Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type:	**Working Group Document**
Title:	**Proposal to Encode Radicals for the Jurchen Script**
Source:	**Andrew West, Sun Bojun (孙伯君), Michael Everson**
Status:	**Individual Contribution**
Action:	**For consideration by JTC1/SC2/WG2 and UTC**
Date:	**2024-06-12**

1. Introduction

This is a proposal to encode a set of fifty-one radicals for the Jurchen script (see WG2 N5261 *Proposal to Encode the Jurchen Script*). The Jurchen script does not incorporate a set of graphic components which act as semantic signifiers, as is the case with the Han script. However, for convenience of character lookup, the modern Chinese dictionary of Jurchen, *Nǚzhēnwén Cídiǎn* 女真文辞典 'Dictionary of Jurchen' (Beijing: Wenwu chubanshe, 1984), compiled by Jīn Qǐzōng 金啓孮 (1918–2004), orders Jurchen characters according to fifty artificial radicals. The ordering of Jurchen ideographs in the *Proposal to Encode the Jurchen Script* follows the fifty radicals used in Jīn Qǐzōng's 1984 dictionary, but disunifies Jin's Radical 冂 into two separate radicals (23 冂 and 24 ⺆), so uses fifty-one radicals in total. It is essential to be able to represent these radicals in the code charts for the ISO/IEC 10646 and Unicode standards, so these Jurchen radicals require encoding at the same time as the set of Jurchen ideographs are encoded. Many of the Jurchen radicals are the same or very similar to CJK radicals in the Kangxi radicals block or the CJK Radicals Supplement block, or to Tangut components, and some of the other Jurchen radicals are the same as CJK unified ideographs. However, CJK ideographs and CJK radicals have a script property of 'Han', and Tangut components have a script property of 'Tangut', so it would not be appropriate to extend their usage to Jurchen radicals. Therefore we propose to encode a set of fifty-one Jurchen radicals separately at 191A0..191D2 in a new 'Jurchen Radicals' block.

Table 1 lists the fifty-one proposed radicals, with the corresponding character range for each radical in the *Proposal to Encode the Jurchen Script*.

Table 1. Proposed Jurchen Radicals

C/P	No.	Glyph	Confusable Han/Tangut Radicals	Character Range	Count
191A0	1	一	U+2F00 一 KANGXI RADICAL ONE U+18800 一 TANGUT COMPONENT-1	18E00..18E35	54
191A1	2	ヲ		18E36..18E48	19
191A2	3	入		18E49..18E4E	6
191A3	4	十	U+2F17 十 KANGXI RADICAL TEN U+1880E 十 TANGUT COMPONENT-15	18E4F..18E71	35
191A4	5	ナ		18E72..18E82	17
191A5	6	厂	U+2F1A 厂 KANGXI RADICAL CLIFF U+1880B 厂 TANGUT COMPONENT-12	18E83..18E8A	8
191A6	7	㇆	U+1881B ㇆ TANGUT COMPONENT-28	18E8B..18EB0	38
191A7	8	ㄥ		18EB1..18EB8	8
191A8	9	‡	U+18829 ‡ TANGUT COMPONENT-42	18EB9..18EC0	8
191A9	10	干	U+2F32 干 KANGXI RADICAL DRY	18EC1..18EC7	7
191AA	11	土	U+2F1F 土 KANGXI RADICAL EARTH	18EC8..18ED7	16
191AB	12	大	U+2F24 大 KANGXI RADICAL BIG	18ED8..18EE0	9
191AC	13	大		18EE1..18EE5	5

C/P	No.	Glyph	Confusable Han/Tangut Radicals	Character Range	Count
191AD	14	王		18EE6..18EEF	10
191AE	15	正		18EF0..18EF1	2
191AF	16	木	U+2F4A 木 KANGXI RADICAL TREE	18EF2..18EFF	14
191B0	17	丨	U+2F01 丨 KANGXI RADICAL LINE U+18801 丨 TANGUT COMPONENT-2	18F00..18F08	9
191B1	18	卜	U+2F18 卜 KANGXI RADICAL DIVINATION	18F09..18F10	8
191B2	19	刂	U+2E89 刂 CJK RADICAL KNIFE TWO	18F11..18F13	3
191B3	20	艹	U+2EBE 艹 CJK RADICAL GRASS ONE U+1882C 艹 TANGUT COMPONENT-45	18F14..18F17	4
191B4	21	艹	U+2EC0 艹 CJK RADICAL GRASS THREE	18F18	1
191B5	22	业		18F19..18F1B	3
191B6	23	冂	U+2F0C 冂 KANGXI RADICAL DOWN BOX	18F1C..18F29	14
191B7	24	冖	U+2F0D 冖 KANGXI RADICAL COVER	18F2A..18F3C	19
191B8	25	小	U+2E8C 小 CJK RADICAL SMALL ONE	18F3D..18F5D	33
191B9	26	山	U+2F2D 山 KANGXI RADICAL MOUNTAIN	18F5E..18F68	11
191BA	27	口	U+2F1D 口 KANGXI RADICAL MOUTH	18F69..18F79	17

C/P	No.	Glyph	Confusable Han/Tangut Radicals	Character Range	Count
191BB	28	日	U+2F47 日 KANGXI RADICAL SUN	18F7A..18F83	10
191BC	29	中		18F84..18F9B	24
191BD	30	田	U+2F65 田 KANGXI RADICAL FIELD	18F9C..18FB2	23
191BE	31	⺀	U+18802 ⺀ TANGUT COMPONENT-3	18FB3..19013	97
191BF	32	丿	U+2F03 丿 KANGXI RADICAL SLASH	19014..1903F	44
191C0	33	⺊		19040..19048	9
191C1	34	⺘		19049..1906C	36
191C2	35	亻	U+2E85 亻 CJK RADICAL PERSON U+18815 亻 TANGUT COMPONENT-22	1906D..19090	36
191C3	36	犭		19091..1909A	10
191C4	37	扌		1909B..190A2	8
191C5	38	厂	U+2E81 厂 CJK RADICAL CLIFF	190A3..190AE	12
191C6	39	人	U+2F08 人 KANGXI RADICAL MAN	190AF..190C9	27
191C7	40	乂		190CA..190D3	10
191C8	41	又	U+2F1C 又 KANGXI RADICAL AGAIN	190D4..190D8	5

C/P	No.	Glyph	Confusable Han/Tangut Radicals	Character Range	Count
191C9	42	夊	U+2F21 夊 KANGXI RADICAL GO	190D9..190DB	3
191CA	43	厶	U+2F1B 厶 KANGXI RADICAL PRIVATE	190DC..190EC	17
191CB	44	帀		190ED..190F9	13
191CC	45	月	U+2F49 月 KANGXI RADICAL MOON	190FA..190FF	6
191CD	46	し	U+2E83 乚 CJK RADICAL SECOND TWO U+18806 乚 TANGUT COMPONENT-7	19100..19110	17
191CE	47	乚	U+18822 乚 TANGUT COMPONENT-35	19111..19121	17
191CF	48	乚		19122..19128	7
191D0	49	、	U+2F02 丶 KANGXI RADICAL DOT	19129..19140	24
191D1	50	⺍	U+18AF3 ⺍ TANGUT COMPONENT-756	19141..19189	73
191D2	51	火	U+2F55 火 KANGXI RADICAL FIRE	1918A..19191	8

2. Technical Details

2.1 Block Allocation

It is suggested to encode the 51 Jurchen radicals in a block named 'Jurchen Radicals' occupying the range 191A0..191DF (64 code points).

2.2 Character Names

The characters should be named consecutively as JURCHEN RADICAL-1 through JURCHEN RADICAL-51.

2.3 Character Properties

All Jurchen radicals should have the following Unicode properties:

Script = Common
General Category = Lo [Letter, Other]
Canonical Combining Class = 0
Bidi Class = L [Left-to-Right]
Line Break = ID [Ideographic]
East Asian Width = W [Wide]
Ideographic = Y
Radical = N

As some of the Jurchen radicals are also used for Khitan ideographs (the as yet unencoded Khitan Large Script), we suggest to give the script property of 'Common' to Jurchen radicals, which would allow for some of the Jurchen radicals to be used to order any future repertoire of Khitan ideographs. Additionally, there are a very few radicals used for Khitan ideographs which are not used for Jurchen, and these could potentially be added to the Jurchen Radicals block in the future.

3. Radicals in Jīn Qǐzōng 1984 and Code Chart.

The radicals as presented in the Table of Contents for *Nǚzhēnwén Cídiǎn* 女真文辞典 are given following the proposal summary form below. Following that we present the Code Chart for the radicals proposed.

4. Proposal Summary Form

<table>
<tr><td colspan="2">
SO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
Please fill all the sections A, B and C below.
Please read Principles and Procedures Document (P & P) from http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html
for guidelines and details before filling this form.
Please ensure you are using the latest Form from http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html.
See also http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html for latest <i>Roadmaps.</i>
</td></tr>
</table>

A. Administrative

1. **Title:**	*__Proposal to Encode the Jurchen Script__*
2. Requester's name:	*Andrew West, Sun Bojun (孙伯君), Michael Everson*
3. Requester type (Member body/Liaison/Individual contribution):	*Individual contribution*
4. Submission date:	*2024-06-12*
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	*YES*
(or) More information will be provided later:	

B. Technical – General

1. Choose one of the following:	
a. This proposal is for a new script (set of characters):	*YES*
Proposed name of script:	*Jurchen Radicals*
b. The proposal is for addition of character(s) to an existing block:	*NO*
Name of the existing block:	
2. Number of characters in proposal:	*51*

3. Proposed category (select one from below - see section 2.2 of P&P document):

A-Contemporary	B.1-Specialized (small collection) X	B.2-Specialized (large collection)
C-Major extinct	D-Attested extinct	E-Minor extinct
F-Archaic Hieroglyphic or Ideographic		G-Obscure or questionable usage symbols

4. Is a repertoire including character names provided?	*YES*
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	*YES*
b. Are the character shapes attached in a legible form suitable for review?	*YES*
5. Fonts related:	
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	
Sun Bojun and Andrew West	
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	
Zhonghua Ziku project and Weifang Beida Jade Bird Huaguang Imagesetter Co., Ltd.	
6. References:	
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	*YES*
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	*YES*
7. Special encoding issues:	
Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?	*NO*

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at http://www.unicode.org for such information on other scripts. Also see Unicode Character Database (http://www.unicode.org/reports/tr44/) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

TP[1]PT Form number: N4102-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)
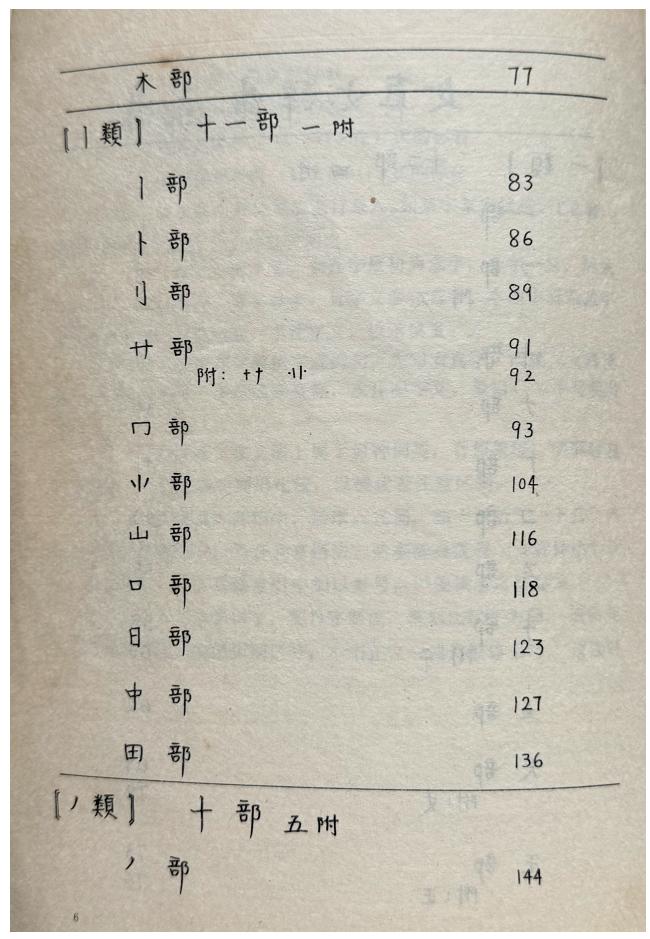
JTC1/SC2/WG2 N5278   Page 7

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	NO
If YES explain	
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?	YES
If YES, with whom?	*Jurchen experts in China*
If YES, available relevant documents:	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	NO
Reference:	
4. The context of use for the proposed characters (type of use; common or rare)	RARE
Reference:	
5. Are the proposed characters in current use by the user community?	YES
If YES, where? Reference:	
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	NO
If YES, is a rationale provided?	
If YES, reference:	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	YES
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	NO
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?	NO
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character?	YES
If YES, is a rationale for its inclusion provided?	YES
If YES, reference:	
11. Does the proposal include use of combining characters and/or use of composite sequences?	NO
If YES, is a rationale for such use provided?	
If YES, reference:	
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	
If YES, reference:	
12. Does the proposal contain characters with any special properties such as control function or similar semantics?	NO
If YES, describe in detail (include attachment if necessary)	
13. Does the proposal contain any Ideographic compatibility characters?	NO
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	

女真文辞典 部首

Nǔzhēnwén Cídiǎn 女真文辞典, page 5.

6

Nǚzhēnwén Cídiǎn 女真文辞典, page 6.

7

Nǚzhēnwén Cídiǎn 女真文辞典, page 7.

[丶類] 二部 一附

丶部

⺍部

　　附 火

計共五類 三十八部 十一附

8

Nǚzhēnwén Cídiǎn 女真文辞典, page 8.

	191A	191B	191C	191D
0	一 191A0	丨 191B0	ㄥ 191C0	丶 191D0
1	乛 191A1	卜 191B1	ㄥ 191C1	⼩ 191D1
2	又 191A2	刂 191B2	亻 191C2	火 191D2
3	十 191A3	廾 191B3	扌 191C3	
4	尹 191A4	艹 191B4	扌 191C4	
5	厂 191A5	㣺 191B5	厂 191C5	
6	乛 191A6	门 191B6	人 191C6	
7	厶 191A7	宀 191B7	乂 191C7	
8	丰 191A8	小 191B8	又 191C8	
9	干 191A9	山 191B9	夂 191C9	
A	土 191AA	口 191BA	厶 191CA	
B	大 191AB	日 191BB	帀 191CB	
C	丈 191AC	中 191BC	月 191CC	
D	王 191AD	田 191BD	し 191CD	
E	正 191AE	丿 191BE	乇 191CE	
F	木 191AF	丿 191BF	乚 191CF	

191A0 一 JURCHEN RADICAL-1
191A1 乛 JURCHEN RADICAL-2
191A2 又 JURCHEN RADICAL-3
191A3 十 JURCHEN RADICAL-4
191A4 尹 JURCHEN RADICAL-5
191A5 厂 JURCHEN RADICAL-6
191A6 乛 JURCHEN RADICAL-7
191A7 厶 JURCHEN RADICAL-8
191A8 丰 JURCHEN RADICAL-9
191A9 干 JURCHEN RADICAL-10
191AA 土 JURCHEN RADICAL-11
191AB 大 JURCHEN RADICAL-12
191AC 丈 JURCHEN RADICAL-13
191AD 王 JURCHEN RADICAL-14
191AE 正 JURCHEN RADICAL-15
191AF 木 JURCHEN RADICAL-16
191B0 丨 JURCHEN RADICAL-17
191B1 卜 JURCHEN RADICAL-18
191B2 刂 JURCHEN RADICAL-19
191B3 廾 JURCHEN RADICAL-20
191B4 艹 JURCHEN RADICAL-21
191B5 㣺 JURCHEN RADICAL-22
191B6 门 JURCHEN RADICAL-23
191B7 宀 JURCHEN RADICAL-24
191B8 小 JURCHEN RADICAL-25
191B9 山 JURCHEN RADICAL-26
191BA 口 JURCHEN RADICAL-27
191BB 日 JURCHEN RADICAL-28
191BC 中 JURCHEN RADICAL-29
191BD 田 JURCHEN RADICAL-30
191BE 丿 JURCHEN RADICAL-31
191BF 丿 JURCHEN RADICAL-32
191C0 ㄥ JURCHEN RADICAL-33
191C1 ㄥ JURCHEN RADICAL-34
191C2 亻 JURCHEN RADICAL-35
191C3 扌 JURCHEN RADICAL-36
191C4 扌 JURCHEN RADICAL-37
191C5 厂 JURCHEN RADICAL-38
191C6 人 JURCHEN RADICAL-39
191C7 乂 JURCHEN RADICAL-40
191C8 又 JURCHEN RADICAL-41
191C9 夂 JURCHEN RADICAL-42
191CA 厶 JURCHEN RADICAL-43
191CB 帀 JURCHEN RADICAL-44
191CC 月 JURCHEN RADICAL-45
191CD し JURCHEN RADICAL-46
191CE 乇 JURCHEN RADICAL-47
191CF 乚 JURCHEN RADICAL-48
191D0 丶 JURCHEN RADICAL-49
191D1 ⼩ JURCHEN RADICAL-50
191D2 火 JURCHEN RADICAL-51