数理解析研究所講究録
第 423 巻 1981 年 156-173

データベースの統合の代数的意味論

北海道大学工学部
田中 譲

1. はじめに

データベースの意味論を，データベースの属性名と，関係間の意味構造を抽象化した射の集合上に定義される代数構造を基に構築し，データベースの統合と検索処理の意味構造を代数構造を用いて明確に定義する。

ここで展開する理論は，先に報告した情報空間モデル(1~3)の改良であり，情報空間モデルを分散データベースの統合構造の記述や，複数概念の統合による概念の一般化（汎化），複数概念の共通概念の抽出（縮退），条件を課すことによる概念の制約等の構造の記述が可能になるように拡張している。

2. 準関係と抽象関係

Dを対象世界の可能な値の集合，Ωを有限集合とする。全関数$\mu: \Omega \to D$をΩ上のタプルと呼ぶ。Ω上のすべてのタプルの集合をD^{Ω}で表わす。$R \subset D^{\Omega}$なるRをΩ上の関係と呼び，ΩをRの属性集合と呼ぶ。Ωが空集合ϕのとき，D^{ϕ}を空タプルεのみからなる集合$\{\varepsilon\}$とし，これを$\mathbb{1}$で表わす。

Ωに対し、

$$r \subset \bigcup_{\Omega' \subset \Omega} D^{\Omega'}$$

なる集合rをΩ上の準関係と呼ぶ。

タプル$\mu : \Omega \to D$に対して、$\mu|_{\Omega'}$で関数μの$\Omega \cap \Omega'$への制限を表わす。Ω上の準関係rに対して、$[\Omega']r$, $\langle \Omega' \rangle r$を

$$[\Omega'] r = \{ \mu|_{\Omega'} \mid \mu \in r \}$$

$$\langle \Omega' \rangle r = \{ \mu|_{\Omega'} \mid \mu \in r,\ \mu|_{\Omega'} \text{は} \Omega' \text{上のタプル} \}$$

と定義する。rに対し、$\langle \Omega' \rangle r \neq \phi$なる$\Omega$の最大の部分集合$\Omega'$を$\Omega(r)$で表わし、$r$の属性集合と呼ぶ。

Ω_0を、対象とするデータベースのすべての準関係の属性集合Ωに対して$\Omega_0 \supset \Omega$なる集合とし、準関係rに対して、関数

$$\wedge r : 2^{\Omega_0} \to \bigcup_{\Omega' \subset \Omega_0} 2^{D^{\Omega'}}$$

を、

$$\wedge r = \lambda x . \langle x \rangle r$$

と定義する。$\wedge r$をrに対する抽象関係と呼ぶことにする。任意の$x \subset \Omega_0$に対して、$\wedge r(x)$は関係であるから、関係代数を適用し得る。

関係代数における1項関係演算pに対して，ω(p)でpに現われる属性の集合を表わすとして，p^rを，

$$p \wedge r = \lambda x.\ [x]\, p\ (\wedge r\,(x \cup \omega(p)))$$

と定義する。

抽象関係に対する2項関係演算として，和，差，積，直積，自然結合を

$$\wedge r + \wedge s = \lambda x.\ \wedge r(x) \cup \wedge s(x)$$

$$\wedge r - \wedge s = \lambda x.\ \wedge r(x) - \wedge s(x)$$

$$\wedge r \circ \wedge s = \lambda x.\ \wedge r(x) \wedge \wedge s(x)$$

$$\wedge r \times \wedge s = \lambda x.\ \wedge r(x \wedge \Omega(r)) \times \wedge s(x \wedge \Omega(s))$$

$$\wedge r * \wedge s = \lambda x.\ \wedge r(x \wedge \Omega(r)) * \wedge s(x \wedge \Omega(s))$$

と定義する。ただし，任意の関係Rに対して

$$R * \mathbb{1} = \mathbb{1} * R = R$$

とする。

準関係と抽象関係の定義を導入することにより，null値を含む関係の演算の意味が定義できたことになる。

3. 情報空間モデル

$\mathcal{R}$を対象データベースを構成している基本準関係の集合とし，条件

$$\forall r \in \mathcal{R}, \forall s \in \mathcal{R} \qquad \Omega(r) \cap \Omega(s) = \phi \ \text{iff} \ r \neq s$$

を満たすように属性名が付けられているものとする。

Ωを

$$\Omega = \cup_{r \in \mathcal{R}} \Omega(r)$$

と定義し，$^\wedge\omega$を

$$^\wedge\omega = \Sigma_{r \in \mathcal{R}} \ ^\wedge r$$

と定義する。$^\wedge\omega$は，関係間の意味構造をまったく考えない場合の，このデータベース全体の抽象関係表現になっている。$^\wedge\omega$のことをこのデータベースの基本世界と呼ぶ。

Lを$L \cap \Omega = \phi$なる可算集合とする。Lはラベル集合と呼ばれる。集合$\mathcal{A}$を

$$\mathcal{A} = \{ L \cup \Omega \text{ の要素の有限系列} \}$$

と定義する。長さ0の系列をϵで表わす。

$X \subset \mathcal{A}$に対する$\mathcal{A}$上の論理式$A(X)$とは，任意の

$$\mu \in D^X$$

に対して，

$$A(\mu(X))$$

が真理値をとり，真偽が決定可能であるような式であると定義する。

$\Omega \cup L$ の各要素 l に対して，$\mathcal{A}$ 上の論理式 $\alpha(l)$ を対応させる写像 α を考える。α は任意の $a \in \Omega$ に対して，$\alpha(a)$ が

$$a\,a = a$$

なる論理式になるものとする。$\alpha(l)$ を l に対する公理と呼ぶ。

$l \in \Omega \cup L$ に対し，$\mathcal{A}l$ で集合

$$\{a l \mid a \in \mathcal{A}\}$$

を表わす。$x \in \mathcal{A}l$ で $x = y l$ のとき，x/l を y と定義する。$\mathcal{A}$ 上の抽象関係 $\hat{}r$ に対し，$l/\hat{}r$ を $\mathcal{A}l$ 上の抽象関係で

$$\lambda x.\ (l/\hat{}r)(x l) = \hat{}r$$

となるものと定義する。

$\Omega \cup L$ の要素に対して，半順序 $l < l'$ を

$$l < l' \quad \text{iff} \quad (l \in L - \Omega) \wedge (l' \in \Omega)$$

と定義し，δ を $x \subset \mathcal{A}$ に対して，$x \cap \mathcal{A} l \neq \phi$ なる極小の $l \in \Omega \cup L$ を1意に1つ決める関数とする。

$\mathcal{A}$ 上の準関係 Δ_δ を，

$$
\begin{aligned}
{}^\wedge\Delta_\delta = \lambda x. \quad (\; & x \subset \Omega \cup \{\varepsilon\} \to {}^\wedge\omega(x - \{\varepsilon\}), \\
& x \subset \mathcal{A}\delta(x) \to (\delta(x)/{}^\wedge\Delta_\delta)(x/\delta(x) - \{\varepsilon\}), \\
& T \to [\alpha(\delta(x))]\Big({}^\wedge[\mathcal{A}\delta(x)]\Delta_\delta \\
& \qquad \times {}^\wedge[\mathcal{A} - \mathcal{A}\delta(x)]\dot{\Delta}_\delta \Big)(x))
\end{aligned}
$$

と定義する。

定理 3.1.

任意の $l \in \Omega \cup L$ に対して，$\alpha(l)$ を $X_l \subset \mathcal{A}$ 上の公理とするとき，$X_l \subset \Omega l \cup \Omega$ ならば，有限集合 $x \subset \mathcal{A}$ に対して ${}^\wedge\Delta(x)$ は計算可能であり，その値は δ に依存しない。このような l を基本的であるという。

証明は x に現われる系列の最大長に関する帰納法で証明で

きる。

定理 3.2.

$x \subset \mathcal{A}$, $y \cap \mathcal{A}l = \phi$ に対して、

$${}^\wedge\Delta(xl \cup y) \neq \phi$$

で、l は上述の定理の条件を満たすならば、基本準関係 $r_l, s_l \in \mathcal{R}$ が存在して、$X_l \subset \Omega(r_l)l \cup \Omega(s_l)$ である。

証明

$x \subset \Omega(r)$ となるような $r \in \mathcal{R}$ が存在しないとき、

$${}^\wedge\omega(x) = (\Sigma_{r \in R}\, {}^\wedge r)(x) = \phi$$

となることより明らか。

このような $L \cup \Omega$ の要素 l に対しては、r_l から s_l への射

$$\sigma_l : {}^\wedge r_l \to {}^\wedge s_l \quad ; \quad \alpha(l)$$

を考えることができる。特に各 $r \in \mathcal{R}$ の各属性 $a \in \Omega(r)$ に対して、射

$$\tilde{a} : {}^\wedge r \to {}^\wedge r \quad ; \quad aa = a$$

を考えることができる。

一般に，r，$s \in \mathcal{R}$ に対して，$\Omega(r)$ に対し，これと同形な $\Omega_\sigma(r)$ を定義し，$\Omega_\sigma(r) \cap \Omega = \phi$ で，$\Omega_\sigma(r) \cup \Omega(s)$ 上の論理式 A_σ を定義するとき，r から s への射を定義したという。この射 σ を

$$\sigma : {}^\wedge r \rightarrow {}^\wedge s \ ; \ A_\sigma$$

と記す。

Σ を射の有限集合とする。ρ を

$$\rho : \Sigma \rightarrow L \cup \Omega \quad , \quad {}^\forall a \in \Omega \quad \rho(\tilde{a}) = a$$

なる単射とする。Σ_0 を

$$\{ \tilde{a} \mid a \in \Omega \}$$

と定義する。$\Omega_\sigma(r) = \Omega(r)\rho(\sigma)$ とし，$\alpha(\rho(\sigma)) \equiv A_\sigma$ としたとき，σ をラベル $\rho(\sigma)$ で表現したという。

Σ を基本射の集合と呼ぶ。以下では，Σ に種々の演算を導入し，射の集合を拡大する。

Σ^* をある段階で定義された拡大された射の集合であるとする。射 $\sigma \in \Sigma^* - \Sigma_0$ に対して，

$$\sigma : {}^{\wedge}r \to {}^{\wedge}S \;;\; \alpha(\rho(\sigma))$$

とする．σの逆σ^{-}を次のような射と定義する．σ^{-}がΣを拡大していく途中で既に定義されているならばその定義に従う。そうでなければ$\rho(\sigma^{-})$を$L-\rho(\Sigma^{*})$より1つ選んで定義し，

$$\sigma^{-} : {}^{\wedge}S \to {}^{\wedge}r \;;\; \rho(\sigma)\backslash\alpha(\rho(\sigma))$$

と定義する。ただし，$\rho(\sigma)\backslash A$は，論理式A中の$X\rho(\sigma)$をXに変更し，$Y \cap \mathcal{A}\rho(\sigma)=\phi$なる$Y$は，$Y\rho(\sigma^{-})$にすることを意味する。$\alpha(\rho(\sigma^{-}))=\rho(\sigma)\backslash\alpha(\rho(\sigma))$と定義する．$\rho(\sigma^{-})$を$\rho(\sigma)^{-}$と表現する．

$a \in \Omega$に対して$\hat{a}^{-}=\hat{a}$と定義する．

$\Omega \cup L$の要素l，l'は有限個の$\Omega \cup L$の要素$l_1, l_2, \cdots, l_h$によって

$$\alpha(l) \quad \text{iff} \quad l_h \backslash l_{h-1} \backslash \cdots \backslash l_1 \backslash \alpha(l')$$

となるとき等価であるといい，$l \sim l'$で表わす．射σ，σ'が$\rho(\sigma) \sim \rho(\sigma')$となるとき，$\sigma$と$\sigma'$は表現が異なるだけで，同じ射である．このとき$\sigma=\sigma'$と書く．

定理 3.3.

$$(\sigma^-)^- = \sigma$$

証明

$$\alpha(\rho((\sigma^-)^-)) = \rho(\sigma^-) \backslash \rho(\sigma) \backslash \alpha(\rho(\sigma))$$

より，$\rho((\sigma^-)^-) \sim \rho(\sigma)$ である。よって $(\sigma^-)^- = \sigma$ である。

Σ が Σ^* まで拡大されたとする。σ, τ を

$$\sigma : s \to t \ ; \ \alpha(\rho(\sigma))$$

$$\tau : r \to s \ ; \ \alpha(\rho(\tau))$$

とするとき，合成 $\sigma \circ \tau$ を次のように定義する。$\sigma \circ \tau$ がすでに定義済みの場合はその定義に従う。そうでない場合は，

$$\sigma \circ \tau : r \to t \ ; \ \rho(\sigma^-) \backslash \alpha(\rho(\tau)) \wedge \alpha(\rho(\sigma))$$

と定義する。$\rho(\sigma \circ \tau)$ は $L - \rho(\Sigma^*)$ より 1 つ選んで定義する。$\alpha(\rho(\sigma \circ \tau))$ を

$$\alpha(\rho(\sigma \circ \tau)) = \rho(\sigma^-) \backslash \alpha(\rho(\tau)) \wedge \alpha(\rho(\sigma))$$

と定義する。$\rho(\sigma \circ \tau) \in L$ を $\rho(\sigma) \circ \rho(\tau)$ と表わす。

定理 3.4

$\ell \in L - \rho(\Sigma^*)$ に対し $\alpha(\ell)$ を恒真と定義する。

$x \subset \mathcal{A}$, $y \subset \mathcal{A} - \mathcal{A}\rho(\sigma)$ とすると,

$$\wedge\Delta(x\rho(\sigma\circ\tau)\vee y) = \wedge\Delta(x\rho(\tau)\rho(\sigma)\vee y)$$

が成立する。

証明省略

定理 3.5

$$(\sigma\circ\tau)^- = \tau^-\circ\sigma^-$$

証明

$\alpha(\rho((\sigma\circ\tau)^-))$ iff $\rho(\sigma\circ\tau)\backslash(\rho(\sigma^-)\backslash\alpha(\rho(\tau))\wedge\alpha(\rho(\sigma)))$

$\alpha(\rho(\tau^-\circ\sigma^-))$ iff $\rho(\tau)\backslash\alpha(\rho(\sigma^-))\wedge\alpha(\rho(\tau^-))$

iff $\rho(\tau)\backslash\rho(\sigma)\backslash\alpha(\rho(\sigma))\wedge\rho(\tau)\backslash\alpha(\rho(\tau))$

$\therefore\ \alpha(\rho((\sigma\circ\tau)^-))$

iff $\rho(\sigma\circ\tau)\backslash\rho(\sigma^-)\backslash\rho(\tau^-)\backslash\alpha(\rho(\tau^-\circ\sigma^-))$

Σ の拡大で逆と合成に関して閉じた集合を $\Sigma(^-,\circ)$ とする。$\ell \in L - \rho(\Sigma(^-,\circ))$ に対し, $\alpha(\ell)$ を恒真と定義する。σ が

基本射であれば σ^- も基本射であることと，定理3.4より，$x \subset \mathscr{A}$ なる有限集合 x に対して，$\wedge\triangle(x)$ は計算可能である。

4. 代数構造

$l, l' \in L \cup \Omega$ に対して，$\langle l/l' \rangle$ を l' の l への置き換えと定義する。

本章では，$\Sigma(-,\circ)$ に $\{+,-,*\}$ の演算を定義し，この演算に関して閉じた集合へと $\Sigma(-,\circ)$ を拡大する。

Σ が Σ^* まで拡大されたとして，σ，$\tau \in \Sigma^*$ の和，差，積を以下のように定義する。これらが以前に定義されているときは，その定義に従う。そうでなければ，$\rho(\sigma+\tau)$，$\rho(\sigma-\tau)$，$\rho(\sigma*\tau)$ を $L-\rho(\Sigma^*)$ の要素より別々の要素を1つづつ選んで，それらに等しいと定義する。$\alpha(\rho(\sigma+\tau))$，$\alpha(\rho(\sigma-\tau))$，$\alpha(\rho(\sigma*\tau))$ を以下のように定義する。

$$\alpha(\rho(\sigma+\tau)) \quad \text{iff} \quad \langle \rho(\sigma+\tau)/\rho(\sigma) \rangle \alpha(\rho(\sigma)) \vee \langle \rho(\sigma+\tau)/\rho(\tau) \rangle \alpha(\rho(\tau))$$

$$\alpha(\rho(\sigma-\tau)) \quad \text{iff} \quad \langle \rho(\sigma-\tau)/\rho(\sigma) \rangle \alpha(\rho(\sigma)) \wedge (\neg \langle \rho(\sigma-\tau)/\rho(\tau) \rangle \alpha(\rho(\tau)))$$

$$\alpha(\rho(\sigma*\tau)) \quad \text{iff} \quad \langle \rho(\sigma*\tau)/\rho(\sigma) \rangle \alpha(\rho(\sigma)) \wedge \langle \rho(\sigma*\tau)/\rho(\tau) \rangle \alpha(\rho(\tau))$$

Lの要素である $\rho(\sigma+\tau)$, $\rho(\sigma-\tau)$, $\rho(\sigma*\tau)$ を $\rho(\sigma)+\rho(\tau)$, $\rho(\sigma)-\rho(\tau)$, $\rho(\sigma)*\rho(\tau)$ と表わすことにする。

準関係 ${}^{\wedge}\gamma$ に対し，A，B を $\mathcal{A}$ 上の論理式とすると，

$$[A \vee B]\,{}^{\wedge}\gamma = [A]\,{}^{\wedge}\gamma + [B]\,{}^{\wedge}\gamma$$

$$[A \wedge \neg B]\,{}^{\wedge}\gamma = [A]\,{}^{\wedge}\gamma - [B]\,{}^{\wedge}\gamma$$

$$[A \wedge B]\,{}^{\wedge}\gamma = [A]\,{}^{\wedge}\gamma \circ [B]\,{}^{\wedge}\gamma = [A][B]\,{}^{\wedge}\gamma$$

と定義する。

$\Sigma(^{-},\circ)$ を $\{^{-},\circ,+,-,*\}$ に関して閉じるように拡大して得られる集合を $\Sigma(^{-},\circ,+,-,*)$ と定義する。$\ell \in L-\rho(\Sigma(^{-},\circ))$ に対し，$\alpha(\ell)$ を恒真としたときの ${}^{\wedge}\Delta$ を ${}^{\wedge}\Delta_2$，$\ell \in L-\rho(\Sigma(^{-},\circ,+,-,*))$ に対して $\alpha(\ell)$ を恒真としたときの ${}^{\wedge}\Delta$ を ${}^{\wedge}\Delta_5$ とする。有限集合 $x \subset \mathcal{A}$ に対して，${}^{\wedge}\Delta_2(x)$ が計算可能であれば，${}^{\wedge}\Delta_5(x)$ も計算可能である。

5. 代数構造と意味構造

本章では，代数構造 $\Sigma(^{-},\circ,+,-,*)$ を用い，あらかじめ定義された基本射の集合 Σ を用いて，データベースの意味構造を構築する過程を例を用いて説明する。

例.

基本準関係が

r (person_#, name, birth-date, sex, parent)

のみからなるデータベースを考察する。基本射としては以下のようなものがある。

Σ : self : ^r → ^r ; person_# person_# = person_#

same-name : ^r → ^r ; name name = name

same-birthdate : ^r → ^r ; birth-date birth-date = birth-date

same-sex : ^r → ^r ; sex sex = sex

same-parent : ^r → ^r ; parent parent = parent

male : ^r → ^r ; person# male = person# ^ sex male = 'male'

female : ^r → ^r ; person# female = person# ^ sex female = 'female'

parent : ^r → ^r ; person# parent' = parent

これらの基本射が、基本語彙となる。基本語彙と、{⁻, ∘, +, −, *}の演算を用いて語彙を構築する。⁻は逆の概念を、∘は概念の修飾を、+は概念の汎化、−は概念の差を、そして*は概念の縮退を表現する。

vocabulary building.

child = parent^{-}

sibling = same_parent - self

father = parent ∘ male

mother = parent ∘ female

son = child ∘ male

daughter = child ∘ female

brother = sibling ∘ male

sister = sibling ∘ female

twin = sibling * same_birth_date

etc.

新しい概念の登録は，それまでに定義した語彙を用いて随時に行うことができる。

検索は，これらの語彙を用いて行う。たとえば，双生児を求めるには，

$^{\wedge}\Delta_5$ (name name twin)

とすればよい。

これは以下のように自動的に展開され実行される。

$\wedge\Delta_5$ (name , name twin)

$= ([\alpha(\text{twin})] (\wedge[\mathcal{A}\,\text{twin}]\Delta_5 \times \wedge[\mathcal{A} - \mathcal{A}\,\text{twin}]\Delta_5)) (\text{name}, \text{name twin})$

($\wedge D = \wedge[\mathcal{A}\,\text{twin}]\Delta_5 \times \wedge[\mathcal{A} - \mathcal{A}\,\text{twin}]\Delta_5$ とおく)

$= ([<\text{twin}/\text{sibling}> \alpha(\text{sibling})] \wedge D$

$* \ [<\text{twin}/\text{birth\_date}> \alpha(\text{birth\_date})] \wedge D) (\text{name}, \text{name twin})$

$= (([<\text{twin}/\text{sibling}><\text{sibling}/\text{parent}> \text{parent}.\text{parent} = \text{parent}] \wedge D$

$- [<\text{twin}/\text{sibling}><\text{sibling}/\text{person\#}> \text{person\#}\ \text{person\#} = \text{person\#}] \wedge D)$

$* [<\text{twin}/\text{birth\_date}> \ \text{birth\_date}\ \text{birth\_date} = \text{birth\_date}] \wedge D)$

$(\text{name}, \text{name twin})$

$= (([<\text{twin}/\text{parent}> \ \text{parent}\ \text{parent} = \text{parent}] \wedge D$

$- [<\text{twin}/\text{person\#}> \ \text{person\#}\ \text{person\#} = \text{person\#}] \wedge D)$

$* [<\text{twin}/\text{birth\_date}> \ \text{birth\_date}\ \text{birth\_date} = \text{birth\_date}] \wedge D).$

$(\text{name}, \text{name twin})$

$= (R_1 - R_2) \wedge R_3$

$R_1 = [\text{name twin}, \text{name}]\, R_1'$

$R_1' = [\text{parent twin} = \text{parent}]$

$(\wedge\Delta_5(\text{parent twin}, \text{name twin}) \times \wedge\Delta_5(\text{name}, \text{parent}))$

$= ((\text{twin}/\wedge r)(\text{parent twin}, \text{name twin}))$

$[\text{parent twin} = \text{parent}] (\wedge r(\text{name}, \text{parent}))$

$R_2 = [\text{name twin}, \text{name}]\, R_2'$

$$R_2' = [\text{person\# twin} = \text{person\#}]\,(\hat{\ }\Delta_S(\text{name twin}, \text{person\# twin}) \times \hat{\ }\Delta_S(\text{name}, \text{person\#}))$$

$$= ((\text{twin}/\hat{\ }r)(\text{name twin}, \text{person\# twin}))\ [\text{person\# twin} = \text{person\#}]\,(\hat{\ }r(\text{name}, \text{person\#}))$$

$$R_3 = [\text{name twin}, \text{name}]\, R_3'$$

$$R_3' = [\text{birth\_date twin} = \text{birth\_date}]\,(\hat{\ }\Delta_S(\text{name twin}, \text{birth\_date twin}) \times \hat{\ }\Delta_S(\text{name}, \text{birth\_date}))$$

$$= ((\text{twin}/\hat{\ }r)(\text{name twin}, \text{birth\_date twin}))\ [\text{birth\_date twin} = \text{birth\_date}]\,(\hat{\ }r(\text{name}, \text{birth\_date}))$$

6. 結論

本論文では、基本射の集合Σより定義されるΣ(¯, ∘, +, −, *)に対し、$\hat{\ }\Delta_S$なる抽象関係を定義し、Σ(¯, ∘, +, −, *)の代数構造を用いて、データベースにおける種々の概念の構築が行えることを示した。今後の課題としては、Σ(¯, ∘, +, −, *)自身の構造の解明、限量記号の導入等が残されている。

参考文献.

(1). Y. Tanaka, "Information Space Model," Proc. Formal Bases for Data Bases, (Toulouse), 1979.

(2) 田中 譲, '情報空間モデルの表示的意味論,' 信学研資. EC80-29, 1980, pp 41-52.

(3) 田中 譲, "関係データベースの設計と意味に関する理論的アプローチ," 信学研資, AL80-51, PRL80-60, 1980, pp55-64.