

# SjisEuc ユーザーズガイド

sjiseuc コマンドは、シフト JIS のテキストファイルを変換して EUC-JP のテキストファイルに変換します。

## 下線で代用

EUC-JP は、Windows ではなく UNIX の日本語です。sjiseuc コマンドがシフト JIS から EUC-JP にテキストファイルを変換するとき、たとえば、丸付きの数字 (①②③⑩⑳)、ローマ数字 (I II III IV X)、1 バイトのカタカナは、EUC-JP で表現しにくいいため、下線に変更されます。

### 変換表を編集

たとえば、①、②、③が、それぞれ 1、2、3 で代用されるように、添付の diy コマンドは、変換表を編集できます。

## 改行

テキストファイルに関して、Windows では、0Dh および 0Ah で改行しますが、UNIX では、0Ah のみで改行します。sjiseuc コマンドは、各行から 0Dh を削除できます。

## 標準入力

パラメータが何も指定されていない場合、sjiseuc コマンドは、標準入力から行を読み込むことができます。

# インストール

SjisEuc は、32 ビット版 Windows のコンソールアプリケーションです。Windows の .NET Framework 3.5 または .NET Framework 4.0 を使用します。

## パスを通す手順の例

1. トリシーカー (<http://tori.tobihiro.jp/PDF-zip-7z.html>) というウェブサイトから sjisEUCjpCS.zip をダウンロードしてください。CS は、C# を意味します (C Sharp)。
2. sjisEUCjpCS.zip を展開してください。
3. C ドライブに sjisEUCjpCS フォルダを移動してください。
4. sjisEUCjpCS フォルダの名前を変更しないでください。
5. Windows ログキー (田キー) を押しながら Pause を押してください。
6. メニューからシステムの詳細設定を選択してください。
7. 詳細設定タブをクリックしてください。
8. [環境変数] ボタンを押してください。キーワード: システム、詳細、環境変数

9. システム環境変数として、Path が見えるまでスクロールしてください。
10. Path をクリックして選択してください。
11. [編集]ボタンを押してください。
12. Windows 10 の場合、[テキストの編集]ボタンを押してください。
13. 右向き矢印(→)キーまたはEndキーを押してください。カーソルが末尾に移動します。
14. フォルダのことをディレクトリとも言います。セミコロン( ; )は、ディレクトリどうしを区切る記号です。 ;C:¥sjisEUCjpCS を入力してください。
15. 各ウィンドウで OK ボタンを押してください。

### フォルダごと削除

アンインストール(プログラムの削除)を行うには、sjisEUCjpCS フォルダごと削除してください。Path から sjisEUCjpCS を削除してください。

## 著作権

Sogaya(そがや)は、SjisEuc の著作権を保有しています。ソースコードを添付しました。

### 免責条項

sjiseuc コマンドのご利用によって発生するいかなる損害も、Sogaya は、責任を負わないものとします。

### 構築

sjiseuc コマンドは、Visual C# 2008 Express または Visual Studio 2010 で構築されています。sjiseuc コマンドは、.NET Framework 3.5 または .NET Framework 4 を使用しません。

### ユーザーズガイド

2019 年 2 月 22 日 Revision 1.012

## sjiseuc コマンドのパラメータ

MS932~EUC-JP.bin というバイナリファイルにしたがって、sjiseuc コマンドは、シフト JIS のテキストファイルを EUC-JP のテキストファイルに変換します。

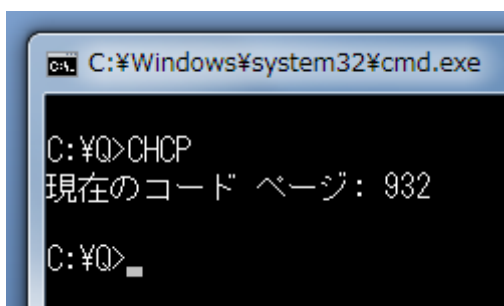
### コマンドプロンプト

Windows のコマンドプロンプトは、シフト JIS が多用されます。MS932 は、いわゆるシフ

ト JIS ですが、規格外です。たとえば、丸付きの数字(①②③⑩⑳)、ローマ数字(ⅠⅡⅢⅣⅤ)、1 バイトのカタカナは、EUC-JP で表現しにくい。

## CHCP コマンド

コマンドプロンプトの文字コードがシフト JIS に設定されていることを確認するには、パラメータを付加しないで、コマンドプロンプトに CHCP を入力して、Enter を押ししてください。932 が表示されます。MS932 が使用されることがわかります。どの Windows も最初から CHCP コマンドを用意しています。MS932 は、CP932 とも言います。



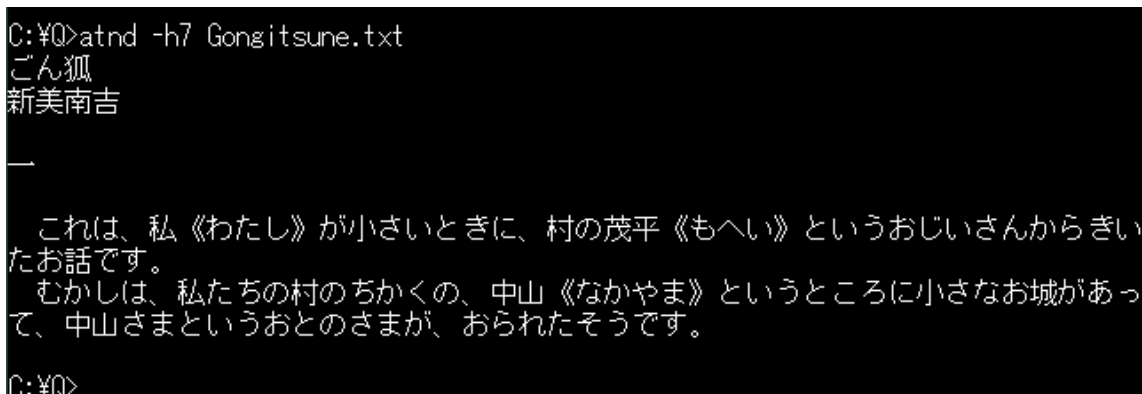
```
C:\Windows\system32\cmd.exe
C:\>CHCP
現在のコード ページ: 932
C:\>
```

## MORE コマンド

コマンドプロンプトでシフト JIS のテキストファイルを開覧する方法として、MORE コマンドが利用できます。スペースキーで進行します。Q で終了します。どの Windows も最初から MORE コマンドを用意しています。

## atnd コマンド

添付の atnd は、シフト JIS 用のコンソールアプリケーションです。シフト JIS のファイルの最初から 7 行を表示するには、コマンドプロンプトに、たとえば、下記のコマンドおよびパラメータを入力して、Enter を押ししてください。-h の直後に 7 を入力してください。-h7 は、-b6 -p7 を意味します。



```
C:\>atnd -h7 Gongitsune.txt
ごん狐
新美南吉
—
これは、私《わたし》が小さいときに、村の茂平《もへい》というおじいさんからきいたお話です。
むかしは、私たちの村のちかくの、中山《なかやま》というところに小さなお城があって、中山さまというおとのさまが、おられたそうです。
C:\>
```

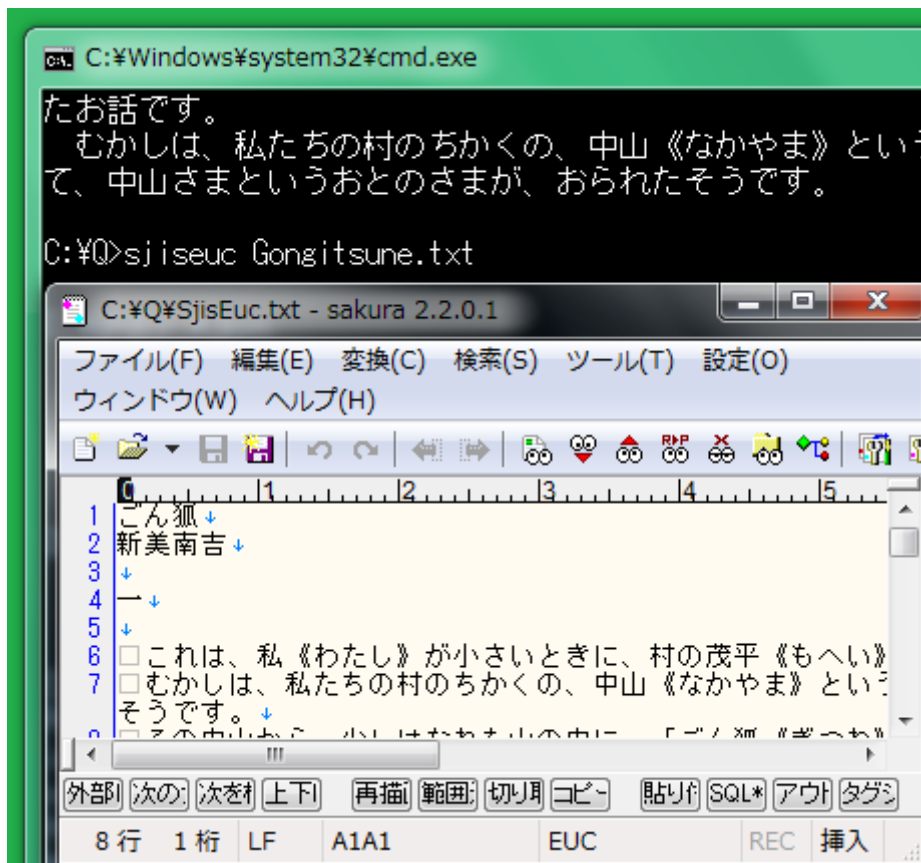
## テキストエディタ

EUC-JP をサポートしているテキストエディタとしては、EmEditor, MIFES, TeraPad, WZ, サクラエディタなどがあります。

## 変換したいファイル

sjiseuc コマンドのパラメータとして、シフト JIS のファイルを指定してください。sjiseuc コマンドは、シフト JIS のファイルを読み込んで、SjisEuc.txt という EUC-JP のファイルを書き出します。

## sjiseuc Gongitsune.txt



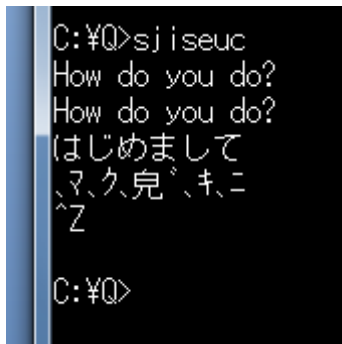
### 2 個のファイル

シフト JIS のファイル、EUC-JP のファイルがこの順に指定された場合、sjiseuc コマンドは、シフト JIS のファイルから行を読み込んで、EUC-JP のファイルに行を書き出します。

## sjiseuc Shift\_JIS.txt EUC-JP.txt

## 標準入力

sjiseuc コマンドのパラメータとして、何も指定されなかった場合、sjiseuc コマンドは、標準入力から行を読み込んで、標準出力に書き出します。下記の例では、英語を入力した場合、英語が表示されますが、日本語を入力した場合、文字化けが発生することがわかります。コマンドプロンプトは、シフト JIS を正しく表示します。EUC-JP は、文字化けが発生します。Ctrl を押しながら Z を押すことでファイルの終わり(EOF)を入力しました。^Z が表示されました。



```
C:\>sjiseuc
How do you do?
How do you do?
はじめまして
マ、ク、兒、キ、ニ
^Z
C:\>
```

### リダイレクト

標準入力から行を読み込む場合、リダイレクトでファイルを作成できます。ただし、0Dh および 0Ah で改行します。

`sjiseuc < Shift_JIS.txt > EUC-JP.txt`

### 0Dh を削除

テキストファイルに関して、Windows では、0Dh および 0Ah で改行しますが、UNIX では、0Ah のみで改行します。標準入力から読み込んで標準出力へ書き出す場合、0Dh が含まれます。さもなければ、sjiseuc コマンドは、各行から 0Dh を削除します。

#### -CRLF

0Dh および 0Ah で改行するには、パラメータとして、-CRLF を付加してください。

#### -LF

リダイレクトではなく sjiseuc コマンドでファイルを作成する場合、0Ah のみで改行するために、パラメータとして、-LF を付加できます。

### iconv コマンド

文字コードを変換するために、iconv コマンドが GNU Win32 に用意されています。インターネットで、たとえば、libiconv-1.9.2-1 などのインストーラーを検索してください。接頭辞の lib は、ライブラリを意味します。このインストーラーをダウンロードして、実行し

て、32ビット版 Windows 7 のパソコンに iconv.exe をインストールしました。

### FC コマンド

テキストファイルを比較するために、FC コマンドを試すことができます。どの Windows も最初から FC コマンドを用意しています。iconv コマンドの結果と、sjiseuc コマンドの結果をテキストファイルとして比較しました。各行が一致しました。

```
C:¥Q>iconv -f Shift_JIS -t EUC-JP Gongitsune.txt > text.tmp
C:¥Q>sjiseuc Gongitsune.txt
C:¥Q>FC text.tmp SjisEuc.txt
ファイル text.tmp と SJISEUC.TXT を比較しています
FC: 相違点は検出されませんでした
C:¥Q>
```

### COMP コマンド

バイナリファイルとして比較するには、FC コマンドのパラメータとして、/B を付加してください。Windows 2000, Windows XP, Windows Vista, Windows 7, Windows 10 では、COMP コマンドを試すことができます。iconv コマンドの結果と、-CRLF 付きの sjiseuc コマンドの結果をバイナリファイルとして比較しました。0Dh および 0Ah による改行も一致しました。

```
C:¥Q>iconv -f Shift_JIS -t EUC-JP Gongitsune.txt > binary.tmp
C:¥Q>sjiseuc -CRLF Gongitsune.txt
C:¥Q>COMP binary.tmp SjisEuc.txt
binary.tmp と SjisEuc.txt を比較しています...
ファイルに違いはありません
(ほかのファイルを比較しますか (Y/N)? n
C:¥Q>
```

## バージョン番号

sjiseuc コマンドのバージョン番号を確認するには、コマンドプロンプトに下記のコマンドおよびパラメータを入力して、Enter を押してください。コマンドおよびパラメータの例も表示されます。

# sjiseuc -V

```
C:¥Q>sjiseuc -V
SjisEuc C# Version 1.02

SjisEuc Shift_JIS.txt
SjisEuc Shift_JIS.txt EUC-JP.txt
SjisEuc Shift_JIS.txt -w
SjisEuc -CRLF Shift_JIS.txt
SjisEuc -CRLF Shift_JIS.txt EUC-JP.txt
SjisEuc < Shift_JIS.txt > EUC-JP.txt
SjisEuc -LF Shift_JIS.txt
SjisEuc -LF Shift_JIS.txt EUC-JP.txt
SjisEuc -V
SjisEuc -VC

C:¥Q>
```

### 短い説明

コマンドおよびパラメータの各行に短い説明を付与して表示するには、コマンドプロンプトに下記のコマンドおよびパラメータを入力して、Enter を押してください。

## sjiseuc -VC

### 実行ファイル

変換表の名前は、MS932~EUC-JP.bin です。sjiseuc コマンドの特徴として、実行ファイル(SjisEuc.exe)の外部から変換表を読み込むことができます。実行ファイルは、自分が**インストールされたディレクトリ**に存在している変換表を読み込みます。実行ファイルのパス名は、以下のとおりです。

## C:¥sjisEUCjpCS¥SjisEuc.exe

### カレントディレクトリ

変換表および実行ファイルが同じディレクトリに存在していない場合、カレントディレクトリから変換表を読み込みます。

### Path に記載の各ディレクトリ

変換表を開くことができるまで、sjiseuc コマンドは、下記の順に変換表を探索します。

1. インストールされたディレクトリ
2. カレントディレクトリ
3. 環境変数(Path)に記載の各ディレクトリ
4. 実行ファイルが起動しているディレクトリ

### どの変換表を使用するか

過去に実行ファイルがインストールされたディレクトリか、いま実行ファイルが起動しているディレクトリか、Path にあるディレクトリか、カレントディレクトリのいずれかに存在している変換表を開くことができたとき、どの変換表を使用するか表示するには、sjiseuc コマンドのパラメータとして、-w を付加してください。パス名ではなく単なるファイル名が表示された場合、カレントディレクトリの変換表を使用します。

## sjiseuc Gongitsune.txt -w

```
C:¥Q>sjiseuc Gongitsune.txt -w
C:¥sjisEUCjpcS¥MS932~EUC-JP.bin is primarily accessible to SjisEuc.
C:¥Q>
```

## 新しいファイル

COPY コマンドで、シフト JIS の新しいテキストファイルを作成するには、コピーされるファイルとして CON を指定してください。この場合、CON は、パソコンのキーボードです。コマンドプロンプトは、CON がファイルであるかのように COPY コマンドを実行します。

### かな漢字変換

コマンドプロンプトで、かな漢字変換を行うには、Alt を押しながら[半角／全角]キーを押してください。Windows 7, Windows 10 の場合、Alt を押す必要がありません。かな漢字変換モードから脱出するには、[半角／全角]キーを押してください。

#### 消去

コマンドプロンプトの画面を消去するには、コマンドプロンプトに下記のコマンドを入力して、Enter を押してください。

## CLS

### ファイルの終わり

コマンドプロンプトに、ファイルの終わり(EOF)を入力するには、Ctrl を押しながら、Z を押してください。Ctrl + Z は、この操作を意味します。^Z が画面に表示される場合があります。

### 中止したい合図

コマンドプロンプトに、中止したい合図を入力するには、Ctrl を押しながら、C を押してください。Ctrl + C は、この操作を意味します。^C が画面に表示される場合があります。



す。

## 変換表を編集する

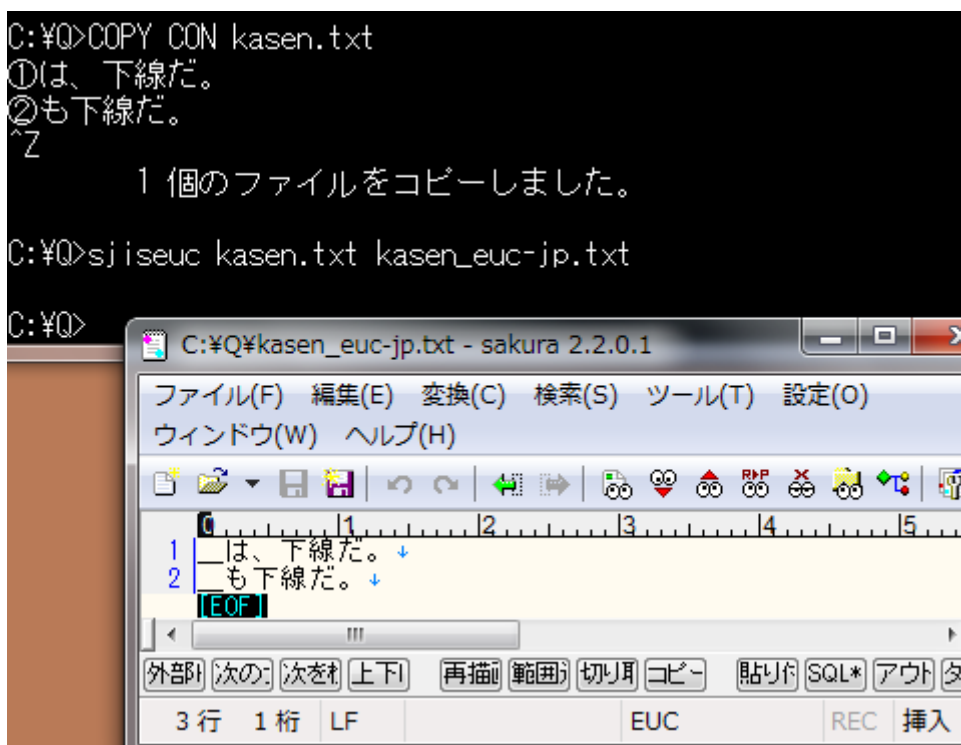
変換表の名前は、MS932~EUC-JP.bin です。通常、変換表および実行ファイルは、同じディレクトリに存在しています。

### diyo コマンド

添付の diyo コマンドは、MS932~EUC-JP.bin というバイナリファイルを編集できます。

### すべて下線になる

EUC-JP は、Windows ではなく UNIX の日本語です。sjiseuc コマンドがシフト JIS から EUC-JP にテキストファイルを変換するとき、丸付きの数字(①②③⑩⑰)、ローマ数字(I II III IV X)などの機種依存文字は、すべて下線(   )に変更されます。1 バイトのカタカナは、ASCII の下線(   )に変更されます。



### 1 文字ずつ編集

たとえば、①を下線ではなく1で代用するには、コマンドプロンプトに下記のコマンドおよびパラメータを入力して、Enter を押してください。

```
diyo -s① -e 1
```

## どの変換表を編集するか

diyo コマンドがアクセスできる変換表のパス名を表示するには、コマンドのパラメータとして、`-w` を付加してください。通常の文字および代用文字のシフト JIS 文字コードが十六進数で表示されます。

# diyo -s① -e1 -w

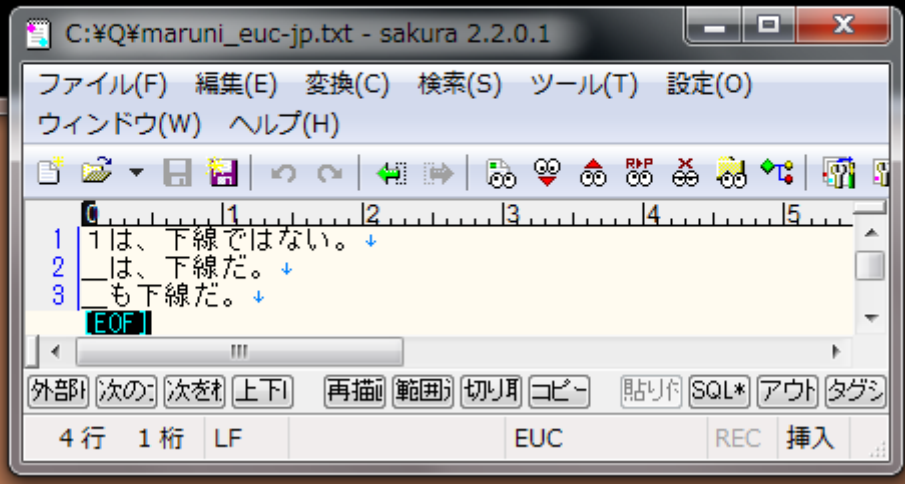
```
C:\>diyo -s① -e1 -w
C:\>dir | FINDSTR "MS932"
2019/02/21 00:11          31,528 MS932~EUC-JP.bin

C:\>copy MS932*.bin C:\sjisEUCjpCS /Y
MS932~EUC-JP.bin
    1 個のファイルをコピーしました。

C:\>copy CON maruni.txt
①は、下線ではない。
②は、下線だ。
③も下線だ。
^Z
    1 個のファイルをコピーしました。

C:\>sjiseuc maruni.txt maruni_euc-jp.txt

C:\>
```



## 1 バイトのカタカナ

コマンドプロンプトでは、1 バイトのカタカナの幅は、漢字の幅の半分になります。1 バイトのカタカナを半角カタカナとも言います。半角カタカナの代用文字として、ASCII の下線 ( \_ ) が sjiseuc コマンドに使用されます。sjiseuc コマンドが、たとえば、ナの代用

として ASCII の下線ではなく N を使用するように変換表を編集するには、コマンドプロンプトに下記のコマンドおよびパラメータを入力して、Enter を押してください。k およびナの間には空白を挿入しないでください。

## diyo -k ナ -aN

パラメータ	用語	目的	説明
-a	ASCII	半角カタカナを廃止	1 バイトの英字、数字、記号
-e	EUC-JP	機種依存文字を廃止	UNIX の日本語
-k	katakana		半角カタカナ
-s	Shift_JIS		Windows の日本語、MS932

## バージョン番号

diyo コマンドのバージョン番号を表示するには、コマンドプロンプトに下記のコマンドおよびパラメータを入力して、Enter を押してください。

## diyo -V

```
C:¥Q>diyo -V
Diyo C# Version 1.02

Diyo -f bars.txt
Diyo -kナ -aN
Diyo -kナ -aN file
Diyo -s① -e1
Diyo -s① -e1 file
Diyo -s① -e1 -w
Diyo -V
Diyo -VC
C:¥Q>
```

### 短い説明

短い説明を表示するには、コマンドプロンプトに下記のコマンドおよびパラメータを入力して、Enter を押してください。

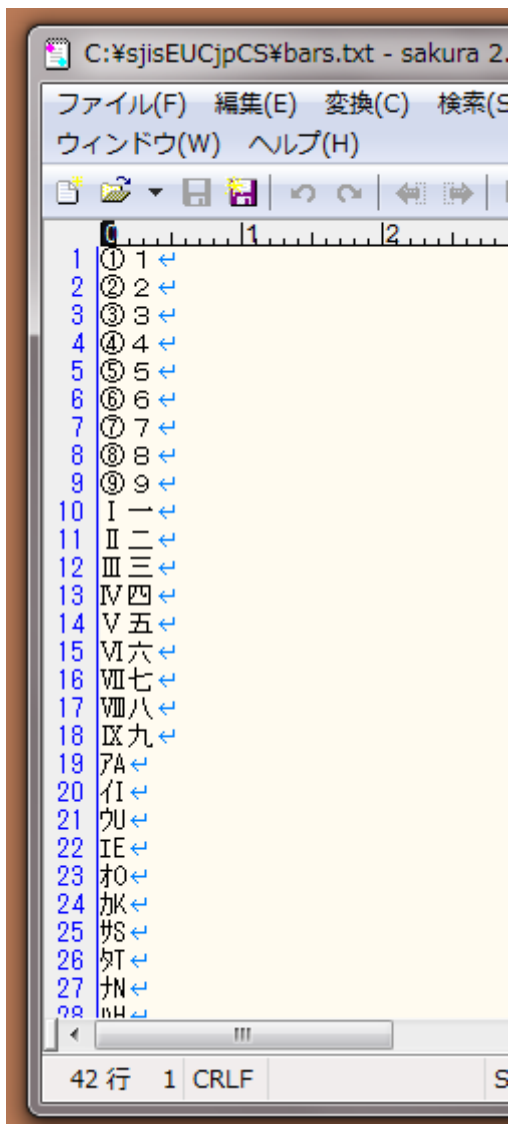
## diyo -VC

## 代用文字のリスト

事前に利用者が代用文字のリストを用意した場合、diyo コマンドは、代用文字のリストにしたがって、一度に多数の文字を編集できます。

### 通常文字および代用文字

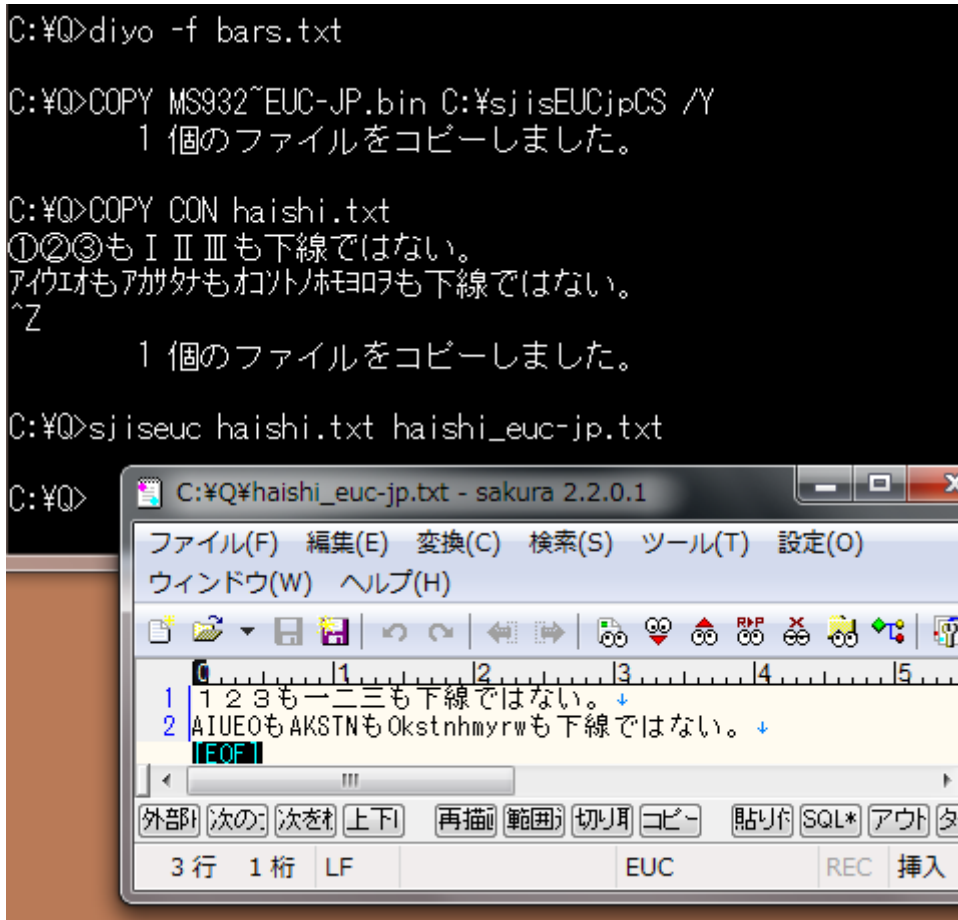
diyo コマンドは、代用文字リストのファイルから 906 行まで読み込むことができます。代用文字リストのファイルの各行は、通常文字と、代用文字とからなります。代用文字もシフト JIS で記入してください。



### ファイル名を指定

利用者が代用文字リストのファイルを用意した場合、diyo コマンドのパラメータとして、`-f` を付加して、空白で区切って、さらにファイル名を付加してください。

**diyo -f bars.txt**



## 1 バイトの文字

1 バイトの文字としては、ASCII および半角カタカナがあります。

!	"	#	\$	%	&	'	(	)	*	+	,	-	.	/
0	1	2	3	4	5	6	7	8	9	:	;	<	=	>
@	A	B	C	D	E	F	G	H	I	J	K	L	M	N
O	P	Q	R	S	T	U	V	W	X	Y	Z	[	\	]
^	_	`	a	b	c	d	e	f	g	h	i	j	k	l
m	n	o	p	q	r	s	t	u	v	w	x	y	z	{
	}	~	。	「	」	、	・	ヲ	ア	イ	ウ	エ	オ	ヤ
ユ	ヨ	ッ	ー	アイ	ウエ	オカ	キク	ケコ	サシ	スセ	ソ	タチ	ツテ	トナ
ニ	ヌ	ネ	ノ	ハ	ヒ	フ	ヘ	ホ	マ	ミ	ム	モ	ヤ	ユ
ヨ	ラ	リ	ル	レ	ロ	ワ	ン	。						

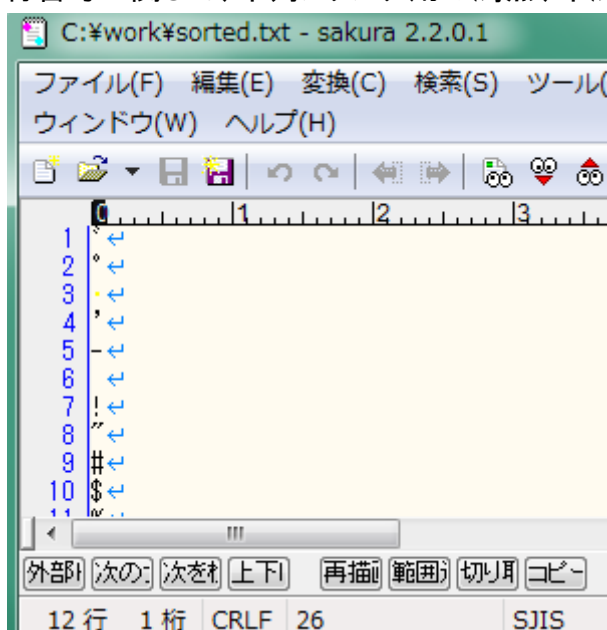
### 半角カタカナ

1 バイトのカタカナは、シフト JIS に含まれています。コマンドプロンプトの表示では、1 バイトのカタカナは、幅が漢字の半分になります。

### SORT コマンド

どの Windows も最初から SORT コマンドを用意しています。上記の 1 バイト文字を各

行に 1 個のみ記入したファイルの行を並べ替えてみました。結果として、sorted.txt の行番号に関して、半角カタカナ用の濁点、半濁点が若い行番号に来ました。



## 並べ替え

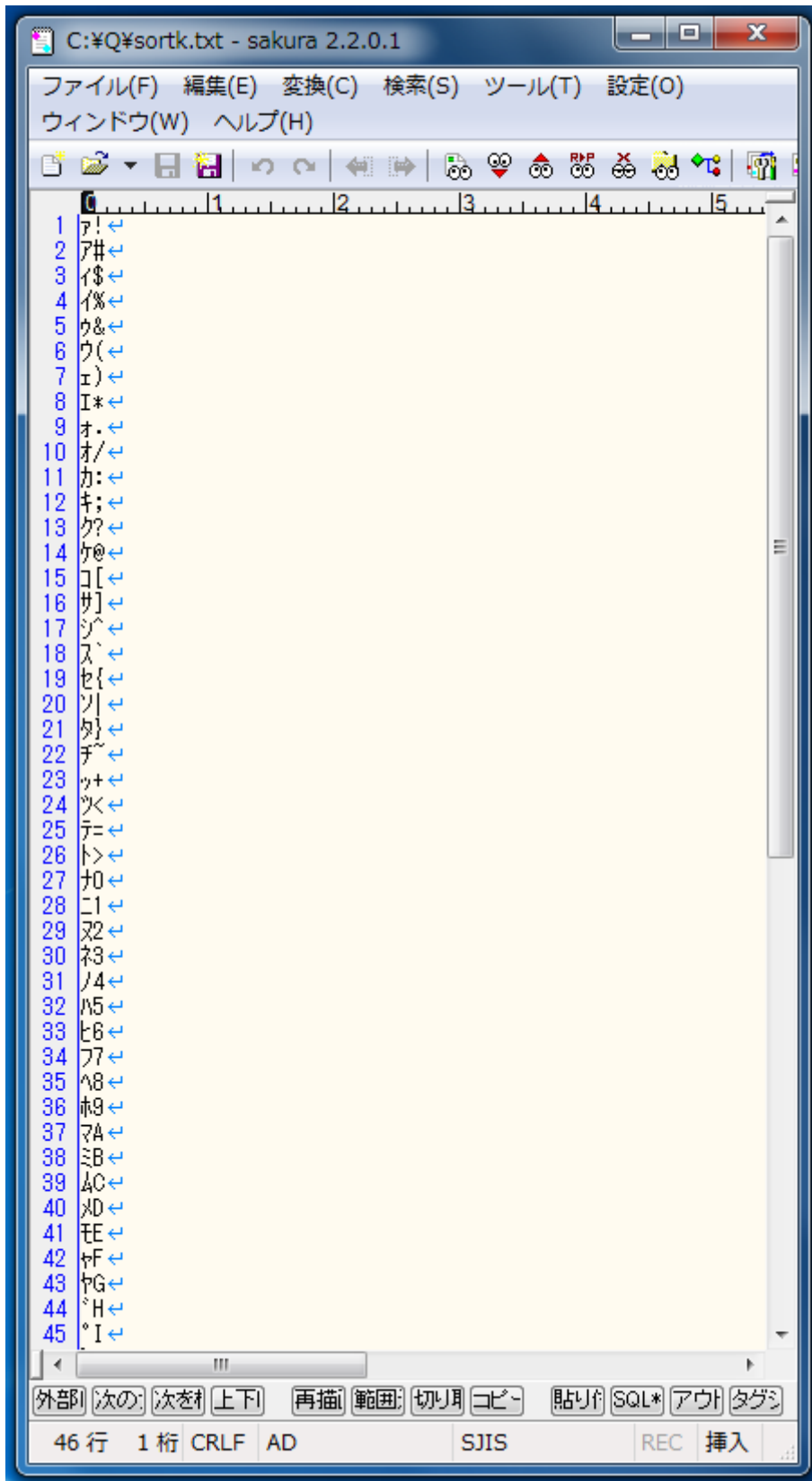
Windows にも UNIX にも、sort コマンドが用意されています。並べ替えに関して、シフト JIS の結果と、EUC-JP の結果が一致しないおそれがあります。上記の bars.txt では、EUC-JP で並べ替えた結果が昇順になりません。代用文字リストのファイルを作成するとき、並べ替えも考慮したい。

## ほぼ同等の位置

機種(キシユ)、著作権(チヨサクケン)など、ヤ行の拗音を表記する小さいカタカナは、直音を表記するカタカナとほぼ同等の位置に並べ替えられます。実行可能(ジッコウカノウ)、接頭辞(セツウジ)など、促音は、直音を表記するカタカナとほぼ同等の位置に並べ替えられます。アルファベットの大文字と小文字を区別しないソフトウェアがあります。

### 無視

合図(アイズ)、英語(エイゴ)などの濁点、添付(テンプ)などの半濁点を無視するソフトウェアがあります。昇順になる代用文字リストを作成するには、並べ替えを行う各ソフトウェアで試行錯誤が必要です。C#版 sortk コマンド用の代用文字リスト(sortk.txt)



## sortk コマンド

添付の sortk コマンドは、テキストファイルの行を並べ替えて、標準出力に書き出します。コマンドプロンプトでは、シフト JIS は、正しく表示されますが、EUC-JP は、文字化けが発生します。ファイルにリダイレクトすることで、適切なテキストエディタで閲覧できます。

### CSV

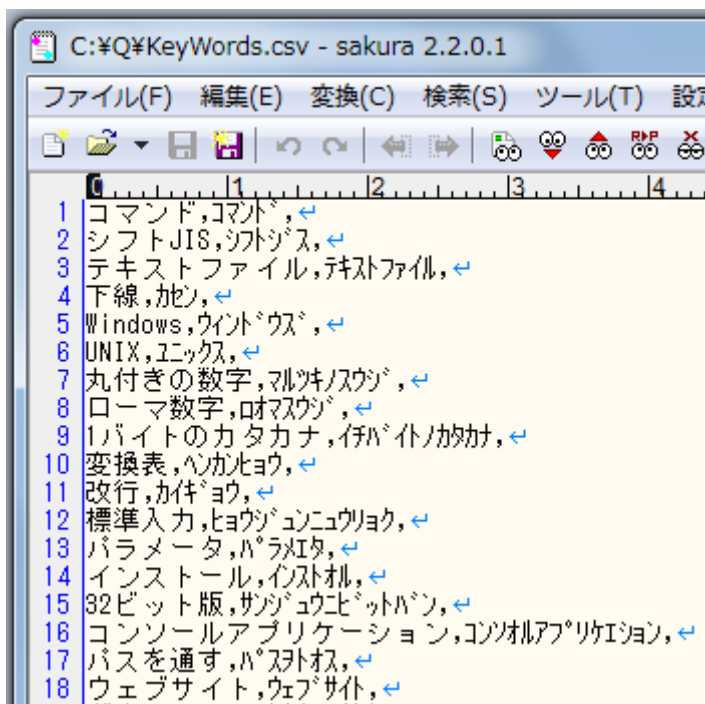
拡張子が.csv であるテキストファイルは、各行が、コンマで分離された多数の語句からなります。CSV のファイルは、Excel などの表計算ソフトウェアで容易に表示できません。

### 区切り文字

CSV の行を並べ替えるには、コマンドプロンプトに sortk コマンドのパラメータとして、-t.(マイナステーコンマ)を付加してください。-t およびコンマの間に空白を挿入しないでください。

### 読み方

キーワードの読み方を KeyWords.csv の第 2 欄に記入しました。



### 濁点を無視

半角カタカナは、清音のカナおよび濁点を組み合わせることで濁音を表現しますが、清音および濁音は、ほぼ同等の位置になるように並べ替えたい。sortk コマンドが、濁点、半濁点を無視するように、sortk コマンドのパラメータとして、-c-を付加してください。diyo コマンドで、半角カタカナの代用として ASCII を使用している場合、さらに、-h



を付加してください。

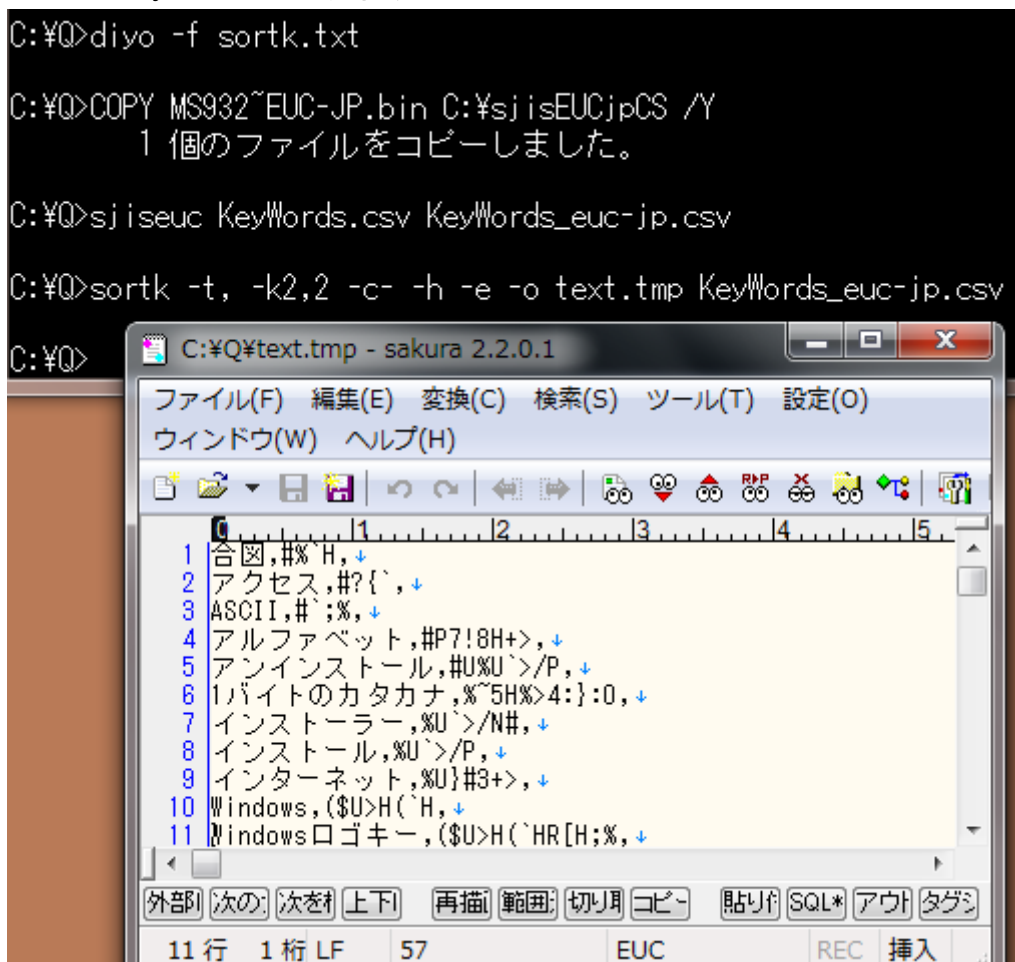
### EUC-JP のファイルを読み込める

EUC-JP のテキストファイルを読み込むには、コマンドのパラメータとして、`-e` を付加してください。UTF-8 のテキストファイルを読み込むには、コマンドのパラメータとして、`-8` を付加してください。新しいファイルを作成するには、`-o text.tmp` を付加してください。`text.tmp` は、新しいファイルの名前です。

### 第 2 欄を比較

`sortk` コマンドは、たとえば、第 2 欄にある語句を比較して、CSV のテキストファイルの行を並べ替えることができます。第 2 欄にある語句を比較するには、`sortk` コマンドのパラメータとして、`-k2,2` を付加してください。`-k` および数字の間に空白を挿入しないでください。`diyo` コマンドが変換表を編集した例

```
C:¥Q>diyo -f sortk.txt
C:¥Q>COPY MS932~EUC-JP.bin C:¥sjisEUCjpCS /Y
1 個のファイルをコピーしました。
C:¥Q>sjiseuc Keywords.csv Keywords_euc-jp.csv
C:¥Q>sortk -t, -k2,2 -c- -h -e -o text.tmp Keywords_euc-jp.csv
C:¥Q>
```



```
1 合図, #H, ↓
2 アクセス, #?{\, ↓
3 ASCII, #%;%, ↓
4 アルファベット, #P7!8H+>, ↓
5 アンインストール, #U`>/P, ↓
6 1バイトのカタカナ, #~5H%>4:}:0, ↓
7 インストーラー, #U`>/N#, ↓
8 インストール, #U`>/P, ↓
9 インターネット, #U}#3+>, ↓
10 Windows, ($U>H(`H, ↓
11 Windows ロゴキー, ($U>H(`HR[H;%, ↓
```

## 五十音順

まず、`sjiseuc` コマンドで、シフト JIS から EUC-JP に変換して、次に、`sortk` コマンドで、第 2 欄に関して並べ替えました。たしかに五十音順に並べ替えることができました。

## 0Ah のみで改行

EUC-JP のテキストファイルを並べ替えた結果として、text.tmp という EUC-JP のテキストファイルを上に図示しました。0Ah のみで改行しています。サクラエディタのウィンドウの下部にあるステータスバーに LF が表示されます。サクラエディタの改行文字は、左向きではなく下向きの矢印になります。

## バージョン番号

sortk コマンドのバージョン番号を確認するには、コマンドのパラメータとして、-V を付加してください。

### コマンドのパラメータを説明

sortk コマンドのパラメータに関する短い説明を表示するには、コマンドプロンプトに、下記のコマンドおよびパラメータを入力して、Enter を押してください。

## sortk -VC

```
C:\>sortk -VC
Sortk C# Version 1.02

Sortk -t, < csv > new      # Read csv to produce a new file.
Sortk -t, -o new csv      # Read csv to produce a new file.
Sortk -t, -j2 csv         # Compare phrases in Field-2.
Sortk -t, -j2 -g csv      # Compare 0.12 with 0.78 in Field-2.
Sortk -t, -j2 -n csv      # Compare 3456 with 9012 in Field-2.
Sortk -t, -k2,2 csv       # Compare phrases in Field-2.
Sortk -t, -k2,3 csv       # Compare phrases from Field-2 to Field-3.
Sortk -t, -k2 csv         # Compare phrases from Field-2 to finish.
Sortk -t, -k2,2 -! csv    # Count ticks.
Sortk -t, -k2,2 -8 csv    # UTF-8
Sortk -t, -k2,2 -b csv    # Skip 20h in the first part of Field-2.
Sortk -t, -k2,2 -c- csv   # Skip combining characters.
Sortk -t, -k2,2 -e csv    # EUC-JP
Sortk -t, -k2,2 -h csv    # H and I are omitted from ASCII characters.
Sortk -t, -k2,2 -r csv    # Arrange in reverse order.
Sortk -t¥| -k2,2 file     # Wiki table
Sortk -t¥x7C -k2,2 file   # Wiki table
Sortk -V                  # version
Sortk -VC                 # version with comments

C:\>
```

パラメータ	ほぼ同等な指定	説明
-!		パルスをカウントして経過時間を表現する。
-8		UTF-8 のテキストファイルを読み込む。
-b		比較開始欄の前半にある空白を無視する。
-c-		半角カタカナの濁点、半濁点を無視する。
-e		EUC-JP のテキストファイルを読み込む。
-g		浮動小数点数を比較する。
-h	-p	濁点の代用文字として、H を無視する。
-j2	-k2, 2	第 2 欄のみ比較する。
-k2		第 2 欄から最終欄まで比較する。
-k2, 3		第 2 欄から第 3 欄まで比較する。
-n		整数を比較する。
-p	-h	濁点の代用文字として、H を無視する。
-o	>で、リダイレクト	新しいファイルを命名する。
-r		降順で並べ替える。
-t,	.csv の拡張子	区切り文字としてコンマを使用する。
-t¥1	-t¥x7C	Wiki の表の区切り記号として縦線を使用する。
-t¥x7C	-t¥1	Wiki の表の区切り記号として縦線を使用する。
-V		バージョン番号および用法を表示する。
-VC		パラメータの短い説明を表示する。

## パラメータからコンマを排除

sortk コマンドのパラメータとして、-t,(マイナスティーコンマ)も-k2,3(マイナスケーニコマサン)もコンマが含まれていますが、コンマの使用を回避できます。

### 十六進数で入力

コンマの十六進数は、2Ch です。-t,ではなく-t¥x2C が指定されても、sortk コマンドは、区切り文字をコンマに設定します。

### .csv の拡張子

何も指定されなければ、sortk コマンドは、区切り文字を空白(20h)に設定します。行が並べ替えられるテキストファイルの拡張子が.csv である場合、sortk コマンドは、、区切り文字をコンマに設定します。

### 範囲を示すハイフン

コンマではなくハイフンを使用できます。-k2,3 ではなく-k2-3 が指定されても、sortk コマンドは、第 2 欄から第 3 欄までにある語句を比較します。

```

C:¥Q>sortk -t, -k10,20 -o text.tmp Field-20.csv
C:¥Q>sortk -t¥x2C -k10-20 -o yen_x2C.csv Field-20.csv
C:¥Q>sortk -k10-20 -o extension.csv Field-20.csv
C:¥Q>FC text.tmp yen_x2C.csv
ファイル text.tmp と YEN_X2C.CSV を比較しています
FC: 相違点は検出されませんでした

C:¥Q>FC text.tmp extension.csv
ファイル text.tmp と EXTENSION.CSV を比較しています
FC: 相違点は検出されませんでした

C:¥Q>

```

### 経過時間を比較

ミリ秒ではなく計数で経過時間を表現するには、sortkコマンドのパラメータとして、-!を付加してください。sortkコマンドは、C#のStopwatchクラスを利用します。sjisutf8 コマンドについては、[後述します](#)。

```

C:¥Q>REM 3-byte katakana (UTF-8)
C:¥Q>sjisutf8 KeyWords.csv 8.csv
C:¥Q>sortk -t, -k2,2 -8 -! -o 3-byte_8s.csv 8.csv
19771 ticks
C:¥Q>REM 1-byte ASCII (EUC-JP)
C:¥Q>sjiseuc KeyWords.csv e.csv
C:¥Q>sortk -t, -k2,2 -e -c -h -! -o 1-byte_es.csv e.csv
19246 ticks

```

### obot コマンド

添付の obot コマンドは、ファイルの冒頭を十六進数で表現します。たとえば、テキストファイルの 5 行ぶんを表示するには、obot コマンドのパラメータとして -m5 を付加してください。まず、GNU Win32 の iconv コマンドで、MS932 のテキストファイルを UTF-8 の一時ファイルに変換しました。次に、obot コマンドで、一時ファイルを読み込んで、十六進数で表現しました。半角カタカナが 1 バイトではなく 3 バイトになったことがわかります。

### 丸付きの数字を変換

リダイレクトで一時ファイルを作成したため、0Dh および 0Ah で改行されています。丸付きの数字を変換するには、iconv コマンドのパラメータとして、Shift\_JIS ではなく CP932 を iconv に入力します。

```
C:¥Q>COPY CON marunana.txt
ナ
ナ
菜
⑦
VII
Z

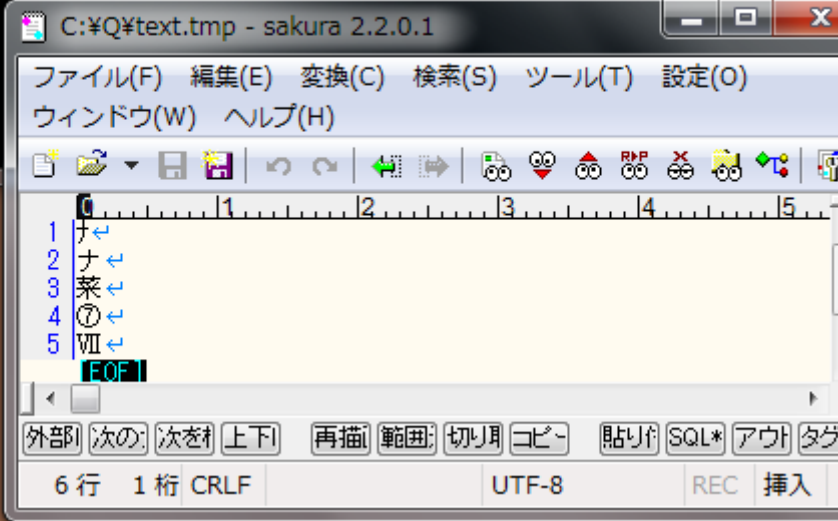
1 個のファイルをコピーしました。

C:¥Q>obot -m5 marunana.txt
C5 0D 0A 83 69 0D 0A 8D D8 0D 0A 87 46 0D 0A 87 5A 0D 0A

C:¥Q>iconv -f CP932 -t UTF-8 marunana.txt > text.tmp

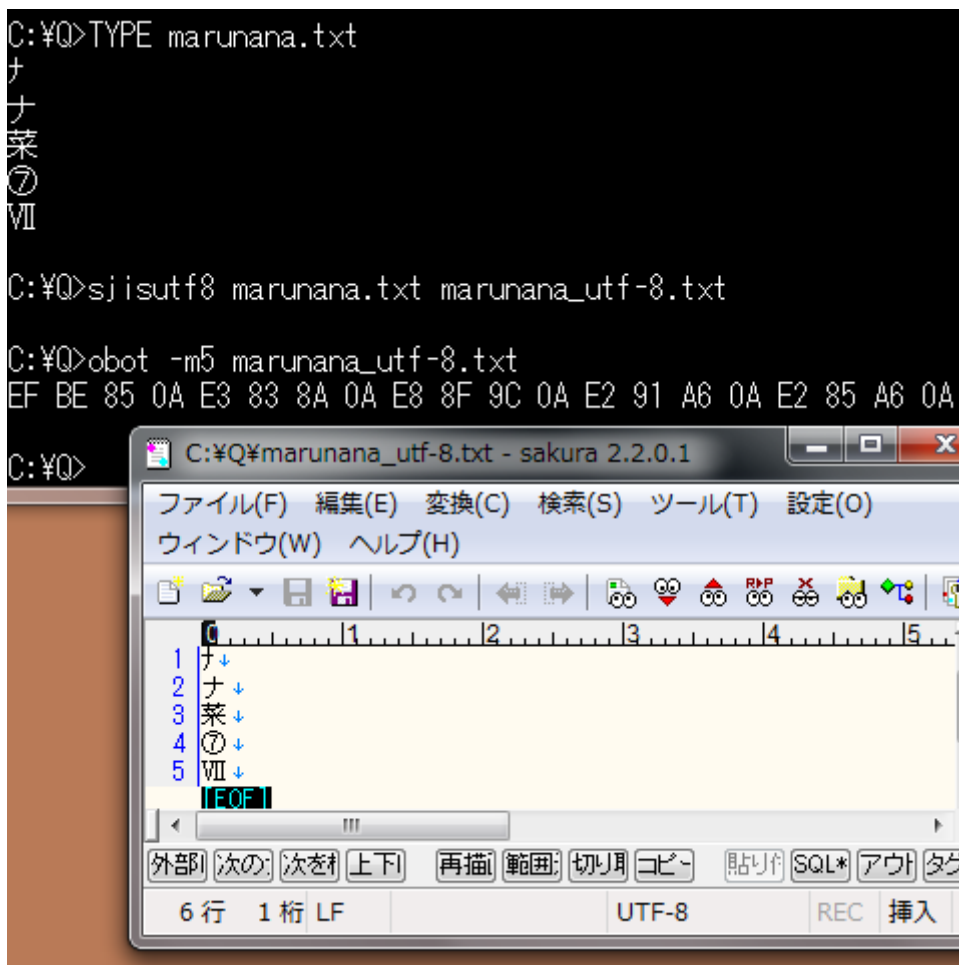
C:¥Q>obot -m5 text.tmp
EF BE 85 0D 0A E3 83 8A 0D 0A E8 8F 9C 0D 0A E2 91 A6 0D 0A
E2 85 A6 0D 0A

C:¥Q>
```



### sjisutf8 コマンド

トリシーカー (<http://tori.tobiipro.jp/PDF-zip-7z.html>) というウェブサイトから sjisUTF8.zip をダウンロードできます。sjisutf8 コマンドは、MS932 のテキストファイルを UTF-8 のテキストファイルに変換します。iconv コマンドと異なり、sjisutf8 コマンドは、0Ah のみで改行できます。半角カタカナが 1 バイトではなく 3 バイトになったことがわかります。



## hanzen コマンド

添付の hanzen コマンドは、シフト JIS のテキストファイルを読み込んで、半角カタカナを全角カタカナに置換して、シフト JIS のテキストファイルに書き出します。

### 清音および濁点から濁音に置換

半角カタカナの濁点があれば、hanzen コマンドは、直前にあるカナと濁点を結合して、濁音のカタカナに置換します。たとえば、hanzen コマンドは、ガを、ガ`ではなくガに置換します。

### 清音および半濁点から半濁音に置換

半角カタカナの半濁点があれば、hanzen コマンドは、直前にあるカナと半濁点を結合して、半濁音のカタカナに置換します。たとえば、hanzen コマンドは、

パを、パ`ではなくパに置換します。

### ファイルのサイズ

bars.txt にしたがって、diyo コマンドは、たとえば、ナを N に置換しますし、①を 1 に置換します。すなわち、1 バイトのカタカナを 1 バイトの英字に置換しますし、2 バイトの機種

依存文字を 2 バイトの文字に置換します。diyو コマンドの結果として、ファイルのサイズは、変化しません。sortk.txt にしたがって、diyو コマンドは、たとえば、ヤを G に置換します。ヌを 2 に置換します。セを{|に置換します。すなわち、1 バイトのカタカナを 1 バイトの ASCII(英字、数字、記号)に置換します。やはり、ファイルのサイズは、変化しません。diyウ コマンドと異なり、hanzen コマンドは、1 バイトのカタカナを 2 バイトのカタカナに置換します。hanzen コマンドの結果として、ファイルのサイズが増大します。カタカナは、1 バイトではなく 2 バイトになったことがわかります。0Dh および 0Ah で改行されています。

```
C:¥Q>TYPE marunana.txt
ナ
ナ
菜
㊦
VI

C:¥Q>hanzen marunana.txt marunana_zen.txt

C:¥Q>obot -m5 marunana_zen.txt
83 69 0D 0A 83 69 0D 0A 8D D8 0D 0A 87 46 0D 0A 87 5A 0D 0A

C:¥Q>TYPE marunana_zen.txt
ナ
ナ
菜
㊦
VI

C:¥Q>■
```

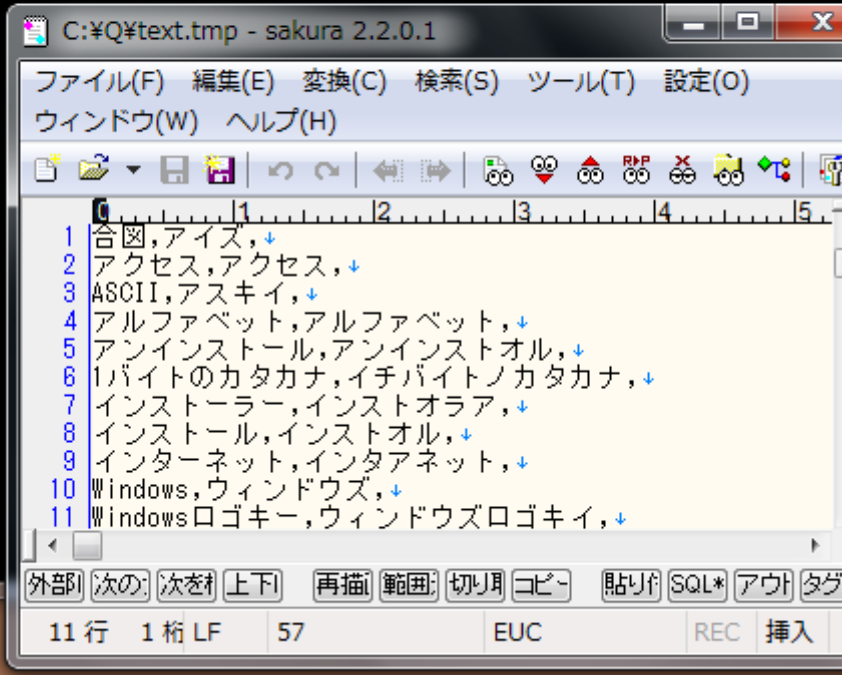
### キーワードの読み方の欄

キーワードの読み方を半角カタカナで入力した欄は、hanzen コマンドで、全角カタカナに変更できます。

### 全角カタカナで並べ替える

まず、hanzen コマンドで、Keywords.csv にある半角カタカナを全角カタカナに置換しました。次に、sjiseuc コマンドで、全角カタカナのファイルを読み込んで、EUC-JP のテキストファイルを作成しました。さらに、キーワードの読み方の欄(第 2 欄)に関して行を並べ替えました。五十音順に並べ替えることができました。

```
C:¥Q>hanzen KeyWords.csv KeyWords_zen.csv
C:¥Q>sjiseuc KeyWords_zen.csv KeyWords_zen_euc-jp.csv
C:¥Q>sortk -t, -k2,2 -e -o text.tmp KeyWords_zen_euc-jp.csv
C:¥Q>
```



```
1 合図,アイス,+
2 アクセス,アクセス,+
3 ASCII,アスキー,+
4 アルファベット,アルファベット,+
5 アンインストール,アンインストール,+
6 1バイトのカタカナ,イチバイトノカタカナ,+
7 インストーラー,インストオラア,+
8 インストール,インストール,+
9 インターネット,インタアネット,+
10 Windows,ウィンドウズ,+
11 Windowsロゴキー,ウィンドウズロゴキイ,+
```

## narfi コマンド

添付の narfi コマンドは、CSV のテキストファイルを読み込んで、選択された欄を新しいファイルに書き出します。新しいファイルの名前は、narfi.txt です。利用者は、新しいファイルを命名できます。

### 短い表

シフト JIS のファイルの最初から 5 行を表示するには、atnd コマンドのパラメータとして、-h5 を付加してください。-h5 は、-b4 -p5 を意味します。リダイレクトすることで、短い CSV ファイルを作成できます。

### 細長い表

narfi コマンドは、幅が広いテキストファイルを細長いファイルに縮小できます。たとえば、21 欄もある CSV で、まず、第 1 欄から第 5 欄までを選択するには、narfi コマンドのパラメータとして、-c5 を付加してください。次に、第 2 欄を選択から除外するには、narfi コマンドのパラメータとして、-b10111 を付加してください。欄を選択することの真偽を示すために、1 または 0 を入力してください。10111 のうち 0 は、第 2 欄を選択しないことを意味します。結果として、4 欄を書き出します。すなわち、4 列の表になります。



ただし、-b10111 が指定されない場合、atnd コマンドは、-b11111 が指定されたとみなして、第 2 欄も書き出します。

```
C:¥Q>atnd -h5 Field-21.csv
1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21
85,52,86,48,68,22,59,44,63,61,92,37,09,83,55,58,51,31,60,02,
18,53,10,36,22,69,60,98,02,20,39,25,29,44,99,60,60,26,03,73,
93,15,38,47,92,70,49,02,38,69,29,09,21,27,86,89,73,25,31,47,
32,43,69,31,69,38,13,47,18,60,16,13,30,66,02,80,88,30,92,30,

C:¥Q>narfi -b10111 -c5 Field-21.csv

C:¥Q>atnd -h5 narfi.txt
1,3,4,5,
85,86,48,68,
18,10,36,22,
93,38,47,92,
32,69,31,69,

C:¥Q>
```

-b10111					
-b	1	0	1	1	1
<b>真偽</b>	第 1 欄選択	第 2 欄を除外する	第 3 欄選択	第 4 欄選択	第 5 欄選択

### バージョンおよびコマンドのパラメータ

narfi コマンドのバージョン番号、コマンドの例、短い説明を表示するには、narfi コマンドのパラメータとして、-VC を付加してください。

```
C:¥Q>narfi -VC
Narfi Version 1.02

Narfi -b10111 wide # Read wide to produce Narfi.txt
Narfi -b10111 wide narrow # Read wide to produce a narrow file.
Narfi -b10111 < wide > narrow # Read wide to produce a narrow file.
Narfi -c2 wide narrow # 2 fields
Narfi -c2 -8 wide narrow # UTF-8
Narfi -c2 -e wide narrow # EUC-JP
Narfi -c2 -m wide narrow # MS932
Narfi -q wide narrow # Commas disapper if placed in quotes.
Narfi -s, wide narrow # CSV
Narfi -s¥l wide narrow # Wiki table
Narfi -s¥x7C wide narrow # Wiki table
Narfi -V # version
Narfi -VC # version with comments
```

### 読み方の欄を削除

もう五十音順に並べ替えることができたので読み方の欄を削除したい場合、narfi コマ

ンドのパラメータとして、たとえば、-b101 を付加してください。

```
C:¥Q>diyo -f sortk.txt  
C:¥Q>COPY MS932~EUC-JP.bin C:¥sjisEUCjpCS /Y  
1 個のファイルをコピーしました。  
C:¥Q>sjiseuc KeyWords.csv KeyWords_euc-jp.csv  
C:¥Q>sortk -t, -k2,2 -c- -h -e -o sorted.csv KeyWords_euc-jp.csv  
C:¥Q>narfi -b101 -e sorted.csv KeyWords_narrow.csv  
C:¥Q>
```

