

# xCJK2uni 宏包

李清

sobenlee@gmail.com

2018/01/28 v0.7\*

## 第 1 节 简介

xCJK2uni 是一个 L<sup>A</sup>T<sub>E</sub>X 宏包, 提供了将 CJK 文字转换成 Unicode 的功能。以宏包形式实现 gbk2uni 程序的功能, 使得使用 CJK 和 hyperref 宏包时, 可以得到正确的书签。并提供了 /ToUnicode 映射文件, 使得使用 GBK 等编码时, 配合 cmap 宏包, 用 pdfL<sup>A</sup>T<sub>E</sub>X 编译得到的 pdf 文件支持查找、复制和粘贴。

xCJK2uni 只适用于下一节说明的有效编码和以 pdfL<sup>A</sup>T<sub>E</sub>X 作为编译引擎的工作方式。

xCJK2uni 依赖 l3kernel。应该与 CJK 或 CCT 等配合使用, 但 xCJK2uni 不会自动引入它们。

## 第 2 节 基本命令

---

`\useCJKencmap` `\useCJKencmap {(Bg5|Bg5+|GB|GBK|JIS|KS)}`

设置当前的编码环境。以上是 xCJK2uni 支持的文字编码, 缺省使用 GBK。xCJK2uni 会在 CJK 环境内自动更新支持的编码。

---

`\CJKchartouni` \* `\CJKchartouni {(单个CJK文字)}`

将文字转换成 16 进制形式的 Unicode。例如 `\CJKchartouni{一}` 将得到 4E00。将它展开两次就可以得到结果。

---

`\CJKsfdtouni` \* `\CJKsfdtouni {(plane)} {(slot)}`

`{plane}` 表示由 .sfd 文件定义的字体的编号, `{slot}` 表示在子字体中的位置, 这个命令将得到该坐标位置的 Unicode。将它展开两次就可以得到结果。如果在当前编码下, 该坐标位置不合法, 将展开为空。

## 第 3 节 代码实现

### 3.1 xCJK2uni

```
1 <(*package)>
2 <@@=CJKtu>
3 \msg_new:nnn { xCJK2uni } { l3-too-old }
4 {
5   Support~package~'expl3'~too~old. \\\
6   Please~update~an~up~to~date~version~of~the~bundles\\
7   'l3kernel'~and~'l3packages'\\
8   using~your~TeX~package~manager~or~from~CTAN.
```

---

\*ctex-kit rev. 87875c2.

```

9 }
10 \ifpackagelater { expl3 } { 2017/12/16 } { }
11 { \msg_error:nn { xCJK2uni } { l3-too-old } }
12 \sys_if_engine_pdfTeX:F
13 {
14   \msg_new:nnn { xCJK2uni } { pdflatex }
15     { The~xCJK2uni~package~is~only~supported~in~pdfTeX. }
16   \msg_critical:nn { xCJK2uni } { pdflatex }
17 }
18 \tl_new:N \l__CJKtu_sfd_tl
19 \prop_new:N \c__CJKtu_sfd_prop
20 \seq_new:N \c__CJKtu_encoding_seq
21 \int_new:N \l__CJKtu_i_min_int
22 \int_new:N \l__CJKtu_i_max_int
23 \int_new:N \l__CJKtu_ii_min_int
24 \int_new:N \l__CJKtu_ii_max_int
25 \int_new:N \l__CJKtu_gap_begin_int
26 \int_new:N \l__CJKtu_gap_end_int
27 \int_new:N \l__CJKtu_plane_int
28 \int_set:Nn \l__CJKtu_i_max_int { "FE }
29 \int_set:Nn \l__CJKtu_ii_max_int { "FE }
\__CJKtu_set_enc:nnnnnn 30 \cs_new_protected_nopar:Npn \__CJKtu_set_enc:nnnnnn #1#2#3#4#5#6
31 {
32   \seq_gput_right:Nn \c__CJKtu_encoding_seq {#1}
33   \prop_gput:Nnn \c__CJKtu_sfd_prop {#1} {#2}
34   \int_const:cn { c__CJKtu_ #1 _i_min_int } {#3}
35   \int_const:cn { c__CJKtu_ #1 _ii_min_int } {#4}
36   \tl_if_empty:nTF {#5}
37     {
38       \int_const:cn { c__CJKtu_ #1 _gap_begin_int }
39         { \l__CJKtu_ii_max_int + \c_one }
40       \int_const:cn { c__CJKtu_ #1 _gap_end_int }
41         { \use:c { c__CJKtu_ #1 _gap_begin_int } }
42       \int_const:cn { c__CJKtu_ #1 _plane_int }
43         { \l__CJKtu_ii_max_int - (#4) + \c_one }
44     }
45     {
46       \int_const:cn { c__CJKtu_ #1 _gap_begin_int } { #5 + \c_one }
47       \int_const:cn { c__CJKtu_ #1 _gap_end_int } {#6}
48       \int_const:cn { c__CJKtu_ #1 _plane_int }
49         { (#5) - (#4) + \l__CJKtu_ii_max_int - (#6) + \c_two }
50     }
51   \cs_if_exist_use:NTF \CJKaddEncHook
52     { {#1} { \__CJKtu_enc_map:nn {#1} {#2} } }
53     {
54       \AtBeginDocument
55       {
56         \cs_if_exist_use:NT \CJKaddEncHook
57           { {#1} { \__CJKtu_enc_map:nn {#1} {#2} } }
58       }
59     }
60 }
\__CJKtu_enc_map:nn 61 \cs_new_protected_nopar:Npn \__CJKtu_enc_map:nn #1#2
62 {
63   \int_set_eq:Nc \l__CJKtu_i_min_int { c__CJKtu_ #1 _i_min_int }
64   \int_set_eq:Nc \l__CJKtu_ii_min_int { c__CJKtu_ #1 _ii_min_int }
65   \int_set_eq:Nc \l__CJKtu_gap_begin_int { c__CJKtu_ #1 _gap_begin_int }
66   \int_set_eq:Nc \l__CJKtu_gap_end_int { c__CJKtu_ #1 _gap_end_int }
67   \int_set_eq:Nc \l__CJKtu_plane_int { c__CJKtu_ #1 _plane_int }
68   \cs_if_exist:cF { CJKtu_ #1 _sfd_map:nn }
69     { \group_begin: \file_input:n { xCJK2uni-#2.def } \group_end: }
70   \cs_set_eq:Nc \CJKtu_sfd_map:nn { CJKtu_ #1 _sfd_map:nn }
71 }
\__CJKtu_use_enc_map:n 72 \cs_new_protected_nopar:Npn \__CJKtu_use_enc_map:n #1
73 {

```

```

74 \prop_get:NxNTF \c__CJKtu_sfd_prop {#1} \l__CJKtu_sfd_tl
75 { \__CJKtu_enc_map:nn {#1} { \l__CJKtu_sfd_tl } }
76 { \msg_error:nnx { xCJK2uni } { invalid-encoding } {#1} }
77 }
78 \msg_new:nnn { xCJK2uni } { invalid-encoding }
79 {
80 The~encoding~`#1'~is~invalid.\\
81 Only~\seq_use:Nnnn \c__CJKtu_encoding_seq { ~and~ } { ,~ } { ,~and~ } ~
82 are~supported.
83 }
84 \prg_generate_conditional_variant:Nnn \prop_get:NnN { Nx } { TF }

85 \__CJKtu_set_enc:nnnnnn { GB } { UGB } { "A1 } { "A1 } { } { }
86 \__CJKtu_set_enc:nnnnnn { JIS } { UJIS } { "A1 } { "A1 } { } { }
87 \__CJKtu_set_enc:nnnnnn { KS } { UKS } { "A1 } { "A1 } { } { }
88 \__CJKtu_set_enc:nnnnnn { Bg5 } { UBig5 } { "A1 } { "40 } { "7E } { "A1 }
89 \__CJKtu_set_enc:nnnnnn { Bg5+ } { UBg5plus } { "81 } { "40 } { "7E } { "80 }
90 \__CJKtu_set_enc:nnnnnn { GBK } { UGBK } { "81 } { "40 } { "7E } { "80 }
\__CJKtu_sfd_to_unicode:nn 91 \cs_new_nopar:Npn \__CJKtu_sfd_to_unicode:nn #1#2
92 {
93 \exp:w \exp_last_unbraced:Nf
94 \__CJKtu_get_unicode:w \CJKtu_sfd_map:nn {#1} {#2} \exp_end:
95 }

\__CJKtu_get_unicode:w 96 \cs_new_nopar:Npn \__CJKtu_get_unicode:w #1 \exp_end:
97 {
98 \tl_if_empty:nTF {#1}
99 { \exp_end: }
100 { \__CJKtu_get_unicode_aux:w #1 \q_stop }
101 }
102 \cs_new_nopar:Npn \__CJKtu_get_unicode_aux:w #1 \or: #2 \q_stop
103 { \or: #2 \exp_end: #1 }

\__CJKtu_char_to_unicode:n 104 \group_begin:
105 \char_set_catcode_active:n { "7F }
106 \cs_new_nopar:Npn \__CJKtu_char_to_unicode:n #1
107 {
108 \exp:w
109 \int_compare:nNnTF { \tl_count:n {#1} } = \c_two
110 { \CJKtu_byte:NN #1 }
111 {
112 \tl_if_head_eq_meaning:nNTF {#1} ^^7f
113 { \__CJKtu_byte:wNwnw #1 }
114 { \__CJKtu_encoding_error: }
115 }
116 }
117 \cs_new_nopar:Npn \__CJKtu_byte:wNwnw ^^7f #1 ^^7f #2 ^^7f
118 { \CJKtu_byte_map:nn { `#1 } {#2} }
119 \cs_new_nopar:Npn \__CJKtu_byte:Nwnw #1 ^^7f #2 ^^7f
120 { \CJKtu_byte_map:nn { `#1 } {#2} }

\CJKtu_bookmarks_hook: 121 \cs_new_protected_nopar:Npn \CJKtu_bookmarks_hook:
122 {
123 \cs_set_eq:NN ^^7f \__CJKtu_byte:Nwnw
124 \cs_set_eq:NN \CJK@XX \CJKtu_byte:NN
125 \cs_set_eq:NN \CJKchar \__CJKtu_byte:wnn
126 \cs_set_eq:NN \CCTchar \CJKtu_byte:nn
127 \cs_set_eq:NN \@CCTSetChar \__CJKtu_CCT_byte:NNN
128 \cs_set_eq:NN \__CJKtu_get_unicode:n \__CJKtu_UTF_xvi_be:n
129 }
130 \group_end:

\CJKtu_byte:NN 131 \cs_new_nopar:Npn \CJKtu_byte:NN #1#2
\__CJKtu_CCT_byte:NNN 132 { \CJKtu_byte:nn { `#1 } { `#2 } }
133 \cs_new_nopar:Npn \__CJKtu_CCT_byte:NNN #1#2#3
134 { \CJKtu_byte:nn { `#1 } { `#3 } }

\__CJKtu_byte:wnn 135 \cs_new_nopar:Npn \__CJKtu_byte:wnn #1# { \CJKtu_byte:nn }

```

```

\CJKtu_byte:nn 136 \cs_new_nopar:Npn \CJKtu_byte:nn #1#2
137 {
138   \int_compare:nTF { \l__CJKtu_i_min_int <= #1 <= \l__CJKtu_i_max_int }
139   {
140     \int_compare:nTF { \l__CJKtu_ii_min_int <= #2 < \l__CJKtu_gap_begin_int }
141     { \CJKtu_byte_map:nn {#1} {#2} }
142     {
143       \int_compare:nTF
144       { \l__CJKtu_gap_end_int <= #2 <= \l__CJKtu_ii_max_int }
145       { \CJKtu_byte_map:nn {#1} {#2} }
146       { \__CJKtu_encoding_error: }
147     }
148   }
149   { \__CJKtu_encoding_error: }
150 }
151 \cs_new_nopar:Npn \__CJKtu_encoding_error:
152 {
153   \msg_expandable_error:nn { xCJK2uni } { encoding-error }
154   \exp_end:
155 }
156 \msg_new:nnn { xCJK2uni } { encoding-error }
157 { Wrong~encoding~scheme. }

\CJKtu_byte_map:nn 158 \cs_new_nopar:Npn \CJKtu_byte_map:nn #1#2
159 {
160   \exp_args:Nf \__CJKtu_index_map:n
161   {
162     \int_eval:n
163     {
164       ( #1 - \l__CJKtu_i_min_int ) * \l__CJKtu_plane_int
165       + ( #2 ) - \l__CJKtu_ii_min_int
166       \int_compare:nNf {#2} < \l__CJKtu_gap_begin_int
167       { + \l__CJKtu_gap_begin_int - \l__CJKtu_gap_end_int }
168     }
169   }
170 }

\__CJKtu_index_map:n 171 \cs_new_nopar:Npn \__CJKtu_index_map:n #1
172 {
173   \__CJKtu_get_unicode:n
174   {
175     \CJKtu_sfd_map:nn
176     { \int_div_truncate:nn {#1} { \c_two_hundred_fifty_six } + \c_one }
177     { \int_mod:nn {#1} { \c_two_hundred_fifty_six } }
178   }
179 }

\__CJKtu_get_unicode:n 180 \cs_new_nopar:Npn \__CJKtu_get_unicode:n #1
181 { \exp_last_unbraced:Nf \__CJKtu_get_unicode:w #1 \exp_end: }

\__CJKtu_UTF_xvi_be:n 182 \cs_new_nopar:Npn \__CJKtu_UTF_xvi_be:n #1
183 { \exp_args:Nf \__CJKtu_UTF_xvi_aux:n { \int_eval:n { "#1" } } }
184 \cs_new_nopar:Npn \__CJKtu_UTF_xvi_aux:n #1
185 {
186   \HyPsd@DecimalToOctalFirst
187   { \int_div_truncate:nn {#1} { \c_two_hundred_fifty_six } }
188   \HyPsd@DecimalToOctalSecond
189   { \int_mod:nn {#1} { \c_two_hundred_fifty_six } }
190 }

191 \cs_if_exist_use:NTF \hypersetup { { unicode , CJKbookmarks = false } }
192 { \PassOptionsToPackage { unicode , CJKbookmarks = false } { hyperref } }
193 \msg_new:nnn { xCJK2uni } { UTF-8-encoding }
194 { It~is~not~necessary~to~load~xCJK2uni~in~UTF-8~encoding. }
195 \tl_if_exist:NF \pdfstringdefPreHook { \tl_new:N \pdfstringdefPreHook }
196 \@ifpackageloaded { CJKutf8 }
197 { \msg_warning:nn { xCJK2uni } { UTF-8-encoding } }
198 {

```

```

199 \cs_new_eq:NN \CJKtu@bookmarks@hook \CJKtu_bookmarks_hook:
200 \tl_gput_right:Nn \pdfstringdefPreHook { \CJKtu@bookmarks@hook }
201 \AtBeginDocument
202 {
203   \@ifpackageloaded { CJKutf8 }
204   {
205     \cs_gset_eq:NN \CJKtu@bookmarks@hook \scan_stop:
206     \msg_warning:nn { xCJK2uni } { UTF-8-encoding }
207   }
208   { \cs_if_exist_use:NT \hypersetup { { CJKbookmarks = false } } }
209 }
210 }
\useCJKencmap 211 \cs_new_eq:NN \useCJKencmap \__CJKtu_use_enc_map:n
\CJKchartouni 212 \cs_new_eq:NN \CJKchartouni \__CJKtu_char_to_unicode:n
\CJKsfdtouni 213 \cs_new_eq:NN \CJKsfdtouni \__CJKtu_sfd_to_unicode:n
\__CJKtu_initial:N 214 \cs_new_protected_nopar:Npn \__CJKtu_initial:N #1
215 {
216   \tl_if_exist:NTF #1
217   {
218     \prop_get:NxNTF \c__CJKtu_sfd_prop {#1} \l__CJKtu_sfd_tl
219     { \__CJKtu_enc_map:nn {#1} { \l__CJKtu_sfd_tl } }
220     { \__CJKtu_use_enc_map:n { GBK } }
221   }
222   { \__CJKtu_use_enc_map:n { GBK } }
223 }
224 \exp_args:Nc \__CJKtu_initial:N { CJK @ @ @ enc }
225 </package>

```

### 3.2 xCJK2uni-make

```

226 <*make>
227 \msg_new:nnn { xCJK2uni } { file-not-found }
228 {
229   SFD file~`#2'~for~encoding~`#1'~not~found.\\
230   xCJK2uni~will~not~work!
231 }
232 \ior_new:N \g__CJKtu_sfd_ior
233 \str_new:N \g__CJKtu_path_str
234 \sys_if_engine luatex:TF
235 {
236   \str_gset:Nx \g__CJKtu_path_str
237   {
238     \lua_now_x:n
239     {
240       kpse.set_program_name("kpsewhich") ~
241       local ~ sfd = kpse.find_file("UGBK.sfd", "subfont~definition~files") ~
242       if ~ sfd ~ then ~
243         tex.write(sfd) ~
244       end
245     }
246   }
247 }
248 {
249   \ior_open:NnTF \g__CJKtu_sfd_ior { |kpsewhich ~ UGBK.sfd }
250   {
251     \ior_str_get:NN \g__CJKtu_sfd_ior \g__CJKtu_path_str
252     \ior_close:N \g__CJKtu_sfd_ior
253   }
254   { \msg_fatal:nmmm { xCJK2uni } { file-not-found } { GBK } { UGBK.sfd } }
255 }
256 \tl_if_blank:oTF { \g__CJKtu_path_str }
257 { \msg_fatal:nmmm { xCJK2uni } { file-not-found } { GBK } { UGBK.sfd } }
258 {
259   \exp_args:No \file_parse_full_name:nNNN { \g__CJKtu_path_str }
260   \l_tmpa_str \l_tmpb_str \l_tmpc_str

```

```

261     \str_gset_eq:NN \g__CJKtu_path_str \l_tmpa_str
262   }
263 \str_const:Nx \c__CJKtu_indent_str
264   { \c_catcode_other_space_tl \c_catcode_other_space_tl }
265 \iow_new:N \g__CJKtu_sfd_map_iow
266 \iow_new:N \g__CJKtu_sfd_cmap_iow
267 \seq_new:N \l__CJKtu_sfd_plane_seq
268 \seq_new:N \l__CJKtu_sfd_line_seq
269 \tl_new:N \l__CJKtu_sfd_plane_tl
270 \tl_new:N \l__CJKtu_cmap_block_tl
271 \int_new:N \l__CJKtu_sfd_index_int
\c__CJKtu_sfd_map_prop 272 \prop_const_from_keyval:Nn \c__CJKtu_sfd_map_prop
273   {
274     GB = { 10 } { UGB } ,
275     JIS = { 40 } { UJIS } ,
276     KS = { 60 } { UKS } ,
277     Bg5 = { 00 } { UBig5 } ,
278     Bg5+ = { 09 } { UBg5plus } ,
279     GBK = { 19 } { UGBK }
280   }
__CJKtu_write_file:nnn 281 \cs_new_protected_nopar:Npn \__CJKtu_write_file:nnn #1#2#3
282   {
283     \group_begin:
284     \seq_put_right:No \l_file_search_path_seq { \g__CJKtu_path_str }
285     \ior_open:NnTF \g__CJKtu_sfd_ior { #3.sfd }
286     {
287       \group_end:
288       \seq_clear:N \l__CJKtu_sfd_plane_seq
289       \__CJKtu_write_sfd_map_header:n {#1}
290       \ior_str_map_inline:Nn \g__CJKtu_sfd_ior
291       { \str_if_eq:nnT {##1} { 00 ~ 0x0000_0x00FF } { \ior_map_break: } }
292       \ior_str_map_inline:Nn \g__CJKtu_sfd_ior
293       {
294         \tl_if_blank:nTF {##1}
295         {
296           \ior_map_break:n
297           {
298             \__CJKtu_write_sfd_map_trailer:n {#1}
299             \__CJKtu_write_cmap_file:n {#2}
300           }
301         }
302         { \__CJKtu_read_sfd_line:nnn {##1} {#1} {#2} }
303       }
304       \iow_close:N \g__CJKtu_sfd_cmap_iow
305       \ior_close:N \g__CJKtu_sfd_ior
306       \seq_clear:N \l__CJKtu_sfd_plane_seq
307       \seq_clear:N \l__CJKtu_sfd_line_seq
308     }
309     {
310       \group_end:
311       \msg_critical:nxxx { xCJK2uni } { file-not-found } {#1} { #3.sfd }
312     }
313   }
__CJKtu_read_sfd_line:n 314 \cs_new_protected_nopar:Npx \__CJKtu_read_sfd_line:nnn #1#2#3
__CJKtu_read_sfd_line:nn 315   {
316     \exp_not:N \__CJKtu_read_sfd_line:nwnn #1
317     \c_backslash_str \exp_not:N \q_nil \exp_not:N \q_stop {#2} {#3}
318   }
319 \use:x
320   {
321     \cs_new_protected_nopar:Npn \exp_not:N \__CJKtu_read_sfd_line:nwnn
322     ##1 0 \token_to_str:N x ##2 \c_backslash_str ##3 \exp_not:N \q_stop
323     ##4##5
324     {
325       \exp_not:N \__CJKtu_read_sfd_line:nnnnn

```

```

326         {##1} { 0 \token_to_str:N x } {##2} {##4} {##5}
327     }
328 }

\__CJKtu_read_sfd_line:nnnnn 329 \cs_new_protected_nopar:Npn \__CJKtu_read_sfd_line:nnnnn #1#2#3#4#5
330 {
331     \tl_if_blank:nF {#1}
332     {
333         \seq_if_empty:NF \l__CJKtu_sfd_plane_seq
334         {
335             \iow_now:Nx \g__CJKtu_sfd_map_iow
336             {
337                 \c__CJKtu_indent_str \c__CJKtu_indent_str
338                 \token_to_str:N \fi: \iow_newline:
339                 \c__CJKtu_indent_str \c_right_brace_str
340             }
341             \__CJKtu_write_cmap_file:n {#5}
342         }
343         \__CJKtu_write_plane_map:nn {#1} {#4}
344         \tl_set:Nx \l__CJKtu_sfd_plane_tl { \tl_trim_spaces:n {#1} }
345     }
346     \seq_set_split:Nnn \l__CJKtu_sfd_line_seq {#2} {#3}
347     \seq_concat:NNN \l__CJKtu_sfd_plane_seq
348     \l__CJKtu_sfd_plane_seq \l__CJKtu_sfd_line_seq
349     \iow_now:Nx \g__CJKtu_sfd_map_iow
350     {
351         \c__CJKtu_indent_str \c__CJKtu_indent_str \c__CJKtu_indent_str
352         \seq_use:Nn \l__CJKtu_sfd_line_seq { ~ \or: } ~
353         \token_to_str:N \or:
354     }
355 }

\__CJKtu_write_plane_map:nn 356 \cs_new_protected_nopar:Npn \__CJKtu_write_plane_map:nn #1#2
357 {
358     \iow_now:Nx \g__CJKtu_sfd_map_iow
359     {
360         \cs_new_nopar:cpn { ~ __CJKtu_ #2 _sfd_map_ \int_eval:n {#1} :n ~ } ~
361         \c_hash_str 1 \iow_newline:
362         \c__CJKtu_indent_str \c_left_brace_str \iow_newline:
363         \c__CJKtu_indent_str \c__CJKtu_indent_str
364         \exp_not:N \if_case:w \etex_numexpr:D
365         \c_hash_str 1 ~ \token_to_str:N \scan_stop:
366     }
367 }

\__CJKtu_write_cmap_file:n 368 \cs_new_protected_nopar:Npn \__CJKtu_write_cmap_file:n #1
369 {
370     \tl_clear:N \l__CJKtu_cmap_block_tl
371     \int_zero:N \l__CJKtu_sfd_index_int
372     \iow_open:Nn \g__CJKtu_sfd_cmap_iow { c #1 \l__CJKtu_sfd_plane_tl .cmap }
373     \__CJKtu_write_cmap_header:nn { C #1 } { \l__CJKtu_sfd_plane_tl }
374     \iow_now:Nx \g__CJKtu_sfd_cmap_iow
375     {
376         \iow_newline:
377         1~begincodespacerange \iow_newline:
378         \c__CJKtu_indent_str < 00 > ~
379         <
380         \int_to_Hex:n
381         { \seq_count:N \l__CJKtu_sfd_plane_seq - \c_one }
382         > \iow_newline:
383         endcodespacerange
384     }
385     \seq_map_function:NN \l__CJKtu_sfd_plane_seq \__CJKtu_write_cmap_body:n
386     \tl_if_empty:NF \l__CJKtu_cmap_block_tl
387     {
388         \iow_now:Nx \g__CJKtu_sfd_cmap_iow
389         {
390             \iow_newline:

```

```

391         \int_mod:nn { \l__CJKtu_sfd_index_int } { \c_one_hundred } ~
392         beginbfchar \iow_newline:
393         \l__CJKtu_cmap_block_tl
394         endbfchar
395     }
396 }
397 \__CJKtu_write_cmap_trailer:
398 \seq_clear:N \l__CJKtu_sfd_plane_seq
399 }

```

```

\__CJKtu_write_cmap_body:n 400 \cs_new_protected_nopar:Npn \__CJKtu_write_cmap_body:n #1
401 {
402     \int_case:nnTF { \l__CJKtu_sfd_index_int }
403     { { 99 } { } { 199 } { } }
404     {
405         \iow_now:Nx \g__CJKtu_sfd_cmap_iow
406         {
407             \iow_newline:
408             100 ~ beginbfchar \iow_newline:
409             \l__CJKtu_cmap_block_tl
410             < \int_to_Hex:n { \l__CJKtu_sfd_index_int } > ~
411             <#1> \iow_newline:
412             endbfchar
413         }
414         \tl_clear:N \l__CJKtu_cmap_block_tl
415     }
416     {
417         \tl_put_right:Nx \l__CJKtu_cmap_block_tl
418         {
419             <
420                 \int_compare:nNf \l__CJKtu_sfd_index_int > \c_fifteen { 0 }
421                 \int_to_Hex:n { \l__CJKtu_sfd_index_int }
422             > ~ <#1> \iow_newline:
423         }
424     }
425     \int_incr:N \l__CJKtu_sfd_index_int
426 }

```

```

\__CJKtu_write_sfd_map_header:n 427 \group_begin:
428 \char_set_catcode_other:N \%
429 \cs_new_protected_nopar:Npn \__CJKtu_write_sfd_map_header:n #1
430 {
431     \iow_now:Nx \g__CJKtu_sfd_map_iow
432     {
433         %<#1> \iow_newline:
434         \cs_new_nopar:cpn { ~ CJKtu_#1_sfd_map:nn ~ } ~ \c_hash_str 1
435         \iow_newline: \c__CJKtu_indent_str \c_left_brace_str \iow_newline:
436         \c__CJKtu_indent_str \c__CJKtu_indent_str
437         \token_to_str:N \cs_if_exist_use:cF
438         \iow_newline: \c__CJKtu_indent_str \c__CJKtu_indent_str \c__CJKtu_indent_str
439         { ~ __CJKtu_#1_sfd_map_ ~ \exp_not:N \int_eval:n { \c_hash_str 1 } ~ :n ~}
440         \iow_newline: \c__CJKtu_indent_str \c__CJKtu_indent_str \c__CJKtu_indent_str
441         { ~ \exp_not:N \use_none:n } \iow_newline:
442         \c__CJKtu_indent_str \c_right_brace_str
443     }
444 }

```

```

\__CJKtu_write_sfd_map_trailer:n 445 \cs_new_protected_nopar:Npn \__CJKtu_write_sfd_map_trailer:n #1
446 {
447     \iow_now:Nx \g__CJKtu_sfd_map_iow
448     {
449         \c__CJKtu_indent_str \c__CJKtu_indent_str
450         \token_to_str:N \fi: \iow_newline:
451         \c__CJKtu_indent_str \c_right_brace_str \iow_newline:
452         %</#1>
453     }
454 }

```



```

__CJktu_write_cmap_header:nn 455 \cs_new_protected_nopar:Npn __CJktu_write_cmap_header:nn #1#2
456 {
457   \iow_now:Nx \g__CJktu_sfd_cmap_iow
458   {
459     %!PS-Adobe-3.0~Resource-CMap \iow_newline:
460     %%DocumentNeededResources:~ProcSet~(CIDInit) \iow_newline:
461     %%IncludeResource:~ProcSet~(CIDInit) \iow_newline:
462     %%BeginResource:~CMap~(TeX-#1-#2-0) \iow_newline:
463     %%Title:~(TeX-#1-#2-0~TeX~#1-#2~0) \iow_newline:
464     %%Version:~1.000 \iow_newline:
465     %%EndComments \iow_newline: \iow_newline:
466     /CIDInit~/ProcSet~findresource~begin \iow_newline: \iow_newline:
467     12~dict~begin \iow_newline: \iow_newline:
468     begincmap \iow_newline: \iow_newline:
469     /CIDSystemInfo~3~dict~dup~begin \iow_newline:
470     \c__CJktu_indent_str /Registry~(TeX)~def \iow_newline:
471     \c__CJktu_indent_str /Ordering~(#1-#2)~def \iow_newline:
472     \c__CJktu_indent_str /Supplement~0~def \iow_newline:
473     end~def \iow_newline: \iow_newline:
474     /CMapName~/TeX-#1-#2-0~def \iow_newline:
475     /CMapVersion~1.000~def \iow_newline:
476     /CMapType~2~def \iow_newline: \iow_newline:
477     /WMode~0~def
478   }
479 }

__CJktu_write_cmap_trailer: 480 \cs_new_protected_nopar:Npx __CJktu_write_cmap_trailer:
481 {
482   \iow_now:Nn \g__CJktu_sfd_cmap_iow
483   {
484     \iow_newline:
485     endcmap \iow_newline:
486     CMapName~currentdict~/CMap~defineresource~pop \iow_newline:
487     end \iow_newline:
488     end \iow_newline: \iow_newline:
489     %%EndResource \iow_newline:
490     %%EOF
491   }
492 }
493 \group_end:

494 \iow_open:Nn \g__CJktu_sfd_map_iow { xCJK2uni-sfd.def }
495 \prop_map_inline:Nn \c__CJktu_sfd_map_prop
496 { \__CJktu_write_file:nnn {#1} #2 }
497 \iow_close:N \g__CJktu_sfd_map_iow
498 \cs_if_exist_use:NF \CJKtuend { \tex_end:D }
499 </make>

```

# 代码索引

意大利体的数字表示描述对应索引项的页码;带下划线的数字表示定义对应索引项的代码行号;罗马字体的数字表示使用对应索引项的代码行号。

## Symbols

\% ..... 428  
 \\ ..... 5, 6, 7, 80, 229

### A

\AtBeginDocument ..... 54, 201

### C

#### catcode commands:

\c\_catcode\_other\_space\_tl ..... 264  
 \CCTchar ..... 126

#### char commands:

\char\_set\_catcode\_active:n ..... 105  
 \char\_set\_catcode\_other:N ..... 428  
 \CJKaddEncHook ..... 51, 56  
 \CJKchar ..... 125  
 \CJKchartouni ..... 1, 211  
 \JKsfdtouni ..... 1, 211

#### CJKtu commands:

\CJKtu\_bookmarks\_hook: ..... 121, 199  
 \CJKtu\_byte:NN ..... 110, 124, 131  
 \CJKtu\_byte:nn ..... 126, 132, 134, 135, 136  
 \CJKtu\_byte\_map:nn ..... 118, 120, 141, 145, 158  
 \CJKtu\_sfd\_map:nn ..... 70, 94, 175

#### CJKtu internal commands:

\_\_CJKtu\_byte:Nwnw ..... 119, 123  
 \_\_CJKtu\_byte:wnn ..... 125, 135  
 \_\_CJKtu\_byte:wNwnw ..... 113, 117  
 \_\_CJKtu\_CCT\_byte:NNN ..... 127, 131  
 \_\_CJKtu\_char\_to\_unicode:n ..... 104, 212  
 \l\_CJKtu\_cmap\_block\_tl 270, 370, 386, 393, 409, 414, 417  
 \_\_CJKtu\_enc\_map:nn ..... 52, 57, 61, 75, 219  
 \_\_CJKtu\_encoding\_error: ..... 114, 146, 149, 151  
 \c\_CJKtu\_encoding\_seq ..... 20, 32, 81  
 \l\_CJKtu\_gap\_begin\_int ..... 25, 65, 140, 166, 167  
 \l\_CJKtu\_gap\_end\_int ..... 26, 66, 144, 167  
 \_\_CJKtu\_get\_unicode:n ..... 128, 173, 180  
 \_\_CJKtu\_get\_unicode:w ..... 94, 96, 181  
 \_\_CJKtu\_get\_unicode\_aux:w ..... 100, 102  
 \l\_CJKtu\_i\_max\_int ..... 22, 28, 138  
 \l\_CJKtu\_i\_min\_int ..... 21, 63, 138, 164  
 \l\_CJKtu\_ii\_max\_int ..... 24, 29, 39, 43, 49, 144  
 \l\_CJKtu\_ii\_min\_int ..... 23, 64, 140, 165  
 \c\_CJKtu\_indent\_str ..... 263, 337, 339, 351, 362,  
 363, 378, 435, 436, 438, 440, 442, 449, 451, 470, 471, 472  
 \_\_CJKtu\_index\_map:n ..... 160, 171  
 \_\_CJKtu\_initial:N ..... 214, 224  
 \g\_CJKtu\_path\_str ..... 233, 236, 251, 256, 259, 261, 284  
 \l\_CJKtu\_plane\_int ..... 27, 67, 164  
 \_\_CJKtu\_read\_sfd\_line:n ..... 314  
 \_\_CJKtu\_read\_sfd\_line:nn ..... 314

\_\_CJKtu\_read\_sfd\_line:nnn ..... 302, 314  
 \_\_CJKtu\_read\_sfd\_line:nnnnn ..... 325, 329  
 \_\_CJKtu\_read\_sfd\_line:nwnn ..... 316, 321  
 \_\_CJKtu\_set\_enc:nnnnnn ..... 30, 85, 86, 87, 88, 89, 90  
 \g\_CJKtu\_sfd\_cmap\_iow .....  
 ..... 266, 304, 372, 374, 388, 405, 457, 482  
 \l\_CJKtu\_sfd\_index\_int .....  
 ..... 271, 371, 391, 402, 410, 420, 421, 425  
 \g\_CJKtu\_sfd\_ior ..... 232, 249, 251, 252, 285, 290, 292, 305  
 \l\_CJKtu\_sfd\_line\_seq ..... 268, 307, 346, 348, 352  
 \g\_CJKtu\_sfd\_map\_iow .....  
 ..... 265, 335, 349, 358, 431, 447, 494, 497  
 \c\_CJKtu\_sfd\_map\_prop ..... 272, 495  
 \l\_CJKtu\_sfd\_plane\_seq .....  
 ..... 267, 288, 306, 333, 347, 348, 381, 385, 398  
 \l\_CJKtu\_sfd\_plane\_tl ..... 269, 344, 372, 373  
 \c\_CJKtu\_sfd\_prop ..... 19, 33, 74, 218  
 \l\_CJKtu\_sfd\_tl ..... 18, 74, 75, 218, 219  
 \_\_CJKtu\_sfd\_to\_unicode:nn ..... 91, 213  
 \_\_CJKtu\_use\_enc\_map:n ..... 72, 211, 220, 222  
 \_\_CJKtu\_UTF\_xvi\_aux:n ..... 183, 184  
 \_\_CJKtu\_UTF\_xvi\_be:n ..... 128, 182  
 \_\_CJKtu\_write\_cmap\_body:n ..... 385, 400  
 \_\_CJKtu\_write\_cmap\_file:n ..... 299, 341, 368  
 \_\_CJKtu\_write\_cmap\_header:nn ..... 373, 455  
 \_\_CJKtu\_write\_cmap\_trailer: ..... 397, 480  
 \_\_CJKtu\_write\_file:nnn ..... 281, 496  
 \_\_CJKtu\_write\_plane\_map:nn ..... 343, 356  
 \_\_CJKtu\_write\_sfd\_map\_header:n ..... 289, 427  
 \_\_CJKtu\_write\_sfd\_map\_trailer:n ..... 298, 445  
 \CJKtuend ..... 498

#### cs commands:

\cs\_gset\_eq:NN ..... 205  
 \cs\_if\_exist:NTF ..... 68  
 \cs\_if\_exist\_use:NTF ..... 51, 56, 191, 208, 437, 498  
 \cs\_new\_eq:NN ..... 199, 211, 212, 213  
 \cs\_new\_nopar:Npn ..... 91, 96, 102, 106, 117, 119,  
 131, 133, 135, 136, 151, 158, 171, 180, 182, 184, 360, 434  
 \cs\_new\_protected\_nopar:Npn ..... 30,  
 61, 72, 121, 214, 281, 321, 329, 356, 368, 400, 429, 445, 455  
 \cs\_new\_protected\_nopar:Npx ..... 314, 480  
 \cs\_set\_eq:NN ..... 70, 123, 124, 125, 126, 127, 128

### E

#### etex commands:

\etex\_numexpr:D ..... 364

#### exp commands:

\exp:w ..... 93, 108  
 \exp\_args:Nc ..... 224  
 \exp\_args:Nf ..... 160, 183  
 \exp\_args:No ..... 259

\exp_end: .....	94, 96, 99, 103, 154, 181	\iow_now:Nn	335, 349, 358, 374, 388, 405, 431, 447, 457, 482
\exp_last_unbraced:Nn .....	93, 181	\iow_open:Nn .....	372, 494
\exp_not:N .....	316, 317, 321, 322, 325, 364, 439, 441		
<b>F</b>			
fi commands:		left commands:	
\fi: .....	338, 450	\c_left_brace_str .....	362, 435
file commands:		lua commands:	
\file_input:n .....	69	\lua_now_x:n .....	238
\file_parse_full_name:nNNN .....	259		
\l_file_search_path_seq .....	284	<b>M</b>	
<b>G</b>			
group commands:		msg commands:	
\group_begin: .....	69, 104, 283, 427	\msg_critical:nn .....	16
\group_end: .....	69, 130, 287, 310, 493	\msg_critical:nnnn .....	311
<b>H</b>			
hundred commands:		\msg_error:nn .....	11
\c_one_hundred .....	391	\msg_error:nnn .....	76
\hypersetup .....	191, 208	\msg_expandable_error:nn .....	153
<b>I</b>			
if commands:		\msg_fatal:nnnn .....	254, 257
\if_case:w .....	364	\msg_new:nnn .....	3, 14, 78, 156, 193, 227
int commands:		\msg_warning:nn .....	197, 206
\c_fifteen .....	420	<b>O</b>	
\int_case:nnTF .....	402	or commands:	
\int_compare:nNnTF .....	109, 166, 420	\or: .....	102, 103, 352, 353
\int_compare:nTF .....	138, 140, 143	<b>P</b>	
\int_const:Nn .....	34, 35, 38, 40, 42, 46, 47, 48	\PassOptionsToPackage .....	192
\int_div_truncate:nn .....	176, 187	\pdfstringdefPreHook .....	195, 200
\int_eval:n .....	162, 183, 360, 439	prg commands:	
\int_incr:N .....	425	\prg_generate_conditional_variant:Nnn .....	84
\int_mod:nn .....	177, 189, 391	prop commands:	
\int_new:N .....	21, 22, 23, 24, 25, 26, 27, 271	\prop_const_from_keyval:Nn .....	272
\int_set:Nn .....	28, 29	\prop_get:NnN .....	84
\int_set_eq:NN .....	63, 64, 65, 66, 67	\prop_get:NnNTF .....	74, 218
\int_to_Hex:n .....	380, 410, 421	\prop_gput:Nnn .....	33
\int_zero:N .....	371	\prop_map_inline:Nn .....	495
\c_one .....	39, 43, 46, 176, 381	\prop_new:N .....	19
\c_two .....	49, 109	<b>Q</b>	
ior commands:		quark commands:	
\ior_close:N .....	252, 305	\q_nil .....	317
\ior_map_break: .....	291	\q_stop .....	100, 102, 317, 322
\ior_map_break:n .....	296	<b>R</b>	
\ior_new:N .....	232	right commands:	
\ior_open:NnTF .....	249, 285	\c_right_brace_str .....	339, 442, 451
\ior_str_get:NN .....	251	<b>S</b>	
\ior_str_map_inline:Nn .....	290, 292	scan commands:	
iow commands:		\scan_stop: .....	205, 365
\iow_close:N .....	304, 497	seq commands:	
\iow_new:N .....	265, 266	\seq_clear:N .....	288, 306, 307, 398
\iow_newline: 338, 361, 362, 376, 377, 382, 390, 392, 407,		\seq_concat:NNN .....	347
408, 411, 422, 433, 435, 438, 440, 441, 450, 451, 459,		\seq_count:N .....	381
460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470,		\seq_gput_right:Nn .....	32
471, 472, 473, 474, 475, 476, 484, 485, 486, 487, 488, 489		\seq_if_empty:NTF .....	333
		\seq_map_function:NN .....	385
		\seq_new:N .....	20, 267, 268
		\seq_put_right:Nn .....	284

`\seq_set_split:Nnn` ..... 346  
`\seq_use:Nn` ..... 352  
`\seq_use:Nnnn` ..... 81

str commands:

`\c_backslash_str` ..... 317, 322  
`\c_hash_str` ..... 361, 365, 434, 439  
`\str_const:Nn` ..... 263  
`\str_gset:Nn` ..... 236  
`\str_gset_eq:NN` ..... 261  
`\str_if_eq:nnTF` ..... 291  
`\str_new:N` ..... 233  
`\l_tmpa_str` ..... 260, 261  
`\l_tmpb_str` ..... 260

sys commands:

`\sys_if_engine luatex:TF` ..... 234  
`\sys_if_engine pdftex:TF` ..... 12

**T**

$\TeX$  and  $\LaTeX_{\epsilon}$  commands:

`\@CCTSetChar` ..... 127  
`\@ifpackagelater` ..... 10  
`\@ifpackageloaded` ..... 196, 203  
`\CJK@XX` ..... 124  
`\CJKtu@bookmarks@hook` ..... 199, 200, 205  
`\HyPsd@DecimalToOctalFirst` ..... 186  
`\HyPsd@DecimalToOctalSecond` ..... 188

tex commands:

`\tex_end:D` ..... 498

tl commands:

`\tl_clear:N` ..... 370, 414  
`\tl_count:n` ..... 109  
`\tl_gput_right:Nn` ..... 200  
`\tl_if_blank:nTF` ..... 256, 294, 331  
`\tl_if_empty:NTF` ..... 386  
`\tl_if_empty:nTF` ..... 36, 98  
`\tl_if_exist:NTF` ..... 195, 216  
`\tl_if_head_eq_meaning:nNTF` ..... 112  
`\tl_new:N` ..... 18, 195, 269, 270  
`\tl_put_right:Nn` ..... 417  
`\tl_set:Nn` ..... 344  
`\tl_trim_spaces:n` ..... 344

token commands:

`\token_to_str:N` ..... 322, 326, 338, 353, 365, 437, 450

two commands:

`\c_two_hundred_fifty_six` ..... 176, 177, 187, 189

**U**

use commands:

`\use:N` ..... 41  
`\use:n` ..... 319  
`\use_none:n` ..... 441  
`\useCJKencmap` ..... 1, 211