

# 名前付き集合モデルを用いたDREAMモデルの定義

中田 充<sup>†</sup> 宝珍輝尚<sup>††</sup> 都司達夫<sup>††</sup>  
<sup>†</sup> 福井大学大学院工学研究科      <sup>††</sup> 福井大学情報工学科

サイエンティフィックデータベース管理システム DREAM のデータモデルである DREAM モデルでは、データはデータエレメントと呼ばれる要素に格納される。名前付きエレメントはデータに名前を付けるための要素であり、視点はその集合として定義される。さらに、オブジェクトは視点の集合であり、オブジェクトの集合はバンドルと呼ばれる。この内、名前付きエレメント、視点、オブジェクト、バンドルの基本的な構造は(識別子, 名前, 集合)の3つ組である。この3つ組のことを名前付き集合と呼ぶ。本論文では、まず名前付き集合に関してその構造や演算を定義し、名前付き集合を用いた名前付き集合モデルを提案する。そのあと、名前付き集合モデルを用いて DREAM のデータモデルの定義を行う。

## Definition of DREAM Model Using Named Set Model

Mitsuru NAKATA<sup>†</sup> Teruhisa HOCHIN<sup>††</sup> Tatsuo TSUJI<sup>††</sup>  
<sup>†</sup> Graduate School of Engineering, Fukui University  
<sup>††</sup> Department of Information Science, Fukui University

DREAM model is a data model for a scientific DBMS *DREAM*. In DREAM model, data is stored into a *data element*. A *named element* is to relate a name with data elements. A set of named elements is called *perspective*. Furthermore, an *object* is a set of perspectives, and a set of objects is called a *bundle*. Structure of every DREAM element except data element can be treated as a triplet (*id*, *name*, *S*), where *id* is the identifier of the triplet, *name* is the name of the triplet and *S* is the corresponding set of DREAM elements. We call this triplet as the *named set* and the data model employing these triplets as the *named set model*. First, this paper proposes a named set model including unified operations on the triplets. Secondly, it defines the DREAM model using the proposed named set model.

## 1 はじめに

データベースの応用分野のうち、データベース化が求められているにもかかわらず、従来のデータベース管理システムではデータベース化が困難な応用分野がある。実験データや測定データなどの科学技術データを扱う応用分野はその一つである。この分野では、得られたデータの研究者の視点や仮定に沿った整理分類、新たな仮説やデータ間に隠れている事実を導き出すための検討分析を円滑に行うためにデータベースが必要とされる。以降では、このような分野のデータをサイエンティフィックデータと呼び、それを格納するデータベースをサイエンティフィックデータベースと呼ぶ。サイエンティフィックデータはデータベース化を前提に生成されておらず、(1)一つのデータに対する複数の視点や考え方の表現やそれらの間でのデータ共有が可能であること、(2)データ格納前にユーザがスキーマを定義しなくてもよいこと、ならびに、(3)試行錯誤を行いながらデータベースをボトムアップに構築できることが要求される。

我々は、これらの要求を満たすサイエンティフィックデータのためのデータモデル (DREAMモデル) を提案し、それをを用いたデータベース管理システムDREAMを設計制作してきている<sup>79)</sup>。

DREAMモデルでは、データはデータエレメント (以下、de) に格納される。データエレメントは、組(id, value)である。ここで、idは識別子、valueはデータ値である。データエレメントは、オブジェクトなどの他の要素から独立して存在できる。名前付きエレメント (以下、ne) は、データエレメントに名前を付ける要素であり、3つ組(id, name, {de})である。対象物の一つの側面を表す視点 (以下、pe) は、(id, name, {ne})であり、一つの対象物を表すオブジェクト (以下、obj) は、(id, name, {pe})である。オブジェクトの集合は、DREAMモデルではバンドルと呼ばれ、3つ組(id, name, {obj})である。これらのうち、名前付きエレメント、視点、オブジェクト、バンドルは、(識別子, 名前, 集合) の3つ組という非常によく似た構造であるにもかかわらず若干の違いがあるため、DREAM

モデルを実現する場合には、それぞれの要素に対して同じ様なプログラムを複数作成しなければならず、煩雑になるという問題がある。また、演算の種類も限定されてしまう。

そこで本論文では、DREAMモデルの実現をより容易とし、かつ、強力な演算能力を持たせることを目的として、核となるデータモデルを明確化し、これを用いてDREAMモデルを再定義する。ここでは、名前付きエレメント、視点、オブジェクト、バンドルの基本となっている3つ組 (識別子, 名前, 集合) をもとにする。この3つ組を“名前付き集合”と呼ぶ。これは、集合が名前を持つことが特徴であり、これにより集合演算を用いた新しい演算が必要となる。本論文では、この演算を提案する。その後、名前付き集合モデルを用いてDREAMのデータモデルの定義を行う。

以降、2章で本論文の前提として、サイエンティフィックデータベース管理システムDREAMの概要と比較演算について述べる。3章で、名前付き集合モデルについて述べた後、4章で名前付き集合モデルを用いてDREAMモデルの定義を行う。最後に5章でまとめる。

## 2 前提

### 2.1 DREAMとDREAMモデル

#### 2.1.1 DREAMの概要

サイエンティフィックデータには大きく分けて2種類のデータがある。一つは実験、測定などによって得られた未整理データ (一次データ) であり、もう一つは、十分な検討や試行錯誤のもとで、一次データから必要とされるデータを同定するなどして研究者が作成したデータ (二次データ) である。実験や測定によって得られた一次データは、データベース化を考慮したデータではないためそれが難しい。そこで、一次データをとりあえず格納し、あとから整理を行いながらデータベースを構築するといった手順が考えられる。そのために、大量の未整理データをデータ型や属性名といったスキーマを意識することなく格納できる機構と二次データを試行錯誤を繰り返しながら作成・管理

できる機構が必要とされる。このため、DREAMは2種類のデータベースを管理する。一つは、一次データを格納するための「基本データベース」であり、もう一つは、二次データを作成、管理するための「導出データベース」である。本論文では、導出データベースのデータモデルであるDREAMモデルについてのみふれる。

### 2.1.2 DREAMモデルの概要

本論文の対象である導出データベースは、データエレメントの集合である。データエレメントは、データ値を保持する以外に、基本データベース中の一次データの一部または全体を参照する。名前付きエレメントは、データエレメントとオブジェクトの集合に名前を付けるための要素であり、その構造は(識別子, 名前, データエレメントまたはオブジェクトの集合)の3つ組である。導出データベースにおいて、一つの意味のある情報単位であるオブジェクトは視点の集合であり、その構造は(識別子, 名前, 視点の集合)の3つ組である。視点は名前付きエレメントの集合に名前を付けたものである。導出データベースでは、オブジェクトの集合はバンドルと呼ばれる。以降では、データエレメント, 名前付きエレメント, オブジェクト, バンドルなどの要素を総称してデータベースエレメントと呼ぶ。

DREAMモデルは、1)データがオブジェクトなどから独立して存在する、2)モデルの構成子の構造に集合を採用している、3)データがモデルの主体でありオブジェクトなどはデータに対する一種のビューと考えられる、といった特徴を持つ。これらの特徴により、導出データベースの構造は容易に変更でき、他に影響を及ぼさない要素の変更や柔軟なデータの共有が可能となっている。また、一つの対象物の多様な側面なども視点を用いて素直に表現することが可能であり、視点間でのデータ共有も可能である。

## 2.2 比較演算

### 2.2.1 $\theta$ -比較と $\sigma$ -比較

$\theta$ は値に関する比較演算子であり、具体的には、

$<, \leq, =, \geq, >, \neq$ のいずれかである。 $x \theta y$ が成立するには、 $\text{dom}(x) = \text{dom}(y)$ でなければならない。ここで、 $\text{dom}(x)$ は $x$ の定義域を表す。 $x, y$ は数値とは限らず文字列なども考えられる。 $\sigma$ は値と集合に関する比較演算子で $\in$ または $\notin$ である。例えば、 $x \sigma Y$ は、 $x$ が集合 $Y$ の要素であれば真である。

### 2.2.2 集合間の比較

集合 $S$ と集合 $R$ の間の比較演算として、通常の比較演算(式(1))がある。

$$S \lambda R \quad \dots (1)$$

ここで、 $\lambda$ は集合同士の比較演算子であり

$\subset, \subseteq, =, \supset, \supseteq, \neq, \eta$ のいずれかである。 $S \eta R$ は、集合 $S$ と $R$ が互いに素であることを表す。本論文では、さらに、集合の要素の値を比較する演算(式(2)~(5))を考える。

$$S \overset{\exists}{\theta} R \equiv (\exists s \in S, \exists r \in R)(s \theta r) \quad \dots (2)$$

$$S \overset{\exists}{\theta^{\vee}} R \equiv (\exists s \in S, \forall r \in R)(s \theta r) \quad \dots (3)$$

$$S \overset{\forall}{\theta} R \equiv (\forall s \in S, \exists r \in R)(s \theta r) \quad \dots (4)$$

$$S \overset{\forall}{\theta^{\vee}} R \equiv (\forall s \in S, \forall r \in R)(s \theta r) \quad \dots (5)$$

例えば、(3)式は、集合 $R$ の全ての要素に対して $\theta$ -比較が成立する集合 $S$ の要素が少なくとも一つ存在する場合に真となる。演算子の左辺の集合(集合 $S$ )に属する少なくとも一つの要素と、右辺の集合(集合 $R$ )の全要素について $\theta$ -比較を行う演算であるので、この演算の演算子は $\overset{\exists}{\theta^{\vee}}$ となる。以降では、(2)~(5)式をまとめて $S \theta R$ と表記する。

## 3 名前付き集合モデル

ここでは、名前付き集合モデルを定義する。

### 3.1 構造

名前付き集合モデルでは、名前付き集合とそれを要素とする集合のみを扱う。名前付き集合とは、識別子(id), 名前(name), 集合(S)からなる3つ組である。名前付き集合 $ns$ の定義を以下に示す。

**[定義 1] 名前付き集合**

$$ns \equiv (id, name, S)$$

但し、以下の条件を満足する。

$$(id_i, id_j \in ID_{all}) (i \neq j \Rightarrow id_i \neq id_j)$$

$$S = \{s | s \in NS_{all} \vee s \in DATA_{all}\}$$

ここで、 $ID_{all}$ はすべての名前付き集合の識別子が

らなる集合であり、 $NS_{all}$ は全ての名前付き集合からなる集合、 $DATA_{all}$ は全ての値からなる集合である。 □

以降では、名前付き集合 $ns$ を $(id, name, S)$ で表す。

### 3.2 演算

名前付き集合の演算のうち特徴的なものについて述べる。以下では、名前付き集合 $ns_i, ns_j$ をそれぞれ、 $(id_i, name_i, S_i), (id_j, name_j, S_j)$ とする。

#### 3.2.1 航行演算子

名前付き集合 $ns$ の航行演算子として“.”(ピリオド)と“→”(矢印)がある。ピリオドは、名前付き集合のそれぞれの要素を指し示す。例えば、 $ns_i.id = id_i$ であり、 $ns_i.name = name_i$ である。

矢印はピリオドと異なり、名前付き集合 $ns$ が別の名前付き集合をその構成要素として含む場合、すべての名前付き集合の集合を表す。矢印を用いた航行演算には、 $ns \rightarrow NS$ と $ns \rightarrow DATA$ がある。例えば、 $ns_i.S_i = \{ns_j, data_1\}, ns_j.S_j = \{data_2\}$ である場合、 $ns_i \rightarrow NS = \{ns_j\}$ であり、 $ns_i \rightarrow DATA = \{data_1, data_2\}$ である。また、航行演算子は名前付き集合の集合に対しても適用可能である。以下に、集合に対する航行演算の定義を示す。

#### [定義 2] 集合に対する航行演算

$$NS.ID \equiv \{id \mid id = ns.id \text{ such that } ns \in NS\}$$

$$NS.NAME \equiv \{name \mid name = ns.name \text{ such that } ns \in NS\}$$

$$NS.S \equiv \{s \mid s = ns.S \text{ such that } ns \in NS\}$$

$$NS \rightarrow NS \equiv \bigcup_{ns'} (ns' = ns \rightarrow NS, ns \in NS)$$

$$NS \rightarrow DATA \equiv \bigcup_{data'} (data' = ns \rightarrow DATA, ns \in NS)$$

ここで、 $NS$ は名前付き集合の集合、 $ID$ は識別子の集合、 $DATA$ は値の集合である。 □

例えば、 $NS.ID$ は、 $NS$ に含まれる名前付き集合の識別子の集合を表し、 $NS \rightarrow DATA$ は $NS$ に属する名前付き集合から“→”で迎れる $data$ の集合を表す。

#### 3.2.2 集合への要素の追加と削除

名前付き集合を $ns$ 、名前付き集合の集合を $NS$ として、 $NS$ に $ns$ を追加する演算 $NS + ns$ と削除する演算 $NS - ns$ を定義する。

$$NS + ns \equiv NS \cup \{ns\}$$

$$NS - ns \equiv NS - \{ns\}$$

これらを用いると、例えば、名前付き集合 $ns_i$ に $ns_j$ を属させる演算は $ns_i.S + ns_j$ と表すことができる。

#### 3.2.3 和と差

名前付き集合の和 $ns_i + ns_j$ は、両方の性質を持つ新たな名前付き集合を作成する。

#### [定義 3] 名前付き集合の和

作成される新しい名前付き集合を $ns_x$  ( $id_x, name_x, S_x$ )とすると、 $ns_x$ の各要素は以下の条件を満たす。

$$(\forall id \in ID_{all}) (id_x \neq id)$$

$$name_x = name_i \quad \dots (6)$$

$$S_x = S_i \cup S_j \quad \dots (7)$$

□

式(6)の条件により、 $ns_i + ns_j$ と $ns_j + ns_i$ は等しくない。以降では、名前が等しい $ns$ の和に限り、一般化して $\sum ns$ と表す。名前付き集合の差の演算は、式(7)の条件が $S_x = S_i - S_j$ となる以外は[定義 3]と同様である。

#### 3.2.4 融合

名前付き集合の第3項 $S$ に別の名前付き集合のそれを属させることを名前付き集合の融合と呼ぶ。

#### [定義 4] 融合

名前付き集合 $ns_i$ に $ns_j$ を融合する演算 $ns_i \leftarrow ns_j$ は、 $ns_i$ を次の条件を満たすように更新する。

$$ns_i = (id_i, name_i, S'_i)$$

$$S'_i = S_i \cup S_j$$

□

#### 3.2.5 名前による名前付き集合の統合

名前付き集合の集合 $NS$ に対して、名前が同じ名前付き集合をまとめる演算を定義する。

#### [定義 5] 名前付き集合の統合

以下のような、名前付き集合の集合 $NS_A, NS_B$ を考える。

$$NS_A : \{c(x_a) \mid x_a \in NS_i, \forall x_b \in NS_j$$

$$\text{such that } (x_a.name \neq x_b.name) \wedge (a \neq b)\}$$

$$NS_B : \{\sum x_a \mid x_a \in NS_i, \exists x_b \in NS_j$$

$$\text{such that } ((\sum x_a).name = x_b.name) \wedge (a \neq b)\}$$

$(\sum x_a).name$  は名前付き集合の和 $\sum x_a$ により作成される名前付き集合の名前を表す。これらの集合により $un(NS_i)$ は $NS_A \cup NS_B$ と定義される。 □

図1に名前付き集合の統合の例を示す。図中の

太枠の四角形一つが一つの名前付き集合を表す。四角形の上の数字が名前付き集合の識別子を表し、下のアルファベットが名前を表す。識別子1+2は、識別子1と2の名前付き集合の和により作成された名前付き集合の識別子を表し、c(3)は識別子3の名前付き集合をコピーすることにより作成された名前付き集合の識別子を表す。

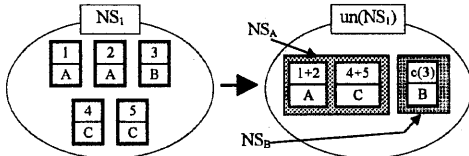


図1 名前付き集合の統合

### 3.2.6 名前付き集合の強和と強差

名前付き集合の強和 $ns_1(+)$  $ns_2$ は、両方の名前付き集合の性質を持つ新たな名前付き集合を作成する。名前付き集合の和の演算との違いは、第3項の集合Sに名前が同じ複数の名前付き集合が属する場合、それらが一つにまとめられる点である。名前付き集合の強差 $ns_1(-)$  $ns_2$ とは、 $ns_1$ と $ns_2$ について、まず、名前が同じ名前付き集合が属している場合はそれらをまとめてから差をとる演算である。以下に、名前付き集合の強和、強差の定義を示す。

#### [定義 6] 名前付き集合の強和

$ns_1(+)$  $ns_2$ の結果作成される新しい名前付き集合を $ns_x : (id_x, name_x, S_x)$ とすると、 $ns_x$ の各要素は以下の条件を満たす。

- $(\forall id \in ID_{all}) (id_x \neq id)$
- $name_x = name_1$
- $S_x = DATA \cup NS$
- $DATA = \{a \mid a \in DATA_{all} \wedge a \in S_i \cup S_j\}$
- $NS = un(X \cup Y)$
- $X = \{x \mid x \in NS_{all} \wedge x \in S_i\}$
- $Y = \{y \mid y \in NS_{all} \wedge x \in S_j\}$

#### [定義 7] 名前付き集合の強差

$ns_1(-)$  $ns_2$ の結果作成される新しい名前付き集合を $ns_x : (id_x, name_x, S_x)$ とすると、 $ns_x$ の各要素は以下の条件を満たす。

- $(\forall id \in ID_{all}) (id_x \neq id)$
- $name_x = name_1$
- $S_x = DATA \cup NS$
- $DATA = \{a \mid a \in DATA_{all} \wedge a \in S_i - S_j\}$

$$NS = NS_A \cup NS_B$$

$$NS_A = \{m \mid m \in un(X), \forall n \in un(Y) \text{ such that } m.name \neq n.name\}$$

$$NS_B = \{m - n \mid m \in un(X), \exists n \in un(Y) \text{ such that } m.name = n.name\}$$

$$X = \{x \mid x \in NS_{all} \wedge x \in S_i\}$$

$$Y = \{y \mid y \in NS_{all} \wedge x \in S_j\}$$

□

以下のような名前付き集合があるとする。

- $ns_a = (id_a, \alpha, \{ns_1, ns_2, ns_3, ns_4\})$ ,
- $ns_p = (id_p, \beta, \{ns_5, ns_6, ns_7, ns_8\})$ ,
- $ns_1 = (id_1, A, S_1)$ ,  $ns_2 = (id_2, A, S_2)$ ,  $ns_3 = (id_3, B, S_3)$ ,
- $ns_4 = (id_4, C, S_4)$ ,  $ns_5 = (id_5, B, S_5)$ ,  $ns_6 = (id_6, D, S_6)$ ,
- $ns_7 = (id_7, D, S_7)$ ,  $ns_8 = (id_8, C, S_8)$

[定義 7]中の $NS_A$ は、 $un(ns_a.S)$ に属する名前付き集合のうち、 $un(ns_p.S)$ に属する名前付き集合が持たない名前を持った名前付き集合の集合であり $\{(1+2, A, S_{1+2})\}$ である。 $NS_B$ は、 $un(ns_p.S)$ に属する名前付き集合のうち、 $un(ns_a.S)$ に属する名前付き集合と同じ名前を持つ物が存在するとき、両者の差をとったものであり、 $\{(c(3)-c(5), B, S_{c(3)-c(5)}), (c(4)-c(8), C, S_{c(4)-c(8)})\}$ となる。 $ns_a(-)ns_p$ は $NS_A + NS_B$ に置き換えられる。従って強差 $ns_a(-)ns_p$ は、 $\{(1+2, A, S_{1+2}), (c(3)-c(5), B, S_{c(3)-c(5)}), (c(4)-c(8), C, S_{c(4)-c(8)})\}$ となる。

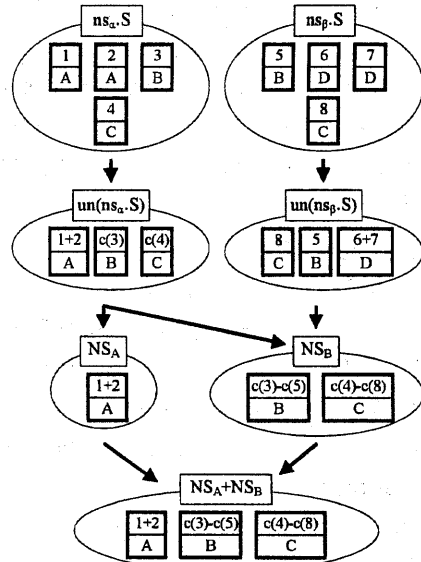


図2  $ns$ の強差  $ns_a(-)ns_p$

### 3.2.7 包含

名前付き集合 $ns_i$ が別の名前付き集合 $ns_j$ の構成要素となっている場合、名前付き集合 $ns_i$ は名前付き集合 $ns_j$ に含まれるといい $ns_i \ll ns_j$ で表す。

### 3.2.8 比較演算

名前付き集合 $ns_1$ と値の集合 $X$ との比較は以下のように定義される。

2つの名前付き集合 $ns_1, ns_2$ の比較演算には、1)名前付き集合の識別子を比較する演算、2)名前付き集合の第3項の集合同士を比較する演算、3)名前付き集合の名前と第3項を比較する演算の3つの演算がある。それらの演算の定義を以下に示す。

#### [定義 8] 比較演算1

$$ns_1 \theta_{id} ns_2 \equiv ns_1.id \theta ns_2.id \quad \dots (8)$$

$$ns_1 \theta ns_2 \equiv ns_1.S \forall \theta ns_2.S \wedge ns_1.S \exists \theta ns_2.S \quad \dots (9)$$

$$ns_1 \theta_{name} ns_2 \equiv ns_1.name \theta ns_2.name \wedge ns_1.S \forall \theta ns_2.S \wedge ns_1.S \exists \theta ns_2.S \dots (10)$$

□

例えば、3つの名前付き集合、 $ns_1=(id_1, name_1, \{1,2,3\})$ 、 $ns_2=(id_2, name_2, \{2,3,4\})$ 、 $ns_3=(id_3, name_3, \{-2,-1,5\})$ について、式(9)の比較で“ $<$ ”を考えた場合、 $ns_1 < ns_2$ が成立する。すなわち、 $ns_1 < ns_2$ は、「 $ns_2$ の最大値が $ns_1$ の全ての要素よりも大きく、 $ns_1$ の最小値が $ns_2$ の全ての要素よりも小さい」とき真となる。従って、 $ns_3$ が“5”を要素として持っているため $ns_3 < ns_1$ は偽である。以降では、これら(8)~(10)式の演算を $ns_1 \theta_x ns_2$ と表す。これを用いて(2)-(5)式の集合の比較演算を名前付き集合の集合に対して拡張する。

#### [定義 9] 比較演算2

$$NS_1 \exists \theta_x NS_2 \equiv (\exists a \in NS_1, \exists b \in NS_2) (a \theta b)$$

$$NS_1 \exists \theta_x NS_2 \equiv (\exists a \in NS_1, \forall b \in NS_2) (a \theta b)$$

$$NS_1 \forall \theta_x NS_2 \equiv (\forall a \in NS_1, \exists b \in NS_2) (a \theta b)$$

$$NS_1 \forall \theta_x NS_2 \equiv (\forall a \in NS_1, \forall b \in NS_2) (a \theta b) \quad \square$$

以降では、これらの式をまとめて $S \theta R$ と表す。

### 3.2.9 選択演算

名前付き集合モデルにおける選択演算について述べる。以下に選択演算を示す。

$$NS[id \theta x] = \{ns \mid ns \in NS \wedge ns.id \theta x\}$$

$$NS[name \theta x] = \{ns \mid ns \in NS \wedge ns.name \theta x\}$$

$$NS[S \sigma x] = \{ns \mid ns \in NS \wedge ns.S \theta x\}$$

$$NS[S \theta R] = \{ns \mid ns \in NS \wedge ns.S \theta R\}$$

$$NS[S \lambda R] = \{ns \mid ns \in NS \wedge ns.S \lambda R\}$$

ここで $NS$ は名前付き集合の集合であり、 $x$ は任意の値、 $R$ は名前付き集合又は値の集合である。

上記の演算のうち、いくつかについて例を用いて説明する。以下に示す名前付き集合を要素とする名前付き集合の集合 $NS: \{ns_1, ns_2, ns_3\}$ を考える。

$$ns_1:(id_1, X, \{ns_4, ns_5, ns_6, ns_7\}),$$

$$ns_2:(id_2, Y, \{ns_8, ns_9, ns_{10}, ns_{11}\}),$$

$$ns_3:(id_3, Z, \{ns_{12}, ns_{13}, ns_{14}, ns_{15}\}),$$

$$ns_4:(id_4, A, \{1\}), ns_5:(id_5, A, \{2\}), ns_6:(id_6, B, \{1\})$$

$$ns_7:(id_7, C, \{3\}), ns_8:(id_8, A, \{1\}), ns_9:(id_9, B, \{2\})$$

$$ns_{10}:(id_{10}, C, \{1\}), ns_{11}:(id_{11}, D, \{8\}), ns_{12}:(id_{12}, A, \{1\}),$$

$$ns_{13}:(id_{13}, A, \{5\}), ns_{14}:(id_{14}, B, \{3\}), ns_{15}:(id_{15}, C, \{5\})$$

ここで、集合 $NS$ に対して“名前 $X$ を持つ名前付き集合を選ぶ”という選択は、 $NS[name="X"]$ により表現される。この演算の結果は $\{ns_1\}$ である。

また、集合 $NS$ に対して“名前 $D$ を持つ名前付き集合を含む名前付き集合を選ぶ”という選択は

$$NS[S[name="D"] \neq \emptyset]$$

となる。この選択の $S[name="D"]$ の部分は、集合 $NS$ のそれぞれの要素を対象として名前が $D$ であるものを選択する。

従って、対象が $ns_1, ns_3$ ではその結果は空集合となり、 $ns_2$ では $\{ns_{11}\}$ が選択される。 $\{ns_2\}$ が演算

$$NS[S[name="D"] \neq \emptyset]$$

の結果となる。集合 $NS$ に対して“名前 $A$ の値と名前 $C$ の値が等しい名前付き集合を選ぶ”という選択は、

$$NS[S[name="A"] \exists \forall S[name="C"]][S[name="A"] \forall \exists S[name="C"]]$$

により表される。 $S[name="A"]$ の部分は、集合 $NS$ の要素から名前が $A$ であるものを選択し、 $S[name="C"]$ の部分は名前が $C$ であるものを選択する。例えば $ns_1$ が対象の場合、前者は $\{ns_4, ns_5\}$ であり後者は $\{ns_7\}$ である。 $NS[S[name="A"] \exists \forall S[name="C"]]$ は、それら選択された名前付き集合の間で比較演算 $\exists \forall$ が成立する要素を含む名前付き集合を集合 $NS$ から選択する。ここでは $\{ns_2, ns_3\}$ となる。例では、この集合にさらに $[S[name="A"] \forall \exists S[name="C"]]$ の条件を用いた選択演算を行って

いる。 $ns_3$ には、名前が $A$ であるものは $ns_{12}$

と $ns_{13}$ であるが、名前がCであるものの中に $ns_{12}$ と同じ値を持つものがない。そのため、 $ns_3$ は除かれて選択の結果は $\{ns_2\}$ となる。

### 3.2.10 結合演算

比較演算を $\phi$ で表すと、名前付き集合の集合 $NS_1$ 、 $NS_2$ の結合(join)演算 $NS_1[\phi] NS_2$ は、 $NS_1 \times NS_2[\phi]$ で表される。 $NS_1 \times NS_2$ は名前付き集合の集合の積であり $\{x+y | x \in NS_1, y \in NS_2\}$ である。これは、リレーショナルモデルにおける直積に相当する。

## 4 名前付き集合モデルを用いた DREAM モデルの定義

次に、名前付き集合モデルを元にして、導出データベースにおけるDREAMモデルを定義する。

### 4.1 準備

データベースエレメントの定義の前にいくつかの準備を行う。

[準備1]  $de$ はデータエレメントを、 $obj$ はオブジェクトを、 $ne$ は名前付きエレメントを、 $pe$ は視点、 $bndl$ はバンドル、 $ddb$ は導出データベースを表す。また、それぞれの集合を、 $DE$ 、 $OBJ$ 、 $NE$ 、 $PE$ 、 $BNDL$ 、 $DDB$ と表す。

[準備2] 全てのオブジェクトからなるオブジェクトの集合を $OBJ_{all}$ と表す。そのほかの要素についても同様である。

### 4.2 データベースエレメント

#### [定義 10] データエレメント

データエレメント $de$ は識別子とデータ値からなる2つ組 $(id, data\_val)$ である。但し $(de_i, de_j \ll ddb)$   $(id(de_i), id(de_j) \in ID_{all}) (i \neq j \Rightarrow id(de_i) \neq id(de_j))$ を満たす。ここで、 $id(de_i)$ は $de_i$ の識別子を表す□

#### [定義 11] オブジェクト、視点、名前付きエレメント、バンドル

オブジェクト $obj$ は3つ組 $(oid, obj\_name, \{pe\})$ である。ここで、 $oid$ はオブジェクトの識別子、 $obj\_name$ はオブジェクトの名前、 $pe$ は視点である。視点 $pe$ は3つ組 $(peid, pe\_name, \{ne\})$ である。 $peid$ は視点の識別子、 $pe\_name$ は視点の名前、 $ne$ 名前付きエレメントである。名前付きエレメント $ne$ は3つ組 $(neid, ne\_name, S)$ である。 $neid$ は名前付きエレ

メントの識別子、 $ne\_name$ は名前付きエレメントの名前である。 $S$ は $\{S | S \in OBJ_{all} \vee S \in DE_{all}\}$ である。名前付きエレメントと視点の間には以下の制約がある。

$(ne \in NE_{all}) (name(ne) \neq NULL)$

$(\forall ne \in NE_{all}, \exists pe \in PE_{all}) (ne \ll pe)$

$(pe_i, pe_j \in PE_{all}, ne \ll pe_i) (i \neq j \Rightarrow \neg(ne \ll pe_j))$

$(ne_i, ne_j \ll pe) (i \neq j \Rightarrow name(ne_i) \neq name(ne_j))$

ここで、 $NULL$ は値が空であることを表す。すなわち、名前付きエレメントは必ず名前を持ち、視点に属さなければならない。しかし、同時に複数の視点に属することはできない。また、同じ視点に属している名前付きエレメントは同じ名前を持つことができない。同様のことが視点とオブジェクトの間にもいえる。

バンドル $bndl$ は3つ組 $(bndlid, bndl\_name, \{obj\})$ である。ここで、 $bndlid$ はバンドルの識別子、 $bndl\_name$ はバンドルの名前である。但し、 $(\forall bndl_i, bndl_j \ll ddb_i) (i \neq j \Rightarrow name(bndl_i) \neq name(bndl_j))$ である。□

### 4.3 演算

#### 4.3.1 航行演算

データエレメントに対する航行演算子には“.”(ピリオド)がある。 $de.id$ はデータエレメントの識別子を、 $de.data$ はデータエレメントのデータを表す。データエレメント以外のデータベースエレメントに対する航行演算子には名前付き集合 $ns$ に対する航行演算子と同様である。例えば、 $ne.name$ は名前付きエレメントの名前を、 $ne \rightarrow OBJ$ は、 $ne$ から辿れる全てのオブジェクトの集合を表す。

#### 4.3.2 和、差、商、積など

データエレメントの和 $de_i + de_j$ は、データエレメント $de_x$ を作成する。 $de_x$ は以下の条件を満たす。

$dom(de_i, data) = dom(de_j, data)$

$(\forall id \in DEID_{all}) (de_x.id \neq id)$

$de_x.data = de_i.data + de_j.data$

ここで、 $dom(de_i, data)$ は $de_i.data$ の定義域を表す。差、積、商などの四則演算やそれ以外の値に関する一般的な演算についても同様のことが成立する。名前付きエレメント、視点、オブジェクト、バ

ンドルは第3項Sにそれぞれデータエレメント又はオブジェクト, 名前付きエレメント, 視点, オブジェクトの集合を持つ名前付き集合であり, 名前付き集合モデルの比較や選択などの演算がほぼそのまま利用できる. ただし, 制約により再定義が必要な演算が幾つかある. ここでは, 最も違いの大きい視点の和の演算について再定義を行う.

#### [定義 12]視点の和

視点 $pe_i, pe_j$ の和 $pe_i+pe_j$ は, 両方の視点に含まれる名前付きエレメントを含む新しい視点を作成する. 演算の結果作成される視点を $pe_x : (peid_x, pe\_name_x, NE_x)$ とすると各要素は以下の条件を満たす.

$$(\forall peid \in PEID_{all}) (peid_x \neq peid)$$

$$pe\_name_x = pe\_name_i$$

$$NE_x = NE_i (+) NE_j \quad \square$$

第3項の集合の和が強和になっている点が, 名前付き集合の和と異なる.

#### 4.3.3 選択演算

DREAMモデルの選択演算は, 名前付き集合モデルの選択演算と同じである. 例えば, オブジェクト集合OBJから, 「視点Aから見て, 名前がBである名前付きエレメントを含み, 値が10であるデータエレメントを含むオブジェクトを選ぶ」という選択は, 次の演算で表現できる.

OBJ[PE[name="A"]][NE[name="B"]][DE[data>10]]

#### 5 まとめ

本論文では, サイエнтиフィックデータベース管理システムDREAMのデータモデルについて述べてきた. まず, 3つ組 (識別子, 名前, 集合) をもとにした名前付き集合モデルを提案した.

名前付き集合は, 名前の付いた集合である, これにより, 集合の和と差において名前が同一の物を一つの要素としてまとめる強和と強差を提案した. また, 集合間の比較において, 要素同士を比較する演算を提案した. 次に, 名前付き集合モデルを用いてDREAMモデルを定義した. その結果, 名前付き集合モデルを用いてDREAMモデルが非常に簡明に定義できることが分かった. これにより, 名前付き集合モデルを実現することで,

DREAMモデルを容易に実現でき, 名前付き集合モデルの強力な演算をDREAMモデルにおいて利用することが可能となる.

今後の課題は, オブジェクトの集合であるバンドルにおいて, 例外的にバンドルに属する例外オブジェクトを認めるためのモデルの拡張, ボトムアップに構築した導出データベースで, 従来のデータ構造を用いたトップダウンな検索をサポートすることなどである.

#### 参考文献

- 1) Christian Kalus, Peter Dadam : Flexible Relations - Operational Support of Variant Relational Structures, Proc. of 21th VLDB, pp. 539-550 (1995).
- 2) François Bancillon, Claude Delobel, Paris Kanellakis : BUILDING AN OBJECT-ORIENTED DATABASE SYTEM: THE STORY OF O<sub>2</sub>, Morgan Kaufmann Publishers (1992).
- 3) Obase Consortium: Obaseプロジェクト第三期研究報告書, 千里国際情報事業財団, pp.205-219 (1994).
- 4) Lieberman, H. : Using Prototypical Objects to Implement Shared Behavior in Object Oriented Systems, Proc. of OOPSLA'86, pp. 214-223 (1986).
- 5) 銭晴 他 : ハイパーテキストデータベースシステム TextLink/Gemにおけるオブジェクトとスキーマの動的, 段階的な構築機能, 信学技報, DE92-39, pp. 1-8 (1993).
- 6) 石丸知之, 植村俊亮 : オブジェクト指向データモデルにおけるオブジェクトの多重表現, 電子情報通信学会D-I, Vol. J78-D-I No.3 pp.349-357 (1995).
- 7) 中田充 他 : サイエнтиフィックデータベースのためのデータモデルの一提案, 情処研報 95-DBS-101, pp. 65-72 (1995).
- 8) 中田充 他 : サイエнтиフィックデータモデルの一評価, 信学技報, DE95-64, pp. 113-120 (1995).
- 9) 中田充 他 : 考古学データの柔軟な管理をめざしたデータベースシステムの設計と実装, 情報考古学Vol.1 (1), pp46-54 (1995).