ISO/IEC JTC1/SC2/WG2
Coded Character Set
Secretariat: Japan (JISC)

Doc. Type: Input to ISO/IEC 10646:2003

Title: Extended collections

Source: Project editor
Project: JTC1 02.18
Status: For review by WG2
Date: 2006-04-25
Distribution: WG2
Reference:
Medium:

To accommodate the need to insert UCS Sequence Identifiers content into collection, a new type of collection: 'extended collection' is introduced to accommodate that need. This requires a slight redefinition of the Collection definition (clause 4) and the introduction of the 'extended collection'.

4.12 Collection
A numbered and named set of entities. For a regular collection, these entities consist only of those coded characters whose code positions lie within one or more identified ranges (see also 4.20 for extended collection).

> NOTE – If any of the identified ranges include code positions to which no character is allocated, the repertoire of the collection will change if an additional character is assigned to any of those positions at a future amendment of this International Standard. However it is intended that the collection number and name will remain unchanged in future editions of this International Standard.

4.20 Extended Collection
A collection for which the entities can also consist of sequences of code positions. The sequences of code positions is referenced by UCS Sequence Identifier (USI) listed in clause 29 (see also 4.12).

The definition for fixed collection is unchanged. As a further consequence, the following note in clause 6.6 UCS Sequence Identifiers is deleted.

> NOTE – UCS Sequences Identifiers cannot be used for specification of subset and collection content. They may be used outside this standard to identify: composite sequences for mapping purposes, font repertoire, etc.

Annex A also needs to be modified as follows:
A.1 Collections
A.1.1 Collections of coded graphic characters
(Content formerly included in A.1)
A .1.2 Extended collections
(New content including the Lithuanian collection(s)).