関係データベースにおける意味制約を反映した非正規形の関係の設計問題[†]

上 林 弥 彦[††] 田 中 克 己[†††*]
武 田 浩 一[††**] 矢 島 脩 三[††]

関係データベースの設計理論における大きな前提の一つに第1正規形の条件がある．しかし，第1正規形の仮定を満たさない非正規形の関係は利用者インタフェースとしてきわめて有用であり，商用の関係データベースシステムにも非正規形の関係の支援機能をもったものがある．本論文では，従来考察されなかった，従属性等の意味制約を反映した非正規関係の設計問題について述べる．すなわち，(i) 正規形→非正規形の関係の変換のための基本操作，ROW-NEST, GROUP-BY, COLUMN-NEST, RELATION-NEST 操作を示し，(ii) 正規形の関係と意味制約に対し，どのような非正規関係が対応づけられるかを検討する．意味制約としては，関数従属性，結合従属性という従属性のクラスと，関連属性集合を考える．これらの意味制約は上記の基本操作により扱えることを示す．とくに実世界の自然な意味制約を表すことができると指摘されている一つの結合従属性と関数従属性の集合，および正規形の関係では直接扱えない関連属性集合を反映した設計法を考えているのが本論文の特色である．本論文の結果は，概念スキーマに種々の利点をもつ正規形の関係を用い，外部スキーマとして利用者にわかりやすい非正規形の関係を用いたシステムの設計等に有用であると考えている．

1. まえがき

関係データベース理論では，操作の単純化と数学的な定式化のために，非正規形を第1正規形に変換して扱うのが一般的である．このために生じる冗長度はさらに高次の正規形を求めるという方法で解決されている．高次の正規形も非正規形もデータのもつ冗長度を減少させるという意味では同じような目的をもっているが，従属性制約と前者との関連についてはよく知られるにもかかわらず，後者との関連については従来ほとんど研究がなされていない．本論文は，関数従属性や結合従属性を含む意味制約を冗長度の少ない非正規形の設計にどのように利用できるかを論じている．このような方法は，従属性と“相互に関連した属性集合”を同時に表示できるため，データベースの利用者インタフェースとしては，正規形よりもわかりやすくなることが多い[9)~11),13)]．

2. 諸定義

関係データベースに対する基本的な定義は，Ullman[16)] に従う．以下，本論文で用いる記号と追加的な定義を与える．

関係 R を構成する属性集合 $\{A_1, A_2, \cdots\}$ を $\underline{R}$ で表し，これを関係スキームと呼ぶ．関係 R の属性集合 X への射影を $R[X]$ で表し，関係 $R_1, \cdots, R_m$ の自然結合を $R_1 * \cdots * R_m$ あるいは $\mathop{*}_{i=1}^{m} R_i$ で表す．

属性集合 X より Y への関数従属性 (FD) を $X \to Y$ で表す．全関数従属性 (FFD) $X \Rightarrow Y$ が R で成立するというのは，$X \to Y$ が成立し，(1) 任意の $A \in Y$ に対し $X' \to A$ かつ $X' \subset X$ となる X' が存在しない，(2) $X = Y = A$ (同一属性である) のいずれかが成立することである．結合従属性 (JD) は，$*[Y_1, \cdots, Y_n]$ で表すものとする．$n=2$ の場合をとくに多値従属性 (MVD) と呼び，$Y_1 \cap Y_2 \to\to Y_1 - Y_2 | Y_2 - Y_1$ と表現することもある．R の射影上で成立する JD を局所結合従属性 (EJD) と呼ぶ．与えられた $\underline{R}$ 上の従属性集合 D に対し，D を満足する関係 R が常に満足する従属性集合を D の閉包といい，D^+ で記す．

JD の有用な図式的表現としてハイパグラフがある[3)]．JD j のハイパグラフとは，j に現れている各属性に一つの節点を対応させ，j の各成分となる属性集合 Y_i に対し，Y_i に含まれる属性の節点集合を輪状に囲んだ (ハイパ) 枝で表現したものである．

関係 R において $\underline{R} \supseteq X$ かつ $X \to \underline{R}$ を満足する属性集合 X を R の超キーといい，もし $X \Rightarrow \underline{R}$ を満足するなら X をキーという．

† Design Problems of Unnormalized Relations Utilizing Semantic Constraints in Relational Databases by YAHIKO KAMBAYASHI (Department of Information Science, Faculty of Engineering, Kyoto University), KATSUMI TANAKA (College of Liberal Arts, Kobe University), KOICHI TAKEDA and SHUZO YAJIMA (Department of Information Science, Faculty of Engineering, Kyoto University).

†† 京都大学工学部情報工学教室
††† 神戸大学教養部
* 現在 南カリフォルニア大学
** 現在 日本アイ・ビー・エム(株)

関連属性集合とは相互に関連し，一つの意味的な単位と考えることのできる属性の集合で，利用者の指定するものや，従属性より得られる(i)，(ii)がある(類似した概念として，“集約[15]”，“連係と対象[12]”，“対象[14]”等がある)．(i)各 FFD X⇒A に対し XA は関連属性集合である，(ii)JD＊$[S_1, \cdots, S_m]$ に対し S_i $(i=1, \cdots, m)$ は関連属性集合である．

本論文でいう従属性は，R が正規であることを前提としている．すなわち，$\underline{R}$ の各属性の定義域は単純値のみから成るというものである．これに対して，定義域値として集合値や関係そのものを許した場合に得られる関係を非正規であるという．

3. 正規形から非正規形への変換を行う基本操作

本節では正規関係から非正規関係への変換を行う四つの操作と，非正規関係の特徴づけについて述べる．

3.1 非正規関係の生成のための変換操作

非正規関係の生成に用いる基本操作として，ROW-NEST，GROUP-BY，COLUMN-NEST，RELATION-NEST の4操作を定義する．

(1) ROW-NEST $[X]$：関係 $R(\underline{R} \supseteq X)$ に対し，ROW-NEST $[X]$ は，各 $y \in R[\underline{R}-X]$ および $W_y = \{t \mid t \in R$ かつ $t[\underline{R}-X]=y\}$ なる組集合 W_y を，$t_y[X]=\{t[X] \mid t \in W_y\}$ かつ $t_y[\underline{R}-X]=y$ なる一つの組 t_y に写像する．

ここで各 W_y を R の ROW-NEST $[X]$ による部分関係，各 $t_y[X]$ を R の ROW-NEST $[X]$ による局所関係と呼ぶ (R, ROW-NEST $[X]$ は誤解のない限り省略する)．直観的には ROW-NEST$[X]$ は R の X 上の属性値を集合値に変換した非正規関係を生成する．ROW-NEST$[X]$ はR が $Y \cap X = \phi$ あるいは $Y \supset X$ なる Y 上の局所関係のみをもつ非正規関係の場合にも適用できる．後者の局所関係が存在する場合には $Y \supset X$ となる最も小さな Y (すなわちある $Y' \supset X$ なら $Y' \supset Y \supset X$ となる) 上の各局所関係に ROW-NEST $[X]$ を施す．これを ROW-NEST 操作の場という．

(2) GROUP-BY $[Y-X] \triangleq$ ROW-NEST $[X]$：ここで，R が正規関係あるいは $Y \supset X$ なる局所関係をもたない関係なら $Y=\underline{R}$ である．ただし Y 上の局所関係において Y-X 上の属性値は単純値であるとする．

ROW-NEST 操作と GROUP-BY 操作は相補的な操作であるが，次に述べる COLUMN-NEST 操作との関連や FD の反映のため，GROUP-BY 操作を区別したほうがよい．

この二つの操作の効果を row-tree によって表現することができる．正規関係 R_1 の row-tree T_R は＊で示される根と，$|\underline{R}_1|$ 個の端節のみから成り，各端節は $\underline{R}_1$ の各属性に対応する．R_1 に対し，GROUP-BY (あるいは ROW-NEST) 操作を適用すると，その結果得られる非正規関係 R_1' の row-tree T_R' は局所関係が生じる属性に対応した端節と，その親となる節点との間に＊ (あるいは Ⓡ) で示される内部節を加えたものとなり，GROUP-BY 操作の場合はさらに根を Ⓖ で置き換える (Ⓖ, Ⓡ はそれぞれ，GROUP-BY, ROW-NEST 操作の頭文字をとっている)．row-tree の根以外の内部節は，非正規関係の局所関係をもつ複合属性に相当する．

(3) COLUMN-NEST $[A_1, \cdots, A_n$ INTO $W]$：本操作は関係 R を表と考えたときに，属性 $A_1, \cdots, A_n$ をこの順に隣接して配置し，それらの属性を要素とする属性集合 W をその上に (一つの欄として) 示した表を生成する．

COLUMN-NEST 操作は表の列方向に階層を生成し，これは関連属性集合の表現となっている．この階層は，column-tree によって表す．正規関係 R_1 は，$\underline{R}_1$ で示す根と，$|\underline{R}_1|$ 個の端節をもつ column-tree Tc をもつ．column-tree は非正規関係の形式を定めるもので，順序木として扱う．属性集合の集合を COLUMN-NEST したり，属性と属性集合を COLUMN-NEST する場合にも本操作を自然に拡張して用いることにする．

(4) RELATION-NEST：関係 R がいくつかの同じ局所関係をもつときには，本操作が適用でき，その結果，その局所関係は独立した関係として分離され，名前がつけられる．もとの関係における局所関係はその名前と置き換えられる．

本操作の適用により，もとの関係と分離された関係間に階層ができるので，これを relation-tree で表現する．同じ関係名が2カ所以上に表れてもよい．

これらの各操作の適用例は5章の図4に示す．

本論文で用いる ROW-NEST と COLUMN-NEST の定義は，他の論文よりも一般的なものとなっている．文献10)では，両操作は独立したものとはなっていない．文献5)では一属性に対する ROW-NEST のみを扱っている．重要な部分クラスとして GROUP-BY

を独立させたのも本論文の特色である．

以後 GROUP-BY 操作等の系列の略記法として，GROUP-BY [X_1], …, GROUP-BY [X_n] を GROUP-BY [X_1; …; X_n] と記す．他の操作についても同様である．

3.2 基本操作間の相互作用

3.1 節で述べたように基本操作のあるものは他の操作の効果を無効にしない限り適用できないものがある．有効な基本操作の系列は，3種の tree 上の変換規則として扱うことができる．関係 S は row-tree T_R, column-tree T_C, relation-tree T_L をもつとする．

GROUP-BY 操作：S に対して GROUP-BY [X] が適用できるのは「ある節点 n が T_R に含まれ，n ラベルは＊または Ⓡ で，n の各部分木 T_i に対し ($i=1, \cdots, m$), T_i の端節に対応する属性の集合 X_i は $X_i \subseteq X$ あるいは $X_i \cap X = \phi$ を満足し，

$$\bigcup_{i=1}^{m} X_i \supset X \quad \text{となっているとき」} \qquad (1)$$

である．GROUP-BY 操作は X が単純値をもつときのみ適用できるので(1)で，$X_i \subseteq X$ となるとき X_i は単一属性である．n が T_R の根でないときはこの GROUP-BY 操作は n に対応する局所関係に作用する．本操作により得られる非正規関係の row-tree は，T_R における n の子として＊のラベルをもつ節点 n' を一つ追加し，$X_i \cap X = \phi$ となった n の部分木 T_i はすべて n' の部分木に改め，n のラベルを Ⓖ (n がラベル Ⓡ をもっていたときは GR) に変えたものとなる．

ROW-NEST 操作：ROW-NEST [X] は(1)と同じ条件が成立し，かつ T_C において X に含まれる属性がすべて隣接しているとき S に適用できる．本操作により得られる非正規関係の row-tree は，T_R における n の子として Ⓡ のラベルをもつ節点 n' を一つ追加し，$X_i \supseteq X$ となった部分木 T_i はすべて n' の部分木に改めたものである．

COLUMN-NEST 操作：COLUMN-NEST [Y_1, …, Y_K INTO W] が S に適用できるのは，(ⅰ)任意の Y_i, Y_j ($i \neq j$, $1 \leq i, j \leq k$) に対し $Y_i \cap Y_j = \phi$ となり，T_C において各 Y_i に対応する部分木をもつ内部節 n が存在する，および(ⅱ) T_C において $Y_1, \cdots, Y_n$ の順で隣接させる（各 Y_i が属性集合であれば，それはすでに順序づけられている）ときに，T_R の Ⓡ または GR をラベルにもつ各節点を根とする木の端節に対応する属性集合の隣接性が保たれる，が成り立つときである．本操作により得られる非正規関係の column-tree は T_C における n の子として W をラベルにもつ節点を加え，各 Y_i を W の部分木に改めたものである．

RELATION-NEST 操作については，R_1 が局所関係をもつときに適用でき，分離された関係については，R_1 に以後適用する操作をそのまま反映させる方法と，R_1 とは独立に操作を施す方法があり，後者は分離された関係においてとくに成立する FD 等を反映するときに有用である．

COLUMN-NEST 操作を実行するときには，たとえば MOVE といった属性の移動を行う操作を必要とするが，本論文では簡単のため省略している．ROW-NEST 操作により得られる局所関係は不可分であり，これは COLUMN-NEST 操作が適用できる条件(ⅱ)で反映されている．しかし GROUP-BY 操作と COLUMN-NEST 操作は図4に示すように互いに独立であるといえる．ここで属性値は行方向に対応しているという暗黙的な仮定を置いている．したがって，GROUP-BY 操作の適用に対しては部分関係の境界を横線で示し，ROW-NEST 操作に対しては局所関係を [ ] で表現する方法をとる．

4章では FD と関連属性集合，5章では JD を中心にこれら三つの意味制約が混在する場合を検討する．

4. 関数従属性を利用した非正規関係の生成

本章では与えられた関係とそこで成立する FD を用いて属性値の階層的な対応[1]を反映した非正規関係（階層表現）を得る方法について考察する．FD に対しては ROW-NEST 操作を用いないので，COLUMN-NEST 操作により任意に（列方向の）階層を反映できる．

4.1 FD 鎖と階層表現

[定理 1][8] GROUP-BY [X] を関係 R に適用して得られる各部分関係では FD $\phi \to X$ が成立する．

[定理 2][8] 関係 R で FD 鎖 $X_0 \to X_1 \to \cdots \to X_n$ が成立するとき，以下の GROUP-BY 操作の系列で FD 鎖を反映した非正規関係を得ることができる．GROUP-BY [X_n]; GROUP-BY [$X_{n-1} - X_n$]; …; GROUP-BY [$X_1 - X_2 - \cdots - X_n$]

関係 R で成立する FD 鎖は，たとえば，ABC→BC→C といった自明な FD や冗長な FD を含みう

るので，FFD をおもに用いる必要がある．また一つの関係でいくつかの FD 鎖が成立するときには，すべての FD 鎖を一つの非正規関係で反映するのは一般に不可能であり，できるだけ非冗長でもとの関係と等価な非正規関係集合を求めることが重要である．

データベース正規化の理論において，以下の等価性が知られている[2),5)]．関係 R と R 上で成立する FD 集合 F が与えられたとき，関係集合 $\{R_1, \cdots, R_n\}$ が R と等価であるとは，(i) 各 $R_i = R[\underline{R}_i]$ $(1 \leq i \leq n)$ で反映されている FD の和と F が等価であり，(ii) ある R_j $(1 \leq j \leq n)$ が $\underline{R}_j \supseteq K$ (K は R のキー) を満足することである．

したがって，R 上で成立する FD 鎖の集合と F が等価であれば (i) の条件は満足でき，もしも各非正規関係が K に相当する属性を含まなければ $\underline{R}_{n+1} = K$ となる関係 R_{n+1} を一つ追加して (ii) を満足させればよい．

［例 1］ R を $\underline{R}$=ABCDEF なる関係とし，FD 集合 {A→BCDE, BC→D} を満足するものとする．この FD 集合を図式的に表現したものを**図 1 (a)** に示す．ただし A→D のような冗長な FD は省略している．このとき最長の FD 鎖は A→BC→D で，まずこの FD 鎖に対応した非正規関係を生成する（図 1 (b)）．あとは FD A→E のみが反映されていないので，AE 上の非正規関係を生成する．R のキーは AF であるから，等価性を保存するために AF 上の関係を追加する．この関係は A, F のいずれに対して GROUP-BY 操作を施してもよい（図 1 (c)，(d)）．

上記の方法において，FD 鎖をどのように選定するかが非常に重要である．このため，次の問題について本節で検討する．

(a) FD 鎖を求めるための図式表現．

(b) FD 鎖における閉路の扱い．

(c) FD 鎖に反映する FF は FFD であることが望ましい．FFD は一般に推移律が成立しないという問題を解決しなければならない．

(d) 複数の FD 鎖を併合して扱う方法．

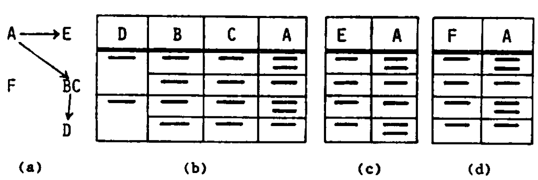


図 1 FD 鎖の図式的表現と FD 鎖を用いた非正規関係の生成

Fig. 1 Graphical representation of FDs and unnormalized relations generated from the FD chains.

4.2 FD 導出グラフ

本節では FD 導出グラフという図 1 (a) のような図式的表現を考察する．関係 R と R 上で成立する FD 集合 F に対する FD 導出グラフ G は，$X_1 \Rightarrow X_2$ が F に含まれる FFD であれば X_1 と X_2 をラベルとする節点および節点 X_1 から節点 X_2 への実線の有向枝を含む．G の任意の節点 X, Y に対し FFD ではない FD $X \rightarrow Y$ が成立するとき節点 X から節点 Y への破線の有向枝が G に含まれる．

手続き P1：節点数，枝数を減少させた FD 導出グラフを生成する．

(1) 与えられた FD 集合より非冗長 FD 集合を求める．

(2) 各 FFD の左辺となる属性集合に節点を対応させる．

(3) もし FFD $X_1 \Rightarrow X_2$ が成立し，$X_3 \subseteq X_2$ となるような節点 X_1, X_3 が存在するなら，節点 X_1 から節点 X_3 へ実線の有向枝を加える．

(4) ステップ (3) で $X_1 \Rightarrow X_2$ に対して，節点 X_1 から節点 $X_{21}, \cdots, X_{2m}$ への枝を加えたとする．このとき，もし $X_2 - X_{21} - \cdots - X_{2m} = X_2'$ が空でなければ，X_2' に対応する節点を追加し，節点 X_1 から節点 X_2' への実線の有向枝を加える．すなわち，$X_1 \Rightarrow X_2 - X_{21} - \cdots - X_{2m}$ を表現する節点と枝を追加したことになる．

(5) FD 集合に現れない $\underline{R}$ の属性集合をラベルとする節点を一つ加える．

(6) いままで得られた節点のうち，$X_1 \supseteq X_2$ となるような節点の対が存在すれば，節点 X_1 より節点 X_2 への破線の有向枝を加える．

4.3 FD 導出グラフにおける閉路

手続き P1 で得られた FD 導出グラフは閉路をもちうるので，次に示す手続きによって閉路を除去する．

手続き P2：FD 導出グラフにおける閉路の除去．

(1) FD 導出グラフにおいて $X_1 \Rightarrow X_2 \Rightarrow \cdots \Rightarrow X_n \Rightarrow X_1$ なる閉路が存在するものとする．

(2) $X = X_1 \cdots X_n$ に相当する節点を生成し，各 X_i $(i = 1, \cdots, n)$ に対応する節点を除去する．このとき，各 X_i と X_j $(i \neq j;\ j = 1, \cdots, n)$ をつないでいた枝はすべて除去し，それ以外の X_i を一端とする枝は，破線に変えて新たに節点 X につなぎ変える．

手続き P2 で閉路を構成する節点に対応した属性集合はすべて X 上の等価キーとなる．すなわち，任意の X_i, X_j に対し，その属性値に 1：1 対応が成立するため，ある X_i に対し GROUP-BY 操作を施すと，結果的には GROUP-BY [X] と同様の効果をもつことになる．手続き P2 は，次のようなキー破壊的な FFD にも応用できる．

[例 2] FFD AB⇒C, C⇒B が成立するものとする．対応する FD 導出グラフは図 2 (a) のようになる．この FFD 集合を取り扱う方法は以下のものがある．

(i) AB→C→B という FD 鎖に注目して，GROUP-BY [B] を適用すると，定理 1 より A→C が各部分関係で成立することがわかる．したがって，次に GROUP-BY [C] を適用し，図 2 (b) の結果を得る[8]．この種の表現方法は文献 4) にも見られる．

(2) GROUP-BY [A] を適用すると，各部分関で B→C，C→B が成立し，B, C が等価キーとなることがわかる．したがって，この場合はこれ以上の GROUP-BY 操作は不要である．結果は図 2 (c) に示す．

いずれの方法をとるかはデータの性質（属性 B の属性値の個数等）に依存するものと考えられる．

4.4 FFD の推移律と節点の併合

手続き P2 で示した等価キーの併合は，もう一つの問題を含んでいる．すなわち，FFD の推移律が成立しない場合の閉路の扱いであり，これは次の例に示すような扱いが可能である．

[例 3] 以下の FFD が成立するものとする．

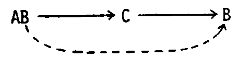
$$X_1 \Rightarrow X_2,\ X_2 \Rightarrow X_1,\ X_1 \Rightarrow X_3$$

いま，2 番目と 3 番目の FFD が推移律を満足しない（すなわち $X_2 \not\Rightarrow X_3$）と仮定すると，$X_2 = X_4 X_5$, $X_5 \neq \phi$ かつ $X_4 \Rightarrow X_3$ を満足する X_4, X_5 が存在することが示せる．したがって FD 導出グラフは図 3 (a) のようになる．GROUP-BY [X_5] の適用により得られる局所関係での FD 導出グラフとして図 3 (b) を得る．ここで，X_1, X_4 が等価キーとなるので GROUP-BY [X_3] を適用するとよい．

節点の併合に関して次の定理が成り立つ．

[定理 3] もし二つの FD 鎖 $U_1 \to U_2 \to \cdots \to U_n$ および $V_1 \to V_2 \to \cdots \to V_m$ が成立するなら，$W_1 \to W_2 \to \cdots \to W_p$ も FD 鎖となる．

ここで，$p = \min(m, n)$, $W_i = U_i V_i$ $(i = 1, \cdots, p)$ さらに $m < n$ なら $W_i = U_i$ $(m < i \leq n)$, $n < m$ なら



(a)

B	C	A
b_1	c_1	a_1
		a_2
	c_2	a_3
b_2	c_3	a_1

(b)

A	B	C
a_1	b_1	c_1
	b_2	c_3
a_2	b_1	c_1
a_3	b_1	c_2

(c)

図 2 キー破壊的な FD に対する非正規関係
Fig. 2 Unnormalized relations for key-breaking FDs.

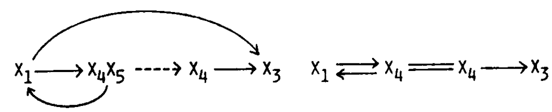


図 3 FFD が推移律を満たさないような FD 導出グラフの縮約
Fig. 3 Reduction of an FD-derivation graph that does not satisfy the transitivity of FFDs.

$W_i = V_i$ $(n < i < m)$ とする． （証明略）

一般に FD 鎖を併合するといくつかの FFD が反映されなくなるため，次のような場合に対して定理 3 を適用するのがよい．

(1) 定理 3 における二つの FD 鎖 $U_1 \to \cdots \to U_n$ と $V_1 \to \cdots \to V_m$ に共通して現れる属性が多い場合に，それぞれの FD 鎖を別々の非正規関係で反映すると結果の表現が大きなものとなる．したがって，小さな非正規関係として $W_1 \to \cdots \to W_p$ を反映したものを用いる．

(2) 関連属性集合が二つの FD 鎖に分割されているとき，それらの FD 鎖を併合して一つの非正規関係でその関連属性集合と併合された FD 鎖を反映する．

(3) 次章で述べるように，JD と FD がともに与えられた場合は先に JD の性質を利用して，その後に本章の方法を用いるほうが一般に効果的である．このときには JD の各成分に対応する局所関係で反映できる FD 鎖がたかだか一つとなるので FD 鎖の併合が有用となる．

たとえば，例 1（図 1）では三つの非正規関係を生成しているが，FD 鎖の併合と自明な FD の利用によ

り FD 鎖 AF→A→EBC→BC→D を得ることができる．したがって GROUP-BY [D; BC; E; A] を適用することで，一つの非正規関係のみを生成することができる．

GROUP-BY 操作と COLUMN-NEST 操作は独立に適用できるため，FD と関連属性集合の両方を反映した非正規関係の設計ができる．

5. 結合従属性を利用した非正規関係の生成

本章では JD に対する非正規関係の生成について述べる．JD j はそのハイパグラフ $H_j=(N, E)$ により図示する．N は節点集合，E は枝集合である．

［例 4］ $\underline{R}$=ABCDEFG 上の $*$ [ABC, ABDE, EFG] なる JD を満足する関係 R を考える．R に対する操作系列として，GROUP-BY [AB; E]; ROW-NEST [C; D; FG]; COLUMN-NEST [ACD INTO X; BE INTO Y] を施して得られる非正規関係をそれぞれ**図 4**（a）～（c）に示す．この結果 FG 上の局所関係がくり返し現れているので RELATION-NEST 操作を施す．

5.1 JD に対する GROUP-BY，ROW-NEST 操作

例 4 により JD に対する基本操作の効果を示した．一般に与えられた JD に対し，GROUP-BY 操作と ROW-NEST 操作系列を生成する方法の基本となる定理を示す．

［定理 4］ 関係 R が JD$*[S_1, \cdots, S_m]$ を満たし $\left(\underline{R}=\bigcup_{i=1}^{m} S_i\right)$，$X\subseteq\underline{R}$ のとき，R の GROUP-BY [X] による各部分関係ではやはり $*[S_1, \cdots, S_m]$ が成立する．

証） 各部分関係を $R_1, \cdots, R_p\left(R=\bigcup_{i=1}^{p} R_i\right)$ とすると，定理 1 より各 $R_i[X]$ は相異なる定数である．いまある R_j が $*[S_1, \cdots, S_m]$ を満足しないと仮定すると，ある組 $t\notin R_j$ があり，各 S_i に対し $t_i\in R_j$ かつ $t[S_i]=t_i[S_i]$ となる $(1\leq j\leq p, i=1, \cdots, m)$．$R$ はこの JD を満足するので $t\in R$ であるが，$t[X]=R_j[X]$ となるので t は R_j 以外の部分関係には含まれず矛盾する．（証明終わり）

［定理 5］ 関係 R が $\phi\rightarrow X$ を満足すると仮定する．このとき R が JD $*[S_1, \cdots, S_m]$ $\left(\underline{R}=\bigcup_{i=1}^{m} S_i\right)$ を満たすことと R が EJD$*[S_1-X, \cdots, S_m-X]$ を満たすことは同値である．

証） t と t' を以下の条件を満たす $\underline{R}$ 上の任意の組とする．（i）ある R 内の組 $t_1, \cdots, t_m$ に対し $t[S_i]=t_i[S_i]$ となる $(i=1, \cdots, m)$．（ii）ある R 内の組 $t_1', \cdots, t_m'$ に対し $t'[S_i-X]=t'_i[S_i-X]$ となる $(i=1, \cdots, m)$．JD $*[S_1, \cdots, S_m]$ が R 上で成立することは t が構成できないか，t が必ず R に含まれていることと同値であり，EJD$*[S_1-X, \cdots, S_m-X]$ が R 上で成立することとは t' が構成できないか $t'[\underline{R}-X]$ が $R[\underline{R}-X]$ に必ず含まれることと同値である．R 上で $\phi\rightarrow X$ が成立するので，$R[X]=x$ とすると，t が構成できることと t' が構成できることとは同値となる．（$t[X]=t'[X]=x$ とするとよい）（証明終わり）

［定理 6］ 関係 R が FD 集合 F と JD 集合 J を満足すると仮定する．このとき，R の GROUP-BY [X] による各部分関係では $F'=\{\phi\rightarrow X\}\cup F$ および $J'=\{*[S_1-X, \cdots, S_m-X]\mid *[S_1, \cdots, S_m]\in J\}$ なる FD 集合および EJD 集合をそれぞれ満足する．

（証明略）

［例 5］ $\underline{R}$=ABCDEFG なる関係 R が JD$*$[ABC, CDE, AEF, AG] を満たしているとする．この JD のハイパグラフを**図 5**（a）に示す．GROUP-BY [A] による R の各部分関係で成立する EJD は $*$[BC, CDE, EF, G]（図 5（b））となり，A の代わりに GROUP-BY [BC] とすれば $*$[DE, AEF, AG]（図 5（c））が成立する．前者は BCDEFG 上の局所関係が G，BCDEF 上の関係の直積として表せることを示し，後者はそのような性質はないことを示している．直積的な属性値の対応は ROW-NEST 操作を用いて小さな非正規関係を生成できるための重要な性質である．

与えられた JD j のハイパグラフを $H_j=(N, E)$ とする．いま，$X\subseteq N$ が H_j の極小の関節集合和（UAS と略す）であるとは（i）各 $A\in X$ に対しある枝 $e_1, e_2\in E$ が存在し $A\in e_1\cap e_2$ となり，（ii）H_j から X を除くと E の互いに素な枝集合（連結成分 CC という）の数がふえ，（iii）$X'\subset X$ で（i），（ii）を満足する X' がないものをいう．

手続き P3：与えられた JD j : $*[S_1, \cdots, S_m]$ に対し，GROUP-BY, ROW-NEST 操作系列を求めて実行する．

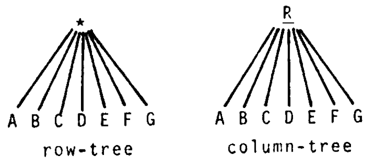
（1） j のハイパグラフ H_j を求める．i=1とする．

（2） H_j より UAS Z_i を求める．そのような Z_i がなければ（4）へ行く．

（3） H_j から Z_i を除いたものを新たに H_j とする．$i=i+1$とし，（2）へ行く．

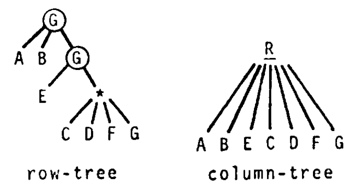
(a) R

A	B	C	D	E	F	G
1	1	1	1	1	1	1
1	1	2	1	1	1	1
1	1	1	1	1	2	2
1	1	2	1	1	2	2
1	1	1	2	2	1	1
1	1	2	2	2	1	1
1	1	1	2	2	2	2
1	1	2	2	2	2	2



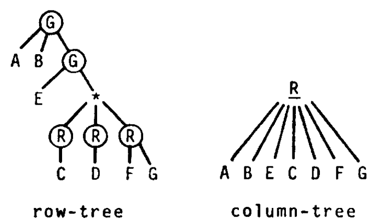
(b) R_1

A	B	E	C	D	F	G
1	1	1	1	1	1	1
			2	1	1	1
			1	1	2	2
			2	1	2	2
		2	1	2	1	1
			2	2	1	1
			1	2	2	2
			2	2	2	2



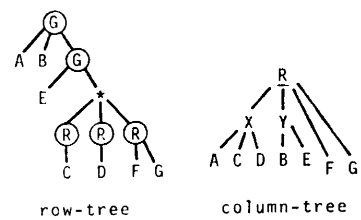
(c) R_2

A	B	E	C	D	F	G
1	1	1	[1 2]	[1]	[1 2]	[1 2]
		2	[1 2]	[2]	[1 2]	[1 2]



(d) R_3

X			Y		F	G
A	C	D	B	E		
1	[1 2]	[1]	1	1	[1 2]	[1 2]
	[1 2]	[2]		2	[1 2]	[1 2]



(e) R_4

X			Y			
A	C	D	B	E	F	G
1	R_5	[1]	1	1	R_6	
	R_5	[2]		2	R_6	

R_5

C
1
2

R_6

F	G
1	1
2	2

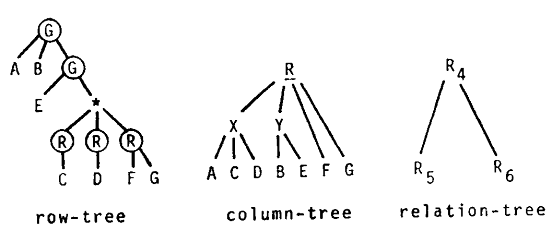


図 4 JD を用いた非正規関係の生成
Fig. 4 Unnormalized relations generated from a JD.

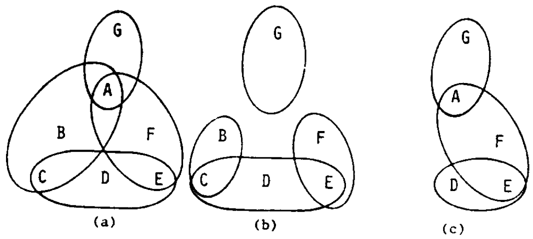


図 5 JD と GROUP-BY 操作の干渉
Fig. 5 Interaction of a JD and GROUP-BY operations.

(4) $X=Z_1\cup\cdots\cup Z_{i-1}$ として GROUP-BY [X] を施す．H_j の CC を $Y_1,\cdots,Y_k$ とすると ROW-NEST [Y_1; ⋯; Y_k] を施す．

(5) 同じ局所関係が現れていれば RELATION-NEST 操作を繰り返し適用する．

5.2 JD, FD, 関連属性集合の階層に対する非正規関係の生成

本節ではいままでの方法をまとめて，文献 14) で自然な制約のクラスと考えられているような，一つの JD と FD 集合から成る従属性集合にさらに関連属性集合の階層を反映した非正規関係を求める手続きを示す．

手続き P 4 : JD, FD, 関連属性集合の階層に対する非正規関係の生成.

(1) 与えられた関連属性集合の階層を column-tree に変換する．与えられた JD を j: $*[S_1,\cdots,S_m]$ とする．

(2) j に対し手続き P 3 を適用し，ステップ(4) で row-tree を求める．(1)の column-tree と row-tree が両立しないときは，一部の関連属性集合を反映しないようにするか，ある局所関係を生成しないようにする．

(3) (2)の row-tree と column-tree に対応する操作系列を適用し，さらに各局所関係で成立する FD 鎖があるときには 4 章の手続きにより GROUP-BY 操作を施す．

(2)の row-tree と column-tree の不整合に対しては，一般に row-tree を優先させて非正規関係のサイズを小さくすることが効果的であると考えられる．

また，手続き P3 のステップ(2)～(4)を，各 UAS Z_i が求まったときに GROUP-BY [Z_i] を施し，H_j の各 CC に対し ROW-NEST 操作を施し，各局所関係で成立する JD のハイパグラフがちょうどその CC と同じになることを利用して，今度は各 CC の UAS を求めるという方法にすることもできる．例 4 にこの方法を施せば GROUP-BY [AB]; ROW-NEST[C; DEFG]; GROUP-BY [E]; ROW-NEST [D; FG] を得る．

謝辞 日頃熱心にご討論いただく矢島研究室の諸氏に感謝する．なお本研究は一部文部省科学研究費による．

参 考 文 献

1) Armstrong, W. W. and Delobel, C.: Decompositions and Functional Dependencies in Relations, *ACM Trans. Database Syst.*, Vol. 5, No. 4, pp. 404-430 (1980).
2) Biskup, J., Dayal, U. and Bernstein, P. A.: Synthesizing Independent Database Schemas, *Proc. ACM SIGMOD Int. Conf. Management of Data*, pp. 143-151 (May-June 1978).
3) Fagin, R., Mendelzon, A. O. and Ullman, J. D.: A Simplified Universal Relation Assumption and Its Properties, *IBM Res. Rep.*, RJ 2900 (1980).
4) Furtado, A. L.: Horizontal Decomposition to Improve a Non-BCNF Scheme, *ACM SIGMOD Rec.*, Vol. 12, No. 1, pp. 26-32 (1981).
5) Jaeschke, G. and Schek, H.-J.: Remarks on the Algebra of Non First Normal Form Relations, *Proc. ACM SIGACT-SIGMOD Symp. Principles of Database Systems*, pp. 124-138 (March 1982).
6) Kambayashi, Y.: A New Synthetic Approach for Relational Database Design, Kyoto Univ. Dept. of Information Science,Yajima Lab. Res. Rep. (presented at AFIPS 1979 National Computer Conference, not contained in the Proceedings), ER 78-02 (Nov. 1978).
7) Kambayashi, Y., Tanaka, K. and Yajima, S.: Semantic Aspects of Data Dependencies and Their Application to Relational Database Design, *Proc. IEEE 3rd Int. Comp. Softw. Appl. Conf.*, pp. 398-403 (Nov. 1979).
8) Kambayashi, Y., Tanaka, K., Takeda, K. and Yajima, S.: Representation of Relations for Database Output Utilizing Data Dependencies, *Proc. 15th Hawaii Int. Conf. Syst. Sci.*, pp. 69-78 (Jan. 1982).
9) Kitagawa, H., Kunii, T. L. and Ishii, Y.: Design and Implementation of a Form Management System APAD Using ADABAS/INQ DBMS, *IEEE 5th Int. Comp. Softw. Appl. Conf.*, pp. 324-334 (Nov. 1981).
10) Kitagawa, H. and Kunii, T. L.: Form Transformer—Formal Aspects of Table Nests Manipulation, *Proc. 15th Hawaii Int. Conf. Syst. Sci.*, pp. 132-141 (Jan. 1982).

11) Luo, D. and Yao, S. B.: Form Operation by Example—A Language for Office Information Processing, *Proc. ACM SIGMOD Int. Conf. Management of Data*, pp. 212-223 (Apr. 1981).
12) Maier, D. and Warren, D. S.: Specifying Connections for a Universal Relation Scheme Database, *Proc. ACM SIGMOD Int. Conf. Management of Data*, pp. 1-7 (June 1982).
13) Makinouchi, A.: A Consideration on Normalized Form of Not-Necessarily-Normalized Relation in the Relational Data Model, *Proc. 3rd Int. Conf. on Very Large Data Bases*, pp. 447-453 (Oct. 1977).
14) Sciore, E.: The Universal Instance and Database Design, *Princeton Univ. Tech. Rep.*, No. 271 (June 1980).
15) Smith, J. M. and Smith, D. C. P.: Database Abstractions: Aggregation and Generalization, *ACM Trans. Database Syst.*, Vol. 2, No. 2, pp. 105-133 (1977).
16) Ullman, J. D.: *Principles of Database Systems*, Computer Science Press, Potomac, Md. (1980).

(昭和 58 年 1 月 26 日受付)
(昭和 58 年 6 月 20 日採録)