The CorrIdx Package

J. K. Fink
Uni Leoben

October 12, 2004

Abstract

The CorrIdx package can be used to sort an index with chemical names containing leading numbers, etc. into a proper alphabetical order. It is intended as a preprocessor to the MakeIndex program.

Contents

The author is not liable for whatever will happen by the use of the material presented here.

NOTE: In the examples below, the INDEX comes always on a new page

1 Introduction

In a text containing chemical names like N,N'-diallyl p-phenyl diamine , 1,2-ethanediol, o-chlorobisphenol A, endo-methylenetetrahydrophthalic acid anhydride, tert-butanol, you may want to include these compounds into the index. If you use the index command \index{...}, LaTeX will produce from the entries

```
\index{1,2-ethanediol}
\index{o-chlorobisphenol~A}
\index{endo-methylenetetrahydrophthalic acid anhydride}
\index{tert-butanol}
\index{N,N'-diallyl p-phenyl diamine}
```

with the LaTeX command `\MakeIndex` a `*.idx` file looking like

```
\indexentry{1,2-ethanediol}{1} \indexentry{o-chlorobisphenol~A}{1}
\indexentry{endo-methylenetetrahydrophthalic acid anhydride}{1}
\indexentry{tert-butanol}{1} \indexentry{N,N'-diallyl p-phenyl
diamine}{1}
```

and the program MakeIndex produces

Index

Note that we differentiate in between the LaTeX command `\MakeIndex` and the MakeIndex program. We are using `\usepackage{makeidx}`.

Now, you may find that the sort order in the example above is not satisfactory. Therefore you must use the `@` in the index command to force a certain sort order. Try now the following index entries

```
\index{ethanediol@1,2-ethanediol}
\index{chlorobisphenol@o-chlorobisphenol~A}
\index{methylenetetrahydrophthalic@endo-methylenetetrahydrophthalic acid anhydride}
\index{butanol@tert-butanol}
\index{diallyl@N,N'-diallyl-p-phenyl diamine}
```

and the MakeIndex program will produce now a more satisfactory result. This is shown now:

Index

The MakeIndex program uses the `*.idx` file and produces a *.ind file. We try now another type of index entries like this

```
\index{chem @1,2-ethanediol}
\index{chem @o-chlorobisphenol~A}
\index{chem @endo-methylenetetrahydrophthalic acid anhydride}
\index{chem @tert-butanol}
\index{chem @N,N'-diallyl p-phenyl diamine}
\index{chem @curing}
\index{chem @synthesis}
```

A single run with the LaTeX command `\MakeIndex` in the text produces

```
\indexentry{chem @1,2-ethanediol}{5}
\indexentry{chem @o-chlorobisphenol~A}{5}
\indexentry{chem @endo-methylenetetrahydrophthalic acid anhydride}{5}
\indexentry{chem @tert-butanol}{5}
\indexentry{chem @N,N'-diallyl p-phenyl diamine}{5}
\indexentry{chem @curing}{5}
\indexentry{chem @synthesis}{5}
```

Now we run the corridx program as preprocessor and get a modified preprocessed file `*.idx` looking like

```
\indexentry{chem ethanediol@1,2-Ethanediol}{5}
\indexentry{chem chlorobisphenol~a@o-Chlorobisphenol~A}{5}
\indexentry{chem methylenetetrahydrophthalicacidanhydride@endo-Methylenetetrahydrophthal
\indexentry{chem butanol@tert-Butanol}{5}
\indexentry{chem diallylphenyldiamine@N,N'-Diallyl p-phenyl diamine}{5}
\indexentry{chem curing@Curing}{5}
\indexentry{chem synthesis@Synthesis}{5}
```

Further, the MakeIndex program produces from this file an *.ind file that is printed out as

Index

This is probably what you want.

2 The Corridx Program

The source code of the program is in Pascal compiled with Free Pascal
(`http://www.freepascal.org/`)

2.1 Using the Program

Corridx stands for correct the `*.idx` file. The simplest call is to run `corridx.exe`.
Then a file `in.txt` must be in the same directory as `corridx.exe` and the output comes in a file `out.txt`. Therefore you may copy your `*.idx` file into the
`in.txt`, run `corridx.exe` and collect the output from `out.txt`. This output
copy back in your `*.idx` file and run immediately the `MakeIndex` program.

If a command line parameter just one meaning a valid file name is there,
then the `corridx.exe` takes this as input file and copies the output back to this
file. If more than one command line is there, or if the filename is not correct, I
did not check what will happen.

I am running from Winedt both programs, as I have changed the index menu
into

```
Run("c:\fullpathtocorridx\corridx.exe %P\%N.idx","%P",0,0,"CorrIdx",1,1);
Prompt('Finished CorrIdx');
Exe('%B\Exec\MiKTeX\MakeIndex.edt');
```

The prompt is necessary to retard the program execution. Or else you may use
instead of the `Run` command the `WinExe` command with the synchronize option.
I did not check out this.

2.2 Program Description

Note that every index entry you create must contain a `@`, otherwise an error will
be created. For example create the entry `\index{chem @1,4-butanediol}`.
The program scans each line in the `*.idx` file. When it finds `chem` just behind
the opening bracket `{`, it ignores what is after `chem` until it finds the `@`. Then it
reads the name string until the closing bracket comes. There should be no `\}`
in between. Finally it reads the string with the page number. The name string
is copied to a sort string which is now processed:

1. Every text in between dollar signs is replaced by a blank.

2. Every non letter besides the tilde `~` is replaced by a blank. So you may
   protect `bisphenol~A`.

3. Every single letter in between blanks is replaced by a blank.

4. Special sequences like exo, endo, tert, sec, cis, trans, syn, anti, cyclo, spiro
   are replaced by blanks.

5. what is left over is made lower case.

The program is looking now in the sort string for the first appearance of a letter and makes in the name string this position as uppercase. Then in the sort string all spaces are eliminated. Finally, the index entry is rebuilt to result in `\indexentry{chem butanediol@1,4-Butanediol}{8}`

2.3 Further Options

Four types of index are allowed:

1. Index of chemicals `\index{chem @1,3-Butadiene}`

2. General Index `\index{gen @Synthesis}`

3. Index of acronyms `\index{acr LCD@LCD!Liquid crystal display}`

4. Index of other items `\index{whatever you want@abc}`

In an index entry of acronyms, a ! is compulsory. The example above produces the index:

Index

3 User Defined Commands

To get a better performance, it is a good idea to define the following commands:

```
\newcommand{\ia}[2]{#1 {\nobreak(#2)}\index{acr #2@#2!#1}}
% is is an acronym index. Example: \ia{Polyalkylene oxide}{PAO} %
\newcommand{\ic}[1]{#1\index{chem #1@#1}}
% ic is a chemical index
\newcommand{\ib}[2]{#1 {\nobreak(#2)}\index{acr #2@#2!#1}\index{chem #1@#1}}
% ib is both acronym and chemical index
\newcommand{\ig}[2][]{#2\index{gen #2@#2#1}}
% ig is a general index
```

We show now the use of these commands in the following sample text. You should look in the original LaTeX file.

Note that we have in the text also the commands

```
\newcommand\swallow[1]{\relax}
\index{acr   @\section*{Acronyms}\swallow|swallow}%
\index{chem  @\section*{Chemicals}\swallow|swallow}%
\index{gen   @\section*{General}\swallow|swallow}%
```

4 Sample Text

There are various types of novolak resins with different ortho to para ratios of the methylene linkages, high ortho novolak resins (HON), general-purpose novolak resins (GPN) and high para novolak resins (HPN).

The liquid-phase oxidation of cumene to cumene hydroperoxide (CHP) results in acetone and phenol. This is used for bisphenol A, bisphenol B, resorcinol, cresols, and xylenols. 2-Cyclohexyl-5-methylphenol is used for photoresists. m-Methoxyphenol, 2-naphthol, cardanol, and cardol, are other suitable phenols.

5 The Sample Text Verbatim

```
There are various types of novolak resins with different \ig[
resins]{ortho} to para ratios of the methylene linkages, \ia{high
ortho novolak resins}{HON}, \ia{general-purpose novolak
resins}{GPN} and \ia{high para novolak resins}{HPN}.

The liquid-phase oxidation of cumene to \ib{cumene
hydroperoxide}{CHP} results in \ic{acetone} and \ic{phenol}. This
is used for \ic{bisphenol~A}, \ic{bisphenol~B}, \ic{resorcinol},
\ic{cresol}s, and \ic{xylenol}s. \ic{2-Cyclohexyl-5-methylphenol}
is used for \ig[!positive]{photoresist}s. \ic{m-Methoxyphenol},
\ic{2-naphthol}, \ic{cardanol}, and \ic{cardol}, are other
suitable \ig[!other]{phenols}.
```

Index

Acronyms

Chemicals

General