リレーショナル・データベースにおける多値従属の推移規則の問題点に関する考察

中村史朗（日立），Peter P. Chen（UCLA）

1. はじめに

Coddは[5,6]、リレーショナル・データベースの提案に当たり、関数従属（Functional Dependency）の概念を導入し、リレーション（データベース）の正規化を行なった。それ以来、データ従属ならびに正規化の理論は、リレーショナル・データベースの設計・解析において、中心的な役割を果たしている。データの操作における望ましくない性質を除くため、Coddは関数従属を基に、3種類の正規形（Normal Form）を定義した。

関数従属は、属性（Attribute）間の関係の一種である。XおよびYを属性の集合とする。この時、関数従属X→Yは、XとYの関係が1対1（one-to-one）ないし多対1（many-to-one）の場合、常に成立する。すなわち、関数従属は、属性間の1対1および多対1の関係を完全に表現できる。この完全性と単一値関数$y=f(x)$との対応性が、関数従属を理解しやすいものとしている。

しかしながら、SchmidとSwenson[13]によって示唆されているように、関数的でない属性間の関係も存在する。Fagin[8]とZaniolo[14]は、それぞれ独立に、多値従属（Multivalued Dependency）という、新しい型のデータ従属を提案した。（Delobel[7]も、同様の概念を関数従属という名前で発表している。）多値従属に基づき、Faginは第4正規形を定義した。その後多値従属に関する研究は活発化している[3,9〜12]。しかしながら、それらの内容は、より構文的（syntactic）となり、意味的（semantic）な考察は、十分には行なわれていない。本報告の目的は、多値従属を意味的な面から考察し、多値従属に対する推移律の持つ問題点を明らかにすることにある。

2. リレーショナル・データベースの基本概念

2.1 リレーション

属性（Attribute）とは、有限集合$U=\{A_1, A_2, \cdots A_n\}$からとられた記号である。各属性には、その属性がとり得る値の集合である領域（Domain）が対応する。属性Aに対応する領域をDOM（A）と記す。1つの領域は、複数の属性と関係づけられて良い。

属性の集合$U=\{A_1, A_2, \cdots A_n\}$上のリレーション（Relation）とは、直積$DOM(A_1) \times DOM(A_2) \times \cdots \times DOM(A_n)$の部分集合のことである。リレーションの要素（行）をタプル（Tuple）と呼ぶ。属性集合$\{A_1, A_2, \cdots A_n\}$上のリレーションRを、$R(A_1, A_2, \cdots A_n)$と記す。リレーションの定義情報をリレーション・スキーム（Relation Scheme）と呼ぶ。

データ従属（Data Dependency）は、リレーション・スキーム上で定義される。あるデータ従属が、あるリレーション・スキーム上で成立するということは、そのスキームの具現であるすべてのリレーションにおいて、当該データ従属が成立することを意味する。したがって、今後誤解が生じる恐れのない場合、"リレーション・スキーム"の代りに"リレーション"を用いる。

リレーションR（U）における1つのタプルを、uとする。YがUの部分集合の時、Yの要素に対応するuの値のみから成るタプルを、u[Y]と記す。RのY上での射影（Projection）R[Y]は、次のように定義される。（アルファベットの後半の大文字は、今後属性の集合を示すものとする。）　$R[Y] = \{u[Y] \mid u \in R\}$

属性の集合Xに対する特定値をxとする。上と同様、Y上のxによる条件付射影は、以下のように定義される。

$$R\{x, Y\} = \{u[Y] \mid u \in R \text{ かつ } u[X] = x\}$$

R(U)をリレーション，XをUの部分集合とする。Rにおけるすべての属性がXに関数従属である時，Xを<u>スーパーキー</u>(Super key)と呼ぶ。Xがスーパーキーであり、Xのいかなる真部分集合もスーパーキーではない時，XはRの<u>キー</u>(key)であるという。

<u>2.2 データ従属ならびに正規形</u>

データ従属は、下記2種類の情報を提供する。

・データの構造情報

・データの完全性(Integrity)に関する制約

(1) <u>関数従属</u>(Functional Dependency，以後FDと略す)

XならびにYをそれぞれ属性の集合とすると、FDは $f: X \to Y$ と表現される。リレーションR(X, Y, …)においてfが成立する時，同一のXの値を持つ2つのタプルはYの値も同一のものを持つ。通常簡略化のため、fを省いて $X \to Y$ と書く。

(2) <u>多値従属</u>(Multivalued Dependency，以後MVDと略す)

リレーションR(U)において、XおよびYをUの部分集合とする。また、$Z = U - XY$ とする(XYは、XとYの和集合を意味する)。XからYへのMVD：$X \twoheadrightarrow Y$ が成立する必要十分条件は、すべてのXZの値 xz に対し、下記が成立することである。

$$R\{xz, Y\} = R\{x, Y\}$$

$X \cap Y \neq \phi$ の時，$X \twoheadrightarrow Y$ が成立することは $X \twoheadrightarrow Y - X$ が成立することと同値である。これは定義から明らかである。非公式に言えば、$X \twoheadrightarrow Y$ が成立する時，YとZはXの下で独立である。

FDの定義より明らかなように、$X \to Y$ はXおよびYのみで定義され、他の属性とは独立である。一方，R(X, Y, Z)におけるMVD：$X \twoheadrightarrow Y$ の妥当性は、Zの存在に依存し、XとYのみからは決定できない。すなわち、$X \twoheadrightarrow Y$ が，R(X, Y, Z)においては成立せず，R(X, Y, Z')においては成立することがあり得る。ここに，$Z' \subset Z$ である。したがって，MVDは<u>文脈依存</u>(Context Sensitive)である。また、MVDは1対多および多対多の関係のすべてを表現することはできない。このことが、正しくMVDを推定することを困難にしている。これは、多くの属性が1つのリレーションに存在する時，顕著となる。

$X \supseteq Y$ の時，FD：$X \to Y$ は<u>trivial</u>であるという。R(X, Y, Z)におけるtrivial MVD：$X \twoheadrightarrow Y$ は、$X \supseteq Y$ ないし $Z = \phi$ の時定義される。Xのいかなる真部分集合X'に対しても，$X' \not\to Y$，$X' \not\twoheadrightarrow Y$ の時 FD：$X \to Y$ およびMVD：$X \twoheadrightarrow Y$ は、それぞれ<u>full</u>であるという。Nontrivial MVD：$X \twoheadrightarrow Y$ がFDでない時，それを<u>strong</u> MVDと呼ぶ。MVDの性質を論じる時はstrong MVDと仮定する。

FD，MVDに関する導出規則(Inference Rules)を下記する。

<u>FD導出規則</u>

FD1 (Reflexivity)：$X \supseteq Y$ ならば，$X \to Y$。

FD2 (Augmentation)：$W \supseteq Z$ かつ $X \to Y$ の時，$XW \to YZ$。

FD3 (Transitivity)：$X \to Y$ かつ $Y \to Z$ ならば，$X \to Z$。

<u>MVD導出規則</u>

MVD0 (Complimentation)：R(U)において $XYZ = U$ かつ $Y \cap Z \subseteq X$ の時，$X \twoheadrightarrow Y$ の必要十分条件は $X \twoheadrightarrow Z$ である。

MVD1 (Reflexivity)：$X \supseteq Y$ ならば，$X \twoheadrightarrow Y$。

MVD2 (Augmentation)：$W \supseteq Z$ かつ $X \twoheadrightarrow Y$ ならば、$XW \twoheadrightarrow YZ$。

MVD3 (Transitivity)：$X \twoheadrightarrow Y$ かつ $Y \twoheadrightarrow Z$ ならば，$X \twoheadrightarrow Z - Y$。

(以下は冗長な規則)

MVD4 (Pseudo-transitivity)：$X \twoheadrightarrow Y$ かつ $YW \twoheadrightarrow Z$ ならば、$XW \twoheadrightarrow Z - YW$。

MVD5 (Union)：$X \twoheadrightarrow Y$ かつ $X \twoheadrightarrow Z$ ならば、$X \twoheadrightarrow YZ$。

MVD6 (Decomposition)：$X \twoheadrightarrow Y$ かつ $X \twoheadrightarrow Z$ ならば、$X \twoheadrightarrow Y \cap Z$, $X \twoheadrightarrow Y-Z$, および $X \twoheadrightarrow Z-Y$。

FD－MVDの導出規則

FD－MVD1：　$X \rightarrow Y$ ならば、$X \twoheadrightarrow Y$。

FD－MVD2：　$Y \cap Z = \phi$ かつ $Z \supseteq Z'$ とした時、$X \twoheadrightarrow Z$ かつ $Y \rightarrow Z'$ ならば、$X \rightarrow Z'$。

(3) リレーションの正規形（Normal Forms）

(a) Boyce-Codd Normal Form（以後BCNFと略）－Nontrivial FD：$X \rightarrow Y$ が成立する時、常にリレーションRのすべての属性Aについて $X \rightarrow A$ が成立するならば、RはBCNFである[8]。言い換えると、XはRのスーパーキーである。BCNFは、従来の第3正規形より強い形である。

(b) 第4正規形（Fourth Normal Form，4NFと略）－Nontrivial MVD：$X \twoheadrightarrow Y$ が成立する時、常にリレーションRのすべての属性Aについて $X \rightarrow A$ が成立するならば、Rは4NFである[9]。したがって、もしRが4NFならば、RはまたBCNFでもある。

3．多値従属の問題点に関する考察

3.1　推移律の問題点の考察

MVD（多値従属）に対する導出規則は、形の上ではFD（関数従属）に対する規則と類似している。しかしながら、FDとMVDとの間には、重要な差が存在する。MVDは文脈依存であるのに対し、FDはそうではない。この文脈依存性が、MVDに対する推移律（Transitivity）－MVD3－において問題を生起する。すなわち、FDとMVDの推移律では、次のような差が存在する。

(1) リレーション $S(X, Y, ---)$ において $X \rightarrow Y$ が、$T(Y, Z, ---)$ において $Y \rightarrow Z$ が成立するとする。この時、リレーション $R(X, Y, Z, \cdots)$ においては、常に、$X \rightarrow Y$，$Y \rightarrow Z$ および推移的に $X \rightarrow Z$ が成立する。例えば、S（EMP，DEPT，AGE）においてEMP→DEPTが、T（DEPT，MANAGER，BUDGET）においてDEPT→MANAGERが成立するとする。この時R（EMP，DEPT，MANAGER）においては、上記FDならびに、推移的にEMP→MANAGERが成立する。

(2) しかし同様の議論をMVDに対して行なうことはできない。例えば，S(PROJ,EMP,PART)においてPROJ↠EMPが、T(EMP,CHILD，PROD)においてEMP↠CHILDが成立するとする。この時、R(PROJ，EMP，CHILD)においては、PROJ↠CHILDは成立しない。その原因は、推移条件PROJ↠EMP，EMP↠CHILDがRにおいて成立しないことによる。すなわち、RにおいてはPROJ↠EMPは成り立たない。上例は、ごく一般的なリレーションを用いている。したがって、この例はMVDの推移条件が成り立たない可能性を示唆している。

$R(X, Y, Z)$ において、推移条件 $X \twoheadrightarrow Y$，$Y \twoheadrightarrow Z$ が成立するとする。この時 $X \twoheadrightarrow Y$ は下記を意味する。

(1) Xの各値は、Yの値の集合を決定する。

(2) YとZとは、Xの下で独立である。

$Y \twoheadrightarrow Z$ は下記を与える。

(3) Yの各値は、Zの値の集合を決定する。

(4) XとZとは、Yの下で独立である。

上記(2)と(3)との間には、意味上の矛盾が存在する。したがって、下記の定理を得る。

注) 2章で述べたように、$S(X, Y, Z, W)$ の射影 $S[X, Y, Z]$ において、$X \twoheadrightarrow Y$ および $Y \twoheadrightarrow Z$ が成立しなければ、Sにおいても成立しない。したがって、以下では X, Y, Z のみを含む場合について考察する。

〔定理1〕 リレーション $R(X, Y, Z)$, $S(X, Y, V)$, $T(Y, Z, W)$ を、それぞれリレーション $U(X, Y, Z, V, W)$ の射影とする。ここに、X, Y, Z, V, W は互いに素である。下記条件が成立する時, Rにおいて $X \twoheadrightarrow Y$ は成り立たない。

(1) Sにおいて $X \twoheadrightarrow Y$, Tにおいて $Y \twoheadrightarrow Z$ と、それぞれ strong MVD が成立する。

(2) Uにおいて, YはXとは独立にZの値を定める。

(証明) 仮にRにおいて $X \twoheadrightarrow Y$ が成立するとして、矛盾を導く。あるXの値 x をとり、$S[x, Y] = Y_0$ とする。Y_0 の要素 y_1, y_2 を選び $T[y_1, Z] = Z_1$, $T[y_2, Z] = Z_2$ とおく。この時条件(2)より $Z_1 - Z_2 \neq \phi$ となるように y_1, y_2 を定めることが可能である。z_1 を $Z_1 - Z_2$ の要素, z_2 を Z_2 の要素とすると, (x, y_1, z_1), (x, y_2, z_2) はRのタプルとなる。しかるに、Rにおいて $X \twoheadrightarrow Y$ を仮定したので, (x, y_1, z_2), (x, y_2, z_1) もRのタプルでなければならない。したがって $R[y_2, Z]$ は、Z_2 の要素でない z_2 を含む。故に $R[Y, Z] \neq T[Y, Z]$。これは、RおよびTが共にUの射影であるという仮定に反する。したがって、Rにおいて $X \twoheadrightarrow Y$ は成立しない。(証明終り)

定理1の条件(2)は大変自然なものである。したがって今度は逆の観点から推移律を検討する。すなわち、次の定理2で推移率が成立する必要十分条件を明らかにする。

〔定理2〕 リレーション $R(X, Y, Z)$ を考える。ここに、X, Y, Z は互いに素である。Rにおいて、$X \twoheadrightarrow Y$ および $Y \twoheadrightarrow Z$ が成り立つ必要十分条件は、すべてのXYの値 xy において、下記の式(1)が成立することである。

$$R[x, Z] = R[x, y, Z] = R[y, Z] \qquad (1)$$

(証明) (1) 十分条件

MVDの定義より、$R[x, Z] = R[x, y, Z]$ を用いて $X \twoheadrightarrow Z$ を得る。規則MVD0 (Complementation)[注1] により、$X \twoheadrightarrow Y$ が得られる。また、$R[x, y, Z] = R[y, Z]$ より $Y \twoheadrightarrow Z$ を得る。

(2) 必要条件

Rにおいて、MVD0により $X \twoheadrightarrow Y$ は $X \twoheadrightarrow Z$ と等価なので、式(1)は $X \twoheadrightarrow Y$ および $Y \twoheadrightarrow Z$ より、直接導かれる。(証明終り)

注1) 以後2章の各導出規則の頭につけた記号(FD1, MVD3など)を頻繁に用いる。

定理2を用いて、下記系1および系2を得る。

〔系1〕 リレーション $R(X, Y, Z, W)$ を考える。ここに、X, Y, Z, W は互いに素である。Rにおいて $X \twoheadrightarrow Y$ および $Y \twoheadrightarrow Z$ が成立すれば、Rは下記条件を満足しなければならない。

(1) $R[x, Y]$ に属するすべてのYの値は、$R[x, Z]$ と同じZの値の集合を定める。

(2) $R[y, X]$ に属するすべてのXの値は、$R[y, Z]$ と同じZの値の集合を定める。

(証明) 条件(1)は、$R[y, Z] = R[x, Z]$ と等価である。ここに $y \in R[x, Y]$。条件(2)は、$R[x, Z] = R[y, Z]$ と等価である。ここに $x \in R[y, X]$。$X \twoheadrightarrow Y$, $Y \twoheadrightarrow Z$ は $R[X, Y, Z]$ においても成立するので、証明は定理2より明らかである。(証明終り)

次の系2は、上林[2)]らにより提案されたものである。

〔系2〕 リレーション $R(X, Y, Z, W)$ を考える。ここに、X, Y, Z, W は互いに素である。Rにおいて $X \twoheadrightarrow Y$ および $Y \twoheadrightarrow Z$ が成立すれば、Rは下記条件を満足しなければならない。

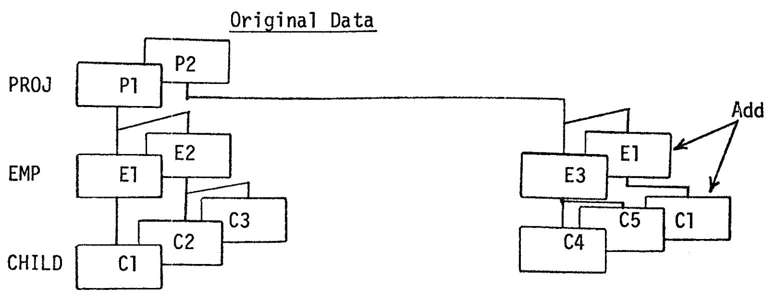
(1) $x \neq x'$ なる X の値 x, x' に対し、$R[x, Y] \cap R[x', Y] \neq \phi$ ならば、
$R[x, Z] = R[x', Z]$。
(2) $y \neq y'$ なる Y の値 y, y' に対し、$R[y, X] \cap R[y', X] \neq \phi$ ならば
$R[y, Z] = R[y', Z]$。

(証明) $X \twoheadrightarrow Y$ および $Y \twoheadrightarrow Z$ は $R[X, Y, Z]$ においても成立するので、証明は下記の通り定理2の式(1)を用いて直接行なうことができる。

(1) $R[x, Z] = R[x, y, Z] = R[y, Z] = R[x', y, Z] = R[x', Z]$。
ここに、$y \in R[x, Y] \cap R[x', Y]$。
(2) $R[y, Z] = R[x, y, Z] = R[x, Z] = R[x, y', Z] = R[y', Z]$。
ここに、$x \in R[y, X] \cap R[y', X]$ (証明終り)

系1および系2は、意味的に受入れることはできない。また実際の場でも発生するとは考えられない。したがって、定理1の条件(2)は実質的には何の制約も課さないと言える。言い換えると、推移条件($X \twoheadrightarrow Y$ および $Y \twoheadrightarrow Z$)は自然な意味では決して起こらない。更に悪いことは、推移律が誤った指定を許す原因になりかねない点である。$S(X, Y, V)$ において $X \twoheadrightarrow Y$ が、$T(Y, Z, W)$ において $Y \twoheadrightarrow Z$ が成立する時、データベース設計者は $R(X, Y, Z)$ において $X \twoheadrightarrow Y$ を指定する可能性がある。その例を図3.1に示す。前述のように、R(PROJ, EMP, CHILD)においては、EMP↠CHILDは成立するが、PROJ↠EMPは成立しない。しかるにPROJ↠EMPがS(PROJ, EMP, PART)で成立するため、誤ってRでPROJ↠EMPをも指定すると、望ましくない(不要な)タプルの存在を許すことになる。図3.1において、正しい元のデータが階層構造で示されている。図3-1-(a)は、タプル(P2, E1, C1)が挿入される前を、(b)は挿入された後のRを示している。(a)では、3個の不要なタプルが存在する。これは、系1-(1)ないし系

R(PROJ, EMP, CHILD) ; PROJ↠EMP, EMP↠CHILD
(incorrect)



(a) Before Addition of E1 & C1 (b) After Addition

(a)

PROJ	EMP	CHILD
P1	E1	C1
P1	E2	C2
P1	E2	C3
P1	E1	C2
P1	E1	C3
P1	E2	C1
P2	E3	C4
P2	E3	C5

Undesired tuples

More Undesired tuples

(b)

PROJ	EMP	CHILD
P1	E1	C1
P1	E2	C2
P1	E2	C3
P1	E1	C2
P1	E1	C3
P1	E2	C1
P1	E1	C4
P1	E1	C5
P1	E2	C4
P1	E2	C5
P2	E3	C4
P2	E3	C5
P2	E1	C1 ← Insert
P2	E1	C4
P2	E1	C5
P2	E3	C1
P2	E1	C2
P2	E1	C3
P2	E3	C2
P2	E3	C3

図3-1. 不正MVDによる不具合

2-(2)によるものである。更に(b)では、(P2, E1, C1)の追加により、11個の不要なタプルが新たに現われている。これは、系1-(2)ないし系2-(1)によるものである。

擬似推移律(Pseudo-transitivity)－MVD4－は、推移律の変形にすぎないので、本節の議論はすべてMVD4に対しても適用できる。

3.2 推移律を用いた新従属導出に関する考察

前節で推移律(MVD3)の問題点を考慮し、推移条件は通常成立しないことを明らかにした。一方我々は、推移律が特に他の導出規則の証明のためにしばしば用いられることを知っている。例えば、和の規則(MVD5)は次のように証明される。

R(X, Y, Z, W)において、$X \twoheadrightarrow Y$および$X \twoheadrightarrow Z$が成り立つとする。ここに、X, Y, Z, Wは互いに素である。この時、$X \twoheadrightarrow YZ$の証明は下記の通りである。

(1)	$X \twoheadrightarrow Y$	;	条件
(2)	$X \twoheadrightarrow Z$	;	条件
(3)	$X \twoheadrightarrow XY$	;	(1), MVD2
(4)	$XY \twoheadrightarrow YZ$	;	(2), MVD2
(5)	$XY \twoheadrightarrow W$	;	(4), MVD0
(6)	$X \twoheadrightarrow W$	;	(3), (5), <u>MVD3</u>
(7)	$X \twoheadrightarrow YZ$	;	(6), MVD0

しかしながら、上例は定理1とは明確に区別する必要がある。上例において、(3)と(5)の組合せは、真の推移条件とは見なせない。(3)において；XとXYとは素でないし、またえR には推移条件は存在しない。推移律(擬似推移律を含む)に合致しない2つのMVDが与えられた時、本節では下記事項を証明する。

(1) 増加律(Augmentation)により常に推移条件を得ることができる。

(2) (1)による推移律を適用して得られる結果は、他の導出規則を用いてすべて導出可能である。

まず、下記補題により(1)を証明する。

〔補題1〕 2つのMVDが与えられた時、MVD2を用いて推移条件を作り出すことが常に可能である。

(証明) リレーションR(X, Y, Z, V, W)において、$X \twoheadrightarrow Y$および$Z \twoheadrightarrow V$が成立すると仮定する。ここに、X, Y, Z, V, Wは必ずしも互いに素である必要はない。$X \twoheadrightarrow Y$をZで増加することにより、$XZ \twoheadrightarrow YZ$を得る。また$Z \twoheadrightarrow V$をYで増加することにより、$YZ \twoheadrightarrow YV$を得る。したがって、推移条件$XZ \twoheadrightarrow YZ$および$YZ \twoheadrightarrow YV$が得られた。(証明終り)

次の補題により、真の推移条件を満たさない条件を明らかにする。

〔補題2〕 リレーションR(X, Y, Z, V, W)が下記を満たすとする。

(1) $X \twoheadrightarrow Y$および$Z \twoheadrightarrow V$

(2) $X \cap Y = \phi$, $Z \cap V = \phi$および$W \cap XYZV = \phi$

(3) $X \not\supseteq Z$および$X \not\subseteq Z$ 注1)

この時、$X \twoheadrightarrow Y$および$Z \twoheadrightarrow V$ならびにそのいかなる射影もMVD3, MVD4に合致しないための必要十分条件は、$Y \cap XZVW = \phi$および$V \cap XYZW = \phi$である。

(証明) 十分条件は明らか。したがって、以下では必要条件を証明する。条件(2)より、$Y \cap ZV = \phi$および$V \cap XY = \phi$を証明すれば十分である。

(1) $Y \cap Z \neq \phi$と仮定する。$A \subseteq Y \cap Z$, $Y' = Y - A$, $Z' = Z - A$とおく。この時、R[X, A, Z', V, W]において$X \twoheadrightarrow A$が成立する。また$Z \twoheadrightarrow V$は$AZ' \twoheadrightarrow V$と同じである。故に、MVD4の推移条件$X \twoheadrightarrow A$および$AZ' \twoheadrightarrow V$を得る。これは矛盾

である。したがって，$Y \cap Z = \phi$でなければならない。

(2) (1)と同様の議論により、$V \cap X = \phi$を得る。

(3) $Y \cap V \neq \phi$と仮定する。$A \subseteq Y \cap V$とすると、$X \twoheadrightarrow A$および$Z \twoheadrightarrow A$が$R[X, A, Z]$において成立する。$R[X, A, Z]$において$X \twoheadrightarrow A$は$X \twoheadrightarrow Z$と等価であるので、これはMVD3の推移条件に合致する。故に$Y \cap V = \phi$でなければならない。

上記(1),(2),(3)および条件(2)より，$Y \cap XZVW = \phi$および$V \cap XYZW = \phi$を得る。(証明終り)

注1) $X \supseteq Z$ないし$X \subseteq Z$が成り立てば，YとVの条件のいかんにかかわらず、MVD3，MVD4には合致しない。

〔定理3〕補題2の条件を満たす2つのMVDが与えられている。この時，補題1により得られる推移律を適用して得られる結果は、すべてMVD0，MVD2およびMVD5により導出可能である。

(証明) 与えられた条件を満たすMVD: $X \twoheadrightarrow Y$および$Z \twoheadrightarrow V$が，$R(X, Y, Z, V, W)$において成立するとする。ここに、$Y \cap XZVW = \phi$および$V \cap XYZW = \phi$である。与えられたMVDから出発して、まず下記の推移条件を得る。

(1) $X \twoheadrightarrow Y$; 条件
(2) $Z \twoheadrightarrow V$; 条件
(3) $XZ \twoheadrightarrow YZ$; (1), MVD2
(4) $YZ \twoheadrightarrow YV$; (2), MVD2
(5) $XZ \twoheadrightarrow XV$; (2), MVD2
(6) $XV \twoheadrightarrow YV$; (1), MVD2

(3)と(4)は、推移条件を生成するための最小の増加(Augmentation)である。(5)と(6)も同じく最小の増加であり、これ以外には最小の増加は存在しない。(3)と(4)から下記の結果を得る。

*(7) $XZ \twoheadrightarrow V$; (3), (4), MVD3, $V \cap XYZ = \phi$
(8) $YZ \twoheadrightarrow XW$; (4), MVD0
(9) $XZ \twoheadrightarrow XW - YZ$; (3), (8), MVD3
*(10) $XZ \twoheadrightarrow W$; (9), $W \cap XYZ = \phi$, 定義

(5)と(6)より下記が得られる。

*(11) $XZ \twoheadrightarrow Y$; (5), (6), MVD3, $Y \cap XZV = \phi$
(12) $XV \twoheadrightarrow ZW$; (6), MVD0
(13) $XZ \twoheadrightarrow ZW - XV$; (5), (12), MVD3
(14) $XZ \twoheadrightarrow W$; (13), $W \cap XZV = \phi$, 定義

(8)−(9)−(10)の操作および(12)−(13)−(14)の操作は可逆である。また、(10)と(14)は同一であるので、結局*を付けた(7),(10)および(11)が推移律を用いないで導けることを示せばよい。証明は以下の通り。

(15) $XZ \twoheadrightarrow Y$; (1), MVD2
(16) $XZ \twoheadrightarrow V$; (2), MVD2
(17) $XZ \twoheadrightarrow YV$; (15), (16), MVD5
(18) $XZ \twoheadrightarrow W$; (17), MVD0, $YV \cap W = \phi$

(15)，(16)および(18)が結果を与える。

推移条件を得る残された唯一の可能性は、冗長な増加による方法である。QをRにおける任意の属性集合とする。この時，冗長性を含む増加により、下記の推移条件を得る。

(19) $XZQ \twoheadrightarrow YZQ$; (1), MVD2

(20)　$YZQ \twoheadrightarrow YVQ$　；　(2), MVD2

(21)　$XZQ \twoheadrightarrow XVQ$　；　(2), MVD2

(22)　$XVQ \twoheadrightarrow YVQ$　；　(1), MVD2

(21), (22)に関する議論も同様であるので、以下では(19), (20)のみを扱う。

*(23)　$XZQ \twoheadrightarrow V-Q$　；　(19), (20), MVD3, $V \cap XYZ = \emptyset$

(24)　$YZQ \twoheadrightarrow XW$　；　(20), MVD0

(25)　$XZQ \twoheadrightarrow XW-YZQ$　；　(19), (24), MVD3

*(26)　$XZQ \twoheadrightarrow W-Q$　；　(25), $W \cap XYZ = \emptyset$, 定義

(23)と(26)の導出は、(16)と(18)に対しMVD2を適用して直接得ることができる。

（証明終り）

3.1節の定理1および2により、真の推移条件は自然な意味では決して成立しないことを明らかにした。また本節において、真の推移条件を持たない2つのMVDに増加律を適用して得られる推移律の結果は、すべて推移律を用いないで導出可能であることを証明した。したがって、推移律(擬似推移律を含む)は実際上無意味なだけではなく、誤まった指定を生みだす原因となって有害ですらある。したがって、推移律は導出規則としてではなく、エラー・チェッカーとして用いるべきである。

3.3　多値従属と関数従属の混在規則の問題に関する考察

本節では、2章で述べた次の導出規則の問題点について考察する。

FD-MVD2：もし $X \twoheadrightarrow Z$ および $Y \rightarrow Z'$ ならば、$X \rightarrow Z'$ である。

ここに、$Z' \subseteq Z$, $Y \cap Z = \emptyset$。

これは、一見推移律と何の関係もなさそうに見える。しかし実は、推移律と同様の議論を行なうことができる。説明の便宜上、以下ではリレーション$R(X, Y, Z)$を考える。ここに、X, Y, Zは互いに素である。また、$Z' = Z$とする。(このような条件は、射影を用いて常に得ることができる。)この時、FD-MVD2は次のように書ける。

FD-MVD2'：$X \twoheadrightarrow Z$ かつ $Y \rightarrow Z$ ならば $X \rightarrow Z$ である。

Rにおいて、$X \twoheadrightarrow Z$ は $X \twoheadrightarrow Y$ と等価であるので、FD-MVD2'は下記のようになる。

FD-MVD2''：$X \twoheadrightarrow Y$ かつ $Y \rightarrow Z$ ならば $X \rightarrow Z$ である。

これにより、推移律とFD-MVD2''との間に極めて強い類似性があることがわかる。$X \twoheadrightarrow Y$ と $Y \rightarrow Z$ は下記を意味する。

(1)　YとZとは、Xの下で独立である。

(2)　各Yの値は、ただ1つのZの値を定める。

(1)と(2)の間に大きな矛盾があることは明らかである。推移律の場合と同様に、次の定理が成り立つ。

〔定理4〕　$R(X, Y, Z)$, $S(X, Y, V)$ および $T(Y, Z, W)$ を、それぞれ$U(X, Y, Z, V, W)$の射影とする。ここに、X, Y, Z, V, Wは互いに素である。下記条件が成立する時、Rにおいて $X \twoheadrightarrow Y$ は成立しない。

(1)　Sにおいて strong MVD：$X \twoheadrightarrow Y$，Tにおいて $Y \rightarrow Z$ が成立する。

(2)　Uにおいて，YはXと独立にZの値を定める。

(証明)　定理1の証明において，$Z_1 = \{z_1\}$, $Z_2 = \{z_2\}$ とおく。以下は定理1と全く同様に証明できる。　(証明終り)

(1)　MVD：$X \twoheadrightarrow Z$ を指定する場合，ZがX以外の属性と独立であることを確認しなければならない。しかしながら，FD：$Y \rightarrow Z$ があるので，ZはYと独立ではない。

(2)　別の観点から言えば，もし $X \rightarrow Z$ が本当に成り立つのであれば，$X \twoheadrightarrow Z$で

はなく $X \to Z$ が常に指定されねばならない。それは、次の理由による。

- FDの存在は、MVDの存在よりずっと容易に検知できる。
- FDの方がMVDより正しい情報を与える。

一旦 $X \to Z$ が成立しないと決定された場合，XとZとの関連は1対1でも多対1でもない。故に、$X \twoheadrightarrow Z$ と $Y \to Z$ から $X \to Z$ を導いてはならない。（もし $Y \to Z$ が成立するならば，$X \twoheadrightarrow Z$ が誤りである。）

(3) 例えば、R(DEPT, EMP, PHONE, ---)において、DEPT ↠ PHONE および EMP → PHONE（あるいはDEPT ↠ EMP およびEMP → PHONE）を指定したとすると、まちがったFD：DEPT → PHONE ―各部にはたった1つの電話しかない― が導かれる。（DEPT ↠ PHONE および DEPT ↠ EMP は誤りである。）多数の属性がある時，このような指定は起こりがちであるので、FD-MVD2は推移律以上に誤りやすい。

(4) 事実，実際の文献にも(3)と同様の誤りが現われている。それを次の例で示す。文献〔2〕において、リレーションR(LANDLORD, ADDR, OCCUPANT, ---)とLANDLORD ↠ ADDR および，OCCUPANT → ADDR を用いて、LANDLORD → ADDR を導出している。しかしながら、文献〔2〕では、同時に、実世界の分析に基づき、各家主は複数の建物を所有可能であると結論づけている。この分析が正しい限り、LANDLORD → ADDR は決して導出されてはならない。この不当な導出は、誤まったMVD：LANDLORD → ADDR に基因している。

4. おわりに

本報告では、多値従属の推移律の持つ問題点につき考察した。データ従属理論の研究は、最近盛んに行なわれているが、一部では、実際のデータベースから、遊離して記号の上だけの解析に偏っていく傾向が見られる。本報告がその反省への一助になれば幸いである。

最後に、海外留学による本研究の推進に当たり，種々御指導いただいたシステム開発研究所三浦所長始め、関係者の方々に厚く御礼申し上げます。

参考文献

1. Armstrong, W.W., "Dependency Structures of Database Relationships," Proc. IFIP '74, North Holland, 1974.

2. Beeri, C., Bernstein, P.A., and Goodman, N., "A sophisticate's Introduction to Database normalization Theory," Proc. 4th VLDB Conf., October 1978.

3. Beeri, C., Fagin, R. and Howard, J.H., "A complete Axiomatization for Functional and Multivalued Dependencies in Database Relations," Proc. SIGMOD Conf., August 1977.

4. Chen, P.P., "The Entity-Relationship Model - Toward a Unified View of Data," ACM Trans. on Database Systems, Vol.1, No.1 (March 1976).

5. Codd, E.F., "A Relational Model for Large Shared Data Banks," Comm. ACM, Vol.13, No.6 (June 1970).

6. Codd, E.F., "Further Normalization of the Database Relational Model," in Database Systems, Prentice-Hall, 1972.

7. Delobel, C., "Normalization and Hierarchical Dependencies in the Relational Data Model," ACM Trans. on Database Systems, Vol.3, No.3 (September 1978).

8. Fagin, R., "Multivalued Dependencies and a New Normal Form for Relational Databases," ACM Trans. on Database Systems, Vol.2, No.3 (September 1977).

9. Kambayashi, Y., Tanaka, K. and Yajima, S., "Semantic Aspects of Data Dependencies and their Application to Relational Database Design," Proc. COMPSAC '79, November 1979.

10. Lien, Y.E., "Multivalued Dependencies with Null Values in Relational Databases," Proc. VLDB Conf., October 1979.

11. Mendelzon, A.O., "On Axiomatizing Multivalued Dependencies in Relational Databases," J. ACM, Vol.26, No.1 (January 1979).

12. Parker, D.S. and Delobel, C., "Algorithmic Applications for a New Result on Multivalued Dependencies," Proc. VLDB Conf., October 1979.

13. Schmid, H.A. and Swenson, J.R., "On the semantics of the Relational Model," Proc. SIGMOD Conf., May 1975.

14. Zaniolo, C., "Analysis and Design of Relational Schemata for Database Systems," Ph.D. Dissertation, Computer Science Department, UCLA, Tech. Rep. UCLA-ENG-7669, July 1976.

15. Nakamura, F. and Chen, P.P., "Problem Areas of Multivalued Dependencies and a New Representation of Data Dependencies," (in preparation).