

Oracle® Solaris 11.3 国際化対応言語環境の 利用ガイド

ORACLE®

Part No: E62626
2016年11月

Part No: E62626

Copyright © 2011, 2016, Oracle and/or its affiliates. All rights reserved.

このソフトウェアおよび関連ドキュメントの使用と開示は、ライセンス契約の制約条件に従うものとし、知的財産に関する法律により保護されています。ライセンス契約で明示的に許諾されている場合もしくは法律によって認められている場合を除き、形式、手段に関係なく、いかなる部分も使用、複写、複製、翻訳、放送、修正、ライセンス供与、送信、配布、発表、実行、公開または表示することはできません。このソフトウェアのリバース・エンジニアリング、逆アセンブル、逆コンパイルは互換性のために法律によって規定されている場合を除き、禁止されています。

ここに記載された情報は予告なしに変更される場合があります。また、誤りが無いことの保証はいたしかねます。誤りを見つけた場合は、オラクルまでご連絡ください。

このソフトウェアまたは関連ドキュメントを、米国政府機関もしくは米国政府機関に代わってこのソフトウェアまたは関連ドキュメントをライセンスされた者に提供する場合は、次の通知が適用されます。

U.S. GOVERNMENT END USERS: Oracle programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, delivered to U.S. Government end users are "commercial computer software" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the programs, including any operating system, integrated software, any programs installed on the hardware, and/or documentation, shall be subject to license terms and license restrictions applicable to the programs. No other rights are granted to the U.S. Government.

このソフトウェアまたはハードウェアは様々な情報管理アプリケーションでの一般的な使用のために開発されたものです。このソフトウェアまたはハードウェアは、危険が伴うアプリケーション(人的傷害を発生させる可能性があるアプリケーションを含む)への用途を目的として開発されていません。このソフトウェアまたはハードウェアを危険が伴うアプリケーションで使用する際、安全に使用するために、適切な安全装置、バックアップ、冗長性(redundancy)、その他の対策を講じることは使用者の責任となります。このソフトウェアまたはハードウェアを危険が伴うアプリケーションで使用したことに起因して損害が発生しても、Oracle Corporationおよびその関連会社は一切の責任を負いかねます。

OracleおよびJavaはオラクル およびその関連会社の登録商標です。その他の社名、商品名等は各社の商標または登録商標である場合があります。

Intel, Intel Xeonは、Intel Corporationの商標または登録商標です。すべてのSPARCの商標はライセンスをもとに使用し、SPARC International, Inc.の商標または登録商標です。AMD, Opteron, AMDロゴ、AMD Opteronロゴは、Advanced Micro Devices, Inc.の商標または登録商標です。UNIXは、The Open Groupの登録商標です。

このソフトウェアまたはハードウェア、そしてドキュメントは、第三者のコンテンツ、製品、サービスへのアクセス、あるいはそれらに関する情報を提供することがあります。適用されるお客様とOracle Corporationとの間の契約に別段の定めがある場合を除いて、Oracle Corporationおよびその関連会社は、第三者のコンテンツ、製品、サービスに関して一切の責任を負わず、いかなる保証もいたしません。適用されるお客様とOracle Corporationとの間の契約に定めがある場合を除いて、Oracle Corporationおよびその関連会社は、第三者のコンテンツ、製品、サービスへのアクセスまたは使用によって損失、費用、あるいは損害が発生しても一切の責任を負いかねます。

ドキュメントのアクセシビリティについて

オラクルのアクセシビリティについての詳細情報は、Oracle Accessibility ProgramのWeb サイト(<http://www.oracle.com/pls/topic/lookup?ctx=acc&id=docacc>)を参照してください。

Oracle Supportへのアクセス

サポートをご契約のお客様には、My Oracle Supportを通して電子支援サービスを提供しています。詳細情報は(<http://www.oracle.com/pls/topic/lookup?ctx=acc&id=info>)か、聴覚に障害のあるお客様は (<http://www.oracle.com/pls/topic/lookup?ctx=acc&id=trs>)を参照してください。

目次

このドキュメントの使用法	9
1 概要	11
Oracle Solaris とグローバルマーケット	11
国際化とローカリゼーションの概要	11
ロケールとは	12
C ロケール	13
ロケールカテゴリ	14
コアロケール	14
ロケールの影響を受ける動作	15
時間の書式	16
日付の書式	16
数値の書式	17
通貨の国際書式	18
言語の単語や文字の違い	18
単語の区切り文字	18
ソートの順序	19
文字セット	19
キーボードの違い	22
用紙サイズの違い	22
2 Unicode と UTF-8 ロケールのサポート	23
Unicode の概要	23
UTF-8 の概要	24
共通ロケールデータリポジトリ	25
UTF-8 以外の文字セットを使用するロケール	25
UTF-8 以外のロケールから UTF-8 ロケールへの移行	26
標準テキストファイル	26
ファイル名とディレクトリ名	26

ZFS	27
NFS	27
3 言語とロケールの作業	29
ユーザーのロケールの管理	29
端末セッションでのロケールの設定	29
GNOME ディスプレイマネージャーでのロケール選択	30
端末セッションでの永続的なロケール設定	32
リモート接続経由のロケールネゴシエーション	32
デフォルトのシステムロケールの管理	33
インストーラでのロケール選択	33
自動インストーラでのロケール選択	34
使用可能なロケールの管理	36
nlsadm を使用したロケールの追加または削除	36
ロケールファセット	37
pkg を使用したロケールの追加または削除	38
▼ パッケージマネージャーで言語サポートを追加する方法	39
旧バージョンのロケール	40
4 デスクトップキーボードの設定と入力方式	43
キーボード配列と入力方式の設定について	43
GNOME デスクトップマネージャーでのキーボード配列の選択	44
▼ GNOME デスクトップマネージャーでキーボード配列を選択する方 法	44
入力方式	45
▼ 入力方式のアクティブ化と非アクティブ化の方法	46
IBus について	47
Internet Intranet Input Method Framework	53
GNOME キーボード設定	54
▼ キーボード設定とカスタマイズを設定する方法	54
キーボード配列のオプションとデフォルトの動作	55
GNOME のキーボード配列インジケータ (キーボードセレクタ)	56
X キーボード拡張 (setxkbmap) を使用したキーボード配列の設定	57
コンソールでのキーボード配列	57
5 フォントの構成	59
fontconfig ライブラリ	59
フォント構成の調整	59

使用可能なフォント	60
インド語派フォント	60
日本語フォント	61
韓国語フォント	62
簡体字中国語および繁体字中国語フォント	63
タイ語フォント	64
6 詳細情報	65
コードセット変換	65
iconv ユーティリティー	65
Unicode 用の国際化コンポーネント	66
ファイル検査機能 (fsexam)	68
エンコーディング自動検出 (auto_ef)	68
国際化ドメイン名のサポート	69
GNU IDN ライブラリ	69
JPRS idnkit-2 ライブラリ	70
印刷の拡張	71
mp ユーティリティー	71
ほかのプラットフォームとの相互運用性	72
NFS サーバーの考慮事項	72
ファイルシステムの考慮事項	73
ASCII 以外のファイル名を含むアーカイブ	73
各国語プロパティの構成	74
カスタムロケールの作成	75
システムロケールを基にした新しいロケールの作成	75
ゼロからのロケールの作成	76
A 使用可能なロケールおよびサポートされている文字セット	77
索引	87

このドキュメントの使用方法

- **概要** – Oracle Solaris OS の国際化機能について説明します。
- **対象読者** – Oracle Solaris OS でグローバルアプリケーションを設計およびサポートするソフトウェア開発者およびシステム管理者。
- **必要な知識** – このガイドの読者は、C プログラミング言語の知識を持っているべきです。

製品ドキュメントライブラリ

この製品および関連製品のドキュメントとリソースは <http://www.oracle.com/pls/topic/lookup?ctx=E62101-01> で入手可能です。

フィードバック

このドキュメントに関するフィードバックを <http://www.oracle.com/goto/docfeedback> からお聞かせください。

◆◆◆ 第 1 章

概要

この章では、Oracle Solaris オペレーティングシステムの国際化とローカリゼーションの概念を紹介します。この章の内容は次のとおりです。

- 11 ページの「Oracle Solaris とグローバルマーケット」
- 11 ページの「国際化とローカリゼーションの概要」
- 12 ページの「ロケールとは」
- 15 ページの「ロケールの影響を受ける動作」
- 18 ページの「言語の単語や文字の違い」

Oracle Solaris とグローバルマーケット

Oracle Solaris OS は、Unicode 6.0 などの最新の国際規格に基づく、国際化されたオペレーティングシステムです。世界中のユーザーは、好みの地域で自身の言語を使ってシステムを操作できます。Oracle Solaris には、約 100 の異なる地域の 60 を超える言語、合計約 250 の異なるロケールのサポートが含まれています。約 140 のロケールが Unicode に基づいているほか、ISO8859、EUC、およびその他のコードセットに基づく各種レガシーロケールのサポートも引き続き利用可能となっています。Unicode やレガシーロケールについては、第2章「Unicode と UTF-8 ロケールのサポート」で詳しく説明します。以降のセクションでは、国際化とローカリゼーションの概要を説明し、ロケールの概念について説明します。

国際化とローカリゼーションの概要

国際化とローカリゼーションは異なる手順です。国際化が言語間や地域間でのソフトウェアの移植を可能にするプロセスであるのに対し、ローカリゼーションはソフトウェアを特定の言語や地域に適応させるプロセスです。国際化されたソフトウェアを開発するには、文化上の特定の要件に従ってプログラムの動作を実行時に変更するイ

インタフェースを使用します。ローカリゼーションでは、ロケールとも呼ばれる言語や地域をサポートするためのオンライン情報を確立する必要があります。

国際化されたソフトウェアは、さまざまなネイティブ言語や慣例の下で動作するほか、ソフトウェアを書き直さなくても、あるロケールから別のロケールに移植できます。Oracle Solaris システムは国際化されており、国際化されたソフトウェアの作成に必要なインフラストラクチャーやインタフェースを提供します。

ロケールとは

アプリケーションプログラムの重要な概念は、プログラムのロケールの概念です。ロケールとは、あるネイティブ言語環境の明示的なモデルおよび定義のことです。ロケールの概念は、<http://opengroup.org> からアクセスできる POSIX 規格に明示的に定義されて含まれています。

ロケールは、国に依存する書式設定やその他の仕様が存在するいくつかのカテゴリから構成されます。プログラムのロケールは、そのコードセット、日付や時間の書式設定規則、通貨規則、10進書式設定規則、および照合(ソート)順序を定義します。

ロケール名は、ベースとなる言語、使用する国(地域)、およびコードセットから構成できます。たとえば、ドイツ語は、Deutsch の略語となる `de` ですが、スイスのドイツ語は `de_CH` になります(`CH` は Confederation Helvetica の略語)。この表記規則により、通貨単位の表記など、国ごとの特定の違いの表現が可能となります。Oracle Solaris 11 のデフォルトのロケールコードセットは、Unicode の ASCII 準拠 8 ビットエンコーディング形式である UTF-8 です。したがって、スイスのドイツ語の完全定義ロケール名は、`de_CH.UTF-8` となります。

ある特定の言語に複数のロケールを関連付けることができるため、地域の違いの表現が可能となります。たとえば、英語を話す米国ユーザーが `en_US.UTF-8` ロケール(米国向けの英語)を選択できる一方、英語を話す英国ユーザーは `en_GB.UTF-8` (英国向けの英語)を選択できます。

一般に、ロケール名は `LANG` 環境変数で指定します。ロケールカテゴリは `LANG` に従属しますが、個別に設定することも可能であり、そうした場合にはそれらの設定が `LANG` をオーバーライドします。`LC_ALL` 環境変数を設定すると、その設定が、`LANG` やすべての個別のロケールカテゴリをオーバーライドします。

ロケールの命名規則は次のとおりです。

```
language[_territory][.codeset][@modifier]
```

ここで、2文字の `language` コードは ISO 639 からのもの、2文字の `territory` コードは ISO 3166 からのもの、`codeset` はこのロケールで使用されるコードセットの名前、`modifier` は、このロケールと修飾子なしのロケールとを区別するための特性の名前です。

すべての Oracle Solaris 製品のロケールは、US-ASCII コード値を持つポータブル文字セットの文字を保持します。

ポータブル文字セットの詳細については、*X/Open CAE Specification: System Interface Definitions, Issue 5* (ISBN 1-85912-186-1) を参照してください。

単一のロケールが複数のロケール名を持つことがあります。たとえば、POSIX は C と同じロケールです。

c ロケール

C ロケールは POSIX ロケールとも呼ばれますが、これは、POSIX に準拠するすべてのシステムのデフォルトのシステムロケールです。Oracle Solaris オペレーティングシステムは POSIX システムです。C ロケールは『Single UNIX Specification, Version 3』で規定されています。<http://www.unix.org/version3/online.html> で登録すれば、この仕様を読んだりダウンロードしたりできます。

国際化されたプログラムが C ロケールで実行されるように指定する方法には、次の 2 つがあります。

- すべてのロケール環境変数の設定を解除します。C ロケールでアプリケーションを実行します。

```
$ unset LC_ALL LANG LC_CTYPE LC_COLLATE LC_NUMERIC LC_TIME LC_MONETARY LC_MESSAGES
```

- ロケールを明示的に C または POSIX に設定します。

```
$ export LC_ALL=C
$ export LANG=C
```

一部のアプリケーションは、[setlocale\(3C\)](#) を実際に呼び出して現在のロケールを参照することなしに、LANG 環境変数をチェックします。この場合、シェルを明示的に C ロケールに設定するには、LC_ALL および LANG ロケール環境変数を指定します。ロケール環境変数間の優先順位の関係については、[setlocale\(3C\)](#) のマニュアルページを参照してください。

ターミナル環境で現在のロケール設定を確認するには、[locale\(1\)](#) コマンドを実行します。

```
$ locale
LANG=C
LC_CTYPE="C"
LC_NUMERIC="C"
LC_TIME="C"
LC_COLLATE="C"
LC_MONETARY="C"
LC_MESSAGES="C"
LC_ALL=
```

ロケールカテゴリ

ロケールカテゴリの種類は次のとおりです。

LC_CTYPE	文字の分類および大文字/小文字の変換。
LC_TIME	月の名前、曜日、一般的な完全表示や短縮表示など、日付や時間の書式を指定します。
LC_MONETARY	ロケールの通貨記号、千の区切り文字、符号の位置、小数点以下の桁数など、通貨の書式を指定します。
LC_NUMERIC	小数区切り文字 (または基数文字)、千の区切り文字、およびグループ化を指定します。
LC_COLLATE	ロケールの照合順序および正規表現の定義を指定します。
LC_MESSAGES	ロケールのローカライズメッセージの記述言語、および肯定と否定の応答 (yes と no の文字列と表現) を指定します。
LO_LTYPE	言語レンダリングに関する情報を提供するレイアウトエンジンを指定します。言語レンダリング (またはテキストレンダリング) は、文字の形状や方向の属性に依存します。

コアロケール

次の表に、Oracle Solaris 11 のコアロケールの一覧を示します。

表 1 言語とコアロケール

言語	コアロケール
中国語 - 簡体字	zh_CN.UTF-8
中国語 - 繁体字	zh_TW.UTF-8
英語	en_US.UTF-8
フランス語	fr_FR.UTF-8
ドイツ語	de_DE.UTF-8
イタリア語	it_IT.UTF-8
日本語	ja_JP.UTF-8
韓国語	ko_KR.UTF-8
ポルトガル語 - ブラジル	pt_BR.UTF-8

言語	コアロケール
スペイン語	es_ES.UTF-8

コアロケールは、追加インストールが可能なロケールよりも、ローカライズメッセージのレベルでより広い範囲をカバーします。インストーラやパッケージマネージャーなどの Oracle Solaris OS コンポーネントはコアロケールでのみローカライズされていますが、GNOME や Firefox といったサードパーティソフトウェアのローカライズメッセージは通常、より多くのロケールで使用可能となっています。

Oracle Solaris 環境のどのロケールでも、関連する言語やアプリケーションのローカライズメッセージが存在していれば、そのローカライズメッセージを表示できます。pkg のファセットプロパティを変更すれば、使用可能なすべてのローカライズメッセージを含む追加ロケールを、インストーリリポジトリからシステムに追加できます。詳細は、[36 ページの「使用可能なロケールの管理」](#)を参照してください。

ロケールの影響を受ける動作

数値を書式設定したり、日付や時間を記述したり、単語や語句を区切ったり、書かれたり話されたりした資料を引用したりする際に使用される表記規則は通常、文化ごとに異なります。ロケールは、次の処理、ファイル、書式、および表現をさまざまな地域で処理する方法を決定します。

- テキストデータのエンコードや処理
- リソースファイルの言語識別やエンコード
- テキスト文字列のレンダリングやレイアウト
- クライアント間でのテキストの交換
- 選択された文字のコードセットやテキスト処理の要件に合う入力方式の選択
- 文化的に固有のフォントやアイコンファイル
- ユーザーインターフェース定義 (UID) ファイル
- 日付や時間の書式
- 数値の書式
- 通貨の書式
- 照合順序
- 正規表現の処理
- 情報メッセージや診断メッセージおよび対話式応答の書式

Oracle Solaris 環境は、言語や文化に依存する情報をアプリケーションから分離し、その情報をアプリケーションの外側に保存します。この方法を使えば、アプリケーションの変換、書き換え、または再コンパイルをマーケットごとに行う必要がなくなります。新しいマーケットに参入するのに必要なのは、外部情報をローカルの言語や慣例にローカライズすることだけです。

以降のセクションでは、ロケールカテゴリごとに存在する言語間の違いや、その他の言語間の違いについて説明します。

時間の書式

次の表に、さまざまなロケールでの 11:59 P.M の出力方法をいくつか示します。現在のロケールでの時間の書式を表示するには、次のコマンドを発行します。

```
$ date +%X
```

表 2 時間の国際書式

ロケール	説明	時間の書式
C	-	23:59:00
en_US.UTF-8	英語、米国	11:59:00 PM
es_US.UTF-8	スペイン語、米国	11:59:00 p.m.
mr_IN.UTF-8	マラーティー語、インド	11-59-00 pm
sq_AL.ISO8859-2	アルバニア語、アルバニア	11.59.00.MD
ja_JP.UTF-8	日本語、日本	23時59分00秒
ko_KR.UTF-8	韓国語、韓国	오후11시 59분 00초
zh_CN.UTF-8	簡体字中国語、中国	23时59分00秒

時間は 12 時間形式、24 時間形式のどちらでも表せます。時間と分の区切り文字は、コロン (:)、ピリオド (.)、ダッシュ (-) のいずれかになります。

タイムゾーンの分割は、国と国との間や国内で発生します。タイムゾーンは、協定世界時 UTC (またはグリニッジ標準時 GMT) から何時間進んでいるか、あるいは何時間遅れているかによって記述できますが、その数値は常に整数になるとは限りません。たとえば、ニューファンドランド島が含まれるタイムゾーンと隣接するタイムゾーンとの時差は、0.5 時間です。

夏時間 (DST) の開始と終了の日付は、国ごとに異なる可能性があります。多くの国では DST をまったく実施しません。さらに、1 つのタイムゾーン内で夏時間が異なる場合があります。たとえば米国では、実施の決定権は州にあります。

日付の書式

次の表に、世界中で使用されている日付の書式をいくつか示します。1 つの国のなかでも違いが存在する場合があります。現在のロケールでの日付の書式を表示するには、次のコマンドを発行します。

```
$ date +%x
```


表 3 日付の国際書式

ロケール	説明	日付の書式
C	-	07/16/11
en_CA.UTF-8	英語、カナダ	7/11/16
en_GB.UTF-8	英語、英国	7/16/11
fi_FI.UTF-8	フィンランド語、フィンランド	7/16/11
ja_JP.UTF-8	日本語、日本	2011年07月16日
ko_KR.UTF-8	韓国語、韓国	2011년07월16일
zh_TW.UTF-8	繁体字中国語、台湾	11年07月16日

数値の書式

ロケールによって指定される数値書式にはさまざまなものが存在しています。たとえば、英国と米国では、ピリオドを使って小数位を示します。ほかの多くの国では、代わりにコンマを使用します。小数区切り文字は基数文字とも呼ばれます。同様に、英国と米国ではコンマを使って千のグループを区切りますが、ほかの多くの国では代わりにピリオドを使用しますし、一部の国では千のグループを細い空白 (Unicode 文字 U+2009) で区切ります。

ロケール固有の書式を含むデータファイルを異なるロケールのシステムに転送した場合、その内容の解釈ミスが頻繁に発生します。たとえば、フランス語の書式の数値を含むファイルは、英国に固有のプログラムでは役に立ちません。

次の表に、一般的に使用される数値書式をいくつか示します。現在のロケールでの数値区切り文字の情報を取得するには、次のコマンドを発行します。

```
$ locale -ck LC_NUMERIC
```

表 4 数値の国際表記規則

ロケール	説明	数値の書式
C	-	4294967.00
ar_SA.UTF-8	アラビア語、サウジアラビア	4967967,00
cs_CZ.UTF-8	チェコ語、チェコ共和国	4 294 967,00
de_DE.UTF-8	ドイツ語、ドイツ	4.294.967,00
de_CH.UTF-8	ドイツ語、スイス	4'294'967.00
en_US.UTF-8	英語、米国	4,294,967.00
hi_IN.UTF-8	ヒンズー語、インド	42,94,967.00

注記 - リスト内の数値を区切る方法を指定するロケール表記規則は、特に存在していません。

通貨の国際書式

通貨の通貨単位、表示順序、およびローカル記号や国際記号は、世界中で大きく異なります。現在のロケールの通貨書式を取得するには、次のコマンドを発行します。

```
$ locale -ck LC_MONETARY
```

次の表に、いくつかの国の通貨書式を示します。

表 5 通貨の国際表記規則

ロケール	通貨	例
C	-	1234.56
da_DK.UTF-8	デンマーククローネ (kr)	1.234,56kr
da_DK.ISO8859-15@euro	ユーロ (EUR)	1.234,56€
en_GB.UTF-8	ポンド (£)	£1,234.56
en_US.UTF-8	米国ドル (\$)	\$1,234.56
fr_FR.UTF-8	ユーロ (EUR)	1 234,56€
ja_JP.UTF-8	日本円 (¥)	¥1,235
th_TH.UTF-8	タイ・バーツ (THB)	#1,234.56
zh_CN.UTF-8	中国元 (¥)	¥1,234.56

ユーロ通貨はすべての UTF-8 ロケールでサポートされています。また、ISO8859-15 コードセットに基づくレガシーロケールも使用可能となっており、ユーロを自国通貨として採用しなかった国向けに、(da_DK.ISO8859-15@euro のように)@euro バリアント付きで存在しています。

言語の単語や文字の違い

このセクションでは、言語間の重要な違いについて説明します。

単語の区切り文字

英語では通常、単語は空白文字で区切られます。一方、中国語、日本語、タイ語などの言語では通常、単語間の区切り文字は一切使用されません。

ソートの順序

特定の文字のソート順は、すべての言語で同じとは限りません。たとえば、文字「ö」は、ドイツでは通常の「o」と一緒にソートされますが、スウェーデンでは別個にソートされ、アルファベットの最後の文字となります。一部の言語では、文字シーケンスの優先順位を決定できるように、各文字がウェイトを持ちます。たとえば、タイ語の辞書では、さまざまなウェイトを持つ文字のシーケンスのソートが定義されています。

文字セット

各文字セットに含まれるアルファベット文字や特殊文字の数は、それぞれ異なります。英語のアルファベットには 26 文字しか含まれていませんが、一部の言語にはより多くの文字が含まれています。たとえば、日本語には 20,000 個を超える文字が含まれ、中国語にはさらに多くの文字が含まれる可能性があります。

西ヨーロッパのアルファベット

西ヨーロッパのほとんどの国のアルファベットは、英語を話す各国で使用されている 26 文字から成る標準アルファベットに似ています。これらのアルファベットにはしばしば、若干の追加基本文字、記号またはアクセント付きの文字、および合字も含まれます。

日本語のテキスト

日本語のテキストは、互いに混在する異なる 3 つの文字から構成されます。

- 中国語から派生した表意文字の漢字
- ひらがなとカタカナという 2 つの表音文字 (または音節文字)

ひらがなの各文字に対応する文字がカタカナにもありますが、ひらがながもっとも一般的な文字であり、角張った字体ではなく丸みを帯びた字体を備えています。漢字は、根源的な単語を書くときに使用されます。カタカナは通常、「外来」語、つまり日本語以外の言語から輸入された単語を表現するために使用されます。

漢字には数万個の文字が含まれますが、一般的に使用される文字の数は、年を経るごとに着実に減少しています。現在では、頻繁に使用される文字は約 3500 個のみですが、平均的な日本人執筆者は、約 2000 個の漢字から成る語彙を持ちます。それでも、コンピュータシステムは日本工業規格 (JIS) の要件に従って、7000 個を超える文字をサポートする必要があります。さらに、ひらがなとカタカナの文字が約 170 個存在し

ています。平均して、日本語テキストの 55% がひらがな、35% が漢字、10% がカタカナとなっています。また、日本語テキストには、アラビア数字やローマ字も存在しています。

漢字の使用を完全に避けることも可能ですが、ほとんどの日本人読者は、漢字以外の文字だけで構成されたテキストを、理解しづらいと感じます。

韓国語のテキスト

韓国語のテキストは、ハングルと呼ばれる表音筆記体系を使って書くことができます。ハングルには 11,000 個を超える文字が含まれていますが、それらの文字は、**jamos** と呼ばれる子音字と母音字から構成されます。韓国語のコンピュータシステムでは通常、ハングル文字の語彙全体のうちの約 3000 文字が使用されます。韓国語では、中国で発明されたセットに基づく、ハンジャと呼ばれる表意文字も使用されます。韓国語のテキストでは、6000 個を超えるハンジャ文字が必要となります。ハンジャは通常、ハングルではあいまいになる場合に、混乱を避けるために使用されます。ハングル文字は、子音字と母音字を組み合わせることで形成されます。これらの文字を組み合わせたものは、1つの音節を構成できるようになり、それが、1つのハングル文字となります。通常、ハングル文字は四角の形に構成されるため、その集合体はハンジャ文字と同じスペースを占有します。また、韓国語テキストには、アラビア数字、ローマ字、および特殊記号文字も存在しています。

タイ語のテキスト

タイ語の文字は、ディスプレイ画面上の、4つのディスプレイセルを備えた1つのコラム位置として定義できます。各コラム位置には最大3つの文字を含めることができます。1つのディスプレイセルの構成は、タイ文字の分類に基づきます。タイ語の文字の中には、別の文字の分類と合成できるものもあります。両方の文字が互いに合成可能な場合は、両方の文字が同じセル内に存在します。それ以外の場合、それらは異なるセル内に存在します。

中国語のテキスト

中国語は通常、漢字と呼ばれる表意文字に含まれる文字だけで構成されます。

- 中華人民共和国の場合、GB2312 (zh_CN.EUC ロケール) に約 7000 個の一般的に使用される漢字が、GBK 文字セット (zh_CN.GBK ロケール) に 20,000 個を超える文字が、そして Unicode 6.0 で規定されたすべての CJK 統合漢字拡張 A 文字を含む GB18030-2000 文字セット (zh_CN.GB18030 ロケール) に約 30,000 個の文字が、それぞれ存在しています。

- 台湾の場合、もっとも頻繁に使用される文字セットは、CNS11643-1992 (zh_TW.EUC ロケール) と Big5 (zh_TW.BIG5 ロケール) です。これらは約 13,000 個の漢字を共有しています。
- 香港では、Big5 文字セットに 4702 個の文字が追加され、Big5-HKSCS 文字セット (zh_HK.BIG5HK ロケール) となりました。

ある文字が根源的な文字でない場合、その文字は通常 2 つ以上の部分から構成されますが、2 つの場合がもっとも一般的です。2 つの部分から成る文字では、一方の部分が一般に意味を表し、他方が発音を表します。時には両方の部分が意味を表す場合もあります。もっとも重要な要素は部首であり、文字は従来部首に基づいて整理されますが、その部首は数百個存在しています。単一の音が多数の異なる文字によって表現される可能性があります。それらの文字は、使用上交換可能ではありません。単一の文字が複数の音を持つこともあります。

ある特定の文脈では、一部の文字がほかの文字よりも適しています。適切な文字は、声調を使用して音声上区別されます。これに対し、話し言葉の日本語や韓国語には、声調はありません。

中国語を表す表音体系はいくつかあります。中華人民共和国でもっとも一般的なのは、ローマ字を使用するピンインであり、Beijing のような地名として西洋で広く採用されています。ウェードギル体系は、Peking のような地名として以前使用されていた古い表音体系です。台湾では通常、独自の字体を持つ表音アルファベットの一種である *zhuyin* (または *bopomofo*) が代わりに使用されます。

ヘブライ語のテキスト

ヘブライ語のテキストは、ヘブライ語やイディッシュ語の文字を書くときに使用されます。ヘブライ語では双方向の文字が使用されます。ヘブライ語では、文字は右から左に読み書きしますが、数字は左から右に読み取ります。ヘブライ語テキストに埋め込まれた英語テキストもすべて、左から右に読み取ります。

ヘブライ語では、27 文字のアルファベットが使用され、句読記号や数字については標準ラテン語 (または英語) のアルファベットから流用されます。ヘブライ語テキストには母音記号や発音記号も含まれます。これらの記号は、基本文字の内側のドット (*dagesh*)、文字の下の母音記号、文字の左上のアクセントのいずれかとして現れます。これらの記号は一般に典礼テキストで使用されるだけであり、日常の用途で見かけることはほとんどありません。ヘブライ語には大文字は一切含まれません。

ヒンズー語のテキスト

ヒンズー語のテキストは、デーバナーガリーと呼ばれる文字で書かれます。ヒンズー語は表音言語であり、一連の音節として書かれます。各音節は、子音字、独立母音字、従属母音記号という、アルファベット形式の 3 種類の部分 (デーバナーガリー文

字)から構築されます。音節自体は子音字と母音字のコアから構成され、オプションでその先頭に子音字が付けられます。デーバナーガリー文字は、基底線から始まる英語と異なり、文字の最上部に書かれる水平線(ヘッドストローク)から垂れ下がります。これらの文字は、文脈に応じて結合したり、形が変わったりする可能性があります。ヒンズー語のテキストではヘブライ語の場合と同じく、大文字と小文字の区別は一切ありません。

キーボードの違い

U.S. キーボード上にあるすべての文字が、ほかのキーボードにもあるとは限りません。同様に、ほかのキーボードには通常、U.S. キーボードには見られない文字が多数含まれています。

入力は Oracle Solaris オペレーティングシステムによって処理されるため、どのキーボードを使用しても任意のロケールの文字を入力できます。

用紙サイズの違い

各国内では、少数の用紙サイズが一般的に使用されています。通常、それらのサイズの1つが、その他のサイズよりもずっと一般的に使用されています。ほとんどの国は、ISO 規格 216:「筆記用紙及び各種印刷物—仕上寸法—A及びBシリーズ」に従います。

国際化されたアプリケーションでは、使用可能なページサイズに関する想定を行わないようにしてください。Oracle Solaris システムは、出力ページサイズを追跡するためのサポートは一切提供しません。この追跡を行うのは、アプリケーションプログラムの責任になります。次の表に、一般的な国際ページサイズを示します。

表 6 一般的な国際ページサイズ

用紙の種類	寸法	国
ISO A4	21.0 cm x 29.7 cm	米国以外のすべて
ISO A5	14.8 cm x 21.0 cm	米国以外のすべて
JIS B4	25.9 cm x 36.65 cm	日本
JIS B5	18.36 cm x 25.9 cm	日本
米国レター	8.5 インチ x 11 インチ	米国とカナダ
米国リーガル	8.5 インチ x 14 インチ	米国とカナダ

Unicode と UTF-8 ロケールのサポート

コンピュータ内のテキスト文字列は、文字コードの連続として表現されます。文字セットは文字と文字コードの間のマッピングです。文字列のエンコーディングに使用される文字セットは、ロケールのもっとも重要な特性の 1 つです。この章では、Oracle Solaris のロケールで使用される文字セットと、その作成に使用されるデータソースについて説明します。内容は次のとおりです。

- 23 ページの「Unicode の概要」
- 25 ページの「共通ロケールデータリポジトリ」
- 25 ページの「UTF-8 以外の文字セットを使用するロケール」
- 26 ページの「UTF-8 以外のロケールから UTF-8 ロケールへの移行」

誤った文字セットを使用して文字列を表示すると、多くの場合、出力が壊れる結果になります。したがって、別の文字セットを使用するために文字列を変換することが必要な場合があります。そのような変換については 26 ページの「UTF-8 以外のロケールから UTF-8 ロケールへの移行」で説明します。

Unicode の概要

Unicode は、コンピュータ処理のためのテキストの表現に使用される汎用の文字エンコーディング規格です。Unicode は、多言語テキストをエンコードする一貫した方法を提供し、国際化対応テキストファイルの交換を容易にします。

多言語テキストをコード化するための規格は *ISO/IEC 10646* です。ISO/IEC 10646 および Unicode 規格には、まったく同じ文字およびエンコーディングポイントが含まれていますが、Unicode 規格は文字とその用法についての追加情報を提供します。

Oracle Solaris 11 は、Unicode 規格バージョン 6.0 および ISO/IEC 10646:2011 をシステムレベルでサポートします。

個々の Unicode 文字は、0 - 1,114,111 の範囲の整数であるコードポイントにマップされます。Unicode コードポイントは、U+nnnn 形式 (nnnn はコードポイントの 16 進数) の

表記法を使用して、またはコードポイントを記述するテキスト文字列によって参照されます。たとえば、小文字の「a」は U+0061 またはテキスト文字列「LATIN SMALL LETTER A」によって表すことができます。

コードポイントはさまざまな文字エンコーディングスキームを使用してエンコードできます。Oracle Solaris の Unicode ロケールでは、UTF-8 形式が使用されます。UTF-8 は Unicode の可変長エンコーディング形式であり、ASCII 文字コード値を透過的に保持します (24 ページの「UTF-8 の概要」を参照)。

Unicode 規格と ISO/IEC 10646、およびこれらの規格の各種表現形式については、次のソースを参照してください。

- [Unicode 規格、バージョン 6.0 \(Unicode Consortium\)](#)
- ISO/IEC 10646:2011, Information Technology-Universal Multiple-Octet Character Set (UCS) - Part 1: Architecture and Basic Multilingual Plane
- [Unicode Consortium Web サイト](#)

UTF-8 の概要

UTF-8 は Unicode の可変長エンコーディング形式です。この形式は Oracle Solaris の Unicode ロケールで使用されます。

この形式の利点は、ASCII エンコーディングスキームと下位互換であり、エンディアンとバイト順序の混乱を避けられることです。Unicode コードポイントは、1-4 個の 8 ビットのバイトによって表される UTF-8 です。次の表は、UTF-8 のビット分布を規定し、1 バイト、2 バイト、3 バイト、および 4 バイトのシーケンスに対応する Unicode コードポイントの範囲を示します。

表 7 UTF-8 のビット分布

コードポイントの範囲	コードポイント (2 進)	第 1 バイト	第 2 バイト	第 3 バイト	第 4 バイト
U+0000..U+007F	0xxxxxxx	0xxxxxxx			
U+0080..U+07FF	0000yyyy yyxxxxxx	110yyyyy	10xxxxxx		
U+0800..U+FFFF	zzzzyyyy yyxxxxxx	1110zzzz	10yyyyyy	10xxxxxx	
U+10000..U+10FFFF	000uuuuu zzzzyyyy yyxxxxxx	11110uuu	10uuzzzz	10yyyyyy	10xxxxxx

UTF-8 エンコーディング形式の詳細は、次のソースを参照してください。

- The Unicode Standard, Version 6.0, Chapter 3 (<http://www.unicode.org/versions/Unicode6.0.0/ch03.pdf>), Section 3.9 “Unicode Encoding Forms”, pp. 93 - 94
- [Unicode Consortium Web サイト](#)

共通ロケールデータリポジトリ

共通ロケールデータリポジトリ (CLDR) プロジェクトは、広範なロケールデータを提供することを目的とした Unicode Consortium のプロジェクトです。CLDR には、通常はオペレーティングシステムがアプリケーションに提供するロケール固有情報が格納されます。この情報には、数値形式、日付、時間、および通貨の文字列、文字分類 (小文字、大文字、印刷可能など)、文字列照合規則などがあります。詳細は、[15 ページの「ロケールの影響を受ける動作」](#)を参照してください。

UTF-8 以外の形式の ja_JP、ko_KR、th_TH、zh_CN、zh_HK、および zh_TW ロケールと、ja_JP.UTF-8 ロケールを除いたすべての Oracle Solaris 11 ロケールデータは、CLDR データバージョン 1.9.0 に合わせて配置されます。これらのロケールはほかのソースに基づいており、CLDR データを使用しません。日本語の UTF-8 に対する CLDR ベースのバリエーションは、ja_JP.UTF-8@cldr という名前で使用できます。

CLDR の詳細は、[Unicode 共通ロケールデータリポジトリ](#) プロジェクトサイトを参照してください。

UTF-8 以外の文字セットを使用するロケール

変換の問題を避けるために、Oracle Solaris のロケールは、UTF-8 Overview で説明する Unicode 文字セットの [24 ページの「UTF-8 の概要」](#) エンコーディング形式を使用します。すべてのサポート対象言語には、推奨およびサポートされる形式として UTF-8 ロケールが存在します。

歴史的、技術的、および法的な理由から、UTF-8 以外のロケールも Oracle Solaris で使用できます。C ロケールは、EMEA 言語用の旧バージョンのシングルバイト (8 ビット) ISO ロケールであり、APAC 言語用の伝統的なロケールです。

1 つの文字を表すために 1 バイト (8 ビット) しか使用しないという理由で、過去にはシングルバイト文字セットが広く利用されていました。しかし、セットのサイズが最大 256 文字と限られていることが原因で、言語ごとに異なる文字セットを使用する必要がありました。このことは、ある文字セットで作成されたファイルがしばしば別の文字セットで読み取れない、多言語ドキュメントの表現が困難である、多くの言語にはシングルバイトで表現しきれない数の文字が存在する、などの多くの問題を引き起こします。中国語などのそうした言語用に、各種の伝統的なマルチバイト文字セットが作成されました。

旧バージョンのロケールまたは従来型ロケールとも呼ばれる UTF-8 以外のロケールは、Oracle Solaris 11 で限定的にサポートされます。これらの限定的にサポートされるロケールは、GDM のログインダイアログで使用できず、デフォルトではインストー

ルされません。UTF-8 ロケールに対して存在するローカリゼーションが、UTF-8 以外のロケールの変種では提供されない場合があります。限定的にサポートされるロケールの一部は、将来の Oracle Solaris リリースで削除される可能性があります。

旧バージョンのロケールは Oracle Solaris のインストーラによってインストールされません。これらのロケールを有効にするには、次の例のように `system/locale/extra` パッケージを手動でインストールする必要があります。

```
# pkg install system/locale/extra
```

ロケールファセットも正しく設定する必要があります。詳細は、[37 ページの「ロケールファセット」](#)を参照してください。

UTF-8 以外のロケールから UTF-8 ロケールへの移行

UTF-8 に移行するとき、データのインポートまたはエクスポートに使用する方法はファイルの種類によって異なります。

標準テキストファイル

標準テキストファイルには、ファイルの文字エンコードの明示的な識別はありません。ファイルが UTF-8 エンコーディングでない場合は、変換が必要です。たとえば、繁体字中国語 `big5` でエンコードされた標準テキストファイルを UTF-8 に変換するには、次のコマンドを実行します。

```
$ iconv -f big5 -t UTF-8 inputfilename > outputfilename
```

テキストエディタアプリケーションは、文字エンコーディングテキストの読み取りと書き込みを自動的に実行できます。あるいは、ファイルを開くとき、または保存するときにエンコーディングを明示的に指定できます。テキストエディタを起動するには、「起動」->「アプリケーション」->「アクセサリ」->「テキストエディタ」の順に選択します。

ファイル名とディレクトリ名

UFS や ZFS などのファイルシステムは、ユーザーが使用する文字セットでファイル名およびディレクトリ名を格納します。UTF-8 以外のロケールを使用するか UTF-8 以外のファイルシステムをマウントし、UTF-8 ロケールに移動すると、ファイル名の文字が正しく表示されない場合があります。この問題を修正するために、`convmv(1)` を使

用して、1つのファイル名、ディレクトリツリーと配下のファイル、またはファイルシステム全体を別のエンコーディングに変換できます。変換されるのはファイル名だけで、ファイルの内容は変換されません。

詳細は、`convmv(1)` のマニュアルページを参照してください。ツールは任意のファイルシステムで機能します。

また、`fsexam(1)` をこの目的に使用することもできます。詳細は、[68 ページの「ファイル検査機能 \(fsexam\)」](#) を参照してください。

ZFS

ZFS は Oracle Solaris 11 で使用されるメインのファイルシステムです。ZFS ではほかのファイルシステムと同じように、ロケールの文字セットを使用してファイル名およびディレクトリ名を格納します。UTF-8 文字セットの場合、`normalization` プロパティは、1つのディレクトリ内に同じファイル名の項目が複数存在することを避けるための名前比較にファイルシステムが使用する正規化アルゴリズムの種類を設定します。

`utf8only` プロパティが有効な場合、ファイルシステムは、UTF-8 文字セットに存在しない文字を含むファイル名を拒絶します。

詳細は、`zfs(1)` のマニュアルページを参照してください。

NFS

詳細は、[72 ページの「ほかのプラットフォームとの相互運用性」](#) を参照してください。

◆◆◆ 第 3 章

言語とロケールの作業

この章では、言語を選択、インストール、および使用方法と、Oracle Solaris 11 システムのロケールについて説明します。内容は次のとおりです。

- 29 ページの「ユーザーのロケールの管理」
- 33 ページの「デフォルトのシステムロケールの管理」
- 36 ページの「使用可能なロケールの管理」

ユーザーのロケールの管理

このセクションでは、GNOME ディスプレイマネージャーでロケールを設定する方法、およびロケールを選択する方法について説明します。

端末セッションでのロケールの設定

次のように LANG 変数を設定することにより、端末セッションでロケールを変更できます。

```
$ export LANG=locale
```

たとえば、de_DE.UTF-8 ロケールに変更するには、次のコマンドを使用します。

```
$ export LANG=de_DE.UTF-8
```

ロケールが正常に変更されたことを確認するには、locale コマンドを実行します。

```
$ locale
LANG=de_DE.UTF-8
LC_CTYPE="de_DE.UTF-8"
LC_NUMERIC="de_DE.UTF-8"
LC_TIME="de_DE.UTF-8"
LC_COLLATE="de_DE.UTF-8"
LC_MONETARY="de_DE.UTF-8"
LC_MESSAGES="de_DE.UTF-8"
LC_ALL=
```

詳細は、[locale\(1\)](#) のマニュアルページを参照してください。

システムで使用可能なロケールの一覧を取得するには、次のコマンドを使用します。

```
$ locale -a
```

さらに多くのロケールをインストールする方法の詳細は、[36 ページ](#)の「[使用可能なロケールの管理](#)」を参照してください。

複合ロケール

[14 ページ](#)の「[ロケールカテゴリ](#)」で説明されている、LC_CTYPE や LC_MESSAGES などの LC* 変数は、LANG 変数とともに端末でも設定できます。LC* 変数が設定されると、特定のカテゴリの LANG 設定がオーバーライドされます。このタイプのロケール設定は複合ロケールと呼ばれます。

```
$ export LANG=de_DE.UTF-8
$ export LC_MESSAGES=en_US.UTF-8
```

この例では、ロケール設定を正しく処理するアプリケーションはドイツ語ロケールで動作しますが、そのローカライズされた出力を英語で表示します。この場合、locale コマンドの出力は次のようになります。

```
$ locale
LANG=de_DE.UTF-8
LC_CTYPE="de_DE.UTF-8"
LC_NUMERIC="de_DE.UTF-8"
LC_TIME="de_DE.UTF-8"
LC_COLLATE="de_DE.UTF-8"
LC_MONETARY="de_DE.UTF-8"
LC_MESSAGES=en_US.UTF-8
LC_ALL=
```

LC_MESSAGES 変数はこの場合、アポストロフィなしで表示され、値が明示的に設定されることを示します。その他の LC* 変数の値は LANG 変数から継承されます。LC_ALL 変数を使用すると、すべての LANG および LC* 設定をオーバーライドできます。詳細は、[locale\(1\)](#) および [setlocale\(3C\)](#) のマニュアルページを参照してください。



注意 - 複合ロケールでは複数の文字セットを使用しないでください。たとえば、LANG=en_US.UTF-8 と LC_MESSAGES=ja_JP.eucJP を設定すると、予期しない結果が発生します。

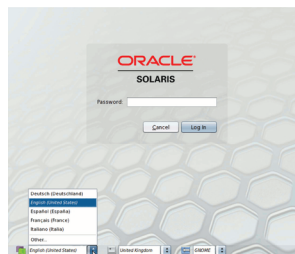
GNOME ディスプレイマネージャーでのロケール選択

Oracle Solaris 11 では、GNOME ディスプレイマネージャー (GDM) でロケールを選択できます。

▼ GNOME ディスプレイマネージャーでロケールを選択する方法

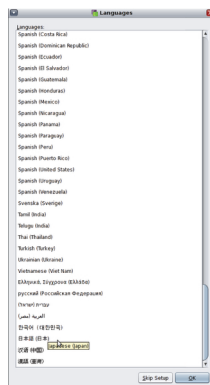
1. ユーザー名を入力し、「ログイン」ボタンをクリックするか、または **Return** キーを押します。
2. 「ログイン」ボタンをクリックするか、または **Return** を押します。

次の図に示すように、GDM ウィンドウの下部で「ロケールの選択」パネルが使用可能になります。



インストール処理中にロケール設定が指定されていない場合、デフォルトのロケール値は C ロケール (POSIX ロケールとも呼びます) に設定されます。

3. 使用可能なすべてのロケールの一覧を表示するには、「その他」を選択します。
システムにインストールされているすべてのロケールの一覧を含むウィンドウが表示されます。



4. 目的のロケールを選択して「了解」をクリックします。

▼ 別のロケールでセッションを開始する方法

1. ユーザー名を入力し、「ログイン」ボタンをクリックするか、または **Return** キーを押します。
2. ログイン画面で、「ロケールの選択」メニューから「ロケール」を選択します。
3. パスワードを入力します。
4. 「ログイン」ボタンをクリックします。

端末セッションでの永続的なロケール設定

ロケール設定を永続的にして、それを新しく開かれた端末セッションに適用するには、ユーザーのシェル初期化ファイル内にエクスポート文を配置します。詳細は、『[Oracle Solaris 11.3 のユーザーアカウントとユーザー環境の管理](#)』の「ユーザーの作業環境について」を参照してください。

リモート接続経由のロケールネゴシエーション

リモートアクセスを提供するアプリケーション (ssh や zlogin など) は、ロケールネゴシエーションを実行する可能性があります。これは、ユーザーの現在のロケール設定を、リモートの場所で使用可能な異なっている可能性のある一連のロケールに対して一致させようとする試みです。このプロセスにより、予期しない結果になる場合があります。次に例を示します。

- ユーザーは、リモートシステムのデフォルトのシステムロケールを持つシェルを取得することを期待しますが、ネゴシエーションが行われた別のロケールを取得します。
- ユーザーは、ネゴシエーションが行われたロケールを取得することを期待しますが、それが使用できないため、シェルにはデフォルトのシステムロケールが設定されます。

Secure Shell でのロケールネゴシエーションメカニズムの詳細は、[ssh\(1\)](#) および [sshd\(1M\)](#) のマニュアルページを参照してください。

詳細は、『[Oracle Solaris 11.3 での Secure Shell アクセスの管理](#)』および『[Oracle Solaris ゾーンの作成と使用](#)』の第 4 章、「非大域ゾーンへのログインについて」を参照してください。

デフォルトのシステムロケールの管理

デフォルトのシステムロケールは、システムがブートおよび動作するロケールです。Oracle Solaris の以前のリリースでは、デフォルトのシステムロケールは `/etc/default/init` ファイルで構成されていました。nlsadm コマンドを `get-system-locale`、`list-locale`、および `set-system-locale` サブコマンドとともに使用できます。詳細は、[nlsadm\(1M\)](#) のマニュアルページを参照してください。

現在のデフォルトのシステムロケールを表示するには、`get-system-locale` コマンドを使用します。

```
# nlsadm get-system-locale
LANG=en_US.UTF-8
LC_CTYPE=
LC_NUMERIC=
LC_TIME=
LC_COLLATE=
LC_MONETARY=
LC_MESSAGES=
LC_ALL=
```

使用可能なシステムロケールを一覧表示するには、次のコマンドを使用します。

```
# nlsadm list-locale
```

デフォルトのシステムロケールを設定するには、`set-system-locale` コマンドを使用します。たとえば、デフォルトのロケールを `fr_FR.UTF-8` に設定するには、次のコマンドを使用します。

```
# nlsadm set-system-locale fr_FR.UTF-8
```

インストーラでのロケール選択

初期のデフォルトのシステムロケールは、インストール中に設定されます。Oracle Solaris のインストーラはローカライズされているため、コア言語の話者であれば、自分の母国語を使用してインストール操作を進めることができます。Oracle Solaris 11 Live CD の 2 番目の画面では、インストール中に使用する言語をユーザーが選択できます。10 種類のコア言語のみを選択可能です。

図 1 インストーラでの言語選択



この言語選択によって、インストール先システムのデフォルトの言語サポートとその他のデータ形式も決定されます。

デフォルト設定にかかわらず、10個のすべてのコアロケールがインストールされ、ユーザーに対して使用可能になります。Automated Installer を使用してデフォルトを変更することにより、言語スコープを拡大または縮小できます。あるいは、`distro_const` コマンドを使用してカスタムの Live CD を生成できます。詳細は、[distro_const\(1M\)](#) のマニュアルページを参照してください。

注記 - テキストインストーラのローカリゼーションは、シリアルコンソール経由でのみ使用可能であり、物理コンソール経由では使用できません。

自動インストーラでのロケール選択

このセクションでは、Automated Installer マニフェストに対する可能な調整のうち、ロケール、キーボード、およびタイムゾーンの設定を変更するものについて説明します。

インストールするロケールの選択

インストールされるロケールの選択は、AI マニフェスト内のファセットによって制御されます。`<software><image>` 要素には、次の構文を持つ `facet` 要素が含まれます。

```
<facet set="true|false">facet-name</facet>
```

次のサンプルマニフェストでは、`facet` 要素を使用して、ドイツ語(ドイツ)と英語(米国)のロケールおよび翻訳のみがターゲットシステムに確実にインストールされるようにします。

```
<!DOCTYPE auto_install SYSTEM "file:///usr/share/install/ai.dtd.1">
<auto_install>
  <ai_instance auto_reboot="true" name="ai-german">
    ...
    <software type="IPS">
      <destination>
        <image>
          <!-- deselect all locales -->
          <facet set="false">facet.locale.*</facet>
          <!-- specify specific locales to install -->
          <!-- install German and English only -->
          <facet set="true">facet.locale.de</facet>
          <facet set="true">facet.locale.de_DE</facet>
          <facet set="true">facet.locale.en</facet>
          <facet set="true">facet.locale.en_US</facet>
        </image>
      </destination>
    </software>
  </ai_instance>
</auto_install>
```

次の例のようなコマンドを実行することにより、使用可能なロケールファセットのリストを `system/locale` パッケージから取得できます。

```
$ pkg contents -m system/locale | /usr/gnu/bin/grep -o facet.locale.[^\ ]* | sort -u
```

マニフェストの作成の詳細は、`ai_manifest(4)` のマニュアルページを参照してください。

UTF-8 以外のロケールのインストール

UTF-8 以外 (レガシー) のロケールはデフォルトインストールの一部ではありませんが、`system/locale/extra` パッケージでも引き続き使用できます。たとえば、`de_DE.ISO8859-1` ロケールをインストールするには、`system/locale/extra` パッケージも AI マニフェストに追加する必要があります。

```
...
<software_data action="install">
  <name>pkg:/entire@release</name>
  <name>pkg:/group/system/solaris-desktop</name>
  <name>pkg:/system/locale/extra</name>
</software_data>
...
```

デフォルトのシステムロケール、キーマップ、およびタイムゾーンの設定

[service_bundle\(4\)](#) および [installadm\(1M\)](#) マニュアルページで説明するように、システム構成 (SC) プロファイルで Automated Installer を使用してインストールされるシステムのデフォルトロケールを指定します。詳細は、『[Oracle Solaris 11.3 システムのインストール](#)』の「[構成プロファイルの指定](#)」を参照してください。

次の例は、デフォルトのシステムロケールをドイツ語に設定するシステム構成プロファイルを示しています。

```
<!DOCTYPE service_bundle SYSTEM "/usr/share/lib/xml/dtd/service_bundle.dtd.1">
<service_bundle type="profile" name="system configuration">
...
  <service name='system/environment' version='1'>
    <instance name='init' enabled='true'>
      <property_group name='environment'>
        <propval name='LANG' value='de_DE.UTF-8' />
      </property_group>
    </instance>
  </service>
...
</service_bundle>
```

サービス管理機構 (SMF) に移行されたその他の国際環境設定を、同様の方法で設定できます。次の例は、英国英語キーボードレイアウトおよび GMT タイムゾーンを使用する方法を示しています。

```
<!DOCTYPE service_bundle SYSTEM "/usr/share/lib/xml/dtd/service_bundle.dtd.1">
<service_bundle type="profile" name="system configuration">
...
<service name='system/keymap' version='1'>
  <instance name='default' enabled='true'>
    <property_group name='keymap'>
      <propval name='layout' value='UK-English' />
    </property_group>
  </instance>
</service>

<service name='system/timezone' version='1'>
  <instance name='init' enabled='true'>
    <property_group name='timezone'>
      <propval name='localtime' value='GMT' />
    </property_group>
  </instance>
</service>
...
</service_bundle>
```

使用可能なロケールの管理

Oracle Solaris のデフォルトインストールでは、次のコアロケールが使用可能です (言語ごとに 1 ロケールのみ)。

- 簡体字中国語 (zh_CN.UTF-8)
- 繁体字中国語 (zh_TW.UTF-8)
- 英語 (en_US.UTF-8)
- フランス語 (fr_FR.UTF-8)
- ドイツ語 (de_DE.UTF-8)
- イタリア語 (it_IT.UTF-8)
- 日本語 (ja_JP.UTF-8)
- 韓国語 (ko_KR.UTF-8)
- ポルトガル語 - ブラジル (pt_BR.UTF-8)
- スペイン語 (es_ES.UTF-8)

このセクションでは、追加言語のサポートを追加するためのさまざまな方法について説明します。

nlsadm を使用したロケールの追加または削除

nlsadm コマンドの `install-locale`、`list-locale`、および `uninstall-locale` サブコマンドを使用して、ロケールを管理したり、各国語のプロパティを管理したりできます。詳細は、[nlsadm\(1M\)](#) のマニュアルページを参照してください。

ロケールおよび使用可能なすべての翻訳をインストールするには、`install-locale` サブコマンドを使用します。たとえば、デンマーク語ロケールをインストールするには、次のコマンドを使用します。

```
# nlsadm install-locale da_DK.UTF-8
```

現在のイメージに基づいて、インストールに使用可能なロケールを一覧表示するには、`list-locale` コマンドを使用します。

```
# nlsadm list-locale
LOCALE          LANG  TERRITORY  CODESET  MODIFIER  FLAGS
af_ZA.UTF-8     af    ZA         UTF-8    -         -
...
```

デフォルトインストールでは、このコマンドは、インストール済みパッケージから使用可能なロケールのみを一覧表示します。パッケージリポジトリでは、さらに多くのロケールを使用できます。これは、`-a` オプションを使用して表示できます。

```
# nlsadm list-locale -a
Reading package information from IPS publisher's repository (it could take couple of
minutes) ...
LOCALE          LANG  TERRITORY  CODESET  MODIFIER  FLAGS
af_ZA.UTF-8     af    ZA         UTF-8    -         -
...
```

ロケール (デンマーク語) および関連するすべての翻訳をアンインストールするには、次のコマンドを使用します。

```
# nlsadm uninstall-locale da_DK.UTF-8
```

注記 - パッケージングの構造のために、要求されたロケール変更を行うことができず、副作用として追加のロケールがインストールまたは削除される場合があります。このような場合、`nlsadm` コマンドは影響を最小限に抑える解決方法を提供し、変更をコミットする前に必要な追加の変更をユーザーに通知します。

ロケールファセット

Oracle Solaris 11 での言語サポート関連コンポーネントのパッケージングを理解するには、ロケールファセットの概念に精通している必要があります。

ファセットとは

Oracle Solaris の以前のリリースでは、ドキュメント、ローカリゼーション、デバッグファイルなどのオプションコンポーネントが個別のパッケージに分割されていました。Oracle Solaris 11 の Image Packaging System を使用すると、Oracle はファセットと呼ばれる特殊なタグを使用して、オプションコンポーネントを同じパッケージ内に

保持できます。ファセットはパッケージ処理を単純化する一方で、追加の機能が必要でない場合にディスク容量の使用を低く抑えます。ファセットの詳細は、『Oracle Solaris 11.3 ソフトウェアの追加と更新』の「パッケージのファセットとバリエーション」および『Oracle Solaris 11.3 でのImage Packaging System を使用したソフトウェアのパッケージ化と配布』の第5章、「バリエーションの許可」を参照してください。

ロケールファセットは、言語またはロケールに固有であるファイルまたは処理にマークを付けるために使用されます。たとえば、web/wget パッケージのマニフェストでは、/usr/share/locale/ja/LC_MESSAGES/wget.mo ファイルに locale.ja=true のタグが付いています。このタグは、wget メッセージの日本語翻訳を含むファイルが、locale.ja ファセットを true に設定することによって日本語のサポートが有効になっている場合にのみインストールされることを示します。

ロケールファセットの構造

ロケールファセットには固定の形式はありません。Oracle IPS リポジトリでは次の規則が使用されます。

```
locale.{language}[_territory]
```

language は ISO 639 規格の 2 文字の言語コード、*territory* は ISO 3166 の 2 文字の地域コードです。

pkg を使用したロケールの追加または削除

言語サポートを追加するには、pkg コマンドを使用して、pkg change-facet コマンドで適切なロケールファセットを設定します。たとえば、フランスで話されるフランス語のサポートをインストールおよび追加するには、次のコマンドを使用できます。

- すべてのフランス語ロケールに共通のファイルをインストールするには

```
# pkg change-facet locale.fr=True
```

- フランス語 (フランス) に固有のファイルをインストールするには

```
# pkg change-facet locale.fr_FR=True
```

すべての使用可能なフランス語の変種のサポートを追加するには、次のコマンドを使用します。

- すべてのフランス語ロケールに共通のファイルをインストールするには

```
# pkg change-facet locale.fr=True
```

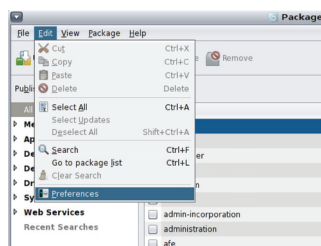
- すべての使用可能なフランス語の変種のサポートを追加するには、次のコマンドを使用します。

```
# pkg change-facet locale.fr_*=True
```

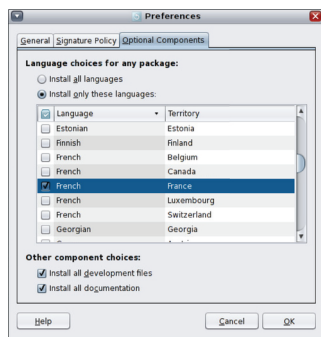
▼ パッケージマネージャーで言語サポートを追加する方法

パッケージマネージャーは、`system/locale` パッケージ内に存在する言語のみを管理します。その他のいずれかの言語のサポートを追加するまたは削除するには、コマンド行を使用します。

1. パッケージマネージャーのウィンドウで、「編集」->「設定」の順に選択します。



2. 「設定」ウィンドウで、「オプションのコンポーネント」タブを選択します。
3. 目的の言語を選択または選択解除します。



4. 「OK」をクリックします。

インストールされたパッケージのすべてのオプションコンポーネントは自動的に更新されます。

旧バージョンのロケール

Oracle Solaris でのデフォルトのシステムロケールは UTF-8 エンコーディングを使用しますが、旧バージョンのロケールもいくつか使用できます。36 ページの「[nlsadm を使用したロケールの追加または削除](#)」で説明されているように、`nlsadm` コマンドを使用して、旧バージョンのロケールをインストールまたは削除できます。

注記 - GNOME デスクトップ環境は UTF-8 ロケールのみをサポートします。

旧バージョンのロケールのサポートは `system/locale/extra` パッケージに含まれています。`nlsadm` コマンドは、必要な場合、このパッケージを自動的にインストールします。`pkg` ファセットの変更を使用してロケールを追加または削除している場合は、次のコマンドを使用してこのパッケージをインストールします。

```
# pkg install pkg:/system/locale/extra
```

特定言語のサポートを有効にするには、対応するロケールファセットを `true` に設定します。たとえば、`da_DK.ISO8859-1` ロケールと、`da_DK` ロケールのすべてをインストールするには、次のコマンドを使用します。

```
# pkg install pkg:/system/locale/extra
# pkg change-facet locale.da=True
# pkg change-facet locale.da_DK=True
```

ロケールの別名化

もっとも一般的なロケールは通常、主要なオペレーティングシステムでサポートされていますが、多くの場合、ロケールの名前は異なります。たとえば、Oracle Solaris では UTF-8 エンコーディングを使用しているフランスで話されるフランス語のロケール名として `fr_FR.UTF-8` を使用しますが、IBM AIX では `FR_FR` を使用し、HP-UX 11.11 および RHEL 5.4 では `fr_FR.utf8` を使用します。この不整合は、異機種システム混在環境において、または Oracle Solaris への移行時に問題となる可能性があります。

この問題に対処するために、ロケール別名のサポートが Oracle Solaris 11 の `libc` で導入されました。[setlocale\(3C\)](#) のマニュアルページで指定されているロケール選択プロセス中、および [gettext\(1\)](#)、[catopen\(3C\)](#)、[gettext\(3C\)](#) の各マニュアルページで指定されているメッセージオブジェクトまたはメッセージカタログ処理中に、ロケー

ル名の別名が受け入れられ、対応する標準のロケール名 (存在する場合) にマップされます。

加えて、以前の Oracle Solaris リリースとの互換性向上のために、メッセージ処理関数では、チェック対象の追加ロケール名として、`fr` や `fr_FR` のような廃止された Solaris ロケール名を使用してメッセージオブジェクトまたはカタログを検索するようになりました。詳細は、[Oracle Solaris 11 - End of Feature Notices](#) ページの「Localization」セクションの「Short Form Locales」を参照してください。

たとえば、代表的なユースケースはほぼ、ユーザーのシェル初期化ファイル内のロケール宣言で Linux スタイルのロケール名が使用される Linux 環境です。たとえば、`setenv LANG ja_JP.utf8` コマンドは `$HOME/.login` に含まれ、ホームディレクトリは NFS マウントされます。このネットワーク環境では、ユーザーが Oracle Solaris 11 システムにログインすると、ロケール別名サポートメカニズムが内部的かつ透過的にそのロケール名を対応する Oracle Solaris ロケール名 (`ja_JP.UTF-8`) にマップし、ユーザー指定のロケール名を尊重およびサポートします。同じように、Solaris 以外のロケール名が [ssh\(1\)](#) を通してリモートの Oracle Solaris 11 システムに渡される時、そのロケール名が認識、尊重、およびサポートされます。

ロケールの別名化の詳細は、[locale_alias\(5\)](#) のマニュアルページを参照してください。ここには、ロケール名マッピングの完全なリストも収められています。

デスクトップキーボードの設定と入力方式

Oracle Solaris 11 デスクトップでキーボードを構成するための、多数のさまざまなオプションがあります。ただし、ほとんどのデスクトップユーザーがキーボードを正しく構成するための設定は、ほんの少ししか必要ありません。この章では、キーボードの設定および入力方式と、用途に合わせてキーボードを正しく構成するプロセスについて説明します。

- [43 ページの「キーボード配列と入力方式の設定について」](#)
- [44 ページの「GNOME デスクトップマネージャーでのキーボード配列の選択」](#)
- [45 ページの「入力方式」](#)
- [57 ページの「コンソールでのキーボード配列」](#)

キーボード配列と入力方式の設定について

キーボード配列やキーボードモデルなどのキーボード配列設定を構成できます。2つの主な領域があります。

- 入力方式 (IM) –たとえば中国語、日本語、韓国語のようなアジア言語など、文字数が多い、または複雑な文字のある言語を通常使用する場合は、この方式を選択します。
- GNOME キーボード配列の設定 –たとえば英語、スペイン語、フランス語、ドイツ語のようなラテン文字を使用する言語など、文字数が少なめの言語を通常使用する場合は、この方式を選択します。

IM には文字を作成するためのより複雑なメカニズムがありますが、ラテン言語にも使用できます。同様に、GNOME キーボード配列設定は、ラテン言語以外にも使用できます。

注記 - さまざまなキーボード配列の設定が可能ですが、キーボードに関連するすべての領域を設定する必要はありません。たとえば、IM をアクティブ化する代わりに、1 つまたは 2 つのラテン語キーボード配列 (たとえば、US/English と French) だけを使用する場合があります。キーボード配列は GNOME キーボード設定で設定することができます。一方、中国語または日本語のキーボード配列を使用する場合は、おそらく IM のほうが適しています。

GNOME デスクトップマネージャーでのキーボード配列の選択

Oracle Solaris 11 は、Gnome デスクトップマネージャー (GDM) でのキーボード選択をサポートしています。

▼ GNOME デスクトップマネージャーでキーボード配列を選択する方法

GNOME セッションのキーボード配列は GDM で選択できますが、必須ではありません。キーボード配列を選択しない場合は、デフォルトのキーボード配列が選択されます。GNOME セッションのあとで、または次回のログイン時に、別の配列を選択できます。

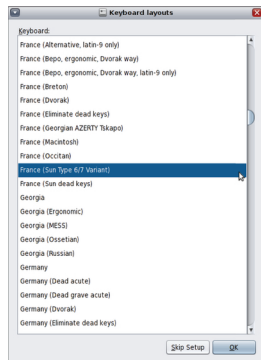
1. ユーザー名を入力します。
2. 「ログイン」 ボタンをクリックするか、**Return** を押します。
「キーボード配列の選択」 パネルが表示可能になります。次の図に示されているように、このパネルは GDM ウィンドウの下部にあります。



キーボード配列のデフォルト値は「USA」に設定されています。

3. 使用可能なすべてのキーボード配列のリストを表示するには、「その他」を選択します。

使用可能なすべてのキーボード配列のバリエーションのリストを含むウィンドウは、次の図に示されているように表示されます。



4. 必要な配列を選択し、「了解」をクリックします。

GDM でキーボード配列を選択する

GDM でキーボード配列が選択された場合、選択されたキーボード配列の値は `$HOME/.dmrc` ファイルに格納され、次回ログイン時に事前選択されます。

注記 - GDM でのキーボード配列の選択は、クライアントとサーバーの両方が同じ物理マシン上にある場合にのみ有効で、リモート接続に対しては無効です。GDM がリモートからアクセスされた場合、キーボード配列のデフォルト値は、`$HOME/.dmrc` ファイルの内容に関係なく USA に設定されます。

入力方式

入力方式 (IM) は、キーボードのような入力デバイスにない特定の文字を、さまざまなデスクトップアプリケーションに入力するメカニズムです。中国語、インド語、日本語、韓国語、およびタイ語などの言語には、入力デバイスで使用可能な文字よりもかなり多くの文字セットが含まれているため、これらの言語には IM が必須です。IM は、キーストロークの組み合わせを、入力デバイスから言語固有の文字に変換し、フォーカスにあるアプリケーションに情報を返します。

入力方式には、IM フレームワークと IM 言語エンジンという 2 つの主要なコンポーネントがあります。IM フレームワークは、IM 言語エンジンとユーザーアプリケーションとの間の連携を可能にする機能を提供する、ソフトウェアコンポーネントです。IM 言語エンジンは、IM フレームワークからキーストロークの組み合わせを取得し、特定の言語の文字に変換して IM フレームワークに返す、ソフトウェアコンポーネントです。

IBus は、Oracle Solaris デスクトップシステム上のデフォルトの IM フレームワークです。*IIIMF* は、セカンダリ IM フレームワークとしてインストーラリポジトリにあります。

デフォルトでは、入力方式は、次の言語でログインしたときにだけアクティブ化されます。

- 中国語 (簡体字)
- 中国語 (繁体字)
- 日本語
- 韓国語
- インド語
- タイ語

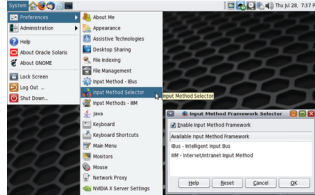
ほかの言語の場合は、必要に応じて、入力方式フレームワークセレクタを使用して IM を手動でアクティブ化できます。

▼ 入力方式のアクティブ化と非アクティブ化の方法

入力方式フレームワークセレクタ (*imf-selector*) は、優先される入力方式フレームワークを選択して、デスクトップセッションで有効または無効にするために使用される構成ツールです。

1. 「システム」->「設定」->「入力方式セレクタ」を選択します。
「入力方式フレームワークセレクタ」ウィンドウが表示されます。

2. 「入力方式フレームワークの有効化」オプションを選択し、リストから優先される入力方式フレームワークを選択します



3. 「OK」ボタンをクリックして、選択を保存します。

IBus について

Linux および Unix オペレーティングシステム用の Intelligent Input Bus (IBus) は、多くのオープンソースの IM 言語エンジンで動作する強力な多言語の入力方式フレームワークです。IBus は、バスに似たアーキテクチャーを使用して、IBus IM フレームワークと IM 言語エンジンとの間の通信を処理します。このプロセスはユーザーセッションごとに実行されます。異なるユーザーデスクトップセッション間で共有されるプロセスはありません。

IBus の構成

IBus は、IBus 設定ツール (ibus-setup) によってユーザーごとに構成されます。このツールにアクセスするには、「システム」->「設定」->「入力方式 - IBus」を選択します。次のタスクでは、変更可能ないくつかの主要な構成について説明します。

▼ 新しい言語エンジンを追加する方法

1. 「IBus 設定」ウィンドウで「入力方式」タブをクリックします。
2. 優先する IM 言語エンジンを選択します。
3. 「追加」をクリックします。

▼ 新しい入力方式トリガーキーを追加する方法

1. 「IBus 設定」ウィンドウで「一般」タブをクリックします。

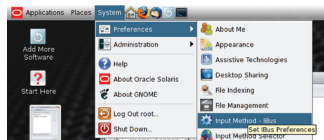
2. 「有効」または「無効」ラベルの「...」ボタンをクリックします。次に、キーコードのラベルと情報ダイアログのポップアップで「...」ボタンをクリックします。ここで、トリガーキーとして優先するキーの組み合わせを押して、情報ダイアログの「閉じる」ボタンをクリックします。次に、トリガーウィンドウの「キーボードの選択」ショートカットで「追加」および「OK」ボタンを押します。
3. トリガーキーは、フォーカスにあるアプリケーションで **IBus** をアクティブ化するために使用されます。**IBus** がアクティブ化されていない場合、文字は入力デバイスからフォーカスがあるアプリケーションに直接送られます。デフォルトのトリガーキーが「有効」または「無効」ラベルに表示されます。

▼ 言語パネルの場所を構成する方法

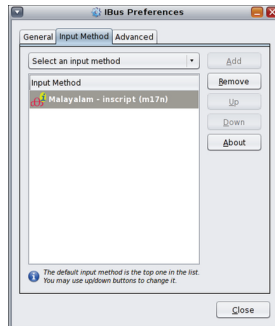
1. **IBus** 設定ツールで、「言語パネルの表示」メニューから構成を選択します。
言語パネルの位置として、3つの可能な位置があります。
 - 「メニューに組み込む」(デフォルト) - 言語パネルは GNOME パネルの通知領域に埋め込まれます。
 - 「アクティブであるとき」 - 言語パネルは、入力方式がアクティブ化されているときにだけ、独立したウィンドウとして表示されます
 - 「常に表示する」 - 言語パネルは、常に独立したウィンドウとして表示されます。
2. 「OK」をクリックします。

▼ **IBus** で新しい言語エンジンを追加する方法

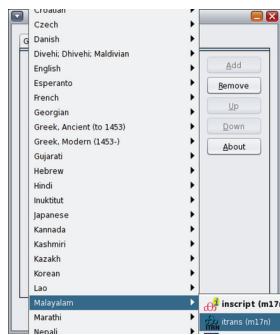
1. 「システム」->「設定」->「入力方式 - **IBus**」を選択します。



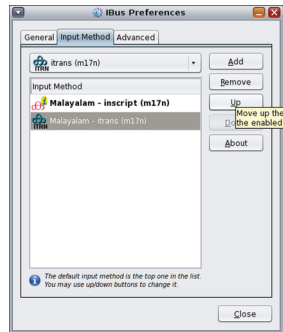
2. 「IBus 設定」ウィンドウが開いたら、「入力方式」タブをクリックします。



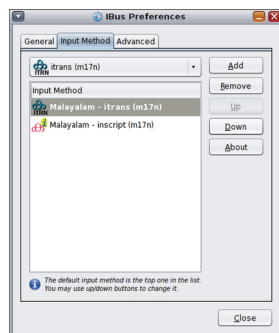
3. 「入力方式の選択」を選択し、必要な入力方式を選択します。



- 「追加」ボタンをクリックして、選択したリストに追加します。



- 新しい入力方式をデフォルトにするには、「上へ」ボタンをクリックします。選択をリストの最上部に移動します。



- 現在選択されている入力方式をデフォルトとして使用するには、ログアウトしてから再度ログインします。

IBus 入力方式フレームワークで使用可能な言語エンジン

IBus 入力方式フレームワークでは、次の言語エンジンを使用できます。

- デスクトップのインストールプロセス中にデフォルトでインストール
 - Anthy - 日本語
 - Chewing - 繁体字中国語
 - Hanguk - 韓国語

- IBus-Sayura -- シンハラ語
- IBus-XKBC - キーボードエミュレーション IM エンジンフレームワーク
- SunPinyin - 簡体字中国語
- 多言語化に基づく各種言語エンジン。
- オプションの言語エンジン
 - Pinyin - 簡体字中国語エンジン
 - IBus table フレームワークに基づく各種言語エンジン

インド語の入力方式

次の表で、インド語に使用できる入力方式の概要を示します。

表 8 インド語の入力方式

ロケール	言語	入力方式
as_IN.UTF-8	アッサム語	Inscript 配列、ITRANS 転写方式、表音配列
bn_IN.UTF-8	ベンガル語	Inscript 配列、ITRANS 転写方式、probhat 配列、Unijoy キーボード配列
en_IN.UTF-8	英語	入力方式は不要
gu_IN.UTF-8	グジャラート語	Inscript 配列、ITRANS 転写方式、表音配列
hi_IN.UTF-8	ヒンズー語	Inscript 配列、ITRANS 転写方式、表音配列、レミントン製タイプライタの配列、`typewriter' を含む入力方式
kn_IN.UTF-8	カナラ語	Inscript 配列、ITRANS 転写方式、KGP 方式
ks_IN.UTF-8	カシミール語	カシミール語のキーボードをシミュレートする入力方式
ml_IN.UTF-8	マラヤーラム語	Inscript 配列、ITRANS 転写方式、Mozhi 入力方式、Swanalekha 入力方式
mr_IN.UTF-8	マラーティー語	Inscript 配列、ITRANS 転写方式、表音配列
or_IN.UTF-8	オリヤー語	Inscript 配列、ITRANS 転写方式、表音配列
pa_IN.UTF-8	パンジャブ語	inscript 配列、ITRANS 転写方式、jhelum 配列、表音配列
sa_IN.UTF-8	サンスクリット語	京都・ハーバード方式による入力方式
ta_IN.UTF-8	タミル語	Inscript 配列、ITRANS 転写方式、表音配列、Renganathan

ロケール	言語	入力方式
te_IN.UTF-8	テルグ語	配列。tamil99 配列、タイプライター配列、"vutam" Type_As_You_Write 配列。 Telugu 用の Apple キーボード配列、inscript 配列、ITRANS 転写方式、pothana Telugu 入力方式バージョン2.0、RTS 方式

IBus XKBC IM エンジン

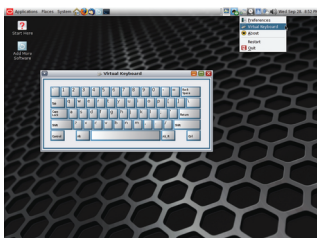
IBus XKBC は、XKeyboard 構成データベースを使用してキーボード配列をエミュレートします。XkeyboardConfig で使用可能な、キーボード配列のバリエーションを含むすべてのキーボード配列は、IBus IM フレームワークでエミュレートできます。IBus XKBC は、IBus 設定ツールの「その他」の言語カテゴリで使用可能です。IBus XKBC ヘルプに、詳細な構成情報があります。

IBus 仮想キーボード (ibus-keyboard)

IBus 仮想キーボードは、IBus XKBC 言語エンジンに基づいたグラフィカルキーボードエミュレータです。これは、グラフィカルキーボードをウィンドウに表示し、キーアイコンをクリックすることによって、フォーカスにあるアプリケーションに文字を入力できます。IBus XKBC と同様に、IBus 仮想キーボードは、エミュレーション用のさまざまなキーボード配列とバリエーションをサポートします。IBus 仮想キーボードを起動するには、「GNOME」パネルの「IBus」メニューから「仮想キーボード」を選択します。

仮想キーボードの構成メニューを表示するには、仮想キーボードアプリケーションのウィンドウを右クリックします。

図 2 仮想キーボード



Internet Intranet Input Method Framework

Internet Intranet Input Method Framework (IIIMF) は、Solaris 9 以降でサポートされている、Oracle Solaris 11 のもう 1 つの IM フレームワークです。IIIMF は、システム規模の共有プロセスとしてではなく、ユーザーごとに実行されます。IIIM 用の構成ツールは、入力方式設定エディタ (iiim-properties) です。このツールにアクセスするには、「システム」->「設定」->「入力方式 - IIIM」を選択します。

IIIMF には、2 つの非常に強力な日本語の言語エンジンである、ATOK と Wnn があります。次の手順で、IIIMF で ATOK または Wnn 言語エンジンを使用方法を説明します。

注記 - このツールについての詳細は、『[ATOK for Solaris ユーザーガイド](#)』を参照し、これらの言語エンジンの詳細情報は『[Wnn8 ユーザーズガイド](#)』を参照してください。これらのドキュメントは日本語版のみ入手可能です。

▼ IIIMF で ATOK および Wnn 言語エンジンを使用方法

1. パッケージマネージャーの GUI または `pkg` コマンドを使用して、次のパッケージをインストールします。

- IIIMF コアパッケージ: `system/input-method/iiim`
- ATOK パッケージ: `system/input-method/iiim/atok`
- Wnn パッケージ: `system/input-method/iiim/wnn`

インストールに使用可能な IIIMF 言語エンジンをリストするには、次のコマンドを実行します。

```
$ pkg list -a system/input-method/iiim/*
```

2. 入力方式フレームワークセレクタを使用して、デスクトップセッションで IIIMF を選択します。

次のログイン時に、IIIMF がデスクトップセッションに使用されます。(デスクトップセッションを再起動する必要があります。)

3. Wnn 言語エンジンを使用している場合は、

- a. 次のコマンドを発行して、`jserver` および `dpkeyserv` サーバーを呼び出します。(サービス管理権限が必要。)

```
# svcadm enable wnn8/server
```

- b. 入力方式設定エディタ (iiim-properties) ツールで Wnn IM 言語エンジンを選択します。

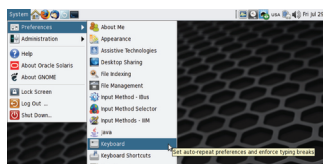
GNOME キーボード設定

Oracle Solaris 11 デスクトップには、100 以上の言語に使用できる 400 以上のキーボード配列があります。ほとんどすべての言語で、複数のキーボード配列のバリエーションが使用可能です。キーボード配列モデルと物理的配列に関係なく、Oracle Solaris 11 デスクトップで使用できるどのキーボード配列でも、常に構成および使用できます。キーボードモデル、配列、バリエーションなどのキーボード設定を設定およびカスタマイズするには、GNOME キーボード設定ツールを使用します。

注記 - キーボード配列の設定ウィンドウは、クライアントとサーバーの両方が同じ物理マシン上にある場合にのみ使用可能で、リモート接続の場合は使用できません。

▼ キーボード設定とカスタマイズを設定する方法

1. 「システム」->「設定」->「キーボード」を選択します。



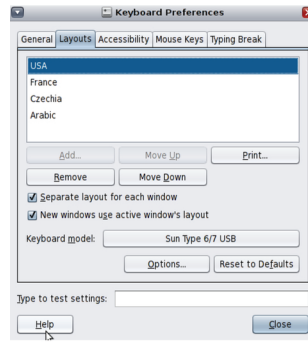
2. 「配列」タブをクリックします。

「レイアウト」タブに、キーボード配列を設定およびカスタマイズするために必要なすべてのツールが表示されます。最大 4 つのキーボード配列を追加でき、パネルの「キーボードのレイアウト表示器」メニューを使用することで、またはそれらを切り替えるキーボードショートカットを設定することで、簡単に切り替えることができます。

「レイアウト」タブでは、キーボードモデル、特定のキーボード配列のバリエーションなどのキーボードのさまざまな設定をカスタマイズすることもでき、**Shift**、**Alt**、**Meta**、**Super**、**Hyper**、**CapsLock** などの特殊キーの特殊動作を設定することもできます。

「レイアウト」タブのリストの最上部にある配列は、デフォルトの配列です。「上へ」および「下へ」ボタンを使用して、リスト内で配列を上下に移動できます。

次の図は「レイアウト」タブを示しています。



キーボード配列のオプションとデフォルトの動作

次のリストで、「キーボード設定」ウィンドウの「レイアウト」タブでもっとも多く実行されるトランザクションを示します。

▼ キーボード配列の追加と削除の方法

1. 国別または言語別に必要なキーボード配列を選択します。
2. 「追加」ボタンをクリックします。
3. 配列を削除するには、リストから削除したいキーボード配列を選択します。
4. 「削除」ボタンをクリックします。

▼ 特定のキーに EuroSign を追加する方法

1. 「オプション」ボタンをクリックします。
2. 「特定のキーに EuroSign を追加」を選択します。
3. 希望する 1 つまたは複数のキーを選択します。
4. 「閉じる」ボタンをクリックします。

▼ X サーバーを終了するキーシーケンスを設定する方法

1. 「オプション」 ボタンをクリックします。
2. X サーバーを終了するキーシーケンスを選択します。
3. **Control + Alt + Backspace** キーを選択します。
4. 「閉じる」 ボタンをクリックします。

▼ 選択したキーボード配列を切り替えるキーボードショートカットを設定する方法

複数のキーボード配列を使用している場合は、簡単に切り替えるためのショートカットを割り当てることができます。

1. 「オプション」 ボタンをクリックします。
2. 配列を変更するキーを選択します。
3. ショートカットとして、希望する 1 つまたは複数のキーを選択します。
4. 「閉じる」 ボタンをクリックします。

GNOME のキーボード配列インジケータ (キーボードセレクト)

次の図のように複数のキーボード配列が選択されると、キーボード配列インジケータが自動的にアクティブ化され、パネルに表示されます。1 つのキーボードだけが選択されているとき (デフォルトの動作)、キーボード配列インジケータはパネルに表示されません。

図 3 キーボード配列インジケータ



「キーボードのレイアウト表示器」メニューがアクティブ化され、パネル上に表示されているときは、マウスを使用して、選択したキーボード配列間で切り替えることが

できます。今すぐ次のキーボード配列に切り替えるには、シングルクリックします。右クリックすると、3つのオプションのあるメニューが開きます。

- 「グループ」 – 選択したキーボード配列のリストを含めます。
- 「キーボード設定」 – 「キーボード配列の設定」ウィンドウを起動します。
- 「現在の配列を表示」 – 対話型アプリケーションでウィンドウを起動して、現在のキーボード配列を表示します。

X キーボード拡張 (setxkbmap) を使用したキーボード配列の設定

GNOME キーボード配列設定ツールを使用するのではなく、setxkbmap コマンドを使用して、コマンド行から X サーバーのすべてのキーボード配列の設定を設定およびカスタマイズできます。

setxkbmap コマンドは、コマンド行に指定されたオプションによって決められるキーボード配列を使用するために、キーボードをマップします。setxkbmap コマンドの設定可能なオプションには、幾何図形、キーボードモデル、配列記号、配列のバリエーション、ルールなどが含まれます。

次の例で、コマンド行から2つのキーボード配列 (US/English と French) を設定する方法を示します。

```
$ /usr/bin/setxkbmap us,fr
```

詳細は、setxkbmap(1) のマニュアルページを参照してください。

コンソールでのキーボード配列

この操作が必要になる頻度は低いと思われませんが、nlsadm コマンドを get-console-keymap、list-console-keymap、および set-console-keymap サブコマンドとともに使用して、コンソールでのキーボード配列を変更できます。Oracle Solaris コンソールでのコンソールキーボード配列には、ASCII 文字のみが含まれます。詳細は、nlsadm(1M) のマニュアルページを参照してください。

現在のコンソールキーボード配列の設定を表示するには、次のコマンドを使用します。

```
# nlsadm get-console-keymap  
keymap=US-English
```

コンソールキーボード配列の設定の使用可能な値を一覧表示するには、次のコマンドを使用します。

```
# nlsadm list-console-keymap
Arabic
...
```

コンソールキーボード配列を日本語に設定するには、次のコマンドを使用します。

```
# nlsadm set-console-keymap Japanese
```

注記 - インストールされているシステム上の `/usr/share/lib/keytables/type_6/kbd_layouts` ファイルでは、使用可能なキーボード配列の一覧が提供されます。

詳細は、[付録A 使用可能なロケールおよびサポートされている文字セット](#)を参照してください。

フォントの構成

この章では、フォント構成のベースとなるライブラリについて説明し、特定のアジア言語で使用可能なフォントの一覧を示します。GNOME デスクトップ環境のフォント構成については、『Oracle Solaris 11.3 デスクトップユーザズガイド』の「フォント設定」で詳細に説明されています。

この章で扱う内容は、次のとおりです。

- 59 ページの「fontconfig ライブラリ」
- 60 ページの「使用可能なフォント」

fontconfig ライブラリ

Oracle Solaris デスクトップ環境では、フォントアクセスの構成とカスタマイズのために fontconfig ライブラリを使用します。fontconfig は、fonts.conf(4) ファイル内の構成に基づいた自動検出を使用して、システムで使用可能なすべてのフォントの一覧を維持します。アプリケーションが名前によって、またオプションでサイズ、フォントの太さ、言語などのほかの属性によって特定のフォントを検索すると、fontconfig がアプリケーションに特定のフォントパスを提供します。fontconfig はフォントをテキストにレイアウトまたは描画しません。

フォント構成の調整

特定のロケールのデフォルトフォントに満足できない場合があります。/etc/fonts.fontconfig ディレクトリ内の fontconfig 構成ファイルに変更を加えることで、フォントの優先順位を変更できます。システム全体の構成は、/etc/fonts/conf.avail ディレクトリ内の個別の xml ファイルにモジュール化されます。/etc/fonts/conf.d ディレクトリにはその後、/etc/fonts/conf.avail ディレクトリで選択された構成ファイルへのシンボリックリンクが置かれます。これらのシンボリックリンクを追加または削除することによって、システム全体の構成を変更できます。

デフォルトフォントやほかの使用可能フォントについての情報を取得するには、`fc-match(1)` および `fc-list(1)` コマンドを使用します。たとえば、次のコマンドは日本語のデフォルトのモノスペースフォントについての情報を表示します。(固定幅フォントとも呼ばれるモノスペースフォントは通常、テキストエディタで使用されます。)

```
$ fc-match Monospace:lang=ja
ipag.otf: "IPAGothic" "Regular"
```

代替の TrueType フォントが使用可能でそちらを使いたい場合に、ビットマップフォントの使用を無効にするには、次の例に示すように、構成ファイル `70-no-bitmaps.conf` へのシンボリックリンクを作成します。

```
# cd /etc/fonts/conf.d
# ln -s ../conf.avail/70-no-bitmaps.conf
```

`fontconfig` は構成ディレクトリ `/etc/fonts/conf.d` をモニターし、そのメモリー内構成を自動的に調整して変更を反映します。詳細は、`fc-cache(1)` のマニュアルページを参照してください。ユーザー単位の構成も、各ユーザーの `$HOME/.fonts.conf` で管理できます。`$HOME/.fonts` からユーザーのカスタムフォントファイルがスキャンされます。`fontconfig` とその構成の詳細は、`fonts.conf(4)`、`fc-match(1)`、および `fc-list(1)` のマニュアルページ、または <http://www.freedesktop.org/wiki/Software/fontconfig> のプロジェクトページを参照してください。

使用可能なフォント

この Oracle Solaris 11 リリースでは、システムで使用可能なフォントが多数更新されており、TrueType または OpenType フォント形式を使用する方向に大きく転換しています。

広く知られる Vera フォントをベースにした [DejaVu フォント](#) など、多くのフォントはオープンソースプロジェクトに由来し、定期的に更新されます。DejaVu フォントは Unicode での使用のために設計されているため、多くのロケールでシステムのデフォルトフォントです。

フォントはロケールファセットの変更時に自動的にシステムに追加されます。詳細は、[37 ページの「ロケールファセット」](#)を参照してください。パッケージマネージャーを使用した追加フォントのインストールのために、フォントは `System/Fonts` にカテゴリ化されます。

インド語派フォント

次の表では、インド語派フォントについて説明します。

表 9 インド語派フォント

ロケール	言語	フォント
as_IN.UTF-8	アッサム語	Mukti Narrow
bn_IN.UTF-8	ベンガル語	Mukti Narrow
en_IN.UTF-8	英語	DejaVu Sans
gu_IN.UTF-8	グジャラート語	Lohit Gujarati
hi_IN.UTF-8	ヒンズー語	Lohit Hindi
kn_IN.UTF-8	カナラ語	Lohit Kannada
ks_IN.UTF-8	カシミール語	Lohit Kashmiri
ml_IN.UTF-8	マラーヤラム語	Lohit Malayalam
mr_IN.UTF-8	マラーティー語	Lohit Marathi
or_IN.UTF-8	オリヤー語	Lohit Oriya
pa_IN.UTF-8	パンジャブ語	Lohit Punjabi
sa_IN.UTF-8	サンスクリット語	Lohit Sanskrit
ta_IN.UTF-8	タミル語	Lohit Tamil
te_IN.UTF-8	テルグ語	Lohit Telugu

日本語フォント

次の表では、日本語 TrueType フォントについて説明します。

表 10 日本語フォント

完全なファミリ名	サブファミリ	表記形式	スタイル	ベンダー	文字セット
IPAGothic	標準	TrueType	Sans serif、固定幅	IPA	JIS X 0201:1976、JIS X 0213:2004
IPAPGothic	標準	TrueType	Sans serif、プロポーショナル	IPA	JIS X 0201:1976、JIS X 0213:2004
IPAMincho	標準	TrueType	Serif、固定幅	IPA	JIS X 0201:1976、JIS X 0213:2004
IPAPMincho	標準	TrueType	Serif、プロポーショナル	IPA	JIS X 0201:1976、JIS X 0213:2004
Sun-Gothic	標準、太字	PCF (12,14,16,20,24)	Sans serif、固定幅	SCR	JIS X 0201:1976、JIS X 0208:1990
Sun-Mincho	標準、太字	PCF(16,20,24)	Serif、固定幅	SCR	JIS X 0201:1976、JIS X 0208:1990

日本語フォントは次のようにパッケージ化されます。

- IPAGothic、IPAPGothic – system/font/truetype/ipafont
- IPAMincho、IPAPMincho – system/font/truetype/ipafont-mincho
- Sun-Gothic、Sun-Minchou – system/font/sun-ja-bitmap-unicode (Unicode インデックスフォント) および system/font/sun-ja-bitmap (JIS インデックスフォント)

Unicode インデックスフォントは、GTK アプリケーションでビットマップフォントを使用するために提供されます。たとえば、日本語 GNOME 環境でビットマップフォントを使用するには、system/font/sun-ja-bitmap-unicode をインストールします。

以前のリリースの CDE 環境で使用されていた日本語ビットマップフォントを使用するには、system/font/sun-ja-bitmap をインストールします。

韓国語フォント

次の表では、韓国語 TrueType フォントについて説明します。

表 11 韓国語フォント

完全なファミリ名	サブファミリ	表記形式	ベンダー	文字セット
RoundedGothic、Kodig、Haeseo、Myeongjo。	標準	TrueType	HanYang Systems Inc	KSC5601-1992
SunDotumChe、SunDotum	標準	TrueType	Oracle Corporation	KSC5601-1992
Baekmuk: Batang、Dotum、Gulim、Headline	標準、太字	TrueType	Baekmuk Font21 Inc.	KSC5601-1987
unfonts	標準、太字	TrueType	KLDP	ISO8859-1、ISO8859-5、KOI8
Unifont	標準	TrueType	GNU Unifont	KSC5601

次の表では、韓国語ビットマップフォントについて説明します。

表 12 韓国語ビットマップフォント

完全なファミリ名	サブファミリ	表記形式	文字セット
Gothic	標準	PCF(16)	KSC5601-1987
Mincho	標準	PCF(16,24)	KSC5601-1987

簡体字中国語および繁体字中国語フォント

次の表では、繁体字中国語および簡体字中国語 TrueType フォントについて説明します。

表 13 簡体字中国語および繁体字中国語フォント

完全なファミリ名	サブファミリ	表記形式	ベンダー	文字セット
AR PL UMinG	標準	TrueType	Arphic	BIG5、BIG5-HKSCS、CNS11643.1992
AR PL UKai	標準	TrueType	Arphic	BIG5、BIG5-HKSCS、CNS11643.1992
WenQuanYi Zen Hei	標準	TrueType	WenQuanYi	BIG5、BIG5-HKSCS、CNS11643.1992
Unifont	標準	TrueType	GNU Unifont	BIG5、BIG5-HKSCS、GB18030-2000、GB2312-1980
uming	標準	TrueType	Arphic	GB18030-2000、GB2312-1980、GBK
ukai	標準	TrueType	Arphic	GB18030-2000、GB2312-1980、GBK
wqy-Zenhei	標準	TrueType	WenQuanYi	ISO8859-1、ISO8859-5
Droid-Sans	標準	TrueType	Google Corporation	BIG5、GB2312-1980、KSC5601、JIS0208

次の表では、繁体字中国語および簡体字中国語ビットマップフォントについて説明します。

表 14 繁体字中国語および簡体字中国語ビットマップフォント

完全なファミリ名	サブファミリ	表記形式	文字セット
Song	標準	PCF(16,24)	GB2312-1980
FangSong	標準	PCF16	GB2312-1980
Unifont	標準	PCF16	BIG5、BIG5-HKSCS、GB18030-2000、GB2312-1980

タイ語フォント

次の表では、タイ語 TrueType フォントについて説明します。

表 15 タイ語フォント

完全なファミリ名	サブファミリ	表記形式	ベンダー	文字セット
Garuda	Book、Bold、Oblique、BoldOblique	TrueType	NECTEC/TLGW	ASCII、ISO8859-1、TIS620
Kinnari	Medium、Oblique、Italic、Bold、BoldOblique、BoldItalic	TrueType	National Electronics and Computer Technology Center	ASCII、ISO8859-1、TIS620
Loma	Book、Oblique、Bold、BoldOblique	TrueType	NECTEC	ASCII、ISO8859-1、TIS620
Norasi	Regular、Oblique、Italic、Bold、BoldOblique、BoldItalic	TrueType	Yannis Haralambous、Virach Sornlertlamvanich、Anutara Tantraporn	ASCII、ISO8859-1、TIS620
Purisa	Medium	TrueType	Poonlap Veerathanabutr	ASCII、ISO8859-1、TIS620
Sawasdee	Regular、Oblique、Bold、BoldOblique	TrueType	Pol Udomwittayanukul	ASCII、ISO8859-1、TIS620、MS-cp1252
TlwgMono	Medium、Oblique、Bold、BoldOblique	TrueType	Poonlap Veerathanabutr	ASCII、ISO8859-1、TIS620
TlwgTypewriter	Medium、Oblique、Bold、BoldOblique	TrueType	Poonlap Veerathanabutr	ASCII、ISO8859-1、TIS620
TlwgTypist	Medium、Oblique、Bold、BoldOblique	TrueType	Theppitak Karoonboonyanan	ASCII、ISO8859-1、TIS620
Umpush	Book、Oblique、Light、LightOblique、Bold、BoldOblique	TrueType	Widhaya Trisarnwadhana	ISO8859-1、TIS620
Waree	Book、Oblique、Bold、BoldOblique	TrueType	Bitstream	ASCII、ISO8859-1、TIS620、MS-cp1252
Unifont	標準	TrueType	GNU Unifont	TIS620、ISO8859-1

詳細情報

この章では、コードセット、印刷、相互運用性、ロケールの作成に関連する、いくつかの厳選された高度なトピックについて説明します。この章で扱う内容は、次のとおりです。

- [65 ページの「コードセット変換」](#)
- [69 ページの「国際化ドメイン名のサポート」](#)
- [71 ページの「印刷の拡張」](#)
- [72 ページの「ほかのプラットフォームとの相互運用性」](#)
- [74 ページの「各国語プロパティの構成」](#)
- [75 ページの「カスタムロケールの作成」](#)

コードセット変換

コードセット変換、つまり文字セット (**charset**) 変換のサポートは、ほとんどのアプリケーションがこの機能に依存することで正しく機能しているため、オペレーティングシステムには不可欠な部分です。

Oracle Solaris の現在のリリースには、Unicode のサポート、ソフトウェアの国際化、およびソフトウェアのグローバル化のために広く使用されているライブラリとツールである、Unicode 用の国際化コンポーネント (ICU) も組み込まれています。

Oracle Solaris 11 には、コードセット変換のためのさまざまなツールとライブラリが組み込まれています。中核となるコードセット変換ユーティリティ、**iconv** は、Oracle Solaris libc 内の **iconv** ライブラリを基に構築されています。

iconv ユーティリティ

iconv(1) コマンド行ユーティリティは、文字または文字のシーケンスを、1つのコードセットから別のコードセットに変換します。これは広範囲のコードセットをサ

ポートします。コードセット名はプラットフォームによって異なることがよくあるため、`iconv` の別名メカニズムによって多くのコードセットが複数の名前をサポートされます。システムで現在使用可能なコードセットのリストを取得するには、次のコマンドを実行します。

```
$ /usr/bin/iconv -l
```

複数のパッケージに `iconv` モジュールが含まれているため、追加のパッケージをインストールすることでデフォルトのリストを拡張できます。デフォルトのインストールには `system/library/iconv/utf-8` パッケージが含まれ、これに、UTF-8 やその他の Unicode コードセットと選択されたほかのコードセットとの間の変換用の、`iconv` モジュールの基本セットが含まれています。その他のパッケージは、パッケージマネージャーの「システム/国際化」カテゴリから、または `pkg(1)` コマンドを使ってインストールする際に `system/library/iconv/*` という名前のパターンを使用することで入手できます。

`iconv -f` オプションはソースコードセットを定義し、`-t` オプションはターゲットコードセットを定義します。次のように、`iconv` を使用すると、ファイル、つまり標準入力を、標準出力に変換できます。

```
$ /usr/bin/iconv -f eucJP -t UTF-8 file.txt
```

この例では、`file.txt` ファイル名が `eucJP` コードセット (日本語 EUC、Extended UNIX Code Packed Format for Japanese) から変換され、結果が UTF-8 で標準出力に書き込まれます。

Oracle Solaris 11 では、`iconv` が拡張されて、これらの特殊な状況での変換の動作を変更するフラグが組み込まれました。

- 不正な文字 – 入力文字が宣言されたソースコードセットで有効ではない
- 同一でない文字 – ターゲットコードセットに一致する文字がない

`//ILLEGAL_DISCARD`、`//NON_IDENTICAL_DISCARD`、`//IGNORE`、および `//TRANSLIT` のようなフラグも、コマンド行で使用できます。詳細は、`iconv_open(3C)` のマニュアルページを参照してください。

注記 - Oracle Solaris の `iconv` モジュールのいくつかは、`iconv_open(3C)` のマニュアルページで説明されているフラグのサブセットのみを実行する場合があります。

`iconv` の詳細は、`iconv(1)`、`iconv(3C)`、`iconv_open(3C)`、および関連するマニュアルページを参照してください。

Unicode 用の国際化コンポーネント

Oracle Solaris 11 では、使用可能なインタフェースに **Unicode 用の国際化コンポーネント (ICU)** の C/C++ ライブラリが追加されます。ICU は、ソフトウェアアプリケーション

ンに Unicode とグローバル化のサポートを提供する、成熟した、幅広く使用されているライブラリのセットです。ICU には移植性があり、すべてのプラットフォーム上で、および C/C++ と Java のソフトウェアの間で、同じ結果が得られます。

ICU が提供するサービスの一部に含まれるもの:

- コードページの変換 – Unicode や、ほとんどすべてのその他の文字セットまたはエンコーディングとの間で、テキストデータを変換します。
- 照合 – 特定の言語、地域、または国の慣例や標準に従って、文字列を比較します。
- 書式設定 – 選択されたロケールに応じて、数値、日付、時間、および通貨金額を書式設定します。
- 時間計算 – 複数の種類のカレンダーとタイムゾーン計算 API の完全なセットが提供されます。
- Unicode のサポート – ICU は、多くの Unicode 文字プロパティ、Unicode 正規化、ケースフォールディング、および Unicode 標準に規定されているその他の基本的な処理のすべてへのアクセスを容易にして、Unicode 標準を詳細に追跡します。
- 正規表現 – ICU の正規表現は、Unicode を完全にサポートすると同時に、非常に競争力のあるパフォーマンスを提供します。
- 双方向テキスト (Bidi) – 左から右および右から左のデータの混在を含むテキストの処理をサポートします。
- テキスト境界 – テキストの範囲内で単語、文、および段落の位置を見つけたり、テキストを表示するときの行の折り返しに適した位置を識別したりします。

Oracle Solaris 11 の ICU は 2 つのパッケージに分けられていて、`library/icu` にはライブラリのみが含まれていますが、`developer/icu` は、ヘッダーファイルと、`uconv(1)` のようないくつかのユーティリティを提供します。

詳細は、プロジェクトの Web サイト (<http://site.icu-project.org>) を参照してください

い。 `libcui18n(3LIB)`、`libicui18n(3LIB)`、`libicudata(3LIB)`、`libicule(3LIB)`、`libiculx(3LIB)`、`libicuuc(3LIB)` のマニュアルページに、Oracle Solaris でライブラリを使用する方法が記載されています。

uconv ユーティリティ

`iconv(1)` に加えて、Unicode 用の国際化コンポーネント (ICU) ツールセットの一部である `uconv(1)` コマンドも、1 つのエンコーディングから別のエンコーディングへの変換に使用できます。 `uconv` は、229 のエンコーディングと、1000 を超える別名をサポートしています。

このツールは、デフォルトではインストールされない `developer/icu` パッケージの一部です。これをインストールするには、次のコマンドを発行します。

```
# pkg install developer/icu
```

cp-1252 エンコーディングのテキストを UTF-8 に変換するには、次のように入力します。

```
$ uconv -f cp1252 -t UTF-8 -o file_in_utf8.txt file_in_cp1252_encoding.txt
```

uconv のもう 1 つの機能は、ベースとなる単語を変換せずに 1 文字ごとに文字を変換する、転写です。次の例では、ギリシア語のテキストの一部がラテン語の文字に変換されます。

```
$ echo "Σολαρίς"| uconv -x Greek-Latin -f utf-8 -t utf-8
Solaris
```

このツールの機能の詳細は、uconv(1) のマニュアルページを参照してください。

ファイル検査機能 (fsexam)

ファイルエンコーディング検査機能 fsexam ユーティリティでは、ファイルの名前、または標準テキストファイルの内容を、古い文字エンコーディングから UTF-8 エンコーディングに変換できます。fsexam ユーティリティには次の新しい機能が含まれています。

- エンコーディングリストのカスタマイズ
- エンコーディングの自動検出
- 予行演習、ログファイル、バッチ変換、ファイルのフィルタリング、シンボリックファイル、コマンド行、および圧縮ファイルのような特殊なファイルタイプのサポート

システムに fsexam を追加するには、storage/fsexam パッケージをインストールします。詳細は、fsexam(1) および fsexam(4) のマニュアルページを参照してください。

エンコーディング自動検出 (auto_ef)

Oracle Solaris には、auto_ef(1) という、ファイルのエンコーディングを識別するためのコマンド行ユーティリティが含まれています。auto_ef は、iconv コード変換を使用してエンコーディングを判定し、ファイルに対して特定のコード変換が成功したかどうかを判断します。これは、ファイル内に出現する文字シーケンスの頻度分析も行います。例:

```
$ auto_ef test_file
eucJP
```

-a オプションを使用すると、指定されたファイルに可能なすべてのエンコーディングが表示されます。

```
$ auto_ef -a test_file
eucJP          0.89
zh_CN.euc     0.40
ko_KR.euc     0.01
```

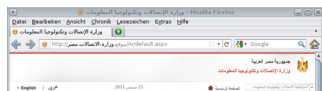
システムに `auto_ef` を追加するには、`text/auto_ef` パッケージをインストールします。詳細は、[auto_ef\(1\)](#) のマニュアルページを参照してください。

国際化ドメイン名のサポート

国際化ドメイン名 (IDN) によって、ホスト名およびドメイン名として英語以外のネイティブ言語を使用できます。英語以外のホスト名およびドメイン名を使用するには、RFC 5890 に規定されているリゾルバルーチンに名前を送信する前に、これらの名前を ASCII 互換エンコーディング (ACE) でエンコードされた名前に変換します。システム管理アプリケーションが IDN をサポートしていない場合、システム管理者はシステムのファイルおよびアプリケーションにも ACE 名を使用する必要があります。

詳細は、RFC 5890 「アプリケーションにおけるドメイン名の国際化 (IDNA)」を参照してください。

図 4 Firefox ブラウザでの IDN の例



Oracle Solaris 11 は、2 組の IDS 実装 (ライブラリと関連のユーティリティーを含む) を提供しており、それは GNU IDN ライブラリと JPRS (Japan Registry Services co., Ltd.) `idnkit-2` ライブラリです。いくつかのアプリケーションには、独自の IDN 実装を持つものもあります。たとえば、Firefox と Thunderbird には、Necko というネットワークプロトコルコンポーネントに独自の IDN サービスがあります。

GNU IDN ライブラリ

GNU-`libidn` は、GNU [Lesser General Public License \(LGPL\)](#) Version 2.1 以降に基づいてライセンス供与された、GNU プロジェクトです。GNU-`libidn` は、さまざまな GNU/Linux ディストリビューションで幅広く採用されています。したがって、デスクトップおよび `pidgin(1)` などの GNOME アプリケーションは、通常は IDN をサポートするために GNU-`libidn` を利用します。

`idn(1)` は、国際化ドメイン名のライブラリへのコマンド行インタフェースです。次の例は、UTF-8 のホスト名を ACE エンコーディングに変換します。結果の URL `http://xn-fsqu00a.xn--0zwm56d` は、このあと、ACE でエンコードされた `http://例子.测试` に相当するものとして使用できます。

```
$ idn --quiet -a 例子.测试
xn-fsqu00a.xn--0zwm56d
```

GNU-`libidn` は、`library/libidn` パッケージとしてインストールできます。オプションの詳細は、`idn(1)` のマニュアルページを参照してください。

注記 - Oracle Solaris 11 に付属する現行バージョンの GNU-`libidn` は、IDNA2003 標準のみをサポートします。

JPRS `idnkit-2` ライブラリ

`idnkit-2` ライブラリは、`idnkit-2` JPRS 公開ライセンスを持つオープンソースの IDN 実装です。専用の `idnkit-2` 変換ユーティリティー `idnconv(1)` では、各種オプションを使用した IDN 変換が可能です。変換の詳細を制御するためのオプションについての詳細は、`idnconv(1)` のマニュアルページを参照してください。

Oracle Solaris 11 は、`libidnkit (3)` の変換ルーチンを利用することによって、`iconv(3C)` インタフェースを介した IDN 変換もサポートしています。次の表に示されているように、`iconv(1)` ユーティリティーは、ACE および UTF-8 との間の変換にも使用できます。

IDNA2008 では 2 つの動作モード `lookup` および `registration` の用語が明示的に定義されているため、対応する `iconv` コード変換名の別名 `IDNA2008-LOOKUP (ACE-ALLOW-UNASSIGNED の別名)` および `IDNA2008-REGIST (ACE の別名)` も提供します。

表 16 `iconv` IDN コード変換

元のコード	変換後のコード
ACE または IDNA2008-REGIST	UTF-8
ACE-ALLOW-UNASSIGNED または IDNA2008-LOOKUP	UTF-8
UTF-8	ACE または IDNA2008-REGIST
UTF-8	ACE-ALLOW-UNASSIGNED または IDNA2008-LOOKUP

ACE および ACE-ALLOW-UNASSIGNED `iconv` コード変換名 (およびその別名) には次の意味があります。

- ACE または IDNA2008-REGIST

ACE は *fromcode* または *tocode* 名であり、RFC 5890 で定義されている ASCII 互換エンコーディングを参照するために、`iconv` コード変換で使用できます。この変換は STD3 ASCII のルールを使用します。割り当てられていない文字は許可されません。通常、ACE はホスト名またはドメイン名を格納または設定するために使用されます。

- ACE-ALLOW-UNASSIGNED または IDNA2008-LOOKUP

ACE-ALLOW-UNASSIGNED は、ACE-ALLOW-UNASSIGNED が割り当てられていない文字を許可する以外は、ACE と同じ処理を実行します。通常、ACE-ALLOW-UNASSIGNED は照会のために使用されます。

次の例は、`hostnames.txt` ファイルからの入力による、ACE から UTF-8 の変換を示します。出力は標準出力に送信されます。

```
$ iconv -f ACE -t UTF-8 hostnames.txt
```

`idnkit-2` ライブラリおよび `iconv` コード変換については、`libidnkit(3)` および `iconv_en_US.UTF-8(5)` のマニュアルページを参照してください。

印刷の拡張

Oracle Solaris 11 では、旧バージョンの LP 印刷サービスが Common UNIX Printing System (CUPS) に置き換えられています。

CUPS は、モジュール化されたオープンソースの印刷システムであり、プリンタ、印刷要求、および印刷キューを管理するための基準としてインターネット印刷プロトコル (IPP) を使用します。CUPS は、ネットワークプリンタのブラウザおよび PostScript Printer Description ベースの印刷オプションもサポートします。CUPS は、ローカルネットワークの共通印刷インタフェースも提供します。

CUPS およびその構成の詳細は、『[Oracle Solaris 11.3 での印刷の構成と管理](#)』の第 1 章、「[CUPS を使用したプリンタの設定と管理 \(概要\)](#)」を参照してください。

このセクションでは、`mp` コマンド行ユーティリティーを使用した印刷の拡張について説明します。

mp ユーティリティー

Oracle Solaris 環境は、UTF-8 で書き込まれたフラットテキストファイルを含む各種の入力ファイル形式を印刷できる、拡張された `mp` 印刷フィルタを提供します。この印刷フィルタは、Oracle Solaris システムで使用可能な TrueType および Type 1 の拡大縮

小可能なフォントと X11 ビットマップフォントを使用します。また、このフィルタはプリンタ常駐フォントも使用できます。

ユーティリティーからの出力は標準の PostScript であり、どの PostScript プリンタにも送信できます。mp ユーティリティーは、どのページ記述言語を出力することもできます。X Print Server クライアントとして構成されている場合、mp は印刷サーバーによってサポートされます。

このユーティリティーを使用するには、次のコマンドを入力します。

```
$ mp filename | lp
```

mp は stdin ストリームを受け入れるため、フィルタとして使用することもできます。

```
$ cat filename | mp | lp
```

mp ユーティリティーは print/mp パッケージから使用でき、デフォルトではインストールされていない可能性のある選択されたフォントパッケージに依存します。不足しているフォントに関するエラーメッセージが表示された場合は、pkg コマンドを使用して、正しいパッケージを検索してインストールできます。

- 不足しているフォントを含むパッケージを検索するには、次のコマンドを使用します。

```
# pkg search -rp TlwgTypist.ttf
PACKAGE
pkg:/system/font/truetype/thai-scalable@...
```

- 不足しているパッケージをインストールするには、次のコマンドを使用します。

```
# pkg install system/font/truetype/thai-scalable
```

詳細は、[mp\(1\)](#) のマニュアルページを参照してください。

ほかのプラットフォームとの相互運用性

以降のセクションでは、マルチプラットフォーム環境についての考慮事項を説明します。

NFS サーバーの考慮事項

NFS バージョン 4 のプロトコル (Oracle Solaris のデフォルト) は、UTF-8 を使用して、ファイル名とその他の文字列を処理します。そのため、ほとんどのユースケースで、

文字セットに関連する調整が必要になります。ただし、一部またはすべての NFS クライアントが指定された文字セットを使用している場合は、`charset` オプションを使用できることに注意してください。

たとえば、ISO8859-1 文字セットを使用して `/export` ディレクトリを共有するには、次のコマンドを使用します。

```
# share -o iso8859-1 /export
```

いくつかのシステムだけのために特定の文字セットを使用するディレクトリを共有するには、`charset=access_list` オプションを使用できます。

```
# share -o iso-8859-1=isosystem.example.com,koi8-r=koisystem.example.com /export
```

クライアントによって作成されたすべてのファイル名とパス名は、サーバーでは UTF-8 に変換されます。

詳細は、[share_nfs\(1M\)](#) のマニュアルページを参照してください。

ファイルシステムの考慮事項

[mount_pcfs\(1M\)](#) は MS-DOS コードページをサポートしていないため、MSDOS、古いバージョンの MS Windows、または Linux の「`msdos`」ドライバで作成された FAT ファイルシステム上では、ASCII 以外の文字は文字化けする場合があります。最新の FAT 実装は文字表示に Unicode を使用していて、読み書きの両方で、Oracle Solaris ではデフォルトで完全にサポートされています。

ASCII 以外のファイル名を含むアーカイブ

状況は改善しているとはいえ、特定のアーカイブ形式の多数の実装における ASCII 以外のファイル名のサポートにはかなりの違いがあるため、ファイル名に ASCII 以外の文字があるファイルをアーカイブすると問題が発生することがあります。

UNIX および Unix 系のシステムの最新の `tar` 実装は、POSIX.1-2001 に規定された POSIX 形式をサポートしているため、ASCII 以外のファイル名は問題なく処理されます。MS Windows プラットフォームでは、多数のアーカイブユーティリティーが現在のコードページを使用してファイル名を格納するため、そのようなアーカイブから抽出されたファイルの名前は文字化けする可能性があります。

その場合、コードページがわかっているときは、`convmv(1)` ツールを使用して修復できます。

```
$ convmv -f cp437 -t utf8 my_extracted_filename
```

Zip ファイルでは、元の仕様により、ファイル名とファイルのコメントのエンコーディングが IBM437 に設定されます。2007 年に PKWare は、UTF-8 も許容するように仕様を拡張しました。その一方で、さまざまな zip 実装が、現在のコードページをファイル名のエンコーディングとして使用する方針を採用しました (通常は MS Windows プラットフォーム上)。

Oracle Solaris 10 および Oracle Solaris 11 で使用される Info-ZIP の Zip 3.0 は、ファイル名を UTF-8 で格納するため、圧縮と圧縮解除のユーティリティーが両方もこのバージョンの場合は、アーカイブの中身が破壊されることはありません。

ファイル名を格納するために UTF-8 以外のエンコーディングを使用した zip アーカイブが Oracle Solaris で抽出されると、ファイル名が文字化けすることがあります。コードページがわかっている場合は、`convmv(1)` ツールを使用して修復できます。

```
$ convmv -f cp437 -t utf8 my-unzipped-filename
```

各国語プロパティの構成

`nlsadm` コマンドを使用して、各国語のプロパティを表示および設定できます。各国語のプロパティには、システムロケールやコンソールキーマップなどのシステムプロパティが含まれます。これらのプロパティはロケールに固有です。これらのプロパティを設定すると、さまざまなロケールでアプリケーションを適切に動作させることができます。`nlsadm` コマンドでは、統合された便利な方法で、各国語のプロパティを管理できます。

`nlsadm` コマンドでは、次のタスクを実行できます。

- システムコンソール、タイムゾーン、またはシステムのコンソールキーマップを設定する。
- システムコンソール、タイムゾーン、またはシステムのコンソールキーマップの値を一覧表示する。
- 地域、状態、言語、タイムゾーン、コードセット、修飾子などの具体的なロケール情報を一覧表示する。
- ロケールの地域および言語情報を設定する。
- すべてのタイムゾーン、コンソールキーマップ、インストールされているロケール、およびインストールに使用可能なロケールを IPS リポジトリから一覧表示する。

詳細は、[nlsadm\(1M\)](#) のマニュアルページを参照してください。

`nlsadm` コマンドは、`locale_description` ファイルを使用してロケール情報を取得します。`locale_description` ファイルには、各国語のプロパティが「キー = 値」形式で格納されています。新しいロケールを作成し、`nlsadm` でこの新しいロケールの

情報を表示する場合は、`locale_description` ファイルを指定する必要があります。詳細は、[locale_description\(4\)](#) のマニュアルページを参照してください。

注記 - `nlsadm` コマンドは、`locale_description` ファイルに依存しません。`locale_description` ファイルが存在しない場合は、`nlsadm` コマンドを実行すると、ロケールの名前または同様のロケールから情報を取得します。

カスタムロケールの作成

このセクションでは、[localedef\(1\)](#) を使用して、システムとともに提供された既存のロケールを基にしたカスタムロケールを作成する方法について説明します。

注記 - Oracle 提供のロケールのみがサポートされます。形式に誤りがあるロケールはエラーの原因になる可能性があります。

システムロケールを基にした新しいロケールの作成

既存のロケールをカスタマイズして、そこから新しいロケールを作成するには、少なくとも 3 つのロケール定義ソースファイルを準備する必要があります。

- `localedef` ソースファイル – ロケールに必要な定義が含まれます。
- `charmap` ソースファイル – コードポイント値と判読可能なシンボリック名との間のマッピングが含まれています。シンボリック名は、`localedef` ソースファイルで使用されます。`charmap` ソースファイルには、ロケールのコードセット名、ロケールのコードポイントで表すことができる最大バイト数などの、その他の定義も含まれます。
- `extension` ソースファイル – `strcoll(3C)` および `fgetwc(3C)` などの標準インタフェースの内部ロケール方式へのマッピングと、ロケールの適切な処理のために必要なその他の情報が含まれます。

システム提供のロケールには、それぞれのロケール定義データファイルがあり、`source/locale/localedef` パッケージから入手できます。これらは、既存のシステムロケールにわずかな変更だけが必要なときに、便利な場合があります。

▼ カスタムロケールを作成する方法

1. 次のコマンドを入力して、必要なパッケージをインストールします。

```
$ sudo pkg install system/header source/locale/localedef
```

- 必ず **PATH** に **Oracle Solaris Studio C コンパイラ**を指定します。

```
$ export PATH=<oracle-studio-path>/bin:$PATH
```

- 次のコマンドを使用して、ワークスペースを準備します。

```
$ mkdir mynewlocale
$ cd mynewlocale
$ mkdir amd64
$ cp /usr/lib/locale/common/methods_unicode.so.3 .
$ cp /usr/lib/locale/common/amd64/methods_unicode.so.3 amd64/
$ cp /usr/lib/localedef/src/charmaps/UTF-8.charmap \
/usr/lib/localedef/src/extensions/UTF-8.x \
/usr/lib/localedef/src/locales/fr_FR.UTF-8.src .
```

- 必要に応じて、**localedef**、**charmap**、または **extension** ソースファイルに変更を加えます。

- 64 ビットのロケールオブジェクトを構築します。

```
$ localedef -m lp64 -f UTF-8.charmap -x UTF-8.x -i fr_FR.UTF-8.src \
-L "-R\\$ORIGIN/../../common/amd64 -Bdirect -M /usr/lib/ld/map.pagealign \
-M /usr/lib/ld/map.noexdata" fr_FR.UTF-8@custom
$ mv fr_FR.UTF-8@custom.so.3 amd64/
```

- 32 ビットのロケールオブジェクトを構築します。

```
$ localedef -m ilp32 -f UTF-8.charmap -x UTF-8.x -i fr_FR.UTF-8.src \
-L "-R\\$ORIGIN/../../common" fr_FR.UTF-8@custom
```

- カスタムロケールをインストールします。

```
$ sudo mkdir -p /usr/lib/locale/fr_FR.UTF-8@custom/amd64
$ sudo cp fr_FR.UTF-8@custom.so.3 /usr/lib/locale/fr_FR.UTF-8@custom/
$ sudo cp amd64/fr_FR.UTF-8@custom.so.3 /usr/lib/locale/fr_FR.UTF-8@custom/amd64/
```

- このロケールを使用して起動するには、次のコマンドを入力します。

```
$ export LANG=fr_FR.UTF-8@custom
```

ゼロからのロケールの作成

ロケールをゼロから作成する必要は、めったにありません。ロケールをゼロから作成する場合も、[75 ページの「システムロケールを基にした新しいロケールの作成」](#)と同じアプローチを使用できます。**localedef**、**charmap**、および **extension** ソースファイルに使用可能なロケールとオプションについての詳細は、[localedef\(1\)](#)、[locale\(5\)](#)、[extensions\(5\)](#)、および [charmap\(5\)](#) のマニュアルページを参照してください。

使用可能なロケールおよびサポートされている文字セット

使用可能なロケールは、推奨ロケールと追加ロケールに分類できます。次の表は Oracle Solaris 11 で使用可能なロケールの概要を示し、該当する箇所では、サポートされる文字セットの詳細も示します。

- [表17](#) - この表では、デフォルトパッケージ `system/locale` で使用可能なロケールの一覧を示します。
- [表18](#) - この表では、オプションパッケージ `system/locale/extra` で使用可能な従来の (レガシー) ロケールの一覧を示します。

インストールされたロケールの管理の詳細は、[36 ページの「nlsadm を使用したロケールの追加または削除」](#)を参照してください。

表 17 推奨ロケール

ロケール	コードセット	説明
C	US-ASCII	C、POSIX
POSIX	US-ASCII	C、POSIX
af_ZA.UTF-8	Unicode 6.0	アフリカーンス語、南アフリカ
ar_AE.UTF-8	Unicode 6.0	アラビア語、アラブ首長国連邦
ar_BH.UTF-8	Unicode 6.0	アラビア語、バーレーン
ar_DZ.UTF-8	Unicode 6.0	アラビア語、アルジェリア
ar_EG.UTF-8	Unicode 6.0	アラビア語、エジプト
ar_IQ.UTF-8	Unicode 6.0	アラビア語、イラク
ar_JO.UTF-8	Unicode 6.0	アラビア語、ヨルダン
ar_KW.UTF-8	Unicode 6.0	アラビア語、クウェート
ar_LY.UTF-8	Unicode 6.0	アラビア語、リビア
ar_MA.UTF-8	Unicode 6.0	アラビア語、モロッコ
ar_OM.UTF-8	Unicode 6.0	アラビア語、オマーン
ar_QA.UTF-8	Unicode 6.0	アラビア語、カタール
ar_SA.UTF-8	Unicode 6.0	アラビア語、サウジアラビア
ar_TN.UTF-8	Unicode 6.0	アラビア語、チュニジア

ロケール	コードセット	説明
ar_YE.UTF-8	Unicode 6.0	アラビア語、イエメン
as_IN.UTF-8	Unicode 6.0	アッサム語、インド
az_AZ.UTF-8	Unicode 6.0	アゼルバイジャン語、アゼルバイジャン
be_BY.UTF-8	Unicode 6.0	ベラルーシ語、ベラルーシ
bg_BG.UTF-8	Unicode 6.0	ブルガリア語、ブルガリア
bn_IN.UTF-8	Unicode 6.0	ベンガル語、インド
bs_BA.UTF-8	Unicode 6.0	ボスニア語、ボスニア・ヘルツェゴビナ
ca_ES.UTF-8	Unicode 6.0	カタロニア語、スペイン
cs_CZ.UTF-8	Unicode 6.0	チェコ語、チェコ共和国
da_DK.UTF-8	Unicode 6.0	デンマーク語、デンマーク
de_AT.UTF-8	Unicode 6.0	ドイツ語、オーストリア
de_BE.UTF-8	Unicode 6.0	ドイツ語、ベルギー
de_CH.UTF-8	Unicode 6.0	ドイツ語、スイス
de_DE.UTF-8	Unicode 6.0	ドイツ語、ドイツ
de_LI.UTF-8	Unicode 6.0	ドイツ語、リヒテンシュタイン
de_LU.UTF-8	Unicode 6.0	ドイツ語、ルクセンブルグ
el_CY.UTF-8	Unicode 6.0	ギリシャ語、キプロス
el_GR.UTF-8	Unicode 6.0	ギリシャ語、ギリシャ
en_AU.UTF-8	Unicode 6.0	英語、オーストラリア
en_BW.UTF-8	Unicode 6.0	英語、ボツワナ
en_CA.UTF-8	Unicode 6.0	英語、カナダ
en_GB.UTF-8	Unicode 6.0	英語、英国
en_HK.UTF-8	Unicode 6.0	英語、香港 SAR 中国
en_IE.UTF-8	Unicode 6.0	英語、アイルランド
en_IN.UTF-8	Unicode 6.0	英語、インド
en_MT.UTF-8	Unicode 6.0	英語、マルタ
en_NZ.UTF-8	Unicode 6.0	英語、ニュージーランド
en_PH.UTF-8	Unicode 6.0	英語、フィリピン
en_SG.UTF-8	Unicode 6.0	英語、シンガポール
en_US.UTF-8	Unicode 6.0	英語、米国
en_ZW.UTF-8	Unicode 6.0	英語、ジンバブエ
es_AR.UTF-8	Unicode 6.0	スペイン語、アルゼンチン
es_BO.UTF-8	Unicode 6.0	スペイン語、ボリビア
es_CL.UTF-8	Unicode 6.0	スペイン語、チリ
es_CO.UTF-8	Unicode 6.0	スペイン語、コロンビア
es_CR.UTF-8	Unicode 6.0	スペイン語、コスタリカ
es_DO.UTF-8	Unicode 6.0	スペイン語、ドミニカ共和国

ロケール	コードセット	説明
es_EC.UTF-8	Unicode 6.0	スペイン語、エクアドル
es_ES.UTF-8	Unicode 6.0	スペイン語、スペイン
es_GT.UTF-8	Unicode 6.0	スペイン語、グアテマラ
es_HN.UTF-8	Unicode 6.0	スペイン語、ホンジュラス
es_MX.UTF-8	Unicode 6.0	スペイン語、メキシコ
es_NI.UTF-8	Unicode 6.0	スペイン語、ニカラグア
es_PA.UTF-8	Unicode 6.0	スペイン語、パナマ
es_PE.UTF-8	Unicode 6.0	スペイン語、ペルー
es_PR.UTF-8	Unicode 6.0	スペイン語、プエルトリコ
es_PY.UTF-8	Unicode 6.0	スペイン語、パラグアイ
es_SV.UTF-8	Unicode 6.0	スペイン語、エルサルバドル
es_US.UTF-8	Unicode 6.0	スペイン語、米国
es_UY.UTF-8	Unicode 6.0	スペイン語、ウルグアイ
es_VE.UTF-8	Unicode 6.0	スペイン語、ベネズエラ
et_EE.UTF-8	Unicode 6.0	エストニア語、エストニア
fi_FI.UTF-8	Unicode 6.0	フィンランド語、フィンランド
fr_BE.UTF-8	Unicode 6.0	フランス語、ベルギー
fr_CA.UTF-8	Unicode 6.0	フランス語、カナダ
fr_CH.UTF-8	Unicode 6.0	フランス語、スイス
fr_FR.UTF-8	Unicode 6.0	フランス語、フランス
fr_LU.UTF-8	Unicode 6.0	フランス語、ルクセンブルク
gu_IN.UTF-8	Unicode 6.0	グジャラート語、インド
he_IL.UTF-8	Unicode 6.0	ヘブライ語、イスラエル
hi_IN.UTF-8	Unicode 6.0	ヒンズー語、インド
hr_HR.UTF-8	Unicode 6.0	クロアチア語、クロアチア
hu_HU.UTF-8	Unicode 6.0	ハンガリー語、ハンガリー
hy_AM.UTF-8	Unicode 6.0	アルメニア語、アルメニア
id_ID.UTF-8	Unicode 6.0	インドネシア語、インドネシア
is_IS.UTF-8	Unicode 6.0	アイスランド語、アイスランド
it_CH.UTF-8	Unicode 6.0	イタリア語、スイス
it_IT.UTF-8	Unicode 6.0	イタリア語、イタリア
ja_JP.UTF-8	Unicode 6.0	日本語、日本 JIS X 0201:1976、 JIS X 0208:1990、JIS X 0212: 1990、JIS X 0213:2004、ベンダー 定義文字 (VDC)、ユーザー定義文 字 (UDC)
ka_GE.UTF-8	Unicode 6.0	ジョージア語、ジョージア
kk_KZ.UTF-8	Unicode 6.0	カザフ語、カザフスタン
kn_IN.UTF-8	Unicode 6.0	カナラ語、インド
ko_KR.UTF-8	Unicode 6.0	韓国語、韓国

ロケール	コードセット	説明
ks_IN.UTF-8	Unicode 6.0	カシミール語、インド
ku_TR.UTF-8	Unicode 6.0	クルド語、トルコ
ku_TR.UTF-8@sorani	Unicode 6.0	クルド語 (ソラニー)、トルコ
ky_KG.UTF-8	Unicode 6.0	キルギス語、キルギスタ
lt_LT.UTF-8	Unicode 6.0	リトアニア語、リトアニア
lv_LV.UTF-8	Unicode 6.0	ラトビア語、ラトビア
mk_MK.UTF-8	Unicode 6.0	マケドニア語、マケドニア
m1_IN.UTF-8	Unicode 6.0	マラヤーラム語、インド
mr_IN.UTF-8	Unicode 6.0	マラーティー語、インド
ms_MY.UTF-8	Unicode 6.0	マレー語、マレーシア
mt_MT.UTF-8	Unicode 6.0	マルタ語、マルタ
nb_NO.UTF-8	Unicode 6.0	ブークモール、ノルウェー
n1_BE.UTF-8	Unicode 6.0	オランダ語、ベルギー
n1_NL.UTF-8	Unicode 6.0	オランダ語、オランダ
nn_NO.UTF-8	Unicode 6.0	ニーノシュク、ノルウェー
or_IN.UTF-8	Unicode 6.0	オリヤー語、インド
pa_IN.UTF-8	Unicode 6.0	パンジャブ語、インド
pl_PL.UTF-8	Unicode 6.0	ポーランド語、ポーランド
pt_BR.UTF-8	Unicode 6.0	ポルトガル語、ブラジル
pt_PT.UTF-8	Unicode 6.0	ポルトガル語、ポルトガル
ro_RO.UTF-8	Unicode 6.0	ルーマニア語、ルーマニア
ru_RU.UTF-8	Unicode 6.0	ロシア語、ロシア
ru_UA.UTF-8	Unicode 6.0	ロシア語、ウクライナ
sa_IN.UTF-8	Unicode 6.0	サンスクリット語、インド
sk_SK.UTF-8	Unicode 6.0	スロバキア語、スロバキア
sl_SI.UTF-8	Unicode 6.0	スロベニア語、スロベニア
sq_AL.UTF-8	Unicode 6.0	アルバニア語、アルバニア
sr_ME.UTF-8	Unicode 6.0	セルビア語、モンテネグロ
sr_ME.UTF-8@latin	Unicode 6.0	セルビア語、モンテネグロ (ラテン文字)
sr_RS.UTF-8	Unicode 6.0	セルビア語、セルビア
sr_RS.UTF-8@latin	Unicode 6.0	セルビア語、セルビア (ラテン文字)
sv_SE.UTF-8	Unicode 6.0	スウェーデン語、スウェーデン
ta_IN.UTF-8	Unicode 6.0	タミル語、インド
te_IN.UTF-8	Unicode 6.0	テルグ語、インド
th_TH.UTF-8	Unicode 6.0	タイ語、タイ
tr_TR.UTF-8	Unicode 6.0	トルコ語、トルコ
uk_UA.UTF-8	Unicode 6.0	ウクライナ語、ウクライナ

ロケール	コードセット	説明
vi_VN.UTF-8	Unicode 6.0	ベトナム語、ベトナム
zh_CN.UTF-8	Unicode 6.0	簡体字中国語、中国
zh_HK.UTF-8	Unicode 6.0	繁体字中国語、香港 SAR 中国
zh_SG.UTF-8	Unicode 6.0	中国語、シンガポール
zh_TW.UTF-8	Unicode 6.0	繁体字中国語、台湾

表 18 追加ロケール

ロケール	コードセット	説明
ar_EG.IS08859-6	IS08859-6	アラビア語、エジプト
bg_BG.IS08859-5	IS08859-5	ブルガリア語、ブルガリア
bs_BA.IS08859-2	IS08859-2	ボスニア語、ボスニア・ヘルツェゴビナ
ca_ES.IS08859-1	IS08859-1	カタロニア語、スペイン
ca_ES.IS08859-15	IS08859-15	カタロニア語、スペイン
cs_CZ.IS08859-2	IS08859-2	チェコ語、チェコ共和国
cs_CZ.UTF-8@euro	Unicode 6.0	チェコ語、チェコ共和国(ユーロ)
da_DK.IS08859-1	IS08859-1	デンマーク語、デンマーク
da_DK.IS08859-15	IS08859-15	デンマーク語、デンマーク
da_DK.IS08859-15@euro	IS08859-15	デンマーク語、デンマーク(ユーロ)
de_AT.IS08859-1	IS08859-1	ドイツ語、オーストリア
de_AT.IS08859-15	IS08859-15	ドイツ語、オーストリア
de_CH.IS08859-1	IS08859-1	ドイツ語、スイス
de_DE.IS08859-1	IS08859-1	ドイツ語、ドイツ
de_DE.IS08859-15	IS08859-15	ドイツ語、ドイツ
el_GR.IS08859-7	IS08859-7	ギリシャ語、ギリシャ
en_AU.IS08859-1	IS08859-1	英語、オーストラリア
en_CA.IS08859-1	IS08859-1	英語、カナダ
en_GB.IS08859-1	IS08859-1	英語、英国
en_GB.IS08859-15	IS08859-15	英語、英国
en_GB.IS08859-15@euro	IS08859-15	英語、英国(ユーロ)
en_IE.IS08859-1	IS08859-1	英語、アイルランド
en_IE.IS08859-15	IS08859-15	英語、アイルランド
en_NZ.IS08859-1	IS08859-1	英語、ニュージーランド
en_US.IS08859-1	IS08859-1	英語、米国
en_US.IS08859-15	IS08859-15	英語、米国
en_US.IS08859-15@euro	IS08859-15	英語、米国(ユーロ)

ロケール	コードセット	説明
es_AR.IS08859-1	IS08859-1	スペイン語、アルゼンチン
es_BO.IS08859-1	IS08859-1	スペイン語、ボリビア
es_CL.IS08859-1	IS08859-1	スペイン語、チリ
es_CO.IS08859-1	IS08859-1	スペイン語、コロンビア
es_CR.IS08859-1	IS08859-1	スペイン語、コスタリカ
es_EC.IS08859-1	IS08859-1	スペイン語、エクアドル
es_ES.IS08859-1	IS08859-1	スペイン語、スペイン
es_ES.IS08859-15	IS08859-15	スペイン語、スペイン
es_GT.IS08859-1	IS08859-1	スペイン語、グアテマラ
es_MX.IS08859-1	IS08859-1	スペイン語、メキシコ
es_NI.IS08859-1	IS08859-1	スペイン語、ニカラグア
es_PA.IS08859-1	IS08859-1	スペイン語、パナマ
es_PE.IS08859-1	IS08859-1	スペイン語、ペルー
es_PY.IS08859-1	IS08859-1	スペイン語、パラグアイ
es_SV.IS08859-1	IS08859-1	スペイン語、エルサルバドル
es_UY.IS08859-1	IS08859-1	スペイン語、ウルグアイ
es_VE.IS08859-1	IS08859-1	スペイン語、ベネズエラ
et_EE.IS08859-15	IS08859-15	エストニア語、エストニア
fi_FI.IS08859-1	IS08859-1	フィンランド語、フィンランド
fi_FI.IS08859-15	IS08859-15	フィンランド語、フィンランド
fr_BE.IS08859-1	IS08859-1	フランス語、ベルギー
fr_BE.IS08859-15	IS08859-15	フランス語、ベルギー
fr_CA.IS08859-1	IS08859-1	フランス語、カナダ
fr_CH.IS08859-1	IS08859-1	フランス語、スイス
fr_FR.IS08859-1	IS08859-1	フランス語、フランス
fr_FR.IS08859-15	IS08859-15	フランス語、フランス
he_IL.IS08859-8	IS08859-8	ヘブライ語、イスラエル
hr_HR.IS08859-2	IS08859-2	クロアチア語、クロアチア
hu_HU.IS08859-2	IS08859-2	ハンガリー語、ハンガリー
is_IS.IS08859-1	IS08859-1	アイスランド語、アイスランド
it_IT.IS08859-1	IS08859-1	イタリア語、イタリア
it_IT.IS08859-15	IS08859-15	イタリア語、イタリア
ja_JP.PCK	PCK	日本語、日本 (PC 漢字コード、別名:Shift-JIS)
ja_JP.eucJP	EUC-JP	JIS X 0201:1976、JIS X 0208:1990、ベンダー定義文字 (VDC)、ユーザー定義文字 (UDC) 日本語 EUC 環境。「UI-OSF 日本語環境実装規約 Version 1.1」に準拠

ロケール	コードセット	説明
		JIS X 0201:1976、JIS X 0208:1990、JIS X 0212:1990 VDC、UDC
ja_JP.UTF-8@cldr	Unicode 6.0	日本語、日本 (CLDR)
		JIS X 0201:1976、JIS X 0208:1990、JIS X 0212:1990、JIS X 0213:2004、VDC、UDC
ko_KR.EUC	KS X 1001	韓国語、韓国
ko_KR.EUC@dict	KS X 1001	韓国語、韓国 (dict)
ko_KR.UTF-8@dict	Unicode 6.0	韓国語、韓国 (dict)
lt_LT.ISO8859-13	ISO8859-13	リトアニア語、リトアニア
lv_LV.ISO8859-13	ISO8859-13	ラトビア語、ラトビア
mk_MK.ISO8859-5	ISO8859-5	マケドニア語、マケドニア
nb_NO.ISO8859-1	ISO8859-1	ノルウェー語 (ブークモール)、ノルウェー
n1_BE.ISO8859-1	ISO8859-1	オランダ語、ベルギー
n1_BE.ISO8859-15	ISO8859-15	オランダ語、ベルギー
n1_NL.ISO8859-1	ISO8859-1	オランダ語、オランダ
n1_NL.ISO8859-15	ISO8859-15	オランダ語、オランダ
nn_NO.ISO8859-1	ISO8859-1	ノルウェー語 (ニーノシュク)、ノルウェー
p1_PL.ISO8859-2	ISO8859-2	ポーランド語、ポーランド
pt_BR.ISO8859-1	ISO8859-1	ポルトガル語、ブラジル
pt_PT.ISO8859-1	ISO8859-1	ポルトガル語、ポルトガル
pt_PT.ISO8859-15	ISO8859-15	ポルトガル語、ポルトガル
ro_RO.ISO8859-2	ISO8859-2	ルーマニア語、ルーマニア
ru_RU.ANSI1251	ANSI1251	ロシア語、ロシア
ru_RU.ISO8859-5	ISO8859-5	ロシア語、ロシア
ru_RU.KOI8-R	KOI8-R	ロシア語、ロシア
sk_SK.ISO8859-2	ISO8859-2	スロバキア語、スロバキア
s1_SI.ISO8859-2	ISO8859-2	スロベニア語、スロベニア
sq_AL.ISO8859-2	ISO8859-2	アルバニア語、アルバニア
sr_ME.ISO8859-5	ISO8859-5	セルビア語、モンテネグロ
sv_SE.ISO8859-1	ISO8859-1	スウェーデン語、スウェーデン
sv_SE.ISO8859-15	ISO8859-15	スウェーデン語、スウェーデン
sv_SE.ISO8859-15@euro	ISO8859-15	スウェーデン語、スウェーデン (ユーロ)
sv_SE.UTF-8@euro	Unicode 6.0	スウェーデン語、スウェーデン (ユーロ)
th_TH.TIS620	TIS-620	タイ語、タイ
tr_TR.ISO8859-9	ISO8859-9	トルコ語、トルコ

ロケール	コードセット	説明
zh_CN.EUC	GB2312	簡体字中国語、中国
zh_CN.EUC@pinyin	GB2312	簡体字中国語、中国 (pinyin)
zh_CN.EUC@radical	GB2312	簡体字中国語、中国 (radical)
zh_CN.EUC@stroke	GB2312	簡体字中国語、中国 (stroke)
zh_CN.GB18030	GB18030	簡体字中国語、中国
zh_CN.GB18030@pinyin	GB18030	簡体字中国語、中国 (pinyin)
zh_CN.GB18030@radical	GB18030	簡体字中国語、中国 (radical)
zh_CN.GB18030@stroke	GB18030	簡体字中国語、中国 (stroke)
zh_CN.GBK	GBK	簡体字中国語、中国
zh_CN.GBK@pinyin	GBK	簡体字中国語、中国 (pinyin)
zh_CN.GBK@radical	GBK	簡体字中国語、中国 (radical)
zh_CN.GBK@stroke	GBK	簡体字中国語、中国 (stroke)
zh_CN.UTF-8@pinyin	Unicode 6.0	簡体字中国語、中国 (pinyin)
zh_CN.UTF-8@radical	Unicode 6.0	簡体字中国語、中国 (radical)
zh_CN.UTF-8@stroke	Unicode 6.0	簡体字中国語、中国 (stroke)
zh_HK.BIG5HK	BIG5-HKSCS	繁体字中国語、香港 SAR 中国
zh_HK.BIG5HK@radical	BIG5-HKSCS	繁体字中国語、香港 SAR 中国 (radical)
zh_HK.BIG5HK@stroke	BIG5-HKSCS	繁体字中国語、香港 SAR 中国 (stroke)
zh_HK.UTF-8@radical	Unicode 6.0	繁体字中国語、香港 SAR 中国 (radical)
zh_HK.UTF-8@stroke	Unicode 6.0	繁体字中国語、香港 SAR 中国 (stroke)
zh_TW.BIG5	BIG5	繁体字中国語、台湾
zh_TW.BIG5@pinyin	BIG5	繁体字中国語、台湾 (pinyin)
zh_TW.BIG5@radical	BIG5	繁体字中国語、台湾 (radical)
zh_TW.BIG5@stroke	BIG5	繁体字中国語、台湾 (stroke)
zh_TW.BIG5@zhuyin	BIG5	繁体字中国語、台湾 (zhuyin)
zh_TW.EUC	CNS11643-1992	繁体字中国語、台湾
zh_TW.EUC@pinyin	CNS11643	繁体字中国語、台湾 (pinyin)
zh_TW.EUC@radical	CNS11643	繁体字中国語、台湾 (radical)
zh_TW.EUC@stroke	CNS11643	繁体字中国語、台湾 (stroke)
zh_TW.EUC@zhuyin	CNS11643	繁体字中国語、台湾 (zhuyin)
zh_TW.UTF-8@pinyin	UTF-8	繁体字中国語、台湾 (pinyin)
zh_TW.UTF-8@radical	UTF-8	繁体字中国語、台湾 (radical)
zh_TW.UTF-8@stroke	UTF-8	繁体字中国語、台湾 (stroke)
zh_TW.UTF-8@zhuyin	UTF-8	繁体字中国語、台湾 (zhuyin)

注記 - 以前の Oracle Solaris リリースでは、ja ロケール (古い Oracle Solaris リリースと互換) の仕様が ja_JP.eucJP ロケールと異なっていましたが、ja ロケールは現在、Oracle Solaris 11 で ja_JP.eucJP に対する別名ロケールです。

索引

あ

- アーカイブ
 - ASCII 以外, 73
- 新しいロケール
 - システムロケール, 75
- アルファベット, 19
- 移行
 - UTF-8, 26
 - UTF-8 以外, 26
 - UTF-8, 23
- 印刷
 - 拡張, 71
- インストーラ
 - ロケール選択, 29
- インストール
 - 追加
 - ロケール, 29
 - ロケール, 36
- インターネット
 - イントラネット, 53
- インド語派
 - フォント, 60
- イントラネット
 - インターネット, 53

か

- 概要
 - UTF-8, 24
- カスタム
 - ロケール, 75
- カスタムロケール
 - 作成, 65
- 仮想
 - キーボード, 52

- カタカナ
 - 文字, 19
- 韓国語
 - ハングル, 20
 - ハンジャ, 20
 - フォント, 62
- 漢字, 20
- 簡体字
 - 中国語, 63
- キーボード
 - GDM, 44
 - 配置, 57
 - 配列, 43, 44, 56
- キーボード配列
 - コンソール, 43
- キーボード配列の選択
 - GNOME ディスプレイマネージャー
 - GDM, 43
- 共通ロケールデータリポジトリ, 25
- 言語
 - エンジン, 50
 - 異なる, 30
 - 文字の違い, 18
- 言語サポート
 - コマンド行, 38
- 言語の単語
 - 文字の違い, 11
- コア
 - ロケール, 14
- 構成
 - 各国語
 - プロパティ, 65, 74
 - フォント, 59
- 構造
 - ファセット, 37

コードセット

変換, 65, 65

国際化

コンポーネント, 66

定義, 11

ローカリゼーション, 11

国際化ドメイン名

IDN

サポート, 65

コンソール

キーボード, 57

さ

作成

新しいロケール, 75

カスタム, 75

ロケール, 76

時間の書式, 16

自動

インストーラ

ロケール選択, 29

自動エンコーディング

検出機能, 68

数値

書式, 17

設定

キーボード配列, 57

キーマップ, 35

タイムゾーン, 35

デフォルトのロケール, 35

相互運用性

プラットフォーム, 65, 72

ソート

順序, 19

た

タイ語, 20

フォント, 64

タイムゾーン, 16

単語

区切り文字, 18

端末

ロケール, 30

中華人民共和国, 21

中国語

bopomofo, 21

zhuyin, 21

漢字, 20

台湾, 21

中華人民共和国, 20

ピンイン, 21

フォント, 63

香港, 21

調整

フォント構成, 59

通貨

書式, 18

ディレクトリ

名前, 26

デーバナーガリー

ヒンズー語, 21

動作

ロケール, 11, 15

ドメイン

名前, 69

な

夏時間 (DST), 16

日本語

テキスト, 19

フォント, 61

入力

方式, 43, 45, 53

入力方式, 43

は

ハングル, 20

ハンジャ, 20

繁体字

中国語, 63

日付の書式, 16

標準

テキスト, 26

ひらがな

- 文字, 19
- ピンイン, 21
- ヒンズー語
 - デーバナーガリー, 21
- ファイル
 - 検査機能, 68
 - システム, 73
 - 名前, 26
- フォント
 - 使用可能, 59, 60
- 複合
 - ロケール, 30
- ページサイズ
 - 一般的なサイズ, 22
- ヘブライ語
 - イディッシュ語, 21

ま

- 文字セット
 - ロケール, 23

ら

- ローカリゼーション
 - 定義, 11
- ロケール, 11
 - C, 13
 - GDM, 30
 - POSIX, 13
 - 移行, 26
 - カテゴリ, 14
 - キーボードの違い, 22
 - 旧バージョン, 40
 - コア, 14
 - 時間の書式, 16
 - 数値の書式, 17
 - ソートの順序, 19
 - 単語の区切り文字, 18
 - 追加, 36
 - 通貨の書式, 18
 - 定義, 12
 - 日付の書式, 16
 - ファセット, 37

- 文化上の表記規則, 15
- ページサイズ, 22
- 別名化, 40
- 文字セット, 19
- ロケール選択
 - Automated Installer, 34
 - GNOME ディスプレイマネージャー, 30
 - GDM, 29
 - インストーラ, 33
- ロケールの設定
 - 端末セッション, 29

A

- ATOK と Wnn
 - 言語エンジン, 53
- auto_ef, 68

C

- C
 - ロケール, 13
- cldr, 23
- Compose キー, 22

F

- fontconfig
 - ライブラリ, 59
- fsexam, 68

G

- GMT からの時差, 16
- GNOME
 - キーボード, 56
- GNU IDN
 - ライブラリ, 69

I

- IBus, 47

キーボード, 52
構成, 47
入力方式, 50
iconv
ユーティリティー, 65

J

JPNIC
idnkit
ライブラリ, 70

M

mp
ユーティリティー, 71

N

nfs, 27
NFS
考慮事項, 72
サーバー, 72

P

POSIX
ロケール, 13

U

uconv
ユーティリティー, 67
Unicode, 23, 66
概要, 23
UTC, 16
UTF-8
概要, 24
UTF-8 以外
文字セット, 25
ロケール, 35

X

X キーボード
拡張, 57
XKBC
IBus, 52

Z

zfs, 27