

## 新字符增收程序和原則

### 背景

《香港增補字符集》在 1999 年 9 月出版後，市民及政府部門仍不時需要把新字符納入字符集。為此，政府與中文界面諮詢委員會（中諮會）於 2000 年 4 月公布《香港增補字符集》字符增收程序和原則，並於 2009 年 9 月加以修訂，全面採用 ISO/IEC 10646 國際編碼標準編配的碼位。

### 建議

2. 為善用 ISO/IEC 10646 國際編碼標準的橫向擴展機制，秘書處現建議修訂《香港增補字符集》的字符增收程序和原則，以便日後《香港增補字符集》除增收尚未納入 ISO/IEC 10646 國際編碼標準字集的字符外，還可為該字集已收錄的字符補充香港字形，務求更全面反映本地用字的需要。

3. 字符增收原則經修訂後，若干過去不獲批准增收的字符，應可按橫向擴展機制重新獲考慮增收。秘書處建議為這些字符作初步分類，以便工作小組根據修訂原則重新審核。

### 徵詢意見

4. 請工作小組成員省覽分別載於附件一和附件二的《香港增補字符集》字符增收程序和原則修訂擬稿，以及載於附件三有待重新審核字符的分類及初步分析結果，並發表意見。秘書處會整理意見，提交中諮會考慮。

2015 年 10 月

附件一：

## 《香港增補字符集》增收程序修訂擬稿

由於《香港增補字符集》日後可以增收 ISO/IEC 10646 國際編碼標準字集已收錄的字符，預計需要更新版本以增收字符的機會將會較頻繁。為求更有效增收字符，現建議以出版補編的方式，把增收的字符正式納入《香港增補字符集》，然後再視乎情況把若干補編合併，出版《香港增補字符集》新版本。現建議大致沿用原有增收程序，主要修訂以藍色字體顯示如下：

1. 市民及政府部門可向中文界面諮詢委員會（中諮會）秘書處提交字符增收申請。申請必須連同字符的字形、讀音及意義等資料提交，如能提供字符的用途、字源和例句更佳。如增收申請涉及符號，申請人宜把相關符號的資料一併提交。
2. 市民提交的字符，會經中諮會秘書處轉交中文界面諮詢委員會工作小組（工作小組）審核。政府部門提交的字符，會由政府內部專責小組初步審議，再提交工作小組審核。
3. 工作小組會根據下文的原則審核字符增收申請。為了避免浪費資源並加快字符審核過程，市民及政府部門應先參考字符增收原則，以免提交不必要的申請。
4. 通過工作小組審核的申請會提交中諮會確認。如申請符合“迫切需要編碼的字符”條件，中諮會秘書處會為字符編配“來源參考”，格式為 HC-0001 至 HC-9999，供申請人參考。字符的字形及相關資料會在互聯網上公布，但必須強調，這些資料只供參考，字符尚待納入《香港增補字符集》（見下文第 6 段）。
5. 如申請不獲接納，申請人可提供進一步資料，要求工作小組重新考慮。由於這些字符大部分已編配 ISO/IEC 10646 碼位，使用者可以在支援 ISO/IEC 10646 國際編碼標準或統一碼（Unicode）的平台上處理這些字符。

## 《香港增補字符集》的正式更新

6. 由於軟件商開發軟件支援增收字符需時，加上字符集更新過於頻密會對使用者造成不便，甚至引起混亂，工作小組會視乎實際需要，在適當時間向中諮會建議公布《香港增補字符集》補編，把增收字符正式納入字符集。在若干補編公布後，工作小組會再視乎實際需要，建議中諮會把補編合併，出版《香港增補字符集》新版本。

7. 所有通過審核但尚未納入 ISO/IEC 10646 國際編碼標準字符集的增收字符，均會提交國際標準化組織轄下的表意文字小組。

~~8. 為進一步與國際標準接軌，中諮會決定由二零零九年九月起，《香港增補字符集》全面採用國際標準化組織接納字符申請時所編配的 ISO/IEC 10646 正式碼位，不再另外加配香港增補字符“標準內碼”，以免字符重複編碼。以往已經公布“標準內碼”的字符，其“標準內碼”仍然生效。~~

附件二：

## 《香港增補字符集》字符增收原則修訂擬稿

《香港增補字符集》日後除增收尚未納入 ISO/IEC 10646 國際編碼標準字集的字符外，還會在字集內以橫向擴展形式補充香港字形。現建議修訂《香港增補字符集》字符增收原則如下：

1. 《香港增補字符集》只收納市民和政府部門在進行電子通訊時有需要使用的字符，增收字符僅限下列範圍：
  - 1.1. 香港社羣用字，包括：
    - (a) 粵語字（須提供依據）；
    - (b) 能證明此前已流通的專有名詞用字（須提供讀音）；以及
    - (c) 作專有名詞用的字而該字見於大型字典（包括《康熙字典》、《漢語大字典》）。
  - 1.2 華人社會中非地區性的新生漢字（例如新生科學名詞用字）。

在上述範圍以外，其他有助於電子通訊的字符會個別考慮。

2. 申請增收的字符須按“字符等同原則”甄別。凡在《香港增補字符集》內已有可視為等同之字符者，概不增收。字符等同原則指 ISO/IEC 10646 等同原則，詳見 ISO/IEC 10646: Annex S - Procedure for the unification and arrangement of CJK Ideographs。等同字符的例子有：青—青、俞—俞、直—直、并—并、植—植、餅—餅。
3. 申請增收的字符如在《香港增補字符集》內並無等同字符，但在 ISO/IEC 10646 國際編碼標準字集內卻有可視為等同者，會根據下列因素考慮是否納入《香港增補字符集》：
  - (a) 該字符的字形是否在香港廣為接納；以及
  - (b) 該字符是否容易與《香港增補字符集》內的字符混淆。

〔註：工作小組文件編號 2015/04 建議把有關字符是否《香港增補字符集》字符的異體字列為考慮因素。由於字形相近的異體字可歸類為容易混淆字符，本修訂擬稿已刪去該因素。〕

字符納入《香港增補字符集》後，將在 ISO/IEC 10646 國際編碼標準字集內加入香港字形。

參考資料：現行的《香港增補字符集》字符增收原則

### 《香港增補字符集》字符增收原則

1. 凡在 ISO/IEC 10646 國際編碼標準字集及《香港增補字符集》內已有可視為等同之字符者，概不增收；因此，增收字符須經“字符等同原則”甄別。根據 ISO/IEC 10646 等同原則(即 ISO/IEC 10646 “Annex S – [Procedure for the unification and arrangement of CJK Ideographs](#)”)認為等同之字符，均視為等同(例如：青 — 青、兪 — 兪、直 — 直、并 — 并、植 — 植、餅 — 餅)。

2. 字符集只收納市民和政府部門在進行電子通訊時有需要使用的字符，增收字符僅限下列範圍：

2.1. 香港社羣用字，包括：

- (a) 粵語字 (須提供依據)；
- (b) 能證明此前已流通的專有名詞用字(須提供讀音)；以及
- (c) 作專有名詞用的字而該字見於大型字典(包括《康熙字典》、《漢語大字典》等)者。

2.2. 華人社會中非地區性的新生漢字(例如：新生科學名詞用字)。

在上述範圍以外，其他有助於電子通訊的字符，會個別考慮。

注：查核字符是否已收錄於 ISO/IEC 10646 國際編碼標準字集及《香港增補字符集》的方法見附件。

(修訂日期 10.2011)

### 附件三： 有待重新審核字符的分類及初步分析結果

由 2008 年 3 月 31 日至 2015 年 8 月 21 日，工作小組處理的《香港增補字符集》字符增收申請共涉及 908 個字符，全部無須增收，當中 268 個並非大五碼或《香港增補字符集—2008》字符。工作小組因應 ISO/IEC 10646 國際編碼標準的橫向擴展機制而修訂《香港增補字符集》字符增收原則後，可考慮重新審議這些字符應否納入該字符集。

為求更有效審議這些字符，秘書處建議把字符分類，並按字源、字形和字義作初步分析，供工作小組參考。

#### 分類及初步分析的參考資料

##### (A) 字源資料

在這 268 個字符中，35 個屬國際表意文字子集 (IICORE) 字符，當中 28 個只有一項字源資料 (G: 9 個、J: 17 個、K: 1 個、P: 1 個)，5 個有兩項，1 個有三項，還有 1 個沒有字源資料。

ISO/IEC 10646:2014 編碼表載列每個字符在字源地區的常用字形 (由一至六個不等)。這些資料可以反映字符在哪些地區使用及其通用程度。

##### (B) 字形資料

字形方面，某些有待重新審議的字符容易與大五碼字符混淆 (例如只有一筆之差)。參考資料包括：

(1) 《容易混淆字符表》

[http://www.ogcio.gov.hk/tc/business/tech\\_promotion/ccli/terms/doc/Easily\\_Confused\\_Chinese\\_Characters.pdf](http://www.ogcio.gov.hk/tc/business/tech_promotion/ccli/terms/doc/Easily_Confused_Chinese_Characters.pdf)

(2) UniHan database

<http://www.unicode.org/charts/unihan.html>

## (C) 字義資料

字符是否見於大型字典及是否簡化字，均是審批的考慮因素。字義資料主要參考：

(1) 《康熙字典》和《漢語大字典》

(2) Unihan database

<http://www.unicode.org/charts/unihan.html>

## 分類建議

秘書處建議按下述分類分批處理 268 個字符：

處理次序	字符類別	數目
1	IICORE 字符	35 個
2	有 5 個字源地區的非 IICORE 字符	1 個
	有 4 個字源地區的非 IICORE 字符	54 個
3	有 3 個字源地區的非 IICORE 字符	91 個
4	有 2 個字源地區的非 IICORE 字符	85 個
	有 1 個字源地區的非 IICORE 字符	2 個

## 初步分析結果

秘書處以首批 35 個 IICORE 字符試行初步分析，結果如下：

序號	ISO/IEC 10646 碼位和字形	提交字形	IICORE 字源資料	ISO/IEC 10646:2014 字源資料	見於大型字典*	容易混淆字符	對應繁體字 (簡化字適用)	說明
1	7523 產	產	AJKP	字源地區：5  產 產 產 產 產 G1-327A T3-3745 J0-3B3A K0-5F27 V1-5F47	2	產	-	見於大型字典  字形與“產”字相近，容易混淆。兩字均廣為使用。 7522 產 產 產 生 100.6 GE-3372 HB1-B2A3 T1-5C28
2	5C19 尙	尙	AKP	字源地區：4  尙 尙 尙 尙 GE-2779 T3-2851 K0-5F46 V1-5328	1	尙	-	見於大型字典  字形與“尙”字相近，容易混淆。後者較廣為使用。 5C1A 尙 尙 尙 尙 尙 小 42.5 G0-4950 HB1-A97C T1-4C7E J0-3E30 V1-5329
3	90DE 郎	郎	AKP	字源地區：4  郎 郎 郎 郎 GE-405D T3-3429 J3-7C67 K0-554D	1	郎	-	見於大型字典  字形與“郎”字相近，容易混淆。後者較廣為使用。 90CE 郎 郎 郎 郎 郎 邑 163.6 G0-4049 HB1-ADA6 T1-5367 J0-4F3A V1-6871

\* 1 代表《康熙字典》；2 代表《漢語大字典》

序號	ISO/IEC 10646 碼位和字形	提交字形	IICORE 字源資料	ISO/IEC 10646:2014 字源資料	見於大型字典*	容易混淆字符	對應繁體字 (簡化字適用)	說明
4	6CA1 没	没	AGJ	字源地區：3  没 没 没 G0-433B T3-2659 J0-4857	2	没	沒	見於大型字典  字形與“沒”字相近，容易混淆。 3CC7 沒 沒 水 85.4 G3-534F T6-284A  “沒”的簡化字
5	60A6 悦	悦	AGJ	字源地區：3  悦 悦 悦 G0-5443 T3-3123 J0-3159	2	悦	-	見於大型字典  字形與“悅”字相近，容易混淆。兩字均廣為使用。 6085 悦 悦 悦 悦 悦 心 61.7 GE-2429 HB1-AEAE T1-5550 KO-666D V1-5542
6	654E 教	教	AKP	字源地區：3  教 教 教 GE-2C36 T3-3636 K0-4E67	1	教	-	見於大型字典  字形與“教”字相近，容易混淆。後者較廣為使用。 6559 教 教 教 教 教 支 66.7 G0-3D4C HB1-B1D0 T1-5A74 J0-3635 K2-3770 V1-5826
7	80A4 肤	肤	AG	字源地區：3  肤 肤 肤 G0-3774 T3-2A56 J4-7532	1, 2	-	膚	見於大型字典  本義與“膚”字相同，也是“膚”的簡化字。

\* 1 代表《康熙字典》；2 代表《漢語大字典》

序號	ISO/IEC 10646 碼位和字形	提交字形	IICORE 字源資料	ISO/IEC 10646:2014 字源資料	見於 大型字典*	容易混淆 字符	對應繁體字 (簡化字 適用)	說明
8	7A33 穩		AG	字源地區：2  穩  G0-4E48 TF-4A64	2	穩	穩	見於大型字典  字形與“穩”字相近，容易混淆。後者較廣為使用。  7A4F 穩 穩 穩 穩 禾 115.11 GE-3638 T3-522C J0-323A K2-4E78  “穩”的簡化字
9	57A6 垦		AG	字源地區：2  垦  G0-3F51 T3-2B6D	2	-	墾	見於大型字典  本義與“墾”字相同，也是“墾”的簡化字。
10	62A5 报		AG	字源地區：2  报  G0-3128 T3-2632	2	-	報	見於大型字典  “報”的簡化字
11	76D8 盘		AG	字源地區：2  盘  G0-454C TF-3723	2	-	盤	見於大型字典  “盤”的簡化字
12	7F57 罗		AG	字源地區：2  罗  G0-425E TF-2B4A	2	-	羅	見於大型字典  “羅”的簡化字

\* 1 代表《康熙字典》；2 代表《漢語大字典》

序號	ISO/IEC 10646 碼位和字形	提交字形	IICORE 字源資料	ISO/IEC 10646:2014 字源資料	見於 大型字典*	容易混淆 字符	對應繁體字 (簡化字 適用)	說明
13	6808 棧		AG	字源地區：1  棧 G0-553B	2	-		見於大型字典  “棧”的簡化字
14	917F 釀		AG	字源地區：1  釀 G0-4470	2	-		見於大型字典  “釀”的簡化字
15	52F2 勳		AJ	字源地區：4  勳 勳 勳 勳 GE-2361 T3-4D2C J0-372E K2-2545	2		-	見於大型字典  字形與“勳”字相近，容易混淆。後者較廣為使用。 52F3 勳 勳 勳 勳 勳 勳 力 19.14 G8-2D53 HB1-8EB1 T1-703A J0-522E K0-7D33 V2-8A58
16	970A 靈		AJ	字源地區：4  靈 靈 靈 靈 GE-4452 T3-4F80 J0-4E6E K2-6C3D	2		-	見於大型字典  字形與“靈”字相近，容易混淆。後者較廣為使用。 9748 靈 靈 靈 靈 靈 靈 雨 173.16 G1-4169 HB1-C646 T1-7C71 J0-704D K0-5644 V1-8B5E
17	95D8 鬪		AJ	字源地區：4  鬪 鬪 鬪 鬪 GE-4377 T3-594A J0-462E V3-355B	-		-	字形與“鬪”字相近，容易混淆。後者較廣為使用。 9B2A 鬪 鬪 鬪 鬪 鬪 鬪 鬥 191.10 GE-4661 H-9F5E T3-5D62 J0-722C K0-7763 V1-6C60

\* 1 代表《康熙字典》；2 代表《漢語大字典》

序號	ISO/IEC 10646 碼位和字形	提交字形	IICORE 字源資料	ISO/IEC 10646:2014 字源資料	見於 大型字典*	容易混淆 字符	對應繁體字 (簡化字 適用)	說明
18	9AEA 髮		AJ	字源地區：4  髮 髮 髮 髮 GE-4653 T3-4A72 J0-4831 K2-6F76	-	髮	-	字形與“髮”字相近，容易混淆。後者較廣為使用。 9AEE 髮 髮 髮 髮 髮 形 190.5 G1-7831 HB1-BE76 T1-7021 J0-717B K0-5B25
19	51EA 凧		AJ	字源地區：4  凧 凧 凧 凧 GE-2323 TF-2259 J0-4664 K2-2451	-	-	-	( GE:GB16500-95 主要包括日韓漢字 )
20	96EB 霽		AJ	字源地區：4  霽 霽 霽 霽 GE-4449 TF-392A J0-3C36 K2-6C35	-	-	-	
21	51E6 処		AJ	字源地區：3  処 処 処 GE-227E J0-3D68 K2-244F	1	处	-	見於大型字典  字形與“处”字相近，容易混淆。後者較廣為使用。 5904 処 処 処 処 文 34.2 G0-3426 H-895D T3-2235 K2-2B45
22	7A42 穗		AJ	字源地區：3  穗 穗 穗 GE-3637 T3-4D57 J0-4A66	2	穗	-	見於大型字典  字形與“穗”字相近，容易混淆。後者較廣為使用。 7A57 穗 穗 穗 穗 穗 禾 115.12 G0-4B6B HB1-C14A T1-7454 J0-634F K0-6234

\* 1 代表《康熙字典》；2 代表《漢語大字典》

序號	ISO/IEC 10646 碼位和字形	提交字形	IICORE 字源資料	ISO/IEC 10646:2014 字源資料	見於大型字典*	容易混淆字符	對應繁體字 (簡化字適用)	說明
23	7DD1 綠	綠	AJ	字源地區：3  綠 G1-424C      綠 綠 T3-486B      J0-4E50	2	綠 绿	-	見於大型字典  字形與“綠”字相近，容易混淆。後者較廣為使用。 7DAO 糸 120.8 綠 綠 綠 綠 綠 綠 GE-383C HB1-BAF1 T1-6A3A J3-7A28 K0-5660 V1-626C  字形與簡化字“绿”字相近，容易混淆。 7EFF 彳 120.8 绿 G0-424C
24	5897 增	增	AJ	字源地區：3  增 GE-2625      增 增 T3-4640      J0-417D	-	增	-	字形與“增”字相近，容易混淆。後者較廣為使用。 589E 土 32.12 增 增 增 增 增 增 G0-5476 HB1-BC57 T1-6C41 J3-2F5D K0-7172 V1-5129
25	6483 擊	擊	AJ	字源地區：3  擊 GE-2872      擊 擊 T3-4C35      J0-3762	-	擊	-	字形與“擊”字相近，容易混淆。後者較廣為使用。 64CA 手 64.13 擊 擊 擊 擊 擊 擊 G1-3B77 HB1-C0BB T1-7364 J1-4123 K0-4C2A V2-8C89
26	9EBA 麵	麵	AJ	字源地區：3  麵 GE-486D      麵 麵 J0-4C4D      麵 V1-6D2A	-	麵	-	字形與“麵”字相近，容易混淆。後者較廣為使用。 9EB5 麥 199.9 麵 麵 麵 麵 麵 G1-7859 HB1-C4D1 T1-7A5C J1-6C70 K0-587D

\* 1 代表《康熙字典》；2 代表《漢語大字典》

序號	ISO/IEC 10646 碼位和字形	提交字形	IICORE 字源資料	ISO/IEC 10646:2014 字源資料	見於 大型字典*	容易混淆 字符	對應繁體字 (簡化字 適用)	說明
27	5C02 專		AJ	字源地區：2  專 專 GE-2775 J0-406C	-	專	-	字形與“專”字相近，容易混淆。後者較廣為使用。 5C08 寸 41.8 專 專 專 專 專 專 G1-5728 HB1-B14D T1-5971 J0-5573 K0-6E76 V1-527D
28	596C 獎		AP	字源地區：5  獎 獎 獎 獎 獎 G1-3D31 T3-4648 J0-547D K0-6D61 V1-5159	1	獎	-	見於大型字典  字形與“獎”字相近，容易混淆。兩字均廣為使用。 734E 犬 94.11 獎 獎 獎 獎 獎 GE-332A HB1-BCFA T1-6D64 J0-6052 K2-4731
29	7872 砧		A	字源地區：3  砧 砧 砧 GE-353A J0-4823 K2-4C6A	-	-	-	
30	6D23 沫		CG	字源地區：3  沫 沫 沫 G3-542E T3-2D45 K2-4037	1, 2	-	-	見於大型字典
31	945A 鑽		CJ	字源地區：4  鑽 鑽 鑽 鑽 GE-4349 T3-6055 J0-6F54 K2-6A75	1, 2	鑽	-	見於大型字典  字形與“鑽”字相近，容易混淆。後者較廣為使用。 947D 金 167.19 鑽 鑽 鑽 鑽 鑽 鑽 G1-576A HB1-C670 T1-7D3D J0-6F53 K0-7348 V2-907B

\* 1 代表《康熙字典》；2 代表《漢語大字典》

序號	ISO/IEC 10646 碼位和字形	提交字形	IICORE 字源資料	ISO/IEC 10646:2014 字源資料	見於 大型字典*	容易混淆 字符	對應繁體字 (簡化字 適用)	說明
32	839F 蒼		CJ	字源地區：4  蒼 蒼 蒼 蒼 G3-692F T3-3857 J0-6831 K1-727E	1, 2	-	-	見於大型字典 ( G3:GB7589-87 補充 GB2312 的漢字 )
33	51D6 準		CJ	字源地區：4  準 準 準 準 GE-2278 T3-397E J0-5245 K2-244A	-	準	-	字形與“準”字相近，容易混淆。後者較廣為使用。 6E96 準 準 準 準 準 準 水 85.10 G1-573C HB1-B7C7 T1-646D J0-3D60 K0-715E V1-5D26
34	9AD9 高		CJ	字源地區：2  高 高 GE-464C T4-362D	2	高	-	見於大型字典  字形與“高”字相近，容易混淆。後者較廣為使用。 9AD8 高 高 高 高 高 高 高 189D G0-385F HB1-B0AA T1-586D J0-3962 K0-4D54 V1-6C56
35	7807 砣		CK	字源地區：2  砣 砣 G5-577A T3-2E3B	1, 2	-	-	見於大型字典  ( G5:GB7590-87 補充 GB2312 的漢字 )

\* 1 代表《康熙字典》；2 代表《漢語大字典》