

データベースシステムの現状と動向

Present Status and Future Trends of Database Systems

現在、データベースは多くの情報システムの重要な構成要素となり、加えて、オフィスオートメーション、エンジニアリング支援システムなど、今後、開発の進む新応用システムに対しても適用が予想されている。このため、データベース技術の重要性は更に増大すると考えられるが、現行のデータベース技術が、現在及び将来の応用システムの要望を十分に満足させている、あるいはさせ得るとは言えない。

本論文では、上記の考え方に立って、まず、データベース管理システムの現状を述べ、次に、現状のデータベース技術のもつ問題点と、その解決のための技術開発の動向、最後に新応用システムのニーズに起因する技術課題について述べる。

吉田 郁三* *Ikuzô Yoshida*
 河崎善司郎** *Zenshirô Kawasaki*
 池田俊明*** *Toshiakira Ikeda*
 齋藤真人*** *Masato Saitô*

1 緒言

1970年前後のDBMS(Database Management System: データベース管理システム)の開発によって、データベースシステムは実用期に入り、現在では多数のデータベースシステムが稼動するに至っている。

このようなDBMSの普及は、DBMSを中心とする現行のデータベース技術が、ユーザーに利益をもたらすことを示している。しかし、現行のデータベース技術が、ユーザーの要望をすべて満たしているわけではなく、不満も表明されている。

さて、上記のユーザー不満の原因は、以下の点にあると考えられる。すなわち、従来から指摘されているとおり、使いやすさの不足、性能面での制約など現在に至るまで未解決の問題点があること、及びOA(Office Automation)、CAE(Computer Aided Engineering: エンジニアリング支援システム)、DSS(Decision Support System: 意思決定支援システム)など、データベース技術の新適用分野の要請に対し、十分に対処できないことである。

本論文では、上記問題点を勘案し、まず、現行DBMSの基本理念、機能及び開発の現状について述べ、次に、現行データベース技術の問題点と、その解決のための技術開発の動向を述べて、本特集に含まれる諸論文の位置づけを明確化し、最後に、新適用分野に対処するための主要技術課題をあげて、今後のデータベース技術の展開方向の参考に供したい。

2 DBMSの出現の背景¹⁾

2.1 DBMSの出現とユーザーニーズ

DBMSの出現の過程は、ユーザーニーズと密接なかかわりをもっている。DBMS出現以前のデータ処理業務では、プログラムは各アプリケーションごとにデータ入出力装置の特性を記述し、また、データファイルを定義、作成しなけりなかつた。このような状況の下では次に述べるような問題点があった。

- (1) 入出力ハードウェアが改まるたびにプログラムを書き替えないければならない。
- (2) 関連のあるデータが別々のファイルに存在することもあり、データの一貫性を保つのが難しい。
- (3) 同じデータを複数プログラムで共用することが難しく、リソースの無駄が生ずる。

データ処理環境の拡大に伴い、アプリケーションプログラ

ムが高度化、複雑化し、量の増大する中でこれらの問題点を解消し、システムの開発及び運用の容易性、信頼性、性能の向上、資源の有効利用などを旨とするために次の点を実現することが望まれた。

- (1) データをアプリケーションプログラムから独立させる。
- (2) データの記述をハードウェア特性から独立させる。
- (3) すべてのデータを集中管理し、複数アプリケーションプログラムがデータを共用できるようにする。

この3点を実現したシステムがDBMSである(図1参照)。一般にDBMSと呼ばれているものに近い機能をもったシステムが現われてきたのは1960年代の後半であり、その歴史は浅いが、この分野のその後の発展は目覚ましいものがある。また、コンピュータユーザーのDBMS利用率も年々拡大する一方であり、ある調査によれば国内のコンピュータユーザー全体のDBMS利用率は約25%であり、1985年には50%、1990年には70~80%に達すると言われている。そのユーザーも当初はほとんど大形機分野に限られていたが、最近では小~中形機分野でもDBMSの利用が促進されてきた。

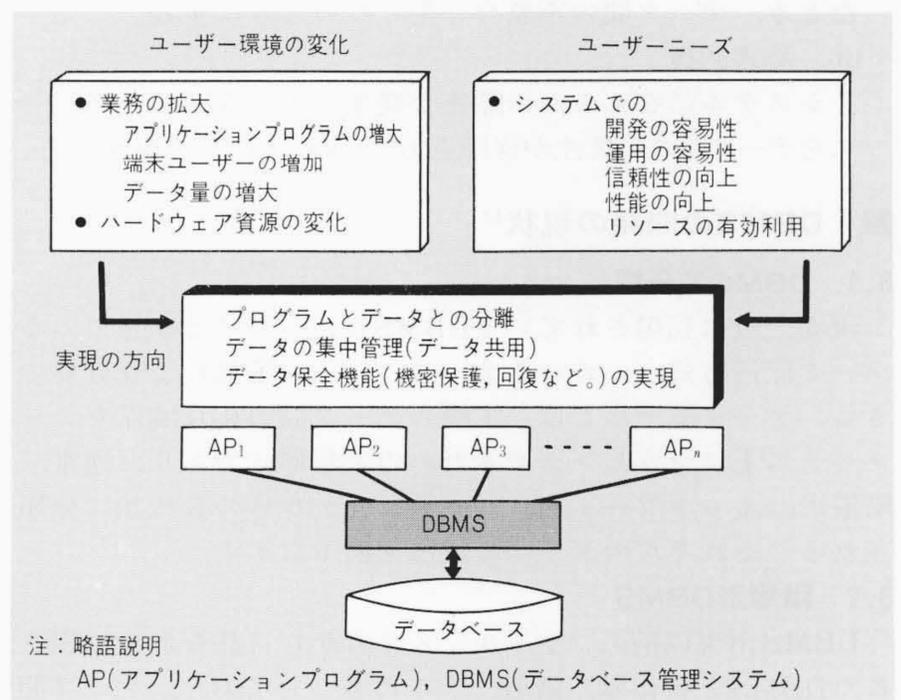
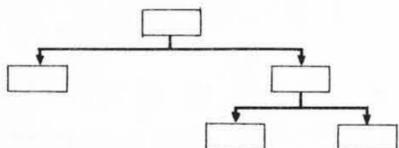
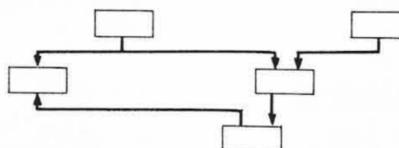
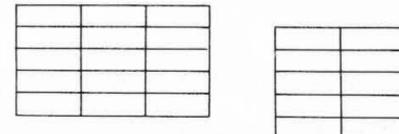


図1 DBMS出現の背景 DBMSは、拡大するデータ処理業務を効率良く対処するために生まれてきた。

* 日立製作所システム開発研究所 ** 日立製作所システム開発研究所 理学博士 *** 日立製作所ソフトウェア工場

表1 データモデルによるDBMSの分類 三つのタイプのDBMSにはそれぞれ特徴があり、データベースの性質によって使い分ける配慮を要する。

	階層形	ネットワーク形	リレーショナル形
ユーザーの見方	<ul style="list-style-type: none"> 木構造 1:n関係 	<ul style="list-style-type: none"> 網目構造 m:n関係 	<ul style="list-style-type: none"> 構造なし 第1正規形フラットファイル 
アクセスパス	<ul style="list-style-type: none"> 原則として一通り 	<ul style="list-style-type: none"> 複数定義できる。 	<ul style="list-style-type: none"> 規定しない。
特徴	<ul style="list-style-type: none"> 各データへのアクセス方法は、一意に決まる。 ユーザーは、論理的データ構造を知らなくてはならない。 性能重視 	<ul style="list-style-type: none"> 各データへのアクセス方法は、複数存在できる。 ユーザーは、論理的データ構造を知らなくてはならない。 多様性重視 	<ul style="list-style-type: none"> アクセス方法を意識する必要はない。 ユーザーは、各データ項目の意味を知っていればよい。 柔軟性重視

2.2 DBMSの基本的機能

DBMSの機能は一般的に次のように分類される。

(1) データ記述機能

これはデータ構造をアプリケーションプログラムとは全く独立に記述できる機能である。DBMSは一般にDDL (Data Definition Language) と呼ばれるデータ定義用の言語を提供しており、この言葉を用いることによってデータベース内のデータを定義することができる。

(2) データ操作機能

データの物理的配置を意識することなく、データベースに対して検索、更新、追加、削除といった操作を提供する機能である。

(3) データ保全機能

複数アプリケーションプログラムにデータの共用を許容し、各種のアクセスを制御することによってデータを保護する機能である。データ保全機能は更に次のように分けられる。

(a) 機密保護機能

不当なアクセスからデータを保護する機能である。

(b) 一貫性制御機能

データベース内のデータ内容の一貫性を保障する機能であり、次の二つの働きをする。

(i) 同時実行制御

複数のユーザーが同時に同じデータベースをアクセスしたとき、データ間の不整合が生じないようにする。

(ii) 障害回復

システムになんらかの障害が発生したとき、データベースをデータ間の一貫性が保障されている時点まで復元する。

3 DBMSの開発の現状¹⁾

3.1 DBMSの分類

現在一般に利用されているDBMSは、ベースとなっているデータ記述方式、いわゆるデータモデルの種類により分類できる。データモデルとは、実際のデータ間の相互関係をデータベース上にマッピングするための表現形式であり、通常、階層形、ネットワーク形、リレーショナル形の3種類に分類される。それぞれのタイプの特徴を表1に示す。

3.2 階層形DBMS

DBMS出現以前のファイルシステムから自然な形で拡張されたDBMSと言える。階層モデルに基づいており、データ間の関係はすべて1:Nの親子関係によって整理されている。レコード間の関連づけが階層構造では極力抑えられているために、データベース内の物理的構造も比較的単純になる。任意のデータレコードに対するアクセスパスは一つに限定されてしまうが、性能が重要でデータ量の多いデータベースに適

している。

3.3 ネットワーク形DBMS

データ間の関連づけを自由に行なえるネットワークモデルを基にしている。DBMSの発展に重要な役割を果たしているCODASYL/DBTG (Conference on Data Systems Languages/Data Base Task Group)^{*1)}が提唱しているのもこの形である。データレコード間の関連づけが柔軟に行なえるために、データベースの物理的構造が複雑になりやすいという欠点があり、性能管理は難しい。しかし、ユーザーがあるデータレコードをアクセスするために様々なアクセスパスを提供することができ、多様なアプリケーションプログラムに比較的容易に対応できる。

3.4 リレーショナル形DBMS

E.F.Codd^{*2)}が1970年に提唱したリレーショナルデータモデルに基づいたもので、操作手順をほとんど意識せずにデータベースへのアクセスができるようにしたDBMSのことである。あらゆるデータを正規化されたフラットテーブルで表わす。データ定義の拡張、変更が自由に行なえるなど、高度の柔軟性を提供しているため、非定型業務の多いデータベースに適している。操作手順が容易であることから、エンドユーザー指向のDBMSであるとも言える。しかし、正規化されたフラットテーブルが必ずしも現実のデータ構造にマッチしているとは言えない。また、大規模データベースに適用した場合、現行技術レベルでは性能面で解決すべき問題が多い。

3.5 製品としてのDBMS

現在までに数多くのDBMSが製品化され実用に供されているが、そのほとんどのものは階層形、あるいはネットワーク形のDBMSである。リレーショナル形DBMSについては、数多くの実験システムが開発されているが、実用上の問題点から製品化されたものは数少ない。DBMSとしてどのタイプのものが優れているかということは、アプリケーションに依存した問題であり、一概には語れない。DBMSの適用には、そのアプリケーションの性質やデータベース規模を十分理解した上で、最も適した製品を選択することが望ましい。

4 データベース技術の動向

以下、4.1節に現行データベース技術の問題点と、これに対処するため進められている技術開発の状況を概説し、4.2節に

*1) CODASYLは、米国の国防省を中心に1959年に発足し、COBOLの仕様を定めるなどした。DBTGは、その下部組織であり、データベースに関する専門委員会である。

*2) IBM San Jose研究所の研究者で、リレーショナルデータモデルに関する研究で先駆的な役割を果たした。

OA, CAE, DSSなどのデータベース技術の新適用分野の特性に起因して顕在化しつつある技術課題について述べる。

4.1 現行データベース技術の問題点と技術開発の状況

(1) 使いやすさ(Usability)の不足への対処

データベースの一つの特徴は、多数の計算機に関する知識レベルの異なるユーザーによって共用される点にある。したがって、全ユーザーの要望を満たすようにデータベースを設計することは、従来の個別ファイルの設計に比べて困難である。

また、現行の主流のデータ操作言語は、手続き形であり、計算機の非専門家にとって、使いやすいとは言い難い。

このような使いやすさの不足に対処するため、現在、データベースの設計技法及びデータ操作言語の研究・開発が進められている。以下に、その概況について述べる。

(a) データベース設計技法、データベース設計は、論理設計、物理設計の二つの領域に大別できる。

論理設計³⁾は、応用システムの対象となるデータの、多数のユーザーからの見方を統合する過程であり、設計に適したデータモデルの設定、このモデルを用いたデータベース設計の方法論、上記のモデル、方法論によって定義したデータベースを既存のDBMSのデータ構造に変換する技法などが開発の対象となっている。

一方、物理設計⁴⁾は、データベースの格納媒体上での編成の過程であり、データベースを物理的記憶構造と格納データへのアクセス方式に対応づける際の最適化問題に帰着する。

したがって、物理的記憶装置の選択、記憶構造の選択、アクセス法の選択などを性能、容量などを評価基準として行なう技法が開発の対象となっている(本特集号別論文「データベースの論理設計技法」参照)。

(b) データベース操作言語に関しては、非手続き化¹⁾、専用化⁵⁾の傾向が著しい。前者は、ユーザーが所要とするデータの満たすべき条件だけを指定すれば、データの入手が行なえること、すなわちプログラミングに関する知識なしに、データの操作を行なえることを目標とし、後者は、例えば、事務、経営などの専門家が、それぞれの分野で習熟している感覚で、データベースの操作が行なえることを目標としている。

この分野では、指定されたデータの入手を効率化するため、システム自身が、最適アクセス手順を生成する処理方式、利用者の特性に親和性のある言語仕様、マンマシンインタフェースの設定などが開発の対象となっている(本特集号別論文「リレーショナルデータベースシステムの動向」参照)。

(2) 性能(Performance)面での制約への対処

データベースシステムでは、性能を決定する要因が多数あり、かつ要因間には複雑な相互関連がある。このため、性能を最適とするように、要因の値を決定することは容易ではない。この問題に対処する技術の一つは、前述したデータベース設計技法である。

しかし、データベース処理での性能の制約には、データベース処理の特性と現行の汎用計算機システムの機能とが不適合であるという根本的な原因がある。すなわち、

(a) 現行の汎用計算機が、基本的には、計算機出現当初の応用システムの要求であった技術計算の効率的実行に適した機能を保有し続けているのに対して、

(b) データベース処理では、大量データの操作(蓄積、検索など)の効率的実行が必要であること。

による。

上記の問題に対処するため、データベース処理に適した専

用計算機、いわゆるデータベースマシンの研究開発が進められている⁶⁾。

データベースマシンは、上記の不適合をハードウェアに用意した機能によって解消し、データベース処理の効率化を目標としている。そして、その実現の方式は、設計者の考え方に応じて異なり、現在までに実験機、商用機を含めて多様なデータベースマシンが発表されている⁶⁾。

データベースマシンの実用化は、今後進展するものと考えられるが、これに伴って、実現方式に関する幾つかの規範が生じ、この規範を満たす方式が将来の主流になっていくものと考えられている⁷⁾。この規範としては、

(a) NBS(National Bureau of Standards: 米国標準局)を中心とするDBMS標準化案

DBMSのアーキテクチャ、データモデル、データ言語などの標準化案に対するデータベースマシンの方式の適合性が問題になる。

(b) VLSI技術による実現容易性

製造コストの低減、信頼性の向上、拡張容易性の向上などには、VLSI技術を用いて作成した同一機能セルの規則的組合せによるデータベースマシンの構築が好都合である。などが考えられる。今後のデータベースマシンの発展の動向を考察する際には、上記の2点に留意する必要があると言えよう。

(3) データの保全性(Data Integrity)不足への対処

データベースシステムは、システムを構成するソフトウェアリソース(データ群及びプログラム群)が共用されるため、既存のソフトウェアリソースを利用して、新規の応用システムの開発が容易になるという利点がある反面、データ群、プログラム群の間に複雑な相互関連が生じている。

したがって、既存のソフトウェアリソースの変更・拡張などを行なう場合、その影響が他の多数のソフトウェアリソースに波及し、このため、変更・拡張などを誤りなく実施することが困難になるという問題がある。

この問題に対処するため、データベースシステムを構成するソフトウェアリソースの定義情報、及びソフトウェアリソース間の関連情報を保持し、ユーザーの要求に応じてその表示を行なうという基本機能をもったソフトウェア、データディクショナリ/ディレクトリシステムが開発されている。

ユーザーは、データベースシステムの変更・拡張などに際して上記の基本機能を利用し、ソフトウェアの関連を把握した上で必要とする変更・拡張などを実施することができるので、誤りを犯す危険性が減少しデータの保全性向上が図れる。

データディクショナリ/ディレクトリシステムは、上述のように、当初はデータベースシステムの保守・運用を目的として開発されたが、その保持する情報の有用性が着目されて、種々の拡張がなされている。すなわち、

(a) データベース設計システムと結合して、データベースの設計・開発・保守の一貫支援を行なう。

(b) データベースの変更・拡張シミュレーションシステムと結合して、変更・拡張などの影響波及箇所を指摘を自動化する。

(c) データの使用実績取得機能を付加し、データのライフサイクルを知り、データ資源の管理に使用する。

などが実施されている(本特集号別論文「データベースシステムの高信頼化技術“HIDDS/HIDRESS”の開発」、「データ資源管理の概念と動向」参照)。

以上、本節で述べた各種の技術の大部分は、既に実用化の段階に達しており、今後は個別技術のいっそうの充実が進む

とともに、そのあるものはDBMSの中に組み込まれて、より高機能なDBMSの実現に寄与するものと考えられる。

4.2 今後の課題^{8),9)}

データベース技術の新しい適用分野として注目を浴びているOA, CAE, DSS, 研究開発支援システムなどでは、単に数値、文字データだけでなく、文章、数式、図形、イメージ、音声など、コード化やレコード概念の設定が困難な情報の取扱いが極めて重要となってくる。これらの点で、従来の汎用データベースシステムでは、データの定義と操作方法が著しく制限されており、今後のアプリケーションに対して十分に満足できるものではない。

これをデータ処理方式の観点からみると、従来のデータベースシステムでは、データ処理目的の達成に必要な知識を利用者があらかじめ保有し、この知識とデータベースシステムのもつ基本操作とが組み合わされて結果が得られるという方式が一般的である。これに対して、今後のシステムではあらかじめデータベースに利用者の仕事に関するモデル、環境、知識、更にデータ処理手続に関する知識を蓄積しておき、利用者側からの要求をシステム自身が基本操作の組みに分解・実行して結果を得る、という方式が主流になるものと思われる。

上記のような高機能データベースシステムに要求される主な機能としては、次のようなものが挙げられるであろう。

(1) 異種類データの統一的定義・操作・管理機能

数値、文字データに加えて、手続(プログラム)、知識(ルール)、文章、数式、図形、イメージ、音声データなどの統一的な処理、管理方式の確立が必要である。特に、プロダクション規則などで表わされた知識を用いて演えき推論を行ない、データベースに対して高レベルのアクセスを可能にすることが重要である。知識の表現、推論機構、知識の無矛盾性制御方式などが重要な技術課題となる。

(2) データベースの進化に即応する柔軟なデータ構造

データベースの変化には、データ値の変化とスキーマ自身の変化の二つの側面がある。データ値の変化に対しては、それに効果的に対処できる機構(最適物理構成、複数利用者による同時更新制御、データの一貫性の保持など。)の高度化、自動化が必要である。スキーマに対しても、その変化(例えば、リレーションや属性の追加、削除、分割、統合など。)が通常データに対するのと同様に扱える柔軟な操作機能の提供が重要である。

(3) 階層的ユーザービュー

データに対する種々の抽象化レベルでの階層的定義と、それらに対するアクセス機能が必要となる。例えば、CAEでの設計プロセスで、設計者は概念設計から詳細設計までの各段階(抽象化レベル)を自由に参照して必要な情報を得たい。このような、いわば概念的拡大・縮小機能の提供が重要である。

(4) 分散データベース管理機能

組織やユーザー権限の分散、システム利用部署ごとの処理方式の多様性、処理の非同期性などの理由で、公共データベースや個人データベースなどの異種データベースを統合利用できるように、分散データベース管理が重要である。これに関しては、各種のデータモデルを内蔵したDBMSの開発、各種のデータモデルを取り扱える統一的操作言語の開発、データの抽出や変換が行なえるデータベース変換システムの開発(本特集号別論文「複数データベースの統合・利用技術」参照)などのアプローチが考えられる。この分散データベース管理機能に関連して、セキュリティ機能の強化が重要である。

(5) エンドユーザー機能

CAEでの設計プロセスやオフィスでの作業は、本質的にイ

ンタラクティブであり、エンドユーザー機能の充実が必須である。これに対しては、データ検索言語の高水準化・非手続化と、高機能入出力端末利用の二つの方向が追求されている。前者に対しては、検索要求を述語論理に基づく形式、自然言語を用いる方法、図的言語を利用する方法などがある。後者に対しては、人間の情報処理機能とマッチした、多様な入出力方式の開発が望まれる。ディスプレイ上での画像情報の移動、拡大、縮小などの編集機能、データベース出力との重ね合せ機能などの高度化が必要である。更に、データベース構築時の一つの課題として、多量データの入力の問題があるが、これには文字や音声認識技術の高度化はもとより、意味処理機能、推論機能を備えた高度な自動インデクシング技術(本特集号別論文「日本語情報検索システムにおけるキーワード自動抽出」参照)の確立が課題である。

5 結 言

現在、データベースは、多くの情報システムの重要な一構成要素となっている。そして、その重要性は、今後、更に増大すると予想される。すなわち、従来からの定型的な事務処理システムに加えて、今後発展が予想されるOAでの帳票類(フォーム)の管理、CAEでの設計図面類の管理、機械翻訳、自然語処理システムなどでの大規模辞書など、データベース技術の必要となる分野は拡大すると考えられるためである。

しかし、本論文の4章で述べたように、現行のデータベース技術は、従来のアプリケーションの要望に対して十分にこたえているとは言い難く、ましてや、上記のような新分野のアプリケーションの要望には対処できないのが現状である。

このような現状を打破していくためには、

(1) 現行データベースの問題点に対処する技術を、従来の成果を踏まえて展開し、その実用性をユーザーと協力して検証していくこと。

(2) 新分野のアプリケーションに必要な基本機能を、各分野の専門家であるユーザーとメーカーが協力して抽出し、その実現に必要な技術課題に対して、着実な研究開発の努力を行なうこと。

が、必要であると考えられる。

参考文献

- 1) C. J. Date : An Introduction to Database Systems 3rd Edition, Addison-Wesley(1981)
- 2) S. B. Yao, et al. : Data-Base Systems, Computer Vol.11, No. 9, 46~60(1978)
- 3) 酒井, 外 : データベース・システムに関する調査—データベースの論理設計—, 日本電子工業振興協会(昭和53年3月)
- 4) 酒井, 外 : データベース・システムに関する調査—データベースの物理設計—, 日本電子工業振興協会(昭和54年3月)
- 5) M. M. Zloof : QBE/OBE A Language for Office and Business Automation, Office Automation Conference 1980 Digest 249~260(1980)
- 6) 吉田, 外 : データベース・システムに関する調査—分散データベースおよびデータベース・マシン—, 日本電子工業振興協会(昭和56年3月)
- 7) S. Y. W. Su, et al. : Database machines and some issues on DBMS standards, Proc. of NCC 191~208(1980)
- 8) 第5世代計算機調査研究委員会 : 第5世代のコンピュータ研究開発課題説明書, 日本情報処理開発協会(昭和56年6月)
- 9) 吉田, 外 : データベース・システムに関する調査—OAとデータベース技術—, 日本電子工業振興協会(昭和57年3月)