

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 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.

ႁႂႃႇႅႆႉႁႃႇႁႃႇႁႃႇႁႃႇႁႃႇႁႃႇႁႃႇႁႃႇႁႃႇႁႃႇႁႃႇႁႃႇႁႃႇႁႃႇ

Font: MPH2BDamase (damase_v.2.ttf) available from:
[http://www.wazu.jp/gallery/views/View MPH2BDamase.html](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 efficient 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

兀 允 尢 允 巳 幺 亓 𠂇 卩 小 才 攴 无 冏 月 夕 母 民 彡 氷 凵 𠂆 𠂇 斗 生 彳 王 正

Kangxi Radicals

幺 广 爻 升 弋 弓 彡 彡 心 戈 户 手 支 支 文 斗 斤 方 无 日 曰 月 木 欠 止 歹

Ideographic Description Characters

CJK Symbols and Punctuation

〈 〉 《 》 「 」 『 』 【 】 ㄟ ㄚ () [] [] [] ~ " " ± ± ± 夕。

Hiragana

あ あ い い う う え え お お か が き ぎ く ぐ け げ こ ご さ ざ

Katakana

= ア ア イ イ ウ ウ エ エ オ オ カ ガ キ ギ ク グ ケ ゲ コ ゴ

Bopomofa

ㄅ ㄆ ㄇ ㄏ ㄒ ㄓ ㄔ ㄕ ㄖ ㄗ ㄘ ㄙ ㄖ ㄗ ㄘ ㄙ ㄖ ㄗ ㄘ ㄙ

Hangul Compatibility Jamo

ㄅ ㄆ ㄇ ㄌ ㄎ ㄔ ㄗ ㄛ ㄝ ㄟ ㄠ ㄡ ㄢ ㄣ

Kanbun

一 二 三 四 上 中 下 甲 乙 丙 丁 天 地 人

Bopomofo Extended

ㄅ ㄆ ㄇ ㄏ ㄙ ㄜ ㄝ ㄚ ㄛ ㄣ ㄤ ㄨ ㄩ ㄟ ㄠ ㄡ ㄢ ㄣ ㄤ ㄨ ㄩ ㄟ ㄠ ㄡ

Katakana Phonetic Extensions

ク シ ス ト ヌ ハ ヒ フ ヘ ホ ム ラ リ ル レ ロ

Enclosed CJK Letters and Months

(口) (ㅂ) (ㅅ) (ㅇ) (ㅈ) (ㅊ) (ㅋ) (ㆁ) (ㄷ) (ㄹ) (ㅎ) (가) (나) (다) (라) (마) (바) (사) (아) (자) (차) (카) (타) (파) (하) (주)

CJK Compatibility

アバート アルファ アンベア アール イニング インチ ウォン エスクード エーカー オン オム カイ カラット カロリー ガン ガン ギガ ギニ ギュリー ギルダナー キ キロ キログラム キロメートル キロワット グラム

CJK Unified Ideographs Extension A

[illegible]

CJK Unified Ideographs

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

Yi Syllables

[illegible]

Yi Radicals

| | | | | | | | | | | S | | C | E | C | | U | V | W | Q | O | O

Hangul Syllables

[illegible]

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

己 方 乙 且 工 历 丰 且 丘 丘 东 丈 卅 卅 卅 屯 中 丙 兮 引

CJK Unified Ideographs Extension C

卅 卅 凡 瓦 事 奇 奔 颯 臉 亡 怦 之 妻 厠 乖 航 稗 仇 李 浩

CJK Compatibility Ideographs Supplement

𠂔 丸 𠂔 𠂔 你 侮 悅 併 偕 備 僧 像 儻 充 免 免 𠂔 具 與 𠂔

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

This is a sample 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 sample 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>
```