要求を明確にすることが望ましいと考えられる。このような見地から、筆者らは既に音声対話をインターフェースに用いた新しい情報検索システムの構想を示している [1]。

対話インターフェースは図 1 に示されるように、音声認識、音声生成のみならず、対話の進め方に関する知識を用いて対話を管理する機能が必要である。また、何を重視するか(短い対話で済ませたい、精密な検索を行いたい、など)によって対話の進め方を変化させることが望ましい。

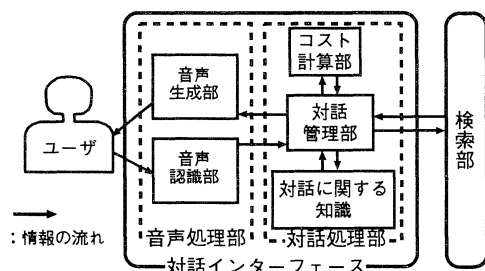


図 1. 対話インターフェースの概略図

本稿では、対話処理部の具体化を目的として、学術情報検索における対話の実例を収集、分析した結果を述べる。

## 2. 対話の収集

前節の目的のための基礎データとして、学術情報検索を目的とした人間対人間の音声対話 62 例・1379 発話を収集した。対話は表 1 に示した条件で行った。

表 1 各セットの条件

セット名	ユーザタイプ 指定	シナリオ 指定	実例数 (発話数)
セット 1	○	○	16(134)
セット 2	○	×	5(42)
セット 3	×	×	41(1203)

また、全ての対話は基本的にはシステムが主導権を持つものとし、ユーザが主導権を持つのは、ユーザが望んだ場合のみとした。なお、表 1 において、ユーザタイプ指定とはユーザの知識や要求をあらかじめ定め、ユーザ役に与えることを表し、シナリオ指定とは、システム役、ユーザ役に随時黒子が発話の指針を与えることを表す。

## 3. 発話の分類

収集した対話を分析するにあたり、本稿では、発話のタイプにのみ着目した分類とそれをさらに細分した分類の二つの段階に分けて分類した。本稿では前者を大別分類と呼び、後者を細分類と呼ぶ。表 2 にシステム役の発話の分類を示す。システム役の発話は、大別分類では 8 種に、細分類では 36 種に分類される。一般にシステム役はまず挨拶し、次にキーワード(以下、KW と略する)を要求し、数回の真偽質問によりユーザ役の検索意図を確かめ、それによる検索結果を伝え、最後に挨拶で締めくくるという順序で対話を行う。また、検索結果の内容についての解説、KW についての詳細な解説、検索方法の変更の提案、などは本稿では対応外とする。

表 2 システム役の発話の大別分類

発話	定義	割合	細分
挨拶	一般的な挨拶や決まり文句	6.6	2
KW の要求	ユーザから KW を得るための発話	10.5	3
真偽質問	はい、いいえで答えられる質問	48.8	14
選択質問	選択肢の中から答えを選択する質問	1.6	3
了解	ユーザの要求を受け入れた事を表す発話	5.7	1
事実の伝達	事実を伝える発話	28.4	7
警告	要求がシステムの仕様を超えていることを表す発話	0.1	1
過剰応答	コンピュータにより実現する上で、本稿では対応しない発話	3.3	5

一方、ユーザ役の発話は、大別分類では 9 種に、細分類では 19 種に分類される。また、表 3 は、ユーザ役の発話の約半数が応答であることを示す。これは

多くの対話においてシステムが主導権をとり、対話を進めているためである。なお、本稿で想定する対話システムの仕様を超える要求は通常の要求とは区別した。

表 3 ユーザ役の発話の大別分類

発話	定義	割合	細分
挨拶	一般的な挨拶や決まり文句	2.1	2
KW の提示	システムに KW を与えるための発話	19.4	4
応答 (肯定)	肯定的な応答	32.5	1
応答 (否定)	否定的な応答	14.9	1
応答 (不明)	迷っていたり、分からない場合の応答	1.1	1
要求	行為や文書、詳細などの要求とその指定のための発話	21.5	5
条件の変更	検索の条件を変更するための発話	3.6	2
検索の中止	検索を中止するための発話	0.8	1
過剰要求	システムの仕様を超える要求	4.2	2

#### 4. 深層状態の分類と分析

対話を管理するためには、現在の対話の状況を的確に把握することが重要である。しかし、発話のみに着目した分析では、その背後にあるシステムやユーザの状態には触れておらず、的確な対話の管理を行うことは難しい。このような観点から、我々は、ある発話を促す要因となる発話者の状態を深層状態と呼び、その分析を行う。システム、ユーザはそれぞれ独自の深層状態を持ち、それは自分に関する情報だけでなく、相手に関する情報や、検索結果など周囲の状況により定まる。深層状態は直接的に発話に表れることは無いが前後の対話の状況 (文脈) で判断できる。

また、深層状態に基づく状態遷移図は、発話を行うことで状態が遷移し、その発話は相手の発話や検索結果など、外部からの情報により誘発される。すなわち、システム役の状態遷移図とユーザ役の状態遷移図は相互に作用し合いながら対話を進めていく。したがって、この状態遷移図は対話の進め方に関する知識として利用できる。さらに、その遷移確率から経路のコストを導き、最適な経路を得るためのコスト計算を行うことができる。

ここでは、システム役の深層状態を分析し、その結果を表 4 に示す。システム役の深層状態は、一般に、まずユーザ役からの KW の獲得により遷移し (KW 受付 1 ~ 3)、それ以上 KW は得られないと判断し (KW 打ち止め)、検索を行う。次に検索結果とそれに対す

るユーザの反応により、状態はさらに遷移し、ユーザ役が満足するか検索する意欲を無くした場合に最終状態となる。

表 4 システム役の深層状態の分類

深層状態	定義	割合
初期状態	対話開始状態	4.7
KW 受付 1	KW を得ていない状態	4.3
KW 受付 2	KW を得ているが、さらに得られる可能性のある状態	12.5
KW 受付 3	手持ちの KW を補強するような KW を要求している状態	2.8
KW 打ち止め	それ以上 KW を得られそうにない状態	13.4
該当なし 1	検索該当数は 0、ユーザ状態は不明	0.1
該当なし 2	検索該当数は 0、ユーザは不満で検索をやり直したい	3.6
該当適正 0	検索該当数は適正で、ユーザ状態は不明	0.1
該当適正 1	検索該当数は適正で、ユーザは満足	22.2
該当適正 2	検索該当数は適正で、ユーザは不満で検索をやり直したい	2.4
該当適正 3	検索該当数は適正で、ユーザ状態は未確定で、詳細情報を希望	20.9
該当過多 0	検索該当数は過多で、ユーザ状態は不明	0.4
該当過多 1	検索該当数は過多で、ユーザは満足	0.8
該当過多 2	検索該当数は過多で、ユーザは不満で検索をやり直したい	1.4
該当過多 3	検索該当数は過多で、ユーザ状態は未確定で、詳細情報を希望	1.5
検索中止	ユーザが継続意欲を無くした状態	4.4
最終状態	対話終了状態	4.4

#### 5. まとめ

本稿では深層状態を考慮した対話管理を行うために、対話の実例を分析し、システム役の深層状態を分類した。今後は、深層状態に基づく状態遷移図を対話の管理に利用する方法を検討する。また、ユーザ役についても同様に分析する。

#### 参考文献

- [1] 藤崎博也, 亀田弘之, 大野澄雄, 阿部賢司, 伊東卓哉, 佐久間聖仁: “キー概念の抽出と未知語の処理に基づく情報検索方式の高度化,” 情報処理学会第 54 回全国大会講演論文集, vol. 3, pp. 23-24 (1997).