ISO/IEC JTC 1/SC 2/WG 2
Multiple-Octet Codes
and

ISO/IEC JTC 1/SC 2/WG 3
7-bit and 8-bit codes and their extension

ISO/IEC JTC 1/SC 2/WG 2 *N 1705*

Date: **March 8, 1998**

Title: Use of Collection Identifiers for Identifying Sub-Repertoires of 10646

Source: USA (ANSI) and Canada

Status: Member body positions

Action: For the consideration of WG 2 and WG 3

References: ISO/IEC JTC 1/SC 2 N 2940 - Resolutions of the 7th Plenary Meeting
of ISO/IEC JTC 1/SC 2; Iraklion-Crete, Greece, 1997-07-08/09)

Distribution: ISO/IEC JTC 1/SC 2/WG 2 and WG 3 members

Background:

The US member body, Canada and some other member bodies have expressed their opinion at
the SC 2 - Crete meeting that for registering of sub-repertoires of ISO/IEC 10646 it should be
sufficient to use the collection identifier mechanism of 10646, and that the project proposal to
expand the scope of ISO 7350 be abandoned. This paper expands on this position.

At the Crete, SC 2 meeting in July 1997, the following resolutions were taken:

**RESOLUTION M 07.22 (Disposition of Comments on CD 7350, Registration of
repertoires of graphic characters from ISO/IEC 10367)**

SC 2 resolves to suspend this project on 7350 Revision (see Resolution M07.28).

UNANIMOUS

RESOLUTION M 07.28 (Collection Identifiers for 10646 subsets/sub-repertoire)

With reference to WG 2 Resolution M33.31 in document N 2927, and WG 3 Resolution
M12.11 in document N 2933, SC 2 instructs WG 2, in corporation with WG 3 to prepare a
proposal to cover the requirements for Collection Identifiers for 10646 subsets and report to
the next SC 2 Plenary. SC 2 further invites National Bodies and Liaison Organizations to
communicate their needs to WG 2. SC 2 invites US and Canadian National Bodies to assist
WG 2 in this effort.

UNANIMOUS

Collections are defined in the latest Technical Corrigendum (COR.2) to 10646-1: 1983 (E), which was recently balloted in SC 2 (Ref. ISO/IEC JTC 1/SC 2/WG 2 N 1664; Draft Technical Corrigendum No. 2 to ISO/IEC 10646-1:1 1983 (E))

> 4.11: collection: A set of coded characters which is numbered and named and which consists of those coded characters whose code positions lie within one or more identified ranges.

> 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.

> 4.17: dense collection: A collection in which every code position within the identified range(s) has a character allocated to it.

> Note - The repertoire of a dense collection can not be extended at a future amendment of this International Standard unless one or more of the identified ranges of code positions is extended.

Annex A of 10646 lists a number of sub-repertoires of 10646. The COR.2 has marked some of these as 'dense collections' based on the definitions cited above.

Note: There is another US / Canada contribution proposing a revisiting the above definitions and proposing removal of the part 'unless one or more ... ' in the Note under the definition 4.17 above. Also, note that the words 'extended at a fture amendment' should be changed to 'extended by a future amendment' in both the above clauses.

Sub-Repertoire of 10646 as Collections

A sub-repertoire of 10646 consists entirely of a set of coded characters defined in the standard. A 'dense collection identifier' can be assigned to any such sub-repertoire by giving it a number, a suitable name and the list of one or more contiguous ranges of the code positions assigned to these characters in the standard. The standard already has a listing of the code positions and names of characters assigned to that code position, and there is no need to replicate this information in the sub-repertoire definition. At present, a request to ISO/IEC JTC 1/SC 2 (which will be assigned to ISO/IEC SC 2/WG 2) to issue a collection identifier and enter that into the Annex A of the standard is all that is needed to identify such a sub-repertoire.

The existing procedure to add a collection identifier is through a 'technical amendment' to the standard, since Annex A is a 'normative' annex of the standard. Even though this procedure is more complex and cumbersome than a simple registration procedure, it can be effectively used to handle the requests for registration of sub-repertoires of 10646. A request for identifying a sub-repertoire can be prepared and submitted by a member body, a liaison organization or an individual expert to SC 2, along with some background information and rationale behind the request.

It is further proposed that the existing 'technical amendment' procedure be retained as it gives the maximum amount of filtering and control of requests for identifying sub-repertoires by ISO at the same time being able to respond responsibly to legitimate requests.