

Universal Multiple-Octet Coded Character Set UCS	
ISO/IEC JTC 1/SC 2/WG 2 IRG N1771R Date: 2011-04-12	
Source:	Japan
Title:	Proposal of Encoding Architecture of Old Hanzi
Actions required:	For consideration at IRG #36
Distribution:	IRG Members and Old Hanzi Experts
Medium :	Electronic

Attached is the sample of encoding architecture of Oracle Bone script: code chart (for first 128 characters) and character name list (for first 16 characters). Also the character-glyph mapping table for first 16 characters is attached. Principles of the sample encoding architecture are follows:

- 0) Character Glyph Model: This sample is based on “Character Glyph Model” in accordance with ISO/IEC TR 15285:1998.
<http://std.dkuug.dk/jtc1/SC2/wg2/docs/TR%2015285%20-%20C027163e.pdf>
- 1) Glyph Database: This sample is based on the Glyph Database IRG N1705 developed by Old Hanzi ad Hoc.
- 2) Character Names: Each character name should be unique. Character names of this sample are mainly based on the glyph determination in Yinxu Jiagu Keci Leizuan (殷墟甲骨刻辭類纂, 姚孝遂主編, 中華書局, ISBN 9787101004779, 1989) and secondary based on the glyph determination in Jiaguwen Zidian (甲骨文字典, 徐中舒主編, 四川辭書出版社, ISBN 9787806822326, 2006). For other characters, there are short descriptions of the glyph structure.
- 3) Selection of Characters: Only stable characters are included in this sample. The stability is evaluated by the number of the evidences with meaningful sentences. Only characters that have ten or more sample sentences in Yinxu Jiagu Keci Leizuan are included.
- 4) Note: The drafting work of this sample was mainly based on Yinxu Jiagu Keci Leizuan, so some of possible characters in the standard are missing in this sample, for example, the very top glyph “一” is missing in Yinxu Jiagu Keci Leizuan because of “too many sentences”. However, the author of this document is convinced that “一” shall be included in the standard. Especially in this issue, further comments are requested, on top of other comments.

Sample code chart

XX000 - XX07F

	0	1	2	3	4	5	6	7
0	 XX000	 XX010	 XX020	 XX030	 XX040	 XX050	 XX060	 XX070
1	 XX001	 XX011	 XX021	 XX031	 XX041	 XX051	 XX061	 XX071
2	 XX002	 XX012	 XX022	 XX032	 XX042	 XX052	 XX062	 XX072
3	 XX003	 XX013	 XX023	 XX033	 XX043	 XX053	 XX063	 XX073
4	 XX004	 XX014	 XX024	 XX034	 XX044	 XX054	 XX064	 XX074
5	 XX005	 XX015	 XX025	 XX035	 XX045	 XX055	 XX065	 XX075
6	 XX006	 XX016	 XX026	 XX036	 XX046	 XX056	 XX066	 XX076
7	 XX007	 XX017	 XX027	 XX037	 XX047	 XX057	 XX067	 XX077
8	 XX008	 XX018	 XX028	 XX038	 XX048	 XX058	 XX068	 XX078
9	 XX009	 XX019	 XX029	 XX039	 XX049	 XX059	 XX069	 XX079
A	 XX00A	 XX01A	 XX02A	 XX03A	 XX04A	 XX05A	 XX06A	 XX07A
B	 XX00B	 XX01B	 XX02B	 XX03B	 XX04B	 XX05B	 XX06B	 XX07B
C	 XX00C	 XX01C	 XX02C	 XX03C	 XX04C	 XX05C	 XX06C	 XX07C
D	 XX00D	 XX01D	 XX02D	 XX03D	 XX04D	 XX05D	 XX06D	 XX07D
E	 XX00E	 XX01E	 XX02E	 XX03E	 XX04E	 XX05E	 XX06E	 XX07E
F	 XX00F	 XX01F	 XX02F	 XX03F	 XX04F	 XX05F	 XX06F	 XX07F

Sample character name list

HEX	NAME
XX000	ORACLE BONE ORIGIN
XX001	ORACLE BONE AFFAIR
XX002	ORACLE BONE TOP
XX003	ORACLE BONE UNDER
XX004	ORACLE BONE EMPEROR
XX005	ORACLE BONE SIDE
XX006	ORACLE BONE SPIRIT 1
XX007	ORACLE BONE SPIRIT 2
XX008	ORACLE BONE ENDURE
XX009	ORACLE BONE SPIRIT AND DIPPER
XX00A	ORACLE BONE WORSHIP 1
XX00B	ORACLE BONE PROTECTION
XX00C	ORACLE BONE BLESSING
XX00D	ORACLE BONE WORSHIP 2
XX00E	ORACLE BONE SPIRIT AND FENCE
XX00F	ORACLE BONE HAPPINESS
...	...

HEX	NAME

Character – Glyph mapping for sample code chart

HEX	NAME	CHAR	GLYPHS
XX000	ORACLE BONE ORIGIN	 0023-0	        
XX001	ORACLE BONE AFFAIR	 2933-0	   
XX002	ORACLE BONE TOP	 1116-0	    
XX003	ORACLE BONE UNDER	 1117-0	    
XX004	ORACLE BONE EMPEROR	 1132-0	                           
XX005	ORACLE BONE SIDE	 3120-0	             
XX006	ORACLE BONE SPIRIT 1	 1118-0	                    
XX007	ORACLE BONE SPIRIT 2	 1119-0	 
XX008	ORACLE BONE ENDURE	 1122-0	        
XX009	ORACLE BONE SPIRIT AND DIPPER	 1124-0	                
XX00A	ORACLE BONE WORSHIP 1	 1856-0	             
XX00B	ORACLE BONE PROTECTION	 0916-0	   
XX00C	ORACLE BONE BLESSING	 0306-0	                   
XX00D	ORACLE BONE WORSHIP 2	 0915-0	                   

XX00E	ORACLE BONE SPIRIT AND FENCE	 2936-0	 0188	 0189	 0190	 0191	 0192			
XX00F	ORACLE BONE HAPPINESS	 1123-0	 0198	 0199	 0215	 0232	 0236	 0237	 0238	 0239
...	...									

(end of document)

ISO/IEC JTC1/SC2/WG2/IRG N1771_ROKFeedback
Doc. #: Korea JTC1/SC2 K2052B

Korea JTC1/SC2, Committee on Character Codes

KIM, Kyongsok, Chairperson of Korea JTC1/SC2
email: gimgs0 AT gmail DOT com, phone: +82-51-510-2292
address: Division of Computer Science and Engineering,
Pusan National University
2 Busandaehagro 63 Beon Gil, Geumjeonggu, BUSAN 609-735, Rep. of KOREA

Author: KIM, Kyongsok
Date: 2011.06.09.
Status: NB position
Subject: R.O.Korea's comments RE: IRGN1771 (Japan's Proposal of Oracle Bone
Coding Framework)
Relevant documents: IRG N1771

1. Related IRG resolution (M36.2):

Resolution IRG M36.9: Old Hanzi Encoding Architecture Unanimous
Action: All Member Bodies

The IRG accepts the recommendation from the Old Hanzi Expert Group and requests its members to review IRGN1771 and provides feedback for discussion in IRG#37.

2. R. O. Korea's feedback:

R. O. Korea discussed this issue within domestic committee and concluded as follows:

R. O. Korea thinks Japan's approach using character glyph model is appropriate.

Some comments RE: IRG N1771 are shown below:

2.1 Criteria for selecting/choosing a representative form/glyph

– In Table in IRG N1771: Character – Glyph mapping for sample code chart on page 4, out of several glyphs, one character is chosen as a representative form or glyph. R.O.Korea wonders if a consensus is made within Old Hanzi group regarding a criterion for selecting a representative form out of several glyphs. (In N1771, there seems no such explanation.)

HEX	NAME	CHAR	GLYPHS										
XX000	ORACLE BONE ORIGIN												
XX001	ORACLE BONE AFFAIR												

2.2 Displaying a character: a mechanism to select a specific variant?

Unlike ordinary CJK ideographs, there seem not so small variations among glyphs belonging to one character. R.O.Korea wonders, once Oracle Bone characters are encoded, if one representative glyph will be used for display or somehow different glyph will be selected using some mechanism.

When we consider encoding Oracle Bone characters, it will be better to consider display issue at the same time. Therefore, R.O.Korea suggests that the current proposal N1771 be augmented with display issue.

(Of course, it is not necessary to use one fixed display method for displaying Oracle Bone characters; however, we need to consider what alternative(s) can be used for displaying Oracle Bone characters.)

2.3 Names of Oracle Bone characters

In N1771, names have meaning of the character such as "ORIGIN", "AFFAIR". If we are going to encode thousands of Oracle Bone characters, will it be desirable to name characters in that way? Or will be better not to include meaning in character names as we did with CJK unified or CJK compatible character names.

* * *