Adding Supplementary Plane Characters to a Font

David J. Perry
August 31, 2003

I have used two methods to add characters in the supplementary planes to a font: Microsoft VOLT and FontLab 4.6 (or higher). I am also told that the Font Creator Program can do this, but have not tried it myself. This information is current as of August 2003.

I. Using VOLT

VOLT is a free tool available from www.microsoft.com/typography. It is mainly designed to add OpenType features to a font but can also be used to add characters in the supplementary planes, which some font editors cannot yet do. This article focuses on adding the characters beyond the BMP. Unfortunately, VOLT is not available for any platform other than Windows.

Important background information: VOLT uses glyph IDs as the basis for all its operations. A glyph ID is simply a number that tells where in the font the character is found (first, tenth, fiftieth, whatever). It has nothing to do with a character's Unicode value or any other meaningful characteristic; it just uniquely identifies each glyph in the font. Such glyph IDs are a part of any TrueType font and are created for you by your font editor. VOLT also stores information in special additional tables that are not part of a standard TT font; these tables allow you to save the font, test it, and return to edit your work. These tables are removed as the last step in creating a finished font.

If you are using FontLab, you need to understand under what circumstances glyph IDs are changed. FL provides many different ways to display the characters in a font: by Unicode value, by name, by glyph ID, etc. You can switch the display around as much as you like; the position of a character in the actual font database (and therefore the glyph ID) is not altered unless you do one of two things. 1) While in glyph index mode (entered by clicking on the little black 1-2-3 icon on the FL toolbar), you highlight one or more characters and drag them to a new position. 2) You right-click on a character and choose Sort . . .

Here are the steps to follow to add supplementary characters.

1. Using FontLab or another editor, set up the font with slots for all the characters needed in the desired order. It's possible to change the order of characters later, but this is not a good idea, especially in a large font (adding additional characters onto the end is not a problem). Outlines can easily be edited at any time as long as the order of glyphs is not altered (thereby altering their glyph IDs).

2. In FontLab it doesn't matter what Unicode values you give the characters that you want to have in Plane 1 or 2. You can leave these blank or use Private Use Area values if you want to test the characters before putting them in their final positions. You will have to enter these Plane 1 or 2 values at a later point in the process (see below).

3. Make sure that the Non-Plane 0 option is checked in FontLab's Encoding and Unicode/Unicode Ranges dialog box. Then generate the font file.

4. Open the font in VOLT. Go to Tools/Options and tell VOLT to generate an additional cmap for Platform = 3, Encoding= 10, Format = 12; this is required for a TrueType font to support supplementary characters under Windows. If you do everything else, but omit this step, your supplementary characters will not work. Save the font under a new name. I like to add the letters *volt* to the font's name; this reminds me that this version of the font has been processed by VOLT and that I will need to remove the special VOLT tables before I produce the final font for distribution.

5. When VOLT opens a font, it ignores the character names in the font and gives generic names of its own (Glyph01, Glyph02, etc.). Assuming that you want to use the names you specified in FontLab, proceed as follows. From the menu bar choose Import/Import PostScript names. Specify the range of values (usually zero to the last glyph in the font). You will now have your names back; save the font.

6. Now comes the tedious part. You must manually enter the Unicode values for all the supplementary characters, and you must do it in the form of surrogate pairs, rather than as a single UTF-32 value. If you know the scalar values and need to determine the surrogate pairs, you can use the convenient calculator provided by Michael Kaplan at http://www.trigeminal.com/16to32AndBack.asp.

7. Click the Edit Glyphs button on the toolbar, or choose Tools/Edit Glyphs from the menu. Select each character destined for Plane 1 or 2 and enter the surrogate pair (no space between the two halves) in the dialog box at the top of the window. Click on the next character to set the value you have typed. Keep going until all are done. Save the font.

8. At this point you can test the font by installing it into Windows. The supplementary characters should be accessible in Word XP or WordPad XP by typing their scalar value followed by ALT-x. You may need to set the font that contains the supplementary characters before entering them; Word sometimes won't display such characters if you change the font afterwards.

9. If you need to do further editing of the character outlines, proceed as follows. In VOLT, export the glyph data to a text file (Export/Export Glyph Data). This saves the names and Unicode values into a separate file. Then open your FontLab database and perform your edits; make sure that you do not alter the order of glyphs in the font (thereby altering the glyph IDs). When finished, generate a new version of the font. Open it in VOLT and import the glyph data that you exported in the previous step. You will also need to tell VOLT generate the format 12 cmap as explained in step 4 above.

10. To add additional characters, export the glyph data as described above. Open your FL database and add the characters you want, then generate a new version of the font. Open this font in VOLT and import the glyph data for the appropriate glyphs. For example, if you had a font with 500 glyphs and added 20 more, you would import the data for glyphs

0 to 499. You can then import the names for the new glyphs (glyph IDs 500 to 519) and manually add the Unicode values as described above, if the new characters are intended to go in one of the supplementary planes. Again, make sure that you add your new characters *at the end* of your font file, in order to preserve order of glyph IDs.

11. When everything is complete, you should "ship" the font. This removes the special tables that VOLT uses to save its work in a form that you can return to and edit again. Choose File/Ship font and be sure to save under a different name.

II. Using FontLab 4.6

FontLab 4.6, released in August 2003, can generate a font with supplementary characters. The user's manual does not mention this; I learned about it from a posting on the FontLab Group at MSN. This message also indicated that both Windows and Mac versions of FontLab can do this, which is useful since VOLT is only available for Windows. (I have experimented only with the Windows version of FL.)

Apparently part of the FontLab code has not yet been rewritten to allow direct entry of Unicode values greater than U+FFFF, a limit which has been in place since the first versions of FontLab. However, the FL developers have added an algorithm, beginning with version 4.6, that will take any characters whose names follow the *uXXXXX* format and put them into the font with their correct Unicode values, as well as writing the necessary cmap table, when a font is generated from the FL database. Since it is best to use the *uXXXXX* format for names anyway, this is not a significant limitation. Here are the steps to follow.

1. Create your glyph outlines and name each one based on the Plane 1 or 2 Unicode value that you want it to have. For example, a character with the Unicode value U+10140 would be named *u10140*. You may leave the Unicode value field blank and use only the name, or you may enter a Unicode value such as a PUA codepoint. If you enter a Unicode value, then the glyph will have two Unicode values in the font you generate. This can be useful since many applications don't yet support supplementary characters; you could use the PUA values for such applications and the real Plane 1 values for those that do.

2. Make sure that the Non-Plane 0 option is checked in FontLab's Encoding and Unicode/Unicode Ranges dialog box. Then generate the font file.

IMPORTANT NOTE. Version 4.6 of FL, with which I recently experimented, has a bug. Instead of creating a cmap with Platform 3, Encoding 10, Format 12, it makes one that has Format 12.256, 12.244, or a similar number. This causes both Word XP and WordPad to crash when you try to use any of the supplementary characters. I opened the font in VOLT and saw this strange format number. Changing it to plain 12 solved the problem. Presumably this bug will be fixed in the next release of FontLab. It is still advantageous to use FL since you don't have to retype all the Unicode values as surrogate pairs in VOLT, you just have to fix this one number.