



Traditional Chinese Solaris User's Guide

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

Part No: 816-0669-10
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 REpondre 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

Preface	7
1 Introduction to Traditional Chinese Solaris Software	11
Introduction	11
Design of Traditional Chinese Solaris Software	11
Traditional Chinese Graphical User Interface	12
Traditional Chinese Input/Output	12
The Localized CDE Desktop	12
2 Starting the Traditional Chinese Solaris Software	15
Setting the Default Locale	15
Using Aliases to Set Locales	16
Adding Font Directories	17
3 Using the Input Method Server	19
Input Method Server Basic Properties	19
Three <code>htt</code> Processes	20
How and When <code>htt</code> Is Started	20
Using the <code>iiim</code> Server in <code>zh_TW.UTF-8</code> Locale	20
4 Entering Traditional Chinese Text	23
About This Chapter	23
Character Sets	24
Input Window Areas	25

Preedit Area	25
Status Area	25
Lookup Choice Area	26
Auxiliary Windows	26
Usage of Traditional Chinese Input Methods	26
Base Operating	26
Input Method Utilities	30
▼ Selecting the Utility Menu	31
Input Method Selection Tool	31
Setting Input Method Options	32
Lookup table	34
Lookup Table for Special Characters	35
Virtual Keyboard	36
User Defined Characters (UDC)	39
Input Method Help	40
Typing ASCII Characters	41
Array Input Mode	41
Typing Array Text	41
Boshiamy Input Mode	43
Typing Boshiamy Text	43
DaYi Input Mode	45
Typing DaYi Text	46
New ChuYin Input Method	48
Typing New ChuYin Text	48
ChuYin Input Mode	59
Typing ChuYin Text	59
I-Tien Input Mode	61
Typing I-Tien Input	61
Telecode Input Mode	63
Typing Telecode Input	63
TsangChieh Input Mode	65
Typing Characters in TsangChieh Mode	65
ChienI Input Mode	67
Typing ChienI Mode Input	67
NeiMa Input Method	69
Typing with MeiMa Input Method	69
English_Chinese input method	71
Typing English_Chinese Characters	71

	CodeTable Input Method Interface	73
	Introduction	73
5	Localized Applications	79
	Using the mailx Utility With Traditional Chinese Characters	79
	Using talk With Traditional Chinese Characters	80
	The xtobdf Utility	80
	The runb5 Utility	80
	Sdtconvtool	80
	iconv	81
6	Fonts	83
	Display PostScript System (DPS)	83
	Using Traditional Chinese PostScript Fonts and DPS Facilities	83
	Creating Composite Roman and Traditional Chinese Fonts	85
	Using Traditional Chinese Fonts in DPS Programming	86
	Converting From BDF to PCF Format	87
	Installing and Checking the Edited Font	87
7	Traditional Chinese Printing Facilities	89
	Printing From a Command Line	89
	Printing With a Line Printer	90
	Printing With the xetops Utility	90
	Printing With the xutops Utility	90
	Printing with the mp Utility	91
A	Binary Compatibility Package	93
	Traditional Chinese Test Utilities	94
	Asian-Specific Utilities	95
	Conversion Utilities	95
	Conversion Within a Codeset	96
	Conversion for Traditional Chinese Character Codes	96

B Mapping Traditional Chinese Keyboard Functions 99

Glossary 101

Index 105

Preface

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

Who Should Use This Book

This user's guide is for someone who wants to use the Traditional Chinese features of Solaris software to manage files, calendar, and e-mail, write or print Traditional Chinese files, and so forth. Tools for these and many other applications run under Traditional 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 Traditional Chinese Solaris software before
- For information on using product features
- If you need instructions for starting up your Traditional Chinese Solaris operating environment (see especially Chapter 2, "Starting the Traditional Chinese Solaris Software.")

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 xiii. This user's guide focuses on using the Traditional Chinese features of the Desktop Tools and other features of Traditional Chinese Solaris software.

How This Book Is Organized

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

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

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

Chapter 3, "Using the `http` Input Method Server," introduces the startup, appearance, and use of `http`.

Chapter 4, "Entering Traditional Chinese Text," describes different Traditional 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 *Traditional Chinese Solaris System Administrator's Guide*.)

Chapter 5, "Localized Applications," describes uses of two desktop tools localized for Chinese: `mailx`, `talk`, and tools to convert file codes.

Chapter 6, "Fonts," explains how to use Display PostScript System (DPS) in Traditional Chinese Solaris environment.

Chapter 7, "Traditional Chinese Printing Facilities," discusses Traditional Chinese Solaris support for line printers with built-in Chinese fonts or using the Traditional Chinese Solaris `xetops`, `xutops` and `mpfilter` utilities.

Appendix A, "Binary Compatibility Package," discusses running compiled binary code of earlier SunOS™ 4.x/Solaris 1.x applications without recompilation.

Appendix C, "Mapping Traditional Chinese Keyboard Functions," discusses how to configure a Sun Chinese keyboard to make selected key functions when you need them.

The Glossary contains a list of words and phrases found in the Traditional 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 wanting to customize their system environment or the operations of their Sun tools will find much pertinent information in *International Language Environments Guide* and *Traditional Chinese Solaris System Administrator's Guide*. These books give 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.

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>
AaBbCc123	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 rm <i>filename</i> .
<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.

Shell	Prompt
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 to Traditional Chinese Solaris Software

Introduction

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

Design of Traditional Chinese Solaris Software

Traditional 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 Traditional Chinese Solaris software. These products introduce Solaris input methods for the input and output of Traditional Chinese. Application programs and CDE use the features of Traditional Chinese CDE to communicate with users in Traditional Chinese.

This Traditional 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 Traditional Chinese.
- A localized development environment, which programmers use to develop localized applications, with `Xlib` and `Motif`, which have been internationalized for this use. Programmers and developers should refer to *International Language*

Traditional Chinese Graphical User Interface

This Traditional Chinese Solaris release uses the Traditional Chinese CDE Motif graphical user interface, which is similar in layout and design to the U.S. release of CDE. Traditional Chinese CDE supports multibyte characters and Traditional 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 can take Traditional Chinese input include a *status area* associated with their input window to show the current conversion mode. With an input conversion mode on, as Traditional Chinese is being typed its entry point becomes a highlighted (reverse video) *preedit area* until the input is converted to Traditional Chinese or special characters and committed. Some input modes also provide conversion choices among several Traditional Chinese characters on menus.

Traditional Chinese Input/Output

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

Traditional Chinese input at 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. All can handle Traditional Chinese input and output. 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.

Application Manager – Contains the applications available on your system. You can launch these applications through the Application 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.

Front Panel – A centrally-located window containing controls for accessing applications and utilities, including the workspace switch. The Front Panel occupies all workspaces.

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[®] commands at a system shell prompt and other terminal operations.

Starting the Traditional Chinese Solaris Software

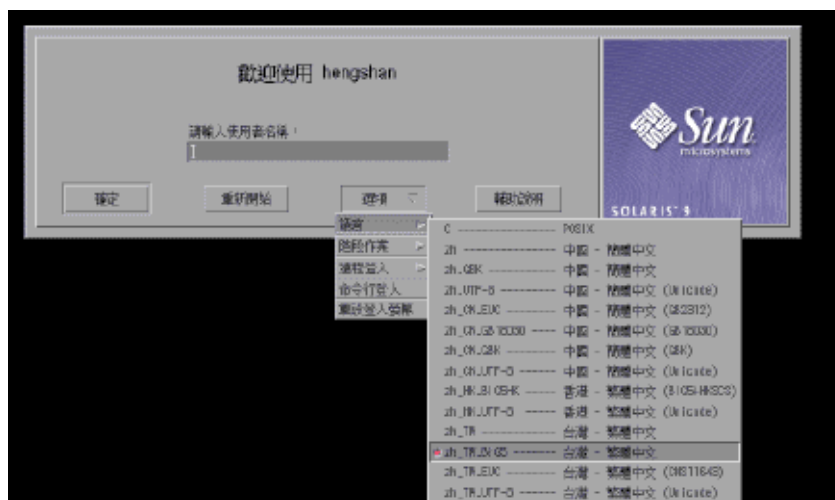
The Traditional Chinese Solaris operating environment must be specially set up for using Traditional Chinese text facilities. This chapter describes the steps required to set up the Traditional Chinese environment and to start Traditional Chinese Solaris operation.

The Traditional Chinese Solaris product includes the following locales:

- `C` – ASCII English environment.
- `zh_TW` – Traditional Chinese environment in extended UNIX code (EUC) CNS 11643-1992 standard.
- `zh_TW.BIG5` – Traditional Chinese environment in the Big5 codeset.
- `zh_HK.BIG5HK` – Traditional Chinese environment in Big5-HKSCS code that supports the Hong Kong Supplementary Character Set (HKSCS), which is a supplementary character set of the Big-5 and ISO 10646 coding schemes.
- `zh_HK.UTF-8` – Traditional Chinese (Hong Kong) environment in Unicode 3.1.
- `zh_TW.EUC` – Symbolic link to `zh_TW` locale.
- `zh_TW.UTF-8` – Traditional Chinese environment in Unicode 3.1 standard.

Setting the Default Locale

You can change your default locale using the following procedure.



- Choose the language button on the dtlogin window.
- Select the C, zh_TW, zh_TW.BIG5, zh_TW.UTF-8, zh_HK.BIG5HK or zh_HK.UTF-8 locale.

Your new locale is now in effect

Using Aliases to Set Locales

You can use aliases to change a terminal-emulation window between the Traditional Chinese locales and ASCII/English locale from time to time without typing long command lines or editing your `.dtlogin` file and running `source` every time.

Setting Up Locale Aliases for the Korn Shell

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

```
alias twn_EUC "export LANG=zh_TW; /bin/stty cs8 -istrip defeucw"
alias twn_BIG5 "export LANG=zh_TW.BIG5; /bin/stty cs8 -istrip defeucw"
alias asc "export LANG=C; /bin/stty cs7 istrip defeucw"
alias twn_UTF-8 "export LANG=zh_TW.UTF-8; /bin/stty cs8 -istrip defeucw"
```

Setting Up Locale Aliases for the C Shell

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

```
alias twn_EUC "setenv LANG zh_TW; /bin/stty cs8 -istrip defeucw"
alias twn_BIG5 "setenv LANG zh_TW.BIG5; /bin/stty cs8 -istrip defeucw"
alias asc "setenv LANG C; /bin/stty cs7 istrip defeucw"
```



```
alias twn_UTF-8 "setenv LANG=zh_TW.UTF-8; /bin/stty cs8 -istrip defeucw"
```

Adding Font Directories

The `Xsession` script (located in `/usr/dt/bin`) that comes with the Traditional Chinese Solaris operating environment includes the following font path: `/usr/openwin/lib/locale/locale/X11/fonts`, where `locale` is either `zh_TW`, `zh_TW.BIG5` or `zh_HK.BIG5HK`. To add a different font directory path dynamically, type:

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

```
system% xset fp rehash
```


Using the Input Method Server

The input method server (IM Server) handles Traditional Chinese input for Traditional Chinese Solaris software. The IM Server receives keyboard input and makes the input available as Traditional Chinese characters to Traditional 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 Traditional Chinese text using Solaris applications, you do not need to read beyond the first section of this chapter.

Input Method Server Basic Properties

The Traditional 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.

Traditional Chinese Solaris applications that use IM Server for Traditional 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_xbe`, and `htt_server`) 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_xbe` and `htt_server` are started automatically.

How and When `htt` Is Started

An application that uses the `htt` input method server to receive Traditional Chinese characters must find `htt` running before it can receive such input. So 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 twle -xim htt_xbe &
```

With the locale set to `zh_TW` or `zh_TW.BIG5`, `htt` is started automatically when the Traditional Chinese windowing environment starts. 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 upon a user’s login from `dtlogin`. This script ensures that `htt` is started before other applications in CDE.

Using the `iiim` Server in `zh_TW.UTF-8` Locale

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
- Hindi
- Thai
- Arabic
- Hebrew
- Unicode Hex
- Unicode Octal
- Lookup
- Traditional Chinese

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

- Japanese — if `ja` locale is installed
- Korean— if `ko` locale is installed.
- Simplified Chinese — if `zh` locale is installed.
- Traditional Chinese (Hong Kong) — if `zh_HK.BIG5HK` locale is installed.

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

- F2 - TsangChieh input mode
- F3 - Array input mode
- F4 - ChuYin input mode
- F5 - Boshiamy input mode
- F6 - ChienI input mode
- F7 - I-Tien input mode
- F8 - DaYi input mode
- F9 - Cantonese input mode
- F10 - English-Chinese input mode
- F11 - TeleCode input mode
- F12 - NewChuYin input mode

For more detailed information about each Input Method, see “Chapter 4, Entering Traditional Chinese Text”.

Entering Traditional Chinese Text

About This Chapter

This chapter describes Traditional Chinese Solaris input modes for typing Traditional Chinese characters. Solaris 9 provides more friendly and extensible input methods and management tools for Chinese Solaris users, including Mainland, HongKong and Taiwan users. Also an Input method auxiliary window has been developed which supports the following new functions and utilities:

- Input methods switching
- Chinese Full_Width/Half_Width Character Mode Switching
- Chinese punctuation/English punctuation Mode Switching
- Input method management panel: (Input method properties setting, Input methods selection by user)
- Lookup tables for BIG5/EUC_TW/Big5–HKSCS/Unicode characters.
- Virtual Keyboard

Two kinds of input methods are supported:

- input method based on codetable: such as TsangChieh, ...
- input method developed by vendor: such as NewChuYin, NeiMa, ...

Input methods with an auxiliary window supports all the Traditional Chinese locales:

- zh_TW/zh_TW.EUC
- zh_TW.BIG5
- zh_TW.UTF-8
- zh_HK.BIG5HK
- zh_HK.UTF-8

In all Traditional Chinese locales, the following input methods are supported:

- New ChuYin
- ChuYin
- TsangChieh
- Array
- Boshiamy
- DaYi
- ChienI
- Cantonese
- EUC NeiMa
- BIG5 NeiMa
- BIG5_HKSCS NeiMa
- English-Chinese
- Optional Codetable Input Methods (such as PinYin, ...)

You can type any Traditional Chinese characters, ASCII/English, and special symbol characters in the input area of the following applications:

- Terminal emulation (TTY) windows, such as a Terminal
- Text entry subwindows, such as those used by Text Editor or Mailer
- Control panel subwindows, like File Manager, for typing a file name
- Other special use subwindows, such as pop-ups

Character Sets

The Traditional Chinese Solaris operating environment conforms to the CNS 11643 standard, updated in late 1992, and the Big5 character sets.

Big5 was defined by five major Taiwanese computer vendors (including the Institute of Information Industry) in May, 1984. Although Big5 is not the national standard, it is more widely used than the CNS 11634-1992.

The total number of characters defined in Big5 is 13,461. It is part of CNS 11643-1992.

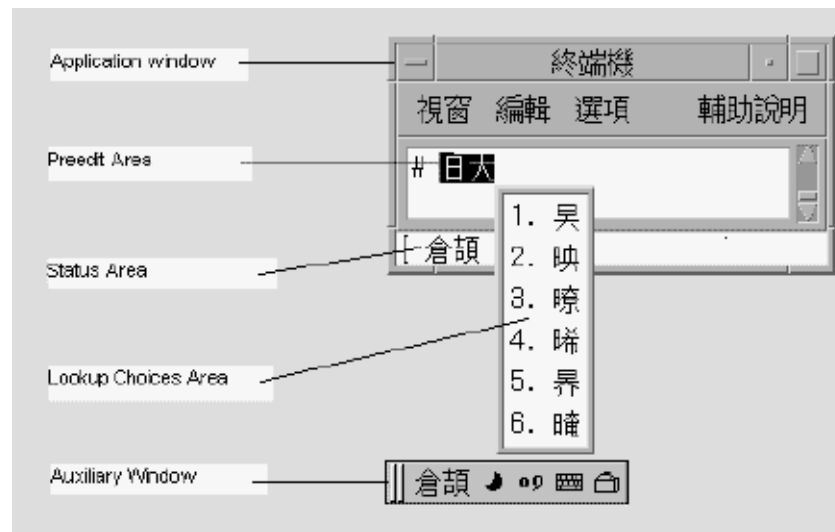
Traditional Chinese Solaris software provides code conversion between Chinese code conventions at three levels of support:

- User commands support file transfers for existing files in different codes.
- Library functions support application development for existing codes.
- STREAMS modules support existing TTY devices using different codes.

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 (reverse video or underlined) entry display area
- Status area indicating the current input/conversion mode
- Auxiliary Windows— Utilities for input method management
- Lookup choice area displaying multiple character choices



Predit Area

The highlighted (reverse video or underlined) preedit area displays characters as they are typed or converted. It holds formations of text before it is converted to an ideogram or symbol and put in the text block being assembled for the application.

Status Area

The status area shows which 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 Chinese or special character choices available for conversion of the character(s) or radical(s) in the preedit area. In the above example, it is a pop-up menu.

Auxiliary Windows

The auxiliary windows provide tools and utilities to manage input methods or to make the input simpler.

Usage of Traditional Chinese Input Methods

Base Operating

▼ Switching Between English Mode and Chinese Mode Windows

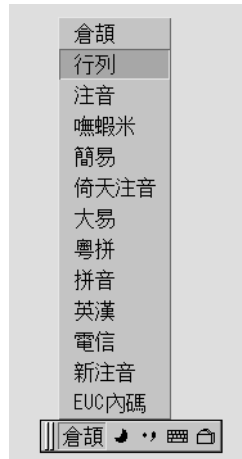
- Type "CTRL+SPACE",
An auxiliary window appears, as below.



▼ Select an Input Method

1. In Chinese status window, type Function key "F2" or "CTRL + Esc" to switch to the first input method, and F3 to switch to the second and so on.

Or click the input method selection button on the auxiliary window. The input method selection menu appears as below:



2. Select the input method you want to use.

▼ Switching Between Half_width Character Mode and Full_width Character Mode

Use 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.

The input method system is in Full_width Character Mode when the button appears as below:



The input method system is in Half_width Character Mode when the button appears as below:



When in Full_width mode, the Full_width character of the input key will be committed to system. For example: inputting 'a' when in Full_width mode, commits the fullwidth character of 'a' to the application as shown below:



▼ Switching Between Chinese Punctuation Mode and English Punctuation Mode

Use the Chinese Status window.

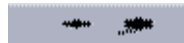
- **Type "CTRL+." to switch between Chinese Punctuation Mode and English Punctuation Mode.**

Or click the Chinese/English Punctuation Button of auxiliary window to switch modes.

indicates the input method system is in Chinese Punctuation Mode.



indicates the input method system is in English Punctuation Mode.



When you type a function key in Chinese Punctuation mode, the corresponding Chinese punctuation character is committed to the application. For example: when you type "\$" in Chinese Punctuation mode, the Traditional Chinese currency symbol character "NT\$" is committed to the application as shown below:



The punctuation key includes these characters: , . / > ; : " \ \$! ^ _ -

The map between English and Chinese punctuation appears as follows:

!	!
"	"
\$	NT\$
&	—
'	'
((
))
,	,
.	o
:	:
;	;
<	<<
>	>>
?	?
@	.
\	\
^	...
_	—

▼ Toggling between Input methods

Use the Chinese Status window.

- Type "CTRL+ESC" to toggle to another input method.

Input Method Utilities

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

The following tools are supported:

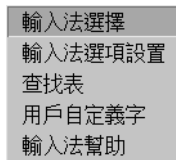
- Input method selection
- Input method option setting
- Lookup table
- User define character
- Input method help
- Virtual Keyboard

▼ Selecting the Utility Menu

- Click the utility button



to display the utilities menu.



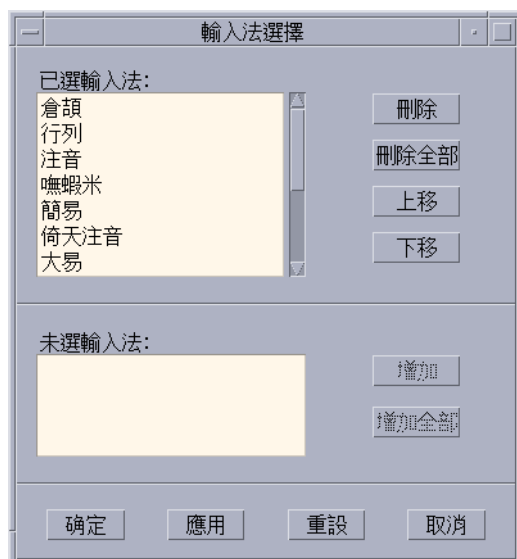
Select one of the input method tools from the menu.

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:

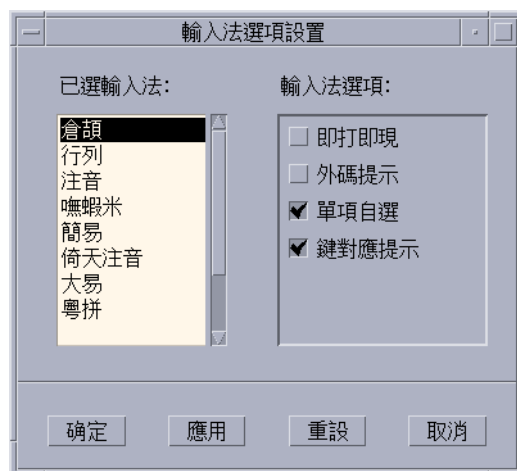


After selecting an input method, click "OK" or "Apply", the setting will be activated. The first input method selected becomes the default input method.

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 input method options setting panel appears as below: Select the Input Method Options Setting item from the utilities menu



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

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

Display candidates key by key:

- If this option is selected, each time a valid key is entered, the input method will search the dictionary table and display the candidates in the Lookup window.
- If this option is not selected, each time a valid key is entered, the input method only displays the key in the preedit area. If the "SPACE" key is pressed, the input method engine searches the dictionary table, and displays the candidates.

Display external codes:

- if this option is selected, the external codes for every candidate will appear in each Lookup window after the candidate is entered.
- if this option is not selected, the external codes for each candidate will not be displayed after the candidate.

This option can be used to study the input method. For example, to visualize the external code of a Chinese character in that input method.

Automatically commit if only one candidate

- if this option is selected, only one candidate for the external code, IM will automatically commit it

- if this option is not selected, IM will display it in Lookup window.

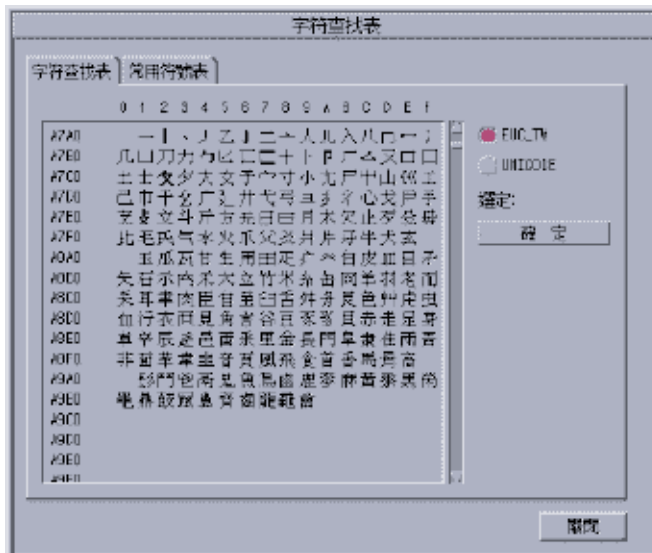
Display keymap character for every external code

- if this option is selected, when a valid key is entered the corresponding keymap character of the key will display in the Preedit area.
- if this option is not selected, the keymap character will not be displayed, only the key.

Lookup table

The lookup table tools can be used to search for and input Chinese characters. Three kinds of lookup tables are provided:

- A lookup table with a native encoding set, such as: in zh_TW/zh_TW.EUC locale, a lookup table with EUC_TW encoding, and in zh_TW.BIG5 locale, a lookup table with BIG5 encoding is provided. The native encoding lookup table panel appears as below:



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



Lookup Table for Special Characters

The lookup table for special characters, such as Greek characters, Mathematical symbols, etc. appears as below:



Virtual Keyboard

Virtual Keyboard tools can be used as a Lookup utilities to simplify the input of some special symbols. They can also can be used to display the keymap of input methods that are based on radicals, such as TsangChieh input methods, which is a typical input method that is based on Chinese radicals.

The Simplified Chinese environment supports these virtual keyboards:

- PC Keyboard
- Tsangchieh Keyboard
- Array Keyboard
- DaYi Keyboard
- ChuYin Keyboard
- I-Tien Keyboard
- Boshiamy Keyboard
- ChienI Keyboard

Virtual keyboards provide input by clicking on a character.

PC Keyboard

The PC Virtual Keyboard looks as below:



TsangChieh Keyboard

The TsangChieh Keyboard looks as below:



Array Keyboard

The Array Virtual Keyboard looks as below:



DaYi Keyboard:

The DaYi Virtual Keyboard looks as below:



ChuYin Keyboard:

The ChuYin Virtual Keyboard looks as below:



I-Tien Keyboard:

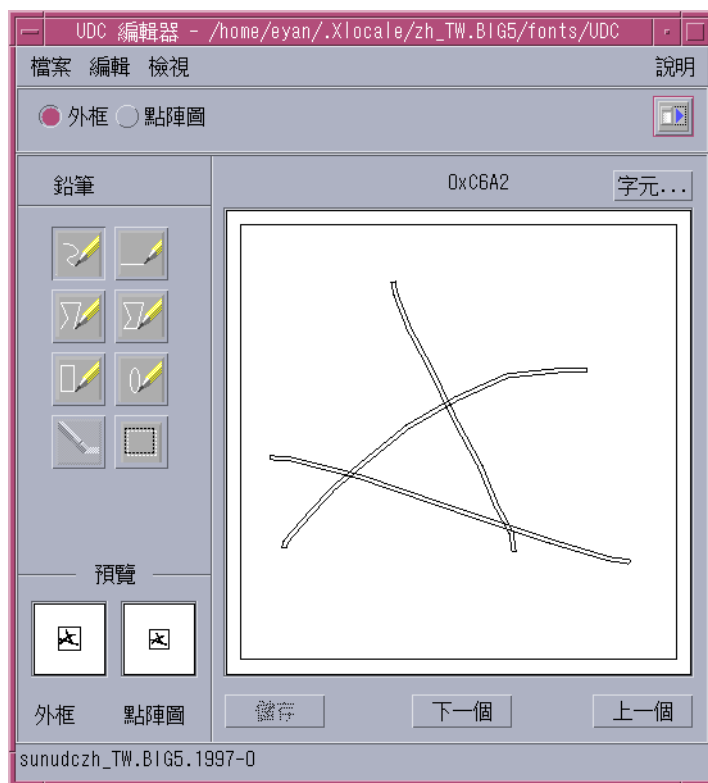
The I-Tien Virtual Keyboard looks as below:



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.**



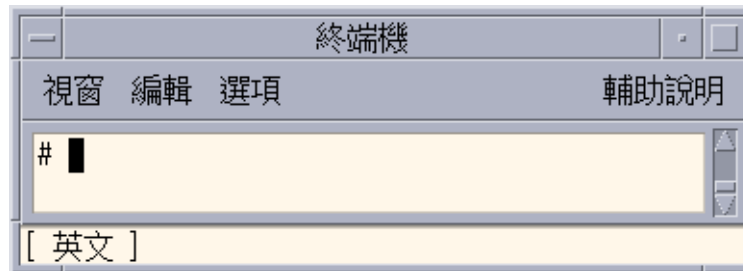
Input Method Help

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.

Typing ASCII Characters

Each tool first starts in ASCII input mode, which you can toggle on or off by pressing Control-spacebar or the Chinese/English key on a Chinese keyboard. Use this mode to type ASCII keyboard characters, such as the echo command shown in the following figure:



[英文]

is displayed in the window's status area when this mode is on, when ASCII input mode is off, their current conversion mode symbol appears.

Array Input Mode

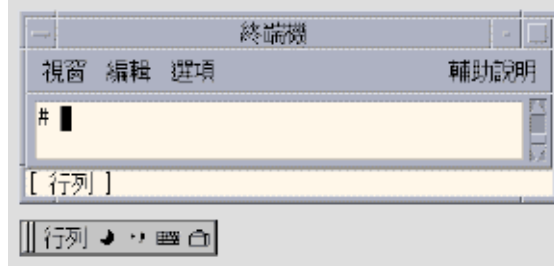
Array input mode works on the radicals that compose Chinese characters. It contains 260 radicals and allocates them on the keyboard in ten columns and three rows. To used this input mode effectively, you need to memorize only ten basic keystrokes. All radicals are naturally allocated on the keyboard by the basic keystrokes.

Typing Array Text

This section contains instructions on how to use the array input method to input Chinese characters.

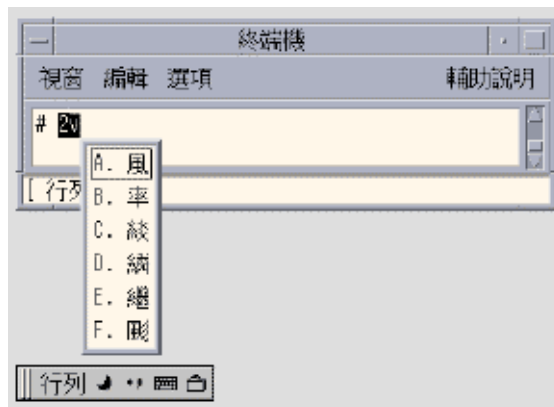
1. Type Control-spacebar to turn on Chinese input conversion.

The mode appears in the status area.



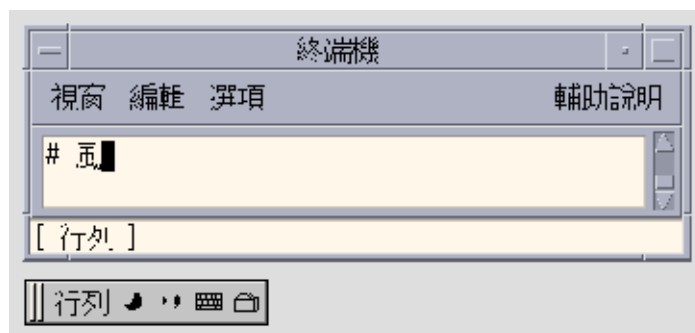
2. Type x.

The Array input converter finds several matching characters, so it automatically displays a lookup choice window:



3. Type [Shift + a] to select the appropriate character.

The radicals are converted and the corresponding character appears:



Boshiamy Input Mode

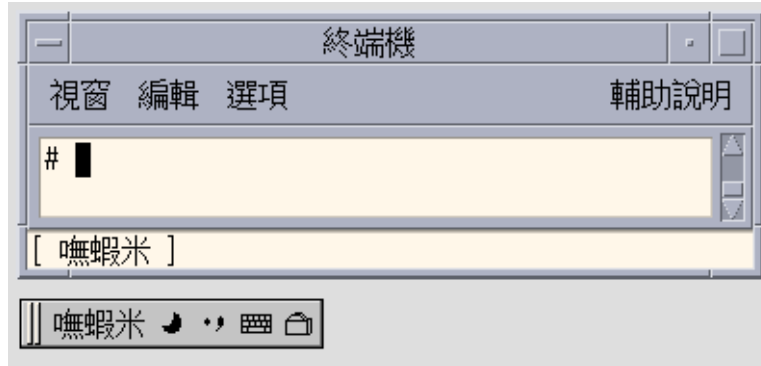
Boshiamy input mode works on the radicals that compose Chinese characters. The radicals in this mode are easy to memorize by their shape, sound, or meaning. For effective use of this mode, memorize the rules: Cut off long and Patch up short.

Typing Boshiamy Text

This section contains instructions on how to use the Boshiamy input method to input Chinese characters.

1. Type Control-spacebar to turn on Chinese input mode.

The mode appears in the status area.



2. Type the sets of radicals: uoo. Press the spacebar to convert it to the corresponding character.

The input converter finds five matching characters and displays a lookup choice window:



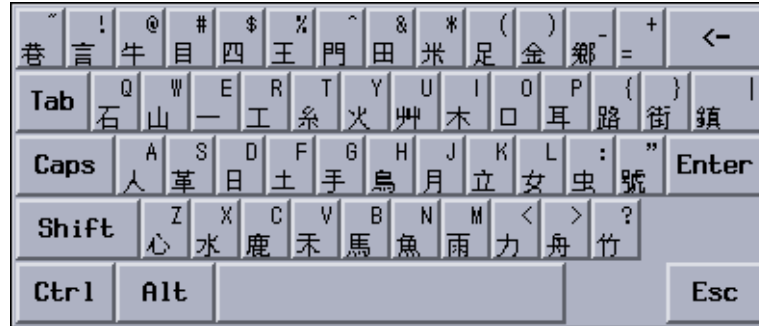
3. Type a number key to select the appropriate character, such as 1 to select the first candidate:



DaYi Input Mode

Similar to the TsangChieh input mode, DaYi conversion works on the radicals that compose Chinese characters. The keystroke sequence of the DaYi input mode completely follows the Chinese writing sequence, and only a few rules need to be memorized.

The keyboard mapping of DaYi input mode appears as below:



Typing DaYi Text

1. Type Control-spacebar to turn on TsangChieh mode.

The mode appears in the status area.

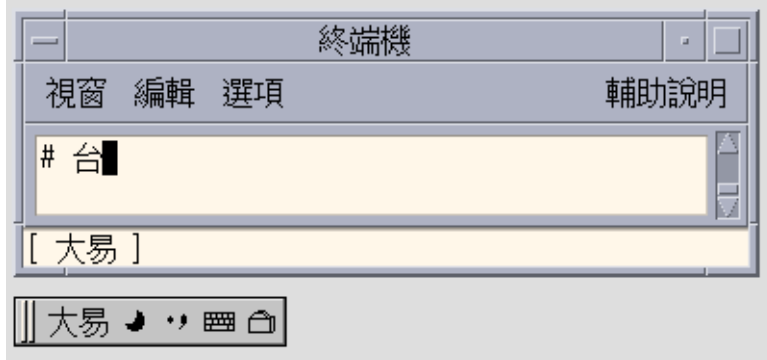


2. Type the set of radicals; o Press the spacebar after each set to convert it to a character.

The DaYi input converter finds six matching characters and displays a lookup choices as below:



3. Type a number key to select the appropriate character, such as 1 to select the first candidate:



New ChuYin Input Method

This section describes the features in the New ChuYin input methods, and how to use some of the features in Traditional Chinese locales, including zh_TW.EUC/zh_TW.BIG5/zh_TW.UTF8/zh_HK.BIG5HK/zh_HK.UTF-8.

The New ChuYin provide the following new features:

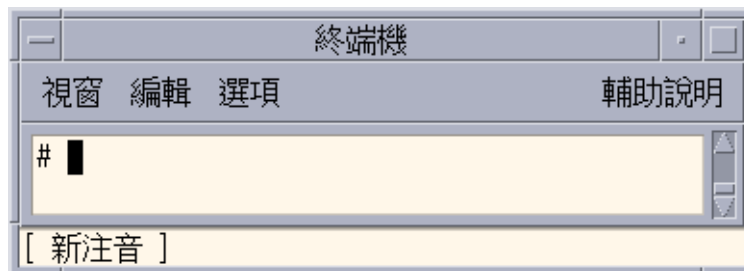
- Support phrase input
- Support memorizing and recalling user defined phrases
- Support simplified ChuYin by only inputting consonant
- Editable pre edit string
- Adjust frequency of phrase and single Chinese character automatically

Typing New ChuYin Text

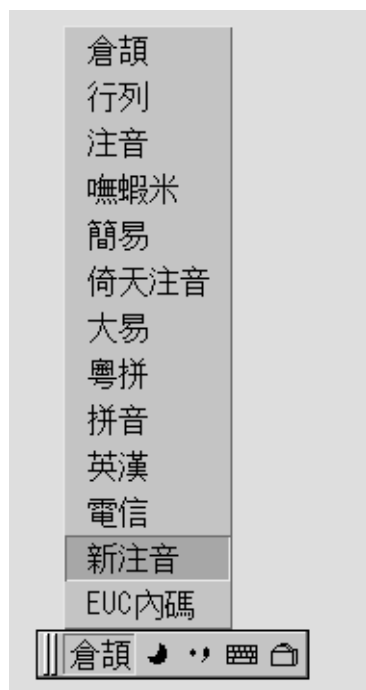
- **Select the New ChuYin Input method**

To switch to English input mode, press [Control] + [Space].

The New ChuYin Status window appears as below:



When another input method is being used, New ChuYin can be invoked by pressing [Control] + [ESC]. Or select New ChuYin from the input method menu on the auxiliary window as shown below:



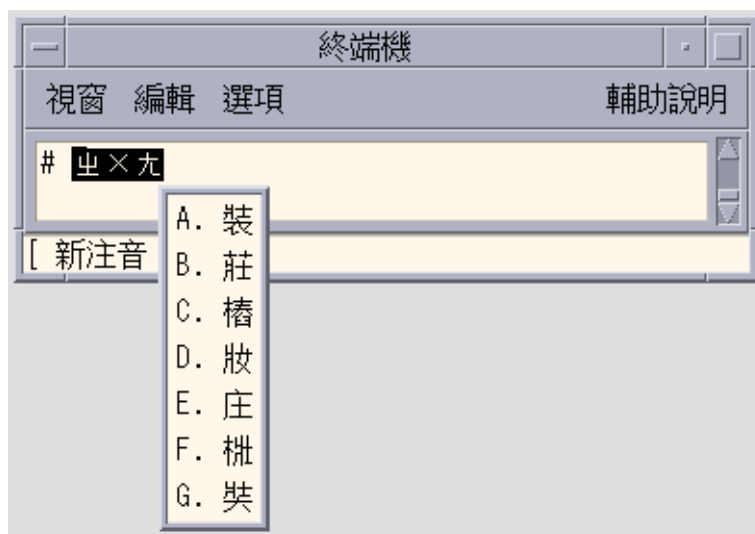
▼ Entering a Single Character

To enter the character:



1. Type key sequence 5j.

The pre-edit and lookup table appear as below:



2. Type the tone 4.

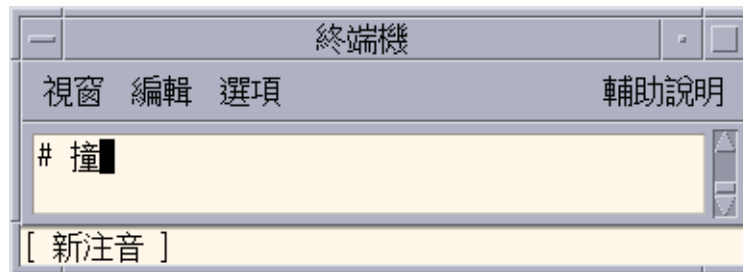
The pre-edit and lookup table changes to the following screen.



Note – Characters can also be input by searching through existing pages instead of inputting the tone. For example, if the tone of a Chinese character is "Tone 1", the choices for "Tone 1" will be shown before each choice. Use the [Space] key to commit the first candidate.

3. Type [Shift] + [D] to commit the forth candidate.

The character appears as below:



▼ Typing Text Phrases

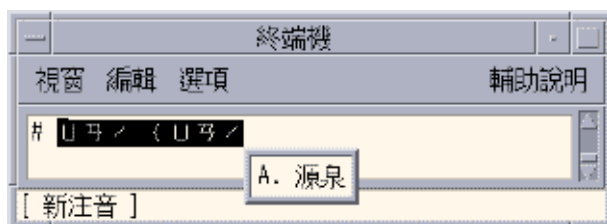
New ChuYin input method provides intelligent word and phrase inputting functions. While typing ChuYin sequences, the input method automatically locates the related words and phrases in a system dictionary.

For example, to enter the phrase

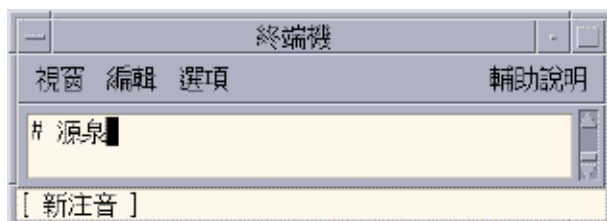
源泉

1. Type the key sequence `m06fm06`

The pre-edit and lookup tables are shown as below:



2. Type [Space] or [Shift] + [A] to commit the selection.



▼ Defining New Phrases

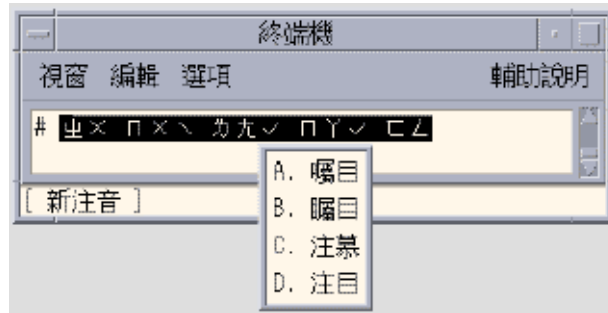
Single phrases can be defined with up to 8 Chinese characters.

For example, to define

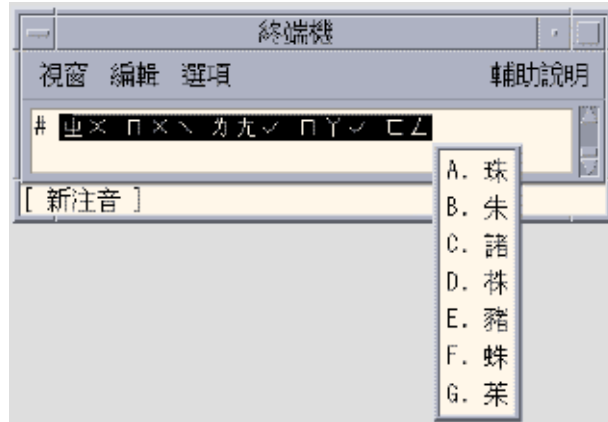
珠穆朗瑪峰

as a new phrase.

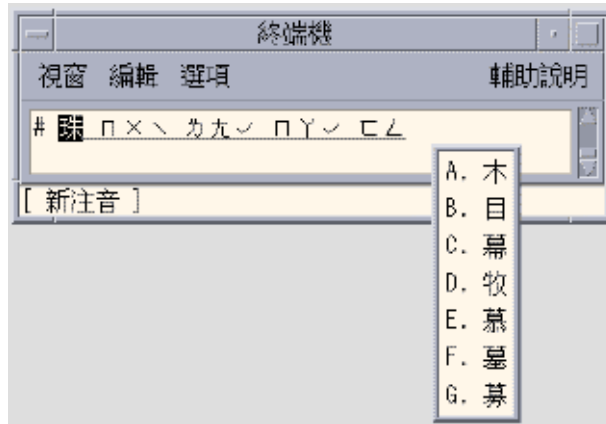
1. Type the key sequence `5jaj4x;3a83z/`.
The new ChuYin input method automatically adds spaces between each syllable.



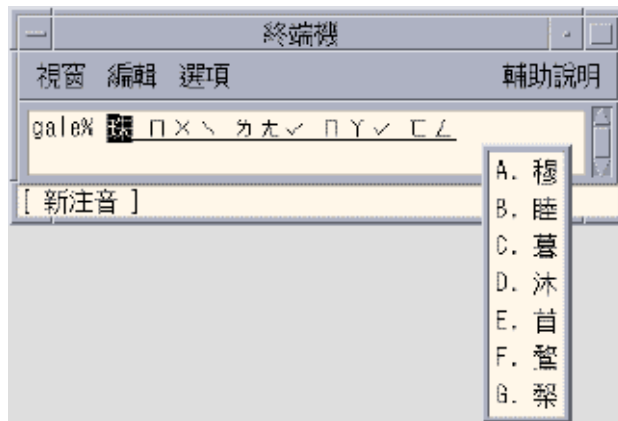
2. Since the first Chinese character is not shown in the lookup window, type [Return] to change to the single characters group.



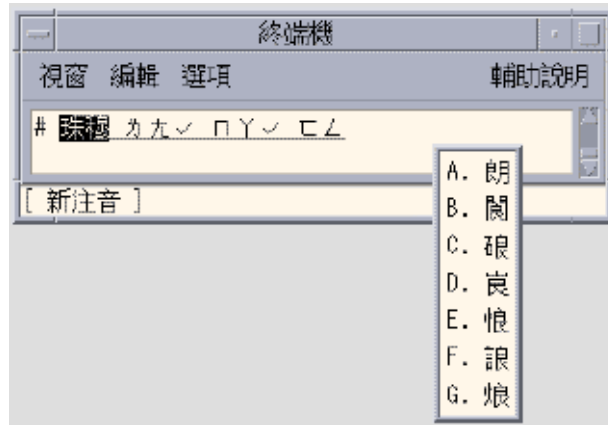
3. Type [Space] to select the first character.
The first character is inserted into the pre-edit string:



4. Since the second character does not appear in the lookup window, press [] to turn to next page which contains the second Chinese character.

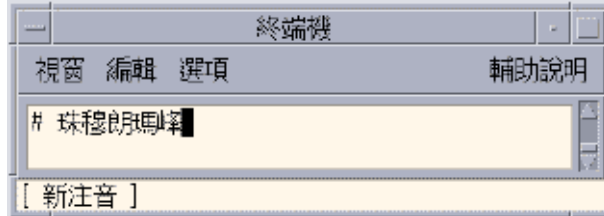


5. Type [Space] to select the second Chinese character.

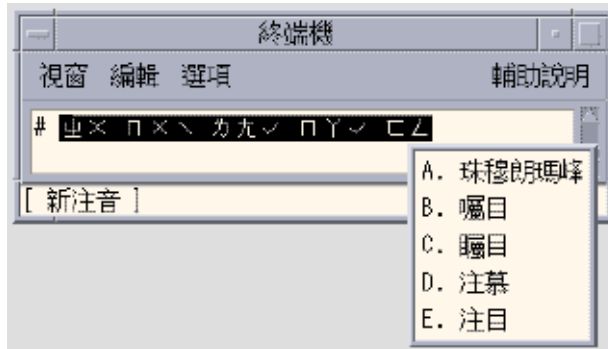


6. Use the same procedure to input the third, fourth and fifth Chinese characters.

The phrase will be committed automatically after the last Chinese character is selected.



The new phrase will be saved. When the string 5jaj4x;3a83z/ is typed again, the new phrase will be shown in the lookup window as shown below.

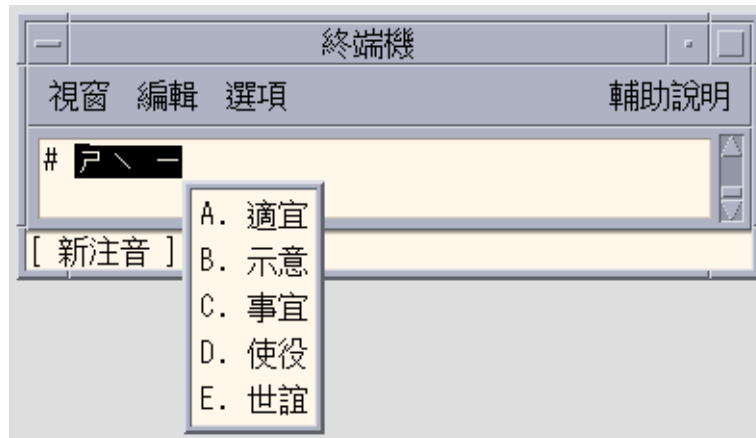


▼ Automatic Presentation of Frequently used Chinese Characters and Phrases

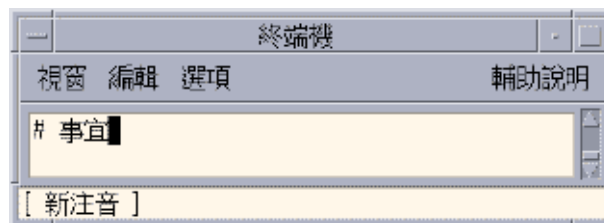
The New ChuYin input method, automatically remembers the most recently selected list items and adds them to the beginning of the list of selectable phrases and characters for easy selection.

1. Type "g4u".

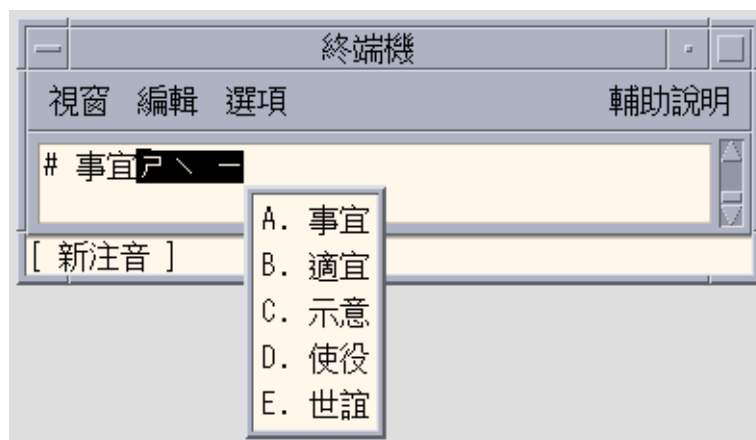
Note the sequence of the choices that appear:



2. Select the third item (C)



3. Type g4u again.



The character selection has been presented as the first in the list because it has been most recently selected previously. Characters or phrases which are used frequently are easier to select.

▼ Displaying Input Phrases by Typing a Consonant

By typing only a consonant, the related Chinese character selections will also be shown.

For example, to input phrase:

事半功倍

- **Type only the consonant of the Chinese character-- g1e1 :**
The complete phrase appears in the lookup window.

▼ Editing the Pre-edit String

- **Press [Shift] + [A-G] Keys to select the pre-edit string.**
The following table shows the definitions of the edit keys.

TABLE 4-1 Edit Key Definitions

Key	Definition
[a-z, 0-9, ,/;'-]	ChuYin Keystroke
Home	Moves to the start of the preedit line
End	Moves to the end of the preedit line
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.

▼ Using Page Scroll Keys

The selections related to a ChuYin pre edit string can be split to two groups, one the phrases group and the other is the single Chinese character Group. The phrase candidates will be shown at first. If there are many candidates that can't be shown in one page, the page scrolling key may be used.

To select a page scroll key.

- Press [Shift] + [A-G] Keys to select the definition from the table below:

TABLE 4-2 Page Scroll Key Definitions

Key	Definition
[Scrolls to previous page
]	Scrolls to next page
Return	More to another group

▼ Switching between PinYin and ChuYin Mode

You can toggle back and forth between PinYin mode and ChuYin mode with a keystroke.

- Press [Alt] + [1] to switch between ChuYin mode and PinYin mode.

In PinYin mode, the PinYin phonetics are same as the keystroke.

ChuYin Input Mode

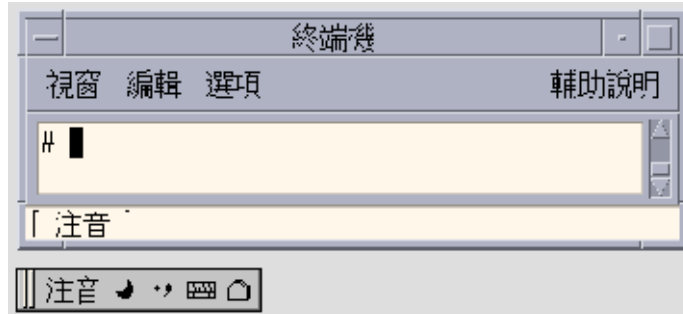
ChuYin is another common input mode. This mode applies the Chinese phonetic symbols to the phonetics of Mandarin Chinese. When a typed sequence of phonetics uniquely matches a Chinese character, the phonetics are automatically converted and the character is displayed. However, if more than one character could represent the phonetics, a lookup selection area appears showing you the available choices. Several pages of lookup choices may be available for some phonetic sequences.

Typing ChuYin Text

The following example uses ChuYin mode to enter Chinese characters representing the statement Taipei's weather is very hot.

1. **Type Control-spacebar to turn on Chinese input mode.**
The mode appears in the status area.
2. **Click the Input method selection button on the auxiliary window and select ChuYin input method.**

The status area shows that ChuYin input mode is on.



3. **Type w96.**

The ChuYin input converter finds several matching characters and displays a lookup choice window:



4. **Press [Shift] + [A] to select the appropriate character.**

The radicals are converted and the correct character appears.



I-Tien Input Mode

I-Tien mode maps Chinese phonetics to the keyboard according to the sound of each Roman keyboard character.

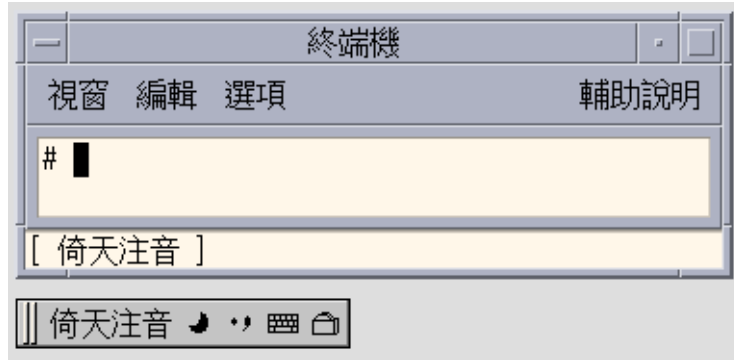
Type the command `itkdb` to display the I-Tien keyboard mapping is shown below:



Typing I-Tien Input

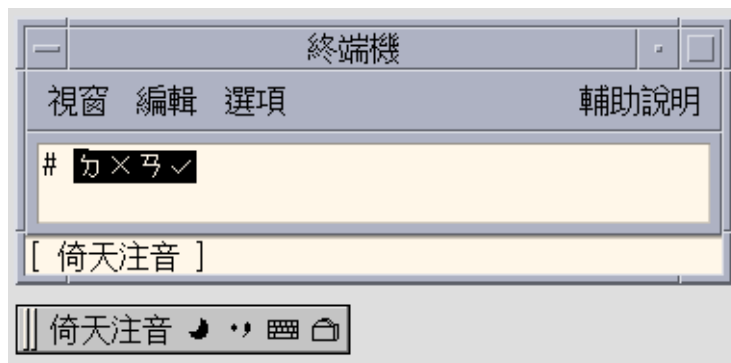
1. Open a new Terminal and turn on Chinese input conversion by pressing Control-spacebar.
2. Press F7 to turn on I-Tien input mode, or click the Input method selection button on the auxiliary window and select I-Tien input method.

The mode appears in the status area.

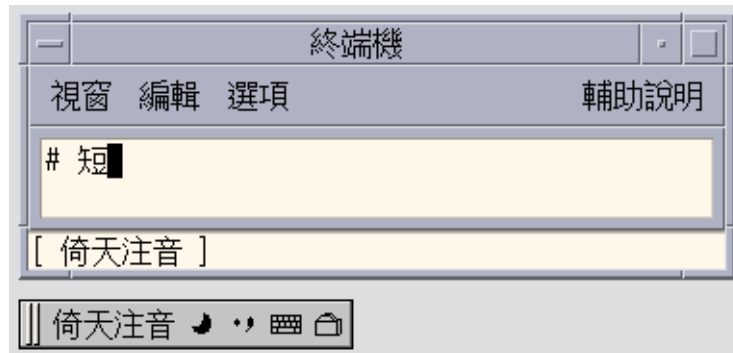


3. Type the set of radicalsdx83.

The mode appears in the status area.



4. Press the spacebar and the I-Tien input converter finds only one matching character and displays it as below.



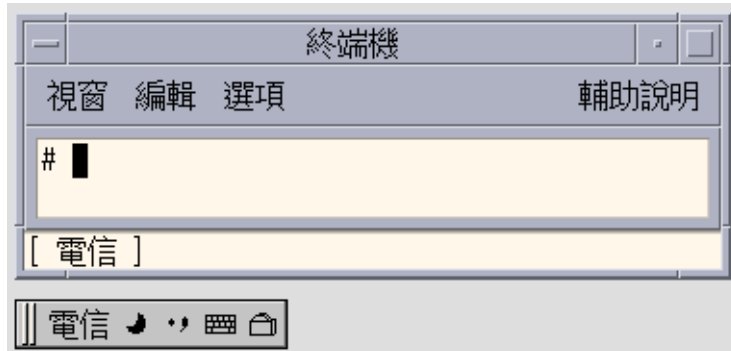
Telecode Input Mode

Telecode conversion uses the old four-digit codes used by the Taiwan Telephone & Telegraph Bureau, these codes uniquely identify 16356 Chinese characters, they are organized into two planes: plane 1 numbered from 0000 to 8045, and plane 2 numbered from .0000 to .8489.

Typing Telecode Input

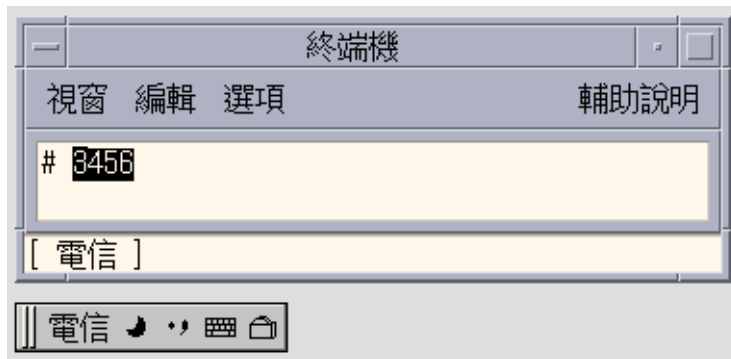
This section contains instructions on how to use the I-Tien input mode to input Chinese characters

1. Open a new Terminal turn on Chinese-input conversion on by pressing Control-spacebar.
2. Press the F12 key to turn on Telecode mode.

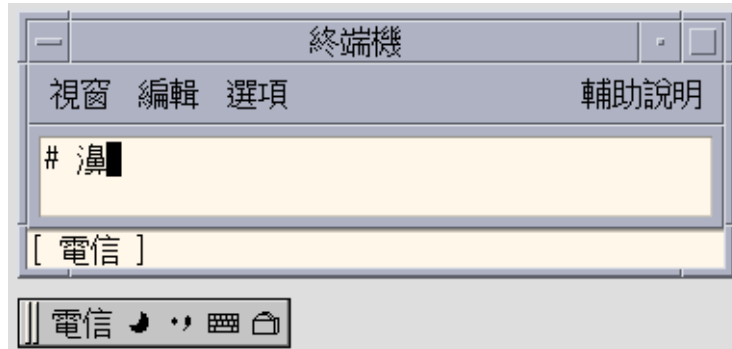


The mode appears in the status area.

3. Type the first three digits of code 3456, as below.



4. Press the spacebar and the Telecode input converter finds only one matching character and displays it as below:



TsangChieh Input Mode

Each tool's first input conversion mode is TsangChieh, a commonly used entry mode. TsangChieh is the default mode for typing Chinese (and other non-ASCII) characters when ASCII mode is toggled off. TsangChieh conversion works on the radicals that compose Chinese characters. Each Chinese character has a keystroke sequence made of the radicals that compose it. A TsangChieh keystroke sequence often uniquely defines a Chinese character, but sometimes a selection from multiple choices is required.

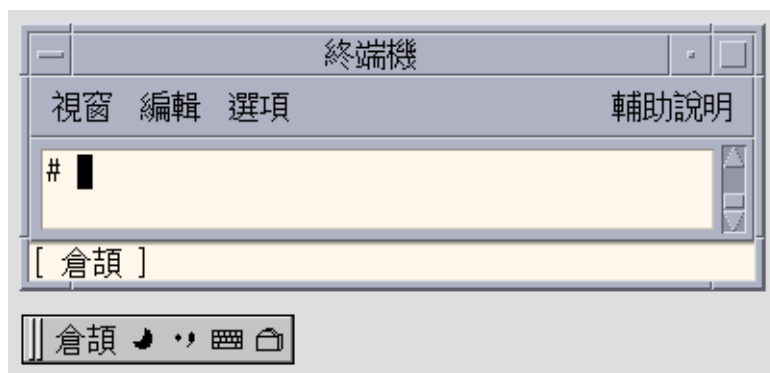
TsangChieh conversion mode can also be used to type the characters in the Chinese National Standard 11643. To type characters in this mode, press Z followed by the three-letter code assigned to the symbol by the standard.

Typing Characters in TsangChieh Mode

To type Chinese characters in this mode, type the Chinese radicals that make up the Chinese characters. If you do not have a Taiwan Chinese keyboard that shows the radicals on the keys, refer to a picture of the Chinese keyboard to see how the Chinese radicals map to your keyboard.

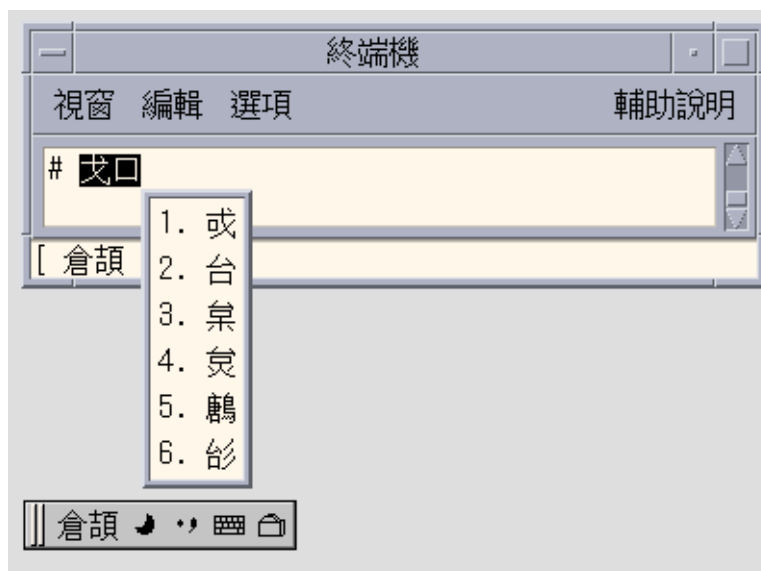
1. **Type Control-spacebar to turn on TsangChieh mode.**

The mode appears in the status area.



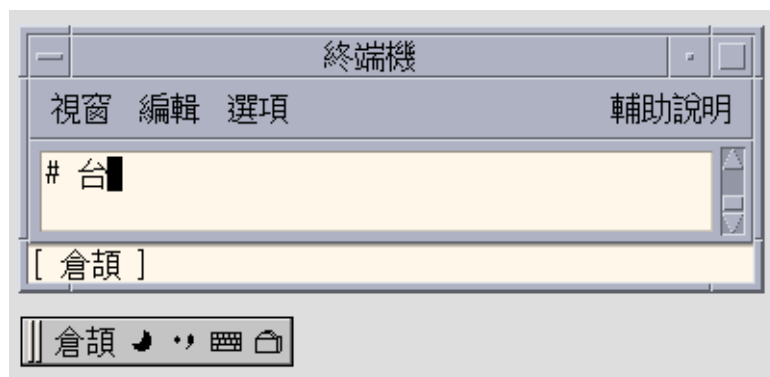
2. Type ir.

The TsangChieh input converter finds two matching characters and displays a lookup choice window:



3. Press 2 to select the appropriate character.

The radicals are converted and the correct character appears:



ChienI Input Mode

ChienI mode is a simplification of TsangChieh. Only the first and last keystrokes of the TsangChieh sequence for the target Chinese character make up the ChienI mode input. This usually displays a lookup selection of Chinese characters, as shown in the following two-character input example.

Typing ChienI Mode Input

This section contains instructions on how to use the ChienI input method to input Chinese characters.

1. **Type Control-spacebar to turn on ChienI mode.**
2. **Click the Input method selection button on the auxiliary window and select ChienI input method.**

The status area shows that ChienI input mode is on.



3. Type a and k.

The ChienI input converter finds several matching characters and displays a lookup choice window:

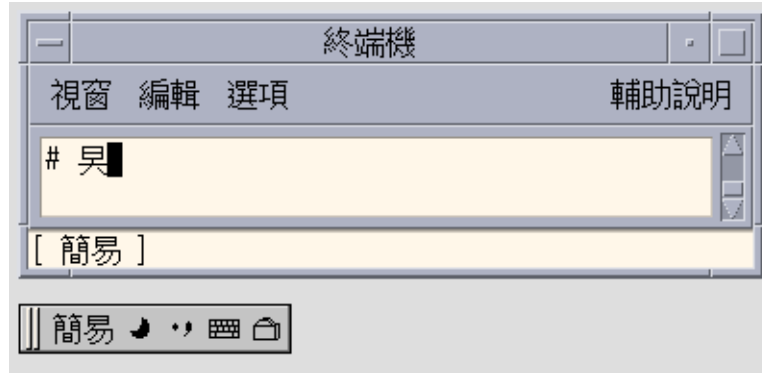


4. Press 1 to select the appropriate character.

The radical is converted and the correct character appears.

5. Type a and h.

The input converter finds several matching characters and displays a lookup choice window:



NeiMa Input Method

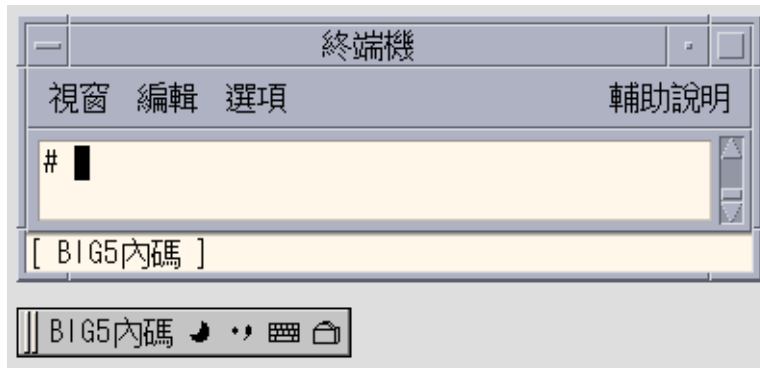
In zh_TW.EUC locale, EUC_TW NeiMa Code input method is available. in zh_TW.BIG5 locale, BIG5 NeiMa input method is available, in zh_HK.BIG5HK locale BIG5+HKSCS input method is available. NeiMa input method uses the internal code to input Chinese characters. Each Chinese character or symbol is identified by a four or eight hexadecimal digital internal code.

Typing with MeiMa Input Method

This section contains instructions on how to use the BIG5 internal codes to input 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 BIG5 NeiMa input method.

The status area shows that BIG5 NeiMa code input mode is on.



3. Press the first three of the four keys that represent the character to display (in this example, b968).

The key remains visible in the preedit area:



4. Type the fourth key.

The character is automatically committed to the window.



English_Chinese input method

English_Chinese input method requires up to fifteen keystrokes to type each Chinese word. English_Chinese maps English word to Chinese phrase. You can use the English_Chinese input method to type Chinese phrase in both zh_TW.EUC and zh_TW.BIG5 locales.

A lookup area showing the characters that match the English_Chinese 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 English_Chinese Characters

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

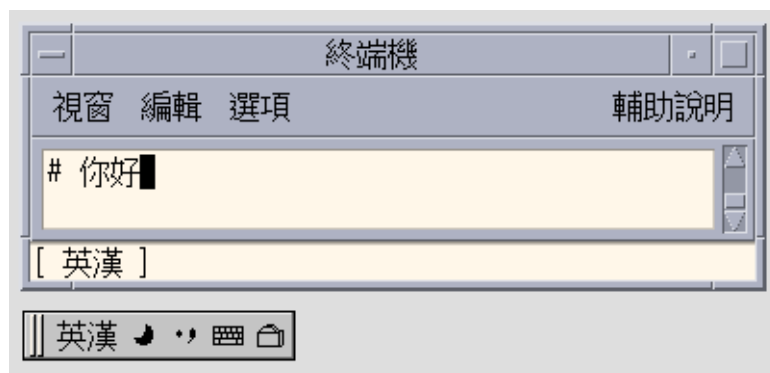
1. **Type the four keystrokes, hello**



2. Type 4 to select the corresponding Traditional Chinese phrase.



The character is automatically committed to the window.



CodeTable Input Method Interface

Introduction

Solaris 9 provides a codetable input method interface which allows Chinese users to add new input methods into their system.

▼ Creating a CodeTable

1. First create a plain text file, in which contain a list Chinese words or phrases along with the key strokes.

Codetable Format:

The codetable file contains the following sections:

- [Description]
- [Comment]
- [Key_Prompt]
- [Function_Key]
- [Phrase]
- [Single]
- [Options]

- For section "[Description]"

This section describes attributes of the codetable, such as encoding, name, valid characters, the maximum number of codes for one input items, and wild characters. The Description contains the following entries:

1. "Name:" specifies the name of this codetable.
2. Encode:" specifies the encoding of this codetable, can be UTF-8, GB, GB2312, GBK, BIG5.
3. WildChar:" specifies the wild character for input codes.
4. "UsedCodes:" specifies the valid characters to input.
5. "MaxCodes:" specifies the maximum number of input codes for one item.

- For section "[Key_Prompt]"

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

- For section "[Function_Key]"

This section specifies the key definition of some function keys, such as Page UP to scroll up the selection items, Page Down to scroll down the selection items, BackSpace to delete a input code, and ClearAll to clear all the input areas, such as preedit area, lookup area. This section contains the following entries:

1. PageUp
2. PageDown
3. BackSpace
4. ClearAll

- For section "[Options]":

This section specifies the options of the codetable input method. The Options section contains the following entries:

1. HelpInfo_Mode: Value: "ON" or "OFF"
2. KeyByKey_Mode: Value: "ON" or "OFF"
3. KeyPrompt_Mode: Value: "ON" or "OFF"
4. AutoSelect_Mode: Value: "ON" or "OFF"
5. SlectKey_Mode: Value: "Number", "Lower" or "Upper"

- For section "[Single]":

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

The format of every line as follows: `keystroke_sequence Characterlist`

- For section "[Phrase]":

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

The format of every line as follows: `keystroke_sequence word1 word2 word3 . . .`

Example of a codetable file:

```

[ Description ]
Name:          倉頡
Encode:       B:G5HK
UsedCodes:    abcdefghijklmnopqrstuvwxyz
WildChar:     ?
MaxCodes:     5

[ Key_Prompt ]
a             日
b             月
c             金
d             木
e             水
f             火
g             土
h             竹
i             戈
j             大
k             中
l             一
m             馬
n             人
o             心
p             手
q             口
r             尸
s             甘
t             山
u             女
v             田
w             雞
x             隹
y             卜
z             乙

[ Function_Key ]
PageUp:       [,< ^P
PageDown:     [,> ^N
BackSpace:    ^H ^?

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

[ Single ]
a             日
aa           胡

```

2. Convert the file to binary format that the codetable input method interface can recognize

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 as follows:

a. bin2txt and txt2bin are under the directory:

```
/usr/lib/im/locale/zh_TW/common/
```

The command syntax is:

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

```
# /usr/lib/im/locale/zh_TW/common/bin2txt binary_codetable_file
source_codetable_file
```

3. Save the file to the system.

▼ Creating a new codetable input method

1. Create and edit the 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 a user wants to have defined. The words are separated by spaces.

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 a Chinese character is typed, all the words associated with it in the dictionary will be shown for 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_TW/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 the input method specification file:

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

The `sysime.cfg` file requires the following information:

For Example: if your new codetable binary file is `newim.data`, then add a line `newim` into the input method specification file `sysime.cfg`. Once the `sysime.cfg` file is updated, the Traditional Chinese Solaris operating environment can interpret the input method.

4. Restart the input method server (htt) and relogin to the system to enable the new input method. To restart the input method server (htt), you need to run the following command as root.

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

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

The new input method will be ready to use after you relogin.

Localized Applications

This chapter describes selected properties you need to use on two localized applications of the Traditional 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 Traditional 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 (8-bit) Traditional 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 Traditional 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 Traditional Chinese characters.

Using `talk` With Traditional Chinese Characters

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

The `xtobdf` Utility

Traditional Chinese Solaris software provides this BDF font generator:

`xtobdf` — Converts font in X server to font in BDF.

The `runb5` Utility

Traditional Chinese Solaris software provides a utility that permits terminal-based Big5 applications to run under Sun's TTY window environment:

`runb5` – Runs Big5 applications under Sun's TTY window environment.

`Sdtconvtool`

`Sdtconvtool` is a graphic user interface utility that enables file conversion between various codesets. Its functionality is similar to `iconv`.

The following steps show how to convert a file encoded in UTF-8 to BIG5 encoding:

- 1. Select the code set of the file to be converted.**

Click on the arrow button to the right of the "Source Code Set" label to reveal a list of available code sets in the system. Scroll through the list and select the code set of the file to be converted. In this case, select "UTF-8."

- 2. Type the path of the file to be converted.**

The path to the file can either be entered manually in the “Source File Path” area, or chosen by selecting the “Browse...” button and selecting the file name from the file selection box. In this case, enter or select `/tmp/en_US.UTF-8`.

The “Clear” button to the right of the “Browse...” button can be used to erase the entered source or path.

3. Select the code set to which the file will be converted.

Select the target file code set information from the pulldown menu to the right of the “Target Code Set:” label, as in Step 1. In this case, select “BIG5.”

4. Type the path for the converted file.

Enter the path to the file in the “Target File Path” area. In this case, enter or select `/tmp/BIG5.all.txt`.

The “Clear” button to the right of the “Browse...” button can be used to erase the entered source or path.

5. Press the “Start Conversion” button.

The “Start Conversion” button will begin the conversion process using the given information. The “Clear All Fields” button will erase both the source and target file path names.

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. Traditional 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 Traditional Chinese Solaris software.

TABLE 5-1

Source Code	Symbol	Target Code	Symbol
CNS 11643	zh_TW-euc	Big-5	zh_TW-big5
CNS 11643	zh_TW-euc	ISO 2022-7	zh_TW-iso2022-7
Big-5	zh_TW-big5	CNS 11643	zh_TW-euc
Big-5	zh_TW-big5	ISO 2022-7	zh_TW-iso2022-7
ISO 2022-7	zh_TW-iso2022-7	CNS 11643	zh_TW-euc
ISO 2022-7	zh_TW-iso2022-7	Big-5	zh_TW-big5
CNS 11643	zh_TW-euc	ISO 2022-CN-EXT	zh_TW-iso2022-CN-EXT
ISO 2022-CN-EXT	zh_TW-iso2022-CN-EXT	CNS 11643	zh_TW-euc
Big-5	zh_TW-big5	ISO 2022-CN	zh_TW-iso2022-CN
ISO 2022-CN	zh_TW-iso2022-CN	Big-5	zh_TW-big5
UTF-8	UTF-8	CNS 11643	zh_TW-euc
CNS 11643	CNS 11643	UTF-8	UTF-8
UTF-8	UTF-8	Big-5	zh_TW-big5
Big-5	zh_TW-big5	UTF-8	UTF-8
UTF-8	UTF-8	ISO 2022-7	zh_TW-iso2022-7
ISO 2022-7	zh_TW-iso2022-7	UTF-8	UTF-8
IBM Cp937	zh_TW-cp937	UTF-8	UTF-8
UTF-8	UTF-8	IBM Cp937	zh_TW-cp937
Big-5+	zh_TW-big5p	UTF-8	UTF-8
UTF-8	UTF-8	Big-5+	zh_TW-big5p
UTF-8	UTF-8	BIG5HK	zh_HK-big5hk
BIG5HK	zh_HK-big5hk	UTF-8	UTF-8

In the following example, an EUC mail file is converted to Big5:

```
system% iconv -f zh_TW-euc -t zh_TW-big5 mail.euc > mail.big5
```

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

Fonts

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

Display PostScript System (DPS)

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

Using Traditional Chinese PostScript Fonts and DPS Facilities

The Traditional Chinese Solaris operating environment DPS provides the fonts listed in the following table.

TABLE 6-1

Font Name	Description
Kai-Medium	Alias of Kai-Medium-EUC-H.
Kai-Medium-EUC-H	Kai-Medium font, EUC encoding, horizontal display; can be used like a Roman font.
Kai-Medium-EUC-V	Kai-Medium font, EUC encoding, vertical display; can be used like a Roman font.

TABLE 6-1 (Continued)

Font Name	Description
Kai-Medium-H	Kai-Medium font, horizontal display, for making a composite with a Roman font.
Kai-Medium-V	Kai-Medium font, vertical display; for making a composite with a Roman font.
Ming-Light	Alias of Ming-Light-EUC-H.
Ming-Light-EUC-H	Ming-Light font, EUC encoding, horizontal display; can be used like a Roman font.
Ming-Light-EUC-V	Ming-Light font, EUC encoding, vertical display; can be used like a Roman font.
Ming-Light-H	Ming-Light font, horizontal display, for making a composite with a Roman font.
Ming-Light-V	Ming-Light font, vertical display; for making a composite with a Roman font.

You can use the following Traditional Chinese fonts just as you would use Roman fonts:

- Kai-Medium
- Kai-Medium-EUC-H
- Kai-Medium-EUC-V
- Ming-Light
- Ming-Light-EUC-H
- Ming-Light-EUC-V

The following figure shows a sample of Kai-Medium and Ming-Light.



Creating Composite Roman and Traditional Chinese Fonts

You can create composite fonts using one Roman font and one of the following Traditional Chinese fonts:

- Ming-Light-H
- Ming-Light-V
- Kai-Medium-H
- Kai-Medium-V

For example, the following PostScript code defines a sample composite font, Times-Italic+Kai-Medium, which uses Times-Italic for ASCII characters and Kai-Medium horizontal font for Traditional Chinese characters:

```
/Times-Italic+Kai-Medium
13 dict begin
    /FontName 1 index def
```

```

        /FMapType 4 def
        /Encoding [ 0 1 ] def
        /WMode 0 def
        /FontType 0 def
        /FontMatrix [1.0 0.0 0.0 1.0 0.0 0.0] def
        /FDepVector [
                /Times-Italic findfont
                /Kai-Medium-H findfont
        ] def
currentdict
end
definefont pop

```

Using Traditional Chinese Fonts in DPS Programming

You can use Traditional Chinese fonts just as you use Roman fonts in DPS wrap definitions. The following sample code creates the above display:

```

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

```

You can tell `PSWDisplayText(Chinese text)` in a C program to display the designated Chinese text; for example, as shown below:



Traditional Chinese Solaris software provides TrueType support in DPS.

Converting From BDF to PCF Format

Before Solaris applications can use a modified BDF file, it must be converted to a file in PCF format. It must then be replaced in the `$OPENWINHOME/lib/locale/zh_TW/fonts` directory as follows:

- Convert the file to PCF format by typing:

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

The `-o` option enables the matrix encoding used for Asian PCF font files. For more information, see the `bdf2pcf(1)` and `mkfontdir(1)` man pages.

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, as follows:

```
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-cns11643-16
MyFont-Medium14
```

```
^D
system% mkfontdir
system% xset +fp `pwd`
```

3. You can view your font by entering:

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


Traditional Chinese Printing Facilities

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

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

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

You can use the Asian Solaris `xetops`, `xutops` or `mp` utilities to print files containing Traditional Chinese text on a PostScript printer, regardless of other printing support.

These printing facilities can be used directly from a command line or from within Traditional Chinese Solaris applications as discussed in the following sections.

Printing From a Command Line

From a command line, you can print one of two ways:

- Directly to a line printer.
- Using the `xetops`, `xutops` or `mp` utilities to convert text to bitmapped graphics.

Printing With a Line Printer

The Traditional 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 following command:

```
system% lp EUC-filename
```

- To print a Big5 format file on a printer that supports this format, use the following command:

```
system% lp -d Big5_printer Big5-filename
```

For more information on setting up the Big5 filter, see *Traditional Chinese Solaris System Administrator's Guide*.

- To print a file that contains Traditional Chinese characters in EUC format on an Epson® dot matrix printer, use the following command:

```
system% lp -d Epson_printer EUC-filename
```

For more information on setting up the filter for dot matrix printers, see *Traditional Chinese Solaris System Administrator's Guide*.

Printing With the xetops Utility

The `xetops` utility allows you to print Traditional Chinese characters using a PostScript-based printer in `zh_TW`, `zh_TW.BIG5` or `zh_HK.BIG5HK` locales. The utility converts Traditional Chinese text into a bitmapped graphics printed image.

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

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

The file may contain ASCII/English characters as well as Traditional Chinese. Refer to the `xetops(1)` man page for more detailed information.

Printing With the xutops Utility

The `xutops` utility allows you to print Traditional Chinese characters in both `zh_TW.UTF-8` and `zh_HK.UTF-8` locale using a PostScript-based printer. The utility converts Traditional Chinese text into a bitmapped graphics printed image.

Note – Starting with the next release of the Solaris environment, `xetops` and `xutops` utilities may no longer be supported.

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

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

The file may contain ASCII/English characters as well as Traditional Chinese. Refer to the `xutops(1)` man page for more detailed information.

Printing with the mp Utility

As a print filter, `mp` generates a pretitled version of contents in PostScript format. The PostScript output file contains glyph images from Solaris-resident scalable or bitmap fonts, depending on each locale's system font configuration for `mp`. Now it is enhanced in the Solaris 9 environment to print files with different encoding text in corresponding asian locales.

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

```
system% mp -L $LANGfilename | lp -d printer
```

The file may contain ASCII/English characters as well as Traditional Chinese. Refer to the `mp(1)` man page for more detailed information.

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.

Binary Compatibility Package

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.

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
```

The following example shows the command for running the compiled binary code of an earlier version of the `textedit` application on a current Traditional Chinese Solaris release system:

```
system% textedit -lc_displaylang tchinese -lc_basiclocale tchinese \  
 -lc_inputlang tchinese
```

Due to incompatibilities between Traditional Chinese Solaris 2.x and 1.x applications, you cannot cut and paste Chinese characters between them.

Traditional Chinese Test Utilities

Every utility listed in this section is supported, but for this version of Solaris, you are encouraged to use the XPG4 internationalization APIs as described in the *International Language Environments Guide*.

The utilities in the following table test various aspects of the Traditional Chinese (CNS 11643) 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 the following table must be a character in WC, wchar_t. For more information, see the *hctype(3x)* man page.

TABLE A-1

Utility	Description
ishalpha	Returns true if it is a Roman character in the CNS 11643 character set.
ishupper	Returns true if it is an uppercase Roman character as defined by the CNS 11643 character set.
ishlower	Returns true if it is a lowercase Roman character in the CNS 11643 character set.
ishdigit	Returns true if it is a number in the CNS 11643 character set.
ishspace	Returns true if it is the space character in the CNS 11643 character set.
ishpunct	Returns true if it is a punctuation character in the CNS 11643 character set.
ishparen	Returns true if it is a left or right parenthesis in the CNS 11643 character set.
ishphontone	Returns true if it is a Mandarin phonetic tone.
ishradical	Returns true if it is a Chinese character radical.
ishline	Returns true if it is a ruled line symbol in the CNS 11643 character set.
ishunit	Returns true if it is a unit character in the CNS 11643 character set.
ishsci	Returns true if it is a scientific symbol in the CNS 11643 character set.
ishgen	Returns true if it is a general symbol in the CNS 11643 character set.
ishgreek	Returns true if it is a Greek character in the CNS 11643 character set.

Asian-Specific Utilities

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

Routines that have Korean or Chinese language-specific dependency are in their own language-specific library, which is linked with the corresponding C compiler option. In Traditional Chinese Solaris, `libhle` is linked with `-lhle`. Refer to the appropriate man page 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 utilities described in this section are available, but you should use `iconv` as a standard function.

Asian 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:
 - Between Combination and Completion code, both KS C 5601-1987 and KS C 5601-1992.
 - Between GB and EUC.
 - Between CNS 11643 code and Big5.
- Between code formats (such as EUC and WC).

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

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

As with classification functions described in the previous section, the use of these utilities can be controlled by the `setlocale` function. Locale-specific routines are contained in a locale-specific library. For Traditional Chinese, that library is `libhl1e`. The library can be linked during compilation using the C compiler option `-lhl1e`.

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)` for all locales and `hconv(3)` for Traditional Chinese.

The following routines are in the regular Chinese C library:

TABLE A-2

Function	Description
<code>tooupper()</code>	Converts codeset 1 Roman lowercase to uppercase.
<code>tohlower()</code>	Converts codeset 1 Roman uppercase to lowercase.

Conversion for Traditional Chinese Character Codes

The following routines perform character-based code conversion on the CNS-11643 character set. They convert characters in the set between CNS-11643, EUC, and Big5 formats. To use these routines, the library `hl1e` must be linked using the C compiler option `-lhl1e`. For more information, see the `hconv(3x)` man page.

TABLE A-3

Function	Description
<code>cbig5toeuc()</code>	Converts Big5 character to EUC.
<code>ccnstoeuc()</code>	Converts CNS character to EUC.
<code>ceuctobig5()</code>	Converts EUC character to Big5.
<code>ceuctocns()</code>	Converts EUC character to CNS.

TABLE A-4

Function	Description
<code>big5toeuc()</code>	Converts Big5 string to EUC.
<code>cnstoeuc()</code>	Converts CNS string to EUC.
<code>euctobig5()</code>	Converts EUC string to Big5.
<code>euctocns()</code>	Converts EUC string to CNS.

Mapping Traditional Chinese Keyboard Functions

This appendix shows how to configure a Sun Chinese keyboard to make selected key functions when you need them.

The Chinese Type-4 keyboard does not have a Compose key. You can use the `xmodmap` command to make any key function as Compose. In the following example, the right Meta key () is set to function as the Compose key:

```
system% xmodmap -e "remove mod1 = Meta_R"
system% xmodmap -e "keysym Meta_R = Multi_key"
```

See the `$OPENWINHOME/share/etc/keytables/Taiwan4.kt` file and the *xmodmap(1)* man page for more information.

The Chinese Type-5 keyboard does not have an AltGraph key. You can use the `xmodmap` command to make any key function as AltGraph. In the following example, the right Meta key () is set to function as the AltGraph key:

```
system% xmodmap -e "remove mod1 = Meta_R"
system% xmodmap -e "keysym Meta_R = Mode_switch"
```

See the `$OPENWINHOME/share/etc/keytables/Taiwan5.kt` file and the *xmodmap(1)* man page for more information.

Glossary

ANSI	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.
ASCII	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.
BIG5	A commonly used code set in Taiwan.
Category	In the Traditional 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.
Character Set	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. For example, the CNS 11643 character

set contains English, Greek, and Chinese character sets in addition to Chinese radicals and many other characters.

CNS

Taiwan's Chinese National Standard. This is the Taiwan analogue to ASCII. In this document set, it refers to the code set defined by CNS 11643. It contains the Chinese characters, phonetic symbols and radicals, control codes, punctuation, and western alphabets, including Roman and Greek characters. Each character is two bytes long, with the highest or most significant bit of each byte set to zero. In other words, it uses the lower seven bits of each byte. Due to the size of the Taiwan Chinese character set, they are divided into multiple codeplanes, with the default plane containing the most commonly used characters. ISO 2022 provides mechanisms for shifting from one codeplane to another.

After its revision in 1992, CNS 11643 defines 48,000 characters, which are divided among codeplanes 1-7, codeplanes 8-16 are undefined, but are included in the code set architecture. codeplanes 1 and 2 (common and rarely used characters) are unaffected by the revision. Characters that were in codeplane 14, a provisional user-defined plane, have been standardized into codeplane 3, with the overflow in codeplane 4.

Code set

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.

EUC

Extended UNIX Code. Describes four code sets modelled 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.

EUC-CNS

The EUC representation of CNS 11643. For Codeset 1, this is the normal CNS code with a one in the most significant bit of each byte. In other words, EUC-CNS equals CNS plus 0x8080. For example, the CNS character 0x212A becomes the EUC-CNS character 0xA1AA. Or in binary, 00100001 00101010 becomes 10100001 10101010. For Codesets 2 and 3, characters are also prefixed by single shift bytes SS2 and SS3. In addition, codeset 2 requires a codeplane byte. The code of a codeset-2

character is SS2 followed by codeplane byte followed by EUC-CNS. The codeplane byte is plane number added to 0xA0; for example plane 2 has codeplane byte 0xA2.

ISO	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.
Locale	A locale describes a language or cultural environment. Its setting affects the display or manipulation of language-dependent features. Traditional Chinese Solaris software provides C for U.S.A, zh_TW for Traditional Chinese extended UNIX code, and zh_TW.BIG5 for the Traditional Chinese Big5 locale.
POSIX	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.
Unicode	The international character set and encoding developed by the Unicode Consortium.
Wide Character Code (WC)	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 <code>wchar_t</code> . 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 <i>X/Open System Interface Portability Guide</i> .

Index

A

Array input, 41
ASCII
 keyboard, 12

B

BDF, xtobdf, 80
Big 5
 number of characters in, 24
Big5 applications, running, 80
BIG5 command, 90
Big5 files, printing, 90
Big5 locale, 15
binary compatibility package (BCP), 93
BoShiaMy
 input, 43

C

changing keyboard mapping, 99
ChienI input, 67
ChuYin input, 59
cns.epson.filter command, 90
converting files, 81
Convtool, 81

D

DaYi
 input mode, 45
desktop tools, 12
dot matrix printers, using with Chinese, 90
dtwm, 11

F

files, conversion, 81

G

Graphical User Interface, 12
GUI, 12

H

htt, 19
 icon, 19
 startup in CDE, 20

I

I-Tien input, 61
icon
 htt, 19
 itkbd command, 61
iconv command, 81

input
 Array, 41
 BoShia My, 43
 ChienI, 67
 DaYi, 46
 status, 25
 Telecode, 63
 TsangChieh, 65
itkbd command, 61

K

keyboard
 Chinese, 61
 entry, 12
 I-Tien mapping, 61
 mapping, 99
ko_UTF-8 locale, 15

L

locale
 overview, 15
 zh_TW.BIG5, 15
locale setting
 in CDE, 16
lookup choice area, 26

M

mailx command, 79
mapping keys, 99
Motif, 12

O

OpenWindows
 V2, 93

P

PCF format, 83

phonetic
 input, 59
phonetics
 ChuYin, 59
 I-Tien, 61
preedit area, 25
printing, xetops print filter, 90

R

runb5 command, 80

S

setting locale
 using aliases in CDE, 16
Space bar
 converting input, 44, 47
status area, 25
symbol input, 65

T

talk command, 80
Telecode input, 63
Telephone Telegraph Bureau code, 63
tools
 desktop, 12
TsangChieh
 input, 65

W

Window Manager, 11

X

X Windows and htt, 19
xetops print filter, 90
XIM, 19
xtobdf command, 80
xutops print filter, 90

Z

zh locale, 15

zh_TW.BIG5 locale, 15

zh_TW locale, 15

