

Universal Multiple-Octet Coded Character Set UCS
---

ISO/IEC JTC1/SC2/WG2 IRG N 1506

Date: 2008-10-28

Doc. Type:	Member body contribution
Title:	Activity Report form TCA
Source:	TCA
Status:	Input to IRG
Action:	For Your Information
Distribution:	IRG Members and Ideographic Experts
No. of pages:	2
Medium:	Electronic

**1. CNS 14754 “Sorting attributes for Chinese data processing” will be published on next month:**

Based on the project from BSMI (“Bureau of Standards, Metrology and Inspection”, the organization for standardization in Taiwan), CMEX (Chinese Foundation for Digitization Technology) drafted the revised version of CNS 14754 “Sorting attributes for Chinese data processing” last year. The reviewing process of BSMI proved the new CNS 14754, and BSMI will publish it on December 2008.

CNS 14754 specifies five attributes of Chinese characters: radix, stroke count without radix, total stroke count, stroke sequence and pronunciation. The radix of a Chinese character is given by means of Kanxi radices. The pronunciation of a Chinese character is presented in the form of Liu’s code. Those attributes are intended used to sort formatted Chinese data, such as names of people. Because CNS 11643-3 specified a new classification to the basic stroke for Chinese characters, the stroke sequences must be revised in according to the new classification.

**2. “The English/Chinese vocabulary for CNS’s information technical standards and communication standards” version 3 has been published:**

Based on the project from BSMI, CMEX drafted the version 3 of “The English/Chinese vocabulary for information technical standards and communication standards” last year. The version 3 was derived from the version 2 of that vocabulary, by means of adding some new terminologies and changing the Chinese translations of some terminologies. The row materials were collected form the CNS standards for information technology and communication areas, which were published after the version 2 vocabulary. The reviewing process of BSMI proved the new English/Chinese vocabulary, and BSMI published it last month.

**3. The Research Center for Information Technology Innovation of Academia Sinica is drafting a document which will be proposed as a relative standard of CNS 11643:**

As an output of the long-term research for describing Chinese character structure, ASRCITI (the Research Center for Information Technology Innovation, Academia Sinica) is drafting a document, named “Ideograph structure description sequence (ISDS) for identifying Chinese character”.

There are more than 70 thousands of Chinese characters has been encoded in CNS 11643, but currently a conventional Chinese PC could only provide a rank of 20 thousands Chinese characters, e.g., 21,204 Chinese characters for the case of Windows XP Big-5 version. It means that more than 40 thousands of Chinese characters in CNS 11643 are “out of Chinese PC.” ASRCITI proposed the ISDS to solve that problem. ISDS is similar to IDS and it is capably to uniquely identify each Chinese character in CNS 11643. By the way, for the case “out of Chinese PC”, one could represent a Chinese character by its ISDS. ASRCITI prepared the document above to describe the ISDS and tend to submit it to the BSMI as a relative standard of CNS 11643.

**4. “The seminar of civil expert’s reports on attending international standardization activities of information & communication technologies” was hosted by the BSMI on Oct. 16, 2008:**

The BSMI supported some civil experts to attend relative international standardization activities on information and communication technology. The BSMI hosted the seminar for asking those civil exports to report their activities, October 16, 2008. The civil expert major in Chinese character encoding gave the speech “the trend of ideographic character encoding on international standardization activities” on the seminar.