

## 履歴オブジェクトに基づく意味的データモデル

田中 克己  
(神戸大学・工学部)

データベースの意味表現能力を高めるために汎化・集約・関数型関連等を用いたオブジェクト指向型の意味的データベースモデルが多く提案され注目されている。又、事物の履歴や時間情報を表現するための履歴データベースモデルの研究も盛んになりつつある。本報告では、これらの意味的データベースモデルで履歴情報を表現する方法について考察している。具体的には、時間をオブジェクトに写像する「履歴オブジェクト」と、オブジェクトを履歴オブジェクトに写像する「履歴型関数」を導入し、これにより単なる履歴情報の表現のみならず、履歴情報の汎化や集約等の扱いが可能となる。また、このような場合に必要となる、履歴オブジェクトに関する基本的な意味制約についても述べる。

### A Semantic Data Model Based on History Objects

Katsumi Tanaka  
Faculty of Engineering, Kobe University

Recently, much attention has been focused on object-based "semantic database models", which capture more meaning of data by the constructs "generalization", "aggregation" and "functional relationships". Also, the study of historical database models is another active field of database research. Historical databases aim to represent and manipulate historical or temporal information. This paper discusses a method to represent historical information on general semantic database models. For this purpose, we introduce a new type of objects called "history object", which map a time to an object. Also, a "history-type" function is introduced, which is a mapping from objects to historical objects. Some fundamental semantic constraints induced by historical objects are also investigated.

## 1. まえがき

データベースのデータの意味表現能力を高めることを目的として、最近種々の「意味的データベースモデル」が提案されている<sup>[1,2]</sup>。これらはデータベーススキーマを、objectやentityと呼ばれる概念とこれらの間の関連を汎化・集約・関数という形で表現しようというものである。また、一方、データベースの応用分野の拡大や利用の高度化にともない、種々の分野において履歴情報の扱いの重要性が指摘されている<sup>[1,3,6]</sup>。

本稿では、意味的データモデルに基づいて履歴情報を表現するための方法について考察する。具体的には、汎化・集約・関数型の関連を扱える標準的な意味的データモデルの1つであるIFOデータモデル<sup>[4]</sup>をもとにして、「履歴オブジェクト型」、「履歴型関数」をさらに導入して、履歴情報の表現を試みる。本稿のアプローチは、Cliffordらの「属性レベルの時刻印(timestamp)付け」のアイデアをオブジェクト指向のデータベースモデル上に導入したものと位置づけられる。

## 2. 準備

Abiteboul, HullのIFOデータモデル<sup>[4]</sup>においては、「オブジェクト型」はPERSONやNAME等のラベルが付され、これらは次の6種類の型に細分類されている。

- (1) 印字可能オブジェクト型 (□で表現)
- (2) 抽象的オブジェクト型 (◇で表現)
- (3) 導出オブジェクト型 (○で表現)
- (4) 集合オブジェクト型 (⊗で表現)
- (5) 集約オブジェクト型 (⊗で表現)
- (6) 直和オブジェクト型 (⊕で表現)

オブジェクト型0のとりうるオブジェクトの集合を0のドメインと呼び、 $Dom(0)$ と表す。上記の(1), (2)にはそれぞれ印字可能なオブジェクトの集合、抽象的なオブジェクトの集合がドメインとして対応する。また、(3), (4), (5), (6)のオブジェクト型はそれぞれ他のいくつかのオブジェクト型を用いて定義される。(3)の導出オブジェクト型はそれぞれ他のいくつかのオブジェクト型との間の汎化関係(is\_a関係)によって定義される。従って、導出オブジェクト型のドメインは汎化関係によって結ばれた上位のオブジェクト型のドメインがこれに対応する。汎化関係は、IFOモデルにおいては、 $A \text{ is\_a } B$ という関係がある時はAからBへ至る矢印( $\rightarrow$ 又は $\Rightarrow$ )で表す。(4), (5), (6)のオブジェクト型は、そのドメインが、下位のオブジェクト型のドメインの有限部分集合族、直積、直和となるようなオブジェクト型である。オブジェクト型0に対するインスタンスとは $Dom(0)$ の有限部分集合のことをいう。IFOデータモデルでは、上記に加えて、オブジェクト型間の関数的関連を定義できる。これにより、オブジェクトの種々の性質を表現できる。

## 3. 履歴オブジェクト型

本節では、前節で述べた各種のオブジェクト型に加えて、履歴情報を表現するための新しいオブジェクト型として、「履歴オブジェクト型」を導入し、その性質について述べる。以後、本稿では、TIMEを離散的時刻の可算無限集合とし、線形順序が定義されているものとする。1つの「履歴オブジェクト」は直観的には、

オブジェクトに関する履歴そのものを表すオブジェクトである。我々のモデルでは、履歴は時間 (TIME) をオブジェクトに写像する関数として表される。「履歴オブジェクト型」は、共通の性質を有する履歴オブジェクトのクラスを指定するものである。より形式的には、

[定義] 与えられたオブジェクト型  $O$  に対して、その履歴オブジェクト型  $O^*$  は、そのドメインが

$$\text{Dom}(O^*) = \{ o^* \mid o^* \text{ は TIME から } \text{Dom}(O) \text{ への部分関数} \}$$

となるようなオブジェクト型である。履歴オブジェクト型  $O^*$  のインスタンスは、 $\text{Dom}(O^*)$  の有限の部分集合である。

履歴オブジェクト型  $O^*$  の 1 つのインスタンスは、上の定義から、関数の有限集合であり、この要素となる各々の関数が時間を (オブジェクト型  $O$  の) オブジェクトに写像するものである。

[例] 下図のように、抽象的オブジェクト型 STUDENT に対して、その履歴オブジェクト型 STUDENT<sup>\*</sup> を考える。



図3.1 抽象的オブジェクト型とその履歴オブジェクト型

オブジェクト型 STUDENT の 1 つのインスタンス  $s$  は、ここでは例えば、学生オブジェクトの集合  $\{s_1, s_2\}$  である。STUDENT のインスタンスは時々刻々その内容が変化するが、その意味は、各時点において存在する学生オブジェクトの集合を表すものと解釈できる。これに対し、例えば STUDENT<sup>\*</sup> のインスタンスを

$$s^* = \{s_1^*, s_2^*, s_3^*\}$$

とする。ただし、

$$s_1^*(t) = s_1 \dots \dots \dots 1975 \leq t \leq \text{now} \quad s_3^*(t) = s_3 \dots \dots \dots 1975 \leq t < 1980$$

未定義...上記以外

未定義...上記以外

$$s_2^*(t) = s_2 \dots \dots \dots 1978 \leq t < 1987$$

未定義...上記以外

とする。このインスタンス  $s^*$  は、現在迄に 3 人の学生オブジェクトが存在したことを示し、さらにこのうち学生オブジェクト  $s_3$  は現在は存在していないことを表している。すなわち、オブジェクト型 STUDENT は、「現在の」学生オブジェクト集合を表し、履歴オブジェクト型 STUDENT<sup>\*</sup> は、現在の時点で登録されている、これまでに出現した学生オブジェクト及び一部将来にも存在するとされている学生オブジェクトからなる集合を表している。明らかに、インスタンス  $s$  とインスタンス  $s^*$  との間には

$$s = \{s_i(\text{now}) \mid s^* \text{ は STUDENT}^* \text{ のインスタンス、} s_i \in s^*\}$$

という関係がある。

履歴データベースでは、格納されている過去の全ての情報よりも、ある特定の時刻あるいは特定の時間区間の内容に対して検索などの操作を行いたい場合が多い。これは、履歴オブジェクトである関数の定義域の制限という形で与えられる。

[定義] 履歴オブジェクト型  $0^*$  の履歴オブジェクト  $oi^*$  の時刻  $t$  (又は時間区間  $I$ ) 上への射影は、関数  $oi^*$  の定義域を  $\{t\}$  (又は  $I$ ) に制限したもので、これを  $oi^*[t]$  (又は  $oi^*[I]$ )

と表す。履歴オブジェクト型  $0^*$  のインスタンス  $o^*$  の時刻  $t$  (又は時間区間  $I$ ) 上への射影は、 $o^*[t]$  (又は  $o^*[I]$ ) と表し、

$$o^*[t] = \{oi^*[t] \mid oi^* \in o^*\}$$

$$o^*[I] = \{oi^*[I] \mid oi^* \in o^*\}$$

と定義する。

[例] 前述の履歴オブジェクト型 STUDENT $^*$  のインスタンス  $s^*$  に対して、 $s^*$  の1981年への射影は

$$s^*[1981] = \{s1^*[1981], s2^*[1981]\}$$

となる。また、 $s^*$  の時間区間 [1975, now] 上への射影は

$$s^*[[1975, now]] = \{s1^*[[1975, now]], s2^*[[1975, now]]\}$$

となる。

履歴オブジェクト型  $0^*$  に対し、そのインスタンスの時刻  $t$  又は時間区間  $I$  上への射影インスタンスにのみ興味がある場合は、このための履歴オブジェクト型を一種のビュー定義の形で与えることが考えられる。

[定義] 履歴オブジェクト型  $0^*$  の時刻  $t$  (又は時間区間  $I$ ) への射影履歴オブジェクト型を  $0^*[t]$  (又は  $0^*[I]$ ) で表す。 $0^*[t]$  (又は  $0^*[I]$ ) のインスタンスは  $0^*$  のインスタンスの各要素の定義域を時刻  $t$  (又は時間区間  $I$ ) 上に制限したものである。

次に、履歴オブジェクト型に付随して定義される関数について述べる。

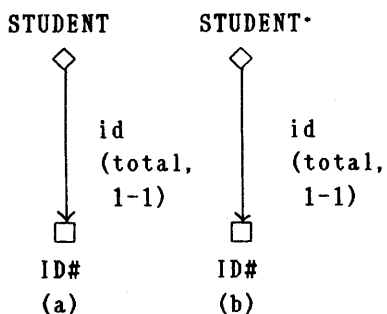


図3.2 履歴オブジェクト型と関数

図3.2(a)は現在の各学生オブジェクトにユニークな学籍番号 (ID#) を与える関数  $id$  を、図3.2(b)では、これまでに出現した各学生オブジェクトにユニークな学籍番号を与える関数として  $id$  を与えている。図3.2(a)では過去に存在して現在は存在しない学生オブジェクトの学籍番号が別の現存する学生オブジェクトに使用される可能性がある。一方、図3.2(b)では、各学生履歴オブジェクトに対して一意的に ID# が与えられるので、過去・現在・将来を通じて同じ ID# が複数の学生に与えられることはない。これ

は、履歴オブジェクト型とそれに付された関数によってimplicitに定義された一種の構造的な意味制約であると考えることができる。

#### 4. 履歴型関数

履歴オブジェクト型は、オブジェクトの履歴自身を表すオブジェクトを定義するものである。一方、各オブジェクトにはその性質を表す情報が関数型関連という形で付随している。オブジェクトの性質も時間的に変化するので、オブジェクトの性質の変化の履歴を表現するために、ここでは「履歴型関数」を導入する。

[定義] 与えられた2つのオブジェクト型 $O_1, O_2$ 及び $\text{Dom}(O_1)$ から $\text{Dom}(O_2)$ への関数 $f$ に対し、 $f$ の履歴型関数を、 $\text{Dom}(O_1)$ から $\text{Dom}(O_2)$ への関数とし、 $f^*$ で表す。

[例] 抽象的オブジェクト型PERSON、印字可能オブジェクト型PERSON-NAMEの間に図4.1に示すように、PERSONオブジェクトの名前履歴を与える履歴型関数 $\text{name}^*$ を考える。

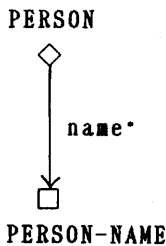


図4.1 履歴型関数  $\text{name}^*$

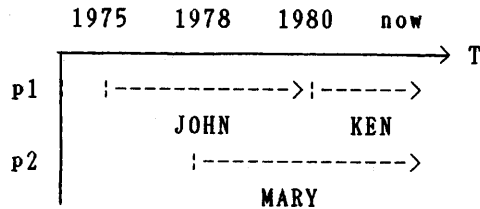


図4.2  $\text{name}^*$ の例

ここで $\text{name}^*$ は $\text{Dom}(\text{PERSON})$ から{TIMEから $\text{Dom}(\text{PERSON-NAME})$ への部分関数}への関数である。例えば、PERSONのインスタンスを $p = \{p_1, p_2\}$ とし、PERSONオブジェクト $p_1$ の名前が時間区間 $[1975, 1980)$ の間が"JOHN",  $[1980, \text{now}]$ の間が"KEN", それ以外が不明(未定義)、 $p_2$ の名前が $[1978, \text{now}]$ の間が"MARY"とすると、

$$\begin{aligned} (\text{name}^*(p_1))(1977) &= \text{"JOHN"}, & (\text{name}^*(p_1))(1980) &= \text{"KEN"} \\ (\text{name}^*(p_1))(1974) & \text{未定義}, & (\text{name}^*(p_2))(\text{now}) &= \text{"MARY"} \end{aligned}$$

となる(図4.2参照)。ここで以下のことに注意を要する。

- (1)  $\text{name}^*$ は現在存在するオブジェクト $p_1, p_2$ にのみ名前履歴を与える。
- (2) 例えば $(\text{name}^*(p_1))(1974)$ や $(\text{name}^*(p_2))(1976)$ は未定義であるが、これは $p_1$ や $p_2$ の名前がこの当時未定義なのであって、 $p_1$ や $p_2$ オブジェクトがこの当時存在したかどうかは不明である。

次に履歴オブジェクト型と履歴型関数の例を示す。図4.1のPERSONを履歴オブジェクト型 $\text{PERSON}^*$ に置き換えると図4.3を得る。

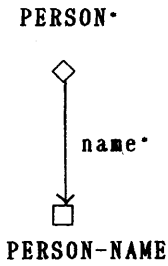


図4.3 履歴オブジェクト型と履歴型関数

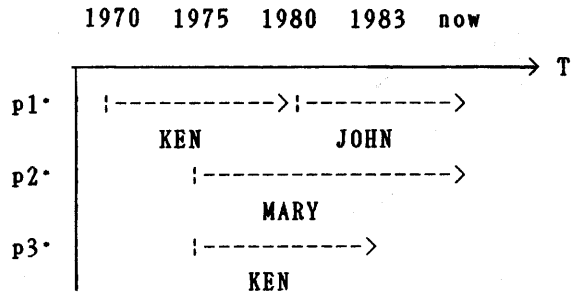


図4.4 pi\*とname\*の値

PERSON\*のインスタンスを  $p^* = \{p1^*, p2^*, p3^*\}$  とし、図4.4のように name\* の値を与えると、 $(name^*(p1^*))(1975) = "KEN"$  である。ここで、履歴オブジェクト  $p1^*$  は例えば  $p1^*(1965)$  は未定義である。そこで、“KEN”（後に“JOHN”）という名前を持った人が1965年当時存在しないとすれば、その当時の名前を与えるのは意味がない。これは履歴オブジェクトに関する存在従属性という意味制約と捉えることができる。すなわち、

$$\forall p_i^*, \forall t \left( (\exists n) (n \in \text{Dom}(\text{PERSON-NAME}) \ \& \ (name^*(p_i^*))(t) = n) \right. \\ \left. \rightarrow (\exists p) (p \in \text{Dom}(\text{PERSON}) \ \& \ p_i^*(t) = p) \right)$$

という形の意味制約という形で表現できる。

## 5. 履歴オブジェクト型の汎化と集約

本節では履歴オブジェクト型の汎化 (is\_a関係)、集約 (⊗で示される part\_of関係、⊕で示される要素-集合関係) について述べる。

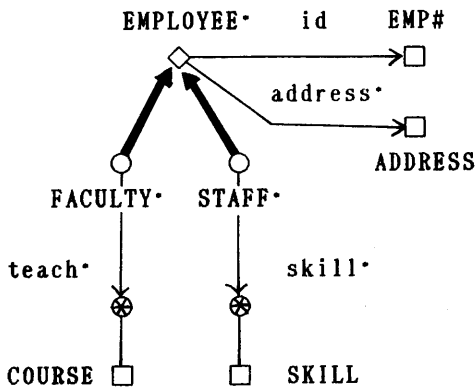


図5.1 履歴オブジェクト型間の汎化

図5.1では、FACULTY\*、STAFF\*という履歴オブジェクト型の汎化がEMPLOYEE\*であることを示している。このスキーマのインスタンスは従業員オブジェクトの履歴及びその従業員番号と住所履歴、FACULTYオブジェクトの履歴及び講義していた科目集合の履歴、STAFFオブジェクトの履歴及びSTAFFの有する技能集合の履歴を表現している。

履歴オブジェクト型  $A^*$ 、 $B^*$  に対して、いま  $A^* \rightarrow B^*$  なる汎化関係が定義されているとする。通常のオブジェクト型間の汎化関係は、下位概念のオブジェクト型のインスタンスに対してその各要素を自分自身に写像する identity function を定義する。履歴オブジェクト型間の汎化関係は、これと同じ意味で定義できるが、この

ような汎化関係に加えて、より弱い意味の汎化関係も考えることができる。ここでは、前者の汎化関係を「強汎化」、後者を「弱汎化」と呼び、以下のように定義する。

**強汎化**：Aのインスタンスaの任意の要素aiに対して、 $f(ai)=ai$ なる関数（identity function）を与える。Bのインスタンスをbとすると、aはbの部分集合である。

**弱汎化**：Aのインスタンスaの任意の要素aiに対して、 $f(ai)=bi$ なる関数（fはtotal, 1-1）を与える。ただし、

$$\forall o, \forall t (o \in \text{Dom}(A)=\text{Dom}(B) \ \& \ ai(t)=o \rightarrow bi(t)=o)$$

である。

後者の弱汎化は、例えば、上の例である従業員がある期間だけFACULTYでもSTAFFでもないというような場合に利用できる。また、明らかに強汎化関係は弱汎化関係である。

次に、図5.2(a), (b)に履歴オブジェクト型のpart\_of関係, 要素-集合関係による集約の例を示す。

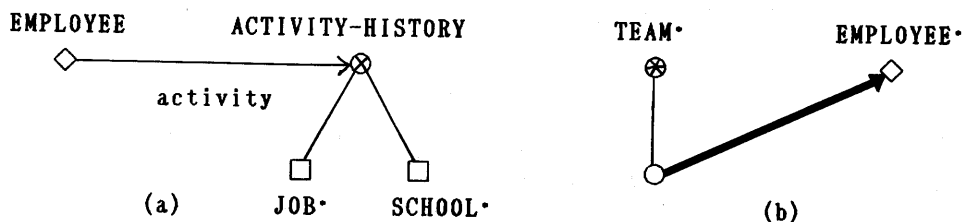


図5.2 履歴オブジェクト型の集約

ここで(a)のACTIVITY-HISTORYという集約オブジェクト型は2つの履歴オブジェクト型を集約したものであり、JOBやSCHOOLは職歴と学歴（学校名履歴）を表す。また、(b)の集合オブジェクト型TEAMは、各チームの構成人員に関する履歴を表している。

## 6. 履歴オブジェクト型の基本的意味制約と、オブジェクトの合併・分裂

これまで、一部（図5.2(a)の例）を除いては、各履歴オブジェクトは時間を同一のオブジェクトに写像するという暗黙の仮定を置いてきた。また、各履歴オブジェクトは相異なるオブジェクトに写像するという仮定も暗黙にしている。これらは履歴オブジェクト型に関する基本的意味制約であると考えられる。以下にこれらをまとめる。

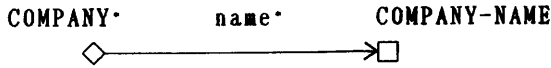
(1) 履歴オブジェクトの写像の一意性

$$\forall t1, \forall t2 (o(t1)=o1 \ \& \ o(t2)=o2 \rightarrow o1=o2)$$

(2) 各履歴オブジェクトは相異なるオブジェクトに写像

$$\forall o1, \forall o2, \forall t (o1(t)=o2(t) \rightarrow o1=o2)$$

ここで上記の(1), (2)を満たさない場合を考えると、上林<sup>[5]</sup>にその重要性が指摘されている「オブジェクトの合併・分裂の履歴」という概念が表現できる。例えば



においてCOMPANY\*が(1), (2)を満たさないとする、このスキーマは会社の合併・分裂の履歴と社名履歴の両方を表現することになる。例えばCOMPANYオブジェクトc1とc2が1980年に合併しc3になり、c1の名前が1970年にx、1975年にy、c2の1970年当時の名前がz、c3の名前がwとする。これは次のように表せる。

$$c1^*, c2^* \in \text{Dom}(\text{COMPANY}^*), c1, c2, c3 \in \text{Dom}(\text{COMPANY})$$

$$c1^*(1970)=c1, c1^*(1980)=c3, c2^*(1970)=c2, c2^*(1980)=c3,$$

$$(\text{name}^*(c1^*))(1970)=x, (\text{name}^*(c1^*))(1975)=y,$$

$$(\text{name}^*(c1^*))(1980)=w, (\text{name}^*(c2^*))(1970)=z, (\text{name}^*(c2^*))(1980)=w$$

ここで各COMPANY\*オブジェクトは各時点において一意的な社名を持つという次のような意味制約を課す必要が有ることに注意を要す。

$$\forall c1^*, \forall c2^*, \forall t (c1^*(t)=c2^*(t) \rightarrow (\text{name}^*(c1^*))(t)=(\text{name}^*(c2^*))(t))$$

## 7. あとがき

本稿では、履歴オブジェクト型と履歴型関数という概念を導入することで既存の意味的データモデルにおいて履歴情報を表現する方法について述べた。意味的データモデルが提供している汎化・集約等の概念を履歴情報に対しても適用できる点が本稿のアプローチの大きな特色である。又、履歴オブジェクト型に関する基本的な意味制約についても述べた。今後は、操作言語、意味制約の種々の性質等にさらに検討を加える必要がある。

## 参考文献

- [1] 田中、「データモデルの高水準化とその応用」、データ組織化の基礎に関する国際会議Tutorial資料、1985年5月
- [2] R.Hull, "A Survey of Research on Semantic Database Models", データ組織化の基礎に関する国際会議Tutorial資料、1985年5月
- [3] J.Clifford and A.U.Tansel, "On An Algebra for Historical Relational Databases: Two Views", Proc. ACM-SIGMOD, May 1985.
- [4] S.Abiteboul and R.Hull, "IFO: A Formal Semantic Database Model", Proc. PODS, April 1984.
- [5] 上林、「Document Management System」、IBMコンピュータサイエンスシンポジウム、1985年10月
- [6] S.Ginsburg and K.Tanaka, "Computation-tuple Sequences and Object Histories", To Appear in ACM TODS, June 1986.