

## 文献データベースの知的インターフェースとその性能評価

# AN INTELLIGENT INTERFACE TO A DOCUMENT DATABASE AND ITS EVALUATION

木下 茂行                      加納 康男                      高橋 友一                      小林 幸雄†  
Shigeyuki KINOSHITA    Yasuo KANOU    Tomoichi TAKAHASHI    Yukio KOBAYASHI

ATR 通信システム研究所  
ATR Communication Systems Research Laboratories

あらまし あいまいな問合せ文を許す文献データベースの知的インターフェースの性能評価の手法を述べる。更に、あいまいな問合せ文からその検索意図を理解する方法を提案し、その方法に対する性能評価の手法を使った実験について述べる。あいまいな問合せ文に基づいて文献を検索するためには、あいまいさを解消して検索意図を明確にするフェーズと、それに基づき文献を検索する二つのフェーズが必要である。前フェーズにはキーワード検索効率、後フェーズには文献検索効率を用い、中間段階と最終段階を評価することにより、うまく検証することができることを示す。問合せ文の理解の方法としては、対話により理解する手法と、対話をしないで理解する手法(類推による理解)を提案し、これを実験により評価する。また、二つの理解手法の問題点もこの評価実験から抽出するとともに、それを解決することにより両手法を統合できることを示す。

Abstract A quantitative evaluation method for intelligent interface to a document database which allows ambiguous queries is discussed. In addition, ambiguous queries understanding methods are proposed, and the understanding methods are validated by the evaluation method. To retrieve documents according to a ambiguous query, two phases are required. Phase 1 clarifies its requirement, and phase 2 achieves the retrieval. Keyword efficiency is used to evaluate the intermediate stage (phase 1), and retrieval efficiency is used to the final stages (phase 2) of document retrieval. Two methods of understanding, understanding through dialogue and understanding by analogy, are proposed, and they are validated by the evaluation method. Their problems are also clarified and approaches to solve them are derived by the evaluation method. The approaches suggest the integration of the two understanding methods.

### 1 まえがき

情報検索システムに対する従来の自然言語インターフェース[1]では、Berkinらが指摘するように、利用者の検索意図が自然言語で完全に表現できると仮定している[3]。文献検索の利用者は、何らかの問題解決を行うのに不足している知識を文献に求めているため、自然言語でそれを完全に表現するのは困難である。また Taylor[2]が指摘するように、検索要求には漠然とした段階から検索文で表現できる段階まであり、利用者が言葉で表現する検索要求は、ふつうかなり漠然とした段階であると考えられる。

従って、文献検索システムの自然言語インターフェースには、このような漠然とした、あるいはあいまい性を持った問合せ文を理解できることが要求される。

ところで、従来の文献検索システムにおいては、その性能評価は文献検索効率(適合率、再現率)で行なわれていた。利用者の検索意図が自然言語で完全に表現できる場合には、この指標での評価で十分であるといえるが、問合せ文があい

まい性持つような場合には、これだけでは不十分である。すなわち、あいまいな問合せ文から、利用者の検索意図を十分に汲み取ったかどうかは、文献検索効率では評価できない。この、問合せ文から利用者の検索意図を抽出している程度を表す指標があれば、あいまいな問合せ文を許すような自然言語インターフェースを十分に性能評価できると考えられる。

我々は、問合せ文からの利用者の検索意図の抽出を検索意図に相当するキーワード集合を抽出することであると捉え、これの性能をキーワード検索効率と呼ぶ指標で評価する。このキーワード検索効率と従来の文献検索効率の二つの指標で、あいまいな問合せ文を許すような自然言語インターフェースの性能評価を試みる。この評価は、あいまいな問合せ文を理解し、それに基づく検索を行なう手法の提案を行ない、その手法の有効性を検証するという形で行なう。あいまいな問合せ文の理解手法の提案と、キーワード検索効率による評価は以前に報告している[9][10]ので、今回は文献検索効率による評価を主体に述べる。

†現在、NTT ヒューマンインターフェース研究所勤務

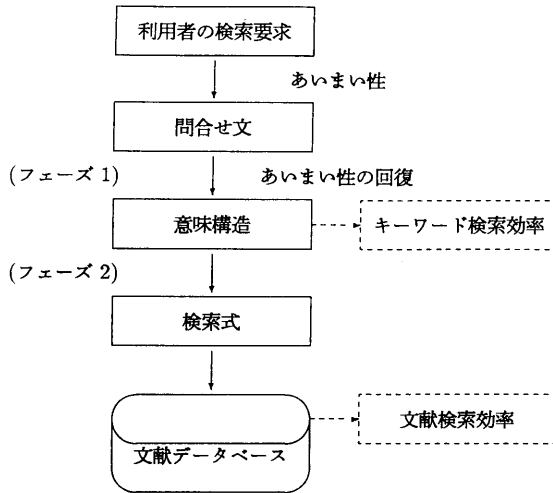


図 1: 文献検索システムのモデル

## 2 文献検索のモデルと性能評価について

### 2.1 文献検索のモデル

問合せ文に含まれる漠然性あるいはあいまい性のうち、ここでは次のものを対象とする。

- (1) 利用者独自の言葉、あるいは特有の言葉の使い方  
問合せ文には、システムの知っている言葉（キーワード）以外の独自の言葉を含むことがある。またキーワードであっても、利用者特有の意味で使用することがある。
- (2) 過剰の限定  
問合せ文で使った言葉の意味を広く解釈するのが望ましい場合がある。例えば、“Dempster-Shafer 理論”と言ってもこれ以外の不確実性を表現するものを含むことがある。
- (3) 限定の不足  
検索意図を問合せ文で十分に限定しない場合がある。

ここでは、文献データベースはキーワードにより索引付けされた、従来型のものと仮定している。このようなあいまい性を含んだ利用者の問合せ文を理解し、それに基づいて文献データベースを検索するために、次のような2つのフェーズで問合せ文を処理するものとして、図1に示すようにモデル化した。

#### (フェーズ1) 意味構造の合成

問合せ文中のあいまい性を回復して、利用者の検索意図を明確化する。その結果は、問合せ文の意味構造として表現する。この意味構造は、キーワードとそれらの間の関係で表現する。

#### (フェーズ2) 文献データベースの検索式への変換

問合せ文の意味構造を、キーワードを文献データベースの検索式へ変換する。

### 2.2 文献検索の性能評価

あいまいな問合せ文に基づく検索を評価するためには、役割の異なった前述の二つのフェーズを独立に評価するとともに、それらの相関関係も評価することが望ましい。ここでは、図1に示すように、フェーズ1の評価指標としてキーワード検索効率、フェーズ2の評価指標として文献検索効率を用いてこれらの定量評価を試みる。

多くの文献検索において、文献検索効率による評価は行なわれてきたが、ここで述べるような問合せ文の理解といったような中間段階での評価、あるいはそれと文献の検索効率との相関の評価はなされていない。

#### (1) キーワード検索効率

フェーズ1では問合せ文の意味構造を作成するが、これを定量評価するために意味構造に含まれるキーワードのみを抽出し、次のように定義するキーワード検索効率を用いる。キーワード適合率 ( $Kp_q$ ) とキーワード再現率 ( $Kr_q$ ) からなり、それぞれ次の式で定義する。

$$Kp_q = \frac{|Eq'|}{|Eq|}$$

$$Kr_q = \frac{|Eq'|}{|Pq'|}$$

ここで、 $|Pq'|$ : 全キーワードのうち利用者の検索意図に適合したキーワードの数

$|Eq|$ : システムが抽出したキーワードの数

$|Eq'|$ : システムが抽出したキーワードのうち利用者の検索意図に適合したキーワードの数

#### (2) 文献検索効率

文献検索効率は文献適合率 ( $Rp_q$ ) と文献再現率 ( $Rr_q$ ) からなり、それぞれ上記のキーワード検索効率の定義の“キーワード”を“文献”と読み換えたものとなる。文献検索効率は順位付け出力を採用するシステムにおいては一般に再現率-適合率曲線で表現される。[6][7]

### 3 提案する理解手法について

#### 3.1 理解手法の概要

これらの問題を解決するために、我々は次のようなアプローチを採用した。

##### (a) ユーザモデルの導入

システムの対象分野の知識を、利用者毎に用意すること（ユーザモデル）により、利用者独自の言葉を理解できるようにした。さらに、ユーザモデルを使用して、問合せ文で使った言葉を手掛かりに隠れた検索意図を推論できるようにした。

##### (b) 問合せ理解の2つの方法

1つの方法、利用者とは対話を行いながら利用者独自の言葉（システムにとっての未知語）やそれ以外のあいまい性を解決する方法（対話による理解と呼ぶ）である[11]。また、

表 1: 問合せ文中の名詞間の関係の分類

種類	関係名	出現回数	使用される言葉 (助詞)
単独		19	
併記	併記	40	とおよび や または ある いは
階層	例示	5	をふくむ などの として の とくに たとえば
	対象	106	の での における への にたいする に対する
限定	手段	24	による をもちいた をつ かった を使用した
	目的	5	のための にもちいる
	因果関係	9	による
	場所	7	での における の
その他		7	

この対話を通してユーザモデルを構築する。この方法では、ユーザモデルが構築された後では、これを使用して推論を行うことにより、少ない対話で適切に検索意図を抽出できることが期待できる。

他の方法は、利用者との対話を行うことなく、利用者独自の言葉を含んだ問合せ文であっても、問合せ文中に含まれる名詞間の関係と、ユーザモデルの構造を照合することによって検索意図を抽出する手法（類推による理解と呼ぶ）である。

提案する問合せ理解の各処理の内容の詳細については以前に報告している[8][9][10][11]ので、ここでは概要を述べるとどめる。

### 3.2 問合せ文とユーザモデルの表現について

利用者の問合せ文と検索対象の分野の知識は、同じものを表現している。すなわち、問合せ文は利用者が何を検索したいかを表現したものであり、検索対象の分野の知識は何が検索されるかを表現したものである。ここでは、問合せ文に含まれる“概念”とそれらの間の関係に基づいて、問合せ文とユーザモデルを表現する。

表1に問合せ文に含まれる概念間の関係を示す。これは、約250の問合せ文の解析結果に基づいている。また、表1中の出現回数は、約250のうちの136の問合せ文でのそれぞれの関係の出現回数である。出現回数の合計が136にならないのは、一つの問合せ文中に複数の関係が存在するからである。

#### (a) 問合せ文の表現について

利用者の問合せ文は「(名詞句表現)が欲しい」という形式で表現されるものとし、この名詞句表現から表1に従って概念を表現している名詞とその間の関係を抽出し表現する。

例: 「意味ネットワークによる対象世界モデルの表現」

(手段 (対象 表現 対象世界モデル)  
意味ネットワーク)

#### (b) ユーザモデルの表現について

ユーザモデルは、あるユーザ独自の分類体系や言葉に基づく個人用のシソーラスである。名詞で表された概念ノードと概念ノード間の関係を表すリンクを持つ意味ネットワークで表現される。概念ノードには、検索キーワードに対応するキーワードノードと、利用者独自の言葉（以下、独自語と呼ぶ）に対応する独自語ノードの2種類がある。リンクには、表1の分類に基づく次の4種類の関係の種類がある。

- ・階層関係
- ・関連関係
- ・同義関係
- ・限定関係 (対象、手段、目的、因果関係)

ユーザモデルは初期状態として典型的な対象世界モデルとして定義されており、利用者との対話によってこれを変形するとともに独自語を取り込む。概念ノード間のリンクには、2つの概念間での利用者の連想の強さを示す3段階のレベル（意図のレベル：初期状態、意図する、意図しない）を持つ。これらのレベルは、隠れた検索意図を推論するために用いる。

### 3.3 対話による理解について

利用者の問合せ文を理解する第一の方法は、問合せ文に含まれるあいまい性を利用者に尋ねることにより明確化しながら、問合せ文をユーザモデルにマッチングする方法である。これは、次のような処理により行なう。

#### (a) あいまいな上位概念の意味の明確化

問合せ文に含まれる言葉に対応する概念ノードに対し、初期状態レベルのリンクで連結されている下位または関連、限定関係の概念ノードがあれば、これらを利用者に問合せリンクの意図のレベルを“意図する”もしくは“意図しない”に修正する。

#### (b) 未知語に対する処理

問合せ文に含まれた言葉が、ユーザモデルにない言葉（未知語）の時は、利用者にその言葉の説明を求める。システムは、対話に基づいてこれを新たなリンクと新たな独自語ノードとしてユーザモデルに組み入れる。

#### (c) 関係構造の付加と削除

利用者がユーザモデルにある概念ノードを用いてそれらの間の新たな関係構造を表明したときには、その関係に基づいたリンクを付加する。一方、利用者がシステムから提示された関係構造を否定した時には、ユーザモデルの関係構造を修正する。

### 3.4 類推による理解について

利用者の問合せ文を理解する第二の方法は、利用者との対話を行なうことなく、問合せ文中の名詞間の関係情報を手がかりとして問合せ文をユーザモデルにマッチングする方法である。これは、次のような処理により行なう。

(a) 未知語を含まない場合の処理

問合せ文の名詞を含む関係構造全体をユーザモデルに照合し、その関係構造に最も近いユーザモデルの部分を意味構造とする。この時、名詞あるいは名詞間の関係が完全に一致しなくても、ある程度の許容範囲内で一致するものを見つ出す。このため、照合の結果は一意ではなく、複数の候補が存在することがある。

(b) 未知語の処理

ある関係構造の片方の名詞が未知語で片方が既知（照合に成功したもの）である場合には、既知の名詞と名詞間の関係を使用して、次の2つの規則に従って処理する。

規則1 未知語を、既知の名詞に対応するユーザモデル上の概念ノードに対してその関係の種類に対応したリンクで繋がれた概念ノードと考える。

規則2 未知語を、既知の名詞と同じ概念ノード、あるいはその上位の概念ノードと考える。

3.5 隠れた意図の推論について

上記の二つの手法は、いずれも直接表明された言葉の理解である。この両手法とも、問合せ文で表明されない隠れた検索意図を抽出する必要がある。そこで、この隠れた検索意図を抽出するとともに、これに相当するキーワードを選択する処理を行なう。この処理は、問合せ文に含まれる名詞に直接に関連するユーザモデルの概念ノードから、“意図する”とラベル付けられたリンクを辿って意図しているキーワードを抽出するものである。

4 文献データベースの検索方法について

文献検索の方法については多くのものが提案されているが、ここでは、比較的単純な次の二つの手法について評価を行なった。単純な検索方法を採用した理由は、評価の主たる目的が問合せ文の理解の評価であり、検索方法についてはこの目的を満足するものであれば良いことによる。

4.1 キーワード照合法

一つ目の手法は、キーワード照合法と呼ぶ最も単純な方法である。フェーズ1で作成した意味構造に含まれるキーワードを抽出し、これを文献に付与されているキーワードと照合し、一致したキーワード数で文献を評点付けする。この評点にしたがって、文献を順位付け出力するものである。

この方式は、ベクトル方式[6][7]の文献検索におけるすべての重みの値を一定にしたものである。また、論理式型の検索式としても記述することができる。

4.2 条件弛緩法

二つ目の手法は、条件弛緩法と呼ぶ論理式型の検索法である。フェーズ1で作成した意味構造が持つキーワードとその間の意味的な関連を、AND/ORの関係に変換することにより、検索式（最適検索式）を作成する。意味構造は、ある概念（キーワード）とそれを修飾する概念（キーワード）を表現していると考えられる。この修飾関係をAND（同じ修

表 2: 実験対象の問合せ文

被験者	問合せ文
1	Q <sub>1</sub> 不完全性またはあいまいさ
	Q <sub>2</sub> 仮説検定処理におけるあいまい処理または確率的扱い
	Q <sub>3</sub> エキスパートシステムにおける信頼性
2	Q <sub>1</sub> ファジィ推論における重み関数
	Q <sub>2</sub> 設計を目的とするエキスパートシステム
	Q <sub>3</sub> バックトラック
3	Q <sub>1</sub> 自律走行ロボットにおける障害物の認識
	Q <sub>2</sub> 認知地図の作成
	Q <sub>3</sub> 自然言語理解におけるメンタルモデルの利用
4	Q <sub>1</sub> ユーザインタフェースにおける言語の役割
	Q <sub>2</sub> 言葉のあいまい性の処理のための推論方式
	Q <sub>3</sub> つかい易さの向上を目的としたユーザインタフェース

飾関係のものが複数あればそれをOR)で表現することにより、検索式を作成する。これは、意味構造中のすべての条件を記述した、最も検索条件のきついものである。

この意味構造から、他の概念（キーワード）を修飾するもののうち最も修飾関係が弱いと思われるものを順次削除して行くことにより、複数の検索式を作成する。この複数の検索式を条件のきついものの順に評点付けし、文献を順位付け出力するものである。

5 評価実験について

5.1 実験方法

5.1.1 実験システム

今回の実験システムでは、検索対象の文献を人工知能分野のものとした。ユーザモデルの初期状態の典型的な対象世界モデルとして、約500のノード（キーワードノード）からなるものとした。これは、[5]に基づいて表現されている。また、文献データベースは約800の文献からなり、キーワードノードの名前に基づくキーワードが付与されている。

5.1.2 実験手順

実験は4人の被験者（当研究所の研究者）に対して行った。1人の被験者に対し3つの問合せ文を作成して貰った。被験者の作成した問合せ文を表2に示す。実験手順は以下の通りである。

(1) 対話による理解の処理

3つの問合せを続けてシステムに提示し、これを3回繰り返した。則ち、3つの問合せをQ<sub>1</sub>、Q<sub>2</sub>、Q<sub>3</sub>とすると、Q<sub>1</sub> ⇒ Q<sub>2</sub> ⇒ Q<sub>3</sub> ⇒ Q<sub>1</sub> ⇒ Q<sub>2</sub> ⇒ Q<sub>3</sub> ⇒ Q<sub>1</sub> ⇒ Q<sub>2</sub> ⇒ Q<sub>3</sub>の順で処理した。処理の終了後に被験者に約500のキーワード集合を表示して、自分の検索意図に相当するキーワード集合を選択してもらった。

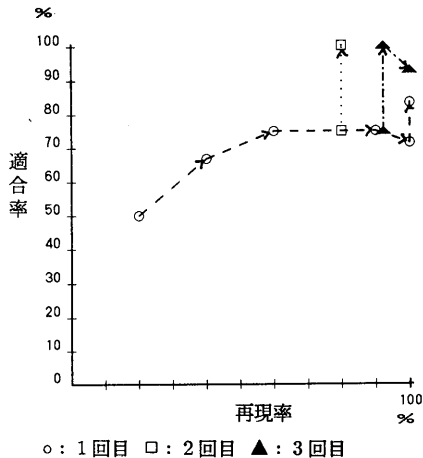


図 2: 対話による理解におけるキーワード検索効率の推移

(2) 類推による理解

各対話による理解処理と同じユーザモデルと同じ問合せ文により処理した。また、被験者が作成した問合せ文の検索意図が同じで言葉による表現を少し変えたものを作成し、ユーザモデルを使用して処理を実施した。

(3) 目的の文献の選択

約 800 の文献の中から自分の検索意図に沿った文献を選択してもらった。

(4) キーワード検索効率、文献検索効率の計算

各処理の終了後に、被験者に選択してもらったキーワード集合および文献集合を元にキーワード検索効率および文献検索効率を求めた。

(1)、(2)、および(4)のキーワード検索効率の計算については以前に実験を行っており、今回はそれに(3)および(4)の文献検索効率の計算を追加した。また、(1)、(2)については[9][10]の報告の後、再実験を行なったため、これらとはキーワード検索効率の値は異なっている。

## 5.2 実験結果

### 5.2.1 対話による理解の評価

被験者 4 の Q<sub>1</sub> に対する対話による理解処理の各対話の時点で抽出できたキーワードを用いて計算したキーワード適合率 / 再現率の対話の進行に伴う推移を図 2 に示す。また、各被験者毎の対話による理解処理における対話の回数を表 3 に示す。図 2 および表 3 から、次のような事が言える。

- キーワード適合率 / 再現率共に増加する方向に対話が進行する。
- 同じ問合せ文では、2 回目以降の対話の回数が減少する。
- 1 回目の対話では、約 10 回の対話により、高いキーワード適合率 / 再現率が得られる。2 回目以降の対話では、約 5 回の対話により、高いキーワード適合率 / 再現率が得られる。

表 3: 対話による理解処理における対話の回数

	被験者 1			被験者 2		
	Q <sub>1</sub>	Q <sub>2</sub>	Q <sub>3</sub>	Q <sub>1</sub>	Q <sub>2</sub>	Q <sub>3</sub>
1 回目	5	8	11	11	15	8
2 回目	4	7	5	4	4	5
3 回目	3	7	4	2	3	5

	被験者 3			被験者 4		
	Q <sub>1</sub>	Q <sub>2</sub>	Q <sub>3</sub>	Q <sub>1</sub>	Q <sub>2</sub>	Q <sub>3</sub>
1 回目	8	7	4	13	6	6
2 回目	3	3	2	3	5	3
3 回目	3	3	2	5	2	3

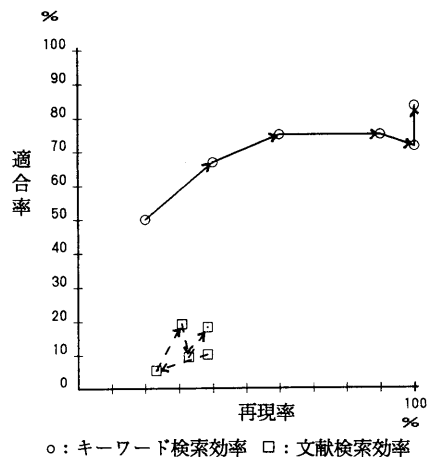


図 3: 対話による理解における文献検索効率の推移

- 2 回目以降の対話の開始時点でのキーワード適合率 / 再現率は、ある程度高い値である。

同じ問合せ文の 1 回目の対話について、キーワード照合法により検索して得られた文献適合率 / 再現率の対話の進行に伴う推移を図 3 に示す。また、被験者 3 の Q<sub>2</sub> の文献検索効率の推移を図 8 に示す。文献適合率 / 再現率は、適合率 - 再現率曲線の各点の平均値で示している。この図から、次のようなことがわかる。

- 文献適合率 / 再現率についても共に増加する方向に対話が進行する。但し、キーワード適合率 / 再現率の変化は敏感には文献適合率 / 再現率に影響しない。また、文献適合率 / 再現率はある程度のところで飽和する。

- 最終的に得られる文献適合率 / 再現率は、キーワード適合率 / 再現率に比べ良くない。

### 5.2.2 類推による理解の評価

図 2 と同じ問合せ文を類推による理解で抽出したキーワードを用いて計算したキーワード適合率 / 再現率を図 4 に示す。このグラフは各処理毎に、照合処理終了時とそれに引き続く推論処理終了時の検索効率をプロットしている。また、この問合せ文の、検索意図は同じで言葉の表現を変えたものを同

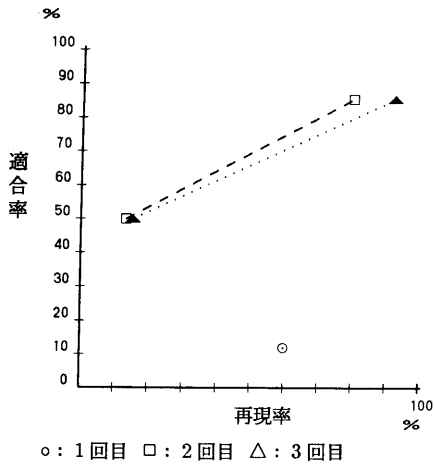


図 4: 類推による理解におけるキーワード検索効率の推移

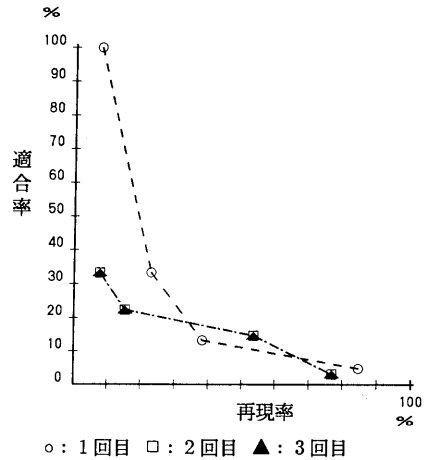


図 6: 類推による理解における文献検索効率の推移

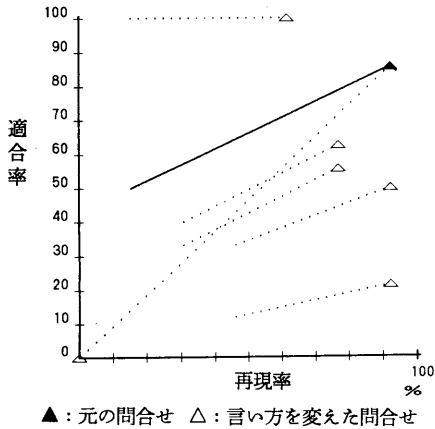


図 5: 同じ意図で異なった問合せ文での類推による理解のキーワード検索効率

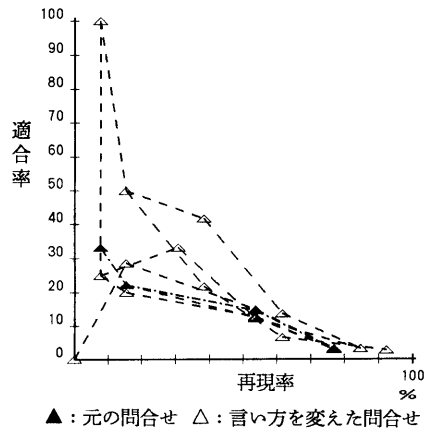


図 7: 同じ意図で異なった問合せ文での類推による理解の文献検索効率

じユーザモデルに対して類推による理解処理を行ったものの結果を図5に示す。

これらから、次のような事が言える。

- ユーザモデルが構築された後(2回目以降の処理)では、高いキーワード適合率/再現率が得られる。但し、1回目の処理を除いて対話による理解でのキーワード検索効率を越えることはない。
- 以前の問合せ文とあまり表現が変わらない(未知語を1つ含む程度)で検索意図が以前のもと同じである場合には、元の問合せ文に比べて再現率が高め、適合率が低めの検索効率となる。

同じ問合せ文の、キーワード照合法により得られた文献適合率/再現率(適合率-再現率曲線により表現)を図6および図7示す。この図から、次のようなことがわかる。

- 得られる文献検索効率は、キーワード検索効率に比べ良くない。

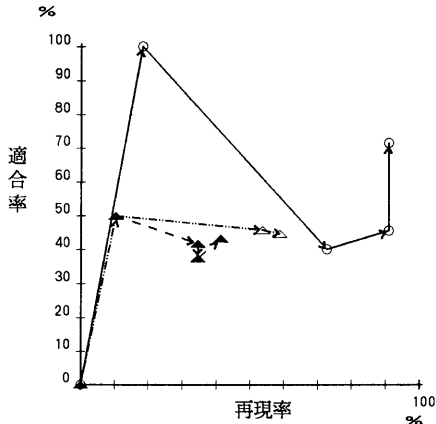
- 図7においては、文献検索効率のばらつきはキーワード検索効率ほど大きくない。更に、図6では1回目の文献検索効率は、2回目以降のものに近い値となっている。これは、文献検索効率は主としてキーワード再現率に依存することを示している。また、図7での最も良い値を示すものは、キーワード検索効率ではほぼ中間の値のものである。

### 5.2.3 文献検索法の比較

文献検索法については、キーワード照合法と条件弛緩法の二つについて実験したが、上記の実験結果は、すべてキーワード照合法でのものである。ここでは、キーワード照合法と条件弛緩法による文献検索効率の比較を行なう。

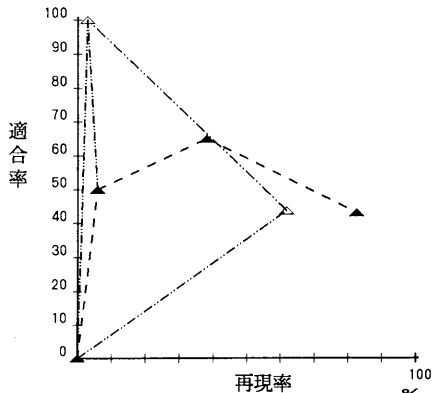
図8および図9に、同じ問合せ文に対するキーワード照合法と条件弛緩法による文献検索効率の比較を示す。これらから、次のようなことがわかった。

- 条件弛緩法での文献検索効率は、キーワード照合法による場合と比べて、検索効率が良いケースと悪いケースに二分される。



○: キーワード検索効率 ▲: キーワード照合法 △: 条件弛緩法

図 8: 被験者 3Q<sub>2</sub>の対話による理解における検索法の比較



▲: キーワード照合法 △: 条件弛緩法

図 9: 被験者 3Q<sub>2</sub>の類推による理解における検索法の比較

- 条件弛緩法での文献検索効率、検索式毎の検索効率のばらつきが大きく、きれいな適合率-再現率曲線を描かない。

## 6 実験結果の考察

今回の実験の結果を前節で述べた。この結果は、キーワード検索効率という評価指標だけで見れば、ほぼ良好な値を示しているといえる。しかしながら、文献検索効率、あるいはキーワード検索効率と文献検索効率の相関で見れば多くの問題点を持っている。ここでは、その考察を行うと共に、解決方向について述べる。

### (1) 文献検索効率の低調さ

今回の実験では、文献検索効率はキーワード検索効率と比べて全般に低い値となっている。条件弛緩法において、前述のような検索効率が良いケースと悪いケースに二分されたり、検索式毎の検索効率のばらつきが大きくなっている。この理由として、次のものが考えられる。

表 4: 被験者が選択した文献中のキーワード

出現頻度		抽出	キーワード
選択文献	全文献		
6	15	yes	自然言語インターフェース
5	48	yes	マンマシンインターフェース
4	12		自然言語処理
3	18	yes	情報検索
3	19		自然言語理解
3	44	yes	データベース
3	129		知識ベース
2	30		人工知能
2	51		モデル
2	248		エキスパートシステム

- キーワードのもれ

検索に必要と思われるキーワードのすべてが必ずしも理解結果の意味構造に含まれていない。表 4 は、被験者が選択した文献中のキーワードのうち、二つ以上の文献に現れるものが、意味構造中に抽出されたかどうかを示す。全文献中の出現頻度に比べて選択文献での出現頻度が高いキーワードは、被験者の検索意図に沿った文献を特定するための重要なキーワードと考えることができる。表 4 の例では、自然言語処理や自然言語理解は、重要なキーワードと考えられるが、抽出されなかった。

- キーワードの重要度

キーワードには、意味的に重要なものと検索上重要なものの二つがある。意味的に重要度とは、検索要求での概念の重要度とである。検索上の重要度とは、検索意図に沿った文献を特定する程度であり、表 4 中の出現頻度に比べて選択文献での出現頻度が高いキーワードは重要度が高い。キーワード照合法では、いずれの重要度も無視しているため、文献検索効率は低い値となっていると考えられる。条件弛緩法では、意味的に重要度のみを考慮している。しかし、今回の実験結果は意味的に重要なキーワードと検索上重要なキーワードは必ずしも一致していないことを示している。このため、両者が一致した場合には検索効率が良くなり、逆の場合には悪くなっている。また、条件を緩めていく際に検索上重要なキーワードを外すと急に検索効率が悪くなり、検索効率のばらつきが大きくなっていると考えられる。

これを解決するためには、次のような方法が考えられる。

- (a) 検索結果に基づくユーザモデルの構築

被験者のあまり意識していないキーワードは、対話により抽出できず、従ってユーザモデルにも反映されない。選択された文献に付与されているキーワードで重要と思われるものをユーザモデルに反映することにより解決できる。

- (b) 文献検索時におけるキーワードの追加

抽出されたキーワードと階層関係で結合されているもの(表 4 の例では、自然言語処理)は、文献検索時に自動的に追加することが可能である。

### (c) キーワードの検索上の重要度の利用

意味構造を持つ意味的な重要度以外に、検索上の重要度を考慮した検索式を作成する。検索上の重要度は、キーワードの出現頻度や、人間のサーチャが持つヒューリスティックな知識を使用することにより知ることができる。

### (2) 文献検索効率の飽和

今回の実験では、キーワード検索効率対話の進行に応じて上昇するのに対して、文献検索効率はあるところから上昇しなくなるという現象が生じている。この原因としては、前述の文献検索効率の低さとともに、対話処理での不必要な細かな対話が挙げられる。これは、問合せ理解処理ですべての検索意図を抽出して検索しようとしていることに起因する。

この問題を解決するためには、検索結果から問合せ理解(対話処理)へのフィードバック[4]が考えられる。すなわち、検索結果を利用者に提示し、(1)での解決法に加えて利用者の対話により必要なキーワードを抽出して文献検索効率高めることで解消できると考えられる。

### (3) フィードバック手法による理解手法の統合

前述のフィードバック手法を用いれば、現在提案している二つの問合せ理解手法を統合することが可能である。さらに、(2)で述べた問題および(1)の検索結果に基づくユーザモデルの構築についても解決することができる。これは、次のような手順で行なう。

- 未知語のみからなるような問合せ文を除いて、類推による理解でまず検索を行う。
- 検索結果を利用者に提示し、欲しい文献を選択してもらう。
- 選択結果からのキーワード抽出と、それに付随した対話からユーザモデルの構築と理解の補足、修正を行ない、再び検索を行なう。

## 7 むすび

あいまいな問合せ文を許す文献データベースの知的インターフェースの性能評価の手法を述べた。文献検索の過程を、問合せ文を理解し、意味構造に変換する過程(フェーズ1)と、意味構造に基づき文献データベースを検索する過程(フェーズ2)に分けて捉えた。フェーズ1に対してはキーワード検索効率という指標で問合せ文の理解の性能を評価し、フェーズ2に対しては文献検索効率という指標で問合せ文に基づく文献検索の性能を評価するものであった。

更に、あいまいな問合せ文からその検索意図を理解するためのユーザモデルを導入した二つの方法、対話による理解と、類推による理解を提案し、その方法に対する性能評価の手法を使った実験について述べた。対話による理解では、キーワード適合率/再現率共に増加するように対話が進行し、1回目の対話で約10回、2回目以降の対話で約5回の対話で十分高いキーワード検索効率を得られることがわかった。また、類推による理解においても、多少の未知語が混在するという状況においても、ほぼ十分なキーワード検索効率を得られることが確認できた。

一方、文献検索効率については、主にキーワード再現率に依存するという傾向は見られたものの、キーワード検索効

率に比べて低い値しか得られなかった。これについては、文献検索効率とキーワード検索効率の相関などを調べることににより、前章で述べたような問題点とそれに対する解決の手がかりを得ることができた。また、提案した二つの問合せ文の理解の手法が統合してこのような問題点を解決しながらうまく働くことを示した。

このように、キーワード検索効率と文献検索効率という二つの指標を使って、あいまいな問合せ文を理解し、それに基づいて検索するメカニズムをうまく検証することができることを示した。人間の使う言葉には、多くの冗長性、漠然性、あいまい性が含まれている。文献データベースに限らず、さまざまな自然言語インターフェースの使い易さを向上するためには、このような人間の使う言葉の特性を許すものが望ましい。冗長性、漠然性、あいまい性等を許した自然言語インターフェースの検討において、本稿で示した二つの評価指標による評価は広く適用が可能である。

最後に、本研究の機会を与えて下さった ATR 通信システム研究所葉原耕平会長、山下紘一社長、本実験に協力して下さいました知能処理研究室の皆様へ感謝します。

## 参考文献

- [1] G. Jakobson, et al. "An Intelligent Database Assistant", IEEE EXPERT, SUMMER, 1986.
- [2] R. S. Taylor, "The Process of Asking Questions", American Documentation, Oct., 1962.
- [3] N. J. Berkin, et al. "ASK for Information Retrieval: Part 1. Background and Theory", The Journal of Documentation, Vol. 38, No. 2, June, 1982.
- [4] J. Yen, et al. "Specification By Reformulation: A Paradigm for Building Integrated User Support Environment", in Proceedings of AAAI-88, 1988.
- [5] David L. Waltz, "Scientific Datalink's Artificial Intelligence Classification Scheme", THE AI MAGAZINE, Spring, 1985.
- [6] サルトン編著, 神保監訳, "SMART 情報検索システム", (株) 企画センター, 1974.
- [7] 伊藤哲郎著, "情報検索", (株) 昭晃堂, 1986.
- [8] S. Kinoshita, Y. Kanou, et. al., "Deep Understanding of Japanese Queries in an Information Retrieval System", Proceedings of the CPCOL-88, 1988.
- [9] 木下, 加納他, "知的情報検索システムにおける問合せ理解とその評価", 情報処理学会知識工学と人工知能研究会報告, 62-2, 1989.
- [10] 木下, 加納他, "知的情報検索システムにおける問合せ理解の評価について", 情報処理学会第38回全国大会, 1989.
- [11] 加納, 木下他, "文献検索システムにおける検索要求意図推論方式の一検討", 情報処理学会第37回全国大会, 1988.