

CODASYL データベースシステムに対する 非手続的更新インターフェース設計の基本概念[†]

滝 沢 誠^{††} 野 口 正一^{†††}

既存の異種データベースシステム(DBS)から成る分散型データベースシステム(DDBS)を実現するためには、まず第一に、DDBS全体の共通モデルに基づいて、各DBS上に視野(ローカル概念スキーマ)を設定する必要がある。その第一段階として、CODASYL DBS上に共通言語の基本更新演算が行える関係型インターフェースを設計する問題を考える。そのためには、次の点を論ずる。(1)実体と関連を表す2種(EとR)の関係を定義し、それに基づいて共通モデルとしての概念モデルを定義する。(2)従来のCODASYLモデルを抽象化した論理データ構造と論理演算の記述(論理 CODASYL モデル)を与える。(3)概念モデルのR関係を新たに型の概念を用いて特徴化し、論理 CODASYL モデルに対応するローカル概念(LC)モデルを与える。そして、LCモデルのデータ構造と共通言語 RQL の基本演算と共に、論理 CODASYL モデルと等価であることを示す。(4)さらに、RQLを正しく従来の COBOL DML で表す方法を与え、RQLと COBOL DML の記述等価性を示す。以上は、CODASYL DBS上に非手続的な共通インターフェースを設計するための論理的基礎を与える。

1. まえがき

分散型データベースシステム(DDBS)実現のための重要な問題の一つは、CODASYL^{1), 5)}、IMS³⁾等の異種の既存データベースシステム(DBS)の統合問題である。本論文では、DBSの異種性をデータモデルの相違とする。統合問題を解決する基本的方法は、まず、DDBS全体の共通モデル(概念モデルと呼ぶ)を定め、これに基づいて各DBS上に視野を設定することである。この視野をローカル概念(LC)スキーマとする。次に、DDBS全体で一意な視野としての全体概念(GC)スキーマをこれらのLCスキーマ上に定義し、各DBSの所在と独立にする。以上は図1の四階層構造^{6), 7)}として示せる。本論文では、概念モデルとして、簡単なデータ構造と非手続的操作言語を有する関係型²⁾のモデルを採用し、既存DBSとして普及しているCODASYLモデルと関係型モデル間での異種性の解決を試みる。従来のCODASYLモデルとして、文献15)等の新仕様もあるが、論理的に本質的な問題の検討は、文献1)と5)で十分であると考える。四階層構造に基づく検索用関係インターフェースとして、ローカルデータベースプロセッサ(LDP-V

1. 5)^{8), 10), 14)}が設計され、JipdecのAIM(M-170F)、ADBS(Acos-700)上で実働している。更新用視野の設計については、セット型のメンバシップクラスをモデル化するために、文献13)は外来キー制約¹¹⁾を用い¹³⁾、12)は親レコード型の主キーを子レコード型の主キーに伝播させる方法¹²⁾を用いている。また、文献16)では、関係DBS上にCODASYLスキーマを設ける問題を論じている。これらに対して、本論文では、実体間の関連を、更新演算の意味を与えて特徴化し、2種のメンバシップクラスに対応した型を与える。

以上の考察に基づき、本論文ではまず、既存データモデルを概念化した概念モデルを定義し(2章)、従来のCODASYLモデルを抽象化した論理 CODASYL モデルの記述を与える(3章)。4章では、概念モデルの実体間の関連を特徴化したローカル概念(LC)モデルの定義と、LCデータ構造と述語論理形式のLC操作言語 RQL の基本演算を与え、これらが論理 CODASYL モデルと等価であることを示す。5章では、RQLのCOBOL DMLによる正しい表現方法を与える。

2. 概念モデル

2.1 データモデルの概念化

DBSは、データモデルに基づいたデータ構造と、この上の操作演算を利用者に提供する処理システムと考えられる。データモデルとしては、関係²⁾、階層³⁾、

[†]Fundamental Concepts on the Design of the Non-procedural Update Interface over the CODASYL Database Systems by MAKOTO TAKIZAWA (Japan Information Processing Development Center) and SHOICHI NOGUCHI (Research Institute of Electrical Communication, Tohoku University).

^{††}(財)日本情報処理開発協会
^{†††}東北大電気通信研究所

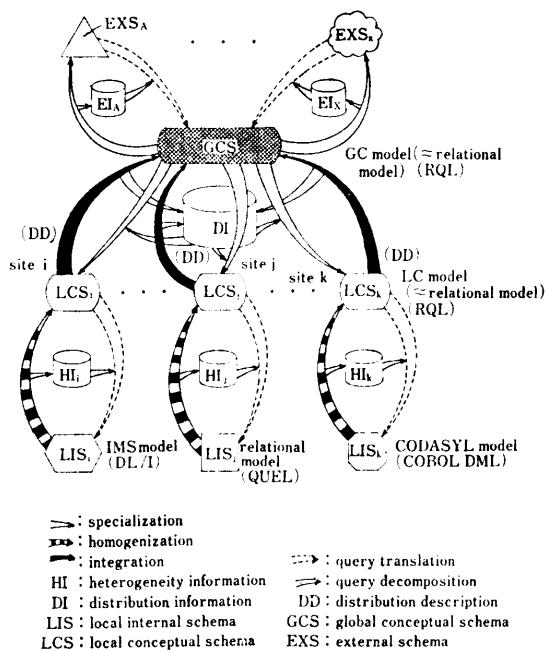


図 1 四層スキーマ構成^{4),7)}
Fig. 1 Four-schema structure.

網^{1),5)}の三つが著名である。本論文では、DBS の異種性を各データモデルの相違とし、データモデルはデータ構造とこれらの操作演算で定まるとする。異種 DBS の統合を行うためには、まず DDBS 全体の共通モデル（概念モデルとする）に基づいて、各 DBS 上に視野を設ける必要がある。

階層、網モデルは、実体と関連との 2 種の概念をしている。また、更新演算の意味を与えて、関連は特性能化され、型づけされている。たとえば、網型の CODASYL モデル^{1),5)}（3 章）では、レコード型は実体集合で与えられ、セット型は二つの実体集合間の 1 対 $m (m \geq 0)$ の関係で表される。さらに、セット型のメンバシップクラスの型は、更新演算の意味する制約条件で特性能化される。階層型の IMS モデル³⁾では、セグメント型が実体集合、階層順序が実体集合間の 1 対 m 関係を表している。以上の考察より、概念モデルは、実体と関連との 2 種の独立な概念をもたねばならず、また各モデルのデータ構造は、概念モデルにおける関連と型の概念を用いて特性能化できることがわかる。本章では、概念モデルのデータ構造と操作言語の定義を与える。

2.2 概念データ構造

概念データ構造は、時間不变な論理構造を与える概念スキーマ \underline{S} と、実際にデータが与えられた（時間可

変な）概念データベース S とから成る。 \underline{S} は、 E と R の 2 種の関係スキームから成る。 E （関係）スキーム E は、属性集合 $\Omega_E = \{\underline{E}_1, \alpha_1, \dots, \alpha_m\} (m \geq 0)$ とインテグリティ条件 Γ_E から成る ($E = (\Omega_E, \Gamma_E)$)。各属性 $\alpha \in \Omega_E$ の可能な値の集合（定義域）を $\text{dom}(\alpha)$ とする。 $\text{dom}(\underline{E})$ は E に固有な識別子の集合で、 \underline{E} を主属性とする。 Ω_E に実際に値を与えた組のなかで、 Γ_E を満足する組を E 組とする。この E 組の集合を E 関係 $E = \{e | e \in \text{dom}(\Omega_E) \wedge \Gamma_E(e)\}$ (ここで $\text{dom}(\Omega_E) = \times_{\alpha \in \Omega_E} \text{dom}(\alpha)$) とする。 E 組 $e \in E$ の属性の部分集合 $A \subseteq \Omega_E$ の値を $A(e)$ とする。 E 内で \underline{E} の値は一意に定めてあるので、 \underline{E} は主キーとなる。 E 関係スキーム Γ_E を省略して、 $E(\underline{E}, \alpha_1, \dots, \alpha_m)$ とも表す。

任意の $n (n \geq 2)$ 個の E 関係スキーム $E_i (i=1, \dots, n)$ と属性集合 $\Omega_{R'} (\Omega_{R'} \cap \Omega_{E_i} = \emptyset)$ があり、関係 $R \subseteq \times_{i=1}^n \text{dom}(\Omega_{E_i}) \times \text{dom}(\Omega_{R'})$ が存在するとき、 R （関係）スキーム $R = (E_1, \dots, E_n, \Omega_{R'}, \Gamma_R)$ が定義される。 $\Omega_{R'} = \{b_1, \dots, b_q\} (q \geq 0)$ と Γ_R は、おののの $\Omega_{E_1}, \dots, \Omega_{E_n}$ より意味的に帰納された属性集合とインテグリティ条件である。 Γ_R は R に与えうる値の組（ R 組）を定める。この R 組の集合を R 関係 $R \triangleq \{r | r \in \times_{i=1}^n \text{dom}(\Omega_{E_i}) \times \text{dom}(\Omega_{R'}) \wedge \Gamma_R(r)\}$ とする。 Ω_{E_i} で \underline{E}_i は主キーなので、 $R(@E_1, \dots, @E_n, \Omega_{R'})$ 、 $R = \{r' | r' = @E_1 \dots @E_n \Omega_{R'}(r) \wedge r \in \times_{i=1}^n \text{dom}(\Omega_{E_i}) \times \text{dom}(\Omega_{R'}) \wedge \Gamma_R(r)\}$ とする。各 R 組 $r \in R$ に対して次の条件が成立立つ。すなわち、すべての E_i について $@E_i(e_i) = @E_i(r)$ なる組 e_i が E_i 内に存在する。

以上の E と R 関係の集合を、概念データベース S と定義する。たんに関係（スキーム）というときは、 $S(\underline{S})$ 内の E または R 関係（スキーム）を表すものとする。

2.3 概念操作言語 (CML)

概念モデルの操作言語 CML によって、 S に対する検索と更新演算を、QUEL⁴⁾ と同様な形式で表せる。

2.3.1 検索演算

検索演算は、(i) 検索条件の設定と、(ii) 目標関係スキームの設定との二つの手順により定義される。(i) の検索条件は、組関係論理言語³⁾ により (1) のよ

うに与えられる。

$$\{(t_1, \dots, t_p) | (\exists u_1) \dots (\exists u_q) \left(\bigwedge_{i=1}^p T_i(t_i) \wedge \bigwedge_{j=1}^q U_j(u_j) \wedge QL(t_1, \dots, t_p, u_1, \dots, u_q) \right) \} \quad (1)$$

$X = \{T_1, \dots, T_p, U_1, \dots, U_q\}$; $x = \{t_1, \dots, t_p, u_1, \dots, u_q\}$ とする。 $x \in X$ は組変数で、 $X \in X$ とすると $X(x)$ は x が関係 X 内の組を値としてとることを表している。 $t_i \in x$ は関係 $T_i \in X$ の組変数で (1) の自由変数、また $u_j \in x$ は関係 $U_j \in X$ の組変数で (1) の束縛変数となる。QL は述語 $\alpha(x) \theta \beta(y), \alpha(x) \theta v$ と結合子 (\vee, \wedge, \sim) から成る式である。ここで、 θ は比較演算子、 v は定数、 $x, y \in x$ はおのおの関係 $X, Y \in X$ の組変数、 α と β はおのおの X と Y の属性である。 $\alpha(x)$ は、関係 X の組の属性値を表す。(ii) では利用者の必要とする目標関係スキームを $T(\Omega_T = \{\alpha_1, \dots, \alpha_k\})$ とすると、 Ω_T は $\tilde{f}(\Omega_{T_1}, \dots, \Omega_{T_p}) = \Omega_T$ なる関係 \tilde{f} を定義することで与えられる。 \tilde{f} から誘導される関数 f を直積 $T_1 \times \dots \times T_p$ に作用させると目標関係 T が与えられる。この f を目標関数とする。

以上をもとに、検索演算は CML で次のように表される。

var $(x_1, X_1) \dots (x_i, X_i); \quad (2)$

get into $T(A)$ where $\Psi; \quad (3)$

(2) は (1) の $X_i(x_i)$ を表し、関係 X_i に組変数 x_i を定める。(3) の A と Ψ はおのおの、目標リストと条件式である。 Ψ は (1) の QL を表す。(1) の $\alpha(x)$ は (3) では $x.\alpha$ と表される。 A は目標関数 f を表し、 $A = (\beta_1, \dots, \beta_k), \beta_j = \alpha_j = \varepsilon_j (j=1, \dots, k)$ である。 ε_j は、 $t.\alpha (t$ は自由変数)、 v 、またはこれらの算術式で、 Ψ を満たす $\{t_1, \dots, t_p\}$ 値から目標属性 α_j 値を与える関数である。たとえば、 E 関係スキーム EMP ($@E, name, age, sal$) に対する「30歳以上の従業員名とその給料の 1.2 倍を求めよ」は CQL で次のように書かれる。

```
var  $(e, EMP);$ 
get into  $T(e.name, sv=e.sal \times 1.2)$  where
 $e.age \geq 30; \quad (4)$ 
```

概念モデルの検索演算では、主属性が識別子をとることから演算子として $=$ と \neq のみが許され、また主属性上の算術演算は許されない。

2.3.2 更新演算

更新演算は、CML では、目標属性集合 Ω を定める目標リスト A と条件式 Ψ 、および Ω に対する次

の六つの基本更新演算 ρ として表現される。ここで、 E と $R (@E_1, \dots, @E_n, b_1, \dots, b_p)$ をおのおの S 内の任意の E と R スキームとする。

- (1) del(e, E) は組 $e \in E$ と同時に、 $@E$ を主属性としてもつ全 R 関係 R から $@E(e) = @E(r)$ なる組 r を除く。
 - (2) del(r, R) は組 $r \in R$ を除く。
 - (3) app(e, E) は組 e を E に加える。
 - (4) app($r, R ; e_1, E_1 ; \dots ; e_n, E_n$) は各 E_i が $@E_i(e_i) = @E_i(r)$ なる組 $e_i (i=1, \dots, n)$ をもつならば、 R に組 r を加える。
 - (5) rep(e, e', E) は組 $e \in E$ の非主属性値を e' に変える。
 - (6) rep(r, r', R) は組 $r \in R$ を r' に変える。
- Ψ と A は以上の e, e', r, r', e_i を定める。

3. 論理 CODASYL モデル

3.1 従来の CODASYL モデルについて

本論文では、従来の CODASYL DBS 上に述語論理形式の操作演算を備えた関係型視野の設計を目指している。このためには、従来の CODASYL モデルの十分な論理的整理がまず必要になる。従来の CODASYL モデルとして、文献 1) のデータ構造と 5) のデータ操作言語 DML とを考察する。文献 15) 等も考えられるが、従来の CODASYL モデルの論理的に本質的な問題を検討するには、1) と 5) で十分であると考える。また、1) と 5) に準拠した DBMS (例、ADBS (日電)、AIM (富士通)) が、高性能大型 DBS に広く用いられているからもある。従来の CODASYL モデルは、物理的側面と論理的側面が未分離であるとともに、論理的な考察がほとんどなされていない。したがって、本章では、従来の CODASYL モデルから、データ構造を抽象化してモデル化し、これから論理基本演算を自然に導出して、論理 CODASYL モデルを定める。また、論理基本更新演算が従来の COBOL DML⁶⁾ によって実現できることを示し、本論文で定めるモデルと従来モデルの論理関係(等価性)を明らかにする。

3.2 論理データ構造

論理 CODASYL モデルの論理データ構造は、時間不变な論理構造を与える論理スキーム C と、これに実際にデータが与えられた(時間可変な)論理データベース C とから成る。 C はレコード (R) 型とセット (S) 型とから成る。 R 型 R は、項目集合 $\Omega_R = \{@R, t_1, \dots, t_m\} (m \geq 0)$ とインテグリティ条件 Γ_R から成る ($R = (\Omega_R, \Gamma_R)$)。各項目 τ の定義域を $\text{dom}(\tau)$ とする。 $@R$ はデータベース (db) キーで、 $\text{dom}(@R)$ は

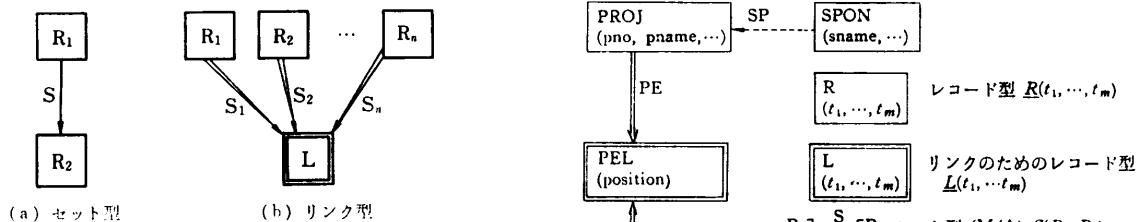


図 2 セット型とリンク型
Fig. 2 Set-type and link-type.

C 内で一意な識別子の集合である。各 t_i はデータ項目である ($i=1, \dots, m$)。 Ω_R に実際に値を与えた組のなかで Γ_R を満たす組（レコード実現値）の集合をレコード実現値(RO)集合 $R = \{r | r \in \text{dom}(\Omega_R) \wedge \Gamma_R(r)\}$ とする。 R 実現値 $r \in R$ の項目の部分集合 $I \subseteq \Omega_R$ で定まる値を $I(r)$ とする。各 $r \in R$ は一般に、データ項目値ではなく、 $@R$ によって識別され、 $@R(r)$ は C 内で一意である。よって、 $@R$ は R の主キーとなる。利用者に $@R$ 値の情報は与えられないで、 R 型を $R(t_1, \dots, t_m)$ (Γ_R は省略) とも書く。また各 $r \in R(r)$ の $@R(r)$ を $\delta(r)$ とも表す。

任意の R 型 R_2 から R_1 ($\neq R_2$) に部分関数が定義されるときに、これをセット (S) 型 $S = (R_1, R_2, \Gamma s)$ として表す(図 2(a)). R_1 と R_2 をおのおの、 S の親と子 R 型とする. Q_{R_1} と Q_{R_2} から意味的に帰納されるインテグリティ条件 Γs を満たす組 $(r_1 \in R_1, r_2 \in R_2)$ (セット (S) 実現値)* の集合をセット実現値(SO)集合 $S = \left\{ s \mid s \in \bigcup_{i=1}^2 \text{dom}(Q_{R_i}) \wedge \Gamma s(s) \right\}$ とする. S は部分関数を表すので、各 $r_2 \in R_2$ は $(r_1, r_2) \in S$ なるたかだか一つの $r_1 \in R_1$ を、また各 $r_1 \in R_1$ は $(r_1, r_2) \in S$ なる $h(\geq 0)$ 個の $r_2 \in R_2$ をもてる. 以下、SO 集合 S を R 実現値の主キー、すなわち db キーで表すことになると、 $S = \left\{ s' \mid s' = (\delta(r_1), \delta(r_2)) \wedge s = (r_1, r_2) \in \bigcup_{i=1}^2 \text{dom}(Q_{R_i}) \wedge \Gamma s(s) \right\}$ と表現できる. また $(\delta(r_1), \delta(r_2)) \in S$ で、 r_1 と r_2 をおのおの、親と子 R 実現値とする. S 型を $S(R_1$ と $R_2)$ (Γs は省略) とも表す.

任意の R 型 $R_1, \dots, R_n (n \geq 2)$, 異なっている必要はない) について, $\Omega_{R_1}, \dots, \Omega_{R_n}$ から意味的に帰納されるデータ項目集合 $\Omega_{R'}$ とインテグリティ条件 Γ_L とが定義され, 一般の関係 $L \subseteq \prod_{i=1}^n \text{dom}(\Omega_{R_i}) \times \text{dom}(\Omega_{R'})$

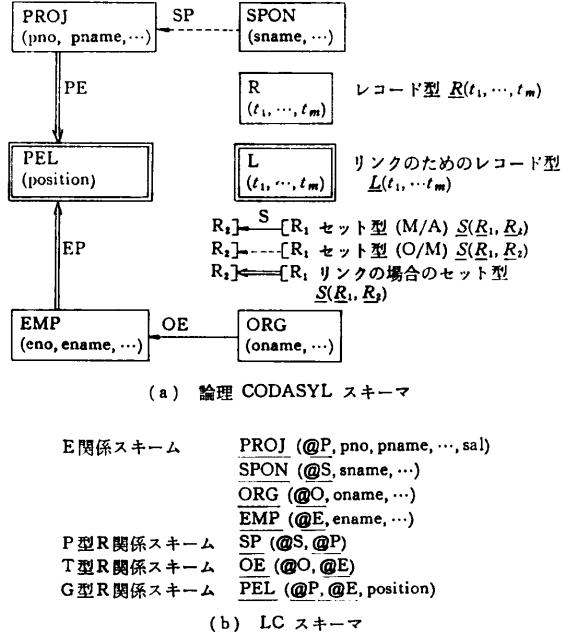


図 3 論理 CODASYL スキーマと LC スキーマ
Fig. 3 Logical CODASYL schema and LC schema.

が定義されるとする。このとき、これを、共通の子 R 型 $\underline{R}(\mathcal{Q}_R)(\mathcal{Q}_R = \mathcal{Q}_{R'} \cup \{@R\})$ をもつ n 個の異なる S 型 $\underline{S}_i(\underline{R}_i, \underline{R})(i=1, \dots, n)$ から成るリンク型 $\underline{L}=(\underline{R}_i, \underline{S}_i(\underline{R}_i, \underline{R}))(i=1, \dots, n), \underline{R}, \Gamma_L)$ と定義する(図 2 (b))。

$$\underline{L} = \left\{ l \mid l = (r_1, \dots, r_m, r) \in \prod_{i=1}^n \text{dom}(\mathcal{Q}_{R_i}) \times \text{dom}(\mathcal{Q}_R) \wedge \right.$$

$$\left. \Gamma_L(l) \wedge (\delta(r_j), \delta(r)) \in S_j \right\}$$

をリンク実現値 (LO) 集合とする。 S_j と \underline{R} をおのおの、リンクの場合の S 型と R 型とする。

以上の RO, SO, LO 集合の集合を論理データベース C とする。

S 型 $S(R_1, R_2)$ には、M/A と O/M の 2 種の型を定めれる(リンクの場合も含む)。これはメンバシップクラス¹³⁾の型であり、 $mc(S)$ と表す^{*}。この意味については、3.3 節で述べる。型付けされた論理スキーマ C では、*S* 型を有向辺、*R* 型を点とした有向閉路内のすべての *S* 型が M/A 型であってはならず、また、リンクの場合の *S* 型は M/A 型であるとする。図 3(a) にプロジェクトと構成員、所属、出資者を表す論理スキーマ例を示す。箱は *R* 型、二重線の箱はリンクの場合の *R* 型、実線と点線の矢はおのおの、M/A と O/M 型の *S* 型を、二重線の矢は、リンクの場合の *S*

* 文献 1) では、 S 実現値は、各 $r_1 \in R_1$ に対して、 r_1 と r_1 を親とする子 $r_2 \in R_2$ との集合である。

型(M/A型)である。

3.3 論理基本更新演算

論理データベース C 上の論理基本演算としては、検索と更新がある。検索演算は、文献 14)で詳論するので、本論文では、更新演算について論じる。まず、 C 内の任意の RO 集合を R とする。各 $r \in R$ に対して、そのすべての M/A 型 S 実現値の子 $r' \in R'$ の集合が定まる。同様に $r' \in R'$ の M/A 型 S 実現値の子 $r'' \in R''$ の集合が定まる。 r から始めて、このように再帰的に定義されていく全 R 実現値の集合を $M(r, R)$ (r も含める) とする。各 $r' \in M(r, R)$ を親または子とする全 S 実現値の集合を $O(r')$ とする。このとき、 $\text{ms}(r, R) \triangleq M(r, R) \cup_{\forall r' \in M(r, R)} O(r')$ とする。また、 C 内の各 R 型 R に対して、 R を子とするすべての M/A 型の S 型 $S'(R', R)$ と R' の対の集合を $Os(R)$ とする。さて、 $R, S(R_1, R_2), L = (R_i, S_i(R_i, R)) (i=1, \dots, m), R$ をおのおの、 C 内の任意の R, S, L 型とすると、次の 9 種の論理基本更新演算が C 上に導ける。

- (1) del(r, R) は C から $\text{ms}(r, R)$ を除く。(2) del($s=(k_1, k_2), S$) は、 $\text{mc}(S)=\text{M/A}$ ならば $\text{mc}(r_2, R_2)(\delta(r_2)=k_2)$ を、O/M ならば s のみを消去する。
- (3) del(l, L) は $l \in L$ および、各 S_i から l を子とする $s_i \in S_i$ を消去する。
- (4) app($r, R; r_1, R_1; \dots; r_n, R_n$) は、 (S_i, R_i) が $Os(R)$ に含まれ $r_i \in R_i$ のとき ($i=1, \dots, n$), R に r を、 S_i に $(\delta(r_i), \delta(r))$ ($i=1, \dots, n$) を加える。
- (5) app($s=(\delta(r_1), \delta(r_2)), S$) は、 $\text{mc}(S)=\text{O/M}$ で $r_1 \in R_1, r_2 \in R_2$ ならば、 S に s を加える。
- (6) app($l, L; r_1, R_1; \dots; r_n, R_n$) は R を L とした(4)と同じである。
- (7) rep(r, r', R) は $r \in R$ のデータ項目値を r' に変える。
- (8) rep($s=(k_1, k_2), s'=(k_1', k_2), S$) は、 $s \in S$ で $r_1' \in R_1(\delta(r_1')=k_1')$ ならば、 s を s' に変える。
- (9) rep(l, l', L) は $l \in L$ のデータ項目値を l' に変える。□

3.4 論理基本更新演算の DML 表現

従来の COBOL DML⁵⁾ (以降 DML と記す) の実行機構を抽象化し、われわれが示した論理データベース C 上で DML に対応した操作演算を実行できる抽象機械 M を定義する。また、論理スキーマ C は R 型 $R_i (i=1, \dots, n)$ と S 型 $S_j (j=1, \dots, m)$ から成るとする。 M は、 $3+2n$ 個のレジスタ $\mu, \nu, \kappa, \kappa_{R_i}, \beta_{R_i} (i=1, \dots, n)$ 、および記憶媒体 D をもつ。 μ と ν にはおのおの現在処理中の R 型と S 型についての情報と

して、その名前がはいっている。 κ と κ_{R_i} には、おののおの C と R_i 内で現在処理中の R 実現値の位置情報としての db キーがいれられる。 β_{R_i} には、 R_i の R 実現値のデータ項目値を記憶できる。レジスタ α の中身を (α) とする。 D には C の各時刻の C がはいる。 $\mu, \nu, \kappa, \kappa_{R_i}, \beta_{R_i}, D$ をおのおの、 $\mu, \nu, \kappa, \kappa_{R_i}, \beta_{R_i}, D$ のとりうる値の集合とし、 $K = \kappa \times \kappa_{R_i}, B = \beta_{R_i}$ とする。 m 上の演算 p は、関数 $p : \mu \times \nu \times B \times K \times D \rightarrow B \times K \times D$ である。 p としては以下のものがある。

- (i) レジスタ $\mu, \nu, \kappa, \kappa_{R_i}, \beta_{R_i}$ に値をセットする。ここでは κ と κ_{R_i} には、ある処理によって与えられるある条件式を満たす R 実現値の db キーがセットされているものとする。このセット方法については、文献 14) で別に論じる。
- (ii) $(\mu)=R_i$ のとき、 (β_{R_i}) をデータ項目値とする R 実現値に、新たな db キーを与える。これを κ と κ_{R_i} にセットして R_i に加える。
- (iii) $(\mu)=R_i$ のとき、 (κ) を db キーとする $r_i \in R_i$ を除く。同時に、 r_i を含む全 S 実現値を除き、 κ と κ_{R_i} をリセットする。
- (iv) $(\mu)=R_i, (\nu)=S_i$ のとき、 (κ) を db キーとしてもつ $r_i \in R_i$ を子とする S 実現値を S_i から除く。
- (v) $(\mu)=R_i, (\nu)=S_i$ のとき、 (κ) を db キーとする $r_i \in R_i$ を、 (κ_{R_i}) を db キーとする $r_i \in R_i$ の子として S_i に加える。
- (vi) (κ) を db キーとする $r_i \in R_i$ のデータ項目値を (β_{R_i}) とする。□

論理基本更新演算(3.3節)を、以上の(i)～(vi)で表せることは示せる。この詳細は、文献 14)で示す。(ii)～(vi)はおのおの DML の store R_i , erase R_i , disconnect R_i from S_i , connect R_i to S_i , modify R_i に対応している。また、DML の modify R_i only S_i は(iii)と(iv)、M/A 型 S 型についての store R_i , erase R_i permanent はおのおの、(ii)と(v)と、 $\text{ms}(r_i, R_i)$ 内の各 R 実現値に(iii)を繰り返し適用することによって表せる。よって、論理基本更新演算は、表 1 に示すように DML で表せる。たとえば、del(r, R) は、(1)まず κ に消去すべき $r \in R$ の db キーを記憶し $(\kappa \leftarrow \delta(r))$ 、(2) erase R permanent で $\text{ms}(r, R)$ を除くことで表せる。また app($r, R; r_1, R_1; \dots; r_n, R_n$) は、(1)追加する r のデータ項目値を β_R に記憶し、(2)各 $r_i \in R_i$ の db キーを κ_{R_i} に記憶し $(\kappa_{R_i} \leftarrow \delta(r_i))$ 、(3) store R で R に r と、各 S_i に $((\kappa_{R_i}), \delta(r))$ とを加えることで表せる。以上より次の命題が成り立つ。

表 1 CODASYL 基本演算の COBOL DML 表現
Table 1 Primitive operations and COBOL DML in the CODASYL model.

基 本 演 算	COBOL DML 表 現
(1) <u>del</u> (r, R);	$[\kappa \leftarrow \delta(r);]$ <u>erase</u> R permanent.
(2) <u>del</u> (s, S); ここで $s=(k_1, k_2)$	$mc(S)=M/A[\kappa \leftarrow \delta(r);]$ <u>erase</u> R_1 permanent $mc(S)=O/M[\kappa \leftarrow \delta(r);]$ <u>disconnect</u> R_2 from S .
(3) <u>del</u> (L, L);	$[\kappa \leftarrow \delta(L);]$ <u>erase</u> L .
(4) <u>app</u> ($r', R; r_1 \in R_1; \dots; r_n \in R_n$); (6) R はレコードまたはリンク型	$[\kappa_{R_i} \leftarrow \delta(r_i); (i=1, \dots, n)]$ $[\beta_R \leftarrow r';]$ <u>store</u> R .
(5) <u>app</u> (s, S);	$mc(S)=O/M[\kappa_{R_i} \leftarrow \delta(r_i); \kappa \leftarrow \delta(r_s);]$ <u>connect</u> R_2 to S .
(7) <u>rep</u> (r, r', R); (9) R はレコードまたはリンク型	$[\kappa \leftarrow \delta(r); \beta_R \leftarrow r';]$ <u>modify</u> R including S_1, \dots, S_n .
(8) <u>rep</u> (s, s', S); ここで $s=(k_1, k_2)$, $s'=(k'_1, k'_2)$	$[\kappa_{R_i} \leftarrow k'_1; \kappa \leftarrow k_2; \kappa \leftarrow k_3;]$ <u>modify</u> R_2 only S .

* 置換されるデータ項目のどれかがセット型 S_1, \dots, S_n のソートまたは DNA 項目である。CALC と項目値の置換はできない。

[命題 3.1] すべての論理基本更新演算は、従来の COBOL DML で表せる。■

従来の COBOL DML⁵⁾ の厳密な意味については、文献 14) で多テープチューリング機械によって表す。

4. 概念モデルによる論理 CODASYL モデル表現

本章では、概念モデルの R スキームを特徴化し、それに従って導かれる基本更新演算を定義し、これを用いてローカル概念 (LC) モデルを与える。また、論理 CODASYL モデルと本 LC モデルとが等価であることを示す。

4.1 LC データ構造

概念スキーム \underline{S} 内の各 R スキーム R を以下のように特徴化したものをローカル概念 (LC) スキーム \underline{L} とする。(1) R スキーム R に、F と G との 2 種の型を設ける。 R が E スキーム E_2 から E_1 への部分関数を表すとき、 R を F 型とする。関数の定義域を表す $@E_2$ が主キーとなり、 $R(@E_1, @E_2)$ と表す。ここで、下線を引いた属性は、主キーを表すとする。 $R = (E_1, \dots, E_n, Q_R)$ が E_1, \dots, E_n 間の一般の関係を表すとき、 R を G 型とする。各 E_i の主キー集合 $\{@E_1, \dots, @E_n\}$ の値によって R 内の各組が定まるので、これ

を主キーとし、 $R(@E_1, \dots, @E_n, Q_R')$ と表す。(2) F 型 R にはさらに、 T (全) または P (部分) の型が与えられ、おののおの更新演算に異なった意味を与える(4.2 節)。□

\underline{L} に実際にデータを与えたものを LC データベース \underline{L} と定義する。

LC と論理 CODASYL データ構造は次のように対応する。(1) R 型 $R(t_1, \dots, t_m)$ と E スキーム $R(@R, t_1, \dots, t_m)$ に対しては、データ項目 t_i と属性 t_i , db キーと $@R$ とが対応する。(2) S 型 $S(R_1, R_2)$ には、F スキーム $S(@R_1, @R_2)$ が対応する。 $@R_i$ は R 型 R_i に対応した E スキームの主キーである($i=1, 2$)。 $S(@R_1, @R_2)$ は、S 型 S が M/A のとき T 型に、O/M のとき P 型になる。(3) L 型 $L=(R_1, S_i(R_1, R), (i=1, \dots, n), R(b_1, \dots, b_q))$ は G スキーム $L(@R_1, \dots, @R_n, b_1, \dots, b_q)$ に対応する*。 $@R_i$

は R 型 R_i に対応したスキームの主キーである。 $(i=1, \dots, n)$ 。□

以上より、次の命題が成り立つことは明らかである。

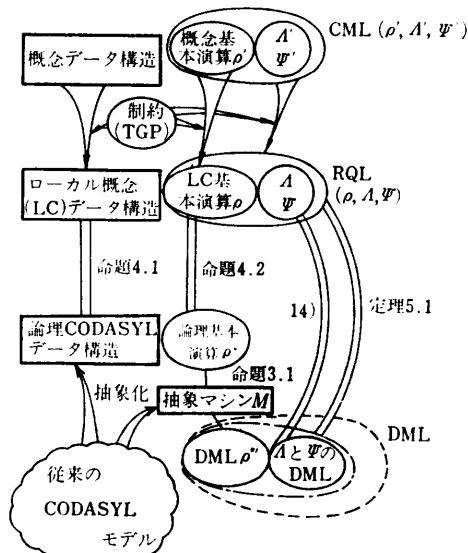


図 4 CODASYL モデルと概念モデル
Fig. 4 CODASYL model and conceptual model.

* $R_i=R_j$ ($i \neq j$) の場合には、 $S_i(@R_i, S_j @R_j)$ を属性名とする。

[命題 4.1] LC データ構造は、論理 CODASYL データ構造と 1 対 1 対応する。■

図 3(b)は、(a)の LC スキーマを示している。

4.2 LC 操作言語 RQL

LC モデルの述語論理形式の操作言語 RQL (relational query language) は、目標リスト Λ 、条件式 Ψ 、および基本演算 ρ によって LC データベース上の演算を表す。 ρ としては基本検索演算と基本更新演算がある。 Λ と Ψ は 2.3 節で述べたものと同じである。本節では、まず基本更新演算を述べ、次に RQL 更新演算表現を述べる。

4.2.1 基本更新演算

R と R をおのおの任意の E スキームとその関係とする。命題 4.1 より、LC と論理 CODASYL データ構造は 1 対 1 対応するので、3.2 節の $ms(r, R)$ と $Os(R)$ とおのおの等価な集合として $ts(r, R)$ と $Ts(R)$ を自然に定義できる。 R スキームが P か G 型だから成る場合には、2.3.2 項の概念基本更新演算と同じである。T型の R スキームを含む場合には、次の LC 基本更新演算が自然に存在する。(i) $\underline{del}(r, R)$ は $ts(r, R)$ を消去する。(ii) $\underline{del}(s, S)$ は、 S が T型のとき $@R_2(s) = @R_2(r_2)$ なる $ts(r_2, R_2)$ を消去する。(iii) $\underline{app}(r, R; r_1, R_1; \dots; r_n, R_n)$ は、 (S_i, R_i) が $Ts(R)$ に含まれ $r_i \in R_i$ ($i = 1, \dots, n$) のとき、 R に r を加えるとともに $@R_i(r_i) = @R_i(s_i) \wedge @R(s_i) = @R(r)$ なる s_i を S_i に加える ($i = 1, \dots, n$)。

以上より、次の命題が成り立つことは明らかである。

[命題 4.2] LC モデルと論理 CODASYL モデルの基本更新演算は、1 対 1 対応する。■

[定理 4.1] LC モデルと論理 CODASYL モデルとは、データ構造と基本演算について 1 対 1 対応し、等価である。

[証明] 命題 4.1 と 4.2 より互いのデータ構造と基本更新演算とが等価であるので、明らかである。■

4.2.2 RQL 更新演算

RQL による更新演算表現について述べる。ここで、 x をある組変数集合とし、 Λ_x と Ψ_x をおのおの、 x を参照する目標リストと条件式とする。

A. 消去演算

X を \mathcal{L} 内の消去対象の関係スキームとし、 X を任意の関係スキームとする。関係 X への消去演算は RQL で、次のように書かれる。

$$\left. \begin{array}{l} \underline{\text{var}}\ (u, X)\ (x_1, X_1) \cdots (x_l, X_l); \\ \underline{\text{delete}}\ u \ \underline{\text{where}}\ \Psi_x; \\ (x = \{u, x_1, \dots, x_l\}) \end{array} \right\} \quad (5)$$

(5) は、 $t \in X \wedge_{i=1}^l x_i \in X_i \wedge \Psi_x(t, w_1, \dots, w_l)$ なる組 $t \in X$ について、LC 基本更新演算 $\underline{del}(t, X)$ することを表している。たとえば、 X が E型ならば $ts(t, X)$ が消去される。

B. 追加演算

(i) E を追加対象の E 関係スキームとする。 E は n 個の T型 $R_i(@E_i, @E)$ で E 関係スキーム E と関連しているとする ($Ts(R) = \{(R_i, E_i) | i = 1, \dots, n\}$)。また、 X を \mathcal{L} 内の任意の関係スキームとする。 E への追加は RQL で次のように書かれる。

$$\left. \begin{array}{l} \underline{\text{var}}\ (e_1, E_1) \cdots (e_n, E_n)\ (x_1, X_1) \cdots (x_l, X_l); \\ \underline{\text{append}}\ \underline{\text{to}}\ E(\Lambda_x) \ \underline{\text{where}}\ \Psi_x; \\ (x = \{e_1, \dots, e_n, x_1, \dots, x_l\}) \end{array} \right\} \quad (6)$$

(6) は、 $w = (t_1, \dots, t_n, w_1, \dots, w_l) \wedge_{i=1}^n t_i \in E_i \wedge_{j=1}^l w_j \in X_j \wedge \Psi_x(w)$ なる w に対する $t' = \Lambda_x(w)$ を求め、組 (t_1, \dots, t_n, t') について $\underline{app}(t', E; t_1, E_1; \dots; t_n, E_n)$ を行うことを表す。

(ii) R を追加対象の R 関係スキームとし、 X を任意の関係スキームとする。 R に対する追加演算は、RQL で次のように表される。

$$\left. \begin{array}{l} \underline{\text{var}}\ (x_1, X_1) \cdots (x_l, X_l); \\ \underline{\text{append}}\ \underline{\text{to}}\ R(\Lambda_x) \ \underline{\text{where}}\ \Psi_x; \\ (x = \{x_1, \dots, x_l\}) \end{array} \right\} \quad (7)$$

(7) は、 $w = (w_1, \dots, w_l) \wedge_{j=1}^l w_j \in X_j \wedge \Psi_x(w)$ なる w について、 $t' = \Lambda_x(w)$ を求め、 $\underline{app}(t', R)$ を行うことを表す。

C. 置換演算

X を置換対象の関係スキーム、 X を任意の関係スキームとする。 X への置換演算は RQL で次のように表される。

$$\left. \begin{array}{l} \underline{\text{var}}\ (u, X)(x_1, X_1) \cdots (x_l, X_l); \\ \underline{\text{replace}}\ u(\Lambda_x) \ \underline{\text{where}}\ \Psi_x; \\ (x = \{u, x_1, \dots, x_l\}) \end{array} \right\} \quad (8)$$

(8) は、 $w = (t, w_1, \dots, w_l) \wedge t \in X \wedge_{j=1}^l w_j \in X_j \wedge \Psi_x(t, w_1, \dots, w_l)$ なる w に対する $t' = \Lambda_x(w)$ を求め、 (t, t') について $\underline{rep}(t, t', X)$ を行うことを表している。すなわち、 $t \in X$ の非主キー値を t' に置換する。

5. 非手続的インターフェース設計

CODASYL DBS 上に、LC モデルの視野と非手続的操作言語 RQL とを利用者に提供するインターフェースを設けるためには、LC データ構造上の RQL 演算を、COBOL DML で実現する必要がある。本章では、RQL を COBOL DML で表せることを示す。まず命題 3.1 (表 1) と命題 4.2 より次の命題が成り立つことは明らかである。

[命題 5.1] すべての LC 基本更新演算を COBOL DML で表現できる。■

次に、4.2.2 項で述べた RQL 更新演算について考える。関係 X に対する RQL 演算が組変数集合 $x = \{y_1, \dots, y_n\}$ (y_i は関係 Y_i の組変数) を参照するとき、この RQL 演算の意味は、次の(9)の形式で表せる。

```
for  $\forall v(v=\lambda_x(w) \wedge \Psi_x(w) \wedge w \in Y_1 \times \dots \times Y_n)$ 
    do op (v, X);      (9)
```

op は LC 基本更新演算子である。条件式 Ψ_x を満足する組 w について、目標リスト λ_x はある目標属性集合 Ω 値の組 v を与える。たとえば 4.2.2 項の B

(i) の $append$ では、 λ_x は組 w から組 (t_1, \dots, t_n, t') を与える。(9)の for...do は更新データ v の集合を導出する目標リスト λ_x と条件式 Ψ_x から成る RQL 検索演算として表せる。RQL 検索演算を DML で正しく表す方法は示せるが、この手順は別¹⁴⁾に論じる。またこの手順に従って設計されたシステムは LDP-V 1.5^{8),10)} として実働している。さて、(9)の LC 基本演算 $op(v, X)$ を命題 5.1 より DML で表せるので、RQL 更新演算の DML 表現を表 2 のように示せる。よって、次の定理が成り立つ。

[定理 5.1] すべての RQL 演算を COBOL DML で表現できる。■

例として図 3 (b)を考えると、消去演算「Jipdec を除け」は RQL で(10)のように表される。

```
var(o, ORG);
delete o where o.oname="Jipdec"; (10)
```

(10)は、表 2 より COBOL DML で表現すると、次の手順に対応する。(i) $oname(t) = "Jipdec"$ なる $t \in ORG$ を見つけ、 t の db キーを κ に記憶し、(ii) $erase ORG permanent.$ を行う。これによって、Jipdec の従業員も同時に消去される。

表 2 RQL の COBOL DML 表現
Table 2 COBOL DML description of RQL operations.

RQL	RQL の COBOL DML 表現
A. 消去	$T \leftarrow \{t w=(t, w_1, \dots, w_n) \in X \times X_1 \times \dots \times X_n \wedge \Psi_x(w)\};$ for $\forall t \in T$ do (i) X は E 型 $\kappa \leftarrow @X(t); erase X permanent. od;$ (ii) X は B 型 $X(@E_1, \dots, @E_n)$ (ii-1) X は T 型 $\kappa \leftarrow @E_1(t); erase E_1 permanent. od;$ (ii-2) X は P 型 $\kappa \leftarrow @E_1(t); disconnect E_1 from X. od;$ (iii) X は G 型 $X(@E_1, \dots, @E_n, Q_x')$ $\kappa \leftarrow t \in X$ の db キー; $erase X. od;$
B. 追加	(1) E 関係 E への追加 $E(@E, Q_E')$ $var(e_1, E_1) \dots (e_n, E_n)$ $(x_1, X_1) \dots (x_n, X_n);$ $append to E(\lambda_x) where \Psi_x;$ $(x = \{e_1, \dots, e_n, x_1, \dots, x_n\})$ (2) R 関係 X への追加 $var(x_1, X_1) \dots (x_n, X_n);$ $append to X(\lambda_x) where \Psi_x;$ $(x = \{x_1, \dots, x_n\})$
C. 置換	$T \leftarrow \{(t, t') w=(t, w_1, \dots, w_n) \in X \times X_1 \times \dots \times X_n \wedge \Psi_x(w) \wedge t' = \lambda_x(w)\};$ for $\forall (t, t') \in T$ do (i) X は E 型 $X(@X, Q_x')$ $\kappa \leftarrow @X(t); \beta_X \leftarrow Q_x'(t'); \kappa \leftarrow @E_1(t) (i=1, \dots, n); store E. od;$ (ii) X は B 型 $X(@E_1, \dots, @E_n)$ $\kappa_{E_1} \leftarrow @E_1(t); \kappa_{E_1} \leftarrow @E_1(t'); connect E_1 to X. od;$ (iii) X は G 型 $X(@E_1, \dots, @E_n, Q_x')$ $\kappa_{E_1} \leftarrow @E_1(t) (i=1, \dots, n); \beta_X \leftarrow Q_x'(t); store X. od;$
	$T \leftarrow \{(t, t') w=(t, w_1, \dots, w_n) \in X \times X_1 \times \dots \times X_n \wedge \Psi_x(w) \wedge t' = \lambda_x(w)\};$ for $\forall (t, t') \in T$ do (i) X は E 型 $X(@X, Q_x')$ $\kappa \leftarrow @X(t); \beta_X \in Q_x'(t'); modify X. od;$ (ii) X は B 型 $X(@E_1, \dots, @E_n)$ $\kappa_{E_1} \leftarrow @E_1(t'); \kappa \leftarrow @E_1(t); modify E_1 only X. od;$ (iii) X は G 型 $X(@E_1, \dots, @E_n, Q_x')$ $\kappa \leftarrow t \in X$ の db キー; $\beta_X \leftarrow Q_x'(t'); modify X including S_1, \dots, S_n. od;$

追加演算「鈴木を（田中と同所属で）EMPに加えよ」は RQL で(11)のように書かれる。

```
var(o, ORG) (e, EMP)( $\alpha$ , OE);
append to EMP (e. name = "鈴木", ...)
where e.ename = "鈴木" and e.@E
=  $\alpha$ .@E and  $\alpha$ .@O = o. @O; (11)
```

(11)は COBOL DML で表すと次の手順に対応する。
(i) $t_1 \in \text{ORG} \wedge w_1 \in \text{EMP} \wedge \text{ename}(w_1) = \text{"田中"}$
 $\wedge (\delta(t_1), \delta(w_1)) \in \text{OE}$ なる t_1 を求め, κ_{ORG} に $\delta(t_1)$ を記憶し, (ii) β_{EMP} に追加する R 実現値 ("鈴木", ...) を記憶し, (iii) store EMP. を行う。これによって, EMP に "鈴木" を加えるとともに, 鈴木を田中と同じ所属にする S 実現値を SO 集合 OE に加えることができる。

6. む す び

本論文では, DDBS の統合化問題の解として, DDBS 全体の共通モデルに基づいて, 従来の CODASYL DBS 上に視野を設計するための基本的問題を論じた。まず, 既存のデータモデル(階層, 網)を抽象化し, これらのモデルの(静的)なデータ構造を表せるモデルとして, 事象と関連概念を有した概念モデルを共通モデルとして定めた。次に, 従来の CODASYL モデルを抽象化した論理 CODASYL データ構造と, それから自然に導ける論理基本更新演算とによって論理 CODASYL モデルを定義した。この論理モデルは, 従来の CODASYL モデルより上位にある。また, 論理基本更新演算を, 従来の DML で表せることを示し, 両モデルの論理的関係を明らかにした。次に, 概念モデルの R 關係スキームを型づけし, 型によって基本更新演算を定めて, ローカル概念(LC) モデルとし, この LC モデルと論理 CODASYL モデルの等価性を示した。さらに, 述語論理に基づいた LC 操作言語 RQL を, LC 基本更新演算に述語論理形式の条件式 Ψ と目標リスト A を付加したものとして定義した。 Ψ と A を DML で表せること^{8), 10), 14)}と, LC 基本更新演算を DML で表せることから, すべての RQL を DML で表せることを示した。以上によって, 従来の CODASYL モデルの論理記述と, 従来の CODASYL DBS 上に述語論理形式の操作言語を備えた視野を提供する論理的基礎とを与えることができたと考える。

以上の設計概念は, すでにローカルデータベースプロセッサ (LDP-V 2) として当協会の AIM (M-170 F)

と ADBS (Acos-700) 上で実働している。LDP-V 2 は PL/I で作成されている(ソース約 15,000 文, 実行形式 250 kB)。現在, 本システムの性能面と利用面からの総合評価を行っているが, 数行の RQL の入力によって, 複雑なデータベース保守を行う 1,000~3,000 行の COBOL DML プログラムを 1 分以内の経過時間 (M-170 F) で生成できる。LDP は DDBS の共通インターフェースとなるとともに, 既存 CODASYL DBS の非定型利用インターフェースおよび COBOL DML プログラムの開発支援ツールとなると考えている。

参 考 文 献

- 1) CODASYL DDL Committee: CODASYL Data Description Language, J. Dev., (1973) (藤中他訳: CODASYL データベース用記述言語, p. 250, 情報処理学会 (1977)).
- 2) Codd, E. F.: A Relational Model of Data for Large Shared Data Banks, CACM, Vol. 13, No. 6, pp. 337-387 (1980).
- 3) Date, C. J.: An Introduction to Database Systems, 3rd ed., p. 574, Addison-Wesley, Reading (1981).
- 4) Held, G. et al.: INGRES—A Relational Data Base System, AFIPS Conf. Proc., pp. 409-416 (1976).
- 5) Olle, T. W.: The CODASYL Approach to Data Base Management, p. 287, John Wiley & Sons, New York (1978) (西村, 植村他訳: CODASYL のデータベース, 共立出版, 東京 (1975)).
- 6) Takizawa, M., Hamanaka, E. and Ito, T.: Resource Integration and Data Sharing on Heterogeneous Resource Sharing System, Proc. ICCC '78, pp. 253-258 (1978).
- 7) Takizawa, M. and Hamanaka, E.: The Four-Schema Concept as the Gross Architecture of Distributed Databases and Heterogeneity Problems, JIP (IPSJ), Vol. 2, No. 3, pp. 134-142 (1979).
- 8) Takizawa, M. and Hamanaka, E.: Query Translation in Distributed Databases, Proc. IFIP '80, pp. 451-456 (1980).
- 9) Takizawa, M.: Distribution Problems in Distributed Databases-Integration and Query Decomposition, JIP (IPSJ), Vol. 5, No. 3, pp. 139-147 (1982).
- 10) 滝沢 誠, 横塚 実, 鈴木 信: CODASYL データベースシステムに対する関係インターフェースシステム—LDP-V 1.5—の設計と実現について, 情報処理学会論文誌, Vol. 23, No. 6, pp. 665-675 (1982).

- 11) Smith, J. M. and Smith, D. C. P.: Database Abstractions: Aggregation and Generalization, *ACM TODS*, Vol. 2, No. 2, pp. 105-133 (1977).
- 12) Wong, E.: Logical Design and Schema Conversion for Relational and DBTG Databases, *Proc. E-R Conf.*, pp. 311-321 (1979).
- 13) Zaniolo, C.: Design of Relational Views over Network Schemas, *Proc. SIGMOD*, pp. 177-190 (1979).
- 14) 滝沢 誠, 野口正一: 非手続的グラフ問合せの DML 表現について, (準備中).
- 15) CODASYL DDL Committee: Report of the CODASYL Data Description Language Committee, *Inf. Syst.*, Vol. 3, No. 4, pp. 247-320 (1978).
- 16) 増永良文: 関係スキーマ上の CODASYL インタフェース設計とその応用, 情報処理学会データベース管理システム研究会資料 29 (1982).
(昭和 57 年 12 月 2 日受付)
(昭和 58 年 6 月 20 日採録)