Doc Type: **Working Group Document**
Title: **New Title**
Source: **Ad-hoc on Mathematical Symbols**
Status: **Ad-hoc report**
Action: **For consideration by JTC1/SC2/WG2 and UTC**
Related: **N2336, N2316, N2343, N2345**

The ad-hoc on mathematical symbols consisted of the following experts:
Christian Cooke, Ireland
Michael Everson, Ireland
Asmus Freytag, Unicode
Prof. Gim, Korea
Hideki Hiura, USA
Shun Ishizaki, Japan
Dr. Kent Karlsson, Sweden
Tatsuo Kobayashi, Japan
Takayuki Sato, Japan
Michel Suignard, USA
Dr. Ken Whistler, USA

Agenda
1. Unification of CJK and math brackets
2. Progressing Math Symbols
3. Roadmap adjustments

The ad-hoc on mathematical symbols met on 2001-04-02.

1. Unification of CJK and math brackets
The status quo is that the Unicode Standard asserts that 2329 and 232A are canonically equivalent with 3008 and 3009 respectively. ISO/IEC 10646 has no such concept. In the second edition of the standard, the two sets are imaged with a different set of glyphs, representing typical appearance (including character width and distribution of white space) for the mathematical or technical usage and CJK usage respectively.

In the context of the normalization forms defined by the Unicode Standard, and their expected wide use on the web, it will become practically impossible to maintain such a distinction based on character code as the normalization will always map to 3008/3009.

In addition, the Unicode Consortium has taken the position that other brackets used in mathematics should be unified with their counterparts in the 3000 block, and rejected encoding additional bracket characters for mathematical use. Note: The parentheses already exist in two forms in ISO/IEC 10646 as 0028/0029 and FE08/FE09. These are not canonically equivalent to each other in Unicode.

Michel Suignard, Tatsuo Kobayashi and other experts in East Asian typography explained the way text layout systems for EA context are dealing with these braces. In short, the software assigns a character property to each punctuation character, and applies kerning based on these character properties. The EA brackets have a large amount of white space (more than half the character cell) on the side of the glyph facing outside. Where two braces are adjacent, ½ unit of space is removed between them. Existing EA (or EA-enabled world-wide) implementations would produce the wrong result even with simple inline mathematical equations. Some implementations disable access to 2329/232A for this reason in their EA profile.

For other characters where CJK and non-CJK usage of character has been unified, such special formatting rules do not apply and their usage in CJK context is infrequent.

The possible actions would be to disunify or to maintain the current unification. Disunifying the characters would require the addition of approximately ten characters. Maintaining the unification on the other hand should lead to 3008/3009 to be imaged with the narrow glyphs in the standard in keeping with the treatment of other unified symbols.

A straw poll was conducted and the majority of experts, including the Japanese experts, expressed preference to disunify the characters. However, some experts were opposed, and some preferred to study the issue further.

Recommendation: Kent Karlsson volunteered to write up a contribution (N2345) to bring this issue and background to WG2 for resolution.

2. Progressing Mathematical Symbols

Document N2336 presents a collection of mathematical characters. Document N2343 requests that the set of long arrows be added immediately to satisfy urgent needs of the MathML community to map existing ISO entity sets.

The ad-hoc reviewed the list of characters in N2336 and came to the conclusion that most of the proposed additions are well documented and not controversial. Some experts suggested that the set of geometrical shapes should be reviewed further. Also considered in need of further review were the on-line dots, the square with large dot, and the slanted parallel.

Three main alternatives were proposed:
1) add only the long arrows (this requires a new block)
2) add only characters needed to fill holes in the existing blocks
3) add all non-controversial characters

Consensus was achieved that alternative 3 would be preferable.

Recommendation: revise N2336 based on alternative 3, i.e. to remove those characters identified for further study, fill the holes left by the character to be withdrawn and present to WG2 for acceptance into the PDAM, with proposed code positions based on the recommendation in item 3 below.

3. Roadmap adjustments

A subgroup looked at the roadmap in document N2316 and estimated the needs for coding additional symbols.

Recommendation: free row 2B for additional future symbols. Use use the last three columns of row 27 for encoding the consensus subset of arrows and mathematical operators from document N2336.