Test font setup for Greek with inputenc/luainputenc

Günter Milde

2020/10/30

Contents

| 1 | LIC | R input 2 |
|---|------|---|
| | 1.1 | Greek alphabet |
| | 1.2 | Diacritics |
| | | 1.2.1 sub-iota |
| | 1.3 | Additional Greek symbols |
| | | 1.3.1 symbols for Greek numbers |
| | | 1.3.2 generic text symbols |
| 2 | LIC | R command input 4 |
| | 2.1 | 0374', ;' 'A ΈΗΤΟΥΩΐ |
| | 2.2 | 0391 ΑΒΓΔΕΖΗΘΙΚΛΜΝΞΟΠΡΣΤΥΦΧΨΩ 4 |
| | 2.3 | 03ΑΑ ΪΫ΄ άξήι΄ |
| | 2.4 | 03Β1 αβγδεζηθικλμνξοπρςστυφχψω |
| | 2.5 | 03 Α ϊϋόύώβθ ϕ π Ω |
| | 2.6 | 1Φ00 ἀδᾶᾶἄἄἄἄἄΑʿAʿAʿAʿAʿAĀĀĀĀĀĀĀĀĀĀĀĀĀĀĀĀĀĀĀĀĀĀĀĀ |
| | 2.7 | 1Φ10 ἐἑἒἒἔἔΈΕΕΕΕ |
| | 2.8 | 1Φ20 ทุ้ทุ้ทุ๊ทุ๊ทุ๊ทุ๊ทุ๊ทุ๊ทุ๊หหหหหหหหหหหหห |
| | 2.9 | 1Φ30 immiTTTTTTT |
| | 2.10 | |
| | 2.11 | $1\Phi50$ ၿပီၿပီၿပီၿပီး ΥΥΥΥΥ |
| | | $1\Phi60$ ἀὰἀἀἄἄά $\Omega\Omega\Omega\Omega\Omega\Omega\Omega\Omega$ |
| | 2.13 | $1\Phi70$ ὰἀὲξὴἡἰἰὸόὺὑὼώ |
| | | $1Φ80$ ἀξάζαζαζαζαζαλ $_1$ 'A $_2$ 'A $_3$ 'A $_4$ 'A $_4$ 'A $_5$ 'A |
| | | $1\Phi90$ ກໍກໍກິກິກິກິກິກິກິກິກິກິກິກິກິກິກິກິກິກ |
| | | 1Φ A0 ψψορφορφορφορφορφορφορφορφορφορφορφορφορφο |
| | 2.17 | 1Φ B0 ἄπαλαάπα ĀĀ'A'AA, $\frac{1}{2}$ |
| | 2.18 | 1Φ 0 ~ π ἡηήῆη ΈΕΗΉΗ $_{\rm r}$ |
| | 2 10 | 1 |
| | 2.20 | 1ΦΕ0 ϋυδοφέρυθταταν |
| | 2 21 | 1ΦΦθ ἀωάδα Ό΄Ο ΌΟΟ Ο΄ 5 |

1 LICR input

The LaTeX internal character representation (LICR) is a verbose, fail-safe 7-bit ASCII encoding that can be used unaltered under both, 8-bit TeX and XeTeX/LuaTeX. Use cases are macro definitions and generated text.

See the source of this document, test-inputenc.tex for the input used in the examples below.

1.1 Greek alphabet

Greek letters via LICR macros:

```
Α Β Γ Δ Ε Ζ Η Θ Ι Κ Λ Μ Ν Ξ Ο Π Ρ Σ Τ Υ Φ Χ Ψ Ω αβγδεζηθικλμνξοπρστυφχψω
```

The small sigma is set with a different glyph if it ends a word:

```
 \sigma \  \, \text{textsigma} \\  \, \varsigma \  \, \text{textfinalsigma} \  \, \text{or} \  \, \text{textvarsigma}
```

The \textautosigma macro, which automatically chooses the glyph according to the position, is not implemented for Unicode fonts.

1.2 Diacritics

Greek accents are tonos = oxia, varia, psili, dasia, dialytika, and perispomeni. Greek diacritics can be input by named macro or symbol macro:

Diacritics as spacing characters:

1.2.1 sub-iota

The sub-iota is input after the base letter.

- \prosgegrammeni sets a spacing GREEK PROSGEGRAMMENI (,): $A_{\rm r}$ $K_{\rm r}$.
- \ypogegrammeni sets a COMBINING GREEK YPOGEGRAMMENI (): $\alpha,\,k,.$

A Greek capital letter followed by COMBINING GREEK YPOGEGRAM-MENI is normalized to the corresponding Greek capital letter WITH [... AND] PROSGEGRAMMENI, if a mapping exists in the Unicode standard: A but K.

```
\begin{array}{ccccc} A_{\scriptscriptstyle 1}A_{\scriptscriptstyle 1}\alpha_{\scriptscriptstyle 1}\alpha_{\scriptscriptstyle 1}&A_{\scriptscriptstyle 1}A_{\scriptscriptstyle 1}\\ \Lambda_{\scriptscriptstyle 1}\Lambda_{\scriptscriptstyle 1}\lambda_{\scriptscriptstyle 1}\lambda_{\scriptscriptstyle 1}&\Lambda_{\scriptscriptstyle 1}\Lambda_{\scriptscriptstyle 1}\\ \alpha_{\scriptscriptstyle 1}\alpha_{\scriptscriptstyle 1}\alpha_{\scriptscriptstyle 2}\alpha_{\scriptscriptstyle 1}&A_{\scriptscriptstyle 1}A_{\scriptscriptstyle 1} \end{array}
```

| text | mathematics | | |
|--------------------|-------------|--------------------|---------------|
| macro | output | macro | output |
| \textpi | π | \pi | π |
| \textvarpi | missing | \varpi | ϖ |
| \textpisymbol | π | | |
| \textrho | ρ | \rho | ρ |
| \textvarrho | missing | \varrho | ϱ |
| \textrhosymbol | ρ | | |
| \texttheta | θ | \theta | θ |
| \textvartheta | missing | \vartheta | ϑ |
| \textthetasymbol | ϑ | | |
| \textepsilon | ε | \epsilon | ϵ |
| \textvarepsilon | missing | $\vert varepsilon$ | ε |
| \textepsilonsymbol | ε | | |
| \textphi | φ | \phi | ϕ |
| \textvarphi | missing | \varphi | φ |
| \textphisymbol | φ | | |
| \textbeta | β | \beta | β |
| \textvarbeta | missing | missing | |
| \textbetasymbol | β | | |
| \textkappa | χ | \kappa | κ |
| \textvarkappa | missing | \varkappa | \varkappa |
| \textkappasymbol | χ | | |
| \textTheta | Θ | \Theta | Θ |
| \textvarTheta | missing | missing | |
| \textThetasymbol | Θ | | |

Table 1: Macros for Greek symbol variants

1.3 Additional Greek symbols

1.3.1 symbols for Greek numbers

- 4 textkoppa
- ч textКoppa
- $m \ref{eq}$ textqoppa (archaic koppa)
- 9 textQoppa (archaic Koppa)
- τ textstigma
- СТ textStigma (Sigma-Tau-Ligature in CB-fonts)¹
- \eth textsampi
- λ textSampi
- f textdigamma
- F textDigamma
- $^{\prime}$ textdexiakeraia
- , textaristerikeraia

 $[\]overline{^1}$ the name "stigma" originally applied to a medieval sigma-tau ligature, whose shape was confusingly similar to the cursive digamma

1.3.2 generic text symbols

LICR macros for some symbols from the 8-bit font encoding LGR that are not confined to Greek but not defined in tuenc.def [2018/08/11 v2.0j].

- ; textsemicolon
- μ textmicro
- ə textschwa

The SI unit prefix MICRO SIGN is not upcased with MakeUppercase:

textmu: $\mu \mapsto M$ but textmicro: $\mu \mapsto M$.

2 LICR command input

- 2.1~~0374 ' , $\,$; ' ' 'A 'E'HT'O' $\Upsilon \Omega t$
- 2.2 0391 ABΓΔΕΖΗΘΙΚΛΜΝΞΟΠΡΣΤΥΦΧΨΩ
- 2.3 03ΑΑ ΪΫάξήίΰ
- 2.4 03Β1 αβγδεζηθικλμνξοπροστυφχψω
- 2.5 03 Α ϊϋόύώβθφπΥΥΠτΕρίηλ ΑκρΘε

Currently, there is no hyperref support for LICR input with non-standard accents or combined diacritics characters.

- 2.7 1Φ10 ἐἑἒεμές Έ΄Ε΄Ε Έ΄ΕΕΕ
- 2.8 1 Φ 20 ทุ้ทุ๊ทุ๊ทุ๊ทุ๊ทุ๊ทุ๊ทุ๊ทุ๊หุ๊ห-H-"H"H"H"H"H"H
- $2.9 \quad 1\Phi 30 \quad \text{in the state of } 1\text{ The state o$
- 2.10 1Φ40 ὀὀὂὂὄὄΟ˙Ο˙Ο˙ϨϨΟ¨Ο

- 2.13 1Φ70 ὰάὲξἡἡλίὸόὑύὼώ
- $2.15 \quad 1\Phi 90$ ทุ่ทุ่ทุ้ทุ้ทุ้ทุ้ทุ้ทุ้ทุ้หุ้ห_เ'H_i'H_i''H_i''H_i''H_i''H_i''
- $2.16 \quad 1\Phi A 0 \;\; \dot{\phi} \dot{\phi} \dot{\ddot{\phi}} \ddot{\ddot{\phi}} \ddot{\ddot{\phi}} \ddot{\ddot{\phi}} \ddot{\ddot{\phi}} \ddot{\tilde{\phi}} \tilde{\gamma} \Omega_{\rm r} \Omega_{\rm r}$
- $2.17 \quad 1\Phi B0 \ \breve{\alpha} \bar{\alpha} \dot{\alpha} \dot{\alpha} \ddot{\alpha} \breve{A} \ddot{A} \dot{A} \dot{A} \dot{A}_{i'i'}$
- 2.18 1Φ ິ ິ ກຸ່ກກຸ້ກຸັ E'E'H'HH. ຶ ຶ ິ
- $2.19 \quad 1\Phi\Delta0$ tititi $\overline{1}$ I'I' " $^{\circ}$
- 2.21 1ΦΦ0 ψωψωωω'Ο'Ο'ΩΩΩ'