

M14294 is a member of the group 書, 書 (both mean “script”). They are so-called “同形異字” (same-shape-but-non-cognate characters). There is a position to distinguish them, but it is not crucial rational to encode them (M13928 and M14294) separately.


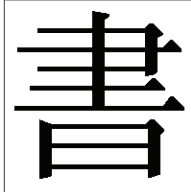
In the revision of JIS X 0208 on 1997, 29-ku-81-ten (J0-3D71) is related with M13928 and M14294.

29 区 81 点～

29-81	書 3d71 8f91 72日6	M13928	
	書 bdf1 66F8 73日6	S6412 M14294	[常] ショ, かく [詳説]

J0-3D71 Description in JIS X 0208:1997, (JIS X 0208:1997, Annex 6, p.104)

On the other hand, 戸籍統一文字 (Koseki-Touitsu-Moji, meaning “Characters for Unified Family Registry System” see IRG N1904 for the outline of Japanese governmental PUA characters and related industrial standards. The search system is reachable at <http://kosekimoji.moj.go.jp/kosekimojodb/mjko/PeopleTop>) has two different codes; 156680 for M13928, 156730 for M14294.

<p>戸籍統一文字番号 156680</p>  <p>親字・正字 昼 晝</p>	<p>音読み チュウ</p> <p>常用音読み</p> <p>訓読み ひる</p> <p>常用訓読み</p> <p>画数 10</p> <p>JIS水準</p> <p>JIS水準(JIS2004)</p> <p>文字コード Unicode Unicode(JIS2004) シフト-JIS シフト-JIS(JIS2004)</p> <p>文字区分</p>	<p>戸籍統一文字番号 156730</p>  <p>親字・正字</p>	<p>音読み ショ</p> <p>常用音読み ショ</p> <p>訓読み かく, ふみ</p> <p>常用訓読み かく</p> <p>画数 10</p> <p>JIS水準 1</p> <p>JIS水準(JIS2004) 1</p> <p>文字コード Unicode 66F8 Unicode(JIS2004) 66F8 シフト-JIS 8F91 シフト-JIS(JIS2004) 8F91</p> <p>文字区分 子の名に使える文字、常用漢字</p>
-------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

戸籍統一文字 156680 (=JMJ-013491≒M13928), 156730 (=JMJ-013492≒M14294)

Similar distinction is found in CNS 11643, 6-3836 (for “noon”) versus 1-5573 (for “script”). The distinction in ISO/IEC 10646 is arguable; U+2F8CC (for “noon”) and U+66F8 (for “script”) are existing, but considering that U+2F8CC is in CJK Compatibility Ideograph Supplement, it can be classified as unified.



<p>全字庫字型</p> <p>書 書 書 明體 楷體 宋體</p> <p>拼音 shū shū zhōu huà</p> <p>筆畫 部首 倉頡</p> <p>10 日 LGA</p> <p>部件 書, 日 部件版本(CNS11643_2_94)</p> <p>筆順序 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 筆順序版本(CNS11643_1_94)</p> <p>漢字國際性</p>	<p>全字庫字型</p> <p>書 書 書 明體 楷體 宋體</p> <p>拼音 shū shū shū shū</p> <p>筆畫 部首 簡體數字 倉頡</p> <p>10 日 書 LGA</p> <p>部件 書, 日 部件版本(CNS11643_2_94)</p> <p>筆順序 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 筆順序版本(CNS11643_1_94)</p> <p>漢字國際性 台灣(T1-5573) 中國(G1-4A69) 日本(J0-3D71) 韓國(K0-5F76) 越南(V1-5878) 北韓(KP0-E1ED)</p>
<p>25 3835 38</p> <p>上 書 夏</p> <p>26 3836 38</p> <p>上 書 夏</p>	<p>2 5572 56</p> <p>上 書 夏</p> <p>3 5573 56</p> <p>上 書 夏</p>

CNS 11643 6-3836 and 1-5573 (www.cns11643.gov.tw (upper), CNS 11643:2007 (lower))

66F8 日 73.6	書 書 書 書 書 書	2F8CC 日 73.6	書
G1-4A69	HB1-AED1 T1-5573 J0-3D71 K0-5F76 V1-5878	T6-3836	

ISO/IEC 10646:2012 U+66F8 and U+2F8CC

Also 戸籍統一文字 has 2 codes; 161600 (classified in “moon” radical) and 329830 (classified in “meat” radical).

<p>戸籍統一文字番号 161600</p>  <p>親字・正字</p> <p>文字区分</p>	<p>音読み セン</p> <p>常用音読み</p> <p>訓読み</p> <p>常用訓読み</p> <p>画数 11</p> <p>JIS水準 4</p> <p>JIS水準(JIS2004) 4</p> <p>文字コード Unicode 8127 Unicode(JIS2004) 8127 シフトJIS シフトJIS(JIS2004) F863</p>	<p>戸籍統一文字番号 329830</p>  <p>親字・正字</p> <p>文字区分</p>	<p>音読み サイ、セ、サ、スイ</p> <p>常用音読み</p> <p>訓読み</p> <p>常用訓読み</p> <p>画数 11</p> <p>JIS水準</p> <p>JIS水準(JIS2004)</p> <p>文字コード Unicode Unicode(JIS2004) シフトJIS シフトJIS(JIS2004)</p>
-----------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

戸籍統一文字 161600 (=JMJ-021069≒M14365), 329830 (=JMJ-021070≒M29531)

In addition, CNS 11643 has 2 codes; 2-3571 (classified in “moon” radical) and 2-3853 (classified in “meat” radical).

<p>全字庫字型</p> <p>脛 脛 脛</p> <p>明體 楷體 宋體</p> <p>拼音 juān</p> <p>漢語 juān</p> <p>注音第二式 juan</p> <p>耶魯 jywan</p> <p>韋式 chwan</p> <p>筆畫 部首 簡體數字 倉頡</p> <p>11 月 脛脛 BICE</p> <p>部件 月, 月, 月, 月 部件版本(CNS11643_2_94)</p> <p>筆順序 ㇇, ㇈, ㇉, ㇊, ㇋, ㇌, ㇍, ㇎, ㇏ 筆順序版本(CNS11643_1_94)</p> <p>漢字區屬性 台灣(T2-3571) 中國(G3-4D2E) 韓國(K2-394D) 北韓(KP1-4A4D)</p>	<p>全字庫字型</p> <p>脛 脛 脛</p> <p>明體 楷體 宋體</p> <p>拼音 juān zūf</p> <p>漢語 juān zūf</p> <p>注音第二式 juan tzuei</p> <p>耶魯 jywan dzwei</p> <p>韋式 chwan tsui</p> <p>筆畫 部首 簡體數字 倉頡</p> <p>11 肉 脛 BICE</p> <p>部件 月, 月, 月, 月 部件版本(CNS11643_2_94)</p> <p>筆順序 ㇇, ㇈, ㇉, ㇊, ㇋, ㇌, ㇍, ㇎, ㇏ 筆順序版本(CNS11643_1_94)</p> <p>漢字區屬性 台灣(T2-3853) 中國(GE-3961) 日本(J1-5631) 韓國(K1-703A) 北韓(KP1-6B1F)</p>
<p>3570</p> <p>脛 脛 脛</p> <p>3571 3ㄝ</p> <p>脛 脛 脛</p>	<p>3852 3</p> <p>脛 脛 脛</p> <p>3853 3</p> <p>脛 脛 脛</p>

CNS 11643 2-3571 and 2-3853 (www.cns11643.gov.tw (upper), CNS 11643:2007 (lower))

Considering CNS 11643 distinction, ISO/IEC 10646 has 2 codes; U+6718 (classified in “moon” radical) and U+8127 (classified in “meat” radical).

6718 月 74.7	脛 脛 脛	脛
	G3-4D2E HB2-D5AF T2-3571	K2-394D
8127 肉 130.7	脛 脛 脛 脛 脛	
	GE-3961 HB2-D74F T2-3853 J1-5631 K1-703A	

ISO/IEC 10646:2012 U+6718 and U+8127

In summary, even if it is hard to distinguish from their shapes, the experts of character encoding want to encode the pair separately, as far as there is a sematic difference of “moon” and “meat” radical. However, it is difficult to make a good documentation to rationalize such preference.

2014/04/13: Daikanwa M1337 “兀”, M7542 “兀”, M7538 “尢” and M7539 “尢” are different characters?

Following my previous entries, more questions would be raised; Morohashi No. 1337 “兀”, No. 7542 “兀”, No. 7538 “尢” and No. 7539 “尢” are different characters? The pronunciations for the characters are described as follows; M1337 = gotsu, M7542 = M7538 = ou, M7539 = yuu. It seems that there is some confusion among these 同形異字.

<p>兀 1337 ゴツ (集韻) 五忽切 囹 ゴチ 囹コツ X wu</p> <p>小 一 たかい。高くして上が平かな 一 在兀上、讀若レ復、茂陵有兀桑里。(段注) 一 在兀上、高而平之意也。凡从兀聲之字、多取孤高之意。● 高くそびえた石。硯(8-24039)に通ず。(正字通) 兀、崖高曰硯、古通用兀。(柳宗元、晉問) 邊突硯兀。(注) 兀、危石也。● はげる。山のはげたさま。岬(4-7897)に通ず。(正字通) 兀、山秃曰岬、古通用兀。(杜牧、阿房宮賦) 蜀山兀阿房出。● あし wu。踞(10-37412)に通ず。(說文通訓定聲) 兀、假借爲踞。(莊子、德充符) 魯有兀者叔山、無趾踵。(釋文) 別足曰兀。● 無知のさま。(孫綽、遊天台山賦) 兀同體於自然。(注) 善曰、兀、無知貌。● 動かないさま。一心に勉強する意。(韓愈、進學解) 常兀兀以窮年。● おちつかないさま。岬(8-2869)・机(8-14457)に通ず。(正字通) 兀、凡不安謂之兀、與岬・机通。● 發語の辭。戲曲などに用ひる。兀的(8)を見よ。● 姓。(集韻) 兀、亦姓。(正字通) 兀、又姓、後魏改樂安王元覽爲兀氏。● 注音符號の聲符の一。舌後帶鼻音。後舌部を上升して軟口蓋と接觸し、氣息の洩出を阻止し、氣息を鼻腔から外洩させると此の音となる。ゴ。m。此の聲符は今用ひず、外國音又は方言を</p>	<p>兀 7542 ワウ 韻) 兀、俗作レ兀。(五音集</p>	<p>尢 7538 ワウ (集韻) 烏光切 囹 X 尢 wu</p> <p>小 姓 古說 篆段 ● すねのまがつた 文 文 文 本 人 尢 (4-7538) に 同 じ。(說 文) 尢、跛 也、曲 脛 人 也、从 大 象 偏 曲 之 形、凡 尢 之 屬 从 尢 尢 尢 篆 文 从 尢。● せのくぐまつて短小の人。(說文 尢、段注) 鄭注、釋爲面鄉天、或云、短小 曰尢、本从圭聲、省作尢。(玉篇) 尢、又 尢也、短小也、俗作尢。● 篆文は尢(4- 7538)に作る。(說文) 尢、尢、篆文从尢。● 注音符號の韻符の一。附聲韻母。單母音 Y を發すると同時に、舌後部を上升せしめ て、舌後帶鼻音の兀を以て收聲すれば、此 の音となる。アン。an</p>	<p>尢 7539 イウ 部。 尢 (4-7538) の本字。(正字 通) 尢、尢本字、說文尢、在乙 部。 參考) 尢(4-7538)は別字。(康熙字典、辨似、二 字相似) 尢、音注、跛曲脛也。尢、尢本字。</p>
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------

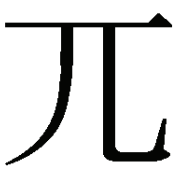
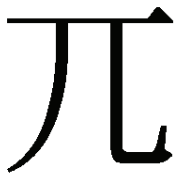




Morohashi M1337, M7542, M7538 and M7539

In JIS X 0208:1997, 49-ku-26-ten (J0-513A) “兀” is annotated with M1337 and M7542, and 53-ku-87-ten (J0-5577) “尢” is annotated with M7538 and M7539. In the other word, it is impossible to distinguish M1337 from M7542, either M7538 from M7539.

49-26	513a 9959	10 兀 1	S445	M1337	ゴツ, ゴツ	地 白兀山(ハツコツ
兀	d1ba 5140	43 尢 0		M7542	サン・福島県)	
53-87	5577 9b97	43 尢 0	S1908	[M7538]66		
尢	d5f7 5C22	43 尢 0		M7539	オウ, (まげあし)	詳説

J0-513A and J0-5577 Description (JIS X 0208:1997, Annex 6, p.161 and p.175)

On the other hand, 戸籍統一文字 has 4 codes; 015460 and 085690 for “兀”, 085670 and 085680 for “尢”.

<p>戶籍統一文字番号 015460</p>  <p>親字・正字</p>	<p>音読み ゴチ、コウ、ゴツ</p> <p>常用音読み</p> <p>訓読み</p> <p>常用訓読み</p> <p>画数 3</p> <p>JIS水準 2</p> <p>JIS水準(JIS2004) 2</p> <p>文字コード Unicode 5140 Unicode(JIS2004) 5140 シフトJIS 9859 シフトJIS(JIS2004) 9859</p> <p>文字区分</p>	<p>戶籍統一文字番号 085690</p>  <p>親字・正字</p> 	<p>音読み オウ</p> <p>常用音読み</p> <p>訓読み</p> <p>常用訓読み</p> <p>画数 3</p> <p>JIS水準</p> <p>JIS水準(JIS2004)</p> <p>文字コード Unicode Unicode(JIS2004) シフトJIS シフトJIS(JIS2004)</p> <p>文字区分</p>
<p>戶籍統一文字番号 085670</p>  <p>親字・正字</p>	<p>音読み オウ</p> <p>常用音読み</p> <p>訓読み</p> <p>常用訓読み</p> <p>画数 3</p> <p>JIS水準</p> <p>JIS水準(JIS2004)</p> <p>文字コード Unicode Unicode(JIS2004) シフトJIS シフトJIS(JIS2004)</p> <p>文字区分</p>	<p>戶籍統一文字番号 085680</p>  <p>親字・正字</p> 	<p>音読み ユウ</p> <p>常用音読み</p> <p>訓読み もっとも</p> <p>常用訓読み</p> <p>画数 3</p> <p>JIS水準 2</p> <p>JIS水準(JIS2004) 2</p> <p>文字コード Unicode 5C22 Unicode(JIS2004) 5C22 シフトJIS 9E87 シフトJIS(JIS2004) 9E87</p> <p>文字区分</p>

戶籍統一文字 015460 (=JMJ-007207≒M1337), 085690 (=JMJ-007208≒M7542),
085670 (≒M7538), 085680 (≒M7539)

CNS 11643 has 3 codes (so looking complex situation); 1-4442 for “兀”, but 1-4452 and 4-2134 for “尢”. It is supposed to be the design of Big5.

<p>全字庫字型</p> <p>兀 兀 兀 明體 指體 宋體</p>	<p>發音</p> 				
拼音	漢語 wu				
字義	筆畫	部首	簡繁體字	倉頡	
相關詞	3	儿	兀	MU	
部件	一, 儿, 部件版本(CNS11643_2_94)				
筆順序	一, 丿, 丨, 筆順序版本(CNS11643_1_94)				
漢字國屬性	台灣(T1-4442) 中國(G0-5823) 日本(J0-513A) 韓國(K0-6834) 越南(V1-4C36) 北韓(KP0-F9F2)				

<p>全字庫字型</p> <p>尢 尢 尢 明體 指體 宋體</p>	<p>發音</p> 				
拼音	漢語 wáng	注音第二式 wáng	耶魯 wáng	韋式 wáng	
字義	筆畫	部首	簡繁體字	倉頡	
相關詞	3	尢	尢	KU	
部件	尢, 部件版本(CNS11643_2_94)				
筆順序	一, 丿, 丨, 筆順序版本(CNS11643_1_94)				
漢字國屬性	台灣(T1-4452) 中國(G0-5E4C) 日本(J0-5577) 韓國(K2-2E2F) 北韓(KP1-3ED5)				

<p>全字庫字型</p> <p>尢 尢 尢 明體 指體 宋體</p>	<p>發音</p> 				
拼音	漢語 wáng yóu	注音第二式 wáng	耶魯 wáng	韋式 wáng	
字義	筆畫	部首	倉頡		
相關詞	3	尢	KU		
部件	尢, 部件版本(CNS11643_2_94)				
筆順序	一, 丿, 丨, 筆順序版本(CNS11643_1_94)				
漢字國屬性					

CNS 11643 1-4442, 1-4452 and 4-2134 (www.cns11643.gov.tw)

Considering that Big5 C940 – F9D5 area was for “次常用國字標準字體表”, Big5 C94A should be considered as identical with M7542, the variant of 尢.

00760	00759	00758	00757	00756	00755	00754	00753	00752	00751	號字
兀	尠	尠	尠	尠	尠	尠	尠	尠	尠	字標 體準
尢	小	小	小	寸	寸	宀	宀	宀	宀	部首
0	10	2	1	9	7	14	13	11	10	筆外部 畫之首
3	13	5	4	12	10	17	16	14	13	畫筆總
										說
										明

A Variant of “尢” in 次常用國字標準字體表 (1982/12/20 edition, p.38)

In ISO/IEC 10646, M1337 and M7542 are coded as U+5140 and U+FA0C, M7538 and M7539 are coded as U+5C22 and U+2F875.

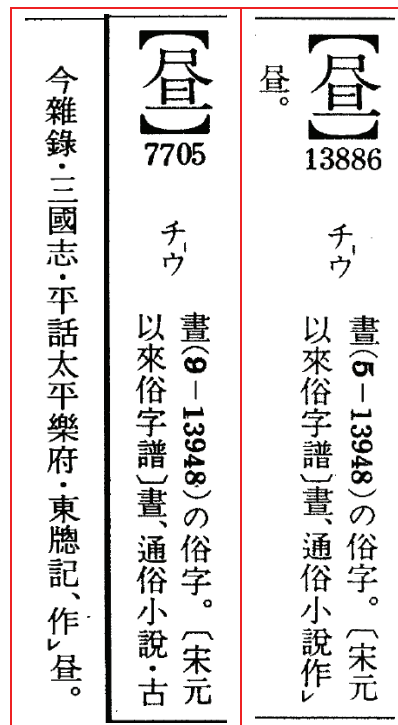
5140 兀 10.1	兀	兀	兀	兀	兀	兀	FA0C 兀 10.1	兀	兀
	G0-5823	HB1-A461	T1-4442	J0-513A	K0-6834	V1-4C36		HB2-C94A	UTC-00915
5C22 尢 43.0	尢	尢	尢	尢	尢		2F875 尢 43.0	尢	
	G0-5E4C	HB1-A471	T1-4452	J0-5577	K2-2E2F			T4-2134	

ISO/IEC 10646:2012 U+5140, U+FA0C, U+5C22 and U+2F875

The situation is similar to my first entry about “書”; encode the pair with the glyphic difference is separately coded, and add compatibility characters for the pair without the glyphic difference. However, we should be careful that the substitution of U+FA0C is standardized as <U+5140 U+FE00>, that of U+2F875 is standardized as <U+5C22 U+FE00> (if VS is ignored, U+FA0C and U+2F875 are replaced by the similar shape but non-cognate characters).

2014/04/15: Daikanwa M7705 “昼” and M13886 “昼” are different characters?

Continued from my recent entries; Morohashi No. 7705 “昼” and No. 13886 “昼” are different characters?



Morohashi M7705 and M13886

M7705 is classified in the radical corpse (尸), and M13886 is classified in the radical sun (日). But no semantic difference is recorded, both are the variant of “晝”. It is questionable whether this pair could be treated as 同形異字, it could be an error in the editing the dictionary (same character is listed twice in different radical). In JIS X 0208:1997, 35-ku-75-ten “昼” (J0-436B) is annotated with both of M7705 and M13886.

35-75	436b 928b	44 尸 6	M7705			
昼	c3eb 663C	72 日 5	S3309 M13886	[常]	チュウ, ひる	58-76 晝

J0-436B Description in JIS X 0208:1997 (JIS X 0208:1997, Annex 6, p.122)

Thus it is impossible to distinguish them in the text coded by JIS X 0208.

On the other hand, 戸籍統一文字 has 2 codes; 087550 (M7705) and 156080 (for M13886).

戸籍統一文字番号	音読み	チュウ	戸籍統一文字番号	音読み	チュウ
087550	常用音読み	チュウ	156080	常用音読み	チュウ
	訓読み	ひる		訓読み	ひる
	常用訓読み			常用訓読み	ひる
	画数	9		画数	9
	JIS水準			JIS水準	1
	JIS水準(JIS2004)			JIS水準(JIS2004)	1
	文字コード	Unicode		文字コード	Unicode 663C
		Unicode(JIS2004)			Unicode(JIS2004) 663C
		シフト-JIS			シフト-JIS 928B
		シフト-JIS(JIS2004)			シフト-JIS(JIS2004) 928B
	文字区分			文字区分	子の名に使える文字、常用漢字

戸籍統一文字 087550 (≒M7705) and 156080 (≒M13886)

The shape of 156080 is slightly different from Morohashi shape, it is more similar to the shape in 常用漢字表 (087550 uses “日”, although M7705 uses “日”).

(Joyo Kanjihyo 1981)

 (Joyo Kanjihyo 2010)

Joyo-Kanji shape of “昼”

CNS 11643 has only 1 code, 3-2C37, which is classified in the radical corpse.

全字庫字型	發音			
拼音	漢語 zhòu			
字義	筆畫	部首	簡/繁體字	倉頡
相關詞	9	尸	晝	SOAM
部 件	尸, 丿, 日, 一, 部件版本 (CNS11643_2_94)			
筆順序	㇇, ㇇, ㇇, ㇇, ㇇, ㇇, ㇇, ㇇, ㇇, 筆順序版本 (CNS11643_1_94)			
漢字國量柱	台灣 (T3-2C37) 中國 (G0-5667) 日本 (J0-436B) 韓國 (K2-3864) 北韓 (KP1-4943)			

CNS 11643 3-2C37 (www.cns11643.gov.tw)

ISO/IEC 10646 has only 1 code, U+663C. Its radical is sun, but its source includes CNS 11643 3-2C37 (which is classified in the radical corpse in its source character set).

663C 日 72.5	晝	晝	晝	晝
	G0-5667	T3-2C37	J0-436B	K2-3864

ISO/IEC 10646:2012 U+663C

In summary, it should be understood as M7705 and M13886 are unified at U+663C. In my opinion, there is no reason to encode M7705 and M13886 separately. What do you think of?

2014/04/16: Daikanwa M247 “二”, M248 “二” are different characters?

Continued my recent entries; Morohashi No. 247 “二” and No. 248 “二” are different characters? Also, No. 249 “二” and No. 250 “二” are different characters?

<p>【二】 250</p> <p>同じ文字を重ねる記號。(容齋隨筆、元二之災)古書、字當三再讀一者、即ニ</p>	<p>【二】 249</p> <p>カ 下(1-14)の古文。(正字通)三、古文下。</p>	<p>【二】 248</p> <p>上(1-13)の古文。(正字通)三、古文上。</p>	<p>【二】 247</p> <p>ニ、ジ (集韻)而至切。眞</p> <p>ニ 小 兼 文 古 文 古</p> <p>●ふたつ。(易、繫辭上) (論語、公治長)賜也、聞レ一以知レ二。(荀子、議兵)權出ニ於一者強、權出ニ於二者弱。●ふたたび。ふたたびする。(齊書、禮志)醮酒ニ辭。(宋史、吳璘傳)一敗而二</p>
---------------------------------------------------------------	----------------------------------------------------	--------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------

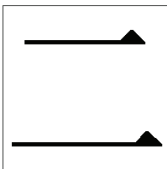
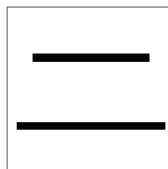
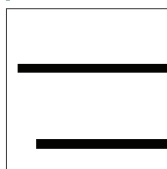

Morohashi M247, M248, M249 and M250

Roughly categorized, M247 is in the group meaning “two”, 二/弌, M 248 is in the group meaning “above”, 二/上/上, M249 is in the group meaning “below”, 二/下/下, M250 is in the group of the repetition marks, 二/々. It is reasonable to recognize all of them are non-cognate. However, the position of JIS X 0208:1997 for these characters is unclear. 38-ku-83-ten (J0-4673) is annotated with M248, but others (M247, M249, M250) are not related with any JIS X 0208 characters.

38-83	4673 93f1	7 二 0	S73	M247	[常] 二, ふた, ふたつ → 式 ⁴⁸⁻¹⁷
	c6f3 4E8C				

J0-4673 Description in JIS X 0208 (JIS X 0208:1997, Annex 6, p.131)

It reflects that JIS X 0208:1997 does not want to discourage the interchange of M248 via 38-ku-83-ten, but does not want to encourage either. On the other hand, 戸籍統一文字 has 4 codes; 002980 (for M247), 002990 (for M248), 003000 (for M249), 003010 (for M250).

<p>戸籍統一文字番号 002980</p>  <p>音読み ニ、ジ 常用音読み ニ 訓読み ふた、ふたたび、ふたつ 常用訓読み ふた、ふたつ 画数 2 JIS水準 1 JIS水準(JIS2004) 1 文字コード Unicode 4E8C Unicode(JIS2004) 4E8C シフトJIS 93F1 シフトJIS(JIS2004) 93F1</p> <p>親字・正字</p> <p>文字区分 子の名に使える文字、常用漢字</p>	<p>戸籍統一文字番号 002990</p>  <p>音読み ショウ 常用音読み ショウ 訓読み ショウ 常用訓読み ショウ 画数 2 JIS水準 2 JIS水準(JIS2004) 2 文字コード Unicode 4E00 Unicode(JIS2004) 4E00 シフトJIS 8140 シフトJIS(JIS2004) 8140</p> <p>親字・正字</p> <p>文字区分</p>
<p>戸籍統一文字番号 003000</p>  <p>音読み カ 常用音読み カ 訓読み カ 常用訓読み カ 画数 2 JIS水準 2 JIS水準(JIS2004) 2 文字コード Unicode 4E0B Unicode(JIS2004) 4E0B シフトJIS 814B シフトJIS(JIS2004) 814B</p> <p>親字・正字</p> <p>文字区分</p>	<p>戸籍統一文字番号 003010</p>  <p>音読み カ 常用音読み カ 訓読み カ 常用訓読み カ 画数 2 JIS水準 2 JIS水準(JIS2004) 2 文字コード Unicode 3007 Unicode(JIS2004) 3007 シフトJIS 8147 シフトJIS(JIS2004) 8147</p> <p>親字・正字</p> <p>文字区分</p>

戸籍統一文字 002980, 002990, 003000 and 003010

Also CNS 11643 has 4 codes; 1-4428, 6-2127, 6-2128, 6-2129 for these characters.

全字庫字型	發音			
拼音	漢語 èr			
字義	筆畫	部首	兩/零體字	倉頡
相關詞	2	二	二	MM
部 件	二_部件版本(CNS11643_2_94)			
筆順序	一,二_筆順序版本(CNS11643_1_94)			
漢字國屬性	台灣(T1-4428) 中國(G0-367E) 日本(J0-4673) 韓國(K0-6C23) 越南(V1-4A50) 北韓(KP0-FCA7)			

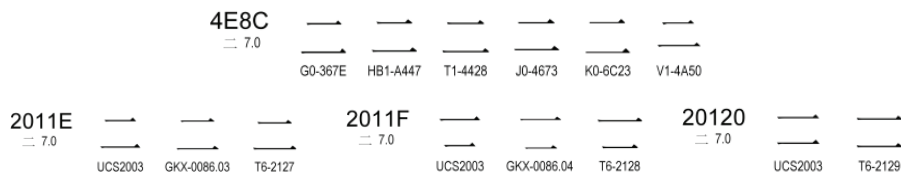
全字庫字型	發音			
拼音	漢語 shàng			
字義	筆畫	部首	倉頡	
相關詞	2	二	MM	
部 件	二_部件版本(CNS11643_2_94)			
筆順序	一,二_筆順序版本(CNS11643_1_94)			
漢字國屬性	台灣(T6-2127) 中國(G_KX)			

全字庫字型	發音			
拼音	漢語 xià			
字義	筆畫	部首	倉頡	
相關詞	2	二	MM	
部 件	一,二_部件版本(CNS11643_2_94)			
筆順序	一,二_筆順序版本(CNS11643_1_94)			
漢字國屬性	台灣(T6-2128) 中國(G_KX)			

全字庫字型	發音			
拼音	漢語 èr			
字義	筆畫	部首	倉頡	
相關詞	2	二	MM	
部 件	一,二_部件版本(CNS11643_2_94)			
筆順序	一,二_筆順序版本(CNS11643_1_94)			
漢字國屬性	台灣(T6-2129)			

CNS 11643 1-4428, 6-2127, 6-2128 and 6-2129 (www.cns11643.gov.tw)

Considering the separation in CNS 11643, ISO/IEC 10646 encodes them separately; U+4E8C, U+2011E, U+2011F, U+20120.

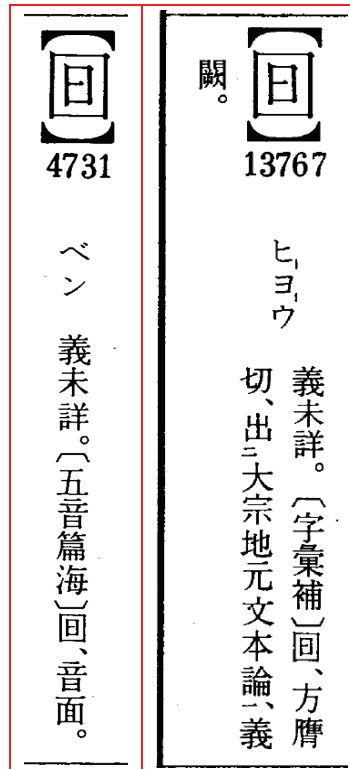


ISO/IEC 10646:2012 U+4E8C, U+2011E, U+2011F and U+20120

Therefore, the major position would be that M247 and M248 should be distinguished in the coded character set. The difference from the cases of M13928/M14294 “書”, M398-1/-2 “假” should be considered.

2014/04/17: Daikanwa M4731 “回”, M13767 “回” are different characters?

Continued my recent entries; Morohashi No. 4731 and No. 13767 are different characters?



Morohashi M4731, M13767


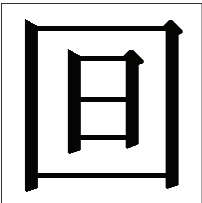
M4731 is noted as “pronounced as ben, meaning is unknown” and classified in the radical enclosure “口”, M13767 is noted as “pronounced as hyou, meaning is unknown” and classified in the radical sun “日”. Both characters lack the meanings, so it is impossible to discuss the group of the glyphic variants. In the methodology to determine the character identity by the comparison of the glyphic variant groups, it is difficult to call this pair as 同形異字.

Both of JIS X 0212 and JIS X 0213 do not include M4731 nor M13767. Either both characters are not found in CNS 11643 (the latest CNS 11643 includes it on the 10th plane, although source mapping table and code chart in ISO/IEC 10646 does not mention).

全字庫字型	發音			
回 回 回 明體 楷體 宋體	ㄅ ㄨㄟ		ㄅ ㄩㄠ	
	男聲 下載	女聲 下載	男聲 下載	女聲 下載
拼音	漢語 bing bian	注音符二式 bing	耶魯 bing	韋式 ping
字義	筆畫	部首	倉頡	
相關詞	7	日	WA	
部 件	口, 日, 部件版本(CNS11643_2_94)			
筆順序	1, 7, 1, 7, 1, 7, 筆順序版本(CNS11643_1_98)			
漢字國屬性	中國(G_KK)			

CNS 11643 10-4D47 (www.cns11643.gov.tw)

On the other hand, 戶籍統一文字 has 2 codes; 054020 for M4731, 154740 for M13767.

<p>戸籍統一文字番号 054020</p>  <p>親字・正字</p>	<p>音読み ベン</p> <p>常用音読み</p> <p>訓読み</p> <p>常用訓読み</p> <p>画数 7</p> <p>JIS水準</p> <p>JIS水準(JIS2004)</p> <p>文字コード</p> <p>Unicode</p> <p>Unicode(JIS2004)</p> <p>シフトJIS</p> <p>シフトJIS(JIS2004)</p> <p>文字区分</p>	<p>戸籍統一文字番号 154740</p>  <p>親字・正字</p>	<p>音読み ヒョウ</p> <p>常用音読み</p> <p>訓読み</p> <p>常用訓読み</p> <p>画数 7</p> <p>JIS水準</p> <p>JIS水準(JIS2004)</p> <p>文字コード</p> <p>Unicode</p> <p>Unicode(JIS2004)</p> <p>シフトJIS</p> <p>シフトJIS(JIS2004)</p> <p>文字区分</p>
-----------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

戸籍統一文字 054020 (=JM~~J~~-037624≐M4731), 154740 (=JM~~J~~-037625≐M13767)

ISO/IEC 10646 encodes “回” at U+2314D, which could be recognized as a unified character of M4731 and M13767 (note: in UAX#38, Unihan database, U+2314D has kHanyu and kIRGHanyuDaZidian properties mapping U+2314D to the 2nd character in Volume 1, p.715. The entry of HDZ unifies both of M4731 and M13767).

2314D
日 72.3

UCS2003 GKX-0490.20

ISO/IEC 10646:2012 U+2314D

To encode them separately, the discussions like the cases for M247/M248 (二/二), M14365/M29531 (腩/腩) should be revisited, but it is quite hard to discuss without the meaning of them. If it is agreed to encode M4731 and M13767 separately, the unification of M398-1/M398-2 (仮/仮) would be problem in future. In my personal opinion, the advantage to separate M4731/M13767 is too unclear in consideration with such future issues.

Appendix: Comments proposing non-cognate unification in Ext. F submissions

00236	伎 JMJ-000583	U+36A 2	36A2 女 38.2 G5-314C	伎 伎 T3-2174	Unificat ion			J, H, K	p
04904	腋 JMJ-003791	U+43E 2	43E2 肉 130.5 GKX-0978.15	腋 腋 T3-2E82	Unificat ion			J	p
03389	脛 JMJ-003833	U+440 B	440B 肉 130.8 GKX-0987.21	脛 脛 脛 T6-497E K3-3021	Unificat ion			J	p
00096	𠂇 JMJ-006371	U+4E4 1	4E41 𠂇 4.0 GK-0875	𠂇 𠂇 𠂇 𠂇 T4-2121 J5-3133 K2-2131	Unificat ion	G: Japan mentioned it in Attributes file. But China invites Japan for more explanation.		J, G, H, K	p
014 09	回 01409 JMJ-037624	U+2314D	2314D 回 72.3 UCS2003 GKX-060.20	回 回	Unificati on			K	Delay
01983	兀 JMJ-007208	U+5140	5140 𠂇 12.1 G6-5122 H6A-661 T1-442 J6-515A H6-681A V1-4236	兀 兀 兀 兀 兀 兀	Unificat ion	G: Although there is evidence, can not see difference by glyphs. Want more expiation.		J, G, H	p
02242	𠂇 JMJ-046074	U+269 E0	269E0 舌 135.8 UCS2003 GKX-1007.17	𠂇 𠂇 𠂇	Unificat ion			J	p
03062	𠂇 JMJ-012948	U+6535	6535 𠂇 90.0 G6-8895 H-88CE T3-218B J6-615E K2-3283	𠂇 𠂇 𠂇 𠂇 𠂇	Unificat ion	G: Want more explanation although there is evidence.		J, G	p
03187	𠂇 JMJ-013491	U+66F8	66F8 𠂇 73.0 G1-4489 H6A-661 T1-673 J6-5171 K6-5171 Y1-5878	𠂇 𠂇 𠂇 𠂇 𠂇 𠂇	Unificat ion			J, H	p
03368	𠂇 JMJ-003782	U+43D 9	43D9 肉 130.4 G3-4CS1	𠂇 𠂇 𠂇 T4-2785 H-8476	Unificat ion			J	p
03471	𠂇 JMJ-013841	U+681 D	681D 𠂇 130.0 G3-681B H62-018B T3-3F6D J1-451C K1-588E	𠂇 𠂇 𠂇 𠂇 𠂇	Unificat ion			J	p
05814	𠂇 JMJ-044844	U+2628 B	2628B 𠂇 122.4 UCS2003 GHZ-21170.04	𠂇 𠂇 𠂇 T5-2752	Unificat ion			J	p
05987	𠂇 JMJ-021070	U+8127	8127 𠂇 160.7 G6-3861 H62-074F T2-3863 J1-4511 K1-703A	𠂇 𠂇 𠂇 𠂇 𠂇	Unificat ion			J	p
06492	𠂇 JMJ-055359	U+29F3 B	29F3B 𠂇 195.15 UCS2003 GKX-1480.08	𠂇 𠂇 𠂇 T7-6335	Unificat ion			J	p
07345	𠂇 JMJ-058872	U+2912 E	2912E 𠂇 162.14 UCS2003 GHZ-64074.10	𠂇 𠂇 𠂇 T7-445C	Unificat ion			J	p
08317	𠂇 JMJ-050235	U+2811 9	28119 足 167.11 UCS2003 GKX-1232.42	𠂇 𠂇 𠂇 T4-5F33	Unificat ion			J	p

see IRG N1395

see IRG N1395

see Yasuoka note

see IRG N1395

see IRG N1395

see Yasuoka note

(end of document)