

國際表意文字子集比較工具

支援 ISO 10646 國際編碼標準

(Linux 平台)

使用手冊

版本： 1.0

二零零七年五月

© 香港特別行政區政府

政府資訊科技總監辦公室

版權所有，不得用任何方式抄襲或翻印

目錄

| | |
|-------------------|---|
| 1. 簡介 | 1 |
| 2. 背景 | 1 |
| 3. 「工具軟件」概覽 | 2 |
| 4. 系統需求 | 3 |
| 4.1 硬件需求 | 3 |
| 4.2 軟件需求 | 3 |
| 5. 安裝步驟 | 4 |
| 6. 操作「工具軟件」 | 5 |
| 7. 移除步驟 | 6 |
| 8. 常見問題 | 6 |

1. 簡介

國際標準化組織 (International Organization for Standardization, ISO) 發表一個名為 International Ideographs Core (IICORE) 的表意文字子集。國際表意文字子集包含 9,810 個字符，可方便用戶於記憶體容量、輸入/輸出能力有限的設備上以中文作日常電子通訊。詳細資料，請瀏覽以下網頁：

<http://www.info.gov.hk/digital21/unicode/structure/iicore.html>。

網站管理員可以利用國際表意文字子集比較工具 (以下簡稱「工具軟件」) 將一些會在資源有限的設備(例如一些支援國際表意文字子集的流動平台) 上瀏覽的網頁內的中文字與 IICORE 作比較。「工具軟件」接納一個 UTF-8 編碼的純文字檔案(包括 UTF-8 編碼的 HTML 檔案)，然後列出檔案中不包括在 IICORE 內的中文字。結果儲存在一個 UTF-8 編碼的 HTML 檔案內。

有關 ISO 10646 國際編碼標準 和《香港增補字符集-2004》的詳細資料，請瀏覽「數碼 21」資訊科技策略網站：

http://www.info.gov.hk/digital21/unicode/structure/intro_iso.html。

2. 背景

由於中文是香港大部分市民所用的語文，因此，為中文電子通訊的使用者制訂一套共通的界面十分重要。

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開放和共通的中文界面。共通中文界面的重點是採用 ISO 10646 國際編碼標準。該標準是由國際標準化組織 (ISO) 發展的一套應用於儲存及交換電子資料的國際編碼標準。

政府積極參與發展 ISO 10646 國際編碼標準的工作。國際標準化組織已於二零零五年十一月在 ISO/IEC 10646:2003 第一修訂版中公布 IICORE。藉著提供一個國際表意文字子集，國際間的中文電子通訊便能得以改善。

3. 「工具軟件」概覽

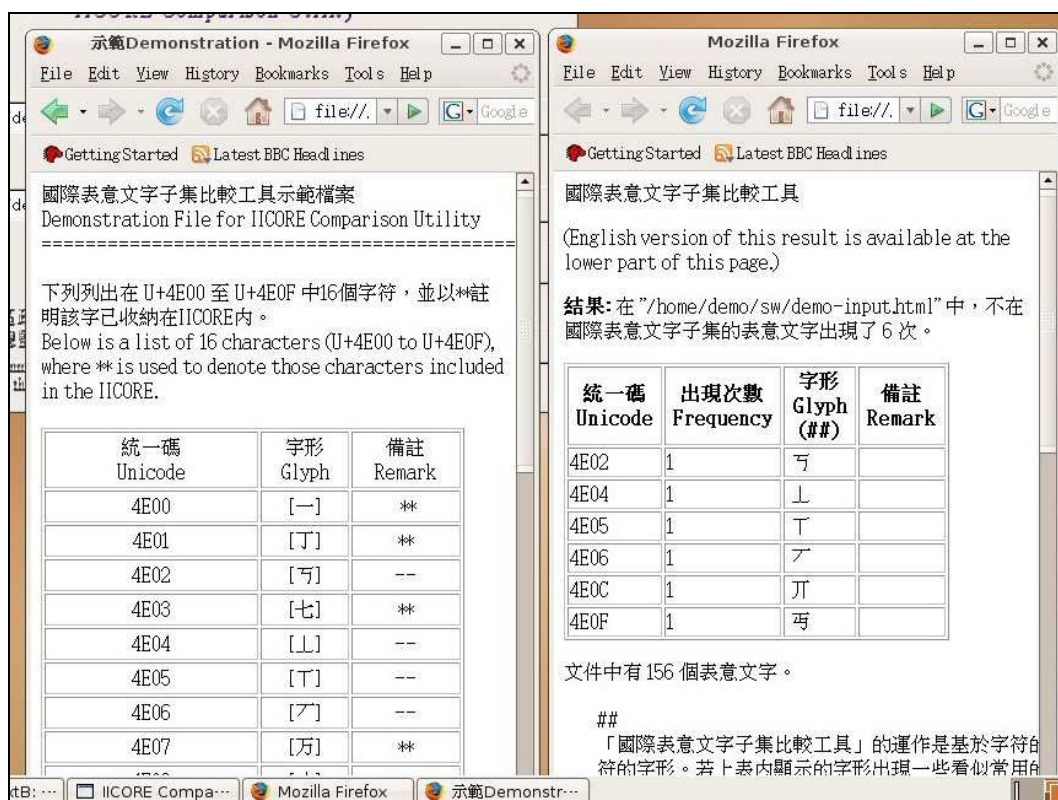
| | |
|--|----------------------|
| 「工具軟件」組件 | 檔案名稱 |
| 裝載着 iicorecompare.jar , demo-input.html 和 notice.html 的檔案 | iicorecompare.tar.gz |
| 「工具軟件」程式 (「工具軟件」必須在 J2SE Runtime Environment 5.0 上執行) | iicorecompare.jar |
| 示範用的 UTF-8 編碼的純文字檔案 | demo-input.html |
| 使用條款的資訊 | notice.html |

本使用手冊提供在 Linux 平台上安裝及使用「工具軟件」的步驟及其他相關資料。

以下顯示一個示範檔案和執行「工具軟件」後的結果以作參考。

示範的輸入檔案

示範輸入檔案的結果



4. 系統需求

4.1 硬件需求

- 等同 Pentium III 300 MHz 的處理器或以上；
- 256 MB 或以上的記憶體；
- 50 MB 硬碟空間；及
- 800x600 或以上的螢幕解像度。

4.2 軟件需求

- Ubuntu Linux (以下簡稱「平台」)¹；
- J2SE Runtime Environment 5.0 (JRE 5.0) ；及
- 支援《香港增補字符集》字型軟件。

注意事項：

關於 JRE 5.0 或應用軟件的使用方法，我們建議用戶參考該軟件的使用手冊，或瀏覽軟件供應商在網站提供的資料。

4.2.1 「平台」

- 下載及安裝 JRE 5.0。這個軟件可在 <http://java.sun.com> 下載； 及
- 下載及安裝支援 Unicode 的網頁瀏覽器，例如 Firefox：
<http://www.mozilla.com/firefox>。

4.2.2 Java Runtime Environment

用戶可參考以下步驟查看「平台」上 Java Runtime Environment 的版本：

- 按以下圖示：應用程式(Applications)>>附屬應用程式(Accessories)>>終端機(Terminal)；
- 輸入命令 `java -version` 按<Enter>鍵；
- 確認版本是「1.5.0」或以上。JRE5.0 的內部版本編號為 1.5.0。詳細資料可瀏覽以下網頁：
<http://java.sun.com/developer/technicalArticles/releases/j2se15>。



¹ 本「工具軟件」在 Ubuntu Desktop 7.04 上作示範。關於 Linux 「平台」安裝和使用方法，我們建議用戶參考該軟件的使用手冊，或瀏覽軟件供應商在網站提供的資料例如：
<http://www.ubuntu.com> 和 <https://help.ubuntu.com/community/Java>。

4.2.3 《香港增補字符集》字型軟件

「工具軟件」支援《香港增補字符集》的最新版本(《香港增補字符集-2004》)。如輸入檔案中包含有《香港增補字符集-2004》新增收的字符，用戶便可能需要安裝支援《香港增補字符集-2004》的字型軟件以正確地檢視這些字符。

技術資料：

在 Java Runtime Environment 中設定中文字體的詳細步驟，可參考相關技術的網頁，例如：

http://docs.sun.com/app/docs/doc/817-7970/6mn8kfs4h?l=zh_TW&a=view 或

<http://java.sun.com/j2se/1.5.0/docs/guide/intl/fontconfig.html> 。

5. 安裝步驟

(在進行安裝前，請再次確定你的平台符合本使用手冊第4章所列的各項需求。)

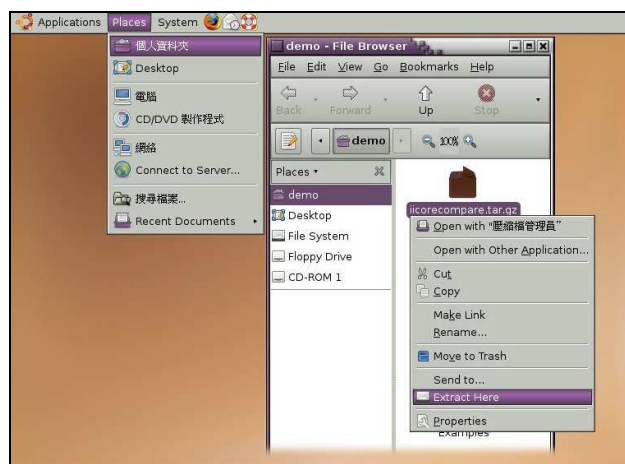
步驟一：在以下網站下載檔案「iicorecompare.tar.gz」至「平台」上：

http://www.info.gov.hk/digital21/unicode/structure/cli_tools.html 。

步驟二：將檔案移到你的資料夾，例如\home\demo。

步驟三：按以下圖示：「位置」(Place)>> 「個人資料夾」(Personal Folder)。在 iicorecompare.tar.gz 上單擊滑鼠右鍵選取「解壓縮至此」(Extract Here)。

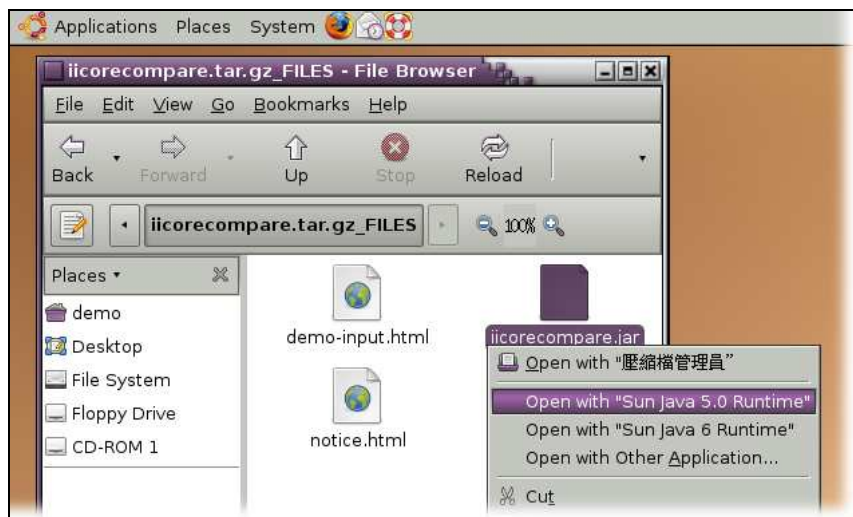
步驟四：雙擊「iicorecompare.tar.gz_FILES」資料夾後有 iicorecompare.jar 即表示已完成。



6. 操作「工具軟件」

步驟一：雙擊「個人資料夾」>>「iicorecompare.tar.gz_FILES」資料夾。

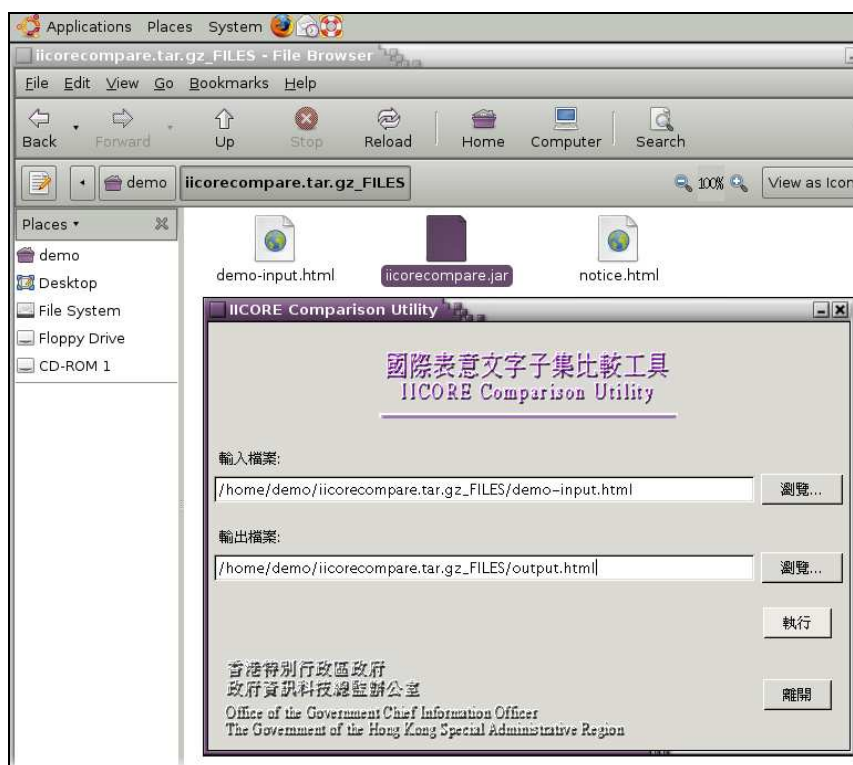
步驟二：在 iicorecompare.jar 上單擊滑鼠右鍵選取「用 Sun Java 5.0 Runtime 打開」(Open with Sun java 5.0 Runtime)。



步驟三：按「瀏覽」按鈕選擇輸入檔案。

步驟四：按另一個「瀏覽」按鈕選擇輸出檔案。

步驟五：按「執行」按鈕。完成後，用戶可用網頁瀏覽器開啟輸出檔案查看結果。



7. 移除步驟

用戶刪除在第三章描述的檔案便可移除「工具軟件」。

8. 常見問題

1. 「工具軟件」有什麼用途？

網站管理員可以利用「工具軟件」將一些會在資源有限的設備(例如一些支援國際表意文字子集的流動平台)上瀏覽的網頁內的中文字與 IICORE 作比較。

2. 「工具軟件」能否處理 UTF-8 編碼的 HTML 檔案？

「工具軟件」能處理 UTF-8 編碼的 HTML 檔案。不過，它不支援 HTML 檔案內以數字描述的文字(例如一)。用戶要將這些檔案轉換成 ISO 10646 編碼的 UTF-8 格式(例如 E4 B8 80)。有關 UTF-8 的詳細技術資料，請瀏覽以下網頁：

<http://www.unicode.org/cgi-bin/GetUnihanData.pl?codepoint=4e00> 和
http://www.unicode.org/faq/utf_bom.html。

3. 「工具軟件」能否處理大五碼編碼的檔案？

用戶要先將這些大五碼編碼的檔案轉換成 ISO 10646 檔案編碼的 UTF-8 格式。遇到不是 UTF-8 編碼的檔案，「工具軟件」會停止處理並會顯示提示訊息。

< 完 >