

L2/02-018

To: UTC
 Re: CJK Numeric Value Data Table
 From: Mark Davis, Helena Shih Chapman
 Date: 2001-01-14

The only property values in Chapter 4 of The Unicode Standard that are not available in machine readable format are *Table 4-7. Primary Numeric Ideographs* and *Table 4-8. Ideographs Used as Accounting Numbers*, copied at the end of this document.

We propose that a data table for these values be added to the next appropriate version of the UCD. The contents could look something like the following, with the values P and A to distinguish Primary from Accounting. The proposed name is CJKNumericValues.txt, although the precise name is not important.

This simply matches the content of the Unicode Standard, so could conceivably be added in 3.2, if the UTC desired. **Note:** the actual ideographs in the comments below would not be present in the file, at least until such time as we decide to allow UTF-8 data in data file comments. They are provided here to help double-check the values.

96F6; P; 0	# 零
4E00; P; 1	# 一
4E8C; P; 2	# 二
4E09; P; 3	# 三
56DB; P; 4	# 四
4E94; P; 5	# 五
516D; P; 6	# 六
4E03; P; 7	# 七
516B; P; 8	# 八
4E5D; P; 9	# 九
5341; P; 10	# 十
767E; P; 100	# 百
5343; P; 1000	# 千
4E07; P; 10000	# 万
5104; P; 100000000	# 億
5146; P; 1000000000000	# 兆
58F9; A; 1	# 壹
58F1; A; 1	# 貳
5F0C; A; 1	# 弍
8CAE; A; 2	# 貳
8D30; A; 2	# 点
5F10; A; 2	# 弍
5F0D; A; 2	# 弍
53C3; A; 3	# 参
53C2; A; 3	# 参
53C1; A; 3	# 点
5F0E; A; 3	# 点
8086; A; 4	# 肆

4F0D; A; 5	# 伍
9678; A; 6	# 陸
9646; A; 6	# •
67D2; A; 7	# 柒
634C; A; 8	# 捌
7396; A; 9	# 玖
62FE; A; 10	# 拾
4F70; A; 100	# 佰
964C; A; 100	# 陌
4EDF; A; 1000	# 仟
842C; A; 10000	# 萬

Future Additions

In addition, we should collect other ideographs that can have numeric values for addition to a future version, such as the following:

5169; N; 2	# 兩 liang3
5EFF; N; 20	# 廿 shorthand used in bookkeeping/newspapers
5345; N; 30	# 卅 shorthand used in bookkeeping/newspapers

The following tables are screen-shots from the Unicode Standard, Chapter 4.

Table 4-7. Primary Numeric Ideographs

U+96F6	0
U+4E00	1
U+4E8C	2
U+4E09	3
U+56DB	4
U+4E94	5
U+516D	6
U+4E03	7
U+516B	8
U+4E5D	9
U+5341	10
U+767E	100
U+5343	1,000
U+4E07	10,000
U+5104	100,000,000 (10,000 × 10,000)
U+5146	1,000,000,000,000 (10,000 × 10,000 × 10,000)

Table 4-8. Ideographs Used as Accounting Numbers

1	U+58F9, U+58F1, U+5F0C ^a
2	U+8CAE ^a , U+8D30 ^a , U+5F10 ^a , U+5F0D ^a
3	U+53C3, U+53C2, U+53C1 ^a , U+5F0E ^a
4	U+8086
5	U+4F0D
6	U+9678, U+9646
7	U+67D2 ^b
8	U+634C
9	U+7396
10	U+62FE
100	U+4F70 ^a , U+964C
1,000	U+4EDF
10,000	U+842C