

附錄1：中文字碼處理說明

- 一、文書以電子方式傳遞時，若需經公文電子交換中心，則交換之文書需轉換為中文標準交換碼 CNS11643 中文標準交換碼（以下簡稱 CNS 碼）。
- 二、線上簽核系統需具備處理 BIG5、BIG5E、Unicode 等中文字碼之文稿。
- 三、各機關辦理檔案目錄、分類表及移轉(交)電子媒體封裝檔之傳(彙)送時，中文字以 CNS 碼、BIG5 或 Unicode 表示之。
- 四、機關(構)應有中文字碼管理人員管理、維護各機關(構)內共用之使用者造字區字集與前置處理軟體建置之單位自用(造)字對照檔之正確性，並可視需要應用 CNS11643 中文標準交換碼全字庫網站 (<http://www.cns11643.gov.tw>) 提供之機制。
- 五、文書傳遞交換中，均需整合搭配以下功能，以使相關作業順利：
 - (一)處理 BIG5、BIG5E、Unicode、CNS 碼。
 - (二)具備 CNS 碼與 BIG5 對照表、CNS 碼與 BIG5E 對照表、CNS 碼與 Unicode 對照表，供轉碼程式使用。
 - (三)依據用戶端之各機關(構)內共用之使用者造字區字集建置「自用(造)字碼與 CNS 碼對照表」檔，供轉碼程式使用。
 - (四)收文端接收之 XML 格式文書，若經轉碼，收文機關(構)之對照表沒有對照之字碼時，應通知機關內部中文字碼管理人員於對照表中新增自用(造)字，提供機關(構)內使用者使用，以確保收文端之電子檔能正確表達發文端所發文書之中文字形。
- 六、中文字碼轉碼原則，相關規範說明如下：

(一)檔案標頭需以<?xml version="1.0" encoding=" " ?>開頭。

- 1.以 UTF-8 編碼的檔案：整份文件需以 UTF-8 編碼，其檔案標頭必須是<?xml version="1.0" encoding="UTF-8"?>開頭。

- 2.以 BIG5 編碼的檔案：整份文件需以 BIG5 編碼，其檔案標頭必須是<?xml version="1.0" encoding="BIG5"?>開頭。
- 3.以 CNS11643 編碼的檔案：整份文件需以 CNS11643 編碼，其檔案標頭必須是<?xml version="1.0" encoding="CNS11643"?>開頭。
- 4.檔案標頭前 5 個 Byte 必須為: 0x3C 0x3F 0x78 0x6D 0x6C。
- 5.檔案標頭不可有 BOM (Byte-order mark) 識別字元，例如 0xEF 0xBB 0xBF 這 3 個位元等。
- 6.檔案標頭不可有 ISO 2022 控制符，例如 SI (0x0f) 等。

(二)CNS11643 編碼規範：

- 1.字集編碼需根據：ISO-2022-CN、ISO-2022-CN-EXT 標準 ISO 2022-CN 屬於 7-bit 編碼，將 ASCII 指定給 G0 字面、中文字指定給 G1 或 G2 字面。
- 2.如果前後 2 個字元同屬一個字面時 (ASCII 也看做一個字面)，後字元可維持與前字元相同的編碼格式 (ASCII 為單位元組、CSIC 為雙位元組)。
- 3.如果前後 2 字元分屬不同字面時，就必須在後字元的前面 (即 2 字元之間) 插入調用控制符甚或逸出順序控制符以切換字面。
- 4.字面的指定與切換需用逸出順序 (ESC sequence) 加上調用控制符。
- 5.第 0、1、2 字面逸出順序 (ESC sequence) 可省略，只用調用控制符，因啟動中文系統時，一併將 ASCII 和 CSIC 的第 1、第 2 字面分別載入 G0、G1 與 G2 字面，預設 G0 字面是 ASCII 字面，G1 是第一字面，G2 是第二字面。
- 6.第 3 ~ 80 字面，先將逸出順序 (ESC sequence) 載入 G3，再用調用控制符 LS3。

(三)字面切換規範：

- 1.切換至 ASCII 字面，使用調用控制符 SI (字元碼為 0Fh)：

當後字元為 ASCII (已載入 G0) 而前字元不是 ASCII 時，兩者之間必須插入 調用控制符 SI (字元碼為 0Fh)。

2. 切換至第一字面，使調用控制符 SO（字元碼為 0Eh）：

若後字元屬於 CSIC 第 1 字面（已載入 G1）而前字元不是，則兩者之間必須插入調用控制符 SO（字元碼為 0Eh）。

3. 切換至第二字面，使用調用控制符 LS2（字元碼為 1Bh 6Eh）：

當後字元屬於 CSIC 第 2 字面（已載入 G2）而前字元屬於其他字面時，兩者之間應插入 調用控制符 LS2（字元碼為 1Bh 6Eh）。

4. 切換至第 3 - 16 字面，是使用調用控制符 LS3（字元碼為 1Bh 6Fh）：

若後字元屬於第 3 至第 16 字面，則兩者之間應先依據該字元所屬字面插入逸出順序“1Bh 24h 2Bh [F]”，再插入調用控制符 LS3（字元碼為 1Bh 6Fh）。

5. 切換至第 17- 80 字面以後的字面，是使用調用控制符 LS3（字元碼為 1Bh 6Fh）：若後字元屬於第 17 至第 80 字面，則兩者之間應先依據該字元所屬字面插入逸出順序“1Bh 24h 2Bh [I] [F]”，再插入調用控制符 LS3（字元碼為 1Bh 6Fh）。

6. 調用控制符列表

表1-1調用控制符列表

調用控制符	字元碼
SI (Shift In)	0F
SO (Shift Out)	0E
LS2 (Locking Shift 2)	1B 6E
LS3 (Locking Shift 3)	1B 6F

7. 字元集庫、逸出順序與調用控制符關係列表：

表1-2字元集庫、逸出順序與調用控制符關係列表

字元集庫	逸出順序 (ESC Sequence)	是否省略 逸出順序(ESC sequence)	調用控制符
ASCII	1B 28 42	必須省略 (載入字元集 G0)	0F (SI)

字元集庫	逸出順序 (ESC Sequence)	是否省略 逸出順序(ESC sequence)	調用控制符
CNS 11643 第 1 字面	1B 24 29 47	必須省略 (載入字元集 G1)	0E (SO)
CNS 11643 第 2 字面	1B 24 2A 48	必須省略 (載入字元集 G2)	1B 6E (LS2)
3-16 字面	1B 24 2B [F]		
CNS 11643 第 3 字面	1B 24 2B [F] F=49	不可省略 (載入字元集 G3)	1B 6F (LS3)
CNS 11643 第 4 字面	1B 24 2B [F] F=4A	不可省略 (載入字元集 G3)	1B 6F (LS3)
CNS 11643 第 5 字面	1B 24 2B [F] F=4B	不可省略 (載入字元集 G3)	1B 6F (LS3)
CNS 11643 第 6 字面	1B 24 2B [F] F=4C	不可省略 (載入字元集 G3)	1B 6F (LS3)
CNS 11643 第 7 字面	1B 24 2B [F] F=4D	不可省略 (載入字元集 G3)	1B 6F (LS3)
CNS 11643 第 8 字面	1B 24 2B [F] F=37	不可省略 (載入字元集 G3)	1B 6F (LS3)
CNS 11643 第 9 字面	1B 24 2B [F] F=38	不可省略 (載入字元集 G3)	1B 6F (LS3)
CNS 11643 第 10 字面	1B 24 2B [F] F=39	不可省略 (載入字元集 G3)	1B 6F (LS3)
CNS 11643 第 11 字面	1B 24 2B [F] F=3A	不可省略 (載入字元集 G3)	1B 6F (LS3)
CNS 11643 第 12 字面	1B 24 2B [F] F=3B	不可省略 (載入字元集 G3)	1B 6F (LS3)
CNS 11643 第 13 字面	1B 24 2B [F] F=3C	不可省略 (載入字元集 G3)	1B 6F (LS3)

字元集庫	逸出順序 (ESC Sequence)	是否省略 逸出順序(ESC sequence)	調用控制符
CNS 11643 第 14 字面	1B 24 2B [F] F=3D	不可省略 (載入字元集 G3)	1B 6F (LS3)
CNS 11643 第 15 字面	1B 24 2B [F] F=3E	不可省略 (載入字元集 G3)	1B 6F (LS3)
CNS 11643 第 16 字面	1B 24 2B [F] F=3F	不可省略 (載入字元集 G3)	1B 6F (LS3)
17-80 字面	1B 24 2B [I] [F]		
CNS 11643 第 17 字面	1Bh 24h 2Bh [I] [F] I=21 h F=30 h	不可省略 (載入字元集 G3)	1B 6F (LS3)
CNS 11643 第 18 字面	1Bh 24h 2Bh [I] [F] I=21 h F=31 h	不可省略 (載入字元集 G3)	1B 6F (LS3)
以此類推	以此類推	以此類推	以此類推
CNS 11643 第 78 字面	1Bh 24h 2Bh [I] [F] I=24 h F=3D h	不可省略 (載入字元集 G3)	1B 6F (LS3)
CNS 11643 第 79 字面	1Bh 24h 2Bh [I] [F] I=24 h F=3E h	不可省略 (載入字元集 G3)	1B 6F (LS3)
CNS 11643 第 80 字面	1Bh 24h 2Bh [I] [F] I=24 h F=3F h	不可省略 (載入字元集 G3)	1B 6F (LS3)

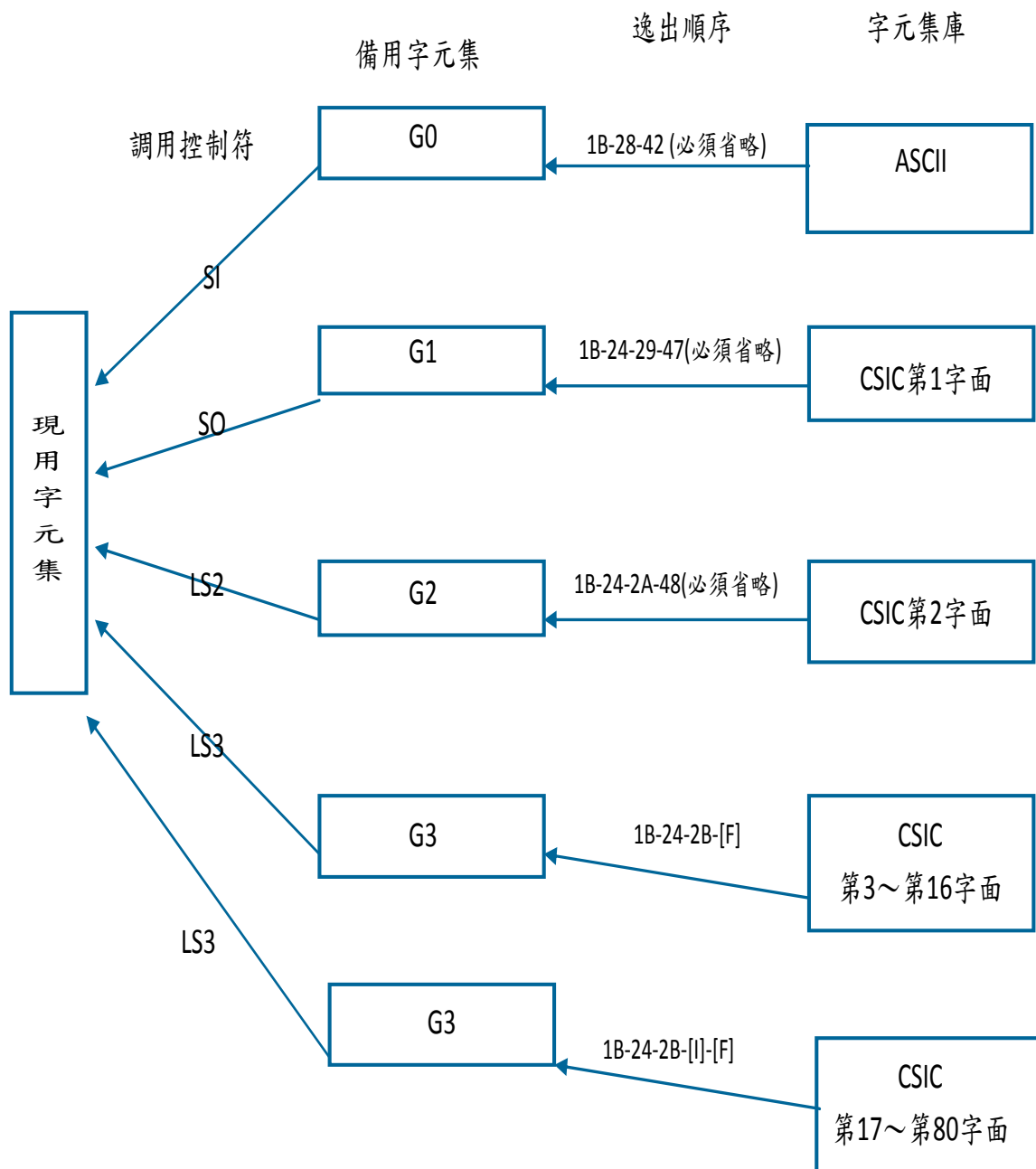


圖1-1字元集庫、逸出順序與調用控制符關係示意圖

8.CSIC 第 1~16 字面的最終位元組[F]

『正式[F]』係指該最終位元組已向 ISO 註冊，『專用』意指該中間字元或最終位元組只能適用於中華民國境內。

表1-3CSIC 第 1~16 字面的最終位元組[F]

字面	[F]
1	47h 正式[F]
2	48h 正式[F]
3	49h 正式[F]
4	4Ah 正式[F]
5	4Bh 正式[F]
6	4Ch 正式[F]
7	4Dh 正式[F]
8	37h 專用[F]
9	38h 專用[F]
10	39h 專用[F]
11	3Ah 專用[F]
12	3Bh 專用[F]
13	3Ch 專用[F]
14	3Dh 專用[F]
15	3Eh 專用[F]
16	3Fh 專用[F]

9.CSIC 第 17~80 字面的專用中間字元[I]及專用最終位元組[F]

表1-4CSIC 第 17~80 字面的專用中間字元[I]及專用最終位元組[F]

字面	[I]-[F]	字面	[I]-[F]	字面	[I]-[F]	字面	[I]-[F]
17	21-30h	33	22-30h	49	23-30h	65	24-30h
18	21-31h	34	22-31h	50	23-31h	66	24-31h
19	21-32h	35	22-32h	51	23-32h	67	24-32h
20	21-33h	36	22-33h	52	23-33h	68	24-33h
21	21-34h	37	22-34h	53	23-34h	69	24-34h
22	21-35h	38	22-35h	54	23-35h	70	24-35h
23	21-36h	39	22-36h	55	23-36h	71	24-36h
24	21-37h	40	22-37h	56	23-37h	72	24-37h
25	21-38h	41	22-38h	57	23-38h	73	24-38h
26	21-39h	42	22-39h	58	23-39h	74	24-39h
27	21-3Ah	43	22-3Ah	59	23-3Ah	75	24-3Ah
28	21-3Bh	44	22-3Bh	60	23-3Bh	76	24-3Bh
29	21-3Ch	45	22-3Ch	61	23-3Ch	77	24-3Ch
30	21-3Dh	46	22-3Dh	62	23-3Dh	78	24-3Dh
31	21-3Eh	47	22-3Eh	63	23-3Eh	79	24-3Eh
32	21-3Fh	48	22-3Fh	64	23-3Fh	80	24-3Fh

10.CSIC 的字面切換應用範例：

表1-5CSIC 的字面切換應用範例

0F (SI)	43 C	4E N	53 S	31 1	31 1	36 6	34 4	33 3	0E (SO)	4463 中	4546 文	6D3A 標	646D 準	4728 交
5F50 換	6E23 碼	5152 是	4B6D 依	0F (SI)	43 C	4E N	53 S	35 5	32 2	30 0	35 5	0E (SO)	452F 及	0F (SI)
43 C	4E N	53 S	37 7	36 6	35 5	34 4	0E (SO)	4466 之	5D3D 規	6E2F 範	4854 而	4C31 制	5354 訂	2122 ，
0F (SI)	43 C	4E N	53 S	31 1	31 1	36 6	34 4	33 3	0E (SO)	5972 將	215A 『	1B-6E (LS2)	2539 俾	0E (SO)
215B 』	6E3E 編	4D75 於	5C49 第	4428 二	4773 字	5372 面	2122 ，	4C27 其	6E3E 編	6E23 碼	5233 為	0F (SI)	32 2	35 5
33 3	39 9	0E (SO)	2122 ，	5972 將	215A 『	1B-24-2B-49 逸出順序：第 3 字面→G3					1B 6F (LS3)	5972 鉢	0E (SO)	215B 』
6E3E 編	4D75 於	5C49 第	4435 三	4773 字	5372 面	2124 。								