

Inclusion of East Asian lists CSS3 module requests

Prepared by Seiichi Kato
Microsoft Development Co., Ltd.
December, 2009

Table of Contents

Inclusion of East Asian lists CSS3 module requests	1
Background and justification:	3
Lists that needed to be corrected from the current CSS3 working draft.....	4
cjk-ideographic algorithm.....	4
Current listing in CSS3 working draft:	4
Corrected version of the cjk-ideographic algorithm:	5
Numbering lists	5
Japanese-informal.....	6
Current listing in CSS3 working draft:	6
Corrected version of the numbering list for Japanese-informal:	6
Japanese-formal	7
Current listing in CSS3 working draft:	7
Corrected version of the numbering list for Japanese-informal:	8
Simplified Chinese-informal.....	8
Current listing in CSS3 working draft:	8
Corrected version of the numbering list for Simplified Chinese-informal:	9
Simplified Chinese-formal	10
Current listing in CSS3 working draft:	10
Corrected version of the numbering list for Simplified Chinese-formal:	10
Traditional Chinese-informal	11
Current listing in CSS3 working draft:	11
Corrected version of the numbering list for Traditional Chinese-informal:	12
Traditional Chinese-formal.....	12
Current listing in CSS3 working draft:	12
Corrected version of the numbering list for Traditional Chinese-formal:	13
Lists to be added to CSS3 working draft.....	15
Hangul-algorithm.....	15
Hangul-legal	15
JapaneseKanji-num	16
Hangul-num.....	17
Hanja-num	17
CjkFull-num.....	17
Katakanahalf.....	18
Katakanahalf-iroha.....	19
Upperfull-alpha	20

This document serves as an addendum/amendment to the CSS3 working draft, at <http://www.w3.org/TR/css3-lists/>, for additional considerations of East Asian lists formats.

Background and justification:

- The lists approved in W3C are only standardized for the alphanumeric sequences of A, B, C; and 1, 2, 3. <http://msdn2.microsoft.com/en-us/library/ms530797.aspx>
- East Asian lists are commonly used in businesses, academics and government, along with alphanumeric numbering lists.
- For HTML output, none of East Asian lists are officially supported as of the current implementation. Lists are standardized in CSS, not in HTML.
- Some East Asian lists are mentioned in the most recent update of CSS3 working draft, though there are some omissions and inaccurate listings. This document is intended to cover the missing lists and to correct the inaccurate listings.

Lists that needed to be corrected from the current CSS3 working draft

cjk-ideographic algorithm

The cjk-ideographic algorithm section in 4.2 should be corrected to indicate the differences of formal versus informal numberings.

Current listing in CSS3 working draft:

The cjk-ideographic algorithm is used by several numbering systems, using different sets of digits. These systems are defined for numbers greater than or equal to 0 and less than 10^{16} . Numbers less than zero or equal to or greater than 10^{16} should use the decimal system. The core algorithm is as follows:

1. Split the decimal number into groups of four digits, starting with the least significant digit.
2. Ignoring groups that have the value zero, append the second group marker to the second group, the third group marker to the third group, and the fourth group marker to the fourth group. These markers are defined in the tables for the specific numbering systems. The first group has no marker.
3. For each group, ignoring digits that have the value zero, append the second digit marker to the second digit, the third digit marker to the third digit, and the fourth digit marker to the fourth digit. These markers are defined in the tables for the specific numbering systems. The first digit has no marker.
4. For any group with a value less than 20, remove the second digit (the 1 in the tens column). Leave any associated markers.
5. Concatenate the groups back into a single string, least significant group last.
6. Collapse any consecutive runs of 0 digits to a single 0.
7. Replace each digit with the relevant character selected from the numbering system's table.

The suffix for the cjk-ideographic numbering systems is a dot . U+002E.

If specified explicitly, the cjk-ideograph keyword should be handled like trad-chinese-informal.

Corrected version of the cjk-ideographic algorithm:

The cjk-ideographic algorithm is used by several numbering systems, using different sets of digits. These systems are defined for numbers greater than or equal to 0 and less than 10^{16} . Numbers less than zero or equal to or greater than 10^{16} should use the decimal system. The core algorithm is as follows:

1. Split the decimal number into groups of four digits, starting with the least significant digit.
2. Ignoring groups that have the value zero, append the second group marker to the second group, the third group marker to the third group, and the fourth group marker to the fourth group. These markers are defined in the tables for the specific numbering systems. The first group has no marker.
3. For each group, ignoring digits that have the value zero, append the second digit marker to the second digit, the third digit marker to the third digit, and the fourth digit marker to the fourth digit. These markers are defined in the tables for the specific numbering systems. The first digit has no marker.
4. For any group with a value less than 20, remove the second digit (the 1 in the tens column). Leave any associated markers. This is applicable to informal lists only. **For formal lists, do not remove the second digit even with a value less than 20.**
5. Concatenate the groups back into a single string, least significant group last.
6. Collapse any consecutive runs of 0 digits to a single 0.
7. Replace each digit with the relevant character selected from the numbering system's table.

The suffix for the cjk-ideographic numbering systems is a dot . U+002E.

If specified explicitly, the cjk-ideograph keyword should be handled like trad-chinese-informal.

Numbering lists

There are several sets of information of numbering lists in section 4.2 that needed to be corrected from the current CSS3 working draft, because the characters mentioned in the tables are inaccurate and misrepresented.

Japanese-informal

Current listing in CSS3 working draft:

Section 4.2 Algorithmic

japanese-informal

This uses the cjk-ideographic system with the following table.

<i>Informal Japanese numbering system</i>		
Values	Codepoints	
Second Group Marker	萬	U+842C
Third Group Marker	億	U+5104
Fourth Group Marker	兆	U+5146
Second Digit Marker	萬	U+842C
Third Digit Marker	億	U+5104
Fourth Digit Marker	兆	U+5146
Digit 0	零	U+96F6
Digit 1	壹	U+58F9
Digit 2	貳	U+8D30
Digit 3	叁	U+53C1
Digit 4	肆	U+8086
Digit 5	伍	U+4F0D
Digit 6	陆	U+9646
Digit 7	柒	U+67D2
Digit 8	捌	U+634C
Digit 9	玖	U+7396

Corrected version of the numbering list for Japanese-informal:

japanese-informal

This uses the cjk-ideographic system with the following table.

<i>Informal Japanese numbering system</i>		
Values	Codepoints	
Second Group Marker	万	U+4E07
Third Group Marker	億	U+5104
Fourth Group Marker	兆	U+5146

Second Digit Marker	十	U+5341
Third Digit Marker	百	U+767E
Fourth Digit Marker	千	U+5343
Digit 0	零	U+96F6
Digit 1	一	U+4E00
Digit 2	二	U+4E8C
Digit 3	三	U+4E09
Digit 4	四	U+56DB
Digit 5	五	U+4E94
Digit 6	六	U+516D
Digit 7	七	U+4E03
Digit 8	八	U+516B
Digit 9	九	U+4E5D

Japanese-formal

Current listing in CSS3 working draft:

Section 4.2 Algorithmic

japanese-formal

This uses the cjk-ideographic system with the following table.

<i>Formal Japanese numbering system</i>		
Values	Codepoints	
Second Group Marker	万	U+4E07
Third Group Marker	億	U+5104
Fourth Group Marker	兆	U+5146
Second Digit Marker	拾	U+62FE
Third Digit Marker	佰	U+4F70
Fourth Digit Marker	仟	U+4EDF
Digit 0	零	U+96F6
Digit 1	壹	U+58F9
Digit 2	貳	U+8CB3
Digit 3	參	U+53C3
Digit 4	肆	U+8086
Digit 5	伍	U+4F0D
Digit 6	陸	U+9678

Digit 7	柒	U+67D2
Digit 8	捌	U+634C
Digit 9	玖	U+7396

Corrected version of the numbering list for Japanese-informal:

Section 4.2 Algorithmic

japanese-formal

This uses the cjk-ideographic system with the following table.

<i>Formal Japanese numbering system</i>		
Values	Codepoints	
Second Group Marker	萬	U+842C
Third Group Marker	億	U+5104
Fourth Group Marker	兆	U+5146
Second Digit Marker	拾	U+62FE
Third Digit Marker	佰	U+4F70
Fourth Digit Marker	仟	U+4EDF
Digit 0	零	U+96F6
Digit 1	壹	U+58F1
Digit 2	貳	U+5F10
Digit 3	参	U+53C2
Digit 4	四	U+56DB
Digit 5	伍	U+4F0D
Digit 6	六	U+516D
Digit 7	七	U+4E03
Digit 8	八	U+516B
Digit 9	九	U+4E5D

Simplified Chinese-informal

Current listing in CSS3 working draft:

Section 4.2 Algorithmic

This uses the cjk-ideographic system with the following table.

<i>Informal simple Chinese numbering system</i>

Values	Codepoints	
Second Group Marker	萬	U+842C
Third Group Marker	億	U+5104
Fourth Group Marker	兆	U+5146
Second Digit Marker	萬	U+842C
Third Digit Marker	億	U+5104
Fourth Digit Marker	兆	U+5146
Digit 0	零	U+96F6
Digit 1	壹	U+58F9
Digit 2	貳	U+8CB3
Digit 3	參	U+53C3
Digit 4	肆	U+8086
Digit 5	伍	U+4F0D
Digit 6	陸	U+9678
Digit 7	柒	U+67D2
Digit 8	捌	U+634C
Digit 9	玖	U+7396

Corrected version of the numbering list for Simplified Chinese-informal:

Section 4.2 Algorithmic

This uses the cjk-ideographic system with the following table.

<i>Informal simplified Chinese numbering system</i>		
Values	Codepoints	
Second Group Marker	万	U+4E07
Third Group Marker	亿	U+4EBF
Fourth Group Marker	兆	U+5146
Second Digit Marker	十	U+5341
Third Digit Marker	百	U+767E
Fourth Digit Marker	千	U+5343
Digit 0	〇	U+3007
Digit 1	一	U+4E00
Digit 2	二	U+4E8C
Digit 3	三	U+4E09
Digit 4	四	U+56DB

Digit 5	五	U+4E94
Digit 6	六	U+516D
Digit 7	七	U+4E03
Digit 8	八	U+516B
Digit 9	九	U+4E5D

Simplified Chinese-formal

Current listing in CSS3 working draft:

Section 4.2 Algorithmic

This uses the cjk-ideographic system with the following table.

<i>Formal simple Chinese numbering system</i>		
Values	Codepoints	
Second Group Marker	万	U+4E07
Third Group Marker	億	U+5104
Fourth Group Marker	兆	U+5146
Second Digit Marker	万	U+4E07
Third Digit Marker	亿	U+4EBF
Fourth Digit Marker	兆	U+5146
Digit 0	零	U+96F6
Digit 1	壹	U+58F9
Digit 2	貳	U+8CB3
Digit 3	參	U+53C3
Digit 4	肆	U+8086
Digit 5	伍	U+4F0D
Digit 6	陸	U+9678
Digit 7	柒	U+67D2
Digit 8	捌	U+634C
Digit 9	玖	U+7396

Corrected version of the numbering list for Simplified Chinese-formal:

Section 4.2 Algorithmic

This uses the cjk-ideographic system with the following table.

<i>Formal simplified Chinese numbering system</i>

Values	Codepoints	
Second Group Marker	万	U+4E07
Third Group Marker	亿	U+4EBF
Fourth Group Marker	兆	U+5146
Second Digit Marker	拾	U+62FE
Third Digit Marker	佰	U+4F70
Fourth Digit Marker	仟	U+4EDF
Digit 0	零	U+96F6
Digit 1	壹	U+58F9
Digit 2	貳	U+8D30
Digit 3	叁	U+53C1
Digit 4	肆	U+8086
Digit 5	伍	U+4F0D
Digit 6	陆	U+9646
Digit 7	柒	U+67D2
Digit 8	捌	U+634C
Digit 9	玖	U+7396

Traditional Chinese-informal

Current listing in CSS3 working draft:

Section 4.2 Algorithmic

This uses the cjk-ideographic system with the following table.

<i>Informal traditional Chinese numbering system</i>		
Values	Codepoints	
Second Group Marker	萬	U+842C
Third Group Marker	億	U+5104
Fourth Group Marker	兆	U+5146
Second Digit Marker	萬	U+842C
Third Digit Marker	億	U+5104
Fourth Digit Marker	兆	U+5146
Digit 0	零	U+96F6
Digit 1	一	U+4E00
Digit 2	二	U+4EBC

Digit 3	三	U+4E09
Digit 4	四	U+56DB
Digit 5	五	U+4E94
Digit 6	六	U+516D
Digit 7	七	U+4E03
Digit 8	八	U+516B
Digit 9	九	U+4E5D

Corrected version of the numbering list for Traditional Chinese-informal:

Section 4.2 Algorithmic

This uses the cjk-ideographic system with the following table.

<i>Informal traditional Chinese numbering system</i>		
Values	Codepoints	
Second Group Marker	萬	U+842C
Third Group Marker	億	U+5104
Fourth Group Marker	兆	U+5146
Second Digit Marker	十	U+5341
Third Digit Marker	百	U+767E
Fourth Digit Marker	千	U+5343
Digit 0	〇	U+3007
Digit 1	一	U+4E00
Digit 2	二	U+4E8C
Digit 3	三	U+4E09
Digit 4	四	U+56DB
Digit 5	五	U+4E94
Digit 6	六	U+516D
Digit 7	七	U+4E03
Digit 8	八	U+516B
Digit 9	九	U+4E5D

Traditional Chinese-formal

Current listing in CSS3 working draft:

Section 4.2 Algorithmic

This uses the cjk-ideographic system with the following table.

<i>Formal Traditional Chinese numbering system</i>		
Values	Codepoints	
Second Group Marker	万	U+4E07
Third Group Marker	亿	U+4EBF
Fourth Group Marker	兆	U+5146
Second Digit Marker	万	U+4E07
Third Digit Marker	亿	U+4EBF
Fourth Digit Marker	兆	U+5146
Digit 0	零	U+96F6
Digit 1	一	U+4E00
Digit 2	二	U+4EBC
Digit 3	三	U+4E09
Digit 4	四	U+56DB
Digit 5	五	U+4E94
Digit 6	六	U+516D
Digit 7	七	U+4E03
Digit 8	八	U+516B
Digit 9	九	U+4E5D

Corrected version of the numbering list for Traditional Chinese-formal:

Section 4.2 Algorithmic

This uses the cjk-ideographic system with the following table.

<i>Formal traditional Chinese numbering system</i>		
Values	Codepoints	
Second Group Marker	萬	U+842C
Third Group Marker	億	U+5104
Fourth Group Marker	兆	U+5146
Second Digit Marker	拾	U+62FE
Third Digit Marker	佰	U+4F70
Fourth Digit Marker	仟	U+4EDF
Digit 0	零	U+96F6
Digit 1	壹	U+58F9

Digit 2	貳	U+8CB3
Digit 3	參	U+53C3
Digit 4	肆	U+8086
Digit 5	伍	U+4F0D
Digit 6	陸	U+9678
Digit 7	柒	U+67D2
Digit 8	捌	U+634C
Digit 9	玖	U+7396

Lists to be added to CSS3 working draft

Hangul-algorithm

To be added to section 4.2 Algorithmic

This uses the cjk-ideographic system with the following table.

<i>hangul-algorithm numbering system</i>		
Values	Codepoints	
Second Group Marker	만	U+B9CC
Third Group Marker	억	U+C5B5
Fourth Group Marker	조	U+C870
Second Digit Marker	십	U+C2ED
Third Digit Marker	백	U+BC31
Fourth Digit Marker	천	U+CC9C
Digit 0	영	U+C601
Digit 1	일	U+C77C
Digit 2	이	U+C774
Digit 3	삼	U+C0BC
Digit 4	사	U+C0AC
Digit 5	오	U+C624
Digit 6	육	U+C721
Digit 7	칠	U+CE60
Digit 8	팔	U+D314
Digit 9	구	U+AD6C

Hangul-legal

To be added to section 4.2 Algorithmic

This uses the cjk-ideographic system for digit markers greater than 100, and group markers with the following table.

<i>hangul-legal numbering system</i>		
Values	Codepoints	
Second Group Marker (at 10,000)	만	U+B9CC
Third Group Marker (at 100,000,000)	억	U+C5B5

Fourth Group Marker (at 1,000,000,000,000)	조	U+C870
Third Digit Marker (at 100)	백	U+BC31
Fourth Digit Marker (at 1,000)	천	U+CC9C

For values between 1 and 99, appropriate digits are picked from the following list (at most one per column) and written in descending order by value.

<i>hangul-legal numbering system</i>					
Tens			Units		
Values	Codepoints		Values	Codepoints	
10	열	U+C2ED	1	하나	U+D558U+B098
20	스물	U+C2A4U+BB3C	2	둘	U+B458
30	서른	U+C11CU+B978	3	셋	U+C14B
40	마흔	U+B9C8U+D754	4	넷	U+B137
50	쉰	U+C270	5	다섯	U+B2E4U+C12F
60	예순	U+C608U+C21C	6	여섯	U+C5ECU+C12F
70	일흔	U+C77CU+D754	7	일곱	U+C77CU+ACF1
80	여든	U+C5ECU+B4E0	8	여덟	U+C5ECU+B35F
90	아흔	U+C544U+D754	9	아홉	U+C544U+D649

JapaneseKanji-num

To be added to section 4.3 Numeric

System	Characters	Codepoints	Base	Suffix	Notes
<i>japanesekanji-num</i>	〇 一 二 三 四 五 六 七 八 九	U+3007, U+4E00, U+4E8C, U+4E09, U+56DB, U+4E94, U+516D, U+4E03, U+516B, U+4E5D	10	. U+002E	The repeating scheme is identical to other base 10 lists mentioned in this numeric

					section.
--	--	--	--	--	----------

Hangul-num

To be added to section 4.3 Numeric

System	Characters	Codepoints	Base	Suffix	Notes
<i>hangul-num</i>	영일이삼사오육칠팔구	U+C601, U+C77C, U+C774, U+C0BC, U+C0AC, U+C624, U+C721, U+CE60, U+D314, U+AD6C	10	. U+002E	The repeating scheme is identical to other base 10 lists mentioned in this numeric section.

Hanja-num

To be added to section 4.3 Numeric

System	Characters	Codepoints	Base	Suffix	Notes
<i>hanja-num</i>	零 一 二 三 四 五 六 七 八 九	U+96F6, U+4E00, U+4E8C, U+4E09, U+56DB, U+4E94, U+516D, U+4E03, U+516B, U+4E5D	10	. U+002E	The repeating scheme is identical to other base 10 lists mentioned in this numeric section.

CjkFull-num

To be added to section 4.3 Numeric

System	Characters	Codepoints	Base	Suffix	Notes
--------	------------	------------	------	--------	-------

<i>eastasiafull-num</i>	0 1 2 3 4 5 6 7 8 9	U+FF10, U+FF11, U+FF12, U+FF13, U+FF14, U+FF15, U+FF16, U+FF17, U+FF18, U+FF19	10	. U+002E	The repeating scheme is identical to other base 10 lists mentioned in this numeric section.
-------------------------	------------------------	---	----	-------------	---

Katakanahalf

To be added to 4.4 Alphabetic.

System	Characters	Codepoints	Base	Suffix	Notes
<i>katanakahalf</i>	アイエウオカキクケ コサシスセソタチツ テトナニヌネノハヒフ ヘホマミムメモヤユ ヨラリルレロワワン	U+FF71, U+FF72, U+FF73, U+FF74, U+FF75, U+FF76, U+FF77, U+FF78, U+FF79, U+FF7A, U+FF7B, U+FF7C, U+FF7D, U+FF7E, U+FF7F, U+FF80, U+FF81, U+FF82, U+FF83, U+FF84, U+FF85, U+FF86, U+FF87, U+FF88, U+FF89, U+FF8A, U+FF8B, U+FF8C, U+FF8D, U+FF8E, U+FF8F, U+FF90, U+FF91, U+FF92, U+FF93, U+FF94, U+FF95, U+FF96, U+FF97, U+FF98, U+FF99, U+FF9A,	46	. U+002E	Half width variation of Japanese katanaka, that already exists in section 4.4.

		U+FF9B, U+FF9C, U+FF66, U+FF9D			
--	--	-----------------------------------	--	--	--

Katakanahalf-iroha

To be added to 4.4 Alphabetic.

System	Characters	Codepoints	Base	Suffix	Notes
<i>katanakahalf-iroha</i>	イロハニホヘトチリ ヌルヲワカヨタレソ ツネナラムウヰノ オクヤマケフコエ テアサキユメシ エヒモセスン	U+FF68, U+FF9B, U+FF8A, U+FF86, U+FF8E, U+FF8D, U+FF84, U+FF81, U+FF98, U+FF87, U+FF99, U+FF66, U+FF9C, U+FF76, U+FF96, U+FF80, U+FF9A, U+FF7F, U+FF82, U+FF88, U+FF85, U+FF97, U+FF91, U+FF73, U+30F0, U+FF89, U+FF75, U+FF78, U+FF94, U+FF8F,	48	. U+002E	Half width variation of Japanese katakana-iroha, that already exists in section 4.4.

		U+FF79, U+FF8C, U+FF7A, U+FF74, U+FF83, U+FF71, U+FF7B, U+FF77, U+FF95, U+FF92, U+FF90, U+FF7C, U+30F1, U+FF8B, U+FF93, U+FF7E, U+FF7D, U+FF9D			
--	--	---	--	--	--

Upperfull-alpha

System	Characters	Codepoints	Base	Suffix	Notes
<i>upperfull-alpha</i>	ABCDEFGH IJKLMNOP QRSTUVW XYZ	U+FF21, U+FF22, U+FF23, U+FF24, U+FF25, U+FF26, U+FF27, U+FF28, U+FF29, U+FF2A, U+FF2B, U+FF2C, U+FF2D, U+FF2E, U+FF2F,	26	. U+002E	Full width variation of Upper-alpha, that already exists in section 4.4.

		U+FF30, U+FF31, U+FF32, U+FF33, U+FF34, U+FF35, U+FF36, U+FF37, U+FF38, U+FF39, U+FF3A			
--	--	--	--	--	--