

L2/02-374

ISO/IEC JTC 1/SC 2
Coded Character Sets
Secretariat: [Japan \(JISC\)](#)

DOC. TYPE	Program Extension Document	
TITLE	WG 2 Proposal for Merging ISO/IEC 10646 Parts 1 & 2 into One Part [WG 2 N 2506, N 2408]	
SOURCE	WG 2	
PROJECT	JTC 1.02.18	
STATUS	For consideration at the 12th SC 2 Plenary Meeting to be held in Tokyo, Japan, 2002-12-12/13.	
ACTION ID	ACT	
DUE DATE		
DISTRIBUTION	P, O and L Members of ISO/IEC JTC 1/SC 2 ; ISO/IEC JTC 1 Secretariat; ISO/IEC ITTF	
ACCESS LEVEL	Open	
ISSUE NO.	150	
FILE	NAME SIZE (KB) PAGES	02n3638.pdf 4

Secretariat ISO/IEC JTC 1/SC 2 - IPSJ/ITSCJ *(Information Processing Society of Japan/Information Technology Standards Commission of Japan) Room 308-3, Kikai-Shinko-Kaikan Bldg., 3-5-8, Shiba-Koen, Minato-ku, Tokyo 105-0011 Japan *Standard Organization Accredited by JISC

Telephone: +81-3-3431-2808; Facsimile: +81-3-3431-6493; E-mail: kimura@itscj.ipsj.or.jp

ISO
INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC 1/ SC 2/ WG 2

JTC1/SC2/WG2 N2506

Date: 2002-09-24

Title:	Merging ISO/IEC 10646 parts 1 & 2 into one part as Edition 3
Source:	Mike Ksar, SC2/WG2 Convener
Action:	SC2 Approval
Distribution:	ISO/IEC JTC 1/ SC 2

1. Scope of Work

SC2/WG2 accepted in principle the proposal in document SC2/WG2 N2408 to combine the two parts of 10646 into a single part and requests SC2 to approve new project sub-division with the rationale from document SC2/WG2 N2408.

2. Detailed Contents

- a. The combined document will be published as ISO/IEC 10646 – 3rd edition. All of the contents of parts 1 & 2 and their respective amendments would have been approved by the time the project is complete.
- b. Having a single part 10646 standard considerably reduces the complexity in its maintenance and ensures the continued synchronization of 10646 with the Unicode Standard.
- c. Document SC2/WG2 N2408 is attached for reference.

**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC 1/SC 2/WG 2**

**Universal Multiple-Octet Coded Character Set
(UCS)**

ISO/IEC JTC 1/SC 2/WG 2 N2408

2001-11-28

Title:	Consolidating the two parts of 10646 into a single part
Source:	USA national body and the Unicode Consortium (Drafted by - V.S. Umamaheswaran - umavs@ca.ibm.com)
References:	None
Action:	For consideration by SC 2/WG 2 and by JTC 1/SC 2
Distribution:	ISO/IEC JTC 1/SC 2/WG 2, ISO/IEC JTC 1/SC 2

Background:

The Unicode standard (TUS) and ISO/IEC 10646 standards have been kept in synchronism with each other extremely well due to the excellent cooperation between the Unicode Technical Committee and ISO/IEC JTC1/SC2/WG2. The same set of tools is used to produce the various charts and names lists in the two standards, ensuring that no discrepancy creeps in. However, one major difference exists – this is in the publication aspects of the two standards.

TUS is published as a single volume and has gone through three major publications, the latest single volume being Version 3. These are supplemented by on-line point-versions like Version 2.1, Version 3.1 etc. made available at www.unicode.org.

10646 is published as two parts. Part 1 contains all the architectural elements plus definition of content of the BMP. Part 2 is a Supplement to part 1 and contains definitions of the supplementary planes. Each part is on a different publication schedule. The second edition of 10646-1 was published in 2000 and the first edition of 10646-2 was (to be) published in November 2001.

The repertoire and encoding defined TUS 3.0 corresponds to 10646-1: 2000. TUS 3.1 corresponds to 10646-1: 2000 plus 10646-2: 2001 (and one more character).

The Problem:

When there was only one part of 10646, changes to it were processed as amendments to that part. It was easier to keep track of a set of these amendments plus the base edition and synchronize that set to a version or point-version of TUS.

With the publication of part 2 of 10646, there are now two amendment tracks -- one for part 1 and another for part 2. Some changes to part 2 will necessitate changes to part 1 even though there may not be any character additions to the BMP, which was the case for amendment 1 to part 1 (which exited the FDAM ballot in October 2001). TUS 3.2 will be in synchronism with 10646 (part 1: 2000, part2: 2001 and Amd.1 to part 1: 2000).

At the October 2001 meeting in Singapore, WG2 initiated work on two project sub-divisions -- Amendment 2 to Part 2, and Amendment 1 to Part 1, both involving character additions to each of

the parts. The intent of Version 4.0 of TUS is to be in synchronism with the result of both these amendments.

Having the same editor for the parts facilitates the progression of such amendments to the two parts on (almost) the same time scale. However, as additions to the repertoire of either part of 10646 are accepted by WG2, future amendments to the two parts can very easily go out of time-synchronization. It also poses a major difficulty to maintain the synchronization between TUS versions and point versions and 10646 parts and their amendments. The same set of tools is also used to generate the glyph charts and names lists for publication of the different amendments and the two standards. The maintenance of these tools is also getting to be complicated with the two amendment tracks, the different publication schedules, and the synchronization of the various sets with TUS.

The UTC and the US national body request the SC2 member bodies and the WG2 to consider and endorse the following towards solving the above problems before it becomes a maintenance nightmare.

- A. At the earliest opportunity – for example, the next edition date of Part 1, merge the two parts into a SINGLE part, incorporating all the amendments published (or approved for publication) to that date. Having a single part 10646 standard considerably reduces the complexity in its maintenance and in maintaining the synchronism between TUS and 10646.

WG2 has already taken some key decisions that will facilitate the single-part publication. The earlier decision (pre-edition 1) by WG2 to publish 10646 in multiple parts was based on considerations for the page count for printed copies of the standard, and due to the lack availability of suitable tools to generate the various charts. The multiple column layout of CJK charts in part 1 occupies majority of the pages in the standard. Publishing the standard in soft-copy media, such as on CD-ROM, has alleviated this problem to a large extent. Another significant WG2 decision that helps in the maintenance was to publish amended code charts as replacement pages instead of a list of characters and glyphs to be added, which was the practice prior to the second edition of 10646. Separating the source-information into machine-readable files made available on the CD-ROM, as was done for part 2 publication, another significant WG2 decision, alleviated the pressure on the page count in part 2, especially for the CJK Extension B charts.

- B. Adopt a single-column CJK layout for part 1 also.

WG2 and national bodies who have been using the multiple-column CJK layout in part 1, should be strongly encouraged to migrate towards the solution of a single-column CJK layout supplemented with a data file containing the equivalent CJK source information, like in part 2. We note that a resolution had been taken in this direction by WG2 at its October 2001 Singapore meeting. Adopting the single column format for CJK also in part 1 will contribute to simplification of the tools, facilitate the merging of the two parts into a single part, and the future maintenance of the standard.

- C. Before the single part publication of 10646 is realized, maintain time synchronization for each of the amendment for the two parts.

For example, Amendment 2 to Part 1 and Amendment 1 to Part 2 should have (almost) the same target publishing date. WG2 and its member body experts are encouraged to work towards achieving this objective. This would be the minimum needed to maintain reasonable synchronism between the future TUS versions and 10646 and its amendments.