

ORとデータ・ベース

3月中旬ごろ、本誌の編集委員の方から、7月号の特集テーマをデータ・ベースとするからORとデータ・ベースについて解説的な記事を書くようにと仰せつかった。その時、私の頭に浮んだのは「モデル人間とデータ人間 [1]」についての森口先生の有名なお話で、この件は森口先生にお願いするのが適切ではないかと編集委員氏に申し上げたものである。しかし、どういふわけか結局は私が書かなければならないことになってしまった。

正直に言って、私はORの教ある手法のいずれについてもエキスパートではないし、また実際問題に直面して生きたデータと取り組んだ経験ももっていない。ただ、情報処理のためのデータ・ベース技術についてほんのわずかに勉強しているにすぎないのであって、こうした輩はORとデータ・ベースについて語る資格が本当はないのだと思う。しかしそのへんのところはご容赦いただき、データ・ベースにまつわるいくつかの話題をご紹介しながら門外漢なりにそれとORとのかかわりについて考えるところをまとめてみようと思う。

1. モデルとデータ

森口先生のお話の一部をもういちど引用させていただくと、「人間のタイプに2種類あって、ひとつはデータ人間、もうひとつがモデル人間である。これまで主としてデータ人間の活躍する場が多かったが、これからはモデル人間が必要とされるようになってきた。しかし世の中すべてがモデル人間ばかりでは困るので、いつでも両者が適当な割合でいてくれることが望ましい」といった趣旨の話がされた。

このとき先生は、「適当な割合で」ということの

一例として50対50という比率をあげられたが、これはそのときの聴衆（それは大部分が企業の人々であった）のことを配慮されての比率であって、その真意はたとえば80対20ぐらゐの気持でおられたのではないかと推察される。そうだとすれば私も同感であり、ORでもっとも大切にしなければならないことは、モデル化しようとする態度であろうと思うからである。そこで、ORの視点にたつて物を見ようとする者、この80対20という比率を自分自身にあてはめて考え、自らをモデル人間として育てる努力をすべきではないだろうか。

一方、諸問題のOR的接近にあたって、いろいろな場面でデータの助けを借りなければならないことも事実である。しっかりしたデータに裏づけされていない論理が説得力をもたないのと同様に、データの充分な分析や種々の条件を反映したデータによるテストをふまえていないORモデルもすぐに馬脚をあらわしてしまう。高橋幸雄先生の言葉を借りれば、こういうのを悪い意味での似非モデルというのであろう [2]。

2. ORとデータ・ベース

他人から似非モデル屋とよばれないためには、モデル人間といえどもデータをないがしろにはならない。しかしこの意味は膨大なデータを覚えるということではなく、データをうまく集めて料理することであって、要するにデータの上手な使い方を知ることにはほかならない。

それではデータそのものは誰がもっていればよいか。この問に対する一つの答えがデータ・ベースである。データ・ベースはいかなればデータの貯蔵庫であり、データ人間が多くのデータを記憶

して状況に応じたデータを取り出してくるのと同じように、モデル人間の要求に合わせてデータを提供するシステムである。ORにとってデータ・ベースとは何かといわれれば、とりあえずこのような定義で充分であろう。したがってこの定義の中には、たとえばお役所のスチールケースにぎっしり詰った紙のファイルも含まれるし、警察に保管されているであろう要注意人物の写真や指紋のリストも含まれるであろう。もちろん、計算機用の磁気テープやその他の記憶装置に納められた記録も該当する。

こう考えるとどんなものもすべてデータ・ベースだということになってしまう。それならことさらデータ・ベースなどといわなくても、データの集まりといってもよいし、従来から使ってきたファイルという言葉を使ってもよいではないか。ファイルとデータ・ベースではどこがちがうのか。あるいはデータ・バンクという言葉もあるが、それとはどちらがうのか、という疑問が当然おこってくるであろう。

私にいわせれば、少なくともORのモデル人間の立場から見れば、それはみな同じものである。モデル人間にとってデータが意味があるのは、それが自分の知らなかったことを教えてくれるもの（情報）だからであり、データそれ自身は一種の媒体にすぎない。これに対して、データを保管する立場からすればデータそのものが対象であり、たとえば数字の列が2倍の長さになれば2倍の苦勞をしてそれを保管することになる。そこで、ファイルとかデータ・ベースという用語の区別は、データを使う人（モデル人間）にとってはあまり重要なことではなく、データを保管する人（データ人間）がその保管方法を改善する過程で技術の違いを識別するために生まれてきたものである。そしてこの立場の典型的な例を、計算機によるデータ処理を専門とする人々（筆者もその一人である）に見ることができる。

しかしORと計算機とは切っても切れない縁に

あるようで、企業においても両方に関係している人は多いし、OR学会員の中にも情報処理学会やACMなど計算機関連の学会にも籍を置く人が相当数いると思われるので、この立場を代弁する意味からもデータ・ベースについて以下でもう少しくわしい説明を加えたいと思う。

3. 計算機とデータ・ベース

言葉を定義することは目的ではないが、まずはこの言葉のおこりをたずねてみることにしよう。

計算機の初期の時代には、プログラムとデータは個々に1対1の組合せでつくられ使われるのが一般的な形態であった。ちょうどわれわれが計算機の使い方を習う場合に、練習問題が与えられるとプログラムをつくりデータをパンチしてカードデッキをつくり、それを計算機にかけてアウトプットを受け取るというのと同じようなものである。これは計算機が個人個人の計算目的のために使われるかぎり自然な使い方であり（当時、計算機を使用できたのは限られた範囲の人たちだけであった）、今日でもこの形がもっとも基本的な使い方なのだということは再認識されてよいことだと思う。

しかし、計算機のすぐれた性能が世の中に宣伝されるようになると、企業などの組織体はその利用を考えるようになっていった。この段階から利用形態にも変化が見られるようになり、限られた人だけが個々の問題を慎重に吟味しながら使っていたそれまでの形態に加えて、組織的な目的のために組織として使う方法が研究されるようになった。これは計算機のその後の歩みにとってきわめて重要な意味をもつことになったといつてよい。

第1に、莫大な資金援助がなされ、新しいことは良いことだといった風潮の中で新技術の援助がなされたことがあげられる（この時代が日本のみならず世界的に見て経済の復興・高度成長期にあたることも見逃せない）。第2に、組織的な利用に適した利用技術や装置の開発が促進され、実際に

多くの成果が実ったことである。今日ではあたりまえと思われるコンパイラやオペレーティングシステム（OS）もこの時代に進歩したし、大量データの蓄積と高速呼出しを可能にするドラムやディスク装置が発達したのもこの時代である。第3に計算機の利用者層の変化をあげることができる。おおざっぱにいうと、計算機を何かの目的のために利用する集団と計算機を扱う集団との分化であり、それぞれの中にもさらにいくつかの層が生まれるようになった。これは利用者人口の増大や組織的利用形態に合わせて、人間も組織化されるようになったということである。

こうした背景のもとでデータ・ベース技術を大きく発展させた原動力は、ディスクに代表される大容量の乱呼出し記憶装置の進歩であったが、これと並んで計算機利用の考え方にも大きな変化があったことを見逃すわけにはいかない。それまでの一般的な利用形態としては、プログラムがあってデータがあり、データがあってプログラムがあるというように常に1対1の相補的な関係がふつうであった。しかし組織的な利用環境のもとでは、たとえば台帳などのようにデータが単独に存在し、それをいろいろな仕事（プログラム）のために使うことが必要になってくる。そこで、データとプログラムは互いに独立した存在と考えられるようになり、データを独自に管理する技術としてデータ・ベース技術が発展してきたわけである。

4. データ・ベースとデータ・バンク

データ・ベースと並んでデータ・バンクといういい方もかなり一般化している。両者が同じ概念を指すものなのか、異なるものなのかは判然としないが、おそらく発生的にはほとんど同じ概念を指すものであったと推察される。たまたまある人がデータ・ベースといい、別の人がデータ・バンクとよんだといった程度の違いだったのではないだろうか。しかしその後、それぞれの用語を使う人々の間で特定の意味合いを含めて解釈されるよ

うになり、今日それらがまったく同義であると断言するにはやや抵抗を覚える。

両者の違いについての明確な定義を知っているわけではないので、あいまいなことをいうのは避けるべきかもしれないが、筆者の個人的なイメージとしてはつぎのように理解している。データ・ベースは、多くのデータを効率よく格納し、検索することに関する諸技術全般、およびその技術にもとづいてつくられたファイル群とその管理機構とを指すものである。その手段は必ずしも限定されないが、ほとんどの場合、計算機ないしその他の電子的・機械的な記憶装置を使うことが前提とされている。これに対して、データ・バンクはデータの内容やその利用形態にさらに深くかかわるものであり、一定の利用環境のもとでデータをどう組織し、どう流通させ、最終利用者にどのような便益を与えるかという問題全般を含むものである。もちろん、この場合でも主たる立場がデータの管理・提供にある点はデータ・ベースの場合と同じであるが、一定の利用環境を考えることによってその流通上の問題までを含める点が大きな違いであるといつてよいであろう。

とはいえ、データを最終的に使う人から見ればこれらの差異をことさら云々することもないであろう。いずれもデータ人間の立場でデータをどう管理するかという問題なのだから。本稿の目的はデータ・ベースについて解説することであるが、ORの専門家（あるいはそれをめざす人）にとってのデータ・ベースという意味であるから、ここではデータ・ベースもデータ・バンクもほとんど同じものと考えて話を進めることにしよう。

5. データ・ベースの構成要素

さて、データ・ベースはどんな要素から構成されるのであろうか。ここでは話を簡単にするために計算機によるデータ・ベースに限定して考えることにする。このときデータ・ベースは以下の三つの要素から構成される。

- データそれ自身
- データ構造の定義
- データ管理の専用ソフトウェア

5.1 データ・ベースのデータ

データ・ベースのデータも通常のファイルであることにはかわりはない。ただし、データの呼出しに関して種々の方法が可能であるように、今日のデータ・ベースは磁気ディスクなどの乱呼出し可能な記憶装置に格納されているのがほとんどである。さらに最近では磁気ディスクを数段上回る大容量の記憶装置も実用化されるようになったので、それとディスクの組合せによって記憶容量と呼出し時間の適切なバランスを考えた大規模データ・ベースが出現することになるであろう。

5.2 データ構造の定義

データは、人間の諸活動にともなって発生しあるいは伝達される情報が一定の約束にしたがって記号化されたものであり、個々のデータは元の環境のもとで解釈されたとき正しい情報が再現されることになる。環境が変わればそれは誤った解釈を生む危険性があるので、データを常に正しい環境の中で使うように配慮しなければならない。

この原則を守るもっとも基本的な方法は、データとプログラムを常に不可分のものとする従来方法である。この場合データとプログラムが同一の思想のもとにつくられ、その枠内で使われるのであるから上記の原則が自然に守られる。しかしこれに抗してデータとプログラムを分離しようとするデータ・ベースにおいては、この原則を守るための別の有効な方法を実現しなければならない。そして、これはデータ自身によってその利用環境を明示するしかない。そこで、データ・ベース中の個々のデータが相互にどのような関係の中で利用されるかという相互関係をデータ構造とよび、データ・ベースのデータは必ずそのデータ構造にしたがって解釈されるものとするのである。データ構造が具体的にどんなものであるかは章を改めて述べることにし、ここでは、データ・ベー

スをつくるためにはそのデータ構造を定義しなければならないということ、そしてデータ・ベースのシステムはこの定義にしたがってスキーマとよぶ定義表をシステム内につくり出すということだけを述べるにとどめておく。なお、データ自身とスキーマとをディスクなどに納めたものを狭い意味でデータ・ベースとよぶことがある。

5.3 データ管理の専用ソフトウェア

データ・ベース中のデータを使うには、データ構造の定義（スキーマ）にしたがった解釈をしなければならないことはすでに述べたとおりである。そこで、データ・ベースを使うプログラムにはスキーマを解釈する部分が必要になってくる。この仕事を個々のプログラムに代わって実行するのがデータ・ベースの専用ソフトウェアであり、通常データ・ベース管理システム（DBMS: Data Base Management System）とよばれているものである。

データ構造の定義とデータ・ベース管理システムの存在によって、データはプログラムから完全に切り離された存在となり、プログラムはデータ・ベース管理システムを介してデータをよび出すことになる。したがって、データ・ベース管理システムは、データとプログラムの仲介者（インタフェイス）としてデータ・ベース利用上の諸機能を備えることになる。これら諸機能は一種の言語として利用できるようにされており、とくにプログラム用にはプログラミング言語の形で、また端末装置などから人間が直接使うためには一種の操作指令（コマンド）の形で使えるようにされている。また、この機能の一部としてデータ構造を記述するための言語（DDL: Data Description Language）も備わっている。

データ・ベース管理システムは、データ・ベースという一種のファイルに対する入出力処理を行なうという意味では、計算機のオペレーティングシステム（OS）の機能の一部と考えることができるし、データ管理のための専用ソフトウェアパ

パッケージであると考えられることもできる。いずれであるにせよ、計算機メーカーやソフトウェア会社から数多くのデータ・ベース管理システムが発表され、また使われている昨今である。

ここで、再びデータの最終的利用者の立場から見一つだけ注意しておく

たいことがある。それは、データ・ベース管理システムはデータ・ベースをつくり使うための単なる道具にすぎないという当たり前のことである。それはちょうど今日の計算機がオペレーティングシステムとよぶ基本ソフトウェアを装備して、プログラミングの手段として COBOL や FORTRAN などの言語が使えるというのとまったく同じであって、それだけではデータを使う人（プログラマではない）にとって何の役にも立たない。この道具を使ってプログラムをつくり、そのプログラムを動かして問題の解決にあたるというのとまったく同じに、データ・ベース管理システムを使ってデータ・ベースをつくり、さらにプログラムをつくってはじめて最終的な利用者がデータを使えることになるのである。「データ・ベースを導入したものの、一向に事態は改善されない」などというときの「導入した」という意味はよくよく吟味されるべきであろう。

6. データ構造の書きかた

データ構造は、データ・ベースに格納するデータの相互関係を規定するものであることはすでに説明した。ここではその表現の一例を具体的に示すことにしよう。ただし表現方法はいく通りもあるので、その一つだけを取り上げることにする。

まずデータは現実世界の実体（事物や事象）に対応して発生するものとする。たとえば給与データは社員という実体に付随して発生するし、売上データは売上傳票ないしは販売行為（事象）に付随すると考えられる。実際のデータはその実体の各個体ごとに発生することになるが、ここでそ

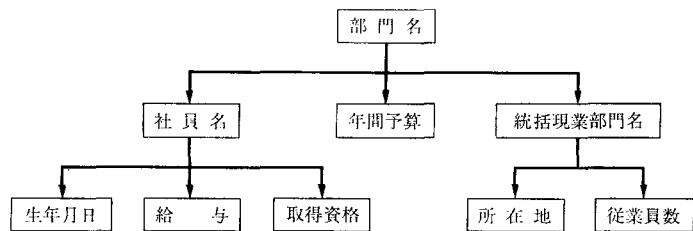


図 1 ある会社についてのデータ構造の一例

の実体を一括して呼称する概念的名称（社員給与とか売上傳票など）をそのデータ全体に対する名称として転用することにする。すでに給与データとか売上データといったのは暗にこの名称を使っているのである。

このようなデータの名称を使えば、あるデータが別のあるデータと常に一定の関係で結ばれているということを簡単な図式によってあらわすことが可能となる。一例として図 1 を見ていただきたい。

この図は、仮想的に簡略化した会社の機構と社員から成るデータ・ベースに対してデータ構造の一例を描いてみたものである。この会社にはいくつもの部門があるであろうが、それらを部門名というデータ名称で代表させている。同じように社員も多数いるが社員名という一つのデータ名称であらわしている。ここで、社員名が部門名の下部構造として描かれている（両者を結ぶ矢がそれをあらわしている）ことはつぎのことを意味している。部門名で代表されている個々のデータ（企画部、営業部、製造第 1 部など）に対して、社員名で代表されているデータのいくつかが従属している、しかも一つの社員名のデータが二つ以上の異なる部門のデータに従属することはないということである。つまり、この部分は図 2-a のようになっているわけで、図 2-b のようなことはない。データ構造の他の部分も同じような約束で書かれている。ただし、社員名に従属している生年月日、給与のように上位のデータ一つに対して必ず一つのデータしかないものもある。

図 1 で示されるデータ構造を一般に階層データ

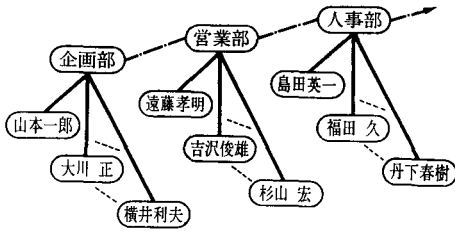


図 2-a 図 1 の構造にしたがうデータの一例

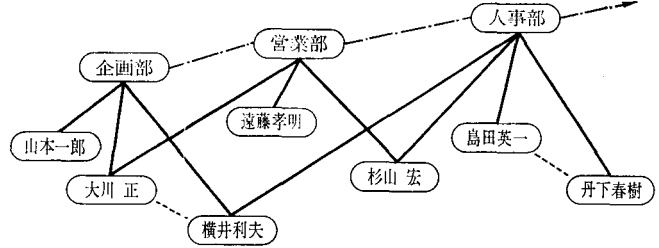


図 2-b 図 1 の構造からは導けないデータの一例

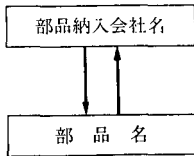


図 3-a 複雑なデータ構造の例(その 1)

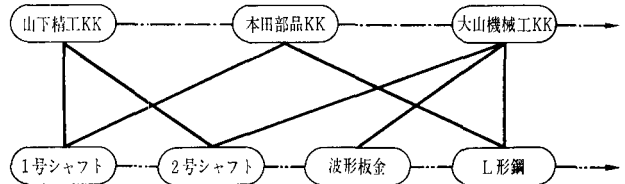


図 3-b 図 3-a の構造にしたがうデータの一例

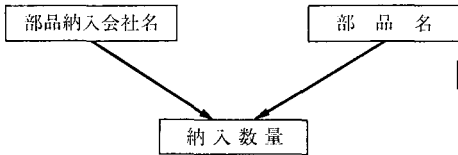


図 4-a 複雑なデータ構造の例(その 2)

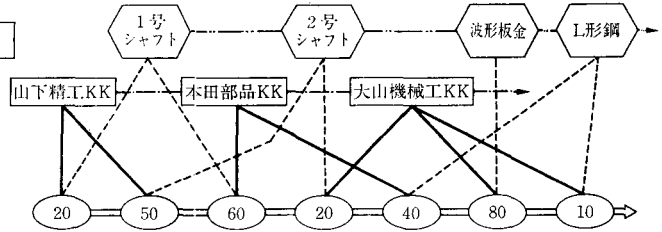


図 4-b 図 4-a の構造にしたがうデータの一例

構造とよんでいる。図 3-a、図 4-a にもっと複雑なデータ構造の例を示す。図 4-a は図 3-a とほとんど同じ内容を少し単純な構造として表現したものである。また 図 3-b、図 4-b は、それぞれ 図 3-a、図 4-a がデータのどんな関連を示しているかを個々のデータによって図解したものである。とくに 図 4-b では、データや相互関係を示す線を種類ごとに区別して示してある。

図 1、図 3-a、図 4-a のように表現されたデータ構造は、この図式のままでは一般に計算機に入力することができない。このためにデータ・ベース管理システムが用意するデータ構造記述用の言語(DDL)を用いて、図式で示された関係を記述する。DDL によるデータ構造の記述では、単にデータの間関係を記述するだけでなく、データ自身の種々の属性や使用上の条件などもあわせて定義する。

データ構造のつくり方にはいくつかの方法があ

り、ある方法によれば簡単明快地表現できることが別の方法では表現しにくい、といったことも起こってくる。データ構造の図式はグラフの一種であるから、一般グラフをその複雑度によって何段階かに分類しておくことによって、あるデータ構造記述方法はどの段階のグラフまでを表現しうるかといった整理が可能であろう。

別の見方をすれば、データ構造をつくることは現実世界の森羅万象をモデル化することにはかならない。したがって OR が常に直面しなければならない現実と理論との相克と妥協がここにも見られるわけで、データ・ベースの研究分野としてももっとも議論の活発なところである。しかし、ここではデータ・ベースの詳細な議論をすることが目的ではないのでこれ以上の細部には立ち入らないことにする。興味をもたれる方は文献などを参照していただきたい [3]~[7]。

7. データ・ベース利用の推奨と警告

本稿を終えるに際して、気がつくことを1, 2述べておくことにしよう。

その第1は、データ・ベースの存在を知って積極的に利用されたいということである。今日では中堅以上のほとんどの企業が計算機を使って業務処理しているので、社内に発生するかなりのデータが計算機用のファイルになっている。もちろんこれが理想的なデータ・ベースの形をしているかどうかはわからないが、とにかくデータの宝庫が眠っているのである。モデルに適したデータはなかなか手に入らないと嘆く前に、これらのデータに当たってみるべきではないだろうか。計算機に載っているデータは何らかの意味で人の手による加工が加わっているので、生のデータより鮮度が落ちるかもしれないが、外部の人間にとっては貴重な情報源と考えるべきではないだろうか。

第2の点は、データ・ベースの限界を心得ることである。データ・ベースは活用すべきであるが、その限界を心得て慎重な取扱いをしなければ、それを使わない以上の危険をとまなうということである。データとモデル、あるいはデータとプログラムは本来適切に組み合わせられるべきであるが、個々の最適性を若干犠牲にしても、より広範なデータの利用を可能にしようとするのがデータ・ベースである。ORにおいてはデータの吟味ということがかなり重要な問題なのだと思う。それと同じで、数値としてのデータはデータ・ベースが提供してくれるであろうが、そのデータを使ってよいかどうかの判断はそれを使う人にまかされているのだということを忘れてはならない。これはデータ・ベースに限った問題ではなく、データを使う者の基本的な心得といったほうがよいであろう。その意味ではデータ・ベースを軽々しく利用することを考えるよりも、自分の手足を使ってデータを集めあるいはつくり出す修練をすることのほうが長い目で見て役に立つことかもしれない。

参 考 文 献

- [1] 森口繁一 「情報処理とOR」情報処理研究No.1, 電力中央研究所情報処理研究センター, 1972. 9.
- [2] 高橋幸雄 「似非モデルの功罪」オペレーションズ・リサーチ Vol.23, No.2, 1978.
- [3] 西村恕彦・中川友康監修「データベースシステム」bit臨時増刊, 1973. 12.
- [4] Martin, J., *Computer Data-Base Organization*, Prentice-Hall, 1975.
- [5] Date, C. J., *An Introduction to Database Systems*, Addison-Wesley, 1977.
- [6] Tsichritzis, D., *Data Base Management Systems*, Academic Press, 1977.
- [7] Jardine, D.A., *The ANSI/SPARC Data Base Management System Model*, North-Holland, 1977.

すずき・みちお 1944年生
電力中央研究所

本誌に事例報告の原稿を

ORの特徴は実践にあるといわれています。実際的な応用をぬきにした理論ということはORでは考えられません。ところがわが国のOR界の現状では理論的な研究発表に比べて実践的な事例の報告がやや少ない感があります。

本誌でも以前から会員の皆さんからの事例報告をお願いしていましたが、まだ十分な成果をあげていないといえませんが、その理由のひとつとしては企業の秘密ということもあると思いますが、ORの実践例というものがないと目新しさがなければ価値が少ないと誤解されていることも一因となっている気がします。

もっと気軽に、「こうやったらこれだけ利益があった」とか、「この問題はこう処理したが、もっとよい方法はないか」というような実例や問題提起をどしどししていただきたいと思います。会員同士の智恵の交換というつもりでこの欄の活用をお願いいたします。

(編集委員会)