

**Title: Follow-up on Action Item 120-A77**

Source: Laurențiu Iancu, Microsoft Corporation

Date: 2009-10-21

Pertaining to proposal [L2/09-273](#) “Proposal to Encode Two Mathematical Symbols,” this document addresses the following Action Item:

[UTC 120 / L2 217](#) Joint Meeting – Redmond, WA – August 10–14, 2009

D.2.2. Proposal to Encode Two Mathematical Symbols [Iancu, [L2/09-273](#) and [WG2/N3677](#)]

[[120-A77](#)] Action Item for Laurențiu Iancu: Follow up with Barbara Beeton re any existing mappings for U+27CE and U+27CF.

## 1. Synopsis

Document [L2/09-273](#) proposed the encoding of two mathematical symbols, U+27CE SQUARED LOGICAL AND and U+27CF SQUARED LOGICAL OR. The Unicode Technical Committee accepted the two symbols for encoding in a future version of the standard (Consensus [120-C18](#)) and assigned Action Item [120-A77](#) to confirm that there are no already encoded Unicode characters being used in existing mappings to represent the proposed symbols. This document examines the presence or absence of the proposed symbols and other similar symbols in existing entity sets and their mappings to Unicode code points, in particular to APL symbols (U+2336–U+237A). The analysis shows that there are no Unicode characters mapped in existing entity sets, which may be used to represent the proposed symbols.

The proposed symbols and other similar-looking symbols are not listed or mapped in any of the entity sets and symbol lists enumerated in the next section. Section 3 analyzes the character collection of the STIX Fonts Project [1]. That collection includes one symbol whose glyph consists of a caret enclosed within a rectangle, which had originally been mapped to U+2353 APL FUNCTIONAL SYMBOL QUAD UP CARET. However, besides being inadequate for the proposed SQUARED LOGICAL AND, recent research has shown that the original mapping had been misattributed and is being rescinded at the time of writing this document. As a result, there are no mappings to Unicode characters (including APL symbols) to use for representing the proposed symbols, thus leaving the proposed symbols eligible for encoding.

## 2. Entity sets not mapping the proposed symbols

The following entity sets and symbol lists do not contain any mappings relevant for the proposed symbols:

- SGML Entity Set ISO/IEC TR 9573-13:1991
  - The most recent version is still an editor's draft available at <http://www.dcarlisle.demon.co.uk/9573/>;
  - Includes the ISO\* public entity sets such as the Added Math Symbols entity sets ISOAMSA (Arrow Relations), ISOAMSB (Binary Operators), ISOAMSC (Delimiters), etc.
- XML Entity Definitions for Characters
  - Latest draft (currently of October 2009) is available at <http://www.w3.org/2003/entities/2007doc/>, with pointers to previous versions;
  - Covers a superset of the ISO\* public entity sets listed in the SGML source above;
  - A consolidated alphabetic list is available at [http://www.w3.org/2003/entities/2007/w3centities-f.ent](http://www.w3.org/2003/entities/2007/w3centities-f.ent;);
- MathML
  - The latest version, <http://www.w3.org/TR/MathML3/>, points to the XML Entity Definitions for Characters, listed above, and includes an Operator Dictionary in Appendix C at <http://www.w3.org/TR/MathML3/appendixc.html>;
  - Entity sets in previous versions of MathML are available at <http://www.w3.org/TR/MathML2/chapter6.html> and <http://www.w3.org/1999/07/REC-MathML-19990707/chapter6.html>;
  - See also <http://www.w3.org/Math/characters/unicode.xml>;
- The Comprehensive LaTeX Symbol List
  - A list of almost 5000 symbols with corresponding LaTeX entities from several LaTeX packages, updated regularly and available online at <http://www.ctan.org/tex-archive/info/symbols/comprehensive/>;
  - Includes several symbols enclosed in squares (with LaTeX names containing 'box' and 'square'); a few APL symbols (in the wasysym package); and a few circled logical-and ('wedge') and logical-or ('vee') symbols, but not squared;
- Elsevier Science Grid in Unicode
  - Available at <http://info.sciencedirect.com/techsupport/xmlsgml/dtd50/esgrid.pdf>;
- OASIS DocBook
  - An XML mark-up language for technical documentation;
  - Character entities are available at <http://www.oasis-open.org/docbook/specs/wd-docbook-xmlcharent-0.3.pdf> and <http://www.oasis-open.org/docbook/specs/wd-docbook-xmlcharent-0.3.html>;
- Adobe Systems Inc. Mathematical Pi Fonts One–Six
  - A set of character charts relayed courtesy of Barbara Beeton (American Mathematical Society).

### 3. The STIX collection

The STIX Fonts Project [1] maintains a consolidated list of symbols originating from several sources submitted by different organizations. A table of symbols with mappings to Unicode and other entities, together with a layout description, are linked from <http://www.ams.org/STIX/>. An update of the table and its description are currently in preparation. Additional characters are being researched for future encoding in a separate proposal [2].

Of all of the lists contributed to the STIX Fonts Project, only two contain characters similar to those proposed. These two lists, submitted by Design Science and Wolfram Research, Inc., are examined in further detail below. The remaining lists neither contain nor map the proposed characters or any similar-looking characters. Scans of the original documents submitted to the STIX Fonts Project are available courtesy of Barbara Beeton (AMS).

- American Chemical Society (ACS)
- American Institute of Physics (AIP)
  - Includes the symbols in Appendix F “Special Symbols Available for Typesetting” of the AIP Style Manual, available online at <http://www.aip.org/pubservs/style/4thed/toc.html>.
- American Mathematical Society (AMS)
- American Physical Society (APS)
- Design Science
  - The list from Design Science contains a symbol that was originally mapped to U+2353;
- Elsevier
- Institute of Electrical and Electronics Engineers (IEEE)
- Kluwer Academic Publishers
- Springer
- Wolfram Research, Inc.
  - The list from Wolfram Research, Inc. contains the same symbol as the one from Design Science, but mapped to U+F767 in the PUA.

The list of symbols contributed by Design Science to the STIX Fonts Project comprises over 1000 characters with mappings to Unicode code points, including some 230 symbols with mappings in the PUA. This collection itself consolidates several lists of characters drawn from various fonts, enumerated at [3] on Design Science’s Web site, including a list of symbols from Wolfram Research, Inc., which was also submitted separately to the STIX Fonts Project.

The Design Science list includes an enclosed up-pointing caret symbol that was originally mapped to U+2353 APL FUNCTIONAL SYMBOL QUAD UP CARET as shown in Figure 1. There is no down-pointing counterpart symbol on the list.

0x232A	>	Right-pointing angle bracket	Symbol	Alt+0241
0x2336	⊥	Apl functional symbol i-beam	DatapageMath1	Alt+0245
0x2353	⊞	Boxed up caret	Mathem3	Alt+0221
<u>0x2370</u>	⊞	Boxed question mark	Mathem3	Alt+0220
<u>0x2423</u>	⋄	Space indicator	Mathem1	Alt+0136

Figure 1: Excerpt from p. 17 of the original printed submission from Design Science to the STIX Fonts Project showing the mapping of an enclosed caret symbol to U+2353 (courtesy of Barbara Beeton, AMS).

The glyph of the “boxed up caret” differs from the proposed SQUARED LOGICAL OR. More importantly, the “boxed up caret” is semantically unrelated to the proposed symbols: it appears in the Design Science list without a down-caret counterpart, whereas the proposed box-max and box-min are dual operators and cannot exist without one another. The presence of the Design Science “boxed up caret” alone on the list implies a different semantics, unrelated to that of the proposed characters. Indeed, the Design Science symbol was recently found to be the same as the Wolfram Research, Inc. “error indicator” symbol described next, which is not mapped to U+2353, but assigned a code point in the PUA.

Wolfram Research, Inc. contributed to the STIX Fonts Project two lists comprising together over 700 letters and symbols with names used in the Mathematica software and mappings to Unicode code points. Private symbols are mapped to Unicode code points in the PUA. In particular, the symbol “[ErrorIndicator]” illustrated in Figure 2 is mapped to U+F767 instead of U+2353 APL FUNCTIONAL SYMBOL QUAD UP CARET. The symbol can viewed online at [4] and [5]; the latter is reproduced in Figure 3. No other relevant symbols or mappings are present in this collection, either.

Barbara Beeton (AMS) concurred with the above analysis and acknowledged that the original mapping of the Design Science symbol to U+2353 solely based on appearance may have been misattributed. As a result, that mapping in the consolidated STIX collection is being rescinded at the time of writing this proposal, leaving no mappings to Unicode APL symbols that may be used for the proposed symbols. Also, the proposed symbols are currently being added to the STIX repertoire with the code points suggested in [L2/09-273](#).

F765	InvisibleComma		,
F766	ReturnKey	RET	ret
F767	ErrorIndicator	⊞	
F768	AliasIndicator	⋮	esc
F769	EscapeKey	ESC	esc

Figure 2: Excerpt from p. 7 of the original printed submission of private symbols from Wolfram Research, Inc. to the STIX Fonts Project showing the same enclosed caret symbol as in Figure 1 but mapped to U+F767 in the PUA rather than the APL symbol U+2353 (courtesy of Barbara Beeton, AMS).

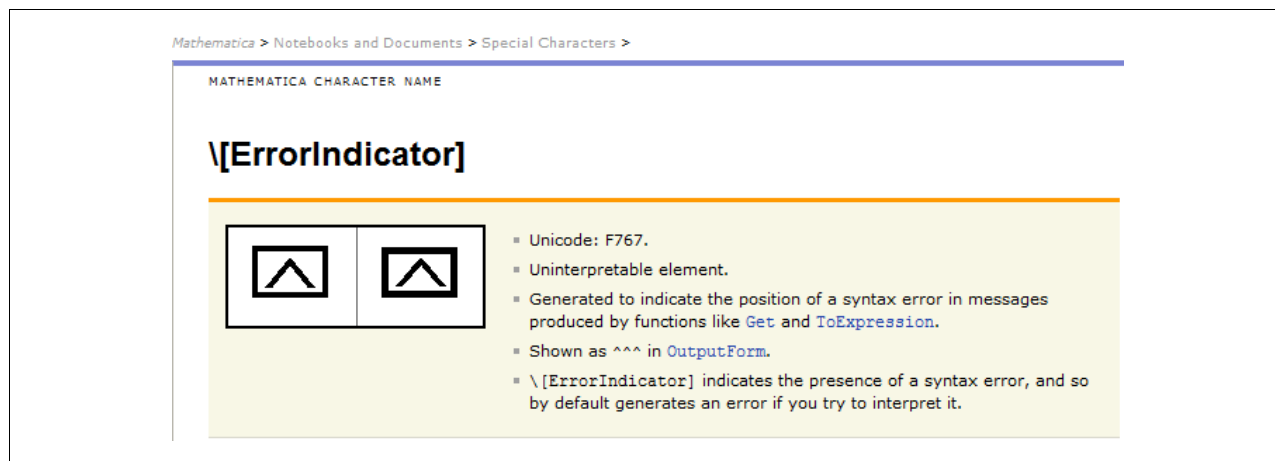


Figure 3: The same symbol from Wolfram Research, Inc. and its mapping to U+F767 online at <http://reference.wolfram.com/mathematica/ref/character/ErrorIndicator.html> (retrieved Oct. 2009).

#### 4. Acknowledgements

The author wishes to express his sincere gratitude to Ms. Barbara Beeton of the American Mathematical Society for her substantial help collecting the information summarized in this document, and to acknowledge Mr. Peter Constable of Microsoft Corporation for his excellent guidance in revising the document.

#### 5. References

- [1] The Scientific and Technical Information Exchange (STIX) Fonts Project, <http://www.stixfonts.org/>, an activity of the STI Pub companies listed at <http://www.stixfonts.org/stipubs.html>.
- [2] D. Anderson et al., "Preliminary Proposal to Encode Characters from the STIX PUA Collection," Unicode Technical Committee and INCITS/L2 Technical Committee documents L2/09-261 & L2/09-262.
- [3] Design Science, "Font Encoding Tables" and "MTCode Encoding Tables," [http://www.dessci.com/en/support/mathtype/tech/encodings/font\\_enc.htm](http://www.dessci.com/en/support/mathtype/tech/encodings/font_enc.htm) and <http://www.dessci.com/en/support/mathtype/tech/encodings/mtcode.htm>.
- [4] Wolfram Research, Inc., Mathematica Documentation, "Listing of Named Characters," <http://reference.wolfram.com/mathematica/guide/ListingOfNamedCharacters.html>. Complete character set linked from <http://www.mathmlcentral.com/characters/>.
- [5] Wolfram Research, Inc., Mathematica Documentation, "\[ErrorIndicator]" private symbol, <http://reference.wolfram.com/mathematica/ref/character/ErrorIndicator.html>.