



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

澳門特別行政區

資訊系統中文編碼統一方案

總體文件

2021年1月



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

目 錄

序言	iii
第一部分 概論	1-1
1.1 《編碼統一方案》簡介	1-1
1.2 《大五碼字符集》及《香港增補字符集》簡介	1-3
1.2.1 概述	1-3
1.2.2 大五碼簡介	1-3
1.2.3 《香港增補字符集》簡介	1-4
第二部分 《編碼統一方案》的內容及編碼架構	2-1
2.1 《編碼統一方案》的內容	2-1
2.2 《編碼統一方案》的編碼架構	2-1
第三部分 《澳門增補字符集》的內容	3-1
第四部分 《澳門增補字符集》的擴展機制	4-1
4.1 漢字擴展機制一：縱向擴展	4-1
4.2 漢字擴展機制二：橫向擴展	4-1
4.3 漢字擴展機制三：註冊 IVS	4-2
4.4 符號擴展機制	4-4
第五部分 推行計劃	5-1



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

序言

為配合澳門特別行政區電子政務的發展，使各公共部門和實體能更準確及高效地以中文進行電子通訊，資訊系統中文編碼工作小組（下稱“工作小組”）制訂了《澳門特別行政區資訊系統中文編碼統一方案》（下稱“《編碼統一方案》”），該方案推出後，將統一各公共部門和實體的中文資訊編碼，作為公共部門間數據互聯的其中一項共通編碼標準，有利於推出更多跨部門的電子化服務，提升整體公共服務的質素和效率。

《編碼統一方案》由三個字符集組成：《大五碼字符集》、《香港增補字符集》及《澳門增補字符集》。多年來澳門各公共部門和實體的資訊系統已採用《大五碼字符集》及《香港增補字符集》，《澳門增補字符集》作為上述兩個字符集的增補，收錄了澳門特別行政區各公共部門需要使用的、與其他部門有流通需求的，但未收納於上述兩個字符集的字符。



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

隨著《編碼統一方案》的推出，使業界主流的系統供應商在支持澳門特別行政區的本地化技術時有統一的字符編碼標準，亦使各公共部門的資訊系統能更全面地採納及支持 ISO/IEC 10646 國際編碼標準。

《澳門增補字符集》設有字符增收機制，字符的增收須按《澳門增補字符集 – 增收原則及程序》為之。增收原則規範了提交的字符可予考慮增收、不予增收的原則，並根據 ISO/IEC 10646 國際編碼標準的認同規則作為參考依據；增收程序方面，制訂了公共部門自查、工作小組的技術跟進小組初步審議、工作小組審定結果的三層審核機制，以確保增收程序的嚴謹性和執行效率。

根據第 299/2018 號行政長官批示核准的《互操作指引規章》第一條第（一）款的規定，《編碼統一方案》為《互操作指引規章》“數據的格式”中“字元碼”的規範標準，各公共部門和



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

實體應按照工作小組發出的指引，執行和實施《編碼統一方案》標準。

《編碼統一方案》的內容詳述於本文件內，並有以下的配套文件及字型檔：

- 《澳門增補字符集 - 2020》字符表
- 《澳門增補字符集 - 增收原則及程序》
- 《資訊技術部門工作指引》
- 《字符增收部門工作指引》
- 《澳門增補字符集 - 2020》字型檔、輸入法及安裝指引
- 《常見問題集》
- 《資訊系統中文編碼統一方案術語表》

上述的文件及字型檔請於《編碼統一方案》網站參閱或下載。



第一部分 概論

1.1 《編碼統一方案》簡介

《編碼統一方案》旨在為澳門各公共部門、實體提供統一的中文編碼方案，該方案包括以下三個字符集，涵蓋澳門現行系統所使用的漢字、有流通需要的葡文符號及其他符號：

- (1) 《大五碼字符集》：這些字符非澳門特有，但澳門的電腦系統已實際使用；
- (2) 《香港增補字符集》：此字符集補充了大五碼字符的不足，方便了香港的中文資訊交換。由於澳門與香港毗鄰，溝通緊密，為了支持兩地資訊系統的互通性，《編碼統一方案》亦包括《香港增補字符集》；
- (3) 《澳門增補字符集》：是對《大五碼字符集》及《香港增補字符集》的增補，收錄澳門特別行政區各公共部門需要使用的、與其他部門有流通需求的，但未收納於上述兩個字符集的字符。

《編碼統一方案》屬編碼字符集，而非字形標準。《編碼



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

《統一方案》包括的三個字符集中，澳門特別行政區政府僅負責維護《澳門增補字符集》中字符的字形。原則上，由於澳門特別行政區政府對《大五碼字符集》、《香港增補字符集》無所有權，因此不維護其全部字符的字形。¹ 但是，為異體字註冊 IVS 時，也須註冊相應的基本字。實際上，澳門特別行政區政府也須維護《大五碼字符集》、《香港增補字符集》中的基本字，因此也對這些基本字進行橫向擴展。

《編碼統一方案》主要是規範各公共部門進行資訊交換時在 Unicode 編碼架構下的編碼標準，對部門內部系統的編碼²，《編碼統一方案》不作規範。

¹ 在實際應用中，對於《大五碼字符集》和《香港增補字符集》的字符，目前市場上有通用的字型產品。

² 比如 Unicode 系統中，部門內部使用的私人使用區碼位，或是部門內部使用的非 Unicode 系統。



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

1.2 《大五碼字符集》及《香港增補字符集》簡介

1.2.1 概述

由於澳門和台灣、香港一樣，都是使用繁體中文，書寫系統與台灣、香港接近，因此，澳門的電腦平台已經沿用《大五碼字符集》及《香港增補字符集》的字符。所以，《編碼統一方案》涵蓋《大五碼字符集》及《香港增補字符集》的漢字和符號，以滿足澳門電腦平台的使用需要。

1.2.2 大五碼簡介

大五碼（Big-5）是由台灣財團法人資訊工業策進會設計的中文共通內碼，1984年，該會發行了大五碼的正式文獻，即《電腦用中文字型與字碼對照表》。之後，大五碼作為工業標準在台灣及其他使用繁體中文的地區普遍使用。

《大五碼字符集》包括 13,053 個漢字和 441 個符號（包括 33 個中文傳輸控制碼）。《大五碼字符集》中的漢字分為常用字集和次常用字集，其中，常用字集包括 5,401 個漢字，編碼



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

區間為 Big-5 0xA440 至 Big-5 0xC67E；次常用字集包括 7,652 個漢字，編碼區間為 Big-5 0xC940 至 Big-5 0xF9D5。

1.2.3 《香港增補字符集》簡介

為了方便市民和政府部門使用中文進行電子通訊及資料交換，香港特別行政區政府於 1999 年 9 月正式公佈了《香港增補字符集 - 1999》，共收錄 4,702 個字符。之後，又陸續公佈了更新版，即《香港增補字符集 - 2001》、《香港增補字符集 - 2004》、《香港增補字符集 - 2008》。香港特別行政區政府於 2017 年 5 月公佈了《香港增補字符集 - 2016》，在《香港增補字符集 - 2008》的 5,009 個字符的基礎上，通過對 ISO/IEC 10646 進行橫向擴展新增 23 個漢字和 1 個歐元符號（€U+20AC）。



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

第二部分 《編碼統一方案》的內容及編碼架構

2.1 《編碼統一方案》的內容

《編碼統一方案》包括以下三個字符集：

- (1) 《大五碼字符集》；
- (2) 《香港增補字符集》；
- (3) 《澳門增補字符集》，包括：
 - (a) 對 ISO/IEC 10646 進行縱向擴展的新漢字及 ISO/IEC 10646 內未編碼、由澳門地區提出申請新編碼的符號；
 - (b) 對 ISO/IEC 10646 進行橫向擴展的漢字及 ISO/IEC 10646 內已編碼、《澳門增補字符集》增收的符號；
 - (c) 註冊 IVS 的異體字。

2.2 《編碼統一方案》的編碼架構

在 ISO/IEC 10646 國際編碼標準的框架下，《編碼統一方案》的編碼方式如下，在 ISO/IEC 10646 中，該編碼將作為澳



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

門地區漢字及符號的字源資料 (Source Reference)。

- (1) “MB-*hhhh*”：代表大五碼字符，並使用“MB0-*hhhh*”、“MB1-*hhhh*”和“MB2-*hhhh*”分別代表大五碼符號、常用字集、次常用字集，其中“*hhhh*”是對應字符在大五碼的碼位；
- (2) “MA-*hhhh*”：代表《香港增補字符集》的字符，“*hhhh*”是對應字符在《香港增補字符集 - 2008》的大五碼碼位。由於《香港增補字符集 - 2008》是《香港增補字符集》大五碼編碼部分的最後版本，之後，香港只繼續為《香港增補字符集》新增收的字符提供 ISO/IEC 10646 編碼，因此，對於《香港增補字符集 - 2008》之後新增收的字符，《澳門增補字符集》將根據澳門的實際使用情況相應地增收，並根據增收方式進行編碼；
- (3) “MC-*nnnnn*”：代表對 ISO/IEC 10646 進行縱向擴展的新漢字及 ISO/IEC 10646 內未編碼、由澳門地區提出申請新編碼的符號，其值“*nnnnn*”在 00001-99999 順



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

序編號；

- (4) “MD-*hhhh[h]*”：代表對 ISO/IEC 10646 進行橫向擴展的漢字及 ISO/IEC 10646 內已編碼、《澳門增補字符集》增收的符號，“*hhhh[h]*”是對應字符在 ISO/IEC 10646 的碼位³。

由於《香港增補字符集 - 2008》是《香港增補字符集》大五碼編碼部分的最後版本，對於《香港增補字符集 - 2008》後續版本中橫向擴展的字符，《澳門增補字符集》亦根據實際需要進行橫向擴展，並通過在“MD”後加標識符“H”來表示該字符為根據《香港增補字符集》相應進行橫向擴展，也就是說，使用“MDH-*hhhh[h]*”對這些字符編碼；

- (5) “ME-*hhhh[h]-nnn*”：代表註冊 IVS 的異體字，“*hhhh[h]*”是對應字符在 ISO/IEC 10646 的碼位⁴。
“*nnn*”是從 001 開始的順序編號，用作區分一個漢字

³ 如字符在 ISO/IEC 10646 的基本平面，編碼為 4 位；如在其他平面，則為 5 位。

⁴ 如字符在 ISO/IEC 10646 的基本平面，編碼為 4 位；如在其他平面，則為 5 位。



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

所可能對應的多個異體字。在為異體字註冊 IVS 時，亦須對其基本字註冊 IVS，由於這些基本字已包括在第（1）至（4）部分中，因此其編碼採用第（1）至（4）部分的編碼方式，“ME-*hhhh[h]-nnn*”僅用於註冊 IVS 的異體字。



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

第三部分 《澳門增補字符集》的內容

《澳門增補字符集》是對《大五碼字符集》及《香港增補字符集》的增補，收錄澳門特別行政區各公共部門需要使用的、與其他部門有流通需求的，但未收納於上述兩個字符集的字符（含漢字及符號），包括三個部分：

- (1) 對 ISO/IEC 10646 進行縱向擴展的新漢字及 ISO/IEC 10646 內未編碼、由澳門地區提出申請新編碼的符號；
- (2) 對 ISO/IEC 10646 進行橫向擴展的漢字及 ISO/IEC 10646 內已編碼、《澳門增補字符集》增收的符號：
如根據 ISO/IEC 10646 的認同規則，澳門使用的漢字與已編碼字符認同，不能分開編碼，則通過橫向擴展的方法補充 ISO/IEC 10646 中的 M-列，保留澳門的特有字形；澳門因流通需要使用的符號如不在《編碼統一方案》內，也可增收到《澳門增補字符集》；
- (3) 註冊 IVS 的異體字：這些亦是澳門使用的、ISO/IEC



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

10646 已編碼的字符。雖然字形與已編碼字符有差異，但根據 ISO/IEC 10646 的認同規則不能分開編碼，通過註冊 IVS，可保留澳門特有的字形，同時亦能使用已編碼的字符。



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

第四部分 《澳門增補字符集》的擴展機制

4.1 漢字擴展機制一：縱向擴展

澳門特別行政區政府在服務過程中為支持互通性實際使用的漢字，如在 ISO/IEC 10646 內無編碼，且根據 ISO/IEC 10646 認同規則並無與之認同的字，則通過縱向擴展的方式向 ISO 轄下的表意文字小組（Ideographic Research Group，簡稱 IRG）申請增收，以期獲得新編碼。

4.2 漢字擴展機制二：橫向擴展

澳門特別行政區政府在服務過程中為支持互通性實際使用的漢字，如未納入《編碼統一方案》，且根據 ISO/IEC 10646 認同規則，在《編碼統一方案》中無與之認同的字，但在 ISO/IEC 10646 內已編碼或有與之認同的字，則可考慮將此字納入《澳門增補字符集》，並通過橫向擴展的方式在 ISO/IEC 10646 中補充澳門字形。

原則上，澳門特別行政區政府僅負責維護《澳門增補字符



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

集》中字符的字形，由於對《大五碼字符集》、《香港增補字符集》無所有權，因此不維護這兩個字符集的全部字形。但是，為異體字註冊 IVS 時，也須註冊相應的基本字。實際上，澳門特別行政區政府也須維護《大五碼字符集》、《香港增補字符集》中的基本字，因此也對這些基本字進行橫向擴展。

4.3 漢字擴展機制三：註冊 IVS

受限於認同規則之下，傳統編碼機制不容許異體字分開編碼，因此對字型廠商無規範其所應遵從。Unicode Consortium 開發了一套為漢字所建立的字形定義技術，在增加了一組用於組合用字的異體選擇符（Variation Selectors）的基礎上，可用表意文字異體字序列（Ideographic Variation Sequence，簡稱 IVS）來定義異體字，並通過註冊添加到表意文字異體字數據庫（Ideographic Variation Database，簡稱 IVD），為須處理異體字的電腦系統提供一套符合統一碼標準而又可互用的方案。



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

IVS 表示異體字的具體格式是一個基本字 (Base Character)⁵ 後面加上一個異體選擇符。IVS 和 IVD 為過去在漢字認同規則下不能編碼的異體字進行編碼提供了解決方案，通過由基本字及異體選擇符組成的異體字序列使未編碼的異體字可與已編碼的基本字相互關聯並編碼，再經 Unicode 的 IVD 註冊，使這些異體字可在 ISO/IEC 10646 體系下編碼。每個表意文字字符最多可關聯 240 個異體字字形。例如，澳門使用的“姪”、“姪”兩字，均與 ISO/IEC 10646 中已編碼、且收錄在《香港增補字符集 - 2008》的“姪” (U+36C7) 認同，為了保留澳門須使用的字形，對已編碼的基本字“姪” (U+36C7) 和兩個異體字“姪”、“姪”註冊 IVS，使這三個字形均可應用於澳門電腦系統。由於該基本字已收錄在《香港增補字符集 - 2008》，因此，在《澳門增補字符集 - 2020》中，基本字的字源資料 (Source Reference) 為“MA-9856” (“9856”是該字符在《香港增補字符集 - 2008》的大五碼碼位)，兩個異體字的字

⁵ 是一個已編碼的中日韓表意文字。



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

源資料 (Source Reference) 分別為“ME-36C7-001”、“ME-36C7-002”。

4.4 符號擴展機制

澳門電腦平台中有流通需要的符號，如在 ISO/IEC 10646 內無編碼，亦可考慮納入《澳門增補字符集》，並申請新編碼。

如符號在 ISO/IEC 10646 內已編碼，但未包括在《編碼統一方案》內，亦可增收到《澳門增補字符集》。由於符號不涉及字形問題，因此可直接使用 ISO/IEC 10646 中的符號。例如，共有 28 個葡文符號在澳門有流通需要，而 13 個已收錄在《香港增補字符集 - 2008》中，因此可直接使用，另外 15 個符號則增收到《澳門增補字符集 - 2020》。

符號的字源資料 (Source Reference) 編碼方式和漢字相同。



第五部分 推行計劃

工作小組已進行並完成以下一系列推行《編碼統一方案》的配套措施，包括：

- (1) 草擬及制訂《編碼統一方案》文件；
- (2) 編訂《澳門增補字符集 - 2020》；
- (3) 外判製作《澳門增補字符集 - 2020》字型檔及輸入法；
- (4) 《編碼統一方案》技術實現的測試及整備；
- (5) 完善《澳門增補字符集 - 增收原則及程序》的機制；
- (6) 為工作小組部門的相關人員進行培訓；
- (7) 建立電子平台，公佈相關訊息及協助工作小組運作；
- (8) 向 IRG 及 Unicode Consortium 提交字符申請。

基於《編碼統一方案》的目的是作為澳門公共部門之間數據互聯的其中一項共通編碼標準，使各部門能準確及高效地以中文進行電子通訊，根據第 299/2018 號行政長官批示核准的《互操作指引規章》第一條第（一）款，由於《編碼統一方



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

案》的標準是屬於“數據的格式”中“字元碼”的規範，因此，《編碼統一方案》將作為《互操作指引規章》數據格式的其中一項規範標準，各公共部門和實體應按照工作小組發出的指引，執行和實施《編碼統一方案》標準。

於《編碼統一方案》公佈後，將分為下列兩階段實施：

- (1) **字符增收機制**：公佈後即實施，以避免各部門的造字不斷增加，並由工作小組統一分配字符的編碼。
- (2) **部門之間使用統一編碼標準進行資訊交換**：公佈首年後實施，主要基於以下兩方面的原因（但不影響已具備條件的部門，可即時應用該統一編碼標準）：
 - (a) 大部分部門需時進行資訊技術的準備及過渡工作，包括：審視及疏理現時的造字、於應用系統測試《澳門增補字符集》字型檔及輸入法、升級或調整現時的應用系統、設置部門的編碼轉換表等，工作小組將提供適當的技術支援。
 - (b) 業界主流的系統供應商將於《編碼統一方案》



澳門特別行政區政府
Governo da Região Administrativa Especial de Macau
行政公職局
Direcção dos Serviços de Administração e Função Pública

公佈後落實技術支持方案，工作小組將促進其
實現技術支持方案。

~~~ 完 ~~~