Gradual Verification of Recursive Heap Data Structures

JENNA WISE, Carnegie Mellon University, USA
JOHANNES BADER, Jane Street, USA
CAMERON WONG, Jane Street, USA
JONATHAN ALDRICH, Carnegie Mellon University, USA
ÉRIC TANTER, University of Chile, Chile
JOSHUA SUNSHINE, Carnegie Mellon University, USA

Current static verification techniques do not provide good support for incrementality, making it difficult for developers to focus on specifying and verifying the properties and components that are most important. Dynamic verification approaches support incrementality, but cannot provide static guarantees. To bridge this gap, prior work proposed gradual verification, which supports incrementality by allowing every assertion to be complete, partial, or omitted, and provides sound verification that smoothly scales from dynamic to static checking. The prior approach to gradual verification, however, was limited to programs without recursive data structures. This paper extends gradual verification to programs that manipulate recursive, mutable data structures on the heap. We address several technical challenges, such as semantically connecting iso- and equi-recursive interpretations of abstract predicates, and supporting gradual verification of heap ownership. This work thus lays the foundation for future tools that work on realistic programs and support verification within an engineering process in which cost-benefit trade-offs can be made.

CCS Concepts: • **Theory of computation → Logic and verification**; **Separation logic**.

Additional Key Words and Phrases: gradual verification, separation logic, implicit dynamic frames, recursive predicates

ACM Reference Format:
Jenna Wise, Johannes Bader, Cameron Wong, Jonathan Aldrich, Éric Tanter, and Joshua Sunshine. 2020. Gradual Verification of Recursive Heap Data Structures. *Proc. ACM Program. Lang.* 4, OOPSLA, Article 228 (November 2020), 28 pages. https://doi.org/10.1145/3428296

1 INTRODUCTION

Hoare proposed a logic for static verification where developers specify method pre- and postconditions [Hoare 1969]. Over time, this work has been extended to support more interesting programs.

*This material is based upon work supported by a Facebook Testing and Verification research award and the National Science Foundation under Grant No. CCF-1901033 and Grant No. DGE1745016. É. Tanter is partially funded by the ANID FONDECYT Regular Project 1190058 and the Millennium Science Initiative Program: code ICN17_002. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, Facebook, ANID, or the Millennium Science Initiative.

Authors' addresses: Jenna Wise, Carnegie Mellon University, USA, jlwise@andrew.cmu.edu; Johannes Bader, Jane Street, USA, johannes-bader@hotmail.de; Cameron Wong, Jane Street, USA, cam@camdar.io; Jonathan Aldrich, Carnegie Mellon University, USA, jonathan.aldrich@cs.cmu.edu; Éric Tanter, University of Chile, Computer Science Department (DCC), Chile, etanter@dcc.uchile.cl; Joshua Sunshine, Carnegie Mellon University, USA, sunshine@cs.cmu.edu.

2475-1421/2020/11-ART228
https://doi.org/10.1145/3428296

Most notably, Reynolds [2002] introduced *separation logic* to support modular verification of programs that manipulate heap data structures. As an extension to separation logic, Parkinson and Bierman [2005] proposed *recursive abstract predicates*, enabling the verification of recursive heap data structures such as graphs, trees, or linked-lists. Shortly after, *implicit dynamic frames* (IDF) was proposed by Smans et al. [2009] as an alternative to separation logic that allows developers to specify heap ownership separately from heap contents.

Unfortunately, these techniques require developers to provide enough specifications to form a complete inductive proof. Consequently, even in very simple programs, a specification that is inductively verifiable may be twice the length of merely specifying the properties the programmer cares about (§3.2). To address this issue, Bader et al. [2018] proposed *gradual verification*, which builds on prior research on gradual typing [Siek and Taha 2007, 2006; Siek et al. 2015], in particular the Abstracting Gradual Typing methodology [Garcia et al. 2016]. Bader et al. [2018] extend a simple Hoare logic static verifier with partial, *imprecise specifications*. Statically, the gradual verifier can optimistically assume any (non-contradictory) strengthening of an imprecise specification. To ensure soundness, dynamic checks are added when partial specifications are optimistically strengthened. Bader *et al.*'s approach smoothly supports the spectrum between static and dynamic verification, as formalized similarly to the refined criteria for gradual typing [Siek et al. 2015].

While promising, the prior work on gradual verification does not support the specification of recursive heap data structures, and thus cannot verify realistic programs. In this paper, we address this limitation by presenting the design, formalization, and meta-theory of a sound gradual verifier for programs that manipulate recursive heap data structures. Our approach follows Bader *et al.*'s methodology, but starts from a static verifier with IDF and recursive abstract predicates. This more sophisticated setting requires us to address the following technical challenges:

- Imprecise specifications may be strengthened not just with boolean assertions about arithmetic expressions, but also with both *abstract predicates* and *accessibility predicates*, which denote ownership of heap locations. Our strengthening definition also includes *self-framing*, a well-formedness condition required by IDF [Smans et al. 2009].
- Both accessibility predicates and abstract predicates must potentially be verified dynamically. Our system verifies accessibility predicates at runtime by tracking and updating a set of owned heap locations. We verify recursive abstract predicates by executing them as recursive boolean functions. This runtime semantics corresponds to an equi-recursive interpretation of abstract predicates, contrasting with the iso-recursive interpretation used in static verifiers [Summers and Drossopoulou 2013]; our theory ensures that these interpretations are consistent.

We show that the resulting gradual verifier is sound, that it is a *conservative extension* of the static verifier—meaning that both coincide on programs with fully-precise specifications—and that it adheres to the *gradual guarantee*. This guarantee, originally formulated for gradual type systems [Siek et al. 2015], captures the intuition that relaxing specifications should not introduce new (static or dynamic) verification errors.

The rest of this paper is outlined as follows. The annotation burden induced by statically verifying linked list insertion is discussed in §2. Section 3 illustrates how this burden can be reduced or eliminated with gradual verification by using examples, and §4 discusses challenges and solutions to supporting such examples. In §5 we formally present a statically verified language supporting a propositional specification logic extended with IDF and recursive heap data structures, before gradualizing the static semantics of this language in §6 and dynamic semantics in §7. §8 discusses the properties of the resulting gradual verifier. Finally, §9 and §10 further relate this paper to prior work and discuss future work, respectively. The supplement of this paper contains full gradual verification examples and supplementary definitions (*e.g.* complete semantics of both the static and

```
1  class Node { int val; Node next; }
2  class List {
3    Node head;
4    void insertLast(int val)
5    {
6      if (this.head == null) {
7        this.head = new Node(val,null);
8      } else {
9        insertLastHelper(val);
10     }
11   }
12   void insertLastHelper(int val)
13   {
14     Node y = this.head;
15     while (y.next != null)
16       { y = y.next; }
17     y.next = new Node(val,null);
18   }
19 }
```

Fig. 1. Linked list with insertion

gradual verifier introduced in this work) [Wise et al. 2020]. Proofs of all propositions and lemmas are also given in the supplement.

2 THE BURDEN OF STATIC VERIFICATION

With static verification tools, ensuring that a component satisfies a given property requires more than specifying the said property: many additional specifications are needed for tools to be able to discharge proof obligations statically. In this section, we show that this additional specification burden can be significant, even for a very simple example.

The program in Figure 1 implements a linked list and two methods for inserting an element at the end of a list. Notice that `insertLastHelper` iteratively traverses a list for insertion, and that both methods diverge if given cyclic lists. Therefore, it is useful to ensure these methods only receive (and produce) acyclic lists. Let us look at how to achieve this with a static verifier.

One way to specify that a list is acyclic is to use the following *abstract predicates* [Parkinson and Bierman 2005], which are essentially pure boolean functions:

```
predicate acyclic(List l) = acc(l.head) ∗ listSeg(l.head,null) and
predicate listSeg(Node from, Node to) = if (from == to) then true
            else acc(from.val) ∗ acc(from.next) ∗ listSeg(from.next,to)
```

Notice that `listSeg` is a *recursive* abstract predicate. Additionally, `acyclic` and `listSeg`'s bodies rely on *accessibility predicates* of the form **acc**(`x.f`) and on the *separating conjunction* $*$, from implicit dynamic frames (IDF) [Smans et al. 2009]. A program can only access a particular heap location if the corresponding accessibility predicate is provided. For example, **acc**(`l.head`) gives permission to access the heap location `o.head` if `l` is bound to the object `o`. The separating conjunction forces accessibility predicates to refer to different heap locations. `listSeg` recursively generates accessibility predicates for every node in a list segment. The accessibility predicates are joined with the separating conjunction. Therefore, the recursive *predicate instance* `acyclic(l)` states that all the heap locations in list `l` are distinct, *i.e.* `l` is acyclic.

A developer can expect to simply specify that both `insertLast` and `insertLastHelper` have `acyclic(this)` as pre- and postconditions. However, they may be disappointed by the many additional specifications required to statically discharge the proof obligations these specifications introduce: loop invariants, fold and unfold statements, and lemmas, as shown in Figure 2, and inspired by Smans et al. [2009]. Although this example is very simple, there is far more specification code (44 lines) than program code (19 lines). Furthermore, this 44:19 ratio only highlights part of the problem: many specifications are far more complex than the program itself, as explained next.

The specification `unfold acyclic(this)` at line 18 expands the abstract predicate `acyclic(this)` into its body. This unfolding exposes the accessibility predicate **acc**(`this.head`), which gives permission to access the heap location of `this.head`. Dually, `fold acyclic(this)` repacks `acyclic(this)`'s body. Figure 2 explicitly uses unfold and fold statements to control the availability of predicate

```
1  class Node { int val; Node next; }
2
3  class List {
4    Node head;
5
6    predicate acyclic(List l) =
7      acc(l.head) * listSeg(l.head,null)
8
9    predicate listSeg(Node from, Node to) =
10     if (from == to) then true else
11       acc(from.val) * acc(from.next) *
12         listSeg(from.next,to)
13
14   void insertLast(int val)
15     requires acyclic(this)
16     ensures acyclic(this)
17   {
18     unfold acyclic(this);
19     if (this.head == null) {
20       this.head = new Node(val,null);
21       fold listSeg(this.head.next,null);
22       fold listSeg(this.head,null);
23       fold acyclic(this);
24     } else {
25       fold acyclic(this);
26       insertLastHelper(val);
27     }
28   }
29
30   void insertLastHelper(int val)
31     requires acyclic(this) *
32       unfolding acyclic(this) in
33         this.head != null
34     ensures acyclic(this)
35   {
36     unfold acyclic(this);
37     Node y = this.head;
38     fold listSeg(this.head,y);
39     unfold listSeg(y,null);
40     while (y.next != null)
41       invariant y != null * acc(this.head) *
42         listSeg(this.head,y) *
43         acc(y.val) * acc(y.next) *
44         listSeg(y.next,null);
45     {
46       Node x = y;
47       y = y.next;
48       unfold listSeg(y,null);
49       fold listSeg(x.next,y);
50       fold listSeg(x,y);
51       appendLemma(this.head, x, y);
52     }
53
54     y.next = new Node(val,null);
55     fold listSeg(y.next.next,null);
56     fold listSeg(y.next,null);
57     fold listSeg(y,null);
58     appendLemma(this.head, y, null);
59     fold acyclic(this);
60   }
61
62   void appendLemma(Node a, Node b, Node c)
63     requires listSeg(a,b) * listSeg(b,c)
64     ensures listSeg(a,c)
65   {
66     if (a == b) {
67     } else {
68       unfold listSeg(a,b);
69       appendLemma(a.next, b, c);
70       fold listSeg(a,c);
71     }
72   }
73 }
```

Static specification ☐ Program code

Fig. 2. Specifying and proving acyclicity for linked list insertion

information. Each predicate instance is an opaque permission to access its body, *i.e.* predicates are *iso-recursive* [Summers and Drossopoulou 2013]. Some *dynamic* verifiers reason about predicate instances *equi-recursively*, *i.e.* treat a predicate instance equal to its complete unfolding. However, completely unfolding recursive predicates often requires statically unknown information, such as the length of the list in our example. Therefore, *static* verifiers reason about predicate instances iso-recursively.

The while loop invariant at lines 41–44 segments a list into three parts using `listSeg`: from the head to the current node (`listSeg(this.head,y)`), the current node (**acc**(`y.val`) * **acc**(`y.next`)), and

the rest of the list (`listSeg(y.next,null)`). The loop body accesses `y.next`, so the loop invariant must expose **`acc`**`(y.next)`. After a new node holding the inserted element is added to the list at line 54, we must show that the `acyclic` predicate holds for the new list. The loop invariant also supports this goal. To build up the `acyclic` predicate we must first construct a `listSeg` predicate from the beginning of the list to the new end of the list. We do this by starting with an empty list segment (line 55) and incrementally extending it with the newly added element (line 56) and the previous end of the list (line 57). This gives us a `listSeg` predicate from the current node to the new end of the list. We then append the `listSeg` predicate from the head of the list to the current node (loop invariant) to the `listSeg` predicate from the current node to the new end of the list (line 57). To achieve this, we need to prove that `listSeg` is transitive. Unfortunately, static tools usually cannot automatically discharge such inductive proofs, so we encode the proof in the `appendLemma` method at 58. Note that such additional proof efforts are part of the barriers to the adoption of static verification, which would be important to get rid of. Finally, we combine the accessibility predicate to the head of the list (loop invariant) with our `listSeg` predicate to reconstruct the `acyclic` predicate (lines 7 and 59).

As the above description makes clear, static verification tools can impose a significant specification burden on developers even for simple programs. Constructing loop invariants and (un)folding predicates can be considerably more complex than program code. Simply ensuring that `insertLast` and `insertLastHelper` receive and produce acyclic lists requires far more specification code than program code. Of course, verifying more properties, for example that some insertion preserves ordering, would require substantially more specification and verification effort.

3 GRADUAL VERIFICATION OF RECURSIVE HEAP DATA STRUCTURES IN ACTION

We now demonstrate how developers can use gradual verification to choose which obligations they want to meet statically and leave the rest to be dynamically checked. They can then incrementally address each proof obligation statically until they reach fully static verification, or stop at any point along the way. As a result, the complexity of verification can be managed in small increments. In the rest of this section, we show different partial specifications of list insertion (§3.1-§3.2), as well as list search (§3.3). These examples illustrate the smooth scaling from dynamic to static checking enabled by gradual verification.

3.1 Gradually Verifying List Insertion: Take 1

Figure 3 presents a possible gradual specification of acyclicity of list insertion. In addition to fully precise formulas (in gray), the specification includes *imprecise formulas* [Lehmann and Tanter 2017] (in yellow), which contain the *unknown formula* `?` in addition to a static part (`true` if omitted).

Here, the developer chooses to completely ignore accessibility predicates, which would be required for full static verification (§2), and only focuses on a partial specification. First, the `acyclic` predicate is kept unknown by using `?` as its body (line 4). Second, only the simple part of the loop invariant—*i.e.* the current node of the list is not null—is statically specified, thanks to the imprecise formula `? * y != null` (line 27). Intuitively, this formula means that only `y != null` is enforced and guaranteed statically, but that other properties can be optimistically assumed. Note that the partial specification explicitly deals with (un)folding the `acyclic` predicate; unfolding `acyclic` implies bringing its imprecision (*i.e.* optimism) in the verification, while folding `acyclic` simply satisfies the declared pre- and postconditions. In general, the only interesting properties that can be verified with this gradual specification are whether `y != null` is preserved by the loop and whether heap accesses are justified with accessibility predicates. We discuss this in more detail.

```
1  class Node { int val; Node next; }
2  class List {
3    Node head;
4    predicate acyclic(List l) = ?
5    void insertLast(int val)
6      requires acyclic(this)
7      ensures acyclic(this)
8    {
9      unfold acyclic(this);
10     if (this.head == null) {
11       this.head = new Node(val,null);
12       fold acyclic(this);
13     } else {
14       fold acyclic(this);
15       insertLastHelper(val);
16     }
17   }
18   void insertLastHelper(int val)
19     requires acyclic(this) *
20       unfolding acyclic(this) in
21         this.head != null
22     ensures acyclic(this)
23   {
24     unfold acyclic(this);
25     Node y = this.head;
26     while (y.next != null)
27       invariant ? * y != null
28       { y = y.next; }
29     y.next = new Node(val,null);
30     fold acyclic(this);
31   }
32 }
```

Imprecise specification Precise specification

Fig. 3. A possible gradual specification of `insertLast` and `insertLastHelper` from Figure 1

```
1  void insertLastHelper(int val)
2    requires acyclic(this) *
3      unfolding acyclic(this) in
4        this.head != null
5    ensures  acyclic(this)
6  {
7    acyclic(this) * unfolding acyclic(this) in
8      this.head != null ⇛
9    ? * acyclic(this)
10   ? * acyclic(this)
11   unfold acyclic(this);
12   ? ⇛ ? * acc(this.head) *
13     this.head != null * acc(this.head.next)
14   ? * acc(this.head) * this.head != null *
15     acc(this.head.next)
16   Node y = this.head;
17   ? * y != null * acc(y.next)
18   while (y.next != null)
19     invariant ? * y != null
20     {
21   ? * y != null * y.next != null * acc(y.next)
22       ⇛ ? * acc(y.next.next)
23           * acc(y.next) * y.next != null
24       ? * acc(y.next.next) * acc(y.next) *
25       y.next != null
26       y = y.next;
27       ? * y != null * acc(y.next)
28     }
29   ? * y != null * y.next == null ⇛
30   ? * acc(y.next)
31   ? * acc(y.next)
32   y.next = new Node(val,null);
33   ?
34   fold acyclic(this);
35   acyclic(this)
36 }
```

Intermediate condition produced by $\widetilde{\text{WLP}}$ Dynamically checked right side of $\widetilde{\Rightarrow}$
Left side of $\widetilde{\Rightarrow}$ Statically checked right side of $\widetilde{\Rightarrow}$

Fig. 4. The gradual verification of insertLastHelper from Figure 3

Figure 4 demonstrates how to gradually verify `insertLastHelper` from Figure 3. The formulas shown in method bodies (highlighted in purple) are the result of applying *gradual weakest liberal precondition* rules $\widetilde{\text{WLP}}$ (defined in §6.5) to each program statement.

$\widetilde{\text{WLP}}$ proceeds from the end of a method body to the beginning, starting with the postcondition on the last line. Then, for each program statement $\widetilde{\text{WLP}}$ calculates a new intermediate condition that

is minimally sufficient to verify the new program statement and the prior intermediate condition. Since ? is the body of the `acyclic` abstract predicate, $\widetilde{\text{WLP}}$ calculates that ? is minimally sufficient for lines 34 and 35. Assigning to `y.next` on line 32 requires an accessibility predicate, so $\widetilde{\text{WLP}}$ joins `acc(y.next)` to ? on line 31.

When $\widetilde{\text{WLP}}$ cannot soundly propagate a condition backwards, a consistent implication ($\widetilde{\Rightarrow}$) check is performed. These implications are necessary under five conditions: at the beginning of a method, at the beginning of a loop body, at the end of a loop with an imprecise invariant, after unfolding an abstract predicate with an imprecise body, and after a method call with an imprecise postcondition. At line 29 the imprecise loop invariant is joined with the negation of the loop guard. The right-hand side of $\widetilde{\Rightarrow}$ is always the next intermediate condition. Since line 29 is not sufficient to statically entail the intermediate condition on 30, but may optimistically do so considering imprecision, it is *optimistically discharged* and therefore highlighted in red. An optimistically-discharged obligation gives rise to a *dynamic check* when running the program. Note that if the left-hand side of a consistent implication cannot possibly imply the right side (*e.g.* as in ? `*` `x == null` $\widetilde{\Rightarrow}$ `x != null`), then the program is *statically rejected*.

The last condition in a loop body is always the loop invariant joined with accessibility predicates needed to evaluate the loop guard. Line 27 contains the loop invariant and an accessibility predicate for `y.next`. When encountering a variable assignment, like the one on line 26, $\widetilde{\text{WLP}}$ substitutes the right-hand side of the assignment (`y.next`) for the left-hand side (`y`) to generate the intermediate condition above the assignment (lines 24 and 25). In addition, accessibility predicates are added for the right-hand side of the assignment (`acc(y.next)`).

As mentioned earlier, a consistent implication is checked at the beginning of a loop body: the left-hand side (line 21) is the loop invariant, the loop guard, and any accessibility predicates necessary for the guard. The right-hand side, as usual, is the next intermediate condition. Observe that here, some of the conditions to prove are *definitely* implied—via standard implication—by the static part of the left-hand side: they can therefore be *discharged statically*, which is highlighted in green (line 23). The others are optimistically discharged, as before.

The condition on line 17 includes the loop invariant and an accessibility predicate to the loop guard. The condition on lines 14 and 15 follows the same pattern as the assignment discussed earlier. The unfold statement generates the consistent implication on lines 12 and 13. The left side is the body of the unfolded abstract predicate, in this case ?. Since ? provides no static information, the entire right-hand side is optimistically discharged.

The condition on line 10 includes the abstract predicate that is unfolded on line 11. This is joined to ? because the body of `acyclic` is an imprecise formula and $\widetilde{\text{WLP}}$ maintains any residual conditions beyond those needed for the unfolding. Finally, the left-hand side of the $\widetilde{\Rightarrow}$ at the beginning of the method is the method precondition (lines 7 and 8). Since `acyclic(this)` is definitely implied, the right-hand side is fully discharged statically.

The complete gradual verification of Figure 3 is given in the supplement [Wise et al. 2020].

3.2 Gradually Verifying List Insertion: Take 2

In Figure 5, we show another, more precise gradual specification of acyclicity for `insertLast` and `insertLastHelper`. The specifications highlighted in gray contain precise formulas, and the ones highlighted in yellow contain imprecise formulas. The darker gray specifications are additional specifications introduced by the developer as an increment over the ones in Figure 3. Here, the developer chooses to fully specify `acyclic`'s body on lines 4 and 5 as **acc**(`l.head`) `*` `listSeg(l.head, null)`. With these predicates, the developer fully specifies `insertLast` for static verification and adds more complete specifications to `insertLastHelper`. The developer uses `listSeg` to write a loop

```
1   class Node { int val; Node next; }
2   class List {
3     Node head;
4     predicate acyclic(List l) =
5       acc(l.head) * listSeg(l.head, null)
6
7     predicate listSeg(Node from, Node to) =
8       if (from == to) then true else
9         acc(from.val) * acc(from.next)
10          * listSeg(from.next, to)
11
12    void insertLast(int val)
13      requires acyclic(this)
14      ensures acyclic(this)
15    {
16      unfold acyclic(this);
17      if (this.head == null) {
18        this.head = new Node(val,null);
19        fold listSeg(this.head.next, null);
20        fold listSeg(this.head, null);
21        fold acyclic(this);
22      } else {
23        fold acyclic(this);
24        insertLastHelper(val);
25      }
26    }
27    void insertLastHelper(int val)
28      requires acyclic(this) *
29        unfolding acyclic(this) in
30          this.head != null
31      ensures ?
32    {
33      unfold acyclic(this);
34      Node y = this.head;
35      unfold listSeg(y, null);
36      while (y.next != null)
37        invariant y != null * acc(y.val) *
38          acc(y.next) * listSeg(y.next, null)
39      {
40        y = y.next;
41        unfold listSeg(y, null);
42      }
43      y.next = new Node(val,null);
44    }
45  }
```

Imprecise specification
Precise specification (from Fig. 3)
New precise specification (increment over Fig. 3)

Fig. 5. Another possible gradual specification of `insertLast` and `insertLastHelper` from Figure 1

invariant, which exposes **acc**(`y.next`) for statically verifying accesses to `y.next` in the loop body and on line 43. However, the developer does not want to build up specifications to statically prove that the new list after insertion is acyclic. They therefore leave the postcondition of `insertLastHelper` unknown. Observe that, in contrast to Figure 2, the programmer does not need to build up a `listSeg` predicate from the previous end of the list to the new one, state and prove a separate lemma about list concatenation, and state a more complex loop invariant. Instead, the gradual verifier ensures at runtime that the new list after insertion is acyclic. This is a major benefit of gradual verification, which can dispense the verification effort from working around certain limitations of static reasoning tools. The detailed verification of Figure 5 with $\widetilde{\text{WLP}}$ is given in the supplement [Wise et al. 2020].

3.3 Gradually Verifying List Search

Let us now consider another helpful method for linked lists, `findMax`, which finds and returns the maximal value of the list. The program in Figure 6 contains an iterative implementation of `findMax`. We discuss how a developer uses gradual verification to ensure that `findMax` indeed returns the maximal value of a list; they incrementally build up specifications as illustrated in Figure 7. In doing so, we show how developers can incrementally address proof obligations of interest and explore the cost-benefit tradeoffs between static reasoning effort and runtime overhead.

```
1  class Node { int val; Node next; }
2
3  class List {
4    Node head;
5
6    int findMax()
7    {
8      int max = this.head.val;
9      Node curr = this.head.next;
10     while (curr != null) {
11       if (curr.val > max) {
12         max = curr.val;
13         curr = curr.next;
14       } else {
15         curr = curr.next;
16       }
17     }
18     result = max;
19   }
20 }
```

Fig. 6. Linked list that iteratively finds and returns its maximal value

The developer begins the first increment (highlighted red, lines 6–20, 23–25, 34) by specifying two properties: whether a value is an upper bound of a list (`upperBound`, lines 6, 7; `upperBoundSeg`, lines 9–12) and whether a value is contained in a list (`contains`, lines 14, 15; `containsSeg`, lines 17–20). The `upperBound` and `contains` predicates are used in `findMax`'s postcondition to ensure that it returns the maximal element (lines 24, 25). The predicates are imprecise to enable heap accesses to `l.head`, `from.val`, and `from.next` without statically-acquired accessibility predicates. The developer specifies that `findMax` not execute on empty lists in its precondition (`this.head != null`). In this first increment, the precondition (line 23) is otherwise imprecise and the loop invariant (line 34) is completely imprecise. As a result, the gradual verifier optimistically assumes—and dynamically checks—accessibility predicates to heap accesses in `findMax`. The invariant also allows the verifier to check `upperBound(this, result)` and `contains(this, result)` at runtime.

To move towards a strengthened version of `findMax`, the developer adds the specifications highlighted in yellow in Figure 7 (lines 30, 31, 35, 55). The developer folds `upperBound(this, result)` on line 55 to show that `findMax` returns an upper bound of the list. The `upperBound` predicate is constructed from an `upperBoundSeg` predicate for the whole list and `result`. To achieve this `upperBoundSeg` predicate, the developer determines that the loop invariant (lines 34, 35) should contain `upperBoundSeg(this.head, curr, max)`. In other words, the loop should produce a value `max` that is the upper bound of the list from its head to the current node at every iteration. Then, when the loop terminates, `max` (`result`) will be an upper bound of the whole list (`upperBoundSeg(this.head, null, max)`). The additional folds before the loop, on lines 30, 31, are used to build up the `upperBoundSeg` for the first loop iteration.

As before, both accessibility predicates and `contains(this, result)` are dynamically verified. However, the verifier now statically establishes that `upperBound(this, result)` holds, at an unfortunate cost. The loop invariant (lines 34, 35) must be preserved for every iteration of the loop, but the developer has only constructed a proof for the first iteration (lines 30, 31). As a result, imprecision introduced by the invariant is used to prove that the invariant holds for the remaining iterations. That is, the invariant is dynamically checked—the list is traversed from its head to the current node—at every iteration beyond the first!

Appalled by this dynamic checking overhead, the developer decides to construct the missing static proofs. The resulting specifications are highlighted in green in Figure 7 (lines 48–50, 58–70). Since the loop's `else` case (lines 46–51) does not modify `max`, the developer focuses their effort here. Their goal is to show that `max` is an upper bound of the list from its head to the next traversed node (line 47). To achieve this, an empty `upperBoundSeg` starting and ending on the next node (line 48) is extended with the previous (current) node (line 49). This creates an `upperBoundSeg` predicate from the current node to the next node. The extension is justified by the negation of the if condition `curr.val` $\leq$ `max` (line 38). Then, the developer achieves their proof goal for the `else` case by appending (lines 50, 58–70) the `upperBoundSeg` predicate from the head of the list to the

```
1  class Node { int val; Node next; }
2
3  class List {
4    Node head;
5
6  predicate upperBound(List l, int bound) =
7      ? * upperBoundSeg(l.head, null, bound)
8
9  predicate upperBoundSeg(Node from, Node to, int bound)
10     = ? * if (from == to) then true else
11       from.val ≤ bound *
12       upperBoundSeg(from.next, to, bound)
13
14 predicate contains(List l, int val) =
15     ? * containsSeg(l.head, null, val)
16
17 predicate containsSeg(Node from, Node to, int val) =
18     ? * if (from == to) then false else
19       if (from.val == val) then true else
20         containsSeg(from.next, to, val)
21
22 int findMax()
23   requires ? * this.head != null
24   ensures upperBound(this, result) *
25     contains(this, result)
26   {
27     int max = this.head.val;
28     Node curr = this.head.next;
29
30     fold upperBoundSeg(this.head.next, curr, max);
31     fold upperBoundSeg(this.head, curr, max);
32
33     while (curr != null)
34     invariant ?
35       * upperBoundSeg(this.head, curr, max)
36     {
37       Node x = curr;
38       if (curr.val > max) {
39         int oldMax = max;
40         max = curr.val;
41         curr = curr.next;
42         fold upperBoundSeg(x.next, curr, max);
43         fold upperBoundSeg(x, curr, max);
44         upperBoundLemma(this.head, x, oldMax, max);
45         appendLemma(this.head, x, curr, max);
46       } else {
47         curr = curr.next;
48         fold upperBoundSeg(x.next, curr, max);
49         fold upperBoundSeg(x, curr, max);
50         appendLemma(this.head, x, curr, max);
51       }
52     }
53
54     result = max;
55     fold upperBound(this, result);
56   }
57
58 void appendLemma(Node a, Node b,
59                  Node c, int val)
60   requires upperBoundSeg(a, b, val) *
61     upperBoundSeg(b, c, val)
62   ensures upperBoundSeg(a, c, val)
63 {
64   if (a == b) {
65   } else {
66     unfold upperBoundSeg(a, b, val);
67     appendLemma(a.next, b, c, val);
68     fold upperBoundSeg(a, c, val);
69   }
70 }
71
72 void upperBoundLemma(Node a, Node b,
73                      int oldVal, int newVal)
74   requires oldVal ≤ newVal *
75     upperBoundSeg(a, b, oldVal)
76   ensures upperBoundSeg(a, b, newVal)
77 {
78   if (a == b) {
79     fold upperBoundSeg(a, b, newVal);
80   } else {
81     unfold upperBoundSeg(a, b, oldVal);
82     appendLemma(a.next, b, oldVal, newVal);
83     fold upperBoundSeg(a, b, newVal);
84   }
85 }
86
87 }
```

1^{st} increment (most imprecise of the 4)
2^{nd} increment
3^{rd} increment
4^{th} increment (most precise of the 4)

Fig. 7. Incrementally more precise ways to gradually verify `findMax` from Figure 6

current node (loop invariant, lines 34, 35) to the upperBoundSeg predicate from the current node to the next node.

Now, the gradual verifier can statically prove that the loop invariant is always preserved by the else branch. However, the verifier still dynamically checks the invariant on each loop iteration executing the then branch. The other dynamic checks for accessibility predicates and the contains predicate also still remain.

Finally, the developer generates specifications for the then branch, highlighted in blue in Figure 7 (lines 42–45, 72–85). As in the else case, the developer's goal is to show that max is an upper bound of the list from its head to the next traversed node (line 41). Here, however, max is assigned the current node's value (line 40). The assignment justfies the build up of the upperBoundSeg predicate from the current node to the next node (lines 42, 43). But, unlike in the else case, the loop invariant's upperBoundSeg predicate applies to an old max value rather than the current (new) one. The old value happens to be less than the current one (then condition, line 38), so the current max is an upper bound of the list from its head to the current node. The developer proves this fact with upperBoundLemma (lines 44, 72–85). Finally, as before, the developer uses appendLemma (lines 45, 58–70) to achieve the proof goal for the then case. This last increment allows the gradual verifier to prove that findMax returns an upper bound of the list completely statically. Only accessibility predicates for heap accesses and contains(this, result) are dynamically checked. The developer can stop here, or work further on either proving contains(this, result) or specifying accessibility predicates.

By using gradual verification on findMax, the developer is able to manage the complexity of meeting proofs obligations incrementally. The developer could have stopped at any of the aforementioned increments and be certain, in the absence of runtime checking errors, that the program returns the greatest element of the list and accesses only owned heap locations. Gradual verification enables the exploration of cost-benefit tradeoffs between static reasoning effort and runtime overhead.

4 CHALLENGES OF RECURSIVE HEAP DATA STRUCTURES

While the basic principles of gradual program verification have already been laid out by Bader et al. [2018], their work only accounts for pre- and postconditions that include basic logical and arithmetic formulas. The contribution of this work is to scale these basic principles to deal with realistic programming scenarios that involve recursive heap data structures.

This section explains the challenges involved in accounting for implicit dynamic frames (IDF) [Smans et al. 2009] and recursive abstract predicates [Parkinson and Bierman 2005] in the context of gradual program verification. We also informally outline our novel solutions to these challenges, which will be formally developed in the following sections.

4.1 Gradual Verification of Heap Ownership

Adapting the Abstracting Gradual Typing approach [Garcia et al. 2016] to the verification setting gives meaning to imprecise formulas such as x > 10 ∧ ? by considering all the *logically consistent strengthenings* of such formulas [Bader et al. 2018; Lehmann and Tanter 2017]. For instance, x > 10 ∧ ? *consistently implies* x > 20, but not x < 0. In the latter case, the formula x < 0 contradicts the static part of the imprecise formula x > 10. In the former case, if we definitely know that x > 10, then it might optimistically be the case that x > 20 as well. Of course, in order to preserve soundness, optimistically assuming x > 20 when one only definitely knows that x > 10 requires a runtime check to corroborate that the value bound to x at runtime is indeed greater than 20.

As we have seen in prior sections, when dealing with heap data structures, the logic—IDF in our case—includes more than arithmetic: we need to be able to talk about heap separation and

ownership of heap cells. How are we to extend the interpretation of imprecise formulas in such a setting, and how can we soundly track optimistic assumptions?

Imprecise Heap Formulas. When using IDF in a static verifier, one must make sure that formulas are self-framed. For instance, `this.head != null` is not self-framed, because it does not explicitly mention the accessibility predicate needed to evaluate the formula. The formula **acc**(`this.head`) ∗ `this.head != null` is self-framed. We want to ensure that programmers can smoothly strengthen specifications, and one logical kind of strengthening is adding accessibility predicates that were previously missing. Accordingly, in our design imprecise formulas must optimistically allow `?` to stand in for accessibility predicates that are necessary for framing. Furthermore, this is true whether the imprecise formula appears directly in an assertion or indirectly in the definition of an abstract predicate. Indeed, in IDF, framing can sometimes come from an abstract predicate. For instance, `acyclic(this)` ∗ `unfolding acyclic(this) in this.head != null` is self-framed if the body of `acyclic(l)` includes **acc**(`l.head`). Thus, our semantics for imprecise formulas must allow `?` to denote not only for predicates such as `acyclic(this)`, but also any unfoldings of them that are necessary to frame the static part of the formula. These semantic choices support different scenarios described in the previous section.

Runtime Checking of Ownership. For a gradual verifier to be sound, optimistic assumptions made statically due to imprecision must be safeguarded dynamically through runtime checks. Extending gradual verification to IDF by allowing imprecision to account for missing accessibility predicates means that we need to keep track of ownership in the runtime system. In particular, we design a runtime that tracks and updates a set of heap locations at every program point, indicating current ownership. Heap locations are added to this set when objects are created. Each time a field is accessed, the set of owned locations is looked up: if the corresponding permission is found, the check succeeds, otherwise a runtime error is raised.

At a call site, if an owned heap location is required by the precondition of the callee, then it is removed from the owned locations of the caller. When the callee finishes executing, all callee owned heap locations are passed to the caller.

The challenge here is how to deal with imprecise preconditions, either directly or via an imprecise abstract predicate. In order to maximize the ability for the callee to execute properly, an imprecise precondition has to require *all* the owned heap locations of the caller. Indeed, said imprecision might potentially denote any location owned by the caller, not already passed statically, and effectively required in the callee. Not transferring its ownership means the callee might error out at runtime.

4.2 Gradual Verification of Recursive Predicates

Recursive predicates can be dealt with in two different manners in program verification [Summers and Drossopoulou 2013]: either *iso-recursively*—in which case to be able to exploit a predicate instance, one needs to explicitly unfold it, and vice versa, to explicitly fold it back to establish it—or *equi-recursively*—in which case a predicate is deemed identical to its unfolding, which need not be specified explicitly. These two approaches have complementary strengths, which, we argue, are particularly relevant when apprehending gradual verification. The iso-recursive approach is critical for making static reasoning manageable for tools (and for humans who must deal with the error messages reported by these tools) because it breaks reasoning into small steps. In contrast, the equi-recursive approach is much more convenient in a dynamic setting, where the runtime system can automatically unfold predicates as needed, and so the user does not have to write explicit folds and unfolds.

In this work, we propose a novel design that achieves the benefits of both approaches. Statically, the gradual verifier treats recursive predicate instances *iso-recursively*: programmers can specify

$$
\begin{array}{rcll}
x, y, z & \in & \mathit{VAR} & \textit{(variables)} \\
v & \in & \mathit{VAL} & \textit{(values)} \\
e & \in & \mathit{EXPR} & \textit{(expressions)} \\
s & \in & \mathit{STMT} & \textit{(statements)} \\
o & \in & \mathit{LOC} & \textit{(object Ids)} \\
\texttt{P} & ::= & \overline{\texttt{cls}}\ s & \\
\texttt{cls} & ::= & \texttt{class}\ C\ \{\ \overline{\texttt{field}}\ \overline{\texttt{pred}}\ \overline{\texttt{method}}\ \} & \\
\texttt{field} & ::= & T\ f; & \\
\texttt{pred} & ::= & \texttt{predicate}\ p(\overline{T\ x}) = \theta & \\
T & ::= & \texttt{int} \mid \texttt{bool} \mid C \mid \top & \\
\texttt{method} & ::= & T\ m(\overline{T\ x})\ \texttt{contract}\ \{s\} & \\
\texttt{contract} & ::= & \texttt{requires}\ \theta\ \texttt{ensures}\ \theta & \\
\oplus & ::= & + \mid - \mid * \mid \backslash & \\
\odot & ::= & \neq \mid = \mid < \mid > \mid \leq \mid \geq &
\end{array}
$$

$$
\begin{array}{rcll}
f & \in & \mathit{FIELDNAME} & \textit{(field names)} \\
m & \in & \mathit{METHODNAME} & \textit{(method names)} \\
C & \in & \mathit{CLASSNAME} & \textit{(class names)} \\
p & \in & \mathit{PREDNAME} & \textit{(predicate names)} \\
s & ::= & \texttt{skip} \mid s\ ;\ s \mid T\ x \mid x := e \mid x.f := y & \\
 & & \mid \texttt{if}\ (e)\ \{\ s\ \}\ \texttt{else}\ \{\ s\ \} & \\
 & & \mid \texttt{while}\ (e)\ \texttt{inv}\ \theta\ \{\ s\ \} \mid x := \texttt{new}\ C & \\
 & & \mid y := z.m(\overline{x}) \mid \texttt{assert}\ \phi \mid \texttt{fold}\ p(\overline{e}) & \\
 & & \mid \texttt{unfold}\ p(\overline{e}) & \\
e & ::= & v \mid x \mid e \oplus e \mid e \odot e \mid e.f & \\
x & ::= & \texttt{result} \mid id \mid \texttt{old}(id) \mid \texttt{this} & \\
v & ::= & n \mid o \mid \texttt{null} \mid \texttt{true} \mid \texttt{false} & \\
\phi & ::= & \texttt{true} \mid \texttt{false} \mid e \odot e \mid p(\overline{e}) \mid \texttt{acc}(e.f) & \\
 & & \mid \texttt{if}\ e\ \texttt{then}\ \phi\ \texttt{else}\ \phi & \\
 & & \mid \texttt{unfolding}\ p(\overline{e})\ \texttt{in}\ \phi \mid \phi \wedge \phi \mid \phi * \phi & \\
\theta & ::= & \text{self-framed}\ \phi &
\end{array}
$$

Fig. 8. SVL_{RP}: Syntax

folds and unfolds in the precise parts of their pre- and postconditions, as well as in program statements, just as they would with mainstream static verifiers. By exploiting syntax, verification becomes simply algorithmic for tools to implement, and visually clear for humans to keep track of the underlying activity of the verifier.

In contrast, dynamically, predicate instances are checked equi-recursively. An equi-recursive evaluation of predicate instances is the natural choice for dynamic checking, as the runtime system can simply execute the predicate as if it were a function. Crucially, an equi-recursive approach to program evaluation allows users to leave out fold and unfold statements, which one can expect to be the default for partially (or un-)verified code. Seen dually, adopting an iso-recursive runtime approach while allowing programmers to omit (un)folding statements would mean trying to automatically infer when to actually perform (un)folding. Known approaches to this are heuristic, meaning that some well-behaved code could be conservatively rejected when made imprecise enough. This would result in a violation of the dynamic gradual guarantee [Siek et al. 2015], whose motto is that *losing precision is harmless*.

Therefore we argue that combining iso- and equi-recursive treatments of recursive predicates is *required* in order to achieve a proper gradual verifier: statically, the iso-recursive approach ensures algorithmic checking, and dynamically, the equi-recursive approach allows imprecise code to run smoothly.

5 SVL_{RP}

Following the AGT methodology [Garcia et al. 2016], gradual verification is built on top of static verification. Therefore, we first formally present a statically verified language supporting a propositional specification logic extended with implicit dynamic frames (IDF) and recursive abstract predicates. This language, called SVL_{RP}, is directly inspired by Summers and Drossopoulou [2013]. Readers familiar with static verification might want to read through this section anyway, because it sets up notations and key concepts used in the gradualization (§6).

5.1 Syntax & Static Semantics

The complete syntax of SVL_{RP} can be found in Figure 8. Programs consist of classes and statements. Classes contain publicly accessible fields, predicates, and methods. Statements include the empty statement, sequences, variable declarations, variable and field assignments, conditionals, while

loops, object allocations, method calls, assertions, as well as fold and unfold statements. Expressions can appear in specifications, and therefore cannot modify the heap. They consist of literal values (integers, objects, `null`, and booleans), variables, arithmetic expressions, comparisons, and field accesses. Methods have contracts consisting of pre- and postconditions, which are formulas represented by ϕ. Formulas join boolean values, comparisons, predicate instances, accessibility predicates, conditionals, and unfoldings via the non-separating conjunction $\wedge$ or the separating conjunction $*$. Note that θ refers to a *self-framed* formula [Smans et al. 2009], formally defined in §5.2.4. We require pre- and postconditions, predicate definitions, and loop invariants to be self-framed.

Looking ahead to gradual verification, we would like formulas to be efficiently evaluable at runtime—and in the presence of accessibility predicates, efficient evaluation requires knowing which branch of a disjunction to evaluate. Therefore, we include a conditional `if` construct in formulas instead of disjunction $\vee$.

As the focus of this work is not on typing, we only consider well-formed and well-typed programs, which is standard and not formalized here. Additionally, variables are declared and initialized before use, and class, predicate, and method names are unique. Finally, contracts should only contain variables that are in scope: a precondition can only contain the method's parameters $\overline{x_i}$ and `this` and a postcondition can only contain the special variable `result`, `this`, and dummy variables $\overline{old(x_i)}$.

5.2 Formula Semantics

In this section, we give meaning to formulas in SVL_{RP}. We also give related definitions for formula satisfiability, implication, footprint computation, and framing. The semantics and related definitions are inspired by Summers and Drossopoulou [2013] and Bader et al. [2018].

5.2.1 Equi-Recursive Evaluation. Evaluating the truth of formulas requires a heap H, a variable environment $\rho \in \textsc{Env}$, and a dynamic footprint $\pi \in \textsc{DynFprint} = \mathcal{P}(\textsc{Loc} \times \textsc{FieldName})$. A heap H is a partial function from heap locations to a value mapping of object fields, *i.e.* $\textsc{Heap} = \textsc{Loc} \rightharpoonup (\textsc{FieldName} \rightharpoonup \textsc{Val})$. Additionally, we introduce a big-step evaluation relation for expressions $H, \rho \vdash e \Downarrow v$, which is standard. An expression e is evaluated according to $H, \rho \vdash e \Downarrow v$ yielding value v. The heap H is used to look up fields and the local variable environment ρ to look up variables.

Then, formula evaluation $\cdot \vDash_E \cdot \subseteq \textsc{Mem} \times \textsc{Formula}$ determines the truth of a formula using heap H, variable environment ρ, dynamic footprint π, and an equi-recursive interpretation of predicate instances. Select rules for formula evaluation are given in Figure 9 (complete rules are in the supplement [Wise et al. 2020]). EvAcc checks whether access demanded by a formula is provided by the dynamic footprint, *e.g.* **acc**(`l.head`) where `l` points to `o` is true when $\langle \texttt{o}, \texttt{head} \rangle \in \pi$. EvSepOp checks two separated subformulas against disjoint partitions of the dynamic footprint. This ensures that access to the same field is not granted twice; for instance, this ensures that **acc**($\texttt{l}_1$`.head`) $*$ **acc**($\texttt{l}_2$`.head`) references two distinct fields. In contrast, the rule for $\wedge$ (EvAndOp) checks both operands against the full dynamic footprint; therefore, **acc**($\texttt{l}_1$`.head`) $\wedge$ **acc**($\texttt{l}_2$`.head`) may reference the same fields. Further, EvPred checks the complete unrolling of a predicate instance using the function $\text{body}_\mu : \textsc{PredName} \to \textsc{Expr}^* \to \textsc{SfrmFormula}$. Given a predicate name and arguments, this function returns the predicate's definition (from the ambient program[1]) after parameter substitution. We make sure that every argument e_i produces a value, only in order to line up with the isorecursive semantics. But we do not need to substitute the values into $\text{body}_\mu(p)(e_1, ..., e_n)$, because it already has the e_i's within it after parameter substitution. Finally, the rule for an unfolding (EvUnfolding) ignores the predicate unfolding, because it is an iso-recursive

[1] Many relations we define are implicitly parameterized over the ambient program.

$$\frac{H, \rho \vdash e \Downarrow o \qquad H, \rho \vdash e.f \Downarrow v \qquad \langle o, f \rangle \in \pi}{\langle H, \rho, \pi \rangle \vDash_E \mathbf{acc}(\mathtt{e.f})} \textsc{EvAcc} \qquad \frac{\langle H, \rho, \pi_1 \rangle \vDash_E \phi_1 \qquad \langle H, \rho, \pi_2 \rangle \vDash_E \phi_2}{\langle H, \rho, \pi_1 \uplus \pi_2 \rangle \vDash_E \phi_1 * \phi_2} \textsc{EvSepOp}$$

$$\frac{H, \rho \vdash e_1 \Downarrow v_1 \qquad \ldots \qquad H, \rho \vdash e_n \Downarrow v_n \qquad \langle H, \rho, \pi \rangle \vDash_E \mathrm{body}_\mu(p)(e_1, \ldots, e_n)}{\langle H, \rho, \pi \rangle \vDash_E p(e_1, \ldots, e_n)} \textsc{EvPred}$$

$$\frac{\langle H, \rho, \pi \rangle \vDash_E \phi}{\langle H, \rho, \pi \rangle \vDash_E \mathtt{unfolding}\ p(e_1, \ldots, e_n)\ \mathtt{in}\ \phi} \textsc{EvUnfolding}$$

Fig. 9. $\mathrm{SVL}_{\mathrm{RP}}$: Formula evaluation (select rules)

$$\frac{H, \rho \vdash e_1 \Downarrow v_1 \qquad \ldots \qquad H, \rho \vdash e_n \Downarrow v_n \qquad \langle p, v_1, \ldots, v_n \rangle \in \Pi}{\langle H, \rho, \Pi \rangle \vDash_I p(e_1, \ldots, e_n)} \textsc{EvPred}$$

Fig. 10. $\mathrm{SVL}_{\mathrm{RP}}$: Iso-recursive formula evaluation (select rule)

only construct. For example, `unfolding acyclic(l) in l.head != null` is true exactly when `l.head != null` is true. Also, all the construct does is provide access to more heap locations in the predicate. The other rules are as expected.

5.2.2 Iso-Recursive Evaluation. We also introduce an iso-recursive formula evaluation semantics, used in static verification (§2). This semantics differs from its equi-recursive counterpart in §5.2.1 on the EvPred rule. Figure 10 presents the iso-recursive version of EvPred. It treats predicate instances as opaque permissions by checking whether a predicate instance demanded by a formula is justified by a dynamic permission set $\Pi \in \textsc{Permissions} = \mathcal{P}((\textsc{Loc} \times \textsc{FieldName}) \cup (\textsc{PredName} \times \textsc{Val}^*))$. Compared to a dynamic footprint, a dynamic permission set can contain dynamic predicate instances in addition to heap locations. For example, `acyclic(l)` where `l` points to `o` is true when $\langle \mathtt{acyclic}, \mathtt{o} \rangle \in \Pi$. Other than EvPred, the iso-recursive semantics is simply defined by replacing π in the equi-recursive rules with Π.

5.2.3 Formula Satisfiability and Implication. Similar to SVL [Bader et al. 2018], formal definitions for formula satisfiability and implication rely on sets of H, ρ, and Π tuples that make formulas true. Definition 5.1 presents a function that produces these sets from formulas. Definitions 5.2 and 5.3 rely on Definition 5.1 to formally state formula satisfiability and implication respectively. Note that these definitions are iso-recursive in order to be implementable in static verification tools (§2).

Definition 5.1 (Denotational Formula Semantics). $[\![\cdot]\!] : \textsc{Formula} \to \mathcal{P}(\textsc{Heap} \times \textsc{Env} \times \textsc{Permissions})$
$[\![\phi]\!] \stackrel{\text{def}}{=} \{\ \langle H, \rho, \Pi \rangle \in \textsc{Heap} \times \textsc{Env} \times \textsc{Permissions} \mid \langle H, \rho, \Pi \rangle \vDash_I \phi\ \}$

Definition 5.2 (Formula Satisfiability). A formula ϕ is satisfiable if and only if $[\![\phi]\!] \neq \emptyset$. Let $\textsc{SatFormula} \subset \textsc{Formula}$ be the set of satisfiable formulas. *Ex.* $\mathbf{acc}(\mathtt{l_1.head}) * \mathbf{acc}(\mathtt{l_2.head})$ is satisfiable since $\mathtt{l_1}$ may not equal $\mathtt{l_2}$. In contrast, $\mathbf{acc}(\mathtt{l_1.head}) * \mathbf{acc}(\mathtt{l_2.head}) * \mathtt{l_1} = \mathtt{l_2}$ is unsatisfiable.

Definition 5.3 (Formula Implication). $\phi_1 \implies \phi_2$ if and only if $[\![\phi_1]\!] \subseteq [\![\phi_2]\!]$.
Ex. $\mathtt{l.head.val} = 6 \implies \mathtt{l.head.val} \geq 5$, and $\mathtt{l.head.val} = 6 \not\Rightarrow \mathbf{acc}(\mathtt{l.head.val}) * \mathtt{l.head.val} \geq 5$ since $\mathbf{acc}(\mathtt{l.head.val})$ is missing on the left-hand side.

5.2.4 Footprints and Framing. A statically-verified language supporting IDF requires formal definitions for the footprint and framing of a formula. These definitions are also iso-recursive.

$$\frac{\langle H,\rho,\Pi\rangle \vDash_I \mathbf{acc}(\mathtt{e.f}) \qquad \langle H,\rho,\Pi\rangle \vdash_{\mathtt{frmI}} e}{\langle H,\rho,\Pi\rangle \vdash_{\mathtt{frmI}} e.f}\ \textsc{FrmField} \qquad \frac{\langle H,\rho,\Pi\rangle \vdash_{\mathtt{frmI}} e}{\langle H,\rho,\Pi\rangle \vdash_{\mathtt{frmI}} \mathbf{acc}(\mathtt{e.f})}\ \textsc{FrmAcc}$$

$$\frac{\langle H,\rho,\Pi\rangle \vdash_{\mathtt{frmI}} \phi_1 \qquad \langle H,\rho,\Pi\rangle \vdash_{\mathtt{frmI}} \phi_2}{\langle H,\rho,\Pi\rangle \vdash_{\mathtt{frmI}} \phi_1 * \phi_2}\ \textsc{FrmSepOp}$$

$$\frac{\langle H,\rho,\Pi\rangle \vdash_{\mathtt{frmI}} e_1 \quad \ldots \quad \langle H,\rho,\Pi\rangle \vdash_{\mathtt{frmI}} e_n}{\langle H,\rho,\Pi\rangle \vdash_{\mathtt{frmI}} p(e_1,\ldots,e_n)}\ \textsc{FrmPred}$$

$$\frac{\begin{array}{c}\langle H,\rho,\Pi\rangle \vDash_I p(e_1,\ldots,e_n) \qquad \langle H,\rho,\Pi\rangle \vdash_{\mathtt{frmI}} e_1 \quad \ldots \quad \langle H,\rho,\Pi\rangle \vdash_{\mathtt{frmI}} e_n \\ \langle H,\rho,\Pi'\rangle \vdash_{\mathtt{frmI}} \phi \qquad \Pi' = \Pi \cup \lfloor \mathrm{body}_\mu(p)(e_1,\ldots,e_n)\rfloor_{H,\rho}\end{array}}{\langle H,\rho,\Pi\rangle \vdash_{\mathtt{frmI}} \mathtt{unfolding}\ p(e_1,\ldots,e_n)\ \mathtt{in}\ \phi}\ \textsc{FrmUnfolding}$$

Fig. 11. SVL$_{\mathrm{RP}}$: Framing (select rules)

The *footprint* of a formula ϕ, denoted $\lfloor\phi\rfloor_{H,\rho}$, is simply the minimum set of permissions Π required to satisfy ϕ given a heap H and variable environment ρ:

$$\lfloor\phi\rfloor_{H,\rho} = \min\{\,\Pi \in \textsc{Permissions} \mid \langle H,\rho,\Pi\rangle \vDash_I \phi\,\}$$

The footprint is defined (*i.e.* there exists a unique minimal set of permission Π) for formulas satisfiable under H and ρ. It can be more directly implemented by simply evaluating ϕ using H and ρ, granting and recording precisely the permissions required. The footprint of `l.head != null` is empty, while the footprint of **acc**(`l.head`) $*$ `l.head != null` is $\{\langle \mathtt{o},\mathtt{head}\rangle\}$ when `l` points to `o`.

A formula is said to be *framed* by permissions Π iff it only mentions fields and unfolds predicates in Π. We give select inference rules for formula framing in Figure 11 and give the rest in the supplement [Wise et al. 2020]. Note that FrmUnfolding allows one unrolling of a predicate to frame a part of a formula. Now, formula ϕ is called *self-framed* (we write $\vdash_{\mathtt{frm}} \phi$) if for all H, ρ, Π $\langle H,\rho,\Pi\rangle \vDash_I \phi$ implies $\langle H,\rho,\Pi\rangle \vdash_{\mathtt{frmI}} \phi$. We define the set of self-framed formulas $\textsc{SfrmFormula} \stackrel{\text{def}}{=} \{\,\phi \in \textsc{Formula} \mid \vdash_{\mathtt{frm}} \phi\,\}$. `l.head != null` is not self-framed, because it can evaluate to `true` even when Π does not contain **acc**(`l.head`). On the other hand, **acc**(`l.head`) $*$ `l.head != null` is self-framed, because it does not evaluate to `true` unless Π contains **acc**(`l.head`). Similarly, `unfolding acyclic(l) in l.head != null` is not self-framed while `acyclic(l)` $*$ `unfolding acyclic(l) in l.head != null` is for `acyclic(l)` with body **acc**(`l.head`). We write θ to denote self-framed formulas.

5.2.5 Relating Permission Sets and Footprints. By using the *footprint* definition in §5.2.4, we can formally relate dynamic permission sets to dynamic footprints via the partial function $\langle\langle\,\cdot\,\rangle\rangle_H$ of type $\textsc{Permissions} \times \textsc{Heap} \rightharpoonup \textsc{DynFprint}$:

$$\langle\langle\,\Pi\,\rangle\rangle_H = \{\langle o,f\rangle \mid \langle o,f\rangle \in \Pi\} \cup \langle\langle\,\Pi'\,\rangle\rangle_H \qquad \mathtt{where}\ \Pi' = \cup_{\langle p,v_1,\ldots,v_n\rangle\in\Pi} \lfloor \mathrm{body}_\mu(p)(v_1,\ldots,v_n)\rfloor_{H,[]}$$

This function completely unrolls the predicate instances in a dynamic permission set gathering owned heap locations on the way. For example, given $\langle$`acyclic`, `o`$\rangle$, with `acyclic` defined precisely as in Figure 2, this function returns all the heap locations ($\{\langle$`o`, `head`$\rangle$, $\langle$`o.head`, `val`$\rangle$, $\langle$`o.head`, `next`$\rangle$, $\ldots\}$) in the list `o`. Note that $\langle\langle\,\cdot\,\rangle\rangle_H$ is only defined when predicates can be finitely unrolled.

5.3 Static Verification

Static verification relies on a *weakest liberal precondition calculus* [Dijkstra 1975] to generate verification conditions. We now present this WLP calculus, which is defined iso-recursively.

5.3.1 WLP Calculus. Select rules for a weakest liberal precondition function $\mathrm{WLP}(s,\theta)$ of type $\textsc{Stmt} \times (\textsc{SatFormula} \cap \textsc{SfrmFormula}) \rightharpoonup (\textsc{SatFormula} \cap \textsc{SfrmFormula})$ are given in Figure 12

$$
\begin{aligned}
\mathrm{WLP}(x := e, \theta) &= \max_{\Rightarrow} \{\, \theta' \mid \theta' \Rightarrow \theta[e/x] \ \wedge\ \theta' \Rightarrow \mathrm{acc}(e) \,\} \\
\mathrm{WLP}(x.f := y, \theta) &= \mathbf{acc}(\mathtt{x.f}) \ \wedge\ \theta[y/x.f] \\
\mathrm{WLP}(y := z.m(\overline{x}), \theta) &= \max_{\Rightarrow} \{\, \theta' \mid\ y \notin \mathrm{FV}(\theta') \ \wedge \\
&\qquad \exists \theta_f \,.\, \theta' \Rightarrow (z \neq \mathtt{null}) \ast \mathrm{mpre}(m)[z/\mathtt{this}, \overline{x/\mathrm{mparam}(m)}] \ast \theta_f \ \wedge \\
&\qquad \theta_f \ast \mathrm{mpost}(m)[z/\mathtt{this}, \overline{x/\mathbf{old}(\mathrm{mparam}(m))}, y/\mathtt{result}] \Rightarrow \theta \,\}
\end{aligned}
$$

Fig. 12. $\mathrm{SVL}_{\mathrm{RP}}$: Weakest liberal precondition calculus (select rules)

(all rules are in the supplement [Wise et al. 2020]). Note, we explicitly restrict the domain and codomain of the WLP function to contain only satisfiable and self-framed formulas. These restrictions are often ensured in Figure 12 by finding a maximum self-framed and satisfiable formula with respect to implication (the weakest formula).

The statement-specific rules for WLP are standard, save for specific care related to field accesses, accessibility predicates, and predicate instances. Rules for variable and field assignment, conditionals, and while loops produce accessibility predicates for field accesses in the program statement, *e.g.* the WLP for `y := l.head` must contain **acc**(`l.head`). Some rules rely on the function acc(e) : Expr $\rightarrow$ Formula, which returns a formula of accessibility predicates corresponding to field accesses in e. More interestingly, the rule for a method call frames off information in the method's postcondition from θ producing the frame θ_f. If the accessibility predicates and predicate instances in θ_f are not in the method's precondition, then θ_f is joined with the precondition to produce the WLP. Consider computing the WLP for the call to `insertLastHelper` on line 26 in Figure 2. In this example, θ = `acyclic(this)`, the precondition is `acyclic(this) * unfolding acyclic(this) in this.head != null`, and the postcondition is `acyclic(this)`. Therefore, θ_f = `true` and the WLP is `this != null * acyclic(this) * unfolding acyclic(this) in this.head != null * true`.

5.3.2 Static Verification. A $\mathrm{SVL}_{\mathrm{RP}}$ program is statically verified if it is a *valid program*:

Definition 5.4 (Valid Method). A method with contract `requires` θ_p `ensures` θ_q, parameters $\overline{x}$, and body s is considered valid if $\theta_p \Rightarrow \mathrm{WLP}(s, \theta_q)[\overline{x/\mathbf{old}(x)}]$ holds.

Definition 5.5 (Valid Program). A program with entry point statement s is considered valid if `true` $\Rightarrow \mathrm{WLP}(s, \mathtt{true})$ holds, $\theta_i \wedge$ `(e = true)` $\Rightarrow \mathrm{WLP}(r, \theta_i)$ and $\theta_i \Rightarrow \mathrm{acc}(e)$ hold for all loops with condition e, body r, and invariant θ_i, and all methods are valid.

5.4 Dynamic Semantics

The soundness of static verification is relative to $\mathrm{SVL}_{\mathrm{RP}}$'s dynamic semantics, which we now expose.

5.4.1 Program States. Program states consist of a heap and a stack, *i.e.* State = Heap × Stack. A stack is made of stack frames that contain a variable environment $\rho \in$ Env, a dynamic footprint $\pi \in$ DynFprint = $\mathcal{P}$(Loc × FieldName), and a program statement $s \in$ Stmt:

$$
S \in \mathrm{Stack} ::= E \cdot S \mid \mathrm{nil} \qquad \text{where} \qquad E \in \mathrm{StackFrame} = \mathrm{Env} \times \mathrm{DynFprint} \times \mathrm{Stmt}
$$

During execution of an $\mathrm{SVL}_{\mathrm{RP}}$ program, expressions and statements operate on the topmost variable environment ρ. Expressions and statements may additionally access and mutate the heap as long as the topmost dynamic footprint contains the corresponding object-field permissions. Thus, the memory accessible at any point of execution can be viewed as a tuple of type Mem = Heap × Env × DynFprint.

$$\frac{\langle H, \rho, \pi\rangle \vDash_E \phi}{\langle H, \langle \rho, \pi, \texttt{assert}\ \phi\ ;\ s\rangle \cdot S\rangle \longrightarrow \langle H, \langle \rho, \pi, s\rangle \cdot S\rangle}\ \textsc{SsAssert}$$

$$\frac{\langle H, \rho, \pi\rangle \vDash_E \mathrm{acc}(e) \qquad H, \rho \vdash e \Downarrow v \qquad \rho' = \rho[x \mapsto v]}{\langle H, \langle \rho, \pi, x := e\ ;\ s\rangle \cdot S\rangle \longrightarrow \langle H, \langle \rho', \pi, s\rangle \cdot S\rangle}\ \textsc{SsAssign}$$

$$\frac{\begin{array}{c}\mathrm{method}(m) = T_r\ m(\overline{T\ x'})\ \texttt{requires}\ \theta_p\ \texttt{ensures}\ \theta_q\ \{\ r\ \} \qquad H, \rho \vdash z \Downarrow o \qquad \overline{H, \rho \vdash x \Downarrow v} \\ \rho' = [\texttt{this} \mapsto o, \overline{x' \mapsto v}, \overline{\texttt{old}(x') \mapsto v}] \qquad \pi' = \langle\langle\ \lfloor \theta_p \rfloor_{H,\rho'}\ \rangle\rangle_H \qquad \pi' \subseteq \pi \qquad \langle H, \rho', \pi'\rangle \vDash_E \theta_p\end{array}}{\langle H, \langle \rho, \pi, y := z.m(\overline{x})\ ;\ s\rangle \cdot S\rangle \longrightarrow \langle H, \langle \rho', \pi', r\ ;\ \texttt{skip}\rangle \cdot \langle \rho, \pi \backslash \pi', y := z.m(\overline{x})\ ;\ s\rangle \cdot S\rangle}\ \textsc{SsCall}$$

$$\frac{\mathrm{mpost}(m) = \theta_q \qquad \langle H, \rho', \pi'\rangle \vDash_E \theta_q \qquad \rho'' = \rho[y \mapsto \rho'(\texttt{result})]}{\langle H, \langle \rho', \pi', \texttt{skip}\rangle \cdot \langle \rho, \pi, y := z.m(\overline{x})\ ;\ s\rangle \cdot S\rangle \longrightarrow \langle H, \langle \rho'', \pi \cup \pi', s\rangle \cdot S\rangle}\ \textsc{SsCallFinish}$$

$$\frac{}{\langle H, \langle \rho, \pi, \texttt{fold}\ p(e_1, \ldots, e_n)\ ;\ s\rangle \cdot S\rangle \longrightarrow \langle H, \langle \rho, \pi, s\rangle \cdot S\rangle}\ \textsc{SsFold}$$

Fig. 13. $\mathrm{SVL}_{\mathrm{RP}}$: Small-step semantics (select rules)

5.4.2 Reduction Rules. Figure 13 presents select rules for $\mathrm{SVL}_{\mathrm{RP}}$'s small-step semantics $\cdot \longrightarrow \cdot \subseteq \textsc{State} \times \textsc{State}$. Complete rules are in the supplement [Wise et al. 2020]. Notably, we structure the rules so as to not require a sequence rule. This aligns the small-step semantics more closely with the WLP calculus, and makes the $\mathrm{SVL}_{\mathrm{RP}}$ soundness proof easier.

The semantics gets stuck when a statement accesses a field that the current state does not own, as specified in SsAssign. Notice that SsAssign relies on acc(e) to check the accessibility of field accesses on the right-hand side. The semantics also gets stuck when preconditions (SsCall), postconditions (SsCallFinish), loop invariants, or assertions (SsAssert) do not hold.

To determine whether a field access is valid at runtime, the semantics tracks a set of owned heap locations π. This set is expanded during allocation with heap locations for the object's fields. At a method call (SsCall) π is split into disjoint caller and callee sets using the method's precondition. The *callee set* π' is derived from the precondition's accessibility predicates and the accessibility predicates gained from unrolling the predicates in the precondition. Ownership of the heap locations in π' is passed to the callee, so the *caller set* is defined as $\pi \setminus \pi'$. After execution of the callee's body finishes (SsCallFinish), execution resumes at the call site. The callee returns to the call site ownership of all received heap locations and new heap locations gained during execution.

Notice that we treat predicates equi-recursively when we track π, determine whether field accesses are valid, and determine whether contracts, loop invariants, or assertions hold. We also treat folds and unfolds equi-recursively as skip statements (SsFold). $\mathrm{SVL}_{\mathrm{RP}}$'s dynamic semantics is equi-recursively defined so the gradual verifier, which builds on $\mathrm{SVL}_{\mathrm{RP}}$'s semantics, adheres to the dynamic gradual guarantee (as discussed in §4.2).

5.5 Soundness

As explained above, the dynamic semantics of $\mathrm{SVL}_{\mathrm{RP}}$ is designed to get stuck when assertions, method contracts, or loop invariants are violated during program execution. The dynamic semantics also gets stuck if a program accesses fields it does not own during execution. Thus informally, *soundness* says that valid $\mathrm{SVL}_{\mathrm{RP}}$ programs do not get stuck, *i.e.* verified programs respect program specifications at runtime. Just as with SVL [Bader et al. 2018], we use a syntactic statement of soundness via progress and preservation.

Now, we introduce the formal definition of a valid state in Definition 5.6. This definition is an invariant that relates the static verification and dynamic semantics of valid $\mathrm{SVL}_{\mathrm{RP}}$ programs. It also relates the formal statements of progress and preservation in Propositions 5.7 and 5.8. Informally,

if the current program state satisfies the WLP of a program, then execution does not get stuck (progress), and after each step of execution, the new state satisfies the WLP of the remaining program (preservation).

Definition 5.6 (Valid State, Final State). We call the state $\langle H, \langle \rho_n, \pi_n, s_n \rangle \cdot \ldots \cdot \langle \rho_1, \pi_1, s_1 \rangle \cdot \mathsf{nil} \rangle \in \textsc{State}$ *valid* if $s_n = s$; skip or skip for some $s \in \textsc{Stmt}$, $s_i = s_i'$; skip for some $s_i' \in \textsc{Stmt}$ for all $1 \leq i < n$, $\pi_i \cap \pi_j = \emptyset$ for all $1 \leq i \leq n, 1 \leq j \leq n$ such that $i \neq j$, and $\langle H, \rho_i, \pi_i \rangle \vDash_E \mathsf{sWLP}_i(s_n \cdot \ldots \cdot s_1 \cdot \mathsf{nil}, \mathtt{true})$ for all $1 \leq i \leq n$ ($\mathsf{sWLP}_i(\overline{s}, \theta)$ is the i-th component of $\mathsf{sWLP}(\overline{s}, \theta)$). A state ψ is *final* if $\psi = \langle H, \langle \rho, \pi, \mathit{skip} \rangle \cdot \mathsf{nil} \rangle$ for some H, ρ, π.

Note that the definition above relies on sWLP, a stack-aware extension of WLP (defined in the supplement [Wise et al. 2020]). sWLP ensures that access permissions are not duplicated in different stack frames. Program validity (Def. 5.5) gives the validity of the initial program state.

Proposition 5.7 (SVL_{RP} Progress). *If ψ is a valid non-final state then $\psi \longrightarrow \psi'$ for some ψ'.*

Proposition 5.8 (SVL_{RP} Preservation). *If ψ is a valid state and $\psi \longrightarrow \psi'$ for some ψ' then ψ' is a valid state.*

6 GVL_{RP}: STATIC SEMANTICS

We now derive GVL_{RP}, the gradually-verified language counterpart of SVL_{RP}, essentially following the Abstracting Gradual Typing methodology [Garcia et al. 2016], whose main principles and mechanisms apply beyond type systems. This section presents the syntax and static semantics of GVL_{RP}. §7 develops the runtime semantics, and §8 establishes the main properties of GVL_{RP}.

6.1 Syntax

The syntax of GVL_{RP} is the same as SVL_{RP} except for the addition of gradual formulas $\widetilde{\phi}$. Gradual formulas replace formulas θ in method contracts, predicate definitions, and loop invariants:

$$\begin{array}{rlrl} \texttt{pred} & ::= \texttt{predicate}\ p(\overline{T\ x}) = \widetilde{\phi} & s & ::= \ldots \mid \texttt{while}\ (e)\ \texttt{inv}\ \widetilde{\phi}\ \{\ s\ \} \\ \texttt{contract} & ::= \texttt{requires}\ \widetilde{\phi}\ \texttt{ensures}\ \widetilde{\phi} & \widetilde{\phi} & ::= \theta \mid ? * \phi \end{array}$$

A gradual formula is either a self-framed *syntactically precise formula* θ or an imprecise formula $? * \phi$. Note that the static part of an imprecise formula does not need to be self-framed (as discussed in §4.1) and ? is syntactic sugar for ? * true. Additionally, the set of all gradual formulas is given by $\widetilde{\textsc{Formula}}$. A *syntactically* precise formula does not contain ? directly, *i.e.* it is not visibly partial. However, it may contain hidden ?s by containing predicates that, when unrolled, expose ?, *e.g.* acyclic(l) where acyclic's body is ?. Self-framing is augmented to handle nested imprecision in GVL_{RP}, and its new definition is given in §6.2. We will refer to formulas that do not contain ?, neither directly nor nested in predicates, as *semantically* precise formulas, *e.g.* acyclic(l) where acyclic's body is **acc**(l.head) * listSeg(l.head, null) (as in Figs. 2 & 5). Note that all semantically precise formulas are syntactically precise, but not all syntactically precise formulas are semantically precise.

6.2 Framing

Definitions for framing and self-framing syntactically precise formulas in GVL_{RP} are redefined to handle imprecise predicate definitions exposed by the FrmUnfolding rule. For example, acyclic(this)'s body is analyzed for the permissions required to frame this.head != null in unfolding acyclic(this) in this.head != null. If acyclic(this)'s body is imprecise, then SVL_{RP}'s framing definition would be undefined for this formula. Therefore, formula framing in GVL_{RP}, $\langle H, \rho, \Pi \rangle \widetilde{\vdash}_{\texttt{frmI}} \phi$, is defined as in SVL_{RP} except for the FrmUnfolding rule:

$$\frac{\langle H,\rho,\Pi\rangle \vDash_I p(e_1,\ldots,e_n) \qquad \langle H,\rho,\Pi\rangle \widetilde{\vdash}_{\texttt{frmI}}\ e_1 \qquad \ldots \qquad \langle H,\rho,\Pi\rangle \widetilde{\vdash}_{\texttt{frmI}}\ e_n \qquad \langle H,\rho,\Pi'\rangle \widetilde{\vdash}_{\texttt{frmI}}\ \phi \qquad \Pi' = \Pi \cup \lfloor \text{body}_\mu(p)(e_1,\ldots,e_n) \rfloor_{\text{TotalFP}(\phi,H,\rho),H,\rho}}{\langle H,\rho,\Pi\rangle \widetilde{\vdash}_{\texttt{frmI}}\ \texttt{unfolding}\ p(e_1,\ldots,e_n)\ \texttt{in}\ \phi}\ \widetilde{\text{FrmUnfolding}}$$

This rule differs from its SVL_{RP} counterpart in computing Π', which aides in framing ϕ. In particular, the retrieval of accessibility predicates and predicate instances from $\text{body}_\mu(p)(e_1,\ldots,e_n)$ now accounts for imprecision. The $\text{TotalFP}(\cdot,\cdot,\cdot) : \text{Formula} \times \text{Heap} \times \text{Env} \to \text{Permissions}$ function returns the explicit and implicit iso-recursive permissions required by ϕ ($\{\langle \texttt{o}, \texttt{head}\rangle\}$ for `this.head != null` where `this` points to `o`). Then, a new footprint definition $\lfloor\widetilde{\phi}\rfloor_{\Pi,H,\rho}$ is used to either frame ϕ optimistically with this maximal permission set or precisely with calculated permissions from $\text{body}_\mu(p)(e_1,\ldots,e_n)$. The result depends on whether $\text{body}_\mu(p)(e_1,\ldots,e_n)$ is imprecise or precise, respectively (`acyclic`'s body is `?` so $\{\langle \texttt{o}, \texttt{head}\rangle\}$ is used):

$$\lfloor\theta\rfloor_{\Pi,H,\rho} = \lfloor\theta\rfloor_{H,\rho} \qquad\qquad \lfloor ? * \phi\rfloor_{\Pi,H,\rho} = \Pi$$

Now, a formula ϕ is called *self-framed* (we write $\widetilde{\vdash}_{\texttt{frm}}\ \phi$) if for all H, ρ, Π, $\langle H,\rho,\Pi\rangle \vDash_I \phi$ implies $\langle H,\rho,\Pi\rangle \widetilde{\vdash}_{\texttt{frmI}}\ \phi$. We redefine the set of self-framed formulas: $\text{SfrmFormula} \stackrel{\text{def}}{=} \{\ \phi \in \text{Formula} \mid \widetilde{\vdash}_{\texttt{frm}}\ \phi\ \}$, and we still write θ to denote self-framed formulas. As a result, `acyclic(this) * unfolding acyclic(this) in this.head != null` is self-framed when `acyclic`'s body is `?`.

6.3 Interpretation of Gradual Formulas

Gradual formulas are given meaning by the set of precise formulas that they represent. The interpretation of gradual formulas is used to define variants of formula evaluation, formula implication, and the WLP calculus that operate over gradual formulas and are *consistent liftings* [Bader et al. 2018; Garcia et al. 2016] of their SVL_{RP} counterparts. Then, the static verification judgment in GVL_{RP} is defined similarly to SVL_{RP} using these lifted definitions. The set denoted by a gradual formula is obtained via a concretization function [Lehmann and Tanter 2017]:

Definition 6.1 (Concretization of Gradual Formulas). $\gamma : \widetilde{\text{Formula}} \rightharpoonup \mathcal{P}^{\text{Formula}}$ is defined as:

$$\gamma(\theta) = \{\ \theta\ \} \qquad \gamma(? * \phi) = \{\ \theta' \in \text{SatFormula} \mid \theta' \Rightarrow \phi\ \} \text{ if } \phi \in \text{SatFormula} \qquad \gamma(? * \phi) \quad \text{undefined otherwise}$$

The concretization of a syntactically precise formula is the singleton set of this formula. The concretization of an imprecise formula is the (infinite) set of syntactically precise formulas that are 1) satisfiable and 2) imply the static part of the imprecise formula. For example, $\gamma(? * x \geq 0) = \{x = 2,\ y = x * x \geq 0,\ \ldots\}$. Notice, $x < 0 * x \geq 0 \notin \gamma(? * x \geq 0)$, because it is not satisfiable.

Novel compared to Bader et al. [2018]'s work is the requirement that all syntactically precise formulas represented by gradual formulas must be *self-framed* (§6.2). This extra condition allows `?` to frame the static part of an imprecise formula, a requirement we motivated in §4.1. Additionally, γ treats predicates opaquely by relying on iso-recursively defined satisfiability, self-framing, and implication. We make this design choice, because γ is an integral part of GVL_{RP}'s static verification system, which we want to be iso-recursive (§4.2). This choice has implications. For example, when both `p(x)` and `q(x)`'s bodies contain **acc**(`x.f`), `p(x)` $*$ `q(x)` is equi-recursively unsatisfiable but iso-recursively satisfiable. Therefore, `p(x)` $*$ `q(x)` $\in \gamma(?$ $*$ `q(x)`$)$. On the other hand, **acc**(`x.f`) $*$ **acc**(`x.f`) $\notin \gamma(?$ $*$ **acc**(`x.f`)$)$, since **acc**(`x.f`) $*$ **acc**(`x.f`) is also iso-recursively unsatisfiable.

Definition 6.1 induces a natural definition of the *(im)precision* of gradual formulas:

Definition 6.2 (Precision of Gradual Formulas). $\widetilde{\phi}_1$ is more precise (*i.e.* less imprecise) than $\widetilde{\phi}_2$, written $\widetilde{\phi}_1 \sqsubseteq \widetilde{\phi}_2$, if and only if $\gamma(\widetilde{\phi}_1) \subseteq \gamma(\widetilde{\phi}_2)$.
Ex. `?` $*$ **acc**(`l.head`) $*$ `listSeg(l.head, null)` $\sqsubseteq$ `?` $*$ **acc**(`l.head`).

Semantic Interpretation of Gradual Formulas. Since Definition 6.1 is interpreted iso-recursively, even if `acyclic`'s body is ?, we can have `acyclic(l)` $*$ `unfolding acyclic(l) in l.head != null` $\in \gamma(? * \texttt{l.head != null})$. That is, γ in Definition 6.1 may give *syntactically* precise, but *semantically* imprecise formulas. We therefore need a semantic interpretation of gradual formulas that extends the concept of concretization to also cover imprecise predicate bodies. As a result, such a *semantic concretization* of gradual formulas would only give *semantically precise* formulas.

A difficulty with writing semantic concretization is that in order to fully interpret formulas, we require an additional function body_μ, which returns predicate bodies from the ambient program given a predicate instance, *e.g.* $\text{body}_\mu(\texttt{acyclic})(\texttt{this}) = ?$. Since body_μ may return imprecise formulas, we cannot use it to interpret formulas that we want to be semantically precise. Instead, we must rely on some new function $\text{body}_\Delta : \textsc{PredName} \to \textsc{Expr}^* \to \textsc{Formula}$, which returns only precise formulas. As a result, we work with *local formulas* $\langle \phi, \text{body}_\Delta \rangle \in \textsc{Formula} \times (\textsc{PredName} \to \textsc{Expr}^* \to \textsc{Formula})$ that explicitly drag along their body function.

Existing rules can easily be adjusted in order to deal with this new parameter, for example:

$$\frac{\text{body}_\Delta(p)(e_1, ..., e_n) = \phi \qquad H, \rho \vdash e_1 \Downarrow v_1 \qquad ... \qquad H, \rho \vdash e_n \Downarrow v_n \qquad \langle H, \rho, \pi \rangle \vDash_E \langle \phi, \text{body}_\Delta \rangle}{\langle H, \rho, \pi \rangle \vDash_E \langle p(e_1, ..., e_n), \text{body}_\Delta \rangle} \textsc{EvPred}$$

The EvPred rule now uses body_Δ to lookup predicate bodies, rather than using the designated body_μ. Notice the function body_Δ is carried around for reference, simply making explicit what was previously assumed as constant and ambient in SVL_{RP}.

Now, we can give an interpretation to *gradual body functions* $\widetilde{\text{body}}_\Delta$ by concretizing them into sets of body_Δ functions that produce precise, self-framed formulas. Given a $\widetilde{\text{body}}_\Delta$, Definition 6.3 returns a set of body_Δ functions constructed from formulas that are in the γ (Def. 6.1) of each gradual formula in $\widetilde{\text{body}}_\Delta$. For example, if $\text{dom}(\widetilde{\text{body}}_\Delta) = \{\texttt{acyclic}\}$, $\widetilde{\text{body}}_\Delta(\texttt{acyclic})(\texttt{l}) = ?$, and $\text{body}_\Delta(\texttt{acyclic})(\texttt{l}) = \textbf{acc}(\texttt{l.head})$, then $\text{body}_\Delta \in \gamma(\widetilde{\text{body}}_\Delta)$. Additionally, each body_Δ function must be well-formed with respect to self-framing, *i.e.* the body that body_Δ returns for each predicate must be self-framed with respect to the body_Δ function itself. For example, if $\text{body}_\Delta(q)(\texttt{l}) = \texttt{acyclic(l)} * \texttt{unfolding acyclic(l) in l.head} \mathrel{!}= \texttt{null}$, then $\text{body}_\Delta(\texttt{acyclic})(\texttt{l})$ must contain $\textbf{acc}(\texttt{l.head})$.

Definition 6.3 (Concretization of Gradual Formulas (continued)). Concretization of a gradual body function $\gamma : (\textsc{PredName} \to \textsc{Expr}^* \to \widetilde{\textsc{Formula}}) \to \mathcal{P}^{\textsc{PredName} \to \textsc{Expr}^* \to \textsc{SfrmFormula}}$ is defined as:

$$\gamma(\widetilde{\text{body}}_\Delta) = \{\ \text{body}_\Delta = \lambda p_i \in \text{dom}(\widetilde{\text{body}}_\Delta).\ \lambda \overline{e} \in \textsc{Expr}^*.\ \theta_{p_i}[\overline{e}/\overline{\texttt{tmp}_i}] \mid \langle \theta_{p_1}, \theta_{p_2}, ... \rangle \in \gamma(\widetilde{\text{body}}_\Delta(p_1)(\overline{\texttt{tmp}_1})) \times \gamma(\widetilde{\text{body}}_\Delta(p_2)(\overline{\texttt{tmp}_2})) \times ..., \forall p_i \in \text{dom}(\widetilde{\text{body}}_\Delta).\ \vdash_{\texttt{frm}} \langle \text{body}_\Delta(p_i)(\overline{\texttt{tmp}_i}), \text{body}_\Delta \rangle\ \}$$

where $\text{dom}(\widetilde{\text{body}}_\Delta) = \{\ p_1, p_2, ...\ \} \subseteq \textsc{PredName}$.

Given this partial function, we can concretize a gradual formula and its gradual body function, yielding a set of semantically precise self-framed formulas:

$$\gamma(\langle \widetilde{\phi}, \widetilde{\text{body}}_\Delta \rangle) = \{\ \langle \theta, \text{body}_\Delta \rangle \mid \theta \in \gamma(\widetilde{\phi}), \text{body}_\Delta \in \gamma(\widetilde{\text{body}}_\Delta),\ \vdash_{\texttt{frm}} \langle \theta, \text{body}_\Delta \rangle\ \}$$

As before, Definition 6.3 allows us to give a natural (semantic) definition for formula precision:

Definition 6.4 (Precision of Formulas (continued)). $\langle \widetilde{\phi}_1, \widetilde{\text{body}}^1_\Delta \rangle$ is more precise than $\langle \widetilde{\phi}_2, \widetilde{\text{body}}^2_\Delta \rangle$, written $\langle \widetilde{\phi}_1, \widetilde{\text{body}}^1_\Delta \rangle \sqsubseteq \langle \widetilde{\phi}_2, \widetilde{\text{body}}^2_\Delta \rangle$ if and only if $\gamma(\langle \widetilde{\phi}_1, \widetilde{\text{body}}^1_\Delta \rangle) \subseteq \gamma(\langle \widetilde{\phi}_2, \widetilde{\text{body}}^2_\Delta \rangle)$.

6.4 Lifting Predicates

We lift predicates on formulas in SVL_{RP} to handle gradual formulas in GVL_{RP} such that they are consistent liftings of corresponding SVL_{RP} predicates. Following AGT [Garcia et al. 2016], the *consistent lifting* $\widetilde{P} \subseteq \widetilde{\textsc{Formula}} \times \widetilde{\textsc{Formula}}$ of predicate $P \subseteq \textsc{Formula} \times \textsc{Formula}$ is defined as:

$$\widetilde{P}(\widetilde{\phi}_1, \widetilde{\phi}_2) \overset{\text{def}}{\iff} \exists \phi_1 \in \gamma(\widetilde{\phi}_1), \phi_2 \in \gamma(\widetilde{\phi}_2).\ P(\phi_1, \phi_2).$$

The existential in this definition expresses the optimistic nature of gradual semantics: we want a gradual predicate to be true if there exists any interpetation of ? that makes the static version of the predicate true.

Since we rely on an equi-recursive dynamic semantics for SVL_{RP} and GVL_{RP} and allow predicate definitions to be imprecise, we now give a semantic definition of gradual formula evaluation:

Definition 6.5 (Consistent Formula Evaluation).
Let $\cdot \mathrel{\widetilde{\models}} \cdot \subseteq \textsc{Mem} \times (\widetilde{\textsc{Formula}} \times (\textsc{PredName} \to \textsc{Expr}^* \to \widetilde{\textsc{Formula}}))$ be defined inductively as

$$\frac{\langle H, \rho, \pi\rangle \models_E \langle \phi, \text{body}_\Delta\rangle \qquad \langle H, \rho, \pi\rangle \vdash_{\texttt{frmE}} \langle \phi, \text{body}_\Delta\rangle}{\langle H, \rho, \pi\rangle \mathrel{\widetilde{\models}} \langle ? * \phi, \widetilde{\text{body}_\Delta}\rangle}$$

$$\frac{\langle H, \rho, \pi\rangle \models_E \langle \theta, \text{body}_\Delta\rangle \qquad \langle H, \rho, \pi\rangle \vdash_{\texttt{frmE}} \langle \theta, \text{body}_\Delta\rangle}{\langle H, \rho, \pi\rangle \mathrel{\widetilde{\models}} \langle \theta, \widetilde{\text{body}_\Delta}\rangle}$$

where $\text{body}_\Delta = \lambda p \in \text{dom}(\widetilde{\text{body}_\Delta}).\ \lambda \overline{e} \in \textsc{Expr}^*.\ \text{static}(\widetilde{\text{body}_\Delta}(p)(\overline{e}))$
and $\text{static} : \widetilde{\textsc{Formula}} \to \textsc{Formula}$ s.t. $\text{static}(\theta) = \theta$ and $\text{static}(? * \phi) = \phi$.

Note that $\cdot \mathrel{\widetilde{\models}} \cdot$ is a consistent lifting of $\cdot \models_E \cdot$ (with γ from Def. 6.3). Our definition is conveniently implementable for equi-recursive dynamic checking: it simply evaluates the static parts of predicates, and ensures that any heap accesses touch only owned locations. For example, if `acyclic`'s body is ? and `l` points to `o`, then `acyclic(l) * unfolding acyclic(l) in l.head != null` evaluates to true when `o.head` is owned and `o.head` $\neq$ `null`. The static part of ? is true, so `acyclic(l)` is ignored.

Additionally, gradual formula evaluation depends on an equi-recursive framing judgment for semantically precise formulas. The framing judgment $\langle H, \rho, \pi\rangle \vdash_{\texttt{frmE}} \phi$ is defined similarly (replacing Π with π and iso-recursive formula evaluation with equi-recursive formula evaluation) to its iso-recursive counterpart in SVL_{RP}, except for FrmPred and FrmUnfolding. Equi-recursive variants of these rules are:

$$\frac{\forall i,\ \langle H, \rho, \pi\rangle \vdash_{\texttt{frmE}} e_i \qquad \langle H, \rho, \pi\rangle \vdash_{\texttt{frmE}} \text{body}_\mu(p)(e_1, ..., e_n)}{\langle H, \rho, \pi\rangle \vdash_{\texttt{frmE}} p(e_1, ..., e_n)} \qquad \frac{\langle H, \rho, \pi\rangle \vdash_{\texttt{frmE}} \phi}{\langle H, \rho, \pi\rangle \vdash_{\texttt{frmE}} \texttt{unfolding}\ p(e_1, ..., e_n)\ \texttt{in}\ \phi}$$

Then, a formula is said to be (equi-recursively) framed by permissions π if its complete unrolling only mentions fields in π. For example, `acyclic(l)`, where `acyclic`'s body is defined as in Figure 2, is framed by π if π contains all of list `l`'s heap locations. We can also easily adjust the equi-recursive framing judgment to pass around and use a body_Δ context, as described in §6.3.

In contrast to gradual formula evaluation (Lemma 6.5), gradual formula implication is a consistent lifting of SVL_{RP} formula implication with the syntactic interpretation of gradual formulas given in Definition 6.1. This is because SVL_{RP} implication is defined iso-recursively, *i.e.* hides imprecision in predicates. We give the definition for gradual formula implication in Lemma 6.6.

Definition 6.6 (Consistent Formula Implication).
Let $\cdot \mathrel{\widetilde{\Rightarrow}} \cdot \subseteq \widetilde{\textsc{Formula}} \times \widetilde{\textsc{Formula}}$ be defined inductively as

$$\frac{\theta_1 \Rightarrow \text{static}(\widetilde{\phi_2})}{\theta_1 \mathrel{\widetilde{\Rightarrow}} \widetilde{\phi_2}}\ \widetilde{\textsc{ImplStatic}} \qquad \frac{\theta \in \textsc{SatFormula} \qquad \theta \Rightarrow \phi_1 \qquad \theta \Rightarrow \text{static}(\widetilde{\phi_2})}{? * \phi_1 \mathrel{\widetilde{\Rightarrow}} \widetilde{\phi_2}}\ \widetilde{\textsc{ImplGrad}}$$

Here also, $\cdot \mathrel{\widetilde{\Rightarrow}} \cdot$ is a consistent lifting of $\cdot \Rightarrow \cdot$ (with γ from Def. 6.3). For example, ? $\mathrel{\widetilde{\Rightarrow}}$? * **acc**(`l.head`) * `l.head != null` because **acc**(`l.head`) * `l.head != null` is satisfiable and implies the static part of both sides of the implication.

$$\widetilde{\mathrm{WLP}}(\texttt{if (}e\texttt{) \{ }s_1\texttt{ \} else \{ }s_2\texttt{ \}}, \widetilde{\phi}) = \alpha(\{\ \max_{\Rightarrow}\ \{\ \phi' \in \textsc{SatFormula} \mid \phi' \Rightarrow \texttt{if } e \texttt{ then } \theta_1 \texttt{ else } \theta_2 \wedge \phi' \Rightarrow \mathrm{acc}(e) \wedge \vdash_{\texttt{frm}} \langle \phi', \mathrm{body}_\Delta{}' \rangle\ \} \mid \theta_1 \in \gamma(\widetilde{\mathrm{WLP}}(s_1, \widetilde{\phi})), \theta_2 \in \gamma(\widetilde{\mathrm{WLP}}(s_2, \widetilde{\phi})), \mathrm{body}_\Delta{}' \in \gamma(\mathrm{body}_\mu),\ \vdash_{\texttt{frm}} \langle \theta_1, \mathrm{body}_\Delta{}' \rangle,\ \vdash_{\texttt{frm}} \langle \theta_2, \mathrm{body}_\Delta{}' \rangle\ \})$$

$$\widetilde{\mathrm{WLP}}(y := z.m(\overline{x}), \widetilde{\phi}) = \alpha(\{\ \max_{\Rightarrow}\ \{\ \phi' \in \textsc{SatFormula} \mid y \notin \mathrm{FV}(\phi') \wedge \vdash_{\texttt{frm}} \langle \phi', \mathrm{body}_\Delta{}' \rangle \wedge \exists \phi_f .\ \phi' \Rightarrow (z \neq \texttt{null}) * \theta_p[z/\mathrm{this}, \overline{x/\mathrm{mparam}(m)}] * \phi_f \wedge \phi_f * \theta_q[z/\mathrm{this}, \overline{x/\texttt{old}(\mathrm{mparam}(m))}, y/\mathrm{result}] \Rightarrow \theta \wedge \vdash_{\texttt{frm}} \langle \phi_f, \mathrm{body}_\Delta{}' \rangle\ \} \mid \theta \in \gamma(\widetilde{\phi}),\ \theta_p \in \gamma(\mathrm{mpre}(m)),\ \theta_q \in \gamma(\mathrm{mpost}(m)),\ \mathrm{body}_\Delta{}' \in \gamma(\mathrm{body}_\mu),\ \vdash_{\texttt{frm}} \langle \theta, \mathrm{body}_\Delta{}' \rangle,\ \vdash_{\texttt{frm}} \langle \theta_p, \mathrm{body}_\Delta{}' \rangle,\ \vdash_{\texttt{frm}} \langle \theta_q, \mathrm{body}_\Delta{}' \rangle\ \})$$

$$\widetilde{\mathrm{WLP}}(\texttt{while (}e\texttt{) inv }\widetilde{\phi}_i\texttt{ \{ }s\texttt{ \}}, \widetilde{\phi}) = \alpha(\{\ \max_{\Rightarrow}\ \{\ \phi' \in \textsc{SatFormula} \mid \phi' \Rightarrow \mathrm{acc}(e) \wedge \vdash_{\texttt{frm}} \langle \phi', \mathrm{body}_\Delta{}' \rangle \wedge \exists \phi_f .\ \phi' \Rightarrow \theta_i * \phi_f \wedge \overline{x_i} \notin \mathrm{FV}(\phi_f) \wedge \vdash_{\texttt{frm}} \langle \phi_f, \mathrm{body}_\Delta{}' \rangle \wedge \phi_f * (\theta_i * (e = \texttt{false}))[\overline{x_i/y_i}] \Rightarrow \theta[\overline{x_i/y_i}]\ \} \mid \theta \in \gamma(\widetilde{\phi}),\ \theta_i \in \gamma(\widetilde{\phi}_i),\ \mathrm{body}_\Delta{}' \in \gamma(\mathrm{body}_\mu),\ \vdash_{\texttt{frm}} \langle \theta, \mathrm{body}_\Delta{}' \rangle,\ \vdash_{\texttt{frm}} \langle \theta_i, \mathrm{body}_\Delta{}' \rangle\ \})$$
$$\text{where } \overline{y_i} \text{ are vars modified by the loop body } s \text{ and } \overline{x_i} \text{ are fresh}$$

$$\widetilde{\mathrm{WLP}}(\texttt{fold } p(\overline{e}), \widetilde{\phi}) = \alpha(\{\ \max_{\Rightarrow}\ \{\ \phi' \in \textsc{SatFormula} \mid \phi' * p(\overline{e}) \Rightarrow \theta \wedge \phi' * p(\overline{e}) \in \textsc{SatFormula} \wedge \vdash_{\texttt{frm}} \langle \phi' * \mathrm{body}_\Delta{}'(p)(\overline{e}), \mathrm{body}_\Delta{}' \rangle\ \} * \mathrm{body}_\Delta{}'(p)(\overline{e}) \in \textsc{SatFormula} \mid \theta \in \gamma(\widetilde{\phi}),\ \mathrm{body}_\Delta{}' \in \gamma(\mathrm{body}_\mu),\ \vdash_{\texttt{frm}} \langle \theta, \mathrm{body}_\Delta{}' \rangle\ \})$$

Fig. 14. $\mathrm{GVL}_{\mathrm{RP}}$: Weakest liberal precondition calculus (select rules).

6.5 Lifting Functions

Functions that operate over formulas in $\mathrm{SVL}_{\mathrm{RP}}$ must also be lifted to handle gradual formulas in $\mathrm{GVL}_{\mathrm{RP}}$. The resulting $\mathrm{GVL}_{\mathrm{RP}}$ functions should approximate consistent liftings of corresponding $\mathrm{SVL}_{\mathrm{RP}}$ functions. Following AGT [Garcia et al. 2016], given a partial function $f : \textsc{Formula} \rightharpoonup \textsc{Formula}$, its *consistent lifting* $\widetilde{f} : \widetilde{\textsc{Formula}} \rightharpoonup \widetilde{\textsc{Formula}}$ is defined as:

$$\widetilde{f}(\widetilde{\phi}) = \alpha(\{\ f(\phi) \mid \phi \in \gamma(\widetilde{\phi})\ \}).$$

Notice, the definition of a *consistent function lifting* requires an abstraction function α, which given a set of formulas produces the most precise gradual formula representing this set. We define $\alpha : \mathcal{P}^{\textsc{Formula}} \rightharpoonup \widetilde{\textsc{Formula}}$ as $\alpha(\overline{\phi}) = \min_{\sqsubseteq}\ \{\ \widetilde{\phi} \in \widetilde{\textsc{Formula}} \mid \overline{\phi} \subseteq \gamma(\widetilde{\phi})\ \}$, *e.g.* $\alpha(\{\texttt{acc}(\texttt{l}_1\texttt{.head}), \texttt{acc}(\texttt{l}_1\texttt{.head}) * \texttt{acc}(\texttt{l}_2\texttt{.head})\}) = \texttt{?} * \texttt{acc}(\texttt{l}_1\texttt{.head})$. Then, α clearly creates a Galois connection with γ from Def. 6.1.

Figure 14 shows select rules for $\widetilde{\mathrm{WLP}}$ (complete rules are in the supplement [Wise et al. 2020]), which approximate the consistent function lifting of WLP. Rules for method call, while loop, and if statements lift the corresponding WLP rules with respect to two (while loop and if statements) or three (method call statements) formula parameters instead of one formula parameter as in other rules. These corresponding WLP rules rely on extra (often implicit) formula parameters that may be imprecise in $\mathrm{GVL}_{\mathrm{RP}}$, and therefore, must be accounted for in the lifting. Similarly, WLP implicitly exposes predicate definitions in body_μ through self-framing (§6.2) and in fold and unfold rules. In $\mathrm{GVL}_{\mathrm{RP}}$, predicate definitions may be imprecise, so non-sequence statement WLP rules are lifted with respect to body_μ. The $\widetilde{\mathrm{WLP}}$ rules are applied to a program in §3.1.

6.6 Lifting the Verification Judgment

We define static verification in $\mathrm{GVL}_{\mathrm{RP}}$ using lifted formula implication ($\widetilde{\Rightarrow}$, §6.4) and lifted WLP ($\widetilde{\mathrm{WLP}}$, §6.5):

Definition 6.7 (Valid Method). A method with contract `requires` $\widetilde{\phi}_p$ `ensures` $\widetilde{\phi}_q$, parameters $\overline{x}$, and body s is considered valid if $\widetilde{\phi}_p \mathrel{\widetilde{\Rightarrow}} \widetilde{\mathrm{WLP}}(s, \widetilde{\phi}_q)[\overline{x/\texttt{old}(x)}]$ holds.

Definition 6.8 (Valid Program). A program with entry point statement s is considered valid if $\texttt{true} \mathrel{\widetilde{\Rightarrow}} \widetilde{\mathrm{WLP}}(s, \texttt{true})$ holds, $\widetilde{\phi}_i \wedge \mathrm{acc}(e) \wedge (\texttt{e = true}) \mathrel{\widetilde{\Rightarrow}} \widetilde{\mathrm{WLP}}(r, \widetilde{\phi}_i \wedge \mathrm{acc}(e))$ holds for all loops with condition e, body r, and invariant $\widetilde{\phi}_i$, and all methods are valid.

7 GVL$_{\mathrm{RP}}$: DYNAMIC SEMANTICS

A valid GVL$_{\mathrm{RP}}$ program will plausibly remain valid during each step of execution. To ensure that it does, the dynamic semantics of SVL$_{\mathrm{RP}}$ are extended with runtime checks and considerations for imprecise specifications.

7.1 Footprint Splitting

To split dynamic footprints at method calls and loop entries in GVL$_{\mathrm{RP}}$'s small-step semantics, we use $\lfloor \widetilde{\phi} \rfloor_{\pi,H,\rho}$:

$$\lfloor \theta \rfloor_{\pi,H,\rho} = \langle\langle\ \lfloor \theta \rfloor_{H,\rho}\ \rangle\rangle_{\pi,H} \qquad\qquad \lfloor ?\ *\ \phi \rfloor_{\pi,H,\rho} = \pi$$

This definition relies on $\langle\langle\ \Pi\ \rangle\rangle_{\pi,H} : \textsc{Permissions} \times \textsc{DynFprint} \times \textsc{Heap} \rightharpoonup \textsc{DynFprint}$, which returns the given dynamic footprint when any predicate bodies analyzed by the function are imprecise. Otherwise, the function returns the dynamic footprint generated from unrolling predicates in Π[2]:

$$\langle\langle\ \Pi\ \rangle\rangle_{\pi,H} = \{\langle o, f\rangle \mid \langle o, f\rangle \in \Pi\} \cup \pi'$$

$$\texttt{where } \pi' = \begin{cases} \pi & \texttt{if } \exists \langle p, v_1, ..., v_n\rangle \in \Pi. \exists \phi \in \textsc{Formula}.\mathrm{body}_\mu(p)(v_1, ..., v_n) = ?\ *\ \phi \\ \langle\langle\ \Pi'\ \rangle\rangle_{\pi,H} & \texttt{otherwise} \\ \texttt{for } \Pi' = \cup_{\langle p, v_1, ..., v_n\rangle \in \Pi} \lfloor \mathrm{body}_\mu(p)(v_1, ..., v_n) \rfloor_{H,[]} \end{cases}$$

Therefore, $\lfloor \widetilde{\phi} \rfloor_{\pi,H,\rho}$ returns the given dynamic footprint π when $\widetilde{\phi}$ is imprecise or contains nested imprecision, and it returns a more precise dynamic footprint computed when $\widetilde{\phi}$ is semantically precise. Example, if `acyclic`'s body is `?`, then $\lfloor$`acyclic(l) * unfolding acyclic(l) in l.head != null`$\rfloor_{\pi,H,\rho}$ will return π. It will return all of list `l`'s heap locations when `acyclic` is defined as in Figures 2 & 5.

7.2 Small-Step Semantics

We give an augmented version of SVL$_{\mathrm{RP}}$'s small-step semantics ($\cdot \mathrel{\widetilde{\longrightarrow}} \cdot \subseteq \textsc{State} \times (\textsc{State} \cup \{\textbf{error}\})$) for GVL$_{\mathrm{RP}}$. We make considerations for imprecision and for runtime verification. Representative rules are given in Figure 15 (complete rules are in the supplement [Wise et al. 2020]).

Imprecision in Specifications. Method preconditions, postconditions, and loop invariants are now checked with gradual formula evaluation (SsCall, SsCallFinish). Asserted formulas must also be checked with gradual formula evaluation due to potentially hidden imprecision (SsAssert). Additionally, we must ensure that introducing imprecision will not introduce a runtime error caused by lack of accessibility (dynamic gradual guarantee, Prop. 8.6). Therefore, if a method precondition in SsCall (or loop invariant) is imprecise or contains nested imprecision, then all owned heap locations are forwarded from the call site to the callee (or loop body) for execution. Otherwise, the call site's owned heap locations can be precisely transferred to the callee (or loop body) as in

[2]Note that $\langle\langle\ \Pi\ \rangle\rangle_{\pi,H}$ is a partial function, as it may not be well-defined if a predicate instance held in Π has an infinite completely unrolling and no nested imprecise predicates.

$$\frac{\langle H, \rho, \pi\rangle \mathrel{\widetilde{\models}} \langle ?\ *\ \phi, \mathsf{body}_\mu\rangle}{\langle H, \langle \rho, \pi, \mathtt{assert}\ \phi\ ;\ s\rangle \cdot S\rangle \mathrel{\widetilde{\longrightarrow}} \langle H, \langle \rho, \pi, s\rangle \cdot S\rangle}\ \textsc{SsAssert}$$

$$\frac{\mathsf{method}(m) = T_r\ m(\overline{T\ x'})\ \mathtt{requires}\ \widetilde{\phi}_p\ \mathtt{ensures}\ \widetilde{\phi}_q\ \{\ r\ \} \qquad H, \rho \vdash z \Downarrow o \qquad \overline{H, \rho \vdash x \Downarrow v}}{\rho' = [\mathtt{this} \mapsto o, \overline{x' \mapsto v}, \mathtt{old}(x') \mapsto v] \qquad \pi' = \lfloor \widetilde{\phi}_p \rfloor_{\pi, H, \rho'} \qquad \pi' \subseteq \pi \qquad \langle H, \rho', \pi'\rangle \mathrel{\widetilde{\models}} \langle \widetilde{\phi}_p, \mathsf{body}_\mu\rangle}{\langle H, \langle \rho, \pi, y := z.m(\overline{x})\ ;\ s\rangle \cdot S\rangle \mathrel{\widetilde{\longrightarrow}} \langle H, \langle \rho', \pi', r\ ;\ \mathtt{skip}\rangle \cdot \langle \rho, \pi \backslash \pi', y := z.m(\overline{x})\ ;\ s\rangle \cdot S\rangle}\ \textsc{SsCall}$$

$$\frac{\mathsf{mpost}(m) = \widetilde{\phi}_q \qquad \langle H, \rho', \pi'\rangle \mathrel{\widetilde{\models}} \langle \widetilde{\phi}_q, \mathsf{body}_\mu\rangle \qquad \rho'' = \rho[y \mapsto \rho'(\mathtt{result})]}{\langle H, \langle \rho', \pi', \mathtt{skip}\rangle \cdot \langle \rho, \pi, y := z.m(\overline{x})\ ;\ s\rangle \cdot S\rangle \mathrel{\widetilde{\longrightarrow}} \langle H, \langle \rho'', \pi \cup \pi', s\rangle \cdot S\rangle}\ \textsc{SsCallFinish}$$

Fig. 15. GVL_{RP}: Small-step semantics adjusted from Fig. 13 (select rules)

SVL_{RP}. Heap locations held after the callee's (or loop body's) execution are returned as usual to the call site.

Runtime Verification. Even for valid GVL_{RP} programs, when specifications are imprecise the formula evaluation premises in GVL_{RP}'s small-step semantics are not guaranteed to hold. Therefore, these premises are turned into runtime checks. If an assertion, accessibility predicate, method precondition, method postcondition, or loop invariant does not hold in a program state where it should, then program execution steps into a dedicated **error** state (extra rules illustrating this can be found in the supplement [Wise et al. 2020]).

8 PROPERTIES OF GVL_{RP}

GVL_{RP} is a sound gradually-verified language that conservatively extends SVL_{RP} and adheres to gradual guarantees. GVL_{RP} is a *conservative extension* of SVL_{RP}—meaning that GVL_{RP} and SVL_{RP} coincide on fully precise programs—by construction following the Abstracting Gradual Typing methodology [Bader et al. 2018; Garcia et al. 2016].

Soundness. Soundness for GVL_{RP} is conceptually similar to soundness for SVL_{RP} except that a GVL_{RP} program may step to a dedicated error state when runtime verification fails. We establish soundness via progress and preservation.

Definition 8.1 (Valid State, Final State). We call the state $\langle H, \langle \rho_n, \pi_n, s_n\rangle \cdot \ldots \cdot \langle \rho_1, \pi_1, s_1\rangle \cdot \mathsf{nil}\rangle \in \textsc{State}$ *valid* if $s_n = s\ ;\ \mathtt{skip}$ or $\mathtt{skip}$ for some $s \in \textsc{Stmt}$, $s_i = s_i'\ ;\ \mathtt{skip}$ for some $s_i' \in \textsc{Stmt}\ \forall\ .\ 1 \leq i < n$, and $s_i = s_i^1\ ;\ s_i^2$ for some $s_i^1, s_i^2 \in \textsc{Stmt}$ where s_i^1 is a method call or while loop statement $\forall\ .\ 1 \leq i < n$. A state ψ is *final* if $\psi = \langle H, \langle \rho, \pi, \mathtt{skip}\rangle \cdot \mathsf{nil}\rangle$ for some H, ρ, π.

Proposition 8.2 (GVL_{RP} Progress). *If ψ is a valid non-final state then $\psi \mathrel{\widetilde{\longrightarrow}} \psi'$ for some ψ' or $\psi \mathrel{\widetilde{\longrightarrow}}$ **error**.*

Proposition 8.3 (GVL_{RP} Preservation). *If ψ is a valid state and $\psi \mathrel{\widetilde{\longrightarrow}} \psi'$ for some ψ' then ψ' is a valid state.*

Gradual Guarantees. GVL_{RP} satisfies both the static and the dynamic gradual guarantees, originally formulated for gradual type systems [Siek et al. 2015], and first adapted to gradual verification by Bader et al. [2018]. These properties ensure in GVL_{RP} that decreasing the precision of specifications never breaks the verifiability and reducibility of a program, *i.e.* losing precision is harmless.

These properties rely on a notion of precision for programs. We say a program p_1 is more precise than program p_2 ($p_1 \sqsubseteq p_2$) if 1) p_1 and p_2 are equivalent except in terms of contracts, loop invariants, and/or predicate definitions, and 2) p_1's contracts, loop invariants, and predicate definitions are more precise than p_2's corresponding contracts, loop invariants, and predicate definitions. A contract $\mathtt{requires}\ \widetilde{\phi}_p^1\ \mathtt{ensures}\ \widetilde{\phi}_q^1$ is more precise than contract $\mathtt{requires}\ \widetilde{\phi}_p^2\ \mathtt{ensures}\ \widetilde{\phi}_q^2$ if $\widetilde{\phi}_p^1 \sqsubseteq \widetilde{\phi}_p^2$

and $\widetilde{\phi}_q^1 \sqsubseteq \widetilde{\phi}_q^2$. Similarly, a loop invariant (predicate definition) $\widetilde{\phi}_i^1$ is more precise than loop invariant (predicate definition) $\widetilde{\phi}_i^2$ if $\widetilde{\phi}_i^1 \sqsubseteq \widetilde{\phi}_i^2$.

Using this notion of program precision, the static gradual guarantee can now be stated as follows:

Proposition 8.4 (GVL_{RP} Static gradual guarantee).
Let $p_1, p_2 \in$ Program such that $p_1 \sqsubseteq p_2$. If p_1 is valid then p_2 is valid.

In general, the static gradual guarantee ensures that reducing the precision of specifications never breaks static verification (*i.e.* makes a valid program invalid).

For the dynamic gradual guarantee, the fact that footprint tracking and splitting is influenced by increasing imprecision (*i.e.* increasing imprecision results in larger parts of footprints being passed up the stack) means that we must define an asymmetric *state precision* relation $\lesssim$:

Definition 8.5 (State Precision). Let $\psi_1, \psi_2 \in$ State. Then ψ_1 is more precise than ψ_2, written $\psi_1 \lesssim \psi_2$, if and only if all of the following applies:
a) ψ_1 and ψ_2 have stacks of size n and identical heaps.
b) ψ_1 and ψ_2 have stacks of variable environments that are identical.
c) Let $s^1_{1..n}$ and $s^2_{1..n}$ be the stack of statements of ψ_1 and ψ_2, respectively. Then for $1 \leq i \leq n$, $s_i^1 \sqsubseteq s_i^2$:
$s \sqsubseteq s'$ if and only if s is a fold or unfold statement and s' is a skip statement or equal to s,
$s = \texttt{while (}e\texttt{) inv } \widetilde{\phi}_i \texttt{ \{ } r \texttt{ \}}$ and $s' = \texttt{while (}e\texttt{) inv } \widetilde{\phi}'_i \texttt{ \{ } r \texttt{ \}}$ where $\widetilde{\phi}_i \sqsubseteq \widetilde{\phi}'_i$,
$s = s_{c_1}; s_{c_2}$ and $s' = s'_{c_1}; s'_{c_2}$ where $s_{c_1} \sqsubseteq s'_{c_1}$ and $s_{c_2} \sqsubseteq s'_{c_2}$, or $s = s'$.
d) Let $\pi^1_{1..n}$ and $\pi^2_{1..n}$ be the stack of footprints of ψ_1 and ψ_2, respectively. Then the following holds for $1 \leq m \leq n$:

$$\bigcup_{i=m}^{n} \pi_i^1 \subseteq \bigcup_{i=m}^{n} \pi_i^2$$

Additionally, as long as it does not break the static gradual guarantee, we allow increased imprecision through dropped fold and unfold statements from one program to the next. This is reflected in condition c) in Definition 8.5 and an adjusted program precision definition $\sqsubseteq_d$. That is, a program p_1 is more precise than a program p_2 if 1) the programs are equivalent except for in terms of contracts, loop invariants, and/or predicate definitions and fold and unfold statements in p_1 may be replaced with `skip` statements in p_2, and 2) p_1's contracts, loop invariants, and predicate definitions are more precise than p_2's corresponding contracts, loop invariants, and predicate definitions. Now, the *dynamic gradual guarantee* can be given:

Proposition 8.6 (GVL_{RP} Dynamic gradual guarantee).
Let $p_1, p_2 \in$ Program such that $p_1 \sqsubseteq_d p_2$, and $\psi_1, \psi_2 \in$ State such that $\psi_1 \lesssim \psi_2$.
If $\psi_1 \xrightarrow{\sim}_{p_1} \psi_1'$, then $\psi_2 \xrightarrow{\sim}_{p_2} \psi_2'$, with $\psi_1' \lesssim \psi_2'$.

Since GVL_{RP} adheres to the dynamic gradual guarantee, reducing the precision of specifications and/or dropping fold and unfold statements does not affect the program's observable behavior.

9 RELATED WORK

We have already discussed the most-closely related research, including the underlying logics [Parkinson and Bierman 2005; Reynolds 2002; Smans et al. 2009] and foundational work on gradual typing and gradual verification [Bader et al. 2018; Garcia et al. 2016; Siek and Taha 2007, 2006; Siek et al. 2015]. The contribution of this work compared to [Bader et al. 2018] is to identify and solve key technical challenges related to recursive heap data structures, namely semantically connecting iso- and equi-recursive interpretations of abstract predicates, and dynamically checking heap ownership.

Lehmann and Tanter [2017] extend the gradual typing paradigm to logical specifications in the form of refinement types. Their language setting is quite different from the one considered here: they deal with higher-order, purely functional programs, while we deal with first-order imperative programs. Therefore they do not have to consider heap ownership. Also, they do not deal with abstract recursive predicates. Combining both approaches in order to account for higher-order stateful programs is a challenging venue for future work.

Prior work on gradual typestate [Garcia et al. 2014; Wolff et al. 2011] and gradual ownership [Sergey and Clarke 2012] integrates static and dynamic checking of ownership of heap data structures. Neither of these efforts considered verifying logical assertions. Both predate the AGT framework that guided our design [Garcia et al. 2016], and the formulation of the gradual guarantees Siek et al. [2015]; it is unclear whether these guarantees are hold in these proposals.

Nguyen et al. [2008] leveraged static information to reduce the overhead of their runtime checking approach for separation logic. They do not try to report static verification failures, because their technique cannot not distinguish between failures due to inconsistent specifications and failures due to incomplete specifications. Also, their runtime checking approach forces developers to specify matching heap footprints in pre- and postconditions to avoid false negatives.

There is also related work focused on making static verification more usable. In particular, Furia and Meyer [2010] infer candidate loop invariants by using heuristics to weaken postconditions into invariants. Their approach cannot infer invariants not expressible as weakenings of postconditions; in contrast, our work can always insert run-time checks where specifications are insufficient for static verification. Additionally, developers can use Dafny's [Leino 2010] **assume** and **assert** statements to debug specifications similar to how they debug programs with print statements [Lucio 2017]. Unlike gradual verification, this approach does not reduce specification burden and requires manual elicitation of missing specifications needed for verification. Similarly, StaDy [Petiot et al. 2014] relies on a combination of static and dynamic analysis techniques to aide developers with debugging specifications. But, it does not reduce specification burden and does not support recursive data structures. Several tools (Smallfoot [Berdine et al. 2005], jStar [Distefano and Parkinson J 2008], Chalice [Leino et al. 2009]) rely on heuristics to infer fold and unfold statements for verification. Incorporating these heuristics in our setting may be challenging due to imprecise specifications, but it is a promising direction for future work.

10 CONCLUSION

Gradual verification is a promising way to enable more incrementality in proofs of programs: developers can focus on the most critical specifications first, benefiting from a combination of static and dynamic checking, and increase the scope of verification over time. By extending sound gradual verification to support programs that manipulate recursive heap data structures, we lay the groundwork for the application of these ideas to realistic programs. Our paper describes how we overcame several key technical challenges, including the semantics of imprecise formulas in the presence of accessibility predicates and recursive predicates, and consistency between iso-recursive static checking and equi-recursive dynamic checking. This opens the door to future work developing prototype gradual verifiers based on our theory, and exploring practical questions such as the efficiency of run-time verification in this setting.

REFERENCES

Johannes Bader, Jonathan Aldrich, and Éric Tanter. 2018. Gradual Program Verification. In *International Conference on Verification, Model Checking, and Abstract Interpretation*. Springer, 25–46.

Josh Berdine, Cristiano Calcagno, and Peter W O'hearn. 2005. Smallfoot: Modular automatic assertion checking with separation logic. In *International Symposium on Formal Methods for Components and Objects*. Springer, 115–137.

Edsger W. Dijkstra. 1975. Guarded Commands, Nondeterminacy and Formal Derivation of Programs. *Commun. ACM* 18, 8 (Aug. 1975), 453–457. https://doi.org/10.1145/360933.360975

Dino Distefano and Matthew J Parkinson J. 2008. jStar: Towards practical verification for Java. *ACM Sigplan Notices* 43, 10 (2008), 213–226.

Carlo Alberto Furia and Bertrand Meyer. 2010. Inferring loop invariants using postconditions. In *Fields of logic and computation*. Springer, 277–300.

Ronald Garcia, Alison M. Clark, and Éric Tanter. 2016. Abstracting Gradual Typing. In *Proceedings of the 43rd Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages* (St. Petersburg, FL, USA) *(POPL '16)*. ACM, New York, NY, USA, 429–442. https://doi.org/10.1145/2837614.2837670

Ronald Garcia, Éric Tanter, Roger Wolff, and Jonathan Aldrich. 2014. Foundations of Typestate-Oriented Programming. 36, 4, Article 12 (Oct. 2014), 12:1–12:44 pages.

Charles Antony Richard Hoare. 1969. An axiomatic basis for computer programming. *Commun. ACM* 12, 10 (1969), 576–580.

Nico Lehmann and Éric Tanter. 2017. Gradual Refinement Types. In *Proceedings of the 44th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages (POPL 2017)*. Paris, France, 775–788.

K Rustan M Leino. 2010. Dafny: An automatic program verifier for functional correctness. In *International Conference on Logic for Programming Artificial Intelligence and Reasoning*. Springer, 348–370.

K Rustan M Leino, Peter Müller, and Jan Smans. 2009. Verification of concurrent programs with Chalice. In *Foundations of Security Analysis and Design V*. Springer, 195–222.

Paqui Lucio. 2017. A Tutorial on Using Dafny to Construct Verified Software. *arXiv preprint arXiv:1701.04481* (2017).

Huu Hai Nguyen, Viktor Kuncak, and Wei-Ngan Chin. 2008. Runtime checking for separation logic. In *International Workshop on Verification, Model Checking, and Abstract Interpretation*. Springer, 203–217.

Matthew Parkinson and Gavin Bierman. 2005. Separation logic and abstraction. In *ACM SIGPLAN Notices*, Vol. 40. ACM, 247–258.

Guillaume Petiot, Nikolai Kosmatov, Alain Giorgetti, and Jacques Julliand. 2014. StaDy: Deep Integration of Static and Dynamic Analysis in Frama-C. (2014).

John C Reynolds. 2002. Separation logic: A logic for shared mutable data structures. In *Logic in Computer Science, 2002. Proceedings. 17th Annual IEEE Symposium on*. IEEE, 55–74.

Ilya Sergey and Dave Clarke. 2012. Gradual Ownership Types. In *Proceedings of the 21st European Conference on Programming Languages and Systems* (Tallinn, Estonia) *(ESOP'12)*. Springer-Verlag, Berlin, Heidelberg, 579–599. https://doi.org/10.1007/978-3-642-28869-2_29

Jeremy Siek and Walid Taha. 2007. Gradual typing for objects. In *European Conference on Object-Oriented Programming*. Springer, 2–27.

Jeremy G Siek and Walid Taha. 2006. Gradual typing for functional languages. In *Scheme and Functional Programming Workshop*, Vol. 6. 81–92.

Jeremy G Siek, Michael M Vitousek, Matteo Cimini, and John Tang Boyland. 2015. Refined criteria for gradual typing. In *LIPIcs-Leibniz International Proceedings in Informatics*, Vol. 32. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik.

Jan Smans, Bart Jacobs, and Frank Piessens. 2009. Implicit dynamic frames: Combining dynamic frames and separation logic. In *European Conference on Object-Oriented Programming*. Springer, 148–172.

Alexander J Summers and Sophia Drossopoulou. 2013. A formal semantics for isorecursive and equirecursive state abstractions. In *European Conference on Object-Oriented Programming*. Springer, 129–153.

Jenna Wise, Johannes Bader, Cameron Wong, Jonathan Aldrich, Éric Tanter, and Joshua Sunshine. 2020. Gradual Verification of Recursive Heap Data Structures. Zenodo. https://doi.org/10.5281/zenodo.4085932

Roger Wolff, Ronald Garcia, Éric Tanter, and Jonathan Aldrich. 2011. Gradual typestate. In *European Conference on Object-Oriented Programming*. Springer, 459–483.