

Updates for Unihan kXHC1983 data

L2/22-184

Andrew West

5 August 2022

The Unihan kXHC1983 field presents the Mandarin Chinese readings given in the 2nd edition of *Xiàndài Hànyǔ Cídiǎn* 现代汉语词典 [*Dictionary of Contemporary Chinese*] (Běijīng: Commercial Press, 1983). When the kXHC1983 data was first incorporated into the Unihan database, a large number of characters included in the dictionary were not yet encoded (mostly derived simplified characters), but since the encoding of CJK Unified Ideographs Extensions C through G, all characters in the dictionary which are not unifiable variants have been encoded. The kXHC1983 data has already been updated with a few Ext. C and Ext. E characters (24 Ext. C and 4 Ext. E), but several hundred entries still need to be updated for characters encoded in Exts. C through G. There are also a number of mistakes in references and readings given in the current (i.e. Unicode 15.0 beta) kXHC1983 data. This document documents the changes and corrections required to bring the kXHC1983 data up to date. The new kXHC1983 data covers 11,072 CJK unified ideographs (URO 10,467; Ext. A 218; Ext. B 130; Ext. C 82; Ext. D 2; Ext. E 95; Ext. F 61; Ext. G 17).

Although the 2nd edition of *Xiàndài Hànyǔ Cídiǎn* was published almost forty years ago, the phonetic data it provides is still valid, other than for a handful of characters for which the official standard reading has been changed. Improving and correcting the kXHC1983 data is important, as the new data should provide a good foundation for creating a new Unihan data set for the 6th (2012) or 7th (2016) editions of *Xiàndài Hànyǔ Cídiǎn* (I have a copy of the 6th edition, and would be willing to collaborate on a project to create a kXHC2012 data set).

Plain text listings of all kXHC1983 entries to be added (249 entries), removed (195 entries), and changed (142 entries) are attached to the pdf. 10,681 entries are unchanged.

1. Spurious entries

The following entry appears to be an error:

U+2A84F 噍 1450.130:zhān

The character 噍 lián is not read zhān, and does not occur in the dictionary. Furthermore, the reference 1450.130 is already assigned to another character (U+9958 饘 zhān), so the entry for U+2A84F should be removed.

In fact, all the Unihan Readings for U+2A84F 噍 appear to refer to U+9958 饘 / U+2B5F4 饘 zhān/zin1 'gruel', and should be replaced with entries corresponding to U+55F9 噍.

U+2A84F	kCantonese	zin1
U+2A84F	kDefinition	gruel
U+2A84F	kMandarin	zhān
U+2A84F	kXHC1983	1450.130:zhān

2. Spurious references

The following five entries include spurious references which do not occur in the dictionary.

Character	Current Entry	New Entry
U+71C8 燈	0225.061,0228.061:dēng	0225.061:dēng
U+807D 聽	1145.031:tīng 1145.032:tìng	1145.031:tīng
U+88B7 裕	0541.133:jiá 0541.134:jié 0906.070:qiā	0541.133:jiá 0906.070:qiā
U+946A 鑪	0736.012,0736.013*:lú	0736.012:lú
U+9AEE 髮	0297.041:fà 0297.042:fǎ	0297.041:fà

3. Incorrect references

The following 26 entries have an error in the dictionary reference for one or more readings. Note that the references for U+8FB9 边 and U+908A 邊 bian (light tone) are currently the same as for 边/邊 biān, but all other light tone characters are given a separate reference (e.g. U+5934 头 1156.030:tóu 1157.010:tou), so the references for 边/邊 bian should be changed, which affects the references for the three following characters on the same page (U+7B3E 筵, U+7C69 籩, U+6241 扁).

Character	Current Entry	New Entry
U+4E2A 个	0373.090:gè 0373.120:gě	0373.090:gě 0373.120:gè
U+507A 偌	1435.072:zá 1439.042:zán 1440.072:zan	1435.072:zá 1439.042:zán 1440.042:zan
U+54B1 咱	1435.070:zá 1439.040:zán 1440.070:zan	1435.070:zá 1439.040:zán 1440.040:zan
U+5592 嚼	1435.071:zá 1439.041:zán 1440.071:zan	1435.071:zá 1439.041:zán 1440.041:zan
U+6241 扁	0064.020:biǎn 0871.020:piān	0064.030:biǎn 0871.020:piān
U+7242 牂	1440.100:zāng	1440.070:zāng
U+74D2 瓚	1440.050:zàn	1440.020:zàn
U+77C7 矇	0778.020:méng 0778.021:mēng	0777.041:mēng 0778.020:méng
U+7B3E 筵	0064.010:biān	0064.020:biān
U+7B74 筴	0110.011:cè 0110.012:jiá 0540.011:jiā	0110.011:cè 0540.011:jiā
U+7C69 籩	0064.011:biān	0064.021:biān
U+807C 聽	1145.033:tīng	1145.032:tīng

Character	Current Entry	New Entry
U+810F 脏	1440.040:zàng 1440.090:zāng	1440.060:zāng 1440.120:zàng
U+81DF 臟	1440.041:zàng	1440.121:zàng
U+81E7 臧	1440.110:zāng	1440.080:zāng
U+846C 葬	1440.020:zàng	1440.100:zàng
U+85CF 藏	0105.060:cáng 1440.030:zàng	0105.060:cáng 1440.110:zàng
U+8863 衣	1350.040:yì 1368.050:yī	1350.040:yī 1368.050:yì
U+8D13 贓	1440.082:zāng	1440.052:zāng
U+8D1C 贓	1440.081:zāng	1440.051:zāng
U+8D43 赃	1440.080:zāng	1440.050:zāng
U+8FB9 边	0063.030:biān 0063.030:bian	0063.030:biān 0064.010:bian
U+908A 邊	0063.031:biān	0063.031:biān 0064.011:bian
U+9142 鄮	0186.020:cuó 1440.060:zàn	0186.020:cuó 1440.030:zàn
U+9A75 𨵿	1440.120:zǎng	1440.090:zǎng
U+9AD2 𨵿	1440.091:zāng	1440.061:zāng

4. Obsolete starred references

The following 12 entries include obsolete starred references which should be removed as the references already have their own entry in the current kXHC1983 data (mostly for Ext. C and Ext. E additions).

Character	Current Entry	New Entry
U+39D1 𠵹	0500.030,0500.031*:huī	0500.030:huī
U+571E 𧯞	0744.030*,0744.031:luán	0744.031:luán
U+5C69 𧯞	0618.100*,0618.101:juē	0618.101:juē
U+65DF 旗	1407.060*,1407.061:yú	1407.061:yú
U+747D 璫	0178.010*,0178.011:cōng	0178.011:cōng
U+8E7B 躑	0618.102*:juē 0920.100*,0920.101:qiāo	0920.101:qiāo
U+93E6 縱	0178.040*,0178.041:cōng	0178.041:cōng
U+9512 銀	0674.080*,0675.030:láng	0675.030:láng
U+9926 銀	1455.050*,1455.051:zhāng	1455.051:zhāng
U+9DB5 鷓	0160.020*,0160.021:chú	0160.021:chú
U+9F9D 𧯞	0939.110*,0939.111:qiū	0939.111:qiū
U+283AA 輶	0286.030*,0286.031:ér	0286.031:ér

5. Unnecessary starred references

The following entry has a reference marked with a star, but this is not required as the glyph form of the character in the dictionary matches the G-source form of the encoded character.

Character	Current Entry	New Entry
U+8199 疆	0561.030*:jiǎng	0561.030*:jiǎng

6. Archaic cross-reference readings

The following 6 characters include readings for archaic cognate characters which are given as cross-references in the dictionary (e.g. for U+9673 陳:<古>又同‘阵’zhèn), but not explicitly given as readings for the character in question. It seems to me that it would be best not to include the readings for the cross-referenced characters as they are not modern readings for the character, and this source is specifically a dictionary of modern Mandarin Chinese. Therefore I recommend removing the readings for archaic cognates.

Character	Current Entry	New Entry	Note in Dictionary
U+7E23 懸	1250.071:xiàn 1250.072: xuán	1250.071:xiàn	<古>又同‘悬’xuán
U+8006 嗜	0898.100:qí 0898.101:shì	0898.100:qí	<古>又同‘嗜’shì
U+83EF 華	0483.031:huá 0483.032:huā 0487.011:huà	0483.031:huá 0487.011:huà	<古>又同‘花’huā
U+8C48 豈	0901.031:qǐ 0901.032:kǎi	0901.031:qǐ	<古>又同‘恺’、‘凯’ [kǎi]
U+9673 陳	0133.021:chén 0133.022:zhèn	0133.021:chén	<古>又同‘阵’zhèn
U+9F4A 齋	0531.121:jì 0894.111:qí 0894.112:zhāi	0531.121:jì 0894.111:qí	<古>又同‘斋’zhāi

7. Missing readings

The following entry has a missing reading (this character is also listed under Section 3).

Character	Current Entry	Corrected Entry
U+908A 邊	0063.031:biān	0063.031:biān 0064.011:bian

8. Incorrect readings

The following 19 entries have one or more readings that do not match the readings given in the dictionary.

Character	Current Entry	Corrected Entry
U+5475 呵	0002.011,0002.061:ā 0002.051:á 0002.071:à 0002.082:a 0448.050,0448.061:hē 0638.100:kē	0002.011:ā 0002.051:á 0002.061:ǎ 0002.071:à 0002.082:a 0448.050,0448.061:hē 0638.100:kē
U+77FD 矝	1232.050:xì	1232.050:xī

Character	Current Entry	Corrected Entry
U+7809 砉	0478.020:huò 1299.070:xū	0478.020:huā 1299.070:xū
U+781D 砉	0297.020:fá	0297.020:fǎ
U+78A1 礪	1505.040:dú	1505.040:zhóu
U+7B69 箊	1155.011:tóng	1155.011:tǒng
U+7BC9 築	1515.051:zhú	1515.051:zhù
U+8052 聒	0422.060:guā	0422.060:guō
U+828E 芎	1293.020:qiōng	1293.020:xiōng
U+83F9 蒞	1544.020:jū	1544.020:zū
U+84C7 膏	0394.030:gǔ	0394.030:gū
U+884C 行	0442.080:háng 0443.050:hàng 0460.010:xìng 1290.030:xíng	0442.080:háng 0443.050:hàng 0460.010:héng 1290.030:xíng
U+8B26 馨	0936.030:qìng	0936.030:qǐng
U+8BE4 诤	1478.040:zhēng	1478.040:zhèng
U+8C49 豉	0144.020:shì	0144.020:chǐ
U+8E31 躅	0280.070:duò	0280.070:duó
U+8FD5 迕	1223.030:wù	1223.030:wǔ
U+9062 邇	1106.060:tà	1106.060:tā
U+9862 顛	0761.051:mán	0761.051:mān

9. Incorrect character forms

The characters for the following 9 entries do not match the character form given in the dictionary, and so the entries should be remapped to the characters shown below. There may be other examples which I have overlooked.

Current Character	Current Entry	New Character	New Entry
U+5002 併	0079.011:bìng	U+4F75 併	0079.011:bìng
U+6329 捩	1170.110:tuō	U+635D 捩	1170.110:tuō
U+7501 甌	0878.080:píng	U+74F6 甌	0878.080:píng
U+75F2 痲	0755.010:má	U+75F3 痲	0755.010:má
U+7A50 穉	0939.110*:qiū	U+2C4EB 穉	0939.110:qiū
U+24C19 旭	0639.020*:kē	U+2AF47 旭	0639.020:kē
U+20A44 所	1104.031:suǒ	U+2B826 所	1104.031:suǒ
U+21323 垞	1444.010*:zào	U+2A8DD 垞	1444.010:zào
U+2310A 旆	1407.060*:yú	U+2D930 旆	1407.060:yú

10. Traditional character forms which should be simplified

The following 166 entries are mapped to traditional character forms where the dictionary shows a simplified character form only. The simplified characters were almost all encoded in Exts. C through G.

Current Character	Current Entry	New Character	New Entry
U+42CE 組	1454.020:zhàn	U+2C61C 组	1454.020:zhàn
U+4308 縵	1106.090:tā	U+26216 縵	1106.090:tā

Current Character	Current Entry	New Character	New Entry
U+4325 纒	1382.040:yǐn	U+26220 纒	1382.040:yǐn
U+493C 錄	1251.040:xiàn	U+2CB63 錄	1251.040:xiàn
U+4955 繖	0876.030*:piě	U+2CB6F 繖	0876.030:piě
U+495B 鏢	0576.020:jiē	U+2CB74 鏢	0576.020:jiē
U+4ACF 顛	0892.020*:qī	U+2CC66 顛	0892.020:qī
U+4B54 鎚	0273.020*:duī	U+2B5F0 鎚	0273.020:duī
U+4C57 鰲	0102.030*:cān	U+2EB1E 鰲	0102.030:cān
U+4C5A 鰐	0740.080*:lù	U+2EB20 鰐	0740.080:lù
U+4C5B 鰻	0515.110*:huò	U+2EB1F 鰻	0515.110:huò
U+4C68 鰱	0669.060*:là	U+2CD9F 鰱	0669.060:là
U+4C75 鰻	1208.080*:wēng	U+2EB22 鰻	1208.080:wēng
U+4C7B 鰮	0483.010*:huá	U+2EB21 鰮	0483.010:huá
U+4C97 鰻	1248.040*:xián	U+2EB23 鰻	1248.040:xián
U+4CCD 鵠	0382.050*:gōng	U+2EB65 鵠	0382.050:gōng
U+4D95 齧	0019.040*:bà	U+2B728 齧	0019.040:bà
U+4D97 齧	0642.021*:kè 0906.100*:qiā	U+2EBD9 齧	0642.021:kè 0906.100:qiā
U+5297 勳	0553.110*:jiǎn	U+2D11B 勳	0553.110:jiǎn
U+52E3 勳	0518.050*:jī	U+2A7DD 勳	0518.050:jī
U+5513 嗶	0130.020*:chē	U+2A833 嗶	0130.020:chē
U+56AA 囁	0212.060*:dàn	U+2BAE6 囁	0212.060:dàn
U+6527 擷	0240.040:diān	U+2D8C7 擷	0240.040:diān
U+665B 睨	1251.020*:xiàn	U+2C02A 睨	1251.020:xiàn
U+6896 榭	0048.040:bèi	U+2AC77 榭	0048.040:bèi
U+6A9F 檣	0542.080:jiǎ	U+69DA 檣	0542.080:jiǎ
U+6BA8 殞	0504.170:huì	U+3C6E 殞	0504.170:huì
U+6F92 瀕	0466.110*:hòng	U+2DC4A 瀕	0466.110:hòng
U+7052 瀆	1440.010:zàn	U+2ADFD 瀆	1440.010:zàn
U+7213 爛	1333.030*:yàn	U+307D8 爛	1333.030:yàn
U+76E8 盪	1302.080:xǔ	U+2AF94 盪	1302.080:xǔ
U+77A4 潤	0977.120:rún	U+251A7 潤	0977.120:rún
U+77B7 潤	0554.030:jiàn	U+2C457 潤	0554.030:jiàn
U+7BE2 簣	0733.050:lǒng	U+2C542 簣	0733.050:lǒng
U+7C00 簣	1444.070:zé	U+7BA6 簣	1444.070:zé
U+7C63 籐	0672.050:lán	U+2E18F 籐	0672.050:lán
U+7D03 紉	1313.110*:xún	U+2C613 紉	1313.110:xún
U+7D1F 紉	0585.080*:jīn	U+2B11B 紉	0585.080:jīn
U+7D29 紉	1493.050*:zhì	U+2E262 紉	1493.050:zhì
U+7D3E 紉	1469.100*:zhěn	U+2C61D 紉	1469.100:zhěn
U+7D41 純	1035.030*:shī	U+2B11F 純	1035.030:shī
U+7D53 絳	0404.030*:guà	U+2E264 絳	0404.030:guà
U+7D56 統	0661.090*:kuàng	U+2C622 統	0661.090:kuàng
U+7D6A 網	1376.020*:yīn	U+2C621 網	1376.020:yīn
U+7D78 覘	0553.040*:jiǎn	U+2C616 覘	0553.040:jiǎn
U+7D9D 緝	0131.040*:chēn 0717.020*:lín	U+2C62D 緝	0131.040:chēn 0717.020:lín
U+7DAA 績	0915.090*:qiàn	U+2C62C 績	0915.090:qiàn

Current Character	Current Entry	New Character	New Entry
U+7DB7 粹	0183.130*:cuì	U+2E26C 粹	0183.130:cuì
U+7DC5 緞	1542.040*:zōu	U+2E26A 緞	1542.040:zōu
U+7DCC 綉	0977.020*:ruì	U+2E26B 綉	0977.020:ruì
U+7E06 緹	0378.010*:gēng	U+2C635 緹	0378.010:gēng
U+7E29 縗	0102.010*:cài	U+2E26F 縗	0102.010:cài
U+7E2F 縑	1328.020*:yǎn	U+2C642 縑	1328.020:yǎn
U+7E59 縑	0299.010*:fān 0300.090*:fán	U+2C646 縑	0299.010:fān 0300.090:fán
U+7E81 纁	1313.090*:xūn	U+2B138 纁	1313.090:xūn
U+7E86 纁	0803.030*:mò	U+2C64A 纁	0803.030:mò
U+858B 蕢	0172.160:cí	U+30C71 蕢	0172.160:cí
U+8621 夔	1385.080:yīng	U+2E428 夔	1385.080:yīng
U+875C 蛭	0348.010:fù	U+2E505 蛭	0348.010:fù
U+87D8 蛭	1125.030:tè	U+2B2CC 蛭	1125.030:tè
U+896D 纈	1274.070:xié	U+2E5B1 纈	1274.070:xié
U+8997 覘	1090.130:sì	U+2C88A 覘	1090.130:sì
U+8A0F 訃	1299.080:xū	U+2C8D9 訃	1299.080:xū
U+8A38 詠	0453.050:hé	U+30D5D 詠	0453.050:hé
U+8AD4 淑	0161.090:chù	U+30D74 淑	0161.090:chù
U+8B0C 譌	0369.020*:gē	U+2C910 譌	0369.020:gē
U+8B0F 謏	1271.010:xiǎo	U+2B372 謏	1271.010:xiǎo
U+8B46 禧	1226.060:xī	U+2B37B 禧	1226.060:xī
U+8B5E 還	1305.040:yuān	U+2B37D 還	1305.040:yuān
U+8B99 灌	0489.010:huān	U+2C930 灌	0489.010:huān
U+8CA4 虺	1363.130*:yí 1369.110*:yì	U+30DE0 虺	1363.130:yí 1369.110:yì
U+8CE8 竇	0178.070:cóng	U+30DE5 竇	0178.070:cóng
U+8D09 贖	0212.080:dàn	U+2B3AB 贖	0212.080:dàn
U+8D11 顛	0359.090:gàn	U+2B3AC 顛	0359.090:gàn
U+8E54 𨾏	1439.110:zàn	U+2B3D0 𨾏	1439.110:zàn
U+8EF1 輻	0394.060:gū	U+2E774 輻	0394.060:gū
U+8F00 輻	0286.030*:ér	U+2E775 輻	0286.030:ér
U+8F0B 峯	1010.030*:shē	U+2AA36 峯	1010.030:shē
U+8F17 輓	0826.080:ní	U+2B410 輓	0826.080:ní
U+8F28 輶	0409.070:guǎn	U+2B411 輶	0409.070:guǎn
U+8F34 輶	0170.060:chūn	U+2E778 輶	0170.060:chūn
U+8F36 輶	1397.030:yóu	U+2CA0E 輶	1397.030:yóu
U+8F47 轆	0566.060*:jiāo	U+2B416 轆	0566.060:jiāo
U+8F55 輶	0371.020:gé	U+2E77A 輶	0371.020:gé
U+8F57 輶	0632.050:kǎn	U+2B418 輶	0632.050:kǎn
U+8F58 輶	0490.060:huán 0491.110:huàn	U+2E779 輶	0490.060:huán 0491.110:huàn
U+91ED 鈺	0359.150:gāng	U+2E8F2 鈺	0359.150:gāng
U+91FD 鈺	0870.100:pì	U+2CB32 鈺	0870.100:pì
U+9207 鈺	0334.040:fū	U+2B4E7 鈺	0334.040:fū
U+921A 鈺	0866.010:pī	U+2CB2B 鈺	0866.010:pī
U+921C 鈺	0466.040:hóng	U+2E8F3 鈺	0466.040:hóng

Current Character	Current Entry	New Character	New Entry
U+9249 鉉	1308.030: xuàn	U+94C9 铉	1308.030: xuàn
U+924F 鉏	0159.090: chú 0613.040: jǔ	U+2CB3A 铏	0159.090: chú 0613.040: jǔ
U+9265 鉃	1070.040: shù	U+2CB38 铋	1070.040: shù
U+9267 鉄	0806.040: mǔ	U+2CB41 铃	0806.040: mǔ
U+9268 鉅	1235.060: xǐ	U+9FED 铄	1235.060: xǐ
U+9277 鉆	0465.050: hóng	U+2B7F9 铅	0465.050: hóng
U+92B6 鉇	0941.020: qiú	U+28C47 铆	0941.020: qiú
U+92CB 鉈	0118.020: chán	U+2E8F4 铇	0118.020: chán
U+931E 鉉	0170.090: chún 0273.030: duì	U+2CB5A 铈	0170.090: chún 0273.030: duì
U+9323 鉊	1525.060: zhuì	U+2E8F5 铉	1525.060: zhuì
U+9324 鉋	0520.020: jī	U+2B4F9 铊	0520.020: jī
U+939D 鉌	0188.040: dā	U+28C4F 铋	0188.040: dā
U+93AF 鉍	0674.080*: láng	U+28C4D 铌	0674.080: láng
U+93C1 鉎	1105.020*: suǒ	U+2CB72 铍	1105.020: suǒ
U+9404 鉏	0497.010: huáng	U+28C51 铋	0497.010: huáng
U+943B 鉃	0614.120: jù	U+2E8F7 铋	0614.120: jù
U+959F 閔	0059.040*: bì	U+2E932 閔	0059.040: bì
U+95A4 閏	0372.010*: gé 0449.031*: hé	U+2CBA4 閏	0372.010: gé 0449.031: hé
U+95C9 閗	1374.030*: yīn	U+2CBB1 閗	1374.030: yīn
U+9767 𨾏	0504.160*: huì	U+2B583 𨾏	0504.160: huì
U+9814 頔	0231.040*: dí	U+2CC56 頔	0231.040: dí
U+981E 頔	0283.070*: è	U+310A8 頔	0283.070: è
U+9820 頔	1198.060*: wěi	U+2CC5F 頔	1198.060: wěi
U+982B 頔	0341.100*: fǔ	U+2B5AF 頔	0341.100: fǔ
U+9847 頔	0183.080*: cuì	U+310AF 頔	0183.080: cuì
U+9851 頔	0632.040: kǎn	U+310B1 頔	0632.040: kǎn
U+9866 顛	0921.120: qiáo	U+310BB 顛	0921.120: qiáo
U+98BF 颯	0298.030*: fān	U+2CCF3 颯	0298.030: fān
U+9902 𨾏	1138.080*: tiǎn	U+310FA 𨾏	1138.080: tiǎn
U+992C 𨾏	0474.060*: hú	U+2B5EB 𨾏	0474.060: hú
U+9958 𨾏	1450.130*: zhān	U+2B5F4 𨾏	1450.130: zhān
U+99C3 𨾏	0620.080*: jué	U+2B61D 𨾏	0620.080: jué
U+99C9 𨾏	0604.060: jiōng	U+2CCF6 𨾏	0604.060: jiōng
U+99DC 𨾏	0059.110: bì	U+31148 𨾏	0059.110: bì
U+99DE 𨾏	1174.050: tuó	U+2B61E 𨾏	1174.050: tuó
U+99EA 𨾏	1017.100*: shēn	U+2CCFD 𨾏	1017.100: shēn
U+9A04 𨾏	0740.090*: lù	U+2B627 𨾏	0740.090: lù
U+9A11 𨾏	0313.070: fēi	U+2CD02 𨾏	0313.070: fēi
U+9A1E 𨾏	0509.060*: huō	U+2CD03 𨾏	0509.060: huō
U+9A20 𨾏	1129.010*: tí	U+2B628 𨾏	1129.010: tí
U+9A23 𨾏	1540.020*: zōng	U+2EAA3 𨾏	1540.020: zōng
U+9A3B 𨾏	1073.060: shuāng	U+2B62D 𨾏	1073.060: shuāng
U+9A50 𨾏	0277.050: dūn	U+2EAA5 𨾏	0277.050: dūn
U+9A69 𨾏	0489.030*: huān	U+2CD10 𨾏	0489.030: huān

Current Character	Current Entry	New Character	New Entry
U+9B86 𩚑	0176.030*:cǐ	U+2B696 𩚑	0176.030:cǐ
U+9B88 𩚒	0610.020*:jū	U+2CD8B 𩚒	0610.020:jū
U+9BA0 𩚓	1196.040*:wéi	U+2CD8F 𩚓	1196.040:wéi
U+9BA1 𩚔	1461.030*:zhào	U+2CD90 𩚔	1461.030:zhào
U+9BA8 𩚕	1371.080*:yì	U+2EB1C 𩚕	1371.080:yì
U+9C15 𩚖	1238.030*:xiā	U+2B6A5 𩚖	1238.030:xiā
U+9C36 𩚗	0535.030*:jì	U+2CDAD 𩚗	0535.030:jì
U+9C5A 𩚘	1234.030*:xǐ	U+2CDAE 𩚘	1234.030:xǐ
U+9C72 𩚙	0715.020*:liè	U+2B6AD 𩚙	0715.020:liè
U+9C75 𩚚	1467.080*:zhēn	U+2EB24 𩚚	1467.080:zhēn
U+9D02 𩚛	0620.050*:jué	U+31254 𩚛	0620.050:jué
U+9D25 𩚜	1416.130*:yù	U+2B6E3 𩚜	1416.130:yù
U+9D4F 𩚝	0086.020*:bǔ	U+2CDD5 𩚝	0086.020:bǔ
U+9D5F 𩚞	0660.020*:kuáng	U+2B6ED 𩚞	0660.020:kuáng
U+9D69 𩚟	0339.030*:fú	U+2B6F3 𩚟	0339.030:fú
U+9D82 𩚠	1371.070*:yì	U+31288 𩚠	1371.070:yì
U+9DC3 𩚡	1332.080*:yàn	U+2EB68 𩚡	1332.080:yàn
U+9DC7 𩚢	0652.030*:kòu	U+2C1AE 𩚢	0652.030:kòu
U+9DC9 𩚣	1127.020:tī	U+4D18 𩚣	1127.020:tī
U+9DDE 𩚤	1073.050:shuāng	U+2EB6A 𩚤	1073.050:shuāng
U+9DDF 𩚥	1527.040*:zhuó	U+2CE26 𩚥	1527.040:zhuó
U+9DED 𩚦	0300.070*:fán	U+2CE2A 𩚦	0300.070:fán
U+9E1C 𩚧	0944.160:qú	U+2CE31 𩚧	0944.160:qú
U+9F58 𩚨	1277.100*:xiè	U+2CE7C 𩚨	1277.100:xiè
U+9F6E 𩚩	1363.170*:yǐ	U+2CE88 𩚩	1363.170:yǐ
U+9F70 𩚪	1444.080*:zé	U+2B72C 𩚪	1444.080:zé
U+9F7C 𩚫	0160.070*:chǔ	U+2CE93 𩚫	0160.070:chǔ
U+279A7 𩚬	1363.090*:yí	U+2B35F 𩚬	1363.090:yí
U+27D2A 𩚭	1103.090*:suǒ	U+2D546 𩚭	1103.090:suǒ
U+29FFF 𩚮	0160.020*:chú	U+2CE05 𩚮	0160.020:chú
U+2A04F 𩚯	0133.010*:chén	U+2EB66 𩚯	0133.010:chén

11. Entries which should be divided into separate simplified/traditional entries

The following 69 entries correspond to pairs of simplified and traditional character forms in the dictionary, but are currently only mapped to the traditional form (because the simplified form was not encoded when the data was created). The current entries should be divided into two entries as shown below.

Current Character	Current Entry	New Characters	New Entries
U+379E 屙	1092.070*, 1092.071:sóng	U+2AA0A 屙 U+379E 屙	1092.070:sóng 1092.071:sóng
U+3A4B 搨	1265.040*, 1265.041:xiāo	U+2AB8B 搨 U+3A4B 搨	1265.040:xiāo 1265.041:xiāo
U+3F08 璿	0749.020*, 0749.021:luó	U+2DE5C 璿 U+3F08 璿	0749.020:luó 0749.021:luó

Current Character	Current Entry	New Characters	New Entries
U+426C 籜	0317.010*, 0317.011:fèi	U+2B088 籜 U+426C 籜	0317.010:fèi 0317.011:fèi
U+45FB 蝻	0580.010*, 0580.011:jié	U+2E502 蝻 U+45FB 蝻	0580.010:jié 0580.011:jié
U+490C 醅	0917.010*, 0917.011:qiāng	U+2E81E 醅 U+490C 醅	0917.010:qiāng 0917.011:qiāng
U+4D1A 𪗇	0362.020*, 0362.021:gǎng	U+2EB70 𪗇 U+4D1A 𪗇	0362.020:gǎng 0362.021:gǎng
U+5131 隴	0733.030*, 0733.031:lǒng	U+2B892 隴 U+5131 隴	0733.030:lǒng 0733.031:lǒng
U+51D4 滄	0167.060*, 0167.061:chuàng	U+300F7 滄 U+51D4 滄	0167.060:chuàng 0167.061:chuàng
U+51D9 澤	0280.100*, 0280.101:duó	U+2A79D 澤 U+51D9 澤	0280.100:duó 0280.101:duó
U+5368 𪗇	1277.110*, 1277.111:xiè	U+2B9EF 𪗇 U+5368 𪗇	1277.110:xiè 1277.111:xiè
U+57E8 垠	0747.050*, 0747.051:lǔn	U+2BB62 垠 U+57E8 垠	0747.050:lǔn 0747.051:lǔn
U+58A0 埴	0999.110*, 0999.111:shàn	U+2BB83 埴 U+58A0 埴	0999.110:shàn 0999.111:shàn
U+5972 𪗇	0281.020*, 0281.021:duǒ	U+2BC02 𪗇 U+5972 𪗇	0281.020:duǒ 0281.021:duǒ
U+5C35 𪗇	1168.010, 1168.011:tuí	U+2AA07 𪗇 U+2BD52 𪗇	1168.010:tuí 1168.011:tuí
U+5D19 崱	0747.020*, 0747.021:lún	U+2AA27 崱 U+5D19 崱	0747.020:lún 0747.021:lún
U+5D7D 𪗇	0251.140, 0251.141:dié	U+2BD87 𪗇 U+5D7D 𪗇	0251.140:dié 0251.141:dié
U+5DD8 𪗇	1329.080*, 1329.081:yǎn	U+2AA58 𪗇 U+5DD8 𪗇	1329.080:yǎn 1329.081:yǎn
U+5E60 𪗇	0470.070*, 0470.071:hū	U+2D613 𪗇 U+5E60 𪗇	0470.070:hū 0470.071:hū
U+5F44 𪗇	0649.060, 0649.061:kōu	U+2BE29 𪗇 U+5F44 𪗇	0649.060:kōu 0649.061:kōu
U+5F4D 𪗇	0422.130*, 0422.131:guō	U+2D6A6 𪗇 U+5F4D 𪗇	0422.130:guō 0422.131:guō
U+640A 𪗇	0152.030*, 0152.031:chōu	U+2BF1D 𪗇 U+640A 𪗇	0152.030:chōu 0152.031:chōu
U+6690 𪗇	1197.030*, 1197.031:wěi	U+2C029 𪗇 U+6690 𪗇	1197.030:wěi 1197.031:wěi
U+66CF 𪗇	1259.040*, 1259.041:xiǎng	U+2C060 𪗇 U+66CF 𪗇	1259.040:xiǎng 1259.041:xiǎng
U+6A9B 𪗇	1516.010*, 1516.011:zhuā	U+2DA86 𪗇 U+6A9B 𪗇	1516.010:zhuā 1516.011:zhuā
U+6AFD 𪗇	1382.030, 1382.032:yǐn	U+2C129 𪗇 U+6AFD 𪗇	1382.030:yǐn 1382.032:yǐn

Current Character	Current Entry	New Characters	New Entries
U+6F0D 澗	0426.010*, 0426.011:guó	U+2C1F9 澗 U+6F0D 澗	0426.010:guó 0426.011:guó
U+6F19 溥	1165.070*, 1165.071:tuán	U+2C1D8 溥 U+6F19 溥	1165.070:tuán 1165.071:tuán
U+6FAB 瀉	1186.010*, 1186.011:wàn	U+2C1D5 瀉 U+6FAB 瀉	1186.010:wàn 1186.011:wàn
U+6FC7 漚	0989.100*, 0989.101:sè	U+2C227 漚 U+6FC7 漚	0989.100:sè 0989.101:sè
U+7416 璣	1451.060*, 1451.061:zhǎn	U+2C359 璣 U+7416 璣	1451.060:zhǎn 1451.061:zhǎn
U+748A 璚	0776.040*, 0776.041:mén	U+2B7A9 璚 U+748A 璚	0776.040:mén 0776.041:mén
U+7495 璠	1315.020*, 1315.021:xún	U+2C364 璠 U+7495 璠	1315.020:xún 1315.021:xún
U+74C5 璣	0698.070*, 0698.071:lì	U+2C35B 璣 U+74C5 璣	0698.070:lì 0698.071:lì
U+77D1 臚	0736.090*, 0736.091:lú	U+2AFA6 臚 U+77D1 臚	0736.090:lú 0736.091:lú
U+7929 磧	1494.010*, 1494.011:zhì	U+309C9 磧 U+7929 磧	1494.010:zhì 1494.011:zhì
U+7DC9 綱	0709.060*, 0709.061:liǎng	U+2E267 綱 U+7DC9 綱	0709.060:liǎng 0709.061:liǎng
U+7DFA 鍋	0402.090*, 0402.091:guā	U+2E268 鍋 U+7DFA 鍋	0402.090:guā 0402.091:guā
U+7E68 縫	0203.060*, 0203.061:da	U+2B124 縫 U+7E68 縫	0203.060:da 0203.061:da
U+7E91 纜	0737.030*, 0737.031:lú	U+2E261 纜 U+7E91 纜	0737.030:lú 0737.031:lú
U+7E9A 纜	0691.150*, 0691.151:lí 1234.040*, 1234.041:xǐ	U+2B125 纜 U+7E9A 纜	0691.150:lí 1234.040:xǐ 0691.151:lí 1234.041:xǐ
U+848D 蕪	1196.060*, 1196.061:wěi	U+2B1ED 蕪 U+848D 蕪	1196.060:wěi 1196.061:wěi
U+8586 菱	0005.030*, 0005.031:ài	U+2B241 菱 U+8586 菱	0005.030:ài 0005.031:ài
U+860B 蘋	0877.010:pín 0883.021:píng	U+2C79F 蘋 U+860B 蘋	0877.010:pín 0883.021:píng
U+873D 蛎	0709.030*, 0709.031:liǎng	U+2E50A 蛎 U+873D 蛎	0709.030:liǎng 0709.031:liǎng
U+8740 蝻	0258.010*, 0258.011:dōng	U+2C7FD 蝻 U+8740 蝻	0258.010:dōng 0258.011:dōng
U+88F2 襉	0709.010*, 0709.011:liǎng	U+2E581 襉 U+88F2 襉	0709.010:liǎng 0709.011:liǎng
U+8A5D 訃	1512.030*, 1512.031:zhǔ	U+2C8DE 訃 U+8A5D 訃	1512.030:zhǔ 1512.031:zhǔ

Current Character	Current Entry	New Characters	New Entries
U+8AD3 諺	0554.060*, 0554.061:jiàn	U+2C8E1 諺 U+8AD3 諺	0554.060:jiàn 0554.061:jiàn
U+8B4A 饒	0819.070*, 0819.071:náo	U+2B362 饶 U+8B4A 饒	0819.070:náo 0819.071:náo
U+8B8F 饗	1201.070*, 1201.071:wèi	U+2C8C0 饗 U+8B8F 饗	1201.070:wèi 1201.071:wèi
U+8B94 饗	1382.020*, 1382.021:yǐn	U+2E64A 饗 U+8B94 饗	1382.020:yǐn 1382.021:yǐn
U+8B9F 饗	0267.010*, 0267.011:dú	U+2E64B 饗 U+8B9F 饗	0267.010:dú 0267.011:dú
U+8E61 躑	0916.070*, 0916.071:qiāng 0919.050*, 0919.051:qiàng	U+2C9C0 躑 U+8E61 躑	0916.070:qiāng 0919.050:qiàng 0916.071:qiāng 0919.051:qiàng
U+9129 鄴	1315.050*, 1315.051:xún	U+2CA7D 鄴 U+9129 鄴	1315.050:xún 1315.051:xún
U+91B2 醲	0840.010*, 0840.011:nóng	U+2CAA9 醲 U+91B2 醲	0840.010:nóng 0840.011:nóng
U+92F9 銀	0124.050*, 0124.051:chǎng	U+2CB2E 钅 U+92F9 銀	0124.050:chǎng 0124.051:chǎng
U+9462 鑢	0742.070*, 0742.071:lù	U+2E8F6 鑢 U+9462 鑢	0742.070:lù 0742.071:lù
U+967F 陘	1238.090*, 1238.091:xiá	U+2CBC5 陘 U+967F 陘	1238.090:xiá 1238.091:xiá
U+9691 陘	0352.080*, 0352.081:gài	U+2CBBF 陘 U+9691 陘	0352.080:gài 0352.081:gài
U+97D4 鞞	0125.030*, 0125.031:chàng	U+2E9F4 鞞 U+97D4 鞞	0125.030:chàng 0125.031:chàng
U+97DB 鞞	0026.020*, 0026.021:bài	U+2B594 鞞 U+97DB 鞞	0026.020:bài 0026.021:bài
U+97E1 鞞	1197.060*, 1197.061:wěi	U+2E9F5 鞞 U+97E1 鞞	1197.060:wěi 1197.061:wěi
U+9857 顛	1364.050*, 1364.051:yǐ	U+2B5AE 顛 U+9857 顛	1364.050:yǐ 1364.051:yǐ
U+9946 饌	0060.150*, 0060.151:bì	U+2EA5B 饌 U+9946 饌	0060.150:bì 0060.151:bì
U+9960 饌	0750.080*, 0750.081:luó	U+2B5E9 饌 U+9960 饌	0750.080:luó 0750.081:luó
U+9B21 鬚	0835.120*, 0835.121:ning	U+2EAC2 鬚 U+9B21 鬚	0835.120:ning 0835.121:ning
U+9C65 鰓	0420.140*, 0420.141:guì	U+2EB1D 鰓 U+9C65 鰓	0420.140:guì 0420.141:guì
U+9E11 鸞	1430.050*, 1430.051:yuè	U+2CE1A 鸞 U+9E11 鸞	1430.050:yuè 1430.051:yuè

12. Entries for variant characters which should be divided into two entries

The following 4 entries correspond to pairs of variant character forms in the dictionary, but are currently only mapped to one character. The current entries should be divided into two entries as shown below.

Current Character	Current Entry	New Characters	New Entries
U+5406 𠵹	1340.050, 1340.051*:yāo	U+5406 𠵹 U+4DB8 𠵹	1340.050:yāo 1340.051:yāo
U+5523 𠵹	1444.020, 1444.021:zào	U+5523 𠵹 U+5515 𠵹	1444.020:zào 1444.021:zào
U+5C44 𠵹	0054.040, 0054.041*:bī	U+5C44 𠵹 U+2BD58 𠵹	0054.040:bī 0054.041:bī
U+96DC 雜	1435.081*, 1435.082:zá	U+305E3 𠵹 U+96DC 雜	1435.081:zá 1435.082:zá

Appendix. Unifiable variants

The dictionary includes 7 variant character forms which are unifiable variants of encoded characters. These are listed below for reference. The character in the left column is the standard G-source glyph form, and the character in the right column is the variant form which is also given in the dictionary (marked with a star in the kXHC1983 entry).

Character	Entry	Variant (*) Form
U+6035 𠵹	0161.040, 0161.041*:chù	𠵹
U+649D 搗	0500.031*, 0500.032:huī	搗
U+70AD 炭	1116.020, 1116.021*:tàn	炭
U+7424 琤	0134.060, 0134.061*:chēng	琤
U+8005 者	1464.090, 1464.091*:zhě	者
U+90A6 邦	0033.050, 0033.051*:bāng	邦
U+9694 隔	0371.090, 0371.091*:gé	隔

The above characters are the only ones in the new kXHC1983 data which still include a starred reference.

Additionally, there are a few head characters in the dictionary which have a significantly different (but unifiable) glyph form compared with the G-source glyph given in the Unicode code charts. There may be other examples that I have overlooked. Potentially, the references for these characters could be marked with a star to indicate that the glyph form is not the expected G-source glyph form.

Character	Entry	Dictionary Form
U+9B2E 𠵹	0604.191:jiū	𠵹 (old style tortoise)
U+9F9C 龜	0418.061:guī 0624.021:jūn 0939.101:qiū	龜 (old style tortoise)
U+9F9D 龜	0939.111:qiū	龜 (old style tortoise)
U+2B6F8 𠵹	1128.010:tí	𠵹 (𠵹 vs. 𠵹)