

自然言語処理を用いた日本語文書自動整形システム

安原 宏[†] 小山 法孝[†]

近年、計算機を用いた文書整形技術は格段に進歩してきた。T_EX と呼ばれる自動組版システムや P_OST_SCRIPT と呼ばれる文字図形の記述言語の出現で電子出版が容易にできるようになった。これらの技術を利用すると従来の日本語行組版を変えるような新しい試みが可能となる。本論文では、日本語のべた詰め表記に対して欧米言語で取られているような単語分かち書きに近い手法の文書整形を提案し、実験システムの概要を述べる。整形処理は整形規則に基づいて実行しており、文書の種類によって異なる整形を施すことが可能となる。整形規則は単語や文字のサイズ、単語や文字の間隔を単語の見出しや品詞、構文構造などを用いて記述するため自然言語処理技術が必須となる。これらの規則を組み合わせると特定の単語を大きくしたり、助詞を小さくしたり、平仮名の続く文節の間には少し隙間を入れたりすることが可能になる。実験の結果、従来のモノスペース組版と比較してプロポーショナル組版の持つ読みやすさや自然さを出すことが可能になった。

Automatic Typesetting System for Japanese Documents Using Natural Language Processing

HIROSHI YASUHARA[†] and NORITAKA KOYAMA[†]

Computer typesetting technology has been progressed rapidly. A typesetting system T_EX and a programming language P_OST_SCRIPT to describe the appearance of text and graphical image made electronic publishing easy. Using these technologies, it will be possible to create new typesetting system for Japanese. In this paper, we propose English-like typesetting for Japanese documents and describe an experimental system to be realized as a new typesetting. The typesetting process is executed by interpreting typesetting rules which specify both the size of words or characters and the space between words or characters. The rules use linguistic informations such as headwords, part of speech and syntactic structures. Therefore the documents have to be processed by Japanese morphological and syntactic analysis before typesetting. The rule set can be selected according to a kind of the documents. The experiment system provided us with readability and natural appearance of Japanese documents.

1. はじめに

印刷は知的生産物のヒューマンインタフェースである。美しく、明瞭で、可読性の高い文書表現法の研究は、文書が社会生活で必須の情報源になっていることを考えると重要な研究テーマである。

日本語文書の一般的な印刷は中国から漢字を輸入して以来、仮名本等の例外はあるものの、ずっとモノスペース組版を取っている。文章を書くときに文節に区切るのは面倒であり、また印刷するさい編集者や植字工が分かち書きするのも効率の悪い作業になる。そもそも日本語は漢字仮名まじり文のため単語に切れ目を入れなくても十分読むことができるというのが一般的な考え方である。

出版印刷業界では電算写植の普及でコンピュータの

利用が進んでいるが、標準的な行組の大幅な見直しはない。最近の行組版の改良の動きとしては約物と呼ばれる字種や和欧文字の混合における字間の制御や禁則処理を定めたものがある^{1),2)}。いずれにしてもこのような行組版で行われていることは字種による行制御であって文章中の言語的役割を考えたものではない。

電子出版の世界では、カーニング(字間調節)やグルー(単語間調節)のきめ細かい組版をコンピュータで作れるようになっており、執筆者、編集者、職人あるいはデザイナーの手作業による新しい組版を目指した実験的試みも出ている³⁾。

コンピュータ技術の向上で、例えば T_EX のような柔軟な組版プログラムや P_OST_SCRIPT*のような文字・図形の記述言語を用いて従来の日本語組版にとらわれない新しい試みを検討してみるのには意義のあることで

[†] 沖電気工業株式会社研究開発本部マルチメディア研究所
Media Laboratory, OKI Electric Industry Co., Ltd.

* P_OST_SCRIPT は Adobe Systems Incorporated 社の登録商標です。

ある。プログラミング言語ではインデントーションを用いて構造を見やすくしたり、フォントを変更したり、全体のレイアウトを変える等の工夫がなされている^{4),5)}。新しい書体や表記法の研究が我々の文書表現に対する固定的な見方を変革することは十分考えられる。

組版の1つの改良方法を手書き文書に求めることができる。手紙などの肉筆でみる書体と印刷で見る書体では大きな違いがある。手書き文書では、一般に助詞を小さくしたり、文節間に適当な空白を挿入する等によって書き手は常に文字情報を確実に伝達することを考慮しているからである。また別の改良方法として、日本語以外の言語表記法を参考にすることができる。例えば英語には日本語にない表記法がある。日本語のモノスペース組版に対して、英語はプロポーショナル組版、固有名詞の先頭は大文字で書く等の特徴がある。もちろん日本語にもルビのような表記法がある。

文書の自然な表現法を追求すると、文書を単に文字列として見るだけでなく、情報伝達の単位となる文の形態素・構文情報および意味・文脈情報を内在したものと捉えることが大切となる。そのためには、自然言語処理技術が有効である。機械処理の観点からは意味処理や文脈処理はまだ完全に自動化の段階に達したわけではないので、しばらくは構文や形態素情報に限定するのが妥当であると考えられる。

本論では、2章で従来の日本語行組版の課題を述べ、3章でマルチメディア時代に適した日本語組版の可能性について考察し、4章で実験システムを説明する。5章では本システムを用いて出力した印刷例を示す。

2. 従来の行組版の限界

現在の行組版は電子化時代を迎えて改良が進みつつある。しかしながらそこで扱われる対象はあくまで文字を中心としたもので欧米語のように単語を基本としたものにはなっていない。我々が文書を読むとき、べた詰め表現は通常の文章であれば大半は問題なく理解できるが、よく注意してみると仮名が連続すると仮名を前方の自立語の付属語とするかそれ自体が自立語なのか曖昧になり頭の中で形態素解析と同様のバックトラックをしていることに気付く。以下は、筆者が本を読んでいて読み間違ったり、一瞬立ち止った例を示している。

「現われたのがもとでございました」

「頭の中からもどけたようです」

「大げんかをしようとかまやしないって気持」

「またたいがいはいつも涙ぐみながら」

「食べものものを通らない」

「顔にかかるもつれた厚い」

「何かののしりの言葉」

「まるでかまいつけなかったと前に申しました」

「経営が思いどおりにはこんでないに決まっている」

「時代からかわらず続いて」

「事務所ではしなかった」

これらに見られる特徴は、

(1) 平仮名の形態素分割が局所的に複数あり音読しても読みにくい。

(2) 速読したら隣の文字を飛ばして別の文字と誤って結合し、読み間違ふ。

のように日本語の仮名表記が原因になっているものが多く、従来の標準的日本語行組版では避けられない課題である。

読みにくい平仮名列を如何に解決するかで、我々は図1のような手紙や手書き文書に対し、べた詰め印字との違いを調査し、以下のような特徴点を見出した。

(1) 平仮名单語と平仮名单語の間には隙間が空けてある。

(2) 文節間には、若干の隙間がある。

(3) 助詞は小さめになっている。特に「の」や「を」は小さい。

(4) 活用語尾(送り仮名)は小さくなっている。

(5) 漢字と平仮名では大きさが異なる。

既に手書き風のフォントは開発されているが、文字のサイズや単語間の隙間(グルー)に関するこれらの現象は標準的な行組版では考慮されていない。文書を文字列として見ているからである。平仮名列をアルファベット列とみなすと、英語の行組版が参考になる。英語では以下のようなことが行われている。

(1) 単語と単語の間にグルーがあり単語の認識が容易である。

(2) 文字間のカーニングによって字間の連続性を保っている。

(3) 文の先頭が大文字で始まるため文頭が明確に

Figure 1 shows a sample of handwritten Japanese text. The characters are written in a cursive, somewhat irregular style. There are noticeable gaps between words and characters, and the sizes of the characters vary, particularly for particles and verb endings, which aligns with the text's description of handwritten features like smaller particles and irregular spacing.

図1 手書き文の例

Fig. 1 Example of handwriting letter.

なる。

(4) 固有名詞を大文字で表記するため、固有名詞の位置が明確になる。

(5) 行末ではハイフネーションが行われる。

つまり英語では文や単語が表記法上明確に区切られていること、品詞についても固有名詞のように特定のものを陽に表現している等の特徴がある。日本語でも小学校低学年の文書では文節間に空白があけてあるが、高学年になるに従ってべた詰めになってくる。アルファベットと仮名漢字という表記上の違いがあるので、どちらがよいとも言えないが、英語の行組版と同じような機能を日本語行組版に導入することの可能性や適切性を検討して見ることは必ずしも無駄なことではない。さらに英語の行組版でも見られないような新機能を追加することも可能となる。

3. 新しい日本語行組版の可能性

電子化時代で高性能なプリンタやいろいろな印刷ソフトが流通している。たとえばワークステーションのプリンタの解像度は、1インチ当たり数百ドットといったきめの細かさを持っているものが市販されている。また印刷の字体も豊富になり、アウトラインフォントといった美しい印刷技術が普及している。また $\text{T}_{\text{E}}\text{X}$ と呼ばれる文書処理システムを使うと使用者は自分の好みの字体やレイアウトで印刷できるようになっている⁹⁾。

我々の目的はこれらの技術を使って2章の課題を解決すると共に、手書き表記の特徴と英語の行組版を参考にして、美しく、明瞭で、可読性の高い日本語表記法を整形規則として組み込んだ日本語文書自動整形システムの枠組みを提示することである。2章で見たように手書き文書と英語の行組版には、単語と単語の間に空白を入れること、単語の持つ言語的役割に応じて文字のサイズを変えるという共通点がある。単語間の空白挿入を2章の例に適用すると以下ようになる。ここでは空白は一律に半角スペースとしている。

「現われたのが もとでございました」
 「頭の中からも どけたようです」
 「大げんかを しようと かまや しないって 気持」
 「また たいがいは いつも 涙ぐみながら」
 「食べものも のどを 通らない」
 「顔に かかる もつれた 厚い」
 「何か ののしりの 言葉」
 「まるで かまいつけなかったと 前に 申しました」
 「経営が 思いどおりに はこんでないに 決まっている」

「時代から かわらず 続いて」

「事務所では しなかった」

このように空白を入れるだけで明らかに文章は読みやすくなっている。さらに次章に示すような整形規則を導入することにより、文字サイズおよびグルー等の各種調整が可能になる。

4. 日本語文書自動整形システム

文章の整形を $\text{T}_{\text{E}}\text{X}$ を用いてユーザが指示すれば3章で示したようなことが可能になる。しかし長い文書を手で整形するには膨大な時間が必要で効率が悪く、広く普及するには任意の文章を自動的に整形するシステムを目指す必要がある。

整形処理には文章の単語見出しや文節情報、あるいは係り受け構造が必要なため自然言語処理技術が必須である。本論では形態素解析⁷⁾と構文解析⁸⁾を用いて整形範囲を単語の形態素情報と文節の係り受け情報に限定する。処理レベルを意味処理や文脈処理レベルにまで拡大すれば、もっときめ細かな整形処理が可能になるが、自然言語処理技術がそこまで自動化していないので使用していない。

4.1 システム構成

システムは図2に示すように文書を解析する日本語文書解析部、解析したデータから整形規則を用いて $\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$ ソースコードを生成し、さらにPOSTSCRIPTコードを生成する文書整形部、整形規則を格納する部分および整形したデータを表示、印刷する出力部からなる。以下、各部分について述べる。

4.2 日本語文書解析部

形態素解析と構文解析からなる。形態素解析の出力は単語分割結果と品詞および文節情報を与える。構文解析は文節間の係り受け情報と係り種類を与える。以

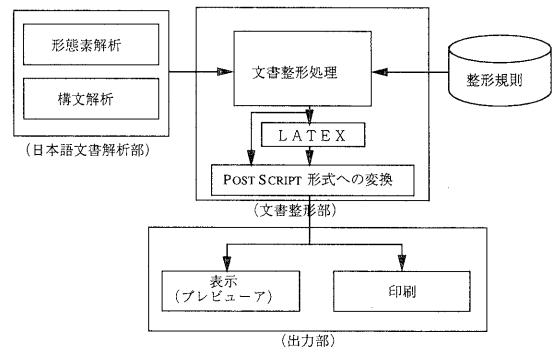


図2 日本語文書自動整形システムの構成

Fig.2 Configuration of automatic typesetting system for Japanese documents.

下に例で示す。

入力文例：「例えば計算機については、一般的には物理的性質の記述になるが、それだけならば問題点は少ない。」

この形態素解析結果，構文解析結果を図3に示す。実際の使用ではこれらの情報は表示しない。

4.3 文書整形部

文書整形部は，形態素解析と構文解析の結果を整形規則を参照しながら，文字サイズおよび文字間隔の情報を L^AT_EX のソースとして使用できる形式で出力する。その後POSTSCRIPTコードに変換する。この処理は L^AT_EX を用いるものと，システムを軽くするために直接 POSTSCRIPT コードに変換するものがある。

規則は 4.4 節で述べる。規則の適用順序は上から順に見て行き最初にマッチしたものを使用する。本システムでは整形規則をまとめてファイルとし，指定したファイルを使用できるようにしている。

4.4 整形規則

整形規則は大きく分けて文字サイズに関する規則および文字間隔に関する規則がある。それらはさらに単語と文字種に細分化されて合計 4 つの規則群に分かれる。規則はユーザが自由に記述できるようになっている。従ってユーザごとに個別に管理したり，用途ごとに設定することが可能である。以下各規則群について説明する。

(1) 単語サイズ規則群

単語を見出しまたは品詞で指定し，指定した単語の文字サイズを決める。

(記述法)

[見出し][品詞][+係り距離][-係り種類]<タブ> サイズ[サイズサイズ…]

※なお，サイズの書式は下記 (5) を参照。

単語は[見出し]，[(品詞)]のいずれか，または両方で指定する。見出しは形態素辞書に登録された見出しと完全に一致する必要がある。

係り距離とは，区切りの直前の文節が，何文節以上先に係っているかを指定するもので，この値以上先の文節に係っている場合にこの規則が適用される。+に続けて数字で指定する。

係り種類は，その文節の，他の文節への係り種類を指定する。-に続けて係り種類のコード (下記 (5) 参照) を指定する。

これらのいずれかを省略すると，無条件，すなわちすべての見出し・品詞または係り距離・種類について適用するとみなされる。したがって，見出し，品詞の両方を省略すると，「あらゆる」単語に対してその規則が適用される。

サイズは，1 文字だけ記述した場合，その単語のすべての文字について適用される。

サイズを 2 文字以上連続して記述すると，見出しの各文字について適用される。

なお，サイズの記述方法については現在，大，基(標準の大きさ)，中，小の順で指定する。

サイズのフィールドの後に，空白またはタブ等で区切った上でコメントを書くことができる。

(記述例)

| | | |
|---------|-----|-------------------------|
| システム | 大 | #「システム」は大サイズ |
| 例えば(副詞) | 基中中 | #「例」が基本サイズ，「え」「ば」が中サイズ。 |
| (助詞) | 小 | #助詞はすべて小サイズ。 |
| (助動詞) | 小 | #同様。 |
| (副詞) | 中 | #同様。ただし「例えば」は上の規則が優先。 |
| (動詞語尾) | 小 | |
| -GAS | 大 | #G格の文節の自立語は大サイズにする |

(2) 文字種サイズ規則群

単語とは独立に，文字種のみによってサイズを決める。これは，単語サイズ規則に一致するものがなかった単語についてのみ，デフォルトとしてこれを適用する。

(記述法)

文字種 <タブ> サイズ

文字種は，ひ(平仮名)，カ(カタカナ)，漢(漢字)，

形態素解析結果：
 例え(副詞) / ば(助詞) / 計算機(名詞) / について(助詞相当語) / は(助詞) / 一(記号) / 般(形容動詞) / 的(形容動詞語尾) / は(助詞) / 物(物理) / 理(名詞) / 的(接尾語) / 性(性質) / 質(名詞) / の(助詞) / 記(記述) / 述(名詞) / に(助詞) / な(動詞) / る(動詞語尾) / が(助詞) / 一(記号) / 般(記号) / 性(名詞) / 質(名詞) / 的(助詞) / なら(助動詞) / ば(助詞) / 問(問題点) / 題(名詞) / は(助詞) / 少(形容詞) / ない(形容詞語尾) / 。(記号)

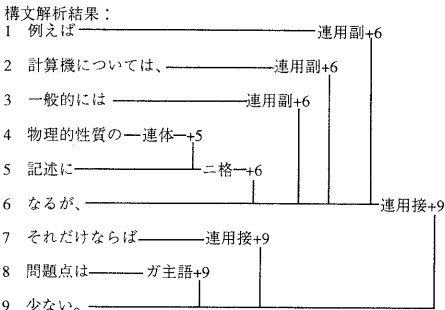


図3 形態素解析，構文解析結果の例

Fig. 3 Output of morphological and syntactic analysis.

A (アルファベット), 数 (数字), 記 (記号) の 6 種類とする。

(記述例)

漢 基 #基本サイズは省略可能
 数 基
 記 基
 A 基
 カ 基
 ひ 中

(3) 単語間隔規則群

連続した前後の単語 (または前のみ) を指定し, それらの間の文字間隔を決める。サイズ規則と異なり, 単語の他に文節区切りがあるかどうか, どの文節に係っているかを指定する。

(記述法)

[見出し 1][品詞 1][+係り距離][-係り種類]
 [/!//][見出し 2][品詞 2]<タブ> 間隔

※なお, 間隔の書式は下記 (5) を参照。

見出し 1 は, 区切りの直前の 1 単語で, 品詞 1 も同様である。

係り距離とは, 区切りの直前の文節が, 何文節以上先に係っているかを指定するもので, この値以上先の文節に係っている場合にこの規則が適用される。+に続けて数字で指定する。

係り種類は, その文節の, 他文節への係り種類を指定する。半角の-に続けて係り種類のコードを指定する。

なお, 係り距離, 係り種類の指定は, 文節区切り (/!//) を指定しないと無効である。

これらのフィールドは, サイズ規則と同様, 省略するとそのフィールドについては無条件に適用される。

ここまでだけを記述すると, 後ろの単語に関わりなく規則が適用されるが, 後ろの単語も指定する時は, [/] に続けて単語 2 を記述する。[/!] とすると, 両単語間に文節区切りがある場合に限り適用される。

※係り種類のコードは, 下記 (5) を参照。

(記述例)

(副詞)//(名詞) 1 #副詞で終わる文節と
 名詞の間は間隔 1
 ,// 0 #[,] で終わる文節は標
 準間隔
 +2-GAS// 2 #ガ格主語で, 2 つ以上
 先の文節に係る文節
 の後は間隔 2

(4) 文字種間隔規則群

単語と単語の境界で, 文字種のみによって間隔を決める。これは, 単語間隔規則に一致するものがなかつ

た単語についてのみ適用される。

(記述法)

文字種[/!//]文字種 <タブ> 間隔

文節区切りの指定は単語間隔規則の場合と同様である。なお, /とした場合は, 文節の区切りであってもなくても適用される。

(記述例)

ひ/ひ 1 #平仮名語と平仮名語の間は,
 間隔 1.

(5) サイズ, 間隔, 係り種類の書式

上記のサイズ, 間隔の記述については, 所定のコードを使用する。

<サイズ>:

大 基 中 小 濃の 5 種類 ; 大は基本+1 ポイント, 基は基本, 中は基本-1 ポイント, 小は基本-2 ポイント, 濃は基本サイズの太字

<間隔> :

0, 1, 2 の 3 種類; 0 は 0 mm, 1 は 0.8 mm, 2 は 1.2 mm とする。

<係り種類>⁸⁾:

GAS ガ格 1 (用言の主語)
 GAO ガ格 2 (目的語; 形容詞のガ)
 GAT ガ格 3 (トピック=ダ文の主語など)
 WO ヲ格
 NI ニ格
 DE デ格
 KARA カラ格
 TO ト格 (内容を表すもの)
 YORI ヨリ格
 HE ヘ格
 MADE マデ格
 NO ノ格
 SEN 文末型
 TNP 連体型 (連体詞)
 TNO 連体型 (ノ)
 TCA 連体型 (動詞の連体修飾)
 TPA 連体型 (並列)
 RAD 連用型/副詞的
 RCO 連用型/接続的
 RST 連用型/様相
 RIN 連用型/独立
 RTP 連用型/主題
 RPA 連用型/並列

4.5 出力部

文書整形部は POSTSCRIPT のコードを出力する。整形結果は POSTSCRIPT 言語が表示できるツール (例え

ば previewer)を持ったコンピュータであれば、画面上で見ることができる。画面を見ながらおかしい整形が見つければ、 \LaTeX ソースを手で直すことになる。

印刷は POSTSCRIPT プリンタを用いる。フォント、マージン、印刷部数等の指定はコマンドオプションで行える。

5. 適用例と効果

上記の文書整形システムを用いるときさまざまな用途の印刷が可能になる。以下にその例を示す。これらの例を見ると、通常の印字とは異なった効果を与えることがわかる。印刷結果に対する意見は人によってゆれがあるが、これは好みの整形規則を選択することで解消できると考えている。本例では、日本語 \TeX の原則であるベースラインに揃えて印刷すること、文字間隔の値は非負とするといった制約はそのまま踏襲したため、横書きの場合は、漢字と小さな仮名が連続するときは若干踊ったように見えたり、平仮名連続では間延びするように感じることもある。今後調整の必要な部

例えば 計算機については、一般的には物理的特性の記述になるが、それだけならば問題点は少ない。

(a) 印刷結果
(a) Printing image

```
##### 単語サイズ規則 #####
@wordsize
(助詞)          小
(助動詞)        小
(接続詞)        小
(動詞語尾)      小
(形容詞語尾)    小
(形容動詞語尾) 小
(形式名詞)      小
(助詞相当語)    小
(補助動詞)      小
(接頭辞)        基
(句読点)        中
(数字)          中
(記号)          中
##### 文字サイズ規則 #####
@mojisize
漢              基
数              基
記              基
A              基
カ              基
ひ              小
##### 単語間隔規則 #####
@worditv
、//           0
+2-TNP//       1
+2-TCA//       1
+2-TNO//       1
+2-TPA//       1
//             1 #すべての文節間隔を1
##### 文字種間隔規則 #####
@mojitv
ひ//ひ        2 #平仮名の文節間隔を2
```

(b) 標準型整形規則
(b) Standard typesetting rule

図4 標準型整形規則による印刷結果
Fig.4 Printing by standard typesetting rule.

分である。

5.1 標準型出力

本標準型出力例の整形規則は、付属語を小さくして、文節間の間隔を1単位とするという標準的な規則である。図4 (a) に印刷結果、図4 (b) にその整形規則を示す。

5.2 強調型出力

本強調型出力例の整形規則は、名詞、動詞、形容詞を強調している。本例では助詞「だけ」、助動詞、活用語尾等も強調した。また特定の単語や人名・社名等の固有名詞を強調することによって、文章中の特定単語を素早く読み取ったり、文書からの情報検索に利用できる。図5 (a) に印刷結果、図5 (b) にその整形規則を示す。

5.3 縦書き手書き型出力

本縦書き手書き型出力例は5.2節と同一の整形規則を用い、縦書き、マージン縮小、手書きフォントを指定して実行した。図6に印刷結果を示す。

6. おわりに

従来、文字列中心で決められていた日本語の行組版に自然言語処理を適用することによって、単語や構文

例えば 計算機については、一般的には物理的特性の記述になるが、それだけならば問題点は少ない。

(a) 印刷結果
(a) Printing image

```
##### 単語サイズ規則 #####
@wordsize
(名詞)          濃
(動詞)          濃
(形容詞)        濃
(動詞語尾)      濃
(形容詞語尾)    濃
(助詞相当語)    濃
(助動詞)        濃
(接尾語)        濃
だけ(助詞)      濃
##### 文字サイズ規則 #####
@mojisize
漢              中
数              中
記              中
A              中
カ              中
ひ              中
##### 単語間隔規則 #####
@worditv
、//           0
+2-TNP//       1
+2-TCA//       1
+2-TNO//       1
+2-TPA//       1
+2//           2 #離れて修飾する文節間隔は2
##### 文字種間隔規則 #####
@mojitv
ひ//ひ        1
```

(b) 強調型整形規則
(b) Stressed typesetting rule

図5 強調型整形規則による印刷結果
Fig.5 Printing by stressed typesetting rule.

例えは計算機については、
一般的には物理的特性の記
述になるが、それだけなら
ば問題点は少ない。

図6 縦書き強調型整形規則による印刷結果

Fig. 6 Vertical printing by stressed typesetting rule.

情報を考慮した整形規則を追加することを提案し、それに基づく日本語文書自動整形システムを試作した。

日本語における漢字仮名まじりの文章をいかに美しく、読みやすい表現に印字するかについて、主に手書き文書に現われる表記法および英語表記法を参考にして整形規則を作成した。

本システムはあくまで実験的なものであり、新しい日本語表記法の検討材料を提供するものである。誰でも使えるようにするためには、用途ごとに豊富な整形規則を作成し、それらをライブラリで利用できるようにしたり、自動整形で出力した整形文書を編集しやすくするヒューマンインタフェースの研究が今後必要である。

応用の観点から見ると、電子化ドキュメントが増大するにつれ、文字をディスプレイ上に見やすく表示することも大切な機能になってくるが本整形システムはそれにも役立つことが期待できる。

本研究は日本語を中心に考えたが、中国語の手書き文書でも日本語の手書き文書とよく似た表記が見られ、同様の手法は中国語、タイ語などのベタ詰め表記を取る言語にも適用することが可能と思われる。

参 考 文 献

- 1) 日本規格協会：日本語文書の行組版方法 JIS X 4051, 電子文書処理システムの標準化に関する調査研究報告書, C 34, pp.109-135 (1994).
- 2) 川崎敏治, 山内 司, 岩元哲夫, 目瀬道弘：文書清書システムにおける行組版機能の一実現方式, 情報処理学会論文誌, Vol.34, No.8, pp.1741-1751 (1993).

- 3) 津野海太郎：本とコンピューター, p.280, 晶文社 (1993).
- 4) 藤原泰行, 繁田英之, 玉木裕二, 並木美太郎, 高橋延匡：言語Cプログラムの読解を助ける出力ツール, 第44回情報処理学会全国大会論文集 (3), pp.267-268 (1992).
- 5) 木原一夫, 綿木 輝, 相沢栄男, 芝野耕司：プログラムパブリッシングシステム, 第46回情報処理学会全国大会論文集 (3), pp.249-250 (1993).
- 6) アスキー出版技術部責任編集：日本語 $T_{E}X$ テクニカルブック I, p.229, アスキー出版局 (1990).
- 7) 小松英二, 安原 宏：コスト最小法形態素解析のコストルールの作成方法, 情報処理学会自然言語処理研究会資料, 85-1, pp.1-8 (1991).
- 8) 山上晃司, 安原 宏：形態素情報による日本語の係り受け解析, 情報処理学会自然言語処理研究会資料, 98-2, pp.9-16 (1993).

(平成6年10月19日受付)

(平成7年3月13日採録)



安原 宏 (正会員)

1946年生, 1969年京都大学理学部数学科卒業, 1972年京都大学大学院理学研究科修士課程修了, 同年沖電気工業(株)入社, 1982~1992年第五世代コンピュータプロジェクトに従事, 1986~1995年(株)日本電子化辞書研究所第6研究室長, 1987~1995年(財)国際情報化協力センター機械翻訳システム研究所研究第6課長, 現在, 沖電気工業(株)マルチメディア研究所に勤務, 電子化辞書, 自然言語処理の研究開発に従事, 人工知能学会, 日本認知科学会, 言語処理学会, ACL, アジア太平洋機械翻訳協会各会員。



小山 法孝 (正会員)

昭和36年生, 昭和58年千葉大学理学部数学科卒業, 昭和63年筑波大学大学院博士課程数学研究科単位取得退学, 同年沖電気工業(株)入社, グラフィックス, 並列処理, 組み込み用マイクロプロセッサのソフトウェア開発環境, の研究開発に従事, 現在, マルチメディア研究所に勤務。