

**Title:** Proposal to encode two Western-style tally marks

**Source:** Ken Lunde (Adobe) & Daisuke MIURA

**Status:** Individual contribution

**Action:** For consideration by UTC

**References:** [L2/15-328](#) & [L2/16-046](#)

**Date:** 2016-03-14

## Abstract

*Tally marks, which form a unary numeral system, have been used for thousands of years as a method of counting or keeping score through the use of stroke clusters, and are still in use today. Tally marks also predate modern digits, and in some cases, served as their basis for some scripts. Furthermore, most tally mark systems are comprised of five digits, which correspond to the number of digits on a single human hand.*

*This proposal introduces the Western-style “tally mark” system, and provides evidence that suggests that the characters that comprise it require a “plain text” representation, and should thus be encoded. The original proposal included five characters for this tally mark system, but this proposal reduces it to two, given that three of the characters, which represent the digits two through four, can be handled as sequences of the character that represents the digit one.*

## Proposal

This document proposes that the two characters, TALLY MARK ONE and TALLY MARK FIVE, be accepted for encoding in a future version of the standard, and that U+1D377 and U+1D378 in the Counting Rod Numerals block serve as tentative code points. It is important to point out that the character in this proposal named TALLY MARK FIVE was to be accepted during UTC #146 (see document [L2/15-328](#)), but I explicitly asked for its acceptance to be delayed so that I could have an opportunity to provide a more compelling argument for encoding TALLY MARK ONE so that both characters, which form a system, could be encoded in a contiguous fashion.

## Introduction

Wikipedia’s very brief [Tally marks](#) article describes four tally mark systems, one of which serves as the basis for this proposal.

The most common tally mark system, which is widely used in Europe, North America, and elsewhere, uses one to four parallel vertical strokes to represent the digits one through four, respectively, along with a horizontal or diagonal stroke that crosses over the four previous vertical strokes—completing the tally cluster—to represent the digit five. This tally mark system is sometimes referred to as the *Barred-Gate Tally Mark System*, because the four vertical strokes represent the posts of a gate, and the fifth stroke represents a bar that closes the gate. **In this proposal, this tally mark system uses character names that begin with TALLY MARK.**

## Examples

The following two tables illustrates how these this tally mark system is used to represent the numbers 1 through 20, and the ideographic tally marks that were accepted during UTC #146 (see document [L2/16-046](#)) are included for comparison and reference purposes:

1	2	3	4	5	6	7	8	9	10
I	II	III	IIII	𠄎	𠄎I	𠄎II	𠄎III	𠄎IIII	𠄎𠄎
一	𠄎	𠄎	𠄎	正	正一	正二	正三	正四	正正

11	12	13	14	15	16	17	18	19	20
𠄎𠄎I	𠄎𠄎II	𠄎𠄎III	𠄎𠄎IIII	𠄎𠄎𠄎	𠄎𠄎𠄎I	𠄎𠄎𠄎II	𠄎𠄎𠄎III	𠄎𠄎𠄎IIII	𠄎𠄎𠄎𠄎
正正一	正正二	正正三	正正四	正正正	正正正一	正正正二	正正正三	正正正四	正正正正

## Representative Glyphs

The representative glyphs for the characters in this proposal are used in the tables of this document, reflect the forms that are in common or modern use, and are also provided in the OpenType font that is attached to this PDF file whose glyphs are encoded from the tentative code points in this proposal. The attached OpenType font also includes the ‘[ccmp](#)’ (*Glyph Composition/Decomposition*) typographic feature that substitutes multiple instances of TALLY MARK ONE into glyphs that represent the digits two through four, along with that of TALLY MARK FIVE for the sake of completeness.

The only important implementation note is that the fifth stroke of TALLY MARK FIVE, which completes the tally cluster, can be a perfectly horizontal stroke or a diagonal stroke with a positive or negative slope. *The representative glyph in this proposal uses a diagonal stroke with a positive slope, which appears to be the most common form.*

## Similar Characters

The table in the next section includes a “Similar Characters” column that draws attention to other characters in Unicode that have a *similar* semantic or shape, but are not deemed suitable due to the following factors:

- The characters function as a system, and encoding them in a contiguous range consisting of only two code points helps to guarantee that implementations will include both characters
- The characters have distinct usage or semantics, specifically that of tally marks with specific numeric attributes
- The glyphs for these characters should be compatible with each other in terms of their shape, height, weight, metrics, placement within the em-box, and other typographic attributes

The third factor is perhaps the strongest for encoding TALLY MARK ONE, specifically to better ensure that its glyph is reasonably compatible with that of TALLY MARK FIVE from a typographic perspective, and that the representative glyphs for this pair of characters serve as an example for developers of font implementations.

## Proposed Block, Character Names & Code Points

The characters fit into the existing block named “Counting Rod Numerals” (1D360 . . 1D37F) in the “Notational systems” zone (1D000 . . 1DFFF). The table below provides representative glyphs, suggested code points, and proposed character names, along with the code points of similar—in shape or meaning—characters, though the list is not exhaustive due to the large number of characters whose appearance is that of a simple vertical stroke:

Glyph	Code Point	Character Name	Similar Characters
	U+1D377	TALLY MARK ONE	U+007E U+16C1 U+2160 U+2223 U+A830 U+A876 U+A8CE U+AAF0 U+FE31 U+FF5C U+FFE8 U+109C0 U+10A40 U+10CFA U+1D360 U+1D369
	U+1D378	TALLY MARK FIVE	U+168E

## Properties

The following are the proposed properties for the two characters in this proposal:

1D377;TALLY MARK ONE;No;0;L;;;1;N;;;;;  
1D378;TALLY MARK FIVE;No;0;L;;;5;N;;;;;

## Emoji Representation

The characters in this proposal are candidates for taking on *emoji style*, such as to use glyphs that appear hand-written, made up of stick-like shapes, or even animated via the new OpenType ‘[SVG](#)’ (*Scalable Vector Graphics*) table.

## References, Evidence & Attestation

This section will provide references, evidence, and attestation of the characters for the tally mark system that is proposed in this document. Links to the references that correspond to the evidence are provided when available.

The following [excerpt](#) from a worksheet illustrates the use of TALLY MARK characters in educational materials:

**ANSWER KEY**

**Tally Marks**

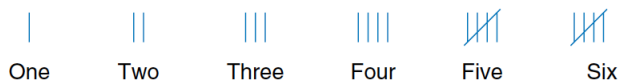
Count the tally marks. Write the value of each set.

a. 26    ||||   ||||   ||||   ||||   ||||   |

b. 8    ||||   |||

Below is an excerpt from [page 54](#) of *Saxon Math 6/5, Third Edition, Student Edition* (Saxon Publishers, 2004) that describes the tally mark system using both TALLY MARK characters:

**Tally marks** Tally marks are used to keep track of a count. Each tally mark counts as one. Here we show the tallies for the numbers one through six.



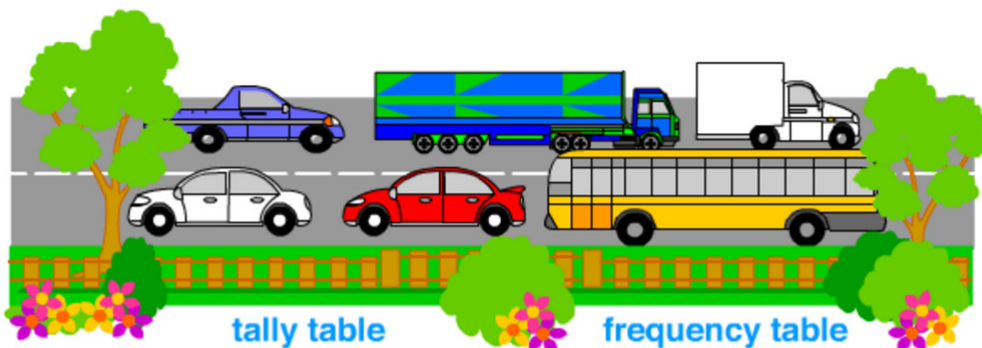
Notice that the tally mark for five is a diagonal mark crossing four vertical marks.

A *Maths Dictionary for Kids* [website](#) shows the two alternate forms of TALLY MARK FIVE (interestingly, the first example uses four instances of U+007C to which a *strikeThrough* style is applied to simulate the TALLY MARK FIVE character):

### tally

- a record of an amount.
- using tally marks to record counting,
- count by 5's to get the total, for example,
  - ~~||||~~ ~~||||~~ ||| = 13.

### traffic frequency tally



vehicle	frequency	vehicle	frequency
cars	<del>    </del>	cars	8
trucks	<del>    </del>	trucks	5
vans		vans	2
utes		utes	2
buses		buses	4
bikes		bikes	0



An [excerpt](#) from *The Big Questions: Mathematics* by Tony Crilly (Quercus Publishing, 2011) refers to the TALLY MARK characters as *Barred-Gate*:

with the act of counting. Tally sticks were still being used in England for the purposes of tax collection in the 13th century, and surprisingly this traditional system lasted until the 1820s, when paper records took over. Today, tally marks still serve a purpose when the aim is to record a continually rising quantity, such as the number of points scored in a game or data for statistical investigations. The familiar method involves counting to five to form an inscription resembling a barred gate, and it seems to have very early origins.



THE BARRED-GATE TALLY SYSTEM



A SOUTH AMERICAN TALLY SCHEME

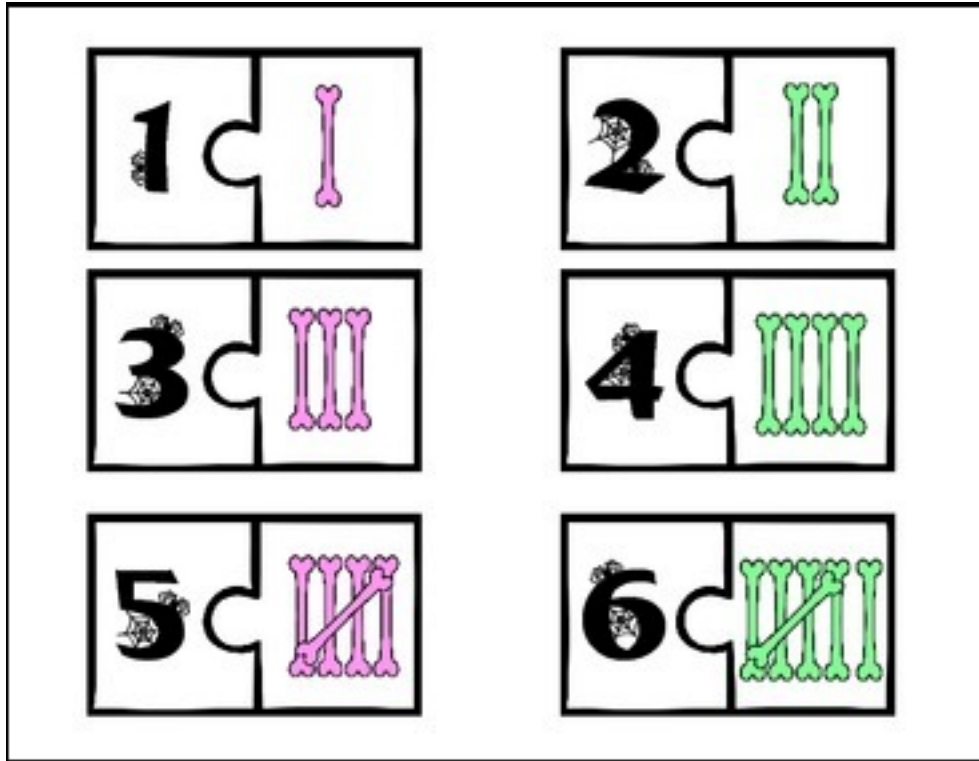
Tally methods are found all over the world. One scheme, from South America, also adopts a system of five lines, albeit producing a different shape.

The image below from the [Dancing Crayon Designs](#) website uses glyphs whose strokes are that of pencil shapes, which helps to bolster the typographic compatibility argument for encoding TALLY MARK ONE:

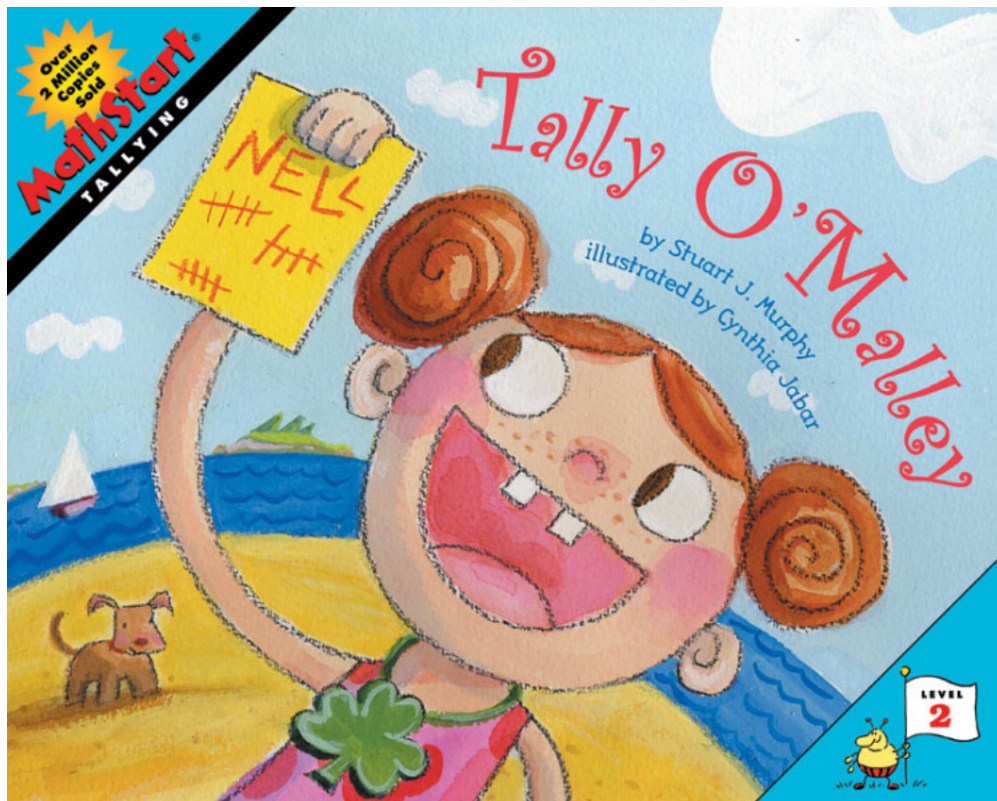




The image below is from an article from [The Best of Teacher Entrepreneurs III](#) website that provides additional typographic compatibility evidence for encoding TALLY MARK ONE:



Finally, [Tally O'Malley, MathStart 2](#) by Stuart J. Murphy (HarperCollins Publishers, 2004) is an entire book dedicated to teaching tally marks:



## Conclusion

Although tally marks are often written by hand, which is true of tens of thousands of other characters already in Unicode, the printed evidence presented in this proposal demonstrates their use as genuine characters in running text or in tables, and with properties that indicate their numeric values. This strongly suggests a need to represent them as characters in “plain text” environments. The tally mark system that is described in this proposal is unique, and forms a well-established system of counting or keeping score.

Although there appears to be no previous formal proposal to encode tally marks, other than my previous proposals (see [L2/15-328](#) and [L2/16-046](#)), there was a lengthy discussion on the [Unicode mailing list](#) over nine years ago, back in February and March of 2007, which began with a post from William J. Poser. No particular conclusion was reached.

That is all.

**ISO/IEC JTC 1/SC 2/WG 2  
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS  
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646<sup>1</sup>**

**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**

<b>1. Title:</b>	<b>Tally Marks</b>
2. Requester's name:	<i>Ken Lunde &amp; Daisuke MIURA</i>
3. Requester type (Member body/Liaison/Individual contribution):	<i>Individual</i>
4. Submission date:	<i>2016-03-14</i>
5. Requester's reference (if applicable):	
6. Choose one of the following:	
This is a complete proposal:	<input checked="" type="checkbox"/>
(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"/>
Proposed name of script:	
b. The proposal is for addition of character(s) to an existing block:	<input checked="" type="checkbox"/>
Name of the existing block:	<i>Counting Rod Numerals</i>
2. Number of characters in proposal:	<i>2</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?	
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<input checked="" type="checkbox"/>
b. Are the character shapes attached in a legible form suitable for review?	<input checked="" type="checkbox"/>
5. Fonts related:	
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	
	<i>Ken Lunde, Adobe Systems Incorporated</i>
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	
	<i>Adobe Systems Incorporated, 345 Park Avenue, San Jose, CA 95110 USA</i>
6. References:	
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<input checked="" type="checkbox"/>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<input checked="" type="checkbox"/>
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"/>

**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.

<sup>1</sup> 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? If YES explain	<i>See text</i>	<i>Yes</i>
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:	<i>See text</i> <i>See text</i>	<i>Yes</i>
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference:	<i>See text</i>	<i>Yes</i>
4. The context of use for the proposed characters (type of use; common or rare) Reference:	<i>See text</i>	<i>common</i>
5. Are the proposed characters in current use by the user community? If YES, where? Reference:	<i>See text</i>	<i>Yes</i>
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:		<i>No</i>
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?		<i>Yes</i>
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:		<i>No</i>
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:		<i>No</i>
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? If YES, is a rationale for its inclusion provided? If YES, reference:	<i>See text</i>	<i>Yes</i> <i>Yes</i>
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:		<i>No</i>
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)		<i>No</i>
13. Does the proposal contain any Ideographic compatibility characters? If YES, are the equivalent corresponding unified ideographic characters identified? If YES, reference:		<i>No</i>