M-N Without Permutations

Ian Dardik

March 31, 2022

1 Introduction

In the past few years there has been ample research into inductive invariant synthesis for parameterized distributed protocols. A key desirable feature for an invariant synthesis tool is the ability to check whether the algorithm terminates with a correct inductive invariant. For tools that remain within the bounds of a decidable logic fragment, this check is feasible. However, any tool that produces a candidate inductive invariant for a system that falls outside of a decidable fragment offers no guarantee that the candidate is indeed correct. In this note, we assume that a candidate inductive invariant is *given* and we exclusively focus on the verification step.

We have discovered a syntactic class of protocols which exhibit a *cutoff* for the number of finite protocol instances which need to be verified. We have captured this result in the M-N Theorem.

In this note we begin by introducing the Sort-Restricted to PNF Language (SRPL), the logic language that we use to encode our class of protocols. We then introduce our encoding of protocols as a transition system in SRPL. Next, we will prove some key lemmas before finally presenting and proving the M-N Theorem.

2 Sort-Restricted to PNF Language

In this section we will define SRPL as the composition of two grammars. SRPL formulas are parameterized by a single sort E of indistinguishable elements. We assume that E is countably infinite.

Definition 1. Let $\mathcal{D}$ be a countable set of domain symbols, $\mathcal{P}$ be a countable set of predicates, $\mathcal{A}$ be a countable set of constants, and $\mathcal{V}$ be a countable set of variables. A parameterized *svf* term is produced by the following grammar:

$$
\begin{array}{lll}
arg(x_1, ..., x_n) & ::= y & \text{for any } y \in \mathcal{V} \\
arg(x_1, ..., x_n) & ::= a & \text{for any } a \in \mathcal{A} \\
arg(x_1, ..., x_n) & ::= x_i & \text{for any } 1 \leq i \leq n \\
arg\_list(x_1, ..., x_n) & ::= arg(x_1, ..., x_n) & \\
arg\_list(x_1, ..., x_n) & ::= arg(x_1, ..., x_n), arg\_list(x_1, ..., x_n) & \\
svf(x_1, ..., x_n) & ::= p(arg\_list(x_1, ..., x_n)) & \text{for any } p \in \mathcal{P} \\
svf(x_1, ..., x_n) & ::= \neg svf(x_1, ..., x_n) & \\
svf(x_1, ..., x_n) & ::= svf(x_1, ..., x_n) \wedge svf(x_1, ..., x_n) & \\
Q & ::= \forall \mid \exists & \\
svf(x_1, ..., x_n) & ::= Q\, x \in D(arg\_list(x_1, ..., x_n)),\ svf(x_1, ..., x_n) & \text{for any } x \in \mathcal{V}, D \in \mathcal{D}
\end{array}
$$

It is important to note that no *svf* term can refer to the sort E directly, and hence cannot quantify over E nor take its cardinality.

Definition 2. Let $\mathcal{V}$ be a countable set of variables. A formula in SRPL is defined by the grammar for the production rule of *srpl*:

arg	$::= x$	for any $x \in \mathcal{V}$
arg_list	$::= arg$	
arg_list	$::= arg, arg\_list$	
Q	$::= \forall \mid \exists$	
$srpl$	$::= Q\, x \in E,\ svf(arg\_list)$	for any $x \in \mathcal{V}$
$srpl$	$::= Q\, x \in E,\ srpl$	for any $x \in \mathcal{V}$

SRPL is quite rich and we should provide some examples to show this.

Definition 3 (Instance)**.** Let ψ be an SRPL formula and let $H \subseteq E$. Then we define $\psi(E \mapsto H)$ by the following rules on the SRPL grammar:

$x(E \mapsto H)$	$:= x$	for any $x \in \mathcal{V}$
$arg(E \mapsto H)$	$:= arg$	
$[arg, arg\_list](E \mapsto H)$	$:= arg, arg\_list$	
$[Q\, x \in E,\ svf(arg\_list)](E \mapsto H)$	$:= Q\, x \in H,\ svf(arg\_list)$	for any $x \in \mathcal{V}$
$[Q\, x \in E,\ srpl](E \mapsto H)$	$:= Q\, x \in H,\ [srpl(E \mapsto H)]$	for any $x \in \mathcal{V}$

In other words, $\psi(E \mapsto H)$ is the formula ψ with E replaced with H. We call $\psi(E \mapsto H)$ an *instance* of ψ, and when H is finite, we call $\psi(E \mapsto H)$ a *finite instance* of ψ.

Definition 4 (Finite Instance Notation)**.** We use a special shorthand for finite instaces that mirrors the notation described in [1]. Let ψ be an SRPL formula and $k > 0$ be given. Then $\psi(k) := \psi(E \mapsto \{e_1, ..., e_k\})$ where each $e_i \in E$ is arbitrary and distinct. We can also write $E(k) := \{e_1, ..., e_k\}$ where each $e_i \in E$ is arbitrary and distinct.

Definition 5 (Valid SPRL Formula)**.** Let ψ be an SRPL formula. Then ψ is valid iff $\psi(E \mapsto H)$ is valid for every $H \subseteq E$.

Lemma 1. Let ψ be an SRPL formula. Then ψ is valid iff $\psi(k)$ is valid for all $k > 0$.

3 *E*-Ground Formulas

Definition 6. Let S be a sort. Then a *ground formula* is generated by the following grammar:

$$\begin{aligned} argument &::= e \text{ for any } e \in S \\ argument\_list &::= argument \\ argument\_list &::= argument, argument\_list \\ ground\_formula &::= p(argument\_list) \text{ for any } n\text{-ary } p \in \mathbf{P}, n \geq 0 \end{aligned}$$

Definition 7. Let F be an RSL formula and $\rho : \mathbf{V} \to E$ be a function. Then we define Replace(F, ρ) recursively

$$\begin{aligned} \text{Replace}(x, \rho) &:= \rho(x) \text{ for any } x \in \mathbf{V} \\ \text{Replace}((argument, argument\_list), \rho) &:= \text{Replace}(argument, rho), \text{Replace}(argument\_list, \rho) \\ \text{Replace}(p(argument\_list), \rho) &:= p(\text{Replace}(argument\_list, \rho)) \text{ for any } n\text{-ary } p \in \mathbf{R}, n \geq 0 \end{aligned}$$

Definition 8 (Ground Instance of *F(k)*)**.** Let F be a quantified PNF formula and $k \in \mathbb{N}$. Then g is a *ground instance* of $F(k)$ iff there exists a mapping $\rho : \mathbf{V} \to E(k)$ and an unquantified formula f such that:

$$g = \text{Replace}(f, \rho) \text{ and } F \in \text{PQF}(f)$$

In other words, g is a ground formula that is identical in structure to F without quantifiers, and with all variables of $F(k)$ replaced by members of $E(k)$.

Example 1. Consider the transition system T with two state variables, $x \in (P \to \mathbb{N})$ and $y \in \mathbb{Z}$. Let $s := (x[1] = 6 \wedge x[2] = 0 \wedge y = -22)$ be a state in the transition system. Let $F := \forall p, q \in P, x[p] \neq x[q]$ and $f := (x[1] \neq x[2])$.

Then $F(2) = \forall p, q \in P(2), x[p] \neq x[q]$. Furthermore, f is a ground instance of $F(2)$, $F(2) \to f$, $s \to F(2)$, and $s \to f$.

Definition 9 (Gr)**.** Let F be a quantified formula and $k \in \mathbb{N}$. Then:

$$\text{Gr}(F, k) := \{f \mid f \text{ is a ground instance of } F(k)\}$$

Example 2. $\text{Gr}([\forall p, q \in P, p = q], 2) = \{(1 = 1), (1 = 2), (2 = 1), (2 = 2)\}$
Note: we sometimes use square braces to wrap formulas when it looks better than parentheses.
Notice that $\text{Gr}([\forall p, q \in P, p = q], 2)$ contains elements that are false. This indicates that the statement $[\forall p, q \in P, p = q](2)$ is not valid.

Example 3. Let sv be a state variable, then:

$$\text{Gr}((\forall p, q \in P, p \neq q \to \text{sv[p]} \neq \text{sv[q]}), 3) = \{(1 \neq 1 \to \text{sv}[1] \neq \text{sv}[1]), (1 \neq 2 \to \text{sv}[1] \neq \text{sv}[2]), ...\}$$

Definition 10 (Elems)**.** Suppose that F is a quantified formula, $k \in \mathbb{N}$, and $f \in \text{Gr}(F, k)$. Then:

$$\text{Elems}(f) := \{e \mid e \in P(k) \wedge e \text{ occurs in } f\}$$

TODO make this definition better.

4 Transition System

We encode a protocol as a transition system $T = (I, \Delta)$ where I is the initial constraint restricted to universal quantification over E and Δ is the transition relation restricted to existential quantification over E, and both are encoded in SRPL. We assume that an inductive invariant candidate Φ is given in SRPL, and is restricted to universal quantification over E. We use the notation $T(k) := (I(k), \Delta(k))$ where $k > 0$.

Definition 11 (States)**.** Let $k > 0$ be given, then:

$$\text{States}(k) := \{s \mid s \text{ is a state of } T(k)\}$$

In this note we consider a "state" $s \in \text{States}(k)$ to be a ground formula. More specifically–under a given interpretation for T–s is a conjunction of constraints that describe a single state in $T(k)$.

Definition 12 (Inductive Invariant)**.** Φ is an inductive invariant iff $\Phi \wedge \Delta \to \Phi'$ is valid.

5 Lemmas

Lemma 2. Let $k \in \mathbb{N}$ such that $s \in \mathrm{States}(k)$ and F is a universally quantified formula. Then:

$$(s \to F(k)) \leftrightarrow (\forall f \in \mathrm{Gr}(F, k), s \to f)$$

Proof. Suppose that $s \to F(k)$. For an arbitrary formula $f \in \mathrm{Gr}(F, k)$, $F(k) \to f$ and hence we see that $s \to F(k) \wedge F(k) \to f$. It follows that $s \to f$.

Now suppose that $\forall f \in \mathrm{Gr}(F, k), s \to f$. Suppose, for the sake of contradiction, that $\neg(s \to F(k))$. Then it must be the case that $s \wedge \neg F(k)$. We know that F is unversally quantified, so let $F(k) := \forall x_1, ..., x_m \in P, \phi(x_1, ..., x_m)$ where $m \geq 1$. Then, because $\neg F(k)$ holds, it must be the case that $\exists x_1, ..., x_m \in P, \neg\phi(x_1, ..., x_m)$. However, $\phi(x_1, ..., x_m) \in \mathrm{Gr}(F, k)$ which, by our original assumption, implies $\neg s$. Hence we have both s and $\neg s$ and we have reached a contradiction. □

6 The M-N Theorem

In this section, we will establish initiation and consecution in two separate lemmas using similar techniques. The M-N Theorem follows immediately from these two lemmas.

Lemma 3 (M-N Initiation)**.** Let m be the number of variables that I quantifies over. Then if $I(m) \to \Phi(m)$ is valid, $I(k) \to \Phi(k)$ is also valid for all $k > m$.

Lemma 4 (M-N Consecution)**.** Let m be the number of variables that Φ quantifies over and n be the number of variables that Δ quantifies over. Then if $\Phi(m+n)$ is inductive, $\Phi(k)$ is also inductive for any $k > m + n$.

Proof. Assume that $[\Phi \wedge \Delta \to \Phi'](m+n)$ is valid. Let $k > m + n$ be given, we want to show that $[\Phi \wedge \Delta \to \Phi'](k)$ is also valid. Let $H = \{e_1, ..., e_k\} \subseteq E$ be an arbitrary finite instance of E. Let $s \in \mathrm{States}(H)$ such that $s \to \Phi(E \mapsto H)$ and let $\delta \in \mathrm{Gr}(\Delta, H)$ such that $\delta \to \Delta(H)$. Then $(s \wedge \delta)$ is a formula that describes the states reachable from s in one "δ step", and it suffices to show that $(s \wedge \delta) \to \Phi'(H)$. Furthermore, let $\phi' \in \mathrm{Gr}(\Phi', H)$ be arbitrary, then, by Lemma 2 and the fact that Φ' is in PNF and universally quantified, it suffices to show that $(s \wedge \delta) \to \phi'$.

Let $\alpha_1, ..., \alpha_i$ be the unique elements of $\{e_1, ..., e_k\}$ in $(\phi \wedge \delta)$, then we know that $i \leq m + n$ because $\phi \in \mathrm{Gr}(\Phi, H)$ where Φ quantifies over m variables and $\delta \in \mathrm{Gr}(\Delta, H)$ where Δ quantifies over n variables. Let $j = m + n - i$, then we can choose $\beta_1, ..., \beta_j$ such that $\{\beta_1, ..., \beta_j\} \subseteq (\{e_1, ..., e_k\} - \{\alpha_1, ..., \alpha_i\})$. Notice that $|\{\alpha_1, ..., \alpha_i, \beta_1, ..., \beta_j\}| = m + n$, and hence, by our initial assumption:

$$[\Phi \wedge \Delta \to \Phi'](E \mapsto \{\alpha_1, ..., \alpha_i, \beta_1, ..., \beta_j\})$$

must be a valid formula.

Now, $s \to \Phi(E \mapsto \{\alpha_1, ..., \alpha_i, \beta_1, ..., \beta_j\})$ because Φ is in PNF and universally quantified (need lemma). Furthermore, $\delta \to \Delta(E \mapsto \{\alpha_1, ..., \alpha_i, \beta_1, ..., \beta_j\})$ because Δ is in PNF and restricted to existential quantification (need lemma). Thus we see:

$$(s \wedge \delta) \to [\Phi \wedge \Delta](E \mapsto \{\alpha_1, ..., \alpha_i, \beta_1, ..., \beta_j\}) \to \Phi'(E \mapsto \{\alpha_1, ..., \alpha_i, \beta_1, ..., \beta_j\}) \to \phi'$$

□

Next we present the M-N Theorem:

Theorem 1 (M-N)**.** Suppose that Φ is in PNF with only universal quantifiers, while Δ is in PNF with only existential quantifiers. Let m be the number of variables that Φ quantifies over and n be the number of variables that Δ quantifies over. If $\Phi(m+n)$ is an inductive invariant, then $\Phi(k)$ is also an inductive invariant for any $k > m + n$.

Proof. This follows immediately from the previous two lemmas. □

References

[1] Aman Goel and Karem Sakallah. On Symmetry and Quantification: A New Approach to Verify Distributed Protocols. In *NASA Formal Methods Symposium*, pages 131–150. Springer, 2021.