Universal Multiple-Octet Coded Character Set (UCS)

ISO/IEC JTC 1/SC 2 N4625 ISO/IEC JTC 1/SC 2/WG 2 N4954 Date: 2018-06-22

Source: WG 2 meeting 67, SOAS, University of London, London, UK; 2018-06-18/22

Title: Recommendations from WG 2 meeting 67 **Action:** For approval by SC 2 and for information to WG 2

Status: Adopted at meeting 67 of WG 2

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

Experts accredited (or invited) by the national bodies or liaison organizations for Canada, China, Czech Republic, Germany, India, Ireland, IRG, Japan, Korea, SEI - UC Berkeley (Liaison), the Unicode Consortium (Liaison), TCA (Liaison), UK, and the USA, were present when the following recommendations were adopted (see attached attendance list).

Character count prior to additions to 5th edition**136411 911** additions in DAM 1 (end of meeting 66) **1531** additions in PDAM 2.2 (end of meeting 66)
Character count **138853**(total allocated till end of meeting 66)

Recommendation M67.01 (Disposition of ballot comments of DAM-1 to 5th Edition):

WG2 recommends that SC2 accept the disposition of DAM-1 ballot comments in document N<u>4952R</u>. The following name changes are noted:

10F45 SOGDIAN PHONOGRAM SHIN to SOGDIAN INDEPENDENT SHIN

10D1D HANIFI ROHINGYA VOWEL MARK A to HANIFI ROHINGYA VOWEL A

10D1E HANIFI ROHINGYA VOWEL MARK I to HANIFI ROHINGYA VOWEL I

10D1F HANIFI ROHINGYA VOWEL MARK U to HANIFI ROHINGYA VOWEL U

10D20 HANIFI ROHINGYA VOWEL MARK E to HANIFI ROHINGYA VOWEL E

10D21 HANIFI ROHINGYA VOWEL MARK O to HANIFI ROHINGYA VOWEL O

10D23 HANIFI ROHINGYA NA KHANNA to HANIFI ROHINGYA NA KHONNA, and,

10F58 SOGDIAN PUNCTUATION CIRCLES WITH DOTS to SOGDIAN PUNCTUATION TWO CIRCLES WITH DOTS

Count in Amd. 1: 911 Unanimous

Recommendation M67.02 (Progression of Amendment 1 to the 5th edition):

WG2 recommends that its project editor prepare the final text of Amendment 1 to the 5th edition of the standard, which will include the changes arising from recommendation M67.01 above, along with the final disposition of comments (document N4952R), and forward it to the SC2 secretariat for processing as an FDAM ballot. The draft code charts are in document N5005. The target starting date is modified to FDAM 1 2018-08.

Count: 911 additions in Amendment 1
Unanimous

Observation M67.03 (Additions in Amendment 2 to 5th edition since meeting 66):

WG2 notes that document N4922 lists 6622 character additions in Amendment 2, in the amendment text that was prepared and sent for PDAM 2.2 ballot after M66. Based on disposition of PDAM 2.2 ballot comments in document N4940, document N4941 lists 605 character additions in the resulting amendment text sent for PDAM 2.3 ballot.

Count: 605 in Amd. 2 to 5th edition Noted

Recommendation M67.04 (Disposition of ballot comments of PDAM-2.3 to 5th Edition):

WG2 recommends that SC2 accept the disposition of PDAM-2.3 ballot comments in document N<u>4995</u>. The following significant changes are noted:

- a. ELYMAIC block is moved from 10ECO..10EDF to 10FEO..10FFF
- b. Following 4 Wancho tone marks are added:

1E2EC WANCHO TONE TUP

1E2ED WANCHO TONE TUP MANG

1E2EE WANCHO TONE OKOI

- 1E2EF WANCHO TONE OKOI MANG
- c. Name of 2E4F MEDIEVAL CORNISH VERSE SEPARATOR is changed to CORNISH VERSE DIVIDER
- d. Several glyphs are changed (see list in document N4995), and
- e. Added cross reference to A723 and A725 for A7BD LATIN SMALL LETTER GLOTTAL I

Count in Amd.2: 605+4=609
Unanimous

Recommendation M67.05 (Additional changes to Amd.2 to 5th Edition):

WG2 recommends that SC2 accept the following name changes for following Egyptian Hieroglyph control characters (based on document N5001):

13432 EGYPTIAN HIEROGLYPH START AT TOP to EGYPTIAN HIEROGLYPH INSERT AT TOP START
13433 EGYPTIAN HIEROGLYPH START AT BOTTOM to EGYPTIAN HIEROGLYPH INSERT AT BOTTOM START
13434 EGYPTIAN HIEROGLYPH END AT TOP to EGYPTIAN HIEROGLYPH INSERT AT TOP END
13435 EGYPTIAN HIEROGLYPH END AT BOTTOM to EGYPTIAN HIEROGLYPH INSERT AT BOTTOM END

The glyphs for Egyptian Hieroglyph Controls will be changed to accommodate smaller size acronyms within the glyphs.

Count in Amd.2: 609 Unanimous

Recommendation M67.06 (IVD registration):

WG2 recommends that SC2 accept to append to sub-clause 16.6.3 in the standard, a sentence like the following, to explain the purpose of IVD registration:

The purpose of IVD registration is to provide a technical solution to represent ideograph variants that are considered unifiable and should not be encoded in CJK Unified Ideographs.

Unanimous

Recommendation M67.07 (Progression of Amendment 2):

WG2 recommends that its project editor prepares the final text of Amendment 2 to the 5th edition of the standard, which will include the changes arising from recommendations M67.04 to M67.06 above, along with the final disposition of comments (document N4995), and forward it to the SC2 secretariat for processing as DAM-2. The draft code charts are in document N5000. The target starting dates are modified to DAM-2 2018-08, FDAM-2 2019-02.

Count: 609 additions in Amd.2. Unanimous

Recommendation M67.08 (Khitan Small Script):

WG2 accepts the ad hoc report in document N<u>5002</u> on Khitan Small Script (Block 18B00-18CFF in document N<u>4982R</u>). WG2 recommends that SC2 accept the repertoire for future encoding, with the following changes:

- a. The name of 18B00 KHITAN SMALL SCRIPT ITERATION MARK is changed to be algorithmic, to KHITAN SMALL SCRIPT CHARACTER-18B00.
- b. Add a new a script-specific format character, 16FE4 KHITAN SMALL SCRIPT FILLER, in the Ideographic Symbols and Punctuation block.

Count: 471 additions for future encoding.
Unanimous

Recommendation M67.09 (Dives Akuru script):

2018-06-22

WG2 recommends that SC2 accept the proposal to encode the Dives Akuru script in document N4929, in a new block named Dives Akuru 11900..1195F, and populate it with 72 characters with their names, code points and glyphs as shown in document N4982R.

Count for future encoding: 471 +72=543

Unanimous

Recommendation M67.10 (Gongche characters for Kunqu Opera):

WG2 recommends that SC2 accept seven ideographs based on document N<u>4967</u> for encoding in a new block 2A6E0..2A6FF named CJK Unified Ideographs Supplement and populate the first seven code points 2A6E0..2A6E6 with the proposed glyphs and with UTC as source for these ideographs.

Count for future encoding: 543 +7=550 Unanimous

Recommendation M67.11 (Moji-Joho-Kiban-Ideographs collection):

WG2 recommends that SC2 accept adding a new collection #391: MOJI-JOHO-KIBAN-IDEOGRAPHS 2018 to the standard (based on document N<u>4955</u>), with a note that this collection should be used instead of the current collection #390.

SOAS, University of London, London, UK

Recommendation M67.12 (New K6 source KS X 1027-5:2014):

WG2 recommends that SC2 accept adding new K6 source corresponding to KS X 1027-5:2014, and corresponding Ideograph glyphs and source code points for 152 CJK Unified Ideographs as proposed in document $N\underline{4968}$ in the standard.

Unanimous

Recommendation M67.13 (K1-6B6B as source for code point 8C6C):

WG2 recommends that SC2 accept changing the glyph for source K1-6B6B to the correct form as reflected by the other horizontal variants for this ideograph at code point 8C6C in the standard.

Adopted with abstention from some experts from Korea and Japan.

Recommendation M67.14 (New T source):

WG2 recommends that SC2 accept adding 23 new T-sources for Chemical terminology, with corresponding glyphs and code points, as proposed in the table starting on page 3 of document N4974 to the standard.

Unanimous

Recommendation M67.15 (New G-source):

WG2 recommends that SC2 accept 14 additional ideograph glyphs with corresponding GCE and GGFZ sources, as proposed in document N4987, and replace UCI-00942 with source reference G9: 96389B35, in the standard.

Unanimous

Recommendation M67.16 (Corrections to G-source glyphs):

WG2 recommends that SC2 accept the 11 glyph corrections to G-source ideographs in document N4988.

Unanimous

Recommendation M67.17 (Emoji additions)

WG2 recommends that SC2 accept the following 63 (60 from document N4960 and 3 moved from PDAM 2.2) characters for future encoding with their glyphs as shown in document N4982R:

1F6D5 HINDU TEMPLE; 1F6FA AUTO RICKSHAW;

1F7E0 LARGE ORANGE CIRCLE; 1F7E1 LARGE YELLOW CIRCLE; 1F7E2 LARGE GREEN CIRCLE; 1F7E3 LARGE PURPLE CIRCLE;

1F7E4 LARGE BROWN CIRCLE; 1F7E5 LARGE RED SQUARE; 1F7E6 LARGE BLUE SQUARE; 1F7E7 LARGE ORANGE SQUARE;

1F7E8 LARGE YELLOW SQUARE; 1F7E9 LARGE GREEN SQUARE; 1F7EA LARGE PURPLE SQUARE;

1F7EB LARGE BROWN SQUARE; 1F90D WHITE HEART; 1F90E BROWN HEART; 1F90F PINCH; 1F93F DIVING MASK:

1F971 YAWNING FACE; 1F979 TROLL; 1F97B SARI; 1F9A3 MAMMOTH; 1F9A4 DODO; 1F9A5 SLOTH; 1F9A6 OTTER;

1F9A7 ORANGUTAN; 1F9A8 SKUNK; 1F9A9 FLAMINGO; 1F9AA OYSTER; 1F9AE GUIDE DOG;1F9AF PROBING CANE;

1F9BA SERVICE ANIMAL VEST; 1F9BB EAR WITH HEARING AID; 1F9BC MOTORIZED WHEELCHAIR; 1F9BD MANUAL WHEELCHAIR; 1F9BE MECHANICAL ARM; 1F9BF MECHANICAL LEG; 1F9C3 BEVERAGE BOX; 1F9C4 GARLIC; 1F9C5 ONION; 1F9C6 FALAFEL; 1F9C7 WAFFLE; 1F9C8 BUTTER; 1F9C9 MATE DRINK;

1F9CA ICE CUBE; 1F9CD STANDING PERSON; 1F9CE KNEELING PERSON; 1FA70 BALLET SHOES;

1FA71 ONE-PIECE SWIMSUIT; 1FA72 BRIEFS; 1FA73 SHORTS; 1FA78 BLOOD DROP;

1FA79 ADHESIVE BANDAGE; 1FA7A STETHOSCOPE; 1FA80 YO-YO TOY; 1FA81 KITE; 1FA82 PARACHUTE; 1FA90 RINGED PLANET; 1FA91 CHAIR; 1FA92 RAZOR; 1FA93 AXE; 1FA94 DIYA LAMP; 1FA95 BANJO.

Count for future encoding: 550+63=613

Adopted with abstention from expert from UK.

Recommendation M67.18 (CJK Ext-G)

WG2 recommends that SC2 accept encoding of 4938 CJK Unified Ideographs in a new block 30000..3134F named CJK Unified Ideograph Extension G, and populate it with their glyphs and sources as shown in document N4982R.

Count for future encoding: 4938+613=5551

Unanimous

Recommendation M67.19 (Miscellaneous additions)

WG2 recommends that SC2 accept the following additional characters with their glyphs as shown in document N4982R for future encoding in the standard:

- a. 111CE SHARADA VOWEL SIGN PRISHTHAMATRA E
- b. 111CF SHARADA SIGN INVERTED CANDRABINDU
- c. 11460 NEWA SIGN JIHVAMULIYA 11461 NEWA SIGN UPADHMANIYA

- d. 1145A NEWA DOUBLE COMMA
- e. Following six circled characters:

1F10D CIRCLED ZERO WITH SLASH

1F10E CIRCLED COUNTERCLOCKWISE ARROW

1F10F CIRCLED DOLLAR SIGN WITH OVERLAID BACKSLASH

1F16D CIRCLED CC

1F16E CIRCLED C WITH OVERLAID BACKSLASH

1F16F CIRCLED HUMAN FIGURE

- f. 11A48 ZANABAZAR SOUARE CLUSTER-INITIAL LETTER LA
 - 11A49 ZANABAZAR SQUARE CLUSTER-INITIAL LETTER SA
- g. Following seven Tangut Component characters:

18AF3 TANGUT COMPONENT-756

18AF4 TANGUT COMPONENT-757

18AF5 TANGUT COMPONENT-758

18AF6 TANGUT COMPONENT-759

18AF7 TANGUT COMPONENT-760

18AF8 TANGUT COMPONENT-761

18AF9 TANGUT COMPONENT-762

h. 08BE ARABIC LETTER PEH WITH SMALL V

08BF ARABIC LETTER TEH WITH SMALL V

08C0 ARABIC LETTER TTEH WITH SMALL V

08C1 ARABIC LETTER TCHEH WITH SMALL V

08C2 ARABIC LETTER KEHEH WITH SMALL V

- i. 08C3 ARABIC LETTER GHAIN WITH THREE DOTS ABOVE
 - 08C4 ARABIC LETTER AFRICAN QAF WITH THREE DOTS ABOVE
- OD81 SINHALA SIGN CANDRABINDU
- k. 0D04 MALAYALAM LETTER VEDIC ANUSVARA
- 31BB BOPOMOFO FINAL LETTER G
- m. Add following ten characters to Latin Extended-D block to support Medieval Cornish:

A7D0 LATIN SMALL LETTER A WITH OVERCURL

A7D1 LATIN SMALL LETTER E WITH OVERCURL

A7D2 LATIN SMALL LETTER I WITH OVERCURL

A7D3 LATIN SMALL LETTER M WITH OVERCURL

A7D4 LATIN SMALL LETTER N WITH OVERCURL

A7D5 LATIN SMALL LETTER R WITH OVERCURL A7D6 LATIN SMALL LETTER S WITH OVERCURL

A7D7 LATIN SMALL LETTER T WITH OVERCURL

A7D8 LATIN SMALL LETTER U WITH OVERCURL

A7D9 LATIN SMALL LETTER Y WITH OVERCURL

n. Add following **five** counting rod numerals (moved out from PDAM 2.2):

1D379 SOUTHERN SONG COUNTING ROD UNIT DIGIT FOUR

1D37A SOUTHERN SONG COUNTING ROD UNIT DIGIT FIVE

1D37B SOUTHERN SONG COUNTING ROD UNIT DIGIT NINE

1D37C SOUTHERN SONG COUNTING ROD TENS DIGIT FIVE 1D37D SOUTHERN SONG COUNTING ROD TENS DIGIT NINE

o. A7C0 LATIN CAPITAL LETTER THORN WITH DIAGONAL STROKE

A7C1 LATIN SMALL LETTER THORN WITH DIAGONAL STROKE

Count for future encoding: 47 +5551=5598 **Unanimous**

Recommendation M67.20 (next Edition)

- a. WG2 recommends that SC2 create a new project proposal for the next edition of the standard.
- b. WG2 recommends that its project editor prepare the text for the next edition of the standard based on contents till end of Amendment 2 to the 5th edition and additions from resolutions M67.08 to M67.19 above, and, submit the text for a CD ballot – with the following target dates for starting the ballots: CD 2018-09: CD-2 2019-01: DIS 2019-07.

Unanimous

Recommendation M67.21 (Consideration for next edition):

WG2 recommends that the project editor adopt the proposed changes from page 3 in document N<u>4991</u>, for the next edition to the organization and referencing of the various data files.

Unanimous

Recommendation M67.22 (Cypro-Minoan script):

WG2 recommends that SC2 accept the proposal on the Cypro-Minoan script in document N<u>4733</u>, as a potential candidate for future encoding, in a new block named Cypro-Minoan in a new block 12700..1278F, and populate it with 140 characters with their names, code points and glyphs based on document N4733.

Unanimous

Recommendation M67.23 (Code point for Square Japanese New Era Name):

WG2 recommends that SC2 accept the request for future encoding of the new Japanese Era name at code point 32FF, pending availability of its name and glyph.

Unanimous

Recommendation M67.24 (Additions to CD before next WG2 meeting):

WG2 recognizes that some scripts and additional characters which are under preparation as potential additions to the standard could become mature and will have consensus among WG2 experts to include in the standard. The project editor should be able to add these to the balloted texts, after exercising due diligence. Candidates include, but are not limited to:

- Cypro-Minoan script
- New Japanese Era name character

Unanimous

Recommendation M67.25 (Assamese script):

After consideration of the proposal in document N<u>4947</u> to encode the Assamese script, WG2 recommends that SC2 accept the ad hoc report on Assamese script in document N4999, which has the following main recommendations:

- a. Add Assamese character names in the nameslist as annotations,
- b. Change the block header from Bengali to Bengali-Assamese, and,
- c. Prepare a revised contribution on new characters to be added.

WG2 encourages the experts on Assamese script to continue the work towards a revised contribution and submit to WG2. WG2 recommends that SC2 invites the national body of India, BIS, to coordinate this effort.

Unanimous

Recommendation M67.26 (Shuishu script):

WG2 invites the authors of Shuishu script (Block 1B300-1B4FF Shuishu Logograms and Block 1B500-1B52F Shuishu Radicals from document N_{4894R}) that was taken out of PDAM 2.2 based on ballot comments, to revise their proposal, taking into account the feedback received at this meeting (see documents N_{4942} , N_{4946} and N_{4956}), working with other experts interested in this script.

Unanimous

Recommendation M67.27 (Mongolian):

WG2 recommends that SC2 notes the progress made by the Mongolian experts in documents N<u>4989</u> and N<u>4998</u>, and invites the Chinese national body, working with Mongolian experts, to prepare and submit to WG2, a separate document containing only the proposed additions with rationale and phonetic information.

Unanimous

Recommendation M67.28 (Small Seal script):

WG2 recommends that SC2 note the progress made by the experts in document N4973, and encourages the experts to continue the work towards a contribution for encoding the script, taking into account the feedback documents and discussion at this meeting. WG2 notes that another ad hoc meeting of experts is planned by TCA in November 2018 in Taipei (tentatively), and invites all the interested experts to take note towards planning to attend the meeting.

Unanimous

Recommendation M67.29 (Emoji Ad hoc):

WG2 recommends that SC2 take note of the report from WG2 Emoji ad hoc meeting in document N<u>5003</u> on improved interworking between WG2 and the Unicode Consortium.

Unanimous

Recommendation M67.30 (IRG Process):

WG2 recommends that SC2 welcome the input from Japanese national body on improving the working of IRG (in document N4948), and invites the Japanese national body to take into consideration the feedback from the discussion in WG2.

Unanimous

Recommendation M67.31 (Future meetings):

WG2 endorses the following schedule for future meetings:

WG2 Meeting 68 - 2019-06-17/21, Redmond, WA, USA

WG2 Meeting 69 - Need Host, Tentative: 2020-06-15/19; (backup Dundee, Scotland, UK)

IRG #52 Hong Kong SAR, China, 2019-05-13/17

IRG #53 China(Tentative), 2019-10-21/25, (location TBD)

Unanimous

Appreciation M67.32 (Appreciation to DKUUG for web site support):

WG2 thanks DKUUG and its staff for its continued support of the web site for the older WG2 documents.

By Acclamation

Appreciation M67.33 (Appreciation to Unicode Consortium for web site support):

WG2 thanks the Unicode Consortium and its staff for providing the new web site and its support for WG2.

By Acclamation

Appreciation M67.34 (Appreciation to Host):

WG2 thanks the national body of UK and SOAS, University of London, for hosting the meeting. In addition, WG 2 thanks Ms. Gema Zamora Fernandez and staff at the Endangered Languages Archive, Dr. Mandana Seyfeddinipur of Endangered Languages Documentation Programme, SOAS, with additional support from Dr. Nathan Hill and staff at the Department of East Asian Languages and Cultures at SOAS.

By Acclamation

Character count prior to additions to 5th edition**136411**

911 additions in FDAM 1 (end of meeting 67)

609 additions in DAM 2 (end of meeting 67)

5598 additions for CD of next edition (end of meeting 67)

Character count **143529** (total allocated till end of meeting 67)

Attendance List

JTC1/SC2/WG2 Meeting 67, SOAS, University of London, London, UK; 2018-06-16/22

The following **39** experts accredited (or invited) by **10** national bodies and **4** liaison organizations were present at different times during the meeting.

Name	Accreditation	Affiliation
Dhruba Jyoti BORAH	.Invited Expert (India)	Assam Hills Medical College
Mahendra Kumar YADAVA	.Invited Expert (India)	Assam Electronics Development Corporation Ltd.
Paramananda RAJBONGSHI	.Invited Expert (India)	Pragjyotish College
Jeremy BURGE	.Invited Expert (Unicode)	Unicode Emoji Sub Committee
Richard ISHIDA	.Invited Expert (w3c)	W3C Consortium
Qin LU	.IRG rapporteur	Hong Kong Polytechnic University
Shuichi TASHIRO	.SC 2 Chairman	Information-technology Promotion Agency
Toshiko KIMURA	.SC 2 Secretariat	IPSJ/ITSCJ
Chui Hsing CHIU	.TCA (Cat. C Liaison)	Bureau of Standards, Ministry of Economic Affairs
Lin Mei WEI	.TCA (Cat. C Liaison)	Chinese Foundation for Digitization Technology
Michel SUIGNARD	.WG 2 Convenor; USA	Unicode Inc.
Alain LABONTÉ	Canada	Independent
Karljürgen FEUERHERM	Canada	Wilfrid Laurier University
V.S. (Uma) UMAMAHESWARAN	Canada	IBM Canada
Guoying LI	China	Beijing Normal University
Huang DAI	China	CESI
Jianchun LU	China	Weifang Beida Jade Bird Huaguang Imagesetter Co. Ltd.
Jinbao LIANG	China	Inner Mongolia Ethnic Affairs Commission of China
Lili HE	China	China Publishing Group Digital Media Co. Ltd.
Nashunwuritu	China	Inner Mongolia University
Xiaowen ZHOU	China	Beijing Normal University
Yifei ZHANG	China	China Publishing Group Digital Media Co. Ltd.
Yingzhe WU	China	Inner Mongolia University
Zhuang CHEN	China	CESI
Jan KUČERA	Czech Republic	Charles University
Karl PENTZLIN	Germany	AC&S Analysis Consulting & Software GmbH
Krishnankutty MANIKANDAN	India	Bureau of Indian Standards
Shikhar Kumar SARMA	India	Gauhati University
Michael EVERSON	Ireland	Evertype
Takada TOMOKAZU	Japan	National Institute for Japanese Language and Linguistics
Toshiya SUZUKI	Japan	Hiroshima University
Wataru TAKAGI	Japan	Hitachi, Ltd.
Gowoon KIM	Korea	Telecommunications Technology Association
Kyongsok KIM	Korea	Pusan National University
Wang Sung YANG	Korea	Hancom
Andrew WEST	UK	Independent
Martin HOSKEN	UK	SIL International
Deborah ANDERSON	USA, UC Berkeley (Cat. C Liaison)	University of California, Berkeley
Peter CONSTABLE	USA, Unicode (Cat. A Liaison)	Microsoft Corp.