ISO/IEC JTC 1/SC 2

CODED CHARACTER SETS

SECRETARIAT: JAPAN (JISC)

DOC TYPE: Officer's Contribution

TITLE: SC 22 Liaison Report

SOURCE: Mr. Arnold Winkler, Liaison Officer

PROJECT: --

STATUS: This document was reviewed at the tenth plenary meeting SC 2 held in Athens, Greece, 2000-09-25/26.

ACTION ID: FYI

DUE DATE:

DISTRIBUTION: P, O and L Members of ISO/IEC JTC 1/SC 2
WG Conveners and Secretariats
Secretariat, ISO/IEC JTC 1
ISO/IEC ITTF

NO. OF PAGES: 4

ACCESS LEVEL: Open

Prepared by Arnold F. Winkler
September 22, 2000

I did not attend the SC22 plenary in Nara, September 2000. My report is based on the resolutions from the plenary meeting, and information from individual working group convenors.

1.   Administrative matters:

1.1.   Management:

Chair: John L. Hill (confirmed in Nara, September 2000)
SUN
Telephone: (717) 221-0584
E-mail: John.Hill@sun.com

Secretariate: Lisa Rajchel
ANSI
11 West 42nd Street
New York, NY 10036
Telephone: (212) 642-4932
Fax: (212) 840-2298
Email: lrajchel@ansi.org <mailto:lrajchel@ansi.org>

1.2.   Current working groups:

WG3	APL
WG4	COBOL
WG5	Fortran
WG9	Ada
WG11	Binding Techniques
WG13	Modula-2
WG14	C
WG15	POSIX
WG16	ISLisp
WG17	Prolog
WG19	Formal Specification Languages
WG20	Internationalization
WG21	C++
JSG	Java Study Group (disbanded September 2000)

2. Technical matters:

2.1. WG3 - APL

In its report from the meeting in Berlin, WG3 points out the discrepancy between the names of 5 APL characters in the APL standards IS 8485 (APL) and IS 13751 (APL Extended) and the Universal Character SET IS 10646. WG3 asks SC22 to request from SC2/WG2 a "correction" of the inconsistent names in IS 10646 (U+22A2 – U+22A5).
The resolution # 00-04 from the SC22 meeting in Nara documents this action.

> Resolution 00-04: APL Tack characters
> Noting the request from WG3 in SC22 N3176, ISO/IEC JTC 1/SC22 requests that ISO/IEC JTC 1/SC2 change the names of 5 APL characters in ISO/IEC 10646-1 as proposed in the said document (proposed action 1) and instructs the ISO/IEC JTC 1/SC22 secretariat to forward this to ISO/IEC JTC 1/SC2.

Proposed solution:
This situation has been discussed in SC2/WG2, based on SC2/WG2 N2260. The result is that Annex P to 10646-1:2000 will be changed to better explain the inconsistency between the APL standard and ISO/IEC 10646-1:2000.

2.2. WG4 – COBOL

The revision of the COBOL standard is under way, the FCD ballot is expected in February 2001.

2.3. WG4 - Request for guidance on "case folding"

In WG4's annual report 2000, the convenor requested an ad-hoc meeting on case folding for case-insensitive identifiers.

COBOL, Pascal, and other languages that have case-insensitive identifiers, need to be able to do reliable case-folding during their parsing/lexing phases of program text interpretation. For that, they need reliable definitions of case-folding as applied to 10646 characters for the domain of characters allowed inside identifiers for each language.

While WG20 has not touched on this issue and the SC22 working groups are starting to search for an answer, the Unicode Technical Committee and the IETF have moved ahead, creating de facto solutions that will see widespread implementation in the near future.

The Unicode Technical Committee has already published CaseFolding.txt, a machine-readable file with recommendations on exactly how to do case-folding for all Unicode 3.0 characters (i.e. 10646-1:2000 characters). The SC22 committees should be reviewing that file, and the associated case mapping information available in UnicodeData.txt and in SpecialCasing.txt—also available on the Unicode website—before concluding that new standardization efforts need to be initiated in SC22 (whether in WG20 or in other working groups), to repeat the work involved in creating those files, which are already freely available to all implementers.

http://www.unicode.org/unicode/reports/tr21/CaseFolding.txt

2.4.   WG20 – Internationalization

WG20's TR 10176, Amendment #1 has been published. It contains the list of characters suitable for identifiers in programming languages. This list (Annex A of TR 10176) will continuously need to be updated to reflect the extended repertoire of ISO/IEC 10646, both parts.

WG20's ISO/IEC 14651 – International string ordering has just been approved as an International Standard (15:1:5). This standard contains a table of all characters encoded in ISO/IEC 10646. This table will have to be amended with the expanding repertoire of both parts of ISO/IEC 10646. This work has been authorized by SC22.

The strong liaison and support from SC2 is appreciated for above tasks.

The convenor of SC2s/WG20 has, in a personal contribution to SC22, suggested that the maintenance of the character tables in TR 10176 and in IS 14651 be assigned to SC2.
SC22 has not accepted the proposals in my personal contribution. Nevertheless, I believe that SC2 is the group that should be prepared to take responsibility for the tables, so that the full repertoire of ISO 10646 characters can be utilized easily in above mentioned standards.

2.5.   WG21 – C++

Uses the character recommendations in TR 10176 for identifiers in its work.

2.6.   JSG – Java Study Group

The Java stuffy group has been disbanded in the SC22 plenary in Nara.

Arnold F. Winkler