PRECIS
Internet-Draft
Intended status: Standards Track
Expires: April 22, 2013

P. Saint-Andre
Cisco Systems, Inc.

October 19, 2012

Preparation and Comparison of Nicknames
draft-ietf-precis-nickname-04

Abstract

This document describes how to prepare and compare Unicode strings representing nicknames, primarily for use within textual chatrooms. This profile is intended to be used by messaging and text conferencing technologies such as the Extensible Messaging and Presence Protocol (XMPP), the Message Session Relay Protocol (MSRP), and Centralized Conferencing (XCON).

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at http://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 22, 2013.

Table of Contents

1. Introduction

1.1. Overview

Technologies for textual chatrooms customarily enable participants to specify a nickname for use in the room; e.g., this is true of Internet Relay Chat [RFC2811] as well as multi-party chat technologies based on the Extensible Messaging and Presence Protocol (XMPP) [RFC6120] [XEP-0045], the Message Session Relay Protocol (MSRP) [RFC4975] [I-D.ietf-simple-chat], and Centralized Conferencing (XCON) [RFC5239] [I-D.boulton-xcon-session-chat]. Recent chatroom technologies also allow internationalized nicknames because they support characters from outside the ASCII range [RFC20], typically by means of the Unicode character set [UNICODE]. Although such nicknames tend to be used primarily for display purposes, they are sometimes used for programmatic purposes as well (e.g., kicking users or avoiding nickname conflicts). Note too that nicknames can be used not only in chatrooms but also more generally as a user's preferred display name (see for instance [XEP-0172]).

To increase the likelihood that nicknames will work in ways that make sense for typical users throughout the world, this document defines rules for preparing and comparing internationalized nicknames.

1.2. Terminology

Many important terms used in this document are defined in [I-D.ietf-precis-framework], [RFC6365], and [UNICODE].

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Rules

A nickname MUST consist only of Unicode code points that conform to the "FreeClass" base string class defined in [I-D.ietf-precis-framework].

For preparation purposes (most commonly, when a chatroom client generates a nickname from user input for inclusion as a protocol element that represents a "nickname slot"), an application MUST at a minimum ensure that the string conforms to the "FreeClass" base string class defined in [I-D.ietf-precis-framework]; however, it MAY in addition perform the normalization and mapping operations specified below for comparison purposes.

For comparison purposes (e.g., when a chatroom server determines if two nicknames are in conflict during the authorization process), an application MUST treat a nickname as follows, where the operations specified MUST be completed in the order shown (in particular, normalization MUST be performed before all other mapping steps and validity checks, consistent with [I-D.ietf-precis-framework]):

1. The string MUST be normalized using Unicode Normalization Form KC (NFKC). Because NFKC is more "aggressive" in finding matches than other normalization forms (in the terminology of Unicode, it performs both canonical and compatibility decomposition before recomposing code points), this rule helps to reduce the possibility of confusion by increasing the number of characters that would match (e.g., U+2163 ROMAN NUMERAL FOUR would match the combination of U+0049 LATIN CAPITAL LETTER I and U+0056 LATIN CAPITAL LETTER V).

2. Uppercase and titlecase characters MUST be mapped to their lowercase equivalents. In applications that prohibit conflicting nicknames, this rule helps to reduce the possibility of confusion by ensuring that nicknames differing only by case (e.g., "stpeter" vs. "StPeter") would not be allowed in a chatroom at the same time.

3. Non-ASCII space characters from the "N" category defined under Section 6.14 of [I-D.ietf-precis-framework] MUST be mapped to U+0020 SPACE.

4. Other mappings MAY be applied, such as those defined in [I-D.yoneya-precis-mappings]. (Note that mapping of fullwidth and halfwidth characters to their decomposition mappings is not necessary, since those mappings are performed as part of normalization using NFKC.)

For both preparation and comparison, the "Bidi Rule" defined in [RFC5893] applies to the directionality of a nickname.

3. Use in Application Protocols

This specification defines only the PRECIS-based rules for handling of nickname strings. It is the responsibility of an application protocol (e.g., MSRP, XCON, or XMPP) to specify the protocol slots in which nickname strings can appear, as well as the entities that are expected to enforce the rules governing nickname strings in that protocol (e.g., chat servers, chat clients, or both). Above and beyond the PRECIS-based rules specified here, application protocols can also define application-specific rules governing nickname strings

(rules regarding the minimum or maximum length of nicknames, further restrictions on allowable characters or character ranges, safeguards to mitigate the effects of visually similar characters, etc.). Naturally, application protocols can also specify rules governing the actual use of nicknames in applications (reserved nicknames, authorization requirements for using nicknames, whether certain nicknames can be prohibited, handling of duplicates, the relationship between nicknames and underlying identifiers such as SIP URIs or Jabber Identifiers, etc.).

4. IANA Considerations

The IANA shall add the following entry to the PRECIS Usage Registry:

Applicability: Nicknames in messaging and text conferencing technologies such as XMPP, MSRP, and XCON.
Base Class: FreeClass.
Subclass: No.
Normalization: NFKC.
Casemapping: Map uppercase and titlecase characters to lowercase.
Additional Mappings: Map non-ASCII space characters to ASCII space.
Directionality: The "Bidi Rule" defined in RFC 5893 applies.
Specification: RFC XXXX. [Note to RFC Editor: please change XXXX to the number issued for this specification.]

5. Security Considerations

5.1. Reuse of PRECIS

The security considerations described in [I-D.ietf-precis-framework] apply to the "FreeClass" base string class used in this document for nicknames.

5.2. Reuse of Unicode

The security considerations described in [UTR39] apply to the use of Unicode characters in nicknames.

5.3. Visually Similar Characters

Section 10.5 of [I-D.ietf-precis-framework] describes some of the security considerations related to visually similar characters, also called "confusable characters" or "confusables".

Although the mapping rules defined under Section 2 of this document are designed in part to reduce the possibility of confusion about

nicknames, this document does not provide more detailed recommendations regarding the handling of visually similar characters, such as those in [UTR39].

6. References

6.1. Normative References

[I-D.ietf-precis-framework]
Saint-Andre, P. and M. Blanchet, "Precis Framework: Handling Internationalized Strings in Protocols", draft-ietf-precis-framework-06 (work in progress), September 2012.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC5893] Alvestrand, H. and C. Karp, "Right-to-Left Scripts for Internationalized Domain Names for Applications (IDNA)", RFC 5893, August 2010.

[UNICODE] The Unicode Consortium, "The Unicode Standard, Version 6.1", 2012, <http://www.unicode.org/versions/Unicode6.1.0/>.

[UTR39] The Unicode Consortium, "Unicode Technical Report #39: Unicode Security Mechanisms", August 2010, <http://unicode.org/reports/tr39/>.

6.2. Informative References

[I-D.boulton-xcon-session-chat]
Barnes, M., Boulton, C., and S. Loreto, "Chatrooms within a Centralized Conferencing (XCON) System", draft-boulton-xcon-session-chat-08 (work in progress), July 2011.

[I-D.ietf-simple-chat]
Niemi, A., Garcia, M., and G. Sandbakken, "Multi-party Chat Using the Message Session Relay Protocol (MSRP)", draft-ietf-simple-chat-16 (work in progress), August 2012.

[I-D.yoneya-precis-mappings]
YONEYA, Y. and T. NEMOTO, "Mapping characters for PRECIS classes", draft-yoneya-precis-mappings-03 (work in progress), October 2012.

[RFC20] Cerf, V., "ASCII format for network interchange", RFC 20, October 1969.

[RFC2811] Kalt, C., "Internet Relay Chat: Channel Management", RFC 2811, April 2000.

[RFC4975] Campbell, B., Mahy, R., and C. Jennings, "The Message Session Relay Protocol (MSRP)", RFC 4975, September 2007.

[RFC5239] Barnes, M., Boulton, C., and O. Levin, "A Framework for Centralized Conferencing", RFC 5239, June 2008.

[RFC6120] Saint-Andre, P., "Extensible Messaging and Presence Protocol (XMPP): Core", RFC 6120, March 2011.

[RFC6365] Hoffman, P. and J. Klensin, "Terminology Used in Internationalization in the IETF", BCP 166, RFC 6365, September 2011.

[XEP-0045] Saint-Andre, P., "Multi-User Chat", XSF XEP 0045, February 2012.

[XEP-0172] Saint-Andre, P. and V. Mercier, "User Nickname", XSF XEP 0172, March 2012.

Appendix A. Acknowledgements

Thanks to Kim Alvefur, Mary Barnes, Dave Cridland, Miguel Garcia, Salvatore Loreto, and Enrico Marocco for their reviews and comments.

Author's Address

Peter Saint-Andre
Cisco Systems, Inc.
1899 Wynkoop Street, Suite 600
Denver, CO 80202
USA

Phone: +1-303-308-3282
Email: psaintan@cisco.com