mPDF

Fonts in mPDF Version 5.x

Note: This example will only display correctly if you have all the fonts installed in mPDF

mPDF version 5 supports Truetype fonts, reading and embedding directly from the .ttf font files. Fonts must follow the Truetype specification and use Unicode mapping to the characters. Truetype collections (.ttc files) and Opentype files (.otf) in Truetype format are also supported.

Easy to add new fonts

1. Upload the Truetype font file to the fonts directory (/ttfonts)
2. Define the font file details in the configuration file (config_fonts.php)
3. Access the font by specifying it in your HTML code as the CSS font-family

These are some examples of Windows fonts:

Arial - The quick, sly fox jumped over the lazy brown dog.
Comic Sans MS - The quick, sly fox jumped over the lazy brown dog.
Trebuchet - The quick, sly fox jumped over the lazy brown dog.
Calibri - The quick, sly fox jumped over the lazy brown dog.
QuillScript - The quick, sly fox jumped over the lazy brown dog.
Lucidaconsole - The quick, sly fox jumped over the lazy brown dog.
Tahoma - The quick, sly fox jumped over the lazy brown dog.

Full Unicode support

The DejaVu fonts distributed with mPDF contain an extensive set of characters, but it is easy to add fonts to access uncommon characters.

Georgian (DejaVuSansCondensed)
ᲐᲧᲨᲝᲩᲣᲑᲒᲘᲜᲑᲐᲠᲥᲣᲧᲐᲖᲔᲥ
Runic (Junicode)
ᚠᚡᚢᚣᚤᚥᚦᚧᚨᚩᚪᚫᚬᚭᚮᚯᚰᚱᚲᚳᚴᚵᚶᚷᚸᚹᚺᚻᚼᚽᚾᚿᛀᛁᛂᛃᛄᛅᛆᛇᛈᛉᛊᛋᛌᛍᛎᛏᛐᛑᛒᛓᛔᛕᛖᛗᛘᛙᛚᛛᛜᛝᛞᛟᛠᛡᛢᛣᛤᛥᛦᛧᛨᛩᛪ᛫᛬᛭
Greek Extended (Quivira)
ἀἁἂἃἄἅἆἇἈἉἊἋἌἍἎἏἐἑἒἓἔἕ
IPA Extensions (Quivira)
ɐɑɒɓɔɕɖɗəɘɚɛɜɝɞʂʄʈʧʉʋʌʍʎʏʐʑʒʓʔʕʖʗʘʙʚʛʜʝʞʟʠʡʢʣʤʥʦʧʨʩʪʫ
Letterlike Symbolsa/ca/s℃℅℆℀℁℈℉ℊℋℌℍℎℏℐℑℒℓ℔№℗℘ℙℚℛℜ
Mathematical Operators (Symbola)
∀∁∂∃∄∅∆∇∈∉∊∋∌∍∎∏∐∑−∓∔∕∖∗∘∙√∛∜∝∞
Miscellaneous Technical (Symbola)
⌀⌁⌂⌃⌄⌅⌆⌇⌈⌉⌊⌋⌌⌍⌎⌏⌐⌑⌒⌓⌔⌕⌖⌗⌘⌙⌚⌛⌜⌝⌞⌟⌠⌡⌢⌣⌤⌥⌦⌧⌨
Enclosed Alphanumerics (Quivira)
①②③④⑤⑥⑦⑧⑨⑩⑪⑫⑬⑭⑮⑯
Miscellaneous Symbols (DejaVuSansCondensed)
☀☁☂☃☄★☆☇☈☉☊☋☌☍☎☏☐☑☒
Supplemental Arrows-B (Symbola)
⤀⤁⤂⤃⤄⤅⤆⤇⤈⤉⤊⤋⤌⤍⤎⤏⤐⤑⤒⤓⤔⤕⤖⤗⤘⤙⤚⤛⤜⤝⤞⤟⤠⤡⤢⤣⤤⤥⤦⤧⤨⤩⤪⤫⤬⤭⤮⤯⤰⤱⤲⤳⤴⤵⤶⤷⤸⤹⤺⤻⤼⤽⤾⤿⥀⥁⥂⥃⥄⥅⥆⥇⥈⥉⥊⥋⥌⥍⥎⥏⥐⥑⥒⥓⥔⥕⥖⥗⥘⥙⥚⥛⥜⥝⥞⥟⥠⥡⥢⥣⥤⥥⥦⥧⥨⥩⥪⥫⥬⥭⥮⥯⥰⥱⥲⥳⥴⥵⥶⥷⥸⥹⥺⥻⥼⥽⥾⥿

The online manual lists a number of open-source fonts available to cover uncommon characters:
http://mpdf1.com/manual/index.php?tid=451

Complex scripts

Although all characters may be displayed, there are various reasons why a script may not appear as expected.

Right-to-left languages (Hebrew, Arabic etc.)

Arabic languages and Hebrew are written in a right-to-left direction (RTL). mPDF recognises both Arabic and Hebrew languages and reverses text direction automatically.

ألسّلام عليكم שלום

Arabic languages (but not Hebrew) also change the form of the letter depending on its position in the text e.g. these are the initial, medial, final, and isolated forms of arabic letter 'ain':

ع ـع ـعـ عـ

The isolated characters are contained in the Unicode block 'Arabic' U+0600 - U+06FF.
The initial, medial and final forms are contained in Unicode Blocks 'Arabic Presentation Forms' A and B (U+FB50 - U+FDFF, U+FE70 - U+FEFE). Note that quite a large number of fonts contain the isolated characters but not the presentation forms. Fonts used with mPDF must contain the 'Arabic Presentation Forms' in order to display arabic text correctly. mPDF automatically converts letters to their initial/medial/final forms in several languages: arabic, persian/farsi, urdu, sindhi and pashto.

Arabic text is used for many different languages e.g. persian/farsi, urdu, pashto etc. These languages often contain letters unique to that language. 'Arabic' fonts do not always contain the full set of arabic charcters necessary for all languages.

Other RTL languages (using other alphabets) are reversed in order, but not otherwise processed, by mPDF e.g. Syriac, Thaana, N'Ko, and Samaritan.

Indic languages

Indic languages are also complex scripts which require some processing of characters before display. For example some vowels consist of 2 characters, to be placed before and after the adjacent consonant e.g.

U+0D1C + U+0D4C [vowel AU] = [written together as ജൌ]

□ + □ = □□

Consonant conjuncts are where two adjacent characters are written as a single 'conjunct' form e.g.

□ + □ = □□

mPDF can support some of these languages, but requires specially prepared font files that are unique to mPDF.
Supported languages: Bengali, Devanāgarī, Gujarāti, Gurmukhi, Kannada, Malayalam, Oriya, Tamil, Telugu

□□□□□□□□ □□□□□□□ □□□□□□ □□□□□□□□ □□□□□□ □□□□□□□!

Complex scripts **not** supported: Khmer, Sinhala, Tibetan, Myanmar (Burmese), Balinese

Vertical writing

Vertical writing is not supported by mPDF (e.g. Mongolian and Phags-pa) although the individual characters can be displayed using suitable fonts.

Combining diacritics

In Unicode, letters with diacritics (e.g. ÁáȦȧÄä) are usually represented as a single character e.g. Unicode U+0196 is an A Umlaut. There are 4 blocks in Unicode of diacritics or 'marks' which can be used to combine with adjacent letters: Combining Diacritical Marks (U+0300 - U+036F), Combining Diacritical Marks Supplement (U+1DC0 - U+1DFF), Combining Marks for Symbols(U+20D0 - U+20FF) and Combining Half Marks (U+FE20 - U+FE2F).

Software applications use special positioning information stored in OpenType font files to resposition the diacritic/mark depending on the context. mPDF does not support this repositioning and is dependent on the font design and original placement of the diacritic:

Á á Ȧ ȧ Ä ä ï (Precomposed characters: DejaVu Sans Condensed)
Á́ á Ȧ́ ȧ Ä́ ä ï (Using diacritics: DejaVu Sans Condensed)
Á́ á Ȧ́ ȧ Ä́ ä ï (Arial Unicode MS)
Á́ á Ȧ́ ȧ Ä́ ä ï (Times New Roman)
A´ a´ Ȧ ȧ A¨ a¨ i¨ (Courier New)

It is recommended to use precomposed characters whenever possible with mPDF.

Unicode Supplementary Planes

The original Unicode allocated characters between x0000 and xFFFF (65,536 characters). This 'Basic Multilingual Plane' supported most characters in common use, including a large number of Unified Chinese-Japanese-Korean characters (CJK). Later the Unicode standard was extended to 16 Planes.

The first plane (plane 0), the Basic Multilingual Plane (BMP), is where most characters have been assigned so far.

Plane 1, the Supplementary Multilingual Plane (SMP), is mostly used for historic scripts such as Linear B, but is also used for musical and mathematical symbols.

Plane 2, the Supplementary Ideographic Plane (SIP), is used for about 40,000 Unified Han (CJK) Ideographs.

mPDF version 5 supports fonts containing characters from all Unicode Planes. By choosing the correct font, almost every single character from Unicode 5 can be displayed in a PDF file.

Unicode Supplementary Multilingual Plane (SMP or Plane 1) U+10000 - U+1FFFF

Gothic text

This paragraph shows Gothic text. These characters lie in the Unicode Supplementary Multilingual Plane U+10330 - U+1034F.

□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□□

Font: MPH2BDamase (damase_v.2.ttf) available from: http://www.wazu.jp/gallery/views/View_MPH2BDamase.html

Egyptian Hieroglyphics

This paragraph shows Egyptian Hieroglyphics. These characters lie in the Unicode Supplementary Multilingual Plane U+13000 - U+1342F.



Font: Aegyptus.otf available from: http://users.teilar.gr/~g1951d/

SMP contains mainly ancient scripts - see http://mpdf1.com/manual/index.php?tid=451 for full list.

mPDF uses a different method to embed fonts in the PDF file if they include characters from SMP or SIP, because the characters cannot be represented by a 4 character hex code 0000-FFFF. This method is less eficient than the default method, and it can be suppressed by adding the font name to the array 'BMPonly' in the config_fonts.php configuration file.

Note that the DejaVu fonts distributed with mPDF and (GNU)FreeSans and FreeSerif fonts do contain a few characters in the SMP plane, but most users will not require them and by default they have been added to the array 'BMPonly'.

CJK CHARACTERS

Below are examples of all the CJK Unicode blocks contained in the Basic Multilingual Plane and Supplemental Ideographic Plane

Plane 0 (BMP)

CJK Radicals Supplement

⺠ ⺡ ⺢ ⺣ ⺤ ⺥ ⺦ ⺧ ⺨ ⺩ ⺪ ⺫ ⺬ ⺭ ⺮ ⺯ ⺰ ⺱ ⺲ ⺳ ⺴ ⺵ ⺶ ⺷ ⺸ ⺹ ⺺ ⺻ ⺼ ⺽ ⺾ ⺿

Kangxi Radicals

⼳ ⼴ ⼵ ⼶ ⼷ ⼸ ⼹ ⼺ ⼻ ⼼ ⼽ ⼾ ⼿ ⽀ ⽁ ⽂ ⽃ ⽄ ⽅ ⽆ ⽇ ⽈ ⽉ ⽊ ⽋ ⽌ ⽍

Ideographic Description Characters

⿰ ⿱ ⿲ ⿳ ⿴ ⿵ ⿶ ⿷ ⿸ ⿹ ⿺ ⿻

CJK Symbols and Punctuation

〈 〉 《 》 「 」 『 』 【 】 〒 〓 〔 〕 〖 〗 〘 〙 〚 〛 〜 〝 〞 〟 〠 〡 〢 〣 。

Hiragana

ぁ あ ぃ い ぅ う ぇ え ぉ お か が き ぎ く ぐ け げ こ ご さ ざ

Katakana

゠ ァ ア ィ イ ゥ ウ ェ エ ォ オ カ ガ キ ギ ク グ ケ ゲ コ ゴ

Bopomofo

ㄅ ㄆ ㄇ ㄈ ㄉ ㄊ ㄋ ㄌ ㄍ ㄎ ㄏ ㄐ ㄑ ㄒ ㄓ ㄔ ㄕ ㄖ ㄗ ㄘ ㄙ

Hangul Compatibility Jamo

ㄱ ㄲ ㄳ ㄴ ㄵ ㄶ ㄷ ㄸ ㄹ ㄺ ㄻ ㄼ ㄽ ㄾ ㄿ ㅀ ㅁ ㅂ

Kanbun

㆖ ㆗ ㆘ ㆙ ㆚ ㆛ ㆜ ㆝ ㆞ ㆟ ㆠ ㆡ ㆢ ㆣ ㆤ ㆥ ㆦ ㆧ

Bopomofo Extended

ㄅ ㄆ ㄇ ㄈ ㄉ ㄊ ㄋ ㄌ ㄍ ㄎ ㄏ ㄐ ㄑ ㄒ ㄓ ㄔ ㄕ ㄖ ㄗ ㄘ ㄙ ㄚ ㄛ ㄜ ㄝ ㄞ ㄟ ㄠ ㄡ ㄢ ㄣ ㄤ ㄥ ㄦ ㄧ ㄨ ㄩ

Katakana Phonetic Extensions

ク シ ス ト ヌ ハ ヒ フ ヘ ホ ム ラ リ ル レ ロ

Enclosed CJK Letters and Months

㈀ ㈁ ㈂ ㈃ ㈄ ㈅ ㈆ ㈇ ㈈ ㈉ ㈊ ㈋ ㈌ ㈍ ㈎ ㈏ ㈐ ㈑ ㈒ ㈓ ㈔ ㈕ ㈖ ㈗ ㈘ ㈙ ㈚ ㈛

CJK Compatibility

㌀ ㌁ ㌂ ㌃ ㌄ ㌅ ㌆ ㌇ ㌈ ㌉ ㌊ ㌋ ㌌ ㌍ ㌎ ㌏ ㌐ ㌑ ㌒ ㌓ ㌔ ㌕ ㌖ ㌗ ㌘ ㌙ ㌚ ㌛

CJK Unified Ideographs Extension A

㐀 㐁 㐂 㐃 㐄 㐅 㐆 㐇 㐈 㐉 㐊 㐋 㐌 㐍 㐎 㐏 㐐 㐑 㐒 㐓 㐔 㐕 㐖 㐗 㐘 㐙 㐚 㐛

CJK Unified Ideographs

一 丁 丂 七 丄 丅 丆 万 丈 三 上 下 丌 不 与 丏 丐 丑 丒 专 且 丕 世 丗 丘

Yi Syllables

ꀀ ꀁ ꀂ ꀃ ꀄ ꀅ ꀆ ꀇ ꀈ ꀉ ꀊ ꀋ ꀌ ꀍ ꀎ ꀏ ꀐ ꀑ ꀒ ꀓ ꀔ ꀕ ꀖ ꀗ ꀘ ꀙ ꀚ ꀛ

Yi Radicals

ꀜ ꀝ ꀞ ꀟ ꀠ ꀡ ꀢ ꀣ ꀤ ꀥ ꀦ ꀧ ꀨ ꀩ ꀪ ꀫ ꀬ ꀭ ꀮ ꀯ ꀰ ꀱ ꀲ ꀳ ꀴ ꀵ ꀶ ꀷ

Hangul Syllables

가 각 갂 갃 간 갅 갆 갇 갈 갉 갊 감 갋 갌 갍 갎 갏 값 갑 값 갓 갔 강 갖 갗 갘

CJK Compatibility Ideographs

豈 更 車 賈 滑 串 句 龜 契 金 喇 奈 懶 癩 羅 蘿 螺 裸 邏 樂 洛 烙 珞 落 酪

Plane 2 (SIP)

CJK Unified Ideographs Extension B

𠀀 𠀁 𠀂 𠀃 𠀄 𠀅 𠀆 𠀇 𠀈 𠀉 𠀊 𠀋 𠀌 𠀍 𠀎 𠀏 𠀐 𠀑 𠀒 𠀓 𠀔 𠀕 𠀖 𠀗 𠀘 𠀙 𠀚 𠀛

CJK Unified Ideographs Extension C

𪜀 𪜁 𪜂 𪜃 𪜄 𪜅 𪜆 𪜇 𪜈 𪜉 𪜊 𪜋 𪜌 𪜍 𪜎 𪜏 𪜐 𪜑 𪜒 𪜓 𪜔 𪜕 𪜖 𪜗 𪜘 𪜙 𪜚 𪜛

CJK Compatibility Ideographs Supplement

丽 丸 乁 亄 你 侮 侻 倂 偺 備 僧 像 僿 充 兗 兎 兤 具 與 𪜈

Using CJK fonts in mPDF

Fonts containing CJK characters are large files, typically 10-30MB. Adobe provides a free download of an 'Asian font pack' allowing you to create PDF files without including (embedding) the font information in the file. This keeps the file size to a minimum and minimises resource usage on your website generating the PDF file. However, users will have to download the Adobe font packs to read the file, and other PDF software will not display the text correctly.

mPDF allows you to embed subsets of CJK fonts keeping file size down, although there is increased memory usage to generate these files.

Some CJK fonts are broken up into 2 files because of the size of the files. One freely available font with almost complete coverage of all CJK characters (in both BMP and SIP) is 'Sun' available from Alan Wood's excellent website: http://www.alanwood.net/unicode/fonts-east-asian.html. This comes as 2 files, Sun-ExtA and Sun-ExtB (both about 20MB in size) containing the characters from BMP and SIP respectively.

mPDF allows you to treat these as one font by defining the second file as an SIP-extension of the first in the config_fonts.php configuration file. The following text includes random characters from the BMP and SIP mixed together:

```
鼂 鼈 鼗 晢 喻 鼦 鼧 鼠 暢 鼹 齄 暱 冕 齅 齐 喵 暅 齔 齚 瞇 瞀 齡 齩 嘰 暤
```

This is the entry in the config_fonts.php configuration file:

```
$this->fontdata = array(
...
      "sun-exta" => array(
            'R' => "Sun-ExtA.ttf",
            'sip-ext' => 'sun-extb',
      ),
      "sun-extb" => array(
            'R' => "Sun-ExtB.ttf",
      ),

...
);
```

This is the HTML code - note only the sun-exta font-family needs to be referenced:

```
<div style="font-family:sun-extA;"> &#40706; &#40712; &#40727; &#x2320f; &#x23225; &#40742;
&#40743; &#x2322f; &#x23231; &#40761; &#40772; &#x23232; &#x23233; &#40773; &#40784; &#x23234;
&#x23256; &#40787; &#40794; &#x23262; &#x23281; &#40802; &#40809; &#x23289; &#x2328a; </div>
```

NB You may also need to edit the value $this->useAdobeCJK=false in config.php or use new mPDF('-aCJK'), and edit the config_cp.php configuration file.

TrueType Collections

TrueType Collections (.ttc files) contain more than one font. mPDF treats each font separately by defining the TTCfontID array in the config_fonts.php configuration file.

This example uses the Windows MingLiU fonts, which consist of 2 files containing 6 fonts (note that mingliub is not a Bold variant):

Font collection file (mingliu.ttc) contains the following fonts:
[1] MingLiU (mingliu) Regular
[2] PMingLiU (pmingliu) Regular (Proportional)
[3] MingLiU_HKSCS (mingliu_hkscs) Regular

Font collection file (mingliub.ttc) contains the following fonts:
[1] MingLiU-ExtB (mingliu-extb) Regular
[2] PMingLiU-ExtB (pmingliu-extb) Regular (Proportional)
[3] MingLiU_HKSCS-ExtB (mingliu_hkscs-extb) Regular

The following text includes characters from both BMP and SIP:

```
□□□□□□□□
□□□□□□□□
□□□□□□□□
```

This is the entry in the config_fonts.php configuration file:

```
$this->fontdata = array(
...
      "mingliu" => array(
            'R' => "mingliu.ttc",
            'TTCfontID' => array (
                'R' => 1,
            ),
            'sip-ext' => 'mingliu-extb',
      ),
      "pmingliu" => array(
            'R' => "mingliu.ttc",
            'TTCfontID' => array (
                'R' => 2,
            ),
            'sip-ext' => 'pmingliu-extb',
      ),
      "mingliu_hkscs" => array(
            'R' => "mingliu.ttc",
            'TTCfontID' => array (
                'R' => 3,
            ),
            'sip-ext' => 'mingliu_hkscs-extb',
      ),
      "mingliu-extb" => array(
            'R' => "mingliub.ttc",
            'TTCfontID' => array (
                'R' => 1,
            ),
      ),
      "pmingliu-extb" => array(
            'R' => "mingliub.ttc",
            'TTCfontID' => array (
                'R' => 2,
            ),
```

```
      ),
      "mingliu_hkscs-extb" => array(
            'R' => "mingliub.ttc",
            'TTCfontID' => array (
                'R' => 3,
            ),
      ),

...
);
```

This is the HTML code:

```
<div style="font-family:mingliu;"> &#40706; &#40742; &#40772; &#40784; &#40802; &#40809;
&#x23289; &#x2328a; </div>
<div style="font-family:mingliu_hkscs;"> &#40706; &#40742; &#40772; &#40784; &#40802; &#40809;
&#x23289; &#x2328a; </div>
<div style="font-family:pmingliu;"> &#40706; &#40742; &#40772; &#40784; &#40802; &#40809;
&#x23289; &#x2328a; </div>
```