

個人情報管理環境

2M-2

盛屋邦彦・三浦孝夫**

*日立ソフトウェアエンジニアリング(株)

**産能大学経営情報学部

1. 背景

大型計算機におけるデータベースでは、不特定多数の利用者が任意のタイミングでこれを用いる環境が当然であった。高価な計算機資源を活用するためには、大量のデータを多数の目的で使用すべきであるという前提が存在したのである。

しかし個人環境では、これらの前提に対してデータベースに対する価値判断もかなり異なったものとなる。最近の計算機処理能力の発展は驚異的で10MIPSを越える速度、10ないし20MBの主記憶、100ないし300MBの容量のハードディスクが200万円前後で購入できる。この変化は、費用対処理規模の変化も引き起こす。つまり大型計算機でしか実現できなかった処理が個人で行えるようになった。これらの状況から、専用目的の個人使用という、およそ逆の狙いをもつデータベースの出現が可能となる。つまり従来のものとは異なった新しいデータベース理論の適用が可能である。

またアーキテクチャ的にも現在の個人レベルの計算機は大型計算機と異なっている。つまり主記憶の増大化である。個人レベルの計算機では1:1以下も珍しくない。つまり、主記憶データベースがごく自然に実現でき、2次記憶はむしろバックアップ域と考えることができる。

本稿ではこのような個人環境でのデータベースの管理、利用方法を考察し、その処理モデルを提案する。

2. 個人情報管理環境への要件

計算機の性能/費用の向上につれて利用層が広がり、必ずしも専門的な知識を持たずプログラミング経験の無い人が計算機の活用を行うようになった。最大の利用は、自らの情報の管理である。電話台帳、住所録、催しのリスト、料理メモ、お歳暮発送リストなど、身の回りにある情報の整理や有機的な活用が対象である。

しかし、これらの情報には共通していくつかの特性がある[LM84, Mo86, A87]。まず扱われるデータ量がきわめて少ない。反面、定型的な利用が無く、どの様なデータの使われ方があるのか、利用者自身を含めて予測がつかない。この結果、与えられたデータを分類する過程でスキーマが定まる。また、ビューとスキーマの区別が無い。利用者はデータの意味はよく理解しているため一貫性制約も異なった方法が考えられる。不注意を未然に防ぐため、常に強要する制約で単純なもの、例えば領域上の制約を考えることがある。しかし複雑な制約については強要するよりむしろ単純に制約を満たしているかを問い合わせられる方がよい。

結局、これらをまとめると個人情報を管理するデータベースに要求される機能は、次のようになる。

a) データの分類整理能力

特に、絶え間ないスキーマの変更を支援する。

b) データ検索能力

これには、与えた条件を満たす情報を得ることと、情報を得るための情報を得ることの、2通りがある。

c) 整合性保証、単純なものでも十分である。
これらをこれまでの大型計算機での手法で実現するのは難しく個人環境向きの手法が必要である。

3. 関連研究

これまで、こういった個人情報の管理を目指す研究がなされている。

PDM (Personal Data Manager) は、意味データモデルに基づくブラウザ指向のシステムである。しかし、n項関連は人工的な主体で表現せねばならぬ。

ハイパーメディア(テキスト)は情報を階層的に表現し、それを使いやすいインターフェイスにより操作する。階層関係に基づかない操作は行いにくい。

ブラウザ方式の操作言語では、共通してスキーマ全体を視覚的に捉え構文的な煩わしさから逃れようとする。スキーマに対する問い合わせや変更についてはまだ議論が不十分である。

関係モデルの上にトークン(名詞)方式の質問を許す提案がなされている[Mo86]。しかしこの方法では得られた質問は一意ではない。

オブジェクト指向パラダイムに基づく提案ではハイパーテキストと同様にモデルの中立性に問題がある。

4. 個人情報の処理モデル

ここでは、個人情報を管理するための基礎となるデータベースモデル(情報処理モデル)を提案する。

4.1 情報表現とデータモデル

個人環境を考えた場合、関係モデルでは主体が値により表現されるため誤解を産みやすく、2項関係に基づくデータモデルでは自然な形でn項関係を表現できない。またオブジェクト指向モデルでは表現の中立性がない。そこでをn項関係に基づく意味論データモデルを採用する[AM84]。これは次の点に特徴をもつ。

a) 主体識別概念を持つ

b) モデル表現の中立性

c) 表現の基本要素が少ない

基本要素

a) 主体(集合): 現実世界の事物の代理物(の集合)

b) 主体型: 主体が属する主体集合名(原子、複合型)

c) 連想(関連): 主体と主体を結び付けるもの

d) 述語: 連想が属する連想集合名

意味操作

汎化/専化、詳細化(集約)/抽象化、構造化/平坦化

4.2 データ操作

個人利用の場合ビュー(部分スキーマ)を通して検索するより、データベース全体を把握できたり、もしくは演繹された情報を得られる方が望ましい。また複

Management Environment of Personal Information

Kunihiko MORIYA¹, Takao MIURA²

¹Hitachi Software Engineering, Ltd. ²SANNO University

雑な式を用いて操作するより基本的な巡航操作を用いて試行錯誤的にデータベースをそのままの形で操作できる方がよい。従ってここではブラウジングに基づくデータ操作を採用する。データベースをそのままの形で見せるため、1で仮定した基本要素及び意味操作がビジュアルに表現されていることが必要である。ここではデータベースの主体及び連想に対応するものとしてそれぞれ主体カードと連想カードを用意する。データベースとはこれらのすべてのカードの集まりである。カードは統一化された操作により参照、生成、併合、削除が行われる。ここではこれをデータ操作と呼ぶ。データ操作はこれらのカードにより巡航的に行われて行く。問い合わせとはデータベース中の部分カードがコピーされることである。問い合わせごとの履歴が残るので前の状態を残したままで検索を絞り込むことが可能である。またできたカード群に対して外延的に操作を加えられる。これはネットワークデータモデルやオブジェクト指向モデル、ハイパーテキストとは本質的に異なる巡航操作である。さらにこのシステムではデータベース全体やスキーマ情報も主体カードとして表現されている。したがってここでの検索はデータ/メタデータの区別のない機構により行われる。

主体カードには様々な情報が記述されている。主体を記述するため主体カードは主体情報と型情報を持っている。これらはラムダ表現により表わされる。この情報によりデータベースの自己記述も行われる。

ラムダ表現に基づく構造情報の記述

a) 主体情報

$\langle \Theta T, \Theta S \rangle$ 主体とそのスペル

b) 型情報

$\langle \Theta T, \Theta \text{LAMBDA} \rangle$ ある型に属する役割とそのときの詳細化情報

ある主体 e とそれが属する主体型 t (例)

$\text{LAMBDA}(t, \Theta E) = \langle \Theta T(\Theta E), \Theta S(\text{"human"}) \rangle$

$\text{LAMBDA}(e, \Theta E) = \langle \Theta T(\Theta E), \Theta S(\text{"yamada"}) \rangle$

e の主体情報

$\text{LAMBDA}(t, \Theta T) = \langle \Theta T(\Theta T), \Theta \text{LAMBDA}(\langle \text{AGE}, \text{SEX} \rangle) \rangle$

$\text{LAMBDA}(e, \Theta T) = \langle \Theta T(\Theta T), \langle \text{AGE}, \text{SEX} \rangle(\langle 30, M \rangle) \rangle$

e の型情報

全ての主体が属する主体型 ΘE

$\text{LAMBDA}(\Theta E, \Theta E) = \langle \Theta T(\Theta E), \Theta S(\text{"entity"}) \rangle$

$\text{LAMBDA}(\Theta E, \Theta T) = \langle \Theta T(\Theta T), \Theta \text{LAMBDA}(\langle \Theta T, \Theta S \rangle) \rangle$

全ての主体型が属する主体型 ΘT

$\text{LAMBDA}(\Theta T, \Theta E) = \langle \Theta T(\Theta E), \Theta S(\text{"type"}) \rangle$

$\text{LAMBDA}(\Theta T, \Theta T) = \langle \Theta T(\Theta T), \Theta \text{LAMBDA}(\langle \Theta T, \Theta \text{LAMBDA} \rangle) \rangle$

全てのスペル主体が属する主体型 ΘS

$\text{LAMBDA}(\Theta S, \Theta E) = \langle \Theta T(\Theta E), \Theta S(\text{"spell"}) \rangle$

$\text{LAMBDA}(\Theta S, \Theta T) = \langle \Theta T(\Theta T), \Theta \text{LAMBDA}() \rangle$

全ての述語が属する主体型 ΘP

$\text{LAMBDA}(\Theta P, \Theta E) = \langle \Theta T(\Theta E), \Theta S(\text{"predicate"}) \rangle$

$\text{LAMBDA}(\Theta P, \Theta T) = \langle \Theta T(\Theta T), \Theta \text{LAMBDA}(\langle \Theta P, \Theta \text{DEF} \rangle) \rangle$

全ての LAMBDA 主体が属する主体型 ΘLAMBDA

$\text{LAMBDA}(\Theta \text{LAMBDA}, \Theta E) = \langle \Theta T(\Theta E), \Theta S(\text{"property"}) \rangle$

$\text{LAMBDA}(\Theta \text{LAMBDA}, \Theta T) = \langle \Theta T(\Theta T), \Theta \text{LAMBDA}(\Theta P) \rangle$

全ての DEF 主体が属する主体型 ΘDEF

$\text{LAMBDA}(\Theta \text{DEF}, \Theta E) = \langle \Theta T(\Theta E), \Theta S(\text{"P-def"}) \rangle$

$\text{LAMBDA}(\Theta \text{DEF}, \Theta T) =$

$\langle \Theta T(\Theta T), \Theta \text{LAMBDA}(\text{complex type of } \Theta T) \rangle$

これらに対応するカードに対してマウス等による直接指示によりデータ操作を行える。以下に基本操作を示す。これらを組み合わせればデータベース中の全て

の情報を問い合わせることが可能である。更新操作については紙面の都合上省略する。

(1) 主体のブラウズ

主体カードに対して主体情報及び型情報(ラムダ値)を問い合わせればよい。複合主体ならばその構成主体が得られる。

(2) 主体の選択

主体型カードの型情報に条件を入れ主体を絞り込む。

(3) 連想の選択

述語カードの型情報に条件を入れ、連想を絞り込む。

(4) 主体のメンバーシップ

主体からそれを含む主体型を求める。

(5) 述語の問い合わせ

主体型からそれが関与する述語を求める。

(6) 制約の問い合わせ

主体からそれに関わる制約を求める。

(7) 補助操作

(集合操作) カード群とカード群の和、積。

(外延操作) 主体カード及び連想カードの直接操作挿入、削除、並び替え等。

他にマクロ機能等がある。

4.3 スキーマ変更とデータ操作

個人データを見て、その利用の方法から動的にスキーマの変更を行いたいという要求が良くある。

これまでのデータベースと違って、データ内容を失うのでない限りある程度スキーマを変更しても良い。

ここで考えるスキーマ変更とは、同じデータ(または等価なデータ)が異なる解釈を有するように変換し、後にデータの外延的な変更を許容するメカニズムを言う。したがってスキーマ変更は一般的にきわめて限られた場合のみ有効である。

最も頻繁に生じるであろう変更は、主体の持つ役割の変更である。つまり汎化及び専化である。汎化は制約条件であり内包的な定義であるが、生成された主体集合は後に外延的にも主体の追加が起こる。また専化は新しい集合の外延的な定義である。

個人情報では、いわゆる構造情報が数多く出現し変更の対象となる。構造情報の抽象化(情報隠蔽)・詳細化(情報開示)もデータ操作におけるスキーマ変更である。

4.4 整合性保証

利用者の制約条件の定式化する能力・範囲が限られているため次の機構を前提として考えている。

a) 主体型の制約に基づく主体の領域チェック

b) スキーマ変更に対し制約条件が可能な限り自動的に追従する。さらにビューを変化させない機構

c) 制約条件について問い合わせられる

より複雑な制約条件は質問処理で対応する。

5. 結び

個人環境におけるデータベースを考察し、それを前提としたデータベースモデルを提案した。

参考文献

[AM84] Arisawa, Miura: Formal approach to database description, COMPCON, 1984

[A87] 有沢: パーソナルデータベース, 信学会報87/12

[LM84] Lyngbaek, Mcleod: A Personal Data Manager, VLDB 1984

[La86] Larson: Visual Approach to Browsing in a Database Environment IEEE COMPUTER, Jun, 1986

[Mo86] Motro: Constructing Query from Token, SIGMOD 1986