DR 09-0040 — WML/DML: Complex scripts

Proposed Solution

Include the following content in a new section under §17.3:

For each Unicode character in a WordprocessingML document, the Office Open XML specification specifies that its font face can be any from any of four font "slots": ascii, hAnsi, cs, or eastAsia. This document describes how Microsoft Office Word determines the font slot to use.

For each Unicode character in a WordprocessingML document, the following font resolution takes place:

- If the run properties contain the *cs* or *rtl* element, then the *cs* (or *csTheme* if defined) attribute is used, regardless of the Unicode character values of the run's content.
- If the *eastAsia* (or *eastAsiaTheme* if defined) attribute's value is "Times New Roman" and the *ascii* (or *asciiTheme*) and *hAnsi* (or *hAnsiTheme*) attributes are equal, then the *ascii* (or *asciiTheme* if defined) attribute is used.
- Otherwise, the following table is used. For all ranges not explicitly called out below, the *hAnsi* (or *hAnsiTheme* if defined) attribute shall be used.

Unicode Block	Range	Classification
Basic Latin	0000 – 007F	*ascii*
Latin-1 Supplement	00A0 – 00FF	*hAnsi*, with the following exceptions: • If *hint* is *eastAsia*, the following characters use *eastAsia* (or *eastAsiaTheme* if defined): A1, A4, A7 – A8, AA, AD, AF, B0 – B4, B6 – BA, BC – BF, D7, F7 • If *hint* is *eastAsia* and the language of the run is either Chinese Traditional or Chinese Simplified, the following characters use *eastAsia* (or *eastAsiaTheme* if defined): E0 – E1, E8 – EA, EC – ED, F2 – F3, F9 – FA, FC
Latin Extended-A	0100 – 017F	*hAnsi*, with the following exception: • If *hint* is *eastAsia*, and the language of the run is either Chinese Traditional or Chinese Simplified, or the character set of the *eastAsia* (or *eastAsiaTheme* if defined) font is Chinese5 or GB2312 then *eastAsia* font is used.
Latin Extended-B	0180 – 024F	*hAnsi*, with the following exception: • If *hint* is *eastAsia*, and the language of the run is either Chinese Traditional or Chinese Simplified, or the character set of the *eastAsia* (or *eastAsiaTheme* if defined) font is Chinese5 or GB2312 then *eastAsia* font is used.
IPA Extensions	0250 – 02AF	*hAnsi*, with the following exception:

		• If *hint* is *eastAsia*, and the language of the run is either Chinese Traditional or Chinese Simplified, or the character set of the *eastAsia* (or *eastAsiaTheme* if defined) font is Chinese5 or GB2312 then *eastAsia* font is used.
Spacing Modifier Letters	02B0 – 02FF	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Combining Diacritical Marks	0300 – 036F	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Greek	0370 – 03CF	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Cyrillic	0400 – 04FF	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Hebrew	0590 – 05FF	*ascii*
Arabic	0600 – 06FF	*ascii*
Syriac	0700 – 074F	*ascii*
Arabic Supplement	0750 – 077F	*ascii*
Thaana	0780 – 07BF	*ascii*
Hangul Jamo	1100 – 11FF	*eastAsia*
Latin Extended Additional	1E00 – 1EFF	*hAnsi*, with the following exception: • If the *hint* is *eastAsia* and the language of the run is Chinese Traditional or Chinese Simplified then that is used.
Greek Extended	1F00 – 1FFF	*hAnsi*
General Punctuation	2000 – 206F	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Superscripts and Subscripts	2070 – 209F	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Currency Symbols	20A0 – 20CF	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Combining Diacritical Marks for Symbols	20D0 – 20FF	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Letter-like Symbols	2100 – 214F	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Number Forms	2150 – 218F	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Arrows	2190 – 21FF	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Mathematical Operators	2200 – 22FF	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Miscellaneous Technical	2300 – 23FF	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Control Pictures	2400 – 243F	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Optical Character Recognition	2440 – 245F	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Enclosed Alphanumerics	2460 – 24FF	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise

		hAnsi is used.
Box Drawing	2500 – 257F	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Block Elements	2580 – 259F	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Geometric Shapes	25A0 – 25FF	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Miscellaneous Symbols	2600 – 26FF	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
Dingbats	2700 – 27BF	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
CJK Radicals Supplement	2E80 – 2EFF	*eastAsia*
Kangxi Radicals	2F00 – 2FDF	*eastAsia*
Ideographic Description Characters	2FF0 – 2FFF	*eastAsia*
CJK Symbols and Punctuation	3000 – 303F	*eastAsia*
Hiragana	3040 – 309F	*eastAsia*
Katakana	30A0 – 30FF	*eastAsia*
Bopomofo	3100 – 312F	*eastAsia*
Hangul Compatibility Jamo	3130 – 318F	*eastAsia*
Kanbun	3190 – 319F	*eastAsia*
Enclosed CJK Letters and Months	3200 – 32FF	*eastAsia*
CJK Compatibility	3300 – 33FF	*eastAsia*
CJK Unified Ideographs Extension A	3400 – 4DBF	*eastAsia*
CJK Unified Ideographs	4E00 – 9FAF	*eastAsia*
Yi Syllables	A000 – A48F	*eastAsia*
Yi Radicals	A490 – A4CF	*eastAsia*
Hangul Syllables	AC00 – D7AF	*eastAsia*
High Surrogates	D800 – DB7F	*eastAsia*
High Private Use Surrogates	DB80 – DBFF	*eastAsia*
Low Surrogates	DC00 – DFFF	*eastAsia*
Private Use Area	E000 – F8FF	If the *hint* is *eastAsia* then *eastAsia* is used, otherwise *hAnsi* is used.
CJK Compatibility Ideographs	F900 – FAFF	*eastAsia*
Alphabetic Presentation Forms	FB00 – FB4F	If the *hint* is *eastAsia* then *eastAsia* is used for characters in the range FB00 – FB1C, otherwise *hAnsi* is used. For the range FB1D – FB4F, *ascii* is used.
Arabic Presentation Forms-A	FB50 – FDFF	*ascii*
CJK Compatibility Forms	FE30 – FE4F	*eastAsia*
Small Form Variants	FE50 – FE6F	*eastAsia*
Arabic Presentation Forms-B	FE70 – FEFE	*ascii*
Halfwidth and Fullwidth Forms	FF00 – FFEF	*eastAsia*

Resolving the
Font.docx

Include the following content in a new section under §21.1.2.3

For each Unicode character in a DML document, the Office Open XML specification specifies that its font face can be any from any of four font "slots": latin, cs, ea, or sym. This document describes how Microsoft Office determines the font slot to use.

For each Unicode character in a DML document, the following table is used. For all ranges not explicitly called out below, the *ea* font shall be used.

Unicode Range	Classification
0000 – 007F	*Use latin font*
F000 – F0FF	*Symbol,* use sym font
South Asian Scripts, such as Telugu	Use cs font
Right to left scripts, such as Arabic or Hebrew	Use cs font
Surrogate Unicode characters	Use ea font
3099 – 309A	Use ea font
2018 – 201E	Quote characters. If the text has an East Asian language ID, then use ea font. Otherwise use latin font.
0080 – 00A6	*Use latin font*
00A9 – 00AF	*Use latin font*
00B2 – 00B3	*Use latin font*
00B5 – 00D6	*Use latin font*
00D8 – 00F6	*Use latin font*
00F8 – 058F	*Use latin font*
10A0 – 10FF	*Use latin font*
1200 – 137F	*Use latin font*
13A0 – 177F	*Use latin font*
1D00 – 1D7F	*Use latin font*
1E00 – 1FFF	*Use latin font*
2000 – 200B	*Use latin font*
200C – 200F	*Use cs font*
2010 – 2029	*Use latin font*
202A – 202F	*Use cs font*
2030 – 2046	*Use latin font*
204A – 245F	*Use latin font*
2670 – 2671	*Use cs font*
27C0 – 2BFF	*Use latin font*

D835	Use latin font
FB00 – FB17	*Use latin font*
FE50 – FE6F	*Use latin font*
Otherwise	*Use ea font*