

中日韓越(CJKV)表意文字筆畫檢索

中日韓越表意文字：

一、普通表意文字塊，範圍 4E00-9FFF，含 20902 字，1991 年隨 Unicode 1.0 版本發佈；

二、表意文字擴展 A，範圍 3400-4DBF，含 6582 字，1999 年隨 Unicode 3.0 版本發佈；

三、表意文字擴展 B，範圍 20000-2A6DF，含 42711 字，2001 年隨 Unicode 3.1 版本發佈；

四、表意文字擴展 C，範圍 2A700-2B73F，含 4149 字，2002-2006 年間提交 IRG，隨 Unicode 5.2 發佈；

五、表意文字擴展 D，範圍 2B740-2B81F，含 222 字，2006-2009 年間提交 IRG，隨 Unicode 6.0 發佈；

六、表意文字擴展 E，範圍 2B820-2CEAF，含 5762 字，2006-2013 年間提交 IRG，隨 Unicode 8.0 發佈；

七、表意文字擴展 F，範圍 2CEB0-2EBEF，含 7473 字，2012-2015 年間提交 IRG，隨 Unicode 10.0 發佈。

定義：

三字節字符：按 UTF-8 編碼規則長度為 3 字節的字符，按 ANSI 編碼規則長度為 2 字節。

如“人”：UTF-8 編碼%E4%BA%BA，GBK 編碼 0xC8CB，UNICODE 編碼 0x4EBA。

四字節字符：按 UTF-8 編碼規則長度為 4 字節的字符。

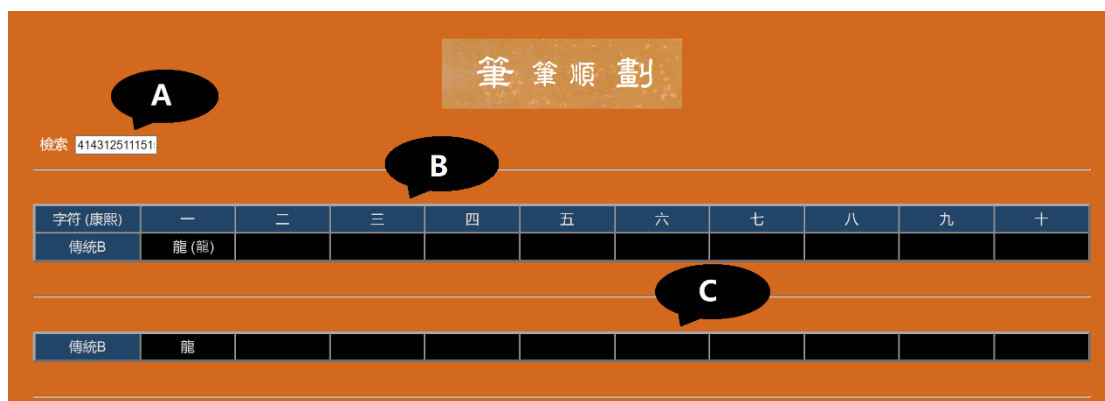
如“刁”：UTF-8 編碼%F0%A0%84%8E，UNICODE 編碼 0x02010E。

字族：在 GBK 字符集內七劃以上唯一存在的，其筆順碼為其他漢字筆畫碼結尾的字符，如“威”。

字元：在 GBK 字符集內七劃以上唯一存在的，其筆順碼為其他漢字筆畫碼開頭的字符，如“蕊”。

漢字可以用筆畫描述，漢字的基本筆畫可分為五種

- 一、橫：用阿拉伯數字 1 表示
- 二、豎：用阿拉伯數字 2 表示
- 三、撇：用阿拉伯數字 3 表示
- 四、捺：用阿拉伯數字 4 表示
- 五、折：用阿拉伯數字 5 表示



圖一

圖一 A 區為筆畫檢索

A 區需錄入檢索漢字基本筆畫的數字編碼，譬如：“龍”字的編碼為 4143125111515111，錄入完畢按回車鍵(Enter)結束。

圖一 B 區為筆順顯示

B 區按照基本漢字、傳統漢字、CJKV 擴展漢字分類，顯示內容為與 A 區檢索數字編碼相同且筆畫順序一致的所有 unicode 字符及其對應的康熙字典部首字符。

圖一 C 區為筆畫顯示

C 區也按照基本漢字、傳統漢字、CJKV 擴展漢字分類，顯示內容為與 A 區檢索數字編碼相同且筆畫數一致的 unicode 字符，限制每個分類至多顯示二十個字符。

版本 1.0.0 支持檢索範圍從普通表意文字至 CJKV 表意文字擴展 D 區之間的字符集。

- 版本 1.0.1 新增個別 CJKV 表意文字擴展 E 區字符的檢索。
- 版本 1.0.2 “雨”字筆順碼 12524444 作字底如“漏”，“雨”字筆順碼 14524444 作字頭如“雷”。
- 版本 1.1.2 新增每日一幅背景圖片，保留顯示 GBK 字符區。
- 版本 1.1.3 鼠標覆蓋基本多文種平面字符，自動複製字符至剪貼板（僅 IE 瀏覽器，5.0 版本以上）。
- 版本 1.1.4 優化筆順區覆蓋字符的顯示及放大功能。
- 版本 1.1.5 筆順區字符單擊放大，字符雙擊複製。
- 版本 1.1.6 筆劃區字符輔以 Unicode 標準內碼顯示。
- 版本 1.2.6 檢索框輸入單個 CJKV 三字節字符，自動轉換成筆順碼輸出，同時顯示檢索字符集。
- 版本 1.3.6 新增四千多箇 GBK 字符的字族/字元分析功能，點擊帶有橘黃色的標題字符執行分區和部首分析。
- 版本 1.3.7 合并顯示基本、傳統 A、傳統 B 三區為 GBK 區。

更新日期：2020 年 10 月 12 日

Stroke Search of CJKV Unified Ideographs

Chinese, Japanese, Korean and Vietnamese Ideographs :

1. The first CJK Unified Ideographs, Range: 4E00-9FFF, including 20902 words, Released with Unicode version 1.0 in 1991;
2. CJK Unified Ideographs Extension A, Range: 3400-4DBF , including 6582 words, Released with Unicode version 3.0 in 1999;
3. CJK Unified Ideographs Extension B, Range: 20000-2A6DF , including 42711 words, Released with Unicode version 3.1 in 2001;
4. CJK Unified Ideographs Extension C, Range: 2A700-2B73F , including 4149 words, Submitted to IRG during 2002-2006, Released with Unicode version 5.2;
5. CJK Unified Ideographs Extension D, Range: 2B740-2B81F , including 222 words, Submitted to IRG during 2006-2009, Released with Unicode version 6.0;
6. CJK Unified Ideographs Extension E, Range: 2B820-2CEAF , including 5762 words, Submitted to IRG during 2006-2013, Released with Unicode version 8.0;
7. CJK Unified Ideographs Extension F, Range: 2CEB0-2EBEF , including 7473 words, Submitted to IRG during 2012-2015, Released with Unicode version 10.0.

Ideographs can be described by strokes, the basic strokes of Ideographs can be divided into five types

1. Horizontal stroke: represented by the Arabic numeral 1
2. Vertical stroke: represented by the Arabic numeral 2
3. Leftfalling stroke: represented by the Arabic numeral 3
4. Rightfalling stroke: represented by the Arabic numeral 4
5. Zigzag stroke: represented by the Arabic numeral 5

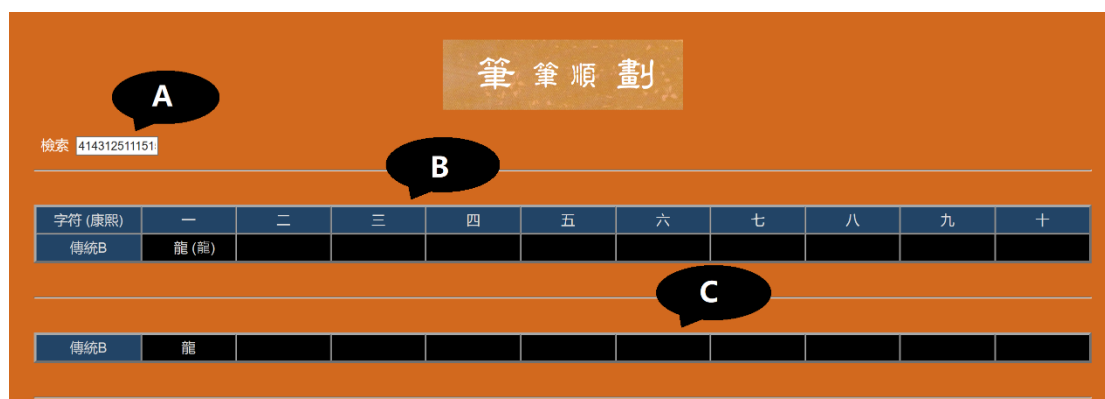


Figure One

Figure One: Area A is stroke input area

In area A, the number code of the basic strokes of Ideographs should be input, for example: The code of the word 龍 is 4143125111515111, Press enter to finish the entry.

Figure One: Area B is stroke order display area

Area B is classified according to basic Chinese characters, traditional Chinese characters and CJKV extended Ideographs, the display contents are all the Unicode characters and corresponding

Kangxi dictionary radical characters with the same number code and stroke order as the search number in area A.

Figure One: Area C is stroke display area

Area C is also classified according to basic Chinese characters, traditional Chinese characters and CJKV extended Ideographs, the display contents are Unicode characters with the same number of strokes as the search number in area A, Limit each category to 20 characters.

Instructions of version update

Version 1.0.0 Support the search range from the first CJK Unified Ideographs to CJK Unified Ideographs Extension D block.

Version 1.0.1 Add individual search of CJK Unified Ideographs Extension E block.

Version 1.0.2 雨 word stroke order code is 12524444 at the bottom of the word 漏, 雨 word stroke order code is 14524444 at the top of the word 雷.

Version 1.1.2 Add a daily background image.

Version 1.1.3 Mouse over automatically copy character to the clipboard (internet Explorer only, version 5.0 or above).

Version 1.1.4 Optimize the display and enlarge function of mouse over the character in stroke order area.

Version 1.1.5 Click to enlarge character in stroke order area, and double-click to copy character.

Version 1.1.6 Characters in stroke area are displayed with Unicode standard inner code.

Version 1.2.6 A single CJKV character is input in the search box, which is automatically converted into stroke order code output, and the retrieval character set is displayed at the same time.

Version 1.3.6 Add more than 4000 GBK character family/meta analysis function, click the title character with orange to perform partition and radical analysis.

Version 1.3.7 Combined display three areas Basic, Traditional A and B are GBK Area.

CJKV 軟體研發技術支持

