数理解析研究所講究録
423巻 1981年 310-322

データベースビューについて

東北大学電気通信研究所　増永良文

1. はじめに

データベースは社会の貴重な共有資源である。しかしながらデータベースを利用しようとすると，一般に蓄積されているデータやその操作言語が利用者の使用したい形で提供されているという保証はない。五人の利用者がいれば五様の利用形態を望むといっても過言ではない。このようなときデータベースシステムがビュー（view）をサポートし，データベースとその操作言語を仮想化して利用者に提示し，データベースに格納されている情報を有効に利用させることができれば大変都合が良い。それが実世界の利用者の視野を実質的にサポートすることになるからである。データベース管理システムのビューサポート機能は現在のところ十分に機能していない。例えば関係データベースシステムとして知られているIBM社のSYSTEM R，QBE，あるい

はカリフォルニア大学バークレー分校のINGRESではビュー定義機能を提供しているが，検索を対象にしているだけで肝心の更新操作はサポートされていない。CODASYL方式のデータベースシステムではいわゆるサブスキーマの定義が利用者視野を定義する概念に相当するが，ビュー定義機能は極めて小さい。尚，良く知られているようにANSI/X3/SPARCはデータベースシステムアーキテクチャとして内部，概念，外部の三スキーマからなる構造を提案しているが，上述ビューサポートの概念はデータベース管理システムが概念スキーマと外部スキーマ間の検索，更新用のインターフェイスを具備することに相当している。

ビューサポートのメカニズムについては1975年頃より研究が開始されてきている。しかしながら数学的平明さを有する関係データベースシステムの枠内でも現在十分なメカニズムの解明がなされているとはいえない。それは本質的に問題解決がシンタクティックな枠組だけでは行なえず，必ずセマンティックな枠組を必要としてくるからである。しかしながら，著者は問題解決に次の三ステップの段階を踏むことを提案し，ビューサポートの問題解決を目指すものである。

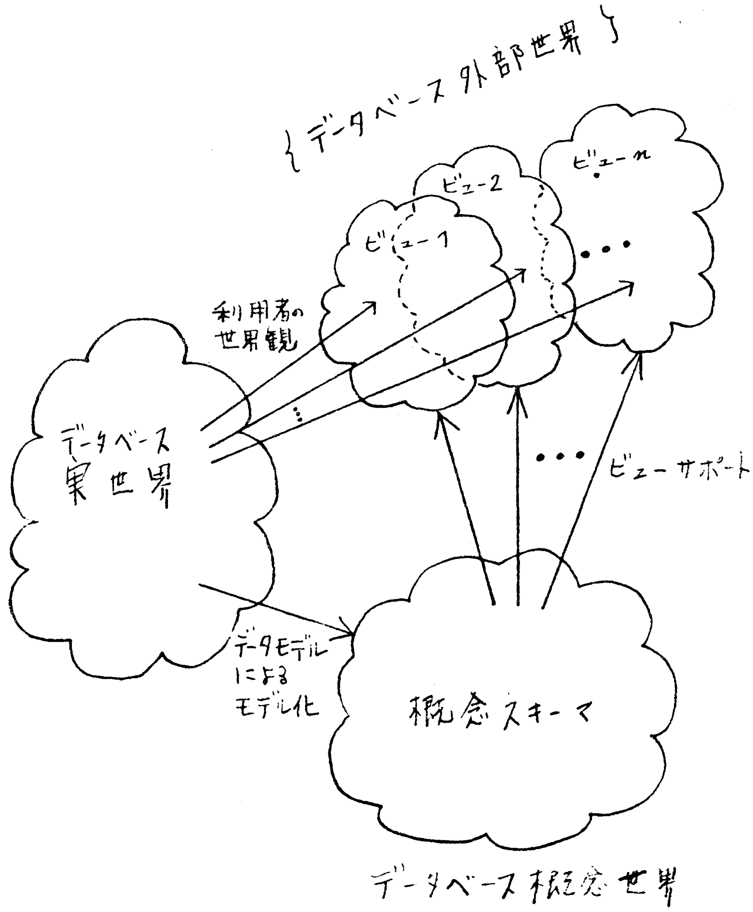
(ステップ1) データベースビュー及びそのサポートとは何

かを与えられたデータモデルの枠組の中で明確に定義する。
(ステップ2) その枠組の中で，どのような機能が十分働き，何がそうでないか，従って何が問題的なのかをシンタクティックかつセマンティックの両側面から明らかにする。
(ステップ3) 問題解決のために何をなすべきかを明確にする。 具体的には次の二つになる。
(i) 与えられたデータモデルの枠内でのサポート限界を明示し，この範囲内で完全に機能するメカニズムを構築する。
(ii) (明らかとなった問題点は満足できないものであるとして)新しいデータモデルを考案し，問題解決をはかる。
本稿ではステップ1及びステップ2の一側面を明らかにすべく関係データモデルの枠組の中で以下若干の議論を展開する。

2. ビューサポート概念

一つの組織体がデータベース構築の対象とした実世界をデータベース実世界(DBRW)と呼ぶ。 DBRWのデータ構造をなるべく忠実にデータベースシステム内にとり込むべく適当なデータモデルが使用されてデータベース概念世界(DBCW)が構築される。 通常組織体には多数のデータベース利用者がいる。 それら各利用者は実世界に対してそれぞれ固

有の世界観によりビューを持ち，データベース外部世界（DBEW）を形成する。図－1にビューサポートの概念を示す。



図－1　ビューサポート概念図

3. ビューとビューサポートの基礎的定式化

関係データモデルの枠組の中で行なう。Rを関係名，A_i（$i=1,\cdots,n$）を属性名，D_i（$=dom(A_i)$）（$i=1,\cdots,n$）をドメイン名とする。$R(A_1/D_1,\cdots,A_n/D_n)$をフリー関係スキーマという。この定義は次のとおり

である。 $R \triangleq \{ r \mid r \subseteq D_1 \times D_2 \times \cdots \times D_n \}$。 c_j を $2^{D_1 \times \cdots \times D_n}$ 上の述語, $C_R \triangleq \bigwedge_j c_j$ をフリー関係スキーマ R 上に課せられた保全制約条件とする。 C_R に関する R の関係スキーマ R_C (以後単に関係 R_C という場合も多い) を次のように定義する。 $R_C \triangleq \{ r \in R \mid r$ は C_R を満す $\}$。写像 $u_R : R \to R$ が関係 $r\ (\in R_C)$ の更新(挿入, 削除, 書換)演算であるとは $u_R(r) \in R_C$ が成立するときをいう。 $U_R \triangleq \{ u_R : R \to R \}$ とする。

次に $R_1, \cdots, R_m$ をフリー関係スキーマとし, $D = (R_1, \cdots, R_m)$ をフリーデータベーススキーマということにする。 $c_k \in 2^{dom(R_1)} \times \cdots \times 2^{dom(R_m)}$ 上の述語, $C_D \triangleq \bigwedge_k c_k$ とする。 C_D は D 上の保全制約条件である。 $D \triangleq \{ (r_1, \cdots, r_m) \mid r_i \in R_i,\ i = 1, \cdots, m \}$ であり, $D_C \triangleq \{ (r_1, \cdots, r_m) \in D \mid (r_1, \cdots, r_m)$ は C_D を満す $\}$ とする。 写像 $u_D : D \to D$; ここに $u_D \triangleq (u_{R_1}, \cdots, u_{R_m})$ で $u_{R_i} : R_i \to R_i$ (つまり $u_D(r_1, \cdots, r_m) \triangleq (u_{R_1}(r_1), \cdots, u_{R_m}(r_m))$ である) と定義する, がデータベース $d\ (\in D_C)$ の更新演算であるとは $u_D(d) \in D_C$ が成立するときをいう。 $U_D \triangleq \{ u_D : D \to D \}$, $U_D^* \triangleq \{1\} \cup U_D \cup U_D^2 \cup \cdots$ とする。 ここに $\forall d \in D,\ 1(d) = d$ である。 $U_D^2 = U_D \times U_D$ である。

そこでVをフリー関係スキーマとする。D_CがVの部分集合としてのビュースキーマ$s(D_C)$をサポートするという概念を明らかにするために、ビュー定義写像を$s : D_C \to V$と定義する。sはD_CからVへの構造的写像である。$s(D_C)$をD_C上のsのもとでのビューということにする。写像$u_v : V \to V$がビューv（$\in s(D_C)$）の更新演算であるとは$u_v(d) \in s(D_C)$が成立するときをいう。$U_V = \{u_v : s(D_C) \to s(D_C)\}$とする。

定義1. $D_C' \subseteq D_C$とする。D_C'が$s(D_C')$をu_v（$\in \{u_v : s(D_C') \to s(D_C')\}$）についてサポートするとは次の条件が成立すること。

$(\exists ! \theta : \{u_v\} \to U_D^*)(\forall d \in D_C')(u_v(s(d)) = s((\theta(u_v))(d))$ かつ $(\theta(u_v))(d) \in D_C')$。

ここにθを操作的写像という。一般にD_Cの部分集合D_C'の規定の仕方には大別して二通りの方法がある。一つはビューの更新操作に依存する場合であり，他の一つはそうではなく専らビューの構造的写像に依存する場合である。ここではこれ以上言及しない。図-2に上の条件式の前半分を図示している。後半分の条件は$\theta(u_v)$がC_Dの更新操作となることを要求している。∃!は唯一の存在を表わしているがこの唯一性は変換の意味論的曖昧さを禁止すること

と表わしている。

$$\begin{tikzcd}[column sep=huge, row sep=huge]
v\ (=s(d)) \arrow[r, "u_v"] & v'\ (=u_v(s(d)))\ (=s((\theta(u_v))(d))) \\
d \arrow[u, "s"] \arrow[r, "u_D(=\theta(u_v))"'] & d'\ (=(\theta(u_v))(d)) \arrow[u, "s"'] \arrow[from=1-1, to=2-2, phantom, "\theta"]
\end{tikzcd}$$

図-2　ビューサポートの可換図

$U_{s(D_c')} \triangleq \{u_v : s(D_c') \to s(D_c')\}$ を分類する。

$U^D_{s(D_c')} \triangleq \{u_v \in U_{s(D_c')} \mid \forall v \in s(D_c'),\ |u_v(v)| \leqq |v|\}$,

$U^I_{s(D_c')} \triangleq \{u_v \in U_{s(D_c')} \mid \forall v \in s(D_c'),\ |u_v(v)| \geqq |v|\}$,

$U^M_{s(D_c')} \triangleq \{u_v \in U_{s(D_c')} \mid \forall v \in s(D_c'),\ |u_v(v)| = |v|\}$。

ここに $|w|$ は関係 w のタップル数を表わす。

定義2　D_c' が $s(D_c')$ をタップル削除操作についてサポートするとは $U^D_{s(D_c')}$ の任意の元 u_v について定義1の条件式が成立すること。

タップル挿入，書換操作のサポートについても同様定義する。

定義3　D_c' が $s(D_c')$ をサポートするとはタップル削除，挿入，書換操作についてサポートすること。

尚詳細な定義関係については機会を更めて述べてみたいと思う。ここではこれ以上行なわない。

4. ビューサポートの問題点

もっとも基本的なビューの定義概念にINGRESやSYSTEM Rがサポートとようとしたものがある。例えばINGRES(1)ではQUELの構文を使って次のようにビューを定義する。

```
RANGE OF X IS ED
RANGE OF Y IS DM
DEFINE EDM(EMP=X.EMP, DEPT =
   X.DEPT, MGR=Y.MGR)
   WHERE X.DEPT=Y.DEPT
```

ここでED(EMP, DEPT), DM(DEPT, MGR)は基本関係または既に定義されているビューであり，この二つの関係のDEPT上の自然結合としてビューEDMが定義された。通常このタイプのビューではビューを定義するのに次の二種類の演算が使用される。

(1) 関係演算（関係代数の用語では直積，和，差，交り，射影，制限，（自然）結合，割り算等）

(2) 非関係演算（関係名，属性名の変換，ドルを円に変換するようなドメインの変換，更に複雑になって平均，最大，最小等のアグリゲート関数とドメイン間の複雑な計算関数など）。

このタイプのビューについては関係データモデルが世に出てしばらくしてから，いちはやくそのサポート可能性が議論されることになった。当初のアプローチは問題点があると直ちにビュー更新を禁止するという考え方であったが，最近[(2),(3)]データ意味論とからみ深い考察がなされるようになったと共に，このようなシンプルな概念のもとに定義されるビューについてすら問題は極めて複雑であることがわかってきている。これら議論の歴史的サーベイは本稿では行なわないが，意味論とからみビューサポートとはこのようなものだということを示す格好の例を挙げてこの章の使命としたい。

関係CP(CHILD，PARENT)を親子関係を表わしている関係とする。親あるいは子供が死亡すれば両者間に関係CPは消滅するとする。このとき子供と祖父母の関係を表わすビューCGを次の様に定義する。

```
RANGE OF X IS CP
RANGE OF Y IS CP
DEFINE CG(CHILD=X.CHILD,
  GRANDPARENT=Y.PARENT)
  WHERE X.PARENT=Y.CHILD
```

いま(a，b)，(b，c)∈CPであったとしよう。ビューCGの定義から(a，c)∈CGである。さて，

bが死亡したとすると(a,b)及び(b,c)はCPから削除される。それら二つの対の消滅は自動的にビューCGからの(a,c)の消滅を誘引する。これはCGがCPとCPの自然結合の射影として定義されているのだから当然である。ビューCGはそのようなものだと理解している限りは問題ない。しかしCGに本来子孫と祖父母の関係を表わしている関係だとユーザの見方を押しつけると問題となる。何故ならもしCGが本来の子孫と祖父母の関係を表わしているならaとcがCGの関係にあることはaの親b(bはcの子孫でもある)の生存,死亡には無関係に維持されねばならないからである。繰り返すがCGでは子孫にとって親が生存しているという制限付の子孫と祖父母の関係が維持されているということなのである。

以上自然結合の射影として定義されたビューCGの持つ意味と,それが本来の子孫-祖父母の関係を表現していると捉えようとすると問題点のあることを述べた。しかしながら問題を更に複雑にするのは場合によっては自然結合の射影として定義されるビューが逆に本来の姿を捉えていると考えられる世界が存在するからである。次式で定義される従業員と上司の関係を表わすビューを考察する。

```
RANGE OF X IS ED
RANGE OF Y IS DM
DEFINE EM(EMP=X.EMP, MGR=Y.
  MGR)
  WHERE X.DEPT=Y.DEPT
```

このとき(e,d)∈EDかつ(d,m)∈DMで，結果として(e,m)∈EMであるとしよう。このとき従業員eとマネージャーmとを結びつけた部局dが廃部になれば，eとmとの関係EMが消滅すると考えるのは極めて自然であろう。

上に示した二つのビューのシンタクティックな構造は同一である。しかし現実世界の意味論と照し合せて，ビューとしての制限が前の例CPでは妥当性に欠き，後の例EMでは妥当のように思われる。このようにビューサポートの問題はセマンティックスに深くかかわっている点がある。これらの問題点の解決が強力なビューサポート機構を実現するうえで望まれていくことになろう。

5. おわりに

ビューサポートをむつかしくしているセマンティカルな一側面を例題を使って第4章で示した。本来ビューはデータ

モデル及びその操作言語の仮想化技法として捉えることが出来，著者は次に示す幾つかのビューをタイプ分けしている。(4)

A．スキーム変換型ビュー

B．DML変換型ビュー

C．データモデル変換型ビュー

D．抽象化型ビュー

E．知識型ビュー

F．時変型ビュー

これらのビューのサポート構造や目的については別の機会に議論を展開したい。

[謝辞]

本研究に関連しこれ迄御討論下さった諸氏に深謝する。

尚，本研究は一部文部省昭和55年度科学研究費補助金，課題番号568006の援助のもとで行なわれたことを付記する。

[文献]

(1) M.R. Stonebraker 他, "The Design and Implementation of INGRES," ACM TODS 1, 3 (1976)

(2) U. Dayal 他, "On the Updatability of Relational Views," Proc. 4th VLDB Conf., p.368-377 (1978)

(3) F. Bancilhon, "Supporting View Updates in Relational Data Base," Proc. IFIP TC-2 Working Conf. on Data Base Architecture, p.198-219 (1979)

(4) 増永, "仮想化技法としてのデータベースビューについて," 情報処理学会データベース管理システム研究会資料 No.24-2, (1981.3月19日).