# The European Computer Modern Fonts — Documentation —

Jörg Knappen
Barbarossaring 43
55118 Mainz
email: knappen@vkpmzd.kph.uni-mainz.de

1–JUN–1996

# 1 Introduction

## 1.1 The dc fonts

In 1990 at the TUG meeting at Cork, Ireland, the european TEX user groups agreed on a 256 character incoding supporting many european languages with latin writing. This encoding is both an *internal encoding* for TEX and a *font encoding*. This double nature is a consequence of the fact, that both kind of encodings cannot be entirely separated within TEX.

The design goals of the Cork encoding are to allow as many languages as possible to be hyphenated correctly and to guarantee correct kerning for those languages. Therefore it includes many ready-made accented letters.

It also includes some innovative features, which have not become very popular yet, though they deserve to become so. First to mention is a special, zero width invisible character, the compound word mark (cwm). Its design purpose is to assist correct hyphenation at morpheme boundaries, which can occur in minimal pairs (e. g. german *Wachs-tube* vs. *Wach-stube*), if the previous example is typeset in *fraktur* different styles of the letter 's' are needed.

In the new version dc 1.3 I have given the cwm a height, namely x-height, in order be function as a carrier of accents which are placed between letters like in the german abbreviature *-b̆g.* (*-burg*).

The second innovative feature is the separation of the two characters <hyphen> and <hyphenchar>. The hyphenchar is designed as a hanging hyphen, giving a smother text boundary.

The final version of the Cork encoded fonts will be called ec (European Computer Modern or Extended Computer Modern) fonts. The current version, dc 1.3, is the last intermediate step towards the final version. Note, that in the cause of bug fixes and improvements, the metrics may change. After the renaming to ec the metrics will remain stable, as the metrics of the cm fonts do.

### 1.2 The tc-fonts

The need for a text companion font was first articulated in the discussion of new 256 character mathematical fonts in 1993. In order to achieve a better orthogonality between text and math, some text symbols stored in the math fonts should be moved to the text companion fonts[1]. The text companion fonts are also the ideal place to store some new characters, like currency symbols.

There are now 108 characters in the `tc` fonts, including special uppercase versions of the accents, oldstyle digits, genealogical symbols, footnote symbols, currency symbols and other custom characters. The selection is unique and superior to most commercially available expert sets.

## 2 Supported languages

The following languages are supported by the Cork encoding: Afrikaans, Albanian, Breton, Croat, Czech, Danish, Dutch, English, Estonian, Faroese, Finnish, French, Frisian, Gaelic, Galician, German, Greenlandic, Hungarian, Icelandic, Irish (modern orthography), Italian, Letzeburgish, Lusatian (Sorbian), Norwegian, Polish, Portuguese, Rhaetian (Rumantsch), Rumanian, Slovak, Slovene, Spanish, Swedish, Turkish. Many non-european languages using the standard latin alphabet (e. g. Bahasa Indonesia, Suaheli) are also supported.

In europe, the following languages aren't supported: Azeri, Basque, Catalan, Esperanto, Irish (old orthography), Latvian, Lithuanian, Maltese, Sami, Welsh. Of course, Greek and all languages with cyrillic writing are outside the scope of the Cork encoding.

## 3 Standard Control Sequences

The following standard control sequences are assigned with LaTeX's `T1` encoding for the `dc` fonts:
`\r` Ring accent (`\r u` gives ů)
`\k` Ogonek (`\k e` gives ę)
`\dh`, `\DH` Icelandic letter edh (ð, Ð)
`\dj`, `\DJ` Letter d with stroke (đ, Đ)
`\ng`, `\NG` Letter eng (ŋ, Ŋ)
`\th`, `\TH` Icelandic letter thorn (þ, Þ).

To load the `tc` fonts, use the `textcomp`-package, which provides control sequences for all the included symbols.

[1]The archives of the math-font-discuss mailing list are available for ftp on `ftp.cogs.susx.ac.uk` in directory `pub/tex/mathfont`.

# 4 Ligatures

In the proportional fonts, the following ligatures are implemented:

`--` – (en dash)  
`---` — (em dash)  
`‘‘` “ (english opening quotes, german closing quotes)  
`’’` ” (english and polish closing quotes)  
`,,` „ (german and polish opening quotes)  
`<<` « (french opening quotes)  
`>>` » (french closing quotes)  
`!‘` ¡ (spanish opening exclamation mark)  
`?‘` ¿ (spanish opening question mark)  
`fi` ﬁ  
`ff` ﬀ  
`fl` ﬂ  
`ffi` ﬃ  
`ffl` ﬄ

In the typewriter fonts, the following ligatures are implemented:

`--` `-` (en dash or number range dash)  
`---` `--` (em dash)  
`‘‘` “ (english opening quotes, german closing quotes)  
`’’` ” (english and polish closing quotes)  
`,,` „ (german and polish opening quotes)  
`<<` « (french opening quotes)  
`>>` » (french closing quotes)  
`!‘` ¡ (spanish opening exclamation mark)  
`?‘` ¿ (spanish opening question mark)

The convention on the dashes suites british usage for number range dashes best and does not interfer with any other known usage. In verbatim mode, all ligatures are switched off.

# 5 Hints on usage

The `dc` fonts are intended for text usage in european languages. The Cork font encoding is selected with the command `\usepackage[T1]{fontenc}` in LaTeX $2_{\varepsilon}$.

The `tc` fonts are a multi-purpose font. Suggested usages include verbatim setting of latin-1 and latin-2 listings, avoiding the so-called “hidden math” in text mode (that’s the reason, why there are footnote symbols in), providing building blocks for virtual fonts (oldstyle digits are included for this reason), or just providing otherwise unavailable symbols (like the permille sign).

Some characters are primarily intended for verbatim listings, in plain text they may be replaced with macros. These characters include the raised digits, the fractions, the trademark sign, and the ordinal indicators.

For text fractions, the following macro is suggested (from the TEXbook, exercise 11.6):

```
\newcommand{\nicefrac}[2]{\leavevmode\kern.1em
\raise.5ex\hbox{\the\scriptfont0 #1}\kern-.1em
/\kern-.15em\lower.25ex\hbox{\the\scriptfont0 #2}}
```

It can produce arbitrary fractions and is not restricted to some simple cases, the output looks $^1/_2$, $^5/_4$, $^{17}/_{42}$.

For the ordinal indicators (º and ª), the following macros are suggested (from spanish.ldf, babel bundle):

```
\newcommand{\ordmale}{\raise1ex\hbox{\underbar{\scriptsize o}}}
\newcommand{\ordfemale}{\raise1ex\hbox{\underbar{\scriptsize a}}}
```

# 6 Naming of the font files

Currently, the extended computer modern font have the prefix dc. This prefix will changew to ec with the final release after another round of bug fixing. I hope to make the transition from dc to ec in about one year. The text companion fonts have the prefix tc, which is not subject to change. However, later releases may included more characters and therefore have different checksums. No characters shall be removed from the tc fonts.

Most of the dc fonts can be generated at any size one want in the range from 5pt to 100pt. For each size, a unique name is needed.

With the release 1.2 of the dc fonts, a new, more precise naming scheme is in effect. Since there are widely used operating sytems limiting the file name to 8 character (plus an extension of 3 characters) the following scheme is used:

- The first two letters (either `dc` or `tc` denote the encoding and the general design of the font.
- The one or two following letters denotes the family, shape, and series attributes of the font. E. g. `r` for roman, `bx` for bold extended, `ti` for text italic, or `bi` for bold extended italic. A complete overview is given at the end of this section.
- The following four digits give the design size in TEX's points multiplied with 100. E. g `1000` denotes tex point, `1440` denotes magstep 2, i. e. 14.4 point, and `0500` denotes five point.

Here are the implemented styles:

**Roman family:** `r` roman, `b` bold, `bx` bold extended, `sl` slanted, `bl` bold extended slanted, `cc` caps and small caps, `xc` bold extended caps and small caps, `sc` slanted caps and small caps, `oc` oblique (bold extended slanted) caps and small caps, `ti` (text) italic, `bi` bold extended italic, `u` unslanted italic, `ci` classical serif italic (new design).

**Sans serif family:** `ss` sans serif, `si` sans serif inclined (slanted), `sx` sans serif bold extended, `so` sans serif bold extended oblique (slanted).

**Typewriter family:** `tt` typewriter, `tc` typewriter caps and small caps, `st` slanted typewriter, `it` italic typewriter.

**Variable width typewriter family** `vt` variable width typewriter, `vi` variable width italic typewriter.

**Various other fonts:** `bm` variant bold roman, `dh` dunhill, `fb` Fibonacci parameters, `ff` funny, `fi` funny italic. Expect errors with the funny fonts, they aren't really worked out.

Here are some examples:

| | |
|---|---|
| `dcr1000` | European computer modern roman at 10pt |
| `tcr1000` | Text companion symbols roman at 10pt |
| `dcss1728` | European computer modern sans serif at 17.28pt |
| `dcbx0900` | European computer modern roman bold extended at 9pt |

Some remaining fonts come at one size only, those are

| | |
|---|---|
| `dcssdc10` | sans serif demi-bold condensed |
| `dcsq8` | sans serif quotation |
| `dcqi8` | sans serif quotation inclined |
| `dclq8` | latex sans serif quotation |
| `dcli8` | latex sans serif quotation inclined |
| `dclb8` | latex sans serif quotation bold |
| `dclo8` | latex sans serif quotation oblique (bold inclined) |
| `idclq8` | invisible latex sans serif quotation |
| `idcli8` | invisible latex sans serif quotation inclined |
| `idclb8` | invisible latex sans serif quotation bold |
| `idclo8` | invisible latex sans serif quotation oblique. |

The last eight fonts are for the slides document class, which replaces old SliTeX. They contain a special version of the capital letter 'I'.

# A The Cork Encoding

| position (octal) | description |
|---|---|
| Accents for lowercase letters | |
| 000 | grave |
| 001 | acute |
| 002 | circumflex |
| 003 | tilde |
| 004 | umlaut |

| | |
|---|---|
| 005 | hungarian |
| 006 | ring |
| 007 | hachek |
| 010 | breve |
| 011 | macron |
| 012 | dot above |
| 013 | cedilla |
| 014 | ogonek |
| Miscellaneous | |
| 015 | single base quote |
| 016 | single opening guillemet |
| 017 | single closing guillemet |
| 020 | english opening quotes |
| 021 | english closing quotes |
| 022 | base quotes |
| 023 | opening guillemets |
| 024 | closing guillemets |
| 025 | en dash |
| 026 | em dash |
| 027 | compound word mark (invisible) |
| 030 | perthousandzero |
| 031 | dotless i |
| 032 | dotless j |
| 033 | ligature ff |
| 034 | ligature fi |
| 035 | ligature fl |
| 036 | ligature ffi |
| 037 | ligature ffl |
| 040 | visible space |
| ASCII | |
| 041 | exclamation mark |
| 042 | straight quotes |
| 043 | hash mark |
| 044 | dollar sign |
| 045 | percent sign |
| 046 | ampersand |
| 047 | apostroph |
| 050 | opening parentheses |
| 051 | closing parentheses |
| 052 | asterisk |
| 053 | plus sign |
| 054 | comma |

| | |
|---|---|
| 055 | hyphen (note: not minus sign) |
| 056 | full stop |
| 057 | solidus |
| 060 | digit 0 |
| ... | |
| 071 | digit 9 |
| 072 | colon |
| 073 | semicolon |
| 074 | less than sign |
| 075 | equals sign |
| 076 | greater than sign |
| 077 | question mark |
| 080 | commercial at |
| 081 | capital letter A |
| ... | |
| 132 | capital letter Z |
| 133 | opening square bracket |
| 134 | backslash |
| 135 | closing square bracket |
| 136 | ASCII circumflex |
| 137 | underscore |
| 140 | opening quote (not ASCII grave!) |
| 141 | lowercase letter a |
| ... | |
| 172 | lowercase letter z |
| 173 | opening curly brace |
| 174 | vertical bar |
| 175 | closing curly brace |
| 176 | ASCII tilde |
| 177 | hyphenchar (hanging) |

Letters for eastern european languages (from latin-2)

| | |
|---|---|
| 200 | capital letter A with breve |
| 201 | capital letter A eith ogonek |
| 202 | capital letter C with acute |
| 203 | capital letter C with hachek |
| 204 | capital letter D with hachek |
| 205 | capital letter E with hachek |
| 206 | capital letter E with ogonek |
| 207 | capital letter G with breve |
| 210 | capital letter L with acute |
| 211 | capital letter L with hachek |
| 212 | capital letter crossed L |
| 213 | capital letter N with acute |
| 214 | capital letter N with hachek |

| | |
|---|---|
| 215 | capital letter Eng |
| 216 | capital letter O with hungarian double acute |
| 217 | capital letter R with acute |
| 220 | capital letter R with hachek |
| 221 | capital letter S with acute |
| 222 | capital letter S with hachek |
| 223 | capital letter S with cedilla |
| 224 | capital letter T with hachek |
| 225 | capital letter T with cedilla |
| 226 | capital letter U with hungarian double acute |
| 227 | capital letter U with ring |
| 230 | capital letter Y with diaeresis |
| 231 | capital letter Z with acute |
| 232 | capital letter Z with hachek |
| 233 | capital letter Z with dot |
| 234 | capital letter IJ |
| 235 | capital letter I with dot |
| 236 | lowercase letter d with bar |
| 237 | section sign |
| 240 | lowercase letter a with breve |
| 241 | lowercase letter a with ogonek |
| 242 | lowercase letter c with acute |
| 243 | lowercase letter c with hachek |
| 244 | lowercase letter d with hachek |
| 245 | lowercase letter e with hachek |
| 246 | lowercase letter e with ogonek |
| 247 | lowercase letter g with breve |
| 250 | lowercase letter l with acute |
| 251 | lowercase letter l with hachek |
| 252 | lowercase letter crossed l |
| 253 | lowercase letter n with acute |
| 254 | lowercase letter n with hachek |
| 255 | lowercase letter eng |
| 256 | lowercase letter o with hungarian double acute |
| 257 | lowercase letter r with acute |
| 260 | lowercase letter r with hachek |
| 261 | lowercase letter s with acute |
| 262 | lowercase letter s with hachek |
| 263 | lowercase letter s with cedilla |
| 264 | lowercase letter t with hachek |
| 265 | lowercase letter t with cedilla |
| 266 | lowercase letter u with hungarain double acute |
| 267 | lowercase letter u with ring |
| 270 | lowercase letter y with diaeresis |
| 271 | lowercase letter z with acute |
| 272 | lowercase letter z with hachek |

| | |
|---|---|
| 273 | lowercase letter z with dot |
| 274 | lowercase letter ij |
| 275 | spanish inverted exclamation mark |
| 276 | spanish inverted question mark |
| 277 | pound sign |

| Letters for western european languages (from latin-1) | |
|---|---|
| 300 | capital letter A with grave |
| 301 | capital letter A with acute |
| 302 | capital letter A with circumflex |
| 303 | capital letter A with tilde |
| 304 | capital letter A with diaeresis |
| 305 | capital letter A with ring |
| 306 | capital letter AE |
| 307 | capital letter C with cedilla |
| 310 | capital letter E with grave |
| 311 | capital letter E with acute |
| 312 | capital letter E with circumflex |
| 313 | capital letter E with diaeresis |
| 314 | capital letter I with grave |
| 315 | capital letter I with acute |
| 316 | capital letter I with circumflex |
| 317 | capital letter I with diaeresis |
| 320 | capital letter Edh (D with bar) |
| 321 | capital letter N with tilde |
| 322 | capital letter O with grave |
| 323 | capital letter O with acute |
| 324 | capital letter O with circumflex |
| 325 | capital letter O with tilde |
| 326 | capital letter O with diaeresis |
| 327 | capital letter OE |
| 330 | capital letter O with slash |
| 331 | capital letter U with grave |
| 332 | capital letter U with acute |
| 333 | capital letter U with circumflex |
| 334 | capital letter U with diaeresis |
| 335 | capital letter Y with acute |
| 336 | capital letter Thorn |
| 337 | capital letter Sharp S (deviating from latin-1) |
| 340 | lowercase letter a with grave |
| 341 | lowercase letter a with acute |
| 342 | lowercase letter a with circumflex |
| 343 | lowercase letter a with tilde |
| 344 | lowercase letter a with diaeresis |
| 345 | lowercase letter a with ring |

| | |
|---|---|
| 346 | lowercase letter ae |
| 347 | lowercase letter c with cedilla |
| 350 | lowercase letter e with grave |
| 351 | lowercase letter e with acute |
| 352 | lowercase letter e with circumflex |
| 353 | lowercase letter e with diaeresis |
| 354 | lowercase letter i with grave |
| 355 | lowercase letter i with acute |
| 356 | lowercase letter i with circumflex |
| 357 | lowercase letter i with diaeresis |
| 360 | lowercase letter edh |
| 361 | lowercase letter n with tilde |
| 362 | lowercase letter o with grave |
| 363 | lowercase letter o with acute |
| 364 | lowercase letter o with circumflex |
| 365 | lowercase letter o with tilde |
| 366 | lowercase letter o with diaeresis |
| 367 | lowercase letter oe |
| 370 | lowercase letter o with slash |
| 371 | lowercase letter u with grave |
| 372 | lowercase letter u with acute |
| 373 | lowercase letter u with circumflex |
| 374 | lowercase letter u with diaeresis |
| 375 | lowercase letter y with acute |
| 376 | lowercase letter thorn |
| 377 | lowercase letter sharp s (deviating from latin-1) |

# B The Text Companion Encoding

| position (octal) | description |
|---|---|
| Accents for capital letters | |
| 000 | grave |
| 001 | acute |
| 002 | circumflex |
| 003 | tilde |
| 004 | umlaut |
| 005 | hungarian |
| 006 | ring |
| 007 | hachek |
| 010 | breve |
| 011 | macron |
| 012 | dot above |

| | |
|---|---|
| 013 | cedilla |
| 014 | ogonek |

| Miscellaneous | |
|---|---|
| 015 | base single straight quote |
| 022 | base double straight quotes |
| 025 | twelve u dash |
| 026 | three quarters emdash |
| 027 | capital cwm |
| 030 | left pointing arrow |
| 031 | right pointing arrow |
| 032 | tie accent (lowercase) |
| 033 | tie accent (capital) |
| 040 | blank symbol |
| 044 | dollar sign |
| 047 | straight quote |
| 052 | centered star |
| 054 | comma |
| 056 | full stop |
| 057 | fraction |

| Oldstyle digits | |
|---|---|
| 060 | oldstyle digit 0 |
| 061 | oldstyle digit 1 |
| 062 | oldstyle digit 2 |
| 063 | oldstyle digit 3 |
| 064 | oldstyle digit 4 |
| 065 | oldstyle digit 5 |
| 066 | oldstyle digit 6 |
| 067 | oldstyle digit 7 |
| 070 | oldstyle digit 8 |
| 071 | oldstyle digit 9 |

| Miscellaneous | |
|---|---|
| 115 | mho sign |
| 117 | big circle |
| 127 | ohm sign |
| 136 | arrow up |
| 137 | arrow down |
| 140 | backtick (ASCII grave) |
| 142 | born |
| 144 | died |
| 154 | leaf |

| | |
|---|---|
| 155 | married |
| 156 | musical note |
| 176 | low tilde |
| 177 | short equals |
| TS1-symbols | |
| 200 | ASCII-style breve |
| 201 | ASCII-style hachek |
| 202 | double tick (ASCII double acute) |
| 203 | double backtick |
| 204 | dagger |
| 205 | ddager |
| 206 | double vert |
| 207 | perthousand |
| 210 | bullet |
| 211 | centigrade |
| 212 | dollaroldstyle |
| 213 | centoldstyle |
| 214 | florin |
| 215 | colon |
| 216 | won |
| 217 | naira |
| 220 | guarani |
| 221 | peso |
| 222 | lira |
| 223 | recipe |
| 224 | interrobang |
| 225 | gnaborretni |
| 226 | dong sign |
| 227 | trademark |
| Symbols from ISO-8859-1 (latin-1) | |
| 242 | cent |
| 243 | sterling |
| 244 | currency sign |
| 245 | yen |
| 246 | broken vertical bar |
| 247 | section sign |
| 250 | high dieresis |
| 251 | copyright |
| 252 | feminine ordinal indicator |
| 254 | logical not |
| 256 | circled R |
| 257 | macron |

| | |
|---|---|
| 260 | degree sign |
| 261 | plus-minus sign |
| 262 | superscript 2 |
| 263 | superscript 3 |
| 264 | tick (ASCII-style acute) |
| 265 | micro sign |
| 266 | pilcrow sign |
| 267 | centered dot |
| 271 | superscript 1 |
| 272 | masculine ordinal indicator |
| 274 | fraction one quarter |
| 275 | fraction one half |
| 276 | fraction three quarters |
| 326 | multiplication sign (times) |
| 366 | division sign |