

Comments and suggestions about the Latin Modern fonts

Karel Píška

Institute of Physics of the ASCR, v.v.i.
CZ-182 21 Prague, Czech Republic
piska (at) fzu dot cz

Abstract

This contribution:

- describes a process of verification of the Latin Modern fonts and lists selected aspects (typographic and technical) tested during this activity;
- summarizes the results of checking and comparisons, mostly for version 0.99.3 (2005) and 1.00 (2006), while information about the current version 1.010/x (2007) is limited;
- documents bugs and their correction, often in a visual form;
- remarks on the crucial changes in the recent versions (2006 and 2007);
- compares individual glyph shapes and finds differences in the last Type 1 and OpenType releases;
- compares metrics, especially character widths and kerning pairs, between the T1 (EC) and CS (CM) encoded subsets, and between ver. 0.99.3 and ver. 1.00, analyzing compatibility and listing the differences;
- studies accents and accented letters, i.e. accent shapes and their positioning, mainly for the accented characters common to CS (Czech and Slovak font collection) and T1 (EC);
- discusses problems with accurate and optimal outline representations, as the Latin Modern font family is a descendant of Computer Modern (designed in METAFONT) and converted into an outline approximation with cubic curves;
- comments on hinting strategies, particularly for accents and slanted fonts.

1 Introduction

The purpose of the activities described in this article was to improve the Latin Modern font package, helping to change unintentional features and fixing mistakes, with a special emphasis on preserving compatibility in typesetting Czech and Slovak documents with L^AT_EX using the Latin Modern outline fonts to substitute for older extensions of the Computer Modern typefaces.

The current version of the paper is a partial excerpt of the technical documentation, a collection of visual documents, mostly in PDF. All are still undergoing revision, trying to reflect the actual updates of the LM package, which is possible only with some time delay. I think it is not unreasonable to include tests of the previous release of LM, performed and stored in “my” archive last year.

I collected my comments, reports and proposals on my web site:

<http://www-hep.fzu.cz/~piska/lm2005.html>
<http://www-hep.fzu.cz/~piska/lm2006.html>

I have specially added some longer tables (OTF glyph names, kerning pairs) to be available in a printed form for investigation and discussion.

For user information about LM, we recommend “An exploration of the Latin Modern fonts”, an article written by Will Robertson [13].

2 Global remarks about font verification

The multistep checking process of an upcoming release can be divided into several stages.

1. We start the first stage from a survey to determine significant changes in comparison with the previous version or releases, important differences, extensions or exclusions. And then we generate “primary” proof printings, for example, the complete proof sheet pages for all fonts and all glyphs present in the font family; all ligature and kerning pairs for all fonts and for selected encodings. After some adaptation of the programs involved, this output can usually be produced automatically with only minimal assistance.

ver.	date	comments about crucial changes
0.99.3	28 Oct 2005	
1.00	13 Apr 2006	metrics completely recalculated
1.010	16 Jan 2007	glyph names changed in OpenType
1.010a	23 Feb 2007	
1.010x	28 Feb 2007	

Table 1: Overview of recent releases of LM.

Quite a lot of disk space to store many huge files is needed. A rich archive of PDF documents is created and ready for human visual scanning with Acrobat Reader to search glyph images for evident errors, to study “usual” (from previously known opinions) weak points, zooming the typical parts of glyphs where artifacts may occur; to search for unsuitable kerning pairs or improper kerns; and to study consistency or compatibility of associated font elements.

The test printings covering the entire glyph repertoire and all kerning pairs for T1 (EC) encoding were generated in 2006 for LM 1.00.

2. In the next stage we compare the actual font release with the preceding release(s), e.g. we find differences between both instances of all glyphs in their outline curve representation. We also compare the major metric data important for typesetting with *TEX*, especially the character widths, kerning and ligature pairs.
3. The subsequent stage depends on the results of the previous analysis. We select, study, select and study again chosen features, potential mistakes and strange events in detail to detect and localize bugs, and give a classification or conclusion, to prepare a report in textual and/or visual form.
4. A comparison with related and other relevant fonts may be important to confirm identity and compatibility or to find differences, intentional or unintentional.
5. Finally, we perform an overall evaluation of the available data, summarize the results, and compile a document with visual demonstration and written comments, conclusions, reports, suggestions and recommendations.

3 Developments and changes in Latin Modern

To begin with, Table 1 shows a concise summary of recent releases of LM.

Text fonts					
	#n	ver. 1.00 #g0	ver. 1.010[x] #k0	ver. 1.010[x] #g1	ver. 1.010[x] #k1
b	10x	701	9399	742	9344
bi	1x	701	12134	742	11963
sc	2x	692	8742	735	8676
r	21x	701	9413	742	9358
ri	6x	701	12148	742	11977
ss	14x	701	8732	742	8677
sq	4x	704	8732	745	8677
tt	14x	662	0	703	0
tc	2x	659	0	702	0

Subtotal text fonts:
 72 49914 551345 52874 547321

Mathematical fonts					
	sy	9x	26	132	26
ex	1x	130	0	130	0
mi	10x	130	164	130	164

Subtotal math fonts:
 20 2618 1874 2618 1874

Total LM fonts:	92	52532	553219	55492	549195
#n	number of fonts with the same counts				
#g	number of glyphs				
#k	number of kerning pairs				

Table 2: Numbers of fonts and glyphs.

Table 2 presents the numbers of glyphs and kerning pairs in LM ver. 1.00 (#g0 and #k0) and the current ver. 1.010[x] (#g1 and #k1), subtotal counts for 72 text fonts and 20 mathematical fonts and the total sums, where:

b = (Roman) Bold, Demi; bi = (Roman) BoldItalic; sc = SmallCaps, r = (Roman) Regular, Oblique, TypewriterVarWd; ri = (Roman) Italic, Unslanted; ss = Sans; sq = SansQuotation; tc = TypewriterCaps; tt = Typewriter; sy = MathSymbols; ex = MathExtension; mi = MathItalic.

One very important modification in LM ver. 1.010 is the change of glyph names in OpenType. In Table 3, Unicode code points are listed together with the OpenType (Unicode) glyph names (in the second column) and PostScript (Type 1) names (column 3). This information has been added to give a quick explanation of the “unintelligible” OTF glyph names in a short and compressed form.

4 Comparison of releases

Because the PostScript and Type 1 glyph names are not identical starting with the version released in 2007 it is impossible to compare the glyphs by name

F6C9 acute.cap Acute	EA14 space_uni0309 hookabove
2116 afii61352 numero	EA13 space_uni0309.cap Hookabove
2217 asterisk.math asteriskmath	EA17 space_uni030A_uni0301 ringacute
EFEF breve.cap Breve	EA16 space_uni030A_uni0301.cap Ringacute
F6CA caron.cap Caron	F6D3 space_uni030F dblgrave
EFF7 circumflex.cap Circumflex	F6D6 space_uni030F.cap dblGrave
- copyright.var varcopyright	EA07 space_uni0311 breveinverted
F6CB dieresis.cap Dieresis	EA06 space_uni0311.cap Breveinverted
EFED dotaccent.cap Dotaccent	EB19 space_uni0323 dotbelow
- dotaccent.var vardotaccent	EA08 space_uni032F breveinvertedlow
1E0C D_uni0323 Ddotbelow	EB69 space_uni0330 tildedown
1E0D d_uni0323 ddotbelow	EB09 star.alt born
FB00 f_f ff	EB2A S_S Germandbls
FB03 f_f_i ffi	EFF5 tilde.cap Tilde
FB04 f_f_l ffl	1E6C T_uni0323 Tdotbelow
FB01 f_i fi	1E6D t_uni0323 tdotbelow
FB02 f_l fl	00A0 uni00A0 nbspace
F6CE grave.cap Grave	00AD uni00AD sphyphen
F6CF hungarumlaut.cap Hungarumlaut	0218 uni0218 Scommaaccent
1E24 H_uni0323 Hdotbelow	0219 uni0219 scommaaccent
1E25 h_uni0323 hdotbelow	021A uni021A Tcommaaccent
- I.var varI	021B uni021B tcommaaccent
- Iogonek.var varIogonek	0300 uni0300 gravecomb
0132 I_J IJ	E300 uni0300.cap Gravecomb
- I_J.var varIJ	0301 uni0301 acutecomb
0133 i_j ij	E301 uni0301.cap Acutecomb
F6BE j.dotless dotlessj	0302 uni0302 circumflexcomb
1E36 L_uni0323 Ldotbelow	E302 uni0302.cap Circumflexcomb
1E37 l_uni0323 ldotbelow	0303 uni0303 tildetildecomb
1E39 l_uni0323_uni0304 ldotbelowmacron	E303 uni0303.cap Tildetildecomb
1E38 L_uni0323_uni0304.cap Ldotbelowmacron	0304 uni0304 macroncomb
F6D0 macron.cap Macron	E304 uni0304.cap Macroncomb
1E42 M_uni0323 Mdotbelow	0306 uni0306 brevecomb
1E43 m_uni0323 mdotbelow	E306 uni0306.cap Brevecomb
1E45 n_uni0307 ndotaccent	0307 uni0307 dotaccentcomb
1E44 N_uni0307.cap Ndotaccent	E307 uni0307.cap Dotaccentcomb
1E46 N_uni0323 Ndotbelow	0308 uni0308 dieresiscomb
1E47 n_uni0323 ndotbelow	E308 uni0308.cap Dieresiscomb
- registered.var varregistered	0309 uni0309 hookabovecomb
EFF3 ring.cap Ring	E309 uni0309.cap Hookabovecomb
1E59 r_uni0307 rdotaccent	030A uni030A ringcomb
1E58 R_uni0307.cap Rdotaccent	E30A uni030A.cap Ringcomb
1E5A R_uni0323 Rdotbelow	030B uni030B hungarumlautcomb
1E5B r_uni0323 rdotbelow	E30B uni030B.cap Hungarumlautcomb
1E5D r_uni0323_uni0304 rdotbelowmacron	030C uni030C caroncomb
1E5C R_uni0323_uni0304.cap Rdotbelowmacron	E30C uni030C.cap Caroncomb
2423 space.visible visiblespace	030F uni030F dblgravecomb
EA0E space_uni0302_uni0300 circumflexgrave	E30F uni030F.cap Dblgravecomb
EA0D space_uni0302_uni0300.cap Circumflexgrave	0311 uni0311 breveinvertedcomb
EA0C space_uni0302_uni0301 circumflexacute	E311 uni0311.cap Breveinvertedcomb
EA0B space_uni0302_uni0301.cap Circumflexacute	0323 uni0323 dotbelowcomb
EA12 space_uni0302_uni0303 circumflextilde	0326 uni0326 commaaccentcomb
EA11 space_uni0302_uni0303.cap Circumflextilde	032E uni032E brevelowcomb
EA10 space_uni0302_uni0309 circumflexhookabove	032F uni032F breveinvertedlowcomb
EA0F space_uni0302_uni0309.cap Circumflexhookabove	F6DE uni2014.alt1 threequartersemidash
EA03 space_uni0306_uni0300 brevegrave	EB6D uni2014.alt2 twelveudash
EA02 space_uni0306_uni0300.cap Brevegrave	2127 uni2127 mho
EA01 space_uni0306_uni0301 breveacute	2190 uni2190 arrowleft
EA00 space_uni0306_uni0301.cap Breveacute	2191 uni2191 arrowup
EA0A space_uni0306_uni0303 brevetilde	2192 uni2192 arrowright
EA09 space_uni0306_uni0303.cap Brevetilde	2193 uni2193 arrowdown
EA05 space_uni0306_uni0309 brevehookabove	266A uni266A musicalnote
EA04 space_uni0306_uni0309.cap Brevehookabove	

Table 3: OTF and PostScript glyph names.

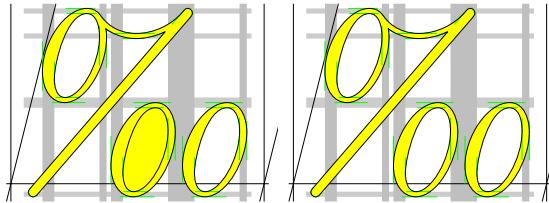


Figure 1: lmri10:perthousand before and after correction of path directions.

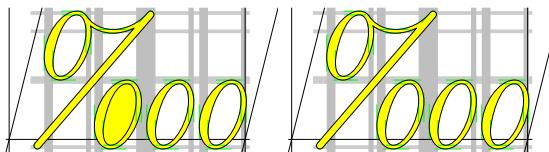


Figure 2: lmri10:permyriad before and after correction of path directions.

without preparation of a new program to do a complete glyph comparison in a different way. The tests performed last year with LM 1.00 may be interesting, but they are out-of-date for this article.

5 Bug reports and corrections

It is important to demonstrate bugs present in the last release and then confirm they have been fixed in the current release. The technical documentation contains a comprehensive list of bugs and other problematic events. Here we will show several examples showing glyph corrections, improvements or other changes (figs. 1 and 2).

Single tests to examine glyphs in the actual fonts can be performed with procedures from my package `tfcpr` [12].

6 Metrics: compatibility and/or quality

One major task, supported by CSTUG, was exploring use of LM for the characters in the common subset of the standard Computer Modern and CS fonts, mainly accented letters belonging to the Czech and Slovak character set.

For all characters from the intersection of the `ec-lm` encoding and CS (which covers all characters from CM) we compared metric data: character widths in the corresponding `tfm` metric files and also with their equivalents in `pfb`, `afm` and `otf`, and kerning pairs in `tfm`.

Generally, good agreement in the glyph widths was found, the differences being negligible (in the last digits), and due to internal numeric representation.

After the complete recalculation of metrics included in LM ver. 1.00 I have found only a couple

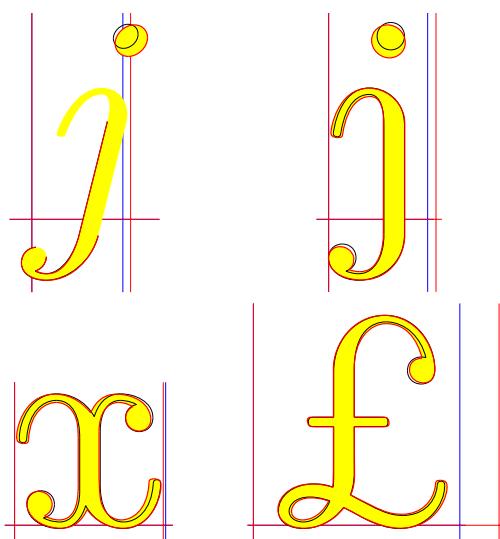


Figure 3: Different widths (and shapes) in lmri10:j and lmu10:j, x, sterling.

of cases of changing character widths: "j" in `lmri*`, and "x" and `sterling` in `lmu10` (fig. 3). And a question: Should have the dots in "i", "j" and "ij" have identical shapes and sizes, or may they differ?

The kerning pair lists were (re)generated by the authors of LM with a semi-automatic algorithm reflecting and adjusting the horizontal spaces between the adjacent characters. All the character pairs from the given subset (defined for the corresponding encoding) were included in the processing. However, the numbers of the kerning pairs in the metric files seem to be extremely large and probably many of these kerning pairs are not relevant to any language. I have no good idea how to exclude *automatically* the irrelevant kerning pairs to reduce the space needed for metric data. I have decided to include the complete list of the ligature and kerning pairs for `lmr10` (Latin Modern Roman at 10 pt); see Tables 4 and 5, in the hope that readers will respond with their comments and suggestions.

We pay special attention to the present or absent kerning pairs and to the kern values in the T1 (`ec-lm`) encoding in a comparison with CM and CS fonts. The kerning pairs can be reordered and divided into groups according their agreement or disagreement between CM, CS and LM (for the `ec-lm` encoding) and we can list cases of their discarding, additions or changes in the associated fonts from these families. The corresponding data are collected in my "technical documentation" which I am gathering step by step. Several selected examples are presented in my articles about font verification and comparison [11, 12].

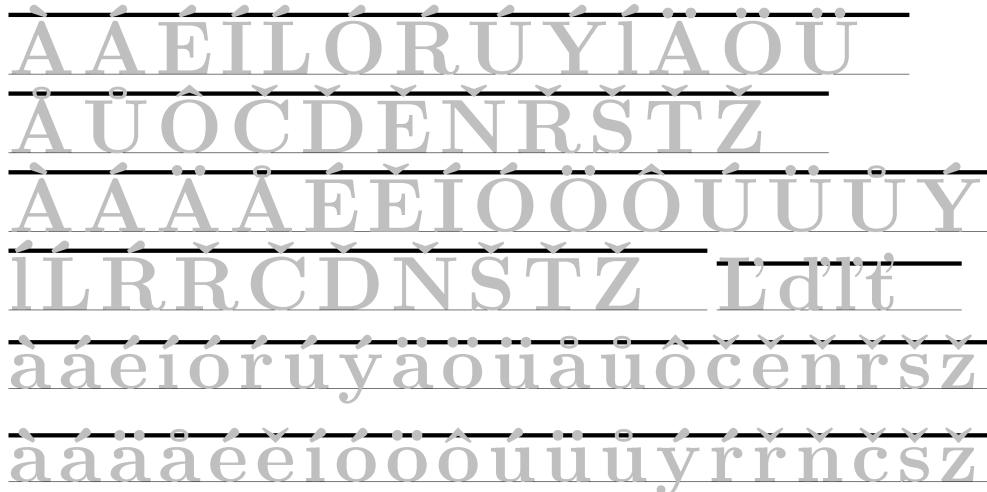


Figure 4: Common optical accent heights.

An important remark: *In OpenType all metric data must be integers* (e.g. metric dimensions). Therefore, metrics converted from T\textit{E}\textit{X} fonts have to be rounded. (Other outline fonts have already been designed on an integer coordinate grid and the T\textit{E}\textit{X} metrics preserve their values.) I cannot estimate the real effects of this incompatibility between OpenType and Type 1 (where the glyph widths are expressed accurately as a ratio of two integers using the `div` command).

Unification of accent heights helps to simplify font production and also to make verification easier. An approach to implementing the unified alignment of optical height for all accents common for all lower case accented letters and also for upper case letters (to another vertical level) is very convenient. Fig. 4 shows a sample test, performed also for all text fonts with a subset of accented letters used in Czech and Slovak, with satisfactory results. The base lines, the boundaries of hinting zones or other vertical levels are marked by the auxiliary horizontal lines.

7 Accents: positioning and shapes

Because LM and CS are both descendants of CM, the glyph images of the common character set are expected to be the same or the differences should be (and in fact are) minimal. Therefore, we concentrated our attention on the accents.

Figures 5 and 6 study several selected accented letters from LM and CS in more detail. In some cases `1mbx10` from LM (stroked outlines) looks better than CS (filled bitmap). Most notably, the uppercase accents in CS, designed in METAFONT more than 15 years ago for dot-matrix printers or used at 300 dpi resolution with laser printers, touch or even overlap

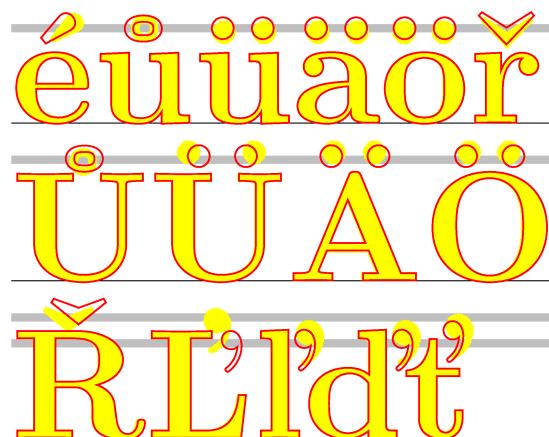


Figure 5: csbx10 (bitmap) vs. lmbx10 (stroked outline).



Figure 6: Additional glyphs, comparing csbx10 and lmbx10.

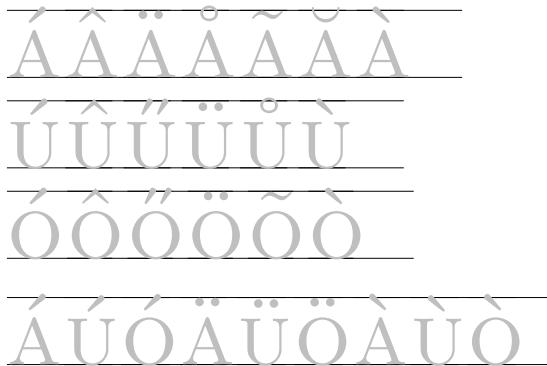


Figure 7: Accents in 1mcsc10 (lower case SmallCaps).

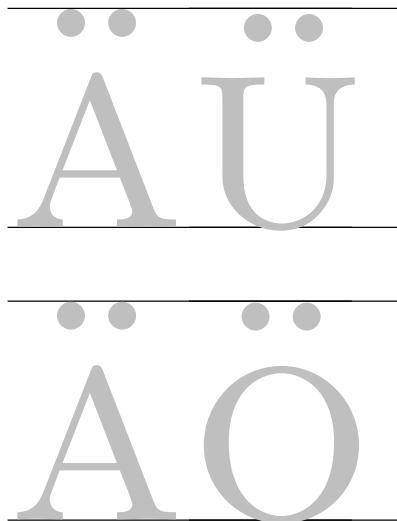


Figure 8: Violation of unified vertical positioning (lower case SmallCaps).

the letter areas. In the current version of LM, probably only Ř needs to adjust the horizontal positioning of the caron.

More complex is the situation with “special” typefaces. For the lower case small capitals (1mcsc10 and its oblique variant) in Fig. 7, with the enlarged detail in Fig. 8, the vertical accent levels are consistent for one letter but are not identical across letters, e.g. the lower case U has evidently had its accents lowered.

And for both SmallCaps and Dunhill (1mdunh10 shown in Fig. 9), I think, the unified optical (lower case) accent level is not the best solution. Accents such as „ and ~ seem to be located too high and the gaps between these narrow (in vertical direction) accents and letters look too big. Additionally, the lower case ring accent in 1mdunh10 (and 1mduno10) is located higher than other accents.



Figure 9: Accents in 1mdunh10.

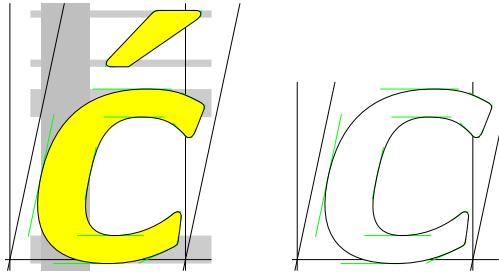


Figure 10: Extrema points and hints in 1mssbo10.

8 Comments on outline font representation

Several glyphs do not fulfill the strongest criteria for the best preciseness or conciseness, and still have tiny defects or small inconsistencies. Their tuning or improvement may be discussed; however, it cannot be considered critical.

Figure 10 demonstrates a few aspects of the font design in LM. In the left part of the picture the technique of hinting the accents is presented. It is widely used in LM and I consider this approach convenient.

The oblique fonts (like 1mssbo10) have been derived from their upright origins and are the result of conversion and transformation. The outline representation is faithful and “optimal”, i.e. it fulfills the “conciseness” consideration. We use fewer Bézier curve segments but, on the other hand, the redundant points at extremes are omitted and the vertical hints look strange or atypical—the boundaries of the hinting zones do not fit in (missing) extrema points. I am interested in comments about this situation.

9 What the tests do not cover

Verification of metrics has been restricted to T1 (EC) encoding (and compared with CM and CS metric data): character widths together with kerning and ligature pairs. Testing of “internal” font matters was foremost; ordinary tests of typesetting real text was not the main goal.

10 Topics for further discussion and conclusion

In the conclusion we summarize some problems that remain to be solved:

- Accent positioning in special cases: Small caps or Dunhill fonts (Fig. 7 and 9).
- Large number of (irrelevant) kerning pairs in metrics (Tables 4 and 5).
- checking of other metric data, e.g. character heights and italic corrections, which are important for typesetting math.
- Points of extrema and hints in (derived) oblique typefaces (Fig. 10).
- tuning of small details in glyph representation.
- proposals for further tests and other improvements (nothing is absolutely perfect).

The Latin Modern fonts fulfill a high quality of technical realization; Type 1 versions are generated by MetaType1 properly, and contain a minimal number of bugs. Remaining tiny defects in online approximation have no practical influence to the printed output of final documents. The metrics (i.e. character widths and kerning pairs) for the CS subset of the T1 (EC) encoding are acceptable; accents and their placement are also acceptable (in most cases).

The LM text fonts could be taken as finished; the LM math fonts in OpenType are (probably) still under development. I have not checked L^AT_EX support or encodings other than T1 (EC). I expect the L^AT_EX users of LM are and will be testing L^AT_EX, dvips, pdfL^AT_EX and other packages during their exploration of LM together with new additions (e.g. new T_EX metrics).

Generally, we can be satisfied with the text LM fonts in the version 1.010(x), and thank the authors for their successful work(s) and wish them further success in the future.

References

- [1] Donald E. Knuth. *The METAFONTbook*. Addison-Wesley, 1986. Volume C of *Computers and Typesetting*.
- [2] Donald E. Knuth. *Computer Modern Typefaces*. Addison-Wesley, 1986. Volume E of *Computers and Typesetting*.
- [3] Computer Modern fonts. CTAN:/fonts/cm.
- [4] CS fonts. ftp://math.feld.cvut.cz/pub/cstex/base/csfonts.tar.gz.
- [5] EC fonts. CTAN:/fonts/ec.
- [6] Latin Modern fonts. CTAN:/fonts/lm.
- [7] Bogusław Jackowski, Janusz M. Nowacki, Piotr Strzelczyk. Programming PostScript Type 1 Fonts Using MetaType1: Auditing, Enhancing, Creating. *Proceedings of EuroTEX 2003*, Brest, France, 24–27 June 2003. *TUGboat* 24:3, pp. 575–581.
- [8] Bogusław Jackowski, Janusz M. Nowacki. Enhancing Computer Modern with accents, accents, accents. *TUGboat* 24:1, 2003.
- [9] Bogusław Jackowski, Janusz M. Nowacki. Latin Modern fonts: How less means more. *Proceedings of the XV EuroTEX 2005 conference*, Pont-à-Mousson, France, March 7–11, 2005.
- [10] Bogusław Jackowski, Janusz M. Nowacki. Rodzina fontów Latin Modern. *Bulletyn Polskiej Grupy Użytkowników Systemu T_EX*, Zeszyt 23:9–12, 2006.
- [11] Karel Píška. Font verification and comparison in examples, *EuroTEX 2006 Proceedings*, *TUGboat* 27:1, pp. 71–75, 2006.
- [12] Karel Píška. Procedures for font comparison, 2007. Proceedings of the EuroTEX 2007 conference, *TUGboat* 29:1, pp. 50–56, 2007.
- [13] Will Robertson. An exploration of the Latin Modern fonts. *TUGboat* 28:2, pp. 177–180, 2007.

Table 4: The complete list of kerning and ligature pairs in ec-lmr10 (beginning).

Table 5: The complete list of kerning and ligature pairs in `ec-lmr10` (end).