JTC1/SC2/WG2 N4747
2016-09-12

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

1. Introduction

Star ratings are commonly used in text and online to rate everything from movies and hotels to mutual funds, Medicare plans, and car safety. The Unicode and ISO/IEC 10646 standards include two star symbols which are commonly used for star ratings (usually one through five stars, indicating popularity or satisfaction):
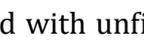
- U+2605 ★ BLACK STAR

- U+2606 ☆ WHITE STAR

When averaging users' star ratings, it is common to round to the nearest half star, and display the average star rating as one or more stars followed by zero or one half star or half-filled star. Some internet sites also allow half stars for users' individual ratings of products or services, but this is less common than for average ratings. There are currently no Unicode characters for such half star or half-filled star symbols, despite being discussed at length on the Unicode mailing list several times, and being frequent topics for discussion on web sites such as Stack Overflow and Yahoo Answers. For background discussion see these threads on the Unicode mailing list:

- Jörg Knappen, "Missing geometric shapes" on 6 Nov 2012
  (http://www.unicode.org/mail-arch/unicode-ml/y2012-m11/0008.html)
- Ken Shirriff, "Adding half-star to Unicode?" on 23 June 2016
  (http://www.unicode.org/mail-arch/unicode-ml/y2016-m06/0094.html)

A star rating is normally shown on a scale of three, four, five or ten stars (five stars being most common), with filled and half-filled stars representing the rating (e.g. ★★★⯪☆ represents 3½ stars out of 5). In LTR contexts, the half star is normally displayed as a

white star with the left half filled-in like ★ (see Figs. 1 through 8 below). In RTL contexts, such as Hebrew or Arabic text, half stars are normally displayed as a white star with the right half filled-in like ☆ (see Fig. 9 and Fig. 10). Other presentations are also possible (e.g. Fig. 15 and Fig. 16).

Less commonly, separate left half-star ◗ and right half-star ◖ symbols may be used in ratings that are not right- or left-justified with unfilled stars (e.g. ★★★◗ represents 3½ stars). See Fig. 11, Fig. 12 and Fig. 13 for examples.

A few web sites have average star ratings that include fractional stars in addition to half stars (see Fig. 14). This usage is not widespread, and other fractional star symbols are not included in this proposal.

It is also possible that half stars or half-filled stars are used in vertical orientation, such as with vertically-written Japanese or Mongolian text, in which case the stars would be halved along the horizontal axis (e.g. ◣ and ◹). However, we have not found any examples of this yet, and so horizontally-halved stars are not included in this proposal.

We propose encoding four half-star and half-filled star symbols in the Miscellaneous Symbols and Arrows block at 2BD8 through 2BD9, as shown below.

Table 1: Proposed Half Star Characters

Code Point	Glyph	Character Name
2BD8	◗	LEFT HALF BLACK STAR
2BD9	◖	RIGHT HALF BLACK STAR
2BDA	★	STAR WITH LEFT HALF BLACK
2BDB	☆	STAR WITH RIGHT HALF BLACK

2. *User Demand*

Many people have asked online for Unicode half-stars, demonstrating the high user demand for this symbol. Stack Overflow has [hundreds of questions about half stars](#), most of which would be solved by a Unicode half star. Other examples of online requests include:

- a [discussion of Unicode stars](#) with three comments asking for half-stars;
- a [Yahoo answers thread](#) asking for Unicode half-star;
- a [question](#) asking for Unicode half star;
- a [Stack Overflow](#) asking how to clip a Unicode star to get a half star;
- a user [asking](#) for a half-star to put in Excel;
- a [Stack Overflow discussion](#) on half-stars that contains a rant about Unicode lacking this character.

Several software workarounds have been developed to display half-stars without their presence in Unicode, such as a [CSS solution](#) and a [JQuery solution](#) using a special font. However, these only work in specific domains, and are not generic solutions that encoded characters would provide.

[HalfStarDemo.ttf](#) is a public-domain demonstration font with these four stars as U+E000 through U+E003. An open-source web font with both half stars ([fa-star-half](#)) and half-filled stars ([fa-star-half-o](#)) in the PUA has also been made available by Font Awesome, although it lacks the RTL versions.

3. *Justification for Encoding Half-Star and Half-Filled Star Symbols*

It has been suggested that only one pair of characters needs to be encoded, and the choice as to whether they are rendered as half-stars or half-filled stars should be left to the font designer. We consider that there are problems with this solution, and that users would be better served by the encoding of all four characters proposed in this document.

The Unicode standard must provide the characters with a representative glyph and a descriptive name, and these cannot encompass both half-star and half-filled star shapes. That is to say, a choice would have to be made whether to call the characters LEFT/RIGHT HALF BLACK STAR (with representative code chart glyphs ◗ and ◖) or whether to call the characters STAR WITH LEFT/RIGHT HALF BLACK (with representative code chart glyphs ☆ and ★). Font designers tend to follow the character name and representative glyph given in the code charts, and so if only characters named LEFT HALF BLACK STAR and RIGHT HALF BLACK STAR are encoded, it is highly unlikely that any font vendor would implement them as half-filled star symbols as that would be confusing and go against user expectations. Conversely, if only characters named STAR WITH LEFT HALF BLACK and STAR WITH RIGHT HALF BLACK are encoded then it is equally unlikely that any font vendor would implement them as half-star symbols as that would also be confusing and go against user expectations. Therefore, if only one pair of half-star or half-filled star symbols are encoded, users will have no choice between half-stars and half-filled stars. If they really want a different style to that provided by font vendors they would have to create their own font, which is not a practical solution for most users.

The Font Awesome web font (http://fontawesome.io/) includes both a half star and a half-filled star in the PUA, which indicates that there is a user demand for both styles of character. Moreover, the presence of both forms in existing fonts argues for the encoding of both half star and half-filled star characters in Unicode for compatibility with these existing fonts, so that users can migrate from legacy PUA fonts to Unicode fonts.

Trying to unify these different shapes will only cause unnecessary confusion and frustration for users. The simplest solution for Unicode and implementers, and the best solution for end users, is to separately encode two half-star symbols and two half-filled star symbols, as shown in Table 1 above.

For further discussion on encoding issues see Section 6 below.

4. Examples of Half Stars and Half-Filled Stars in Print and Online

Fig. 1: Advertisement in *The Times*, 2 July 2016

Fig. 2: Advertisement in the *Metro*, 29 July 2016

Fig. 3: Review for Reading Museum on Google Maps
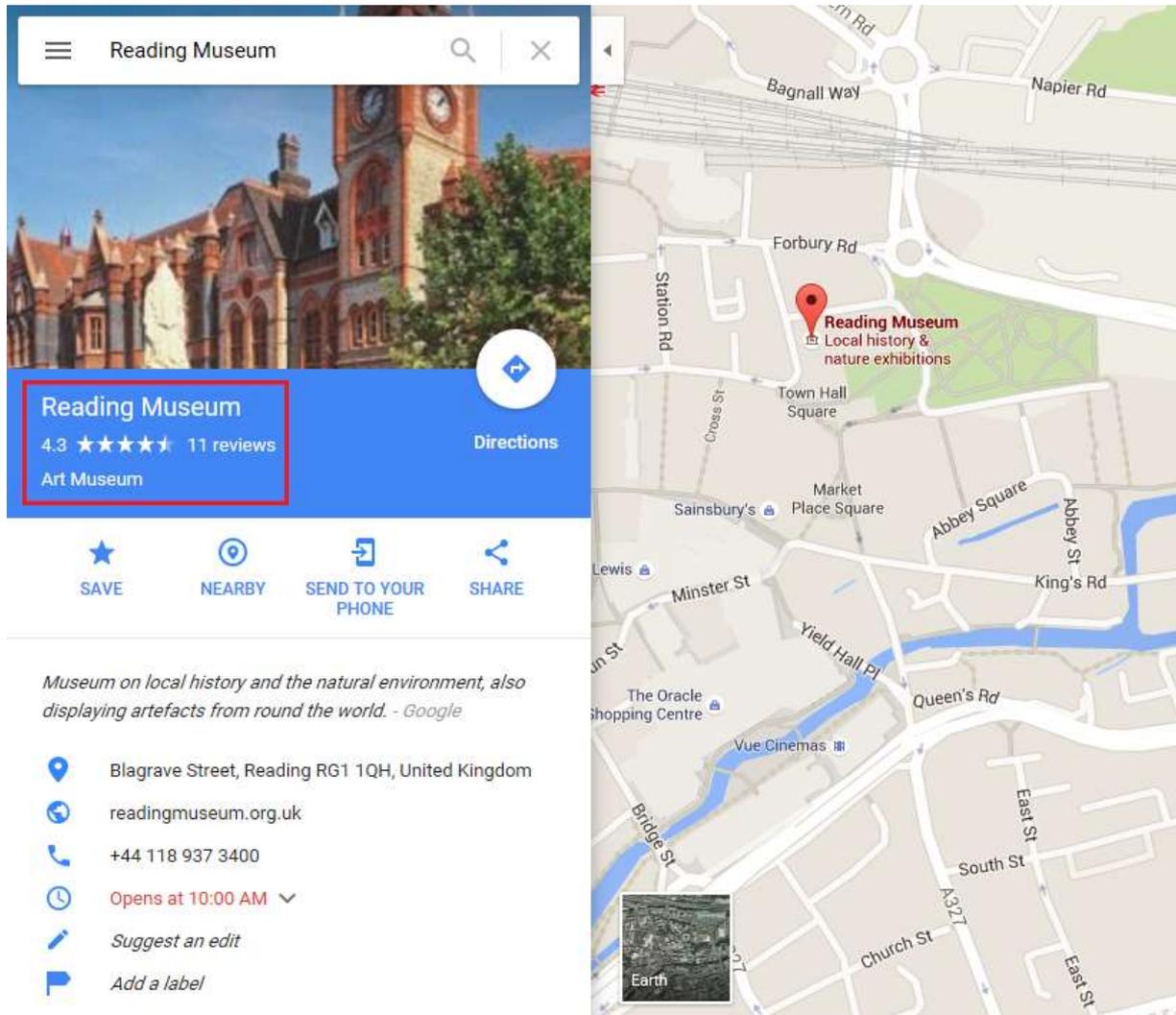
Fig. 4: Google search for French restaurants in San Jose

Fig. 5: Camera rating at amazon.co.uk

Sony DSCW800 Digital Compact Camera (20.1 MP, 5x Optical Zoom) - Black
by Sony
★★★★½ ▼   555 customer reviews | 272 answered questions
#1 Best Seller in Compact System Cameras

Price: £68.97 & FREE Delivery in the UK. Details

Promotion Message Promotion Available   1 Promotion(s) ▼

Fig. 6: Movie Reviews at www.rollingstone.com

JUNE 24, 2016
★½★★

Independence Day: Resurgence

Aliens attack yet again in this so-so sequel to the 1996 sci-fi blockbuster

JUNE 23, 2016
★★★★

Free State of Jones

Matthew McConaughey wants to free the slaves or die tryin' in this civics lesson of a Civil War drama

JUNE 23, 2016
★★½★

The Shallows

Blake Lively finds herself stranded in shark-infested waters in this summer-movie nailbiter

Fig. 7: Key to star ratings on filmphonic.com

Fig. 8: Star ratings at pactarcanum.com

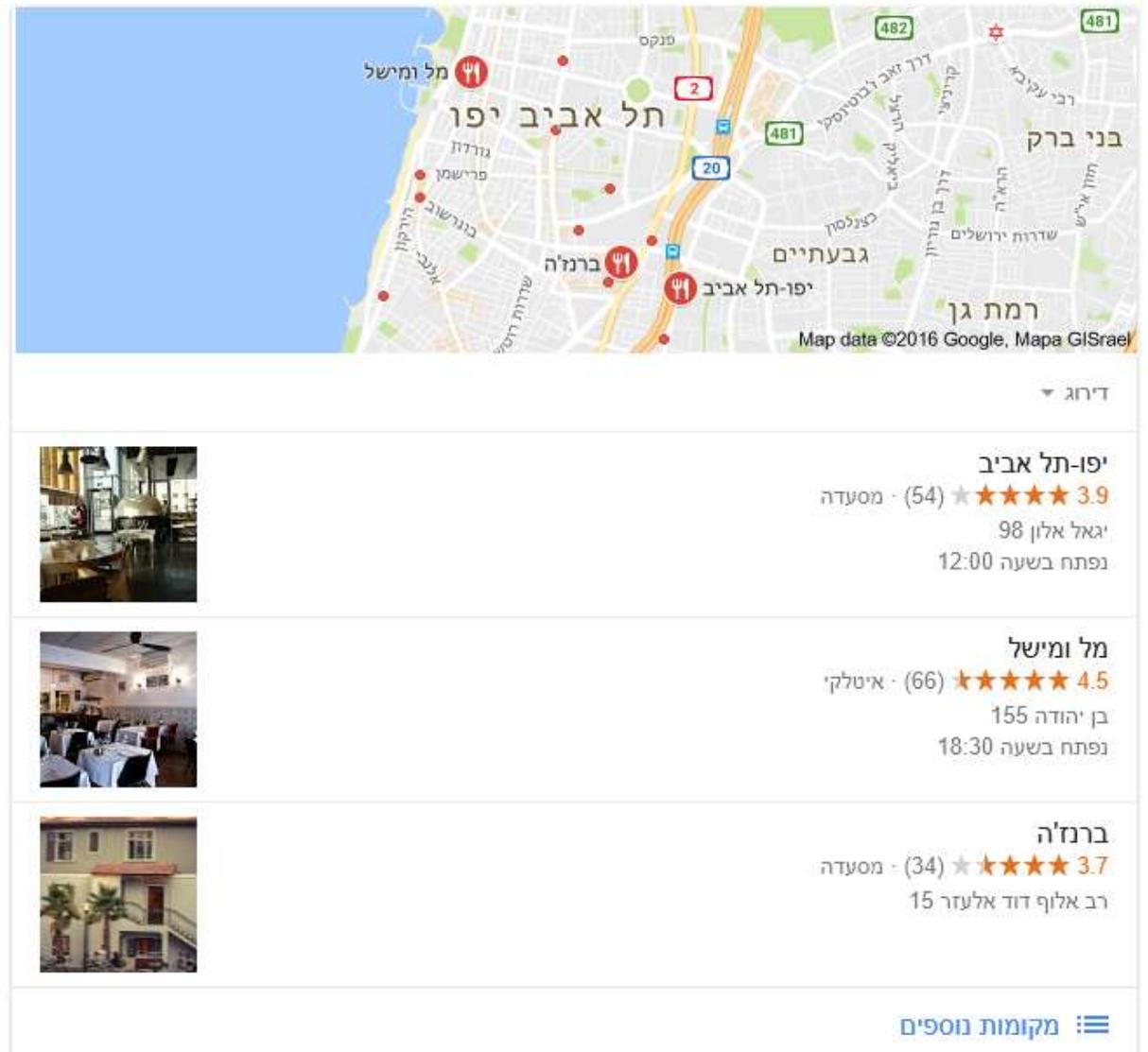**Fig. 9: Search for restaurants on www.google.co.il**

Fig. 10: Advertisement in *Haaretz*, 2 November 2012

(Image posted to Unicode mailing list by Simon Montagu)

Fig. 11: "Pennsylvania Agencies Outperform National Average on New Home Health Star Ratings", 2 January 2016

Quality of Patient Care Star Rating	Percent (approximate)
★★★★★	2.5
★★★★⯪	8.6
★★★★	23.4
★★★⯪	27.7
★★★	21.3
★★⯪	9
★★	5.4
★⯪	1.8

back | printer-friendly version

Fig. 12: Expedia: San Jose Hotels Search

Today 25% off

Hotel Presidente - San Jose City Center, Costa Rica ★★★⯨

San Jose   Map ⦿
020 3564 4852 (Call cost)
Expedia Rate ✔ Free Cancellation
3 people booked this hotel in the last 48 hours
Booked in the last 3 hours

Very good! 4.1/5
(806 reviews)
4 nights from
£379 £284
Sale!

Today 42% off

Country Inn & Suites By Carlson, San Jose Aeropuerto ★★★⯨

Ciudad Cariari   Map ⦿
020 3564 4852 (Call cost)
Expedia Rate ✔ Free Cancellation
5 people booked this hotel in the last 48 hours
Booked in the last 5 hours

Very good! 4.0/5
(403 reviews)
4 nights from
£414 £240
Sale!
✔ Book Now, Pay Later

Hotel El Portico ★★⯨

San Jose De La Montana   Map ⦿
020 3564 4852 (Call cost)
Expedia Rate

4 nights from
We have 1 left at
£230

Fig. 13: Ratings of Cruise ships at www.vacationstogo.com

Cruise Ships & Ratings

| Choose a Ship | ∨ |

Show Me the Deals!

All of the world's major cruise ships are listed below, by cruise line. The ship rating is based on a 6-star system, where 1 star = Limited Appeal and 6 stars = Exceptional In Every Way. The key to the ratings is at the bottom of the page.

For detailed information about a cruise ship or line, click the name below. For a list of special discounts on any of these ships, select any ship and click "Show Me the Deals!".

Azamara Club Cruises
Azamara Journey (2000) ★★★★★⤍
Azamara Quest (2000) ★★★★★⤍

Celebrity
Celebrity Constellation (2002) ★★★★★
Celebrity Eclipse (2010) ★★★★★⤍
Celebrity Equinox (2009) ★★★★★⤍
Celebrity Infinity (2001) ★★★★★
Celebrity Millennium (2000) ★★★★★
Celebrity Reflection (2012) ★★★★★⤍
Celebrity Silhouette (2011) ★★★★★⤍
Celebrity Solstice (2008) ★★★★★⤍
Celebrity Summit (2001) ★★★★★
Celebrity Xpedition (2001) ★★★★⤍
Celebrity Xperience (2017) ★★★★⤍

Blount Small Ship Adventures
Grande Caribe (1997) ★★★
Grande Mariner (1998) ★★★

Carnival
Carnival Breeze (2012) ★★★★⤍
Carnival Conquest (2002) ★★★★
Carnival Dream (2009) ★★★★⤍
Carnival Ecstasy (1991) ★★★⤍
Carnival Elation (1998) ★★★⤍
Carnival Fantasy (1990) ★★★⤍
Carnival Fascination (1994) ★★★⤍
Carnival Freedom (2007) ★★★★⤍
Carnival Glory (2003) ★★★★⤍
Carnival Imagination (1995) ★★★⤍
Carnival Inspiration (1996) ★★★⤍

Fig. 14: Ratings of websites showing 7.4 and 9.6 stars

Fig. 15: Australian Health Star Ratings on food products

Fig. 16: Yelp London Arts & Entertainment Ratings

Fig. 17: Various other examples of star ratings

5. Unicode Properties

```
2BD8;LEFT HALF BLACK STAR;So;0;ON;;;;;N;;;;;
2BD9;RIGHT HALF BLACK STAR;So;0;ON;;;;;N;;;;;
2BDA;STAR WITH LEFT HALF BLACK;So;0;ON;;;;;N;;;;;
2BDB;STAR WITH RIGHT HALF BLACK;So;0;ON;;;;;N;;;;;
```

Range: Miscellaneous Symbols and Arrows (2B00..2BFF), with proposed code points 2BD8-2BDB. The symbols are suitable for the Basic Monolingual Plane (BMP); they are high utility due to the size of the community and their heavy use of the symbol. Code points in the Supplemental Multilingual Plane (SMP) are not recommended because these characters are likely to be used by a variety of mobile devices and apps, with variable levels of support for SMP characters.

Collation: The stars are not part of an existing collating sequence and does not require a particular collation order.

Bidi mirroring: As with other existing half shape symbols such as U+25D6 LEFT HALF BLACK CIRCLE, the proposed characters should have bidi mirrored = No.

6. Further Discussion

Can an existing Unicode symbol be used?

There are dozens of existing Unicode stars, but none function well as half stars. We have seen documents that use workarounds such as ✸ (HEAVY OUTLINED BLACK STAR) or (★) (a star in parentheses). These workarounds are not good substitutes for a half star and usually require additional explanatory text. For example, Malcolm McDonald's *Market Segmentation* (Elsevier, 2004) p. 318:

> In those cases where the relative importance of a segment's DBCs remain expressed as symbols, such as stars and half-stars, they now need to be expressed numerically and 100 points distributed between them. This requires the following steps to be carried out for each segment;
>
> 1 Convert the symbols to numbers with each 'half-star', for example, being equivalent to a single unit, thus '★ ★ ★ ☆' (three and a half stars) is equivalent to 'seven'.
> 2 Add up the individual numbers to arrive at a total for the segment.
> 3 Divide each individual number by the total and multiply the answer by 100.
>
> A worked example appears in Table 10.3.

Table 10.3 Expressing the relative importance of DBCs numerically

	Segment 2		
DBCs	DBC rating	As a number	Reworked
1. Product technical performance	★ ★ ★ ★ ☆	9	41
2. Leading edge image	★ ★ ★	6	27
3. Delivery reliability	★ ★ ☆	5	23
4. Price	★	2	9
	Total	22	100

The ½ symbol is sometimes used in place of the half star in contexts where a half-star image would be difficult to use, such as in text. Here is a published example from *Leonard Maltin's 2015 Movie Guide*:

> **About Schmidt** (2002) C-124m. *** ½ D: Alexander Payne. Jack Nicholson, Hope Davis, Dermot Mulroney, Len Cariou, Howard Hesseman, Kathy Bates, June Squibb, Matt Winston, Harry Groener. Superb human comedy about a newly retired insurance actuary in Nebraska who begins—for the first time—to question the choices he has made in life. Subtle, deliberately paced, and splendidly acted, with as many sad moments as funny ones. Nicholson's magnificently quiet, true performance may be the best of his career. Enhanced by Rolfe Kent's perfectly nuanced score. Payne and Jim Taylor (loosely) adapted Louis Begley's novel. [R] ▼▶

An interesting use of ½ for the half-star is in the title of a paper in *Popular Music* (2001) by Kembrew McLeod:

'★½': a critique of rock criticism in North America

KEMBREW McLEOD

A major reason not to use ½ for a half star is it doesn't work well in an interactive setting. Replacing ☆ with ★ when clicked works more smoothly than replacing ☆ with ½.

What about other half-symbols for ratings?

The question comes up in discussion: What if someone wants to use half-tomatoes or half-hearts? There's no demonstrated demand for other half-symbols, while half-stars are commonly used. We don't anticipate a flood of other half-symbols.

For concrete evidence, a Google search for [unicode half] turns up a variety of related search terms, including "unicode half star" and "unicode half filled star". This indicates that there is a demand for the half star and half-filled star symbols, but no untapped demand for half tomato symbols. However, there may be a demand for half heart, but that is probably related to broken heart 💔 rather than to ratings.

unicode half

unicode half
unicode half **star**
unicode half **space**
unicode half **circle**
unicode half **filled circle**
unicode half **block**
unicode half **filled star**
unicode half **note**
unicode half **bracket**
unicode half **heart**

Google Search   I'm Feeling Lucky
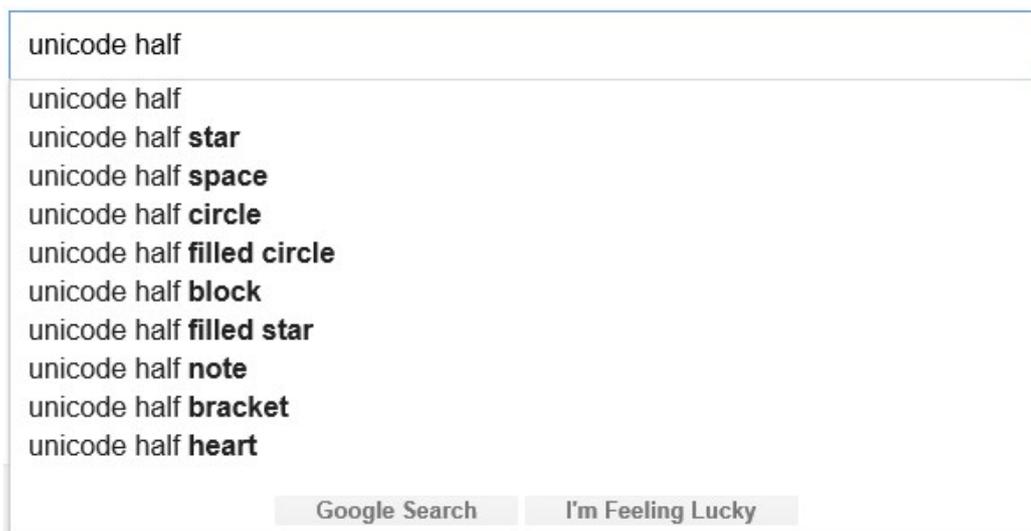
What about fancy styling?

Some sites use colored stars — half yellow, half gray is popular. Other sites use highly decorative stars. We don't think Unicode should try to support these uses; sites can use images, as they do now if they want more styling than is possible with characters.

Why just half? What about quarters or 13%?

Should other fractions be supported, opening a can of worms? There's no demonstrated demand for other fractions. The half-star is a very common rating symbol. Many systems (e.g. Amazon) display a numeric value such as 4.3, but round the displayed stars to the half-star. Use of other star fractions is very rare.

One suggestion was to use a combining character to permit an arbitrary fraction to be applied to an arbitrary character. This would allow a 13% filled heart for instance. While very flexible, this would be a nightmare to implement. There is no evidence of demand for this.

Why not double the number of stars and get rid of halves?

The suggestion sometimes comes up to use a scale of 0-10 stars, and then the half star won't arise. It was pointed out that this makes subitizing more difficult. That is, people can understand 4 stars at a glance, but larger numbers would require much slower counting.

Conclusion

Half stars are commonly used for ratings on websites and publications. There is user demand for half stars and the lack of Unicode support requires inconvenient workarounds.

Looking at the Unicode Criteria for Encoding Symbols, the half stars are a good fit for addition to Unicode. They occur in running text. They have a well defined user community. Being able to search for half-stars in text would be useful. They will complete the class of star symbols already in the standard. Finally, they are letter-like in the sense that they often should match the surrounding font style. For these reasons, the four suggested half stars are proposed for addition to Unicode.

7. Acknowledgments

8. *Proposal Summary Form*

<table>
<tr><td colspan="2" align="center">SO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646[1]
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*.</td></tr>
</table>

A. Administrative

1. **Title:**	*Final proposal to encode four half star symbols*
2. Requester's name:	*Ken Shirriff, Andrew West*
3. Requester type (Member body/Liaison/Individual contribution):	*Individual contribution*
4. Submission date:	*2016-09-12*
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	*YES*
(or) More information will be provided later:	

B. Technical – General

1. Choose one of the following:	
a. This proposal is for a new script (set of characters):	*NO*
Proposed name of script:	
b. The proposal is for addition of character(s) to an existing block:	*YES*
Name of the existing block:	*Miscellaneous Symbols and Arrows*
2. Number of characters in proposal:	*4*

3. Proposed category (select one from below - see section 2.2 of P&P document):

A-Contemporary   **X**   B.1-Specialized (small collection)   B.2-Specialized (large collection)
C-Major extinct   D-Attested extinct   E-Minor extinct
F-Archaic Hieroglyphic or Ideographic   G-Obscure or questionable usage symbols

4. Is a repertoire including character names provided?	*YES*
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	*YES*
b. Are the character shapes attached in a legible form suitable for review?	*YES*
5. Fonts related:	
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	
Andrew West	
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	
Andrew West	
6. References:	
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	*YES*
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	*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)?	*YES*

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: N4102-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, 2012-01)

C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	*YES*
If YES explain *L2/16-186, L2/16-230*	
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.)?	*YES*
If YES, with whom? *Unicode mailing list*	
If YES, available relevant documents:	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	*NO*
Reference:	
4. The context of use for the proposed characters (type of use; common or rare)	*common*
Reference:	
5. Are the proposed characters in current use by the user community?	YES
If YES, where? Reference:	
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	*NO*
If YES, is a rationale provided?	
If YES, reference:	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	*NO*
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	*NO*
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?	*NO*
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character?	*NO*
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
11. Does the proposal include use of combining characters and/or use of composite sequences?	*NO*
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:	
12. Does the proposal contain characters with any special properties such as control function or similar semantics?	*NO*
If YES, describe in detail (include attachment if necessary)	
13. Does the proposal contain any Ideographic compatibility characters?	*NO*
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	