

知識処理を指向した 数値情報データベースの利用支援

— 仲介プログラムの自動コーディング —

磯本征雄

仁木克巳

石田卓也

(名古屋市立大学計算センター)

(横浜国立大学工学部)

〒467 名古屋市瑞穂区瑞穂町山の畑1

〒240 横浜市保土ヶ谷区常盤台156番地

概要： 本稿では、一般のコンピュータ利用者が自分の応用プログラムの情報源として数値情報データベースを手軽に利用することを助ける数値情報データベース利用支援システムについて議論する。学術研究用の数値情報データベースでは、予め用意されたソフトウェアの利用だけでなく、そのデータを自分で作成した応用プログラムから直接利用することを望む利用者が少なくない。ここでの議論の目的は、利用者の持っている応用プログラムを解析して、データベースと応用プログラムをつなぐ仲介プログラムを自動生成し、これを実行してデータベースから応用プログラムに数値データを受け渡す手順を支援システムとして実現することである。この論文における筆者らの考え方は、応用プログラムの中には既にデータ入力に関する知識が記述されており、これを解析すればデータベースへのアクセスの方法を引き出せるということである。そして、これらのシステム化が本稿の課題である。

Knowledge based Support System for Users of Numerical Databases

— Automatic Coding of Data Transforming Program from Database to Application Program —

Yukuo ISOMOTO,

Katumi MIKI, and Takuya ISHIDA

Computation Center,
Nagoya City University,Technical Department
of Yokohama National University,Yamanohata 1, Mizuho-cho, Mizuhoku,
Nagoya, 467

Tokiwadai 156, Hodogaya-ku, Yokohama, 240

ABSTRACT: This paper discusses a database-user support system which gives numerical data of a database to a computer user in input format of his application program. The system analyzes his application program to know its input-data format, and generates automatically its data transferring program from a database to a data file. In actual usage of a database, some of database users do not only use it for information retrieval, but also want to use it as source data of their application programs. By the aid of this support system, they can use data in the database as his own data file. The aim of this paper is the systematization of the supporting method in which data of a database are transferred into a data file of an application program.

1. はじめに

学術研究に使われるデータベースには、他の分野に見られない特質がある。その一つは、多くのものが開発途上にあり、また規模が小さく経済的な効用が少ないために、特別の管理要員を置くのが困難なことである。もう一つは、特に数値情報データベースにおいて顕著なことであるが、コマンドによる情報検索だけでなく、利用者が各自に開発した応用プログラムのデータ入力源としても利用したいことである。これまでに、バンキング・システムや経済情報システムなどの数値情報データベースが情報処理専門家によって開発されたソフトウェアによって管理・利用されてきたことに較べると、多くの点でかなりの違いがある。これらの学術研究に特有の課題を解決するには、データベース利用の支援システムを開発するのが最も現実的な方法であると考えている。本論文では、このような学術研究用の数値情報データベース利用上の障害になっている事柄を改善するための利用支援システムについて議論する。

ここで述べるデータベース利用支援システムの対象者は、学術研究用に自分の応用プログラムのデータ入力源として数値情報データベースを利用したいと望む人達である。ところで、学術研究用の数値情報データベースの利用法には、次の2種類がある。その一つは、たん白質結晶データベース^{1,2)}のように「望みのたん白質の3次元結晶構造図をコマンドによって図示する」といったふうに、システム側で準備した既成の機能のみを利用する方法である。もうひとつは、データベースの末端利用者が、データベースの数値データを自分で開発した応用プログラムのデータ・ファイルとして利用する方法である。学術研究では、内容の独創性が重視され、したがって後者のように独自に開発した応用プログラムから直接データベースのデータを利用する要求は多い。本稿の議論の目的も、そのような利用者へのデータベース利用支援の方法を示すことにある。

ところで、データベースを応用プログラムの数値情報源として利用する場合、利用者はデータ操作言語とデータ記述言語の文法、データ構造の記述法、

データベース利用に必要なコンピュータ操作法などの様々な知識を持つ必要がある。しかし、一般のコンピュータ利用者にこれらに関する総ての知識を要求するのは余りにも負担が大きく、しかもこの状況は今後も余り変わらないであろう。このような事柄も踏まえたうえで筆者らは、学術研究者へのデータベースによる数値情報の提供を円滑にする手段として本利用支援システムの開発を開始した。

一方、最近の知識情報処理技術との関係から本支援システムを捉えたと、上記の課題は、比較的容易に解決の期待のもてる興味ある内容であることに気付く。筆者のひとり磯本は、既にコンピュータ末端利用者向けのデータベース構築支援システム³⁾を開発して一部を実用化してきた。本稿で議論する利用支援システムは、この構築支援システムと対にしたものである。これらの議論の中で筆者らは、応用プログラムや概念スキーマをデータ処理やデータ構造に関する知識の記述であると考えて、これらの知識処理を基にすれば数値情報データベース利用支援のシステム化も可能であると理解している。

2. 数値情報データベースの利用支援

本論文で議論する支援システムの役割は、出来る限り簡単な手続きで、大型コンピュータやパソコンで稼働している応用プログラムにデータベースの数値データを提供することである。図2. 1は、応用プログラムからデータ・ファイルのデータを編集するまでの過程と環境を示す。まず本支援システムはパソコン端末から応用プログラムのコピー（以下、コピー・プログラムと呼ぶ）を受け取り、これを解析してデータベースからデータ・ファイルにデータ

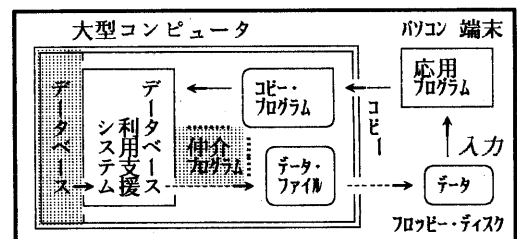


図2. 1 本支援システムの稼働環境。

を編集するプログラム（以下、仲介プログラムと呼ぶ）を生成する。次に仲介プログラムを実行してデータ・ファイル上に応用プログラムの入力データを編集し、それをパソコン端末のフロッピー・ディスクに転送する

この過程でパソコン利用者は、手元のパソコンをTSS端末として目的のデータベースがサービスされている大型コンピュータに接続し、パソコン上の応用プログラムのコピーを大型コンピュータに転送する。その後は、上記の説明の通り本支援システムが自動処理する。その後はデータ・ファイルのデータをフロッピー・ディスク上に転送し、パソコン上の応用プログラムをスタンドアロンに実行してデータを処理する。図2. 1で、応用プログラムが既に大型コンピュータの中にある場合には、本支援システムが編集した外部記憶装置上のデータ・ファイルを直接利用することになる。

これらの事柄は、熟練したコンピュータ利用者ならば自分でできる内容である。しかし問題なのは、対象にしている利用者がコンピュータの専門家ではないということである。しかも応用プログラムのコーディングに必要な知識については、彼が最もよく知っているであろう。本論文で述べる支援システムでは、このような状況で有効に働くことを前提に構成を考える必要がある。

本支援システムの動きを外部から見る限りでは、データベースからデータ・ファイルへのデータ書式変換が主たる機能である。そのためには、利用者の要求する書式のデータを編集する総ての方法を規則の形式で整理して、これに従ってシステムを起動させることも可能であるが、この方法ではシステム構成が複雑になる割にはその効果は薄いであろう。そこで本支援システムでは、データ編集の知識を応用プログラムの解析によって引き出し、この結果に基づいて仲介プログラムと外部スキーマを自動生成することに帰着させた。つまり、応用プログラムを仲介プログラムに変換する規則をシステムに組み込んで自動生成し、これら以外の処理はコマンド・プロシージャで実行する。3章では、その詳細について

説明する。

3. 仲介プログラムと外部スキーマの自動生成

自動生成される仲介プログラムは、図3. 1に示すように、少なくとも以下の4つの機能を含む必要がある：

- ① データベースおよびデータ・ファイルを開く（OPEN）機能。
- ② 利用者の求めているデータをデータベースから検索（RETRIEVE）する機能。
- ③ 検索されたデータをデータ・ファイル上に編集（EDIT）する機能。
- ④ データベースおよびデータ・ファイルを閉じる（CLOSE）機能。

この内で、①と②と④は、応用プログラムの構造に関係なく、データベース管理システムのデータ記述言語とデータ操作言語、およびプログラミング言語

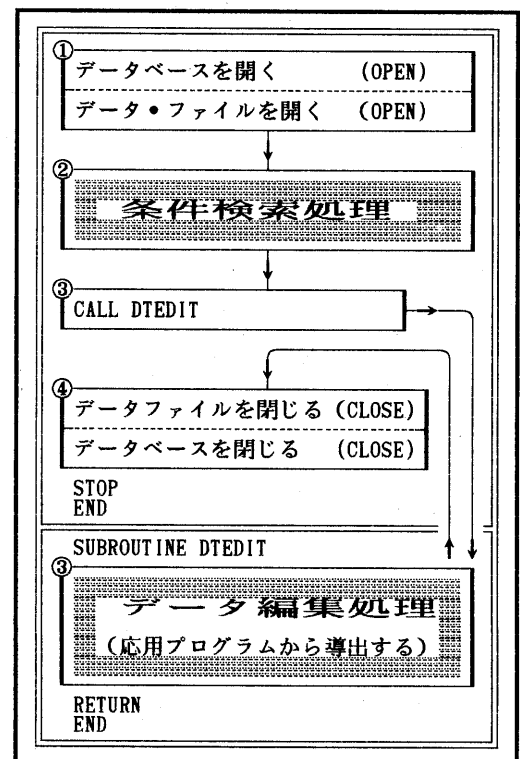


図3. 1 仲介プログラムの構造

やホスト・コンピュータのオペレーティング・システム（OS）のみに依存して決まる。

データ編集処理を行う③のサブルーチンは、プログラム変換による自動プログラミング⁴⁾同様に、コピー・プログラムの変換を通して自動生成される。この中で、データ編集処理の部分をわざわざサブルーチンとして結合した理由は、応用プログラムから導かれる部分の構造を独立に扱うことによって、変数名や文番号などが主プログラムの部分と重複するのを避けることができ、またプログラムの構造としても見易いからである。このようにして「①+②+③+④」を組み立てる手順は、部品合成による自動プログラミング⁵⁾に似ている。

図3. 2は、コピー・プログラムを受け取ってこ

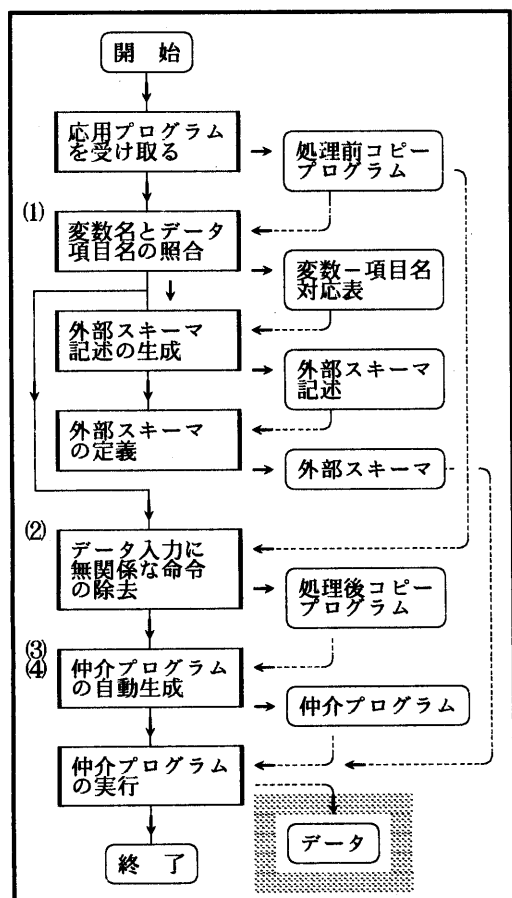


図3. 2 応用プログラムの解析から仲介プログラムの編集を経てデータ編集に至る手順。

れを解析し、その結果によって仲介プログラムと外部スキーマを自動生成して、これらの実行後に所定のデータを得るまでの処理の流れを示す。図の右側は、各段階で生成されたものである。以下に本章では、これら仲介プログラムと外部スキーマの自動生成の手順を説明する。

3. 1 応用プログラムの構文解析と前処理

仲介プログラムを自動生成するには、応用プログラム中の「データ入力文」と「その入力並び」、および「データ入力に対する制御文」の情報が必要である。ここでは、図3. 2の流れ図に沿って、仲介プログラム自動生成の前処理について説明する。

データ入力並びの変数名と

概念スキーマのデータ項目名の照合

仲介プログラムの自動生成には、概念スキーマのデータ項目名と応用プログラムの変数名の対応一覧表（以下、対応表と呼ぶ）が不可欠である。この対応表は、データ項目名と変数名をディスプレイ画面の左右に並べて表示し、利用者の指示にしたがって作成する（図3. 6参照）。このようにして作成された対応表は、仲介プログラム実行の時点まで保存されて、仲介プログラムと外部スキーマの自動生成に必要な情報として使われる。

仲介プログラムにとって不要な文の除去

応用プログラムの入力データを外部記憶装置上に編集するには、これと同じ書式のデータを同じ構造でデータ・ファイル上に編集するプログラムをつくらなければならない。この論法にしたがうならば、「応用プログラム内のデータ入力命令をデータベース・アクセスとデータ出力に読み換える」という方法で仲介プログラムのデータ編集処理サブルーチンを自動生成できる。したがって、応用プログラムを仲介プログラムに導くには、まず応用プログラムのデータ入力以外の部分を除去することから始める。

応用プログラム中のデータ入力に関係の無い文を次の規則に従って除去する。

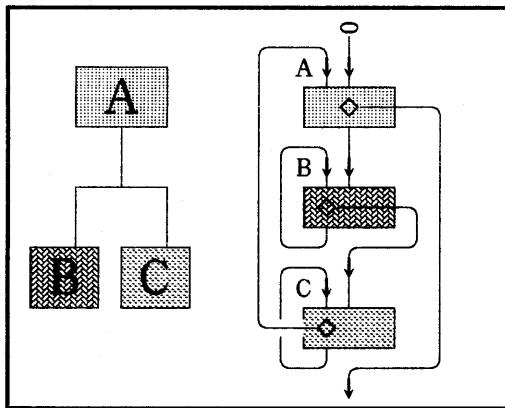


図3. 5 階層型データ構造とデータ転送手続きループの入れ子構造との関係。

順は、仲介プログラムでは入れ子構造のループ処理を形成する。仲介プログラム自動生成におけるこれら入れ子のループは、「②における不要な文の除去の操作」や、「図3. 3に示した置き換え操作」の後の結果として自動的に導かれるので、特に新たな取り扱いはいらない。

ここまでの文の除去や置き換え操作の過程で、飛越し先やループ処理に使われた文番号に間違った処理をしないように注意する必要がある。この点に関して本システムでは、文の書き換えのタイプに合わせて、文番号の位置のチェックをしている。

3. 3 条件検索機能及びファイルの OPEN/CLOSE 処理の追加と外部スキーマの定義

前節までの議論で、データ編集処理のサブルーチン部分が生成される。さらに本章の始めの図3. 1で示した構造にするためには、主プログラムを合成する必要がある。

条件検索機能及び OPEN/CLOSE 処理の追加

最初にも述べたように、データベースやファイルの OPEN/CLOSE は、実際に使用したコンピュータやデータベース管理システムに直接依存する形式で記述する。この詳細は、一般性がなくかつ冗長になるので説明を割愛する。

編集すべきデータを選び出す条件検索処理は、条

件検索式、和集合・積集合の演算式などの入力とその処理を簡易な会話で実行できるようにシステムを作成すればよい。次ページで本支援システムの会話手順（図3. 6参照）を説明することで、検索方法の具体的な内容説明に変える。

外部スキーマの定義

一般に、概念スキーマの中で当面使われる項目名が決まれば、外部スキーマの記述は一意に決まる。そこで本支援システムでは、概念スキーマの内の対応表にある項目名を含んだ部分構造を抜き出して、これを外部スキーマ定義用ジョブ・デック記述に変換する。続いてこのジョブ・デックを実行して外部スキーマを定義する。外部スキーマの定義は、仲介プログラムと同様に、本支援システムが生成するコマンド・プロシージャによって処理される。

SYSTEM ? <u>ECDB</u> ←①	
***** * WELCOME TO ECDB DATABASE * *****	
MODULE ? <u>DATA CONVERSION</u> ←②	
TYPE IN THE FILE NAME OF YOUR PROGRAM? I*? <u>/SIMULATION</u> ←③	
PLEASE IDENTIFY DATA ITEM AND VARIABLE NAME.	
VARIABLE NAME	DATA ITEM NAME
XX	SNO
YY	CS
ZZ	TEMP
{	PH
	{
I*? <u>XX=CS; YY=PH; ZZ=TEMP</u> ←④	
ITEM TABLE IS GENERATED.	
START INFORMATION RETRIEVAL.	
CONDITION ? <u>/CS > 1.0 AND PH < 4/</u> ←⑤	
RECORD COUNT; 12	
DATA THE DATA FILE (YES OR NO)? <u>YES</u> ⑥	
BEGIN DATA EDDITION. DATA EDITION HAS FINISHED.	
* ← 通常のシステム・コマンドを入力する。	

図3. 6 本利用支援システムの利用時の会話。

3. 4 本支援システムの稼働

図3. 6には、本支援システムの稼働の様子を示した。図の中で、下線のついた文字列は、利用者の入力文字であり、番号①～⑥では次の指示をする：

- ① 利用するデータベースを呼び出す。
- ② データベースに対してデータ書式変換処理の要求をするコマンドの入力である。
- ③ 応用プログラムの保存されているファイル名を入力する。
- ④ 直前に示された変数名と項目名の一覧表を参考にして変数名と項目名の対応を入力する。この入力を受けて、本支援システムは対応表を作成する。
- ⑤ 条件検索処理のための検索条件を入力する。
この検索条件式は、別に定めるデータ操作言語の文法に従って入力される。このほかにも、検索済みサブ・セット間の和集合や積集合の演算式の入力も可能である。
- ⑥ データ編集の是非に対する応答。ここで NO を入力すると⑤に戻り、YES を入力すると、データ編集処理に移る。

図3. 6の④と⑤の入力は、本支援システムに固有の文法規則に従う必要がある。この詳細な説明は、紙数も制限されているので、ここでは割愛する。

4. 本支援システムの構成上の考察

本支援システムは、3章の図3. 1の左に並んだ各段階の個々の事柄を処理する副モジュール群と、それら全体を制御する主モジュールで構成される。副モジュール群には、仲介プログラムの自動生成とその実行をホスト・コンピュータやデータベース管理システムの下で支障なく処理するのに必要な規則が組み入れられている。そして図4. 1は、このようなシステム構成の背後にある、本支援システムに関与するソフトウェア群とその階層関係を示す。

図4. 1の下半分の網掛け部分は、利用者からは直接には見えない本支援システム内部に包み隠された部分である。そして上側は、本支援システムが生成した仲介プログラムと外部スキーマである。本支

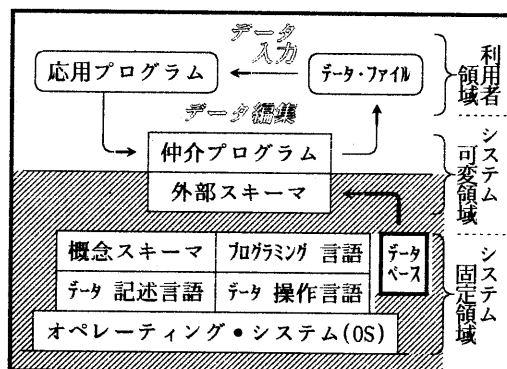


図4. 1 本利用支援システムを支える諸要素

援システムは、これらのソフトウェア構成を土台にしたうえで、データの編集と転送に必要な図2. 1示した処理を逐次実行していく。図右は、本支援システム稼働の際に関与する様々なソフトウェアの区分である。この中で、固定領域のみが本支援システム本体として実行される部分であると考えてよい。

このように本支援システムには、仲介プログラムと外部スキーマの自動生成のための規則や背後で稼働しているオペレーティング・システム(OS)やデータベース管理システムの実行に関わる様々な事柄を規則的形式で間接的に組み込んでいる。したがって本支援システムの構成は、処理している個々の内容が大掛かりであるにも関わらず、システム・プログラムは比較的小規模になっている。

社会の情報化が進展する中で、プログラミング言語の標準化^{6,7)}やデータベース管理システムの標準化^{8,9)}が進められており、これらがコンピュータ利用者に共通な標準言語として役立つことは、一般的に見て自然な考え方であろう。そこで、本支援システムの構成についても、FORTRAN や BASIC を応用プログラム記述用のプログラミング言語とみなす。データベース管理システムに関しても、プロトタイプには階層型のINQを使用した¹⁰⁾が、この点についても仕様の標準化の動向に従うであろうメーカーの対応に追従すればよいであろう。

5. インプリメンテーション

本支援システムは、大型コンピュータ ACOS システム 2000（日本電気製）、オペレーティング・システム ACOS-6 の下で稼働している。データベース管理システムには、キーワード検索機能を備えた階層型の INQ¹⁰⁾ を使った。これらは、4 章でも述べたように、仲介プログラムの諸規則の中に吸収される。

本支援システムを記述したシステム・プログラムには、FORTRAN と LISP を使った。FORTRAN は、ディスプレイ画面への表示やキーボードからの入力の方が比較的容易なので、図 3. 6 の会話手順とその周辺の処理に使っている。一方、3 章で議論した諸規則に関する処理部分のモジュールは、条件と実行といった組合せの処理が多いので、UTILIST¹⁰⁾ を使って記述した。

6. まとめ

本支援システムは、応用プログラムの要求に合わせた書式の入力データを数値情報データベースから再編集することを役割とする。本文で議論した仲介プログラムは、このデータ再編集規則の内の手続き的内容を FORTRAN プログラムの形式で自動生成したものである。本支援システムには、固定領域として仲介プログラム自動生成の規則を LISP 言語で記述した。つまり、一般的知識を固定領域に置き、利用者ごとに変化するデータ再編集の知識を可変領域部とした。このように一般的知識とその時々に変わる状況に依存する個別的な知識を区別することで、本支援システムを簡潔な構造にすることが可能になった。

一般に、支援システムが親切であるためには、利用者の個別的な特殊事情に対して出来る限り具体的に対応する必要がある。ここでは、データベース利用支援を実例に、上記の考え方で知識処理の一適用例として試みてきた。幸い、知識処理の道具としてのソフトウェアが整備されてきた現状では、これらを活用した一層高度なコンピュータ利用やその支援を促進するための機能の充実が可能になっている。情報化の円滑な推進の一助として本利用支援システ

ムのようなものを整備することには、それだけの意義があるであろう。

参 考 文 献

- 1) 磯本征雄, 他: 学術研究用たん白質データベース PROTEIN-DB, 情報処理学会誌, Vol.21, No.1, pp.15-22 (1980).
- 2) Isomoto, Y., et al.: Knowledge based DBMS for Non-professional Database Administrators, Proceedings of Delivering Computer Power to End Users, COMPCON 83 fall, IEEE COMPUTER SOCIETY, pp.514-522 (1983).
- 3) 全国共同利用大型計算機センター: オンライン・データベース利用ガイド, 全国共同利用大型計算機センターデータベース連絡会版, (1987).
- 4) 吉田紀彦: プログラム変換による自動プログラミング—特集号: 自動プログラミング—, 情報処理, Vol.28, No.10, pp.1320-1328 (1987).
- 5) 古宮誠一, 原田実: 部品合成による自動プログラミング—特集号: 自動プログラミング—, 情報処理, Vol.28, No.10, pp.1329-1345 (1987).
- 6) 菅 忠義: FORTRAN—特集号: プログラミング言語の最近の動向—, 情報処理, Vol.22, No.6, pp.452-456 (1981).
- 7) 木下 旬: BASIC—特集号: プログラミング言語の最近の動向—, 情報処理, Vol.22, No.6, pp.452-456 (1981).
- 8) 横山一郎: データベース言語NDL—小特集: データベースの標準化動向—, 情報処理, Vol.29, No.3, pp.201-207 (1988).
- 9) 芝野耕司: データベース言語SQL—小特集: データベースの標準化動向—, 情報処理, Vol.29, No.3, pp.208-214 (1988).
- 10) 日本電気: ACOS ソフトウェア, データベース管理, INQデータ操作言語説明書.
- 11) 日本電気: ACOS ソフトウェア, プログラム管理, UTILISP説明書.