



# Simplified Chinese Solaris User's Guide

---

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

Part No: 816-0667-10  
May 2002

Copyright 2002 Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 U.S.A. All rights reserved.

This product or document is protected by copyright and distributed under licenses restricting its use, copying, distribution, and decompilation. No part of this product or document may be reproduced in any form by any means without prior written authorization of Sun and its licensors, if any. Third-party software, including font technology, is copyrighted and licensed from Sun suppliers.

Parts of the product may be derived from Berkeley BSD systems, licensed from the University of California. UNIX is a registered trademark in the U.S. and other countries, exclusively licensed through X/Open Company, Ltd.

Sun, Sun Microsystems, the Sun logo, docs.sun.com, AnswerBook, AnswerBook2, and Solaris are trademarks, registered trademarks, or service marks of Sun Microsystems, Inc. in the U.S. and other countries. All SPARC trademarks are used under license and are trademarks or registered trademarks of SPARC International, Inc. in the U.S. and other countries. Products bearing SPARC trademarks are based upon an architecture developed by Sun Microsystems, Inc.

The OPEN LOOK and Sun™ Graphical User Interface was developed by Sun Microsystems, Inc. for its users and licensees. Sun acknowledges the pioneering efforts of Xerox in researching and developing the concept of visual or graphical user interfaces for the computer industry. Sun holds a non-exclusive license from Xerox to the Xerox Graphical User Interface, which license also covers Sun's licensees who implement OPEN LOOK GUIs and otherwise comply with Sun's written license agreements.

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

DOCUMENTATION IS PROVIDED "AS IS" AND ALL EXPRESS OR IMPLIED CONDITIONS, REPRESENTATIONS AND WARRANTIES, INCLUDING ANY IMPLIED WARRANTY OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NON-INFRINGEMENT, ARE DISCLAIMED, EXCEPT TO THE EXTENT THAT SUCH DISCLAIMERS ARE HELD TO BE LEGALLY INVALID.

---

Copyright 2002 Sun Microsystems, Inc. 4150 Network Circle, Santa Clara, CA 95054 U.S.A. Tous droits réservés

Ce produit ou document est protégé par un copyright et distribué avec des licences qui en restreignent l'utilisation, la copie, la distribution, et la décompilation. Aucune partie de ce produit ou document ne peut être reproduite sous aucune forme, par quelque moyen que ce soit, sans l'autorisation préalable et écrite de Sun et de ses bailleurs de licence, s'il y en a. Le logiciel détenu par des tiers, et qui comprend la technologie relative aux polices de caractères, est protégé par un copyright et licencié par des fournisseurs de Sun.

Des parties de ce produit pourront être dérivées du système Berkeley BSD licenciés par l'Université de Californie. UNIX est une marque déposée aux Etats-Unis et dans d'autres pays et licenciée exclusivement par X/Open Company, Ltd.

Sun, Sun Microsystems, le logo Sun, docs.sun.com, AnswerBook, AnswerBook2, et Solaris sont des marques de fabrique ou des marques déposées, ou marques de service, de Sun Microsystems, Inc. aux Etats-Unis et dans d'autres pays. Toutes les marques SPARC sont utilisées sous licence et sont des marques de fabrique ou des marques déposées de SPARC International, Inc. aux Etats-Unis et dans d'autres pays. Les produits portant les marques SPARC sont basés sur une architecture développée par Sun Microsystems, Inc.

L'interface d'utilisation graphique OPEN LOOK et Sun™ a été développée par Sun Microsystems, Inc. pour ses utilisateurs et licenciés. Sun reconnaît les efforts de pionniers de Xerox pour la recherche et le développement du concept des interfaces d'utilisation visuelle ou graphique pour l'industrie de l'informatique. Sun détient une licence non exclusive de Xerox sur l'interface d'utilisation graphique Xerox, cette licence couvrant également les licenciés de Sun qui mettent en place l'interface d'utilisation graphique OPEN LOOK et qui en outre se conforment aux licences écrites de Sun.

CETTE PUBLICATION EST FOURNIE "EN L'ETAT" ET AUCUNE GARANTIE, EXPRESSE OU IMPLICITE, N'EST ACCORDEE, Y COMPRIS DES GARANTIES CONCERNANT LA VALEUR MARCHANDE, L'APTITUDE DE LA PUBLICATION A REPOUDRE A UNE UTILISATION PARTICULIERE, OU LE FAIT QU'ELLE NE SOIT PAS CONTREFAISANTE DE PRODUIT DE TIERS. CE DENI DE GARANTIE NE S'APPLIQUERAIT PAS, DANS LA MESURE OU IL SERAIT TENU JURIDIQUEMENT NUL ET NON AVENU.



020207@3062



# Contents

---

<b>Preface</b>	<b>7</b>
<b>1 Introduction</b>	<b>11</b>
Design of the Simplified Chinese Solaris Software	11
Simplified Chinese Graphical User Interface	12
Simplified Chinese Input/Output	12
The Localized CDE Desktop	12
<b>2 Starting the Simplified Chinese Solaris Software</b>	<b>15</b>
Introduction	15
Setting the Default Locale	16
Using Aliases to Set Locales	16
Changing Font Directories	17
<b>3 Using the <code>htt</code> Input Method Server</b>	<b>19</b>
Introduction	19
Input Method Server Basic Properties	19
Three <code>htt</code> Processes	20
<code>htt</code> Application Display	20
How and When <code>htt</code> Is Started	21
Using the <code>iiim</code> Server in <code>zh.UTF-8</code> and <code>zh_CN.UTF-8</code> Locales	21
<b>4 Entering Simplified Chinese Text</b>	<b>23</b>
About This Chapter	23

Input Window Areas	23
Preedit Area	24
Status Area	24
Lookup Choice Area	24
Auxiliary Window	24
Input Method Utilities	25
▼ Selecting the Utility Menu	25
Input Method SelectionTool	25
Setting Input Method Options	27
Lookup table	28
Virtual Keyboards	31
User Defined Characters (UDC)	36
Input Method Help	37
Input Methods and Conversion Modes for Entering Text	38
Typing ASCII Text	38
Switching Between English Status and Chinese Status:	39
Select Input Method	39
Toggling Input Methods	40
Switching Between Half_width Character Mode and Full_width Character Mode	40
Switching Between Chinese Punctuation Mode and English Punctuation Mode:	41
Language Input Methods	43
New QuanPin and New ShuangPin Input Methods	44
▼ Defining Phrases for Later Use	44
▼ Selecting Frequently-Used Candidates	47
Other Features	49
Keyboard Definition	51
Dictionary Files	52
New ShuangPin Features	53
New ShuangPin Keyboard Mapping	53
GBK Code Input Method	56
Typing GBK Code Text	57
GB2312 Code Input Method	59
Typing GB2312 Code Text	59
GB18030 Code Input Method	61
Typing GB18030 Code Text	61
QuanPin Input Method	64

	Typing QuanPin Text	64
	English_Chinese Input Method	66
	Typing English_Chinese Text	66
<b>5</b>	<b>Code Table Input Method Interface</b>	<b>71</b>
	Introduction	71
	▼ Creating a Code Table	71
	▼ Converting the Source Codetable File from Text Format to Binary Format or from Binary Format to Text Format:	74
	▼ Creating a new codetable input method	74
<b>6</b>	<b>Localized Applications</b>	<b>77</b>
	About This Chapter	77
	Using the mailx Utility With Simplified Chinese Characters	77
	Using talk With Simplified Chinese Characters	78
	The xtobdf Utility	78
	stdconvtool	78
	iconv	79
<b>7</b>	<b>Fonts</b>	<b>83</b>
	About This Chapter	83
	Display PostScript System (DPS)	83
	Using Simplified Chinese PostScript Fonts and DPS Facilities	83
	Converting BDF to PCF Format	87
	Installing and Checking the Edited Font	88
	Creating Characters With sdtudctool	88
	▼ Using sdtudctool	89
	▼ Specifying Options	90
	▼ Viewing the Character Reference	91
	▼ Register the UDC to NewPinYin input method	93
	▼ Typing the New Characters	94
	▼ Printing the New Characters	95
<b>8</b>	<b>Simplified Chinese Printing Facilities</b>	<b>97</b>
	Introduction	97
	Printing Chinese Output From a Command Line	98

Printing With a Line Printer	98
Using <code>xetops</code> and <code>xutops</code> Utilities	98
Printing with the <code>mp</code> Utility	100

**A Backward Compatibility 103**

Simplified Chinese Test Utilities	103
Simplified Chinese Conversion Utilities	105
Conversion Utilities	105
Conversion for Simplified Chinese Character Codes	106
Binary Compatibility Package (BCP)	107
Running Earlier Binary Code	107

**Glossary 109**

**Index 113**

# Preface

---

---

## About This Book

*Simplified Chinese Solaris User's Guide* describes product behavior unique to the Simplified Chinese Solaris™ operating environment and answers many questions commonly asked about the software. This guide introduces you to the general appearance and properties of a variety of localized Desktop Tools™ and utilities offered with the Simplified Chinese Common Desktop Environment (CDE).

---

## Who Should Use This Book

This user's guide is for someone who wants to use the Simplified Chinese features in the Solaris operating environment to manage files, calendar, and e-mail, write or print Simplified Chinese files, and so forth. Tools for these and many other applications run under Simplified Chinese Solaris software. This guide helps you easily find, access, and get started with these tools. You should read this guide:

- If you have not used Simplified Chinese Solaris software before
- If you need information on using product features
- If you need instructions for starting the Simplified Chinese Solaris operating environment (see especially Chapter 2)

---

## Before You Read This Book

Before you read this book, become familiar with the basics of the Solaris base release user documents, particularly the ones listed under “Related Books” on page 9. This user’s guide focuses on using the Simplified Chinese features of the Desktop Tools and other features of Simplified Chinese Solaris software.

---

## How This Book Is Organized

Each chapter of this guide addresses a different aspect of using Simplified Chinese Solaris software. The chapters tell how to check your set up before you begin using the facilities of Simplified Chinese Solaris software and give step-by-step instructions for using Simplified Chinese facilities.

Chapter 1 briefly describes general modifications made to Solaris software, including CDE, to internationalize and localize it for Simplified Chinese.

Chapter 2 gives the step-by-step instructions you must follow to start your Solaris user environment. It also describes Simplified Chinese Solaris-specific features you must use to turn Simplified Chinese facilities OFF/ON by using `dtlogin`.

Chapter 3 introduces the startup, appearance, and use of `http`.

Chapter 4 describes different Simplified Chinese character entry modes and provides a step-by-step tutorial in their use. (Further information on customizing commands and other advanced user topics are covered in *International Language Environments Guide* and *Simplified Chinese Solaris System Administrator’s Guide*.)

Chapter 5 describes the interface that allows the user to add new input.

Chapter 6 describes uses of two desktop tools localized for Chinese `mailx`, `talk`, and tools to convert file codes.

Chapter 7 explains how to customize fonts used in your Simplified Chinese Solaris applications.

Chapter 8 discusses Simplified Chinese Solaris support for line printers with built-in Chinese fonts or using `xetops`, `xutops`, or `mp` print filters for a PostScript printer.

The Glossary contains a list of words and phrases found in the Simplified Chinese Solaris documentation set, and their definitions.



---

## Related Books

You should become familiar with the following basic documentation:

- *Solaris Common Desktop Environment: User's Guide*

Advanced users may want to read *Solaris Advanced User's Guide*. Advanced users who want to customize their system environment or the operations of their Sun tools will find pertinent information in *International Language Environments Guide* and *Simplified Chinese Solaris System Administrator's Guide*. These books contain information on setting up, administering, programming, and customizing product features for advanced users, developers/programmers, and system administrators.

---

## What Typographic Changes Mean

The following table describes the typographic changes used in this book.

**TABLE P-1** Typographic Conventions

Typeface or Symbol	Meaning	Example
AaBbCc123	The names of commands, files, and directories; on-screen computer output	Edit your <code>.login</code> file. Use <code>ls -a</code> to list all files. <code>machine_name% You have mail.</code>
<b>AaBbCc123</b>	What you type, contrasted with on-screen computer output	<code>machine_name% su</code> Password:
<i>AaBbCc123</i>	Command-line placeholder: replace with a real name or value	To delete a file, type <code>rm filename</code> .
<i>AaBbCc123</i>	Book titles, new words or terms, or words to be emphasized	Read Chapter 6 in <i>User's Guide</i> . These are called <i>class</i> options. You <i>must</i> be root to do this.

---

## Shell Prompts in Command Examples

The following table shows the default system prompt and superuser prompt for the C shell, Bourne shell, and Korn shell.

**TABLE P-2** Shell Prompts

<b>Shell</b>	<b>Prompt</b>
C shell prompt	machine_name%
C shell superuser prompt	machine_name#
Bourne shell and Korn shell prompt	\$
Bourne shell and Korn shell superuser prompt	#

# Introduction

---

---

## Design of the Simplified Chinese Solaris Software

This software is a Simplified Chinese localization of the Sun™ Solaris operating environment. The Simplified Chinese Solaris software includes the Simplified Chinese Common Desktop Environment (CDE) for windowed applications that are built on Sun's Solaris operating environment.

Simplified Chinese Solaris software is an extension of base Solaris software. Virtually all utilities and features of the U.S. and International Solaris standard releases are incorporated in Simplified Chinese Solaris software. These products introduce Solaris input methods for the input and output of Simplified Chinese. Application programs and CDE use the features of Simplified Chinese CDE to communicate with users in Simplified Chinese.

This Simplified Chinese localization of Sun's internationalized CDE includes enhancements for handling appropriate linguistic and cultural conventions, which it provides to two broad working environments:

- A localized user environment, which includes localized desktop tools and window manager (`dtwm`) that communicate with users in Simplified Chinese.
- A localized development environment, which programmers use to develop localized applications, with Xlib and Motif, which have been internationalized for this use. Programmer and developers should refer to *International Language Environments Guide*.

---

## Simplified Chinese Graphical User Interface

This Simplified Chinese Solaris release uses the Simplified Chinese CDE Motif graphical user interface, which is similar in layout and design to the U.S. release of CDE. Simplified Chinese CDE supports multibyte characters and Simplified Chinese messages with Motif objects. Differences in character width and proportional spacing cause minor differences in the exact layout of some Motif objects.

All application windows that accept Simplified Chinese input include a *status area* associated with their input window to show the current conversion mode. With an input conversion mode on, as Simplified Chinese is being typed its entry point becomes a highlighted (reverse video) *preedit area* until the input is converted to Simplified Chinese or special characters and committed. Some input modes also provide conversion choices (*candidates*) among several Simplified Chinese characters on menus.

## Simplified Chinese Input/Output

To accommodate the diversity of Simplified Chinese, this Solaris software provides several different input methods for entering Simplified Chinese characters. With these methods you can type ASCII/English characters, Simplified Chinese radicals, and Simplified Chinese characters using an ASCII keyboard.

Simplified Chinese input from the keyboard is stored temporarily in an intermediate representation. The conversion manager program, with the help of user interaction, then transforms this intermediate representation into a displayed character string.

---

## The Localized CDE Desktop

The following desktop tools are provided in this Solaris release. Each tool can accept and display Simplified Chinese text. A manual page is provided for each.

- Address Manager – Carries out remote operations and finds information about the systems and users on your network. Can speed up such tasks as sending email, logging in remotely, and setting appointments on someone else's calendar.
- App Manager – Contains the applications available on your system. You can launch these applications through the App Manager interface.

- Audio Tool – Tool for recording, playing, editing, and controlling workstation audio parameters.
- Calculator – Visual calculator for use with mouse or keyboard.
- Calendar Manager – Manages business and social appointments; can use electronic mail to send automatic reminders.
- Clock – Displays current analog or digital time.
- Console – Standard Motif scrolling window terminal emulator.
- File Finder – Tool for searching for a folder and subfolders that returns a list of files or folders that match your search criteria. You can also specify the size, owner, date modified, type, and permissions in your search criteria.
- File Manager – Graphical tool for accessing files and directories. Represents file types with varying colors and icons. Navigates through the file system with the mouse.
- Help – On-line searchable help for CDE.
- Icon Editor – Visual tool for editing icon appearance and creating new icons.
- Image Tool – Interactive image viewer. Image Tool can be used to view the contents of file types such as GIF, TIFF, JPEG, PostScript, and others.
- Mailer – Tool for handling electronic mail.
- Performance Meter – Real-time system performance meter that can display a variety of data.
- Print Manager – Graphical front-end to the print command. It supports *drag-and-drop* file transfer operations.
- Process Manager – Tool for displaying and performing actions on the processes that are currently running on your workstation.
- Snapshot – Tool to *snap* or capture picture of a window or region of a screen in a bitmap (raster file). Used for capturing screen image displays in this user's guide.
- Text Editor – Visual text editor used in CDE tools such as the Mailer composition window.
- Style Manager – Tool for setting workstation preferences, such as audio feedback from keyboard, mouse response, and so on.
- Terminal – Standard Motif window terminal emulator. The window behaves like an ASCII character terminal for entry of UNIX<sup>®</sup> commands at a system shell prompt and other terminal operations.



# Starting the Simplified Chinese Solaris Software

---

---

## Introduction

The Simplified Chinese Solaris operating environment must be set up for using Simplified Chinese text facilities. This chapter describes the steps required to set up the Simplified Chinese environment and to start Simplified Chinese Solaris operation.

The Simplified Chinese Solaris product includes the following locales:

- `zh.GBK` – Simplified Chinese environment GBK, an extension of GB2312-80 ( that is Guo Biao Kuo in Chinese PinYin, which supports all CJK characters that are in Unicode 2.0).
- `zh` — Simplified Chinese environment in extended UNIX code (EUC).
- `zh_CN.EUC` — Symbolic link to `zh` locale.
- `zh_CN.GB18030` — Simplified Chinese environment in GB18030-2000. This standard is a superset of GBK and supports CJK Unified Ideographs Extension A and Yi, Mongolian, Tibetan and Uigur minority scripts in Unicode 3.1.
- `zh_CN.GBK` — Symbolic link to `zh.GBK` locale.
- `zh.UTF-8` — Simplified Chinese environment in Unicode 3.1.
- `zh_CN.UTF-8` — Symbolic link to `zh.UTF-8` locale.





## ▼ Setting Up Locale Aliases for the C Shell

- Put aliases like the following in your `.cshrc` file.

```
alias chn_EUC "setenv LANG zh; /bin/stty cs8 -istrip defeucw"
alias chn_GBK "setenv LANG zh.GBK; /bin/stty cs8 -istrip defeucw"
alias asc "setenv LANG C; /bin/stty cs7 istrip defeucw"
alias chn_UTF-8 "setenv LANG=zh.UTF-8; /bin/stty cs8 -istrip defeucw"
```

---

## Changing Font Directories

The Xsession script (in `/usr/dt/bin`) that comes with the Simplified Chinese operating environment includes the following font path:  
`/usr/openwin/lib/locale/zh/X11/fonts`. To add a different font directory path dynamically, type:

```
system% xset +fp font_directory-path
system% xset fp rehash
```



## Using the `htt` Input Method Server

---

---

### Introduction

The input method server (IM Server) handles Simplified Chinese input for Simplified Chinese Solaris software. The IM Server receives keyboard input and makes the input available as Simplified Chinese characters to Simplified Chinese Solaris system applications. It can serve any internationalized X Window application that uses X Window Input Method (XIM) application program interfaces (API) to receive language input.

This chapter explains the basic display features of the IM Server.

You need to be familiar with the input method terminology in the chapter 4 before reading this chapter. However, to create Simplified Chinese text using Solaris applications, you do not need to read beyond the first section of this chapter.

---

### Input Method Server Basic Properties

The Simplified Chinese Solaris operation environment starts the IM server automatically when you start the solaris operating environment in a locale that requires an input method. The IM server continues to run and service applications that are started and connected to it.

Simplified Chinese Solaris applications that use IM Server for Simplified Chinese character input typically find the IM server running when they started. To get the IM server's service for the language input, the IM Server should be running before an application is started. If the IM Server is not running when an application is started, the application may not be able to get the input service, even if IM Server is started later.

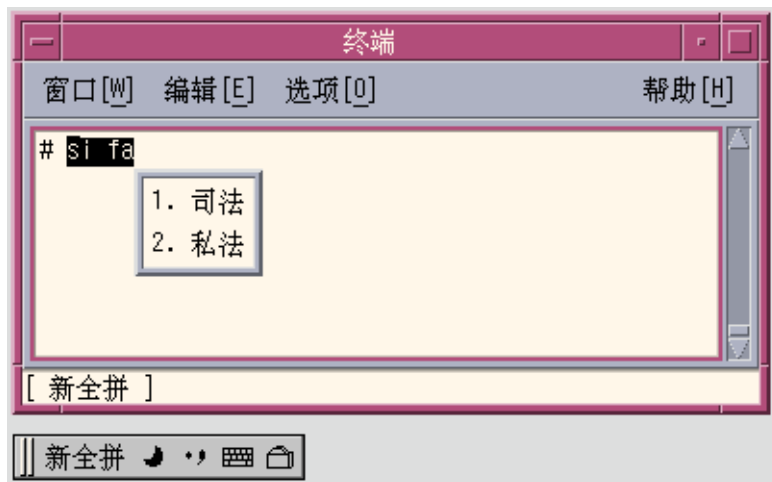
## Three `htt` Processes

The input method server comprises three related programs (`htt`, `htt_server`, and `htt_xbe`) processes running together. One process controls the input method server properties, another controls the population of the input method server (that is, a "watch dog" process), and the third handles input methods of clients.

Usually, you don't need to know the details of these three processes. When `htt` is started, `htt_server` and `htt_xbe` start also.

## `htt` Application Display

The following figure shows a typical application display when using the `htt` input method.



---

## How and When `htt` Is Started

Any application that uses the `htt` input method server to receive Simplified Chinese characters must find `htt` running before it can receive input. If your default locale is the `C` locale (which does not require `htt`) and you then change from `C` to another locale in a single Terminal, for example, you must start `htt` with a command line like the following:

```
system% /usr/openwin/bin/htt -if cle -xim htt_xbe &
```

With the locale set to `zh`, `htt` starts automatically when the Simplified Chinese windowing environment is started. In CDE, `htt` is started from a script, `/usr/dt/config/Xsession.d/0020.dtims`. This script is executed by `Xsession`, which runs at every initialization stage when a user logs in from `dtlogin`. This script ensures that `htt` is started before other applications in CDE.

---

## Using the `iiim` Server in `zh.UTF-8` and `zh_CN.UTF-8` Locales

The Internet Intranet Input Method Server (`iiim`) supports both European Local Input Methods using Compose Key and Asian Remote Input Methods using IIIM Protocol. The default Language Engine is English/European which supports ASCII and some European Languages such as German and French.

To switch to other Language Engines, press CTRL+Space, as with other Asian Locales.

To switch to Chinese Input Methods, click the left mouse button in the Status Area. The following Language Engine List appears:

- Cyrillic
- Greek
- Thai
- Arabic
- Hebrew
- Hindi
- Unicode Hex
- Unicode Octal
- Lookup
- Simplified Chinese

The following Language Engines appear if their corresponding locales are installed:

- Japanese — if `ja` locale is installed.

- Korean— if ko locale is installed.
- Traditional Chinese — if zh\_TW locale is installed.
- Traditional Chinese (Hong Kong)— if zh\_HK locale is installed.

After the Simplified Chinese Language Engine is selected, you can use Function Keys to switch between different Input Methods:

- F2 — New Pinyin
- F3 — New Shuang Pin
- F4 — QuanPin
- F5 — ShuangPin
- F6 — English\_to\_Chinese

For more detailed information about Input Methods, see “Chapter 4, Entering Simplified Text”.

## Entering Simplified Chinese Text

---

---

### About This Chapter

This chapter describes the Simplified Chinese Solaris input modes for typing Simplified Chinese characters with Simplified Chinese Solaris software. You can type any of the following kinds of characters:

- ASCII/English characters
- Simplified Chinese ideograms (characters or symbols representing ideas/actions/things)
- Special symbol characters

You can type all of these characters in the input areas of the following application windows:

- In terminal emulation (TTY) windows such as Terminal
- In text entry subwindows such as used by Text Editor or Mailer
- In control panel subwindows such as used by File Manager for typing a file name
- In other special use (sometimes pop-up) subwindows

For information about creating your own input method, see Chapter 5.

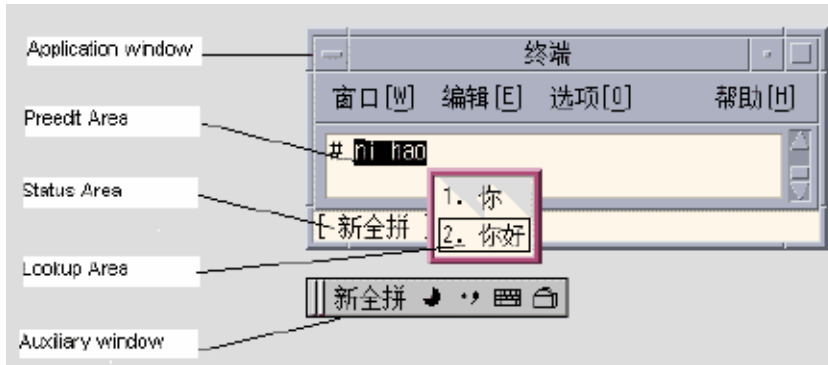
---

### Input Window Areas

Three separate areas of an application subwindow are involved in entering characters. These areas are typically displayed, named, and used as follows:

- Predit area—Highlighted (such as inverse video or underlined) entry display area

- Status area—Area indicating the current input/conversion mode
- Lookup choice area—Area displaying multiple character choices
- Auxiliary Windows— Utilities for input method management



## Preedit Area

The highlighted (for example, reverse video and underlined) preedit area displays characters as they are typed or converted. It holds formations of text before they are converted to Simplified Chinese characters or symbols and put in the text block being assembled for the application.

## Status Area

The status area is shows what input conversion mode is in effect. In the above example, it is located in the lower left corner of the window margin.

## Lookup Choice Area

The lookup choice area displays multiple Simplified Chinese or special character choices available for conversion of the character(s)/radical(s) in the preedit area. In the above example, it is a pop-up.

## Auxiliary Window

The auxiliary window provides tools and utilities to manage the input methods or to make the input simpler.



---

## Input Method Utilities

Solaris 9 provides graphics interface tools and utilities to manage input methods, set the properties of input methods, and to facilitate the input of special characters.

The following tools are supported:

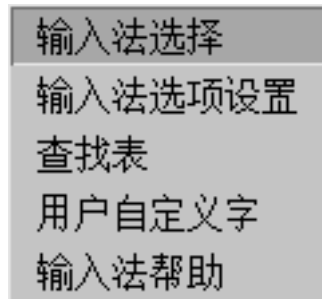
- Input method selection
- Input method option setting
- Lookup table .
- User define character
- Input method help
- Virtual Keyboard.

### ▼ Selecting the Utility Menu

- Click the utility button



to display the utilities menu.



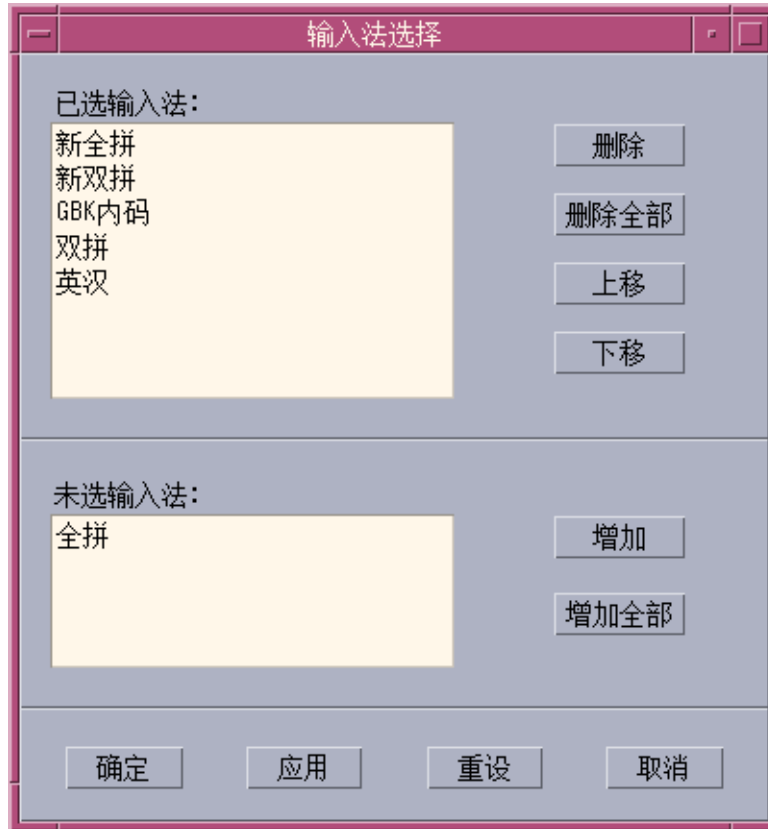
Select one of the input method tools from the menu.

### Input Method Selection Tool

The input method selection tool allows you to select a list of input method. You can also set the default input method and the sequence of the input methods.

- **Select input method selection item of the utility menu**

The input method selection panel appears as below:

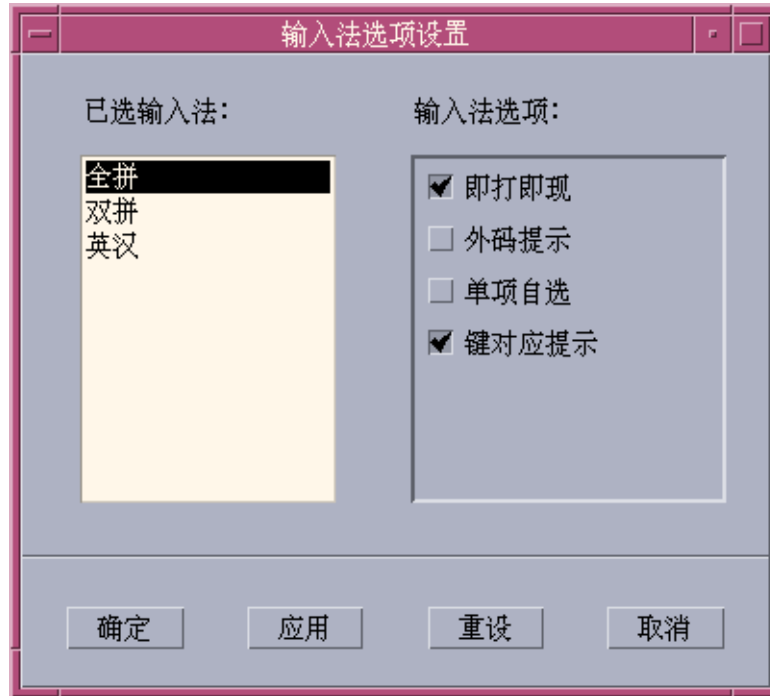


After selecting an input method, click "OK" or "Apply", the setting will be activated. The first input method selected becomes the default input method.

Press "CTRL+Space" in the application window to activate Chinese input, the default input method will be selected. Press "F2" to switch to the first input method selected. Press "F3" to switch to the second one, and so on.

## Setting Input Method Options

The properties of Simplified Chinese input methods can be set from the input method options setting screen, which appears as follows:



With this options setting tool, user can set the options of input methods. After setting the options in this panel, then click "OK" or "Apply", the setting will be activated.

For the input methods that based on code table, there are 4 options can be set by user:

1. "Display candidates key by key"  
if TRUE: when a valid key is entered for this input method, IM will search the dictionary table and display the candidates in Lookup window.  
if FALSE: when a valid key is entered for this input method, IM does not search the dictionary table, but displays the key in the preedit area. The "SPACE" key must be pressed for IM to begin to search the dictionary table, and display the candidates.
2. "Display external codes":  
if TRUE: in every Lookup window, the external codes for every candidate will follow after the candidate.

if FALSE: the external codes for every candidate will not follow after that candidate.

This option provides a way to study the input method and view the external code of a Chinese character in that input method.

3. "Automatically commit if only one candidate":

if TRUE: if there is only one candidate for the external code, IM will automatically commit it.

if FALSE: IM will display it in the Lookup window.

4. "Display keymap character for every external code"

if TRUE: when a valid key is entered, the corresponding keymap character of the key will display in Preedit area,

if FALSE: only the key is displayed without the keymap character,

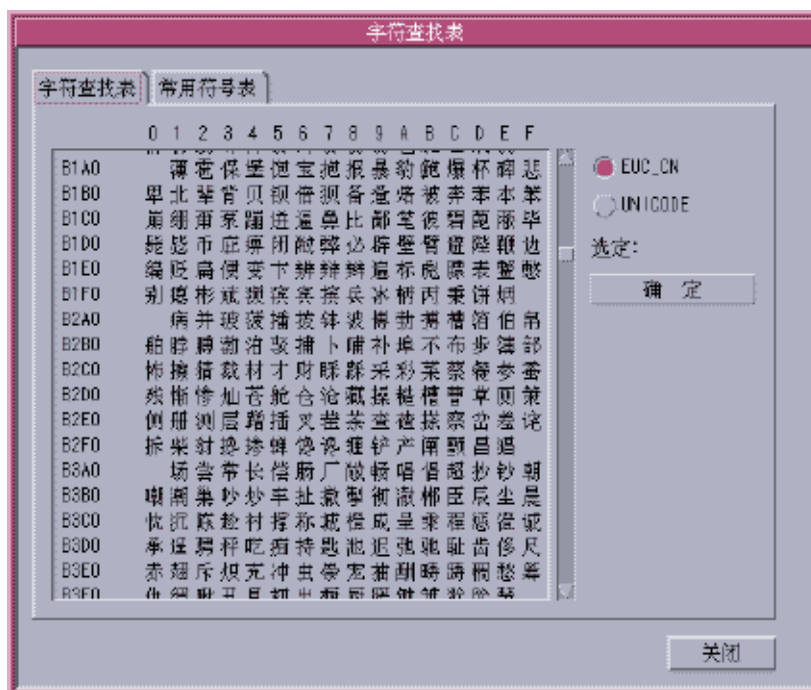
## Lookup table

User can use the the lookup table tools to search a Chinese characters and input it.

There are three kinds of lookup table provided:

- Lookup table with native encoding set. such as: in zh\_CN.EUC/zh\_CN/zh locale, lookup table with EUC\_CN encoding is provided, and in zh\_CN.GBK/zh.GBK, lookup table with GBK encoding is provided, and in zh\_CN.GB18030 locale, GB18030 encoding lookup table is provided.

The lookup table panel with native encoding appears as below:



- Lookup table with UNICODE encoding.  
 The lookup table panel with UNICODE encoding appears as below:



- The Lookup table for special characters, such as Greek characters, Mathmatic symbols, etc. is shown below:





## Greek Keyboard

The Greek virtual keyboard appears as below:



## Russian Keyboard:

The Russian Virtual Keyboard appears as below:





### ZhuYin Keyboard:

The ZhuYin Virtual Keyboard appears as below:



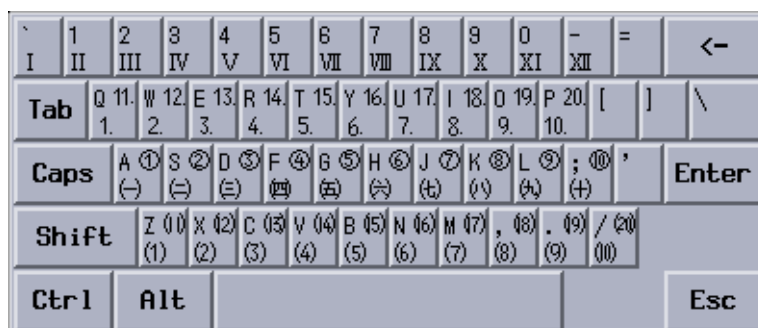
### Chinese Punctuation Characters Keyboard:

The Chinese Punctuation Characters Keyboard appears as below:



### Number Symbol Lookup Keyboard:

The Number Symbol virtual keyboard appears as below



### Mathmatic Symbol Lookup Keyboard:

The Mathmatic Symbol virtual keyboard appears as below:



### Special Symbols Lookup Keyboard:

The Special Symbol virtual keyboard appears as below:



### Table Symbol Lookup Keyboard:

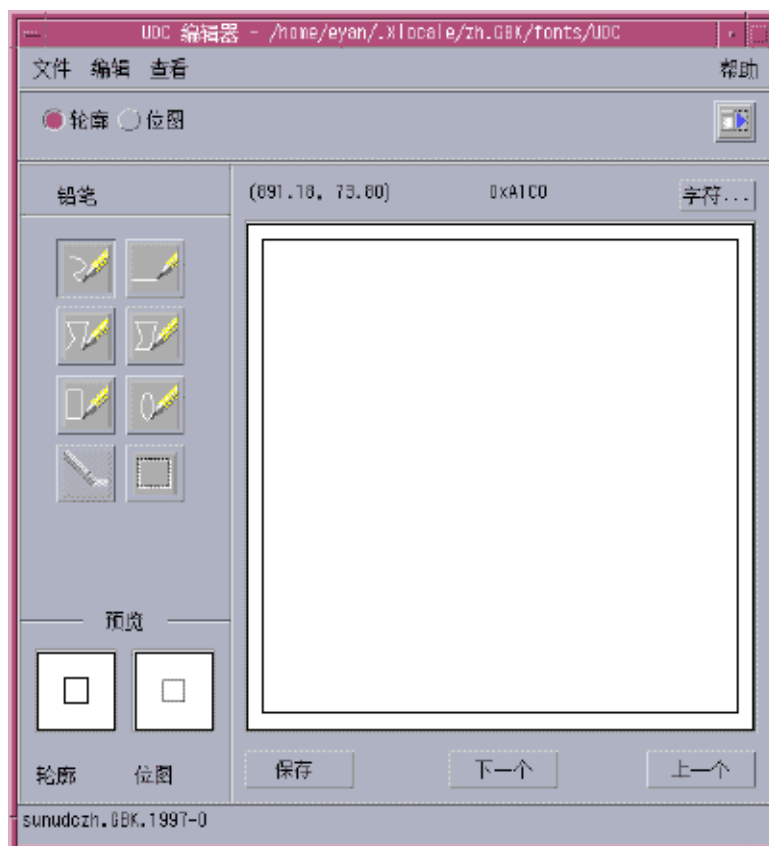
The Table Symbol virtual keyboard appears as below:



## User Defined Characters (UDC)

The UDC editor tool allows you to draw and save new characters. After ascribing the character to an input method, it can be displayed in an application.

- Select the user defined character item on the utility menu to activate the UDC tool.



---

**Note** – Chapter 7 provides more information about user defined characters.

---

## Input Method Help

Help pages are displayed in a default browser such as Netscape or Hotjava.

- Select the input method help item of the utility menu to activate help pages in a browser.

---

# Input Methods and Conversion Modes for Entering Text

The following input methods and conversion modes are available for entering ASCII/English, Simplified Chinese and other text:

In `zh/zh_CN/zh_CN.EUC` locale:

- NewQuanPin input method, the default method (F2).
- NewShuangPin input method (F3).
- QuanPin input method (F4).
- ShuangPin input method (F5).
- GB2312 NeiMa input method (F6).
- English\_Chinese Input method (F7).

In `zh.GBK/zh_CN.GBK` locales:

- NewQuanPin input method, the default method (F2).
- NewShuangPin input method (F3).
- QuanPin input method (F4).
- ShuangPin input method (F5).
- GBK NeiMa input method (F6).
- English\_Chinese Input method (F7).

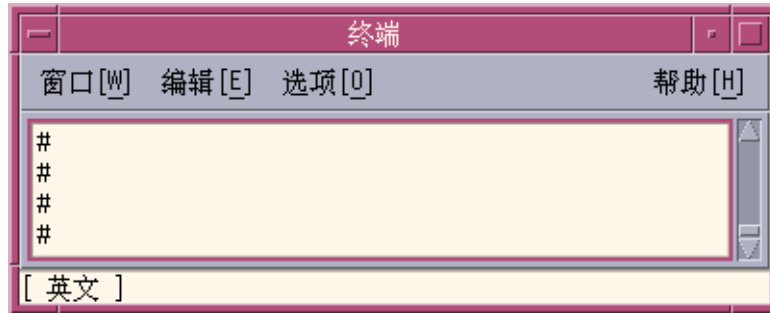
In `zh_CN.GB18030/zh.UTF-8/zh_CN.UTF-8` locales:

- NewQuanPin input method, the default method (F2).
- NewShuangPin input method (F3).
- QuanPin input method (F4).
- ShuangPin input method (F5).
- GB18030 NeiMa input method (F6).
- English\_Chinese Input method (F7).

Press Control-spacebar to toggle on or off the Simplified Chinese input conversion. The function keys listed above (for example: F2, F3) turn on the corresponding input methods.

## Typing ASCII Text

Each tool first starts with all Simplified Chinese input modes off, the window's status area blank. This mode is for typing ASCII text:



Simplified Chinese input conversion mode is toggled on and off by pressing Control-spacebar. After Simplified Chinese input conversion has been turned on once and input conversion is then turned off, the status area is no longer blank, but instead shows that conversion is off.

## Switching Between English Status and Chinese Status:

- Type "CTRL+SPACE".

An auxiliary window appears, as shown below:



## Select Input Method

- In the Chinese status window, type Function key "Fn". For example, type F2 to switch to the first input method. F3 to switch to the second input method, and so on. OR click the input method selection button on the auxiliary window.

The input method selection menu appears as below:



## Toggling Input Methods

This procedure allows you to toggle between 6 input methods. Text is entered in the Chinese status window.

- Type "CTRL+ESC" to switch input methods.

## Switching Between Half\_width Character Mode and Full\_width Character Mode

This method is entered in the Chinese status window.

- Type "SHIFT+SPACE" to switch between Half\_width Character Mode and Full\_width Character Mode. OR click the Half\_width/Full\_width button of auxiliary window to toggle between modes.

The



indicates the input method system is in Half\_width Character Mode.



The



indicates the input method system is in Full\_width Character Mode.

When in Full\_width mode, the Full\_width character of the input key will be committed to system. For example: if you input 'a' when in Full\_width mode, the fullwidth character of 'a' is committed to the application as in the figure below:



## Switching Between Chinese Punctuation Mode and English Punctuation Mode:

This method is entered in the Chinese status window.

- Type "CTRL+." to switch between Chinese Punctuation Mode and English Punctuation Mode. Or click the Chinese/English Punctuation Button on the auxiliary window to toggle between modes.

The



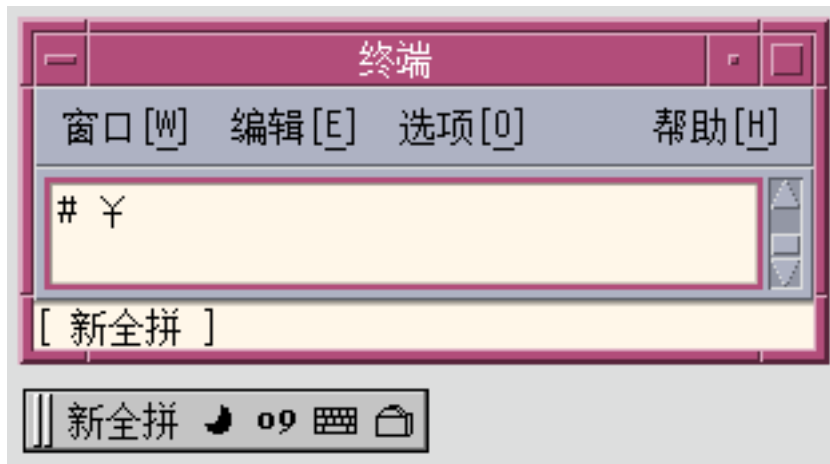
icon indicates the input method system is in Chinese Punctuation Mode.

The



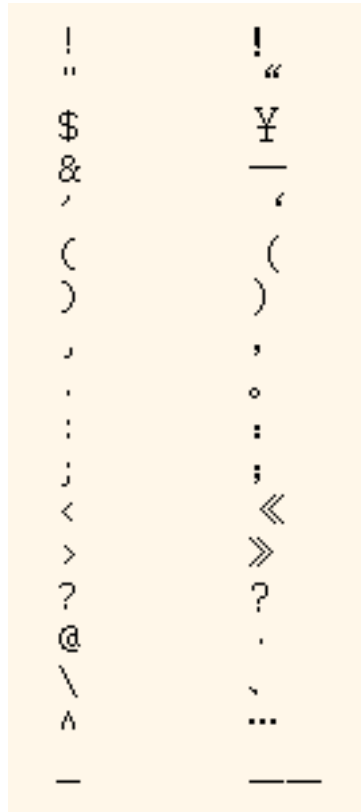
icon indicates the input method system is in English Punctuation Mode.

When the punctuation key is selected in Chinese Punctuation mode, the corresponding Chinese punctuation character is committed to the application. For example: when in Chinese Punctuation mode and the "\$" symbol is selected, the Chinese currency symbol character is committed to the application as in the figure below:



The punctuation keys include the following: , . / <> ; ' " \ \$ ! ^ & \_ -

The correspondence between English and Chinese punctuation is mapped below:



---

## Language Input Methods

Solaris 9 supports the following input methods for the Simplified Chinese environment:

1. NewQuanPin and NewShuangPin Input Method
2. GB2312 Code
3. GBK Code
4. GB18030 Code
5. ShuangPin
6. QuanPin
7. English\_Chinese

---

## New QuanPin and New ShuangPin Input Methods

This section describes the features in the New QuanPin and New ShuangPin input methods, and how to use some of the features in the `zh_CN.EUC` and `zh_CN.GBK` locales.

PinYin is a popular input method in PRC, and there are various PinYin-based input methods. Two of them, New QuanPin and New ShuangPin, contain the following features:

- Supports user-defined phrases for later use.
- Dynamically adjusts the frequency of candidates.
- Accepts PinYin strings up to 222 characters long.
- Supports input by ShengMu.
- Supports GBK Chinese.

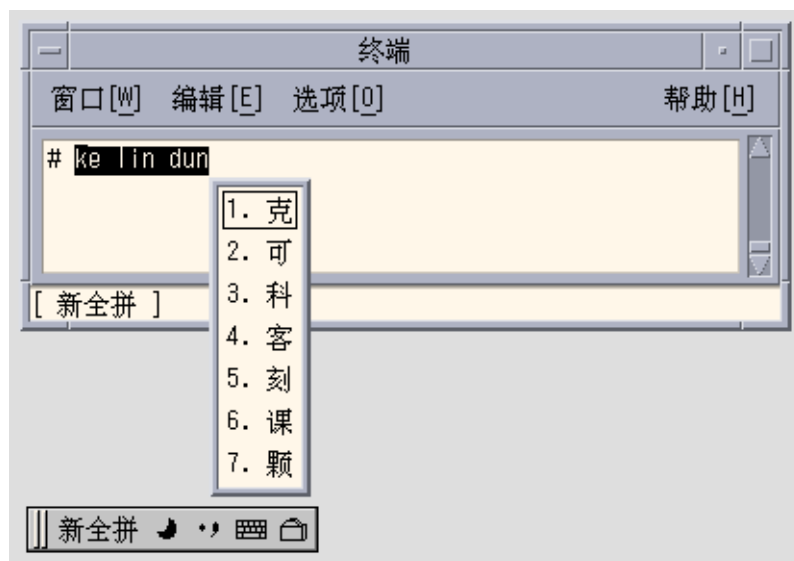
These features are described in detail in the following sections.

### ▼ Defining Phrases for Later Use

The following example shows how to define the phrase “ke lin dun” and store it for later use.

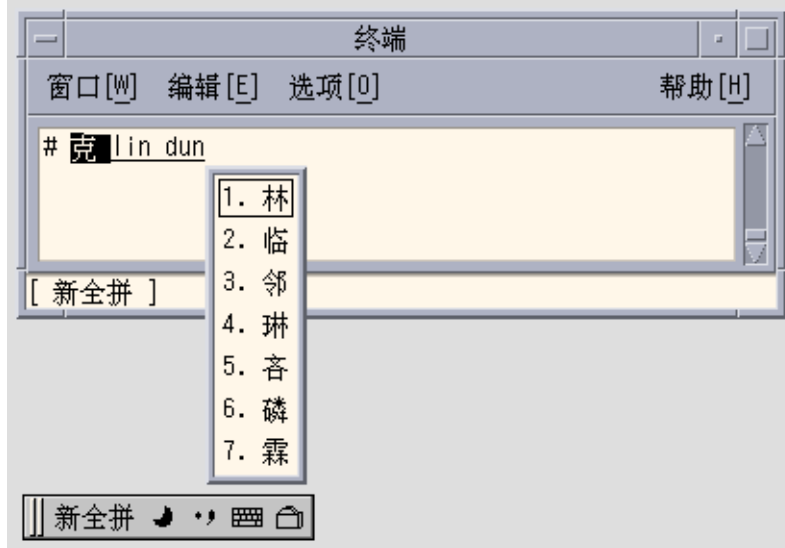
- 1. Type the phrase `kelindun` without spaces.**

The New QuanPin and New ShuangPin input methods will insert spaces for you automatically.

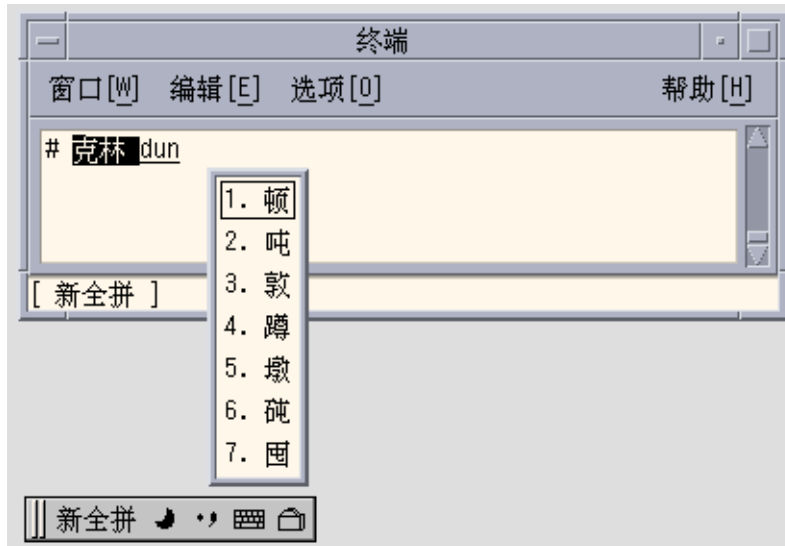


2. Type the number representing the first character you want to select.

The following example shows the second character selected.

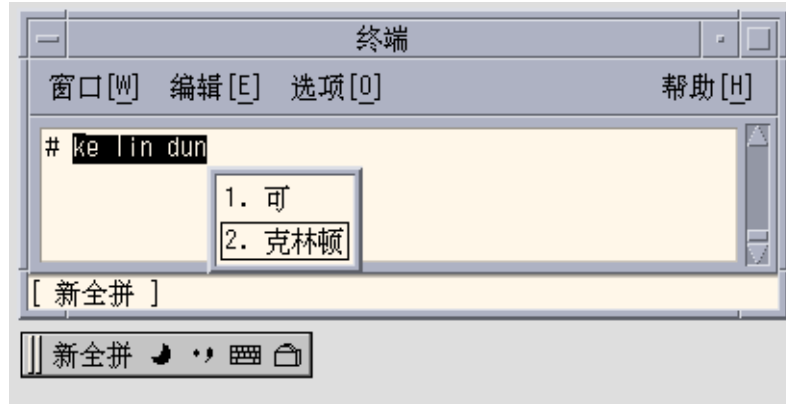


3. Select Chinese for the second and third parts of the phrase.



The new phrase is defined and added to the user dictionary file. The next time you

type **ke lin dun**, you will see the phrase you defined in the candidate area.



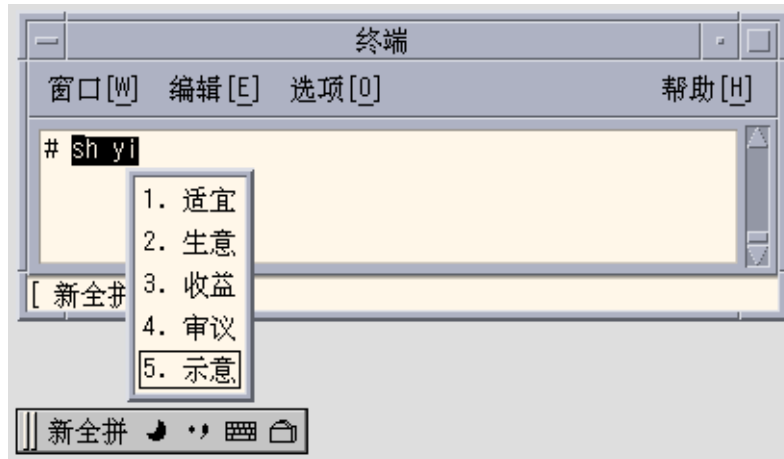
## ▼ Selecting Frequently-Used Candidates

In these input methods, candidates that have been selected are moved to the start of the list to facilitate repeated use.

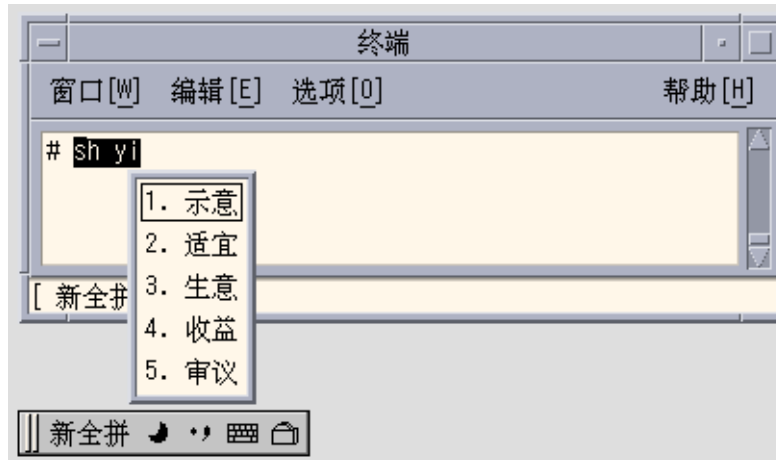
### 1. Type **sh yi**.

Notice the order of the five available candidates.

2. Select the fifth candidate.



3. Type sh yi again.



Notice that the fifth candidate has moved to the first position because you previously selected it. Frequently-used candidates are promoted for faster selection.

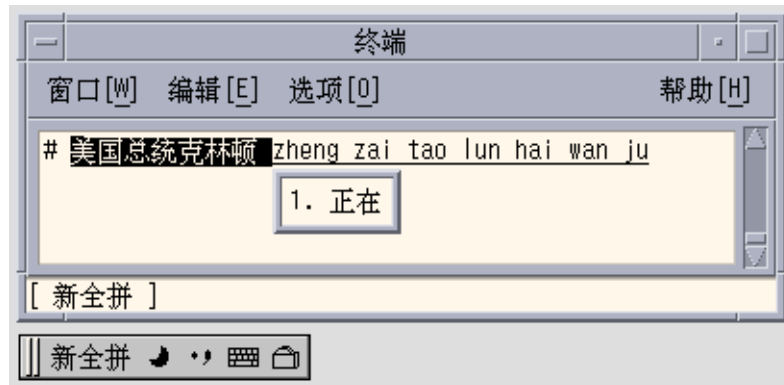


## Other Features

### Typing Long PinYin Strings

The New QuanPin input methods accepts PinYin strings up to 222 characters long. The following illustrations use the string below:

```
>>meiguozhongtongkelindunzhengzaitaolunhaiwanjushiwenti<<
```



The result is the following Chinese string:

美国总统克林顿正在讨论海湾局势问题

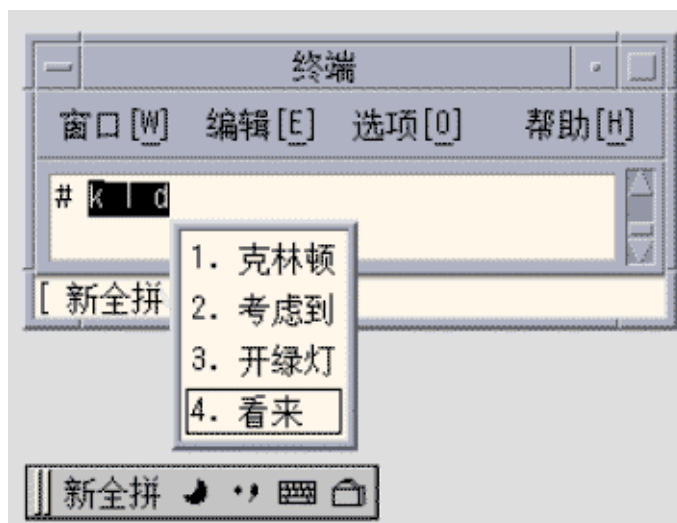
---

**Note** – The New ShuangPin input method supports up to 30-character strings.

---

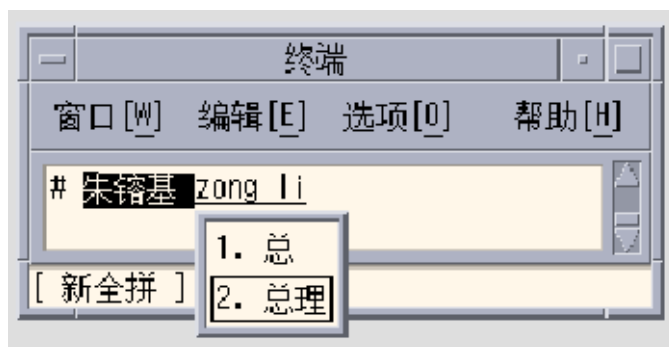
### Typing ShengMu

You can also type ShengMu only. Candidates are supplied for ShengMu, as shown in the following illustration:



## GBK Support

The zh\_CN.GBK locale supports GBK by default, as shown in the following illustration:



The second Chinese character in the following illustration is defined only in the GBK standard.

朱镕基

Single GBK candidates are placed at the end of the list of candidates. Press Return to scroll to the GBK area. For easier selection next time, you can define the GBK candidate as a phrase (for more information, see “Defining Phrases for Later Use” on page 44). Once a phrase is defined, you can insert it easily.

Both New QuanPin and New ShuangPin support GBK Hanzi by default in the zh.GBK locale. However, because several Hanzi have the same ShengMu (the first part of Pinyin), New QuanPin and New ShuangPin do not display GBK candidates if you provide only the ShengMu.

For example, typing the string rong will display GBK candidates because it is a complete Pinyin string. However, typing r alone will not display any GBK candidates because it is only a ShengMu.

## Keyboard Definition

### Edit Keys

The following table shows the definitions of the edit keys.

---

**Note** – The preedit line is a normal X text field.

---

**TABLE 4-1** Edit Key Definitions

Key	Definition
[a-z]	PinYin character.
Home	Moves to the start of the preedit line.
End	Moves to the end of the preedit line.
Left	Moves the caret in the preedit line to the left. If left is Hanzi, the original PinYin is recovered.
Right	Moves the caret in the preedit line to the right.
Delete	Deletes the PinYin character following the caret on the preedit line.
Backspace	Deletes the PinYin character preceding the caret on the preedit line.

### Page Scroll Keys

The candidates of a Pinyin string belong to the following groups:

- G1 – Highest frequency Hanzi + Long (3 or more) Cizu + Double Chinese Cizu
- G2 – GB Single Hanzi

- G3 – GBK Single Hanzi (in the zh\_CN.GBK locale)

Some Pinyin strings may have more candidates than can be displayed in the same window. In that case, use the keys described in the following table to scroll through the candidates.

**TABLE 4-2** Page Scroll Key Definitions

Key	Definition
- =	Scrolls to previous/next candidate(s)
[ ]	Scrolls to previous/next candidate(s)
, .	Scrolls to previous/next candidate(s)
Return	Quickly scrolls through all candidates

## Select Keys

New QuanPin and New ShuangPin use the numeric selection keys.

## Separators

In accord with the national Pinyin standard, the separator (') is supported to avoid ambiguous interpretations of Pinyin strings. For example, the Pinyin string [jiang] can be interpreted as [jiang] or [ji] [ang]; both are valid. In New QuanPin, however, [jiang] is interpreted only as [jiang]. You must use the separator and enter [ji'ang] for it to be interpreted as [ji] and [ang]. New ShuangPin does not require the use of separators.

## Dictionary Files

New QuanPin and New ShuangPin share two dictionary files: PyCiku.dat and Ud.Ciku.dat. In the zh\_CN.EUC and zh\_CN.GBKlocale, the default path names are /usr/lib/im/locale/zh\_CN/data/PyCiku.dat and /usr/lib/im/locale/zh\_CN/data/UdCiku.dat.

Users cannot normally write to these files. However, since users can affect the way New QuanPin and New ShuangPin work through features such as frequency adjustment and user-defined phrases, it is necessary to update the dictionary files frequently.

A user's dictionary is normally located in ~/.Xlocale/PyCiku.dat or ~/.Xlocale/UdCiku.dat (~ indicates the home directory of the user who starts the http command). When New QuanPin and New ShuangPin are started, they locate and read the dictionary files in the user's home directory. If a dictionary file is not found, the system default path is used (that is, /usr/lib/im/locale/zh\_CN/...).

## New ShuangPin Features

ShuangPin is an abbreviated form of QuanPin. It is faster but more difficult to use than QuanPin. New ShuangPin supports all of the features, keyboard definitions, and dictionary files of New QuanPin.

There are various ShuangPin keyboard mapping designs in PRC. The most popular three are ZiRanMa, Chinese Star, and Intelligent\_ABC. The New ShuangPin input method supports all three of these keyboard mappings.

## New ShuangPin Keyboard Mapping

The following tables contain keyboard mappings for the ZiRanMa, Chinese Star, and Intelligent\_ABC keyboards.

**TABLE 4-3** ZiRanMa Keyboard Mapping

Key	Definition
i	ch
u	sh
v	zh
a	a
b	ou
c	iao
d	uang, iang
e	e
f	en
g	eng
h	ang
i	i
j	an
k	ao
l	ai
m	ian
n	in

**TABLE 4-3** ZiRanMa Keyboard Mapping (Continued)

Key	Definition
o	o, uo
p	un
q	iu
r	uan, er
s	iong, ong
t	ue
u	u
v	v, ui
w	ua, ia
x	ie
y	uai, ing
z	ei

**TABLE 4-4** CStar2.97 Keyboard Mapping

Key	Definition
u	ch
i	sh
v	zh
a	a
b	ia, ua
c	uan
d	ao
e	e
f	an
g	ang
h	iang, uang
i	i
j	ian
k	iao

**TABLE 4-4** CStar2.97 Keyboard Mapping (Continued)

Key	Definition
l	in
m	ie
n	iu
o	o, uo
p	ou
q	er, ing
r	en
s	ai
t	eng
u	u
v	v, ui
w	ei
x	uai, ue
y	iong, ong
z	un

**TABLE 4-5** Intelligent ABC Keyboard Mapping

Key	Definition
i	ch
u	sh
v	zh
a	a
b	ou
c	in, uai
d	ua, ia
e	e
f	en
g	eng
h	ang

**TABLE 4-5** Intelligent ABC Keyboard Mapping (Continued)

Key	Definition
i	i
j	an
k	ao
l	ai
m	ue, ui
n	un
o	o, uo
p	uan
q	ei
r	iu, er
s	ong, iong
t	uang, iang
u	u
v	v
w	ian
x	ie
y	ing
z	iao

---

## GBK Code Input Method

This method uses the GBK code defined by the Chinese Internal Code Specification. It includes all of the Chinese characters and symbols in GB2312-80, and other CJK Chinese characters in GB 13000-1. Each Chinese character or symbol is identified by a four hexadecimal digital internal code defined in the Chinese Internal Code Specification.

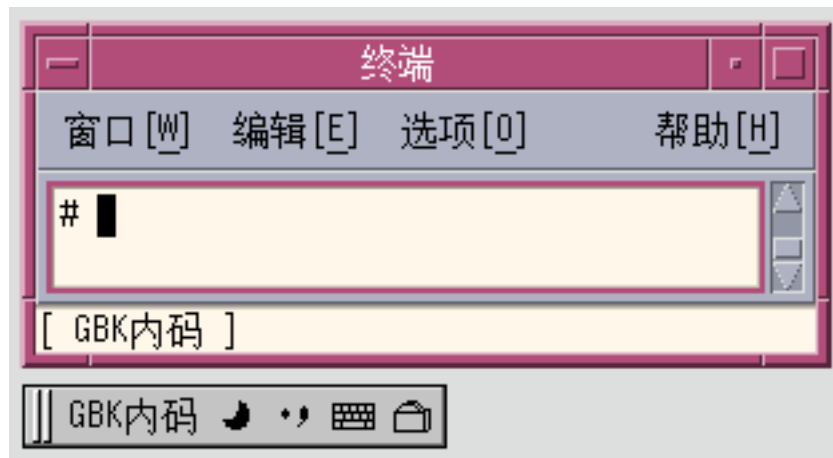


## Typing GBK Code Text

This section contains instructions on how to use the GBK codes to type Chinese characters and symbols.

1. In a new Terminal, turn Chinese input conversion on by pressing Control-spacebar.
2. Press F4 to turn on the GBK code input method.

The status area shows that GBK code input mode is on.



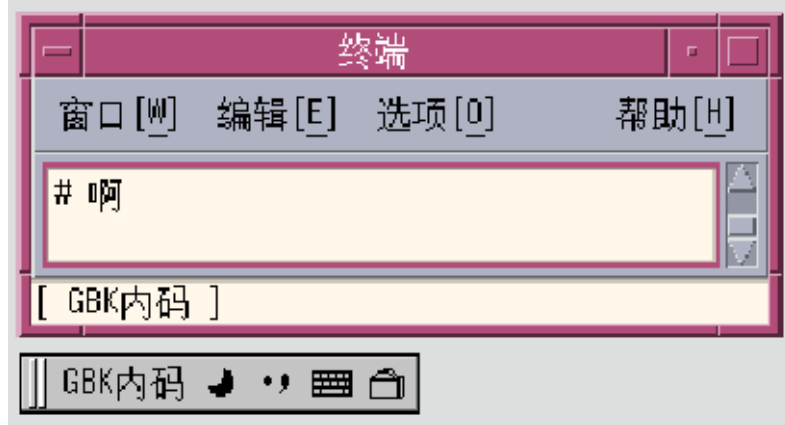
3. Press the first three of the four keys that represent the character to display (in this example, b0a1).

The key remains visible in the preedit area.



4. Type the fourth key.

The character automatically replaces the preedit area.



---

## GB2312 Code Input Method

This method uses the GBK code defined by the Chinese Internal Code Specification. It includes all of the Chinese characters and symbols in GB2312-80, and other CJK Chinese characters in GB 13000-1. Each Chinese character or symbol is identified by a four hexadecimal digital internal code defined in the Chinese Internal Code Specification.

### Typing GB2312 Code Text

This section contains instructions on how to use the GB2312 codes to type Chinese characters and symbols.

1. In a new Terminal, turn Chinese input conversion on by pressing Control-spacebar.
2. Click the Input method selection button on the auxiliary window and select GB2312 input method.

The status area shows that GB2312 code input mode is on.



3. Press the first three of the four keys that represent the character to display (in this example, b0a1).

The key remains visible in the preedit area.



4. Type the fourth key.

The character automatically replaces the preedit area.



---

## GB18030 Code Input Method

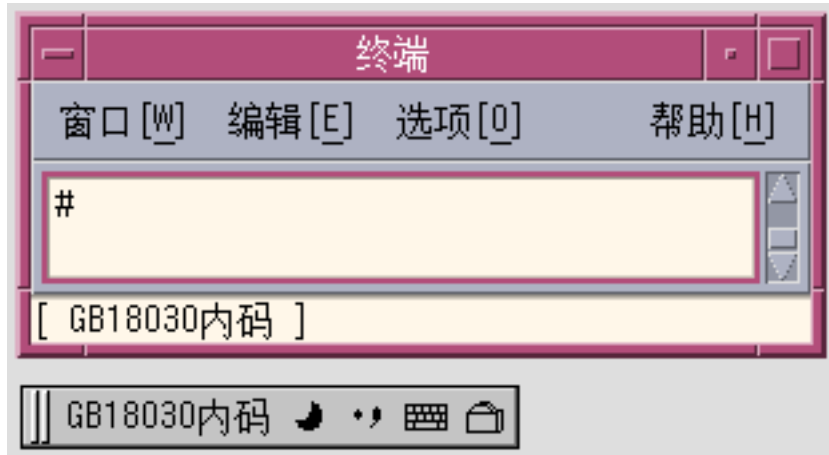
This method uses the GB18030 code defined by the Chinese Internal Code Specification. It includes all of the Chinese characters and symbols in GB2312-80, and other CJK Chinese characters in GB 18030. Each Chinese character or symbol is identified by a four or eight hexadecimal digital internal code defined in the Chinese Internal Code Specification.

### Typing GB18030 Code Text

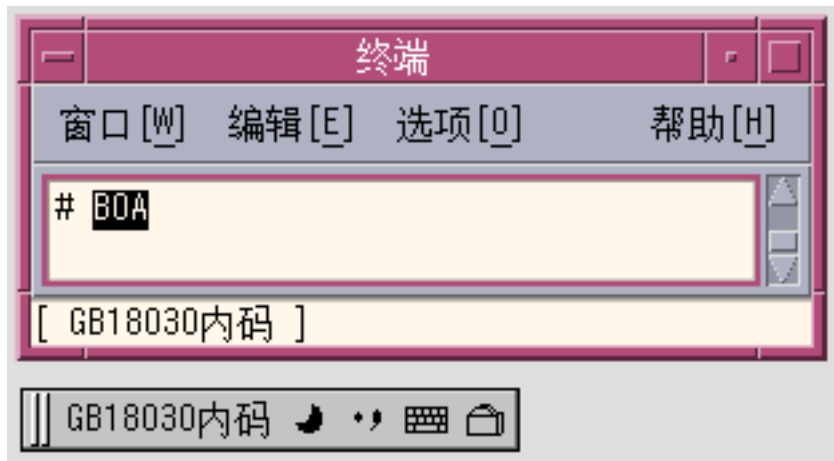
This section contains instructions on how to use the GB18030 codes to type Chinese characters and symbols.

1. In a new Terminal, turn Chinese input conversion on by pressing Control-spacebar.
2. Click the Input method selection button on the auxiliary window and select GB18030 input method.

The status area shows that GB18030 code input mode is on.



3. For example: To input Chinese GB18030 character with code 0xb0a1, press the first three of the four keys that represent the character to display (in this example, b0a). The key remains visible in the preedit area.



4. Type the fourth key.  
The character automatically replaces the preedit area.

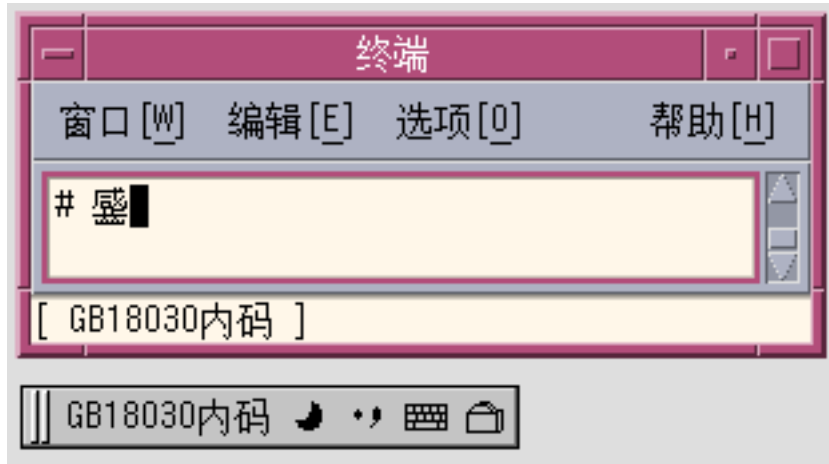


5. For example: To input Chinese GB18030 character with code 0x82358538, press the first three of the four keys that represent the character to display (in this example, 8235853).

The key remains visible in the preedit area.



6. Type the last key. The character is automatically committed to the window.



---

## QuanPin Input Method

The QuanPin input method requires up to six keystrokes to type each Chinese Pinyin character. Quanpy maps Pinyin phonetics to single lowercase Roman letters. You can use the QuanPin input method to type individual Chinese characters in both zh\_CN.EUC and zh\_CN.GBK.

A lookup area showing the characters that match the QuanPin input is displayed with each keystroke. If more than one option is available, you can type a period (.) to display the next page moving forward through the lookup choices and typing a comma (,) to display the next page moving backward. You can select the character you want by typing the label letter corresponding to the character in the lookup area.

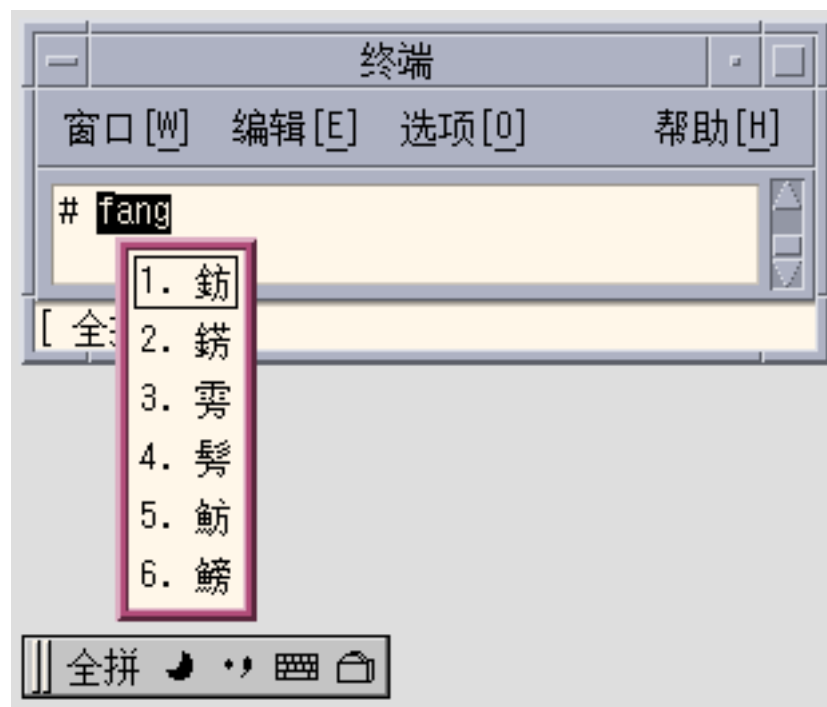
### Typing QuanPin Text

This section describes how to create QuanPin text.

The following figure shows how to use this input method to type the character representing the Full Pinyin word *fang*. The word requires four keystrokes. Type them and select the text as follows:

1. **Type the four keystrokes** *fang*.





2. Type 1 to select the corresponding GBK Chinese character in the lookup choice list.

Your choice is substituted for the Full Pinyin string in the preedit area.



---

## English\_Chinese Input Method

The English\_Chinese input method requires up to fifteen keystrokes to type each Chinese word. English\_Chinese maps the English word to a Chinese phrase. You can use the English\_Chinese input method to type a Chinese phrase in both zh\_CN.EUC and zh\_CN.GBK locales.

A lookup area showing the characters that match the QuanPin input is displayed with each keystroke. If more than one option is available, you can type a period (.) to display the next page. Moving forward through the lookup choices and typing a comma (,) displays the next page moving backward. You can select the character you want by typing the label letter corresponding to the character in the lookup area.

## Typing English\_Chinese Text

This section describes how to create English\_Chinese text.

The following figure shows how to use this input method to type the character representing the English word "world". The word requires five keystrokes. Type them and select the text as follows:

1. Type the five keystrokes world.



2. Type 3 to select the corresponding Chinese phrase in the lookup choice list.  
Your choice is substituted for the English string in the preedit area.

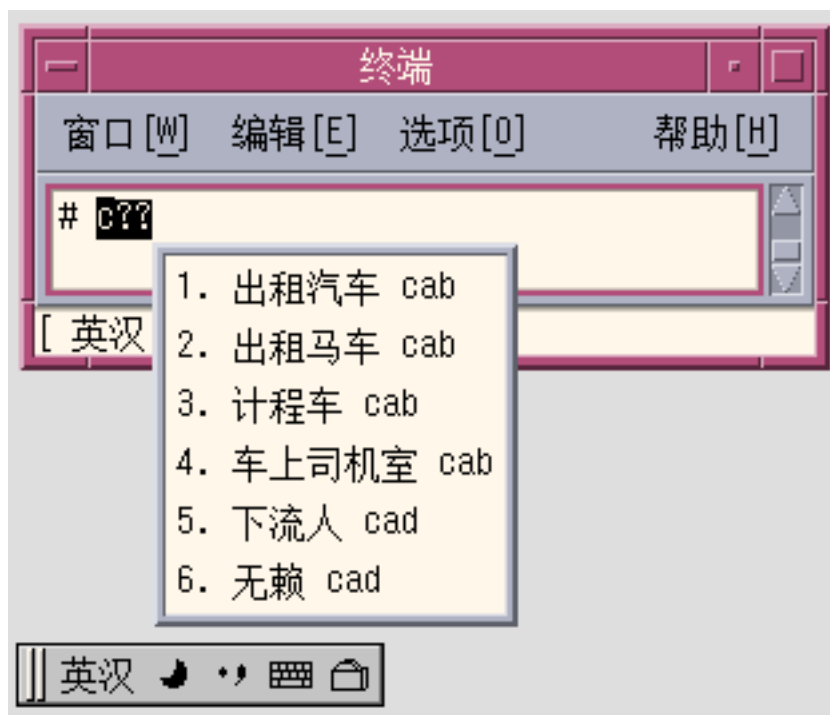


- Wild characters "\*" or "?" can be used to search in the dictionary; "\*" stands for one or several letters, and ? represents only one letter. For example, to search all English

words which end with lution, you can input “\*lution” and the lookup choice window appears as below.



Or to search all English words which begin with “c”, and only three letters, you can input “c??”, the lookup choice window appears as below:



# Code Table Input Method Interface

---

---

## Introduction

The Simplified Chinese Solaris software provides code table interface for adding new input methods. In a plain text file, you can create a list Chinese words or phrases along with the key strokes you have selected to create them.

### ▼ Creating a Code Table

#### 1. Create and edit code table source file:

First create a plain text file, which contains a list Chinese words or phrases along with the key strokes, then convert it to binary format that the codetable input method interface can recognized, then user can add it into system.

There are several sections in codetable file:

```
#  
# Description  
# Comment  
# Key_Prompt  
# Function_Key  
# Phrase  
# Single  
# Options  
#
```

The following figure shows an example source file:

```

HANZI codetable input table

[Description]
Name:          英汉
Encode:        GB2312
UsedCodes:     -ABCDEFGHIJKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz
WildChar:      *?
MaxCodes:      15

[Function_Key]
PageUp:        ,<
PageDown:      .>

[Options]
KeyByKey_Mode:      ON
HelpInto_Mode:      ON
AutoSelect_Mode:    OFF
KeyPrompt_Mode:     ON
SelectKey_Mode:     Number

[Phrase]
Chinese 中国人 中国人的 中国的 中文 汉语
Christ  基督 救世主
Christian 信徒 基督徒 基督教徒 基督教的

```

## 2. Enter the Description

This section specifies attributes of the codetable, such as encoding, name, valid characters, the maximum number of codes for one input item, and wild characters.

- a. **Name:** specify the name of this codetable.
- b. **Encode:** specify the encoding of this codetable, can be UTF-8, GB, GB2312, GBK, BIG5.
- c. **WildChar:** specify the wild character for input codes.
- d. **UsedCodes:** specify the valid characters to input.
- e. **MaxCodes:** specify the maximum number of input codes for one item.

## 3. Enter the Comment

This section can be used to keep the codetable writer's comments or explanations.

## 4. Enter the Key\_Prompt

This section specifies the prompt string of an input key. The prompt string will be displayed in the Preedit Area of application software.

## 5. Enter the Function\_Key



This section specifies the key definitions of some function keys, such as: Page UP key to scroll the candidate items up, Page Down key to scroll the candidate items down, BackSpace Key to delete an input code, and the ClearAll key to clear all input areas, such as the preedit area and the lookup area.

This section has following entry items:

- a. "PageUp:"
- b. "PageDown:"
- c. "BackSpace:"
- d. "ClearAll:"

#### 6. Enter Options

This section specify the options of the codetable input method, such as: whether to display help information for each candidate items, or whether to display the prompt string of the input key in preedit area, whether to display the lookup candidates key by key or only display the lookup candidates when user input space key, whether to commit the candidate when there is only one lookup result, and the select key modes: Number mode, Lower case mode or Upper case mode.

The Options section has following entry items:

- a. "HelpInfo\_Mode:" Value: "ON" or "OFF"
- b. "KeyByKey\_Mode:" Value: "ON" or "OFF"
- c. "KeyPrompt\_Mode:" Value: "ON" or "OFF"
- d. "AutoSelect\_Mode:" Value: "ON" or "OFF"
- e. "SelectKey\_Mode:" Value: "Number", "Lower" or "Upper"

#### 7. Enter the Single Section

This section specify the input codes and its corresponding single Chinese characters. these Chinese characters must not be separated by space key.

The format of every line as follows:

```
eystroke_sequence Characterlist
```

CharacterList is a list of Chinese characters with no separate spaces.

#### 8. Enter the Phrase Section

This section specifies the input codes and their corresponding phrase words. These Chinese phrase words must be separated by the space key.

The format of every line is as follows:

```
keystroke_sequence word1 word2 word3 . . .
```

## ▼ Converting the Source Codetable File from Text Format to Binary Format or from Binary Format to Text Format:

Use the utility tools `txt2bin` to convert a text codetable file to binary file. and `bin2txt` to convert binary format file to text format file.

`bin2txt` and `txt2bin` are under directory:  
`/usr/lib/im/locale/zh_CN/common`.

The command syntax is:

```
# /usr/lib/im/locale/zh_CN/common/txt2bin
source_codetable_file
binary_codetable_file
# /usr/lib/im/locale/zh_CN/common/bin2txt
binary_codetable_file
source_codetable_file
```

## ▼ Creating a new codetable input method

### 1. Create and edit code table source file:

Prepare the code table source file to present the new input method.

This dictionary file is a plain text file that contains the words user wants to have associated. The words are separated by spaces.

You need to list characters, words, or phrases and the corresponding key strokes required to create them. Each row defines the mapping between one Chinese word and its input key stroke. Thus, when the user types a Chinese character, all the words associated with it in the dictionary will be shown for a selection.

### 2. Convert the source codetable file to binary format:

Use the utility tools "`txt2bin`" to convert the prepared text codetable file to binary file.

The command syntax is:

```
# /usr/lib/im/locale/zh_CN/common/txt2bin
source_codetable_file binary_codetable_file
```

### 3. Add the code table into the input method specification file.

After the binary codetable file is ready, the input method can be added to the system by updating input method specification file:

```
/usr/lib/im/locale/zh_CN/sysime.cfg
```

The "`sysime.cfg`" file requires the following information: Input method name

For example: If your new codetable binary file is `newim.data`, then add a line `newim` to the input method specification file `sysime.cfg`: Once the `sysime.cfg` file is updated, the Simplified Chinese Solaris operating environment can interpret the input

method.

**4. Restart the input method server (htt) or relogin system so that your new input method can be recognized.**

To restart the htt server, login as root, and run:

```
# /etc/init.d/IIim stop  
# /etc/init.d/IIim start
```

Then logout from the CDE desktop and relogin. The new input method will be ready for use.



## Localized Applications

---

---

### About This Chapter

This chapter describes selected properties you need to use with two localized applications of the Simplified Chinese Solaris operating environment. This chapter also provides lists of code conversion utilities.

These tools (and the commands to invoke them) include:

- `mailx (/usr/SUNWale/bin/mailx)`
- `talk (/usr/SUNWale/bin/talk)`

---

### Using the `mailx` Utility With Simplified Chinese Characters

The encoding variable in `.mailrc` does not affect Mailer behavior, but it does set `mailx` function. To send e-mail in formats other than 7-bit ASCII, such as to send (8-bit) Simplified Chinese characters you must use `/usr/SUNWale/bin/mailx`. The encoding variable in `.mailrc` sets encoding formats for `mailx`.

International transmission conventions require that header information use only ASCII characters. So Simplified Chinese characters should not be used in the header (including the Subject line) with Mailer or with `mailx`.

---

**Note** – The `/bin/mailx` application has not been localized and cannot send or receive Simplified Chinese characters.

---

---

## Using `talk` With Simplified Chinese Characters

To use Simplified Chinese characters with `talk`, the `/usr/SUNWale/bin/talk` application is required because `/bin/talk` is not localized.

---

## The `xtobdf` Utility

Simplified Chinese Solaris software provides this BDF font generator:

`xtobdf` — Convert from font in X server to font in BDF

---

## `stdconvtool`

`Stdconvtool` is a graphic user interface utility that enables file conversion between various code sets. Its functionality is the same as `iconv`.



- Please select both source Code set and target code set from pull-down list respectively and type source file into Source File Path field.

---

## iconv

The `iconv` command converts the characters or sequences of characters in a file from one code set to another and writes the results to standard output. Simplified Chinese Solaris software includes special filters for the `iconv` command.

If no conversion exists for a particular character, it is converted to the underscore “\_” in the target codeset. The following options are supported:

- f from code — Symbol of the input code set.
- t to code — Symbol of the output code set.

The following code set conversion modules are supported in Simplified Chinese Solaris software.

**TABLE 6-1** Simplified Chinese i conv Code Conversion Modules (zh locale)

<b>Code</b>	<b>Symbol</b>	<b>Target Code</b>	<b>Symbol</b>
GB2312-80	zh_CN.euc	ISO 2022-7	zh_CN.iso2022-7
ISO 2022-7	zh_CN.iso2022-7	GB2312-80	zh_CN.euc
GB2312-80	zh_CN.euc	ISO 2022-CN	zh_CN.iso2022-CN
ISO-2022-CN	zh_CN.iso2022-CN	GB2312-80	zh_CN.euc
UTF-8	UTF-8	GB2312-80	zh_CN.euc
GB2312-80	zh_CN.euc	UTF-8	UTF-8
GB2312-80	zh_CN.euc	BIG5	zh_TW-big5
BIG5	zh_TW.big5	GB2312-80	zh_CN.euc
HZ-GB-2312	HZ-GB-2312	GB2312-80	zh_CN.euc
GB2312-80	zh_CN.euc	HZ-GB-2312	zh_CN.euc

**TABLE 6-2** Simplified Chinese i conv Code Conversion Modules (zh.GBK locale)

<b>Code</b>	<b>Symbol</b>	<b>Target Code</b>	<b>Symbol</b>
UTF-8	UTF-8	GBK	zh_CN.gbk
GBK	zh_CN.gbk	UTF-8	UTF-8
GBK	zh_CN.gbk	BIG5P	zh_TW-big5p
GBK	zh_CN.gbk	BIG5HK	zh_TW-big5hk
GBK	zh_CN.gbk	ISO-2022-CN	zh_CN.iso2022-CN
ISO-2022-CN	zh_CN.iso2022-CN	GBK	zh_CN.gbk
GBK	zh_CN.gbk	BIG5	zh_TW-big5
BIG5	zh_TW-big5	GBK	zh_CN.gbk
BIG5P	zh_TW-big5p	GBK	zh_CN.gbk
BIG5HK	zh_TW-big5hk	GBK	zh_CN.gbk
HZ-GB-2312	HZ-GB-2312	GBK	zh_CN.gbk
GBK	zh_CN.gbk	HZ-GB-2312	zh_CN.gbk
HZ-GB-2312	HZ-GB-2312	UTF-8	UTF-8
UTF-8	UTF-8	HZ-GB-2312	HZ-GB-2312



**TABLE 6-3** Simplified Chinese `iconv` Code Conversion Modules (`zh_CN.GB18030` locale)

Code	Symbol	Target Code	Symbol
UTF-8	UTF-8	GB18030-2000	zh_CN.gb18030
GB18030-2000	zh_CN.gb18030	UTF-8	UTF-8
GB18030-2000	zh_CN.gb18030	BIG5HK	zh_HK-big5hk
GB18030-2000	zh_CN.gb18030	BIG5P	zh_TW-big5p
BIG5HK	zh_HK-big5hk	GB18030-2000	zh_CN.gb18030
BIG5P	zh_TW-big5p	GB18030-2000	zh_CN.gb18030

The following `iconv` code conversion modules are located in `/usr/lib/iconv`:

For the `zh` locale:

- `zh_CN.euc%zh_TW-big5.so`
- `zh_TW-big5%zh_CN.euc.so`

For the `zh.GBK` locale:

- `UTF-8%zh_CN.gbk.so`
- `zh_CN.gbk%UTF-8.so`
- `zh_CN.gbk%zh_CN.iso2022-CN.so`
- `zh_CN.iso2022-CN%zh_CN.gbk.so`
- `zh_CN.gbk%zh_TW-big5.so`
- `zh_TW-big5%zh_CN.gbk.so`

In the following example, an EUC mail file is converted to ISO 2022-CN:

```
system% iconv -f zh_CN.euc -t zh_CN.iso2022-CN mail.euc > mail.iso2022-CN
```

For further information, see the `iconv(3)` and `iconv_zh(5)` man pages. These utilities can be used for converting files for printing. See *Simplified Chinese Solaris System Administrator's Guide* for more information.



## Fonts

---

---

### About This Chapter

This chapter describes the PostScript fonts included in the Simplified Chinese Solaris operating environment, what you need to use them, and how to edit them.

---

### Display PostScript System (DPS)

The Simplified Chinese Solaris operating environment provides PostScript fonts in the Display PostScript System (DPS). This section describes what you need to use DPS in Simplified Chinese Solaris software. For further details, see *Programming the Display PostScript System with X*, published by Adobe® Systems.

### Using Simplified Chinese PostScript Fonts and DPS Facilities

The Simplified Chinese Solaris operating environment includes the fonts listed in the following table. You can use any of the Simplified Chinese fonts just as you would use Roman fonts.

**TABLE 7-1** Simplified Chinese Solaris Operating Environment PostScript Fonts

<b>Font Name</b>	<b>Description</b>
Song-Medium	Alias of Song-Medium-EUC; can be used like a Roman font.
Song-Medium-EUC	Song-Medium font, EUC encoding, horizontal display; can be used like a Roman font.
Kai-Medium	Alias of Kai-Medium-EUC; can be used like a Roman font.
Kai-Medium-EUC	Kai-Medium font, EUC encoding, horizontal display; can be used like a Roman font.
FangSong-Medium	Alias of FangSong-Medium-EUC; can be used like a Roman font.
FangSong-Medium-EUC	FangSong-Medium font, EUC encoding, horizontal display; can be used like a Roman font.
Hei-Medium	Alias of Hei-Medium-EUC; can be used like a Roman font.
Hei-Medium-EUC	Hei-Medium font, EUC encoding, horizontal display; can be used like a Roman font.

The following figure shows a sample of Song-Medium.



FIGURE 7-1 Sample Simplified Chinese Text Display PostScript Output

## Creating Composite Roman and Simplified Chinese Fonts

You can create composite fonts using any one Roman font and the Simplified Chinese fonts. For example, the following PostScript code defines a composite font, Times-Italic+Kai-Medium, which uses Times-Italic for ASCII characters and Kai-Medium font for Simplified Chinese characters:

```

/makeEUCfont {
  /AsianFont exch def
  /WestFont exch def
  /NewFont exch def

  /CIDInit /ProcSet findresource
  begin

```

```

NewFont [ AsianFont WestFont ]
beginrearrangedfont
  1 usefont
  1 beginbfrange
    <00> <7e> <00>
  endbfrange
endrearrangedfont
end
} bind def

/LC_Times-Roman /Times-Roman /Kai-Medium makeEUCfont

```

## Using Simplified Chinese Fonts in DPS Programming

You can use Simplified Chinese fonts just as you use Roman fonts in DPS wrap definitions. The following code sample creates the display in the next figure.

```

defineps PSWDisplayText(char *text)
  /pointSize 50 def
  /Helvetica pointSize selectfont
  (Hello World) stringwidth pop 2 div neg 0 moveto
  (Hello World) show

  /cpSize 40 def
  /Song-Medium cpSize selectfont
  (text) stringwidth pop 2 div neg pointSize neg moveto
  (text) show
endps

```

You can call `PSWDisplayText` (*Chinese text*) in a C program to display the designated Chinese text. For an example see the following figure.



**FIGURE 7-2** Using Simplified Chinese Solaris Operating Environment DPS

Simplified Chinese Solaris software provides TrueType support in DPS.

---

## Converting BDF to PCF Format

Before applications can use the modified BDF file, it must be converted to a PCF format file.

- **Make the edited font usable by the Simplified Chinese Solaris operating environment by converting to X11 PCF format as follows:**

```
system% bdf2pcf -o myfont14.pcf myfont.bdf
```

For more information, see the `bdf2pcf(1)` man page.

---

## Installing and Checking the Edited Font

1. To add a new bitmap, move the `.pcf` font file into your font directory. You may compress the `.pcf` font file before moving it to your font directory:

```
system% compress myfont14.pcf
```

2. Run the following commands in your font directory.

The `.bdf` file should not be in the font directory.

```
system% cat >> fonts.alias
-new-myfont-medium-r-normal--16-140-75-75-c-140-gb2312.1980-0
Myfont-Medium14'
^D
system% mkfontdir
system% xset +fp `pwd`
```

3. You can view your font using the X Font Displayer (`xfd`) by typing:

```
system% xfd -fn Myfont-Medium14
```

All Simplified Chinese fonts have XLFD font names, and you can use the `xlsfonts` utility to display these names as follows:

```
system% xlsfonts | grep gb2312
```

---

## Creating Characters With `sdtudctool`

You can create new characters using the drawing tools in the `sdtudctool` utility. (For information on modifying a font, see `Editing Fonts With fontedit` on page 83.) Following is the user defined characters code range for different SCH locales:

**TABLE 7-2** Code Points for locales

zh	0xD7FA-0xD7FE
zh.GBK	0xAAA1-0xAFFE
zh_CN.GB18030	0xAAA1-0xAFFE
zh.UTF-8	0xAAA1 - 0xAFFE

The `sdtudctool` utility supports bitmap, Type 1, and CID fonts. You can also specify a font size for the new characters. This section contains procedures for using `sdtudctool`.



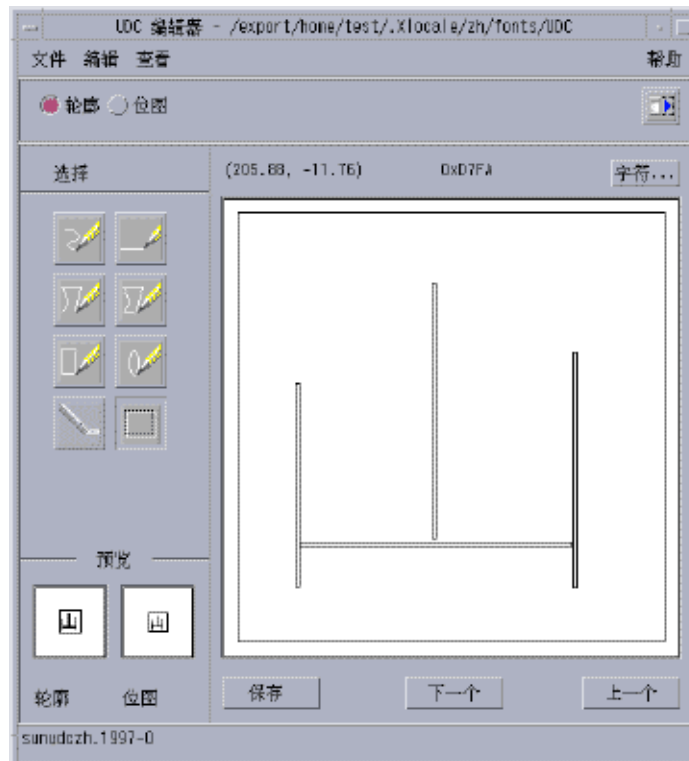
## ▼ Using sdtudctool

This section contains instructions for creating, inputting and printing new characters with the `sdtudctool` utility and other tools. For example inzh locale, use the following procedure.

1. To start the utility, type the following command:

```
system% sdtudctool
```

The following window appears.



2. Click the Character button.

The following dialog box appears:



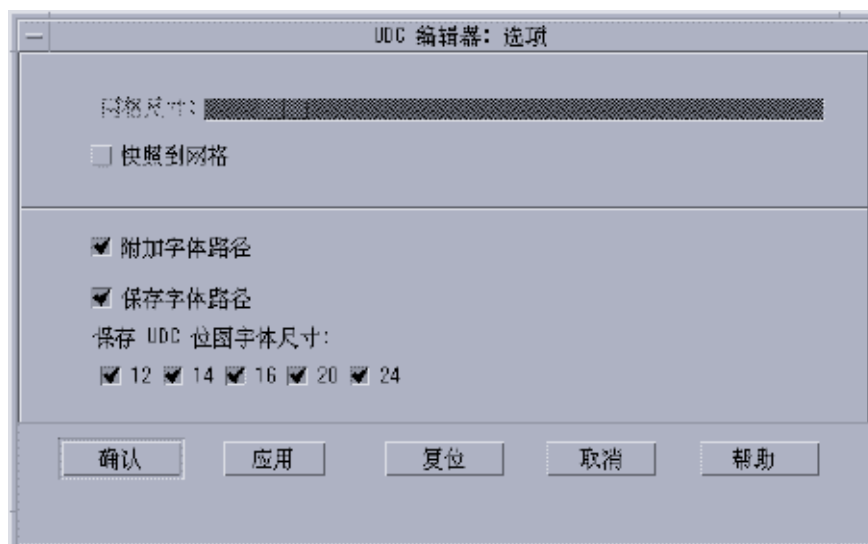
3. Specify the code position of the new character.
4. Create the outline or bitmap for the character using the drawing tools.
5. Click Save to save the character.
6. Continue creating new characters.
7. When you are finished, choose Save from the File menu to save the changes to your environment.

The new character(s) are saved in `$HOME/.Xlocale/{locale}/fonts/UDC...`

## ▼ Specifying Options

1. Choose Options from the File menu.

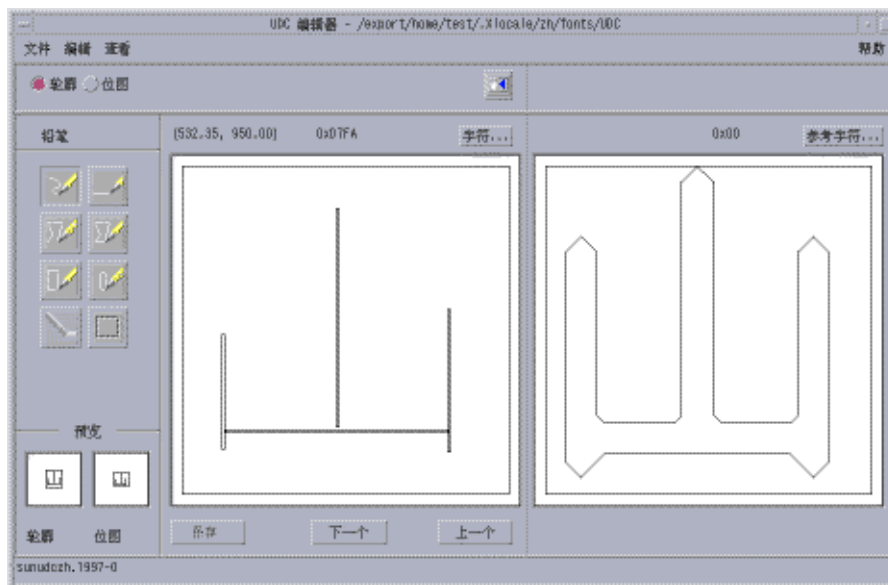
The Options dialog box appears.



2. Turn on **Append Font Path** to append the UDC font aliases to your font path.
3. Turn on **Save Font Path** to add the font files to the system font path.  
The new font files (located in `$HOME/.Xlocale/{locale}/fonts/UDC`) will be listed in the `$HOME/.OWfontpath` file and made available to the X server the next time you log in.  
If **Save Font Path** is off, the font files will still be saved in `$HOME/.Xlocale/{locale}/fonts/UDC`, but will not be added to the system fontpath the next time you log in.
4. Choose one or more of the following font sizes (in points) for the bitmap fonts: 12, 14, 16, 20, 24.

## ▼ Viewing the Character Reference

1. To view the character reference, click the icon below the Help menu.  
The following window appears.



2. To close the window, click the icon again.

## ▼ Register the UDC to NewPinYin input method

1. Choose Save from the File menu or choose "Open dictionary tool" from the File menu. The sdtudc\_register tool will be invoked.

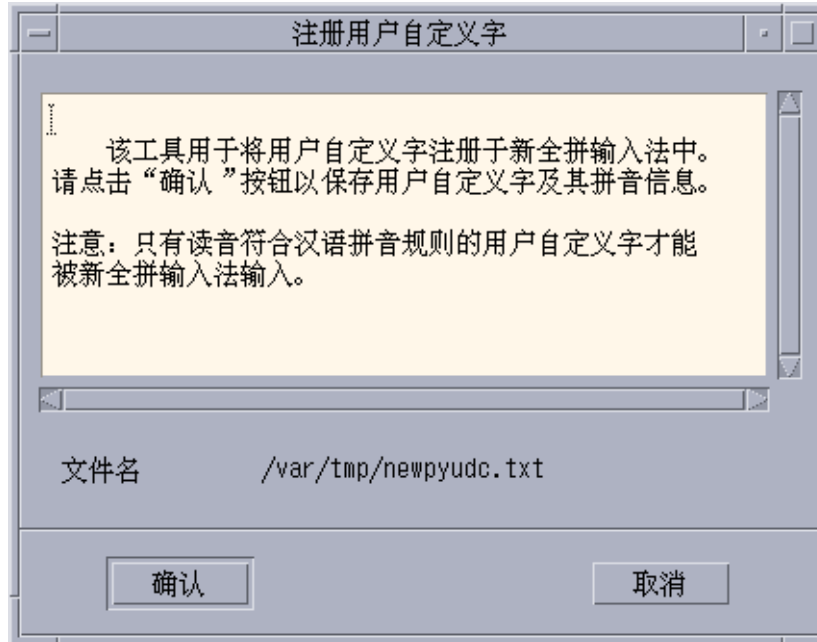


2. Input the pinyin in "reading" column. Please note that the only the valid pinyin can be registered to NewPinYin input method.



3. Click "Confirm" button

Following dialog is displayed.

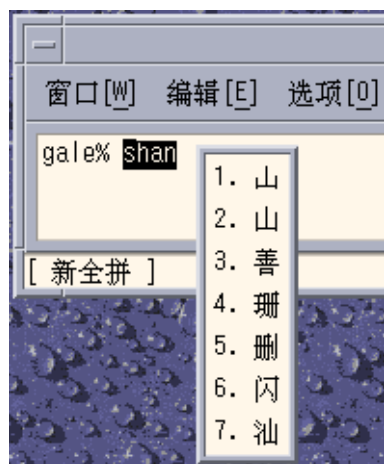


4. The UDC and its pinyin will be registered in `/var/tmp/newpyudc.txt`. Click "OK" in the dialog to confirm

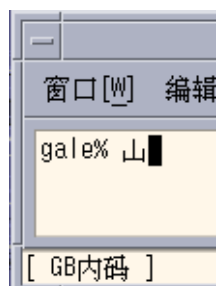
## ▼ Typing the New Characters

To type the characters you created, use the NewPinyin input method with the UDC's pinyin or use Neima input method.

1. Type "shan" in NewPinyin mode.



2. Type "D7FA" in GB Neima mode.



## ▼ Printing the New Characters

You can use "mp" utility to print the characters you created.

1. Set the fontpath for "mp" utility.

```
system% setenv UDC_BASE $HOME/.Xlocale/zh
```

2. Print a text file which contains UDC(s) by specifying the configuration file for UDC.

```
system% /usr/openwin/bin/mp -u/usr/lib/lp/locale/zh/mp/mp.conf.U  
udc.txt > udc.ps
```





# Simplified Chinese Printing Facilities

---

---

## Introduction

The Simplified Chinese Solaris operating environment supports printing Simplified Chinese output through the following types of printing facilities:

- Line printer containing built-in Simplified Chinese fonts
- PostScript-based printer

---

**Note** – Before you can print Simplified Chinese, a system administrator must set up your Simplified Chinese printing support as described in *Simplified Chinese Solaris System Administrator's Guide*.

---

You can use the Simplified Chinese Solaris `xetops`, `xutops` or `mp` utilities to print files containing Chinese text on a PostScript printer. These printing facilities can be used directly from a command line or from within Simplified Chinese Solaris applications as discussed in the following sections.

---

# Printing Chinese Output From a Command Line

## Printing With a Line Printer

The Simplified Chinese Solaris operating environment uses EUC code sets. Its printing applications (such as desktop tools) generate PostScript code that uses EUC. If you use different PostScript printing, make sure it has the same capabilities.

- To print an EUC file, use the command in the following example:

```
system% lp EUC_filename
```

- To print a GB format file on a printer that supports this format, use the following command:

```
system% lp -T GB GB_filename
```

For more information on setting up the GB filter, see *Simplified Chinese Solaris System Administrator's Guide*.

## Using xetops and xutops Utilities

The `xetops` and `xutops` utilities produce bitmapped graphics as printed images. Simplified Chinese Solaris software includes the `xetops` and `xutops` utilities so any system can print Simplified Chinese text on a PostScript printer. The `xetops` utility in EUC locale and `xutops` Utility in UTF-8 locale may no longer be supported in the future.

- `xetops` handles files in the EUC locale, such as zh
- `xutops` handles files in the UTF-8 locale, such as zh.UTF-8.

A typical command line for printing a file named `filename` containing Simplified Chinese characters with `xetops`, would be as follows:

```
system% pr filename | xetops | lp
```

```
system% xetops -h
```

The following syntax information is returned:

```
xetops [-r] [-B|[-T [-t S|F|K|H] [-s [i|m|c]
        [-p A3|A4|A5|B5|USLtr|USLg1]
        [-o length=i|c] [,width=i|c]]
[filename]
```

The options for `xetops` are as following:

- `-r`: prints the file in landscape mode.
- `-B`: prints the file using bitmap fonts. (The default is TrueType SongTi font.)
- `-T`: prints the file using TrueType fonts.
- `-t`: use this option to specify up to four typefaces. *S* indicates SongTi, *F* indicates FangSongTi, *K* indicates KaiTi, and *H* indicates HeiTi.
- `-s`: designates the size of each character. *i* stands for inches, *m* stands for millimeters, and *c* stands for centimeters. (The default is inches.)
- `-p`: specifies the paper size (A3, A4, A5, B5, US Letter, or US Legal).
- `-o`: specifies the length and width of the paper. *i* stands for inches, and *c* stands for centimeters. (The default is the number of the lines.)

Here are some examples of `pr` and `xetops`. The following command:

```
system% pr filename | xetops | lp
```

will print the file using the SongTi TrueType font. The printed characters will be 4mm in size.

The following command:

```
system% pr filename | xetops -t K -s 8m | lp
```

will print the file using the KaiTi TrueType font. The printed characters will be 8mm in size.

The following command:

```
system% xetops -t H -s 1.2 filename | lp
```

will print the file using the HeiTi TrueType font. The printed characters will be 1.2 inches in size.

The following command:

```
system% xetops -B filename | lp
```

will print the file using a bitmap font.

The syntax for `xutops` is similar:

```
system% pr filename | xutops | lp
```

Make *filename* the name of the file to print. This file can contain ASCII/English characters as well as Simplified Chinese.

## Printing with the mp Utility

### *Using the mp Utility*

The mp utility supports all asian locales. As a printing filter, mp generates a pretitfied version of contents in PostScript format. The Postscript output file contains glyph images from Solaris system-resident scalable or bitmap fonts, depending on each locale's system font configuration for mp. As a print filter, mp(1) is enhanced in the Solaris 9 environment to print either zh file in zh locale or UTF-8 files in zh.UTF-8 locales.

A typical command line for printing a file named *filename* containing Simplified Chinese characters with or without ASCII/English characters, would be as follows:

```
system% mp filename | lp -d printer
```

Make *filename* the name of the file to print. This file can contain ASCII/English characters as well as Simplified Chinese.

To print with the mp utility, type the following:

```
system% mp filename | lp
```

You can also use the utility as a filter, as the utility accepts stdin stream:

```
system% cat filename | mp | lp
```

You can set the utility as a printing filter for a line printer. For example, the following command sequence tells the printer service LP that the printer lp1 accepts only mp format files. This command line also installs the printer lp1 on port /dev/ttya. See the lpadmin (1m) man page for more details.

```
system% lpadmin -p lp1 -v /dev/ttya -I MP
system% accept lp1
system% enable lp1
```

You can add the lpfilter utility for a filter by using the lpfilter(1M) command as follows:

```
system% lpfilter -f filtername -F pathname
```

The lpfilter command tells LP that a converter (in this case, xutops) is available through the filter description file named pathname. The pathname can be determined as follows:

```
Input types: simple
Output types: MP
Command: /usr/bin/mp
```

The filter converts the default type file input to PostScript output using /usr/bin/mp.

To print a UTF-8 text file, use the following command:

```
system% lp -T MP UTF-8-file
```

For more details on the mp(1) command, refer to the mp(1) man page.



---

## Backward Compatibility

---

This appendix contains information for making programs backward-compatible with earlier versions of Asian Solaris Software. Every utility described is supported, but for this version of Solaris, you are encouraged to use the XPG4 internationalization APIs as described in *International Language Environments Guide*.

---

## Simplified Chinese Test Utilities

These utilities test various aspects of the Simplified Chinese (GB-2312-80) national standard character set. They also assume that the character being tested is part of the national standard character set.

The arguments for the functions in these tables must be a character in `WC`, `wchar_t`. For more information, see the `cctype(3x)` man page.

**TABLE A-1** Simplified Chinese Character Classification Functions

<b>Routine</b>	<b>Description</b>
<code>ischanzi</code>	Returns true if it is a Hanzi ideogram in GB-2312-80.
<code>iscaccent</code>	Returns true if it is an accent notation in GB-2312-80.
<code>iscphonetic</code>	Returns true if it is a phonetic symbol in GB-2312-80.
<code>iscpinyin</code>	Returns true if it is a Pinyin symbol in GB-2312-80.
<code>iscalpha</code>	Returns true if it is a Roman alphabetic in GB-2312-80.
<code>iscdigit</code>	Returns true if it is a Roman digit in GB-2312-80.
<code>iscnumber</code>	Returns true if it is a number in GB-2312-80.

**TABLE A-1** Simplified Chinese Character Classification Functions (Continued)

<b>Routine</b>	<b>Description</b>
<code>isclower</code>	Returns true if it is a Roman lowercase in GB-2312-80.
<code>iscupper</code>	Returns true if it is a Roman uppercase in GB-2312-80.
<code>iscblank</code>	Returns true if it is a white space character from GB-2312-80.
<code>iscspace</code>	Returns true if it is a space character from GB-2312-80.
<code>iscgen</code>	Returns true if it is a graphic or general symbol in GB-2312-80.
<code>iscsci</code>	Returns true if it is a scientific symbol in GB-2312-80.
<code>iscline</code>	Returns true if it is a ruled line symbol in GB-2312-80.
<code>iscunit</code>	Returns true if it is a unit character in GB-2312-80.
<code>iscparen</code>	Returns true if it is a right or left parenthesis in GB-2312-80.
<code>iscpunct</code>	Returns true if it is a punctuation character in GB-2312-80.
<code>iscgreek</code>	Returns true if it is a Greek character in GB-2312-80.
<code>iscrussian</code>	Returns true if it is a Russian character in GB-2312-80.
<code>iscspecial</code>	Returns true if it is a Greek or Russian character in GB-2312-80.
<code>ischira</code>	Returns true if it is a Japanese Hiragana character in GB-2312-80.
<code>isckata</code>	Returns true if it is a Japanese Katakana character in GB-2312-80.

Two additional routines for Simplified Chinese, `iscgb` and `isceuc`, test for characters from the GB-2312-80 character set. The `iscgb` routine expects a wide character, and `isceuc` expects a GB-2312-80 character in EUC format. For more information, see the `cctype(3x)` man page.

**TABLE A-2** General Simplified Chinese General Character Classification Functions

<b>Routine</b>	<b>Description</b>
<code>iscgb</code>	Returns true if it is in GB-2312-80.
<code>isceuc</code>	Returns true if it is a GB-2312-80 character in EUC format.



---

## Simplified Chinese Conversion Utilities

This section describes functions for wide character and string input and output, character classification, and conversion functions for the Simplified Chinese character sets. Solaris 2.7 software implements a wide character library for handling Simplified Chinese character codes according to industry standards.

Routines that have Chinese language-specific dependency are in their own language-specific library, which is linked with the corresponding C compiler option. Simplified Chinese Solaris `libc1e` is linked with `-lc1e`

Refer to the appropriate man pages for more information.

Asian Solaris software defines WC as a constant-width, four-byte code. WC uses the ANSI C data type `wchar_t`, which Solaris software defines in `wchar.h` as follows:

```
typedef long wchar_h;
```

In Solaris software, `long` is four bytes.

## Conversion Utilities

The conversion functions described in this section are available, but you should use `iconv()` as a standard function.

Simplified Chinese Solaris software provides facilities for various conversions, for example:

- Characters within a codeset, such as converting uppercase ASCII to lowercase.
- Between different conventions for national standard character sets, such as GB and EUC.
- Between code formats (such as converting between EUC and WC).

Programs using the general multibyte conversion utilities should include the header files `widec.h` and `wctype.h`. Simplified Chinese Solaris specific routines (such as `iscxxx`) are declared in `zh/xctype.h`.

Programs using general multibyte conversion utilities should include three header files: `wctype.h`, `widec.h`, and `zh/xctype.h`.

The `locale/xctype.h` file declares the Chinese locale-specific routines, which have names of the form `iscxxxx`:

As with the classification functions described in the previous section, the use of these functions can be controlled by the `setlocale` function (described elsewhere in this and other chapters).

Locale-specific conversion routines (such as Chinese `cgbtœuc`) are contained in the `libc1e` library:

This library can be linked during compilation using the C compiler option `-lc1e`.

## Conversion Within a Codeset

The multibyte conversion functions are similar to the one-byte conversion functions `toupper` and `tolower`. These functions convert wide-characters to other wide characters. For more information on conversion routines, see the man pages for `wconv(3)` and `cconv(3)`.

The following routines are in the regular Chinese C library:

**TABLE A-3** Simplified Chinese Case Conversion Functions (declared in `zh/xctype.h`)

Function	Description
<code>tocupper</code>	Converts codeset 1 Roman lowercase to uppercase
<code>toclower</code>	Converts codeset 1 Roman uppercase to lowercase

## Conversion Between Simplified Chinese Codesets

In the Simplified Chinese character sets, the Roman characters and numbers in codeset 0 are repeated in codeset 1. The following functions test wide characters.

**TABLE A-4** Simplified Chinese Codeset Conversion Functions

Function	Description
<code>atocgb</code>	Converts alphabetic or numeric characters in ASCII (codeset 0) to the corresponding characters in GB-2312-80 (codeset 1).
<code>cgbtoa</code>	Converts alphabetic or numeric characters in GB-2312-80 (codeset 1) to the corresponding characters in ASCII (codeset 0).

For further information on these functions, see the man page for `cconv(3)`.

## Conversion for Simplified Chinese Character Codes

The following routines do character-based code conversion on the GB-2312-80 character set. They convert characters and strings between EUC format and GB-2312-80 format. To use these routines, the library `libc1e` must be linked using the C compiler option `-lc1e`. For further information, see the `cconv(3)` man page.

TABLE A-5 Simplified Chinese Character-Based Functions

Function	Description
cgbtoeuc	Converts a character in GB-2312-80 format (7 bit) to EUC format
scgbitoec	Converts a string in GB-2312-80 format (7 bit) to EUC format
sncgbitoec	Converts part of a string in GB-2312-80 format (7 bit) to EUC format
euctocgb	Converts a character in EUC format to GB-2312-80 format (7 bit)
seuctocgb	Converts a string in EUC format in GB-2312-80 format (7 bit)
sneuctocgb	Converts a part of a string in EUC to GB-2312-80 format (7 bit)

---

## Binary Compatibility Package (BCP)

Applications compiled under Chinese OpenWindows 2.x or Solaris 1.x or SunOS 4.x systems have different binary formats than the current Chinese Solaris release. Older applications can nevertheless be run under the current Chinese release without being recompiled by using its included binary compatibility package (BCP).

---

**Note** – SUNWobcp must be included in your system configuration in order for you to run the following commands. See your system administrator for installation.

---

---

## Running Earlier Binary Code

The following BCP command runs the compiled binary code of earlier SunOS4.x /Solaris 1.x /Chinese OpenWindows 2.x applications without recompilation, although OpenWindows V2 Chinese applications display no input server status region. As shown in the following examples, the command calls the application by its old name (*old\_application\_name*) and sets the basic locale, input language, and display language using the older version's specific locale name (*oldlocale*):

```
system% old_application_name -lc_basiclocale oldlocale -lc_inputlang oldlocale \  
-lc_displaylang oldlocale
```

- For a current Simplified Chinese Solaris release system to run the compiled binary code of an earlier version of the `textedit` application enter, for example:

```
system% textedit -lc_displaylang chinese -lc_basiclocale chinese -lc_inputlang chinese
```

Incompatibilities between Simplified Chinese Solaris 2.x and 1.x applications do not permit cutting and pasting of Chinese characters between them.

# Glossary

---

<b>ANSI</b>	American National Standards Institute. ANSI proposes standard definitions for different computing languages. The most recent standard for the C language, prepared by the ANSI C X3J11 Committee, includes library functions for computing with multibyte characters for international usage, as well as a new data type, <code>wchar_t</code> , for dealing with four-byte characters. This standard is not completed, so it is referred to as the “proposed ANSI C standard,” or ANSI C-X3J11.
<b>ASCII</b>	American Standard Code for Information Interchange. A seven bit code containing English upper and lowercase letters, punctuation, numbers and control codes. The eighth bit in each byte is used by different applications for parity checking, communication and message passing protocols, compacting data, or other purposes. Applications that are intended to be internationalized cannot utilize this bit if they are going to use multiple code sets or multibyte characters, and utilities that handle multiple code sets or multibyte characters.
<b>Category</b>	In the Simplified Chinese Solaris documentation set, category is related to localization. A category is a portion of a country’s language representation and cultural conventions. For instance, the date is often represented in the U.S. as <i>Month, Day, Year</i> ; while in another country it might be <i>Day, Month, Year</i> . The date and time can be thought of as one category of a local language. Categories also refer to the program categories, the environment variables that are related to categories, and the ANSI localization tables for each category.
<b>Character Set</b>	A character set is defined as a set of elements used for the organization, control, or representation of data. Character sets may be composed of alphabets, ideograms, or other units. This may seem a bit open-ended, but character sets may contain other character sets, which makes the boundaries unclear.

<b>Code set</b>	Also called a coded character set, this is a set of unambiguous rules that establishes a character set and the one-to-one relationship between each character in the character set and its bit representation. For example, the English character set, including punctuation and numbers, can be mapped to the ASCII code set in such a way that each character corresponds to only one bit code, and no bit code corresponds to more than one character.
<b>EUC</b>	Extended UNIX Code. Describes four code sets modeled on ISO-2022. Each code set can contain one or more different character sets, like the Hangul and Hanja character sets in KS C 5601. The four code sets are referred to as codesets 0, 1, 2, and 3, and in this text they are sometimes abbreviated as cs0, cs1, cs2, and cs3. Other internationalization efforts sometimes call these g0, g1, g2, and g3. Codeset 0 is also called the primary code set, and codesets 1, 2, and 3 are called the supplementary code sets. In the Korean and Chinese implementations of the EUC codes, the primary code set (cs0) contains ASCII and begins with a zero in the most significant bit.
<b>ISO</b>	International Standards Organization. Composed of a number of professional societies and companies, this organization studies and makes recommendations on internationalization issues. ISO 2022 proposes and describes the Extended UNIX Codes. Other ISO proposals include the European 8-bit code and communication protocols for internationalization.
<b>Locale</b>	A locale describes a language or cultural environment. Its setting affects the display or manipulation of language-dependent features. Simplified Chinese Solaris software provides C for U.S.A and zh for Simplified Chinese.
<b>POSIX</b>	Portable Operating System for Computer Environments. An IEEE standards group comprising seven committees that create documents for standardizing and internationalizing UNIX. POSIX document 1003.1 deals with the kernel and system calls. 1003.2 concerns the C-shell and standard libraries. The other five deal with real-time computing, communications and networking, and other issues.
<b>Unicode</b>	The international character set and encoding developed by the Unicode Consortium.
<b>Wide Character Code (WC)</b>	A constant-width four-byte code, called WC in Asian Solaris documentation, for the internal representation of EUC codes using the new ANSI-C data type wchar_t. Although EUC does not specify limits on the size of the supplementary code sets (codeset 0 is always one byte), WC specifies a character as four bytes. Standardizing on four bytes takes up more memory space than necessary if the environment is primarily ASCII, but it also speeds processing time for strings of

mixed characters; the 1000th character always begins at byte 4000 (and the 0th character starts at byte 0). This is useful for any type of indexing in applications.

**X/Open**

X/Open started as a consortium of international UNIX vendors from Europe, USA, and Asia. It is now one of the major standards organizations like POSIX and ANSI; source of the X/Open System Interface Portability Guide.





# Index

---

## A

Address Manager, 12  
aliases  
    setting locales with, 16  
App Manager, 12  
ASCII  
    typing, 38  
ASCII keyboard, 12  
Audio Tool, 13

## B

backward compatibility, 103  
BDF, xtobdf, 78  
binary compatibility package (BCP), 107  
bitmap fonts, printing with, 99

## C

Calculator, 13  
Calendar Manager, 13  
candidates, 12  
characters, 15  
Chinese characters, 15  
Chinese Star keyboard mapping, 53  
CID fonts, 88  
Clock, 13  
code tables, 71  
codeset conversion, 106  
Common Desktop Environment (CDE), 15  
compatibility, backward, 103

Console, 13  
Control keys  
    for input modes, 38  
Control-spacebar toggle, 38  
conversion  
    in codesets, 106  
conversion manager, 12  
conversion utilities, 105  
converting  
    GB-2312-80 character codes, 106  
converting files, 79  
Convtool, 78  
.cshrc file, 17

## D

default locale, setting, 16  
desktop tools, 12  
dictionary files, 52  
Display PostScript System (DPS), 83  
.dtlogin file, 16  
dtwm, 11

## F

FangSongTi font, 99  
file conversion, 78  
File Finder, 13  
File Manager, 13  
files, conversion, 79  
font directories, changing, 17

fonts, 84  
  sdtudctool utility, 88  
  specifying in xetops, 99  
  UDC directory, 90  
Function keys  
  for input modes, 38

## G

GB-2312-80 code conversion, 106  
GB13000-1, 15  
GB2312 Code input method, 59  
GBK Code input method, 57, 61  
Golden input conversion mode, 38  
Graphical User Interface, 12  
GUI, 12

## H

HeiTi font, 99  
Help, 13  
htt  
  display, 20  
  startup script, 21  
  startup sequence, 21  
htt\_props, 20  
htt\_server, 19  
htt\_server, 20

## I

Icon Editor, 13  
iconv, 78  
iconv command, 79  
ideograms, 23  
Image Tool, 13  
input code tables, 71  
input method  
  creating new, 71  
  Golden, 38  
  New QuanPin, 38, 44  
  New ShuangPin, 44  
  ShuangPin, 53  
input method server (htt), 19

input methods  
  Location, 93, 95  
Intelligent\_ABC keyboard mapping, 53  
intermediate representation, 12  
ISO/IEC10646-1 characters, 15

## K

KaiTi font, 99  
keyboard mappings, 53  
keyboards, 12  
keys, input conversion control, 38

## L

locales  
  overview, 15  
  setting, 16  
  setting default, 16  
Location code input mode, 93, 95  
lookup choice area, 24

## M

Mailer, 13  
mailx utility, 77  
Motif, 11

## N

New QuanPin, 38  
New QuanPin input method, 44  
New ShuangPin input method, 44

## O

OpenWindows  
  V2 applications, 107

## P

paper size, with xetops utility, 99

Performance Meter, 13  
PostScript fonts, 83  
preedit area, 12, 23  
Print Manager, 13  
printing, 97  
    using xetops print filter, 98  
    with xetops utility, 99  
Process Manager, 13  
.profile file, 16  
PSWDisplayText(), 86  
PyCiku.dat dictionary file, 52

## Q

QuWei input method, *See* Location code input mode

## R

reverse video, 12

## S

scripts  
    htt startup script, 21  
sdtudctool, using, 88  
separators, 52  
ShuangPin input method, 53  
Simplified Chinese text, input modes, 23  
Snapshot, 13  
Song fonts, 84  
SongTi font, 99  
status area, 12, 24  
Style Manager, 13

## T

talk command, 78  
TCA\_CNS 11643-92, 15  
terminal window, 13  
text  
    Chinese, 23  
    intermediate representation of, 12  
Text Editor, 13

tools  
    desktop, 12  
TrueType fonts, printing with, 99

## U

Ud.Ciku.dat dictionary file, 52  
UDC directory, 90  
utilities  
    conversion, 105

## W

window environments, 15  
Window Manager, 11  
windows  
    Simplified Chinese characters allowed  
        in, 23  
    status area, 12

## X

xetops print filter, 98, 100  
    examples, 99  
Xlib, 11  
xtobdf command, 78  
xutops print filter, 98

## Z

zh locale, 15  
ZiRanMa keyboard mapping, 53

