**ISO/IEC JTC 1/SC 2/WG 2**
**PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS**
**FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646**
**Please fill all the sections A, B and C below.**
**Please read Principles and Procedures Document (P & P) from** http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html **for guidelines and details before filling this form.**
**Please ensure you are using the latest Form from** http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html**.**
**See also** http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html **for latest *Roadmaps*.**

**A. Administrative**

1. **Title: TWENTY SIX MATHEMATICAL CHARACTERS**
2. Requester's name: _____ Unicode Consortium / US NB_____
3. Requester type (Member body/Liaison/Individual contribution): _____liaison and NB contribution_____
4. Submission date: ___2004-Nov_____
5. Requester's reference (if applicable): _____
6. Choose one of the following:
   This is a complete proposal: _____YES___
   or, More information will be provided later: _____

**B. Technical - General**

1. Choose one of the following:
   a. This proposal is for a new script (set of characters): _____
      Proposed name of script: _____
   . b. The proposal is for addition of character(s) to an existing block: _____
      Name of the existing block: ___ Miscellaneous Mathematical Symbols-A_____
2. Number of characters in proposal: ____26_____
3. Proposed category (select one from below - see section 2.2 of P&P document):
   A-Contemporary _____ B.1-Specialized (small collection) ___X_ 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
4. Proposed Level of Implementation (1, 2 or 3) (see Annex K in P&P document): _____3_____
   Is a rationale provided for the choice? _____no_____
   If Yes, reference: _____
5. Is a repertoire including character names provided? _____YES_____
   a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document? _YES_
   b. Are the character shapes attached in a legible form suitable for review? _____YES____
6. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard? _____already available to editor_____
   If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used: _____
7. References:
   a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided? __yes_____
   b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached? ___yes____
8. Special encoding issues: none
   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)?
   _____
9. Additional Information:
   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 http://www.unicode.org/Public/UNIDATA/UCD.html and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

**C. Technical - Justification**

1. Has this proposal for addition of character(s) been submitted before? ___NO_____

   If YES explain _____

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

   If YES, with whom? ____Math. Publishers, MathML (W3C), Math Implementers_____

   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? _____

   Reference: ____users of mathematical notation, ISO entity sets_____

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

   Reference: _____

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

   If YES, where? Reference: _____MathML_____

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

   If YES, is a rationale provided? __no_____

   If YES, reference: _____

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

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

   If YES, is a rationale for its inclusion provided? _____

   If YES, reference: _____

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

   If YES, is a rationale for its inclusion provided? _____

   If YES, reference: _____

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

    If YES, is a rationale for its inclusion provided? _____

    If YES, reference: _____

11. Does the proposal include use of combining characters and/or use of composite sequences? __yes_____

    If YES, is a rationale for such use provided? ____no_____

    If YES, reference: _____

    Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? ___N/A_____

    If YES, reference: _____

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

    If YES, describe in detail (include attachment if necessary) _____

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

    If YES, is the equivalent corresponding unified ideographic character(s) identified? _____

    If YES, reference: _____

**Proposal to add 26 characters for mathematical and technical usage**

**Summary**
This proposal requests the addition of 26 characters for mathematical and technical usage. These characters complete the set of mathematical and technical characters that was expanded in Unicode 3.2, and to a lesser extent 4.0, and the corresponding amendments to ISO/IEC 10646 as consolidated in ISO/IEC 10646:2003. That expanded repertoire has been enabling the MathML project at W3C to map what used to be entitites to character code points. Recently, this mapping effort was extended and there is now an ISO project to publish a revision of ISO 9573 containing formal mappings to 10646.

In the course of this mapping project, several characters were found missing, leaving a few entities without mapping, while other entities were mapped to characters that only approximate the intended usage of the entities. This proposal contains requests to add the missing characters so that this mapping effort can be completed.

In a separate effort, a consortium of technical and scientific publishers, STIX, is engaged in creating a font encompassing the characters needed in the publication of technical and mathematical journal articles and similar works. The balance of the requested characters comes from this STIX project of reviewing the mathematical and technical literature. The characters requested in this document are primarily complements of characters already existing in the standard and their rationale for their inclusion is therefore based on inherent regularities in mathematical notation and usage.

The list of characters has been reviewed by several experts. The reviewers feel that the both the user community and the academic publishers have an interest in being able to encode the character in question.

The following tables listed the requested characters with suggested code position, shape and name. For each proposed character, a description field gives information about the character, in particular any known mappings to entity sets. As comment field contains information relating the character to other, related characters that are already encoded and in some cases gives more detailed rationale for its inclusion. A character annotated with 'source: ISO…' was reported in connection with mapping to ISO 9573 entity sets. All others were reported in connection with completing the STIX review.

**References:**

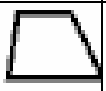
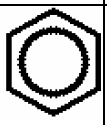
STIX project home page: http://www.ams.org/STIX (this site has links to the STIX font project and other sites).

MathML specification can be accessed at http://www.w3.org

ISO 9573 mapping project, see for example http://www.w3.org/2003/entities

**Table 1: Proposed additions of Mathematical Operators, and Technical symbols**

Note: **Bold** ID number used for character from 9573 mappings

| ID | Code | Shape | Name | Description | Comments & Notes |
|---|---|---|---|---|---|
| 1 | 27C7 |  | OR WITH DOT INSIDE | logical or with dot inside | Complements the existing 27D1 ⟑ AND WITH DOT INSIDE |
| **2** | 27C8 |  | REVERSE SOLIDUS PRECEDING SUBSET | reverse solidus followed by subset = ⟈ (afii DBF4 | operators are usually single characters or at best combining sequences. Source: ISOAMSA |
| **3** | 27C9 |  | SUPERSET PRECEDING SOLIDUS | Superset followed by solidus = ⟉ (afii D95C) | operators are usually single characters or at best combining sequences. Source: ISOAMSA |
| **4** | 23E1 |  | ELECTRICAL INTERSECTION | - electrical intersection = ⏧ (afii DB4E) | Source: ISOTECH |
| **5** | 23E2 |  | WHITE TRAPEZIUM | - trapezium = ⏢ (afii DBB8) | Source: ISOTECH |
| 6 | 23E3 |  | BENZENE RING WITH CIRCLE | - benzene ring [hexagon] with circle = &benzenr; (afii D8DC) | This is a variant of 232C ⌬ BENZENE RING, but it should not be unified. (see note at end). ISOCHEM |
| **7** | 23E4 |  | STRAIGHTNESS | - "straightness" = ¯ (afii EE49) | **Drafting Symbol.** Possibly unifiable with some existing horizontal line – but this is **not** a dash |
| **8** | 23E5 |  | FLATNESS | - "flatness" = ▱ (afii EE4A) | **Drafting symbol.** Generally not unifiable with 25B1 Parallelogram. Source: ISOTECH |
| **9** | 23E6 |  | AC CURRENT | - ac current = ∿ (afii DB3B) | While 223F ∿ SINE WAVE may be used to express AC semantics, as a symbol its not unifiable with this character. Source: ISOTECH |
| 10 | 26B2 |  | NEUTER | - neuter [circle with short vertical below] | While the semantics of neuter can be represented with MEDIUM WHITE CIRCLE, this symbol can- |

| | | | | | |
|---|---|---|---|---|---|
| | | | | | not be unified with 26AA |
| **11** | 1D7CA | Ϝ | MATHEMATICAL BOLD CAPITAL DI-GAMMA | b.Gammad(9573-2003-isogrk4) | Source: ISOGRK4 U+03DC is mapped to Gammad (9573-2003-isogrk3) |
| **12** | 1D7CB | ϝ | MATHEMATICAL BOLD SMALL DI-GAMMA | b.gammad(9573-2003-isogrk4) | Source: ISOGRK4 U+03DD is mapped to gammad(9573-2003-isogrk3): |

**Table 2: Proposed additions of combining diacritics for symbols**

These characters are generically needed to place a harpoon or arrow above, resp. below a variable or mathematical expression.

| ID | Code | Shape | Name | Description | Comments & Notes |
|---|---|---|---|---|---|
| 13 | 20EC |  | COMBINING RIGHT-WARDS HARPOON WITH BARB DOWN-WARDS | - combining over right harpoon down | Compare U+21C1 RIGHT-WARDS HARPOON WITH BARB DOWN-WARDS |
| 14 | 20ED |  | COMBINING LEFT-WARDS HARPOON WITH BARB DOWN-WARDS | - combining over left harpoon down | Compare U+21BD LEFT-WARDS HARPOON WITH BARB DOWN-WARDS |
| 15 | 20EE |  | COMBINING LEFT ARROW BELOW | - combining under left arrow | Complements 20D6 COMBINING LEFT AR-ROW ABOVE |
| 16 | 20EF |  | COMBINING RIGHT ARROW BELOW | - combining under right arrow | Complements 20D7 COM-BINING RIGHT ARROW ABOVE |

**Table 3: Half filled shapes from STIX review**

| ID | Code | Shape | Name | Description | Comments & Notes |
|---|---|---|---|---|---|
| 17 | 2B14 |  | SQUARE WITH UPPER RIGHT DIAGONAL HALF BLACK | - square, filled top right corner = &squarftr; (no afii) | See U+25E9 ◩ SQUARE WITH UPPER LEFT DI-AGONAL HALF BLACK |
| 18 | 2B15 |  | SQUARE WITH LOWER LEFT DI-AGONAL HALF | - square, filled bottom left corner = &squarfbl; (no | See U+25EA ◪ SQUARE WITH LOWER RIGHT DIAGONAL HALF |

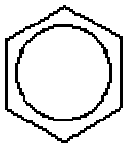
| | | | | | |
|---|---|---|---|---|---|
| | | | BLACK | afii) | BLACK |
| 19 | 2B16 |  | DIAMOND WITH LEFT HALF BLACK | - diamond, filled left half = &diamonfl; | Compare 25E7 ◧ SQUARE … 25D0 ◐ CIRCLE…and 25ED ◭ UP-POINTING TRIANGLE WITH LEFT HALF BLACK |
| 20 | 2B17 |  | DIAMOND WITH RIGHT HALF BLACK | - diamond, filled right half = &diamonfr; | Compare 25E8 ◨ SQUARE WITH LEFT HALF BLACK, etc. |
| 21 | 2B18 |  | DIAMOND WITH TOP HALF BLACK | - diamond, filled bottom half = &diamonfb; | Compare 2B12 S-QUARE WITH TOP HALF BLACK, etc. |
| 22 | 2B19 |  | DIAMOND WITH BOTTOM HALF BLACK | - diamond, filled top half = &diamonft; | Compare 2B13 SQUARE WITH BOTTOM HALF BLACK, etc. |

**Table 4: Geometric shapes from STIX review**

| ID | Code | Shape | Name | Description | Comments & Notes |
|---|---|---|---|---|---|
| 23 | 2B20 |  | WHITE PENTAGON | - open pentagon (afii DB2D) | |
| 24 | 2B21 |  | WHITE HEXAGON | - benzene ring [open hexagon] = &benzen; (no afii) | The shape requested is a hexagon, even though it's labeled &benzen; here |
| 25 | 2B22 |  | BLACK HEXAGON | - filled hexagon | |
| 26 | 2B23 |  | HORIZONTAL BLACK HEXAGON | - horizontal filled hexagon | |

**The symbol for benzene**

The Kekulé structure for benzene, consisting of a hexagon denoting the ring of six carbon atoms, each of which has one hydrogen attached, and three doubled lines, denoting alternating single and double bonds is the reference glyph for 232C BENZENE RING.



Some authors prefer it, but many others deliberately replace it by the more modern symbol, shown here on the right and in Table 1, which shows a hexagon with an inscribed circle. While hydrogen and carbon atoms are implied by the corners of the diagram in the usual manner, the circle represents the delocalized electrons. The modern diagram is felt to better represent the actual physical structure of benzene, which has six equal bonds of average length, not three shorter double bonds and three longer single bonds.

Without lines or circle, that is as a bare hexagon, the symbol represents cyclohexane and not benzene.

Unlike the Kekulé structure, it is not possible to deduce the number of hydrogen atoms from the benzene symbol with the circle. On the other hand, the chemical bonding of Benzene is quite different from a series of alternating single and double bonds as suggested by the Kekulé structure. This is because the electrons are delocalized due to a process called resonance.

While both forms of the symbol unambiguously represent the same chemical molecule, it appears that the choice of the particular representation is often quite deliberate, as each symbol emphasizes different aspects of the structure. Even a cursory examination of the subject will lead to a paper or website where authors give and defend opposite preferences, and almost all introductory texts indeed present both symbols, until establishing a convention in favor of one or the other.

These two forms should therefore be disunified. Unlike the differences in shape captured by standardized variation sequences for mathematical symbols, the differences in shape and identifiable motivation in usage seem pronounced enough that there would be little benefit over adding a separate character.

The BENZENE RING WITH CIRCLE should be added at 23B3.

**Notes on certain characters**

The rationale for including characters listed with bold ID numbers in the table is that they are needed to map the entities in ISO 9573. This information is not repeated here.

#1 OR WITH DOT INSIDE

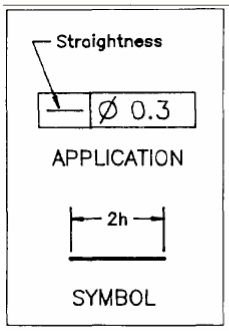
The rationale for adding this character is that it is an obvious pair with 27D1 AND WITH DOT INSIDE. As one reviewer wrote “if you have the one you need the other: I'm sure I've seen it, as is my logic colleague”

#3 SUPERSET FOLLOWED BY SOLIDUS

Note that the AFII glyph database incorrectly contains a glyph with a reverse solidus for this character.

#7, 8 STRAIGHTNESS, FLATNESS

These characters (and a number of other symbols, many of which were already in Unicode/10646, for example 232D - 2335) are from a group of dimensioning and tolerancing symbols that are defined in ASME Y 14.5-1994 and are widely referenced, for example in manuals. As defined in the standard, all of these symbols have specific dimensions which makes unification with existing symbols of similar basic geometry somewhat problematic.




**3.2.1** *Straightness* - A straight line is used to indicate a straightness requirement. See Figure 19. It is only applied in a feature control frame, and may be used to control straightness of surface elements. It may also be used to control the straightness of an axis or center plane.

**Figure 19. Straightness Symbol**

**3.2.2** *Flatness* - The flatness symbol appears as an oblique view of a square surface. See Figure 20. This symbol is used in feature control frames and is only used to control the form variations on flat features.

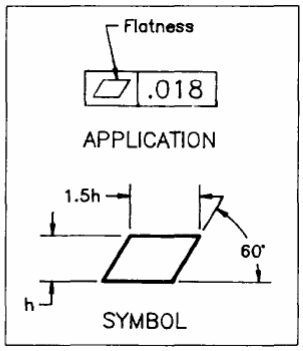



**Figure 20. Flatness Symbol**

#10 NEUTER

This symbol was submitted to the STIX collection by two of the scientific societies, ACS (American Chemical Society), with an entity name of “tmpneuter” and described as a “medium circle with short vertical bar” and AIP (American Institute of Physics) which called it “cirmid”.

The Unicode Standard encodes a medium white circle at 26AA which is aliased as ‘sexless’ or ‘genderless’ in the context of gender symbols. A search of biology references on the web shows the use of U+2205 ∅ EMPTY SET as an indication of ‘neuter’ or ‘sterile’ (See for example ‘Signs and Measures’ in: http://floraseries.landcareresearch.co.nz/pages/Book.aspx?fileName=Flora%201.xml) In light of this apparent non-uniformity of usage, the name for the proposed character should perhaps be based on the character shape, not its imputed function.

Related Symbols:

⚲ This symbol, with a vertical line as long as the diameter of the circle, is cited as an old symbol for Aphrodite. One source (symbols.com) notes that “The sign ⚲ is commonly used in modern technical, electrical, and military systems…sometimes ⚲ is used for mine on maps.”

⚲ In MathML, the “&cirmid” entity, which is also found in ISOAMSA, is mapped to the Unicode symbol U+2AEF ⫯ VERTICAL LINE WITH CIRCLE ABOVE.

#13-16 COMBINING HARPOON AND ARROWS

In mathematical notation the decoration of variables or formulas with arrows and other symbols is common. In mathematical layout systems such as TeX, combining characters can be defined on the fly, allowing unrestricted use of over/under arrows and harpoons. The proposed characters are the obvious completion of a set of characters already encoded in Unicode/10646 (combining arrows) that are genuinely useful in technical notation.

Of the characters proposed here, one reviewer wrote: “The use of left or right harpoons above is 'relatively common' in the theoretical physics literature by those who wish to be classier than using arrows. Under arrows I would agree to having seen too. I'm pretty sure I should be able to find over harpoons in my personal library, probably in something dealing with the mathematical end of current algebras (algebras of currents).”

Another reviewer expressed this desire for a complete repertoire more directly: “why aren't you [also] adding a combining left right arrow below to go with 20E1?”

#17-22 Half-filled shapes

The proposed characters are found in the STIX master list of symbols requested by member societies. The diamonds with filled top and bottom half were submitted to STIX by AIP (American Institute of Physics) the ones with filled left and right half were submitted by Elsevier.

The entity names are from an ISOPUB entity set and were corroborated as such in 1997 by the then editor of 9573-13. These characters are completing the set of such characters in the Geometrical Shapes and Miscellaneous Symbols and Arrows blocks by using typical symmetry operations; rotations, reflections and exchange of black / white.

#23 PENTAGON
The pentagon came from Elsevier.

#24 WHITE HEXAGON

The white hexagon provides the generic shape in contrast to the benzene ring. In responding to a question whether a further form, i.e. a white horizontal hexagon was also needed, one reviewer wrote: “In my view these shapes, if there, will be used, and if the 'natural completion' of a set is not there, there will come complaints.”

The vertical open hexagon was submitted byACS (American Chemical Society), and was also in a collection submitted by Design Science. (U+2394, the horizontal hexagon, was submitted by AIP and Elsevier.)

#25-26 Black Hexagons

Both orientations of filled hexagon were submitted to STIX by ACS (American Chemical Society). Also included in a collection sent to STIX from kluwer,

## Proposed Character Properties

As these are additions to existing blocks, most characters properties will be the same as for similar symbols in the given blocks. The following properties deserve explicit note.

### *General Category*

Mn: 20EC
20ED
20EE
20EF
Sm: 27C7
27C8
27C9
So: 23E1..23E6
26B2
2B14..2B19
2B20..2B23
Lu: 1D7CA
Ll: 1D7CB

### *Decompositions*

1D7CA <font> 03DC
1D7CB <font> 03DD

### *Other_Math*

1D7CA
1D7CB

### *Bidi*

ON: <all, except as listed>
L: 1D7CA
1D7CB

### *LineBreak*

AL: <all, except as listed>
CM: 20EC
20ED
20EE
20EF

### *East Asian Width*

N: <all>