



DECwindows Motif Supplemental Guide for Traditional Chinese Support

April 2007

This guide provides the Traditional Chinese specific information and describes the Traditional Chinese features supported on the DECwindows Motif system.

Revision/Update Information: This is a revised manual.

Operating System: HP OpenVMS Alpha Version 8.3

Software Version: HP DECwindows Motif/Hanyu for OpenVMS Alpha Version 1.6

First Printing, May 1994

Revised, April 2007

**Hewlett-Packard Company
Palo Alto, California**

© 2007 Hewlett-Packard Development Company, L.P.

Confidential computer software. Valid license from HP required for possession, use or copying. Consistent with FAR 12.211 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Intel and Itanium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group.

The X device is a trademark of X/Open Company Ltd. in the UK and other countries.

The HP DECwindows Motif/Hanyu for OpenVMS Version 1.6 documentation set is available on CD-ROM.

Contents

Preface	ii
Audience	ii
Organization	ii
Related Documents	ii
Conventions	iii
1. Character Sets	1
1.1. CNS 11643	1
1.2. DTSCS	4
1.3. Big-5	5
2. Codesets	7
2.1. DEC Hanyu	7
2.2. Taiwanese EUC	10
2.3. Big-5	12
3. Locales	14
3.1. Supported Locales	14
4. Fonts	15
4.1. DECwindows Fonts	15
4.2. XLFD Font Names	15
4.3. Bitmap Font Samples	17
4.4. Font Encodings	17
4.5. Specifying Fonts in DECwindows Applications	20
5. Keyboards	21
5.1. Supported Keyboard Types	21
5.2. Keyboard Layouts	21
6. Chinese Input Methods	22
6.1. Activating Chinese Input Methods	22
6.2. Switching Input Method	24
6.3. Motif Interface for Input Methods	25
6.4. Operations of Input Server	28
6.5. Customizing Input Server	29
6.6. Tsang-Chi Input Method	32
6.7. Quick Tsang-Chi Input Method	43
6.8. Phonetic Input Method	45
6.9. Internal Code Input Method	49
6.10. Phrase Input Method	51
6.11. Symbol Input	53
7. Other Chinese Features	54
7.1. Hanyu DECterm	54

Preface

This manual provides information for end users and programmers who want to use and develop internationalized applications in Traditional Chinese (Hanyu) locales on the HP DECwindows Motif system.

Audience

This manual is for both new and experienced end users and programmers who are interested in the Traditional Chinese variant of the DECwindows Motif system.

Organization

This manual is organized as follows:

Chapter 1. Character Sets

It describes the Traditional Chinese character sets supported in DECwindows Motif.

Chapter 2. Codesets

It describes the Traditional Chinese codesets.

Chapter 3. Locales

It describes the Traditional Chinese locales.

Chapter 4. Fonts

It provides information on Traditional Chinese fonts.

Chapter 5. Keyboards

It provides information on Traditional Chinese keyboards.

Chapter 6. Chinese Input Methods

It describes how to input Traditional Chinese characters.

Chapter 7. Other Chinese Features

It provides descriptions on other Traditional Chinese features.

Related Documents

The following documents supplement information in this manual:

1. *DECwindows Motif Internationalization Guide*

2. *OSF/Motif User's Guide Revision 1.1, Open Software Foundation, Prentice Hall, Englewood Cliffs, New Jersey 07632.*
3. *OSF/Motif Style Guide Revision 1.1, Open Software Foundation, Prentice Hall, Englewood Cliffs, New Jersey 07632.*
4. *X Window System, Third Edition, Robert W. Scheifler and James Gettys, Digital Press, 1992.*
5. *Programmer's Supplement for Release 5 of the X Window System, Version 11, David Flanagan, O'Reilly & Associates, Inc.*

Conventions

The following typographical conventions are used in this manual:

mouse	The term <i>mouse</i> is used to refer to any pointing device, such as a mouse, a trackball, a puck, or a stylus.
MB1 (Select) MB2 (Drag) MB3 (Menu)	MB1 indicates the left mouse button, MB2 indicates the middle mouse button, and MB3 indicates the right mouse button.
\$	A dollar sign represents the DCL prompt.
PF1 x	A sequence such as PF1 x indicates that you must first press and release the key labeled PF1, then press and release another key or a pointing device button.
[Return]	A key name enclosed in brackets indicates that you press that key on the keyboard.
[Ctrl/C]	This symbol indicates that you hold down the first named key (in this case, [Ctrl]) while pressing the key that follows the slash (in this case, [C]).

1. Character Sets

DECwindows Motif supports the following Traditional Chinese (Hanyu) character sets:

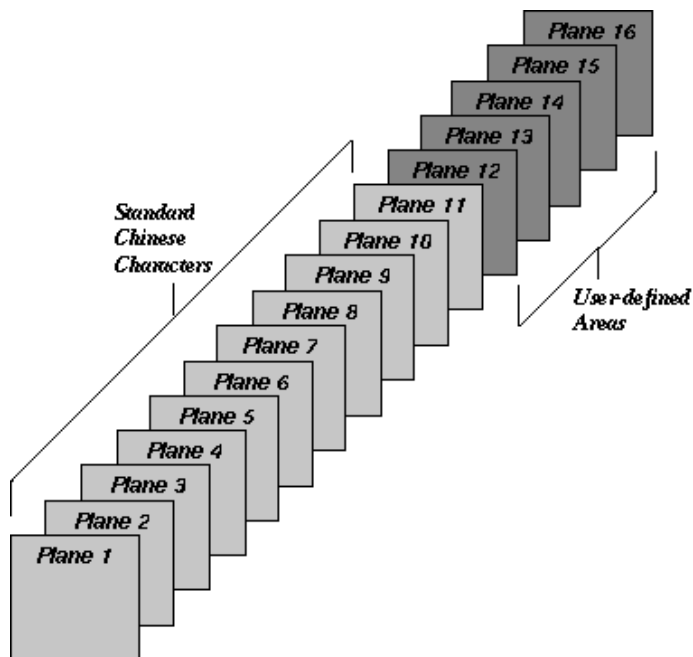
- CNS 11643
- DTSCS
- Big-5

1.1. CNS 11643

The CNS (Chinese National Standard) 11643 character set standard was published by the National Bureau of Standards of Taiwan in 1986 and was updated in 1992. It is also called "Standard Interchange Code for Generally-used Chinese Character" (SICGCC).

CNS 11643 provides 16 character planes for defining Chinese characters. Each character plane is divided into 94 rows and each row has 94 columns. Altogether, a total number of 8,836 characters can be accommodated in each plane. Character planes 1-11 are reserved for defining standard Chinese characters while character planes 12-16 are user-defined areas.

Figure 1–1: CNS 11643 Character Planes



The original CNS 11643 standard, published in 1986, defines certain groups of characters only on the first and second character planes. Table 1-1 describes these groups of characters.

Table 1-1: Characters Defined in CNS 11643-1986

Character Plane	Character Type	Number of Character
Plane 1	Special characters	651
	Control characters	33
	Frequently used characters	5,401
Plane 2	Less frequently used characters	7,650

Figure 1-2 and Figure 1-3 illustrate the positions of these characters in the first and second character planes.

Figure 1–2: CNS 11643 First Character Planes

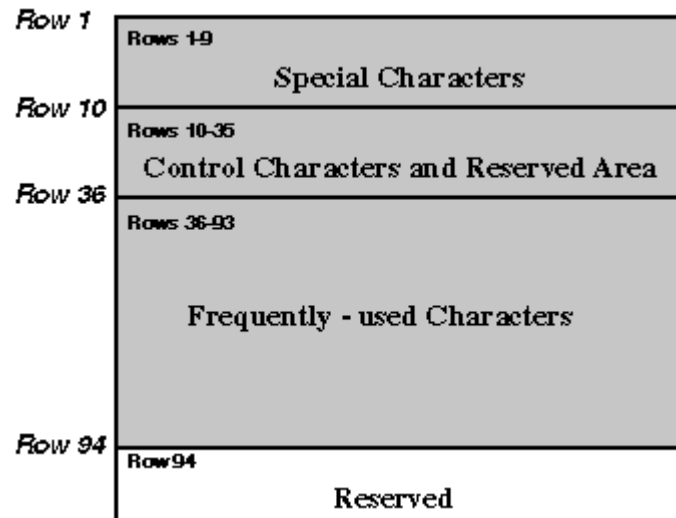
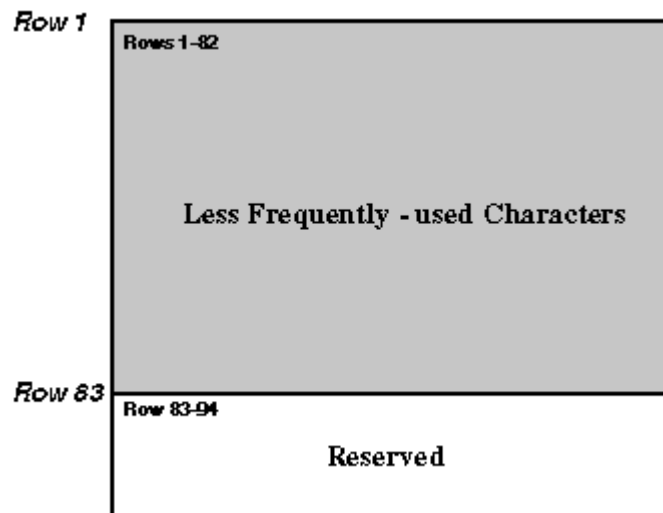


Figure 1–3: CNS 11643 Second Character Plane



Because the CNS11643-1986 character set was not rich enough to fully meet application requirements, such as names and addresses, the information industry in Taiwan requested an expansion of the character set. In 1991, the Bureau of National Standard formed a team to study how to expand CNS 11643. On August 4, 1992, the Bureau of National Standard published the revised CNS 11643 - Chinese Standard Interchange Code (CSIC).

The revised CNS 11643, called CNS 11643-1992, defined 651 special characters, 33 control characters and 48,027 Chinese characters, as shown in Error! Reference source not found.

Table 1-2: Characters Defined in CNS 11643-1992

Character Plane	Character Type	Number of Characters
Plane 1	Special characters	651
	Control characters	33
	Frequently used characters	5,401
Plane 2	Less frequently used characters	7,650
Plane 3	Rarely used characters (EDPC Part I)	6,148
Plane 4	Used for residency system, ISO 2 nd edition DIS 10646 Han characters, 171 EDPC Part II Characters	7,298
Plane 5	Rarely used characters	8,603
Plane 6	Variants based on the Ministry of Education publications (less than, or equal to, 14 strokes)	6,388
Plane 7	Variants based on the Ministry of Education publications (greater than 14 strokes)	6,539

Planes 5, 6, and 7 are based on Taiwanese Ministry of Education publications.

Because the number of characters defined in CNS 11643-1992 is far greater than those required for general use, the revised CNS 11643 is called "Chinese Standard Interchange Code (CSIC)".

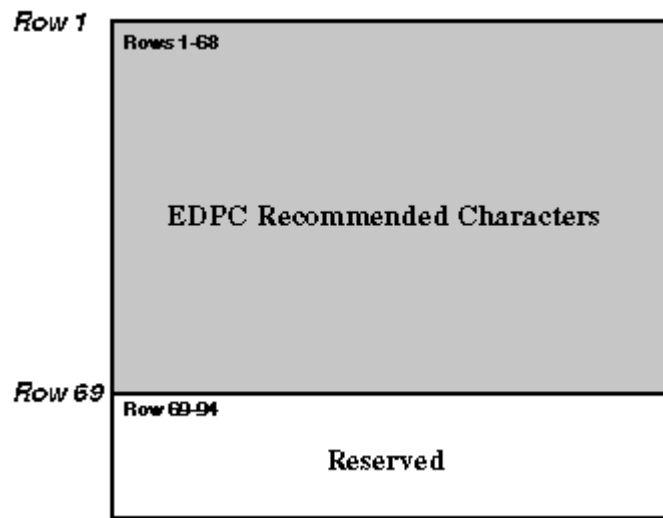
Note

In this release, the new characters added to CNS 11643-1992 are not supported. Only the characters defined in CNS 11643-1986 and DTSCS (which will be described in the next section) are supported.

1.2. DTSCS

In addition to CNS 11643, the operating system supports the DIGITAL Taiwan Supplemental Character Set (DTSCS). Currently, only the EDPC Recommended Character Set, which defines a total of 6,319 characters, is included in DTSCS. EDPC Recommended Character Set was first published by the Electronic Data Processing Center of Executive Yuen in June, 1988.

Figure 1-4: EDPC Recommended Character Set



As a de facto standard, computer vendors support the EDPC Recommended Character Set and assign it to CNS 11643 character plane 14.

In the revised CNS 11643-1992, the 6,319 characters in the EDPC Recommended Character Set are assigned to the third and fourth character planes of CNS 11643, as described in Table 1-3 and shown in Figure 1-4.

Table 1-3: Mapping of EDPC Recommended Character Set to CNS 11643-1992

EDPC Characters	Character Plane	Number of Characters
Part I	Plane 3	6,148
Part II	Plane 4	171

1.3. Big-5

The Big-5 character set, though not a national standard, is commonly used by the Taiwan information industry, particularly in the PC and workstation market. The Big-5 character set was designed to meet the requirements of five major software vendors in Taiwan. Since its publication, much software and hardware, and many peripheral devices have been developed to support Big-5.

Big-5 is very similar to the first two planes of CNS 11643-1992. The frequently used Chinese characters (5,401)

defined in the two character sets are exactly the same except that their positions in the code table are different. For the less frequently used Chinese characters, Big-5 defines two more characters in addition to the 7,650 characters defined in the second character plane of CNS 11643, and their positions in the code table are also different.

2. Codesets

DECwindows Motif supports the following Traditional Chinese codesets:

- DEC Hanyu
- Taiwanese EUC
- Big-5

2.1. DEC Hanyu

The DEC Hanyu codeset, denoted by `dechanyu`, consists of the following character sets:

- ASCII
- CNS11643, the first and second character planes
- DTSCS
- User-Defined Characters

DEC Hanyu uses a combination of single-byte, 2-byte, and 4-byte data to represent ASCII characters, symbols, or ideographic characters.

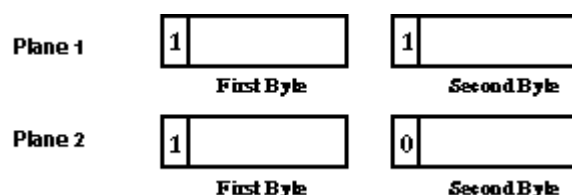
2.1.1. ASCII Code

All ASCII characters can be represented in the form of single-byte 7-bit data in DEC Hanyu. That is, the Most Significant Bit (MSB) of ASCII characters is always set off.

2.1.2. CNS 11643 Code

Each CNS 11643 character is represented by a 2-byte code in DEC Hanyu, which complies with the CNS 11643 standard. The MSB of the first byte is always set on while that of the second byte can be on for the first character plane or off for the second character plane. See Figure 2–1.

Figure 2–1: DEC Hanyu Encoding of CNS 11643 Planes



The first byte of a CNS 11643 code determines the row number of the character, while the second byte determines its column number. Table 2-1 illustrates the code range of a CNS 11643 code.

Table 2-1: CNS 11643 Code Range in DEC Hanyu

Character Plane	1st Byte (hexadecimal)	2nd Byte (hexadecimal)
Plane 1	A1 to FE	A1 to FE
Plane 2	A1 to FE	21 to 7E

The following formulas illustrate the code of a CNS 11643 character in relation to its row and column numbers.

CNS 11643 Plane 1 character:

First byte = A0 + row number
 Second byte = A0 + column number

CNS 11643 Plane 2 character:

First byte = A0 + row number
 Second byte = 20 + column number

For example, if a character is positioned at the first column of the 36th row on CNS 11643 Plane 1, its encoding value is calculated as follows:

First byte = A0 (hex) + 36 = C4 (hex)
 Second byte = A0 (hex) + 01 = A1 (hex)

Its encoded value is C4A1.

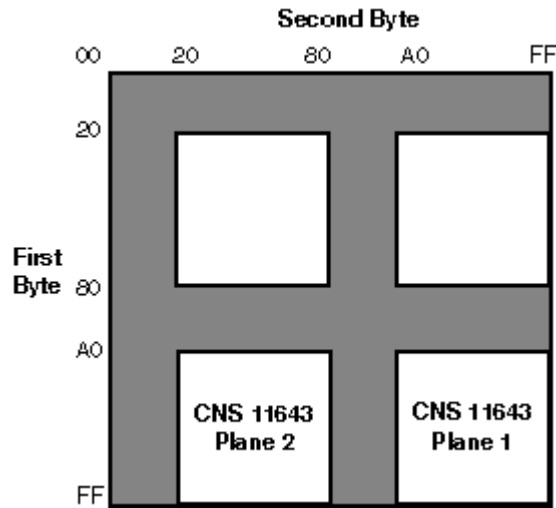
Similarly, if a character is positioned at the first column of the 36th row on CNS 11643 Plane 2, its encoding value is calculated as follows:

First byte = A0 (hex) + 36 = C4 (hex)
 Second byte = 20 (hex) + 01 = 21 (hex)

Its encoded value is C421.

Figure 2–2 illustrates the division of a 2-byte code space and the position of CNS 11643 characters.

Figure 2–2: Code Space for CNS 11643 in DEC Hanyu



2.1.3. DTSCS Code

Each DTSCS character is represented by a 4-byte code in DEC Hanyu. The first two bytes are the leading codes, namely 0xC2 and 0xCB, which are used as a designator sequence for the DTSCS character set. The MSB of the third and fourth bytes is set on for the EDPC Recommended Character Set. See Figure 2–4.

Figure 2–3: DEC Hanyu Encoding of DTSCS Characters

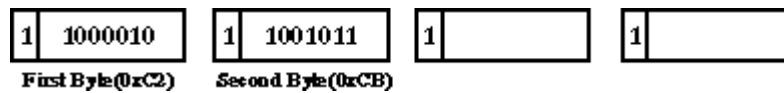
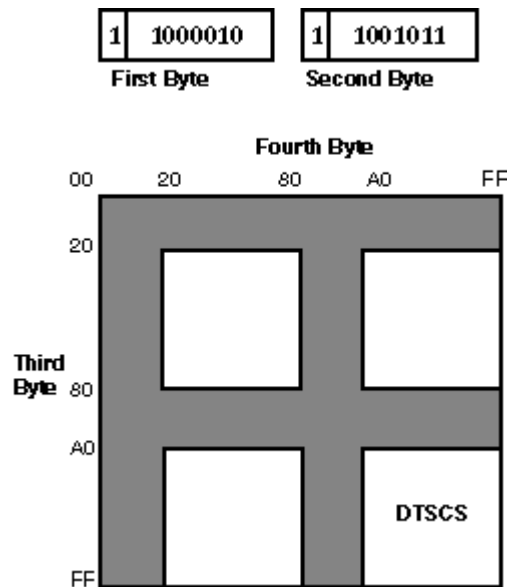


Figure 2-4 illustrates the 4-byte code space and the position of DTSCS characters.

Figure 2-4: Code Space for DTSCS in DEC Hanyu



2.1.4. User Defined Characters

In addition to the CNS11643 and the DTSCS character sets described above, DEC Hanyu provides 3,587 positions for User Defined Characters (UDC). The positions for UDCs are those unused (but not reserved) code points on the CNS 11643 first and second character planes. Therefore, the encoding of UDC is exactly the same as that of CNS 11643 except that they occupy different regions, as shown in Table 2-2.

Table 2-2: UDC Code Range in DEC Hanyu CNS

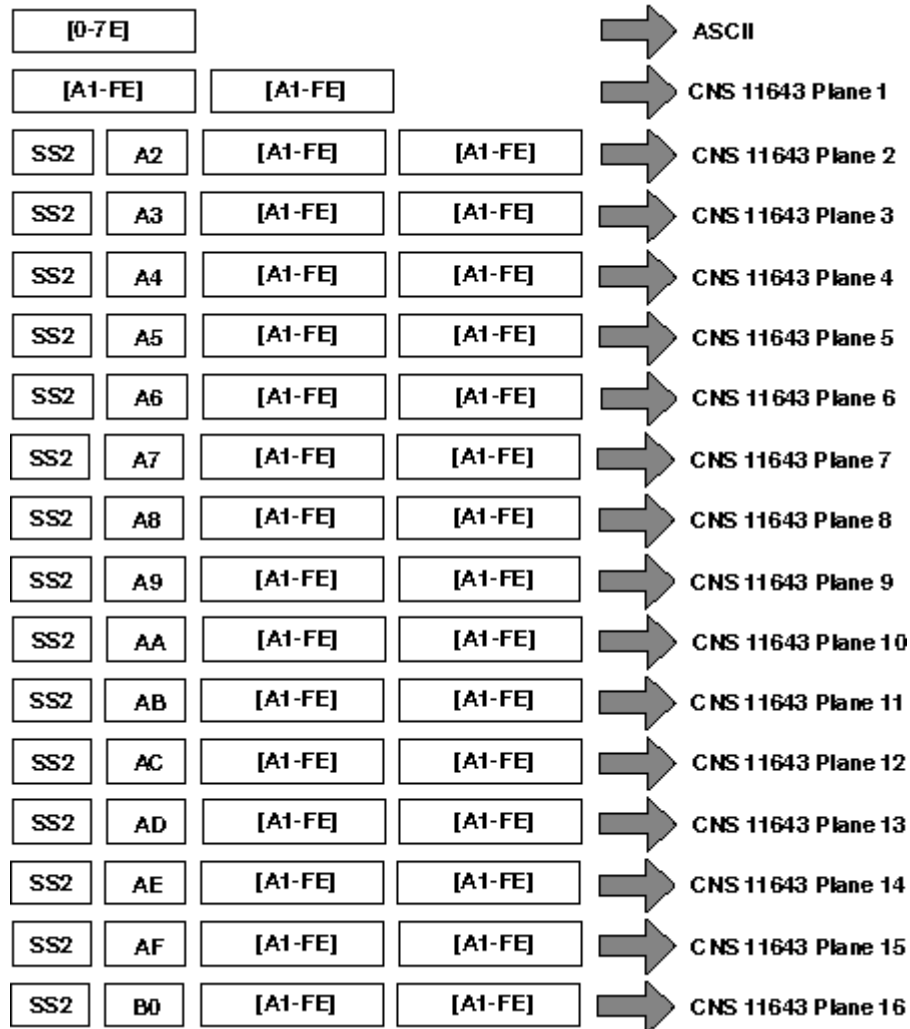
Character Plane	Number of UDC	Code Range
Plane 1	145	FDCC-FEFE
Plane 1	2,256	AAA1-C1FE
Plane 2	1,186	F245-FE7E

2.2. Taiwanese EUC

Taiwanese EUC (Extended UNIX Code), denoted as eucTW, is another codeset to support CNS 11643. The

design of Taiwanese EUC allows the 16 character planes of CNS 11643 to be encoded in a unified way. A stream of data encoded in Taiwanese EUC can contain characters defined in ASCII and the 16 character planes. Figure 2–5 illustrates the encoding of Taiwanese EUC.

Figure 2–5: Encoding of Taiwanese EUC



Taiwanese EUC uses the Single-Shift 2 control character (SS2) and an additional byte to specify a character plane. The only exception is the first plane, which does not require leading codes. Instead, two bytes specify a character's position on the first plane. The first byte determines its row number, while the second determines its column number. The MSBs of the two bytes are set on.

In this release, only the characters defined in the first and second planes of CNS 11643 and those in the EDP Recommended Character Set that have been remapped into the third and fourth character planes of the revised CNS 11643-1992 are supported in Taiwanese EUC. Other characters that were added to the CNS 11643-1992 standard are not supported.

2.3. Big-5

The Big-5 codeset, denoted as big5, is the only codeset that supports the Big-5 character set. The encoding of the Big-5 codeset is similar to that of CNS 11643 in DEC Hanyu. Each Big-5 character is represented by a 2-byte code, which complies with the Big-5 standard. The MSB of the first byte is always set on while that of the second byte can be set on or off.

The Big-5 code range is defined as shown in Table 2-3.

Table 2-3: Big-5 Code Range

Character	Number of Characters	Code Range
Special Symbols	408	A140-A3BF
Level 1 characters	5,401	A440-C67E
Level 2 characters	7,652	C940-F9D5

The operating system supports codeset conversion for HKSCS (the Hong Kong Supplementary Character Set) and uses Big-5 encoding for HKSCS representation. HKSCS characters map to Big-5 in the range of 8840 to FEFE.

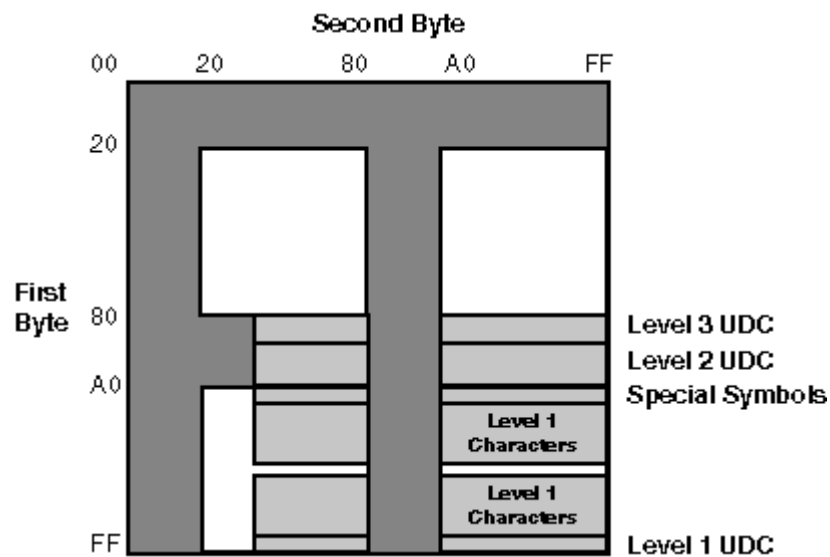
In addition to the code points for special symbols and Chinese characters shown in Table 2-4, three areas are set aside for user defined spaces. Some vendors in Taiwan support user defined characters in the code ranges shown in Table 2-4.

Table 2-4: Big-5 User Defined Spaces

Character	Number of Characters	Code Range
Level 1 user defined space	785	FA40-FEFE
Level 2 user defined space	2,983	8E40-A0FE
Level 3 user defined space	2,041	8140-8DFE

Figure 2-6 illustrates the encoding of the Big-5 codeset in a 2-byte code space.

Figure 2-6: Code Space for Big-5



3. Locales

3.1. Supported Locales

DECwindows Motif supports different Traditional Chinese locales for different countries and areas. These include Taiwan and Hong Kong. Table 3-1 shows the valid Traditional Chinese locales.

Table 3-1: Traditional Chinese Locales

Codeset	Locale
DEC Hanyu	zh_TW
	zh_TW.dechanyu
	zh_HK.dechanyu
Taiwanese EUC	zh_TW.eucTW
	zh_HK.eucTW
Big-5	Zh_TW.big5
	Zh_HK.big5

Note

zh_TW is an alias for zh_TW.dechanyu.

You can select the locale through the Language Menu of Session Manager. If you are using the Common Desktop Environment (CDE), you can select the locale using the language menu on the CDE login screen.

The applicable locales are shown in Table 3-2.

Table 3-2: Chinese Language Names

Locale	Language Name
zh_TW	Traditional Chinese Taiwan
zh_TW.dechanyu	Traditional Chinese Taiwan (DEC Hanyu)
zh_TW.eucTW	Traditional Chinese Taiwan (EUC)
zh_TW.big5	Traditional Chinese Taiwan (Big-5)
Zh_HK.dechanyu	Traditional Chinese Hong Kong (DEC Hanyu)
zh_HK.eucTW	Traditional Chinese Hong Kong (EUC)
Zh_HK.big5	Traditional Chinese Hong Kong (Big-5)

4. Fonts

4.1. DECwindows Fonts

DECwindows Motif provides Chinese DECwindows fonts in various sizes and typefaces for 75 dpi (dot-per-inch) display devices. Table 4-1 lists the screen fonts for Traditional Chinese.

Table 4-1: Traditional Chinese Screen Fonts

Typeface	Glyph Size	Bounding Box	Remarks
Screen	15 x 16	16 x 18	Mandatory font
	22 x 22	24 x 24	Mandatory font
Sung	22 x 22	24 x 24	Optional font
	30 x 30	32 x 32	Optional font
Hei	15 x 16	16 x 16	Optional font
	22 x 22	24 x 24	Optional font

There are two sets of DECwindows fonts, one for CNS 11643-1986 and one for DTSCS.

4.2. XLFD Font Names

You must specify the DECwindows font names in X Logical Font Description (XLFD) format in your application programs or in the application resource files. You can specify wildcards "*" for any fields in the font names.

You can use the following font names for both 75 dpi and 100 dpi display devices. If you want to state the display resolution explicitly, you can specify 75 or 100 in the X- and Y-resolution fields, that is, the second and third asterisks in the following XLFD names.

- Screen family font names in XLFD format:

CNS 11643-1986 Fonts

```
-ADECW-Screen-Medium-R-Normal--*-180-*-*M-160-DEC.CNS11643.1986-2  
-ADECW-Screen-Medium-R-Normal--*-240-*-*M-240-DEC.CNS11643.1986-2
```

DTSCS Fonts

-ADECW-Screen-Medium-R-Normal--*-180-*-*M-160-DEC.DTSCS.1990-2

-ADECW-Screen-Medium-R-Normal--*-240-*-*M-240-DEC.DTSCS.1990-2

- Sung family font names in XLFD format:

CNS 11643-1986 Fonts

-ADECW-Sung-Medium-R-Normal--*-240-*-*M-240-DEC.CNS11643.1986-2

-ADECW-Sung-Medium-R-Normal--*-320-*-*M-320-DEC.CNS11643.1986-2

DTSCS Fonts

-ADECW-Sung-Medium-R-Normal--*-240-*-*M-240-DEC.DTSCS.1990-2

-ADECW-Sung-Medium-R-Normal--*-320-*-*M-320-DEC.DTSCS.1990-2

- Hei family font names in XLFD format:

CNS 11643-1986 Fonts

-ADECW-Hei-Medium-R-Normal--*-160-*-*M-160-DEC.CNS11643.1986-2

-ADECW-Hei-Medium-R-Normal--*-240-*-*M-240-DEC.CNS11643.1986-2

DTSCS Fonts

-ADECW-Hei-Medium-R-Normal--*-160-*-*M-160-DEC.DTSCS.1990-2

-ADECW-Hei-Medium-R-Normal--*-240-*-*M-240-DEC.DTSCS.1990-2

Table 4-2 shows the font names, in XLFD format, of several miscellaneous Chinese screen fonts.

Table 4-2: XLFD of Miscellaneous Chinese Screen Fonts

XLFD Font Name	Character Set
-ADECW-Screen-Medium-R-Normal--*-180-*-*M-80-ISO8859-1	ISO Latin-1
-ADECW-Screen-Medium-R-Normal--*-180-*-*M-80-DEC-DECctrl	DEC Display Control
-ADECW-Screen-Medium-R-Normal--*-180-*-*M-80-DEC-DECsuppl	DEC Supplemental
-ADECW-Screen-Medium-R-Normal--*-180-*-*M-80-DEC-DECtech	DEC Technical
-ADECW-Screen-Medium-R-Normal--*-180-*-*M-80-DEC-DRCS	DEC DRCS
-ADECW-Screen-Medium-R-Normal--*-240-*-*M-	ISO Latin-1

120-ISO8859-1	
-ADECW-Screen-Medium-R-Normal--*-240--*-*-M-	DEC Display
120-DEC-DECctrl	Control
-ADECW-Screen-Medium-R-Normal--*-240--*-*-M-	DEC
120-DEC-DECsuppl	Supplemental
-ADECW-Screen-Medium-R-Normal--*-240--*-*-M-	DEC Technical
120-DEC-DECtech	
-ADECW-Screen-Medium-R-Normal--*-240--*-*-M-	DEC DRCS
120-DEC-DRCS	

4.3. Bitmap Font Samples

Figure 4-1 and Figure 4-2 illustrate samples of Chinese fonts.

Figure 4-1: Sung Font Sample

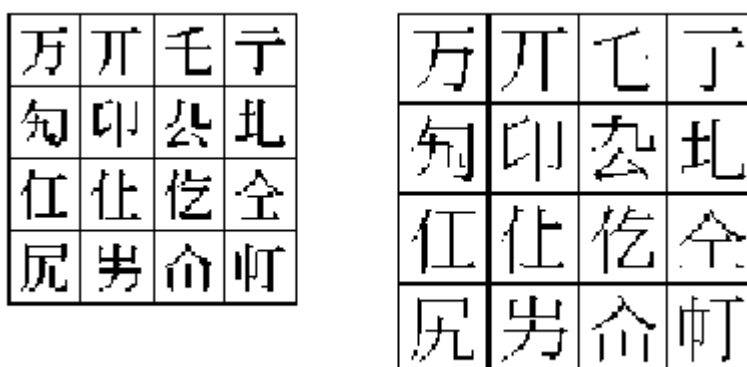
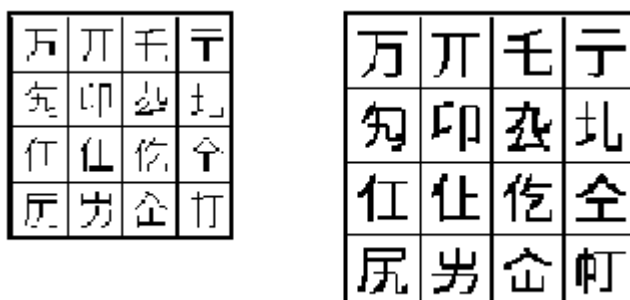


Figure 4-2: Hei Font Sample



4.4. Font Encodings

The X Consortium registers names for font encodings that are used in XLFDs. However, no names currently are

registered for CNS 11643 and DTSCS. Therefore, they are currently supported as private encodings as shown in Table 4-3.

Table 4-3: Chinese DECwindows Font Encodings

Encoding	Character Set Registry
CNS 11643-1986	DEC.CNS11643.1986-1
DSTSC	DEC.DTSCS.1990-2

Because the X Window System provides only basic Xlib functions for handling 8-bit and 16-bit characters, the 4-byte data representation of DTSCS is trimmed to remove the two leading bytes, C2 CB, to form a 2-byte encoding. DECwindows applications should either preprocess the 4-byte data and then handle them with the low level Xlib functions or handle Chinese strings with the internationalized text drawing functions provided by X11R6 Xlib or Motif Toolkit.

Figure 4-3 and Figure 4-4 illustrate these two encoding schemes.

Figure 4-3: CNS 11643-1986 Font Encoding Scheme

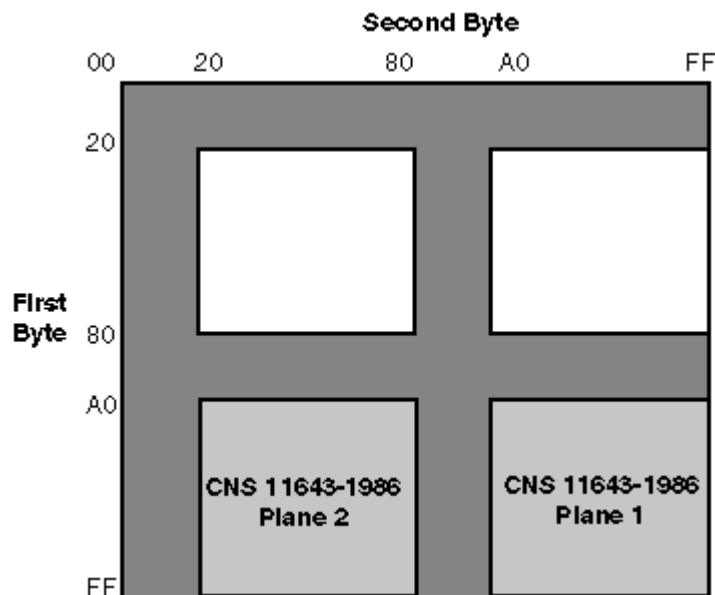
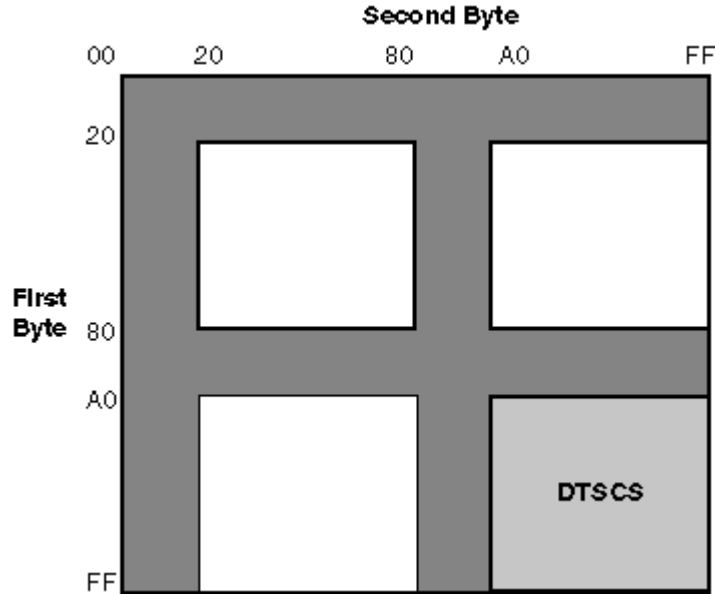


Figure 4-4: DTSCS Font Encoding Schemes



Vendors may adopt different encoding schemes or even different character sets to produce their fonts. The Chinese DECwindows fonts supplied by DECwindows Motif are all in GR encoding. To allow you to run applications on third-party workstations on which only GL-encoded fonts are installed, the DECwindows implementation of X11R6 Xlib supports the conversion of GR encoding to GL encoding for text drawing and measurement, as shown in Table 4-4.

Table 4-4: Font Encoding Conversion

Character Set	Convert From	Covert To
Taiwanese EUC	euctw-1 (plane 1)	dec.cns11643.1986-2
	euctw-1 (plane 2)	dec.cns11643.1986-2
	euctw-1 (plane 3)	dec.dtscs.1990-2
	euctw-1 (plane 4)	dec.dtscs.1990-2
Big-5	big5-0	dec.cns11643.1986-2

4.5. Specifying Fonts in DECwindows Applications

Table 4-5 shows the default fonts used in the Motif Toolkit:

Table 4-5: Traditional Chinese Default Fonts

XLFD Font Name	Character Set
-ADECW-Screen-Medium-R-Normal--*-180- *-*-M-80-iso8859-1	ISO8859-1
-ADECW-Screen-Medium-R-Normal--*-180- *-*-M-160-DEC.CNS11643.1986-2	DEC.CNS11643.1986-2
-ADECW-Screen-Medium-R-Normal--*-180- *-*-M-160-DEC.CNS11643.1986-2-UDC	DEC.CNS11643.1986-2-UDC
-ADECW-Screen-Medium-R-Normal--*-180- *-*-M-160-DEC.DTSCS.1992-2	DEC.DTSCS.1990-2
-ADECW-Screen-Medium-R-Normal--*-180- *_*_*_*_*	Fontset

To override the default fonts of a Traditional Chinese DECwindows application, you should specify the ISO Latin-1, DTSCS, and CNS11643 (UDC) fonts as well as the Chinese fontset when creating widget instances.

5. Keyboards

5.1. Supported Keyboard Types

DECwindows Motif supports the following types of keyboards for Traditional Chinese input methods:

- LK201-D
- LK401-D

5.2. Keyboard Layouts

The following figures illustrate the layouts of the Traditional Chinese keyboards.

Figure 5-1: LK201-D Keyboard Layout

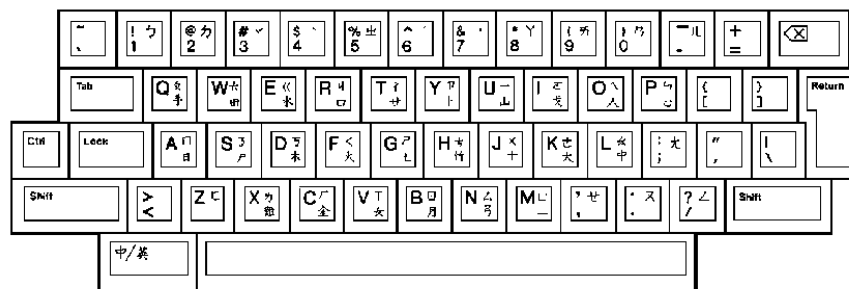
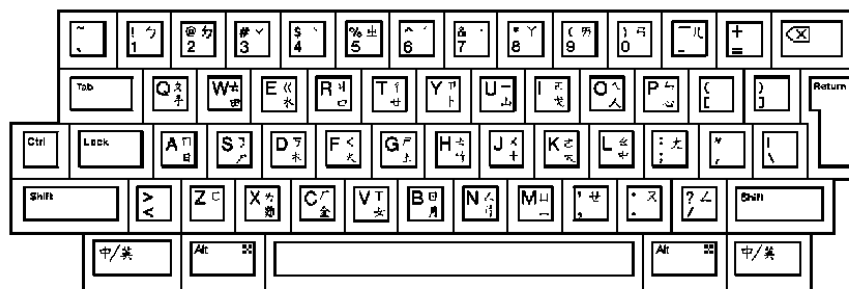


Figure 5-2: LK401-D Keyboard Layout



6. Chinese Input Methods

DECwindows Motif supports the following input methods for entering Traditional Chinese characters:

- Tsang-Chi (倉頡)
- Quick Tsang-Chi, also known as Easy (速成)
- Phonetic (注音)
- Internal Code (內碼)
- Phrase input method (片語)
- Symbol (符號)

In addition to these input methods, DECwindows Motif provides the DECW\$IM input server. It allows you to configure, manage and use the Traditional Chinese input methods.

This chapter describes:

- Each of the input methods listed above.
- How to activate, deactivate and customize the input methods and the input server.
- How to switch between input methods.
- How to work with input methods under the DECwindows Motif interface.

6.1. Activating Chinese Input Methods

You must activate a Traditional Chinese input method for response to the requirements of applications that are internationalized for the Traditional Chinese language. This section describes the steps you follow to activate a Traditional Chinese input method.

6.1.1. Input Server

Input methods are implemented by the DECW\$IM input server. The DECW\$IM input server is an X client process that runs on a standard X server. This means that the DECW\$IM input server can run on any system that can access your X display device.

The `DECW$IM` input server provides input method services to the R6 X library (Xlib) supported by DECwindows Motif. You can write internationalized applications using the standard R6 application programming interface and communicate with the input server.

The `DECW$IM` input server is a multilingual input server because it gives you the means to use and manage input methods for Simplified Chinese, Japanese and Korean, as well as Traditional Chinese. This manual focuses on the Traditional Chinese input methods.

6.1.2. Starting Input Server

The `DECW$IM` input server requires that you perform the following prerequisite steps:

- Set the current locale to Traditional Chinese.
- Stop any other input server process running on the system.
- Set the pre-editing style, or interaction style, to either Off-the-Spot or Root Window as described in [Section 6.3.2](#).

There are several ways to start the `DECW$IM` input server:

- Automatic Startup

If the Traditional Chinese language is selected on the CDE login menu, the `DECW$IM` input server starts automatically by default. When you log in, the following command procedure runs:

```
CDE$SYSTEM_DEFAULTS: [CONFIG.XSESSION_D] 0
020IMS_START.COM
```

The value of the `DTSTARTIMS` symbol determines whether the command procedure will automatically start the input server.

- Using a Command

If you want to start the `DECW$IM` input server by hand, enter the following command:

```
$ SPAWN/NOWAIT/INPUT=_NL: MC DECW$IM
```

If you want to start `DECW$IM` from a remote system, log in to the remote system and enter the following command:

```
$ SET DISPLAY/CREATE/NODE=<display-name>  
$ SPAWN/NOWAIT/INPUT=_NL: MC DECW$IM
```

In the `<display-name>` field, enter the display name for your workstation.

After you activate an input method, applications that have been internationalized to support that input method can communicate with the input server to obtain input method services. You must start these applications after the input server starts.

Note

Applications that are started before the input server cannot connect to it. However, if the application contains an `XmText` or `XmTextField` widget with the `reconnectable` resource set to `True`, the application is able to establish a connection with the input server if the application starts before the input server or when the application is running and the input server stops and restarts. For more information, see `XmText(3X)` and `XmTextField(3X)`.

6.1.3. Input Server Window

The `DECW$IM` input server window provides an options menu, including the `Class/Input Method Customization` menu item. If you select `Class/Input Method Customization`, it displays the `Customize Class` dialog box. The `Customize Class` dialog box and its sub-dialog boxes are described in [Section 6.5](#).

The title bar of the input server window is used to indicate the current state of the input server as described in [Section 6.4](#).

6.2. Switching Input Method

By default, a client application starts in the English language input mode. If the `DECW$IM` input server is

active on your system, use the [Shift/Space] key sequence to shift from English input mode to non-English input mode. When you use [Shift/Space] to shift to non-English input mode, keyboard character input is transmitted to `DECW$IM`, which converts the input using a selected input method and transmits the converted input to the client application.

The key sequences in this section can be customized by means of the Customize Input Method dialog box.

To select the Traditional Chinese input method under `DECW$IM`, use the appropriate key sequence from Table 6-1. First, press the invocation key for the input method class ([Shift/Space] for Traditional Chinese), and then an input method invocation key (any of [F6] through [F10]). Keep in mind that you do not repeat the invocation key for the input method class if you are switching input methods within that class. For example, use the [Shift/Space][F7] key sequence to invoke the Quick Tsang-Chi input method under Traditional Chinese, but to switch from Quick Tsang-Chi to Phonetic within Traditional Chinese, use only [F10].

Table 6-1: Key Sequences Used to Select `DECW$IM` Input Methods

Input Method Class	Input Method	Default Key Sequence
Traditional Chinese		[Shift/Space]
	Tsang-Chi	[F6]
	Quick Tsang-Chi	[F7]
	Phonetic	[F10]
	Phrase Input	[F9]
	Internal Code	[F8]

6.3. Motif Interface for Input Methods

You can interact with the `DECW$IM` input server through a Motif-style user interface. This interface allows an input method to provide feedback about the data being edited to compose a character, list choices for selection, provide options for customizing the input server, and so on.

6.3.1. Input Areas

The X Input Method specification defines the three input areas shown in the following table:

Table 6-2: Window Input Areas

Region	Description
Auxiliary area	An option menu that helps you to customize the input methods and the input method window.
Status area	Displays critical information about the internal state of the input methods.
Pre-edit area	Displays the intermediate text that is being composed. Also displays a list of valid candidates for the input key sequences.

6.3.2 Interaction Style

The use of the input areas depends on the interaction style (or pre-edit style) selected for the application. The Traditional Chinese input server supports two interaction styles:

- Root Window
- Off-the-Spot

You can specify the priority of the interaction styles for DECwindows Motif applications by using the `VendorShell` resource, `XmNpreeditType`. By default, the resource value is set to:

```
"offthespot , root , overthespot , onthespot "
```

This list is in priority order. The first style in the list is used in an input method. If that style is not available, the second is used, and so on.

To set a preferred interaction style, you can do one of the following:

- Use the Session Manager's Options menu in the DECwindows session. From the Session Manager's Options menu, select the "Input Method..." item, and

then click on one of the pre-edit styles in the Input Method Options dialog box.

- Use the Style Manager's Keyboard control in the CDE session. Click on the Style Manager Keyboard control, then click on the Input Methods button in the Style Manager Keyboard dialog box, and then click on one of the pre-edit styles in the Input Methods dialog box.

The `XmNpreeditType` resource is set to a priority list beginning with the pre-edit style that you have chosen. After that, the applications you invoke start up with the new setting.

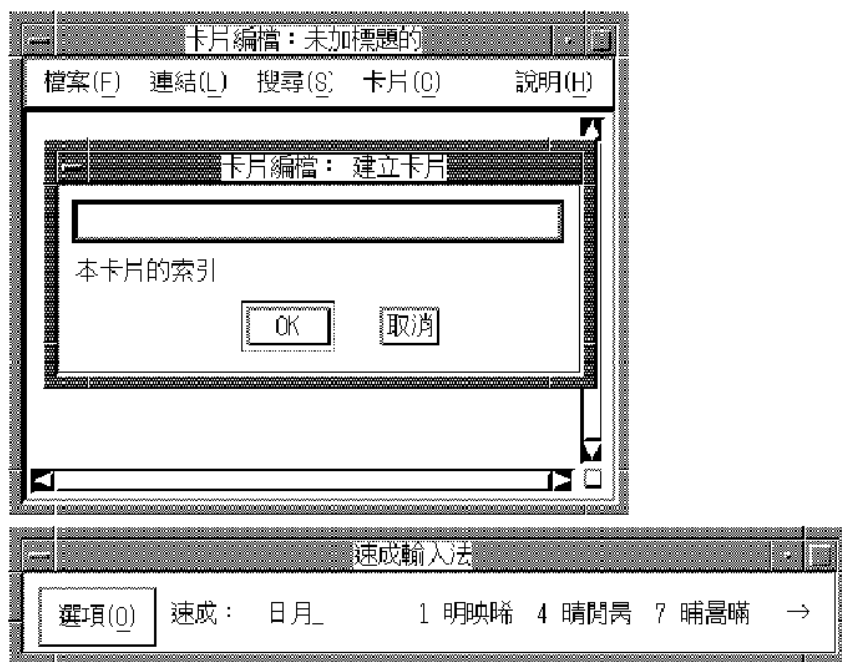
Note

Some applications, such as DECterm, may provide their own user-interface to handle interaction styles. Those mechanisms may override the methods described here.

6.3.2.1. Root Window Interaction

Choose the Root Window interaction style if you want to display the pre-edit data in a separate window from the application window.

Figure 6-1: Root Window Interaction Style



6.3.2.2. Off-The-Spot Interaction

To display the pre-edit data in a fixed location of the application window, choose the Off-the-Spot interaction style. With this interaction style, the input server creates the input area at the bottom of the application window.

Figure 6-2: Off-the-Spot Interaction Style



6.4. Operations of Input Server

When the input server is started, no application is connected to it and the input server window displays "無連接" (no connection) in its title bar. When an internationalized application is started in a Traditional Chinese locale and gets focused, the name of the current interaction style is displayed in the title bar of the input server window to indicate that the application is connected to the input server, and the string "英語:" is displayed on the status area to indicate that the input mode for the application is English. If you invoke a Traditional Chinese input method, the input state displayed on the status area is updated accordingly. If you shift the input focus to a non-internationalized application window, the title of the input server window changes to indicate there is no connection again.

The input server can maintain multiple input states for each input context or application window. Thus, when you shift the input focus to an application window, the input server restores the previous input state for that

application window. For example, if you are using an application window under the Tsang-Chi input method, then shift the input focus to another application window under the Phonetic input method, when you bring the previous window back in focus, the Tsang-Chi input method will be active.

6.5. Customizing Input Server

The input server has different levels and different methods of customization. Under the `DECW$IM` input server, the input server window provides an options menu containing the following selections:

- Class/Input Method Customization that displays a dialog box for customizing input methods
- Help that displays online help for the `DECW$IM` input server
- Save Current Settings that saves values for use by the `DECW$IM` input server
- Exit

If you select Class/Input Method Customization, it displays the Customize Class dialog box. The Customize Class dialog box has further sub-dialog boxes. This series of dialog boxes enables you to do the following:

- Select a class of input methods that is appropriate to the locale of the client application. For a Traditional Chinese locale, you select and activate the Traditional Chinese class.
- Select and activate one or more input methods within a class. The `DECW$IM` input server supports multiple input methods for Traditional Chinese.
- Establish an input method class as the default.
- Establish an input method as the default for its class.
- Customize the Tsang-Chi and Phonetic input methods.
- Customize key sequences used to invoke input methods.
- Customize error bell volume.

Changes you make to the `DECW$IM` input server settings are written to the `DECW$IM.DAT` file in your login directory.

Systemwide default settings for the DECW\$IM input server are stored in the resource file, DECW\$SYSTEM_DEFAULTS:DECW\$IM.DAT, and in the configuration file, DECW\$SYSCOMMON:[DECW\$DEFAULTS.SYSTEM.XIM]DECW\$IM_MODULES.CONF. The system administrator can modify these files to make systemwide changes.

6.5.1. Customizing Input Method Class

If you select Class/Input Method Customization from the DECW\$IM options menu, it displays the Customize Class dialog box. The Customize Class dialog box enables you to activate, deactivate, and customize the Traditional Chinese and the other input method classes. The following table briefly describes each of the Customize Class dialog box selections. For a full description of the dialog box selections, see the DECW\$IM input server online help.

Table 6-3: Customize Class Selections

Selection	Description
Active Class	Displays the input method classes that are currently active on the system. Select one or more of these classes to customize or delete.
Invocation Key	Displays the key sequence used to invoke the selected input method class. Use the checkboxes under the display to change the invocation key.
Bell Volume Slider	A bell rings when an error is made during character composition. Drag the slider to adjust the volume of the bell.
Options pull-down menu	Select "Get System Defaults" to reset the Customize Class dialog box.
Add...	Opens a dialog box that allows you to select and activate a currently inactive input method class.
Delete	Deactivates a selected input method class. The class and its methods are made inactive and not available to applications.
Customize...	Opens a dialog box that allows you to activate, deactivate, and customize

	the input method modules associated with a selected input method class. The Customize Input Method dialog box selections are similar to the Customize Class dialog box selections described in this table.
Set Default	Establishes the selected input method class as the default.
OK	Saves the customizations to become effective the next time you start the input server. Closes the dialog box.
Apply	Same as OK, except that the dialog box does not close.
Cancel	Closes the dialog box without making any changes.
Help	Displays help on this dialog box.

6.5.2. Customizing Input Method Module

If you click on the Customize button in the Customize Class dialog box, the Customize Input Method dialog box is displayed. The following table briefly describes each of the Customize Input Method dialog box selections. For a full description of the dialog box selections, see the DECW\$IM input server online help.

Table 6-4: Customize Input Method Selections

Selection	Description
Active IM	modules that are currently active on more of these input methods to
Invocation Key	Displays the key sequence used to invoke the selected input method module. Use the checkboxes under the display to change the invocation key.
Options pull-down menu	Select "Get System Defaults" to reset the Customize Input Method dialog box.
Add...	Opens a dialog box that allows you to select and activate a currently inactive input method module.
Delete	Deactivates a selected input method

	module. The method is made inactive and not available to applications.
Customize...	Opens a dialog box to change settings depending on the selected input method.
Set Default	Establishes the selected input method module as the default.
OK	Saves the customizations to become effective the next time you start the input server. Closes the dialog box.
Apply	Same as OK, except that the dialog box does not close.
Cancel	Closes the dialog box without making any changes.
Help	Displays help on this dialog box.

6.6. Tsang-Chi Input Method

To understand the Tsang-Chi input method, you must understand the concepts of Tsang-Chi root radicals, auxiliary forms, and character-splitting.

6.6.1. Tsang-Chi Root Radicals

The Tsang-Chi input method is based on the concept of root radicals. The input method requires a Chinese character to be broken down into various root radicals according to the character's shape.

Altogether 24 Tsang-chi root radicals have been defined from which almost all existing Chinese characters can be composed. The root radicals are divided into four groups and assigned to the alphabet key [A]-[Y] on the main keyboard (the [X] and [Z] keys are not assigned). The following tables illustrate the classification of the root radicals, their corresponding English keys, their auxiliary forms, and the way that they are derived.

- Philosophical classification – Table 6-5a
- Stroke classification – Table 6-5b
- Human Body classification – Table 6-5c
- Form classification – Table 6-5d

Table 6-5 is a quick reference to all of the root radical classification.

Table 6-5: Tsang-Chi Root Radicals Classification

a. Philosophical:

類別	字根 (口訣)	英文鍵	輔助字形	聯想方式	例子
哲	日	A	日、日	取日的變形。	巳、眉、胃、見
	月	B	冂、冂 心、夕	取月的外形。 取月的變形。	冂、巾、冂、冂 采、受、炎、然
	金	C	丩、儿 八、儿	取金的部分形。 取丩的變形。	鼎、Y、曾 兌、父、匹
理	木	D	十 卩 九	取木的主形。 取十的變形。 取木的主形。	寸、才、干、子 也、皮、五、夫 虎
	水	E	彡 又 水	取水做偏旁之形。 取水的左右重合之形。 取水的變形。	江、河、海、漢 支、雙、又、叔 朱、泰、隸、米
類	火	F	灬 小 小	取火做字尾之形。 取火的上半形。 取小的變形。	杰、焦、無、馬 尚、肖、半、紀 赤、不、少、紀
	土	G	土 士	取土做偏旁之形。 取土的變形。	坎、灯、坦、地 吉、寺、志、壬

b. Stroke:

類別	字根 (口訣)	英文鍵	輔助字形	聯想方式	例子
筆	竹 (斜)	H	ノ 厂	斜的定義為ノ。 取ノ的變形。	么、朱、天、八 反、派
	戈 (點)	I	・ 厶、厂	點的定義為・。 取・的變形。	犬、之、南、太 台、公、店、庫
	十 (交)	J	十	取十的變形。	宀、完、宋、安 寶、資、賓、客
畫	大 (又)	K	乂 厂 厂	又的定義為乂。 取大的部分形。 取厂的變形。	爻、刈、艾、交 友、有、左、光 痘、深、疾、痛
	中 (橫)	L	丨 中 示	縱的定義為丨，即古上而下。 中亦為取向，意義相近。 取丨的變形。	川、片、介、畫 幸、書、初、被
類	一 (橫)	M	一、厂 厂、一	皆取一的變形。	承、亞、仄、原 无、石、可
	弓 (鈞)	N	丿 一 乙、乙	鈞的定義為丿。 取丿的變形。 取一的變形。 取一的變形。	了、丁、小、可 乙、尺、丑、子 角、久、久、夕 几、屈、朵、屯

c. Human Body:

類別	字根 (口訣)	英文鍵	輔助字形	聯想方式	例子
人 身 類	人	O	亻、人 𠂇、𠂆	取人做偏旁之形。 取人做字首之形。 取人的部分形。 取人的變形。	仁、什、竹、倍 舞、臨、全、肉 八、爪、子、之 豕、兆、丰、家
	心	P	忄、小 匕、勺 匕、匕	取心做偏旁之形。 取心做字尾之形。 取心的變形。 取匕的變形。	忄、忙、怪、怡 恭、慕、恭、添 官、比、自、自 屯、切、武、民
	手	Q	扌、丰 扌、丰	取手做偏旁之形。 取手的主形。 取丰的變形。	扌、扎、扒、把 牛、毛、生、年 夫、春、年、件
	口	R			關、吧、邑、員

d. Form:

類別	字根 (口訣)	英文鍵	輔助字形	聯想方式	例子
字 形 類	尸 (側)	S	コ、匚 丁、尸 冂	側的定義為左右向開口，所以為コ、匚。 取コ、匚的變形。 取尸的部分形。	尹、邑、巨、臣 區、凶、臣、璋 力、刁、長、非、聽
	艹 (並)	T	艹 廿、卄 卄	並的定義為兩形相並，所以取艹。 取艹的變形。	華、卉、苗、曲 昔、興、前、羊 聯、盛、血、並
	山 (仰)	U	凵 凵 凵	仰之定義為向上開口，故為凵。 取凵的變形。 取凵的變形。	凶、函、扇、幻 逆、臂、民、氏
	女 (紐)	V	ㄥ ㄥ、ㄥ	紐的定義為曲紐，所以為ㄥ。 取ㄥ的變形。	幻、迥、臘、玄 亡、衣、長
	田 (方)	W	口 冂	方的定義為一方框，內含其他形，所以為口。 取口的變形。	囙、因、國、獨 四、百、西、首 母、毋、每、貢
	卜 (卜)	Y	一、卜 一、一 一	取卜的變形。 取一、一的變形。 取一、一的變形。	市、卞、上、卓 冬、斗、邇、連 冷、冽

Table 6-6: Quick Reference Table of the Tsang-Chi Root Radicals

哲理類		筆畫類		人身類		字形類		
日	日 日	竹	ノ 尸	人	人	尸	冂 匚 冫	冂 冂 冂
	A	(斜)	H	イ	〇	(側)	S	
月	冂 一	戈	ノ 厶	心	匕 匕	廿	冂 冂 冂	冂 冂 冂
	B	(點)	I		P	(逆)	T	
金	ノ 八	十	ノ	手	手 手	山	冂 冂	冂 冂
	C	(交)	J	才	Q	(仰)	U	
木	十 木	大	又 ナ	口		女	ノ 冂	冂 冂
	D	(又)	K		R	(紐)	V	
水	又 水	中	丨 中			田	口 冂	冂 冂
彳	E	(縱)	L			(方)	W	
火	灬 小	一	工 尸			ト	一 一 一	冂 冂 冂
灬	F	(橫)	M				Y	
土	士	弓	丨 一 冂					
	G	(鉤)	N					

6.6.2. Tsang-Chi Code Generation

To input a Chinese character using the Tsang-Chi input method, you must generate the character's Tsang-Chi code based on character decomposition. Most Chinese characters can be divided into two categories: the composite form (Table 6-7) and the connected form (Table 6-8). The composite form can be split into a character head and a character body while the connected form cannot be split.

Table 6-7: Composition Form Characters

Composite Form	Examples
Left-right form	針、憶、轉、謝
Top-bottom form	哲、靈
Inclusion form	國、圓

Table 6-8: Connected Form Characters

Connected Form	Examples
Connected form	亞、兩、爾

6.6.2.1. General Rules

The general rules for generating Tsang-Chi codes are:

- The character category must be composite or connected.
- The code according to the writing order is usually one of the following:
 - From outside to inside
 - From top to bottom
 - From left to right
- For composite character decomposition, the maximum number of radicals is five. The character head can be decomposed into a maximum of two radicals and the character body can be decomposed into a maximum of three radicals.

For connected character decomposition, the maximum number of radicals is four.

- If more than one Tsang-Chi code exists for a character, enter the one with fewer radicals.
- If several Tsang-Chi codes have the same number of radicals, enter the one that better represents the character.

6.6.2.2. Connected Characters

Connected characters are those which cannot be split due to the existence of crossed or connected strokes, such as 武, 面, 貞. Each character can be input by entering at most four radicals. If more than four radicals can be derived, enter the first three radicals and the last to

generate the Tsang-Chi code. Table 6-9 illustrates some examples of decomposing connected characters.

Table 6-9: Examples of Connected Character Decomposition

字例	拆字	取碼	字首	取碼	字身	取碼
亡	卜女	YV	卜	Y	女	V
也	心木	FD	心	P	木	D
回	尸尸山	SSU	尸	S	尸山	SU
勿	心竹竹	FHE	心	P	竹竹	HH
斤	竹一中	HML	竹	H	一中	ML
斥	竹一卜	HMY	竹	H	一卜	MY
商	卜金月口	YCBR	卜	Y	金月口	CBR
世	心甘	PT	心	P	甘	T
甚	廿一一女	TMMV	廿	T	一一女	MMV
弟	金弓中竹	CNLH	金	C	弓中竹	NLH
壘	一口口土	MRRG	一	N	口口土	RRG
冏	月金口	BCR	月	B	金口	CR

6.6.2.3. Connected Characters

Composite characters are those that can be split from top to bottom, left to right, and outside to inside, such as 思, 珠, 圓. You can decompose the character head into one to two radicals. If more than two radicals are generated, enter the first and the last radicals.

The character body can be decomposed into one to three radicals. If it is made up of three or fewer radicals, you should enter all the radicals. If it is made up of more than three radicals that are connected, enter the first two radicals and the last.

If the character body is itself a composite character, you can further decompose the character body into subhead and subbody. Enter the first and last radicals of the subhead, and the last radical of the subbody.

Table 6-10 illustrates examples of decomposing composite characters. In the "Shape" column, a solid square represents a character head while a square represents a radical of the character body.

Table 6-10: Examples of Composite Character Decomposition

	外觀	字例	字首	取碼	字身	取碼
邊		明計	日金	A C	月十	B J
		卦讀	土土 卜口	GG YR	卜 土田金	Y GWC
		額	人人	OO	一月金	MBC
列		勳察	女木 口月	VD LB	大尸 弓中	KS NL
		刺	卜人	YO	中弓	LN
形		川棚	中手	L Q	中中 廿山月	LL TUB
		洲	水	E	中中中	LLL
上		曷占	日卜	A Y	日口	A R
		裁登	土土 竹人	GI HO	口日	R A
		皇	女女	YV	田木	WD
下		直奎	土土	G K	口廿 土土	RT GG
		登	弓人	NO	一口廿	NR
包舍形		國區區	田尸 尸	W S S	戈口一 中一口 口口	JRB LMC RRR

6.6.2.4. Exceptional Characters

Approximately 95 percent of Chinese characters can be decomposed according to the rules described in Section

6.6.2.1. The remaining 5 percent are exceptional characters that need to be entered in different ways. The exceptional characters can be divided into the following groups:

- Compound Characters

The Tsang-Chi input method has defined nine compound characters. A compound character can be a connected character, or the character head, or body of a composite character. In any case, compound characters must be represented by their first and last radicals. Table 6-11 lists compound characters.

Table 6-11: Compound Characters

複合字	拆字	取碼	舉例	拆字	取碼	備註
門	日弓	AN	關們	日弓尸口十 人日弓	ANSRJ OAN	門當字首使用，兩碼。 門當字身使用只須兩碼
目	月山	BU	晴盲	月山手一月 卜女月山	BUQMB YVBJ	目當字首使用。 目當字身使用。
鬼	竹戈	HI	魁槐	竹戈卜十 竹日竹戈	HIYJ HARI	鬼當字首使用。 鬼當字身，只須取二碼
門	中弓	LN	關開	中弓卜中月 中弓廿金	LNYLEB LNTC	門當字首使用，取首字 形" "及尾字形"。"
冫	弓巾	NL	降邢	弓中竹水手 一廿弓中	NLHEQ MTNL	冫當字首使用。 冫當字身，也是取兩碼
佳	人土	UG	集離	人土木 心山人土	UGII POOG	佳當字首使用。 佳當字身只取首尾兩碼
扌	卜心	YP	扌慮	卜心尸一 口卜心一	YPSM RYPW	扌不分割，當字首使用 扌取"卜"及"七"。
幾	女戈	VI	幾機	女戈田 木女戈人	VIW DVIO	"幾"字形，視作與下面 字形分離，當字首。
廛	卜弓	YN	廛廛	卜弓女 卜弓廿手	YNV YNTQ	"廛"字形，整個當字首 使用。

- Difficult Characters

Difficult Characters are those which are difficult to decompose in the Tsang-Chi input method. Usually, these characters are composed of some special root radicals that are neither the Tsang-Chi root radicals nor their auxiliary forms. Using the Tsang-Chi input method you can press the [X] key (which is labeled with 難 and therefore will be referred to as the [難]

in this document) to access the special root radicals and use them to compose difficult characters. Table 6-12 lists difficult characters.

The rules of decomposing difficult characters are:

- If it is easy to identify the first and the last radicals, enter the first radical, the [難] key, and the last radical.

For example, “身” can be decomposed into “竹難竹” (HXH).

- If the first radical is easy to identify while the others are difficult, enter the first radical and then press the [難] key for the rest.

For example, “齊” can be decomposed into “卜難” (YX), and “白” can be decomposed into “竹難” (HX).

- Never use the [難] key for the first radical.

Table 6-12: Difficult Characters

舉例	拆字	取碼	舉例	拆字	取碼
身	竹難竹	HXH	鷓	弓火竹難竹	NPHXH
鳶	戈難火	IXF	薦	廿戈難火	TIXF
慶	戈難水	IXE	癩	廿手戈難水	TQIXE
鹿	戈難心	IXP	婦	女中難竹	VLXH
龜	弓難山	NXU	淵	水中難中	FXXL
兼	廿難金	TXC	風	中弓弓難山	LNNXU
颺	口難山	RXU	謙	卜口廿難金	YRTXC
白	竹難	HX	蠅	中戈口難山	LIRXU
肅	中難	LX	兒	竹難竹山	HXEU
齊	卜難	YX	蕭	廿中難	TLX
巳	弓難	NX	濟	水卜難	EYX

註：凡是字的外形與白字相近者，其拆字方式均相同，即只取白左上部的“𠂇”（竹），其餘均以〈難〉鍵代替。例如：
 興-----→ 竹難月金 HXBC 興-----→ 竹難卜金 HXYC
 興-----→ 竹難十金 HXJC 叟-----→ 竹難中水 HXLE
 藍-----→ 竹難月廿 HXBT 齒-----→ 一十竹難 MJHX

• Special Characters

Some special characters are composed by superimposing the root radicals. “木”, “大”, and “火”

on other stroke or radicals. To keep the decomposed radicals as simple as possible, enter the root radicals first, before the rest of the character is entered. Table 6-13 lists special characters.

Table 6-13: Special Characters

特殊字	拆字	取碼	取形順序	說明
東	木田	DW	木田	先取木，再取相疊的田
束	木中	DL	木中	"中"字形較口完整。
束	木田火	DWF	木口灬	先取木，再取口、灬。
束	木月	DB	木冂	取木，再把冂重疊上去
束	木十	DJ	木十	先取木，再取重疊的十
束	木人人	DOO	木人人	先取木，再取人、人。
乘	竹木中	HDL	ノ木中	乘為連體字。
乘	竹木中心	HDLP	ノ木丨匕	ノ取完，再取木之主形
典	大大大大	KKKK	大又又又	大主形先取，再取又。
拳	火人手	FOQ	火人手	火主形先取，再取人。
背	火金月	FCB	火八月	火主形先取，八再補上

6.6.3. Invoking Tsang-Chi Input Method

When you invoke the Tsang-Chi input method, the Chinese string "倉頡" is displayed in the status area, as shown in Figure 6-3.

Figure 6-3: Invocation of the Tsang-Chi Input Method

倉頡：

The radicals that you enter with the Tsang-Chi input method are displayed in the pre-edit area, as shown in Figure 6-4. To correct the data, do one of the following:

- Press the Delete key and reenter the correct radical
- Press the key (that is, F6 on a standard LK201 or LK401 keyboard) to erase all radicals in the pre-edit buffer

To signal end of input, press the Return key or Space bar.

Figure 6-4: Invocation of the Tsang-Chi Input Method

倉頡：木

6.6.4. Tsang-Chi Multiple Candidates

If there is exactly one character represented by a Tsang-Chi code, the character is sent directly to the application. Sometimes, multiple candidates for a Tsang-Chi code are available for selection when the code represents more than one Chinese character. In this case, the candidates are displayed in the pre-edit area in the following order:

1. Most frequently used
2. Less frequently used
3. Seldom used

The pre-edit area can display up to nine candidates at a time, as shown in Figure 6-5.

Figure 6-5: Multiple Candidates

1 XXX 4 XXX 7 XXX →

The number 1, 4 and 7 divide the nine characters into three groups so that you can easily select the desired candidate. To select a character that is displayed in the pre-edit area, press the corresponding numeric key on the main keyboard.

When there are more than nine candidates for selection, the indications, "→", "↔", and "←" are displayed in the pre-edit area. Table 6-14 lists the indicators and their definitions.

Table 6-14: Meaning of Arrow Characters

Indicator	Definition
→	The current row is the first row and you can press [Space] or [⇒] to move to the next row.
↔	The current row is somewhere between the first and the last row. You can press: <ul style="list-style-type: none">• [Space] or [⇒] — move to the next row• [⇐] — move to the previous row• [↑] — move to the first row
←	The current row is the last row and you can press [⇐] to move to the previous row or [↑] to the first row.

If you enter another Tsang-Chi code without selecting a candidate, the first candidate in the list is sent to the application.

If you do not want to select any candidate, but want to clear the Tsang-Chi code, press the Return key or the 倉頡 key (that is, F6).

6.6.5. Tsang-Chi Repeat Character Input

If you want to repeat input of the same character, press the equals (=) keys. Press the Delete key and reenter the correct radical

6.6.6. Tsang-Chi Error Handling

If you input incorrect data, the bell will ring. If no character is generated after you enter a Tsang-Chi code, this indicates that there is no character for the code. The radicals already entered remain in the pre-edit buffer. To handle the error, do one of the following:

1. Press the Delete key to erase the radicals, one at a time and then reenter the correct radical.
2. Press the Return key or the 倉頡 key (that is, F6 on a standard LK201 or LK401 keyboard) to erase all radicals in the pre-edit buffer and then reenter the correct radicals.
3. Enter new radicals without pressing the Return key. The radicals in the pre-edit buffer are replaced by the newly entered radicals.

6.7. Quick Tsang-Chi Input Method

The Quick Tsang-Chi input method, also known as the Quick input method or the Easy input method, is a variant of the Tsang-Chi input method and follows the same principles and rules for decomposing characters into radicals as described in Section 6.6.2.1. However, the process for entering radicals is simplified and requires only the first and the last radicals.

For example, the character “商” is decomposed into the following:

- Under the Tsang-Chi input method: “卜”, “金”, “月”, “口”
- Under the Quick Tsang-Chi input method: “卜” and “口”

This section discusses the mechanism of the Quick Tsang-Chi input method. For details about character decomposition, see Section 6.6.2.

6.7.1. Quick Tsang-Chi Code Generation

As in the Tsang-Chi input method, the character decomposition in Quick Tsang-Chi is based on whether a character is of the composite form or the connected form. However, the Quick Tsang-Chi input method requires only the first and the last radicals regardless of the number of radicals obtained.

6.7.2. Invoking Quick Tsang-Chi Input Method

When you invoke the Quick Tsang-Chi input method, the Chinese string “速成” is displayed in the status area, as shown in Figure 6-6.

Figure 6-6: The Quick Tsang-Chi Input Method

速成 :

6.7.3. Entering Quick Tsang-Chi Code

The radical that you enter is displayed in the pre-edit area, as shown in Figure 6-7. To correct the data, press the Delete key and reenter the correct radical. Alternatively, you can press the 速成 key (that is, F7 on a standard LK201 or LK401 keyboard) to erase all radicals in the pre-edit buffer. If only one radical is required to input a character, press the Return key or Space bar to signal the end of input.

Figure 6-7: Entering a Quick Tsang-Chi Code

速成 : 月

6.7.4. Quick Tsang-Chi Multiple Candidates

If exactly one character is represented by a Quick Tsang-Chi code, the character is sent directly to the application.

Frequently, multiple candidates for a Quick Tsang-Chi code are available for selection when the Quick Tsang-Chi code represents more than one Chinese character. In this case, the candidates are displayed in the pre-edit area.

To clear the Tsang-Chi code and not select any candidate, press the Return or the 速成 key (that is, F7).

6.7.5. Quick Tsang-Chi Repeat Character Input

If you want to repeat the input of the same character, press the equals (=) key.

6.7.6. Quick Tsang-Chi Error Handling

If you input incorrect data, the bell will ring. If no character is generated after you enter a Quick Tsang-Chi code, the bell indicates that there is no character for the code. The radicals you entered remain in the pre-edit buffer. To handle the error, you can do one of the following:

- Press the Delete key to erase the radicals, one at a line, and then reenter the correct radicals.
- Press the Return key or the 速成 key to erase all radicals in the pre-edit buffer, and then reenter the correct radicals.
- Enter new radicals without pressing the Return Key. The radicals in the pre-edit buffer are replaced by the newly entered radicals.

6.8. Phonetic Input Method

The Phonetic input method is based on Chinese phonetic symbols (bopomofo) that represent the pronunciation of Chinese characters.

6.8.1. Phonetic Symbol Categories

Phonetic symbols can be divided into three categories: consonants, vowels, and tone marks. There are 21 consonants, 16 vowels, and 5 tone marks. The 5 tone marks for Chinese pronunciation are the first, the second,

phonetic code can be generated according to the following rules:

- The phonetic symbols are entered in the following order:
 1. Consonant
 2. Vowel
 3. Tone marks
- A phonetic representation must have at least one consonant or one vowel.
- The end of input must be indicated with a tone mark or a Return.

Table 6–16: Examples of Phonetic Input

Code Format	Phonetic Symbols	Characters
Consonant + vowel + tone mark	尸 ㄨ ㄛ	誰
Consonant + [vowel] + tone mark	虫	知, 織
Vowel + tone mark	ㄛ	安, 庵
Consonant + semivowel + vowel + tone mark	ㄊ ㄟ ㄨ ㄛ	曉, 小
Semivowel + vowel + tone mark	ㄨ ㄛ ㄨ	我

6.8.3. Invoking Phonetic Input Method

When you invoke the Phonetic input method, the Chinese string “注音” is displayed in the status area, as shown in Figure 6-8.

The example in Figure 6-8 shows how to input the character “吉” by entering the phonetic symbols ㄐ ㄟ ㄛ (RU6) at the main keyboard.

Figure 6–8: Entering the Phonetic Symbols for “吉”

注音：ㄐ ㄟ ㄛ 1 及吉吃 4 即急級 7 集極擊 →

6.8.4. Entering Phonetic Code

The phonetic symbols that you enter are displayed in the pre-edit area. To correct the data, press the Delete key and reenter the correct symbol. Alternatively, you can press the 注音 key (that is, F10 on a standard LK201 or

LK401 keyboard) to erase all phonetic symbols in the pre-edit buffer.

You can press various termination keys to signal that you are done entering the phonetic symbols. Table 6-17 shows the results of entering phonetic symbols with different termination keys.

Table 6–17: Phonetic Symbols with Different Termination Keys

Tone	Key	Description	Example
1 st	Space	Lists characters with the same consonant or vowel in the order of the first, second, third, fourth, and light tone marks.	Type ㄅ, then press the Space bar, “ㄅ、之、支...、直、姪、值...、止、只、旨...、至、志、...” will be listed.
2 nd	6	Lists characters of the second tone.	Type ㄛ一ㄛ ㄛ, “年、黏、啉...” will be listed.
3 rd	3	Lists characters of the third tone.	Type ㄉ一ㄝˇ, “了、瞭、瞭...” will be listed.
4 th	4	Lists characters of the fourth tone.	Type ㄉ一ㄝˋ, “列、劣、冽...” will be listed.
Light	7	Lists characters of the light tone.	Type ㄩˊ, “匙...” will be listed.
None	Return	Characters corresponding to the phonetic symbols are displayed according to the order of tone marks. If only a consonant is entered before pressing Return, characters corresponding to any valid combinations of this consonant and other vowels are also displayed.	Type ㄅ, then press the Return key, characters corresponding to the order “ㄅ、ㄅㄛ、ㄅㄛˇ、ㄅㄛˋ、...、ㄅㄜ、ㄅㄜˇ、...、ㄅㄨ、ㄅㄨˇ、...、ㄅㄨˋ...” will be listed.

6.8.5. Phonetic Symbol Multiple Candidates

If a phonetic string matches exactly one character, the character is sent directly to the application. Frequently, a phonetic string matches multiple Chinese characters. In this case, the candidates are displayed in the pre-edit area.

If you do not want to select any candidate, but want to clear the phonetic code, press the Return key or the 注音 key (that is, F10).

6.8.6. Phonetic Symbol Repeat Character Input

If you want to repeat the input of the same character, press the equals (=) key.

6.8.7. Phonetic Symbol Error Handling

If you input incorrect data, the bell will ring. If no character is generated after you enter a phonetic code, this indicates that there is no character for the code. The phonetic symbols you entered remain in the pre-edit buffer. To handle the error, you can do one of the following:

- Press the Delete key to erase the phonetic symbols, one at a time, and then re-enter the correct symbols.
- Press the Return key or the 注音 key to erase all phonetic symbols in the pre-edit buffer, and then reenter the correct phonetic symbols.
- Enter new phonetic symbols without pressing the Return key. The phonetic symbols in the pre-edit buffer are replaced by the newly entered symbols.

6.9. Internal Code Input Method

Each character in DEC Hanyu has been assigned a unique internal code, just like the ID number of a company employee. For a complete list of the characters and their internal codes, see *The DEC Chinese Code Book* (Part Number EK-VT38D-CB-001).

Note

In this release, the Internal Code input method supports only the DEC Hanyu internal code. The internal codes for Taiwanese EUC and Big-5 are not supported. Even if you set the locale to one of the Taiwanese EUC or Big-5 locales, this input method still requires you to specify the DEC Hanyu internal code.

6.9.1. Internal Code Input Procedure

When you invoke the Internal Code input method, the Chinese string “內碼” is displayed in the status area.

To enter an internal code, you can optionally enter a character set number to be followed by a 4-digit hexadecimal number that specifies the position of the character with respect to the character set. The character set number can be 1 for the CNS 11643-1986 character set, or 2 for the DTSCS character set.

If you omit the character set number and enter only the 4-digit hexadecimal code, you must press the Return key or the Space bar to signal the end of input. If you enter the character set number and the 4-digit hexadecimal code, the respective character is sent automatically without pressing the Return key. Figure 6-9 shows the input of “例” using the Internal Code input method.

Figure 6—9: Input of “例” Using the Internal Code Input Method

租賃契約條_
內碼: 1CBF3

The internal code that you enter is displayed in the pre-edit area. To correct the data, press the Delete key and re-enter the correct code. Alternatively, you can press the 內碼 key (that is, F8 on a standard LK201 or LK401 keyboard) to erase all characters in the pre-edit buffer.

Since internal codes are unique for any symbols or Chinese character in DEC Hanyu, there are no multiple candidates for an internal code.

6.9.2. Internal Code Repeat Character Input

If you want to repeat the input of the same character, press the equals (=) key.

6.9.3. Internal Code Error Handling

If you input incorrect data, the bell will ring. If no character is generated after you enter an internal code,

this indicates that there is no character for the code. The internal code you entered remains in the pre-edit buffer. To handle the error, you can do one of the following:

- Press the Delete key to erase the characters representing the internal code, one at a time, and then reenter the correct internal code.
- Press the Return key or the 內碼 key to erase all characters in the pre-edit buffer, and then reenter the correct internal code.
- Enter a new internal code without pressing the Return key. The characters in the pre-edit buffer are replaced by the newly entered internal code.

6.10. Phrase Input Method

The Phrase input method allows you to define a database of frequently used phrases and store that database in a user or systemwide directory.

You associate each of the phrases in the phrase database with a unique phrase code. The phrase code can be any alphabetic label. The associated phrase can be any string of characters that is appropriate to the application's locale. You create the phrase database as a simple text file containing phrase codes and phrases and store it in a user directory or a system directory on a path that is locale dependent.

For example, if you create a phrase database file for use with applications running in the zh_TW.dechanyu locale, you copy the file to one of the following directories:

- For a systemwide database file that will be used with Traditional Chinese applications using the DEC Hanyu codeset, copy the phrase database to `DECW$SYSCOMMON:[DECW$DEFAULTS.SYSTEM.XIM.ZH_TW_DECHANYU] PHRASE_DATA.TXT`.
- For a personal database file that will be used with Traditional Chinese applications using the DEC Hanyu codeset, copy the phrase database to `PHRASE_DATA.TXT` in the `[.DECW$IM.ZH_TW_DECHANYU]` subdirectory of your login directory.

Use the following rules when populating the phrase database:

- Use a text editor to create and populate the Phrase input method database files.
- Enter the phrase code and its associated phrase on the same line. Separate the code and phrase with spaces. However, the phrase code itself cannot contain spaces. The associated phrase can contain spaces, but not in the first or last character position.
- A phrase code consists of alphanumeric characters and underscores. The associated phrase can contain any character that is legal in the target locale.

To invoke the Phrase input method, you must use the `DECW$IM` input server to activate the input method and use the default key sequence (or an invocation key sequence that you have defined). For example, if the Phrase input method is active and the client application locale supports it, press the default invocation key, `[F9]`, or a redefined invocation key to invoke the Phrase input method.

When you invoke the Phrase input method, the Chinese string “片語” is displayed in the status area, as shown in Figure 6-10.

Figure 6-10: The Phrase Input Method

片語：

When you use the Phrase input method and enter a phrase code as input to the application, `DECW$IM` searches your personal database file, if it exists, and the systemwide database file along the locale-dependent path. If `DECW$IM` finds the same phrase in both databases, it uses the definition in your personal database.

When you enter a phrase code, it appears in the pre-edit area. To correct erroneous input, use the `[Delete]` key and re-enter the code. Press the `[Space]` or `[Return]` key to signal the end of phrase code input. If you enter an incorrect code, the error bell rings. The message,

"Requested phrase does not exist", may also display. To correct the problem, make sure that the phrase is in the database file and that the file path and application locale are consistent.

6.11. Symbol Input

The Symbol input method is a simple and straightforward way to input 2-byte alphabets, numerals and other symbols. All alphabets, numerals and other symbols that you enter are immediately converted to corresponding 2-byte alphabets, numerals and other symbols.

To enter the Symbol input mode, press Z key when invoking Tsang-Chi or the Quick Tsang-Chi input method. When the Symbol input method is invoked, the string "符號" is displayed in the status area. Every character typed from the main keyboard is converted to its 2-byte form immediately. The pre-edit area is not used.

7. Other Chinese Features

7.1. Hanyu DECterm

Hanyu DECterm is a VT382-D terminal emulator. This section describes the Chinese features which are specific to Hanyu DECterm. For the details about the common internationalization features provided by DECterm, please see *DECwindows Motif Internationalization Guide*.

This section describes the Hanyu DECterm in the following areas:

- How to create a Hanyu DECterm
- Terminal emulator features
- Chinese character input/output
- Other VT382-D functionalities

7.1.1. Creating a Hanyu DECterm

The terminal type that DECterm will emulate is sensitive to the session language.

To create a Hanyu DECterm through Session Manager, you can set the session language to one of the Hanyu locales, say Chinese China, and then select DECterm from the Applications menu of Session Manager.

The user interface language of Hanyu DECterm always follows the terminal type. The language is independent of the language selection.

7.1.2. Customizing DECterm

You may apply to your Hanyu DECterm windows any of the customization features that are currently applicable to the ISO Latin-1 DECterm window except for the customization of the NRCS character sets.

Customized features can be saved and restored in the same way as in ISO Latin-1 DECterm.

7.1.3. Font Sizes

Choosing Big Font or Little Font option from Window... of the Options menu will let you choose either the 24x24 or 16x18 Chinese fonts.

7.1.4. Terminal ID

Choosing General... from the Options menu will allow you to change the general features, such as the terminal type, for the Hanyu DECterm from the dialog box displayed. Besides, you may also select the following information from the dialog box:

- VT382 ID

7.1.5. Interaction Style

Choosing Input Method... from the Options menu will allow you to select the interaction style for Hanyu DECterm. For example, if you want to select the root window interaction style, you can click on the Root Window button and then apply the change. If you click on the ISO Latin-1 Input button, Hanyu DECterm will disable the input of Chinese data until another style is selected.

7.1.6. Input Server

Choosing Input Method... from the Options menu also allows you to switch to use another input server for Hanyu DECterm. By default, the `DECW$IM` input server is used. To select another input server, you can click on the Other button and then enter the input server name in the input field.

7.1.7. Copying Information

You can choose the Edit menu to copy information within or between DECterm windows. The Cut-and-Paste operation is enhanced to handle mixed ASCII and Chinese characters.

7.1.8. Default Character Set

Hanyu DECterm supports CNS 11643 (first and second planes), DTSCS, and all character sets supported by the ISO Latin-1 DECterm.

ISO Latin-1 DECterm uses the DEC Multinational Character Set (MCS) as the default character set. This can be overridden by choosing other options in General... from the Options menu. For Hanyu DECterm, the default character set for 8-bit data is the Hanyu character set (CNS 11643 and DTSCS).

In general, Hanyu DECterm cannot display mixed MCS (or ISO Latin-1 Supplemental Graphic Characters) and Hanyu characters. If you really want to achieve this, you can output the data together with the designated character set escape sequences.

7.1.9. Chinese Character Input/Output

You can enter Chinese characters in Hanyu DECterm by invoking any of the Chinese input modes as described in Chapter 6. Mixed ASCII and Chinese characters can be displayed properly in Hanyu DECterm without any special settings.

7.1.10. Reconnecting Input Server

The Hanyu input server provides the Chinese character input capability. If this process does not exist or terminates for some reasons, a message "Hanyu input method does not exist" will display. You can restart the input server again and then use the Reset Terminal option from the Commands menu to reconnect the Hanyu DECterm to the input server.

7.1.11. VT382-D Functionalities

The following functionalities of VT382-D are implemented in the Hanyu DECterm terminal emulator:

- Display characteristics and capabilities
- Text capabilities
 - Level 3 terminal compatibility

- VT300 mode
 - VT100 mode
 - VT52 mode
- ANSI compatible control functions
- Supporting Terminal State Interrogation (TSI)
- All of the Chinese input methods
- Supporting the following character sets:
 - DEC Special Graphics Character Set (line drawing)
 - DEC Supplemental Character Set
 - DEC Technical Character Set
 - ISO Latin 1 Supplemental Character Set
 - CNS11643-1986 and DTSCS-1990 Character Sets
- Control Representation mode
- Supporting sixel graphics
- UDK editing function
- Chinese character display attributes: reverse, underline, bold, blink, double height/width
- Display/Suppress leading code

A selection button is added in the Display ... item under the Options menu for users to enable or disable the display of a symbol for the leading code in a 4-byte EDPC character.

- The escape sequence DECLCSM, that is, Leading Code Suppression Mode, is also supported in Hanyu DECterm.

For the details of the VT382-D functions, please refer to *VT382-D Programming Reference Manual* and *VT382-D User's Manual*.