

ISO
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 N 1033

Date: 1994-06-01

Title:	Unconfirmed Minutes of ISO/IEC JTC 1/SC 2/WG 2 Meeting 25 Falez Hotel, Konyaalti Falez Mevkil, 07050 Antalya, Turkey, 1994-04-18--22
Source:	V.S. UMaMaheswaran, Meeting Secretary; Mike Ksar, Convener
Action:	WG 2 Members and Member Bodies
Distribution:	ISO/IEC JTC 1/SC 2/WG 2 Members, ISO/IEC JTC 1/SC 2

1 Opening and roll call

1.1 Opening

The convener opened the meeting at 9:20h. He thanked the Turkish Standards Institution (TSE) for organizing the meeting and for the excellent facilities at the Falez Hotel, Konyaalti falez mevkil, 07050 Antalya, Turkey. The target is to close the meeting by Thursday PM.

Mr. Ahmet Kurter, TSE, welcomed the delegates on behalf of TSE. The work of WG 2 and WG 3 will be beneficial to the world. TSE has been an active member of this valuable international effort. Mr. Kurter welcomed the delegates to the city of Antalya. The delegates were also invited to a cocktail reception on Monday at 7 p.m., and a gala dinner on Friday evening. All delegates were requested to inform TSE staff about their participation in a bus or boat trip in the Antalya area planned for Sunday the 24th (subsequently changed to 23rd).

1.2 Roll call

The attendance list was circulated - all delegates were requested to mark any corrections to the mailing list of WG 2 distributed before the meeting. New attendees were requested to give full information - address, affiliation, telephone and fax numbers and e-mail address where available.

Thirty three (33) delegates representing 17 member bodies were present. The convener conveyed apologies from Mr. Mark Davis, the Editor, who could not attend the meeting for personal reasons. Mr. Bruce Paterson pointed out (towards the end of the meeting) that the number of member bodies attending this WG 2 meeting is a new record: 17 member bodies were present, compared to 12 at Washington DC, and 14 or 15 at Athens. The attendance list given below is ordered alphabetically according to the country represented.

Name	Country	Affiliation
Alain LaBonté	CANADA	Government of Quebec
V. S. UMaMaheswaran	CANADA, Meeting Secretary	IBM Canada
M. Enghebatu	CHINA	Mongolian Language Institute
Nyima Trashi	CHINA	Tibet University
Shamalayi	CHINA	National Institute of Chengdu
Wang Chuk	CHINA	Tibetan Bureau of Translation
Wushur Slamu	CHINA	Xianjiang University
YongJun Zhou	CHINA	China Great-Wall Computer Group
Bo Jensen	DENMARK	IBM
Keld Simonsen	DENMARK	Consultant
Alaa Ghoneim	EGYPT	IBM Egypt
Michel Suignard	FRANCE	Microsoft
Joachim Friemelt	GERMANY	Siemens
N. Subramanian	INDIA	CDAC
Stefan Fuchs	ISRAEL	Bezeq
Akio Kido	JAPAN	IBM Japan
Kohji Shibano	JAPAN	Tokyo International University
Shigenobu Kato	JAPAN, IRG Rapporteur	Toppan Printing
Takayuki K. Sato	JAPAN	HP
Kyo Won Yoon	KOREA	KBS
Kyongsok Kim	KOREA	Pusan National University
Won Sun Lim	KOREA	Ministry of Culture and Sports
Johan van Wingen	NETHERLANDS	Independent Consultant
Kolbjørn Aambø	NORWAY	University of Oslo

Sten Lindberg	SWEDEN	IBM Svenska
Jan van den Beld	SWITZERLAND SC 2 Secretary, WG 3 convener	ECMA
Jürgen Bettels	SWITZERLAND	Digital Equipment Europe
Bülent Ayasli	TURKEY	Consultant
Ümit Karakas	TURKEY	Hecettepe Universitesi
Yilmaz Akyol	TURKEY	Consultant Independent
Bruce Paterson	UK	Consultant
John Clews	UK t	Consultant
Arnold Winkler	USA	UNISYS
Asmus Freytag	USA, UNICODE	Microsoft
Mike Ksar	USA, Convener	HP

2 Approval of the agenda

Input Documents:

N 958 Venue and Call for meeting #25 in Antalya, Turkey; Mike Ksar; 1994-02-01

N 962 Second call for WG 2 meet #25 in Antalya, Turkey; Mike Ksar; 1994-02-01

N 962R Agenda items and corresponding document numbers, WG 2 meeting #25, V.S. Umamaheswaran, 1994-04-18

Objectives of the meeting as stated in N 962 were as follows:

"The WG 2 meeting will focus on finalizing the discussion on the basic architectural issues that WG 2 developed in the last meeting in Washington DC in November 1993 (based on documents N 946 and N 947). The meeting will also review the relevant action items on the basic architecture before proceeding with a discussion on repertoire issues."

Documents identified on the agenda had been distributed to WG 2 membership, and some copies were made available at the meeting (courtesy of TSE). New documents since the last WG 2 mailing (beyond N 975) (excepting N 977, N 978, N 980, N 981, N 982 and N 983 which will be mailed out later) were made available during the meeting. Document N 962R itemizing the various documents to be considered under different agenda items, (prepared by Dr. V.S. Umamaheswaran, the meeting secretary) was used for the discussion. The following further modifications were made to the agenda:

- Under Item 7: Add Presentations on ISCII (by delegate from India), on Tibetan, Mongolian, Yi and other scripts (by delegates from China).
- Under Item 8: Add Defect reports - several papers including China's input related to Arabic script, France's input related to Armenian, NNI input related to Greek.
- Item 5 was postponed till Mr. Kato, IRG rapporteur arrived.

The agenda was approved as amended -- in N 962R. Other documents generated during the meeting were added to the appropriate agenda items as the meeting progressed. The final revised agenda is given below. The convener emphasized that new contributions and all repertoires-related contributions will be considered by the meeting only if time permitted.

Item Number	Title	Page
1	Opening and roll call	1
1.1	Opening	1
1.2	Roll call	1
2	Approval of the agenda	2
3	Approval of the minutes of meeting 24 (Washington)	3
4	Review action items from previous meetings	3
5	Report from WG 2/IRG Documents: N 979, N 980	5
6	Finalize basic architecture	6
6.1	UTF-16 Related	6
6.1.1	N 976 - Mr. B. Paterson	6
6.1.2	N 972 - Professor K. Kim	7
6.1.3	N 988 - Mr. B. Paterson	7
6.1.4	N 1008, N 1009 - Mr. YongJun Zhou	8
6.1.5	N 1010: YongJun Zhou	8
6.1.6	N 998 - Mr. T.K. Sato	8
6.1.7	N 969 - Mr. Bruce Paterson	8
6.1.8	N 991 - Mr. Ed Hart's personal comments	9
6.1.9	Renaming of UTF-16 back to UCS-2E	9
6.1.10	Relevant Resolutions	9
6.2	UTF-8 Related	10
6.2.1	N 993 - pDAM2	10
6.2.2	Proposal to Remove UTF-1 - Mr. J. Bettels	10
6.2.3	Relevant Resolutions	11
7	Repertoire issues	12

Item Number	Title	Page
7.1	Procedures and principles of allocation	12
7.1.1	N 1002 - Mr. T.K. Sato, Japan	12
7.1.2	N 973 - Professor K. Kim	12
7.1.3	Relevant Resolutions	13
7.2	Additions to repertoire	13
8	Other business	13
8.1	Defect reports	13
8.1.1	Discussion on Procedures for Dealing with Defect Reports	14
8.1.2	N 974 - Clarification of the term Conjoining - Korea	14
8.1.3	N 975 - Incorrect Glyphs - Korea	14
8.1.4	N 989 - UK Several Defect Reports	14
8.1.5	N 992 - Canada	15
8.1.6	N 999 - Japan	15
8.1.7	N 1004 - Japan	16
8.1.8	N 1006 - Japan	16
8.1.9	N 1014 - Japan	16
8.1.10	N 1018 - Mr. J.W. van Wingen	16
8.1.11	N 1022 - Mr. Michel Suignard	17
8.1.12	N 1007 - Turkey	17
8.1.13	N 1023 - Israel	17
8.1.14	N 1025 - Mr. Stefan Fuchs	18
8.1.15	N 994 - Corrigendum 1 on Æ	18
8.1.16	Document N 995 - pDAM 3 on C1 Defect	18
8.1.17	N 1030 - Mr. John Clews and Mr. N. Subramanian	18
8.1.18	N 1031 - Ad hoc group on editorial defect reports	19
8.1.19	Relevant Resolutions	19
8.2	Liaison reports	19
8.2.1	N 1005 - Mr. T.K. Sato	20
8.2.2	N 1029 - Dr. Asmus Freytag, The Unicode Consortium	20
8.2.3	N 1024 - Mr. Alain LaBonté - SC 18/WG 9	20
8.2.4	N 960 - ITU TS SG 8 - Mr. Joachim Friemelt	20
8.2.5	N 1020 - Mr. Stefan Fuchs	20
8.2.6	N 1026 - The Unicode consortium - Dr. Asmus Freytag	20
8.2.7	N 1032 - Mr. Keld Simonsen	21
8.2.8	Liaison from WG 2 to SC 29 - Mr. Jan van den Beld	21
8.3	WG 2 program of work	21
8.4	8.4. Presentations	Error! Bookmark not defined.
8.4.1	ISCI - N. Subramanian	21
8.4.2	N 1011 - Mongolian, TODO and XIBE (incl. Manchu) - by Professor Enghebatu	23
8.4.3	N 964 - Tibetan - Professor Nyima Trashi	23
8.4.4	N 1012 - Uighur, Kazakh and Kirghiz supplements for Arabic script - Professor Wushur Slamun	23
8.4.5	N 965: Yi Script - by Mr. YongJun Zhou and Mr. Shamalayi	24
8.5	Future meetings	24
8.6	Other Concerns	24
9	Closing	24
9.1	Approval of Resolutions	24
9.1.1	Administrative Resolutions:	24
9.2	Adjournment	25
10	Action Items	25
10.1	From WG 2 Meeting 24, Washington DC	25
10.2	From WG 2 Meeting 25 Antalya, Turkey	25
10.3	List of output documents from the meeting	28

3 Approval of the minutes of meeting 24 (Washington)

Input Documents:

N 955 Unconfirmed minutes of WG 2 meeting 24; Mike Ksar and V.S. Umamaheswaran; 1993-11-30

Dr. V.S. Umamaheswaran gave a brief presentation of the minutes of meeting 24. The minutes were approved with minor corrections. Some noted corrections are:

- Item 2.1c should be a sub item of 2.1b, and 2.1d should be renumbered 2.1c.
- Item 9.2.2 h: 2nd line, '... proposal in N 908.' should be '...proposal in N 918.'
- List of action items, item 1 e is duplicated elsewhere and should be deleted
- Inconsistency in method of referencing delegate names in the minutes was pointed out, along with some spelling errors.

4 Review action items from previous meetings

Input Documents:

N 955 Unconfirmed minutes of WG 2 meeting 24; Mike Ksar and V.S. Umamaheswaran; 1993-11-30

N 962R Agenda items and corresponding document numbers, WG 2 meeting #25, V.S. Umamaheswaran, 1994-04-18

The status of the different action items listed in document N 955 (and attached to N962R for convenient reference at the meeting) is shown below.

Item	Assigned to / action (from N 955)	Reference Output/Status
1.	All national bodies	
a.	to review the documents N 946 and N 947 and submit written contributions to 'refine the starting list of characters under different categories' by 1994-01-31; (reference resolution SP-7 in section 6.1).	N 983 UK N 984 Canada N 1002 Japan
b.	to contribute input on source of characters and cross-reference to other standards; (reference resolution TKS-2).	N 963 China N 964 China N 965 China N 966 China N 967 China N 968 Mongolia N 973 Korea N 984 Canada N 1011 China
c.	to contribute input towards developing an informative Annex on Equivalence between pre-composed characters and their corresponding composite sequences; (reference resolution TKS-3).	N 1005 SC 22/WG 20 N 1029 UNICODE N 1032 Denmark
d.	to review all existing WG 2 documents relevant to additional characters, including comments / disposition of comments -- prior to publication of 10646-1; do all the homework needed so that we do not miss any characters from previous submissions - before next WG 2 meeting; (reference item 8.1 of minutes).	N 990 USA
2.	Editor:	
a.	to prepare the text for a Normative Annex on UTF-16 by 1993-12-01 based on the items in resolutions UTF16-4 to UTF16-8; (reference resolution UTF16 -2).	N 970R Editor
b.	to prepare an accompanying background paper showing clearly the benefits and disadvantages of the proposed UTF-16, addressing the comments and concerns brought up during the WG 2 meeting no. 24 by 1994-01-05; (reference resolution UTF16-3)	TO DO
c.	to draft a proposed text for a new normative Annex on UTF-8; (reference resolution UTF8-2).	N 993 Editor
d.	to include some text on 'what criteria can be used for defect report on character names' as part of the Set of Principles document (reference section 9-a-i.3).	TO DO
e.	to include the corrections by changing 'LIGATURE' to 'LETTER' to the following non-presentation form of characters in the next CORRIGENDUM: - 00C6 - LATIN CAPITAL LIGATURE AE - 00E6 - LATIN SMALL LIGATURE AE - 01E2 - LATIN CAPITAL LIGATURE AE WITH MACRON - 01E3 - LATIN SMALL LIGATURE AE WITH MACRON - 01FC - LATIN CAPITAL LIGATURE AE WITH ACUTE - 01FD - LATIN SMALL LIGATURE AE WITH ACUTE.	N 994 Editor
f.	to prepare a defect report to correct the names for all the non-presentation form characters that have the term LIGATURE in their names (reference resolution AE-1), excluding those under resolution AE-2; (reference resolution AE-3).	TO DO
f.	to prepare a draft text for an addendum regarding C1 controls with N 917 as input document, for consideration at the next WG 2 meeting; (reference resolution C0C1-2).	N 995 Editor
h.	with assistance from the Japanese member body, to prepare appropriate text for a CORRIGENDUM to address the defects reported in document N 939 under items 1, 3, 4 and 5; (reference resolution DFR-1).	TO DO
3.	Convener (Mike Ksar):	
a.	to forward document N 915 to SC 2/WG 3; (reference section 7.1.1).	DONE
b.	to forward the liaison statement in document N 944 to SC 2 for distribution to SC 2 and SC 2 / WG 3 members for their review and comments; (reference resolution CGM-5).	DONE
c.	to forward NP on addenda on UTF-8 contained in N 952R, to SC 2 for SC 2 letter ballot; (reference resolution NP-1).	DONE
d.	to request SC 2 to establish a formal relationship of liaison, Category C, between SC 2/WG 2 and Taipei Computer Association and forward the statements of benefits and responsibilities (reference resolution L-1).	DONE
4.	IRG:	
a.	to proceed with drafting a suitable text for a TR on 'User's guide to CJK Unified Ideographs' documented in N 931; (reference resolution TKS-4).	N 979 IRG N 980 IRG
b.	to proceed with preparing a suitable draft text for an informative Annex to ISO/IEC 10646-1 for rules on unification, indexing and establishing relationships amongst the variants of unified CJK ideographs described in N 932 (with the word 'ordering' changed to 'indexing'); (reference resolution TKS-5).	N 979 IRG N 980 IRG
5.	Takayuki K. Sato:	
	to forward liaison report N 948 R to SC 22 / WG 20; (reference resolution TKS-1).	DONE

6. Keld Simonsen and Convener:		
a.	Produce revised document N952-R containing a request for sub-division of project. N 952-R will make old N 952 and N 954 obsolete	N 952R Simonsen, Convener
b.	Update document N 957-R containing the WG 2 program of work - capturing all the items that were accepted at this meeting, with dates showing the next milestone.	N 957R Simonsen, Convener
7. US member body & Korea member body:		
	"US needs to have Korea and US, SC 2, and SC 18 experts develop a proposal for a single writing system for Hanguk"; (reference section 5, item pending from Seoul meeting).	Character Glyph Model deals with this item. Action Item is dropped.
8. Japan member body:		
	to get further clarification on the contrast between existing and proposed shapes for regarding "ditto mark" (3003x) and "wave dash" (301Cx) reported as defects in document N 939 item 6; (reference resolution DFR-2).	N 1014 Japan
9. The meeting secretary (V.S. Umamaheswaran)		
	To incorporate all the comments on draft resolutions in N 949 and produce the final text of resolutions.	N949R Secretary/Convener N955 Secretary/Convener

5 Report from WG 2/IRG Documents: N 979, N 980

Input Documents:

N 979 Resolutions of the 2nd IRG Meeting (IRG N 080); KATO Shigenobu, IRG Rapporteur; 1994-03-16

N 980 Minutes of the 1st IRG Meeting (IRG N 035); KATO Shigenobu, IRG Rapporteur; 1994-03-16

(Note: Document N 980 was not available at the meeting - will be mailed to members.)

Mr. S. Kato presented the 2nd IRG meeting report held in Hanoi 1994-02-28--03.04. See document N 979. The document contains 21 paragraphs.

- a. Two annexes are planned:
 - Rules of unification (editor Mr. T. Koike, Japan), and
 - Identification of source standards for each ideograph (IRG, China, USA - Mr. L. Collins).
- b. New terms have been defined for IRG use to assist in preparation of additions to the repertoire. Several contributions being discussed by IRG were proposed to be organized into different categories based on this new terminology.
- c. Editor Group Appointments:
 - Chief project editor -Mr. YongJun Zhou, China.
 - Editorial WG - Mr. Koike, Japan; Prof. K.S. Hyeon, Korea; Dr. Thanh Nhan Ngo, Viet Nam, Mr. T.C. Kao, TCA, Prof. B. Tsou, HKITF and Mr. L. Collins (Tentative), USA.
- d. Mr. Akio Kido, Japan, is nominated new IRG rapporteur, replacing Mr. Kato. Mr. Kato will remain an IRG member.
- e. A fifth column (V-column) is proposed to be added to the ideographic tables in 10646-1 to accommodate Vietnamese ideographs in the unified Han set. About 200 characters are anticipated to be added.
- f. Future meetings:
 - 3rd IRG - Beijing, 1994-08-29--09-02
 - 4th IRG - Taipei; 1995-02-07--03-03
 - 5th IRG - Seoul; 1995-05 -- 06-xx

Discussion:

- a. Mr. John Clews - Is IRG responsible for code position assignments? NO. IRG Role is clearly documented in N 888 and N 904 (Athens WG 2 minutes??)
- b. The convener requested that the proposed annexes be sent to him in time for distribution to WG 2, if they are to be considered at the WG 2 October 94 meeting.
- c. Clarification was sought on item 5 in N 979 - IRG's decision on adding a V column. Question was raised if IRG considered the practical problem the V column may cause in the publication of the 10646. IRG may use as many columns as needed in their report to WG 2. It should not assume that the number of columns will also appear in tact in 10646. May be other alternatives than a V column will be possible.
- d. Mr. T.K. Sato: The V column is a logical extension of the current CJK columns. There may be a need for additional columns as the requirements unfold.

Disposition:

- a. Annexes on Unification Rules and Supplemental Repertoires will be provided by IRG to WG 2 prior to the next WG 2 meeting.
- b. National bodies who are interested in attending IRG meeting to inform the new rapporteur. See resolutions ADMIN-1 and ADMIN-2 under item

6 Finalize basic architecture

Input Documents:

- N 969 Interaction of proposed UTF-16 and UTF-8; Bruce Paterson; UK; 1994-02-07
- N 970R UTF-16 Proposed draft amendment to 10646; Mark Davis; 1994-02-08
- N 971 Interaction between UTF-16 and UTF-8; Misha Wolf, UK; 1994-02-09
- N 972 Comments on UTF-16; Kyongsok Kim, Tae-Jin Kang, Dachyuk Ahn; KBS, Korea; 1994-01-20
- N 976 Editorial Comment on PDAM1 for 10646 (WG 2 N 970); Bruce Paterson, UK; 1994.03.04
- N 988 A UTF-16 with no adverse side-effects; Peter Fenwick; U.K.; March 1994
- N 991 Personal comments on n 970, "ISO/IEC 10646-1 Proposed Draft Amendment 1"; Ed Hart; 1994.04.14
- N 993 10646 Proposed Draft Amendment 2; Mark Davis, WG 2 Project Editor; 1994.03.03
- N 994 10646-1 Proposed Draft Corrigendum 1; Mark Davis, WG 2 Project Editor; 1994.03.03
- N 995 10646-1 Proposed Draft Amendment 3; Mark Davis, WG 2 Project Editor; 1994.03.03
- N 996 Background C code sample implementation for 10646-1 Proposed Draft Amendments 1 and 2; Mark Davis, WG 2 Project Editor; 1994.01.15
- N 997 Comments on UTF-16; Takayuki K. Sato, Japan; 1994.03.28
- N 998 Additions on BMP; Japan (TKS); 1994.03.28
- N 1008 Comments on architecture of 10646; China; 1994.04.11
- N 1009 Redefining UTF16 for UCS-4 with compression; YongJun Zhou, China; 1994.04.12
- N 1010 UTF16 with UTF16 Swapping seems better; YongJun Zhou, China; 1994.04.12

Output Documents:

- N 1021 Applicability of UTF-16 (clarification); Bruce Paterson, Kyongsok Kim, Sato; 1994.04.18
- N 1027 UTF-16 Specification (Re: N 970R and N 976); Bruce Paterson, UK; 1994.04.20

6.1 UTF-16 Related

UTF-16 Related Documents: N 970R (base), N 972 (Korea), N 976 (UK), N 988 (P. Fenwick), N 991 (Ed Hart), N 997 (Japan), N 998 (Japan), N 1008 (China), N 1009 (China), N1010 (China). In the absence of the editor Mr. Mark Davis, the convener volunteered to take the questions and get back to the editor.

6.1.1 N 976 - Mr. B. Paterson

Mr. Bruce Paterson presented N 976. At the time of writing this document was a Personal Contribution. It has subsequently been adopted by UK without any change. N 976 contains a summary of the substantial editorial changes proposed on N 970R the pDAM 1 draft.

Item a. *The concept of 'surrogate' is not appropriate.*

The term RC-element (a two-octet R-octet and C-octet sequence) is proposed instead of 'surrogate' in N 970R. Terms high-half and low-half are proposed instead of High Half (without the hyphen) and Low Half in N 970R. Alternative text for the pDAM 1 is included in N 976.

Discussion:

- a. Dr. Asmus Freytag: Suggest perhaps RC-element definition should come before its use in proposed 4.34 and 4.35. The more correct English use will perhaps be acceptable.
- b. Mr. Bruce Paterson: Even though RC-element is logically before, standards editing calls for Alphabetical ordering of terms defined. The term '(if any)' in RC-element definition is used in the context that BMP is currently used. Removing 'if any' may not do any harm.
- c. Several delegates: There was a discussion on the potential confusion of the definition of RC-element. The intent was to reference 'a new element which identifies the least significant 16-bits of a UCS-4 code point or the entire UCS-2 code point'.

Disposition: Mr. Paterson and Dr. Freytag (after discussions outside the meeting) proposed the following:

- Change the definition 4.36 under item 1 of N 976 to:
4.36 RC-element: a two-octet sequence comprising the R-octet and the C-octet (see 6.2) from the four-octet sequence that corresponds to a cell in the coding space of this coded character set.
- Add the following in the main body of 10646-1:
Clause 14.1 "Two-octet BMP form".
Add at the end of paragraph 2:
"(i.e. its RC-element)".
See resolution pDAM-UTF16 in section 6.1.10 below.

Item b. *Terms like 'code', 'character' etc. are not used strictly in accordance with the terms defined in the standard. The term 'system' is used without a definition in the standard.*

A proposed alternative text Annex X is attached to N 976. Delegates were requested to review the proposals.

Disposition: See resolution pDAM-UTF16 in section 6.1.10 below.

Item c. *M.6 - Interpretation of malformed sequences:*

This section is an informative text dealing with error conditions in receiving devices, and should not be in the normative text. The proposal is to remove it.

Discussion:

- a. Dr. Asmus Freytag: Due to the nature of the change and its potential impact on an existing UCS-2 implementation some level of conforming behaviour is recommended.
- b. Mr. Mike Ksar: See 10646-1 p2 "2.3c Receiving devices, Notes" and p 746 Appendix H. Error conditions due to an unrecognized character are in the standard now. An incomplete character is not treated in today's standard. Because UTF-16 introduces this new error condition we have to give it some attention, though the specific wordings in the proposal in N 970R from the editor may not be quite what we need.
- c. Dr. Umamaheswaran proposed the development of an informative annex on error handling. The convener supports this. The problem of missing RC-elements will occur for every UTF variant, so a text, if any, should be generic.
- d. Mr. Joachim Friemelt suggested that we be careful in dealing with error handling.

Disposition: Mr. Bruce Paterson and Dr. Asmus Freytag formulated the following wording as amended text for clause X.6 in N 976 (and is documented in the output document N1027 from this meeting):

X.6 Incorrect Sequences of S-Zone RC-Elements:

interpretation by receiving devices

If a CC-data-element includes either:

- an RC-element from the high half zone that is not immediately followed by an RC-element from the low half zone, or
- an RC-element from the low half zone that is not immediately preceded by an RC-element from the high half zone,

then according to X.1 (6) such an RC-element is not in conformance with the requirements of UTF-16. It is known as an unpaired RC-element, and it cannot be transmitted by an originating device that is in conformance with the requirements of UTF-16.

If a receiving device receives an unpaired RC-element, because of error conditions either:

- in the originating device, or
- in the interchange between the originating and the receiving device, or
- in the receiving device itself,

then it shall interpret that unpaired RC-element in the same way that it interprets a character that is outside the adopted subset that has been identified for the device (see 2.3c).

See resolution pDAM-UTF16 in section 6.1.10 below.

Item d. Refers to M-7: use of UTF-16 Process Code.

Suggestion is to either delete or to make it an advisory note.

The convener: Probably will be acceptable to the editor.

Disposition: See resolution pDAM-UTF16 in section 6.1.10 below.

Item e. Various omissions and inaccuracies -- items 2A, 2bis, 3 and 4 are proposed in N 976.

Disposition: See resolution pDAM-UTF16 in section 6.1.10 below.

6.1.2 N 972 - Professor K. Kim

Professor K. Kim explained that the UCS-4 and UTF-16 common number of code points will be much smaller than the total number in UCS-4. Proposal in N 972 is to define different zones of UCS-4 -- BMP, EMP and HMP.

Discussion:

- a. Mr. B. Paterson: Korean proposal will be a useful addition that can be added to section 9, as clarification of the additional planes into groups that can be handled by UTF-16 and other planes that cannot be handled.
- b. Mr. T.K. Sato -- referenced N 997. Because of the potential misinterpretation that UTFs will cover all of UCS-4, the terminology UTF-16 may be misleading. It proposes that we go back to UCS-2E. We may have to get the clarification that NOT ALL UTF-s can cover the UCS-4 one for one correspondence, if we are going to keep the name UTF-16.
- c. Dr. Asmus Freytag: we should consider going back to the term UCS-2E. Needs discussion.

Disposition: Professor K. Kim, Mr. B. Paterson and Mr. T.K. Sato met outside the meeting and proposed the following text as editorial instructions to address Korea's concerns (documented in output document N 1021):

9.1 Planes reserved for future standardization

.. In clause 9, add a new section as follows:

9.2 Planes accessible by the UTF-16 method:

Code positions in planes 01 to 10 in group 00 may be transformed to the UTF-16 representation (see Annex X). This representation is compatible with the two-octet BMP form of UCS-2 (see 14.1).

Code positions in planes 11 to FF in group 00, or in other groups, cannot be transformed to the UTF-16 representation.

See resolution pDAM-UTF16 in section 6.1.10 below.

6.1.3 N 988 - Mr. B. Paterson

The paper authored by Mr. P.W. Fenwick, UK, was focusing on character processing related items. It is a private contribution - not yet considered by UK.

Discussion:

- a. In Washington similar papers were considered and UTF-16 was arrived at on the basis of UCS-2 interchange considerations.
- b. The editor, Mr. Mark Davis had reviewed this contribution - the proposal has 'overloading' - i.e. a given code position can be part of single or double pairs use, and hence cannot be entertained.

6.1.4 N 1008, N 1009 - Mr. YongJun Zhou

Mr. YongJun Zhou briefly presented N 1008 and N 1009. China wants UCS-4 instead of UTF-16. for the following reasons:

- availability of code space
- transition difficulty
- development complexity in the UTF-16 case.

A 6-octet UCS-16 code is described as a compression of UCS-4 code in N 1009. This would use a zone 'D000'x to 'DFFF'x divided into a High, Medium and Low zone to transform values corresponding to all 2 GBytes of UCS-4 space. For processing code UCS-4 would be used. N 1008 claims that it is better than current UTF-16 proposal.

Discussion:

Dr. Asmus Freytag: Disadvantage of 6 octets; Upper limit of 32 bits for integers in many computers. It cannot be easily done with 6 octets. This seems to be an extension of the concepts of the current proposal for UTF-16. 'Unequal treatment of BMP versus non-BMP characters' mentioned in N 1008 is not addressed in N 1009. Justification for going beyond about a million characters is not there. As to the claimed 'unfairness' -- current standard makes UCS-2 not being able to carry any of UCS-4 at all (beyond BMP). A million characters may be adequate for now - but there is requirement for different treatment of low and high bytes in UTF-16. UCS-2 becomes disallowed in the proposed compression scheme. Also, COMPRESSION is a topic outside the scope of the standard.

Disposition: The convener - since compression is not within the scope of WG 2, suggested that China take the proposal to other committees who are involved in defining compression and similar mechanisms.

6.1.5 N 1010: YongJun Zhou

Four questions have been raised in this personal contribution from Mr. YongJun Zhou, and to address these questions, UTF16 with swapping is proposed as a solution.

Discussion:

- a. Mr. B. Paterson - Swapping was rejected at the Washington WG 2 meeting.
- b. Dr. Asmus Freytag: Swapping is limited in its scope in the document N 1010.
"Q 1: How about the users using the characters encoded in supplementary planes?"
There is more room still BMP today - chances are one will be using UCS-2. Additional planes are accessible via UCS-4. The concern seems to be the same rationale as UTF-16.
- c. Mr. Jürgen Bettels: the core of the proposal is on swapping which has been considered to be not acceptable.
- d. Professor K. Kim: The title says 'seems better' - better than what? Answer: Page 6 says it clearly -- better than basic UTF-16.

Disposition: Proposal in N 1010 uses swapping which goes against the Washington decision of not to use swapping. Not accepted.

6.1.6 N 998 - Mr. T.K. Sato

N 998 - Prohibition of Characters in O-Zone - also refer to N 928 and N 938. Japan wants a clarification that O-zone code positions will not be used for adding new characters in BMP. If UTF-16 is agreed upon, Japan does not object to use of O-zone for UTF-16 use.

Discussion:

- a. Mr. Joachim Friemelt: This applies to the whole of the BMP.
- b. Dr. Asmus Freytag: The committee has agreed to define the architecture before any allocations are decided.

Disposition: Record that no new BMP characters will be added, until the architecture extension is completed, excepting the use for UTF-16 High and Low 16-bits. See resolution PRIN-1 in section 6.1.10

6.1.7 N 969 - Mr. Bruce Paterson

Interaction of Proposed UTF-16 and UTF-8: Mr. Paterson raised the question - can UTF-8 be used directly against UTF-16, for example, as if it is the same as UCS-2? UTF-FSS had defined the algorithm in terms of Character - being a UCS Character as the input to the UTF-FSS transform.

Discussion:

- a. Dr. Asmus Freytag: Refer to proposed UTF-8 definition from Mr. Mark Davis in document N 993. UK's concern seems to have been addressed at least partly.
- b. Dr. V.S. Umamaheswaran: There is still a need for a statement addressing whether successive transforms are allowed (e.g. UTF-1 + UTF-8).
- c. Mr. B. Paterson: There seems to be still some concerns - editorial clarification may resolve the concerns.

Disposition: Dr. Asmus Freytag and Mr. Bruce Paterson proposed clarification text to address the above point. It was further enhanced at Dr. V.S. Umamaheswaran's request.

In N.1 add a new bullet:

- The transformation defined for UTF-8 is applicable to characters coded according to UCS-4 or UCS-2, but not for characters coded according to UTF-16. If it is desired to transform data coded in UTF-16, this can be accomplished by transforming from UTF-16 to UCS-4 before applying UTF-8. See resolution pDAM-UTF16 in section 6.1.10 below.

6.1.8 N 991 - Mr. Ed Hart's personal comments

Mr. Ed. Hart has posed two questions - one related to how actual distribution of amendment will be done along with some stated preferences, and the second one related to the ordering of high surrogate before the low surrogate. Several editorial comments on pDAM 1 on UTF-16 are also included.

Discussion on pDAM processing:

- a. Mr. Jürgen Bettels: prefers possibility of whole sections being replaced - than words or sentences being shown as deleted - it becomes unreadable otherwise.
- b. Mr. Sten Lindberg (Chair, JTC 1/SC 2): JTC 1 did address this issue, there are some guidelines formulated. The numbering of pDAMs was also discussed by JTC 1 so that the consecutive numbers do reflect the changes in a cohesive way.
- c. The normative text of the standard is only 14 pages - so that we should be able to change all of it at once.

Disposition: The convener: We should look at JTC 1 procedures addressing Mr. Hart's concerns. Editor is to be informed about the different options.

Discussion on order of high and low surrogates:

Section 6.3 of the standard already addresses this question. The pDAM wordings should not clash with section 6.3. The words 'surrogates' will probably be replaced by 'RC element' based on acceptance of input from UK.

Disposition: No action is needed.

Disposition of Editorial comments: Forward document N 991 to editor for consideration during further editing of pDAM1. See resolution pDAM-UTF16 in section 6.1.10 below.

6.1.9 Renaming of UTF-16 back to UCS-2E

Towards avoiding having to add an explanatory clause about UTF-s in general in the main body of 10646-1, Mr. Jürgen Bettels proposed that UTF-1 be removed from the standard, and that UTF-16 be renamed as UCS-2E, leaving UTF-8 as the only UTF (see also discussion under item 6.2.2.).

Discussion:

- a. Mr. Jürgen Bettels: the rationale for proposing the renaming of UTF-16 as UCS-2E is that the transformation in UTF-16 is Restricted - unlike UTF-8. Some editorial changes still have to be made.
- b. Dr. Asmus Freytag: Supported the proposal. It does not matter what it is called -- can be UCS2B4. The distinction is to make difference between transformation between codes within the standard versus between 10646 and others. There is a transformation. It is not universal. Would prefer to see it more integrated;
- c. Mr. Michel Suignard: Supported the idea.
- d. Prof. K. Kim: Objected to the name change to UCS-2E... UCS-2 and UCS-4 are basic architecture. Procedural question - in Washington DC we have a resolution in place.
- e. Dr. V.S. Umamaheswaran: Do not think it is a good idea. Referred to the minutes of the Washington meeting - one of the reasons WG 2 even accepted to proceed with the pDAM was on the basis of the name change from UCS-2E to UTF-16. There is still a transformation involved. if the word Universal in UTF causes the problem use TF.
- f. Professor K. Shibano: We balloted the sub-division of work and the progress of pDAM on the basis of the resolution in Washington.
- g. Mr. T.K. Sato - may be significant editorial work.
- h. Mr. Joachim Friemelt - Even if it is a transformation between codes within the standard there is a transformation. It should be avoided. If neither UTF nor UCS is acceptable may be we have to invent some other term.
- i. Mr. Mike Ksar - we are still reviewing the proposal that has been based on the Washington resolution. Proposes changing the processing of the "it" (current UTF-16) as a transformation between UCS-2 and UCS-4. Would prefer to see the text related to the current UTF-16 integrated into the main body of the standard than being a normative annex.
- j. Mr. Bruce Paterson: UK needs more time to think about it - leave it to national ballot.

Disposition: We will not make a decision at this meeting - can await national ballot comments.

6.1.10 Relevant Resolutions

pDAM-UTF-16

WG 2 instructs its editor to revise and update pDAM 1 in N 970R as follows:

UNANIMOUS

- a. Editorial suggestions and Annex M of document N 976 from UK with the following further modifications to the definition of RC element:
 - i. 4.36 RC-element: a two-octet sequence comprising the R-octet and the C-octet (see 6.2) from the four-octet sequence that corresponds to a cell in the coding space of this coded character set. Clause 14.1 "Two-octet BMP form":
Add at the end of paragraph 2:
"(i.e. its RC-element)".
 - ii. Paragraphs related to error handling as documented in N 1027 - 'UTF-16 Specification - X.6 Incorrect Sequences of S-Zone RC Elements'
- b. Comments expressed in document N 991
- c. Editorial clarification in Section 9 of the standard as drafted in document N 1021, addressing the concerns expressed in N 972 from Korea.

Further the editor is requested to forward the updated pDAM 1 to SC 2 secretariat for further processing, (by 1994-06-01 if possible, to enable SC 2 ballot results to be available to WG 2 by its meeting on 1994-10-10).

PRIN-1

UNANIMOUS

WG 2 resolves that it will not add new characters to the code positions of ISO/IEC 10646-1, until pDAM 1 is formally approved by SC 2 letter ballot, except for the use as defined in pDAM 1.

6.2 UTF-8 Related

The relevant documents are: N993 and N 996. N 996 is for Information only -- reference was made to the 'copyright and patent statements' in the document. We cannot use the material as text in the standard - can only be used for information to WG 2 members.

6.2.1 N 993 - pDAM2

Discussion:

- a. Mr. Bruce Paterson: Several editorial errors are there. Suggest including consideration for answering the concerns re: UTF transforms between UCS and UTF etc. (see discussion and proposed text under 6.1.7. above). If UTF-8 is applied directly to UTF 16 rather than UCS-2 or UCS-4, we may end up in possible dual coding. One way may be to exclude s-zone 16 bits.
- b. Definition - of UTF - should spell out in such a way that it will make it clear that it is transformation from and to UCS-4 (or UCS-2) as canonical form.
- c. Error handling related to 'incomplete' character, much similar to the problem discussed under UTF 16, also needs to be addressed.
- d. There is an explanatory first paragraph in N 993. Do we need an explanatory text or preamble as to why an annex is included in the final text? The informative section is to be moved to a footnote.
- e. **Disposition:** Appropriate text was drafted (following the text proposed in N 1027) by Dr. Asmus Freytag, and was accepted. There are several inconsistent uses of 'characters' - Mr. Bruce Paterson provided appropriate suggested rewordings (see text in resolution pDAM-UTF-8 in item 6.2.3. below).

6.2.2 Proposal to Remove UTF-1 - Mr. J. Bettels

Towards avoiding having to add an explanatory clause about UTF-s in general in the main body of 10646-1, Mr. Jürgen Bettels proposed that UTF-1 be removed from the standard, and that UTF-16 be renamed as UCS-2E, leaving UTF-8 as the only UTF.

Discussion:

- a. Dr. Asmus Freytag: Supported the proposal. Proliferation of UTFs is not good. If this annex cannot be withdrawn strong statements to deprecate this must be used. Will support removing it UTF-1 from standard and move to registry.
- b. Mr. Michel Suignard - Supported the idea.
- c. Dr. V.S. Umamaheswaran - Opposed. Annex 1 withdrawal may be considered as a major change; need to get confirmation from ITU-TS national body (Canada) before Canada can support the proposal. It is significant enough that it will need a national ballot. Where would be the home for UTF-1 if it is moved to a registry and taken out of the standard?
- d. Mr. T.K. Sato - it will need a national ballot.
- e. Mr. J. van den Beld - if we remove annex on UTF-1 we also should withdraw its registration. An alternative would be to move the text of annex 1 on UTF-1 to a Registry for those people who are using it.
- f. Mr. YongJun Zhou: some people in China use UCS-1. However, China will verify whether UTF-1 implementation was only experimental and whether UTF-1 is replaced with UTF-8 in their implementation? While UTF-1 may not be suitable for UNIX file system it is still a candidate for Communications.
- g. Mr. A. Kido - SC 22 WG 20 / POSIX - discussing use of code independent mode of operation. UTF-1 is still a candidate for interchange.

- h. Mr. Joachim Friemelt: Suggests that UTF-1 could be removed. A Registry may contain a statement that the registered item is not used in a standardized way.
- i. Prof. K. Kim - If UTF-1 is moved to a registry, can we remove a registered entity?
- j. Mr. John Clews - Several registered items in the past have been withdrawn.
- k. Professor K. Shibano: Does not UTF-1 also cater for 7-bit use? Answer: no it is an 8-bit UTF.
- l. **Disposition:** Proposed to draft a resolution: UTF-1 will be withdrawn from Annex of 10646. Move the text of Annex 1 on UTF-1 from 10646-1 to ISO Register. See resolution UTF-1 under item 6.2.3 below.

Other Concerns:

- Mr. J. Bettels pointed out the missing is the ASN.1 identifiers for UTF-8 and UTF-16. These are to be taken up with ASN.1 group.
- Dr. V.S. Umamaheswaran - Definition of UTFs in the main body is still OPEN... Currently they are embodied in the appropriate UTF-16 and UTF-8 pDAM texts.
- X Open consortium is a user of UTF-8. Liaison could be Mr. Jürgen Bettels. SC 22 WG 20 - is another channel for X Open requirements. No decision was taken at this meeting.

6.2.3 Relevant Resolutions

pDAM-UTF-8:

UNANIMOUS with Amendments

WG 2 instructs its editor to revise, update and forward to SC 2 secretariat for further processing pDAM 2 in N 993 as follows:

- i. Amendments to Draft Annex N, as in N 993
 - a. Move three paragraphs, after the title, to a Note at the end of the Annex.
 - b. In N.2 replace first sentence by:
In UTF-8 the characters from this standard shall have coded representations that comprise sequences of octets of lengths 1, 2, 3, 4, 5 and 6 octets.
 - c. In para 2, 3rd line of the table, add after 0000 FFFF:
(excluding 0000 D800 to 0000 DFFF)
 - d. In para 3, first item, add:
Coded representations in the range 0000 D800 to 0000 DFFF shall be excluded from this conversion process.
 - e. Para 2, line 1, add "coded" before "character"
Para 3, line 3, delete "character".
Para 4, item 1, line 3, replace "character code" by "coded representation"
 - f. Amend all text of N.2 to Normative style, using "shall".
 - g. In N.1 add new bullet:
The transformation defined for UTF-8 is applicable to characters coded according to UCS-4 or UCS-2, but not for characters coded according to UTF-16.

If it is desired to transform data coded in UTF-16, this can be accomplished first by transforming from UTF-16 to UCS-4 (see _____) before applying UTF-8.

- ii. Remove all references to UTF-1 from pDAM 2 subject to the acceptance of Resolution UTF-1.
- iii. Add the following text to clarify handling of incorrect octet sequences in the UTF-8 form:

nn Incorrect Sequences of octets: Interpretation by receiving devices

If a CC-data-element includes either:

- a sequence of a first octet that is not immediately followed by correct number of following octets,
- or
- one or more adjacent following octets that are not required to complete a sequence of first and following octets,

then according to Y.nn (nn) such a sequence of octets is not in conformance with the requirements of UTF-8. It is known as a malformed sequence, and it shall not be transmitted by an originating device that is in conformance with the requirements of UTF-8.

If a receiving device receives a malformed sequence, because of error conditions either:

- in the originating device, or
- in the interchange between the originating and the receiving device, or
- in the receiving device itself,

then it shall interpret that malformed sequence in the same way that it interprets a character that is outside the adopted subset that has been identified for the device (see 2.3c).

UTF-1:

APPROVED - CANADA Abstains

WG 2 instructs its editor to

- a. Prepare a new corrigendum to remove the informative Annex G on UCS transformation format (UTF-1)
- b. Transfer the current content of Annex G to ISO Registry following ISO 2375.

7 Repertoire issues

7.1 Procedures and principles of allocation

Input Documents:

- N 946** Proposed principles and procedures for allocation of new characters and scripts; Davis/Hart/Lindberg; 1993-11-03
N 947 A proposed initial list of character allocations; Davis/Hart/Lindberg; 1993-11-03
N 973 National position on N 884; K. Kim, Kang and Ahn; KBS, Korea; 1994-01-20
N 1002 Comments on N 947 "Proposed categorization and allocation of characters"; Japan (TKS); 1994.03.28

Documents N 946 and N 947 were not discussed. They were re-confirmed as the principal documents on principles and starting classification list - see resolution PRIN-3 under section 7.1.3.

7.1.1 N 1002 - Mr. T.K. Sato, Japan

- The first two paragraphs are only aimed at WG 2. The third proposal may be applicable everywhere.
- Some minimum set of requirements are to be laid on the requester to enable the processing of a national body request easier at the publication stage. For example, rules like Machine Readable Information be made available.
- Will information related to how to process the data be made available? Algorithms re: sorting etc. are optional. Would like to see if a combination is involved what would be the combinations used to produce typical shapes, especially if it is a complex case.

Discussion:

- a. Mr. John Clews: Some proforma will be needed to ask for information that will help WG 2 decide better as to what goes into Category A for example. Some rules such as 'should be a national standard at the national level etc.', information such as how the script is used, and some information about the characteristics were suggested.
- b. Dr. Asmus Freytag: prefer to have the definition as is for the different categories. The machine readable form should not be asked at the early phase.
- c. The convener - Machine readable form of a glyph associated with a character is required when the character is accepted for inclusion in the standard. If we have a proforma we should have a clear identification of what the form is to be used for.
- d. China: is there a constraint based on the number of users of the script? Answer is NO.

Disposition: Accept N 1002 as input for inclusion in N 946. See resolution PRIN-2 in section 7.1.3. Invite Mr. John Clews to draft and propose a proforma incorporating his suggestions.

7.1.2 N 973 - Professor K. Kim

N 973 - Professor K. Kim explained the contents of this paper. This paper relates several characters that have been considered in the past in the categorization.

- Item 1 - of N 973 - Two Bangeom characters in Category B to be added - repertoire issue ref: N 935.
- Item 2 - of N 973 - Gugyeoul characters to be added to Category B - repertoire issue ref: N 936.
- Items 3 to 6 - of N 973 are categorization issues. N 973 Item 4 - is not relevant since N 884 is overridden by N 974.

Discussions on N 973 Item 2: (About 3000 characters...)

- a. Mr. John Clews: there is some confusion in 946. For example, the contemporary Hanja characters are part of Category A. Numbers like 3000 is in the list - before we see the actual input how can we decide whether the characters indeed is Category A.
- b. Mr. B. Paterson: referring to Point 3 in Professor K. Kim's paper - some clarification is needed.
- c. Mr. T.K. Sato - looks like the discussion is mixing up the Guidelines and the Starting List proposed in 947.
- d. Mr. Akio Kido - IRG has a discussion item in the next meeting - in terms of prioritizing IRG repertoires. At the present time they cannot specify whether Han characters are Category A.
- e. Professor K. Kim: - informed WG 2 that the proposed 3000 characters are in current use in word processors.
- f. Mr. YongJun Zhou: we need principles on how to deal with Han characters, for IRG to go by.
- g. Dr. Asmus Freytag - In principle WG 2 should expect IRG to suggest a prioritization for all ideograph characters.

Disposition: IRG is requested to suggest a prioritization of all ideographic characters defined by it.

N 973 Item 5 - Cut off number of characters be changed from 160 to 1 row.

Discussion:

- a. China: Suggest that the cut off is 1 row.
- b. Dr. Asmus Freytag: 160 is just a guideline as an example.

Disposition: The number 160 is only an example, it is not a FIRM cut off number. No need to change.

N 973 Item 6: Remove TeX from Category A.

Discussion:

Dr. Asmus Freytag: TeX is cited only as an EXAMPLE. American mathematical society uses TeX as the means for publishing Scientific documents. The character repertoires required by TeX - those that are

deemed to be CHARACTERS - not presentation variants - are candidates for consideration. There are about 5000 or so symbols used, many of which are already coded as mathematical operators.

Disposition: No change is needed.

7.1.3 Relevant Resolutions

PRIN-2

UNANIMOUS

WG 2 accepts the proposed categorization of characters as proposed in document N 1002 and instructs its editor to revise and update N 946 accordingly.

PRIN-3

UNANIMOUS

WG 2 reconfirms the guidelines, principles and procedures in document N 946 for allocation of new characters and scripts to ISO/IEC 10646-1. All requesters for additional characters in the standard should take note.

7.2 Additions to repertoire

Input Documents:

- N 963 Proposal for encoding Mongolian script; China; 1994-01
- N 964 Proposal for encoding Tibetan script; China; 1994-01
- N 965 Proposal for encoding Yi script; China; 1994-01
- N 966 Proposal for encoding De Hong Dai script; China; 1994-01
- N 967 Proposal for encoding New Xishuangbanna Dai script; China; 1994-01
- N 968 Proposal for Mongol Character set; Mongolian National Institute; 1993-12-22
- N 977 Full width Macron in 10646 (FFE3); Takayuki K. Sato, Japan; 1994-01-03
- N 978 Comment on document N 956, Report on the Encoding of Canadian Aboriginal Syllabic languages; Michael Everson, Ireland; 1994-03-04
- N 981 Coding of Indic Conjuncts; Hugh McG Ross, U.K.; 1994-03-23
- N 982 On Truly Syllabic Scripts; Hugh McG Ross, U.K.; 1994-03-17
- N 983 Individual Characters for First Addendum; Hugh McG Ross, U.K.; 1994-03-23
- N 984 Canadian Aboriginal Syllabics: Character Set Coding Requirements; Canada; 1994-04-18
- N 985 Electrotechnical Symbols for First Addendum; Hugh McG Ross, U.K.; 1994-03-28
- N 986 Proposal for Tibetan Script in the BMP; Bruce Paterson, Peter Lofting, U.K.; 1994.03.24
- N 987 A real test for updating of 10646-1; Peter Fenwick, U.K.; March 1994
- N 990 Cross-reference of WG 2 documents to scripts and repertoire issues; Ed Hart, US; 1994.04.14
- N 1003 Subsetting of 10646 using sub repertoires; Japan (TKS); 1994.03.28
- N 1005 List of equivalencies; Alain LaBonté, Canada; 1994.04.14
- N 1011 A proposal about installing Mongolian; China; 1994.04.18
- N 1012 Proposal to supplement repertoire of Uighur, Kazakh and Arabic; China; 1994.04.18
- N 1013 Motion on coding of old Xishuang Banna Dai Writing in 10646; China; 1994.04.18
- N 1016 Digraphs in Dutch; J.W. van Wingen; 1994.04.10

N 963 is withdrawn in favour of N 1011 - Mongolian. Document N 984 replaces N 956 on Canadian Aboriginal Syllabics. N 990 attempts to cross reference all the documents prior to those listed above. Member bodies are invited to verify and enhance document N 990.

8 Other business

8.1 Defect reports

Input Documents:

- N 974 Defect Report on the term "conjoining"; Kyongsok Kim, KBS, Korea, 1994-01-20
- N 975 Defect Report on incorrect Korean glyphs; Ta-Jin Kang, KBS-Korea; 1994-01-20
- N 989 Several Defect Reports; BSI U.K.; 1994-04-14
- N 992 Editorial Correction - U+2351 - ... APL function character; Canada; 1994.04.18
- N 994 10646-1 Proposed Draft Corrigendum 1; Mark Davis, WG 2 Project Editor; 1994.03.03
- N 995 10646-1 Proposed Draft Amendment 3; Mark Davis, WG 2 Project Editor; 1994.03.03
- N 999 Defect report on 10646-1, clause 17; Japan (TKS); 1994.03.28
- N 1004 Defect Report on 10646-1, Annex F, Definitions; Character Names; Japan (TKS); 1994.03.28
- N 1006 Defect report of J-column in the CJK Unified Ideographs; Japan; 1994.04.05
- N 1014 Defects of Japanese graphic characters (revised N 939); Japan (K. Shibano); 1994.04.18
- N 1015 Report of SC 2 liaison to SC 22; J. van Wingen; 1994.04.18
- N 1017 Contribution to the discussion of character names containing "Ligature"; J.W. van Wingen; 1994.04.15
- N 1018 Comments of NNI on Æ; NNI, J.W. van Wingen; 1993.12.09
- N 1019 Draft of defect report on Greek; Yannis Haralambous, via NNI (The Netherlands); 1994.04.06
- N 1022 Defect Report on Armenian; Michel Suignard, AFNOR, France; 1994.04.19
- N 1023 Correction of Some Hebrew Characters Name; Technical Correction; Israel, Stefan Fuchs; 1994.04.19
- N 1025 Hebrew Character names in IS 10646-1; Annotation of names of characters; Stefan Fuchs - (Standards Institution of Israel); 1994.04.20
- N 1030 Defects in tables 17-25, and suggested defect report strategy; N. Subramanian, Centre for Development of Advanced Computing, India, John Clews, SESAME Computer Projects, UK; 1994.04.20

Output Documents:

- N 1028 Clarification on the UK Defect Report WG 2 N 989 concerning tables 39 and 40; Michel Suignard, AFNOR, France, Expert Contribution, April 20, 1994.
- N 1031 Disposition of Defect Reports and Instructions to the editor of ISO/IEC 10646-1, Antalya, April 18-22, 1994

An ad hoc group - Mr. J.W. van Wingen, Mr. Michel Suignard, Mr. Jürgen Bettels and Mr. Arnold Winkler as volunteers - was created and charged with drafting editing Instructions to be discussed and approved by WG 2 meeting, on those defect report items that are deemed to be Editorial Defects.

8.1.1 Discussion on Procedures for Dealing with Defect Reports

- a. JTC 1 Procedures are being drafted to replace the 1992 procedures. These classify the defects as Editorial or Technical in nature. National bodies are requested to follow these procedures and classify their defect reports.
- b. For editorial defects we can form a small editorial sub group to come up with proposed changes and present to WG 2. The technical defect reports should be discussed by WG 2 and WG 2 has to formally respond back on each technical defect and report back to SC 2.
- c. Both types of defects require a response and proposed correction text. In addition technical defects require a CORRIGENDUM to be prepared and processed.
- d. What makes a defect a TECHNICAL or an EDITORIAL? We need a formal guideline.
- e. Glyph Shapes - can be interpreted to be Editorial. Names of Characters - may be interpreted as TECHNICAL. Consequences to the implementer is important no matter how inadvertently the defect was introduced. A third category called TECHNICAL Addition can also exist.
- f. Classification of a defect as editorial or technical can be done based on JTC 1 procedures. However, acceptance and processing within the committee is to be done on a case by case basis.

8.1.2 N 974 - Clarification of the term Conjoining - Korea

Category: EDITORIAL.

Reference: Section 24 on page 14 of 10646-1.

Presented by Professor K. Kim.

The proposal is to reword '... Hangul Jamos are conjoining characters' to '... Hangul Jamos are not combining characters ...'. The term 'conjoining' and 'combining' are confusing. Word conjoining is not defined in 10646. Use the term 'non-combining' instead of conjoining. There is an alternative proposal to define perhaps three categories: Combining, Non-Combining or Conjoining. Hangul Jamos are different from the normal combining characters and non-combining characters.

Discussion:

- a. Mr. Bruce Paterson: Remove the sentence that causes the confusion. Remove the FIRST sentence. Keep the Level 3 statement. Add a Note: Hangul-Jamos are Not Combining Characters. The standard calls for 'Level 3' for dealing with Jamos. Level 1 explicitly excludes Jamos, Level 2 also does that.
- b. Mr. Michel Suignard - structure the sentence to say 'Jamos do not require combining characters ...'.

Disposition: Mr. Paterson's suggestion was accepted (see document N 1031 for corrigendum text). The open question about whether to add a third category is now mute.

8.1.3 N 975 - Incorrect Glyphs - Korea

Presented by Professor K. Kim.

Category: EDITORIAL

There is discrepancy between a few Character name and corresponding Glyphs. The Glyph is INCORRECT. Character name is CORRECT.

Disposition: Accepted (see document N 1031 for corrigendum text). Korea is requested to forward the set of bit maps, and or the outline of the shapes along with a blown-up (96x96 bits) hard copy is needed by the editor.

8.1.4 N 989 - UK Several Defect Reports

Presented by Mr. B. Paterson.

Correct a defect in the defect report: - Page 012 - suggested change is
In Name at 0189 change character name from LATIN CAPITAL LETTER AFRICAN D
to LATIN CAPITAL LETTER OF LETTER D WITH TAIL

N 989 - First Item; PU Zone definition.

Category: EDITORIAL

Change the title of Clause 11 --'and zones' to be added to the title.

Disposition: Accepted (see document N 1031 for corrigendum text).

N 989 - Second item: 01E0 and 01E1, Replace Dot Above with Ring Above in name

Category: TECHNICAL

Change name of 'Dot Above' with 'Ring Above' in positions 01E0 and 01E1.
The error came about possibly from the fact the 'RING' appeared as a 'DOT' because it was small - and got called 'DOT' above.

Discussion:

The statement 'are used in Skolt Lapp' in the Defect report was pointed to be an error by Norway.

Disposition: Because it is a proposal to change 'normative text of the standard' and is considered not desirable, UK accepts that the defect report will be withdrawn and will submit two additional characters to be added with the correct name and correct shape.

N 989 - Third item: Several Glyphs - draw Capital Upsilon with straight arms

Category: EDITORIAL.

Disposition: Accepted (see document N 1031 for corrigendum text).

N 989 -Fourth item: Glyph changes to Malayalam script

Category: EDITORIAL.

Disposition: Accepted (see document N 1031 for corrigendum text).

N 989 - Fifth item: Greek character name change

Category: TECHNICAL

Disposition: The proposed changes can be controversial - ON HOLD. The convener will write a letter to the Greek member body inviting them to comment on the UK proposal.

N 989 - Sixth item: Arrows - transposed glyphs

Category: EDITORIAL.

Disposition: Accepted after verification with UNICODE representative (see N 1028) (see document N 1031 for corrigendum text).

N 989 - Seventh item: - Mathematical Operators 22A2 and 22A3 - Transposed Glyphs

Category: EDITORIAL.

Disposition: Not accepted, after verification with UNICODE representative (see N 1028)

N 989 - Eighth item: 22A4 and 22A5 - mathematical operators 22A4 and 22A5 - Transpose Glyphs

Category: EDITORIAL.

A change from proposal in defect report.

Disposition: Not accepted, after verification with UNICODE representative (see N 1028).

N 989 - Ninth item: 234A - Overbar missing in Glyph

Category: EDITORIAL.

Disposition: Accepted (missing from list in N 1031 - N 1031 will be revised))

N 989 - Tenth item: 0189 - 'African D stroke' - proposed name change

WITHDRAWN.

N 1028 - Clarification on the UK Defect report - was prepared after discussion between Mr. B. Paterson, Mr. Michel Suignard and Dr. Asmus Freytag on some of the UK defect reports. This document was distributed and dispositions of the defect reports (in the above list) are based on this document. Except for the last sentence on 'consistent naming for characters with tacks' the rest of the text in N 1028 was accepted.

8.1.5 N 992 - Canada

Category: EDITORIAL.

Missing overbar on the glyph shown for character 2351 APL FUNCTION SYMBOL UP TACK OVERBAR.

Dr. V.S. Umamaheswaran presented.

Disposition: Accepted (missing from the list in N 1031)

8.1.6 N 999 - Japan

Mr. T. K. Sato presented.

Three clarification requests and proposed editorial corrections to address items 1 to 3 were also included in the document. A fourth item was TECHNICAL in nature and was WITHDRAWN.

Item 1: Clarification of clauses 17.2, 17.3, 17.4 and G.5.

Category: EDITORIAL.

Regarding 17.2, the request is to add a clarification that ESC sequences are not part of CC-Data Element.

Item 2: Editorial clarification of clause 17.5

Category: EDITORIAL.

Discussion:

- a. Clarification is needed to remove any ambiguity about the use of ESC sequences. The ESC sequences can be used for example, in tape labels etc. not necessarily in an ISO 2022 environment. The word 'return' adds to the confusion - gives the impression that 2022 is always the default status.
- b. Mr. B. Paterson and Mr. T.K. Sato worked ad hoc on this item. The word 'return' causes confusion because one is not sure if we started from 2022. If the switch took place from 2022 in the first place, then the word 'return' makes sense. In section 17.5 it is proposed to use " When the escape sequences from clause 17.2 are used, then "... instead of " ... from ISO 2022 ... ".

- c. Problem of switching from 10646 to other coding systems may still be there. However, it is a subject matter for a technical enhancement. This defect report addresses only the restriction to return to 2022. Japan is satisfied with the proposed wording changes.

Item 3: ESC sequences specified in 17.3 and 17.4

Category: EDITORIAL.

The sequences are aimed at permitting the use of the escape sequences in any environment - within a CC Data Element of 10646 or any other standard.

Disposition of items 1 to 3 above: Japan agreed to processing the defects in items 1 to 3 in N 999 in such a way as not to change the intent of current clause 17. See document N 1031 for proposed text as instructions to the editor.

8.1.7 N 1004 - Japan

Presented by Professor K. Shibano.

Category: Each item is categorized differently in N 1004.

30 different items are included. Of these some items were discussed during the meeting. Others were delegated to the Ad Hoc group on Defect Reports to consider. See N 1031 for the disposition of each item.

Item 19 - Square Root - the glyphs are different - with and without a BAR - but they represent the same character SQUARE ROOT.

Disposition: They were clarified to be the same. No action.

Item 20 - 'SIMILAR TO' operator is used in Japanese Ministry. The name 'TILDE' implies some connotation of an accent. Request also change in name.

Discussion:

- a. Dr. Asmus Freytag: There is a note in UNICODE document that the SIMILAR TO and REVERSE TILDE are really Glyph Variants. No need to change the name.
- b. Mr. Bruce Paterson: there is a possibility to use in parenthesis explanatory note.
- c. The convener - the parenthesis method should be used exclusively as a clarification about the character itself than an alternative name.
- d. Dr. Asmus Freytag: In this particular case it may be a border line case. If Japan is satisfied with the clarification that the two characters are synonymous there is no need to change.

Disposition: Japan accepts the clarification.

Item 22 FULL WIDTH MACRON and OVERLINE

The name is considered a mistake.

Disposition: China, Korea and Japan met with the UNICODE representative to determine if there is a need for a separate OVERLINE character or not. Compatibility characters are usually for those that are in current standards but may be glyph variants of others. We have not had a clarification as to their use in Korea and China... still on HOLD.

Item 12 - Clause 16: Use of control functions: A clarification statement about who is responsible for dealing with control function semantics.

Disposition: No action for now. This is for further discussion in WG 2.

8.1.8 N 1006 - Japan

Category: EDITORIAL

9 Ideograph Character shapes are shown to be in error.

Delegated to Ad Hoc group

Disposition: The proposed corrections were accepted. See document N 1031 for corresponding instructions to the editor. Japanese member body is requested to forward a set of bit maps and /or the outline of the shapes along with a blown-up (96x96) hard copy to the editor.

8.1.9 N 1014 - Japan

Category: EDITORIAL

Professor K. Shibano presented.

Response to an Action Item from Washington. This paper is update of N 939 discussed in Washington. Items 1 to 5 have been dealt with in Washington. Items 6 and 7 are delegated to the ad hoc group on Editorial Defect Reports.

Disposition: All the items in N 1014 were accepted (see document N 1031 for editorial instructions on items 6 and 7). Items 1 through 5 were accepted in Washington. Japanese member body is requested to forward a set of bit maps and /or the outline of the shapes along with a blown-up (96x96) hard copy to the editor.

8.1.10 N 1018 - Mr. J.W. van Wingen

Mr. J.W. van Wingen presented the document. The Netherlands member body questions the decision of WG 2 to rename some of the Ligatures as Letters. Wants WG 2 to reverse the direction accepted in Washington. Mr. van Wingen claims that the 'Letter' in DIS was a mistake and the editor had corrected the DIS at that time to 'Ligature'.

Discussion:

- a. Dr. Asmus Freytag: we had an in depth discussion at Washington. The DIS ballot was also using word 'Letter'.
- b. It was observed that the concern is about the procedure on how to process the 'draft corrections prepared' by the editor in document N 994.

Disposition: See resolution DR-2 under item 8.1.19. Document N 994 containing the corrigenda related to AE Ligature will be processed as a TECHNICAL CORRIGENDUM. This will give the Netherlands member body and other member bodies an opportunity to ballot on the proposed corrigendum.

8.1.11 N 1022 - Mr. Michel Suignard

Category: EDITORIAL

Mr. Michel Suignard presented.

Some shapes for Armenian are incorrect. The document proposes some changes especially to the graphic symbols for small letters.

Discussion:

- a. The convener: Some opinion from Armenian experts should be attached to defect report.
- b. Mr. J.W. van Wingen - there are still some errors in the French proposal.

Disposition: A formal defect report is requested to be generated by France. Mr. J.W. van Wingen is invited to contribute.

8.1.12 N 1007 - Turkey

Mr. Ümit Karakas presented.

Category: TECHNICAL

Towards consistency in locating characters in similar groups together, this document proposes relocation of some characters such as Superscript 1, 2 etc. to the appropriate group of characters - for example, Table 36 Row 20 location 113.

Discussion:

- a. The convener: the standard has been implemented at least in one system and being implemented by many. Any changes to allocations of characters to hex values will NOT be received very well within the UNICODE community or by the US member body.
- b. Mr. Michel Suignard - questioned the proposed cedilla reallocation. Will cause problems for Latin-1 interaction with 10646.
- c. Dr. V.S. Umamaheswaran: stated that any proposal to change existing hex value allocations will not be well-received in Canada. Already it is too late to move any characters around without invalidating some database or applications.
- d. Dr. Asmus Freytag: re-emphasized the break in implementation or data using 10646 cannot be accepted.
- e. Mr. Bruce Paterson supported the view that reallocation of characters is not desirable. Also, duplications are not desirable and should not be permitted within 10646.
- f. Mr. Jürgen Bettels - was not supportive of reallocation of characters - lesson we learn early on is not to mess with data already encoded.
- g. Mr. J.W. van Wingen - pointed out that deprecation is used in other standards, and that standards are known to change technically.
- h. TSE requested a discussion by the delegates attending the WG 2 meeting.

Disposition: The convener summarized: 10646-1 is already published; we are not in a position to change code positions at this time. There are several implementations in progress of this standard and WG 2 should be very careful not to impact the existing implementations. Reemphasized that duplicates should not be permitted. Deprecation as a concept may be entertained in 10646 will be to declare a code point as a permanently de-allocated. The majority opinions expressed seems to be not to entertain the proposed TSE changes.

8.1.13 N 1023 - Israel

Stefan Fuchs - Standards Institution of Israel, presented.

The character names in 10646-1 are not in alignment with those adopted by SII (Standards institution of Israel).

Discussion:

- a. Code points and Names are Normative in 10646. If they are entertained by WG 2 they have to be treated as Technical Corrigendum. 1311-2 is the Israel National Standard. The identified six characters are not part of current ISO 8859-9.
- b. For the time being the national body has adopted the method of reflecting the 10646 names as equivalents in the national standard. But their preference is to get the names adopted by the national body also be correctly reflected in the international standard.
- c. Mr. J.W. van Wingen - one can propose a technical change to be processed anytime using the existing process.

- d. The word 'point' in the name is less appropriate to use than the word 'mark' in the name of the character. The criticality of the requested change is to be decided by WG 2.
- e. The names - are unique - for the purposes of the standard. They may not always reflect the intended use of the characters. There are no semantics in the name itself.
- f. Mr. Joachim Friemelt - if the customer comes up with names used in the national standard only, and we have to look for them in 10646 - there may not be a correspondence if they are significantly different.
- g. Mr. Jürgen Bettels - the names are used in charmap resources for example in XOpen Locales. If Israel submits a locale to XOpen the 10646 names will still be used.

Disposition: Since the change requested is not considered 'Mandatory' or 'Critical' WG 2 will not process this request for change.

8.1.14 N 1025 - Mr. Stefan Fuchs

Mr. Stefan Fuchs presented.

Subsequent to the rejection of N 1023 name change request, Mr. B. Paterson and Mr. S. Fuchs worked on an alternate proposal to add annotations to existing names for six Hebrew characters in 10646-1. This is considered adequate to satisfy Israel's requirements. N 1025R was proposed as an enhancement to 1025. Some missing characters have been added.

Discussion:

- a. Mr. J.W. van Wingen - Reluctance to change the names are well understood. Would like to get a complete set of corrections in one package. (These remarks do not apply specifically to document N 1025).
- b. Mr. John Clews: Clarification question - why 'mark' is used instead of 'accent legarme'. Answer: It is in line with the usage of Israeli national standard.
- c. Dr. V.S. Umamaheswaran: the proposal looks like a reasonable compromise between Israel national requirements and the WG 2's desire not to change.
- d. Dr. Asmus Freytag: There is a potential impact on people who maintains a database. It is convenient for some people and inconvenient to others. Israel is asked -- 'is this really necessary?'. If the answer is NO then the preference should be not to change.
- e. Professor K. Shibano: There is confusion as to whether annotations such as those proposed in N 1025R are permitted under current naming guidelines in 10646-1.
- f. Mr. Joachim Friemelt: Concern is ' what are the rules regarding determining when we have to change the names of a character in the standard?'.
The convener: The rule is NOT to change any name -- exception is when it is discovered that we have made an error in the document. And if it is deemed to be an error, it will be processed as a TECHNICAL DEFECT and processed as such.

Disposition: The changes proposed in N 1025R are accepted as editorial changes. (Note this set of changes were not in N 1031. Should be added in N 1031R).

- The objection from NNI on the grounds that this change should be considered a Technical Change and not an editorial change is noted.
- Japanese member body was invited to prepare a defect report to remove the inconsistency regarding annotations in character names from Rule 12 in the annex on naming rules in 10646-1.

8.1.15 N 994 - Corrigendum 1 on Æ

The document prepared by the editor, Mr. Mark Davis, contains the proposed text for the corrigendum related to Æ ligature resolutions in Washington meeting.

Discussion:

- a. Mr. Bruce Paterson: The word 'editorial correction' should be changed to 'correction' in the corrigendum text.
- b. Mr. Bo Jensen: Danish member body requests that a statement justifying the corrigendum be added prior to the corrigendum text. The corresponding defect reports in N 910, 911, 912 and 923 should be referenced and attached along with the corrigendum - for letter ballot to SC 2.

Disposition: See resolution DR-2 under 8.1.19. The modified corrigendum is to be processed as a TECHNICAL CORRIGENDUM and sent to the SC 2 secretariat for ballot processing.

8.1.16 Document N 995 - pDAM 3 on C1 Defect

The document prepared by the editor, Mr. Mark Davis, contains the draft amendment text on C1 controls following the Washington meeting resolutions. The text is for a Draft Amendment (though being discussed under Defect Report).

Disposition: The editor is to be instructed to process a New Work item on the C1 controls and attach the Proposed Solution to the work item. See resolution pDAM3 under item 8.1.19.

8.1.17 N 1030 - Mr. John Clews and Mr. N. Subramanian

Mr. John Clews presented. The situation is that there is an ISCII 1991 standard. 10646-1 has incorporated an earlier working document of the ISCII standard. The background of the situation that has led to the current differences is outlined in the document. As a minimum the users of 10646 should be informed that the contents of ISCII 1991 and 10646 differ. Some options for addressing the situation are offered. The preference is towards minimizing the processing complexity.

Discussion:

- a. Dr. Asmus Freytag, Mr. Michel Suignard: The paper as it stands does not have sufficient information at hand to make a judgment or being able to evaluate the problem and the proposed options.
- b. The convener: it was unfortunate that the latest ISCII standard was brought to the attention of WG 2 Seoul 10646-1 editing meeting at the last minute when one could not do anything about the contents of 10646 standard.
- c. Mr. John Clews: A revised contribution with supportive material following the guidelines of WG 2 can be submitted for consideration at the next meeting.
- d. Mr. Bruce Paterson: ISCII is an 8-bit code; it is not a replacement for 10646. The question raised is -- whether or not a simple algorithm versus a complex one should exist for mapping between the two. The repertoire differences may be reconciled -- a few characters are missing or are found in 10646.
- e. Mr. Michel Suignard: Section 18 of the standard already states that a transformation employing table look up is needed.
- f. Mr. N. Subramanian: it is to be pointed out that the reference in the document be changed to assist the developers understand the document is not quite in line with ISCII -1991.

Disposition: The Indian member body is invited to prepare a defect report asking for the appropriate editorial change - following the JTC 1 Defect Reporting procedures (and forms).

8.1.18 N 1031 - Ad hoc group on editorial defect reports

- A list of suggested editing instructions was prepared by the ad hoc group. Mr. Jürgen Bettels presented the document. One item from Mr. Michel Suignard and editorial corrections from Mr. Bruce Paterson were noted as missing. N 1031R will be prepared and submitted along with a supporting Resolution (see resolution DR-1 under item 8.1.19.)
- (N 1031R will be sent by the convener by mail - was not ready by the end of the meeting).
- (Meeting secretary's note: several other editorial defects accepted and itemized under the dispositions listed in the items 8.1.xx above were also not in N 1031. These have been pointed out under the dispositions above. I believe these should be also included in N 1031R).

8.1.19 Relevant Resolutions

DR-1

UNANIMOUS

WG 2 instructs its editor to prepare editorial corrigenda forms from the instructions to the editor contained in N 1031R. Further the editor is instructed to create the appropriate response documents and forward them together with the Corrigenda Forms to the SC 2 secretariat for forwarding to ITTF.

DR-2

UNANIMOUS

WG 2 instructs its editor to revise and update the Corrigendum 1 in N 994 as follows:

- a. Change the words 'editorial correction' to 'correction' in the corrigendum text.
- b. Add a statement justifying the corrigendum prior to the corrigendum text, referencing and attaching the defect reports in N 910, 911, 912 and 923 to the text.

Further the editor is instructed to forward it to SC 2 secretariat for further processing as a Technical Corrigendum, along with the response forms for the defect report.

pDAM3

UNANIMOUS

WG 2 instructs its editor to prepare text for pDAM 3 based on the text in N 995, and forward it to SC 2 secretariat for further processing as a pDAM.

8.2 Liaison reports

Input Documents:

- N 959** Application for Liaison member to WG 2; Hong Kong Information Technology Federation Ltd. (HKITF); 1993-12-17
- N 960** Liaison statement to SC 2 re 10646; ITU-TSS Study Group 8; 1993-11-25
- N 961** Status and font samples of digitized Chinese (Hanzi) font manufacturers in Taiwan; jack Kai-tung Huang; 1994-01
- N 1005** List of equivalences between precomposed and composite sequences - List of lexical character sequences in general - Feedback and new request from SC 22/WG 20; Alain LaBonté, SC 22/WG 20; 1994.04.14
- N 1020** Appendix to N 960; Clarification to Liaison from ITU-TSS SG 8 relating to naming rules in 10646; Stefan Fuchs, Israel (personal contribution); 1994.04.18
- N 1026** Liaison Report; 'Encoding Newsletter', The Unicode Consortium; Spring 1994
- N 1029** Equivalence between pre-composed characters and their corresponding composite sequences; The Unicode Consortium; 1994.04.20
- N 1032** Keld Simonsen - WITHDRAWN

N 959 - Liaison from HKITF.

Liaison request is under SC 2 ballot. Another liaison request from Taipei is also under SC 2 ballot.

N 961 - is an information document listing a set of fonts available in support of 10646? WG 2 does not pass any value judgments on the fonts that are identified in this document. Dr. Asmus Freytag informed the meeting that the Unicode consortium is currently in the process of mapping to several available fonts. This information will be made available to anyone requesting from Unicode consortium.

8.2.1 N 1005 - Mr. T.K. Sato

- Mr. T.K. Sato presented the liaison report from SC 22/WG 20. The original request from SC 22 / WG 20 for normalization was considered by WG 2 in Washington. While normalization was not feasible, equivalency table was suggested and national body contribution was sought. A sample format was also considered desirable. These messages were sent to SC 22/WG 20 as liaison response.
- Document N 1005 contains SC 22/WG 20 response to the above. They accepted the concept of equivalency table. A sample format has been attached to N 1005. There is also the notion of non-equivalences.
- Mr. Alain LaBonté explained further: there is a need to process what is known as 'text elements' in the work of SC 22/WG 20. The equivalences are both to characters that are fully coded and possible equivalent combined character sequences.

Discussion:

- a. Dr. Asmus Freytag: The Unicode consortium has constructed a list of 'de-compositions' for use within the scope of encoding - in a language independent manner. Use of these sequences in a particular language etc. are outside the scope of 10646. There will be a contribution made by the Unicode consortium on 'decomposition equivalents'.
- b. Resolution TKS-3 - from Washington is referenced (reproduced below for convenience).

Disposition: Accept N 1005 as an input document to the editor. The working group recognizes the need for equivalence table (as requested in N 926 item 2) between pre-composed characters and their corresponding composite sequences of UCS characters in UCS Level 3. All member bodies are invited to contribute input towards developing an informative Annex to ISO/IEC 10646-1 for consideration at the next SC 2/WG 2 meeting. The editor is encouraged to prepare a draft text based on contributions received.

8.2.2 N 1029 - Dr. Asmus Freytag, The Unicode Consortium

- The document is presented as an information document, as an input towards the Informative Annex on Equivalence between pre-composed characters and their corresponding composite sequences.
- Each precomposed character is broken down into its components and each one is given an equivalent sequence. Attention was also given to the sequence of the combining marks, with principles of 'if there is a substantive change in its appearance from a presentation process' then they are to be treated different. In the tables attached to N 1029 the first name is the ISO name, the second name is the Unicode 1.0 (historical) reference name.
- Additional information can be obtained from Unicode Consortium - contact Dr. Asmus Freytag.

Disposition: Accept N 1029 as an input document to the Editor.

8.2.3 N 1024 - Mr. Alain LaBonté - SC 18/WG 9

- The document explains how JTC 1/SC 18/WG 9 makes use of characters in the work of SC 18/WG 9. Part 7 of 9995 uses the characters for symbols on keyboards. A liaison statement is already in the repertoire collection of documents yet to be acted on by WG 2.
- Complementary keyboards in part 3 also relate to work of WG 2. Part 8 deals with allocation of letters on numeric key pad - intent is to use these for text input also -- non European character use will be interesting. A new work item proposal for input methods to enter graphic characters of 10646 is also attached to the liaison report. It was explained that the input method is not to replace existing methods of entry from national keyboards for national characters. The proposal is to work on a method to be keyboard independent and not entering a numeric value of the code position but a character.

Discussion:

It was suggested that WG 9 should provide some more information where the icons or symbols are used and if they are only picture symbols that may never appear in a textual string. The response was that some of the symbols (11 or so) are already included in 10646-1.

Disposition: Input from WG 9 was included in the list of characters for consideration for additions to the repertoire.

8.2.4 N 960 - ITU TS SG 8 - Mr. Joachim Friemelt

ITU/TS SG 8 develops several recommendations involving coded character sets. They followed the ISO 10646-1 standard character naming. They found that -- even following the ISO rules - there are some inconsistencies in the names in 10646 with the rules.

Disposition: It was pointed out that we may be able to make the rules more consistent. However, existing names cannot be changed to follow the new rules.

8.2.5 N 1020 - Mr. Stefan Fuchs

This document was presented as an information document. The naming rules were modified from ISO to be able to satisfy the existing character names in current ITU recommendations.

Disposition: Mr. Stefan Fuchs and Mr. Joachim Friemelt were invited to prepare a paper containing ITU/T suggestions for clarification to the current naming rules in 10646-1.

8.2.6 N 1026 - The Unicode consortium - Dr. Asmus Freytag

- N 1026 is a newsletter called 'Encoding' Spring 1994 from the Unicode consortium.
- There are several action items that Unicode consortium has on its plate in support of WG 2 as potential contributions to its work.

- Unicode Workshop was held in Japan. One of the highlights of this workshop was that there were several working implementations using Unicode. We are no longer in the stage of a standard to be implemented.
- The newsletter contains a list of members of the consortium, and articles on Character and Glyph Model, Joiners and Ligatures, in addition to the Unicode workshop in Japan and IRG in Hanoi.
- Unicode implementers are asking the consortium questions on dealing with ligatures etc. Refer to the article on Joiners and Ligatures.
- There are six scripts with almost final coding assigned under public review at the present time. This document will be made available to WG 2. An editor is being hired to do the work by Unicode. Their work is proposed to be submitted to WG 2.
- Relationship between Unicode and WG 2 was compared with the relationship of IRG to WG 2.
- The consortium is free to encode in the Unicode and make proposals to WG 2. WG 2 is not obliged to accept that. Unicode will also work to keep their code exactly aligned with 10646 with respect to the code position allocations. IRG does not code. However they are also free to make recommendations to WG 2 as to their desired coding.
- Output from Unicode is available to anyone who wants to participate in the work of additional coding work.

Additional Information:

Mr. Sten Lindberg: Liaison matters from higher levels of JTC 1 - 'C' Liaison - are being handled formally. JTC 1 in its February 1994, Washington plenary, has recommended that "ITTF grants Category C liaison status" to Unicode and AFII (amongst others). The appropriate letters will be going out soon from ISO. -- Reference Resolution 45 of JTC 1 in JTC 1 N 2899 Attachment 1. The liaison approval from SC 2 are to be found in JTC 1 N 2828 and N 2834.

8.2.7 N 1032 - Mr. Keld Simonsen

Contributions were requested from resolution TKS-3 from the Washington WG 2 meeting. Document N 1032 was presented as a Danish contribution entitled 'Character Mnemonics, Coded Characters Sets, Information Preserving Transformation' is an input towards this work item for an informative annex.

Discussion:

- Mr. T.K. Sato: JIS has been quoted as a source in N 1032. There are some errors in the document. Denmark is requested to remove the reference to JIS. This document is to be considered as an input to the 'correspondence between existing standards and 10646'.
- Mr. Michel Suignard: the document contains several mnemonics; it adds to the confusion from reading the document. There are several characters that are missing or are incorrectly mapped. The document date is even older than 10646 publication.
- Professor K. Shibano: Source of information from JIS has been quoted as ISO Registry. Japan is in the process of revising the JIS. Would Denmark respect the responsibility and role of individual national bodies responsible for standards?
- Dr. Asmus Freytag: The intent of the resolution in Washington was to get the input on the sources of characters for the informative annex. If the document is known to contain incomplete or erroneous information it may be appropriate to request that this document be withdrawn as a WG 2 document.
- The convener: Encouraged the Danish member body to update the information removing the parts that are not relevant to the Annex under resolution TKS-3 from Washington. The request is to contribute the material that is directly related to TKS-3. The convener suggested that this document should be withdrawn by Denmark.
- Mr. Keld Simonsen: There is an updated document that may be made available to WG 2.

Disposition: N 1032 will be considered as a TEMPORARY document for this meeting only. The Danish member body is invited to update the information and contribute only the material that is relevant to the work item from resolution TKS-3 from Washington meeting.

8.2.8 Liaison from WG 2 to SC 29 - Mr. Jan van den Beld

There is an outstanding invitation from SC 29 for Liaison representative from WG 2. Member bodies are invited to nominate a WG 2 expert to perform this role.

8.3 WG 2 program of work

Input Documents:

N 957R WG 2 Program of Work

- Some useful comments were received from Mr. Bruce Paterson.
- Mr. J. W. van Wingen: Subset Registration as an item should be added. A method to describe the collections in 10646 is needed. It should be better than what is in the standard now. Sweden had taken on the role of coordinating this item. Resolution 15 in Athens - addressed that item.

8.4 Presentations

8.4.1 ISCII - N. Subramanian

- A copy of ISCII 1991 document and copies of material used for the presentation were made available at the meeting.
- The origin of several scripts in India are derived from Brahmi Script. Common code has been designated 9 Indian scripts based on Brahmi script. ISCII standard only talks about only these 9. There are also other languages used in India that use Persian scripts.

- IS 10401-1982 '8-bit coded set for information interchange' equivalent to 4873.
- About 84 characters - vowels, consonants and maatras (vowel signs). 15 languages using common phonetic script. ISCII is a super set of the requirements of all the languages. Attribute mechanisms are used for distinguishing between different scripts. An EXT is used to extend and encode ancient Vedic characters (used for writing Sanskrit - additional 32 Vedic signs following EXT). ATR is used to switch for FONT designers -- similar to the use of Shifting of 2022. Control codes are used to switch equivalent of FONTS for each Script.
- Mr. Subramanian went on to describe the various character classifications and their attributes, and how they are combined to form conjuncts (a description using the Backus Nauer form was also presented).
- Properties of ISCII code -- Typed in the order of Phonetic sequence; Direct Sorting is possible -- existing algorithms for sorting will work straight from the code points and give correct results. Transliteration is easy between the multiple languages. Many people sing in 5 or more languages - but transliterate the words in their own script. Roman transliteration is also included.
- Simply changing the font gives the transliteration automatically.
- Keyboards - the standard English QWERTY language keyboard is used. The ISCII alphabet is used on lower and upper shifts of layer two. With the help of composition in the computer the fonts for all Glyphs are composed. Additional layer is switched in. The keyboard arrangement is done in such a way that there is no need to change the KB to ASCII code mapping - a shift in the layer puts you into the right half of the ISCII code with Key positions generating the corresponding ISCII codes.
- Using the same technique, there are also codes for TIBETAN and BHUTANESE. TISCII and ISCII are related directly to each other. Allows Indian to Tibetan interchange directly. Tibetan has an additional six consonants. Six tiers of characters are used for glyph presentation. Similar work has been used on SINHALA also. Sinhalese and Tamil can be used interchangeably. Some work on Thai has also been done. There are several implementations and products available in India. A list of 'Graphics and Intelligence based Technology' from Department of Electronics of India was also given. Urdu is written from Right to Left - the implementations take care of the right to left also. Add on card is available by which you can get support for Indian scripts. A GIST Terminal with VT 220 orientation is available. An ASIC chip inside the GIST or PC automatically takes care of representing up to 17 scripts in the system. Other scripts can easily be down-loaded.
- Graphics based systems such as COREL draw etc. require different types of solutions. Fonts are generated on the fly by Font conversion algorithms.
- Electronic Subtitle Receiver - TV Signals are received in the SUBTITLE in the language of your choice. Encoding ISCII text in the invisible region of TV broadcast. Previously translated messages in 17 languages are sent on the wire - up to the user to pick the correct language on their VCR/TV set. The appropriate script for that language will display the words meaningfully. Other scripts if selected will only Transliterate.
- ISO 10646-1 has used an older internal discussion draft of ISCII. There have been communications from Indian sources through Unicode consortium in the past.

8.4.2 N 1011 - Mongolian, TODO and XIBE (incl. Manchu) - by Professor Enghebatu

- Document N 1011 (replaces document N 963) contains details on the material presented by Professor Enghebatu.
- The proposal describes the origin, historical background and current state of three scripts - Mongolian, Todo and Xibe (including Manchu).
- Present Mongolian script can be traced to 10th century. The characters are imported from Arabic script. A sample of Mongolian script is shown -- Top to Bottom Left to Right is the direction. There are 3 different forms. Head form, Middle form and Tail form. Some characters have the same glyphs but different pronunciations. Some characters have same pronunciation but different glyphs - a has 11 different forms for different positions. If the characters are based on shapes - up to 15 percent of information can be lost. Only four characters are there with no loss of information. Mongolian uses special punctuation marks and a special space.
- the 3 scripts have different origins - probably will not be used in the same document. It may be possible to define a unification and separation rule. Unification rules can save a lot of coding space. Only 160 coding positions are needed.
- If unification is not used, we will need 145 + 114 + 108 code points for Mongolian Todo and Xibe.
- We have to code on the basis of pronunciation as well as glyph. Glyph based will only satisfy the Publishing industry. Applications such as sorting or others will not be satisfied.

Discussion:

- Will there be a need for separately coding the different forms - head and middle and tail?
- What is the relationship between existing GB standard and the proposal in N 1011. XIBE (including Manchu), Todo and Mongolian? GB standard is only used for publishing and not processing.
- N 968 from Mongolian standards institute needs more characters and lacks some symbols compared to Chinese input.
- Can China and Mongolian standards institute work together to agree upon and arrive at a single input to WG 2?
- Mr. T. K. Sato: N 963 contains some presentation forms - somehow these should be considered in the new proposal.

Suggestion to Chinese body - classify the characters into canonical form separating strictly presentation forms. China is requested that China and Mongolia get together and come up with an agreed upon joint proposal between the two by the next WG 2 meeting. All WG 2 members and liaison organizations to contact WG 2 convener if they are interested in participating in the discussions.

8.4.3 N 964 - Tibetan - Professor Nyima Trashy

Document N 964 was presented. The document contains a description and background on Tibetan use. A description of the Script was presented. Has spelling procedure on both horizontal and vertical axes. 30 consonants and 4 vowels.

Discussion:

- Professor K. Shibano: The precomposed approach of China for Tibetan language is preferred. Are the characters in code positions 51 -- 62 are combining? If so can you use the current rules in 10646 for combinations?
- Some experts think - precomposed is better, others think the combining method is better. There are 705 in the present set. To write Sanskrit in Tibetan script we need composition techniques.
- The names of characters in N 964 are not fully according to the 10646 naming guidelines.
- Mr. Michel Suignard Suggested some coordination between Indian proposal to deal with Tibetan with China so that there is consistency between similar writing systems. Also the principles in N 946 and N 947 should be followed.
- Mr. John Clews: Are there any practical implementations on which the coding proposal is based? Some implementations in China use pre-composed characters. There are no GB (Chinese National) standard on Tibetan.
- Professor K. Kim: Is there no Half-Digit Zero? Answer: NO.
- Mr. T.K. Sato: If it is true that the 'combining characters' are needed only to write the Sanskrit in Tibetan. Can you separate the Tibetan combining characters separately?
- Mr. N. Subramanian: India has some work done on both Tibetan and Bhutanese. Bhutan is similar to Tibetan and needs an additional character. India is willing to work with those who are interested in Tibetan language. Is there a similar effort needed for WG 2 work on Sinhala? India is interested in participating with other groups interested in them also. We do have some information from the Sri Lankan government.
- The convener: A Tibetan Ad Hoc group with interested member bodies and liaison organizations can be formed in order to arrive at a consensus position. In the past there have been differences of opinion.

8.4.4 N 1012 - Uighur, Kazakh and Kirghiz supplements for Arabic script - Professor Wushur Slamu

Document N 1012 was presented. Several additions of presentation forms have been asked to be added to 10646-1 to deal with Uighur, Kazakh and Kirghiz use of Arabic script. There are five proposals.

Discussion:

- a. China is advised to follow the new set of guidelines out of the Washington meeting from document N 946 and N 947.
- b. Also item such as number 5, we request the Chinese member body submit a Defect Report.
- c. The compatibility zone area is meant to include only those characters that are necessary to convert from existing standards. Any addition of presentation forms to the compatibility zone is based on existing national standards. A request to add more presentation forms may not be acceptable according to the guidelines document for future additions to 10646.
- d. Is it true that these languages were written some time ago in Latin? Mr. John Clews advised that there were some representation in Latin and Cyrillic in recent times. Historically the Kazak and Uzbek script were using Arabic scripts.

8.4.5 N 965: Yi Script - by Mr. YongJun Zhou and Mr. Shamalayi

Paper N 965 was presented. Nature of Yi and demographic information are given in this document. The ancient Yi script uses about 10,000 ideographic characters. The more recently it has been normalized into syllabic version - 1165 characters.

Discussion:

- a. Has China considered composition as a technique to reduce the number of characters?
- b. A number of the characters include in position 0311 - COMBINING INVERTED BREVE of 10646 for example.
- c. China is requested to study this possibility of composition to reduce the number of characters that needs coding in the BMP.

8.5 Future meetings

- a. WG 2 meeting no. 26:
Week of 1994-10-10 (October 10, 1994) (Monday to Friday 5 days) in San Francisco Bay Area, USA;
Hosted by The Unicode Consortium.
(Note: WG 2/IRG is planning to meet during the week of 1994-08-29 in Beijing, and SC 22/ WG 20 is meeting the week of 1994-10-03 in the US West Coast.)
- b. WG 2 meeting no. 27:
Week of 1995-06-26 (June 26, 1995) (Monday to Wednesday, 3 days), along with SC 2 plenary and WG 3 meetings in Helsinki, Finland.
- c. WG 2 meeting no. 28:
Week of 1995-11-06 (November 6, 1995) in Japan (possibly Kyoto)
- d. WG 2 meeting no. 29:
Week of 1996-04-xx (April , 1996) (Monday to Friday, 5 days); Canada or Denmark

8.6 Other Concerns

- Mr. J.W. van Wingen: Netherlands received several SC 2 ballots too late from SC 2 secretariat.
- Mr. J. van den Beld: It is unfortunate that some of the documents slipped in the mailing. They have been mailed with revised deadlines. Because the next plenary is in 1995, even if the ballot responses were received by the Secretariat, he has no idea as to how to process the responses. Some comments cannot be acted upon. On the Character Glyph model - we may have to get the final work item approved at the JTC 1 level.

9 Closing

9.1 Approval of Resolutions

- An editing committee consisting of the meeting secretary, the convener and Mr. Jan van den Beld prepared the draft resolutions. These were discussed at the plenary.
- Several editorial corrections were pointed out to the secretary for inclusion in the final document of resolutions. Draft resolution pDAM-UTF-8 was amended to add a paragraph to further clarify interaction between UTF-8 and UTF-16. Wording changes were accepted in draft resolutions pDAM-UTF-8, PRIN-1, PRIN-2 (also correct a document reference from 947 to 946), and PRIN-3. Some wrongly spelt names were pointed out. Resolution ADMIN-6, thanking the presenters was added later.
- All the resolutions except Resolution UTF-1 (Canada abstained), were accepted (some with agreed upon amendments) Unanimously.
- The corrections were incorporated by the convener and the secretary and the final text of the resolutions are contained in the document N 1034. The final text of the resolutions are also embedded as 'relevant resolutions' in the various sections in this minutes document. All the ADMIN resolutions are grouped under item 9.1.1.

9.1.1 Administrative Resolutions:

ADMIN-1

WG 2 expresses its sincere thanks to Mr. Shigenobu Kato for his role as the rapporteur of the Ideographic Rapporteur Group (IRG) and wishes him all success.

UNANIMOUS

ADMIN-2

WG 2 endorses the nomination of Mr. Akio Kido as the new rapporteur of IRG, and wishes him all success.

UNANIMOUS

ADMIN-3

UNANIMOUS

WG 2 expresses its sincere thanks to Mr. Ahmet Kurter and all other members of the staff of Turkish Standards Institution(TSE) for the excellent arrangements, their kind hospitality and continued support during the WG 2 meeting.

ADMIN-4

UNANIMOUS

WG 2 thanks Mr. Joachim Friemelt for his valuable contributions to the work of WG 2 over the past several years and wishes him a happy retirement.

ADMIN-5

UNANIMOUS

WG 2 thanks Dr. V.S. UMAmaheswaran for serving as meeting secretary for the Antalya WG 2 meeting.

ADMIN-6

WG 2 thanks the following for their informative presentations on different scripts at the meeting:

- a. Mr. N. Subramanian, CDAC, Pune, India, on Indian Standard Code for Information Interchange, ISCII, and its handling of the different scripts used in India
- b. Professor Enghebatu, Mongolian Language Institute, China, on Mongolian script
- c. Professor Nyima Trashi, Tibet University, China, on Tibetan script
- d. Professor Wushur Slamu, Xianjiang University, China, on Uighur, Kazhak and Kirghiz scripts, and,
- e. Mr. YongJun Zhou, China Great-Wall Computer Group and Mr. Shamalayi, National Institute of Chengdu, China, on Yi script.

9.2 Adjournment

The meeting adjourned at 18:00h on Thursday, 1994-04-21, half a day ahead of schedule.

10 Action Items

10.1 From WG 2 Meeting 24, Washington DC

Item	Assigned to / action (Reference N 954R - resolutions, and N 955 - minutes of Washington DC WG 2 meeting WG 2 - 24)	Reference Output/Status
2. Editor - Mr. Mark Davis		
b.	to prepare an accompanying background paper showing clearly the benefits and disadvantages of the proposed UTF-16, addressing the comments and concerns brought up during the WG 2 meeting no. 24 by 1994-01-05; (reference resolution UTF16-.3)	
d.	to include some text on 'what criteria can be used for defect report on character names' as part of the Set of Principles document (reference section 9-a-i.3).	
f.	to prepare a defect report to correct the names for all the non-presentation form characters that have the term LIGATURE in their names (reference resolution AE-1), excluding those under resolution AE-2; (reference resolution AE-3).	
h.	with assistance from the Japanese member body, to prepare appropriate text for a CORRIGENDUM to address the defects reported in document N 939 under items 1, 3, 4 and 5; (reference resolution DFR-1)	

10.2 From WG 2 Meeting 25 Antalya, Turkey

Item	Assigned to / action	Reference Output/Status
1. Meeting Secretary (Dr. V.S. UMAmaheswaran):		
a.	To revise N 1034 to incorporate all agreed to amendments and editorial corrections to in the draft resolutions.	
b.	To prepare the meeting minutes (unconfirmed) N 1033. Both these documents are to be sent to the convener as soon as possible for convener's distribution to WG 2 members and member bodies.	
2. Convener (Mr. Mike Ksar):		
a.	To write a letter to the Greek member body inviting them to comment on the UK proposal on defect item number 5 on Greek character name change in N 998.	
b.	Update document N 957-R containing the WG 2 program of work - capturing all the comments received under item 8.3 in this set of minutes.	
c.	Look at JCT 1 procedures regarding processing pDAMs, and advise the editor on the different options.	
3. Project Editor (Mr. Mark Davis):		
a.	Act on resolution pDAM-UTF-16.	
b.	Act on resolution pDAM-UTF-8.	
c.	Act on resolution UTF-1.	
d.	Act on resolutions DR-1, DR-2 and DR-3.	
e.	Act on resolution pDAM3.	
f.	Act on resolution PRIN-2.	
g.	Take note of contributions in N 1005 from SC 22/WG 20 and N 1029 from the Unicode Consortium towards work on the annex on Equivalence of combined sequences per resolution TKS-3 from Washington WG 2 meeting 24. Editor is invited to draft the annex text based on contributions received.	
4. French and the Netherlands member bodies.		
	Invited to prepare a formal defect report on Armenian characters (described in document N	

	1022). Mr. J.W. van Wingen had volunteered to contribute.	
5.	Japanese member body	
	Requested to forward a set of bit maps and /or the outline of the corrected shapes reported in document N 1006 and N1014 along with a blown-up (96x96) hard copy to the editor.	
	Invited to prepare a defect report to remove the inconsistency regarding annotations in character names from Rule 12 in the annex on naming rules in 10646-1.	
6.	Korean member body	
	Korea is requested to forward the set of bit maps, and or the outline of the corrected shapes of the characters in defects in document N 975, along with a blown-up (96x96 bits) hard copy is needed by the editor.	
7.	Messrs. Joachim Friemelt and Stefan Fuchs	
	Invited to prepare a paper containing ITU/T suggestions for clarification to the current naming rules in 10646-1.	
8.	Indian member body	
	The Indian member body is invited to prepare a defect report asking for the appropriate editorial change regarding references to ISCII 1991 - following the JCT 1 Defect Reporting procedures (and forms).	
9.	Danish member body (Mr. Keld Simonsen)	
	The Danish member body is encouraged to update the information in N 1032 removing the parts that are not relevant to the Annex under resolution TKS-3 from the WG 2 Washington meeting no. 24.	
10.	Chinese member body	
	China is requested to study this possibility of composition to reduce the number of characters of the Yi script in document N 965 that needs coding in the BMP.	
	Request that China and Mongolia get together and come up with an agreed upon joint proposal between the two by the next WG 2 meeting	
	China is requested to coordinate the Tibetan script requirements by working with other member bodies such as India and other countries such as Bhutan and Sikkim using the Tibetan language and script.	
11.	Mr. John Clews	
	Mr. John Clews is invited to draft and propose a proforma incorporating his suggestions for guidelines for submission of new repertoires to be coded.	
12.	Mr. Jürgen Bettels	
	To update document 1031 (as 1031 R) to capture all the missing editorial defect reports and associated editing instructions.	
13.	All member bodies	
a.	National bodies who are interested in attending IRG meetings should contact the new IRG rapporteur Mr. A. Kido.	
b.	N 990 attempts to cross reference all the documents prior to those listed in this meeting report.. Member bodies are invited to verify and enhance document N 990.	
c.	JTC 1 Procedures are being drafted to replace the 1992 procedures to deal with handling of defect reports. These classify the defects as Editorial or Technical in nature. National bodies are requested to follow these procedures and classify their defect reports.	
d.	Member bodies are invited to nominate a WG 2 expert to be the liaison representative from SC 2 WG 2 to SC 29.	
e.	To inform the convener if they are interested in participating on the Mongolian script discussions to be initiated by China and Mongolia.	
f.	WG 2 reconfirms the guidelines, principles and procedures in document N 946 for allocation of new characters and scripts to ISO/IEC 10646-1. All requesters for additional characters in the standard should take note	
g.	Inform the convener if they are interested in participating on the Tibetan script discussions to be initiated between China, India and other countries using the Tibetan script.	
h.	All member bodies are invited to contribute input towards developing an informative Annex to ISO/IEC 10646-1 for consideration at the next SC 2/WG 2 meeting.	
14.	IRG Rapporteur (Mr. Akio Kido)	
a.	IRG is requested to suggest a prioritization of all ideographic characters defined by it.	
b.	Annexes on Unification Rules and Supplemental Repertoires are to be forwarded to WG 2 convener in time for circulation prior to the next WG 2 meeting if they have to be considered by the October WG 2 meeting.	

1.	All national bodies	
a.	to review the documents N 946 and N 947 and submit written contributions to 'refine the starting list of characters under different categories' by 1994-01-31; (reference resolution SP-7 in section 6.1).	N 983 UK N 984 Canada N 1002 Japan
b.	to contribute input on source of characters and cross-reference to other standards; (reference resolution TKS-2).	N 963 China N 964 China N 965 China

		N 966 China N 967 China N 968 Mongolia N 973 Korea N 984 Canada N 1011 China
c.	to contribute input towards developing an informative Annex on Equivalence between pre-composed characters and their corresponding composite sequences; (reference resolution TKS-3).	N 1005 SC 22/WG 20 N 1029 UNICODE N 1032 Denmark
d.	to review all existing WG 2 documents relevant to additional characters, including comments / disposition of comments -- prior to publication of 10646-1; do all the homework needed so that we do not miss any characters from previous submissions - before next WG 2 meeting; (reference item 8.1 of minutes).	N 990 USA
2. Editor:		
a.	to prepare the text for a Normative Annex on UTF-16 by 1993-12-01 based on the items in resolutions UTF16-4 to UTF16-8; (reference resolution UTF16 -2).	N 970R Editor
b.	to prepare an accompanying background paper showing clearly the benefits and disadvantages of the proposed UTF-16, addressing the comments and concerns brought up during the WG 2 meeting no. 24 by 1994-01-05; (reference resolution UTF16-3)	TO DO
c.	to draft a proposed text for a new normative Annex on UTF-8; (reference resolution UTF8-2).	N 993 Editor
d.	to include some text on 'what criteria can be used for defect report on character names' as part of the Set of Principles document (reference section 9-a-i.3).	TO DO
e.	to include the corrections by changing 'LIGATURE' to 'LETTER' to the following non-presentation form of characters in the next CORRIGENDUM: - 00C6 - LATIN CAPITAL LIGATURE AE - 00E6 - LATIN SMALL LIGATURE AE - 01E2 - LATIN CAPITAL LIGATURE AE WITH MACRON - 01E3 - LATIN SMALL LIGATURE AE WITH MACRON - 01FC - LATIN CAPITAL LIGATURE AE WITH ACUTE - 01FD - LATIN SMALL LIGATURE AE WITH ACUTE.	N 994 Editor
f.	to prepare a defect report to correct the names for all the non-presentation form characters that have the term LIGATURE in their names (reference resolution AE-1), excluding those under resolution AE-2; (reference resolution AE-3).	TO DO
f.	to prepare a draft text for an addendum regarding C1 controls with N 917 as input document, for consideration at the next WG 2 meeting; (reference resolution C0C1-2).	N 995 Editor
h.	with assistance from the Japanese member body, to prepare appropriate text for a CORRIGENDUM to address the defects reported in document N 939 under items 1, 3, 4 and 5; (reference resolution DFR-1).	TO DO
3. Convener (Mike Ksar):		
a.	to forward document N 915 to SC 2/WG 3; (reference section 7.1.1).	DONE
b.	to forward the liaison statement in document N 944 to SC 2 for distribution to SC 2 and SC 2 / WG 3 members for their review and comments; (reference resolution CGM-5).	DONE
c.	to forward NP on addenda on UTF-8 contained in N 952R, to SC 2 for SC 2 letter ballot; (reference resolution NP-1).	DONE
d.	to request SC 2 to establish a formal relationship of liaison, Category C, between SC 2/WG 2 and Taipei Computer Association and forward the statements of benefits and responsibilities (reference resolution L-1).	DONE
4. IRG:		
a.	to proceed with drafting a suitable text for a TR on 'User's guide to CJK Unified Ideographs' documented in N 931; (reference resolution TKS-4).	N 979 IRG N 980 IRG
b.	to proceed with preparing a suitable draft text for an informative Annex to ISO/IEC 10646-1 for rules on unification, indexing and establishing relationships amongst the variants of unified CJK ideographs described in N 932 (with the word 'ordering' changed to 'indexing'); (reference resolution TKS-5).	N 979 IRG N 980 IRG
5. Takayuki K. Sato:		
	to forward liaison report N 948 R to SC 22 / WG 20; (reference resolution TKS-1).	DONE
6. Keld Simonsen and Convener:		
a.	Produce revised document N952-R containing a request for sub-division of project. N 952-R will make old N 952 and N 954 obsolete	N 952R Simonsen, Convener
b.	Update document N 957-R containing the WG 2 program of work - capturing all the items that were accepted at this meeting, with dates showing the next milestone.	N 957R Simonsen, Convener
7. US member body & Korea member body:		
	"US needs to have Korea and US, SC 2, and SC 18 experts develop a proposal for a single writing system for Hangul"; (reference section 5, item pending from Seoul meeting).	Character Glyph Model deals with this item. Action Item is dropped.
8. Japan member body:		
	to get further clarification on the contrast between existing and proposed shapes for regarding "ditto mark" (3003x) and "wave dash" (301Cx) reported as defects in document N 939 item 6;	N 1014 Japan

	(reference resolution DFR-2).	
9.	The meeting secretary (V.S. Umamaheswaran)	
	To incorporate all the comments on draft resolutions in N 949 and produce the final text of resolutions.	N949R Secretary/Convener N955 Secretary/Convener

10.3 List of output documents from the meeting

- N 1000R Document Register (N 950 - N 1000); Mike Ksar, Convener; 1994-04-21
- N 1020 Appendix to N 960; Clarification to Liaison from ITU-TSS SG 8 relating to naming rules in 10646; Stefan Fuchs, Israel (personal contribution); 1994.04.18
- N 1021 Applicability of UTF-16 (clarification); Bruce Paterson, Kyongsok Kim, Sato; 1994.04.18
- N 1025R Hebrew Character names in IS 10646-1; Annotation of names of characters; Stefan Fuchs - (Standards Institution of Israel); 1994.04.20
- N 1027 UTF-16 Specification (Re: N 970R and N 976); Bruce Paterson, UK; 1994.04.20
- N 1028 Clarification on the UK Defect Report WG 2 N 989 concerning tables 39 and 40, Michel Suignard, AFNOR, France; Expert Contribution; 1994.04.20
- N 1031 Defect Report Responses; Jürgen Bettels, et al; 1994-04-20
- N 1033 Unconfirmed minutes of WG 2 meeting 25, Antalya, Turkey;
- N 1034 Resolutions of WG 2 meeting 25, 18-22 April 1994, Antalya, Turkey; 1994-04-21

----- END OF DOCUMENT -----