The NewComputerModern font family

Antonis Tsolomitis

Abstract

We present the NewComputerModern font family (NewCM for short), an extension of the default Computer Modern fonts at (currently) 10pt for Unicode TeX engines.

1   Introduction

Back in the mid-1980s you could hardly typeset anything in a non-Latin language. Thanks to the work of many people and the Babel package, TeX was extended so it could typeset Cyrillic, Greek and many other scripts and languages that needed different fonts and typesetting rules. For Greek in the early nineties it was difficult just to type in Greek (it was a pain merely to set up a Greek keyboard under GNU/Linux even in the late nineties).

Things progressed over the years and the TeX world gained the Unicode engines, promising to solve access to thousands of glyphs outside the Latin blocks. However, we are in 2021 and still: You install a TeX system (any), you start a simple document, you run `xelatex` or `lualatex`, you fire up your PDF reader . . . And you realize that the old frustration is still there. There is nothing in the PDF except Latin glyphs.

After all these years I find this not satisfactory, to say the least. There is no fallback mechanism (as there is for Office apps) and the default fonts contain only Latin glyphs (plus math). So, the user must make choices. Select a Cyrillic font, a Greek font, a Hebrew one etc. But the user must know how to do it, and it is not trivial because s/he must find fonts that match in style and weight; and if math is needed the task is even more difficult.

NewCM was born with these thoughts in mind. If TeX is Unicode-enabled, where is its default Unicode font? Should a default font support all the planet's languages? Most probably not. But why not support at least the large communities of spoken languages whose members have a proven interest for their language being supported?

NewComputerModern is an attempt to expand the Latin Modern (`lm`) fonts to common non-Latin scripts, while keeping metric compatibility with `lm`.

2   The `fontsetup` package

So if one (i.e., me) merges in the glyphs from existing fonts is he done? Not at all. Why? Because your new font is not the default, and people will not easily switch away from the system's default. Now we are at the same point. It is not easy to properly set up fonts that support math too. What is needed is a simple way to do that. A "one liner". This need gave birth to the `fontsetup` package: An easy way to properly load fonts and matching accompanying mathematics. The line

`\usepackage[olddefault]{fontsetup}`

will load all you need, from the font side, to typeset in all languages covered by the `lm` fonts plus Cyrillic, Greek, Hebrew and Cherokee. All with matching math, sans serif and typewriter fonts. It will also provide access to several other Unicode blocks such as Braille patterns and more, to be discussed below. Of course hyphenation and label strings must be loaded for the main language (for example, the `xgreek` package for Greek).

But why "olddefault"? What is "default" and what are other options?

3   A Book weight for NewCM

An old problem of ComputerModern is the fact that it is a light font. And this problem is the same with Claudio Beccari's Greek which was added in NewCM, and Cyrillic from `cmu`, also added, because a goal for those fonts was to match the weight of Knuth's original fonts.

To design heavier fonts from scratch would be a huge undertaking, given the thousands of glyphs involved. To do it automatically with a font editor is known to create problems with the glyphs. But there is a catch with that last sentence. We do not want to create bold versions with the font editor. That would be bad. We just need a little bit of heaviness to be added so that the fonts look good at both low and high resolutions; the existing bold can stay untouched.

So this gave birth to NewCM Book weight. Was it as simple as it sounds? No, because we need math too. So the Book weight math font must carry all the information needed to properly typeset math, and this is many weeks of work for just one font. But in the end, there we have it: this is the "default" version of the `fontsetup` package. That is the Book weight for NewCM, supporting all languages the project supports, and all the features that will be presented below. The amount of added "boldness" is such that it matches in color with the GFSDidot family, which I have used in my books in the past and which looks good at both high resolution printing and low resolution screens.

The `fontsetup` package has more options to easily load many other font families with matching math. Please check its documentation (`ctan.org/pkg/fontsetup`).

4 Bold Sans

The BoldSans in the `lm` and `cmu` fonts is merely a stroke-extension of the Sans, with rounded corners. NewCM provides a true BoldSans:

LM	NewCM
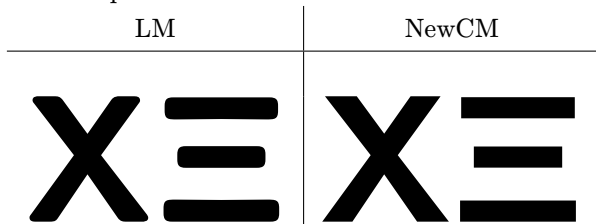
This currently covers Latin and Greek, but soon it will cover Cyrillic too.

5 New languages added

Cyrillic has been added from the `cmu` package and Greek monotonic and polytonic from Claudio Beccari's fonts.

Greek:

Θεώρημα 5.1 (Πυθαγόρειον) *Ἐν τοῖς ὀρθογωνίοις τριγώνοις τὸ ἀπὸ τῆς τὴν ὀρθὴν γωνίαν ὑποτεινούσης πλευρᾶς τετράγωνον ἴσον ἐστὶ τοῖς ἀπὸ τῶν τὴν ὀρθὴν γωνίαν περιεχουσῶν πλευρῶν τετραγώνοις.*

Russian:

Я помню чудное мгновенье:
Передо мной явилась ты,
Как мимолетное виденье,
Как гений чистой красоты.

(Пушкинъ)

Hebrew and Cherokee were designed from scratch:

A few letters from Hebrew:

צלשׁשׂוהדגבא

A few letters from Cherokee:

ᏦᎤᏛᏏᏳᏟ ᏦᎤᎶᎦᎩᏍ

Back to Greek, Small Caps is included (in Mono font too) and all polytonic accents of Greek. Ypogegrammeni is the default for all characters including Small Caps and prosgegrammeni is offered as an alternative shape in the `ss01` lookup table:

	ypogegrammeni	prosgegrammeni
regular	ᾌ ᾜ ᾬ ᾼΗῼ	Ἄ, Ἤ, Ὤ, Α,Η,Ω,
sans	ᾌ ᾜ ᾬ ᾼΗῼ	Ἄ, Ἤ, Ὤ, Α,Η,Ω,
mono	ᾌ ᾜ ᾬ ᾼΗῼ	Ἄ,Ἤ,Ὤ, Α,Η,Ω,

The prosgegrammeni alternates can be accessed with commands from the `fontsetup` package, either of:

`\textprosgegrammeni{⟨text⟩}`
`{\prosgegrammeni ⟨text⟩}`

6 More Unicode blocks

Braille, both 6dot (uni2801–uni283F) as well as 8dot (uni2840–uni28FF) patterns are included in two versions. The Regular font provides the characters for sighted persons (such as teachers) so they can easily see which dots are on and which off. The Sans font contains the true Braille characters. I decided to have the sighted version in the Regular font since a blind person does not need the real Braille pattern, as those are produced by embossers. The Braille patterns here are meant for use as fonts to typeset text mainly for sighted persons.

	6dot	8dot
Regular version	⠿⠿⠿⠿⠿	⣿⣿⣿⣿⣿
Sans version	⠿⠿⠿⠿⠿	⣿⣿⣿⣿⣿

IPA symbols are included, and following a suggestion of Huanyu Liu the kerning found in the `tipa` package has been added here and further improved. Moreover the letters eth, eng, beta, theta and chi exist in IPA-style in the fonts and are accessible in the ss05 lookup table since they have a different design than the Latin and Greek letters. You can access this lookup table using the `\textipa` and `\textsansipa` commands of the `fontsetup` package.

	Non-IPA	IPA
Regular	ð ŋ β ϑ χ	ð ŋ β θ χ
Sans	ð ŋ β θ χ	ð ŋ β θ χ

7 Ligatures and stylistic alternatives in Latin

The Serif font includes additional ligatures fb ffb ffh ffj ffk fft fh fj ft fk, and the same with long-s instead of f in the default liga table (in addition to the default fi fl ffi ffl ff). It also includes an alternative k (in the cv01 table) and sp ch ck ct st in the dlig table. Finally it also includes "end" versions for the letters a, e, m, n and r in the cv02 table.

Regular	k	a e m n r	sp ch ck ct st
cv01	k		
cv02		a, e, m, n, r	
dlig			sp ch ck ct st

8 Archaic Greek writing for scholars and others

The Sans Serif Regular font provides access to 6th century BCE and 4th century BCE Greek capitals in ss04 and ss03 lookups. The `fontsetup` package provides commands such as:

`\textivbce{}, \ivbce, \textvibce{}, \vibce`.

6th century BCE:

ᛗΘΔΕΙΣ ΑΓΕΩᛗΕΤΡΒΤΟΣ ΕΙΣΙΤΩ

4th century BCE:

ΜΗΔΕΙϟ ΑΓΕΩΜΕΤΡΗΤΟϟ ΕΙϟΙΤΩ

Moreover, all fonts (except Mono and Math) support Ancient Greek Numerals (the full Unicode block of Greek digits u10140–u1018E is supported), with most symbols designed from scratch (those that did not exist in C. Beccari's original fonts). A few of the new symbols:

$$\text{Ⱶ} \text{Ⱶ}^{\prime} \text{Ꭓ} \text{Ⱶ}^{\prime} \text{Ⱶ}^{\prime} \overset{\text{M}}{\text{Δ}} \text{Ϙ}^{\prime}$$

The four numerals that already existed in this range (that is u10144–u10147) in Beccari's fonts have been altered to a new design matching the style of cm but also provide some Ancient Greek flair. The new designs in Serifed and SansSerifed are:

$$\boxed{\Delta} \boxed{H}^{\prime} \boxed{X}^{\prime} \boxed{M}^{\prime} \qquad \boxed{\Delta} \boxed{H} \boxed{X} \boxed{M}$$

The `fontsetup` package provides commands for all of the above symbols. The commands follow the Unicode name of each slot (minus the "Greek Acrophonic"). So the Unicode slot u1014F named "Greek Acrophonic Attic Five Staters" can be accessed with the command `\atticfivestaters` and gives $\boxed{\Sigma}^{\prime}$; and the slot u10182 named "Greek Kyathos Base Sign" can be accessed with the command `\greekkyathosbasesign` and gives Ⱪ.

9   Old Italic

The fonts also fully support the Old Italic Unicode block (u10300–u1032F) in the Sans font. For example, the slots u10307, u10310, u10312, u10314, u1031F and u1032F are 𐌇𐌓𐌒𐌔𐌟𐌯.

10   Unicode Math coverage and options

NewCM provides full Unicode math support, that is all mathematics Unicode slots presented in `unicode.org/charts/` in both Math weights, Regular and Book. These blocks are:

Mathematical Symbols
  Arrows (uni2190–uni21FF)
  Supplemental Arrows-A (uni27F0–uni27FF)
  Supplemental Arrows-B (uni2900–uni297F)
  Supplemental Arrows-C (u1F800–u1F8FF)
  Additional Arrows (uni2B00–uni2BFF)
  Miscellaneous Symbols and Arrows
    (uni2B00–uni2BFF)

Mathematical Alphanumeric Symbols
    (u1D400–u1D7FF)
  Arabic Mathematical Alphabetic Symbols
    (u1EE00–u1EEFF)
  Letterlike Symbols (uni2100–uni214F)

Mathematical Operators
    (uni2200–uni22FF)
  Basic operators: Plus, Factorial
    (uni0000–uni007F)
  Division, Multiplication (uni0080–uni00FF)
  Supplemental Mathematical Operators
    (uni2A00–uni2AFF)
  Miscellaneous Mathematical Symbols-A
    (uni27C0–uni27EF)
  Miscellaneous Mathematical Symbols-B
    (uni2980–uni29FF)
  Floors and Ceilings (uni2308–uni230B)
  Invisible Operators (uni2061–uni2064)

Geometric Shapes (uni25A0–uni25FF)
  Additional Shapes (uni2B00–uni2BFF)
  Box Drawing (uni2500–uni257F)
  Block Elements (uni2580–uni259F)
  Geometric Shapes Extended (u1F780–u1F7FF)

Unfortunately, the `unicode-math` package does not currently provide commands for the hundreds of extra glyphs that have been added in order to fully cover the above Unicode ranges. The user can consult the Unicode charts at the above link and access the required glyph with `\char"#` where # is the Unicode number of the slot the glyph belongs to. For example, `\char"2BDA` will give the Hygeia symbol (uni2BDA) the Rod of Asclepius as shown above.

10.1   Upright and extensible integrals

The Math fonts (both Regular and Book weights) include upright integrals in the ss02 StylisticSet.

Normal integrals	Upright integrals
$\int \iiint \int \oint_a^b$	$\int \iiint \int \oint_a^b$

Use with
```
\setmathfont[StylisticSet=2]
                    {NewCMMath-Book.otf}
```
or
```
\setmathfont[StylisticSet=2]
                  {NewCMMath-Regular.otf}
```
or use the `upint` option of the `fontsetup` package; for the Book weight:

```
\usepackage[upint,default]{fontsetup}
```
or for the regular weight:

```
\usepackage[upint,olddefault]{fontsetup}
```

Moreover, extensible integrals are supported by the fonts but *not* by the Unicode TeX engines. The following code is a trick so that extensible integrals

can be constructed using LuaLaTeX. The result is shown at the end of the article. What the code below does is define the slot uni222B (Integral) as a delimiter. Then, this is extended as a delimiter with the mechanism that the font provides.

```
\documentclass{article}
\usepackage[default]{fontsetup}
\begin{document}
\[
\Uleft\Udelimiter 0 0 "222B
\begin{pmatrix}
  1
\end{pmatrix}
\Uright.
\rightarrow
\Uleft\Udelimiter 0 0 "222B
\begin{pmatrix}
  1\\2
\end{pmatrix}
\Uright.
\rightarrow
\Uleft\Udelimiter 0 0 "222B
\begin{pmatrix}
  1\\2\\3\\4
\end{pmatrix}
\Uright.
\rightarrow
\Uleft\Udelimiter 0 0 "222B
\begin{pmatrix}
  1\\2\\3\\4\\5\\6
\end{pmatrix}
\Uright.
\]
\end{document}
```

10.2 Non-Unicode symbols

It seems that Unicode has forgotten to include slots for the negation of uniform convergence. The fonts include two extra slots for $\nrightrightarrows$ and $\nleftleftarrows$ that can be accessed in math mode with the commands `\nrightrightarrows` and `\nleftleftarrows` of the `fontsetup` package.

Unicode seems to have also forgotten to include MathSansGreek. These are included in the Math fonts and they are and accessible with commands such as `\msansAlpha` or `\mitsansAlpha`.

$$\mathsf{AB\gamma\delta} \quad \mathit{AB\gamma\delta}$$

11 Future work

The immediate plans for NewCM are to provide the fonts at 8pt and to provide support for accent (diacritics) stacking. Work on the 8pt version has already begun. The 8pt size of the `lm` fonts looks lighter at 8pt than the 10pt font. This should not happen in

my opinion, so the 8pt design (set at 8pt) will match the weight of the 10pt design in the latin glyphs too. Nonetheless, metric compatibility will be preserved.

12 Thanks

There are many people I would like to thank who have reported bugs of the fonts. Special thanks go to Karl Berry, Claudio Beccari, David Carlisle, Robert Alessi, Huanyu Liu and Manuel Boni for supporting this project with their help and suggestions.

⬦ Antonis Tsolomitis
Department of Mathematics
University of the Aegean
Karlovassi, 832 00 Samos
Greece
atsol (at) aegean dot gr
https://myria.math.aegean.gr/~atsol/

$$\int\begin{pmatrix}1\\2\end{pmatrix} \rightarrow \int\begin{pmatrix}1\\2\\3\\4\end{pmatrix} \rightarrow \int\begin{pmatrix}1\\2\\3\\4\\5\\6\end{pmatrix} \rightarrow \int\begin{pmatrix}1\\2\\3\\4\\5\\6\\7\\8\\9\\10\\11\\12\end{pmatrix} \rightarrow \int\begin{pmatrix}1\\2\\3\\4\\5\\6\\7\\8\\9\\10\\11\\12\\2\\3\\4\\5\\6\\7\\8\\9\\10\\11\\12\end{pmatrix}$$