1 The Cyrillic font encodings: T2A, T2B, T2C, and X2

Since the number of Cyrillic glyphs exceeds the limit for a ‘T’ encoding, it is necessary to create multiple glyph containers. The output encodings ‘T2A’, ‘T2B’, and ‘T2C’ are the result of long discussions in various Internet mailing lists; they emerged from the ‘X2’ glyph container which was previously called ‘T2’. Please consult section 6.4 (‘Naming conventions’) in the file fntguide.tex of the base \texttt{LATeX} 2\epsilon distribution for details on the differences between \texttt{LATeX} font encodings.

When typesetting Cyrillic texts, there is a tradition to use Cyrillic letters (in some situations) inside math formulae just like Latin letters. By default, this does not work, because symbols declared with \texttt{\DeclareTextSymbol} may not be used inside math. If you need to ‘transparently’ typeset symbols in math formulae which are declared in encoding definition files, you could use the \texttt{mathtext} package available at \texttt{CTAN:macros/latex/contrib/supported/t2}. Note that this package uses at least one additional math alphabet per font encoding.

To make code more compact, we use the \texttt{\LastDeclaredEncoding} macro to share the same lines in the DTX file for different encodings. This macro appeared in the December 1998 \texttt{LATeX} 2\epsilon release (together with the integration of Cyrillic stuff), thus we use \texttt{\NeedsTeXFormat} below. We will probably change this ‘sharing mechanism’ to use a yet-to-be-written \texttt{docstrip} extension (namely, substring substitution) in the future.

\begin{verbatim}
1 ⟨X2 | T2A | T2B | T2C⟩
2 \NeedsTeXFormat{LaTeX2e}[1998/12/01]
3 ⟨X2⟩\ProvidesFile{x2enc.def}
4 ⟨T2A⟩\ProvidesFile{t2aenc.def}
5 ⟨T2B⟩\ProvidesFile{t2benc.def}
6 ⟨T2C⟩\ProvidesFile{t2cenc.def}
7 [2005/09/27 v1.0i Cyrillic encoding definition file]

To assure that \texttt{\MakeUppercase} and \texttt{\MakeLowercase} will work correctly, \texttt{\@uclclist} (defined in the \texttt{LATeX} kernel) contains all Cyrillic uppercase-lowercase letter pairs. This also needs at least December 1998 \texttt{LATeX} 2\epsilon release.

8 ⟨X2⟩\DeclareFontEncoding{X2}{}{}
9 ⟨T2A⟩\DeclareFontEncoding{T2A}{}{}
10 ⟨T2B⟩\DeclareFontEncoding{T2B}{}{}
11 ⟨T2C⟩\DeclareFontEncoding{T2C}{}{}
12 \DeclareFontSubstitution{\LastDeclaredEncoding}{cmr}{m}{n}

Common accents.

13 \DeclareTextAccent{\'}{\LastDeclaredEncoding}{0}
14 \DeclareTextAccent{\'}{\LastDeclaredEncoding}{1}
15 \DeclareTextAccent{\'}{\LastDeclaredEncoding}{2}
16 \DeclareTextAccent{\'}{\LastDeclaredEncoding}{3}
17 \DeclareTextAccent{\'}{\LastDeclaredEncoding}{4}
18 \DeclareTextAccent{\H}{\LastDeclaredEncoding}{5}
19 \DeclareTextAccent{\r}{\LastDeclaredEncoding}{6}
20 \DeclareTextAccent{\v}{\LastDeclaredEncoding}{7}
21 \DeclareTextAccent{\u}{\LastDeclaredEncoding}{8}
22 \DeclareTextAccent{\=}{\LastDeclaredEncoding}{9}
23 \DeclareTextAccent{\.-}{\LastDeclaredEncoding}{10}
\end{verbatim}

1
\textperthousand\% 11 cedilla
\textpertenthousand\% 12 ogonek

\% Cyrillic flex (inverted breve) accent:
\DeclareTextAccent{\f}{\LastDeclaredEncoding}{18}
\% double grave (‘Cyrillic umlaut’) accent:
\DeclareTextAccent{\C}{\LastDeclaredEncoding}{19}
\% Cyrillic breve accent:
\DeclareTextAccent{\U}{\LastDeclaredEncoding}{20}

In these definitions we no longer use the helper function \sh@ft from plain.tex since that now has two incompatible definitions.
\DeclareTextCommand{\b}{\LastDeclaredEncoding}{1}
\hmode@bgroup\o@lign{\relax#1\crcr\hidewidth\ltx@sh@ft{-3ex}\%\vbox to.2ex{\hbox{\char9}\vss}\hidewidth}\egroup
\DeclareTextCommand{\c}{\LastDeclaredEncoding}{1}
\leavevmode\setbox\z@hbox{#1}\ifdim\ht\z@=1ex\accent11 #1\%\else{\ooalign{\hidewidth\char11\hidewidth\crcr\unhbox\z@}}\fi
\DeclareTextCommand{\d}{\LastDeclaredEncoding}{1}
\hmode@bgroup\o@lign{\relax#1\crcr\hidewidth\ltx@sh@ft{-1ex}.\hidewidth}\egroup
\DeclareTextCommand{\k}{\LastDeclaredEncoding}{1}
\oalign{\null#1\crcr\hidewidth\char12}
\DeclareTextCommand{\textperthousand}{\LastDeclaredEncoding}{1}
\%\char 24
\DeclareTextCommand{\textpertenthousand}{\LastDeclaredEncoding}{1}
\%\char 24\char 24
\DeclareTextCompositeCommand{\r}{\LastDeclaredEncoding}{1}
\leavevmode\setbox\z@hbox{!}\dimen@\ht\z@\advance\dimen@-1ex%\rlap{\raise.67\dimen@\hbox{\char6}}A
\textendash{\LastDeclaredEncoding}{21}

Slot 22 is a ‘cyrillic emdash’ — \cyrdash which is longer than endash, but shorter than english emdash (but we define an alias too).
\textemdash{\LastDeclaredEncoding}{22}
\textcompwordmark{\LastDeclaredEncoding}{23}
\% 24 perthousandzero
\textvisiblespace{\LastDeclaredEncoding}{32}
\textdollar{\LastDeclaredEncoding}{36}
\textless{\LastDeclaredEncoding}{60}
\textgreater{\LastDeclaredEncoding}{62}
\textbackslash{\LastDeclaredEncoding}{92}
\textasciicircum{\LastDeclaredEncoding}{94}
\textunderscore{\LastDeclaredEncoding}{95}
\textbraceleft{\LastDeclaredEncoding}{123}
\textbar{\LastDeclaredEncoding}{124}
\textrighthand{\LastDeclaredEncoding}{125}
\textasciitilde{\LastDeclaredEncoding}{126}
\textendash{\LastDeclaredEncoding}{21}
\i and \j are unused in Cyrillic languages, so are absent in X2. Slots 25, 26, 27 in the X2 encoding are ‘cross grave’, ‘cross acute’, and ‘hcross’ modifiers, respectively.

\begin{itemize}
\item \texttt{(X2)\textsymbol{\i}} (LastDeclaredEncoding\{25\})
\item \texttt{(X2)\textsymbol{\j}} (LastDeclaredEncoding\{26\})
\end{itemize}
\texttt{\textsymbol{\i}}(\textsymbol{\i})

Here are the letters common to all font encodings.

\begin{itemize}
\item \texttt{\textsymbol{\CYRA}} (LastDeclaredEncoding\{192\})
\item \texttt{\textsymbol{\cyra}} (LastDeclaredEncoding\{224\})
\item \texttt{\textsymbol{\CYRB}} (LastDeclaredEncoding\{193\})
\item \texttt{\textsymbol{\cyrb}} (LastDeclaredEncoding\{225\})
\item \texttt{\textsymbol{\CYRV}} (LastDeclaredEncoding\{194\})
\item \texttt{\textsymbol{\cyrv}} (LastDeclaredEncoding\{226\})
\item \texttt{\textsymbol{\CYRG}} (LastDeclaredEncoding\{195\})
\item \texttt{\textsymbol{\cyrg}} (LastDeclaredEncoding\{227\})
\item \texttt{\textsymbol{\CYRD}} (LastDeclaredEncoding\{196\})
\item \texttt{\textsymbol{\cyrd}} (LastDeclaredEncoding\{228\})
\item \texttt{\textsymbol{\CYREW}} (LastDeclaredEncoding\{197\})
\item \texttt{\textsymbol{\cyre}} (LastDeclaredEncoding\{229\})
\item \texttt{\textsymbol{\CYRZH}} (LastDeclaredEncoding\{198\})
\item \texttt{\textsymbol{\cyrzh}} (LastDeclaredEncoding\{230\})
\item \texttt{\textsymbol{\CYRZ}} (LastDeclaredEncoding\{199\})
\item \texttt{\textsymbol{\cyrz}} (LastDeclaredEncoding\{231\})
\item \texttt{\textsymbol{\CYRI}} (LastDeclaredEncoding\{200\})
\item \texttt{\textsymbol{\cyrri}} (LastDeclaredEncoding\{232\})
\item \texttt{\textsymbol{\CYRISHRT}} (LastDeclaredEncoding\{201\})
\item \texttt{\textsymbol{\cyrishrt}} (LastDeclaredEncoding\{233\})
\item \texttt{\textsymbol{\CYRK}} (LastDeclaredEncoding\{202\})
\item \texttt{\textsymbol{\cyrk}} (LastDeclaredEncoding\{234\})
\item \texttt{\textsymbol{\CYRL}} (LastDeclaredEncoding\{203\})
\item \texttt{\textsymbol{\cyrl}} (LastDeclaredEncoding\{235\})
\item \texttt{\textsymbol{\CYRM}} (LastDeclaredEncoding\{204\})
\item \texttt{\textsymbol{\cyrm}} (LastDeclaredEncoding\{236\})
\item \texttt{\textsymbol{\CYRN}} (LastDeclaredEncoding\{205\})
\item \texttt{\textsymbol{\cyrn}} (LastDeclaredEncoding\{237\})
\item \texttt{\textsymbol{\CYRB}} (LastDeclaredEncoding\{206\})
\item \texttt{\textsymbol{\cyrb}} (LastDeclaredEncoding\{238\})
\end{itemize}
And now the remaining letters.

\[
\text{\textcopyright{\textsc{	extregistered}}} X2
\]

\[
\text{\textcopyright{\textregistered}} X2\{128\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{160\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{129\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{161\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{130\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{162\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{131\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{163\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{132\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{164\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{133\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{165\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{134\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{166\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{135\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{167\}
\]

\[
\text{\textcopyright{\textregistered}} X2\{136\}
\]
\DeclareTextSymbol{\textcyryi}{X2}{168}
\DeclareTextSymbol{\textcyrkds}{X2}{137}
\DeclareTextSymbol{\textcyrkds}{X2}{169}
\DeclareTextSymbol{\textcyrkbeak}{X2}{138}
\DeclareTextSymbol{\textcyrkbeak}{X2}{170}
\DeclareTextSymbol{\textcyrkhh}{X2}{139}
\DeclareTextSymbol{\textcyrkhh}{X2}{171}
\DeclareTextSymbol{\textcyrlhh}{X2}{140}
\DeclareTextSymbol{\textcyrlhh}{X2}{172}
\DeclareTextSymbol{\textcyrndsc}{X2}{141}
\DeclareTextSymbol{\textcyrndsc}{X2}{173}
\DeclareTextSymbol{\textcyrnu}{X2}{142}
\DeclareTextSymbol{\textcyrnu}{X2}{174}
\DeclareTextSymbol{\textcyryn}{X2}{143}
\DeclareTextSymbol{\textcyryn}{X2}{175}
\DeclareTextSymbol{\textcyrotld}{X2}{144}
\DeclareTextSymbol{\textcyrotld}{X2}{176}
\DeclareTextSymbol{\textcyrbsdsc}{X2}{145}
\DeclareTextSymbol{\textcyrbsdsc}{X2}{177}
\DeclareTextSymbol{\textcyrshrt}{X2}{146}
\DeclareTextSymbol{\textcyrshrt}{X2}{178}
\DeclareTextSymbol{\textcyr}{X2}{147}
\DeclareTextSymbol{\textcyr}{X2}{179}
\DeclareTextSymbol{\textcyryhcrs}{X2}{148}
\DeclareTextSymbol{\textcyryhcrs}{X2}{180}
\DeclareTextSymbol{\textcyrhdsc}{X2}{149}
\DeclareTextSymbol{\textcyrhdsc}{X2}{181}
\DeclareTextSymbol{\textcyrhh}{X2}{150}
\DeclareTextSymbol{\textcyrhh}{X2}{182}
\DeclareTextSymbol{\textcyrchldsc}{X2}{151}
\DeclareTextSymbol{\textcyrchldsc}{X2}{183}
\DeclareTextSymbol{\textcyrhrds}{X2}{152}
\DeclareTextSymbol{\textcyrhrds}{X2}{184}
\DeclareTextSymbol{\textcyrie}{X2}{153}
\DeclareTextSymbol{\textcyrie}{X2}{185}
\DeclareTextSymbol{\textcyrschwa}{X2}{154}
\DeclareTextSymbol{\textcyrschwa}{X2}{186}
\DeclareTextSymbol{\textcyreps}{X2}{155}
\DeclareTextSymbol{\textcyreps}{X2}{187}
\langle X2 \rangle
\% \langle \textCYRGUP}{T2A}{128}
\DeclareTextSymbol{\textCYRGUP}{T2A}{128}
\DeclareTextSymbol{\textCYRGUP}{T2A}{160}
\DeclareTextSymbol{\textCYRGHC}{T2A}{129}
\DeclareTextSymbol{\textCYRGHC}{T2A}{161}
\DeclareTextSymbol{\textCYRDJE}{T2A}{130}
\DeclareTextSymbol{\textCYRDJE}{T2A}{162}
\DeclareTextSymbol{\textCYRTSHE}{T2A}{131}
\DeclareTextSymbol{\textCYRTSHE}{T2A}{163}
\DeclareTextSymbol{\textCYRSHHA}{T2A}{132}
\DeclareTextSymbol{\textCYRSHHA}{T2A}{164}
\DeclareTextSymbol{\textCYRZHDSC}{T2A}{133}
\DeclareTextSymbol{\textCYRZHDSC}{T2A}{165}
The final block of Cyrillic letters occupies the space covered with Latin letters in ASCII. These letters are only present in X2 encoding (which does not contain Latin letters at all). However, some Cyrillic glyphs here look exactly like some Latin analogues, and we declare them also for T2* encodings, so that they are accessible from inputenc files. In T2* encodings, these letters share the same slots with the corresponding Latin letters. The affected letters are: \CYRII, \CYRJE, \CYRQ, \CYRDZE, \CYR (and lowercase forms).
\DeclareTextSymbol{\CYRDZE}{\LastDeclaredEncoding}{83}
\DeclareTextSymbol{\cyrdze}{\LastDeclaredEncoding}{115}
\DeclareTextSymbol{\CYRTETSE}{X2}{84}
\DeclareTextSymbol{\cyrtetse}{X2}{116}
\DeclareTextSymbol{\CYRDZH}{X2}{85}
\DeclareTextSymbol{\cyrdzh}{X2}{117}
\DeclareTextSymbol{\CYRVCRS}{X2}{86}
\DeclareTextSymbol{\cyrchvcrs}{X2}{118}
\DeclareTextSymbol{\CYRW}{\LastDeclaredEncoding}{87}
\DeclareTextSymbol{\cyrw}{\LastDeclaredEncoding}{119}
\DeclareTextSymbol{\CYRYAT}{X2}{88}
\DeclareTextSymbol{\cyryat}{X2}{120}
\DeclareTextSymbol{\CYRBYUS}{X2}{89}
\DeclareTextSymbol{\cyrbyus}{X2}{121}
\DeclareTextSymbol{\CYRIZH}{X2}{90}
\DeclareTextSymbol{\cyrizh}{X2}{122}
\DeclareTextSymbol{\CYRDELTA}{X2}{28}
\DeclareTextSymbol{\cyrdelta}{X2}{29}
\DeclareTextSymbol{\CYRNLHK}{X2}{30}
\DeclareTextSymbol{\cyrnlhk}{X2}{31}
\DeclareTextSymbol{\CYRABHCH}{X2}{69}
\DeclareTextSymbol{\cyrabhch}{X2}{101}
\DeclareTextSymbol{\CYRABHCH}{T2C}{146}
\DeclareTextSymbol{\cyrabhch}{T2C}{178}

Text composites. The following declarations will not work for 8-bit chars generated
via \inputenc unless a dblacnt package is used.
\DeclareTextComposite{"}{\LastDeclaredEncoding}{\CYRE}{156}
\DeclareTextComposite{"}{\LastDeclaredEncoding}{\cyre}{188}
\DeclareTextComposite{U}{\LastDeclaredEncoding}{\CYRI}{201}
\DeclareTextComposite{U}{\LastDeclaredEncoding}{\cyri}{233}
\DeclareTextComposite{c}{\LastDeclaredEncoding}{\CYRZ}{134}
\DeclareTextComposite{c}{\LastDeclaredEncoding}{\cyrz}{166}
\DeclareTextComposite{K}{\LastDeclaredEncoding}{\CYRS}{145}
\DeclareTextComposite{K}{\LastDeclaredEncoding}{\cyrs}{177}
\DeclareTextComposite{U}{\LastDeclaredEncoding}{\CYRU}{146}
\DeclareTextComposite{U}{\LastDeclaredEncoding}{\cyru}{178}
\DeclareTextComposite{K}{\LastDeclaredEncoding}{\cyrabhch}{69}
\DeclareTextComposite{K}{\LastDeclaredEncoding}{\cyrabhch}{101}
\DeclareTextComposite{K}{T2C}{\cyrabhch}{146}
\DeclareTextComposite{K}{T2C}{\cyrabhch}{178}