Windows版
辞書ブラウザ
(EDBroW)
利用マニュアル

EDR

本ドキュメントに記載されているすべてのソフトウエア，辞書データ，およびそれに付随するデータ類は契約に許された目的以外での使用は認められておりません．

©Japan Electronic Dictionary Research Institute, Ltd. 1995 All rights reserved.

本利用マニュアルおよび利用マニュアル内に記載されているソフトウエア(EDBroW, 辞書トランスレータ等)は，(株)日本電子化辞書研究所にすべての権利が帰属します．
当社の書面による同意なしには本利用マニュアルの一部または全部を問わず，複製および転載することを禁じます．

UNIX は米国AT&T の登録商標です．

Sun-workstationおよび SPARCstationは米国 Sun Microsystem社の登録商標です．

Microsoftとそのロゴ，MS，MS-DOS は米国マイクロソフト社の登録商標です．
Windows, Visual C++ は米国マイクロソフト社の商標です．

Intelは米国インテル社の登録商標です．

本ドキュメントに記載されている辞書データ，および記載内容は事前の予告なく変更されることがあります．

EDR 株式会社 日本電子化辞書研究所
東京都千代田区神田佐久間河岸 78-1 Tel:03-3851-5521 Fax:03-3851-5840

目次

第 1 章　はじめに 1

1.1　システム概要 1
1.2　システム構成 1
1.3　動作環境 2
1.4　用語，表記法の定義 3
1.4.1　用語 3
1.4.2　表記法 5

第 2 章　インストール手順 6

2.1　ファイルの展開 6
2.2　ブラウザのインストール 6
2.3　辞書トランスレータのインストール 7
2.4　ソフトウェア構築手順 7
2.4.1　ディレクトリの構成 8
2.4.2　辞書トランスレータの構築 8
2.4.3　辞書検索ライブラリの構築 8
2.4.4　ブラウザの構築 9

第 3 章　辞書構築手順 10

3.1　辞書構築の概要 10
3.1.1　テキスト辞書のマージについて 10
3.1.2　辞書構築の流れ 11
3.3　Windows上でのシステム辞書構築 11
3.3　UNIX 上でのシステム辞書構築 14
3.4　注意点 15
3.4.1　共起辞書の構築手順 15
3.4.2　概念見出し辞書の構築 17

第 4 章　利用マニュアル 18

4.1　起動の準備 18
4.2　システムのモード 20
4.2.1　表示モード 20
4.2.2　その他のモード 21
4.3　システムの起動 22

4.4 辞書の検索 25
4.4.1 単語辞書検索 26
4.4.2 対訳辞書検索 30
4.4.3 共起辞書検索 33
4.4.4 概念体系辞書検索 36
4.4.5 概念記述辞書検索 37
4.4.6 概念見出し辞書検索 38
4.4.7 上位概念検索 40
4.4.8 可能関係子検索 41
4.4.9 関連検索 44
4.5 ユーザ情報の操作 46
4.5.1 ユーザ情報レコードの変更 47
4.5.2 ユーザ情報レコードの新規作成 52
4.5.3 ユーザ情報レコードの削除 52
4.5.4 ユーザ情報レコードのコピー 53
4.5.5 取消 53
4.5.6 キーワードの設定 53
4.5.7 ユーザ情報のクリア 53
4.5.8 ユーザ情報の保存 54
4.5.9 ユーザ情報の読み込み 54
4.5.10 ユーザ情報の印刷 55
4.6 その他の操作 55
4.6.1 モードの変更 55
4.6.2 辞書の切換 56

第 5 章 システム辞書仕様 57

5.1 各辞書に共通な事柄 57
5.2 単語辞書 58
5.3 対訳辞書 58
5.4 共起辞書 59
5.5 概念見出し辞書 60
5.6 概念体系辞書 60
5.7 概念記述辞書 60
5.8 コード一覧 60
5.8.1 日本語左連接情報コード 61
5.8.2 日本語右連接情報コード 61
5.8.3 日本語品詞コード 62
5.8.4 英語左連接情報コード 63
5.8.5 英語右連接情報コード 63
5.8.6 英語品詞コード 64

5.8.7 共起辞書品詞コード 65
5.8.8 概念関係子コード 65
第 6 章 ライブラリ解説 66
6.1 概要 66
6.2 関数一覧 68
6.2.1 辞書のオープン / クローズ用関数 68
6.2.2 辞書検索関数 68
6.3.3 フィルタキー用関数 69
6.3.4 辞書レコードアクセス関数 69
第 7 章 ファイルフォーマット 74
7.1 システム辞書 74
7.1.1 基本構造 74
7.1.2 単語辞書 76
7.1.3 対訳辞書 78
7.1.4 共起辞書 78
7.1.5 概念体系辞書 79
7.1.6 概念記述辞書 80
7.1.7 概念見出し辞書 81
7.2 ユーザ情報ファイル 81
7.2.1 基本構造 81
7.2.2 レコード本体の構造 82

第 1 章

はじめに

1.1 システム概要

Windows版辞書ブラウザ (Electronic Dictionary Browser for Windows以下，EDBroW と呼ぶ) は，Microsoft Windows 3.1上で動作する，EDR 電子化辞書のブラウジングシステムです．

EDBroW は，以下のような特徴を持っています．

- CDROM によりテキスト形式で提供される EDR 電子化辞書を，専用の辞書トランスレータによりパソコン上から検索できる形に形式変換します．また，変換された辞書を検索するライブラリも提供しますので，これをパソコン上での EDR 辞書検索 (応用システムなど) に利用することができます．
- すべての辞書を直接検索できるだけでなく，ある辞書の検索結果をもとに，他の辞書を関連検索することが可能です．
- EDR 電子化辞書の特徴の一つである概念辞書を使った，複合検索が可能です．すなわち，ある概念が，概念体系上でどのような位置にあるかを表示する「上位概念検索」や，概念記述と概念体系の継承機能を利用して，特定の 2 概念間にどのような概念関係が成り立ち得るかを調べる「可能関係子検索」が利用できます．
- 辞書の検索結果を，「ユーザ情報」としてバッファに取り込むことができます．取り込まれた個々の辞書データは編集することが可能であり，また，各辞書データ各々に対して，参照情報，覚書情報を付加することができます．ユーザ情報は，ひとまとめにしてファイルに保存することが可能です．

1.2 システム構成

図1.1に，システム構成を示します．

EDR 電子化辞書はテキスト形式で配布されます．これを，EDBroW から利用できる形式 (システム辞書) に変換する必要があります．このために用意されているのが，辞書トランスレータです．辞書トランスレータは，UNIX ワークステーション，または Windows環境下で，オフラインで動作します．

辞書の変換方法については，第 3 章を参照して下さい．

EDBroW の基本機能は，システム辞書の検索です．Windows上に実現されている検索用インタフェースから，辞書検索ライブラリを呼び出し，システム辞書を検索します．検索結果は，表示用インタフェースにより表示されます．

EDBroW のもう一つの機能は，辞書レコードの取り込みと編集，保存などの機能です．EDBroW では，辞書検索の結果をシステムの内部領域に複数取り込むことができます．この取り込まれた辞書レコードのセットに，キーワードを付けた一まとまりのことを，ユーザ情報と呼びます．

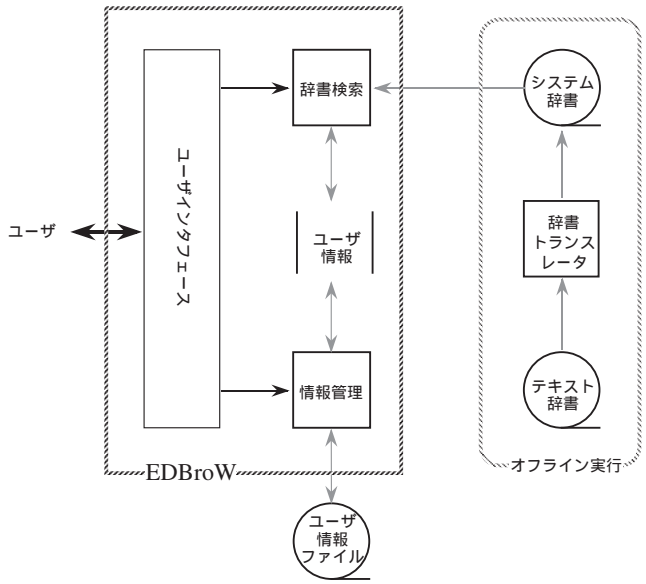


図 1.1 システム構成

ユーザ情報は，以下のような目的で利用することができます．

- 辞書の検索結果をユーザ情報として保存し，これをもとに，他の辞書の関連項目を検索する．
- 辞書データにエラーがある場合，ユーザ情報中で修正したデータを保存しておき，これをフィードバック情報として辞書の改良に利用する．
- あるテーマに関連する辞書情報をまとめて保存する．

EDBroW では，上記の機能を実現するため，ユーザ情報レコードの変更 / 削除 / コピー機能や，他のレコードに対する参照情報の付与機能，レコード自体への覚書情報の付与機能などを有しています．また，ユーザ情報は，ファイルに保存したり，ファイルから読み込むことも可能です．

1.3 動作環境

パソコンにおけるシステムの動作環境は，以下の通りです．

ハードウェア

本体　　Microsoft Windows 3.が動作するもの
（CPU: 486DX2 以降を推奨）
メモリ　　8MB以上（16MB以上を推奨）
HDD　　空き領域最大 500MB（作成する辞書の構成により異なる）
CDROM ドライブ

または UNIX ワークステーション上のファイルを取得 / 利用できる環境

ワークステーション上で辞書を構築する場合には，それをパソコン上で利用

する手段 (ftp などで取得できる，あるいは NFS マウントできる) が必要です．そうでない場合には，パソコン上で CDROM を利用できる環境が必要となります．

ソフトウェア

Microsoft Windows 3.1
Microsoft Visual C++ 1.0 Professional Edition（ソフトウェア構築を行う場合にのみ必要）

また，ワークステーション上のツールは，以下の環境で動作確認を行っています．

ハードウェア

本体	Sun SparcStation SS10
メモリ	32MB
OS	SunOS 4.1.3+ JLE 1.1.3 (JLE はなくても良い)
HDD	空き領域最大 500MB（作成する辞書の構成により異なる）

ソフトウェア

SunOS 4.1.3附属の C コンパイラ もしくは gcc 2.5.8
（ソフトウェア構築をする場合にのみ必要）
※ SparcStation SS10+ Soralis2.3でも動作を確認しています

辞書のサイズ，作業に必要な領域等については，「第 4 章 辞書構築手順」を参照して下さい．

1.4 用語，表記法の定義

以下に，本マニュアルにおける基本的な用語，表記法について解説します．なお，Windowsの基本的な操作についてはここでは触れません．詳しくは，Windowsのマニュアル等を参照して下さい．

1.4.1 用語

操作

クリック	EDBroW では，クリックはすべてマウスの左ボタンを使います．
コンボボックス	通常のテキスト入力も可能ですが，テキストボックスの右側にある【↓】ボタンをクリックすると，入力可能な文字列の一覧が表示され，マウスで選択することにより入力が可能です．EDBroW では，品詞，連接情報，概念識別子，概念関係子などの入力にコンボボックスを利用しています．
リストボックス	EDBroW では，辞書検索結果の表示のほとんどに，リストボックスを利用しています．辞書の種類により，リストボックスの動作が異なります．詳しくは第 4 章を参照して下さい．

ブラウザに固有の用語

検索キー　辞書を検索する際に，最低一つは入力しなければならないキーです．複数指定することも可能です．検索キーは，システム辞書中に作成されているインデックスに対応します．各辞書の検索キーについては，第４章または第７章を参照して下さい．

フィルタキー　検索キーの補助として，検索結果の絞り込みを行う際に利用可能なものです．各辞書によってキーの種類が異なります．詳しくは第４章，第７章を参照して下さい．

レコード　一つの語および一つの概念に対応するデータの単位です．EDR 電子化辞書では，一つの単語に対して複数のレコードが存在することが多くあります．

なお，辞書検索の結果表示の際のデータ単位を，「辞書レコード」と呼び，これがユーザ情報に取り込まれた際のデータ単位を，「ユーザ情報レコード」あるいは単に「レコード」と呼びます．

フィールド　辞書レコードや，ユーザ情報レコード中の一つの情報（例えば単語辞書レコードの単語見出し）のことをフィールドと呼びます．

不変化部分　単語において，規則変化語が語形変化する場合の，純粋に文字列としての不変化部分であって，必ずしも語根を意味しません．たとえば，'wife/wivesは規則変化として扱われており，その不変化部分は　'wi' です．これに対して，規則変化部分の基本形（終止形・単数形・現在形など）のことをEDBroW では「語尾」と呼びます．また，「語尾つき」とは，'不変化部分＋語尾の文字列' を意味します．

例

単語（語尾つき）	不変化部分	語尾
wife	wi	fe
come	com	e
引く	引	く

始点 / 終点　概念記述レコードは，以下のような形式で記述します．

〈概念識別子１〉−〈関係子〉→〈概念識別子２〉

ここで，概念識別子１のことを「始点側」，概念識別子２のことを「終点側」と呼びます．始点 / 終点のそれぞれの意味づけは，関係子によって異なりますが，例えば関係子 "agent"の場合には，始点側が動作に相当する概念，終点側がその動作を行う主体に相当する概念となります．

概念体系レコードの場合にも同様の表記を使います．この場合，始点側が下位概念，終点側が上位概念となります．

また，共起辞書の「表層共起情報」においても，同様の表記を用いますが，

この場合には，受け側 / 係り側と呼びます．なお共起辞書においても，意味情報では，始点 / 終点という表現を使いますが，表層の受け側 / 係り側と，意味情報の始点 / 終点は，必ずしも対応しません．

センシティブ　選択可能なボタン / メニュー，入力 / 参照可能なテキストボックスなどを "センシティブである" といいます．逆に，選択できないボタン / メニュー，入力 / 参照できないテキストボックスなどを "センシティブでない" といいます．センシティブでないものは，淡色表示となります．

1.4.2 表記法

メニュー　EDBroW は，メインウィンドウのメニュー操作によるコマンド入力で動作します．このコマンド入力を以下のように記します．

[ファイル] – [終了]

これは，トップの [ファイル] メニューをマウスで選択し，メニュー一覧の中から，[終了] をマウスで選択することを意味します．

さらにサブメニューがある場合には，

[ユーザ情報編集] – [新規作成] – [単語レコード作成]

のように表記します．

ボタン　ダイアログボックスなどで利用するボタンは，【OK】のように表記します．

第 2 章

インストール手順

2.1 ファイルの展開

まず，インターネットから取得したファイルを展開します．ディレクトリには，以下のものが含まれています．

ディレクトリ名	内容
browser	ブラウザ本体 + 設定ファイル
dtrans	辞書トランスレータ
src	ソースコード一式

Windowsからのシステム起動には，以下の 2 つの方法があります．

ファイルマネージャから起動

プログラムマネージャから起動

ファイルマネージャから起動する場合には，適当なディレクトリに実行ファイルと設定ファイル (edbrow.exe, edbrow.ini, code.tbl) をコピーするだけで，他に必要な作業はありません．ただし，起動には，まずファイルマネージャを起動して，必要な実行ファイルを探す，という手間が必要です．頻繁に起動する必要のあるツールは，プログラムマネージャに登録し，ここから起動する方が便利です．

以下のインストール手順では，システムをプログラムマネージャに登録する方法を述べます．

2.2 ブラウザのインストール

browserディレクトリにある，edbrow.exeが実行形式です．

最初に，ファイルマネージャなどで，browserディレクトリにある全てのファイルを適当なディレクトリにコピーした後，Windows 3.1上でアプリケーション登録を行って下さい．

まず，必要があればグループを作成します．グループ作成の手順は，Windowsのマニュアルを参照して下さい．その後，EDBroW を登録します．手順は以下の通りです．

- プログラムマネージャを起動し，登録したいグループを開く．
- [アイコン] – [登録とグループの作成] を選択する．
- ダイアログボックス上で，「アイコンの登録」を選び，【OK】をクリックする．
- 図 2.1 に示すダイアログボックスで，必要な情報を入力する．

最低限，「コマンドライン」に，edbrow.exeのフルパス名を入力すれば OK です．必要ならば，タイトル，ショートカットキーなども指定して下さい．その際，「実行時のディレクトリ」は

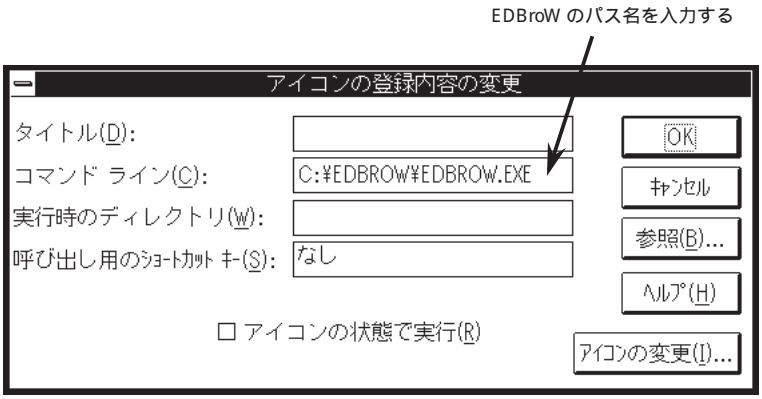


図 2.1 アイコンの登録

入力しないで下さい（自動的に，edbrow.exeのあるディレクトリになります）．実行時のディレクトリを指定することも可能ですが，その場合，指定されたディレクトリに，システムの初期化ファイル（EDBROW.INI ）がなければなりません．また，ユーザ情報ファイルも，実行時のディレクトリに格納されることになります．

最後に，【OK】をクリックすれば，登録は完了です．

以降は，プログラムマネージャからアイコンをダブルクリックすることにより，システムを起動できるようになります．

2.3 辞書トランスレータのインストール

Windows版では，dtransディレクトリにある，dtrans.exeを適当なディレクトリにコピーし，ブラウザのインストールと同様の手順で，アプリケーション登録を行って下さい．その際，実行時のディレクトリには注意して下さい．Windows版辞書トランスレータでは，作成される辞書のパス名を指定することもできますが，デフォルトでは，辞書は dtrans.exeがあるディレクトリに作成されます．

UNIX 版では，dtransディレクトリにあるファイル dtransを適当なディレクトリにコピーし，このディレクトリにパスを通して下さい．なお，このディレクトリにある dtransは，Sun SparcStation + SunOS-4.1.用にコンパイルしたものです．これは，Sun SparcStation + Solaris 2.3 でも動作は確認しています．それ以外の UNIX ワークステーションでは，トランスレータをコンパイルし直す必要があります．

トランスレータの構築方法については，2.4を参照して下さい．

2.4 ソフトウェア構築手順

本節では，ソフトウェアの構築手順について記します．通常は，提供されている実行形式のファイルで問題なく動作しますので，本節の説明は読まなくても結構です．

ソースコードを変更したい場合，あるいは，Sun SparcStaion以外の UNIX ワークステーションで dtransを利用したい場合には，ソフトウェアを再構築する必要があります．その場合は，以下の手順に従って構築して下さい．

2.4.1 ディレクトリの構成

srcディレクトリには，以下のものが含まれています．

ディレクトリ名	内容
browser	EDBroW ソースコード一式
dtrans	辞書トランレータ・ソースコード一式
→ unix	UNIX 版トランスレータ
→ windows	Windows版トランスレータ
search	辞書検索ライブラリ・ソースコード一式

Windows上でのソフトウェア構築には，Microsoft Visual C++ Ver 1.0 Professional Editionが必要です．

2.4.2 辞書トランスレータの構築

辞書トランスレータは，UNIX 版と Windows版の2種類があります．それぞれにソースを用意してありますが，ファイルが MS-DOS 形式 (SJIS) であるか，UNIX 形式 (EUC) であるか，の違いしかありません．

UNIX 版

makeで，dtransコマンドが作成されます．SunOS-4.1.3附属の Cコンパイラ，および gcc-2.5.8 で動作確認を行っています．

Windows 版

- Windows 3.1上で，Visual C++ を起動する．
- プロジェクト dtrans.makを開く
- ビルドする．

※ プロジェクトを開く際に，「プロジェクトが元のディレクトリから移動しています．依存関係を生成します」というメッセージが表示されますが，問題ありません．

2.4.3 辞書検索ライブラリの構築

必要があれば，ブラウザ構築の前に辞書検索ライブラリを構築します．提供されるソースでは，すでに構築されたライブラリ (edsrch.lib) がディレクトリに含まれていますので，ブラウザのみの再構築の場合は，ライブラリの構築は必要ありません．

ライブラリの作成手順は，以下の通りです．

- Windows 3.1上で，Visual C++ を起動する．
- プロジェクト edsrch.makを開く．
- ビルドする．

これで，edsrch.libが作成されます．

※ なお，edsrch.libは，EDBroW からの利用を前提に作られていますので，それ以外の用途で利用する場合には不都合が生じることがあります．したがって，他の用途で利用する場合には，ライブラリの再構築が必要です．詳しくは，ライブラリのディレクトリにある，readme.{euc,sj}を参照して下さい．

2.4.4 ブラウザの構築

以下の通りです．(先に，辞書検索ライブラリが作成されていることを確認して下さい)

- Windows 3.1上で，Visual C++ を起動する．
- プロジェクト edbrow.maを開く
- ビルドする．

※プロジェクトを開く時に，「外部メイクファイルですか」と聞かれますので，【はい】をクリックして下さい．

第 3 章

辞書構築手順

EDBroW 利用にあたっては，最初に CD-ROM で用意されたテキスト形式の辞書（以下，テキスト辞書と呼ぶ）を，EDBroW から利用できる形式の辞書（以下，システム辞書と呼ぶ）に変換します．変換ツールである辞書トランスレータは，UNIX 上，もしくは Windows上で動作します．ツールは，dtrans以下にあります．dtrans.exeが，パソコン用の辞書トランスレータ，dtransが Sun SparcStation用の辞書トランスレータです．

辞書トランスレータは，当初 UNIX 上でのみ動作することを念頭に開発し，その後，パソコン上に移植しました．MS-DOS では，利用可能なメモリに制限があるため，パソコン用トランスレータは，QuickWin アプリケーション*として開発しました．したがって，辞書構築手順は，UNIX 版，パソコン版で大差はありません．メモリ容量や CPU の速度にもよりますが，UNIX 版の方が比較的高速に処理できます．

3.1 辞書構築の概要

3.1.1 テキスト辞書のマージについて

提供されるテキスト辞書は，格納されているデータにより，単語辞書，対訳辞書，共起辞書，概念辞書に大別されます．概念辞書はさらに，概念体系辞書，概念記述辞書，概念見出し辞書の 3 種類に分かれています．

概念辞書を除く各辞書は，その言語識別（日本語 / 英語，対訳の場合は日英 / 英日）に応じて別ファイルとなります．また，専門用語（情報処理）の辞書も別ファイルとなっています．したがって，単語辞書であれば，最大で，日本語単語辞書，英語単語辞書，日本語専門用語単語辞書，英語専門用語単語辞書，の 4 種類となるわけです．ユーザによっては，日本語単語辞書だけを購入した，という場合もあるでしょう．

EDBroW では，一度に扱えるシステム辞書は種類ごとに 1 つです．すなわち，単語辞書が 1 つ，対訳辞書が 1 つという形です．このため，辞書トランスレータでは，複数ある同種のテキスト辞書をマージして，一つのシステム辞書として作成する機能を有しています．

概念辞書を除くシステム辞書は，言語識別，基本語 / 専門用語の別を格納するように設計されており，辞書をマージした場合でも情報は失われません．また，ブラウザでそれらの種別を指定して検索することも可能です．

> ※ 概念見出し辞書，概念体系辞書は，基本語 / 専門用語の別に応じて，2 種類のテキスト辞書が提供されますが，この種別を格納するように設計されてはいません．これは，基本語 / 専門用語で概念識別子が重複することはないため，ユーザの混乱は招かない，という理由によります．

* Windows上で，stdin/stdout/stderrが利用可能なアプリケーションで，コマンドラインオプションが利用できないことを除いて，GUI を持たないアプリケーションをそのまま Windows上で動作させることが可能なものです．

辞書をマージして作成するか，テキスト辞書ごとに別の辞書として作成するか，はユーザの自由です．EDBroW では，起動時に初期化ファイルに記述されたデフォルトのシステム辞書を開くようになっていますが，起動後にシステム辞書を切り換えることもできます．

辞書をマージした場合，例えば単語辞書を概念識別子で検索することにより，その概念につながる日本語 / 英語双方の単語を一度に表示することができるようになります（対訳辞書の場合も同様）．逆に，辞書をマージすると，辞書サイズが大きくなる分，検索速度が若干低下しますが，さほど問題はありません．

3.1.2 辞書構築の流れ

デフォルトでは，システム辞書は辞書トランスレータを実行しているカレントディレクトリに作成されます（ただし，オプションなしの起動では，パス名を直接指定することが可能）．また，作業領域も，デフォルトではカレントディレクトリですが，環境変数により，別のディレクトリを指定することもできます．指定する環境変数は，DTRANS_TMP　です．

UNIX の場合は，コマンドラインから，

```
% setenv DTRANS_TMP /work
```

などと指定するか，.cshrcなどで指定して下さい．また，Windowsの場合には，起動ドライブの autoexec.batに，以下のような一行

```
set DTRANS_TMP D:\TMP
```

を追加し，パソコンを再起動して下さい．

必要な作業領域は，辞書の種類にもよりますが，最大でも作成されるシステム辞書サイズの 2 倍以内です．また，システム辞書サイズは，最大でも元となるテキスト辞書サイズを越えることはありません．

> ※ 概念見出し辞書の場合のみ，システム辞書サイズはテキスト辞書サイズ（複数ある場合は合計）の 75% 程度になりますが，それ以外の辞書では，40% ~ 60% 程度になります．

辞書の構築手順は，おおまかに言って以下のようになっています．

- 最初に作成するシステム辞書の種類を指定する．
- システム辞書の種類に対応するテキスト辞書を，必要な数だけ指定する．
- テキスト辞書の読み込みが完了すると，検索のためのインデックスなどを作成し，システム辞書が完成する．

共起辞書のみ，他の辞書と構築手順が若干違います．これについては，3.4.1を参照して下さい．

次節に，単語辞書を例にとって，システム辞書構築手順を詳説します．

3.3 Windows 上でのシステム辞書構築

Windows上の辞書トランスレータは，インタラクティブに動作します．辞書の種類，パス名な

どを指定するプロンプトが表示されますので，必要な情報を入力していって下さい．

なお，途中で終了する場合には，q を入力して下さい．処理の途中で，Ctrl+C や [ファイル] – [終了] コマンドで終了すると，作業ファイルや作成途中のシステム辞書ファイルが消去されずに残ってしまうことがあります．

※途中で中断した場合には， fdc?????.tmp, idx?????.tm0, idx?????.tm1, srt?????.tmpなどのファイル (????? はプロセスID) が残ることがあります．これらはかなり大きなファイルとなりますので，消去して下さい．

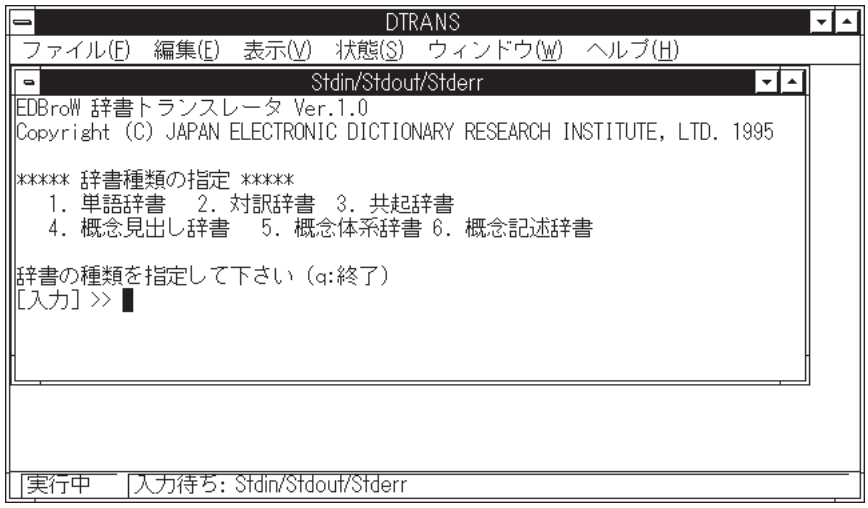


図 3.1 辞書トランスレータ起動画面

それでは，実際に単語辞書を作成してみます．

まず Windowsを起動し，プログラムマネージャで，dtransのアイコンをダブルクリックして下さい (または，ファイルマネージャから dtrans.exeをダブルクリック) ．辞書トランスレータが起動し，図3.1 のウィンドウが表示されます．

最初に，

辞書の種類を指定して下さい (q : 終了)
[入力] >>

のプロンプトで，どのシステム辞書を作成するかを選択します．作成したい辞書の番号を入力し，リターンキーを押して下さい．ここでは，単語辞書作成なので，1 を入力します．

続いて，

作成する辞書のパス名を入力して下さい
(デフォルト：カレントディレクトリに wd_sys.dicを作成)
[入力] >>

というプロンプトが表示されます．リターンのみを入力すると，カレントディレクトリに，wd_sys.dicというシステム辞書ファイルが作成されますが，パス名を指定すると，そのパス名でシステム辞書を作成します．パスの指定が不正である場合には，再度上記プロンプトが表示されますので，正しいパス名を入力して下さい．なお，辞書ファイルのデフォルト名は，辞書

の種類により異ります（→3.3）．

また，すでに同じパス名のファイルがある場合，ファイルを削除してよいかどうかの確認を求めます．ここで y を入力すると，既存のファイルは削除され，新たにシステム辞書ファイルを作成します．n を入力した場合には，システムは処理を終了します．

次に，

テキスト辞書のパス名を入力して下さい (q:指定終了)
[入力] >>

というプロンプトが表示されます．単語辞書の CDROM をセットし，テキスト辞書（CDROM 上）のパス名を入力してリターンキーを押して下さい．テキスト辞書の読み込みを開始します．テキスト辞書の読み込みにかかる時間は，作成する辞書の種類や，計算機のパフォーマンスにより異なります（DX2-66MHz / メモリ24MB，日本語単語辞書のみの場合で 15分程度）．

なお，指定したパスに誤りがある場合や，別の種類のテキスト辞書を指定した場合にはエラーが表示されますので，入力をやり直してください

テキスト辞書の読み込みが完了すると，再び上記のプロンプトが表示されます．別の単語辞書をマージする場合には，CDROM を入れかえ，再度テキスト辞書のパス名を入力して下さい．マージする必要がない場合，あるいはマージするすべてのテキスト辞書の読み込みが完了した場合には，q を入力します．

テキスト辞書の読み込みが完了すると，自動的にシステム辞書の構築にに移ります．ここで，読み込みに要するのと同じくらいの時間がかかります．

辞書の構築が終了すると，ウィンドウ上に以下のようなメッセージが表示されます．

*******単語辞書の作成完了 *******
変換レコード数　390164個
異常レコード数　　0 個

これで，辞書の変換は完了です．処理が終了しても，ウィンドウは消去されませんので，[ファイル] − [終了] コマンドを選択してウィンドウを消去して下さい．

以下に，4 種類の単語辞書をマージして，システム辞書（単語）を作成する場合の表示と入力の例を示します（CDROM ドライブのデバイスを E: ，作成する辞書のパス名を D:¥DICT¥WD_SYS.DIC と仮定，下線部分が入力を示す）．

```
*****辞書種類の指定 *****
  1. 単語辞書  2. 対訳辞書  3. 共起辞書
  4. 概念見出し辞書  5. 概念体系辞書 6. 概念記述辞書
辞書の種類を指定して下さい (q:終了)
[入力] >> 1                              ←辞書種類
*****単語辞書の作成開始 *****
作成する辞書のパス名を入力して下さい
（デフォルト：カレントディレクトリに wd_sys.dicを作成）
[入力] >> D:¥DICT¥WD_SYS.DIC        ←システム辞書パス名
ファイル D:¥DICT¥WD_SYS.DIC を消去しても良いですか [Y/N] y
システム辞書 D:¥DICT¥WD_SYS.DIC を作成します

テキスト辞書のパス名を入力して下さい (q:指定終了)
```

```
[入力] >> E:¥JWD.DIC                    ←テキスト辞書パス名
テキスト辞書のパス名を入力して下さい (q:指定終了)
[入力] >> E:¥EWD.DIC                    ←テキスト辞書パス名
テキスト辞書のパス名を入力して下さい (q:指定終了)
[入力] >> C:¥TJWD.DIC                   ←テキスト辞書パス名 ( 誤り )

[ERROR!]  ファイルがオープンできません : C:¥TJWD.DIC

テキスト辞書のパス名を入力して下さい (q:指定終了)
[入力] >> E:¥TJWD.DIC                   ←テキスト辞書パス名
テキスト辞書のパス名を入力して下さい (q:指定終了)
[入力] >> E:¥TEWD.DIC                   ←テキスト辞書パス名
テキスト辞書のパス名を入力して下さい (q:指定終了)
[入力] >> q                             ←指定終了コマンド
データ読み込み完了

固定長パートの書き込み中...終了
インデックス(#1)の作成開始
  ソート中...終了
  書き込み中...終了
インデックス(#2)の作成開始
  ソート中...終了
  書き込み中...終了
インデックス(#3)の作成開始
  ソート中...終了
  書き込み中...終了

*******単語辞書の作成完了 *******
  変換レコード数    927658個
  異常レコード数          0個
```

対訳辞書，概念辞書の場合も，作成方法はまったく同じです．ただし，インデックスを 3 種類作成するのは単語辞書だけです．

概念辞書の場合，一つの CD-ROM に複数の辞書 (概念体系，概念記述，概念見出し) が入っていますので，パス名に注意して下さい．専門用語の概念辞書 (概念体系，概念見出しのみ) をマージする場合の方法は，単語辞書と同様です．

共起辞書の場合には，基本的には上記と同じ手順ですが，いくつか情報を追加する必要があります．詳しくは，3.4.1に記します．

3.3 UNIX 上でのシステム辞書構築

UNIX 版では，コマンドラインオプションを指定することができます．オプションを指定しない場合はインタラクティブに動作し，辞書の指定方法などはパソコン版とまったく同じですので，省略します．

コマンドラインオプションでは，辞書種別を指定します．この場合，テキスト辞書は stdinから入力します．

オプションと，作成されるシステム辞書の関係は以下のようになっています．

-WD　単語辞書
-BI　対訳辞書
-CC　共起辞書
-CI　概念見出し辞書
-CH　概念体系辞書
-CR　概念記述辞書

同種別ファイルのテキスト辞書をマージする場合には，

```
% cat JWD.TAB EWD.TAB | dtrans -WD
```

のようにして起動して下さい．後は，自動的にシステム辞書が作成されます．なお，コマンドラインオプションを指定して起動した場合には，辞書はカレントディレクトリに以下のファイル名で作成されます（これらは，Windows版 / UNIX 版のオプションなしによる起動時 のデフォルトファイル名でもあります）．

単語辞書	wd_sys.dic
対訳辞書	bi_sys.dic
共起辞書	cc_sys.dic
概念体系辞書	ch_sys.dic
概念記述辞書	cr_sys.dic
概念見出し辞書	ci_sys.dic

すでに同じ名前のファイルがあった場合には，辞書トランスレータは処理を中断します．

3.4　注意点

3.4.1　共起辞書の構築手順

共起辞書には，概念識別子の代わりに「補足付き概念説明」と呼ばれるデータが入っているレコードが多くあります．このデータを通常の概念識別子と同等に扱うためには，共起辞書構築時に，他の辞書とは異なる操作が必要です．すなわち，補足付き概念説明に対して概念識別子を自動付与し，概念識別子と補足付き概念説明文字列との対応を，概念見出し辞書のテキスト形式と同じフォーマットで格納することにより，補足付き概念説明を通常の概念識別子と同じ枠組みで利用しようというもので，これを実現するため若干の操作が必要となるわけです．以下に，その手順を説明します．

辞書トランスレータを起動し，辞書種類として 3 (共起辞書) を入力すると，最初に以下のプロンプトが表示されます．

補足付概念説明の格納方法を指定して下さい
1. 格納しない (0x0L として格納)
2. 重複を許して格納 (識別子は自動付与)
3. 重複を除いて格納 (識別子は自動付与)

[入力] >>

ここで，格納方式を指定します．各方式の意味は以下の通りです．

格納しない　補足付き概念説明が概念識別子の代わりにはいっている場合，概念識別子を 0 として格納します．したがって，補足付き概念説明の持つ情報は失われ，意味情報が不完全になりますが，概念見出し辞書に対する変更はありません．

重複を許して格納　補足付き概念説明があった場合，他の辞書の概念識別子とぶつからない範囲の概念識別子を自動生成し，これを格納します．テキスト辞書中に同じ補足付き概念説明が現れた場合でも，別の概念識別子を付与します．概念識別子は，補足付き概念説明がテキスト辞書中に出現した順に，c00001から一つづつ付与されます．

重複を除いて格納　補足付き概念説明があった場合，他の辞書の概念識別子とぶつからない範囲の概念識別子を自動生成し，これを格納します．テキスト辞書中に同じ補足付き概念説明が現れた場合には，同じ概念識別子を付与します．概念識別子は，補足付き概念説明がテキスト辞書中に出現した順（同じものなら最初に出現したもの）に，c00001から一つづつ付与されます．

1 を選択した場合には，その他の辞書変換の場合とと同じ処理に戻ります． 2 または 3 を選択した場合，以下のプロンプトが表示されます．

```
補足付概念説明ファイルのパス名を入力して下さい
（リターンのみ：カレントディレクトリの cctoci.txtに格納）
[入力] >>
```

ここで，パス名を入力します．このファイルは，辞書トランスレータが自動付与した概念識別子と，補足付き概念説明文字列との対応をとるもので，実際には概念見出し辞書（テキスト）と同じ形式で，概念識別子のフィールドに自動生成した識別子を，日本語概念説明のフィールドに補足付き概念説明文字列を格納します．

ファイル名入力が終われば，その後は他の辞書構築手順と同様の処理に戻ります．

ここで作成された，補足付き概念説明格納ファイルは，概念見出し辞書を作成する際にマージして下さい．補足付き概念説明が，他の概念識別子とほとんど同様に利用できるようになります．ただし，c0000l以降の概念識別子は共起辞書のみに有効であり，これを使って他の辞書を検索することはできませんので，ご注意下さい．

共起辞書構築時の注意点

1. 共起辞書では，システム辞書サイズ縮小のために，変換時にかなりのメモリを必要とします．補足付き概念説明の格納方法として，1 または 2 を選択した場合で，パソコン本体のメモリが 8MB 程度の場合，変換はできますが相当の時間がかかりますのでご注意下さい．
2. 補足付き概念説明の格納方法として，重複を許す / 重複を除くのどちらを選択するか，は作成されるファイルサイズとシステムのメモリ容量とのトレードオフとなります．重複を許す場合，メモリは 8MB 程度でも処理できますが，作成されるテキストファイルの

大きさは倍以上になります（日本語共起辞書のみの場合で 20MB 程度）．逆に，重複を許す場合，メモリは少なくとも 20MB は必要ですが，テキストファイルののサイズは半分以下になります（日本語共起辞書の場合で 9MB 程度）．

3 UNIX 版トランスレータで，オプションつき（-CC）で起動した場合，補足付き概念説明の格納方式は，自動的に 3（重複を除いて格納）になり，データはカレントディレクトリの cctoci.txtに格納されます．

3.4.2 概念見出し辞書の構築

EDBroW では，単語辞書，対訳辞書，共起辞書には概念識別子のみを格納し，概念識別子の説明であるところの概念見出しは，概念見出し辞書として別に格納しています．このため，概念見出し辞書がないと，各辞書レコードの語義が表示されません．

したがいまして，本システムの利用にあたっては，なるべく概念辞書一式を購入されることをお勧めします．

ただし，事情により，概念辞書一式が購入できない場合もあるかと思います．辞書トランスレータのオプショとンして，mkcitxtというユーティリティを用意してあります．これは，単語 / 対訳辞書の中から，概念見出しに相当するエントリのみを抜き出し，テキスト形式の概念見出し辞書と同等のフォーマットで出力するツールです．このツールにより作成されたテキストファイルを，辞書トランスレータで概念見出し辞書に変換することができます．

詳しくは，dtransディレクトリの中にある，util ディレクトリ以下の，readme.{euc, sj}を参照して下さい．

なお，共起辞書の変換中に「補足付き概念説明」を格納するファイルを作成した場合には，このファイルもマージして，概念見出し辞書を作成して下さい．

第 4 章

利用マニュアル

本章では，EDBroW の操作方法について説明します．

4.1 起動の準備

システムを起動する前には，何点か準備が必要です．

最初に，第 3 章の手順に従って，必要なシステム辞書を作成して下さい．必要があれば，ここで作成された辞書を適当なディレクトリにコピーして下さい．

次に，システム初期化ファイルを編集します．システム初期化ファイルは，EDBROW.INI という固定名で，EDBroW の実行時ディレクトリになければなりません．EDBROW.INI で，システム辞書のパス名の指定と，デフォルトのモードを設定できます．

EDBROW.INI の書式は，以下の通りです．

- 先頭に #のある行は，コメントとして扱われる．
- 〈識別子〉=〈データ〉の形式で記述する．

システム辞書パス名は，表 4.1 に示す識別子で指定します．

識別子	辞書種類
WD	単語辞書
BI	対訳辞書
CC	共起辞書
CH	概念体系辞書
CR	概念記述辞書
CI	概念見出し辞書

表 4.1 辞書指定用識別子

また，表示モードとして，「コード説明表示モード」を利用したい場合には，識別子 'CODETBL' に，コード表ファイルのパス名を指定する必要があります．EDBroW のシステム一式に含まれている，CODE.TBL というファイルがコード表ファイルです．

辞書，コード表ファイルのパス名が設定されていない，または不正なパス名である場合には，その辞書（または「コード説明表示モード」）が利用できなくなります（その辞書の検索メニューが選択不可になります）．また，不要な辞書の指定行は，削除するかコメントアウトしてください．実際に存在しないパス名などが指定されていると，EDBroW は起動時にエラーメッセージを表示します．

デフォルトモード設定は，表 4.2 に示す識別子で指定します．

モード設定と，その設定時の動作については，次節を参照して下さい．ここでの設定は，シス

識別子	内容	可能な指定	備考
DISPLAY_UINFO	ユーザ情報 表示モード	DETAIL ABBREV SIMPLE	詳細表示 省略表示 簡約表示
DISPLAY_CHDW	概念見出し 表示モード	JPN ENG ALL	日本語優先 英語優先 日英表示
DISPLAY_WEND	語尾表示 モード	ON OFF	
DISPLAY_CODEX	コード説明 表示モード	ON OFF	
DISPLAY_SRCHDLG	関連検索 モード	ON OFF	ダイアログ起動 〃 しない

表 4.2　モード指定識別子

テム起動時のデフォルトのモードを決めるだけで，モード自体は，システムの動作中に変更することができます．

以下に，EDBROW.INI の例を示します．

```
#単語辞書
WD = C:¥EDBROW¥DICT¥WD_SYS.DIC
#概念体系辞書
CH = C:¥EDBROW¥DICT¥CH_SYS.DIC
#概念記述辞書
CR = C:¥EDBROW¥DICT¥CR_SYS.DIC
#概念説明辞書
CI = C:¥EDBROW¥DICT¥CI_SYS.DIC
#対訳辞書
BI = C:¥EDBROW¥DICT¥BI_SYS.DIC
#共起辞書
CC = C:¥EDBROW¥DICT¥CC_SYS.DIC
#コード表ファイル
CODETBL = CODE.TBL
#表示モード
DISPLAY_UINFO = ABBREV
DISPLAY_CHDW = ALL
DISPLAY_WEND = ON
DISPLAY_CODEX = ON
#関連検索時の　検索ダイアログ起動モード
DISPLAY_SRCHDLG = ON
```

なお上記の例は，EDBroW 配布時の EDBROW.INI の内容と同じです．したがって，各辞書のパス名を上記と同じにすれば，そのまま EDBroW を利用できます．

4.2 システムのモード

4.2.1 表示モード

表示モードには，4 種類あります．

ユーザ情報表示モード

辞書検索の結果は，ユーザ情報として取り込むことができ，取り込まれた情報は，ユーザ情報表示領域に表示されます．この表示形式を，以下の 3 種類から選ぶことができます．

詳細表示　ユーザ情報の各レコードについて，すべての情報を表示する．参照情報，覚書情報は，このモード時にのみ表示される．

省略表示　各レコードの一部の情報を省略して，見易い形式で表示する．

簡約表示　各レコードの，最低限必要な情報のみを表示する．

各辞書レコードの，それぞれのモードにおける表示形式は，各辞書検索の項を参照して下さい．

概念見出し表示モード

概念見出し辞書を除くすべての辞書レコードには，概念識別子のみが格納されており，これに対する説明文は，概念見出し辞書を参照して表示するように設計されています．概念見出し表示モードは，この表示形式を指定するもので，以下の 3 種類から選ぶことができます．

日本語優先　概念見出し辞書から得られたフィールド（4 種類）のうち，日本語のデータを優先して，以下の形式で表示します．

{〈日本語概念見出し〉}〈日本語概念説明〉

日本語概念見出しがない場合には，{} は表示されず，日本語概念説明のみが表示されます．

概念見出しレコードに，日本語概念見出し，日本語概念説明が両方とも無い場合には，英語のデータを，以下の形式で表示します．

{〈英語概念見出し〉}〈英語概念説明〉

英語優先　英語概念見出し / 英語概念説明を優先して表示します．形式は，日本語優先モード時と同様です．

日英表示　概念見出し辞書レコードの，すべてのデータを，以下の形式で表示します．

{〈日本語概念見出し〉}〈日本語概念説明〉 /
{〈英語概念見出し〉}〈英語概念説明〉

語尾表示モード

単語レコードにのみ有効なモードです．

システムの単語辞書では，単語見出しとして単語の不変化部分を格納しています（動詞「引く」であれば "引"）．辞書レコード表示の際，語尾表示モードが ON になっていれば，単語の語尾

（終止形）を付加して表示します．これは，単語見出しだけでなく，読みのフィールド（日本語単語辞書のみ）にも有効です．

例

OFF 【引】(ヒ) (JVE) <3cf7b1> {J}

ON 【引・く】(ヒ・ク) (JVE) <3cf7b1> {J}

コード説明表示モード

単語レコード，対訳レコードに有効なモードです．

コード表示された品詞，左右連接属性に，そのコードの説明を日本語で付加して表示します．品詞，左右連接属性以外の文法コードには無効です．

例

OFF 品詞： JN1

ON 品詞： JN1（普通名詞）

なお，このモードは，辞書検索結果の詳細表示時，および，ユーザ情報では，表示モードが「詳細表示」になっている時にのみ有効です．

4.2.2 その他のモード

デフォルト検索モード

EDBroW には，検索モードとして，以下の 3 種類が用意されています．

完全一致 検索キーと完全に一致するすべてのレコードが検索されます．

部分一致 検索キーの先頭から 1 文字づつ伸長させたキーを完全一致で検索した結果の和集合が検索されます．すなわち，検索キーが "ABCDE" である場合には，

A
AB
ABC
ABCD
ABCDE

が検索されます．

前方一致 検索キーと，先頭から検索キー長さ分までが一致するレコードが検索されます．すなわち，検索キーが "man"である場合には，

man
manager
mankind
manner
manual

などが検索されます．

これらのモードは，可変長の検索キーにのみ有効であり，固定長の検索キー（概念識別子）は完全一致による検索しかできません．この場合には，他のモードが選択されていても，自動的に完全一致で検索を実行します．

また，単語辞書に限り，「語尾つき検索モード」が利用可能ですが，これは，部分一致検索とフィルタキーを利用することにより，検索キーを「語尾つき」でも検索できるようにしたものです．つまり，単語 "dictionary"が，検索キー "dictionary"で検索できます（完全一致検索の場合には "dictionar を検索キーとする）．

これらの検索モードは，各辞書検索ダイアログボックス内で指定できますが，ダイアログボックス起動時のデフォルト検索モード（起動時にチェックされているラジオボタン）を指定するのが，「デフォルト検索モード」の役目です．

デフォルト検索モードでは，「語尾つき検索」も指定できるようになっていますが，これが指定された場合，単語辞書以外では，「完全一致」がデフォルト検索モードとなります．

関連検索時のモード

関連検索とは，ユーザ情報中の辞書レコードから，関連する他の辞書のレコードを検索する機能です．関連検索の詳細については，4.4.9で説明しますが，関連検索ダイアログボックスで，検索する辞書の種類と検索キーを指定した後，【OK】をクリックすると，対象となる辞書の検索ダイアログボックスが起動するようになっています．関連検索時のモード設定で，辞書検索用のダイアログボックスは起動せずに，直接辞書を検索するようにすることが可能です．モードがOFF の時には，対象となる辞書検索ダイアログボックスは起動しません．

4.3 システムの起動

システムは，プログラムマネージャから，EDBroW のアイコン（図4.1参照）をクリックして，[ファイル] – [開く] を選択するか，アイコンをダブルクリックすれば起動します．

図 4.1 アイコン

システムが起動すると，図4.2に示すメインウィンドウが表示されます．

ユーザ情報表示エリアには，起動時にはなにも表示されません．辞書検索結果を取り込むなどして，ユーザ情報が作成されると，このエリアに

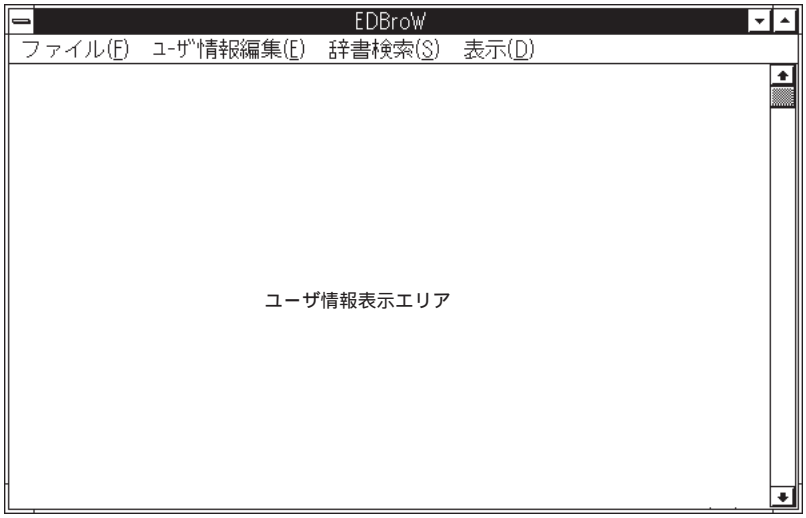


図 4.2 EDBroW メインウィンドウ

表示されるようになります．

EDBroW は，メインウィンドウのメニューを選択することにより，検索などのコマンドを実行します．以下に，各メニュー項目と，その動作概要を示します．なお，ショートカットキーが定義されている場合には，[メニュー]（ショートカット）のように記します．

[ファイル(F)] メニュー

システム全体の動作にかかわるコマンドをサブメニューとして持ちます．

[ユーザ情報読込(L)...]（Ctrl+L）

EDBroW の実行時ディレクトリにある，ユーザ情報ファイル（拡張子 .uif）を読み込みます．

[ユーザ情報保存(S)...]（Ctrl+S）

現在バッファにあるユーザ情報を，ユーザ情報ファイルに保存します．

[ユーザ情報クリア(C)]（Ctrl+N）

現在バッファにあるユーザ情報をクリアし，バッファをシステム起動時の状態に戻します．

[辞書切換え(J)...]（Ctrl+J）

システム辞書を切りかえます．

[モード設定(M)...]（Ctrl+D）

システムのモードを変更します．

[印刷(P)...]

現在バッファにあるユーザ情報を印刷します．

[終了(X)]

システムを終了します．

[ユーザ情報編集(E)] メニュー

ユーザ情報の編集に関連するコマンドをサブメニューとして持ちます．

[新規(N)]

ユーザ情報レコードを新規作成します．レコードの種類に応じて，以下のサブメニューがあります．

[単語レコード(W)...]

[対訳レコード(B)...]

[共起レコード(C)...]

[概念体系レコード(H)...]

[概念記述レコード(R)...]

[概念見出しレコード(I)...]

[変更(M)...]（Ctrl+E）

選択されたユーザ情報レコードの内容を変更します．

[削除(X)...]（Del）

選択されたユーザ情報レコードを削除します．

[コピー(C)...]（Ctrl+V）

選択されたユーザ情報レコードをコピーします．

[取消(U)...]（Ctrl+U）

ユーザ情報への変更を取り消し，バッファを一つ前の状態に戻します．

[キーワード(K)...]

ユーザ情報のキーワードを設定 / 変更します．

[辞書検索(S)] メニュー

辞書検索用のコマンドをサブメニューとして持ちます．システム辞書を辞書レコード単位で検索するコマンドとして，以下のサブメニューがあります．

[単語辞書検索(W)...]（Ctrl+W）

[対訳辞書検索(B)...]（Ctrl+B）

[共起辞書検索(C)...]（Ctrl+C）

[概念体系辞書検索(H)...]（Ctrl+T）

[概念記述辞書検索(R)...]（Ctrl+R）

[概念見出し辞書検索(I)...]（Ctrl+F）

その他に，以下の検索コマンドが用意されています．

[上位概念検索(U)...]（Ctrl+O）

概念体系辞書を使い，検索キーとなる概念識別子が体系上のどの位置にあるか，を表示します．

[可能関係子検索(P)...]（Ctrl+P）

概念記述辞書と概念体系辞書の継承機能を利用し，検索キーとなる 2 概念間にどのような概念関係が成り立ち得るかを調べ，表示します．

[関連検索(A)...]（Ctrl+A）

指定されたユーザ情報レコードをもとに，関連する他の辞書を検索するためのコマンドです．

[表示(D)] メニュー

ユーザ情報の表示モードを変更します．

[ユーザ情報詳細表示(D)]

ユーザ情報表示モードを「詳細表示」に設定します．

[ユーザ情報省略表示(A)]

ユーザ情報表示モードを「省略表示」に設定します．

[ユーザ情報簡約表示(S)]

ユーザ情報表示モードを「簡約表示」に設定します．

4.4 辞書の検索

各辞書の検索方法の説明に入る前に，各辞書に共通な操作，注意点等について説明します．

概念識別子の入力

検索ダイアログボックスなどにおける概念識別子の入力は，16進数6桁（半角）の文字列により行いますが，入力がちょっと面倒です．このため，EDBroW では，概念識別子の入力を必要とするすべてのダイアログボックスに，「概念識別子入力コンボボックス」を利用しています．

概念識別子入力コンボボックスは，現在ユーザ情報中に取り込まれている，すべてのレコードに含まれる概念識別子をリスト表示し，リストからマウスで選択することにより概念識別子の入力が可能となるものです．

例えば，ある単語の 1 語義にあたる概念識別子で，別の辞書を引きたい場合には，まず単語辞書のレコードをユーザ情報中に取り込みます．次に，検索したい辞書のダイアログボックスを立ち上げれば，概念識別子入力コンボボックスを利用して，マウスにより概念識別子が入力できます（図 4.3参照）．

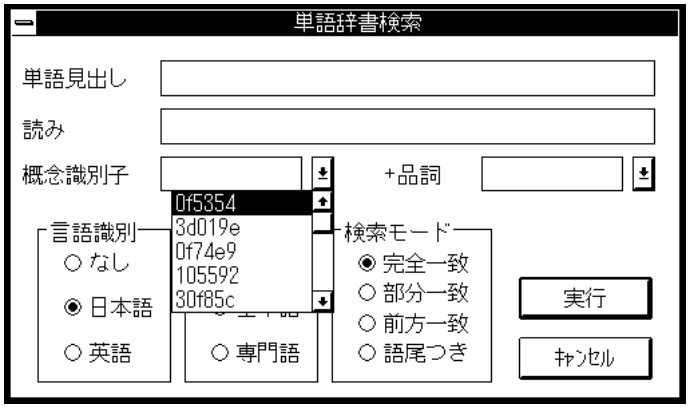


図 4.3 概念識別子入力コンボボックス

※ 概念識別子は，ユーザ情報に取り込まれた順序の逆順でリスト表示されます．なお，ユーザ情報が削除されても，概念識別子は残ります．

フィルタキーについて

辞書検索は，検索キーとフィルタキーとの組み合わせで行うことができます．ただし，検索キーは，最低でも 1 つは指定されていなければなりません．各辞書の検索ダイアログボックスでは，検索キーとフィルタキーの区別をつけるため，フィルタキーについては，フィールド名の前に '+' を表示しています．

検索時間と結果表示

辞書検索は，システム辞書中のインデックス部分を二分探索することにより実現されています．このため，辞書検索の結果が大量になる場合，ディスクアクセスが数多く発生し，検索にかなりの時間がかかることがあります．その間，中断はできません．

検索中は，マウスカーソルが砂時計の形になり，クリックは無効です．数分間戻ってこないこともありますが，ハングアップしているわけではありませんので，しばらくお待ち下さい．

特に以下のような検索を行なった場合，検索にかなりの時間がかかることがありますので，ご注意下さい．

- 概念体系辞書を，「上位概念」のみを指定して検索した場合
- 単語 / 対訳 / 共起辞書を，「前方一致」で検索した場合
- 共起辞書を，エントリ数の多い単語見出しで検索した場合
- 専門用語辞書などで，先頭数文字が同じであるレコードが多数あるものを検索した場合

また，検索結果の格納領域は，メモリ容量の許す限り確保できますが，結果が大量にある場合，すべての情報を表示できないことがあります．その場合は，複数検索キーや，フィルタキーなどを使い，検索結果ができるだけ少なくするようにして下さい．

検索結果が大量になると，「...情報が多すぎます」というメッセージボックスが表示されますが，可能なレコードまでは表示されます．

4.4.1 単語辞書検索

検索キーの入力

検索キー，フィルタキー，検索モードの指定は，単語辞書検索ダイアログボックスで行います．単語辞書検索ダイアログボックスは，[検索]- [単語辞書検索](Ctrl+W) コマンドで起動します．図 4.4にダイアログボックスを示します．

ここで，検索キー，フィルタキー，検索モードを入力します．検索キーとして指定できるのは，

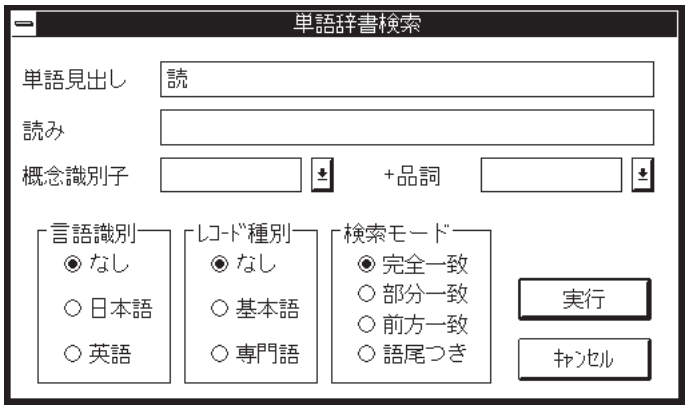


図 4.4 単語辞書検索ダイアログボックス

以下の３種類です

単語見出し
読み
概念識別子

上記３種類の検索キーのうち，必ず一つは指定するようにして下さい．複数指定することも可能ですが，その場合，２つ目以降のキーは，検索結果に対する制約条件となります．

なお，単語見出し / 読みは，通常単語の不変化部分を入力します．語尾つき単語で検索したい場合には，検索モードを「語尾つき」にして下さい (語尾つき検索は，完全一致検索に比べ，若干時間がかかります) ．

フィルタキーとして指定できるのは，以下の３種類です．

品詞
言語識別
レコード種別 (基本語 / 専門用語の別)

品詞はコンボボックスで入力，言語識別とレコード種別は，ラジオボタンで指定します．言語識別，レコード種別ラジオボタンは，ダイアログボックス起動時には，「なし」(種別を指定しない) が選択されています．

ただし，品詞はコードで入力するようになっているので，言語識別が指定されていないと，フィルタキーとしての意味を持ちません．また，品詞入力にはコンボボックスを利用できますが，言語識別が指定されないと，品詞コードの一覧が表示されません．品詞をフィルタキーとして指定する場合は，先に言語識別を指定するようにして下さい．

最後に，検索モードを指定します．各検索モードと，その動作については，4.2.2に示した通りです．

検索キー入力，フィルタキー指定が完了したら，【実行】をクリックして下さい．辞書の検索を開始します．

検索結果の表示

検索が終了すると，図 4.5に示す検索結果表示ダイアログボックスが起動します．

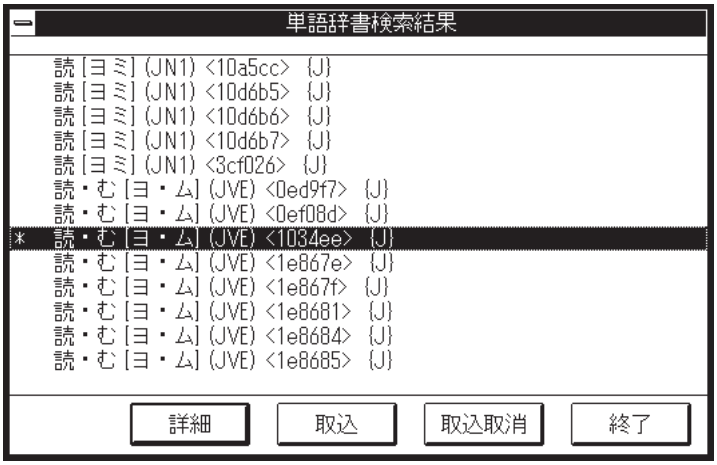


図 4.5 単語辞書検索結果表示

ここでは，検索結果がリスト表示されます．リストボックス内で，表示されている辞書レコードをクリックすると，反転表示となり，「詳細表示」「取込」「取込取消」操作の対象となります．また，レコードをダブルクリックすると，そのレコードを詳細表示します（選択後に【詳細】をクリックした場合と同様）．

各ボタンの動作は，以下の通りです．

【詳細】　選択されている辞書レコードを詳細表示します．

【取込】　選択されている辞書レコードを，ユーザ情報への取り込みの対象とします．

【取込取消】　選択されている辞書レコードを，ユーザ情報への取り込みの対象から除外します．

【終了】　表示ダイアログボックスを消去し，メインウィンドウの入力待ち状態に戻ります．取り込みが指定されている場合には，該当する辞書レコードをユーザ情報に取り込みます．

【取込】がクリックされると，辞書レコードの先頭に '*' が表示され，当該辞書レコードがユーザ情報への取り込みの対象となったことを示します．また，取り込み対象である辞書レコードを選択し，【取込取消】をクリックすると，先頭の '*' が消え，取込対象から除外されます．

取込は，複数指定することが可能です．また，辞書レコードが選択されていない状態で【取込】がクリックされると，先頭の辞書レコードを取り込みの対象にします．

【詳細】がクリックされると，図 4.6 に示す詳細表示ダイアログボックスが起動します．辞書レコードが選択されていない場合には，先頭レコードが詳細表示の対象となります．

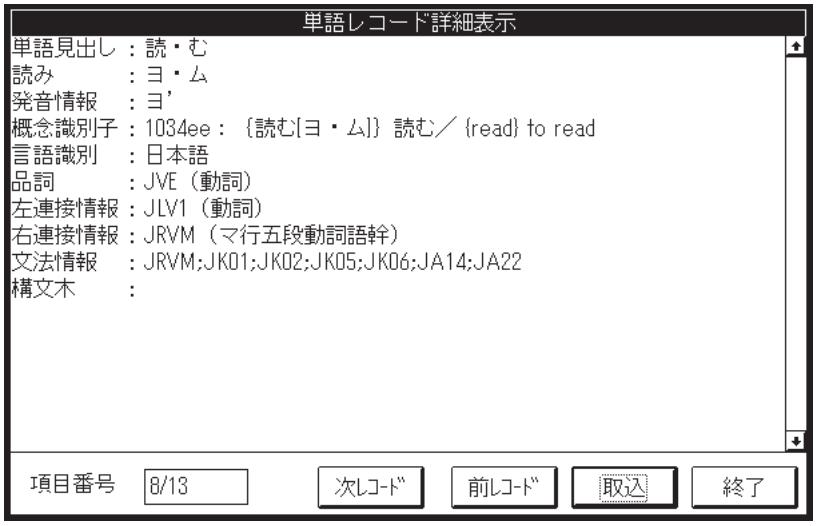


図 4.6　単語レコード詳細表示

ここでは，指定された辞書レコードの全ての情報を表示します．また，次のレコードや前のレコードを詳細表示したり，現在のレコードを取込の対象に指定することも可能です．

各ボタンの動作は，以下の通りです．

【次レコード】　次の辞書レコードを詳細表示します．次のレコードがない場合には，ボタンはセンシティブになりません．

【前レコード】　一つ前の辞書レコードを詳細表示します．一つ前のレコードがない場合には，ボタンはセンシティブになりません．

【取込】　現在詳細表示している辞書レコードを取込対象とします．

【終了】　詳細表示ダイアログボックスを消して，検索結果表示ダイアログボックスに戻ります．

詳細表示，取込の指定などが完了したら，【終了】ボタンをクリックすると，検索結果ダイアログボックスが消去されます．その際，取込対象に指定されていた辞書レコードが，ユーザ情報中に取りこまれます．

注意点

システム辞書では，慣用句に対する単語見出しとして，構成語の先頭 1 単語のみを格納しています．このため，例えば検索キー「足」で，完全一致による検索を実行すると，「足を棒にする」のような慣用句のレコードも一緒に検索されてきます．これらの区別は，「読み」のフィールドで行ってください (日本語の慣用句の場合「アシ/ヲ/ボウ/ニ/ス」のようなカナ表記が，英語の慣用句の場合，音節区切りのかわりに構成語見出しが，読みのフィールドに格納されています)．

表示形式

各ユーザ情報表示モードにおける，単語レコードの表示形式は，以下に示す通りです．なお，検索結果表示ダイアログボックスでは，各辞書レコードは「簡約表示」モード時と同様の形式で，詳細表示ダイアログボックスでは，「詳細表示」モード時と同様の形式で表示されます．

簡約表示モード時

〈単語見出し〉[〈読み〉](〈品詞〉)<〈概念識別子〉> {〈言語識別〉}¥n

〈言語識別〉 ::= J | E | TJ | TE

※ J は日本語，E は英語を表す．また T は専門用語であることを表す．

省略表示モード時

単語見出し：〈単語見出し〉[〈読み〉]{〈言語識別〉}¥n
〈発音 / 構成語表示〉¥n
概念識別子：〈概念識別子〉:〈概念見出し文字列〉¥n
品詞 / 連接：〈品詞〉(〈左連接情報〉,〈右連接情報〉)¥n
文法情報　：〈文法情報〉
構文木　：〈構文木〉

〈発音 / 構成語表示〉::= 発音情報：〈発音情報〉 | 構成語情報：〈構成語情報〉
〈言語識別〉 ::= J | E | TJ | TE

※〈文法情報〉〈構文木〉は，データが無い場合にはフィールド名も含めて表示されない．

詳細表示モード時

単語見出し：〈単語見出し〉
読み　　：〈読み〉
〈発音 / 構成語表示〉¥n
概念識別子：〈概念識別子〉:〈概念見出し文字列〉¥n
言語識別　：〈言語識別〉
品詞　　：〈コード表示〉¥n
左連接情報：〈コード表示〉¥n
右連接情報：〈コード表示〉¥n
文法情報　：〈文法情報〉
構文木　：〈構文木〉

〈発音 / 構成語表示〉::= 発音情報：〈発音情報〉 | 構成語情報：〈構成語情報〉
〈言語識別〉::= "日本語" | "英語" | "日本語 (専門用語) " | "英語 (専門用語) "
〈コード表示〉::=〈コード文字列〉 | 〈コード文字列〉 (〈コード説明〉)

※ 〈コード表示〉は，「コード説明表示モード」の ON/OFF により異なる．

4.4.2 対訳辞書検索

検索キーの入力

検索キー，フィルタキー，検索モードなどの指定は，対訳辞書検索ダイアログボックスで行います．対訳辞書検索ダイアログボックスは，[検索]-[対訳辞書検索](Ctrl+B)コマンドで起動します．図 4.7 にダイアログボックスを示します．

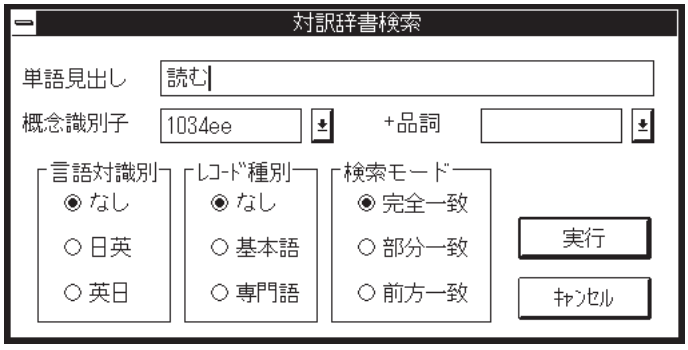


図 4.7 対訳辞書検索ダイアログボックス

ここで，検索キー，フィルタキー，検索モードなどを入力します．検索キーとして指定できるのは，以下の 2 種類です

単語見出し
概念識別子

上記 2 種類の検索キーのうち，必ず一つは指定するようにして下さい．両方指定することも可能ですが，その場合，概念識別子は検索結果に対する制約条件となります．

なお，対訳辞書では，単語見出しは「語尾つき」で格納されていますので，検索キーはは語尾つき単語を入力して下さい．

フィルタキーとして指定できるのは，以下の 3 種類です．

品詞
言語対識別
レコード種別（基本語 / 専門用語の別）

フィルタキーの指定，検索モードの指定は，単語辞書検索ダイアログボックスと同様です．

検索キー入力，フィルタキー指定が完了したら，【実行】をクリックして下さい．辞書の検索を開始します．

検索結果の表示

検索が終了すると，図 4.8に示す検索結果表示ダイアログボックスが起動します．

図 4.8　対訳辞書検索結果表示

以降の動作は，単語辞書検索の場合とまったく同様です．また，図 4.9に示す詳細表示ダイアログボックスの動作も，単語辞書検索時と同様です．

表示形式

各ユーザ情報表示モードにおける，共起レコードの表示形式は，以下に示す通りです．なお，検索結果表示ダイアログボックスでは，各辞書レコードは「簡約表示」モード時と同様の形式で，詳細表示ダイアログボックスでは，「詳細表示」モード時と同様の形式で表示されます．

なお，対訳情報の格納形式は，テキスト辞書の形式とまったく同じです．すなわち，以下の通りです．

〈対訳情報〉 ::= 〈訳語情報〉 | 〈対訳情報〉//〈訳語情報〉
〈訳語情報〉 ::= 〈訳語種別〉'|'〈訳語表記〉'|'〈訳語品詞〉

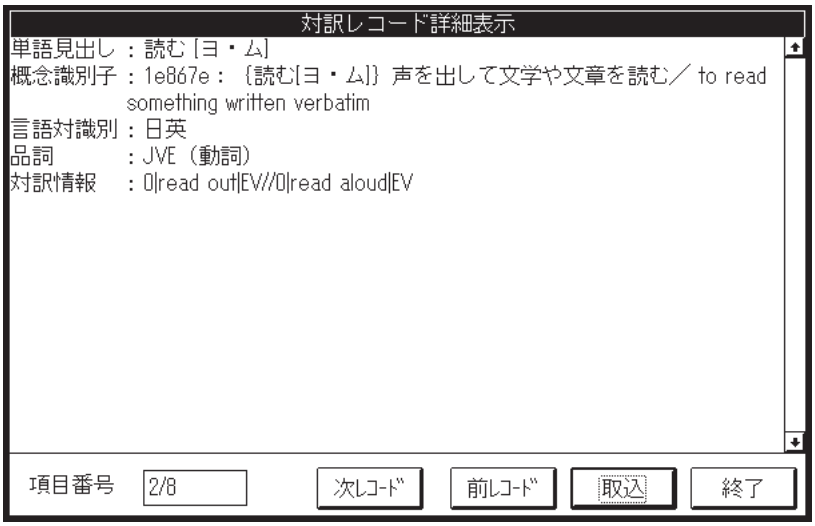


図 4.9 対訳レコード詳細表示

簡約表示モード時

〈単語見出し〉(〈品詞〉)<〈概念識別子〉>{〈言語対識別〉}→ 〈対訳情報〉¥n

〈言語対識別〉::= "JE" | "EJ" | "TJE" | "TEJ "

※ JE は，EJ は英日を表す．また T は専門用語であることを表す．

省略表示モード時

単語見出し：〈単語見出し〉[〈読み〉]{〈言語対識別〉}¥n
概念識別子：〈概念識別子〉:〈概念見出し文字列〉¥n
対訳情報 ：〈対訳情報〉

〈言語対識別〉::= "JE" | "EJ" | "TJE" | "TEJ "

※ [〈読み〉] が表示されるのは，日英の場合のみ

詳細表示モード時

単語見出し：〈単語見出し〉[〈読み〉]
概念識別子：〈概念識別子〉:〈概念見出し文字列〉¥n
言語対識別：〈言語対識別〉
品詞 ：〈コード表示〉¥n
対訳情報 ：〈対訳情報〉

〈言語対識別〉::= "日英" | "英日" | "日英（専門用語）" | "英日（専門用語）"
〈コード表示〉::=〈コード文字列〉 | 〈コード文字列〉(〈コード説明〉)

※ 〈コード表示〉は，「コード説明表示モード」の ON/OFF により異なる．

4.4.3 共起辞書検索

検索キーの入力

検索キー，フィルタキー，検索モードなどの指定は，共起辞書検索ダイアログボックスで行います．共起辞書検索ダイアログボックスは，[検索]ー[共起辞書検索](Ctrl+C)コマンドで起動します．図 4.10にダイアログボックスを示します．

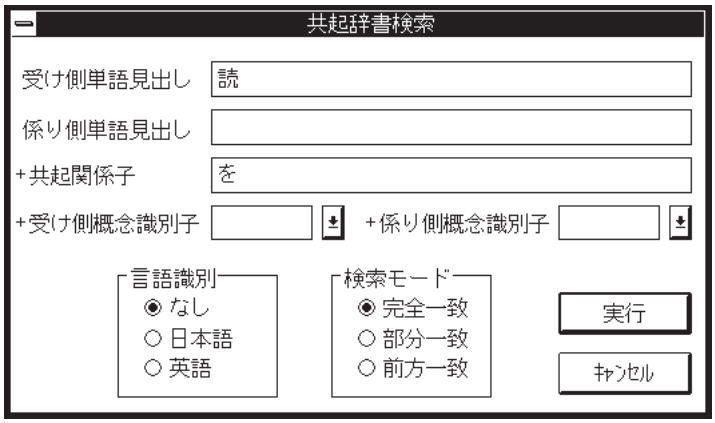


図 4.10 共起辞書検索ダイアログボックス

ここで，検索キー，フィルタキー，検索モードなどを入力します．検索キーとして指定できるのは，以下の 2 種類です．

受け側単語見出し
係り側単語見出し

上記 2 種類の検索キーのうち，必ず一つは指定するようにして下さい．両方指定することも可能ですが，その場合，係り側単語見出しは検索結果に対する制約条件となります．

なお，共起辞書では，各単語見出しは単語の不変化部分を格納していますので，検索キーとしては不変化部分を入力して下さい．

フィルタキーとして指定できるのは，以下の 4 種類です．

共起関係子
言語識別
受け側概念識別子
係り側概念識別子

共起関係子は，テキスト文字列で入力します．言語識別は，ラジオボタンで指定して下さい．その他の動作は，単語辞書検索の場合と同様です．

検索キー入力，フィルタキー指定が完了したら，【実行】をクリックして下さい．辞書の検索を開始します．

検索結果の表示

検索が終了すると，図 4.11に示す検索結果表示ダイアログボックスが起動します．

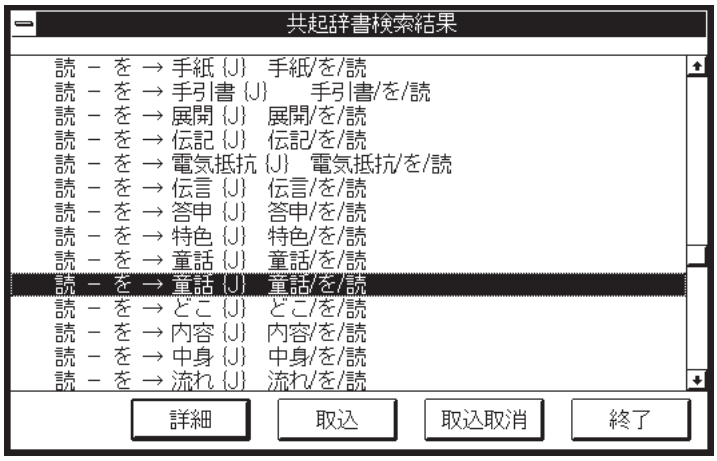


図 4.11 共起辞書検索結果表示

以降の動作は，単語辞書検索の場合とまったく同様です．また，図 4.12に示す詳細表示ダイアログボックスの動作も，単語辞書検索時と同様です．

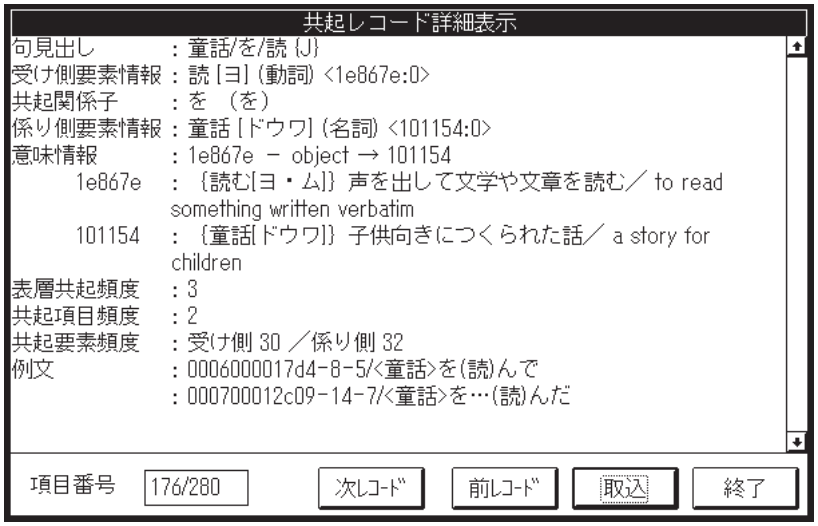


図 4.12 共起レコード詳細表示

表示形式

各ユーザ情報表示モードにおける，共起レコードの表示形式は，以下に示す通りです．なお，検索結果表示ダイアログボックスでは，各辞書レコードは「簡約表示」モード時と同様の形式で，詳細表示ダイアログボックスでは，「詳細表示」モード時と同様の形式で表示されます．

なお，共起辞書中には，概念関係子が "nil" であるレコードがあります．これは，意味情報が格

納されていないレコードですので，表示時には，意味情報のフィールドが "(nil)" と表示されます．

簡約表示モード時

〈受け側単語見出し〉−〈共起関係子〉→〈係り側単語見出し〉
{〈言語識別〉} 〈句見出し〉

〈言語識別〉 ::= J | E
※ J は日本語，E は英語を表す．

省略表示モード時

句見出し : 〈句見出し〉{〈言語識別〉}¥n
表層共起情報：〈受け側単語見出し〉<〈受け側概念識別子〉> −
〈共起関係子〉→〈係り側単語見出し〉<〈係り側概念識別子〉>¥n
意味情報 : 〈受け側概念識別子〉−〈概念関係子〉→ 〈係り側概念識別子〉¥n
〈始点概念識別子〉：〈概念見出し文字列〉¥n
〈終点概念識別子〉：〈概念見出し文字列〉¥n
頻度情報 ：表層共起頻度〈表層共起頻度〉/ 共起項目頻度〈共起項目頻度〉
/ 共起要素頻度〈受け側共起要素頻度〉/〈係り側共起要素頻度〉¥n
例文 :〈例文〉¥n

〈言語識別〉 ::= J | E

※意味情報中の始点 / 終点概念識別子と，受け側 / 係り側概念識別子との対応は，共起レコード中の「関係子方向」データ (陽には表示されない) の内容に応じて異なる．

詳細表示モード時

句見出し : 〈句見出し〉{〈言語識別〉}¥n
受け側要素情報：〈受け側単語見出し〉[〈受け側単語補助情報〉]
(〈受け側品詞〉)<〈受け側概念識別子〉:〈受け側慣用句フラグ〉>¥n
共起関係子 : 〈共起関係子〉(〈共起関係子補助情報〉) ¥n
係り側要素情報：〈係り側単語見出し〉[〈係り側単語補助情報〉]
(〈係り側品詞〉)<〈係り側概念識別子〉:〈係り側慣用句フラグ〉>¥n
意味情報 : 〈受け側概念識別子〉−〈概念関係子〉→ 〈係り側概念識別子〉¥n
〈始点概念識別子〉：〈概念見出し文字列〉¥n
〈終点概念識別子〉：〈概念見出し文字列〉¥n
表層共起頻度 ：〈表層共起頻度〉¥n
共起項目頻度 ：〈共起項目頻度〉¥n
共起要素頻度 ：受け側〈受け側共起要素頻度〉/ 係り側〈係り側共起要素頻度〉¥n
例文 :〈例文〉¥n

〈言語識別〉 ::= J | E

※例文は，複数ある場合には，一例文ごとに改行して表示．

4.4.4 概念体系辞書検索

検索キーの入力

検索キーの入力は，概念体系辞書検索ダイアログボックスで行います．概念体系辞書検索ダイアログボックスは，[検索]-[概念体系辞書検索](Ctrl+T)コマンドで起動します．図 4.13にダイアログボックスを示します．

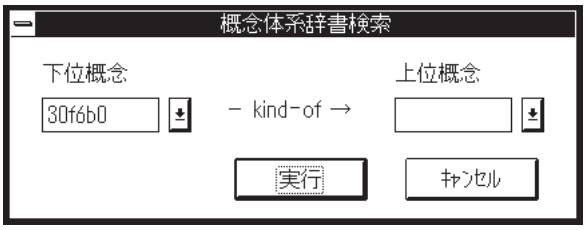


図 4.13 概念体系辞書検索ダイアログボックス

検索キーは，以下の 2 種類が指定できます．

下位概念
上位概念

入力が完了したら，【実行】をクリックすると検索を開始します．

検索結果の表示

検索が終了すると，図 4.14に示す検索結果表示ダイアログボックスが起動します．

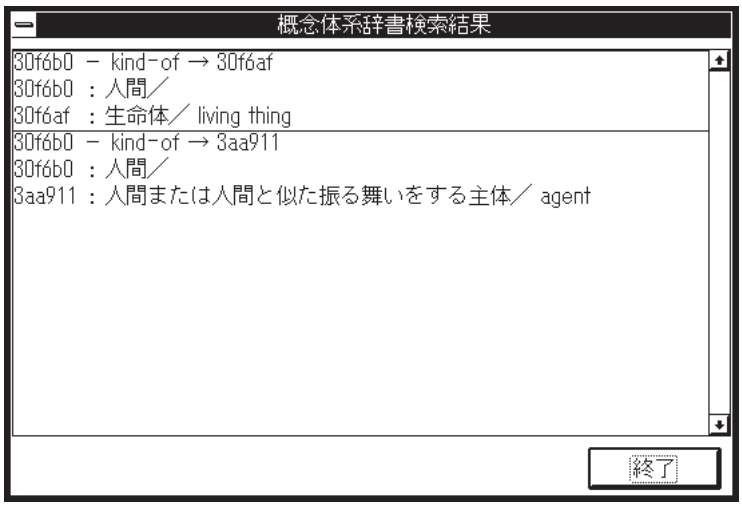


図 4.14 概念体系辞書検索結果表示

リストボックス内では，各レコードを以下の形式で表示しています．

〈下位概念〉- kind-of→ 〈上位概念〉¥n
〈下位概念〉:〈概念見出し文字列〉¥n
〈上位概念〉:〈概念見出し文字列〉¥n

ただし，概念見出し文字列が一行におさまらない場合には，折り返し表示されます．各レコードの間には，仕切り線が引かれます．

ここで，レコードをクリックすると，反転表示となり，そのレコードがユーザ情報への取り込みの対象となったことを示します．再度クリックすると，通常表示に戻り，取込対象から除外されます．

【終了】をクリックすると，ダイアログボックスは消去され，取込対象に指定されたレコードがユーザ情報中に取り込まれます．

表示形式

ユーザ情報中の概念体系レコードは，簡約表示モード時には，概念見出しの行を省略し，1 行で表示しますが，それ以外のモードの場合には，検索結果表示ダイアログボックスでの表示形式と同じです．

4.4.5 概念記述辞書検索

検索キーの入力

検索キーの入力は，概念記述辞書検索ダイアログボックスで行います．概念記述辞書検索ダイアログボックスは，[検索]- [概念記述辞書検索](Ctrl+R)コマンドで起動します．図 4.15にダイアログボックスを示します．

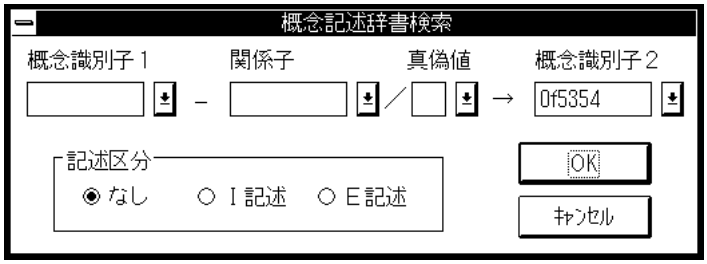


図 4.15 概念体系辞書検索ダイアログボックス

検索キーは，以下の 2 種類が指定できます．

概念識別子 1（始点側）
概念識別子 2（終点側）

両方指定することも可能ですが，その場合，概念識別子 2 は，フィルタキーとして働きます．

フィルタキーとして指定できるのは，以下の 3 種類です．

関係子
真偽値
記述区分

関係子，真偽値は，コンボボックスにより入力できます．真偽値として入力できるのは，0,1 のみです．記述区分は，ラジオボタンにより入力します．このラジオボタンは，起動時には「なし」に設定されています．

入力が完了したら，【実行】をクリックすると検索を開始します．

検索結果の表示

検索が終了すると，図 4.16に示す検索結果表示ダイアログボックスが起動します．

図 4.16　概念記述辞書検索結果表示

リストボックス内では，各レコードを以下の形式で表示しています．

〈概念識別子１〉−〈関係子〉 /〈真偽値〉→〈概念識別子２〉(〈記述区分〉)¥n
〈概念識別子１〉:〈概念見出し文字列〉¥n
〈概念識別子２〉:〈概念見出し文字列〉¥n

ダイアログボックスの動作は，概念体系辞書の場合とまったく同様です．

表示形式

ユーザ情報中の概念記述レコードは，簡約表示モード時には，概念見出しの行を省略し，１行で表示しますが，それ以外のモードの場合には，検索結果表示ダイアログボックスでの表示形式と同じです．

4.4.6　概念見出し辞書検索

各辞書レコード表示時には，そのレコードに格納されている概念識別子に対し，概念見出し辞書を参照して，概念見出し文字列を表示するように設計されていますので，概念見出し辞書の検索は通常あまり必要ありません．ただし，概念見出し辞書レコードも，その他の辞書レコードと同様にユーザ情報中に取り込んで編集することができますので，ユーザ情報中に概念見出しレコードを 取り込みたい場合などに，この検索が有効です．

検索キーの入力

検索キーの入力は，概念見出し辞書検索ダイアログボックスで行います．概念見出し辞書検索ダイアログボックスは ,[検索]− [概念見出し辞書検索](Ctrl+F)コマンドで起動します．図 4.17 にダイアログボックスを示します．

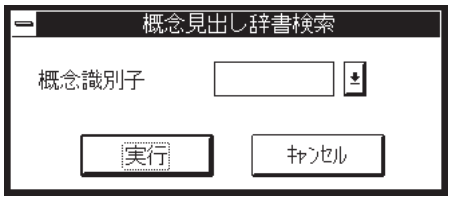


図 4.17　概念見出し辞書検索ダイアログボックス

検索キーとして指定できるのは，概念識別子のみです．

入力が完了したら，【実行】をクリックすると検索を開始します．

検索結果の表示

検索が終了すると，図 4.18に示す検索結果表示ダイアログボックスが起動します．

図 4.18　概念見出し辞書検索結果表示

概念見出し辞書では，概念識別子 1 つに対して，複数のレコードが存在することはありません．
概念見出し辞書レコードをユーザ情報へ取り込む場合には，ここで【取込】をクリックします．
概念見出しの場合，一度取り込みが設定されると，取消はできませんので，ご注意下さい．

【終了】をクリックすると，ダイアログボックスが消去され，レコードが取込指定されている場合には，ユーザ情報中に取り込まれます．

表示形式

各ユーザ情報表示モードにおける，概念見出しレコードの表示形式は，以下に示す通りです．
なお，検索結果表示ダイアログボックスでは，各辞書レコードは「詳細表示」モード時と同様の形式で表示されます．

簡約表示モード時

〈概念識別子〉

※概念識別子のみを表示する．

省略表示モード時

〈概念識別子〉：〈概念見出し文字列〉¥n

詳細表示モード時

概念識別子　　：〈概念識別子〉¥n
日本語概念見出し：〈日本語概念見出し〉¥n
英語概念見出し　：〈英語概念見出し〉¥n
日本語概念説明　：〈日本語概念説明〉¥n
英語概念説明　　：〈英語概念説明〉¥n

4.4.7　上位概念検索

上位概念検索とは，概念体系辞書を使い，ある概念識別子が，概念体系中でどのような分類の下にあるか，を表示するものです．検索結果は表示のみで，ユーザ情報への取り込み機能はありません．

検索キーの入力

検索キーの入力は，上位概念検索ダイアログボックスで行います．上位概念検索ダイアログボックスは，[検索]- [上位概念検索](Ctrl+O)コマンドで起動します．図 4.19にダイアログボックスを示します．

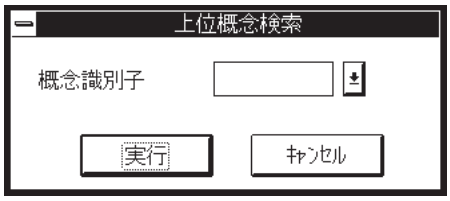


図 4.19　上位概念検索ダイアログボックス

検索キーとして指定できるのは，概念識別子のみです．

入力が完了したら，【実行】をクリックすると検索を開始します．

検索結果の表示

検索が終了すると，図 4.20に示す検索結果表示ダイアログボックスが起動します．

検索結果は，以下のフォーマットで表示されます．

〈番号〉〈概念識別子〉:〈概念見出し文字列〉

先頭の番号は，入力された概念識別子を 0 として，そこから体系上でいくつ上位の概念であるか，を表示するものです．また，各行の先頭は，一つ上位の概念になるごとに，一定文字数分インデントして表示されます．

※上位概念検索では，検索結果のすべての概念識別子に対して，概念見出し辞書を検索します．このため，検索終了までに多少の時間がかかります．

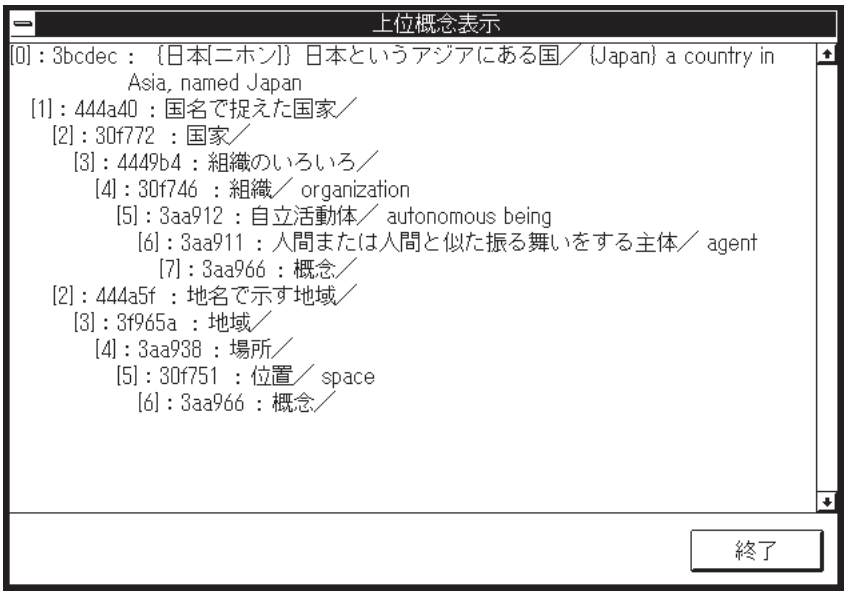
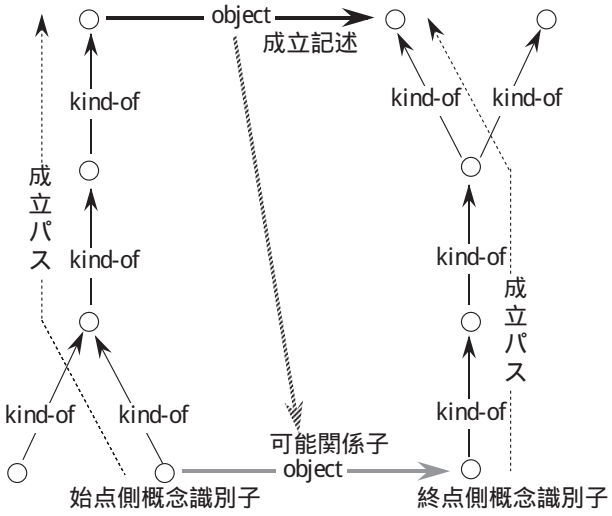


図 4.20　上位概念検索結果表示

4.4.8　可能関係子検索

可能関係子検索とは，概念記述辞書と概念体系辞書を利用して，ある特定の 2 概念間に，どのような概念関係が成り立ち得るか，を調べるものです．



すなわち，検索キーとして指定された 2 つの概念識別子（始点 / 終点）の上位概念を検索し，両者の上位概念をそれぞれ始点 / 終点として持つ概念記述があるかどうか，を検索します．概念体系の継承機能により，ここで検索された概念記述における関係子が，検索キーとして指定した 2 概念間にも成り立ち得る「可能関係子」となります．概念記述が複数検索されれば，複数の関係が 2 概念間に成り立つことになります．

なお，真偽値 0 の概念記述は，「その関係が成立しない」ことを意味しますので，可能関係子検索では，真偽値 1 の概念記述に到達するまでの経路（成立パス）上に，同じ関係子を持つ真

偽値 0 の概念記述が存在する場合には，その関係は成り立たないものとしています．

検索の実行

[検索]- [可能関係子検索](Ctrl+P) コマンドにより，図 4.21 に示すダイアログボックスが起動します．

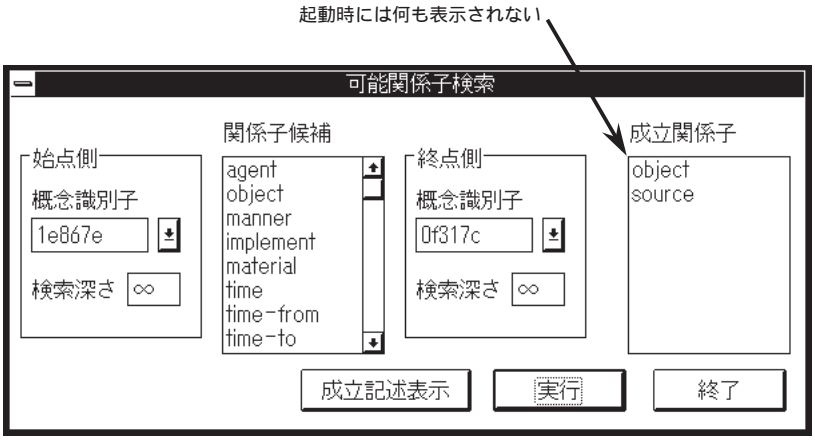


図 4.21　可能関係子検索ダイアログボックス

このダイアログボックスの動作は，他の検索ダイアログと多少異なります．まず，関係を調べたい 2 つの概念識別子 (始点，終点) を入力します．片方のみの指定では検索できません．

可能関係子検索では，あらかじめ関係子の候補を指定することができます．例えば，agent, object, implementのうちで 2 概念間に成り立つ関係はどれか，という検索が可能です．この場合には，中央の「関係子候補」リストボックスで，候補としたい関係子をクリックして反転させて下さい．候補は複数指定できます．なお，ここで何も指定しなかった場合には，すべての関係子を候補として検索を実行します．

最後に，「検索深さ」を設定します．検索深さとは，検索対象となる概念から，概念体系上で，何レベル上位の概念までを検索範囲とするか，を示します．デフォルトでは "∞" が表示されています．これは，深さ指定なし (すなわち概念体系上をすべて検索すること) を示しています．例えば，上位 5 レベルまでを検索範囲としたければ，始点 / 終点側の検索深さフィールドに 5 を入力して下さい．始点，終点に別々の深さを指定することも可能です．なお，数値は必ず半角文字で入力して下さい．

すべての指定が完了したら，【実行】をクリックします．検索が完了すると，右端の「成立関係子」リストボックス中に，対象となった 2 概念間に成立し得る関係子を一覧表示します．

ここで，【終了】をクリックすると，可能関係子検索を終了してメインウィンドウに戻りますが，可能関係子検索では，単に成立した関係子を表示するだけでなく，より詳細な情報を表示することができます．以下にその方法を述べます．

なお，ここで再度【実行】をクリックすると，同じ検索を再度実行しますので，ご注意下さい．

詳細情報の表示

実行が終了すると，【成立記述表示】ボタンがセンシティブになります．このボタンは，成立

した関係子が，どの概念記述レコードから導かれたものか，を表示するためのものです．一つの関係子が，複数の概念記述レコードから導かれた場合もありますが，その場合には，複数の概念記述を表示します．

最初に「成立関係子」リストボックス中で，概念記述レコードを表示させたい関係子を指定します．関係子は複数指定できます．また，何も指定しなかった場合には，すべての関係子を表示の対象とします．

指定が完了したら，【成立記述表示】をクリックして下さい．図 4.22に示す成立記述表示ダイアログボックスが起動します．

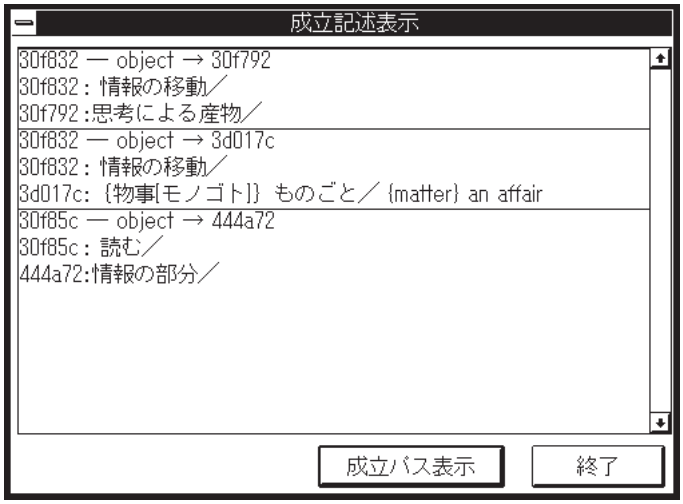


図 4.22　成立記述表示

ここでの操作は，「概念記述辞書検索結果表示」に似ています．レコードは，以下の形式でリスト表示されます．

〈概念識別子１〉－〈関係子〉→〈概念識別子２〉¥n
〈概念識別子１〉:〈概念見出し文字列〉¥n
〈概念識別子２〉:〈概念見出し文字列〉¥n

成立した関係子は真偽値が1であるもののみですので，真偽値の表示を省略しています．

ここで，【終了】をクリックすると，ダイアログボックスが消去され，可能関係子検索ダイアログボックスに戻ります．

【成立パス表示】ボタンは，表示されている概念記述から，検索キーとなった 2 つの概念まで，概念体系上のどのようなパスを経由しているか，を表示するためのものです．

まず，パスを調べたい概念記述レコードをクリックし，反転表示させます．一度に複数指定することも可能です．また，概念記述レコードが一つも指定されなかった場合には，全ての記述をパス表示対象とします．

指定が完了したら，【成立パス表示】をクリックして下さい．図 4.23に示す，成立パス表示ダイアログボックスが起動します．

ここでの表示は，〈関係子〉より上側が，検索キーとして指定した始点側の概念識別子から，概念記述の始点概念にたどりつくまでの，概念体系上のパスを示しています．下側は，終点側

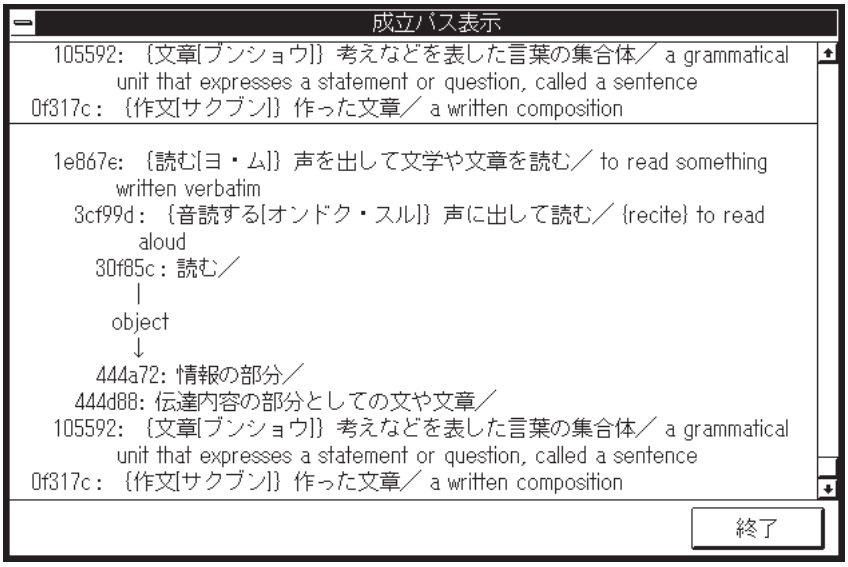


図 4.23　成立パス表示

のパスです．概念体系上で 1 つ上位になるごとに，右側に一定文字数分インデントして表示します．

複数の概念記述について，成立パス表示を実行した場合には，上記形式のパス情報が複数個表示され，個々のパス情報の間に仕切り線が引かれます．

【終了】をクリックすると，ダイアログボックスが消去され，成立記述表示ダイアログボックスに戻ります．

4.4.9　関連検索

関連検索とは，ユーザ情報に取り込んだレコードから，検索キー / フィルタキーとなり得るものを抽出して，これをもとに他の辞書を検索するためのインタフェースです．

まず，取りこまれているユーザ情報レコードのうち，関連検索のキーとしたいレコードをマウスで選択し，レコードIDを反転表示させます．次に ,[辞書検索]- [関連検索](Ctrl+A) を選択，あるいは レコードIDをダブルクリックすると，図 4.24に示す関連検索ダイアログボックスが起動します．

このダイアログボックスでは，3 種類の指定を行います．

検索対象辞書指定

ラジオボタンで，検索対象となる辞書を指定します．他の辞書だけでなく，ユーザ情報レコード自身が格納されていた辞書を指定することも可能です．起動時には，デフォルトで単語辞書が指定されています．

なお，ユーザ情報レコードが，概念体系，概念記述，概念見出しのいずれかである場合には，共起辞書は検索できませんので，共起辞書ラジオボタンはセンシティブではなくなります．

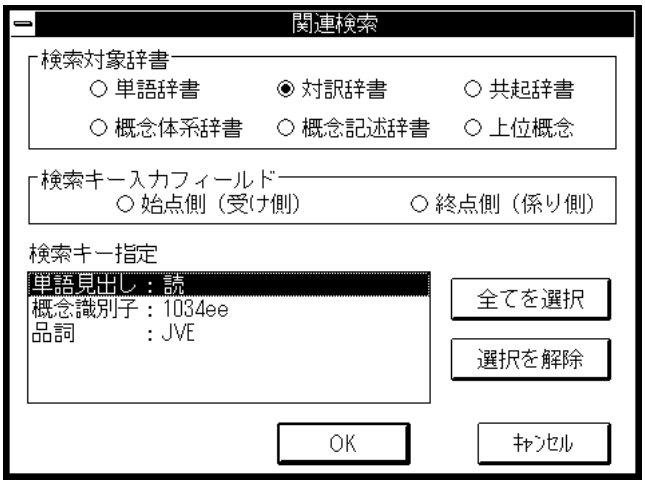


図 4.24　関連検索ダイアログボックス

また，辞書の指定は必ず最初に行って下さい．検索キーなどを指定してから，辞書指定を行なうと，指定した検索キーがリセットされてしまいます．

検索対象フィールド指定

このラジオボタンは，検索対象辞書が，概念体系辞書，概念記述辞書，共起辞書のいずれかの場合に限り，センシティブになります．例えば，単語辞書レコードの単語見出しから共起辞書を検索する場合に，受け側，係り側のどちらに検索キーを入力するのか，を指定するのがこのボタンです．

検索キー指定

検索キー指定リストボックスには，現在選択されているユーザ情報レコードの中から，検索キー／フィルタキーとなり得るものがリスト表示されます．どの辞書を検索対象とするかにより，検索キーとなり得るものが変化します．

ここで，どの情報を検索キー／フィルタキーとして採用するか，を指定します．リスト表示されているデータから，検索キーとしたいものをクリックして反転表示させて下さい．複数の指定が可能です．また，リストボックス横にある【全てを選択】ボタンは，一度に表示されているすべてのキーを選択します．【選択を解除】ボタンは，反転表示をクリアして，未選択状態に戻します．

指定の対象となる検索キーが一つしかない場合には，あらかじめ選択状態となります．

なお，ユーザ情報レコードが，概念記述／概念体系／共起のいずれかのレコードである場合には注意が必要です．リストボックスには，始点／終点（受け側／係り側）の両方の情報が表示されていますが，この双方を用いて他の辞書を検索することはできません．必ず，一つだけを指定（共起辞書の場合，受け側のペアか係り側のペアのみを指定）して下さい．また，これらのユーザ情報レコードの場合には，【全てを選択】ボタンはセンシティブにはなりません．

検索の実行

最後に，【OK】をクリックすると，検索に移ります．その際，検索キーの指定が不正であったり，指定が足りなかったりした場合には，エラーメッセージが表示されますので，再度指定をやり直して下さい．

検索実行時の動作は，「モード設定」ダイアログボックスで指定しているモードにより異なります．すなわち，「実行時に辞書検索ダイアログボックスを起動」が ON になっている場合には，指定された検索キーが各フィールドに入力された状態で，各辞書の検索ダイアログボックスが起動します．ここでより細かな指定をすることも可能です．その後の動作は，各辞書の検索手順と全く同様です．

モードが OFF の場合には，検索ダイアログボックスは起動せず，すぐに検索を開始します．検索結果が表示されてからの動作は，各辞書検索手順と同様です．

※ 共起辞書レコードから対訳辞書の検索（あるいはその逆）を実行する場合には注意して下さい．選択されているレコードが語尾を持つ場合，共起辞書には語尾なしで，対訳辞書には語尾つきでデータが格納されているため，検索に失敗します．この場合には，検索ダイアログボックスを起動するモードにしておき，起動したダイアログボックスで語尾を削って（あるいは付け足して）から，検索を実行して下さい．

なお，単語←→対訳の検索では，この問題はありません．単語→対訳では，語尾を自動的に追加し，対訳→単語では，語尾つき検索モードで検索します．

4.5 ユーザ情報の操作

辞書レコードが取り込まれると，メインウィンドウの表示が図4.25のようになります．

キーワード表示エリアには，現在のユーザ情報のキーワードが表示されます．これは，キーワー

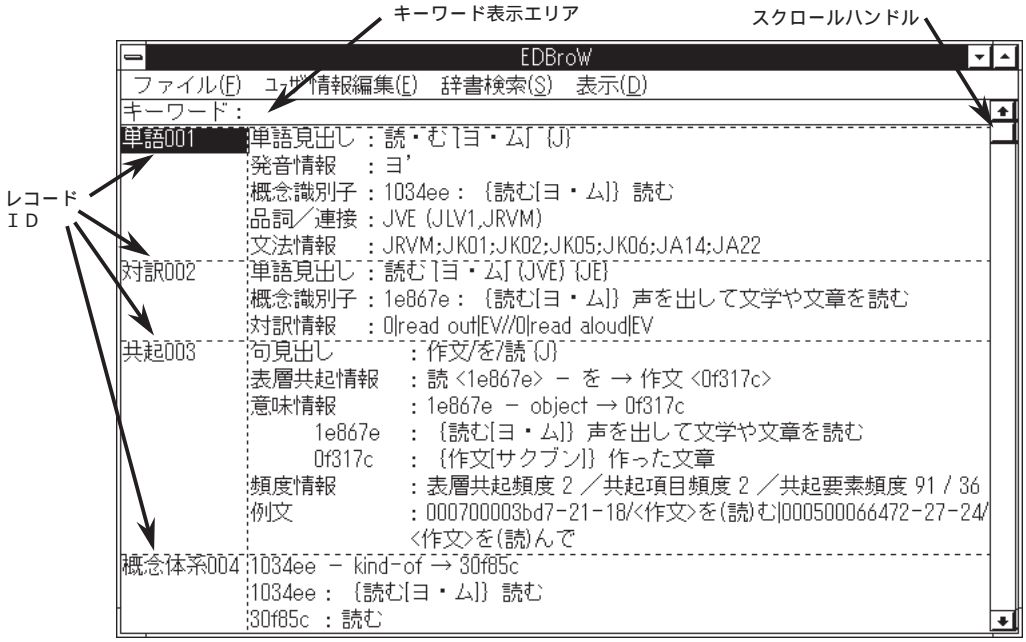


図 4.25 ユーザ情報の表示

ド設定ダイアログボックスにより，変更することが可能です．

また，レコードIDは，ユーザ情報の選択 / 非選択の対象となります．すなわち，ここでマウスをクリックすると，レコードIDが反転表示となり，変更 / コピー / 削除 / 関連検索の対象となります．再度クリックすると反転表示が解除され，非選択となります．なお，レコードIDをダブルクリックすると，関連検索が起動します．

ユーザ情報表示エリアでの，スクロールバーの動作は，以下のようになります．

- 【↑】をクリック
  ユーザ情報を，レコードの 1 フィールドづつ，上方向にスクロールします．ユーザ情報表示エリアの先頭行がユーザ情報の最初のレコードの第一フィールドである場合には，それ以上スクロールされません．
- 【↓】をクリック
  ユーザ情報を，レコードの 1 フィールドづつ，下方向にスクロールします．ユーザ情報表示エリアの先頭行がユーザ情報の最後のレコードの最終フィールドである場合には，それ以上スクロールされません．
- スクロールハンドルと【↑】の間をクリック
  ユーザ情報を，1 レコードづつ，上方向にスクロールします．ユーザ情報表示エリアの先頭行がユーザ情報の最初のレコードである場合には，それ以上スクロールされません．
- スクロールハンドルと【↓】の間をクリック
  ユーザ情報を，1 レコードづつ，下方向にスクロールします．ユーザ情報表示エリアの先頭行がユーザ情報の最後のレコードである場合には，それ以上スクロールされません．
- スクロールハンドルをドラッグ
  ユーザ情報を，1 レコードづつ，スクロールして表示します．スクロールハンドルを離した位置によって，先頭に表示されるユーザ情報が変化します．

4.5.1 ユーザ情報レコードの変更

最初に，ユーザ情報表示エリアで変更したいユーザ情報を選択します．次に，[ユーザ情報編集] – [変更](Ctrl+E) コマンドを選択すると，ユーザ情報の種類に応じて，各レコードの編集ダイアログボックスが起動します．

各ダイアログボックスの動作はほとんど同じです．ダイアログボックスには，以下の 4 つのボタンが用意されています．

【参照】　当該レコードに付与する参照情報を入力するための，「参照」ダイアログボックスを起動します．

【覚書】　当該レコードに付与する覚書情報を入力するための，「覚書」ダイアログボックスを起動します．

【OK】　ダイアログボックス上での変更事項をユーザ情報レコードに反映し，ダイア

ログボックスを終了します．ユーザ情報表示エリアが，変更の結果を受けて更新されます．

【キャンセル】　変更を中断します．

編集ダイアログボックスで，各フィールドの情報を編集して下さい．以下に，各レコードの編集ダイアログボックスごとに，入力時の注意点を記します．

単語レコード編集

図 4.26に，単語レコード編集ダイアログボックスを示します．

図 4.26　単語レコード編集ダイアログボックス

このダイアログボックスは，対象レコードが慣用句であるかないか（慣用句指定チェックボックスにより指定）によって，動作が異なります．

慣用句でない場合　「読み」の下の入力フィールドは，「発音情報」になります．また，「構文木」のフィールドはセンシティブになりません．

慣用句の場合　「読み」の下の入力フィールドは，「構成語情報」になります．また「構文木」のフィールドがセンシティブになりますが，「左連接情報」「右連接情報」のフィールドはセンシティブになりません．

言語識別はラジオボタンで，慣用句であるかないか，専門語であるかないか，はチェックボックスにより指定します．言語識別の指定により，「品詞」「左連接情報」「右連接情報」のコンボボックスのリスト表示が，それぞれの言語の文法コードに変わります．

概念識別子は，16進数6桁（半角）で入力して下さい．

対訳レコード編集

図 4.27に，対訳レコード編集ダイアログボックスを示します．

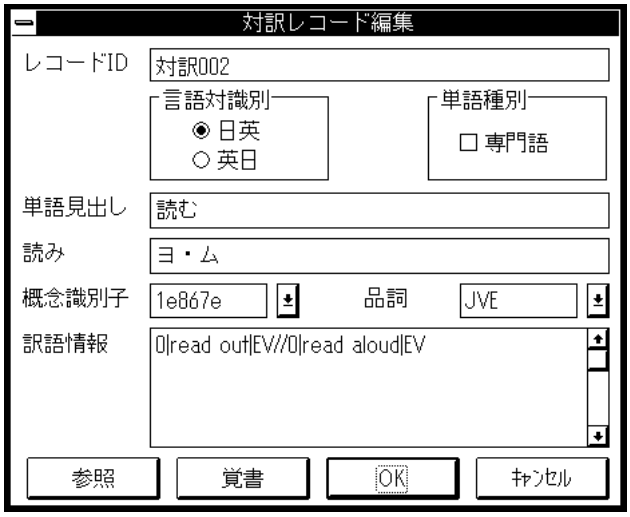


図 4.27　対訳レコード編集ダイアログボックス

言語対識別はラジオボタンで，専門用語であるかないか，はチェックボックスで，それぞれ指定します．言語対識別の指定によって，「品詞」コンボボックスの品詞リストが，各言語対に対応するものに変わります．概念識別子は，16進数6桁（半角）で入力して下さい．

共起レコード編集

図 4.28に，共起レコード編集ダイアログボックスを示します．

図 4.28　共起レコード編集ダイアログボックス

言語識別はラジオボタンで，受け側 / 係り側の慣用句フラグはチェックボックスで指定します．共起レコードの場合，言語識別を変更しても，品詞入力コンボボックスのリスト表示は変更されません．

関係子方向とは，表層共起情報（構文情報）と，意味情報との対応をとるためのものです．受け側概念識別子と，意味情報中の始点概念識別子が一致する場合には，↓ を，そうでない場合には↑ を指定します．

各概念識別子は，16進数6桁（半角）で，各頻度情報は10進整数（半角）で入力して下さい．また，例文中では，必ず受け側単語見出しは( )で，係り側単語見出しは< >でくくって下さい．

概念体系レコード編集

図 4.29に，概念体系レコード編集ダイアログボックスを示します．

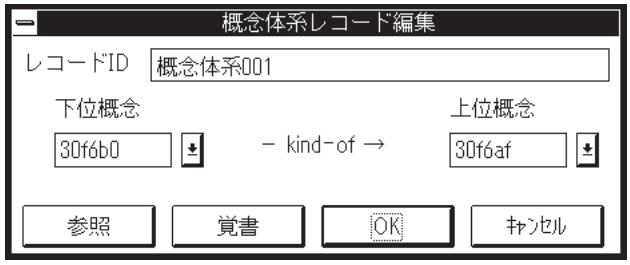


図 4.29 概念体系レコード編集ダイアログボックス

各概念識別子は，16進数6桁（半角）で入力して下さい．

概念記述レコード編集

図 4.30に，概念記述レコード編集ダイアログボックスを示します．

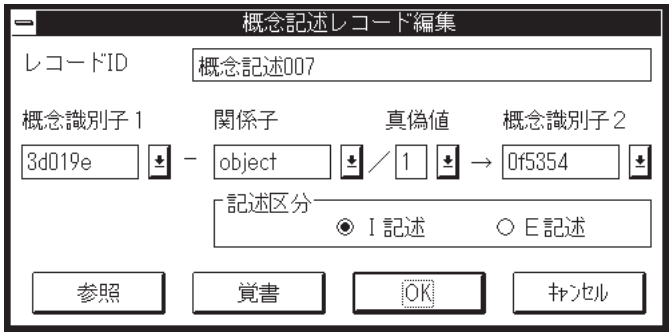


図 4.30 概念記述レコード編集ダイアログボックス

記述区分はラジオボタンで指定します．真偽値は，0 または 1（半角）を，概念識別子は，16進数6桁（半角）で，それぞれ入力して下さい．

概念見出しレコード編集

図 4.31に，概念見出しレコード編集ダイアログボックスを示します．

図 4.31　概念見出しレコード編集ダイアログボックス

概念見出しレコードは，ユーザ情報中の他のレコードが，表示の際に利用しています．したがって，ここでの変更結果は他のユーザ情報にも影響します（概念見出し文字列が変更されます）．

上記の理由により，データの整合性を保つため，レコード変更処理の場合には，このダイアログボックスの概念識別子のフィールドはセンシティブになりません（レコード新規作成の場合にはセンシティブになります）．

覚書情報の付与

各レコードの編集ダイアログで，【覚書】をクリックすると，図 4.32に示す覚書編集ダイアログボックスが起動します．

図 4.32　覚書入力ダイアログボックス

「覚書情報編集エリア」で，レコードに付与するメモなどを入力した後，【OK】をクリックすると，当該ユーザ情報レコードに覚書が付与されます．

参照情報の付与

各レコードの編集ダイアログで，【参照】をクリックすると，図 4.33に示す参照情報編集ダイアログボックスが起動します．

図 4.33 参照入力ダイアログボックス

「参照指定リストボックス」中には，現在ユーザ情報中にあるレコードIDの一覧が表示されています．レコードを参照情報として指定するには，レコードIDをクリックして反転表示させて下さい．また，参照を解除するには，再度クリックして，通常表示に戻して下さい．

なお，ダイアログ起動時にすでに参照情報として指定されているレコードIDは，最初から反転表示となっています．

参照情報の指定が終了したら，【OK】をクリックすれば，反転表示になっているレコードIDが参照情報として付与されます．

4.5.2 ユーザ情報レコードの新規作成

[ユーザ情報編集]- [新規] のサブメニューから，新規作成したいレコードを選択すると，「レコードの変更」で起動したものと同じ編集ダイアログボックスが起動します．起動時にデータがまったく入っていないことを除き，操作等は変更の際と同様です．

新規作成した場合には， 新しいレコードIDを付与した新規レコードが，ユーザ情報中に作成されます．

4.5.3 ユーザ情報レコードの削除

ユーザ情報表示エリアでレコードを選択し ,[ユーザ情報編集]- [削除](Del) コマンドを選択します．選択されたレコードが削除されます．

なお，レコードが削除された場合に，レコードIDの付けかえは行われません．したがって，例えば「単語003」というレコードを削除した場合，それ以降に「 ~003」というレコードは作成されません．

4.5.4 ユーザ情報レコードのコピー

ユーザ情報表示エリアでレコードを選択し ,[ユーザ情報編集]- [コピー](Ctrl+V) コマンドを選択します . 選択されたレコードの内容をコピーし , 新たなレコードID を付与した新規レコードが作成されます .

4.5.5 取消

[ユーザ情報編集]- [取消](Ctrl+U) コマンドを選択することにより , ユーザ情報の状態を , 一つ前に戻すことができます . すなわち , 一つ前にユーザ情報に対して行なった動作により , 以下のような処理となります .

「ユーザ情報クリア」の後	クリアする前の状態に戻します .
「レコード削除」の後	削除したレコードを復活します .
「レコードコピー」の後	コピーしたレコードを消去します .
「辞書検索結果の取込」の後	取り込んだレコードを消去します .
「レコード変更」の後	変更したレコードを元の状態に戻します .

一旦「取消」を実行すると , 取消実行前の状態に戻すことはできません . また , ユーザ情報に対して変更がなされていない場合には , このメニュー項目はセンシティブにはなりません .

4.5.6 キーワードの設定

ユーザ情報に付けられるキーワードは ,[ユーザ情報編集]- [キーワード] コマンドにより起動する , キーワード設定ダイアログボックスで設定することができます . 図 4.34に , ダイアログボックスを示します .

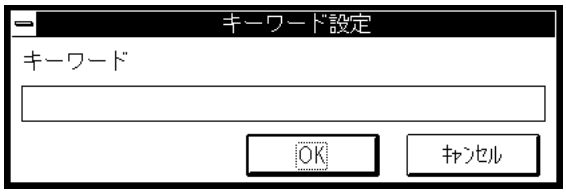


図 4.34 キーワード入力ダイアログボックス

「キーワード入力エリア」でキーワードを指定し , OK をクリックします . 以降 , ユーザ情報表示エリアの「キーワード : 」欄に , ここで設定したキーワードが表示されるようになります .

なお , キーワードは , ユーザ情報保存時にも変更することができます . 詳しくは , 「ユーザ情報の保存」の項を参照して下さい .

4.5.7 ユーザ情報のクリア

現在あるユーザ情報をクリアして , 空の状態にします . その際 , ユーザ情報のファイルへの保存を確認するメッセージボックスが起動しますので , 【はい】【いいえ】のどちらかをクリックして下さい .

ユーザ情報がクリアされ , システム起動時と同じ状態に戻ります .

4.5.8 ユーザ情報の保存

ユーザ情報をファイルに保存するには，[ファイル]ー[ユーザ情報保存](Ctrl+S) コマンドを選択します．ここで，図 4.35に示すユーザ情報保存ダイアログボックスが起動します．

図 4.35 ユーザ情報保存ダイアログボックス

ファイル名入力エリアには，デフォルトのファイル名が表示されています．現在のユーザ情報が，既存のユーザ情報ファイルから読み込んだものであればそのファイル名が，そうでない場合には，デフォルトファイル名である "UINFO.UIF" が表示されています．ファイル名を変更したい場合は，入力し直して下さい．なお，ユーザ情報ファイルの拡張子は，".UIF" で固定です．拡張子が違う場合には，自動的に ".UIF" に変更します．

キーワード入力エリアには，現在のユーザ情報に付与されているキーワードが表示されます．変更したい場合は，このエリアで変更して下さい．

設定が終わったら，【OK】をクリックして下さい．ユーザ情報が，指定されたファイル名で，EDBroW 実行時のディレクトリに保存されます．

なお，このダイアログボックスは，ユーザ情報クリア時や，システム終了時に，「現在のユーザ情報を保存する」を選択した場合にも起動します．動作はここで述べた場合と同様です．

4.5.9 ユーザ情報の読み込み

ユーザ情報をファイルから読み込むには，[ファイル]ー[ユーザ情報読込](Ctrl+L)コマンドを選択します．ここで，図 4.36に示すユーザ情報読込ダイアログボックスが起動します．

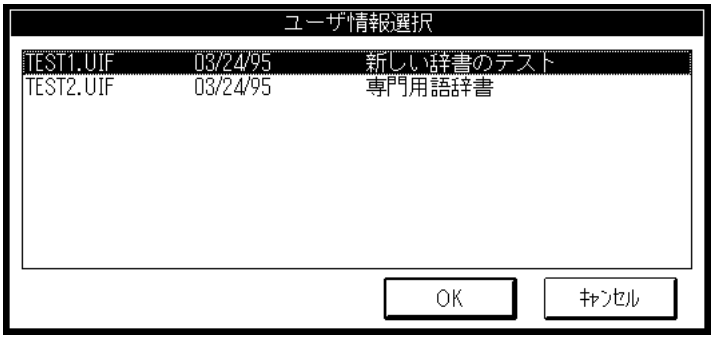


図 4.36 ユーザ情報読込ダイアログボックス

各ユーザ情報ファイルが，以下の形式でリスト表示されます．

〈ファイル名〉¥t〈日付〉¥t〈キーワード〉

読み込みたいファイルをマウスでクリックして反転表示させ，【OK】をクリックすれば，ファ

イル中のユーザ情報を読み込み，ユーザ情報表示エリアに表示します．

なお，ユーザ情報ファイルは，EDBroW 実行時のディレクトリにある，拡張子 が ".UIF" であるファイルのみが対象です．

4.5.10 ユーザ情報の印刷

現在のユーザ情報を，プリンタに出力します．出力の形式は，「詳細表示」モード時の形式と同様です．

なお，起動するダイアログボックスの内容は，システムにより異なります．

4.6 その他の操作

4.6.1 モードの変更

[ファイル]-[モード変更](Ctrl+D)コマンドを実行すると，図4.37に示すモード変更ダイアログボックスが表示されます．

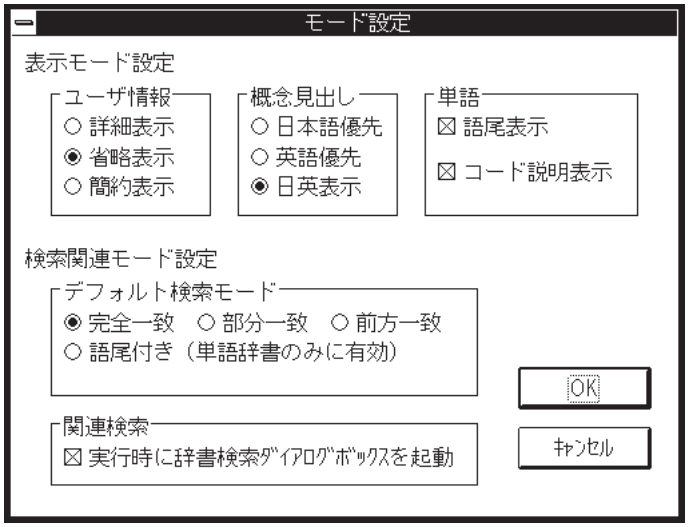


図 4.37 モード設定ダイアログボックス

モード変更は，「ユーザ情報表示モード」「概念見出し表示モード」「デフォルト検索モード」はラジオボタンでモードを選択，その他のモードはチェックボックスで ON/OFF を選択できるようになっています．

モードの選択が終了したら，OK をクリックして下さい．モードが変更されます．

ユーザ情報表示モードについては，メニューから直接モード変更することもできます．その方法は以下の通りです．

- [表示] メニューをクリックすると，3 種類のモードがメニュー上に表示され，現在選択されているモードにチェックマークが付いている．
- 選択したいモードの上までマウスを移動し，再度クリックすると，モードが変更される．

4.6.2 辞書の切換

EDBroW では，システムの動作中に，システム辞書ファイルを切りかえることもできます．デフォルトでは，システム初期化ファイルに記述されているパス名の辞書を利用しますが，それ以外の辞書を利用したい場合には，[ファイル]- [辞書切換](Ctrl+J) を選択して下さい．図 4.38に示すダイアログボックスが表示されます．

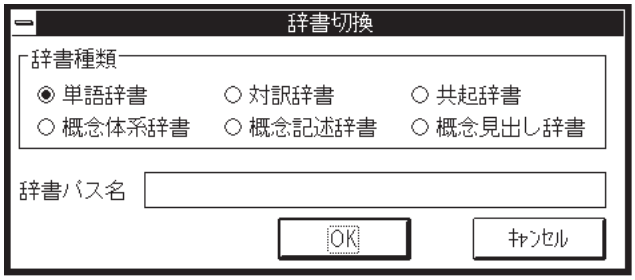


図 4.38 辞書切換ダイアログボックス

ここで，ラジオボタンにより変更したいシステム辞書の種類を入力し，対応するシステム辞書のパス名を入力します．最後に【OK】をクリックすれば，指定されたシステム辞書が切りかわります．

ここで，辞書のパス名が間違っている，あるいは，指定した辞書種類と，パス名で指定した辞書ファイルとの整合がとれない (種類が異なる) 場合には，エラーメッセージが表示されますが，そうでない場合には，何も表示されません．

第 5 章

システム辞書仕様

本章では，EDR から提供されるテキスト形式の辞書 (以下，テキスト辞書と呼ぶ) と，辞書トランスレータにより作成される EDBroW 用の辞書 (以下，システム辞書と呼ぶ) について，その対応関係を示す．

5.1 各辞書に共通な事柄

概念識別子

テキスト辞書では，概念識別子はそのほとんどが16進数6桁の文字列で記述されているが，一部の概念識別子 (数字を表す概念) では，概念識別子が16進数で6桁より大きくなることがある．

また，システム辞書では，概念識別子は unsigned long で格納している．このため，テキスト辞書中で16進数6桁を越える概念識別子については，別途用意するテーブルにより変換した後に格納する．

変換の対象となっているのは，以下の 6 つの概念識別子である．

テキスト辞書中の 概念識別子	システム辞書中の 概念識別子
000103e8	a00001
00012710	a00002
00010f4240	a00003
000105f5e100	a00004
0001e8d4a51000	a00005
00012386f26fc10000	a00006

なお，上記以外で，テキスト辞書中で6桁を越える概念識別子を含むレコードは，エラーとしてシステム辞書には格納されない．

文法コード

品詞コード，左右連接属性コードについては，テキスト辞書中のコード文字列を，本章末に示すテーブルにより，数値としてシステム辞書に格納する．

また，日本語単語辞書，日英対訳辞書において，"JN1;JVE" という品詞は，サ変名詞を表わすものであり，システム辞書では，"JSA" (品詞番号 38) として格納する．

品詞，連接以外の文法コードについては，文字列をそのまま格納する．

以下に，システム辞書に格納されるデータの名称と，対応するテキスト辞書の名称との対応を示す．

5.2 単語辞書

システム辞書名称	テキスト辞書名称	備考
単語見出し *1	不変化部分	
読み *2	カナ表記 (日本語) 音節区切り (英語:慣用句以外) 見出し表記 (英語:慣用句)	
発音情報 *3	発音情報 (慣用句以外)	
構成語情報 *3	単語見出し (慣用句のみ)	
概念識別子	概念識別子	
左連接情報	左連接情報	慣用句の場合には
右連接情報	右連接情報	格納しない
品詞	品詞	
言語識別 *4	–	
レコード種別 *5	–	
慣用句フラグ *6	–	
文法情報	品詞 / 構文木以外の 文法情報 + 用法	各文法情報コードを ';' で連結して格納
構文木	構文木	慣用句以外は空

*1 慣用句の場合には，先頭の 1 単語のみを格納する．

*2 英語単語の慣用句の場合，テキスト辞書では音節区切りのフィールドが空であるため，かわりに見出し表記を格納する．

*3 発音情報 / 構成語情報 は，慣用句とそれ以外のレコードでフィールドが重複しないため，システム辞書中では同フィールドとして格納する．

*4 レコードIDの別により判定．"JWD" を含む場合には 0 を，"EWD" を含む場合には 1 を格納

*5 レコードIDの別により判定．先頭が 'T' で始まる場合には 1 を，それ以外は 0 を格納

*6 主に管理のための情報．レコードが構文木を持つ場合に 1 を，そうでない場合に 0 を格納

※ただし，テキスト辞書の頻度情報は，システム辞書には格納しない

5.3 対訳辞書

システム辞書名称	テキスト辞書名称	備考
単語見出し	単語見出し	
読み	カナ表記 (日英)	英日の場合は空
品詞	品詞	
概念識別子	概念識別子	
言語対識別 *1	–	
レコード種別 *2	–	
対訳情報	対訳情報	

*1 レコードIDの別により判定．"JEB" を含む場合には 0 を，"EJB" を含む場合には 1 を格納

*2 レコードIDの別により判定．先頭が 'T' で始まる場合には 1 を，それ以外は 0 を格納

5.4 共起辞書

システム辞書名称	テキスト辞書名称	備考
受け側単語見出し	受け側単語表記	
受け側表記情報	表記情報 *1	カナ表記もしくは原形
受け側品詞	品詞 *1	
受け側概念識別子	概念識別子 *2	
受け側慣用句フラグ	慣用句フラグ *1	
共起関係子	共起関係子	
関係単語表記	関係単語表記	
係り側単語見出し	係り側単語表記	
係り側表記情報	表記情報 *1	カナ表記もしくは原形
係り側品詞	品詞 *1	
係り側概念識別子	概念識別子 *2	
係り側慣用句フラグ	慣用句フラグ *1	
概念関係子	概念関係子 *3	
関係子方向 *4	–	
表記順 *4	–	
言語識別 *5	–	
表層共起頻度	表層共起頻度	
共起項目頻度	共起項目頻度	
受け側共起要素頻度	受け側共起要素頻度	
係り側共起要素頻度	係り側共起要素頻度	
例文	例文	情報を省略して格納 *6

*1 受け側 / 係り側の各情報は，共起句要素情報の中から，受け側 / 係り側に対応するもののみを格納

*2 概念識別子が「補足付き概念説明」である場合には，他の概念識別子と重ならない識別子を新設して格納することが可能．ただし，新設した概念識別子については，「補足付き概念説明」文字列を「日本語概念説明」のフィールドに格納した形式の概念見出し補助ファイルを作成し，他の概念見出し辞書ファイルとマージする (辞書トランスレータにより，格納方法を選択可能)

*3 概念関係子が "" である場合には，関係子 "nil"(関係子番号 80) を格納

*4 要素番号をすべて格納しないかわりに，意味情報における始点 / 終点側と，受け側 / 係り側との対応をとるために関係子方向を利用．始点 = 受け側の場合に 0 を，始点 = 係り側の場合に 1 を格納．また，句見出しを格納しないかわりに，句見出しにおける受け側 / 係り側単語の出現順を表記順とし

て格納する．受け側が先行する場合に 1 を，係り側が先行する場合に 0 を格納．

*5 レコードIDの別により判定．"JCC" を含む場合には 0 を，"ECC" を含む場合には 1 を格納

*6 例文は，1 例文につき文IDが複数個ある場合には，先頭の文IDのみを格納し，最大 5 例文を格納する．

※句見出しは格納しないが，受け側 / 係り側単語見出し，関係単語表記，表記順により，再構成が可能である．

5.5 概念見出し辞書

システム辞書名称	テキスト辞書名称	備考
概念識別子	概念識別子	
日本語概念見出し	日本語概念見出し	
英語概念見出し	英語概念見出し	
日本語概念説明	日本語概念説明	
英語概念説明	英語概念説明	

5.6 概念体系辞書

システム辞書名称	テキスト辞書名称	備考
上位概念	上位概念	
下位概念	下位概念	

5.7 概念記述辞書

システム辞書名称	テキスト辞書名称	備考
概念識別子 1	概念識別子 1	
概念識別子 2	概念識別子 2	
関係子	関係子	
真偽値	真偽値	
記述区分 *1	記述区分	

*1 J記述を 1，E記述を 0 として格納

5.8 コード一覧

次ページ以降に，左右連接情報，品詞，概念関係子のコード番号と，対応するコードを示す．

5.8.1 日本語左連接情報コード

番号	コード
0	–
1	JLN1
2	JLN2
3	JLN3
4	JLN4
5	JLN5
6	JLN6
7	JLN7
8	JLN8
9	JLV1
10	JLA1
11	JLM1
12	JLD2
13	JLR1
14	JLC1
15	JLC3
16	JLT1
17	JSF1
18	JLJ6
19	JSF4
20	JSF5
21	JSF7
22	JSF8
23	JSF9
24	JSFA
25	JLS1
26	JLS2
27	JLJ1
28	JLJ2
29	JLJ3
30	JLJ4
31	JLJ5
32	JLJ7
33	JLJ8
34	JLJ9
35	JLJA
36	JLJB
37	JLWA
38	JLJD
39	JLJE
40	JLJF
41	JLJG
42	JLJH
43	JLJI
44	JLJJ
45	JLJK
46	JLJL
47	JLJM
48	JLJN
49	JSD1
50	JSD2
51	JSD4
52	JLW1
53	JLW2
54	JLW3
55	JLW4
56	JLW5
57	JLW6
58	JLW7
59	JLW8
60	JLW9
61	JLWA
62	JLWB
63	JLWC
64	JLWD
65	JLWE
66	JLWF
67	JLWG
68	JLWH
69	JLWI
70	JLWJ
71	JLWK
72	JSVK
73	JSVG
74	JSVS
75	JSVT
76	JSVN
77	JSVB
78	JSVM
79	JSVR
80	JSVW
81	JSV1
82	JSV2
83	JSV3
84	JSV4
85	JSV5
86	JSV6
87	JSV7
88	JSA0
89	JSA1
90	JSA2
91	JSA3
92	JSM0
93	JSM1
94	JSM2
95	JSM3
96	JSM4
97	JSMA
98	JSM5
99	JSM7
100	JSM8

5.8.2 日本語右連接情報コード

番号	コード
0	–
1	JRN1
2	JRN2
3	JRN3
4	JRN4
5	JRN5
6	JRN6
7	JRN7
8	JRN8
9	JRN9
10	JRNA
11	JRNB
12	JRND
13	JRNE
14	JRNF
15	JRVK
16	JRVG
17	JRVS
18	JRVT
19	JRVN
20	JRVB
21	JRVM
22	JRVR
23	JRVW
24	JRV1
25	JRV2
26	JRV3
27	JRV4
28	JRV5
29	JRV6
30	JRV7
31	JRA1
32	JRA2
33	JRA3
34	JRM1
35	JRM2
36	JRM3
37	JRM4
38	JRMA
39	JRM5
40	JRM9
41	JRM7
42	JRM8
43	JRD2
44	JRD3
45	JRD4
46	JRD5
47	JRD6

48	JRD7
49	JRD8
50	JRD9
51	JRDA
52	JRDB
53	JRDC
54	JRDD
55	JRDE
56	JRDF
57	JRDG
58	JRDH
59	JRR1
60	JRC1
61	JRC2
62	JRT1
63	JRP1
64	JRP2
65	JRP3
66	JRP4
67	JRP5
68	JRP6
69	JRP7
70	JRS1
71	JRS2
72	JRJ1
73	JRJ2
74	JRJ3
75	JRJ4
76	JRJ5
77	JRJ6
78	JRJ7
79	JRJ8
80	JRJ9
81	JRJA
82	JEM1
83	JEM2
84	JEM3
85	JEM4
86	JEM5
87	JEM6
88	JEM7
89	JEM8
90	JED1
91	JED2
92	JED3
93	JED4
94	JED5
95	JED6
96	JED7
97	JED8
98	JED9
99	JEDA
100	JEE1
101	JEE2
102	JEE3
103	JEE4
104	JEE5
105	JEE6
106	JEA1
107	JEA2
108	JEC1
109	JEC2
110	JEI1
111	JEI2

5.8.3 日本語品詞コード

番号	コード
0	–
1	JN1
2	JN2
3	JN3
4	JN4
5	JN7
6	JVE
7	JAJ
8	JAM
9	JD1
10	JD2
11	JNM
12	JC1
13	JC3
14	JT1
15	JT2
16	JT3
17	JT4
18	JN5
19	JB1
20	JUN
21	JN6
22	JEV
23	JEA
24	JEM
25	JNP
26	JPR
27	JAP
28	JMP
29	JIP
30	JSE
31	JJO
32	JJ1
33	JJD
34	JJP
35	JAX
36	JIT
37	JSY
38	JSA

5.8.4　英語左連接情報コード

番号	コード
0	–
1	ELB1
2	ELN1
3	ELN2
4	ELP1
5	ELP2
6	ELA1
7	ELA3
8	ELD1
9	ELD3
10	ELV1
11	ELV2
12	ELT1
13	ELT2
14	ELW1
15	ESN1
16	ESN2
17	ESN3
18	ESN4
19	ESN5
20	ESN6
21	ESN7
22	ESN8
23	ESF1
24	ESF2
25	ESF3
26	ESF8
27	ESF4
28	ESF5
29	ESF6
30	ESF7
31	ESV1
32	ESV2
33	ESV3
34	ESV4
35	ESV5
36	ESV6
37	ESV7
38	ESV8
39	ESV9
40	ESVA
41	ESVB
42	ESPB
43	ESPD
44	ESPF
45	ESPG
46	ESPK
47	ESPL
48	ESPM
49	ESPN
50	ESPP
51	ESPR
52	ESPT
53	ESPV
54	ESA1
55	ESA2
56	ESA3
57	ESAB
58	ESAD
59	ESAF
60	ESAG
61	ESAK
62	ESAL
63	ESAM
64	ESAN
65	ESAP
66	ESAR
67	ESAS
68	ESAT
69	ESAV
70	ESAZ
71	ESD1
72	ESD2
73	ESD3
74	ESDB
75	ESDD
76	ESDF
77	ESDG
78	ESDK
79	ESDL
80	ESDM
81	ESDN
82	ESDP
83	ESDR
84	ESDS
85	ESDT
86	ESDV
87	ESDZ
88	ELS1
89	ELS2
90	ELS3
91	ELS4
92	ELS5
93	ELS6
94	ELS7
95	ELS8
96	ELS9
97	ELSA
98	ELSB
99	ELSC

5.8.5　英語右連接情報コード

番号	コード
0	–
1	ERB1
2	ERR1
3	ERR2
4	ERR3
5	ECF1
6	ECN1
7	ECN2
8	ECN3
9	ECN4
10	ECN5
11	ECN6
12	ECN7
13	ECN8
14	ERN1
15	ERN2
16	ERN3
17	ERN4
18	ERN5
19	ERN6
20	ERN7
21	ERP1
22	ERP2
23	ERP3
24	ERP4
25	ERP5
26	ERP7
27	ERP8
28	ERP9
29	ERPA
30	ERPB
31	ERPC
32	ERPD
33	ERPE
34	ERL1
35	ERL2
36	ERL3
37	ERL4
38	ERL5
39	ERQ1
40	ERQ2
41	ERQ3
42	ERQ4
43	ERQ6
44	ERM1
45	ERM2
46	ERE1
47	ERE2

48	ERE3
49	ERE4
50	ERE5
51	ERE6
52	ERE7
53	ERE8
54	ERO1
55	ERO2
56	ERO3
57	ERH1
58	ERH2
59	ERH3
60	ERH4
61	ERX1
62	ERC1
63	ERC2
64	ECV1
65	ECV2
66	ECV3
67	ECV4
68	ECV5
69	ECV6
70	ECV7
71	ECV8
72	ECV9
73	ECVA
74	ECVB
75	EPVB
76	EPVD
77	EPVF
78	EPVG
79	EPVK
80	EPVL
81	EPVM
82	EPVN
83	EPVP
84	EPVR
85	EPVT
86	EPVV
87	ERV1
88	ERV2
89	ERV3
90	ERV4
91	ERV5
92	ECA1
93	ECA2
94	ECA3
95	ECAB
96	ECAD
97	ECAF
98	ECAG
99	ECAK
100	ECAL
101	ECAM
102	ECAN
103	ECAP
104	ECAR
105	ECAS
106	ECAT
107	ECAV
108	ECAZ
109	ERA1
110	ERA2
111	ERA3
112	ECD1
113	ECD2
114	ECD3
115	ECDB
116	ECDD
117	ECDF
118	ECDG
119	ECDK
120	ECDL
121	ECDM
122	ECDN
123	ECDP
124	ECDR
125	ECDS
126	ECDT
127	ECDV
128	ECDZ
129	ERD1
130	ERD2
131	ERD3
132	ERD5
133	ERIT
134	ERPP
135	ERS1
136	ERS2
137	ERS3
138	ERS4
139	ERS5
140	ERS6
141	ERS7
142	ERS8
143	ERS9
144	ERSA
145	ERSB
146	ERPF
147	ERTO
148	ERSC

5.8.6 英語品詞コード

番号	コード
0	–
1	EN1
2	EN2
3	EN3
4	EN4
5	EN5
6	EP1
7	EP2
8	EP3
9	EP4
10	EP5
11	EVE
12	EBE
13	EAJ
14	ED1
15	ED2
16	ED3
17	ED4
18	ED5
19	EPR
20	EPR2
21	ET1
22	ET2
23	EAR
24	EAV
25	EAV2
26	EIT
27	EC1
28	EC12
29	EC2
30	EC22
31	EPF
32	EEN
33	EEV
34	EEA
35	EED
36	ENP
37	EVP
38	EAP
39	EDP
40	EPP
41	EIP
42	ESE
43	EUN
44	ESY

5.8.7 共起辞書品詞コード

番号	文字列
0	–
1	名詞
2	動詞
3	形容詞
4	形容動詞
5	副詞
6	接尾語
7	NOUN
8	PRON
9	DEMO
10	INDEF
11	VERB
12	ADJ
13	ADV
14	UNIT
15	PTCL

※ 1 ~ 6 までは日本語共起辞書用，8 以降は英語共起辞書用

5.8.8 概念関係子コード

番号	文字列
0	–
1	agent
2	object
3	manner
4	implement
5	material
6	time
7	time-from
8	time-to
9	duration
10	location
11	place
12	source
13	goal
14	scene
15	beneficiary
16	a-object
21	condition
22	cooccurrence
23	sequence
31	part-of
32	element-of
42	unit
43	quantity
44	degree
45	number
46	modifier
51	possessor
52	purpose
53	cause
54	basis
55	and
56	or
57	from-to
60	kind-of
61	equivalent
62	similar
71	statement
80	nil
101	all

※ nil は共起辞書用

第 6 章

ライブラリ解説

本章では，システム辞書検索用のライブラリについて，その仕様と，検索関数，アクセス関数について説明する．

6.1 概要

辞書検索ライブラリは，検索用関数と辞書レコードアクセス関数とから構成される．詳細な関数の仕様については，6.2 に示す．

検索キーを用いた検索の手順は，以下の通りである．

- 検索キーを用いて，インデックスパートを二分探索し，該当する最初のエントリと最後のエントリを見つける．
- 各エントリの項目番号から固定長辞書パートでのレコードのオフセットを求め，固定長レコードにアクセスする．
- レコードが検索条件に合致していれば，検索結果として保存し，相違していれば，棄却する．保存の際，可変長フィールドがある場合には，オフセットから可変長パートにアクセスし，情報を得る．

検索モードとしては，「完全一致」「部分一致」「前方一致」の3種類を提供する．各々の検索モードと検索結果は，以下の通りである．

検索キー	検索モード	検索結果
東京	完全一致	「東京」のみ
	部分一致	「東」「東京」
	前方一致	「東京」「東京都」「東京音頭」など

ただし，概念識別子など，検索キーが可変長でないもの場合には，検索モードとして完全一致のみが有効であり，その他のモードが指定された場合にも，完全一致により検索を行う．

検索結果は，以下に示す辞書レコードのリストとなる．リストの末端では，nextPtrには NULL が格納される．

```
typedef struct EDRecList {
        struct EDRecList    *nextPtr;     //次のレコードへのポインタ
        unsigned int        dicType :4;   //辞書レコードの種類
        unsigned int        selected:1;   //選択済みフラグ
        unsigned int        marked  :1;   //マーク付きフラグ
        unsigned char       *dictRec;     //辞書レコードへのポインタ
} EDRecList;
```

※ 選択済みフラグ，マーク付きフラグは，EDBroW でのみ利用されるものであり，通常は必要ない．

辞書ファイルのオープン / クローズ

各辞書ファイルは構造を持っている．このため，辞書構造体（構造体名 DictPtr ）を用意し，辞書固有の情報をここに格納するとともに，検索のための辞書指定には，この辞書構造体を利用する．

辞書構造体に格納されるデータは，システム辞書の先頭に格納されているものを読み込んだものである．また，辞書構造体にデータをセットするための，辞書ファイルオープン関数，および辞書クローズ関数が，辞書検索ライブラリの一部として提供される．

複数キー / フィルタキー指定による検索

インデックスに対応する複数のキーと，インデックスには対応しないが，フィルタとして利用できるキーとを与え，必要な情報のみを取り出す関数を用意する．この関数は，どれか一つのキーを利用して辞書を検索し，結果をその他のキーによりフィルタリングすることにより実現される．複数キー指定の場合の，検索モードとフィルタリング方法の対応は以下の通りである．

キーの種類	フィルタリング方法
インデックスに対応する検索キー	
可変長キーの場合	与えられた検索モードに依存
固定長キーの場合	完全一致するもののみ
インデックスに対応しないフィルタキー	完全一致するもののみ

いくつかの特殊な検索は，検索モードと複数 / フィルタキーの組み合わせにより実現できる．

例えば，単語辞書検索に利用される「語尾付き単語見出し検索」は，単語見出しによる部分一致検索に，「語尾付き単語」をフィルタキーとして与えることにより実現可能である．

フィルタキーとして利用できるのは，以下の１０種類である

FLT_LANG	0	言語識別（単語 / 対訳 / 共起）
FLT_KIND	1	レコード種別（単語 / 対訳）
FLT_POS	2	品詞（単語 / 対訳）
FLT_CRARC	3	関係子（概念記述）
FLT_CRTF	4	真偽値（概念記述）
FLT_CRTYPE	5	記述区分（概念記述）
FLT_WDHDW	6	語尾付き見出し（単語）
FLT_CCARC	7	共起関係子（共起）
FLT_CCCID1	8	受け側概念識別子（共起）
FLT_CCCID2	9	係り側概念識別子（共起）

なお，フィルタキーは，以下に示すフィルタキー構造体（フィルタキー作成関数により作成される）を，複数ある場合にはリストにして関数に渡す．リストの末端では，nextKeyに NULL を格納する．

```
typedef     struct FilterKey {
        struct FilterKey *nextKey; //次のフィルタキーへのポインタ
        int     filterType;        //フィルタキー種別
        char    *filterKey;        //フィルタキー
} FilterKey;
```

6.2 関数一覧

6.2.1 辞書のオープン / クローズ用関数

DictPtr *DictOpen (char *dictName)

機能： システム辞書をオープンして，辞書構造体を作成する．
引数： dictName システム辞書ファイルのパス名
返値： 正常 ... 辞書構造体の先頭アドレス
異常 ... NULL

void DictClose (DictPtr *dictPtr)

機能： システム辞書をクローズする．
引数： dictPtr 辞書構造体の先頭アドレス
返値： なし

6.2.2 辞書検索関数

EDRecList *EDsearch (DictPtr *dictPtr, unsigned char *keyPtr, int idxNum, int srchMode)

機能： システム辞書を検索する
引数： dictPtr 辞書構造体の先頭アドレス
keyPtr 検索キー
idxNum インデックス番号
srchMode 検索モード
返値： 正常 ... 辞書レコード構造体の先頭アドレス
異常 ... NULL（検索キーを含むレコードが無かった場合を含む）
備考： 概念識別子の場合には，unsigned long の概念識別子へのポインタを，unsigned char にキャストして検索キーとして渡す．

※インデックス番号，検索モードは，以下の数値により指定する．

インデックス番号

IDX_1	0	単語見出し（単語・対訳），概念識別子 1（概念記述・概念体系），受け側単語見出し（共起），概念識別子（概念見出し）
IDX_2	1	読み（単語），概念識別子（対訳），係り側単語見出し（共起），概念識別子 2（概念記述，概念体系）
IDX_3	2	概念識別子（単語）

検索モード

SRCH_FULL	0	完全一致
SRCH_PART	1	部分一致
SRCH_HEAD	2	前方一致

EDRecList *EDmltKeySearch (DictPtr *dictPtr, char *keyPtr1, char *keyPtr2, char *keyPtr3, FilterKey *fltKeyPtr, int srchMode)

機能： システム辞書を複数キーで検索する
引数： dictPtr 辞書構造体の先頭アドレス
keyPtr1~3 検索キー（インデックス #1 ~ #3に対応）
fltKeyPtr フィルタキーリストの先頭アドレス
srchMode 検索モード

返値： 正常 ... 辞書レコード構造体の先頭アドレス
異常 ... NULL（検索キーを含むレコードが無かった場合を含む）

備考： 概念識別子の場合には，unsigned longの概念識別子へのポインタを，unsigned charにキャストして検索キーとして渡す．
指定する必要のない検索キーには，長さ 0 の文字列を渡す．
フィルタキーは，makeFltKey()で作成されたものを，複数ある場合はリストにして渡す．また，フィルタキーの指定が必要ない場合には，NULL を指定する．

6.3.3 フィルタキー用関数

FilterKey *makeFltKey (int filterType, char *filterKey)

機能： フィルタキーを作成する

引数： filterType フィルタキーの種類
filterKey フィルタキー文字列
※真偽値，言語識別など 0/1の値を持つフィールドをフィルタキーとする場合には，"0"，"1" の文字列を渡す

返値： 正常 ... フィルタキーの先頭アドレス
異常 ... NULL

void *freeFltKey (FilterKey *filterKey)

機能： フィルタキー格納領域を解放する

引数： filterKey フィルタキー構造体リストの先頭アドレス

返値： なし

6.3.4 辞書レコードアクセス関数

辞書レコードアクセス関数は，そのほとんどがほぼ同様の方法で利用できる．特に断わらない限り，以下の手順を基本とする

- アクセス関数には，辞書レコード構造体の，レコード本体の先頭アドレスを引数として与える（以下，変数名 recPtrで統一）
- アクセス異常の場合 -1 を返す
- 固定長の情報（概念識別子，各種フラグなど）はレコード本体の先頭アドレスのみを引数として取り，関数の返り値として，データの値を返す．
- 可変長の情報の場合，関数の第 2 引数に，可変長文字列を格納する領域の先頭アドレスを渡し，関数は可変長文字列の長さを返す．ただし，文字列を格納する領域は，呼び出し側で確保する．

概念辞書アクセス関数

unsigned long CIgetCID (unsigned char *recPtr)

機能： 概念見出しレコードの概念識別子を返す

備考： エラー時は 0 を返す

int CIgetECExpl (unsigned char *recPtr, char *strPtr)

機能： 概念見出しレコードの英語概念説明を得る

int CIgetECWord (unsigned char *recPtr, char *strPtr)
機能： 概念見出しレコードの英語概念見出しを得る

int CIgetJCExpl (unsigned char *recPtr, char *strPtr)
機能： 概念見出しレコードの日本語概念説明を得る

int CIgetJCWord (unsigned char *recPtr, char *strPtr)
機能： 概念見出しレコードの日本語概念見出しを得る

unsigned long CDgetCID1 (unsigned char *recPtr)
機能： 概念体系・記述レコードの概念識別子を返す
備考： エラー時は 0 を返す

unsigned long CDgetCID2 (unsigned char *recPtr)
機能： 概念体系・記述レコードの概念識別子を返す
備考： エラー時は 0 を返す

int CDgetArc (unsigned char *recPtr, char *relPtr)
機能： 概念記述レコードの関係子を得る

int CDgetTFVal (unsigned char *recPtr)
機能： 概念記述レコードの真偽値を返す

int CDgetType (unsigned char *recPtr)
機能： 概念記述レコードの記述区分を返す
備考： I 記述の場合 1 を，E記述の場合 0 を返す

単語辞書アクセス関数

unsigned long WDgetCID (unsigned char *recPtr)
機能： 単語辞書レコードの概念識別子を返す
備考： エラー時は 0 を返す

int WDgetHWord (unsigned char *recPtr, char *strPtr)
機能： 単語辞書レコードの単語見出しを得る

int WDgetRead (unsigned char *recPtr, char *strPtr)
機能： 単語辞書レコードの読みを得る

int WDgetLAdj (unsigned char *recPtr, char *strPtr)
機能： 単語辞書レコードの左連接情報を得る

int WDgetRAdj (unsigned char *recPtr, char *strPtr)
機能： 単語辞書レコードの右連接情報を得る

int WDgetLang (unsigned char *recPtr)
機能： 単語辞書レコードの言語識別を返す
備考： 日本語の場合 0 を，英語の場合 1 を返す

int WDgetIdmFlg (unsigned char *recPtr)
機能： 単語辞書レコードの慣用句フラグを返す
備考： 慣用句の場合 1 を，そうでない場合 0 を返す

int WDgetTechFlg (unsigned char *recPtr)
機能： 単語辞書レコードのレコード種別を返す
備考： 専門用語の場合 1 を，そうでない場合 0 を返す

int WDgetPOS (unsigned char *recPtr, char *strPtr)
機能： 単語辞書レコードの品詞を得る

int WDgetGram (unsigned char *recPtr, char *strPtr)
機能： 単語辞書レコードの文法情報を得る

int WDgetPron (unsigned char *recPtr, char *strPtr)
機能： 単語辞書レコードの発音情報 / 構成語情報を得る
備考： 構成語情報 / 発音情報は格納しているフィールドが同じであるため，
アクセスには同じ関数を利用する．

int WDgetSTree (unsigned char *recPtr, char *strPtr)
機能： 単語辞書レコードの構文木を得る

int WDgetWEnd (unsigned char *recPtr, char *wendPtr)
機能： 単語辞書レコードの単語見出し語尾を得る

int WDgetREnd (unsigned char *recPtr, char *wendPtr)
機能： 単語辞書レコードの読み語尾を得る
備考： 日本語単語辞書の場合のみ

対訳辞書アクセス関数

unsigned long BIgetCID (unsigned char *recPtr)
機能： 対訳辞書レコードの概念識別子を返す
備考： エラー時は 0 を返す

int BIgetHWord (unsigned char *recPtr, char *strPtr)
機能： 対訳辞書レコードの単語見出しを得る

int BIgetRead (unsigned char *recPtr, char *strPtr)
機能： 対訳辞書レコードの読みを得る

int BIgetLang (unsigned char *recPtr)
機能： 対訳辞書レコードの言語対識別を返す
備考： 日英の場合 0 を，英日の場合 1 を返す

int BIgetTechFlg (unsigned char *recPtr)
機能： 対訳辞書レコードのレコード種別を返す
備考： 専門用語の場合 1 を，そうでない場合 0 を返す

int BIgetPOS (unsigned char *recPtr, char *strPtr)
機能： 対訳辞書レコードの品詞を得る

int BIgetTrans (unsigned char *recPtr, char *strPtr)
機能： 対訳辞書レコードの対訳情報を得る

共起辞書アクセス関数

int CCgetCCHWord (unsigned char *recPtr, char *strPtr)
機能： 共起辞書レコードの句見出しを得る
備考： 句見出しは，単語見出しと関係単語表現から合成する

int CCgetHWord1 (unsigned char *recPtr, char *strPtr)
機能： 共起辞書レコードの受け側単語見出しを得る

int CCgetHWord2 (unsigned char *recPtr, char *strPtr)
機能： 共起辞書レコードの係り側単語見出しを得る

int CCgetHwInfo1 (unsigned char *recPtr, char *strPtr)
機能： 共起辞書レコードの受け側表記情報を得る

int CCgetHwInfo2 (unsigned char *recPtr, char *strPtr)
機能： 共起辞書レコードの係り側表記情報を得る

int CCgetCCArc (unsigned char *recPtr, char *strPtr)
機能： 共起辞書レコードの共起関係子を得る

int CCgetArcHWord (unsigned char *recPtr, char *strPtr)
機能： 共起辞書レコードの関係単語表記を得る

int CCgetText (unsigned char *recPtr, char *strPtr)
機能： 共起辞書レコードの例文を得る

unsigned long CCgetCID1 (unsigned char *recPtr)
機能： 共起辞書レコードの受け側概念識別子を返す
備考： エラー時は 0 を返す

unsigned long CCgetCID2 (unsigned char *recPtr)
機能： 共起辞書レコードの係り側概念識別子を返す
備考： エラー時は 0 を返す

int CCgetArc (unsigned char *recPtr, char *strPtr)
機能： 共起辞書レコードの概念関係子を得る

int CCgetPOS1 (unsigned char *recPtr, char *strPtr)
機能： 共起辞書レコードの受け側品詞を得る

int CCgetPOS2 (unsigned char *recPtr, char *strPtr)
機能： 共起辞書レコードの係り側品詞を得る

int CCgetLang (unsigned char *recPtr)
機能： 共起辞書レコードの言語識別を返す
備考： 日本語の場合 0 を，英語の場合 1 を返す

int CCgetHwOrder (unsigned char *recPtr)
機能： 共起辞書レコードの表記順を返す
備考： 係り側→受け側の順の時 0 を，逆の場合 1 を返す

int CCgetArcDir (unsigned char *recPtr)
機能： 共起辞書レコードの関係子方向を返す
備考： 受け側 = 始点側の時 0 を，そうでない場合 1 を返す

int CCgetIdmFlg1 (unsigned char *recPtr)
機能： 共起辞書レコードの受け側慣用句フラグを返す

int CCgetIdmFlg2 (unsigned char *recPtr)
機能： 共起辞書レコードの係り側慣用句フラグを返す

int CCgetSfFreq (unsigned char *recPtr)
機能： 共起辞書レコードの表層共起頻度を返す

int CCgetDpFreq (unsigned char *recPtr)
機能： 共起辞書レコードの共起項目頻度を返す

int CCgetHwFreq1 (unsigned char *recPtr)
機能： 共起辞書レコードの受け側共起要素頻度を返す

int CCgetHwFreq2 (unsigned char *recPtr)
機能： 共起辞書レコードの受け側共起要素頻度を返す

第 7 章

ファイルフォーマット

7.1 システム辞書

7.1.1 基本構造

システム辞書とは，本システムが直接参照する辞書であり，以下に示す 6 種類がある．

単語辞書
対訳辞書
共起辞書
概念体系辞書
概念記述辞書
概念見出し辞書

システム辞書は読み出し専用とし，EDBroW の動作によってシステム辞書を書き換えることはない．また，単語 / 対訳 / 共起辞書については，言語識別，基本語 / 専門用語の区別によらず，同一ファイルとすることが可能である*．

各システム辞書は，それぞれ1つのファイルとして作成される．辞書ファイルは，以下の6つのパートにより構成される．

辞書情報パート
可変長辞書パート
固定長辞書パート
インデックスパート #1 ~ #3

インデックスパートは，辞書を検索するためのインデックス部分であり，一つの辞書に対して3つまでのインデックスパートを持つことができる．辞書情報パートには，辞書識別子，各辞書 / インデックスパートのレコードサイズ，先頭オフセットなどが格納される．

図7.1に，例として単語辞書の構造を示す．

* ただし，共起辞書の専門用語辞書，および日本語共起パータン副辞書については，基本語辞書と形式がまったく異なるため，本システムではサポートしない．

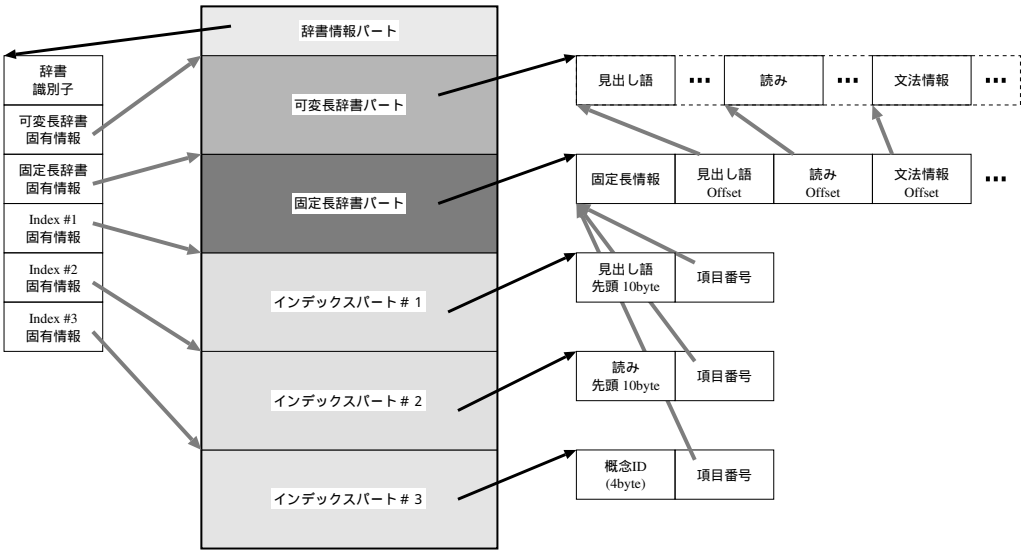


図 7.1 単語辞書の構造

辞書情報パート

フィールド名	格納データ	データ型	データサイズ
辞書固有情報	辞書識別子	char	1byte
	項目数	long	4byte
可変長パート	先頭オフセット	long	4byte
固定長パート	先頭オフセット	long	4byte
固有情報	レコードサイズ	char	1byte
	固定長情報サイズ	char	1byte
インデックス	先頭オフセット	long	4byte
#1 ~ #3	インデックスサイズ	char	1byte
固有情報	インデックスタイプ	char	1byte
	データフラグ	char	1byte
	キー格納位置	char	1byte

表 7.1　辞書情報パートのレコード形式

表 7.1 に，辞書情報パートのレコード形式を示す．

ここで，各インデックス固有情報における，"データフラグ"とは，そのインデックスパートに，辞書情報がそのまま格納されている（固定長辞書のかわりになる）ことを示し，辞書情報が格納されている場合は 1 ，そうでない場合には 0 を格納する* ．また，"インデックスタイプ"とは，格納されているキーが固定長か可変長（の固定長先頭バイト）か，を示すものであり，可変長の場合には 1 を，固定長の場合には 0 を格納する．"キー格納位置" は，固定長辞書中で，当該キー（可変長キーの場合は先頭からのオフセット）が格納されている位置を示す．

* これは，検索速度の向上と，辞書サイズの圧縮を意図したものであり，システム辞書では，特に頻繁に検索される概念見出し辞書と，概念体系辞書がこの形式で作成される．

辞書識別子は，表 7.2 に示すコードを格納する．

システム辞書名	コード	略称
単語辞書	1	WD
概念体系辞書	2	CH
概念記述辞書	3	CR
概念見出し辞書	4	CI
対訳辞書	5	BI
共起辞書	6	CC

表 7.2　辞書識別子

表 7.3に，各システム辞書の各属性を示す．表中，各インデックスパートにおいて，"○ " は可変長キー，"☆" は固定長キーの，それぞれインデックスとなること (インデックスタイプに対応) を示し，"★" はさらに，固定長辞書の役割を果すこと (データフラグに対応) を示す．また，固定長辞書 / インデックスの，記号横の ( ) 内の数字は，そのレコードのデータサイズ (byte) を示す．

辞書名	識別子	可変長パート	固定長パート	Index #1	Index #2	Index #3
単語辞書	WD	○	○(27)	○(14)	○(14)	☆(8)
概念体系辞書	CH	−	−	★(8)	★(8)	−
概念記述辞書	CR	−	○(14)	☆(8)	☆(8)	☆(6)
概念見出し辞書	CI	○	−	★(8)	−	−
対訳辞書	BI	○	○(21)	○(14)	☆(8)	−
共起辞書	CC	○	○(47)	○(14)	○(14)	○(14)

表 7.3 各辞書情報のサイズ

以下に，各辞書ごとに，その詳細な構造を示す．

7.1.2　単語辞書

表 7.4に，単語辞書のレコード形式を示す．

可変長のフィールドについては，可変長辞書パートに格納し，固定長辞書パートには，そのオフセット，すなわち辞書の先頭からの可変長データ格納位置のバイト数を格納する．また，可変長辞書パート中では，同一の文字列が頻繁に格納されることのないよう，辞書トランスレータにおいて圧縮処理を行う．なお，テキスト辞書中のEUC コードは，Shift-JIS コードに変換して格納する*．

可変長フィールドのうち，「発音情報 / 構成語情報」のフィールドは，当該辞書レコードの種別によって，格納される情報が異なる．すなわち，慣用句レコードの場合 (慣用句フラグが1である場合) には，構成語情報が，そうでない場合には，発音情報が格納される．

※ここに示した，可変長フィールドの可変長辞書パートへの格納方法は，以下に示すすべてのシステム辞書に共通である．

フィールド名	データ型	データサイズ	検索キー	備考
単語見出し	文字列	可変長	○	Index #1
読み	文字列	可変長	○	Index #2
発音情報 / 構成語情報	文字列	可変長	×	
概念識別子	unsigned long	4byte	○	Index #3
固定長情報	ビットフィールド	3byte		
左連接情報		8ビット	×	
右連接情報		8ビット	×	
品詞情報		6ビット	△	
言語識別		1ビット	△	
レコード種別		1ビット	△	
慣用句フラグ		1ビット	×	
可変長文法情報	文字列	可変長	×	
構文木	文字列	可変長	×	

表 7.4　単語辞書のレコード形式

固定長情報中の，左連接情報・右連接情報・品詞は，コード数値 (→ 5.8) により格納する．言語識別は，日本語の場合0を，英語の場合1を格納し，慣用句フラグは，当該レコードが慣用句である場合に1を，そうでない場合には0を格納する．

「検索キー」の欄において，"○ " は，システム辞書の検索において検索キーとなることを示す．検索キーとなるフィールドに対しては，これによって辞書を昇順にソートしたインデックスパートを作成し，検索を高速に行えるようにする．また，"× " は検索キーとはならないことを示し，"△ " は，直接の検索キーにはならないが，フィルタキーとして利用可能であることを示す．この表記は，以下に述べる辞書の構造すべてに共通である．

各インデックスパートのレコード形式を 表 7.5 に示す．

インデックス	フィールド名	データ型	データサイズ	備考
#1	単語見出し	文字列	10byte	単語見出しの先頭 10byteを格納
	項目番号	unsigned long	4byte	負でない整数
#2	読み	文字列	10byte	読みの先頭 10byteを格納
	項目番号	unsigned long	4byte	負でない整数
#3	概念識別子	unsigned long	4byte	概念識別子を格納
	項目番号	unsigned long	4byte	負でない整数

表 7.5　単語辞書のインデックスパート レコード形式

単語見出し / 読みは，単語辞書の単語見出し / 読みフィールドの先頭10バイトと等しい．長さが10バイト未満の場合は，'¥0'で終端する．10バイト以上の場合は，終端しない．また，概念識別子は単語辞書の概念識別子と等しく固定長である．

項目番号は，単語辞書の各エントリーに付与した非負の整数である．項目番号に固定長辞書パートのレコードサイズを掛けたものに，固定長辞書パートの先頭オフセットを加えると，そのエントリーのオフセットとなる．

インデックスパートは，キーとなるシステム辞書のフィールドに対して昇順にソートする．すなわち，単語見出し / 読みの10バイトを超える部分に対しても，昇順であることを保証する．ここでいう昇順とは，単純に文字コードの比較によるものであり，テキスト辞書の順番は考慮しなくて良い．また，インデックスパートのレコードは，単語辞書のレコードと1対1に対応するため，それらの総数は等しい．

7.1.3 対訳辞書

表 7.6 に対訳辞書のレコード形式を示す．

フィールド名	データ型	データサイズ	検索キー	備考
単語見出し	文字列	可変長	○	Index #1
概念識別子	unsigned long	固定長	○	Index #2
固定長情報	ビットフィールド	1byte		
品詞		6ビット	△	
レコード種別		1ビット	×	
言語対識別		1ビット	△	
対訳情報	文字列	可変長	×	

表 7.6　対訳辞書のレコード形式

言語対識別とは，日本語→英語の対訳レコードであるか，英語→日本語であるか，を示すものであり．日英の場合は 0 を，英日の場合は 1 を格納する．また，レコード種別は，当該レコードが専門用語であるかどうか，を格納するものであり，値としては 1/0をとる．

訳語情報は，テキスト形式のテキスト辞書に格納されている文字列を，そのまま格納する．

インデックスパートのレコード形式を，表 7.7 に示す．

インデックス	フィールド名	データ型	データサイズ	備考
#1	単語見出し	文字列	10byte	単語見出しの先頭10byteを格納
	項目番号	unsigned long	4byte	負でない整数
#2	概念識別子	unsigned long	4byte	概念識別子を格納
	項目番号	unsigned long	4byte	負でない整数

表 7.7　対訳辞書のインデックスパート レコード形式

7.1.4 共起辞書

表 7.8 に共起辞書のレコード形式を示す．

言語識別は，日本語の場合 0 を，英語の場合 1 を格納する．品詞，関係子については，5.8 に示すコード番号を格納する．

表記順とは，句見出しと受け側 / 係り側単語との順序の対応をとるもので，句見出し中で係り側→受け側の順に並ぶ場合には0 を，逆の場合には 1 を格納する．

関係子方向とは，表層の係り受け関係と，意味情報における関係との対応をとるもので，受け側概念識別子が意味情報中の始点概念となる場合に 0 を，そうでない場合に 1 を格納する．

フィールド名	データ型	データサイズ	検索キー	備考
受け側単語見出し	文字列	可変長	○	Index #1
受け側補助情報	文字列	可変長	×	
係り側単語見出し	文字列	可変長	○	Index #2
係り側補助情報	文字列	可変長	×	
共起関係子	文字列	可変長	△	
関係単語表記	文字列	可変長	×	
受け側概念識別子	unsigned long	4byte	△	
係り側概念識別子	unsigned long	4byte	△	
固定長情報	ビットフィールド	3byte	×	
言語識別		1bit		
係り側品詞		4bit		
受け側品詞		4bit		
概念関係子		8bit		
概念関係子方向		1bit		
表記順		1bit		
受け側慣用句フラグ		1bit		
係り側慣用句フラグ		1bit		
表層共起頻度	int	2byte	×	
共起項目頻度	int	2byte	×	
受け側共起要素頻度	int	2byte	×	
係り側共起要素頻度	int	2byte	×	
例文	文字列	可変長	×	

表 7.8　共起辞書のレコード形式

その他フラグに関しては，コードもしくは 1/0に変換し，ビットフィールドとして格納する．

なお，例文は，テキスト辞書では複数例文が格納されているが，システム辞書では，'|'で区切った形式により，最大 5 例文を格納する．

インデックスパートのレコード形式を，表 7.9 に示す．

インデックス	フィールド名	データ型	データサイズ	備考
#1	受け側単語見出し	文字列	10byte	受け側単語見出しの先頭10byteを格納
	項目番号	unsigned long	4byte	負でない整数
#2	係り側単語見出し	文字列	10byte	係り側単語見出しの先頭10byteを格納
	項目番号	unsigned long	4byte	負でない整数

表 7.9　共起辞書のインデックスパート レコード形式

7.1.5　概念体系辞書

表 7.10 に概念体系辞書のレコード形式を示す．

この形式のレコードは，以下のような概念体系のデータを表現したものである．

概念識別子1(下位概念) − kind-of→ 概念識別子2(上位概念)

フィールド名	データ型	データサイズ	検索キー	備考
概念識別子1	unsigned long	4byte	○	Index #1
概念識別子2	unsigned long	4byte	○	Index #2

表 7.10　概念体系辞書のレコード形式

インデックスパートのレコード形式を，表 7.11に示す．概念体系辞書では，検索の高速化と，辞書ファイルのサイズ縮小のために，インデックスパートが固定長辞書の役割を果たす．

インデックス	フィールド名	データ型	データサイズ	備考
#1	概念識別子1	unsigned long	4byte	概念識別子1を格納
	概念識別子2	unsigned long	4byte	概念識別子2を格納
#2	概念識別子2	unsigned long	4byte	概念識別子2を格納
	概念識別子1	unsigned long	4byte	概念識別子1を格納

表 7.11　概念体系辞書のインデックスパート レコード形式

7.1.6　概念記述辞書

表 7.12 に，概念記述辞書のレコード形式を示す．

フィールド名	データ型	データサイズ	検索キー	備考
概念識別子1	unsigned long	4byte	○	Index #1
概念識別子2	unsigned long	4byte	○	Index #2
概念関係子	char	1byte	△	
真偽値	ビットフィールド	1bit	△	
記述区分	ビットフィールド	1bit	△	

表 7.12　概念記述辞書のレコード形式

この形式のレコードは，以下のような概念記述のタプルを表現したものである．

概念識別子1 (始点概念) − 概念関係子 / 真偽値 → 概念識別子2 (終点概念)

記述区分とは，当該概念記述が，実例としてEDR コーパス中に出現したもの(E記述)であるか，これに対して概念体系上で抽象化を行なったもの(I記述)であるかを示し，E記述の場合 0 を，I記述の場合 1 を格納する．

インデックスパートのレコード形式を，表 7.13に示す．

インデックス	フィールド名	データ型	データサイズ	備考
#1	概念識別子1	unsigned long	4byte	概念識別子1を格納
	項目番号	unsigned long	4byte	負でない整数
#2	概念識別子2	unsigned long	4byte	概念識別子2を格納
	項目番号	unsigned long	4byte	負でない整数

表 7.13　概念記述辞書のインデックスパート レコード形式

7.1.7 概念見出し辞書

表 7.14に，概念見出し辞書のレコード形式を示す．

フィールド名	データ型	データサイズ	検索キー	備考
概念識別子	unsigned long	4byte	○	Index #1
日本語概念見出し	文字列	可変長	×	
英語概念見出し	文字列	可変長	×	
日本語概念説明	文字列	可変長	×	
英語概念説明	文字列	可変長	×	

表 7.14 概念見出し辞書のレコード形式

可変長辞書パートでは，各概念識別子 に対応する，日本語概念見出し，日本語概念説明，英語概念見出し，英語概念説明を，TAB で区切った形式以下の形式で格納し，NULL 文字により終端する．

〈日本語概念見出し〉¥t〈日本語概念説明〉¥t〈英語概念見出し〉¥t〈英語概念説明〉

インデックスパートのレコード形式を表 7.15に示す．概念見出し辞書では，インデクスパートが固定長辞書としての役割も果たす．

インデックス	フィールド名	データ型	データサイズ	備考
#1	概念識別子 オフセット	unsigned long unsigned long	4byte 4byte	概念識別子を格納 可変長辞書パートの オフセットを格納

表 7.15 概念見出し辞書のインデックスパート レコード形式

7.2 ユーザ情報ファイル

7.2.1 基本構造

ユーザ情報ファイルは，レコード・バッファ内のユーザ情報を1レコードとし，これをテキスト形式で格納しているファイルである．

ユーザ情報ファイル一つに格納されるユーザ情報は一つである．また，ユーザ情報ファイルであることを明示するため，ファイルの拡張子は .uif に統一する．

ユーザ情報ファイルの形式を以下に示す．

〈ユーザ情報ファイル〉::=〈キーワード〉¥n〈ユーザ情報レコード〉。。。

〈ユーザ情報レコード〉::=〈レコード固有情報〉¥t〈レコード本体〉¥n

〈レコード固有情報〉::=〈レコードID〉'|'〈レコード識別〉'|'〈参照IDリスト〉'|'〈覚書〉

〈レコード本体〉::=〈単語レコード〉|〈対訳レコード〉|〈共起レコード〉|
〈概念体系レコード〉|〈概念記述レコード〉|〈概念見出しレコード〉

〈レコードID〉::=〈１０進整数〉

〈レコード識別〉::=〈１０進整数〉
〈参照IDリスト〉::=〈参照ID〉|〈参照ID〉;〈参照IDリスト〉
〈参照ID〉::=〈１０進整数〉
〈覚書〉::=〈文字列〉

なお，レコード固有情報において，情報が無い場合には，長さ0の文字列を格納する．ただし，これが許されるのは，〈参照IDリスト〉および〈覚書〉のみである．

7.2.2 レコード本体の構造

レコード本体では，基本的に辞書レコード中の情報をそのまま格納する．ただし，品詞，左右連接情報，概念関係子については，辞書レコード中の数値コードに対応するコード文字列を格納する．また，各辞書レコード本体においては，概念識別子のみを格納するかわりに，以下の形式による意味情報を格納する．

〈意味情報〉::=〈概念見出しデータ〉|?〈概念識別子〉
〈概念見出しデータ〉::=〈概念識別子〉¥t〈日本語概念見出し〉¥t〈英語概念見出し〉¥t
〈日本語概念説明〉¥t〈英語概念説明〉
〈概念識別子〉::=〈１６進整数〉
〈日本語概念見出し〉::=〈文字列〉
〈英語概念見出し〉::=〈文字列〉
〈日本語概念説明〉::=〈文字列〉
〈英語概念説明〉::=〈文字列〉

通常は，意味情報として概念見出しデータを格納するが，概念見出し辞書にエントリのない概念識別子の場合には，"?〈概念識別子〉" の文字列として格納する．
なお，レコードの各項目において，情報が無い場合には長さ0の文字列を格納する．

単語レコード

〈単語レコード〉::=〈単語見出し〉¥t〈読み〉¥t〈発音情報〉¥t〈品詞〉¥t
〈左連接属性〉¥t〈右連接属性〉¥t〈意味情報〉¥t〈言語識別〉¥t
〈慣用句フラグ〉¥t〈レコード種別〉¥t〈文法情報〉¥t〈構文木〉
〈単語見出し〉::=〈文字列〉
〈読み〉::=〈文字列〉
〈発音情報〉::=〈文字列〉
〈品詞〉::=〈文字列〉

〈左連接属性〉 ::= 〈文字列〉
〈右連接属性〉 ::= 〈文字列〉
〈言語識別〉 ::= J | E
〈慣用句フラグ〉 ::= 1 | 0
〈レコード種別〉 ::= 1 | 0
〈文法情報〉 ::= 〈文字列〉
〈構文木〉 ::= 〈文字列〉

対訳レコード

〈対訳レコード〉 ::= 〈単語見出し〉¥t〈読み〉¥t〈品詞〉¥t〈言語対識別〉¥t
〈レコード種別〉¥t〈意味情報〉¥t〈対訳情報〉

〈単語見出し〉 ::= 〈文字列〉
〈読み〉 ::= 〈文字列〉
〈品詞〉 ::= 〈文字列〉
〈言語対識別〉 ::= "JE" | "EJ"
〈レコード種別〉 ::= 1 | 0
〈対訳情報〉 ::= 〈文字列〉

共起レコード

〈共起レコード〉 ::= 〈受け側単語見出し〉¥t〈受け側表記情報〉¥t〈受け側品詞〉¥t
〈共起関係子〉¥t〈関係単語表記〉¥t
〈係り側単語見出し〉¥t〈係り側表記情報〉¥t〈係り側品詞〉¥t
〈受け側意味情報〉¥t〈受け側慣用句フラグ〉¥t〈概念関係子〉¥t
〈関係子方向〉〈係り側意味情報〉¥t〈係り側慣用句フラグ〉¥t
〈言語識別〉¥t〈表層共起頻度〉¥t〈共起項目頻度〉¥t
〈受け側共起要素頻度〉¥t〈係り側共起要素頻度〉¥t〈例文〉

〈受け側単語見出し〉 ::= 〈文字列〉
〈受け側表記情報〉 ::= 〈文字列〉
〈受け側単語品詞〉 ::= 〈文字列〉
〈共起関係子〉 ::= 〈文字列〉
〈関係単語表記〉 ::= 〈文字列〉
〈係り側単語見出し〉 ::= 〈文字列〉
〈係り側表記情報〉 ::= 〈文字列〉
〈係り側単語品詞〉 ::= 〈文字列〉
〈受け側意味情報〉 ::= 〈意味情報〉

〈受け側慣用句フラグ〉 ::= 1 | 0
〈概念関係子〉 ::= 〈文字列〉
〈関係子方向〉 ::= 1 | 0
〈係り側意味情報〉 ::= 〈意味情報〉
〈係り側慣用句フラグ〉 ::= 1 | 0
〈言語識別〉 ::= J | E
〈表層共起頻度〉 ::= 〈１０進整数〉
〈共起項目頻度〉 ::= 〈１０進整数〉
〈受け側共起要素頻度〉 ::= 〈１０進整数〉
〈係り側共起要素頻度〉 ::= 〈１０進整数〉
〈例文〉 ::= 〈文字列〉

概念体系レコード

〈概念体系レコード〉 ::= 〈下位概念〉¥t〈上位概念〉

〈下位概念〉 ::= 〈意味情報〉
〈上位概念〉 ::= 〈意味情報〉

概念記述レコード

〈概念記述レコード〉 ::=
　〈始点概念〉¥t〈終点概念〉¥t〈概念関係子〉¥t〈真偽値〉¥t〈記述区分〉

〈始点概念〉 ::= 〈意味情報〉
〈終点概念〉 ::= 〈意味情報〉
〈関係子〉 ::= 〈文字列〉
〈真偽値〉 ::= 1 | 0
〈記述区分〉 ::= 1 | 0

概念見出しレコード

〈概念見出しレコード〉 ::= 〈意味情報〉