

# 香港教育學院

## 資訊與應用科技系

### 中級程度

#### I7 Chinese Character Input

#### 中文輸入

資訊與應用科技系：黎耀志

教育學院學生：吳韻華

---

# 課程目錄

5. 漢字編碼.....	1
5.1 大五碼.....	1
5.2 GB.....	1
5.3 「香港增補字符集」與「標準香港字庫開發計劃」.....	2

## 5. 漢字編碼

漢字的信息處理系統與通訊系統之間進行漢字信息交換的代碼稱為漢字信息交換碼，為了使漢字信息的交換與信息處理得到穩定的應用，所以有必要把漢字信息交換碼加以統一，而現在在互聯網上用以把漢字信息傳遞的漢字信息交換碼多為大五碼(BIG5)和 GB(GB2312)。

大五碼是用以傳遞繁體漢字的信息，它多應用於香港和台灣地區；而 GB 則是用於簡體漢字的信息傳遞，多應用於中國內地。

### 5.1 大五碼

大五碼的漢字是由兩個位元組 (byte) 所組成。其中第一位元組的值在 16 進制的 A0 ~ FE 之間，第二位元組則在 40 ~ 7E 及 A1 ~ FE 之間。其餘的 A040 ~ A0FE、C6A1 ~ C8FE、FA40 ~ FEFE 為空白區域，經常用作用戶造字區。

編碼的範圍如下：

第一位元組	第二位元組	說明
A1~A3	40~7E,A1~FE	標點、符號、注音符號(408 個)
A4~C6	40~7E,A1~FE	中文常用字(5401 個)
C9~F9	40~7E,A1~FE	中文次常用字(7652 個)

### 5.2 GB

GB 2312 收錄簡化漢字及一般符號、序號、數字、拉丁字母、日文假名、希臘字母、俄文字母、漢語拼音符號、漢語注音字母，共 7445 個圖形字符。其中漢字以外的圖形字符 682 個，漢字 6763 個。

GB2312 碼是中華人民共和國國家標準漢信息交換用編碼、全稱「信息交換用漢字編碼字符集基本集」。

## 5.3 「香港增補字符集」與「標準香港字庫開發計劃」

現在香港的互聯網多使用大五碼作為漢字信息傳遞的交換碼，但由於香港本地文化的特別，而以上兩個交換碼的系統是以中國與台灣的需要而設計，因此一些香港常用但獨有的漢字並未包含其中，因此在香港的互聯網上觀看網上的信息時，會出現亂碼或缺字的情況。

有見及此，香港政府便製定一套「香港增補字符集」，它的目的是收納香港特區政府及市民在中文電子通訊中有需要使用的字符，來補充現時大五碼的編碼標準內並未包含的字符，以作為一個共通的中文界面，方便大家能準確地以中文進行電子通訊。



「香港增補字符集」網頁

另外設立「標準香港字庫開發計劃」，目的是製作一個統一及全備的香港特殊中文字表，給香港各出版機構、互聯網服務供應商和電子媒介讀者等使用，使出版業以致讀者可以運用一套統一和最新的香港字庫。「標準香港字庫開發計劃」可以避免發生電子出版「缺字」的難題，並使香港出版機構和讀者可以透過電子數碼媒介交換資訊。此外，這個計劃亦會特別設計一些電子濾字系統(Electronic Fonts Filter System)，讓出版機構可以把原有的檔案融入這個香港特殊中文字表。



「標準香港字庫開發計劃」網頁



## 網上資源

有關「香港增補字符集」和「標準香港字庫開發計劃」的詳細資料可到以下網址瀏覽：

「香港增補字符集」：<http://www.info.gov.hk/gccs/>

「標準香港字庫開發計劃」：<http://www.5c.org/chinesef.htm>