The `sanitize-umlaut` package

Manual for version 1.00 (2016/05/28)

Thomas F. Sturm[1]

http://www.ctan.org/pkg/sanitize-umlaut

https://github.com/T-F-S/sanitize-umlaut

Abstract

The packages sanitizes umlauts to be used directly in index entries for MakeIndex and friends with `pdflatex`. This means, that inside `\index` an umlaut can be used as `"U` or `Ü`. In both cases, the letter is written as `"U` into the raw index file for correct processing with MakeIndex and `pdflatex`.

Contents

---

[1]Prof. Dr. Dr. Thomas F. Sturm, Institut für Mathematik und Informatik, Universität der Bundeswehr München, D-85577 Neubiberg, Germany; email: thomas.sturm@unibw.de

1 Purpose of the Package

The packages sanitizes umlauts to be used directly in index entries for `makeindex` and friends with `pdflatex`. This means, that inside `\index` an umlaut can be used as `"U` or `Ü`. In both cases, the letter is written as `"U` into the raw index file for correct processing with `makeindex` and `pdflatex`.

The package is intended

- for documents in german language using the babel package with a setting identical or similar to `\usepackage[ngerman]{babel}`.

- for documents which are processed by `latex` or `pdflatex` (not `lualatex` or `xelatex`).

- for documents with an index which is processed using the MakeIndex program.

- for authors who like to use `\index{Übermaß}` instead of `\index{"Uberma"s}`.

All these conditions are satisfiable by simply including the `sanitize-umlaut` package.

An alternative would be to filter the resulting raw `.idx` index *before* `makeindex` is applied to create the final `.ind` index. Another alternative is to replace MakeIndex by Xindy or another index processor.

2 Package Usage

2.1 Prerequisites

The source document needs some encoding by `inputenc` since `pdflatex` is assumed as engine. For example:

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

Just some few encodings are supported by `sanitize-umlaut`. These are the most important for German language texts:

encoding	recognized as
utf8	utf8
latin1, ansinew, cp1252	latin1
applemac	applemac

Further, the `babel` package with German settings is needed:

```
\usepackage[ngerman]{babel}
```

2.2 Package Application

Now, the package application is simple. You just put

```
\usepackage{sanitize-umlaut}
```

into your document preamble *after* `inputenc` and, maybe, after `babel`. That is all.

2.3 Sanitized Characters

The umlauts and the sharp s are replaced by their `babel` shorthand codes which are written to the `.idx` file.

character	replacement
ä	"a
ö	"o
ü	"u
Ä	"A
Ö	"O
Ü	"U
ß	"s

2.4 Technical Information

The package uses `\inputencodingname` (set by the `inputencoding` package) to determine the current encoding.

The package redefines the `\@sanitize` macro at the begin of the document. It adds some encoding redefinitions to this macro. `\@sanitize` is used inside `\index` in a local group. If another package (besides `babel`) also changes this macro or uses it outside `\index`, strange things may happen.

3 Application Examples

file "german.ist" for the examples

```
actual '='    %  instead of @
quote  '!'    %  instead of "
level  '>'    %  instead of !
```

```
% !TeX encoding=UTF-8
% arara: pdflatex
% arara: makeindex: { style: german.ist, german: true }
% arara: pdflatex
\documentclass[a4paper,12pt]{article}
\usepackage[T1]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[ngerman]{babel}
\usepackage{makeidx}
\usepackage{sanitize-umlaut}
\makeindex
\begin{document}
\section{Basic Example}
Test äöüÄÖÜß.
\index{Aber}   \index{Arg}   \index{Ärger}
\index{Ofen}   \index{Ö - wie schön}   \index{oberhalb}
\index{Ufer}   \index{Übermaß}
\index{Latex=\LaTeX}   \index{Ärger>Index}
Test äöüÄÖÜß.
\printindex
\end{document}
```

1 Basic Example

Test äöüÄÖÜß.   Test äöüÄÖÜß.

1

Index

Aber, 1
Ärger, 1
   Index, 1
Arg, 1

LᴬTᴇX, 1

oberhalb, 1
Ö - wie schön, 1
Ofen, 1

Übermaß, 1
Ufer, 1

2

```
% !TeX encoding=UTF-8
% arara: pdflatex
% arara: makeindex: { style: german.ist, german: true }
% arara: pdflatex
\documentclass[a4paper,12pt]{article}
\usepackage[T1]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[ngerman]{babel}
\usepackage{makeidx}
\usepackage{sanitize-umlaut}
\usepackage[hyperindex,colorlinks]{hyperref}
\makeindex
\begin{document}
\section{Example with hyperref}
Test äöüÄÖÜß.
\index{Aber}   \index{Arg}   \index{Ärger}
\index{Ofen}   \index{Ö - wie schön}   \index{oberhalb}
\index{Ufer}   \index{Übermaß}
\index{Latex=\LaTeX}   \index{Ärger>Index}
Test äöüÄÖÜß.
\printindex
\end{document}
```

1 Example with hyperref

Test äöüÄÖÜß.   Test äöüÄÖÜß.

Index

Aber, 1
Ärger, 1
    Index, 1
Arg, 1

LaTeX, 1

oberhalb, 1
Ö - wie schön, 1
Ofen, 1

Übermaß, 1
Ufer, 1

1

2

```
% !TeX encoding=UTF-8
% arara: pdflatex
\documentclass[a4paper,12pt]{article}
\usepackage[T1]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[ngerman]{babel}
\usepackage[makeindex]{imakeidx}
\makeindex[options=-s german.ist -g]
\usepackage{sanitize-umlaut}
\begin{document}
\section{Example with imakeidx}
Test äöüÄÖÜß.
\index{Aber}   \index{Arg}   \index{Ärger}
\index{Ofen}   \index{Ö - wie schön}   \index{oberhalb}
\index{Ufer}   \index{Übermaß}
\index{Latex=\LaTeX}   \index{Ärger>Index}
Test äöüÄÖÜß.
\printindex
\end{document}
```

---

1 Example with imakeidx

Test äöüÄÖÜß.   Test äöüÄÖÜß.

Index

Aber, 1
Ärger, 1
   Index, 1
Arg, 1

LaTeX, 1

oberhalb, 1
Ö - wie schön, 1
Ofen, 1

Übermaß, 1
Ufer, 1

```
% !TeX encoding=UTF-8
% arara: pdflatex
\documentclass[a4paper,12pt]{article}
\usepackage[T1]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[ngerman]{babel}
\usepackage[makeindex]{imakeidx}
\makeindex[options=-s german.ist -g]
\usepackage{sanitize-umlaut}
\usepackage[hyperindex,colorlinks]{hyperref}
\begin{document}
\section{Example with imakeidx and hyperref}
Test äöüÄÖÜß.
\index{Aber}   \index{Arg}   \index{Ärger}
\index{Ofen}   \index{Ö - wie schön}   \index{oberhalb}
\index{Ufer}   \index{Übermaß}
\index{Latex=\LaTeX}   \index{Ärger>Index}
Test äöüÄÖÜß.
\printindex
\end{document}
```

1 Example with imakeidx and hyperref

Test äöüÄÖÜß.   Test äöüÄÖÜß.

Index

Aber, 1
Ärger, 1
  Index, 1
Arg, 1
LaTeX, 1

oberhalb, 1
Ö - wie schön, 1
Ofen, 1

Übermaß, 1
Ufer, 1

```
% !TeX encoding=UTF-8
% arara: pdflatex
\documentclass[a4paper,12pt]{article}
\usepackage[T1]{fontenc}
\usepackage[utf8]{inputenc}
\usepackage[ngerman]{babel}
\usepackage[makeindex]{imakeidx}
\indexsetup{level=\section*,noclearpage}
\makeindex[name=personen,title=Personenregister,options=-s german.ist -g]
\makeindex[name=allgemein,title=Allgemeines Register,options=-s german.ist -g]
\usepackage{sanitize-umlaut}
\begin{document}
\section{Example with multiple indexes}
Test äöüÄÖÜß.
\index[personen]{Huber, Hans}     \index[personen]{Hübner, Jörg}
\index[allgemein]{Aber}           \index[allgemein]{Arg}
\index[allgemein]{Ärger}          \index[allgemein]{Ofen}
\index[allgemein]{Ö - wie schön}  \index[allgemein]{oberhalb}
\index[allgemein]{Ufer}           \index[allgemein]{Übermaß}
\index[allgemein]{Latex=\LaTeX}   \index[allgemein]{Ärger>Index}
Test äöüÄÖÜß.
\clearpage
\printindex[allgemein]
\printindex[personen]
\end{document}
```

1 Example with multiple indexes

Test äöüÄÖÜß.   Test äöüÄÖÜß.

Allgemeines Register

Aber, 1   oberhalb, 1
Ärger, 1   Ö - wie schön, 1
   Index, 1   Ofen, 1
Arg, 1
                           Übermaß, 1
LaTeX, 1   Ufer, 1

Personenregister

Huber, Hans, 1   Hübner, Jörg, 1