U n i E d i t

MULTILINGUAL TEXT EDITOR

USER'S GUIDE

HUMANITIES COMPUTING FACILITY

DUKE UNIVERSITY

Copyright Information

SEND YOUR COMPLETED UNIEDIT REGISTRATION FORM, OR DIRECT ANY QUESTIONS, SUGGESTIONS, PRODUCT ORDERS, AND COMMENTS TO:

HUMANITIES COMPUTING FACILITY
319 NORTH BUILDING
BOX 90269
DUKE UNIVERSITY
DURHAM, NC 27708-0269 USA
PHONE: (919) 660-3190
FAX: (919) 660-3191
E-MAIL: HCF@DUKE.EDU

Table of Contents

Introduction: About the UniEdit Manual

To accompany the recent flourish of improvements in CALIS technology, we are proud to offer this new UniEdit documentation. In the following paragraphs, the general layout and methodology employed will be explained.

For further introductory information, see the following section.

To look up the definition of any unfamiliar term, see the Glossary.

Section Summaries

Introduction: What is UniEdit?

The introduction explains the basic concept behind UniEdit, from its strengths to its structure. You'll also find important information on how to contact us at the Humanities Computing Facility.

Getting Started with UniEdit

This section will move you right along from the introduction to UniEdit through the UniEdit menu arrangement to working with the screen elements and producing accents and diacritics.

Specific Language Reference Manuals

This section is dedicated to the special options and keyboard shortcuts for specific languages, especially for the "CJK" languages—Chinese, Japanese, and Korean—and the right-to-left script languages—Hebrew, Arabic, Persian (Farsi), Urdu, Yiddish, etc. Contact us at our address listed on page 3 of the printed manual and in the section of On-line Help "What is UniEdit" for more information on purchasing the CJK (Chinese-Japanese-Korean) module or the Right-to-Left Scripts module.

Configuring and Customizing UniEdit

This section gives pointers on how to customize your copy of UniEdit so that it best serves your individual needs.

UniEdit Conversions

This section deals with information necessary to convert files to and from UniEdit Unicode and other coding standards, including older versions of text CALIS and WinCALIS 1.x files.

Chapter 1: Getting Started with UniEdit

What is UniEdit?

UniEdit is a by-product of the development at Duke University of a Microsoft Windows version of CALIS, called WinCALIS. CALIS is the "Computer-Assisted Language Instruction System." It combines the most innovative concepts of Computer-Assisted Learning with the soundest pedagogical principles to equip language teachers with a stimulating educational tool. With CALIS, teachers and courseware designers can prepare language lessons which supplement classroom learning, i.e. create exercises to be administered, corrected, and scored individually for each student. With the successful integration of the popular Microsoft Windows environment, our text-based version is now available, new and improved, as WinCALIS. Successful language education now comes with the click of a mouse.

UniEdit is the "editor" part of WinCALIS Author, without the authoring tools.

The Strengths of UniEdit

Easy to Use

Microsoft Windows and a friendly user environment allow you to write quickly and easily.

Multilingual

UniEdit can communicate in just about any language. Presently supporting hundreds of languages, written in more than twenty orthographic scripts (see below for sample), the open-ended design of UniEdit's WinCALIS technology means the sky's the limit in terms of linguistic capability. We are working in conjunction with our users to develop the necessary fonts, character sets, and keyboards to support even more languages.

Amharic (Ethiopic)	Finnish	Italian
Arabic	French	Japanese
Armenian	Gaelic	Korean
Belarus	Georgian	Lappish
Bulgarian	German	Latin
Chinese	Greek	Latvian
Croatian	Hausa	Lithuanian
Czech	Hebrew	Macedonian
Danish	Hungarian	Malay
Dutch	Icelandic	Mongolian (Cyrillic)
English	Indonesian	

Norwegian	Russian	Ukrainian
Old English	Serbian	Urdu
Persian (Farsi)	Slovak	Vietnamese
Phonetic (IPA)	Spanish	Yiddish
Polish	Swedish	Yoruba
Portuguese	Tagalog	
Romanian	Turkish	

Flexible

Specific keyboard mappings are presented as defaults but may be customized.

HELP-ful

On-line help means you are never more than a mouse-click away from the answer to your question. Complete with commands and menus, the Windows help system is fast and informative.

Structure

The UniEdit workstation package consists of three main modules: the UniEdit program itself, the Multilingual Printer, and the TrueType Printer. All three components come under the rubric of "WinCALIS Pangaea."

The UniEdit component is that part of the UniEdit package in which you create and view your documents. The printers allow you to print your document files.

Copyright Information

Unicode is a registered trademark of the Unicode Consortium.

UniEdit Registration

Be sure to fill out and mail the registration form that is added to this manual so we can alert you to all new UniEdit developments and opportunities.

Send your completed UniEdit Registration Form to:

Humanities Computing Facility
319 North Building
Box 90269
Duke University
Durham, N.C. 27708-0269 USA

You must be a registered user to access Humanities Computing Facility's or its agents' Support Services. Please make sure to mail or fax your registration form today.

UniEdit Support Services

You may contact Duke University (Humanities Computing Facility) or its agents for Technical Assistance regarding the functionality of the program, its subset of commands, and the way the software works with the hardware, all as stated in the documentation.

No matter how you choose to contact us, please be prepared to provide the following information:

- Your Registration Number

- Product version number listed found by choosing About from the Help menu in UniEdit.

- The type of computer hardware you are using.

    Optionally, and according to the type of question you may have, also provide us with the following:

- Software version numbers of MS-DOS, and Microsoft Windows.

- Contents of the WINCALIS.INI file.

- Exact wording of any messages on the screen.

- What you were doing when the problem occurred.

- How you tried to solve the problem.

Electronic Mail

Ask questions and receive help from the Technical Support Department via e-mail, which is by far the most flexible way of getting support. Just send the electronic mail to the account listed below. This option is given the highest priority.

hcf@duke.edu

World Wide Web

http://www.lang.duke.edu, the Humanities Computing Facility Home Page, provides information about products, events, and technical support. Demonstration WinCALIS Exercises, the UniEdit Tryout Edition, and supplemental files may be downloaded from this site.

Fax Machine

Feel free to send us your product orders, questions or comments. Our fax number is:

919-660-3191

Telephone

If you would like to talk to us in person, or if you have questions you would like to ask us, use these telephone numbers. You can reach Technical Support Services by phone between 8:30 A.M. and 5:00 P.M., Monday through Friday. When you call, you should be at your computer and have your documentation at hand. Be prepared to provide the information as requested on the previous page.

919-660-3190

Address

Please feel free to write us at the following address:

Humanities Computing Facility
319 North Building
Box 90269
Duke University
Durham, N.C. 27708-0269 USA

You must be a registered user to access Humanities Computing Facility's or any of its agents' Support Services. Support Services are subject to prices, terms, and conditions in place at the time the service is used.

System Requirements

Before you begin the installation process, make sure that you have all the hardware and software you need to run UniEdit. At minimum, you will need the following:

- A personal computer with an Intel 80386 or higher processor, a minimum of 4M memory, standard VGA graphics or above.

- A hard disk with 5M of free disk space and one of the following: CD-ROM drive or 3 ½ (1.44M) disk drive.

- Windows 3.1 or higher running on your machine.

Installation

The files on the UniEdit program diskettes are in compressed format. The Setup Program decompresses the files as it copies them to your hard drive. UniEdit cannot be run from the

diskettes included in your package. You must first install the software using the UniEdit Setup program.

With Windows 3.1 running on your machine, select File, then Run from the Program Manager menu bar. In Windows 95, click the Start button, then choose Run. In the dialog box that appears, type **a:\setup.exe** (where a: corresponds to the 3.5" floppy disk drive of your machine). This will start the installation program.

Note: You may exit at any time during the setup procedure.

The first thing you will be asked to do is enter your name and organization or affiliation. When you press ENTER, you will proceed to the next screen which will verify that the organization name you just entered is correct. Next, setup will inform you of the file(s) that it will install automatically.

Next, setup will assign a default drive and directory where the programs will be installed, unless you specify an alternate drive and/or directory. You will also be provided with information on space required and space available on the hard drive.

Note: If you click on Install at this point, only the UniEdit core module will be installed. To install additional modules, click on the Optional Modules button.

Choose Individual Files

If you clicked on the Optional Modules button, a dialog box will appear listing various file options. You can choose to install only that part of UniEdit you will be working with by clicking on the desired files with the mouse, or select the All button. When you have made your selections, press the Continue button. If the All button is chosen, Setup will install all modules available in UniEdit. In order to install the Chinese-Japanese-Korean option, you <u>must</u> have setup disks 5, 6, 9, and 11. In order to install the Right-to-Left Scripts option, you <u>must</u> have setup disk 7.

Once the files are chosen, click on the Install button. UniEdit will verify that you have sufficient memory on the hard drive before beginning installation.

An overview of UniEdit fundamentals is available as a part of the installation package.

Starting UniEdit

When viewing the UniEdit group window, you will see three program-item icons, one for each of the components of the package. We will now examine how to start UniEdit and move around in the UniEdit environment. If you are unfamiliar with Microsoft Windows, please go through the demonstration program provided for you on Microsoft Windows or check your Microsoft Windows User's Guide.

After you have Windows up and running on your machine and have installed UniEdit, select the UniEdit program icon to start it. You will be presented with a blank window for your document, the "Document1" window. To help you understand all of your options, the on-line help system is always available. Any changes you make and save as part of your desktop will be in effect until you change them again, so you do not have to reset them every time you run UniEdit.

Chapter 2: Overview of UniEdit

UniEdit and Unicode

UniEdit editor is a relatively simple text processor, but a powerful multilingual editing system. By incorporating support for Unicode™ the editor combines all of the world's alphabets into one program, making UniEdit truly multilingual. The following will offer a quick overview of some of the basic rules and methodologies you must know in order to understand how the UniEdit editor works.

The UniEdit Editor is geared towards ease of multilingual editing and foreign language learning. It does this by incorporating support from Unicode, a standard code that can map the combination of all the world's character sets onto your standard keyboard. Using two bytes for each character representation, rather than one which the standard computer ASCII format uses, 65,536 unique characters can be represented, allowing UniEdit to be multilingual and international. The aim of UniEdit was fourfold. The first was to offer completeness in all of its languages. Unicode provides this by allowing UniEdit to support up to 65,000 individual characters. The second aim was simplicity, such that every unique character is represented with no duplication. The third was correctness to ensure that every character is a real character as defined by experts. Finally, we wanted to ensure fidelity so that no information would be lost when converting text to or from existing character sets.

See also "Unicode Character Selection Dialog Box."

The Editor

The UniEdit Editor is a simple word-processing application that you can use to create and print Unicode documents. To write documents, you need to know how to type, edit, and format text in the UniEdit windows, save the file, and print document file on your printer.

All the editing windows in UniEdit share the following items:

Control Menu

The Control-menu box is in the upper-left corner of each window. The Control menu is most useful if you use your keyboard to work with Windows. By using Control menu commands, you can restore, move, size, maximize, minimize, and switch to other applications. (If you use a mouse, you can perform these tasks by clicking and dragging.)

Title Bar

The title bar shows the names of the editor windows that are in the Restore form along with the task number you are currently working on.

Maximize/Minimize Buttons

Using a mouse, you can click the Maximize button in the upper right-hand corner to enlarge the active window to fill the entire screen, or you can click the Minimize button next to it to reduce the window to an icon.

Restore Button

The Restore button takes the place of the Maximize button once a window has been enlarged. The Restore button contains both an up arrow and a down arrow. If you click the Restore button, the window will return to its previous size.

Vertical/Horizontal Scroll Bar

These can be used to move parts of a document into view when it doesn't fit into a window.

Window Border/Window Corner

The window border is the outside of the window. By moving the mouse pointer to the window border, you will notice the mouse pointer changes from a vertical "I-beam" to one with horizontal double arrows. You may shorten or lengthen each side of a window by holding down the mouse button and dragging the border when the mouse pointer shows the horizontal double arrows.

The small box at the corner of each window, called the window corner, also may be used to adjust the size of the window. It works like the window border, but adjusts the sides of the windows at the same time.

Insertion Point

The insertion point is the location in which the text will appear when you start typing, where there is a blinking "caret."

Mouse Pointer

The mouse pointer changes with the movement of the mouse. When you select Context-Sensitive Help from the Help menu, or press Ctrl+F1, the mouse pointer changes into a Help pointer. You may click the Help pointer on any menu item or CALIS command, special response or option keyword to get on-line help about it.

For a detailed description of the main menu bar, see the section UniEdit Menus.

See the Glossary of this manual for definitions of other Windows and UniEdit terms and concepts.

UniEdit Editor Limits

Note: Users should be aware of the following size limitations, which affect the editing and display of text in UniEdit windows at all times:

- A maximum of 6000 lines of text is permitted in <u>each</u> UniEdit document window (though limited Windows system resources may lower or slow this specification).

Tip: The TrueType Printer program provides a handy place for opening and viewing the first few pages of UniEdit Unicode files of virtually unlimited length. And since the file is opened in read-only mode, it has some of the speed and safety of other file viewers for DOS and Windows, like "LIST" or Xtree. You can edit a very long file in segments of less than 6000 lines, with all but the first segment saved as type "Unicode/No Header." Then use the DOS command "copy /b"file1 + file2 + file3 hugefile" to concatenate the separate segments into one big file. This is the procedure used for creating the large on-line UniEdit dictionaries.

- For word-wrapped text (with soft returns), there is a limit of approximately 5000 characters per paragraph (text between hard returns).

- UniEdit will handle, but not display properly, the ends of very long lines (greater than about 260 "Medium" 21-pixels-high characters in length. Such lines can be edited by breaking them into two or more shorter lines, then performing needed editing, then recombining the lines by deleting the hard returns. Alternatively, Word Wrap can be selected from the Edit menu, and the long line will be temporarily wrapped, if it is edited (e.g., insert a space character, then delete it) or the file is opened from disk with Word Wrap selected.

- When mixing right-to-left text (Arabic or Hebrew scripts) and left-to-right text, a maximum of 13 changes of direction is permitted per line (between hard returns).

UniEdit Document Windows

UniEdit arranges its windows and information like any other Windows application. Menus pull down, windows are maximized (opened full-screen) and minimized (reduced to icons), sized, and moved, all with the click of the mouse. Under the Window menu, options allow you to "tile" your windows, i.e. have them all displayed at once like tiles in a floor, or "cascade" them, i.e. have them overlap like folders with the most recently used ones on top. At any time you may also "arrange" your minimized window icons with the Arrange Icons menu option.

The active window, that is, the one in which you are able to enter information at any given moment, will be denoted by a change in intensity or hue of its title bar. To activate a window, you may click in it once or select it under the Window menu. Moving from window to window is that simple.

UniEdit has six separate document windows, each of which is like a Unicode Notepad. To switch from document to document, select the document window you desire from the Window menu. UniEdit works much like a plain-text editor like the Windows Notepad or the DOS Editor, but uses Unicode exclusively.

Accents and Diacritics

In order to produce diacritics (i.e. a modification of a character, such as an accent) and to accommodate more characters than the standard English keyboard normally allows, combinations of keys are sometimes used. In order to modify characters with diacritics, one must gain a familiarity with the keyboard layout.

The Add Accent Key = F2

The "add accent key" F2 is used to make accents and diacritics. It functions much like the Ctrl, Shift, and Alt keys on the standard keyboard. Pressing F2, or selecting the menu item "Add Accent," under Edit will highlight the letter to the left of the insertion point (with blinking caret) and the next letter pressed will overstrike the highlighted letter as a diacritical modifying accent (grave, acute, circumflex, tilde, dieresis, etc.).

Accent Keys = Pale Blue

Each diacritic is associated with a particular character on the keyboard, appearing in pale blue in the Keyboard Help virtual keyboard display. An individual diacritic may modify several characters, such as a single accent working with several vowels. In order to display the character

set associated with a particular diacritic, click on one of the light blue accent keys or press the letter on the keyboard.

Note: Because accents like the circumflex (^), the tilde (~), and the double quote ("), used to insert the umlaut, are shifted keys, Shift must be pressed along with the accent key.

The accent key will turn red and the associated characters will appear in green. In some cases there are more diacritics than displayed on the original Help keyboard. Clicking on or pressing the Shift key will display additional accents and diacritics.

Thus to see the full set of characters modified by the tilde (~), you must first press Shift + `, which will display the upper-case characters (shifted keys) modified by the tilde, then click on or press Shift again to display the lower-case characters (unshifted keys) modified by the tilde.
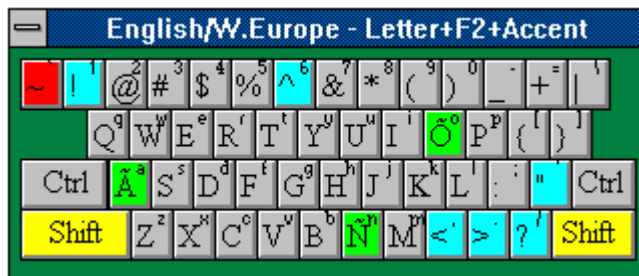
Putting the Keys Together

After the overstrike key, accent keys, and their associated characters have been identified, the three may be used in conjunction to create an accented character. The general procedure is to type the character itself, the overstrike key, and the accent key in that order. Notice when the overstrike key has been pressed, the character appears in reverse video, ready to be modified with the diacritic. For example:

Using the English/Western Europe keyboard (which supports all Roman languages) to create an ê:

1. Type the lower case "e"

2. Press the "overstrike key" F2

3. Press Shift+6

In order to create capital letters simply type the capital letter and then proceed as described above.

UniEdit Shortcut Keys

To	Press
Get on-line help	F1
Add Accent	F2
Switch between two currently used languages	F3
Look Up Word in Dictionary	F5

Chapter 3: UniEdit Menus

The File Menu

The items under the File menu allow you to create, open, or save any Unicode text file in the currently active document window. You may also import or export files in ANSI, IBM extended ASCII (DOS), or other formats.

New   Ctrl+N

When you are already working in a document window and you wish to quit working on the current document and begin working on a new Unicode text file, you must first open the File menu and select New (Ctrl+N), in order to tell the program that it must prepare a new file to store your commands. If you are working on another document at the time and you select New, UniEdit will ask if you want to save the current document first before opening a new file.

Open...   Ctrl+O

The Open menu item allows you to pick a previously saved file from a specified directory so that it may be examined or edited. When you are finished with the file, it may be Saved and reopened at any time. You must specify a path name for the document or scroll through directories to make the selection.

To open an existing document:
- From the File menu select Open (Alt+F,N) option.

- In the File Name box, type or select the document you want to open. If the document you want to open is not listed in the File Name box, do one or more of the following:

a.   In the Drives box, select a new drive.

b.   In the Directories box, select a new directory.

c.   In the List Files of Type box choose the type of heading which matches the format of the file. You can open the file as a Unicode/WC 2.0 file, a BOM-headed file, a No Header file, or an ANSI/WinCALIS 1.1 file, among others.

d.   Select OK.

UniEdit can import into UniEdit format files in the following file formats:

Unicode/WC 2.0

Unicode/BOM

Unicode/No-Header

ANSI/WinCALIS 1.1

Other conversions (many formats)

Unicode/WC 2.0

- For saving files in Unicode format with the UniEdit header (including the Unicode "Byte Order Mark" or BOM), identifying it as a file compatible with UniEdit.

Unicode/BOM

- For saving files in Unicode format preceded by the Unicode "Byte Order Mark," (to facilitate compatibility between machines with Intel ("little-endian") processors and machines with Motorola ("big-endian") processors).

Unicode/No-Header

- For saving files in Unicode format, without the UniEdit identifying header or the BOM.

ANSI/WinCALIS 1.1

- For saving ANSI-standard (ISO 8859-1) single-byte plain-text files to be read into a Windows/DOS-based product (for example, Microsoft Write for Windows or any DOS Text editor).

Other conversions

- For saving files in a variety of other single-byte, double-byte, or mixed single-and-double-byte national coding standards, such as are used in localized versions of DOS and Microsoft Windows.

For more information on "Other conversions," see the section UniEdit Conversions.

Save   Ctrl+S

When you open an existing document, UniEdit copies it from the disk where it is stored into RAM memory and displays it. As you work, you are actually making changes to a copy of the document in RAM. To ensure that your latest work is stored safely on the disk, you should frequently save your document. (This is easy to do—just press the shortcut key Shift+F4!) To permanently store the contents of your document, you must save it. This means that the information becomes recorded on some permanent media, such as a floppy disk, hard drive, or network storage system, and thus is retrievable for future use or modification. The document will be saved to the specified directory under the current name of the file. When selecting Save for the first time, you must provide certain information in order to tell UniEdit where to save the file.

To save a new, unnamed document:
- From the File menu, choose Save (Ctrl+S) OR

- From the File menu, choose Save As (Alt+F,V) or (Ctrl+F4).

- Type a name of up to eight characters in the File Name box which pops up. (UniEdit will automatically add an extension if you do not specify one.)

- Click on the OK button.

To save the document in a different drive or directory:
- From the File menu, choose Save (Ctrl+S) OR

- From the File menu, choose Save As (Alt+F,V) or (Ctrl+F4).

- In the Drives box, select a drive.

- In the Directories box, select a directory.

- Type the filename or the complete path in the File Name box.

- Select OK.

Save as...   Ctrl+F4

If you want to rename a previously saved document, make a backup copy, or save it in a new place, use the Save As... option. You may then use the Windows File Manager program to delete the old version if you so choose.

To save a UniEdit document with a different file format:
- From the file menu, under Text File select the menu item Save As (Alt+F,V or Ctrl+F4).

- In the File Name box, type a new name for the document.

- In the Save File as Type box, select the file format.

- Select OK.

UniEdit can convert files in UniEdit format to the following file formats:

Unicode/WC 2.0

Unicode/BOM

Unicode/No-Header

ANSI/WinCALIS 1.1

Other conversions (many formats)

For a detailed explanation of "Other conversions," see the section "UniEdit Conversions."

Print   Ctrl+P

Selecting the File, Text File, Print menu option transfers you to one of the two UniEdit printing utilities in order to print a document generated in the Scratchpad window or other window and saved as a text file on disk. If you wish to print directly from memory (and you are using only TrueType fonts), highlight the text to print, and choose the "Print Selection" menu option instead.

When you select "Print," a dialog box appears which allows you to choose which utility to use: the TrueType Printer program (MLPTT.EXE) or the Multilingual Printer program (MLP.EXE). Press Enter for the default selection, the TrueType Printer program.

The UniEdit package comes with two printer programs, the TrueType Printer Program (MLPTT.EXE), for text written in UniEdit entirely using a scalable TrueType font, and the Multilingual Printer program (MLP.EXE), for text in any language and script. (Currently UniEdit does not support TrueType fonts for Chinese, Japanese, Korean, a few other scripts such as Amharic, and miscellaneous symbol and exotic Roman character fonts.)

If your text has been written using a TrueType font (most alphabetic languages), you are advised to print with the TrueType Printer Program (MLPTT.EXE), which is faster, more powerful, and produces superior-quality printing at any font size. If you are not sure, experiment by printing a short test document with the TrueType Printer Program. Non-TrueType characters will merely not display properly or not display at all on the printed page.

If you cannot print your print job successfully with the TrueType Printer, you should resort to printing with the Multilingual Printer. This printer program is capable of printing any text which is displayed on your UniEdit screen, in any language or script, by printing all text, whether originally a raster font or a TrueType font, as graphics images.

You can also start either printer program by clicking on its icon in the UniEdit Pangaea group.

For more information on printing, see also "Printing a UniEdit Document."

Pr<u>i</u>nt Selection

Selected Text (marked and highlighted in reverse video) will be printed directly from memory by the TrueType Multilingual Printer program.

Note: The Print Selection feature will not work properly for non-TrueType fonts. This includes Chinese-Japanese-Korean, Amharic, text with IPA and other exotic Roman letters, and certain other languages supported by UniEdit for which TrueType printing fonts are not available.

Create <u>K</u>eyboard...

Creates a *.WCK binary language keyboard file from its ASCII text counterpart *.MAP file opened in the Scratchpad as an ASCII/ANSI file.

See Also "Defining Language Keyboard

Create <u>G</u>lobal File...

Creates a new WINCALIS.WGI binary global information file from its ASCII text counterpart WINCALIS.INF, opened in the Scratchpad as an ASCII/ANSI file.

See also "The UniEdit Global Information File

E<u>x</u>it   Alt+F4

When you are finished working with UniEdit, you should select the Exit command to return to the program manager. If you have made any changes to the current document, UniEdit will ask if you want to save the file.

The Edit Menu

<u>U</u>ndo   Alt+BkSp

Undo allows you to delete the last editing change made to the text and recall any affected text. For example, if you delete a paragraph from your text and then change your mind, you can employ the undo feature to bring back the paragraph. Undo may be used to undo an entire series of editing changes.

The shortcut key for Undo is Alt+BackSpace.

R<u>e</u>do   Alt+Shift+BkSp

Undoes the effect of Undo and restores the previous editing change, or series of editing changes.

The shortcut key for Redo is Alt+Shift+BackSpace.

Cu<u>t</u>   Ctrl+X

To cut a highlighted text block means to delete the block from the document and copy it to the clipboard, a temporary place where text is stored for later reinsertion using the Paste command.

The shortcut keys for Cut are Ctrl+X or Shift+Delete.

<u>C</u>opy   Ctrl+C

Copying a highlighted text block is like cutting it, except the block remains in place in the activated window. A copy of the selected text is placed on the clipboard so that it may be replaced in another location or window.

The shortcut keys for Copy are Ctrl+C or Ctrl+Insert.

<u>P</u>aste   Ctrl+V

Paste takes whatever information is on the clipboard put there by the cut or copy commands and inserts the text at the insertion point (where the blinking "caret" is located).

The shortcut keys for Paste are Ctrl+V or Shift+Insert.

<u>D</u>elete   Del

Delete removes the highlighted block from the active window, but does not place it on the clipboard. Text may immediately be restored using the Undo command.

Select <u>A</u>ll

Select All highlights all the text in the active window so it may be cut, copied or deleted. You may also highlight text by placing the arrow at the starting point, clicking and dragging the highlight to the end point of the text block you wish to select.

Find...

You can search for each occurrence of a combination of any characters, including uppercase and lowercase characters, whole words, or parts of words using the Find option.

To find text:
- From the Edit menu, choose Find (Alt+E,F).

- In the Find box type the text, word(s), or combination of characters you want to find.

- Set the options you want to control the search.

- Choose the Find Next button. If an occurrence is found, the screen scrolls so you can see the text in context.

Match Whole Word Only

Finds occurrences of the text as a whole word (no partial words). For example, the input "Tube" would not match up with the text "Tubes" if the "Match Whole Word only" option had been selected.

Match Case

Finds text having a certain pattern of uppercase and lowercase letters.

Note: A case sensitive search is much faster than a non-case sensitive search, since in the latter case, UniEdit must consult the UniEdit Global Information File for information on upper- or lower-case equivalents for every character in the text—even for those scripts which have no case distinction. Thus Match Case is checked by default.

Up

Searches the document above the insertion point.

Down

Searches the document below the insertion point.

Cancel

Cancels a search. Closes the Find dialog box when you are finished searching.

Note: To search for text in a non-ANSI character set, first select the appropriate keyboard from the L: [Language] menu.

Replace...

With the Replace command, you can find and replace text. You can automatically replace all occurrences of the found text, or selectively replace only certain instances. For example, if you wanted to replace the word "beast" with the word "monster" you would simply type "beast" in the Find What box, go to the Replace With box using the mouse or Tab key and enter the word "monster." If you wanted to replace every occurrence of the word, you would select Replace All.

To replace text:

- Select the portion of the document you want to search. UniEdit editor searches the entire document if you do not select a portion of the document.

- From the Edit menu, choose Replace (Alt+E,R).

- In the Replace box, type the text to be replaced.

- Set the options you want to control the search.

- To replace text, choose Replace.

- To leave text unchanged and search for next occurrence, choose Find Next.

- To cancel a Replace in progress or when you have finished replacing, select Cancel.

Match Whole Word Only

Finds occurrences of the text as a whole word (no partial words)

Match Case

Finds text having a certain pattern of uppercase and lowercase letters (a case-sensitive search).

Note: A case sensitive search is much faster than a non-case sensitive search, since in the latter case, UniEdit must consult the UniEdit Global Information File for information on upper- or lower-case equivalents for every character in the text—even for those scripts which have no case distinction. Thus Match Case is checked by default.

To replace the search text with nothing, make sure the Replace With box is empty by deleting the contents of the box.

To undo the effects of the Replace command, choose Undo from the Edit menu as the first action after completing the change.

Add Accent   F2

Selecting the menu item Add Accent or pressing the F2 function key highlights the character to the left of the insertion point (blinking caret) and prepares to overstrike that character with an accent determined by the next key pressed. See the section Accents and Diacritics.

Go to Line...

Moves the cursor to the line number (Alt+E,L) you specify in the Go to Line # edit box.

Switch Languages   F3

Switches between the currently selected language and the "top" language keyboard for the workstation (by default English/W.Europe) by using the function key F3.

Note: The "top" language is determined by the default language .INI file defined in the "DefaultIni=" line in the [WinCALIS] section of the main UniEdit initialization file "wincalis.ini," in the UniEdit program directory. The default language .INI file is set by default for the United States and Western Europe to "latin1.ini." If the default .INI file is changed to some other language .INI file, the "top" language will also be changed accordingly. The initial language for use on startup in UniEdit (also by default set to English/W.Europe) is also customizable by the UniEdit

user. The initial language is determined by the setting of the InitialLanguage= line in the [Initial] section of the "wincalis.ini" file. The out-of-the-box value for this line is Initial Language=English/W.Europe.

Unicode Input...   (Ctrl+I)

Allows you to browse through the Unicode character tables and insert a character in your text (Alt+E,I or Ctrl+I).

Unicode Character Selection Dialog Box

The Unicode Character Selection dialog box enables you to browse through the Unicode tables for your general information. You may also insert a character in your text by going to its location in the table, either by scrolling through Unicode character blocks or by typing all or some of its hexadecimal Unicode character code in the edit box in the lower right corner.

Hint: Clicking above or below the slider (or "scroll box")on the vertical scroll bar will scroll up or down a full "page" of 256 characters. To scroll up or down a half-page, so as to see all the characters on a page, press the up or down scroll arrows. If you move the focus to the window by clicking anywhere in the window, you may also move about with the cursor control keys. Pressing PageUp or PageDown will scroll up or down a full page, while the arrow keys will move about a row or column at a time.

The shortcut key to open the dialog box in UniEdit is Ctrl+I.
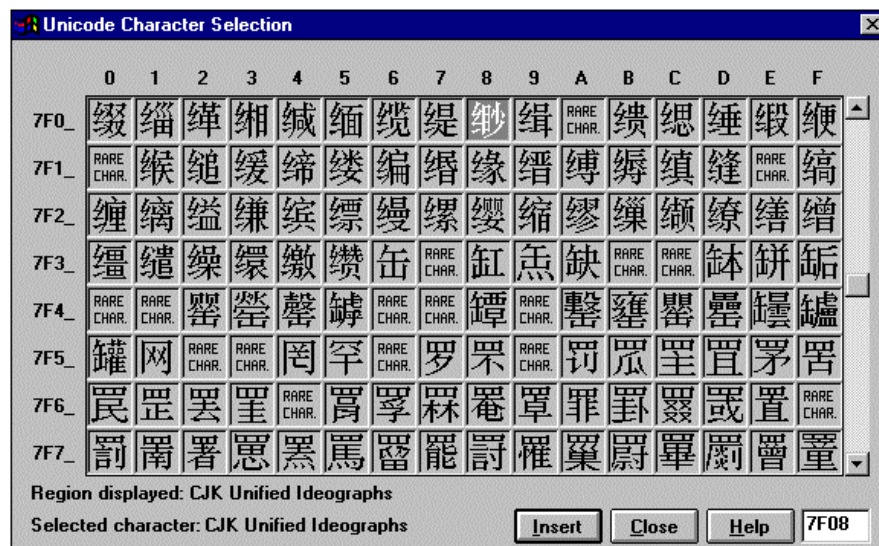
Unicode Character Blocks

Here is a list of Unicode character "blocks," with their starting and ending hexadecimal Unicodes:

Start	End	Block Name
0020	007E	BASIC LATIN
00A0	00FF	LATIN-1 SUPPLEMENT

0100	017F	LATIN EXTENDED-A
0180	024F	LATIN EXTENDED-B
0250	02AF	IPA EXTENSIONS
02B0	02FF	SPACING MODIFIER LETTERS
0300	036F	COMBINING DIACRITICAL MARKS
0370	03CF	BASIC GREEK
03D0	03FF	GREEK SYMBOLS AND COPTIC
0400	04FF	CYRILLIC
0530	058F	ARMENIAN
0590	05CF	HEBREW EXTENDED-A
05D0	05EA	BASIC HEBREW
05EB	05FF	HEBREW EXTENDED-B
0600	0652	BASIC ARABIC
0653	06FF	ARABIC EXTENDED
0900	097F	DEVANAGARI
0980	09FF	BENGALI
0A00	0A7F	GURMUKHI
0A80	0AFF	GUJARATI
0B00	0B7F	ORIYA
0B80	0BFF	TAMIL
0C00	0C7F	TELUGU
0C80	0CFF	KANNADA
0D00	0D7F	MALAYALAM
0E00	0E7F	THAI
0E80	0EFF	LAO
10A0	10CF	GEORGIAN EXTENDED
10D0	10FF	BASIC GEORGIAN
1100	11FF	KOREAN HANGUL JAMO
1E00	1EFF	LATIN EXTENDED ADDITIONAL
1F00	1FFF	GREEK EXTENDED
2000	206F	GENERAL PUNCTUATION
2070	209F	SUPERSCRIPTS AND SUBSCRIPTS
20A0	20CF	CURRENCY SYMBOLS
20D0	20FF	COMBINING DIACRITICAL MARKS FOR SYMBOLS
2100	214F	LETTERLIKE SYMBOLS
2150	218F	NUMBER FORMS
2190	21FF	ARROWS
2200	22FF	MATHEMATICAL OPERATORS
2300	23FF	MISCELLANEOUS TECHNICAL
2400	243F	CONTROL PICTURES
2440	245F	OPTICAL CHARACTER RECOGNITION
2460	24FF	ENCLOSED ALPHANUMERICS
2500	257F	BOX DRAWING
2580	259F	BLOCK ELEMENTS
25A0	25FF	GEOMETRIC SHAPES
2600	26FF	MISCELLANEOUS SYMBOLS
2700	27BF	DINGBATS
3000	303F	CJK* SYMBOLS AND PUNCTUATION
3040	309F	HIRAGANA
30A0	30FF	KATAKANA
3100	312F	BOPOMOFO
3130	318F	HANGUL COMPATIBILITY JAMO
3190	319F	CJK* MISCELLANEOUS
3200	32FF	ENCLOSED CJK LETTERS AND MONTHS

3300	33FF	CJK* COMPATIBILITY
3400	3D2D	KOREAN HANGUL
3D2E	44B7	HANGUL SUPPLEMENTARY-A
44B8	4DFF	HANGUL SUPPLEMENTARY-B
4E00	9FFF	CJK* UNIFIED IDEOGRAPHS
E000	F8FF	PRIVATE USE AREA
F900	FAFF	CJK COMPATIBILITY IDEOGRAPHS
FB00	FB4F	ALPHABETIC PRESENTATION FORMS
FB50	FDFF	ARABIC PRESENTATION FORMS-A
FE20	FE2F	COMBINING HALF MARKS
FE30	FE4F	CJK* COMPATIBILITY FORMS
FE50	FE6F	SMALL FORM VARIANTS
FE70	FEFE	ARABIC PRESENTATION FORMS-B
FF00	FFEF	HALFWIDTH AND FULLWIDTH FORMS
FFF0	FFFD	SPECIALS

* "CJK" stands for Chinese-Japanese-Korean. For further Chinese-Japanese-Korean help, see the Kangxi Radical Chart in Chapter 5.

Insert

The Insert submenu allows the user to place in text various non-displaying control characters, such as joiners, zero-width spaces, and other special characters, to achieve certain special effects like overriding language directionality.

Joiner   Shift+F5

This character can be used to maintain certain contextual forms that would not otherwise occur. For instance, if a character such as an Arabic letter has an initial, medial, and final form, you may insert joiners in order to force the system to maintain an independent medial or final form.

The shortcut key to insert a Joiner is Shift+F5.

Non-Joiner   Shift+F6

The Non-Joiner works in much the same way as the joiner. You may use this option to prevent two characters from joining when they otherwise would.

The shortcut key to insert a Non-Joiner is Shift+F6.

Left to Right Override   Shift+F7

Right to Left Override   Shift+F8

These options allow you to manually determine the direction in which text will appear on screen. This could be useful if, for instance, you were writing an English sentence that started with a Hebrew word, and you wished to force the sentence to be left-justified, or conversely, a Hebrew sentence that began with an English word, that you wished to be right-justified.

The shortcut key to insert a Left-to-Right Override is Shift+F7.

The shortcut key to insert a Right-to-Left Override is Shift+F8.

Zero Width Space   Shift+F9

This character is a no-width space character which can be used to identify word boundaries in languages that do not have spaces between words.

The shortcut key to insert a zero-width space is Shift+F9.

Word Wrap

Allows you to wrap text (Alt+E,W); text retrieved from disk or in any line being edited is wrapped at the right margin of the maximum window size for the screen resolution of your system, as reported by Windows.

The L: [Language] Menu

General Language and Keyboard Information

The language pull-down menu (Alt+L) allows you to select any of the language keyboards available in UniEdit. Depending on the number of fonts and keyboard layouts available on your system, the number of languages may vary from just a few to most of the languages of the world.

Western European languages such as Danish, English, French, German, Italian, Portuguese, Spanish, and Swedish are all available with the ANSI (Latin1) Roman character set and default "English/W.Europe" keyboard (no keyboard change required). All accents are typed with three keystrokes (base letter + F2 + accent).
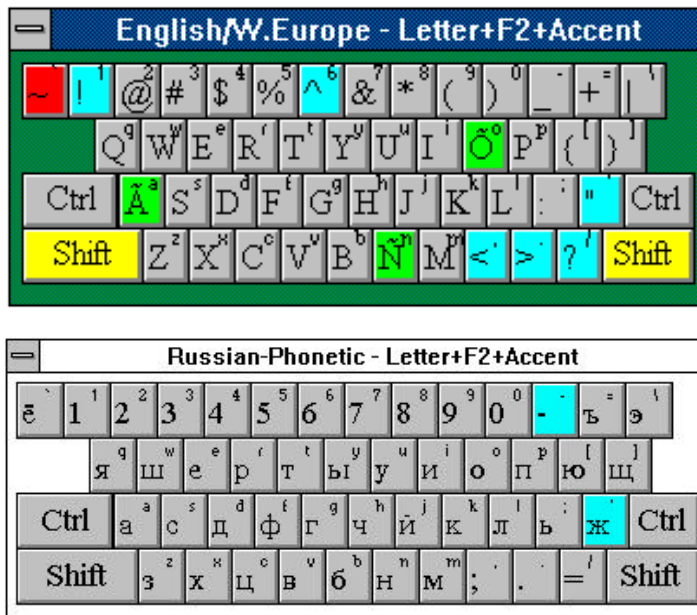
Other languages, such as Russian or Ukrainian, with the Cyrillic character set, or Arabic or Persian (Farsi), with the Arabic character set, require entirely different fonts and keyboards. A

symbols !, @, #, $, %, ^, &, *, or |.

When you change languages in the middle of a document, only the new text that you type will appear in the selected language. Previously written text will not be affected.

<u>A</u>dd Language...

Adds a language keyboard item to your L:[Language] menu in UniEdit from among those available in your UniEdit program directory (i.e., for which there exist *.INI, *.MAP, and *.WCK files).

<u>D</u>elete Language...

Deletes a language keyboard item from your L:[language] menu. Only the menu listing is deleted. The *.INI, *.MAP, and *.WCK files for each language keyboard are not deleted from your UniEdit program directory.

Note: When you install UniEdit, it installs *.INI, *.MAP, and *.WCK files for all available languages into your program directory by default. Add Language allows you to access any available language. If you need to economize on disk space, after deleting a language from your menu, you may also use the File Manager to delete its *.INI, *.MAP, and *.WCK files from your hard disk, or move them to a backup storage location. *.MAP files are simply the text versions of binary keyboard *.WCK files, which are those used by UniEdit,. If you do not anticipate the need to edit or add any keyboard file, you may delete *all* the *.MAP files from your system. See Deleting Nonessential Files to Save on Disk Space.

The Tools Menu

Language <u>O</u>ptions

Opens additional menus for options pertaining to specific language input (e.g.. Chinese, Japanese, Korean)

<u>C</u>hinese Options...   Alt+F11

Opens the dialog box for setting Chinese dictionary search options. See Chinese (requires separate CJK module).

Chinese S<u>t</u>atus Bar

When you choose Chinese Pinyin-to Hanzi as the active language keyboard, the Chinese Status Bar will pop up by default, displaying the current settings for Chinese input.

The Status Bar may be turned off by deselecting it here. If Chinese Pinyin-to-Hanzi is not the currently selected language, the menu item is dimmed.

See Chinese Status Bar (requires separate CJK module).

<u>J</u>apanese Options...

The View Japanese Dictionary allows you to view and edit the Japanese internal Kana-Kanji conversion dictionary. (Currently disabled.) See Japanese (requires separate CJK module).

<u>K</u>orean Options...

Holds the Hanja-to-Hangul (F7) and Hangul-to-Hanja (F8) conversion functions. See Korean (requires separate CJK module).

<u>S</u>ort Lines...

Sorts the selected lines of text in Unicode (alphabetical) order. To browse the Unicode character selection and order, choose Edit, then Unicode Input (or Ctrl+I), then use the scroll bar.

G<u>e</u>nerate Index

This menu option in WinCALIS Author or UniEdit runs the Unicode index builder to build an index for a dictionary file in the default "Dictionaries" subdirectory. The dictionary should be saved as type "Unicode/No Header".

See How To Create Your Own Dictionary

<u>L</u>ook Up Word

If one or more on-line dictionaries have been supplied to accompany your copy of UniEdit, selecting this menu option will initiate an on-line dictionary search. A dialog box will open, allowing you to look up the meaning or pronunciation of a selected word or phrase. Depending on how the dictionary was installed, it is either accessible only from specific language keyboards (e.g, for a Korean-English dictionary, the current keyboard must be one of the Korean keyboards), or accessible from any language keyboard which does not have a specific dictionary linked to it (e.g, for a Korean-English dictionary, the current keyboard could be either the default English/W. Europe keyboard or any of the Korean keyboards).

In order to select a word or phrase to look up, drag the mouse across it (or double click on it if it is a single word) to highlight it. Then press F5 or click on the menu item Too<u>l</u>s, <u>L</u>ook Up Word... to look up the word. Repeat as often as desired.

You can press the Escape key at any time during a search to cancel further searching and display the matches found thus far.

See Also: User-Supplied On-line Dictionaries

Set <u>T</u>ools

This menu option displays the Set Tools dialog box, enabling you to add, delete, or move menu items on the Tools menu.

Check <u>S</u>pelling

If a spelling checker has been installed on your workstation and linked to UniEdit, selecting this option will start the spelling checker. A copy of the spelling checker WinSpell® 3.0 is included with the UniEdit workstation setup.

Print File (<u>T</u>rueType)

Selecting this item launches the UniEdit TrueType Multilingual Printer Program.

Print File (<u>B</u>itmap)

Selecting this item launches the UniEdit Graphics Multilingual Printer Program.

<u>K</u>eyboard Macro

Selecting this item launches the Windows Macro Recorder, enabling you to define a keyboard macro to type automatically and rapidly a complex or repetitive series of keystrokes.

Tools Menu <u>R</u>eadme

Selecting this item displays information about the Tools menu and some of its functions.

The Window Menu

Tile <u>H</u>orizontally

Arranges open windows above and below each other in the UniEdit editor screen.

Tile <u>V</u>ertically

Arranges open windows side by side in the UniEdit editor screen.

C<u>a</u>scade

Arranges open windows partially overlapping, with their title bars visible.

Arrange <u>I</u>cons

Arranges the icons for minimized windows across the bottom of the screen

<u>D</u>esktop

UniEdit provides you with a customizable user environment. The Desktop menu options allow you to customize your UniEdit desktop, including such aspects as which windows are open, their

size and shape, whether or not to display the status bar at the bottom of the screen. the default startup language keyboard, and Chinese editing and dictionary search options.

S**a**ve

Allows you to save the current desktop.

Restore   Ctrl+7

Restores the Initial desktop.

Def**a**ult   Ctrl+8

Restores the default desktop configuration as installed by UniEdit Setup.

Status Bar

Selecting this item displays a Status Bar at the bottom of the screen.

1 Document1 (Ctrl+1)

Clicking on this menu option or pressing the shortcut key Ctrl+1 switches to the window for Document 1.

2 Document2 (Ctrl+2)

Clicking on this menu option or pressing the shortcut key Ctrl+2 switches to the window for Document 2.

3 Document3 (Ctrl+3)

Clicking on this menu option or pressing the shortcut key Ctrl+3 switches to the window for Document 3.

4 Document4 (Ctrl+4)

Clicking on this menu option or pressing the shortcut key Ctrl+4 switches to the window for Document 4.

5 Document5 (Ctrl+5)

Clicking on this menu option or pressing the shortcut key Ctrl+5 switches to the window for Document 5.

6 Document6 (Ctrl+6)

Clicking on this menu option or pressing the shortcut key Ctrl+6 switches to the window for Document 6.

The "Unsaved" Asterisk *

When changes are made to any document, an asterisk * will be added to the title bar to indicate that changes have not yet been saved. Once the document is saved, then the asterisk * will disappear.

Arranging Windows and Icons

UniEdit enables you to arrange your desktop so that windows and icons are easy to see. The Tile Horizontally option under the Window menu arranges open windows above and below each other in the UniEdit editor window. The Tile Vertically option arranges the windows side by side in the UniEdit editor window. The Cascade command resizes and layers windows so that each title bar is visible.

To arrange windows on your desktop:
- Choose Cascade or Tile from the Window menu.

- Use the Arrange Icons command to evenly arrange the icons in a group window.

To arrange program-item icons:
- Choose Arrange Icons from the Window menu.

To activate the editing window in UniEdit:
- Using the mouse, click anywhere in one of the UniEdit windows OR

- Select Window from the menu bar, and click on one of the six available editing windows (Ctrl + 1-6)

The Help Menu

In addition to the menu options explained below, UniEdit has the following additional context-sensitive Help features, available from the keyboard:

- Keyword Search (Shift+F1)

Put your caret anywhere on a CALIS command, task option, special response, or global variable keyword (in capital letters), or any other defined Windows Help keyword and press Shift+F1 to call up the Help Topic(s) on that item.

- Context-Sensitive Help

Highlight any menu item and press F1 (or Shift+F1 or Ctrl+F1) to go directly to the Help topic(s) for that menu item. See also the Context-Sensitive Help menu item below.

<u>C</u>ontents   F1

The Contents menu option provides a list of main topics in the on-line User's Manual available in On-line Help, to help you understand various aspects of UniEdit. The gray Search button at the top of the Help window is particularly useful. You can search for any keywords or topics using the Search feature and Help will find it and take you right to it. Click on "Full Table of Contents"

to open a separate Table of Contents window. Or press the "Glossary" button for a glossary of Windows and UniEdit terms, or the "Index" button for a book-style index to the on-line User's Manual.

Index

Opens UniEdit On-line Help and takes you directly to the Index.

Search for Help on...

Opens UniEdit On-line Help and takes you directly to the Search dialog box.

Context-Sensitive Help   Ctrl+F1

Turns the mouse pointer into a Question-mark context-sensitive help pointer. You can then click with the Question-mark mouse pointer on any keyword in a UniEdit command, task option, special response, or global variable keyword or on any pull-down menu item for which information is desired.

Keyboard Help

Keyboard Help displays on screen a picture of the current keyboard layout, with color-coded accent overstrike keys and the letters which they affect.

About Multilingual Text Editor (UniEdit)...

About Multilingual Text Editor (UniEdit) shows the current version, date, and registration number of your copy of UniEdit.

Chapter 4: Printing a Document

You can use UniEdit printing facilities to print any document created in UniEdit and saved as a file on disk. The UniEdit Workstation package comes with two printer programs, the Multilingual Printer program (MLP.EXE), for text in any language and script, and the TrueType Printer Program (MLPTT.EXE), for text written in UniEdit entirely using a scalable TrueType font. (Currently UniEdit does not support TrueType fonts for Chinese, Japanese, Korean, a few other scripts such as Amharic, and miscellaneous symbol and exotic Roman character fonts.)

If your text is by default written in a TrueType font (most alphabetic languages), you are advised to print with the TrueType Multilingual Printer Program (MLPTT.EXE), which is faster, more powerful, and produces superior-quality printing at any font size. You can start this program by clicking on its icon in the UniEdit Pangaea group. If you are not sure, experiment by printing a short test document with the TrueType Printer Program. Non-TrueType characters will merely not display properly or at all on the printed page. If this affects your print job, you should resort to printing with the Multilingual Printer.

Multilingual Printer Program (MLP.EXE)

Selecting the File, Print menu item in UniEdit the displays the Select Printer Utility dialog box. Selecting "Graphics Printer" launches the Multilingual Printer program (MLP.EXE). Because the printing process takes a snapshot of every page it sends to the printer, the process is a bit slow. However, this method allows UniEdit to work with virtually any printer, and enables you to generate printed material in any language supported by UniEdit.

When you choose Print (Ctrl+P) from the File menu, the Select Printer Utility dialog box opens. Choose "Graphics Printer." Within the Multilingual Printing dialog box, you will be presented with the following menu items:

Settings
Print
Printer Setup
Print Help

<u>S</u>ettings

The default settings for margins, paper size, and page orientation are set so you can produce attractive documents without having to alter the settings.

Margins

The settings you choose for margins affect all pages of a document. The default settings for margins are fixed so you can produce attractive documents without having to change the settings.

Margin settings determine the point at which text appears on a page. If you want to emphasize certain paragraphs, you can indent them to set them off from the margins. You can change the script's length and improve readability of a document by altering its margin settings.

Type the measurements for the margins you want to adjust: Top, Bottom, Left, and Right. The measurements must be set in inches.

Header

Prints the filename as a header at the top of each page.

Text

You can use Text formatting to align text and adjust spacing in your script. You can apply Font Size, Line Spacing, and Kerning.

Font Size

Choosing the appropriate font size for text can set the tone for a document and affect its readability by varying the pixel size in your script. Your selection of sizes is determined by your printer.

Line Spacing

The default setting of 0.02 inches prints text approximately as spaced on the screen. However, you can increase or decrease the amount of white space between lines in a document.

Kerning

You can change the spacing or kerning of characters to make the script more readable.

Print Scaling

You can change the size of the font in which your document is printed at a given printer resolution, such as 150 dots per inch, from teeny-tiny (Print Scaling 1 -- i.e., 1 pixel on the screen is printed as 1 dot on the paper), to medium (Print Scaling 2 -- i.e., 1 pixel on the screen is doubled both horizontally and vertically as 4 dots on the paper), to magnified in size, but not quality, through further multiplication of pixels (Print Scaling 3 and above). Default is Print Scaling 2.

Print

Using the Print option you may print any text file.

To print a document:
- From the UniEdit File menu, choose Print (Ctrl+P) to open the Select Printer Utility dialog box. Choose "Graphics Printer." Or start the program by clicking on its icon in the UniEdit Pangaea group.

- Next, in the Print menu find and select the file you wish to print.

- Then press the Print button to print the document or press the Preview button to view the document before printing.

Pr**inter Setup**

The Multilingual Printer will print on the printer installed as the Windows default printer. If more than one printer is installed, use the Windows Control Panel, Printers "applet" to set the desired printer as the default printer. Printing options are printer-specific settings, including

resolution, color, orientation, text quality, and others. For additional information on changing the settings, consult your Windows manual.

TrueType Printer Program (MLPTT.EXE)

All of the UniEdit fonts can be printed using the Multilingual Printer (MLP.EXE). However, UniEdit also supports most alphabets in TrueType fonts. Thus when printing text as TrueType, you may want to try the TrueType Printer (MLPTT.EXE). You may use the TrueType Printer for alphabets such as Arabic, Armenian, Cyrillic, Greek, Georgian, Hebrew, and Thai, as well as certain other alphabetic languages. To receive updates on character sets, please contact Humanities Computing Facility or its agents.

The TrueType Printer program also provides a handy place for quickly opening and viewing UniEdit Unicode files. And since the file is opened in read-only mode, it has some of the speed and safety of other file viewers for DOS and Windows, like "LIST" or XTree.

When you choose Print (Ctrl+P) from the UniEdit File menu, the Select Printer Utility dialog box appears. Choose "TrueType Printer." Then in the TrueType Multilingual Printer program open the File menu. You will be presented with the following menu items:

New
Open...
Print...
Print Preview
Print Setup...
Print Settings...

You will also see a list of the four most recently opened document files.

Print...

Using the Print menu option you may print any document file saved to disk.

- **To print a document:**
    - From the UniEdit File menu, choose Print (Ctrl+P) to open the Select Printer Utility dialog box. Choose "TrueType Printer." Or start the program by clicking on its icon in the UniEdit Pangaea group.

    - Next, using the Open option in the TrueType Multilingual Printer File menu, open the file you wish to print.

    - From the File menu, choose Print to print the document or Print Preview to view the document before printing.

Print Preview

Print Preview shows an entire page at a reduced size. You can check page breaks, position of text, and layout. When you choose Print Preview (Alt+F,P) from the File menu, UniEdit printing module displays a Document dialog box with the following menu items:

Menu Item:	**Function:**
Print	Prints the document.
Next Page	Displays the next page.
Previous Page	Displays the previous page.
Two Page	Displays two pages at once.
Zoom In	Controls portion of the page you see in your document window. Allows you to magnify part of a page.
Zoom Out	Controls portion of the page you see in your document window. Allows you to reduce part of a page.
Close	Closes the File menu and returns you to the Multilingual Printing dialog box.

Print Setup...

Printing options are printer-specific settings, including resolution, color, orientation, text quality, and others. For additional information on changing the settings, consult your Windows manual.

Print Settings...

The default settings for margins, paper size, and page orientation are set so you can produce attractive documents without having to alter the settings.

Chapter 5: Language Reference Manuals

Arabic and Hebrew (Right-to-Left) Scripts

Right-to-Left and Left-to-Right Languages Basic Information

Cursor Position

The position of the cursor at the beginning of a line is determined by the direction of the chosen language. For example, if you choose Arabic the cursor will appear on the right side of the window. If before typing, you switch to English the cursor will flip to the left of the window.

Cursor Movement

The direction the cursor moves is determined by the first non-neutral/directional character of the line. If the first character encountered by the UniEdit editor is of type left to right, the cursor will follow that direction until the line break (you press ENTER) even if the language type has changed in the middle of the line. The same holds true if the first character the UniEdit editor encounters is of the form right to left. In those instances in which a neutral character begins a line UniEdit will not make a judgment until it encounters the first non-neutral character.

Text Order

Although the directional movement of the cursor does not change within a line, all text will appear in its proper order. For example, if the cursor is moving from left to right and you change to a language written from right to left, the text will be 'pushed' out from the cursor point. If you switch languages again the cursor will jump over the text it 'pushed' out and continue the line. This will occur in the same manner if the cursor is moving right to left

Switching Directions

When mixing right-to-left text and left-to-right text, a maximum of 13 changes of direction is permitted per line (between hard returns).

Chinese

There are three different Chinese input methods available:

• Pinyin Phonetic Conversion

• Cangjie Input Based on Character Shape

• Character Input through Browsing Unicode

Of these, the normal method for typing and editing text is the first, pinyin phonetic method, which has many sophisticated features to enhance efficiency. The other methods are intended to supplement pinyin input, when you need to type a rare character, or a character the pronunciation of which you are uncertain of.

There is also a keyboard available for typing the Chinese Phonetic Alphabet ("Zhuyin Fuhao" or "Bopomofo") characters. This is not a method for entering Chinese characters.

See Also

"The Cangjie Input Method"
"Kangxi Radical Chart"

For users from Taiwan who do not know the pinyin romanization, see "Zhuyin Fuhao-Pinyin

For information on inserting characters by browsing through the Unicode tables, see "Unicode

The Pinyin Phonetic Input Method

This technique is a phonetic conversion method, based on entering Modern Standard Chinese in phonetic transcription in the Pinyin romanization, with or without tones. WinCALIS Author editor takes the pinyin text and converts it into the accurate Chinese characters each time you press the space bar or another conversion function key after typing the phonetic version of a Chinese word or character.

Note: When mixing Chinese with any Roman-script language (such as English), a space, parenthesis, or other punctuation mark should separate the Roman language from the pinyin in order to obtain a proper conversion.

Conversion Preference

WinCALIS Author editor takes any unconverted pinyin to the left of the cursor and converts it when the space bar is pressed. If the string of characters to the left of the cursor does not constitute a valid pinyin syllable or syllables, a message "Invalid Pinyin" will appear on the screen.

WinCALIS Author handles the conversion process through a lookup table connected to an internal dictionary of Pinyin-Hanzi equivalents, with over 40,000 entries, including about 17,000 Chinese characters in both simple (modern) and complex (traditional) styles, as well as polysyllabic words. If the correct character(s) cannot be determined, because there are several homophones

with the same pinyin, a Chinese Selection dialog box will appear and display some homophone alternatives, arranged in order of descending frequency.

To Convert or Not to Convert?

In order to convert a collection of pinyin words previously typed, place the cursor at the end of the pinyin phrase to be converted and press the space bar or other appropriate function key. If a match can not be found within the dictionary, a message "Invalid pinyin" is displayed. Either press the space bar or use the mouse and click on the OK button to clear the message. You can move the cursor back one syllable at a time, until part of the phrase is converted, then convert the remaining phrase or choose individual syllables and convert them one at a time with the [F7] or [F8] function key.

Watch Your Tones!

When typing Chinese in WinCALIS, the words must not contain any spaces or hyphens in pinyin transcription. Enter tone marks with the number keys 1-4, then press the most suitable of several different Convert keys. The internal dictionary will search for a correct character translation and try to match what you have typed. A Chinese translation is either inserted at the cursor position of your text, or if more than one possible word is detected the Chinese Selection dialog box will appear and you can choose the appropriate homophone. To indicate your selection click the mouse on the word of your choice, or press the space bar to make the Selection Window active, move up and down the list with the arrow keys until the desired word is highlighted, then press ENTER. You can cancel the window without making a selection by pressing the ESC key.

What is Your Pinyin Status?

After each conversion, the correct toned pinyin is displayed in the Chinese Status Bar at the bottom of the screen. The display of the pinyin text is intended to help you become familiar with the entry that was successfully converted by WinCALIS, so that you may make an informed choice in situations where there is some uncertainty as to the correct pinyin form, tone, and so on. This is particularly handy when you type a pinyin form without tones and press the F8 key for an all-tone search of the dictionary. By typing a toned form you can usually avoid the need to select from the homophone selection window.

You have a number of options for converting pinyin to characters. Most of the options are selected with either the space bar or one of the function keys.

Functions of the Space Bar

When Chinese is the selected language in WinCALIS or WinCALIS Author, the space bar functions as follows:

- **Chinese Space**

When there is no pinyin text preceding the cursor, the space bar inserts a blank space, the length of one Chinese character, which is equal to over two English or pinyin-size characters, at the current cursor position.

- **Convert-Word Key**

When the space bar is pressed after typing a word in pinyin transcription including the proper tones, WinCALIS will search its internal dictionary for a free-form word which fully matches the

typed pinyin pronunciation, including tone: e.g., , , etc. All characters which are only bound forms (parts of words, e.g.) are excluded, with the exception of the special case of capitalized proper nouns. If more than one word with the same pronunciation (homophone) is found, the choices are displayed in the Chinese Selection dialog box. If you do not see the desired word in the Chinese Selection window, particularly in the case of single characters, cancel the window by pressing the ESC key, then press function key F7 for a full character search. Or simply press F8 for an all-tone search (tone is ignored).

If a proper noun, such as a surname (e.g., Zhèng), the name of a dynasty (e.g., Táng), or a place name (e.g., Beijing), is typed with an initial capital letter, it is converted directly when the space bar is pressed, even if it is a bound form or there are other non-proper-noun homophones. If there are two or more such matches (e.g. Zhang, Zhang), they are displayed in a selection window.

Selecting Options from the Chinese Selection Dialog Box

When a Chinese Selection box is displayed in the screen, a word can be selected by using the arrow keys to move from the first initial highlighted pinyin to the desired word as an alternative to using the mouse.

The vowel ü (u-dieresis) (u with two dots over it) is typed with the 'v' key or CTRL + u. The rare vowel ê (e-circumflex) is typed with CTRL-e. Each of these vowels is also modifiable by the tone marks 1-4.

When tone marks are added to a syllable, they are automatically placed over the correct vowel in each syllable, as long as they are typed after the first occurring vowel in each syllable. By convention, students should be encouraged to enter the tone right after the vowel it will appear over (like dotting an "i" or crossing a "t" immediately in printing) or at the end of each syllable (like dotting "i"s and crossing "t"s after the end of a word in handwriting).

Certain phrase-final punctuation marks also serve as convert keys during Chinese input, thus making it unnecessary to first press the space bar, then press the desired punctuation mark key. (Not yet implemented.) These marks include the period (。), clause comma (，), serial comma (、), exclamation point (！), question mark (？), and "yuandian" (●). Each of these marks is typed by pressing the designated key on the keyboard, which is also used for the smaller corresponding English punctuation marks during English input. This is also true of the Chinese-spaced quotation mark and Chinese space.

Other punctuation marks unique to Chinese are entered by pressing special keys. These are as follows:

Chinese serial comma (、)	Grave accent key (`)
Angle-bracket open quote (「)	Shift+2
Angle-bracket close quote (」)	Shift+3
Double angle-bracket open quote (『)	Shift+4
Double angle-bracket close quote (』)	Shift+5
"Yuandian" 圆点 separator mark (‧)	Shift+6
Chinese ditto mark (夕)	Shift+7.

All keyboard key assignments can be displayed by clicking on main menu Help then selecting menu item Keyboard Help. A virtual keyboard picture will come up on the screen.

Chinese Function Keys

You can use the function keys on your keyboard to perform many tasks faster:

F1   Get help on the selected command, dialog box, or message

F2   Highlight the character to be accented (Accent Overstrike Key)

F3   Switch between two currently used languages

F4   Reselect ("Oops") key. (WinCALIS)

If a word is incorrectly converted, press F4 immediately to display the Selection Window dialog box. Reselecting the proper word from the Selection Window will automatically replace the incorrectly converted word.

Save current script file and continue editing (WinCALIS Author)

F5   Look up highlighted word in On-line Dictionary

F6   Expand Scope of Selection (Selection Window)

Broadens the range of choice in the Chinese Selection Window when pressed during display of a Chinese Selection window.

Allows student to give up (WinCALIS)

F7   Character Search

Initiates a dictionary search for all entries in the dictionary which match the typed pinyin, including tone, regardless of whether they are free words or bound morphemes.

F8   All-tone Dictionary Search

Initiates a dictionary search or all characters or words which match typed pinyin, regardless of tone; any tone marks displayed will be ignored. In All-Tone mode, initiates a search for only those characters or words with the tone typed.

F9   Reselect ("Oops") key. (WinCALIS Author and UniEdit)

If a word is incorrectly converted, press F9 immediately to display the Selection Window dialog box. Reselecting the proper word from the Selection Window will automatically replace the incorrectly converted word.

F10   Default Selection/Duplicates the last word entered.

Avoids selection from the Chinese Selection window by automatically selecting and displaying either the most common among two or more homophone words or the one most recently selected from a Chinese Selection window (ignores bound forms); each time a different selection is made, that becomes the new "default" value. If pressed without any pinyin at the cursor, F10 repeats the last selection made from a selection window (useful for reduplication, or for quick conversions of the same word in multiple locations).

F11   Play Audio (WinCALIS)

F12   Play Video (WinCALIS)

Alt+F3   Set Cache/No-Cache Mode

In Cache mode, entries selected from the main dictionary on disk are added to a temporary cache, and displayed even in No-Disk mode along with memory-dictionary entries. No-Cache mode disables this feature.

Alt+F4   Close the application

Alt+F5   Set Convert/No-Convert (literal Pinyin)

In Convert mode, pinyin text at the cursor is converted into Chinese characters. In No-Convert mode, pinyin text is left unconverted.

Alt+F7   Alternate between Word and Character Mode.

In Word Mode, pressing the space bar searches for words (free forms) only. In Character Mode, pressing the space bar searches for all forms, both free and bound.

Alt+F8   Alternate between using the space bar for Tone Mode and All-Tone Mode.

In Tone Mode, pressing the space bar searches for toned forms only. In All-Tone Mode, pressing the space bar searches for all forms, regardless of tone.

Alt+F9   Alternate between Simple and Complex character style.

Changes the conversion of the space bar and other encode keys between complex-style (traditional) and simple-style (modern) characters.

Alt+F10   Alternate between Default and No-Default Search.

In Default Mode pressing the Space Bar converts the default form either the most frequently occurring form or the most recently selected form) without any Selection Window. In No-Default Mode pressing the space bar displays a Selection Window.

Alt+F11   Displays the Chinese Dictionary Search Option Dialog Box.

Alt+F12   Alternate between Disk Mode and No-Disk Mode.

In Disk Mode, pressing the space bar immediately searches both the memory and disk dictionaries, producing a larger selection window, with more selections. In No-Disk Mode, pressing the space bar searches the memory dictionary and the cache for matching entries. Single entries are directly converted. If there is a Selection Window, further entries from the main disk dictionary are displayed by pressing the "More" button or F6.

Your favorite settings may be saved as part of the WinCALIS "desktop" in WinCALIS Author, by clicking on Script, Desktop, Save. The original settings are "Word, Tone, No-Default, Cache, Convert, Simple," which for the typical author or student user should produce the most efficient and accurate encoding of pinyin phonetic input into Chinese characters.

Chinese Status Bar

When Pinyin-to-Hanzi is the current language keyboard, a Status Bar at the lower left corner of the screen displays the current settings for Chinese input.
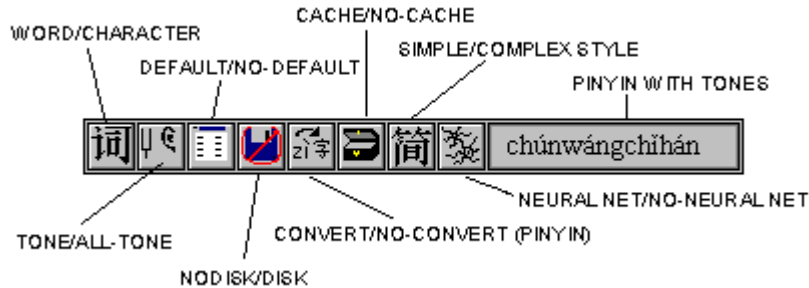
These settings can be changed by clicking on the buttons with the mouse, by pressing the corresponding Alt + function key combination, or by displaying the Chinese Dictionary Search options dialog box with Alt+F11 and clicking on the desired check box. You can view the following in the Status Bar:

the WORD/CHARACTER button,

the TONE/ALL-TONE button,

the DEFAULT/NO-DEFAULT button,

the NODISK/DISK button,

the CONVERT/NO-CONVERT (PINYIN) button,

the CACHE/NO-CACHE button,

the SIMPLE STYLE/COMPLEX STYLE button,

the NEURAL NETWORK/NO-NEURAL NETWORK button and

the PINYIN word, showing the correct toned form of the last word converted from Pinyin to Hanzi.

Special Pinyin Forms

Certain words occur so often that in addition to their regular forms, they have special alternative toneless or non-standard pinyin forms which make them easy to enter, without the need for selection from the Homophone Selection Window. Here is a list of some of these ad hoc forms. Note: This list includes just a sample of the neutral-tone "particles" which are normally typed without tone, such as "le," "ne," "zhe," etc.

不	b, bu (and bú, bù)	她	taa (and tā)
的	d (and de)	他们 (他們)	tamen (and tǎmen)
地	di (and de, dì)	她们 (她們)	taamen (and tǎmen)
个 (個)	ge (and gè)	它	tuo (and tā)
过 (過)	guo (and guò)	也	ye (and yě)
很	hen (and hěn)	一	yi (and yī)
就	jiu (and jiù)	在	zai (and zài)
里 (裏)	li (and lǐ)	再	zay (and zài)
儿 (兒)	r (and ér)	着 (著)	zhe (and zháo, zhāo, zhuó, zhù)
是	shi (and shì)	子	zi (and zǐ)
他	ta (and tā)		

Text in pinyin transcription also can be entered in your document, with or without tonal diacritic marks, by clicking on the NO-CONVERT button or pressing Alt+F5. This should facilitate the production of printed teaching materials, glossaries, and dictionaries.

National Phonetic Alphabet ("Zhuyin Fuhao

Under the Chinese language submenu, selection of the "Bopomofo" language enables entry of the National Phonetic Alphabet symbols widely used in Taiwan, and commonly known as bo-po-mo-fo, in your text. The keyboard mapping is the standard one used in Taiwan. Note that this is not a bo-po-mo-fo input method, but rather a way of entering these symbols in your text, such as might be desired in a glossary or annotated text.

Zhuyin Fuhao-Pinyin Correspondence Table

For users familiar with the Chinese Phonetic Alphabet ("Zhuyin Fuhao" or "Bopomofo"), but not with Pinyin romanization, the following chart of rough equivalents may be useful:

ㄅ b	ㄉ d	ㄍ g	ㄐ j	ㄓ zh(i)	ㄗ z(i)	ㄧ (y)i	ㄚ a	ㄞ ai	ㄢ an	ㄦ (e)r	ㄧㄡ iu
ㄆ p	ㄊ t	ㄎ k	ㄑ q	ㄔ ch(i)	ㄘ c(i)	ㄨ (w)u	ㄛ o	ㄟ ei	ㄣ en	ㄧ i	ㄨㄟ ui
ㄇ m	ㄋ n	ㄏ h	ㄒ x	ㄕ sh(i)	ㄙ s(i)	ㄩ (y)u	ㄜ e	ㄠ ao	ㄤ ang	ㄨ u	ㄨㄣ un
ㄈ f	ㄌ l			ㄖ r(i)			ㄝ e	ㄡ ou	ㄥ eng	ㄩ ü	ㄨㄥ ong

The Cangjie Input Method

Summary of Main Principles

The Cangjie method of computer input, which is also known as the Dragon method, is a fast, reliable alphabetic typing method for Chinese characters, based on their shape, without regard to their pronunciation. It was named by the inventor of the method, Chu Pang-fu, in Taiwan, after

the legendary inventor of writing in Chinese antiquity, Cangjie (in Japanese, pronounced Sôketsu).

We have created for WinCALIS and WinCALIS Author an extended, more user-friendly version of the original Cangjie method, which makes better use of the Windows graphical interface and mouse. You can select this form of character shape-based entry under the L:[Language] menu, by selecting the "Cangjie (by shape)" keyboard in the Chinese submenu, the "Kanji by shape" keyboard in the Japanese submenu, or the "Hanja by shape" keyboard in the Korean submenu.

When you select one of these keyboards, a dialog box appears. You can enter the letter-symbols representing component shapes either by typing them from the keyboard in the text box or by clicking on the piano-keyboard-style buttons, which show the letters and their corresponding shapes.

Here is a description of the essentials of the original Cangjie method:

The Cangjie method allows use of the ordinary computer keyboard to encode Chinese characters, by assigning 24 of the 26 letters of the Roman alphabet to represent the most common component elements of Chinese characters, grouped into 24 archetypal categories, each with a full-character label. Every Chinese character has a unique Cangjie code of from 1 to 5 letters. A space typed after a sequence of letters signals the computer system that the code is complete and initiates a lookup routine. (No space is required after a five-letter code.)

The letter "X" is a special symbol used in a small number of cases to disambiguate two characters which would have the same Cangjie input code. The letter "Z" is used in some versions of the Cangjie method to signal a user-defined character or phrase, or at the end of a Cangjie, as a disambiguator, like an "X" at the beginning. (WinCALIS uses both of these systems somewhat arbitrarily. Originally, "X" at the beginning was used uniformly, but now "Z" at the end is used instead for recent additions, since this fits the new user-friendly interface better.)

Mnemonic symbols

The 24 letter-symbols belong to four broad mnemonic categories:

- traditional Chinese natural philosophy (the seven heavenly luminaries and their correlates in the five "elements")

    (A) ("sun")   (B) ("moon")   (C) ("metal")   (D) ("wood")

        water")   (F) ("fire")   (G) ("earth")

- character stroke-types

    (H) ("left-downward sloping stroke")   (I) ("dot")   (J) ("cross")

        criss-cross")   (L) ("vertical")   (M) ("horizontal")

- the human body

    (O) ("man")   (P) ("heart/mind")   (Q) ("hand")   (R) ("mouth")

- typical character shapes

    (S) ("box open on the side")   (T) ("pair")

(U) ("box open on top")   (V) ("angle `V' shape")

     large box")   (Y) ("dot plus line")

As may be seen from some of the above examples, several of the letters also have a mnemonic association through their own shapes:

- "V" represents V-shaped components

- components looking like half a square "S" are represented by "S"

- U-shaped components are represented by "U"

- a small "l" is an example of a vertical stroke, represented by "L"

- a "dot" is suggested by the "dot" of the lower-case "i"

- the hook or angle of the "N" represents hook-shaped components

Order of Encoding

In entering the Cangjie code, the order of letters is similar to, but not completely identical with, the traditional stroke-order of Chinese characters. A character is built up left-to-right, top-to-bottom, and from outside to inside. When a character is so complex that it has more than the necessary number of components to uniquely identify it, the left-hand side and the first and last components of the right-hand side are usually encoded, with the middle of the right-hand side omitted.

In a typical character with a "radical" component on the left side, the radical is encoded with either one letter for the entire radical, or two letters, representing its first (top) and last (bottom) component parts. The remaining part of the character is encoded with the minimum number of letters to distinguish it from any other characters.

Chart Showing Cangjie Composition Principles

取碼規則： 取碼順序爲從上至下，從左至右，從外至內

整體字取碼：

商
YCBR

複合字取碼：
字首取碼，

字首

硫
MRYIU

字首

資
YOBUC

字首

建
NKLQ

字身取碼，
字身爲整體字者：

字首　字身

頭
MTMHC

字身可分爲次字首次字身者，
次字首一碼取足者：

字首　次字首

次字身

權
DTRG

次字首不能一碼取足者，

字首　次字首

次字身

轉
JJJII

次字首

次字身

緻
VFHJE

重複字：×爲代表

日	A	日	XA	皮	DHE	版	XDHE

難字取法，×代表過爲瑣屑部份

身	HXH	鳥	IXF	夃	LXH	龜	NXU
兼	TXC	慶	IXE	腍	IXP	卅	LXL
飛	RXU	臼	HX	廉	LX	齊	YX

只取首尾的常用字形

門	AN	目	BU	鬼	HI	几	HU
鬥	LN	臣	NL	隹	OG	气	OU
虎	VI	尸	YP	百	YR		

特殊字：木大火的特殊組合

東	DW	柬	DL	棗	DWF	棘	DB
末	DJ	棗	HDL	本	DM	夷	KN
爽	KKKK	券	FQSH				

Kangxi Radical Chart

Rad. No.	Kangxi Radical	Starts at Unicode	Rad. No.	Kangxi Radical	Starts at Unicode	Rad. No.	Kangxi Radical	Starts at Unicode
[1]	一	4E00	[2]	丨	4E28	[3]	丶	4E36
[4]	丿	4E3F	[5]	乙	4E59	[6]	亅	4E85
[7]	二	4E8C	[8]	亠	4EA0	[9]	人, 亻	4EBA
[10]	儿	513F	[11]	入	5165	[12]	八, 丷	516B
[13]	冂	5182	[14]	冖	5196	[15]	冫	51AB
[16]	几	51E0	[17]	凵	51F5	[18]	刀	5200
[19]	力	529B	[20]	勹	52F9	[21]	匕	5315
[22]	匚	531A	[23]	匸	5338	[24]	十	5341
[25]	卜	535C	[26]	卩	5369	[27]	厂	5382
[28]	厶	53B6	[29]	又	53C8	[30]	口	53E3
[31]	囗	56D7	[32]	土	571F	[33]	士	58EB
[34]	夂	5902	[35]	夊	590A	[36]	夕	5915
[37]	大	5927	[38]	女	5973	[39]	子, 孑, 孒	5B50

[40]	宀	5B80	[41]	寸	5BF8	[42]	小	5C0F
[43]	尢	5C22	[44]	尸	5C38	[45]	屮	5C6E
[46]	山	5C72	[47]	巛,川	5DDB	[48]	工	5DE5
[49]	己,已,巳	5DF1	[50]	巾	5DFE	[51]	干	5E72
[52]	幺,幺	5E7A	[53]	广	5E7F	[54]	廴	5EF4
[55]	廾	5EFE	[56]	弋	5F0B	[57]	弓	5F13
[58]	彐,彑	5F50	[59]	彡	5F61	[60]	彳	5F73
[61]	心,忄	5FC3	[62]	戈	6208	[63]	戶,户,戸	6236
[64]	手,扌,才	624B	[65]	支	652F	[66]	攴,攵	6534
[67]	文	6587	[68]	斗	6597	[69]	斤	65A4
[70]	方	65B9	[71]	无	65E0	[72]	日	65E5
[73]	曰	66F0	[74]	月	6708	[75]	木	6728
[76]	欠	6B20	[77]	止	6B62	[78]	歹	6B79
[79]	殳	6BB3	[80]	毋,母	6BCB	[81]	比	6BD4
[82]	毛	6BDB	[83]	氏	6C0F	[84]	气	6C14
[85]	水,氵	6C34	[86]	火,灬	706B	[87]	爪	722A
[88]	父	7236	[89]	爻	723B	[90]	爿,丬	723F
[91]	片	7247	[92]	牙	7529	[93]	牛	725B
[94]	犬,犭	72AC	[95]	玄	7384	[96]	玉,王	7389
[97]	瓜	74DC	[98]	瓦	74E6	[99]	甘	7518
[100]	生	751F	[101]	用,甩	7528	[102]	田,由,甲	7530
[103]	疋	758B	[104]	疒	7592	[105]	癶	7676
[106]	白	767D	[107]	皮	76AE	[108]	皿	76BF
[109]	目	76EE	[110]	矛	77DB	[111]	矢	77E2
[112]	石	77F3	[113]	示,礻	793A	[114]	禸	79B8
[115]	禾	79BE	[116]	穴	7A74	[117]	立	7ACB
[118]	竹	7AF9	[119]	米	7C73	[120]	糸,糹	7CF8
[121]	缶	7F36	[122]	网	7F51	[123]	羊	7F8A
[124]	羽	7FBD	[125]	老,考	8001	[126]	而	800C
[127]	耒	8012	[128]	耳	8033	[129]	聿,聿	807F

[130] 肉,月	8089	[131] 臣	81E3	[132] 自	81EA		
[133] 至	81F3	[134] 臼	81FC	[135] 舌	820C		
[136] 舛	821B	[137] 舟	821F	[138] 艮	826E		
[139] 色	8272	[140] 艸,艹	8278	[141] 虍	864D		
[142] 虫	866B	[143] 血	8840	[144] 行	884C		
[145] 衣,衤	8663	[146] 襾,西	897E	[147] 見,见	898B		
[148] 角	89D2	[149] 言,讠	8A00	[150] 谷	8C37		
[151] 豆	8C46	[152] 豕	8C55	[153] 豸	8C78		
[154] 貝,贝	8C9D	[155] 赤	8D64	[156] 走	8D70		
[157] 足	8DB3	[158] 身	8EAB	[159] 車,车	8ECA		
[160] 辛	8F9B	[161] 辰	8FB0	[162] 辶	8FB6		
[163] 邑,阝	9091	[164] 酉	9149	[165] 采	91C6		
[166] 里	91CC	[167] 金,钅	91D1	[168] 長,长	9577		
[169] 門,门	9580	[170] 阜,阝	961C	[171] 隶	96B6		
[172] 隹	96B9	[173] 雨	96E8	[174] 靑,青	9751		
[175] 非	975E	[176] 面	9762	[177] 革	9769		
[178] 韋,韦	97CB	[179] 韭	97ED	[180] 音	97F3		
[181] 頁,页	9801	[182] 風,风	98A8	[183] 飛,飞	98DB		
[184] 食,饣	98DF	[185] 首	9996	[186] 香	9999		
[187] 馬,马	99AC	[188] 骨,骨	9AA8	[189] 高	9AD8		
[190] 髟	9ADF	[191] 鬥	9B25	[192] 鬯	9B2F		
[193] 鬲	9B32	[194] 鬼	9B3C	[195] 魚,鱼	9B5A		
[196] 鳥,鸟	9CE6	[197] 鹵	9E75	[198] 鹿	9E7F		
[199] 麥,麦	9EA5	[200] 麻	9EBB	[201] 黃,黄	9EC3		
[202] 黍	9ECD	[203] 黑,黑	9ED1	[204] 黹	9EF9		
[205] 黽,黾	9EFD	[206] 鼎	9F0E	[207] 鼓	9F13		
[208] 鼠,鼠	9F20	[209] 鼻	9F3B	[210] 齊,齐	9F4A		
[211] 齒,齿	9F52	[212] 龍,龙	9F8D	[213] 龜,龟	9F9F		
[214] 龠	9FA0						

Japanese

Keyboards

Once Japanese is selected as a language in WinCALIS Author or Student Workstation, five keyboard choices appear in the pull-down menu list. They are: Roman-to-Kanji, Roman-to-Kana Only, JIS Hiragana Keyboard, JIS Katakana Keyboard and Kanji by shape. Following are descriptions of each keyboard.

Roman-to-Kanji

This keyboard, which provides the user with the most efficient and complete access to Japanese characters, will transform Roman characters into kana or kanji. WinCALIS and WinCALIS Author support the two most common systems for transliterating Japanese in Roman characters: 1) the Hepburn system, and 2) the *Kunrei* system. Some Japanese word-processing systems allow many non-standard ways of entering Japanese in romanized form (e.g., "ca" for "ka," "jya" for "ja/zya," etc.). Refer to the Kana Chart to see how each character may be accessed in WinCALIS. Keep in mind also the peculiarities of transliterating Japanese, detailed in this section.

Roman-to-Kana Only

This keyboard lets you type in kana characters without going through the buffer, so you need not press Enter for each *bunsetsu* (refer to the Kana Chart to see how it accepts transliteration). It will automatically change Roman characters into hiragana; katakana may be accessed by pressing Shift or CapsLock.

JIS Kana Keyboards

These keyboards transform the QWERTY keyboard into one that directly enters kana characters, according to the Japanese Industrial Standard. For a display of the corresponding keyboard while typing in WinCALIS or WinCALIS Author, click on Help, and choose Keyboard Help.

Hiragana keyboard

To type a voiced kana (marked with *dakuten*, two dots), first type the key for its unmarked unvoiced counterpart, then press the overstrike key, F2, then press the *dakuten* accent key, which is the key marked with double quote " (the shifted single quote ' key). See the chart for an example.

To type a kana beginning with the "p" sound (marked with *handakuten* circle), first type the key for its unmarked counterpart, then press the overstrike key, F2, then press the *handakuten* accent

key, which is the "o" key (the letter "o" suggests the little circle diacritic). See the chart for an example.

Kanji by shape

Kanji characters are composed of "root" shapes. Instead of retrieving a character by its pronunciation, this keyboard allows you to locate a character by its looks, by elements of its root structure. This is especially helpful in situations where you know what the character looks like, but not how it is pronounced. The root structures are accessed by choosing the Kanji-by-shape keyboard, and clicking on the piano-keyboard-style buttons. Notice that the database requires you to preface some duplicate characters with the "X" key. To clear the Kanji by Shape Window, simply change your Language keyboard.

For more details see the section "The Cangjie Input Method."

For information on inserting characters by browsing through the Unicode tables, see "Unicode

Peculiarities of Typing in Japanese

In Roman-to-Kanji keyboard,

to get:	*type:*
私	w-a-t-a-s-h-i
を	w-o
は	h-a
へ	h-e
いって	i-t-t-e
ほん	h-o-n-n
きょう	k-y-o-u
百	1-0-0 OR h-y-a-k-u
ABC	Ctrl + Shift + a b c
・(中黒)	/
. (bold period)	^

Particles

Particles must be entered by typing their literal romanization, that is, "w-o," "h-a," and "h-e." The program won't distinguish between the character when it is part of a word, or when it is a particle.

WinCALIS will include the "n" character within most words; the user may need to type an additional "n" within other words and always at the ends of words. So, WinCALIS will produce the proper hiragana if "k-o-n-b-a-n" is input, but notice you must type "h-o-n-n" to close the end of that word.

Long Vowels

In order to get the long "o" sound, the user must actually type in "u," e.g.: "k-y-o-u." For katakana, use the hyphen "-" for vowel extensions.

Typing in Katakana

To type in katakana, simply hold Shift or press CapsLock in the Roman-to-Kanji or Roman to Kana Only keyboard, or type in hiragana and then convert to katakana with the spacebar. As you may have noticed on the kana chart, there are a few unconventional combinations of characters. These will probably be used more for foreign words, e.g.: type in "f-o- - -m-u" (include the hyphen for the long vowel). To convert, press the space bar, then Enter to accept the katakana conversion to the English word "form," which is in the program dictionary (not in the internal dictionary). You must type in capitals to get proper nouns in katakana.

Numbers

WinCALIS will convert Arabic numbers to kanji. Simply type the digits while you are in the Roman-to-Kanji keyboard, press the spacebar to convert, then highlight and enter your choice of character(s).

The Active Buffer

There are important things to know about buffers when you are typing Japanese. If the characters are not black, you are in an active buffer and you must do something to clear the buffer. If the computer beeps at you, it means that you are trying to do something outside of the active buffer, and you must clear it (by pressing Enter) before you can continue.

When in the Roman-to-Kanji keyboard, this is the process for entering text:

- When you type hiragana, the letters will appear blue (active buffer). If you want the words to remain in hiragana, simply press Enter.

- If you want to convert the words into kanji, press the space bar, while the characters are still blue. Kanji will replace the words in blue and the sentence will appear red (active buffer). If all conversions are correct, press Enter (Note: The maximum length of a phrase to be converted is 32 characters. If you type the 32nd character, the phrase will automatically be converted).

- If any conversions are incorrect, do the following while the characters are still red. The reverse video block is your *bunsetsu*. Simply move the *bunsetsu* over the inappropriate

conversion and press the Spacebar. The homophone list will appear on the screen. Move the arrows to the desired kanji, then press Enter. The selected kanji will appear in the phrase in reverse video. If this is correct, press Enter again.

- WinCALIS divides your sentence into *bunsetsu* blocks as its dictionary sees fit, but you may change these blocks. First, convert back to hiragana by choosing it from the list (if you try to change the *bunsetsu* while there is still Kanji within the active buffer, you will get a pop up message reminding you to convert). Use the shifted left arrow to mark the first *bunsetsu*. At this point, you may convert that block by pressing the spacebar, or use the arrow keys to "jump" from block to block. To continue to change *bunsetsu* blocking, move to the block that you would like to adjust, and use the shifted right and left arrow keys to shrink or expand it.

- The user may find it helpful to clear the buffer (press Enter) early and often while working in the Roman-to-Kanji keyboard so the *bunsetsu* stay manageable. This will be convenient both for maintaining kana, and in locating the desired kanji. If the homophone list does not display the character you want, try changing the blocking of the *bunsetsu*. With use, you will become more familiar with how the transliteration process occurs within WinCALIS.

Look at the following passage, and notice the *bunsetsu* blocks—the bars indicate the points where conversions may be most efficiently placed (where the spacebar should be pressed):

「ユーザーフレンドリ」 | と | 呼ばれる |
理由は、ごく　　 | 普通の | 人がごく |
簡単に | 使用 | 方法を | マスター | でき、 |
毎日の | タイプ | の | 仕事にすぐ |
役立てることができることからきている。 |

This passage includes hiragana, katakana, kanji, and punctuation, so it is a nice exercise routine for typing Japanese. Notice how each bunsetsu is entered:

よばれる	—	呼ばれる
りゅうは、ごく	—	理由は、ごく
ふつうの	—	普通の
ひとがごく	—	人がごく
かんたんに	—	簡単に
しよう	—	使用
ほうほうを	—	方法を
まいにちの	—	毎日の
しごとにすぐ	—	仕事にすぐ
やくだてることができることからきている。	—	
役立てることができることからきている。		

The Homophone List

When you choose a kanji from the homophone list, the program will "remember" that one and present it next time the same kana are entered. The kanji most frequently used will be given priority over the other homophones (this does not rearrange the homophone list, which is constant. Rather, a separate user history file is being created and constantly updated).

Typing in Two Languages

Toggling between two languages can be done either by: clicking and dragging the mouse on the language choice from the options list; pressing F3 to toggle to the language at the top of the language list (usually English); or by clicking and dragging with the right-hand mouse button.

Punctuation in Japanese

Setting aside keyboard exploration time may make entering Japanese text easier. Because Japanese uses some punctuation that is exclusive to Japanese, the characters may be hidden from view on the standard English keyboard. Most of the punctuation parallels the basic keyboard, but it is in a *zenkaku*, full-width form. The keyboard layout is displayed by clicking on Help, Keyboard Help. See the chart for a sample of available marks.

Kana Kode Key

In order to obtain a kana character below, you must type in one set of Roman letter(s) listed beneath it. The first line of rômaji is the *Kunrei* system. Where applicable, the second line is the Hepburn system, and the third is nonstandard. Keep in mind the peculiarities of typing in Japanese, detailed in this section.

あ	い	う	え	お			
a	i	u	e	o			
あ	い	う	え	お			
xa	xi	xu	xe	xo			
la	li	lu	le	lo			
か	き	く	け	こ	きゃ	きゅ	きょ
ka	ki	ku	ke	ko	kya	kyu	kyo
			co				
さ	し	す	せ	そ	しゃ	しゅ しぇ しょ
sa	si	su	se	so	sya	syu she syo
	shi	ce			sha	shu she sho
	ci						
た	ち	つ	て	と	ちゃ	ちゅ ちぇ ちょ
ta	ti	tu	te	to	tya	tyu tyo
	chi	tsu			cha	chu che cho
な	に	ぬ	ね	の	にゃ	にゅ にょ
na	ni	nu	ne	no	nya	nyu nyo

は	ひ	ふ	へ	ほ	ひゃ	ひゅ	ひょ
ha	hi	hu fu	he	ho	hya	hyu	hyo

					ふぁ	ふぃ	ふぇ	ふぉ
					fa	fi	fe	fo

ま	み	む	め	も	みゃ	みゅ	みょ
ma	mi	mu	me	mo	mya	myu	myo

や	(い)	ゆ	(え)	よ	や	ゆ	よ
ya		yu		yo	xya lya	xyu lyu	xyo lyo

ら	り	る	れ	ろ	りゃ	りゅ	りょ
ra	ri	ru	re	ro	rya	ryu	ryo

わ	ゐ	(う)	ゑ	を	ゎ	ヰ	ヱ
wa	wi		we	wo	Wa	WI	WE

ん		う			っ	カ	ケ
nn n'		vu			xtu ltu xtsu	XKA LKA	XKE LKE

が	ぎ	ぐ	げ	ご	ぎゃ	ぎゅ	ぎょ
ga	gi	gu	ge	go	gya	gyu	gyo

ぐゎ	ぐぇ
gwa	gwe

ざ	じ	ず	ぜ	ぞ	じゃ	じゅ	じぇ	じょ
za	zi	zu	ze	zo	zya	zyu	zye	zyo
	ji				ja	ju	je	jo

だ	ぢ	づ	で	ど	ぢゃ	ぢぃ	ぢゅ	ぢぇ	ぢぉ
da	di	du	de	do	dya	dyi	dyu	dye	dyo
		dzu							

ば	び	ぶ	べ	ぼ	びゃ	びゅ	びょ
ba	bi	bu	be	bo	bya	byu	byo

ぱ	ぴ	ぷ	ぺ	ぽ	ぴゃ	ぴゅ	ぴょ
pa	pi	pu	pe	po	pya	pyu	pyo

ワ	ウィ	ウ	ウェ	ウォ	ヴァ	ヴィ	ヴ	ヴェ	ヴォ
wa	uxi	u	uxe	uxo	VA	VI	VU	VE	VO
	uli		ule	ulo					

Full-width Characters

Full width Roman characters are accessed by typing Ctrl + Shift + the character. This provides aesthetic appeal in situations where Roman characters are included within Japanese text.

Spelling Variation In Japanese

Note the potential for confusion on the student's part (in the WinCALIS language instruction system) because of spelling variation allowable in written Japanese. The careful courseware author will anticipate alternative spellings along the lines of those listed below in writing correct and incorrect responses.

Preferred form?   Alternate acceptable form 1   Alternate form 2 (questionable)

あさ	朝	
あした	明日	あす、みょうにち
後で	あとで	
現れる	表れる	
ある	有る	在る
いい	良い	善い、佳い、好い
行く	いく	ゆく
いたす	致す	
いる	居る	
いる	要る	
うち	内	
映る	写る	
おもしろい	面白い	
代える	替える	換える
かえる	帰る	返る

方　かた
彼　かれ
川　河
昨日　きのう　さくじつ
今日　きょう
今　ぐあい　具合　具合い
下さい　ください
位　ぐらい　くらい
くる　来る
今朝　けさ
御　ご、お
極　ごく　こんや　極く
今夜　こと
事　ころ、ごろ
頃

歳　才
次第　しだい
随分　ずいぶん
だれ　誰
付ける　つける
出来る　できる
と言う　という　とう
など　等
とき　時
ところ　所
無い　ない
何　なに
何か　なにか
成る　なる
爲る　なる

そのば	その場		
ばあい	場合		
はいる	入る		
初めて	始めて		
放す	離す		
早い	速い		
ひとつ	一つ		
ひとり	一人		
ほう	方		
他	ほか	た	
本当に	ほんとうに		
まいる	参る		
もうしわけ	申し訳		
もとづく	基づく		
もの	物		
よろこぶ	悦ぶ	喜ぶ	慶ぶ
わかる	解る	分かる	
わけ	訳		
わたし、わたくし	私		

Note: The path to the "history" file, which stores your most recent choice among homophone alternatives, "wcjpnhst.dic," can be redefined by adding the following line to the [WinCALIS] section of the main initialization file "wincalis.ini":

```
Jpnhistory=c:\wincalis\
```

Korean

How to Type Hangul

Select one of the Hangul keyboards in the "Korean" submenu of the "Language" menu. There are four different Hangul keyboards to choose from: the Phonetic keyboard, the Standard keyboard, the Kong keyboard, and the Morse keyboard.

The Hangul (Phonetic) Keyboard

The Phonetic keyboard is the typical student keyboard. The consonants and the vowels are arranged to reflect the similar sounds of the Roman alphabet. For example, the consonant ㅍ is located on the "p" key, and the vowel ㅏ is located on the "a" key. Thus, it may be the easiest keyboard for a native English speaker to use. However, some Korean letters do not have an equivalent in English, and thus are placed on unused keys on the keyboard. For example, ㅇ is located on the "q" key with no relevance to sound (but a mnemonic association with the shape of the letter "Q"). Likewise, ㅐ looks a little like a "Z" on its side, so it is on the "z" key; and ㅡ , which is romanized ŭ (a "double u") is on the "w" key. The user must become familiar with these "exception" letters.

Phonetic keyboard (unshifted)

The Hangul (Standard) Keyboard

The Standard keyboard, as suggested by its name, is the standard keyboard in Korea (the Korean typewriter keyboard). Thus, it is probably the keyboard that a native Korean would be familiar

with. It is divided into two parts: The vowels are located on the right half of the keyboard, and the consonants are located on the left half of the keyboard.

Standard keyboard (unshifted)

The Hangul (Kong) Keyboard

The Kong keyboard is separated into three parts: The initial letters (consonants, such as ㄱ , ㄷ , etc.) are placed on the right third of the keyboard. The medial letters (vowels such as ㅏ , ㅐ , etc.) are placed on the middle third of the keyboard. The final letters (consonants) are on the left third of the keyboard. The user must be aware of these differences. For example, although the initial and the final consonants look the same (ㄴ , an initial on the "h" key, and ㄴ , a final on the "s" key), they cannot be used interchangeably.

Kong keyboard (unshifted)

The Hangul (Morse) Keyboard

The Morse keyboard, among the four, requires the greatest number of keystrokes to type. There are no compound vowels, only simple. Thus, the user must combine the simple vowels to make compound vowels. For example, ㅏ + ㅣ = ㅐ . In addition, the vowels and consonants are scattered about the keyboard in relation to frequency of use and convenience of layout of the particular letter (similar to the arrangement of the English keyboard).

Morse keyboard (unshifted)

Hangul-Hanja Conversion

Hangul-to-Hanja (F8)

To convert the syllable or word to the left of the insertion point (the blinking caret) written in Hangul into its Hanja equivalent, press the F8 function key, or click on the menu item Tools, Language Options, Korean Options, Hangul-to-Hanja. (One of the Korean language keyboards must be the currently selected keyboard.) The conversion engine scans to the left from the insertion point attempting to match up to four Hangul characters. If there are homophone equivalents, they will appear in a Selection Window. Choose the desired equivalent.

For example, to convert 도서관 to 圖書館 , press F8 with the insertion point to the right of 도서관 .

Hanja-to-Hangul (F7)

To convert the syllable or word to the left of the insertion point (the blinking caret) written in Hanja into its Hangul equivalent, press the F7 function key, or click on the menu item Tools, Language Options, Korean Options, Hanja-to-Hangul. (One of the Korean language keyboards must be the currently selected keyboard.) The conversion engine scans to the left from the insertion point attempting to match up to four Hanja characters. If there is more than one Hangul equivalent (i.e., there is more than one reading for the Hanja), they will appear in a Selection Window. Choose the desired equivalent.

For example, to convert 韓國 to 한국 , press F7 with the insertion point to the right of 韓國 .

You can also enter Korean Hanja characters through two other methods:

Shape-based coding: The Hanja By Shape Keyboard

Under the Language or L:[Language] menu, select the "Hanja by shape" keyboard in the Korean submenu. For details on this method see "The Cangjie Input Method". Cangjie, the legendary inventor of writing in ancient China, is known in Korean as Ch'anggyôl. To clear the Cangjie Window, simply change your Language keyboard.

Hanja by shape keyboard
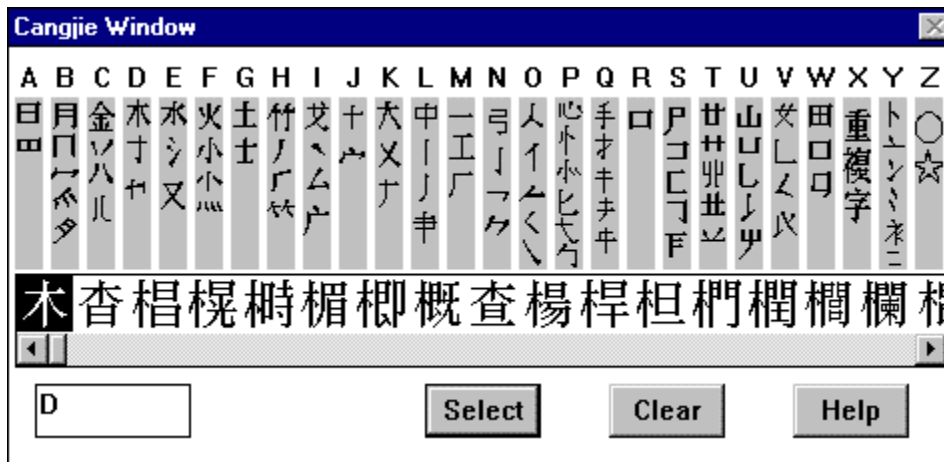
HOW TO TYPE KOREAN TRANSLITERATION

The Hangul (Morse) Keyboard

In order to type the special o with breve ($\breve{o}$) and u with breve ($\breve{u}$) characters necessary for the Korean Ministry of Education or McCune-Reischauer romanization, you may choose the "Korean Transliteration" keyboard in the Korean submenu, as well as one of the other general transliteration keyboards. You can type the base vowel "o" or "u," then press the overstrike key F2, then type the mnemonic accent key for the breve "u"; or you can type them directly by pressing "q" or "f" for " $\breve{o}$ " and "x" or "v" for " $\breve{u}$ "

Korean Transliteration keyboard
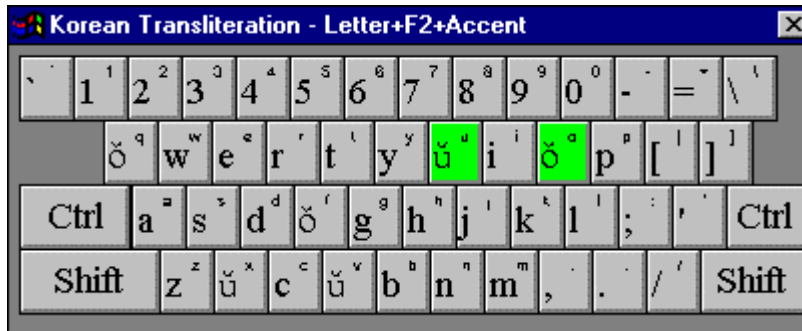
Chapter 6: On-Line Dictionaries

User-Supplied On-line Dictionaries

Tools Menu Option "Look Up Word...   F5"

You can type or highlight a word or phrase on the screen and look it up in an on-line dictionary, if one has been installed on your workstation.

If one or more on-line dictionaries have been supplied to accompany your copy of UniEdit, selecting this menu option will open a dialog box, allowing you to look up the meaning or pronunciation of a selected word or phrase. Depending on how the dictionary was installed, it can be either accessible only from specific language keyboards (e.g, for a Korean-English dictionary, the current keyboard must be one of the Korean keyboards), or accessible from any language keyboard which does not have a specific dictionary linked to it (e.g, for a Korean-English dictionary, the current keyboard could be either the default English/W. Europe keyboard or any of the Korean keyboards).

Full-scale bilingual dictionaries are also available or under development for various languages. See: Japanese-and-English On-line Dictionary and Korean-and-English On-line Dictionary.

How to Look Up a Word

In order to select a word or phrase to look up, drag the mouse across it (or double click on it if it is a single word) to highlight it. Then press F5 or click on the menu item Tools, Look Up Word... to look up the word. Repeat as often as desired.

You can press the Escape key at any time during a search to cancel further searching and display the matches found thus far.

The On-line Dictionary Dialog Box

Search In:
    Main Entries
    Everywhere

Source:
    Highlighted Text
    Clipboard

List Box Window

Highlighted string for
Look Up

Match:
    Any Part of An Entry
    Start of Entry Fields
    Whole Words
    Whole Entry Fields

Search Again

Save Settings

Close

Help

Match Case
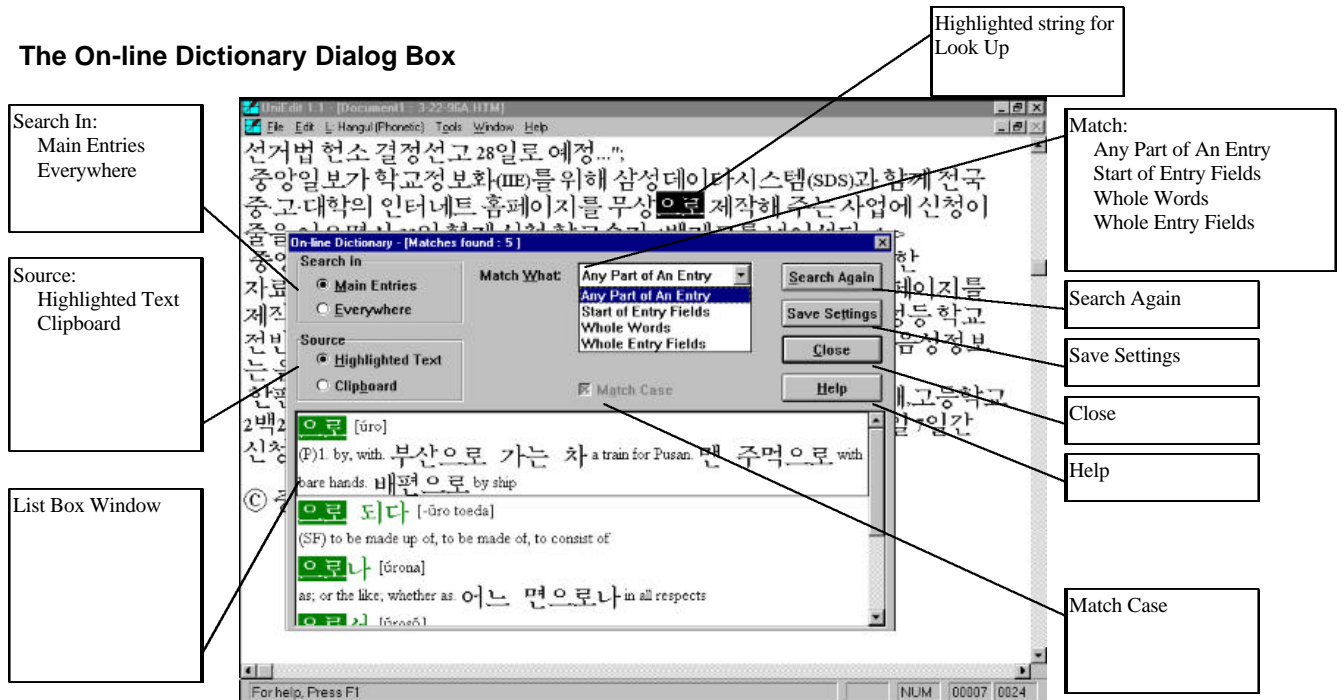
The On-line Dictionary Dialog Box contains various option settings in its upper part and a list box displaying the items found matching the word looked up in its lower part. Options are explained below. The words in the dictionary matching the word looked up are highlighted in reverse video. The total number of matches found is displayed in the title bar of the dialog box.

Options in the On-line Dictionary Dialog Box

Note: In order to understand the following options, note that an "entry" refers to an entire dictionary entry, which is enclosed in an outline box within the dictionary list box. You can scroll down through the displayed entries using the mouse and the scroll bar or using the PageDown, PageUp, and arrow keys. The term "field" is used to refer to the separate sections of a whole entry:

- the main entry field, highlighted in dark green, left justified on the first line of an entry

- the pronunciation or gender tag field, enclosed in square brackets, to the right on the first line

- the definition field(s), beginning on the second line of an entry

Escape Key

You can press the Escape key at any time during a search to cancel further searching and display the matches found thus far. However, since the Escape key has a special meaning in WinCALIS (quit an exercise), Escape can not be used in WinCALIS for the first search, only from subsequent searches when the On-line dictionary dialog is already displayed.

Radio Button and Checkbox Options

Search in:

- <u>M</u>ain Entries

Displays only the matches found in the main entry portion (field) of a full entry. The main entry field is displayed in dark green at the left margin of the first line of a full entry. Main entries are what the user of a dictionary most commonly searches for. This is the system startup default option.

- <u>E</u>verywhere

Displays matches found in all fields of a full entry (displayed in dark green or in black).

Source

- <u>H</u>ighlighted Text

Uses the highlighted text as the word or phrase to look up. This is the system startup default option.

- Clip<u>b</u>oard

Uses the Unicode text currently on the Windows clipboard as the word or phrase to look up. This feature is primarily designed as a way to use the WinCALIS on-line dictionary system with other Unicode-compliant applications. For looking up words from non-Unicode applications, use the "Paste As..." option in the Edit Menu of the WinCALIS Composition Editor, WinCALIS Author, or UniEdit and select the appropriate coding standard to convert from. Then highlight the word or phrase to look up in the normal way, using the default "Highlighted Text" as Source.

Match Case

Will match lower case or upper case by default. If unchecked, will ignore case in finding matches. Since the WinCALIS system understands case relations in all the world's alphabets by consulting a table, case-sensitive searches are considerably faster than searches which ignore case.

Drop-down Listbox Options

Match <u>W</u>hat:

- Any Part of An Entry

The search will match any part of an entry, including phrases, words, and sequences of as few as three letters occurring anywhere in a word. This is the broadest search scope. This is useful when looking up words which are not likely to occur at the beginning of an entry, such as suffixes like English "-ment" or "-tion," or when using the dictionary "in reverse," to look up words which occur in definition fields. For example, in a Spanish-English dictionary, one would use this option along with "Search In: Everywhere" to look up the English word "house" in order to find its Spanish equivalent. It will find all cases where the word "house" occurs.

When looking up *hanja* Chinese characters in the Korean-English dictionary, this option should be used along with either the "Search in: Main Entries" option (will match only *hanja* main entries) or the "Search in: Everywhere" option (will also find any *hanja* occurring within the definition fields).

- Start of Entry Fields

The search will match the start of any entry field specified. For example, if "Main Entries" only are searched in (the default option), then "Start of Entry Fields" will match only the first letters of the Main Entry. This is the next broadest search scope and is the system startup default option. If all entry fields are searched in (the option "Everywhere"), then the "Start of Entry Fields" option will match the first letters of any field.

This is the best mode for handling language inflection, i.e., the variation of endings for verbs and nouns--highlight only the stem portion of a verb or noun to look up. This is also the typical search mode we use when looking up words in print dictionaries and glossaries. In other words, when we look up the English word "pruning," we scan with our eyes the main entries in a print dictionary for "prun...," until we find the appropriate match "prune," having regularized or "lemmatized" in our mind the "ing" form of "prune." The computer will not perform this lemmatization automatically, so that it will not find the match "prune," if the search string is "pruning." But it will find "prune" if the search string is "prun."

.

- Whole Words

The search will match any "whole word," which is defined as a sequence of letters bounded by spaces or punctuation, including the hyphen. This is the second narrowest search scope. Specific language dictionaries may also apply special principles. For example, in the WinCALIS Korean-English dictionary, bound forms have been tagged with an invisible Unicode Non-Spacing Joiner, which behaves like a regular character, so that bound forms are ignored during a "whole word" search.

- Whole Entry Fields

The search will succeed only if the word or phrase looked up matches an entire entry field precisely. This is the narrowest search scope. Matching whole entry fields has limited usefulness, but can be helpful, for example, in matching full main entry or pronunciation fields.

Button Options

<u>S</u>earch Again

Reinitiates a search using the current search options. It is often advisable during a lengthy search for too broad a search string to press the Escape key and narrow the search somewhat. For example the on-line dictionary search engine would find over 9000 matches for the search item "giv" in the Japanese-English dictionary, since "giv" is the marker for a Japanese given name.

Save Se<u>t</u>tings

Retains the current search options for subsequent searches during the current WinCALIS session.

<u>C</u>lose

Closes the dialog box. You can also close the dialog box by pressing the Escape key.

<u>H</u>elp

Calls up On-line Help.

Japanese-and-English Dictionary

Key to Abbreviations

an	adjectival nouns or quasi-adjectives (keiyodoshi)
a-no	nouns which may take the genitive case particle "no"
vs	noun or participle which takes the aux. verb suru
vt	transitive verb
vi	intransitive verb
id	idiomatic expression
col	colloquialism
vul	vulgar expression or word
pn	person name (family or given)
pl	place name
giv	given name
fam	familiar language
pol	polite (teineigo) language
hum	humble (kenjougo) language
hon	honorific or respectful (sonkeigo) language
pref	prefix
suf	suffix
uk	word usually written using kana alone
uK	word usually written using kanji alone

oK	word containing out-dated kanji
io	irregular okurigana usage
arch	archaism
X	rude or vulgar term

The WinCALIS Japanese<-->English on-line dictionary has been adapted to the WinCALIS on-line dictionary system and converted to <u>Unicode</u>/ISO 10646 from the "Jim Breen Japanese/English Dictionary." It is used with the permission of the copyright owner, Prof. Jim Breen of Monash University.

We are indebted to Prof. Breen not only for his permission to adapt and distribute his Japanese/English dictionary, but also because the simple but elegant architecture of his dictionary provided a model for our on-line dictionary and automatic indexing system in general.

Korean-and-English Dictionary

Key to Abbreviations

(ABBR) Abbreviation

(AD)   Adverb

(ADJ)   Adjective

(BF)   Bound form, bound noun

Ch.   Chinese

(CJ)   Conjunction

Eng.   English

(HS)   ROK Ministry of Education "High School" *hanja* (*hanja* learned by Korean high school students)

(IN)   Interjection (including onomatopoeia)

(MS)   ROK Ministry of Education "Middle School" *hanja* (*hanja* learned by Korean middle school students)

(N)   Noun

(NC)   Numeral classifier (numerical counter, measure)

(Obs.) Obsolete or archaic form

(PA)   Particle, auxiliary word

(PF)   Prefix

(PN)   Proper noun (personal name, place name)

(PR)   Pronoun

(SF)   Suffix

(VE)   Verb

(VC)   Verb causative

(VG)   Verb gerund

(VI)   Verb intransitive

(VN)   Verbal noun

(VT)   Verb transitive

Key to Symbols

⇨   An arrow indicates a cross-reference, synonym, or other relevant usage.

【】   Solid lenticular brackets enclose *hanja* Chinese character information in an "Okp'yon"
               character-dictionary-style field. See the item "Hanja *Okp'yon* Field" below for more
               information.

〖〗   Hollow lenticular brackets enclose the variant inflected stem forms of common verbs.

〔〕   Tortoise shell brackets enclose *hanja* Chinese character equivalents for Sino-Korean
               noun and bound morpheme (BF) entries.

<   A less-than sign means "(etymologically) derived from." Etymologies have been
               included for many recent loan words of European origin.


( )   Parentheses enclose usage tags, meaning clarifications, and specialized terminology.
               E.g., "(Music)" indicates a term used especially in musical terminology; "(Medicine)"

indicates a term used especially in medical terminology; and "(Obs.)" indicates an obsolete or archaic form.

General Features

Layout

A sample dictionary entry is pictured below, with labels for each field. See below for further details on each field and on other dictionary options.
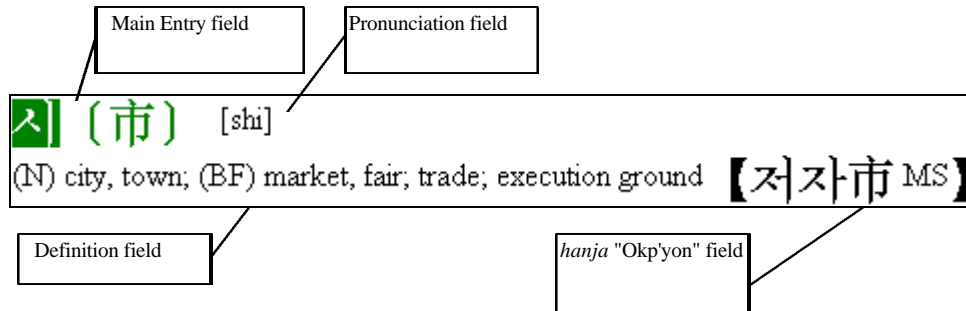
Main entry field

Main entries are highlighted in dark green. For verbs the main entry includes the dictionary citation form, as well as variant inflected forms for the most common and short verbs, enclosed in hollow lenticular brackets. The purpose of this is both for reference and to make it easier to locate a verb by highlighting and looking up its stem element. For Sino-Korean nouns, the *hanja* Chinese character equivalent is given, enclosed in tortoise shell brackets.

Pronunciation field

The second field of each entry, after the main entry on the same line, is the pronunciation field, enclosed in square brackets. Pronunciation is given according to the McCune-Reischauer pronunciation.

Alternate pronunciations are provided within the definition field, preceded by "Also":

Note example entry:
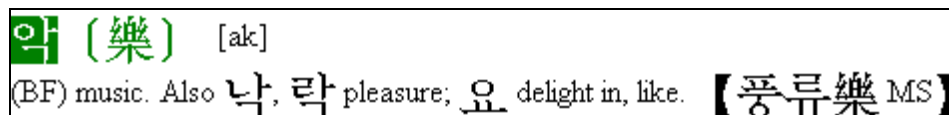
Definition field

Definitions in English, derived forms (e.g., adjectives or verbs formed from nouns), examples with English translations, and usage information are all presented in the definition field,

beginning on the line following the main entry field. Distinct definitions are grouped under separate Arabic numerals, in accord with common Korean lexicographic practice. For most verbal nouns ending with *hada*, the noun is given as the main entry, and the *hada* form is given within the definition field.

Hanja "Okp'yon" *field*

The entries for the most common morphemes of Chinese origin end with a *hanja* "Okp'yon" Chinese character dictionary-style field, enclosed in solid lenticular brackets. This usually contains first a Korean "tag meaning" for the morpheme (*hanja* character). Second is the *hanja* character itself. Third is an approximate indication of frequency of usage, if it is one of the 1800 characters designated as the most important for educational purposes by the Ministry of Education of the Republic of Korea. "MS" indicates that it is one of the most common 900 characters which are taught in middle school in the Republic of Korea. HS indicates the next most common 900 characters, which are taught in high school in the Republic of Korea.

Serial Arrangement

Entries are arranged in ascending *hangul* alphabetical order, in the order of the Unicode/ISO 10646 encoding standard. This order is in turn based on the Korean Standard KS C 5601-1989.

For more information on Unicode, consult The Unicode Consortium, *The Unicode Standard: Worldwide Character Encoding*. Version 1.0, Vol. 2 (Reading , Mass: Addison-Wesley, 1992); for more information on ISO 10646, consult the document *ISO/IEC 10646-1:1993 Part 1: Architecture and Basic Multilingual Plane* (available from ANSI, 11 West 42nd St., New York, NY 10036).

Note: In 1996 the Unicode/ISO 10646 code points for all Korean *hangul* shifted from their location in the code range 3400-4D00 under the existing Unicode 1.1 standard to their new location in the code range AC00-D7A3 in the proposed revised standard Unicode 2.0/ISO10646-?:199?. In order to both accommodate the existing standard and anticipate the future standard within the WinCALIS Korean-English dictionary, the dictionary lookup engine has been designed to handle *either* range of code points. The dictionary is distributed in two different versions, one for Unicode 1.1 and one for Unicode 2.0. Either can be installed on the user's system, or both can be installed simultaneously, if they are linked to one or more separate language keyboards under the Language menu--for example the existing "Hangul (Phonetic)" keyboard can be linked to the dictionary using the Unicode 1.1 standard, while a new "Hangul (Phonetic) Alt." keyboard can be linked to the dictionary using the Unicode 2.0 standard. In order to access a specific dictionary, then, the user would need to switch keyboards to the desired one before looking up a word. In theory, these two dictionaries could be combined and reindexed together, so as to allow looking up Korean *hangul* encoded in *either* Unicode 1.1 or Unicode 2.0, but this not only would produce a dictionary double the current size, but also would have the undesirable effect of displaying duplicate entries for all *hanja* and English words looked up. It would also be more difficult to maintain and extend.

Grammar

The system of grammar employed is generally that of Fred Lukoff, as found in his textbook, *An Introductory Course in Korean* (2nd edition, Seoul:Yonsei University Press, 1993). The monumental dictionary *A Korean-English Dictionary* by Samuel E. Martin et al. (New Haven: Yale University Press, 1967) was also extensively consulted, as well as most of the common Korean-English dictionaries published in Korea for learners of English.

Scope

The dictionary contains about 22,000 main lexical entries, with several thousand more derived forms given in the body of the entries. Pains have been taken to ensure that virtually all the vocabulary items presented in the Korean textbooks commonly used in the United States have been incorporated here. Other special areas of coverage which have been incorporated include:

• over four hundred lexical items which are unique to North Korean usage, with equivalent South Korean usage added as a cross-reference; indicated with the usage tag "(N. Korea)"

• common geographic names, both Korean and foreign

• common Korean surnames (marked "surname")

• common Chinese surnames (marked "Ch. surname")

Entries have been provided for all of the 4888 *hanja* Chinese characters from the Korean Standard KS C 5601-1989, plus others not included in the KS standard which are used to write common vocabulary. Each entry gives the *hangul* reading and McCune-Reischauer transcription. Many of these *hanja* have obscure meanings or no meaning at all when used independently, so in order to avoid distorting their function, they have been simply marked in the definition field with the phrase "[rare combining form]." However, it is hoped that the pronunciation provided will be of value in identifying and reading them when they occur in the names of places and persons.

Acknowledgments:

We are very grateful to the following teachers and scholars for their advice in the design and compilation of this Korean-English dictionary: Mrs. Yang S. Camus, Prof. Kay Kang, Mrs. Young Ah Lee (Chief Editor), Mr. Bo Park, Dr. Bruce Phinney, Prof. Han Sohn, Prof. Janna Son-Yarborough, and Mr. Sukmoon Yoon.

How to Create Your Own On-line Dictionary

The item "Look Up Word..." in the Tools menu is a powerful new feature in WinCALIS. Depending on what is installed on a particular workstation, WinCALIS, WinCALIS Author, or UniEdit (the WinCALIS Multilingual Text Editor) can access an unlimited number of Unicode-text on-line dictionaries. These dictionaries can be either multilingual (e.g., Russian-Arabic, Arabic-Russian), or monolingual (e.g., a dictionary of specialized vocabulary for advanced students of English). Any author can generate a dictionary by creating a (Unicode) text file in WinCALIS Author or UniEdit in a fixed format, like:

EntryField/Pron.orGenderField/UsageField/DefinitionField1....

There should be one entry per line, separated by hard returns, and with fields separated by slashes (/). Note that because the slash has a special meaning, it is a reserved character and should not be used within fields. The EntryField will start a new entry and will be highlighted in dark green color. The second field ("Pron.orGenderField") will be placed on the same line surrounded by square brackets. If the second field is empty—that is, there are two successive slashes (//), the brackets will be replaced by a colon (:). Succeeding fields will be placed on new lines. The maximum length for an entire entry is practically speaking about 512 characters, since that is about how much can be viewed in the on-line dictionary list box at one time.

When you have created a small pilot dictionary file in a text window, save the file to disk in the ..\WINCALIS\DICTS subdirectory (or whatever path you have specified in the "Dictionaries=" line of WINCALIS.INI), with the file extender ".DIC", specifying the file type "Unicode/No Header" in the "Save File as Type:" box.

Then you need to generate the index to your dictionary. In WinCALIS Author or UniEdit, in the Tools menu, click on "Generate Index" and specify the name of the dictionary to generate the index for. Usually you should specify the same 1-8 letter filename for both the dictionary and the index to be generated. E.g., if your dictionary file is named "wcrusfrn.dic," then the index should be named "wcrusfrn.idx."

At this time you can also select a language keyboard to associate the new dictionary with, if you wish to use more than one language dictionary. When that keyboard is the active keyboard in WinCALIS, WinCALIS Author, or UniEdit, this dictionary will be the one which is used. WinCALIS Author or UniEdit will create or update a [Dictionary] section in the corresponding language-keyboard .INI file on your workstation. Anytime after you restart WinCALIS and WinCALIS Author, or UniEdit, if that language-keyboard is the current keyboard, the programs will look for your dictionary and index file in the ..\WINCALIS\DICTS directory or other directory specified in the "Dictionaries=" line of WINCALIS.INI.

Note: In order to associate your dictionary with more than one keyboard, you may use the Windows Notepad to open the .INI file of the keyboard you first selected (check the time/date stamp on the file in the File Manager to determine which one it is, if you can't spot it), then copy the entire [Dictionary] section to the Clipboard and paste it into as many other .INI files as you wish. (See LATIN1.INI for a sample.) See also Language Keyboard .INI Files.

When you press the "OK" button in the Generate Index dialog box, Author or UniEdit will begin to create the index file. For a small pilot dictionary of just a few items, this process takes only a few seconds or minutes. But for a real glossary or dictionary of thousands of entries, and with

many items to index, expect the index generation process to take 1-3 hours, even on a fast machine. It generally works well to start the index generator at the end of a day and let it run during the night. Also, be sure that there is plenty of free space on the hard drive for the several hundred temporary files which are created during the indexing process--at least 2-3 times the total size of the dictionary file itself.

Every alphabetic sequence of 3 letters will be fully indexed, as will every single Han character for Chinese-Japanese-Korean-Vietnamese. Note that at the moment there is no way to index 2-letter words such as English "to" or "by." There is also not yet a "stop-word" function, which will exclude certain 3-letter combinations, like "and," from being indexed.

When index generation is completed, the dictionary and index files are finished and ready to use. If you wish students to have access to your dictionary, you will need to arrange to install the dictionary files on the student workstation(s) and link them to keyboards as needed. If you would like to contribute your dictionary as a general enhancement to the WinCALIS system, we can pre-install your dictionary in student workstations.

Note that the default dictionary and index names are WINCALIS.DIC and WINCALIS.IDX. A dictionary with that name will be found regardless of the current keyboard, unless another dictionary is specified for the current keyboard.

Users who are developing large dictionaries should review editing limits imposed by the Author/UniEdit editor. In particular, note that there is a limit of 6000 lines of text in each window in WinCALIS Author and UniEdit. There can be up to six concurrent 6000-line texts, one in each window. You will need to develop your dictionary in component parts of ca. 5000 lines each. Then you can use the DOS COPY command or a DOS batch file to concatenate each of the component files into one big file (hence the "Unicode/No-Header" file type). There is a sample batch file, UPDATE.BAT, included in the ..\WINCALIS\DICTS directory during setup.

The dictionary file does not need to be in any particular order, and sorting it (with the "Sort Lines" item under the "Tools" menu) is useful mostly to avoid duplicate entries and so that display of items in the dictionary is in ascending order, when the word looked up is not a main entry. You can sort multiple component files in the same way you might sort a stack of cards to be alphabetized by first dividing it into piles by initial letter, then proceeding to alphabetize each of the piles, then recombining all the piles in order. You would proceed by sorting each of your files, then chopping each of them into eighth-length files (or quarters or halves or 1/20ths, depending on the total length of your dictionary--you will need as many divisions as you have component files) at logical divisions (e.g, in the Latin alphabet, a-c, d-f, g-i, j-l, m-o, p-r, s-u, v-z, then combining all the first-eighth files and resorting them; then all the second-eighth files and resorting; and so on. When all the sub-files are reassembled and sorted, they can be reassembled. The entire dictionary should now be in sorted Unicode order. A sample DOS batch file to do the concatenating operation mechanically is included in the ..\WINCALIS\DICTS subdirectory. Be sure that you keep master copies of all component sub-files--remember that full dictionary files exceeding 6000 lines are not directly editable or even viewable in the WinCALIS editor. If you accidentally combine all your component files and fail to keep the components, you

can, in a pinch, use a file utility such as "SPLIT49.EXE" to split the big file again at even byte boundaries into any number of components. SPLIT49.EXE and its documentation, SPLIT49.DOC can be downloaded from the Humanities Computing Facility website, http://www.lang.duke.edu.

We are in the process of developing large bilingual dictionaries for Korean-and-English and Japanese-and-English. We hope that with the collaboration of interested WinCALIS authors and users, we will be able to distribute useful dictionaries to accompany WinCALIS for a wide range of languages.

Tools Menu Option "Generate Index"

This menu option in WinCALIS Author or UniEdit runs the automatic Unicode index builder to build an index for a dictionary file in the default "Dictionaries" subdirectory. The dictionary should be saved as type "Unicode/No Header".

Chapter 7: Configuring and Customizing UniEdit

These are various topics related to the general subject of customizing your UniEdit Workstation so it serves you the best, maximizing your productivity.

To learn how to set the font size for use in your UniEdit, set default directories, change the virtual keyboard picture, change the Language Menu, and much more, see "The Main UniEdit Initialization File WINCALIS.INI."

To learn how to change or fine-tune the default font(s) for use in your UniEdit, see "The Default

To learn about how to make your own language keyboard or modify an existing one, see "Defining Language Keyboard Files."

To learn about how the Unicode-smart UniEdit system *globally* defines punctuation, capitalization, and accented characters, see "The UniEdit Global Information File

To learn how to add an on-line dictionary, see "User-Supplied Dictionaries."

To learn more about UniEdit fonts and typefaces, see "Fonts."

To find out about setting up a Spelling Checker, see "Linking to a Spelling Checker."

To find out which UniEdit files are essential and which you can get rid of without worry, see "Deleting Nonessential Files to Save on Disk Space."

For tips on network installation and operation of UniEdit, see "Installing UniEdit on a Network."

To find out how to remove the UniEdit Workstation from your system, see "How to Uninstall

The Main UniEdit Initialization File WINCALIS.INI

General Tips

The main UniEdit initialization file WINCALIS.INI, which contains all kinds of customizable settings, is to be found in the main UniEdit program directory (usually the parent directory of the default scripts directory).—e.g., C:\WINCALIS. It is created dynamically at the time of installation. It is a typical ASCII Windows *.INI file. It can be edited with a plain-text editor like the Windows Notepad. Be sure to use the Windows File Manager or other tool to make a backup copy of WINCALIS.INI, before you make any changes in it.

If you are currently running UniEdit, you will have to exit and restart it for any new settings to take effect.

If you wish to make only a temporary change to one or another setting in your WINCALIS.INI file, you can leave the Notepad running in the background, with the WINCALIS.INI file still open in it. Then after you have finished using UniEdit, you can return to the Notepad and change the setting back again.

To change the default font for use in UniEdit, you will need to edit another UniEdit system file, The Default Character-to-Font Mapping File WCFNT.INF.

Changing the Default Font Size for Use in UniEdit

(For how to change the font or typeface itself, see The Default Character-to-Font Mapping File WCFNT.INF.)

UniEdit always displays text in "Medium" font size. But you can define Medium as any pixel height you choose, provided the font is available. Here's how to do it. UniEdit will still use "Medium" for any font it needs. But you can redefine the value of "Medium" to any pixel height.

The actual point size of the "Medium" used by UniEdit is determined by the pixel height settings in the default language .INI file defined in the "DefaultIni=" line in the [WinCALIS] section of WINCALIS.INI. The default language .INI file is set for the United States and Western Europe to "LATIN1.INI." Changing the line "Medium=21" in the [FontSize] section of "LATIN1.INI" to some greater or smaller pixel height than 21 pixels will adjust the size of the default font accordingly in UniEdit. All pixel heights are set by default for standard VGA resolution (640x480 pixels) as follows:

Tiny=12
Fine=16
Small=19
Medium=21
Large=32
Xlarge=48
Huge=66.

So to change the font size for ANSI (all West European languages), edit the LATIN1.INI file. (The procedure is identical for other character sets—Cyrillic, Greek, Hebrew, Thai, etc.) Redefine the following line to any desired pixel height.

Medium=19

Then exit UniEdit and restart UniEdit.

You can leave the LATIN1.INI file running in Notepad in the background. Then when you no longer need a smaller or larger font size, restore the original settings and save the file again.

Special versions of "latin1.ini" for use with EGA monitors and with VGA monitors at higher than standard VGA resolution are distributed as part of the UniEdit and Student Workstation packages.

The pixel heights for the fonts in many other character sets can be adjusted in the same way, by specifying the pixel heights in the language keyboard *.INI file you will first use to type that character set, generating that font on your screen.

Setting the Virtual Keyboard Type

The virtual keyboard picture displayed when Keyboard Help is selected is defined in the WINCALIS.INI file in the main UniEdit subdirectory. It can be redefined by carefully editing the line in the [WinCALIS] section

KeyboardType=0

Keyboard arrangements are perhaps best described by the position of the backslash (reverse solidus) key (at least for U.S. keyboards). Substitute a number corresponding to one of the following available pictures:

- Type 0: Backslash key (\) on top row, next to Backspace key (default)

- Type 1: Backslash key (\) on 2nd row from top, below Backspace key

- Type 2: Backslash key (\) on bottom row at extreme right, below Enter key

- Type 3: Backslash key (\) on bottom row at extreme left

Changing the UniEdit Desktop

In the Desktop submenu under the Window menu you can "Save" the current desktop arrangement, "Restore" the previous arrangement, or go back to the "Default" arrangement, as it was set in WINCALIS.INI during installation of the UniEdit Workstation. Here's how each of these sections looked in your WINCALIS.INI out of the box:

```
[Initial]
Word=Yes
Tone=Yes
Def=No
NoDisk=Yes
Simple=Yes
NoConvt=No
NoCache=No
Net=Yes
InitialLanguage=English/W.Europe
Document1=MAX
Document2=ICON
Document3=ICON
Document4=ICON
Document5=ICON
Document6=ICON
Active=Document1
Wordwrap=Yes
Statusbar=Yes

[Default]
Word=Yes
Tone=Yes
Def=No
NoDisk=Yes
Simple=Yes
NoConvt=No
NoCache=No
Net=Yes
InitialLanguage=English/W.Europe
Document1=MAX
Document2=ICON
```

Document3=ICON
Document4=ICON
Document5=ICON
Document6=ICON
Active=Document1
Wordwrap=Yes
Statusbar=Yes

The first eight lines in each section refer to Chinese internal dictionary search options, which are also a part of the desktop configuration.

Setting the Top and Initial Languages in your Language Menus

The "top" language in your UniEdit L: menu is determined by the default language .INI file defined in the "DefaultIni=" line in the [WinCALIS] section of WINCALIS.INI. The default language .INI file is set by default for the United States and Western Europe to "LATIN1.INI" If the default .INI file is changed to some other language .INI file, the "top" language will also be changed accordingly.

The initial language for use on startup in UniEdit (also by default set to "English/W.Europe") is also customizable by the user. The initial language is determined by the setting of the InitialLanguage= line in the [Initial] section of the WINCALIS.INI file. The out-of-the-box value for this line is InitialLanguage=English/W.Europe.

Rearranging the Language Menus

The list of languages and keyboards in the L:[Language] menu in UniEdit is constructed dynamically out of the language keyboard *.INI and *.WCK files present in the UniEdit program directory each time you run either of these programs.

The list is arranged into language or script grouping submenus, each one containing various language keyboards, based on the list of Submenus and *.INI files in the [WinCALIS] section of WINCALIS.INI reproduced below.

A typical Language Menu-related section of WINCALIS.INI

[WinCALIS]
.
.
.
SubMenu1=Arabic script
SubMenu2=African
SubMenu3=Asian (other)
SubMenu4=Chinese
SubMenu5=Europe (other)
SubMenu6=Japanese
SubMenu7=Special Fonts
SubMenu8=Cyrillic script
SubMenu9=Hebrew script
SubMenu10=Korean
SubMenu11=Indic scripts
Language20=AMHARIC
Language60=ARMENIAN

```
Language30=ARABICC
Language50=ARABICX
Language40=ARABUSG
Language90=BELARUS
Language140=BOPOMOFO
Language100=BULGARIA
Language150=CANGJIE
Language120=CENTASIA
Language130=CHINESE
Language180=CROATIAN
Language190=CZECH
Language220=DUTCH
Language230=EEUR
Language240=ENGLISH
Language250=FORMSYMB
Language260=FRENCH
Language280=GEORGIAN
Language300=GERMAN
Language320=GREEKCAL
Language330=GREEKPC
Language360=HAUSA
Language400=HEBREWI
Language420=HEBREWK
Language430=HEBREWT
Language410=HEBREWU
Language510=HIRAGANA
Language460=IBM-PC
Language470=IPA
Language480=ITALIAN
Language540=KANJICJ
Language530=KATAKANA
Language560=KOREANK
Language570=KOREANM
Language580=KOREANS
Language590=KOREANU
Language600=KORHANJA
Language640=LATIN
Language650=LATVIAN
Language660=LITHUAN
Language680=MACEDON
Language450=MAGYAR
Language700=MONGOL
Language710=PASHTO
Language720=PERSIANC
Language730=PERSIANP
Language740=PERSIANX
Language750=POLISH
Language760=PORTUG
Language490=ROMAJI
Language500=ROMAJINC
Language810=ROMANIAN
Language780=RUSSIANC
Language790=RUSSIANS
Language800=RUSSTECH
```

Language820=SCANDIN
Language830=SERBIAN
Language840=SINDHI
Language850=SLOVAK
Language870=SPANISH
Language890=TAGALOG
Language920=THAI
Language930=TRANSLIT
Language940=TURKISH
Language950=UKRAINE
Language960=URDU
Language970=VIETNAME
Language980=YIDDISH
Language990=YORUBA

For example, the line "Language100=BULGARIA" tells UniEdit to look for a file named BULGARIA.INI. Its sequence in the list of ascending numbers "Language#" more or less determines the alphabetical sequence in the language menu or submenu of the language keyboard identified within BULGARIA.INI, which happens to be Bulgarian.

UniEdit then looks in the file BULGARIA.INI. It looks like this:

A typical language keyboard *.INI file

[Language]
Name=Bulgarian
Face=Times
SubMenuID=8
RightLeft=0
CharSet=0
Keyboard=Bulgaria
LanguageID=100
Pitch=2
DeadKey=
Font1=Kiev
Font2=Moscow
Font3=Sofia

[FontSize]
Tiny=12
Fine=16
Small=19
Medium=21
Large=32
ExtraLarge=48
Huge=66

The "Name=Bulgarian" line determines the menu listing, which will be just that, "Bulgarian," and the SubMenuID=8 line places the listing for Bulgarian in the "Cyrillic script" subdirectory.

UniEdit then gets the name of the *.WCK keyboard file from the "Keyboard=Bulgaria" line and looks for the file BULGARIA.WCK. If it finds such a file, it will fully enable the menu listing. If it fails to find the file, it will display the menu listing for Bulgarian dimmed.

For the ANSI character set (Unicode page 0) you can freely add as many fonts here as are available (and which you believe will also be available to students working your lessons). For other character sets, fonts must be consistent with the UniEdit Unicode-character-to-font mapping tables UNI2FNT1.WCF....UNI2FNT6.WCF. Fonts designed for use in non-Unicode compliant systems, such as Windows 3.1 word processing programs, will not work with UniEdit. If you have questions, feel free to contact the Duke University Humanities Computing Facility or its agents.

You can rearrange your own Language menus as you please. For example, you can promote your favorite language keyboard to membership in the main menu by removing the SubMenuID= line, or just "commenting it out" with a semicolon thus:

;SubMenuiD=8

and you can move it up or down the menu by changing its LanguageID#.

The [FontSize] section of the language keyboard *.INI file determines the actual height in pixels of each font size used by UniEdit. See Changing the Default Font Size for Use in UniEdit.

(The lines Face=, CharSet=, Pitch=, and DeadKey= are no longer used in the UniEdit system.)

Setting Default Directories

Setting the Conversions Subdirectory

The "CONVERT" subdirectory holds code conversion utilities and filters, if during Setup you selected the option to install "Code Conversion Utilities." If you did not select this option and you wish to convert files from one coding standard to another, you will need to run Setup again and select this option to install the necessary files.

The path to the "CONVERT" subdirectory is defined in the WINCALIS.INI file in the main UniEdit subdirectory. It can be redefined to whatever path you choose by carefully editing the line in the [WinCALIS] section

   Conversions= ...

Setting the Documents Subdirectory

The "DOCS" subdirectory can be used for your document text files. It is not essential to the operation of UniEdit programs and may be deleted if you wish.

The path to the "DOCS" subdirectory can be redefined to whatever path you choose by carefully editing the line in the [CalAuth] section

   Documents= ...

Setting the "Home" Directory

You will not ordinarily need to change the program directory setting. But if you need to move the entire UniEdit program directory and all its subdirectories to another drive or directory, you will then need to change the "Home=" line in WINCALIS.INI, along with a number of other lines. If UniEdit was installed in the default directory location "WINCALIS" on drive C:, the "Home=" line would look like this.

Home=C:\WINCALIS\

It is also important for network installations of UniEdit to be aware of the drive letter for the "Home" and other directories in WINCALIS.INI. For example, it is sometimes the case with installation on a network server that the DOS drive letter by which that server is known to a local workstation changes. If that happens, UniEdit will not function properly on the local workstation. See Installing UniEdit on a Network.

Be careful in editing the "Home=" line to retain the final backslash (reverse solidus).

Setting the Dictionaries Directory

The "DICTS" subdirectory holds any preinstalled on-line dictionaries.

The path to the "DICTS" subdirectory is defined in the WINCALIS.INI file in the main UniEdit subdirectory. It can be redefined to whatever path you choose by carefully editing the line in the [WinCALIS] section

   Dictionaries= ...

Linking to a Spelling Checker

The UniEdit Workstation Setup package includes a shareware Windows spelling checker program, WinSpell 3.0®, which may be freely distributed.

WinSpell or another user-supplied spelling checker can be "linked" to UniEdit and UniEdit menus by specifying the program name and the DOS path to it in the "SpellChecker=" line in the [WinCALIS] section of WINCALIS.INI. For example:

SpellChecker=c:\winspell\winspell.exe

will enable the "Check Spelling" menu item in the UniEdit "Utilities" menu. If that menu item is selected, the WinSpell program will be executed.

The Default Character-to-Font Mapping File WCFNT.INF

UniEdit always displays text in the typeface defined in the file "WCFNT.INF" in the UniEdit program directory. This typeface is by default "Times New Roman" for the ANSI character set (Unicode page 00) and extended Latin character set (Unicode pages 01-02, etc.). The default typeface for use in UniEdit may be changed to Arial, Courier New, Fixedsys, MS Sans Serif (the font you are reading in now! -- a raster screen font which is preferred to TrueType fonts for Windows Help because of its superior legibility at low resolutions such as on computer monitors), or any other font available to the user.

Just edit the file "WCFNT.INF" with a plain text editor, such as the Windows Notepad. (Keep a backup copy of the original file.) For example, to change the font for just the basic Windows ANSI (Latin1) character set to Arial, the first line in the [FaceNames] section should be changed to read:

00=Arial.

In a similar fashion the default font for any character set may be changed by editing the line(s) for its corresponding code pages in "WCFNT.INF."

If you are currently running UniEdit, you will have to exit and restart UniEdit for the new settings to take effect.

If you wish to make only a temporary change in your UniEdit font, you can leave the Notepad with running in the background, with the WCFNT.INF file still open in it. Then after you have finished using UniEdit, you can return to the Notepad and change the font back again.

See "Rearranging the Language Menus" for how to change the list of fonts displayed in the Fonts dialog box.

(For how to change the ANSI font size, you will need to edit LATIN1.INI. See The Main WinCALIS Initialization File WINCALIS.INI.

Sample Portion of First Part of WCFNT.INF File

```
[FaceNames]
00=Times New Roman
01=Times New Roman
02=Times New Roman
03=Athens
04=Kiev
05=Tel Aviv
06=Sahara
.
.
```

The latter part of the WCFNT.INF file contains "magnification" percentages for the fonts used for each character set. E.g., the "Athens" font for Greek is by default magnified to 110% of its basic size.

Sample Portion of Second Part of WCFNT.INF File

```
[03]
Athens=110

[04]
Kiev = 120
Sofia = 120

[05]
Tel Aviv = 120

[06]
Baghdad = 200
Sahara = 200
```

New Configuration Dialog Box Coming Down the Pike:

Author Workstation WinCALIS Startup:

  o Start Author Workstation WinCALIS in Text Conversion (*.CAL) mode

  o Start Author Workstation WinCALIS in No Conversion (*.WCL) mode

Location of Files:

  Main Scripts Directory: c:\wincalis\scripts

  Shared Graphics Directory: c:\wincalis\graphics

  Shared Multimedia Directory: c:\wincalis\mltmedia

  Conversion Utilities Directory: c:\wincalis\convert

  Default Documents Directory: c:\wincalis\documents

Keyboard Arrangement (for display in keyboard help):

  o Backslash key (\) on top row, next to Backspace key

  o Backslash key (\) on 2nd row from top, below Backspace key

  o Backslash key (\) on bottom row at extreme right, below Enter key

  o Backslash key (\) on bottom row at extreme left

L:[Language] Menu Top Language: English/W.Europe

   (change by setting Default Language Keyboard)

Default Language Keyboard: LATIN1.INI

Author Fonts:

  Author ANSI Font: Times New Roman

    (change by setting WCFNT.INF 00= line below)

  Author Font Size Pixel Height: 21

   (change by setting Font Size Medium= below)

  LATIN1.INI current settings:

  Font Size Pixel Heights:

    Tiny=12
    Fine=16
    Small=19
    Medium=21

```
    Large=32
    ExtraLarge=48
    Huge=66
  Fonts:
    Font1=Times New Roman
    Font2=Arial
    Font3=Courier New
    Font4=Fixedsys
    Font5=MS Sans Serif
    Font6=
```

WCFNT.INF current settings:

```
 [FaceNames]
 00=Times New Roman
 01=Times New Roman
 02=Times New Roman
 03=Athens
 04=Kiev
 05=Tel Aviv
 06=Sahara
 .
 .
```

Defining Language Keyboard Files

The UniEdit program allows for the remapping of keyboard layouts and the display of non-Roman character sets. The user may select a keyboard mapping and a character set from the pull-down "L:[anguage]" menu.

Once a language is selected, choosing "keyboard" from the "Help" menu will display the layout of the currently selected font. Accents and key combinations for producing characters are also displayed.

In order to make a character set and a keyboard layout available to the user, three files are required.

1.   *LANGUAGENAME*.**INI**

2.   *LANGUAGENAME.WCK*

3.   The appropriate character font, contained in font files such as *FONTFILE*.**FON**

Font Files

Font files contain character sets in a format which can be displayed in Windows.

They should be placed in the subdirectory \FONTS located below the Windows 95 directory.

UniEdit comes bundled with all the fonts necessary to display all character sets.

Language Keyboard ".INI" Files

There is a file ending with the extender .INI for each UniEdit language keyboard. The .INI file is the initialization file for a particular language keyboard, which specifies the keyboard arrangement, name, fonts, on-line dictionary and other attributes of the language keyboard.

It must be placed in the UniEdit home directory--by default \WINCALIS.

You may edit .INI files with a plain text editor such as the Windows Notepad or the DOS EDIT program. For example, you may wish to specify a different on-line dictionary or a new font available for a keyboard. Be sure you understand the particular structure documented below.

The following file "RUSSIANS.INI" is an example of such a file

Comments on the file are preceded with a semicolon (;)

[Language]

Name=Russian-Standard

;The name of the language keyboard that

;appears in the "L:[Language]" pull-down menu.

SubmenuID=8

;This value determines which submenu, if any, the

;keyboard appears in, in this case "Cyrillic"

Face=WC2Uni04

;Not currently used.

RightLeft=0

;The zero value for RightLeft indicate the

;character set is written from left-to-right.

;A value of "1" means right-to-left.

CharSet=0

;This value is either zero, 2, or 255 - ANSI, Symbol or OEM

;(Not currently used)

Keyboard=RussianS

;This is the DOS filename of the .WCK file

;which specifies the keyboard arrangement.

LanguageID=790

;This number must correspond with the ID number

;specified for the RussianS keyboard in the WINCALIS.INI

;main initialization file. The specific number assigned

;will determine the order of the name in the

;L[anguage]: menu

;The highest possible value for this number is 999.

;See "Rearranging the Language Menus"

Pitch=2

;This value indicates a proportional font.

;Zero would indicate a fixed font. (Not currently used.)

DeadKey=

;(Not currently used--overstrike (dead) key is always F2

Font1=Kiev

Font2=Moscow

Font3=Sofia

;Typefaces available for use with this keyboard and character

;set are specified here. They will appear as choices when the

;user selects the Set Font menu option in the WinCALIS

;Author Format menu.

[FontSize]

Tiny=12

Fine=16

Small=19

Medium=21

Large=32

ExtraLarge=48

Huge=66

;The adjectives on the left are the names given

;to font point sizes in WinCALIS. The numeric values

;on the right indicate the pixel height for each size.

[Dictionary]

DictFile=wcruseng.dic

IndexFile=wcruseng.idx

;These lines specify the filenames of the on-line dictionary

;and its index file to be accessed in the directory defined in the

;Dictionaries= line of WINCALIS.INI when the user presses F5 or selects

;the menu option Tools, Look Up Word while "Russian-Standard"

;is the current language keyboard. If no filenames are specified,

;the default dictionary with the filenames WINCALIS.DIC and

;WINCALIS.IDX will be used instead.

;See How to Create Your Own On-line Dictionary

;and Setting the Dictionaries Directory

".WCK" Files

(NOTE: Some of the information below is now obsolete.)

This is a binary format file which specifies which keystrokes produce which characters.

It cannot be edited directly. To create a ***.WCK** file you must edit a text file with the same name and the extension ***.MAP**. After this file is edited it is processed into the binary file format of a ***.WCK** file.

After editing such a file select the "Create Keyboard" option from the "File" pull-down menu in UniEdit. The "*.MAP" file will be compiled into a "*.WCK" file.

While the ".MAP" file is not needed to map characters once the ".WCK" file is created, you should save it in case you wish to make changes to the keyboard mapping.

The map file consist of several headings which are written inside of square brackets. ([ ])

Below each of these topics are the actual characters OR their numeric character values, enclosed in curly braces ({ }), affected by the headings under which they are listed.

Comments can be included in a ".MAP" file by preceding them with a semicolon (;).

Text which appears on a line after a semicolon will be ignored.

If you wish to remap a square brace ([,]), a curly brace ({, }), or a semicolon (;) you must use the character values for these characters rather than the characters themselves.

To make a ".MAP" file as readable as possible it is recommended that you list each combination of characters on a single line. This is not required.

A simple debugging feature in the program which compiles your ".MAP" file into a ".WCK" file will give an error message telling you which line of the ".MAP" file caused a problem.

Because creating a new ".WCK" file overwrites the previous one (unless you give it a new name) it is recommended that you keep backups.

Also, it is a good idea to develop ".MAP" files incrementally, and convert them to ".WCK" files regularly. This process takes only a few seconds and will make pinpointing errors a simple task.

Note that if you use the UniEdit you can type in character from foreign character sets instead of their character values.

This is not recommended because such files appear unintelligible when other fonts are used to display them or when they are edited with DOS-based word processors.

At the end of the conversion process a message box will appear to confirm that the ".WCK" file has been successfully generated.

Click on one of the following topics for more information on the contents of a ".MAP" file.

All of the headings listed below are required in every ".MAP" file, even if they are not used, with the exception of [Position]

The [MAP] heading

The [ALPHA] heading

The [PUNCTUATION] heading

The [SPACE] heading

The [DIGIT] heading

The [VOWEL] heading

The [UPPER/LOWER] heading

The [ACCENT/NOACCENT] heading

The [POSITION] heading

The [ACCENTS] heading

The [ENDV1] heading

Sample ".MAP" File HEBREWI.MAP

```
;hebrewi.map - Israeli keyboard

[map]
t {05D0} ;alef
c {05D1} ;bet
d {05D2} ;gimel
s {05D3} ;dalet
v {05D4} ;he
u {05D5} ;vav
z {05D6} ;zayin
j {05D7} ;het
y {05D8} ;tet
h {05D9} ;yod
l {05DA} ;final kaf
f {05DB} ;kaf
k {05DC} ;lamed
o {05DD} ;final mem
n {05DE} ;mem
i {05DF} ;final nun
b {05E0} ;nun
x {05E1} ;samekh
g {05E2} ;ayin
{003B}{05E3} ;final pe ;mapped to semicolon
p {05E4} ;pe
. {05E5} ;final tsadi
m {05E6} ;tsadi
e {05E7} ;qof
r {05E8} ;resh
a {05E9} ;shin
, {05EA} ;tav
: {05C3} ; Hebrew punctuation sof pasuq
/ .
q /
w ,
S {05B0} ;HEBREW POINT SHEVA
```

O {05B1} ;HEBREW POINT HATAF SEGOL
I {05B2} ;HEBREW POINT HATAF PATAH
B {05B3} ;HEBREW POINT HATAF QAMATS
A {05B4} ;HEBREW POINT HIRIQ
F {05B5} ;HEBREW POINT TSERE
C {05B6} ;HEBREW POINT SEGOL
H {05B7} ;HEBREW POINT PATAH
D {05B8} ;HEBREW POINT QAMATS
E {05B9} ;HEBREW POINT HOLAM
G {05BB} ;HEBREW POINT QUBUTS
J {05BD} ;HEBREW POINT METEG
K {05BE} ;HEBREW PUNCTUATION MAQAF
L {05BF} ;HEBREW POINT RAFE
M {05C0} ;HEBREW PUNCTUATION PASEQ;;HEBREW POINT PAS
N {FFFF}
P {FFFF}
Q {FFFF}
R {FFFF}
T {FFFF}
U {FFFF}
V {FFFF}
W {FFFF}
X {FFFF}
Y {FFFF}
Z {FFFF}

[alpha]
;Unicode 1.1; WinCALIS 1.0 draft Hebrew; Hebrew Text CALIS codes; description
{05B0} ;HEBREW POINT SHEVA
{05B1} ;HEBREW POINT HATAF SEGOL
{05B2} ;HEBREW POINT HATAF PATAH
{05B3} ;HEBREW POINT HATAF QAMATS
{05B4} ;HEBREW POINT HIRIQ
{05B5} ;HEBREW POINT TSERE
{05B6} ;HEBREW POINT SEGOL
{05B7} ;HEBREW POINT PATAH
{05B8} ;HEBREW POINT QAMATS
{05B9} ;HEBREW POINT HOLAM
{05BB} ;HEBREW POINT QUBUTS
{05BC} ;HEBREW POINT DAGESH OR MAPIQ;;HEBREW POINT DAGESH
{05BD} ;HEBREW POINT METEG
{05BE} ;HEBREW PUNCTUATION MAQAF
{05BF} ;HEBREW POINT RAFE
{05C0} ;HEBREW PUNCTUATION PASEQ;;HEBREW POINT PAS
{05C1} ;HEBREW POINT SHIN DOT
{05C2} ;HEBREW POINT SIN DOT
{05C3} ;HEBREW PUNCTUATION SOF PASUQ
{05D0} ;alef
{05D1} ;bet
{05D2} ;gimel
{05D3} ;dalet
{05D4} ;he
{05D5} ;vav
{05D6} ;zayin

{05D7} ;het
{05D8} ;tet
{05D9} ;yod
{05DA} ;final kaf
{05DB} ;kaf
{05DC} ;lamed
{05DD} ;final mem
{05DE} ;mem
{05DF} ;final nun
{05E0} ;nun
{05E1} ;samekh
{05E2} ;ayin
{05E3} ;final pe
{05E4} ;pe
{05E5} ;final tsadi
{05E6} ;tsadi
{05E7} ;qof
{05E8} ;resh
{05E9} ;shin
{05EA} ;tav
{05F0} ;HEBREW LIGATURE YIDDISH DOUBLE VAV
{05F1} ;HEBREW LIGATURE YIDDISH VAV YOD
{05F2} ;HEBREW LIGATURE YIDDISH DOUBLE YOD
{05F3} ;HEBREW PUNCTUATION GERESH
{05F4} ;HEBREW PUNCTUATION GERSHAYIM
{FB1E} ;HEBREW POINT JUDEO-SPANISH VARIKA;HEBREW POINT VARIKA;
{FB1F} ;HEBREW LIGATURE YIDDISH YOD YOD PATAH; [05D9] & [05D9] & [05B7]
{FB20} ;HEBREW LETTER ALTERNATIVE AYIN;[05E2]
{FB21} ;HEBREW LETTER WIDE ALEF;[05D0]
{FB22} ;HEBREW LETTER WIDE DALET;[05D3]
{FB23} ;HEBREW LETTER WIDE HE;[05D4]
{FB24} ;HEBREW LETTER WIDE KAF;[05DB]
{FB25} ;HEBREW LETTER WIDE LAMED;[05DC]
{FB26} ;HEBREW LETTER WIDE FINAL MEM;[05DD]
{FB27} ;HEBREW LETTER WIDE RESH;[05E8]
{FB28} ;HEBREW LETTER WIDE TAV;[05EA]
{FB29} ;HEBREW LETTER ALTERNATIVE PLUS SIGN;[002B]
{05CA} ;HEBREW LETTER SHIN WITH SHIN DOT;[05E9] & [05C1]
{05CB} ;HEBREW LETTER SHIN WITH SIN DOT;[05E9] & [05C2]
{FB2C} ;HEBREW LETTER SHIN WITH DAGESH AND SHIN DOT;[05E9] & [05BC] & [05C1]
{FB2D} ;HEBREW LETTER SHIN WITH DAGESH AND SIN DOT;[05E9] & [05BC] & [05C2]
{FB2E} ;HEBREW LETTER ALEF WITH PATAH;[05D0] & [05B7]
{FB2F} ;HEBREW LETTER ALEF WITH QAMATS;[05D0] & [05B8]
{FB30} ;HEBREW LETTER ALEF WITH MAPIQ;[05D0] & [05BC]
{05CC} ;HEBREW LETTER BET WITH DAGESH;[05D1] & [05BC]
{FB32} ;HEBREW LETTER GIMEL WITH DAGESH;[05D2] & [05BC]
{FB33} ;HEBREW LETTER DALET WITH DAGESH;[05D3] & [05BC]
{FB34} ;HEBREW LETTER HE WITH MAPIQ;[05D4] & [05BC]
{FB35} ;HEBREW LETTER VAV WITH DAGESH;[05D5] & [05BC]
{FB36} ;HEBREW LETTER ZAYIN WITH DAGESH;[05D6] & [05BC]
{FB38} ;HEBREW LETTER TET WITH DAGESH;[05D8] & [05BC]
{FB39} ;HEBREW LETTER YOD WITH DAGESH;[05D9] & [05BC]

{FB3A} ;HEBREW LETTER FINAL KAF WITH DAGESH;[05DA] & [05BC]
{05CF} ;HEBREW LETTER KAF WITH DAGESH;[05DB] & [05BC]
{FB3C} ;HEBREW LETTER LAMED WITH DAGESH;[05DC] & [05BC]
{FB3E} ;HEBREW LETTER MEM WITH DAGESH;[05DE] & [05BC]
{FB40} ;HEBREW LETTER NUN WITH DAGESH;[05E0] & [05BC]
{FB41} ;HEBREW LETTER SAMEKH WITH DAGESH;[05E1] & [05BC]
{FB43} ;HEBREW LETTER FINAL PE WITH DAGESH;[05E3] & [05BC]
{05CD} ;HEBREW LETTER PE WITH DAGESH;[05E4] & [05BC]
{FB46} ;HEBREW LETTER TSADI WITH DAGESH;[05E6] & [05BC]
{FB47} ;HEBREW LETTER QOF WITH DAGESH;[05E7] & [05BC]
{FB48} ;HEBREW LETTER RESH WITH DAGESH;[05E8] & [05BC]
{FB49} ;HEBREW LETTER SHIN WITH DAGESH;[05E9] & [05BC]
{05CE} ;HEBREW LETTER TAV WITH DAGESH;[05EA] & [05BC]
{FB4B} ;HEBREW LETTER VAV WITH HOLAM;[05D5] & [05B9]
{FB4C} ;HEBREW LETTER BET WITH RAFE;[05D1] & [05BF]
{FB4D} ;HEBREW LETTER KAF WITH RAFE;[05DB] & [05BF]
{FB4E} ;HEBREW LETTER PE WITH RAFE;[05E4] & [05BF]
{FB4F} ;HEBREW LIGATURE ALEF LAMED;[05D0] & [05DC]

[punctuation]
!@#$%^&*()_+-=\',./|:”<>?
{003B}{005B}{005D}{007B}{007D} ;i.e. ;[]{}
{05C3} ;Hebrew punctuation sof pasuq

[space]
{0020} ; {32}

[digit]
1234567890

[vowel]

[upper/lower]

[accent/noaccent]
{FB2A}{05E9} ;shin shindot
{FB2B}{05E9} ;shin sindot
{FB31}{05D1} ;bet dagesh
{FB44}{05E4} ;pe dagesh
{FB4A}{05EA} ;tav dagesh

[Position]
;isolate, initial, medial, final;
{05D0}{05D0}{05D0}{05D0} ;alef
{05D1}{05D1}{05D1}{05D1} ;bet
{05D2}{05D2}{05D2}{05D2} ;gimel
{05D3}{05D3}{05D3}{05D3} ;dalet
{05D4}{05D4}{05D4}{05D4} ;heh
{05D5}{05D5}{05D5}{05D5} ;vav
{05D6}{05D6}{05D6}{05D6} ;zayin
{05D7}{05D7}{05D7}{05D7} ;het
{05D8}{05D8}{05D8}{05D8} ;tet
{05D9}{05D9}{05D9}{05D9} ;yud
{05DB}{05DB}{05DB}{05DA} ;kaf

```
{05DC}{05DC}{05DC}{05DC} ;lamed
{05DE}{05DE}{05DE}{05DD} ;mem
{05E0}{05E0}{05E0}{05DF} ;nun
{05E1}{05E1}{05E1}{05E1} ;sameh
{05E2}{05E2}{05E2}{05E2} ;ayin
{05E4}{05E4}{05E4}{05E3} ;peh
{05E6}{05E6}{05E6}{05E5} ;tsadi
{05E7}{05E7}{05E7}{05E7} ;qof
{05E8}{05E8}{05E8}{05E8} ;resh
{05E9}{05E9}{05E9}{05E9} ;shin
{05EA}{05EA}{05EA}{05EA} ;tav

[accents]
shindot ' ; {05C1}
sindot ` ; {05C2}
dagesh , ; {05BC}

[shindot]
{05E9}{FB2A} ;shin

[sindot]
{05E9}{FB2B} ;shin

[dagesh]
{05D1}{FB31} ;bet
{05E4}{FB44} ;pe
{05EA}{FB4A} ;tav

[endv1]
```

".MAP" File Heading [MAPPING]

This heading is used to map or remap characters to the keyboard.

Mappings are listed in pairs.

First, the keyboard character which is to be replaced is listed.

Second, the character which will replace it is listed.

For example

[MAPPING]

a f

{0064} {009B}

As specified above, pressing the "a" key on the PC keyboard will no longer produce an "a" but

In the same way, pressing the "d" key (character value 0064) will produce character 009B instead.

You may specify characters by entering them directly or by their character values enclosed in curly braces.

Separate each entry with a least one blank space.

Note that the first value or character must be one of the keys located on the PC keyboard.

Depending on what font is used to display these files, you may see Cyrillic, Greek, or Roman characters, among others.

Files created using a keyboard mapping intended for Russian will look like garbage if they are displayed with another font.

When users select the keyboard help feature in UniEdit, the keyboard window will display the characters as you have remapped them.

".MAP" File Heading [ALPHA]

This heading is used to list all of the characters in a font which are alphabetic.

This feature is useful because it allows the script author to write diagnostic messages based on the presence or absence of an alphabetic character.

Characters may be listed under this heading by entering them directly or by typing in their character value enclosed in curly braces.

".MAP" File Heading [PUNCTUATION]

This heading is used to specify all of the characters in a font which are used for punctuation.

WinCALIS uses this information to determine if a student response is correct or incorrect based on the improper use of punctuation.

Characters under this heading may be entered directly or by listing their character value in curly braces.

".MAP" File Heading [SPACE]

Specify what character (or possibly characters) is used as the space character.

WinCALIS uses this information to ignore multiple spaces between words in a student response.

This character must be listed by its character value. (You can't display an empty character!)

This character is nearly always {0020}.

".MAP" File Heading [DIGIT]

Specify which characters are numeric.

At this time this feature is not used. Potentially it could be used to distinguish answers which contain numeric characters from other answers. Call us if you need this for your courseware development.

Although unused, every ".MAP" file must contain this heading. No characters need follow it.

Characters can be entered directly or by listing their character value inside of curly braces.

".MAP" File Heading [VOWEL]

This heading is only useful for Chinese at present. If you feel that the ability of WinCALIS to distinguish consonants and vowels would be useful for some other language, this feature could easily be added.

This heading must appear in all "

Characters may be entered directly or by listing their character value inside of curly braces.

".MAP" File Heading [UPPER/LOWER]

The information in this and the following section is no longer used by WinCALIS/UniEdit. The global information file WINCALIS.WGI (q.v.) is used instead.

This heading is used to list uppercase/lowercase equivalents.

This information is used by WinCALIS to determine whether case is significant in judging student responses.

Characters under this heading are listed in pairs.

First the upper case character or its character value in curly braces.

Second the lower case character or its character value in curly braces.

".MAP" File Heading [ACCENT/NOACCENT]

This heading is used to list accented/unaccented equivalents.

This information is used by WinCALIS to determine whether the proper use of accents is significant in judging student responses.

Characters under this heading are listed in pairs.

First the accented character or its character value in curly braces.

Second the unaccented character or its character value in curly braces.

".MAP" File Heading [POSITION]

This is the only optional heading in a "

It is used to specify what characters are affected by their position in a word.

Characters are listed in sets of four in the following order:

{[independent]}{{initial]}{[medial]}{[final]}

Characters may be listed by entering the characters themselves or by listing their character values in curly braces.

This feature is used with the UniEdit Persian and Arabic fonts.

".MAP" File Heading [ACCENTS]

This heading is used to precede a listing of all of the names given to diacritic marks.

After the name of the diacritic, list the single character which will be used in combination with the overstrike key F2 to form an accented character.

Diacritic names must not contains spaces.

For example:

[ACCENTS]

umlaut "

nonsense {0078}

Two accent keys are now defined with the names "umlaut" and "nonsense."

When keyboard help is selected from within UniEdit, a keyboard will display with the double quote (") and character 0078 - a lowercase "x" in pale blue.

This indicates to the user that these keys can be used in combination with the dead key to produce a new character.

Beneath the [ACCENTS] heading each accent is listed a second time as a heading itself.

Following this is a listing of all characters which can be typed and then altered or changed completely with the overstrike key/accent key combination.

For example:

[umlaut]

a {00E4}

{0075} {00FC}

[nonsense]

a z

{0064} b

As specified above, the key combination "a" + the dead key + the "umlaut" accent key (a double quote as defined in the [accents] section) will produce character 00E4 - an "a" with two dots over it.

In the same way, a lower case "u" (character 0075) becomes a "u" with two dots over it (character 00FC).

The "nonsense" accent key changes a lower case "a" into a lower case "z."

Character 0064 (a lower case "d") becomes a lower case "b."

With the keyboard help window open the user can click with the secondary mouse button on an accent key to see which keys it affects and what characters they produce. The selected accent key turns red and the affected keys display in green.

According to the examples above, if the user opens the keyboard window and clicks with the secondary mouse button the "nonsense" accent key the lower case "x" turns from pale blue to green. The two characters it affects - the "a" and the "d" are redisplayed as a "z" and a "b."

Clicking on or pressing the "umlaut" accent key (the double quote) changes the "a" and the "u" to umlauted characters.

All listings under accent headings must be paired - original character and character to be displayed after the overstrike key/accent key combination is struck.

Characters may be entered directly or listed by character value within curly braces.

".MAP" File Heading [ENDV1]

No characters are listed beneath this heading. It simply marks the end of the file. However, there should be at least one new line after the heading.

The UniEdit Global Information File WINCALIS.WGI

The binary global information file WINCALIS.WGI in UniEdit replaces much of the language-specific information which was formerly in the individual language keyboard *.WCK files WinCALIS 1x regarding what constitutes punctuation, capitalization (upper-and-lower case relationships), and accented-unaccented character relationships. WinCALIS now is able to evaluate these linguistic and orthographic characteristics properly, no matter which keyboard is currently selected.

Note that the [upper/lower] section of WINCALIS.WGI is used for case-independent searches in the Find and Replace dialog boxes. It could also be used in the future for sorting purposes, along with the [accent/noaccent] section.

The name WINCALIS.WGI is arbitrary and is identified to the UniEdit system by the "GlobalInfoFile=" line in WINCALIS.INI. By default, the file is called WINCALIS.WGI and is in the main program directory.

WINCALIS.WGI is created from its ASCII text counterpart WINCALIS.INF by opening WINCALIS.INF in the Scratchpad as an ASCII/ANSI file, then clicking on the Create Global File menu item in the File menu. WINCALIS.INF is included in the UniEdit Workstation, for browsing and editing on the part of users. Please notify the Humanities Computing Facility of Duke University if you find errors or omissions in the file.

Samples entries under each subheading in WINCALIS.WGI

[non-spacingmarks] (unused)
[punctuation]
{0021} ;!
{0022} ;"
{0023} ;#
{0024} ;$
.

.
{037E} ;Greek question mark (semicolon)
{05C3} ;Hebrew
{060C} ;Arabic comma
{061B} ;Arabic semicolon
{061F} ;Arabic question mark
{066A} ;Arabic percent sign
{06D4} ;Arabic period (Farsi period) (Urdu)?
{1255} ;Ethiopian (Amharic) wordspace
{1256} ;Ethiopian (Amharic) comma
.
.
[upper/lower]
A a
B b
C c
D d
E e
.
.
{00DC}{00FC} ;U diaeresis u diaeresis
{00DD}{00FD} ;Y acute y acute
{00DE}{00FE} ;capital thorn small thorn
{00DF}{00DF} ;esszed esszed
{0100}{0101} ;A macron a macron
{0102}{0103} ;A breve a breve
{0104}{0105} ;A ogonek a ogonek
{0106}{0107} ;C acute c acute
{0108}{0109} ;C circumflex c circumflex
.
.
[accent/noaccent]
{00C0} A   ;A grave
{00C1} A   ;A acute
{00C2} A   ;A circumflex
{00C3} A   ;A tilde
.
.
{FB44}{05E4} ;pe dagesh
{FB4A}{05EA} ;tav dagesh
{FD15}{00EA} ;e circumflex grave e circumflex
{FD39}{00EA} ;e circumflex acute e circumflex
{FD60}{00EA} ;e circumflex macron e circumflex
{FD61}{00EA} ;e circumflex hacek e circumflex
.
.
[endv1]

Deleting Nonessential Files to Save on Disk Space

When you install UniEdit, it installs *.INI, *.MAP, and *.WCK files for all available languages into your program directory by default. The Delete Language menu item in the L:[Language] menu deletes languages from the listing in the menu, but it does not delete any files from your disk.

If you need to economize on disk space, after deleting a language from your menu, you may also use the Windows File Manager, DOS commands, or another tool to delete its *.INI, *.MAP, and *.WCK files from your hard disk, or move them to a backup storage location.

*.MAP files are simply the text versions of binary keyboard *.WCK files, which are those used by UniEdit. If you do not anticipate the need to edit or add any keyboard file, you may delete *all* the *.MAP files from your system. See the section below for other nonessential files you may delete to conserve disk space.

If necessary, you can always recover deleted files by reinstalling the UniEdit Workstation.

Files Installed During UniEdit Setup

The following list shows all the files installed during a *complete* setup. Essential files are marked with an asterisk. Some other files, usually noted, may be essential to *your* specific needs.

UniEdit Program directory and subdirectories:

amharic.ini	;you need an *.ini file and *.wck file	calgloss.hlp	;(Part of Users Manual)
		cangjie.ini	
amharic.wck	;for each language keyboard you use	centasia.ini	
		centasia.wck	
amharic.map	;*.map files are only needed if you modify	centasia.map	
		chinese.ini	
arabicc.ini	';or create keyboard (See	croatian.ini	
arabicc.map	;Defining Language Keyboard Files)	croatian.wck	
		croatian.map	
arabicc.wck		czech.ini	
arabicx.ini		czech.wck	
arabicx.map		czech.map	
arabicx.wck		dutch.ini	
arabusg.ini		dutch.wck	
arabusg.map		dutch.map	
arabusg.wck		eeur.ini	
armenian.ini		eeur.wck	
armenian.wck		eeur.map	
armenian.map		english.ini	
belarus.ini		english.wck	
belarus.wck		english.map	
belarus.map		formsymb.ini	
bopomofo.ini		formsymb.wck	
bopomofo.map		formsymb.map	
bopomofo.wck		french.ini	
bug.txt		french.wck	
bulgaria.ini		french.map	
bulgaria.wck		gencyril.map	
bulgaria.map		generic.map	

georgian.ini

georgian.wck

georgian.map

german.ini

german.wck

german.map

greekcal.ini

greekcal.wck

greekcal.map

greekpc.ini

greekpc.wck

greekpc.map

hausa.ini

hausa.wck

hausa.map

hebrewi.ini

hebrewi.map

hebrewi.wck

hebrewk.ini

hebrewk.map

hebrewk.wck

hebrewt.ini

hebrewt.map

hebrewt.wck

hebrewu.ini

hebrewu.map

hebrewu.wck

hiragana.ini

hiragana.map

hiragana.wck

ibm-pc.ini

ibm-pc.wck

ibm-pc.map

ipa.ini

ipa.wck

ipa.map

*irreg.idx   ;essential for CJK (Chinese-

*irreg.ptr   ;Japanese-Korean) module

italian.ini

italian.wck

italian.map

kanjicj.ini

katakana.ini

katakana.map

katakana.wck

koreank.ini

koreank.map

koreank.wck

koreanm.ini

koreanm.map

koreanm.wck

koreans.ini

koreans.map

koreans.wck

koreanu.ini

koreanu.map

koreanu.wck

korhanja.ini

latin.ini

latin.wck

latin.map

latin1.ega

*latin1.ini   ;this is sine qua non
                      keyboard

*latin1.wck   ; "

latin1.map

latvian.ini

latvian.wck

latvian.map

lithuan.ini

lithuan.wck

lithuan.map

macedon.ini

macedon.wck

macedon.map

magyar.ini

magyar.wck

magyar.map

*mlp.dat

*mlp.exe   ;Multilingual Printer
                      Program

mlp.hlp

*mlptt.exe   ;TrueType Printer Program

mlptt.hlp

mongol.ini

mongol.wck

mongol.map

pashto.ini

pashto.map

pashto.wck

persianc.ini

persianc.map

persianc.wck

persianp.ini

persianp.map

persianp.wck

persianx.ini

persianx.map

persianx.wck

*pinyin.idx   ;essential for CJK module

pinyin.map

*pinyin.ptr   ; "

pinyin.wck

polish.ini

polish.wck

polish.map

portug.ini

portug.wck

portug.map	*uni2fnt5.wcf	
readme.txt	*uni2fnt6.wcf	
romaji.ini	uni2sjis.tab	
romaji.map	*unicjcmp.inx	;essential for CJK module
romaji.wck	*unicjjpn.inx	; "
romajinc.ini	*unicjsmp.inx	; "
romanian.ini	unigrep.exe	;See readme.txt
romanian.wck	unisort.exe	; "
romanian.map	urdu.ini	
russianc.ini	urdu.map	
russianc.wck	urdu.wck	
russianc.map	vietname.ini	
russians.ini	vietname.wck	
russians.wck	vietname.map	
russians.map	*wcchn.dll	;essential *.dll Dynamic
russtech.ini		Link Libraries
russtech.wck	*wcchnbpw.dat	;essential for CJK module
russtech.map	*wcchndsk.dic	; "
scandin.ini	*wcchnmem.dic	; "
scandin.wck	*wcchnptr.dic	; "
scandin.map	*wcdef.dll	
serbian.ini	*wcfnt.inf	;See The Default Character-to-Font Mapping File WCFNT.INF
serbian.wck		
serbian.map		
sindhi.ini	*wcjpn.dll	
sindhi.map	*wcjpn.dic	;essential for CJK module
sindhi.wck	*wcjpnhst.dic	; "
*sjis2wc2.tab ;essential for CJK module	*wckey.dll	
slovak.ini	*wckrn.dll	
slovak.wck	*wckrn.dic	;essential for CJK module
slovak.map	*wckrn.idx	;essential for CJK module
spanish.ini	*wckrn.wcf	;essential for CJK module
spanish.wck	*wcr2l.dll	
spanish.map	*wcr2l.wcf	;essential for Arabic script
tagalog.ini	*wcrtw.dll	
tagalog.wck	*wctxt.dll	
tagalog.map	wincalis.inf	;See The UniEdit Global Information File WINCALIS.WGI
thai.ini		
thai.wck		
thai.map	wincalis.wci	
translit.ini	*wincalis.wgi	;See The UniEdit Global Information File WINCALIS.WGI
translit.wck		
translit.map		
turkish.ini	*wincalis.ini	;See The Main UniEdit Initialization File WINCALIS.INI
turkish.wck		
turkish.map		
ukraine.ini	yiddish.ini	
ukraine.wck	yiddish.map	
ukraine.map	yiddish.wck	
*uni2fnt1.wcf ;Unicode-font mapping tables	yoruba.ini	
	yoruba.wck	
*uni2fnt2.wcf	yoruba.map	
*uni2fnt3.wcf		
*uni2fnt4.wcf		

convert\convtnd.rec ;optional conversion filters
convert\lst2tab3.exe ;only *.tab files are essential
convert\lst2tab5.exe ;unless you want to edit or create new ones
convert\readme.txt
convert\wc2&1004.lst ;*.lst files are text versions
convert\wc2&1004.tab ;of *.tab files
convert\wc2&1040.lst ;See Eight-bit Code Conversions
convert\wc2&1040.tab
convert\wc2&1041.lst
convert\wc2&1041.tab
convert\wc2&1043.lst
convert\wc2&1043.tab
convert\wc2&1250.lst
convert\wc2&1250.tab
convert\wc2&1251.lst
convert\wc2&1251.tab
convert\wc2&1252.lst
convert\wc2&1252.tab
convert\wc2&1253.lst
convert\wc2&1253.tab
convert\wc2&1254.lst
convert\wc2&1254.tab
convert\wc2&1255.lst
convert\wc2&1255.tab
convert\wc2&1256.lst
convert\wc2&1256.tab
convert\wc2&8bit.doc
convert\wc2&8bit.exe
convert\wc2&ansi.lst
convert\wc2&ansi.tab
convert\wc2&brav.lst
convert\wc2&brav.tab
convert\wc2&brov.lst
convert\wc2&brov.tab
convert\wc2&cale.lst
convert\wc2&cale.tab
convert\wc2&calg.lst
convert\wc2&calg.tab
convert\wc2&calh.lst
convert\wc2&calh.tab
convert\wc2&calp.lst
convert\wc2&calp.tab
convert\wc2&calr.lst

convert\wc2&calr.tab
convert\wc2&d437.lst
convert\wc2&d437.tab
convert\wc2&d850.lst
convert\wc2&d850.tab
convert\wc2&d852.lst
convert\wc2&d852.tab
convert\wc2&d855.lst
convert\wc2&d855.tab
convert\wc2&d857.lst
convert\wc2&d857.tab
convert\wc2&d860.lst
convert\wc2&d860.tab
convert\wc2&d861.lst
convert\wc2&d861.tab
convert\wc2&d863.lst
convert\wc2&d863.tab
convert\wc2&d864.lst
convert\wc2&d864.tab
convert\wc2&d865.lst
convert\wc2&d865.tab
convert\wc2&d866.lst
convert\wc2&d866.tab
convert\wc2&d869.lst
convert\wc2&d869.tab
convert\wc2&go74.lst
convert\wc2&go74.tab
convert\wc2&iso1.lst
convert\wc2&iso1.tab
convert\wc2&iso2.lst
convert\wc2&iso2.tab
convert\wc2&iso3.lst
convert\wc2&iso3.tab
convert\wc2&iso4.lst
convert\wc2&iso4.tab
convert\wc2&iso5.lst
convert\wc2&iso5.tab
convert\wc2&iso6.lst
convert\wc2&iso6.tab
convert\wc2&iso7.lst
convert\wc2&iso7.tab
convert\wc2&iso8.lst
convert\wc2&iso8.tab
convert\wc2&iso9.lst
convert\wc2&iso9.tab
convert\wc2&koi7.lst
convert\wc2&koi7.tab
convert\wc2&macc.lst
convert\wc2&macc.tab
convert\wc2&pcm4.lst
convert\wc2&pcm4.tab
convert\wc2&vsci.lst
convert\wc2&vsci.tab
convert\wc2&wc1a.lst

convert\wc2&wc1a.tab		convert\cjk\wca&wcb.exe
convert\wc2&wc1b.lst		convert\cjk\wca&wcb.tab
convert\wc2&wc1b.tab		convert\cjk\wca2wcb.tab
convert\wc2&wc1c.lst		convert\cjk\wcb2wca.tab
convert\wc2&wc1c.tab		docs\readme.txt
convert\wc2&wc1e.lst		docs\wincalis.dir
convert\wc2&wc1e.tab		docs\wincalis.wci
convert\wc2&wc1g.lst		
convert\wc2&wc1g.tab		**Windows System directory:**
convert\wc2&wc1j.lst		
convert\wc2&wc1j.tab		*c:\windows\system\ctl3d.dll ;3-D dialog boxes
convert\wc2&wc1k.lst		
convert\wc2&wc1k.tab		
convert\wc2&wc1v.lst		c:\windows\system\wc03_14a.ttf ;UniEdit fonts
convert\wc2&wc1v.tab		
convert\wincalis.wci		c:\windows\system\wc03_14a.fot
convert\wincalis.dir		c:\windows\system\wc04_14w.ttf
convert\cjk\b5&wcuni.lst	;CJK (Chinese-Japanese-	c:\windows\system\wc04_14w.fot
		c:\windows\system\wc05_14x.ttf
	;Korean)	c:\windows\system\wc05_14x.fot
convert\cjk\b5&wcuni.tab	related	c:\windows\system\wc06_14s.ttf
	conversion	c:\windows\system\wc06_14s.fot
		c:\windows\system\wc06_15b.ttf
convert\cjk\cmp2smp.lst	;filters (See	c:\windows\system\wc06_15b.fot
convert\cjk\cmp2smp.tab	;Sixteen-bit	c:\windows\system\wc0e_14t.ttf
	Code	c:\windows\system\wc0e_14t.fot
	Conversions)	c:\windows\system\wc10_14g.ttf
convert\cjk\dct2uni.exe		c:\windows\system\wc10_14g.fot
convert\cjk\dct2uni.lst		c:\windows\system\wc20_14t.ttf
convert\cjk\dct2uni.tab		c:\windows\system\wc20_14t.fot
convert\cjk\east2uni.exe		c:\windows\system\wc21_14t.ttf
convert\cjk\eucj&uni.lst		c:\windows\system\wc21_14t.fot
convert\cjk\eucj&uni.tab		c:\windows\system\wc22_14t.ttf
convert\cjk\gb&wcuni.lst		c:\windows\system\wc22_14t.fot
convert\cjk\gb&wcuni.tab		c:\windows\system\wc23_14t.ttf
convert\cjk\ks&wcuni.lst		c:\windows\system\wc23_14t.fot
convert\cjk\ks&wcuni.tab		c:\windows\system\wc25_14t.ttf
convert\cjk\lst2tabd.exe		c:\windows\system\wc25_14t.fot
convert\cjk\lst2tabe.exe		c:\windows\system\wc26_14t.ttf
convert\cjk\lst2tabh.exe		c:\windows\system\wc26_14t.fot
convert\cjk\readmeb5.doc		c:\windows\system\wc27_14t.ttf
convert\cjk\readmegb.doc		c:\windows\system\wc27_14t.fot
convert\cjk\readmesj.doc		c:\windows\system\wc2alphb.fon
convert\cjk\sjis&uni.lst		c:\windows\system\wc2alphr.fon
convert\cjk\sjis&uni.tab		c:\windows\system\wc2cjk1.fon
convert\cjk\sjis2wc2.exe		c:\windows\system\wc2cjk2.fon
convert\cjk\smp2cmp.lst		c:\windows\system\wc2cjk3.fon
convert\cjk\smp2cmp.tab		c:\windows\system\wcc1_14t.ttf
convert\cjk\u10&u11.lst		c:\windows\system\wcc1_14t.fot
convert\cjk\u10&u11.tab		c:\windows\system\wcfe_14s.ttf
convert\cjk\uni2east.exe		c:\windows\system\wcfe_14s.fot
convert\cjk\wc2&east.exe		c:\windows\system\wcfe_15b.ttf
convert\cjk\wc2&east.doc		c:\windows\system\wcfe_15b.fot

Installing UniEdit on a Network

Note: If you are installing UniEdit on a workstation which is running a network copy of Windows, you must have the necessary network privileges to write to the network Windows system directory, even if UniEdit, is being installed on a local drive, since UniEdit must install its fonts and a few other system files in the common Windows directory.

UniEdit performs well in a network setting. But there are certain aspects of UniEdit performance which you should be aware of.

UniEdit Home Directory

It is important to be aware of the drive letter for the "Home" and other directories specified in WINCALIS.INI. See the various sections on changing subdirectory defaults in The Main UniEdit Initialization File WINCALIS.INI. It is sometimes the case with installation on a network server that the DOS drive letter by which that server is known to a local workstation changes. If that happens, UniEdit will not function properly on the local workstation. You will need to bring the drive letters specified in WINCALIS.INI into line with the expectations of the local workstations.

Tips on Installing UniEdit on a Shared Network Drive from a Local Workstation

See the note at the beginning of this topic on necessary network privileges.

Install UniEdit in the desired directory on the shared network drive from one workstation. This will create a complete installation in that workstation's copy of Windows, including the Program Manager group with icons and the fonts installed by UniEdit Setup.

Then in order to give UniEdit icons for the copies of Windows running on the other workstations, you will need to copy the group file WINCALIS.GRP from that workstation's Windows directory to every other workstation's Windows directory, and add the following line to each progman.ini file:

Group8=C:=\WINDOWS\WINCALIS.GRP

(or whatever group number has been created in the original PROGMAN.INI).

You will also need to install the UniEdit fonts which were installed in the Windows System directory (e.g., \windows\system) of the workstation used for the initial Setup (assuming that it is running a local copy of Windows) into the copies of Windows running on the other workstations. Copy all the files wc2*.fon and wc*.ttf into the other windows\system directories.

Then run Control Panel in the Main group, then the Fonts applet, then Add, and change to the \windows\system directory. After logging the fonts in that directory, holding down Ctrl, click on all fonts beginning with "WinCALIS 2.0...". If you alternatively copy the lines from the original machine's win.ini file into the other machines' win.ini files, you will also need to copy the respective "wc*.fot" files from the original machine.

There are two other required files to copy into the Windows System directories:

ctl3d.dll

pxengwin.dll

If the subdirectory and Windows desktop arrangements are the same on all the workstations, you can make a little batch file to do these copying jobs effortlessly. Put all the above files and the batch file, e.g., "TRANSFER.BAT," in a temporary location somewhere on the network, say, the Z: drive. Then run that batch file from each workstation. Your batch file might look like this:

```
copy z:wincalis.grp c:\windows
copy c:\windows\progman.ini c:\windows\progmanb.ini
copy z:progman.ini c:\windows
copy z:wc2*.fon c:\windows\system
copy z:wc*.ttf c:\windows\system
copy z:ctl3d.dll c:\windows\system
copy z:pxengwin.dll c:\windows\system
```

From each workstation, you would type "z:transfer" to perform the needed transfer of files.

How to Uninstall UniEdit

To uninstall UniEdit you will need to remove the directories and files listed below from your machine. You can use the Windows File Manager, or you can use DOS commands. UniEdit also adds fonts to your Windows system, causing Windows to add them to the [fonts] section of your WIN.INI file. UniEdit makes no other modifications directly to any of your Windows initialization files, such as WIN.INI, SYSTEM.INI, etc.

Uninstalling Using the Windows File Manager:

1) Click on the File Manager icon in the Windows 3.1 Main group to start the File Manager.

2) Select the drive which contains the directory which you specified as the UniEdit program directory during setup, e.g., the C: drive.

3) Highlight the UniEdit program directory (by default C:\WINCALIS). Then press the Delete key. When asked, press "Delete All".

   WARNING!! This will delete ALL files in ALL subdirectories under the UniEdit program directory—over 600 files in the case of a full installation. Be sure to exercise caution that you have specified the correct directory, and make backup copies first of any of your own script or document files which you wish to keep.

4) Still using the File Manager, locate the main Windows directory (e.g., C:\WINDOWS) and click on it in order to display the files in it. Delete the following file by highlighting it and pressing the Delete key:

MLPTT.INI

5) Now locate the Windows System directory (e.g., C:\WINDOWS\SYSTEM) and click on it in order to display the files in it. Delete the following files:

   WARNING!! Other applications may have also installed one or more of these files and still need them. If you are in doubt, do NOT delete them. They will not affect the performance of your system.

PXENGWIN.DLL   Paradox Engine library for record-keeping

CTL3D.DLL   3-D Effects in dialog boxes

If the following files were added through the Control Panel, Drivers, Add Driver procedure described in a separate section below, they should also be removed through the Control Panel.

AAPLAY.DLL   Autodesk Animation Player
MCIAAP.DRV   "
MCIAWI.DRV   Gold Disk Animation
MCIAWI1D.DLL   "
MCIAWI1O.DLL   "
MCIAWI1T.TSK   "
MCIMMP.DRV   Microsoft Multimedia Movie Player
MCIPANAS.DRV   Panasonic Videodisc Player
MCIPIONR.DRV   Pioneer Videodisc Player
MCIVISCA.DRV   ViSCA-compatible VCRs
MCIVSONY.DRV   Sony Videodisc Player
MMP.DLL   Microsoft Multimedia Movie Player

6) Now close the Windows File Manager, and click on the icon for the Control Panel, also in the Main group, to open the Windows Control Panel.

7) Click on the Fonts applet.

8) Scroll through the list of "Installed Fonts" until you see the group of fonts described as "WinCALIS 2.0......". All such fonts were added only as part of the UniEdit Setup, and may be safely removed as part of the Uninstall process. Highlight all these font names by either of the following procedures:

a) While holding down the Shift key, click on the first font beginning with "WinCALIS 2.0...", then scroll to the last font beginning with "WinCALIS 2.0..." and click on it.

b) While holding down the Ctrl key, click on each font which begins with "WinCALIS 2.0...".

(This procedure may also be used by any UniEdit user to selectively remove unwanted fonts from his Windows system, in order to conserve disk space, speed up Windows startup time, etc. For example, a UniEdit who deals only with French, German, or Spanish, and is willing to confine himself to the ANSI/Latin1 character set, may choose to remove all of these extra "WinCALIS 2.0" fonts, or at least any of the fonts for Unicode pages with exotic alphabets.)

9) Click on the Remove button. When asked for confirmation, click in the check box which says "Delete font files from disk" and press the "Yes to All" button to uninstall all fonts installed during the UniEdit Setup.

10) Close the Fonts Applet and the Windows Control Panel.

UniEdit is now completely uninstalled.

Chapter 8: UniEdit Conversions

Eight-bit Code Conversions

Conversion of WinCALIS 1.0, WinCALIS 1.1, Text CALIS *.CAL Script Files and other 8-bit Code Text Files to and from WinCALIS 2.x

The stand-alone DOS utility WC2&8BIT.EXE may be used to convert single or multiple text files, including *.CAL script files, from 8-bit (single-byte) codes to WinCALIS 2.x 16-bit Unicode (ISO 10646) code points. It can also perform the reverse function of converting WinCALIS 2.x 16-bit Unicodes to 8-bit (single-byte) codes.

Single WinCALIS 1.1 ANSI (Latin1) text files may also be converted upon opening them in a UniEdit document window by selecting "ANSI/WinCALIS 1.1" in the "List Files of Type" box in the File, Open dialog box, then saving them as UniEdit Unicode files (select the appropriate file type in the File, Save "Save File As Type" box). Or they may be converted by copying them to the Windows clipboard from Notepad, Microsoft Write, or other ANSI text editor, then pasting them into UniEdit and saving them as UniEdit Unicode files.

Non-ANSI 8-bit-coded files may be converted upon opening them in UniEdit by selecting "Other conversions" in the File, Open "List Files of Type" box. Or they may similarly be converted by copying them to the Windows clipboard from another application, then using the Paste As... option in the UniEdit Edit menu to convert codes "on the fly" upon pasting them into UniEdit, and saving them as WinCALIS 2.x Unicode files or another Unicode file type.

Screens of DOS-IBM-extended ASCII coded files, such as generic Text CALIS scripts, will be converted to ANSI in the process of copying from a DOS window to the Windows Clipboard, but this procedure is slow for files longer than a few screens.

WC2&8BIT.EXE can be found in the "CONVERT" subdirectory under your UniEdit program directory if you installed the option "Code Conversion Utilities" during Setup. In order to access this utility complete the following steps:

- Go to a DOS prompt and change directories to the "CONVERT" subdirectory.

- At the prompt type:

    wc2&8bit sourcefile targetfile [/n] [/r] [/t:tablefile]

The /n switch ignores all file headers (such as the WinCALIS/Uniedit Unicode file header), in both reading and writing files.

The direction of conversion is by default from an 8-bit-code source file to a UNICODE target file. The /r switch reverses the direction of the conversion, that is, from a UniEdit 2 source file to an 8-bit-code target file.

The /t:tablefile switch causes the specified table file to be used for conversion. If no table name is specified on the command line, WC2&8BIT.EXE will use any table it finds in the current directory with the name WC2&8BIT.TAB. If a table name is specified, WC2&8BIT.EXE will use that table instead.

There is limited support for DOS wild cards and paths, e.g.

WC2&8BIT *.CAL NEW

where 'new' is a subdirectory under the current directory to which converted versions of all files with matching filename *.CAL will be written. Or:

WC2&8BIT *.TXT *.UNI

where all source files with file extension .TXT are converted to target files with file extension .UNI.

Many conversion tables are supplied in the CONVERT directory. It is also possible to compile your own table file. The table should be a binary file with codes 1-255 (no 0) using the following structure:

bytes 1-2 = UNICODE 1 (Little-endian [Intel integer] order)

bytes 3-4   = UNICODE 2

.

.

bytes 509-510 = UNICODE 255.

A *.TAB table with this structure can be produced with the utilities LST2TAB3.EXE or LST2TAB5.EXE, in the CONVERT subdirectory, from a plain text file *.LST, with one code correspondence per line and with the codes on each line separated by tabs. Use LST2TAB3.EXE if the first code is expressed in decimal and the second code (UNICODE) is expressed in hexadecimal, e.g.,

001(tab)0001

.

.

255(tab)00A0.

Use LST2TAB5.EXE if the first code in each line is expressed in hexadecimal, rather than in decimal, e.g.,

01(tab)0001

.

.

FF(tab)00A0.

If you name your *.TAB file WC2&USER.TAB, it will also be available from within WinCALIS Author and UniEdit, via the File,Open, File,Save and Edit,Paste As... options, using the "User-Supplied 8-bit Char. Set" item in the "Convert From" or "Convert To" dialog box.

Key to supplied conversion tables to be used with WC2&8BIT.EXE, with filename of type WC2&????.TAB:

WC2&ANSI.TAB   ASCII/ANSI(=ISO 8859/1 "Latin1", ECMA-94)

WC2&D437.TAB   ASCII/DOS Code Page 437 (U.S. DOS)/IBM

                              "Extended ASCII" (Generic Text CALIS)

WC2&D850.TAB   DOS Code Page 850 (Multilingual)

WC2&D852.TAB   DOS Code Page 852 (Latin Slavic)

WC2&D855.TAB   IBM Code Page 855 (Cyrillic)

WC2&D857.TAB   DOS Code Page 857 (Turkish)

WC2&D860.TAB	DOS Code Page 860 (Portugal)
WC2&D861.TAB	DOS Code Page 861 (Icelandic)
WC2&D863.TAB	DOS Code Page 863 (Canadian-French)
WC2&D864.TAB	IBM Code Page 864 (Arabic)
WC2&D865.TAB	DOS Code Page 865 (Nordic)
WC2&D866.TAB	DOS Code Page 866 (Cyrillic)
WC2&D869.TAB	IBM Code Page 869 (Greek)
WC2&PCM4.TAB	PC Code Page M4
WC2&1004.TAB	PC Code Page 1004
WC2&1040.TAB	PC Code Page 1040 (Korean Hangul alphabet)
WC2&1041.TAB	PC Code Page 1041 (Japanese Katakana)
WC2&1043.TAB	PC Code Page 1043 (other East Asian SBCS)
WC2&1250.TAB	Microsoft Windows Code Page 1250 (E.Europe)
WC2&1251.TAB	Microsoft Windows Code Page 1251 (Cyrillic)
WC2&1252.TAB	Microsoft Windows Code Page 1252 (Windows ANSI)
WC2&1253.TAB	Microsoft Windows Code Page 1253 (Greek)
WC2&1254.TAB	Microsoft Windows Code Page 1254 (Turkish)
WC2&1255.TAB	Microsoft Windows Code Page 1255 (Hebrew)
WC2&1256.TAB	Microsoft Windows Code Page 1256 (Arabic)
WC2&1257.TAB	Microsoft Windows Code Page 1257 (Baltic)
WC2&KOI7.TAB	KOI-7 (Cyrillic)
WC2&GO74.TAB	GOST 19768-74 (Old KOI-8 Cyrillic)
WC2&BRAV.TAB	Brjabrin's Alternativnyj Variant (Cyrillic)
WC2&BROV.TAB	Brjabrin's Osnovnoj Variant (Cyrillic)
WC2&RSKR.TAB	AATSEEL "REX" (Cyrillic)
WC2&GERD.TAB	"Gerdsooz" encoding for Persian Farsi
WC2&ISCB.TAB	ISCII 1988 Indian Standard (Bengali)
WC2&ISCD.TAB	ISCII 1988 Indian Standard (Devanagari)
WC2&ISCG.TAB	ISCII 1988 Indian Standard (Gurmukhi)

WC2&ISCJ.TAB ISCII 1988 Indian Standard (Gujarati)

WC2&ISCK.TAB   ISCII 1988 Indian Standard (Kannada)

WC2&ISCM.TAB   ISCII 1988 Indian Standard (Malayalam)

WC2&ISCO.TAB   ISCII 1988 Indian Standard (Oriya)

WC2&ISCT.TAB   ISCII 1988 Indian Standard (Tamil)

WC2&ISCU.TAB   ISCII 1988 Indian Standard (Telugu)

WC2&ISO1.TAB   ISO 8859/1 ("Latin1") (same as WC2&ANSI.TAB)

WC2&ISO2.TAB   ISO 8859/2 ("Latin2"-Eastern Europe, etc.)

WC2&ISO3.TAB   ISO 8859/3 ("Latin3"-Eastern Europe, etc.)

WC2&ISO4.TAB   ISO 8859/4 ("Latin4"-Eastern Europe, etc.)

WC2&ISO5.TAB   ISO 8859/5 (=ECMA-113,GOST 19768-87,

           New KOI-8 Cyrillic)

WC2&ISO6.TAB   ISO 8859/6 (=ECMA-114,ASMO 449)

           (Arabic, Farsi, etc.)

WC2&ISO7.TAB   ISO 8859/7 (=ECMA-118,ELOT 928) (Greek)

WC2&ISO8.TAB   ISO 8859/8 (Hebrew)

WC2&ISO9.TAB   ISO 8859/9

WC2&ISOA.TAB   ISO 6438 (=DIN 31625) (African Extended Latin)

WC2&ISOG.TAB   ISO 5428-1980 (Bibliographic Greek)

WC2&MACA.TAB   Macintosh Arabic Character Set

WC2&MACC.TAB   Macintosh Cyrillic Character Set

WC2&MACE.TAB   Macintosh Central Europe Character Set

WC2&MACG.TAB   Macintosh Greek Character Set

WC2&MACI.TAB   Macintosh Icelandic Character Set

WC2&MACL.TAB   Macintosh Latin (Roman) Character Set

WC2&MACN.TAB   Macintosh Croatian Character Set

WC2&MACR.TAB   Macintosh Romanian Character Set

WC2&MACS.TAB   Macintosh (Adobe) Symbol Character Set

WC2&MACT.TAB   Macintosh Turkish Character Set

WC2&MACU.TAB	Macintosh Ukrainian Character Set
WC2&TIS6.TAB	Thai Industrial Standard 620-2529
WC2&USER.TAB	User-supplied 8-bit Character Set
WC2&VSCI.TAB	VISCII Vietnamese Standard
WC2&WC1B.TAB	WINCALIS 1.0 BASE.FON (English,French, German,Italian,Portuguese,Spanish)
WC2&WC1C.TAB	WINCALIS 1.0 CYRILLIC.FON (Russian)
WC2&WC1E.TAB	WINCALIS 1.0 EEUR.FON (Polish, Hungarian, generic East Europe, etc.)
WC2&WC1G.TAB	WINCALIS 1.0 PGREEK.FON (Greek)
WC2&WC1J.TAB	WINCALIS 1.0 HIRAGANA.FON (Japanese hiragana)
WC2&WC1K.TAB	WINCALIS 1.0 KATAKANA.FON (Japanese katakana)
WC2&WC1V.TAB	WINCALIS 1.0 VIETNAME.FON (Vietnamese)
WC2&CALA.TAB	Amharic Text CALIS
WC2&CALE.TAB	East European Text CALIS
WC2&CALG.TAB	Greek Text CALIS
WC2&CALH.TAB	Hebrew Text CALIS
WC2&CALP.TAB	Portuguese Text CALIS
WC2&CALR.TAB	Russian Text CALIS

The files WC2&D437.LST (and LST2TAB3.EXE) and WC2&ISO1.LST (and LST2TAB5.EXE) can serve as templates from which to create more conversion tables for use with WC2&8BIT.EXE.

For conversion between various 2-byte coding standards (e.g., for Chinese, Japanese, Korean, Unicode 1.0 or 1.1) and UniEdit Unicode 2.0, use the following separate utilities (requires UniEdit Workstation Setup CJK Conversions Disk):

- WC2&16BT.EXE for 16-bit fixed-width coded files

- DCT2UNI.EXE for Duke Chinese Typist 2.0 coded files

- WC2&EAST.EXE for JIS, Shift-JIS, EUC, Big5, GB, KS, and other 7-bit modal or mixed 16-bit and 8-bit coded files.

- WC2&24BT.EXE for 24-bit fixed-width coded files such as CCCII or EACC/ANSI Z39.64

- UCS2&UTF.EXE for UTF-7 and UTF-8 coded Unicode files.

Sixteen-bit Code Conversions

Conversion of East Asian double-byte files consisting of modal 7-bit or mixed 16-bit and 8-bit codes to and from UniEdit

The stand-alone DOS utility WC2&EAST.EXE may be used to convert single or multiple text files from modal 7-bit or mixed 16-bit/8-bit coding standards such as JIS, SHIFT-JIS, KS, Big5, or GB code to UniEdit 16-bit Unicode (ISO 10646) files. It can also perform the reverse function of converting UniEdit 16-bit Unicodes to modal 7-bit or mixed 16-bit and 8-bit coding standards. For modal 7-bit coded files, use the "EUC" coding table (e.g.,for Japanese, use the table EUCJ&UNI.TAB).

Single files can be converted by opening or saving them in UniEdit and selecting "Other conversions" in the "List files of type" or "Save file as type" box. Or they may be converted by copying them to the Windows clipboard from the source application, then pasting them into UniEdit using the Paste As... option in the Edit menu to convert codes "on the fly" during pasting, and saving them as UniEdit Unicode files.

WC2&EAST.EXE can be found in the CONVERT subdirectory under your UniEdit program directory. In order to access this utility, do the following:

- Go to a DOS prompt and change to "CONVERT" subdirectory

- At a DOS prompt type the command line:

    wc2&east sourcefile targetfile [/n] [/r] [/m] [/t:tablefile]

The /n switch ignores all file headers (such as the WinCALIS/Uniedit Unicode file header), in both reading and writing files.

The direction of conversion is from a modal 7-bit or mixed 16-bit/8-bit source file to a UniEdit Unicode target file by default. The /r switch reverses the direction of the conversion, that is, from a UniEdit source file to a modal 7-bit or mixed 16-bit/8-bit-code target file.

The /m switch, when used in conjunction with the /r switch, specifies conversion from UniEdit Unicode source file to a modal 7-bit target file (i.e., a file in which all codes are 7-bit codes, with language shifts marked by Escape codes).

The /t:tablefile switch causes the specified table file to be used for conversion. If no table name is specified on the command line, WC2&EAST.EXE will use any table it finds in the current directory with the name WC2&EAST.TAB. If a table name is specified, WC2&EAST.EXE will use that table instead.

There is limited support for DOS wild cards and paths ,e.g.,

     WC2&EAST *.GB NEW

where 'new' is a subdirectory under the current directory to which converted versions of all files with matching filename *.GB will be written. Or:

     WC2&EAST *.GB *.UNI

where all source files with file extension .GB are converted to target files with file extension .UNI.

Several conversion tables are supplied in the CONVERT subdirectory. It is also possible to compile your own table file. The table should be a binary file with the following structure:

bytes 1-2   = Source (GB, Big5, etc.) code 1 (Little-endian byte-reversed)

bytes 3-4   = Unicode 1 (Big-endian)

bytes 5-6   = Source (GB, Big5,etc.) code 2

bytes 7-8   = Unicode 2

.

.

A *.TAB table with this structure can be produced with the utility SJIS2WC2.EXE, in the CONVERT subdirectory, from a plain text file *.LST, with one code correspondence per line and with the codes on each line separated by tabs, e.g.,

8140(tab)3000

.

.

EAA4(tab)7199.

Usage: SJIS2WC2 LISTFILE.LST

Key to supplied conversion tables to be used with WC2&EAST.EXE, (rename as WC2&EAST.TAB or type filename after /t:tablefile switch):

B5&WCUNI.TAB	Big-5 codes (Taiwan Chinese)
EUCJ&UNI.TAB	Extended Unix Codes (Japanese)
	(use this table for JIS modal 7-bit conversion)
GB&WCUNI.TAB	GB-2312-80 codes (simple-style Chinese)
KSJH&UN1.TAB	Korean Johab codes --> Unicode 1.1
KSJH&UN2.TAB	Korean Johab codes --> Unicode 2.0
KSWS&UN1.TAB	KS-C5601 Korean Standard Wansung --> Unicode 1.1
KSWS&UN2.TAB	KS-C5601 Korean Standard Wansung --> Unicode 2.0
KSUH&UN1.TAB	Korean "Unified Hangul" codes --> Unicode 1.1
KSUH&UN2.TAB	Korean "Unified Hangul" codes --> Unicode 2.0
SJIS&UNI.TAB	Shift-JIS codes (Japanese)

The corresponding files B5&WCUNI.LST, EUCJ&UNI.LST, GB&WCUNI.LST, KSJH&UN1.LST, etc. can serve as templates from which to create more conversion tables for use with WC2&EAST.EXE.

For conversion between various single-byte (8-bit) coding standards, use the parallel utility WC2&8BIT.EXE in the CONVERT subdirectory.

Other DBCS (double-byte character set) utilities are explained below.

Conversion of fixed-width double-byte (16-bit) files to and from UniEdit

Conversion of Duke Chinese Typist files

Use DCT2UNI.EXE with the table DCT2UNI.TAB to convert Duke Chinese Typist 2.0 coded files to Unicode.

In order to access this utility, do the following:

- Go to a DOS prompt and change to "CONVERT" subdirectory

- At a DOS prompt type the command line:

    dct2uni sourcefile targetfile [/n] [/r] [/t:tablefile]

The /n switch ignores all file headers (such as the WinCALIS/Uniedit Unicode file header), in both reading and writing files.

The direction of conversion is from a DCT source file to a UniEdit Unicode target file by default. The /r switch reverses the direction of the conversion, that is, from a UniEdit source file to a DCT target file.

The /t:tablefile switch causes the specified table file to be used for conversion. If no table name is specified on the command line, DCT2UNI.EXE will use any table it finds in the current directory with the name DCT2UNI.TAB. If a table name is specified, DCT2UNI.EXE will use that table instead.

There is limited support for DOS wild cards and paths, e.g.,

    DCT2UNI *.DC2 NEW

where 'new' is a subdirectory under the current directory to which converted versions of all files with matching filename *.DC2 will be written. Or:

DCT2UNI *.DC2 *.UNI

where all source files with file extension .DC2 are converted to target files with file extension .UNI.

Conversion of Miscellaneous Fixed-width 16-bit files

Types of fixed-width 16-bit codes you may wish to convert include the following:

- conversion of text files from various other East Asian word-processing programs

- conversion of older or non-standard Unicode-format files, e.g., of Unicode 1.0 or 1.1-coded files to Unicode 2.0

  Use the table U10&U11.TAB and/or the table U11&U20.TAB, or modified versions of them (see below).

- conversion between simple-style Chinese and complex-style Chinese.

  Use the tables SMP2CMP.TAB or CMP2SMP.TAB.

Note: Since a simple-style-->complex-style Chinese conversion is one-->many and no intelligence is employed, it will require post-editing by someone familiar with proper traditional usage.

- conversion of files with Korean KS Hanja Compatibility Codes to "pure Unicode"

  Use the table KORCOMP.TAB.

- conversion of files in the WinCALIS/UniEdit full code set (including private codes, "compatibility" character codes, and Arabic-script Presentation Characters) to "pure Unicode"

  Use the table WC&PURUN.TAB.

Note: UniEdit maintains some distinctions among Han characters which are not maintained in the Unicode Han character set, by mapping one of the glyphic forms to a code in part of the private-use code space, in the range F000-F1FF. It is our desire to maintain distinctions of significance to language teachers and students within Unicode plain text (without formatting information) and without "escape codes" in the character stream, indicating that the following character is to be understood as "Japanese," or "Korean," etc. This affects approximately 400 Han characters, usually glyphic variations between the Chinese and Japanese character sets or the "standard" P.R.C. and Taiwan character sets, such as the distinct simple-style Chinese and Japanese *jôyôkanji* glyphs for U+4E0E, virtually all characters with the "silk radical" (Kangxi radical 120), and most characters with the "bone" radical (Kangxi

radical 188). Here is a sample of the code equivalences (see the file "wc&purun.lst" in the ..\convert subdirectory for the complete list):

UniEdit Unicode	"Pure" Unicode
F001	4E0E
F002	50B2
F003	533E
F004	5448
F005	548E
F006	55F7
...	

When you convert files to "pure" Unicode, these distinctions are lost and cannot be recovered.

In addition, certain other composite characters are also mapped to private-use codes, such as Cyrillic vowels with acute accents (F700-F715), Roman letters with exotic diacritic combinations (F600-F6FF), etc. These characters cannot yet be decomposed automatically during conversion into their component parts.

Single files can be converted by opening or saving them in UniEdit and selecting "Other conversions" in the "List files of type" or "Save file as type" box. Or they may be converted by copying them to the Windows clipboard from the source application, then pasting them into UniEdit using the Paste As... option in the Edit menu to convert codes "on the fly" during pasting, and saving them as UniEdit Unicode files.

To convert many files at once, you may use the DOS utility WC2&16BT.EXE at the DOS prompt.

WC2&16BT.EXE can be found in the CONVERT subdirectory under your UniEdit program directory. In order to access this utility, do the following:

- Go to a DOS prompt and change to "CONVERT" subdirectory

- At a DOS prompt type the command line:

  wc2&16bt sourcefile targetfile [/n] [/r] [/t:tablefile]

The /n switch ignores all file headers (such as the WinCALIS/UniEdit Unicode file header), in both reading and writing files.

The direction of conversion is from a non-Unicode 2.0 16-bit source file to a UniEdit Unicode target file by default. The /r switch reverses the direction of the conversion, that is, from a UniEdit source file to a non-Unicode 2.0 16-bit target file.

The /t:tablefile switch causes the specified table file to be used for conversion. If no table name is specified on the command line, WC2&16BT.EXE will use any table it finds in the current directory with the name WC2&16BT.TAB. If a table name is specified, WC2&16BT.EXE will use that table instead.

There is limited support for DOS wild cards and paths ,e.g.,

      WC2&16BT *.U10 NEW

where 'new' is a subdirectory under the current directory to which converted versions of all files with matching filename *.GB will be written. Or:

      WC2&16BT *.U11 *.U20

where all source files with file extension .U11 are converted to target files with file extension .U20.

You can modify any of the binary *.TAB tables above or construct a new one with the included utility LST2TABH.EXE by modifying or imitating the *.LST text source files of the above *.TAB files. Each *.LST file has one code correspondence per line, with the codes on each line separated by tabs, e.g.,

4E26(tab)5E76

.

.

EAA4(tab)7199.

Usage: LST2TABH LISTFILE.LST

Twenty-four-bit Code Conversions

Twenty-four-bit (triple-byte) fixed-width coding standards include CCCII and EACC/ANSI Z39.64. You may convert files using these codes to UniEdit files with 16-bit Unicode UCS-2 character codes in one of the following ways.

Single files can be converted by opening or saving them in UniEdit and selecting "Other conversions" in the "List files of type" or "Save file as type" box. Or they may be converted by copying them to the Windows clipboard from the source application, then pasting them into UniEdit using the Paste As... option in the Edit menu to convert codes "on the fly" during pasting, and saving them as UniEdit Unicode files.

To convert many files in "batch mode" rapidly at once, you may use the DOS utility WC2&24BT.EXE at the DOS prompt.

WC2&24BT.EXE can be found in the CONVERT subdirectory under your UniEdit program directory. In order to access this utility, do the following:

- Go to a DOS prompt and change to "CONVERT" subdirectory

- At a DOS prompt type the command line:

    wc2&24bt sourcefile targetfile [/n] [/r] [/t:tablefile]

The /n switch ignores all file headers (such as the WinCALIS/Uniedit Unicode file header), in both reading and writing files.

The direction of conversion is from a 24-bit (3-byte) coded source file to a UniEdit Unicode target file by default. The /r switch reverses the direction of the conversion, that is, from a UniEdit source file to a 24-bit target file.

The /t:tablefile switch causes the specified table file to be used for conversion. If no table name is specified on the command line, WC2&24BT.EXE will use any table it finds in the current directory with the name WC2&24BT.TAB. If a table name is specified, WC2&24BT.EXE will use that table instead.

There is limited support for DOS wild cards and paths ,e.g.,

    WC2&24BT *.EAC NEW

where 'new' is a subdirectory under the current directory to which converted versions of all files with matching filename *.EAC will be written. Or:

WC2&24BT *.EAC *.UNI

where all source files with file extension .EAC are converted to target files with file extension .UNI.

Key to supplied conversion tables to be used with WC2&24BT.EXE, (rename as WC2&24BT.TAB or type filename after /t:tablefile switch):

CCCI&UNI.TAB   CCCII codes (Taiwan Chinese)

EACC&UNI.TAB   EACC/ANSI Z39.64 (REACC) codes (library MARC records, such as Research Libraries Network, Library of Congress etc.)

You can modify any of the binary *.TAB tables above or construct a new one with the included utility LST2TAB6.EXE by modifying or imitating the *.LST text source files of the above *.TAB files. Each *.LST file has one code correspondence per line, with the codes on each line separated by tabs, e.g.,

213021(tab)4E00

.

.

70775D(tab)9B0F.

Usage: LST2TAB6 LISTFILE.LST

UTF-7 and UTF-8 Conversions

You may convert Unicode files with codes in the "Universal Coded Character Set Transformation Format" UTF-7 (7-bit codes) or UTF-8 (8-bit codes) to UniEdit files with 16-bit UCS-2 character codes in one of the following ways.

Single files can be converted by opening or saving them in UniEdit and selecting "UTF-7/UTF-8" in the "List Files of Type" or "Save File as Type" box.

To convert many files in "batch mode" rapidly at once, you may use the DOS utility UCS2&UTF.EXE at the DOS prompt.

UCS2&UTF.EXE can be found in the CONVERT subdirectory under your UniEdit program directory. In order to access this utility, do the following:

- Go to a DOS prompt and change to "CONVERT" subdirectory

- At a DOS prompt type the command line:

    ucs2&utf sourcefile targetfile [/n] [/r] [/7]

The /n switch ignores all file headers (such as the WinCALIS/Uniedit Unicode file header), in both reading and writing files.

The direction of conversion is from a Unicode UTF-7 or UTF-8 coded source file to a UniEdit Unicode UCS-2 coded target file by default. The /r switch reverses the direction of the conversion, that is, from a UniEdit source file to a UTF-7 or UTF-8 target file.

The /7 switch, used in conjunction with the /r switch, converts from UCS-2 to UTF-7. By default conversion will be to UTF-8.

There is limited support for DOS wild cards and paths ,e.g.,

    UCS2&UTF *.UTF NEW

where 'new' is a subdirectory under the current directory to which converted versions of all files with matching filename *.UTF will be written. Or:

    UCS2&UTF *.UTF *.UCS

where all source files with file extension .UTF are converted to target files with file extension .UCS.

Index