

零壹中文電腦叢書之5

中文資訊碼對照表

—字母碼、倉頡碼、內碼、三角碼、
神通碼、主計碼、IBM碼、電報碼—

字	字母碼	倉頡碼	內碼	三角碼	神通碼	主計碼	IBM碼	電報碼
晶	AAA	AAA	084001	616161	02157	34A9	6371	9561
間	AABUU	AABUU	0844AA	E16	388754	01267	40FB	2699
醒	AAHM	AAH	IOYV	CE7568				
道	YYWS	IOYWS	CE79C8	382163	07819	2CED		



零壹科技股份有限公司編著



全華科技圖書股份有限公司印行

零壹中文電腦叢書之5

中文資訊碼對照表

—字母碼、倉頡碼、內碼、三角碼、
神通碼、主計碼、IBM碼、電報碼—

字	字母碼	倉頡碼	內碼	三角碼	神通碼	主計碼	IBM碼	電報碼
日	A	A	080000	610000	00556	040B	1378	248
晶	AAA	AAA	084001	616161	02157	34A9	6371	9561
疊	AAAM	AAAM	08413A	61106	157	1AE3		1
靦	AABUU	AABUU	0844AA	5E16	38A4	1A67	40FB	2699
晧	AAFU	AAFU	084			1477		9555
曜	AAHM	AAHM		IOYV	CE7568			
暇	AAIE	AAIE		IOYVW	CE77A5			
明	YYVW	YYVW	IOYVW	CE77A5				
道	YYWS	YYWS	IOYWS	CE79C8	38A4	1A67	20ED	66



零壹科技股份有限公司編著



全華科技圖書股份有限公司印行



全華圖書 版權所有 翻印必究

局版台業字第0223號 法律顧問：陳培豪律師

中文資訊碼對照表

字母碼、倉頡碼、內碼、三角碼
神通碼、主計碼、IBM碼、電報碼

零壹科技股份有限公司
電話：(02)7992897~8

出版者 全華科技圖書股份有限公司
北市龍江路76巷20-2號
電話：581-1300・541-5342
581-1362・581-1347
郵撥帳號：100836

發行者 陳 本 源
印刷者 華一彩色印刷廠
定價 新台幣 280 元
再版 中華民國72年12月

感謝您

感謝您選購全華圖書！

希望本書能滿足您求知的慾望！

圖書之可貴在其量也在其質

量指圖書內容充實、質指資料新穎够水準，我們就是本著這個原則，竭心盡力地為國家科學中文化努力，貢獻給您這一本全是精華的全華圖書。

序 言

國家興亡，匹夫有責，值此知識爆炸的時代，資訊處理效率的後果，已不僅是國家的興亡而已。民族文化，思想傳統都面臨絕滅的挑戰。

軍事政治以及經濟上的危機，不論是誰都能切身感受，但文化傳統却不然，其重要性有如日光空氣，人類生存不能一日缺乏。正因如此，一般人在日常生活中無法體認其價值。

人之生存，絕非僅僅物質可以滿足，至少，在基本的需求解決後，繼之而起的就是精神層次的追求。自初民起，這種精神層次的精華，逐日累積孕育成爲思想及概念，加上文字符號的記載傳播，源遠流長，我們稱之爲文化，也就是我們精神生活的根源，處世接物，思考行爲的準則。

近百年來，雖然我們物質條件落後，但仍能保持和諧的身心，實拜前人遺教之助。因之，如果希望後世仍然能像我們一樣沐浴在博大精深的中華文化之中；如果我們相信未來的人類能因古聖先哲的睿智而蒙受恩澤，我們就應該面對時代，承先啟後，將中國文字的優點發揮在資訊工具上，讓中華文化永垂不朽。

本公司矢志於中文資訊研究，完全著眼在獻身文化的立場，既無任何公、私財團之支援，也沒有知名之學者專家參與，草路濫縷，歷經艱辛。雖然現今採用我們輸出入方法的系統不下數十家之多，但本著前述宗旨，僅以最低售價，供各界採用。

研究機構如無後盾，很難生存下去，本公司也不例外，幸而同仁們基於對家國的熱愛，對文化的信念，無計個人得失而默默地盡心竭力，始能勉強維持。這次編印本書純屬服務，原已見肘的困境，更是難堪，以致人事星散，剩下的全部同仁爲了迎接七十一年資訊展，以不眠不休的精神，剋服萬難，堅持到底，由十一月初開始動員，到廿九日才告完成，再交全華付印，其中疏漏錯誤唯有懇求各界諒解與指正。

當今經濟掛帥，我們的價值觀或爲人所不屑，但這正是中華文化有待考驗之處，自古即有顏回之簞食瓢飲，只要我們能够生存，我們也就不改此樂。

爲表示負責，茲將發展系統及參與本書作業同仁姓名列後

系 統：倉頡中文電腦001 (MCZ改裝)

印字機：OKIDATA SL125

總編輯：沈紅蓮

程 式：熊黎民 林曉星 姜有謨

系 統：林嘉勳

校 對：蔡東龍 林福星 陳秋麗 汪蘭英

企 劃：董魯生

建 檔：全體同仁

總 務：劉世文

零壹公司 朱邦復誌於71年11月29日

于台北內湖

一. 各 資 訊 碼 之 選 取

中文輸入方法甚多，以本公司人力財力自不可能全部收齊，加以很多資料未見公開，是以僅能就各家已經有商品上市，或曾經發佈之資料著手，計有本公司之中文字母碼，倉頡字母碼，（以康熙字典為藍本），王安公司之三角號碼，主計處之中文資訊交換標準碼，IBM 碼，神通碼，電報明碼等，茲分別解說如次：

1. 中文字母碼：

係以倉頡字母改進第三代之產品，由於筆者最初研究中文輸入時，僅以能解決常用字之應用為目的，所以只選取國語日報大辭典所收之字彙為對象，但一旦擴充字彙之容量，即面臨重覆字增加的難題，如丿及衣部同置卜鍵，在國語日報辭典中，丿部僅有20字，衣部有87字，其重覆字僅僅2個。現以康熙字典為根據，丿部為112字，衣部有455字，結果取碼重覆之字數提高為35字，足証原安排之不妥當，故將衣部調入“中”鍵。

在編碼者的立場既要簡單方便又要合情合理，以往我們爲了推廣，力求做到這兩者，然而爲了中文長遠的發展，字彙的包容能力以及重覆字率的重要性就相對提昇。因此，改碼雖然不利於市場，但爲了責任，我們不得不將倉頡碼作了適當的小量修正。

2. 倉頡字母碼：

倉頡碼目前使用系統甚多，尤以家用及商用之微型電腦，以及中文端末機。基於成本之考慮，絕大部份採用本法，如果不涉及中國文化之整理以及全部的人名建檔，本法足能應付裕如。

此外，倉頡與中文字母碼最基本的差異是在資料結構上，由於取碼不同，字形產生器中的結構就有很大的分別。換句話說，要收容相同數量的字彙，用倉頡碼要比中文字母碼多十分之一的空間。目前，用64KB的頁區，已到極限。僅管再擴充至128K並無困難，但是對目前很多小系統形成太重的負擔。因此我們建議，倉頡碼儘量放心地應用在工商企業以及一般的應用上，如果要兼顧圖書館等大量資料處理，我們將設法用軟體程式來解決各種轉換的問題。

3. 三角號碼：

“數字化中文字彙”係胡立人，張源渭，黃克東三位教授合作發明的編碼方式，由王安公司推廣發行，曾是中文輸入系統中最受歡迎者，其唯一的缺點是未能解決中文字形輸出的問

題。對微電腦系統而言，無法加以利用。不論如何，本輸入法之優點甚多，且輸入碼等長，可視作內碼使用。

4. 中文資訊交換標準碼：

係行政院主計處於民國70年2月所公佈之資料，主計處是我國主管電腦的行政機構，故主計碼當然有其重要性，只是主計碼屬於交換性質，與輸入輸出無關。故使用者需自建一龐大之對照表，以與其他不同系統交換訊息。

5. IBM 碼：

IBM 雖執世界電腦界之牛耳，但在中文電腦上，並非有資金就能解決問題。至少，中文電腦應由中國人來設計規劃，目前IBM 採用大鍵盤，字形全部來自日本，這次採用之資料得自年初，據說已有16000 字之新碼，但經多方探尋，仍無法得到

6. 神通碼：

神通電腦公司為國內首先發展中文系統之一，且卓著成效，但初期採用大鍵盤，成本過高，近改用小鍵盤，唯其輸入方式與輸出字形無關，故必須用表轉換，且在輸入時需多次複選，不適用於大量資料輸入，且無自然排序功能。

7. 電報明碼：

電報明碼是在民國元年由政府頒佈的標準，當時只為電報通訊之用，沿用迄今已成為資碼中採用最廣泛的一種，但其中收字不全，傳將擴充為16000 字，因一萬碼以前各系統早已收齊，本次編碼係採用研議中之版本，尚未定案，茲為中文資訊

前途，廣徵博引，未得授權，先行刊登，一切責任本公司自負與他人無涉。

8. 綜合統計：

由於中文字庫有組字功能，在理論上，組合字有 $25 \times 25 \times 25$ 種之多，茲假設為 N ，單字首有24種，計 $24N$ ，複字首有105種，計 $105N$ ，圖形符號3種，為 $3N$ 外加固定字形6080個，全部總計共2,068,630個。在這些字形中，未被採用者佔百之九九，是以唯有收集別家字形，加以對照，才有意義。這次所收家數不多，已有33,200字，概估以現有各系統及字典計，應有四萬字，如得再加增訂，可望收全。

再由各家所收之字，可見差別甚大，很多不知出自何典，本字庫仍有空間限制，無法全部納入，茲統計如下：

◆◆	中文字母碼	三角碼	神通碼	主計碼	IBM碼	電報碼
原有字數	2,068,630	10,332	10,883	15,980	8,381	15,241
表列字數	33,200	10,332	17,402	15,155	8,195	14,971

◆◆註：因三角碼全係常用字，故已全部收齊，IBM碼雖在常用字區，但日式異體字甚多，故未收全，原主計要有16,290個，其中符號120個未計，故僅列入15980個。

二. 中文字庫之發展歷程

~~~~~

中文字庫之原始構想始於民國62年，當筆者旅居巴西時，目睹拼音文字應用之功效，決心回國發展。迨至66年五月，始獲大周建設公司之贊助，出版“中文輸入研究”乙書，次年，得當時三軍大學校長蔣上將緯國將軍之賞識，義務性地為國軍研展“倉頡一號”電傳系統，但未受採用。67年10月若干友人斥資組成豪邦電腦公司，與筆者之零壹科技顧問公司合作開發中文電腦，惜因資金不足，豪邦中途解散。時宏碁公司得悉本計劃，幾經磋商，同意協助發展，並應用在其所代理之MC2系統上，於69年9月推出問世，定名為“天龍中文電腦”，是為中文字庫之第一代產品。

天龍中文電腦的劃時代成就證明中文之編碼及字形組合切實可行，但字形之美觀及字彙之收容去理想尚遠，本公司遂獨力繼續研究發展。

民國70年資訊展，本公司利用全亞PA8000改裝成“倉頡8000中文電腦”，其中字形產生器為第二代，而且加上中文初級程式語言，深令資訊界振奮，先後有十餘家廠商與本公司簽約，將其系統加裝中文功能，自是中文電腦才日漸普及。

年來本公司不斷突破創新，已完成第三代24\*24明體字庫，收容字彙高達35000個，組合字形2000000個以上，較採用日本方式貯存字形空間節省達四十倍之多，且適用於任何機種。至於更高密度以及各種字體之字形俱在發展中，可供打字機及排版之用。

在輸入技術上，第一代及第二代我們定名為倉頡字母輸入法，第三代由於字母的安排位置略有改變，(其原委詳見下章)，為示分別特更名為中文字母輸入法。在當初第一代推出之時，我們已經發覺有改碼的必要，然因商業產品推廣不易，遽然改碼將徒增用戶困擾，現經兩年努力，第二代產品已足敷一般資料處理之用，但為整個中文資訊的未來發展以及圖書館的應用，古籍整理等設想，如不改碼，勢難應付實際需求。謹此，我們懇求各界的諒解與支持。

### 三 中 文 字 母 之 應 用 功 能



在資訊處理上，不外乎傳輸，處理，列印，貯存。由於英文系統應用於先，全世界之既有設施俱已定型，是以欲求便於應用，絕不能與英文字母之結構相違。因之，中文字母即以保持中文之特性以及配合英文之架構作各種角度之配合考慮，並以經濟實用為唯一前提。(附加生產成本目前約新台幣三千元，如大量使用可降至一千元以下。)

#### 1. 傳輸：

在傳統之電傳技術上，僅採用五位元之方式，現亦有改用八位元之趨勢，但由於標準之制定受制於外人，故最安全可靠之方式，仍然應與ASCII 架構相符。中文字母法係利用ASCII 之41H 至5AH 為輸入碼，故只要加以辨識碼，即可與任何現有之傳輸系統通用。

## 2. 處理：

### 2.1. 排序：

資料處理中最重要的一項因素便是排序的功能，所謂排序，就是文字編碼的邏輯順序。最理想者，要達到電腦與人腦之順序合一，也就是說，輸入碼之邏輯理念必需讓電腦接受。目前我們能應用電腦，是基於電腦二進位的數序能與人的十進位數序作直接的轉換，而且ASCII碼的制定，即根據二進位理論加以排列。我們可以說：拼音文字系統能領先圖形文字系統之唯一因素，就是其字母本身具備排序之規則。

由於中文無順序排列之法則，前人幾經摸索研究，始建立字首及筆劃的系統。然而應用起來異常困難，尤其在一般微型電腦上，除非只用限定的某些字，否則就得採用與輸入碼毫不相關且擴充有限的內碼。（仍然需要龐大的轉換對照表）

中文字母碼為採用拼音字母順序觀念，配合中文字形結構而成。由於字母具有順序，是以各字碼秩序井然，與拼音字母相較，在排序功能上，毫無差別。

### 2.2. 處理：

由於中文字母法之輸入碼即為輸出字形之唯一代號，故毋需轉換表，在系統中處理極為方便。又由於輸入碼可經由一公式轉換為三字元之內碼，如需等長度之場合，亦可任意採用。

2.3. 中文字形之供應為一64KB之頁區，可經由輸出埠或跳頁技術所控制，任何微電腦皆可使用，且不需修改系統。加以空間小，電流消耗，提供之字數，舉世無匹，組字速度利用16位元技術，可望在1MS 以下。

2.4. 中文字形產生器具有多種功能如次：

2.4.1. 二百萬個組合字形可供商業及軍事入密，以及新創字之選擇，簡縮字之代碼。

2.4.2. 具有造字機能，平均每字只需8 字元，較存點矩陣約少四倍。

2.4.3. 具有大量之圖案圖形，供各種特殊場合用。如同文字處理之花邊，表格，以及圖表設計。且各種圖案皆有輸入規則，易記易用。

2.4.4. 另有圖形語言，以供圖形之設計。

### 3. 內碼：

為求應用方便，令字碼輸入後立即轉換為唯一對應之內碼，且具有相同之順序。其目的不僅為求節省空間，兼顧螢幕編輯以及多種文字夾雜應用之功能，更重要的是必需提供最大彈性的擴充能力。準此，中文字母法採用三字元之內碼。(註：若採用其他方式之內碼亦可自建內碼轉換表。)

### 3.1. 內碼轉換式：

由於中文字母碼最大長度為五碼，每碼序位佔五個位元，共計為廿五個位元。以三字元尚不足一位元，因之，前二碼各取五位元，共為十位元，後三碼則用公式轉為二進位碼形式之25倍碼。即可轉換為14位元，加上前十位元，共計24位元，恰為三字元，是為內碼。

在轉換中，輸入碼位之排列係以第一碼為基準，意即凡不足五碼者，需將輸入碼後面之空碼視為零。此法與數字觀念相反，但與文字觀念能完全配合。且具有相同之順位。

例：以x 代表一值，在數字上之排列方為

個位值      00x

十位值      0xx

百位值      xxx

如以文字觀念排列，則為

個位值      x00

十位值      0xx

百位值      xxx

其數值不同，但順位不變。

故此，內碼轉換公式必須加位置修正變數，設此變數為N1, N2, N3，當該置有字碼時，其值為1，否則為0。

$$N1(651X-650)+N2(26Y-25)+N3(Z)$$



此式中， $X, Y, Z$  為相對應於  $N_1, N_2, N_3$  之碼值。

由此式即可得25進位之文字順位值。

### 3.2. 內碼之頁區應用：

由於前述之方式三字节之內碼並未全部為中文字母碼所佔滿。因此，我們把各內碼之第一碼視為頁區碼，其後二字节視為順位碼，每頁區含64KB之順位，共有256個頁區。

3.2.1 頁區01至06未用，因中文字母最高順位為“日”字，數序為1，轉為內碼時，頁區為08，故00至07尚可保留為其他用途。

頁區00：專供中文文字處理時作空格用。

頁區07：ASCII 碼區，在07後有二碼，因為英文及數字每二字對應中文一字，故二碼可作一碼或二碼共用。

頁區08至CE：為中文字碼區，其中仍然保留有十二區供擴充用，且有圖形區（詳見後面說明）等與文字併用。

頁區CF：注音碼。注音中，子音有21個，佔五位元，母音15個，佔四位元。但因母音可能出現在第二第三位置（子音僅在第一位置），故母音分二組，共八位元，另有三位元供四聲及輕音用，共為16位元。茲以圖形說明如次：