関係質問と従属性の干渉†

田中克己†† 上林弥彦†††*

ある種の意味制約を満足するデータを加工するとどのような意味制約を満足するかという問題は，データベースにおけるデータの加工を考えるうえで非常に重要である．本論文では，関係モデルのデータベースを対象として，関数従属性や結合従属性と呼ばれる意味制約と関係代数質問との間の干渉問題を扱っている．このため，結合従属性が質問によって保存されるか否かを検査する拡張追跡法という手法を導入し，また元の従属性が保存されるための諸条件やそのための属性名変更の方法等について検討している．

1. まえがき

現実世界のデータはある種の意味制約を満足しており，このようなデータを加工した場合に満足される意味制約を求めることは重要な問題である．本論文では関係モデルを対象としてこの問題を扱っている．

意味制約としては，関数従属性 (FD)[4] と結合従属性 (JD)[16] (多値従属性 (MVD)[5]を含む) を考え，データの加工操作として関係代数に基づく質問を考えている．主要結果は次のとおりである．

(1) FD と JD の成立している複数の関係を結合した場合に，与えられた局所的な JD が結果の関係で成立するかどうかを調べるための拡張追跡法の提案．

(2) ある関係で成立する JD が，結合後も成立するための，元の従属性集合に関する，諸条件．

(3) 制約演算と FD, JD との干渉問題に対する検討．

(4) 上記(2)で，ある関係で成立する JD が結果の関係でつねに成立するようにするための，属性名変更の方法に対する検討．

上記(1)については，著者らの結果[17)~19)]とは独立に，質問結果が与えられた従属性を満足するかどうかに関する証明法が，Klug[11),12)], Hull[8)], Casanova[3)], 伊藤ら[9)] によって示されている．したがって，本論文では，われわれの提案した拡張追跡法[18)]をまとめなおすとともに，(2)，(3)，(4)について，証明法とは異なる新しい成果を示している．

図1は，従属性と質問の干渉の簡単な例を示したものである．図1(a)は，MVD $A \rightarrow\rightarrow B|C$ を満足しているが，「r_1 の中で $B=C$ となる組をすべて取り出せ」という質問の結果(図1(b))では，これは満たされない．しかし質問によって，FD $B \rightarrow C$ と $C \rightarrow B$ が満たされるようになっている．図1(a)と図1(c)の関係を結合した結果(図1(d))でも，やはり上記の MVD は満たされていない．

このように，質問によって，消されてしまう従属性と新たに加わる従属性があり，この性質を検討するのが本論文の主目的である．この結果は次のような場合に利用できる．

(1) 質問の正しさの判定[15)]: FD については Klug[11)] が検討している．分散データベースについては文献 19)．

(2) 質問が正しくない場合の対処: 本論文では属性名変更による方法を示す．

(3) ビュー上で成立する従属性の計算．

(4) 質問処理の効率化への応用．

(5) 質問結果のもつ従属性を用いた，利用者にわかりやすい非正規表示[10)]．

(6) 普遍関係[7)]を仮定しないデータベース設計．

2. 基本的事項

属性集合Rに対し，組 t はRから属性の定義域への写像，関係 r は組 t の有限集合とする[20)]．$A, B, \cdots$ で属性を，$\cdots X, Y, Z$ で属性集合を表す．また，属性集合やその和集合は連接で表すものとする．関係代数の基本演算を次のように定義する ($A, B \in R$, c: 定数，関係 r_i は属性集合 R_i 上の関係)．

射影: $r[X]=\{t[X] \mid t \in r\}$ $(X \subseteq R)$

(ただし $t[X]$ は t の X 上への制限)

制約: $r[A='c']=\{t \mid t \in r, t[A]=c\}$

$r[A=B]=\{t \mid t \in r, t[A]=t[B]\}$

† Interference Problems between Relational Queries and Dependency Constraints by KATSUMI TANAKA (College of Liberal Arts, Kobe University; Department of Computer Science, University of Southern California) and YAHIKO KAMBAYASHI (Faculty of Engineering, Kyoto University).

†† 神戸大学教養部，南カリフォルニア大学計算機科学科
††† 京都大学工学部
* 現在 九州大学工学部

r_1

A	B	C
0	0	0
0	0	1
0	1	0
0	1	1

(a)

r_1'

A	B	C
0	0	0
0	1	1

(b)

r_2

B	C	D
0	0	1
1	1	0

(c)

$r_1 * r_2$

A	B	C	D
0	0	0	1
0	1	1	0

(d)

図 1 従属性と質問との非干渉
Fig. 1 Interaction between dependencies and queries.

結合：$r_1 * \cdots * r_n = \mathop{*}\limits_{i=1}^{n} r_i$

$= \{t | t$ は $\bigcup\limits_{i=1}^{n} R_i$ 上の組，各 i に対し $t[R_i] \in r_i\}$

R 上の関係 r 内の任意の組 t, t' に対し $t[X]=t'[X]$ なら $t[Y]=t'[Y]$ となるとき $(X, Y \subseteq Y)r$ は関数従属性 (FD) $X \rightarrow Y$ を満たすという．また，$r[\bigcup\limits_{j=1}^{m} S_j] = \mathop{*}\limits_{j=1}^{m} r[S_j]$ (各 $S_j \subseteq R$) のとき，r は局所結合従属性 (EJD) $*[S_1, \cdots, S_m]$ を満たすという．$m=2$ の EJD を局所多値従属性 (EMVD) と呼び $S_1 \cap S_2 \rightarrow\rightarrow (S_1 - S_2) | (S_2 - S_1)$ で表す．$R = \bigcup\limits_{j=1}^{m} S_j$ のとき，EJD や EMVD をとくに結合従属性 (JD)，多値従属性 (MVD) という．R 上の従属性集合 Σ (内のすべての従属性) を満たすいかなる関係も従属性 σ を満たすとき，Σ は σ を論理的に含意するといい，$\Sigma \vDash \sigma$ で表す．関係スキームは，$\underline{R}_i = \langle R_i, \Sigma_i \rangle$ (R_i：属性集合，Σ_i：R_i 上の従属性集合) で与えられる．Σ_i を満たす R_i 上の関係すべてからなる集合を SAT (R_i, Σ_i) と記す．

各 SAT (R_i, Σ_i) 内の任意の関係 r_i に対して

(1) $r_i[A='c'] \in$ SAT (R_i, σ) $(A \in R_i)$,

(2) $r_i[A=B] \in$ SAT (R_i, σ) $(A, B \in R_i)$,

(3) $r_i[X] \in$ SAT (X, σ) $(X \subseteq R_i)$,

(4) $\mathop{*}\limits_{i=1}^{n} r_i \in$ SAT $(\bigcup\limits_{i=1}^{n} R_i, \sigma)$

のおのおのが成り立つとき，それぞれ，σ が $\underline{R}_i[A='c']$, $\underline{R}_i[A=B]$, $\underline{R}_i[X]$, $\mathop{*}\limits_{i=1}^{n} \underline{R}_i$ 上で成立するという．

たとえば，$\underline{R}_1 = \langle R_1 = ABCD$, $\Sigma_1 = \{*[AB, ACD]\} \rangle$, $\underline{R}_2 = \langle R_2 = BCE$, $\Sigma_2 = \{BC \rightarrow E\} \rangle$ とする．**図 2** (a) の関係 r_1, r_2 はおのおの Σ_1, Σ_2 を満たす．容易に確かめられるように，$\underline{R}_2$ の FD: $BC \rightarrow E$ は $\underline{R}_1 * \underline{R}_2$ 上で成立する．たとえば，図 2 (b) の結合 $r_1 * r_2$ は $BC \rightarrow E$ を満たしている．しかし，$r = r_1 * r_2$ とすると，

$$r[AB] * r[ACD] \supsetneq r[ABCD]$$

r_1

A	B	C	D
1	0	1	0
1	0	0	1
1	1	1	0
1	1	0	1

r_2

B	C	E
0	1	0
0	0	0
1	1	0

(a)

$r_1 * r_2$

A	B	C	D	E
1	0	1	0	0
1	0	0	1	0
1	1	1	0	0

(b)

r_2'

B	C	E
0	1	0
0	0	0

(c)

図 2 結合上の FD と EJD
Fig. 2 FD and EJD on a join relation.

となるため，EJD: $*[AB, ACD]$ は $\underline{R}_1 * \underline{R}_2$ 上では成立しない．このことは，この EJD をどの $r_1 * r_2$ も満たさないことを意味するのではない．図 2 (c) の r_2' に対して，$r_1 * r_2'$ はこの EJD を満たしている．

関係代数の諸演算と従属性の増減との関係を直観的に要約すると次のようになる．

(1) 制約：$r[A='c']$ FD は増加，JD は不変．
$r[A=B]$ FD は増加，JD は減少．

(2) 射影：FD は減少，元のある EJD が JD となる (Beeri, Rissanen[2])．

(3) 結合：FD は不変，元の JD の一部が EJD となり成立，結合に対応する JD が追加．

(4) 集合演算：集合積 FD は不変，JD は減少．
集合和 FD, JD ともに減少．

(5) 属性名変更：これについては 6 章で考察する．

3. 結合演算と従属性との干渉

関係スキーム $\underline{R}_1, \cdots, \underline{R}_n$ において，各 Σ_i を FD および EJD の集合とすると，著者らは，σ (FD か EJD) が $\mathop{*}\limits_{i=1}^{n} \underline{R}_i$ で成立するならば，

$$\bigcup_{i=1}^{n} \Sigma_i \cup \{*[R_1, \cdots, R_n]\} \vDash \sigma$$

となることを示した[19]．これは結合で保存できる従属性は必ず上記の左辺から導けるものに限ることを示している．類似の結果はより強い仮定のもとで文献 2)，14) でも示された．上記の式の成立を Maier らの追跡法[13]を用いて調べることができるが，その際，Σ_i 内の JD は $R = \bigcup\limits_{i=1}^{n} R_i$ 上では一般に EJD となり，Maier らの追跡法では EJD に対して停止性が保証されないという問題がある．また，上記は σ が $\mathop{*}\limits_{i=1}^{n} \underline{R}_i$ で成立するた

めの必要条件であり，十分条件とはならない．これは，結合によりもともと成立していた MVD や JD が，$\overset{n}{\underset{i=1}{*}}\underline{R}_i$ 上では EMVD や EJD という形ですら成立しないことがあるためである．

しかし次の(A)〜(C)のいずれかの場合には，上記の式を十分条件としても利用できる．

(A) 各 Σ_i $(i=1, \cdots, n)$ はすべて FD のみからなる．

(B) 普遍関係仮定[7]が成立する．すなわち，$\{(r_1, \cdots, r_n) | r_i \in \mathrm{SAT}\ (R_i, \Sigma_i)$ かつある $\bigcup_{i=1}^{n} R_i$ 上の r に対して $r[R_i]=r_i$ $(i=1, \cdots, n)\}$ のみを対象とする．

この(A)は FD は結合でつねに保存されることから明らか．(B)は $\overset{n}{\underset{i=1}{*}} r_i$ がつねに Σ_i 内の従属性を満たすことを保証している．さらに，(A)を満たすように各 $\underline{R}_i$ を変形できるための必要十分条件として次の(C)がある．

(C) 各 $\underline{R}_i$ を次の(ⅰ)〜(ⅲ)を満たす関係スキーム集合 $\{\underline{R}_{i_1}, \cdots, \underline{R}_{i_k}\}$ $(k \geqq 1)$ に変形できる．(ⅰ)各 $\Sigma_{i_j} (j=1, \cdots, k)$ は FD 集合である．(ⅱ) $\Sigma_i \models \bigcup_{j=1}^{k} \Sigma_{i_j} \cup \{*[R_{i_1}, \cdots, R_{i_k}]\}$，(ⅲ) $\bigcup_{j=1}^{k} \Sigma_{i_j} \cup \{*[R_{i_1}, \cdots, R_{i_k}]\} \models \Sigma_i$．すなわち，この分解は Σ_i をそれと等価な一つの JD と FD 集合に置きかえることに相当する．

また，次の条件は σ が $\overset{n}{\underset{i=1}{*}}\underline{R}_i$ 上で成り立つための十分条件として利用できる（必要条件とはならない）．

(D) 各 $\underline{R}_i=\langle R_i, \Sigma_i \rangle$ の Σ_i から論理的に含意されるすべての FD からなる集合を F_i とすると，

$$\bigcup_{i=1}^{n} F_i \cup \{*[R_1, \cdots, R_n]\} \models \sigma.$$

さらに，σ が FD の場合には次の定理を得る．

［定理1］ 関係スキーム $\underline{R}_i=\langle R_i, \Sigma_i \rangle$ $(i=1, \cdots, n)$ に対し，Σ_i は FD, EJD の集合とする．$\bigcup_{i=1}^{n} R_i$ 上の FD f に対し，f が $\overset{n}{\underset{i=1}{*}}\underline{R}_i$ 上で成立すれば，$\bigcup_{i=1}^{n} \Sigma_i \models f$ である．

（証明） $R=\bigcup_{i=1}^{n} R_i$ とする．Fagin の結果[6]から，従属性集合 $\bigcup_{i=1}^{n} \Sigma_i$ に対する R 上の「アームストロング関係」r が存在する．ここでアームストロング関係 r とは，$\bigcup_{i=1}^{n} \Sigma_i \models \sigma$ なるいかなる σ (FD か EJD) も満たすが，それ以外の FD, EJD は満たさないような有限の関係である[6]．よって，r が満たすいかなる FD も $\bigcup_{i=1}^{n} \Sigma_i$ から論理的に含意される．r の各射影 $r[R_i]$ は Σ_i を満たすから，$r[R_i] \in \mathrm{SAT}\ (R_i, \Sigma_i)$ $(i=1, \cdots, n)$ である．f が $\overset{n}{\underset{i=1}{*}}\underline{R}_i$ 上で成立するから，$\overset{n}{\underset{i=1}{*}} r[R_i]$ は f を満たす．ここで $\bigcup_{i=1}^{n} \Sigma_i \models f$ でないとする．$\overset{n}{\underset{i=1}{*}} r[R_i] \supseteq r$ だから，r もまた f を満たす．しかし，r が満たす任意の FD は $\bigcup_{i=1}^{n} \Sigma_i$ から論理的に含意されるはずである．よって矛盾が得られた．

（証明終）

さて，上記の(A)〜(D)はいずれも制限が強すぎて一般性に乏しい．そこでここでは一般の場合に適用できる拡張追跡法[18]を導入する．Maier らの追跡法[13]はもともと一つの属性集合上での従属性の論理的含意関係を調べるものである．よって，複数の関係の結合による元の JD 等の破壊といった問題を扱うには不十分である．また上記の(B)の場合で，かつ各 Σ_i が FD, JD のみからなるときでも，先に述べたように Maier らの追跡法ではその停止性が保証されない．拡張追跡法は複数の関係間の操作を陽に扱うために，Maier の追跡法に射影，結合演算を取り入れて拡張したものである．

属性集合 $R=\{A_1, \cdots, A_k\}$ に対する要約表[13]は変数の k 個組の集合である．各変数は各属性に対応し，識別変数[13]と非識別変数とがあり，おのおの a_i, b_j で示す．FD: $X \to Y$ に対し，X の等しい要約表の二つの組の Y の部分を等しくさせる操作を FD 則[13]といい，Y の値が a と b のときは a で，b_i と b_j のときは添字の小さいほうで置き換える．また，JD: $*[Y_1, \cdots, Y_m]$ に対し，要約表の組 $w_1, \cdots, w_m$ から $w[Y_j]=w_j[Y_j]$ $(j=1, \cdots, m)$ なる組 w が構成でき，かつ w が要約表になければこれを追加する操作を JD 則[13]という．

［拡張追跡法］[18] （各 Σ_i は FD, JD 集合とする．）

(1) 与えられたEJD: $*[S_1, \cdots, S_m]$($\bigcup_{j=1}^{m} S_j=S$, $S \subseteq \bigcup_{i=1}^{n} R_i$) に対し，初期要約表 $T=\{w_1, \cdots, w_m\}$ を構成する．ここで T の各列は $\bigcup_{i=1}^{n} R_i$ の各属性に対応し，各組 w_j は S_j に対応する $(j=1, \cdots, m)$．S_j 内の属性 A_l に対し w_j の第 l 列を a_l とし，他の列には相異なる非識別変数を与える．

（2） $T_i=T[R_i]$ $(i=1,\cdots,n)$ とする.

（3） 各 T_i に Σ_i を用いて, FD 則, JD 則を適用する. また, ある T_i の非識別変数 b が FD 則で変更されれば他の T_j 内の同一の b を同様に変更する(要約表間の波及という). いかなる FD 則, JD 則の適用および要約表間の波及も行えなくなったものを $\mathrm{chase}_{\Sigma_i}(T_i)$ $(i=1,\cdots,n)$ で表す.

（4） 結合 $\mathop{*}\limits_{i=1}^{n}\mathrm{chase}_{\Sigma_i}(T_i)$ が S に対応する列すべてに識別変数をもつような組を有しているときかつそのときに限り, $\mathop{*}\limits_{i=1}^{n}\underline{R}_i$ 上で $*[S_1,\cdots,S_m]$ が成立するとする.

拡張追跡法の正しさ[18]は次のようにして示せる. もし $\mathop{*}\limits_{i=1}^{n}\mathrm{chase}_{\Sigma_i}(T_i)$ が（4）に述べた組をもたないなら, 明らかに $\mathrm{chase}_{\Sigma_i}(T_i)\in\mathrm{SAT}(R_i,\Sigma_i)$ だから, これが $*[S_1,\cdots,S_m]$ を満たさない例となる. また, ある $\mathop{*}\limits_{i=1}^{n}r_i$ 内に $t[S_j]=t_j[S_j]$ なる組 t が作れる組 $t_1,\cdots,t_m$ があれば, T の各 w_j を t_j に対応させ, かつ各 $\mathrm{chase}_{\Sigma_i}(T_i)$ を r_i の部分関係に対応させることができる (各 Σ_i 内の FD はすべて $\mathop{*}\limits_{i=1}^{n}r_i$ が満たすことに注意). そこで（4）から, t が必ず $\mathop{*}\limits_{i=1}^{n}r_i$ に含まれることが示せる. 停止性も文献 13) の結果から容易に導ける.

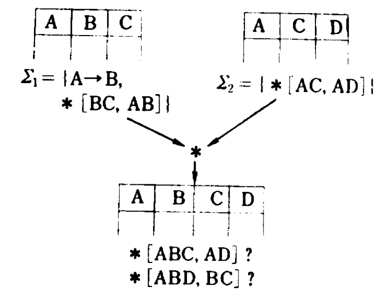

図 3　結合上の JD の検査

Fig. 3 Testing satisfaction of JD on a join relation.

[例]　図 3 の関係スキーム $\underline{R}_1=\langle ABC, \Sigma_1\rangle$, $\underline{R}_2=\langle ACD, \Sigma_2\rangle$ に対し, 図 4（a）は拡張追跡法により $*[ABC, AD]$ が $\underline{R}_1*\underline{R}_2$ 上で成り立つことを示している. ここでは, $\mathop{*}\limits_{i=1}^{2}\mathrm{chase}_{\Sigma_i}(T_i)$ が組 (a_1, a_2, a_3, a_4) を含むのでこのことがいえる. 図 4（b）は $\underline{R}_1*\underline{R}_2$ 上では $*[ABD, BC]$ が成り立たないことを示している. これは, $\mathop{*}\limits_{i=1}^{2}\mathrm{chase}_{\Sigma_i}(T_i)$ が (a_1, a_2, a_3, a_4) を含まないためである. 実際, $\mathrm{chase}_{\Sigma_i}(T_i)\in\mathrm{SAT}(R_i,\Sigma_i)$ $(i=1,2)$ であるのに, $\mathop{*}\limits_{i=1}^{2}\mathrm{chase}_{\Sigma_i}(T_i)$ は $*[ABD, BC]$ を満たしていない. この例の場合, Maier らの追跡法[13]によれば,

$$\Sigma_1\cup\Sigma_2\cup\{*[R_1,R_2]\}\models *[ABD, BC]$$

となってしまうことがわかる. これは先に述べたよう

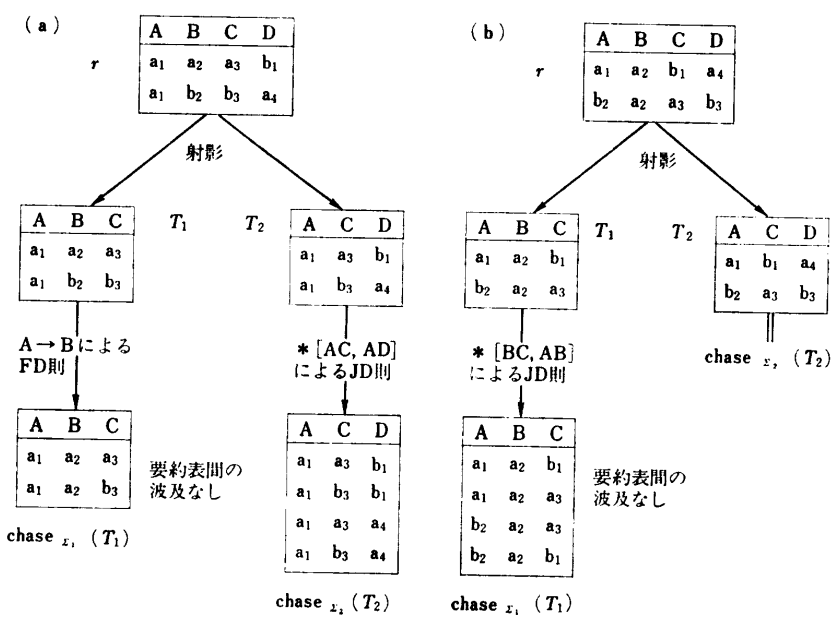

図 4　拡張追跡法の適用例

Fig. 4 Example to illustrate the modified chase method.

に上記の式が必要条件にしかなっていないことを示している.

4. 制約演算と従属性との干渉

関係スキーム $\underline{R}=\langle R, \Sigma\rangle$ に対し，Σ は FD, JD 集合であり，かつ，Σ は R 内のある A に対し $X\to A$ ($A\notin X$, X は空集合でない) なる形のいかなる FD も論理的に含意しないものとしよう．このとき，任意の $r\in \mathrm{SAT}$ (R, Σ) に対して，$r[A='c']$ ($A\in R$, c: 定数) 上ではいかなる組の A の値も c であるかまたは $r[A='c']=\phi$ であるので，$\phi\to A$ という FD がつねに満足される．さらに，この種の制約演算から，元の Σ 内の FD や JD は影響を受けないこと，および $\underline{R}[A='c']$ で成立する FD, JD 集合は Σ と $\phi\to A$ から導けるものに限ることが次の定理よりわかる.

［定理2］ 関係スキーム $\underline{R}=\langle R, \Sigma\rangle$ (Σ は FD, JD 集合)，$A\in R$, 定数 c に対して

$$\{r[A='c']\mid r\in \mathrm{SAT}\ (R, \Sigma)\}$$
$$=\mathrm{SAT}\ (R, \Sigma\cup\{\phi\to A\}).$$

ただし，各属性の定義域はおのおの2個以上の値からなり，Σ は $X\to A$ ($A\notin X$ かつ $X\neq\phi$) という形のいかなる FD も論理的に含意しないもの*とする.

（証明） $\underline{R}[A='c']$ 上で $\phi\to A$ が成り立つのは明らか．また Σ 内の任意の FD や JD も，A の値が定数になってもその定義から，やはり $\underline{R}[A='c']$ 上で成り立つ．よって，あとは SAT (R, Σ) 内の関係 r で，$r[A='c']$ が $\Sigma\cup\{\phi\to A\}$ から論理的に含意されるもののみを満たすものがつねに構成できればよい．$\Sigma\cup\{\phi\to A\}$ のアームストロング関係[6] r_1 に対して，いま $r_1[A]=\{c\}$ とする．いま，R 上の関係 r_2 が $r_2[A]=\{c'\}$ $(c'\neq c)$, $r_2[R-A]=r_1[R-A]$, $B\neq A$, $B\in R$ なる任意の B に対し，$c'\notin r_2[B]$ とすると，r_2 も $\Sigma\cup\{\phi\to A\}$ のアームストロング関係である．$r=r_1\cup r_2$ とすると次の(ｉ)，(ⅱ)がいえる.

(ｉ) Σ 内の任意の FD: $X\to B$ $(A\neq B)$ を r は満たす．これは，$t_1[X]=t_2[X]$ なる r の任意の組 t_1, t_2 に対し，$X\ni A$ のときは t_1, t_2 ともに r_1 または r_2 の一方に属し，$X\not\ni A$ のときは $r_2[R-A]=r_1[R-A]$ であることから明らかである.

(ⅱ) Σ 内の任意の JD: $*[S_1,\cdots,S_m]$ を r は満たす．これは，r 内の組 $t_1,\cdots,t_m$ が $t[S_j]=t_j[S_j]$ $(j=1,\cdots,m)$ なる組 t が構成できれば，$t[A]$ の値は c か c' のいずれかであることと，$r_1[R-A]=r_2[R-A]$ であることから明らかである.

よって $r\in \mathrm{SAT}$ (R, Σ) であり $r[A='c']$ は $\Sigma\cup\{\phi\to A\}$ のアームストロング関係である．(証明終)

次に，$r[A=B]$ という形の制約については1章でも述べたように，r が満たしていたある JD が，$r[A=B]$ では満たされないことがあり注意を要する．これに関して，次の定理を得る.

［定理3］ 関係スキーム $\underline{R}=\langle R, \Sigma\rangle$ において，$\Sigma=F\cup J$, F は FD 集合，J は JD 集合とする．R 内の任意の相異なる属性 A, B に対し，

$$F'=F\cup\{A\to B, B\to A\},$$
$$J'=\{*[S_1,\cdots,S_m, AB]\mid *[S_1,\cdots,S_m]\in J\}$$

とすると，$F'\cup J'$ が $\underline{R}[A=B]$ 上で成り立つ.

（証明） $\underline{R}[A=B]$ 上で F' が成り立つことは明らか．J 内の JD: $*[S_1,\cdots,S_m]$, $r\in \mathrm{SAT}$ (R, Σ) に対し，いま $r[A=B]$ 内に組 $t_1,\cdots,t_{m+1}$ があり，$t[S_j]=t_j[S_j]$ $(j=1,\cdots,m)$, $t[AB]=t_{m+1}[AB]$ なる組 t が構成可能とする．$t_1,\cdots,t_m$ は r にも属し，r は $*[S_1,\cdots,S_m]$ を満たすから，t は r に属し，かつ $t[A]=t[B]$ である．よって t は $r[A=B]$ にも属す．すなわち，$r[A=B]$ は $*[S_1,\cdots,S_m, AB]$ を満たす.

(証明終)

上記の定理で，たとえば Σ 内の JD $*[AB, AC, CD]$ は，$r[B=C]$ 上では $*[AB, AC, CD, BC]$ の形に変形される．$*[AB, AC, CD]\models *[AB, AC, CD, BC]$ であるが，$*[AB, AC, CD, BC]\models *[AB, AC, CD]$ ではないことに注意が必要である．また，上記の定理よりただちに，JD の保存に関する次の条件が得られる.

$\underline{R}=\langle R, \Sigma\rangle$ 内の JD $*[S_1,\cdots,S_m]$ が $\underline{R}[A=B]$ 上で成立するための十分条件は，少なくとも一つの S_j $(j\in\{1,\cdots,m\})$ が属性 A と B を含むことである.

さらに，元の MVD $A\to\to B|C$ が縮退 MVD (degenerated MVD)[1] である場合には，つねに $\underline{R}[A=B]$ 上でも $A\to\to B|C$ が成立することが容易に示せる.

5. JD 保存のための諸条件

3章では，関係スキーム $\underline{R}_1,\cdots,\underline{R}_n$ に対して，ある従属性 (FD, EJD) が $\mathop{*}\limits_{i=1}^{n}\underline{R}_i$ 上で成立するか否かという問題について論じた．本章では，元の関係スキームの JD が結合演算で保存されるための諸条件を示す．このことに関して，まず3章の拡張追跡法から次の定理を得る.

* 最近，武田と著者らは，この制限をとり除けることを文献21) に示している.

［定理4］ 関係スキーム $\underline{R}_1, \cdots, \underline{R}_n$ のうち，ある $\underline{R}_j=\langle R_j, \Sigma_j\rangle$ において，Σ_j が JD：$*[S_1, \cdots, S_m]$ を含むとする．このとき，（1）$n=2$ または（2）$\{1, \cdots, n\}$ 内のすべての $k, l\ (k\neq l)$ に対し $(R_k-R_j)\cap(R_l-R_j)=\phi$ のいずれかが成り立てば，$\mathop{*}\limits_{i=1}^{n}\underline{R}_i$ 上で $*[S_1, \cdots, S_m]$ が成立するための必要十分条件として次を得る：3章の拡張追跡法のステップ（3）において，各 $\mathrm{chase}_{\Sigma_p}(T_p)\ (p=1, \cdots, n, p\neq j)$ が $R_j\cap R_p$ にすべて識別変数をもつような組を含むことである．

（略証） （1）または（2）が成り立つときは，各 $\mathrm{chase}_{\Sigma_p}(T_p)\ (p=1, \cdots, n, p\neq j)$ において，R_p-R_j 内の属性に対応する部分は結合にいっさい関与しないことと，3章の拡張追跡法の正しさから明らかである．

定理4は拡張追跡法そのものにもとづくものであるが，次に，元の関係スキームの JD が $\mathop{*}\limits_{i=1}^{n}\underline{R}_i$ 上で EJDとして成り立つためには各 $\underline{R}_i$ にどのような従属性が課されていればよいかという問題が考えられる．関係スキームの個数を二つに制限したとき，次の諸条件が得られる．

［定理5］ 関係スキーム $\underline{R}_1=\langle R_1, \Sigma_1\rangle$，$\underline{R}_2=\langle R_2, \Sigma_2\rangle$（$\Sigma_1, \Sigma_2$ はおのおの R_1, R_2 上の FD, JD 集合）において，Σ_1 内の JD を $*[S_1, \cdots, S_m]$ とし，さらにある $S_1, \cdots, S_k\ (k\leq m)$ に対し，

$$R_1\cap R_2=(S_1\cap R_2)\cup\cdots\cup(S_k\cap R_2)$$

とする．ここで，$T_i=S_i\cap R_2\ (i=1, \cdots, m)$ とすると，次の条件のいずれかが満足されれば，$*[S_1, \cdots, S_m]$ は $\underline{R}_1*\underline{R}_2$ 上で成立する．

（1） $\Sigma_2\models *[T_1, \cdots, T_m]$

（2） $(\Sigma_2\models T_1\cap T_2\rightarrow T_1$ または $\Sigma_2\models T_1\cap T_2\rightarrow T_2)$
かつ
$(\Sigma_2\models T_2\cap T_3\rightarrow T_2$ または $\Sigma_2\models T_2\cap T_3\rightarrow T_3)$
かつ
⋮
$(\Sigma_2\models T_{k-1}\cap T_k\rightarrow T_{k-1}$ または
$\Sigma_2\models T_{k-1}\cap T_k\rightarrow T_k)$

（3） $(\Sigma_1\models S_1\cap S_2\rightarrow T_1$ または $\Sigma_2\models S_1\cap S_2\rightarrow T_2)$
かつ
$(\Sigma_1\models S_2\cap S_3\rightarrow T_2$ または $\Sigma_1\models S_2\cap S_3\rightarrow T_3)$
かつ
⋮
$(\Sigma_1\models S_{k-1}\cap S_k\rightarrow T_{k-1}$ または
$\Sigma_1\models S_{k-1}\cap S_k\rightarrow T_k)$

（略証） $\mathrm{SAT}(R_1, \Sigma_1)\ni r_1$，$\mathrm{SAT}(R_2, \Sigma_2)\ni r_2$ なる任意の関係 r_1, r_2 に対し，r_1*r_2 が組 $t_1, \cdots, t_m$ を含み，$t[S_j]=t_j[S_j]\ (j=1, \cdots, m)$ なる組 t が構成可能とする．r_1 は $*[S_1, \cdots, S_m]$ を満たすから，明らかに $t\in r_1$ である．そこで，上記の（1）～（3）のいずれかが成り立てば $t[R_1\cap R_2]\in r_2[R_1\cap R_2]$ となることを示せばよい．仮定から，部分組 $t_1[S_1\cap R_2], \cdots, t_m[S_m\cap R_2]$ はすべて r_2 に現れる．そこで，（1）が成り立てば，$*[T_1, \cdots, T_m]=*[S_1\cap R_2, \cdots, S_m\cap R_2]$ だから，$t[R_1\cap R_2]\in r_2[R_1\cap R_2]$ は明らか．次に（2）が成り立つ場合，部分組 $t_1[T_1], \cdots, t_k[T_k]$ が r_2 に現れ，$t_1[T_1\cap T_2]=t_2[T_1\cap T_2], \cdots, T_{k-1}[T_{k-1}\cap T_k]=t_k[T_{k-1}\cap T_k]$ であることに注目すると，$t[R_1\cap R_2]=t[T_1\cdots T_k]\in r_2[R_1\cap R_2]$ が示せる．これはたとえば，$T_1\cap T_2\rightarrow T_1$ であれば，$t_1[T_1]=t_2[T_1]$ なので，$t_2[T_1T_2]=t[T_1T_2]$ となる．これを繰り返すことで，ある $t_j(j\in\{1, \cdots, k\})$ に対して $t_j[T_1T_2\cdots T_k]=t[T_1T_2\cdots T_k]$ となることが容易に示せるためである．次に（3）が成り立つ場合，$t_1[S_1\cap S_2]=t_2[S_1\cap S_2], \cdots, t_{k-1}[S_{k-1}\cap S_k]=t_k[S_{k-1}\cap S_k]$ であり，（3）の FD の組合せを r_1 が満たすから，（2）の場合と同様の方法で，ある $t_j[T_1\cdots T_k]$ が $t[T_1\cdots T_k]$ と等しいことがわかる．$t_j[T_1\cdots T_k]\in r_2[R_1\cap R_2]$ だから $t[T_1\cdots T_k]\in r_2[R_1\cap R_2]$．よって結論が得られた． （証明終）

定理5は，われわれの文献17）の結果の拡張となっている．上記の定理で，条件（2）と（3）が独立した形で与えられていることには注意を要する．たとえば，（2）の FD $T_1\cap T_2\rightarrow T_1$ と $S_1\cap S_2\supseteq T_1\cap T_2$ であることから，FD の推論則[20]を用いれば，（3）の FD $S_1\cap S_2\rightarrow T_1$ が導出できる．しかし本論文では関係スキーム $\underline{R}_1, \underline{R}_2$ は互いに独立した形で与えられることを前提にしているため，たとえば，$R_1=ABC$, $R_2=BCD$ で，$\underline{R}_1$ 上で $B\rightarrow C$ が成立しても $\underline{R}_2$ 上では $B\rightarrow C$ が必ずしも成立しない．これは，文献7）の普遍関係スキーム仮定すら設けないことを意味し，より一般的な結果となっている．もちろん，この仮定を設ければ，条件（3）は（2）から導くことができる．

定理5の系として，R_1 上の MVD $X\rightarrow\rightarrow Y|Z$（$R_1\cup R_2$ 上の EMVD）が $\underline{R}_1*\underline{R}_2$ 上で成立するための次の十分条件が容易に得られる．

［系］[17] 関係スキーム $\underline{R}_1=\langle R_1, \Sigma_1\rangle$，$\underline{R}_2=\langle R_2, \Sigma_2\rangle$ に対し，$R_1=XYZ$, $X\rightarrow\rightarrow Y|Z\in\Sigma_1$ なる任意の従属性は次の条件のいずれかが満足されれば，$\underline{R}_1*\underline{R}_2$ 上でも成立する：（1）$\Sigma_2\models *[XY\cap R_2, XZ\cap R_2]$，（2）$\Sigma_2\models X\cap R_2\rightarrow Y\cap R_2$ または $\Sigma_2\models X\cap R_2\rightarrow Z\cap R_2$，（3）

$\Sigma_1 \models X \to Y \cap R_2$ または $\Sigma_1 \models X \to Z \cap R_2$.

［例］ $\underline{R}_1 = \langle R_1 = ABCDEF,\ \Sigma_1 \rangle$, $\underline{R}_2 = \langle R_2 = ACEG, \Sigma_2 \rangle$ とし，$*[ABC, CDE, EFA] \in \Sigma_1$ とする．$S_1 = ABC$, $S_2 = CDE$, $S_3 = EFA$ とすれば，$R_1 \cap R_2 = ACE$ だから $R_1 \cap R_2 = (S_1 \cap R_2) \cup (S_2 \cap R_2)$ である．よって，定理5で，$T_1 = AC$, $T_2 = CE$ となるので，次の(1)〜(3)のいずれかが成り立てば，$*[ABC, CDE, EFA]$ が $\underline{R}_1 * \underline{R}_2$ 上で成り立つ：(1) $\Sigma_2 \models *[AC, CE]$, (2) $\Sigma_2 \models C \to A$ または $\Sigma_2 \models C \to E$, (3) $\Sigma_1 \models C \to A$ または $\Sigma_1 \models C \to E$.

さらに，JD$*[ABC, CDE, EFA]$ は各属性集合の順序を変えた $*[ABC, EFA, CDE]$ や $*[CDE, EFA, ABC]$ とは明らかに等価であり，この二つの JD に対し，おのおの定理5を適用すると次の条件(1)′〜(3)′，(1)″〜(3)″ を得る：(1)′ $\Sigma_2 \models *[AC, AE]$, (2)′ $\Sigma_2 \models A \to C$ または $\Sigma_2 \models A \to E$, (3)′ $\Sigma_1 \models A \to C$ または $\Sigma_1 \models A \to E$, (1)″ $\Sigma_2 \models *[CE, AE]$, (2)″ $\Sigma_2 \models E \to A$ または $\Sigma_2 \models E \to C$, (3)″ $\Sigma_1 \models E \to A$ または $\Sigma_1 \models E \to C$. よって(1)〜(3)，(1)′〜(3)′，(1)″〜(3)″ の九つのうち少なくとも一つが成立すれば $*[ABC, CDE, EFA]$ が $\underline{R}_1 * \underline{R}_2$ 上で成立することがわかる．このことは3章で示した拡張追跡法を用いても容易に確かめることができる．

6. 属性名変更による JD の保存

本章では，ある関係スキーム上の JD が結合演算ののちに保存されない場合に，いくつかの属性名を変更することで保存されるようにする基本的な方法について述べる．

関係スキーム $\underline{R}_i = \langle R_i, \Sigma_i \rangle$ $(i = 1, \cdots, n$, Σ_i は R_i 上の FD, JD の集合), Σ_j 内の JD σ に対して，いま σ が $\mathop{*}\limits_{i=1}^{n} \underline{R}_i$ 上で成立しないとする．以下に示す方法は，属性名変更により σ を $\mathop{*}\limits_{i=1}^{n} \underline{R}_i$ 上で成立するようにし，かつ変更する属性数を低減させるものである．

［属性名変更による JD の保存］

(1) $\underline{R}_1, \cdots, \underline{R}_n$ を［定理4］の条件(2)を満足するように属性名変更を行い，この結果を改めて $\underline{R}_1, \cdots, \underline{R}_n$ とする．ただし変更する属性数を最小にする．

(2) σ と $\{\underline{R}_1, \cdots, \underline{R}_n\}$ に対し，拡張追跡法を適用．σ が $\mathop{*}\limits_{i=1}^{n} \underline{R}_i$ 上で成立することがわかれば終了．

(3) 各 R_i $(i = 1, \cdots, n,\ i \neq j)$ の部分集合 X のうち，要約表 T_i の組の $R_j \cap (R_i - X)$ に識別変数のみが現れるような極小な X の集合を CANDIDATE (T_i) とする．

(4) すべての CANDIDATE (T_i) $(i = 1, \cdots, n,\ i \neq j)$ に対して，次の条件を満たす属性集合 Y を求める：(a) Y は各 CANDIDATE (T_i) 内の少なくとも一つの要素を含む．(b)(a)を満たすもののうち，Y の要素数が最小である．

(5) 各 T_i $(i \neq j)$ に対する R_i について，$R_i \cap Y$ 内の属性名を変更する．

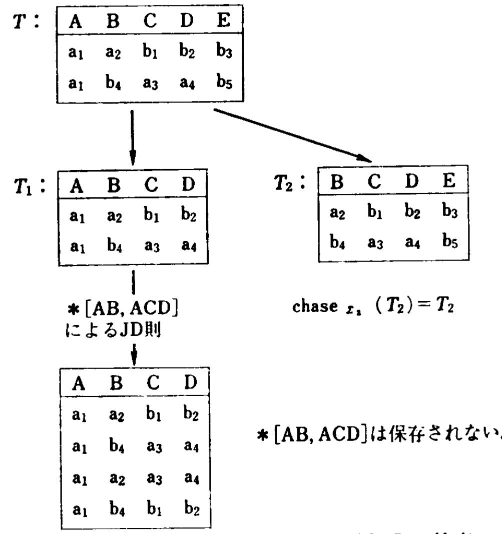

図5 (a) $\underline{R}_1 * \underline{R}_2$ 上の $*[AB, ACD]$ の検査

Fig. 5 (a) Testing satisfaction of $*[AB, ACD]$ on $\underline{R}_1 * \underline{R}_2$.

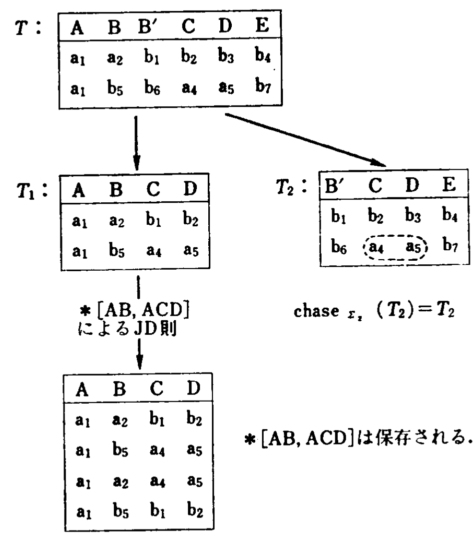

図5 (b) R_2 の属性を B' とした場合の $*[AB, ACD]$ の検査

Fig. 5 (b) Testing satisfaction of $*[AB, ACD]$ when attribute B in R_2 is replaced by B'.

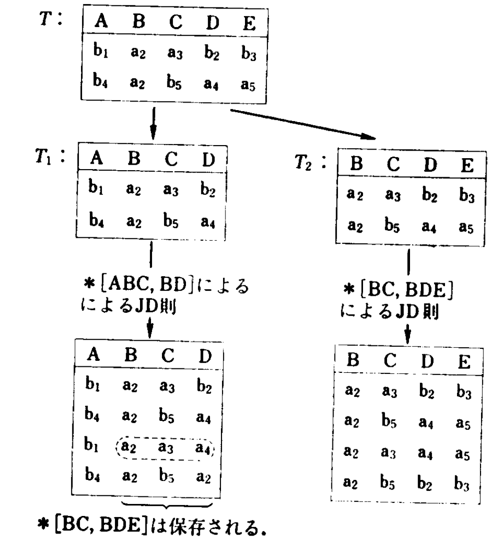

図 5 （c） $\underline{R}_1*\underline{R}_2$ 上の $*[BC,BDE]$ の検査
Fig. 5 (c) Testing satisfaction of $*[BC,BDE]$ on $\underline{R}_1*\underline{R}_2$.

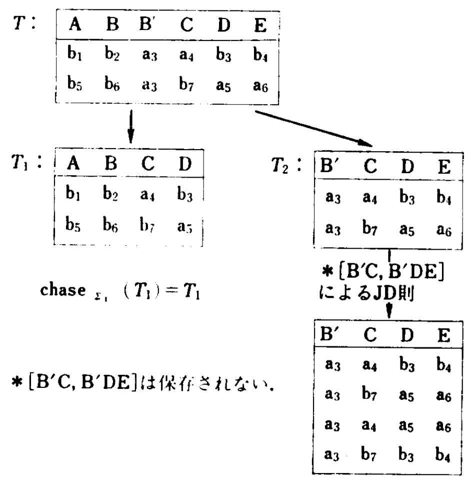

図 5 （d） $*[B'C,B'DE]$ の検査
Fig. 5 (d) Testing satisfaction of $*[B'C,B'DE]$.

［例］ $\underline{R}_1=\langle R_1=ABCD, \Sigma_1=\{*[AB,ACD], *[ABC,BD]\}\rangle$, $\underline{R}_2=\langle R_2=BCDE, \Sigma_2=\{*[BC,BDE]\}\rangle$ とする．図5（a）に示すように，$\underline{R}_1$ の JD $*[AB,ACD]$ は $\underline{R}_1*\underline{R}_2$ 上では成立しない．これは図5（a）の $chase_{\Sigma_2}(T_2)$ が (a_2,a_3,a_4,b)（b は任意の非識別変数）という組を含まないためである．この例では上記の方法のステップ(1)は意味がなく，そこでステップ(3)で CANDIDATE $(T_2)=\{\{B\}\}$ を得る．よって，R_2 の属性 B を B' に変更すれば，$*[AB,ACD]$ は $\underline{R}_1*\underline{R}_2'$ 上で成立する（図5（b）参照）．この方法で注意が必要なのは，ある σ を保存するためにある属性の名前を変更すると，それまで保存できていた別のJD σ' が保存できなくなることがあることである．たとえば，$\underline{R}_2$ の JD$*[BC,BDE]$ は元の $\underline{R}_1$, $\underline{R}_2$ に対しては $\underline{R}_1*\underline{R}_2$ 上で成立する（図5（c）参照）．しかし，図5（d）に示すように，R_2 の属性 B を B' に変更したあとでは，$\underline{R}_2$ の JD$*[B'C,B'DE]$ はもはや $\underline{R}_1*\underline{R}_2$ 上では成立しない（図5（d）参照）．したがって，すべての元の JD を保存するために属性名変更を行い，かつ，変更する属性数を最小にする問題は，今後の課題として残っている．

7. あとがき

本論文では，FD･EJD といった従属性制約と関係質問との干渉問題について論じた．結合や制約という演算を用いる質問では，もともとある関係でつねに満たされている JD や MVD という従属性が，その結果上では EJD という形でさえも必ずしも満たされなくなったり，また新たな従属性が追加される場合がある．このことは，1章でも触れたように，質問処理のみならず，データベースやビューの設計においても重要な問題となる．たとえば，関係の縦切り分解による従来の設計理論は元の情報の復元に結合演算を想定しており[20]，与えられた従属性がこの設計にどの程度反映されたかという問題はやはり従属性と結合演算との干渉問題に帰着できるためである．

本論文の主要な結果と今後の課題は以下のとおりである：

（1） 各関係に FD･JD が与えられたとき，これらの結合後にある EJD がつねに満たされるかどうかを検査する手法として，拡張追跡法を提案した．この方法は Maier らの追跡法[13]に結合・射影演算を導入して拡張したものである．他の演算に対する一般的検査法は今後の課題である．また，この問題が，元の従属性集合と結合によって生じる JD による従属性の論理的含意という形で扱えるための諸条件も示した．

（2） 制約演算と FD, JD との干渉問題に対する検討．

（3） ある関係で成立する JD が結合後も EJD の形で成立するための諸条件を，元の従属性集合の満たすべき性質という形で得た．また，この問題に対しては，結合後も成立するように属性名変更を行う方法も提案した．属性数の最小化等は今後の課題である．

謝辞 ご検討いただいた，京都大学工学部矢島脩三教授ならびに武田浩一氏（現，日本 IBM）に深謝します．

参 考 文 献

1) Armstrong, W. W. and Delobel, C.: Decompositions and Functional Dependencies in Relations, ACMTODS, Vol. 5, No. 4, pp. 404-430 (1980).
2) Beeri, C. and Rissanen, J.: Faithful Representations of Relational Database Schemes, IBM Res. Rep., RJ 2722 (1980).
3) Casanova, M. A.: A Theory of Data Dependencies over Relational Expressions, ACM Symp. PODS, pp. 189-198 (1982).
4) Codd, E. F.: Further Normalization of the Data Base Relational Model, Courant Computer Science Symp. 6, pp. 33-64 (1971).
5) Fagin, R.: Multivalued Dependencies and a New Normal Form for Relational Databases, ACMTODS, Vol. 2, No. 3, pp. 262-278 (1977).
6) Fagin, R.: Horn Clauses and Database Dependencies, ACM SIGACT Symp. Theory of Computing, pp. 123-134 (1980).
7) Fagin, R., Mendelzon, A. O. and Ullman, J. D.: A Simplified Universal Relation Assumption and Its Properties, IBM Res. Rep., RJ 2900 (1980).
8) Hull, R.: Implicational Dependency and Finite Specification, Univ. of Southern California, Computer Science Dept., Tech. Rep. (1981).
9) 伊藤, 岩崎, 谷口, 嵩: 関係データベースにおけるビュー上の従属性, 信学技報, AL 81-25 (1981. 6).
10) Kambayashi, Y., Tanaka, K., Takeda, K. and Yajima, S.: Representation of Relations for Database Output Utilizing Data Dependencies, Proc. 15 th Hawaii International Conference on System Sciences, pp. 69-78 (1982).
11) Klug, A.: Calculating Constraints on Relational Expressions, ACMTODS, Vol. 5, No. 3, pp. 260-290 (1980).
12) Klug, A.: On Inequality Tableaux, Univ. of Wisconsin-Madison, Computer Science, Tech. Rep. #403 (1980).
13) Maier, D., Mendelzon, A. O. and Sagiv, Y.: Testing Implications of Data Dependencies, ACMTODS, Vol. 4, No. 4, pp. 455-469 (1979).
14) Maier, D., Mendelzon, A. O., Sadri, F. and Ullman, J. D.: Adequacy of Decompositions of Relational Databases, *J. Comput. Syst. Sci.*, Vol. 21, No. 3, pp. 368-379 (1980).
15) 増永: 関係データベースにおけるカジュアルユーザの質問意図の解析について, 情報処理学会データベース管理システム研究会, 12-1 (1979. 3).
16) Rissanen, J.: Theory of Relations for Databases—A Tutorial Survey, Lecture Notes in Computer Science 64, pp. 536-551 (1978).
17) Tanaka, K., Kambayashi, Y. and Yajima, S.: Preservability of Data Dependencies for Relational Database Operations, Proc. JIPDEC Information Systems Seminar on Semantic Aspects of Databases, pp. 151-174 (1980).
18) Tanaka, K. and Kambayashi, Y.: Testing Join Dependency Preserving by a Modified Chase Method, Lecture Notes in Computer Science 118, Springer-Verlag, Berlin, Heidelberg, New York, Proc. 10 th Symp. on MFCS '81, pp. 524-533 (1981).
19) Tanaka, K. and Kambayashi, Y.: Logical Integration of Locally Independent Relational Databases into a Distributed Database, Proc. 7 th International Conference on VLDB, pp. 131-141 (1981).
20) Ullman, J. D.: *Principles of Database Systems*, 2nd ed., Computer Science Press, Potomac, Maryland (1982).
21) 武田, 上林, 田中: 質問結果に現われる従属性集合の計算について, 情報処理学会第 26 回全国大会, 1 F-2 (1983. 3).

(昭和 57 年 12 月 27 日受付)
(昭和 59 年 12 月 20 日採録)