

高水準データ・ベース・システムの現況と将来*

國井 利泰 (東京大学・理学部・情報科学科)**

1. 高水準データ・ベース・システムの定義

高水準データ・ベース・システムは、いわゆる“ユーザー向システム”^{1,2)} (user oriented systems) ”と云われるものの一種である。ユーザー向システムというのは、ユーザーがシステムの詳細に立入りプログラム作製等の手順を経たりせず、非手続的に直接システムと相互作用しつつ、問題そのものに対する経験と発見とを使って、充分構造の分っていない問題を直接解決することを可能にするようなシステムである。高水準データ・ベース・システムは、同時に、次のような重要な機能を持つ。特定のコンピュータの構成から独立した超表現 (meta representation) で表わしたデータの処理を可能にするような、機械独立システム (machine independent systems) としての機能である。データ・ベースと呼ばれる高コストの共用データ群が、特定のコンピュータでしか使えないようでは、その価値が激減するからである。もちろん、機械独立システムでは、データ表現のためのデータ定義言語³⁾ (data definition languages, 略してDDL) として、超表現の代わりに利用者の検索言語 (query languages) を直接用いることもできるが⁴⁾、目的の相異なる両者を共用することは無理であることが当然のことながら幾多の実験の結論である。

高水準データ・ベース・システムは、すべてのデータ・ベース・システムがそうあるべきように、絶えず最新の内容が追加され不要なものが削除されているから、それが自由に行えるようデータ独立システム (data independent systems) でなければならぬ⁵⁾。同じデータ共用向のシステムであるコンピュータ・グラフィクスを取り上げてみると、データは通常数多くのポイントで相互に多重に結合合わされ、このデータ独立性は通常保たれていない。現実のデータ・ベース・システムでは、まだディスク等のファイルでの蓄積構造やそれへのアクセス方式に依存したデータの追加・削除・更新方式を採用せねばならないことが多く、その意味でデータの操作はその方式に縛られて自由では無く、データ独立システムにはなっていないものがほとんどである⁶⁾。

データ・ベース・システムは、当然のことながらそれが本来の目的に使われてみるまでは、すなわち実際に様々の応用に共用されてみるまでは、どのような応用があるか明確には分らないのが現実である。したがって数多くの原理上相異なる応用にシステムの変更無しに利用し得るものでなければならぬ。データ・ベース・システムの現状は、データがある利用目的に沿って表 (すなわちn項関係、ただしnは正の整数である) として、あるいはネットワーク、トリイ等のようなデータ構造化して入力されることが多く、このような応用上の自由さは残念ながら達成されて居らず、したがって応用独立システム (application independent systems))

* Present Status and Future of High Level Data Base Systems

** Toshiyasu L. Kunii (Department of Information Science, Faculty of Science, The University of Tokyo)

にはなっていない。高水準データ・ベース・システムは、したがって、当然の目標として、応用独立性を追求することが望ましい。応用上のモジュラリティを保つことによりこれを達成している例が一つある（例えば、アメリカ合衆国スタンフォード大学情報処理センターのSPIRES⁸⁾）。図式-副図式 (Schema-Subschema) 方式によるCODASYLのDBTG提案⁹⁾は、応用独立性をある程度考慮に入れようとする萌芽と見なせるが、まだ程遠い。几項関係としての表の論理的操作を中心武器として登場したリレーショナル・モデル¹⁰⁾は、高水準データ・ベース・システムであることを売りものにしているが、それは機械独立かつデータ独立ではあっても、データ構造化の水準のシステムの一つであるために応用独立になっていない。例えばある応用上の観点から作製された表の集まりとして構造化されたデータを別の応用に用いるためには、新しい応用の観点 (views) を表す副図式と従って再構造化せねばならず、このようにデータと応用上の観点が未分離であるというデータ構造水準¹¹⁾ (data structure level) のデータ・ベース・システムに特有の制約から脱け切っていない。したがってここで定義している高水準データ・ベース・システムの域にはリレーショナル・データ・ベース・システムはまだ到達していない。

2. "ユーザー向システム"

上の冒頭に、高水準データ・ベース・システムは、いわゆる"ユーザー向システム"^{1,2)}の一種であると述べたが、"ユーザー向システム"についてもう少し考察を進めよう。ユーザーにはプログラム作製等による手順を経ずに直接データと非手続的に相互作用できることが望ましいこと、一般にユーザーが日常の情報活動に用いる情報の形態としては通常のデータ・ベース・システムの扱う文字、数字以外に画像、音響等のパターン・データがあるのでこれ等が自由に混合して使えなければならぬということ、この二つが"ユーザー向システム"であることの要件であると考えられる。したがって、

"ユーザー向システム"

- a. 人間・データ直接相互作用システム (非手続的)
これを対話型データ機械 (interactive data machines),
あるいは単にデータ機械 (略してDM)とよぶことにする。
- b. 拡張データ・タイプ・システム (extended data type systems). 文字、数字、絵画、パターン、音響等の信号データ等の任意のタイプデータ¹⁵⁻¹⁷⁾を自由に混合して扱えるシステム。

対話型データ機械では、複雑な手続的プログラム群により事前に高度に構造化する通常のデータ処理方式を、複雑な構造化プログラムとデータ構造を次のもう一つで置き換えることによりユーザー向にする：

- 1) . 単純な対話用コマンド (interactive commands) の集まりと複雑なデータの組合せの多様性。
- 2) . 上記の組合せ順序の多様性。

3. 文献

1. E. D. Carlson, J. L. Bennett, G. M. Gidding and P. E. Mantey, "The Design and Evaluation of Interactive Geo-data Analysis and Display System", Information Processing 74, pp. 1057 - 1061 (1974).
2. T. L. Kunii, T. Amano, H. Arisawa and S. Okada, "An Interactive Fashion Design System 'INFADS'", Computer and Graphics, 1, pp. 297 - 302 (1975).
3. 例文は, C. J. Date, "An Introduction to Database Systems" (Addison-Wesley Pub., Mass., 1975).
4. R. F. Boyce, D. D. Chamberlin, "Using a Structured English Query Language as Data Definition Facility", IBM Research Report, RJ 1318, San Jose, California (1973).
5. 例文は, J. Martin, "Computer Data-Base Organization" (Prentice-Hall, New Jersey, 1975).
6. R. Williams, "A Survey of Data Structures for Computer Graphics systems", Com. Surveys, 3, pp. 1 - 21 (1971).
7. 3, 5 および R. Rustin, ed. "Data Models: Data Structure Set versus Relational", Proc. 1974 ACM-SIGMOD Debate. (1974).
8. "Design of SPIRES", Vol I, II (Stanford Center for Information Processing, Stanford University, California, 1973).
9. "CODASYL Data Description Language", Journal of Development (National Bureau of Standards, Washington, 1973).
10. E. F. Codd "A Relational Model for Large Shared Data Banks," CACM, 13; pp. 377 - 387 (1970).
11. ibid, "Recent Investigations in Relational Data Base Systems", Information Processing 74, pp. 1017 - 1021 (1974).
12. M. Stonebraker, "A Functional View of Data Independence", Proc. ACM-SIGFIDET Workshop on Data Description, Access, and Control (1974)
13. M. M. Astrahan, et al, "System R", to appear in ACM Trans. on Database Systems.
14. T. L. Kunii and H. Kunii, "A Virtual Information System for Environmental Applications", Preprints of IFIP Working Conference on Modelling of Environmental Systems, pp. 129 - 131 (Tokyo, 1976).
15. M. Kikuchi and T. Yamasaki, "A Nonprocedural Language for Business Applications", Proc. Int. Computer Symposium, pp. 66 - 74 (1975).
16. M. Hammer, W. G. Howe, V. J. Kruskal and I. Wladawsky, "A Very High Level Programming Language for Data Processing Applications", IBM Res. Rep. RC 5583 (1975)
17. M. Hammer, "The Design of Usable Programming Languages", ACM Nat. Conf. Proc., pp. 225 - 229 (1975).
18. T. L. Kunii, S. Weyl, J. M. Tenenbaum, "A Relational Database Schema for Describing Pictures with Color and Texture", Proc. 2nd Joint Conf. on Pattern Recognition, p. 330 (1974).