

カスタム化・進化可能なデータベース管理システムについて

宝珍 輝尚

NTT情報通信網研究所

概要

データモデルのカスタム化と進化について議論する。データモデルのカスタム化とは応用毎にデータモデルを専用化することであり、データモデルの進化とはデータベースのロード／アンロードなしにデータモデルを変更できることである。まず、データモデルのカスタム化と進化の必要性を示す。次に、現在の研究動向を概説し、研究課題を示す。最後に一実現方法を示す。

On Customizable and Evolvable Database Management Systems

Teruhisa Houchin

NTT Network Information Systems Laboratories
1-2356, Take, Yokosuka-shi, Kanagawa 238-03 Japan

Abstract

This paper describes data model customization and evolution. Data model customization means that a data model can be altered according to an application. Data model evolution is the ability that a data model can be changed without loading / unloading a database. First, requirements of these are shown. Next, current research efforts are surveyed and research topics are listed. Finally, an implementation is presented.

1 はじめに

DB 応用分野の急激な広がりや急速な進歩により従来の汎用 DBMS では各々の分野に十分対応できなくなっている。これを解決するために拡張可能 DBMS やオブジェクト指向 DBMS の研究が行われ、データ型やデータアクセス方法の拡張可能化が研究されている。

また、同じ背景からデータモデルの拡張可能化が要求されている。事実、次世代 DBMS としてのオブジェクト指向 DBMS では应用到に依存する様々なデータモデリングのためにデータモデルの拡張性が必要であると指摘されており [1]、オブジェクト指向データモデルがネットワークデータモデルや関係モデルから発展してきたことも示されている [2]。データモデルの拡張可能化については、いくつかの研究が行われ [3-8]、一部には製品も出てきている [9] が、研究は緒についたばかりである。

そこで、本稿ではデータモデルの拡張可能化について体系化した議論を試みる。まず、2. でデータモデルの拡張可能化の要求をカスタム化と進化の両面からまとめる。次に、3. で研究動向と問題点を示す。4. ではデータモデル、実装方法、DB の標準化の観点から検討課題を示す。最後に、5. で一実現方法と今後の課題について述べる。

2 要求

CAD、VLSI 設計やネットワーク管理といった先進的 DB 応用分野からは様々な要求がある。

- 高機能化 複雑なデータを素直にモデリングしたり、版管理を行いたい。
- 高性能化 複雑なデータの高速なアクセスや高スループットを得たい。

- 迅速な DB 構築 時期を逃さずに応用システムを構築したい。
- 柔軟化 DB 設計時に完全に仕様が決定できない場合や当初の設計では対処できなくなる場合があるので、将来の変更に対して十分に対処可能であってほしい。
- ノンストップ化 システムを停止せずに運用したい。

これらの要求をデータモデルの拡張可能性という観点から考える。まず必要なのは、DB 応用毎にデータモデルを専用化したいというデータモデルのカスタム化である。これには、特別な機能を付加したいという機能追加要求と高性能や高柔軟性を得るために不要な機能を削除したいという機能削除要求が考えられる。

さらに、このカスタム化は一度設定したら変更できないような固定的なものではなく、柔軟化、ノンストップ化の要求から変更可能であることが望まれる。ここでは、迅速にデータベースシステム (DBS) を変更できるように、DB のロード/アンロードなしにデータモデルを変更してゆく機能 (データモデル進化機能) が必要である。データモデルの進化にはオフラインとオンラインでの進化が考えられるが、ノンストップ運転の場合にはオンラインでの進化が不可欠である。

3 従来の研究と問題点

3.1 研究動向

データモデルコンパイラ (DMC) [3-5] は、データモデルの仕様からの DB 設計ツールの自動生成を目標としている。ここではデータモデルをメタ実体とメタ関連によって表現する。メタ実体では、属性の型、データ演算の型、デー

タ制約の型とグラフィカルな表現を指定できる。メタ関連では、メタ実体で指定可能な項目の他に、関連に関与する実体、該実体の関連における役割や継承の指定を行うモデル演算集合、カーディナリティ制約や一貫性制約等を指定するモデル制約を指定できる。

KIDS[6]は、DMCと同様に、仕様からDBMSを導出するアプローチを採用している。データモデルの記述には実体-属性-値(EAV)モデルが採用されている。EAVモデルで記述された仕様をもとにreduction rulesとclustering rulesにより物理的な格納構造が決定される。物理的な実体はオブジェクトサーバにより操作される。また、KIDSではトランザクションの原子性に関する考察も行われている。

多言語DBS[7]のDBMSカーネルでは、属性に基づくデータモデル(ABDM)を使用している。ネットワークデータモデルや関数データモデルに存在するような関連もABDMの枠組みの中で記述されている。新データモデルを実現するには、該データモデルからABDMへの写像プログラムを実装する必要がある。

MATISSE[9]はデータモデルのカスタム化が可能な商用DBMSである。MATISSEでは、内部データ管理モデルに値と関連のみという単純なMICRO-MODELを採用し、データモデルをテンプレートとして提供する。

Cooperは、データモデルならびにユーザインタフェースが形成可能なシステムを永続的プログラミング言語PS-algolを使用して実現しようとしている[8]。DB利用者は利用開始時に使用するデータモデルとユーザインタフェースを指定する。データモデルは、基底クラス、(強、弱)複合クラス、関連(属性、集約化、set

of, 汎化、特化)、演算ならびに制約をメニューから選択して指定する。これにより、あらかじめ用意された一種の雛形に、指定された項目に相当するコードが埋め込まれ、カスタム化されたデータモデルが実現される。また、データの格納等を行うデータ管理部では上記の特徴を持つ意味的なデータモデルをサポートしている。データモデルやユーザインタフェースの変更は、PS-algolのコンパイラ呼び出し機能を使用して動的に自プログラムをリコンパイルし再リンクすることで実現されている。

この他に、マルチデータモデルの研究[10, 11]が行われている。これらの研究では、概念スキーマのモデリングツールとしての意味的なデータモデルから実働システムが存在するネットワーク/関係データモデルへの変換を目的に相異なるデータモデルのスキーマ間の変換可能性が議論され、データモデル変換のためにハイパーグラフ[10]やERモデル[11]が使用されている。

3.2 従来の研究の問題点

前節で示した研究はデータモデルのカスタム化には有効である。しかしながら、Cooper以外のアプローチではDBS構築時にデータモデルが決定かつ固定されてしまう。従って、オフラインでしかデータモデルを進化できず、ノンストップのシステムには適用できないという問題がある。

一方、Cooperのアプローチでは、データモデルの記述に基底クラス、2種の複合クラス、5種の関連ならびに制約を使用しているが、これらのみで将来のデータモデルが記述可能か疑問である。例えば、時間のモデリングはこの枠組みでは記述困難と考えられる。

4 研究課題

ここでは、データモデル、実装方法と標準化等に分けて研究課題を示す。

4.1 データモデル

(1) データモデルの構成要素の明確化

データモデルを記述するにはデータモデルの構成要素を明確化する必要がある。実体や関連の記述に関しては多くの研究がなされている [5-10] が、制約の記述についての研究はまだ不十分であり、振舞いの記述についてはほとんど研究されていない。

(2) データモデルの記述方法の検討

データモデルを利用者に開放するためには、データモデルの記述方法を提供しなければならない。DMCやKIDSでは仕様(すなわちデータモデル記述言語)を提供している [3, 6]。メニューによる選択方式 [8] も考えられている。

(3) データモデル記述能力の明確化

(1),(2)によりどの程度データモデルが記述可能かを明確化できると良い。しかしながら、定量的な比較や評価は困難であろう。

4.2 実装方法

(1) カーネル実現方法の検討

従来の研究からも明らかなように単純化したDBMSカーネルの上にデータモデル依存部を構築するアプローチが妥当と考えられる。これはオブジェクト指向DBMSにおける川越の指摘 [12] とも合致する。DBMSカーネルは、あらゆるデータモデルのデータをも操作できる汎用で基本的な機能を持つ必要がある。また、DBMSカーネルの提供するサービスは固定的でなく拡張/カスタム化可能でなければならない。さらに、データモデル進化のために、既

存データと新たなデータ間の整合を考慮しなければならない。

(2) スキーマ管理方法の検討

データモデルのカスタム化と進化をサポートすると、1DBを複数のデータモデルにより使用したいという要求も考えられる。このためのスキーマの管理方法を検討する必要がある。また、スキーマでは新たに追加されたデータモデリング機能に対する情報を管理しなければならない。このために、データモデルの進化に伴って自動的に変更可能なスキーマ管理方法を検討する必要がある。

(3) DBMSの他の構成要素との整合性の検討

スキーマ管理以外にも、データモデルが拡張されるとその影響はDBMSの様々な所に現れる。例えば、利用者インタフェースである。言語に基づくものであれ視覚的なものであれ、新データモデリング機能を扱う利用者インタフェースを進化させる必要がある。

(4) 動的な変更方法の検討

オンラインでデータモデル進化を可能とするためには、動的にシステムを変更してゆく機能が必要である。PS-algolのコンパイラ呼び出し機能は該機能実現のために使用できる機能の一つと考えられる。また、既存AP(変更時に実行途中の可能性はある)との整合性をも考慮しなければならない。更に、変更が既存システムと整合しているか試験できるデータモデル変更環境を考えなければならない。

4.3 DBの標準化・データの流通

(1) DBの標準化

従来、DBの標準化はデータモデル(もしくはデータベース言語)のレベルで行われてきた。データモデルのカスタム化や進化が可能になる

と従来のレベルではDBは標準化できない。DBの標準化のためにはDBMSカーネルの標準化等を考えてゆく必要がある。

(2) データ流通への対処

異種DBS間でデータを流通させるための研究が行われている[13, 14, 6]。これらの研究において、本稿で考えているデータモデルのカスタム化や進化を考慮する必要がある。

5 一実現方法と今後の課題

ここでは、カスタム化・進化可能なDBMSの一実現方法として拡張可能DBMS:COMMONのアプローチと今後の課題を示す。

DBMSカーネルのデータモデルにデータグラフ[16]を採用したCOMMONの構成を図1に示す。データグラフはグラフに基づいてデータを論理的に表現する。応用毎に異なるデータモデルは統一したデータグラフに変換する。データグラフで表現されたデータは、論理-物理変換手続きにより応用に応じた特性が得られるようなデータ構造に変換され二次記憶媒体に格納される。DBMSカーネルにグラフという汎用的な構造表現を採用することで、あらゆる

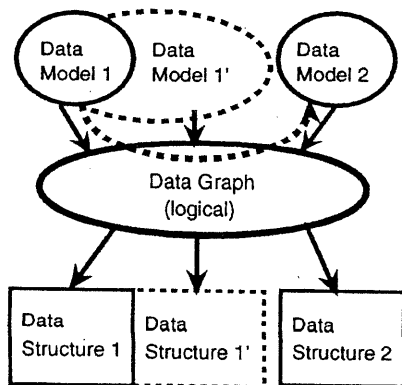


図1 COMMONのデータモデルとデータグラフ

データを表現できるのではないかと考えている。また、論理-物理変換手続きはDBMSカーネルの拡張性を実現している。主なデータモデルからデータグラフへの変換を試みている[17]が、これは構造と演算についての変換方法のみであり、データモデルの持つ意味と制約の表現の検討が必要である。この表現には、手続き的に動作を記述する振舞いによる記述方法や宣言的に正常な状態を記述する制約による記述方法が考えられる。ここでは、データグラフの定義情報から他データモデルのスキーマを導出可能とするために、データモデルとデータグラフ間の双方向変換を考慮する必要がある。

次に、COMMONにおけるデータモデルの定義と問合せ処理について述べる。データモデルの定義情報はメタスキーマとして格納する。また、メタスキーマ上でのデータグラフへの変換指定も格納する。データ定義情報はスキーマとして格納し、データ定義の際に変換指定を基にスキーマとデータグラフとの対応情報を作成し格納する。このスキーマ・データグラフ対応情報は問合せ時に問合せ表現の変換に使用する。この概要を図2に示す。本アプローチでは、

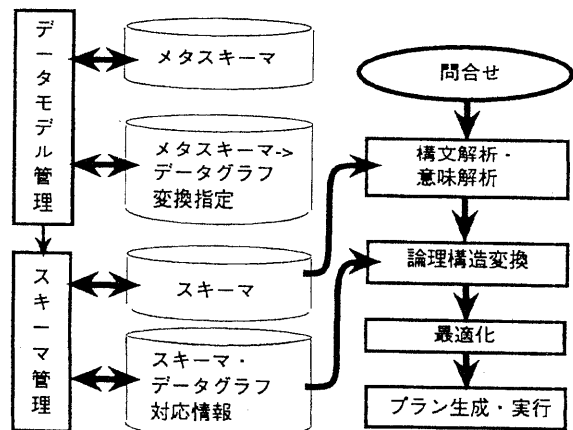


図2 データモデル管理、スキーマ管理と問合せ処理

あるデータモデルを通してある DB をアクセスする際に、必要ならばスキーマを作成するという方法を採用する予定である。また、スキーマ管理方法や他 DBMS 構成要素との整合性、動的な変更方法についても今後の課題である。

6 おわりに

カスタム化・進化可能な DBMS について、特に、データモデリング機能に関する考察を行った。まず、要求を示し、研究動向を概説した。次に、研究課題を示し、一実現方法を示した。今後は、データモデルのカスタム化に留まらず進化に対する研究が必要であろう。

謝辞

日頃議論して頂く井上潮主幹研究員、佐藤哲司主任研究員をはじめとする NTT 情報通信網研究所データベース研究部の皆様に感謝致します。

参考文献

- [1] 増永、田中：”次世代データベースシステムの展望”，情報処理，Vol. 32, No. 5, pp.602-613 (1991).
- [2] Scheck H-J and Scholl M. H. : "Evolution of Data Models", Proc. of Int'l Symposium on Database Systems of the 90's; Lect. Notes in Comp. Sci. 468, pp.135-153 (1990).
- [3] Hong. S. and Maryanski F. : "Representation of Object-Oriented Data Models", Inform. Sciences, No.52, pp.247-284 (1990).
- [4] Hong. S. and Maryanski F. : "Using a meta model to represent object-oriented data models", Proc. of 6th Intl. Conf. on DATA Engineering, pp.11-19 (1990).
- [5] Maryanski F. : "The Data Model Compiler: A Tool for Generating Object-oriented Database Systems", Proc. of Int'l Workshop on Object-Oriented Database Sys., pp.73-84 (1986).
- [6] Dittrich K. R. et al : "Database Technology Research at the University of Zurich: Using and Engineering Object-oriented, Active, and Heterogeneous DBMS", 信学技報 DE91-60 (1992).
- [7] Demurjian S. A. and Hsiao D. K. : "Towards a Better Understanding of Data Models Through the Multilingual Database System", IEEE Trans. Softw. Eng., Vol. 14, No. 7, pp.946-958 (1988).
- [8] Cooper R. : "Configurable Data Modelling Systems", Entity-Relationship Approach: The Core of Conceptual Modelling (H. Kangassaro ed.), Elsevier Science Publishers B. V.(North-Holland), pp.57-73 (1991).
- [9] S G N 社 : MATISSE 発表会資料 (1992).
- [10] Morgenstern M. : "A unifying approach for conceptual schema to support multiple data models", Entity-Relationship Approach to Information Modeling and Analysis (P. P. Chen ed.), Elsevier Science Publishers B. V.(North-Holland), pp.279-297 (1983).
- [11] Hwang H.-Y. and Dayal U. : "Using the Entity-Relationship model for implementing multimodel database systems", ([10] と同じ), pp.235-257 (1983).
- [12] 川越 : "OODB とデータベース応用", 情処研報 90-DBS-78, Vol. 90, No. 63, pp.177-178 (1990).
- [13] Hsiao D. K. and Kamel M. N. : "Heterogeneous Database: Proliferations, Issues, and Solutions", IEEE Trans. on Know. and Data Eng., Vol. 1, No. 1, pp.45-62(1990).
- [14] Gangopadhyay D. and Barsalou T. : "On the Semantic Equivalence of Heterogeneous Representations in Multimodel Multidatabase Systems, SIGMOD RECORD, Vol. 20, No. 4 (1991).
- [15] Scheck H.-J. et al : "The DASDBS Project", IEEE Trans. on Know. and Data Eng., Vol. 2, No. 1, pp.25-43 (1990).
- [16] 宝珍 : "拡張可能 DBMS:COMMON の格納構造と基本演算について", 情処研報データベース・システム 82-6 (1991).
- [17] 宝珍 : "データモデルコンパイラの実現に向けて", 信学第 3 回データベースワークショップ, pp. 23-28(1992).