

Internet Draft  
[draft-ietf-krb-wg-utf8-profile-01.txt](#)  
February 24, 2003  
Expires in six months

Preparation of Internationalized Strings Profile for Kerberos UTF-8 Strings  
J. Altman

Status of this memo

This document is an Internet-Draft and is in full conformance with all provisions of [Section 10 of RFC2026](#).

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

To view the list Internet-Draft Shadow Directories, see <http://www.ietf.org/shadow.html>.

## Abstract

This document describes how to prepare UTF-8 strings for use with Kerberos protocols in order to increase the likelihood that name input and name comparison work in ways that make sense for typical users throughout the world. This is a profile of "Preparation of Internationalized Strings" [[RFC3454](#)].

## **1. Introduction**

This document specifies processing rules that will allow users to enter Kerberos Principal Names and input to cryptographic String to Key functions. It is a profile of stringprep [[RFC3454](#)].

This profile defines the following, as required by [[RFC3454](#)]

- The intended applicability of the profile: internationalized host name parts
- The character repertoire that is the input and output to stringprep: defined in [Section 2](#)
- The list of unassigned code points for the repertoire: defined in [Appendix D](#).

- The mappings used: defined in [Section 3](#).
- The Unicode normalization used: defined in [Section 4](#)
- The characters that are prohibited as output: Defined in [section 5](#)

## **[1.2 Terminology](#)**

The key words "MUST", "SHALL", "REQUIRED", "SHOULD", "RECOMMENDED", and "MAY" in this document are to be interpreted as described in [RFC 2119](#) [[RFC2119](#)].

Examples in this document use the notation for code points and names from the Unicode Standard [Unicode3.1] and ISO/IEC 10646 [[ISO10646](#)]. For example, the letter "a" may be represented as either "U+0061" or "LATIN SMALL LETTER A". In the lists of prohibited characters, the "U+" is left off to make the lists easier to read. The comments for character ranges are shown in square brackets (such as "[SYMBOLS]") and do not come from the standards.

## **[2. Character Repertoire](#)**

Unicode 3.2 [[Unicode3.2](#)] is the repertoire used in this profile. The reason Unicode 3.2 was chosen instead of a version of ISO/IEC 10646 is that Unicode 3.2 is the basis for [[RFC3454](#)].

## **[3. Mapping](#)**

This profile specifies stringprep mapping using the mapping table in [Appendix C](#). That table includes all the steps described in this section.

Note that text in this section describe how [Appendix C](#) was formed. It is there for people who want to understand more, but it should be ignored by implementors. Implementations of this profile MUST map based on Appendix C, not based on the descriptions in this section of how Appendix C was created.

### **[3.1 Mapped to nothing](#)**

The following characters are simply deleted from the input (that is, they are mapped to nothing) because their presence or absence should not make two strings different.

Some characters are only useful in line-based text, and are otherwise invisible and ignored.

00AD; SOFT HYPHEN

1806; MONGOLIAN TODO SOFT HYPHEN

200B; ZERO WIDTH SPACE  
2060; WORD JOINER  
FEFF; ZERO WIDTH NO-BREAK SPACE

Variation selectors and cursive connectors select different glyphs, but do not bear semantics.

034F; COMBINING GRAPHEME JOINER  
180B; MONGOLIAN FREE VARIATION SELECTOR ONE  
180C; MONGOLIAN FREE VARIATION SELECTOR TWO  
180D; MONGOLIAN FREE VARIATION SELECTOR THREE  
200C; ZERO WIDTH NON-JOINER  
200D; ZERO WIDTH JOINER  
FE00; VARIATION SELECTOR-1  
FE01; VARIATION SELECTOR-2  
FE02; VARIATION SELECTOR-3  
FE03; VARIATION SELECTOR-4  
FE04; VARIATION SELECTOR-5  
FE05; VARIATION SELECTOR-6  
FE06; VARIATION SELECTOR-7  
FE07; VARIATION SELECTOR-8  
FE08; VARIATION SELECTOR-9  
FE09; VARIATION SELECTOR-10  
FE0A; VARIATION SELECTOR-11  
FE0B; VARIATION SELECTOR-12  
FE0C; VARIATION SELECTOR-13  
FE0D; VARIATION SELECTOR-14  
FE0E; VARIATION SELECTOR-15  
FE0F; VARIATION SELECTOR-16

### **3.2 Space Character Conversions**

Space characters can make accurate visual transcription of names nearly impossible and could lead to user entry errors in many ways. The following Unicode spaces are to be mapped to 0020; SPACE:

0020; SPACE  
00A0; NO-BREAK SPACE  
1680; OGHAM SPACE MARK  
2000; EN QUAD  
2001; EM QUAD  
2002; EN SPACE  
2003; EM SPACE  
2004; THREE-PER-EM SPACE  
2005; FOUR-PER-EM SPACE  
2006; SIX-PER-EM SPACE  
2007; FIGURE SPACE  
2008; PUNCTUATION SPACE  
2009; THIN SPACE  
200A; HAIR SPACE  
202F; NARROW NO-BREAK SPACE

205F; MEDIUM MATHEMATICAL SPACE  
3000; IDEOGRAPHIC SPACE

#### **4. Normalization**

This profile specifies using Unicode normalization form KC, as described in [\[UAX15\]](#).

NOTE: There was some discussion on the mailing list that would suggest that Unicode NFKC does not properly handle the composition of normalized Hangul strings. Following the lead of the IDN working group, the Kerberos working group will not attempt to second-guess the the authors of Unicode 3.1 Annex 15 (formerly Technical Report 15) [\[UAX15\]](#), which specifies the normalization methods, or the Ideographic Rappaorteur Group (IRG), which is the formal subgroup of ISO/IEC JTC1/SC2/WG2 charged with approving all CJKV elements of the Unicode standards. Such issues are outside the working group's charter and its area of expertise.

#### **5. Prohibited Output**

This profile specifies using the prohibition table in [Appendix D](#).

Note that the subsections below describe how [Appendix D](#) was formed. They are there for people who want to understand more, but they should be ignored by implementors. Implementations of this profile MUST map based on [Appendix D](#), not based on the descriptions in this section of how Appendix D was created.

The collected lists of prohibited code points can be found in [Appendix D](#) of this document. The lists in [Appendix D](#) MUST be used by implementations of this specification. If there are any discrepancies between the lists in [Appendix D](#) and subsections below, the lists in [Appendix D](#) always takes precedence.

Some code points listed in one section would also appear in other sections. Each code point is only listed once in the tables in [Appendix D](#).

##### **5.1 Control characters**

Control characters (or characters with control function) cannot be seen and can cause unpredictable results when displayed.

0000-001F; [CONTROL CHARACTERS]  
007F; DELETE  
0080-009F; [CONTROL CHARACTERS]  
06DD; ARABIC END OF AYAH  
070F; SYRIAC ABBREVIATION MARK  
180E; MONGOLIAN VOWEL SEPARATOR  
200C; ZERO WIDTH NON-JOINER

200D; ZERO WIDTH JOINER  
2028; LINE SEPARATOR  
2029; PARAGRAPH SEPARATOR  
2060; WORD JOINER  
2061; FUNCTION APPLICATION  
2062; INVISIBLE TIMES  
2063; INVISIBLE SEPARATOR  
206A-206F; [CONTROL CHARACTERS]  
FEFF; ZERO WIDTH NO-BREAK SPACE  
FFF9-FFFC; [CONTROL CHARACTERS]  
1D173-1D17A; [MUSICAL CONTROL CHARACTERS]

## **5.2 Private use and replacement characters**

Because private-use characters do not have defined meanings, they are prohibited. The private-use characters are:

E000-F8FF; [PRIVATE USE, PLANE 0]  
F0000-FFFFD; [PRIVATE USE, PLANE 15]  
100000-10FFFD; [PRIVATE USE, PLANE 16]

## **5.3 Non-character code points**

Non-character code points are code points that have been allocated in ISO/IEC 10646 but are not characters. Because they are already assigned, they are guaranteed not to later change into characters.

FDD0-FDEF; [NONCHARACTER CODE POINTS]  
FFFE-FFFF; [NONCHARACTER CODE POINTS]  
1FFFE-1FFFF; [NONCHARACTER CODE POINTS]  
2FFFE-2FFFF; [NONCHARACTER CODE POINTS]  
3FFFE-3FFFF; [NONCHARACTER CODE POINTS]  
4FFFE-4FFFF; [NONCHARACTER CODE POINTS]  
5FFFE-5FFFF; [NONCHARACTER CODE POINTS]  
6FFFE-6FFFF; [NONCHARACTER CODE POINTS]  
7FFFE-7FFFF; [NONCHARACTER CODE POINTS]  
8FFFE-8FFFF; [NONCHARACTER CODE POINTS]  
9FFFE-9FFFF; [NONCHARACTER CODE POINTS]  
AFFFE-AFFFF; [NONCHARACTER CODE POINTS]  
BFFFE-BFFFF; [NONCHARACTER CODE POINTS]  
CFFFE-CFFFF; [NONCHARACTER CODE POINTS]  
DFFFE-DFFFF; [NONCHARACTER CODE POINTS]  
EFFFE-EFFFF; [NONCHARACTER CODE POINTS]  
FFFFE-FFFFF; [NONCHARACTER CODE POINTS]  
10FFFE-10FFFF; [NONCHARACTER CODE POINTS]

The non-character code points are listed the PropList.txt file from the Unicode database.

## **5.4 Surrogate codes**

The following code points are permanently reserved for use as surrogate code values in the UTF-16 encoding, will never be assigned to characters, and are therefore prohibited:

D800-DFFF; [SURROGATE CODES]

### **5.5 Inappropriate for plain text**

The following characters should not appear in regular text.

FFF9; INTERLINEAR ANNOTATION ANCHOR  
FFFA; INTERLINEAR ANNOTATION SEPARATOR  
FFFB; INTERLINEAR ANNOTATION TERMINATOR  
FFFC; OBJECT REPLACEMENT CHARACTER

Although the replacement character (U+FFFD) might be used when a name is displayed, it doesn't make sense for it to be part of the name itself. It is often displayed by renderers to indicate "there would be some character here, but it cannot be rendered". For example, on a computer with no Asian fonts, a name with three ideographs might be rendered with three replacement characters.

FFFD; REPLACEMENT CHARACTER

### **5.6 Inappropriate for canonical representation**

The ideographic description characters allow different sequences of characters to be rendered the same way, which makes them inappropriate for host names that must have a single canonical representation.

2FF0-2FFB; [IDEOGRAPHIC DESCRIPTION CHARACTERS]

### **5.7 Change display properties**

The following characters can cause changes in display or the order in which characters appear when rendered, or are deprecated in Unicode.

0340; COMBINING GRAVE TONE MARK  
0341; COMBINING ACUTE TONE MARK  
200E; LEFT-TO-RIGHT MARK  
200F; RIGHT-TO-LEFT MARK  
202A; LEFT-TO-RIGHT EMBEDDING  
202B; RIGHT-TO-LEFT EMBEDDING  
202C; POP DIRECTIONAL FORMATTING  
202D; LEFT-TO-RIGHT OVERRIDE  
202E; RIGHT-TO-LEFT OVERRIDE  
206A; INHIBIT SYMMETRIC SWAPPING  
206B; ACTIVATE SYMMETRIC SWAPPING  
206C; INHIBIT ARABIC FORM SHAPING  
206D; ACTIVATE ARABIC FORM SHAPING  
206E; NATIONAL DIGIT SHAPES  
206F; NOMINAL DIGIT SHAPES

## **5.8 Tagging characters**

The following characters are used for tagging text and are invisible.

E0001; LANGUAGE TAG  
E0020-E007F; [TAGGING CHARACTERS]

## **6. Bidirectional Characters**

This profile specifies checking bidirectional strings as described in [\[RFC3454\] section 6](#).

## **7. Unassigned Code Points**

This profile lists the unassigned code points for Unicode 3.2 in [Appendix E](#). The list in [Appendix E](#) **MUST** be used by implementations of this specification. If there are any discrepancies between the list in Appendix E and the Unicode 3.2 specification, the list [Appendix E](#) always takes precedence.

## **8. Security Considerations**

ISO/IEC 10646 has many characters that look similar. In many cases, users of security protocols might do visual matching, such as when comparing the names of trusted third parties. This profile does nothing to map similar-looking characters together.

Principal names and passwords are entered by users and used within the Kerberos protocol. The security of the Internet would be compromised if a user entering a single internationalized string could be connected to different servers or denied access based on different interpretations of internationalized strings.

## **9. IANA Considerations**

IANA is to register this profile as described in [\[RFC3454\]](#).

## **10. References**

[CharModel] Unicode Technical Report;17, Character Encoding Model.  
<<http://www.unicode.org/unicode/reports/tr17/>>.

[Glossary] Unicode Glossary, <<http://www.unicode.org/glossary/>>.

[ISO10646] ISO/IEC 10646-1:2000. International Standard -- Information technology -- Universal Multiple-Octet Coded Character Set (UCS) -- Part 1: Architecture and Basic Multilingual Plane.

[RFC2119] Scott Bradner, "Key words for use in RFCs to Indicate Requirement Levels", March 1997, [RFC 2119](#).

[RFC3454] Paul Hoffman and Marc Blanchet, "Preparation of Internationalized Strings ("stringprep")", [draft-hoffman-stringprep](#), work in progress

[Unicode3.2] The Unicode Standard, Version 3.2.0: The Unicode Consortium. The Unicode Standard, Version 3.0. Reading, MA, Addison-Wesley Developers Press, 2000. ISBN 0-201-61633-5, as amended by: Unicode Standard Annex #27: Unicode 3.1  
<<http://www.unicode.org/unicode/reports/tr27/>>; and  
by: Unicode Standard Annex #28: Unicode 3.2  
<<http://www.unicode.org/unicode/reports/tr28/>>

[UAX9] The Unicode Consortium. Unicode Standard Annex #9, The Bidirectional Algorithm, <<http://www.unicode.org/unicode/reports/tr9/>>.

[UAX15] Mark Davis and Martin Duerst. Unicode Standard Annex #15: Unicode Normalization Forms, Version 3.1.0.  
<<http://www.unicode.org/unicode/reports/tr15/tr15-21.html>>

## **[A. Acknowledgements](#)**

This draft is based upon the work of the IETF IDN Working Group's IDN Nameprep design team.

This profile is the work of the Kerberos Working Group. Significant contributions were provided by Jeffrey Hutzelman, Sam Hartman, Tom Yu, Ken Raeburn, and Jeffrey Altman.

## **[B. Editor Contact Information](#)**

Jeffrey Altman  
Internet Access Methods  
**200 W 72ND ST STE 48**  
NEW YORK NY 10024  
+1 (212) 580-2700  
e-mail: [jaltman@iamx.com](mailto:jaltman@iamx.com)

## **[C. Mapping Tables](#)**

The following is the mapping table from [Section 3](#). The table has three columns:

- the character that is mapped from
- the zero or more characters that it is mapped to
- the reason for the mapping

The columns are separated by semicolons. Note that the second column may be empty, or it may have one character, or it may have more than one character, with each character separated by a space.



```

----- Start Mapping Table -----
00A0; 0020; NO-BREAK SPACE
00AD; ; Map to nothing
034F; ; Map to nothing
1680; 0020; OGHAM SPACE MARK
1806; ; Map to nothing
180B; ; Map to nothing
180C; ; Map to nothing
180D; ; Map to nothing
2000; 0020; EN QUAD
2001; 0020; EM QUAD
2002; 0020; EN SPACE
2003; 0020; EM SPACE
2004; 0020; THREE-PER-EM SPACE
2005; 0020; FOUR-PER-EM SPACE
2006; 0020; SIX-PER-EM SPACE
2007; 0020; FIGURE SPACE
2008; 0020; PUNCTUATION SPACE
2009; 0020; THIN SPACE
200A; 0020; HAIR SPACE
200B; ; Map to nothing
200C; ; Map to nothing
200D; ; Map to nothing
2060; ; Map to nothing
202F; 0020; NARROW NO-BREAK SPACE
205F; 0020; MEDIUM MATHEMATICAL SPACE
3000; 0020; IDEOGRAPHIC SPACE
FE00; ; Map to nothing
FE01; ; Map to nothing
FE02; ; Map to nothing
FE03; ; Map to nothing
FE04; ; Map to nothing
FE05; ; Map to nothing
FE06; ; Map to nothing
FE07; ; Map to nothing
FE08; ; Map to nothing
FE09; ; Map to nothing
FE0A; ; Map to nothing
FE0B; ; Map to nothing
FE0C; ; Map to nothing
FE0D; ; Map to nothing
FE0E; ; Map to nothing
FE0F; ; Map to nothing
FEFF; ; Map to nothing
----- End Mapping Table -----

```

#### **D. Prohibited Code Point List**

```

----- Start Prohibited Table -----

```

0000-001F; [CONTROL CHARACTERS]  
007F; DELETE  
0080-009F; [CONTROL CHARACTERS]  
0340; COMBINING GRAVE TONE MARK  
0341; COMBINING ACUTE TONE MARK  
06DD; ARABIC END OF AYAH  
070F; SYRIAC ABBREVIATION MARK  
100000-10FFFD; [PRIVATE USE, PLANE 16]  
10FFFE-10FFFF; [NONCHARACTER CODE POINTS]  
180E; MONGOLIAN VOWEL SEPARATOR  
1D173-1D17A; [MUSICAL CONTROL CHARACTERS]  
1FFFE-1FFFF; [NONCHARACTER CODE POINTS]  
200C; ZERO WIDTH NON-JOINER  
200D; ZERO WIDTH JOINER  
200E; LEFT-TO-RIGHT MARK  
200F; RIGHT-TO-LEFT MARK  
2028; LINE SEPARATOR  
2029; PARAGRAPH SEPARATOR  
202A; LEFT-TO-RIGHT EMBEDDING  
202B; RIGHT-TO-LEFT EMBEDDING  
202C; POP DIRECTIONAL FORMATTING  
202D; LEFT-TO-RIGHT OVERRIDE  
202E; RIGHT-TO-LEFT OVERRIDE  
2060; WORD JOINER  
2061; FUNCTION APPLICATION  
2062; INVISIBLE TIMES  
2063; INVISIBLE SEPARATOR  
206A-206F; [CONTROL CHARACTERS]  
206A; INHIBIT SYMMETRIC SWAPPING  
206B; ACTIVATE SYMMETRIC SWAPPING  
206C; INHIBIT ARABIC FORM SHAPING  
206D; ACTIVATE ARABIC FORM SHAPING  
206E; NATIONAL DIGIT SHAPES  
206F; NOMINAL DIGIT SHAPES  
2FF0-2FFB; [IDEOGRAPHIC DESCRIPTION CHARACTERS]  
2FFFE-2FFFF; [NONCHARACTER CODE POINTS]  
3FFFE-3FFFF; [NONCHARACTER CODE POINTS]  
4FFFE-4FFFF; [NONCHARACTER CODE POINTS]  
5FFFE-5FFFF; [NONCHARACTER CODE POINTS]  
6FFFE-6FFFF; [NONCHARACTER CODE POINTS]  
7FFFE-7FFFF; [NONCHARACTER CODE POINTS]  
8FFFE-8FFFF; [NONCHARACTER CODE POINTS]  
9FFFE-9FFFF; [NONCHARACTER CODE POINTS]  
AFFFE-AFFFF; [NONCHARACTER CODE POINTS]  
BFFFE-BFFFF; [NONCHARACTER CODE POINTS]  
CFFFE-CFFFF; [NONCHARACTER CODE POINTS]  
D800-DFFF; [SURROGATE CODES]  
DFFFE-DFFFF; [NONCHARACTER CODE POINTS]  
E000-F8FF; [PRIVATE USE, PLANE 0]  
E0001; LANGUAGE TAG  
E0020-E007F; [TAGGING CHARACTERS]

EFFFE-EFFFF; [NONCHARACTER CODE POINTS]  
F0000-FFFFD; [PRIVATE USE, PLANE 15]  
FDD0-FDEF; [NONCHARACTER CODE POINTS]  
FEFF; ZERO WIDTH NO-BREAK SPACE  
FFF9-FFFC; [CONTROL CHARACTERS]  
FFF9; INTERLINEAR ANNOTATION ANCHOR  
FFFA; INTERLINEAR ANNOTATION SEPARATOR  
FFFB; INTERLINEAR ANNOTATION TERMINATOR  
FFFC; OBJECT REPLACEMENT CHARACTER  
FFFD; REPLACEMENT CHARACTER  
FFFE-FFFF; [NONCHARACTER CODE POINTS]  
FFFFE-FFFFF; [NONCHARACTER CODE POINTS]  
----- End Prohibited Table -----

NOTE WELL: Software that follows this specification that will be used to check names before they are put in authoritative name servers MUST add all unassigned code points to the list of characters that are prohibited. See [Section 6 of \[RFC3454\]](#) for more details.

#### **E. Unassigned Code Point List**

----- Start Unassigned Table -----  
0221  
0234-024F  
02AE-02AF  
02EF-02FF  
0350-035F  
0370-0373  
0376-0379  
037B-037D  
037F-0383  
038B  
038D  
03A2  
03CF  
03F7-03FF  
0487  
04CF  
04F6-04F7  
04FA-04FF  
0510-0530  
0557-0558  
0560  
0588  
058B-0590  
05A2  
05BA  
05C5-05CF  
05EB-05EF  
05F5-060B

060D-061A  
061C-061E  
0620  
063B-063F  
0656-065F  
06EE-06EF  
06FF  
070E  
072D-072F  
074B-077F  
07B2-0900  
0904  
093A-093B  
094E-094F  
0955-0957  
0971-0980  
0984  
098D-098E  
0991-0992  
09A9  
09B1  
09B3-09B5  
09BA-09BB  
09BD  
09C5-09C6  
09C9-09CA  
09CE-09D6  
09D8-09DB  
09DE  
09E4-09E5  
09FB-0A01  
0A03-0A04  
0A0B-0A0E  
0A11-0A12  
0A29  
0A31  
0A34  
0A37  
0A3A-0A3B  
0A3D  
0A43-0A46  
0A49-0A4A  
0A4E-0A58  
0A5D  
0A5F-0A65  
0A75-0A80  
0A84  
0A8C  
0A8E  
0A92  
0AA9

0AB1  
0AB4  
0ABA-0ABB  
0AC6  
0ACA  
0ACE-0ACF  
0AD1-0ADF  
0AE1-0AE5  
0AF0-0B00  
0B04  
0B0D-0B0E  
0B11-0B12  
0B29  
0B31  
0B34-0B35  
0B3A-0B3B  
0B44-0B46  
0B49-0B4A  
0B4E-0B55  
0B58-0B5B  
0B5E  
0B62-0B65  
0B71-0B81  
0B84  
0B8B-0B8D  
0B91  
0B96-0B98  
0B9B  
0B9D  
0BA0-0BA2  
0BA5-0BA7  
0BAB-0BAD  
0BB6  
0BBA-0BBD  
0BC3-0BC5  
0BC9  
0BCE-0BD6  
0BD8-0BE6  
0BF3-0C00  
0C04  
0C0D  
0C11  
0C29  
0C34  
0C3A-0C3D  
0C45  
0C49  
0C4E-0C54  
0C57-0C5F  
0C62-0C65  
0C70-0C81

0C84  
0C8D  
0C91  
0CA9  
0CB4  
0CBA-0CBD  
0CC5  
0CC9  
0CCE-0CD4  
0CD7-0CDD  
0CDF  
0CE2-0CE5  
0CF0-0D01  
0D04  
0D0D  
0D11  
0D29  
0D3A-0D3D  
0D44-0D45  
0D49  
0D4E-0D56  
0D58-0D5F  
0D62-0D65  
0D70-0D81  
0D84  
0D97-0D99  
0DB2  
0DBC  
0DBE-0DBF  
0DC7-0DC9  
0DCB-0DCE  
0DD5  
0DD7  
0DE0-0DF1  
0DF5-0E00  
0E3B-0E3E  
0E5C-0E80  
0E83  
0E85-0E86  
0E89  
0E8B-0E8C  
0E8E-0E93  
0E98  
0EA0  
0EA4  
0EA6  
0EA8-0EA9  
0EAC  
0EBA  
0EBE-0EBF  
0EC5

0EC7  
0ECE-0ECF  
0EDA-0EDB  
0EDE-0EFF  
0F48  
0F6B-0F70  
0F8C-0F8F  
0F98  
0FBD  
0FCD-0FCE  
0FD0-0FFF  
1022  
1028  
102B  
1033-1035  
103A-103F  
105A-109F  
10C6-10CF  
10F9-10FA  
10FC-10FF  
115A-115E  
11A3-11A7  
11FA-11FF  
1207  
1247  
1249  
124E-124F  
1257  
1259  
125E-125F  
1287  
1289  
128E-128F  
12AF  
12B1  
12B6-12B7  
12BF  
12C1  
12C6-12C7  
12CF  
12D7  
12EF  
130F  
1311  
1316-1317  
131F  
1347  
135B-1360  
137D-139F  
13F5-1400  
1677-167F

169D-169F  
16F1-16FF  
170D  
1715-171F  
1737-173F  
1754-175F  
176D  
1771  
1774-177F  
17DD-17DF  
17EA-17FF  
180F  
181A-181F  
1878-187F  
18AA-1DFF  
1E9C-1E9F  
1EFA-1EFF  
1F16-1F17  
1F1E-1F1F  
1F46-1F47  
1F4E-1F4F  
1F58  
1F5A  
1F5C  
1F5E  
1F7E-1F7F  
1FB5  
1FC5  
1FD4-1FD5  
1FDC  
1FF0-1FF1  
1FF5  
1FFF  
2053-2056  
2058-205E  
2064-2069  
2072-2073  
208F-209F  
20B2-20CF  
20EB-20FF  
213B-213C  
214C-2152  
2184-218F  
23CF-23FF  
2427-243F  
244B-245F  
24FF  
2614-2615  
2618  
267E-267F  
268A-2700



2705  
270A-270B  
2728  
274C  
274E  
2753-2755  
2757  
275F-2760  
2795-2797  
27B0  
27BF-27CF  
27EC-27EF  
2B00-2E7F  
2E9A  
2EF4-2EFF  
2FD6-2FEF  
2FFC-2FFF  
3040  
3097-3098  
3100-3104  
312D-3130  
318F  
31B8-31EF  
321D-321F  
3244-3250  
327C-327E  
32CC-32CF  
32FF  
3377-337A  
33DE-33DF  
33FF  
4DB6-4DFF  
9FA6-9FFF  
A48D-A48F  
A4C7-ABFF  
D7A4-D7FF  
FA2E-FA2F  
FA6B-FAFF  
FB07-FB12  
FB18-FB1C  
FB37  
FB3D  
FB3F  
FB42  
FB45  
FBB2-FBD2  
FD40-FD4F  
FD90-FD91  
FDC8-FDCF  
FDFD-FDFF  
FE10-FE1F

FE24-FE2F  
FE47-FE48  
FE53  
FE67  
FE6C-FE6F  
FE75  
FEFD-FEFE  
FF00  
FFBF-FFC1  
FFC8-FFC9  
FFD0-FFD1  
FFD8-FFD9  
FFDD-FFDF  
FFE7  
FFE7-FFF8  
10000-102FF  
1031F  
10324-1032F  
1034B-103FF  
10426-10427  
1044E-1CFFF  
1D0F6-1D0FF  
1D127-1D129  
1D1DE-1D3FF  
1D455  
1D49D  
1D4A0-1D4A1  
1D4A3-1D4A4  
1D4A7-1D4A8  
1D4AD  
1D4BA  
1D4BC  
1D4C1  
1D4C4  
1D506  
1D50B-1D50C  
1D515  
1D51D  
1D53A  
1D53F  
1D545  
1D547-1D549  
1D551  
1D6A4-1D6A7  
1D7CA-1D7CD  
1D800-1FFFF  
2A6D7-2F7FF  
2FA1E-2FFFF  
30000-3FFFF  
40000-4FFFF  
50000-5FFFF

60000-6FFFFD  
70000-7FFFFD  
80000-8FFFFD  
90000-9FFFFD  
A0000-AFFFFD  
B0000-BFFFFD  
C0000-CFFFFD  
D0000-DFFFFD  
E0000  
E0002-E001F  
E0080-EFFFFD  
----- End Unassigned Table -----