

Sourcebrowser

update

Taco Hoekwater

September 9, 2020

<https://source.contextgarden.net>:

- browser reimplemented in Lua
- new indexing script in Lua
- uses SCITE lexers

Screen layout:

- Search box
- Navigation pane
- File source

Search box:

- Search input area
- Search button
- Reset link

Optional search indexes:

- Comments
- Docs
- MkII
- Data

Search options:

- Whole words only
- Lua string pattern matches

Navigation pane:

- doubles as search results overview

File Source:

- File header
- File contents

File header:

- File name
- Download link
- File size
- File modification date (currently useless)

File contents:

- line numbers with search links
- highlighted sources (SCITE lexer)

There are two `mtxrun` scripts:

- ❑ `mtx-source.lua` is the web server's CGI script
- ❑ `mtx-index.lua` creates the search index

- the search indices are created offline, then uploaded
- these are Lua tables with the search term as key
- the values contain a list of file ids, line numbers, and match type for each occurrence
- actual file names are in a separate table connecting ids and names

- the default search index and the file name table are reloaded for each server request
- that is why the default search index has to be relatively small (8Mb right now)
- First attempt was more than 300Mb

The search index was minimised by:

- skipping some files (like hyphenation patterns)
- only indexing ASCII
- requiring a length of at least 2
- only indexing terms with at least one letter
- splitting strings into words
- dropping terms with a high occurrence

- navigation pane is sometimes hard to use
(it needs independent scrolling from the source)
- there is no 'back to top' link
- search indices are far from perfect

- AND/OR searches
- Run the search index database as a system service
- write a dedicated tokeniser to replace the SCITE lexers for indexing
- hyperlinks to definitions would be really nice ...

Demo time!