The selinput package

Heiko Oberdiek <heiko.oberdiek at googlemail.com>

2007/09/09 v1.2

Abstract

This package selects the input encoding by specifying between input characters and their glyph names.

Contents

1	Doo	cumentation	1
	1.1	Introduction	1
	1.2	User interface	2
	1.3	Options	3
	1.4	Encodings	3
2	Imp	plementation	3
3	Tes	t	7
4	Inst	tallation	9
	4.1	Download	9
	4.2	Bundle installation	9
	4.3	Package installation	10
	4.4	Refresh file name databases	10
	4.5	Some details for the interested	10
5	Cat	calogue	11
6	Ref	erences	11
7		201,	12
	[200	7/06/16 v1.0]	12
		7/06/20 v1.1	12
		7/09/09 v1.2]	12
8	Ind	ey	12

1 Documentation

1.1 Introduction

 \LaTeX supports the direct use of 8-bit characters by means of package inputenc. However you must know and specify the encoding, e.g.:

```
\documentclass{article}
\usepackage[latin1]{inputenc}
% or \usepackage[utf8]{inputenc}
% or \usepackage[??]{inputenc}
```

```
\begin{document}
    Umlauts: ÄÖÜäöüß
\end{document}
```

If the document is transferred in an environment that uses a different encoding, then there are programs that convert the input characters. Examples for conversion of file test.tex from encoding latin1 (ISO-8859-1) to UTF-8:

However, the encoding name for package inputenc must be changed:

```
\usepackage[latin1]{inputenc} \rightarrow \usepackage[utf8]{inputenc}
```

Of course, unless you are using some clever editor that knows package inputenc, recodes the file and adjusts the option at the same time. But most editors can perhaps recode the file, but they let the option untouched.

Therefore package selinput chooses another way for specifying the input encoding. The encoding name is not needed at all. Some 8-bit characters are identified by their glyph name and the package chooses an appropriate encoding, example:

1.2 User interface

```
\SelectInputEncodingList \{\langle encoding \ list \rangle\}
```

\SelectInputEncodingList expects a comma separated list of encoding names. Example:

```
\SelectInputEncodingList{utf8,ansinew,mac-roman}
```

The encodings of package inputenx are used as default.

```
\SelectInputMappings \{\langle mapping pairs \rangle\}
```

A mapping pair consists of a glyph name and its input character:

```
\SelectInputMappings{
  adieresis={ä},
  germandbls={ß},
  Euro={€},
```

The supported glyph names can be found in file ix-name.def of project inputenx [1]. The names are basically taken from Adobe's glyphlists [2, 3]. As many pairs are needed as necessary to identify the encoding. Example with insufficient pairs:

```
\SelectInputEncodingSet{latin1,latin9}
\SelectInputMappings{
  adieresis={\vec{a}},
  germandbls={\vec{b}},
}
Umlauts: \vec{A\vec{O}\vec{U}\vec{a}\vec{O}\vec{U}\vec{a}\vec{O}\vec{U}\vec{a}\vec{U}\vec{a}\vec{U}\vec{a}\vec{U}\vec{a}\vec{U}\vec{a}\vec{U}\vec{A}\vec{U}\vec{U}\vec{U}\vec{A}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\vec{U}\ve
```

The first encoding latin1 passes the constraints given by the mapping pairs. However the Euro symbol is not part of the encoding. Thus a mapping pair with the Euro symbol solves the problem. In fact the symbol alone already succeeds in selecting between latin1 and latin9:

```
\SelectInputEncodingSet{latin1,latin9}
\SelectInputMappings{
  Euro={€},
}
Umlauts: ÄÖÜäöüß and Euro: €
```

1.3 Options

warning: The selected encoding is written by \PackageInfo into the .log file only. Option warning changes it to \PackageWarning. Then the selected encoding is shown on the terminal as well.

ucs: The encoding file utf8x of package \ucs requires that the package itself is loaded before. If the package is not loaded, then the option ucs will load package ucs if the detected encoding is UTF-8 (limited to the preamble, packages cannot be loaded later).

utf8=...: The option allows to specify other encoding files for UTF-8 than IATEX's utf8.def. For example, utf8=utf-8 will load utf-8.def instead.

1.4 Encodings

Package stringenc [4] is used for testing the encoding. Thus the encoding name must be known by this package. Then the found encoding is loaded by \inputencoding by package inputenc or \InputEncoding if package inputenx is loaded.

The supported encodings are present in the encoding list, thus usually the encoding names do not matter. If the list is set by \SelectInputEncodingList, then you can use the names that work for package inputenc and are known by package stringenc, for example: latin1, x-iso-8859-1. Encoding file names of package inputenx are prefixed with x-. The prefix can be dropped, if package inputenx is loaded.

2 Implementation

```
1 (*package)
2 \NeedsTeXFormat{LaTeX2e}
3 \ProvidesPackage{selinput}
4   [2007/09/09 v1.2 Semi-automatic input encoding detection (HO)]%
5 \RequirePackage{inputenc}
6 \RequirePackage{kvsetkeys}[2006/10/19]
7 \RequirePackage{stringenc}[2007/06/16]
8 \RequirePackage{kvoptions}
```

```
\SelectInputEncodingList
```

```
9 \newcommand*{\SelectInputEncodingList}{%
                            \let\SIE@EncodingList\@empty
                            \kvsetkeys{SelInputEnc}%
                        12 }
\SelectInputMappings
                        13 \newcommand*{\SelectInputMappings}[1]{%
                            \SIE@LoadNameDefs
                        14
                        15
                           \let\SIE@StringUnicode\@empty
                           \let\SIE@StringDest\@empty
                        16
                            \kvsetkeys{SelInputMap}{#1}%
                        17
                            \ifx\\SIE@StringUnicode\SIE@StringDest\\%
                              \PackageError{selinput}{%
                        19
                        20
                                No mappings specified%
                              }\@ehc
                        21
                            \else
                        22
                              \EdefUnescapeHex\SIE@StringUnicode\SIE@StringUnicode
                        23
                              \let\SIE@Encoding\@empty
                        24
                              \@for\SIE@EncodingTest:=\SIE@EncodingList\do{%
                        25
                        26
                                \ifx\SIE@Encoding\@empty
                        27
                                  \StringEncodingConvertTest\SIE@temp\SIE@StringUnicode
                        28
                                                              {utf16be}\SIE@EncodingTest{%
                        29
                                     \ifx\SIE@temp\SIE@StringDest
                        30
                                       \let\SIE@Encoding\SIE@EncodingTest
                                     \fi
                        31
                                  }{}%
                        32
                                \fi
                        33
                              }%
                        34
                              \ifx\SIE@Encoding\@empty
                        35
                                \StringEncodingConvertTest\SIE@temp\SIE@StringDest
                        36
                                                            {ascii}{utf16be}{%
                        37
                                  \def\SIE@Encoding{ascii}%
                        38
                        39
                                  \SIE@Info{selinput}{%
                        40
                                    Matching encoding not found, but input characters%
                        41
                                    \MessageBreak
                        42
                                    are 7-bit (possibly editor replacements).%
                        43
                                    \MessageBreak
                                    Hence using ascii encoding%
                        44
                                  ጉ%
                        45
                                }{}%
                        46
                        47
                              \ifx\SIE@Encoding\@empty
                        48
                                 \PackageError{selinput}{%
                        49
                                  Cannot find a matching encoding%
                        50
                                \ \ (ehd)
                        51
                        52
                              \else
                                \ifx\SIE@Encoding\SIE@EncodingUTFviii
                        53
                                  \SIE@LoadUnicodePackage
                        54
                                  \ifx\SIE@UseUTFviii\@empty
                        55
                                  \else
                        56
                                     \let\SIE@Encoding\SIE@UseUTFviii
                        57
                                  \fi
                        58
                                \fi
                        59
                        60
                                 \begingroup\expandafter\expandafter\expandafter\endgroup
                        61
                                 \expandafter\ifx\csname InputEncoding\endcsname\relax
                        62
                                  \inputencoding\SIE@Encoding
                                \else
                        63
                                  \InputEncoding\SIE@Encoding
                        64
                        65
                                 \SIE@Info{selinput}{Encoding `\SIE@Encoding' selected}%
                        66
                        67
                              \fi
```

```
\fi
                              68
                              69 }
        \SIE@LoadNameDefs
                              70 \def\SIE@LoadNameDefs{%
                              71
                                   \begingroup
                                     \endlinechar=\m@ne
                              72
                              73
                                     \catcode92=0 % backslash
                              74
                                     \catcode123=1 % left curly brace/beginning of group
                              75
                                     \catcode125=2 % right curly brace/end of group
                              76
                                     \catcode37=14 % percent/comment character
                              77
                                     \@makeother\[%
                                     \@makeother\]%
                              78
                                     \@makeother\.%
                              79
                                     \@makeother\(%
                              80
                                     \@makeother\)%
                              81
                              82
                                     \@makeother\/%
                              83
                                     \ensuremath{\tt @makeother}\-\%
                              84
                                     \let\InputenxName\SelectInputDefineMapping
                              85
                                     \InputIfFileExists{ix-name.def}{}{%
                              86
                                       \PackageError{selinput}{%
                              87
                                         Missing `ix-name.def' (part of package `inputenx')%
                              88
                                       \ \ (dehd)
                              89
                                     }%
                                     \global\let\SIE@LoadNameDefs\relax
                              90
                              91
                                   \endgroup
                              92 }
\SelectInputDefineMapping
                              93 \newcommand*{\SelectInputDefineMapping}[1]{%
                                   \expandafter\gdef\csname SIE@@#1\endcsname
                              94
                              95 }
                              96 \kv@set@family@handler{SelInputMap}{%
                                   \@onelevel@sanitize\kv@key
                              97
                              98
                                   \ifx\kv@value\relax
                              99
                                     \PackageError{selinput}{%
                                       Missing input character for `\kv@key'%
                             100
                             101
                                     }\@ehc
                             102
                                   \else
                                     \@onelevel@sanitize\kv@value
                             103
                                     \ifx\kv@value\@empty
                             104
                                       \PackageError{selinput}{%
                             105
                             106
                                         Input character got lost?\MessageBreak
                             107
                                         Missing input character for `\kv@key'%
                             108
                                       \ \ensuremath{\mbox{Qehc}}
                             109
                                     \else
                             110
                                       \@ifundefined{SIE@@\kv@key}{%
                             111
                                         \PackageWarning{selinput}{%
                                           Missing definition for `\kv@key'%
                             112
                                         }%
                             113
                                       }{%
                             114
                                          \edef\SIE@StringDest{%
                             115
                                            \SIE@StringDest
                             116
                                            \kv@value
                             117
                             118
                                         }%
                             119
                                          \edef\SIE@StringUnicode{%
                             120
                                            \SIE@StringUnicode
                             121
                                            \csname SIE@@\kv@key\endcsname
                                         }%
                             122
                                       }%
                             123
                             124
                                     \fi
```

\fi

125

```
126 }
                          127 \kv@set@family@handler{SelInputEnc}{%
                          128
                               \@onelevel@sanitize\kv@key
                               \ifx\kv@value\relax
                          129
                                 \ifx\SIE@EncodingList\@empty
                          130
                                    \let\SIE@EncodingList\kv@key
                          131
                          132
                          133
                                    \edef\SIE@EncodingList{\SIE@EncodingList,\kv@key}%
                          134
                                 \fi
                          135
                               \else
                                 \@onelevel@sanitize\kv@value
                          136
                          137
                                 \PackageError{selinput}{%
                          138
                                   Illegal key value pair (\kv@key=\kv@value)\MessagBreak
                          139
                                   in encoding list%
                                 \ \ \@ehc
                          140
                               \fi
                          141
                          142 }
\SIE@LoadUnicodePackage
                          143 \def\SIE@LoadUnicodePackage{%
                               \@ifpackageloaded\SIE@UnicodePackage{}{%
                                 \RequirePackage\SIE@UnicodePackage\relax
                          145
                          146
                               \SIE@PatchUCS
                          147
                               \global\let\SIE@LoadUnicodePackage\relax
                          148
                          149 }
                          150 \let\SIE@show\show
                          151 \def\SIE@PatchUCS{%
                          152
                               \AtBeginDocument{%
                          153
                                 \expandafter\ifx\csname ver@ucsencs.def\endcsname\relax
                          154
                                 \else
                                    \let\show\SIE@show
                          155
                                 \fi
                          156
                               }%
                          157
                          158 }
                          159 \SIE@PatchUCS
                          160 \AtBeginDocument{%
                          161
                               \let\SIE@LoadUnicodePackage\relax
                          162 }
   \SIE@EncodingUTFviii
                          163 \def\SIE@EncodingUTFviii{utf8}
                          164 \@onelevel@sanitize\SIE@EncodingUTFviii
  \SIE@EncodingUTFviiix
                          165 \def\SIE@EncodingUTFviiix{utf8x}
                          166 \ConelevelCsanitize\SIECEncodingUTFviiix
                          167 \let\SIE@UnicodePackage\@empty
                          168 \let\SIE@UseUTFviii\@empty
                          169 \let\SIE@Info\PackageInfo
                          170 \SetupKeyvalOptions{%
                              family=SelInput,%
                              prefix=SelInput@%
                          173 }
                          174 \define@key{SelInput}{utf8}{%
                              \def\SIE@UseUTFviii{#1}%
                               \@onelevel@sanitize\SIE@UseUTFviii
                          176
                          177 }
                          178 \DeclareBoolOption{ucs}
                          179 \DeclareVoidOption{warning}{%
```

```
\let\SIE@Info\PackageWarning
                   180
                   181 }
                   182 \ProcessKeyvalOptions{SelInput}
                   183 \ifSelInput@ucs
                         \def\SIE@UnicodePackage{ucs}%
                   185
                         \ifx\SIE@UseUTFviii\@empty
                   186
                          \let\SIE@UseUTFviii\SIE@EncodingUTFviiix
                        \fi
                   187
                   188 \ensuremath{\setminus} else
                        \ifx\SIE@UseUTFviii\@empty
                   189
                           \@ifpackageloaded{ucs}{%
                   190
                             \let\SIE@UseUTFviii\SIE@EncodingUTFviiix
                   191
                           }{%
                   192
                   193
                             \let\SIE@UseUTFviii\SIE@EncodingUTFviii
                   194
                           }%
                   195
                        \fi
                   196 \fi
\SIE@EncodingList
                   197 \edef\SIE@EncodingList{%
                       utf8,%
                   198
                   199 x-iso-8859-1,%
                   200 x-iso-8859-15,%
                   201 x-cp1252,% ansinew
                   202 x-mac-roman,%
                   203 x-iso-8859-2,%
                   204 x-iso-8859-3,%
                   205 x-iso-8859-4,%
                   206 x-iso-8859-5,%
                       x-iso-8859-6,%
                   207
                   208 x-iso-8859-7,%
                   209 x-iso-8859-8,%
                   210
                        x-iso-8859-9,%
                       x-iso-8859-10,%
                   211
                   212
                        x-iso-8859-11,%
                   213 x-iso-8859-13,%
                   214 x-iso-8859-14,%
                   215 x-iso-8859-15,%
                   216 x-mac-centeuro,%
                   217 x-mac-cyrillic,%
                   218 x-koi8-r,%
                   219 x-cp1250,%
                   220 x-cp1251,%
                   221 x-cp1257,%
                   222 x-cp437,%
                   223 x-cp850,%
                   224 x-cp852,%
                   225 x-cp855,%
                   226 x-cp858,%
                   227 x-cp865,%
                        x-cp866,%
                   228
                   229
                        x-nextstep,%
                        x-dec-mcs%
                   231 }%
                   232 \verb|\cone| evel@sanitize\\| SIE@EncodingList\\|
                   233 (/package)
                   3
                        Test
                   234 (*test)
                   235 \NeedsTeXFormat{LaTeX2e}
```

```
236 \documentclass{minimal}
237 \usepackage{textcomp}
238 \usepackage{qstest}
239 (*test1 j test2 j test3)
240 \makeatletter
241 \let\BeginDocumentText\@empty
242 \def\TestEncoding#1#2{%
243
     \SelectInputMappings{#2}%
244
     \Expect*{\SIE@Encoding}{#1}%
245
     \Expect*{\inputencodingname}{#1}%
     \g@addto@macro\BeginDocumentText{%
246
247
        \SelectInputMappings{#2}%
248
        \Expect*{\SIE@Encoding}{#1}%
249
       \textbf{\SIE@Encoding:} %
250
        \kvsetkeys{test}{#2}\par
     }%
251
252 }
253 \def\TestKey#1#2{%
    \define@key{test}{#1}{%
254
255
       \sbox0{##1}%
256
        \sbox2{#2}%
257
        \Expect*{wd:\the\wd0, ht:\the\ht0, dp:\the\dp0}%
258
               *{wd:\the\wd2, ht:\the\ht2, dp:\the\dp2}%
259
        [#1=##1] % hash-ok
     3%
260
261 }
262 \RequirePackage{keyval}
263 \TestKey{adieresis}{\"a}
264 \TestKey{germandbls}{\ss}
265 \TestKey{Euro}{\texteuro}
266 \makeatother
267 \usepackage[
268 warning,%
269 (test2) utf8=utf-8,
270 (test3) ucs,
271 ]{selinput}
272 (test1 j test3) \inputencoding{ascii}
273 (test2)\inputencoding{utf-8}
274 (test3)\usepackage{ucs}
275 \begin{qstest}{preamble}{}
276
     \TestEncoding{x-iso-8859-15}{\%}
       adieresis=^^e4,%
277
       germandbls=^^df,%
278
       Euro=^^a4,%
279
280
281
     \TestEncoding{x-cp1252}{%}
       adieresis=^^e4,%
282
       germandbls=^^df,%
283
       Euro=^^80,%
284
     7%
285
286 (test1)
           \TestEncoding{utf8}{%
287 (test2)
           \TestEncoding{utf-8}{%
288 (test3) \TestEncoding{utf8x}{%
       adieresis=^^c3^^a4,%
289
290
       germandbls=^^c3^^9f,%
291 (!test2) Euro=^^e2^^82^^ac,
292
    }%
293 \end{qstest}
294 \langle test3 \rangle \cdot let \cdot ifUnicodeOptiongraphics \cdot iffalse
295 \begin{document}
296 \begin{qstest}{document}{}
297 (test3)\makeatletter
```

```
\BeginDocumentText
298
299 \end{qstest}
300 (/test1 j test2 j test3)
301 (*test4)
302 \usepackage[warning,ucs]{selinput}
303 \SelectInputMappings{%
       adieresis=^^c3^^a4,%
304
       germandbls=^^c3^^9f,%
305
       Euro=^^e2^^82^^ac,%
306
307 }
308 \begin{qstest}{encoding}{}
    \Expect*{\inputencodingname}{utf8x}%
310 \end{qstest}
311 \begin{document}
312 adieresis=^^c3^^a4, %
    germandbls=^^c3^^9f, %
313
    Euro=^^e2^^82^^ac%
314
315 (/test4)
316 (*test5)
317 \usepackage[warning,ucs]{selinput}
318 \SelectInputMappings{%
       adieresis={\"a},%
319
320
       germandbls={{\ss}},%
321
       Euro=\texteuro{},%
322 }
323 \begin{qstest}{encoding}{}
324 \Expect*{\inputencodingname}{ascii}%
325 \end{qstest}
326 \begin{document}
327 adieresis=\{\"a\}, \%
328 germandbls={{\ss}}, %
    Euro=\texteuro{}%
329
330 (/test5)
331 \end{document}
332 (/test)
```

4 Installation

4.1 Download

Package. This package is available on CTAN¹:

CTAN:macros/latex/contrib/oberdiek/selinput.dtx The source file.

CTAN:macros/latex/contrib/oberdiek/selinput.pdf Documentation.

Bundle. All the packages of the bundle 'oberdiek' are also available in a TDS compliant ZIP archive. There the packages are already unpacked and the documentation files are generated. The files and directories obey the TDS standard.

CTAN:install/macros/latex/contrib/oberdiek.tds.zip

TDS refers to the standard "A Directory Structure for TEX Files" (CTAN:tds/tds.pdf). Directories with texmf in their name are usually organized this way.

4.2 Bundle installation

Unpacking. Unpack the oberdiek.tds.zip in the TDS tree (also known as texmf tree) of your choice. Example (linux):

```
unzip oberdiek.tds.zip -d ~/texmf
```

¹ftp://ftp.ctan.org/tex-archive/

Script installation. Check the directory TDS:scripts/oberdiek/ for scripts that need further installation steps. Package attachfile2 comes with the Perl script pdfatfi.pl that should be installed in such a way that it can be called as pdfatfi. Example (linux):

```
chmod +x scripts/oberdiek/pdfatfi.pl
cp scripts/oberdiek/pdfatfi.pl /usr/local/bin/
```

4.3 Package installation

Unpacking. The .dtx file is a self-extracting docstrip archive. The files are extracted by running the .dtx through plain T_EX:

```
tex selinput.dtx
```

TDS. Now the different files must be moved into the different directories in your installation TDS tree (also known as texmf tree):

If you have a docstrip.cfg that configures and enables docstrip's TDS installing feature, then some files can already be in the right place, see the documentation of docstrip.

4.4 Refresh file name databases

If your T_EX distribution (te T_EX , mik T_EX , ...) relies on file name databases, you must refresh these. For example, te T_EX users run texhash or mktexlsr.

4.5 Some details for the interested

Attached source. The PDF documentation on CTAN also includes the .dtx source file. It can be extracted by AcrobatReader 6 or higher. Another option is pdftk, e.g. unpack the file into the current directory:

```
pdftk selinput.pdf unpack_files output .
```

Unpacking with IATEX. The .dtx chooses its action depending on the format:

plain TeX: Run docstrip and extract the files.

LATEX: Generate the documentation.

If you insist on using \LaTeX for docstrip (really, docstrip does not need \LaTeX), then inform the autodetect routine about your intention:

```
latex \let\install=y\input{selinput.dtx}
```

Do not forget to quote the argument according to the demands of your shell.

Generating the documentation. You can use both the .dtx or the .drv to generate the documentation. The process can be configured by the configuration file ltxdoc.cfg. For instance, put this line into this file, if you want to have A4 as paper format:

\PassOptionsToClass{a4paper}{article}

An example follows how to generate the documentation with pdfIAT_FX:

```
pdflatex selinput.dtx
makeindex -s gind.ist selinput.idx
pdflatex selinput.dtx
makeindex -s gind.ist selinput.idx
pdflatex selinput.dtx
```

5 Catalogue

The following XML file can be used as source for the TeX Catalogue. The elements caption and description are imported from the original XML file from the Catalogue. The name of the XML file in the Catalogue is selinput.xml.

```
333 (*catalogue)
334 <?xml version='1.0' encoding='us-ascii'?>
335 <!DOCTYPE entry SYSTEM 'catalogue.dtd'>
336 <entry datestamp='$Date$' modifier='$Author$' id='selinput'>
    <name>selinput</name>
337
     <caption>Semi-automatic detection of input encoding.</caption>
338
339
     <authorref id='auth:oberdiek'/>
     <copyright owner='Heiko Oberdiek' year='2007'/>
340
     <license type='lppl1.3'/>
341
     <version number='1.2'/>
342
343
     <description>
       This package selects the input encoding by specifying pairs
344
       of input characters and their glyph names.
345
346
       The package is part of the <xref refid='oberdiek'>oberdiek</xref>
347
       bundle.
348
    </description>
349
    <documentation details='Package documentation'</pre>
350
         href='ctan:/macros/latex/contrib/oberdiek/selinput.pdf'/>
351
352 <ctan file='true' path='/macros/latex/contrib/oberdiek/selinput.dtx'/>
353 <miktex location='oberdiek'/>
354 <texlive location='oberdiek'/>
355 <install path='/macros/latex/contrib/oberdiek/oberdiek.tds.zip'/>
356 </entry>
357 (/catalogue)
```

6 References

- [1] Heiko Oberdiek: *The inputenx package*; 2007-04-11 v1.1; CTAN:macros/latex/contrib/oberdiek/inputenx.pdf.
- [2] Adobe: Adobe Glyph List; 2002-09-20 v2.0; http://partners.adobe.com/public/developer/en/opentype/glyphlist.txt.
- [3] Adobe: Adobe Glyph List For New Fonts; 2005-11-18 v1.5; http://partners.adobe.com/public/developer/en/opentype/aglfn13.txt.
- [4] Heiko Oberdiek: *The stringenc package*; 2007-06-16 v1.1; CTAN:macros/latex/contrib/oberdiek/stringenc.pdf.

7 History

[2007/06/16 v1.0]

 $\bullet~$ First version.

[2007/06/20 v1.1]

• Requested date for package stringenc fixed.

$[2007/09/09\ v1.2]$

• Line end fixed.

8 Index

Numbers written in italic refer to the page where the corresponding entry is described; numbers underlined refer to the code line of the definition; plain numbers refer to the code lines where the entry is used.

Symbols	\end 293, 299, 310, 325, 331
\"	\endcsname 61, 94, 121, 153 \endlinechar 72
\)	\Expect 244, 245, 248, 257, 309, 324
\	(Expect 211, 210, 210, 201, 503, 521
\ 79	${f G}$
V 82	\g@addto@macro 246
\@ehc 21, 101, 108, 140	\gdef 94
\@ehd 51, 88	
$\ensuremath{\texttt{Qempty}}$ $10, 15, 16, 24, 26, 35, 48,$	Н
55, 104, 130, 167, 168, 185, 189, 241	\ht 257, 258
\@for	T
\@ifpackageloaded 144, 190 \@ifundefined 110	I 204
\Qmakeother 77, 78, 79, 80, 81, 82, 83	\iffalse 294 \ifSelInput@ucs 183
\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\	\ifUnicodeOptiongraphics 294
103, 128, 136, 164, 166, 176, 232	\ifx 18, 26, 29, 35, 48, 53, 55,
\[61, 98, 104, 129, 130, 153, 185, 189
\\ 18	\InputEncoding 64
\] 78	\inputencoding 62, 272, 273
	\inputencodingname 245, 309, 324
\mathbf{A}	
	\InputenxName 84
\AtBeginDocument 152, 160	\InputIfFileExists
	\InputIfFileExists 85
\AtBeginDocument 152, 160	\InputIfFileExists
$\begin{tabular}{lllllllllllllllllllllllllllllllllll$	\InputIfFileExists
\AtBeginDocument	K K K K K K K K K K
\AtBeginDocument	K K K Kv@key
\AtBeginDocument	K K We We We We We We
AtBeginDocument	K K We will be w
AtBeginDocument	K K We We We We We We
AtBeginDocument	K K Wkv@key
AtBeginDocument	K K We We We We We We
AtBeginDocument	K K Wkv@key
AtBeginDocument	K K K K K K K K K K
AtBeginDocument	K K Kv@key

P	\SIE@LoadNameDefs 14, <u>70</u>
$\verb \PackageError 19, 49, 86, 99, 105, 137 $	$\SIE@LoadUnicodePackage$. 54, $\underline{143}$, 161
\PackageInfo 169	\SIE@PatchUCS 147, 151, 159
\PackageWarning 111, 180	\SIE@show 150, 155
\par 250	\SIE@StringDest 16, 18, 29, 36, 115, 116
\ProcessKeyvalOptions 182	\SIE@StringUnicode 15, 23, 27, 119, 120
\ProvidesPackage 3	\SIE@temp 27, 29, 36
D	\SIE@UnicodePackage 144, 145, 167, 184
R	\SIE@UseUTFviii 55, 57, 168,
\RequirePackage 5, 6, 7, 8, 145, 262	175, 176, 185, 186, 189, 191, 193
S	\ss 264, 320, 328
\sbox 255, 256	\StringEncodingConvertTest 27, 36
\SelectInputDefineMapping 84, 93	
(bottoodinpadbottmonapping Oi, vo	
\SelectInputEncodingList 2, 9	${f T}$
	\TestEncoding
$\verb \SelectInputEncodingList 2, \underline{9}$	_
$\label{eq:localization} $$\operatorname{SelectInputEncodingList} \dots \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ $	\TestEncoding
$\label{eq:localization} $$ \SelectInputEncodingList \dots 2, \underline{9} $$ \SelectInputMappings \dots \\ \dots 2, \underline{13}, 243, 247, 303, 318 $$$	\TestEncoding 242, 276, 281, 286, 287, 288
$\label{eq:localization} $$ \end{array} $$\end{array} $$\end{array} $$array$	\TestEncoding 242, 276, 281, 286, 287, 288 \TestKey 253, 263, 264, 265
eq:localized-loca	\TestEncoding 242, 276, 281, 286, 287, 288 \TestKey 253, 263, 264, 265 \textbf 249
$\label{eq:linear_select_select} $$ \end{array} $$$	\TestEncoding 242, 276, 281, 286, 287, 288 \TestKey 253, 263, 264, 265 \textbf 249 \texteuro 265, 321, 329
$\label{eq:linear_select_select} $$ \end{array} $$$	\TestEncoding 242, 276, 281, 286, 287, 288 \TestKey 253, 263, 264, 265 \textbf 249 \texteuro 265, 321, 329
\SelectInputEncodingList 2, 9 \SelectInputMappings	\TestEncoding 242, 276, 281, 286, 287, 288 \TestKey 253, 263, 264, 265 \textbf 265, 321, 329 \the 257, 258
$\label{eq:linear_select_select} $$ \end{array} $$$	\TestEncoding
$\label{eq:linear_select_select_select} $$ \end{array} $$ ar$	\TestEncoding