



## 電腦字碼的故事（完結篇）

我仔頂一篇文章講著「電腦字碼的故事」，講到尾仔，咱的字碼定佇『Big5』這套，算是萬國歸一統。毋過咱人就袂用得掛無事牌（kuà-bô-sū-pài）咧，因為有網路的出現，各國的電腦攏會當相連（sio-liâm），電腦內的資料攏提會著，看會著。

小等一下，資料是提會著，毋過看毋捌，就是講網路提著的資料，除了英文字母佮數字以外，攏總是亂碼，攏總看無。所以電腦界就制定一个全世界各種語言共用的字碼，號做『Unicode』，就是『universal code』鬥起來的新字，意思是環球通用的字碼。咱的電腦有裝『Unicode』，外國資料咱就看有，咱的漢字外國電腦嘛看會著。

『Unicode』將各國字母攏囵在內，每一个字母用兩個位元（uī-guân）。位元是電腦用詞，兩個位元會當囵 65,536 字。各國語言大部份是字母鬥起來，親像拉丁語系的文字。全世界干焦漢字超過六萬字，袂使攏囵入去，所以漢字干焦囵三萬外字。咱的常用字是五千字，囵三萬字會使講解決 99.999% 的問題囉，所以咱嘛用甲真順序（sūn-sī）。

電腦是美國人發明的，佢的字是二十六個字母鬥起來的，所以初時的電腦用的是『Ansi』編碼，一个字母用一个位元表示，照『Unicode』編碼，佢嘛是需要一个字母用兩個位元表示，按呢是重倍（tîng-puē）討債（thó-tsè）。

問題來矣，電腦的資料，或者是網路往來的資料，八成外是英文。別國語言，親像德語、法語、華語等的資料總共才占一成外爾爾。所以有一寡實用主義的人就想講，這八成外猶原用『Ansi』，其他採用『Unicode』，按呢應該是省空間，閣猶原環球通用。是啦，這就是後來的『UTF8』編碼，現時網路上蓋通行的編碼，算是實用型，改良型的『Unicode』。

電腦字碼行到遮，已經解決 99.999% 的問題，所以這篇文章應該算是完結篇。

作者：邱文錫

### ☆詞彙學習☆

【掛無事牌】kuà-bô-sū-pài：不出錯、不發生問題；高枕無憂。

【鬥】tàu：裝上、組合。

【順序】sūn-sī：形容事情能夠按照程序進行而沒有阻礙。

【重倍】tîng-puē：雙倍。照原數增加一倍。也引申為「極為」、「極其」。

【討債】thó-tsè：浪費、糟蹋。沒有節制、無益的耗費。

【外】guā：數字之後表示「餘」、「多一點」的意思。超出某一個數量單位的。

【猶原】iu-guân：仍然、仍舊。表示狀況不變，持續下去。