

從檢查碼到修正碼

一、國際標準書碼 (ISBN)

國際標準書碼 (ISBN) 是一套國際統一的書號編碼規則，這個書碼由 10 個數字所組成，而每一個 ISBN 書碼唯一決定了一本書，這 10 個數字被分成 4 個部份，型式如下：

$$\text{ISBN } d_1 d_2 d_3 d_4 d_5 d_6 d_7 d_8 d_9 - d_{10}.$$

其中， d_{10} 是該本書的檢核號碼，而前面 9 個數字以連字號區分成 3 個部份，從第一部份依序分別用來表示區域語言、出版社及書名資料。例如：ISBN 957—603—139—7。

查詢此 ISBN 書碼即可得知該書為九章出版社出版之「生活的數學」一書。數字 603 與 139 分別表示出版社及書名之代號。若想知道 ISBN 所表示的書目是否有誤，只須依下列程序檢驗最後的檢核號碼是否相符，便可得知。方法如下：

將數字 $d_1, d_2, d_3, \dots, d_{10}$ 代入下面檢核函數中，所得之數字如果等於 0，那麼即表示該 ISBN 編號是正確的。如果 d_{10} 是一個兩位數值，則用 X 表示。

$$F(d_1, d_2, \dots, d_{10}) = 10d_1 + 9d_2 + 8d_3 + 7d_4 + 6d_5 + 5d_6 + 4d_7 + 3d_8 + 2d_9 + d_{10} \equiv 0 \pmod{11}$$

其實，最後的數字作為檢核之用，具有重要的意義。因為對一個數字的疏忽，若沒有最後的檢核數字作複核，結果將會「書目全非」。

下表為一些國際知名出版商的 ISBN 代號：

出 版 商	代 號
九章出版社	957-603
Addison-Wesley	0-201
Amer. Math. Soc.	0-821
Cambridge University Press	0-521
Dover	0-486
Oxford University Press	0-198
Springer-Verlag Berlin	3-540

在電腦中預設 ISBN 碼檢查的運算公式，當一位圖書館管理員鍵入一本書的 ISBN 書號時，如果按錯數字，則電腦大多可以提醒。但是如果不幸鍵入的數與正確的數正好相差 11 的倍數，則無法辨識，不過這種情況發生的機率不大。如果確知那一位數字錯誤或遺漏，則有辦法補正。

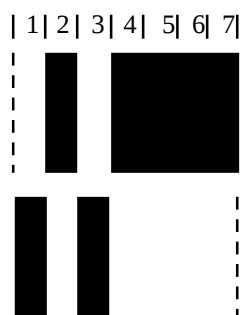
例如：957-63 | -139-7，在 | 內的數字是可以求出的。

二、國際產品碼 (UPC 條碼)

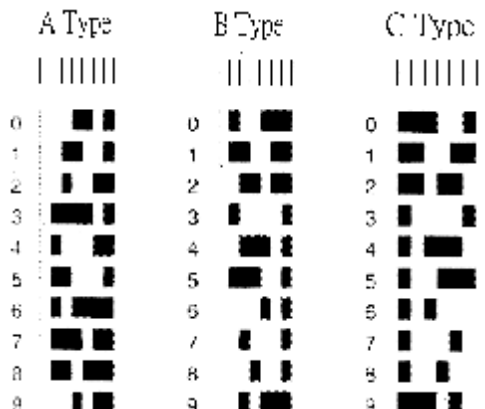
國際產品碼 UPC 條碼，全名為 Universal Product Code (又譯為統一商品條碼)，普遍使用於美洲一帶，算是 EAN 碼的前身，因此兩者有著類似的特性及編碼方式，只有 0~9 數字

字元而已。UPC 碼分為標準型 UPC-A 碼與簡易型 UPC-E 碼兩種，UPC-A 條碼與 EAN-13 碼同為 13 個數字的條碼，而 UPC-E 條碼與 EAN-8 碼同為 8 個數字的條碼。

首先說明 UPC 及 EAN 碼每一個字元的編碼原則，如圖一所示：



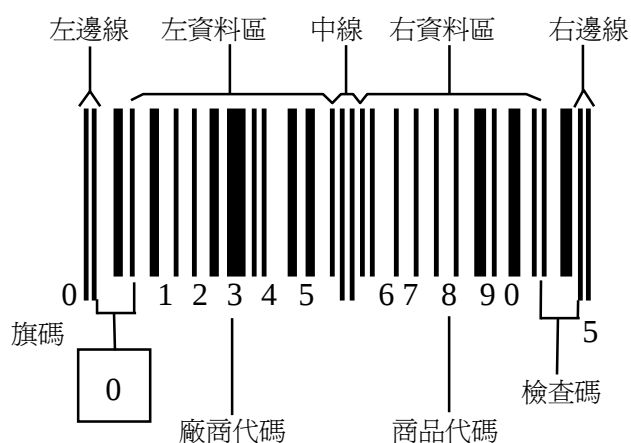
圖一



圖二

每一字元由 2 條 Space 及 2 條 Bar 所組合而成，而又細分成 7 等份，因此能組合不同粗細比例的線條字元，這就是 UPC 及 EAN 碼字元編碼的基礎。如圖一的上方組合成 0101111，為 A Type 編碼方式的數字“6”，圖一的下方組合成 1010000，為 C Type 編碼方式的數字“6”，兩者雖同是“6”，但編碼的型態卻不同，圖二所示為三種編碼型態的差異。

下圖為 UPC-A 條碼的基本架構：



左邊線：固定為 101。

旗碼：佔 2 個字元，第一個字元固定為 0，通常不列印出來，第二個字元代表商品種類。

廠商代碼：佔 5 個字元，須提出申請。

中線：固定為 01010。

商品代碼：佔 5 個字元，由廠商自行訂定。

檢查碼：佔 1 個字元。

右邊線：固定為 101，與左邊線相同。

UPC-A 檢查碼的計算方式：

例如有一 UPC-A 碼為：001234567890C，C 為檢查碼。設此編碼由左至右的 13 個數字依序為 $n_{13}, n_{12}, n_{11}, n_{10}, n_9, n_8, n_7, n_6, n_5, n_4, n_3, n_2, n_1$ ，則檢查碼公式為

$$n_1 = 10 - \left[\left(3 \sum_{\substack{i=2 \\ \text{even}}}^{12} n_i + \sum_{\substack{i=3 \\ \text{odd}}}^{13} n_i \right) \pmod{10} \right] \pmod{10},$$

代入公式，得

$$\begin{aligned}
 n_1 &= 10 - [3(0 + 8 + 6 + 4 + 2 + 0) + (9 + 7 + 5 + 3 + 1 + 0) \pmod{10}] \pmod{10} \\
 &= 10 - [85 \pmod{10}] \pmod{10} \\
 &= 10 - 5 \pmod{10} \\
 &= 5
 \end{aligned}$$

所以，檢查碼 $C = 5$ 。

EAN-UCC(國際貨品編碼協會)檢查碼：

使用 EAN 編碼有什麼好處？所有號碼均是全球獨一無二；EAN 編號本身並不包含任意思，它們只作為從數據庫提取詳盡資料的鑰匙；其不含任何意思、獨一無二的特質使其適用於各行各業及跨越國界限制；每個 EAN 編碼末的終檢碼，確保安全可靠的資料套取。

怎樣計算終檢碼？

計算13位數的終檢碼例子：

	數位位置																	
EAN UCC-8											N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8
UCC-12							N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	N_{11}	N_{12}
EAN UCC-13						N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	N_{11}	N_{12}	N_{13}
EAN UCC-14					N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	N_{11}	N_{12}	N_{13}	N_{14}
SSCC	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	N_{11}	N_{12}	N_{13}	N_{14}	N_{15}	N_{16}	N_{17}	N_{18}
每一數位乘以下的倍數																		
	$\times 3$	$\times 1$	$\times 3$	$\times 1$	$\times 3$	$\times 1$	$\times 3$	$\times 1$	$\times 3$	$\times 1$	$\times 3$	$\times 1$	$\times 3$	$\times 1$	$\times 3$	$\times 1$	$\times 3$	
相加起來																		
以最接近10的倍數減去以上的結果=終檢碼																		

舉例

位置	N_1	N_2	N_3	N_4	N_5	N_6	N_7	N_8	N_9	N_{10}	N_{11}	N_{12}	N_{13}
沒有終檢碼的編號碼	6	2	9	1	0	4	1	5	0	0	2	1	
步驟1: 乘數	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	$\times$	
倍數	1	3	1	3	1	3	1	3	1	3	1	3	
步驟2: 相加	=	=	=	=	=	=	=	=	=	=	=	=	
總數	6	6	9	3	0	12	1	15	0	0	2	3	= 57
步驟3: 把最接近10的倍數(60)減去以上的結果=終檢碼(3)													
附有終檢碼的編號碼	6	2	9	1	0	4	1	5	0	0	2	1	3

三、 身份證號碼檢驗規則

中華民國的國民身份證字號共有 10 位，第一位為英文字母，第二個數字是男女生之分，男生為 1，女生為 2，最後一位數字為檢查碼。

檢查碼的運算原則：

1. 英文代號係依據下表轉換成數字：

A = 10	台北市	B = 11	台中市	C = 12	基隆市	D = 13	台南市	E = 14	高雄市
F = 15	台北縣	G = 16	宜蘭縣	H = 17	桃園縣	*I = 34	嘉義市	J = 18	新竹縣
K = 19	苗栗縣	L = 20	台中縣	M = 21	南投縣	N = 22	彰化縣	*O = 35	新竹市
P = 23	雲林縣	Q = 24	嘉義縣	R = 25	台南縣	S = 26	高雄縣	T = 27	屏東縣
U = 28	花蓮縣	V = 29	台東縣	*W = 30	金門縣	X = 31	澎湖縣	Y = 32	陽明山
*Z = 33	連江縣								

2. 設身份證字號為 $yhgfedcbax$ ，其中 y 表英文字母 A~Z， $a, b, c, d, e, f, g, h, x$ 為數字，若 $y = 10m + n$ ，則檢查碼 x 的公式為：

$$x = 10 - [(m + 9 \times n + a \times 1 + b \times 2 + c \times 3 + d \times 4 + e \times 5 + f \times 6 + g \times 7 + h \times 8) \pmod{10}] \pmod{10}$$

若餘數為 0，則檢查碼就是 0。

例如：身份證字號為 Y100232736，其檢查碼 6 是否符合？

因為 $Y = 32$ ，運用公式有

$$\begin{aligned} x &= 10 - [(3 + 9 \times 2 + 3 \times 1 + 7 \times 2 + 2 \times 3 + 3 \times 4 + 2 \times 5 + 0 \times 6 + 0 \times 7 + 1 \times 8) \pmod{10}] \\ &= 10 - [(21 + 3 + 14 + 6 + 12 + 10 + 0 + 0 + 8) \pmod{10}] \\ &= 10 - 74 \pmod{10} \equiv 10 - 4 \pmod{10} \\ &= 6 \end{aligned}$$

結果符合。

研究問題：

1. 為什麼檢查碼的運算公式通常是用線性運算？

註：所謂線性運算是指： $f(x_1, x_2, \dots, x_n) = a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_nx_n$ ，其中 $a_1, a_2, \dots, a_n$ 為整係數。

2. 如果已知檢查碼的運算公式是線性運算，如何求出運算公式？如果某銀行帳號連同檢查碼共有 8 位，那麼至少要多少資料才能求出檢查碼的運算公式？

3. 在修正碼一文中，如果傳送的不是 0, 1 碼，而是數字碼，情形會如何？

4. 如果可能錯二碼，則情形會如何？