

Advice on the Compatibility Issue of existing application software with New Platforms

1. Introduction

Currently, many application software can handle data coded in ISO/IEC 10646-1:2000 (equivalent to Unicode 3.0) and HKSCS-2001 (Hong Kong Supplementary Character Set -2001). HKSCS-2001 contains 4,818 characters that are commonly used in Hong Kong. More than 1,600 of them are assigned temporarily in the Private Use Area (PUA) of the ISO/IEC 10646-1:2000. These characters in the PUA are coded in 16-bit. In ISO/IEC 10646:2003 with Amendment 1 (equivalent to Unicode 5.0), these characters are assigned with proper code points, most of them are coded in 32-bit.

2. Compatibility issue of existing application software with the New Platforms

Many application software can handle data coded in Unicode 3.0 and HKSCS-2001. Apart from increase in Chinese characters newly included, Unicode 5.0-enabled platforms (the “New Platforms”) also bundle Chinese input methods that enable users to input data coded in Unicode 5.0 accordingly. However, they do not provide Chinese input methods that are compatible to Unicode 3.0 and HKSCS-2001. As application software developers may take time and effort to gradually update their software to support Unicode 5.0, hence these application software may not be able to handle all data input on these New Platforms, particularly those 32-bit characters, including HKSCS.

3. Suggestion

Software platform vendors and developers of Chinese input methods should consider the following suggestions to facilitate users to gradually migrate to the New Platforms:

- New Platforms provide Chinese input methods that can enable user to input Unicode 3.0 and HKSCS-2001 characters. This can let users of the New Platforms temporary avoid inputting characters which may not be properly displayed on the existing platforms or handled by these application software.
- New Platforms provide Chinese font that supports Unicode 3.0 and

HKSCS-2001. This allows users of the New Platform to display documents or file created on the existing platforms properly.

- New Platforms provide code conversion software
- When more and more application software can handle data coded in Unicode 5.0, users of the New platforms can then switch to the Unicode 5.0-enabled Chinese input methods.

4. Reference

Cangjie Input Code Table of HKSCS-2001 for ISO/IEC 10646-1:2000 is available for download at

<http://www.ogcio.gov.hk/ccli/eng/hkscs/terms/terms7.html>

There are sample tests for Chinese input method that support ISO/IEC 10646-1:2000 and HKSCS-2001

HKSCS-2001 Character	Cangjie	ISO/IEC 10646-1:2000 Code point	Remarks
嘑	RJRB	35C5	
莱	TJD	82FF	
引	ONL	E141	The code point of this HKSCS-2001 character in ISO/IEC 10646-1:2000 is E141. The code point of this character in Unicode 5.0 is 201C1.

The code point of the Character [𠵹] (Cangjie : UHXU) in Unicode 5.0 is 21E23. Chinese Input method developed for compatible to Unicode 3.0 and HKSCS does not required to include this character newly included in the HKSCS-2004.