

Java Application: CharMap

Written by: Keith Fenske, <http://www.psc-consulting.ca/fenske/>


First version: Monday, 19 May 2008

Document revised: Saturday, 9 May 2009

Copyright © 2008 by Keith Fenske. Released under the GNU General Public License (GPL).

Description

CharMap is a Java 5.0 graphical (GUI) application to display Unicode characters or glyphs in text fonts, and copy those characters to the system clipboard. Its major purpose is as a visual accessory for word processors such as Microsoft Word. The “character map” utility that comes with Windows suffers from several problems. This Java application can be resized, for text and the program window, which is important in many languages. Features are limited to make the application faster and simpler to use. A single click adds a character to the sample text, and the sample text is automatically copied to the system clipboard on each click.

Menu	font name, size					caption
	!	"	#	\$	%	
&	'	()	*	+	
,	-	.	/	0	1	
2	3	4	5	6	7	
8	9	:	;	<	=	
>	?	@	A	B	C	
Clear	sample text					Copy All

You may choose the font to be displayed and the size of the characters or glyphs. (Glyphs are bits and pieces that a font combines to produce the characters you see. In most cases, one character maps to one glyph.) You may edit the sample text, erase it with the “Clear” button, or copy it to the system clipboard with the “Copy All” button. Paste the text into your word processor in the normal manner, which is usually a Control-V key combination. Editing the sample text and pressing the Enter key also copies to the clipboard. Specific characters can be

copied from the sample text by selection and with the usual Control-C combination. More characters are available via the scroll bar on the right. A description is shown in the “caption” field when characters have a particular name or meaning. Common readings or sounds are given for Chinese, Japanese, and Korean characters. Cantonese is prefixed with “C”, Japanese “Kun” with “J”, Korean with “K”, Mandarin with “M”, and Sino-Japanese “On” with “S”. An English translation of CJK character definitions would have been more amusing but less practical.

Keyboard shortcuts are provided to mimic the scroll bar: the Control-Home key combination goes to the very first character, Control-End goes to the last character, Page Down and Page Up scroll one screen at a time, and the arrow keys scroll one line at a time. You need to combine the End and Home keys with the Control (Ctrl) key when the sample text has keyboard focus. The F1 key is the only helpful undocumented feature.

GNU General Public License (GPL)

CharMap is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License or (at your option) any later version. This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY, without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details. You should have received a copy of the GNU General Public License along with this program. If not, see the <http://www.gnu.org/licenses/> web page.

Installation

You must have the Java run-time environment (JRE) installed on your computer. CharMap was developed with Java 5.0 (1.5) and should run on later versions. It won't run on earlier versions such as Java 1.4 without being edited and recompiled. For Macintosh computers, the version of Java is determined by your version of MacOS. For Windows, Linux, and Solaris, you can download the JRE from Sun Microsystems:

Sun Java

JRE for end users: <http://www.java.com/getjava/>

SDK for programmers: <http://developers.sun.com/downloads/>

IDE for programmers: <http://www.netbeans.org/>

Once Java is installed, you need to put the program files for CharMap into a folder (directory) on your hard drive. The name of the folder and the location are your choice, except it is easier if the name does not include spaces. Assume that files will go into a C:\JAVA folder. Then create the folder and unpack the Java *.class files into this folder (if you received the program as a ZIP file). The files look something like this:

CharMap4.class (37 KB, executable program)
CharMap4.doc (43 KB, this documentation in Microsoft Word format)
CharMap4.gif (18 KB, sample program image)
CharMap4.ico (14 KB, icon for Windows)
CharMap4.java (120 KB, source code)
CharMap4.pdf (80 KB, this documentation in Adobe Acrobat format)
CharMap4.txt (2.3 MB, data file with character names)
CharMap4Grid.class (10 KB, helper class for main program)
CharMap4User.class (1 KB)
CharMapParse1.java (4 KB, extra program to parse Unicode character data)
CharMapParse5.java (19 KB)
GnuPublicLicense3.txt (36 KB, legal notice)
RunJavaPrograms.pdf (87 KB, more notes about running Java)

Some Apple Macintosh computers have older versions of Java that are difficult to update. Included in this distribution file is the same program re-compiled for Java 1.4 (commonly used from about 2003 to 2007) instead of Java 5.0 (about 2005 to 2008). The significant differences are that extended Unicode characters higher than U+FFFF can not be displayed, and that Java 1.4 was known for incorrect rendering of non-text or symbol fonts. On the other hand, since the program only has to enumerate the first 65,536 characters in any font, it does run faster, even on computers with Java 5.0 or later. Very little software currently makes use of the extended Unicode range.

CharMap4J14.class (37 KB, executable program for Java 1.4)
CharMap4J14.java (121 KB, source code for Java 1.4)
CharMap4J14Grid.class (10 KB)
CharMap4J14User.class (1 KB)

To run the program on Windows, start a DOS command prompt, which is Start button, Programs, Accessories, Command Prompt on Windows 2000/XP. Change to the folder with the program files and run the program with a “java” command:

```
c:  
cd \java  
java CharMap4
```

The program name “CharMap4” must appear exactly as shown; uppercase and lowercase letters are different in Java names. Should you find this program to be popular, you can create a Start menu item or desktop shortcut on Windows 2000/XP with a target of “java.exe CharMap4” starting in the “c:\java” folder. One complication may arise when trying to run this program. Java looks for an environment variable called CLASSPATH. If it finds this variable, then that is a list of folders where it looks for *.class files. It won’t look anywhere else, not even in the

current directory, unless the path contains “.” as one of the choices. The symptom is an error message that says:

```
Exception in thread "main" java.lang.NoClassDefFoundError: CharMap4
```

To find out if your system has a CLASSPATH variable defined, type the following command in a DOS window:

```
set CLASSPATH
```

To temporarily change the CLASSPATH variable to the current directory, use the following command line:

```
java -cp . CharMap4
```

To permanently change the CLASSPATH, you must find where it is being set. This may be in an old AUTOEXEC.* file in the root directory of your system disk (usually the C:\ folder), or it may be in Control Panel, System, Advanced, Environment Variables on Windows 2000/XP.

Removal or Uninstall

To remove this program from your computer, delete the installation files listed above. If the folder that contained the files is now empty, you may also delete the folder ... if you created the folder, of course, not the system. If you created desktop shortcuts or Start menu items, then delete those too. There are no hidden configuration or preference files, and no information is stored in the Windows system registry. You don't need an “uninstall” program.

Graphical Versus Console Application

The Java command line may contain options for the initial display font, the size of the text, and the window position. See the “-?” option for a help summary:

```
java CharMap4 -?
```

The command line has more options than are visible in the graphical interface. An option such as -u14 or -u16 is recommended because the default Java font is too small.

Restrictions and Limitations

Which fonts will work with this program depends upon the operating system and version of the Java run-time environment. Java 5.0 on Windows 2000/XP will show installed TrueType fonts, that is, fonts that have been added with the Control Panel, Fonts icon. (Temporary fonts are not

shown if opened with the Windows Font Viewer by double clicking on a font file name.) If you think this program is not working correctly on your computer, then “Lucida Console” is a good font for testing the spacing and positioning, because its glyphs are tightly packed. Version 4 of CharMap supports extended Unicode (up to 1,114,112 characters) and is noticeably slower than version 3, which only supports the standard range of 65,536 characters. Version 4 also tends to run out of memory for very large fonts; see the -Xmx option on the Java command line.

This program contains character data from the Unicode Consortium; please visit their web site at <http://www.unicode.org/> for more information. Korean character names were converted from Korean standards document KS X 1001:2002 with the title “Hangeul Syllables in Unicode 4.0” and dated 25 March 2004. A plain text file called “CharMap4.txt” is expected to be in the current working directory with mappings from Unicode character numbers to caption strings. You may edit this file to produce whatever captions you wish. Please read comments in the file for further instructions.

file: CharMap4.doc 2009-05-09