

ウィンドウズ用特殊サンスクリットフォント のためのエディタ作成

逢坂雄美, 鈴木享

仙台電波工業高等専門学校
〒989-3124 仙台市青葉区上愛子北原

ウィンドウズPCには拡張ASCIIコード上で定義されたフォント(特殊サンスクリットフォント, Pali96 or Norman)用の簡便な入力法がないので, 2年前にその為のテキストエディタ(EDISAN1.0)を開発した. このエディタは特殊フォント入力用に特化しており, 日本語フォントの入力処理等を考慮していなかった. その結果, 一度保存したファイルを開いた際に2バイトコードの日本語が1バイトコードに変換されてしまう. 我々は, これらの問題点を解決したエディタ, EDISAN3.0を開発した. 本エディタでは, 通常の検索/置換機能に加えて, 電子ブックの代用ができるように, リンク機能を追加した.

Simple Editor for Sanskrit Font on Windows

Yumi OUSAKA and Tohru SUZUKI

Sendai National College of Technology
Kami-Ayashi, Aobaku, Sendai 989-3124

We don't have any editor for Sanskrit fonts on Windows, which are defined over the extended ASCII code. We made ourselves the editor for Sanskrit font, EDISAN 1.0 two years ago. When making EDISAN 1.0, we took accounts of only the treatment of the Sanskrit font, and couldn't take accounts of the treatment of the so-called two-byte characters such as Chinese character and Kana-character. We revised EDISAN 1.0 up to 3.0 so that we can type such characters as well as the Sanskrit font. We added the lynx-function to the new editor, EDISAN 3.0 inclusive of the usual Find/Replace function.

1. 序論

これまで, 我々はブラークリット語・パーリ語・サンスクリット語用のフォント" Pali96"をMacintosh上で作成してきた. このフォントは, 拡張ASCIIコード上で定義されたフォントである. これを元に, 中期インド・アリアン聖典の解析ツールの構築等をMacintoshを使用して行ってきた. これらの研究ツールをWindows

PCでも使用できるように拡張中である. 最初に, 特殊フォントをWindows用True Type Fontに拡張し, その後このフォントを使って作成されたテキストを移植した. 日本語版 Windows OSを搭載しているパーソナルコンピュータ上で, 適当なワードプロセッサを使用してこのテキストを読んでみると文字化けを起し, 正確に読むことができないと言う厄介な事態が発生し

た。これは、2バイトコードフォントを標準にする日本語ワードプロセッサには、特殊フォントを表示するのに不都合な点があることに起因する。このバグを回避できたら、印刷と画面上の表示が一致する、所謂 WYSIWYG (what you see is what you get) 表示のテキストを利用できることになり、言語学的な研究の能率も向上し、1バイトの特殊フォントを基に作成された大事な基礎資料を有効に活用することができ、さらに、語彙の索引作成等のプログラム開発も容易にできるという、多くの利点がある。なお、我々のフォントと同種の特殊フォントとしては Norman Font があるが、これに対しても同様な状態になっている。

Windows PC 上では特殊フォント (Pali96 or Norman) の簡便な入力法がなく (入力法がない文字もある) 不便であるということから、これまで Delphi3.0 を使い、特殊フォント用の簡便なテキストエディタ (EDISAN1.0) を開発した。このエディタは特殊フォント入力用に特化しており、日本語フォントの入力処理等を考慮していなかった。その結果、一度保存したファイルを開いた際に2バイトコードの日本語が1バイトコードに変換されてしまう。又、フォントサイズを変更すると入力画面上でのカーソルの位置表示が異常になった。本論文では、これらの問題点を解決したエディタの開発について議論する。

エラーが起きる条件を詳しく調べてみると、1バイトコードで書かれた文字は正常に表示されるが、2バイトコードで書かれた文字のみに文字化けが起こっていることが分かった。ソースプログラムをチェックをしたが、文字化けが起こるようなエラーはなかった。1バイトコード特殊フォントのキー入力処理のプログラムをカットし、実行してみると、2バイトコード文字は正常に表示されることが分かった。このエラーは Delphi3.0 のリッチエディットコントロール (RichEditControl: 書式付編集コントロール) の不完全さに起因する可能性がある。Delphi3.0 上で文字化けの現象を解決するのは難しいと判断し、他の開発ソフトで EDISAN を開発することを検討した。

EDISAN 開発用のツールは特殊フォントを全

て正しく表示できることが、必要十分条件となる。VisualC++ では、言語として C++ を使っている。C++ 上での文字コードは普通のアルファベットには ASCII コード、日本語には SJIS コードを使っており、2バイト文字である日本語を表示させたいときは、文字コードの前後に SJIS コードであることを示すコードを書くことになっている。つまり、VisualC++ では2バイトコードが1バイトコードよりきちんと分離されており、127~255番目の拡張 ASCII コードを含めて、0~255番目のすべてのコード上の1バイト文字を正常に表示できるようになっている。この結果、拡張 ASCII コード上に定義されている特殊フォントを VisualC++ では正常に表示できる。それと同時に、SJIS コード上に展開されている日本語も正常に表示できる。以上のことから、Visual C++ で EDISAN の開発をすることにした。

マイクロソフト社より、ある程度のワープロ機能を備えた Word Pad のプログラムが公開されている。我々はこの Word Pad に、特殊フォント入力機能等の必要機能を付加することにした。この方法により、Word Pad の通常の機能を活かした上で、簡易エディタを迅速に作成できる。

ところで、通常の検索・置換機能に問題がある。この機能は特殊フォントに対応していないので、ダイアログボックスでは特殊フォントの表示と入力を正確にできなかった。表示機能を追加して、特殊フォント表示ができるように改良した。入力機能については現在検討中であり、コピー/ペースト機能により代用している。本論文ではこの機能のアルゴリズムについての議論を割愛する。

本エディタでは、これらの機能に加えて、リンク機能を追加することにした。我々は、これまで電子テキスト化された文献や解析データ等を、Macintosh 用の電子ブックとしてインターネット上に公開している。Windows PC 用にも電子ブックを作成してみたが、Windows PC 用作成ツールを使うと特殊フォントの一部が文字化けしてしまう。この結果、電子ブックを公開できない状態にあり、テキストデータのみの公開になっている。実際にテキストを閲覧するに

は、テキストデータという形では非常に効率が悪い。Windows用の電子ブックの代用ができるように、新しいEDISANに文章を読みやすくするためのリンク機能を付けることにした。

本論文では、第2節で特殊フォントのキーボード入力法とリンク機能について述べる。第3節ではキーボード入力法のアルゴリズムについて議論する。第4節で今後の課題等について述べる。

2. キーボード入力法とリンク機能

最初に特殊フォントPali96等の入力法について記述し、次いでリンク機能の使用法につい

て概説する。

2.1 特殊フォント入力機能

Windows PCで特殊フォントを表示するには、例えばASCIIコード140番目の‘ā’の場合、Alt key + 0140と入力する必要がある。それに比べMacintoshでは、Option key + aという簡単なキーストロークによりできる。このようにWindows PCでは、複雑なキーストロークが必要であり、また特殊フォントに対応したキーストロークが無い場合もある。MacintoshでのOption KeyをWindows PCでのAlt Keyに対応させ、特殊フォントの簡便な入力を可能に

表1. パーリフォントPali96のキー操作法とアスキーコード割り当て。

正規の文字 a, i 等は、勿論通常通りにタイプできる。特殊文字“ā”はAlt keyと正規の文字“a”を同時に押すことによりタイプできる。同様にして他の特殊文字“m”, “kh”等を打つことができる。特殊文字“ī”はAlt keyと正規の文字“i”を同時に押した後で引続き文字“i”を打つことによりタイプできる。同様にして“ū”を打つことができる。最後に、特殊文字“th”は3つのキー、Alt key, Shift key, 正規の文字“y”を同時に押すことによりタイプできる。

character	key stroke	numeric code	character	key stroke	numeric code
ā	<A>+n, a	136	ṭ	<A>+t	160
ā	<A>+a	140	ṭh	<A>+<S>+y	231
ī	<A>+` , i	147	ḍ	<A>+d	182
ī	<A>+i, i	148	ḍh	<A>+i, U	243
ū	<A>+e, u	156	ṇ	<A>+n, n	150
ū	<A>+u, u	159	th	<A>+y	180
ṛ	<A>+r	168	dh	<A>+z	189
ṁ	<A>+m	181	ph	<A>+p	185
kh	<A>+k	251	bh	<A>+b	186
gh	<A>+9	187	ḷ	<A>+l	194
ṅ	<A>+g	169	ḷ	<A>+<S>+l	241
ch	<A>+c	141	ś	<A>+s	167
jh	<A>+j	198	ṣ	<A>+x	197
ñ	<A>+<S>+i	246	ḥ	<A>+h	250

表2. 韻律図式の表示で使用される特殊記号の入力法とアスキーコード割り当て。キーボード入力法は、表1で述べたパーリフォントの場合と同様にできる。

character	key stroke	numeric code	character	key stroke	numeric code
/	<A>+1	193	—	<A>+5	176
//	<A>+2	170	∪	<A>+7	166
	<A>+3	163	∪	<A>+8	165
	<A>+4	162	ε	<A>+0	188
§	<A>+6	164			

表3. パーリフォントPali96の大文字のキー操作法とアスキーコード割り当て.

character	key stroke	numeric code	character	key stroke	numeric code
Ā	<A>+<S>+a	129	Ō	<A>+u, O	133
Ū	<A>+u, U	134	Ā	<A>+`, A	203
Ō	<A>+n, O	205	Ḍ	<A>+<S>+d	235
Ṭ	<A>+i, E	230	Ḣ	<A>+<S>+h	238
Ṣ	<A>+`, U	244	Ñ	<A>+n, <S>+n	132
Ṛ	<A>+u, i	149			

表4. ノーマンフォントのキー操作法とアスキーコード割り当て.

character	key stroke	numeric code	character	key stroke	numeric code
ã	<A>+n, a	139	ñ	<A>+n, n	150
ā	<A>+a	140	ı	<A>+ı	160
ì	<A>+<S>+`, i	147	đ	<A>+đ	182
ī	<A>+y	180	ñ	<A>+<S>+i	246
ũ	<A>+z	189	ı	<A>+ı	194
ū	<A>+u, E	232	š	<A>+s	167
ṃ	<A>+m	181	š	<A>+w	183
ñ	<A>+b	186	h	<A>+h	250

表5. ノーマンフォントの大文字のキー操作法とアスキーコード割り当て.

character	key stroke	numeric code	character	key stroke	numeric code
Ā	<A>+n, <S>+a	204	Ñ	<A>+n, <S>+n	132
Ā	<A>+<S>+h	238	Ṭ	<A>+`, <S>+e	233
Ī	<A>+<S>+y	231	Ḍ	<A>+<S>+q	223
Ū	<A>+<S>+u	242	Ñ	<A>+<S>+j	239
Ō	<A>+n, <S>+o	205	Ļ	<A>+<S>+x	226
Ṛ	<A>+`, <S>+u	244	š	<A>+<S>+s	234
Ṣ	<A>+<S>+c	200	š	<A>+<S>+z	221
Ṛ	<A>+<S>+t	220	Ḣ	<A>+<S>+g	222

するエディタを作成した。

表1～3に特殊フォントPali96の入力法を示す([A]はAltキー,[S]はShiftキーを表す)。

表4, 5にはNorman font に対するキー操作法を示している。

なお, Norman font の5個のフォント, 'Ḍ, Ñ, Ļ, š, Ḣ' に対するキー入力, キー操作の割り当て上の問題から, マッキントッシュの場合と完全には対応していない。

これらのキー入力法, 特にNorman font に対しては実際に言語学研究者のチェックにより, 正常に動作しており, 簡単に入力できるとの評価を得ている。

2.2 リンク機能

序論で述べたように, Windows PC では特殊フォントの一部が文字化けしてしまうために電子ブックを作成できない。最近 Windows PC 用 Acrobat 4.0 により, 2バイトコードの文字体系に対してもファイルをPDF化できるようになった。それを試したところ, 一応特殊フォントを使用したファイルを表面上, PDF化できるようになった。しかし, マッキントッシュ用の同じバージョンのAcrobatでは, 特殊フォントを画像処理してしまう。作成元のAdobe社に確認したところ, 拡張アスキーコードを使用したフォントへの対応は, 現在のと

ころ考えていないとのことである。Windows PCでPDF化できたとしても、その動作保証はできないとのことであった。それゆえ、今もってWindowsPC用の信頼できる電子ブック作成ツールは、残念ながら存在しない。我々はEDISAN3.0に文章閲覧の際に便利な機能であるリンク機能を付加し、このEDISAN3.0が電子ブック作成ツールを代用できるようにした。

リンク機能を使うには、メニューの[Link]から[View]を選択するか、ツールバーの一番右端のボタンをクリックし、リンクダイアログを表示させる(図1参照)。その後、以下の操作を実行する。

1. リンクをする項目を\$マークで囲んでおく。

2. Loadボタンをクリックすると、\$で囲んだ項目がリストボックスに自動的に読み込まれる。このとき、Readと書かれたチェックボックスをオフにしておく。
3. チェックボックスをオンにすると、自動的にテキスト上の\$マークが消え、リンク可能な状態になる。
4. リストボックス上の項目を選択すると、その項目の場所に自動的に飛ぶ。
5. チェックボックスを再びオフにすると、\$マークが元の位置に表示されて、リンク不可になる。

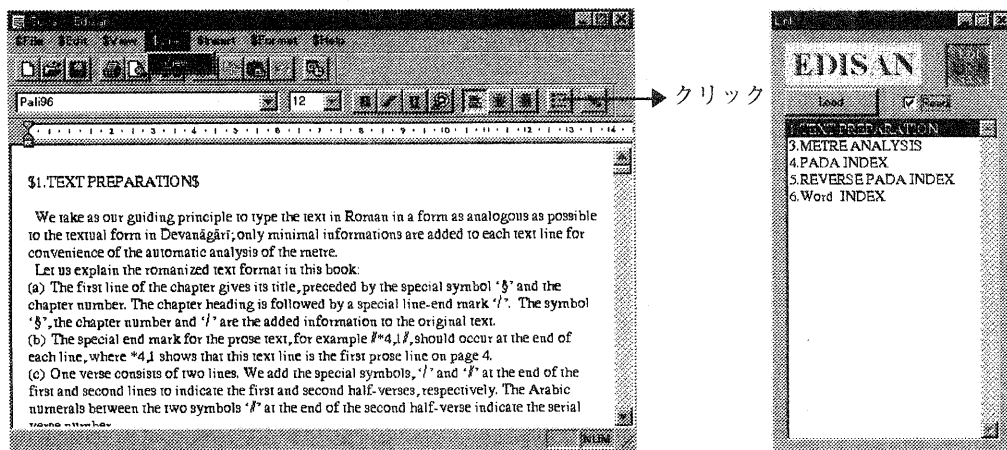


図1. EDISAN3.0とリンクダイアログボックス

3. 入力処理プログラム

表1~5から分かるように特殊フォントには、1ストロークと2ストロークの文字がある。これらの特殊フォントの入力処理をOnChar関数とOnSysChar関数により実現する。この処理の流れを図2に示し、図3、4にOnSysChar関数とOnChar関数の実際の処理を示している。

1ストローク文字'a'の場合、Alt Key+aを入力すると処理プログラムにはOnSysChar関数が発生する。そこでaの文字コードを取得し、その文字コードに対応している特殊文字'a'を表示する。2ストローク文字'dh'の場合は、1ストロー

ク目でAlt Key + iを入力するとOnSysChar関数が発生し、iの文字コードを取得する。その文字コードを2ストローク目で用いるために記憶しておく。そして、次の2ストローク目にUを入力しOnChar関数が発生したら、Uの文字コードを取得し、その文字コードと1ストローク目の文字コードに対応する特殊文字'dh'を表示する。ここで間違ったキーを押すと1ストローク目にOnSysChar関数の中で記憶した文字コードは無効になる。

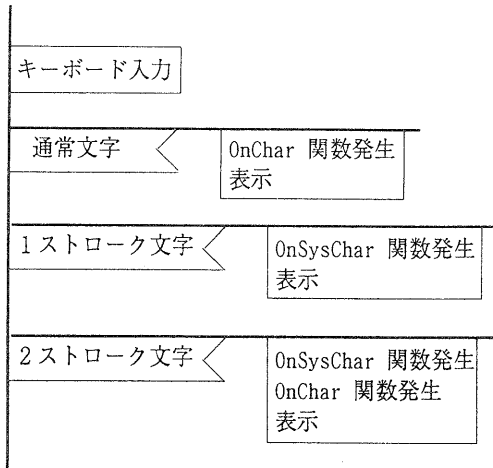


図2. キーアサインの流れ

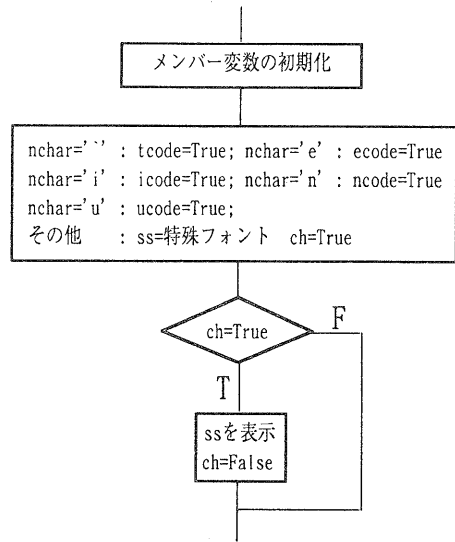


図3. OnSysChar関数.

この関数では、1 ストローク文字の表示 (ss) 又は 2 ストローク文字の判定をして、その情報 (nchar, tcode 等) を保存する。

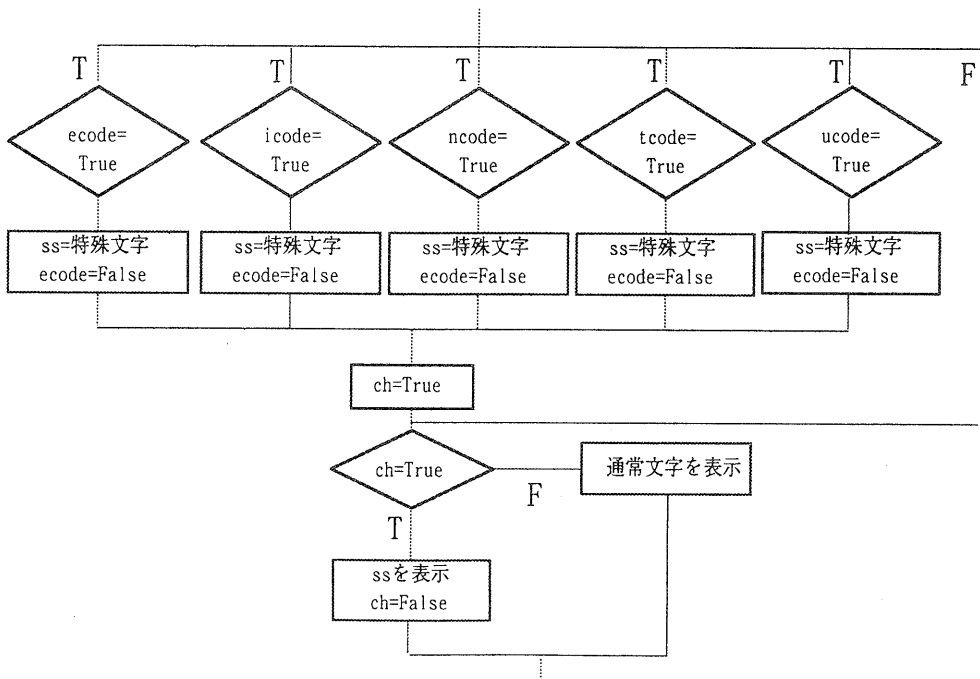


図4. OnChar関数

2 ストローク文字と通常文字の表示をする。2 ストローク文字に対しては、OnSysChar関数で獲得されたnchar, tcode等の情報を使用している。

4. 今後の課題

我々の開発したエディタの1番の目的である特殊フォント入力機能は正常に動作し、又日本語も正常に入力できるようになり、従来の問題点は解消された。しかし、EDISAN3.0には以下の問題点が残されている。

EDISAN3.0のテキストエディタは、元々WordPadをベースにして開発したため、検索・置換ダイアログの検索(置換)文字列に特殊文字を、キーボードから入力できない状態にある。そのため、検索文字列を入力するには、コピー&ペーストをしなければならず、若干煩雑である。

EDISAN3.0で付加したリンク機能(リストボックス上のある項目を選択するとその項目の場所に自動的に飛ぶ機能)は実現はできたが、以下の問題点がある。

1. リンクする項目を設定(\$マークの設定)するのが面倒である。
2. Readの状態でファイルを保存するとリンクデータが消えてしまう。
3. ファイルを開く度にリンクデータをロードしなければならない。

その他の問題点としては、ページ番号挿入機能が不完全であり、又ヘッダー・フッター、脚注、上付き文字・下付き文字等の機能がなく、論文作成エディタとしては、機能が足りないことである。

これらの問題点については、今後解決していく予定である。

文 献

- (1) 逢坂雄美, 宮尾正大, 「インターネットとパーリテキストの活用」『仏教研究』第27号(1998年3月), pp. 205-215.
- (2) 逢坂雄美, 「ウインドーズ版パーリフォント用テキストエディタ」『パーリ学仏教文化学』1998年, pp. 53-59.
- (3) デービス・チャップマン, 「3週間完全マスター Visual C++6.0」(日経BP社).
- (4) 林晴比古, 「新Visual C++6.0入門シニア編」(ソフトバンク).