



日本語入力方式の概要と移行

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95054
U.S.A.

Part No: 816-3950-12
2003 年 8 月

Copyright 2003 Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 U.S.A. All rights reserved.

本製品およびそれに関連する文書は著作権法により保護されており、その使用、複製、頒布および逆コンパイルを制限するライセンスのもとにおいて頒布されます。サン・マイクロシステムズ株式会社の書面による事前の許可なく、本製品および関連する文書のいかなる部分も、いかなる方法によっても複製することが禁じられます。

本製品の一部は、カリフォルニア大学からライセンスされている Berkeley BSD システムに基づいていることがあります。UNIX は、X/Open Company, Ltd. が独占的にライセンスしている米国ならびに他の国における登録商標です。フォント技術を含む第三者のソフトウェアは、著作権により保護されており、提供者からライセンスを受けているものです。

Federal Acquisitions: Commercial Software—Government Users Subject to Standard License Terms and Conditions.

本製品に含まれる HG 明朝 L、HG-MincyoL-Sun、HG ゴシック B、および HG-GothicB-Sun は、株式会社リコーがリコービマジクス株式会社からライセンス供与されたタイプフェイスマスタをもとに作成されたものです。HG 平成明朝体 W3@X12 は、株式会社リコーが財団法人日本規格協会からライセンス供与されたタイプフェイスマスタをもとに作成されたものです。フォントとして無断複製することは禁止されています。

Sun、Sun Microsystems、docs.sun.com、AnswerBook、AnswerBook2 は、米国およびその他の国における米国 Sun Microsystems, Inc. (以下、米国 Sun Microsystems 社とします) の商標もしくは登録商標です。

サンのロゴマークおよび Solaris は、米国 Sun Microsystems 社の登録商標です。

すべての SPARC 商標は、米国 SPARC International, Inc. のライセンスを受けて使用している同社の米国およびその他の国における商標または登録商標です。SPARC 商標が付いた製品は、米国 Sun Microsystems 社が開発したアーキテクチャに基づくものです。

OPENLOOK、OpenBoot、JLE は、サン・マイクロシステムズ株式会社の登録商標です。

Wnn は、京都大学、株式会社アステック、オムロン株式会社で共同開発されたソフトウェアです。

Wnn6 は、オムロン株式会社、オムロンソフトウェア株式会社で共同開発されたソフトウェアです。© Copyright OMRON Co., Ltd. 1995-2000. All Rights Reserved. © Copyright OMRON SOFTWARE Co., Ltd. 1995-2002 All Rights Reserved.

「ATOK」は、株式会社ジャストシステムの登録商標です。

「ATOK Server/ATOK12」は、株式会社ジャストシステムの著作物であり、「ATOK Server/ATOK12」にかかる著作権その他の権利は、株式会社ジャストシステムおよび各権利者に帰属します。

本製品に含まれる郵便番号辞書 (7 桁/5 桁) は郵政事業庁が公開したデータを元に制作された物です (一部データの加工を行なっています)。

本製品に含まれるフェイスマーク辞書は、株式会社ビレッジセンターの許諾のもと、同社が発行する『インターネット・パソコン通信フェイスマークガイド '98』に添付のものを使用しています。© 1997 ビレッジセンター

Unicode は、Unicode, Inc. の商標です。

本書で参照されている製品やサービスに関しては、該当する会社または組織に直接お問い合わせください。

OPEN LOOK および Sun Graphical User Interface は、米国 Sun Microsystems 社が自社のユーザおよびライセンス実施権者向けに開発しました。米国 Sun Microsystems 社は、コンピュータ産業用のビジュアルまたはグラフィカル・ユーザインタフェースの概念の研究開発における米国 Xerox 社の先駆者としての成果を認めるものです。米国 Sun Microsystems 社は米国 Xerox 社から Xerox Graphical User Interface の非独占的ライセンスを取得しており、このライセンスは米国 Sun Microsystems 社のライセンス実施権者にも適用されます。

DtComboBox ウィジェットと DtSpinBox ウィジェットのプログラムおよびドキュメントは、Interleaf, Inc. から提供されたものです。(© 1993 Interleaf, Inc.)

本書は、「現状のまま」をベースとして提供され、商品性、特定目的への適合性または第三者の権利の非侵害の黙示の保証を含みそれに限定されない、明示的であるか黙示的であるかを問わない、なんらの保証も行われぬものとします。

本製品が、外国為替および外国貿易管理法 (外為法) に定められる戦略物資等 (貨物または役務) に該当する場合、本製品を輸出または日本国外へ持ち出す際には、サン・マイクロシステムズ株式会社の事前の書面による承諾を得ることのほか、外為法および関連法規に基づく輸出手続き、また場合によっては、米国商務省または米国所轄官庁の許可を得ることが必要です。



030609@5943



目次

はじめに 5

パート I 「入力方式の概要」

1 概要 11

Solaris の入力方式 11

複数言語文字入力 11

ATOK12 11

Wnn6 12

ATOK12 の概要 12

「それ」がわかる ATOK 12

指示詞の照応関係の解析 13

快適な変換を実現する高度な日本語処理技術 13

リアルタイムに文書校正を支援 14

入力をサポートする便利な機能 15

複数の辞書を同時に使用する 16

充実したヘルプ 16

Wnn6 の概要 16

特徴 16

かな漢字変換機能 18

学習機能 22

Wnn6 の構成 23

パート II 「ファイル構成」

2	ATOK12 のファイル構成	29
	ファイル構成	29

3	Wnn6 のファイル構成	33
	ファイル構成	33

パート III 「ユーザー登録単語および設定ファイルの移行」

4	ATOK12 への移行	41
	ATOK8 の辞書に登録した単語を ATOK12 で使用する	41
	ATOK8 の環境設定ファイルを ATOK12 に移す	42
	cs00 の辞書に登録した単語を ATOK12 で利用する	42

5	Wnn6 への移行	45
	ATOK8、ATOK7、cs00、VJE-Delta、EGBRIDGE の単語情報を Wnn6 で使用する	45
	ファイル形式	46
	単語ファイル	47
	書式ファイル	50
	Wnn6 テキスト形式辞書	56
	ログファイル	57

A	Solaris 1.x アプリケーション	59
---	----------------------	----

	索引	61
--	----	----

はじめに

このマニュアルでは、Solaris™ 9 の日本語環境で使用できる入力方式 (複数言語文字入力、ATOK12、Wnn6) の概要とユーザー登録単語などの移行方法について説明します。各入力方式を使った入力方法、カスタマイズ方法などの詳細については、各入力方式のマニュアルを参照してください。

対象読者

Solaris 9 の日本語入力方式の利用者と、日本語入力方式を運用するシステムの管理者 (スーパーユーザー) を対象としています。

内容の紹介

パート I, 入力方式の概要

第 1 章

入力方式の概要と、日本語入力方式 (ATOK12、Wnn6) の特徴を説明します。

パート II, ファイル構成

第 2 章

ATOK12 のファイル構成について説明します。

第 3 章

Wnn6 のファイル構成について説明します。

パート III, ユーザー登録単語および設定ファイルの移行

第 4 章

ATOK8 および cs00 でユーザーが登録した単語情報を ATOK12 で使用する方法について説明します。また、ATOK8 の環境設定を ATOK12 に移す方法についても説明します。

第 5 章

ATOK8、ATOK7、cs00、VJE-Delta、EGBRIDGE でユーザーが登録した単語情報を Wnn6 で使用する方法と、移行時に使用するファイルの形式について説明します。

付録 A

Solaris 1.x アプリケーションで日本語を入力するための設定について説明します。

関連マニュアル

- 『ATOK12 ユーザーズガイド』
- 『Wnn6 ユーザーズガイド』
- 『Wnn6 上級ユーザーおよびシステム管理者ガイド』

Sun のオンラインマニュアル

docs.sun.com では、Sun が提供しているオンラインマニュアルを参照することができます。マニュアルのタイトルや特定の主題などをキーワードとして、検索を行うこともできます。URL は、<http://docs.sun.com> です。

表記上の規則

このマニュアルでは、次のような字体や記号を特別な意味を持つものとして使用します。

表 P-1 表記上の規則

字体または記号	意味	例
AaBbCc123	コマンド名、ファイル名、ディレクトリ名、画面上のコンピュータ出力、コード例を示します。	.login ファイルを編集します。 ls -a を使用してすべてのファイルを表示します。 system%
AaBbCc123	ユーザーが入力する文字を、画面上のコンピュータ出力と区別して示します。	system% su password:
AaBbCc123	変数を示します。実際に使用する特定の名前または値で置き換えます。	ファイルを削除するには、rm <i>filename</i> と入力します。
『 』	参照する書名を示します。	『コードマネージャ・ユーザーズガイド』を参照してください。
[]	参照する章、節、ボタンやメニュー名、強調する単語を示します。	第 5 章「衝突の回避」を参照してください。 この操作ができるのは、「スーパーユーザー」だけです。
\	枠で囲まれたコード例で、テキストがページ行幅を超える場合に、継続を示します。	sun% grep `^#define \ XV_VERSION_STRING'

コード例は次のように表示されます。

■ C シェル

```
machine_name% command y|n [filename]
```

■ C シェルのスーパーユーザー

```
machine_name# command y|n [filename]
```

■ Bourne シェルおよび Korn シェル

```
$ command y|n [filename]
```

■ Bourne シェルおよび Korn シェルのスーパーユーザー

```
# command y|n [filename]
```

[] は省略可能な項目を示します。上記の例は、*filename* は省略してもよいことを示しています。

| は区切り文字 (セパレータ) です。この文字で分割されている引数のうち 1 つだけを指定します。

キーボードのキー名は英文で、頭文字を大文字で示します (例: Shift キーを押します)。ただし、キーボードによっては Enter キーが Return キーの動作をします。

ダッシュ (-) は 2 つのキーを同時に押すことを示します。たとえば、Ctrl-D は Control キーを押したまま D キーを押すことを意味します。

パートⅠ 入力方式の概要

第 1 章

概要

この章では、Solaris の日本語入力方式の概要と、ATOK12 および Wnn6 の、両日本語入力方式の特徴を説明します。

Solaris の入力方式

Solaris 共通デスクトップ環境 (以降「Solaris CDE」とします) 上では、ログイン画面の「オプション」で選択した入力方式を使用して、地域環境に応じた文字を入力することができます。Solaris CDE の日本語環境では、以下の入力方式を使用することができます。

複数言語文字入力

日本語、韓国語、簡体字中国語、繁体字中国語、繁体字中国語(香港)、タイ語、英語、ヨーロッパ系各種言語、キリル文字、ギリシア語、アラビア語、ヒンディー語、ヘブライ語等の文字を入力することができます。ただし、日本語、韓国語、簡体字中国語、繁体字中国語、繁体字中国語(香港)、およびタイ語の文字を入力するためには、対応する地域環境のパッケージがインストールされている必要があります。地域環境とパッケージの対応に関しては、『国際言語環境ガイド』を参照してください。なお、日本語の入力には ATOK12 が使用されます。

この入力方式は、ja_JP.UTF-8 ロケールで使用できます。

ATOK12

日本語を入力するための入力方式です。

この入力方式は、ja、ja_JP.eucJP、ja_JP.PCK および ja_JP.UTF-8 ロケールで使用できます。

Wnn6

日本語を入力するための入力方式です。Wnn6 には、日本語の入力機能を持たない日本語の表示が可能な端末上で日本語の入力を可能にする uum(1) コマンドが含まれます。

今回のリリースで提供される Wnn6 は、最大 3 つまでのクライアント (htt や uum など、Wnn6 のかな漢字変換サーバーである jserver に直接接続するプログラム) の同時接続をサポートします。

追加のクライアントライセンスは、別途購入することができます。詳細は、本製品のご購入先にお問い合わせください。

この入力方式は、ja、ja_JP.eucJP、ja_JP.PCK および ja_JP.UTF-8 ロケールで使用できます。

ATOK12 の概要

「それ」がわかる ATOK

文脈解析変換

入力してきた文章の分野を自動的に判断し、文脈情報に基づいた同音語処理を実現しています。

現在、コンサートホールではピアノコンクールが行われている。 各国から集まったピアニストたちが熱演を繰り広げているところである。 いよいよ次は最後の奏者だ。
このレースも終局を迎え、ゴール地点の騎技場もすぐそこに迫っている。 しかし依然勝負は見えず、各チームの監督たちも気をもんでいる。 いよいよ次は最後の走者だ。

指示詞の照応関係の解析

文脈解析変換をさらに強化し、文章中の指示詞(これ、それ、あれ、この、その、あの、など)と、それらが指し示す言葉との関係を適切に判断して、同音語を正しく識別します。

今日の授業の問題で一つだけとても難しいものがありました。
○ はじめにそれを解いたのは私の友人でした。
× はじめにそれを説いたのは私の友人でした。

「それ」が「授業の問題」を指すことを認識し、「解いた」に正しく変換します。

来年度中に大幅な減税が実施される予定です。
○ その政策で国民の負担は多少は軽くなるはずですが。
× その制作で国民の負担は多少は軽くなるはずですが。

「その」が「大幅な減税」を指すことを認識し、「政策」に正しく変換します。

車の窓から遠くに姫路城が見えます。
○ あの城はどれくらい有名なのでしょうか。
× あの日はどれくらい有名なのでしょうか。

「あの」が「姫路城」を指すことを認識し、「城」に正しく変換します。

快適な変換を実現する高度な日本語処理技術

格フレーム処理

文章の差異による同音語の違いを的確に判断して変換します。

そのせいふくがかのじょにあう → その制服が彼女に合う
そのせんせいがかのじょにあう → その先生が彼女に合う

係り受け関係の解析

それぞれの文節が文章の中でどのような係り受け関係になっているかを解析し、高い変換効率を実現しています。

窓を 夜が 明けるまでに 開けておきました

「窓を」が「開けておきました」に、「夜が」が「明けるまでに」に係っているのを正しく認識します。

文節区切りの判断

構文意味解析により、判断しづらい文節の区切りも的確に判断します。

きょうははをみがいた → 今日は歯を磨いた
きょうははをたずねた → 今日母を訪ねた

「磨いた」、「訪ねた」の動詞の違いにより、文節の区切りを正しく判断しています。

ATOK12の日本語処理技術について

最先端技術の搭載だけでなく、地味に思われがちな処理も着実に強化して、より一層快適な変換を実現しています。

たとえば、ATOK12 では、数字交じりの文章もより正確に変換できます。

ミュージカルの第1次選考会が午後1時より始まります。
ホテルの2階で第2回懇親会が開かれます。

リアルタイムに文書校正を支援

JUST MEDDLER2 (ジャストメドラー 2)

入力・変換する時点で文章をチェックし、正確な文章の表現をサポートする JUST MEDDLER2 を搭載しています。「ら抜き表現の使用」、「文体の統一」、「同一助詞の過剰な連続」、「修飾関係のあいまい性」、「わかりにくい否定表現」、「誤用」、「仮名遣い誤り」、「商標名」、「機種依存文字」をチェックし、メッセージで注意を促します。

入力をサポートする便利な機能

以前入力した文字列を短い読みで変換する

一度入力した文字列を短い読みで変換して、効率的に入力することができます。文字列を入力したとき、途中で変換した・しないにかかわらず、意味のまとまりのある候補から選択して入力できます。確定した候補は学習されるので、コンピュータを起動し直しても、引き続き短い読みで変換できます。

入力した文字列を履歴から繰り返し入力する

文字列を確定したあとに同じ文字列を未確定の状態を入力できます。ATOK12では、直前に入力した文字列だけでなく、最近確定した文字列も履歴から選択して入力できます。

変換候補を並べ替えて目的の文字列をすばやく検索する

候補ウィンドウに表示される変換候補を、先頭の文字列または末尾の文字列が同じ単語で並べ替えることができます。人名など変換候補が多い場合に便利な機能です。

文字を半角または全角に統一する

文字の種類ごとに、半角・全角のどちらに変換するかを指定できます。入力中の文字や、後変換での表記を統一することもできます。

単語にコメントや置換候補を設定する

特定の語句が入力された場合に、メッセージが表示されるように設定しておくことができます。禁止用語や登録商標、常用漢字外などの入力の注意や、置換候補として電子メールの宛先を登録しておくなど、いろいろな応用が可能です。

わかりやすい漢字検索

部首、総画数、読みなどの検索条件をもとに、文字パレットを使いながら視覚的に漢字検索を行います。複数の検索条件を組み合わせて設定することも可能です。

入力ミスを自動的に修正する

ローマ字入力時のアルファベットの入力ミスや、カナ入力時の濁点・半濁点の入力ミスを自動的に修正し、正しいかな漢字変換を行います。

- ローマ字入力時

入力文字列	変換後の文字列
まっっちゃをのむ	抹茶を飲む
ちゅごくりょこう	中国旅行
こうえんおいす	公園のいす

■ カナ入力時

入力文字列	変換後の文字列
らくか° き	落書き
らつきー	ラッキー
てがみわだす	手紙を出す

複数の辞書を同時に使用する

複数の辞書をまとめて、まるでひとつの辞書のような感覚で運用することができます。1個のシステム辞書、1個のユーザー辞書、そして4個までの補助辞書をセットで使うことが可能です。たとえば、補助辞書には医学用辞書や法律用辞書など分野別の辞書を組み込んでおくと便利です。これが1つの辞書セットで、辞書セットは最大10セットまで設定しておくことができます。

充実したヘルプ

ユーティリティの操作と機能に関してだけでなく、日本語入力・変換など全般にわたる情報をヘルプを使って参照することができます。

Wnn6 の概要

特徴

Wnn6 には次のような特徴があります。

高度な変換機能と学習機能

Wnn6 は、FI (Flexible Intelligence) 変換機能と FI 学習機能により、高水準のかな漢字変換効率を実現しています。

- FI 変換
文節間の接続状態を見て、最も使用頻度の高い変換を行います。
- FI 学習
確定した文節間の関係規則が、FI 関係辞書 (文節間接続関係辞書) に登録されていない場合、各ユーザーの FI 関係辞書に登録します。

クライアント / サーバー方式

Wnn6 のシステム構成は、クライアント / サーバー方式を採用しています。これにより、複数のクライアントからサーバー側にある同一の辞書を使用することができるので、どのクライアントからも同一の学習効果を得ることができます。

また、辞書の追加などのメンテナンス作業を効率良く行なったり、効果的なオフライン学習機能をサポートすることができます。

さらに、この方式を使用することにより、サーバーにアクセスできるクライアントを制限することもできます。

GUI ユーティリティ

辞書操作や環境設定には、GUI のツールが用意されています。Wnn6 に対して特別な知識がなくても、すべての環境の設定を簡単に行うことができます。

入力スタイル

ユーザーの好みに合わせて、キーの割り当てを他の日本語入力システム (cs00, ATOK8 など) に合わせることができます。

辞書

辞書は、システム (固定形式) 辞書とユーザー (登録可能形式) 辞書に分類されます。辞書のエントリである変換文字列に対して、各ユーザーが個別の頻度情報を持つことにより、変換効率が向上します。

各辞書の特徴は次のとおりです。

- システム辞書語彙数 (約 20 万語)
システム辞書は、『岩波国語辞典第四版』をもとにしています。また、次のような独自の語彙が加えられています。
 - 現代語、話しことば
 - 複合語、複合品詞
 - 全国の主要地名、世界の国名、都市名

- 人名 (姓と名の分離)
- FI 関係用例 (約 300 万)
- ユーザー辞書
 - 最大登録語数
システムからの制限はありません。マシンのディスク容量によります。
 - 登録読み、登録漢字文字数 (253 文字)
 - 登録読み文字種 (全文字種)
 - 登録漢字文字種 (全文字種)

単語登録、単語削除、単語検索などの辞書操作は、辞書ユーティリティを利用して簡単にを行うことができます。また、他の日本語入力システム (cs00、ATOK7、ATOK8、VJE-Delta、EGBRIDGE) で作成された辞書から、Wnn6 の辞書フォーマットへ変換する辞書コンバータもあります。これらを活用することにより、他の日本語入力システムから Wnn6 への移行がスムーズに行えます。

かな漢字変換機能

FI 変換

Wnn6 の FI 変換とは、FI 関係辞書 (文節間接続関係辞書) を用いて、変換文字列中の各文節間の接続度合いを調査し、接続度合いの高い候補を優先して変換する機能をいいます。これにより、高いかな漢字変換効率を実現しています。

FI 変換には次の変換機能があります。

- 格係り受け変換
変換文字列中に、格情報を持つ文節と述語文節との接続関係を考慮します。

表 1-1 Wnn6 の格係り受け変換

が格	手が挙がる / 株が騰がる / てんぶらが揚がる
を格	身長を測る / 時間を計る / 暗殺を謀る
に、には格	宿に泊まる / 駅に停まる
で格	車で帰る / 貴社で会う
へ格	京へ向かった / 今日へ持ち越した
より格	車より速い / 予定より早い
から格	敵から奪回する / 会から脱会する

表 1-1 Wnn6 の格係り受け変換 (続き)

まで格	誤解まで招いた / 五階まで昇る
-----	------------------

■ 所有格変換

変換文字列中に、所有格情報 (の格) を持つ文節と対象文節との接続関係を考慮します。

表 1-2 Wnn6 の所有格変換

かいとう	⇒	会頭の回答
かいじょう	⇒	会場の開場

■ 受身変換

変換文字列中の述語文節に助動詞「れる」、「られる」がついている場合、「受身」、「可能」、「自発」、「尊敬」の効果を判別して、格係り受け接続関係を考慮します。

表 1-3 Wnn6 の受身変換

	FI 接続関係	変換可能文字列
受身	交通を - 規制する	交通が規制される
可能	テレビを - 見る	テレビが [を] 見られる
自発	故郷を - 偲ぶ	故郷が偲ばれる
尊敬	社長が - 読む	社長が読まれる

■ 使役変換

変換文字列中の述語文節に助動詞「せる」、「させる」がついている場合、「使役」の効果を判別して、格係り受け接続関係を考慮します。

表 1-4 Wnn6 の使役変換

	FI 接続関係	変換可能文字列
使役	彼が - 答える	彼に答えさせる

■ 合成語変換

変換文字列中に、2 種類の文節が接続して 1 つの意味を持つもの (複合語または人名) の接続関係を考慮します。

表 1-5 Wnn6 の合成語変換

複合語	⇒	集団 - 登校	会社 - 訪問
人名	⇒	福沢 - 諭吉	夏目 - 漱石

■ 修飾語変換

変換文字列中に、修飾を表す文節 (形容詞、形容動詞、副詞、連体詞) と被修飾文節との接続関係を考慮します。

表 1-6 Wnn6 の修飾語変換

形容詞	あつい△△△	熱い湯 / 厚い本 / 暑い夏
形容動詞	ていちょうな△△△ ふしんな△△△	丁寧な挨拶 / 低調な作品 不審な人影 / 不振な成績
副詞	▲▲▲もる	ぼたぼた漏る / こんもり盛る
連体詞	▲▲▲きのう	小さな機能 / 楽しかった昨日

■ 複文変換

変換文字列中に (格情報文節+述語文節) の組合せが複数存在する場合、それぞれの接続状態を判別して、接続関係を考慮します。

表 1-7 Wnn6 の複文変換

▽	▼
家が建ち	人が立つ

その他の変換

FI 変換以外には、次のような変換機能があります。

■ 複合語変換

付属語を含まない候補を優先します。

■ 最終使用最優先変換

同音異義語のなかで、直前に使用した単語を必ず先頭の候補に挙げます。次候補は使用した順に 6 個までを提示します。

■ 単漢字変換

単漢字だけを使って変換します。

■ 揺らぎ処理

入力文字列が辞書に登録されていない場合、その文字列と表記は異なっても、同じ発音になる別の単語が辞書に登録されていれば、その単語を候補として提示します。

表 1-8 Wnn6 の揺らぎ処理

こおり / こうり ⇒ 氷

■ 長音変換

上記の「揺らぎ処理」で、表記の違いが音引き記号 (一) を含む場合をいいます。

表 1-9 Wnn6 の長音変換

ういんどう / ういんどー ⇒ ウィンドウ

■ 送りがな基準処理

送りがなの基準 (本則¹ / 送る / 送らない) を指定しておく、指定された基準で送りがなを決定します。

表 1-10 Wnn6 の送りがな基準処理

	本則	送る	送らない
おこなう ⇒	行う	行なう	行う

■ Aa 変換

アルファベットを次の順で変換します。



■ 異形字処理

新字 ⇔ 旧字 / 略字 ⇔ 正字の変換を行います。

表 1-11 Wnn6 の異形字処理

渡辺 ⇔ 渡邊

■ 郵便番号変換

郵便番号 (3 桁、5 桁、7 桁) からその住所に変換します。

表 1-12 Wnn6 の郵便番号変換

600 ⇒ 京都府京都市下京区

■ 電話番号変換

電話番号からその住所に変換します。

1 「送り仮名の付け方」(昭和 48 年 6 月内閣公示文二号)の本則をさします。

表 1-13 Wnn6 の電話番号変換

075 ⇒ 京都府京都市
■ 英単語日本語変換 英単語を日本語の単語に変換します。

表 1-14 Wnn6 の英単語日本語変換

COMPUTER ⇒ コンピュータ
■ かな / カナ / ローマ字後変換 入力文字列をかな / カタカナ / ローマ字に変換します。
■ 連濁処理 連濁した読みを発音通りに入力すると、正しい漢字に変換します。連濁の条件が満たされていない場合には変換しません。

表 1-15 Wnn6 の濁音処理

べんきょうづくえ	⇒	勉強机
づくえ	⇒×	机

学習機能

FI 学習

Wnn6 には、FI 変換 (格係り受け変換、所有格変換、受身変換、使役変換、合成語変換、修飾語変換、複文変換) で、現在の FI 関係辞書に登録されていない文節間の関係をユーザーが確定した場合に、新しくユーザーごとの FI 関係辞書に登録する機能があります (FI 学習機能)。

FI 学習機能は、「する / しない / 一時的」の各設定に切り換えることができます。

頻度学習

単語の使用頻度をユーザーごとに管理して、使用頻度の高い単語を変換時に優先的に表示することができます。システム辞書内の単語と、FI 関係システム辞書内の単語を管理します。

学習レベルは、「じわじわ / 基準 / すぐ / 必ず / しない」の学習レベルで切り換えることができます。

その他の学習機能

- 無変換学習

辞書に登録されていない「ひらがな / カタカナ / ローマ字」の候補が確定された場合に、自動的に辞書へ登録します。「する / しない / 一時的」の各設定に切り換えることができます。

- 文節切り学習
文節の切り直しを行なって変換文字列を確定した場合に、切り直した前後の文節を学習します。「する / しない / 一時的」の各設定に切り換えることができます。
- 送りがな基準学習
送りがな基準がある単語は、単語ごとに確定した規則「本則 / 送る / 送らない」を学習して、直前に使用した規則を最優先します。「する / しない」の各設定に切り換えることができます。
- 接頭語学習
接頭語に「お / ご」「御」のどれを使用するか、直後の名詞ごとに学習します。「する / しない」の各設定に切り換えることができます。
- 接尾語学習
単語ごとに確定した文字列の接尾語の規則「カナ / 漢字 (送る) / 漢字 (送らない)」を学習して、直前に使用した規則を最優先します。「する / しない」の各設定に切り換えることができます。
- 一般語学習
確定した一般語の情報「ひらがな / カタカナ / 漢字」を学習します。「する / しない」の各設定に切り替えることができます。
- 数字学習
確定した数字表記の種類「漢数字 / 半角 / 全角 / カンマ付き」を学習し、直前に使用した種類を最優先します。この学習は無効にできません。
- 登録語自動反映
複数のユーザー辞書で登録されている同一単語を、共用辞書に反映させます。この機能はシステム管理者が設定します。
- オフライン学習
変換効率の向上とディスクとメモリー資源の削減を目的に、オフラインで動く機能をいいます。不必要な単語の削除や頻度の再配置を行い、登録語自動反映モジュールを起動させます。この機能はシステム管理者を対象としています。

Wnn6 の構成

Wnn6 は、ユーザーの入力処理を行うクライアント (Wnn6/Http) と、かな漢字変換を行うサーバー (jsrver) からなるクライアント / サーバー方式で構成されています。jsrver は、辞書引き専用サーバー (wnnds) と共に使用することができます。

クライアント / サーバー方式により、複数のクライアント (Wnn6/Http など) からのかな漢字変換要求を、1 台のマシンで動いている jsrver で処理することができます。

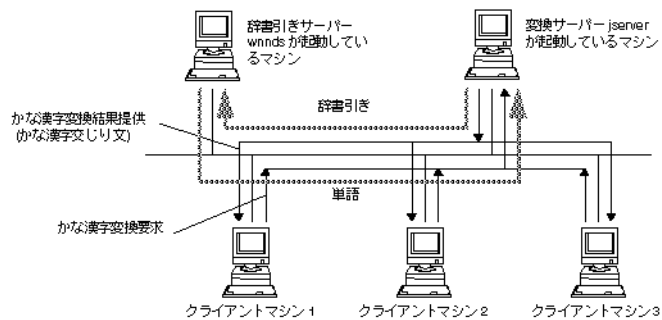


図 1-1 サーバー / クライアント方式の Wnn6

jserver のかな漢字変換機能を使って日本語入力処理を行うには、ユーザーとのインタフェース部分でアプリケーションプログラムが必要になります。このアプリケーションプログラムがクライアントに相当します。Solaris では、アプリケーションプログラムとして Wnn6/Htt と uum を提供しています。

Wnn6/Htt は X ウィンドウシステムで動く複数のクライアントに対して、日本語入力環境を提供します。Wnn6/Htt からは、Wnn6 辞書ユーティリティ、Wnn6 設定ユーティリティなどのユーティリティプログラムを起動することもできます。uum は、漢字端末や X ウィンドウシステムの `extern` などの個々のウィンドウ上で動作します。

辞書への単語登録と単語削除は、辞書ユーティリティ `wnndictutil` で簡単に操作できます。また、辞書コンバータ `wnotow` を使って、ATOK8、ATOK7、cs00、VJEdelta、EGBRIDGE のユーザー辞書ファイルを、Wnn6 でも活用することができます。

Wnn6/Htt から、キーの割り当ての設定や、かな漢字変換実行の環境 (辞書の指定、変換パラメータ値の設定など) のカスタマイズを行うことができます。その他の環境についても、Wnn6 設定ユーティリティ `wnnenvutil` を使って、カスタマイズを行うことができます。

Wnn6 のシステム構成の概略図を図 1-2 に示します。

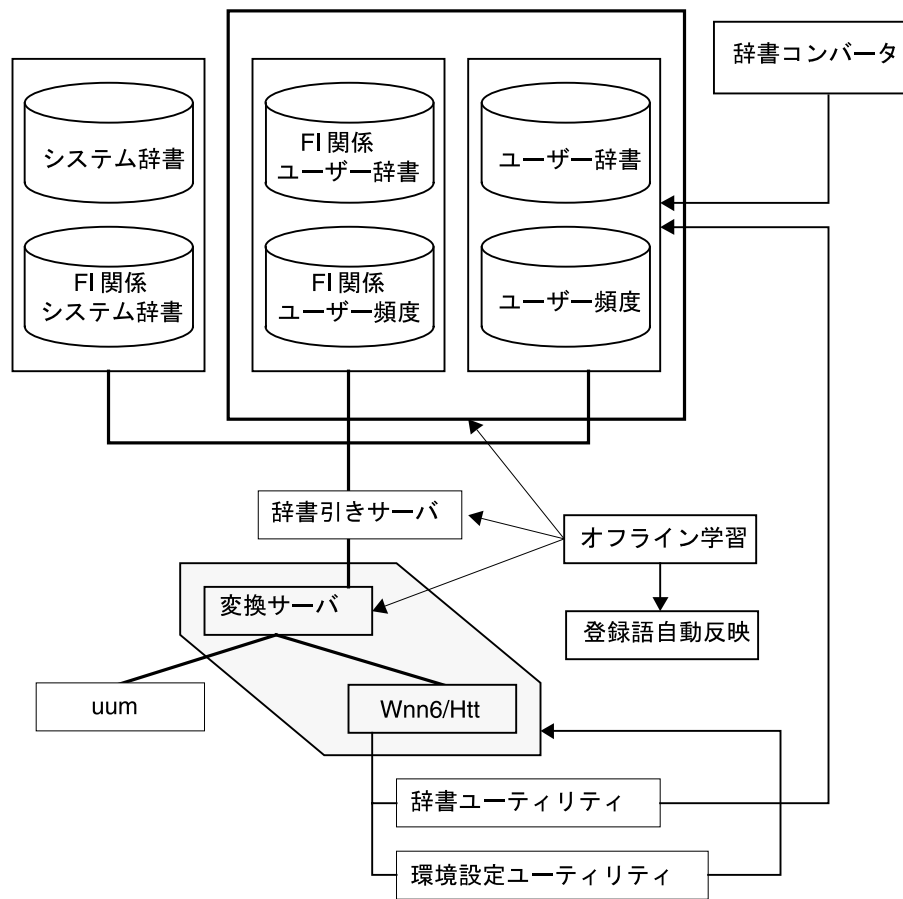


図 1-2 Wnn6 のシステム構成

パート II ファイル構成

第 2 章

ATOK12 のファイル構成

この章では、ATOK12 のファイル構成について説明します。

ファイル構成

ATOK12 のディレクトリ構成を図 2-1 に示します。

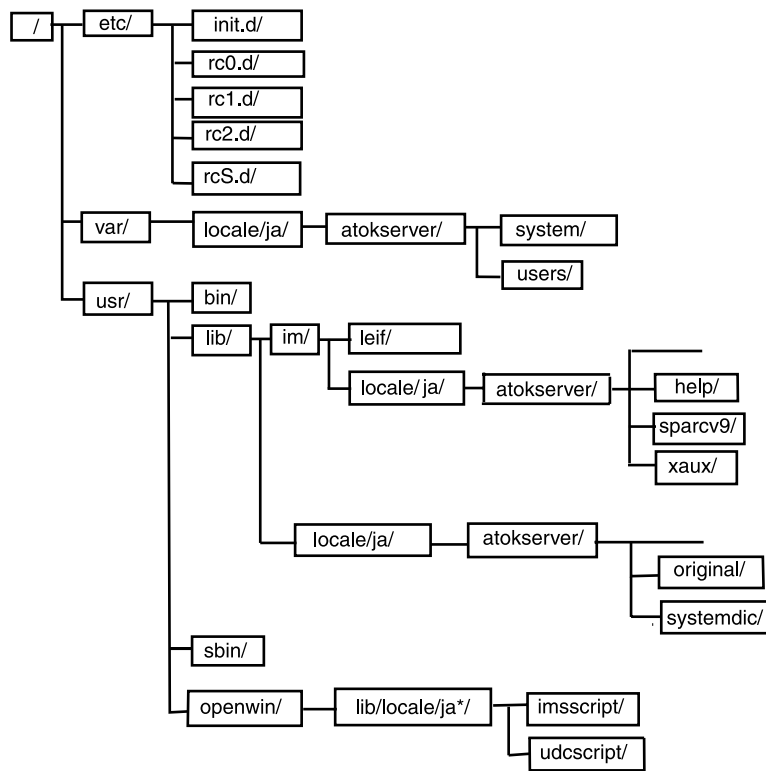


図 2-1 図 2-1 ATOK12 のディレクトリ構造

図 2-1 の各ディレクトリの内容、またはディレクトリの下に設置されるファイルの内容を表 2-1 で示します。

表 2-1 ATOK12 のファイル一覧

ファイル名	内容
/etc/init.d/atstv	ATOK12 ユーザ管理デーモン起動、停止用スクリプト
/etc/rc0.d/K33atstv	ATOK12 ユーザ管理デーモン停止用スクリプト
/etc/rc1.d/K33atstv	ATOK12 ユーザ管理デーモン停止用スクリプト
/etc/rc2.d/S99atstv	ATOK12 ユーザ管理デーモン起動用スクリプト
/etc/rcS.d/K33atstv	ATOK12 ユーザ管理デーモン停止用スクリプト

表 2-1 ATOK12 のファイル一覧 (続き)

ファイル名	内容
/var/locale/ja/atokserver/original	/usr の下の、雛形の置かれたディレクトリへのシンボリックリンク
/var/locale/ja/atokserver/system/atokuserinfo.dat	ユーザー管理ファイル
/var/locale/ja/atokserver/systemdic	/usr の下の、システム辞書の置かれたディレクトリへのシンボリックリンク
/var/locale/ja/atokserver/users/	ユーザー毎の設定ファイル、スタイルファイル、辞書ファイルを置くディレクトリ
/usr/bin/atok12migd	ATOK8 から ATOK12 への辞書移行コマンド
/usr/bin/atok12migs	ATOK8 から ATOK12 への設定移行コマンド
/usr/lib/im/leif/atok12.so	ATOK12 モジュール
/usr/lib/im/locale/ja/atokserver/atok12aux.jar	Java クライアント用 AUX モジュール
/usr/lib/im/locale/ja/atokserver/atok12aux.so	X クライアント用 AUX モジュール
/usr/lib/im/locale/ja/atokserver/help/*	ヘルプファイル
/usr/lib/im/locale/ja/atokserver/sparcv9/atok12aux.so	X クライアント用 AUX モジュール (64ビット)
/usr/lib/im/locale/ja/atokserver/xaux/*	X クライアント用ユーティリティ
/usr/lib/locale/ja/atokserver/atok12.dkd	検索辞書ファイル
/usr/lib/locale/ja/atokserver/*.so	ATOK12 モジュール
/usr/lib/locale/ja/atokserver/atokmngdaemon	ユーザー管理デーモン
/usr/lib/locale/ja/atokserver/original/atok12.conf	設定ファイル (雛形)
/usr/lib/locale/ja/atokserver/original/*.sty	スタイルファイル (雛形)
/usr/lib/locale/ja/atokserver/original/*.dic	ユーザー辞書 (雛形)

表 2-1 ATOK12 のファイル一覧 (続き)

ファイル名	内容
/usr/lib/locale/ja/atokserver/systemdic/atok12.dic	基本辞書
/usr/lib/locale/ja/atokserver/systemdic/atok12tk.dic	単漢字辞書
/usr/lib/locale/ja/atokserver/systemdic/atok12y5.dic	郵便番号辞書 (5桁)
/usr/lib/locale/ja/atokserver/systemdic/atok12y7.dic	郵便番号辞書 (7桁)
/usr/lib/locale/ja/atokserver/systemdic/atok12yc.dic	郵便番号辞書 (企業)
/usr/lib/locale/ja/atokserver/systemdic/facem12.dic	フェイスマーク辞書
/usr/lib/locale/ja/atokserver/systemdic/facemil12.dic	フェイスマーク辞書
/usr/sbin/atok12mngtool	ユーザー管理ツール
/usr/openwin/lib/locale/ja*/imsscript/S507atok12	入力方式起動スクリプト
/usr/openwin/lib/locale/ja*/udcscript/ATOK12_udcregister	外字ツール単語登録支援機能用スクリプト

第 3 章

Wnn6 のファイル構成

この章では、Wnn6 のファイル構成について説明します。

ファイル構成

Wnn6 システムのディレクトリ構成を図 3-1 に示します。

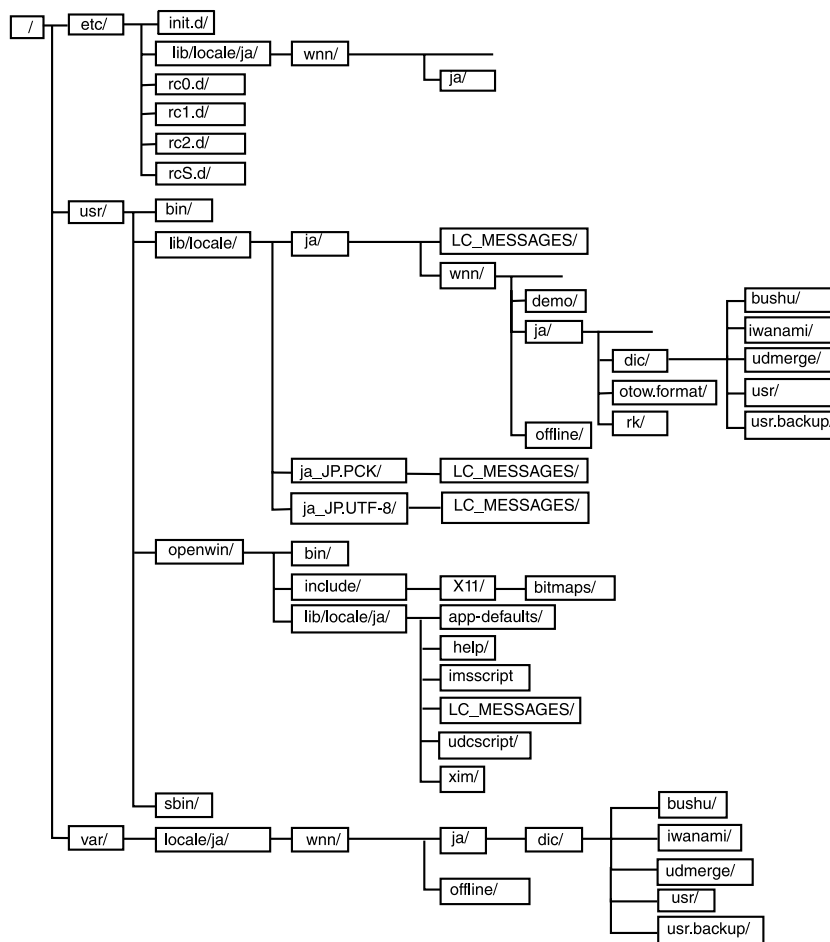


図 3-1 Wnn6 システムのディレクトリ構成

図 3-1 の各ディレクトリの内容、またはディレクトリの下に設置されるファイルの内容を表 3-2 で示します。

表 3-1 Wnn6 システムのファイル一覧 (1)

ファイル名	内容
/etc/init.d/Wnn6	jservice 起動、停止用スクリプト
/etc/lib/locale/ja/wnn/dpkeyallow	Wnn6 ライセンスサーバーへのアクセス制御
/etc/lib/locale/ja/wnn/dpkeylist	Wnn6 使用許諾情報

表 3-1 Wnn6 システムのファイル一覧 (1) (続き)

ファイル名	内容
/etc/lib/locale/ja/wnn/dpkeyservlist	Wnn6 ライセンスサーバーのリスト
/etc/lib/locale/ja/wnn/offlinerc	オフライン学習を初期化する
/etc/lib/locale/ja/wnn/serverdefs	Wnn6 ライブラリのサーバーとの接続環境設定
/etc/lib/locale/ja/wnn/udmergerc	登録語自動反映を初期化する
/etc/lib/locale/ja/wnn/wnnhosts	サーバーへのアクセス許可
/etc/lib/locale/ja/wnn/ja/jserverrc	jservice を初期化する
/etc/rc0.d/K15wnn6	jservice 停止用スクリプト
/etc/rc1.d/K15Wnn6	jservice 停止用スクリプト
/etc/rc2.d/S94Wnn6	jservice 起動用スクリプト
/etc/rcS.d/K15Wnn6	jservice 停止用スクリプト
/usr/bin/uum	かな漢字変換フロントエンドプロセッサ
/usr/bin/wnnatod	テキスト形式辞書⇒バイナリ形式辞書
/usr/bin/wnndtoa	バイナリ形式辞書⇒テキスト形式辞書
/usr/bin/wnnotow	辞書コンバータ
/usr/bin/wnnstat	変換サーバーの使用者状況を見る
/usr/bin/wnntouch	辞書の inode を合わせる
/usr/lib/locale/ja/LC_MESSAGES/*.cat	日本語 EUC 環境用メッセージ
/usr/lib/locale/ja/wnn/cvt_key_tbl*	キーコード変換表
/usr/lib/locale/ja/wnn/cvt_xim_tbl	ファンクション / メタキー変換データベース
/usr/lib/locale/ja/wnn/dpkeyserv	Wnn6 ライセンスサーバー
/usr/lib/locale/ja/wnn/dpkeystat	Wnn6 ライセンスの使用状況を見る
/usr/lib/locale/ja/wnn/jserver	かな漢字変換サーバー
/usr/lib/locale/ja/wnn/wnnds	辞書引きサーバー

表 3-1 Wnn6 システムのファイル一覧 (1) (続き)

ファイル名	内容
/usr/lib/locale/ja/wnn/wnnoffline	オフライン学習
/usr/lib/locale/ja/wnn/wnnudmerge	登録語自動反映
/usr/lib/locale/ja/wnn/ximrc	Wnn6/Htt を初期化する
/usr/lib/locale/ja/wnn/demo/	デモ用ライブラリのディレクトリ
/usr/lib/locale/ja/wnn/ja/hinsi.data	品詞管理
/usr/lib/locale/ja/wnn/ja/rk/mode	オートマトン変換 / モード定義表
/usr/lib/locale/ja/wnn/ja/uumrc	uum を初期化する
/usr/lib/locale/ja/wnn/ja/uumkey	キーバインド定義
/usr/lib/locale/ja/wnn/ja/wnnenvrc	かな漢字変換環境初期化
/usr/lib/locale/ja/wnn/ja/dic/bushu/bushu.dic	部首入力用辞書
/usr/lib/locale/ja/wnn/ja/dic/bushu/bushu.dic.no_JISX0212	部首入力用辞書 (補助漢字を含まない)
/usr/lib/locale/ja/wnn/ja/dic/bushu/bushu.fzk	部首入力用接続ベクトル定義
/usr/lib/locale/ja/wnn/ja/dic/iwanami/fisd	FI 関係システム辞書 (固定形式)
/usr/lib/locale/ja/wnn/ja/dic/iwanami/*.dic	システム辞書 (固定形式)
/usr/lib/locale/ja/wnn/ja/dic/iwanami/*.fzk	付属語情報データ

表 3-2 Wnn6 システムのファイル一覧 (2)

ファイル名	内容
/usr/lib/locale/ja/wnn/ja/dic/udmerge	/var の下にある共用辞書用ディレクトリへのリンク
/usr/lib/locale/ja/wnn/ja/dic/usr/	/var の下にあるユーザー辞書用ディレクトリへのリンク
/usr/lib/locale/ja/wnn/ja/dic/usr.backup/	/var の下にあるユーザー辞書バックアップ用ディレクトリへのリンク
/usr/lib/locale/ja/wnn/ja/otow.format/*.fmt	辞書コンバータで使用する各日本語入力システムの書式
/usr/lib/locale/ja/wnn/ja/rk/1*	オートマトン変換: 対応表の前処理部分

表 3-2 Wnn6 システムのファイル一覧 (2) (続き)

ファイル名	内容
/usr/lib/locale/ja/wnn/ja/rk/2*	オートマトン変換: 対応表の本処理部分
/usr/lib/locale/ja/wnn/ja/rk/3*	オートマトン変換: 対応表の後処理部分
/usr/lib/locale/ja/wnn/offline/	/var の下にあるオフライン学習用ディレクトリへのリンク
/usr/lib/locale/ja_JP.PCK/LC_MESSAGES/*.cat	PC 漢字コード環境用メッセージ
/usr/lib/locale/ja_JP.UTF-8/LC-MESSAGES/*.cat	UTF-8 環境用メッセージ
/usr/openwin/bin/wnnbushu	部首入力
/usr/openwin/bin/wnnenvutil	設定ユーティリティ
/usr/openwin/bin/wnndictutil	辞書ユーティリティ
/usr/openwin/include/X11/bitmaps/Wnn6.bitmap	ビットマップ
/usr/openwin/lib/locale/ja/app-defaults/Wnnbushu	Wnn6/Htt の部首入力リソース
/usr/openwin/lib/locale/ja/app-defaults/Wnndictutil	辞書ユーティリティのリソース
/usr/openwin/lib/locale/ja/app-defaults/Wnnenvutil	環境設定ユーティリティのリソース
/usr/openwin/lib/locale/ja/app-defaults/Xjsi	Wnn6/Htt のリソース
/usr/openwin/lib/locale/ja/help/wnnenvutil.hlp	環境設定ユーティリティのヘルプ
/usr/openwin/lib/locale/ja/imsscript/S509Wnn6	Wnn6/Htt 起動用スクリプト
/usr/openwin/lib/locale/ja/LC_MESSAGES/*.cat	メッセージ
/usr/openwin/lib/locale/ja/udcscript/Wnn6_udcregister	外字ツール単語登録支援機能用スクリプト
/usr/openwin/lib/locale/ja/xim/xjsi.so	htt 用インタフェースモジュール
/usr/sbin/wnnaccess	サーバーへのアクセスを制御する
/usr/sbin/wnnkill	jservice を終了する

表 3-2 Wnn6 システムのファイル一覧 (2) (続き)

ファイル名	内容
/var/locale/ja/wnn/ja/dic/bushu	/usr の下にある部首入力用ディレクトリへのシンボリックリンク
/var/locale/ja/wnn/ja/dic/iwanami	/usr の下にある辞書用ディレクトリ (iwanami) へのシンボリックリンク
/var/locale/ja/wnn/ja/dic/udmerge/	共用辞書用ディレクトリ
/var/locale/ja/wnn/ja/dic/usr/<ユーザー名>/fiud	FI 関係ユーザー辞書 (登録可能形式)
/var/locale/ja/wnn/ja/dic/usr/<ユーザー名>/muhenkan	ユーザーごとの無変換学習辞書 (登録可能)
/var/locale/ja/wnn/ja/dic/usr/<ユーザー名>/bunsetsu	ユーザーごとの文節切り学習辞書 (登録可能)
/var/locale/ja/wnn/ja/dic/usr/<ユーザー名>/ud	ユーザー辞書 (登録可能)
/var/locale/ja/wnn/ja/dic/usr/<ユーザー名>/*.h	ユーザーごとの頻度情報
/var/locale/ja/wnn/ja/dic/usr.backup/	ユーザー辞書バックアップ用ディレクトリ
/var/locale/ja/wnn/offline/	オンライン学習用データディレクトリ

ここで示したファイルのパス名は初期設定のものであります。

初期化ファイル jserverrc、uumrc、wnnenvrc の中で、次のパスを設定していません。

- /usr/lib/locale/ja/wnn/ja/dic
jserverrc のエントリ jserver_dir で指定されています。
- ユーザーが新しく辞書を作成する場合の辞書と頻度ファイル
uumrc のエントリ setjishopath/sethindopath で指定されています。
- /usr/lib/locale/ja/wnn/ja/dic/iwanami/*.fzk
wnnenvrc のエントリ setfuzokugo で指定されています。
- /usr/lib/locale/ja/wnn/ja/dic/usr/<ユーザー名>/*.h、
/usr/lib/locale/ja/wnn/ja/dic/usr/<ユーザー名>/ud、
/usr/lib/locale/ja/wnn/ja/dic/usr/<ユーザー名>/fiud、
wnnenvrc のエントリ setdic で指定されています。

パート III ユーザー登録単語および設定ファイル の移行

第 4 章

ATOK12 への移行

この章では、ATOK8 および cs00 でユーザーが登録した単語情報を ATOK12 で使用する方法について説明します。また、ATOK8 の環境設定を ATOK12 に移す方法についても説明します。

ATOK8 の辞書に登録した単語を ATOK12 で使用する

ATOK8 の辞書は、`atok12migd` コマンドを使用して、ATOK12 の辞書に併合することができます。この際、ユーザーによって登録された単語が併合の対象となります。ATOK8 の初期状態の辞書に含まれていた単語は、併合の対象にはなりません。また、ATOK8 の辞書は読み出し元となりますので、この併合によって ATOK8 の辞書の内容が変わることはありません。

例 - ATOK12 の辞書 (`atok12u0.dic`) に、ATOK8 の辞書 (`atok8.dic`) を併合する。

```
sun% atok12migd atok8.dic atok12u0.dic
```

併合の結果の辞書 (`atok12u0.dic`) は、システム管理者 (`root`) の権限で、`/var/locale/ja/atokserver/user/<ユーザー名>` の下に移します。

例 - ユーザー (`yamada`) が作成した ATOK12 の辞書 (`atok12u0.dic`) を、システムのディレクトリに移す。

```
sun# cp -p atok12u0.dic /var/locale/ja/atokserver/users/yamada
sun# chown yamada /var/locale/ja/atokserver/users/yamada/atok12u0.dic
sun# chmod 600 /var/locale/ja/atokserver/users/yamada/atok12u0.dic
```

`atok12migd` コマンドの詳細に関しては、`atok12migs(1)` のマニュアルページを参照してください。

ATOK8 の環境設定ファイルを ATOK12 に移す

ATOK8 の環境設定ファイルは、`atok12migs` コマンドを使用して、ATOK12 のスタイルファイルに変換することができます。

例 - ATOK8 の環境設定ファイル (`atok8.ucf`) を、ATOK12 のスタイルファイル (`atok12new.sty`) に変換する。

```
sun% atok12migs atok8.ucf atok12new.sty
```

作成したスタイルファイルは (`atok12new.sty`)、システム管理者 (`root`) の権限で、`/var/locale/ja/atokserver/user/<ユーザー名>` の下に移します。

例 - ユーザー (`yamada`) が作成した ATOK12 のスタイルファイル (`atok12new.sty`) を、システムのディレクトリに移す。

```
sun# cp -p atok12new.sty /var/locale/ja/atokserver/users/yamada
sun# chown yamada /var/locale/ja/atokserver/users/yamada/atok12new.sty
sun# chmod 600 /var/locale/ja/atokserver/users/yamada/atok12new.sty
```

`atok12migs` コマンドの詳細については、`atok12migs(1)` のマニュアルページを参照してください。

cs00 の辞書に登録した単語を ATOK12 で利用する

cs00 の辞書に登録した単語を ATOK12 で使用する場合、以下の手順で行います。

1. `udicm`、`sdtudicm` (Solaris CDE の場合)、または `udicmtool` (OpenWindows の場合) を使用して、既存の `cs00` の辞書から単語リストのファイルを作成します。この作業は、`cs00` が動作しているシステム上で行います。Solaris 9 では、`cs00` は提供されていません。

例 `-udicm` を使用して、cs00 のユーザー辞書 (`cs00_u.dic`) を単語リストのファイル (`cs00_u.list`) に変換する

```
sun% udicm show cs00_u.dic > cs00_u.list
```

`udicm` コマンドの詳細に関しては、Solaris 9 の `udicm(1)` のマニュアルページを参照してください。

`sdtudicm` または `udicmtool` を使用方法については、Solaris 8 の『`cs00` ユーザーズガイド』を参照してください。

2. **cs00toatok** コマンドを使用して、**cs00** の単語リストのファイルを **ATOK12** が扱える単語ファイルに変換します。

例 - **cs00** の単語リストのファイル (**cs00_u.dic**) を **ATOK12** が扱える単語ファイルリスト (**atok12.list**) に変換する

```
sun% cs00toatok cs00_u.list > atok12.list
```

3. **ATOK12** 辞書ユーティリティを使用して、手順 2 で作成した単語ファイルから単語を登録します。
 - a. **ATOK** パレットから **ATOK12** 辞書ユーティリティを起動します。
 - b. **ATOK12** 辞書ユーティリティの「一括処理」ボタンを押し、「単語一括処理」を選択します。
 - c. 「単語ファイル」として、作成した単語ファイル (この例では**atok12.list**) を指定します。
 - d. 必要に応じて、「**ATOK** 辞書」と「出力ファイル」を設定します。
 - e. 「登録」ボタンを押しします。

第 5 章

Wnn6 への移行

この章では、ATOK8、ATOK7、cs00、VJE-Delta、EGBRIDGE でユーザーが登録した単語情報を Wnn6 で使用する方法について説明します。また、移行の際に利用されるファイルの形式についても説明します。

ATOK8、ATOK7、cs00、VJE-Delta、EGBRIDGE の単語情報を Wnn6 で使用する

次の手順により単語情報を Wnn6 で使用できるようになります。

1. **ATOK8、ATOK7、cs00、VJE-Delta、EGBRIDGE** でユーザーが登録した単語情報を、テキスト形式で単語ファイルに保存します。この方法については、各日本語入力システムのマニュアルを参照してください。なお、作成した単語ファイルの文字コードが日本語 **EUC** でない場合、文字コードを日本語 **EUC** に変換する作業が必要です。変換方法については、**iconv(1)** および **iconv_ja(5)** のマニュアルページを参照してください。

- 例 - ATOK8 で単語ファイルを作成する場合

- a. ATOK8 の「辞書ツール」ボタンを押し、辞書メンテナンスツールを起動します。
- b. 「一覧表示」ボタンを押し、一覧表示ウィンドウを開きます。
- c. 「単語一覧ファイル作成」を「する」に設定し、一覧表示を実行して、`atok8.txt` に単語情報を保存します。

- 例 - cs00 で単語ファイルを作成する場合

```
sun% udicm show cs00_u.dic > cs00.txt
```

- 例 - PC 漢字コードで作成されたテキストファイルの文字コードを変換する場合

```
sun% iconv -f PCK -t eucJP pck-word-file >word-file
```

- 作成した単語ファイルを **Wnn6** のテキスト形式辞書に変換します。

```
sun% wnnotow -i word-file -o wnn6-dict -f format -l log-file
```

各指定の意味は次のとおりです。詳細については、後述の「ファイル形式」の節を参照してください。

- i *word-file* 変換の対象となる単語ファイル
- f *format* 変換の対象となる単語ファイルの形式を示す書式ファイル
- o *wnn6-dict* 作成する Wnn6 テキスト形式辞書
- l *log-file* 変換に失敗した単語を記録するログファイル

- 例 - ATOK8 で登録した単語情報を使用する場合

```
sun% wnnotow -i atok8.txt -o wnn6.u -f \  
/usr/lib/locale/ja/wnn/ja/otow.format/atok8-wnn6.fmt
```

- 例 - ATOK7 で登録した単語情報を使用する場合

```
sun% wnnotow -i atok7.txt -o wnn6.u -f \  
/usr/lib/locale/ja/wnn/ja/otow.format/atok7-wnn6.fmt
```

- 例 - cs00 で登録した単語情報を使用する場合

```
sun% wnnotow -i cs00.txt -o wnn6.u -f \  
/usr/lib/locale/ja/wnn/ja/otow.format/cs00-wnn6.fmt
```

- 例 - VJE-Delta で登録した単語情報を使用する場合

```
sun% wnnotow -i vje.txt -o wnn6.u -f \  
/usr/lib/locale/ja/wnn/ja/otow.format/vje-wnn6.fmt
```

- 例 - EGBRIDGE で登録した単語情報を使用する場合

```
sun% wnnotow -i egb.txt -o wnn6.u -f \  
/usr/lib/locale/ja/wnn/ja/otow.format/egbridge-wnn6.fmt
```

- 作成した **Wnn6** テキスト形式辞書を **Wnn6** 辞書ユーティリティ (**wnndictutil**) で登録します。**Wnn6** 辞書ユーティリティの使用方法については、『**Wnn6 ユーザーズガイド**』を参照してください。

ファイル形式

この節では、**Wnn6** のテキスト形式辞書を作成する場合に使われるファイルの形式について説明します。各ファイルでは、日本語 EUC を使用します。

単語ファイル

変換の対象となる単語ファイル (ATOK8、ATOK7、cs00、VJE-Delta、EGBRIDGE) を指定します。それぞれの「読み」と「単語」の長さは、253 文字までです。

ATOK8 辞書形式

!コメント		
読み 1,	単語 1,	品詞 1
読み 2,	単語 2,	品詞 2
読み 3,	単語 3,	品詞 3
.	.	.
.	.	.

例:

!ATOK8UT;単語一覧;A:ATOK8 ATOK8.DIC		
book,	本,	一般名詞
ア,	有,	ラ行五段
ア,	亜,	単漢字
ア,	阿,	単漢字
ア,	あ,	カ行五段
.	.	.
.	.	.

ATOK7 辞書形式

【コメント】		
読み 1,	単語 1,	品詞 1
読み 2,	単語 2,	品詞 2
読み 3,	単語 3,	品詞 3

・	・	・
・	・	・

例：

ア,	亜,	単漢字
ア,	あ,	力行五段
ア,	あ,	ワ行五段
ア,	阿,	単漢字
ア,	開,	力行五段
・	・	・
・	・	・

cs00 辞書形式

#コメント		
読み 1 「Tab」	単語 1 「Tab」	品詞 1
読み 2 「Tab」	単語 2 「Tab」	品詞 2
読み 3 「Tab」	単語 3 「Tab」	品詞 3
・	・	・
・	・	・

例：

あ	有	:RV
あ	飽	:KV
あ	遭	:MV
・	・	・
・	・	・

VJE-Delta 辞書形式

#コメント		
読み 1 「Tab」	単語 1 「Tab」	品詞 1
読み 2 「Tab」	単語 2 「Tab」	品詞 2
読み 3 「Tab」	単語 3 「Tab」	品詞 3
・	・	・
・	・	・

例：

あ	編	[マ五]
あ	あ	[ワ五]
あ	あ	[ラ五ある]
あ	開	[カ五]
あ	空	[カ五]
・	・	・
・	・	・

EGBRIDGE 辞書形式

#コメント		
読み 1 「Tab」	単語 1 「Tab」	品詞 1
読み 2 「Tab」	単語 2 「Tab」	品詞 2
読み 3 「Tab」	単語 3 「Tab」	品詞 3
・	・	・
・	・	・

例：

あー	アーカイブ	1+20
----	-------	------

あきはばら	秋葉原	1+3
あせ	(^_^;	1
あたらし	新し	1
あとかぎ	』	1
あとかく]	1
あとてん	"	1
あとば	】	1
アプリ	アプリケーション	1
・	・	・
・	・	・

書式ファイル

単語ファイルの書式を定義します。

書式ファイルの内容は、「書式データ」部分と「品詞」部分に分かれています。

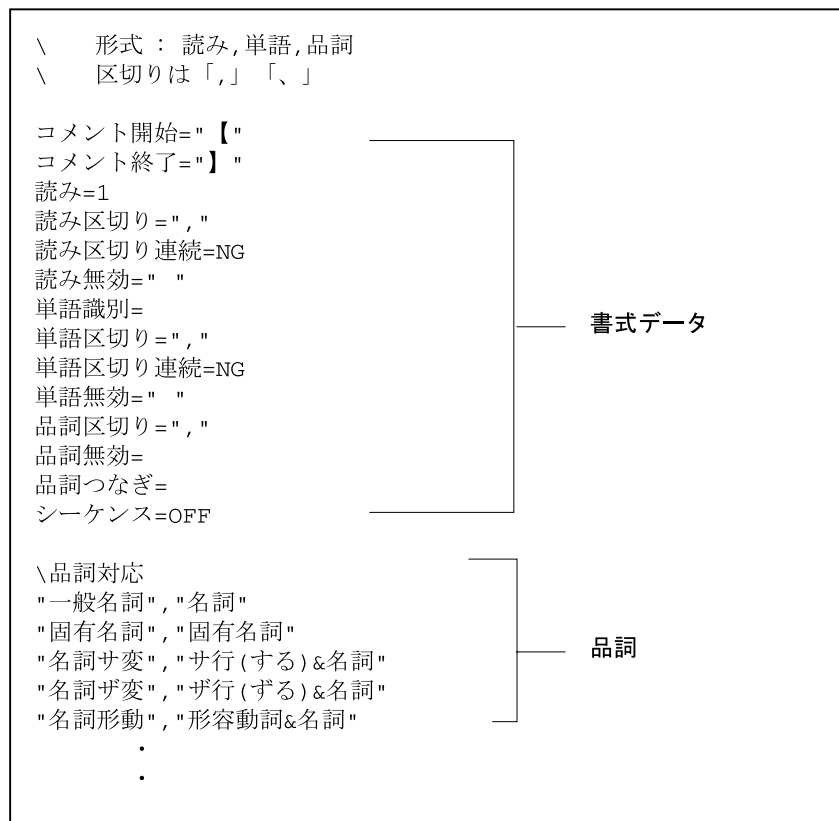
書式ファイルは /usr/lib/locale/ja/wnn/ja/otow.format/ の下に、接尾辞 .fmt が付いて置かれています。

各日本語入力システムの書式ファイル名は表 5-1 のとおりです。

表 5-1 各日本語入力システムの書式ファイル

ATOK8 用書式ファイル	atok8-wnn6.fmt
ATOK7 用書式ファイル	atok7-wnn6.fmt
cs00 用書式ファイル	cs00-wnn6.fmt
VJE-Delta 用書式ファイル	vje-wnn6.fmt
EGBRIDGE 用書式ファイル	egbridge-wnn6.fmt

以下は ATOK7 の書式ファイル例です。「書式データ」と「品詞」は次のように分類されています。



"\" で始まる行はコメント行になり、解釈されません。

書式データ

書式データは、単語ファイルの単語データについて、個々の要素の区切り文字列や最大文字列、読み / 単語に使用される文字で無効とする文字などを定義しています。

書式データのフォーマットは次のとおりです。

キーワード	=	パラメタ
キーワード	=	パラメタ
キーワード	=	パラメタ
.	.	.

・ ・ ・

書式キーワードに対応する定義を ASCII 文字 "=" で指定します。「タブ文字」などの制御文字を使用する場合、実際の制御文字を使用することができます。

また、次の例のように、"\ "に続く一文字、または"\ "に続く 3 桁の 8 進数で、ASCII コードを指定することもできます。

例：

制御コード	8 進数
\a	\007
\t	\011
\n	\012
\v	\013
\f	\014
\r	\015
\"	\042
\'	\047
\\	\104

書式データでキーワードが定義されていない場合は、ATOK7 用の定義が適用されます。

表 5-2 日本語入力システム各キーワードの意味

キーワード	意味
コメント開始	コメント開始部分の文字を指定。最大 4 バイトでコメント開始文字を指定することができる。単語ファイルにコメントがない場合は、パラメタを記述しない
コメント終了	コメント終了部分の文字を指定。最大 4 バイトでコメント終了文字を指定することができる。単語ファイルにコメントがない場合は、パラメタを記述しない。行末は必ずコメントの終了になる
読み	読みの文字幅 (半角、全角) を指定。半角の場合は 1、全角の場合は 2 を指定
読み区切り	読みの終了を示す文字を指定。最大 4 バイトで終了文字を指定することができる。「読み区切り」は複数設定することができる
読み区切り連続	読み区切りの文字を連続して入力できるかを指定。できる場合は OK、できない場合は NG を指定
読み無効	読みに使用される文字で、無効となる文字を指定。最大 4 バイトで無効となる文字を指定することができる。単語ファイルに「読み無効」の文字がない場合は、パラメタを記述しない。「読み無効」は複数指定することができる

表 5-2 日本語入力システム各キーワードの意味 (続き)

キーワード	意味
単語識別	単語を囲む文字を指定。最大 4 バイトで文字を指定することができる。単語ファイルに「単語識別」の文字がない場合は、パラメタを記述しない
単語区切り	単語の終了を示す文字を指定。最大 4 バイトで終了文字を指定することができる。単語ファイルに「単語区切り」の文字がない場合は、パラメタを記述しない。「単語区切り」は複数指定することができる
単語区切り連続	単語区切りの文字を連続して入力できるかを指定。できる場合は OK、できない場合は NG を指定
単語無効	単語に使用される文字で、無効となる文字を指定。最大 4 バイトで無効となる文字を指定することができる。単語ファイルに「単語無効」の文字がない場合は、パラメタを記述しない。「単語無効」は複数指定することができる
品詞区切り	品詞の終了を示す文字を指定。最大 4 バイトで終了文字を指定することができる。単語ファイルに「品詞区切り」の文字がない場合は、パラメタを記述しない。行末は必ず品詞の終了
品詞無効	品詞に使用される文字で、無効となる文字を指定。最大 4 バイトで無効となる文字を指定することができる。単語ファイルに「品詞無効」の文字がない場合は、パラメタを記述しない。「品詞無効」は複数指定することができる
品詞つなぎ	複数の品詞を指定する場合に使用。最大 4 バイトでつなぎとなる文字を指定することができる。単語ファイルに「品詞つなぎ」の文字がない場合は、パラメタを記述しない
シーケンス	エスケープシーケンスの解釈が必要であるかを指定。必要がある場合は ON、必要がない場合は OFF を指定

注 - 「読み区切り」から「単語識別」の間にある文字と、「単語識別」から「単語区切り」の間ある文字は無視されます。「読み区切り」と「読み無効」に同じ文字を指定することはできません。「単語区切り」と「単語無効」、「単語識別」と「単語無効」、「単語識別」と「単語区切り」でも、同様に同じ文字を指定することはできません。

書式データの例：

```
読み区切り=" "
読み区切り="\t"
読み区切り連続=OK
```

上記のように指定した場合、「空白文字」または「タブ文字」で、読みを区切りま
す。

図 5-1 に書式データと単語ファイルの対応例を示します。

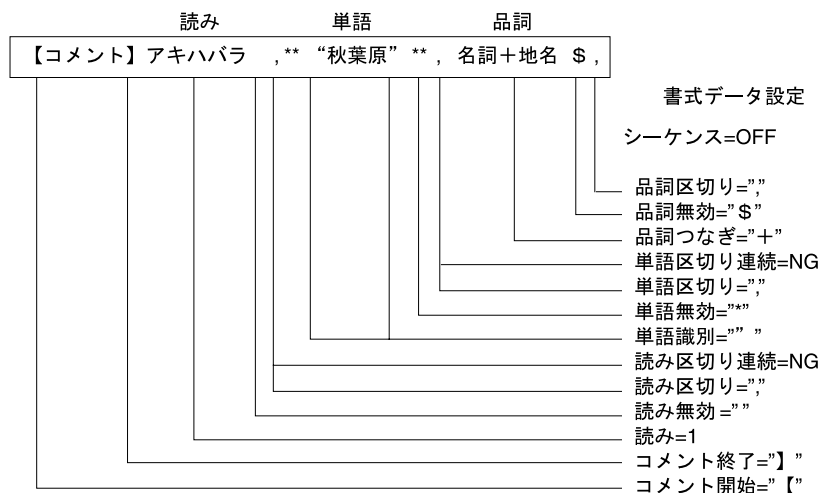


図 5-1 書式データと単語ファイルの対応例

ATOK8、ATOK7、cs00、VJE-Delta、EGBRIDGE の各書式データの設定は次のとおり
です。

表 5-3 各日本語入力システムの書式データ設定一覧

キーワード	ATOK8	ATOK7 (初期設 定)	cs00	VJE-Delta	EGBRIDGE
コメント開始	"!"	" ["	"#"	指定なし	" / "
コメント終了	指定なし	"] "	指定なし	指定なし	指定なし
読み	1	1	2	2	2
読み区切り	","	","	" "	「タブ文字」	「タブ文字」
読み区切り連 続	NG	NG	OK	OK	NG
読み無効	" "	" "	指定なし	指定なし	指定なし
単語識別	"\""	指定なし	指定なし	指定なし	指定なし
単語区切り	","、"	","	「タブ文字」	「タブ文字」	「タブ文字」

表 5-3 各日本語入力システムの書式データ設定一覧 (続き)

キーワード	ATOK8	ATOK7 (初期設定)	cs00	VJE-Delta	EGBRIDGE
単語区切り連続	NG	NG	OK	OK	NG
単語無効	" "	" "	指定なし	指定なし	指定なし
品詞区切り	","	","	指定なし	","	「タブ文字」
品詞無効	"\$" "*"	指定なし	指定なし	"*"	指定なし
品詞つなぎ	指定なし	指定なし	":"	指定なし	"+"
シーケンス	OFF	OFF	OFF	OFF	OFF

品詞

単語ファイルで指定した辞書の品詞と Wnn6 での辞書の品詞の対応を示します。

品詞対応の書式は次のとおりです。

入力辞書の品詞,	Wnn6 の品詞
入力辞書の品詞,	Wnn6 の品詞
入力辞書の品詞,	Wnn6 の品詞
.	.

単語ファイルの品詞名称に対応する Wnn6 の品詞名称を、ASCII 文字 "," で指定します。

■ 例：単語ファイル ATOK7 の場合

"一般名詞",	"名詞"
"固有名詞",	"固有名詞"
"名詞サ変",	"サ行(する)&名詞"
.	.

■ 例：単語ファイル VJE-Delta の場合

"【カ五】",	"力行五段"
---------	--------

"【ガ五】",	"ガ行五段"
"【サ五】",	"サ行五段"
.	.

Wnn6 テキスト形式辞書

Wnn6 で使用できるテキスト形式辞書です。

Wnn6 テキスト形式辞書の書式は、次のとおりです。

\comment			
\total			
\hinsi			
読み	単語	品詞	頻度
読み	単語	品詞	頻度
読み	単語	品詞	頻度
.	.	.	.

例：

\comment			
\total			
\hinsi			
あ	亜	単漢字	1
あ	あ	カ行五段	1
あ	あ	ワ行五段	1
あ	会	ワ行五段	1
あ	合	ワ行五段	1
あ	阿	単漢字	1
.	.	.	.
.	.	.	.

ログファイル

単語情報の変換に失敗した単語を記録します。

ログファイルの書式は、次のとおりです。

変換できない行	←	エラー番号
変換できない行	←	エラー番号
変換できない行	←	エラー番号
・	・	・
・	・	・

「変換できない行」には、単語ファイル中の変換不可能な単語が存在する行が入りません。

エラー番号とその原因は表 5-4 のとおりです。

表 5-4 単語情報の変換 エラー番号と原因対応表

エラー番号	原因
1	「読み」の長さが 253 を超えている
2	「読み」の長さが 0
3	単語部分が識別できない。 書式データ：単語区切りなどを確認する
4	「単語」の長さが 253 を超えている
5	「単語」の長さが 0
6	単語識別が書式データと不適合
7	品詞部分が識別できない。 書式データ：品詞区切りなどを確認する
8	品詞が不適切
9	品詞定義の形式が不適合

付録 A

Solaris 1.x アプリケーション

Solaris 1.x アプリケーションで日本語を入力する場合は、入力方式として `kkcv/ccv` を使用します。Solaris 1.x 用アプリケーションでは、複数言語文字入力や ATOK12、Wnn6 は使用できません。

`kkcv/ccv` は標準の状態では起動しません。Solaris 1.x アプリケーションで日本語を入力するには、以下の手順で設定します。

1. システム管理者 (**root**) の権限で `/etc/css.conf` を編集し、`/usr/lib/css.d` の行を挿入します。

編集後の `/etc/css.conf` は、次のようになります。

```
# "css.conf"      CSs configuration file
#
/etc/css.d
/usr/lib/css.d
# End of "css.conf"
```

2. `/usr/sbin/cssd` を起動します。

```
# sh /etc/init.d/loc.ja.cssd stop
# sh /etc/init.d/loc.ja.cssd start
```


索引

数字・記号

*.dic, 36
*.fmt, 36
*.fzk, 36
*.h, 38

A

Aa 変換, 21
ATOK7, 45
 ATOK7 辞書形式, 47
 書式データ設定一覧, 54
 書式ファイル, 50
atok7-wnn6.fmt, 50
ATOK7 辞書形式, 47
ATOK7 用書式ファイル, 50
ATOK8, 41, 45
 ATOK8 辞書形式, 47
 書式データ設定一覧, 54
 書式ファイル, 50
atok8migd, 41
atok8-wnn6.fmt, 50
ATOK8 用書式ファイル, 50
ATOK12, 11, 29
 移行, 41
 概要, 12
 ファイル構成, 29
atok12migd, 41
atok12migs, 42

B

bunsetsu, 38
bushu, 38
bushu.dic, 36
bushu.dic.no_JISX0202, 36
bushu.fzk, 36

C

cs00, 41, 45
 cs00 辞書形式, 48
 書式データ設定一覧, 54
 書式ファイル, 50
cs00-wnn6.fmt, 50
cs00toatok8, 42, 43
cs00 辞書形式, 48
cs00 用書式ファイル, 50
cvt_key_tbl, 35
cvt_xim_tbl, 35

D

VJE-Delta 辞書形式, 49
dpkeyallow, 34
dpkeylist, 34
dpkeyserv, 35
dpkeyservlist, 35
dpkeystat, 35

E

EGBRIDGE, 45
EGBRIDGE 辞書形式, 49
書式データ設定一覧, 54
書式ファイル, 50
egbridge-wnn6.fmt, 50
EGBRIDGE 辞書形式, 49
EGBRIDGE 用書式ファイル, 50
/etc/css.conf, 59

F

fisd, 36
fiud, 38
FI 学習, 17, 22
FI 学習機能, 16
FI 関係辞書, 17, 18
FI 関係用例, 18
FI 変換, 17, 18

G

GUI ユーティリティ, 17

H

hinsi.data, 36

I

iconv, 46
iwanami, 38

J

jserver, 23, 35
jserver_dir, 38
jserverrc, 35, 38

K

K15Wnn6, 35

M

mode, 36
muhenkan, 38

O

offlinerc, 35

S

S94Wnn6, 35
S509Wnn6, 37
sdtudicm, 42
serverdefs, 35
setdic, 38
setfuzokugo, 38
sethindopath, 38
setjishopath, 38
Solaris CDE, 11
Solaris 共通デスクトップ環境, 11
Solaris の入力方式, 11

U

ud, 38
udicm, 42, 45
udicmtool, 42
udmerge, 36
udmergerc, 35
/usr/lib/css.d, 59
/usr/sbin/cssd, 59
uum, 24, 35
uumkey, 36
uumrc, 36, 38

V

VJE-Delta, 45

VJE-Delta (続き)
VJE-Delta 辞書形式, 49
書式データ設定一覧, 54
書式ファイル, 50
VJE-Delta 用書式ファイル, 50
vje-wnn6.fmt, 50

W

Wnn6, 11
Wnn6 設定ユーティリティ, 24
移行, 45
概要, 16
カスタマイズ, 24
クライアント / サーバー方式, 17
辞書コンバータ, 35
辞書の追加, 17
システムのディレクトリ構成, 33
システムのファイル一覧, 34
テキスト形式辞書, 56
ファイル構成, 33
Wnn6.bitmap, 37
Wnn6/Http, 23, 24
Wnn6 辞書ユーティリティ, 46
Wnn6 システムのディレクトリ構成, 33
Wnn6 システムのファイル一覧, 34
Wnn6 テキスト形式辞書, 56
Wnn6 の概要, 16
Wnn6 の構成, 23
Wnn6 のシステム構成, 24
Wnn6 ライセンスサーバー, 34
wnnaccess, 37
wnnatod, 35
wnnbushu, 37
Wnnbushu, 37
Wnndictutil, 37
wnndictutil, 24, 46
wnnds, 23, 24
wnndtoa, 35
wnnenvrc, 36, 38
Wnnenvutil, 37
wnnenvutil, 24, 37
wnnenvutil.hlp, 37
wnnhosts, 35
wnnkill, 37
wnnoffline, 36
wnnotow, 24, 35, 46

wnnstat, 35
wnntouch, 35
wnnudmerge, 36

X

ximrc, 36
xjsi.so, 37

い

異形字処理, 21
移行
ATOK8 への移行, 41
ATOK12 への移行, 41
Wnn6 への移行, 45

う

受身変換, 19

え

英単語日本語変換, 22

お

送りがな基準学習, 23
送りがな基準処理, 21
オフライン学習, 17, 23, 35

か

格係り受け変換, 18
学習機能, 22
カスタマイズ, 24
かな / カナ / ローマ字変換, 22
かな漢字変換機能, 18
かな漢字変換サーバー, 35
環境設定ファイル, 42

き

キーバインド, 17
キーバインド定義, 36

こ

合成語変換, 19
固定形式, 17
コメント開始, 52
コメント終了, 52

さ

最終使用最優先変換, 20

し

シーケンス, 53
使役変換, 19
辞書, 17
辞書コンバータ, 24, 25
辞書引きサーバー, 35
辞書ユーティリティ, 25, 37
システム辞書, 語彙数, 17
修飾語変換, 19
書式データ, 51, 54
書式データ設定一覧, 54
書式ファイル, 50
所有格変換, 19

す

数字学習, 23

せ

接頭語学習, 23
接尾語学習, 23

た

単漢字変換, 20
単語区切り, 53
単語区切り連続, 53
単語識別, 53
単語ファイル, 47
単語無効, 53

ち

長音変換, 20

て

電話番号変換, 21

と

登録可能形式, 17
登録語自動反映, 23, 35

に

入力スタイル, 17

ひ

品詞, 55
品詞区切り, 53
品詞つながり, 53
品詞無効, 53
頻度学習, 22
頻度情報, 17

ふ

ファイル形式, 46
ファイル構成, ATOK12, 29
複合語変換, 20
複文変換, 20
文節間接続関係辞書, 17, 18

文節切り学習, 23

む

無変換学習, 22

ゆ

ユーザーが登録した単語情報, 45

ユーザー辞書, 18, 38

ユーザー辞書ファイル, 24

郵便番号変換, 21

揺らぎ処理, 20

よ

読み, 52

読み区切り, 52

読み区切り連続, 52

読み無効, 52

れ

連濁処理, 22

ろ

ログファイル, 57

わ

ワークスペースメニュー, Wnn6, 33

