

国立国語研究所学術情報リポジトリ

Chinese character indexes making program by COBOL

メタデータ	言語: jpn 出版者: 公開日: 2017-03-31 キーワード (Ja): キーワード (En): 作成者: 石綿, 敏雄, ISIWATA, Tosio メールアドレス: 所属:
URL	https://doi.org/10.15084/00000998

COBOLによる漢字索引作成

石 綿 敏 雄

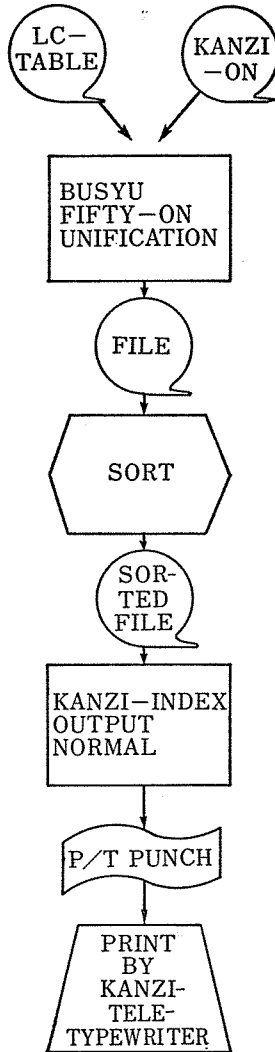
国語研究所では漢字テレタイプを用いて、漢字かなまじり文の処理を行なっている。この処理プログラム作成に際して、あるいは、アウト・プットのポスト・エディティングの作業のために、漢字コードの各種（漢字コード、漢字部首、漢字50音など）の索引を備えておくのが便利である。このうちのいくつかのものでできているが、すでにできている二つの漢字テーブルを用いて、各種の漢字索引を作成するプログラムを作成したので、これについて報告する。

（漢字テーブル二種については、ともに国語研報告31に収めた次の二編、松本昭「国研用漢字テレタイプと同機利用の言語情報処理」、田中章夫「電子計算機によるワードリスト作成上の一問題」を参照）。

国語研究所でできている漢字テーブルには、長単位語処理のプログラムで用いられる漢字代表音テーブルと（“KANZ I-ON”）漢字集計のときに用いられる漢字部首コード（“LC-TABLE”）のテーブルがあり、それぞれ一巻ずつの磁気テープになっている。この両テーブルをまとめて一巻とし、その各部分についてソート（分類）し、これをハンドブックとして使用しやすいように編集して高速紙テープさん孔装置で紙テープにさん孔し、これをオフラインの漢テレで印字するのである。プログラムはこの三部分からなるが、紙テープ出力のプログラムは二本用意し、実際は二本目につくった、ページ式印字のプログラムを使用した。アウトプットとしては部首順漢字索引、漢テレのなかでの漢字自身のコードによる索引、代表音による漢字索引である。出力前にそれぞれの部分によるソートを行なっておく。出力プログラムは同一のものを用いるので、フォーマットは同一であるが、データの順序だけが異なるのである。¹⁾

プログラム言語としてはCOBOLを使用した。この言語は最近次第に広く用いられてきており、言語情報処理のためにも COBOL が便利であるといわれる（「数理科学」1968年9月号に収められた各論文）。ただ、使用計算機によって

GENERAL FLOW CHART



は、その命令体系や、記憶装置の構造から、アセンブラや（あるいは、ばあいによっては、本当の機械語）で書いた方が扱いやすいこともないではない（たとえば、HITAC3010を用いたばあい可変長データの扱いに関しては、COBOLで書くよりもマシン・コードやアセンブラで書く方がずっとわかりやすい）。しかし異機種間の互換性があること、ドキュメンテーションがよいこと、プログラム作成者以外のものにもプログラムがわかること、そのために、プログラムの内容の変更ができること、などの理由から、一般的に言えば、業務用のプログラムは、コンパイラ言語で書いた方がよいと、筆者は考えている。このプログラムをコンパイラで書いたのは、そのようなことをためしてみるという目的もあったのである。アセンブラで書いたプログラムは公開しても意味がない

```

001010 IDENTIFICATION DIVISION.
001020 PROGRAM-ID. BUSYU-FIFTY-ON-UNIFICATION.
002010 ENVIRONMENT DIVISION.
002020 CONFIGURATION SECTION.
002030 SOURCE-COMPUTER. 3010, MEMORY SIZE 20000 CHARACTERS, PAPER-READER
002040 PRINTER 1, SIMULTANEOUS-CONTROL.
002050 OBJECT-COMPUTER. 304, NO ROLLBACK.
002060 INPUT-OUTPUT SECTION.
002070 FILE-CONTROL.
002080 SELECT FIFTY-ON-ORDER-FILE ASSIGN TO TAPES 1.
002090 SELECT BUSYU-ORDER-FILE ASSIGN TO TAPES 2.
002100 SELECT OUTPUT-TAPE-FILE ASSIGN TO TAPES 3.
003010 DATA DIVISION.
003020 FILE SECTION.
003030 FD FIFTY-ON-ORDER-FILE BLOCK CONTAINS 200 RECORDS,
003040 RECORD CONTAINS 10 CHARACTERS, LABEL RECORDS ARE STANDARD,
003050 VALUE OF ID "KANZI-ON", ACTIVE-TIME IS 9998, DATA RECORD IS FIFTY-ON-ORDER.
003060 01 FIFTY-ON-ORDER.
003070 02 CHIN-CH-FO-FILE, PICTURE IS X(2).
003080 02 FIFTY-ON-FO-FILE, PICTURE IS X(8).
003090 FD BUSYU-ORDER-FILE RECORD CONTAINS 6 CHARACTERS, LABEL RECORDS
003100 ARE STANDARD, VALUE OF IDENTIFICATION IS "LC-TABLE", ACTIVE-TIME IS 9998,
003110 DATA RECORD IS BUSYU-ORDER.
003120 01 BUSYU-ORDER.
003130 02 CHIN-CH-BO-FILE, PICTURE IS X(2).
003140 02 BUSYU-BO-FILE, PICTURE IS X(4).
003150 FD OUTPUT-TAPE-FILE BLOCK CONTAINS 100 RECORDS, RECORD CONTAINS
003160 14 CHARACTERS, LABEL RECORDS ARE STANDARD, VALUE OF ID IS
003170 "TABLE-01", ACTIVE-TIME IS 9998, DATA RECORD IS OUTPUT-AREA.
003180 01 OUTPUT-AREA.
003190 02 CHINESE-CHARACTER, PICTURE IS X(2).
003200 02 BUSYU-OP-FILE, PICTURE IS X(4).
003210 02 FIFTY-ON-OP-FILE, PICTURE IS X(8).
004010 PROCEDURE DIVISION.
004020 START. OPEN INPUT FIFTY-ON-ORDER-FILE, BUSYU-ORDER-FILE,
004030 OUTPUT OUTPUT-TAPE-FILE.
004040 READ-FO-FILE. READ FIFTY-ON-ORDER-FILE, AT END GO TO FINISH-R.
004050 READ-BO-FILE. READ BUSYU-ORDER-FILE, AT END GO TO FINISH-R.
004060 IF CHIN-CH-FO-FILE IS EQUAL TO CHIN-CH-BO-FILE, GO TO
004070 NORMAL-EDITING, OTHERWISE GO TO SPECIAL-EDITING.
004080 NORMAL-EDITING. MOVE SPACES TO OUTPUT-AREA, MOVE CHIN-CH-FO-FILE
004090 TO CHINESE-CHARACTER, MOVE FIFTY-ON-FO-FILE TO
004100 FIFTY-ON-OP-FILE, MOVE BUSYU-BO-FILE TO BUSYU-OP-FILE,
004110 WRITE OUTPUT-AREA, GO TO READ-FO-FILE.
004120 SPECIAL-EDITING. MOVE SPACES TO OUTPUT-AREA MOVE CHIN-CH-BO-FILE
004130 TO CHINESE-CHARACTER, MOVE BUSYU-BO-FILE TO BUSYU-OP-FILE,
004140 WRITE OUTPUT-AREA, GO TO READ-BO-FILE.
004150 FINISH-R. CLOSE FIFTY-ON-ORDER-FILE, BUSYU-ORDER-FILE,
004160 OUTPUT-TAPE-FILE, STOP RUN.

```

が、コンハイラで書いたものは、人に見せることもできよう。

次にプログラム言語それ自体についての関心のあり方もある。プログラム言語自体も一つの「言語」であり、一つの *langue artificielle* である。(計算機の普及にともない、大衆化にともなって、その「言語」の問題はいよいよ重要な意味をもってくると思う。その「言語」の問題は、じつは言語学者ももっと関心をもってよい、重要なことがらであると思われるが、実情は、ごく一部の^{人々を除いて}まだほとんど関心がないというところである)。 *langue naturelle* 自然語(に近い言語)でプログラムを作成するということは、「巨視的」にみればあい、電子計算機のある種の使い方にとっては一つの理想であろう。その意味で、同じく COBOL 言語を使用しても、プログラム作成上できるだけ自然

```
010100 IDENTIFICATION DIVISION.
010200 PROGRAM-ID. KANZI-INDEX-OUTPUT-SIMPLIFIED.
010300 ENVIRONMENT DIVISION.
010400 CONFIGURATION SECTION.
010500 SOURCE-COMPUTER. 3010, MEMORY SIZE 20000 CHARACTERS,
010600 PAPER-READER, PRINTER 1, SIMULTANEOUS-CONTROL.
010700 OBJECT-COMPUTER. 304, NO ROLLBACK.
010800 INPUT-OUTPUT SECTION.
010810 FILE-CONTROL.
010900 SELECT INPUT-FILE ASSIGN TO TAPES 1.
011000 SELECT OUTPUT-FILE ASSIGN TO PAPER-PUNCH.
020100 DATA DIVISION.
020200 FILE SECTION.
020300 FD INPUT-FILE, BLOCK CONTAINS 100 RECORDS, RECORD CONTAINS
020400 14 CHARACTERS, LABEL RECORDS ARE STANDARD, VALUE OF ID IS
020500 "TABLE-02", ACTIVE-TIME IS 9998, DATA RECORD IS INPUT-AREA.
020600 01 INPUT-AREA.
020700 02 KANZI, PICTURE IS X(2).
020800 02 BUSYU-CODE, PICTURE IS X(4).
020900 02 GOZYUON, PICTURE IS X(8).
021000 FD OUTPUT-FILE, RECORD CONTAINS 20 CHARACTERS, LABEL RECORD IS
021100 OMITTED, DATA RECORD IS OUTPUT-AREA.
021200 01 OUTPUT-AREA.
021300 02 KANZI, PICTURE IS X(2).
021400 02 FILLER, PICTURE IS X(2).
021500 02 BUSYU-CODE, PICTURE IS X(4).
021600 02 FILLER, PICTURE IS X(2).
021700 02 GOZYUON, PICTURE IS X(8).
021800 02 CARRIAGE-RETURN-POSITION, PICTURE IS X(2).
022000 CONSTANT SECTION.
022100 77 CARRIAGE-RETURN-SYMBOL, PICTURE IS X(2), VALUE IS "*"J".
030100 PROCEDURE DIVISION.
030200 START. OPEN INPUT INPUT-FILE, OUTPUT OUTPUT-FILE.
030300 READ-INPUT-FILE. MOVE SPACES TO INPUT-AREA, MOVE SPACES TO OUTPUT-AREA.
030400 READ INPUT-FILE, AT END GO TO END-ROUTINE.
030500 MOVE KANZI IN INPUT-AREA TO KANZI IN OUTPUT-AREA.
030600 MOVE BUSYU-CODE IN INPUT-AREA TO BUSYU-CODE IN OUTPUT-AREA.
030700 MOVE GOZYUON IN INPUT-AREA TO GOZYUON IN OUTPUT-AREA.
030800 MOVE CARRIAGE-RETURN-SYMBOL TO CARRIAGE-RETURN-POSITION.
030900 WRITE OUTPUT-AREA, GO TO READ-INPUT-FILE.
031000 END-ROUTINE. CLOSE INPUT-FILE, OUTPUT-FILE.
031100 STOP RUN.
```

語に近づけ、自然の文章に近づけて書いてみることも、一つのころみとして無益ではないと思われる。もちろんこれにとどまらず、コンパイラ言語としても更に便利なものを考えてゆくのが当然であろうが、そのためには COBOL で書ける自然さを追及してみることも必要であろう（その一つとして、ここでは、一般の人、計算機に縁のない人でもわかりやすいように、くりかえしの部分はなるべくいちいち展開して書くということを徹底してみた。もちろんそれではプログラムのうまみはないのだが）。このような、コンパイラを使うこと、および、実用実務に供することの二つの目的をもって作成したが、もちろん後者の実用という方が優先するから、コンパイラの書き方の検討ということとはあまり十分にはためされていない。

```

000100 IDENTIFICATION DIVISION.
000200 PROGRAM-ID. KANZI-INDEX-OUTPUT-NORMAL.
000300 ENVIRONMENT DIVISION.
000400 CONFIGURATION SECTION.
000500 SOURCE-COMPUTER. 3010, MEMORY SIZE 20000 CHARACTERS,
000600 PAPER-READER, PRINTER 1, SIMULTANEOUS-CONTROL.
000700 OBJECT-COMPUTER. 304, NO ROLLBACK.
000800 INPUT-OUTPUT SECTION.
000900 FILE-CONTROL.
001000 SELECT INPUT-FILE ASSIGN TO TAPES 1.
001100 SELECT OUTPUT-FILE ASSIGN TO PAPER-PUNCH.
010100 DATA DIVISION.
010200 FILE SECTION.
010300 FD INPUT-FILE, BLOCK CONTAINS 100 RECORDS, RECORD CONTAINS 14
010400 CHARACTERS, LABEL RECORDS ARE STANDARD, VALUE OF ID IS "TABLE-02",
010500 ACTIVE-TIME IS 9998, DATA RECORD IS INPUT-AREA.
010600 01 INPUT-AREA.
010700 02 KANZI.
010800 03 KANZI-CODE-1, PICTURE IS X(1).
010900 03 KANZI-CODE-2, PICTURE IS X(1).
011000 02 BUSYU-CODE.
011100 03 BUSYU-CODE-1, PICTURE IS X(1).
011200 03 BUSYU-CODE-2, PICTURE IS X(1).
011300 03 BUSYU-CODE-3, PICTURE IS X(1).
011400 03 BUSYU-CODE-4, PICTURE IS X(1).
011410 02 GOZYUON.
011420 03 KANA-1, PICTURE IS X(2).
011430 03 KANA-2, PICTURE IS X(2).
011440 03 KANA-3, PICTURE IS X(2).
011450 03 KANA-4, PICTURE IS X(2).
011500 FD OUTPUT-FILE, RECORD CONTAINS 36 CHARACTERS, LABEL RECORD IS OMITTED,
011600 DATA RECORD IS OUTPUT-AREA.
011650 01 OUTPUT-AREA.
011700 02 KANZI, PICTURE IS X(2).
011800 02 FILLAR, PICTURE IS X(2).
011900 02 BUSYU-CODE, PICTURE IS X(4).
012000 02 P-POSITION, PICTURE IS X(2).
012100 02 GOZYUON.
012200 03 KANA-1, PICTURE IS X(2).
012300 03 KANA-2, PICTURE IS X(2).
012400 03 KANA-3, PICTURE IS X(2).
012500 03 KANA-4, PICTURE IS X(2).
012600 02 FILLAR, PICTURE IS X(2).
012700 02 CODE-AREA.
012800 03 CODE-1, PICTURE IS X(1).
012900 03 KANZI-CODE-1, PICTURE IS X(1).
013000 03 CODE-2, PICTURE IS X(1).
013100 03 KANZI-CODE-2, PICTURE IS X(1).
013200 03 FILLAR, PICTURE IS X(2).
013300 03 CODE-3, PICTURE IS X(1).
013400 03 BUSYU-CODE-1, PICTURE IS X(1).
013500 03 CODE-4, PICTURE IS X(1).
013600 03 BUSYU-CODE-2, PICTURE IS X(1).
013700 03 CODE-5, PICTURE IS X(1).
013800 03 BUSYU-CODE-3, PICTURE IS X(1).

```

次に各プログラムの説明にうつる。

BUSYU-FIFTY-ON-UNIFICATION

部首コード漢字テーブルの磁気テープ・ファイルと、代表音漢字テーブルの磁気テープファイルをまとめて一つの磁気テープ・ファイルを作成する。

部首コードのファイルには漢字以外の文字がはいっているので、処理対象が漢字のときには NORMAL-EDITING を行なうが、そうでないばあいには SPECIAL-EDITING を行なう。

KANZI-INDEX-OUTPUT-SIMPLIFIED

このプログラムはアウトプット内容が簡単でページごとの切れ目のないものができるので、実際には使用しなかった。しかし次のプログラムの内容の骨子

```
013900      03 CODE-6, PICTURE IS X(1).
014000      03 BUSYU-CODE-4, PICTURE IS X(1).
014100      02 CARRIAGE-RETURN-POSITION, PICTURE IS X(2).
015100 WORKING-STORAGE SECTION.
015200 01 PAGE-AREA.
015300      02 FILLAR, PICTURE IS X(1).
015400      02 PAGE-1, PICTURE IS X(1).
015500      02 FILLAR, PICTURE IS X(1).
015600      02 PAGE-2, PICTURE IS X(1).
015700      02 FILLAR, PICTURE IS X(1).
015800      02 PAGE-3, PICTURE IS X(1).
015900      02 FILLAR, PICTURE IS X(1).
016000      02 PAGE-4, PICTURE IS X(1).
016100 01 PAGE-COUNT-AREA.
016200      02 PAGE-1, PICTURE IS X(1).
016300      02 PAGE-2, PICTURE IS X(1).
016400      02 PAGE-3, PICTURE IS X(1).
016500      02 PAGE-4, PICTURE IS X(1).
016600 01 PAGE-COUNT, SIZE IS 4 NUMERIC.
016700 01 PAGE-FULL-COUNT, SIZE IS 2 NUMERIC CHARACTERS.
017000 CONSTANT SECTION.
017100      77 CARRIAGE-RETURN, PICTURE IS X(2), VALUE IS "%J".
017200      77 TWENTY-EIGHT, SIZE IS 2, NUMERIC, VALUE IS 28.
017300      77 HYPHEN, PICTURE IS X(2), VALUE IS "-".
017400      77 ONE, SIZE IS 4, NUMERIC, VALUE IS 0001.
017500      77 ONE-FOR-PFC, SIZE IS 2, NUMERIC, VALUE IS 01.
017600      77 N, PICTURE IS X(2), VALUE IS "N".
017700      77 P, PICTURE IS X(2), VALUE IS "P".
017800      77 APOSTROPHE-FRAME, PICTURE IS X(8), VALUE IS "  ".
017900      77 EQUAL-SIGN, PICTURE IS X(1), VALUE IS "=" .
018000      77 DOLLAR-SIGN, PICTURE IS X(1), VALUE IS "$".
018100      77 NUMBER-SIGN, PICTURE IS X(1), VALUE IS "#".
018200      77 APOSTROPHE, PICTURE IS X(1), VALUE IS "'".
018300      77 SPACE-CODE, PICTURE IS X(1), VALUE IS " ".
020100 PROCEDURE DIVISION.
020200 START. OPEN INPUT-FILE, OUTPUT-FILE. MOVE ALL ZEROS
020300 TO PAGE-COUNT. MOVE ALL ZEROS TO PAGE-FULL-COUNT.
020400 PAGE-CHANGE. MOVE SPACES TO OUTPUT-AREA. MOVE HYPHEN TO KANZI IN OUTPUT-AREA. MOVE
020500 CARRIAGE-RETURN TO CARRIAGE-RETURN-POSITION. MOVE CARRIAGE-RETURN
020600 TO KANA-1 IN OUTPUT-AREA. MOVE CARRIAGE-RETURN TO KANA-2 IN
020700 OUTPUT-AREA. MOVE CARRIAGE-RETURN TO KANA-3 IN OUTPUT-AREA.
020800 MOVE CARRIAGE-RETURN TO KANA-4 IN OUTPUT-AREA. WRITE OUTPUT-AREA.
020900 MOVE SPACES TO OUTPUT-AREA. MOVE APOSTROPHE-FRAME TO PAGE-AREA.
021000 ADD ONE, PAGE-COUNT. MOVE PAGE-COUNT TO PAGE-COUNT-AREA.
021100 MOVE PAGE-1 IN PAGE-COUNT-AREA TO PAGE-1 IN PAGE-AREA.
021200 MOVE PAGE-2 IN PAGE-COUNT-AREA TO PAGE-2 IN PAGE-AREA.
021300 MOVE PAGE-3 IN PAGE-COUNT-AREA TO PAGE-3 IN PAGE-AREA.
021400 MOVE PAGE-4 IN PAGE-COUNT-AREA TO PAGE-4 IN PAGE-AREA.
021500 MOVE P TO P-POSITION. MOVE PAGE-AREA TO GOZYUON IN OUTPUT-AREA. MOVE
021600 CARRIAGE-RETURN TO CARRIAGE-RETURN-POSITION. WRITE OUTPUT-AREA.
021700 ADD ONE-FOR-PFC. PAGE-FULL-COUNT.
021800 DATA-READ. MOVE SPACES TO INPUT-AREA. MOVE SPACES TO OUTPUT-AREA.
021900 READ INPUT-FILE. AT END GO TO END-ROUTINE.
```

となるものがふくまれている。

KANZI-INDEX-OUTPUT-NORMAL

このプログラムについては、プログラムが冗長であることを最初にことわっておかなければならない。じっさい、全体的に、いたるところで、くりかえしをいちいち展開して書いている。たとえば50音コードを漢テレコードにあらためるばあいには、DATA DIVISION で可変長の部分に OCCUR を、PROCEDURE DIVISION で PERFORM 動詞を使用し、DATA MOVING の順序を少しく入れかえればかなりまとまったプログラムになる。MOVE 動詞もいちいち書かずCORRESPONDING を用いれば1つの SENTENCE ですむ。しかし、前述のような理由で、いちおうこの形で作ってみた。簡潔な形のプロ

```
022000 DATA-EDITING. MOVE KANZI IN INPUT-AREA TO KANZI IN OUTPUT-AREA.
022100 MOVE BUSYU-CODE IN INPUT-AREA TO BUSYU-CODE IN OUTPUT-AREA.
022200 MOVE KANA-1 IN INPUT-AREA TO KANA-1 IN OUTPUT-AREA.
022300 IF KANA-2 IS EQUAL TO N, GO TO CODE-EDITING, OTHERWISE MOVE KANA-2
022400 IN INPUT-AREA TO KANA-2 IN OUTPUT-AREA.
022500 IF KANA-3 IS EQUAL TO N, GO TO CODE-EDITING, OTHERWISE MOVE KANA-3
022600 IN INPUT-AREA TO KANA-3 IN OUTPUT-AREA.
022700 IF KANA-4 IS EQUAL TO N, GO TO CODE-EDITING, OTHERWISE MOVE KANA-4
022800 IN INPUT-AREA TO KANA-4 IN OUTPUT-AREA.
022900 CODE-EDITING. MOVE KANZI-CODE-1 IN INPUT-AREA TO KANZI-CODE-1 IN
023000 CODE-AREA. IF KANZI-CODE-1 IN INPUT-AREA IS EQUAL TO NUMBER-SIGN OR
023100 APOSTROPHE OR DOLLAR-SIGN, MOVE EQUAL-SIGN TO CODE-1, OTHERWISE
023150 IF EQUAL TO SPACE-CODE, MOVE SPACE TO CODE-1, OTHERWISE
023200 MOVE APOSTROPHE TO CODE-1.
023300 MOVE KANZI-CODE-2 IN INPUT-AREA TO KANZI-CODE-2 IN CODE-AREA.
023400 IF KANZI-CODE-2 IN INPUT-AREA IS EQUAL TO NUMBER-SIGN OR APOSTROPHE
023500 OR DOLLAR-SIGN, MOVE EQUAL-SIGN TO CODE-2, OTHERWISE IF EQUAL TO SPACE-CODE,
023510 MOVE SPACE TO CODE-2, OTHERWISE MOVE
023600 APOSTROPHE TO CODE-2.
023700 MOVE BUSYU-CODE-1 IN INPUT-AREA TO BUSYU-CODE-1 IN CODE-AREA.
023800 IF BUSYU-CODE-1 IN INPUT-AREA IS EQUAL TO NUMBER-SIGN OR APOSTROPHE
023900 OR DOLLAR-SIGN, MOVE EQUAL-SIGN TO CODE-3, OTHERWISE IF EQUAL TO SPACE-CODE,
023910 MOVE SPACE TO CODE-3, OTHERWISE MOVE
024000 APOSTROPHE TO CODE-3.
024100 MOVE BUSYU-CODE-2 IN INPUT-AREA TO BUSYU-CODE-2 IN CODE-AREA.
024200 IF BUSYU-CODE-2 IN INPUT-AREA IS EQUAL TO NUMBER-SIGN OR APOSTROPHE
024300 OR DOLLAR-SIGN, MOVE EQUAL-SIGN TO CODE-4, OTHERWISE IF EQUAL TO SPACE-CODE,
024310 MOVE SPACE TO CODE-4, OTHERWISE MOVE
024400 APOSTROPHE TO CODE-4.
024500 MOVE BUSYU-CODE-3 IN INPUT-AREA TO BUSYU-CODE-3 IN CODE-AREA.
024600 IF BUSYU-CODE-3 IN INPUT-AREA IS EQUAL TO NUMBER-SIGN OR APOSTROPHE
024700 OR DOLLAR-SIGN, MOVE EQUAL-SIGN TO CODE-5, OTHERWISE IF EQUAL TO SPACE-CODE,
024710 MOVE SPACE TO CODE-5, OTHERWISE MOVE
024800 APOSTROPHE TO CODE-5.
024900 MOVE BUSYU-CODE-4 IN INPUT-AREA TO BUSYU-CODE-4 IN CODE-AREA.
025000 IF BUSYU-CODE-4 IN INPUT-AREA IS EQUAL TO NUMBER-SIGN OR APOSTROPHE
025100 OR DOLLAR-SIGN, MOVE EQUAL-SIGN TO CODE-6, OTHERWISE IF EQUAL TO SPACE-CODE,
025110 MOVE SPACE TO CODE-6, OTHERWISE MOVE
025200 APOSTROPHE TO CODE-6.
025210 MOVE CARRIAGE-RETURN TO CARRIAGE-RETURN-POSITION.
025300 ADD ONE-FOR-PFC, PAGE-FULL-COUNT.
025400 DATA-PUNCH. WRITE OUTPUT-AREA. IF PAGE-FULL-COUNT IS LESS THAN
025500 TWENTY-EIGHT GO TO DATA-READ, OTHERWISE IF EQUAL TO TWENTY-EIGHT,
025600 GO TO PAGE-END, OTHERWISE IF GREATER THAN TWENTY-EIGHT, STOP
025700 "PROGRAM-ERROR", GO TO PAGE-END, OTHERWISE NEXT SENTENCE.
025800 PAGE-END. MOVE ALL ZEROS TO PAGE-FULL-COUNT MOVE SPACES TO OUTPUT-AREA.
025900 MOVE CARRIAGE-RETURN TO CARRIAGE-RETURN-POSITION. MOVE
026000 CARRIAGE-RETURN-POSITION TO KANZI IN OUTPUT-AREA.
026100 MOVE CARRIAGE-RETURN TO P-POSITION. WRITE OUTPUT-AREA.
026200 GO TO PAGE-CHANGE.
026300 END-ROUTINE. CLOSE INPUT-FILE, OUTPUT-FILE. STOP RUN.
```


グラムに直すことはまだしていない。少しでも早く実用に供すること、という目的によって、あえてそうしたのである。

この処理プログラムの主要な目的は先述の KANZI-INDEX-OUTPUT-SIMPLIFIED と同じであるが、ページ式に印字すること、および50音の可変長データを処理して、うしろにある（印字すると目ざわりな）コードを取り除くこと、ページづけのカウンタおよびそれと、漢テレコード自体を漢テレコードに直して印字するために種々の必要な手続きをとっている（前者は PAGE-CHANGE のパラグラフ、後者は CODE-EDITING のパラグラフ）。CODE-EDITING のためになぜこのような処理をしなければならないかは、ここでは説明を省略する。国語研報告31の松本昭氏の前掲論文にあるコード表を利用すれば見当がつくと思う。

HITAC 3010 の COBOL を用いて紙テープにさん孔し、これを漢テレにかけるばあいには、一つ問題がある。それは HITAC 3010 の COBOL においては、アウトプット（WRITE 動詞）するばあい、画一的に、レコードの終わりに E/I 記号をつけてしまうことである。磁気テープに書くばあいにはそれでもよいのだが、紙テープ上にさん孔するばあいには、オフラインの、別設計の組織をもった機械にかけるのだから、これではまずい。とにかく通常の COBOL 動詞 WRITE を使うと、この E/I 記号をつけないようにすることはできず、もう一行あながあいてしまう。そこで、このプログラムではいったん紙テープにアウト・プットしたのち、もういちど別の簡単なプログラムを通して E/I を取り去り、それを漢テレにかけた。実は漢テレは E/I がついていても、実験してみるとこれを無視しパスするのだが、長時間それを行なわせることには疑問があったので、そうしたのである。別のプログラムというのはステップ数も少ないし、これに COBOL を使うことはできないわけだから、マシン・コードで作成した。

このようなプログラムを別につくることにも問題がある。COBOL の ENTER 動詞を使って own coding する手もある。大量のデータを長時間扱うばあいにはその方がよいだろう。しかしこのばあいは、いったん索引をつくったらあとそれを長期間使えるわけだから、このような方法を用いたのである。

この漢字索引作成プログラムを作成したあとでの感想をさいごに書いておく。

- ① COBOL 語はたいへんよい言語であると思うが、可変長データの扱いがさらに便利であることが望ましい。これは「巨視的」にみなくてもそうである。HITAC 3010 のような計算機の命令語の体系はむしろ参考にしてもよいのではないかと思う。もちろん、このプログラムでの可変長データの扱いはもっとも原始的で、幼稚なものであるが、それは度外視しての提言である。
- ② オフ・ラインの他機種のことと考えてほしい。(ただしこれは現状でもやむを得ないし、現状の方がいいことがあることもみとめる)。

P 0 0 0 1

[印字見本]

一	計計	いち	IL	0101
丁	計形	ちよう	QL	0103
七	計頭	しち	IV	010)
万	計建	まん	EJ	010G
丈	計見	しょう	H.	010I
三	計件	さん	IP	010.
上	計県	しょう	HI	010:
下	計伝	か	4Y	010J
不	計当	ふ	BP	010/
与	計扱	よ	W)	010T
丑	計島	うし	(R	010Z
且	計以	かつ)Q	0129
世	計位	せ	2)	012A
丘	計逸	きゅう	R4	012E
丙	計引	へい	TQ	012I
丞	計新	しょう	ZY	012N
岡	計水	りよう	HL	012X
並	計石	へい	CO	013)
中	形奥	ちゅう	8/	0341
丸	型先	かん	78	054X
丹	型歌	たん	PM	0555
主	型画	しゅ	FA	0557
乃	経県	の	%Y	070:
久	経度	きゅう	Y:	070R
之	経競	の)T	0721
乏	経逸	ぼう	US	072E
衆	経影	しょう	I3	0739