IANA Charset MIB draft−01

Table of Contents

1 Introduction

This IANA Charset MIB [CHARMIB] module defines the single textual convention 'IANACharset'. Once adopted, all future versions of the IANA Charset MIB [CHARMIB] may be machine−generated, whenever the IANA Charset Registry [CHARSET] is updated by IANA staff according to the procedures defined in [RFC2978], using the utility [IANACHAR] described in section 3 of this document (or any other machine−generation method).

It is strongly recommended that future updates to the IANA Charset MIB [CHARMIB] be machine generated (rather than hand−edited) to avoid asynchrony between the IANA Charset Registry [CHARSET] and the IANA Charset MIB [CHARMIB].

Note: Usage questions and comments on this IANA Charset MIB [CHARMIB] should be sent directly to the editor (imcdonald@sharplabs.com).

1.1 Conformance Terminology

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

2 SNMP Network Management Framework

See section 2 'The Internet Standard Management Framework' of [RFC2570].

3 Generation of IANA Charset MIB

Intellectual Property: The C language utility 'ianachar.c' [IANACHAR] is hereby donated by the author (Ira McDonald) to IANA, in perpetuity, free of license or any other restraint.

The [IANACHAR] utility may be used to generate an updated version of the 'IANACharset' textual convention by reading and parsing the plaintext IANA Charset Registry [CHARSET].

This utility parses each charset registration, finding (in order):

```
1)  The 'Name' field (which is saved for a fallback - see below);

2)  The 'MIBenum' field (which contains the IANA-assigned positive
    decimal enum value); and

3)  The (usually present) 'Alias' field that begins with 'cs' (that
    contains the IANA-assigned enum label).  If an 'Alias' field is
    not found, the utility constructs one from the 'Name' field by:

    - Beginning the enum label with a lowercase 'cs' prefix;

    - Copying _only_ alpha/numeric characters from the 'Name' field
      to the enum label (ignoring punctuation, whitespace, etc.).
```

4 Definition of IANA Charset MIB

```
IANA-CHARSET-MIB DEFINITIONS ::= BEGIN
--  http://www.iana.org/assignments/ianacharset-mib

IMPORTS
    MODULE-IDENTITY,
        mib-2
            FROM SNMPv2-SMI
    TEXTUAL-CONVENTION
            FROM SNMPv2-TC;

ianaCharset MODULE-IDENTITY
    LAST-UPDATED  "200210080000Z"
    ORGANIZATION    "IANA"
    CONTACT-INFO    "       Internet Assigned Numbers Authority

                    Postal: ICANN
                            4676 Admiralty Way, Suite 330
                            Marina del Rey, CA 90292

                    Tel:    +1 310 823 9358
                    E-Mail: iana@iana.org"

    DESCRIPTION     "This MIB module defines the IANACharset Textual
                    Convention.  The IANACharset TC is used to specify
                    the encoding of string objects defined in a MIB.

                    Note:  The IANACharset TC, originally defined in
                    RFC 1759, was inaccurately named CodedCharSet.

                    Note:  Best practice is to define new MIB string
                    objects with invariant UTF-8 (RFC 2279) syntax
                    using the SnmpAdminString TC (defined in RFC 2571)
                    in accordance with IETF Policy on Character Sets and
                    Languages (RFC 2277)."

    --  revision history

    REVISION      "200210080000Z" -- October 8, 2002
    DESCRIPTION   "Original version transferred from Printer MIB."

    ::= { mib-2 999 } -- [[to be assigned by IANA - standards track]]

IANACharset ::= TEXTUAL-CONVENTION
    STATUS  current
    DESCRIPTION
        "Specifies an IANA registered 'charset' - coded character set
        (CCS) plus optional character encoding scheme (CES) - terms
        defined in 'IANA Charset Registration Procedures' (RFC 2978).

        Objects of this syntax are used to specify the encoding for
        string objects defined in one or more MIBs.  For example, the
        prtLocalizationCharacterSet, prtInterpreterDefaultCharSetIn, and
        prtInterpreterDefaultCharSetOut objects defined in Printer MIB.

        The current list of 'charset' names and enumerated values
        is contained in the IANA Character Set Registry at:

           http://www.iana.org/assignments/character-sets
```

```
    Enum names are derived from the IANA Charset Registry 'Alias'
    fields that begin with 'cs' (for character set).
    Enum values are derived from the parallel 'MIBenum' fields."
SYNTAX  INTEGER {
    other(1),   -- used if the designated
                -- character set is not currently
                -- registered by IANA
    unknown(2), -- used as a default value
    csASCII(3),
    csISOLatin1(4),
    csISOLatin2(5),
    csISOLatin3(6),
    csISOLatin4(7),
    csISOLatinCyrillic(8),
    csISOLatinArabic(9),
    csISOLatinGreek(10),
    csISOLatinHebrew(11),
    csISOLatin5(12),
    csISOLatin6(13),
    csISOTextComm(14),
    csHalfWidthKatakana(15),
    csShiftJIS(17),
    csEUCPkdFmtJapanese(18),
    csEUCFixWidJapanese(19),
    csISO4UnitedKingdom(20),
    csISO11SwedishForNames(21),
    csISO15Italian(22),
    csISO17Spanish(23),
    csISO21German(24),
    csISO60DanishNorwegian(25),
    csISO69French(26),
    csISO10646UTF1(27),
    csISO646basic1983(28),
    csINVARIANT(29),
    csISO2IntlRefVersion(30),
    csNATSSEFI(31),
    csNATSSEFIADD(32),
    csNATSDANO(33),
    csNATSDANOADD(34),
    csISO10Swedish(35),
    csKSC56011987(36),
    csISO2022KR(37),
    csEUCKR(38),
    csISO2022JP(39),
    csISO2022JP2(40),
    csISO13JISC6220jp(41),
    csISO14JISC6220ro(42),
    csISO16Portuguese(43),
    csISO18Greek7Old(44),
    csISO19LatinGreek(45),
    csISO25French(46),
    csISO27LatinGreek1(47),
    csISO5427Cyrillic(48),
    csISO42JISC62261978(49),
    csISO47BSViewdata(50),
    csISO49INIS(51),
    csISO50INIS8(52),
    csISO51INISCyrillic(53),
    csISO54271981(54),
    csISO57GB1988(56),
```

```
csISO58GB231280(57),
csISO61Norwegian2(58),
csISO70VideotexSupp1(59),
csISO84Portuguese2(60),
csISO85Spanish2(61),
csISO86Hungarian(62),
csISO87JISX0208(63),
csISO88Greek7(64),
csISO89ASMO449(65),
csISO90(66),
csISO91JISC62291984a(67),
csISO92JISC62991984b(68),
csISO93JIS62291984badd(69),
csISO94JIS62291984hand(70),
csISO95JIS62291984handadd(71),
csISO96JISC62291984kana(72),
csISO2033(73),
csISO99NAPLPS(74),
csISO102T617bit(75),
csISO103T618bit(76),
csISO111ECMACyrillic(77),
csa71(78),
csa72(79),
csISO123CSAZ24341985gr(80),
csISO88596E(81),
csISO88596I(82),
csISO128T101G2(83),
csISO88598E(84),
csISO88598I(85),
csISO139CSN369103(86),
csISO141JUSIB1002(87),
csISO143IECP271(88),
csISO146Serbian(89),
csISO147Macedonian(90),
csISO150(91),
csISO151Cuba(92),
csISO6937Add(93),
csISO153GOST1976874(94),
csISO8859Supp(95),
csISO10367Box(96),
csISO158Lap(97),
csISO159JISX02121990(98),
csISO646Danish(99),
csUSDK(100),
csDKUS(101),
csKSC5636(102),
csUnicode11UTF7(103),
csISO2022CN(104),
csUTF8(106),
csISO885914(110),
csISO885916(112),
csGB18030(114),
csUnicode(1000),
csUCS4(1001),
csUnicodeASCII(1002),
csUnicodeLatin1(1003),
csUnicodeIBM1261(1005),
csUnicodeIBM1268(1006),
csUnicodeIBM1276(1007),
csUnicodeIBM1264(1008),
csUnicodeIBM1265(1009),
```

```
csUnicode11(1010),
csSCSU(1011),
csUTF16BE(1013),
csUTF16(1015),
csUTF32(1017),
csUTF32LE(1019),
csWindows30Latin1(2000),
csWindows31Latin1(2001),
csWindows31Latin2(2002),
csWindows31Latin5(2003),
csHPRoman8(2004),
csAdobeStandardEncoding(2005),
csVenturaUS(2006),
csVenturaInternational(2007),
csDECMCS(2008),
csPC850Multilingual(2009),
csPCp852(2010),
csPC8CodePage437(2011),
csPC8DanishNorwegian(2012),
csPC862LatinHebrew(2013),
csPC8Turkish(2014),
csIBMSymbols(2015),
csIBMThai(2016),
csHPLegal(2017),
csHPPiFont(2018),
csHPMath8(2019),
csHPPSMath(2020),
csHPDesktop(2021),
csVenturaMath(2022),
csMicrosoftPublishing(2023),
csWindows31J(2024),
csGB2312(2025),
csBig5(2026),
csMacintosh(2027),
csIBM037(2028),
csIBM038(2029),
csIBM273(2030),
csIBM274(2031),
csIBM275(2032),
csIBM277(2033),
csIBM278(2034),
csIBM280(2035),
csIBM281(2036),
csIBM284(2037),
csIBM285(2038),
csIBM290(2039),
csIBM297(2040),
csIBM420(2041),
csIBM423(2042),
csIBM424(2043),
csIBM500(2044),
csIBM851(2045),
csIBM855(2046),
csIBM857(2047),
csIBM860(2048),
csIBM861(2049),
csIBM863(2050),
csIBM864(2051),
csIBM865(2052),
csIBM868(2053),
csIBM869(2054),
```

```
        csIBM870(2055),
        csIBM871(2056),
        csIBM880(2057),
        csIBM891(2058),
        csIBM903(2059),
        csIBBM904(2060),
        csIBM905(2061),
        csIBM918(2062),
        csIBM1026(2063),
        csIBMEBCDICATDE(2064),
        csEBCDICATDEA(2065),
        csEBCDICCAFR(2066),
        csEBCDICDKNO(2067),
        csEBCDICDKNOA(2068),
        csEBCDICFISE(2069),
        csEBCDICFISEA(2070),
        csEBCDICFR(2071),
        csEBCDICIT(2072),
        csEBCDICPT(2073),
        csEBCDICES(2074),
        csEBCDICESA(2075),
        csEBCDICESS(2076),
        csEBCDICUK(2077),
        csEBCDICUS(2078),
        csUnknown8BiT(2079),
        csMnemonic(2080),
        csMnem(2081),
        csVISCII(2082),
        csVIQR(2083),
        csKOI8R(2084),
        csIBM866(2086),
        csPC775Baltic(2087),
        csKOI8U(2088),
        csIBM00924(2090),
        csIBM01141(2092),
        csIBM01143(2094),
        csIBM01145(2096),
        csIBM01147(2098),
        csIBM01149(2100),
        csIBM1047(2102),
        cswindows1250(2250),
        cswindows1252(2252),
        cswindows1254(2254),
        cswindows1256(2256),
        cswindows1258(2258),
        reserved(3000)
    }
END
```

5 IANA Considerations

IANA should assign a base arc in the 'mgmt' (standards track) OID tree for the 'ianaCharset' MODULE−IDENTITY defined in the IANA Charset MIB [CHARMIB].

Whenever any 'charset' is added to the plaintext IANA Charset Registry [CHARSET], a new version of the IANA Charset MIB [CHARMIB] may be machine−generated with the C language utility [IANACHAR] described in section 3 of this document.

6 Internationalization Considerations

The IANA Charset MIB [CHARMIB] defines the 'IANACharset' textual convention that may be used in a given MIB module to supply explicit character set labels for one or more text string objects defined in that MIB module.

For example, the Printer MIB [RFC1759] defines the three character set label objects 'prtLocalizationCharacterSet' (for description and console strings), 'prtInterpreterDefaultCharSetIn' (for received print job input data), and 'prtIntpreterDefaultCharSetOut' (for processed print job output data).

The IANA Charset MIB [CHARMIB] supports implementation of the best practices specified in "IETF Policy on Character Sets and Languages" [RFC2277].

Note: The use of the 'SnmpAdminString' textual convention [RFC2571], which has a fixed character set of UTF−8 [RFC2779], is strongly encouraged in defining new MIB modules. The IANA Charset MIB [CHARMIB] supports locale−specific MIB objects with variable character sets.

7 Security Considerations

There are no management objects defined in this IANA Charset MIB [CHARMIB]. Only one textual convention (IANACharset) is defined.

8 Acknowledgements

The editor would like to thank: Bert Wijnen (Lucent) for his original suggestion that the 'IANACharset' textual convention should be extracted from Printer MIB v2 (work−in−progress) [RFC1759bis]; Ron Bergman (Hitachi Printing Solutions) and Harry Lewis (IBM) for their many years of effort as editors of Printer MIB v2 (work−in−progress) [RFC1759bis].

9 Normative References

[CHARMIB] IANA Charset MIB (in the future, to be archived at):
http://www.iana.org/assignments/ianacharset−mib

[CHARSET] IANA Charset Registry (currently archived at):
http://www.iana.org/assignments/character−sets

[IANACHAR] IANA Charset MIB Generation Utility (archived at):
http://www.pwg.org/pub/pwg/pmp/tools/ianachar.c −− Uses an IANA Charset MIB template file (archived at): http://www.pwg.org/pub/pwg/pmp/tools/ianachar.dat

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

[RFC2277] Alvestrand. "IETF Policy on Character Sets and Languages", RFC 2277, January 1998.

[RFC2279] Yergeau. "UTF−8, a Transformation of ISO 10646", RFC 2279, January 1998.

[RFC2571] Wijnen, Harrington, Presuhn. "An Architecture for Describing SNMP Network Management Frameworks", RFC 2571, April 1999.

[RFC2978] Freed, Postel. "IANA Charset Registration Procedures", RFC 2978, October 2000.

10 Informative References

[RFC1759] Smith, Wright, Hastings, Zilles, Gyllenskog. "Printer MIB", RFC 1759, March 1995.

[RFC1759bis] Bergman, Lewis, McDonald. "Printer MIB v2" (work−in−progress), <draft−ietf−printmib−mib−info−13.txt>, October 2002.

[RFC2570] Case, Mundy, Partain, Stewart. "Introduction to Version 3 of the Internet−standard Network Management Framework", RFC 2570, April 1999.

11 Author Address


```
Editor:     Ira McDonald
Postal:     High North Inc
            221 Ridge Ave
            Grand Marais, MI  49839
            USA
Tel:        +1 906-494-2434
Email:      imcdonald@sharplabs.com
```


Note: Usage questions and comments on this IANA Charset MIB [CHARMIB] should be sent directly to the editor (imcdonald@sharplabs.com).

12 Full Copyright Statement