

文書レイアウト方式の検討

4J-10

木谷 強

NTT 電気通信研究所

1. はじめに

近年、オフィスにおいても、ビットマップディスプレイや高精細プリンタを具備したワークステーションが普及し始め、文字のほかに図形が混在するマルチメディア文書の高品位な出力が可能になってきた。このような背景において、ワークステーションを利用した簡易な企業内印刷の要求が高まっており、文書を美しくかつ読みやすくする観点から、印刷業界の組版ルールを文書レイアウト処理にとり入れた、文書編集支援システムの実現が望まれている。

文書編集支援機能には、大きく分けて文章を推敲する機能⁽¹⁾と文書のレイアウトを整える機能がある。文書レイアウトに関して印刷業界の組版ルールを調査した結果、文字の並べ方を統一することと、全ページを通して文書の構成を統一することが重要であることが明らかになった⁽²⁾。そこで、文字間隔の調整方法と見出しの整形方法に関して組版ルールを分析した。本報告では、その分析結果と、ワークステーション上で動作する文書処理プログラムにその組版ルールを適用する際の実現方式について述べる。

2. 文字間隔調整方法

2.1 組版ルール

文字間隔の調整方法について印刷業界の組版ルールを分析した結果、以下の3点が重要であることが明らかになった。

(1) アキ

印刷業界では、文章を読みやすくするために、英字・数字と和字の間、および括弧類や記号類(約物と呼ばれる)の前後に「アキ」と呼ばれるスペースを挿入している。たとえば、文献(2)では、数字と和字の間は四分相当(全角の1/4)のスペースを挿入することを推奨している。

(2) 禁則処理

行頭・行末禁則処理のほかに、行分離禁止処理が必要である。行分離禁止処理とは、英単語など分割してはならない語を2行に分割しない処理である。また、行頭・行末禁則処理では、禁則文字が連続した場合の多重禁則処理も考慮する必要がある。

(3) アケ・ツメの優先順位

アキの挿入と禁則処理の結果、一般的には行末の位置が不揃いになるため、行の右端を揃える処理が必要となる。この処理を端数処理と呼び、文字間に適当なアキを配分し行末文字を追い出す場合と、存在するアキをツメることにより次行から文字を追い込む場合がある。アケ・ツメ処理において、アケ・ツメする文字間の優先順位は、アキの増減が目立

ないように個所から優先することを基本としており、表1に示すルールがある⁽²⁾。

表1 アケ・ツメする文字間の優先順位

優先順位	アケ	ツメ
1	漢字とかなの間	かなとかなの間
2	かなとかなの間	漢字とかなの間
3	漢字と漢字の間	欧文の語間
4	欧文の語間	欧文と和文の間
5	欧文と和文の間	約物の前後
6	約物の前後	漢字と漢字の間

2.2 実現方式

ワークステーション上で動作する文書処理プログラムに、上記の組版ルールを適用する際の実現方式を述べる。

(1) 処理形態

アケ・ツメ処理の結果は、ユーザによる文字の入力、修正と同時に表示されることが望ましい。しかし、複雑な組版ルールを文書処理プログラムに適用した場合、文字編集の基本である文字の挿入、削除処理の処理速度を低下させる可能性が高い。したがって、本機能は、インタラクティブな編集機能から独立させ、文書処理の一機能としてバッチ処理的に動作させることが望ましい。

(2) 文字フォント

従来の日本語ワードプロセッサは、全ての文字を全角、半角などの定まった大きさの単位でデザインしていた。しかし、英数字や約物は、レターフェイスに比べてボディーフェイスが小さいため、それらの文字が連続した場合は間延びした感じを与えている。したがって、英数字と約物については、そのボディーフェイスに応じてデザインされた文字フォントを使用する必要がある。

(3) アキの管理方式

アキの管理方式として、アキをスペース文字として表現する方法が提案されている⁽³⁾。しかし、この方式ではスペース文字によってデータサイズが増加することと、文字編集時に普通文字とスペース文字の両方を編集する必要があり、文字編集処理が頻繁に発生する文書処理では処理が煩雑になる。これに対し、アキを直前文字の文字ピッチに含ませる方法が考えられる。この方法は、可変長の文字ピッチが扱える文書処理プログラムにおいては、編集処理を変更する必要がないため有効である。

(4) 端数処理方式

端数処理の原則は、禁則文字を考慮しながら、アキをツメて文字がその行に入る場合は追い込み処理を行い、入り切らない場合はアキを追加して追い出し処理を行うものである。

アケ・ツメ量の優先順位とアケ・ツメ量の限界値は前後の文字種類によって異なるため、アケ・ツメ処理時に頻繁に前後の文字種類を判定する処理が必要となる。そこで、アケ・ツメ量の優先順位とアケ・ツメ量の限界値をテーブルに定義し(図1)、さらに文字データ構造に前後の文字の種類を示す識別番号の属性を持たせ、この識別番号によりテーブルを参照する方法を採用する。識別番号を利用することにより、アケ・ツメ処理時に、毎回前後の文字種別を判定する処理を省略することができる。また、テーブル内の数値を変化させることにより、アキの取り方を容易に変更することができる。

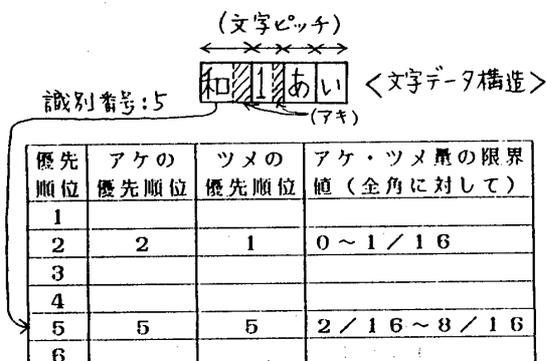


図1 識別番号を利用したアケ・ツメ処理の効率化

3. 見出しの整形方法

3.1 目的

見出しの周辺には、見出し行の前後に置かれる改行と、見出しと本文の先頭に置かれるインデントによって、空白が生じる場合が多い。見出し整形機能は、この空白部分を文書全体にわたり統一することを目的とし、これを自動的に生成する機能である。

次に、見出しによって発生する空白部分を自動的に生成する方法を述べる。

3.2 実現方式

見出しを判定するためにユーザに見出しを示すマークを付けさせる方法は、ユーザにマークの付与操作をさせるという点で好ましくない。そこで、見出しを自動的に判定する処理方法を検討した。

(1) 見出しの判定方法

見出しの表現上の形式には、以下の特徴がある。

- ① 行頭に記述順序や記述レベルを示す識別記号が付けられる。
- ② 見出しが複数行にまたがることはほとんど無い。
- ③ 見出し行の行末は句点で終わることは無い。

上記の特徴を利用して見出し行を判定する。すなわち、行頭は見出しの識別記号で始まり、行末に句点がなく復改で終わっているとき、見出し行と判定する。

(2) 識別記号と空白の取り方の指定方法

見出しの識別記号には、数字、英字、アラビア数字や、それらに括弧が付いたものなど、さまざまなものが利用される。また、改行やインデントの取り方もユーザの好みや文書の種類によってまちまちである。さらに、識別記号と空白の取り方は、見出しのレベルによっても異なる場合がある。したがって、標準的な見出し形式をユーザに提示し選択可能にするほか、ユーザ特有の形式も指定可能にする必要がある。識別記号と空白の取り方の指定方法の例を図2に示す。

本機能により、見出しによって発生する空白を文書全体にわたり統一できるだけでなく、空白を取るという作業を不要にできる効果がある。

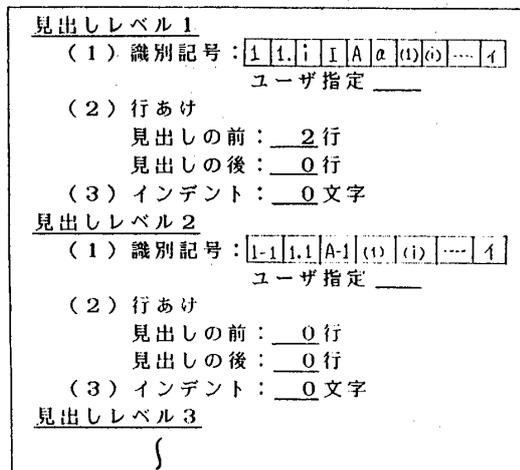


図2 識別記号と空白の取り方

4. おわりに

文書のレイアウトを整える機能のうち、文字間隔の調整機能と見出しの整形機能について、印刷業界の組版ルールを分析し、そのルールをワークステーション上で動作する文書処理プログラムにインプリメントする際の実現方式について述べた。

参考文献

- (1) 木谷 他: 「文書処理編集支援方式の検討」、昭和61年度電子通信学会総合全国大会、No. 1729
- (2) 大塚: 「写真植字の15章」、印刷学会出版部、1983
- (3) 長谷部 他: 「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」、日本語文書処理研究会資料、21-4、1985