Structuring the Synthesis of Heap-Manipulating Programs

NADIA POLIKARPOVA, University of California, San Diego, USA
ILYA SERGEY, Yale-NUS College, Singapore and National University of Singapore, Singapore

This paper describes a deductive approach to synthesizing imperative programs with pointers from declarative specifications expressed in Separation Logic. Our synthesis algorithm takes as input a pair of assertions—a pre- and a postcondition—which describe two states of the symbolic heap, and derives a program that transforms one state into the other, guided by the shape of the heap. Our approach to program synthesis is grounded in proof theory: we introduce the novel framework of Synthetic Separation Logic (SSL), which generalises the classical notion of heap entailment \( P \vdash Q \) to incorporate a possibility of transforming a heap satisfying an assertion \( P \) into a heap satisfying an assertion \( Q \). A synthesized program represents a proof term for a transforming entailment statement \( P \leadsto Q \), and the synthesis procedure corresponds to a proof search. The derived programs are, thus, correct by construction, in the sense that they satisfy the ascribed pre/postconditions, and are accompanied by complete proof derivations, which can be checked independently.

We have implemented a proof search engine for SSL in a form of the program synthesizer called SuSLik. For efficiency, the engine exploits properties of SSL rules, such as invertibility and commutativity of rule applications on separate heaps, to prune the space of derivations it has to consider. We explain and showcase the use of SSL on characteristic examples, describe the design of SuSLik, and report on our experience of using it to synthesize a series of benchmark programs manipulating heap-based linked data structures.

CCS Concepts: • Theory of computation → Logic and verification; • Software and its engineering → Automatic programming;

Additional Key Words and Phrases: Program Synthesis, Separation Logic, Proof Systems, Type Theory

ACM Reference Format:

1 INTRODUCTION

Consider the task of implementing a procedure \( \text{swap}(x, y) \), which swaps the values stored at two distinct heap locations, \( x \) and \( y \). The desired effect of \( \text{swap} \) can be concisely captured via pre/postconditions expressed in Separation Logic (SL)—a Hoare-style program logic for specifying and verifying stateful programs with pointers (O’Hearn et al. 2001; Reynolds 2002):

\[
\{ x \mapsto a \land y \mapsto b \} \text{} \text{void} \text{} \text{swap} (\text{loc } x, \text{ loc } y) \{ x \mapsto b \land y \mapsto a \}
\] (1)

This specification is declarative: it describes what the heap should look like before and after executing \( \text{swap} \) without saying how to get from one to the other. Specifically, it states that the program takes as input two pointers, \( x \) and \( y \), and runs in a heap where \( x \) points to an unspecified value \( a \), and \( y \) points to \( b \). Both \( a \) and \( b \) here are logical (ghost) variables, whose scope captures both pre- and postcondition (Kleymann 1999). Because these variables are ghosts, we cannot use them
directly to update the values in x and y as prescribed by the postcondition; the program must first
“materialize” them by reading them into local variables, a2 and b2 (cf. lines 2–3 of the code below).

In our minimalistic C-like language, loc denotes untyped
pointers, and let introduces a local variable. Unlike in C,
both formals and locals are immutable (the only mutation
is allowed on the heap).

As a result of the two reads, the ghost variables in the
post-condition can now be substituted with equal program-level variables: x ↦→ b2 ∗ y ↦→ a2. This
updated postcondition can be realized by the two writes on lines 4–5, which conclude our imple-
mentation, so the whole program can now be verified against the specification (1).

Having done this exercise in program derivation, let us now observe that the SL specification has
been giving us guidance on what effectful commands (e.g., reads and writes) should be emitted next.
In other words, the synthesis of swap has been governed by the given specification in the same way
proof search is guided by a goal in ordinary logics. In this work, we make this connection explicit
and employ it for efficiently synthesizing imperative programs from SL pre- and postconditions.

**Motivation.** The goal of this work is to advance the state of the art in synthesizing provably
correct heap-manipulating programs from declarative functional
specifications. Fully-automated
program synthesis has been an active area of research in the past years, but recent techniques
mostly targeted simple DSLs (Gulwani et al. 2011; Le and Gulwani 2014; Polozov and Gulwani 2015)
or purely functional languages (Feser et al. 2015; Kneuss et al. 2013; Osera and Zdancewic 2015;
Polikarpova et al. 2016). The primary reason is that those computational models impose strong
structural constraints on the space of programs, either by means of restricted syntax or through a
strong type system. These structural constraints enable the synthesizer to discard many candidate
terms a-priori, before constructing the whole program, leading to efficient synthesis.

Low-level heap-manipulating programs in general-purpose languages like C lack inherent struc-
tural constraints wrt. control- and data-flow, and as a result the research in synthesizing such
programs has been limited to cases when such constraints can be imposed by the programmer.
From the few existing approaches we are aware of, SIMPL (So and Oh 2017) and IMPSYNT (Qiu and
Solar-Lezama 2017) require the programmer to provide rather substantial
sketches of the control-
flow structure, which help restrict the search space; JENNISYS by Leino and Milicevic (2012) does
not require sketches, but also does not scale to functions with destructive heap updates.

**Key Ideas.** Our theoretical insight is that the structural constraints missing from an imperative
language itself, can be recovered from the program logic used to reason about programs in that
language. We observe that synthesis of heap-manipulating programs can be formulated as a proof
search in a generalized proof system that combines entailment with Hoare-style reasoning for
unknown programs. In this generalized proof system, a statement \( P \leadsto Q \) means that there exists a
program c, such that the Hoare triple \( \{P\} c \{Q\} \) holds; the witness program c serves as a proof term
for the statement. In order to be useful, the system must satisfy a number of practical restrictions.
First, it should be expressive enough to (automatically) verify the programs with non-trivial heap
manipulation. Second, it should be restrictive enough to make the synthesis problem tractable.
Finally, it must ensure the termination of the (possibly recursive) synthesized programs, to avoid
vacuous proofs of partial correctness.

In this paper we design such a generalized proof system based on the symbolic heap fragment of malloc/free Separation Logic\(^1\) with inductive predicates, to which we will further refer as just Separation Logic or SL (O’Hearn et al. 2009; Reynolds 2002). Separation Logic has been immensely

\(^1\)This nomenclature is due to Cao et al. (2017), who provide it as a rigorous alternative to the folklore notion of classical SL.
successful at specifying and verifying many kinds of heap-manipulating programs, both interactively and automatically (Appel et al. 2014; Berdine et al. 2011; Charguéraud 2010; Chen et al. 2015; Chin et al. 2012; Chlipala 2011; Distefano and Parkinson 2008; Nanevski et al. 2010; Piskac et al. 2014a), and is employed in modern symbolic execution tools (Berdine et al. 2005; Rowe and Brotherston 2017). We demonstrate how to harness all this power for program synthesis, devise the corresponding search procedure and apply it to synthesize a number of non-trivial programs that manipulate linked data structures. Finally, we show how to exploit laws of SL and properties of our proof system to prune the search space and make the synthesis machinery efficient for realistic examples.

Contributions. The central theoretical contribution of the paper is Synthetic Separation Logic (SSL): a system of deductive synthesis rules, which prescribe how to decompose specifications for complex programs into specifications for simpler programs, while synthesizing the corresponding computations compositionally. In essence, SSL is a proof system for a new transforming entailment judgment $P \leadsto Q | c$ (reads as “the assertion $P$ transforms into $Q$ via a program $c$”), which unifies SL entailment $P \vdash Q$ and verification $\{P\} c \{Q\}$, with the former expressible as $P \leadsto Q | \text{skip}$.

The central practical contribution is the design and implementation of SuSLik—a deductive synthesizer for heap-manipulating programs, based on SSL. SuSLik takes as its input a library of inductive predicates, a (typically empty) list of auxiliary function specifications, and an SL specification of the function to be synthesized. It returns a—possibly recursive, but loop-free—program (in a minimalistic C-like language), which provably satisfies the given specification.

Our evaluation shows that SuSLik can synthesize all structurally-recursive benchmarks from previous work on heap-based synthesis (Qiu and Solar-Lezama 2017), without any sketches and in most cases much faster. To the best of our knowledge, it is also the first synthesizer to automatically discover the implementations of copying linked lists and trees, and flattening a tree to a list.

The essence of SuSLik’s synthesis algorithm is a backtracking search in the space of SSL derivations. Even though the structural constraints (i.e., the shape of the heap) embodied in the synthesis rules already prune the search space significantly (as shown by our swap example), a naïve backtracking search is still impractical, especially in the presence of inductive heap predicates. To eliminate redundant backtracking, we develop several principled optimizations. In particular, we draw inspiration from focusing proof search (Pfenning 2010) to identify invertible synthesis rules that do not require backtracking, and exploit the frame rule of SL, observing that the order of rule applications is irrelevant whenever their subderivations have disjoint footprints.

Paper outline. In the remainder of the paper we give an overview of SSL reasoning principles (Sec. 2), describe its rules and the meta-theory (Sec. 3), outline the design of our synthesis tool (Sec. 4), present the optimizations and extensions of the basic search algorithm (Sec. 5), and report on the evaluation of the approach on a set of case studies involving various linked structures (Sec. 6), concluding with a discussion of limitations (Sec. 7) and a comparison with the related work (Sec. 8).

2 DEDUCTIVE SYNTHESIS FROM SEPARATION LOGIC SPECIFICATIONS

In Separation Logic, assertions capture the program state, represented by a symbolic heap. An SL assertion (ranged over by symbols $P$ and $Q$ in the remainder of the paper) is customarily represented as a pair $\{\phi; P\}$ of a pure part $\phi$ and a spatial part $P$. The pure part (ranged over by $\phi$, $\psi$, and $\chi$) is a quantifier-free boolean formula, which describes the constraints over symbolic values (represented by variables $x$, $y$, etc.). The spatial part (denoted $P$, $Q$, and $R$) is represented by a collection of primitive heap assertions describing disjoint symbolic heaps (heaplets), conjoined by the separating conjunction operation $\ast$, which is commutative and associative (Reynolds 2002). For example, in the assertion $\{a \neq b; x \mapsto a \ast y \mapsto b\}$ the spatial part describes two disjoint memory cells that store symbolic values $a$ and $b$, while the pure part states that these values are distinct.
Our development is agnostic to the exact logic of pure formulae, as long as it supports standard boolean connectives and equality, and comes equipped with a validity oracle and a synthesis oracle (we elaborate on both later in this section). The soundness of our development depends only on the soundness of the validity oracle. Our implementation uses the quantifier-free logic of arrays, uninterpreted functions, and linear integer arithmetic, which is sufficient to express all examples in this paper, and is efficiently decidable by SMT solvers, thus providing a sound and complete validity oracle.

To begin with our demonstration, the only kinds of heaplets we are going to consider are the empty heap assertion emp and points-to assertions of the form \( (x, i) \mapsto e \), where \( x \) is a variable or a pointer constant (e.g., 0), \( i \) is a non-negative integer offset (necessary to represent records), and \( e \) is a symbolic value stored in the memory cell addressed via the value of \((x + i)\). In most cases, the offset is 0, so we will abbreviate heap assertions \((x, 0) \mapsto e \) as \( x \mapsto e \).

Our programming component (to be presented formally in Sec. 3) is a simple imperative language that supports reading from pointers to (immutable) local variables (let \( x = *y \)), storing values into pointers (\( *y = e \)), conditionals, recursive calls, and pure expressions. The language has no return statement; instead, a function stores its result into an explicitly passed pointer.

### 2.1 Specifications for Synthesis

A synthesis goal is a triple \( \Gamma; \mathcal{P} \leadsto \mathcal{Q} \), where \( \Gamma \) is an environment, i.e., a set of immutable program variables, \( \mathcal{P} \) is a precondition, and \( \mathcal{Q} \) is a postcondition. Solving a synthesis goal amounts to finding a program \( c \) and a derivation of the SSL assertion \( \Gamma; \mathcal{P} \leadsto \mathcal{Q} \vdash c \). To avoid clutter, we employ the following naming conventions:

- (a) the symbols \( \mathcal{P} \), \( \phi \), and \( \mathcal{P} \) refer to the goal’s precondition, its pure, and spatial part;
- (b) similarly, the symbols \( \mathcal{Q} \), \( \psi \), and \( \mathcal{Q} \) refer to the goal’s postcondition, its pure and spatial part;
- (c) whenever the pure part of a SL assertion is true (\( \top \)), it is omitted from the presentation.

In addition, we will use the following macros to express the scope and the quantification over variables of a goal \( \Gamma; \{ \mathcal{P} \} \leadsto \{ \mathcal{Q} \} \). First, by \( \text{Vars}(A) \) we will denote all variables occurring in \( A \), which might be an assertion, a logical formula, or a program. Ghosts (universally-quantified logical variables), whose scope is both the pre- and the postcondition, are defined as \( \text{GV}(\Gamma, \mathcal{P}, \mathcal{Q}) = \text{Vars}(\mathcal{P}) \setminus \Gamma \). Existentials are defined as \( \text{EV}(\Gamma, \mathcal{P}, \mathcal{Q}) = \text{Vars}(\mathcal{Q}) \setminus (\Gamma \cup \text{Vars}(\mathcal{P})) \). For instance, taking \( \Gamma = \{ x \} \), \( \mathcal{P} = \{ x \neq y; x \mapsto y \} \), \( \mathcal{Q} = \{ x \mapsto z \} \), we have the ghosts \( \text{GV}(\Gamma, \mathcal{P}, \mathcal{Q}) = \{ y \} \), the existentials \( \text{EV}(\Gamma, \mathcal{P}, \mathcal{Q}) = \{ z \} \), and the pure part of the postcondition \( \mathcal{Q} \) is implicitly true.

### 2.2 Basic Inference Rules

To get an intuition on how to represent program synthesis as a proof derivation in SSL, consider Fig. 1, which shows four basic rules of the logic, targeted to synthesize programs with constant user-defined inductive predicates.

The \text{Emp} rule is applicable when spatial pre- and postcondition are both empty. It requires that no existentials remains in the goal, and the pure precondition implies the pure postcondition: the premise \( \Gamma \vdash \phi \Rightarrow \psi \) represents an invocation of the pure validity oracle. \text{Emp} has no subgoals in its premises (i.e., it is a terminal rule), and no computational effect: its witness program is simply \( \text{skip} \).

The \text{Read} rule turns a ghost variable \( a \) into a program variable \( y \) (fresh in the original goal). That is, the newly assigned immutable program variable \( y \) is added to the environment of the sub-goal, and all occurrences of \( a \) are substituted by \( y \) in both the pre- and postcondition. As a side-effect, the rule prepends the read statement \text{let} \( y = *(x + i) \) to the remainder of the program to be synthesized.

---

[2] Further in the paper, we will extend the language of heap assertions to support dynamically allocated memory blocks and user-defined inductive predicates.
Structuring the Synthesis of Heap-Manipulating Programs

The rule WRITE allows for writing a symbolic expression \( e \) into a memory cell \( \langle x, i \rangle \), provided all \( e \)’s variables are program-level. Notice that our WRITE rule, unlike the one in classical SL (O’Hearn et al. 2001), requires that a matching points-to heaplet be present in the postcondition, in order to decide which expression \( e \) to write. In the interest of modularity, the rule does not remove the resulting identical heaplets \( \langle x, i \rangle \mapsto e \) from the sub-goal, leaving this to a more general rule FRAME—SSL’s version of Separation Logic’s ‘frame rule’. This rule enables “framing out” a shared sub-heap \( R \) from the pre- and postcondition, as long as this does not create new existential variables. Notice, that unlike the classical SL’s FRAME rule, our version does not require a side condition saying that \( R \) must not contain program variables that are modified by the program \( c \), because all variables in our language are immutable.

Synthesizing swap. Armed with the basic SSL inference rules from Fig. 1, let us revisit our initial example: the swap function (1). Fig. 2 shows the derivation of the program using the rules, and should be read bottom-up. For convenience, we name each subgoal’s witness program, starting from \( c_1 \) (which corresponds to swap’s body). Furthermore, each intermediate sub-goal highlights via gray boxes a part of the specification, which “triggers” the corresponding SSL rule. Intuitively, the goal of the synthesis process is to “empty” the spatial parts of the pre- and the postcondition, so that the derivation can eventually be closed via EMP; to this end, READ and WRITE work together to create matching heaplets between the two assertions, which are then eliminated by FRAME.

![Fig. 1. Simplified basic rules of SSL.](image1)

2.3 Spatial Unification and Backtracking

Now, consider the synthesis goal induced by the following SL specification:

\[
\{ x \mapsto 239 \land y \mapsto 30 \} \text{ void pick(loc } x, \text{ loc } y \} \{ x \mapsto z \land y \mapsto z \}
\]

(2)

Since \( z \) does not appear among the formals or in the precondition, it is treated as an existential. The postcondition thus allows \( x \) and \( y \) to point to any value, as long as it is the same value.

To deal with existentials in the heap, we introduce the rule UNIFYHEAPS, which attempts to find a unifying substitution \( \sigma \) for some sub-heaps of the pre- and the heap.

\[
\text{UNIFYHEAPS } \begin{array}{l}
\sigma \in R' = R \\
\theta \neq \text{dom} (\sigma) \in \text{EV}(\Gamma, \mathcal{P}, Q) \\
\Gamma; \{P \mapsto R\} \Rightarrow [\sigma | \psi; Q \mapsto R'] c \\
\Gamma; \{\psi; P \mapsto R\} \Rightarrow [\psi; Q \mapsto R'] c
\end{array}
\]

Fig. 3. SSL rule for heap unification.

postcondition. The domain of $\sigma$ must only contain existentials. For example, applying UnifyHeaps to the spec (2) with $R \triangleq x \mapsto 239$ and $R' \triangleq x \mapsto z$ results in the substitution $\sigma = [z \mapsto 239]$, and the residual synthesis goal $\{x, y\} \{x \mapsto 239 * y \mapsto 30\} \leadsto \{x \mapsto 239 * y \mapsto 239\}$, which can be now synthesized by using the Frame, Write, and Emp rules.

Due to its freedom to choose a sub-heap and a unifying substitution, UnifyHeaps introduces non-determinism into the synthesis procedure and might require backtracking—a fact also widely observed in interactive verification community (Gonthier et al. 2011; McCreight 2009) wrt. SL assertions. For instance, consider the spec below:

$$\begin{align*}
\{x \mapsto a * y \mapsto b\} \text{void notSure}(\text{loc } x, \text{ loc } y) \{x \mapsto c * c \mapsto 0\}
\end{align*}$$

(3)

One way to approach the spec (3) is to first read from $x$, making $a$ a program-level variable $a2$ (via Read), then use UnifyHeaps and Frame on the $x \mapsto \bullet$ heaplets in the pre-/post-condition, substituting the existential $c$ by $a2$.

That, however, leaves us with an unsolvable goal $\{x, y, a2\} \{y \mapsto b\} \leadsto \{a2 \mapsto 0\}$. Hence we have to backtrack, and instead unify $c$ with $y$, eventually deriving the correct program notSure.

### 2.4 Reasoning with Pure Constraints

So far we have only looked at SL specifications whose *pure* parts were trivially true. Let us now turn our attention to the goals that make use of non-trivial pure boolean assertions.

#### 2.4.1 Preconditions

To leverage pure preconditions, we adopt a number of the traditional Smallfoot-style rules (Berdine et al. 2006), whose SSL counterparts are shown in the top part of Fig. 4. In the nomenclature of Berdine et al., all those rules are *non-operational*, i.e., correspond to constructing the proofs of symbolic heap entailment and involve no programming component. Note that the original rules in Berdine et al. (2005) assume a restricted pure logic with only equalities; we adapt these rules to our logic-agnostic style, relying on the oracle for pure validity instead of original syntactic premises. For instance, the rule SubstLeft makes use of a precondition that implies equality between two universal variables, $x = y$, substituting all occurrences of $x$ in the subgoal by $y$. The rule StarPartial makes explicit the fundamental assumption of SL: disjointness of symbolic heaplets connected by the $*$ operator. Most commonly, this rule’s effect is observable in combination with another rule, Inconsistency, which identifies an inconsistent precondition, and emits an always-failing program *error*.

These three SSL rules can be observed in action via the following examples:

$$\begin{align*}
\{a = x \land y = a; x \mapsto y * y \mapsto z\} \text{ void urk}(\text{loc } x, \text{ loc } y) \{\text{true}; y \mapsto a * x \mapsto y\}
\end{align*}$$

(4)

After applying SubstLeft, the goal transforms to $\{x, y\} \{x \mapsto x * x \mapsto z\} \leadsto \{x \mapsto x * x \mapsto x\}$, which is a trivial, as the precondition requires two *disjoint* points-to heaplets with the same source—a fact, which converted into a pure sub-formula $x \neq x$ by StarPartial (i.e., false), resulting in the never-occurring *error* body (via Inconsistency), as no pre-state would satisfy its precondition.

#### 2.4.2 Postconditions

In the presence of non-trivial pure postconditions, we face the problem of finding suitable terms to substitute for their existentials. This is essentially a program synthesis problem where both the specifications and the solution are terms from the pure logic. This is a challenging, yet well-studied problem (Alur et al. 2013), which we consider orthogonal to our agenda of deriving pointer-manipulating programs. Hence, SSL relies on a *pure synthesis oracle* to pick appropriate instantiations for existentials; formally, the oracle is represented by a non-deterministic

---

3Inconsistent preconditions are unlikely to appear in top-level goals, but frequently arise as a result of a case split, where they signify “dead” (i.e., unreachable) branches.
The first clause states that, if $v \neq y + i'$, $\phi' = \phi \land (x + i \neq y + i')$ is obtained by making a substitution $e^* + P \leadsto \{Q\}|c$.

In the case of logical overlap, the conditions for different clauses are checked in the order the clauses are defined. Following the rule Pick with $\psi$ restricted to program variables—be quite effective, if theoretically incomplete, at discharging such synthesis goals.

As an example, consider the following goal (where $S$ and $S_1$ are finite sets):

$\{S = \{v\} \cup S_1; x \mapsto a\} \textbf{ void elem(loc x, int v)} \{S = \{v_1\} \cup S_1; x \mapsto v_1 + 1\}$

Following the rule UnifyPure, one can unify the two facts about sets in the pre- and the postcondition, obtaining the substitution $[v_1 \mapsto v]$. The rest is accomplished by the rule Write, which emits the only necessary statement for elem’s body: $\ast x = v + 1$.

### 2.5 Synthesis with Inductive Predicates

The real power of Separation Logic stems from its ability to compositionally reason about linked heap-based data structures, such as lists and trees, whose shape is defined recursively via inductive heap predicates. The most traditional example of a data structure defined this way is a linked list segment (Reynolds 2002), whose definition is given by the two-clause predicate below:

$$\text{lseg}(x, y, S) \triangleq x = y \land \{S = \emptyset; \text{emp}\}$$

$$\quad | \ x \neq y \land \{S = \{v\} \cup S_1; [x, 2] * x \mapsto v * (x, 1) \mapsto \text{nxt} * \text{lseg}(\text{nxt}, y, S_1)\}$$

The predicate $\text{lseg}(x, y, S)$ describes a linked list that starts at location $x$, ends at location $y$, and contains a set of elements $S$.

The first clause states that, if $x$ and $y$ are equal, the list contains no elements and occupies an empty heap emp. The complementary second clause postulates the existence of an allocated memory block (or just block) of two consecutive locations rooted at $x ([x, 2] * x \mapsto v * (x, 1) \mapsto \text{nxt})$, such that the first location stores the payload $v$, while the second one stores a pointer nxt to the tail of the list, whose shape is defined recursively as $\text{lseg}(\text{nxt}, y, S_1)$. In general, a predicate definition $D \triangleq p(\overline{x}_i)\langle e_j, \{\chi_j, R_j\}\rangle_{j \in 1, \ldots, N'}$ starts with the name $p$ and a vector or formal parameters $\overline{x}_i$, followed by a sequence of $N$ inductive clauses. The $j$th clause consist of a guard $e_j$—a boolean expression over the predicate’s formals, followed by the clause body—a SL assertion with a spatial part $R_j$ and pure part $\chi_j$, describing the shape of the heap and pure

---

4The predicate can be parametrized by different abstractions: the length of the list, its set/multiset of elements, or even an algebraic list datatype. The choice of abstraction is orthogonal to the story of this paper. In our examples, we chose the set abstraction, which supports an efficient SMT encoding and results in sufficiently strong specifications for synthesis.

5Program-level expressions are a subset of pure logic terms, free from non-executable constructs, such as set operations. In the case of logical overlap, the conditions for different clauses are checked in the order the clauses are defined.

---

constraints, correspondingly. Free variables of the clauses (e.g., v, nxt above) are treated as ghosts or existentials, depending on whether the predicate instance is in a pre- or postcondition of a goal. We require that all predicates be well-founded—i.e., have their recursive applications only on strictly smaller sub-heaps (Brotherston et al. 2008); we ensure well-foundedness via a syntactic check that each recursive clause contains at least one points-to heaplet. From now on, we extend the definition of the goal, with a context Σ, which stores the definitions of inductive predicates and specified functions, which are accessible in the derivation.

2.5.1 Dynamic Memory. In order to support dynamically allocated linked structures, as demonstrated by definition (6), we extend the language of symbolic heaps with two new kinds of assertions: blocks and predicate instances. Symbolic blocks are a well-established way to add to SL support for consecutive memory locations (Brotherston et al. 2017; Jacobs et al. 2011), which are allocated and disposed all together.6 Two SSL rules, Alloc and Free (presented in Sec. 3), make use of blocks, as those appear in the post- and the preconditions of their corresponding goals. Conceptually, Alloc looks for a block in the postcondition rooted at an existential and allocates a block of the same size (by emitting the command \[ \text{let } x = \text{malloc}(n) \]), adding it to the subgoal’s precondition. Free is triggered by an un-matched block in the goal’s precondition, rooted at some program variable x, which it then disposes by emitting the call to \( \text{free}(x) \), removing it from the subgoal’s precondition.

2.5.2 Induction. Let us now synthesize our first recursive heap-manipulating function, a linked list’s destructor \( \text{listfree}(x) \), which expects a linked list starting from its argument \( x \) and ending with the null-pointer, and leaves an empty heap as its result (we explain the meaning of the tag 0 in \( \text{lseg}^0(x,0,S) \) below):

\[
\{\text{lseg}^0(x,0,S)\} \text{ void listfree(loc x) } \{\text{emp}\}
\]  

(7)

The first synthesis step is carried out by the SSL rule INDUCTION. We postpone its formal description until the next section, conveying the basic intuition here. INDUCTION only applies to the initial synthesis goal whose precondition contains an inductive predicate instance, and its effect is to add a new function symbol to the goal’s context, such that an invocation of this function would correspond to a recursive call. In our example (7) INDUCTION extends the context \( \Sigma \) with a “recursive hypothesis” as follows:7

\[
\Sigma_1 \triangleq \Sigma, \text{listfree}(x') : \{\text{lseg}^1(x',0,S')\}\{\text{emp}\}
\]  

(8)

2.5.3 Unfolding Predicates. The top-level rule INDUCTION is complemented by the rule OPEN (defined in Sec. 3), which unfolds a predicate instance in the goal’s precondition according to its definition, and creates a subgoal for each inductive clause. For instance, invoked immediately on our goal (7), it has the following effect on the derivation:

(a) Two sub-goals, one for each of the clauses of \( \text{lseg} \), are generated to solve:

(i) \( \Sigma; \{ x \} : \{ x = 0 \wedge S = \emptyset ; \text{emp} \} \leadsto \{ \text{emp} \} \)

(ii) \( \Sigma; \{ x \} : \{ x \neq 0 \wedge S = \{ v \} \cup S_1 ; [x,2] * x \mapsto v * (x,1) \mapsto \text{nxt} * \text{lseg}^1(\text{nxt},y,S_1) \} \leadsto \{ \text{emp} \} \)

(b) Assuming \( c_1 \) and \( c_2 \) are the programs solving the sub-goals (i) and (ii), the final program is obtained by combining them as \( \text{if } (x = 0) \{ c_1 \} \text{ else } \{ c_2 \} \).

Thus, OPEN performs case-analysis according to the predicate definition. Note how the precondition of each generated sub-goal is refined by the corresponding clause’s guard and body. The resulting sub-programs, once synthesized, are combined with a conditional statement that branches on the

6An alternative would be to adopt an object model with fields, which is more verbose (Berdine et al. 2005).

7In SSL, a context \( \Sigma \) can also store user-provided specifications of auxiliary functions synthesized earlier. We will elaborate on case studies relying on user-provided auxiliary functions in Sec. 6.2.
The second subgoal (ii), contains another instance of the inductive predicate—\textit{lseg}(\textit{nxt}, \textit{y}, \textit{S}_1). In principle, this instance can be unfolded again, yielding in turn a third instance, and so on. To avoid infinite unfolding of predicate instances we introduce \textit{level tags} (natural numbers, ranged over by $\ell$), which now annotate some predicate instances in the pre- and postcondition of the goal and the context functions. All predicates in the initial goal have their level tag set as $\ell = 0$. The rule \textit{OPEN} only applies to instances with $0 \leq \ell < \text{MaxUnfold}$, incrementing their tags (hence, the number of "telescopic" unfoldings is bounded by the global parameter \text{MaxUnfold}).

2.5.4 Recursive calls. Synthesizing recursive programs requires extra care in order to avoid vacuously correct (in the sense of partial program correctness) \textit{non-terminating} programs that simply call themselves. To ensure termination of synthesized programs, we require that the pre-heap $P_f$ of any recursive call be strictly smaller than the pre-heap $P$ of the top-level synthesis goal. Following ideas from Cyclic Termination Proofs (Brotherston et al. 2012), we enforce this restriction syntactically: since all inductive predicates are well-founded, it is sufficient to check that one of the predicate instances in $P_f$ has been derived from the corresponding instance in $P$ by unfolding. To perform this check, we piggy-back on level tags: intuitively, if every level tag in $P$ starts out as 0 and is only incremented as a result of \textit{OPEN}, a recursive call on an instance with $\ell > 0$ is guaranteed to terminate.

To see how tag control function calls, consider the rule \textit{CALL} in Fig. 5. It fires when the goal contains in its precondition a symbolic sub-heap $R$, which can be unified with the precondition $P_f$ of a function symbol $f$ from the goal’s context $\Sigma$. This unification is similar to the effect of $\text{UNIFYHEAP}$, with the difference that \textit{CALL} takes level tags into the account (reflected in the tag-aware equality predicate $\equiv^\ell$), while $\text{UNIFYHEAP}$ (Fig. 3) and $\text{FRAME}$ (Fig. 1) ignore them when comparing sub-heaps for equality.

Returning to the \textit{listfree} example, the remaining subgoal (ii)

\begin{align*}
\{x\} & ; \ x \neq 0 \land S = \{v\} \cup S_1; [x, 2] \ast x \mapsto v \ast \langle x, 1 \rangle \mapsto \textit{nxt} \ast \textit{lseg}^1(\textit{nxt}, \textit{y}, \textit{S}_1) \} \leadsto \{\text{emp}\} \tag{9}
\end{align*}

can be now transformed, via \textit{READ} (focused on $\langle x, 1 \rangle \mapsto \textit{nxt}$), into

\begin{align*}
\{x, \textit{nxt2}\} & ; \ x \neq 0 \land S = \{v\} \cup S_1; [x, 2] \ast x \mapsto v \ast \langle x, 1 \rangle \mapsto \textit{nxt2} \ast \textit{lseg}^1(\textit{nxt2}, \textit{y}, \textit{S}_1) \} \leadsto \{\text{emp}\} \tag{10}
\end{align*}

The grayed fragment in (10) can now be unified with the precondition of \textit{listfree} (8) following \textit{CALL}'s premise. As the tags match (both indicate the \textit{first} unfolding of the predicate), unification succeeds with the substitution $\sigma = [x' \mapsto \textit{nxt2}, S' \mapsto S_1]$ from $f$’s parameters and ghosts to the goal variables. The same rule produces, from the $f$’s postcondition, a new symbolic heap $R'$, which replaces the targeted fragment in the precondition, recording the effect

\begin{verbatim}
void listfree(loc x) {
    if ((x = 0) () else {
        let nxt2 = *(x + 1);
        listfree(nxt2);
        free(x);
    }
}
\end{verbatim}

Fig. 6. Synthesized \textit{listfree} (7).
of the call. In this example, the function’s postcondition is \{\text{emp}\}, so the goal becomes:

\[
\{x, \text{nxt}2\}; \{x \neq 0 \land S = \{v\} \cup S_1; [x, 2] \ast x \mapsto v \ast (x, 1) \mapsto \text{nxt}2 \ast \text{emp}\} \leadsto \{\text{emp}\}
\] (11)

The remaining steps are carried out by the rule \text{Free}, followed by \text{Emp}, with the former disposing the remaining block, thus, completing the derivation with program \text{listfree} shown in Fig. 6.

### 2.5.5 Unfolding in the postcondition

Whereas \text{Open} unfolds predicate instances in a goal’s precondition, a complementary rule \text{Close} (Fig. 5) performs a similar operation on the goal’s postcondition. The main difference is that instead of performing a case-split and emitting several subgoals, \text{Close} non-deterministically picks a single clause \(k\) from the predicate’s definition (the intuition being that the required case split has already been performed by \text{Open}). Upon unfolding, the clause’s adapted guard \([\sigma]e_k\) and pure part \([\sigma]x_k\) are added to the subgoal’s postcondition, while its spatial part also gets its level tags increased by one \(\lceil \lceil \lceil \sigma \rceil \rceil_{\ell+1}\), in order to account for the depth of unfoldings.

To showcase the use of \text{Close}, let us define a new predicate for a linked null-terminating structure \lseg{2}, which stores in each node the payload \(v\) and \(v + 1\):

\[
lseg{2}(x, S) \triangleq x = 0 \land \{S = \emptyset; \text{emp}\}
\]  
\[
| \quad x \neq 0 \land \{S = \{v\} \cup S_1; [x, 3] \ast x \mapsto v \ast (x, 1) \mapsto v + 1 \ast (x, 2) \mapsto \text{nxt} \ast \lseg{2}\}
\] (12)

We now synthesize an implementation for the following specification, requiring to morph a regular list \lseg{x, 0, S} to \lseg{y, S}, both parameterized by the same set \(S\):

\[
\{r \mapsto 0 \ast \lseg{0}(x, 0, S)\} \quad \text{void listmorph(loc x, loc r)} \{r \mapsto y \ast \lseg{0}(y, S)\}
\] (13)

The derivation starts with \text{Induction}, which produces the following function symbol (note that the postcondition always has all level tags \textit{erased} in order to prevent chaining of recursive calls):

\[
\text{listmorph}(x', r') : \{r' \mapsto 0 \ast \lseg{1}(x', 0, S')\}\{r' \mapsto y' \ast \lseg{2}(y', S')\}
\] (14)

Next, we \text{Open} \lseg{x, 0, S}, producing two sub-goals. The first one:

\[
\{x, r\}; \{S = \emptyset \land x = 0; r \mapsto 0\} \leadsto \{r \mapsto y \ast \lseg{0}(y, S)\}
\] (15)

is easy to solve via \text{Close}, which should pick the first clause from \lseg{2}’s definition (12) (corresponding to \text{emp}), followed by \text{Frame} to \(r \mapsto y\) in the postcondition. The second subgoal, after having read the value of \(x + 1\) into a program variable \text{nxt}2, looks as follows:

\[
\{x, r, \text{nxt}2\};
\[
\{S = \{v\} \cup S_1 \land x \neq 0; r \mapsto 0 \ast [x, 2] \ast x \mapsto v \ast (x, 1) \mapsto \text{nxt}2 \ast \lseg{1}(\text{nxt}2, 0, S_1)\} \leadsto \{r \mapsto y \ast \lseg{2}(y, S)\}
\] (16)

Now, \text{Close} comes to the rescue, unfolding the grayed instance in (16)’s postcondition:

\[
\{x, r, \text{nxt}2\};
\[
\{S = \{v\} \cup S_1 \land x \neq 0; r \mapsto 0 \ast [x, 2] \ast x \mapsto v \ast (x, 1) \mapsto \text{nxt}2 \ast \lseg{1}(\text{nxt}2, 0, S_1)\} \leadsto \{r \mapsto y \ast \lseg{0}(y, S)\}
\] (17)

\[
\{S = \{v_1\} \cup S_2; r \mapsto y \ast [y, 3] \ast y \mapsto v_1 \ast (y, 1) \mapsto v_1 + 1 \ast (y, 2) \mapsto \text{nxt}1 \ast \lseg{2}(\text{nxt}1, S_2)\}
\]
We can now use the CALL rule, unifying the precondition of the induction hypothesis (14) with the grayed parts in the goal (17), obtaining the following subgoal:

\[
\begin{align*}
S &= \{v_1\} \cup S_2; r \mapsto y *\left[\begin{array}{l}
\vdots
\end{array}\right] y \mapsto v_1 *\left[\begin{array}{l}
\vdots
\end{array}\right] v_1 + 1 *\left[\begin{array}{l}
\vdots
\end{array}\right] lseg2(nxt1, S_2) \\
&\sim
\end{align*}
\]

The instances of \( \text{lseg}2 \) in the pre- and the postcondition can now be unified via UnifyHeaps, instantiating \([\text{nxt1} \mapsto y_1, S_2 \mapsto S_1]\), and then framed via Frame. The remaining derivation is done by Reading from \( r \) and \( x \), subsequent disposing of a two-cell block (grayed in the precondition) and allocation of a three-cell block in order to match the grayed block in the postcondition. Finally, the exact payload for cells of the newly-allocated 3-pointer block is determined by unifying the set assertions in the pure pre- and postcondition (via UnifyPure), and then Write records the right values to satisfy the constraints imposed for the head of \( \text{lseg}2 \)-like list by Definition (12). The resulted synthesized implementation of listmorph is shown in Fig. 7.

### 2.6 Enabling Procedure Calls by Means of Call Abduction

We conclude this overview with one last example—a recursive procedure for copying a linked list:

\[
\begin{align*}
\{r \mapsto x *\left[\begin{array}{l}
lseg0(x, 0, S)\end{array}\right]\} &\textbf{void} \ \text{listcopy}(\text{loc} \ r) \ \{r \mapsto y *\left[\begin{array}{l}
lseg0(x, 0, S) *\left[\begin{array}{l}
lseg0(y, 0, S)\end{array}\right]\end{array}\right]\} \\
&\text{(18)}
\end{align*}
\]

To make things more fun, we pass the pointer to the head of the list via another pointer \( r \), which is also used to record the result of the function—an address \( y \) of a freshly allocated list copy. The synthesis begins by using Induction, producing the function symbol

\[
\textbf{void} \ \text{listcopy}(\text{loc} \ r'): \ \{r' \mapsto x' *\left[\begin{array}{l}
lseg1(x', 0, S')\end{array}\right]\} \{r' \mapsto y' *\left[\begin{array}{l}
lseg1(x', 0, S') *\left[\begin{array}{l}
lseg1(y', 0, S')\end{array}\right]\end{array}\right]\} \\
&\text{(19)}
\]

It follows by Read (from \( r \) into \( x2 \)) and Open, resulting in two subgoals, the first of which (an empty list) is trivial. The synthesis proceeds, reading from \( x2 \) into \( v2 \) and from \( x2 + 1 \) into \( \text{nxt2} \), so after using Close (on either of \( \text{lseg}0(x, 0, S) \) heaplets) in the postcondition, UnifyHeaps and Frame, we reach the following subgoal:

\[
\begin{align*}
\{x, r, x2, v2, \text{nxt2}2\}; \ \{S = \{v2\} \cup S_1 \wedge x2 \neq 0; r \mapsto x2 *\left[\begin{array}{l}
lseg1(\text{nxt2}, 0, S_1)\end{array}\right]\} \\
&\sim
\end{align*}
\]

\[
\begin{align*}
\{S = \{v2\} \cup S_2; r \mapsto y *\left[\begin{array}{l}
lseg1(\text{nxt2}, 0, S_2) *\left[\begin{array}{l}
lseg0(y, 0, S)\end{array}\right]\end{array}\right]\} \\
&\sim
\end{align*}
\]

\[
\text{(20)}
\]

At this point of our derivation, we run into an issue. Ideally, we would like to use the grayed fragment of the goal (20)’s precondition, to fire the rule CALL with the spec (19), i.e. to make a recursive call on the tail list. However, the (19)’s precondition requires \( r \) to point to the start of that list \( (\text{nxt2}) \), whereas in our case it still points to the start of the original list \( (x2) \).

Any programmer would know a solution to this conundrum: we have to to write \( \text{nxt2} \) into \( r \), in order to provide a suitable symbolic heap to make a recursive call. The Write rule, however, cannot make this change on its own, since its application is guided by the desired post-heap (in order to decide what to write), which in this case is determined by the callee’s precondition and not by the current goal. To overcome this limitation, we introduce a novel rule \text{AbduceCall} (shown in Fig. 8), which emits a separate synthesis sub-goal to \emph{prepare} the symbolic pre-heap for the recursive call.
First, the rule splits the precondition of the candidate callee $F$ from $\Sigma$ into two symbolic sub-heaps, $P_F$ and $F_F$, such that all predicate instances are in $P_F$, while the rest of the heaplets (i.e., blocks and points-to assertions) are in $F_F$. Next, it tries to unify $P_F$ with some sub-heap $P$ from the goal’s precondition, finding a suitable substitution $\sigma$, such that $P = [\sigma]P_F$. While doing so, it does not account for the “remainder” $[\sigma]F_F$, which might not be immediately matched by anything in the goal’s precondition. In order to make it match, the rule emits a sub-goal $\Sigma; F; \{\phi; F\} \leadsto \{\phi, F'\}|c_1$, whose purpose is to synthesize a program $c_1$, which will serve as an impedance matcher between some symbolic subheap $F$ from the original goal’s precondition and $F' = [\sigma]F_F$.\footnote{Our implementation is smarter than that: it ensures that $F$ and $F'$ have the same shape and differ only in pointers’ values.}

For instance, in the specification (19), $P_F = \text{lseg}^4(x', 0, S')$ and $F_F = r' \mapsto x'$, so an attempt to unify the former with the predicate instance in the grayed fragment of the goal (20) results in the substitution $\sigma = [x' \mapsto \text{nxt}2, S' \mapsto S_1]$. Applying it to the remainder of the function’s precondition, we obtain $F' = [\sigma]F_F = r' \mapsto \text{nxt}2$. One of the candidates for the role of $F$ from the goal’s precondition is the heaplet $r \mapsto x_2$, so the corresponding subgoal will be of the form $\{r, \ldots \}; \{\ldots; r \mapsto x_2\} \leadsto \{r' \mapsto \text{nxt}2\}$, which will produce the write $\star r = \text{nxt}2$. Fig. 9 shows the eventually synthesized implementation, with the abduced call-enabling write on line 6.

3 SYNTHETIC SEPARATION LOGIC IN A NUTSHELL

Having shown SSL in action, we now proceed with giving a complete set of its inference rules, along with statements of the formal guarantees SSL provides \textit{wrt.} synthesized imperative programs.

The syntax for the imperative language supported by SSL is given in Fig. 10. The set of values includes at least true and false, integers, and locations (isomorphic to non-negative integers). Expressions include variables, values, boolean equality checks and additional theory-specific expressions (e.g., integer arithmetic). The language of commands does not include loops, which are modelled via recursive procedure calls ($f(\overline{e})$). For simplicity we do not provide a mechanism to return a variable from a procedure (so the language is missing the \textit{return} command); result-returning discipline for a procedure can be encoded via passing a result-storing additional pointer, as, e.g., in Example (18). A function dictionary $\Delta$ is simply a list of function definitions of the form $f(\overline{\alpha})$ \{ $c$ \}. For brevity, we omit the type annotations (e.g., in function signatures) in the formalism.

The complete syntax of SSL assertions is shown in Fig. 11, and their meaning was explained in detail throughout Sec. 2. We only notice here that, syntactically, pure logic terms are a superset of program-level expressions $e$, and can contain theory-specific non-executable constructs, such as set operations. In a SSL assertion $\{\phi, P\}$, the term $\phi$ must be \textit{boolean}; our implementation enforces this through a simple type system, which we omit from the formalization for simplicity.

3.1 The Zoo of SSL Rules

Fig. 12 presents \textit{all} rules of core SSL.\footnote{Sec. 5 extends core SSL with several optimizations.} Since most of them have already made an appearance in Sec. 2, here, we only elaborate on the new ones, and highlight some important aspects of their interaction. It is convenient to split the set of rules into the following six categories:
C1 *Top-level rules* are represented by just one rule: **INDUCTION.** This rule is only applicable at the very first stage of the derivation, and it produces a function symbol \( f \), with the specification identical to the top-level goal, modulo the level tag of a predicate instance \( p^0(\vec{e}) \) in the precondition. In the case of *several* predicate instances \( p^0(\vec{e}) \) in the goal’s precondition, the rule non-deterministically picks one recursion scheme, whereas other predicate instances in \( f \)’s precondition are “sealed” via tag erasure \([P]\).

C2 *Terminals* include **EMP** and **INCONSISTENCY.** These rules conclude a successful derivation by emitting **skip** or **error**, respectively.

C3 *Normalization rules* include **NULLNOTLVAL**, **SUBSTLEFT**, **STARPARTIAL**, and **READ.** The role of these rules is to normalize the precondition: eliminate ghosts and equal variables, and make explicit the assumptions encoded in the spatial part. As we discuss in Sec. 5.1, a characteristic feature of those rules is that they cannot cause their subderivation to fail.

C4 *Unfolding rules* are the rules targeted specifically to predicate instances. The four basic unfolding rules are **OPEN**, **CLOSE**, **ABDUCECALL**, and **CALL.** As hinted before, the rule **OPEN** picks an instance \( p^\ell(\vec{e}) \) in the goal’s precondition, and produces a number of subgoals, emitting a conditional that accounts for all of \( p \)’s clauses. **CLOSE** acts symmetrically on the postcondition. Both rules increment tags for the instances in the clause bodies, and are not applicable for \( \ell \geq \text{MaxUnfold} \). Having the search depth bounded by the global parameter MaxUnfold affects the completeness but not soundness of SSL; MaxUnfold \( \leq 2 \) suffices for all examples in the paper. The rule **ABDUCECALL** prepares the heap for a call application (via **CALL**), as described in Sec. 2.6. Note that the rules **UNIFYHEAPS** and **FRAME** have been slightly generalized w.r.t. to what was shown in Fig. 1 and Fig. 3, and now include a parametric premise frameable. If we define frameable to return true only for predicate instances, we obtain *unfolding versions* of these rules, which deal specifically with predicate instances.

C5 *Flat* rules deal with with the flat part of the heap (blocks and points-to heaplets). They include **ALLOC**, **FREE**,\(^{10}\) **WRITE**, as well as the flat versions of **UNIFYHEAPS** and **FRAME**, with frameable defined to return true for flat heaplets.

C6 *Pure synthesis* rules are responsible for instantiating existentials. In Fig. 12 this category is represented by a single non-deterministic rule **PICK**, which replaces an existential with an arbitrary pure term \( \psi \). Note that the choice of \( \psi \) does not affect the soundness of SSL, but enumerating all possible terms would make the proof search intractable. In practice, we assume that \( \psi \) is suggested by a pure synthesis oracle. We describe a concrete oracle used in our implementation in Sec. 5.6.

\(^{10}\)The iterated \(*\) operator in the premises of **ALLOC** and **FREE** is a syntactic sugar for records of fixed size.
Fig. 12. All core SSL rules. Grayed parts are parameters; instantiating them differently yields different rules.
Let us refer to rules that emit a sub-program as operational and to the rest (i.e., to the rules that only change the assertions) as non-operational. The rules from Fig. 12 form the basis of SSL as a proof system, allowing for possible extensions for the sake of optimization or handling pure constraints. We make them intentionally declarative rather than algorithmic, which is essential for establishing the logic’s soundness, leaving a lot of freedom for possible implementations. Such decisions have to be made, for instance, when engineering an implementation of ABDUCECALL or Pick. The algorithmic aspects of SSL, e.g., non-deterministic choice of a frame or a unifying substitution, are handled by the procedures from Sec. 4.

3.2 Formal Guarantees for the Synthesized Programs

The programs resulting from the SSL derivations enjoy both (a) validity (i.e., they obey their ascribed Hoare-style specifications in a sence of partial correctness) and (b) termination. Therefore, the soundness result for SSL entails the Hoare-style total correctness. This result is formally stated by Theorem 3.6, which builds on the two components, establishing the validity and termination separately, as we describe in Sec. 3.2.1 and 3.2.2.

Preliminaries. The definition of the operational semantics of the SSL language (Fig. 10) follows the standard RAM model. Heaps (ranged over by $h$) are represented as partial finite maps from pointers to values, with support for pointer arithmetic (via offsets). A function call is executed within its own stack frame $(c, s)$, where $c$ is the next command to reduce and $s$ is a store recording the values of the function’s local variables and parameters. A stack $S$ is a sequence of stack frames, and a configuration is a pair of a heap and a stack. The small-step operational semantics relates a function dictionary $\Delta$, and a pair of configurations: $\Delta; (h, S) \rightsquigarrow (h', S')$, with $\rightsquigarrow^*$ meaning its reflexive-transitive closure. We elide the transition rules for brevity; similar rules can by found, e.g., in the work by Rowe and Brotherston (2017).

3.2.1 Validity. Let us denote the valuation of an expression $e$ under a store $s$ as $[e]_s$. $I$ ranges over interpretations— mappings from user-provided predicates $D$ to the relations on heaps and vectors of values. To formally define the validity of Hoare-style specs in SSL, we use the standard definition of the satisfaction relation $\models_I^\Sigma$ on pairs of heaps and stores, contexts, interpretations, and SSL assertions without ghosts. For instance, the following SSL definitions are traditional for Separation Logics with interpreted predicates (Berdine et al. 2005; Nguyen et al. 2007):

- $\langle h, s \rangle \models_I^\Sigma \{ \phi; \text{emp} \} \text{ iff } [\phi]_s = \text{true and } \text{dom}(h) = \emptyset$.
- $\langle h, s \rangle \models_I^\Sigma \{ \phi; [x, n] \} \text{ iff } [\phi]_s = \text{true and } \text{dom}(h) = \emptyset$.
- $\langle h, s \rangle \models_I^\Sigma \{ \phi; (e_1, i) \mapsto e_2 \} \text{ iff } [\phi]_s = \text{true and } \text{dom}(h) = [e_1]_s + i$ and $h([e_1]_s + i) = [e_2]_s$.
- $\langle h, s \rangle \models_I^\Sigma \{ \phi; P_1 * P_2 \} \text{ iff } \exists h_1, h_2, h = h_1 \cup h_2 \text{ and } \langle h_1, s \rangle \models_I^\Sigma \{ \phi; P_1 \} \text{ and } \langle h_2, s \rangle \models_I^\Sigma \{ \phi; P_2 \}$.
- $\langle h, s \rangle \models_I^\Sigma \{ \phi; p(\chi_i) \} \text{ iff } [\phi]_s = \text{true and } D \triangleq p(\chi_i)(e_j, \{ \chi_j, R_j \}) \in \Sigma \text{ and } \langle h, [\chi_i]_s \rangle \in I(D)$.

Therefore, blocks have no spatial meaning, except for serving as an indicator on the memory fragments that we allocated and can be disposed.

The notion of Hoare-style validity is standard and is given by Definition 3.1.

Definition 3.1 (Validity). We say that a Hoare-style specification $\Sigma; \Gamma; \{ P \} \ c \ \{ Q \}$ is valid wrt. the function dictionary $\Delta$ if whenever $\text{dom}(s) = \Gamma$, $\exists \sigma_vy = [x_i \mapsto d_i]_{x_i \in \text{GV}(\Gamma, P, Q)}$ such that $\langle h, s \rangle \models_I^\Sigma [\sigma_vy]\{ P \}$, and $\Delta; (h, (c, s) \cdot e) \rightsquigarrow^* (h', (\text{skip}, s') \cdot e)$, it is also the case that $\langle h', s' \rangle \models_I^\Sigma [\sigma_vy \cup \sigma_vy]Q$ for some $\sigma_vy = [y_j \mapsto d_j]_{y_j \in \text{EV}(\Gamma, P, Q)}$.

That is, the definition assumes an initial stack frame $(c, s)$ and a concrete heap $h$, both consistent with an environment $\Gamma$ and the precondition $P$ (with $\sigma_vy$ providing concrete values for $P$’s ghost
variables). In the case if the program terminates with the store \( s' \) and a heap \( h' \), they satisfy the postcondition \( Q \), with \( \sigma_{ev} \) instantiating \( Q \)'s existentials.

Proving validity in the absence of auxiliary and recursive function calls would be almost trivial, thanks to similarities between SSL and traditional Separation Logic. Unfortunately, dealing with (self-)recursion requires a well-founded inductive argument when reasoning about validity of nested function calls in a verified procedure.\(^{11}\) Therefore, we conduct our proof by well-founded induction, establishing validity of a synthesized top-level recursive procedure in an assumption that all of its recursive sub-calls are valid and are “smaller” in some sense. To equip ourselves for such an argument, we provide a series of auxiliary definitions, the key one being of sized specification validity, which relies on the notion of a heap size \(|h| \triangleq |\text{dom}(h)|\):

**Definition 3.2 (Sized validity).** We say a specification \( \Sigma; \Gamma; \{\mathcal{P}\} \vdash \{Q\} \) is \( n \)-valid wrt. the function dictionary \( \Delta \) whenever for any \( h, h', s, s' \), whenever (a) \(|h| \leq n \), (b) \( \Delta; \langle h, (c, s) \cdot e \rangle \sim^* \langle h', (\text{skip}, s') \cdot e \rangle \), and (c) \( \text{dom}(s) = \Gamma \) and \( \exists \sigma_{ev} = [x_i \mapsto d_i]_{x_i \in \text{EV}(\Gamma, \mathcal{P}, Q)} \) such that \( \langle h, s \rangle \models \Sigma [\sigma_{ev}] \mathcal{P} \), it is the case that \( \exists \sigma_{ev} = [y_j \mapsto d_j]_{y_j \in \text{EV}(\Gamma, \mathcal{P}, Q)} \), such that \( \langle h', s' \rangle \models \Sigma [\sigma_{ev} \cup \sigma_{ev}] Q \).

Definition 3.2 is rather peculiar in that it defines a standard Hoare-style soundness wrt. pre/postcondition, while doing that only for heaps of size smaller than \( n \). This is an important requirement to stage a well-founded inductive argument for our soundness proof of SSL-based synthesis in the presence of recursive calls. By introducing the explicit sizes to the definition of validity, we can make sure that we only deal with calls on strictly decreasing subheaps wrt. the heap size when invoking functions (auxiliary ones or recursive self). To make full use of this idea, and also account for the possibility of having auxiliary functions, we lift Definition 3.2 to contexts, stratifying the shape of function dictionaries \( \mathcal{D} \)’s specifications.

**Definition 3.3 (Coherence).** A dictionary \( \Delta \) is \( n \)-coherent wrt. a context \( \Sigma (\text{coh} (\Delta, \Sigma, n)) \) iff

- \( \Delta = \varepsilon \) and \( \text{functions}(\Sigma) = \varepsilon \), or
- \( \Delta = \Delta', \mathcal{f} \{ \bar{t}_i \bar{x}_i \} \{ c \}, \) and \( \Sigma = \Sigma', \mathcal{f} \{ \bar{x} \} : \{\mathcal{P}\} \vdash \{Q\}, \) and \( \text{coh} (\Delta', \Sigma', n) \), and \( \Sigma'; \{\bar{x}\} : \{\mathcal{P}\} \vdash \{Q\} \) is \( n \)-valid wrt. \( \Delta' \), or
- \( \Delta = \Delta', \mathcal{f} \{ \bar{t}_i \bar{x}_i \} \{ c \}, \) and \( \Sigma = \Sigma', \mathcal{f} \{ \bar{x} \} : \{\psi; [\mathcal{P}] + \mathcal{P}(\bar{e})\} \{\{Q\}\}, \) and \( \text{coh} (\Delta', \Sigma', n) \), and \( \Sigma; \{\bar{x}\} : \{[\mathcal{P}] + \mathcal{P}(\bar{e})\} \{\{Q\}\} \) is \( n \)-valid wrt. \( \Delta \) for all \( n' < n \).

That is, coherence is defined inductively on the dictionary/context shape (regarding specified functions and ignoring predicate definitions). A possibility of a (single) recursive definition \( f \) is taken into the account in its last option. In that last clause, recursive calls to the function \( f \) from \( \Delta \) may only take place on heaps of size strictly smaller than \( n \), whereas there is no such restriction for the calls to user-provided auxiliary functions, that can be invoked on heaps of sizes up to \( n \).

Validity of synthesized programs is stated by Lemma 3.4, which defines validity of the synthesized program for any size of the input heap \( n \), assuming that the validity of all recursive calls of the synthesized programs is only established for \( n' < n \), where the case \( n = 0 \) means no calls take place, so it forms the base of the inductive reasoning.

**Lemma 3.4 (Sized validity of SSL-derived programs).** For any \( n, \Delta' \), if

(i) \( \Sigma'; \Gamma; \{\mathcal{P}\} \rightarrow \{Q\} \) \( c \) for a goal named \( f \) with formal parameters \( \Gamma \triangleq \bar{x} \), and
(ii) \( \Sigma' \) is such that \( \text{coh} (\Delta', \Sigma', n) \), and
(iii) for all \( p^0(\bar{e}), \psi; \mathcal{P} \), such that \( \{\mathcal{P}\} = \{\psi; p^0(\bar{e}) \cdot \mathcal{P}\} \), taking \( \mathcal{F} \triangleq f(\bar{e}) : \{\psi; p^1(\bar{e}) \cdot [\mathcal{P}]\} \{\{Q\}\}, \Sigma', \mathcal{F}; \Gamma; \{\mathcal{P}\} \vdash \{Q\} \) is \( n' \)-valid for all \( n' < n \) wrt. \( \Delta \triangleq \Delta', f (\bar{t}_i \bar{x}_i \} \{ c \}, \) then \( \Sigma'; \Gamma; \{\mathcal{P}\} \vdash \{Q\} \) is \( n \)-valid wrt. \( \Delta \).

\(^{11}\)This challenge is well-known when reasoning about correctness of recursive programs (Le and Hobor 2018).

Proof. By the top-level induction on \( n \) and by inner induction on the structure of derivation \( \Sigma'; \Gamma; \{P\} \leadsto \{Q\} | c \). We refer the reader to Appendix A of the extended version of the paper (Polikarpova and Sergey 2018) for the details of the proof.

The assumptions (ii) and (iii) postulate the validity of calls to a function (auxiliary and the one being synthesized) on strictly smaller heaps, akin to the generalized principle of mathematical induction. Overall, Lemma 3.4 states that a program derived via SSL constitutes a valid spec with its goal (i.e., all writes, reads and deallocations in it are safe wrt. accessing heap pointers), assuming that recursive calls, if present, are made on reduced sub-heaps.

3.2.2 Termination. The technique we used for stating and proving Lemma 3.4—allowing for safe calls done only on smaller sub-heaps—is reminiscent to the one employed in size-change termination-based analyses (Lee et al. 2001), which ensure that every infinite sequence of calls would cause infinite descent of some values, leading to the following result.

**Lemma 3.5 (Termination of SSL-synthesized programs).** A program, which is derived via SSL rules from a spec that uses only well-founded predicates, terminates, assuming any of its auxiliary functions terminates.

Proof. The only source of non-termination is self-recursive calls. Due to the tagging discipline, every recursive self-call is applicable after opening a well-founded instance predicate, and hence is done on a smaller sub-heap. Both self-recursive and auxiliary function calls erase the tags in the post-heap, preventing the “expansion” of the symbolic heaps that can be used for further calls.

The argument for termination, supplied by the tagging discipline and used in the proof of Lemma 3.5, is somewhat similar to the syntactic termination criterion, used in proof assistants such as Coq (Coq Development Team 2018). In fact, by parameterizing inductive predicates with pure algebraic data types for describing the heap contents (e.g., using algebraic lists instead of sets in the definition (6) of lseg), we could have employed them as a termination measure instead. We will explore this opportunity in the future work.

3.2.3 Soundness. We conclude this section with the main SSL soundness result. We say that a function dictionary \( \Delta \) is consistent with a context \( \Sigma \) iff any function \( f \in \Delta \) is valid wrt. a corresponding specification in \( \Sigma \) with \( f \)’s formals.

**Theorem 3.6.** If \( \Sigma; \Gamma; \{P\} \leadsto \{Q\} | c \) for a goal named \( f \) with formal parameters \( \Gamma \triangleq \Pi_i \), then:

(a) \( \Sigma; \Gamma; \{P\} c \{Q\} \) is valid wrt. any function dictionary \( \Delta \), consistent with \( \Sigma \), and
(b) \( c \) terminates for any input store-heap pair, satisfying its precondition.

Proof. Follows immediately from Lemma 3.4 taken for a heap of any size, and Lemma 3.5.

4 BASIC SSL-POWERED SYNTHESIS ALGORITHM

In this section we show how to turn SSL from a declaratively defined inference system (Fig. 12) into an algorithm for deriving provably correct imperative programs. The core algorithm is a fairly standard goal-directed backtracking proof search (Kneuss et al. 2013; Mellish and Hardy 1984).

**Encoding Rules and Derivations.** The display on the right shows an algorithmic representation of SSL derivations and rules. To account for the top-level goal (which mandates the program synthesizer to generate a runnable function), we include the function name \( f \)

12We thank François Pottier for this suggestion.
Algorithm 4.1: synthesize (G : Goal, rules : Rule\*).

<table>
<thead>
<tr>
<th>Input:</th>
<th>Goal G = (f, Σ, Γ, {P}, {Q})</th>
</tr>
</thead>
<tbody>
<tr>
<td>List rules of available rules to try</td>
<td></td>
</tr>
<tr>
<td>Result:</td>
<td>Terminating program c, such that Σ, Γ, {P} \in {Q} is valid</td>
</tr>
</tbody>
</table>

1. function synthesize (G, rules) =
2. withRules(rules, G)
3. function withRules (rs, G) =
4. match rs
5. case [] ⇒ Fail
6. case R :: rs' ⇒
7. match filterComm (R(G))
8. case ⊥ ⇒ withRules(rs')
9. case subderivs ⇒
10. tryAlts(subderivs, R, rs', G)

12. function tryAlts (derivs, R, rs, G) =
13. match derivs
14. case [] ⇒ if isInvert(R) then Fail else withRules(rs, G)
15. case \langle goals, \mathcal{K} :: derivs' ⇒
16. match solveSubgoals(goals, \mathcal{K})
17. case Fail ⇒ tryAlts(derivs', R, rs, G)
18. case c ⇒ c
19. function solveSubgoals (goals, \mathcal{K}) =
20. cs := []
21. pickRules = λG. phasesEnabled ? nextRules(G) : AllRules
22. for G ← goals; c = synthesize(G, pickRules(G)); c ≠ Fail do
23. cs := cs ++ [c]
24. if |cs| < |goals| then Fail else \mathcal{K}(cs)

The algorithm. The pseudocode of our synthesis procedure is depicted by Algorithm 4.1. Let us ignore the grayed fragments in the pseudocode for now and agree to interpret the code as if they were not present. Those fragments correspond to optimizations, which we describe in detail in Sec. 5. The algorithm is represented by four mutually-recursive functions:

- synthesize (G, rules) is invoked initially on a top-level goal, with rules instantiated with AllRules—all rules from Fig. 12. It immediately passes control to the first auxiliary function, withRules.
- withRules (rs, G) iterates through the list rs of remaining rules, trying to apply each one to the goal G. If a rule R is applicable, it emits a set of alternative sub-derivations subderivs, which are passed to tryAlts. Otherwise \mathcal{R}(G) returns ⊥, and the next rule form the list is attempted, until no more rules remain (line 5), at which point the synthesis for the current goal fails.
- tryAlts (derivs, R, rs, G) recursively processes the list of alternative sub-derivations derivs, generated by the rule R. If the list is exhausted (line 14), withRules is invoked to try the rest of the rules rs. Otherwise, solveSubgoals is invoked for an alternative to solve all its sub-goals goals and apply the continuation \mathcal{K}. In the case of success (line 18), the resulting program c is returned.
solveSubgoals \( (goals, K) \) tries to solve all subgoals given to it, by invoking synthesize recursively with a suitable (full) list of rules, essentially, restarting the search problem “one level deeper” into the derivation. Unless some of the goals failed, their results are combined via \( K \).

The algorithm explores the space of all valid SSL derivations rooted at the initial synthesis goal. The search proceeds in a depth-first manner: it starts from the root (the initial goal) and always extends the current incomplete derivation by applying a rule to its leftmost open leaf (i.e., a leaf that is not a terminal application). The algorithm terminates when the derivation is complete, i.e., it has no open leaves.

5 Optimizations and Extensions

The basic synthesis algorithm presented in Sec. 4 is a naïve backtracking search in the space of all valid SSL derivations. Note that this is already an improvement over a blind search in the space of all programs: some incorrect programs are excluded from consideration a-priori, such as, e.g., programs that read from unallocated heap locations. In this section we show how to further prune the search space by identifying unsolvable goals and avoiding their exploration.

5.1 Invertible Rules

Our first optimization relies on a well-known fact from proof theory (Liang and Miller 2009) that certain proof rules are invertible: applying such a rule to any derivable goal produces a goal that is still derivable. In other words, applying invertible rules eagerly without backtracking does not affect completeness. Algorithm 4.1 leverages this fact in line 14: when all sub-derivations of an invertible rule \( R \) fail, the algorithm need not backtrack and try other rules, since the failure cannot be due to \( R \) and must have been caused by a choice made earlier in the search.

In SSL, the normalization rules—Read, StarPartial, NullNotLVal, and SubstLeft—are invertible. The effect of these rules on the goal is either to change a ghost into a program-level variable or to strengthen the precondition; no rule that is applicable to the original goal can become inapplicable as a result of this modification, which is confirmed by inspection of all rules in Fig. 12.

5.2 Multi-Phased Search

Among the rules of SSL described in Sec. 3, the unfolding rules are focused on transforming (and eventually eliminating) instances of inductive predicates, while flat rules are focused on other types of heaplets (i.e., points-to and blocks). We observe that if the unfolding rules failed to eliminate a predicate instance from the goal, there is no point to apply flat rules to that goal. It is easy to show that the flat rules can neither eliminate predicates from the goal, nor enable previously disabled unfolding rules: the only unfolding rule that matches on flat heaplets is CALL, but those heaplets are taken care of separately by ABDUCECALL.

Following this observation, without loss of completeness, we can split the synthesis process into two phases: the unfolding phase, where flat rules are disabled, and the flat phase, which only starts when the goal contains no more predicate instances, and hence unfolding rules are inapplicable. This optimization is implemented in line 21 of Algorithm 4.1. As a result, some unsolvable goals will be identified early, in the unfolding phase. For example, the following goal:

\[
\{y, a, b\}; \{y \mapsto e \; a \mapsto 0;\} \leadsto \{y \mapsto u \; u \mapsto 0 \; \text{lseg}\{u, 0, S\}\}
\]

will fail immediately without exploring fruitless transformations on its flat heap, since no unfolding rules are applicable (assuming MaxUnfold = 1).

\[\text{The actual implementation is more efficient and breaks out of the loop as soon as one of the subgoals fails.}\]
4.1 and the way it interacts with reject a new rule application of Fig. 12. Fig. 13. Algorithm

5.5 Auxiliary Functions

The version of INDUCTION rule from Fig. 12 and the way it interacts with CALL are harmful for the framework’s completeness, due to the erasure of the tags from the postcondition of the function $f$.
Structuring the Synthesis of Heap-Manipulating Programs

Fig. 14. Pure synthesis rules.

to be applied recursively ($Q_f = [Q]$), so the subsequent applications of $f$ on the symbolic heap, overlapping with the “sealed” part $[Q]$ are prohibited. While it is not unsafe to employ the predicates from the procedure call’s postcondition in further procedure calls, one cannot ensure that they denote “smaller heaplets” and thus that the program terminates.\footnote{A similar issue is reported in the work by Rowe and Brotherston (2017) on verifying termination of procedural programs.}

To circumvent this limitation for common scenarios, we had to introduce one divergence between SuSLik and SSL as shown in Fig. 12. Specifically, in the tool, we implemented support for stratified\footnote{A similar issue is reported in the work by Rowe and Brotherston (2017) on verifying termination of procedural programs.} chained auxiliary function calls, \textit{i.e.}, calls on a heap resulted from another call (see flatten w/append and insertion sort benchmarks in Sec. 6). To allow for them, we had to implement a slightly different version of \textsc{induction} rule, which instead of erasing tags of the corresponding part of the postcondition ($Q_f \triangleq [Q]$) would increment them: $Q_f \triangleq [Q_f]^{++1}$. This would prevent a call of the same function on the resulting heap fragment, but would enable calls of auxiliary functions, whose specs’ preconditions feature predicates with matching higher-level tags. Specifically, each auxiliary function is assumed to have all predicate instances tagged with some $\ell > 0$ in its precondition, and with $\ell + 1$ in its postcondition. The discipline for assigning specific tag levels to auxiliary functions’ specifications is fixed to follow the order in which their signatures are introduced in the goal. Eventually, due to incrementation of level tags with each call (to a recursive self or to an auxiliary function), no applicable functions would be left in the context $\Sigma$. While this extension is unlikely to break the SSL soundness and termination results (due to the limit of chained applications, enabled by growing tags), it would require us to generalize the well-foundedness argument in Sec. 3.2, and in the interest of time we did not carry out this exercise.

5.6 Pure Synthesis and Branch Abduction

The core SSL relies on a pure synthesis oracle to suggest suitable applications of the \textsc{pick} rule. In our implementation, the declarative \textsc{pick} rule is replaced with the four rules depicted in Fig. 14, which together play the role of an incomplete yet practically effective pure synthesis oracle. The first three rules have been explained in Sec. 2.4.2; it is easy to see that each of them is a specialization of \textsc{pick}, and therefore is sound. The fourth rule, \textsc{branch}, comes to the rescue whenever the pure postcondition cannot be satisfied without introducing a case split, and generates a conditional statement with some guard $e$. Although \textsc{branch} is sound for any choice of guard, to curb the non-determinism, we pick $e$ from the set of learned guards $\mathcal{L}$.

The set $\mathcal{L}$ is populated through a mechanism we call \textit{branch abduction}, reminiscent of similar mechanisms in other synthesis techniques (Alur et al. 2017; Kneuss et al. 2013; Leino and Milicevic 2012; Polikarpova et al. 2016). The mechanism piggy-backs on the failure rule \textsc{postinvalid} (Sec. 5.4), which detects a goal whose pure postcondition $\psi$ does not follow from the precondition $\phi$. While the goal is still rejected as unsolvable, branch abduction also searches a small set of expressions (all
atomic boolean expressions over program variables in \( \Gamma \) trying to find an expression \( e \) such that \( \phi \land e \Rightarrow \psi \); if such an \( e \) exists, it is added to \( \mathcal{L} \).

For a simple example, consider the following synthesis goal that corresponds to computing a lower bound of two integers \( x \) and \( y \):

\[
\{ x, y \}; \{ r \mapsto 0 \} \leadsto \{ m \leq x \land m \leq y; r \mapsto m \}
\]

We first apply \texttt{PickVar} with \([m \mapsto x]\), arriving at the goal \( \{ r \mapsto 0 \} \leadsto \{ x \leq x \land x \leq y; r \mapsto x \} \). At this point \texttt{PostInvalid} fires, and branch abduction infers a guard \( x \leq y \) and adds it to \( \mathcal{L} \). Upon backtracking from the unsolvable goal, \texttt{Branch} fires since \( \mathcal{L} \) is non-empty. Both of its subgoals can be solved by suitable applications of \texttt{PickVar}.

6 IMPLEMENTATION AND EVALUATION

We implemented SSL-based synthesis as a tool, called \texttt{SuSLik}, in Scala, using Z3 (de Moura and Björner 2008) as the back-end SMT solver via the \texttt{SCALASMT} library (Cassez and Sloane 2017).

We evaluated our implementation with the goal of answering the following research questions:

1. **Generality**: Is \texttt{SuSLik} general enough to synthesize a range of nontrivial programs with pointers?
2. **Utility**: How does the size of the inputs required by \texttt{SuSLik} compare to the size of the generated programs? Does \texttt{SuSLik} require any additional hints apart from pre- and postconditions? What is the quality of the generated programs?
3. **Efficiency**: Is it efficient? What is the effect of optimizations from Sec. 5 on synthesis times?
4. **Comparison with existing tools**: How does \texttt{SuSLik} fare in comparison with existing tools for synthesizing heap-manipulating programs, specifically, \texttt{ImpSynt} (Qiu and Solar-Lezama 2017)?

6.1 Benchmarks

In order to answer these questions, we assembled a suite of 22 programs listed in Tab. 1. The benchmarks are grouped by the main data structure they manipulate: integer pointers, singly linked lists, sorted singly linked lists, binary trees, and binary search trees.

To facilitate comparison with existing work, most of the programs are taken from the literature on synthesis and verification of heap-manipulating programs: the \texttt{ImpSynt} synthesis benchmarks (Qiu and Solar-Lezama 2017), the \texttt{Jennisys} synthesis benchmarks (Leino and Milicevic 2012), and the \texttt{Dryad} verification benchmarks (Qiu et al. 2013). We manually translated these benchmarks into the input language of \texttt{SuSLik}, taking care to preserve their semantics. \texttt{Dryad} and \texttt{ImpSynt} use the \texttt{Dryad} dialect of separation logic as their specification language, hence the translation in this case was relatively straightforward. As an example, consider an \texttt{ImpSynt} specification and its \texttt{SuSLik} equivalent in Fig. 15. In addition to a pre-/postcondition, the former also contains a program \texttt{sketch}, where the constructs \texttt{?}, \texttt{cond}, and \texttt{statement} denote, respectively, an unknown expression, conditional guard, and statement, to be filled by the synthesizer. The main difference between the two pre-/postcondition pairs is that the \texttt{Dryad} logic supports recursive functions such as \texttt{len}, \texttt{min}, and \texttt{max}; in \texttt{SuSLik} this information is encoded in more traditional SL style: by passing additional ghost parameters to the inductive predicate \texttt{srt1}. The extra precondition \( 0 \leq n \land 0 \leq k \land k \leq 7 \) in \texttt{SuSLik} corresponds to implicit axioms in \texttt{ImpSynt} (in particular, the condition on \( k \) is due to its encoding of list elements as unsigned 3-bit integers—there is no such restriction in \texttt{SuSLik}). In addition to benchmarks from the literature, we also added several new programs that show-case interesting features of \texttt{SuSLik}.

\[\text{[15]The tool sources can be found at https://github.com/TyGuS/suslik.}\]
Table 1. Benchmarks and SuSLik results. For each benchmark, we report the size of the synthesized Code (in AST nodes) and the ratio Code/Spec of code to specification; as well as synthesis times (in seconds): with all optimizations enabled (Time), without phase distinction (T-phase), without invertible rules (T-inv), without early failure rules (T-fail), without the commutativity optimization (T-com), and without any optimizations (T-all). T-IS reports the ratio of synthesis time in ImpSYNT to Time. ‘-’ denotes timeout of 120 seconds.

<table>
<thead>
<tr>
<th>Group</th>
<th>Description</th>
<th>Code</th>
<th>Code/Spec</th>
<th>Time</th>
<th>T-phase</th>
<th>T-inv</th>
<th>T-fail</th>
<th>T-com</th>
<th>T-all</th>
<th>T-IS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integers</td>
<td>swap two</td>
<td>12</td>
<td>0.9x</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>1.4x</td>
</tr>
<tr>
<td></td>
<td>min of two(^2)</td>
<td>10</td>
<td>0.7x</td>
<td>0.1</td>
<td>0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>0.1</td>
<td>0.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>length(^1,2)</td>
<td>21</td>
<td>1.2x</td>
<td>0.4</td>
<td>0.9</td>
<td>0.5</td>
<td>0.4</td>
<td>0.6</td>
<td>1.4</td>
<td>29x</td>
</tr>
<tr>
<td></td>
<td>max(^1)</td>
<td>27</td>
<td>1.7x</td>
<td>0.6</td>
<td>0.8</td>
<td>0.5</td>
<td>0.4</td>
<td>0.4</td>
<td>0.8</td>
<td>20x</td>
</tr>
<tr>
<td></td>
<td>min(^1)</td>
<td>27</td>
<td>1.7x</td>
<td>0.5</td>
<td>0.9</td>
<td>0.5</td>
<td>0.4</td>
<td>0.5</td>
<td>1.2</td>
<td>49x</td>
</tr>
<tr>
<td></td>
<td>singleton(^2)</td>
<td>11</td>
<td>0.8x</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>dispose</td>
<td>11</td>
<td>2.8x</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>initialize</td>
<td>13</td>
<td>1.4x</td>
<td>&lt; 0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>&lt; 0.1</td>
<td>0.1</td>
<td>&lt; 0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>copy(^3)</td>
<td>35</td>
<td>2.5x</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
<td>0.1</td>
<td>0.2</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>append(^3)</td>
<td>19</td>
<td>1.1x</td>
<td>0.2</td>
<td>0.3</td>
<td>0.3</td>
<td>0.2</td>
<td>0.3</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>delete(^3)</td>
<td>44</td>
<td>2.6x</td>
<td>0.7</td>
<td>0.5</td>
<td>0.3</td>
<td>0.2</td>
<td>0.3</td>
<td>0.7</td>
<td></td>
</tr>
<tr>
<td>Linked List</td>
<td>prepend(^1)</td>
<td>11</td>
<td>0.3x</td>
<td>0.2</td>
<td>1.4</td>
<td>83.5</td>
<td>0.1</td>
<td>0.1</td>
<td>-</td>
<td>48x</td>
</tr>
<tr>
<td></td>
<td>insert(^3)</td>
<td>58</td>
<td>1.2x</td>
<td>4.8</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5.0</td>
<td>-</td>
<td>6x</td>
</tr>
<tr>
<td></td>
<td>insertion sort(^1)</td>
<td>28</td>
<td>1.3x</td>
<td>1.1</td>
<td>1.8</td>
<td>1.3</td>
<td>1.2</td>
<td>1.2</td>
<td>74.2</td>
<td>82x</td>
</tr>
<tr>
<td>Sorted list</td>
<td>size</td>
<td>38</td>
<td>2.7x</td>
<td>0.2</td>
<td>0.3</td>
<td>0.2</td>
<td>0.2</td>
<td>0.2</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>dispose</td>
<td>16</td>
<td>4.0x</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td>&lt; 0.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>copy</td>
<td>55</td>
<td>3.9x</td>
<td>0.4</td>
<td>49.8</td>
<td>-</td>
<td>0.8</td>
<td>1.4</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>flatten w/append</td>
<td>48</td>
<td>4.0x</td>
<td>0.4</td>
<td>0.6</td>
<td>0.5</td>
<td>0.4</td>
<td>0.6</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td>flatten w/acc</td>
<td>35</td>
<td>1.9x</td>
<td>0.6</td>
<td>1.7</td>
<td>0.7</td>
<td>0.5</td>
<td>0.6</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>BST</td>
<td>insert(^2)</td>
<td>58</td>
<td>1.2x</td>
<td>31.9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>11x</td>
</tr>
<tr>
<td></td>
<td>rotate left(^1)</td>
<td>15</td>
<td>0.1x</td>
<td>37.7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.5x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>rotate right(^3)</td>
<td>15</td>
<td>0.1x</td>
<td>17.2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.8x</td>
<td></td>
</tr>
</tbody>
</table>

\(^1\) From (Qiu and Solar-Lezama 2017)  \(^2\) From (Leino and Milicevic 2012)  \(^3\) From (Qiu et al. 2013)

6.2 Results

Evaluation results are summarized in Table 1. All experiments were conducted on a commodity laptop (2.7 GHz Intel Core i7 Lenovo Thinkpad with 16GB RAM).

6.2.1 Generality and Utility. Our experiment confirms that SuSLik is capable of synthesizing programs that manipulate a range of heap data structures, including nontrivial manipulations that require reasoning about both the shape and the content of the data structure, such as insertion into a binary search tree. We manually inspected all generated solutions, as well as their accompanying SSL derivations, and confirmed that they are indeed correct.\(^\text{16}\) Perhaps unsurprisingly, some of the solutions were not entirely intuitive: as one example, the synthesized version of list copy, in a bizarre yet valid move, swaps the tails of the original list and the copy at each recursive call!

Two of the programs in Table 1 make use of auxiliary functions: “insertion sort” calls “insert”, and “tree flatten w/append” calls the “append” function on linked lists. The specifications of auxiliary functions have to be supplied by the user (while their implementations can, of course, be synthesized independently). Alternatively, tree flattening can be synthesized without using an auxiliary function, if the user supplies an additional list argument that plays the role of an accumulator (see “tree flatten w/acc”). As such, SuSLik shares a common limitation of existing synthesizers for recursive functions: they require the initial synthesis goal to be inductive, and do not try to discover recursive auxiliary functions (we discuss this in more detail in Sec. 7).

\(^\text{16}\)In the future, we plan to output SSL derivations as SL proofs, checkable by a third-party system such as VST (Appel 2011).
\begin{verbatim}
lloc srtl_insert(lloc x, lloc k)
requires srtl(x)
ensures srtl(ret) ∧
  len(ret) = old(len(x)) + 1 ∧
  min(ret) = (old(k) < old(min(x)) ? old(k) : old(min(x))) ∧
  max(ret) = (old(max(x)) < old(k) ? old(k) : old(max^*(x)))
  if (cond(1)) {
    loc ?? := new;
    return ??;
  } else {
    statement(1);
    loc ?? := srtl_insert(??, ??);
    statement(1);
    return ??;
  }
}
\end{verbatim}

Fig. 15. (left) The ImpSynt input for the sorted list insertion; (right) The SuSLik input for the same benchmark.

For simple programs specification sizes are mostly comparable with the size of the synthesized code, whereas more complex benchmarks is where declarative specifications really shine: for example, for all Tree programs, the specification is at most half the size of the generated code. Three notable outliers are \texttt{prepend}, \texttt{rotate left}, and \texttt{rotate right}, whose implementations are relatively short, while the specification we inherited from ImpSynt describes the effects of the functions on the minimum and maximum of the list/tree. Note that the specification sizes we report exclude the definitions of inductive predicates, which are reusable, and are shared between the benchmarks.

6.2.2 Efficiency. SuSLik has proven to be efficient in synthesizing a variety of programs: all 22 benchmarks are synthesized within 40 seconds, and all but four of them take less than a second.

In order to assess the impact on performance of various optimizations described in Sec. 5, Tab. 1 also reports synthesis times with each optimization disabled: the column \textit{T-phase} corresponds to eliminating the distinction between phases; \textit{T-inv} corresponds to ignoring rule invertibility; \textit{T-fail} corresponds to dropping all failure rules; \textit{T-com} corresponds to disabling the symmetry reduction; finally, \textit{T-all} corresponds to a variant of SuSLik with all the above optimizations disabled. The results demonstrate the importance of optimizations for nontrivial programs: 8 out of 22 benchmarks time out when all optimizations are disabled. The simpler benchmarks (e.g., \texttt{swap}) do not benefit from the optimizations at all, since they do not exhibit a lot of backtracking. At the same time, all three BST benchmarks time-out as a result of disabling even a single optimization.

The copy benchmark is peculiar in that disabling any single optimization makes little difference, while disabling all of them together leads to a timeout. A closer examination of all possible configurations reveals that the speedup is due to interplay between phase distinction and invertible rules; this is not entirely surprising, since both optimizations prevent their fair share of particularly disastrous (wrt. performance) rule applications, such as early incorrect unification on flat heaps.

6.2.3 Comparison with Existing Synthesis Tools. We compare SuSLik with the most closely related prior work on ImpSynt \cite{Qiu:2017}. Out of the 14 benchmarks from \cite{Qiu:2017} successfully synthesized by ImpSynt, we excluded 5 that are not structurally recursive;\footnote{3 out of 5 are iterative versions of their recursive benchmarks, so we synthesize an equivalent recursive program for each of those; the other 2 use an internal loop or non-structural recursion, and are currently beyond the scope of SuSLik (Sec. 7).} the remaining 9 were successfully synthesized by SuSLik. The qualitative difference in terms of the required user input is immediately obvious from the representative example in Fig. 15:
in addition to the declarative specification, ImpSynt requires the user to provide an implementation
*sketch*, which fixes the control structure of the program, the positions of function calls, and the
number of other statements. These additional structural constraints are vital for reducing the size
of the search space in ImpSynt. Instead, SuSLik prunes the search space by leveraging the structure
inherent in separation logic proofs, allowing for more concise, purely declarative specifications.

Despite the additional hints from the user, ImpSynt is also less efficient: as shown in the column
*T-IS* of Tab. 1, on 6 out of 9 common benchmarks, ImpSynt takes at least an order of magnitude
longer than SuSLik, even though it has been evaluated on a 10-core server with 96GB of RAM.

7 LIMITATIONS AND DISCUSSION

The synthesis problem tackled in this work is clearly undecidable, since it subsumes several well-
known undecidable problems: termination of recursive programs, SL entailment in the presence of
general inductive predicates ([Antonopoulos et al. 2014](#)), and pure synthesis from first-order logic
specifications ([Reynolds et al. 2015](#)). Hence, no complete and terminating synthesis algorithm exists,
and in SuSLik, we made a design decision to sacrifice completeness for the sake of termination.

Sources of Incompleteness. SSL and SuSLik currently cannot synthesize programs that are not
structurally recursive wrt. some inductive predicate, such as, for instance, merging sorted lists or
merge sort. To relax this restriction we could adopt a more flexible termination argument for the
synthesized programs, such as showing that each recursive call decreases the value of a custom
termination metric; this technique is used in several automated verifiers and synthesizers ([Leino 2013](#); [Polikarpova et al. 2016](#)). A termination metric maps the tuple of function’s arguments into an
element of some set with a pre-defined well-founded order (usually a tuple of natural numbers),
and can be either provided by the user or inferred by the synthesizer. Custom termination metrics
would also enable support for mutually-recursive inductive predicates.

Another source of incompleteness is the limit MaxUnfold on the number of predicate unfoldings
via Open and Close rules. This addresses the undecidability of reasoning with inductive predicates,
but also precludes synthesis of some useful programs, e.g., allocating a large constant-sized list.

Finally, SuSLik’s pure synthesis oracle is obviously incomplete. This source of incompleteness
can be mitigated by delegating to an off-the-shelf pure synthesizer ([Kneuss et al. 2013](#); [Polikarpova et al. 2016](#); [Reynolds et al. 2015](#)), or eliminated entirely—for a restricted fragment of pure logic—by
leveraging complete synthesis procedures ([Jacobs et al. 2013](#); [Kuncak et al. 2010](#)).

Termination. Algorithm 4.1 terminates, as long as the pure synthesis oracle suggest finitely many
alternatives for a given goal. This can be established by considering the following tuples, ordered
lexicographically, as a termination measure for a given goal \( \langle \# 0\text{- or 1-tagged predicate instances;}
\# heaplets in pre- and postcondition, for which there is no matching one in the post-/precondition;
\# existentials; \# “flat” heaplets; \# conjuncts in the precondition; \# of points-to heaplets, whose
 disjointness or non-null-ness is not captured in the precondition \rangle). Notice that each rule from
Fig. 12, except for Open and Close reduces this value for the emitted sub-goals. Applicability of
those two rules is handled via MaxUnfold parameter.

Logic and Language Limitations. Although our programming component has no support for
*while*-loops, this is not a fundamental limitation: any program with a top-level loop can be rewritten
as a recursive program, and any internal loop can be rewritten as a *recursive auxiliary function*. Hence, the fundamental challenge is the synthesis of auxiliary functions, which SSL currently does
not support; instead, the initial goal is considered as inductive and handled via *Induction* rule.
Discovering specifications of auxiliary functions (or, equivalently, invariants of internal loops)
significantly increases the search space. To our knowledge, all existing synthesis techniques that
support this feature are quite limited, and rely on built-in or user-provided templates (Eguchi et al. 2018; Qiu and Solar-Lezama 2017; Si et al. 2018; Srivastava et al. 2010).

Some of SSL limitations are inherent for Separation Logics in general: SLs are known to work well with disjoint tree-like linked structures, and programs whose recursion scheme matches the data definition, but not so well with ramified data structures, e.g., graphs. To address those, one could integrate a more powerful, ramified version of FRAME rule (Hobor and Villard 2013) into SSL, but this would likely require more hints from the user, thus reducing the utility of the approach.

8 RELATED WORK

There are two main directions in the area of program synthesis: synthesis from informal descriptions (such as examples, natural language, or hints) (Albarghouthi et al. 2013; Feng et al. 2017; Feser et al. 2015; Murali et al. 2018; Osera and Zdancewic 2015; Polozov and Gulwani 2015; Smith and Albarghouthi 2016; Yaghmazadeh et al. 2017) and synthesis from formal specifications. We will only discuss the more relevant latter direction. The goal of this type of program synthesis is to obtain a provably correct program.

In this area, there is a well-known trade-off between three dimensions: how complex the synthesized programs are, how strong the correctness guarantees are, and how much input is required form the user. On one end of the spectrum there are interactive synthesizers (Delaware et al. 2015; Itzhaky et al. 2016), which can be very expressive and provide strong guarantees, but the user is expected to guide the synthesis process (although, usually, with aid of dedicated proof tactics). On the other end, there is fully automated synthesis for loop- and recursion-free programs over simple domains, like arithmetic and bit-vectors (Alur et al. 2017; Gulwani et al. 2011). Our work lies in the middle of this spectrum, where synthesis is automated but programs are more expressive.

In the presence of loops or recursion, verifying candidates becomes nontrivial. Synthesizers like SKETCH (Solar-Lezama 2013) and ROSETTE (Torlak and Bodík 2014) circumvent this problem by resorting to bounded verification, which only provides restricted guarantees and has scalability issues due to path explosion. In contrast, our work relies on unbounded deductive verification.

Among synthesis approaches that use unbounded verification, synthesizers like LEO (Kneuss et al. 2013) and SYNQUID (Polikarpova et al. 2016) focus on pure functional (recursive) programs, which are an easier target for unbounded verification. Proof-theoretic synthesis (Srivastava et al. 2010) is capable of synthesizing imperative programs with loops and arrays, but no linked structures; they pioneered the idea of synthesizing provably-correct programs by performing symbolic (SMT-based) search over programs and their verification conditions simultaneously.

Finally, the two pieces of prior work that are most closely related to ours in terms of scope are JENNISYS (Leino and Milicevic 2012) and Natural Synthesis (Qiu and Solar-Lezama 2017), both of which generate provably-correct heap-manipulating programs. Both of them are essentially instances of proof-theoretic synthesis with a program logic for reasoning about the heap. To that end, JENNISYS uses the DAFNY verifier (Leino 2013), which supports expressive yet undecidable specifications, and often requires hints from the user, so in practice the tool doesn’t scale to complex examples (for example, none of their benchmarks performs mutation). Natural Synthesis uses DRYAD (Madhusudan et al. 2012; Qiu et al. 2013), an SL-style program logic for reasoning about heap-manipulating programs. The downside of this approach is that whole-program symbolic search doesn’t scale to larger programs; to mitigate this, they require the user to provide sketches with substantial restrictions on the structure of the program. Our approach does not require sketches (but on the other hand, we do not support loops).

The recent tool FootPatch by van Tonder and Le Goues (2018) is very close in its methods and goals to SuSLIK. FootPatch builds on INFER (Calcagno and Distefano 2011), an open-source SL-based static analyzer by Facebook, using it for automated program repair. It takes the intermediate
assertions, provided by Infer for programs with bugs, such as resource and memory leaks, and null dereferences, and constructs additive patches based on the observed discrepancy. In this, it acts similarly to our AbduceCall rule. FootPatch does not synthesize patches that would involve recursion or complex control flow.

Instead of whole-program symbolic search, like in proof-theoretic synthesis, our work follows the tradition of deductive synthesis, i.e., backtracking search in the space of program derivation composed of synthesis rules, which gradually transform a specification into a program. This tradition originates from the work by Manna and Waldinger (1980), and similar ideas have been used in more recent synthesis work (Delaware et al. 2015; Kneuss et al. 2013; Polikarpova et al. 2016). In particular, the overall structure of our synthesis algorithm (backtracking and-or search) is similar to LEON (Kneuss et al. 2013), but our rules focus on heap manipulation, whereas their rules focus on synthesizing pure terms (so in fact Leon can be used as a component by our algorithm). Recent work on Optitian (Miltner et al. 2018) is very different in scope—they synthesize bijective string lenses from regular expression specifications and examples—but has interesting similarities in the technique. Their pre- and postcondition are regexes, and their technique tries to “align” them by e.g., unfolding the Kleene star; this is similar to how SuSLik tries to align the spatial pre- and postcondition by unfolding predicates.

Deductive synthesis is closely related to proof search, and there has been recent resurgence in applying proof-theoretic techniques, like focusing, to program synthesis (Frankle et al. 2016; Scherer 2017). But none of them do it for a complex logic that can reason about stateful programs.

Despite the vast space of available tools for symbolic verification based on Separation Logic: SMALLFOOT (Berdine et al. 2006), HTT (Nanevski et al. 2010), BEDROCK (Chlipala 2011), SLAYER (Berdine et al. 2011), HIP/SLEEK (Chin et al. 2011), VERIFAST (Jacobs et al. 2011), SLAD (Bouajjani et al. 2012), GRASSHOPPER (Piskac et al. 2014b), VIPER (Müller et al. 2016), CYCLIST (Rowe and Brotherston 2017), to name just a few, to the best of our knowledge none of them have been employed for deriving programs from specifications. It is certainly our hope that this work will bring new synergies between the research done in verification, theorem proving, and program synthesis communities.

For instance, in our approach to establish termination of SSL-synthesized programs, we used techniques close in spirit to the methods for proving total correctness in type/SL-based frameworks. E.g., SSL’s tags might be seen as a variant of resource capacities used in HIP/SLEEK (Le et al. 2014). Our use of Definition 3.2 of sized validity is similar to the induction on the finiteness of the heap used by Le and Hobor (2018) in their work on a logic for fractional shares. The way we use tagged predicates for establishing validity in Lemma 3.4 is reminiscent to the purpose of the ▷-modality in type theories for state and recursion (Appel et al. 2007).

9 CONCLUSION

In their seminal paper, Manna and Waldinger (1980) set forth an agenda for deductive synthesis of functional programs: “theorem provers have been exhibiting a steady increase in their effectiveness, and program synthesis is one of the most natural application of those systems”.

In this work, we moved this endeavour to an uncharted territory of stateful computations. For this, we employed a proof system which, instead of a pure type theory (Martin-Löf 1984), is based on Separation Logic—a Type Theory of State (Nanevski 2016). Taking this vision as a guiding principle, we designed Synthetic Separation Logic—a modest extension of Separation Logic, tailored for program synthesis, and implemented a proof search algorithm for it. In doing so, we took full advantage of the power of local reasoning about state (O’Hearn et al. 2001), which resulted in a principled and fast approach for synthesizing provably correct heap-manipulating programs.
ACKNOWLEDGMENTS

We wish to thank Aquinas Hobor and Reuben Rowe for their insightful and extremely detailed comments on presentation, formalism, and examples. Their technical feedback has helped immensely to bring out the best of this paper. We also wish to thank Shachar Itzhaky and Ranjit Jhala for their comments on the draft. We are very grateful to Franck Cassez for his help with configuring and using ScalaSMT. We thank the POPL’19 PC and AEC reviewers for the careful reading and many constructive suggestions on the logic, algorithm, and the implementation.

Sergey’s research was supported by a generous gift from Google and by the grant by the UK Research Institute in Verified Trustworthy Software Systems (VeTSS).

REFERENCES


Structuring the Synthesis of Heap-Manipulating Programs


Georges Gonthier, Beta Ziliani, Aleksandar Nanevski, and Derek Dreyer. 2011. How to make ad hoc proof automation less ad hoc. In ICFP. ACM.


