

To: UTC and ISO/IEC JTC1/SC2 WG2
Title: Proposal to encode GREEK CAPITAL LETTER YOT
From: Michael Bobeck
Date: 12 December 2010

I wish to propose the addition of this character to the “Greek and Coptic” block in isolated U+037F codepoint.

Notes

- isolated U+03F3 is already used by small GREEK LETTER YOT
- U+0378 & U+0379; U+0380 & U+0381; U+0382 & U+0383 are more suitable for contiguous cased pairs
- isolated U+03A2 is reserved for basic Greek alphabet charting purposes
- semiisolated U+038B & U+038D are more suitable for semicontiguous cased pair

Proposed Characters

	037
0	
1	
2	
3	
4	
5	
6	
7	
8	
9	
A	
B	
C	
D	
E	
F	J

Character Names

037F	GREEK CAPITAL LETTER YOT • Lowercase is U+03F3
------	---

Annotation additions

03F3	GREEK LETTER YOT • Uppercase is U+037F
------	---

Unicode Character Properties

Properties for the new character:

```
037F;GREEK CAPITAL LETTER YOT;Lu;0;L;;;;;N;;;03F3;
```

Properties for U+03F3 should now reference the new uppercase letter:

```
03F3;GREEK LETTER YOT;Ll;0;L;;;;;N;;;037F;;037F
```

Other Information

Greek and Coptic

The creators of fonts like Vusillus and New Athena Unicode had since 2003/2004 established implementations of GREEK CAPITAL LETTER YOT, which were put into **Private Use Area** block. These precarious implementations include both cases of Yot. Each user of these fonts is prone to nonstable codepoint value of capital Yot, which thwarts interoperation between Greek text creators, if they use mutually incompatible codepoints. Capital Yot is the part of the etymological Greek orthography, and is required for at least ALL CAPS typesetting of the Greek language at least in chapter titles and in epigraphic specimens when Greek language is subjected to the etymological treatment. The lowercase Yot is already in Unicode at the U+03F3. ALL CAPS spelling of Greek language as used both in chapter titles and in etymologically treated epigraphy, obviously needs to have the uppercase variant of the U+03F3 and could theoretically use this proposed uppercase Yot if it would be accepted into Unicode. Currently, etymologies of Greek epigraphy are using other options, such as Latin J (jay) or Cyrillic J (je) for the uppercase equivalent to U+03F3, but at expense of frequent typographic errors in processed text. It is obvious that people are searching for solutions to uppercase, and since there are existing at least two fonts which are actually using the GREEK CAPITAL LETTER YOT, this missing uppercase should be added to the Unicode Standard into still unallocated U+037F codepoint. Examples of unofficial implementation and usage are in figures below.

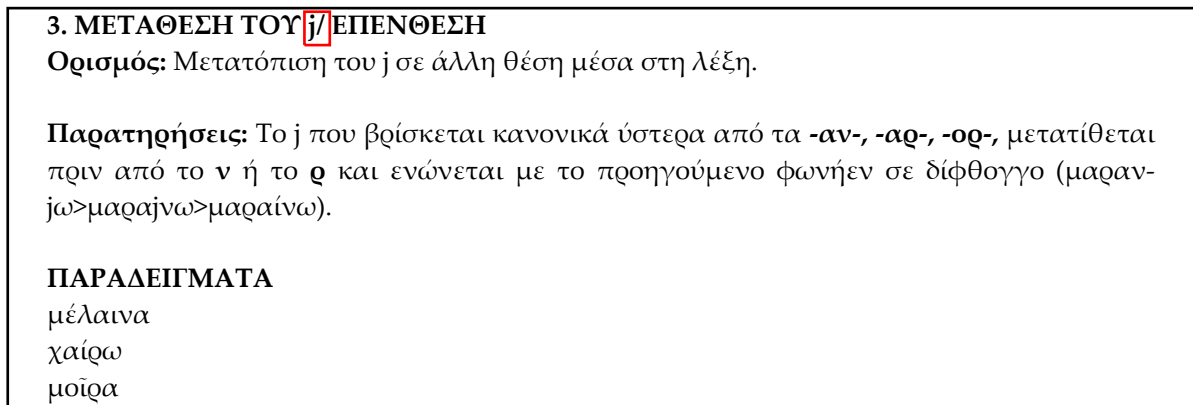


Figure 1. GREEK CAPITAL LETTER YOT as it is not correctly rendered in ALL CAPS Greek text
Source: Εβίνα Σιστάκου, ΠΑΘΗ ΦΘΟΓΓΩΝ & Εβίνα Σιστάκου, ΧΕΙΜΕΡΙΝΟ ΕΞΑΜΗΝΟ
<http://www-old.lit.auth.gr/public/aef-101-sistakou.pdf> (page 7)
http://www.lit.auth.gr/sites/default/files/documents/AEF_101_SISTAKOU-Corpus_0.pdf (page 14)



Figure 2. GREEK CAPITAL LETTER YOT as it should be rendered correctly in ALL CAPS Greek text

(Α) ὅτε τὰ(ν) πτόλιν Ἐδάλιον κατέφορον Μᾶδοι κὰς Κετιῆφες ἰ(ν) τῷ Φιλοκύπρῳ φέτει τῷ Ὀνασαγό|ραν, βασιλεὺς Στασίκυπρος κὰς ἁ πτόλις Ἐδαλιῆφες ἄνωγον Ὀνάσιλον τὸν Ὀνασικύπ|ρῳ τὸν ἵ|ατῆραν κὰς τὸς κασιγνέτος ἵ|ασθαι τὸς ἁ(ν)θρόπος τὸς ἰ(ν) τᾷ μάχαι ἰγ|μαμένος ἄνευ μισθῶν· κὰς παι εὐφρετάσατν βασιλεὺς κὰς ἁ πτόλις Ὀνασί⁵λῳι κὰς τοῖς κασιγνέτοις ἁ(ν)τι τῷ μισθῶν κὰ ἁ(ν)τι τᾷ ὑχέρῳν δοφέναι ἐξ τῷ | φοῖκῳι τῷ βασιλῆφος κὰς ἐξ τᾷ πτόλιφι ἄργύρῳ τὰ(λαντον) ἁ' τὰ(λαντον)· ἔ θυφάνοι νν ἁ(ν)τι τῷ | ἄργύρῳ τῷδε, τῷ ταλά(ν)τῶν, βασιλεὺς κὰς ἁ πτόλις Ὀνασίλῳι κὰς τοῖς κασιγνέτοις ἁπὺ τᾷ ζᾷ τᾷ βασιλῆφος τᾷ ἰ(ν) τῷρῳνι τῷ Ἄλα(μ)πρῳ|άται τὸ(ν) χῶρον | τὸν ἰ(ν) τῷ ἔλει τὸ(ν) χρανόμενον Ὀ(γ)κα(ν)τος ἄλφῳ κὰς τὰ τέρχν|α τὰ ἐπιό(ν)τα ¹⁰πά(ν)τα ἔχεν πανόνιον ὑφαις ζαν ἄτελέν. ἔ κέ σις Ὀνάσιλον ἔ τὸς | κασιγνέτος ἔ τὸς παῖδας τῷ(ν) παῖδῶν τὸν Ὀνασικύπρῳ ἐξ τῷ χῶρῳι τῷδε | ἐξ ὀρύξῃ, ἰδέ παι ὃ ἐξ ὀρύξῃ, πείσει Ὀνασίλῳι κὰς τοῖς κασιγνέτοις ἔ τοῖς παισὶ τὸν ἄργυρον τὸ(ν)δε, ἄργύρῳ τὰ(λαντον) ἁ' τὰ(λαντον). | κὰς Ὀνασίλῳι οἴφῳι ἄνευ τῷ(ν) κασιγνέτῶν τὸν ἄλλῶν ἐφρετάσατν βασιλεὺς¹⁵ κὰς ἁ πτόλις δοφέναι ἁ(ν)τι τᾷ ὑχέρῳν τῷ μισθῶν ἄργύρῳ πε δ' πε | β' δι Ἐ(δάλια)· ἔ δῶκοι νν βασιλεὺς κὰς ἁ πτόλις Ὀνασί(Β)λῳι ἁ(ν)τι τῷ ἄργύρῳ τῷδε ἁπὺ τᾷ ζᾷ τᾷ βασιλῆφος τᾷ ἰ(ν) Μαλαν|ἵα τᾷ πεδῳ|αι τὸ(ν) χῶρον τὸ(ν) χρανζόμενον Ἀμεν|ἵα ἄλφῳ κὰς τὰ τέρχν|α τὰ ἐπιό(ν)τα πά(ν)τα, τὸ(ν) ποεχόμενον πὸς τὸ(ν) ῥόφο(ν) τὸ(ν) Δρύμιον κὰς πὸ²⁰ τᾶν ἱερέφ|ῳαν τᾶς Ἀθάνας, κὰς τὸ(ν) κᾶπον τὸν ἰ(ν) Σίμιδος ἄρουρα|ι, τὸ(ν) Διφείθεις ὃ Ἀρμανεὺς ἔχε ἄλφο(ν), τὸν ποεχόμενον πὸς Πα(ν)σαγόρα|ν τὸν Ὀνασαγόραν κὰς τὰ τέρχν|α τὰ ἐπιό(ν)τα πά(ν)τα ἔχεν πανόνιος ὑ|φαις ζαν ἄτελ|ἵα ἰό(ν)τα. ἔ κέ σις Ὀνάσιλον ἔ τὸς παῖδας τὸς Ὀνασίλῳν ἐξ τᾷ ζᾷ τᾷδε ἰ ἐξ τῷ κάπῳι τῷδε ἐξ ὀρύξῃ, ἰ²⁵δὲ ὃ ἐξ ὀρύξῃ, πείσει Ὀνασίλῳι ἔ τοῖς παισὶ τὸν ἄργυρον τὸ(ν)δε, ἄργύρῳ|ν πε δ' πε β' δι Ἐ(δάλια). ἰδέ τὰ(ν) δάλτον τὰ(ν)δε, τὰ φέπ|ῳα τὰδε ἰναλαλισμένα | βασιλεὺς

Figure 3. GREEK LETTER YOT rendered in mixed case Greek text derived from original ancient ALL CAPS Cypriot inscription (Tituli Cyprii)

In case of ALL CAPS verbatim etymological rendition of source epigraph,
GREEK CAPITAL LETTER YOT is needed

Source: Eduard Schwyzer, *Dialectorum Graecarum exempla epigraphica potiora*, Lipsiae, in aedibus Salomonis Hirszelii, 1923. (page 679)

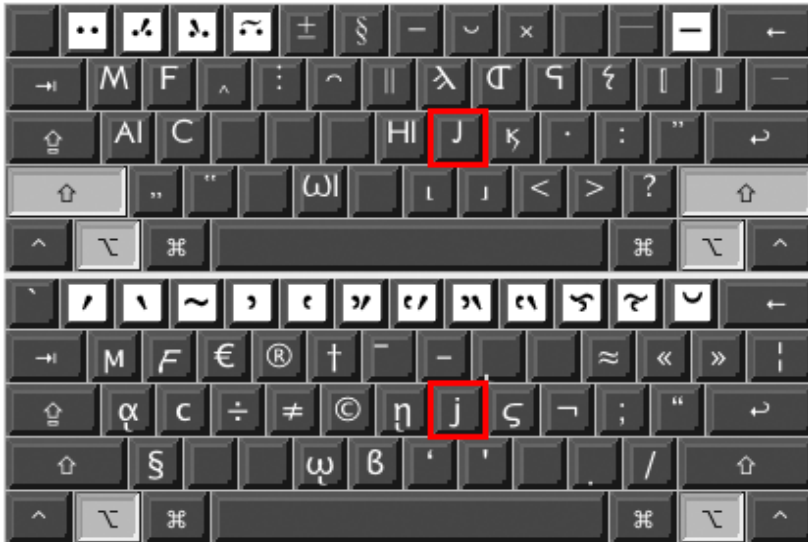


Figure 4. GREEK CAPITAL LETTER YOT and GREEK LETTER YOT in GreekKeys Unicode keyboard and font package

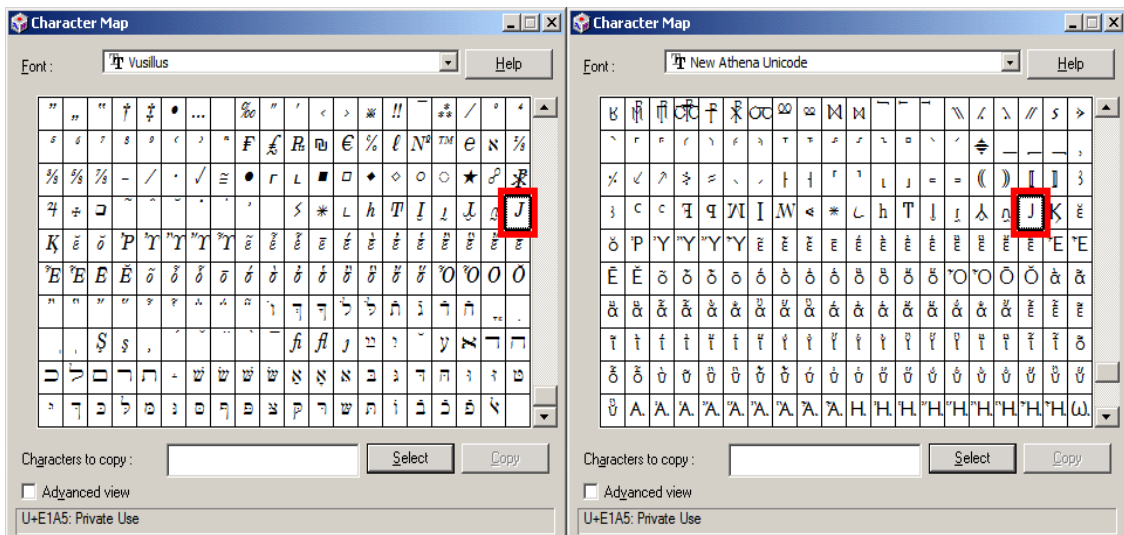


Figure 5. GREEK CAPITAL LETTER YOT in New Athena Unicode and Vusillus fonts

Technical references

- <http://apagreekkeys.org/NAUdownload.html>
- http://apagreekkeys.org/fonts/NAU3_71.zip/NAU3_71/newathu.ttf (New Athena Unicode font inside this ZIP)
- <http://www.users.dircon.co.uk/~hancock/antioch.htm>
- http://www.users.dircon.co.uk/~hancock/an2setup.exe/vu3i____.ttf (Vusillus font inside this SFX CAB)
- <http://ist-socrates.berkeley.edu/~pinax/greekkeys/AboutGK2008.html>
- <http://ist-socrates.berkeley.edu/~pinax/greekkeys/technicalDetails.html>
- <http://isites.harvard.edu/fs/docs/icb.topic537340.files/GKUKeyboard.pdf>
- <http://omega.enstb.org/yannis/pdf/boston99.pdf> (section 1.2.2)
- <http://www.tlg.uci.edu/~opoudjis/unicode/yot.html>
- http://www.tlg.uci.edu/~opoudjis/unicode/unicode_gaps.html (section 2.2.3)
- http://www.tlg.uci.edu/~opoudjis/unicode/unicode_mixing.html (section 5)
- <http://www.tlg.uci.edu/~opoudjis/unicode/maybes.html>

**ISO/IEC JTC 1/SC 2/WG 2
PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS
FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646¹**

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:	<u>Proposal to Encode Greek Capital Letter Yot</u>
2. Requester's name:	<u>Michael Bobeck</u>
3. Requester type (Member body/Liaison/Individual contribution):	<u>Individual</u>
4. Submission date:	<u>12 December 2010</u>
5. Requester's reference (if applicable):	_____
6. Choose one of the following:	_____
This is a complete proposal:	<u>Yes</u>
or, More information will be provided later:	<u>No</u>

B. Technical – General

1. Choose one of the following:	
a. This proposal is for a new script (set of characters):	<u>No</u>
Proposed name of script:	_____
b. The proposal is for addition of character(s) to an existing block:	<u>Yes</u>
Name of the existing block:	<u>Greek and Coptic</u>
2. Number of characters in proposal:	<u>1</u>
3. Proposed category (select one from below - see section 2.2 of P&P document):	
A-Contemporary <u>X</u>	B.1-Specialized (small 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):	<u>1</u>
Is a rationale provided for the choice?	<u>No</u>
If Yes, reference:	_____
5. Is a repertoire including character names provided?	<u>Yes</u>
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?	<u>Yes</u>
b. Are the character shapes attached in a legible form suitable for review?	<u>Yes</u>
6. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?	_____
If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:	<u>Use Latin J from Fiscella font already used inside existing Unicode PDFs for publishing purposes</u>
7. References:	
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?	<u>Yes</u>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?	<u>Yes</u>
8. Special encoding issues:	
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)?	<u>Suggested character properties are included.</u>

¹ Form number: N2652-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11)

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? If YES explain _____	<u>No</u>
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.)? If YES, with whom? _____ If YES, available relevant documents: _____	<u>No</u>
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included? Reference: <u>See proposal</u>	<u>No</u>
4. The context of use for the proposed characters (type of use; common or rare) Reference: <u>See proposal</u>	<u>common</u>
5. Are the proposed characters in current use by the user community? If YES, where? Reference: <u>See proposal</u>	<u>Yes</u>
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP? If YES, is a rationale provided? _____ If YES, reference: <u>Both lettercases preferably should be inside the same block</u>	<u>Yes</u> <u>Yes</u>
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)? <u>Now, both lettercases can be only scattered within the same block</u>	
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence? If YES, is a rationale for its inclusion provided? _____ If YES, reference: _____	<u>No</u>
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters? If YES, is a rationale for its inclusion provided? _____ If YES, reference: _____	<u>No</u>
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character? If YES, is a rationale for its inclusion provided? _____ If YES, reference: _____	<u>No</u>
11. Does the proposal include use of combining characters and/or use of composite sequences? If YES, is a rationale for such use provided? _____ If YES, reference: _____ Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided? _____	<u>No</u> <u>No</u>
12. Does the proposal contain characters with any special properties such as control function or similar semantics? If YES, describe in detail (include attachment if necessary) _____	<u>No</u>
13. Does the proposal contain any Ideographic compatibility character(s)? If YES, is the equivalent corresponding unified ideographic character(s) identified? _____ If YES, reference: _____	<u>No</u>