数理解析研究所講究録
第 494 巻 1983 年 1-12

横切り分解によるデータ従属性の保持

京都大学工学部　上林弥彦　Yahiko Kambayashi
神戸大学教養部　田中克己　Katsumi Tanaka

1．まえがき

データベース設計において残された一つの重要な問題は，従属性集合の持つ閉路の扱いである。この問題に対しては従来よく用いられている射影分解では扱うことが困難であるため別の方法を検討する必要がある。関数従属性に対してはSciore が属性追加による方法を発表している。[1] 本稿ではデータ表現のＦＯＲＭの問題とも関係する[2]関係の横切りによる従属性集合の持つ閉路の取り扱いについて考察する。従属性の持つ閉路は質問処理の問題とも大きく関係しており重要である。[3]

2．基本的定義

Ａ，Ｂ，…は一つの属性を表わし，…，Ｘ，Ｙ，Ｚは属性集合を表わすものとする。関係 R_i は属性集合 $\underline{R_i}$ に対して定義された組の有限集合である。$\underline{R_i}$ を R_i に対する関係スキーマと呼ぶ。簡単のため，和集合を連接で示すものとする。たとえば，$AB = \{A, B\}$，$XY = X \cup Y$ である。

組tのうち属性集合Xに対応する部分のみをとる射影を，t[X]で表わす。関係の射影，選択，自然結合は次のように定義される。

射影：　$R[X]=\{t[X] \mid t\in R\}$

選択：　$R[X=c]=\{t \mid t[X]=c,\ t\in R\}$

自然結合：　$R*S=\{tu[\underline{R}\cup\underline{S}] \mid t[\underline{R}\cap\underline{S}]=u[\underline{R}\cap\underline{S}],\ t\in R,\ u\in S\}$

関数従属性（FD）：$X\to Y$が関係Rで成立するといわれるのは，すべての組t，sの組合せにおいて，$t[X]=s[X]$ならば$t[Y]=s[Y]$となる場合である。FDのうちYが一属性のものを基本FDという。全関数従属性（FFD）は，$X\Rightarrow Y$と表わされ，$\forall A\in Y$について$X\to A$かつXが最小集合となるものである。

結合従属性（JD）：$*[S_1, S_2, \ldots, S_m]$が関係Rで成立するとは，Rが常に次の条件を満足する場合である。

$$R=R[S_1]*R[S_2]*\ldots*R[S_m]$$

上記のJDにおいて，$S_i \supset S_j$となるようなi, jが存在しないとき，このJDを既約であるという。またJDで$m=2$の場合を特に多値従属性（MVD）と呼ぶこともある。

従属性の集合をDで表わし，$D=F\cup J$とする。ここで，FはFD集合，JはJD集合である。D，F，Jの要素を，

それぞれ，d，f，j で表わす。従属性集合Dより導出可能なすべての従属性をDの閉包と呼び D^* で表わす。$d \in D^*$ の場合，d はDより導出可能であるといい，$D \models d$ で表わす。

3．関係の横切りに対する基本的考察

関係の分解には大きく分けて，射影分解と横切りとがある。射影分解は，Rを $R = R_1 * R_2 * \ldots * R_m$ を満足する m 個の関係に分けるものである。Rにおいて，FD：$X \to Y$ が成立すると，次のような射影分解ができる。

$$R = R[XY] * R[\underline{R} - Y] \qquad \text{(i)}$$

JD：$[S_1, S_2, \ldots, S_m]$ が成立すると次のような射影分解が可能である。

$$R = R[S_1] * R[S_2] * \ldots * R[S_m] \qquad \text{(ii)}$$

1つのFDの管理は(i)の場合 $R[XY]$ で管理すればよく，(ii)の場合は各関係が結合によって失なわれる組を持たないように管理するとよい。しかし，FDやJDが1個以上ある場合には，射影分解で管理するのがむつかしいことがある。このため，横切りとの併用について考察する。関係の横切りは次のように定義される。

$$R = R_1 \cup R_2 \cup \ldots \cup R_m$$

ここで，すべての i と j の組合せについて $R_i \cap R_j = \phi$ が成立すれば，基本横切りという。次のような性質が成立する。

P1：　関係RがFD集合Fを満足するとき，横切りによって得られた関係もFを満足する。もとのFFDがFFDではないFDとなったり，新しいFDが加わることもある。

P2：　関係RがJD集合Jを満足するとき，横切りによって得られた関係ではJが成立するとは限らない。

基本横切りのうち次の条件を満足するものを，GROUP-BY [X]と呼ぶ。

$\forall i \quad R_i[X] = $ 定数

P3：　関係RにGROUP-BY [X]を適用すると各 R_i はFD：$\phi \rightarrow X$ も満足する。FFD：$Y \rightarrow Z$ は $(Y-X) \rightarrow Z$ となる。

P4：　JD：$*[S_1, S_2, \ldots, S_m]$ を満足する関係Rに，GROUP-BY [X] を適用すると，各 R_i は JD′：$*[S_1-X, S_2-X, \ldots, S_m-X]$ を満足する。

P5：　次に示す D_1 と D_2 は，$D_1^* = D_2^*$ が成立する。

$$D_1 = \{ *[S_1, S_2, \ldots, S_m],\ \phi \rightarrow X \}$$

$$D_2 = \{ *[S_1-X, S_2-X, \ldots, S_m-X],\ \phi \rightarrow X \}$$

4、横切りによる関数従属性の保持

FD：$Y \rightarrow A$ をGROUP-BY [X]を適用した後に保持する方法を考える。各部分関係を R_i とし，$Z = X \cap Y$ とする。

(1) $A \in X$ の場合　　$R_i[ZA] = R_j[ZA]$　となる関係をすべてまとめてS_kとする。

$S_k[A] \neq S_h[A]$, $S_k[Z] = S_h[Z]$　となるS_kとS_hは次の条件を満足する。

$$S_k[Y-X] \cap S_h[Y-X] = \phi$$

(2) $A \notin X$ の場合　　この場合はさらに次の2つの場合に分けられる。

(2－1) $X \subseteq Y$ の場合　　この場合各部分関係でFD：$(Y-X) \rightarrow A$を管理すればよい。

(2－2) $X \not\subseteq Y$　　$Y \rightarrow A$は各部分関係における管理のみでは不充分で，$R_i[Z]$の等しいすべての部分関係上で成立するかどうかの管理が必要である。

（証明） (1) まず次を証明する。「関係RでY→Aが成立することと，RをGROUP－BY［A］で横切り分解してできた$T_1, T_2, \ldots, T_m$において，すべてのiとj ($i \neq j$)の組合せに対して，$T_i[Y] \cap T_j[Y] = \phi$となることとは等価である。」$T_i[Y] \cap T_j[Y] \neq \phi$の場合，共通に含まれるYの値を$y$とする。$T_i[YA]$は$ya_i$を含み，$T_j[YA]$は$ya_j$を含み，GROUP－BYの条件から$a_i \neq a_j$である。すなわち，$R[YA]$は$ya_i$, ya_j ($a_i \neq a_j$)という組を含むため$Y \rightarrow A$に反する。逆に条件が満足されればこの

ようなことはない。

上記の$S_1, S_2, \ldots$, はGROUP-BY[ZA]を行ってできた関係となる。$S_k[A]=a_k$, $S_h[A]=a_h$, $S_k[Z]=S_h[Z]=z$とすると, $S_k[ZA]=z a_k$, $S_h[ZA]=z a_h$となる。上記と同様にして，この場合はRで$Y \to A$が成立するのと，次の条件は等価となる。

$$S_k[Y-X] \cap S_h[Y-X]=\phi .$$

(2-1) $X \subseteq Y$かつ$A \notin X$の場合. 各部分関係において$R_i[X] \cap R_j[X]=\phi$が成立するため, $(Y-X) \to A$が各部分関係で成立することと，Rで$Y \to A$が成立することは等価になる。

(2-2) $X \nsubseteq Y$ $A \notin X$の場合. まず$Z=X \cap Y \neq \phi$の場合を考える。$R_i[Z]$の値の等しい部分関係をすべてまとめると等価的にGROUP-BY[Z]をしたことになり，$Z \subseteq Y$であるため(2-1)と同じ方法で管理できる。$Z=\phi$の場合は$X \cap YA=\phi$となり，GROUP-BYに使われたXはこのFDと全く共通属性を持たない。したがって，元のRにもどしてFDの管理を行なわなければならない。$Z=\phi$の場合，すべての$R_i[Z]$は等しいと定義すれば上記の特殊な場合に含めることができる。　　(証明終り)

与えられたFDに対して，管理が容易になるようにXを選

び，GROUP－BY［X］をほどこすことになる。管理の容易なのは次の場合である。

a．Y→Aに対し，X⊆YであるXでGROUP－BY［X］を行なう（上記の（2－1））。

b．Y→Aに対し，GROUP－BY［A］を適用する。（上記（1）の特殊な場合）。

P6：　すべてのFD：$Y_i \to A_i$ を部分関係におけるFD管理（上記a）のみで保持できるための必要十分条件は，$\cap Y_i \neq \phi$ であることである。この場合，$X = \cap Y_i$ によってGROUP－BY［X］を適用するとよい。

［例］R（R＝ABCD）で，2つのFD：AB→DとBC→Dが成立しているとする。射影分解によるFDの管理では，R［ABD］，R［BCD］，R［ABC］の3つの関係を作り，前の2つで2つのFDを管理することになる。3番目の関係はすべての関係を結合してRへもどすために必要となる。横切り分解では次のようになる。

まず，GROUP－BY［B］をほどこし，$R_1, R_2, \ldots, R_m$ を得る。P6により，各 R_i でA→DとC→Dを管理するとよい。これは，各 R_i にさらにGROUP－BY［D］をほどこすと上記bの場合となり，各 R_i を分けた部分関係でAとCの値集合が互いに素になるようにすればよい。

5. 関数従属性によるキー破壊の扱い

FDによるキー破壊のある場合は，射影分解のみではFDの保持のできないことが知られている。本節では，著者ら[2]とFurtado[4]によって独立に与えられた場合を含めた一般的なキー破壊の扱いについて考察する。

関係R（R＝AB C）において，次の2つのFD f_1 と f_2 が成立している場合を考える。

f_1：AB→C　　　　f_2：C→B

（方法1）　GROUP-BY [B]を適用する。

各R_iでは次の2つのFDが成立する。A→C　φ→B

f_1 はaの場合に対応するので各部分関係でA→Cを管理するだけで保持できる。f_2 はbの場合に対応しているので，各R_i において R_i[C] の値が互いに素という条件で管理できる。

（方法2）　GROUP-BY [A]を適用する。

各R_iでは次の2つのFDが成立する。φ→A　B↔C

f_1 はaの場合に対応するので各部分関係でB→Cを管理すればよい。f_2 はaでもbでもないため，Rに対する管理を必要とする。

（方法3）　GROUP-BY [C]を適用する。

各R_iではφ→BCのみが成立する。f_1 はbの場合に相当し

ているため，各部分関係でABの値が互いに素になるよう管理すればよい。f_2 は a に対応しているため，各部分関係で $\phi \to B$ が成立するように管理する。すなわち，各部分関係においてBの値は定数である。

（方法4）　2つの属性によるGROUP-BYの適用

GROUP-BY［BC］は結果として方法3と同じになる。GROUP-BY［AB］は f_1 によりGROUP-BY［ABC］と同等となり，関係を一組づつの部分関係に分ける操作となり無意味である。

GROUP-BY［AC］は，f_1 に対しては(1)で $Z \neq \phi$ の場合となり，f_2 に対しては(2-2)となるため，両方とも簡単には管理できない。

6、横切りによる結合従属性の保持

本節では，JD：$*[S_1, \ldots, S_m]$ とGROUP-BY[X]との関係について考察する。　$*[S_1, \ldots, S_m] \models *[S_i, Y_i]$ ($Y_i = \bigcup_{j \neq i} S_j$) であるため，まず $*[S_i, Y_i]$ の保持について考察する。

P7：　Rに対してGROUP-BY[X]を適用してできた関係を $R_1, R_2, \ldots, R_n$ とする。Rで $*[S_i, Y]$ が成立することと，各 R_i で $*[S_i, Y]$ が成立することが等価となるための必要十分条件は，$S_i \cap Y \supseteq X$ が成立することである。

（証明）Xが$S_i \cap Y$にない属性を含めば，同じ$S_i \cap Y$の値が複数のR_iに出現することになる。したがって，各R_iで$*[S_i, Y]$が成立しないことがある。$X = S_i \cap Y$であれば，各R_iで，$*[S_i - X, Y - X]$と$\phi \rightarrow X$が成立する。前者は$(S_i - X) \cap (Y - X) = \phi$であるため直積である。P5により各$R_i$で$*[S_i, Y]$が成立していることになる。$S_i \cap Y \supsetneq X$の場合は，各関係は上記の部分関係の和集合となるが，やはり$*[S_i, Y]$を満足する。GROUP－BYをしない場合に相当する$X = \phi$の場合も含まれている。（証明終り）

P8： Rに対してGROUP－BY[X]を適用してできた関係を$R_1, R_2, \ldots, R_n$とする。Rで$j: *[S_1, S_2, \ldots, S_m]$が成立することと，各$R_i$で同じ$j$が成立することとが等価になるための必要十分条件は，$\cap S_i \supseteq X$となることである。

（証明）P7が$Y = Y_1, Y_2, \ldots, Y_m$のすべての場合適用できる条件を求めると$\cap S_i \supseteq X$となる。$\cap S_i \supseteq X$の場合の等価性の証明はP7の場合とほぼ同じである。（証明終り）

P9： P8で，FD：$Y \rightarrow X$が成立する場合は，条件は$\cap S_i \supseteq Y$となる。

（証明）GROUP－BY[X]で得られた部分関係では，Yの値が互いに素である。したがって，各R_iでjが成立するのとRでjが成立するのとは等価になる。（証明終り）

P8は横切り分解によってJDが完全に保持できる条件を示しているが非常に制限が強い。P9を用いると巡回JDも扱うことができる。

[例] JD j：*[AB, AC, ADE, EF, EGH, HI]を考える。JDの各成分を枝に，各属性を節点に対応させたハイパーグラフ表現は図1に示されている。ここで，A→H，E→HというFDが成立すれば，

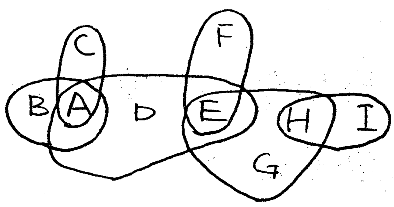

図1

GROUP-BY[H]によって，次のJDは部分関係における管理で保持できる。

*[AB, AC, ADEFGHI]　　A→Hによる。

*[EABCD, EF, EGHI]　　E→Hによる。

*[HABCDEFG, HI]　　GROUP-BY[H]による

jはこれらのJDより導びける。この場合，FDの保持のため，各部分関係におけるAやEの値集合が互いに素になるように管理しなければならない。

[例] JD j：*[ABC, CDE, AEF]に対応するハイパーグラフは図2に示されている。ここでA→E，C→EというFDが成立していれば，GROUP-BY[E]によってできた部分関係で同じJDを管理するとJDの保持ができる。

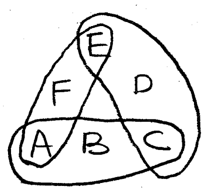

図2

7. むすび

本稿では，従属性の保持と横切り分解に関して基本的な考察を行った。単なるGROUP-BYでは能力が限られているため，より一般的な横切り分解を考えるのが今後の課題である。

謝辞　御指導，御検討いただいた京都大学工学部矢島脩三教授に深謝します。

文献

(1) E.Sciore,"Improving Database Schemes by Adding Attributes," ACM Conference on Principles of Database Systems, pp.379-383 (1983).

(2) Y.Kambayashi,K.Tanaka,K.Takeda and S.Yajima,"Representation of Relations for Database Output Utilizing Data Dependencies, Hawaii International Conf. System Sciences, pp.69-78, (1982).

(3) Y.Kambayashi and M.Yoshikawa,"Query Processing Utilizing Dependencies and Horizontal Decomposition," ACM SIGMOD International Conference on Management of Data, (1983).

(4) A.L.Furtado,"Horizontal Decomposition to Improve a non-BCNF Scheme," ACM SIGMOD Record 12,1, pp.26-32 (1981).