

# The soul package

Heiko Oberdiek\*

2023-06-14 v3.1

## Abstract

This version is a merge of the original package `soul` (version 2.5 which contains some small corrections compared to the version 2.4 from 2003) and the `soulutf8` which added some support for UTF-8. Namely the input encodings `utf8.def` from package `inputenc` and package `ucs`'s `utf8x.def` are supported. The original `soul` package has been renamed to `soul-ori` and is still loaded. Its documentation is still relevant.

## Contents

<b>1</b>	<b>Documentation</b>	<b>2</b>
1.1	Patch	2
1.2	Future	2
<b>2</b>	<b>Implementation</b>	<b>2</b>
2.1	Reload check and package identification	2
2.2	Catcodes	4
2.3	Loading packages	5
2.3.1	plain $\TeX$	5
2.3.2	$\LaTeX$	6
2.3.3	$\varepsilon\text{-}\TeX$	6
2.4	Macro for redefinitions	6
2.5	Redefinition of <code>\SOUL@eval</code>	6
2.6	UTF-8 analysis	9
2.6.1	Help strings	9
2.6.2	Support for <code>utf8.def</code>	10
2.6.3	Support for <code>utf8x.def</code>	10
2.7	Actions for UTF-8 sequences	11
2.7.1	Redefinition of <code>\SOUL@splittoken</code>	12
2.8	Patches	12
<b>3</b>	<b>Installation</b>	<b>15</b>
3.1	Download	15
3.2	Package installation	15
3.3	Refresh file name databases	15
<b>4</b>	<b>References</b>	<b>15</b>
<b>5</b>	<b>History</b>	<b>16</b>
	[2007/09/09 v1.0]	16
	[2016/05/16 v1.1]	16
	[2019/12/15 v1.2]	16
	[2023-02-18 v3.0]	16
	[2023-06-14 v3.1]	16

---

\*Please report any issues at <https://github.com/ho-tex/soul/issues>

## 1 Documentation

This package `soul` does not have own options and does not define new user commands. Any option is passed to package `soul-ori` [1] that is loaded first. Then some internal macros of `soul-ori` are redefined to add support for UTF-8. The following input encodings are supported:

```
utf8      LATEX base    TDS:tex/latex/base/utf8.def [3]
utf8x     Package ucs   TDS:tex/latex/ucs/utf8x.def [2]
```

UTF-8 byte sequences are added as token group to a word, even if these UTF-8 characters are some kind of hyphen or space. As exception the following three Unicode characters are handled specially:

Slot	Name	Action
U+00A0	NO-BREAK SPACE	like ~
U+2013	EN DASH	--
U+2014	EM DASH	---

### 1.1 Patch

Also package `soul` tries to patch package `soul-ori` to improve its behaviour:

- A problem with additional levels of curly braces is fixed. As advantage more implicate kernings are detected. However, the result may be incompatible with the original behaviour of package `soul-ori` because of these respected implicate kernings.
- $\varepsilon$ -T<sub>E</sub>X, especially `\unexpanded` is supported. This allows a better protection of token groups (`\mbox{...}`, `math, ...`).

### 1.2 Future

Currently package `soul-ori` does not seem to be maintained. Nevertheless if there will be a new version that adds support for UTF-8, then this package may become obsolete.

## 2 Implementation

```
1 (*soulutf8)
2 \NeedsTeXFormat{LaTeX2e}
3 \ProvidesPackage{soulutf8}%
4 [2023-06-14 v3.1 Permit use of UTF-8 characters in soul (HO)]%
5 \PackageWarning{soulutf8}
6   {This package is obsolete,\MessageBreak
7    use the soul package directly. \MessageBreak
8    }{}%
9 \DeclareOption*{\PassOptionsToPackage{\CurrentOption}{soul}}%
10 \ProcessOptions\relax
11 \RequirePackage{soul}
12 </soulutf8>
13 (*package)
```

### 2.1 Reload check and package identification

Reload check, especially if the package is not used with L<sup>A</sup>T<sub>E</sub>X.

```
14 \begingroup\catcode61\catcode48\catcode32=10\relax%
```

```

15 \catcode13=5 % ^~M
16 \endlinechar=13 %
17 \catcode35=6 % #
18 \catcode39=12 % '
19 \catcode44=12 % ,
20 \catcode45=12 % -
21 \catcode46=12 % .
22 \catcode58=12 % :
23 \catcode64=11 % @
24 \catcode123=1 % {
25 \catcode125=2 % }
26 \expandafter\let\expandafter\x\csname ver@soul.sty\endcsname
27 \ifx\x\relax % plain-TeX, first loading
28 \else
29   \def\empty{}%
30   \ifx\x\empty % LaTeX, first loading,
31     % variable is initialized, but \ProvidesPackage not yet seen
32   \else
33     \expandafter\ifx\csname PackageInfo\endcsname\relax
34       \def\x#1#2{%
35         \immediate\write-1{Package #1 Info: #2.}%
36       }%
37     \else
38       \def\x#1#2{\PackageInfo{#1}{#2, stopped}}%
39     \fi
40     \x{soul}{The package is already loaded}%
41     \aftergroup\endinput
42   \fi
43 \fi
44 \endgroup%

```

Package identification:

```

45 \begingroup\catcode61\catcode48\catcode32=10\relax%
46 \catcode13=5 % ^~M
47 \endlinechar=13 %
48 \catcode35=6 % #
49 \catcode39=12 % '
50 \catcode40=12 % (
51 \catcode41=12 % )
52 \catcode44=12 % ,
53 \catcode45=12 % -
54 \catcode46=12 % .
55 \catcode47=12 % /
56 \catcode58=12 % :
57 \catcode64=11 % @
58 \catcode91=12 % [
59 \catcode93=12 % ]
60 \catcode123=1 % {
61 \catcode125=2 % }
62 \expandafter\ifx\csname ProvidesPackage\endcsname\relax
63   \def\x#1#2#3[#4]{\endgroup
64     \immediate\write-1{Package: #3 #4}%
65     \xdef#1{#4}%
66   }%
67 \else
68   \def\x#1#2[#3]{\endgroup
69     #2[#{#3}]%
70     \ifx#1\undefined
71       \xdef#1{#3}%
72     \fi
73     \ifx#1\relax
74       \xdef#1{#3}%
75     \fi

```

```

76   }%
77   \fi
78 \expandafter\x\csname ver@soul.sty\endcsname
79 \ProvidesPackage{soul}%
80 [2023-06-14 v3.1 Permit use of UTF-8 characters in soul (HO)]%

```

## 2.2 Catcodes

```

81 \begingroup\catcode61\catcode48\catcode32=10\relax%
82 \catcode13=5 % ^^M
83 \endlinechar=13 %
84 \catcode123=1 % {
85 \catcode125=2 % }
86 \catcode64=11 % @
87 \def\x{\endgroup
88   \expandafter\edef\csname SOuL@AtEnd\endcsname{%
89     \endlinechar=\the\endlinechar\relax
90     \catcode13=\the\catcode13\relax
91     \catcode32=\the\catcode32\relax
92     \catcode35=\the\catcode35\relax
93     \catcode61=\the\catcode61\relax
94     \catcode64=\the\catcode64\relax
95     \catcode123=\the\catcode123\relax
96     \catcode125=\the\catcode125\relax
97   }%
98 }%
99 \x\catcode61\catcode48\catcode32=10\relax%
100 \catcode13=5 % ^^M
101 \endlinechar=13 %
102 \catcode35=6 % #
103 \catcode64=11 % @
104 \catcode123=1 % {
105 \catcode125=2 % }
106 \def\TMP@EnsureCode#1#2{%
107   \edef\SOU@AtEnd{%
108     \SOU@AtEnd
109     \catcode#1=\the\catcode#1\relax
110   }%
111   \catcode#1=#2\relax
112 }
113 \TMP@EnsureCode{10}{12}% ^^J
114 \TMP@EnsureCode{33}{12}% !
115 \TMP@EnsureCode{34}{12}% "
116 \TMP@EnsureCode{36}{3}% $
117 \TMP@EnsureCode{39}{12}% '
118 \TMP@EnsureCode{40}{12}% (
119 \TMP@EnsureCode{41}{12}% )
120 \TMP@EnsureCode{42}{12}% *
121 \TMP@EnsureCode{43}{12}% +
122 \TMP@EnsureCode{44}{12}% ,
123 \TMP@EnsureCode{45}{12}% -
124 \TMP@EnsureCode{46}{12}% .
125 \TMP@EnsureCode{47}{12}% /
126 \TMP@EnsureCode{58}{12}% :
127 \TMP@EnsureCode{60}{12}% <
128 \TMP@EnsureCode{62}{12}% >
129 \TMP@EnsureCode{91}{12}% [
130 \TMP@EnsureCode{93}{12}% ]
131 \TMP@EnsureCode{94}{7}% ^
132 \TMP@EnsureCode{96}{12}% ‘
133 \TMP@EnsureCode{126}\active % ~
134 \TMP@EnsureCode{128}{12}% ^^80

```

```

135 \TMP@EnsureCode{147}{12}% ^^93
136 \TMP@EnsureCode{148}{12}% ^^94
137 \TMP@EnsureCode{160}{12}% ^^a0
138 \TMP@EnsureCode{194}{12}% ^^c2
139 \TMP@EnsureCode{226}{12}% ^^e2
140 \edef\Soul@AtEnd{\Soul@AtEnd\noexpand\endinput}

```

## 2.3 Loading packages

Package soul-ori uses `\documentclass` to detect L<sup>A</sup>T<sub>E</sub>X.

```
141 \ifx\documentclass\undefined
```

### 2.3.1 plain T<sub>E</sub>X

First we check, whether package soul-ori is already loaded.

```
142 \expandafter\ifx\csname SOUL@\endcsname\relax
```

In case of plain T<sub>E</sub>X package soul-ori defines some macros in a simple manner that will break the definitions of `miniltx.tex`, for example. Therefore these macros are first saved and restored afterwards.

```

143 \let\Soul@orgDeclareRobustCommand\DeclareRobustCommand
144 \let\Soul@orgnewcommand \newcommand
145 \let\Soul@orgDeclareOption \DeclareOption
146 \let\Soul@orgPackageError \PackageError
147 \def\Soul@restorelatexcmds{%
148 \let\DeclareRobustCommand\Soul@orgDeclareRobustCommand
149 \let\newcommand \Soul@orgnewcommand
150 \let\DeclareOption \Soul@orgDeclareOption
151 \let\PackageError \Soul@orgPackageError
152 }%
153 \input soul-ori.sty\relax
154 \Soul@restorelatexcmds
155 \fi

```

`\SOUL@error` Package soul-ori's use of `\PackageError` is replaced by `\@PackageError` of package `infwarerr`.

```

156 \input infwarerr.sty\relax
157 \let\Soul@orgSOUL@error\SOUL@error
158 \def\SOUL@error{%
159 \begingroup
160 \let\PackageError\@PackageError
161 \Soul@orgSOUL@error
162 \endgroup
163 }%

```

```
164 \input etexcmds.sty\relax
```

`\@onelevel@sanitize` Define L<sup>A</sup>T<sub>E</sub>X's `\@onelevel@sanitize` if not already available.

```

165 \expandafter\ifx\csname @onelevel@sanitize\endcsname\relax
166 \def\@onelevel@sanitize#1{%
167 \edef#1{%
168 \expandafter\strip@prefix\meaning#1%
169 }%
170 }%

```

`\strip@prefix`

```

171 \def\strip@prefix#1>{}%
172 \fi
173 \else

```

### 2.3.2 L<sup>A</sup>T<sub>E</sub>X

```
174 \DeclareOption*{\PassOptionsToPackage{\CurrentOption}{soul-ori}}%
175 \ProcessOptions\relax
176 \RequirePackage{soul-ori}[2023-02-18]%
177 \RequirePackage{infwarerr}[2019/12/03]%
178 \RequirePackage{etexcmds}[2019/12/15]%
179 \fi
```

### 2.3.3 ε-<sub>T</sub>E<sub>X</sub>

In plain T<sub>E</sub>X command \+ is an *outer* macro. Therefore numbers are used to avoid problems.

```
180 \ifetex@unexpanded
181 \catcode33=14 % '!': comment
182 \catcode43=9 % '+': ignore
183 \else
184 \catcode33=9 % '!': ignore
185 \catcode43=14 % '+': comment
186 \fi
```

## 2.4 Macro for redefinitions

\SOuL@redefine

```
187 \def\SOuL@redefine#1{%
188 \begingroup
189 \def\SOuL@cmd{#1}%
190 \afterassignment\SOuL@cmdcheck
191 \def\SOuL@temp
192 }
```

\SOuL@cmdcheck

```
193 \def\SOuL@cmdcheck{%
194 \expandafter\ifx\SOuL@cmd\SOuL@temp
195 \else
196 \edef\SOuL@temp*{\expandafter\string\SOuL@cmd}%
197 \@PackageWarningNoLine{soul}{%
198 Command \SOuL@temp* has changed.\MessageBreak
199 Supported versions of package 'soul-ori': 2021/12/14.\MessageBreak
200 Depending on the unknown changes the redefinition\MessageBreak
201 of \SOuL@temp* may not behave correctly%
202 }%
203 \fi
204 \expandafter\endgroup
205 \expandafter\def\SOuL@cmd
206 }
```

## 2.5 Redefinition of \SOUL@eval

\SOUL@eval Macro \SOUL@eval is redefined to add detection of the first byte of a UTF-8 sequence. Because \SOUL@eval is overwritten, a warning is issued, if the contents of \SOUL@eval is not as expected.

```
207 \SOuL@redefine\SOUL@eval{%
First the expected definition.
208 \def\SOUL@n*##1{\SOUL@scan}%
209 \if\noexpand\SOUL@@\SOUL@spc
210 \else
211 \SOUL@ignorespacesfalse
212 \fi
213 \ifnum\SOUL@minus=\thr@@
214 \SOUL@flushminus
215 \else\ifnum\SOUL@comma=\tw@
```

```

216 \SOUL@flushcomma
217 \else\ifnum\SOUL@apo=\tw@
218 \SOUL@flushapo
219 \else\ifnum\SOUL@grave=\tw@
220 \SOUL@flushgrave
221 \fi\fi\fi\fi
222 \ifx\SOUL@@-\else\SOUL@flushminus\fi
223 \ifx\SOUL@@,\else\SOUL@flushcomma\fi
224 \ifx\SOUL@@'\else\SOUL@flushapo\fi
225 \ifx\SOUL@@'\else\SOUL@flushgrave\fi
226 \ifx\SOUL@@-%
227 \advance\SOUL@minus\@ne
228 \else\ifx\SOUL@@,%
229 \advance\SOUL@comma\@ne
230 \else\ifx\SOUL@@'%
231 \advance\SOUL@apo\@ne
232 \else\ifx\SOUL@@'%
233 \advance\SOUL@grave\@ne
234 \else
235 \SOUL@flushminus
236 \SOUL@flushcomma
237 \SOUL@flushapo
238 \SOUL@flushgrave
239 \ifx\SOUL@@\SOUL@stop
240 \def\SOUL@n*{%
241 \SOUL@doword
242 \SOUL@eventuallyexhyphen\null
243 }%
244 \else\ifx\SOUL@@\par
245 \def\SOUL@n*\par{\par\leavevmode\SOUL@scan}%
246 \else\if\noexpand\SOUL@@\SOUL@spc
247 \SOUL@doword
248 \SOUL@eventuallyexhyphen\null
249 \ifSOUL@ignorespaces
250 \else
251 \SOUL@everySPACE{}%
252 \fi
253 \def\SOUL@n* {\SOUL@scan}%
254 \else\ifx\SOUL@@\%
255 \SOUL@doword
256 \SOUL@eventuallyexhyphen\null
257 \SOUL@everySPACE{\unskip\nobreak\hfil\break}%
258 \SOUL@ignorespacestrue
259 \else\ifx\SOUL@@~%
260 \SOUL@doword
261 \SOUL@eventuallyexhyphen\null
262 \SOUL@everySPACE{\nobreak}%
263 \else\ifx\SOUL@@\slash
264 \SOUL@doword
265 \SOUL@eventuallyexhyphen{/}%
266 \SOUL@exhyphen{/}%
267 \else\ifx\SOUL@@\mbox
268 \def\SOUL@n*{\SOUL@addprotect}%
269 \else\ifx\SOUL@@\hbox
270 \def\SOUL@n*{\SOUL@addprotect}%
271 \else\ifx\SOUL@@\soulomit
272 \def\SOUL@n*\soulomit##1{%
273 \SOUL@doword
274 {\spaceskip\SOUL@spaceskip##1}%
275 \SOUL@scan
276 }%
277 \else\ifx\SOUL@@\break

```







```

395 \Soul@defsanitizedstring{UTFviii}{UTFviii@}
396 \Soul@defsanitizedstring{octets}{@octets}
397 \Soul@defsanitizedstring{two}{two}
398 \Soul@defsanitizedstring{three}{three}
399 \Soul@defsanitizedstring{four}{four}
400 \Soul@defsanitizedstring{macrocolon}{macro:}
401 \Soul@defsanitizedstring{csnameu}{csname u8-}
402 \Soul@defsanitizedstring{undeferr}{utf@viii@undeferr}
403 \def\Soul@stringendash{^^e2^^80^^93}
404 \def\Soul@stringemdash{^^e2^^80^^94}
405 \def\Soul@stringnobreakspace{^^c2^^a0}
406 \edef\Soul@charhash{\string #}
407 \edef\Soul@chartwo{\string 2}
408 \edef\Soul@charthree{\string 3}
409 \def\Soul@empty{}

```

## 2.6.2 Support for utf8.def

\Soul@analyzeutfviii

```

410 \begingroup
411   \edef\x{\endgroup
412     \def\noexpand\Soul@analyzeutfviii{%
413       \noexpand\expandafter\noexpand\Soul@checkutfviii
414       \noexpand\meaning\noexpand\Soul@@
415       \Soul@stringUTFviii\Soul@stringoctets
416       \noexpand@nil
417     }%
418     \def\noexpand\Soul@checkutfviii
419       ##1\Soul@stringUTFviii##2\Soul@stringoctets##3\noexpand@nil
420   }%
421 \x{%
422   \def\Soul@temp{#2}%
423   \chardef\Soul@octets=%
424     \ifx\Soul@temp\Soul@stringtwo
425       \tw@
426     \else\ifx\Soul@temp\Soul@stringthree
427       \thr@@
428     \else\ifx\Soul@temp\Soul@stringfour
429       4 %
430     \else
431       \z@
432     \fi\fi\fi
433 }

```

## 2.6.3 Support for utf8x.def

\Soul@analyzeutfviiix

```

434 \begingroup
435   \edef\x{\endgroup
436     \def\noexpand\Soul@analyzeutfviiix{%
437       \noexpand\expandafter\noexpand\Soul@checkutfviiix
438       \noexpand\meaning\noexpand\Soul@@
439       \Soul@stringmacrocolon\Soul@charhash1{}{}{}{}%
440       \Soul@stringcsnameu\Soul@stringundeferr
441       \noexpand@nil
442     }%

```

\Soul@checkutfviiix

```

443   \def\noexpand\Soul@checkutfviiix
444     ##1\Soul@stringmacrocolon\Soul@charhash1##2##3##4##5##6%
445     \Soul@stringcsnameu##7\Soul@stringundeferr##8\noexpand@nil
446   }%
447 \x{%

```

```

448 \def\Soul@temp{#7}%
449 \ifx\Soul@temp\Soul@empty
450   \chardef\Soul@octets=\z@
451 \else
452   \def\Soul@temp{#5}%
453   \ifx\Soul@temp\Soul@charthree
454     \chardef\Soul@octets=4 %
455   \else
456     \def\Soul@temp{#3}%
457     \ifx\Soul@temp\Soul@chartwo
458       \chardef\Soul@octets=\thr@@
459     \else
460       \chardef\Soul@octets=\tw@
461     \fi
462   \fi
463 \fi
464 }

```

## 2.7 Actions for UTF-8 sequences

\Soul@addtwooctets

```

465 \def\Soul@addtwooctets#1#2{%
466   \def\Soul@temp{#1#2}%
467   \@onelevel@sanitize\Soul@temp
468   \ifx\Soul@temp\Soul@stringnobreakspace
469     \SOUL@doword
470     \SOUL@eventuallyexhyphen\null
471     \SOUL@everySPACE{\nobreak}%
472     \let\Soul@next\Soul@scan
473   \else
474     \def\Soul@next{%
475 !       \SOUL@addtoken{\noexpand#1\noexpand#2}}%
476 +       \SOUL@addtoken{\etex@unexpanded{#1#2}}}%
477   }%
478   \fi
479   \Soul@next
480 }

```

\Soul@addthreectets

```

481 \def\Soul@addthreectets#1#2#3{%
482   \def\Soul@temp{#1#2#3}%
483   \@onelevel@sanitize\Soul@temp
484   \ifx\Soul@temp\Soul@stringendash
485     \SOUL@doword
486     \SOUL@eventuallyexhyphen{-}%
487     \SOUL@exhyphen{--}%
488     \let\Soul@next\Soul@scan
489   \else
490     \ifx\Soul@temp\Soul@stringemdash
491       \SOUL@doword
492       \SOUL@eventuallyexhyphen{-}%
493       \SOUL@exhyphen{---}%
494       \let\Soul@next\Soul@scan
495     \else
496       \def\Soul@next{%
497 !       \SOUL@addtoken{\noexpand#1\noexpand#2\noexpand#3}}%
498 +       \SOUL@addtoken{\etex@unexpanded{#1#2#3}}}%
499     }%
500     \fi
501   \fi
502   \Soul@next
503 }

```

`\SOUL@addfouroctets`

```
504 \def\SOUL@addfouroctets#1#2#3#4{%
505 ! \SOUL@addtoken{\noexpand#1\noexpand#2\noexpand#3\noexpand#4}}%
506 + \SOUL@addtoken{\etex@unexpanded{#1#2#3#4}}}%
507 }
```

### 2.7.1 Redefinition of `\SOUL@splittoken`

`\SOUL@splittoken` Macro `\SOUL@splittoken` separates the first token or token group from a word and redefines the word to contain the remaining tokens. However if the remaining tokens are a token group, then the curly braces will be removed and the token group is splitted by the next call of `\SOUL@splittoken`. The redefinition avoids the removal of curly braces around the remaining tokens.

```
508 \SOUL@redefine\SOUL@splittoken#1#2\SOUL@stop{%
509 \global\SOUL@token={#1}%
510 \global\SOUL@word={#2}%
511 }#1{%
512 \global\SOUL@token={#1}%
513 \SOUL@remainingtoken\relax
514 }
```

`\SOUL@remainingtoken`

```
515 \long\def@gobble#1{}
516 \def\SOUL@remainingtoken#1\SOUL@stop{%
517 \global\SOUL@word=\expandafter{\@gobble#1}%
518 }
```

## 2.8 Patches

The fixed `\SOUL@splittoken` allows to remove the double sets of curly braces in other macros of package `soul-ori`. The benefit is that implicate kernings are more often detected and fixes a bug in package `soul-ori`. The disadvantage is incompatibility. The width of the resulting strings may change.

`\SOUL@flushcomma`

```
519 \SOUL@redefine\SOUL@flushcomma{%
520 \ifcase\SOUL@comma
521 \or
522 \edef\x{\SOUL@word={\the\SOUL@word,}}\x
523 \or
524 \edef\x{\SOUL@word={\the\SOUL@word{,}}}\x
525 \fi
526 \SOUL@comma\z@
527 }{%
528 \ifcase\SOUL@comma
529 \or
530 \edef\x{\SOUL@word={\the\SOUL@word,}}\x
531 \or
532 \edef\x{\SOUL@word={\the\SOUL@word{,}}}\x
533 \fi
534 \SOUL@comma\z@
535 }
```

`\SOUL@flushapo`

```
536 \SOUL@redefine\SOUL@flushapo{%
537 \ifcase\SOUL@apo
538 \or
539 \edef\x{\SOUL@word={\the\SOUL@word'}}\x
540 \or
541 \edef\x{\SOUL@word={\the\SOUL@word{''}}}\x
542 \fi
```

```

543 \SOUL@apo\z@
544 }{%
545 \ifcase\SOUL@apo
546 \or
547 \edef\x{\SOUL@word={\the\SOUL@word'}}\x
548 \or
549 \edef\x{\SOUL@word={\the\SOUL@word{''}}}\x
550 \fi
551 \SOUL@apo\z@
552 }

```

\SOUL@flushgrave

```

553 \SOUL@redefine\SOUL@flushgrave{%
554 \ifcase\SOUL@grave
555 \or
556 \edef\x{\SOUL@word={\the\SOUL@word'}}\x
557 \or
558 \edef\x{\SOUL@word={\the\SOUL@word{''}}}\x
559 \fi
560 \SOUL@grave\z@
561 }{%
562 \ifcase\SOUL@grave
563 \or
564 \edef\x{\SOUL@word={\the\SOUL@word'}}\x
565 \or
566 \edef\x{\SOUL@word={\the\SOUL@word{''}}}\x
567 \fi
568 \SOUL@grave\z@
569 }

```

\SOUL@addgroup

```

570 \SOUL@redefine\SOUL@addgroup#1#2{%
571 {%
572 \let\protect\noexpand
573 \edef\x{%
574 \global\SOUL@word={%
575 \the\SOUL@word
576 {\noexpand#1#2}}%
577 }%
578 }%
579 \x
580 }%
581 \SOUL@scan
582 }#1#2{%
583 \begingroup
584 \let\protect\noexpand
585 \edef\x{\endgroup
586 \SOUL@word={%
587 \the\SOUL@word
588 ! {\noexpand#1#2}}%
589 + {\etex@unexpanded{#1#2}}}%
590 }%
591 }%
592 \x
593 \SOUL@scan
594 }

```

\SOUL@addmath

```

595 \SOUL@redefine\SOUL@addmath$#1${%
596 {%
597 \let\protect\noexpand
598 \edef\x{%

```

```

599     \global\SOUL@word={%
600     \the\SOUL@word
601     {\hbox{#1$}}}%
602     }%
603     }%
604     \x
605     }%
606     \SOUL@scan
607 }$#1${%
608 \begingroup
609 \let\protect\noexpand
610 \edef\x{\endgroup
611 \SOUL@word={%
612 \the\SOUL@word
613 !   {\hbox{#1$}}}%
614 +   {\etex@unexpanded{\hbox{#1$}}}%
615     }%
616     }%
617     \x
618     \SOUL@scan
619 }

```

\SOUL@addprotect

```

620 \SOuL@redefine\SOUL@addprotect#1#2{%
621   {%
622     \let\protect\noexpand
623     \edef\x{%
624       \global\SOUL@word={%
625       \the\SOUL@word
626       {\hbox{#2}}}%
627     }%
628     }%
629     \x
630   }%
631   \SOUL@scan
632 }#1#2{%
633 \begingroup
634 \let\protect\noexpand
635 \edef\x{\endgroup
636 \SOUL@word={%
637 \the\SOUL@word
638 !   {\hbox{#2}}}%
639 +   {\etex@unexpanded{\hbox{#2}}}%
640     }%
641     }%
642     \x
643     \SOUL@scan
644 }

```

\SOUL@addtoken

```

645 + \SOuL@redefine\SOUL@addtoken#1{%
646 + \edef\x{%
647 + \SOUL@word={%
648 + \the\SOUL@word
649 + \noexpand#1%
650 + }%
651 + }%
652 + \x
653 + \SOUL@scan
654 + }#1{%
655 + \edef\x{%
656 + \SOUL@word={%
657 + \the\SOUL@word

```

```

658 +      \etex@unexpanded{#1}%
659 +    }%
660 +  }%
661 +  \x
662 +  \SOUL@scan
663 + }%

664 \SOuL@AtEnd%
665 \</package>

```

## 3 Installation

### 3.1 Download

**Package.** This package is available on CTAN<sup>1</sup>:

[CTAN:macros/latex/contrib/soul/soul.dtx](#) The source file.

[CTAN:macros/latex/contrib/soul/soul-ori.dtx](#) The source file.

[CTAN:macros/latex/contrib/soul/soul.pdf](#) Documentation.

### 3.2 Package installation

Install the package with the package manager of your TeXsystem. If you want to do it manually:

**Unpacking.** Run `tex` or `LATEX` on the `.ins` file

```
tex soul.ins
```

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

```

soul.sty      → tex/generic/soul/soul.sty
soul-ori.sty → tex/generic/soul/soul-ori.sty
soulutf8.sty → tex/latex/soul/soulutf8.sty
soul.pdf      → doc/generic/soul/soul.pdf
soul-ori.pdf  → doc/generic/soul/soul-ori.pdf
soul.dtx      → source/generic/soul/soul.dtx
soulori.dtx   → source/generic/soul/soul-ori.dtx

```

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`.

### 3.3 Refresh file name databases

If your TeX distribution (TeX Live, MiKTeX, ...) relies on file name databases, you must refresh these. For example, TeX Live users run `texhash` or `mktexlsr`.

## 4 References

- [1] Melchior Franz: *The soul package*; 2003/11/17;  
[CTAN:pkg/soul](#).
- [2] Dominique P. G. Unruh: *ucs.sty – Unicode Support*; 2004/10/17;  
[CTAN:pkg/unicode](#).

---

<sup>1</sup>[CTAN:pkg/soul](#)

- [3] Frank Mittelbach, Chris Rowley: *Providing some UTF-8 support via inputenc*; 2006/03/30;  
[CTAN:macros/latex/base/utf8ienc.dtx](#).

## 5 History

[2007/09/09 v1.0]

- First version.

[2016/05/16 v1.1]

- Documentation updates.

[2019/12/15 v1.2]

- Documentation updates.

[2023-02-18 v3.0]

- merged soulutf8 and the original soul into one package.

[2023-06-14 v3.1]

- Updated

## 6 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.

	<b>Symbols</b>		<code>\chardef</code> . . . . . 423, 450, 454, 458, 460
<code>\@PackageError</code> . . . . .	160	<code>\csname</code> . . . . .	26,
<code>\@PackageWarningNoLine</code> . . . . .	197		33, 62, 78, 88, 142, 165, 392, 393
<code>\@gobble</code> . . . . .	515, 517	<code>\CurrentOption</code> . . . . .	9, 174
<code>\@ne</code> 227, 229, 231, 233, 312, 314, 316, 318			
<code>\@nil</code> . . . . .	416, 419, 441, 445		
<code>\@onelevel@sanitize</code> <u>165</u> , 393, 467, 483			
<code>\@undefined</code> . . . . .	70, 141		
<code>\@</code> . . . . .	254, 339		
	<b>A</b>		<b>D</b>
<code>\active</code> . . . . .	133	<code>\DeclareOption</code> . . . . .	9, 145, 150, 174
<code>\advance</code> . . . . .	227,	<code>\DeclareRobustCommand</code> . . . . .	143, 148
	229, 231, 233, 312, 314, 316, 318	<code>\documentclass</code> . . . . .	141
<code>\afterassignment</code> . . . . .	190		
<code>\aftergroup</code> . . . . .	41		
	<b>B</b>		<b>E</b>
<code>\break</code> . . . . .	257, 277, 279, 342, 362, 364	<code>\empty</code> . . . . .	29, 30
		<code>\endcsname</code> . . . . .	26,
			33, 62, 78, 88, 142, 165, 392, 393
		<code>\endinput</code> . . . . .	41, 140
		<code>\endlinechar</code> . . . . .	16, 47, 83, 89, 101
		<code>\etex@unexpanded</code> . . . . .	
			476, 498, 506, 589, 614, 639, 658
	<b>C</b>		<b>H</b>
<code>\catcode</code> . . . . .	14,	<code>\hbox</code> . . . . .	269,
	15, 17, 18, 19, 20, 21, 22, 23, 24,		354, 601, 613, 614, 626, 638, 639
	25, 45, 46, 48, 49, 50, 51, 52, 53,	<code>\hfil</code> . . . . .	257, 342
	54, 55, 56, 57, 58, 59, 60, 61, 81,		
	82, 84, 85, 86, 90, 91, 92, 93, 94,		
	95, 96, 99, 100, 102, 103, 104,		
	105, 109, 111, 181, 182, 184, 185		
			<b>I</b>
		<code>\if</code> . . . . .	209, 246, 294, 331
		<code>\ifcase</code> . . . . .	374,
			377, 520, 528, 537, 545, 554, 562





\SOUL@scan	208, 245, 253, 275, 293, 330, 338, 360, 472, 488, 494, 581, 593, 606, 618, 631, 643, 653, 662	
\SOUL@spaceskip	274, 359	
\SOUL@spc	209, 246, 294, 331	
\SOUL@splittoken	508	
\SOUL@stop	239, 324, 508, 516	
\SOuL@stringcsnameu	440, 445	
\SOuL@stringemdash	404, 490	
\SOuL@stringendash	403, 484	
\SOuL@stringfour	428	
\SOuL@stringmacrocolon	439, 444	
\SOuL@stringnobreakspace	405, 468	
\SOuL@stringoctets	415, 419	
\SOuL@stringthree	426	
\SOuL@stringtwo	424	
\SOuL@stringundeferr	440, 445	
\SOuL@stringUTFviii	415, 419	
\SOuL@temp	191, 194, 196, 198, 201, 422, 424, 426, 428, 448, 449, 452, 453, 456, 457, 466, 467, 468, 482, 483, 484, 490	
\SOUL@token	509, 512	
\SOUL@word	510, 517, 522, 524, 530, 532, 539, 541, 547, 549, 556, 558, 564, 566, 574, 575, 586, 587, 599, 600, 611, 612, 624, 625, 636, 637, 647, 648, 656, 657	
\soulomit	271, 272, 356, 357	
\spaceskip	274, 359	
\strip@prefix	168, 171	
		<b>T</b>
\the	89, 90, 91, 92, 93, 94, 95, 96, 109, 522, 524, 530, 532, 539, 541, 547, 549, 556, 558, 564, 566, 575, 587, 600, 612, 625, 637, 648, 657	
\thr@@	213, 298, 427, 458	
\TMP@EnsureCode	106, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139	
\tw@	215, 217, 219, 300, 302, 304, 425, 460	
		<b>U</b>
\unskip	257, 342	
		<b>W</b>
\write	35, 64	
		<b>X</b>
\x	26, 27, 30, 34, 38, 40, 63, 68, 78, 87, 99, 411, 421, 435, 447, 522, 524, 530, 532, 539, 541, 547, 549, 556, 558, 564, 566, 573, 579, 585, 592, 598, 604, 610, 617, 623, 629, 635, 642, 646, 652, 655, 661	
		<b>Z</b>
\z@	431, 450, 526, 534, 543, 551, 560, 568	