Test font setup for Greek with inputenc/luainputenc

Günter Milde

2020/10/30

Contents

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

α β γ δ ε ζ η ϑ ι κ λ μ ν ξ ο π ρ σ τ υ φ χ ψ ω

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

σ `textsigma`
ς `textfinalsigma` or `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:

ά́ ή́ ὰ ὴ ΐ ῒ ΰ ῢ ᾶ ῆ̃ ἄ ἤ ὰ ἤ ᾶ ἤ

Diacritics as spacing characters:
with empty argument: ´ ´ ` ̓ ̔ ¨ ~ ι ̲ ̆ ̊ ̋ ̌ ̑ ̒ ̾
with space as argument: ´ ´ ` ̓ ̔ ¨ ~ ι ̲ ̆ ̊ ̋ ̌ ̑ ̒ ̾
with protected space: ´ ´ ` ̓ ̔ ¨ ~ ι ̲ ̆ ̊ ̋ ̌ ̑ ̒ ̾
with nobreakspace: ´ ´ ` ̓ ̔ ¨ ~ ι ̲ ̆ ̊ ̋ ̌ ̑ ̒ ̾

1.2.1 sub-iota

The sub-iota is input after the base letter.

- `\prosgegrammeni` sets a spacing GREEK PROSGEGRAMMENI (ͅ): Aͅ Kͅ.
- `\ypogegrammeni` sets a COMBINING GREEK YPOGEGRAMMENI (̥): α̥ k̥.

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

ḀAͅαḀ AͅA
Λ̥ΛͅλΛ̥ ΛͅΛ
ᾳᾳαḀ AͅA

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	\varepsilon	ε
\textepsilonsymbol	ε		
\textphi	φ	\phi	ϕ
\textvarphi	missing	\varphi	φ
\textphisymbol	φ		
\textbeta	β	\beta	β
\textvarbeta	missing	*missing*	
\textbetasymbol	β		
\textkappa	ϰ	\kappa	κ
\textvarkappa	missing	\varkappa	$\varkappa$
\textkappasymbol	ϰ		
\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

ϟ textkoppa
Ϟ textKoppa
ϙ textqoppa (archaic koppa)
Ϙ textQoppa (archaic Koppa)
ϛ textstigma
Ϛ textStigma (Sigma-Tau-Ligature in CB-fonts)[1]
ϡ textsampi
Ϡ textSampi
ϝ textdigamma
Ϝ textDigamma
ʹ textdexiakeraia
͵ textaristerikeraia

[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: µ ↦ M but textmicro: µ ↦ M.

2 LICR command input

2.1 0374 ʹ ͵ ; ΄ ΅ Ά · Έ Ή Ί Ό Ύ Ώ ΐ

2.2 0391 ΑΒΓΔΕΖΗΘΙΚΛΜΝΞΟΠΡΣΤΥΦΧΨΩ

2.3 03AA Ϊ Ϋ ά έ ή ί ύ

2.4 03B1 αβγδεζηθικλμνξοπρςστυφχψω

2.5 03ˆA ϊϋόύώβϑφϖϚϜϞϺϽϾϿЌЍЎѐёђ

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

2.6 1Φ00 ȁȃȅȇȉȋȍȏȁ̂ȃ̂ȅ̂ȇ̂ȉ̂ȋ̂ȍ̂ȏ̂A

2.7 1Φ10 ȅȇȅ̂ȇ̂ȇ̌E E̋E̋E̋E̋E

2.8 1Φ20 ȟȟȟ̂ȟ̂ȟ̌ȟ̌H H̋H̋H̋H̋H̋H̋H

2.9 1Φ30 ıȉȋ̂ȋ̂ȋ̌ȋ̌I I̋I̋I̋I̋I̋I

2.10 1Φ40 ȍȍ̂ȏ̂ȏ̌O ŐŐŐŐO

2.11 1Φ50 ȕȕ̂ȗ̂ȗ̌ȗ̌Υ Υ̋Υ̋Υ̋Υ

2.12 1Φ60 ѡѡ̂ѡ̂ѡ̌ѡ̌ѡ̌Ω ΩΩΩΩΩΩΩ

2.13 1Φ70 ȁȃȅȇȟıȍȕȗѡ

2.14 1Φ80 ạ̏ạ̑ẹ̏ẹ̑ị̏ị̑ọ̏ọ̑A I̋A I̋A I̋A I̋A I̋A I

2.15 1Φ90 ḥ̌ḥ̌ḥ̌ḥ̌ḥ̌ḥ̌ḥ̌H I̋H I̋H I̋H I̋H I̋H I

2.16 1ΦA0 ọ̏ọ̏ọ̑ọ̑ọ̑ọ̑ọ̑Ω I̋Ω I̋Ω I̋Ω I̋Ω I̋Ω I

2.17 1ΦB0 ăāạ̏ạ̑ȃ̂ăĂĀȁ́AA I̋ I̋

2.18 1Φ̋0 ̃ ̋ ȟ̃ȟ̃ȟ̃ȟ̃ÉE̋H́HH I̋ ̋ ̃

2.19 1ΦΔ0 ŧŧŧĬĪ̀ÍI̋ ̋ ̃

2.20 1ΦE0 ŏŏȍọ̏ọ̑ŏŪŏŸŪ̀ΎP̋ ́ ̀

2.21 1ΦΦ0 ọ̏ọ̏ȍ̂ѡ̃ȍ̀ÓŐΩΩI ́ ̀