2000-03-15

ISO/IEC JTC1/SC2/WG2 Universal Multiple-Octet Coded Character Set (UCS) Secretariat: ANSI

Title: Rationale for non-Kanji characters proposed by JCS committee

Doc. Type: national body contribution

Source: Japan

Project:

Status: To be discussed at the WG2 meeting in Beijing

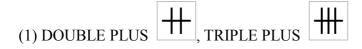
Date: 2000-03-15

Distribution: ISO/IEC JTC1/SC2/WG2

Reference:

As mentioned in the IRG N690, UTC commented on the Japanese requirements to add characters of JIS X0213. This document shows some evidences and explains usages.

1. Rationale and Evidence of usage



These symbols are used, to represent some "strength" or "levels," in discipline including clinical medicine. They represent twice and treble levels, respectively, of the level the PLUS SIGN does.

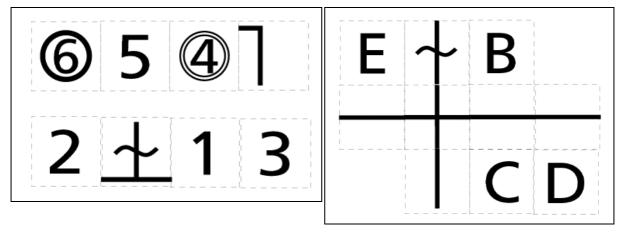
(2) Dentist's symbols

These symbols are used in dentistry when drawing XXX together with some BOX DRAWING characters. The proposal includes two types of characters; those used in single-line drawing and those in triple-line drawing.

	Symbols to be used in single-line drawing:		ΦΦ	$\overline{\Lambda}$	7		
--	--	--	----	----------------------	---	--	--

Symbols to be used in triple-line drawing:





Example single line drawing

Example triple-line drawing

(3) DOUBLE HYPHEN

Recommended in KOUSEI-HIKKEI and Japan Book Publishers Association. Historically, this symbol was derived from duplex form of two HYPHENs or EN DASHes. It is used in katakana-written foreign names to indicate a break between surname and given name, to preserve a hyphen in alphabet-written form of a compound surname, etc. It has a totally different semantics than EQUALS SIGN (U+003D.)

統領
シャトーブリアン Chateaubriand, François I (1768-1848)(仏); 文学者,政治家
シャネル Chanel, Coco (1883-1971)(仏); デナジャンヌーダルク Saint Joan of Arc (仏) Jean (1412? 1431)(仏); 愛国者・聖女シャンポリオン Champollion, Jean François 1832)(仏); エジナド学者シュアン=ツァン 玄奘〉 Xuanzang ((w) tsang) (600? - 664)(中国); 僧しゅうおんらい(周恩来)・チョウ=エンライシューマン Schumann, Robert (1810-56)(東家しゅき(朱熹)・チュー=シー・ファン (1825-99)(オーストリア); 作曲家 ② Strauss, Johann (Wa (1825-99)(オーストリア); 作曲家 ③ Richard (1864-1949)(独); 作曲家 ③ Richard (1864-1949)(独); 作曲家 ③ zバイツァー Schweitzer, Albert (仏); 神学者,哲学者,医者,音楽家シュバリエ Chevalier, Maurice (1888-1972) 手、映画俳優シュペングラー Spengler, Oswald (1880-19哲学者シュミット Schmidt, Helmut (1918-)(独)シュムペーター Schumpeter, Joseph (1883-

(4) LEFT WHITE PARENTHESIS





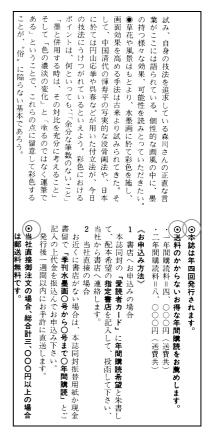
Recommended in KOUSEI-HIKKEI and Japan Book Publishers Association.

nigori 濁り 1 muddiness; impurity. 2
[濁点] a voiced-sound mark. ~を打つ
put a voiced-sound mark (alongside a
'kana' letter))
||濁り水 muddy water.
nigoru 濁る 1 [水などが] become muddy
[impure, cloudy (飲物が)). 2 [音が] have
a flat sound.
nigoshirae 荷拵え packing. — 荷拵え
する pack (up).
nigosu 濁す make (water) muddy
[turbid]. ¶ことばを~ say an ambiguous
thing. 返事を~ give a vague (noncommittal) answer.

(5) CIRCLED BULLET

Recommended in KOUSEI-HIKKEI and Japan Book Publishers Association.

In the example shown rightward, FISHEYE (U+25C9) is used in the upper column and CIRCLED BULLET is used lower column. It can be seen that these characters are used distinguished.



(6) DOUBLE ASTERISK



Recommended in KOUSEI-HIKKEI and Japan Book Publishers Association.

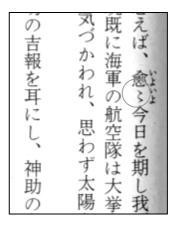
(*A, a [ei] n. (pl. A's or As, a's or as) 1 英語アルファベットの第1字 (母音字). 2 A, a 字によって表される音: [ei] (ate), [oi] (fall), [ei] (cat), [oi] (allow). 3 A字形のもの. 4 [数] 第一の既知数 [量]. from A to B どこからどこへ (行くの)でも. (from) A to Z [or izzard] 全部; 完全に (completely): I know baseball from A to Z. 私は野球のことならすみからすみまで知っている. 中 A - Z not know A from B ものを知らない。学がない、(be igno.)

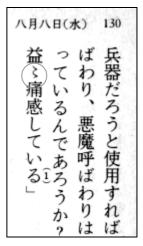
(7) ITERATION MARK



ITERATION MARK is used to indicate to repeat one character.

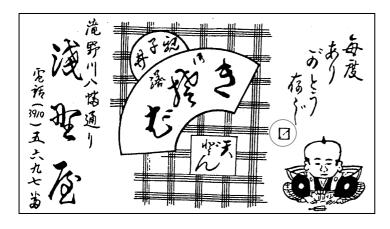
This character is recommended in KOUSEI-HIKKEI and Japan Book Publishers Association.





(8) MASU MARK

This symbols is derived from a pictogram for a 桝 (pronounced as "MASU"; a traditional square measure-cup made with woods.) It is used [in informal contexts] as an abbreviation for a Japanese word "MASU" [which frequently appears at the end of a sentence.]



(9) KATAKANA DIGRAPH KOTO

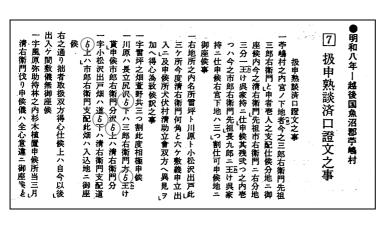


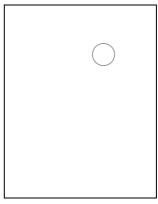
This is a digraph of KATAKANA KO followed vertically by KATAKANA TO. It is actually used in several used law texts.

(10) HIRAGANA DIGRAPH YORI



This is a digraph of HIRAGANA YO followed vertically by HIRAGANA RI. It is used in newspapers.

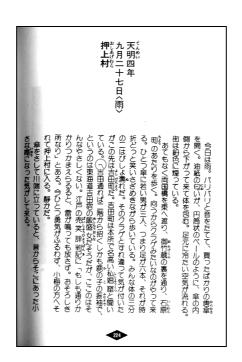




(11) PART-ALTERNATION MARK



兵衛は箒のさきに樽屋橋、黒船獄門は新町橋で、立引きしたのが出入の湊橋。田橋、浦里禿は緑橋、忠兵衛故郷は大和橋、羅城門には渡辺橋、秀吉幼名が日吉橋、後藤又田橋、浦里禿は緑橋、忠兵衛故郷は大和橋、羅城門には渡辺橋、秀吉幼名が日吉橋、後藤又李稙の御供は天満橋、梅王松王は桜橋、鹿ノ子の振袖シャンと着て、二階から招くは吉



(12) WHITE SESAME DOT



, SESAME DOT



くらましている感じがする。

These characters are adopted in the ISO/IEC 1541-1 DAM.3 and IOS/IEC 9541-2 DAM.1. See the documents for explanation about these characters.

る。しかし、一般に女真文字は、この例ほど明確でないものが多く、なんとなく漢字との類似性を放意的に使用しながら、点をくわえて区別している。また「尽」 \rightarrow abka 〈天〉も、天に加画したものであえて区別している。また「尽」 \rightarrow abka 〈天〉も、天に加画したものであるものは、「乱」 \rightarrow inenggi 〈日〉、「魚」 \rightarrow biya 〈目〉のように、明らかに漢字を表しくは加画によってつくりだしたものもある。この種のものとしては女真文字があげられる。すでに述しくは加画によってつくりだい。漢字を変造というよりも、漢字の簡略化、も【10】 意識的に国字として新しい文字をつくりながら、漢字を変造というよりも、漢字の簡略化、も

る。もともと、和語に漢字というごろもは必要ではなかったからである。れば、和語から漢字をのぞくのは、熱帯地方の住民に、モーニングの着用を禁ずるようなものであい。もっとも、なかには、それでも、かぜひとつひかない、じょうぶな漢語もある。それにくらべ、襲語から漢字をうばうのは、極北の民族に、けがわのコートをぬげというようなものかもしれな

2. Sources

Some characters shown above are recommended in KOUSEI-HIKKEI and HENSYUU-HIKKEI. These are widely referenced in Japan. Hence, proposed characters taken from them are regarded as actually being used in information interchange.

Kyodo News Services' K-JIS: Kyodo News Services' proprietary character codeset for news delivery. It is used when news or other information is delivered to news publishers in Japan. Based on JIS X 0208, it includes additional symbols as proprietary extension.

3. Pre-Composed characters

3.1 Enclosed numbers

As UTC has never proposed concrete composition methods for additional circled numerals, we considered that we cannot represent those characters by composition in UCS/Unicode. So, we keep proposing to assign independent code points for those proposed characters.

3.2 KANA extensions

Kana extension for Ainu is an integral part of the writing system used by aborigines in Japan. Such an extension has already been done for many other scripts in UCS, so the proposed addition must be reasonable.

Also, characters in this group represent independent sounds from their _base_ characters, so it is inappropriate to implement using combining methods for accented Latin characters.

3.3 RISING SYMBOL, FALLING SYMBOL

divLの審議の過程で、豊島先生から以下のコメントが寄せられている。

「Unicode 2.0 book の該当部分は、誤解に基づく記述だと思います。IPA の意図は、pitch + contour で high + rising -> high-rising 等を表記しようというものなのに、Unicode(2.0) は、low + high == rising だと誤解しています。これでは、high-rising と low-rising を区別する事が出来ず、こうした区別を持つ言語(e.g. 広東語)の声調表記が不可能になります。」

In Unicode book 2.0, these characters can be produced with another two characters such as low

But according to linguist's comment in Japan, it is misunderstanding. IPA aims (pitch) + (pitch) => (rising)

--- end of document ---