ManBearPig. A study on how modern UNIX sytem search their manual pages.

spiros thanasoulas st19@illinois.edu

October 27, 2020

Description

The goal of this project would be to create a report and possibly code improvements for the mandoc project (https://mandoc.bsd.lv/) and more specifically for its textual search features.

Background

UNIX system provide their documentation to the user through a set of tools collectively referred to as the Manual Page system. The well known man(1) command exists today on all UNIX systems but even on other platforms like MacOSX and android. Searching efficiently keywords and semantics has been of paramount importance for the user to quickly get to the relevant manual page and the command apropos(1) traditionally served that purpose, meaning doing database lookups. The databases are built with the makewhatis(1) tool.

Project proposal

We will investigate the C source code of the mandoc project, targeted on the modules of searching and database building. We will understand how it works and them potentially propose improvements or patches based on modern techniques learned from CS410: Text Information Systems @ UIUC course. One area that seems easily applicable is to enrich the semantic search capabilities with root word extraction.

Proposed Workflow

We propose that the analysis and development will be split across 5 6hr man - days of work

Day 1	Code and Documentation analysis. Understanding of how apropos searches keywords, what semantic capabilities it has and how makewhatis builds the database
Day 2	Analysis and Design of possible improvements to the data structure, and investigation of how an improved semantic search can be incorporated in the codebase.
Day 3	Development and Documentation
Day 4	Development and Documentation
Day 5	Final report

members

st19 / solo project