

Title: Zones

Status: Expert Contribution

Source: Asmus Freytag and John Clews

Action: For consideration by WG2

Distribution: Members of SC2/WG2

ISO/IEC 10646 in clause 4.33 defines a Zone as a sequence of rows for characters with a common characteristic. In Clause 8, 4 zones are defined in the original edition:

Alphabetic
Ideographic
Open
Restricted

No zones were ever defined outside the BMP.

One of the first amendments introduced the S -Zone, divided into the high-half and low-half zones.

Clause 10 defines a Restricted Use zone. There is an implied relation between R-Zone and Restricted Use zone "the R-Zone shall be used for the Restricted Use Zone," but they are not explicitly made identical, leaving open the question of what exactly is the Restricted Use zone.

The Contents of the Restricted Use zone are listed as

- a) Private Use characters
- b) Presentation Forms
- c) Compatibility characters

If the R-Zone and the Restricted Use zone are identical, the Restricted Use zone also contains the Specials (FFFC and FFFD, as well as FEFF) which are not members of a) b), or c). There is no information given on which characters in the R-Zone are of type b) or c). In addition, many characters outside the R-Zone are now recognized as having been coded "to provide compatibility with other standards."

Reading the text it becomes apparent that the original intent of the zones was mainly for guiding future allocation, rather than to give definite information about the characteristics of characters to the implementer. The former role has now been subsumed successfully, and more detailed by the road maps.

The latter role is served by the more specific information elsewhere in the standard. The current text, suggesting that no characters outside the R-Zone can be presentation forms or compatibility characters is highly misleading.

With FPDAM 17, ideographs are coded outside the I-Zone as currently defined. This motivates a closer look at the zone concept. Also now that the reserved areas of the BMP are being filled, it makes little sense to retain an O-Zone.

In this expert's opinion, it would do no harm to remove the zone allocation in clause 8 that attempts to cover the entire BMP. However, the S-Zone, with its high-half and low-half zone could be retained.

The Restricted Use Zone, as defined today includes characters (from the specials block) which are neither Private use characters, Compatibility character or Presentation forms, which are the three types of characters that (according to clause 10) form the Restricted Use Zone.

Furthermore, contrary to what section 10 implies, there are many characters in the BMP that are compatibility characters, but exist outside the R-Zone.

PART 1: SUGGESTED EDITORIAL CORRECTIONS REGARDING ZONES IN UCS

The zones listed in Clause 8, and in the relevant diagrams in Clause 4.38, clause 8 and Figure 2 etc are now incorrect.

For Section 4.39, I propose that this section be deleted.

Zones do not play a functional role, nor are they a functional class in UCS - their presence or absence in the standard is not relevant, except in clause 10: The restricted use zone, which is also incorrect as in section 10.3 the class of characters described are also outside of the Restricted use zone.

For section 10, this anomaly might be overcome by using text something like the following:

10 Characters with special uses

Some graphic characters (predominantly those in the range D800-FFFF) are used in particular ways. These characters include:

(a) Private use... (text to continue as is through to the end of section 10.3).

It may be useful to add a further note that:

In addition to compatibility characters included in the range D800-FFFF, there are also further East Asian compatibility characters listed in the range 2F00-D7FF (unless the CJK compatibility ideographs in F9-FA are different to CJK compatibility characters listed among those in 2F00-D7FF).

For section 8, the text describing zones either needs to be deleted, or changed. I favour the former, but if the latter is favoured, I suggest that it may perhaps be useful to use the term areas rather than zones, to avoid confusion with any previous use which may be different.