

6H-1

## 従属性制約の保持と質問処理効率化による ネットワークデータベース設計法の比較

古川 哲也 上林 彌彦

(九州大学工学部)

### 1. まえがき

データベースにおける重要な問題として、意味制約の保持と質問処理の効率化がある。ネットワークモデルに基づくデータベースではデータの対応をリンクで表しているため、そのスキーマ設計では両方を考慮しなければならない。これまでの設計法は経験的なものであったが、体系的に設計を行う方法として、

(A)意味制約を反映した非冗長な構造に対し質問処理効率化のための変換を行う。

(B)質問処理のみを考慮した構造に対し意味制約の保持のための変換を行う。

が考えられる。(A)については、関数従属性集合による設計<sup>[1]</sup>、多値従属性集合による設計<sup>[2]</sup>や質問処理効率化のためのスキーマの変換<sup>[3]</sup>がある。

本稿では、(A)の方法の問題点について触れたあと、(B)の方法について質問処理のための設計法として対象集合を用いたものを提案し、意味制約(関数従属性、結合従属性)を保持するためにスキーマを変換する方法を示す。また、これら2つの方法の特徴を比較して最適なネットワークデータベースの設計法について検討する。

### 2. 基本的事項

ネットワークモデルは、同じデータ項目(本稿では属性と呼ぶ)からなるレコードの集合であるレコード型と、2つのレコード型間のレコードの1対多の対応を表す親子集合型の集合によって表現される。この構造はバックマン線図と呼ばれる有向グラフ $B(V, A)$ で表される。ここで、 $V$ は各レコード型に対応する節点集合、 $A$ は親レコード型から子レコード型へ向かう有向枝の集合である。親レコード型が $R_0$ 、子レコード型が $R_m$ である親子集合型を $\langle R_0, R_m \rangle$ で表す。値が定めればレコード型 $R$ のレコードがただ1つ定まるような最小の属性集合 $K$ を $R$ のキーという。

本稿で対象とする質問は、属性集合 $X$ の値を指定し、それに対応する属性集合 $Y$ の値を求めるとい

うもので、 $Q(X, Y)$ と記す。 $X, Y$ はそれぞれ関係代数における選択、射影演算に対応しており、ネットワークモデルでは巡航操作により自然結合演算に対応する処理が行われるので、このクラスの質問は非常に一般的な質問である。バックマン線図の部分グラフで表される質問 $Q$ の処理に必要なネットワーク構造を $Q$ の質問スキーマ $B_Q$ と呼ぶ。

### 3. 質問処理効率化のための変換<sup>[3]</sup>

ネットワーク構造を与えられた質問 $Q(X, Y)$ を効率良く処理できるように変換する。質問処理の効率の良さとして、質問スキーマ $B_Q$ が満たすべき条件を次のようにする。

(1)  $X$ と共通属性を持たないレコード型を巡航せずに $X$ の選択条件の検査ができる。

(2)  $X, Y$ と共通属性を持たないレコード型を含まない。

(3) 得られた結果に重複がない。

文献<sup>[3]</sup>の方法では、指定された条件を満たすような質問スキーマが存在するようにネットワーク構造を変換できる。この方法を用いると(A)の型の設計を行えるが、次のような問題もある。

1. 適用できる規則が最適な構造に対して非決定的である。

2. 個々の質問に対して変換を行うので、複数の質問に対して変換を行うと結果が複雑になる可能性がある。

### 4. 質問処理のための構造

質問集合を効率良く処理できるネットワーク構造を考える。意味的につながりのある属性集合を対象と呼ぶ。質問 $Q(X, Y)$ では、属性集合 $XY$ の対応を求める必要があるため、 $XY$ を1つの対象と見ることが出来る。また、 $X$ の選択条件は巡航操作で早く検査される方がよいので、 $X$ も1つの対象とする。

対象集合 $S$ に対するバックマン線図を次のようにして作る。

i) すべての $X_1, X_2 (\in S)$ に対し、 $X_1 \cap X_2$ も $S$ の要素とする。

Comparison of Network Database Design Methods  
by Preservation of Dependencies and Efficient Query Processing

Tetsuya FURUKAWA and Yahiko KAMBAYASHI

Kyushu University

ii) Sの各要素Xをレコード型R(X)とする。

iii) レコード型 $R_1(X_1), R_2(X_2)$  ( $X_1 \subset X_2$ ) に対し、

$X_1 \subset X_3 \subset X_2$ となるレコード型 $R_3(X_3)$ が存在しなければ親子集合型 $\langle R_1, R_2 \rangle$ を作る。

一般に $R_i(X_i)$ のキーは $X_i$ となる。i)はSが共通集合演算で閉じるようにするためのもので、バックマン線図で同じ属性を含むレコード型をそれのみで連結とし、一貫性を保つために必要である。このとき親子集合型には質問スキーマに含まれるものと一貫性を保持するためのみのものがあり、これを区別すると処理効率を変えずに冗長性を下げることができる。

iv) 各レコード型R(X)に対し、質問スキーマでの親レコード型が唯一( $R'(X')$ とする)のとき、Rの属性集合Xを $X-X'$ で置き換える。

さらに索引構造の付加が行える。

質問集合 $Q_1(X_1, Y_1), Q_2(X_2, Y_2), \dots, Q_n(X_n, Y_n)$ に対しては、 $S = \{X | X = X_i \text{ 又は } X = X_i Y_i\}$ として設計を行えばよい。

この方法により設計された構造では、各質問 $Q(X, Y)$ に対し次の条件を満足する質問スキーマが存在する。

レコード型の並び $R_1, R_2, \dots, R_n$ で、親子集合型 $\langle R_i, R_{i+1} \rangle$  ( $1 \leq i < n$ )が存在し、あるkに対して $R_k$ の属性集合はXであり、 $R_n$ の属性集合はYである。

この構造は、3節での効率良い処理のための3つの条件を満足する鎖構造となる。巡航操作はすべて親レコード型から子レコード型に向かって行われ、最初のk個のレコード型で選択条件を満たすレコードの対応がただ1つ定まる。 $R_k$ のそのレコードつながる $R_{k+1}, \dots, R_n$ のレコードの対応をもとめればそれが解となる。

### 5. 従属性制約の保持のための変換

4節で得られたネットワーク構造に関数従属性: $X \rightarrow Y$ を反映させる。レコード型のキーに反映するには、レコード型 $R(XY)$ を付加し( $XY = X_i$ となるレコード型 $R_i(X_i)$ が存在すれば不要)、属性集合の包含関係により適宜親子集合型を付加する。親子集合型に反映するには、さらにレコード型 $R'(Y)$ を付加するが、Yの意味的つながりや等価な関数従属性集合と構造の対応を考慮しなければならない。

結合従属性の保持では、基本的には結合従属性の属性集合からなるレコード型の付加によるが、もとの構造が表す結合従属性との関係に注意を要する<sup>[4]</sup>。バックマン線図Bで表されるネットワーク構造が表す結合従属性を $JD(B)$ 、与えられた結合従属性

をJとする。 $JD(B) = *[X_1 Y_1, X_2 Y_2, \dots, X_n Y_n]$ (正確にはこれを既約化したもの)である。

1) Jの方が制約が強い部分では、もとのバックマン線図で表される対応は、付加された部分から得られる対応に一致しなければならない。

2)  $JD(B)$ の方が制約が強い部分では、その部分の巡航は情報損失結合に対応するが、4節での設計で考慮された質問ではそのような巡航はない。

3) 比較不能のとき、1)と2)の組み合わせであるが、Jを $JD(B)$ より弱くしたJ'の構造を付加した方が一貫性のためにはよい。

以上のように、関数従属性、結合従属性共にそのレコード型を追加するので、質問処理のための構造の設計時にそれらを含めることもできる。関数従属性: $X \rightarrow Y$ は $\{XY, Y\}$ を、結合従属性: $*[X_1, X_2, \dots, X_n]$ は $\{X_1, X_2, \dots, X_n\}$ を対象集合に加える。

### 6. 方法の比較

ネットワークデータベース設計のための2つの方法を比較する。(A)は質問に対し、処理効率のクラスを選ぶことができ、冗長性との関係を考慮することができる。(B)ではすべての質問は効率良く処理されるが、冗長性の問題、従属性制約の管理の問題がある。従って、ネットワークデータベースの設計は、次のように行くとよい。

1. 特に効率良く処理できるようにしたい質問、構造に反映させたい従属性制約を対象集合とする構造を設計する。
2. 代表的な質問に対し冗長性との兼ね合いを考えながらそれを効率よく処理できるように構造を変換する。

### 7. あとがき

質問処理のための設計として対象集合からネットワークデータベースを設計する方法を示し、従属性制約の反映を基にする方法と比較した。本稿で示した従属性制約保持のための変換は質問処理効率を変えないが、効率が悪くなくても冗長性を減らし従属性制約を反映する変換も必要と思われる。

### 参考文献

- [1] Kuck, S.M., Sagiv, Y., "Designing Globally Consistent Network Schemas", Proc. ACM SIGMOD, pp.185-195, May 1983.
- [2] Lien, Y.E., "On the Equivalence of Database Models", JACM, Vol.29, No.2, pp.333-362, April 1982.
- [3] 古川, 斎藤, 上林, "質問処理効率化のためのネットワークデータベースのスキーマ変換", 情報処理学会第32回全国大会, 3B-1, 1986年3月.
- [4] 古川, 上林, "ネットワークモデルにおける木質問", 信学会技報, AL84-62, 1985年1月.