Network Working Group
Request for Comments: 1556
Category: Informational

H. Nussbacher
Israeli Inter-University
Computer Center

December 1993

Handling of Bi-directional Texts in MIME

Status of this Memo

This memo provides information for the Internet community. This memo does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

Abstract

This document describes the format and syntax of the "direction" keyword to be used with bi-directional texts in MIME.

Description

The MIME standards (RFC 1521 and 1522) defined methods for transporting non-ASCII data via a standard RFC822 e-mail system. Specifically, the Content-type field allows for the inclusion of any ISO language such as Arabic (ISO-8859-6) or Hebrew (ISO-8859-8). The problem is that the these two languages are read from right to left and can have bi-directional data such as mixed Hebrew and English on the same line.

Fortunately, ECMA (European Computer Manufacturers Association) has tackled this problem previously and has issued a technical report called "Handling of Bi-Directional Texts". ECMA TR/53, as it is called, was used to update the Standard ECMA-48 which in turn was used as the basis for ISO/IEC 6429 which was adopted under a special "fast track procedure". It is based on this information that a new character set is being defined which will indicate that the bi-directional message is either encoded in implicit mode or explicit mode. The default is visual mode which requires no special character set other than the standard ones previously defined by ISO-8859.

Examples of new character sets for bi-directionality support:

```
Content-type: text/plain; charset=ISO-8859-6-e
Content-type: text/plain; charset=ISO-8859-6-i
Content-type: text/plain; charset=ISO-8859-8-e
Content-type: text/plain; charset=ISO-8859-8-i
```

The "i" suffix refers to implicit mode and the "e" suffix refers to explicit mode.

Implicit

Implicit directionality is a presentation method in which the direction is determined by an algorithm according to the type of characters and their position relative to the adjacent characters and according to their primary direction. The complete algorithm is quite complex and sites wishing to implement it should refer to the ECMA Technical Report for further details.

Explicit

Explicit directionality is a presentation method in which the direction is explicitly defined by using control sequences which are interleaved within the text and are used for direction determination. This presentation method is also defined in ECMA TR/53, which defines three new control functions and updates 22 existing control functions in the ECMA-48 standard.

Visual

Visual directionality is a presentation method that displays text according to the primary display direction only, which is left to right. All text is viewed in the same direction which is the primary display direction. The displaying application is not aware of the contents direction and displays the text as if it were a uni-directional text. The composing application needs to prepare the text in such a way that it will be displayed correctly. No control characters or algorithms are used to determine how the data is to be displayed. This is the simplest of all methods and the default method for use with MIME encoded texts.

References

[ECMA TR/53] Handling of Bi-Directional Texts, European Computer Manufacturers Association, 114 Rue du Rhone, CH-1204, Geneva, Switzerland, June 1992.

[ISO-6429] Information Technology - Control Functions for Coded Character Sets, 3rd edition, December 15, 1992.

[ISO-8859] Information Processing -- 8-bit Single-Byte Coded Graphic Character Sets, Part 6: Arabic alphabet, ISO 8859-6, 1988.

[ISO-8859] Information Processing -- 8-bit Single-Byte Coded Graphic Character Sets, Part 8: Latin/Hebrew alphabet, ISO 8859-8, 1988.

[RFC822] Crocker, D., "Standard for the Format of ARPA Internet Text Messages", STD 11, RFC 822, UDEL, August 1982.

[RFC1521] Borenstein N., and N. Freed, "MIME (Multipurpose Internet Mail Extensions) Part One: Mechanisms for Specifying and Describing the Format of Internet Message Bodies", Bellcore, Innosoft, September 1993.

[RFC1522] Moore K., "MIME Part Two: Message Header Extensions for Non-ASCII Text", University of Tennessee, September 1993.

Security Considerations

Security issues are not discussed in this memo.

Author's Address

Hank Nussbacher
Computer Center
Tel Aviv University
Ramat Aviv
Israel

Fax: +972 3 6409118
Phone: +972 3 6408309
EMail: hank@vm.tau.ac.il