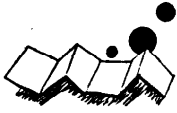


## 解説

## UNIX の文書処理機能†



長谷部 紀元††

## 1. はじめに

## 1.1 UNIX の文書処理機能

文書処理機能といった場合、一般的には自然言語処理に近い分野までを含めて考える場合もあるが、本論文では UNIX オペレーティング・システム上でよく使用されている、高品質の文書作成機能を取りあげる。UNIX の文書処理機能は本文ばかりでなく、図・表・式・参考文献・索引などを一貫して処理することが可能で、その品質は出版目的の使用に耐える。学術雑誌の投稿論文の版下作成に使用された<sup>1),2)</sup>ばかりでなく、雑誌の定常的な刊行において<sup>3)</sup>、また多数の専門書<sup>4)</sup>の出版において使用されている。

UNIX オペレーティング・システムの応用分野の中で、文書処理の占める重要さは特筆すべきものがある。UNIX の開発の初期の段階で研究費を獲得する決め手となったのは、ベル研究所における特許文書処理への応用であった<sup>5)</sup>。またこの機能は UNIX の初期からマニュアルの作成にも利用され、UNIX の普及に寄与した。

UNIX の文書処理機能全般については既にいくつもの解説<sup>6),7)</sup>があり、他の文書処理システムとの綿密な比較<sup>8)</sup>も行われている。日本においても **troff** と類似のシステムに対しての解説<sup>9),10)</sup>や整形プログラムの基礎となるアルゴリズムに関する解説<sup>10)</sup>がある。本稿では実例に即した形で、UNIX の高度の文書処理機能の実現方法を具体的にみていくことにする。本稿の版下は題字などごく一部を除いて、**troff** を日本語化した **JTROFF**<sup>11),12)</sup> を使用して作成したものであり、全体が実例にもなっている。引用する例には日本語が含まれているが、原稿の記述法は本来の **troff** と同じである。なお UNIX に関する一般的な文献については本号の他の論文を見られたい。

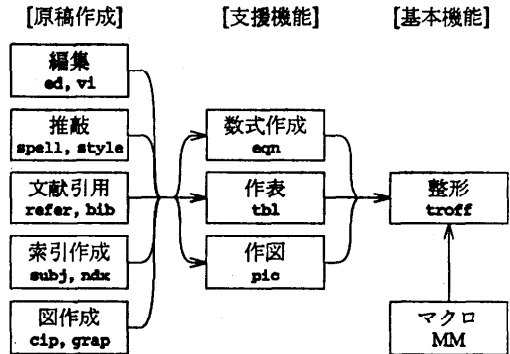


図-1 UNIX の文書処理機能

## 1.2 UNIX の文書処理機能の性格

UNIX の文書処理機能は、図-1 に示すようなプログラム群から構成されている。その中心となっているのが整形プログラム **troff**<sup>13)</sup> である。これと同一の原稿を整形できるものに **nroff** があって、こちらはタイプライタのような均等な字幅を持つものを対象としている。UNIX の文書処理機能の特徴は高品質の出力が可能などところにあると考えられるので、以後の解説は専ら **troff** について行う。UNIX の文書処理機能全体の働きはその品質のレベルと総合性からみて、単なる整形と言うよりは組み版と呼んだ方が適当であるとされる。以後の記述を簡潔にするためにこの機能を構成するプログラム群全体を **TROFF** と総称することにする。

**troff** の名称は "typesetter runoff" に由来している。当初は字母がネガ・フィルムである、光学式の写真植字機を使用していた。この写植機は日本でも UNIX の影響によって使用されたこともあるが<sup>14)</sup>、**troff** の普及が遅かったのはこの出力装置が特殊なものであったためであろう。その後 CRT での文字表示を利用する電子式写真植字機の普及ともなって、設計の見直しが行われ、任意の植字機への適用が可能となった<sup>15)</sup>。この結果ドット・プリンタへの適用可能性も大幅に強化された。現在研究環境で **troff** はレーザ・ビーム・

† Document Preparation Facilities of UNIX by Kigen HASEBE (University of Library and Information Science).

†† 図書館情報大学

プリンタを出力装置として盛んに使用されている。JTROFF の出力装置もレーザ・ビーム・プリンタである。

### 1.3 他の文書処理システム

TROFF と同様の目的を持った文書処理システムに T<sub>E</sub>X<sup>16),17)</sup> と SCRIBE<sup>18)</sup> がある。T<sub>E</sub>X は数式の組み版に特に優れ、troff と同様の高品質出力を行う。SCRIBE は使用者インタフェースの面で高い評価を受けている。T<sub>E</sub>X はこの点については特によいものではない。これを補って SCRIBE に近い使用者インタフェースを作り出すマクロ・パッケージとして、L<sup>A</sup>T<sub>E</sub>X<sup>19)</sup> が開発されている。

## 2. 基本機能

ここでは troff の基本的な整形機能について述べる。

### 2.1 整形指令

troff の処理の方式は事前に用意済みの原稿を、パッチ的に整形・出力するものである。原稿中で整形に関わる指令を、印刷しようとしている本文に加えて与える必要がある。ワードプロセッサと呼ばれるものが対話的に全ての処理を行うのとは大きく異なる。

指令は形態上、本文とは別の行に書かれるリクエストと、本文の中に埋め込んだ形で書かれるエスケープ・シーケンスの 2 種が存在する。リクエストは行の先頭に特別な文字 (普通には ‘.’) が存在することで本文と区別される。これは 1 ないし 2 文字のリクエスト名で識別される。エスケープ・シーケンスは特別な文字 (通常は ‘\’) で先導される。こちらは 1 文字の機能名で区別され、形態は多様である。

### 2.2 単純な例

troff の基本的な機能を理解するために文献 9) と同じ本文を使用して、単純な例を見ることにする。図-2 に出力例を、図-3 にその原稿を示す。ここでは出力装置が植字機であるために特に必要となる機能に焦点を当てて解説する。なお、本稿で使用する例の左端にあるイタリックの数字は説明のための行番号であって、実際の原稿や出力には現れない。説明中でイタリックの数字で引用するのはこの番号である。

整形プログラムの最も基本的な機能は、入力原稿中のランダムな長さの行を追い込んで (fill), 指定された長さの出力行を作り出すことである。しかし、追込みを行うのが不適当な場合もある。それで追込みの停止 (1), 再開 (10) のためのリクエストが存在する。

LOCK WILLOW  
SEPT 19TH

Dear Daddy,

Something has happened and I need your advice. I need it from you, and from nobody else in the world. Wouldn't it be possible for me to see you? It's much easier to talk than to write; and I'm afraid your secretary might open the letter.

Judy

P.S. I'm very unhappy.

図-2 troff の出力例

```

1 .nf
2 .in 2i
3 .ps 7p
4 .vs 9p
5 LOCK WILLOW
6 SEPT 19TH
7 .in 0
8 .ps 10p
9 .vs 12p
10 .fi
11 .sp 0.5v
12 .ft I
13 Dear Daddy,
14 .ft R
15 .ti 0.3i
16 Something has happened and I need your
17 advice. I need it from \fiyou, \fR
18 and from nobody else in the world.
19 Wouldn't it be possible for me to see you?
20 It's much easier to talk than to write;
21 and I'm afraid your secretary might open
22 the letter.
23 .sp 0.3v
24 .ti 2i
25 \fiJudy\fR
26 .sp 0.3v
27 .ti 0.3i
28 \fB\sP.S.\s10 \fII'm very unhappy\fR.
```

図-3 troff の原稿例

パラグラフの先頭などでは頭下げ (indent) を行う。これには持続的な効果を持つもの (2) と臨時的なもの (15) とがある。植字機では紙面上の位置の指定について自由度が大きいため、各種の指定方法が必要となる。ここでの単位 ‘i’ は inch である。印刷文字のサイズは 10p がよく使用されるが (8), 印刷効果のために自由に変更を行うことができる (3)<sup>†1</sup>。単位 ‘p’ は point である。行送りの量は文字サイズとは独立であるので、適切な値を別に指定する (4, 9)。行間取る余白の量も自由な指定ができる (11)。ここで使用している単位 ‘v’ は行送りの単位量として先に指定したもので (9), 垂直方向における基本単位である。水平方向について

†1 本稿の実際の文字のサイズは、印刷の都合で原稿上の指定と同じではない。

も文字サイズに比例する基本単位として、全角に相当する 'm' などが用意されている。

植字機の利点は豊富なフォントの使い分けによって、読者に与える印象を強め、伝達する情報の量を増加できることにある (12, 14)。この 'I' と 'R', それに 28 の 'B' はそれぞれ *Italic*, *Roman*, それに **bold** のフォントを指定している。フォントと文字サイズの指定は、いま見たような独立した行の形式をしたリクエストの他、テキスト中のエスケープ・シーケンスでも可能である。28 の '\f' と '\s' がその例である。

以上のような機能によって組み版出力の情報伝達能力に関する品質が維持されるのであるが、整形言語に要求される記述能力はそれだけ複雑になる。

### 2.3 マクロ機能

#### 2.3.1 物理的対象と論理的対象

上でみたとおり、**troff** の整形の指示に関わる機能は物理的なレベルに設定されている。使用者が原稿を作成するとき意識する論理的な構成要素、つまり章節・パラグラフ・脚注などの概念は提供されていない。したがって、**troff** の物理的レベルの指令を一般の使用者がそのまま使用することは、不適当であることが多い。論理的な対象の実現はマクロによって行うことが前提とされている。さらに、より高度の論理的対象、つまり数式や表などは、**TROFF** に用意される独立の支援プログラムに任されている。

マクロについてここで詳しく説明することはできない。簡単にいってしまうと、マクロはテキストとリクエストの集まりとして、整形処理の過程で動的に定義される。定義されたマクロはリクエストとほとんど同じ方法で使用できる。マクロの定義方法はアセンブラ的で、使い易いとは決して言えない。しかし論理的対象を実現するための道具は強力に用意されている。

中でも面白くまた他の整形プログラムに例を見ないのが、コンピュータの割り込みに類似した **trap**<sup>20)</sup> である。これは脚注や図版などの論理的対象を原稿上の位置とは非同期に、ページ境界などの場所に配置するために使用される。普通ならば、この種の論理的対象をそのまま整形プログラムの機能として組み込むところであろう。**troff** は **trap** という低レベルの機能によって、使用者レベルでの一般的な処理を可能にしている。

整形処理における物理的対象を記述するために開発された、強力によく整理された言語に **PostScript**<sup>21)</sup> がある。この言語の処理系は植字機の中に組み込まれ

```

1 .TL
2 UNIXの文書処理機能
3 .AU "長谷部紀元"
4 .MT 4
5 .2C
6 .FS †
7 Document Preparation ...
8 .FE
9 .H 1 "はじめに"
10 .H 2 "UNIXの文書処理機能"
11 文書処理機能と...
12 .P
    ...
13 .H 2 "UNIXの文書処理機能の性格"
    ...

```

図-4 MMの使用例

て、各種の物理的対象の機械独立化を実現する。整形プログラムの機械独立性に大きく寄与するものと期待されている。

#### 2.3.2 マクロ・パッケージ

論理的対象を体系的に実現するには、マクロのパッケージが必要である。ここではそのなかで充実度の高い **Memorandum Macro**<sup>22)</sup> (**MM**) を紹介する。**MM** には使用者が組み版の形式を制御するためのパラメタが多数ある。パラメタ値のセットを定義するファイルを用意しておけば、誰もが容易に目標の出版物の組み版形式に合わせた出力を得ることができる。これに合わない特殊な表現を必要とする場合には、パッケージ中のマクロの手直しが必要になる。

図-4 に本稿のような論文を組み版するための、**MM** マクロのおよその使用例を示す。少数のマクロで組み版の指示を済ませることができる。最初は表題部で、1-3 はタイトルと著者名、4 は文書の形式 (ここでは論文) である。5 は二段組みを指示している。6-8 は脚注、9 以降が本文である。章節はそのレベルを指定するだけで、階層的な番号が自動的に付けられる。パラグラフの開始は 12 のように指定する。

## 3. 支援機能

ここでは数式・表・図形などの論理的対象を実現している支援プログラム、それに原稿作成を支援している参考文献のためのプログラムを解説する。

### 3.1 数式の組み版

#### 3.1.1 数式の記述

数式は 2 次元的な構造を持っており、しかも構成要素が不均等なサイズであるために、その組み版は専門家の間でも困難なものとされている。**eqn** は数式に対して柔軟な文法を設定するという巧妙な方法により、素人が容易に組み版できるようにすることに成功し

た。プログラミングの手段としてパーザ・ジェネレータである **yacc** を使用していることは有名である。アメリカの学会で通常のタイピストが使用して、学会誌の版下を作成するのに大きな効果を挙げたことが報告されている<sup>3)</sup>。図-5 に **eqn** による簡単な数式の出力の例を示す。図-6 はその原稿である。数式を頭から(英語式に)読み下すままに入力すればよく、使い勝手がよい。機能はここに示した分数・積分のほか、総和・根号・マトリックスなどおよそその数式の構成方法や、数学位の記号の指定方法が用意されている。

### 3.1.2 eqn の処理

数式の開始と終了はマクロの形式をした '**.EQ**' と '**.EN**' によって指示される。**eqn** はその間に挟まれた数式を **troff** 原稿の形式に変換する。一方、マクロの形をした '**.EQ**' などの方は、他の部分と同じくそのまま出力する。**eqn** は数式の番号をどう付けるか、ページの上でどのように配置するかなどについては関与しない。これらは別途に与えるマクロの定義によって自由に制御できる。この処理の手法は **troff** のプリプロセッサの全てに共通している。

図-5 の(1)式を **eqn** で処理した結果を簡略化して図-7 に示す。これを例として、数式のような複雑な対象に対処するために **troff** が用意している手段を以下に眺めておく。

### 3.1.3 マクロ変数

ここで **troff** のマクロ機能について補足しておく。**troff** の演算機能の変数として、値として数値を取るもの (number register) と文字列をとるもの (string) の2種が存在する。ここでは読み易さのために変数の表記を抽象化して数値変数をギリシャ大文字で、文字列変数をギリシャ小文字で示す。実際の表記例は文字列変数に関するものが図-13 の2以下にある。

### 3.1.4 数式の組み版

さて、図-7 の説明を始めよう。目的の数式は最後の15で出力される。それまでは文字列変数  $\phi$  の中に数式を組み立てる過程である。まず文字列変数  $\phi$  に文字列 '分数' を代入し(1)、その後 '\(eq' (等号) を追加する(2)。3, 4も同様の文字列の操作である。ついで各文字列の長さを求め、数値変数に代入する(5)。そのために、関数と呼ばれるエスケープ・シーケンスの一種の '\w' を使用している。ここで '' で囲まれているのは関数の引数である。7, 8では目的の分数式の幅を、分子と分母の長さの最大値として求めている。

$$\text{分数} = \frac{\text{分子}}{\text{分母}} \quad (1)$$

$$\text{定積分} = \int_{\text{下限値}}^{\text{上限値}} \text{被積分関数}(x) dx \quad (2)$$

図-5 eqn による出力例

```
1 .EQ (1)
2 分数 = 分子 over 分母
3 .EN
4 .EQ (2)
5 定積分 =- int sub 下限値
6 sup 上限値 被積分関数 (x) -dx
7 .EN
```

図-6 eqn の原稿例

```
1 .ds \phi 分数
2 .as \phi \(eq
3 .ds \gamma 分子
4 .ds \delta 分母
5 .nr \Gamma \w'\gamma'
6 .nr \Delta \w'\delta'
7 .nr \Lambda \Gamma
8 .if \Delta>\Lambda .nr \Lambda \Delta
9 .ds \tau \
10 \v'30u'\h'\Lambda u-\Delta u/2u'\delta\
11 \v'-54u'\h'-\Delta u-\Gamma u/2u'\gamma\
12 \v'12u'\h'-\Gamma u-\Lambda u/2u'\l'1'\Lambda u'\v'12u'
13 .as \phi \tau
14 .ds \phi \fI\phi\fr
15 \phi
```

図-7 eqn による数式の組み版

'**.if**' リクエストは各種の条件判断を行い、結果に基づいて入力を選択を行う。

9 から 12 までは 1 個のリクエストで、分数式を組み立てる核心のところである。行末の '\h' はリクエストの次行への継続を表す。組み立ての手順は、先ず下がって水平位置を分数の中央に調節した後で分母を置き、次いで上がって分子を中央に置き、最後に元の垂直位置へ戻りながら横線を引いている。'\v' と '\h' は垂直と水平の方向の印字位置の移動を指示する関数で、移動の方向は引数の値が正のときにそれぞれ下と右になる。'\h' 関数の引数中の演算の意味は、2 個の量の差の 1/2 である。通常の算術演算の優先順位は存在しない。数値の最後に付加されている 'u' は長さの単位で、出力先の植字機の分解能である。**troff** は演算に先立って数値に各種の単位間の変換を施す。これによる混乱を避けるためにここでは単位を明示している。この出力例は 300 dot/inch で作成されている。'\l' は指定された長さの水平線分を引く関数である。出来上がった分数式を '分数=' に連結し(13)、全体をイタリックにして完成する(14)。

ここで注意を引くことは、**eqn** が文字列長の計算など物理的な事項の処理を全て **troff** に委ねているこ

とである。

### 3.1.5 弱点

eqn は通常の使用法の限りでは簡便で使いやすい。しかし大きな構成の式は仕上がりが美しくないことがある。原因は troff が光学式の植字機用に設計されていることにある。積分記号や根号のような記号についても大きな字形が準備されておらず、全植字サイズに共通の字母を単純に拡大して使用する。そのため記号の字形のバランスが崩れてしまう。本格的な数式の組み版を行うには T<sub>E</sub>X の方が優れている。

### 3.2 表の組み版

#### 3.2.1 表の記述

表の作成は、行や欄の寸法の概念がはっきりしているのので、数式に比較すれば簡単である。しかし各欄にはいるデータの性質が異なる、複数の欄を臨時に融合させる、罫線の引き方が複雑である、といったことがあるとやはり面倒になる。ことに行の縦幅や欄の横幅の決定は文字のサイズや字幅が可変になると手におえない。tbl<sup>2)</sup> はこれらの問題を手際よく解決している。eqn と組み合わせて使用すれば数式の表の作成が可能であるなど、実用性は高い。

図-8 に表の組み版出力の例を、図-9 にはこの表を作成するための原稿を示す。以下に tbl のおよその機能を説明する。1 個の表は二つの部分からなり、全体が '.TS' と '.TE' で囲まれる。前半の '.' で終る部分が表の性質を定義している (5)。これは更に ';' で終るオプションと (2)、その後の形式指定の部分とに分かれる (3-5)。**'box'** は表全体を箱にいれる。形式指定では、行毎に各欄の持つ属性を指定する。欄の基本的な属性は位置揃えに関するもので、左 (l)・中央 (c)・右 (r)・数値 (n) などがある。**'n'** の欄では小数点の位置が表全体を通して縦に揃えられる。もう一つ面白いのは互り欄 (s, span) である (3)。1 個の 'c' の欄が右の 3 個の欄に互って広げられている。形式指定の 'l' は欄の間に縦の罫線を引くことを指定している (4)。

さて 6 から 12 までが表を構成するデータである。データは表の行毎に 1 行とし、欄は ASCII の TAB 文字 (図-9 では ⊙ で表示) で区切る。各データ行は形式指定中の対応する行に従って作表される。7 と 9 のように水平の罫線を引くためのものがあるが、これは形式指定中の行には対応しない。形式定義よりもデータの行数が多い場合には、定義の最後の行が繰り返し使用される。各欄の文字サイズやフォントなどの細かな属性は形式定義で行うほか、データから与えることが可

ジョブ統計			
区分	TSS	バッチ	合計
教員	9,837	12,290	22,127
学生	83,658	11,514	95,172
その他	22,463	16,112	38,575

図-8 tbl による出力例

```

1 .TS
2 box;
3 c s s s
4 c l c l c
5 l i n i n i n .
6 \s12ジョブ統計\s10
7
8 区分⊙TSS⊙バッチ⊙合計.
9 =
10 教員⊙9,837⊙12,290⊙22,127
11 学生⊙83,658⊙11,514⊙95,172
12 その他⊙22,463⊙16,112⊙38,575
13 .TE

```

図-9 tbl の原稿例

能である (6)<sup>12)</sup>。

#### 3.2.2 tbl の処理

tbl は与えられた定義とデータから、自動的に各欄の幅を決定する機能を持っている。使用者は欄の個数とデータの間には矛盾がないことにだけ注意すればよく、原稿の作成は容易である。

### 3.3 図形の組み版

コンピュータで図形を出力させること自体は何でもないので、普通のやり方では版下を作成する段階でプロット出力を本文と合成する手順を要する。pic<sup>2)</sup> では、本文と同一の原稿中に埋め込まれた記述から図形を組み版し、完成したページとして出力することができる。図-1 も pic による出力であるが、簡単な出力例を図-10 に、その原稿を図-11 に示す。

pic もパーザ・ジェネレータである yacc を使ってプログラムされているので、図形記述の文法が明快・柔軟である。基本図形がいくつか用意されており、それとテキストを組み合わせて指定することで、容易に図を組み立てることができる (2-6, 9-12)。テキストによる記述だけで、作り出す図形の位置関係を事前に確定的に指定することには困難が多い。pic では基本図形毎にサイズと作図後の移動量に対して標準値を設定することによって、この問題の解決を図っている。図形間の相対的な位置の指定を可能にする各種の記述法も強力に用意されている (7, 12)。この例には入っていない

<sup>12)</sup> JTROFF における漢字のサイズの指定方法はこれとわずかに異なっている。なお図-6 の 5, 6 では漢字のサイズ指定を省略してある。

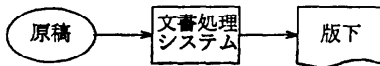


図-10 pic による出力例

```

1 .PS
2 ellipse "原稿"
3 arrow
4 box "文書処理" "システム"
5 arrow
6 box invis "版下"
7 move to end of last arrow
8 move down 0.25
9 line up 0.5
10 line right 0.75
11 line down 0.5
12 spline from end of last line\
13   left 0.25 up 0.05 then\
14   left 0.25 down 0.1 then\
15   left 0.25 up 0.05
16 .PE

```

図-11 pic の原稿例

いが、図形にラベルを付けることができ、それを位置指定の中で引用することも可能である。サイズを指定する単位の標準は inch であるが(8-15)、『.ps』マクロ上の指定で図全体のサイズの決定できるなど、ここも柔軟になっている。図形の取り扱いを容易にするための構造化の手段として、各種の演算・関数、ブロック化、マクロ、さらには‘if’や‘for’のような制御文が用意されている。

troff は専用の関数で図形の処理を行う。例えば 12 のスプライン曲線は

```
\D'~ Δx1 Δy1 Δx2 Δy2 ...'
```

のようになる。ここで (Δx<sub>i</sub>, Δy<sub>i</sub>) は隣接する 2 頂点の平面座標の差である。

pic は随分と頑張っているが、複雑な図を書き下すことはやはり難しい。それで対話的に作成する図を pic への入力に変換するプログラムが必要になる<sup>24)</sup>。別の応用として pic への入力を自動作成するプログラムが多数開発されている。例えば grap<sup>25)</sup> は数値データから折れ線など各種のグラフを作成する。

図形の位置を決定するための全く違った手法として、連立方程式を解く機能を組み込んだ ideal というおもしろいソフトウェアも存在する<sup>26)</sup>。

### 3.4 参照文献の引用

研究論文などの原稿の準備において、参照文献の書誌情報を記述するには面倒なことが多い。使用する文献の書誌情報をデータ・ベース化しておき、原稿の中で簡単なキーワードで呼び出して、適切な場所に自動的に埋め込むようにしたのが refer<sup>27)</sup> と bib<sup>28)</sup> で

```

1 %T PIC \ (em A Language for
   Typesetting Graphics
2 %A Brian W. Kernighan
3 %J SP&E
4 %V 12
5 %N 1
6 %P 1-21
7 %D 1982
8 %A 長谷部紀元
9 %A 龜山豊久
10 %T 和欧混合文書組み版システムの試作と
   組み版規則の検討
11 %P 10
12 %N 21-4
13 %R 情報処理学会日本語文書処理研究会資料
14 %D 1985
15 %K jtroff kenkyuukai

```

図-12 bib のデータベースの例

```

1 .[-
2 .ds [F 1
3 .ds [T PIC \ (em A Language for
   Typesetting Graphics
4 .ds [A Kernighan&B&W&B
5 .ds [J Softw. Pract. Exper.
6 .ds [V 12
7 .ds [N 1
8 .ds [P 1-21
9 .ds [D 1982.
10 .][
11 .[-
12 .ds [F 2
13 .ds [A 長谷部紀元
14 .ds [A 龜山豊久
15 .ds [T 和欧混合文書組み版システムの試作と
   組み版規則の検討
16 .ds [P 10
17 .ds [N 21-4
18 .ds [R 情報処理学会日本語文書処理研究会資料
19 .ds [D 1985
20 .][

```

図-13 bib の中間出力例

ある。何れも転置ファイルを使用している。bib では掲載雑誌毎の習慣に合わせた組み版結果を同一の書誌情報から得られるよう、機能が強化されている。以下これについて述べる。

図-12 にデータ・ベースに蓄える書誌情報の例を示す。原稿の中ではキーワードを

```
... pic[ kernighan pic 1982 .]では...
```

```
... 必要がある[ jtroff kenkyuukai .]...
```

のように指定すればよい。これの組み版出力は参照文献中の 2) と 11) の項にある。bib による中間出力を図-13 に示す。内容はデータベース中の該当する行を変換したものを、タイトルなど各種の書誌項目に当たる文字変数へ代入するリクエストの列である。変換機能には著者の名前の頭文字の抽出や転置(4)、雑誌名の略称からの展開(5)などがある。1、10 はこの中間出力を処理するための troff のマクロ。4 と 14 のギリシャ小文字はマクロによって変換すべき文字列変数

である。変換指定とマクロのセットを雑誌毎に用意しておけば、使用者は雑誌名の指定だけで容易に、目的の形式での組み版結果を得ることができる。

書誌情報の取扱のレベルになると、自然言語としての考慮が多少なりとも必要になる。漢字名には頭文字の習慣はない。14の8は英語ならば'and'になるが、日本語では','にする必要がある。bibにはこのような必要に対処するための仕掛が部分的にはあるが、もとより非ヨーロッパ系の言語に対する準備があるわけではない。

## 4. UNIX 環境

### 4.1 ソフトウェア・ツールと文書処理

以上述べてきたように UNIX の文書処理機能が充実したのは、UNIX が本来持っている特質が文書処理に好適であったためと考えられる。文書処理の対象は言うまでもなくテキストであり、これは UNIX の最大の応用分野であるソフトウェア開発作業におけるもっとも重要な対象である。したがって UNIX の環境とソフトウェア・ツールの恩恵を最大限に受けることができた。UNIX 環境の中で使用できることが TROFF の最大の特徴にもなっている。

原稿を準備する段階で有効なツールは数え切れない。高度のプログラミングが可能なテキスト変換ツールである `awk` を使えば、数値データのファイルから作図プログラム用の複雑な入力を自動作成するようなこともできる<sup>29)</sup>。複雑なファイル構成の文書を組み版するときには、シナリオに基づいて必要なコマンドを実行する機能を持ったツールである `make` が役に立つ。

### 4.2 推敲支援機能

もちろん文書処理専用のプログラムも多数存在するが、それらもまた UNIX 環境の中でスマートに試作・開発が行われる。改訂部分に印を付加した文書の原稿を自動的に作成する `diffmark` は `diff` を利用している<sup>30)</sup>。また索引を作成するためのツールである `ndx, subj`<sup>22)</sup> など同様に既存のコマンドを最大限に利用して、新たにコーディングする部分を必要最小限にとどめている。

このような環境では、単純なテキスト処理のツールの域をこえた文章推敲のためのソフトウェアも自然に着想されたのであろう。それらの例として英語の綴り誤りのチェックを行う `spell`<sup>31)</sup>、文体に関する統計的な量を計測する `style` や常用句に関して妥当性を調べる `diction` など<sup>32),33)</sup> がある。これらは実用的で

あるばかりでなく、内容的にも興味深い。

## 5. 研究環境における文書処理

### 5.1 パッチ処理と対話型処理

研究者・技術者にとって文書作成システムが興味を引く理由は、原稿作成の効率化と高品質の表現手段の自由な制御にあると考えられる。この観点からみて、どのようなシステムが望ましいであろうか。

対話的手法による文書作成の利点はよく知られている。これに対して、出力装置に対する独立性の不足や原稿に対する系統的な操作の弱点が指摘されている<sup>6)</sup>。しかし汎用コンピュータ上の文書作成システムとしてより重要であるのは、プログラムの実行結果などコンピュータから得られる情報を利用する機能である。これを実現するには、TROFF のように支援プログラムと協同するパッチ型の処理の方が明らかに優れている。

更にパッチ型の処理を行うシステムは汎用の高品質出力サブシステムとして使用することもできる。例えば統計計算パッケージ `S`<sup>4)</sup> は結果のグラフを `pic` への原稿の形で出力する。プログラムを解説文書と一体のものとして作成しようとする `WEB`<sup>34),35)</sup> は正にそのようなサブシステムの存在を前提としている。高品質の出力を直接に作り出すことの可能性が、コンピュータでの情報処理の新たな展開のきっかけになることも多いと考えられる。

文書作成システムのあるべき方向としては、同一のシステムがこれら二つの方式を兼ね備えることであろう。そのためには対話型システムが、外部との情報交換が可能で十分に強力な論理的構造を持った文書データを処理するようになればよい<sup>36)</sup>。

### 5.2 日本語への応用

日本語の表記法はヨーロッパ系言語と大きく異なるため、`troff` のような整形プログラムを日本語に応用することは簡単ではない。まず出力行への追込み処理の単位がアルファベットの連続した「語」ではなく文字である。これは日本語の文字を語と同一視すれば基本的に解決できる。さらに、組み版規則に関わる記号の種類は多く、また漢字とアルファベットのように設計原理の異なる文字の隣接の問題もある。各種の文字の間に挿入するアキを前後関係に依存して調節することが、印刷効果の上で不可欠である。マクロの処理によってテキストが変換されることを考慮すると、隣接する文字の種類は整形の実行時にしか決まらない。し

たがってアキの幅の決定など、日本語文書の高品質の整形を行うには、整形処理のアルゴリズムをかなり立ち入ったところで変更する必要がある<sup>11)</sup>。TROFFを各国語の組み版システムを応用する研究は中東の言語などに対しても行われている<sup>37)</sup>。

UNIX という柔軟で強力なシステムの上で日本語の処理が可能になり、また高品質の文書作成が可能になったので、推敲などを含む文書処理の研究<sup>38)</sup>が今後盛んになると考えられる。

### 5.3 文書処理の学術出版への応用

日本でもワークステーションとレーザ・ビーム・プリンタ、それにTROFFのような組み版ソフトウェアの普及によって、研究者が出版物と変わらぬ表現能力・品質の文書を手元で作成することが可能になろうとしている。現在のワープロによる清書とは質的に異なることが可能になるわけである。研究の各段階において高品質の文書をこの手段で作成することが普通となるであろう。そうなれば最終的な研究論文あるいは専門書も同一の手段によって作成することが自然である。日本においても学会誌などでこの傾向に対処することが必要になることが考えられる。

これに対する大きな障害は、多くの経費を必要とする漢字の高品質デジタル・フォントの準備である。本稿の版下も、この点について強く制約を受けている。METAFONT<sup>16)</sup>によって漢字フォントを統一的に設計する研究<sup>39)</sup>もあるが、字種の多さの困難を克服するには時間を必要とするようである。早急な解決策が望まれる。

## 6. ま と め

troff を中心とする UNIX の文書処理機能は、汎用コンピュータ上のものとしては早い時期に開発されたものであり、歴史的な制約を受けていることは否めない。しかしながら強力な支援ソフトウェアが多数蓄積されており、実用性は極めて高い。また、新たなソフトウェアとの協働動作を可能にしている柔軟な拡張性は、現在でも十分に生産的であると考えられる。

### 謝辞

本稿の版下をJTROFFで準備することができたのは編集委員会と事務局の協力によるものである。版下の作成作業には図書館情報大学の亀山豊久氏の協力を得た。使用した漢字フォントは(株)リコーの提供による。ここに記して各位に深く感謝する。

## 参 考 文 献

- 1) Kernighan, B. W. and Cherry, L. L.: *A System for Typesetting Mathematics*, *Comm. ACM*, Vol. 18, No. 3, pp. 151-157 (1975).
- 2) Kernighan, B. W.: *PIC — A Language for Typesetting Graphics*, *Softw. Pract. Exper.*, Vol. 12, No. 1, pp. 1-21 (1982).
- 3) Kernighan, B. W. and Lesk, M. E.: *UNIX Document Preparation, Document Preparation Systems*, Nievergelt, J., et al. (Ed.), pp. 1-20, North-Holland (1982).
- 4) Becker, R. A. and Chambers, J. M.: *S: An Interactive Environment for Data Analysis and Graphics*, p. 550, Wadsworth (1984).
- 5) Ritchie, D. M.: *The Evolution of the UNIX Time-sharing System*, *AT&T Bell Lab. Tech. J.*, Vol. 63, No. 8 Part 2, pp. 1577-1593 (1984).
- 6) Gehani, N.: *Tutorial: UNIX Document Formatting and Typesetting*, *IEEE Software*, Vol. 3, No. 5, pp. 15-24 (1986).
- 7) Furuta, R., Scofield, J. and Shaw, A. C.: *Document Formatting Systems: Survey, Concepts, and Issues*, *Comput. Surv.*, Vol. 14, No. 3, pp. 417-472 (1982).
- 8) 和田英一: エディタとテキスト処理 (13) *Runoff, bit*, Vol. 15, No. 4, pp. 372-379 (1983).
- 9) 木村泉: 文書整形言語, 情報処理, Vol. 22, No. 6, pp. 559-564 (1981).
- 10) 古郡延治: 英文清書法の基本操作とアルゴリズム, 情報処理, Vol. 24, No. 4, pp. 499-506 (1983).
- 11) 長谷部紀元, 亀山豊久: 和欧混合文書組み版システムの試作と組み版規則の検討, 情報処理学会日本語文書処理研究会資料, No. 21-4, p. 10 (1985).
- 12) 亀山豊久, 長谷部紀元: 和欧混合文書組み版システム JTROFF の開発, 情報処理学会 第31回全国大会講演論文集, pp. 1413-4 (1985).
- 13) Ossanna, J. F.: *NROFF/TROFF User's Manual*, *Computer Science Technical Report No. 54*, Bell Laboratories (1976).
- 14) Hasebe, K., Nomoto, S., Ishida, H. and Ishihata, K.: *An On-Line Phototypesetter Support System*, 情報処理学会 第19回全国大会講演論文集, pp. 771-772 (1978).
- 15) Kernighan, B. W.: *A Typesetter-independent TROFF*, *Computer Science Technical Report No. 97*, Bell Laboratories (1982).
- 16) Knuth, D. E.: *T<sub>E</sub>X and METAFONT: New Directions in typesetting*, Digital Press (1979).
- 17) Knuth, D. E.: *The T<sub>E</sub>Xbook*, p. 483, Addison Wesley (1986).
- 18) Reid, B. K. and Walker, J. H.: *SCRIBE Introductory User Manual, and SCRIBE Format Designer's Guide*, p. 332, Carnegie-Mellon Uni-



- versity (1979).
- 19) Lampport, L.: *ÉTeX: A Document Preparation System*, p. 242, Addison Wesley (1986).
  - 20) Witten, I. H., Bonham, M. and Strong, E.: *On the Power of Traps and Diversions in a Document Preparation Language*, *Softw. Pract. Exper.*, Vol. 12, No. 12, pp. 1119-1131 (1982).
  - 21) Adobe Systems Inc.: *PostScript Language Reference Manual*, p. 321, Addison-Wesley (1985).
  - 22) *UNIX System V, DOCUMENTER's WORKBENCH, Software, Technical Discussion and Reference Manual*, AT&T (1986).
  - 23) Lesk, M.E.: *Tbl — A Program to Format Tables, UNIX Programmer's Manual, Seventh Edition, Vol. 2*, Bell Laboratories (1979).
  - 24) *CIP Graphics Program, 5620 Dot-Mapped Display Text/Graphics Guide*, AT&T.
  - 25) *UNIX System V, DOCUMENTER's WORKBENCH, Software, User's Guide*, AT&T (1986).
  - 26) Van Wyk, C. J.: *A High-Level Language for Specifying Pictures*, *ACM Trans. Gr.*, Vol. 1, No. 2, pp. 163-182 (1982).
  - 27) Lesk, M. E.: *Some Applications of Inverted Indexes on the UNIX System, UNIX Programmer's Manual, Seventh Edition, Vol. 2*, Bell Laboratories (1979).
  - 28) Budd, T. A.: *BIB — A Program for Formatting Bibliographies, User Contributed Software Supplemental Manual*, University of California (1983).
  - 29) Van Wyk, C. J.: *AWK as Glue for Programs, Softw. Pract. Exper.*, Vol. 16, No. 4, pp. 369-388 (1986).
  - 30) Kernighan, B.W.: *The Unix System and Software Reusability, IEEE Trans. Softw. Eng.*, Vol. SE-10, No. 5, pp. 513-518 (1984).
  - 31) McIlroy, M.D.: *Development of a Spelling List, IEEE Trans. on Commun.*, Vol. COM-30, No. 1, pp. 91-99 (1982).
  - 32) Cherry, L.: *Writing Tools, IEEE Trans. Commun.*, Vol. COM-30, No. 1, pp. 100-105 (1982).
  - 33) Macdonald, N. H., Frase, L. T., Gingrich, P. S. and Keeman, S. A.: *The Writer's Workbench: Computer Aids for Text Analysis, IEEE Trans. Commun.*, Vol. COM-30, No. 1, pp. 105-110 (1982).
  - 34) Knuth, D.E.: *Literate Programming, Comput. J.*, Vol. 27, No. 2, pp. 97-111 (1984).
  - 35) Thimbleby, H.W.: *Literate Programming in C: Cweb Manual & Small Example*, University of York (1984).
  - 36) Bond, S. J.: *The Andrew System*, Information Technology Center, Carnegie-Mellon University (1986).
  - 37) Buchman, C., Berry, D.M. and Gonczarowski, J.: *DITROFF/FFORTID, An Adaptation of the UNIX DITROFF for Formatting Bi-Directional Text, ACM Trans. Office Info. Sys.*, Vol. 3, No. 4 (1985).
  - 38) 牛島和夫: *Writer's WorkBench と推蔽支援ツール, bit*, Vol. 17, No. 4 別冊 ワープロと日本語処理, pp. 224-230 (1985).
  - 39) Hobby, J.D. and Guoan, G.: *A Chinese META-FONT, Proc. ICTP '83*, pp. 62-67 (1983).

(昭和61年11月14日受付)