Network Working Group
Internet-Draft
Updates: 5234 (if approved)
Intended Status: Standards Track
Expires: October 17, 2015

S. Leonard
Penango, Inc.

April 15, 2015

Comprehensive Core Rules for ABNF
draft-seantek-abnf-more-core-rules-01

Abstract

This document extends the base definition of ABNF (Augmented Backus-Naur Form) to include comprehensive support for certain symbols, namely those in the US-ASCII standard.

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 October 17, 2015.

1. Comprehensive Core Rule Update

Augmented Backus-Naur Form (ABNF) [RFC5234] is a formal syntax that is popular among many Internet specifications. Many Internet documents employ this syntax along with the Core Rules defined in Appendix B.1 of [RFC5234]. However, the Core Rules do not specify many symbols in the US-ASCII range that are also needed by these relying documents, forcing document authors to define them as local rules. Sometimes different documents define these common symbols in different ways, resulting in confusion or incompatibility when the rules are misread or are combined with other sets of rules. This document extends the [RFC5234] to include comprehensive support for certain symbols, namely those in [US-ASCII].

Appendix A of this document is meant as a drop-in replacement for Appendix B.1 of [RFC5234]. I.e., these Core Rules are no more or less useful or normative than those in [RFC5234]. Future document authors should use these rules, and should not attempt to redefine or augment them (except for backwards compatibility with prior documents).

2. IANA Considerations

This document implies no IANA considerations.

3. Security Considerations

Security is truly believed to be irrelevant to this document.

4. Normative References

[US-ASCII] American National Standards Institute, "Coded Character Set -- 7-bit American Standard Code for Information Interchange", ANSI X3.4, 1986.

[RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.

Appendix A. Comprehensive Core Rules

Certain basic rules are in uppercase, such as SP, HTAB, CRLF, DIGIT, ALPHA, etc.

```
      ALPHA          =  %x41-5A / %x61-7A   ; A-Z / a-z

      BIT            =  "0" / "1"

      CHAR           =  %x01-7F
                            ; any 7-bit US-ASCII character,
                            ;  excluding NUL

      CR             =  %x0D
                            ; carriage return

      CRLF           =  CR LF
                            ; Internet standard newline

      CTL            =  %x00-1F / %x7F
                            ; controls

      DIGIT          =  %x30-39
                            ; 0-9

      DQUOTE         =  %x22
                            ; " (Double Quote)

      HEXDIG         =  DIGIT / "A" / "B" / "C" / "D" / "E" / "F"

      HTAB           =  %x09
                            ; horizontal tab

      LF             =  %x0A
                            ; linefeed

      LWSP           =  *(WSP / CRLF WSP)
                            ; Use of this linear-white-space rule
                            ;  permits lines containing only white
                            ;  space that are no longer legal in
                            ;  mail headers and have caused
                            ;  interoperability problems in other
                            ;  contexts.
                            ; Do not use when defining mail
                            ;  headers and use with caution in
                            ;  other contexts.

      OCTET          =  %x00-FF
```

```
                              ; 8 bits of data

        SP              =  %x20

        VCHAR           =  %x21-7E
                              ; visible (printing) characters

        WSP             =  SP / HTAB
                              ; white space

        NUL             =  %d0
        SOH             =  %d1
        STX             =  %d2
        ETX             =  %d3
        EOT             =  %d4
        ENQ             =  %d5
        ACK             =  %d6
        BEL             =  %d7
        BS              =  %d8
        HT              =  %d9   ; also defined as HTAB
        ; LF            =  %d10  ; already defined
        VT              =  %d11
        FF              =  %d12  ; (literally used in every RFC)
        ; CR            =  %d13  ; already defined
        SO              =  %d14
        SI              =  %d15
        DLE             =  %d16
        DC1             =  %d17
        DC2             =  %d18
        DC3             =  %d19
        DC4             =  %d20
        NAK             =  %d21
        SYN             =  %d22
        ETB             =  %d23
        CAN             =  %d24
        EM              =  %d25
        SUB             =  %d26
        ESC             =  %d27
        FS              =  %d28
        GS              =  %d29
        RS              =  %d30
        US              =  %d31
        ; SP            =  %d32  ; already defined
        DEL             =  %d127
```

Author's Address

Sean Leonard
Penango, Inc.
5900 Wilshire Boulevard
21st Floor
Los Angeles, CA 90036
USA

EMail: dev+ietf@seantek.com
URI: http://www.penango.com/