Proposal for addition of half stars

Ken Shirriff
Aug 4, 2016

Abstract

Star ratings are commonly used in text and online to rate movies, products and businesses. Ratings often use half stars, but unfortunately Unicode does not include a half star. There is user demand for half stars in Unicode. This proposal requests four types of half stars: solid and outlined, with left and right half stars for RTL support.

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. Ratings often use half stars, but unfortunately Unicode does not include a half star. There is substantial user demand for half stars in Unicode. This proposal requests four types of half stars: solid and outlined, with left and right half stars for bidirectional support.

Examples of star ratings

Ratings with half-stars are commonly used online, for example Amazon product reviews:

Google ratings:

Goodreads.com book reviews:

The half star is heavily used on movie review sites. Some examples follow, in several languages.

bigpicturebigsound.com:

By Matthew Passantino ★★★★

blu-ray.com

Independence Day: Resurgence Review
Preview by Brian Orndorf, June 24, 2016

★★★★★ 3 / 10

bollywoodhungama.com:

| 1. | Raman Raghav 2.0 | ★★★★★ |
| 2. | Udta Punjab | ★★★★★ |

cinemaglitz.com:

Amma Kanakku Movie Review: The
calculation in Tamil went terribly
wrong...

★★★★★

biosagenda.nl:

★★★★★
Waardering: 3.3/5

rollingstone.com:

JUNE 23, 2016
★★★★

The Shallows

Blake Lively finds herself stranded in shark-infested waters

rogerebert.com:

★★★ Free State of Jones

★★★★ Hunt for the
Wilderpeople

theglobeandmail.com:

★★★★
KATE TAYLOR
The Globe and Mail

Proposal for addition of half stars, 2 of 10

The government system for rating Medicare plans uses half stars, resulting in many plan documents with half stars. For example:

We received the following Summary Star Rating for HealthSpan's health/drug plan services:

Health Plan Services: ★★★★⯪
4.5 Stars

Drug Plan Services: ★★★★★
5 Stars

The number of stars shows how well our plan performs.

Half stars are used in contexts as diverse as rating windows or global competition:

| 13 | 3/12/4 pyrolytic low-e IG, argon fill | Al, thermally improved | ★★⯪ | ★★★★⯪ |
| 14 | 3/12/4 pyrolytic low-e IG, argon fill | Timber or uPVC | ★★⯪ | ★★★★★ |

| Brazil's CADE | ★★★★⯪ |
| Canada's Competition Bureau | ★★★★⯪ |

RTL

Simon Montagu provided the following example of half-stars in RTL text, published in Haaretz 2 November 2012. This shows that usage of half stars isn't limited to left-to-right languages. Note that the line of stars is reversed: the half-star is on the left and has the solid part on the right, opposite of left-to-right usage.

⯪ ★ ★ ★ ★
"השאיר על פני
חיוך בלתי נגמר"
יאיר הוכנר – אתר 'סרט'

★ ★ ★ ★
"ממתק אמיתי"
ישי קיצ'לס – ישראל היום

User demand for the symbol

Many people have asked online for Unicode half-stars, showing the 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 are this [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, and 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

The suggested glyphs

From above, it can be seen that half stars are used in both solid and outlined forms. The four proposed glyphs are:

◗ LEFT HALF BLACK STAR

◖ RIGHT HALF BLACK STAR

⯪ STAR WITH LEFT HALF BLACK

⯫ STAR WITH RIGHT HALF BLACK

The glyphs are based on 'BLACK STAR' ★ (U+2605) and 'WHITE STAR' ☆ (U+2606) with the appropriate half filled in.

A typical use for these characters would be representing 3.5 of 5:

★★★⯨ (standard)

⯪★★★ (RTL language)

★★★⯨☆ (empty stars outlined)

☆⯫★★★ (empty stars outlined, RTL)

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

Properties

The proposed [properties](#) for the new symbols are:

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

Name: The names LEFT HALF BLACK STAR, RIGHT HALF BLACK STAR, STAR WITH LEFT HALF BLACK, STAR WITH RIGHT HALF BLACK are suggested. The names parallel LEFT HALF BLACK CIRCLE, CIRCLE WITH LEFT HALF BLACK, etc. These names follow the shape-based character naming guidelines of ISO/IEC 10646:2014 Annex L Guideline 1.

Range: Miscellaneous Symbols and Arrows (2B00-2BFF), with proposed code points 2BD8-2BDB – the next available code points in the range. 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: Existing stars are in the "Other Neutrals" class, so half stars should probably use the ON bidirectional class. The half stars have the obvious mirrored counterparts, so they can be Bidi mirrored. However, similar characters such as LEFT HALF BLACK CIRCLE are not marked as mirrored. I'll leave it up to the Unicode experts to determine if Bidi Mirrored would be appropriate or not.

Property	Suggested Value
Code point	2BD8-2BDB
Name	LEFT HALF BLACK STAR
	RIGHT HALF BLACK STAR
	STAR WITH LEFT HALF BLACK
	STAR WITH RIGHT HALF BLACK
General Category	So
Canonical Combining Class	0
Bidirectional Class	ON
Decomposition Type /	
Decomposition Mapping	
Numeric Type	
Numeric Value	
Bidi Mirrored	Y?
Unicode 1 Name	
ISO Comment	
Simple Uppercase Mapping	
Simple Lowercase Mapping	
Simple Titlecase Mapping	

Table 1: Suggested character properties for the half stars.

The stars are uncased and have no special line-breaking behavior. The character is not meant for use in identifiers. It is not a white-space character and has no numeric value. It is not a combining character or punctuation.

Potential issues

Can an existing Unicode symbol be used?

There are dozens of existing Unicode stars, but none function well as half stars. I've seen documents 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 explanatory text. For example (*Market segmentation*, by Malcolm McDonald):

Convert the symbols to numbers with each 'half-star', for example, being equivalent to a single unit, thus '★ ★ ★ ☆' (three and a half stars) is

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. A published example from *Leonard Maltin's 2009 Movie Guide*:

About Schmidt (2002) C-124m. ***½ D: Alexander Payne. Jack Nicholson, Hope

An interesting use of ½ for the half-star is in the tile of a [journal paper](#):

'★$\frac{1}{2}$': a critique of rock criticism
in North America

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. I don't anticipate a flood of other half-symbols.

For concrete evidence, a Google search for [unicode half] turns up a variety of related searches. All the Unicode symbols already exist except for "unicode half star". This suggests there is a demand for the half star, but no other untapped demand for half characters.

Searches related to unicode half

unicode half **block**   unicode half **width**

unicode half **note**   unicode half **circle**

unicode half **star**   **special** unicode **characters**

unicode half **space**   one half **character**

What about fancy styling?

Some sites use colored stars - half yellow, half gray is popular. Other sites use highly decorative stars. I 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.

ISO/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://std.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://std.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html **.**
See also http://std.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

A. Administrative

1. **Title:** *Proposal for addition of half stars*
2. Requester's name: *Ken Shirriff*
3. Requester type (Member body/Liaison/Individual contribution): *individual*
4. Submission date: *08/04/2016*
5. Requester's reference (if applicable):
6. Choose one of the following:
     This is a complete proposal: *X*
     (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?
     See text
     b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):
     public domain font
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*
     see text

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: N4502-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? *No*
   If YES explain
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.)? *No*
   If YES, with whom?
   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? *Yes*
   Reference: *see text*
4. The context of use for the proposed characters (type of use; common or rare) *common*
   Reference: *see text*
5. Are the proposed characters in current use by the user community? *Yes*
   If YES, where? Reference: *see text*
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely
   in the BMP? *Yes*
   If YES, is a rationale provided? *Yes*
   If YES, reference: *see text*
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? *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? *n/a*
    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: