

LiquidRust: Refinement Types for Imperative Languages with Ownership

Master's Thesis of

Carsten Csiky

at the Department of Informatics
Institute of Theoretical Informatics (ITI)

Reviewer: Prof. Dr. Bernhard Beckert
Second reviewer:
Advisor: Dr. Mattias Ulbrich
Second advisor: M.Sc. Florian Lanzinger, M.Sc. Sebastian Graf

4.4.2022 – 4.10.2022

Karlsruher Institut für Technologie
Fakultät für Informatik
Postfach 6980
76128 Karlsruhe

I declare that I have developed and written the enclosed thesis completely by myself, and have not used sources or means without declaration in the text.

PLACE, DATE

.....
(Carsten Csiky)

Abstract

Software correctness is a central goal for software development. One way to improve correctness is using strong and descriptive types, that express correctness properties as types, and verifying their compatibility with the programs by using type systems. In particular Refinement Types demonstrated, that even with a verification system that is restricted to a decidable logic, intricate properties can be expressed in a lightweight and gradual way and verified automatically. However existing approaches for adapting Refinement Types from functional to imperative languages proved hard without compromising on at least one of core features of Refinement Types. This thesis aims to design a Refinement Type system without such compromises by taking advantage of Rust's restriction to unique mutable references. I will define a Refinement Type language and typing rules and argue for their soundness. Additionally I will implement a prototype verifier to evaluate the effectiveness of the approach on selected examples.

Zusammenfassung

Deutsche Zusammenfassung

Contents

Abstract	i
Zusammenfassung	iii
1 Introduction	1
2 Foundations	3
2.1 Refinement Types	3
2.2 Rust	3
2.2.1 Insights into Rust as a Verification Target	6
2.3 Rustc	6
3 Empirical Analysis of Use-Cases	9
3.1 Unsafe Rust	9
3.2 Mutability	11
3.3 Conclusions	12
4 The MiniCorten Language	15
4.1 Syntax	15
4.2 Semantics	17
5 The Refinement Type System	19
5.1 Features	19
5.1.1 Decidable, Conservative Subtyping	19
5.1.2 Mutable Values	19
5.1.3 Mutable References	20
5.1.4 Path Sensitivity	21
5.1.5 Strong Type Updates	22
5.1.6 Modularity	22
5.1.7 Atomic Updates	22
5.2 Function Declaration Type Checking	24
5.3 Sub-Typing Rule: $\Gamma \vdash \tau \leq \tau'$	24
5.4 Sub-Context Rules: $\Gamma \leq \Gamma'$	25
5.5 Expressions Typing: $\Gamma \vdash e : \tau$	25
5.6 Statement Type Checking $\Gamma \vdash s \Rightarrow \Gamma'$	25
5.7 Soundness of the Type System	26
5.8 Extensions	28
5.8.1 Records / structs	28
5.8.2 Algebraic Data Types	29

5.8.3	Inference	29
5.8.4	Predicate Generics	29
5.8.5	Convenience Improvements	29
6	Implementation: CortenC	31
7	Evaluation	37
7.1	Maximum using Path Conditions	37
7.2	Recursion: Fibonacci-Numbers	38
7.3	Loop Invariants: Proof of the Gauß Summation Formula	39
7.4	Complex Mutable References	40
7.5	Rephrasing built-ins in terms of Refinement Types	40
7.6	Interoperability with Other Tools	41
8	Related Work	43
8.0.1	Alternative Approaches for Handling Mutability	43
8.1	Comparison to Flux	44
9	Conclusion & Future Work	47
	Bibliography	49

List of Figures

2.1	Diagram providing an overview of the Rustc Compiler Stages	7
3.1	Scatter plot of a point representing creates, relating the number of unsafe occurrences to the number of unsafe LOC	10
3.2	Cumulative, Logarithmic Histogram of the Amount of unsafe Uses in each Category	11
3.3	Ratio of Immutable to Mutable Items of Different AST Nodes. Ratio is Relative to Total Number Of Occurrences	12
8.1	Figure showing a value space, where inclusion can not be shown	46
.1	Cumulative, Logarithmic Histogram of the Amount of unsafe Uses in each Category	51
.2	Logarithmic Histogram of the Number of Crates that contain n % immutable / mutable items	52

List of Tables

3.1	Number of Unsafe Uses by Crate Category	11
-----	---------------------------------------------------	----

1 Introduction

With increasing amount and reliance on software, ensuring the correctness of programs is a vital concern for the future of software development. Although research in this area has made good progress, most approaches are not accessible enough for general adoption by the developers. Especially in light of a predicted shortage of developers[3], it is not sufficient to require developers to undergo year-long training in specialized and complex verification methods to ensure the correctness of their software. It is therefore crucial to integrate with their existing tooling and workflows to ensure the future high quality of software. One avenue for improving accessibility for functional verification is extending the expressiveness of the type system to cover more of the correctness properties. Using type systems for correctness was traditionally prevalent in purely functional languages where evolving states are often represented by evolving types, offering approachable and gradual adoption of verification methods. Tracing evolving states in the type system would be especially useful for languages with mutability, since substantial parts of the behaviour of the program is expressed as mutation of state. In particular Rust seems like a promising target language, because mutability is already tracked precisely and thus promising functional verification for relatively minimal effort on the programmer's part.

The goal of the thesis is to show that Refinement Types can be idiomatically adapted to languages with unique mutable references. The type system presented in this thesis enables gradual adoption of lightweight verification methods in mutable languages.

The type system is sound and effective in the identified use-cases. A feasibility study on minimal examples of challenging use cases shows how useful the proposed verification system is.

Specifying or verifying complete Rust modules or the entire Rust language is not the goal of the thesis. In particular `unsafe` Rust will not be taken into account in specification nor implementation. Implementing Liquid Type inference is also not a goal of the thesis.

The accompanying implementation extends the Rust compiler, enables automatic parsing of the refinement type language and automated type checking of a subset of Rust, as well as limited inference and error reporting. The thesis also gives a description of the syntax and semantics of the subset of Rust as well as the refinement type system.

The contributions of this thesis are:

1. Automatic empirical analysis of the usage of mutability and `unsafe` in Rust using syntactic information
2. Extension to the Rust type system to allow for refinement type specifications
3. Description and implementation of a type checker for the introduced type system
4. Evaluation of the type system on minimal benchmarks

The thesis is structured as follows: chapter 2 will give an overview of the foundation the thesis build upon. Thereafter in chapter 3 an empirical analysis of `unsafe` and mutability uses is performed to gain an understanding of how Rust is used. Chapter 4 defines the subset of Rust, that will be the basis for the type system. Next chapter 5 explains the actual type system and type checking, as well as justifying its correctness, followed by chapter chapter 6, which will provide additional information about how the type system was implemented. The implementation will then be tested on minimal benchmarks in chapter 7. Chapter 8 reports on related work and alternative approaches. Finally chapter 9 concludes the thesis and gives an overview over possible future work.

2 Foundations

We start by giving a basic overview over the fundamental concepts, that are relevant for our work. The semantics of our subset of Rust are discussed in chapter 4.

2.1 Refinement Types

The type system, that this thesis will adapt, is based on the Refinement Type system described by Vazou et al. [22] and Rondon et al. [18]. The central idea of Refinement Types is adding predicates to a language's type: For example, a type `i32` might be refined with the predicate $v > 0$. The refined type is written as $\{v : \text{i32} \mid v > 0\}$. In terms of semantics this means, that any inhabitant of that type must also satisfy the predicate.

The notion of refinements is embedded into the base language using subtyping: A refined type τ is a subtype of τ' if τ 's predicate implies the predicate of τ' (in the current typing context Γ). As is common with subtyping systems, the subtype can then be used in the place where the super type is expected. The following rule shows how subtyping is handled by Liquid types [18, p. 6]:

$$\frac{\text{SMT-Valid}(\llbracket \Gamma \rrbracket \wedge \llbracket e_1 \rrbracket \Rightarrow \llbracket e_2 \rrbracket)}{\Gamma \vdash \{v : b \mid e_1\} \leq \{v : b \mid e_2\}}$$

Refinement Types are related to dependent types, distinguished by the limited expressiveness of the type. Refinement types trade off expressiveness for automatic type checking. Allowed formulas are chosen from a SMT decidable logic. Rondon et al. [18] chose the EUFA as the logic fragment, meaning a propositional logic with the theories of equality, uninterpreted functions and linear arithmetic.

2.2 Rust

The Rust programming language [13], is a multi-paradigm systems programming language, originally developed by Graydon Hoare at Mozilla Research. Even though Rust is quite young, the first stable release published in 2015, it has attracted interest in a wide area of applications and organization. It is used everywhere from Volvo building embedded software for cars [4] to Western Digital implementing a NVMe driver for Linux in Rust [11] and of course the Rust compiler itself. It draws inspiration from functional and imperative programming languages. The core forms a set of features usually found in functional programming languages: Functions, closures, algebraic data types, type classes (called traits in Rust) and pattern matching, immutability are all common place in Rust. However, Rust is also influenced by C: Rust extends its ML base with optional mutability

of data and references and manual memory management. Many design decisions of Rust can be explained by Rust's ambitions to offer safe and clear ML-like abstractions without compromising on performance compared to C. For example, Rust elected to use manual memory management like C, which guarantees faster and predictable execution times when compared to garbage collection. Crucial, Rust does not compromise on safety to accomplish this: To this end, Rust incorporates the notion of affine types into its type system (called *ownership system* in Rust). This additional type system guarantees, Inspired by research language Cyclone introduced by Grossman et al. [6], this type system guarantees the absence of memory errors. Rust also offers a special language subset called unsafe Rust, which deactivates the static ownership checking and places the responsibility of ensuring memory safety on the programmer, analogous to C. unsafe Rust is used to provide low-level primitives, that expose a safe interface, that can be used in safe Rust. From here on out, we will focus on safe Rust.

Among other things, the ownership system is used to handle memory management, safe concurrency and simplifies reasoning about program behaviour. The next paragraph will explain how the ownership system works.

Aliasing XOR Mutability The ownership system is an implementation of affine types. It uniquely associates each memory object with an owner. The owner of a Rust object is the binding, that is responsible for freeing the object. In the listing 1, first `a` is the owner of the `User` object. An object may only be used once, at which point it is consumed and ownership is transferred to its new binding. In the example `greet(user)` uses `a` and transferring ownership to `greet`. At the end of `greet`, there is no new owner, as a result the user object is freed.

```
struct User { name: String, age: u32 }

fn greet(user: User) -> String { // greet takes ownership for user
    format!("Hi {}", user.name)
} // End of scope: Owned value user is dropped
...
let a = User { name: "Alice", age: 25 };
let result = greet(a); // Transfer ownership of user to greet
println!(a.name); // error: use of moved value `a`
```

Listing 1: Example demonstrating the Ownership System: `greet(a)` transfers ownership of `a` to `greet`

On closer inspection, handing over the ownership of `a` to `greet` is not sensible: `greet` just needs to read the value, but should not take ownership of `a`. For that purpose, Rust offers a different kind of value transfer called *references*, which have similar execution semantics to pointers. Because creating a reference does not transfer the ownership of the pointee, it is called a *borrow*. Listing 2 rephrases `greet` to only take a borrow of the user, instead of taking ownership. Because `user` is not owned by `greet`, it is not permitted to extract `user.name` (of Type `String`), instead only borrows to the name are allowed.

```
fn greet(user: &User) -> String { // greet borrows user
    let name_ref : &String = &user.name; // borrow name from user
    format!("Hi {}", name_ref)
} // End of scope: The borrows are dropped, but the user object is not
...
let a = User { name: "Alice", age: 25 };
let result = greet(&a); // Lend a to greet
println!(a.name); // No error
```

Listing 2: Example demonstrating borrowing: `greet(&a)` lends `a` to `greet`

Borrowing, just like pointers, entail the risk of dangling pointers. Rust protects from this risk, by tracking the lifetimes (i.e. the required duration, where a reference must be valid) of objects. While interesting in its own right, it is not relevant for this thesis.

Mutable references can exacerbate the problem, which leads to the key insight of Rust: Memory unsafety can arise from the mere presence of two references where at least one is mutable. Thus, Rust ensures, that for every variable at every point during the execution, that variable can either be aliased or mutable, but never both at the same time. Listing 3 demonstrates that Rust will not accept a program where aliased and mutable access is possible at a time. To resolve the issue, the mutable reference should be borrowed after the call to `greet`.

```
fn age(user: &mut User) {
    (*user).age = user.age + 1;
}
...
let a = User { name: "Alice", age: 25 };
let b = &a;
let c = &mut a; // error: cannot borrow `a` as mutable because it is also borrowed as immutable
age(c);
let result = greet(&b); // Lend a to greet
println!(a.name);
```

Listing 3: Example demonstrating mutable borrowing: `age(a)` needs a mutable borrow of `user` to increase the age, but `a` is borrowed immutably, which is forbidden

In conclusion ownership makes both aliasing as well as mutability quite tame: If a variable is aliased, it must be immutable and as a result, represents just a value (like in pure functional languages). If a variable is mutable, it can not be aliased and as a result, any effect of the mutation is well known and locally visible. Although the ownership system was originally introduced to allow for safe, manual memory management, it turned out to be useful for other problems as well: For example concurrency, hardware access and automated reasoning all profit from the strong guarantees of safe Rust. The consequences for the latter will be explained in the next subsection.

2.2.1 Insights into Rust as a Verification Target

Rust justifies a new approach to verification, because it is possible to take advantage of guarantees provided by Rust's ownership system.

Rust features a few exceptional design decisions that profoundly influence the design of verification systems for it. The following paragraphs will elaborate on this.

Opaque Generics Rust does not provide a way to check a generic parameter for its instantiation. This means that we cannot extract any additional information from a generic parameter T . A function from T to T can therefore only be the identity function. Wadler [23] shows, that it is possible to derive facts about the behaviour of such polymorphic functions.

In contrast, languages with instance-of-checks allow an implementation of a polymorphic function to distinguish between different instances of the generic parameter, precluding this extensive reasoning.

Explicit Mutable Access A consequence of the ownership rules is, that a function can only mutate (part of) the state, that is passed to it as a parameter. There is no global state and there is no implicit access to an object instance. Thus any intention of mutating state must already be expressed in the function signature, which makes the specification of this mutation quite natural.

Minimal subtyping In contrast to object oriented languages, Rust has no user-extendable notion of object subtyping. Rust does have a notion of subtyping for its permission system, which serves the purpose of coercing mutable references $\&\text{mut } T$ to immutable references $\& T$ etc. where needed. At the type level Rust also has a concept of subtyping for traits. Because of the limited, well-known scope, these subtyping relations create no ambiguity with a refinement type system.

What Rust does not solve Even though Rust simplifies reasoning about mutability and aliasing of mutable data, Rust does not eliminate the need to reason about multiple references to mutable data. For example, listing 4 shows how `cell`'s type is influenced by changes to `r`. This does not mean, that the ownership rule *mutability XOR aliasing* are broken: Even though both `cell` and `r` are mutable, but only one is active.

```
let mut cell = 2;
let r = &mut cell;
*r += 1;
cell += 1;
assert_eq!(cell, 3);
```

Listing 4: Example of an apparent violation of ownership rules

2.3 Rustc

Since the implementation interacts with Rustc, the reference implementation of a Rust compiler, this section will summarize the relevant parts of Rustc mainly based on the Rustc dev guide [15].

Rustc offers an APIs for writing compiler plugins, which can access the intermediate languages of Rust at different stages during the compilation and allows the plugin to emit warnings and errors. At a high level, rustc is built with stages, that incrementally lower the source code to an executable binary. Rustc has multiple stages and associated data structures. Relevant for the thesis are the first few stages: Firstly the AST, which is a representation of the source code after lexing and parsing,

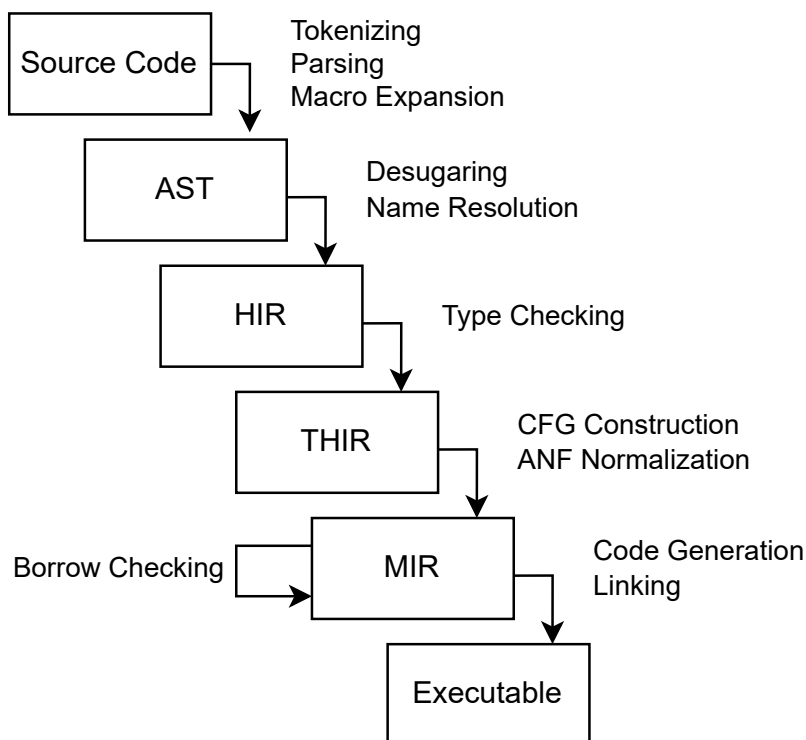


Figure 2.1: Diagram providing an overview of the Rustc Compiler Stages

Figure 2.1 shows the the main compiler stages of Rustc and the associated data structure. The first data structure, that is extracted from the source code is the AST (short for abstract syntax tree), which is the result of parsing and expanding macros. Based on the AST, the high-level intermediate representation (short HIR) is created by resolving the names and desugaring some language constructs. The HIR contains accurate source labels, which map all nodes in the HIR to their source code locations. The HIR is then used as the basis for type checking returning the typed high-level intermediate representation (short THIR), which annotates every node with its type. THIR also lowers some additional constructs like automatic dereferencing and overloading and more. In contrast to the previous representations, "[t]he THIR only represents bodies, i.e. «executable code»; this includes function bodies [...]. Consequently, the THIR has no representation for items like structs or traits." [21, p. 1] THIR will then be used to create the mid-level IR by drastically simplifying the language: The tree structure with complex expression of the THIR is turned in to a control-flow graph (short CFG) with basic blocks containing simple (i.e. non-nested) expressions.

Finally Rustc uses external tools like llvm to generate executable files and link them.

3 Empirical Analysis of Use-Cases

Before designing a system for Rust, it makes sense to gain some understanding of how Rust is used. For this purpose we will look at two key features of Rust, that influence how an approachable verification should look like. Firstly `unsafe` Rust, with a similar lack of guarantees to C, would make verification and specification significantly harder. But if the use of `unsafe` is limited, like intended by the language designers, it would allow us to focus on the safe part of Rust and leave the verification of `unsafe` Rust to more complex verification systems, like Prusti [2]. Secondly with mutability being the main difference to the traditional domain of Refinement Types, estimating the need for covering this language feature is interesting. The prevalence of mutability should also inform the acceptable level of user effort.

Thus, the use-case analysis should answer the following questions about Rust code:

1. How rare are uses of `unsafe`?
2. How much effort is acceptable when specifying `unsafe` code?
3. How common are mutable variables and references in Rust?
4. How much effort is acceptable when specifying mutable variables and references in Rust?

To check these assumptions an analysis of existing Rust code was performed¹. As a basis for the analysis, the Rust package registry `crates.io` was used. It contains the source code for both Rust libraries as well as various applications written in Rust (e.g. `ripgrep`).

We analyzed all published Rust crates (Rust's version packages) on February 2nd 2022 on `crates.io` with at least 10 versions², which totals 11 882 crates, containing 228 263 files with a combined code-base size of over 64 million lines of Rust code (without comments and white space lines)³.

The analysis parses these files and searches for certain AST patterns, which are subsequently extracted and saved. Thanks to using the tree-sitter parsing framework, the analysis framework can easily be extended to other queries and languages.

3.1 Unsafe Rust

Firstly we will be answering the question of `unsafe` usage in Rust. There is already some research on how `unsafe` is used in Rust. For example Astrauskas et al. [1] found, that

¹The source code is available at <https://gitlab.com/csicar/crates-analysis/>

²The limit of 10 versions is used to eliminate inactive and placeholder packages

³Calculated with `cloc`

about 76% of crates did not use any unsafe. On top of that, unsafe signatures are only exposed by 34.7% of crates, that use unsafe.

”The majority of crates (76.4%) contain no unsafe features at all. Even in most crates that do contain unsafe blocks or functions, only a small fraction of the code is unsafe: for 92.3% of all crates, the unsafe statement ratio is at most 10%, i.e., up to 10% of the codebase consists of unsafe blocks and unsafe functions.” [1, p. 13] Our data seems to confirm this: 8 044 of the 11 882 crates (67.7%) did not use any unsafe.

Astrauskas et al. also found, that ”however, with 21.3% of crates containing some unsafe statements and – out of those crates – 24.6% having an unsafe statement ratio of at least 20%, we cannot claim that developers use unsafe Rust sparingly, i.e., they do not always follow the first principle of the Rust hypothesis.” [1, p. 14]

Although true, when analyzing unsafety for our use case, it makes sense to further distinguish between libraries and executables crates: Libraries are intended to be used by other Rust programs: Usage of unsafe in libraries may not be as problematic as in executables, because libraries are written once but used by many applications, justifying higher verification effort.

In our analysis we found, that expectedly crates.io contains significantly more libraries than binaries⁴: About 81,7% of crates contained just a library, 10,3% contained just an executable target and 10,3% contained both. We also found, that libraries are much more likely to use unsafe Rust.

Figure 3.1 gives an overview over how unsafe sections are distributed in the crates.

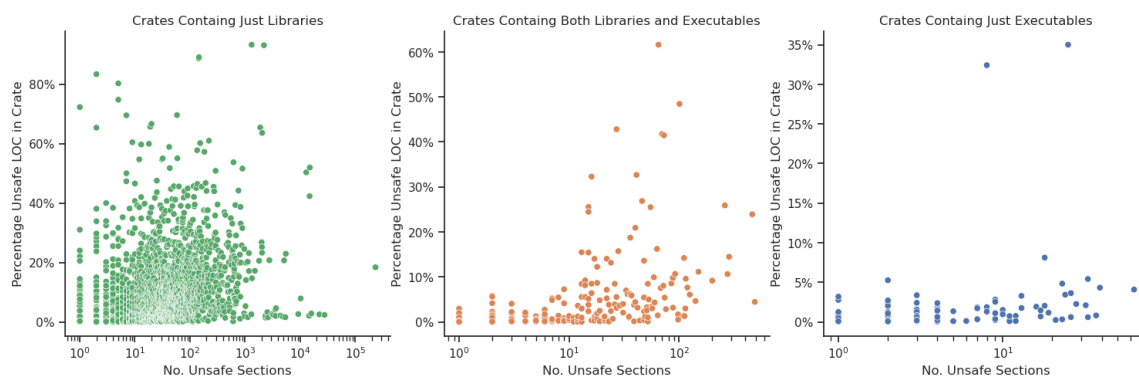


Figure 3.1: Scatter plot of a point representing crates, relating the number of unsafe occurrences to the number of unsafe LOC

Table 3.1 shows the result of our analysis: The total number of unsafe uses in an order of magnitude higher for binaries compared to libraries. The data in the table includes all crates except the outlier windows-0.32.0, which alone contains 233 608 uses of unsafe. Nearly 2/3 of all other library uses of unsafe combined. Looking at the distribution of unsafe uses in Figure 3.2, we can see, that this is an exception: Most other libraries do not use that much unsafe statements. We can also see, that even if a executable crate uses unsafe, it uses few. Note the difference in x- and y-axis scaling: No executable contains

⁴Libraries and Executables are distinguished by checking if they contain `bin` or `lib` target or one of the corresponding files according to the cargo naming convention.

Plots in Figure 3.1 sind ein bisschen klein - Außerdem ist in Figure 3.1 vielleicht eher die durchschnittliche unsafe-Block-Größe interessant? Dann bräuchte die y-Achse keine Log-Skala

more than 63 occurrences on unsafe blocks, but more than 10% of libraries contain more than 63 unsafe sections. On average, a library contains about 63 unsafe sections, while crates with binaries contain on average about 6 unsafe sections.

Crate contains	No. Crates	Total No. of Unsafe Sections	Total Unsafe LOC
Library	9 707	382 997	2 166 213
Both	1 224	7 720	51 004
Binary	951	940	5 873

Table 3.1: Number of Unsafe Uses by Crate Category

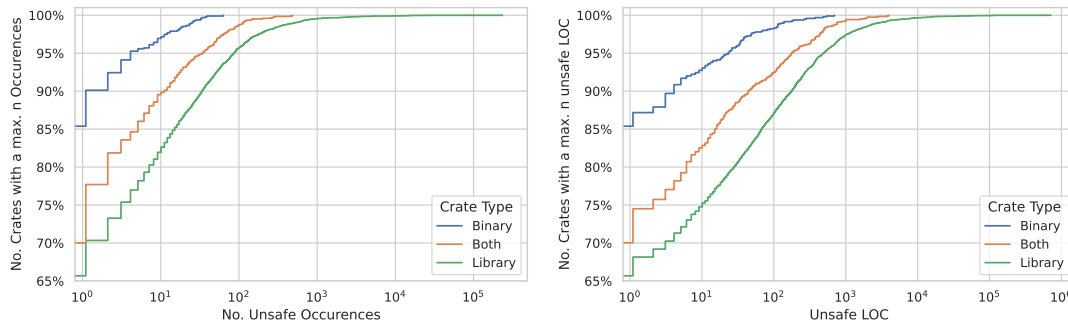


Figure 3.2: Cumulative, Logarithmic Histogram of the Amount of unsafe Uses in each Category

3.2 Mutability

Finally we will analyse how common mutable variables and references are used in Rust. The frequency of usage will inform the acceptable level of specification effort.

To analyse the dataset for usage patterns, we search the dataset for certain syntactical structures to infer mutability information about the following AST items:

- **Local Variable Definitions** can be tracked with high confidence. They occur in function bodies and take the form: `let mut a = <expr>`
- **Parameters**, which are considered immutable if they are passed as immutable references or owned. They take the syntactic form: `mut a: i32` or `b: &mut i32`
- **Function Definitions**, which are considered immutable, if all parameters considered immutable. They take the syntactic form: `fn f(mut a: i32, b: &mut i32) { ... }`
- **Arguments** are parts of a function call and can be arbitrary expression, which makes the tracking hard.

- **Function Calls**, which are considered immutable, if all arguments are considered immutable.

A total of around 52 million of these items were found in the dataset.

Figure 3.3 shows the ratio of mutable to immutable items. For each syntactical category, the percentages are relative to the total number of occurrences in the dataset.

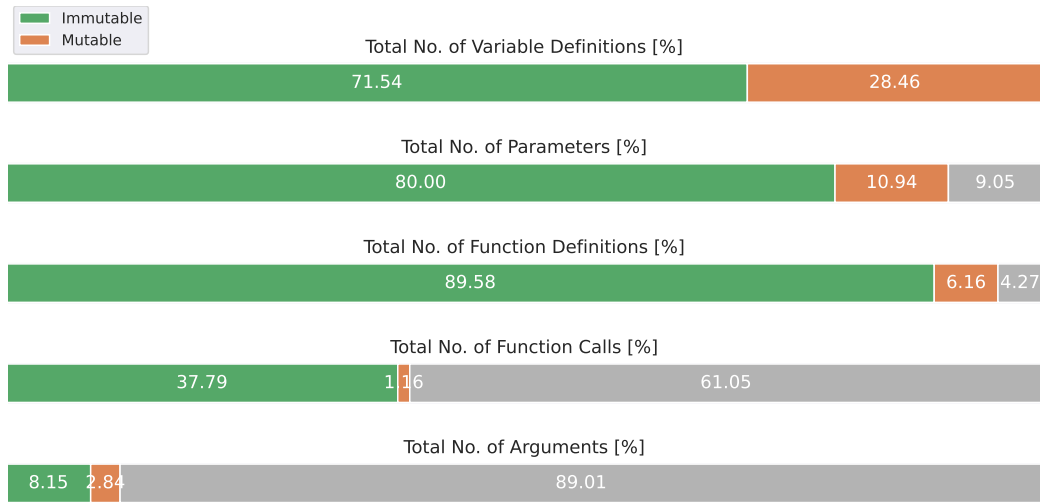


Figure 3.3: Ratio of Immutable to Mutable Items of Different AST Nodes. Ratio is Relative to Total Number Of Occurrences

Unfortunately there are some areas, where the syntactic analysis is not sufficient. Namely the analysis of function calls and arguments, which is mainly caused by the uncertainty of mutability of complex arguments. Luckily the data from function definitions and parameters can complete the picture: About 80% of parameters are immutable and between about 10 and 20% of parameters are mutable. And less than 10% of functions have mutable parameters at all. When it comes to local variables, Rust users are more liberal in their use of mutability: About 30% of local variables are defined mutable.

For verification this means, that the use of mutability is wide spread. Especially local variables are often mutable and should therefore a verification system should try to minimize the effort for the user. Mutable parameters are less common, but still need to be accounted for in verification.

3.3 Conclusions

For this thesis the following conclusion can be drawn:

- Even though uses of `unsafe` are not rare, it is acceptable to ignore in favour of a simpler type system.
- Mutable variables are used very often and should therefore have very little associated specification effort.

- Mutable parameters are used less frequently justifying a higher specification effort, but their specification must still be possible.

4 The MiniCorten Language

Rust's main disadvantage as a target language is its size: There is a lot of syntax and semantics that would need to be accounted for. A lot of it is even incidental to the verification. To reduce the complexity and amount of work that needs to be done, we will focus on a subset of Rust described in this section.

The goal is to remove as much incidental complexity as possible without compromising the central topic of research: How to extend LiquidTypes to mutability under the presence of Rust's ownership model.

4.1 Syntax

This subsection will introduce the syntax of MiniCorten, a language modelled after a simplified version of Rust, with the addition of refinement types. To simplify the formal definitions and proofs, the language is restricted to A-Normal Form (short ANF), meaning arguments of expressions must be variables. ANF was described by [20]. Note that the implementation does not have this restriction.

There are two namespaces for identifiers: $x \in \text{PVar}$ for variables of program variables and $\alpha \in \text{LVar}$ for variable in the refinement predicates. The syntax is defined by the following BNF-Grammar:

<i>program</i>	$::=$	<i>func_decl</i> *	<i>function declarations</i>
<i>func_decl</i>	$::=$	<i>ident</i> (<i>param</i> *) \rightarrow τ { <i>s</i> }	
<i>param</i>	$::=$	<i>x</i> : τ	
<i>s</i> : <i>stmt</i>	$::=$	<i>e</i>	<i>expression</i>
		<i>s</i> ₁ ; <i>s</i> ₂	<i>sequence</i>
		let mut? <i>x</i> = <i>e</i>	<i>declaration</i>
		<i>x</i> = <i>e</i>	<i>assignment</i>
		let <i>x</i> _v = <i>x</i> _f (<i>x</i> ₁ ,., <i>x</i> _n)	<i>function call</i>
		while (<i>x</i>) { <i>s</i> }	<i>while loop</i>
		if <i>e</i> _c { <i>s</i> _t } else { <i>s</i> _e }	<i>if statement</i>
		relax_ctx!{ φ * ; (<i>x</i> : τ)* }	<i>context relaxation</i>
<i>e</i> : <i>expression</i>	$::=$	<i>x</i>	<i>variable reference</i>
		<i>lit</i>	<i>constant</i>
		<i>e</i> \odot <i>e</i>	<i>binary operation</i>
		* <i>x</i>	<i>dereference</i>
		& <i>x</i>	<i>immutable reference</i>
		&mut <i>x</i>	<i>mutable reference</i>
		<i>e</i> as τ	<i>type relaxation</i>
τ : <i>type</i>	$::=$	ty!{ α : <i>b</i> φ }	<i>refinement type</i>
φ : <i>pred</i>	$::=$	<i>ref_pred</i>	<i>predicate for a reference type</i>
		<i>value_pred</i>	<i>predicate for a value type</i>
<i>ref_pred</i>	$::=$	<i>logic_ident</i> = & <i>ident</i>	
		<i>ref_pred</i> <i>ref_pred</i>	<i>mutable reference</i>
<i>value_pred</i>	$::=$	α	<i>variable</i>
		<i>v</i>	<i>literal</i>
		<i>value_pred</i> \odot <i>value_pred</i>	<i>binary op</i>
		! <i>value_pred</i>	<i>negation</i>
<i>b</i> : <i>base_ty</i>	$::=$	i32	<i>integer</i>
		unit	<i>unit type</i>
		bool	<i>boolean</i>
		& <i>b</i>	<i>immutable reference</i>
		&mut <i>b</i>	<i>mutable reference</i>
<i>v</i> : <i>lit</i>	$::=$	0,1,...,n	<i>integer</i>
		true	<i>boolean true</i>
		false	<i>boolean false</i>
		()	<i>unit value</i>
\odot : <i>bin_op</i>	$::=$	\wedge	
		\vee	
		\geq	
		+	
		=	

Most constructs are standard, the main difference being the addition of refinement types, that consist of a logic variable α , a base type b (from the target language) and a

predicate φ . Intuitively this means, that the value inhabiting that type satisfied the φ in α represents the value.

4.2 Semantics

The semantics are mostly standard. The main difference being that Rust and our simplified language enable most statements to be used in place of an expression. The value is determined by the last statement in the sequence. For example If-Expressions, sequences of statements and function all follow this rule: Their (return) value is determined by the last statement or expression in the sequence.

The semantics loosely based on Jung's MiniRust [7].

Another difference between Rust and MiniCorten is of course the addition of refinement types.

In terms of the formal description, the rules are similar to Pierce's [16, p. 166f] "Reference" language. The main difference is, that in Rust, every piece of data has a unique, known owner. This fact makes the concept of locations redundant. Instead we treat $\text{ref}(x)$ as a value itself. The following definitions show the new execution rules.

Definition 4.2.1 (Execution-State). The execution state is a partial function from program variables to values: $\sigma : \text{PVar} \rightarrow \text{Value}$.

Definition 4.2.2 (Evaluation of Expressions: $\llbracket e \rrbracket \sigma$).

$$\begin{aligned}
 \llbracket v \rrbracket \sigma &= v \\
 \llbracket x \rrbracket \sigma &= \sigma(x) \\
 \llbracket e_1 + e_2 \rrbracket \sigma &= \llbracket e_1 \rrbracket \sigma + \llbracket e_2 \rrbracket \sigma \\
 \llbracket *x \rrbracket \sigma &= \sigma(y) \quad \text{if } \sigma(x) = \text{ref}(y) \\
 \llbracket \&\text{mut } x \rrbracket \sigma &= \text{ref}(x) \\
 \llbracket e \text{ as } \tau \rrbracket \sigma &= \llbracket e \rrbracket \sigma
 \end{aligned}$$

Definition 4.2.3 (Declaration Environment). The environment of function declarations is constant and globally known. It is defined as a partial function $\Sigma : \text{Fn-Name} \rightarrow (n, s)$ where n is the number of arguments and s is the body of the function

Definition 4.2.4 (Small-Step Semantics of MiniCorten: $\langle e \mid \sigma \rangle \rightsquigarrow \langle e' \mid \sigma' \rangle$).

$$\begin{array}{c}
 \text{SS-ASSIGN} \frac{\star}{\langle x = e \mid \sigma \rangle \rightsquigarrow \langle \text{unit} \mid \sigma[x \mapsto \llbracket e \rrbracket \sigma] \rangle} \\
 \\
 \text{SS-DECL} \frac{\star}{\langle \text{let } x = e \mid \sigma \rangle \rightsquigarrow \langle \text{unit} \mid \sigma[x \mapsto \llbracket e \rrbracket \sigma] \rangle} \\
 \\
 \text{SS-SEQ-INNER} \frac{\langle e_1 \mid \sigma \rangle \rightsquigarrow \langle e'_1 \mid \sigma' \rangle}{\langle e_1; e_2 \mid \sigma \rangle \rightsquigarrow \langle e'_1; e_2 \mid \sigma' \rangle} \quad \text{SS-SEQ-N} \frac{\star}{\langle \text{unit}; e_2 \mid \sigma \rangle \rightsquigarrow \langle e_2 \mid \sigma' \rangle} \\
 \\
 \text{SS-IF-T} \frac{\llbracket x \rrbracket \sigma = \text{true}}{\langle \text{if } x \{s_t\} \text{ else } \{s_e\} \mid \sigma \rangle \rightsquigarrow \langle s_t \mid \sigma \rangle} \quad \text{SS-IF-F} \frac{\llbracket x \rrbracket \sigma = \text{false}}{\langle \text{if } x \{s_t\} \text{ else } \{s_e\} \mid \sigma \rangle \rightsquigarrow \langle s_e \mid \sigma \rangle} \\
 \\
 \text{SS-WHILE} \frac{\star}{\langle \text{while } e_c \{s_b\} \mid \sigma \rangle \rightsquigarrow \langle \text{if } e_c \{s_b; \text{while } e_c \{s_b\}\} \text{ else } \{\text{unit}\} \mid \sigma \rangle} \\
 \\
 \text{SS-CALL} \frac{\langle \Sigma(f)[arg_1 \triangleright x_1, \dots, arg_n \triangleright x_n] \mid \sigma \rangle \rightsquigarrow \langle v_r \mid \sigma' \rangle}{\langle f(x_1, \dots, x_n) \mid \sigma \rangle \rightsquigarrow \langle v_r \mid \sigma' \oplus \sigma \rangle}
 \end{array}$$

where

$$(\sigma' \oplus \sigma)(x) = \begin{cases} \sigma(x) & \text{if } x = arg_i \\ \sigma'(x) & \text{if } x \in \text{dom}(\sigma) \cap \text{dom}(\sigma') \\ \sigma(x) & \text{otherwise} \end{cases}$$

The SS-CALL rule does not scope the variables from the callee to the caller. This is actually fine, because by definition variable names are distinct. Therefore only the arguments need to be replaced in the resulting state.

5 The Refinement Type System

5.1 Features

This subsection will explain some of the key features of the Corten type system. Corten is directly embedded in Rust using two macros. Firstly the macro `ty!` can be used in place of a Rust type and adds a predicate to the Rust type, that any inhabitant of that type must satisfy. For example, the type `ty!{ v : i32 | v >= 0 }` stipulates, that a value of Rust type `i32` is positive. The second macro is `relax_ctx!{ ... }` which will be explained in subsection 5.1.6.

5.1.1 Decidable, Conservative Subtyping

Having a decidable subtyping system is essential to making it feasible in practice: A developer expects type checking to be automatic and part of compilation without additional effort. There is no need for user interaction, if the SMT solver can automatically verify the program.

5.1.2 Mutable Values

As seen in section 3.1, mutable declarations are quite common in Rust. Therefore handling them is essential for Corten. To explain the problems that mutable values can cause, consider the problem seen in ??: Suppose the value given to `i` is known to be positive, which is expressed by the type $\{v_1 : i32 \mid v_1 > 0\}$, then assigning a new value to `i` may change the type. Notably, the new value's type might depend on the old value, which is the case here: `i + 1` has type $\{v_2 : i32 \mid v_2 = v_1 + 1\}$. Corten elected to treat predicates as immutable, meaning assigning to a variable will not change the old predicates, that was associated with the variable. Therefore assigning a new value to a variable will give that variable a new type, but crucially keep the logic variable in the typing context, which means that types of other variables can still refer to it. Logic variables, that are no longer associated with a program variable (like v_1), are called unassociated variables. That means, that the number of predicates in the context increases with every new value assignment. In the example, the typing context initially contains the predicate $v_1 > 0$ and the association of $\{i \mapsto v_1\}$. After the assignment, the predicate $v_2 = v_1 + 1$ is *added* to the context and the association is *changed* to $\{i \mapsto v_2\}$.

There are other approaches for handling changing values in refinement types, which will be described and compared in 8.0.1.

```

fn decr() -> ty!{ v: i32 | v >= 0 } {
  let mut i = ...;    // i : {v1 : i32 | v1 > 0}
  i = i - 1;          // i : {v2 : i32 | v2 == v1 - 1}
  i
}

```

Listing 5: Example demonstrating why predicates and mutable values may cause problems

5.1.3 Mutable References

Besides the mutation of values, mutable references are also common in Rust, which is underpinned by the analysis, which found more than 10% of parameters to be mutable.

Corten has two ways of dealing with assignment to mutable references: If the reference destination is known, the destination’s type will be updated with the assigned values type (i.e. strong update). If there are multiple possible reference destinations, Corten will require the assigned value to satisfy the predicates of all possible destinations (i.e. weak updates). This is a standard approach (e.g. [9]), but more precise in Rust: The set of possible destinations only grows, when it depends on the execution of an optional control flow path. This means most of the time, strong updates are possible and weak updates are only needed if the reference destination is actually dynamic (i.e. dependent on the execution)¹

The problem with mutable references is possibility of aliasing. Aliasing is problematic, because it might create interdependencies between the types of different variables: If one variable is changed, it might affect whether another variable is typed correctly. In listing 6 the return value `a` is affected by changes done to a different variable `a`. Conservative approximation requires, that all possible effects must be tracked.

The key idea is, that reference types constrain the reference target, while the value owner’s type constrains the value. Because of the ownership system, Corten can accurately track references, which are phrased as predicates. In the example, `b = &mut a` and `&mut a` will have inferred type $\{ r : \&\text{mut} \mid r == \& a \}$, meaning any inhabitant of this type can at most refer to the variable called `y`. When a new value is assigned to `*b`, the type system will look at the typing context to find out what `b` might refer to. In this case `a` is the *only* possible target and we can therefore *update* its type. I.e. in type system, `*b = 0` will change the type of `a` to $\{ s : i32 \mid s == 0 \}$, but the type of `b` stays the same (because it still refers to the same location). We call this kind of assignment a strong assignment, as it can change the type.

This is sensible, because in Rust’s ownership system, `a` must be the unique owner of the memory belonging to `a`, meaning no other value predicate can be affected by the change.

Note, that the type can only be changed because we know exactly what `b` refers to. If there is ambiguity about the destination of a reference, strong updates are no longer possible.

¹It would also be possible to encode the path condition in the reference type, but this was decided against for the stated goal of simplicity for the user

```

fn client() -> ty!{ v : i32 | v == 4 } {
  let a = 2;          // a : {v1 : i32 | v1 == 2}
  let b = &mut a;     // b : {v2 : &i32 | v2 == &a}
  *b = 0; // changes a's value and type
  let c = &mut b;     // c : {v3 : &i32 | v3 == &b}
  **c = 4; // changes a's value and type
  a
}

```

Listing 6: Example demonstrating interdependencies between mutable references

To support these use cases, Corten also supports weak updates, which can not change types, but allow assigning to ambiguous reference destinations. Listing 7 demonstrates how ambiguity about the reference destination may emerge: Depending on the if condition, `res` could refer to either `y` or `z`. Naturally, we can weaken the reference type to `ty!{ r1 : &mut i32 | r1 == &y || r1 == &z }`, meaning `res` could refer to either `y` or `z`. Because the destination is ambiguous, assigning to `*res` can not change the type of `y` or `z`, which is the case if the type of the assigned value is at least as specific as the types of all possible reference destinations.

```

fn weak_updates(x : ty!{ x1 : i32 }) -> ty!{ v : i32 | v > 2 * x1 + 10 } {
  let mut y = x as ty!{ y1 : i32 | y1 >= x1 };
  let mut z = x + 10 as ty!{ z1 : i32 | z1 >= x1 + 10 };
  let mut res;
  if x > 0 {
    res = &mut y as ty!{ r : &mut | r == &y || r == &z };
    // branches of if need same type => weaken
  } else {
    res = &mut z as ty!{ r : &mut | r == &y || r == &z };
  }
  *res = x + 11; // res could refer to b or c
                // -> assigned value must satisfy both types
  y + z
}

```

Listing 7: Example demonstrating weak updates

An intricacy about conservative reference tracking are functions, that return references. In Rust this is possible, if the reference was passed as an argument: The function signature `fn f(a : &mut T, b : &mut T) -> &mut T` would allow `f` to return either a reference to `a` or `b`. For the callee knowing where the returned value points to, is important because the type of dereferencing it depends on that knowledge. Consequently, for the return type to be conservative, it must state every possible reference destination.

5.1.4 Path Sensitivity

Firstly, the type system is path sensitive, meaning that the type system is aware of necessary XXX that need to be passed for an expression to be evaluated. For example listing

8 shows a function computing the maximum of its inputs. In the then branch, a is only a maximum of a and b , because the condition $a > b$ implies it. Corten will symbolically evaluate the condition and store it in its typing context.

```
fn max(a : ty!{ av: i32 }, b: ty!{ bv: i32 })
  -> ty!{ v: i32 | v >= av && v >= bv } {
  if a > b {
    a as ty!{ x : i32 | x >= av && x >= bv }
  } else {
    b
  }
}
```

Listing 8: Function computing the maximum of its inputs; guaranteeing that the returned value is larger than its inputs

- inference - also with mutations

5.1.5 Strong Type Updates

```
fn strong_updates(b: ty!{ bv: i32 | bv > 0 }) -> ty!{ v: i32 | v > 2 } {
  let mut a = 2;
  a = a + b;
  a
}
```

Listing 9: Example of changes to a 's value affecting its type

5.1.6 Modularity

For a verification system to be scalable, it needs to be able modularize a proof. Corten can propagate type information across function calls and by taking advantage of Rust's ownership system, it can do so very accurately. Listing 10 shows, how an incrementing function `inc` can be specified: `inc` signature stipulates, that it can be called with any `i32` (given the name `a1`) and will return a value `a2`, which equals $a1 + 1$. Only the signature of `inc` is necessary for the type checking of `client`.

Notice, that all of the type information about y is preserved when `inc(x)` is called. There are no further annotations needed to type check this program. This is possible, because in safe Rust, any (externally observable) mutation done by a function must be part of the function signature. Corten expands on this, by enabling the user specify exactly how a referenced value is mutated.

5.1.7 Atomic Updates

Complex mutation patterns can result in complex interdependencies. We deem it necessary to allow different types to refer to each other. The properties of arguments might

```

fn inc(a: &mut ty!{ a1: i32 => a2 | a2 == a1 + 1 }) {
    *a = *a + 1;
}
fn client(mut x: ty!{ xv: i32 | xv > 2 }) -> ty!{ v: i32 | v > 7 } {
    let mut y = 2;
    inc(&mut x); inc(&mut x);
    inc(&mut y)
    x + y
}

```

Listing 10: Example showing how Corten allows for accurate type checking in the presence of function calls

depend on each other, consequently their refinement types should be able to refer to each other.

The listing 11 is extract from the evaluation example `??`. It requires establishing an invariant, where `i` has the predicate ≥ 0 and `sum` has the predicate $= i \cdot n$.

Attempting to update one type at a time, will not yield the desired result. If `i` is updated first, then it is not possible to establish `sum`'s type, because `i`'s type is no longer expressive enough. If `sum` is updated first, we can establish its type, but once `i` is updated, it will refer to a different logic variable for `i` than `sum`. If we tried to carry the information, that the two logic variables are the same in the type, it would no longer be an invariant, because reassigning the value would invalidate it².

```

let sum = 0;
let i = 0;

relax_ctx!{
    i: ty!{ iv : i32 | iv > 0 },
    sum: ty!{ sv: i32 | 2 * sv == iv * (iv + 1) }
}

i = i + 1;
sum = sum + i;

```

Listing 11: Demonstration of interdependent types

For this purpose, Corten provides a mechanism for the programmer to instruct Corten to relax all variables at once named `relax_ctx`. The syntax is similar to the refinement type syntax, but allows for multiple variables to be described at a time. The type system will check all refinement types at once, resulting in the successful type checking of the example.

²For this use-case, it might also be possible to relax a single variable at a time, by renaming the logic variables to their old names after the subtyping relation is shown. This approach was not chosen, because its influence on correctness are hard to assess.

5.2 Function Declaration Type Checking

The following sections will introduce the type checking rules of Corten, starting with the rule for checking function declarations. We assume the program has passed the Rust type checking rules, meaning (non-refined) type- and ownership-checking was successful.

Definition 5.2.1 (Typing Context $\Gamma = (\mu, \Phi)$). is the typing context, consisting of a partial function $\mu : \text{PVar} \rightarrow \text{LVar}$ used for tracking the current logic variable associated with a program variables and a set of predicates Φ , which constrains the logic variables in the image of μ .

Without loss of generality, we assume arguments are ordered by their type: First immutable / owned arguments and then mutable arguments.

When handling mutable references in parameters some subtleties need to be considered. The function can change both the referenced *value* as well as the reference *location*. To describe the referenced value is normally done using the $\{\alpha \mid \alpha \doteq \&b\}$ syntax. The question arises: If α was the logic variable for a parameter, what should be the analog for b ? In contrast to local variables, there is no variable representing the referenced value for parameters. As seen in section 5.1.2, using the dereference operator would come with a lot of complications. Instead we introduce arg_n^i , a special variable denoting the initial abstract value (i.e. stack location), that the mutable reference of argument n points to. i denotes the level of nesting: For the n th parameter with the Rust type `&mut &mut i32` we would generate arg_n^1 and arg_n^2 .

$$\text{FN-DECL} \frac{(\{a_1 \mapsto \alpha_1, b_1 \mapsto \delta_1\}, \varphi_1^\alpha \wedge \dots \wedge \delta_1 = arg_1^1 \wedge \varphi_1^\beta \wedge \dots) \vdash \bar{s} \Rightarrow \Gamma' \quad \Gamma' \vdash s_{res} : \tau_{res} \Rightarrow \Gamma'' \quad \Gamma'' \vdash \tau_{res} \leq \tau \quad \Gamma'' \leq (\{\beta_1 \mapsto \gamma_1\}, \varphi_1^\gamma)}{fn \ f(a_1 : \{\alpha_1 \mid \varphi_1^\alpha\}, \dots, b_1 : \{\beta_1 \mid \varphi_1^\beta \Rightarrow \gamma_1 \mid \varphi_1^\gamma\}) \rightarrow \tau\{\bar{s}; s_{res}\}}$$

5.3 Sub-Typing Rule: $\Gamma \vdash \tau \leq \tau'$

Next, we introduce the Rule for checking the subtype relation: a type τ is a subtype of τ' in the context Γ , if the predicate of the supertype imply the predicate of the subtype. Corten uses a decidable, propositional logic with the theories of equality and linear arithmetic. Before dispatching a request to the SMT solver, the supertype's logic variable is substituted for the subtype's variable to the predicate to refer to the same variable.

$$\leq\text{-Ty} \frac{\text{SMT-Valid}(\Phi \wedge \varphi'[\beta \triangleright \alpha] \implies \varphi)}{\Gamma = (\mu, \Phi) \vdash \{\alpha \mid \varphi\} \leq \{\beta \mid \varphi'\}}$$

Note that, $\Gamma \vdash \tau \leq \tau'$ can be rephrased in terms of $\Gamma \leq \Gamma'$ by introducing a fresh variable with the predicates τ and τ' . The rephrased form is equivalent, but rephrasing the sub context rule in terms of sub-typing is in general not possible: Because all constraints need to be satisfied at the same time when checking for sub-contexts, just checking each variable at a time would be a weaker proposition.

5.4 Sub-Context Rules: $\Gamma \leq \Gamma'$

In contrast to other refinement type systems, Corten allows types in the context to refer to one another. For example a type specifications $a : \{ \alpha : \text{!32} \mid \alpha > \beta \}$, $b : \{ \beta : \text{!32} \mid \beta \neq \alpha \}$ would be valid and result in the context $\Gamma = (\emptyset, \alpha \neq \beta \wedge \beta \geq \alpha)$. In the example, the value 0 is a valid inhabitant of a 's type as long as $b \geq 1$.

The rule $\leq\text{-CTX}$ serves the purpose of allowing these interdependencies between predicates while still providing an applicable concept of subtyping: As long as the set of predicates in Γ' imply the predicates for Γ , the sub-context relation is satisfied and Γ' (Γ) is called the super-context (sub-context) respectively. To be more precise: The rule needs to relate the predicates for each *program* variable with the opposing predicates for that program variable. The substitution $\Phi'[\mu'(\alpha) \triangleright \mu(\alpha) \mid \alpha \in \text{dom}(\mu)] \implies \Phi$ forces the logic variable associated with x in Γ' to be substituted by the logic variable $\mu(x)$, which is associated with x in Γ .

The domain restriction ensures, that the sub-context does not contain additional program variables, which would leave unconstrained variables in the antecedent

$$\leq\text{-CTX} \frac{\text{SMT-Valid}(\Phi'[\mu'(\alpha) \triangleright \mu(\alpha) \mid \alpha \in \text{dom}(\mu)] \implies \Phi) \quad \text{dom}(\mu) \subseteq \text{dom}(\mu')}{(\Phi, \mu) \leq (\Phi', \mu')}$$

5.5 Expressions Typing: $\Gamma \vdash e : \tau$

$$\begin{array}{c} \text{LIT} \frac{l \text{ fresh} \quad \text{base_ty}(v) = b}{\Gamma \vdash v : \{l : b \mid l \doteq v\}} \quad \text{VAR} \frac{\alpha \text{ fresh} \quad \mu(x) = \beta}{(\mu, \Phi) \vdash x : \{\alpha : b \mid \beta \doteq \alpha\}} \\ \text{VAR-DEREF} \frac{\Gamma \vdash y : \tau \quad \Gamma \vdash x : \{\beta : \&b \mid \beta \doteq \&y\}}{\Gamma \vdash *x : \tau} \\ \text{ADD} \frac{\gamma \text{ fresh} \quad \mu(x_1) = \alpha \quad \mu(x_2) = \beta}{\Gamma \vdash x_1 \cdot x_2 : \{\gamma : b \mid \gamma \doteq \alpha \cdot \beta\}} \\ \text{REF} \frac{\star}{\Gamma \vdash \&x : \{\alpha : \&b \mid \alpha \doteq \&x\}} \quad \text{INTRO-SUB} \frac{\Gamma \vdash e : \tau \quad \Gamma \vdash \tau \leq \tau'}{\Gamma \vdash e \text{ as } \tau' : \tau'} \end{array}$$

5.6 Statement Type Checking $\Gamma \vdash s \Rightarrow \Gamma'$

Abbreviations We write:

- Γ, c for $(\mu, \Phi \wedge c)$
- $\Gamma[a \mapsto \alpha]$ for $(\mu[a \mapsto \alpha], \Phi)$
- $\Gamma(x)$ for $\mu(x)$

$$\begin{array}{c}
 \text{IF} \frac{\Gamma \vdash \Gamma(x) : \text{bool} \Rightarrow \Gamma \quad \Gamma, \Gamma(x) \doteq \text{true} \vdash s_t \Rightarrow \Gamma' \quad \Gamma, \Gamma(x) \doteq \text{false} \vdash s_e \Rightarrow \Gamma'}{\Gamma \vdash \text{if } x \text{ then } s_t \text{ else } s_e \Rightarrow \Gamma'} \\
 \\
 \text{WHILE} \frac{\Gamma_I, \Gamma_I(x) \doteq \text{true} \vdash s \Rightarrow \Gamma'_I \quad \Gamma'_I \leq \Gamma_I}{\Gamma_I \vdash \text{while } x\{s\} \Rightarrow \Gamma_I, \Gamma_I(x) \doteq \text{false}} \\
 \\
 \text{SEQ} \frac{\Gamma \vdash s_1 \Rightarrow \Gamma' \quad \Gamma' \vdash s_2 \Rightarrow \Gamma''}{\Gamma \vdash s_1; s_2 \Rightarrow \Gamma''} \\
 \\
 \text{DECL} \frac{\Gamma \vdash e : \{\beta : b \mid \varphi\}}{\Gamma \vdash \text{let } x = e \Rightarrow \Gamma[x \mapsto \beta], \varphi} \quad \text{ASSIGN} \frac{\Gamma \vdash e : \{\beta : b \mid \varphi\}}{\Gamma \vdash x = e \Rightarrow \Gamma[x \mapsto \beta], \varphi} \\
 \\
 \text{ASSIGN-STRONG} \frac{\Gamma \vdash e : \{\beta \mid \varphi\} \quad \Gamma \vdash x : \{\alpha : \&b \mid \alpha \doteq \&y\}}{\Gamma \vdash *x = e \Rightarrow \Gamma[y \mapsto \beta], \varphi} \\
 \\
 \text{ASSIGN-WEAK} \frac{\begin{array}{c} \Gamma \vdash e : \tau \quad \Gamma \vdash x : \{\alpha : \&b \mid \alpha \doteq \&y_1 \vee \dots \vee \alpha \doteq \&y_n\} \\ \Gamma \vdash \tau \leq \{\beta_1 \mid \varphi_1\} \quad \dots \quad \Gamma \vdash \tau \leq \{\beta_n \mid \varphi_n\} \\ \Gamma[x \mapsto \beta_1], \varphi_1 \leq \Gamma \quad \dots \quad \Gamma[x \mapsto \beta_n], \varphi_n \leq \Gamma \end{array}}{\Gamma \vdash *x = e \Rightarrow \Gamma} \\
 \\
 \text{FN-CALL} \frac{\begin{array}{c} \text{subst} = \dots \quad (\mu[\text{subst}], \alpha \doteq \mu(a) \wedge \dots \wedge \varphi_\alpha \wedge \dots) \leq (\mu, \Phi) \\ f : (\{\alpha \mid \varphi_\alpha\} \Rightarrow \{\alpha' \mid \varphi'_\alpha\}, \dots) \rightarrow \tau \end{array}}{(\mu, \Phi) \vdash \text{let } \text{res} = f(a, \dots, \&\text{mut}b, \dots) \Rightarrow (\mu[a \mapsto \alpha', \dots, \text{subst}^{-1}], \Phi \wedge \varphi'_\alpha \wedge \dots)}
 \end{array}$$

The $\Gamma \vdash \alpha$ must reference y rule stipulates, that the context Γ requires α to refer to exactly y . It is necessary to follow basic indirections from equality constraints, that other rules (like FN-CALL or VAR) introduce. Because of the how limited the subset of the predicate language used for references is, even a very basic, conservative approximation is adequate for this purpose: A fixpoint construction of an equivalence class containing all logic variables equal to α is incrementally constructed. If any of these variables reference y , then α must reference y . Of course a more sophisticated approach is also possible: Dispatching a request to SMT, that asserts, that α references some (unconstrained) y and that there is second y' with this a property. Retrieving the value for y from SMT is the result of the rule application.

$$\text{MUST-REF} \frac{\begin{array}{c} \exists y. \forall \alpha, \beta, \dots. \Phi \rightarrow \alpha \doteq \&y \\ \models \Phi \rightarrow \alpha \doteq \&y \quad \models \Phi \wedge \alpha \doteq \&y \wedge \alpha \doteq \&y' \rightarrow y \doteq y' \end{array}}{(\mu, \Phi) \vdash \alpha \text{ must reference } y}$$

In think the first precondition and second precondition should serve the same purpose. Second one should be nicer and correct

5.7 Soundness of the Type System

As usual, we consider the three properties (see Pierce [17, p. 95, p.167]) of a type system when assessing the correctness of MiniCorten:

Definition 5.7.1 (State Conformance). A state σ is conformant with respect to a typing context $\Gamma = (\mu, \varphi)$ (written as $\sigma : \Gamma$), iff:

$$\models \left(\bigwedge_{x \in \text{Var}} \mu(x) \doteq \sigma(x) \right) \rightarrow \varphi$$

I.e. the execution's value assignments imply the type refinements

Definition 5.7.2. Progress: If t is closed and well-typed, then t is a value or $t \rightsquigarrow t'$, where t' is a value.

Definition 5.7.3 (Preservation). If $\Gamma \vdash e : \tau \Rightarrow \Gamma$ and $e \rightsquigarrow e'$, then $\Gamma \vdash e' : \tau \Rightarrow \Gamma$

The state conformance rule differs from the usual rule in two ways. Firstly we consider the runtime value instead of the runtime type. Secondly there needs to be one set of variable assignments that satisfies all predicates in Γ .

Because MiniCorten is based on Rust, we will assume, that progress, as well as preservation for the base-types. This includes preservation of type- and ownership-safety. Thus it is sufficient to prove, that assuming these properties hold, preservation of the refinement types is holds.

Lemma 5.7.4 (Conformance of Evaluation). If $\sigma : \Gamma$ and $\llbracket e \rrbracket \sigma = v$ then $\models \varphi \rightarrow \llbracket e \rrbracket \Gamma$

Proof. □

Lemma 5.7.5 (Preservation of State Conformance). If $\Gamma \vdash s \Rightarrow \Gamma_2$, $\sigma : \Gamma$ and $\langle s \mid \sigma \rangle \rightsquigarrow \langle s_1 \mid \sigma_1 \rangle$, then $\sigma_1 : \Gamma_1$ and $\Gamma_1 \vdash s_1 \Rightarrow \Gamma_2$ for some Γ_1

Proof. Rule Induction over $\langle t \mid \sigma \rangle \rightsquigarrow \langle t' \mid \sigma' \rangle$

- CASE SS-ASSIGN: Given $\sigma :: \Gamma, \Gamma \vdash x = e \Rightarrow \Gamma_2$ show $\exists \Gamma_1, \sigma[x \mapsto \llbracket e \rrbracket \sigma] : \Gamma_1 \wedge \Gamma_1 \vdash \text{unit} \Rightarrow \Gamma_2$

Rule inversion over $\Gamma \vdash x = e \Rightarrow \Gamma_2$ gives $\Gamma \vdash e : \{\alpha : b \mid \alpha \doteq \llbracket e \rrbracket\}$ and $\Gamma_2 = (\Gamma[x \mapsto \alpha], \alpha \doteq \llbracket e \rrbracket)$

For $\Gamma_1 = \Gamma_2$, $\sigma[x \mapsto \llbracket e \rrbracket \sigma] : (\Gamma[x \mapsto \alpha], \alpha \doteq \llbracket e \rrbracket)$, α is fresh and therefore monotonicity applies.

- CASE SS-SEQ-INNER: Given: $\langle c_1 \mid \sigma \rangle \rightsquigarrow \langle c'_1 \mid \sigma' \rangle$,
 $\Gamma \vdash c_1; c_2 \Rightarrow \Gamma_2$,
 $\forall \Gamma_2, \Gamma \vdash c_1 \Rightarrow \Gamma_2 \wedge \sigma : \Gamma \rightarrow \exists \Gamma_1, \sigma' : \Gamma_1 \wedge \Gamma_1 \vdash c'_1 \Rightarrow \Gamma_2$
 show $\exists \Gamma_1, \sigma' : \Gamma_1 \wedge \Gamma_1 \vdash c'_1; c_2 \Rightarrow \Gamma_2$.

Rule inversion over $\Gamma \vdash c_1; c_2 \Rightarrow \Gamma_2$ yields $\Gamma \vdash c_1 \Rightarrow \Gamma_1, \Gamma_1 \vdash c_2 \Rightarrow \Gamma_2$

The preconditions for the third induction hypothesis are satisfied for Γ_1 delivers $\exists \Gamma'_1, \sigma' : \Gamma'_1 \wedge \Gamma'_1 \vdash c'_1 \Rightarrow \Gamma_1$

State conformance $\sigma' : \Gamma'_1$ follows directly from this.

For Γ'_1 the preconditions for the SEQ rule are satisfied.

- CASE SS-SEQ-N: Given: $\Gamma \vdash \text{unit}; c \Rightarrow \Gamma_2, \sigma : \Gamma$ show $\exists \Gamma_1, \sigma : \Gamma_1 \wedge \Gamma \vdash c \Rightarrow \Gamma_2$.
Rule inversion over $\Gamma \vdash \text{unit}; c \Rightarrow \Gamma_2$ yields $\Gamma \vdash c \Rightarrow \Gamma_2$. Together with the second induction hypothesis this implies the goal.
- CASE SS-DECL: analogous to CASE SS-ASSIGN
- CASE Deref $\sigma(x) = v, e' = v, \sigma' = \sigma$
For $\Gamma' = \Gamma, \sigma : \Gamma'$ true by assumption
Show: $\Gamma \vdash v : \tau' \Rightarrow \Gamma$. By rule inversion LIT, $\tau' = \{l \mid l \doteq v\}$

□

Lemma 5.7.6. *State Conformance is monotone: If $\sigma : \Gamma, (\beta = v) \rightarrow \varphi$ then $\sigma[x \mapsto v] : \Gamma[x \mapsto \beta] \wedge \varphi$*

Proof.

$$\models \left(\bigwedge_{x \in \text{Var}} \mu(x) \doteq \sigma(x) \right) \rightarrow \varphi$$

...?

□

5.8 Extensions

To limit the complexity a few features are not defined as part of the Corten or the implementation. Even though these extensions are not part of the scope of the thesis, these extensions were taken into consideration when designing the type system. To show, that the type system is capable of handling the additional challenges, we will shortly describe how these extensions are planned to be added to the type system.

5.8.1 Records / structs

Firstly, a key part of realistic programs are data structures that comprise multiple basic data types. In Rust these are called structs and work similar to records or product types in functional languages.

Once again, we can take advantage of Rust ownership system: Any part of a struct (even nested fields) can only belong to one variable. This means, that the proposed system for handling mutable references extends seamlessly to structs: The variable mapping in Γ is generalized: $\mu : \text{Path} \rightarrow \text{LVar}$. The relevant typing rules will work without major changes:

$$\text{VAR} \frac{\beta \text{ fresh} \quad \mu(p.x.y.z) = \alpha}{\Gamma \vdash p.x.y.z : \{\beta : b \mid \beta \doteq \alpha\} \Rightarrow \Gamma}$$

$$\text{ASSIGN-STRONG} \frac{\Gamma \vdash e : \tau}{\Gamma \vdash p.x.y.z = e \Rightarrow \Gamma[p.x.z.y \mapsto \tau]}$$

5.8.2 Algebraic Data Types

With records added, the only thing missing for support of algebraic data types are sum types. Sum types allow the programmer to express, that an inhabitant of a type may be one variant of a set of multiple fixed options: For example the result of a fallible computation may either be $\text{Ok}(V)$ meaning a successful computation with the result V or a failure $\text{Err}(E)$ with a description of the error E . In rust these sum types are called enums.

Sum types influence the type system in two key ways:

Firstly, the specification of its values. Suppose a programmer would like an authentication function signature to state, that the function returns an `Err` code "403" if the password was incorrect. This requires the type language to assert what variant is expected as well as access to its fields (if the variant is known).

Secondly, path sensitivity needs to be extended to cover match expressions (called case in most functional programming languages). match expression allow the programmer to branch depending on the variant of a value.

5.8.3 Inference

While the current implementation is able to type expressions without explicit type annotations, solving a set of type constraints is not implemented. Rondon et al. [18] describe a mechanism to infer complex refined types by combining known predicates from the context. This approach should be adaptable to Corten to reduce the amount of needed type annotations even further.

5.8.4 Predicate Generics

Vazou et al. [22] found, that the expressiveness of refinement types can still be expanded without leaving a decidable fragment by adding uninterpreted functions to the logic. In the type system, these uninterpreted functions represent "abstract predicates". At the definition site, abstract predicates can not be inspected and restricted, but the caller can instantiate them with a concrete predicate.

5.8.5 Convenience Improvements

For practical use, it would make sense to relax the distinction between logic variables and program variables for immutable variables. Because these variables can not change, there is no need to introduce replace their types. This also means, that relaxing the predicate for immutable variables is never necessary.

In some cases, the programmer might want to choose an invariant for a variable, that any assigned value has to satisfy. In terms of the type system, this would not be challenging, only requiring

6 Implementation: CortenC

The type system described in the previous chapter was implemented to test the practical feasibility of our approach. There are a few differences and details between the type system described above and the implementation CortenC (short for Corten Checker) that will be highlighted in this chapter.

In contrast to the MiniCorten and the described type system, CortenC uses actual Rust as its target language. In addition to MiniCorten’s features, CortenC also covers expressions with side-effects, non-ANF expressions, statements as well as basic inference. These features do not change the expressiveness of the type system, but make it a lot more practical.

Target Selection The architecture of the rustc compiler makes it possible to implement a refinement type system quite cleanly: Rustc’s plugin system allows CortenC to access rustc’s intermediate representations and also emit diagnostics with source locations. Since the diagnostics from CortenC use the same interface as Rust’s diagnostics, other tools like IDEs can handle them without any adaptation.

CortenC uses the HIR. It is a good candidate for extensions to the type system, because it contains source labels, allowing accurate problem reporting to the user. Secondly, in HIR all names are resolved making and types can be queried for HIR nodes, which means the Rustc plugin can reuse a lot of work done in the Rustc Compiler. Thirdly, non-executable segments of the program, like type declarations, are represented in the HIR.

Using MIR is also a sensible option: The MIR is a drastically simplified representation of just the executable segments of the program as a CFG with non-nested expressions. Because of the simplicity, it would be appealing to use the MIR as the basis for the implementation. An additional advantage of MIR could be, that ownership analysis is done on MIR, which would allow the refinement type system to use that information. Surprisingly, we did not encounter any situation, where ownership information is necessary for typing refined types.

Ultimately, it was decided against using the MIR for the following reasons: Firstly a concern for the quality of diagnostics: The MIR code is quite distant from the user-provided code. For example, when trying to explain a error when typing a while-loop, the implementation may need to reconstruct the original source code structure from the CFG, which could be inaccurate and error-prone. In addition, source locations may be less accurate or unavailable. Finally the MIR does not contain type declarations, which could make it hard to extend the implementation for algebraic data type or predicate generics (see section 5.7)

Language Embedding CortenC exposes a minimal interface to the programmer: The interface consists of two macros: `ty!{ ... }` which allows the programmer to specify

a refined type for a rust base type and `relax_ctx!{ ... }` to allow the programmer to relax the typing context (mostly used for introducing a loop-invariant).

Listing 12 shows how the `ty!` macros is defined: Most important are the first two macro forms, which handle immutable types like `ty!{ v : i32 | v > 0 }` and mutable parameter types like `ty!{ v1: i32 | v1 > 0 => v2 | v2 < 0 }`. There are additional short forms of these macro calls, which were cut from the listing for brevity. Both macro calls are translated into the type alias `Refinement` (`MutRefinement` respectively). For the examples above, it would be `Refinement<i32, "v", "v > 0">` and `MutRefinement<i32, "v1", "v1 > 0", "v2", "v2 > 0">` respectively. These type aliases are quite useful for CortenC: They ensure, that CortenC keeps compatibility with Rust. I.e. a program with CortenC type annotations can be compiled in normal Rust without problems and the type aliases also simplify the implementation of CortenC, because the refined type can be extracted directly from the Rust type alias. Consequently, CortenC does not need to perform name resolution, even for external function declarations.

Note that, because the `ty!` macro takes the place of a Rust type, normal IDE features, like "Goto Type Definition" still work on the base types without any adaptation.

```
pub type Refinement<
    T,                                // Rust type
    const B: &'static str,           // Logic Variable
    const R: &'static str           // Predicate
> = T;

pub type MutRefinement<
    T,                                // Rust type
    const B1: &'static str,          // Logic Variable Before
    const R1: &'static str,          // Predicate Before
    const B2: &'static str,          // Logic Variable After
    const R2: &'static str,          // Predicate After
> = T;

#[macro_export]
macro_rules! ty {
    ($i:ident : $base_ty:ty | $pred:expr) => {
        $crate::Refinement< $base_ty, {stringify! { $i }}, {stringify! { $pred }}>
    };
    ($i:ident : $base_ty:ty | $pred:expr => $i2: ident | $pred2:expr) => {
        $crate::MutRefinement< $base_ty, {stringify! { $i }}, {stringify! { $pred }}, {stringify! { $i2 }}, {stringify! { $pred2 }}>
    };
    // -- snip --
}
```

Listing 12: Definition of CortenC's macros

For the second macro `relax_ctx!` takes the place of a statement. It is translated to a function call, that can then be used to reconstruct the described context.

Architecture CortenC follows the structure implied by the type system: All concepts and type checking rules have a correspondence in CortenC: `RContext` is the representation of Γ in CortenC, with the main difference being, that predicates stay associated

with a logic variable (and potentially a program variable). The purpose being, that in the future, CortenC could create better error messages, by associating “blame” to a type checking failure: If the predicate of a logic variable is part of the reason why a subtyping check fails, it would be useful to convey that information to the user. `RefinementType` is the representation of τ in CortenC.

CortenC registers a callback in Rustc to run after HIR construction. The callback calls `type_check_function` for each function in the HIR, which either returns `Result::Ok` indicating, that the function type checks or `Result::Err` if type checking failed. In that case, the callback will emit a diagnostic to rustc to inform the user about the refinement type error. Listing 13 shows the core functions signatures involved in type checking. `type_check_function` corresponds to the FN-DECL rule by constructing the initial context and delegating checking of the body to `type_expr`. Since Rust expressions can contain statements, `type_expr` will use `transition_stmt` to type check these statements before returning the final type context. Both `type_expr` and `transition_stmt` carry a collection of parameters, that contain the Rust type checking data (The global type context `TyCtxt` and the local type context `TypeckResults`), the refinement type context `RContext`, a handle to the SMT solver `SmtSolver` and a way to generate fresh logic variable names `Fresh`.

These function utilize `require_is_sub_context` and `require_is_subtype_of` to dispatch the actual SMT solver requests. These functions are also responsible for encoding the predicates and substituting variable names where necessary. The following paragraph will give an example for how these SMT solver requests look like.

Simple Example Consider the example program 14, which returns the increment of the argument. Rustc will call the CortenC callback with the item `fn inc` once the HIR construction is completed. Next, the callback calls `type_check_function` where the initial context $\Gamma = (a \mapsto n, \{n > 0\})$ is constructed and calls `type_expr` with the the function body expression. Base the function body is a sequence `transition_stmt` is called for all expect the last expression in the sequence. `let b = 1` updates the context to $\Gamma_1 = (a \mapsto n, b \mapsto v_1, \{n > 0, v_1 = 1\})$ (v_1 is a name generated by `Fresh`). Finally `type_expr` types the expression `a + b` according to the typing rules described above resulting in the type $\{v_2 : i32 \mid v_2 = n + v_1\}$. This type is then returned as the overall type of the body expression to `type_check_function`, which has the responsibility of checking that in the context Γ_1 , the actual return type is a subtype of the specified type in the signature. Thusly `require_is_subtype_of` is called with the two types and subsequently dispatches the SMT shown in listing ???. Because the SMT call returned `unsat`, the subtype relation is valid and the function type checking was successful.

```
fn type_check_function(
  function: &hir::Item,
  tcx: &TyCtxt,
) -> Result<RefinementType> { .. }

fn transition_stmt(
  stmts: &a [hir::Stmt],
  tcx: & TyCtxt,
  ctx: & RContext,
  local_ctx: & TypeckResults,
  solver: &mut SmtSolver,
  fresh: &mut Fresh,
) -> Result<RContext> { .. }

fn type_expr(
  expr: &a Expr,
  tcx: & TyCtxt,
  ctx: & RContext,
  local_ctx: & TypeckResults,
  solver: &mut SmtSolver,
  fresh: &mut Fresh,
) -> Result<(RefinementType, RContext)> { .. }

fn require_is_sub_context(
  super_ctx: &RContext,
  sub_ctx: &RContext,
  tcx: &TyCtxt,
  solver: &mut SmtSolver,
) -> anyhow::Result<()> { .. }

fn require_is_subtype_of(
  sub_ty: &RefinementType,
  super_ty: &RefinementType,
  ctx: &RContext,
  tcx: &TyCtxt,
  solver: &mut SmtSolver,
) -> anyhow::Result<()> { .. }
```

Listing 13: Overview of the central function CortenC is built from. (Lifetimes were removed for clarity)

```
fn inc(a : ty!{ n : i32 | n > 0 }) -> ty!{v : i32 | v > 1} {
  let b = 1;
  a + b
}
```

Listing 14: Simple Example Program used for demonstrating the operation of CortenC

```

(declare-datatypes () ((Unit unit)))
; <Context>
; decl for <fud.rs>:79:73: 79:74 (#0) local b
(declare-const _0 Int)

; decl for <fud.rs>:79:9: 79:10 (#0) local a
(declare-const n Int)

; predicate for _0: local b      ty!{ _0 : i32 | _0 == 1 }
(assert (= |_0| 1))

; predicate for n: local a      ty!{ n : i32 | n > 0 }
(assert (> |n| 0))
; </Context>

(declare-const _1 Int)
(assert (= (+ |n| |_0|) |_1|))

(assert (not (> |_1| 1)))

; checking: ty!{ _1 : i32 | n + _0 == _1 } <= ty!{ v : i32 | v > 1 }
(check-sat)
; done checking is_subtype_of! is sat: false

```

Listing 15: SMT Requests dispatched by CortenC for checking that the returned type matches the specified type

7 Evaluation

In this chapter Corten will be tested on a selection of examples to test the practicality.

7.1 Maximum using Path Conditions

The first example will show how a the max function can be implemented and fully verified in CortenC. The mathematical maximum is defined as $r = \max(a, b)$ iff $r \geq a \wedge r \geq b \wedge (r = a \vee r = b)$. Listing 16 shows how this can be expressed in CortenC.

```
fn max(a: ty!{ a: i32 }, b: ty!{ b: i32 })
-> ty!{ v: i32 | v >= a && v >= b && (v == a || v == b) } {
  if a > b {
    a as ty!{ x: i32 | x >= a && x >= b && (x == a || x == b) }
  } else {
    b
  }
}
```

Listing 16: Example demonstrating a fully specified max function using Corten’s path sensitivity

Of course an incorrect implementation will produce a type judgement error. If, for example, the else branch returned a instead of b, the system would return the error shown in listing 17. The exact location of where the error occurred is also the current implementation does not expose this data yet.

```
Subtyping judgement failed:
ty!{ _1 : i32 | true && _1 == a } is not a sub_ty of
ty!{ v : i32 | v >= a && v >= b && (v == a || v == b) }

in ctx RContext {
  // formulas
  // types
  local a : ty!{ a : i32 | true }
  local b : ty!{ b : i32 | true }
}
For example this assignment, satisfies the sub type, but not the super type:
{ _1 = 0, b = 1, a = 0 }
```

Listing 17: Example of an error message created by CortenC

Path sensitivity is not limited to the return value: Effect on the type context can also be path sensitive as demonstrated by listing 18, which implements a clamping function.

The function `clamp` ensures, that the reference passed to it will be at most `max` or stay the same. `client` uses `clamp` and can use the facts from `clamp`'s mutation specification when proving its own return type specification.

Note, that when typing `client`, CortenC needs to match `&mut x` with `a` and thusly `x` with `s` to correctly handle the effects of calling `clamp`. Because of this reason, function calls require all arguments to be variables or reference to variables. This is not a restriction of the type system, but a simplification for the implementation.

```
fn clamp(
  a: &mut ty!{ a1: i32 | true => s | (s <= max) && (s == a1 || s == max) },
  max: ty!{ max: i32 }
) -> ty!{ v: () } {
  if *a > max {
    *a = max as ty!{ r | (r <= max) && (r == a1 || r == max) }; ()
  } else {};
  ()
}

fn client() -> ty!{ v : i32 | v == 42 } {
  let mut x = 1337; let max = 42;
  clamp(&mut x, max);
  x
}
```

Listing 18: Example demonstrating optional mutation of an external location

7.2 Recursion: Fibonacci-Numbers

Thanks to Corten's modular approach, handling recursive functions is strait forward in CortenC. Listing 19 demonstrates this by implementing a function, that returns the n th fibonacci number F_n and asserting that $F_n \geq n^2$. Because n is arbitrary, this proves that the fibonacci sequence grows faster than the n^2 .

```
fn fib(n: ty!{ nv: i32 | nv >= 0 }) -> ty!{ v: i32 | 10 * v >= nv * nv } {
  if n >= 2 {
    let n1 = n - 1; let n2 = n - 2;

    let f1 = fib(n1); let f2 = fib(n2);
    (f1 + f2) as ty!{ r : i32 | 10 * r >= nv * nv }
  } else {
    1
  }
}
```

Listing 19: Example demonstrating recursive function calls by proving a divergence property of the fibonacci sequence

7.3 Loop Invariants: Proof of the Gauß Summation Formula

The next example demonstrates, that CortenC can also verify loop invariants. Listing 20 shows a function, that calculates the sum $\sum_{i=1}^N i$ by repeatedly increment a variable by i . Of course there exists a closed formula computing the same result: $\frac{n \cdot (n+1)}{2}$, which is given as the specification. CortenC will therefore proof, that the closed form return the same result compared to the iterative summation.

This example is quite challenging, because of interdependencies between i and sum in the invariant as well as the old value of i and old values of sum in the loop body.

To proof, that the invariant holds after each loop body execution, we need the exact information, that i was incremented by one and sum by i as well as the knowledge, that i and sum satisfied the predicates. Corten solves this problem by treating predicates as immutable and retaining unassociated predicates (i.e. no longer associated with a program variable).

The second challenge is establishing the loop invariant in the first place: Relaxing the type of i from $= 0$ to $\leq nv$ is only valid, because we know relaxing the type of sum from $= 0$ to $2 * sv == iv * (iv + 1)$ is valid if $iv = 0$ and vice versa: To update the type of sum first, we need the knowledge, that $iv = 0$, but once we relax the type of i , i gets a new logic variable, which is distinct from the logic variable used in the type of sum . Thus the loop invariant needs to be established atomically.

For this purpose, CortenC provides the `relax_ctx!` macro, that allows the programmer to change the type of multiple variables at a time, as long as the previous state *as a whole* implies the newly chosen predicates.

```
fn gauss(n: ty!{ nv : i32 | nv > 0 })
-> ty!{ v : i32 | 2 * v == nv * (nv + 1) } {
  let mut i = 0;
  let mut sum = 0;

  // Loop Invariant:
  relax_ctx!{
    n |-> nv | nv > 0,
    i |-> iv | iv <= nv,
    sum |-> sv | 2 * sv == iv * (iv + 1)
  }
  while i < n {
    i = (i + 1);
    sum = (sum + i);
  }
  sum
}
```

Listing 20: Example loops with complex loop invariants and value updates affecting the invariant

7.4 Complex Mutable References

Listing 21 shows, that CortenC can handle complex mutation patterns in a modular way. The `swap` function replaces the value of `x` with `y` and vice versa. The specification states, that for any values `x1` and `y1` at the reference target of `x` and `y`, `swap` will update their values to satisfy the predicates `== y1` and `== x1`. Notable is, that CortenC can use these facts in the `client` function: For `*a` to resolve to the required type, we have to infer, that `b` is a reference to `j` with is associated with logic variable `j1` and `swap` ensures that the value referenced by `a` now references the value of `j1`. Therefore `a` must now refer to `j`, which is associated with logic variable `j1`.

```
fn swap(
  x : &mut ty!{ x1: i32 | true => x2 | x2 == y1 },
  y : &mut ty!{ y1: i32 | true => y2 | y2 == x1 }
) -> ty!{ v: () } {
  let tmp = *x;
  *x = *y;
  *y = tmp;
  ()
}

fn client(
  mut i: ty!{ i1 : i32 }, mut j: ty!{ j1 : i32 }
) -> ty!{ v: i32 | v == j1 - j1 } {
  let mut a = &mut i;
  let mut b = &mut j;
  swap(a, b);
  *a - *b
}
```

Listing 21: Example demonstrating modularity and ease of specification for complex mutation patterns

7.5 Rephrasing built-ins in terms of Refinement Types

The following example demonstrates, that Corten naturally extends to unusual edge cases and still allows easy verification based on them. Rust has built-in functions for aborting execution called `panic` and a function for conditionally aborting execution called `assert`. Listing 22 shows a reimplementation of these functions with full refined type specifications. `panic` asserts that after a call to it, `false` is valid. This naturally follows from inverting the path condition after exiting the loop. `assert` also uses the path condition to assert the correctness of `cond` and uses `panic` for the else case. The return type of `assert` can use the logic variable `c` from the input directly as the predicate, that is satisfied after the execution. The `client` just needs to make sure, that the proof of `c` is stored in the context by storing the value in the variable `_witness`: If the returned value is discarded, it will not be stored in the context and therefore will not be available when proving the return value specification.

```

fn panic() -> ty!{ v : () | false } {
    while(true) { () }
}

fn assert(cond: ty!{ c : bool }) -> ty!{ v: () | c } {
    if cond {
        () as ty!{ v: () | c }
    } else {
        panic()
    }
}

fn client(a : ty!{ av: i32 }) -> ty!{ v: i32 | v > 0 } {
    let arg = a > 0;
    let _witness = assert(arg);
    a
}

```

Listing 22: Example showing how panic and assert can be naturally specified and verified in CortenC

7.6 Interoperability with Other Tools

- Corten: only safe - Expend Corten by using external tools to verify unsafe - use assert to tell Corten what to assume and external tool what to verify

8 Related Work

There currently does not exist an implementation of refinement types for Rust.

Relevant papers originate from two lines of work. Firstly additions to refinement types for mutability, asynchronous execution etc. and secondly other verification frameworks for Rust.

For example, Lanzinger [10] successfully adapted refinement types to Java, which allows the user to check, that property types described by java annotations hold true throughout the program. At this point in time, specification and verification is limited to immutable (`final`) data.

Kloos et al. [9] extended refinement types to mutable and asynchronous programs. The paper explores how changes to possibly aliased memory cells can be tracked throughout a OCaml program. For that purpose the types are extended by a set of requirements on memory objects, which track distinctness and refined types of these memory cells. In contrast to OCaml, Rust already guarantees that mutable memory is not aliased and in particular *all mutable memory locations must be accessible by a variable name in the current context*, which offers substantial advantages in terms of simplicity to specification and verification of Rust programs.

Refinement types are also used in other applications. For example Graf et al. [5] use refinement types to check the exhaustiveness of pattern matching rules over complex (G)ADT types in Haskell. To check the exhaustiveness of patterns in Liquid Rust with ADTs may require similar approaches.

In terms of alternative verification approaches, Prusti[2] is notable, because of their work on formalizing the full Rust semantics, including `unsafe`. Prusti is a heavy-weight functional verification framework for Rust; based on separation logic.

Alternative verification approaches also exists: For example RustHorn[14] employs constrained horn clauses based verification to Rust. Particularity relevant for this thesis is the novel formalization for mutable references used in the paper. The authors stipulate that mutable references should be specified by a pre- and post-state from before a reference is borrowed to after it is returned.

The notion of Abstract Refinement Types that this thesis is based on is defined by Vazou et al. [22]. The basic idea is to allow the programmer to "refine" language types with predicates from a decidable logic. The type system has a notion of subtyping for refined types, where one type is a subtype of another if one predicate implies the other.

8.0.1 Alternative Approaches for Handling Mutability

Of course the design space permits choosing other tradeoffs. For example, Rondon et al. [19] chooses - only interdependence in single record using fold mechanism - this is not the only way to deal with mutability:

8.1 Comparison to Flux

During the writing of the thesis, Lehmann et al. [12] published a preprint for a paper, that is especially relevant. Lehmann et al. describe a implementation of Liquid Types for Rust called Flux, covering the same solution and problem space to Corten. For the purpose of plurality in design and implementation of such a system, it was decided to only read the paper after finishing the Corten type system.

Flux distinguished between three different kinds of references: Shared (immutable) references, mutable weak (unique) references and strong references.

```
#[lr::sig(fn(i32<@n>)  
-> bool<0 < n>)]  
fn is_pos(n: i32) -> bool {  
  if 0 < n { true } else { false }  
}  
  
#[lr::sig(fn(i32<@x>)  
-> bool{v: x <= x && 0 <= v})]  
fn add(x: i32) -> i32 {  
  if x < 0 { -x } else { x }  
}
```

```
fn is_pos(n: ty!{ n : i32})  
-> ty!{ v: bool | v == 0 < n } {  
  if 0 < n { true as ... } else { false }  
}  
  
fn add(x: ty!{ x: i32})  
-> ty!{ v: i32 | x <= x && 0 <= v } {  
  if x < 0 { -x as ... } else { x }  
}
```

Listing 23: Example demonstrating the Ownership System: `greet(a)` transfers ownership of `a` to `greet`

- Flux: explicit variable bind `@x`. Corten: every logic var implicitly - Flux: separate attribute. Corten: embed in types - Flux same observations: Exclusive Ownership => strong updates - Flux distinguish between strong and mut references - Corten: same, but mut passing references explicit modification

To demonstrate the difference in the handling of strong updates, consider the signatures of incrementing and decrementing functions of a datatype `Nat`. The functions are fully specifications and ensure the invariant, that a natural number is a positive.

```
fn increment(&strg v : Nat<n>) -> ()  
  ensures *self: Nat<n+1>;  
fn decrement(&strg v : Nat<n>) -> ()  
  requires n > 0  
  ensures *self: Nat<n-1>;
```

```
fn increment(n: &mut ty!{  
  n1: Nat => n1 | n1 == n1+1 }  
) -> ();  
  
fn decrement(n: &mut ty!{  
  v1: Nat | v1 > 0 => v2 | v2 == v1-1 }  
) -> ();
```

Listing 24: Comparison of specifying Type Changes Caused by a strong mutation. Flux on the left; Corten on the right

Flux type system has the function body judgement $\Sigma, \Delta \mid K; T \vdash F$, which "checks if a function body is well typed under a global environment Σ " [12, p. 11]

Δ in Flux is analogous to Corten's Φ . Both contain the constraints on refinement variables and are initialized with the function preconditions. Flux's T maps locations to types, and serves a similar purpose to μ in Corten: Keeping track of the current properties, that are known for a location. Corten's type system does not associate constraints with

locations with specific predicate, like Flux does, because in the typing rules of Corten, the origin of the constraint is irrelevant¹.

Flux has a separate type syntax for references: $\&_{\text{mut}}\tau$ that denotes that the referenced location has type τ . Corten uses a similar syntax for specifying mutable references, but in terms of the semantics delegates the type constraint to the referenced value and even disallows reference types to constrain the referenced type. Because the reference location can change during the program execution, constraining a reference can have many possible meanings and unintuitive consequences. Flux avoids this problem by restricting to weak references, which can not change their type. Presumably for the same reasons, Flux like Corten chose to forbid type specification on strong references, but their handling differs a bit: Even though strong reference destinations are specified with a similar syntax $\text{ptr}(l)$ (equivalent to $\&l$ in Corten), Corten allows multiple possible destinations to be specified. In that case, the reference may only be update weakly. Importantly, Corten allows the type to change, meaning a reference can be unambiguous and therefore strongly updatable and later ambiguous in its target meaning only weakly updatable.

Function signatures are expressed quite similarly. Flux function signatures have the form $\forall v : \sigma. \text{fn}(r; \bar{x}. T_i) \rightarrow \rho_0. T_0$. Consider the function decrease from listing 24. In Flux the function signature would be encoded as:

$$\forall n : \text{int}, \rho : \text{loc}. \text{fn}(n > 0, x. \{\rho \mapsto \text{nat}_\tau\langle n \rangle, x \mapsto \text{prt}(\rho)\} \rightarrow \rho_0. \{\rho_0 \mapsto (), \rho \mapsto \text{nat}_\tau\langle n - 1 \rangle\})$$

The rule T-DEF initializes the context to $\Delta = (n > 0, v : \text{loc}, n : \text{int})$, $K = (x \mapsto \text{prt}(\rho), \rho \mapsto \text{nat}_\tau\langle n \rangle)$ and the expected return value $T_0 = \{\rho_0 \mapsto (), \rho \mapsto \text{nat}_\tau\langle n - 1 \rangle\}$.

Corten would generate an initial context $\Gamma = (\mu, \Phi)$ with $\mu = \{n \mapsto \&\text{arg}_0, \text{arg}_0 \mapsto v_1\}$, $\Phi = \{v_1 > 0\}$ and expected return type $\{v_r : () \mid \text{true}\}$ and end context $\mu' = \{\text{arg}_0 \mapsto v_2\}$, $\Phi' = \{v_2 \doteq v_1 - 1, v_1 > 0\}$.

What Corten calls a sub-context relation, is referred to as context inclusion in Flux. Though Flux is significantly more conservative with its rules: Only allowing permutations (C-PERM), removal (C-WEAKEN) of predicates and subtyping of just individual types (C-SUB) and only if they can be separated (C-FRAME), meaning they do not depend on other predicates. Flux should thusly seem that Flux, can only show inclusion for a subset of what the logic and theories allow. As an example consider the sets displayed in fig. 8.1, which are described by the constraints super-context $x > y - 1 \wedge x > y + 1$, which allows x to vary from y by at most 1. A sensible sub-context would be $x = y$, which satisfies the constraints, but there is no possible proof for this fact. Because Corten delegates sub-context checking to SMT, handling these dependencies is no problem.

The subtyping rules are quite similar. To the degree applicable, S-EXISTS is analogous to $\leq\text{-TY}$.

Another difference is the choice of target language: Flux chose λ_{Rust} by Jung et al. [8] as a basis for the formalization, which expresses the semantics in a continuation passing style over a language based on Rustc's MIR. On the other hand Corten based the formalization on a language based on Rustc's HIR and semantics expressed in a small-spec semantic.

¹interestingly our implementation, like Flux, does have that distinction to improve error messages

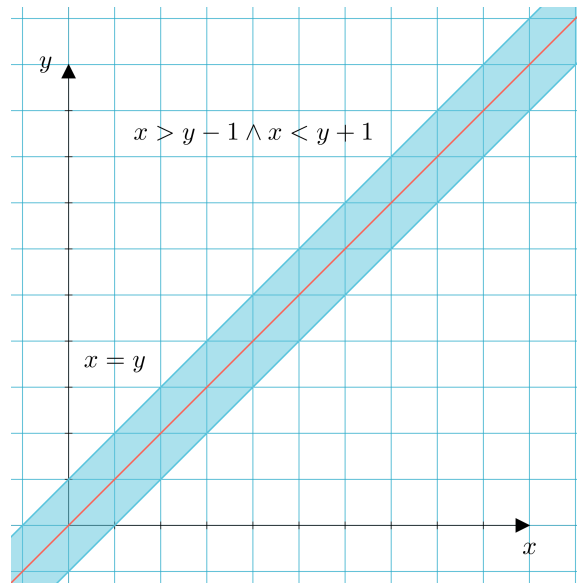


Figure 8.1: Figure showing a value space, where inclusion can not be shown

The corresponding implementations also use the MIR and HIR respectively. The advantages and disadvantages are discussed in chapter 6. Flux takes advantage of the MIR to cover a greater variety of control-flow scenarios. Lehmann et al. also found, that “the refinement annotations [...] do not appear in Rust’s MIR” [12, p. 17], which poses a challenge for the implementation.

Lastly Lehmann et al. evaluate Flux against Prusti proofing properties like “the absence of index-overflow error in a suite of vector-manipulating programs.” [12, p. 2] Lehmann et al. found, that the liquid type inference system is powerful enough to infer necessary invariants for this use case, partly because “loop invariants express either simple inequalities or tedious bookkeeping” [12, p. 21] In these cases Prusti still required invariants to be specified, which Lehmann et al. partly attribute to the fact, for every unchanged mutable reference, the invariant needs to state that the reference was not modified. By using `mut`, but not `strg` references, Flux could avoid that problem. Even though Corten does not share the the strong references mechanism with Flux, the argument still applies: Corten would also not require additional invariants for the unchanging mutable references. Because Corten does not support vectors (yet), a direct comparison is not possible.

9 Conclusion & Future Work

The goal of this thesis is to show, that Refinement Types can be idiomatically adapted to languages with unique mutable references. The thesis has show, that CortenC can type check Refinement Types in Rust, enforcing properties over mutable data and references and that these types fit idiomatically into Rust, if an inference system was added to CortenC. The type system was sufficiently expressive for verifying complex properties about mutable data, recursive functions and loop invariants. In addition, the thesis analyzed the use of mutability in Rust and reaffirmed the limited prevalence of unsafe in Rust code. For a subset of Rust (named Corten), a formal description of the syntax and small-step semantics was described. A syntax for Refinement Types for Rust was devised and an accompanying type system was designed. Based in the formal description, a proof-of-concept implementation shows, that automatic type checking of this type system is possible and feasible. The implementation also shows, that using the interface provided by Rustc is practical and beneficial for user-friendliness.

The evaluations has shown, that meaningful specifications can be automatically type checked. As planned and suspected, the lack of an inference system, calls for some avoidable specification, otherwise the evaluation has demonstrated, that the required specification effort is limited. The breath of features implemented in the thesis is limited, mainly because getting the foundations right took some experimenting, though the foundations layed in the thesis are purposefully designed with these extensions in mind. Therefore covering more language features should be unproblematic.

In terms of future work, there are two axis to work on: On one hand covering more language features, like records or algebraic data types, would greatly improve the practicality of CortenC.

On the other hand expanding the expressiveness of the type system would also be interesting: Introducing abstract predicates would allow a whole new set of properties to be specified. In particular developing a model of concurrency patterns could be a great application for them in Rust.

Bibliography

- [1] Vytautas Astrauskas et al. “How do programmers use unsafe rust?” In: *Proceedings of the ACM on Programming Languages* 4 (OOPSLA Nov. 13, 2020), pp. 1–27. ISSN: 2475-1421. DOI: 10.1145/3428204. URL: <https://dl.acm.org/doi/10.1145/3428204> (visited on 02/23/2022).
- [2] Vytautas Astrauskas et al. “Leveraging rust types for modular specification and verification”. In: *Proceedings of the ACM on Programming Languages* 3 (OOPSLA Oct. 10, 2019), pp. 1–30. ISSN: 2475-1421. DOI: 10.1145/3360573. URL: <https://dl.acm.org/doi/10.1145/3360573> (visited on 02/23/2022).
- [3] Travis Breaux and Jennifer Moritz. “The 2021 Software Developer Shortage is Coming”. In: *Commun. ACM* 64.7 (June 2021). Place: New York, NY, USA Publisher: Association for Computing Machinery, pp. 39–41. ISSN: 0001-0782. DOI: 10.1145/3440753. URL: <https://doi.org/10.1145/3440753>.
- [4] Johannes Foufas. *Why Volvo thinks you should have Rust in your car*. Volvo Cars Engineering. Sept. 23, 2022. URL: <https://medium.com/volvo-cars-engineering/why-volvo-thinks-you-should-have-rust-in-your-car-4320bd639e09> (visited on 09/26/2022).
- [5] Sebastian Graf, Simon Peyton Jones, and Ryan G. Scott. “Lower your guards: a compositional pattern-match coverage checker”. In: *Proceedings of the ACM on Programming Languages* 4 (ICFP Aug. 2, 2020), pp. 1–30. ISSN: 2475-1421. DOI: 10.1145/3408989. URL: <https://dl.acm.org/doi/10.1145/3408989> (visited on 03/15/2022).
- [6] Dan Grossman et al. “Region-Based Memory Management in Cyclone”. In: (), p. 12.
- [7] Ralf Jung. *MiniRust*. original-date: 2022-05-21T15:30:40Z. Sept. 28, 2022. URL: <https://github.com/RalfJung/minirust> (visited on 09/28/2022).
- [8] Ralf Jung et al. “RustBelt: securing the foundations of the Rust programming language”. In: *Proceedings of the ACM on Programming Languages* 2 (POPL Jan. 2018), pp. 1–34. ISSN: 2475-1421. DOI: 10.1145/3158154. URL: <https://dl.acm.org/doi/10.1145/3158154> (visited on 02/17/2022).
- [9] Johannes Kloos, Rupak Majumdar, and Viktor Vafeiadis. “Asynchronous Liquid Separation Types”. In: (2015). In collab. with Marc Herbstritt. Artwork Size: 25 pages Medium: application/pdf Publisher: Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik GmbH, Wadern/Saarbruecken, Germany, 25 pages. DOI: 10.4230/LIPICS.ECOOP.2015.396. URL: <http://drops.dagstuhl.de/opus/volltexte/2015/5223/> (visited on 01/27/2022).

- [10] Florian Lanzinger. “Property Types in Java: Combining Type Systems and Deductive Verification”. Master Thesis. Karlsruher Institut für Technologie, Feb. 2021.
- [11] Michael Larabel. *LPC 2022: Rust Linux Drivers Capable Of Achieving Performance Comparable To C Code*. phoronix. Sept. 12, 2022. URL: <https://www.phoronix.com/news/LPC-2022-Rust-Linux> (visited on 09/26/2022).
- [12] Nico Lehmann et al. *Flux: Liquid Types for Rust*. Number: arXiv:2207.04034. July 8, 2022. arXiv: 2207 . 04034[cs]. URL: <http://arxiv.org/abs/2207.04034> (visited on 07/14/2022).
- [13] Nicholas D. Matsakis and Felix S. Klock. “The rust language”. In: *ACM SIGAda Ada Letters* 34.3 (Oct. 18, 2014), pp. 103–104. ISSN: 1094-3641. DOI: 10.1145/2692956.2663188. URL: <https://doi.org/10.1145/2692956.2663188> (visited on 09/26/2022).
- [14] Yusuke Matsushita, Takeshi Tsukada, and Naoki Kobayashi. “RustHorn: CHC-based verification for Rust programs”. In: *European Symposium on Programming*. Springer, Cham, 2020, pp. 484–514.
- [15] *Overview of the Compiler - Guide to Rustc Development*. URL: <https://rustc-dev-guide.rust-lang.org/overview.html> (visited on 09/19/2022).
- [16] Benjamin C. Pierce. *Types and Programming Languages*. Google-Books-ID: ti6zoAC9Ph8C. MIT Press, Jan. 4, 2002. 656 pp. ISBN: 978-0-262-16209-8.
- [17] Benjamin C. Pierce. *Types and programming languages*. Cambridge, Mass: MIT Press, 2002. 623 pp. ISBN: 978-0-262-16209-8.
- [18] Patrick M. Rondon, Ming Kawaguchi, and Ranjit Jhala. “Liquid types”. In: *Proceedings of the 2008 ACM SIGPLAN conference on Programming language design and implementation - PLDI '08*. the 2008 ACM SIGPLAN conference. Tucson, AZ, USA: ACM Press, 2008, p. 159. ISBN: 978-1-59593-860-2. DOI: 10.1145/1375581.1375602. URL: <http://portal.acm.org/citation.cfm?doid=1375581.1375602> (visited on 06/30/2022).
- [19] Patrick Maxim Rondon, Ming Kawaguchi, and Ranjit Jhala. “Low-level liquid types”. In: *Proceedings of the 37th annual ACM SIGPLAN-SIGACT symposium on Principles of programming languages*. POPL '10. New York, NY, USA: Association for Computing Machinery, Jan. 17, 2010, pp. 131–144. ISBN: 978-1-60558-479-9. DOI: 10.1145/1706299.1706316. URL: <https://doi.org/10.1145/1706299.1706316> (visited on 09/16/2022).
- [20] Amr Sabry and Matthias Felleisen. “Reasoning about Programs in Continuation-Passing Style”. In: *Lisp and Symbolic Computation*. 1993, pp. 288–298.
- [21] *The THIR (Typed High-level IR) - Guide to Rustc Development*. URL: <https://rustc-dev-guide.rust-lang.org/thir.html> (visited on 09/19/2022).
- [22] Niki Vazou, Patrick M. Rondon, and Ranjit Jhala. “Abstract Refinement Types”. In: *Programming Languages and Systems*. Ed. by Matthias Felleisen and Philippa Gardner. Berlin, Heidelberg: Springer, 2013, pp. 209–228. ISBN: 978-3-642-37036-6. DOI: 10.1007/978-3-642-37036-6_13.

-
- [23] Philip Wadler. “Theorems for free!” In: *Proceedings of the fourth international conference on Functional programming languages and computer architecture*. FPCA ’89. New York, NY, USA: Association for Computing Machinery, Nov. 1, 1989, pp. 347–359. ISBN: 978-0-89791-328-7. DOI: 10.1145/99370.99404. URL: <https://doi.org/10.1145/99370.99404> (visited on 08/29/2022).

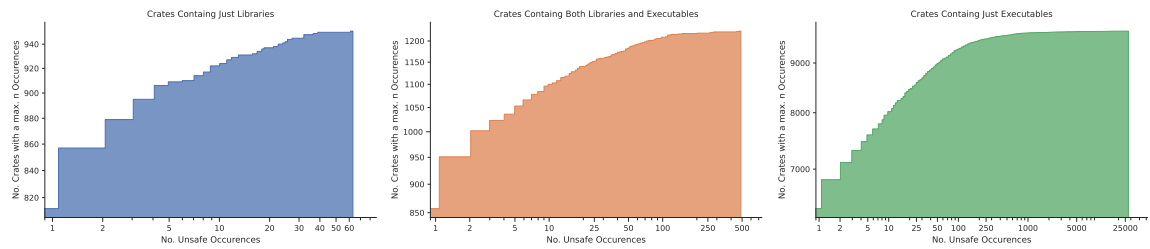


Figure .1: Cumulative, Logarithmic Histogram of the Amount of unsafe Uses in each Category

