
On T_EX and Greek...

Yannis Haralambous

1 Some historical background

The first person who made greek characters for T_EX was D. E. Knuth; each CM font contains the greek uppercase characters. These are excellent for greek text as well as for mathematics. Only two remarks could be made: first, the Υ is actually a calligraphic form of Y and the latter is more frequently used in Greece; second, the typewriter Γ, Δ, ..., Ω don't have much in common with greek typewriting.

Knuth also did the well known lowercase greek characters $\alpha, \beta \dots \omega$ for use in mathematics. These are not suitable for greek text.

The American Mathematical Society added a digamma (*F*), probably for the sake of completeness (the author would be interested to find out in which area of mathematics a digamma is or could be used).

Then came Silvio Levy [1], who created an entirely new family of greek fonts (roman, *slanted*, **boldface**, *typewriter*). These fonts are partly coded in 8-bit and Silvio uses ligatures and macros to obtain the characters in the ASCII range 129–256.

As a matter of fact, Silvio's idea was to have as many combinations of accents, breathings and subscript iota as possible and at the same time integrate a new feature concerning sigmas, namely an automatic choice between the medial σ and the final ς . For this, he included every possible combination σ + *(character)* as characters of the font (obtained by ligatures) and defined ς to be ς in all other cases. This is a brilliant idea, but has a major disadvantage: instead of 2 font positions σ, ς it needs 58. This forces him, because of lack of space, to use the `\accent` primitive in the case of the combination *(breathing)* + *(grave accent)* (`^`, `~`). Of course, as he already points out, this doesn't harm hyphenation since these accent combinations appear only on monosyllables. Also, Silvio announced an hyphenation table, without specifying for which kind of Greek.

A year later, Klaus Thull and the author [2] made some changes to Silvio's fonts:

- reducing them to 128 characters, because of problems with some drivers,
- making a new family of one-accent fonts (the official system in Greece),
- adding a SMALL CAPITALS font, a few ancient greek characters ($\Upsilon, \Phi, \Psi, \Omega, \Lambda$) and hyphenation patterns for modern "post-1974" greek ($\delta\eta\mu\sigma\tau\iota\kappa\acute{\eta}$).

2 What comes next

Since the release of T_EX 3.0, many efforts have been undertaken to establish standards for 256-character CM fonts. These fonts – or at least tentative versions of them – have been presented in Cork 1990 and in Vienna 1991 and contain a maximum of accented characters to suit most of the occidental languages.

Besides supporting 8-bit input, the new T_EX also allows hyphenation of different languages in the same document. In this way, soon we will have different pattern tables included in the plain format of the standard T_EX-distribution. As a matter of fact, Mike Ferguson has proposed to coordinate such a collection of hyphenation patterns.

And what about greek? Before considering hyphenation problems, a standard font table should be fixed! I think it is time to take this step: choosing a standard greek font table and starting to work with it. In the appendix of this note, I propose a font table. All combinations of characters, accents, breathings and subscript iota are present as individual characters, reached by ligatures. Special symbols by Oxford in epigraphical texts [3] are included. Note that positions '013 and '040 are blank on purpose to maintain compatibility with PostScript fonts¹.

Of course this is just a proposal and I would like to collect your opinions and ideas. Other possible choices are:

1. Silvio's font table (see [1, p. 24]),
2. the greek standard 8-bit ASCII code,
3. greek PC (for example [4, p. 212]) or Macintosh (greek family script) 8-bit ASCII code, and finally
4. an entirely new one.

Choices 2 and 3 are more "open to the outside world" but have a major disadvantage: they don't contain all characters needed for classical greek.

Concerning multi-accent and one-accent systems, I think it is only natural to share the same font table (in the second case all breathings and the subscript iota will be missing and all accents replaced by the "universal accent", which also needs to be standardized!). Of course many greek T_EX-users will say that 'multi-accented characters are only old junk'; I would answer that the first principle of T_EX is universality, therefore the fonts used for classical or modern greek should be compatible.

¹ Have you ever tried to use `cmr10` in Illustrator 3.0?

3 Problems.

Besides a choice of the font table, there are some problems which should be discussed, concerning the transliteration of Greek in 7-bit ASCII. The latter can still be useful for electronic mail and other media or devices which don't allow 8-bit. I think that Silvio's transliteration

| | | | | | | | | | | | |
|---|---|---|---|---|------|---|---|---|---|---|---|
| α | β | γ | δ | ε | ζ | η | θ | ι | κ | λ | μ |
| a | b | g | d | e | z | h | j | i | k | l | m |
| ν | ξ | ο | π | ρ | σ, ζ | τ | υ | φ | χ | ψ | ω |
| n | x | o | p | r | s | t | u | f | q | y | w |

has become a standard (I still get messages from people using x for χ, th for θ, y for υ, which can cause quite a confusion...). In [2] a slight modification to Silvio's transliteration is proposed: c for ζ. I have another modification to propose: v, V for the letter digamma Ϝ, ϝ. This is some kind of rehabilitation of this letter, which was rather common in some ancient greek idioms (who would deny the beauty of Sappho's

ἀλλὰ καὶ μὲν γλῶσσα φέαγε, λέπτον
δ' αὐτικὰ χρωῖ πῦρ ὑπαδερόμακον,
ὀππάτεσι δ' οὐδ' ἐν ὄρρημ' ἐπιρρόμβεισι
[δ' ἄκουαι)

But there is a problem: the circumflex accent. Silvio encodes it as ~ which—at least visually—is the most natural choice. But ~ is active in T_EX and plays an important rôle in line breaking. Silvio has to enclose greek text inside a group where ~ is non-active; this brings several inconveniences, which could be avoided. I propose to change the transliteration of the circumflex accent. The next most natural choice would be ^ (in french “*accent circonflexe*”). But unfortunately, the same problem would appear. If you look at the standard ASCII table (see [5, p. 367]) only the four characters *, +, =, | are possible choices. The characters * and | should be eliminated because of their uses in various T_EX-constructions. Only + and = remain unused. Therefore I propose = as a 7-bit transliteration of the circumflex accent. A sequence like

Ποῦ πῆγε ὁ ζῆλος τῶν παλιῶν ἡμερῶν;

would be typed as

```
Po=u p=hge <o z=hloc t=wn pali=wn
      <hmer=wn?
```

For the minor inconvenience of typing \$=\$ when you really want a =² you will be using greek fonts in a straightforward manner, like CM fonts.

² according to Knuth ([5, p. 51]): You can also type + and =, to get the corresponding symbols + and =; but it's much better to use such characters only in math mode, i.e., enclosed between two \$

Similar problems arise with diaeresis (διαλυτικά) and subscript iota (ὑπογεγραμμένη). All this has to be discussed.

4 A Greek T_EX Users (sub-) Group

There is a lot of work which remains to be done —and more and more people wanting to use T_EX for greek, and potentially being able to contribute—:

- hyphenation tables should be made according to T_EX 3.0 and the new standard font (once it's established). The recent release of a greek dictionary on electronic media should be exploited;
- L^AT_EX and $\mathcal{A}\mathcal{M}\mathcal{S}$ -T_EX styles should be created;
- new fonts, like a typewriter font which would look like greek typewriting, and sans-serifs should be made;
- greek PostScript fonts should be gathered and the corresponding .t_fm files created. They then could be used as virtual fonts;
- introductory and T_EX-educational material should be translated to greek; and finally the most important:
- all this should be gathered, ordered, documented and made available by the maximum number of servers. Some part of it could also be included in the standard T_EX distribution.

And before even starting, there should be a solid discussion on the standardization problems I mentioned in previous sections.

Therefore I solemnly propose the creation of some structure (or sub-structure of TUG) which will take the responsibility for distributing and coordinating the tasks. Please send me your opinions, ideas and proposals, either to my bitnet address or to my physical address. I propose that the deadline (minus a week) for submission of technical papers for the first regular issue of *TUGboat*, volume 13 (November 19), be a (tentative) deadline for your answers, after which I'll collect the informations and statistics, and announce them through the *TUGboat* (that is, after agreement of the editorial board).

I especially invite T_EX users in Greece to contact me, because of my lacunary knowledge on T_EX-activities in Greece.

5 Say it in Greek...

Καί σὰ ἐλληνικά τώρα: πρώτα ἀπ' ὅλα θὰ ἤθελα νὰ εὐχαριστήσω τοὺς ἐκδότες τοῦ *TUGboat* ποὺ μοῦ ἐπέτρεψαν νὰ ἐκφραστῶ σὰ ἐλληνικά μέσα ἀπ' τὶς στήλες τοῦ περιοδικοῦ. Κι αὐτὸ γιὰ νὰ σᾶς προτείνω, μὲ πιὸ πολλὰ

signs, since that tells T_EX to insert the proper spacing for mathematics.

έμφαση να συνεργασθούμε όλοι μαζί για να φτιάξουμε ένα ελληνικό T_EX που να ξεπερνάει τις νόρμες των διαφόρων χωρών, να είναι κοινό για τα άρχαία και τα νέα, και που να έκμεταλλεύεται όλες τις δυνατότητες των έκπεκτάσεων του T_EX: L^AT_EX, A_MS-T_EX κ.δ.κ.

Προηγουμένως ανέφερα μια σειρά από έργασίες που πρέπει ακόμα να γίνουν. Νομίζω ότι μόνο με μια οργανωμένη προσπάθεια μπορούμε να ανταπεξέλθουμε σ' αυτό τó έργο, σε πεπερασμένο χρόνο. Προτείνω λοιπόν πρώτα να κάνω μια καταγραφή

- των απόψεών σας πάνω στα προβλήματα που προ- ανέφερα (κώδικα των γραμματοσειρών, μεταγλώτ- τιση τής περισπωμένης, των διαλυτικῶν και τής ύπογεγραμμένης),
- του τι έχει γίνει κιόλας, που να μπορούσε να χρησι- μοποιηθεί,
- του κατά πόσον θέλετε, αλλά και μπορείτε να βοηθήσετε.

Επίσης προτείνω την ίδρυση μιᾶς ελληνικής ομάδας φίλων³ του T_EX, είτε σάν ύποομάδας του T_EX Users Group, είτε ανεξάρτητης (τò όνομα θά μπορούσε να είναι ΕΦΤ_EX: "Ελληνες Φίλοι του T_EX ἢ κάτι άλλο). Αυτό θά προωθούσε την συνεργασία⁴ και θά διευκόλυνε την πληροφόρηση των όλοένα και πιο πολλῶν -Ελλήνων και ξένων- που θέλουν να γράφουν στα Έλληνικά με τó T_EX.

Σᾶς προσκαλῶ λοιπόν να δώσετε δυναμικά τó παρὸν και να μού μεταφέρετε τις απόψεις, ιδέες, και προπαντός τις προτάσεις σας. Επίσης προβλέπονται συζητήσεις πάνω σ' αυτό τó θέμα στα Συνέδρια του T_EX που θά γίνουν στο Dedham (HΠΑ) στις 15 με 18 Ἰουλίου, στο Παρίσι στις 23 με 26 Σεπτεμβρίου και στο Ἀμβούργο (κατά πάσα πιθανότητα) κάποτε τó φθινόπωρο του 1991.

6 And a motto

Following the tradition of the *The T_EXbook* of plac- ing the motto at the end, here is mine:⁵

Που πᾶς γυμνός καλὲ μου Μιγκέλ;
γέμισε ὁ τόπος με κηλίδες ἄπουσίας.

Τις τεφροδόχους κλείσαν σὲ κρυφὴ σηηλιά
και μᾶς πουλᾶνε ἐνέσεις εὐθυμίας.

(Ἄλκαϊος)

◊ Yannis Haralambous
101/11, rue Breughel
59650 Villeneuve d'Ascq
France
FAX: +33 20 91 05 64
Bitnet: yannis@frcit181

³ εἶσαι φίλος; γίνε μέλος!...

⁴ ἦ θά τὴν χάλαγε τελείως...

⁵ sorry for the lack of greek sans-serif quotation fonts, it's one of the things which have to be done.

7 Appendix: The font table

| | '0 | '1 | '2 | '3 | '4 | '5 | '6 | '7 | |
|------|----|----|----|----|----|----|----|----|-----|
| '02x | | | [|] | { | } | < | > | "1x |
| '03x | ˆ | ˜ | ˘ | ˙ | ˚ | ˛ | ˜ | ˚ | |
| '04x | | ! | " | # | \$ | % | & | ' | "2x |
| '05x | (|) | * | . | , | - | . | / | |
| '06x | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | "3x |
| '07x | 8 | 9 | : | . | ' | = | ' | ; | |
| '10x | @ | A | B | Σ | Δ | E | Φ | Γ | "4x |
| '11x | H | I | Θ | K | Λ | M | N | O | |
| '12x | Π | X | P | Σ | T | Y | F | Ω | "5x |
| '13x | Ξ | Ψ | Z | [| § |] | Δ | Ϛ | |
| '14x | ` | α | β | ς | δ | ε | φ | γ | "6x |
| '15x | η | ι | θ | κ | λ | μ | ν | ο | |
| '16x | π | χ | ρ | σ | τ | υ | φ | ω | "7x |
| '17x | ξ | ψ | ζ | « | ' | » | - | — | |
| '20x | ά | ά | ά | ά | ά | ά | ά | ά | "8x |
| '21x | ά | ά | ά | ά | ά | ά | ά | ά | |
| '22x | ά | ά | ά | λ | ά | ά | ά | ά | "9x |
| '23x | ή | ή | ή | ή | ή | ή | ή | ή | |
| '24x | ή | ή | ή | ή | ή | ή | ή | ή | "Ax |
| '25x | ή | ή | ή | " | ή | ή | ή | - | |
| '26x | ώ | ώ | ώ | ώ | ώ | ώ | ώ | ώ | "Bx |
| '27x | ώ | ώ | ώ | ώ | ώ | ώ | ώ | ώ | |
| '30x | ώ | ώ | ώ | | ώ | ώ | ώ | | "Cx |
| '31x | ι | ι | ι | ι | ι | ι | ι | ι | |
| '32x | ι | ι | ι | ι | ι | ι | ι | ι | "Dx |
| '33x | ι | ι | ι | | ι | ι | ι | | |
| '34x | ε | ε | ε | ε | ο | ο | ο | ο | "Ex |
| '35x | ε | ε | ε | ε | ο | ο | ο | ο | |
| '36x | ι | ι | ι | ι | ι | ι | ι | ι | "Fx |
| '37x | α | η | φ | ρ | ρ | ! | ~ | - | |
| | "8 | "9 | "A | "B | "C | "D | "E | "F | |

References

- [1] Silvio Levy: Using Greek Fonts with T_EX, *TUGboat*, 9 (1988) 20-24.
- [2] Klaus Thull and Yannis Haralambous: Type- setting Modern Greek with 128 Character Codes, *TUGboat*, 10 (1989) 354-358.
- [3] Marcus N. Tod (ed), A Selection of Greek His- torical Inscriptions, Oxford 1948.
- [4] Χρήστος Κοιλίας, Ὀδηγὸς Χρήστη τοῦ MS-DOS v. 3.3, Ἀθήνα 1988.
- [5] Donald E. Knuth, *The T_EXbook* (9th printing), Reading 1989.