

ISO/IEC JTC1/SC2/WG2
Coded Character Set
Secretariat: Japan (JISC)

Doc. Type: Draft disposition of comments

Title: Draft disposition of comments on PDAM1.3 to ISO/IEC 10646 5th edition

Source: Michel Suignard (project editor)

Project: JTC1 02.10646.00.01.00.05

Status: For review by WG2

Date: 2017-09-11

Distribution: WG2

Reference: SC2 N4546

Medium: Paper, PDF file

Comments were received from the following members: China, Ireland, UK, and USA. The following document is the disposition of those comments. The disposition is organized per country.

Note – With some minor exceptions, the full content of the ballot comments has been included in this document to facilitate the reading. The dispositions are inserted in between these comments and are marked in **Underlined Bold Serif text**, with explanatory text in *italicized serif*.

China: Positive with comments

General comments:

G1. Pdam 1.3

China is in favor of ISO/IEC JTC 1/SC 2 N 4529 (ISO/IEC 10646 (Ed.5)/PDAM 1.3).

Noted

G2. Zanabazar Square and Soyombo

China does not comment on Zanabazar Square and Soyombo because we need further study.

Noted

It will still be possible to comment on the next phase of this amendment (DAM).

and repertoire of several orthographies is at skytower.org/~ernstjtreml/downloadableKurdishFiles/Different_Kurdish_Scripts_Comparison.pdf

Курдский (курманджи) может дополняться гласным *ʷ* (перевернутое *ʷ*).

Турецкий: окончания *-ʷʷ ~ -ʷʷʷ*, *-ʷʷʷ ~ -ʷʷʷʷ*,
-ʷʷʷ ~ -ʷʷʷʷ, *-ʷʷʷ ~ -ʷʷʷʷ ~ -ʷʷʷ ~ -ʷʷʷ*.

Fig. T1.3. Example of TURNED AYB used for Kurdish from Н. В. Юшманов. 1941. Опреде-литель языков ('Language identification guide'). Москва & Ленинград: Издательство Академии Наук СССР. The text reads "Kurdish (Kurmanji) can be supplemented by the vowel *ʷ* (inverted *ʷ*)". The second paragraph refers to Armenian script used for Turkish word endings.

Proposed acceptance

T2. Page 114, Row 18B0: Khitan Small Script

Ireland requests that the invented modern radicals which have no historical attestation, and the annotations that attested Khitan Small Script characters can also be used as radicals, be removed from the PDAM, with the remaining characters shifted up to remove holes. Khitan Small Script characters have specific clustering behaviour, and users and implementers could be confused if characters which should not cluster be encoded in various places within the code table. Moreover, we judge that the argument is strong for a unified radical repertoire which will serve the closely related Khitan Small Script, Khitan Large Script, and Jurchen. Basic attested Khitan Small Script characters, however, are not controversial. We attach to these comments a code table and names list that accords with our recommendations and those in our T3

(see document SC2/N4546 for these code charts and names lists)

Proposed acceptance in principle

See also comment T4 and E5 from UK and TE4 from US. Given that amendment 1 is likely to go in enquiry phase (DAM) it seems prudent, while accepting these comments, to move Khitan Small Script in Amendment 2 to give experts more time to refine the model. This new names list also removes the annotation for 8 characters as requested by UK (comment E5).

T3. Page 115, Row 18B0: Khitan Small Script

The character 18CFD KHITAN SMALL SCRIPT ITERATION MARK acts like any other ordinary Small Khitan character and participates in cluster formation, so separating it from the rest of the range in the code table is unnecessary. It could be ordered as a member of the Radical-13 range, but we think it makes more sense to encode it at the beginning of the block at 18B00.

Propose Partially accepted

See also comment T6 from UK. Typically, modern software has no dependency on code point position or arrangement for a given character. If the character is moved it may make more sense to move it to another block: Ideographic Symbols and Punctuation (16FEF-16FF) which already contains two Asian iteration marks (Tangut and Nushu). Code point position could be 16FE2. If Khitan Small Script is moved in Amendment 2, this character should also be moved to Amendment 2.

T4. Page 138, Row 1F30: Miscellaneous Symbols and Pictographs, 1F3B1 BILLIARDS

Ireland objects strenuously to the glyph change which has been accepted for 1F3B1 BILLIARDS. The source for this character in both SoftBank and KDDI is, properly *ビリヤード*, *biriyādo*, that is 'billiards'. The reference glyphs for both of those is a green game table with coloured balls, with a cue in the SoftBank glyph. It has always been supposed that this character could be used for pocketed or pocketless billiards, snooker, pool, carom, and other cue games. No indication of the "eight ball" is given in the Japanese sources or the original reference glyph, and indeed, numbered balls are not used in many cue sports, like snooker and carom.

e-80E U+1F49B proposed	BILLIARDS	【ビリヤード】	#470 ビリヤード 「biriyaado」 U+EADD SJIS-F3B1 JIS-7A33	#159 #old391 ビリヤード 「biriyaado」 U+E42C SJIS-FB6C	U+FE80E
------------------------------	-----------	---------	---	--	---------

We understand that some (but not all) vendors have represented this character with an eightball, and in fact that the CLDR short name for this character is now “pool 8 ball”. Billiards is not pool. We believe that this constitutes an unauthorized change by vendors of the semantics of this character. Their glyph and the CLDR short name do not represent “billiards” but rather the “Magic 8 Ball” oracle toy, popular especially in the United States, where it seems to have been introduced in 1950, though a “magic ball” was used in a 1940 Three Stooges short called “You Nazty Spy”. Ireland believes that such a reassignment of a UCS character’s semantics by vendors is not acceptable. If a new symbol character is required, a new character should be proposed. We would not have objected to the encoding of a new character 🎱 EIGHT BALL. We also object to the editor’s disposition of comments on the Irish comment for PDAM 1.2:

“A clear majority of the implementations represents this character with an eight ball. While the original glyphs of the historic sources are a good hint, they do not create an absolute reference concerning the glyph appearance, and more so given the vast growth of the Emoji concept beyond its Japanese root.”

This does not address the issue. Characters in the UCS, including symbol characters, have specific semantics. The Japanese legacy characters were encoded with due consideration given to their precise semantics, and just because vendors think that it is expedient to change the meaning of a character in order to save time in encoding, this is not the way the Universal Character Set works. The game of billiards, ビリヤード, biriyādo, is not the same as a Magic 8 Ball. The UTC has acted improperly in not protecting the semantics and base glyph for this character and encouraging its members to respect the International Standard. This causes, in our view, an unnecessary and unfavourable tension between the two committees which maintain the one character set. Since vendor action has resulted in the effective removal of a character from the UCS, Ireland requests the addition of a new character, 🎱 BILLIARD GAMES, on the next available PDAM, at 1F93F in the Supplemental Symbols and Pictographs block.

1F93F 

WG2 discussion

See also comment T7 from UK.

It seems reasonable to entertain the possibility of encoding a new symbol, whether it should be Emoji or not is open to debate.

T5. Page 162, Row 1F90: Supplemental Symbols and Pictographs

In the event that our requests in E8 are not satisfied, Ireland requests the addition of three new characters, 🦌 DEER IN PROFILE, 🦍 GORILLA IN PROFILE, and 🦏 RHINOCEROS IN PROFILE, on the next available PDAM, in the Supplemental Symbols and Pictographs block. A note about the relations between ANIMAL, ANIMAL FACE, and ANIMAL IN PROFILE should be made clear in the names list and in the Principles and Procedures document. Specific expectations about glyph presentation of this class of symbols are being instilled in users via emoji, and this should be taken seriously and formally in the International Standard.

1F9C9 

1F9CA 

1F9CB 

WG2 discussion

It would preferable to accept E8. It should also be noted that in the platform/font used to create this comment, the deer, gorilla and rhinoceros appear as profile in their current proposed code positions.

T6. Page 162, Row 1F90: Supplemental Symbols and Pictographs

In our comments to PDAM 1.2, Ireland objected to the encoding of 1F995 SAUROPOD and 1F996 T-REX as presented there. We understand that these have already been published by the Unicode Consortium—without National Bodies having ever even seen a proposal for either—and that they are, with their names, an evident *fait accompli*. However, in terms of the processes which SC2 and the UTC make use of as we jointly administer a single character set, Ireland must point out that SC2 National Bodies never entered into an agreement with the UTC that we would rubberstamp everything—or anything—proposed by the UTC. In particular, we did not do so when the first large group of emoji were proposed for the standard on the basis of compatibility with existing Japanese character sets. SC2 treated the emoji repertoire as what they are—a collection of symbols—and National Bodies insisted on additions, deletions, and alterations. This application of SC2 National Body expertise was accepted by the UTC as a reasonable response to adventitious repertoire additions without a clear prior-usage rationale. The resulting additions, deletions, and alterations improved the UCS and were the product of effective and collegial collaboration between our two committees. Such collaboration must be encouraged. Currently the mood is more adversarial, to no good purpose.

Now that the Consortium has adopted special procedures for the invention of new symbol characters, it is clear that SC2 oversight is even more important than it has been in the past. The UTC has been permitting vendors to alter the semantics of symbols, rather than encouraging a uniform treatment of them (as with BILLIARDS, or with ANIMAL vs ANIMAL FACE mentioned above); the UTC has devised inconsistent methods for indicating gender and coloration, and has been devising symbol characters with little real rationale in terms of completeness of sets or even in terms of prediction of genuinely likely utility (given the way that text-frequency citations are prioritized over more semantic or metaphorical concerns. In our view, SC2's oversight and input is necessary in order to ensure the Emoji Subcommittee and the UTC do a better job than they have been doing. It should be pointed out that the guidelines on the Unicode website "Submitting Emoji Proposals" provides input to the Emoji Subcommittee. That Subcommittee is composed of experts at large, and has nothing to do with experts appointed by National Bodies to SC2 and SC2/WG2, and the procedures of that Subcommittee have neither been reviewed nor adopted by SC2. We are not bound by them. We have, in the past, augmented Emoji proposals by additions, deletions, and alterations, adding new symbol characters via ballot comments. This procedure was satisfactory and beneficial in the past. We apply it here per that precedent. We expect serious discussion and contributions from the UTC should they feel that this balance between our two committees be done away with. Accordingly, with regard to SAUROPOD and T-REX, we object that their encoding is a woefully inadequate representation of a class of animals that is both important and extremely popular.

Our extensive ballot comments on dinosaur symbols from PDAM 1.2 have discussed this previously and remain valid. We have repeated many of them here, as they seem to us to have been dismissed without due consideration, due to the imminent publication of Unicode 10.0 at the time PDAM 1.2 was balloted. To put it frankly, if the UTC is going to invent symbols and encode them, there must be a price to pay. That price is simple: accept reasonable additions, deletions, and alterations of the things which they propose. We do not insist or expect that the emoji property will be given to any of the symbols we propose. We do think that the UTC have invented a drastically incomplete set of symbols, without any rationale as to its incompleteness and so, to receive our support, the repertoire must be expanded somewhat to be complete. (In his disposition of comments to PDAM 1.2, the editor suggested that "Both characters are to be used in an Emoji context, not as part of a Dinosaur taxonomy". This cannot be accepted. All characters in the UCS are available for use by everyone, and all symbol characters which are valid symbol characters whether or not the UTC applies the emoji property to some of them.)

A group like Sauropod and an individual species like Tyrannosaurus rex do not form anything like a coherent group that epitomizes "dinosaur". Many millions of people admire dinosaurs, and it's quite common to find that people have had one or more favourite dinosaurs from childhood. "Where's my Ticeratops?" "Why isn't there an Iguanodon?" will surely be some of the first reactions to the standardization of only two pictographs in this block. On the other hand, a properly complete set will certainly be very popular indeed.

Evidently some very basic proposals had been made to encode some “dinosaurs” as emoji, but singling out two simply makes no sense. The UCS includes many mammal symbols, and work seems to be ongoing to identify a larger and larger set of them, based on evident familiarity, metaphor, and desirability indicating some expected use. We understand that many successful emoji proposals, at least in part, have not been based on systematic analysis or even on internet discussions about missing emoji, but rather on Instagram and Google Trends data based on word frequency. This does not seem to be an entirely sufficient criterion, particularly as emoji are often used metaphorically, and outside of metaphor words may be used for all sorts of ordinary reasons. It may be useful to note that the word “cricket” is probably far, far more commonly used on the internet for the sport than for the insect. Do many people use the word “Sauropod” in speech? Quite likely they do not, but the class of “dinosaurs” is comprised of a number of very familiar groups, and, in our view, a relatively small number of encoded pictographs would suffice to represent those groups.

The set of existing UCS emoji symbols for the kingdom Animalia is not very well balanced. This is nobody’s fault. The set began with animals implemented in late-1990s Japanese telecom sets. The set was augmented by German and Irish National Body comments adding more animals for, for example, a complete set of characters used in the Asian Zodiac. (That is why there is a crocodile encoded, for instance.) Since then some more animals have been added. The current UCS (including the content of this PDAM) has, in the emoji classification:

- 1 amphibian
- 12 birds
- 10 “bugs” (8 arthropods, 1 mollusc, and 1 architectural device made out of a proteinaceous extrusion)
- 48 mammals (47 mammals and 1 pair of paw prints)
- 12 “marine” creatures (including 4 fish, 3 mammals, 2 arthropods, 2 molluscs, and 1 mollusc shell)
- 8 reptiles (including 2 dragons and 2 dinosaurs).













There are, according to *Mammal Species of the World*, 5,416 species of mammals identified in 2006. These were grouped into 1,229 genera, 153 families and 29 orders. While it is likely that more mammal pictographs could be added to the UCS, 47 isn’t a bad start. It’s unlikely that symbols for 1200 genera would be needed. For dinosaurs, however, the number of genera is much smaller (about 500) and there too, it is unlikely that a great many symbols would be needed. But given their popularity, Ireland believes that certainly more than two is necessary.

Described as “dragons” in the Western Jin Dynasty, dinosaurs have fascinated our culture for a very long time. Modern study of dinosaurs has done so no less, and popular culture is permeated by them: noteworthy are Jules Verne’s 1864 *Journey to the Centre of the Earth* (Ichthyosaurs, Plesiosaurs; Dimetrodon was in the 1959 film of this book); Arthur Conan Doyle’s 1912 *The Lost World* (Ichthyosaurs, Iguanodon, Plesiosaurs, Pterosaurs, Sauropods, Stegosaurians, some carnivorous Therapods); the 1933 film *King Kong* (Ceratopsians, Plesiosaurs, Pterosaurs, Sauropods, Stegosaurians, Tyrannosaurids); many others, until more modern scientific findings about dinosaurs found their way into Michael Crichton’s 1990 novel *Jurassic Park* and the films that were based on it. Some non-dinosaur characters also have high visibility in popular culture. Two of these, the MAMMOTH and DODO, are commonly used as metaphors in ordinary phrases: “a mammoth sale”, “as dead as a dodo”. (This is a lot more than can be said about BROCCOLI.)

The character names given below are chosen from the standard scientific taxonomy, and so the most identifiable species in each class of dinosaur are reflected with accurate nomenclature. Thus there are some genera, families, superfamilies, suborders, orders, and clades represented. Informative notes assist the users in knowing what groupings the symbol characters cover. Because the characters proposed here represent the most iconic and popularly identifiable dinosaurs, we would not expect further requests to encode additional ones. Ireland believes that a set of 18 symbols (including the 2 on this PDAM) representing dinosaurs and some other prehistoric creatures would do well to fill in the gaps implied by SAUROPOD and T-REX. We believe that encoding only those 2 characters at this time would simply lead to calls to fill the gaps, and we think that we have done filled those gaps here. A species-based nomenclature would be possible but less advantageous. If TRICERATOPS were encoded, the glyph should really not be of a Protoceratops or Styracosaurus. CERATOPSIAN gives glyph designers more choice.




Ireland requests the addition of the following characters on the next on the next available PDAM, in the Supplemental Symbols and Pictographs bloc. We suggest the following code positions:

Dinosaurs and other prehistoric reptiles¹

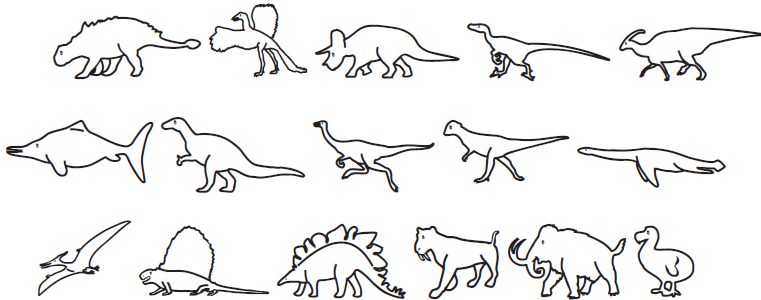
	1F9A1	ANKYLOSAUR	<ul style="list-style-type: none"> • a suborder of the order Ornithischia
	1F9A2	ARCHAEOPTERYX	<ul style="list-style-type: none"> • a genus of the suborder Theropoda of the order Saurischia
	1F9A3	CERATOPSIAN	<ul style="list-style-type: none"> • a suborder of the order Ornithischia • includes Protoceratops, Styracosaurus, Triceratops
	1F9A4	DROMAEOSAURID	<ul style="list-style-type: none"> • a family of the suborder Theropoda of the order Saurischia • includes Deinonychus, Utahraptor, Velociraptor
	1F9A5	HARDROSAURID	<ul style="list-style-type: none"> • a family of the suborder Ornithopoda of the order Ornithischia • includes Edmontosaurus, Hadrosaurus, Parasaurolophus
	1F9A6	ICHTHYOSAUR	<ul style="list-style-type: none"> • a member of the order Ichthyosauria
	1F9A7	IGUANODON	<ul style="list-style-type: none"> • a genus of the suborder Ornithopoda of the order Ornithischia
	1F9A8	ORNITHOMIMID	<ul style="list-style-type: none"> • a family of the suborder Theropoda of the order Saurischia • includes Gallimimus, Ornithomimus, Struthiomimus
	1F9A9	PACHYCEPHALOSAUR	<ul style="list-style-type: none"> • a family of the suborder Pachycephalosauria of the order Ornithischia • includes Pachycephalosaurus, Stegoceras
	1F9AA	PLESIOSAUR	<ul style="list-style-type: none"> • a member of the order Plesiosauria • includes Elasmosaurus, Liopleurodon, Plesiosaurus • represents Nessie, the Loch Ness Monster
	1F9AB	PTEROSAUR	<ul style="list-style-type: none"> • a member of the order Pterosauria • includes Pteranodon, Pterodactylus, Quetzalcoatlus
	1F9AC	SPHENACODONTID	<ul style="list-style-type: none"> • a genus of the family Sphenacodontidae • includes Ctenospondylus, Dimetrodon, Secodontosaurus, Sphenacodon
	1F9AD	STEGOSAURIAN	<ul style="list-style-type: none"> • a suborder of the order Ornithischia • includes Huayangosaurus, Kentrosaurus, Stegosaurus

¹ Dinosaurs are animals which belong to the orders Ornithischia and Saurischia. Ichthyosaurs, Plesiosaurs, and Sphenacodontids are not taxonomically classified as dinosaurs, and in places the term “dinosaurs” has been used loosely here to refer to dinosaurs and other prehistoric reptiles, as indicated by the informative header to the names list.

Extinct creatures

-  1F9AE SABRE-TOOTHED CAT
= Smilodon
-  1F9AF MAMMOTH
 - indicates great size
-  1F9CC DODO
 - indicates extinction

Examples of the glyphs at a larger size:



WG2 discussion


See also comment T10 from UK.

This comment conflates two comments made by Ireland in the ballot concerning PDAM1.2 (see document N4823 comment T4 (pages 10-11) and T9 (page 12-16)). As before it covers JTC1/SC2 and Unicode/UTC synchronization process issues which should be addressed by a separate contribution outside of the ballot comment context. As before, these considerations will be noted and will not be given, again, any further consideration by the project editor in the context of these ballot comment disposition, being out of scope.

The second part concerns the addition of dinosaur symbols. As before the editor will insist that a separate proposal be made to propose the encoding of these additional symbols.

It should also be noted that the list proposed by Ireland differ somewhat from the list proposed by UK in its comment T10. It would be advisable for Ireland and UK to prepare a joint contribution for further discussion. Finally, one of the name in the list from Ireland has a typo: Hardrosaurid should be 'Hadrosaurid'.

T7. Page 162, Row 1F90: Supplemental Symbols and Pictographs

In keeping with our belief that incomplete sets of symbol characters should be improved, Ireland requests the addition of a new Fantasy Being character,  TROLL, to the next available PDAM, in the Supplemental Symbols and Pictographs block. While we recognize that SC2 has no oversight in terms of what gets assigned the emoji property, there is certainly evidence that this particular character has been requested by many for such a purpose.



1F9CF

(COLTER HETTICH/NEW YORK DAILY NEWS)

This isn't the 1990s bejeweled, treasure variety. This is the annoying, instigating, Internet low-life-with-no-life, the keyboard-peddling jerk who spouts sexist, racist or otherwise hateful disparaging nonsense on Twitter or in the comment sections of YouTube videos just to start an argument.



Like the mythical monsters of legend, Internet trolls are an ancient cyberspace concept, but these annoying little gremlins are still out there. Rather than giving them the satisfaction of a back and forth feud they can just brag about under bridges with their troll buddies, it would be much more convenient to stamp them with the label that says just what they are.




WG2 discussion

This again requires a separate contribution.

T8. Page 162, Row 1F90: Supplemental Symbols and Pictographs

In keeping with our belief that incomplete sets of symbol characters should be improved, Ireland requests the addition of two new Animal symbol characters,  BADGER and  SQUIRREL, to the next available PDAM, in the Supplemental Symbols and Pictographs block..

1F9CD 

1F9CE 

WG2 discussion

See also comment T8 from UK.

The UK comment includes these two symbols, plus the SWAN. This again requires a separate contribution.

Editorial comments:

(All accepted glyph changes are based on receiving a font from Ireland with the appropriate glyphs.)

E1. Page 12, Row 059: Hebrew

Ireland requests that the glyph for 05EF be centred.

Accepted

E2. Page 59, Row 260: Miscellaneous Symbols.

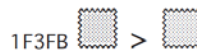
Ireland has reviewed a number of glyphs for this block with regard to vendor glyphs commonly used for them, and recommends the following glyph change to better align the black-and-white chart glyphs with them..

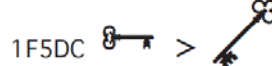


Proposed acceptance

E3. Page 138, Row 1F30: Miscellaneous Symbols and Pictographs

Ireland has reviewed a number of vendor glyphs for this block, and recommends a number of glyph changes to better align the black-and-white chart glyphs with the images that most users will find in a colour-glyph environment. For many of these symbols the rationale for alteration has simply to do with orientation. For some, the recent specific changes relating to gender and coloration have had to be taken into account in order to support sequences even in black-and-white fonts. Some glyphs, like those for the Fitzpatrick swatches, have simply been cleaned up and made clearer.



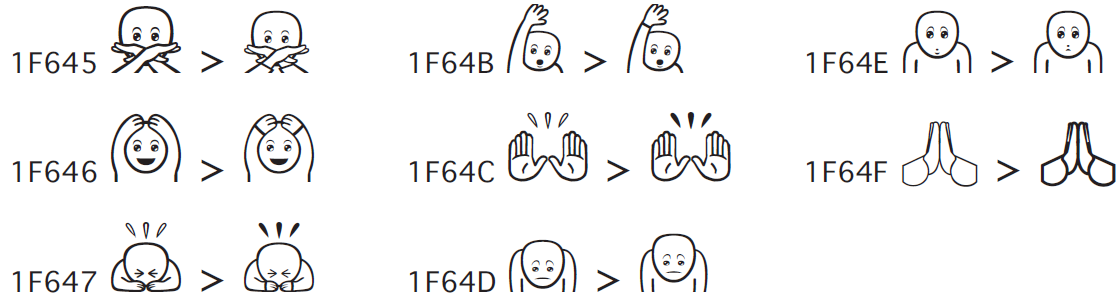


Propose partially accepted

The project editor will work with Irish experts and other interested parties to determine which of these glyphs changes should be propagated and in which amendment.

E4. Page 153, Row 1F60: Emoticons

Ireland has reviewed a number of vendor glyphs for this block, and recommends a number of glyph changes to facilitate a support of the recent specific changes relating to gender and coloration to sequences making use of these symbols. This includes adding sleeves and making other small adjustments.

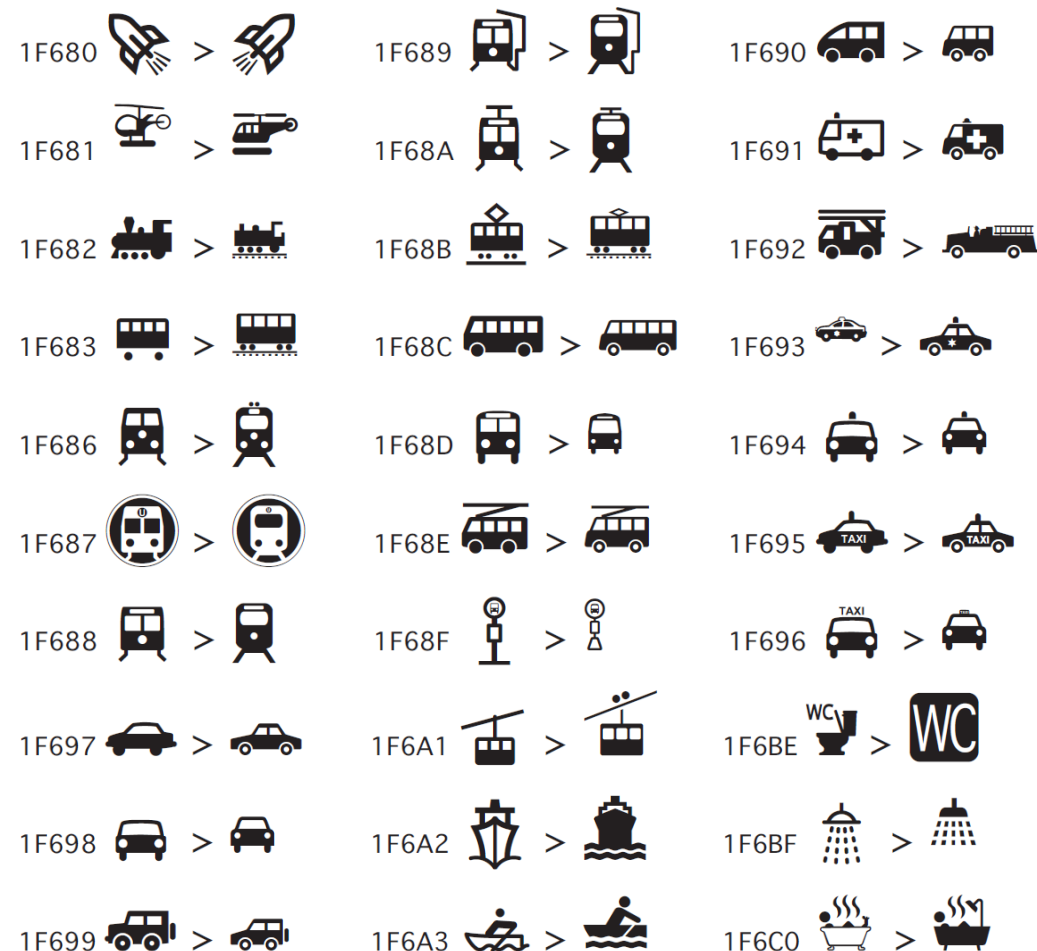


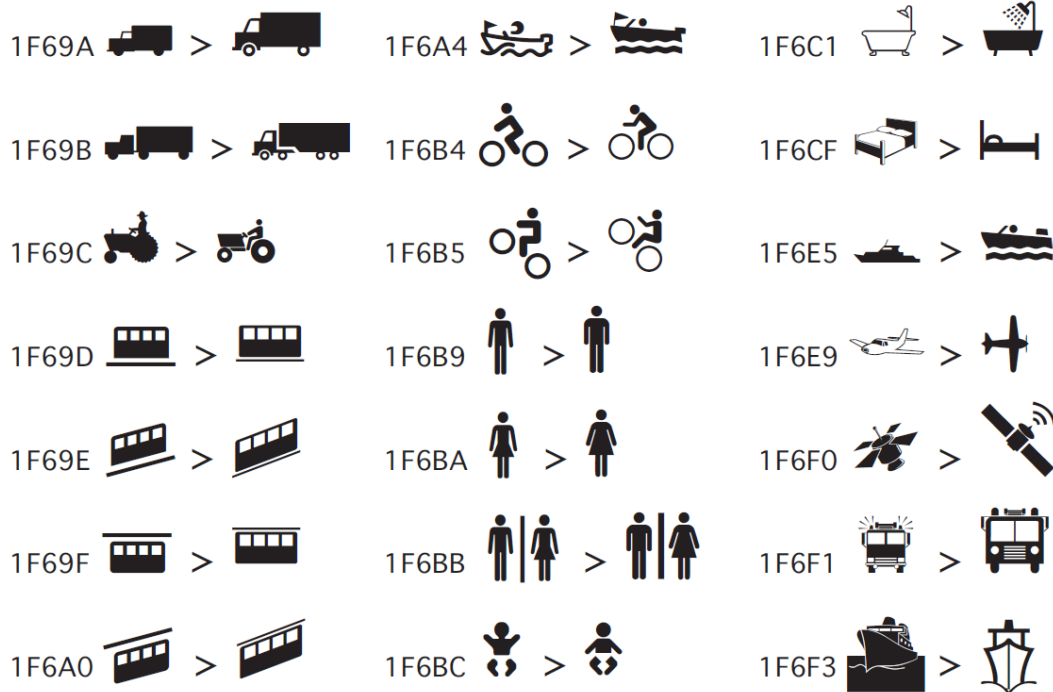
Propose Partially Accepted

Adding sleeves and some of the small adjustments. However, it seems that 1F64C is already as requested, and it is not clear that the emotion traits conveyed in 1F647 (deep bowing) and 1F64C (celebration) should use the same upper part symbol.

E5. Page 156, Row 1F68: Transport and Map Symbols

Ireland has reviewed a number of vendor glyphs for this block, and recommends a number of glyph changes to harmonize better with the more common vendor glyphs for these symbols. For a few of these symbols, the recent specific changes relating to gender and coloration have had to be taken into account in order to support sequences even in black-and-white fonts.



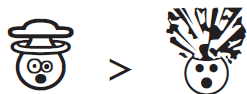


Propose Partially Accepted

Many of the proposed changes are probably acceptable but the project editor would like to work with Irish experts and other interested parties to determine which of these glyphs changes should be propagated and in which amendment.

E6. Page 162, Row 1F90: Supplemental Symbols and Pictographs

Ireland requests that the reference glyph for 1F92F SHOCKED FACE WITH EXPLODING HEAD be based upon the blast of the international warning sign for explosives. We understand this symbol to mean “Mind blown”. Ordinary explosives express this well. We do not believe the UCS needs to include symbols whose glyphs feature “cute” atomic mushroom clouds.



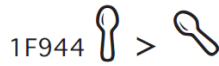
Accepted in principle

However, the sign should be closer to that international warning sign, with more visible black rays like:



E7. Page 162, Row 1F90: Supplemental Symbols and Pictographs

Ireland has reviewed a number of vendor glyphs for this block, and recommends a number of glyph changes to harmonize better with the more common vendor glyphs for these symbols. With regard to 1F94A, many vendors have the vertical orientation, but this looks more like an oven mitt than a boxing glove, and the horizontal orientation “punches” better. A number of the symbols representing human beings have been altered to facilitate the application of symbol sequences expressing gender and coloration (as with similar characters given above).



Propose Partially Accepted

Many of the proposed changes are probably acceptable but the project editor would like to work with Irish experts and other interested parties to determine which of these glyphs changes should be propagated and in which amendment.

E8. Page 162, Row 1F90: Supplemental Symbols and Pictographs

With respect to a number of characters which are shown by many vendors with just the face of the animal, we believe that as symbols, a face only should be shown when the word FACE appears in the character name. Compare 1F42D MOUSE FACE with 1F401 MOUSE.




Such distinctions in symbols is important. In our view, where vendor practice is divided on this issue, a principled decision should be taken for each one, and vendors encouraged to implement uniform representation. Where a name does not contain FACE, as in 1F42B BACTRIAN CAMEL, the whole creature should be depicted (as all vendors do).

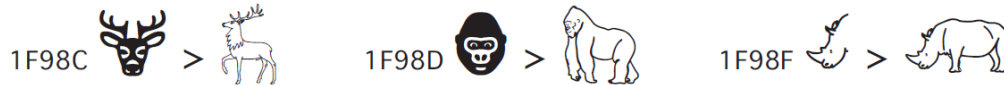


Users of symbols in the UCS are not only users of emoji, and users of both kinds may rightly expect a certain uniformity in presentation of such symbols whether in a monochrome or polychrome font. We believe that where the name is just that of the animal, the whole animal should be shown. Where an animal face is required, the word FACE should be part of the name. Unfortunately, for a number of characters, some vendors are showing a face for the animal where FACE does not appear in the name.

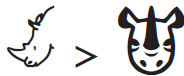


U+1F98D							—			—	—
U+1F98F							—			—	—

This is a clear example of how symbols in the UCS are not being treated uniformly by vendors. Now in the first place, the emoji property is of little consequence. Will emoji continue to be used in ten years' time? Twenty? No one can say. But symbols are symbols, and whether used in black and white or in colour, users have a right to think they are sending a gorilla and not have it turn up as only a gorilla face on the receiving end. Since ANIMAL and ANIMAL FACE are distinguished in the UCS, it cannot be said that the glyph shape of these three is irrelevant. Thus we request the following glyph changes:



If this request is not satisfied, and it is decided to retain the rhinoceros face, then we request the following glyph change to be made:



WG2 discussion

See also comment T9 from UK

The project editor would prefer the first option (changing the three glyphs).

E9. Page 162, Row 1F90: Supplemental Symbols and Pictographs

Ireland recommends the following annotations for two characters on this PDAM.



1F995 SAUROPOD

- a clade of the suborder of Sauropodomorpha of the order Saurischia
- includes Apatosaurus, Brachiosaurus, Brontosaurus, Diplodocus



1F996 T-REX

- * Tyrannosaurid
- a superfamily of the suborder of Theropoda of the order Saurischia
- includes Albertosaurus, Gorgosaurus, Tyrannosaurus (T. rex)

WG2 discussion

This is related to the discussion concerning the comment T6.

We give below a chart of the Supplemental Symbols and Pictographs block with our proposed amendments marked in yellow and our proposed additions marked in blue.

(Ed. code table inserted next page but without the names list, for the names list see SC2 N4546)

	1F90	1F91	1F92	1F93	1F94	1F95	1F96	1F97	1F98	1F99	1F9A	1F9B	1F9C	1F9D	1F9E	1F9F
0	1F900	1F910	1F920	1F930	1F940	1F950	1F960		1F980	1F990			1F9C0	1F9D0	1F9E0	
1	1F901	1F911	1F921	1F931	1F941	1F951	1F961		1F981	1F991	1F9A1			1F9D1	1F9E1	
2	1F902	1F912	1F922	1F932	1F942	1F952	1F962		1F982	1F992	1F9A2			1F9D2	1F9E2	
3	1F903	1F913	1F923	1F933	1F943	1F953	1F963		1F983	1F993	1F9A3			1F9D3	1F9E3	
4	1F904	1F914	1F924	1F934	1F944	1F954	1F964		1F984	1F994	1F9A4			1F9D4	1F9E4	
5	1F905	1F915	1F925	1F935	1F945	1F955	1F965		1F985	1F995	1F9A5			1F9D5	1F9E5	
6	1F906	1F916	1F926	1F936	1F946	1F956	1F966		1F986	1F996	1F9A6			1F9D6	1F9E6	
7	1F907	1F917	1F927	1F937	1F947	1F957	1F967		1F987	1F997	1F9A7			1F9D7		
8	1F908	1F918	1F928	1F938	1F948	1F958	1F968		1F988		1F9A8			1F9D8		
9	1F909	1F919	1F929	1F939	1F949	1F959	1F969		1F989		1F9A9		1F9C9	1F9D9		
A	1F90A	1F91A	1F92A	1F93A	1F94A	1F95A	1F96A		1F98A		1F9AA		1F9CA	1F9DA		
B	1F90B	1F91B	1F92B	1F93B	1F94B	1F95B	1F96B		1F98B		1F9AB		1F9CB	1F9DB		
C		1F91C	1F92C	1F93C	1F94C	1F95C			1F98C		1F9AC		1F9CC	1F9DC		
D		1F91D	1F92D	1F93D		1F95D			1F98D		1F9AD		1F9CD	1F9DD		
E		1F91E	1F92E	1F93E		1F95E			1F98E		1F9AE		1F9CE	1F9DE		
F		1F91F	1F92F	1F93F		1F95F			1F98F		1F9AF		1F9CF	1F9DF		

UK: Negative

General/Technical/Editorial comment (G, T, or E prefix)

G1. Emoji inclusion process

We are very concerned that a large number of emoji characters were added to Amendment 1 at the discretion of the project editor, but our technical comments on some of these characters were ignored, and these characters were fast-tracked into Unicode version 10.0.

In particular, we are very disappointed that our comment to change the name of 1F996 from T-REX to TYRANNOSAURUS REX was not accepted. We note that the project editor's draft disposition of comments accepted our comment, but the final disposition of comments did not accept it, apparently because an officer of the UTC refused to make the necessary changes to the Unicode 10.0 data files. We do not understand why the name of 1F996 could not have been changed in time for the release of Unicode 10.0, especially as the name only occurs in four Unicode data files (DerivedName.txt, Index.txt, NamesList.txt and UnicodeData.txt), and in one Emoji data file (emoji-test.txt where it occurs as an informative comment).

It seems that the UTC considered Unicode 10.0 to be fixed and stable before the PDAM ballot which first introduced these emoji characters to SC2 had even closed, effectively bypassing the ballot process for these characters, and depriving national bodies of the opportunity to make meaningful technical comments..

Proposed change by UK:

Improve communication and cooperation between SC2 and the UTC, so that this situation does not occur again the future.

Impress upon the UTC liaison representative to SC2 that characters should not be fast-tracked into the Unicode Standard unless they have completed at least one round of ballot at Committee stage with no unresolved comments.

Ask the UTC to ensure that in future any characters added to a PDAM ballot are not fast-tracked into the Unicode Standard until the ballot has completed and any ballot comments have been resolved.

Noted

There is clearly tension between the pressure of adding Emoji characters quickly and the need of due process. The project editor is not adding Emoji characters at his own discretion, but instead as a result of the requirement to keep the Unicode Standard and ISO/IEC synchronized. While not officially introduced in 10646 repertoire until pdam 1.2, the name in question was known at least six months in advance within the UTC. The project editor is working at improving the communication between the two committees to decrease that tension for future Emoji contents.

G2. Process concerning the addition of related characters

Over the last few years many arbitrary emoji characters have been added to the standard at the request of the UTC and/or US national body. As ISO/IEC 10646 does not have a formal concept of emoji, these characters are represented as monochrome glyphs in the code charts, and we can consider them as pictographic symbol characters.

We believe it is perfectly appropriate for national bodies to ask in ballot comments for additional related symbol characters to be added to the ballot where the proposed characters provide incomplete or one-sided coverage of the category of symbols they represent. Whether these additional characters are treated as emoji or not is a matter for the UTC and/or vendors to decide.

Proposed change by UK:

Do not dismiss ballot requests to add additional symbol characters related to emoji characters under ballot.

Noted

Request for addition have not been dismissed. However, the editor has repeatedly requested on such instances that the additions requests should be done through separate contributions with their own rationale and justification, not solely through ballot comments. In some rare cases, additions have been done that way, but typically not in a situation where the additions would multiply by 7 times the size of the original repertoire. While ISO/IEC 10646

does not have a formal concept of Emoji, it is well known that Emoji characters have a high implementation cost for vendors, so this needs to be considered as well. Below, the disposition for the request of additional Dinosaurs in pdam1.2 (Document N4823 page 16 is shown:

The two proposed emoji characters are a representation of two popular, if not scientific exact, representation of the historic reptiles: the generic docile dinosaur (named here SAUROPOD) and the generic aggressive version (named here T-REX). Both characters are to be used in an Emoji context, not as part of a Dinosaur taxonomy.

Furthermore, there is no requirement at having a proposal to be comprehensive and the UCS encoding model is additive. It is perfectly possible to create a proposal to create additional symbols dinosaurs and other prehistoric reptiles, but this should be covered by a separate contribution.

E3. Clause 24, Source reference file for Tangut ideographs

Table 6 has not been updated for the new source references used for 187ED..187F1.

Proposed change by UK:

In 24.2 Table 6 add these additional kTGT_MergedSrc formats: (L2012-dddd) and (WG2N4724-d).

Accepted

T4. Clause 33, Khitan Small Script

The radical system for Small Khitan is a modern invention and does not reflect native usage in original Liao and Jin dynasty texts. The inclusion of radical characters that are only used in modern sources in the main block of Khitan characters is not appropriate, and would be confusing to users. As these radical characters should not participate in cluster formation, their inclusion among Khitan characters that are used in cluster formation adds an unnecessary burden on implementers, who would be unable to use a single code point range to determine cluster-forming characters. Some fonts may simply allow all characters in the Khitan Small Script block to be cluster-forming, which would result in unexpected behaviour for end users.

Khitan Small, Khitan Large, and Jurchen are related scripts for which an overlapping set of radicals have been used in modern sources, and it may be better to define a separate block for unified Khitan and Jurchen radicals. As modern radicals are not required for writing Small Khitan text, we request that the twelve characters used only as radicals are removed from the Khitan Small Script block pending further study..

Proposed change by UK:

Remove the following twelve characters, and move the remaining characters up to fill the gaps.

18B69 (Radical-03)

18BD2 (Radical-06)

18C02 (Radical-07)

18C15 (Radical-08)

18C32 (Radical-10)

18C3B (Radical-11)

18C57 (Radical-13)

18C6A (Radical-14)

18C86 (Radical-15)

18C94 (Radical-16)

18C9E (Radical-17)

18CDD (Radical-20).

Propose acceptance in principle

See comment T2 from Ireland, and TE4 from US and disposition in T2 from Ireland.

E5. Clause 33, Khitan Small Script

As it may be appropriate to disunify Khitan radicals from Khitan characters, and encode a unified block for Khitan and Jurchen radicals, the code chart annotation for characters that can also be used as radicals should be removed.

Proposed change by UK:

Remove the annotations for:

18B00 (Radical-01)
18B35 (Radical-02)
18B94 (Radical-04)
18BAD (Radical-05)
18C2B (Radical-09)
18C51 (Radical-12)
18CCA (Radical-18)
18CD6 (Radical-19).

Propose acceptance

See also comment T2 from Ireland.

T6. Clause 33, Khitan Small Script

18CFD acts like an ordinary Khitan character and participates in cluster formation, and it is more convenient for implementers if all cluster-forming characters are in a single code point range. Therefore it is best not leave a gap between it and the other Khitan characters. We suggest moving 18CFD to the start of the block.

Proposed change by UK:

Move 18CFD to 18B00, and move 18B00-18CE0 down.

Proposed Partially accepted

See comment T3 From Ireland and its disposition.

T7. Clause 33, Miscellaneous Symbols and Pictographs, 1F3B1 BILLIARDS

Despite our request in the ballot comments for PDAM 1.2, the UTC has published Unicode 10.0 with a new unacceptable glyph for 1F3B1 BILLIARDS. The new glyph is an “eight ball”, which is only used in the North American game of Pool, and therefore cannot be used to represent a generic billiards game. Moreover, this glyph is not normally used to represent the actual sport of Billiards, but is mostly used on the internet to represent a fortune-telling gimmick known as the “Magic 8-Ball” (see https://en.wikipedia.org/wiki/Magic_8-Ball and <https://xkcd.com/1525/>).

1F3B1 was encoded in order to map to Japanese carrier symbols that show a billiard cue and a set of coloured (unnumbered) billiard balls (<http://www.unicode.org/L2/L2009/09027r2-emoji-backgrnd.pdf>). Allowing vendors to radically change the glyph for this character so that it no longer reflects its original semantics, has subverted the ISO/IEC 10646 and Unicode standards, and has destabilized any existing data that use this character to represent billiards as a sport.

The recent history of emoji encoding demonstrates a strong desire to represent all major international sports with dedicated emoji characters. Snooker is a major international sport, and is a very popular spectator and participant sport in the UK, which we strongly believe should be represented by an appropriate emoji character. We recognize that it is now too late to change the glyph of 1F3B1 back to show a generic billiards game, but to ameliorate the unfortunate situation we have arrived at, we request that:

1. 1F3B1 is given a formal alias of MAGIC 8-BALL.
2. A new character with a glyph showing a billiards cue and a frame of billiards balls, is added to this amendment.

Proposed change by UK:

1. Add the formal alias MAGIC 8-BALL to 1F3B1.

2. Add a new character BILLIARDS CUE AND BALLS to this amendment. The glyph should show a billiards cue and a triangular set of coloured but unnumbered billiards balls. Add the informative alias “snooker” to this character.

WG2 discussion

See also comment T4 from Ireland.

The non-acceptance was based on the following argument (document N4823 page 11):

A clear majority of the implementations represents this character with an eight ball. While the original glyphs of the historic sources are a good hint, they do not create an absolute reference concerning the glyph appearance, and more so given the vast growth of the Emoji concept beyond its Japanese root..

While adding a name alias to the character showing ‘magic 8-ball’ is probably desirable, it is probably not necessary to elevate it to a formal name alias (which have little weight in the context of ISO/IEC 10646).

Adding a character is certainly worth considering, although not in this amendment but rather in Amendment 2.

T8. Clause 33, Supplemental Symbols and Pictographs, additional small mammals

Hedgehog is only one of several iconic and distinctive European small mammals that are not currently encoded as emoji. A 2013 BBC poll to find the national species for Britain

(<http://www.discoverwildlife.com/british-wildlife/britains-national-species-revealed>) chose Hedgehog in first place, Badger in second place, Oak tree in third place, and Red Squirrel in third place. This shows the importance of Hedgehog, Badger and Squirrel to the British public, and we therefore request that characters for all three of these animals be added to the current amendment. Swan is an iconic and distinctive British bird that is required to improve coverage of water fowl, and so we request that Swan is also added to the current amendment.

Proposed change by UK:

Add the following characters to the Supplemental Symbols and Pictographs block:

BADGER

SQUIRREL

SWAN.

WG2 discussion

See also comment T8 from Ireland and its disposition. It was not clear from the UK comment, but HEDGEHOG is already part of this amendment proposed at code point U+1F994.

T9. Clause 33, Supplemental Symbols and Pictographs, GIRAFFE FACE and ZEBRA FACE

The distinctive features of a giraffe are its long legs and long neck, so the character for a giraffe should show the full giraffe body, not just its face, which may not be easily recognized as belonging to a giraffe at small character sizes.

The distinctive feature of a zebra is its striped equine body, so the character for a zebra should show the full zebra body, not just its face, which may not be easily recognized as belonging to a zebra at small character sizes.

The character names GIRAFFE FACE and ZEBRA FACE are inappropriate as character designers may wish to draw characters depicting the whole animal body. This is already the case for the Emojidex designs for Giraffe Face and Zebra Face (see <https://emojipedia.org/giraffe-face/> and <https://emojipedia.org/zebra-face/>).

As a general principle it is best to avoid adding FACE to the names of animal characters unless it is contrastive (e.g. RABBIT and RABBIT FACE), as it should be up to designers to decide whether to depict the whole animal or only its head.

Proposed change by UK:

Add a note to the standard stating that 1F992 (GIRAFFE FACE) and 1F993 (ZEBRA FACE) may be depicted as the head or the whole body.

Partially accepted

See also comments T5 and E8 from Ireland.

If the face is part of the name it is probably better to keep the preferred glyphs as there are. The Irish comment E8 tends to prefer that animal glyphs corresponding to character names with no face in them should be shown as full body, not as just face or fragment like DEER, GORILLA, and RHINOCEROS. It is not clear that variations as shown by the ‘emojidex’ should be encouraged through an annotation.

T10. Clause 33, Supplemental Symbols and Pictographs, T-REX and SAUROPOD

Dinosaurs are by far the most well-known and well-loved group of extinct animals, and encoding dinosaur emoji will fill a welcome gap in the current emoji coverage. However, the two characters, T-Rex and SAUROPOD, only partially fill this gap. The UK has a particular interest in dinosaurs, as dinosaurs were first recognized and described in Great Britain, and fossils of many species of dinosaurs and related prehistoric animals have been discovered in various parts of the UK. We would therefore like to see the standard include a more representative range of characters representing the major groups of dinosaurs and related animals, as well as iconic (and metaphorical) extinct animals such as Mammoth and Dodo.

Proposed change by UK:

Add the following characters to the Supplemental Symbols and Pictographs block:

ANKYLOSAURUS
GALLIMIMUS
IGUANODON
PACHYCEPHALOSAURUS
PARASAUROLOPHUS
STEGOSAURUS
TRICERATOPS
VELOCIRAPTOR
ARCHAEOPTERYX
DIMETRODON
ICHTHYOSAUR
PLESIOSAUR
PTEROSAUR
MAMMOTH
SMILODON
DODO

WG2 discussion

See comment T6 and related disposition from Ireland. The two proposed sets intersect, but names are sometimes different.

T11. Clause 33, Chess Symbols,

The names RED XIANGQI XXX and BLACK XIANGQI XXX are not ideal, as they make it seem as if there are two games called “Red Xiangqi” and “Black Xiangqi”. In fact they are all symbols for “Xiangqi”, and the colour (RED or BLACK) modifies the name of the piece not the name of the game. Therefore XIANGQI RED XXX and XIANGQI BLACK XXX are more appropriate names.

Proposed change by UK:

Rename the characters at 1FA60-1FA6D to:

1FA60 XIANGQI RED GENERAL
1FA61 XIANGQI RED MANDARIN
1FA62 XIANGQI RED ELEPHANT
1FA63 XIANGQI RED HORSE
1FA64 XIANGQI RED CHARIOT
1FA65 XIANGQI RED CANNON

1FA66 XIANGQI RED SOLDIER
1FA67 XIANGQI BLACK GENERAL
1FA68 XIANGQI BLACK MANDARIN
1FA69 XIANGQI BLACK ELEPHANT
1FA6A XIANGQI BLACK HORSE
1FA6B XIANGQI BLACK CHARIOT
1FA6C XIANGQI BLACK CANNON
1FA6D XIANGQI BLACK SOLDIER

Propose Acceptance

E12. Clause 33, Chess Symbols, 1FA65 annotation

1FA65: The annotation “design sometimes shows 5352 卒 (Ming dynasty)” is incorrect, and should refer to 5305 包.

Proposed change by UK:

For 1FA65, change "design sometimes shows 5352 卒 (Ming dynasty)" to "design sometimes shows 5305 包 (Ming dynasty)".

Accepted

E13. Clause 33, Chess Symbols, 1F6AB annotation

1FA6B: alias "hēi ju" is missing a tone mark.

Proposed change by UK:

1FA6B: change alias "hēi ju" to "hēi jū".

WG2 discussion

This has to do with code chart production limits. Any characters beyond Latin-1 repertoire would be shown with a different type face (in fact the code chart version of the character) and would like the following:

hēi jū

If that is acceptable, the requested change could be done. Another solution is to go for a Serif style for all annotations comparable to the ‘Times’ style used for Latin code points in the code chart.

USA: Positive with comments

Technical comments:

TE.1. Kana Supplement, Kana Extended-A, Transport and Map Symbols, Supplemental Symbols and Pictographs, Zanabazar Square

The large number of the new characters in these blocks have already been published in Unicode 10.0. (Hentaigana, emoji, and three Zanabazar Square characters). We strongly recommend that these characters not be removed from this ballot.

Proposed change by US:

We request none of these characters be removed or otherwise changed because of their publication status in Unicode 10.0.

Noted

TE.2. Miscellaneous Symbols and Arrows (2B00-2BFF)

Add “PAWNS”/”PAWN” to the chess notation names for U+2BFA through U+2BFD to clarify what the symbols mean.

Proposed change by US:

Change the names to:

U+2BFA UNITED PAWNS SYMBOL
U+2BFB SEPARATED PAWNS SYMBOL
U+2BFC DOUBLED PAWNS SYMBOL
U+2BFD PASSED PAWN SYMBOL.

Propose Acceptance

The current annotations for these symbols should also be removed.

TE.3. Miscellaneous Symbols and Arrows (2B00-2BFF)

Change the names for three chess notation symbols (U+2BBA..U+2BBC) from “INTERLOCKED” to “OVERLAPPING” to more accurately describe the glyphs.

Proposed change by US:

Change the names to:

U+2BBA OVERLAPPING WHITE SQUARES
U+2BBB OVERLAPPING WHITE AND BLACK SQUARES
U+2BBC OVERLAPPING BLACK SQUARES

Propose Acceptance

TE.4. Khitan Small Script (2B00-2BFF)

The model for the Khitan Small Script is incomplete without the format control characters. As a result, we recommend it be moved to Amendment 2.

Proposed change by US:

Remove Khitan Small Script from PDAM 1.3, and move into Amendment 2.

Propose Acceptance

See also comments T4, E5, T6 from UK and comments T2 and T3 from Ireland.

Editorial comments:

ED.1. Latin Extended-D (A720-A7FF)

To more accurately describe the use of the character, change the subheading for U+A7AF LATIN LETTER SMALL CAPITAL Q and add a usage annotation.

Proposed change by US:

Change the subheading above A7AF LATIN LETTER SMALL CAPITAL Q to "Letter for Japanese linguistics" or "Letter for Japanese phonemic transcription" and add an annotation: "used to represent gemination in Japanese phonetic transcription", or simply "gemination".

Accepted

The subheading will read: "Letter for Japanese phonemic transcription" and the annotation will simply be (gemination').*

ED.2. Armenian (0530-058F)

To more accurately describe the use of the new characters U+058B and U+058C, change the Armenian heading above U+058B "Small letters" to "Lowercase letters" and add a note in the subheader that the characters are letters for phonetic notation.

Proposed change by US:

Make the changes as described.

Propose accepted in principle

See also comment T1 from Ireland.

If the comment T1 from Ireland is accepted, there is no need to change any heading, but annotation can be added to these two moved characters.

ED.3. Miscellaneous Symbols and Arrows (2B00-2BFF)

Add an annotation "True Black Moon Lilith" to U+2BDE BLACK DIAMOND ON CROSS to clarify what astrological symbol it is.

Proposed change by US:

Add the annotation as described.

Accepted

ED.4. Miscellaneous Symbols and Arrows (2B00-2BFF)

Change the subheader above U+2BF3 from "Russian Astrological aspects" to "Russian astrological aspects" for consistency.

Proposed change by US:

Make the change as described.

Accepted

ED.5. Miscellaneous Symbols and Arrows (2B00-2BFF)

The note for U+2BF6 RUSSIAN ASTROLOGICAL SYMBOL BINOVILE says "the letters N²", but ² is not a letter. Similarly, the note for U+2BF8 RUSSIAN ASTROLOGICAL SYMBOL TREDECILE says "the letters D³".

Proposed change by US:

In the notes for U+2BF6 and U+2BF8, change "the letters" to "the term".

Accepted

ED.6. Hanifi Rohingya (10D00-10D3F)

The note “not used at present” for U+10D1C HANIFI ROHINGYA LETTER VA is vague.

Proposed change by US:

Change the note for U+10D1C HANIFI ROHINGYA LETTER VA to: “was not used widely and its usage has been discontinued”.

Accepted

ED.7. Hebrew (0590-05FF)

Change the heading for U+05EF HEBREW YOD TRIANGLE from “Hebrew letter” since it is not a letter.

Proposed change by US:

Change the heading for U+05EF HEBREW YOD TRIANGLE to “Hebrew sign”.

Accepted

ED.8. Grantha (11300-1137F)

Modify the note for U+1133B COMBINING BINDU BELOW, “used as nukta for Badaga”, to one that reflects its use with other languages, as described in WG2 N4840 (=Unicode document L2/15-256).

Proposed change by US:

Change the note for U+1133B COMBINING BINDU BELOW to: “used as nukta for Betta Kurumba and other languages”.

Accepted

Note that the original request to add a Tamil Sign Nukta in N4840 was transformed in a request to this character “COMBINING BINDU BELOW”.

ED.9. Indic Siyaq Numbers (1EC70-1ECBF)

Remove quotes around 'Raqm' and 'Rakam' in the note at the top of the names list for Indic Siyaq Numbers so it is consistent with the format elsewhere.

Proposed change by US:

Make the changes as described.

Accepted