

2H-3

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

加納 康男      木下 茂行      高橋 友一      小林 幸雄

(株) A T R 通信システム研究所

1. まえがき

文献検索システムにおける利用者の検索要求意図を理解する方法について述べる。利用者にとって、文献検索要求文に、検索のために必要な概念を完全に表現することは難しい。また利用者は、一般に、文献データベースに付けられたキーワードではなく自分自身の言葉で検索要求を行う。このため、利用者の用いた用語だけでは、意図に適合したキーワードを選択することは困難であるという問題点がある。

従って、利用者との対話によって曖昧な検索要求から要求概念が明確化でき、利用者自身の言葉を用いた検索要求文から必要なキーワードを抽出して、意図する文献が検索できるシステムが望ましい。

このために、我々は、利用者毎の概念構造(ユーザモデル)を持ち、利用者との対話によってユーザモデルを修正し、ユーザモデルを利用して利用者が表明しなかったキーワードを推論するシステム(図1)を検討している。

本論文では、「検索意図の理解」を、ユーザモデルを用いて利用者の意図する文献検索のために必要十分なキーワード集合を抽出することと考え、ユーザモデルの表現方法、対話を通したユーザモデル構築のための規則、利用者が表明しなかった概念を推論してキーワードを抽出する手法を提案する。

2. ユーザモデルの表現方法

ユーザモデルは、名詞(句)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の独自語に対応する独自語ノードの2種類から構成される。リンクは、①階層関係、②関連関係、③同義関係、④限定関係(目的限定、手段限定、対象限定)の4種の関係の種類があり、リンクの種類は検索要求文や利用者との対話から決定される。限定関係リンクと関連関係リンクは、概念間で意図する強さを示す3段階のレベル(初期状態、意図する、意図しない)を持つ。これらのリンクは、システムが①利用者の意図をユーザモデルに反映する、②利用者への問い合わせ時にユーザモデルの内容を表現する、③検索意図を推論するために用いる。

ユーザモデルは、初期状態においては、キーワードノードとキーワードノード間を初期状態のレベルにある階層関係リンクと関連関係リンクで結んだシソーラスとして定義されている。

3. 利用者との対話によるユーザモデル修正

利用者との対話によってその検索意図をユーザモデルに反映させるための手順について述べる。

(1) 構文解析

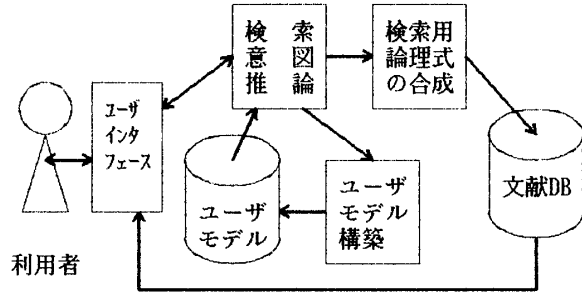


図1 システム構成の概略

検索要求文において表明された名詞と名詞に結び付けられた助詞の意味に基づく係り受け構造を、格表現に変換する。これによって関係リンクの種類が決定される。

(2) 曖昧な上位概念の意味の明確化

検索要求文に含まれ、またはリンクを用いて推論された概念ノードに対して、初期状態レベルの関係リンクで連結されている下位または関連する概念ノードがあれば、これらを利用者に問い合わせる関係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。この対話例を図2-1、図2-2に示す。

(3) 未知語に対する処理

検索要求文に含まれた言葉が、ユーザモデルに登録されていない言葉(未知語)の時は、利用者による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関係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。このとき同時に、新しい独自語ノードに対して、その定義に用いた既存の概念ノード(定義指示ノード)の情報を与える。この対話例を図2-3に示す。

(4) 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。この対話例を図2-2、図2-4に示す。

以上のユーザモデルの修正操作をまとめて表1に示す。

表1 ユーザモデル修正操作

操作名	操作
OP1	既存ノードに新しい関係リンクと新しい独自語ノードを付加する。新しい独自語ノードには、定義指示ノードの情報を与える。
OP2	既存ノード間に新しい関係リンクを付加する。
OP3	2つの既存ノード間の関係リンクを削除する。独自語が関係する場合には定義指示情報を修正する。
OP4	階層関係、関連関係リンクの意図のレベルを修正する。

4. ユーザモデル上での意図の推論

検索要求文に含まれる言葉と一致するユーザモデルの概念ノードから、リンクを辿って意図しているキーワードを抽出する。その時適切なキーワード集合を抽出するために、①独自語からキーワードへのマッピングについては、方向性として独自語の定義指示ノードの情報を、独自語の意味の定義として限定関係以外の関係リンクを利用する。

②キーワードについては、関係リンクごとに設定された推論の方向性および推論の強さを表す値 $W_i$ を利用する。

各リンク、ノードの種類毎の推論方向、推論後に活性化されるノード、推論の強さ $W_i$ を表2にまとめて示す。

これらを用いた推論規則として次の規則を用いる。

**ルール1** 一つの推論を行うパスにおいて、各リンクの推論の強さの値 $W_i$ を所定の閾値で掛け合わせた値が閾値を越えないようにして推論範囲を制約しながら関連するキーワードを抽出する。

**ルール2** 関連するキーワードが検索を行うに十分な数となったとき推論を終了する。さもなければ問い合わせによって処理を繰り返す。

なお、推論におけるループ性を排除するために、各概念ノードに推論の履歴情報を持つ。最終的には、このようにして抽出されたキーワードをシステムの推論結果として利用者に表示して確認を求める。図2において網かけされたノードは、意図しているものと推論されたノードを示す。

初期状態

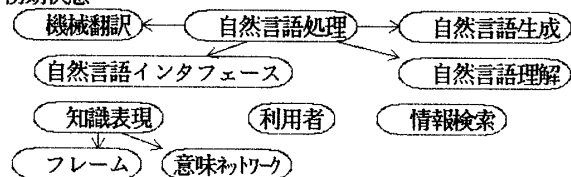


図 2-1

1回目の対話

U; 情報検索における自然言語処理について調べてほしい。  
 S; 自然言語処理には、機械翻訳、自然言語インタフェース、自然言語理解、自然言語生成という概念がありますがこれを含みますか？  
 U; 自然言語インタフェースと自然言語理解だけを含みます。

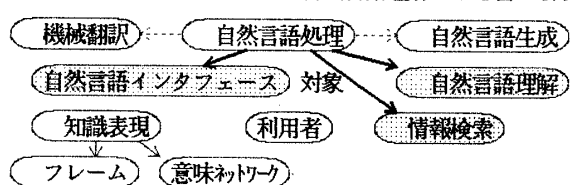


図 2-2

5. 今後の課題

本方式で抽出されたキーワードが、キーワードのブラウジングによって得られた利用者の真意の中にあるキーワードと一致する度合いを、キーワードの適合率、キーワードの再現率という観点から評価することを考えている。また、検索された文献に対する利用者の評価からユーザモデルを修正する方式の検討、抽出されたキーワードと関係リンクの種類に基づいて論理式を合成する方法の検討を行う。そして、本手法の有用性を示すためのプロトタイプシステムを作成して評価を進める。

【参考文献】

(1) 加納, 木下他: "情報検索システムにおけるユーザモデル構築の一検討", 電子情報通信学会, 第4回交換・情報ネットワークワークショップ資料, PP. 185-186, 1988

2回目の対話

U; 情報検索において意味ネットワークを用いたユーザモデルについて述べた論文を調べてほしい。  
 S; ユーザモデルとはどのようなものですか？  
 U; 利用者の知識表現のことです。

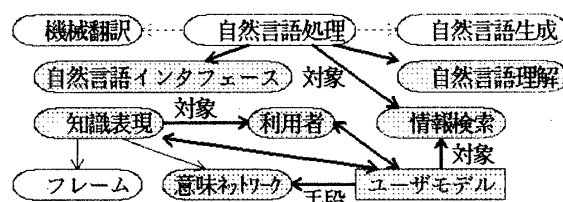


図 2-3

3回目の対話

U; ユーザモデルについて書いた論文ある？  
 S; 情報検索における自然言語インタフェースと自然言語理解で、意味ネットワークを用いたユーザモデルについてですか？  
 U; 意味ネットワークは含みません。

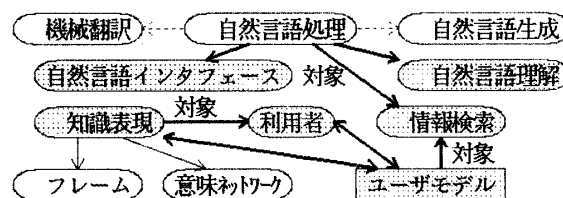


図 2-4

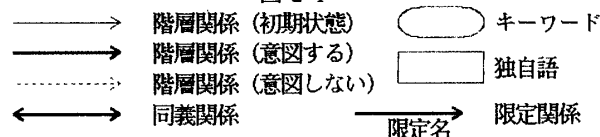


図 2 ユーザモデルの表現と対話例

表 2 関係リンクの種類ごとの推論における性質

関係リンクの種類		推論の方向性	推論後に活性化されるノード	推論の強さ $W_i$
階層関係 意図している	独自語含む	独自語⇒キーワード(1)	定義指定された概念ノード	必ず連想する 1
	キーワードのみ	上位語⇒下位語	下位語のみ	$W_1$
関連関係 意図している	独自語含む	独自語⇒キーワード(1)	定義指定された概念ノード	必ず連想する 1
	キーワードのみ	両方向	両方のキーワード	$W_2$
同義関係	独自語含む	独自語⇒キーワード(1)	定義指定された概念ノード	必ず連想する 1
	キーワードのみ	両方向	両方のキーワード	必ず連想する 1
限定関係	独自語含む	独自語⇒キーワード(1)	定義指定された概念ノード	$W_3$
	キーワードのみ	両方向	両方のキーワード	$W_4$
階層, 関連関係 意図していない	独自語含む	---	---	---
	キーワードのみ	推論しない	表明された方のノードのみ	推論しない 0

(1) 独自語⇒独自語が定義指示された概念ノード