

基本関係データベース演算による自然言語表現の変換†

上林 弥彦** 天野 浩文**

一般利用者にとってより使いやすいデータベースシステムを実現するため、形式的な質問言語で書かれた質問をその意味に対応する自然言語表現に変換する方法について考察した。この問題は、質問作成支援システムのほか、自然言語によるデータベース質問の処理にも応用できる。本稿で述べる方式は、自然言語を扱うための複雑な処理を避けるため、制限された英文の集合に単純な文字列操作を施すことにより、質問の解となるデータの意味に対応する表現を得るものである。そのため、主として、この英文集合の満たすべき構文上の制限と、これらの英文の質問に応じた変換法について検討した。本稿では、関係代数の基本的な3つの演算(選択・射影・結合)を扱う。いくつかの表を結び付け(結合)、それによってできた表からある条件を満たす組を選び出し(選択)、必要な列を切り出す(射影)という質問に対して、用意された文集合から選び出された表現集合を連結して、それらに修飾語句を加え、出力の部分強調する、という処理を提案する。この方法で生成される表現によれば、元の関係がどのような操作を受けるかを推定するのも容易である。さらに、上記の出力文に関する考察から、質問の構造との明確な対応をもつ単純な表現が作りにくいような質問のクラスが存在することも示す。

1. ま え が き

近年普及が著しい関係データベース^{1),2)}の長所を生

も修正しなければならない可能性があり、自然言語の利点を十分に生かせるとは限らない。後者の方法では、ある特定のデータベースに適合させるために与え

門家でない利用者でも容易に作成できなければならない。しかし、計算機の処理しやすい形式的な言語を用いることは、人間にとって必ずしも容易ではない。

そこで、本論文では、形式的な言語で作成した質問に対応する自然言語表現を単純な方法で生成することを考え、第一段階として、ある制限されたクラスの質問から自然言語表現を生成する方法について述べる。この手法で生成される表現は形式的な言語で記述された質問の意味を表し、利用者が自分の質問の意味を確認するのに役立つ。

対象となるデータベースの構成を変更したい場合の仕事量が増加する傾向にあり、システム自体のオーバーヘッドも大きい。いずれの方法でも、自然言語インタフェースはまだ必ずしも成功しているとは言えない。

一方、初心者に自然言語で指示を与え、初心者向けの形式的なデータベース言語として知られるQBE^{1),2)}で質問を書かせた場合、指示どおりの答を得られない質問が33%も存在するにもかかわらず、そのうちの82%が構文的には正しいという報告⁶⁾もある。すなわち、利用者の意図する質問であるかどうかの判定には

が、これらの情報がデータベース全体をひとまとめにして扱うため、大規模なデータベースにおける関係の挿入/削除といった変更の際の保守が複雑になる。文献 8) は、個々のデータベースに依存する意味的な情報を、答を求める条件に関する属性および答として出力される属性との対に対応する二項関連の表現の形で記憶するため、多項関連の表現に問題がある。

これに対し、われわれの方法は、データの意味を表す英文の集合を質問に応じて機械的に変換することにより求める英文を得るので、自然言語処理の困難な問題を避けることができる。あらかじめ作成しておく表現を1つの関係の中に局所化している点は文献 7) の手法と異なっている。また、1つの単純な文で表せるような属性集合として、対象 (object) を導入しているため、多項関連の表現も二項関連と同様に自然な形で導入できる。本論文では関係データベース質問の重要な部分クラスである SPJ 質問について扱うものとする。

2. 基本的事項

2.1 関係データベースとそれに対する質問

関係データベースは関係の集まりであり、関係は1枚の表に対応させることができる (図 1)。表に見立てた関係の列はそれぞれ1つの属性に対応し、同じ表の行は組 (関係の要素) に対応する。関係 R を構成する属性集合を関係スキーマと呼び、 R で表現する。

関係代数は、関係に対する演算の系列で質問を記述する。基本的な演算には次のようなものがある。これらの例を図 1 にまとめて示す。

[射影 (projection)] 表の列のうちから、必要なものだけを取り出す操作を射影という。属性集合 $X \subseteq R$ に対し、 R の組 r の X の部分のみを集めたものを $r[X]$

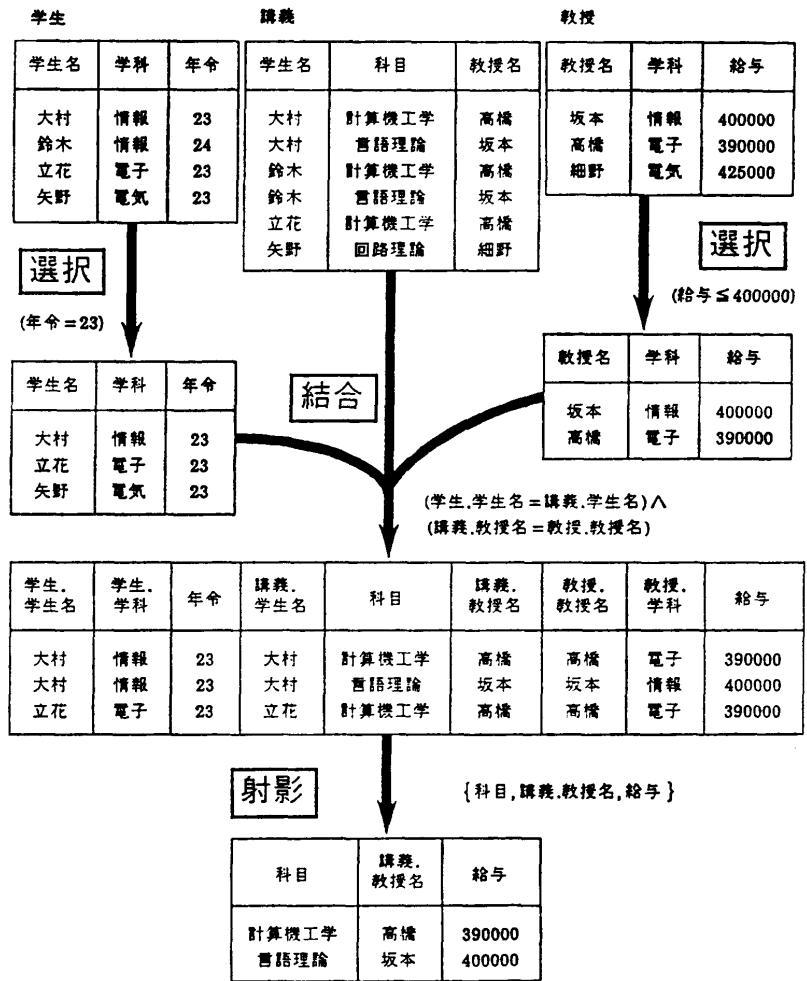


図 1 関係および関係データベース演算の例
 Fig. 1 An example of relations and relational database operations.

で表すとき、 R の X への射影は次の式で定義される：
 $R[X] = \{r[X] | r \in R\}$.

[選択 (selection)] 表の中から、ある条件を満たすものだけを取り出す操作を選択という。 θ を比較演算子 $=, \neq, <, \leq, >, \geq$ のいずれかとするとき、関係 R の属性 A に関する θ 選択を次のように定義する：

$$R[A\theta B] = \{r | r \in R \wedge (r[A]\theta r[B])\}.$$

ただし、 A と B とは比較可能な定義域をもつ属性でなければならない。

[結合 (join)] 2つの関係の組をつなぎあわせて得られる新たな組を集め、幅の広い新たな関係を得る操作を結合という。2つの組を結び付けるのは、それぞれの組の間に与えられた条件が成立する場合に限定する。この条件を結合条件と呼び、2つの組の属性値の比較で表す。同じ関係どうしの結合を特に自己結合と

呼ぶ。属性 $A(\in R)$ と属性 $B(\in S)$ が同一の定義域をもつとき、 R と S の θ 結合を次のように定義する：

$$R[A\theta B]S = \{t \mid t[R] \in R \wedge t[S] \in S \\ \wedge (t[A]\theta t[B])\}.$$

ただし、 T を結合後の関係とすると、 T は R と S の直和である。

選択や結合は複数個の属性についても適用できる。2つの関係に共通する全属性について等結合した場合、同じ値になる列の対ができる。このような値の重複の除去まで行う演算を自然結合といい、 $R * S$ で表す。

選択・射影・結合のほか、関係除算や和集合、共通集合、直積などの集合演算がある。関係除算は他の演算で表現できることが知られている。選択・射影・結合の3演算のみで表現できる質問を選択・射影・結合 (selection/projection/join, SPJ) 質問という。

2.2 SPJ 質問と超グラフ

超グラフ (hypergraph) は、節点集合 N と枝集合 E の対 (N, E) で定義され、枝は節点集合の空でない部分集合、 N は各枝の節点の和集合である。

x_1, \dots, x_m が相異なる節点、 S_1, \dots, S_m が相異なる枝であるとき、次の3つの条件を満たす連鎖 $(S_1, x_1, S_2, x_2, \dots, S_m, x_m, S_{m+1})$ を Berge-閉路と呼ぶ。

- (1) $m \geq 2$
- (2) $\forall i(1 \leq i \leq m); x_i \in S_i \wedge x_i \in S_{i+1}$
- (3) $S_1 = S_{m+1}$

超グラフが Berge-閉路をもつとき (図2 (a), (b) において破線で示す)、Berge-巡回であるといい、そうでないとき (図2 (c)), Berge-非巡回であるという。

関係を枝、属性を節点としたときの超グラフは、

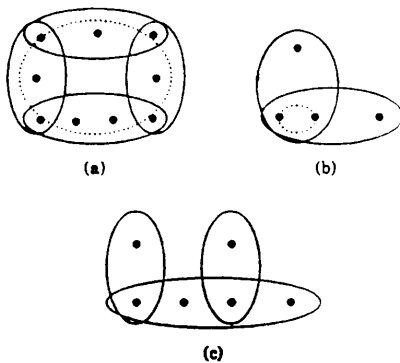


図2 超グラフの例
Fig. 2 Examples of hypergraphs.

SPJ 質問における結合演算の構造を表すことができる。

2.3 従属性と正規形

データは実世界の事物のもつ何らかの制約を反映していなければならない。このような意味的制約の中で特に重要なものは、従属性^{1),2)}である。

属性集合 X の値を決めると属性集合 Y の値が一意に決まる場合に、 X から Y への関数従属性 (functional dependency, FD) が成立する、といい $X \rightarrow Y$ で表す。 $X \rightarrow X'(X' \subseteq X)$ のような FD を自明であるという。 $X \rightarrow Y$ が $R(XYZ)$ で成立するとき、 R を $R[XY]$ と $R[XZ]$ に分割しても情報は失われない ($R \equiv R[XY] * R[XZ]$, XY 等は属性の和集合を表す)。

関係 R において、 $K \rightarrow R$ となるような極小の属性集合 $K(\subseteq R)$ を R のキーと呼ぶ。

関係 R が $R[X_1] * R[X_2] * \dots * R[X_n](\cup X_i = R)$ と常に等しいとき、結合従属性 (join dependency, JD) $*[X_1, \dots, X_n]$ が成立するという。したがって、FD が成立すると、対応する JD が導ける。 $X_i = R$ となるような i が存在する場合、その JD は自明であるという。以下では、 $X_i \subseteq R$ で $X_i \subseteq X_j (i \neq j)$ であるような $i, j (1 \leq i, j \leq n)$ のない JD を考える。

非キー属性が“ \rightarrow ”の右辺に現れるすべての自明でない FD の左辺がキーもしくはキーを含む属性集合になっており、自明でなく FD でも表せない JD を含まない関係は FD-JD 正規形であるという。このような関係は、更新に際し FD や JD の保持が容易になるので、データベース設計の1つの指針となっている。

3. 関係の意味を表す標準表現群

本章では、あらかじめ人間が作成しておき、システムに与えておく表現集合の作成法について述べる。これらの表現は、質問の意味を表す英文を作成する材料になる。ただし、本稿では、関係はすべて FD-JD 正規形に分解されているものとする。

関係の属性に対応する名詞句を含む自然な文は、それらの属性の意味を表していると考えることができる。ただし、関係の意味を表すためには、少なくとも、その関係の全属性に対応する名詞句を含む表現が必要である。また、それらの表現をできるだけ分解し、必要に応じてその中から選び出して連結できるようにしておく、後に述べる変換操作が単純になる。

本稿では、表現の分解、すなわち、属性集合の分解のための道具として、対象 (object) を導入する。対象

は1つの単純な文に対応するような意味のある属性集合であり、次の条件のいずれかによって定義される。ただし、 K はキー、 A は単一の非キー属性を表す。

(O1) K が単一の属性であるとき、 $K \rightarrow A$ で与えられる属性集合 KA

(O2) K が複数の属性からなり、

(O2-1) K のすべての属性名を含む名詞句が主語となるような自然な文が存在するとき、 $K \rightarrow A$ で与えられる KA

(O2-2) 上記の条件が成り立たないとき、 R

(O3) $O_1 \cap O_2 \neq \emptyset$ となる対象 O_1, O_2 の和集合特に、(O1) または (O2) による対象を基底対象 (base object), (O3) による対象を派生対象 (derived object) と呼ぶ。

関係 R における $FD: K \rightarrow (R-K)$ を考える。キーは複数存在しうるが、そのうち1つを選べば R の全属性が “ \rightarrow ” の左右どちらかに現れる。 $K \rightarrow A_1 \dots A_i (A_i: \text{単一の属性})$ のような FD は、 $K \rightarrow A_1 \dots A_{i-1} A_{i+1} \dots A_j$ と $K \rightarrow A_i$ から導出できる。 $i=1, \dots, j$ について、 $K \rightarrow A_i$ を表す文が KA_i 属性に対応する名詞句を含んだ全んでいれば、これらのすべての重ね合わせによって、 R のすべての属性に対応する名詞句を含む表現集合が得られる。

ただし、関係の意味を表す文が、 KA に対応する文の重ね合わせとなる場合と R に対応する文となる場合とに分かれているのは、次のような理由による。

R のキー K を主語にできれば、右辺が単一の属性であるような各 FD の意味を表すすべての文が、 K に対応する主語を共有することができる。このとき、各属性に対応する名詞句は1度しか現れない。一方、 K の属性すべてを含む自然な名詞句がないときは、 K の一部の属性だけを共有させることになる。たとえば、 $R = \{A, B, C, D\}$, $K = \{A, B\}$ で主語が A のみを含むとすると、 $AB \rightarrow C$, $AB \rightarrow D$ の文を分割して考えると、

〈主語〉 〈 B と C に対応する動詞句〉,
and 〈 B と D に対応する動詞句〉.

のような文を用いなければならない。この場合、後で述べる変換操作が属性に対応する名詞句の後に修飾語句を挿入することがあるので、変換後の表現が読みにくくなるおそれがある。したがって、文を分割して考えるかわりに R 全体に対応する文を1つのまとまりとして扱う。

関係の意味を表し、質問の意味表現を生成するもと

になる標準表現は、データベース管理者等が次のような規則にもとづいてあらかじめ作成しておく。

(C1) 標準表現は1つの基底対象について1つの文を用いるものとし、句点を用いずに表現する。

(C2) 属性名のうち、省略形を用いてあるものや、何か他の語句を加えた方がわかりやすいものは、属性名にそれらの語句を加えて1つの名詞句とする。

(C3) 条件 (O1), (O2-1) による基底対象の標準表現の主語はキー K に対応する名詞句とし、 K を共有する対象が複数個存在するときは、それらの文で共通になるようにする。条件 (O2-2) による基底対象の標準表現の主語は、 K の空でない部分集合に対応する名詞句とする。

(C4) 各基底対象の主語でない属性が単一の述語動詞の目的語、補語、あるいは、修飾語句の一部となるように適当に動詞を選び、標準表現とする。

(C5) 関係詞節の使用はなるべく避ける。常に大文字となる場合 (Prof. 等) を除きすべて小文字を用いる。

次の (a) ~ (c) は関係スキーマおよびそれに対応する標準表現の例で、属性名に下線が引かれているのは、それがキー属性であることを示している。

(a) STUDENT (SNAME, DEPT, AGE)

student {SNAME}
: belongs to the department {DEPT}
: is {AGE} years old

(b) PROFESSOR (PNAME, DEPT, SAL)

Prof. {PNAME}
: belongs to the department {DEPT}
: earns a salary of {SAL}

(c) LECTURE (SNAME, SUBJECT, PNAME)

student {SNAME}
: takes the lecture {SUBJECT}
given by Prof. {PNAME}

規則 (C1), (C2) は、FD-JD 正規形における各基底対象に、単純でわかりやすい文を対応させることに相当する。規則 (C2) は属性名が語句の省略形である場合等にそれが指す事物をはっきりさせるためのもので、“student {SNAME}” 等がその例である。(C3) により、(a), (b) ではキーに対応する語句が、(c) ではキー属性の1つに対応する語句が、それぞれ主語に選ばれている。(C3) により、多項関連に対する表現も自然な形で導入できる (例文 (c))。規則 (C4) により、主語に選ばれなかった属性に関する記述が標準

表現の中に組み込まれる。規則 (C5) により、最終的な出力文の頭に小文字が来たら機械的に大文字にするだけでよい。

続いて、各対象に関して、各属性に対応する名詞句を先行詞とする副節（関係副節、前置詞句）を考える。標準表現の主語を先行詞とする副節は、主語の直後に“that”を挿入するだけで得られる。主語以外の部分に含まれる属性については、その名詞句を先行詞とする副節を作成しておく。以下に副節の例を示す。

- (a') STUDENT (SNAME, DEPT, AGE)
 the department {DEPT}
 : that student {SNAME} belongs to
 an {AGE}
 : of student {SNAME}
- (b') PROFESSOR (PNAME, DEPT, SAL)
 the department {DEPT}
 : that Prof. {PNAME} belongs to
 a salary {SAL}
 : that Prof. {PNAME} earns
- (c') LECTURE (SNAME, SUBJECT, PNAME)
 the lecture {SUBJECT}
 : given by Prof. {PNAME}
 and taken by student {SNAME}
 Prof. {PNAME}
 : who gives the lecture {SUBJECT}
 taken by student {SNAME}

データベース管理者等によって作成されたこのような表現集合は個々のデータベースに依存する意味的な情報を表しており、次章で述べる変換・連結の操作の基本的な単位となる。

4. 選択・射影・結合演算に対応する変換

一般に、SPJ 質問は、いくつかの関係を結合し、その結果の属性の属性に条件を加えて選択を行い、必要な属性だけを射影する、という形で表せる。以下、この形の、自己結合を含まない SPJ 質問だけを考え、結合、選択の条件となった属性集合をそれぞれ J , S 、最終的に射影される属性集合を P とする。

このような SPJ 質問に対応する表現の生成手順の概要は次のようになる。

①対象の決定

$SUPUJ$ を被覆する基底対象（あるいは基底対象集合）を決定する。これに対応する文を基礎文と呼ぶ。

②結合に対応する変換

基礎文に対し、結合される関係に含まれる対象に対応する副節を埋め込む。

③選択に対応する変換

S に対応する語句の後に選択条件を示す修飾語句を埋め込む。

④射影に対応する変換

出力属性 P を強調するため、[] でくくる。

⑤出力形式の最終的な調整

自然な英文になるように、主語を共有する文が複数個ならぶときにはこの間に“and”などの接続詞を付加する。また、読みやすいように書式を整える。

段階(1) 対象の決定

SPJ 質問の意味を表すためには、質問に関与する属性集合 $SUPUJ$ のすべてに対応する名詞句の形で含む表現集合を用いる必要がある。すなわち、 $SUPUJ$ を含むような派生対象を求めなければならない。このための規則は次のように表せる。

(OS) 各関係 R_i について、 $(SUPUJ) \cap R_i$ を含み、かつ、できるだけ小さな対象を求める。複数個ある場合には基底対象の数の小さなものを1つ選ぶ。

例1 (((STUDENT [SNAME=SNAME]
 LECTURE) [PNAME=
 PNAME] PROFESSOR)
 [STUDENT. AGE=23,
 SAL \leq 400000]) [SUBJECT,
 LECTURE. PNAME, SAL]

この質問において、 $(SUPUJ) \cap STUDENT$, $(SUPUJ) \cap PROFESSOR$, $(SUPUJ) \cap LECTURE$ を被覆する対象は、それぞれ、{SNAME, AGE}, {PNAME, SAL}, {SNAME, PNAME, SUBJECT} である。

段階(2) 結合に対応する変換

この段階で、質問の意味を表す英文の骨格ができあがる。このための規則は次のようになる。

(J1) 段階(1)で得られた対象のうち、属性数の最大のもを1つ選び、その標準表現を土台の文とする。

(J2) 上記以外の対象（を構成する基底対象）を求め、これらのうち、(J1)の対象と共通部分をもつものの副節（共通部分の属性を先行詞とするもの）を括弧付きで埋め込む。ただし、 $O_1 \cap O_2$ は結合条件 $A \theta B$ ($A \in O_1$, $B \in O_2$) となった属性であり、 O_1 が土台であるとき、 $\langle A \rangle$ の後に次のような語句を挿入する ($\langle A \rangle$)

表 1 関数 $E(\theta)$
Table 1 Function $E(\theta)$.

θ	$E(\theta)$
=	equal to same...as (that)
≠	greater or less than different from
>	greater than alphabetically ordered after
<	less than alphabetically ordered before
≧	not less than equal to or alphabetically ordered after
≦	not greater than equal to or alphabetically ordered before

は A に対応する名詞句の部分を表す).

(J2-1) A と B の名前が同じとき

θ が = のとき (O_2 の副節)

θ が = 以外のとき ($E(\theta)$ one O_2 の副節)

(J2-2) A と B の名前が異なるとき

$E(\theta)\langle B \rangle$ (O_2 の副節)

ただし, $E(\theta)$ は比較演算子 θ に応じて例えば表 1 のような文字列を与える関数である.

(J3) すべての基底対象の表現が連結されるまで, (J2)と同様の操作を繰り返す.

表 1 の $E(\theta)$ は給与, 得点などの数値属性に対応する名詞句の修飾に用いる文字列の例である. しかし, 年令や重量のようにそれ特有の修飾語句を用いた方がよい場合がある. また, 等号で結合される属性 A と B が同一名である場合には「 B に等しい A 」ではなく「 B と同一の A 」と解釈すべきであるので規則(J2-1)を設けている.

例 2 例 1 の質問に対し, 埋め込みの土台になる文は,

student {SNAME} takes the lecture {SUBJECT}
given by Prof. {PNAME}

であり, 埋め込まれる副節は,

{SNAME} <- that is {AGE} years old
{PNAME} <- that earns a salary of {SAL}

の 2 つである. 変換後の表現は, 次のようになる.

student {SNAME} (that is {AGE} years old)
takes the lecture {SUBJECT} given by Prof.
{PNAME} (that earns a salary of {SAL})

段階(3) 選択に対応する変換

選択に対応する変換規則は, 上記の $E(\theta)$ を用いる

と, 次のようになる.

(S1) 属性-属性の比較: $r[A]\theta r[B]$

(S1-a) A, B が数値データの時
 $\langle A \rangle$ の後に次の語句を挿入する.

$E(\theta)\langle B \rangle$

(S1-b) 非数値データの時

$\langle A \rangle$ を次の語句で置換する.

= the same $\langle A \rangle$ as $\langle B \rangle$

≠ $\langle A \rangle$ different from $\langle B \rangle$

(S2) 属性-定数の比較: $r[A]\theta 'c'$

= $\langle A \rangle$ をその定数で置換する.

それ以外 (S1) のときの変換を行い $\langle B \rangle$ をその定数で置換する.

(S1)により標準表現中の A に関する記述に, θ に応じた修飾語が付く. ただし, この変換が適合しない文もありうるので, そのような文については, 3章で述べた表現作成の段階で, あらかじめ代替表現を作っておかなければならない. “=” の場合は, そのまま置換する方がわかりやすい.

例 3 例 2 の表現に対して, (STUDENT. AGE = 23) \wedge (SAL \leq 400000) なる選択演算に対応する変換を施した後の表現は次のようになる.

student {SNAME} (that is 23 years old) takes
the lecture {SUBJECT} given by Prof. {PNA-
ME} (that earns a salary of {SAL} (less than
400000))

段階(4) 射影に対応する変換

(P) 最終的に出力される属性集合 P を強調するため, 属性名を囲む { } を [] に替える.

例 4 例 3 の表現に, SUBJECT, PNAME, SAL への射影演算に対応する変換を加えた後の表現は次のようになる.

student [SNAME] (that is 23 years old) takes
the lecture [SUBJECT] given by Prof. [PNA-
ME] (that earns a salary of [SAL] (less than
400000))

段階(5) 出力形式の調整

出力文の生成にはさらに以下の変換が必要である.

(T1) 次の表現を先頭に配置する.

List [出力属性] (の並び) such that :

(T2) 各文の文頭が小文字であれば大文字にする.

規則にしたがって生成される英文の例を以下に示す.

例 5 例 1 の質問は, 23 歳の学生の受講する科目と

その担当教官で給与 400000 以下の者の名前と給与を求める質問である。

List [SUBJECT], [PNAME], [SAL] such that:

Student [SNAME] (that is 23 years old) takes the lecture {SUBJECT} given by Prof. [PNAME] (that earns a salary of [SAL] (less than 400000)).

5. 質問の構造と自然言語の構文との関連

本稿で提案した方式は、構文の制限された文集合に対し選択・射影・結合演算に応じた変換を加えるので、単純で平明な表現が得られる。さらに、この表現は元の質問の構造をある程度反映するという特徴がある。この性質から、各関係とその間の関連を明示した英文で表現しにくい質問のクラスを考える。

本方式では対象の構造によって文の骨格を決定するため、本章では属性を節点、対象を枝としたときの超グラフを考える。この超グラフが Berge-巡回となる時、次の2種類の不都合が生じる。

(1) 超グラフが図2(a)のような形状になると、結合される2つの属性に対応する名詞句が離れる。

(2) 超グラフが図2(b)のような形状になると、2つ以上の先行詞を同時に修飾できる自然な副節が存在しないことがある。これらを無理に分解すると上と同様の問題が起こる。

これらの問題点は、遠方の語句を参照する表現が必要となる場合に対応している。このような指示語があいまいにならないようにするためには、特別な配慮が必要になる。文献7)には、処理の困難な自然言語質問が存在するとの指摘があった。しかし逆に、質問の対象構造を反映した単純な英文によって記述しにくい質問も存在することを、上記の問題点は示唆している。

例6 次の文には2つの DEPT が現れるが、これらが結合されることをこれだけでは明示できない。

Student {SNAME} (that belongs to the department {DEPT}) takes the lecture {SUBJECT} given by Prof. {PNAME} (that belongs to the department {DEPT}).

例7 前述の関係 STUDENT と属性集合 {DEPT, AGE} を共有する対象が存在するとき、次の文の DEPT と AGE を同時に修飾できる副節は存在しない。

い。

Student {SNAME} is {AGE} years old, and belongs to the department {DEPT}.

6. むすび

本稿の方法により、自然言語処理の困難さを避けた上で、自己結合を含まず、対象の構造が Berge-非巡回となる質問の意味に対応する平明なあいまいでない表現を生成できる。この方式の利点は、システムがあまり複雑にならないこと、質問修正のオーバーヘッドもあまり大きにならないこと、個々のデータベースへの個別化のための情報を個々の関係について局所的に与えればよいこと、などである。

日本語は語順が必ずしも一定でなく、自然言語処理においては、特別な配慮が必要である^{11),12)}が、本方式を日本語に応用する場合、元になる日本語文に制限を加えるので、困難な問題は避けることができる。

本稿では、自己結合を含まず、対象の超グラフが Berge-非巡回となるような選択・射影・結合演算のみを検討した。さらに広範囲の質問を扱えるようにするためには、今後、上記の制限を緩和する必要がある。Berge-巡回超グラフをもつような質問については、その超グラフを Berge-非巡回な複数の超グラフに分割して処理する方法が考えられる。また、よりわかりやすい表現を簡単に生成できるようにするため、名詞の単数/複数形の変換や、動詞の否定形の扱いなども今後の検討課題である。

謝辞 本研究に関し、熱心に御討論・御助言いただく本学情報工学科今井浩助教授、古川哲也助手、ならびに大学院生 Mohamed El-Sharkawi 氏をはじめとする研究室の皆様へ深謝いたします。

また、本研究は、昭和63年度文部省科学研究費補助金特定研究「言語情報処理の高度化」のうちの「情報ドキュメンテーションのための言語の研究」の援助を受けている。

参考文献

- 1) Ullman, J. D.: *Principles of Database Systems*, 2nd ed., Computer Science Press (1982).
- 2) Date, C. J.: *An Introduction to Database Systems Vol. I*, 4th ed., Addison-Wesley (1986).
- 3) Codd, E. F.: How about Recently? (English Dialog with Relational Databases Using RENDEZVOUS Version 1), *Databases: Improving Usability and Responsiveness*, pp. 3-28, Academic Press (1978).

- 4) Mylopoulos, J. et al.: TORUS—A Natural Language Understanding System for Data Management, *Proc. 4th Int. Joint Conf. on Artificial Intelligence*, pp. 414-421 (Sept. 1975).
- 5) Damerau, F. J.: Problems and Some Solutions in Customization of Natural Language Database Front Ends, *ACM Trans. Office Inf. Syst.*, Vol. 3, No. 2, pp. 165-184 (April 1985).
- 6) Thomas, J. C. and Gould, J. D.: A Psychological Study of Query by Example, *Proc. AFIPS National Computer Conf.*, pp. 439-445 (May 1975).
- 7) Lowden, B. G. T. and De Roeck, A. N.: The REMIT System for Paraphrasing Relational Query Expressions into Natural Language, *Proc. 12th Int. Conf. on Very Large Data Bases*, pp. 365-371 (Aug. 1986).
- 8) Luk, W. S. and Kloster, S.: ELFS: English Language from SQL, *ACM Trans. Database Syst.*, Vol. 11, No. 4, pp. 447-472 (Dec. 1986).
- 9) 上林, 天野: 自然言語表現に対するデータベース操作, 情報処理学会自然言語処理研究会資料, 85-NL-52 (1985. 11).
- 10) Kambayashi, Y.: An Overview of a Natural Language-Assisted Database User Interface: ENLI, *Proc. IFIP 10th World Computer Congress*, pp. 1055-1060 (Sept. 1986).
- 11) 長尾: 機械翻訳, 電子通信学会誌, Vol. 65, No. 4, pp. 386-392 (1982. 4).
- 12) 吉田, 日高: 機械翻訳のための構文解析手法, 情報処理, Vol. 26, No. 10, pp. 1157-1164 (1985. 10).

(昭和63年9月1日受付)
(平成元年7月18日採録)



上林 弥彦 (正会員)

昭和18年生。昭和45年京都大学大学院博士課程電子工学修了。京都大学助手, イリノイ大リサーチアシェイト, 京都大学助教授を経て昭和59年より九州大学工学部教授。

論理回路・オートマトン, データベースの研究に従事。この間, カナダマッギル大学, クェート大学, 中国武漢大学の客員教授。「Database—A Bibliography」(Computer Science Press)「データベース」(昭晃堂)など。50年米沢賞, 58年丹羽賞, 62年電子情報通信学会著述賞。電子情報通信学会, ACM, IEEE 各会員, 本会理事。



天野 浩文 (正会員)

昭和38年生。昭和61年九州大学工学部電子工学科卒業。昭和63年同大学大学院工学研究科修士課程情報工学専攻修了。現在同博士後期課程在学中。関係データベースの利用

者インタフェース, 自然言語処理に興味を持つ。ACM会員。