

Unicode Alphabets

Mikkel Eide Eriksen

July 16, 2019

1 Preface

While Unicode supports the vast majority of use cases, there are certain specialized niches which require characters and glyphs not (yet) represented in the standard.

Thus the Private Use Area (PUA) at code points E000–F8FF, which enables third parties to define arbitrary character sets.

This package allows configuring a number of macros to enter characters from the PUA by name or code point.

2 Setup

The package is configured in the following manner:

```
\usepackage[options]{unicode-alphabets}
```

Where **options** must be one or more of the following character sets. See references for further detail on each, as well as usable fonts.

CYFI Early Cyrillic glyphs[1].

LINCUA Shortcut for enabling the character sets CYFI, MUFI, and TITUS[5].

MUFI The Medieval Unicode Font Initiative[3].

SIL SIL International[2].

TITUS Thesaurus Indogermanischer Text- und Sprachmaterialien[6].

UNZ Normung von Sonderzeichen[4].

Figure 1: Character sets

There is no default, since future versions of this package may supply more character sets that are incompatible with the above.

Additionally, one may configure different fonts for each character set, as in the following example:

```
\usepackage[
    MUI,
    TITUS,
    fonts={
        MUI=\mufifont,
        TITUS=\titusfont,
    },
]{unicode-alphabets}
```

Figure 2: Example setup

If no fonts are configured, the document font will be used (note that this may give in undesirable results, as few fonts support multiple character sets¹).

Finally, the MUI character set has defined entity names, which result in the creation of macros for each character (see `msignfour` in the following example). These can be suppressed with the `disableentitymacros` option.

3 Usage

Each set defines a macro in the following manner. Let’s use MUI as an example.

By default, a macro with the lower-case name of the character set is defined: `\mufi{}`. It can then be used to display characters from the given set:

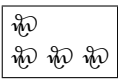
	¹ <code>\mufi{FLOURISHED SMALL LETTER M SIGN}\\</code>
	² <code>\msignfour{}</code>
	³ <code>\mufi{msignfour}</code>
	⁴ <code>\mufi{F2F3}</code>

Figure 3: Example usage

Additionally, starred versions of each macro are defined, which suppress using the configured font, falling back to the document font.

References

- [1] Victor Baranov et al. “Proposal for a unified encoding of Early Cyrillic glyphs in the Unicode Private Use Area”. In: *Scripta & e-scripta* 8 (2010), pp. 9–26.

¹I believe Andreas Stötzner’s *Andron Mega* does, albeit I haven’t tried it as it is somewhat expensive.

- [2] Lorna Evans. *SIL PUA 9.0*. July 5, 2016. URL: <https://scripts.sil.org/SILPUAassignments>.
- [3] *MUFI: The Medieval Unicode Font Initiative*. July 12, 2019. URL: <https://mufi.info/> (visited on 07/12/2019).
- [4] Bund für deutsche Schrift und Sprache e. V. *Normung von Sonderzeichen*. July 12, 2019. URL: <https://www.bfds.de/der-bund/normung-von-sonderzeichen-unicode/> (visited on 07/12/2019).
- [5] Andreas Stötzner. “LINCUA: A Unicode-PUA harmonization scheme”. In: (2012).
- [6] *Thesaurus Indogermanischer Text- und Sprachmaterialien*. July 10, 2017. URL: <http://titus.uni-frankfurt.de/indexe.htm> (visited on 07/12/2019).