中諮文件編號 2001/09

中文界面諮詢委員會

中文資訊科技工作小組

進展報告

2001.11.23

1. 工作小組在中文界面諮詢委員會第 9 次會議後召開了兩次會議。
   - 第 11 次會議（2001.07.11），主要有三項議題：

   A. 互選工作小組召集人
   B. 確認工作小組成員
   C. 設立 ISO 10646 資訊中心網站的跟進事項

   - 第 12 次會議 (2001.11.16)，主要有二項議題：

   A. 建議制訂《香港增補字符集》的新版本（邀請「中文電腦用字工作小組」成員參加討論）
   B. 設立 ISO 10646 資訊中心網站的跟進事項

2. 互選工作小組召集人並確認工作小組成員

   巫偉明先生建議由陸勤博士出任工作小組的召集人，與會者表示贊成。陸勤博士同意出任工作小組的召集人。工作小組並確認李健康教授及李友強先生為小組成員。

3. 設立 ISO 10646 資訊中心網站的跟進事項

   A. 香港生產力促進局已根據早前資訊科技署發放的資料，完成了網站的原型。網站包括 ISO 10646 國際編碼標準的簡介、連結至相關網站的超連結、開發支援 ISO 10646 國際編碼標準的產品供應商資料及提供產品查詢功能。此外，網站亦提供表格給產品供應商填報有關產品資料，香港生產力促進局會根據表格的資料，更新網站的資料庫。

   B. 與會者同意在網站發放的資料應包括中文及英文版文。此外，就產品供應商填報的產品資料表格，與會者建議對其作出以下修訂：

   - 表格內容應同時以中文及英文印製。
   - 產品供應商可選擇以中文或英文填報產品資料。
   - 加入相關問題以得悉產品可否支援《香港增補字符集- 2001》和 ISO 10646-2:2001。
   - 表格內可容許產品供應商指明這次填報的是新產名資料還是更新以前的

產品資料。

秘書處會著手更新產品資料表格，並將新表格以電郵送交給各組員作進一步討論。

4. 建議制訂《香港增補字符集》的新版本

鑑於《香港增補字符集》發表後，中諮會又通過新增收了 113 個字符，且 ISO 10646-2:2001 計劃於十一月發佈，小組成員一致認為有必要公佈新的版本。名稱為《香港增補字符集－２００１》以對應 ISO 10646-2:2001。修訂內容包括：

A. 新增字符

B. 包括 ISO 10646-2:2001 的 Big5 與 ISO 10646 字碼對應表

C. 更正錯誤

新版仍以下載形式在“數碼 21" 網站公佈。

詳細建議請見附錄。

附錄

中文資訊科技工作小組

制訂《香港增補字符集》的修訂版本

目的

1 本文件旨在就制訂《香港增補字符集》的修訂版本，徵詢各組員的意見。

背景

2 香港特區政府在「數碼 21」資訊科技策略下，為香港特別行政區制訂一套共通的電子中文界面，以方便市民能更有效地使用中文進行電子通訊。共通電子中文界面的基礎包括採用 ISO 10646 國際編碼標準及制訂《香港增補字符集》。

3 在一九九九年九月，香港特別行政區政府與「中文界面諮詢委員會」(中諮會)公布所合作編製的《香港增補字符集》。《香港增補字符集》包括 4,702 個中文字符，該等字符是市民及政府部門在以中文處理電子資訊時需要使用，而不包括在大五碼內的字符。制訂《香港增補字符集》的目的，是令市民及政府共同使用一套包含香港特有用字的字符集，以更有效及準確地進行中文電子通訊及資料交換。

4 目前，《香港增補字符集》內的 4,702 個字符中，有 3,054 個字符已獲納入 ISO 10646 國際編碼標準的 ISO/IEC 10646-1:2000 版本內。其餘的 1,648 個字符，則被放在該版本的私人使用區，而其中的 1,622 個字符將獲納入 ISO 10646 國際編碼標準的下一個版本內。

5 鑑於市民及政府部門不時有需要申請把新字符納入《香港增補字符集》內，政府與中諮會於二零零零年四月公布了《香港增補字符集》字符增收程序和原則。自公布該程序和原則以來，中諮會經處理市民及政府部門提交的新增字符申請後，共增收了 113 個字符(請參閱附件一)。其中 75 個字符已獲收納在 ISO 10646 國際編碼標準的 ISO/IEC 10646-1:2000 版本內，在餘下的 38 個字符中，有 29 個將獲收納在 ISO 10646 國際編碼標準的下一個版本內，剩下的 9 個則放在其私人使用區。

6 由於軟件商開發軟件支援增收字符需時，而且《香港增補字符集》更

新過於頻密會對用戶造成不便，中諮會在二零零零年四月舉行的第六次會議上，決定在適當時間才公布《香港增補字符集》的修訂版本，把增收字符正式納入《香港增補字符集》內。

7 ISO 10646 國際編碼標準的下一個版本預計會在二零零一年年底前獲得公布。這個包括擴展區 B 的新版本將會增收大約四萬多個表意文字，其中包括所有收錄在《康熙字典》、《漢語大字典》和《漢語大詞典》內的漢字，以及包括在本文第四段提及的 1,622 個《香港增補字符集》字符。

修訂《香港增補字符集》的建議

8 鑑於 ISO 10646 國際編碼標準的新版本將於短期內獲得公布，而且《香港增補字符集》自公布至今，已超過兩年，我們認為現時是適當時候將《香港增補字符集》作出修訂，正式將中諮會所通過增收的 113 個字符，納入《香港增補字符集》內。此外，修訂的另一個目的是反映本文第四段提及的 1,622 個獲納入 ISO 10646 國際編碼標準的《香港增補字符集》字符的新碼位編排。

9 經修訂的《香港增補字符集》，其載列的字符編碼表將在每個字符下列出其大五碼及在 ISO 10646 國際編碼標準下各個版本的碼位。這樣的編排可清楚地列出每個字符的有關編碼，以方便查閱。字符編碼表的建議樣本，請參閱附件二。

10 有兩個在《香港增補字符集》現行版本內的字符錯誤對應個別的 ISO 10646 碼位。我們建議在《香港增補字符集》的修訂版本內作出有關更正，詳情請參閱附件三。

11 經修訂的《香港增補字符集》，其建議的名稱是《香港增補字符集—2001》，而一九九九年九月公布的版本則沿用《香港增補字符集》的名稱。

12 我們將會把《香港增補字符集—2001》的技術性文件及樣版文字輸入工具，存放在「數碼 21」資訊科技策略網站上供市民下載。為方便查閱，我們亦會將有關《香港增補字符集》的技術性文件繼續保留在該網站上。

13 為拉近《香港增補字符集—2001》及 ISO 10646 國際編碼標準新版本各自的公布日期，我們會盡快將《香港增補字符集—2001》文件草稿提

供給「中文電腦用字工作小組」及「中文資訊科技工作小組」各成員審閱。待收納兩個工作小組成員的意見後，我們會將《香港增補字符集—2001》提交給中諮會審批，並安排向公眾發布獲中諮會通過的《香港增補字符集—2001》。

資訊科技署
二零零一年十一月

附件一

《香港增補字符集》新增字符

自二零零零年四月公布《香港增補字符集》字符增收程序和原則以來，中諮會經處理市民及政府部門提交的新增字符申請後，共增收了 113 個字符。其中 75 個字符已獲收納在 ISO 10646 國際編碼標準的 ISO/IEC 10646-1:2000版本內，在餘下的38個字符中，有29個將獲收納在ISO 10646國際編碼標準的下一個新版本內，剩下的 9 個字符則將放在其私人使用區。有關 113 個新增字符及它們在 ISO 10646 國際編碼標準內的分佈，請參閱附載的列表。

附件一

《香港增補字符集》新增字符

序號	字形	大五碼	ISO 10646-1:1993	ISO 10646-1:2000	ISO 10646-2:2001	碼位區
1	倻	8C40	503B	503B	0-503B	CJK
2	兞	8C41	6DFE	6DFE	0-6DFE	CJK
3	鬻	8C42	F579	F579	2-9C73	Ext-B
4	熚	8C43	F57A	F57A	F57A	PUA
5	尉	8C44	F57B	3DC9	0-3DC9	Ext-A
6	祜	8C45	888F	888F	0-888F	CJK
7	灦	8C46	F57D	F57D	2-414E	Ext-B
8	弁	8C47	7077	7077	0-7077	CJK
9	嵿	8C48	5CF5	5CF5	0-5CF5	CJK
10	飛	8C49	F580	4B20	0-4B20	Ext-A
11	晫	8C4A	F581	F581	2-51CD	Ext-B
12	兟	8C4B	F582	3559	0-3559	Ext-A
13	簪	8C4C	F583	F583	2-5D30	Ext-B
14	愚	8C4D	6122	6122	0-6122	CJK
15	銙	8C4E	F585	F585	2-8A32	Ext-B
16	辦	8C4F	8FA7	8FA7	0-8FA7	CJK
17	鈊	8C50	91F6	91F6	0-91F6	CJK
18	燅	8C51	7191	7191	0-7191	CJK
19	䏌	8C52	6719	6719	0-6719	CJK
20	垔	8C53	73BA	73BA	0-73BA	CJK
21	瞽	8C54	F58B	F58B	2-3281	Ext-B
22	鷂	8C55	F58C	F58C	2-A107	Ext-B
23	毘	8C56	F58D	3C8B	0-3C8B	Ext-A
24	鼒	8C57	F58E	F58E	2-1980	Ext-B
25	飋	8C58	F58F	4B10	0-4B10	Ext-A
26	礙	8C59	78E4	78E4	0-78E4	CJK
27	琯	8C5A	7402	7402	0-7402	CJK
28	江	8C5B	51AE	51AE	0-51AE	CJK
29	鄑	8C5C	F593	F593	2-870F	Ext-B
30	盡	8C5D	F594	4009	0-4009	Ext-A
31	檜	8C5E	6A63	6A63	0-6A63	CJK
32	齬	8C5F	F596	F596	2-A2BA	Ext-B
33	篟	8C60	F597	4223	4223	Ext-A
34	藾	8C61	860F	860F	0-860F	CJK
35	稷	8C63	7A2A	7A2A	0-7A2A	CJK
36	驟	8C64	F59B	F59B	2-9947	Ext-B
37	鏌	8C65	F59C	F59C	2-8AEA	Ext-B
38	靖	8C66	9755	9755	0-9755	CJK
39	瀾	8C67	704D	704D	0-704D	CJK
40	匡	8C68	5324	5324	0-5324	CJK
41	帒	8C69	F5A0	F5A0	2-207E	Ext-B
42	鏘	8C6A	93F4	93F4	0-93F4	CJK
43	盡	8C6B	76D9	76D9	0-76D9	CJK
44	鐃	8C6C	F5A3	F5A3	2-89E3	Ext-B
45	曙	8C6D	F5A4	F5A4	F5A4	PUA

附件一

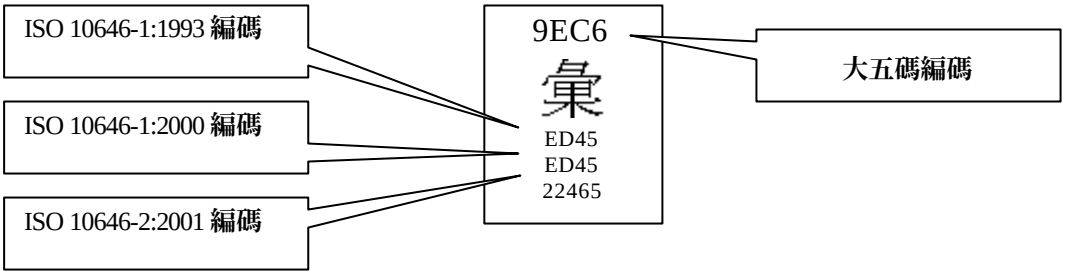
序號	字形	大五碼	ISO 10646-1:1993	ISO 10646-1:2000	ISO 10646-2:2001	碼位區
46	矜	8C6E	77DD	77DD	0-77DD	CJK
47	亢	8C6F	4EA3	4EA3	0-4EA3	CJK
48	俪	8C70	4FF0	4FF0	0-4FF0	CJK
49	僕	8C71	50BC	50BC	0-50BC	CJK
50	丰	8C72	4E2F	4E2F	0-4E2F	CJK
51	众	8C73	4F17	4F17	0-4F17	CJK
52	匯	8C74	F5AB	F5AB	F5AB	PUA
53	吴	8C75	5434	5434	0-5434	CJK
54	綋	8C76	7D8B	7D8B	0-7D8B	CJK
55	墒	8C77	5892	5892	0-5892	CJK
56	壅	8C78	58D0	58D0	0-58D0	CJK
57	峰	8C79	F5B0	F5B0	2-1DB6	Ext-B
58	庄	8C7A	5E92	5E92	0-5E92	CJK
59	庙	8C7B	5E99	5E99	0-5E99	CJK
60	禳	8C7C	5FC2	5FC2	0-5FC2	CJK
61	愈	8C7D	F5B4	F5B4	2-2712	Ext-B
62	斋	8C7E	658B	658B	0-658B	CJK
63	杈	8CA1	F5B6	F5B6	2-33F9	Ext-B
64	椙	8CA2	6919	6919	0-6919	CJK
65	檄	8CA3	6A43	6A43	0-6A43	CJK
66	氩	8CA4	F5B9	F5B9	2-3C63	Ext-B
67	浪	8CA5	6CFF	6CFF	0-6CFF	CJK
68	爀	8CA7	7200	7200	0-7200	CJK
69	爬	8CA8	F5BD	F5BD	2-4505	Ext-B
70	玔	8CA9	738C	738C	0-738C	CJK
71	瑒	8CAA	F5BF	3EDB	0-3EDB	Ext-A
72	琛	8CAB	F5C0	F5C0	2-4A13	Ext-B
73	嬅	8CAC	5B15	5B15	0-5B15	CJK
74	璹	8CAD	74B9	74B9	0-74B9	CJK
75	讚	8CAE	8B83	8B83	0-8B83	CJK
76	簾	8CAF	F5C4	F5C4	2-5CA4	Ext-B
77	禖	8CB0	F5C5	F5C5	2-5695	Ext-B
78	窓	8CB1	7A93	7A93	0-7A93	CJK
79	箸	8CB2	7BEC	7BEC	0-7BEC	CJK
80	糁	8CB3	7CC3	7CC3	0-7CC3	CJK
81	繕	8CB4	7E6C	7E6C	0-7E6C	CJK
82	莘	8CB5	82F8	82F8	0-82F8	CJK
83	蔺	8CB6	8597	8597	0-8597	CJK
84	蕲	8CB7	F5CC	F5CC	F5CC	PUA
85	祕	8CB8	8890	8890	0-8890	CJK
86	赀	8CB9	F5CE	F5CE	F5CE	PUA
87	躺	8CBA	8EB9	8EB9	0-8EB9	CJK
88	幹	8CBB	F5D0	F5D0	F5D0	PUA
89	迖	8CBC	8FCF	8FCF	0-8FCF	CJK
90	發	8CBD	855F	855F	0-855F	CJK

序號	字形	大五碼	ISO 10646-1:1993	ISO 10646-1:2000	ISO 10646-2:2001	碼位區
91	駺	8CBE	99E0	99E0	0-99E0	CJK
92	鉀	8CBF	9221	9221	0-9221	CJK
93	鋼	8CC0	F5D5	F5D5	F5D5	PUA
94	關	8CC1	F5D6	F5D6	2-8DB9	Ext-B
95	墶	8CC2	F5D7	F5D7	2-143F	Ext-B
96	曉	8CC3	F5D8	4071	0-4071	Ext-A
97	漿	8CC4	F5D9	42A2	0-42A2	Ext-A
98	娛	8CC5	5A1A	5A1A	0-5A1A	CJK
99	奠	8CC9	9868	9868	0-9868	CJK
100	杜	8CCA	676B	676B	0-676B	CJK
101	簍	8CCB	F5E0	4276	0-4276	Ext-A
102	坳	8CCC	573D	573D	0-573D	CJK
103	贙	8CCE	85D6	85D6	0-85D6	CJK
104	斑	8CCF	F5E4	F5E4	2-497B	Ext-B
105	芿	8CD0	82BF	82BF	0-82BF	CJK
106	蘊	8CD1	F5E6	F5E6	2-710D	Ext-B
107	鰥	8CD2	F5E7	4C81	0-4C81	Ext-A
108	蘖	8CD3	F5E8	F5E8	2-6D74	Ext-B
109	嵣	8CD4	5D7B	5D7B	0-5D7B	CJK
110	芚	8CD5	F5EA	F5EA	2-6B15	Ext-B
111	蓬	8CD6	F5EB	F5EB	2-6FBE	Ext-B
112	騛	8CD7	F5EC	F5EC	F5EC	PUA
113	驥	8CD8	F5ED	F5ED	F5ED	PUA

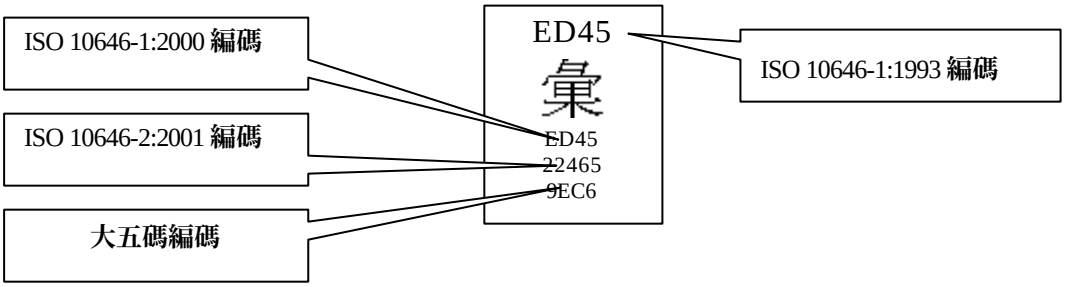
附件二

《香港增補字符集—2001》
字符編碼表的建議樣本

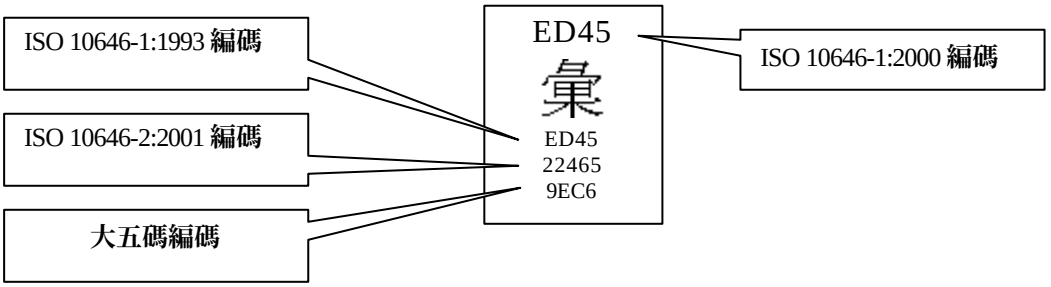
圖例一 ： 香港增補字符集對應大五碼的內碼編排



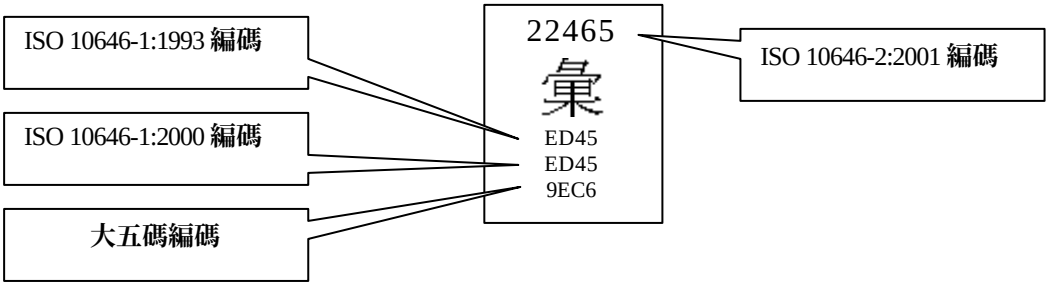
圖例二 ： 香港增補字符集對應 ISO 10646-1:1993 的內碼編排



圖例三 ： 香港增補字符集對應 ISO 10646-1: 2000 的內碼編排



圖例四: 香港增補字符集對應 ISO 10646-2: 2001 的內碼編排



附件三

《香港增補字符集》字符的編碼更正

就《香港增補字符集》已收納的下列二個字符，需要在《香港增補字符集》的修訂版本內作出更正如下:

1 在大五碼字集內的“龜”字，其對應的 ISO 10646 編碼是 9F9C。《香港增補字符集》內收納了另一類似的“𪓐”字，其對應的 ISO 10646 編碼亦是 9F9C。當上述兩字符在 ISO 10646 國際編碼標準及大五碼間進行碼位轉換時，將會產生錯誤的轉換結果。

現建議將已收納在《香港增補字符集》內的“𪓐”字的 ISO 10646 對應編碼由 9F9C 改為 F907。建議更正上述字符的編碼如下：

大五碼	在現時《香港增補字符集》內對應的 ISO 10646 編碼 (ISO/IEC 10646-1:1993) (ISO/IEC 10646-1:2000)	在《香港增補字符集 - 2001》內對應的 ISO 10646 編碼 (ISO/IEC 10646-1:1993) (ISO/IEC 10646-1:2000) (ISO/IEC 10646-2:2001)
C074 龜	9F9C	9F9C
8BF8 𪓐	9F9C	F907

2 現時在《香港增補字符集》內的“欒”字，其對應大五碼和 ISO 10646 (ISO10646-1:2000 版本)的編碼分別是 A0E8 和 EEA1。由於在 ISO10646-1:2000 版本內此字的編碼其實是 3C18，現建議將《香港增補字符集》內就該字符對應的 ISO 10646 編碼由 EEA1 更正為 3C18。