

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

Doc Type: Working Group Document
Title: Defect Report on MATHEMATICAL SCRIPT SMALL L
Source: Ireland
Status: National Body Contribution
Date: 2002-05-15
Distribution: WG2 and UTC

Defect

When the Mathematical Alphanumeric characters were encoded in ISO/IEC 10646-2:2001, the character U+2113 was incorrectly unified with MATHEMATICAL SCRIPT SMALL L; therefore U+1D4C1 MATHEMATICAL SCRIPT SMALL L is missing.

Background

The character at U+2113, called SCRIPT SMALL L in ISO/IEC 10646, is used as a variant of the italic SMALL LETTER *l* in order to distinguish it from the DIGIT 1 and other characters of similar appearance. It is a well-known mathematical symbol with the name *ell*, and is found, with or without this name, in mathematical documents, directories of mathematical symbols and early character sets.

There have been claims that U+2113 comes into the standard from the ANSI/NISO Z39.47 standard, “Extended Latin Alphabetic Character Set for Bibliographic Use,” also known as “ANSEL”, the Extended Latin character set in the MARC 21 exchange formats published by the Library of Congress. It was used as the abbreviation for “leaf/leaves” until 1981, when revised cataloging rules specified the use of the words “leaf” or “leaves”. The script form was used to avoid confusion with the DIGIT 1. However, the direct sources for U+2113 in the UCS were actually two others:






XCCS 2.0 (1990): 357B/151B “Liter”
IBM CDRA GCGID SM160000: “Liter Symbol”

Both of these sources have the *looped l* glyph ℓ shown continuously in ISO/IEC 10646-1 and the Unicode Standard since Version 1.0.

Although not officially sanctioned by the SI (which simply recommends the use of U+0061 LATIN SMALL LETTER L), it has been common practice to use the *looped l* shown in U+2113 to indicate *litre*; many documents exist with this usage, and the meaning “litre” is commonly given to this symbol in listings of symbols and their usage. While the shape of U+2113 matches that existing usage, it does not readily match the lower case *l* in common styles of script fonts used for mathematical purposes.

When the contributing editors for ISO/IEC 10646 collected the glyphs for the mathematical alphanumeric symbols, they found that none of the script fonts investigated had a lower-case *script l* which looked like the symbol encoded at U+2113. In fact, the editors had to “cheat” and redesign the font so that U+2113 would fit the existing character while all the other “SCRIPT” letters were taken unmodified from the same script font. From these facts, it appears questionable to identify U+2113 with the SCRIPT SMALL L in the sense used in the mathematical alphanumeric symbols.

A comparison of the characters in the standard and their representative glyphs illustrates the issue further:

Code	Character	Glyph	Notes
U+006C	LATIN SMALL LETTER L		Upright l (as shown at U+006C currently), typically standing on a sturdy serif, and with a top serif, but easily confused in shape with U+0031 “1”.
U+2113	SCRIPT SMALL L		Looped l (as shown at U+2113 currently), leaning to the right, but based on cursive handwriting looped form, and unlike the formal script shape.
U+1D459	MATHEMATICAL ITALIC SMALL L		Italic l (as shown at U+1D459 currently), leaning to the right, typically with a curled under base, and with a short top serif.
U+1D4F5	MATHEMATICAL BOLD SCRIPT SMALL L		Script l (as shown at U+1D4F5 currently, in bold form), leaning to the right, and shaped similarly to an italic l, but with a long handwriting-based uptake stroke, instead of a top serif.
U+1D4C1	MATHEMATICAL SCRIPT SMALL L		Missing: This should be the non-bold version of U+1D4F5

U+006C, as the basic UCS character for the Latin small letter l, obviously can be given any style whatsoever, so it can appear with *any* of these glyphs, depending on markup and font.

It is quite clear that the *looped l* required for the U+2113 letterlike symbol is not actually the mathematical script form required for the two sets of mathematical script alphanumeric symbols (plain and bold). Therefore the gap at U+1D4C1 is a mistake, and we should add U+1D4C1 MATHEMATICAL SCRIPT SMALL L.

That leaves us with the problem of what is U+2113 SCRIPT SMALL L, other than the recognition that it isn’t actually a “MATHEMATICAL SCRIPT” form. How do we reconcile the widespread use of this symbol as “litre” with the ANSEL admonition:

In bibliographic work, the script l, *ℓ*, is commonly used as the abbreviation for the term “leaves”. It shall not be used as a symbol for the unit of measure “liter.”

The answer is fairly simple; ANSEL discourages the use of *ℓ* as a symbol for litre in bibliographic records, preferring the official SI recommendation, regardless of whether *ℓ* is used as a symbol for litre by others in other contexts.

Recommendations

Based on the discussion presented we conclude with these recommendations:

1. MATHEMATICAL SCRIPT SMALL L is missing from the standard, and should be encoded at U+1D4C1. (This should be fixed quickly, by the way, since it has an impact on MathML. We recommend adding it to the current PDAM.)
2. U+2113 SCRIPT SMALL L is not really a misnomer – there is a standards precedent for referring to the form in question as a *script l*, even if the glyph required for it is not that appropriate for the *mathematical script* form.

3. Current mappings and interpretations of U+2113 are otherwise correct. It is encoded for legacy interchange with the XCCS and IBM SM160000 litre symbol characters. It is also encoded for legacy interchange with the ANSEL leaf/leaves symbol character.
4. U+2113 should be further annotated in the Unicode Standard, to make it clear that it should not be confused with the MATHEMATICAL SCRIPT SMALL L (once that is encoded), and to make it clear that it is also the obsolete bibliographic leaf/leaves symbol.
5. If people want to abbreviate “litre” *according to SI recommendations*, they should just use U+006C. Otherwise they should use U+2113.