

## 解説

技術文書整形出力システム:  $\text{T}_{\text{E}}\text{X}$ \*

藤田 博†

## 1. まえがき

最近、手紙など通常文書の作成にワードプロセッサが手軽に使われるようになったが、数式や算譜などを含む技術文書の作成にはまだものたりない。計算機周辺にレーザビームプリンタ等、高解像度のハードコピー装置やビットマップディスプレイが現われ始めるに伴い、エディタで作成した技術文書を高品位に出力するシステムへの要求が高まっている。

本文では、そのような要求に応えるシステムとして注目を集めている $\text{T}_{\text{E}}\text{X}$ \*<sup>1)</sup>について解説した後、他のいくつかの同様なシステムについても簡単に触れる。詳細については参考文献<sup>1),2),3)</sup>を参照されたい。

2.  $\text{T}_{\text{E}}\text{X}$ 

$\text{T}_{\text{E}}\text{X}$ はKnuth<sup>4)</sup>によって開発された。特に数式の整形について入念に考えられており、その出力の美しさには目を見張るものがある。

## 2.1 文章整形

基本的な文章整形機能として、行分け、頁分け、右揃え、センタリングなどを自動的に行う他、フォントも多くの字体やサイズの中から選択できる。

図-1に例を示す。 $\backslash$ は $\text{T}_{\text{E}}\text{X}$ のコントロールシーケンス開始記号、{, }は区切り記号である。 $\backslash$ hsize 199ptは頁の横幅を199ポイントにし、 $\backslash$ ctrline{...}はセンタリングを行い、 $\backslash$ vskip 0.10inは縦方向に0.10インチの空白をあげ、 $\backslash$ slはslantedフォントを使い、 $\backslash$ "は次の一文字にウムラウトを付け、 $\backslash$ cは同じくcedillaアクセント、 $\backslash$ parは段落の終わり、 $\backslash$ vfillはその頁の縦方向の余白をこの位置に詰めることを各々指示している。---はem-dashと称する1本の線となる。

```
\hsize 199pt
\ctrline(ANOTHER STORY)
\vskip 0.10in
\ctrline(\sl by A. N. onym)
\vskip 0.5cm
Once upon a time, in another
distant galaxy called  $\backslash$ "U $\backslash$ "u $\backslash$ c c,
there lived a computer
named I. H. Sibustim.\par
Mr. Sibustim---or "I. H.," as
he preferred to be called---
was good at typesetting, but he
had other nasty qualities.
For example, he gave up printing pages
when a typist entered a long text.
\par\vfill\end
```

## ANOTHER STORY

by A. N. Onym

Once upon a time, in another distant galaxy called Üüç, there lived a computer named I. H. Sibustim.

Mr. Sibustim—or "I. H.," as he preferred to be called—was good at typesetting, but he had other nasty qualities. For example, he gave up printing pages when a typist entered a long text.

図-1  $\text{T}_{\text{E}}\text{X}$ の文章入力とその整形出力

他に、頁の縦幅( $\backslash$ vsize)、横方向の空白( $\backslash$ hskip)、行間隔( $\backslash$ baselineskip)、段落間隔( $\backslash$ parskip)などを指定することができる。

## 2.1.1 箱と糊

$\text{T}_{\text{E}}\text{X}$ は、文書要素を箱(box)と糊(glue)という対象に抽象化して統一的に処理する。文字や縦線( $\backslash$ vrule)、横線( $\backslash$ hrule)はアトミックな1つの箱である。箱は横方向に並べたり( $\backslash$ hbox)、縦方向に並べたり( $\backslash$ vbox)してひとまとめにされ、上位の箱が構成される。例えば行は単語のhbox、段落や頁は行のvboxである。箱

† A Typesetting System for Technical Text:  $\text{T}_{\text{E}}\text{X}$   
by Hiroshi FUJITA (Central Research Laboratory  
Mitsubishi Electric Co.)

†† 三菱電機(株)中央研究所

\*  $\text{T}_{\text{E}}\text{X}$ はアメリカ数学会の登録商標である。



```


$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} \left( \sum_{i=1}^m \cos^2 x_i(t) \right) (f(t) + g(t)) dt$$


```

図-4 TeX の数式入力とその整形出力

数式を組む場合、指数、添字、分数等、大小の文字を使い分けたり文字を上げたり下げたりするが、TeXでは基本的に次に示す4つ、及びそれらの変形4つ、計8つのスタイルをもつことにより対処している。

- display style (独立行として組む)
- text style (文章中に埋め込む)
- script style (上付き, 下付き)
- scriptscript style (2次の上付き, 下付き)

また、フォントは3つのサイズが使われる。これら8つのスタイルと3つのサイズのフォントを組合せることにより、たいていの数式がうまく組み上がる。

もう1つ数式の整形で大切なのは、スペーシングの問題である。まず数式の構成要素は以下の7つのタイプに分類される。

- |           |                  |
|-----------|------------------|
| 1) 一般箱    | (変数 部分式)         |
| 2) 演算子箱   | ( $\sum$ log 等)  |
| 3) 算術演算子箱 | (+ - $\times$ 等) |
| 4) 関係演算子箱 | (= < 等)          |
| 5) 開き箱    | ( { ( [ 等)       |
| 6) 閉じ箱    | ( } ) ] 等)       |
| 7) 区切り箱   | ( , ; 等)         |

そしてこれらの箱の隣合せ方に応じてスペーシングが決められる。しかし、いろいろな数式の総ての場合に適切なスペーシングを行うような一般規則を与えることは困難なので、出力結果に満足できないときのために幾つか異なるサイズのスペースが用意され、ユーザが指定できるようにしてある。

2.3 作表

表は、次のコントロールシーケンスを与えて組む。

```

\halign (幅指定) { (書式行)(行1)... (行n) }

```

各行は列要素を@で区切り、\crで終了する。表の枠や仕切線は、横線(\hrule)、縦線(\vrule)を行列要素として並べることにより構成できる。例を図-6に示す。

```


$$1 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4}}}}$$


```

$$1 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \frac{1}{a_4}}}}$$

```


$$\sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + x}}}}$$


```

$$\sqrt{1 + \sqrt{1 + \sqrt{1 + \sqrt{1 + x}}}}$$

```


$$A = \begin{pmatrix} x - \lambda & 1 & 0 \\ 0 & x - \lambda & 1 \\ 0 & 0 & x - \lambda \end{pmatrix}$$


```

$$A = \begin{pmatrix} x - \lambda & 1 & 0 \\ 0 & x - \lambda & 1 \\ 0 & 0 & x - \lambda \end{pmatrix}$$

```


$$a_1 + b_1 w + c_1 w^2 = \alpha + \beta; \quad b_2 x + c_2 x^2 = 0.$$


```

$$a_1 + b_1 w + c_1 w^2 = \alpha + \beta; \quad b_2 x + c_2 x^2 = 0.$$

図-5 いろいろな数式の入出力例

2.4 マクロ

TeXのマクロは比較的単純な文字列置換マクロである。例えば、TeXのロゴは次のように定義する。

```

\def\TeX{\def\b{\hskip-1.5pt}
\hbox{T\b\lower2pt\hbox{E}\b X}}

```

マクロはパラメタをもつことができ、一般には次のように定義する。

```

\def (マクロ名)(パラメタ列) { (マクロ展開形) }

```

マクロをうまく用いることにより、コントロールシーケンスの量を少なくし、宣言的な書式指定のみ(後出のScribe的な使い方)で望みの文書出力を得るようにすることもできる。例えば、論文を作成する際、投稿する雑誌ごとの書式をマクロ定義したファイルデータベース化しておき、必要に応じて\inputコント



```

\font G=go \:G \lineskipOpt
---MM-----\Tpar
++e+?0?0++e|\par
+e++?0+00e|\par
+++?0+0++e|\par
++++e?e+++\par
+++++*****|\par

```

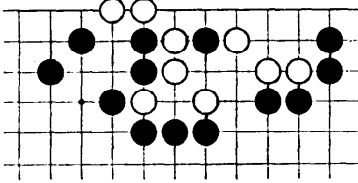


図-10 ユーザ定義のフォントの使用例

格的に画像を扱う拡張も行われており、フォントをラスタ形式ではなくベクタ形式でもつことにより任意のscaling, rotationを可能としている例もある。

### 2.7 TeXの拡張

現在TeXはPascal版のTeX82が標準とされている。KnuthはTeX82のためにWEBシステムを開発した。WEBソースファイルをTANGLEプログラムに通すとPascalソースファイルが得られ、WEAVEプログラムを通すとTeXソースファイルが得られる仕組みになっている。TeXのデバッグや使用サイトの環境に応じたカスタマイズ、機能拡張などの際に、TeX82のWEBソース本体はいじらず、変更ファイルをWEBで書いてTANGLE/WEAVEに“merge-read”させる方式をとるようKnuthは勧めている。

```

"The letter A";
call charbegin("A, 13, 2, 2, ph, 0, 0);
hpen;
lft0x1 = round 1.5u; bot0y1 = 0;
rt5x4 = round(r-1.5u); bot5y4=0;
top0y3 = top5y2 = h+0;
x3-x1 = x4-x2; rt5x2 = rt0x3;
w5 draw 2..4;
y5 = y6 = e;
x5-1 = (y5-y1)/(y3-y1) [x1, x3];
x6+1 = (y6-y4)/(y2-y4) [x4, x2];
w0 draw 5..6;
lpen#; w5 draw 3..5;
hpen; w0 draw 3..1;
if ucs #0;
call `a serif (1, 0, 3, -.5ucs);
call `b serif (1, 0, 3, +ucs);
call `c serif (4, 5, 2, -ucs);
call `d serif (4, 5, 2, +.5ucs);
fi.

```

図-9 METAFONTのプログラム例(文献<sup>9)</sup>より引用)

```

.EQ
1 over (2 pi)
int from (sqrt y) to (+ infty)
left ( sum from i=1 to m
      cos sup 2 x sub i (t) right )
left ( f(t) + g(t) right ) dt
.EN

```

図-11 EQNの数式入力(出力は図-4と同じ)

直ちに試みたい拡張の1つとして、かなや漢字のフォントを加え、日本文用の禁則処理を組み込んで日本語版を作ること等が考えられる<sup>8)</sup>。

TeXはおよそ0.5MBもの主記憶を必要とする巨大プログラムである。ユーザの間にはTeXを小さなマシンで使いたいという要求も強く、Z8000, M68000等をCPUとする、ワークステーションクラスのマシンへのインプリメントが数カ所で行われている。

### 3. 他の文書整形システム

TeXに比肩しうるシステムとして、UNIX\*\*のRoffを核とするプログラム群と、Scribe\*\*\*がある。

#### 3.1 Roff

UNIXにはNroff/Troff<sup>11)</sup>という文書整形プログラムがある。前者はタイプライタ型端末用、後者は写植機用と使い分けるが、いずれも文章整形のみを行う。

数式や表を含む文書に対しては、EQN<sup>12)</sup>, TBL<sup>13)</sup>というプログラムがNroff/Troffのプリプロセッサとして用意されている。EQNの数式入力(図-11)はTeXとほぼ同じ要領であるが、TBLの表入力(図-12)はTeXより簡単になっている。その他、文献リストの作成や引用を簡単化するREFERや、簡単な線画を文書に組み込むPIC<sup>14)</sup>というプリプロセッサもある。

このように文書の構成要素に応じて独立したプログラムが用意され、適宜これらをパイプでつないで処理する方式はいかにもUNIXらしい。

#### 3.2 Scribe

ScribeはReid<sup>15)</sup>によって開発された。TeXやRoff系との大きな違いは、整形のためのいろいろなパラメータや手続きが既にパッケージ化され、ユーザがその選択を宣言的に記述できるように設計してあることで、

\* 本文は、TeX78(Sail版)を日本語用に拡張したJ-TeX(仮称)によって作成した。

\*\* UNIXはBell研究所の登録商標である。

\*\*\* ScribeはUnilogic社の登録商標である。

```

.TS
box:
c s s
c | c | c
c | c | c
l | n | n.
Program Size
-
File@Source@Binary
name@(lines)@(words)
MAIN@480@5,693
DEFINE@778@7,408
SIMPL@1,860@20,703
OTHERS@743@7,214
INDUCT@791@8,638
UTIL@589@6,256
-
.T@
c | n | n.
total@5,242@55,912
.TE

```

図-12 TBLの表入力(出力は図-6と同じ)

定型的な文書を手軽に整形出力したいユーザには使い易いシステムといえる。

Scribeに比べ、T<sub>E</sub>Xはとっつきにくいとよく言われる。その主要な論点は、“ユーザを書式化の詳細に深入りさせすぎる”ことである。そこで、Fácil T<sub>E</sub>Xというユーザフロントエンドや、AMS-T<sub>E</sub>X<sup>TM</sup>というT<sub>E</sub>Xを数学者向けに使い易くした版が作られている。また、Scribeの思想を盛込んで文書スタイルの設定を宣言的に行えるようにした、B<sub>T</sub>E<sub>X</sub>という版もある。

Scribeの欠点は、T<sub>E</sub>XやRoff系に比べ数式や図表の整形出力機能が弱いことである。

#### 4. あとがき

現在T<sub>E</sub>Xは開発当初のStanfordにおけるDEC-20とXGPのホスト計算機/出力装置以外にIBMなどメインフレームからVAX, PDP-11, M68000に至る多くの計算機、並びに、Xerox-9700, レーザビームプリンタCanon LBP-10, 写植機Alphatype CRS, その他多くの出力装置において稼働している。

また、T<sub>E</sub>Xユーザ会(TUG)が結成されており、その会報からはdebug情報、拡張機能やマクロパッケージ等有益な情報を得ることができる。

終わりに、T<sub>E</sub>Xの日本語版への拡張にあたり、三菱電機(株)通信機製作所情報機器部の山内部長には漢字フォントの件でお世話になったことを感謝致します。

#### 参考文献

- 1) 木村 泉: 文書整形言語, 情報処理, Vol. 22, No. 6, pp. 559-564 (1981).
- 2) 和田 英一: エディタとテキスト処理(連載) 19-15, *Runoff* pp. 372-379, *Scribe* pp. 460-466, *T<sub>E</sub>X* pp. 589-595, *bit*, Vol. 15, No. 4-6 (1983).
- 3) Furuta, R., Scofield, J. and Shaw, A.: *Document Formatting Systems: Survey, Concepts, and Issues*, *Comput. Surv.*, Vol. 14, No. 3, pp. 417-472 (1982), 邦訳 金岡 恭: 文書書式化系: その調査, 概念, 論点について, *bit* 別冊コンピュータ・サイエンス, pp. 167-216 (1983).
- 4) Knuth, D. E.: *T<sub>E</sub>X and METAFONT: New Directions in Typesetting*, Digital Press and American Mathematical Society (1979).
- 5) Knuth, D. E. and Plass, M. E.: *Breaking Paragraph into Lines*, *Softw. Pract. Expr.*, Vol. 11, No. 11, pp. 1119-1184 (1981).
- 6) Knuth, D. E.: *The T<sub>E</sub>Xbook*, Addison-Wesley (1984).
- 7) Spivak, M.: *The Joy of T<sub>E</sub>X: A gourmet guide to typesetting technical text by computer*, American Mathematical Society (1980).
- 8) Beeton, B. eds.: *TUGBOAT The T<sub>E</sub>X Users Group Newsletter*, Vol. 4, No. 2, (1983).
- 9) 工藤 司, 藤田 博: T<sub>E</sub>X 出力システム, 情報処理学会第28回全国大会論文集 (1984).
- 10) Mei, T. Y.: *LCCD, A Language for Chinese Character Design*, STAN-CS-80-824, Department of Computer Science, Stanford University, (1980).
- 11) Ossanna, J. F.: *NROFF/TROFF User's Manual*, Programmer's Manual for UNIX<sup>TM</sup> SYSTEM III, Vol. 2A, 7 (1981).
- 12) Kernighan, B. W. and Cherry, L. L.: *Typesetting Mathematics — User's Guide (Second Edition)*, *ibid.* 10.
- 13) Lesk, M. E.: *Tbl — A Program to Format Tables*, *ibid.* 9.
- 14) Kernighan, B. W.: *PIC — A Language for Typesetting Graphics*, *Softw. Pract. Expr.*, Vol. 12, pp. 1-21 (1982).
- 15) Reid, B. K. and Walker, J. H.: *SCRIBE Introductory User's Manual*, Unilogic, Ltd. (1980).

(昭和59年4月20日受付)