

**Title:** Proposal to Redefine the Long Property Alias STerm to Sentence\_Terminal  
**Source:** Laurențiu Iancu (Microsoft Corporation) and Ken Whistler  
**Status:** Individual contribution  
**Action:** For consideration by the Unicode Technical Committee  
**Date:** 2015-10-09

## Abstract

The UCD STerm property indicates whether a character is a punctuation mark that appears in sentence terminal position. As of Unicode 8.0, both the long and the abbreviated names of that property are STerm. As a long name, STerm is not self-descriptive and leads to a rather awkward definition of the Sentence\_Break property value STerm [UAX29]. This document proposes to redefine the long name alias to a better suited term.

## Proposal statement

Redefine the long name alias of the STerm property from STerm to Sentence\_Terminal. Expressed in the syntax used in PropertyAliases.txt, the proposed change is from

```
STerm                ; STerm
to
STerm                ; Sentence_Terminal
```

## Rationale

The STerm property is an informative, binary property in the UCD. Characters assigned the STerm property are punctuation marks (General\_Category=Other\_Punctuation) used in closing a sentence or equivalent span of text. The STerm property is used in the derivation of the Sentence\_Break property value STerm [UAX29]. The set of characters with the STerm property ([Unicode set](#)) is strictly included in the set of characters with the broader Terminal\_Punctuation property [14-010].

As a long property name, STerm has disadvantages:

1. It is not self-descriptive, as it should be according to UAX #44: *"The long symbolic name alias is self-descriptive, and is treated as the official name of a Unicode character property. For clarity it is used whenever possible when referring to that property in [...] the Unicode Standard."* [UAX44]. The lack of a self-descriptive trait is compensated for in the description of the STerm property in UAX #44. Cf. the description of Terminal\_Punctuation:

<a href="#">STerm</a>	B	I	Sentence Terminal. Used in Unicode...
<a href="#">Terminal_Punctuation</a>	B	I	Punctuation characters that generally mark...

2. Although the names of properties and property values belong to different namespaces, the use of STerm as a property name in the definition of the Sentence\_Break property value STerm is rather awkward [UAX29]:

<u>STerm</u>	STerm = Yes
--------------	-------------

These disadvantages can be remedied by redefining the long alias STerm to Sentence\_Terminal. Such a change is possible because there are no formal stability policies to prevent it. It also seems preferable compared to keeping the long alias as is and introducing an additional alias. Because the current long alias is identical to the short alias of the property, the short alias STerm will continue to match after redefining the long alias to be Sentence\_Terminal.

## References

- [14-010] Laurențiu Iancu, Ken Whistler, *Proposal to make STerm a proper subset of Terminal\_Punctuation*, L2/14-010R, February 2014, <http://www.unicode.org/L2/L2014/14010r-terminal-punct.pdf>.
- [UAX29] Mark Davis, Laurențiu Iancu, *Unicode® Standard Annex #29: Unicode Text Segmentation*, Version 8.0, June 2015, <http://www.unicode.org/reports/tr29/tr29-27.html>.
- [UAX44] Mark Davis, Laurențiu Iancu, Ken Whistler, *Unicode® Standard Annex #44: Unicode Character Database*, Version 8.0, June 2015, <http://www.unicode.org/reports/tr44/tr44-16.html>.