

Gradual Program Verification with Implicit Dynamic Frames

Master's Thesis of

Johannes Bader

at the Department of Informatics
Institute for Program Structures and Data Organization (IPD)

Reviewer: Prof. Dr.-Ing. Gregor Snelting, Karlsruhe Institute of Technology - Karlsruhe, Germany

Advisors: Assoc. Prof. Jonathan Aldrich, Carnegie Mellon University - Pittsburgh, USA
Assoc. Prof. Éric Tanter, University of Chile - Santiago, Chile

Duration: 2016-05-10 – 2016-09-28

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, and have followed the rules of the KIT for upholding good scientific practice.

Karlsruhe, 2016-09-??

.....
(Johannes Bader)

Abstract

Both static and dynamic program verification approaches have disadvantages potentially disqualifying them as a single methodology to rely on. Motivated by gradual type systems which solve a very similar dilemma in the world of type systems, we propose *gradual verification*, an approach that seamlessly combines static and dynamic verification. Drawing on principles from abstract interpretation and recent work on *abstracting gradual typing* by Garcia, Clark and Tanter, we formalize steps to obtain a gradual verification system in terms of a static one.

This approach yields *by construction* a verification system that is compatible with the original static system, but overcomes its rigidity by resorting to methods of dynamic verification if necessary. In a case study, we show the flexibility of our approach by applying it to a statically verified language that uses implicit dynamic frames to enable race-free reasoning.

Acknowledgments

I wish to thank my advisors Jonathan Aldrich and Éric Tanter for offering me this topic and for their patient assistance throughout the past few months. In moments of uncertainty, their remarks and thoughts guided me in the right direction.

Also I am very grateful to all my family and friends who encouraged and supported me during .

Contents

1	Introduction	3
1.1	Motivation	5
1.1.1	As extension to unverified setting	5
1.1.2	As extension to fully verified setting	6
2	Background	9
2.1	Hoare Logic	9
2.2	Abstracting Gradual Typing	10
2.3	Implicit Dynamic Frames	10
3	Gradualization of a Statically Verified Language	13
3.1	A Generic Statically Verified Language (SVL)	13
3.2	Gradual Formulas	17
3.2.1	Dedicated wildcard formula	17
3.2.2	Wildcard with upper bound	17
3.2.3	Precision	18
3.2.4	UNSORTED	18
3.3	Gradual Statements	19
3.4	Gradual Program State	19
3.5	Lifting Predicates and Functions	20
3.5.1	Lifting Predicates	20
3.5.2	Lifting Functions	22
3.5.3	Generalized Lifting	23
3.6	Abstracting Static Semantics	23
3.6.1	The Problem with a Predicate Lifting	24
3.6.2	The Deterministic Approach	24
3.7	Gradual Soundness vs Gradual Guarantee	25
3.8	Abstracting Dynamic Semantics	26
3.8.1	Perfect Knowledge	26
3.8.2	Partial Knowledge	28
4	Case Study: Implicit Dynamic Frames	29
4.1	Language	29
4.1.1	Syntax	29
4.1.2	Program State	30
4.1.3	Formula Semantics	30
4.1.4	Static Semantics	32
4.1.5	Well-Formedness	34
4.1.6	Dynamic Semantics	35
4.1.7	Soundness	35
4.2	Gradualization	35
4.2.1	Extension: Statements	35
4.2.2	Extension: Program State	35

4.3	Gradualize Hoare Rules	36
4.4	Gradual Dyn. Semantics	36
4.5	Enhancing an Unverified Language	36
5	Evaluation/Analysis	41
6	Conclusion	43
6.1	Conceptual Nugget: Comparison/Implication to AGT!	43
6.2	Limitations	43
6.3	Future Work	43
7	Appendix	45
8	UNSORTED	47
8.1	HoareMotivEx	47
8.2	NPC formula	47
8.2.1	Impact of NP-hard Verification Predicates	48

1 Introduction

Program verification aims to check a compute program against its specification. Automated methods require this specification to be formalized, e.g. using annotations in the source code. Common examples are method contracts, loop invariants and assertions.

Approaches to check whether program behavior complies with given annotations can be divided into two categories:

	Static verification	Dynamic verification
Approach	The program is not executed. Instead formal methods (like Hoare logic or separation logic) are used, trying to derive a proof for given assertions.	The specification is turned into runtime checks , making sure that the program adheres to its specification during execution. Violations cause a runtime exception to be thrown, effectively preventing the program from entering a state that contradicts its specification. Note that in practice this approach is often combined with control flow based testing techniques to detect misbehavior as early as possible.
Drawbacks	The syntax available for static verification is naturally limited by the underlying formal logic. Complex properties might thus not be expressible, resulting in inability to prove subsequent goals. Furthermore, the logic itself might be unable to prove certain goals due to code complexity and undecidability in general. Using static verification usually requires rigorous annotation of the entire source code, as otherwise there might be too little information to find a proof. While fully annotating own code can be tedious (there are supporting tools), using unannotated libraries can become a problem: Even if it is possible annotate the API afterwards, lacking the source code the verifier is unable to prove those annotations. In case the annotation are wrong this results in inconsistent proves.	Violations are only detected at runtime, with the risk of going unnoticed before software is released. To minimize this risk, testing methods are required, i.e. more time has to be spent after compilation. The usage of runtime checks naturally imposes a runtime overhead which is not always acceptable.

1 Introduction

The goal of this work is to formalize “gradual verification”, an approach that seamlessly combines static and dynamic verification in order to weaken or even avoid above drawbacks. The resulting system provides a continuum between traditional static and dynamic verification, meaning that both extremes are compatible with, but only special cases of the gradual verification system. Section 1.1 gives example scenarios of both static and dynamic verification suffering from their drawbacks, but illustrating how gradual verification could avoid them.

Our approach is based on recent formalizations regarding gradual typing, using the concept of abstract interpretation to define a gradual system in terms of a static one (this process is called “gradualization”). Gradual typing arose from drawbacks of static and dynamic type systems which are very similar to the drawbacks identified above. From a theoretical perspective, type systems are even a special case of program verification. These similarities motivated our idea of reinterpreting and adapting the gradual typing approach to the verification setting.

Chapter 2 provides the background of our approach, introducing the concepts motivating and driving our approach. Furthermore it categorizes existing work that goes in a similar direction, pointing out how it differs from our work. In chapter 3 we describe our approach of gradualization in a generic way, meant to be used as a manual or template for designing gradual verification systems. We do just that in form of case study in chapter 4, applying the approach to a statically verified language that uses implicit dynamic in order to enable race-free static reasoning.

~\\ % more detail???

Most modern programming languages use static methods to some degree, ruling out at least
%% static typing

Static typing disciplines are among the most common representatives, guaranteeing type safety. Yet, the rigidity and limitations of static type systems resulted in the introduction of casts (e.g. as implemented in C# or Java) overrule purely static reasoning, allowing the programmer to bypass the type system. At this location, a runtime check is introduced, resulting in a cast exception should the cast fail. Note that such deviations from a purely static type system (one where there is no need for casts) are not ideal. It is still guaranteed that execution does not enter an inconsistent state by simply introducing casts.

Note that casts are necessary only because of a typical drawback of static systems, namely the lack of flexibility. More sophisticated type systems (e.g. the one in Haskell) might have been able to deduce the need for casts automatically.

%% dynamic typing

%At the other end of the spectrum are dynamically typed languages.

%In scenarios where the limitations of a static type system would clutter up the source code, dynamically typed languages are a better choice.

%% static verification

In contrast, general purpose static verification techniques are not common amongst popular programming languages. Note that such languages are usually driven by cost-benefit and usability considerations, rather than by theoretical concerns.

%% static verification

% example?

% research Eiffel!

% Design-by-Contract!!! Eiffel!

```
% D even has both

% this is more of a consequence of the "deep roots" of dynamic verification!!!
%But even preconditions at expression level are implemented as runtime checks, reflected
%Examples:
%\begin{description}
%  \item[Division by zero]~\\
%    Integer division performs a dynamic check...
%
%\end{description}
```

-

What is the thesis about?
 Why is it relevant or important?
 What are the issues or problems?
 What is the proposed solution or approach?
 What can one expect in the rest of the thesis?

"Static verification checks that properties are always true, but it can be difficult and

1.1 Motivation

1.1.1 As extension to unverified setting

Motivating example:

```
boolean hasLegalDriver(Car c)
{
    return c.driver.age >= 18;
}
```

Motivating example with argument validation:

```
boolean hasLegalDriver(Car c)
{
    if (!(c != null))
        throw new IllegalArgumentException("expected c != null");
    if (!(c.driver != null))
        throw new IllegalArgumentException("expected c.driver != null");

    // business logic (requires 'c.driver.age' to evaluate)
}
```

Motivating example with declarative approach (JML syntax):

```
//@ requires c != null && c.driver != null;
boolean hasLegalDriver(Car c)
{
    // business logic (requires 'c.driver.age' to evaluate)
}
```

There are two basic ways to turn this annotation into a guarantee:

Static Verification (e.g. ESC/Java [12])

In the unlikely event that the verifier can prove the precondition at all call sites, our problem is solved. Otherwise, we have to enhance the call sites in order to convince the verifier. Choices:

- Add parameter validation, effectively duplicating the original runtime check across the program.
- Add further annotations, guiding the verifier towards a proof. This might not always work due to limitations of the verifier or decidability in general.

There are obvious limitations to this approach, static verification tends to be invasive. At least there is a performance benefit: Runtime checks (originally part of every call) are now only necessary in places where verification would not succeed otherwise.

Runtime Assertion Checking (e.g. run JML4c [18])

This approach basically converts the annotation back into a runtime check equivalent to our manual argument validation. It is therefore less invasive, not requiring further changes to the code, but also lacks the advantages (no perf. impact, static guarantee) of static verification.

1.1.2 As extension to fully verified setting

```
int collatzIterations(int iter, int start)
  requires 0 < start;
  ensures 0 < result;
{
  // ...
}

int myRandom(int seed)
  requires 0 < seed && seed < 10000;
  ensures 0 < result && result < 4;      // not provable
{
  int result = collatzIterations(300, seed);
  // we know: result ∈ { 1, 2, 4 }

  if (result == 4) result = 3;
  return result;
}
```

Non-solution:

```
int collatzIterations(int iter, int start)
  requires 0 < start;
  ensures 0 < result;
{
  // ...
}

int myRandom(int seed)
  requires 0 < seed && seed < 10000;
  ensures 0 < result && result < 4;
{
```

```

int result = collatzIterations(300, seed);
// we know: result ∈ { 1, 2, 4 }

// "cast"
if (!(result < 5))
    throw new IllegalStateException("expected result < 5");

// verifier now knows: 0 < result && result < 5

if (result == 4) result = 3;
return result;
}

```

This solution is not satisfying, - much to write, have to think about what to write (requires you to kind of think from verifiers perspective) - intuitively the problem is with the method's postcondition being too weak, i.e. we solved the problem at the wrong place!

```

int collatzIterations(int iter, int start)
    requires 0 < start;
    ensures 0 < result && ?;
{
    // ...
}

int myRandom(int seed)
    requires 0 < seed && seed < 10000;
    ensures 0 < result && result < 4;
{
    int result = collatzIterations(300, seed);
    // we know: result ∈ { 1, 2, 4 }

    // verifier allowed to
    // assume 0 < result && result < 5
    // from 0 < result && ?
    // (adding runtime check)

    if (result == 4) result = 3;
    return result;
}

```


2 Background

[1] GraVy: metric of progress of the verification process and allows the verification engineer to focus on the remaining statements. Gradual verification is not a new static verification technique. It is an extension that can be applied to any existing static verification techniques to provide additional information to the verification engineer. Thus, issues, such as handling of loops or aliasing are not addressed in this paper. These are problems related to sound verification, but gradual verification is about how to make the use of such verification more traceable and quantifiable

[16] ESC/Java Software development and maintenance are costly endeavors. The cost can be reduced if more software defects are detected earlier in the development cycle. This paper introduces the Extended Static Checker for Java (ESC/Java), an experimental compile-time program checker that finds common programming errors. The checker is powered by verification-condition generation and automatic theorem proving techniques. It provides programmers with a simple annotation language with which programmer design decisions can be expressed formally. ESC/Java examines the annotated software and warns of inconsistencies between the design decisions recorded in the annotations and the actual code, and also warns of potential runtime errors in the code. This paper gives an overview of the checker architecture and annotation language and describes our experience applying the checker to tens of thousands of lines of Java programs.

[10] JML => static verification

[4] JML => RAC ...lead up to [18] JML4c

[15] Spec#

[3] Spec# extension (concurrent OO)

[14] Design-by-Contract then also: Eiffel by Bertrand Meyer

[13] Code Contracts! Combines runtime and static checking

[6] “verified design-by-contract”

[5] = static verification plus directed dynamic verification In this paper, we present a technique to complement partial verification results by automatic test case generation. In contrast to existing work, our technique supports the common case that the verification results are based on unsound assumptions. We annotate programs to reflect which executions have been verified, and under which assumptions. These annotations are then used to guide dynamic symbolic execution toward unverified program executions. Our main technical contribution is a code instrumentation that causes dynamic symbolic execution to abort tests that lead to verified executions, to prune parts of the search space, and to prioritize tests that cover more properties that are not fully verified.

2.1 Hoare Logic

...for static semantics

[9]

2.2 Abstracting Gradual Typing

[19] Gradual Typing for Functional Languages

apply their gradual typing approach to other areas

[20] Refined criteria for gradual typing gradual guarantee: The gradual guarantee says that if a gradually typed program is well typed, then removing type annotations always produces a program that is still well typed. Further, if a gradually typed program evaluates to a value, then removing type annotations always produces a program that evaluates to an equivalent value.

[7] AGT In this paper, we propose a new formal foundation for gradual typing, drawing on principles from abstract interpretation to give gradual types a semantics in terms of preexisting static types. Abstracting Gradual Typing (AGT for short) yields a formal account of consistency—one of the cornerstones of the gradual typing approach—that subsumes existing notions of consistency, which were developed through intuition and ad hoc reasoning.

[8] Abstracting Gradual Typing (AGT) is an approach to systematically deriving gradual counterparts to static type disciplines (Garcia et al. 2016). The approach consists of defining the semantics of gradual types by interpreting them as sets of static types, and then defining an optimal abstraction back to gradual types. These operations are used to lift the static discipline to the gradual setting. The runtime semantics of the gradual language then arises as reductions on gradual typing derivations. To demonstrate the flexibility of AGT, we gradualize a prototypical security-typed language with respect to only security labels rather than entire types, yielding a type system that ranges gradually from simply-typed to securely typed. We establish noninterference for our gradual language using Zdancewic’s logical relation proof method. Whereas prior work presents gradual security cast languages, which require explicit security casts, this work yields the first gradual security source language, which requires no explicit casts.

prior to AGT [24] the language extends the notion of gradual typing to account for typestate: gradual typestate checking seamlessly combines static and dynamic checking by automatically inserting runtime checks into programs.

[2] develop a theory of gradual effect checking, which makes it possible to incrementally annotate and statically check effects, while still rejecting statically inconsistent programs. We extend the generic type-and-effect framework of Marino and Millstein with a notion of unknown effects, which turns out to be significantly more subtle than unknown types in traditional gradual typing. We appeal to abstract interpretation to develop and validate the concepts of gradual effect checking.

[23] Grad Effects in Scala, benchmarks on runtime impact!

2.3 Implicit Dynamic Frames

Race-free Assertion language! => static verification tool able to reason soundly about concurrent programs

[21] IDF

[11] Chalice, a verification methodology based on implicit dynamic frames

Chalice’s verification methodology centers around permissions and permission transfer. In particular, a memory location may be accessed by a thread only if that thread has permission to do so. Proper use of permissions allows Chalice to deduce upper bounds on the set of locations modifiable by a method and guarantees the absence of data races for

concurrent programs. The lecture notes informally explain how Chalice works through various examples.

also: Viper (Verification Infrastructure for Permission-based Reasoning; is a suite of tools developed at ETH Zurich, providing an architecture on which new verification tools and prototypes can be developed simply and quickly.) has Chalice as front-end

[22] In this paper, we provide both an isorecursive and an equirecursive formal semantics for recursive definitions in the context of Chalice

[17] VERY IMPORTANT: chapter 2.2

Finally, we show that we can encode the separation logic fragment of our logic into the implicit dynamic frames fragment, preserving semantics. For the connectives typically supported by tools, this shows that separation logic can be faithfully encoded in a first-order automatic verification tool (Chalice).

Although IDF was partially inspired by separation logic, there are many differences between SL and IDF that make understanding their relationship difficult. SL does not allow expressions that refer to the heap, while IDF does. SL is defined on partial heaps, while IDF is defined using total heaps and permission masks. The semantics of IDF are only defined by its translation to first-order verification conditions, while SL has a direct Kripke semantics for its assertions.

3 Gradualization of a Statically Verified Language

As illustrated earlier gradual verification can be seen as an extension of both static and dynamic verification. Yet, our approach of “gradualization” formalizes the introduction of the dynamic aspect into a fully static system. Later we will show how a programming language without static verification features can be approached.

3.1 A Generic Statically Verified Language (SVL)

Requirements:

Syntax

We assume the existence of the following two syntactic categories:

$$\begin{aligned} s &\in \text{STMT} \\ \phi &\in \text{FORMULA} \end{aligned}$$

We assume that there is a sequence operator $;$ such that:

$$\forall s_1, s_2 \in \text{STMT}. \quad s_1; s_2 \in \text{STMT}$$

Let STMT_s be the set of all statements having s as prefix.

Program State

Operational semantics (see below) are formalized as discrete transitions between program states. Therefore a program state contains all information necessary to evaluate expressions and determine the next program state. We assume that PROGRAMSTATE is the set of all possible program states in SVL.

In order to determine the next program state (or detect termination), a state must have a notion of “upcoming work”, usually represented by a statement internally. TODO: reasonable to call that “continuation”?

Let PROGRAMSTATE_s (with $s \in \text{STMT}$) be the set of program states having s as upcoming work. This notion will be necessary to define soundness of SVL’s static semantics.

Examples:

Primitive

$$\text{PROGRAMSTATE} = \underbrace{(\text{VAR} \rightarrow \mathbb{Z})}_{\text{variable memory}} \times \text{STMT}$$

$$\text{PROGRAMSTATE}_s = (\text{VAR} \rightarrow \mathbb{Z}) \times \text{STMT}_s$$

Stack

$$\begin{aligned} \text{PROGRAMSTATE} &= \bigcup_{i \in \mathbb{N}_+} \underbrace{\left((\text{VAR} \rightarrow \mathbb{Z}) \times \text{STMT} \right)^i}_{\text{stack frame}} \\ \text{PROGRAMSTATE}_s &= \left((\text{VAR} \rightarrow \mathbb{Z}) \times \text{STMT}_s \right) \times \underbrace{\bigcup_{i \in \mathbb{N}_0} \left((\text{VAR} \rightarrow \mathbb{Z}) \times \text{STMT} \right)^i}_{\text{lower frames}} \end{aligned}$$

TODO: notion of initial state?

Formula Semantics

Formulas are used to describe program states. For example, a method contract stating `arg > 4` as precondition is supposed to make sure that the method is only entered, if `arg` evaluates to a value larger than 4 in the program state at the call site.

We assume that we are given a computable predicate

$$\cdot \models \cdot \subseteq \text{PROGRAMSTATE} \times \text{FORMULA}$$

that decides, whether a formula is satisfied given a concrete program state.

We can derive a notion of satisfiability, implication and equivalence from this evaluation predicate.

Definition 3.1.1 (Formula Satisfiability).

A formula ϕ is **satisfiable** iff

$$\exists \pi \in \text{PROGRAMSTATE}. \pi \models \phi$$

Let $\text{SATFORMULA} \subseteq \text{FORMULA}$ be the set of satisfiable formulas.

Definition 3.1.2 (Formula Implication).

A formula ϕ_1 **implies** formula ϕ_2 (written $\phi_1 \xRightarrow{\phi} \phi_2$) iff

$$\forall \pi \in \text{PROGRAMSTATE}. \pi \models \phi_1 \implies \pi \models \phi_2$$

Definition 3.1.3 (Formula Equivalence).

Two formulas ϕ_1 and ϕ_2 are **equivalent** (written $\phi_1 \equiv \phi_2$) iff

$$\phi_1 \xRightarrow{\phi} \phi_2 \wedge \phi_2 \xRightarrow{\phi} \phi_1$$

Lemma 3.1.4 (Partial Order of Formulas).

The implication predicate is a partial order on FORMULA .

We assume that there is a largest element $\text{true} \in \text{FORMULA}$. Note that the presence of an unsatisfiable formula (as invariant, pre-/postcondition, assertion, ...) in a sound verification system implies that the corresponding source code location is unreachable: Preservation guarantees that any reachable program state satisfies potentially annotated formulas, trivially ensuring that the formula is satisfiable.

This property is true regardless of whether SVL forbids usage of unsatisfiable formulas entirely or whether it only fails when trying to use the corresponding code (which would involve proving that a satisfiable formula implies an unsatisfiable one). Therefore we will often restrict our reasoning on the satisfiable formulas SATFORMULA , without explicitly stating that the presence of an unsatisfiable formula would result in failure.

Dynamic Semantics

We assume that there is a small-step semantics $\mathcal{S} \subseteq \text{PROGRAMSTATE} \rightarrow \text{PROGRAMSTATE}$ describing precisely how program state can be updated. Taking n steps at once can be abbreviated as \mathcal{S}^n , where undefinedness is propagated.

$$\mathcal{S}^s \subseteq \text{PROGRAMSTATE}_s \times \text{PROGRAMSTATE}$$

$$\mathcal{S}^s(\pi_s, \pi) \stackrel{\text{def}}{\iff} \exists n \in \mathbb{N}_+. \mathcal{S}^n(\pi_s) = \pi \wedge \pi \text{ is the first state after } s \text{ is fully consumed}$$

We further assume that there is a designated non-empty set $\text{PROGRAMSTATEFIN} \subseteq \text{PROGRAMSTATE}$ of states indicating regular termination of the program. W.l.o.g. we assume $\text{dom}(\mathcal{S}) \cap \text{PROGRAMSTATEFIN} = \emptyset$, e.g. final states are stuck. Optionally, there may be a subset $\text{PROGRAMSTATEEX} \subseteq \text{PROGRAMSTATEFIN}$ of states indicating exceptional termination of the program. To simplify reasoning about exceptional states, we assume

$$\forall \pi_X \in \text{PROGRAMSTATEEX}, \phi \in \text{FORMULA}. \pi_X \models \phi$$

and something with special statement set?

With this semantics we can formalize the notion of valid Hoare triples:

$$\begin{aligned} \models \{\cdot\} \cdot \{\cdot\} &\subseteq \text{FORMULA} \times \text{STMT} \times \text{FORMULA} \\ \models \{\phi_{pre}\} s \{\phi_{post}\} &\stackrel{\text{def}}{\iff} \forall \langle \pi_{pre}, \pi_{post} \rangle \in \mathcal{S}^s. \pi_{pre} \models \phi_{pre} \implies \pi_{post} \models \phi_{post} \end{aligned}$$

Static Semantics

We assume that there is a Hoare logic (HL)

$$\vdash \{\cdot\} \cdot \{\cdot\} \subseteq \text{SATFORMULA} \times \text{STMT} \times \text{SATFORMULA}$$

describing which programs (together with pre- and postconditions about the program state) are accepted. While the Hoare logic might be defined for arbitrary formulas in practice, we only ever reason about it in presence of satisfiable formulas, hence the “restricted domain”???

In practice, this predicate might also have further parameters. For instance, a statically typed language might require a type context to safely deduce

$$x : \text{int} \vdash \{\text{true}\} x := 3 \{ (x = 3) \}$$

As we will see later, further parameters are generally irrelevant for and immune to gradualization, so it is reasonable to omit them for now.

We assume that

$$\frac{\vdash \{\phi_p\} s_1 \{\phi_q\} \quad \vdash \{\phi_q\} s_2 \{\phi_r\}}{\vdash \{\phi_p\} s_1 ; s_2 \{\phi_r\}} \text{HOARESEQUENCE}$$

is derivable from given Hoare rules.

3 Gradualization of a Statically Verified Language

We further assume that this predicate is monotonic in the precondition w.r.t. implication:

$$\begin{aligned}
& \forall s \in \text{STMT}. \\
& \forall \phi_1, \phi_2 \in \text{FORMULA}. \\
& \quad \forall \phi'_1 \in \text{FORMULA}. (\phi_1 \xRightarrow{\phi} \phi_2) \wedge \vdash \{\phi_1\} s \{\phi'_1\} \\
& \implies \exists \phi'_2 \in \text{FORMULA}. (\phi'_1 \xRightarrow{\phi} \phi'_2) \wedge \vdash \{\phi_2\} s \{\phi'_2\}
\end{aligned}$$

Intuitively, this means that more knowledge about the initial program state can not result in a loss of information about the final state.

Definition 3.1.5 (Weakest Static Precondition).

Let $\text{wsp} : \text{STMT} \rightarrow \mathcal{P}(\text{PROGRAMSTATE})$ be defined as

$$\text{wsp}(s) = \{ \pi \in \text{PROGRAMSTATE}_s \mid \exists \phi_1, \phi_2 \in \text{FORMULA}. \vdash \{\phi_1\} s \{\phi_2\} \wedge \pi \models \phi_1 \}$$

Intuitively, the $\text{wsp}(s)$ is a predicate on program states, indicating whether we could deduce anything about the state after executing s , using only our Hoare rules.

Example:

- Given that

$$\frac{}{\vdash \{\phi[e/x]\} x := e \{\phi\}} \text{HOAREASSIGN}$$

is the only Hoare rule for assignment, it follows that

$$\text{wsp}(x := e) = \text{PROGRAMSTATE}$$

- Given that

$$\frac{\phi \xRightarrow{\phi} \phi_a}{\vdash \{\phi\} \text{assert } \phi_a \{\phi\}} \text{HOARESTATICASSERT}$$

is the only Hoare rule for assertions, it follows that

$$\text{wsp}(\text{assert } \phi_a) = \{ \pi \in \text{PROGRAMSTATE} \mid \pi \models \phi_a \}$$

Soundness

We expect that given static semantics are sound w.r.t. given dynamic semantics.

$$\frac{\pi \in \text{wsp}(s_1)}{\exists n \in \mathbb{N}_+, s_2 \in \text{STMT}. \mathcal{S}^n(\pi) \in \text{PROGRAMSTATE}_{s_2}} \text{PROGRESS}$$

$$\frac{\vdash \{\phi_1\} s \{\phi_2\}}{\models \{\phi_1\} s \{\phi_2\}} \text{PRESERVATION}$$

3.2 Gradual Formulas

We introduce the concepts of gradual verification by introducing a wildcard formula $?$ into the formula syntax, resulting in a new set of gradual formulas GFORMULA . There are different ways to introduce the wildcard, we will describe two common options in the following sections.

Note that we want to strictly extend the existing formula syntax in order to maintain compatibility with the static system, i.e. $\text{FORMULA} \subset \text{GFORMULA}$ holds. This design goal ensures that any program considered syntactically valid by the static system will still be syntactically valid in the gradual system (motivated by gradual guarantee 2.2).

We decorate formulas $\tilde{\phi} \in \text{GFORMULA}$ to distinguish them from formulas drawn from FORMULA . Using the concept of abstract interpretation, we want to reason about gradual formulas by mapping them back to a set of satisfiable static formulas (called “concretization”) and then applying static reasoning to that set. Intuitively, a program state satisfies a gradual formula iff it satisfies (at least) one of the static formulas of its concretization. (This intuition is formalized in section 3.5.1.)

Without knowing specifics of the syntax extension, we can already formalize this approach for static formulas:

Definition 3.2.1 (Concretization).

Let $\gamma : \text{GFORMULA} \rightarrow \mathcal{P}(\text{SATFORMULA})$ be defined as follows:

$$\gamma(\phi) = \begin{cases} \{ \phi \} & \phi \in \text{SATFORMULA} \\ \emptyset & \text{otherwise} \end{cases}$$

$$\gamma(\tilde{\phi}) = \text{to be defined when extending the syntax} \quad \forall \phi \in \text{GFORMULA}$$

There are two typical ways of extending the formula syntax.

3.2.1 Dedicated wildcard formula

The most straight forward way to extend the syntax is by simply adding $?$ as a dedicated formula:

$$\tilde{\phi} ::= \phi \mid ?$$

This is analogous to how most gradually typed languages are realized (e.g. `dynamic-type` in C# 4.0 and upward).

Since $?$ is supposed to be a placeholder for an arbitrary formula, its concretization is defined as.

$$\gamma(?) = \text{SATFORMULA}$$

This approach is limited since programmers cannot express any additional static knowledge they might have. For example, a programmer might resort to using the wildcard lacking some knowledge about variable x (or being unable to express it), whereas he could give a static formula for y , say $(y = 3)$. Yet, there is no way to express this information as soon as the wildcard is used.

3.2.2 Wildcard with upper bound

To allow combining wildcards with static knowledge, we might view $?$ merely as an unknown conjunctive term within a formula:

$$\tilde{\phi} ::= \phi \mid \phi \wedge ?$$

3 Gradualization of a Statically Verified Language

We pose $? \stackrel{\text{def}}{=} \text{true} \wedge ?$.

We expect $\phi \wedge ?$ to be a placeholder for a formula that also “contains” ϕ . There are two ways to express this containment, resulting in different concretizations.

Syntactic $\gamma_1(\phi \wedge ?) = \{ \phi \wedge \phi' \mid \phi' \in \text{SATFORMULA} \}$

Semantic $\gamma_2(\phi \wedge ?) = \{ \phi' \in \text{SATFORMULA} \mid \phi' \xRightarrow[\phi]{} \phi \}$

Lemma 3.2.2. $\forall \tilde{\phi} \in \text{GFORMULA}. \gamma_1(\tilde{\phi}) \subseteq \gamma_2(\tilde{\phi})$

Lemma 3.2.3. $\forall \tilde{\phi} \in \text{GFORMULA}. \gamma_1(\tilde{\phi}) = \gamma_2(\tilde{\phi}) \text{ modulo equivalence}$

Note that $\gamma_1(?) = \gamma_2(?) = \text{SATFORMULA}$, meaning that this approach of extending the formula syntax is compatible with (but superior to) the approach introduced in the previous section.

3.2.3 Precision

Comparing gradual formulas (e.g. $\mathbf{x} = 3$, $x = 3 \wedge ?$, $?$) gives rise to a notion of “precision”. Intuitively, $\mathbf{x} = 3$ is more precise than $x = 3 \wedge ?$ which is still more precise than $?$. Using concretization, we can formalize this intuition.

Definition 3.2.4 (Formula Precision).

$$\tilde{\phi}_a \sqsubseteq \tilde{\phi}_b \iff \gamma(\tilde{\phi}_a) \subseteq \gamma(\tilde{\phi}_b)$$

Read: Formula $\tilde{\phi}_a$ is “at least as precise as” $\tilde{\phi}_b$.

The strict version \sqsubset is defined accordingly.

With the notion of precision, we can give a formal definition of the gradual guarantee (TODO: ref) that we are aiming to satisfy.

Definition 3.2.5 (Gradual Guarantee (for Gradual Verification Systems)).

3.2.4 UNSORTED

PROBABLY UNNECESSARY:

Because of its generality, we will pursue the approach introduced in section 3.2.2 for the remainder of this chapter. As concretization we chose the semantic version, as it is more flexible than the syntactic one in practice. For reference, the full definitions:

Syntax:

$$\tilde{\phi} ::= \phi \mid \phi \wedge ?$$

Concretization:

$$\begin{aligned} \gamma(\phi) &= \{ \phi \} \quad \forall \phi \in \text{SATFORMULA} \\ \gamma(\phi \wedge ?) &= \{ \phi' \in \text{SATFORMULA} \mid \phi' \xRightarrow[\phi]{} \phi \} \\ \gamma(\tilde{\phi}) &= \emptyset \quad \text{otherwise} \end{aligned}$$

3.3 Gradual Statements

Formulas play a role for some statements, extending their syntax may thus also affect the syntax of statements. A common example where formulas are even part of the syntax is the assertion statement `assert ϕ` . Having a gradual formula syntax available does not necessary mean that all statements have to adopt it. In case of the assertion statement there might be little benefit in allowing gradual formulas.

A more complex example affected by gradualization of formulas is a call statement `m()`; in presence of method contracts. Although not directly visible, this statement's semantics (static and dynamic) is affected by the contract of m , consisting of pre- and postcondition. One can think of m as a reference to some method definition including method contract. Note that in practice such method definitions usually reside in some "program context" that is then passed to static and dynamic semantics. As the full meaning of such a statement is unknown without context, it is hard to reason about it abstractly. W.l.o.g. we will thus think of m as syntactic sugar for

```
assert  $\phi_{m_{pre}}$ ;
// body of m
assume  $\phi_{m_{post}}$ ;
```

As one of the main goals of gradual verification is to allow for gradual method contracts, it makes sense to extend the syntax accordingly. This means that the syntax of our desugared call statement is affected:

```
assert  $\widetilde{\phi_{m_{pre}}}$ ;
// body of m
assume  $\widetilde{\phi_{m_{post}}}$ ;
```

In general, statement syntax is extended, resulting in a superset $\text{GSTMT} \supseteq \text{STMT}$ of gradual statements. Note that the superset is induced merely by allowing GFORMULA instead of FORMULA in certain places (chosen freely by the gradual language designer). We give meaning to gradual statements using a concretization function.

Definition 3.3.1 (Concretization of Gradual Statements). *Let $\gamma_s : \text{GSTMT} \rightarrow \mathcal{P}(\text{STMT})$ be defined as*

$$\gamma_s(\widetilde{s}) = \{ s \in \text{STMT} \mid s \text{ is } \widetilde{s} \text{ with all gradual formulas replaced by some concretizations} \}$$

Definition 3.3.2 (Precision of Gradual Statement). *Let $\sqsubseteq_s \subseteq \text{GSTMT} \times \text{GSTMT}$ be a predicate defined as*

$$\widetilde{s}_a \sqsubseteq_s \widetilde{s}_b \iff \gamma_s(\widetilde{s}_a) \subseteq \gamma_s(\widetilde{s}_b)$$

The notion of gradual statements will become important for gradual static and dynamic semantics.

3.4 Gradual Program State

...continuation...

Therefore the introduction of gradual statements GSTMT leads to a notion of gradual program states $\text{GPROGRAMSTATE} \supseteq \text{PROGRAMSTATE}$. TODO: GPROGRAMSTATE_s

Again, we give meaning to gradual program states using concretization.

3 Gradualization of a Statically Verified Language

Definition 3.4.1 (Concretization of Gradual Program States). *Let $\gamma_\pi : \text{GPROGRAMSTATE} \rightarrow \mathcal{P}(\text{PROGRAMSTATE})$ be defined as*

$$\gamma_\pi(\tilde{\pi}) = \{ \pi \in \text{PROGRAMSTATE} \mid \pi \text{ is } \tilde{\pi} \text{ with all continuations??? replaced by a concretization} \}$$

Definition 3.4.2 (Precision of Gradual Program States). *Let $\sqsubseteq_\pi \subseteq \text{GPROGRAMSTATE} \times \text{GPROGRAMSTATE}$ be a predicate defined as*

$$\tilde{\pi}_a \sqsubseteq_\pi \tilde{\pi}_b \iff \gamma_\pi(\tilde{\pi}_a) \subseteq \gamma_\pi(\tilde{\pi}_b)$$

Consequence:

$$\forall \tilde{\pi}_s \in \text{GPROGRAMSTATE}_{\tilde{s}}, \pi \in \gamma_\pi(\tilde{\pi}_s). \exists s \in \gamma_s(\tilde{s}). \pi \in \text{PROGRAMSTATE}_s$$

We demand that formula semantics are not affected by this extension, which is trivially the case if evaluation does not depend on the remaining work in the first place. Formally:

$$\forall \phi \in \text{FORMULA}, \tilde{\pi} \in \text{GPROGRAMSTATE}, \pi \in \gamma_\pi(\tilde{\pi}). \tilde{\pi} \models \phi \iff \pi \models \phi$$

3.5 Lifting Predicates and Functions

The Hoare logic of our language are defined in terms of predicates and functions that operate on formulas. Examples:

After introducing and giving meaning to gradual formulas, we will now describe how to redefine existing predicates and functions in order for them to deal with gradual formulas.

Definition 3.5.1 (Gradual Lifting). *The process of extending an existing predicate/function in order to deal with gradual formulas. The resulting predicate/function has the same signature as the original one, with occurrences of FORMULA replaced by GFORMULA.*

Lifted predicates and functions are not allowed to deal with gradual formulas arbitrarily but must do so in a sound way. What soundness means is a direct consequence of the gradual guarantee (definition 3.2.5), i.e. an unsound predicate/function may cause the gradual verification system to break the gradual guarantee. This idea is formalized in the following sections.

3.5.1 Lifting Predicates

In this section, we assume that we are dealing with a binary predicate $P \subseteq \text{FORMULA} \times \text{FORMULA}$. The concepts are directly applicable to predicates with different arity or with additional non-formula parameters. The lifted version we are targeting has signature $\tilde{P} \subseteq \text{GFORMULA} \times \text{GFORMULA}$. W.l.o.g. we further assume that P appears unnegated in the axiomatic semantics (otherwise we simply regard the negation of that predicate as P).

Rules emerging from the gradual guarantee:

Introduction

Having source code that is considered valid by the static verification system, the same source code must be considered valid by the gradual verification system. In

other words, switching to the gradual system may never “break the code”. This means that arguments satisfying P must satisfy \tilde{P} :

$$\frac{P(\phi_1, \phi_2)}{\tilde{P}(\phi_1, \phi_2)} \text{GPREDINTRO}$$

Or equivalently, using set notation

$$P \subseteq \tilde{P}$$

Monotonicity

A central point of a gradual verification system is enabling programmers to specify contracts with less precision. Source code that is rejected by the verifier might get accepted after reducing precision. If the opposite would happen, though, that would be highly counter-intuitive and ...??? workflow. To prevent such behavior, we expect satisfied predicates to still be satisfied after reducing the precision of arguments:

$$\frac{\tilde{P}(\widetilde{\phi_1}, \widetilde{\phi_2}) \quad \widetilde{\phi_1} \sqsubseteq \widetilde{\phi'_1} \quad \widetilde{\phi_2} \sqsubseteq \widetilde{\phi'_2}}{\tilde{P}(\widetilde{\phi'_1}, \widetilde{\phi'_2})} \text{GPREDMON}$$

or equivalently, thinking of predicates as boolean functions

$$\tilde{P} \text{ is monotonic w.r.t. } \sqsubseteq$$

or something with set terminology!???

$$\tilde{P} \text{ is somewhat closed under weakening}$$

Definition 3.5.2 (Sound Predicate Lifting). *A lifted predicate is **sound/valid** if it is closed under the above rules.*

Note that the rules for sound lifting only give a lower bound for the predicate. Thus $\tilde{P} = \text{GFORMULA} \times \text{GFORMULA}$ is a sound predicate lifting of any binary predicate $P \subseteq \text{FORMULA} \times \text{FORMULA}$.

Definition 3.5.3 (Optimal Predicate Lifting). *A sound lifted predicate is **optimal** if it is the smallest set closed under the above rules.*

This definition coincides with the definition of consistent predicate lifting in AGT:

Lemma 3.5.4 (Optimal Predicate Lifting (Direct Definition)). *Let $\tilde{P} \subseteq \text{GFORMULA} \times \text{GFORMULA}$ be defined as*

$$\tilde{P}(\widetilde{\phi_1}, \widetilde{\phi_2}) \stackrel{\text{def}}{\iff} \exists \phi_1 \in \gamma(\widetilde{\phi_1}), \phi_2 \in \gamma(\widetilde{\phi_2}). P(\phi_1, \phi_2)$$

Then \tilde{P} is an optimal lifting of P .

Consistent lifting of common predicates:

3 Gradualization of a Statically Verified Language

Lemma 3.5.5 (Consistent Lifting of Evaluation).

Let $\cdot \tilde{\models} \cdot \subseteq \text{PROGRAMSTATE} \times \text{GFORMULA}$ be defined as

$$\pi \tilde{\models} \tilde{\phi} \stackrel{\text{def}}{\iff} \pi \models \text{static}(()\phi)$$

Then $\cdot \tilde{\models} \cdot$ is a consistent lifting of $\cdot \models \cdot$.

We define $\text{SATGFORMULA} = \{ \tilde{\phi} \in \text{GFORMULA} \mid \exists \pi. \pi \tilde{\models} \tilde{\phi} \}$ as the set of satisfiable gradual formulas.

Lemma 3.5.6 (Restricted Domain of Concretization).

$\gamma|_{\text{SATGFORMULA}}$ never returns the empty set.

Lemma 3.5.7 (Consistent Lifting of Implication).

Let $\cdot \xRightarrow[\phi]{} \cdot \subseteq \text{GFORMULA} \times \text{GFORMULA}$ be defined as

$$\tilde{\phi}_1 \xRightarrow[\phi]{} \tilde{\phi}_2 \stackrel{\text{def}}{\iff} \exists \phi_1 \in \gamma(\tilde{\phi}_1), \phi_2 \in \gamma(\tilde{\phi}_2). \phi_1 \xRightarrow[\phi]{} \phi_2$$

Then $\cdot \xRightarrow[\phi]{} \cdot$ is a consistent lifting of $\cdot \xRightarrow[\phi]{} \cdot$.

3.5.2 Lifting Functions

Static verification rules may contain functions manipulating formulas. We can also derive rules for lifting such functions from the gradual guarantee. In this section, we assume that we are dealing with a function $f : \text{FORMULA} \rightarrow \text{FORMULA}$. Again, the concepts are directly applicable to functions with higher arity.

Restrictions imposed by the gradual guarantee:

Introduction

We ensure that our verification system is “immune” to reduction of precision. Thus, when passing a static formula ϕ to \tilde{f} , the result must be the same or less precise than $f(\phi)$.

$$\forall \phi \in \text{FORMULA}. f(\phi) \sqsubseteq \tilde{f}(\phi)$$

Monotonicity

Reducing precision of a parameter may only result in a loss of precision of the result. In other words, the function must be monotonic w.r.t. \sqsubseteq (in every argument).

$$\forall \tilde{\phi}_1, \tilde{\phi}_2 \in \text{GFORMULA}. \tilde{\phi}_1 \sqsubseteq \tilde{\phi}_2 \implies \tilde{f}(\tilde{\phi}_1) \sqsubseteq \tilde{f}(\tilde{\phi}_2)$$

Definition 3.5.8 (Sound Function Lifting). A lifted function is **sound/valid** if it adheres to the above rules.

Note that the rules for sound lifting only give a lower bound for the gradual return values. Thus a function $\tilde{f} : \text{GFORMULA} \rightarrow \text{GFORMULA}$ constantly returning $?$ is a sound lifting of any function $f : \text{FORMULA} \rightarrow \text{FORMULA}$.

Definition 3.5.9 (Optimal Function Lifting). A sound lifted function is **optimal** if its return values are as precise as possible.

This definition coincides with the definition of consistent function lifting in AGT:

Lemma 3.5.10 (Optimal Function Lifting (Direct Definition)).

Let $\alpha : \mathcal{P}(\text{SATFORMULA}) \rightarrow \text{GFORMULA}$ be a partial function such that $\langle \gamma, \alpha \rangle$ is a $\{f\}$ -partial Galois connection.

Let $\tilde{f} : \text{GFORMULA} \rightarrow \text{GFORMULA}$ be defined as

$$\tilde{f}(\tilde{\phi}) \stackrel{\text{def}}{=} \alpha(\bar{f}(\gamma(\tilde{\phi})))$$

Then \tilde{f} is an optimal lifting of f .

Examples

$$\alpha(\bar{\phi}) = \min_{\sqsubseteq} \{ \tilde{\phi} \mid \bar{\phi} \subseteq \gamma(\tilde{\phi}) \}$$

The logical and operator $\cdot \wedge \cdot$ of our formula syntax can be viewed as a binary function on formulas.

$$\tilde{f}(\tilde{\phi}_1, \tilde{\phi}_2) = \alpha(\{ \phi_1 \wedge \phi_2 \mid \phi_1 \in \gamma(\tilde{\phi}_1) \wedge \phi_2 \in \gamma(\tilde{\phi}_2) \})$$

PARTIAL:

$$\forall \phi \in \text{FORMULA} \cap \text{dom}(f). f(\phi) \sqsubseteq \tilde{f}(\phi)$$

$$\forall \tilde{\phi}_1, \tilde{\phi}_2 \in \text{GFORMULA}. \tilde{\phi}_1 \sqsubseteq \tilde{\phi}_2 \wedge \tilde{\phi}_1 \in \text{dom}(\tilde{f}) \implies \tilde{f}(\tilde{\phi}_1) \sqsubseteq \tilde{f}(\tilde{\phi}_2)$$

3.5.3 Generalized Lifting

The previous sections describe how lifting is performed in order to deal with GFORMULA instead of FORMULA. In general, the same rules apply to any gradual extension of an existing set that comes with a concretization function.

Example: The signature of Hoare rules contain STMT and can therefore be lifted w.r.t. this parameter using the definitions in section 3.3.

3.6 Abstracting Static Semantics

With the rules for lifting set up we can apply them to the static verification predicate: Lifting

$$\vdash \{\cdot\} \cdot \{\cdot\} \subseteq \text{FORMULA} \times \text{STMT} \times \text{FORMULA}$$

w.r.t. all parameters yields

$$\tilde{\vdash} \{\cdot\} \cdot \{\cdot\} \subseteq \text{GFORMULA} \times \text{GSTMT} \times \text{GFORMULA}$$

Optimality discussion:

```
{i = 10000}
n = collatzIterations(300, i);
{1 <= n * n <= 4}
{n = 4}
staticAssert (n = 4);
{n = 4}
```

...not verifiable with optimal lifting!

3.6.1 The Problem with a Predicate Lifting

As seen in section 3.7, the lifted Hoare predicate in general requires an additional assertion to guarantee preservation. Yet, there is a more fundamental design issue connected to the gradual lifting approach which we will illustrate in this section.

...rule-wise lifting yields overall lifting... neat.

Problem: non-deterministic! Compiler has to find “good” intermediate formulas

too weak could always choose ?

too strong could choose stuff that is not guaranteed by runtime... (so: inject runtime assertions? yes: could be wrong! no: could enter method violating precondition)

3.6.2 The Deterministic Approach

The approach we propose is based on the idea to treat the Hoare predicate as a (multivalued) function, mapping preconditions to the set of possible/verifiable postconditions. We can obtain a lifted version of this hypothetical construct and demand certain properties similar to the ones defined in section ??:

Definition 3.6.1 (Deterministic Lifting). *Given a binary predicate $P \subseteq \text{FORMULA} \times \text{FORMULA}$ we call a partial function $\vec{P} : \text{FORMULA} \rightarrow \text{FORMULA}$ **deterministic lifting** of P if the following conditions are met:*

Introduction

$$\forall (\phi_1, \phi_2) \in P. \phi_1 \in \text{dom}(\vec{P})$$

Preservation

$$\forall \widetilde{\phi}_1, \widetilde{\phi}_2 \in \text{GFORMULA}. \vec{P}(\widetilde{\phi}_1) = \widetilde{\phi}_2$$

$$\implies$$

$$\forall \phi_1 \in \gamma(\widetilde{\phi}_1), \phi_2 \in \text{FORMULA}. P(\phi_1, \phi_2) \implies \exists \phi \in \gamma(\widetilde{\phi}_2). P(\phi_1, \phi) \wedge \phi \xRightarrow{\phi} \phi_2$$

Monotonicity

Note: Identical to monotonicity condition of lifted partial functions.

$$\forall \widetilde{\phi}_1, \widetilde{\phi}_2 \in \text{GFORMULA}. \widetilde{\phi}_1 \sqsubseteq \widetilde{\phi}_2 \wedge \widetilde{\phi}_1 \in \text{dom}(\vec{P}) \implies \vec{P}(\widetilde{\phi}_1) \sqsubseteq \vec{P}(\widetilde{\phi}_2)$$

...assume we have obtained deterministic lifting $\vec{P} \vdash \{\cdot\} \cdot \{\cdot\}$ of our Hoare triple. This gradual partial function has desirable properties:

Obtaining a Sound Gradual Lifting

Lemma 3.6.2 (Deterministic Gradual Lifting).

Let \vec{P} be a deterministic lifting of P . Then

$$\tilde{P}(\widetilde{\phi}_1, \widetilde{\phi}_2) \stackrel{\text{def}}{\iff} \exists \phi'_2. \vec{P}(\widetilde{\phi}_1) = \widetilde{\phi'_2} \wedge \widetilde{\phi'_2} \xRightarrow{\phi} \widetilde{\phi}_2$$

is a sound gradual lifting of P .

Determinism

A verifier dealing with deterministic liftings has no more obligation of finding good intermediate formulas.

Preservation

A (gradual) postcondition returned by the lifted function is guaranteed to reflect the execution state after executing the statements in question (given that the precondition was met). Almost. Combines all the knowledge of static rules.

Composability

Lemma 3.6.3 (Composability of Deterministic Lifting).

Let \vec{P}_1, \vec{P}_2 be deterministic liftings of predicates P_1, P_2 . Then

$$\vec{P}_3 \stackrel{\text{def}}{=} \vec{P}_2 \circ \vec{P}_1$$

is a deterministic lifting of $P_3(\phi_1, \phi_3) = \exists \phi_2. P_1(\phi_1, \phi_2) \wedge P_2(\phi_2, \phi_3)$.

3.7 Gradual Soundness vs Gradual Guarantee

Valid Hoare triples for gradual system

$$\begin{aligned} \widetilde{\models} \{ \cdot \} \cdot \{ \cdot \} &\subseteq \text{GFORMULA} \times \text{GSTMT} \times \text{GFORMULA} \\ \widetilde{\models} \{ \widetilde{\phi}_{pre} \} \widetilde{s} \{ \widetilde{\phi}_{post} \} &\stackrel{\text{def}}{\iff} \forall \langle \widetilde{\pi}_{pre}, \widetilde{\pi}_{post} \rangle \in \widetilde{\mathcal{S}}^s. \widetilde{\pi}_{pre} \widetilde{\models} \widetilde{\phi}_{pre} \implies \widetilde{\pi}_{post} \widetilde{\models} \widetilde{\phi}_{post} \end{aligned}$$

(Note: NOT A sound gradual lifting! Would accept $\widetilde{\models} \{ ? \} x := 3 \{ (y = 4) \wedge ? \}$)

Soundness of gradual system:

$$\begin{aligned} &\frac{\widetilde{\pi} \in \widetilde{\text{wsp}}(\widetilde{s}_1)}{\exists n \in \mathbb{N}_+, \widetilde{s}_2 \in \text{GSTMT}. \widetilde{\mathcal{S}}^n(\widetilde{\pi}) \in \text{PROGRAMSTATE}_{\widetilde{s}_2}} \text{GPROGRESS} \\ &\frac{\widetilde{\vdash} \{ \widetilde{\phi}_1 \} \widetilde{s} \{ \widetilde{\phi}_2 \}}{\widetilde{\models} \{ \widetilde{\phi}_1 \} \widetilde{s} \{ \widetilde{\phi}_2 \}} \text{GPRESERVATION} \end{aligned}$$

Gradual guarantee: Let $\widetilde{\vdash} \{ \cdot \} \cdot \{ \cdot \}$ be gradual lifting of $\vdash \{ \cdot \} \cdot \{ \cdot \}$. Then:

$$\begin{aligned} &\vdash \{ (x = 2) \} y := 3 \{ (x = 2) \wedge (y = 3) \} \\ \xRightarrow{\text{Introduction}} &\widetilde{\vdash} \{ (x = 2) \} y := 3 \{ (x = 2) \wedge (y = 3) \} \\ \xRightarrow{\text{Monotonicity}} &\widetilde{\vdash} \{ ? \} y := 3 \{ (x = 2) \wedge (y = 3) \} \end{aligned}$$

Preservation is obviously not satisfied!

Reiteration:

$$\begin{aligned} &\frac{\vdash \{ \phi_1 \} s \{ \phi_2 \}}{\models \{ \phi_1 \} s; \text{assert } \phi_2 \{ \phi_2 \}} \text{PRESERVATION}' \\ &\frac{\widetilde{\vdash} \{ \widetilde{\phi}_1 \} \widetilde{s} \{ \widetilde{\phi}_2 \}}{\widetilde{\models} \{ \widetilde{\phi}_1 \} \widetilde{s}; \text{assert } \widetilde{\phi}_2 \{ \widetilde{\phi}_2 \}} \text{GPRESERVATION}' \end{aligned}$$

TODO: more bla, like “there is fundamentally no way around this - the programmer can specify postconditions that...”

3.8 Abstracting Dynamic Semantics

Let $\tilde{\mathcal{S}}$ be gradual lifting of \mathcal{S} .

Progress: Note that premise is tautology. So we artificially make conclusion true by demanding that lifting is total. This always works since the lifting can be defined arbitrarily wherever the original function is undefined.

Preservation: Conclusion is already a tautology. This is not really satisfying: An arbitrary verification predicate would satisfy this kind of preservation. Also, this is no guarantee for all the formulas describing intermediate program states. A stronger notion of preservation gives this guarantee:

$$\frac{\vec{\vdash} \{\tilde{\phi}_1\} \tilde{s} \{\tilde{\phi}_2\}}{\tilde{\vDash} \{\tilde{\phi}_1\} \tilde{s} \{\tilde{\phi}_2\}} \text{GPRESERVATION}$$

Making this guarantee work is trickier and there are different trade-offs available. Without further assumptions, $\vec{\vdash} \{\cdot\} \cdot \{\cdot\}$ is not a subset of $\tilde{\vDash} \{\cdot\} \cdot \{\cdot\}$.

Running example:

$$\vec{\vdash} \{?\} \text{ assert } (x = 3) \{(x = 3) \wedge ?\}$$

holds but not

$$\tilde{\vDash} \{?\} \text{ assert } (x = 3) \{(x = 3) \wedge ?\}$$

So far, our definition of $\tilde{\mathcal{S}}$ as a total lifting of \mathcal{S} may be too weak, breaking the subset relationship:

\mathcal{S} too weak It is possible that the dynamic semantics of SVL defines

$$\mathcal{S}^{\text{assert } (x = 3)}(\pi_{(x = 4)}) = \pi'_{(x = 4)}$$

This is not unreasonable, since this function is guaranteed to be only called with “valid” program states in the static system! An additional runtime check would be overhead.

$\tilde{\mathcal{S}}$ too weak If $\mathcal{S}^{\text{assert } (x = 3)}(\pi_{(x = 4)})$ is undefined due to runtime checks. Yet, the lifting is supposed to be total, so passing along the program state unchecked is again a valid realization:

$$\tilde{\mathcal{S}}^{\text{assert } (x = 3)}(\pi_{(x = 4)}) = \pi'_{(x = 4)}$$

Mapping to an exception would have been better in this case.

Note that both problems are unrelated to optimality of the lifting.

3.8.1 Perfect Knowledge

Choose $\tilde{\mathcal{S}} : \text{GPROGRAMSTATE} \rightarrow \text{GPROGRAMSTATE}$ as lifted version of $\mathcal{S} : \text{PROGRAMSTATE} \rightarrow \text{PROGRAMSTATE}$ with $\tilde{\mathcal{S}}(\tilde{\pi}) = \pi_{EX}$ if stuck for all concretizations.

$$\text{wsp} : \text{GSTMT} \rightarrow \text{PROGRAMSTATE}$$

$$\text{wsp}(\tilde{s}) \stackrel{\text{def}}{=} \bigcup_{s \in \gamma_s(\tilde{s})} \text{wsp}(s)$$

$$\forall \tilde{s} \in \text{GSTMT}.$$

$$\tilde{\pi}_s \in \text{GPROGRAMSTATE}_{\tilde{s}}. \text{wsp}(\tilde{s}) \cap \gamma_{\pi}(\tilde{\pi}_s) = \emptyset \implies \tilde{\mathcal{S}}^s(\tilde{\pi}_s) = \pi_{EX}$$

MINUS: - need above knowledge... - not always desirable

```
{i = 10000}
n = collatzIterations(300, i);
{1 <= n * n <= 4}
{n = 4}
staticAssert (n = 4);
{n = 4}
```

would throw exception!?

Proof:

$$\tilde{s} \in \text{GSTMT}$$

$$\tilde{\phi}_1, \tilde{\phi}_2 \in \text{GFORMULA}$$

$$\tilde{\pi}_1, \tilde{\pi}_2 \in \text{GPROGRAMSTATE}$$

$$1 = \text{Premise} \quad \vec{\vdash} \{\tilde{\phi}_1\} \tilde{s} \{\tilde{\phi}_2\}$$

$$2 = \text{HoareIntrosA} \quad \tilde{\mathcal{S}}^s(\tilde{\pi}_1, \tilde{\pi}_2)$$

$$3 = \text{HoareIntrosB} \quad \tilde{\pi}_1 \tilde{\models} \tilde{\phi}_1$$

$$4 = \text{Case} \quad \exists \pi_s \in \gamma(\tilde{\pi}_1). \pi_s \in \text{wsp}(s)$$

$$5 = 4 + \text{wsp def} \quad \exists \phi'_1, \phi' \in \text{FORMULA}. \pi_s \models \phi'_1 \wedge \vdash \{\phi'_1\} s \{\phi'\}$$

$$6 = 4 + 5 + \text{rule42} \quad \exists \phi_1 \in \gamma(\tilde{\phi}_1). \phi_1 \xRightarrow{\phi} \phi'_1 \wedge \pi_s \models \phi_1$$

$$7 = 5 + 6 + \text{mono} \quad \exists \phi \in \text{FORMULA}. \vdash \{\phi_1\} s \{\phi\}$$

$$8 = 7 + \text{intro} \quad \exists \tilde{\phi} \in \text{GFORMULA}. \vec{\vdash} \{\phi_1\} s \{\tilde{\phi}\}$$

$$9 = 1 + 6 + 8 + \text{mono_det_hoare} \quad \tilde{\phi} \sqsubseteq \tilde{\phi}_2$$

$$10 = 8 + \text{pres} \quad \exists \phi_2 \in \gamma(\tilde{\phi}). \vdash \{\phi_1\} s \{\phi_2\}$$

$$11 = 6 + 10 + \text{snd} \quad \mathcal{S}^s(\pi_s) \models \phi_2$$

$$12 = 11 + \text{intro} \quad \tilde{\mathcal{S}}^s(\pi_s) \models \phi_2$$

$$13 = 3 + 12 + \text{mono} \quad \tilde{\mathcal{S}}^s(\tilde{\pi}_s) \models \phi_2$$

$$14 = 13 + \text{intro} \quad \tilde{\mathcal{S}}^s(\tilde{\pi}_s) \tilde{\models} \phi_2$$

$$15 = 10 + 14 + \text{mono} \quad \tilde{\mathcal{S}}^s(\tilde{\pi}_s) \tilde{\models} \tilde{\phi}$$

$$16 = 9 + 15 + \text{mono} \quad \tilde{\mathcal{S}}^s(\tilde{\pi}_s) \tilde{\models} \tilde{\phi}_2$$

3 Gradualization of a Statically Verified Language

$\tilde{s} \in \text{GSTMT}$

$\widetilde{\phi_1}, \widetilde{\phi_2} \in \text{GFORMULA}$

$\widetilde{\pi_s} \in \text{GPROGRAMSTATE}_{\tilde{s}}$

1 = PremiseA $\vec{\vdash} \{\widetilde{\phi_1}\} \tilde{s} \{\widetilde{\phi_2}\}$

2 = PremiseB $\widetilde{\pi_s} \tilde{\vDash} \widetilde{\phi_1}$

3 = Case $\neg \exists \pi_s \in \gamma(\widetilde{\pi_s}). \pi_s \in \text{wsp}(s)$

4 = 3 + completeness $\forall \pi_s \in \gamma(\widetilde{\pi_s}). \mathcal{S}^s(\pi_s) \text{ stuck}$

5 = 4 + def $\widetilde{\mathcal{S}}^s(\widetilde{\pi_s}) = \pi_{EX}$

6 = 5 + precision $\widetilde{\mathcal{S}}^s(\widetilde{\pi_s}) \tilde{\vDash} \widetilde{\phi_2}$

3.8.2 Partial Knowledge

wsp not always known, think of sequence operator. Turns out we don't need it for sequence operator. Assume approach of previous section, but not for sequences. Preservation still holds:

$$\begin{array}{c}
 \vec{\vdash} \{\widetilde{\phi_1}\} \tilde{s}_1; \tilde{s}_2 \{\widetilde{\phi_3}\} \\
 \hline
 \vec{\vdash} \{\widetilde{\phi_1}\} \tilde{s}_1 \{\widetilde{\phi_2}\} \quad \vec{\vdash} \{\widetilde{\phi_2}\} \tilde{s}_2 \{\widetilde{\phi_3}\} \quad \text{INVERSION} \\
 \hline
 \tilde{\vDash} \{\widetilde{\phi_1}\} \tilde{s}_1 \{\widetilde{\phi_2}\} \quad \tilde{\vDash} \{\widetilde{\phi_2}\} \tilde{s}_2 \{\widetilde{\phi_3}\} \quad \text{GSOUNDNESS} \\
 \hline
 \tilde{\vDash} \{\widetilde{\phi_1}\} \tilde{s}_1; \tilde{s}_2 \{\widetilde{\phi_3}\} \quad \text{SEQ}
 \end{array}$$

4 Case Study: Implicit Dynamic Frames

4.1 Language

We now introduce a simplified Java-like statically verified language SVLidf that uses Chalice/Eiffel/Spec# sub-syntax to express method contracts.

4.1.1 Syntax

$program \in \text{PROGRAM}$	$::= \overline{cls} \ s$
$cls \in \text{CLASS}$	$::= \text{class } C \{ \overline{field} \ \overline{method} \}$
$field \in \text{FIELD}$	$::= T \ f;$
$method \in \text{METHOD}$	$::= T \ m(T \ x) \ \text{contract} \{ s \}$
$contract \in \text{CONTRACT}$	$::= \text{requires } \phi; \text{ ensures } \phi;;$
$T \in \text{TYPE}$	$::= \text{int} \mid C$
$s \in \text{STMT}$	$::= \text{skip} \mid T \ x \mid x.f := y \mid x := e \mid x := \text{new } C \mid x := y.m(z) \\ \mid \text{return } x \mid \text{assert } \phi \mid \text{release } \phi \mid \text{hold } \phi \{ s \} \mid s_1; s_2$
$\phi \in \text{FORMULA}$	$::= \text{true} \mid (e = e) \mid (e \neq e) \mid \text{acc}(e.f) \mid \phi * \phi$
$e \in \text{EXPR}$	$::= v \mid x \mid e.f$
$x, y, z \in \text{VAR}$	$::= \text{this} \mid \text{result} \mid \text{name}$
$v \in \text{VAL}$	$::= o \mid n \mid \text{null}$
$o \in \text{LOC}$	
$n \in \mathbb{Z}$	
$C \in \text{CLASSNAME}$	$::= \text{name}$
$f \in \text{FIELDNAME}$	$::= \text{name}$
$m \in \text{METHODNAME}$	$::= \text{name}$

Figure 4.1. SVL: Syntax

We pose $\text{false} \stackrel{\text{def}}{=} (\text{null} \neq \text{null})$.

$$\boxed{\lfloor \phi \rfloor_{H,\rho} = A_d}$$

$$\begin{aligned} \lfloor \text{true} \rfloor_{H,\rho} &= \emptyset \\ \lfloor (e_1 = e_2) \rfloor_{H,\rho} &= \emptyset \\ \lfloor (e_1 \neq e_2) \rfloor_{H,\rho} &= \emptyset \\ \lfloor \text{acc}(x.f) \rfloor_{H,\rho} &= \{(o, f)\} \text{ where } H, \rho \vdash x \Downarrow o \\ \lfloor \phi_1 * \phi_2 \rfloor_{H,\rho} &= \lfloor \phi_1 \rfloor_{H,\rho} \cup \lfloor \phi_2 \rfloor_{H,\rho} \end{aligned}$$

What about undefinedness of acc case? Guess: propagates to undefinedness of small-step rule \Rightarrow covered by soundness

Figure 4.2. SVL: Dynamic Footprint

4.1.2 Program State

The program state of SVL is defined as $\text{PROGRAMSTATE} = \text{HEAP} \times \text{STACK}$ with

$$\begin{aligned} H \in \text{HEAP} &= \text{LOC} \rightarrow (\text{CLASSNAME} \times (\text{FIELDNAME} \rightarrow \text{VAL})) \\ \rho \in \text{VARENV} &= \text{VAR} \rightarrow \text{VAL} \\ \Gamma \in \text{TYPEENV} &= \text{VAR} \rightarrow \text{TYPE} \\ A_s \in \text{STATICFOOTPRINT} &= \mathcal{P}^{\text{EXPR} \times \text{FIELDNAME}} \\ A_d \in \text{DYNAMICFOOTPRINT} &= \mathcal{P}^{\text{LOC} \times \text{FIELDNAME}} \\ E \in \text{STACKENTRY} &= \text{VARENV} \times \text{DYNAMICFOOTPRINT} \times \text{STMT} \\ S \in \text{STACK} &::= E \cdot S \mid \text{nil} \end{aligned}$$

REQUIRED?

Definition 4.1.1 (Topmost Stack Entry). *Let $\text{topmost} : \text{STACK} \rightarrow \text{STACKENTRY}$ be defined as*

$$\begin{aligned} \text{topmost}(E \cdot S) &= E \\ \text{topmost}(\text{nil}) &\text{ undefined} \end{aligned}$$

Program states with scheduled statement s are defined as

$$\text{PROGRAMSTATE}_s \stackrel{\text{def}}{=} \text{HEAP} \times \{ (\rho, A_d, s) \cdot S \mid \rho \in \text{VARENV}, A_d \in \text{DYNAMICFOOTPRINT}, S \in \text{STACK} \}$$

4.1.3 Formula Semantics

Framing

SVL uses the concepts of implicit dynamic frames to ensure that a statement can only access memory locations (more specifically: fields) which it is guaranteed to have exclusive access to. This is achieved by explicitly tracking access tokens $\text{acc}(\langle \text{expression} \rangle. \langle \text{field} \rangle)$ as part of formulas throughout the entire program during verification.

The Hoare rules of SVL also make sure that access is never duplicated within or across stack frames, effectively ruling out concurrent access to any field during runtime.

Implicit dynamic frames also allows static reasoning about the values of fields during verification, i.e. as part of verification formulas. In order to guarantee that such formulas

$$\boxed{H, \rho \vdash e \Downarrow v}$$

$$\frac{}{H, \rho \vdash x \Downarrow \rho(x)} \text{EEVAR}$$

$$\frac{}{H, \rho \vdash v \Downarrow v} \text{EEVALUE}$$

$$\frac{H, \rho \vdash e \Downarrow o}{H, \rho \vdash e.f \Downarrow H(o)(f)} \text{EEACC}$$

Figure 4.3. SVL: Evaluating Expressions

$$\boxed{H, \rho, A \models \phi}$$

$$\frac{}{H, \rho, A \models \mathbf{true}} \text{EATRUE}$$

$$\frac{H, \rho \vdash e_1 \Downarrow v_1 \quad H, \rho \vdash e_2 \Downarrow v_2 \quad v_1 = v_2}{H, \rho, A \models (e_1 = e_2)} \text{EAEQUAL}$$

$$\frac{H, \rho \vdash e_1 \Downarrow v_1 \quad H, \rho \vdash e_2 \Downarrow v_2 \quad v_1 \neq v_2}{H, \rho, A \models (e_1 \neq e_2)} \text{EANEQUAL}$$

$$\frac{H, \rho \vdash e \Downarrow o \quad H, \rho \vdash e.f \Downarrow v \quad (o, f) \in A}{H, \rho, A \models \mathbf{acc}(e.f)} \text{EAACC}$$

$$\frac{A_1 = A \setminus A_2 \quad H, \rho, A_1 \models \phi_1 \quad H, \rho, A_2 \models \phi_2}{H, \rho, A \models \phi_1 * \phi_2} \text{EASEPOP}$$

Figure 4.4. SVL: Evaluating Expressions

$$\boxed{A_s \vdash_{\text{frm}} e}$$

$$\frac{}{A \vdash_{\text{frm}} x} \text{WFVAR}$$

$$\frac{}{A \vdash_{\text{frm}} v} \text{WFVALUE}$$

$$\frac{(e, f) \in A \quad A \vdash_{\text{frm}} e}{A \vdash_{\text{frm}} e.f} \text{WFFIELD}$$

Figure 4.5. SVL: Framing Expressions

always reflect the program state (preservation), formulas mentioning a certain field must also contain the access token to that very field:

Definition 4.1.2 (Self-Framing). *A formula is **self-framing/self-framed** if it contains access to all fields it mentions.*

We omit the emptyset...

Definition 4.1.3 (Formula Self-Framedness). *A formula ϕ is **self-framed** iff*

$$\vdash_{\text{sfrm}} \phi$$

*Let $\text{SFRMFORMULA} \subseteq \text{SATFORMULA}$ be the set of **self-framed and satisfiable** formulas.*

As illustrated in example??? self-framed formulas are required for race-free verification.

SVL will thus only consider method contracts using self-framed and satisfiable formulas well-formed (see section 4.1.5).

4.1.4 Static Semantics

The static semantics of SVL consist of typing rules and a Hoare calculus making use of those typing rules. All the rules are implicitly parameterized over some program $p \in \text{PROGRAM}$, necessary for example to extract the type of a field in the following typing rules.

$$\boxed{A_s \vdash_{\text{sfrm}} \phi}$$

$$\frac{}{A \vdash_{\text{sfrm}} \text{true}} \text{WFT}_{\text{TRUE}}$$

$$\frac{A \vdash_{\text{frm}} e_1 \quad A \vdash_{\text{frm}} e_2}{A \vdash_{\text{sfrm}} (e_1 = e_2)} \text{WFE}_{\text{EQUAL}}$$

$$\frac{A \vdash_{\text{frm}} e_1 \quad A \vdash_{\text{frm}} e_2}{A \vdash_{\text{sfrm}} (e_1 \neq e_2)} \text{WFNE}_{\text{EQUAL}}$$

$$\frac{A \vdash_{\text{frm}} e}{A \vdash_{\text{sfrm}} \text{acc}(e.f)} \text{WFA}_{\text{ACC}}$$

Figure 4.6. SVL: Framing Formulas

$$\boxed{[\phi] = A_s}$$

$$\begin{array}{ll} [\text{true}] & = \emptyset \\ [(e_1 = e_2)] & = \emptyset \\ [(e_1 \neq e_2)] & = \emptyset \\ [\text{acc}(e.f)] & = \{(e, f)\} \\ [\phi_1 * \phi_2] & = [\phi_1] \cup [\phi_2] \end{array}$$

Figure 4.7. SVL: Static Footprint

$$\boxed{\Gamma \vdash e : T}$$

$$\frac{}{\Gamma \vdash n : \text{int}} \text{STVALNUM}$$

$$\frac{}{\Gamma \vdash \text{null} : C} \text{STVALNULL}$$

$$\frac{\Gamma(x) = T}{\Gamma \vdash x : T} \text{STVAR}$$

$$\frac{\Gamma \vdash e : C \quad \text{fieldType}_p(C, f) = T}{\Gamma \vdash e.f : T} \text{STFIELD}$$

Figure 4.8. SVL: Static Typing of Expressions

Typing

Verification

Let $\text{wsp} : \text{STMT} \rightarrow \mathcal{P}(\text{PROGRAMSTATE})$ be defined as

$$\begin{aligned} \text{wsp}(s) &= \{ \pi \in \text{PROGRAMSTATE}_s \mid \exists \phi_1, \phi_2 \in \text{FORMULA}, \Gamma \in \text{TYPEENV}. \Gamma \vdash \{\phi_1\} s \{\phi_2\} \wedge \pi \models \phi_1 \} \\ \text{wsp}(s) &= \begin{cases} \text{PROGRAMSTATE}_s & \text{if } s = x := \text{new } C \\ \{ \pi \in \text{PROGRAMSTATE}_s \mid \pi \models \text{acc}(x.f) \} & \text{if } s = x.f := y \\ \{ \pi \in \text{PROGRAMSTATE}_s \mid \pi \models \text{acc}(e) \} & \text{if } s = x := e \\ \text{PROGRAMSTATE}_s & \text{if } s = \text{return } x \\ \{ \pi \in \text{PROGRAMSTATE}_s \mid \pi \models (y \neq \text{null}) * \text{mpre}_p(m) \} & \text{if } s = x := y.m(z) \\ \{ \pi \in \text{PROGRAMSTATE}_s \mid \pi \models \phi \} & \text{if } s = \text{assert } \phi \\ \{ \pi \in \text{PROGRAMSTATE}_s \mid \pi \models \phi \} & \text{if } s = \text{release } \phi \end{cases} \end{aligned}$$

4.1.5 Well-Formedness

With static semantics in place, we can define what makes programs well-formed. Well-formedness is required to ... The following predicates

A program is well-formed if both its classes and main method are. For the main method to be well-formed, it must satisfy our Hoare predicate, given no assumptions.

$$\frac{\overline{cls_i \text{ OK}} \quad \vdash \{\text{true}\} s \{\text{true}\}}{(\overline{cls_i} s) \text{ OK}} \text{OKPROGRAM}$$

$$\begin{array}{c}
\frac{\text{unique } \overline{\text{field-names}} \quad \text{unique } \overline{\text{method-names}} \quad \overline{\text{method}_i \text{ OK in } C}}{(\text{class } C \{ \overline{\text{field}_i} \overline{\text{method}_i} \}) \text{ OK}} \text{ OKCLASS} \\
\\
\frac{\begin{array}{c} FV(\phi_1) \subseteq \{x, \text{this}\} \\ FV(\phi_2) \subseteq \{x, \text{this}, \text{result}\} \quad x : T_x, \text{this} : C, \text{result} : T_m \vdash \{\phi_1\} s \{\phi_2\} \\ \phi_1, \phi_2 \in \text{SFRMFORMULA} \quad \neg \text{writesTo}(s, x) \end{array}}{(T_m \ m(T_x \ x) \ \text{requires } \phi_1; \ \text{ensures } \phi_2; \ \{s\}) \text{ OK in } C} \text{ OKMETHOD}
\end{array}$$

4.1.6 Dynamic Semantics

4.1.7 Soundness

4.2 Gradualization

We will now follow along the procedure introduced in chapter 3 to design a gradually verified language “GVL ” based on SVL.

The path we take:

Syntax:

$$\tilde{\phi} ::= \phi \mid ? * \phi$$

Concretization:

$$\begin{aligned}
\gamma(\phi) &= \{ \phi \} \quad \forall \phi \in \text{SFRMFORMULA} \\
\gamma(? * \phi) &= \{ \phi' \in \text{SFRMFORMULA} \mid \phi' \xRightarrow{\phi} \phi \} \\
\gamma(\tilde{\phi}) &= \emptyset \quad \text{otherwise}
\end{aligned}$$

4.2.1 Extension: Statements

In GVL we want the programmer to specify gradual method contracts. Therefore we extend their syntax as follows.

$$\widetilde{\text{contract}} \in \text{GCONTRACT} \quad ::= \text{requires } \tilde{\phi}; \ \text{ensures } \tilde{\phi};$$

This extension is propagated to method declarations (now accepting gradual contracts but not changing otherwise), yielding GMETHOD. Carrying on with the same logic, we get an extended set of class definitions GCLASS and finally an extended set of programs GPROGRAM. Again, note that the only syntactical difference is the acceptance of gradual formulas in method contracts.

We see no motive to extend the syntax of statements themselves and define GSTMT = STMT. As postulated in section 3.3, the call statement hides away gradualized syntax by referencing a method with gradual contract. This becomes obvious when looking at its static or dynamic semantics (see HCALL and ESCALL??/ESCALLFINISH) where the method name is effectively dereferenced.

4.2.2 Extension: Program State

GPROGRAMSTATE = PROGRAMSTATE

4.3 Gradualize Hoare Rules

4.4 Gradual Dyn. Semantics

4.5 Enhancing an Unverified Language

$$\boxed{\Gamma \vdash \{\phi_{pre}\} \text{ } s \text{ } \{\phi_{post}\}}$$

$$\frac{}{\Gamma \vdash \{\phi\} \text{ skip } \{\phi\}} \text{HSKIP}$$

$$\frac{\phi \xRightarrow{\phi} \phi' \quad \vdash_{\text{sfrm}} \phi' \quad x \notin FV(\phi') \quad \Gamma \vdash x : C \quad \text{fields}_p(C) = \overline{T \text{ } f};}{\Gamma \vdash \{\phi\} \text{ } x := \text{new } C \text{ } \{\phi' * (x \neq \text{null}) * \text{acc}(x.f_i) * (x.f_i = \text{defaultValue}(T_i))\}} \text{HALLOC}$$

$$\frac{\phi \xRightarrow{\phi} \text{acc}(x.f) * \phi' \quad \vdash_{\text{sfrm}} \phi' \quad \Gamma \vdash x : C \quad \Gamma \vdash y : T \quad \vdash C.f : T}{\Gamma \vdash \{\phi\} \text{ } x.f := y \text{ } \{\phi' * \text{acc}(x.f) * (x \neq \text{null}) * (x.f = y)\}} \text{HFIELDASSIGN}$$

$$\frac{\phi \xRightarrow{\phi} \text{acc}(e) \quad \vdash_{\text{sfrm}} \phi' \quad \phi \xRightarrow{\phi} \phi' \quad x \notin FV(\phi') \quad x \notin FV(e) \quad \Gamma \vdash x : T \quad \Gamma \vdash e : T}{\Gamma \vdash \{\phi\} \text{ } x := e \text{ } \{\phi' * (x = e)\}} \text{HVARASSIGN}$$

$$\frac{\phi \xRightarrow{\phi} \phi' \quad \vdash_{\text{sfrm}} \phi' \quad \text{result} \notin FV(\phi') \quad \Gamma \vdash x : T \quad \Gamma \vdash \text{result} : T}{\Gamma \vdash \{\phi\} \text{ return } x \text{ } \{\phi' * (\text{result} = x)\}} \text{HRETURN}$$

$$\frac{\begin{array}{l} \Gamma \vdash y : C \quad \text{method}_p(C, m) = T_r \text{ } m(T_p \text{ } z) \text{ requires } \phi_{pre}; \text{ ensures } \phi_{post}; \{ _ \} \\ \Gamma \vdash x : T_r \quad \Gamma \vdash z' : T_p \quad \phi \xRightarrow{\phi} (y \neq \text{null}) * \phi_p * \phi' \quad \vdash_{\text{sfrm}} \phi' \quad x \notin FV(\phi') \\ x \neq y \wedge x \neq z' \quad \phi_p = \phi_{pre}[y, z' / \text{this}, z] \quad \phi_q = \phi_{post}[y, z', x / \text{this}, z, \text{result}] \end{array}}{\Gamma \vdash \{\phi\} \text{ } x := y.m(z') \text{ } \{\phi' * \phi_q\}} \text{HCALL}$$

$$\frac{\phi \xRightarrow{\phi} \phi'}{\Gamma \vdash \{\phi\} \text{ assert } \phi' \text{ } \{\phi\}} \text{HASSERT}$$

$$\frac{\phi \xRightarrow{\phi} \phi_r * \phi' \quad \vdash_{\text{sfrm}} \phi'}{\Gamma \vdash \{\phi\} \text{ release } \phi_r \text{ } \{\phi'\}} \text{HRELEASE}$$

$$\frac{x \notin \text{dom}(\Gamma) \quad \Gamma, x : T \vdash \{(x = \text{defaultValue}(T)) * \phi\} \text{ } s \text{ } \{\phi'\}}{\Gamma \vdash \{\phi\} \text{ } T \text{ } xs \text{ } \{\phi'\}} \text{HDECLARE}$$

$$\frac{\begin{array}{l} \vdash_{\text{sfrm}} \phi \quad \phi_f \xRightarrow{\phi} \phi_r * \phi' \\ \phi' \xRightarrow{\phi} \phi \quad FV(\phi') = FV(\phi) \quad \neg \text{writesTo}(FV(\phi), s) \quad \Gamma \vdash \{\phi_r\} \text{ } s \text{ } \{\phi'_r\} \end{array}}{\Gamma \vdash \{\phi_f\} \text{ hold } \phi \text{ } \{ \text{ } s \text{ } \} \text{ } \{\phi'_r * \phi'\}} \text{HHOLD}$$

$$\boxed{(H, S) \rightarrow (H, S)}$$

$$\frac{}{(H, (\rho, A, \mathbf{skip}) \cdot S) \rightarrow (H, (\rho, A, s) \cdot S)} \text{ESSKIP}$$

$$\frac{H, \rho \vdash x \Downarrow o \quad H, \rho \vdash y \Downarrow v_y \quad (o, f) \in A \quad H' = H[o \mapsto [f \mapsto v_y]]}{(H, (\rho, A, x.f := ys) \cdot S) \rightarrow (H', (\rho, A, s) \cdot S)} \text{ESFIELDASSIGN}$$

$$\frac{H, \rho \vdash e \Downarrow v \quad \rho' = \rho[x \mapsto v]}{(H, (\rho, A, x := es) \cdot S) \rightarrow (H, (\rho', A, s) \cdot S)} \text{ESVARASSIGN}$$

$$\frac{\rho' = \rho[x \mapsto o] \quad o \notin \text{dom}(H) \quad \text{fields}_p(C) = \overline{T \ f}; \quad A' = A \cup \overline{(o, f_i)} \quad H' = H[o \mapsto [f_i \mapsto \text{defaultValue}(T_i)]]}{(H, (\rho, A, x := \mathbf{new} \ Cs) \cdot S) \rightarrow (H', (\rho', A', s) \cdot S)} \text{ESALLOC}$$

$$\frac{H, \rho \vdash x \Downarrow v_x \quad \rho' = \rho[\mathbf{result} \mapsto v_x]}{(H, (\rho, A, \mathbf{return} \ xs) \cdot S) \rightarrow (H, (\rho', A, s) \cdot S)} \text{ESRETURN}$$

$$\frac{H, \rho \vdash y \Downarrow o \quad H, \rho \vdash z \Downarrow v \quad H(o) = (C, _) \quad \text{method}_p(C, m) = T_r \ m(T \ w) \ \mathbf{requires} \ \phi; \ \mathbf{ensures} \ _; \ \{ \bar{r} \} \quad \rho' = [\mathbf{result} \mapsto \text{defaultValue}(T_r), \mathbf{this} \mapsto o, w \mapsto v] \quad H, \rho', A \models \phi \quad A' = \lfloor \phi \rfloor_{H, \rho'}}{(H, (\rho, A, x := y.m(z)s) \cdot S) \rightarrow (H, (\rho', A', \bar{r}) \cdot (\rho, A \setminus A', x := y.m(z)s) \cdot S)} \text{ESCALL}$$

$$\frac{\text{mpost}_p((_)C, m) = \phi \quad H, \rho \vdash y \Downarrow o \quad H(o) = (C, _) \quad H, \rho', A' \models \phi \quad A'' = \lfloor \phi \rfloor_{H, \rho'} \quad H, \rho' \vdash \mathbf{result} \Downarrow v_r}{(H, (\rho', A', \emptyset) \cdot (\rho, A, x := y.m(z)s) \cdot S) \rightarrow (H, (\rho[x \mapsto v_r], A \cup A'', s) \cdot S)} \text{ESCALLFINISH}$$

$$\frac{H, \rho, A \models \phi}{(H, (\rho, A, \mathbf{assert} \ \phi s) \cdot S) \rightarrow (H, (\rho, A, s) \cdot S)} \text{ESASSERT}$$

$$\frac{H, \rho, A \models \phi \quad A' = A \setminus \lfloor \phi \rfloor_{H, \rho}}{(H, (\rho, A, \mathbf{release} \ \phi s) \cdot S) \rightarrow (H, (\rho, A', s) \cdot S)} \text{ESRELEASE}$$

$$\frac{\rho' = \rho[x \mapsto \text{defaultValue}(T)]}{(H, (\rho, A, T \ xs) \cdot S) \rightarrow (H, (\rho', A, s) \cdot S)} \text{ESDECLARE}$$

$$38 \quad \frac{H, \rho, A \models \phi \quad A' = \lfloor \phi \rfloor_{H, \rho}}{(H, (\rho, A, \mathbf{hold} \ \phi \ \{ \bar{s}' \} s) \cdot S) \rightarrow (H, (\rho, A \setminus A', \bar{s}') \cdot (\rho, A', \mathbf{hold} \ \phi \ \{ \bar{s}' \} s) \cdot S)} \text{ESHOLD}$$

$$\boxed{\Gamma \vdash \{\widetilde{\phi}_{pre}\} \ s \ \{\widetilde{\phi}_{post}\}}$$

$$\frac{\widetilde{\phi} \div x = \widetilde{\phi}' \quad \Gamma \vdash x : C \quad \text{fields}_p(C) = \overline{T \ f};}{\Gamma \vec{\vdash} \{\widetilde{\phi}\} \ x := \text{new } C \ \{\widetilde{\phi}' \ * (x \neq \text{null}) \ * \overline{\text{acc}(x.f_i)} \ * (x.f_i = \text{defaultValue}(T_i)) \}} \text{GHALLOC}$$

$$\frac{\widetilde{\phi} \div \text{acc}(x.f) = \widetilde{\phi}' \quad \Gamma \vdash x : C \quad \Gamma \vdash y : T \quad \vdash C.f : T}{\Gamma \vec{\vdash} \{\widetilde{\phi}\} \ x.f := y \ \{\widetilde{\phi}' \ * \text{acc}(x.f) \ * (x \neq \text{null}) \ * (x.f = y)\}} \text{GHFIELDASSIGN}$$

$$\frac{\widetilde{\phi} \xRightarrow[\phi]{\text{acc}(e)} \quad \widetilde{\phi} \div x = \widetilde{\phi}' \quad x \notin FV(e) \quad \Gamma \vdash x : T \quad \Gamma \vdash e : T}{\Gamma \vec{\vdash} \{\widetilde{\phi}\} \ x := e \ \{\widetilde{\phi}' \ * (x = e)\}} \text{GHVARASSIGN}$$

$$\frac{\widetilde{\phi} \div \text{result} = \widetilde{\phi}' \quad \Gamma \vdash x : T \quad \Gamma \vdash \text{result} : T}{\Gamma \vec{\vdash} \{\widetilde{\phi}\} \ \text{return } x \ \{\widetilde{\phi}' \ * (\text{result} = x)\}} \text{GHRETURN}$$

$$\frac{\begin{array}{l} \Gamma \vdash y : C \quad \text{method}_p(C, m) = T_r \ m(T_p \ z) \ \text{requires } \phi_{pre}; \ \text{ensures } \phi_{post}; \ \{ _ \} \\ \Gamma \vdash x : T_r \quad \Gamma \vdash z' : T_p \quad \widetilde{\phi} \xRightarrow[\phi]{(y \neq \text{null}) \ * \widetilde{\phi}_p} \\ x \neq y \wedge x \neq z' \quad \widetilde{\phi}_p = \widetilde{\phi}_{pre}[y, z'/\text{this}, z] \quad \widetilde{\phi}_q = \widetilde{\phi}_{post}[y, z', x/\text{this}, z, \text{result}] \end{array}}{\Gamma \vec{\vdash} \{\widetilde{\phi}\} \ x := y.m(z') \ \{\widetilde{\phi}' \ * \widetilde{\phi}_q\}} \text{GHCALL}$$

$$\frac{\widetilde{\phi}' \vdash \widetilde{\phi} \xRightarrow[\phi]{\phi_a}}{\Gamma \vec{\vdash} \{\widetilde{\phi}\} \ \text{assert } \phi' \ \{\widetilde{\phi}'\}} \text{GHASSERT}$$

$$\frac{\widetilde{\phi}' \vdash \widetilde{\phi} \xRightarrow[\phi]{\phi_r} \quad \widetilde{\phi}' \div [\phi_r] = \widetilde{\phi}''}{\Gamma \vec{\vdash} \{\widetilde{\phi}\} \ \text{release } \phi_r \ \{\widetilde{\phi}'\}} \text{GHRELEASE}$$

$$\frac{x \notin \text{dom}(\Gamma) \quad \Gamma, x : T \vec{\vdash} \{(x = \text{defaultValue}(T)) \ * \widetilde{\phi}\} \ s \ \{\widetilde{\phi}'\}}{\Gamma \vec{\vdash} \{\widetilde{\phi}\} \ T \ x; s \ \{\widetilde{\phi}'\}} \text{GHDECLARE}$$

$$\frac{\begin{array}{l} \vdash_{\text{sfrm}} \phi \quad \widetilde{\phi}'_f \vdash \widetilde{\phi}_f \xRightarrow[\phi]{} \phi \quad \widetilde{\phi}'_f \div [\phi] = \widetilde{\phi}_r \\ \widetilde{\phi}'_f \div [\widetilde{\phi}_r] \div FV(\widetilde{\phi}'_f) \setminus FV(\widetilde{\phi}) = \widetilde{\phi}' \quad \neg \text{writesTo}(FV(\phi), s) \quad \Gamma \vdash \{\widetilde{\phi}_r\} \ s \ \{\widetilde{\phi}'_r\} \end{array}}{\Gamma \vec{\vdash} \{\widetilde{\phi}_f\} \ \text{hold } \phi \ \{ s \} \ \{\widetilde{\phi}'_r \ * \widetilde{\phi}'\}} \text{GHHOLD}$$

$$\frac{\Gamma \vec{\vdash} \{\widetilde{\phi}_p\} \ s_1 \ \{\widetilde{\phi}_q\} \quad \Gamma \vec{\vdash} \{\widetilde{\phi}_q\} \ s_2 \ \{\widetilde{\phi}_r\}}{\Gamma \vec{\vdash} \{\widetilde{\phi}_p\} \ s_1; s_2 \ \{\widetilde{\phi}_r\}} \text{GHSEQ}$$

5 Evaluation/Analysis

> E: with gradual tpestates the same problem happened: as soon as the potential for unknown annotations was accepted, there was a “baseline cost” just to maintain the necessary infrastructure. With simple gradual types, it’s almost nothing. With gradual effects, we’ve shown that it can boil down to very little (a thread-local variable with little overhead, see OOPSLA’15).

6 Conclusion

Recap, remind reader what big picture was. Briefly outline your thesis, motivation, problem, and proposed solution.

6.1 Conceptual Nugget: Comparison/Implication to AGT!

6.2 Limitations

no shared access...

6.3 Future Work

$$\text{wlp}(\text{"x := a.f"}, \text{acc}(\text{b.f})) = \begin{cases} \text{acc}(\text{b.f}) * \text{acc}(\text{a.f}) \\ \text{acc}(\text{b.f}) * (\text{a} = \text{b}) \end{cases}$$

7 Appendix

8 UNSORTED

8.1 HoareMotivEx

Hoare Logic as formal setting

```
class Point
{
    int manhattanDistance(Point p)
        requires \phi_{pre};
        ensures  \phi_{post};
    {
        s1;
        s2;
        .
        .
        .
    }
}
```

$$\text{this} : \text{Point}, p : \text{Point}, \text{result} : \text{int} \vdash \{\phi_{pre}\} \text{ s1; s2; } \dots \{\phi_{post}\}$$

8.2 NPC formula

Checking a formula at runtime, i.e. performing a runtime assertion check, is the integral part of dynamic verification and thus plays a role in gradual verification. Formally, a runtime assertion check corresponds to evaluating a closed formula since the environment provides an instantiation of the formula's free variables. It is reasonable to demand that this check can be performed in a time polynomial, if not linear to the formula's length (the specifics are up to the language designer, of course).

Such a requirement effectively restricts the formula syntax. For example, a syntax containing universal quantification generally violates above runtime limitations: A formula $\forall x_1 \in M, x_2 \in M, \dots, x_n \in M. P(x_1, x_2, \dots, x_n)$ would require $|M|^n$ steps to evaluate. As a result, the execution time is already exponential if M is finite – and unbounded otherwise.

Putting quantification (and therefore the introduction of new variables) aside, there are little restrictions to formula syntax, essentially allowing any predicates or operations that can be evaluated in linear/polynomial time. This includes equality/inequality relations, arithmetic and even own predicates that might be recursive to some extent.

Nevertheless such “easily” evaluable formulas are also subject to higher order reasoning in the static verification rules, including checks like satisfiability of or implication between formulas. Those judgments basically introduce quantification of the free variables, whereas evaluation works on a concrete instantiation. This makes static verification NP-hard in general:

NPC One can easily encode SAT instances as formulas, either directly (if the syntax covers boolean variables, conjunction and disjunction) or using arithmetic (if the syntax covers addition and a comparison relation like “greater-than”). Note that although evaluating such formulas is trivial, checking for satisfiability is NP-complete.

Undecidability ...Paeno-arithmetic

We chose the formula syntax of ... specifically to ensure that even static semantics are decidable in polynomial time. This allowed applying the procedures of AGT directly, as they are based on a decidable type system, i.e. decidable .

8.2.1 Impact of NP-hard Verification Predicates

Let’s assume that our rules for static verification indeed contain an NP-hard predicate P . (NOTE: need positive occurrence for following reasoning!) The immediate consequence is that any working verifier would have to realize a conservative approximation of the actual predicate.

Under-approximation: for static guarantees to hold, verifier must under-approximate P ... blabla

Over-approximation: for (det.) gradual lifting to be ?sound?, it must over-approximate P ... blabla

Bibliography

- [1] Stephan Arlt, Cindy Rubio-González, Philipp Rümmer, Martin Schäfer, and Natarajan Shankar. The gradual verifier. In *NASA Formal Methods Symposium*, pages 313–327. Springer, 2014.
- [2] Felipe Bañados Schwerter, Ronald Garcia, and Éric Tanter. A theory of gradual effect systems. In *ACM SIGPLAN Notices*, volume 49, pages 283–295. ACM, 2014.
- [3] Frank Piessens Wolfram Schulte Bart Jacobs, Jan Smans. A statically verifiable programming model for concurrent object-oriented programs. In *ICFEM*, volume 4260, pages 420–439. Springer, January 2006.
- [4] Yoonsik Cheon and Gary T Leavens. A runtime assertion checker for the java modeling language (jml). 2002.
- [5] M. Christakis, P. Müller, and V. Wüstholtz. Guiding dynamic symbolic execution toward unverified program executions. In L. K. Dillon, W. Visser, and L. Williams, editors, *International Conference on Software Engineering (ICSE)*, pages 144–155. ACM, 2016.
- [6] David Crocker. Safe object-oriented software: the verified design-by-contract paradigm. In *Practical Elements of Safety*, pages 19–41. Springer, 2004.
- [7] Ronald Garcia, Alison M Clark, and Éric Tanter. Abstracting gradual typing. *ACM SIGPLAN Notices*, 51(1):429–442, 2016.
- [8] Ronald Garcia and Eric Tanter. Deriving a simple gradual security language. *arXiv preprint arXiv:1511.01399*, 2015.
- [9] Charles Antony Richard Hoare. An axiomatic basis for computer programming. *Communications of the ACM*, 12(10):576–580, 1969.
- [10] Bart Jacobs and Erik Poll. A logic for the java modeling language jml. In *International Conference on Fundamental Approaches to Software Engineering*, pages 284–299. Springer, 2001.
- [11] K Rustan M Leino, Peter Müller, and Jan Smans. Verification of concurrent programs with chalice. In *Foundations of Security Analysis and Design V*, pages 195–222. Springer, 2009.
- [12] K Rustan M Leino, Greg Nelson, and James B Saxe. Esc/java user’s manual. *ESC*, 2000:002, 2000.
- [13] Francesco Logozzo Manuel Fahndrich, Mike Barnett. Embedded contract languages. In *ACM SAC - OOPS*. Association for Computing Machinery, Inc., March 2010.
- [14] Bertrand Meyer. *Design by contract*. Prentice Hall, 2002.

Bibliography

- [15] Wolfram Schulte Mike Barnett, Rustan Leino. The spec# programming system: An overview. In *CASSIS 2004, Construction and Analysis of Safe, Secure and Interoperable Smart devices*, volume 3362, pages 49–69. Springer, January 2005.
- [16] Greg Nelson. Extended static checking for java. In *International Conference on Mathematics of Program Construction*, pages 1–1. Springer, 2004.
- [17] Matthew J Parkinson and Alexander J Summers. The relationship between separation logic and implicit dynamic frames. In *European Symposium on Programming*, pages 439–458. Springer, 2011.
- [18] Amritam Sarcar and Yoonsik Cheon. A new eclipse-based jml compiler built using ast merging. In *Software Engineering (WCSE), 2010 Second World Congress on*, volume 2, pages 287–292. IEEE, 2010.
- [19] Jeremy G Siek and Walid Taha. Gradual typing for functional languages. In *Scheme and Functional Programming Workshop*, volume 6, pages 81–92, 2006.
- [20] Jeremy G Siek, Michael M Vitousek, Matteo Cimini, and John Tang Boyland. Refined criteria for gradual typing. In *LIPICs-Leibniz International Proceedings in Informatics*, volume 32. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2015.
- [21] Jan Smans, Bart Jacobs, and Frank Piessens. Implicit dynamic frames: Combining dynamic frames and separation logic. In *European Conference on Object-Oriented Programming*, pages 148–172. Springer, 2009.
- [22] Alexander J Summers and Sophia Drossopoulou. A formal semantics for isorecursive and equirecursive state abstractions. In *European Conference on Object-Oriented Programming*, pages 129–153. Springer, 2013.
- [23] Matías Toro and Eric Tanter. Customizable gradual polymorphic effects for scala. In *ACM SIGPLAN Notices*, volume 50, pages 935–953. ACM, 2015.
- [24] Roger Wolff, Ronald Garcia, Éric Tanter, and Jonathan Aldrich. Gradual typestate. In *European Conference on Object-Oriented Programming*, pages 459–483. Springer, 2011.