

## データベース照会システム「ヤチマタ」と 名詞句データ模型†

藤崎 哲之助†† 間下 浩之†† 諸橋 正幸††  
渋谷 政昭†† 鷹尾 洋一††

ヤチマタは、日常使用に近い日本語によるデータベース照会システムである。照会文を片仮名により端末より入力すれば、回答結果が表の形で端末上に表示される。照会文としては十分に多様な表現が許されるときもに不必要な詳細は書かずにすむため、使用者は自然に習熟することができる。仮名書きにおける分かち書きの自由さ、同音異義語の多さ、は致命的な障害ではない。曖昧さが生じたときには対話により解消される。

ヤチマタは汎用に設計されており、できるだけ多くの部分を諸応用から独立させる一方、各応用に依存する部分は容易に付加できる。また、1つのデータベース管理システムの下で、他の形式的な照会・更新言語と共存することができる。

ヤチマタ設計の基礎は名詞句データ模型の概念である。これは関係形式データ模型の変形であって、データベースと自然言語との間を橋渡しするものである。すべてのデータは名詞句に対応し、名詞句の修飾は関係代数演算と類似の演算に対応する。

汎用性を示すため、部品の在庫、航空便の時刻、地域統計、の3つの実験的な照会システムが稼動しており、明確に規定された範囲の中では日本語で正確に質問できるし、日本語はむしろ照会言語向きの面もあることを実証している。

### 1. ま え が き

データベース利用の拡大にともない、調査・開発・意志決定のためそれを対話的に利用する要求が強くなる。これに応じて多くの照会用言語が開発されてきたが、形式的言語では、たまさかの利用者 (the casual user) が容易に習熟できるほど簡便で、しかも必要な機能をもつものは見当たらない。より容易な方式として、たとえば Query-By-Example<sup>1)</sup> ではメニュー方式を採用している。

一方、照会言語として自然言語を用いる試みが、他方面での利用と並んで研究されてきた。たとえば、LUNAR<sup>2)</sup>、REQUEST<sup>3)</sup> が代表的であり、最近では TORUS<sup>4)</sup> LADDER/INLAND<sup>5)</sup>、ROBOT<sup>6)</sup> が開発されている。「ヤチマタ」もこれらと同方向を目指す研究用実験システムであるが、(1)日本語用、(2)応用分野が一般的、(3)他の形式言語と共存可能、という特長をもっている。

本論文では、ヤチマタの設計目標と実現のための方針を述べ(2章)、設計のための基礎概念として新しく

導入した名詞句データ模型について論じ(3章)、ヤチマタの構成を概観し実働化の諸経験を記す(4章)。なお、ヤチマタの名称は本居春庭の文法書「詞の八衢(ことばのやちまた)」1808 によっている<sup>7)</sup>。

### 2. ヤチマタの設計目標と方針

ヤチマタの目標の第1は、日常使用に近い片仮名文による照会である。入力文を仮名とするのは、将来においても末端利用者が漢字入力装置に熟達するとは期待できないからである。仮名文では分かち書きの正書法が確立しておらず、多くの同音異義語があることが障害となるが、後述のようにこれらは克服することができた。検索結果はデータの表の形で表示する。

目標の第2は、システムのできるだけ多くの部分を諸応用と独立となるように切り離し、各応用に依存する部分は容易に付加できるようにすることである。LUNAR、REQUESTなどのように特定の目的のために個々に照会システムを構築したのでは費用がかさむし、小規模の外部仕様変更が、大規模なシステム改造をもたらす心配がある。

第3の目標は、物理データベースを管理する1つのシステムの下で、日本語照会言語だけでなく、他の複数の照会言語も共存させることである。情報システムの中では、システム・プログラムのため、日常業務

† Yachimata, a Data Base Query System, and the Noun-phrase Data Model, by TETSUNOSUKE FUJISAKI, HIROYUKI MASHITA, MASAYUKI MOROHASHI, MASAOKI SIBUYA, and YOICHI TAKAO (Tokyo Scientific Center, IBM Japan, Ltd.)  
†† 日本アイ・ビー・エム(株)東京サイエンティフィック・センター

処理のために、形式的な照会言語（データ準言語）が必要である。データの更新も多くの場合は日常業務であり、一定の規則に従うので日本語向きではない。この機能とともに、末端利用者ごとに異なるサブスキーマを使用することも許したい。

さて、第2の汎用性の目標を満たすためには、諸情報を使用者の目からは一貫した立場で眺めねばならないが、データ模型という概念は、まさしくそのような枠組を設けるものである。諸模型の中では、E. F. Coddの関係形式<sup>8),9)</sup>が単純・強力である。しかもこれにたいしては各種の形式言語が考えられており、これに基づくシステムでは複数の照会言語を積極的に考慮している<sup>10)</sup>、第3のデータベース共用の目標も満たしやすい。第1目標の自然言語使用も、関係形式研究の初期より提案されていて、Codd自身が実験的対話システムを開発している<sup>11)</sup>。

自然言語と関係形式とはそのままでは互いになじまない。たとえば照会言語 SEQUEL<sup>10)</sup>の文をそのまま日本語に替えるとどうなるであろうか。まず、データベース、あるいはそのサブスキーマを構成する関係形式の名称と、その構造についての知識が前提となる。照会文の中では、関係形式の名称および当該列の名称をすべて指定し、複数の関係形式の結合(join)演算を指定しなければならない。これでは自然言語を用いることが末端利用者にとって何の助けにもならない。これを解決するために、われわれは関係形式を少し変形した「名詞句データ模型」の概念を導入する。

### 3. 名詞句データ模型

#### 3.1 名詞表と動詞表

「名詞表」は一般に図1の構造をもったデータの組であり、普通名詞に対応している。関係形式に類似しているが少し異なり、まず列の並びに関して対称でなく、表の名前である普通名詞が同時に代表的な第1列を表わしている。この第1列を、「対象属性」、第2列以降を「限定属性」と呼ぶことにする。対象属性が個体・事物を指すならば、限定属性はその諸特徴を示し、前者が動作・処理を指すならば、後者は当事者・物を示す。限定属性をもたず、対象属性だけから成る1項関係式の名詞表も考えることにする。各属性(列)に現われ得る要素の集合「定義域」は、関係形式の場合と同じ意味をもち、大小・等不等の比較演算可能な規準で定めることによりデータの型を細分する。

名詞表はもうひとつ関係形式と異なっていて、後者

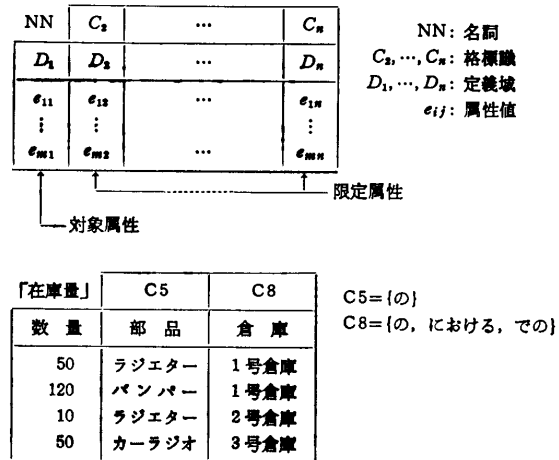


図1 名詞表の構造と例

Fig. 1 A noun table and its example.

の役割語 (role word) の機能を増してある。図1の例では、限定属性が「(倉庫)における(部品)の在庫量」という連関を表わし、第1行は「1号倉庫におけるラジエターの在庫量は50である」というデータを示している。つまり、名詞表の限定属性は〈名詞句(限定属性)〉〈後置詞〉〈名詞表=対象属性〉という形の修飾句を許すことに対応する。「後置詞」とはここでの造語で上例の「における, の」を表わし、格助詞および類似の機能の辞句の総称である。後置詞の適当な集りを「格標識」と呼び、名詞表の限定属性の見出し行では、図1のように、定義域とともに格標識を指定する。これにより修飾句に現われる名詞句と後置詞の組合せを規定する。格標識の役割は動詞表において、より明確となる。

格標識は上述のように後置詞の集合であり、逆に1つの後置詞は複数の格標識に属してよい。1つの格標識は同じ名詞表・動詞表に複数回現われてよく、複数の表に現われてもよい。定義域についても同様である。ただし、(定義域, 格標識)のひとつの対(つ)い)が同じ表に2回以上現われてはならない。この制約は、1つの名詞に、同じ意味の複数の修飾句が係ることはない、という前提に基づいている。

「動詞表」は、用言(動詞, 形容詞, 形容動詞)に対応し、図2のような構造のデータの組である。これは用言を名前とする関係形式にほぼ等しく、列の並べ換えに関して対称である。用言は基本的に要素間の連関を表わし、したがって動詞表は限定属性だけから成り、C. J. Fillmoreの格文法<sup>12)</sup>における文の基本構造、〈格〉…〈格〉〈動詞〉にも対応している。ただし1つの限定属性が〈名詞句〉〈後置詞〉〜という修飾句を規定す

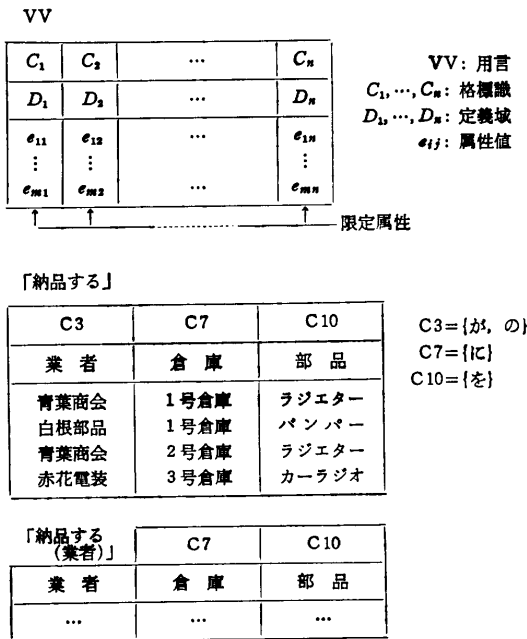


図2 動詞表の構造と例, 例から導かれる名詞表  
 Fig. 2 A verb table, its example and a derived noun table.

る点は名詞表の場合と同じで、〈格〉は表層的な意味しかもっていない。

用言を導入すると照会文の表現が豊かとなるが、名詞句主体のデータ模型では使い方を限定する。つまり〈名詞句=限定属性〉〈後置詞〉〈連体形用言〉〜という形の連体修飾句の中でだけ使える。そのため実際の処理過程も名詞表の処理に帰着し、動詞表は補助的な役割を果たすものである。

図2の例は「納品する」の動詞表である。その第1行は「青葉商会が1号倉庫にラジエターを納品する」という事実を表わしている。この表の修飾句としての役割は「(倉庫)に(部品)を納品する業者」、「(業者)が(倉庫)に納品する部品」、「(業者)が(部品)を納品する倉庫」の3つの名詞表と同等である。「納品する業者」に対応する名詞表を図2に示す。その第1行の意味は「1号倉庫にラジエターを納入する業者は青葉商会である」となる。

### 3.2 名詞表とデータベース, 照会文

われわれのデータ模型では、データである固有名はもちろん、集合名の普通名詞、用言、これらを受ける詞辞の意味はすべて業務ごとに異なり、それぞれの意味はデータベースに基づいて正確に定義されなければならないと考える。逆に定義さえすれば語彙を豊富にし、同じ意味内容を様々な表現できるし、同じ詞辞で

異なる意味を表わすこともできる。そのための手段が名詞表・動詞表であったが、これらは互いに重複し、関連し、データベースの経時的変化と共に内容が変わるものであり、したがって仮想的に存在する。実在するものは表の見出し行と、表とデータベースとの対応関係である。この対応関係は、関係形式データベースならば、その形式言語によりサブスキーマ<sup>10)</sup>(implied relation<sup>10)</sup>とも呼ぶ)として表わされる。適用分野に依存しない詞辞は、次節以降に示すように数多くはない。

名詞表は、基本的な普通名詞の意味を定義するばかりでなく、照会文の構文規則、つまり自立語のつながり方の意味の規定に役立つ。名詞句主体のデータ模型では、照会文は本質的には〈名詞句〉〈は?〉の形をしている。諸種の特性をもつ事物などを指定する名詞句は、階層的な修飾句をもっているが、名詞句を修飾して新しい名詞句を作る過程は、1つあるいは2つの名詞表から新しい名詞表を導くことに相当する。関係形式に関する検索演算の過程・結果が常に関係形式であるように、名詞表に関する検索過程・結果も常に名詞表となるように構文と意味を次節で定義する。したがって今後名詞表の名前は普通名詞とは限らず名詞句一般となる。これまで述べてきた「普通名詞」にしる文法通りのものに限らず、名詞句で支障はない。なおデータである固有名、数も対象属性だけの1行の名詞表とみなすことにする。

データベースからある知識を得るには、ふつう一連の関連した照会を行うため、前の照会結果を後から指すことが不可欠である。そのための簡便な方法として〈連体詞(その, この, ……)〉〈名詞句〉により、〈名詞句〉と同一定義域をもつ最新の結果を指す、などの技法を用いる。

### 3.3 データ検索演算

名詞表についての基本諸演算は、制約演算、集合演算、組込関数の3種に大別できる。これらについて順次述べる前に、一般的な前提を注意しておく。

諸演算は主として名詞表の対象属性にたいする作用であるが、その際に限定諸属性を考慮しなければならぬ。2つの名詞の対象属性間の演算を行うとき、(定義域、格標識)の対が等しい限定属性が両名詞表に共通にあれば、これらの属性値が一致するもの間だけで演算を行う。つまり共通限定属性値が等しいという条件の上で、対象属性に関する条件・演算に従う2名詞表の結合演算が行われ、結果の表では共通限定属性

が1つにまとめられる。このような操作を「限定属性の統合」と呼び、統合可能な両表の属性を「統合可能属性」と呼ぶ。1つの名詞表の中には(定義域, 格標識)の対が一通りしかないから、統合は一意的に行われ、統合によって結果の表における一意性が保たれる。

「制約演算」は名詞表からある条件を満たす行を選出する演算で、関係代数の結合演算で実行できる。名詞表の非対称性により、限定属性を条件付ける「正順制約」と、対象属性を条件付ける「逆順制約」とがある。

正順制約は〈名詞句1(名詞表1の対象属性=名詞表2の限定属性)〉〈後置詞〉〈名詞句2(名詞表2の対象属性)〉という構文に対応している。この名詞表2の限定属性の定義域は〈名詞句1〉つまり名詞表1の対象属性のものと一致し、格標識は〈後置詞〉を含むものとする。図3は図1の例における正順制約の例で、「ラジエターの在庫量」および、これにさらに正順制約を適用した「1号倉庫における(ラジエターの在庫量)」の名詞表である。

正順制約における2つの名詞表が、上記構文に現われる属性以外の統合可能属性をもつときには、両属性値間の等号条件が結合演算の条件に付加される。図4は図1の例の他に名詞表「特売用部品」があるとき、正順制約「特売部品の在庫量」の結果である。

逆順制約は〈名詞句1(名詞表1の対象属性)〉〈が〉〈名詞句2(名詞表2の対象属性)〉〈である〉〜という構文の修飾句に対応している。両対象属性の定義域は等しいものとする。両表が統合可能な限定属性をもつときの処置は前述の注意の通り。結合演算の結果から対象属性を除くと、動詞表の形の表となるが、この修飾句を受ける名詞句と同じ定義域をもつ限定属性が対象属性に変換されて結果の名詞表となる。このとき同じ定義域をもつ限定属性が複数個あれば、曖昧となることは避けられない。この曖昧さは動詞表を名詞表に

「ラジエターの在庫量」		
	C5	C8
数量	部品	倉庫
50	ラジエター	1号倉庫
10	ラジエター	2号倉庫

「1号倉庫におけるラジエターの在庫量」		
	C5	C8
数量	部品	倉庫
50	ラジエター	1号倉庫

図3 正順制約の例

Fig. 3 Examples of ordinary restriction.

「特売用部品」	
部品	倉庫
ラジエター	1号倉庫
カーラジオ	3号倉庫

「特売用部品の在庫量」		
	C5	C8
数量	部品	倉庫
50	ラジエター	1号倉庫
50	カーラジオ	3号倉庫

図4 統合可能属性のある正順制約の例

Fig. 4 Examples of ordinary restriction with compatible attributes.

変換するときにも生じる。

図1の例で、「在庫量が50である(部品)」という逆順制約を行った結果が図5である。同図には新しい名詞表「適正在庫」が与えられたとき不等号制約「在庫量が適正在庫以下である(部品)」を行った結果も示す。両名詞表が同じ形であるため、すべての限定属性について統合が行われる。

「集合演算」和・差・積のうち、和・積は並列句に差は否定にはほぼ対応する。和は各詞句の並列〈名詞句1〉〈接続詞(と, および, か, または, …)〉〈名詞句2〉に対応する。この演算は2つの名詞表の対象属性の和集合を作るが、両名詞表に統合可能な限定属性がある場合には、対象属性とこれらとの $n$ 項関係( $n$ 個組)の集合についての和を作る。結果の名詞表の限定属性は、元の2表に共通のものだけである。

積は、2つの名詞表の対象属性の積集合を作る演算で、2種の構文、〈名詞句1〉〈であり〉〈名詞句2〉

「在庫量が50である(部品)」	
部品	倉庫
ラジエター	1号倉庫
カーラジオ	3号倉庫

「適正在庫」		
	C5	C8
数量	部品	倉庫
20	ラジエター	1号倉庫
100	カーラジオ	1号倉庫
30	ラジエター	2号倉庫

「在庫量が適正在庫以下である(部品)」	
部品	倉庫
ラジエター	2号倉庫

図5 逆順制約の例

Fig. 5 Examples of reverse restriction.

〈であり(る)〉～, および 〈名詞句 1〉〈である〉〈名詞句 2〉に対応する. いずれの構文でも2つの名詞表を対象属性および統合可能限定属性の等号条件により結合する. 第1の対称構文ではこれがそのまま結果となるが, 第2の非対称構文では〈名詞句 2〉の名詞表と同じ構造をもつようにさらに射影演算を行う.

英語の関係代名詞に相当する〈名詞句 1〉〈のうちで〉〈名詞句 2〉〈のもの〉などの類の構文は, 前後を入れ替えて非対称な積として扱う.

名詞句の並列は和の, 動詞句の並列ならば積の演算を行えばよいが, 連用修飾句内に名詞並列句が現われる場合には特別の考慮が必要である. 接続詞「か, または」などに属するものは文脈に関係なく和に対応する. 「と, および」などは連用修飾句内のときにだけ積とし, 他の場合は和として扱う. 次の例を見よ.

「ラジエーターとバンパーの在庫量」:

(ラジエーターの在庫量) U (バンパーの在庫量)

「ラジエーターとバンパーの在庫量が10以下の倉庫」:

(ラジエーターの在庫量が10以下の倉庫) ∩

(バンパーの在庫量が10以下の倉庫)

名詞句の否定, 〈名詞句〉〈でない (以外の)〉～, などの構文の処理は基本的には集合差の演算である. 否定演算を  $\neg$ , 2項検索演算を  $\odot$ , 名詞表を  $N1, N2$ , 関係形式としての非対称差を  $-$  とすると,

$$N1 \odot (\neg N2) \Rightarrow N1 - (N1 \odot N2)$$

が一般的処理であるが, 文脈により少し扱いを変える必要がある. たとえば

逆順制約: 「 $N1$  が  $N2$  でない」  $\Rightarrow$

$$N1 * - (N1 \text{ が } N2 \text{ である})$$

ただし  $N1*$  は  $N1$  の対象属性を落とした動詞表類似の表である. 上記の演算で限定属性の統合が行われる場合にも, 否定の結果が肯定の場合の相補となるように扱う.

組込関数としては, 単項算術演算, 2項算術演算, 統計関数の3種類が標準的であろう. 単項演算は, 絶対値, 初等関数, 等で, 名詞表の対象属性が数であるときに, 要素ごとに関数を用いて関数値に変える. 限定属性はすべてそのままである.

2項算術演算は, 4則演算, 等である. 2つの名詞表の対象属性が数であるとき, 要素のすべての対にたいして演算を行い, 結果を対象属性値とする名詞表を作る. いわば外積を求めることになる. ただし限定属性がある

「在庫量-適正在庫」	C5	C8
数量	部品	倉庫
30	ラジエーター	1号倉庫
-20	ラジエーター	2号倉庫

図6 限定属性のある表の2項算術演算

Fig. 6 Binary arithmetic operation with attributes.

場合には, 統合が行われるべき対にたいしてだけ演算する. たとえば図1の「在庫量」から図5の「適正在庫」を引いた結果は図6の通りである. 統計関数は合計, 平均, 最大, 最小など単一の要素値を求める演算で, 同一の対象属性値が複数個あるときにそれを考慮する.

#### 4. ヤチマタの処理系

##### 4.1 ヤチマタの概略

ヤチマタにおける照会文の処理過程を図7に示す. 実線は処理の流れ, 点線は作業用データ構造の作成, 2重線は参照である. 影をつけたのは適用分野に依存する部分, その他がヤチマタの主要部分である.

辞書の主要部分は構文規則に現われる末端記号であり, 付加部分は, 名詞表・動詞表に現われるすべての詞辞を見出しとし, 品詞, 定義域, 動詞変化, 格標識, それに限定属性の見出し行, を項目としている. 分かち書きを自由にするため, (1)辞書参照のとき, 空白の後から始まる綴りだけでなく, 各文字から始まる綴りを扱う, とともに, (2)辞書の見出しに空白の可能な位置を明記しておく. たとえば「収納・する」と記載しておき, 「収納する. 収納する.」とは照合するが他の区切りとはしないようにする. なお数と部品コードなど, 数字・記号を含み一般単語と区別できるものは, 文字列規則に従って辞書参照の前に選出する.

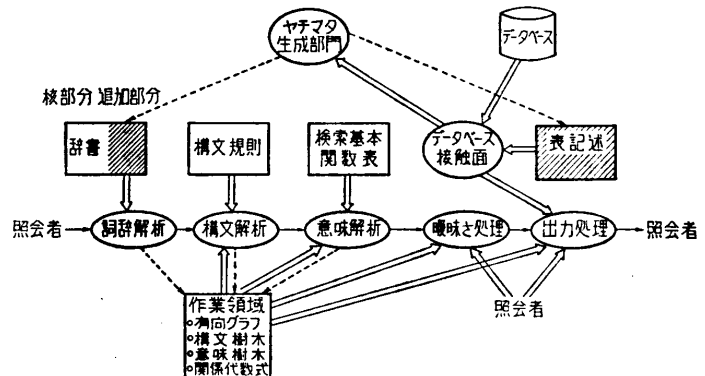


図7 照会文の処理過程

Fig. 7 Diagram of the query processing.

辞書参照の結果は一種の有向グラフであり、各辺には辞書から得た全情報が付記されている。構文解析法は REL<sup>13)</sup> でも用いられている Kay アルゴリズムで、上昇型シンタックス・ディレクティブ型である。解析過程も同様のグラフであり、最終結果は文全体に渡る樹木構造である。

構文規則の記述法は文脈自由文法に準ずるが、右辺の項には構文特徴の制約条件を付加することを許し、左辺には複数個の項を許し、その各項に構文特徴を与えることができる。書き換え以上の変形を許すため、構文規則適用の前または後に呼び出すサブルーチン名を左辺に指定できる。構文解析結果の樹木構造は同時に意味を表わし、節はデータ検索の基本関数に、枝はその引数に対応している。意味解析では、この樹木構造を関係代数式（一般には形式的なデータ準言語）に翻訳する。以上、詞辞解析から意味解析までは、より詳しく文献 14) に説明してある。

最後に、次節で述べる曖昧さ処理の後、出力処理が関係代数式をデータベースとの接触面に渡し、検索結果を編集表示する。必要ならば照会者は、表示する列、ソートの昇降順などを表示前に指定できる。

辞書の追加部分と、名詞表・動詞表とデータベースとの対応を記した表記述、などは応用システム生成のとき登録される。ヤチマタ生成部門は端末から対話的にデータを受け付けて、この登録をする。また生成部門は、データベースにあるデータとしての固有名を自動的に辞書に登録することも行う。

#### 4.2 流暢さと曖昧さ

照会文の流暢さ、つまり表現の自由度を増すために名詞句データ模型では辞詞の多重な定義を許した。たとえば「倉庫、所在地、ある」を適当に定義すれば、「港区の倉庫」「港区にある倉庫」「所在地が港区の倉庫」「所在地が港区である倉庫」「倉庫で港区にあるもの」などの表現が可能である。階層のある概念も、システム生成の段階でそれぞれを定義すれば、容易に明確に扱える。たとえば「東京」という名詞表を、「千代田区、港区、…」を要素とする対象属性だけから成るものと定義すれば、「所在地が東京である倉庫」も許される。

照会中に語彙を増すためには、すでに定義されている詞辞で新しい詞辞を定義する機能が要求される。このとき単に同義語を

定義するだけならば簡単であるが、「(太郎) 祖父=(太郎) 父の父」のように、パラメータを含む定義を許そうとすると種々の問題を発生するので、設計はしたが実動化は試みなかった。

「ラジエターの在庫量が倉庫 1 (のラジエターの在庫量) よりも多い倉庫」のような省略文を許そうとすると、〈名詞句 1〉〈が〉〈名詞句 2〉〈よりも多い〉〜という構文の〈名詞句 1〉の範囲を定めねばならないが、名詞句模型だけでは完全な策は得られない。その他、流暢さの困難について文献 15) で若干論じている。

照会文の曖昧さの原因には次の 4 つがある。(1) 修飾や否定の範囲が不明で、人間をも迷わす場合。特に並列句を含む場合に生じる。(2) 詞辞に複数個の意味をもたせるため。特に「ある、もつ」などの基本動詞は多様な意味を与えることになる。(3) 仮名使用による同音異義。期せずして「高知の耕地」という例に出会った。(4) 分かち書きの自由によるもの。「の、に、と」などの字は、どこに現われても助詞とみなされる。「小学生」という名詞だけが定義されているとき「小学生徒 1 人当り」を入力してしまったところ、妙な分析がされ当惑した。「在る、居る」を動詞として定義すると、「である、している」などの中に現われる。

これらの曖昧さの処理のために、詞辞・構文・意味の解析過程ではすべての可能性を残しながら進み、適合しなくなった点で捨てる。構文分析で品詞、活用形の検査と、格標識、定義域、修飾形のゆるい検査を行う。これにより上記の(1)以外はほとんどの場合に解決される。意味解析が終わったときに複数個の解釈があったときには、できるだけ分かり易い形で表示

```

シフ"ヤク ニ アリ ラジ"イター ノ アル ソウコ ニ ノウニョウシテイル キ"ョウシヤ ハ?
***コノ フ"ンショウ ハ アイマイ テ"ス、 タタ"シイ モノ ヲ イランテ" クタ"サイ。***

*'? 'ガ" フ"イタ フ"フ"ン ハ アイマイ テ"ス。

シフ"ヤク ニ アリ ラジ"イター ノ アル ソウコ ニ ノウニョウシテイル キ"ョウシヤ ハ?
???????

1:
シフ"ヤク ニ アリ ラジ"イター ノ アル ソウコ ニ ノウニョウシテイル キ"ョウシヤ ハ?
>>>>>>>>>.....****

2:
シフ"ヤク ニ アリ ラジ"イター ノ アル ソウコ ニ ノウニョウシテイル キ"ョウシヤ ハ?
>>>>>>>>>.....*****

タタ"シイ モノ ノ ハ"ンコ"ウ ヲ イレテ クタ"サイ。

```

図 8 曖昧さを含む場合の選択メニュー  
Fig. 8 A menu for resolving ambiguity.

し、照会者に選択を任せる。図8はその例で、曖昧の原因は(2)の型であり(1)の型でもある。

同音の名詞表が複数個あるとき、構文分析の段階で曖昧さが確定することもある。たとえば「(航空便)の(空港)発の時刻」と「(航空便)の(空港)着の時刻」をあえて定義し、「JAL 6 便の時刻は」と尋ねる場合である。名詞の定義段階で準備しておけば、どちらかを照会者に選択させることができる。

#### 4.3 実働化

ヤチマタはほとんど PL/I で書かれ、アイ・ビー・エム仮想計算機能/370・会話型モニター・システム (VM/CMS) の下で開発され、稼動している。データベースには PRTV<sup>16)</sup> を用いている。その照会言語はほとんど関係代数そのものであるが、除算はない。汎用性を示すために、実験的な応用として、(1)自動車部品の在庫、(2)国内航空便の発着時刻、(3)都道府県別地域統計、の照会システムが動いている。これらの照会例文などは文献 17) に譲る。

実験のため短大・大学卒の、情報処理について何の知識もない女性タイピスト3人に、都道府県別地域統計の照会を練習してもらった。機械の操作も含めて約3時間の説明の後、午後2回、例題を含め自由に照会回答を試みてもらった結果、何をどのように聞けるかほとんど理解してくれた。

最近日本語による質問解答システムが発表されているが<sup>18)~21)</sup>、データベース照会という汎用性を旨とした点においてヤチマタは新しい。現時点で実用化上の難点は処理時間と、作業領域の大きさである。

#### 5. むすび

日本語、特に仮名書き文による機械との対話には、種々の困難がある。しかしデータベースの照会に限るならば十分に正確な多様な照会が可能であること、その基礎として名詞句データ模型が有効であることを示した。データベースの場合にはデータ管理者により扱う情報の範囲・内容が明確に規定されている、と期待できる好条件が備わっている。

本稿では日本語照会文についてだけ議論したが、名詞句データ模型の概念そのものは個別言語によらない。言語ごとの修正は当然必要であるが、日本語はむしろ名詞句模型向きであると予想している。

#### 参考文献

1) M. M. Zloof: Query-By-Example: The Innovation and Definition of Tables and Forms,

- Proc. Conf. VLDB, Vol. 1, No. 1, pp. 1-24 (1975).
- 2) W. A. Woods: Progress in Natural Language Understanding—An Application to Lunar Geology, Natnl. Comp. Conf., 28, pp. 441-450 (1973).
- 3) W. J. Plath: The REQUEST System, IBM Res. Tech. Repart, RC 5604, p. 20 (Aug. 1975).
- 4) J. Mylopoulos, et al. TORUS: A Step towards Bridging the Gap between Data Bases and the Casual User, Inf. Syst., Vol. 2, pp. 49-64 (1976).
- 5) G. G. Hendrix, et al.: Developing a Natural Language Interface to Complex Data, Proc. 3rd Conf. VLDB, p. 292 (1977).
- 6) L. R. Harris: User-oriented Data Base Query with the ROBOT Natural-language Query System, Proc. 3rd. Conf. VLDB, pp. 303-311 (1977).
- 7) 足立巻一: やちまた, 河出書房新社 (1974).
- 8) E. F. Codd: A Relational Model of Data for Large Shared Data Banks, Commun. ACM, Vol. 13, No. 6, pp. 377-387 (1970).
- 9) 穂鷹良介, 渋谷政昭: データベースの関係形式, 情報処理, Vol. 17, No. 10, pp. 904-910 (1976).
- 10) M. M. Astrahan, et al.: System R: Relational Approach to Database Management, ACM TODS, Vol. 1, No. 2, pp. 97-137 (1976).
- 11) E. F. Codd: Seven Steps to RENDEZVOUS with the Casual User, Proc. IFIP Working Conf. DBMS, Cargese, Corsica, North-Holland (April 1974).
- 12) C. J. フィルモア (田中春美・船城道雄訳): 格文法の原理, 三省堂 (1975).
- 13) F. B. Thompson: REL—A Rapidly Extensible Language System, ACM Natnl. Conf., 24, pp. 399-417 (1969).
- 14) 諸橋正幸, ほか: ヤチマタにおける擬似日本語, 情報処理学会計算言語学研資 8 (1976).
- 15) T. Fujisaki et al.: A System for Querying Regional Statistics, Tech. Conf. Relational DB (June 1976), Bari Sci. Ctr., IBM Italy.
- 16) S. Todd: PRTV, An Efficient Implementation for Large Relational Data, Proc. Conf. VLDB, Vol. 1, No. 1, pp. 554-556 (1975).
- 17) 藤崎哲之助, ほか: 日本語によるデータベース照会回答システム「ヤチマタ」, IBM Review, No. 63, pp. 69-76 (1976).
- 18) 長尾 真, 辻井潤一, 田中一敏: 意味および文脈情報を用いた日本語文の解析, 情報処理, Vol. 17, No. 1, pp. 10-18, pp. 19-28 (1976).
- 19) 雨宮真人, 島津 明, 若菜 忠, 若山忠雄, 酒井士朗: 図形世界を話題とした質問解答システム, 情報処理, Vol. 18, No. 8, pp. 799-807

- (1977).
- 20) 大須賀節雄: 質問応答のための知識表現と推論方式, 昭和51年電気四学会連合大会, pp.7-117-120 (1976).  
の支援システム, 昭和51年電気四学会連合大会, pp. 7-121-124 (1976).  
(昭和53年3月15日受付)  
(昭和53年6月5日採録)
- 21) 田中穂積, 佐藤泰介: 日本語応答システムとそ
-