

漢字情報処理システムの開発

(漢字コード紙テープ編集処理システム)

(昭和54年5月31日 原稿受付)

情報処理施設	中	村	為	雄
安川情報システム株式会社	宮	崎	義	隆
情報処理教育センター	矢	鳴	虎	夫
情報処理教育センター	磯		泰	行

Development of Kanji Information Processing System (Editor of the paper-tape punched by Kanji-codes)

by Tameo NAKAMURA
Yoshitaka MIYAZAKI
Torao YANARU
Yasuyuki ISO

ABSTRACT

We describe Kanji Code Paper-tape Edit System recently accomplished, which added a new function to the former Kanji Code Paper-tape Processing System.

The function is as follows.

At first, users make the source Kanji code paper-tape, then the system catalogues the source code with the line number in the magnetic tape.

When users certify the list printed out from the magnetic tape and find some mistakes, users make the edit paper tape to correct the source code by merely adding the corresponding line number.

Moreover this system contains DUMP function to proceed easily.

1. はじめに

九州工業大学情報処理施設では昭和52年度より漢字情報処理システムの開発を行なっている。52年度末に基本システムが完成し、53年度に漢字タブレットを購入し漢字コード紙テープ処理が行なえるようになった。今回は漢字コード紙テープの編集を行なうシステムが完成したので報告する。

2. 漢字コード紙テープ編集処理

漢字コード紙テープ処理システムを使用することにより、利用者は複雑な書式も簡単に処理できるようになった。

しかし、漢字タブレット装置を用いて漢字コードを紙テープに出力する場合に、誤った漢字を指定してそう入削除を行わなければならない時には紙テープのコピー停止、修正等をくり返し行なうのはとても不便である。

そこで、JIS漢字コードとコントロールに分離した漢字データを磁気テープに入れ、修正用漢字コード紙テープを用いて、編集を行なうようにした。

2.1. 処理方式の設計

編集処理を行なう上で必要である機能として次のものが考えられる。

- (1) 文字単位のそう入、削除、変更ができる。
- (2) 行単位のそう入、削除、変更ができる。

(3) 漢字コード紙テープ・コントロールの変更ができる。

以上の処理を行なうにはランダム処理の可能な磁気ディスク装置が適当と考えられるがシステム機器構成上、磁気ディスク装置に作業エリアを取ることができないので磁気テープ装置を使用する。

その結果シーケンシャル処理となったが、それを捕う意味で各行の漢字データに行番号をつけ、プリンタ/プロットに出力するようにした。利用者は、その出力を参照しながら修正漢字コード紙テープを作成することになる。

2.2. 編集キー・コマンドの種類

編集のために必要なキー・コマンドを次のように定めた。

- (1) 編集用紙テープの始まりを示す <編始>
- (2) " 終りを示す <編終>
- (3) 編集コマンドの句切りを示す <編、>
- (4) 編集一行分の終りを示す <編: >
- (5) 漢字データ一行分の削除を示す <削行>
- (6) 漢字データ文字の削除を示す <削字>
- (7) 漢字データ一行分のそう入を示す <挿行>
- (8) 漢字データ文字のそう入を示す <挿字>
- (9) 文字列のサーチ <捜字>
- (10) 文字列の変更 <変更>

2.3. 修正漢字コード紙テープの形式

修正漢字コード紙テープは次に示す形式で作られていなければならない。

- (1) 紙テープの最初は <編始> がなければならない。
- (2) 紙テープの最後は <編終> がなければならない。
- (3) 編集一行分の最初には <行番号> が、最後には <編: > がなければならない。
- (4) 各コマンドは <編、> で句切られる。
- (5) コマンドでデータの必要なものは直後にデータを続ける。データが必要ないものはキーのみで有効である。

2.4. 処理の流れ

処理の流れは次のようになる。ただし磁気テープ装置の01（以後MT01と言う）には修正する漢字データファイルがセットされ、磁気テープ装置02（以後MT02と言う）には新しいテープがセットされ、紙テープ読取装置（以後PTRと言う）には修正漢字コード紙テープがセットされているものとする。

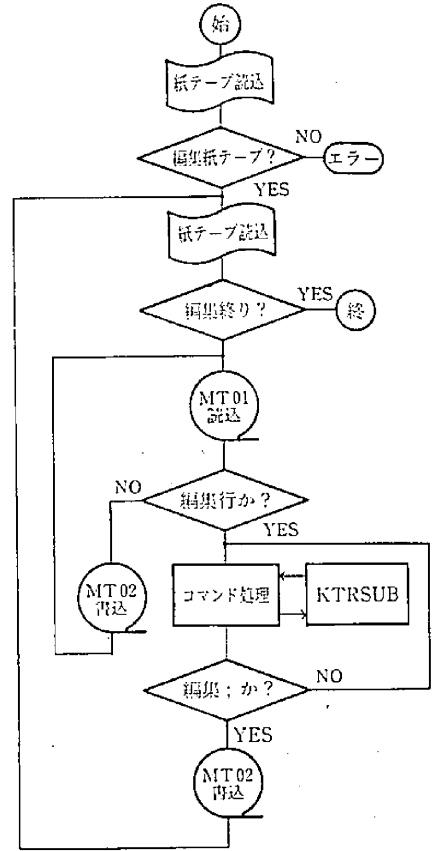


図-1 編集処理の流れ

1ブロック = 256語		COMAND	
0	COMAND (25)	0	LINE NO.
25	MODE (67)	1	N MAX
92	TAB POSITION (15)	2	SET N
107	TAB SET (15)	3	ZEN BAI
122	GAJI (67)	4	TATE BAI
	0 = 'FFFF' (16)	5	YOKO BAI
189	DATA (67)	6	GYO KAN
	0 = '2121' (16)	7	JJ KAN
		8	ZEN KAN
		9	
		10	
		11	CRLF
		12	END
		13	GAJI N
		14	TAB N
		15	LAST CHAR
		16	
		17	
		18	
		19	
		20	
		21	
		22	
		23	
		24	

図-2 漢字データ・ブロック構成

- (1) PTRより紙テープを読み込み修正用テープであることを確認する。
- (2) 修正すべき行までMT 01のデータをMT 02に移す。
- (3) 各キーコマンドの処理を行ない漢字データの編集をする。その時二行分の漢字データの長さを越えた場合はその分のデータを無視する。
- (4) 〈編;〉を見つけると、編集した漢字データをMT 02に移す。
- (5) 以下〈編終〉まで処理を繰返す。

この処理を流れ図にしたものが図-1である。この処理を行なうサブルーチンは KDEDIT である。

また、漢字データのブロック構成を図-2に示す。

2.5. ユーティリティ・ルーチン

本システムをサポートするためのユーティリティ・ルーチンは漢字コード紙テープ・ダンプ・ルーチン、ドット・パターン作成ルーチン、漢字パターン登録ルーチンである。

2.5.1. 漢字コード紙テープ・ダンプ・ルーチン

本ルーチンは、漢字コード・データが正しく紙テープに穿孔されているかまたは編集処理のために、コントロールと漢字を拡大処理などを行わずに、そのままプリンタ/プロッタに出力する。これにより利用者は漢字データの誤りが見つかりやすくなった。

< 01 20 > 26
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

< 01 30 > 27
a b c d e f g h i j k l m n o p q r s t u v w x y z あ

< 01 40 > 32
いうえお かきくけこ さしすせそ たちつと なにぬ

< 01 50 > 15
へほまみむめも やゆよ

< 01 60 > 10
らりるれろ わるゑをん

< 01 70 > 27
かぎくけこ さしすせそ たちつと ばびぶべ ぽびびよ へほがぎ

< 01 80 > 23
クゲゴザジスゼソダヂツテド バビブベボバビブベボ

< 02 90 > 01
<

2 1 3 0

< 03 00 > 00

図-3 RTDUMP出力例

編集・修正内容

- (1) <20> に“あいうえお”を挿入。
- (2) <30> より“MNO”を削除。
- (3) <45> を挿入。
- (4) <50> に“サシスセソ”を挿入。
- (5) <60> に“大中小”を挿入。

```

< 00 10 >          桁 10
1 2 3 4 5 6 7 8 9 0
< 00 20 >          桁 10
A B C D E F G H I J
< 00 30 >          桁 10
K L M N O P Q R S T
< 00 40 >          桁 05
U V W X Y Z
< 00 50 >          桁 04
a b c e
< 00 60 >          桁 10
k l m n o p q r s t
< 00 70 >          桁 07
モ ヴ ウ ヲ ヱ ズ
< 00 80 >          桁 10
1 1 1 1 1 1 1 1 1 1
< 00 90 >          桁 10
一 二 三 四 五 六 七 八 九 十
< 01 00 >          桁 05
昭和年月日
< 01 10 >          桁 00
    
```

編集・修正前

```

< 00 10 >          桁 10
1 2 3 4 5 6 7 8 9 0
< 00 20 >          桁 15
A B C あ い う え お D E F G H I J
< 00 30 >          桁 07
K L P Q R S T
< 00 40 >          桁 05
U V W X Y Z
< 00 45 >          桁 05
ア イ ウ エ オ
< 00 45 >          桁 05
カ キ ク ケ コ
< 00 50 >          桁 09
サ シ ス セ ソ a b c e
< 00 60 >          桁 12
k 大 中 小 m n o p q r s t
< 00 70 >          桁 07
モ ヴ ウ ヲ ヱ ズ
< 00 80 >          桁 10
1 1 1 1 1 1 1 1 1 1
< 00 90 >          桁 10
一 二 三 四 五 六 七 八 九 十
< 01 00 >          桁 05
昭和年月日
< 01 10 >          桁 00
    
```

編集・修正後

図-4 漢字コード紙テープ編集・修正例

2.5.2. ドット・パターン作成ルーチン

本ルーチンは OKITAC-4500 の FORTRAN コンパイラが、実行時に参照する時しか FORMAT 文のエラーチェックを行なわないことを利用して、FORMAT 文を用いたドット・パターンの作成を行なうものである。

ドット・パターン用に ORMAT 文の構成

- (1) ビットを ON にする <S>
- (2) ビットを OFF にする <R>
- (3) ビット単位で SKIP する <X>
- (4) 以降全てのビットを OFF にする <I>

記述方法は nSm と書く。n,m は省略が可能で、省略時

の値は 1 である。その他の点については FORTRAN 文法に従っている。カッコを使用した場合最大 5 コまで、一つの FORMAT 文では一行文 (2112 ビット) の指定しかできず、それ以降の記述は無視される。

例 100 FORMAT (S, 100 (R9,S), 1)

実際にはあらかじめ、使用する FORMAT 文の文番号を指定しておき (FMTSET), その指定順序に従って決まる番号をパンチカードにより指定することでドット・パターンが作成される。(DPTRCV)

例 CALL FMTSET (\$100, \$200, \$300)
CALL DPTRCV (IR, S99)

IR: ドット・パターン・セット・エリア

S 99: カード終了時の飛び先文番号

上記のように指定することにより文番号100, 200, 300がそれぞれ1, 2, 3と決まり, 1Rに8行分(2112×8ビット)のドット・パターンがセットされる。

次に上記のパンチ・カードの指定方法について述べる。

パンチ・カードによる FORMAT 文の指定方法

基本的には指定された文番号に対応した数字をコマで句切って指定すればよい。一枚のカードに対する指定数に制限はなく, 任意のブランクを置くことができる。また同じ番号の指定のくりかえしには〈*〉を使ってまとめることができる。すなわち, 次のように書かれたaとbは全く同じである。

a 1, 2, 2, 3, 3, 3, 3, 1, 2, 2, 2

b 1, 2*2, 4*3, 1, 3*2

また, 同一内容のカードが続く時はカードの先頭にN nと書くことにより, 同一カードをn回, 参照することができる。

例 N3, 1, 2

流れ図を図-5に, 出力例を図-6に示す。

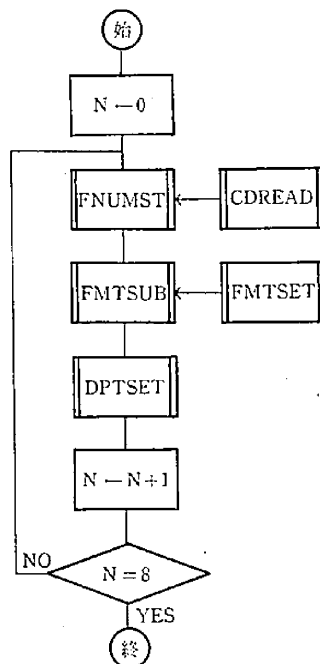


図-5 ドット・パターン作成ルーチンの流れ

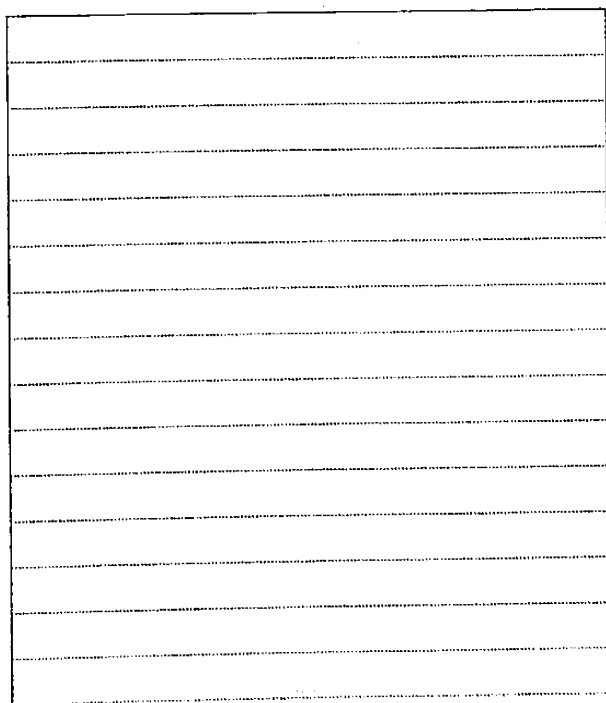
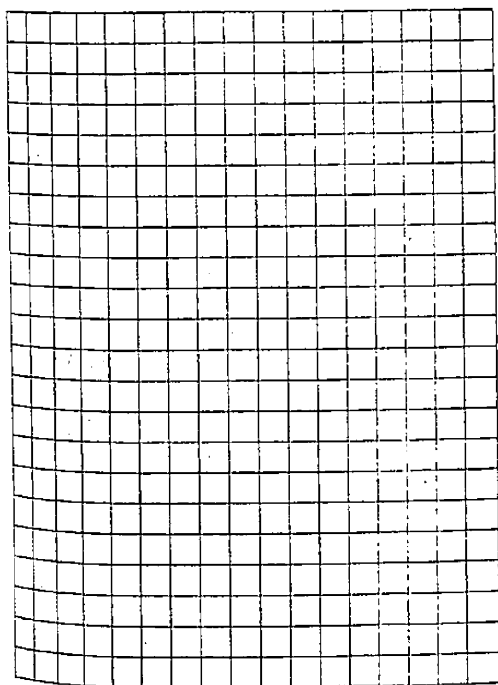


図-6 ドット・パターン出力例 (実物大)

2.5.3. 漢字パターン・データ登録ルーチン

漢字処理を行なう上で漢字パターンの修正、追加等が必要になってくる。これは既に基本システムにおいて完成していたが、6000字余もある漢字を扱うため、処理速度が問題となっていた。そこで FORTRAN で書かれていたプログラムをアセンブラに書き換えた。

本ルーチンは現在ある漢字パターンの読み出し、修正、新しい漢字パターンの追加の3つの部分よりなる。このルーチンの作成により処理速度が約4倍となった。

3. おわりに

今回作成したシステムにより、漢字タブレット装置を用いたシステムはほぼ実結したと考えられる。システム機器の制約から編集についてはシーケンシャル処理にとどまったが、磁気ディスク装置に余裕があれば、ランダム・アクセス処理が可能となり、より使い易いシステムになるであろう。

これからの展望としては、漢字処理用語の開発が考えられる。また、漢字処理で一番問題となるのは数千字にも及ぶ漢字の入力方法、誤りの訂正等である。それには本システムのようにオフライン方式の漢字入力装置を持ったものと、解像度の高いグラフィック・ディスプレイ装置を用い、対話形式でスクリーン・エディタによる修正機能を持ったシステムの併用が必要と思われる。

漢字情報処理はまだ比較的新しい分野に属するもので、漢字コードについては情報交換用漢字符号系が発表され統一の動きがみられるが、機能、漢字パターンの形式、漢字コードの入出力方法等については、各システムによりまちまちであり、互換性は全くといっていい程ない。早急なる規格の統一を望むとともに、より良い漢字情報処理システムの開発を行ないたい。

参考文献

- 1) 中村, 安永, 竹中 "漢字情報処理システムの開発 I" 九工大研究報告 (工学) No. 36昭和53年3月.
- 2) 中村, 竹中 "プリンタ/ブロッック・コントロール・システムの開発 (OKITAC-4300)" 九工大研究報告 (工学) No. 37昭和53年9月.
- 3) 中村 "プリンタ/ブロッック・コントロール・システムの開発 (OKITAC-4500)" 九工大研究報告 (工学) No. 37昭和53年9月.
- 4) 中村, 安永, 竹中 "漢字情報処理システムの開発" 電気四学会九州支部連合大会講演論文集第30回昭和52年10月.
- 5) 中村・竹中・宮崎 "漢字情報処理システムの開発 (基本システム)" 電気四学会九州支部連合大会講演論文集第31回昭和53年10月.

- 6) 宮崎, 中村 "漢字情報処理システムの開発 (応用システム)" 電気四学会九州支部連合大会講演論文集第31回昭和53年10月.
- 7) 中村, 山口 "磁気テープ・コントロール・サブルーチン・コントロール・システムの開発" 電気四学会九州支部連合大会講演論文集第31回昭和53年10月.
- 8) 日本工業技術センター "漢字情報処理資料集" 昭和54年1月.
- 9) 日本工業標準調査会 "情報交換用漢字符号系" JIS C6226-1978 昭和53年1月.
- 10) 日本電気漢字タブレット装置 (C-5122J) 取扱説明書" 昭和53年1月.
- 11) XEROX "VERSATEC OPERATION AND MAINTENANCE MANUAL D1200A Printer/Plotter" August 1976.
- 12) アルプス・コンピュータ・エンジニアリング "漢字情報処理装置 CT-2000 概説書".
- 13) 沖電気株式会社 "OKITAC-4500 DOS-45 説明書".
- 14) 沖電気株式会社 "OKITAC-4500 DOS-45 MACRO-ASSEMBLER 説明書".

付 録

このシステムのために開発したサブルーチンの任様を以下に説明する。

1. TPRINT

(I) 機能

漢字コード紙テープをテープ・リーダー (PTR) より読み込み、プリンタ/ブロッックに漢字を出力する。

(II) FORTRAN CALL 形式

CALL TPRINT (N)

(III) パラメータ

N: 1行の読み込みコード数 ($1 \leq N \leq 66$)

ただし、改行コードを読み込むとNに満たない場合 SPACE をつめてNとする。

(IV) 補助サブルーチン

KTREAD, KWIDES, PPKRCV, PPDPLT, PPDCTL

(V) 注意

本ルーチンを CALL する前に PPOPEN サブルーチンが CALL されていなければならない。

終のコードを読み込むとこの RETURN する。

2. KTREAD

(I) 機能

漢字コード紙テープ (NEC C-5122J) を PTR より読み込み、JIS 漢字コードに変換する。

(II) FORTRAN CALL 形式

CALL KTREAD (ID, N, NN, \$E)

CALL KTREAD (ID, N, NN, \$E, IT, IY, IG,
IJ, IZ)

(III) パラメータ

ID: 読み込みエリア (整数一次元配列名)

N: 読み込みコード数

NN: RETRUN 時の実読み込みコード数

\$E: 読み込み中にエラー、テープ・エンド、終
のコードのいずれかを検出した時の飛び先
文番号

IT: 縦方向の拡大倍数

IY: 横方向の拡大倍数

IG: 行間 (8ビット単位, 約 1mm)

IJ: 字間

IZ: 前間

(1) 改行コードを検出すると、残り指定エリアにブラ
ンク・コードがつめられてリターンする。

(2) TAB コードを検出すると、TABSET ルーチンに
より指定された TAB 位置までスペース・コード
がつめられ次の読み込みをつづける。

TAB 位置が指定最大値を越えると、改行と同
様の処理を行なう。

(3) 改 TAB-ON のコードを検出すると一行が最大
値をこえると次の行の最初には、自動的にTABコ
ードがそう入される。改TAB-OFFコードは、その
指定の解除である。この機能は論文等を指定最大
値を使ってベタ打ちとし行の最初をアケル場合に
使用する。

(4) 縦あるいは横のコードを検出すると、それ以降は
そのモードにセットされる。

(5) 全倍、縦倍、横倍、行間、字間、前間のコードも
(4)と同様である。

(6) 終コードを検出すると \$E の文番号へ JUMP す
る。

(IV) 補助サブルーチン

TASET, MBTRED, KCONVB, KTRSUB

(V) 注意

本ルーチンを CALL する以前に TABSET がされて
ない時は、TAB 位置は 1, 6, 11, 16, 21, 26, 31, 36,
41, 46, 51, 56, 61, 66にセットされている。

パラメータ項における5つのコードは、行の最初にあ
ることが望ましい。

3. KCONVB

(I) 機能

漢字タブレット・コードを JIS 漢字コードに変換す
る。

(II) FORTRAN CALL 形式

CALL KCONVB (ID, N, MODE)

(III) パラメータ

ID: 整数型一次元配列名, 漢字コード

N: 変換コード数

MODE: 0 = 縦字, 1 = 横字

(IV) 補助サブルーチン

なし

(V) 注意

変換するデータが、漢字タブレット・コード以外の場
合、JIS 漢字コードのブランク ("2121"₍₁₆₎) に変換され
る。

4. MBTRED

(I) 機能

紙テープより一語 (2バイト) バイナリ形式で読み込
む。

(II) FORTRAN CALL 形式

CALL MBTRED (I, \$N)

(III) パラメータ

I: 読み込みエリア (整数数)

\$N: 読み込み中にエラーが発生した時の飛び先
文番号。

(IV) 補助サブルーチン

なし

(V) 注意

本ルーチンは、先読みを行なうため、紙テープの終わ
りは、少なくとも40バイト分の FEED が必要である。
コードの OO₍₁₅₎ 及び FF₍₁₆₎ は無視する

5. KWDRCV

(I) 機能

指定されたエリアのデータをパラメータにより変換
し、パターン of 整数倍拡大を行なう。

(II) FORTRAN CALL 形式

CALL KWDRCV (IS, IR, N, \$E, IT IY, IG, IJ, IZ)

(III) パラメータ

IS : 変換したデータがセットされるエリア (1056)

IR : 変換するデータが入っているエリア (3168)

N : 漢字データ数 ($1 \leq N \leq 66$)

\$E : 変換終了時の飛び先文番号

IT, IY, IG, IJ, JZ : KTREAD サブルーチン参照

(IV) 補助サブルーチン

KWIDES

(V) 注意

本ルーチンは変換したデータを1056語ずつ受け渡すため、ISを除くパラメータを変えることなく、繰り返しCALLしなければならない。

使用例

```
10  CONTINUE
    CALL KWDRVC (IS, IR, N, $99, IT, IY,
                IG, IJ, IZ)
    CALL PPDPLT (IS, 2112)
    GO TO 10
99  CONTINUE
```

IR (3168) にセットされた漢字データを変換パラメータにより変換して、プリンタ/プロッタに出力する例。

6. KWIDES

(I) 機能

指定されたエリアのデータの横の拡大を行なう。

(II) FORTRAN CALL 形式

CALL KWIDES (ID, N)

(III) パラメータ

ID : 拡大するデータが入っている配列名

N : 倍率 $1 \leq 132$

(IV) 補助サブルーチン

なし

(V) 注意

拡大によりID (132/N-1) からID (132) までのデータは失なわれる。

7. TABSET

(I) 機能

漢字コードの読み込み位置を設定する。

(II) FORTRAN CALL 形式

CALL TABSET (I1, I2,IN)

(III) パラメータ

IN : TAB セットの位置 $1 \leq IN \leq 66$

個数は最大14個

(IV) 補助サブルーチン

TABSUB

(V) 注意

本ルーチンはKTREAD サブルーチンを使用する前に漢字コードの読み込み位置の制御を行なう。

8. KTDPRT

(I) 機能

漢字コード・テープ及びドット・パターン処理を行なう。マス目等を漢字出力時に付ける。

(II) FORTRAN CALL 形式

CALL KTDPRT (N)

(III) パラメータ

N : 整定数または整変数、漢字コード一行読み込み数

(IV) 補助サブルーチン

PPDPLT, KTREAD, PPKRCV, DPTRCV

(V) 注意

本ルーチンをCALLする前にPPOPENがCALLされてなければならない。またFMTSET サブルーチンによりFORMAT番号が指定されてなければならない。

9. DPTRCV

(I) 機能

カードに指定されたフォーマットの番号によりドット・パターンをIRにセットする。

(II) FORTRAN CALL 形式

CALL DPTRCV (IR, \$E)

(III) パラメータ

IR : 整数型配列名 (1056) ドット・パターンをセットする配列名

\$E : サブルーチン終了時の飛び先文番号

(IV) 補助ルーチン

FNUMST, FMTSUB, DPTSET

(V) 注意

使用例

```
10  CONTINUE
    CALL DPTRCV (IR, $99)
    CALL PPDPLT (IR, 2112)
    GO TO 10
```


99 CONTINUE

10. FMTSUB

(I) 機能

指定された番号により、ドット・パターンフォーマット・アドレスをセットする。

(II) FORTRAN CALL 形式

CALL FMTSUB (N, \$F)

(III) パラメータ

N: FMTSET により作られたフォーマットの番号 (1~n)

\$F: フォーマット文の文番号

(IV) 補助サブルーチン

FMTSET

(V) 注意

使用例

CALL FMTSET (\$ 10, \$ 20, \$ 30, \$ 40, \$ 50)

CALL FMTSUB (3, \$F)

とすると、SF = \$ 30となる。

11. FNUMST

(I) 機能

カードに指定されたフォーマット番号をセットする。

(II) FORTRAN CALL 形式

CALL FNUMST (IS, SE)

(III) パラメータ

IS: フォーマット番号セット・エリア

SE: カード・エンドのときの飛び先文番号

(IV) 補助ルーチン

CDREAD

(V) 注意

カードの形式。

例 N₂, 1, 2, 3, 2*5, 3

と書かれたカードは本ルーチンから呼ばれるたびに、

IS に次のようにセットされる。

1, 2, 3, 5, 5, 3, 1, 2, 3, 5, 5, 3

12. FMTSET

(I) 機能

フォーマット文のアドレスをセットする。

(II) FORTRAN CALL 形式

CALL FMTSET (\$1, \$2, ..., \$N)

(III) パラメータ

\$1, \$2, ..., \$N: FORMAT 文の文番号

個数は10個まで

(IV) 補助ルーチン

なし。

(V) 注意

本ルーチンは、FMTSUB に対する補助ルーチンである。FMTSUB を使用したルーチンを呼ぶ前に、本ルーチンを CALL しなければならない。

13. DPTSET

(I) 機能

フォーマット文に従ってドット・パターンを指定エリアにセットする。

(II) FORTRAN CALL 形式

CALL DPTSET (ID, SF)

(III) パラメータ

ID: 整数型配列名 ドット・パターンをセットする配列名 エリア ≥ 132

SF: FORMAT 文のアドレス

(IV) 補助ルーチン

なし。

(V) 注意

フォーマットの指定。

S = セット・ビット

R = リセット・ビット

X = スキップ

I = 改行

使用例

CALL DPTSET (ID, \$100)

100 FORMAT (S, 10 (R8, S), 5X, S, 1)

14. KTDUMP

(I) 機能

漢字コード紙テープをPTRより読み込み内容をコントロール・コード付でプリンタ/ブロックに出力する。

(II) FORTRAN CALL 形式

CALL KTDUMP (N)

(III) パラメータ

N: 1行の読み込み漢字数 (1 ≤ N ≤ 66)

(IV) 補助サブルーチン

KTREAD, PPKRCV, PPDPLT, PPDCTL

(V) 注意

本ルーチンを CALL する前に PPOPEN が CALL されていないとエラーになる。

漢字コードの各コントロールも 1 つの字として扱われる。