Modular Reasoning about Heap Paths
via Effectively Propositional Formulas

Shachar Itzhaky
Tel Aviv University
shachar@tau.ac.il

Anindya Banerjee
IMDEA Software Institute, Madrid,
Spain
anindya.banerjee@imdea.org

Neil Immerman
University of Massachusetts, Amherst,
USA
immerman@cs.umass.edu

Ori Lahav
Tel Aviv University
orilahav@post.tau.ac.il

Aleksandar Nanevski
IMDEA Software Institute, Madrid,
Spain
aleks.nanevski@imdea.org

Mooly Sagiv
Tel Aviv University
msagiv@acm.org

Abstract
First order logic with transitive closure, and separation logic enable elegant interactive verification of heap-manipulating programs. However, undecidability results and high asymptotic complexity of checking validity preclude complete automatic verification of such programs, even when loop invariants and procedure contracts are specified as formulas in these logics. This paper tackles the problem of procedure-modular verification of reachability properties of heap-manipulating programs using efficient decision procedures that are complete: that is, a SAT solver must generate a counterexample whenever a program does not satisfy its specification. By (a) requiring each procedure modifies a fixed set of heap partitions and creates a bounded amount of heap sharing, and (b) restricting program contracts and loop invariants to use only deterministic paths in the heap, we show that heap reachability updates can be described in a simple manner. The restrictions force program specifications and verification conditions to lie within a fragment of first-order logic with transitive closure that is reducible to effectively propositional logic, and hence facilitate sound, complete and efficient verification. We implemented a tool atop Z3 and report on preliminary experiments that establish the correctness of several programs that manipulate linked data structures.

1. Introduction
This paper shows how to harness existing SAT solvers for proving that a potentially recursive procedure satisfies its specification and for automatically producing counterexamples when it does not. We concentrate on proving safety properties of imperative programs manipulating linked data structures which is challenging since we need to reason about unbounded memory and destructive pointer updates. The tricky part is to identify a logic which is expressive enough to enable the modular verification of interesting procedures and properties and weak enough to enable sound and complete verification using SAT solvers.

Recently it was shown \cite{11} how to employ effectively propositional logic (or BSR logic \cite{1}) for verifying programs manipulating linked lists. It decides the validity of formulas of the form $\forall^* \exists^* q$ using SAT solvers where $q$ is a quantifier free relational formula (or equivalently decides the satisfiability of $\exists^* \forall^* q$ formulas). It has been successfully used in many other contexts \cite{18}.

In this paper we show that effectively propositional logic does not suffice to naturally express the effect on the global heap when the local heap of a procedure is accessible via shared nodes from outside. For example, Fig. 1 shows a pre- and post-heap before and after a call and yet any $AE_{AR}$ formula can be simulated by a BSR formula. This adaptation rule is expressed in a generalized version of BSR called $AE_{AR}$. $AE_{AR}$ allows an extra entry function symbol which maps each node $u$ in the global heap into the first node accessible from $u$ in the local heap. In Fig. 1, $z_1$, $z_2$, and $z_3$ are mapped to 2, 3, and 4, respectively. The key facts are that $AE_{AR}$ suffices to precisely define the global reachability relationship after each procedure call and yet any $AE_{AR}$ formula can be simulated by a BSR formula. Thus the automatic methods of \cite{11} still apply to this significantly more general setting.

This paper shows that in many cases, including the above example, reachability can be checked precisely using SAT solvers. Our solution is based on the following principles:

\begin{itemize}
  \item We follow the standard techniques (e.g., see \cite{2, 14, 25, 29}) by requiring that the programmer defines the set of potentially modified elements.
  \item The programmer only specifies postconditions on local heaps and ignores the effect of paths from the global heap.
  \item We provide a general and exact adaptation rule for adapting postconditions to the global heap. This adaptation rule is expressed in a generalized version of BSR called $AE_{AR}$. $AE_{AR}$ allows an extra entry function symbol which maps each node $u$ in the global heap into the first node accessible from $u$ in the local heap. In Fig. 1, $z_1$, $z_2$, and $z_3$ are mapped to 2, 3, and 4, respectively. The key facts are that $AE_{AR}$ suffices to precisely define the global reachability relationship after each procedure call and yet any $AE_{AR}$ formula can be simulated by a BSR formula. Thus the automatic methods of \cite{11} still apply to this significantly more general setting.
  \item We restrict the verified procedures in order to guarantee that the generated verification condition of every procedure remains in
\end{itemize}
AE^{AR}. The main restrictions are: type correctness, deterministic paths in the heap, limited number of changed list segments in the local heap (each of which may be unbounded) and limited amount of newly created heap sharing by each procedure call. These restrictions are enforced by the generated verification condition in AE^{AR}. This formula is automatically checked by the SAT solver.

1.1 Main Results

The results in this paper can be summarized as follows:

- We define a new logic, AE^{AR}, which extends BSR with a limited idempotent function and yet is equi-satisfiable with BSR.
- We provide a precise adaptation rule in AE^{AR}, which expresses the locality property of the change, and in conjunction with the postcondition on the local heap, precisely updates the reachability relation of the global heap.
- We generate a modular verification formula in AE^{AR} for each procedure, asserting that the procedure satisfies its pre- and post-conditions and the above restrictions. This verification condition is sound and complete, i.e., it is valid if and only if the procedure adheres to the restrictions and satisfies its requirements. We implemented this tool on top of Z3.
- We show that many programs can be modularly verified using our methods. They satisfy our restrictions and their BSR invariants can be naturally expressed.

1.2 A Running Example

To make the discussion more clear, we start with an example program. We use the union-find data structure\(^2\) which maintains a forest using a parent pointer at each node (see Fig. 2).\(^2\)

The method find requires that the argument \(x\) is not null. The formula tail\((x, r_s)\) asserts that the auxiliary variable \(r_s\) is equal to the root of \(x\). The procedure changes the pointers of some nodes in the closed interval \([x, r_s]\) to point directly to \(r_s\). Intervals are formally defined later (Definition 5). Intuitively, the closed interval \([a, b]\) denotes the set of nodes pointed to by \(a, a.f, a.f.f\) and so on up until \(b\) inclusive.

The return value of find (denoted by retval) is \(r_s\). The postcondition uses the symbol \(f\) that denotes the value of \(f\) before the method was invoked. Since find compresses paths from ancestors of \(x\) to single edges to \(r_s\), this root may be shared via new parent pointers. Fig. 3 depicts a typical run of find.

The method union requires that both its arguments are not null. It potentially modifies the ancestors of \(x\) and \(y\), i.e., \([x, r_s]\) \(\cup\) \([y, r_y]\). Fig. 4 depicts a typical run of union. Notice that we support an unbounded number of cutpoints\(^2\) (see Section 3).

\(^2\) We have simplified union by not keeping track of the sizes of sets in order to attach the smaller set to the larger.

---

**Figure 1.** Reversing a list pointed to by a head \(h\) with many shared nodes accessible from outside the local heap (surrounded by a rounded rectangle).

**Figure 2.** An annotated implementation of Union-Find in Java. \(f\) is the backbone field denoting the parent of a tree node.

**Figure 3.** An example scenario of running find (\(\square\)=return value).

**Figure 4.** An example scenario of running union.
1.3 Working Assumptions

Type correct The procedure manipulates references to dynamically created objects in a type-safe way. For example, we do not support pointer arithmetic.

Deterministic Reachability The specification may use arbitrary uninterpreted relations. It may also use the reachability formula \( \alpha(f^*)\beta \) meaning that \( \beta \) is reachable from \( \alpha \) via zero or more steps along the functional backbone field \( f \). It may not use \( f \) in any other way. Until Section 6 we require \( f \) to be acyclic and we restrict our attention to only one backbone field.

Precondition There is a requires clause defining the precondition which is written in alternation-free relational first-order logic \( (AF^R) \) and may use the relation \( f^* \) \[1\].

Mod-set There is a modifies clause defining the mod-set (mod\( _f \)), which is the set of potentially changed memory locations (We include both source and target of every edge that is added or deleted). The modified set may have an unbounded number of vertices, but we require it to be the union of a bounded number of \( f \)-intervals, that is chains of vertices through \( f \)-pointers.

Postcondition There is an ensures clause which exactly defines the new reachability relation \( f^* \) restricted to mod\( _f \). The ensures clause, written in \( AF^R \), may use two vocabularies (employing both \( f \) and \( f^* \) to refer to the reachability relations before and after.

Bounded new sharing All the new shared nodes — nodes pointed to by more than one node — must be pointed to by local variables at the end of the procedure’s execution. This requires that only a bounded number of new shared nodes can be introduced by each procedure call. Note that many heap-manipulating programs exhibit limited sharing as noted in the experimental measurements of Mitchell \[16\]. A similar restriction is also used in shape analysis techniques for device driver programs \[27\].

Loop-free We assume that all code is loop free, with loops replaced by recursive calls.

1.4 Outline of the rest of this paper

Section 2 provides a rule for adapting local changes to states containing a global heap. The idea is that the programmer only specifies changes in a small, local area of the heap. Section 3 introduces a new logic called \( AE^{AR} \). Section 4 formalizes the requirements for specifying the meaning of commands and procedures. The technique for generating verification conditions is presented in Section 5. Extensions to the frameworks are discussed in Section 6. Our preliminary verification experience appears in Section 7. Section 8 discusses related work and Section 9 concludes. Details of the logical proof are contained in \[10\] Appendix A.

2. Adaptation of Local Effect to the Global Heap

Our goal is to reason modularly about a procedure that modifies a subset of the heap. We wish to automatically update the reachability relation in the entire heap based on the changes to the modified subset. We remark that in this paper we are concerned with reachability between any two nodes in the heap, as opposed to only those pointed to by program variables. When we discuss sharing we mean sharing via pointer fields in the heap as opposed to aliasing from stack variables, which does not concern us in this paper.

2.1 Non-Local Effects

Reachability is inherently non-local: a single edge mutation can affect the reachability of an unbounded number of points that are an unbounded distance from the point of change. Fig. 5 contains a typical run of find. Two kinds of “frames” are depicted: (i) mod\( _f = \{x, y, c\} \) specified by the programmer, denotes the nodes whose edges can be directly changed by find — this is the standard notion of a frame condition; (ii) mod\( _f \) denotes nodes for which \( f^* \), the path relation, has changed. We do not and, in general we cannot, specify mod\( _f \) in a modular way because it usually depends on variables outside the scope of this function such as \( y \) in Fig. 5. In the example shown, there is a path from \( y \) to \( c \) before the call which does not exist after the call. Furthermore, mod\( _f \) can be an arbitrarily large set: in particular, it may not be expressible as the union of a bounded set of intervals: for example, when adding a subtree as a child of some node in another tree, mod\( _f \) spans only one edge, whereas mod\( _f \) is the entire subtree added — which may contain an unbounded number of branches.

The postcondition of find is sound (every execution of find satisfies it), but incomplete: it does not provide a way to determine information concerning paths outside mod, such as from \( y \) to \( c \) in Fig. 5. Therefore, this rule is often not enough in order to verify the correctness of programs that invoke find in larger contexts.

Notice the difficulty of updating the global heap, especially the part mod\( _f \). We express this below, first letting mod\( _f \) be an arbitrary set of nodes. This rule is often not enough in order to verify the correctness of programs that invoke find in larger contexts.

The problem with this is that it uses a logic that is too expressive and thus hard for automated reasoning: FO(TC) is not decidable (in fact, not even recursively enumerable). The first problem is that the \( \beta \) relation is not usually first order expressible and generally requires transitive...
The memory states (c) and (d) are obtained from the memory states (a) and (b), respectively, by deleting the edge $\langle 3, 2 \rangle$. The reachability in (c) is not the same as in (d), which shows it is impossible to update reachability in general w.r.t. edge deletion, without using the edge relation.

Figure 7. The function $en^{mod}$ maps every node $\sigma$ to the first node in mod reachable from $\sigma$. Notice that for any $\alpha \in mod$, $en^{mod}(\alpha) = \alpha$ by definition.

Figure 6. Memory states with non-unique pointers where global reasoning about reachability is hard. In the memory state (a), there is one edge from $\alpha$ into the modified-set $\{1, 2, 3, 4\}$, and in memory state (b), there are two edges from $\alpha$ into the same modified-set, $\{1, 2, 3, 4\}$. The two memory states have the same reachability relation and therefore are indistinguishable in terms of reachability. The memory states (c) and (d) are obtained from the memory states (a) and (b), respectively, by deleting the edge $(2, 3)$. The reachability in (c) is not the same as in (d), which shows it is impossible to update reachability in general w.r.t. edge deletion, without using the edge relation.

We now present an equivalent adaptation rule in a restricted logic, without transitive closure or extra quantifier-alternations. This is possible due to our assumptions from Section 1.3 and it greatly simplifies reasoning about modified paths in the entire heap. We require a new function symbol, $en^{mod}$. We call $en^{mod}(\sigma)$ the entry point of $\sigma$ in mod, i.e., the first node on the (unique) path from $\sigma$ that enters mod, and null if no such node exists (see Figure 7).

Note that since transitive closure is only applied to functions, entry points such as $\alpha$ in eq (1) are uniquely determined by $\sigma$, the origin of the path. A key property of $en^{mod}$ is that on mod itself, $en^{mod}$ is the identity, and therefore for any $\sigma \in V$ it holds that $en^{mod}(en^{mod}(\sigma)) = en^{mod}(\sigma)$ — that is, the function $en^{mod}$ is idempotent. It is important to note that $en^{mod}$ does not change as a result of local modifications in mod. Hence, we do not need to worry about $en^{mod}$ in the pre-state as opposed to the post-state. Formally, $en^{mod}$ is characterized by the following formula:

$$\forall \sigma : (en^{mod}(\sigma) = null \land \forall \alpha \in mod : \neg (f^*(\alpha) \land (\sigma(f^*) \in mod \land \exists \tau \in mod : \sigma(\tau) \land \neg \exists \alpha \in mod : \sigma(\tau) \land \sigma(f^*) \land \alpha(f^*) \in mod) \land (\sigma(\tau) \land \exists \alpha \in mod : \sigma(\tau) \land \sigma(f^*) \land \alpha(f^*) \in mod))$$

Using $en^{mod}$ the new adaptation rule adapt[mod] is obtained by considering, for every source and target, the following three cases:

Out-In: The source is out of mod; the target is in; In-Out: The source is in mod; the target is out; Out-Out: The source and target are both out of mod.

The full adaptation rule is obtained by taking the conjunction of the formulas for each case (eq (3), eq (4), eq (5)), that are described below, and the formula defining $en^{mod}$ (eq (2)).

Out-In Paths Using $en^{mod}$ we can easily handle paths that enter mod. Such paths originate at some $\sigma \notin mod$ and terminate at some $\tau \in mod$. Any such path therefore has to go through $en^{mod}(\sigma)$ as depicted in Fig. 8. Thus, the following simple formula can be used:

$$\forall \sigma \notin mod, \tau \in mod : \sigma(f^*) \Leftrightarrow en^{mod}(\sigma)(f^*) \tau$$

Observe that for any $\beta \in mod$, the atomic formula used above, $en^{mod}(\sigma)(f^*) \beta$, corresponds to the FO(TC) sub-formula $\exists \alpha \in mod : \sigma(\alpha) \land \alpha(\beta) \in mod$ from eq (1).

In-Out Paths We now shift attention to paths that exit mod. Exit points, that is, last points on some path that belong to mod, are more subtle since both ends of the path are needed to determine them. The end of the path is not enough since it can be shared, and the origin of the path is not enough since it can exit the set multiple times, because a path may exit mod and enter it again later. Therefore, we cannot define a function in a similar manner to $en^{mod}$. The fact that transitive closure is only applied to functions is useful here: every interval $[\alpha, \beta]$ has at most one exit $\beta$. We therefore utilize the fact that mod is expressed as a bounded union of intervals — which bounds the potential exit points to a bounded set of terms. We will denote the exit points of mod by $t_i$.

For example, in the procedure swap shown in Fig. 9 $mod = \{x, f(x)\}$ and there is one exit point $t_i = f(x)(f(x))$ is a constant set by the pre-condition to have the value of $f(f(x))$ using the inversion formula eq (6) to be introduced formally in Section 3.1.

Any path that originates in mod and terminates outside mod must leave through a last exit point $t_i$. Notice that the exit points also do not change as a result of modifying edges between nodes in mod. Let $p$ be a path from $\sigma$ to $\tau$ and let $t_i$ be the last exit point along $p$. Note that the part of the path from $t_i$ to $\tau$ consists only of unchanged edges — since they are all outside of mod. We can therefore safely use $f^*$, rather than $q^*$, to characterize

Figure 8. This diagram depicts how an arbitrary path from $\sigma \notin mod$ to $\tau \notin mod$ is constructed from three segments: $[\sigma, \alpha]$, $[\alpha, t_i]$, and $[t_i, \tau]$ (here $i = 2$). Arrows in the diagram denote paths; thick arrows entering and exiting the box denote paths that were not modified since they are outside of mod. Here, $\alpha = en^{mod}(\sigma)$ is an entry-point and $t_1, t_2$ are exit-points.
void swap_two(Node a, Node b) {
    swap(a); swap(b);
}

Figure 9. A simple function that swaps two adjacent elements following $x$ in a singly-linked list. Dotted lines denote the new state after the swap. The notation e.g. $E_{f}(x, f^1)$ denotes the single edge from $x$ to $f^1$ following the $f$ field.

Figure 10. A subtle situation occurs when the path from $\sigma$ passes through multiple exit-points. In such a case, the relevant exit-point for $\sigma(f^*)\tau_1$ is $t_1$, whereas for $\sigma(f^*)\tau_2$ and $\tau_1(f^*)\tau_2$ it would be $t_2$.

Figure 11. An example of a procedure where the mod-set is not (essentially) convex.

void swap_two(Node x) {
    Node t = x.f;
    t.f = x.f.f;
    x.f = t.f;
    return;
}

In conclusion, our adaptation rule, adapt[mod], is the conjunction of the three formulas in eq (3), eq (4), eq (5), and the formula defining $en^{\text{mod}}(\sigma)$ (eq (2)). We need some more formalism, introduced in the next section, before we show that adapt[mod] meets our needs.

3. Adaptable Heap Reachability Logic

In this section we introduce an extension of $AE^R$ from [11], called adaptable heap reachability logic, and denoted by $AE^{RA}$. This extension still has the attractive property of $AE^R$, as it is essentially reducible to the function-free $\forall^*\exists^*$-fragment of first-order logic, and thus its validity can be checked by a SAT-solver.

3.1 Preliminaries

This section reviews the $AF^R$ (alternation free) and $AE^R (\forall\exists)$ logics from [11]. They are decidable for validity since their negation corresponds to the BSR fragment [18]. These logics include the relation $f^*$ (the reflexive transitive closure of $f$) but forbid the explicit use of function symbols including $f$. Until Section 6 we will use at most one designated backbone function ($f$) per formula.

**Definition 1.** A vocabulary $V = (\mathcal{C}, \{f\}, \mathcal{R})$ is a triple of constant symbols, function symbol, relation symbols.

A term, $t$, is a variable or constant symbol.

An atomic formula is one of the following:

(i) $t_1 = t_2$;

(ii) $t_1, t_2, \ldots, t_n$ where $\tau$ is a relation symbol of arity $n$;

(iii) $t_1(f^*) t_2$.

A quantifier-free formula ($QF^R$) is a boolean combination of atomic formulas. A universal formula begins with zero or more universal quantifiers followed by a quantifier-free formula.

An alternation-free formula ($AF^R$) is a boolean combination of universal formulas. $AE^R$ consists of formulas with quantifier-prefix $\forall^*\exists^*$.

In particular, $QF^R \subseteq AF^R \subseteq AE^R$. The preconditions and the postconditions in Fig. 2 are all $AF^R$ formulas.

**Decidability and Inversion** Every $AE^R$ formula can be translated to a first-order $\forall^*\exists^*$ formula via the following steps [11].

(i) Add a new uninterpreted relation $R_f$ which is intended to represent ($f^*$), the reflexive transitive closure of reachability via $f$.

(ii) Add the consistency rule $\Gamma_{\text{linOrd}}$ shown in Table 4 which requires that $R_f$ is a total order, i.e., reflexive, transitive, acyclic, and linear, and (iii) Replace all occurrences of $t_1(f^*) t_2$ by $R_f(t_1, t_2)$.

**Proposition 1 (Simulation of $AE^R$).** [12] Proposition 3, Appendix A.1] Consider a $AE^R$ formula $\varphi$ over vocabulary $V = (\mathcal{C}, \{f\}, \mathcal{R})$. Let $\varphi' \equiv \varphi[R_f(t_1, t_2)/t_1(f^*) t_2]$. Then $\varphi'$ is a first-order formula over vocabulary $V' = (\mathcal{C}, \emptyset, \mathcal{R} \cup \{R_f\})$ and $\Gamma_{\text{linOrd}} \rightarrow \varphi'$ is valid if and only if the original formula $\varphi$ is valid.

When the graph of $f$ is acyclic, the relation $E_f$ characterizing the function $f$ can be recovered from its reflexive transitive closure, $f^*$, at the cost of an extra universal quantifier:

$$E_f(\alpha, \beta) \equiv \alpha(f^*) \beta \land \forall \gamma: \alpha(f^*) \gamma \rightarrow \beta(f^*) \gamma$$

(6)
Table 1. A universal formula $\Gamma_{\text{linOrd}}$ requiring that all points reachable from a given point are linearly ordered.

| $\forall \alpha : R_f(\alpha, \alpha)$ | reflexivity |
| $\forall \beta, \gamma : R_f(\alpha, \beta) \land R_f(\beta, \gamma) \rightarrow R_f(\alpha, \gamma)$ | transitivity |
| $\forall \alpha, \beta : R_f(\alpha, \beta) \land R_f(\beta, \alpha) \rightarrow \alpha = \beta$ | acyclicity |
| $\forall \alpha, \beta, \gamma : R_f(\alpha, \beta) \land R_f(\alpha, \gamma) \rightarrow (R_f(\beta, \gamma) \lor R_f(\gamma, \beta))$ | linearity |

Here $\alpha(f^+) \beta \equiv \alpha(f^+) \beta \land \alpha \neq \beta$.

3.2 Adaptable Heap Reachability Logic

The new logic $AE^{AR}$ is obtained by augmenting $AE^R$ with unary function symbols, denoted by $g, h_1, \ldots, h_n$, where:

- $g$ must be interpreted as an idempotent function.
- The images $h_1, \ldots, h_n$ are all bounded by some pre-determined parameter $N$, that is: each $h_i$ takes at most $N$ distinct values.
- Function symbols may not be nested, i.e., all terms involving function symbols have the form $f(z)$, where $z$ is a variable.

We later show that $AE^{AR}$ suffices for expressing the verification conditions of the programs discussed above. In the typical use case, the function $g$ assigns the entry point in the mod-set for every node (called $en^\text{mod}$) above, and the functions $h_1, \ldots, h_n$ are used for expressing the entry points in inner mod-sets. The main attractive feature of this logic is given in the following theorem.

Theorem 1. Any $AE^{AR}$-formula $\varphi$ can be translated to an equivalent (first-order) function-free $\forall^* \exists^*$-formula.

The proof of Theorem 1 given in [10] Appendix A1, begins by translating $\varphi$ to a $\forall^* \exists^*$-formula $\varphi'$ as described in Proposition 1 without modifying the function symbols $g, h_1, \ldots, h_n$. The function symbols are then replaced by new relation and constant symbols. We add new universal formulas to express the above semantic restrictions on the functions.

4. Modular Specifications of Procedure Behaviours

4.1 Notations

Definition 2. Let $V = (C, \{f\}, R)$ be a vocabulary including the constant symbol null. A state, $M$, is a logical structure with domain $|M|$, including null, and $null^M = null = f^{null}(null)$, where $s^M$ is the interpretation of the symbol $s$ in the structure $M$. A state is appropriate for an annotated procedure proc if its vocabulary includes every symbol occurring in its annotations, and constants corresponding to all of the program variables occurring in proc.

The diagrams in this paper denote states. For example, Fig. 4 shows a transition between two states.

Below we define the notion of two-vocabulary structures, which are useful to describe relations between pre- and post-states.

Definition 3 (Two-vocabulary Structure). For states $M$ and $M'$ over the same vocabulary $V = (C, \{f\}, R)$ and with the same domain $|M| = |M'|$, we denote by $M \oplus M'$ the structure over the two-vocabulary $V' = (C \cup C', \{null\}, \{f, f\}', R \cup R)$ obtained by combining $M, M'$ in the following way: $f$ is interpreted as $f^M$ and $f'$ is interpreted as $f'^{null}$, similarly for all the other symbols.

Definition 4 (Backbone Differences). For states $M$ and $M'$ over the same vocabulary $V = (C, \{f\}, R)$ and with the same domain $|M| = |M'|$, the set $M \ominus M'$ consists of the “differences between $M$ and $M'$ w.r.t. $f$, excluding null”, i.e. all elements $u$ of $|M|$ such that $f^M(u) \neq f'^{null}(u)$, as well as their non-null images in $f^M$ and $f'^{null}$.

4.2 Modification Set

We must specify the mod-set, mod, containing all the endpoints of edges that are modified. Therefore when adding or deleting an edge $(s, t)$, both ends — the source, $s$, and the target, $t$ — are included in mod. Often in programming language semantics, only the source is considered modified. However, thinking of the heap as a graph, it is useful to consider both ends of a modified edge as modified. For example, in the running example program find (Fig. 2), since new references to $r_s$ may be introduced as a result of path compression, the root $r_s$ is also considered as part of mod.

Our mod-sets are built from two kinds of intervals:

Definition 5 (Intervals). The closed interval $[a, b]$ is

$[a, b] \equiv \{ \gamma | a(f^+) \gamma \land \gamma(f^+) b \}$

and the half-open interval $[a, null]$ is:

$[a, null] \equiv \{ \gamma | a(f^+) \gamma \land \gamma(f^+) null \}$

(notice that $\gamma(f^+)$ null is always true in acyclic heaps).

Definition 6 (mod-set). The mod-set, mod, of a procedure is a union $I_1 \cup I_2 \cup \ldots \cup I_k$, where each $I_i$ may be $[s_i, t_i]$ or $[s_i, null]$, $s_i, t_i$ are parameters of the procedure or constant symbols occurring in the pre-condition.

In our examples, the mod-sets of find and union are written above each procedure, preceded by the symbol “@ mod” (Fig. 2). Note that it follows from Definition 6 that $a \in \text{mod}$, is expressible as a quantifier-free formula.

Definition 7. Given an appropriate state $M$ for proc with modset mod, mod$^M$ is the set of all elements in $|M|$ that are in one of the intervals defining mod (see Definition 6).

4.3 Pre- and Post-Conditions

The programmer specifies $AER$ pre- and post-conditions. Two-vocabulary formulas may be used in the post-conditions where $f$ denotes the value of $f$ before the call.

4.4 Specifying Atomic Commands

Table 2 provides specification of atomic commands. They describe the memory changed by atomic statements and the changes on the local heap.

Accessing a pointer field The statement ret = y.f reads the content of the $f$-field of $y$, into ret. It requires that $y$ is not null and that an auxiliary variable $a$ points to the $f$-field of $y$ (which may be null). It does not modify the heap at all. It sets ret to $a$. It is interesting to note that the postcondition is quantifier free and much simpler than the one provided in [11]. The reason is that we do not need to specify the effect on the whole heap.

Edge Removal The statement $y.f = null$ sets the content of the $f$-field of $y$, to null. It requires that $y$ is not null and that an auxiliary variable $a$ points to the $f$-field of $y$ (which may be
null). It modifies the node pointed-to by y and potentially the node pointed-to by s. Notice that the modset includes the elements pointed to by y and s, the two end-points of the edge. It removes paths between y and s. The postcondition asserts that there are no paths from y to s. Also, since s is potentially modified, it asserts that there are no paths from s to y.

**Edge Additions** The statement \( y.f = x \) is specified assuming without loss of generality, that the statement \( y.f = null \) was applied before it. Thus, it only handles edge additions. It therefore requires that y is not null and its f-field is null. It modifies the node pointed-to by y and potentially the node pointed-to by x. It creates a new path from y to x and asserts the absence of new paths from x back to y. Again the absence of back paths (denoted by \( ¬(f^∗)y \)) is needed for completeness. The reason is that both the node pointed-to by x and y are potentially modified. Since x is potentially modified, without this assertion, a post-state in which \( x.f == y \) will be allowed by the postcondition.

### 4.4.1 Soundness and Completeness

We now formalize the notion of soundness and completeness of modular specifications and assert that the specifications of atomic commands are sound and complete.

**Definition 8** (Soundness and Completeness of Procedure Specification). Consider a procedure \( \text{proc} \) with precondition \( P \), modset \( \text{mod} \), post-condition \( Q \). We say that \((P, \text{mod}, Q)\) is sound with respect to \( \text{proc} \) if for every appropriate pre-state \( M \) such that \( M \models P \) and appropriate post-state \( M \) which is a potential outcome of the body of \( \text{proc} \) when executed on \( M \): (i) \( M \models \text{mod} \), (ii) \( M/M \models Q \). Such a triple \((P, \text{mod}, Q)\) is complete with respect to \( \text{proc} \) if for every appropriate states \( M, M' \) such that (i) \( M \models P \), (ii) \( M'/M \models Q \), and (iii) \( M \models \text{mod} \), then there exists an execution of the body of \( \text{proc} \) on \( M \) whose outcome is \( M' \).

The following proposition establishes the correctness of atomic statements.

**Proposition 2** (Soundness and Completeness of Atomic Commands). The specifications of atomic commands given in Table 2 are sound and complete.

The following lemma establishes the correctness of \( \text{find} \) and \( \text{union} \), which is interesting since they update an unbounded amount of memory.

**Lemma 1** (Soundness and Completeness of Union-Find). The specification of \( \text{find} \) and \( \text{union} \) in Fig. 13 is sound and complete.

We can now state the following proposition:

**Proposition 3** (Soundness and Completeness of \( \text{adapt}[] \)). Let \( \text{mod} \) be a mod-set of some procedure \( \text{proc} \). Let \( M \) and \( M' \) be two appropriate states for \( \text{proc} \). Then, \( M \models \text{mod} \) if and only if \( M/M \) augmented with some interpretation for the function symbol \( \text{en}^\text{mod} \) is a model of \( \text{adapt}[\text{mod}] \).

### 5. Generating Verification Condition for Procedure With Sub-calls in \( AE^{AR} \)

We follow the standard procedures, e.g. [26], which generates a verification condition for a procedure annotated with specification using weakest (liberal) preconditions. Roughly speaking, for a Hoare triple \( \{P\} \text{proc} \{Q\} \), we generate the usual verification condition \( \text{vc}[\text{proc}] = P \rightarrow \text{wp}[\text{proc}][Q] \).

For the basic commands (assignment, composition, and conditional) we employ the standard definitions, given in Table 3.

#### 5.1 Modular Verification Conditions

The modular verification condition would also contain a conjunct for checking that mod affected by the invoked procedure is a subset of the “outer” mod. This way the specified restriction can be checked in \( AE^{AR} \) and the SMT solver can therefore be used to enforce it automatically.

#### 5.2 Weakest-precondition of Call Statements

As discussed in Section 2 the specification as it appears in the “ensures” clause of a procedure’s contract is a local one, and in order to make reasoning complete we need to adapt it in order to handle arbitrary contexts. This is done by conjoining \( Q_{\text{proc}} \) occurring in the “ensures” clause from the specification of \( \text{proc} \) with the universal adaptation rule adapt[mod], where mod is replaced with the mod-set as specified in the “modifies” clause of \( \text{proc} \).

Table 4 presents a formula for the weakest-precondition of a statement containing the single procedure call, where the invoked procedure has the specifications as in Fig. 13 where “\( \text{proc} \)” has the formal parameters \( \pi = x_{1}, \ldots, x_{k} \), and it is used with \( \pi = a_{1}, \ldots, a_{k} \) (used in the formula) as the actual arguments for a specific procedure call; we assume w.l.o.g. that each \( a_{i} \) is a local variable of the calling procedure.
\[
\text{wp}[r := \text{proc}(\pi)](Q) \leftarrow \text{P}_{\text{proc}}[\pi]\]
\[
\forall \alpha : \alpha \in \text{mod}_{\text{proc}}[\pi/\tau] \rightarrow \alpha \in \text{mod}_{\text{prog}} \wedge
\forall \zeta : Q_{\text{proc}}[\tau/\xi, \xi/f, f/f, \xi/\text{retval}] \wedge
\text{adapt}[\text{mod}_{\text{prog}}[f/f]][\pi/\tau, f/f, f/f] \rightarrow Q[f/f, \xi/\zeta]
\]

Table 4. Computing the weakest (liberal) precondition for a statement containing a procedure call. \( r \) is a local variable that is assigned the return value; \( \pi \) are the actual arguments passed. \( \hat{f} \) is a fresh function symbol.

In general it is not obvious how to enforce that the set of locations modified by inner calls is a subset of the set of locations declared by the outer procedure. Moreover, this can be tricky to check since it depends on aliasing and paths between nodes pointed to by different variables. Fortunately, the sub-formula \( \forall \alpha : \alpha \in \text{mod}_{\text{proc}}[\pi/\tau] \rightarrow \alpha \in \text{mod}_{\text{prog}} \) captures this property, ensuring that the outer procedure does not exceed its own mod specification, taking advantage of the interval-union structure of the mod. Since all the modifications (even atomic ones) are done by means of procedure calls, this ensures that no edges incident to nodes outside mod are changed.

**Proposition 4.** The rule for \( \text{wp}[] \) of call statements is sound and complete, that is, when \( \text{proc} \) is a procedure with specification as in Fig. 13 called in the context of \( \text{prog} \) whose mod-set is \( \text{mod} \):

\[
M \models \text{wp}[r := \text{proc}(\pi)](Q) \downarrow
\]
\[
M \models \text{P}_{\text{proc}}[\pi/\tau] \wedge \text{mod}_{\text{prog}}^M \subseteq \text{mod}_{\text{mod}}^M \wedge
\forall M : (M/M \models Q_{\text{proc}}[\tau/\xi, M \oplus M \subseteq \text{mod}_{\text{mod}}^M] \rightarrow M[r \rightarrow \text{retval}^M] \models Q
\]

5.3 Reducing Function Symbols

Notice that when we apply the adaptation rule for \( \text{AE}^{\text{AR}} \), as discussed above, it introduces a new function symbol \( \text{en}^{\text{mod}} \) depending on the concrete mod-set of the called procedure. This introduces a complication: the mod-sets of separate procedure calls may differ depending on the concrete mod-set of the called procedure. This introduces a new function symbol \( \text{en}^{\text{mod}} \) as well as the entire call, is bounded, and we replace all of the function symbols \( \text{en}^{\text{mod}} \) by \( \text{en}^{\text{mod}} \) defined for every such invocation. Consider for example the situation of \( \text{union} \) making two invocations to \( \text{find} \). In Fig. 14 one can see that the mod of \( \text{union} \) is \([x, r_x] \cup [y, r_y] \), while the mod of the first call \( t := \text{find}(x) \) is \([x, r_x] \), which may be a proper subset of the former. The mod of the second invocation is \([y, r_y] \), which may overlap with \([x, r_x] \).

To meet the requirement of \( \text{AE}^{\text{AR}} \) concerning the function symbols, we observe that: (a) the amount of sharing any particular function call creates, as well as the entire call, is bounded, and we can pre-determine a bound for it; (b) the modification set of any sub-call must be a subset of the top call, as otherwise it violates the obligation not to change any edge outside mod. These two properties allow us to express the functions \( \text{en}^{\text{mod}} \) of the sub-calls using \( \text{en}^{\text{mod}} \) of the top procedure and intermediate functions with bounded image. Thus, we replace all of the function symbols \( \text{en}^{\text{mod}} \) introduced by \( \text{adapt}[S] \) for different \( S \)'s with a single (global, idempotent) function symbol together with a set of bounded function symbols.

Consider a statement \( r := \text{proc}(\pi) \) in a procedure \( \text{prog} \). Let \( A \) denote the mod-set of \( \text{prog} \), and \( B \) — the mod-set of \( \text{proc} \). We show how \( \text{en}^{\text{B}} \) can be expressed using \( \text{en}^{\text{A}} \) and one more function, \( \text{en}^{\text{A}} \)

\[\text{en}^{\text{B}}(\sigma) \equiv \begin{cases} \text{en}^{\text{A}}(\sigma) & \text{en}^{\text{A}}(\sigma) \in B \\ \text{en}^{\text{B}}(\sigma) & \text{otherwise} \end{cases}\]

Using equality the nesting of function symbols can be reduced (without affecting the quantifier alternation).

Consult Fig. 15 notice that \( \text{en}^{\text{B/A}}(\sigma) \) is always either:

- The beginning \( s_1 \) of the intervals \([s_i, t_k] \) of \( B \) (such as \( B \) in the figure);
- A node that is \( \text{shared} \) by backbone pointers from two nodes in \( A \) (such as \( A \));
- The value \( \text{null} \).

A bound on the number of \( s_i \)’s is given in the modular specification of \( \text{proc} \). A bound on the number of shared nodes is given in the next subsection. This bound is effective for all the procedure calls in \( \text{prog} \); hence \( \text{en}^{\text{B/A}} \) can be used in the restricted logic \( \text{AE}^{\text{AR}} \).

5.4 Bounding the Amount of Sharing

We show that under the restrictions of the specification given in Section 4 and Section 4 the number of shared nodes inside \( \text{mod} \) — that is, nodes in mod that are pointed to by more than one f-pointer of other nodes in mod — has a fixed bound throughout the procedure’s execution.

Consider a typical loop-free program \( \text{prog} \) containing calls of the form \( v_i := \text{proc}(\pi_i) \). Assume that the mod-set of \( \text{prog} \) is a union of \( k \) intervals. We show how to compute a bound on the number of shared nodes inside the mod-set. Since there are \( k \) start points, at most \( \binom{k}{2} \) elements can be shared when \( \text{prog} \) starts.
executing. Each sub-procedure invoked from \textit{prog} may introduce, by our restriction, at most as many shared nodes as there are local variables in the sub-procedure. Therefore, by computing the sum over all invocation statements in \textit{prog}, plus \( (\frac{k}{2}) \), we get a fixed bound on the number of shared nodes inside the mod-set.

\[
N_{\text{shared}} = k + \left( \frac{k}{2} \right) + \sum_{\text{proc}_i} |\text{Pvar}_{\text{proc}_i}| 
\]

\text{Pvar}_{\text{proc}_i} \text{ signifies the set of local variables in the procedure \textit{proc}_i. Notice that if the same procedure is invoked twice, it has to be included twice in the sum as well.}

5.5 Verification Condition for the Entire Procedure

Since every procedure on its own is loop-free, the verification condition is straightforward:

\[
v_e[\text{prog}] = P_{\text{prog}} \rightarrow wp[\text{prog}] (Q_{\text{prog}} \land \text{“shared }\subseteq\text{ Pvar”})
\]

where “\text{shared }\subseteq\text{ Pvar}” is a shorthand for the \((AE^R)\) formula:

\[
\forall \alpha, \beta, \gamma \in \text{mod} : E_f(\alpha, \gamma) \land E_f(\beta, \gamma) \rightarrow \\
\alpha = \beta \lor \bigvee_{v \in \text{Pvar}} \gamma = v
\]  \hspace{1cm} (9)

See eq (6) for the definition of \(E_f\). It expresses the obligation mentioned in Section 1.3 that all the shared nodes in \text{mod} should be pointed to by local variables, effectively limiting newly introduced sharing to a bounded number of memory locations.

Now \(v_e[\text{prog}]\) is expressed in \(AE^R\), and it is valid if-and-only-if the program meets its specification. Its validity can be checked using effectively-propositional logic according to Section 5.

6. Extensions

6.1 Explicit Memory Management

This section sketches how to handle explicit allocation and reclamation of memory and exemplifies it on simple procedures shown in Fig. 16 and Fig. 17. The procedures \text{push} and \text{pushMany} extend a list at the beginning by an unbounded of fresh elements allocated using \text{new}. The procedure \text{deleteAll} takes as argument a list that has no foreign pointers into it, and explicitly frees all elements. We verify that the procedures do not introduce dangling pointers.

Table 5 updates the specification of atomic commands (provided in Table 3) to handle explicit memory management operations. For simplicity, we follow ANSI C semantics but do not handle arrays and pointer arithmetic. The allocation statement assigns a freed location denoted by \(s\) to \text{retval} and sets its value to be non-freed. All accesses to memory locations pointed-to by \(y\) in statements \text{ retval } = y.f, y.f = x, and \(x = y\) are required to access non-freed memory. Finally, \text{free}(y) sets the free predicate to true for the node pointed-to by \(y\). As a result, all the nodes reachable from \(y\) cannot be accessed.

The adaptation rule needs to be augmented in order to accommodate the change. Since nodes that are about to be allocated do not have names, the mod shall refer only to allocated nodes; free nodes can always be changed and they need not be specified in the mod. Of course, the procedure’s post-condition should describe the new structure of the allocated area in terms of reachability, if modular reasoning is desired.

The change would be as follows: whenever there is a reference to some \(\sigma \not\in \text{mod}\) (in eq (3), eq (4), and eq (5)), we would now consider only \(\sigma \not\in \text{mod} \cup \text{free}\). This way the adaptation rule makes no claims regarding the \text{free} nodes. Everything else remains just the same.

\text{Node push}(\text{Node } h) \{ \\
\text{Node } e = \text{new} \text{ Node}(); \\
\text{e.f} = h; \\
\text{return } e; \\
\}\n
\text{Node pushMany}(\text{Node } h) \{ \\
\text{if} (\text{?}) \{ \\
\text{h} = \text{push}(h); \\
\text{h} = \text{pushMany}(h); \\
\}\n\text{return } h; \\
\}\n
Figure 16. The procedure \text{push} allocates a new element and inserts it to the beginning of the list. The procedure \text{pushMany} calls \text{push} the same list an arbitrary number of times.

\text{Node deleteAll}(\text{Node } h) \{ \\
\text{if} (\text{?}) \{ \\
\text{j} = h.f; \\
\text{h.f} = \text{null}; \\
\text{free}(h); \\
\text{deleteAll}(j); \\
\}\n\}

Figure 17. The procedure \text{deleteAll} explicitly reclaims the elements of a list dominated by its head, where no other pointers to this list exist.

6.2 Cyclic Linked-Lists

For data structures with a single pointer, the acyclicity restriction may be lifted by using an alternative formulation that keeps and makes no claims regarding the free nodes. Everything else remains just the same.

\text{forall } \alpha, \beta : h(f^*a) \land h(f^*b) \rightarrow (\alpha(f^*b) \leftrightarrow \beta(b^*)a).

\text{forall } \alpha : \text{free}(a) \leftrightarrow \text{free}(a) \land \alpha \neq \text{retval} \\
\text{retval}(f^*)h \land \neg h(f^*)retval
Unfortunately modularity presents another challenge: how the modset should be specified, and how to formulate the adaptation rule. Since there are two pointer fields (forward and backward), the adaptation rule (eq (1)) has to be instantiated twice. However that would require mod to be defined as a union of intervals also according to $\beta$ in addition to its being defined as such using $\alpha$; otherwise our logical arguments from Section 2.3 no longer hold.

When the input is a valid doubly-linked list this can always be done, since $[\alpha, \beta]_{\alpha} \subseteq [\beta, \alpha]_{\beta}$. In cases such where the input is somewhat altered or corrupt (for example $\beta = (x, y)$), the logic will have to be modified to incorporate the volatile exit points of back-pointers potentially pointing to arbitrary nodes. This extension is out of the current scope.

### 7. Experimental Results

#### 7.1 Implementation Details

A VC generator described in Section 5 is implemented in Python, and PLY (Python Lex-Yacc) is employed at the front-end to parse modular recursive procedure specifications as defined in Section 4. The tool checks that the pre and the post-conditions are specified in $AF^{R}$ and that the modset is defined. SMT-LIB v2 [4] standard notation is used to format the VC and to invoke Z3. The validity of the VC can be checked by providing its negation to Z3. If Z3 exhibits a satisfying assignment then that serves as counterexample for the correctness of the assertions. If no satisfying assignment exists, then the generated VC is valid, and therefore the program satisfies the assertions.

The output model/counterexample (S-Expression), if one is generated, is then also parsed and $f^*$ is evaluated on all pairs of nodes. This structure represents the state of the program either at the input or at the beginning of a loop iteration: running the program from this point will violate one or more invariants. To provide feedback to the user, $f$ is recovered by computing eq (2), and then the pygraphviz tool is used to visualize and present to the user a directed graph, whose vertices are nodes in the heap, and whose edges are the $f$ pointer fields.

#### 7.2 Verification Examples

We have written modular specifications for the example procedures shown in Table 7. We are encouraged by the fact that it was not difficult to express assertions in $AF^{R}$ for these procedures. The annotated examples and the VC generation tool are publicly available with the submission. We only picked interesting examples with interesting cutpoints to show the benefits of our approach in contrast to [11].

To give some account of the programs’ sizes, we observe the program summary specification given as pre- and postcondition, count the number of atomic formulas in each of them, and note the depth of quantifier nesting: all our samples had only universal quantifiers in the specification. We did the same for the generated VC: naturally, the depth of quantifier nesting is much larger than the size of the generated domain — which reflects the number of nodes in the heap. As expected, Z3 was able to produce concrete counterexample of reasonable size, producing output which is readable for the programmer and useful for debugging. In fact, our tool converts Z3 models into directed graph diagrams which facil-

<table>
<thead>
<tr>
<th>Command</th>
<th>Pre</th>
<th>Mod</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>retval = new()</td>
<td>free(s)</td>
<td>$\emptyset$</td>
<td>retval = s $\land \neg$free(s)</td>
</tr>
<tr>
<td>access y</td>
<td>$\neg$free(y)</td>
<td>$\emptyset$</td>
<td>$\emptyset$</td>
</tr>
<tr>
<td>free(y)</td>
<td>$y \neq$ null $\land \neg$free(y)</td>
<td>$\emptyset$</td>
<td>free(y)</td>
</tr>
</tbody>
</table>

Table 5. The specifications of atomic commands for resource allocations in a C-like language.

<table>
<thead>
<tr>
<th>Benchmark</th>
<th>Formula size</th>
<th>Solving time (Z3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLL: filter</td>
<td>$\forall P, Q$ $\exists \alpha, \beta [s \land \neg y \rightarrow free(s) \land \neg free(y)]$</td>
<td>0.48s</td>
</tr>
<tr>
<td>SLL: quicksort</td>
<td>$\forall P, Q$ $\exists \alpha, \beta [s \land \neg y \rightarrow free(s) \land \neg free(y)]$</td>
<td>1.06s</td>
</tr>
<tr>
<td>SLL: insert-sort</td>
<td>$\forall P, Q$ $\exists \alpha, \beta [s \land \neg y \rightarrow free(s) \land \neg free(y)]$</td>
<td>0.37s</td>
</tr>
<tr>
<td>UF: find</td>
<td>$\exists \alpha, \beta [s \land \neg y \rightarrow free(s) \land \neg free(y)]$</td>
<td>0.40s</td>
</tr>
<tr>
<td>UF: union</td>
<td>$\exists \alpha, \beta [s \land \neg y \rightarrow free(s) \land \neg free(y)]$</td>
<td>1.39s</td>
</tr>
</tbody>
</table>

Table 6. Description of some pointer manipulating programs verified by our tool.

Table 7. Implementation Benchmarks; $P, Q$ — program’s specification given as pre- and post-condition, mod — mod-set, VC — verification condition, $\#$ — number of atomic formulas/intervals, $\forall$ — quantifier nesting. The tests were conducted on a 1.7GHz Intel Core i5 machine with 4GB of RAM, running OS X 10.7.5. The version of Z3 used was 4.2, complied for 64-bit Intel architecture (using gcc 4.2, LLVM). The solving time reported is wall clock time of the execution of Z3.

### 7.3 Buggy Examples

We also applied the tool to erroneous programs and programs with incorrect assertions. The results, including run-time statistics and formula sizes, are reported in Table [7]. The table lists four kinds of deliberately-introduced bugs that were provided as input to the tool. Formula sizes are measured in the same way as in Table 7.

For Table 8, apart from find and union, we also checked full functional correctness of the other examples (filter, quicksort, insertion sort). For filter, we checked that elements remain in the same order and that only the elements satisfying the filtering predicate were removed. For the sorting routines, we checked that the resulting list contains the same elements and is indeed sorted (via an order relation).

**Checked properties.** For Table 8, apart from find and union, we also checked full functional correctness of the other examples (filter, quicksort, insertion sort). For filter, we checked that elements remain in the same order and that only the elements satisfying the filtering predicate were removed. For the sorting routines, we checked that the resulting list contains the same elements and is indeed sorted (via an order relation).


### 8. Related Work

#### 8.1 Modular Verification

The area of modular procedure specification is heavily studied. Many of these works require that the user declare potential changes similar to the modset (e.g., see \[2\][14][25][29]). The frame rule of separation logic \([9]\) naturally supports modular reasoning where the separating conjunction combines the local postcondition with the assertion at the call site. Unlike separation, reachability is a higher abstraction which relies on type correctness and naturally abstracts operations such as garbage collection. Nevertheless, in Section 6.1 we show that it can also deal with explicit memory relocations.

We believe that our work in this paper pioneers the usage of an effectively propositional logic which is a weak logic to perform modular reasoning in a sound and complete way. Our adaptation rule is more complex than the frame rule as it automatically updates reachability. The idea of using the idempotent entry point function to enable local reasoning about list-manipulating programs (including an EPR reduction) has been explored independently by Piskac et al. \([19]\), where it was used to automate the frame rule in separation logic. In this paper we identify a general EPR fragment of assertions for which this idea of the idempotent entry point function is sound and complete.

<table>
<thead>
<tr>
<th>Benchmark (+ Nature of defect)</th>
<th>Formula size</th>
<th>Solving time (Z3)</th>
<th>C.e. size (l4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>UF: find Incorrect handling of corner case</td>
<td>27 3 201 6</td>
<td>1.60s</td>
<td>2</td>
</tr>
<tr>
<td>UF: union Incorrect specification</td>
<td>19 2 186 6</td>
<td>0.70s</td>
<td>8</td>
</tr>
<tr>
<td>SLL: filter Uncontrolled sharing</td>
<td>36 4 317 6</td>
<td>0.49s</td>
<td>14</td>
</tr>
<tr>
<td>SLL: insert-sort Unmet call precondition</td>
<td>21 2 283 9</td>
<td>0.88s</td>
<td>8</td>
</tr>
</tbody>
</table>

Table 8. Information about benchmarks that demonstrate detection of several kinds of bugs in pointer programs. In addition to the previous measurements, the last column lists the size of the generated counterexample in terms of the number of vertices — linked-list or tree nodes.

Table 8. Information about benchmarks that demonstrate detection of several kinds of bugs in pointer programs. In addition to the previous measurements, the last column lists the size of the generated counterexample in terms of the number of vertices — linked-list or tree nodes.

27 3 201 6 1.60s 2
19 2 186 6 0.70s 8
36 4 317 6 0.49s 14
21 2 283 9 0.88s 8

#### 8.2 Decision Procedures

Many decision procedures for reasoning about linked lists have been proposed \([4][15][7][28]\). All these logics are based on monadic second-order logic on trees which has a non-elementary time (and space) asymptotic complexity. We follow \([11]\) in using a weak logic which permits sound and complete reasoning using off the shelf SAT solvers which are efficient in practice and can be implemented in polynomial space. Indeed our preliminary experimental results reported Section \([7]\) show that Z3 is fast enough and may be even useful for automatically generating abstract interpreters as suggested by \([20]\).

Interestingly, the adaptation rule drastically simplifies the Weakest-Precondition rules given in \([11]\). Notice the specifications in Table \(2\) do not use quantifiers at all, whereas in \([11]\) the formulas contain quantifiers with alternations. Indeed the appropriate quantifiers are added in a generic manner by the adaptation rule and weakest-precondition.

#### 8.3 Incremental Reachability Update

Formulas for updating transitive closure w.r.t. graph mutations have been developed by various authors (e.g., \([5][7][8][13]\)). These works assume that a single edge is added and deleted. This submission generalizes these results to procedures which perform unbounded mutations. Indeed our adaptation rule generalizes \([7][11][13]\) which provides reachability update formula w.r.t. the removal of a single edge.

#### 9. Conclusion

A crucial method for simplifying the reasoning about linked data structures is partitioning them into basic blocks, where each basic block has only one entry point and one exit point. This paper slightly generalizes by reasoning about blocks with a potentially unbounded number of entry points, as demonstrated by \([\text{find}]\) and \([\text{union}]\). Notice that this unboundedness supports modularity: even in the case where in every particular call context there is a bounded number of paths (e.g. when there is a bounded number of roots in the heap), the bound is not known in advance, therefore the programmer has to prepare for an unbounded number of cases.

It is important to note that the adaptation rule adds expressive power to verifying programs: it is in general impossible for the programmer to define, in \(AP^B\), a modular specification for all the procedures. Generation of a verification condition requires coordination between the separate call sites as mentioned above, in particular taking note of potential sharing. This coordination requires per-call-site instantiation, which, thanks to having the adaptation rule in the framework, is done automatically.

Finally we remark that there is a trade-off between \(mod\) and the post-condition: defining a simpler, but larger \(mod\) may cause the post-condition to become more complicated, sometimes not even \(AP^B\)-expressible. Also notice that if \(mod = V\) (the entire heap), modular reasoning becomes trivial since it can be done by relational composition, but this puts the burden of writing the most complete post-conditions on the programmer, which sometimes is not even possible in a limited logic.

Therefore, we believe that this paper takes a step towards modular reasoning about reachability in programs that manipulate linked lists. Lifting such reasoning to more complex data structures such as trees and graphs remains future work.

**Acknowledgement.** Thanks to the anonymous referees for their comments. Itzhaky, Lahav and Sagiv were funded by the European Research Council under the European Union’s Seventh Framework Program (FP7/2007-2013) / ERC grant agreement no. [321174-VSSC] and by a grant from the Israel Science Foundation (652/11). Banerjee and Nanveski were partially supported by
by Spanish MINECO projects TIN2009-14599-C03-02 Desafios, TIN2010-20639 Paran10, TIN2012-39391-C04-01 Strongsoft, EU NoE Project 256980 Nessos, AMAROUT grant PCOFUND-GA-2008-229599, and Ramon y Cajal grant RYC-2010-0743. Immerman was partially supported by NSF grant CCF 1115448.

References


