

Universal Multiple-Octet Coded Character Set
International Organization for Standardization
Organisation Internationale de Normalisation
Международная организация по стандартизации

Doc Type: Working Group Document**Title: Proposal for encoding Last Resort Pictures in Plane 14 of the UCS****Source: Michael Everson****Status: Individual Contribution****Action: For consideration by JTC1/SC2/WG2 and UTC****Date: 2008-04-01**

1. Introduction. The Last Resort font is a collection of glyphs which represents types of UCS characters. These glyphs are designed to allow users to recognize that an encoded value is a specific type of UCS character, a Private Use Area character, an unassigned character, or one of the illegal character codes. Apple Computer and SIL are two organizations which have shipped Last Resort fonts for some time now. Recently Apple decided to make its Last Resort font available for public distribution via the Unicode Consortium's website (sometime after the publication of Unicode 5.1).

One of the principles of a Last Resort font is that is normally displayed when no other font is available for the character in question. Similarly, the Control characters encoded at U+0000-001F cannot be represented in text, since they are intended "do" things (even if few of them do much on modern operating systems. At 2400-243F the CONTROL PICTURES block provides glyphs for those characters so that they can be discussed and displayed in text. This facility would be as valuable for the Last Resort functionality as it is for the control characters. Accordingly, this proposal requests the encoding of a new block, U+E0200-E03FF LAST RESORT PICTURES, in Plane 14 of the UCS.

The scheme for encoding in Plane 14 proposed here is block-based. As new blocks are added to the standard, new characters would be added to the Last Resort Pictures block. This maintenance burden is not very great, since the Consortium's Last Resort font has to be updated in any case. In fact maintaining the Plane 14 block could well ensure a more timely updating of the Last Resort font.

2. Character names. The character names proposed begin with the prefix "LR-", followed by the character range and block name.

3. Unicode Character Properties.

```
E0200;LR-0000-FFFF PLANE-0 UNDEFINED;So;0;ON;;;;;N;;;;;
E0201;LR-0000-007F BASIC LATIN;So;0;ON;;;;;N;;;;;
E0202;LR-0080-00FF LATIN-1 SUPPLEMENT;So;0;ON;;;;;N;;;;;
E0203;LR-0100-017F LATIN EXTENDED-A;So;0;ON;;;;;N;;;;;
...
```

4. Issues. The BMP is nearly full, and it is not hard to leave a bit of space where between blocks for future expansion. Whether this should be done for the SMP or not may merit some discussion.

A. Administrative

1. Title

Proposal for encoding Last Resort Pictures in Plane 14 of the UCS

2. Requester's name

Michael Everson

3. Requester type (Member body/Liaison/Individual contribution)

Individual contribution.

4. Submission date

2008-04-01

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

Yes.

6b. More information will be provided later

No.

B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

Yes.

1b. Proposed name of script

Last Resort Pictures.

1c. The proposal is for addition of character(s) to an existing block

No.

1d. Name of the existing block

2. Number of characters in proposal

201.

3. Proposed category (A-Contemporary; B.1-Specialized (small collection); B.2-Specialized (large collection); C-Major extinct; D-Attested extinct; E-Minor extinct; F-Archaic Hieroglyphic or Ideographic; G-Obscure or questionable usage symbols)

Category G.

4a. Is a repertoire including character names provided?

Yes.

4b. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?

Yes.

4c. Are the character shapes attached in a legible form suitable for review?

Yes.

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Michael Everson.

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, Fontographer.

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

7. Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?

Yes.

8. Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database <http://www.unicode.org/Public/UNIDATA/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

See above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

No.

2a. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?

No.

2b. If YES, with whom?

2c. If YES, available relevant documents

3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?

No.

4a. The context of use for the proposed characters (type of use; common or rare)

Rare.

4b. Reference

5a. Are the proposed characters in current use by the user community?

The glyphs are, but they are unavailable for use in text.

5b. If YES, where?

In Last Resort fonts.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

No.

6b. If YES, is a rationale provided?

6c. If YES, reference

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

Yes.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

No.

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?

	E020	E021	E022	E023	E024	E025	E026	E027	E028	E029	E02A	E02B	E02C	E02D	E02E	E02F
0																
1																
2																
3																
4																
5																
6																
7																
8																
9																
A																
B																
C																
D																
E																
F																

Plane 0 Blocks (BMP)































E0200		LR-0000-FFFF PLANE-0 UNDEFINED	E0238		LR-1A00-1A1F BUGINESE
E0201		LR-0000-007F BASIC LATIN	E0239		LR-1B00-1B7F TAI THAM
E0202		LR-0080-00FF LATIN-1 SUPPLEMENT	E023A		<reserved>
E0203		LR-0100-017F LATIN EXTENDED-A	E023B		<reserved>
E0204		LR-0180-024F LATIN EXTENDED-B	E023C		LR-1B00-1B7F BALINESE
E0205		LR-0250-02AF IPA EXTENSIONS	E023D		LR-1B80-1BBF SUNDAHESE
E0206		LR-02B0-02FF SPACING MODIFIER LETTERS	E023E		<reserved>
E0207		LR-0300-036F COMBINING DIACRITICAL MARKS	E023F		LR-1C00-1C4F LEPCHA
E0208		LR-0370-03FF GREEK AND COPTIC	E0240		LR-1C50-1C7F OL CHIKI
E0209		LR-0400-04FF CYRILLIC	E0241		LR-1C80-1CDF MEITEI MAYEK
E020A		LR-0500-052F CYRILLIC SUPPLEMENT	E0242		LR-1CE0-1CFF VEDIC EXTENSIONS
E020B		LR-0530-058F ARMENIAN	E0243		LR-1D00-1D7F PHONETIC EXTENSIONS
E020C		LR-0590-05FF HEBREW	E0244		LR-1D80-1DBF PHONETIC EXTENSIONS SUPPLEMENT
E020D		LR-0600-06FF ARABIC	E0245		LR-1DC0-1DFF COMBINING DIACRITICAL MARKS SUPPLEMENT
E020E		LR-0700-074F SYRIAC	E0246		LR-1E00-1EFF LATIN EXTENDED ADDITIONAL
E020F		LR-0750-077F ARABIC SUPPLEMENT	E0247		LR-1F00-1FFF GREEK EXTENDED
E0210		LR-0780-07BF THAANA	E0248		LR-2000-206F GENERAL PUNCTUATION
E0211		LR-07C0-07FF NKO	E0249		LR-2070-209F SUPERSCRIPITS AND SUBSCRIPITS
E0212		LR-0800-083F SAMARITAN	E024A		LR-20A0-20CF CURRENCY SYMBOLS
E0213		<reserved>	E024B		LR-20D0-20FF COMBINING DIACRITICAL MARKS FOR SYMBOLS
E0214		<reserved>	E024C		LR-2100-214F LETTERLIKE SYMBOLS
E0215		<reserved>	E024D		LR-2150-218F NUMBER FORMS
E0216		<reserved>	E024E		LR-2190-21FF ARROWS
E0217		LR-0900-097F DEVANAGARI	E024F		LR-2200-22FF MATHEMATICAL OPERATORS
E0218		LR-0980-09FF BENGALI	E0250		LR-2300-23FF MISCELLANEOUS TECHNICAL
E0219		LR-0A00-0A7F GURMUKHI	E0251		LR-2400-243F CONTROL PICTURES
E021A		LR-0A80-0AFF GUJARATI	E0252		LR-2440-245F OPTICAL CHARACTER RECOGNITION
E021B		LR-0B00-0B7F ORIYA	E0253		LR-2460-24FF ENCLOSED ALPHANUMERICS
E021C		LR-0B80-0BFF TAMIL	E0254		LR-2500-257F BOX DRAWING
E021D		LR-0C00-0C7F TELUGU	E0255		LR-2580-259F BLOCK ELEMENTS
E021E		LR-0C80-0CFF KANNADA	E0256		LR-25A0-25FF GEOMETRIC SHAPES
E021F		LR-0D00-0D7F MALAYALAM	E0257		LR-2600-26FF MISCELLANEOUS SYMBOLS
E0220		LR-0D80-0DFF SINHALA	E0258		LR-2700-27BF DINGBATS
E0221		LR-0E00-0E7F THAI	E0259		LR-27C0-27EF MISCELLANEOUS MATHEMATICAL SYMBOLS-A
E0222		LR-0E80-0EFF LAO	E025A		LR-27F0-27FF SUPPLEMENTAL ARROWS-A
E0223		LR-0F00-0FFF TIBETAN	E025B		LR-2800-28FF BRAILLE PATTERNS
E0224		LR-1000-109F MYANMAR	E025C		LR-2900-297F SUPPLEMENTAL ARROWS-B
E0225		LR-10A0-10FF GEORGIAN	E025D		LR-2980-29FF MISCELLANEOUS MATHEMATICAL SYMBOLS-B
E0226		LR-1100-11FF HANGUL JAMO	E025E		LR-2A00-2AFF SUPPLEMENTAL MATHEMATICAL OPERATORS
E0227		LR-1200-137F ETHIOPIC	E025F		LR-2B00-2BFF MISCELLANEOUS SYMBOLS AND ARROWS
E0228		LR-1380-139F ETHIOPIC SUPPLEMENT	E0260		LR-2C00-2C5F GLAGOLITIC
E0229		LR-13A0-13FF CHEROKEE	E0261		LR-2C60-2C7F LATIN EXTENDED-C
E022A		LR-1400-167F UNIFIED CANADIAN ABORIGINAL SYLLABICS	E0262		LR-2C80-2CFF COPTIC
E022B		LR-1680-169F OGHAM	E0263		LR-2D00-2D2F GEORGIAN SUPPLEMENT
E022C		LR-16A0-16FF RUNIC	E0264		LR-2D30-2D7F TIFINAGH
E022D		LR-1700-171F TAGALOG	E0265		LR-2D80-2DDF ETHIOPIC EXTENDED
E022E		LR-1720-173F HANUNOO	E0266		LR-2DE0-2DFF CYRILLIC EXTENDED-A
E022F		LR-1740-175F BUHID	E0267		LR-2E00-2E7F SUPPLEMENTAL PUNCTUATION
E0230		LR-1760-177F TAGBANWA			
E0231		LR-1780-17FF KHMER			
E0232		LR-1800-18AF MONGOLIAN			
E0233		<reserved>			
E0234		LR-1900-194F LIMBU			
E0235		LR-1950-197F TAI LE			
E0236		LR-1980-19DF NEW TAI LUE			
E0237		LR-19E0-19FF KHMER SYMBOLS			

E0268 LR-2E80-2EFF CJK RADICALS SUPPLEMENT
 E0269 LR-2F00-2FDF KANGXI RADICALS
 E026A <reserved>
 E026B LR-2FF0-2FFF IDEOGRAPHIC DESCRIPTION CHARACTERS
 E026C LR-3000-303F CJK SYMBOLS AND PUNCTUATION
 E026D LR-3040-309F HIRAGANA
 E026E LR-30A0-30FF KATAKANA
 E026F LR-3100-312F BOPOMOFO
 E0270 LR-3130-318F HANGUL COMPATIBILITY JAMO
 E0271 LR-3190-319F KANBUN
 E0272 LR-31A0-31BF BOPOMOFO EXTENDED
 E0273 LR-31C0-31EF CJK STROKES
 E0274 LR-31F0-31FF KATAKANA PHONETIC EXTENSIONS
 E0275 LR-3200-32FF ENCLOSED CJK LETTERS AND MONTHS
 E0276 LR-3300-33FF CJK COMPATIBILITY
 E0277 LR-3400-4DBF CJK UNIFIED IDEOGRAPHS EXTENSION A
 E0278 LR-4DC0-4DFF YIJING HEXAGRAM SYMBOLS
 E0279 LR-4E00-9FFF CJK UNIFIED IDEOGRAPHS
 E027A LR-A000-A48F YI SYLLABLES
 E027B LR-A490-A4CF YI RADICALS
 E027C LR-A4D0-A4FF LISU
 E027D LR-A500-A63F VAI
 E027E LR-A640-A69F CYRILLIC EXTENDED-B
 E027F <reserved>
 E0280 LR-A700-A71F MODIFIER TONE LETTERS
 E0281 LR-A720-A7FF LATIN EXTENDED-D
 E0282 LR-A800-A82F SYLOTI NAGRI
 E0283 LR-A830-A83F COMMON INDIC NUMBER FORMS
 E0284 LR-A840-A87F PHAGS-PA
 E0285 LR-A880-A8DF SAURASHTRA
 E0286 LR-A8E0-A8FF DEVANAGARI EXTENSIONS
 E0287 LR-A900-A92F KAYAH LI
 E0288 LR-A930-A95F REJANG
 E0289 LR-A960-A97F HANGUL JAMO EXTENDED-A
 E028A LR-A980-A9DF JAVANESE
 E028B <reserved>
 E028C LR-AA00-AA5F CHAM
 E028D <reserved>
 E028E LR-AA80-AADF TAI VIET
 E028F <reserved>
 E0290 <reserved>
 E0291 <reserved>
 E0292 <reserved>
 E0293 <reserved>
 E0294 <reserved>
 E0295 LR-AC00-D7AF HANGUL SYLLABLES
 E0296 LR-D7B0-D7FF HANGUL JAMO EXTENDED-B
 E0297 LR-D800-DB7F HIGH SURROGATES
 E0298 LR-DB80-DBFF HIGH PRIVATE USE SURROGATES






E0299 LR-DC00-DFFF LOW SURROGATES
 E029A LR-E000-F8FF PRIVATE USE AREA
 E029B LR-F900-FAFF CJK COMPATIBILITY IDEOGRAPHS
 E029C LR-FB00-FB4F ALPHABETIC PRESENTATION FORMS
 E029D LR-FB50-FDFF ARABIC PRESENTATION FORMS-A
 • fb50-fdcf, fdf0-fdff
 E029E LR-FDD0-FDEF NONCHARACTERS
 E029F LR-FE00-FE0F VARIATION SELECTORS
 E02A0 LR-FE10-FE1F VERTICAL FORMS
 E02A1 LR-FE20-FE2F COMBINING HALF MARKS
 E02A2 LR-FE30-FE4F CJK COMPATIBILITY FORMS
 E02A3 LR-FE50-FE6F SMALL FORM VARIANTS
 E02A4 LR-FE70-FEFF ARABIC PRESENTATION FORMS-B
 E02A5 LR-FF00-FFEF HALFWIDTH AND FULLWIDTH FORMS
 E02A6 LR-FFF0-FFFF SPECIALS
 E02A7 LR-FFFE-FFFF NONCHARACTERS

Plane 1 Blocks (SMP)






E02A8 LR-0000-FFFF PLANE-1 UNDEFINED
 E02A9 LR-10000-1007F LINEAR B SYLLABARY
 E02AA LR-10080-100FF LINEAR B IDEOGRAMS
 E02AB LR-10100-1013F AEGEAN NUMBERS
 E02AC LR-10140-1018F ANCIENT GREEK NUMBERS
 E02AD LR-10190-101CF ANCIENT SYMBOLS
 E02AE LR-101D0-101FF PHAISTOS DISC
 E02AF <reserved>
 E02B0 <reserved>
 E02B1 LR-10280-1029F LYCIAN
 E02B2 LR-102A0-102DF CARIAN
 E02B3 <reserved>
 E02B4 LR-10300-1032F OLD ITALIC
 E02B5 LR-10330-1034F GOTHIC
 E02B6 <reserved>
 E02B7 LR-10380-1039F UGARITIC
 E02B8 LR-103A0-103DF OLD PERSIAN
 E02B9 <reserved>
 E02BA LR-10400-1044F DESERET
 E02BB LR-10450-1047F SHAVIAN
 E02BC LR-10480-104AF OSMANYA
 E02BD <reserved>
 E02BE <reserved>
 E02BF <reserved>
 E02C0 LR-10800-1083F CYPRIOT SYLLABARY
 E02C1 LR-10840-1085F IMPERIAL ARAMAIC
 E02C2 <reserved>
 E02C3 <reserved>
 E02C4 <reserved>
 E02C5 <reserved>
 E02C6 <reserved>
 E02C7 LR-10900-1091F PHOENICIAN
 E02C8 LR-10920-1093F LYDIAN
 E02C9 <reserved>
 E02CA <reserved>
 E02CB <reserved>
 E02CC LR-10A00-10A5F KHAROSHTHI

E02CD		<reserved>
E02CE		LR-10A80-10A9F OLD SOUTH ARABIAN
E02CF		<reserved>
E02D0		LR-10B00-10B3F AVESTAN
E02D1		LR-10B40-10B5F INSCRIPTIONAL PARTHIAN
E02D2		LR-10B60-10B7F INSCRIPTIONAL PAHLAVI
E02D3		LR-10B80-10B9F PSALTER PAHLAVI
E02D4		<reserved>
E02D5		<reserved>
E02D6		<reserved>
E02D7		<reserved>
E02D8		<reserved>
E02D9		<reserved>
E02DA		<reserved>
E02DB		<reserved>
E02DC		<reserved>
E02DD		<reserved>
E02DE		<reserved>
E02DF		<reserved>
E02E0		LR-12000-123FF CUNEIFORM
E02E1		LR-12400-1247F CUNEIFORM NUMBERS AND PUNCTUATION
E02E2		LR-1D000-1D0FF BYZANTINE MUSICAL SYMBOLS
E02E3		LR-1D100-1D1FF MUSICAL SYMBOLS
E02E4		LR-1D200-1D24F ANCIENT GREEK MUSICAL NOTATION
E02E5		LR-1D300-1D35F TAI XUAN JING SYMBOLS
E02E6		LR-1D360-1D37F COUNTING ROD NUMERALS
E02E7		LR-1D400-1D7FF MATHEMATICAL ALPHANUMERIC SYMBOLS
E02E8		LR-1F000-1F02F MAHJONG TILES
E02E9		LR-1F030-1F09F DOMINO TILES
E02EA		LR-1FFFE-1FFFF NONCHARACTERS



Plane 2 Blocks (SIP)

E02EB		LR-20000-2FFFD PLANE-2 UNDEFINED
E02EC		LR-20000-2A6DF CJK UNIFIED IDEOGRAPHS EXTENSION B
E02ED		LR-2A700-2B77F CJK UNIFIED IDEOGRAPHS EXTENSION C
E02EE		LR-2F800-2FA1F CJK COMPATIBILITY IDEOGRAPHS SUPPLEMENT
E02EF		LR-2FFFE-2FFFF NONCHARACTERS



Plane 14 Blocks (SSP)

E02F0		LR-E0000-EFFFD PLANE-14 UNDEFINED
E02F1		LR-E0000-E007F TAGS
E02F2		LR-E0100-E01EF VARIATION SELECTORS SUPPLEMENT
E02F3		LR-E0200-E03FF LAST RESORT PICTURES
E02F4		LR-EFFFE-EFFFF NONCHARACTERS

Plane 15 Blocks (SPA)

E02F5		LR-F0000-FFFFF SUPPLEMENTARY PRIVATE USE AREA-A
E02F6		LR-FFFFE-FFFFF NONCHARACTERS

Plane 16 Blocks (SPB)

E02F7		LR-100000-10FFFF SUPPLEMENTARY PRIVATE USE AREA-B
E02F8		LR-10FFFE-10FFFF NONCHARACTERS