



Case Study on an Elektra storage plugin

BACHELOR'S THESIS

submitted in partial fulfillment of the requirements for the degree of

Bachelor of Science

in

Software and Information Engineering

by

Jakob Fischer

Registration Number 01427178

to the Faculty of Informatics

at the TU Wien

Advisor: Dr.techn. Dipl.-Ing. Markus Raab

Vienna, 27th November, 2019

Jakob Fischer

Markus Raab

Erklärung zur Verfassung der Arbeit

Jakob Fischer

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 27. November 2019

Jakob Fischer

Abstract

Elektra is a library for handling configuration data, which also can mount external configuration files into it's system with the help of storage plugins. These storage plugin convert the file specific format into Elektra's format. As part of this thesis, we wrote such a storage plugin, for enabling the usage of TOML configuration files within Elektra. We evaluated the written plugin on the FLOSS project LCDproc, where we checked, if the plugin correctly reads and writes the configuration. Additionally, we compared the reading speed of our plugin with two other existing storage plugins within LCDproc. The newly written plugin could read and write TOML configuration without any loss of information. It's reading speed of TOML files was as good as the two other plugins, we matched it against.

Contents

Abstract	v
Contents	vii
1 Introduction	1
1.1 TOML	1
1.2 Elektra	1
1.3 LCDproc	2
1.4 Research Question	2
2 Implementation	3
2.1 Plugin	3
2.2 Tools	3
2.3 Grammar	3
3 Methodology	5
4 Evaluation	7
4.1 Hardware Setup	7
4.2 Software Setup	7
4.3 LCDd	8
4.4 lcdproc	10
4.5 lcdexec	10
4.6 lcdvc	11
4.7 Benchmarks	13
4.8 Discussion	14
5 Future Work	15
6 Related Work	17
7 Conclusion	19
List of Figures	21
	vii

List of Tables	21
Acronyms	23
Bibliography	25

Introduction

The practical part of this thesis consisted of writing an Elektra storage plugin, which can handle the reading and writing of TOML (Tom's Obvious, Minimal Language) configuration files from and into the key database Elektra handles.

1.1 TOML

Tom's Obvious, Minimal Language is a configuration file format developed since February 2013 [1]. It claims to be a format that is easily to read and parse [2] and finds use in many different projects [3], such as Rust's project manager `cargo` [4] or Python's package installer `pip` [5].

1.2 Elektra

Elektra is a configuration framework for accessing configuration settings in a global key database [6]. An application using Elektra can read and write key/value pairs via calls to the Elektra library, making the implementation of an own configuration system obsolete. Furthermore, elektrified applications can access each others configuration settings, to provide better application interoperability.

Although Elektra is mainly written in C, there are bindings for other languages like java, python or ruby [7].

Elektra can also read from and write to different configuration file formats, like JSON, XML or ini [8]. When needed, applications can easily switch to a different file format, because configuration access is done with the Elektra API. Developers or Administrators no longer have to commit to one file format for configuration.

Support for different languages and file formats is implemented by the use of plugins, which can be enabled if necessary. With this modularity, Elektra can provide a wide

range of functionality, while simultaneously avoiding being a bloated library. Developers can compile Elektra with the exact set of needed functionality.

1.3 LCDproc

LCDproc is an open source project for displaying stats like CPU/RAM usage of a system on different kinds of LCD (Liquid Crystal Display) devices [9] [10]. It has a client/server architecture, where the clients provide their system stats to the server, which can display them on a LCD device.

1.4 Research Question

As part of this thesis, we wrote an Elektra storage plugin for reading and writing TOML files. We will test this newly written TOML plugin on ElektraInitiative's fork of LCDproc, where it gets modified to use Elektra for it's handling of configuration. The TOML plugin will be used as the storage plugin for the LCDproc configuration. For evaluation, TOML configuration files for each tested part of LCDproc were created.

There are three questions we want to answer for the TOML plugin:

RQ1 Does the plugin correctly read the LCDproc configuration?

RQ2 Does the plugin correctly write the LCDproc configuration without any loss of information?

RQ3 Does the plugin put an unreasonable overhead on LCDproc loading times?

When answering RQ1, we also must ensure that we are reading the configuration from our supplied file, and not the default values supplied by the Elektra specification.

Ad RQ2 we also have to check, if comments and newlines - which don't affect the LCDproc execution practically, but are only for the user to read - are preserved.

Implementation

The plugin is implemented in C, using the C99 standard. It requires POSIX for some checks using regular expressions.

2.1 Plugin

The TOML parser is realized as an Elektra storage plugin. The plugin exposes two functions, **ElektraTomlGet** for reading, and **ElektraTomlSet** for writing, as expected of an Elektra storage plugin.

2.2 Tools

It uses two external tools for generating a parser for the toml file format: **Flex** (Fast Lexical Analyzer) [11] for lexical analysis and **bison** [12] for parsing. The key generation is done in the bison parser with the use of actions.

In order to generate the appropriate Elektra keys, the parser has access to a driver, which contains all needed Elektra logic. We adapted this driver oriented approach from the **yambi** [13] plugin. The driver contains functions in the style of `driverEnterKey/driverExitTable`, to clarify, on which point of parsing a certain function shall be invoked.

2.3 Grammar

Since bison generates LR (Left-to-Right, Rightmost derivation in reverse) parsers, grammar rules should be left-recursive. Otherwise, when using right-recursive rules, the parser would first need to shift the whole recursive expression onto the stack before it can start reducing it. The parser could run out of stack space, if it must parse right-recursive rules

too deep, as it may happen when parsing a huge array. Therefore, the written parser only contains left-recursive rules.

The parser makes use of so-called midrule-actions, which are parser actions that are not placed at the end of a grammar rule. This can affect the resulting grammar, since after executing a mid-rule action, the parser has to commit to the chosen parse branch [14]. We tried to minimize the usage of midrule-actions, to make the grammar as independent of actions as possible. However, especially for possibly nested structures like arrays or inline tables they were unavoidable.

Methodology

In this evaluation, we will check four parts of LCDproc: The server daemon LCDd and three clients, namely `lcdproc`, `lcdexec`, `lcdvc`. To valid the way, the TOML plugin reads/writes configuration, we will force a roundtrip of the file by changing a value within that file with the Elektra command `kdb set`. When we call `kdb set`, the configuration first gets read from the file, the value to set is updated, and then the keys are written back into the file. After we roundtripped a file, we compare it with its original file by the help of the `diff` command. This `diff` command is always invoked with the `-w` switch, so that whitespace differences get ignored. The TOML plugin currently preserves whitespace only before comments (tabs get converted into four spaces).

The TOML files copied during installation either differ from the default configuration loaded by the specification, or we will alter the keys of the file by a call of `kdb set` before we execute the program. During execution, we check if the program starts without any errors concerning the reading of the TOML file. We also check, if we can observe the expected difference from the default configuration, to ensure that we really are using the values from our configuration file.

In order to check, if the plugin imposes an unreasonable overhead in the loading times of LCDproc, we conducted benchmarks. There, we compared the loading time of LCDproc while using the existing storage plugins `yamlepp`, `ni` and our TOML for loading configuration.

Evaluation

4.1 Hardware Setup

The evaluation was executed on the following hardware setup:

```
Processor: Intel i7-4790K @ 4.20 GHz  
RAM: 16 GB (2 x 4GB, 1 x 8GB), DDR-3 @ 1600 MHz  
HDD: Seagate Desktop ST2000DM001, 2TB, 6 GB/s
```

4.2 Software Setup

The following versions of Elektra and LCDproc were used for evaluation:

```
Elektra:  
repository: github.com/bauhaus93/libelektra.git  
branch: 'plugin_toml'  
commit: #e75d60f8699d8ea80f0f099364b3b8f03ec4413f  
LCDproc:  
repository: github.com/bauhaus93/lcdproc.git  
branch: 'toml_test'  
commit: #70bf70135190c270c6a75e0b975f591df323bb65
```

The evaluation was done within a docker container, which was built as follows.

```
cd libelektra  
docker build -t buildElektra-buster \
```

```
--build-arg JENKINS_USERID='id -u' \  
--build-arg JENKINS_GROUPID='id -g' \  
-f scripts/docker/debian/buster/Dockerfile \  
scripts/docker/debian/buster/
```

We extended this image by building another image on top of it. This dockerfile mainly installs/sets up sudo for installing Elektra and LCDproc.

```
FROM buildelektra-buster  
USER root  
RUN apt-get update && apt-get -y install sudo vim tmux  
RUN adduser jenkins sudo  
RUN echo '%sudo ALL=(ALL) NOPASSWD:ALL' >> /etc/sudoers  
USER jenkins  
WORKDIR /home/jenkins
```

As a last step, we invoke a setup script. It installs Elektra and LCDdproc, and then copies LCDproc TOML files to the `./config` directory.

```
./lcdproc/toml_setup_install.sh
```

This script also invokes the `post-install.sh` script, which will mount the configuration scripts at the point expected by LCDproc.

4.3 LCDd

The LCDd server displays stats sent by the client applications to LCD displays. For our purposes, we configured it to display it to curses. Since the default configuration LCDd file is very minimal, and most of the configuration is loaded by the specification, we created a LCDd TOML file for evaluation purposes. It is based on the old default configuration file of LCDd [15]. We converted this old configuration from its ini-like format into TOML. All driver configuration settings, which are not referenced in the specification [16], were removed.

The resulting TOML file has 606 lines including 369 lines of comments and 143 empty lines. It contains a table and a nested array. It also contains multiple TOML table arrays - one for each LCDproc driver.

We force a full read/write roundtrip of the configuration file by setting the currently used driver. This driver is already the active one in the file, but we just set it again for the roundtrip. Afterwards, we check if there are any differences between the original file and the newly written file.


```
kdb set '/sw/lcdproc/lcdd/#0/current/server/drivers/#0' '@/curses/#0'  
diff -ws ./lcdproc/LCDd.toml .config/LCDd.toml  
Files ./lcdproc/LCDd.toml and .config/LCDd.toml are identical
```

After the roundtrip, the resulting LCDd.toml is identical to the original file under disregard of whitespace.

We start and stop the LCDd daemon once, change some parameters of the curses driver, and start the daemon again.

```
LCDd -f  
kdb set '/sw/lcdproc/lcdd/#0/current/curses/#0/foreground' 'yellow'  
kdb set '/sw/lcdproc/lcdd/#0/current/curses/#0/background' 'blue'  
kdb set '/sw/lcdproc/lcdd/#0/current/curses/#0/backlight' 'blue'  
kdb set '/sw/lcdproc/lcdd/#0/current/curses/#0/size' '100x40'  
LCDd -f
```

The resulting output of LCDd is:

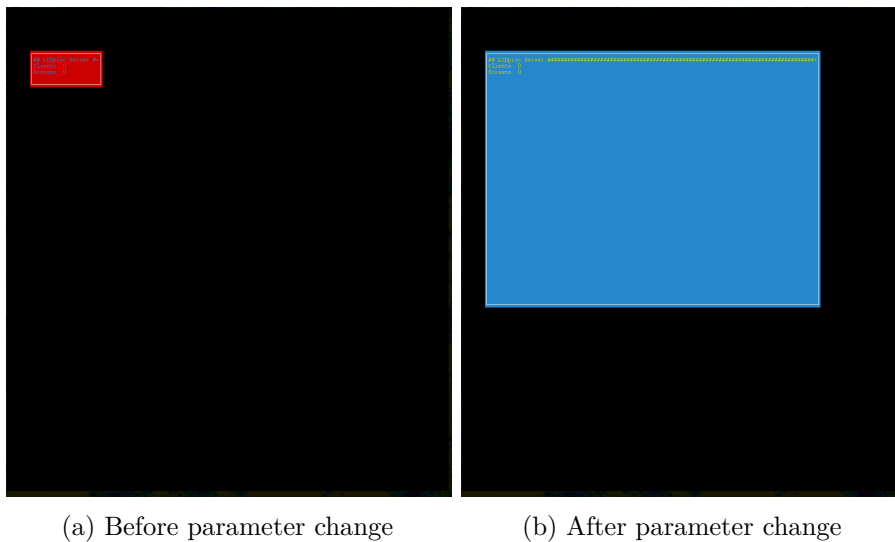


Figure 4.1: LCDd output

We observe, that we successfully changed the foreground/background color and window size of the daemon.

4.4 lcdproc

The lcdproc client is the main client for LCDproc. It sends the stats of it's host to the connected LCDd server. The configuration file used for the evaluation of lcdproc contains has 156 lines, including 58 lines of comment and 46 empty lines. The lcdproc configuration gets roundtripped by changing the used port from the default value of 13666 to 12345. Afterwards the file difference without whitespace gets checked.

```
kdb set 'user/sw/lcdproc/lcdproc/#0/current/lcdproc/port' '12345'
diff -w ./lcdproc/lcdproc.toml .config/lcdproc.toml
9c9
< port=13666
---
> port = 12345
```

The only difference found is the line with the port number, for which we changed the value. The rest of the file remained unchanged.

Then, the lcdproc client gets started without an active LCDd-daemon and is not expected to start completely, since it cannot connect to the server.

```
lcdproc -f
sock_connect: connect failed
Error connecting to LCD server localhost on port 12345.
Check to see that the server is running and operating normally.
```

In the error message, we see that the client tried to connect to port 12345, and not the default 13666.

4.5 lcdexec

With lcdexec, the LCDd server can execute preconfigured commands on the machine where lcdexec is running. We can add menus to the lcdexec configuration, which can contain further menus or the aforementioned commands, which can be any line of shell code. For the evaluation, we created a configuration file containing one menu with two options. The menu options execute 'echo a' and 'echo b', respectively. The lcdexec configuration file has 125 lines, including 36 lines of comment and 25 empty lines. We first roundtrip the configuration file by setting the main menu key (which already contains the now set value). Afterwards, the file is checked without whitespaces.

```
kdb set 'user/sw/lcdproc/lcdexec/#0/current/menu/main' \
      '@/menu/menu/#0'
```

```
diff -ws ./lcdproc/lcdexec.toml .config/lcdexec.toml
Files ./lcdproc/lcdexec.toml and .config/lcdexec.toml are identical
```

We see, there are no differences between the original and the roundtripped configuration file.

In a tmux (Terminal Multiplexer) session, LCDd and lcdexec get executed simultaneously. In LCDd, we navigate into our configured lcdexec menu and execute the first option, which echoes the letter a.



Figure 4.2: Menu navigation in LCDd while lcdexec is connected

We see both of our configured menus, and, on selecting the first menu option, we also see that a gets printed in the lcdexec window. We also see a warning message, that an item id could not be found, which gets printed, when the server has sent an unknown command.

4.6 lcdvc

The lcdvc client can display a system console on a display which is controlled by LCDd. Our configuration file only contains two assignments, the rest of the 33 lines consist of comments and empty lines. In the created TOML configuration we set the devices to read from to /dev/urandom.

4. EVALUATION

```
kdb set user/sw/lcdproc/lcdvc/#0/current/lcdvc/vcsdevice \  
/dev/urandom  
kdb set user/sw/lcdproc/lcdvc/#0/current/lcdvc/vcsadevice \  
/dev/urandom  
diff -ws lcdproc/lcdvc.toml .config/lcdvc.toml  
31,32c31,32  
< vcsdevice="/dev/null"  
< vcsadevice="/dev/null"  
---  
> vcsdevice = "/dev/urandom"  
> vcsadevice = "/dev/urandom"
```

We see, that the only file differences are the values we changed by `kdb set` before.

When we start LCDd and `lcdexec` in a `tmux` session, we see that `lcdexec` forwards the stream of `/dev/urandom` to LCDd.

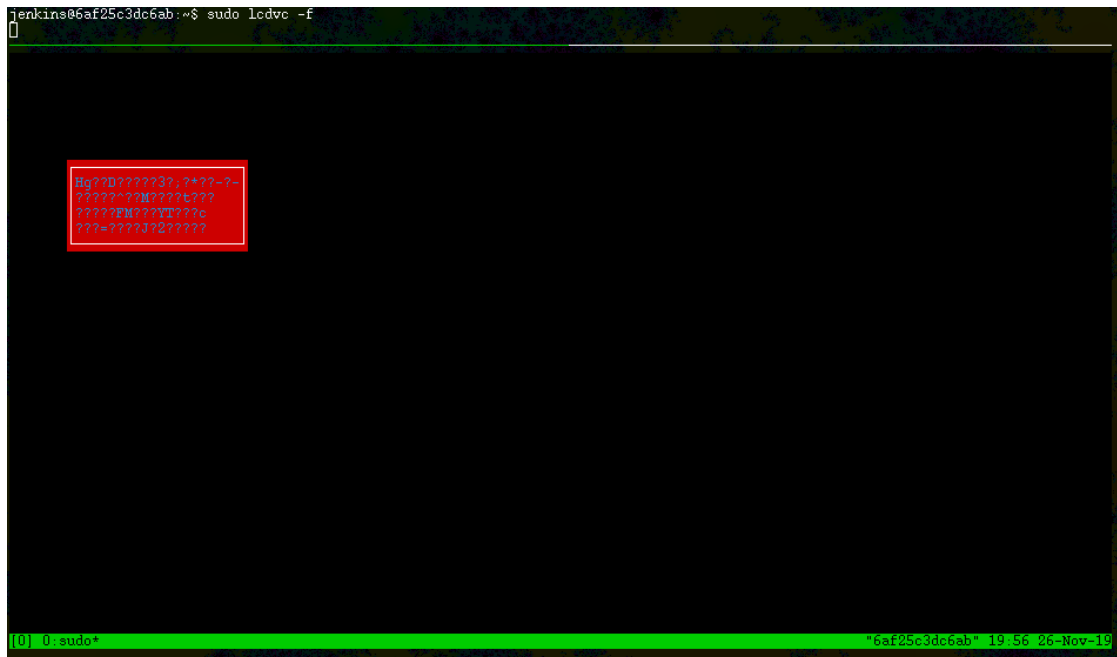


Figure 4.3: LCDd receiving `/dev/random` from `lcdexec`

4.7 Benchmarks

4.7.1 Additional Setup

For Elektra, we used our fork of Elektra [17]. For LCDproc, we used our fork of Elektra's fork of LCDproc [18].

TOML

```
elektra:
  branch 'thesis_bench'
  commit #0052eba6de6d46ec5265fdc01a1fd83046afe104
lcdproc:
  branch 'toml_test'
  commit #adadaaeclede9a5ecf840020ec5eee5585ecbf29
```

YAML

```
elektra:
  branch 'thesis_bench'
  commit #0052eba6de6d46ec5265fdc01a1fd83046afe104
lcdproc:
  branch 'bench_yaml'
  commit #e9e9cd97c1a5116214b53bffa4649aea00726d37c
```

ni

```
elektra:
  branch 'thesis_bench'
  commit #0052eba6de6d46ec5265fdc01a1fd83046afe104
lcdproc:
  branch 'bench_ni'
  commit #1d06f0c65cd7252bbb4a43aeb7bbd59e6b9118f0
```

We benchmarked the TOML plugin against two other Elektra storage plugins, yamlcpp for YAML (YAML Ain't Markup Language) and ni for INI. To use these plugin for comparison, we had to convert our TOML configuration file used in the previous evaluation into YAML and INI.

We also had to disable two required plugins from the yamlcpp plugin, because otherwise we would get an error from Elektra for too many mounted plugins on the start of LCDd. The disabled plugins were the `directoryvalue` and the `base64` plugin. The `directoryvalue` plugin would be needed if we wanted to write non-leaf keys into the file (and subsequently, convert them back on reading). The `base64` plugin would be needed if we had to handle binary key/values. For reading the configuration files, neither of those functionalities are need. For benchmarking purposes, we also modified

the LCDd main function, to not execute it's mainloop after initialization, but instead exit the application.

4.7.2 Execution

We benchmarked the three plugins by executing them repeatedly within a python script:

```
RUNS = 100
times = []
for _ in range(RUNS):
    start = time.time()
    subprocess.call(["LCDd", "-f"])
    times.append(1000. * (time.time() - start))

avg = sum(times) / RUNS
median = times[round(RUNS/ 2)]
print(f"runs = {RUNS}\navg = {avg}ms\n
      median = {median} ms\nmin = {min(times)} ms\n
      max = {max(times)} ms")
```

4.7.3 Results

plugin	avg	median	min	max
toml	75.03 ms	75.46 ms	66.01 ms	166.86 ms
yamlcpp	80.72 ms	83.55 ms	67.32 ms	134.60 ms
ni	74.16 ms	75.28 ms	65.89 ms	160.39 ms

Table 4.1: Benchmarks for LCDd

4.8 Discussion

In this evaluation, we tested not only basic key/value assignments, but also TOML features like tables and table arrays, and even nested arrays, which all could be read and restored correctly. We could observe that in all four programs of LCDproc, we would get identical files after reading/writing them when ignoring whitespace. This also means, that no comments got lost or misordered during roundtripping a file. Only during lcdexec evaluation we got a warning for a missing menu id, but since our configured command got executed anyway, we assumed success. This may be no error attributable to the written plugin.

The benchmarks showed, that the reading speed of the plugin can match with two other existing Elektra storage plugins, with the ni plugin being the current default for LCDprocs configuration storage.

Future Work

One thing, that the TOML plugin currently cannot store is Elektra metadata. Other plugins, like `yamlcpp` and `ni` already support this feature. The `yamlcpp` plugin does this by assigning a special meaning to the value `'!elektra/meta'`, which indicates that the key must be handled in a different matter [19]. For the TOML plugin we could maybe store metadata info in the comments before key with metadata, also by introducing a special sequence which indicate the presence of metadata.

As an example, a key `'a'` with a type metakey of value `unsigned_short` could look like:

```
#!ELEKTRA_META! type=unsigned_short
a = 3
```

When reading a file with this special tagged comments, we could differentiate them already at the stage of lexing from normal comments. However, there is the question of how this would affect the readability of a TOML file, when it contains many many metadata.

Another thing that could be improved in the future, would be a more complete preservation of whitespace characters. Especially whitespace used as indentation in front of keys/tables/table arrays/array elements would be very nice, since they improve human readability of the file. We actually tried to preserve this kind of whitespace, but turned it back, since the grammar got too error prone. When implemented properly, two points need to be taken care of: On the one side, the bison grammar should emit no reduce/reduce or shift/reduce ambiguities. On the other side, a valid TOML file should not be invalid because a valid spot for whitespace got overlooked in the grammar. When we first tried to implement this feature, we struggled with these two problems, which seemed to amplify each other. If we would eliminate more grammar ambiguities, we would get more files parsed as invalid and vice-versa.

Related Work

In his master's thesis, Markus Raab laid the groundwork for Elektras plugin based structure [20]. He also describes the problems that a storage plugin has to solve, like preserving the order of the data in the file or avoiding any loss of information between roundtrips of a file.

A general overview on how to create a parser with both flex and bison can be found in the book written by John Levine [21]. There he also elaborates on which kinds of conflicts can occur when writing a parser with bison and how to resolve them. He also writes about how to emit useful error messages from within the lexer/parser.

A more theoretical background on how formal languages could be parsed can be found in Stefano Reghizzi's book on formal languages [22].

Conclusion

In this paper, we wanted to verify, if our written TOML plugin reads and writes the LCDproc configuration correctly. We also wanted to survey, if the plugin puts an unreasonable overhead on the LCDproc startup time.

In our evaluation, we saw that the plugin could handle the configuration needed by LCDproc. It would do so in both reading and writing direction. No information got lost when reading / writing different configuration files of LCDproc.

Furthermore, the benchmarks showed that it has a decent reading speed of the LCDd configuration file, compared to two other storage plugins.

List of Figures

4.1	LCDd output	9
4.2	Menu navigation in LCDd while lcdexec is connected	11
4.3	LCDd receiving /dev/random from lcdexec	12

List of Tables

4.1	Benchmarks for LCDd	14
-----	-------------------------------	----

Acronyms

API Application Programming Interface. 1

Flex Fast Lexical Analyzer. 3

FLOSS Free/Libre/Open Source Software. v

INI Initalization. 13

LCD Liquid Crystal Display. 2, 8

LR Left-to-Right, Rightmost derivation in reverse. 3

POSIX Portable Operating System Interface. 3

tmux Terminal Multiplexer. 11, 12

TOML Tom’s Obvious, Minimal Language. v, 1–3, 5, 8, 11, 13–15, 19

YAML YAML Ain’t Markup Language. 13

Bibliography

- [1] Toml git contributions. <https://github.com/toml-lang/toml/graphs/contributors>. Accessed 2019-10-23.
- [2] Toml readme. <https://github.com/toml-lang/toml/blob/master/README.md>. Accessed 2019-10-23.
- [3] Toml wiki. <https://github.com/toml-lang/toml/wiki>. Accessed 2019-10-23.
- [4] Cargo github. <https://github.com/rust-lang/cargo/>. Accessed 2019-10-23.
- [5] Pip reference guide. <https://pip.pypa.io/en/stable/reference/pip/>. Accessed 2019-10-23.
- [6] libelektra. <https://www.libelektra.org/home>. Accessed 2019-10-23.
- [7] Elektra bindings. <https://www.libelektra.org/bindings/readme>. Accessed 2019-10-23.
- [8] Elektra storage plugins. <https://www.libelektra.org/plugins/readme#storage>. Accessed 2019-10-23.
- [9] Lcdproc website. <http://lcdproc.omnipotent.net>. Accessed 2019-11-19.
- [10] Lcdproc github. <https://github.com/lcdproc/lcdproc>. Accessed 2019-11-19.
- [11] Flex github. <https://github.com/westes/flex>. Accessed 2019-11-19.
- [12] Bison website. <https://www.gnu.org/software/bison>. Accessed 2019-11-19.
- [13] Elektra yambi plugin. <https://github.com/ElektraInitiative/libelektra/tree/master/src/plugins/yambi>. Accessed 2019-11-19.
- [14] Bison midrule conflicts. https://www.gnu.org/software/bison/manual/html_node/Midrule-Conflicts.html. Accessed 2019-11-19.

- [15] Lcdproc.conf. <https://github.com/lcdproc/lcdproc/blob/master/LCDd.conf>. Commit #d68c7e4da03362795edef12373cc80b3e553e29c.
- [16] Lcdd spec. <https://github.com/bauhaus93/lcdproc/blob/master/server/specification/LCDd-spec.ini>. Commit #3e2b9d3f5cade3bbe21fdb5a9ealc9ffc1c994fa.
- [17] Bauhaus93 elektra fork. <https://github.com/bauhaus93/libelektra>. Accessed 2019-11-23.
- [18] Bauhaus93 lcdproc fork. <https://github.com/bauhaus93/lcdproc>. Accessed 2019-11-23.
- [19] Elektra yamlcpp plugin. <https://www.libelektra.org/plugins/yamlcpp>. Accessed 2019-11-27.
- [20] Markus Raab. A modular approach to configuration storage. Master's thesis, Technische Universität Wien, 2010.
- [21] John R. Levine. *flex & bison*. O'Reilly Media, 1 edition, 2009.
- [22] Stefano Crespi Reghizzi. *Formal Languages and Compilation*. Springer-Verlag London, 2009.