

A L^AT_EX Addition for Formatting Indexes

Thomas Hofmann
Ciba-Geigy, Basle, Switzerland

A query by Jim Ludden (TUGboat, Vol. 7, No. 2, page 111) asked for a L^AT_EX addition for formatting indexes. Here is my proposal for UNIX systems: `latexindex`, a Bourne Shell script based on `texindex` by Robert Plamondon (location unknown). When using `latexindex`, there is no manual work on the `.idx` file.

`Latexindex` needs the four files `index.awk`, `index1.awk`, `index.sed`, and `index1.sed` whose full pathnames have to be set in the variables `INDEXAWK`, `INDEXAWK1`, `INDEXSED`, and `INDEXSED1` at the top of the file `latexindex`.

Description of `latexindex`

`Latexindex` processes an `.idx` file created by L^AT_EX and generates a `theindex` environment. The extension `.idx` for the input file name may be omitted. The result is left on the standard output. Example: If `test.tex` includes a command `\input test.ind`, call `latexindex test >test.ind` after having L^AT_EXed `test`, and then start L^AT_EX again.

The page numbers for entries of the form `\index{<entry>}` appear in roman font, for entries of the form `\index{<entry>*}`, in bold-face. The asterisk only marks a bold-faced page number at the end of an entry, otherwise it is regarded as part of the entry.

If an index entry appears on the same page both with and without an asterisk, the entry without the asterisk will be ignored. Consecutive page numbers are turned into ranges; however, bold-faced page numbers always appear separately. (Example: Index, 2-5, **6**, **7**, 10-11)

For sorting there is no difference between uppercase and lowercase letters. Digits fall before the letter 'A'. Special characters are ignored. However, two entries which differ only in special characters are not combined. Therefore `\index{Index}` and `\index{In\dex}` yield two different entries.

Subentries in index commands must have the form `\index{<main entry>,<subentry>}`. A subentry appears indented under the main entry. If a main entry has several subentries, the main entry appears only once, followed by a sorted list of subentries. It is not possible to create subsubentries. An entry without subentry may not include commas.

An index command having the form `\index{[<entry1>]<entry2>}` yields an index entry `<entry1>` but sorted by `<entry2>`. In this way, font switches and other L^AT_EX commands can be included in the index without using the command names for sorting. Example: `\index{[<it Index>]Index}` allows the italic index entry *Index* to be sorted by "In", not by "it". When using subentries, both the main entry and the subentry have to appear twice (e.g. `\index{[Document style,<tt article>]Document style,article*}`). For a bold-faced page number in this case the asterisk has to be placed, in the same way as described above, immediately before the closing curly brace (and not within the square brackets). Commands changing the environment which appear within brackets or curly braces lose their effect outside. A command within the main entry that changes the environment is not effective for the subentry (for `\index{[<it Index,Fonts>]Index,Fonts}` only "Index" appears in italic, "Fonts" appears in Roman).

Bugs

- An opening curly brace must not immediately follow a closing one "}-{".
- Neither the symbol '@' nor brackets may appear in index entries, except to mark the sorting entry. If these symbols have to appear in the index, they must be defined by appropriate commands. (Example: Instead of `\index{Program[me]}` you must type `\newcommand{\lbr}{[}`
`\newcommand{\rbr}{]}`
`\index{Program\lbr me\rbr}.`)
- There may be only *one* comma within a sorting entry (for separating the subentry), and there may also be only one comma in the brackets.
- Main and subentries must not be nested in common curly braces (except the outer braces of the `\index` command). E.g. the command `\index{{\it Index,Bugs}}` is not allowed; the correct form is `\index{{\it Index},{\it Bugs}}`, or simply `\index{{\it Index},{\it Bugs}},`

Editor's note: Arrangements are being made to include the files comprising `latexindex` in the Unix distribution, and perhaps to make them available separately through TUG.