

## 6H-2 日本語質問応答システムにおけるメタ質問処理

加藤 恒昭 中川 優

NTT情報通信処理研究所

### 1. はじめに

関係データベースの検索を自然言語によって行うことの利点のひとつとして、データの格納構造についての知識が不要である事が挙げられる。しかし、対象とするDBにどのような情報が格納されているかという情報なしでの利用は、ユーザにとって使い勝手のよいものではないし、そのような状況で入力される質問について適切に回答する事も困難である。ユーザが必要に応じて「何が質問できるのか」という質問(メタ質問と呼ぶ)をする事ができ、それによって何が格納されているかの知識を得られることが必要である。

このためのメタ質問処理機能としては、1. 対話の中でメタ質問と一般の質問を区別する事[1]、2. メタ質問について適切に解答する事が必要である。本稿では2.の問題について述べる。また、メタ質問は、「これ(検索結果)は何ですか」のように、現在の話題を確認するためのメタ質問と、「～の何が検索できますか」など分野知識を知るためのメタ質問とに分類されるが、本稿では後者について述べる。これは分野知識がユーザにとって不明確で、対話を続けられなくなった場合の割込み[1]としてなされたり、解答がユーザの予想に反したものである場合になされるものである。

本稿の構成は以下の通り。まず、検索対象となる分野知識を説明する。次に、統語的な特徴からメタ質問を分類する。その後、メタ質問に対する適切な解答について論じ、最後にメタ質問処理機能の実現について述べる。

### 2. 分野知識

我々が試作している自然言語インタフェースシステムでは、構文意味解析部や問題解決部での利用法に対応して、分野知識に以下の階層構造を持たせている[2]。

第一層(最下層) RDBと基本的な概念(意味素と呼ぶ)の対応づけ

第二層 意味素と概念(クラス)との対応づけ

第三層 概念の構造の記述(日本語との対応づけ)

図1aの関係表を例に考えると、ここから社員、部などのクラス(定義域)や社員と役職、社員と年齢などの関係が意味素として抽出され(図1b)、両者が第一層で関係づけられる。これらの意味素から、社員を役職で制約し

社員表				
社員	役職	年齢	所属	……

(a) 検索対象となるRDB

社員, 役職名, 歳, 部  
 役職(社員, 役職名)  
 年齢(社員, 歳)  
 所属(社員, 部)  
 (b) 抽出される意味素

管理職, 組合員(社員のサブクラス)  
 営業, 事務, 開発(部のサブクラス)  
 若い(歳, Deg), …  
 (c) 意味素から構成される概念  
 (意味素と重複する部分は省略)

図1 分野知識によって得られる概念

管理職の概念(クラス)が社員のサブクラスとして定義され、また、年齢がある値以下であることが「若い」と定義される(図1c)。これらの定義が第二層に記述される。第三層では、管理職と社員の間の変化(IS-A)階層や「若い」が年齢を主格にとることなどが、格関係やIS-A関係などをアークとし、普通名詞や用言に対応するクラスをノードとする意味ネット形式で表現される。

この階層化は、分野知識をその種類に応じて分類したと見ることができる。つまり、第三層は概念の構造を与えている。第二層は概念の定義を与えている。第一層はDBの物理構造に関連するもので、ユーザに見せる必要は無い。従って、この階層化はメタ質問処理にも適したものである。つまり、メタ質問の解答として何が必要かわかれば、分野知識のどの層を検索すればよいか明らかになる。

### 3. メタ質問の分類

分野知識に関するメタ質問は、構文意味解析部によって、表1に示すメタ質問タイプ分類される。この分類は、解答内容にほぼ対応する。また、前節で述べた様に、検索内容も解答内容に応じて以下のように決定できる。

関連する属性	→	第三層
分類	→	第三層 汎化階層
定義	→	第二層
具象値	→	分野辞書
概念間の関係	→	第二、第三層

ABタイプ間の相違は、解答形式に反映される。Aは、ある項目について何を条件に検索できるかを尋ねており、Bはある項目のどんな属性が検索できるかを尋ねている。従って、それに応じた解答をしなければならない。

表1 メタ質問の分類

メタ質問の文型(例)	解答内容
A ~を検索するための条件は	ある概念に関連する属性
B ~の何が検索できるか	ある概念に関連する属性
C ~の分類は	ある概念の分類
D ~の定義は/~とは	ある概念の定義
E ~にはどんなものがあるか	ある概念の分類/具象値
F ~と~の関係は	概念間の関係

#### 4. メタ質問への解答

以上のように解答内容、解答形式はメタ質問タイプを基にほぼ決定できるが、以下の2つの問題がある。

- ・ 適当な解答内容が存在しない場合
  - ・ メタ質問タイプから解答内容が一意に決定できない場合
- 前者は、意味素に対してタイプDのメタ質問をするような場合である。一般の質問の場合、解答が存在しないことが情報となるが、メタ質問はユーザへの支援であるから、解答なしという状況を防ぎ、何らかの情報を与える必要がある。

後者の例として、タイプCのメタ質問は具象値とサブクラスのどちらを解答するかが曖昧である。これは、ユーザの語感にも依存するが、クラスの特徴から決まる場合も多い。この例でも、サブクラスを持つクラスの場合、それを解答した方が適切であると思われる。従って、クラスの特徴によって処理を切り分けることを考える。

これらを考慮して、解答内容の決定は以下の様に行う。まず、メタ質問タイプと、クラスがどの品詞に相当するかというクラスの特徴に応じて処理が切り分けられる。各処理では、より適当な解答が優先されるように検索すべき項目が並べられており、解答内容が存在するまで、検索を行う。これは解答が存在しない場合の対処と同時に、サブクラスを持つかなど分野知識の検索後明らかとなるクラスの特徴に応じた解答内容の分類となる。このような2段階の分類を行うことで、適切な解答内容が得られる。例として、メタ質問がタイプCで、クラスが名詞相当である場合をとると、図2aに示す①②③の処理を①より順に実行し、解答が得られた時点で終了する。②③は概念の定義を答えていないので、不適切であるとも

[名詞]の定義を知りたい

- ① 第二階層を検索して、その定義があれば、解答とする
  - ② ISAアークをたどり、下位クラスがあれば、解答とする
  - ③ それを属性とするクラスを解答とする
- (a) 解答内容の選択

- ①管理職 役職が課長か役職が部長の社員です。
  - ②社員 管理職、組合員などです。
  - ③年齢 社員の年齢です。
- (b) 解答内容の例

図2 解答項目決定処理の例

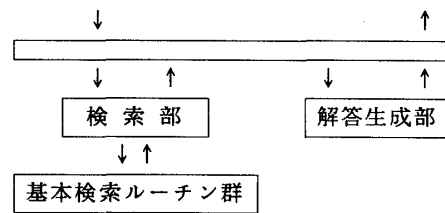


図3 メタ質問処理部の構成

いえるが、ユーザに何らかの情報を与えるための代替的な解答である。それぞれの場合の例を図2bに示す。

#### 5. メタ質問処理の実現

メタ質問処理部の構成を図3に示す。入力の意味解析部によって判定されたメタ質問タイプとクラス名である。検索部は、前節で述べた方法に従って、適切な基本検索ルーチン呼び、分野知識から解答内容を収集する。基本検索ルーチンには、分野知識の第二層を検索して定義を求めるとIS-Aリンクをたどってサブクラスを求めなどの意味ネットワーク検索用プリミティブとがある。

解答内容は、解答生成部に渡され、ここで解答文が生成される。日本語生成は基本的にはテンプレート方式を用いており、メタ質問タイプに応じて、適切なテンプレートが選択されると共に、解答内容を名詞句とするか、文とするが決定される。また、入力クラスの代名詞化なども行われる。図4に幾つかの解答例を示す。下線部が解答テンプレート、破線部が代名詞化された部分である。

社員はどんな条件で検索できますか  
 => その役職、その年齢、彼が属す部で検索できます。  
 管理職、組合員があります。

社員の何が検索できますか  
 => その役職、その年齢、彼が属す部の情報があります。  
 属すとは => 社員が部に属すことです。  
 若いとは => 社員の年齢が25歳以下のことです。  
 社員の分類は => 管理職、組合員に分類されています。

図4 メタ質問とその解答の例

#### 6. おわりに

メタ質問処理の実現について述べた。本方式では、解析に利用する以外の情報をなるべく利用しないように設計を行ったが、アクセス効率、情報の精度の観点から不十分な点がある。後者の例として、サブクラスの和が親クラスと一致するかの情報が挙げられる。この情報は、解析では不要であるが、「人間」のサブクラスが「社員」だけの場合、人間の分類に対する解答は「社員に分類されている」となる。「社員とそれ以外に…」としたい所である。今後は、メタ質問に十分な解答ができるような分野知識内容の検討と、その形式について検討してゆく。

#### 参考文献

- [1]加藤他：「質問応答における意図の把握と話題の管理」情処 自然言語処理研究会 58-6 1986
- [2]加藤他：「日本語質問応答システムにおける知識ソースの利用方式」情処 第33回全国大会 5M-3 1986