

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: Working Group Document

Title: Preliminary proposal to disunify Emoji symbols for the western zodiac

Author: Karl Pentzlin

Status: Individual Contribution

Action: For consideration by JTC1/SC2/WG2 and UTC

Date: 2011-11-01

1. Introduction

The document L2/11-414 "Emoji Variation Sequences" by Peter Edberg (2011-10-30) discusses the necessity of selecting different presentations for some Emoji symbols which were unified with existing characters.

L2/11-415 "Unified Emoji Reference" by Peter Edberg (2011-10-28) lists all these Emoji.

While L2/11-414 is focused on the problem that Emoji are usually displayed colorful while Unicode characters encoded earlier, especially the targets of the unifications, are usually displayed black and white, this proposal addresses another problem which fig. 1 (an excerpt of the table in L2/11-415) shows.

Looking at the 12 symbols for the Western Zodiac, the DoCoMo and SoftBank symbols are colorful versions of the abstract astrological symbols which are encoded as U+2648...U+2653, and thus they were correctly unified (and subject to the possible solutions outlined in L2/11-414).

However, the KDDI symbols are completely different, being pictographs rather than abstract astrological symbols.

This proposal claims that here applies what Asmus Freytag had written in his mail from 2011-08-08 on the Unicore list (re "windings/ guaranteeing plain text is unaffected by richer layers /tarot"):

It's just that when you do this [unification] for one context, but a shape is used in different contexts where different rules exist, you have to do the right thing so that users in the other context have predictable behavior as well. For symbols, that required coding the cancellation and multiplication x separately.

E.g., being able to read the abstract astrological symbols requires special knowledge, while the pictographs are understandable for a considerably wider audience (who only has to know the Western Zodiac items by name).

In consequence, here the disunification of the Western Zodiac pictographs from the abstract astrological symbols is proposed.

This e.g. would enable KDDI users to ensure that symbols sent to DoCoMo or SoftBank users can be perceived exactly, rather than being translated into abstract symbols based on a special semantic concept.

2. Encoding Considerations

The set of pictographs for the Western Zodiac is stable since at least 1,500 years (see fig. 3 ff.).

When the Emoji were encoded, there were the following decisions made regarding the Eastern astrological symbols:

- The pictographs were included, rather than confining to the abstract symbols (which are already encoded Han characters in this case).
- Where an astrological unit (year of a 12-year cycle, in this case) is expressed by different pictographs in different cultural contexts, all of these pictographs were encoded as different characters.

Here, it is proposed to follow this way.

Therefore, not only the 10 missing pictographs directly representing the KDDI symbols (2 of these 12 symbols can be unified with already encoded symbols) are proposed, but also a small set of pictographs enabling to correctly represent the complete Western zodiac by pictographs in the most common affected cultural contexts. This especially applies if the KDDI symbol differs from the common traditional pictograph, which is proposed as different symbol in such cases.

Remarks on some of the Western zodiac items (enumerated by the Latin name shown in fig. 2, which are identical to the names of the abstract symbols already encoded at U+2648 ... 2653:

Taurus: The usual pictograph shows a full lateral view of the bull.

However, a "Bull" is neither an "Ox" nor a "Cow", thus it is not unified with U+1F402 OX or U+1F404 COW, thus a similar symbol BULL is proposed.

The KDDI symbols shows a bull's face, therefore this is proposed separately.

Gemini: The traditional pictograph shows two young men touching each other's bodies with one or both hands. Sometimes, two little boys are shown instead of men. Modern pictographs often show two women showing this posture. These forms are proposed separately.

The KDDI symbol in fact is a smiley not in any way related to the traditional pictograph.

Therefore, it is proposed separately as smiley in the appropriate block.

Cancer: Pictographs showing a crab (like the KDDI symbol) are found as well as such showing a crayfish. Both forms are proposed, as they are clearly different when regarded as animal symbols rather than zodiac symbols.

Virgo: The traditional pictograph shows a young woman, most commonly in a sitting posture, holding sheaves (typically three) or a tree branch (typically with three twigs) in her hand.

The KDDI symbol simply is a girl's picture (showing the face and the upper part of the body). As such a pictograph is also found elsewhere, this is proposed separately.

Libra: Mostly depicted simply as scales, but pictographs showing a man holding the scales also are found (see e.g. fig. 3). The first form can be unified with U+2696 SCALES, while the second form is proposed here.

Sagittarius: Mostly, the pictograph shows a centaur using a bow and arrow. A common variant shows a man using a bow and arrow, typically in a kneeling posture. The KDDI symbol shows only the bow and arrow, thus reducing the pictograph to a part symbolizing the whole. Consequently, all three symbols are proposed.

Capricorn: The Latin word "capricornus" denotes the alpine ibex (scientific name: capra ibex), also known by its German name "Steinbock" ("stone goat"), which occurs in the Alps and some mountainous areas of Southern Europe.

Therefore, European pictographs usually show this animal.

In other areas, the pictographs show a "sea goat", a mythological creature being composed of the front part of a goat (with its two forelegs) and the rear part of a fish which postures its tail in a loop. According to the table shown in Fig. 2, this interpretation goes back to the

Sumero-Babylonian one.

The KDDI symbol simply shows a goat face, which is considered a valid interpretation of the "sea goat" as, by its composition, the face of this creature is the one of a goat.

Consequently, all three symbols are proposed.

Aquarius: The usual pictograph shows a water bearer (usually in a sitting posture), a man pouring water out of a jar or amphora. The KDDI symbol shows only an amphora, thus reducing the pictograph to a part symbolizing the whole.

Consequently, both symbols are proposed.

3. Proposed Characters

NOTE: *The "representative glyphs" shown in this preliminary document are only placeholders. They have to be replaced by glyphs drawn in the style of related symbols in the Unicode charts.*

The placeholder pictures are taken from Google image searches and may be copyrighted.

Block: *Miscellaneous Symbols and Pictographs (1F300...1F5FF), or a new block "Miscellaneous Symbols and Pictographs Supplement" (e.g. at 1F900...1F9FF)*

Western Zodiac Pictographs

Other symbols belonging to this group are:

→ 1F40F RAM
(KDDI F667) – (ARIES)

→ 2696 SCALES
(KDDI F66D) – (LIBRA)



U+1xxx0

BULL
→ 2649 TAURUS
→ 1F402 OX
→ 1F404 COW



U+1xxx1

BULL FACE
→ 2649 TAURUS
→ 1F42E COW FACE
(KDDI F668)



U+1xxx2

TWO YOUNG MEN TOUCHING
→ 264A GEMINI
→ 1F46C TWO MEN HOLDING HANDS



U+1xxx3

TWO BOYS TOUCHING
→ 264A GEMINI



U+1xxx4

TWO YOUNG WOMEN TOUCHING
→ 264A GEMINI
→ 1F46D TWO WOMEN HOLDING HANDS

	U+1xxx5	CRAB → 264B CANCER (KDDI F66A)
	U+1xxx6	CRAYFISH → 264B CANCER
	U+1xxx7	LION → 264C LEO (KDDI F66B)
	U+1xxx8	YOUNG WOMAN HOLDING SHEAVES → 264D VIRGO
	U+1xxx9	FRONT VIEW OF A YOUNG WOMAN → 264D VIRGO → 1F467 GIRL → 1F479 WOMAN (KDDI F66C)
	U+1xxxA	SCORPION → 264F SCORPIUS (KDDI F66E)
	U+1xxxB	MAN HOLDING SCALES → 264E LIBRA
	U+1xxxC	CENTAUR USING BOW AND ARROW → 2650 SAGITTARIUS
	U+1xxxD	MAN USING BOW AND ARROW → 2650 SAGITTARIUS
	U+1xxxE	BOW AND ARROW → 2650 SAGITTARIUS (KDDI F66F)
	U+1xxxF	ALPINE IBEX = steinbock → 2651 CAPRICORN
	U+1xxy0	SEA GOAT WITH LOOPED FISH TAIL → 2651 CAPRICORN



U+1xxy1 GOAT FACE
→ 2651 CAPRICORN
(KDDI F670)



U+1xxy2 WATER BEARER POURING OUT JAR
→ 2652 AQUARIUS
(KDDI F66E)



U+1xxy3 AMPHORA
→ 2652 AQUARIUS
(KDDI F671)



U+1xxy4 PAIR OF FISH FACING IN OPPOSITE DIRECTION
→ 2653 PISCES
(KDDI F672)

Block: Emoticons



U+1F644 SMILING TWIN FACES
→ 264A GEMINI
→ 263A WHITE SMILING FACE
(KDDI F669)

Properties:

1xxx0;BULL;So;0;ON;;;;;N;;;;;

--- Rest TBD (All other proposed characters follow this pattern.)

4. Proposed changes for annotations of existing characters

U+2696 SCALES
add: → 264E LIBRA
add: · pictograph for the seventh of the signs of the Western zodiac
(KDDI F66D)

U+1F40F RAM
add: → 2648 ARIES
add: · pictograph for the first of the signs of the Western zodiac
(KDDI F667)

5. Proposed changes in EmojiSources.TXT

TBD

6. Examples and Figures

Fig. 1: From L2/11-415 "Unified Emoji Reference" by Peter Edberg (2011-10-28):
Page header from p.1, excerpt from the table from p.3 .

DoCo Mo	KDDI	Soft-Bankk	Apple glyphs: std(s),emoji	Unicd glyph	Unicode codepoint, name, props
F8A7 	F667 	F7DF 	 , 		2648;ARIES;So; = ram
F8A8 	F668 	F7E0 	 , 		2649;TAURUS;So; = bull
F8A9 	F669 	F7E1 	 , 		264A;GEMINI;So; = twins
F8AA 	F66A 	F7E2 	 , 		264B;CANCER;So; = crab
F8AB 	F66B 	F7E3 	 , 		264C;LEO;So; = lion
F8AC 	F66C 	F7E4 	 , 		264D;VIRGO;So; = virgin maiden
F8AD 	F66D 	F7E5 	 , 		264E;LIBRA;So; = scales
F8AE 	F66E 	F7E6 	 , 		264F;SCORPIUS;So; = scorpion
F8AF 	F66F 	F7E7 	 , 		2650;SAGITTARIUS;So; = archer
F8B0 	F670 	F7E8 	 , 		2651;CAPRICORN;So; = goat
F8B1 	F671 	F7E9 	 , 		2652;AQUARIUS;So; = water carrier
F8B2 	F672 	F7EA 	 , 		2653;PISCES;So; = fish

Fig. 2: Listing the Western Zodiac on Wikipedia.
Retrieved 2011-11-01 from <http://en.wikipedia.org/wiki/Zodiac>

No	Symbol	Long.	Latin name	English translation	Greek name	Sanskrit name	Sumero-Babylonian name ^[10]
1	♈	0°	Aries	The Ram	Κριός/Kriós	Meṣa (मेष)	MUL LU ₄ HUN.GA "The Agrarian Worker", Dumuzi
2	♉	30°	Taurus	The Bull	Ταῦρος/Tauros	Vriṣabha (वृषभ)	MULGU ₄ .AN.NA "The Steer of Heaven"
3	♊	60°	Gemini	The Twins	Δίδυμοι/Didymoe	Mithuna (मिथुन)	MUL-MAŠ.TAB.BA.GAL.GAL "The Great Twins" (Castor and Pollux)
4	♋	90°	Cancer	The Crab	Καρκίνος/Karkinos	Karkaṭa (कर्कट)	MUL-AL.LUL "The Crayfish"
5	♌	120°	Leo	The Lion	Λέων/Léōn	Siṃha (सिंह)	MUL-UR.GU.LA "The Lion"
6	♍	150°	Virgo	The Maiden	Παρθένος/Parthénos	Kanyā (कन्या)	MUL-AB.SIN "The Furrow"; "The Furrow, the goddess Shala's ear of corn"
7	♎	180°	Libra	The Scales	Ζυγός/Zygós	Tula (तुला)	zibanitum "The Scales"
8	♏	210°	Scorpio	The Scorpion	Σκορπιός/Skorpiós	Vṛścika (वृश्चिक)	MUL-GIR.TAB "The Scorpion"
9	♐	240°	Sagittarius	The (Centaur) Archer	Τοξότης/Toxótēs	Dhanus (धनुष)	MUL-PA.BIL.SAG, Nedu "soldier"
10	♑	270°	Capricorn	"Goat-horned" (The Sea-Goat)	Αιγόκερωσ/Aegókerōs	Makara (मकर)	MUL-SUḪUR.MAŠ "The Goat-Fish"
11	♒	300°	Aquarius	The Water-Bearer	Υδροχόος/Hydrokhóos	Kumbha (कुम्भ)	MUL-GU.LA "The Great One", later qā "pitcher"
12	♓	330°	Pisces	The Fish	Ιχθύες/Igthues	Mīna (मीन)	MUL-SIM.MAḪ "The Tail of the Swallow", later DU.NU.NU "fish-cord"

Fig. 3: Mosaic pavement of a 6th century synagogue at Beit Alpha, Jezreel Valley, northern Israel – Retrieved 2011-10-31 from: http://en.wikipedia.org/wiki/File:Beit_Alpha.jpg



Fig. 4: Zodiac signs, 16th century , medieval woodcuts. – Retrieved 2011-10-31 from: http://en.wikipedia.org/wiki/File:Zodiac_woodcut.png

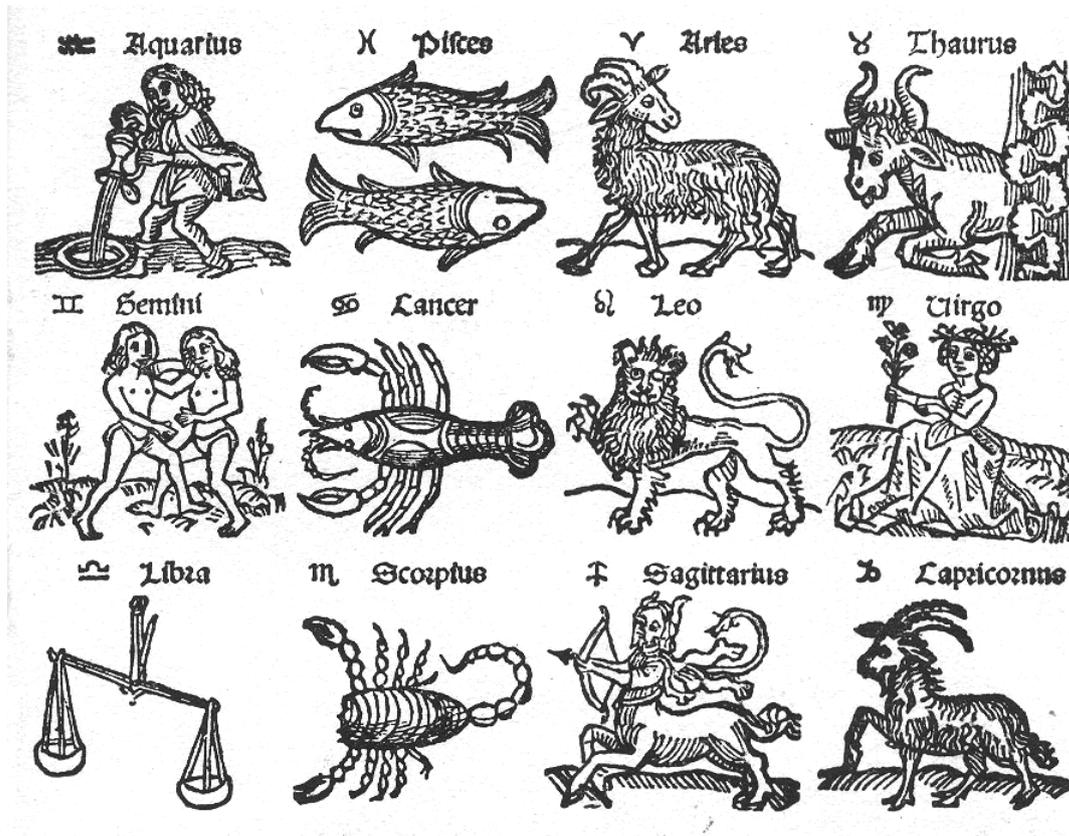


Fig. 5: Gemini and Virgo, reproduction from a (possibly normand) medieval book of astrology (15th century) – Retrieved 2011-10-31 from: <http://en.wikipedia.org/wiki/File:Gemini2.jpg> , <http://en.wikipedia.org/wiki/File:Virgo2.jpg>



Fig. 6: From a book by Barocius, 1585. – Retrieved 2011-10-31 from: <http://www.horoscoper.net/astrology/zodiac.htm>



Fig. 7: From a book from about 1800. – Retrieved 2011-10-31 from: <http://karenswhimsy.com/zodiac-symbols.shtm>

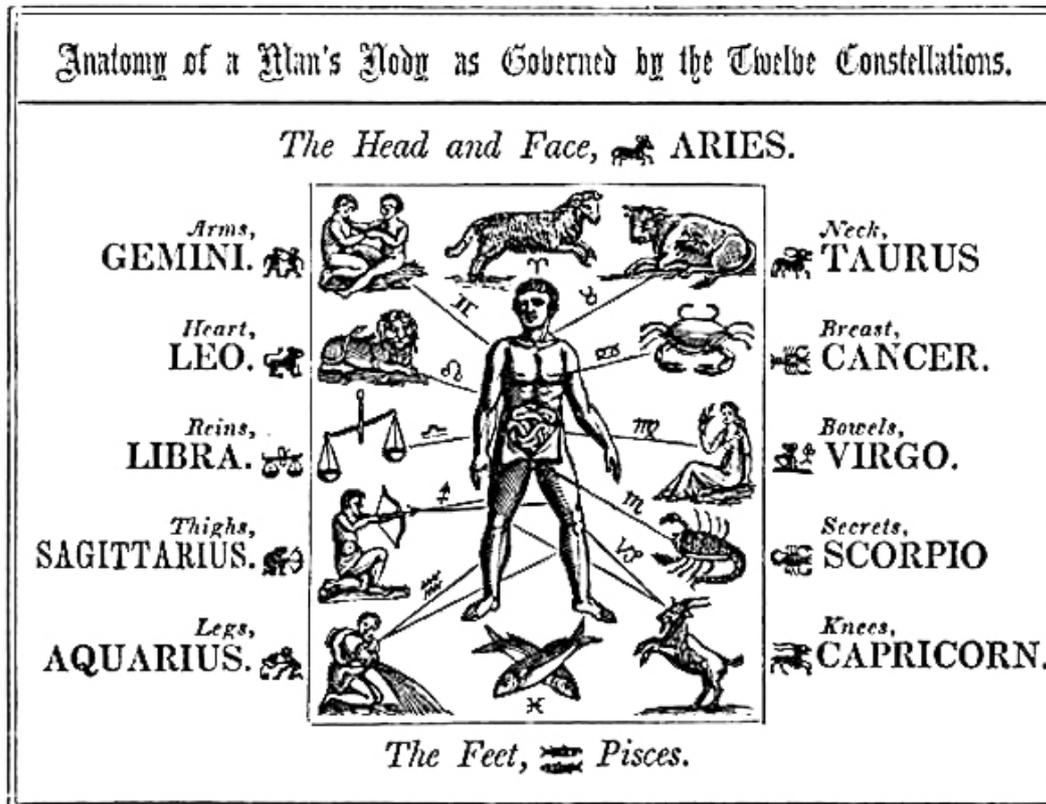


Fig. 8: Illustration from an Ethiopian book. – Retrieved 2011-10-31 from:
<http://tseday.wordpress.com/2008/09/14/ethiopian-astronomy/>

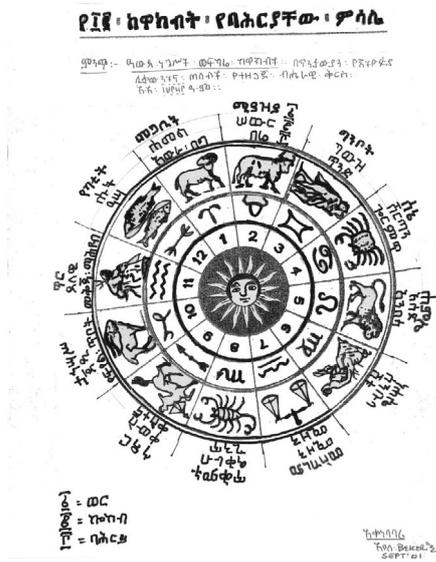


Fig. 9: Illustration from an astrology webpage, original source not identified.
Retrieved 2011-10-31 from:
<http://www.astrologycom.com/zodiac.html>



Fig. 10: Illustration from an astrology-critical webpage, original source of the constituents not identified. – Retrieved 2011-10-31 from:
<http://freddy.com/new-zodiac-sign-ophiuchus-why-astrology-is-even-sillier-than-we-thought/uncategorized/>



Fig. 11: Icons. – Retrieved 2011-10-31 from:
http://www.123rf.com/photo_9720690_abbildung-der-tierkreis-ikonen.html



Fig. 12: Tattoo designs. – Retrieved 2011-10-31 from:
http://www.tattoodonkey.com/astrology-zodiac-tattoo-designs-designsimg/tattoos.org/galleries/Astrology_Zodiac_Tattoos/astrology-zodiac-tattoo-designs.jpg/



Fig. 13: Tattoo designs. – Retrieved 2011-10-31 from:
<http://www.ryory.com/zodiac-tattoo/zodiak2-2>



**ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹**

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

A. Administrative

1. Title:	Preliminary proposal to disunify Emoji symbols for the western zodiac
2. Requester's name:	<i>Karl Pentzlin</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Individual contribution</i>
4. Submission date:	<i>2011-11-01</i>
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<input checked="" type="checkbox"/> Yes
(or) More information will be provided later:	<input type="checkbox"/>

B. Technical – General

1. Choose one of the following:		
a. This proposal is for a new script (set of characters):	<input type="checkbox"/> No	
Proposed name of script:		
b. The proposal is for addition of character(s) to an existing block:	<input checked="" type="checkbox"/> Yes	
Name of the existing block:	<i>Miscellaneous Symbols and Pictographs; Emoticons</i>	
2. Number of characters in proposal:	<i>22</i>	
3. Proposed category (select one from below - see section 2.2 of P&P document):		
A-Contemporary <input checked="" type="checkbox"/>	B.1-Specialized (small collection) <input type="checkbox"/>	B.2-Specialized (large collection) <input type="checkbox"/>
C-Major extinct <input type="checkbox"/>	D-Attested extinct <input type="checkbox"/>	E-Minor extinct <input type="checkbox"/>
F-Archaic Hieroglyphic or Ideographic <input type="checkbox"/>	G-Obscure or questionable usage symbols <input type="checkbox"/>	
4. Is a repertoire including character names provided?	<input checked="" type="checkbox"/> Yes	
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<input checked="" type="checkbox"/> Yes	
b. Are the character shapes attached in a legible form suitable for review?	<input checked="" type="checkbox"/> Yes	
5. Fonts related:		
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<i>TBD</i>	
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<i>TBD</i>	
6. References:		
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<input checked="" type="checkbox"/> Yes	
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<input checked="" type="checkbox"/> Yes	
7. Special encoding issues:		
Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?	<input type="checkbox"/> No	

8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database (<http://www.unicode.org/reports/tr44/>) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

¹ Form number: N3902-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before? If YES explain		No
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)? If YES, with whom? If YES, available relevant documents:	TBD t	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	Everybody interested in astrology referring to the Western Zodiac	Yes
4. The context of use for the proposed characters (type of use; common or rare) Reference:	see text	common
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	see text	Yes
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? If YES, reference:		No
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?		Yes
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? If YES, reference:		No
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? If YES, reference:		No
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? If YES, is a rationale for its inclusion provided? If YES, reference:		No
11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? If YES, reference: Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? If YES, reference:		No
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary)		No
13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference:		No