Chapter 1

Related Work

Library-specific linters are rare because they require a non-trivial amount of work to implement. Even more niche is linters for embedded DSLs. This thesis can also be seen as a rough framework for implementing a linter for an eDSL (esp. Parser datatype to represent DSL model). afaik nothing this complex has been done in scalafix

Related to scalafix, so it could be possible to do very similar things as I've done: Rust (I can see this as possible, ask Boogle people for thoughts?) – Clippy lint passes C# – Roslyn

Library-specific linters: xunit.analyzers – seems to be relatively simple one-off rules though, not a DSL

Things that could help: Squid quasiquotes: type-safe and hygienic quasiquotes – but only for scala 2.11/12 * squid quasiquotes have rudimentary function inlining features https://infoscience.epfl.ch/record/231700 Scala 3 macros: I wonder if this could help?

Related concepts to the Function stuff: Metaprogramming and multi-staged programming? a dual: instead of for optimisation/code-generation purposes, we do it just for stringifying Actually I feel like I've borrowed a lot of stuff from metaprogramming optimisation techniques – Haskell Parsley itself

Chapter 2

Evaluation Plan

A large corpus of student WACC parsers is available to evaluate the real-world applicability of `parsley-garnish`. This can be utilised to assess the tool's ability to achieve quantitative and qualitative metrics.

The large number of parsers available to test can help evaluate the correctness and accuracy of each lint rule. We might find that certain rules are more prone to false positives than others, or that some automated refactorings may introduce new issues such as compilation errors. These quantitative metrics can be counted and used to evaluate rules against each other, as well as the overall effectiveness of the tool.

We can also estimate the proportion of issues that `parsley-garnish` is able to detect automatically. This can be done by comparing the number of issues detected automatically against the number of issues found manually by a code reviewer.

Qualitative metrics such as the quality of automated fixes can also be assessed by manually reviewing the output of the tool. For example, an issue I pointed out in the project plan is that the output of the left-recursion transformation may be in a form that is "ugly" when it could be simplified to a much more human-readable form. The quality of fixes would affect the practicality of the `parsley-garnish`, since users may be reluctant to apply automated fixes if the output quality can be poor.

Furthermore, the performance of `parsley-garnish` can be evaluated in terms of execution time and resource consumption. There are no direct competitors to benchmark against, but we can still evaluate the tool's performance against a baseline of acceptable performance, or perhaps a collection of other Scala static analysis tools deemed to be of similar complexity. I am thinking of benchmarking the tool during development as well, since it may affect design decisions such as implementing some rules as purely syntactic rather than semantic rules. I am not sure how much resource overhead a semantic rule would incur, so this could be insignificant or a major tradeoff (between speed, correctness, and accuracy) to consider.

Finally, it is possible to gather feedback from students interested in using `parsley-garnish` to improve their WACC parsers. User surveys and interviews would provide insights into the tool's usability, helpfulness, and areas for improvement. Although the timing of the WACC Lab does not align with the project timeline, I am hoping that some students may still be interested in trying out the tool on their final WACC submissions or possibly personal projects using `parsley`.