From Traces To Proofs: Proving Concurrent Programs Safe

Chinmay Narayan*, Subodh Sharma†, Shibashis Guha‡ and S.Arun-Kumar§
Department of Computer Science and Engineering,
Indian Institute of Technology Delhi
Email: *chinmay@cse.iitd.ac.in, †svs@cse.iitd.ac.in, ‡shibashis@cse.iitd.ac.in, §sak@cse.iitd.ac.in

Abstract—Nondeterminism in scheduling is the cardinal reason for difficulty in proving correctness of concurrent programs. A powerful proof strategy was recently proposed [6] to show the correctness of such programs. The approach captured data-flow dependencies among the instructions of an interleaved and error-free execution of threads. These data-flow dependencies were represented by an *inductive data-flow graph* (iDFG), which, in a nutshell, denotes a set of executions of the concurrent program that gave rise to the discovered data-flow dependencies. The iDFGs were further transformed in to alternative finite automatons (AFAs) in order to utilize efficient automata-theoretic tools to solve the problem. In this paper, we give a novel and efficient algorithm to directly construct AFAs that capture the data-flow dependencies in a concurrent program execution. We implemented the algorithm in a tool called `ProofTraPar` to prove the correctness of finite state cyclic programs under the sequentially consistent memory model. Our results are encouraging and compare favorably to existing state-of-the-art tools.

I. INTRODUCTION

The problem of checking whether or not a correctness property (specification) is violated by the program (implementation) is already known to be challenging in a sequential set-up, let alone when programs are implemented exploiting concurrency. The central reason for greater complexity in verification of concurrent implementations is due to the exponential increase in the number of executions. A concurrent program with n threads and k instructions per thread can have $(nk)!/(k!)^n$ executions under a sequentially consistent (SC)[12] memory model. A common approach to address the complexity due to the exponential number of executions is *trace partitioning*.

In [6], a powerful proof strategy was presented which utilized the notion of trace partitioning. Let us take Peterson's algorithm in Figure 1 to convey the central idea behind trace partitioning approach. In this algorithm, two processes, P_i and P_j, coordinate to achieve an exclusive access to a critical section (CS) using shared variables. A process P_i will busy-wait if P_j has expressed interest to enter its CS and t is j.

In order to prove the mutual exclusion (ME) property of Peterson's algorithm, we must consider the boolean conditions of the while loops at control locations 3 and 8. ME property is established only when at most one of these conditions is false under any interpretation of the program, *i.e.*, ME must be shown to hold true on unbounded number of traces (trace is "a sequence of events corresponding to an interleaved execution

flag$_i$ = $false$, flag$_j$ = $false$, t = 0;

P_i
$While(true)\{$
1. flag$_i$:=true;
2. t:=j;
3. while(flag$_j$ = true & t = j);
4. //Critical Section
5. flag$_i$:=false;
}

P_j
$While(true)\{$
6. flag$_j$:=true;
7. t:=i;
8. while(flag$_i$ = true & t = i);
9. //Critical Section
10. flag$_j$:=false;
}

Fig. 1: Peterson's algorithm for two processes P_1 and P_2

of processes in the program"[9]) generated due to unbounded number of unfoldings of the loops. Notice that events at control locations 3 and 8 are *data-dependent* on events from control locations $2, 6, 7$ and $1, 2, 7$, respectively. In any finite prefix of a trace of $P_i \| P_j$ (interleaved execution of P_i and P_j) up to the events corresponding to control location 3 or 8, the last instance of event at control location 2, `lst2`, and the last instance of event at control location 7, `lst7`, can be ordered in only one of the following two ways; either `lst2` appears before `lst7` or `lst2` appears after `lst7`. This has resulted in partitioning of an unbounded set of traces to a set with mere two traces.

When `lst2` appears before `lst7`, then the final value of the variable t is i, thus making the condition at control location 8 to be true. In the other case, when `lst2` appears after `lst7`, the final value of the variable t is j, thereby making the condition at control location 3 evaluate to true. Hence, in no trace both the conditions are false simultaneously. This informal reasoning indicates that both processes can never simultaneously enter in their critical sections. Thus, proof of correctness for Peterson's algorithm can be demonstrated by picking two traces, as mentioned above, from the set of infinite traces and proving them correct. In general, the intuition is that a proof for a single trace of a program can result in pruning of a large set of traces from consideration. To convert this intuition to a feasible verification method, there is a need to construct a formal structure from a proof of a trace σ such that the semantics of this structure includes a set of all those traces that have proof arguments equivalent to proof of σ. The approach in [6] presented *Inductive Data Flow Graphs* (iDFG) as a formal structure to capture data-dependencies among the events of a trace and performs trace partitioning. All traces that have the same iDFG must have the same proof of correctness.
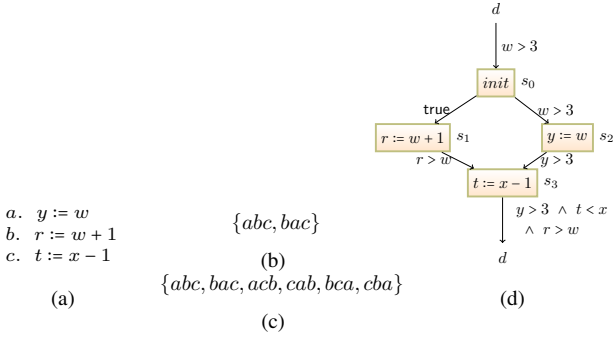
Fig. 2: Comparison with [6]

In every iteration of their approach, a trace is picked from the set of *all* traces that is yet to be covered by the *iDFG*. An iDFG is constructed from its proof. The process is repeated until all the traces are either covered in the *iDFG* or a counter-example is found. An intervening step is involved where the iDFG is converted to an *alternating finite automaton* (AFA). While we explain AFA in later sections, it suffices to understand at this stage that the language accepted by this AFA and the set of traces captured by the corresponding iDFG is the same. Their reason for this conversion is to leverage the use of automata-theoretic operations such as subtraction, complement *etc.*, on the set of traces.

Though the goal of paper [6] is verification of concurrent programs which is the same as in this work, our work has some crucial differences: (i) AFA is constructed directly from the proof of a trace without requiring the iDFG construction, (ii) the verification procedure built on directly constructed AFA is shown to be sound and complete (weakest-preconditions are used to obtain the proof of correctness of a trace), (iii) to the best of our knowledge, we provide the first implementation of the proof strategy discussed in [6]. The example trace of Figure 2(a) highlights the key difference between iDFG to AFA conversion of [6] and the direct approach presented in this work. Note that all three events a, b, and c are data independent, hence every resulting trace after permuting the events in abc also satisfy the same set of pre- and post-conditions. For a Hoare triple $\{w > 3\}\ abc\ \{x = t + 1\ \wedge\ z > 5\}$, Figure 2(b) shows the set of traces admitted by an AFA (obtained from iDFG shown in Figure 2(d)) after the first iteration, as computed by [6]. This set clearly does not represent every permutation of abc; consequently, more iterations are required to converge to an AFA that represents all traces admissible under the same set of pre- and post-conditions. In contrast, the AFA that is constructed directly by our approach from the Hoare triple $\{w > 3\}\ abc\ \{y > 3\ \wedge\ t < x\ \wedge\ r > w\}$, admits the set of traces shown in Figure 2(c). Hence, on this example, our strategy terminates in a single iteration.

To summarize, the contributions of this work are as follows:

- we present a novel algorithm to directly construct an AFA from a proof of a sequential trace of a finite state (possibly cyclic) concurrent program. This construction is used to give a sound and complete verification procedure along the lines of [6] and is potentially faster.

- While [6] allowed the use of any sequential verification method to construct a proof of a given trace, the paper does not comment on the performance and the feasibility of their approach due to the lack of an implementation. The second contribution of this paper is an implementation in the form of a tool, `ProofTraPar`. We compare our implementation against other state-of-the-art tools in this domain, such as THREADER [10] and Lazy-CSeq [11] (winners in the concurrency category of the software verification competitions held in 2013, 2014, and 2015). `ProofTraPar`, on an average, performed an order of magnitude better than THREADER and 3 times better than Lazy-CSeq.

The paper is organized as follows: Section II covers the notations, definitions and programming model used in this paper; Section III presents our approach with the help of an example to convey the overall idea and describes in detail the algorithms for constructing the proposed alternating finite automaton along with their correctness proofs. This section ends with the overall verification algorithm with the proof of its soundness and completeness for finite state concurrent programs. Section IV presents the experimental results and comparison with existing tools namely THREADER [10] and Lazy-CSeq [11]. Section V presents the related work and Section VI concludes with possible future directions.

II. PRELIMINARIES

A. Program Model

We consider shared-memory concurrent programs composed of a fixed number of deterministic sequential processes and a finite set of shared variables SV. A *concurrent program* is a quadruple $\mathcal{P} = (P, A, \mathcal{I}, \mathcal{D})$ where P is a finite set of processes, $A = \{A_p \mid p \in P\}$ is a set of automata, one for each process specifying their behaviour, $\mathcal{D}$ is a finite set of constants appearing in the syntax of processes and $\mathcal{I}$ is a function from variables to their initial values. Each process $p \in P$ has a disjoint set of local variables LV_p. Let Exp_p (BExp_p) denote the set of expressions (boolean expressions), ranged over by $\exp$ (ϕ) and constructed using shared variables, local variables, $\mathcal{D}$, and standard mathematical operators. Each specification automaton A_p is a quadruple $\langle Q_p, q_p^{init}, \delta_p, \mathrm{Assrn}_p \rangle$ where Q_p is a finite set of control states, q_p^{init} is the initial state, and $\mathrm{Assrn}_p \subseteq Q_p \times \mathsf{BExp}_p$ is a relation specifying the assertions that must hold at some control state. Each transition in δ_p is of the form (q, op_p, q') where $op_p \in \{\mathrm{x{:}{=}exp}, \mathtt{assume}(\phi), \mathtt{lock}(x)\}$. Here x:=exp evaluates $\exp$ in the current state and assign the value to x where $\mathrm{x} \in \mathsf{SV} \cup \mathsf{LV}_p$. $\mathtt{assume}(\phi)$ is a blocking operation that suspends the execution if the boolean expression ϕ evaluates to false otherwise it acts as nop. This instruction is used to encode control path conditions of a program. $\mathtt{lock}(x)$, where $x \in \mathsf{SV}$, is a blocking operation that suspends the execution if the value of x is not equal to 0 otherwise it assigns 1 to x. Operation unlock is achieved by Assigning 0 to this shared variable. Each of these operations are deterministic in nature, i.e. execution of any two same operations from same states always give the same behaviour. As an
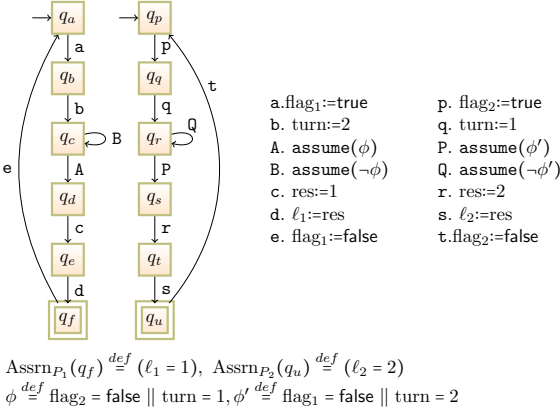
$$\text{Assrn}_{P_1}(q_f) \stackrel{def}{=} (\ell_1 = 1),\ \text{Assrn}_{P_2}(q_u) \stackrel{def}{=} (\ell_2 = 2)$$
$$\phi \stackrel{def}{=} \text{flag}_2 = \text{false} \ ||\ \text{turn} = 1, \phi' \stackrel{def}{=} \text{flag}_1 = \text{false} \ ||\ \text{turn} = 2$$

Fig. 3: Specification of Peterson's algorithm

example, Figure 3 shows the specification of two processes in Peterson's algorithm. Variable res is introduced to specify the mutual exclusion property as a safety property. A process P_i sets this variable to i inside its critical section. Assertions $\text{assert}(\text{res} = \text{i})$ is checked in P_i before leaving its critical seciton. If these assertions hold in every execution of these two processes then the mutual exclusion property holds. A tuple, say t, of n elements can be represented as a function such that $t[k]$ returns the kth element of this tuple. Given a function fun, $fun[a \leftarrow b]$ denote another function same as of fun except at a where it returns b.

a) Parallel Composition in the SC memory model:
Given a concurrent program $\mathcal{P} = (P, A, \mathcal{I}, \mathcal{D})$ consisting of n processes $P = \{p_1, \cdots, p_n\}$ we define an automaton $\mathcal{A}(\mathcal{P}) = (\overline{Q}, \overline{q}^{init}, \overline{\delta}, \overline{\text{Assrn}})$ to represent the parallel composition of $\mathcal{P}$ in the SC memory model. Here $\overline{Q} = Q_{p_1} \times \cdots \times Q_{p_n}$ is the set of states ranged over by $\overline{q}$, $\overline{q}^{init} = (q_{p_1}^{init}, \cdots, q_{p_n}^{init})$ is the initial state, and transition relation $\overline{\delta}$ models the interleaving semantics. Formally, $(\overline{q}, op_j, \overline{q}') \in \overline{\delta}$ iff there exists a $j \in \{1 \cdots n\}$ such that $\overline{q}[j] = q_{p_j}$, $\overline{q}' = \overline{q}[j \leftarrow q'_{p_j}]$ and $(q_{p_j}, op_j, q'_{p_j}) \in \delta_{p_j}$. For a state $\overline{q}$, let $T(\overline{q}) = \{\text{Assrn}_{p_i}(\overline{q}[i]) \mid i \in \{1 \cdots n\}\}$. If $T(\overline{q})$ is not empty then $\overline{\text{Assrn}}(\overline{q})$ is the conjunction of assertions in the set $T(\overline{q})$. Relation $\overline{\text{Assrn}}$ captures the assertions which need to be checked in the interleaved traces of P. As our interest lies in analyzing those traces which reach those control points where assertions are specified hence we mark all those states where the relation $\overline{\text{Assrn}}$ is defined as accepting states. Every word accepted by $\mathcal{A}(\mathcal{P})$ represents one SC execution leading to a control location where at least one assertion is to be checked.

B. Weakest Precondition

Given an operation $op \in \mathcal{OP}(P)$ and a postcondition formula ϕ, the weakest precondition *of op with respect to* ϕ, denoted by $\text{wp}(op, \phi)$, is the weakest formula ψ such that, starting from any program state s that satisfies ψ, the execution of the operation op terminates and the resulting program state s' satisfies ϕ.

Given a formula ϕ, variable X and expression e, let $\phi[\text{X}/e]$ denote the formula obtained after substituting all free occurrences of X by e in ϕ. We assume an equality operator over formulae that represents syntactic equality. Every formula

$$\text{wp}(op, \phi) = \begin{cases} \phi[\text{x}/\text{exp}] & \text{if } op \stackrel{def}{=} \text{x:=exp} \\ \phi \ \wedge \ \phi' & \text{if } op \stackrel{def}{=} \text{assert}(\phi') \\ \text{wp}(\text{assume}(\text{x}=0), \text{wp}(\text{x:=1}, \phi)) & \text{if } op \stackrel{def}{=} \text{lock}(x) \\ \text{wp}(op_1, \text{wp}(op_2, \phi)) & \text{if } op \stackrel{def}{=} op_1.op_2 \\ \phi' \Rightarrow \phi & \text{if } op \stackrel{def}{=} \text{assume}(\phi') \end{cases}$$

Fig. 4: Weakest precondition axioms

is assumed to be normalized in a conjunctive normal form (CNF). We use true (false) to syntactically represent a logically valid (unsatisfiable) formula. Weakest precondition axioms for different program statements are shown in Figure 4. We have the following properties about weakest preconditions.

Property 1: If $\text{wp}(S, \phi_1) = \psi_1$ and $\text{wp}(S, \phi_2) = \psi_2$ then,
- $\text{wp}(S, \phi_1 \ \wedge \ \phi_2) = \psi_1 \ \wedge \ \psi_2$, and
- $\text{wp}(S, \phi_1 \ \vee \ \phi_2) = \psi_1 \ \vee \ \psi_2$. Note that this property holds only when S is a deterministic operation which is true in our programming model.

Property 2: Let ϕ_1 and ϕ_2 be the formulas such that ϕ_1 logically implies ϕ_2 then for every operation op, the formula $\text{wp}(op, \phi_1)$ logically implies $\text{wp}(op, \phi_2)$.

We say that a formula ϕ is *stable* with respect to a statement S if $\text{wp}(S, \phi)$ is logically equivalent to ϕ. In this paper, we use weakest preconditions to check the correctness of a trace with respect to some safety assertion. A trace σ reaching upto a safety assertion ϕ is safe if the exeuction of σ starting from the initial state $\mathcal{I}$ either 1) blocks(does not terminate) because of not satisfying some path conditions, or 2) terminates and the resulting state satisfies ϕ. Following lemmas clearly define the conditions, using weakest precondition axioms, for declaring a trace σ either safe or unsafe. Detailed proofs of these are given in Appendix A and in B. Here $\sigma[\text{assume/assert}]$ denote the trace obtained by replacing every instruction of the form $\text{assume}(\phi)$ by $\text{assert}(\phi)$ in σ.

Lemma 1: For a trace σ, an initial program state $\mathcal{I}$ and a safety property ϕ, if $\text{wp}(\sigma[\text{assume/assert}], \neg\phi) \ \wedge \ \mathcal{I}$ is unsatisfiable then the execution of σ, starting from $\mathcal{I}$, either does not terminate or terminates in a state satisfying ϕ.

Lemma 2: For a trace σ, an initial program state $\mathcal{I}$ and a safety property ϕ, if $\text{wp}(\sigma[\text{assume/assert}], \neg\phi) \ \wedge \ \mathcal{I}$ is satisfiable then the execution of σ, starting from $\mathcal{I}$, terminates in a state not satisfying ϕ.

C. Alternating Finite Automata (AFA)

Alternating finite automata [1],[3] are a generalization of nondeterministic finite automata (NFA). An NFA is a five tuple $\langle S, \Sigma \cup \{\epsilon\}, \delta, s_0, S_F \rangle$ with a set of states S, ranged over by s, an initial state s_0, a set of accepting states S_F and a transition function $\delta : S \times \Sigma \cup \{\epsilon\} \rightarrow \mathbb{P}(S)$. For any state s of this NFA, the set of words accepted by s is inductively defined as $\text{acc}(s) = \{a.\sigma \mid a \in \Sigma \cup \{\epsilon\}, \sigma \in \Sigma^*, \exists s'.\ s' \in \delta(s,a).\ \sigma \in \text{acc}(s')\}$ where $\epsilon \in \text{acc}(s)$ for all $s \in S_F$. Here, the existential quantifier represents the fact that there should exist at least one outgoing transition from s along which $a.\sigma$ gets accepted. An AFA is a six tuple $\langle S_\forall, S_\exists, \Sigma \cup \{\epsilon\}, \delta, s_0, S_F \rangle$ with Σ, s_0 and $S_F \subseteq S$ denoting the alphabet, initial state and the set of accepting states respectively. $S = S_\forall \bigcup S_\exists$ is the set of all states, ranged

over by s and $\delta : S \times \Sigma \cup \{\epsilon\} \to \mathbb{P}(S)$ is the transition function. The set of words accepted by a state of an AFA depends on whether that state is an *existential state* (from the set $S_\exists$) or a *universal state* (from the set $S_\forall$). For an existential state $s \in S_\exists$, the set of accepted words is inductively defined in the same way as in NFA. For a universal state $s \in S_\forall$ the set of accepted words are $\text{acc}(s) = \{a.\sigma \mid a \in \Sigma \cup \{\epsilon\}, \forall s' \in \delta(s, a). \ \sigma \in \text{acc}(s')\}$ with $\epsilon \in \text{acc}(s)$ for all $s \in S_F$. Notice the change in the quantifier from $\exists$ to $\forall$. In the diagrams of AFA used in this paper, we annotate universal states with $\forall$ symbol and existential states with $\exists$ symbol. For a state s, let $\text{succ}(s, a) = \{S \mid (s, a, S) \in \delta\}$ be the set of a-successors of s. For an automaton $\mathcal{A}$, let $\mathcal{L}(\mathcal{A})$ be the language accepted by the initial state of that automaton. For any $\sigma \in \Sigma^* \ |\sigma|$ denote the length of σ and $\text{rev}(\sigma)$ denote the reverse of σ.

III. Our Approach

The overall approach of this paper can be broadly described in the following steps: (i) Given a concurrent program $\mathcal{P}$, construct its interleaved traces, say Tr as in Subsection II-A; (ii) Pick a trace σ and a safety property, say ϕ, to prove for this trace; (iii) Prove σ correct with respect to ϕ using weakest precondition method, Lemma 1 and Lemma 2, and generate a set of traces which are also *provably correct* following the same reasoning as of σ. Let us call this set Tr'; (iv) Remove set Tr' from Tr and repeat from Step (ii) until either all the interleaved traces of $\mathcal{P}$ are proved correct or an erroneous trace is found.

Step (iii) of this procedure, correctness of σ, can be achieved by checking the unsatisfiability of $\text{wp}(\sigma[\text{assume/assert}], \neg\phi) \wedge \mathcal{I}$. However, we are not only interested in checking the correctness of σ but also in constructing a set of traces which have a similar reasoning as of σ. Therefore, instead of computing $\text{wp}(\sigma[\text{assume/assert}], \neg\phi)$ directly from the weakest precondition axioms of Figure 4, we construct an AFA from σ and $\neg\phi$. Step (iv) is then achieved by applying automata-theoretic operations such as complementation and subtraction on this AFA. Notion of universal and existential states of AFA helps us in finding a set of sufficient dependencies used in the weakest precondition computation so that any other trace satisfying those dependencies gets captured by AFA. Subsequent subsections covers the construction, properties and use of this AFA in detail.

A. Constructing the AFA from a Trace and a Formula

Definition 1: 1 An AFA constructed from a trace σ of a Program P and a formula ϕ is $\hat{\mathcal{A}}_{\sigma,\phi} = \langle S_\forall, S_\exists, \mathcal{OP}_\epsilon, s_0, S_F, \delta, \text{AMap}, \text{RMap} \rangle$, where,

1) $(\mathcal{OP}_\epsilon = \mathcal{OP} \cup \{\epsilon\})$ is the alphabet ranged over by op. Here $\mathcal{OP}$ is the set of instructions used in program P. Symbol ϵ acts as an identity element of concatenation and $\text{wp}(\epsilon, \phi) = \phi$.
2) $S = S_\forall \bigcup S_\exists$ is the largest set of states, ranged over by s, s.t.
   a) Every state is annotated with a formula and a prefix of σ denoted by $\text{AMap}(s)$ and $\text{RMap}(s)$ respectively.

$\delta(s, op) =$

$$\{s'\} \quad \text{if} \quad \begin{cases} 1.\text{AMap}(s') = \text{wp}(op[\text{assume/assert}], \text{AMap}(s)), \\ 2.\text{RMap}(s) = \text{RMap}(s').op.\sigma'' \\ \textbf{where } \sigma'' \text{ is the longest sequence s.t.} \\ \text{wp}(\sigma''[\text{assume/assert}], \text{AMap}(s)) = \text{AMap}(s) \end{cases}$$
$$\text{(Literal-Assn)}$$

$$\{s\} \quad \text{if} \quad \begin{cases} 1.\text{AMap}(s) = \text{wp}(op[\text{assume/assert}], \text{AMap}(s)), \\ 2.(s) \text{ is an existential state, and} \\ 3.s \text{ is an existential state} \end{cases}$$
$$\text{(Literal-Self-Assn)}$$

$$\{s_1, \cdots, s_k\} \quad \text{if} \quad \begin{cases} 1.\text{AMap}(s) = \bigwedge_k \phi_k \ \textbf{or } \text{AMap}(s) = \bigvee_k \phi_k, \\ 2.\text{AMap}(s_k) = \phi_k, \\ 3.\forall k, \text{RMap}(s) = \text{RMap}(s_k), \\ 4.op = \epsilon \end{cases}$$
$$\text{(Compound-Assn)}$$

$$\{\} \quad \text{otherwise}$$

Fig. 5: Transition function used in the Definition 1

   State s_0 is the initial state such that $\text{AMap}(s_0) = \phi$, $\text{RMap}(s_0) = \sigma$.
   b) $s' \in S$ iff either of the following two conditions hold,
   - $\exists s \in S$ such that $\text{AMap}(s')$ is $\text{wp}(op[\text{assume/assert}], \text{AMap}(s))$, $\text{RMap}(s) = \text{RMap}(s').op.\sigma'$ and σ' is the largest suffix of $\text{RMap}(s)$ such that formaula $\text{AMap}(s)$ is stable with respect to $\sigma'[\text{assume/assert}]$.
   - $\exists s \in S$ such that $\text{AMap}(s) = \bigwedge \Phi$ or $\text{AMap}(s) = \bigvee \Phi$, $\text{RMap}(s) = \text{RMap}(s')$, $\text{AMap}(s') = \phi'$ and $\phi' \in \Phi$.
   c) A state $s \in S$ is an existential state (universal state) iff $\text{AMap}(s)$ is a literal (compound formula).
3) $s_F \subseteq S$ is a set of accepting states such that $s \in S_F$ iff $\text{wp}(\text{RMap}(s)[\text{assume/assert}], \text{AMap}(s))$ is same as $\text{AMap}(s)$, i.e. $\text{AMap}(s)$ is *stable* with respect to $\text{RMap}(s)[\text{assume/assert}]$, and
4) Function $\delta : S \times \mathcal{OP}_\epsilon \to \mathbb{P}(S)$ is defined in Figure 5.

Every state added to S, following Point 2b, is annotated with either a smaller RMap or if same RMap then a smaller formula compared to the states already present in S. Further, every formula and trace σ is of finite length. Hence the set of states S is finite. By Point 2c of this construction, a state s where $\text{AMap}(s)$ is a compound formula, is always a universal state irrespective of whether $\text{AMap}(s)$ is a conjunction or a disjunction of clauses. The reason behind this decision will be clear shortly when we will use this AFA to inductively construct the weakest precondition $\text{wp}(\sigma[\text{assume/assert}], \phi)$. Note that we assume every formula is normalized in CNF.

Let us explain the construction of $\hat{\mathcal{A}}_{\sigma,\phi}$ with an example trace $\sigma = \text{abApqPrcs}$ of Peterson's algorithm, shown in Figure 7. This trace is picked from the Peterson's specification in Figure 3. To prove σ correct with respect to the safety formula $\phi \stackrel{def}{=} (\ell_2 = 2)$ we first construct $\hat{\mathcal{A}}_{\sigma,\neg\phi}$ for trace σ and $\neg\phi$ which will later help us to derive $\text{wp}(\sigma[\text{assume/assert}], \neg\phi)$. This AFA is shown in Figure 6. For a state s, $\text{AMap}(s)$ is written inside the rectangle representing that state and $\text{RMap}(s)$ is written inside an ellipse next to that state. We show here some of the steps illustrating this construction.

1) By Definition 1, we have $\text{AMap}(s_0) = \neg(\ell_2 = 2)$ and $\text{RMap}(s_0) = \sigma = \text{abApqPrcs}$ for initial state s_0.
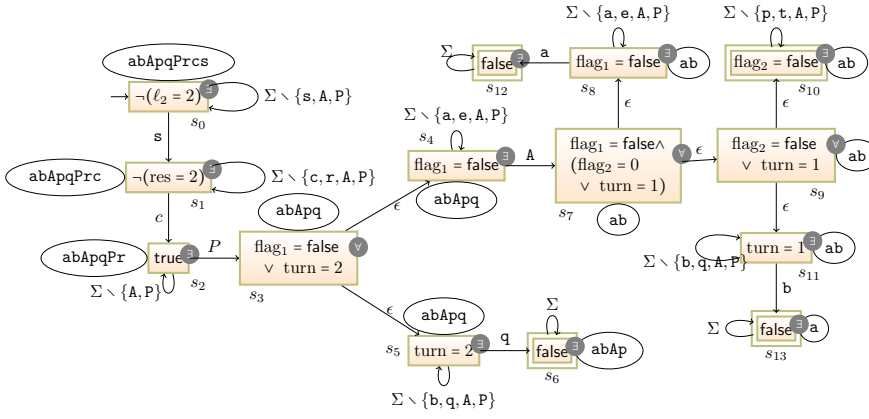
a. $\text{flag}_1 := \text{true}$
b. $\text{turn} := 2$
A. $\text{assume}\left(\begin{array}{l}\text{flag}_2 = \text{false}\\ \parallel\ \text{turn} = 1\end{array}\right)$
p. $\text{flag}_2 := \text{true}$
q. $\text{turn} := 1$
P. $\text{assume}\left(\begin{array}{l}\text{flag}_1 = \text{false}\\ \parallel\ \text{turn} = 2\end{array}\right)$
r. $\text{res} := 2$
c. $\text{res} := 1$
s. $\ell_2 := \text{res}$

2) In any transition created by Rule LITERAL-ASSN, say from s to s' on op, the state s' is annotated with the weakest precondition of an operation op, taken from $\text{RMap}(s)$, with respect to $\text{AMap}(s)$. Operation op is picked in such a way that $\text{AMap}(s)$ is *stable* with respect to every other operation present after op in $\text{RMap}(s)$. Such transitions capture the inductive construction of the weakest precondition for a given ϕ and trace σ. Transitions $\delta(s_0, \text{s}) = \{s_1\}$ in Figure 6 is created by this rule as $\text{wp}(\text{s}[\text{assume/assert}], \text{AMap}(s_0)) = \text{AMap}(s_1)$, and $\text{RMap}(s_0) = \text{RMap}(s_1).\text{s}$ hold. Similarly, the transition $\delta(s_2, \text{P}) = s_3$ follows from the same rule and the facts that $\text{wp}(\text{P}[\text{assume/assert}], \text{AMap}(s_2)) = \text{AMap}(s_3)$, $\text{RMap}(s_2) = \text{RMap}(s_3).\text{P}.\text{r}$, and $\text{wp}(\text{r}, \text{AMap}(s_2)) = \text{AMap}(s_2)$ hold.

3) In any transition created by Rule COMPOUND-ASSN, say from s to $s_1, \cdots, s_k$, the states $s_1, \cdots s_k$ are annotated with the subformulae of $\text{AMap}(s)$. For example, transitions $\delta(s_3, \epsilon) = \{s_4, s_5\}$ and $\delta(s_7, \epsilon) = \{s_8, s_9\}$.

4) Transition $\delta(s_8, \text{a}) = \{s_{12}\}$ follows from the rule LITERAL-ASSN. Note that $\text{RMap}(s_{12})$ is empty and hence by Point 3 of Definition 1, s_{12} is an accepting state. Following the same reasoning, states s_6, s_{10} and s_{13} are also set as accepting states.

5) Rule LITERAL-SELF-ASSN adds a self transition at a state s on a symbol $op \in \mathcal{OP}_\epsilon$ such that $\text{AMap}(s)$ is *stable* with respect to $op[\text{assume/assert}]$. Transition $\delta(s_0, \mathcal{OP}_\epsilon \smallsetminus \{\text{s}, \text{A}, \text{P}\}) = \{s_0\}$ is one such example.

Following lemma relates $\text{RMap}(s)$ at any state to the set of words accepted by s in this AFA.

Lemma 3: Given a $\sigma \in \mathcal{L}(\mathcal{A}(\mathcal{P}))$ and ϕ, let $\hat{\mathcal{A}}_{\sigma,\phi}$ be the AFA satisfying Definition 1. For every state s of this AFA, the condition $\text{rev}(\text{RMap}(s)) \in \text{acc}(s)$ holds.

Detailed proof of this lemma is given in Appendix C. This lemma uses the reverse of $\text{RMap}(s)$ in its statement because the weakest precondition of a sequence is constructed by scanning it from the end. This can be seen in the transition rule LITERAL-ASSN. As a corollary, $\text{rev}(\sigma)$ is also accepted by this AFA because by Definition 1, $\text{RMap}(s_0)$ is σ. ∎

$$\text{HMap}(s) = $$
$$\begin{cases}
\text{AMap}(s) & \text{if } s \in S_F \\
& \text{(BASE-CASE)} \\
\bigwedge_k \text{HMap}(s_k) & \text{if } \delta(s, \epsilon) = \{s_1, \cdots, s_k\} \text{ and } \text{AMap}(s) = \bigwedge_k \text{AMap}(s_k) \\
& \text{(CONJ-CASE)} \\
\bigvee_k \text{HMap}(s_k) & \text{if } \delta(s, \epsilon) = \{s_1, \cdots, s_k\} \text{ and } \text{AMap}(s) = \bigvee_k \text{AMap}(s_k) \\
& \text{(DISJ-CASE)} \\
\text{HMap}(s') & \text{if } (s, op, \{s'\}) \in \delta \\
& \text{(LIT-CASE)}
\end{cases}$$

Fig. 8: Rules for HMap construction

B. Constructing the weakest precondition from $\hat{\mathcal{A}}_{\sigma,\phi}$

After constructing $\hat{\mathcal{A}}_{\sigma,\phi}$ the rules given in Figure 8 are used to inductively construct and assign a formula, $\text{HMap}(s)$, to every state s of $\hat{\mathcal{A}}_{\sigma,\phi}$. Figure 9 shows the afa of Figure 6 where states are annotated with formula $\text{HMap}(s)$. This formula is shown in the ellipse beside every state. For better readability we do not show $\text{RMap}(s)$ in this figure.

Following Rule BASE-CASE, HMap of s_6, s_{12}, and s_{13} are set to false whereas $\text{HMap}(s_{10})$ is set to $\text{flag}_2 = \text{false}$. By Rule LIT-CASE, HMap of s_5, s_8 and s_{11} are also set to false. After applying Rule DISJ-CASE for transition $\delta(s_9, \epsilon) = \{s_{10}, s_{11}\}$, $\text{HMap}(s_9)$ is set to $\text{flag}_2 = \text{false}$. Similarly, using Rule CONJ-CASE we get $\text{HMap}(s_7)$ as false. Finally, $\text{HMap}(s_0)$ is also set to false. HMap constructed inductively in this manner satisfies the following property;

Lemma 4: Let $\hat{\mathcal{A}}$ be an AFA constructed from a trace and a post condition as in Definition 1 then for every state s of this AFA and for every word σ accepted by state s, $\text{HMap}(s)$ is logically equivalent to $\text{wp}(\text{rev}(\sigma)[\text{assume/assert}], \text{AMap}(s))$.

Here we present the proof outline. Detailed proof is given in Appendix E. First consider accepting states of $\hat{\mathcal{A}}$. For example, states s_6, s_{10}, s_{12} and s_{13} of Figure 9. Following the definition of an accepting state and by the self-loop adding transition rule LITERAL-SELF-ASSN, every word σ accepted by such an accepting state s satisfies $\text{wp}(\text{rev}(\sigma)[\text{assume/assert}], \text{AMap}(s)) = \text{AMap}(s)$. Therefore, setting $\text{HMap}(s)$ as $\text{AMap}(s)$ for these accepting states, as done in Rule BASE-CASE completes the proof for accepting states.
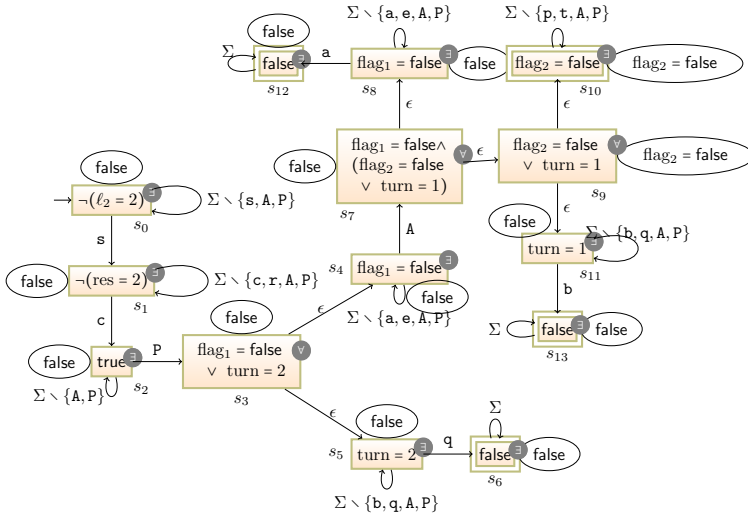
5

Fig. 9: HMap construction for the running example

Algorithm 1: Converting universal to existential states while preserving Lemma 4

Data: Input AFA $\langle S_\forall, S_\exists, \mathcal{OP} \cup \{\epsilon\}, s_0, S_F, \texttt{AMap}, \texttt{RMap}\rangle$
Result: Modified AFA

1. Let s be a state in AFA such that $s \in S_\forall$, $\delta(s, \epsilon) = \{s_1, \cdots, s_k\}$, HMap$(s)$ is unsatisfiable, and AMap$(s) = \bigwedge_k \texttt{AMap}(s_k)$;
2. Let Unsatcore$(s) \subseteq \mathbb{P}(\{s_1, \cdots, s_k\})$ such that $\{s'_1, \cdots, s'_n\} \in$ Unsatcore(s) iff $\{\texttt{HMap}(s'_1), \cdots, \texttt{HMap}(s'_n)\}$ is a minimal unsat core of $\bigwedge_k \texttt{HMap}(s_k)$;
3. Create an empty set U;
4. **foreach** $\{s'_1, \cdots, s'_n\} \in$ Unsatcore(s) **do**
5.   create a new universal state $s_u \in S_\forall$ and add it to the set U;
6.   Set AMap$(s_u) = \bigwedge_i \texttt{AMap}(s'_i)$;
7.   Set HMap$(s_u) = \bigwedge_i \texttt{HMap}(s'_i)$;
8.   Add a transition by setting $\delta(s_u, \epsilon) = \{s'_1, \cdots, s'_n\}$;
9. **end**
10. Remove transition $\delta(s, \epsilon) = \{s_1, \cdots, s_k\}$;
11. Convert s to an existential state;
12. Add a transition from s on ϵ by setting $\delta(s, \epsilon) = $ U where U is the set of universal states created one for each element of Unsatcore(s);

Now consider a state s with transition $\delta(s, \epsilon) = \{s_1, \cdots, s_k\}$, created using Rule COMPOUND-ASSN, and let σ be a word accepted by s. By construction, s must be a universal state and hence σ must be accepted by each of $s_1, \cdots, s_k$ as well. Using this lemma inductively on successor states $s_1, \cdots, s_k$ (induction on the formula size) we get wp$(\sigma[\texttt{assume/assert}], \texttt{AMap}(s_i)) = \texttt{HMap}(s_i)$ for all $i \in \{1 \cdots k\}$. Now we can apply Property 1 depending on whether AMap(s) is a conjunction or a disjunction of AMap(s_k). Thereafter replacing AMap(s) with $\bigvee_k \texttt{AMap}(s_k)(\bigwedge_k \texttt{AMap}(s_k))$ and HMap(s) with $\bigvee_k \texttt{HMap}(s_k)(\bigwedge_k \texttt{HMap}(s_k))$, as done in HMap construction rules CONJ-CASE and DISJ-CASE, completes the proof. Note that, keeping s as a universal state for conjunction as well as disjunction allowed us to use Property 1 in this proof. Otherwise, if we create an existential state s when AMap(s) is a disjunction of formulae then we can not prove this lemma by taking the disjunction of HMap of its successor states. ∎

Above lemma serves two main purposes in our verification algorithm; first it checks the correctness of a trace σ w.r.t. a safety property for which this AFA was constructed. If HMap$(s_0) \wedge \mathcal{I}$ is unsatisfiable, as is true with our peterson's example trace, then σ is declared as correct. Second, this lemma guarantees that every trace accepted by this AFA, that is also present in the set of all traces of $\mathcal{P}$, is also safe and hence we can skip proving their correctness altogether (Removing such traces is equivalent to subtracting the language of this AFA from the language representing the set of all traces). A natural question then is to ask if we can increase the set of accepted words of this AFA while preserving Lemma 4.

C. Enlarging the set of words accepted by $\hat{\mathcal{A}}_{\sigma,\phi}$

Converting Universal States to Existential States Figure 10 shows an example trace $\sigma = $ abcde obtained from the parallel composition of some program P. Figure 10 shows the AFA constructed for $\sigma = $ abcde and $\phi = S < t \wedge z < x$. Using Lemma 4 we get wp$(\texttt{abcde}, \phi) = $ false. Further, note

that the conditions wp$(\texttt{abcde}, S < t) = $ false and wp$(\texttt{abcde}, z < x) = $ false hold, i.e. we have two ways to derive the unsatisfiability of wp$(\texttt{abcde}, \phi)$; one is due to the operation d, and the other is due to the operation a followed (may not be immediately followed) by operation e. In this example, any word that enforces either of these two ways will derive false as the weakest precondition. For example, the sequence $\sigma' = $ adcbe is not accepted by the AFA of Figure 11 but the condition wp$(\texttt{rev}(\texttt{adcbe}), \neg\phi) = $ false follows from wp$(\texttt{d}, \neg\phi) = $ false which is already captured in the AFA of Figure 11. After looking at this AFA, we find two states s_1 and s_2 in Figure 10 with unsatisfiable HMap assertion. It seems sufficient to take any one of these branches to argue the unsatisfiability of HMap(s_0) because HMap(s_0), by definition, is a *conjunction* of HMap(s_1) and HMap(s_2). Therefore, if we convert s_0, a universal state, to an existential state then the modified AFA will accept adcbe. Before arguing the correctness of this modification let us look at Algorithm 1 to see the steps involved in this modification. This algorithm picks a universal state s such that AMap(s) is a *conjunction* of clauses and only a subset of its successors are sufficient to make HMap(s) unsatisfiable. State s_0 of Figure 11 is one such state. For each such minimal subsets of its successors, this algorithm creates a universal state, as shown in Line 5 of this algorithm. It is easy to see that HMap(s_u) is also unsatisfiable. Before adding $\delta(s_u, \epsilon) = \{s'_1, \cdots, s'_n\}$ transition in AFA this algorithm sets AMap(s_u) as $\bigwedge_i \texttt{AMap}(s'_i)$. By construction, every word accepted by s_u must be accepted by $s'_1, \cdots, s'_n$. Each of these states $s'_1, \cdots, s'_n$ satisfy Lemma 4 hence Lemma 4 continues to hold for these newly created universal states as well. Now consider a newly created transition $(s, \epsilon, \texttt{U})$ in Line 8. For any state $s'' \in \texttt{U}$, AMap(s) logically implies AMap(s'') because s'' represents a subset of the original successors of s, viz. $s_1, \cdots, s_k$. As s is now an existential state, any word accepted by s, say σ', is accepted by at least one state in U, say s'. Using Lemma 4 on s', HMap(s') is logically
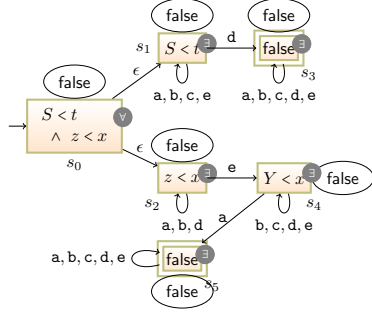
Fig. 10: Example Trace

Fig. 11: AFA for σ given in Figure 10
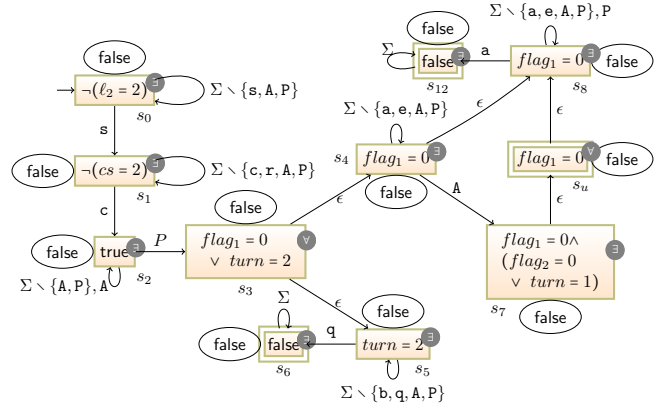

Fig. 12: AFA of Figure 9 after Modification

$$
\begin{cases}
\text{Set } \delta(s, op) = & \text{if HMap}(s) \text{ and HMap}(s') \text{ are unsatisfiable,} \\
\delta(s, op) \cup \{s'\} & \text{AMap}(s) \text{ is a literal, and} \\
& \text{wp}(op[\text{assume/assert}], \text{AMap}(s)) \Rightarrow \text{AMap}(s') \\
& \quad\quad (\text{RULE-UNSAT}) \\
\text{Set } \delta(s, op) = & \text{if HMap}(s) \text{ and HMap}(s') \text{ are valid} \\
\delta(s, op) \cup \{s'\} & \text{AMap}(s) \text{ is a literal, and} \\
& \text{AMap}(s') \Rightarrow \text{wp}(op[\text{assume/assert}], \text{AMap}(s)) \\
& \quad\quad (\text{RULE-VALID})
\end{cases}
$$

Fig. 13: Rules for adding more edges

equivalent to $\text{wp}(\text{rev}(\sigma')[\text{assume/assert}], \text{AMap}(s'))$. Using the knowledge that $\text{HMap}(s)$ and $\text{HMap}(s')$ are unsatisfiable and the monotonicity property of the weakest precondition, Property 2, we get that $\text{HMap}(s)$ is logically equivalent to $\text{wp}(\text{rev}(\sigma')[\text{assume/assert}], \text{AMap}(s))$. This transformation is formally proved correct in Appendix E.

Adding More transitions to $\hat{\mathcal{A}}_{\sigma,\phi}$ using the Monotonicity Property of the Weakest Precondition We further modify $\hat{\mathcal{A}}_{\sigma,\phi}$ by adding more transitions. For any two states s and s' such that $\text{AMap}(s)$ and $\text{AMap}(s')$ are literals, both $\text{HMap}(s)$ and $\text{HMap}(s')$ are unsatisfiable, and there exists a symbol a (can be ϵ as well) such that $\text{wp}(a[\text{assume/assert}], \text{AMap}(s))$ logically implies $\text{AMap}(s')$, an edge labeled a is added from s to s'. This modification also preserves Lemma 4 following the same monotonicity property, Property 2 used in the previous optimization. Similar argument holds when $\text{HMap}(s)$ and $\text{HMap}(s')$ are valid and $\text{AMap}(s') \Rightarrow \text{wp}(a, \text{AMap}(s))$ holds. The rules of adding edges are shown in Figure 13.

Consider the AFA of Figure 9 that was constructed for a trace $\sigma = \text{abApqPrcs}$ of Peterson's example. Figure 12 shows the modified AFA. An edge from s_4 to s_8 on symbol ϵ gets added following Rule RULE-UNSAT because $\text{HMap}(s_4)$ and $\text{HMap}(s_8)$ are unsatisfiable and $\text{wp}(\epsilon, \text{AMap}(s_4))$ logically implies $\text{AMap}(s_8)$. Similarly, because $\text{wp}(\text{P}[\text{assume/assert}], \text{AMap}(s_8))$ logically implies $\text{AMap}(s_8)$ and $\text{HMap}(s_8)$ is unsatisfiable, a *self loop* at s_8 is added on P. By the same argument, a *self loop* at s_2 is added on symbol A. After applying previous transformation, converting universal to existential state, we can safely remove the transition from s_7 to s_9 and all other states

---

Algorithm 2: Algorithm to check the safety assertions of a concurrent program P

Input: A concurrent program $\mathcal{P} = \{p_1, \cdots, p_n\}$ with safety property map $\overline{\text{Assrn}}$

Result: yes, if program is safe else a counterexample

1 Construct the automaton $\mathcal{A}(\mathcal{P})$ of all the SC executions of P as defined in Section II;
2 Set $\texttt{tmp} := \mathcal{L}(\overline{\mathcal{A}(P)})$;
3 **while** $\texttt{tmp}$ *is not empty* **do**
4   Let $\sigma \in \texttt{tmp}$ with ϕ as a safety assertion to be checked;
5   Let $\hat{\mathcal{A}}_{\sigma,\neg\phi}$ be the AFA constructed from σ and $\neg\phi$;
6   **if** $\mathcal{I} \wedge \text{HMap}(s_0)$ *is satisfiable* **then**
7   σ is a valid counterexample violating ϕ;
8   **return** (σ);
9   **else**
10   Let $\hat{\mathcal{A}}'$ be the AFA modified by optimizations;
11          $\texttt{tmp} := \texttt{tmp} \setminus Rev$, where $Rev = \{\text{rev}(\sigma) \mid \sigma \in \mathcal{L}(\hat{\mathcal{A}}')\}$;
12   **end**
13 **end**
14 **return** (yes);

---

reachable from s_9. Now consider a trace $\text{rev}(\text{abpqPArcs})$ that is accepted by this modified AFA in Figure 12 but was not accepted by the original AFA of Figure 9. Note that $\text{wp}(\text{abpqPArcs}, \neg(\ell_2 = 2))$ is unsatisfiable and this is a direct consequence of Lemma 4. Because of the transformation, addition of extra edges and conversion of universal state to an existential state, we do not need to reason about this trace separately. This transformation is formally proved correct in Appendix F.

D. Putting All Things Together For Safety Verification

In Algorithm 2, all the above steps are combined to check if all the SC executions of a concurrent program P satisfy the safety properties specified as assertions. Proof of the following theorem is given in Appendix G.

Theorem 1: Let $P = (p_1, \cdots, p_n)$ be a finite state program (with or without loops) with associated assertion maps Assrn_{p_i}. All assertions of this program hold iff Algorithm 2 returns yes. If the algorithm returns a word σ then at least one assertion fails in the execution of σ.

Program	ProofTraPar	THREADER[10]	Lazy-CSeq[11]
Peterson.safe	**0.3**	3.2	3.1
Dekker.safe	**1.1**	1.7	4.2
Lamport.safe	**2.4**	47	5.1
Szymanksi.safe	**3**	12.8	4
TimeVarMutex.safe	**0.76**	8.56	4.2
RWLock.safe (2R+2W)	8.8	140	**6.7**
RWLock.unsafe (2R+2W)	3.8	153	**0.7**
Qrcu.safe (2R+1W)	**20**	–	41
Qrcu.unsafe (2R+1W)	13.8	76	**1.1**

Fig. 14: Comparison with THREADER[10], and Lazy-CSeq [11] (Time in seconds)

IV. EXPERIMENTAL EVALUATION

We implemented our approach in a prototype tool, `ProofTraPar`. This tool reads the input program written in a custom format. In future, we plan to use off-the-shelf parsers such as CIL or LLVM to remove this dependency. Individual processes are represented using finite state automata. Shuffled language of these automata is computed using standard language theoretic operation. We use an automata library, libFAUDES [5] to carry out operations on automata. As this library does not provide operations on AFA, mainly complementation and intersection, we implemented them in our tool. After constructing the AFA from a trace we first remove ϵ transitions from this AFA. This is followed by adding additional edges in AFA using proposed transformations. Complementation of this AFA involves interchanging existential and universal states with appropriate changes in their transitions. Instead of reversing this AFA (as in Line 11 of Algorithm 2) we intersect it with an NFA that represents the reversed language of the set of all traces. This avoids the need of reversing an AFA. Note that we do not convert our AFA to NFA but rather carry out intersection and complementation operations directly on AFA. Our tool uses uses Z3 [4] theorem prover to check the validity of formulae during AFA construction. `ProofTraPar` can be accessed from the repository https://github.com/chinuhub/ProofTraPar.git.

Figure 14 tabulates the result of verifying *pthread-atomic* category of SV-COMP benchmarks using our tool, THREADER [10] and Lazy-CSeq [11]. These tools were the winners in the concurrency category of the software verification competition of 2013 (THREADER), 2014 and 2015 (Lazy-CSeq). Dash (–) denotes that the tool did not finish the analysis within 15 minutes. Numbers in bold text denote the best time of that experiment. Safe/Unsafe versions of these programs are labeled with *.safe/.unsafe*. Except on Reader-Writer Lock and on unsafe version of QRCU(Quick Read Copy Update), our tool performed better than the other two tools. On unsafe versions, our approach took more time to find out an erroneous trace as compared to Lazy-CSeq [11]. Context-bounded exploration by Lazy-CSeq [11] and the presence of bugs at a shallow depth seems to be a possible reason behind this performance difference. Introducing priorities while picking traces in order to make our approach efficient in bug-finding is left open for future work.

V. RELATED WORK

Verifying the safety properties of a concurrent program is a well studied area. Automated verification tools which use model checking based approaches employ optimizations such as Partial Order Reductions (POR) [13], [8], [7] to handle larger number of interleavings. These optimizations also selectively check a representative set of traces among the set of all interleavings. POR based methods were traditionally used in bug finding but recently they have been extended efficiently, using abstraction and interpolants, for proving programs correct [14]. The technique presented in this paper, using AFA, can possibly be used to keep track of partial orders in POR based methods. In [15], a formalism called concurrent trace program (CTP) is defined to capture a set of interleavings derived from a concurrent trace and safety properties. CTP captures the partial orders encoded in that trace. Corresponding to a CTP, a formula ϕ_{ctp} is generated such that ϕ_{ctp} is satisfiable iff there is a feasible linearization of the partial order encoded in CTP that violates the given property. Our AFA is also constructed from a trace but unlike CTP it only captures those different interleavings which guarantee the same proof outline as of the given trace. Recently in [9], a formalism called *HB-formula* has been proposed to capture the set of happens-before relations in a set of executions. This relation is then used for multiple tasks such as synchronization synthesis, bug summarization and predicate refinement. Since the AFA constructed by our algorithm can also be represented as a boolean formula (universal states correspond to conjunction and existential states correspond to disjunction) that encodes the ordering relations among the participating events, it will be interesting to explore other usages of this AFA along the lines of [9] and [2].

VI. CONCLUSION AND FUTURE WORK

We have presented a trace partitioning based approach for verifying the safety properties of a concurrent program. To this end, we have introduced a novel construction of an alternating finite automaton to capture the proof of correctness of an execution of the program. The AFA constructed in this paper is general enough to represent a proof and a set of executions. We also presented an implementation of our algorithm which compared competitively with existing state-of-the-art tools. We plan to extend this approach for parameterized programs and programs under relaxed memory models. We also plan to investigate the use of interpolants with weakest precondition axioms to incorporate abstraction for handling infinite state programs.

REFERENCES

[1] J.A. Brzozowski and L. L. Ernst. On equations for regular languages, finite automata, and sequential networks. *TCS*, 10:19–35, 1980.

[2] P. Cerný, E.M. Clarke, T. A. Henzinger, A. Radhakrishna, L. Ryzhyk, R. Samanta, and T. Tarrach. From non-preemptive to preemptive scheduling using synchronization synthesis. In *CAV*, pages 180–197, 2015.

[3] A. K. Chandra, D. C. Kozen, and L. J. Stockmeyer. Alternation. *J. ACM*, 28(1):114–133, January 1981.

[4] L. De Moura and N. Bjørner. Z3: An efficient smt solver. In *TACAS*, pages 337–340. Springer-Verlag, 2008.

[5] Bernd Opitz et al. libfaudes-discrete event system library.

[6] A. Farzan, Z. Kincaid, and A. Podelski. Inductive data flow graphs. In *POPL*, pages 129–142, 2013.

[7] C. Flanagan and P. Godefroid. Dynamic partial-order reduction for model checking software. In *POPL*, pages 110–121, 2005.

[8] P. Godefroid. *Partial-Order Methods for the Verification of Concurrent Systems: An Approach to the State-Explosion Problem*. Springer, 1996.

[9] A. Gupta, T. A. Henzinger, A. Radhakrishna, R. Samanta, and T. Tarrach. Succinct representation of concurrent trace sets. In *POPL*, pages 433–444, 2015.

[10] A. Gupta, C. Popeea, and A. Rybalchenko. Threader: A constraint-based verifier for multi-threaded programs. In *CAV*, pages 412–417, 2011.

[11] O. Inverso, E. Tomasco, B. Fischer, S. La Torre, and G. Parlato. Bounded model checking of multi-threaded c programs via lazy sequentialization. In *CAV*, volume 8559 of *LNCS*, pages 585–602. Springer, 2014.

[12] L. Lamport. How to make a multiprocessor computer that correctly executes multiprocess programs. *IEEE Trans. Comput.*, 28(9):690–691, September 1979.

[13] D. Peled. All from one, one for all: On model checking using representatives. In *CAV*, pages 409–423. Springer-Verlag, 1993.

[14] B. Wachter, D. Kroening, and J. Ouaknine. Verifying multi-threaded software with Impact. In *FMCAD*, pages 210–217. IEEE, 2013.

[15] C. Wang, S. Kundu, M. Ganai, and A. Gupta. Symbolic predictive analysis for concurrent programs. In Ana Cavalcanti and DennisR. Dams, editors, *FM 2009: Formal Methods*, volume 5850 of *LNCS*, pages 256–272. Springer Berlin Heidelberg, 2009.

APPENDIX

A. Proof of Lemma 1

We prove it by induction on n.

1) Base case $|\sigma| = 0$: If $|\sigma| = 0$ then $\mathtt{wp}(\sigma[\mathtt{assume/assert}], \neg\phi) = \neg\phi$. If $\neg\phi \wedge \mathcal{I}$ is unsatisfiable then $\mathcal{I}$ satisfies ϕ. Hence proved.

2) Induction step, $|\sigma| = n + 1$: Let $\sigma = \sigma'.a$. If $\mathtt{wp}(\sigma'.a[\mathtt{assume/assert}], \neg\phi) \wedge \mathcal{I}$ is unsatisfiable then following cases can happen based on a.

   - $a : x := E$:- If $\mathtt{wp}(\sigma'.a[\mathtt{assume/assert}], \neg\phi) \wedge \mathcal{I}$ is unsatisfiable then $\mathtt{wp}(\sigma'[\mathtt{assume/assert}], \mathtt{wp}(a, \neg\phi)) \wedge \mathcal{I}$ is also unsatisfiable. By substituting $\mathtt{wp}(a, \neg\phi)$ with $\neg\phi[E/x]$ we get that $\mathtt{wp}(\sigma'[\mathtt{assume/assert}], \neg\phi[E/x]) \wedge \mathcal{I}$ is unsatisfiable. Using IH on σ' it implies that after executing σ' from $\mathcal{I}$ the resultant state either does not terminate or terminates in a state satisfying $\phi[E/x]$. If σ' does not terminate then so does the execution of σ starting from $\mathcal{I}$. If σ' terminates in a state satisfying $\phi[E/x]$ then by the definition of the weakest precondition, execution of a from this state will satisfy ϕ. Hence proved.

   - $a : \mathtt{assume}(\phi')$:-If $\mathtt{wp}(\sigma'.a[\mathtt{assume/assert}], \neg\phi) \wedge \mathcal{I}$ is unsatisfiable then $\mathtt{wp}(\sigma'[\mathtt{assume/assert}], \mathtt{wp}(a, \neg\phi)) \wedge \mathcal{I}$ is also unsatisfiable. By substituting $\mathtt{wp}(a, \neg\phi)$ with $\phi' \wedge \neg\phi$ we get that $\mathtt{wp}(\sigma'[\mathtt{assume/assert}], \phi' \wedge \neg\phi) \wedge \mathcal{I}$ is unsatisfiable. Using IH on σ' it implies that after executing σ' from $\mathcal{I}$ the resultant state either does not terminate or terminates in a state satisfying $\neg\phi \vee \phi'$. If σ' does not terminate then the execution of σ from $\mathcal{I}$ does not terminate as well. If σ' terminates in a state satisfying $\neg\phi$ then the execution of a blocks and hence the execution of σ does not terminate. If σ' terminates in a state satisfying ϕ' but $\neg\phi$ does not hold then $\phi \wedge \phi'$ must hold. Execution of $\mathtt{assume}(\phi')$ acts as $\mathtt{nop}$ instruction and the resultant state satisfies ϕ. hence proved.

   - $a : \mathtt{lock}(x)$:- As weakest precondition of $\mathtt{lock}(x)$ is obtained from the weakest precondition of assignment and assume instruction hence the similar reasoning works for this case.

B. Proof of Lemma 2

Proof Let us prove it by induction on the length of σ.

1) Base case, $|\sigma| = 0$: When the length of σ is 0 and $\mathcal{I} \wedge \neg\phi$ is satisfiable then $\mathcal{I}$ does not satisfy ϕ. Hence proved.

2) Induction Step, $|\sigma| = n + 1$: Let $\sigma = \sigma'.a$. Following case can happen based on the type of a.

   - $a : x := E$:- If $\mathtt{wp}(\sigma[\mathtt{assume/assert}], \neg\phi) \wedge \mathcal{I}$ is satisfiable then $\mathtt{wp}(\sigma'[\mathtt{assume/assert}], \mathtt{wp}(a, \neg\phi)) \wedge \mathcal{I}$ is also satisfiable. By substituting $\mathtt{wp}(a, \neg\phi) = \neg\phi[E/x]$ we get that $\mathtt{wp}(\sigma'[\mathtt{assume/assert}], \neg\phi[E/x]) \wedge \mathcal{I}$ is satisfiable. By IH on σ', execution of σ' from $\mathcal{I}$

terminates in a state not satisfying $\phi[E/x]$. By definition of the weakest precondition, the state reached after executing a from this state does not satisfy ϕ. Hence proved.

- a : $\text{assume}(\phi')$:-If $\text{wp}(\sigma[\text{assume/assert}], \neg\phi) \wedge \mathcal{I}$ is satisfiable then $\text{wp}(\sigma'[\text{assume/assert}], \text{wp}(\text{assume}(\phi')[\text{assume/assert}], \neg\phi)) \wedge \mathcal{I}$ is also satisfiable. By substituting $\text{wp}(\text{assume}(\phi')[\text{assume/assert}], \neg\phi) = \phi' \wedge \neg\phi$ we get that $\text{wp}(\sigma'[\text{assume/assert}], \neg(\neg\phi' \vee \phi)) \wedge \mathcal{I}$ is satisfiable. By IH on σ', execution of σ' from $\mathcal{I}$ terminates in a state not satisfying $\neg\phi' \vee \phi$. In other words, ϕ' and $\neg\phi$ holds in the state reached after executing σ' from $\mathcal{I}$. Therefore, after executing $\text{assume}(\phi')$, the resultant state satisfies $\neg\phi$ and hence proved.

- $a : \text{lock}(x)$:-Similar to the combination of above two cases.

C. Proof of Lemma 3

Proof We use induction for this proof. Let us use the following ordering on the states of $\hat{\mathcal{A}}_\sigma\phi$. For any two states s and s', $s < s'$ if $|\text{RMap}(s)| < |\text{RMap}(s')|$ or if lengths are same then $\text{AMap}(s)$ is a sub formula of $\text{AMap}(s')$. Any two states which are not related by this order, put them in any order to make $<$ as a total order. It is clear that the smallest state in this total order must be one of the accepting state. Now we are ready to proceed by induction using this total order.

- Base case; For every accepting state $s \in S_F$, by Point 3 of Definition 1, the condition $\text{wp}(op, \text{AMap}(s)) = \text{AMap}(s)$ holds for every $op \in \mathcal{EL}(\text{RMap}(s))$. Further, By transition rule LITERAL-SELF-ASSN of this AFA, a self transition must be there for all such $op \in \mathcal{EL}(\text{RMap}(s))$ and hence the condition $\text{rev}(\text{RMap}(s)) \in \text{acc}(s)$ holds (because these transitions can be taken in any order to construct the required word).

- Induction step; Following possibilities exist for the state s,
  - s is a universal state; By construction, there should be states $s_1, \cdots, s_k$ such that $(s, \epsilon, \{s_1, \cdots, s_k\})$ is a transition. By our induction ordering, $s_1, \cdots, s_k$ are smaller than s and hence we apply IH on them to get that $\text{rev}(\text{RMap}(s_i)) \in \text{acc}(s_i)$ for $i \in \{1 \cdots k\}$. However, by the transition rule COMPOUND-ASSN, $\text{RMap}(s) = \text{RMap}(s_1) = \cdots = \text{RMap}(s_k)$ and hence $\text{rev}(\text{RMap}(s) \in \text{acc}(s_i)$ for $i \in \{1 \cdots k\}$. By the definition of $acc(s)$ for a universal state, $\text{acc}(s)$ is intersection of the sets $\text{acc}(s_i)$ for $i \in \{1 \cdots k\}$ and hence we get the required result, viz. $\text{rev}(\text{RMap}(s)) \in \text{acc}(s)$.
  - s is an existential state; If s is an accepting state then Base case holds here. Consider the case when s is not an accepting state. It should have a successor state s' such that $(s, op, \{s'\})$ is a transition. By transition rule LITERAL-ASSN $\text{RMap}(s) = \text{RMap}(s').op.\sigma''$ such that $\text{wp}(\sigma''[\text{assume/assert}], \text{AMap}(s)) = \text{AMap}(s)$. By transition rule LITERAL-

SELF-ASSN, s will have self loop transitions on all symbols in $\sigma''(*)$. Applying IH on s' gives that $\text{rev}(\text{RMap}(s')) \in \text{acc}(s')(\#)$. Because of the transition $(s, op, \{s'\})$, $op.\text{acc}(s') \subseteq \text{acc}(s)$. This along with (#) gives us $op.\text{rev}(\text{RMap}(s')) \in \text{acc}(s)(**)$. Rearranging this and using (*) we get $\text{rev}(\text{RMap}(s')).op.\sigma'') \in \text{acc}(s)$ or equivalently $\text{rev}(\text{RMap}(s)) \in \text{acc}(s)$. Hence proved.

D. Proof of Lemma 4

Proof We use induction for this proof. Same as in the previous proof, let us use the following ordering on the states of $\hat{\mathcal{A}}$. For any two states s and s', $s < s'$ if $|\text{RMap}(s)| < |\text{RMap}(s')|$ or if lengths are same then $\text{AMap}(s)$ is a sub formula of $\text{AMap}(s')$. Any two states which are not related by this order, put them in any order to make $<$ as a total order. It is clear that the smallest state in this total order must be one of the accepting state. Now we are ready to proceed by induction using this total order.

- Base case, By definition of the accepting state in AFA construction, Point 3 of Definition 1, and the self loop transition rule, Rule LITERAL-SELF-ASSN, we know that for every word $\sigma' \in \text{acc}(s)$, $\text{wp}(\sigma'[\text{assume/assert}], \text{AMap}(s)) = \text{AMap}(s)$. Rule BASE-CASE of Figure 8 sets $\text{HMap}(s)$ same as $\text{AMap}(s)$ for such states hence the statement of this lemma follows for the accepting states.

- Induction step; we pick a state s such that one of the following holds,
  1) s is a universal state;By construction, there should be states $s_1, \cdots, s_k$ such that $(s, \epsilon, \{s_1, \cdots, s_k\})$ is a transition. Let σ be a word accepted by s then by the definition of accepting set of words of a universal states, σ must be accepted by each of $s_1, \cdots s_k$. By our induction ordering, $s_1, \cdots, s_k$ are smaller than s and hence we apply IH on them to get that $\text{wp}(\text{rev}(\sigma)[\text{assume/assert}], \text{AMap}(s_i)) = \text{HMap}(s_i)$ for $i \in \{1 \cdots k\}$. Two cases arise based on whether
     - $\text{AMap}(s)$ is a conjunction of $\text{AMap}(s_i)$ for $i \in \{1 \cdots k\}$; Following Rule CONJ-CASE we set $\text{HMap}(s) = \bigwedge_i \text{HMap}(s_i)$ and $\text{wp}(\text{rev}(\sigma)[\text{assume/assert}], \text{AMap}(s)) = \text{HMap}(s)$ then follows from the Property 1, using conjunction, of the weakest precondition.
     - $\text{AMap}(s)$ is a disjunction of $\text{AMap}(s_i)$ for $i \in \{1 \cdots k\}$; Following Rule CONJ-CASE we set $\text{HMap}(s) = \bigwedge_i \text{HMap}(s_i)$ and $\text{wp}(\text{rev}(\sigma)[\text{assume/assert}], \text{AMap}(s)) = \text{HMap}(s)$ then follows from the Property 1, using disjunction, of the weakest precondition.
  2) s is an existential state; If s is an accepting state then the same argument as used in the Base case holds. If s is not an accepting state then the only outgoing transition from s is of the form $(s, op, \{s'\})$, By rule LITERAL-ASSN(*). Now consider a word $\sigma \in \text{acc}(s)$. σ must be of the form $\sigma''.op.\sigma'$ where $\text{wp}(\sigma''[\text{assume/assert}], \text{AMap}(s)) = \text{AMap}(s)(*)$

(because of the self transitions constructed from Rule LITERAL-SELF-ASSN) and $\sigma' \in \mathrm{acc}(s')$. Therefore, $\mathrm{wp}(\mathrm{rev}(\sigma)[\mathrm{assume/assert}], \mathrm{AMap}(s)) =$
$=\mathrm{wp}(\mathrm{rev}(\sigma''.op.\sigma')[\mathrm{assume/assert}], \mathrm{AMap}(s))$
$=\mathrm{wp}(\mathrm{rev}(\sigma').op.\mathrm{rev}(\sigma'')[\mathrm{assume/assert}], \mathrm{AMap}(s))$
$=\mathrm{wp}(\mathrm{rev}(\sigma').op[\mathrm{assume/assert}], \mathrm{AMap}(s))$ (using (*))
$=\mathrm{wp}(\mathrm{rev}(\sigma')[\mathrm{assume/assert}],$ $\mathrm{wp}(op[\mathrm{assume/assert}], \mathrm{AMap}(s)))$ (using weakest precondition definition)
$=\mathrm{wp}(\mathrm{rev}(\sigma')[\mathrm{assume/assert}], \mathrm{AMap}(s'))$ (using Transition rule LITERAL-ASSN)

As $\sigma' \in \mathrm{acc}(s')$ this is same as $\mathrm{HMap}(s')$ by applying IH on s'. As $\mathrm{HMap}(s)$ is same as $\mathrm{HMap}(s')$, as done in Rule LIT-CASE, we prove this case as well.

E. Proof of Correctness of Optimization-I

Lemma 5: Let $\hat{\mathcal{A}}$ be an automaton constructed from a trace and a post condition as defined in Definition 1 and further modified by Algorithm 1 then for every state s of this AFA and for every word σ accepted by state s, $\mathrm{HMap}(s)$ is logically equivalent to $\mathrm{wp}(\mathrm{rev}(\sigma)[\mathrm{assume/assert}], \mathrm{AMap}(s))$.

Proof Proof of this lemma is very similar to the proof of Lemma 4 given in Appendix . Here we only highlight the changes in the proof. Note that this optimization converts some universal states to existential states. Let s be one such state that was converted from universal to existential state. Let $(s, \epsilon, \{s_1, \cdots, s_k\})$ was the original transition in the AFA which got modified to $(s, \epsilon, \{s_{u_1}, \cdots, s_{u_n}\}$ where s_{u_i} are newly created universal states in Line 5 of Algorithm 1. By construction, $\mathrm{HMap}(s_{u_i})$ is unsatisfiable for each of these $s_{u_1}, \cdots, s_{u_n}$(*). Let σ be a word accepted by s after converting it to existential state. By acceptance conditions, σ must be accepted by at least one state, say s_{u_m} in the set $\{s_{u_1}, \cdots, s_{u_n}\}$. By IH on s_{u_m} we get $\mathrm{wp}(\sigma[\mathrm{assume/assert}], \mathrm{AMap}(s_{u_m})) = \mathrm{HMap}(s_{u_m})$(**). Further, by construction $\mathrm{AMap}(s)$ implies $\mathrm{AMap}(s_{u_m})$. This fact, along with the monotonicity property of the weakest precondition, Property 2, we get that $\mathrm{wp}(\sigma[\mathrm{assume/assert}], \mathrm{AMap}(s))$ is unsatisfiable and hence same as $\mathrm{HMap}(s)$.

F. Proof of Correctness of Optimization-II

Lemma 6: Let $\hat{\mathcal{A}}$ be an automaton constructed from a trace and a post condition as defined in Definition 1 and further modified by adding edges as discussed above then for every state s of this AFA and for every word σ accepted by state s, $\mathrm{HMap}(s)$ is logically equivalent to $\mathrm{wp}(\mathrm{rev}(\sigma)[\mathrm{assume/assert}], \mathrm{AMap}(s))$.

Proof As a result of adding edges in this optimization, we can not use the ordering among states as done for earlier proofs. This is because, now a transition (s, op, S) does not guarantee that the states in the set S are smaller then s and hence it will not be possible to apply IH directly. Therefore in this proof we apply induction on the length of σ' accepted by some state s.

- Induction step; Let $s \in \hat{\mathcal{A}}$ and $\sigma \in \mathrm{acc}(s)$ such that $|\sigma| = m + 1$. Either $s \in S_\exists$ or $s \in S_\forall$. If $s \in S_\exists$ and $\sigma \in \mathrm{acc}(s)$ then there exists a state s' such that $(s, op, \{s'\}) \in \delta$ and $\sigma' \in \mathrm{acc}(s')$, where $\sigma = \sigma''.op.\sigma'$ and $\mathrm{wp}(\sigma''[\mathrm{assume/assert}], \mathrm{AMap}(s)) = \mathrm{AMap}(s)$(**). Based on this transition $(s, op, \{s'\}) \in \delta$ we have the following sub-cases,
  - $(s, op, \{s'\})$ was added by the this optimization by virtue of one of the following conditions,
    * $\mathrm{HMap}(s)$ and $\mathrm{HMap}(s')$ are unsatisfiable and $\mathrm{wp}(op[\mathrm{assume/assert}], \mathrm{AMap}(s)) \Rightarrow \mathrm{AMap}(s')$ (Rule RULE-UNSAT); By IH on σ' we have $\mathrm{wp}(\mathrm{rev}(\sigma')[\mathrm{assume/assert}], \mathrm{AMap}(s'))$ is logically equivalent to $\mathrm{HMap}(s')$. Using Property 2 (conjunction part) and the assumption $\mathrm{wp}(op[\mathrm{assume/assert}], \mathrm{AMap}(s)) \Rightarrow \mathrm{AMap}(s')$ we get $\mathrm{wp}(\mathrm{rev}(\sigma')[\mathrm{assume/assert}], \mathrm{wp}(op, \mathrm{AMap}(s)))$ is unsatisfiable and same as $\mathrm{HMap}(s)$. Using (**), $\mathrm{wp}(\mathrm{rev}(\sigma')[\mathrm{assume/assert}], \mathrm{wp}(op.\mathrm{rev}(\sigma''), \mathrm{AMap}(s)))$ is unsatisfiable and same as $\mathrm{HMap}(s)$. By replacing $\sigma = \sigma''.op.\sigma'$ we get the required proof.
    * $\mathrm{HMap}(s)$ and $\mathrm{HMap}(s')$ are valid and $\mathrm{AMap}(s') \Rightarrow \mathrm{wp}(op[\mathrm{assume/assert}], \mathrm{AMap}(s))$ (Rule RULE-VALID); By IH on σ' we have $\mathrm{wp}(\mathrm{rev}(\sigma')[\mathrm{assume/assert}], \mathrm{AMap}(s'))$ is logically equivalent to $\mathrm{HMap}(s')$. Using property 2 (disjunction part) and the assumption $\mathrm{AMap}(s') \Rightarrow \mathrm{wp}(op[\mathrm{assume/assert}], \mathrm{AMap}(s))$ we get $\mathrm{wp}(\mathrm{rev}[\mathrm{assume/assert}](\sigma'), \mathrm{wp}(op[\mathrm{assume/assert}], \mathrm{AMap}(s)))$ is valid and same as $\mathrm{HMap}(s)$. Using (**) and or replacing $\sigma = \sigma''.op.\sigma'$ we get the required result and hence proved.
  - If this transition was already in δ; we can use the same reasoning as used in the proof of Lemma 4 to show that $\mathrm{wp}(\mathrm{rev}(\sigma)[\mathrm{assume/assert}], \mathrm{AMap}(s))$ is logically equivalent to $\mathrm{HMap}(s)$
- If $s \in S_\forall$ then similar argument goes as in the proof of Lemma 4 because no new transition gets added from these states as a result of this optimization.

G. Proof of Theorem 1

Proof
- Let us first prove that this algorithm terminates for finite state programs. For finite state programs the number of possible assertions used in the construction of AFA are finite and hence only a finite number of different AFA are possible. It implies the termination of this algorithm.
- Following Lemma 4 and the fact that $\mathrm{AMap}(s_0) = \neg\phi$, every word σ' accepted by this AFA, equivalently written as $\sigma' \in \mathrm{acc}(s_0)$, satisfies $\mathrm{wp}(\mathrm{rev}(\sigma')[\mathrm{assume/assert}], \neg\phi) = \mathrm{HMap}(s_0)$(*). By Lemma 3 and the fact that $\mathrm{RMap}(s_0) = \sigma$ we get $\mathrm{rev}(\sigma) \in \mathrm{acc}(s_0)$(**). Combining (**) and (*), we get $\mathrm{wp}(\mathrm{rev}(\mathrm{rev}(\sigma))[\mathrm{assume/assert}],$

$\neg\phi) = \text{HMap}(s_0)$ or equivalently $\text{wp}(\sigma[\texttt{assume/assert}], \neg\phi) = \text{HMap}(s_0)$.

– If $\mathcal{I} \wedge \text{HMap}(s_0)$ is satisfiable (Line 6) then $\mathcal{I} \wedge \text{wp}(\sigma[\texttt{assume/asssert}], \neg\phi)$ is satisfiable as well. Following Lemma 2 we got a valid error trace which is returned in Line 8.

– If $\mathcal{I} \wedge \text{HMap}(s_0)$ is unsatisfiable then by Lemma 1 this trace is provably correct. Now we apply optimizations to the AFA to increase the set of words accepted by it. The final AFA is then reversed and subtracted from the set of executions seen so far. Lemma 4 ensures that for all such words σ' the condition $\mathcal{I} \nRightarrow \text{wp}(\sigma', \neg\phi)$ holds and therefore none of them violate ϕ starting from the initial state. Therefore in every iteration only correct set of executions are being removed from the set of all executions. Therefore when this loop terminates then all the executions have been proved as correct.