Universal Multiple-Octet Coded Character Set
(UCS)

ISO/IEC JTC1/SC2/WG2 *N*
IRG N 378
Date: 1996-07-31

Source: ISO/IEC JTC1/SC2/WG2/**IRG**:
Ideographic Rapporteur Group

Title: **Draft Text of General Description for Extension A**

Action: To be submitted to WG2
for Inclusion in ISO/IEC 10646-1

Status: Supporting Paper for IRG #7 (Resolution 3)

Distribution: ISO/IEC JTC1/SC2/WG2 and WG2/IRG

JTC 1/SC2/WG2/IRG **N378**

DATE: **1996.07.12**
Supersedes:

ISO/IEC JTC 1/SC2/WG2/IRG
Ideographic Rapporteur Group
Secretariat: China

Source / Contribution Identifier : **tk9626**

Meeting : @ / for **SC2/WG2 #31**

Issue Number : Project Number :

Title : **Draft text of general description for Eextension A**

Keywords :

Status :**ISO/IEC JTC1/SC2/WG2/IRG contribution**

Short Description :
This draft text is general description of IRG Unified CJK Ideograph Extension A. When Extension A is agreed the BMP allocation by SC2/WG2, this draft text should be merged in to Clause 26 of ISO/IEC 1646-1.

Proposed Conclusion / Requested Action :
To be submitted at SC2/WG2 #31

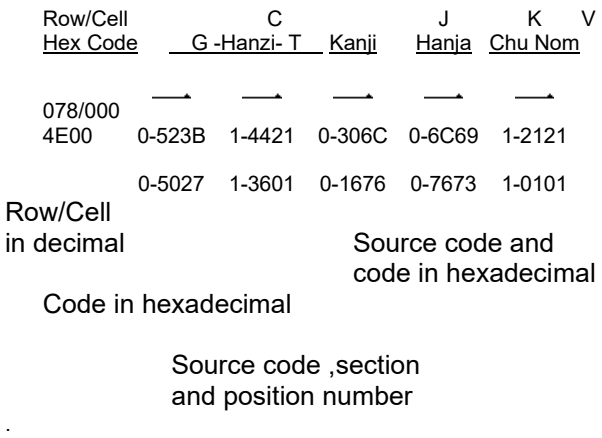
Arguments / Text of Contribution :
See attachment

Contact :
Mr. Zhang Zhoucai, Rapporteur of ISO/IEC JTC 1/SC2/WG2/IRG
Center of Computer & Microelectronics Industry Development Research
Apartment 620-621, DaTun Road, ChaoYang District, Beijing, 100101, China
Tel: +8610 6201 1118-300 Fax: +8610 6496 0428 E-mail: joezhg@public.bta.net.cn

Draft text of general description for CJK Unified Ideograph Extension A

xx CJK Unified Ideograph Extension A

Entries in the code tables for CJK (Chinese/Japanese/Korean) unified ideograph extension A are arranged as follows.



.

The leftmost column shows the code position in ISO/IEC 10646, giving the code representation both in decimal and in hexadecimal notation.

Each of the other columns shows the graphic symbol for the character, and its coded representation, as specified in a source standard for character sets that is also identified in the table entry. Each of these source standards is assigned to one of five groups indicated by G, T , J, K, or V as shown in the lists below. In each table entry, a separate column is assigned for corresponding character (if any) from each of those groups of source standards.

An entry in any of the G, T, J, K, or V columns includes a sample graphic symbol from the source character set standard, together with its coded representation in that standard. The first line below the graphic symbol shows the coded representation in hexadecimal notation. The second line shows the coded representation in

decimal notation which comprises two digits for section number followed by two digit for position number. Each of the coded representations is prefixed by a one-character source code identification followed by a hyphen. This source code character identifies the coded character set standard from which the character is taken as shown in the lists below.

Hanzi G sources are
G3 GB7589-87 unsimplified forms
G5 GB7590-87 unsimplified forms
G7 General purpose Hanzi list for modern Chinese Language
GS Singapore characters

Hanzi T sources are
T3 TCA-CNS 11643-1992 3rd plane with some additional characters
T4 TCA-CNS 11643-1992 4th plane
T5 TCA-CNS 11643-1992 5th plane
T6 TCA-CNS 11643-1992 6th plane
T7 TCA-CNS 11643-1992 7th plane
TF TCA-CNS 11643-1992 15th plane

Kanji J source is
JA Unified Japanese IT Vendors Contemporary Ideographs 1993

Hanja K source is
K3 PKS C 5700-2 1994

Chu Nom V source is
V0 TCVN 5773:1993

For CJK (Chinese/Japanese/Korean) ideographs in the BMP, the names shall be algorithmically constructed by appending their two-octet coded representation in hexadecimal notation to “CJK UNIFIED IDEOGRAPH-”. For example, the first CJK ideograph character in the BMP has the name “CJK UNIFIED IDEOGRAPH-4E00”.