

同値律をみたす述語による
関係データモデル従属性の一般化について

富士通・国際情報社会科学研究所 竹島 卓

1. まえがき

関係データモデル (Codd [1]) の提案以来、いわゆる従属性の問題が種々の意図をもって研究されている。それらは、(1) data anomaly の回避ないしは正規化 (2) 関係の情報無損失分解 (3) 検索 (内包せ) に対するアクセスパスの決定 (4) 簡易な統合制約 (5) 数学的興味 などに代表される。

そこにおける課題としては、(1°) 情報無損失分解の必要十分条件 (2°) 適切な正規形の設定 (3°) 従属性についての適切な推論体系 などがあげられる。 (1°) に対しては結合従属 (JD), (2°) に対しては射影結合正規形 (PJNF) によって一応の結着がついている。 (3°) については潜在形の従属性 (embedded dependency) には有限の公理系が存在しないという研究結果があり、適切な推論体系が望まれている。そこで本稿では課題 (3°) に対するひとつの解答を示す。

従属性を一階述語論理によつて形式化しようとする研究は
已に Nicolas [2] や Fagin [3] らによつて試みられてゐる。
彼等のアプローチでは、(i) 関係にひとつつの述語を対応させる。
(ii) タップルの成分を走る成分変数を用ひる。 (iii) 成分値に対する等号 (=) の述語を用ひる。 とくに特徴をもつてゐる。
彼等の形式化はかなりうまくいってゐるが、それは少し乱暴な形式化によつて(i)の方法を採用してゐるために、実は一階述語論理の意味論そのものが展開されてゐるからに他ならぬ。
しかしながら次の二点は問題がある。(a) 関係モデルは本来 many sorted なものである。(b) 上記性質を無視したためドメイン関数の選択によらない不变な構造(同値関係や
からみ合構造[4]など)が表現できない。

これに対し、本稿で述べる形式化は、(i') 各属性にひとつの二項述語を割当てる。 (ii') タップルの上を走るタップル変数を用ひる。 (iii') 属性の数だけの同値律集合(LE)を用ひる。
とくに特徴をもち、上記(a)および(b)の欠点を解消してゐる。
また本稿で述べる結果は、独立に行なわれたカテゴリー論によるアプローチ(加藤[4])の結果ともよほど対応し、それによつて代数的な肉付けが得られてゐる。

2 関係モデル

従属性の形式化にとって最小限必要な関係モデルでの概念を定義する。その定義は、従来のものを多少整理洗練した形で述べてある。それは、空な属性集合やその上の関係、一般には非可算無限の濃度をもつ関係などを含んでいる。

したがって必ずしも計算機×モリ上に展開できるとはかぎらない、数学的な関係を扱うが、これは統一的な議論のためににはたらのよ“特別なものばかり”を扱うわけにはいかないからである。

関係 R は 3 項組 $\langle A, \mathcal{D}, R \rangle$ のことである。ここに、 A は高々可算な集合で関係 R の属性集合と呼ばれる。 \mathcal{D} は空でない、任意の集合 $D_a (a \in A)$ から成る A 上の集合族で関係 R のドメイン関数と呼ばれる。また R は集合族 $\mathcal{D} = \{D_a\}_{a \in A}$ の直積 $\prod \mathcal{D}$ の任意の部分集合で関係 R の本体呼ばれる。こには集合族 \mathcal{D} の直積 $\prod \mathcal{D}$ は次式で定義される。

$$\prod \mathcal{D} \triangleq \{\lambda \mid \lambda: A \rightarrow \bigcup_{a \in A} D_a \ \& \ \forall a \in A (\lambda(a) \in D_a)\} \quad (1)$$

$\prod \mathcal{D}$ は \mathcal{D} の定義域である属性集合 A 上の（いとつの）関係エンバースと呼ばれることがあり、そのときは $R(A, \mathcal{D})$ とも書かれる。関数 $\lambda \in R(A, \mathcal{D})$ は R のタップル又は \mathcal{D} 上のタップルと呼ばれる。また $\lambda(a) \in D_a$ はタップル λ の a -成分と呼ばれ、従来 $\lambda.a$ と書かれたものである。

我々は A を任意にひとつ固定したとき、その上で可能なすべての関係 R に关心がある。ドメイン関数 ϕ は属性 $a \in A$ の意味を与えるものと解釈できるが、従属性といわれるものは実はこの意味付けに依存しない不变な構造の表現を意図したものであつたと理解される。

3 従属性の形式化

3.1 構文

属性集合 A を任意にひとつ固定し、その上の言語（従属性言語） $L(A)$ をつきのように定める。これは等号なしの一階言語の小さな部分集合である。ここではやや非形式的に述べるがその意図するところは了解できると思う。

(1) 基本記号

- ・ 変数記号の集合 $X = \{x, y, z, \dots\}$.
- ・ 2項述語記号の集合 $P = \{p_a\}_{a \in A}$,
ただし $a \neq b$ ならば $p_a \neq p_b$.
- ・ 基本論理記号の集合 $\{\sim(\text{not}), \wedge(\text{and}), \forall(\text{for all})\}$.

(2) 原子式

$x, y \in X$ で $p_a \in P$ のとき $p_a(x, y)$ は原子式。

(3) 式

(i) α が原子式 $\Rightarrow \alpha$ は式。

(ii) α が式 $\Rightarrow \neg\alpha$ は式.

(iii) α, β が式 $\Rightarrow \alpha \wedge \beta$ は式.

(iv) α が式, $x \in X \Rightarrow \forall x \alpha$ は式.

(4) 閉じた式

α が式でかつ α の中のすべての変数記号の出現が束縛されているとき, α は 閉じていい といわれる. 閉じた式のことを 文ともいう.

3.2 意味

(1) モデル

文の意味を定めるためにモデルを定義する. モデル M は組 $\langle D, A_p \rangle$ である. ここに D は空でない集合で M の基礎領域と呼ばれる. また A_p は 2 項述語記号 $p_a \in P$ 達に D 上の 2 項述語 $\pi_a \in 2^{D \times D}$ 達を割当てる関数で P 上の割当と呼ばれる.

(2) 変数記号への割当

モデル $M = \langle D, A \rangle$ 上の変数記号への割当て A_v とは、変数記号 $x \in X$ に D の元を割当てる関数である. すなわち.

$A_v(x) \in D$ である.

(3) 割当ての変種

割当て A_v の変種 $A_v[x/d]$ とはやはり M 上の変数記号への割当てである. つまりのようく定義される. 任意の $y \in X$

$$\text{に対して } A_\nu[x/d](\xi) = \begin{cases} d & \text{if } \xi = x \\ A_\nu(\xi) & \text{otherwise} \end{cases}$$

(4) 付値

式 γ の付値関数とは、モデル $M = \langle D, A_p \rangle$ と割当て A_ν に基づく定まる式の意味 — 1(真), 0(偽) — を計算する関数である。これを $V_{A_\nu}^M(\gamma)$ と書く。 $V_{A_\nu}^M(\gamma)$ はつきのように帰納的に定義される。

(i) 原子式 $p_a(x, y)$ に対して、

$$V_{A_\nu}^M(p_a(x, y)) = A_p(p_a)(A_\nu(x), A_\nu(y)).$$

(ii) 式 $\sim x$ に対して、

$$V_{A_\nu}^M(\sim x) = \neg V_{A_\nu}^M(x). \quad (*)$$

(iii) 式 $x \wedge \beta$ に対して、

$$V_{A_\nu}^M(x \wedge \beta) = V_{A_\nu}^M(x) \wedge V_{A_\nu}^M(\beta). \quad (*)$$

(iv) 式 $\forall x \alpha$ に対して、

$$V_{A_\nu}^M(\forall x \alpha) = \bigvee_{d \in D} V_{A_\nu}^M(\alpha), \quad (*)$$

ここで $A_\nu = A_\nu[x/\bar{d}]$.

この付値関数が論理結合記号の通常の解釈に沿つたものであることは、脚注の $\times \neg$ 記号の定義によつてわかる。

(*) 注： \neg , \wedge , $\&$ は意味領域での演算記号でそれぞれ、

$\neg 0 = 1$, $\neg 1 = 0$, $0 \wedge 0 = 0$, $0 \wedge 1 = 1$, $1 \wedge 0 = 0$, $1 \wedge 1 = 1$, $\bigvee_{d \in D} v_d = \text{if}$
for all $d \in D$ $v_d = 1$ then 1 else 0, と定められる。

3.3 二次論理記号

記法上の便宜のために, \vee (or), \rightarrow (implies), \exists (for some) の3論理結合記号を定義する。定義は構文的に行なうが、その意味が通常の解釈と変わらないことは、前項(4)の付値関数の定義にもどして確かめられる。定義はつきのとおり。

$$(i) \alpha \vee \beta \triangleq \sim(\sim\alpha \wedge \sim\beta)$$

$$(ii) \alpha \rightarrow \beta \triangleq \sim\alpha \vee \beta$$

$$(iii) \exists x \alpha \triangleq \sim \forall x \sim \alpha$$

3.4 モデルによる式の充足

式 γ が割当 Au によってモデル M において充足されたとは、

$V_{Au}^M(\gamma) = 1$ となることである。このことを, $M, Au \text{ sat } \gamma$

と書く。また式の集合 Γ に対しては, $M, Au \text{ sat } \Gamma$

とは、任意の $\gamma \in \Gamma$ について $M, Au \text{ sat } \gamma$ であることをいう。

特に γ が閉じていい場合、すなわち、文 α である場合には、付値 $V_{Au}^M(\alpha)$ の値は Au に依存しないので、 Au を省略して、 $M \text{ sat } \alpha$ と書く。同様に、文の集合 Σ に対しても Au を省略して、 $M \text{ sat } \Sigma$ と書く。

3.5 関係による式の充足

(1) 誘導モデル

本体が空でない関係 $R = \langle A, D, R \rangle$ はつきのようにひとつのモデル $M^* = \langle D^*, Ap^* \rangle$ を誘導する。すなわち、

$D^\# = R$, $A_p^\# \rightarrow 2^{R \times R}$ s.t. $A_p^\#(p_a) = \pi_a$ とおくとき,

$$\lambda, \mu \in R \text{ に対して } \pi_a(\lambda, \mu) = \begin{cases} 1 & \text{if } \lambda(a) = \mu(a) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

(2) 関係による充足

式 γ が割当て A_u によって関係 R に於いて充足されるということを $R, A_u \text{ sat } \gamma$ と書くことにすれば、それは誘導モデルによってつきのように定義される。

$$\underline{R, A_u \text{ sat } \gamma} \stackrel{\text{def}}{=} M^\#, A_u \text{ sat } \gamma.$$

ただし、 A_u は変数記号への R の元の割当てである。同じく式、すなわち文 α や、文の集合 Σ に対してはモデルでの解釈同様 A_u を省略し、 $R \text{ sat } \alpha$ や $R \text{ sat } \Sigma$ のように書く。

これが「従属性 α が関係 R で成立する。」という宣言のモデル論的定義である。ここで、 α は従来の肯定的な従属性のみならず、それの否定であってもかまわないことに注意を換起しておく。

4 同値律集合 LE

関係 R の誘導するモデル $M^\# = \langle D^\#, A_p^\# \rangle$ では、 $A_p^\#$ の割当ては R 上の 2 項述語がどれも同値律とみなすという特徴がある。このことは式(2)によって確かめられる。これに対し、勝

手なモデル $M = \langle D, A_p \rangle$ では、その A_p の割当てる 2 項述語は必ずしも同値律を満足しない。そこで一般のモデルのうち、それによつて割当たられる 2 項述語が同値律をみにすような特別なモデルが、式の意味として従属性を考察する上で重要であろうと推察でする。この考え方を実現するため (A 上の) 同値律集合と呼ばれる特別な文の集合を導入しておく。この集合を LE であらわすことにして、 LE はすべての $a \in A$ に対して次の形の（同値律をあらわすと解釈できます）特別な文の全体から成る。

- (i) $\forall x \, p_a(x, x)$,
- (ii) $\forall x \forall y \, (p_a(x, y) \rightarrow p_a(y, x))$,
- (iii) $\forall x \forall y \forall z \, (p_a(x, y) \wedge p_a(y, z) \rightarrow p_a(x, z))$.

5 特徴関係

LE をみたす任意のモデルから、特徴関係と呼ばれる関係が構成できることを示す。さらに特徴関係では a のモデルと丁度同じ文達が充足されることを示す。この特徴関係は、いわゆる Armstrong の関係 [5] といわれたものの拡張になつてゐる。

5.1 特徴関係の構成

LE を満たすモデルの任意の M と $M = \langle D, A_p \rangle$ とする。
 M の特徴関係 $R^* = \langle A, \mathcal{B}^*, R^* \rangle$ はつきの 4 段階で構成される。

(1) $D \times I$ の関数 $\mathcal{B}^*: A \rightarrow Sets$, s.t. $\mathcal{B}^*(a) = D/\pi_a$.

ここで $\pi_a = A_p(p_a)$ であり仮定によりこれは同値律を満たすといふ。また D/π_a は D の π_a による商集合をあらわす。

(2) 関係エ = バース $\mathcal{R}(A, \mathcal{B}^*) = \prod \mathcal{B}^*$.

(3) $\varphi: D \rightarrow \mathcal{R}$ なる特別な関数 φ s.t.

$$d \in D, a \in A \text{ に対して } \varphi(d)(a) = d/\pi_a, \quad (3)$$

ここで d/π_a は d の π_a による同値類をあらわす。

(4) $R^* \triangleq \varphi(D) \subset D$ の φ による像。

以上の \mathcal{B}^* と R^* とくっつてでてくる $R^* = \langle A, \mathcal{B}^*, R^* \rangle \in M$ の特徴関係といふ。(図 1)

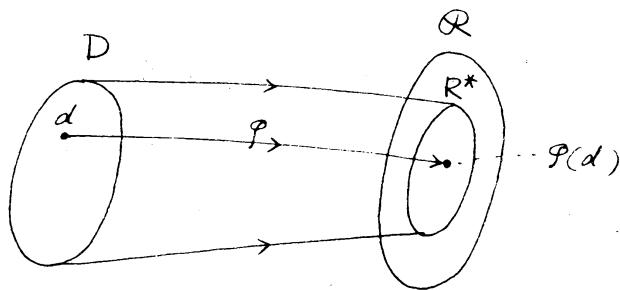


図 1 特徴関係

5.2 特徴関係の誘導モデル

モデル $M = \langle D, A_p \rangle$ が LE を満たすものとする。この M

の特徴関係 R^* を前項 5.1 のように定める。今その R^* の誘導するモデル M^* を $M^* = \langle D^*, A_p^* \rangle$ であらわそう。 M と M^* との対応、とくに A_p によって割当てられる 2 項述語と A_p^* によって割当てられるそれとの間の対応を調べよう。今 $p_a \in P$ に対して、 $A_p(p_a) = \pi_a$, $A_p^*(p_a) = \pi_a^*$ とおこう。

$\pi_a : D \times D \rightarrow 2$, $\pi_a^* : D^* \times D^* \rightarrow 2$ である。ここで、 M^* の定義から $D^* = R^*$ であり、 R^* の定義 $R^* = \varphi(D)$ から $\pi_a^* : \varphi(D) \times \varphi(D) \rightarrow 2$ となる。ただし、 φ は式(3)で与えられる。さて、任意の $d, c \in D$ に対して、

$$\begin{aligned} \pi_a^*(\varphi(d), \varphi(c)) = 1 &\iff \varphi(d)(a) = \varphi(c)(a) \\ &\iff d/\pi_a = c/\pi_a \\ &\iff \pi_a(d, c) = 1 \end{aligned}$$

である。すなはち、任意の $d, c \in D$ に対して、

$$\pi_a^*(\varphi(d), \varphi(c)) = 1 \iff \pi_a(d, c) = 1 \quad (4)$$

が成立する。(φ は全射であるから、任意の $\lambda, \mu \in R^*$ に対して、適当な $d, c \in D$ が存在し、 $\lambda = \varphi(d)$, $\mu = \varphi(c)$ となりかつ、そのような任意の d, c に対して

$$\pi_a^*(\lambda, \mu) = 1 \iff \pi_a(d, c)$$

である。) そこで、 φ によって引き起される $D \times D$ 上の述語から $D^* \times D^*$ 上のそれへの写像で、 $\pi_a \mapsto \pi_a^*$ となるものを φ であらわすよう拡大すると、 $\pi_a^* = \varphi(\pi_a)$ と書け、式(4)は、

$$\varphi(\pi_a)(\varphi(d), \varphi(c)) = \pi_a(d, c) \quad (5)$$

となる。この様子を図示すると図2のようになる。

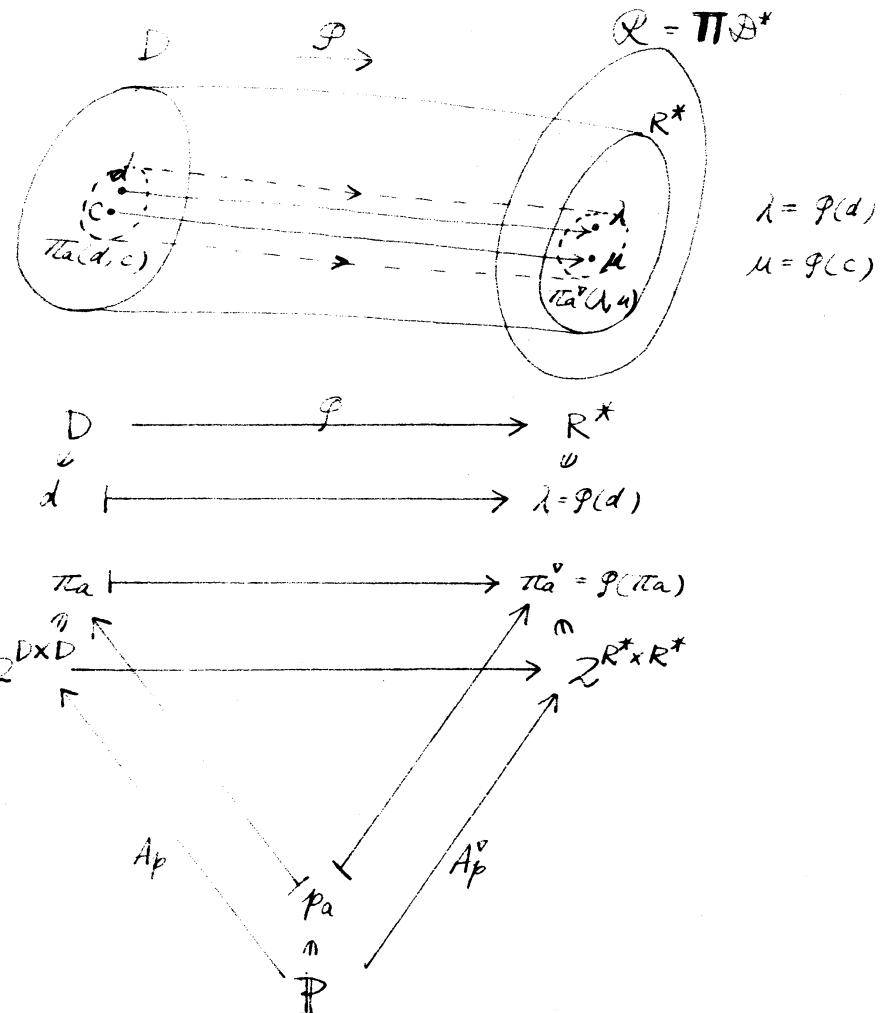


図2 M と M^* との対応

5.3 特徴関係の定理

M を LE を満たす任意のモデルとし、 R^* を 5.1 項で定めた特徴関係とする。このときつきの定理が成立つ。

[定理1] 任意の文 σ に対して, $M \text{ sat } \sigma$ と

$R^* \text{ sat } \sigma$ とは同等である。□

(証明) Au をモデル M 上の任意の割当てとする。 R^* 上の割当て Au^* をつぎのように定める。すなはち、任意の $\xi \in X$ に対して $Au^*(\xi) = \mathcal{P}(Au(\xi)) \in R^*$, ただし、 \mathcal{P} は式(3)で与えられる。このとき、定理は次の補助定理の特別な場合となる。

[補助定理] 任意の式 γ に対して, $M, Au \text{ sat } \gamma$ と $R^*, Au^* \text{ sat } \gamma$ とは同等である。□

補助定理を証明する。証明は、 $V_{Au}^{M^*}(\gamma) = V_{Au}^M(\gamma)$ であることを、 γ の構成に関する帰納法を用いて行なう。ここで M^* は R^* の誘導モデルで、5.2項で与えられるものとする。

(case 1) γ が原子式 $p_a(x, y)$ のとき

$$\begin{aligned} V_{Au}^{M^*}(p_a(x, y)) &= Ap(p_a)(Au^*(x), Au^*(y)) \\ &= \mathcal{P}(Ap(p_a))(\mathcal{P}(Au(x)), \mathcal{P}(Au(y))) \quad \cdots (*) \\ &= Ap(p_a)(Au(x), Au(y)) \quad \cdots \text{式(5)より.} \\ &= V_{Au}^M(p_a(x, y)). \end{aligned}$$

(*) 注. 5.2項 \mathcal{P} の拡大により, $Ap(p_a) = \pi_a^* = \mathcal{P}(\pi_a) = \mathcal{P}(Ap(p_a))$.

すなはち, Au^* の定義より $Au^*(x) = \mathcal{P}(Au(x))$.

(case 2) γ が式 $\sim \alpha$ のとき.

$$V_{A_n^*}^{M^*}(\sim \alpha) = \rightarrow V_{A_n^*}^{M^*}(\alpha)$$

$$= \rightarrow V_{A_n}^M(\alpha) \quad \cdots \text{帰納法の仮定より.}$$

$$= V_{A_n}^M(\sim \alpha).$$

(case 3) γ が式 $\alpha \wedge \beta$ のとき. (case 2) と同様.

(case 4) γ が式 $\forall x \alpha$ のとき.

$$V_{A_n^*}^{M^*}(\forall x \alpha) = \bigvee_{\lambda \in R^*} V_{A_n^*}^{M^*}[x/\lambda](\alpha)$$

$$= \bigvee_{d \in D} V_{A_n^*}^{M^*}[x/f(d)](\alpha) \quad \cdots R^* = \mathcal{P}(D) \text{ の元.}$$

$$= \bigvee_{d \in D} V_{(A_n[\exists/x])^*}^{M^*}(\alpha) \quad \cdots (*)$$

$$= \bigvee_{d \in D} V_{A_n}^M[x/d](\alpha) \quad \cdots \text{帰納法の仮定.}$$

$$= V_{A_n}^M(\forall x \alpha).$$

(case 1) から (case 4) まで補助定理の証明が完了した.

定理は 補助定理における式 γ が文 α である場合として既に得られる. (証明終) □

[系] M, R^* は定理 1 と同様とする.

任意の文の集合 Σ に対して、 $M \models \text{sat } \Sigma$ と $R^* \models \text{sat } \Sigma$ とは同等である。□

$$(*) \text{ 注: } A_n^*[x/f(d)](\xi) = \begin{cases} f(d) & \text{if } x = \xi \\ A_n^*(\xi) = f(A_n(\xi)) & \text{otherwise.} \end{cases}$$

$$\therefore A_n^*[x/f(d)](\xi) = f(A_n[x/d](\xi)) = (A_n[x/d])^*(\xi).$$

6 推論と形式化の完全性

α を任意の文、 Σ を任意の文の集合とする。

[定義] 関係での推論 (\Rightarrow)

α が Σ の関係モデルでの帰結であるとは、任意の関係
 $R = \langle A, \in, R \rangle$ 、ただし R は空でない、について

$R \text{ sat } \Sigma$ ならば $R \text{ sat } \alpha$ となることであり、

そのことを $\Sigma \Rightarrow \alpha$ と書く。□

[定義] モデルでの推論 (\vdash_{FOL})

α が Σ の意味論的帰結であるとは、任意のモデル M について
 $M \text{ sat } \Sigma^{\text{LE}}$ ならば $M \text{ sat } \alpha$ となることあり、それを $\Sigma \vdash_{\text{FOL}} \alpha$ と書く。

上記の 2 つの定義におけるそれそれの推論 \Rightarrow 、 \vdash_{FOL} と
 について互に同値であることが証明できる。

[定理 2] \Rightarrow に対する \vdash_{FOL} の健全性

$\Sigma \vdash_{\text{FOL}} \alpha$ ならば $\Sigma \Rightarrow \alpha$ 。□

[定理 3] \Rightarrow に対する \vdash_{FOL} の完全性

$\Sigma \Rightarrow \alpha$ ならば $\Sigma \vdash_{\text{FOL}} \alpha$ 。□

特徵関係の定理（定理 1）を用いて定理 3 が証明できる。
 また誘導モデルを用いて同様に定理 2 が証明できる。ここ
 では定理 3 の証明の方を上げるが定理 2 の証明もほぼ同様
 である。

(定理2の証明) 証明は帰理法による。そこで

$$\Sigma \Rightarrow \sigma \quad (6)$$

であるともかかわらず

$$\text{not}(\Sigma \vdash_{\text{WLE}} \sigma) \quad (7)$$

であると仮定する。式(7)より、

$$M \text{ sat } \Sigma^{\sim} \text{LE}^{\sim} \{\sim\sigma\}$$

となるモデルMが存在する。MはLEをみたすので、その特徴関係が構成でき、それは定理1によってMと同じ文集合をみたす。特に $\Sigma^{\sim} \{\sim\sigma\}$ をもみたすから、

$$R^* \text{ sat } \Sigma^{\sim} \{\sim\sigma\}$$

となる。これから

$$(R^* \text{ sat } \Sigma) \text{ and not}(R^* \text{ sat } \sigma)$$

となるが、これは式(6)のいう

任意の関係Rについて R sat Σ ならば R sat σ に反する。したがって仮定は反証され、したがって定理が証明された。□

7 あとがき

本稿で定義した従属性の文の集合は、従来の FD, JD グループ (JD, MD, MVD) よりもそれらの潜在形 (EJD, EMD, EMVD) あるいは Template 従属 (TD) [6] を含包し、十分大きな

表現力をもつていい。また、第6節において、関係モデルにおける従属性の推論が一階述語論理の意味論における推論によって等価におさかえられることがわかった。このことは、本稿の形式化による、2. 一階述語論理の形式的演算子体系が従属性の推論に対して自由に適用できることを示していい。したがって、先にあげたような従属性の推論には、少くとも部分決定アルゴリズムが存在することは直ちに了解される。さらにこれら従来の従属性はいずれもホーン節から由来されるので詳しく調べれば色々と興味ある性質をもつていい。竹島他の属性集合条件[7]などの解析が可能な点とともに、ホーン節の性質によつていい。

本形式化の鍵となる概念は第4節で定義した同値律集合LEであるが、カテゴリー論的接続を試みた研究[4]によつても、関係を特徴づけるものが、LEと同等のものがあることが指摘されている。この事実に示唆されて、関係に対する実世界のモデル化ということをあらためて見直すことも重要な点である。

(REFERENCES)

1. Codd, E. F. "A Relational Model of Data for Large Shared Data Banks," CACM, Vol.13, No.6, pp.377-387(1970).
2. Nicolas, J. M. "First Order Formalization for Functional, Multivalued, and Mutual Dependencies," Proc. ACM SIGMOD, Jul. 1978, pp.40-46.
3. Fagin, R. "Horn clauses and Database Dependencies," Proc. 12th ACM Symp. on Theory of Computing, Apr. 1980, pp.123-134.
4. 加藤昭彦 "射の導入による関係データモデルの表現独立性への接続と従属性への応用," 本講究録.
5. Armstrong, W. W. "Dependency Structures of Data Base Relationships," Proc. IFIP 74, pp.580-583, North Holland(1974).
6. Sadri, F. & Ullman, J. "A Complete Axiomatization for a Large Class of Dependencies in Relational Database," Proc. 12th ACM Symp. on Theory of Computing, Apr. 1980, pp.117-122.
7. 竹島卓, 国藤進, 小林要 "従属性の属性集合条件," 情報処理学会第21回全国大会予稿集 19-2 (1980)