

人名に付属する多数のアトリビュート毎の仮想

ランダム・ファイルと同時にOPENでき、処理系

(株) 佐々木彬夫事務所      佐々木 彬夫  
 (株) エー・ピー・アイ      有水 章  
 (株) ジャパン・コミュニケーション・インスティテュート  
    若槻 史郎、吉田 等

## 1. 概要

パーソナル・コンピュータ PC-9800 に位速のマーク・カード・リーダーを接続し、マーク並びにキー入力により人名に付属するアトリビュート(住所、生年月日、.....)を128個入力した上処理を行なうシステムを開発した。

128個のアトリビュートのうち30個はマーク・カードから直接読み取り、49個はキー入力、残りの49個はマークから値を計算する様になっている。

処理は1000個までの人名とそれに付属するアトリビュートの登録と、複数のアトリビュートの値に関する論理的な条件が成立する人名の抽出、並びにその人名に付属する住所などの印字出力が主な内容である。

アトリビュートの値に関する論理的な条件としては、30個までのアトリビュートに関して

アトリビュートM=ある範囲 AND アトリビュートN=ある範囲 AND

D.....

を指定して検索/抽出を行なう事ができる様になっている。図1にユーザによる検索の具体例を示す。

本システムの特徴は、アトリビュートの格納方式の工夫により図1の様な検索が高速である事と、マーク・カード・リーダーのインターフェースの工夫により図2の様なマークでも設けりなく読み取れる事であるが、ここではアトリビュートの格納方式と検索のスピードについて報告する。

本システムで採用したアトリビュートの格納方式とアクセスの方式により、プログラムの「業務処理部分」では各アトリビュートがそれぞれ独立の仮想ランダム・ファイルに格納されており、かつそれらのファイルが全てOPENの状態と見做して、1000名分を12万8000個の独立したデータにランダム・アクセスを行なう事ができる。

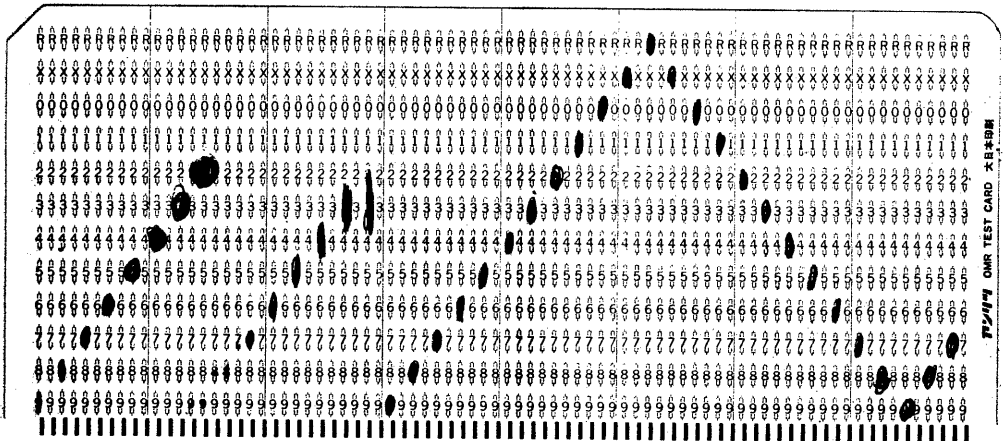
シヨウケン ノ シテイ

- |                         |                                    |                 |
|-------------------------|------------------------------------|-----------------|
| 1. セイネン = . 年           | 2. タンシヨウヒ = 月 日                    | 3. ネンダクイ = 3    |
| 4. セイノツ = 1             | 5. カソク コウセイ >                      | 6. オキヤクサマ =     |
| 7. オスマイ =               | 8. セタイ =                           | 9. ショクキヨウ =     |
| 10. シュエ =               | 11. チイキコト =                        | 12. コキヤク フンルイ = |
| 13. ショウヒン =             |                                    |                 |
| 14. オカイケ ( ) = 年 月 イセ   | 15. 21セイキ テスト ( ) >                | 16. ハイテツト ( ) > |
| 17. コウニウ ヱコト ( ) >      | 18. シヨウキヨウ イコウ ノ ヱコト ( ) ... (Y/N) |                 |
| 19. オウハン ヱウイト、ソノタ ( ) > | 20. △シヨウケン(センプ) ... (Y/N)          |                 |

コレノ シヨウケンヲ ミタス オキヤクサマハ、8マイ サマ デス!

1. OK      2. サラニ シヨウケンヲ シテイ.....(ハ)ンコウ=(

四一. ユーザによる検査の具体例。性別が1(男)で毎分が3ごちよ人名が3個ちの事を表している。



四二. 本システムで読み取れるマークの例(一番ひどいマーク以外は正しく読み取)事がごさい)

## 2. 本システムの必要性

パーソナル・コンピュータではフロッピー・ディスクにデータを格納するのにはシーケンシャル・ファイルとランダム・ファイルを使用する事ができるが、本システムの様に不特定多数のデータの照合/検索を行なう場合にはランダム・ファイルを使用しなければならぬ。しかし、次の様な問題があるため、人名並びにそれに付属する各アトリビュートを単なるランダム・ファイルに格納して処理する事は現実的ではない。

例えば、PC-9800 28インチ(2D)のフロッピー・ディスクを使用する場合に例えば、ディレクトリのエントリは最大151個、同時にオープンできるファイルは15個までであり、ディスク上のエリアの割当は1トラック=26セクタ=6656Bをその単位とし、ディスクとプログラムの間の出入力は256B(固定)のレコードを単位とする様になっている。

この事は、プログラム上で同時に使用できるファイルは15個までであり、あるファイルの内容がただの1Bであったとしてもディスク上ではそのファイルが1トラックを占有する事を意味する。

本システムでは1000個の人名(16B)に付属する各アトリビュートが例えば住所(60B)、特定の商品の所持の有無(1B)などの様に、それぞれのバイト長が異なるので、上記の様なファイルシステムでは実用的なシステムを作るのが困難である。

この問題に対処するため、本システムでは図3に示す様にデータ格納用にディスク全面を占有する"DATA"という1MBのデータ・ファイルを用い、"DATA"のエリアを128個のアトリビュートの格納用に分割して使用している。

```
FILES 2
DATA      151
OK
```

図3. 本システムではただ1個のデータ・ファイル "DATA" のエリアを分割してアトリビュートを格納している。

従って、特定のアトリビュートのN番目( $N=1\sim 1000$ )のエントリにアクセスするためにはその度毎に"DATA"内における格納場所(レコード番号とバイト位置)を計算している。

本システムにおけるこのマッピングもBASICで書かれておき、「業務処理部」からは見えないうちになっている。

この様なマッピングをBASICで書くことも実用上差支えないという事は

重要な事実であり、この事とマッピングの具体的方法、並びにシステムとして実現した性能がこの報告の骨子である。

### 3. 本システムの仮想ランダム・ファイル

上述の様に本システムでは "DATA" というただ1個のランダム・ファイル上のエリアを分割して128個のそれぞれ長さが異なるアトリビュートの格納に用い、アクセスの度毎に格納場所を計算している。そして、「業務処理部分」からはこのマッピングが見えない様になっている。

具体的には、例えば `VR% = 325: GOSUB *R.ADRS: PRINT INF$` という処理によって325番目の人名に付属する住所をCRTの画面に表示できる( `VR%` は *Virtual Record* の番号を表示する変数)。

書き込みの時には、 `VR% = 325: INF$ = "ヨコハマ シ....."` : `GOSUB *W.ADRS` によって325番目の住所の所に文字列を書く事ができる。

`R` は Read、 `W` は Write であり、 `INF$` は Information 授受用の変数である。

"DATA" 自体はランダム・ファイルであるが、上記の様方 Read/Write も具体的にはランダム・ファイルのレコードの `R/W` の実行であり、「業務処理部分」では各アトリビュート毎に専用のランダム・ファイルがあり、そのエントリにアクセスしては種り種りフロッピー・ディスクを使用する事ができる。

これが本システムの仮想ランダム・ファイルの意味で、プログラム起動時に `OPEN "2:DATA" AS #1` を実行するのみで個々の仮想ランダム・ファイルの `OPEN/CLOSE` は不要であり、この事も実用上大きな便益となっている。

本システムの名レコードは人名とそれに付属するアトリビュート全体の集合であるが、各アトリビュートが "DATA" 内で分散格納されている様子を図4に示す。

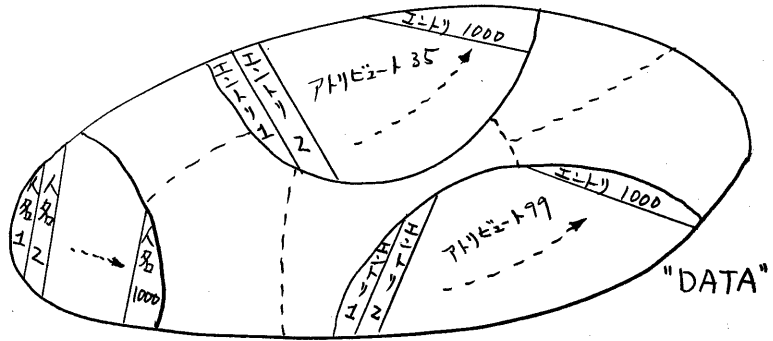


図4. 各アトリビュートの "DATA" 内の格納状態

## 4. マッピングとアトリビュートのアクセスの詳細

本システムでは各アトリビュートを "DATA" 内に分散して格納していき、  
 上記の通りであり、例えば電話番号はレコード52から1エントリ当たり  
 12Bを割り当てて1000個を格納する様になっている。

アクセスの具体的な方法は、例えば433番目の電話番号として "123-  
 456-7890" を書いた後で読み出してCRTに表示して見ると処理は図5  
 の通りである。

```

VR% = 433: INF$ = "123-456-7890": GOSUB *W.TEL'.... カキコミ
VR% = 433: GOSUB *R.TEL: PRINT INF$'.... ヨミ出し
    END
    
```

```

50260 *MAP.TEL
50270 PR%=52+(VR%-1)*21: N1%=12*((VR%-1) MOD 21): N2%=12: RETURN
50730
50740 *ACCESS
50750 FIELD #2, N1% AS DUMMY$, N2% AS DD$: GET #2, PR%: RETURN
50760
50770 *WDISK
50780 GOSUB *ACCESS: LSET DD$=INF$: PUT #2, PR%: RETURN
50790
50800 *RDISK
50810 GOSUB *ACCESS: INF$=DD$: RETURN
50820
51080 *W.TEL
51090 GOSUB *MAP.TEL: GOSUB *WDISK: RETURN
51680 *R.TEL
51690 GOSUB *MAP.TEL: GOSUB *RDISK: RETURN
    
```

図5. "DATA" のレコード52から始まる格納領域内の電話番号  
 にアクセスする処理

上のリスト中 VR% は前述の通り *Virtual Record* の番号を表示する変数であり、  
 PR% は *Physical Record* ("DATA" のレコード) の番号を表示する変数である。

サブルーチン \*MAP.TEL が電話番号に関するマッピングの計算であり、  
 図6にその意味を示す。

図5と図6は本システムで行っているマッピングの一例を示すものであるが、  
 一般には次の様にしてマッピングを行えばよい。

$$PR\% = \text{先頭物理レコード番号} + (VR\% - 1) * (\text{最大データ收容個数} / \text{レコード})$$

$$N1\% = \text{データ長} * ((VR\% - 1) \text{ MOD } (\text{最大データ收容個数} / \text{レコード}))$$

この様なマッピングを行なう場合、一般に "DATA" のレコード長である

よ256Bはデータ長の整数倍ではなから、各レコードの終りに無駄が生ずる。マッピングの形を更に工夫してこの無駄をなくする事も可能であるが、実用上の便益は左程大きくないと考えられる。

VR%	PR%	N1%	N1%
1	52	0	12
2	52	12	12
3	52	24	12
4	52	36	12
5	52	48	12
6	52	60	12
7	52	72	12
8	52	84	12
9	52	96	12
10	52	108	12
11	52	120	12
12	52	132	12
13	52	144	12
14	52	156	12
15	52	168	12
16	52	180	12
17	52	192	12
18	52	204	12
19	52	216	12
20	52	228	12
21	52	240	12
22	53	0	12
23	53	12	12
24	53	24	12
25	53	36	12
26	53	48	12
27	53	60	12
28	53	72	12
29	53	84	12
30	53	96	12

図6 マッピングの計算ルーチン \*MAP, TEL の処理

本システムでは前述の様に特定の属性の値が範囲に入っている人名を抽出する事が主要な機能となっているので、その属性の値を土着目のエントリから1000番目のエントリまで、しらみつぶしに調べなければならぬ。

人名のみであればハッシングなどの手法も使えようが、他の属性は端から調べて行かなければならず、システム全体としての処理速度はこの事により上限が抑えられる。

そしてその際には決定的な要因は属性1000個の値を調べよの"DATA"のレコードへのアクセスが何回必要かという事である。

本システムでは例えば人名(16B)は"DATA"の1レコード当り16個格納されているから、"SASAKI AKI O"が登録済みか否かを1000個の人名をしらみつぶしに調べる場合のディスク・アクセスの所要回

数は1/16となる。

更に、大多数を占める1エントリが1Bのアトリビュートでは僅か4回のVコードを調べるのみで照合/検索が終了する。

この事が本システムにおいて処理速度の大幅な向上が実現した基本的な理由である。

## 5. 実現した性能

上述の様な工夫の結果、本システムでは1000個の人名のチェックが25秒、アトリビュート1000個のチェックが35~90秒で終了する。

本システムの開発は1983年6月に終了し、現在既に全国各所で実用に供されている。

## 6. 謝辞

本システムの開発には多くの方々が必要な立場で協力された。これらの方々に心からの謝意を表す。