修正碼

劉江楓 · 劉翰亘

『修正碼』, 是一個有重要的實用價值, 有深遠的教育意義, 及有強烈的趣味引力, 不可多得的精彩研究題, 適合各程度的學生及老師探討。

我們有一部傳達機, 每次發出十五個0或1的數字, 不過有時不準確, 一個0會變爲1, 或一個1變爲0, 但不會錯多於一個數字。如果我們對這毛病置之不理, 仍全用十五個數字表達訊息, 將引至誤會, 甚至有嚴重的後果。

假如錯誤不常出現, 可用前十四個數字表達訊息, 最後一個, 如前面1的總數是奇數, 則選爲1, 如是偶數, 則選爲0, 所以十五個數字內, 1的總數必定是偶數。收到傳達時, 可檢查這總數, 如是偶數, 便知無誤, 如是奇數, 便回傳要求重發。這叫『奇偶碼』, 是『檢查碼』的基本例子, 效率是14/15, 不能再高了。

假如錯誤經常出現,『檢查碼』就不如理想了, 我們希望收方不但可以知道出了問題, 還能判別錯在那裡, 這便是『修正碼』的主題。

例一: Alt 碼 (1948/1949) 年刊登 [1])

前五個數字表達訊息, 然後複述兩次, 如訊息是00110, 則發出0011000110001 10。如收到000100011000110, 因第二及第三份相同, 所以知道正確的訊息是00110。

這是『修正碼』最早的實用例子,『三盤兩勝』 的想法十分簡單, 但1/3的效率非常低。當他是國中二年級學生時, Mark Rabenstein 便指出, 因至多有一個錯, 所以兩份訊息便足夠了, 問題是怎樣知道那一份是對的。

例二: Rabenstein 碼 (1984年發明, 1985年刊登 [4])

前七個數字表達訊息, 然後複述一次, 最後一個數字的選擇, 使後八個數字內, 1的總數是偶數, 如訊息是0011001, 則發出001100100110011。如收到001100101110011, 因後八個數字內, 1的總數是奇數, 所以最後那個數字『選錯』了, 這表示前面沒錯, 正確的訊息是0011001。

這個例子, 把『奇偶碼』的想法用上了, 7/15的效率, 比1/3高多了。

例三: 劉翰亘九宮碼 (1997年發明, 本文首次刊登)

前九個數字表達訊息, 分別填進九宮格, A, B, C, D, E, F, G, H 及 J 內 (見圖一), 後六個數字, 是每行每列的『奇偶碼』, K 的選擇, 使 A, B, C 和 K 之間, 1的總數是偶數。L 依 D, E 及 F 選, M 依 G, H 和 J 選, N 依 A, D 和 G 選, P 依 B, E 和 H 選, Q 依 C, F 和 J 選。如訊息是 011101001, 則依圖一填數, 使每行每列, 1的總數都是偶數, 發出 011101001001111。如收到 010101001001111, 則依圖二填數, 便發現 K 和 Q『選錯』了, 所以錯在這兩行列的交點 C, 正確訊息是 011101001。如收到 011101001000111, 則依圖三填數, 便發現祇有 M『選錯』, 所以錯就在 M 本身, 正確息是 011101001。

例三的效率是3/5, 比例二有所提高, 雖然也是用『奇偶碼』的想法, 但將『一字長蛇』變爲『縱橫交錯』, 從狹窄的直線, 擴展到空闊的平面, 應用上靈活多了。但還有沒有改進的餘地呢? 應該是有的。

爲什麼我們可以這樣說呢? 因爲 K, L, M, N, P 和 Q 六個數字中, 1的總數必定是偶數, 所以其中一個是浪費了, 現在我們證明這個結論。

A 0	D 1	G 0	N 1
B 1	E 0	H 0	P 1
C 1	F 1	J 1	Q 1
K 0	L 0	M 1	

圖一

A 0	D 1	G 0	N 1
B 1	E 0	H 0	P 1
C 0	F 1	J 1	Q 1
K 0	L 0	M 1	

圖二

A 0	D 1	G 0	N 1
B 1	E 0	H 0	P 1
C 1	F 1	J 1	Q 1
K 0	L 0	M 0	

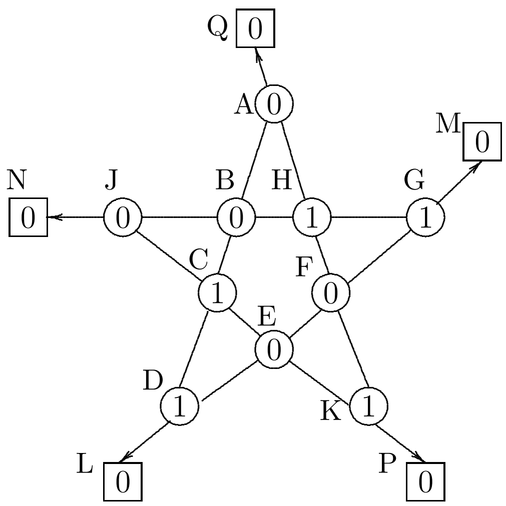
圖三

如 A 到 J 中有一個數字由1變0或由0變1, 則 K, L 和 M 中有一個改變, N, P 和 Q 中也有一個改變, 所以這六個數字中, 可能兩個1變了0, 可能兩個0變1, 亦可能一個1變了0, 而一個0變了1, 但1的總數的奇偶性不會改變。當 A 到 J 全變爲0後, K 到 Q 也全變爲0, 這時1的總數是偶數, 所以開始時也是。

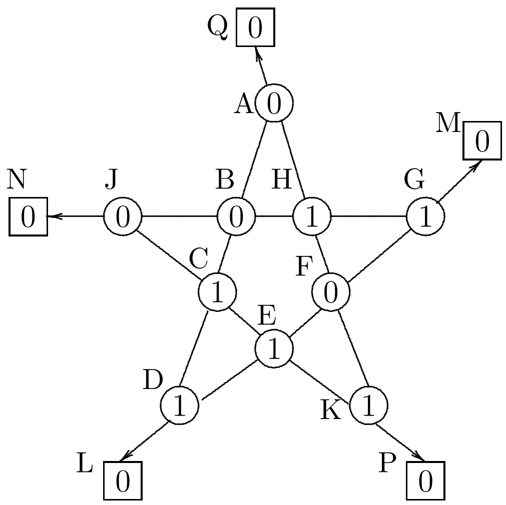
例四: 劉翰亘五星碼 (1997年發明, 本文首次刊登)

前十個數字表達訊息, 分別填進五角星的交點 $A, B, C, D, E, F, G, H, J$ 及 K 上, 後五個數字, 是每線的『奇偶碼』, L 依 A, B, C 及 D 選, M 依 D, E, F 及 G 選, N 依 G, H, B 及 J 選, P 依 J, C, E 及 K 選, Q 依 K, F, H 及 A 選。如訊息是 0011001101, 則依圖四填數, 發出 001100110100000。如收到

001110110100000, 則依圖五填數, 便發現 M 和 P『選錯』了, 所以錯在這兩線的交點 E, 正確訊息是0011001101。如收到 001100110100100, 則依圖六填數, 便發現祇有 N『選錯』, 所以錯就在 N 本身, 正確訊息是0011001101。

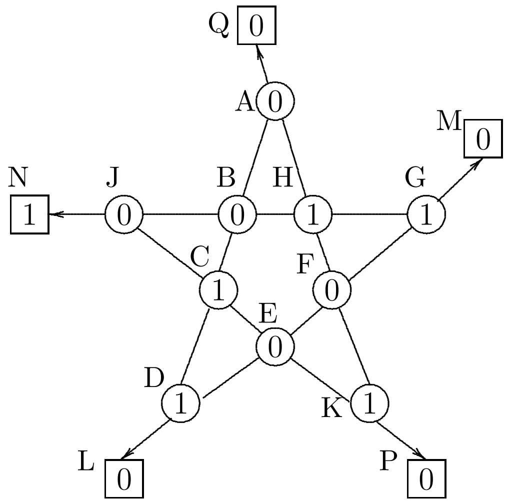


圖四



圖五

以上兩個例子, 是第二作者國小六年級時, 參加香港保良局比賽後想出來的。例四的效率是2/3, 比例一足足高了一倍, 但還是可以改進的, 因爲和例三一樣, L, M, N, P 和 Q 五個數字中, 1的總數必定是偶數, 所以仍有浪費。還有一個毛病, 就是圖形不甚好劃, 我們可以嘗試, 用抽象符號代替幾何圖形。



圖六

A	B	C	D	E	F	G	H	J	K	L	M	N	P	Q
a	a	a	a							a				
b				b	b	b					b			
	c			c			c	c				c		
		d			d		d		d				d	
			e			e		e	e					e
0	0	1	1	0	0	1	1	0	1	0	0	0	0	0

圖七

圖七應用了一個有五個元素的集 $\{a, b, c, d, e\}$, 這個集的一部份稱爲它的子集, 其中一個是沒有元素的空集 $\{\ \}$, 不過我們對它興趣不大, 有五個單元子集: $\{a\}$, $\{b\}$, $\{c\}$, $\{d\}$ 和 $\{e\}$, 及十個雙子集: $\{a,b\}$, $\{a,c\}$, $\{a,d\}$, $\{a,e\}$, $\{b,c\}$, $\{b,d\}$, $\{b,e\}$, $\{c,d\}$, $\{c,e\}$ 及 $\{d,e\}$。圖四的五

條線, 在圖七變成上列五個單元子集, 十個點變成上列十個雙元子集, 很容易看出, 這兩個圖形是基本上一樣的。我們沒有用到三元, 四元和五元的子集, 這就是改進的途徑了。

例五: Hamming-Golay 碼 (1949/50 年刊登 [2], [3])

前十一個數字表達訊息, 分別塡在 $\{a, b, c, d\}$ 非空子集 $A, B, C, D, E, F, G, H, J, K$ 和 L 之下 (見圖八), M 的選擇, 使所有包含 $\{a\}$ 的子集 A, B, C, D, F, G, H 和 M 之下, 1的總數成爲偶數。N 依 A, B, C, E, F, J 和 K 選, P 依 A, B, D, E, G, J 和 L 選, Q 依 A, C, D, E, H, K 和 L 選。如訊息是 10111100011, 則依圖八塡數, 發出 101111000110100。如收到 101110000110100, 則依圖九塡數, 算出含 $\{a\}$ 的子集下有三個1, 含 $\{b\}$ 的有五個, 含 $\{c\}$ 的有四個, 含 $\{d\}$ 的有六個, 所以錯在 $F = \{a, b\}$ 之下, 正確訊息是10111100011。

A	B	C	D	E	F	G	H	J	K	L	M	N	P	Q
a	a	a	a		a	a	a				a			
b	b	b		b	b			b	b			b		
c	c		c	c		c		c		c			c	
d		d	d	d			d		d	d				d
1	0	1	1	1	1	0	0	0	1	1	0	1	0	0

圖八

A	B	C	D	E	F	G	H	J	K	L	M	N	P	Q
a	a	a	a		a	a	a				a			
b	b	b		b	b			b	b			b		
c	c		c	c		c		c		c			c	
d		d	d	d			d		d	d				d
1	0	1	1	1	0	0	0	0	1	1	0	1	0	0

圖九

例五的效率高達11/15, 亦在早期便發明及應用的, 但這並不表例二, 例三和例四毫無意義, 而是學生研究的寶貴成果。例二在 Mathematics Magazine 發表後, 甚獲好評, 雖然 Mark 沒有在大學進修數學, 但這經驗, 對他在1997年在 Berkeley 大學拿到化學博士的學位, 是有一定的影響。

例五還能改進嗎? 這回不成了, 因爲傳達可能無誤, 亦可能錯在十五個數字中任何一個, 並有十六種不同的情況, 必須由訊息後補選的數字判別。每個數字祇有兩種選擇, 僅能判別兩種情況, 兩個數字判別四種, 三個八種, 所以必須補選至少四個數字, 才足夠判別十六種情況, 效率就不能高於11/15了。

本文僅涉及『修正碼』問題的皮毛, 接著的研究, 自然是找能檢查或修正兩個錯誤或以上的碼, 可參閱 Dennis Shasha [5]的名著第七章的第一節及第三章的第一節。

嗚謝: 第二作者, 謹向台北市九章數學圖書文具公司負責人孫文先先生致意, 感謝他出錢出力, 組織台灣各地小朋友參加有意義的課外數學活動, 讓本人有機會接觸到『修正碼』的問題, 並介紹本人和第一作者認識。

參考資料

1. F. L. Alt, A Bell Telephone Laboratories' Computing Machine (I) Math Comput 3 (1948-49), 1-13.

2. M. J. E. Golay, Notes on Digital Coding, Proc. I.E.E.E., 37(1949), 657.
3. R. W. Hamming, Error Detecting and Correcting Codes, Bell System Tech. J., 29 (1950), 147-160.
4. M. Rabenstein, An Example of an Error Correcting Code, Math Mag, 58 (1985), 225-226.
5. D. Shasha,「The Puzzling Adventures of Dr. Ecco」, Dover Publications Inc New York (1998).

—劉江楓任教於加拿大Alberta 大學數學系, 劉翰亘就讀於台北市介壽國民中學三年級—