A Cutoff Rule For A Special Class Of Parameterized Distributed Protocols

Ian Dardik

March 25, 2022

1 Introduction

In this note, we consider the verification problem of a transition system $T = (I, \Delta)$ where I is the initial constraint, Δ is the transition relation, and the system is parameterized by a single sort P of indistinguishable elements (We make the notion of "indistinguishable" precise in Assumption 1 below).

To begin, we will introduce notation for the template and finite instances of a transition system. We adopt the convention of [3] where $T(P)$ is the template of T and $T(|P|)$ is a finite instance. We can also refer to the template or a finite instance of a quantified formula F and the sort P. For example, suppose F is in Prenex Normal Form (PNF) and univerally quantifies over j variables, i.e. F can be written as:

$$F := \forall x_1, ..., x_j \in P, \phi(x_1, ..., x_j)$$

where ϕ is a non-quantified statement whose only free variables are $x_1, ..., x_j$. Then $F(k)$ is identical to the formula F, except P is replaced by $P(k) \subseteq P$, where $|P(k)| = k$. Without loss of generality–because we assume each element of P is indistinguishable–we will assume that $P(k) = \{1, ..., k\}$. Thus we see:

$$F(k) = \forall x_1, ..., x_j \in P(k), \phi(x_1, ..., x_j)$$

In this note, we are concerned with the specific scenario in which we are given a candidate inductive invariant Φ, and the finite instances $\Phi(1), ..., \Phi(k)$ have been proved to be inductive invariants for $T(1), ..., T(k)$; we want to know whether Φ is an inductive invariant for T. We are specifically concerned with the case in which Φ is restricted to PNF with only universal quantifiers, and Δ is restricted to PNF with only existential quantifiers.

Throughout this note, we will build several lemmas that lead to an interesting result: let m be the number of variables that Φ quantifies over and n be the number of variables that Δ quantifies over; if we suppose that $\Phi(m + n)$ is an inductive invariant for $T(m + n)$, then $\Phi(k)$ is also an inductive invariant for $T(k)$ for all $k > m + n$. We will refer to this as the M-N Theorem in this note. This result is useful because it reduces the verification problem on T to model checking a finite number of instances $T(1), ..., T(m + n)$. Essentially, $m + n$ is a cutoff instance size for proving that our inductive invariant holds.

Note: I think it is likely that if $\Phi(m + n)$ is an inductive invariant, then it is *also* the case for $\Phi(k)$ for all $k < m + n$, but I left this out of this note for the time being to focus on the $k > m + n$ case.

2 Preliminaries

In this section we introduce several assumptions, definitions, and notation that we will use to prove the M-N Theorem.

2.1 Inductive Invariant Of T

Definition 1. A formula F is an inductive invariant for T iff $\forall k \in \mathbb{N}$, $F(k)$ is an inductive invariant for $T(k)$.

2.2 Sorted Logic

In this section we introduce Sorted Logic, including its syntax, based on the grammar for FOL defined in [2]. We do not explicitly include semantics since they are clear.

Definition 2. Let $\mathbf{R}$ be a countable set of predicate symbols. Then an *unquantified formula* is generated by the following grammar:

$$\begin{aligned}
argument &::= x \text{ for any } x \in \mathbf{V} \\
argument_list &::= argument \\
argument_list &::= argument, argument_list \\
atomic_formula &::= p(argument_list) \text{ for any } n\text{-ary } p \in \mathbf{R}, n \geq 1 \\
formula &::= atomic_formula \\
formula &::= \neg formula \\
formula &::= formula \wedge formula
\end{aligned}$$

Where predicates are n-ary functions that return boolean values.

Definition 3. Let $\mathbf{V}$ be a countable set of variables and let the formula $< \mathit{param} >$ be given. Then a *parameterized quantified formula* is generated by the following grammar:

$$\begin{aligned}
par_formula &::= \forall x \in D \;\; < \mathit{param} > \;\text{ for any } x \in \mathbf{V} \\
par_formula &::= \exists x \in D \;\; < \mathit{param} > \;\text{ for any } x \in \mathbf{V} \\
par_formula &::= \forall x \in D \; par_formula \text{ for any } x \in \mathbf{V} \\
par_formula &::= \exists x \in D \; par_formula \text{ for any } x \in \mathbf{V}
\end{aligned}$$

Where D is the domain given in an interpretation.

We use the notation $par\_formula(< \mathit{param} >)$ to denote the set of parameterized quantified formulas generated by $< \mathit{param} >$. If the parameter $< \mathit{param} >$ is an unquantified formla, then $par\_formula(< \mathit{param} >)$ will generate quantified formulas in Prenex Normal Form (PNF).

Example 1. Let $f = p(x, y) \wedge q(z)$ and $F = \forall x, y, z \in D, p(x, y) \wedge q(z)$. Then f is an unquantified formula, F is a quantified PNF formula, and $F \in par\_formula(f)$.

TODO give an example for an interpretation.

Definition 4. Let F be a formula where $\{p_1, ..., p_m\}$ are all the predicates appearing in F. An *interpretation* $\mathbf{I}_A$ is the pair:

$$(D, \{r_1, ..., r_m\})$$

Where D is a non-empty domain and each r_i is a relation. We refer to a relation very loosely; in this context, a relation may be a quantified formula. However, we require that each r_i cannot refer to D, e.g. no relation can quantify over D or refer to its cardinality.

The formulas I, Δ, and Φ are written in Sorted Logic and must be closed (not have any free variables). Furthermore, we will implicitly assume that any formula is a closed formula for the remainder of this note, unless explicitly mentioned. TODO comment about the fact that we will assume an arbitrary interpretation for all formulas (i.e. formulas *are* interpreted).

2.3 States And Ground Formulas

Definition 5 (States)**.** Let $k \in \mathbb{N}$, 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 formula. More specifically, s is a non-quantified conjunction of constraints that describe a single state in $T(k)$.

Definition 6 (Satisfaction)**.** Let f and g be formulas in First Order Logic (FOL). Then we write $f \models g$ iff $f \to g$. Alternatively, f satisfies g iff f is stronger than g.

Definition 7 (Ground Formula)**.** A *ground formula* is an unquantified formula whose variables are replaced by members of P. In other words, a ground formula has no free variables, but may contain members of P.

Definition 8 (Ground Formula *of F(k)*)**.** Let F be a quantified PNF formula and $k \in \mathbb{N}$. We say that g is a ground formula *of* $F(k)$ iff there exists a formula f such that $g = f[\mathbf{V} \mapsto P(k)]$ and $F(k) \in \mathit{par\_formula}(f)$.

In otherwords, f is a ground formula that is identical in structure to F without quantifiers, and with all variables of $F(k)$ replaced by members of $P(k)$.

Example 2. 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 formula of $F(2)$, $F(2) \models f$, $s \models F(2)$, and $s \models f$.

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

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

Example 3. $\text{Gr}([\forall p, q \in P, p = q], 2) = \{(1 = 1), (1 = 2), (2 = 1), (2 = 2)\}$
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 4. 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.

2.4 Permutation Transformations

Definition 11 (Permutation Transformation)**.** Let $k \in \mathbb{N}$, $\pi : P(k) \to P(k)$ be a permutation on $P(k)$, and G be the set of all possible formulas. Then $M_\pi : G \to G$ is the *permutation transformation* on π, a syntactic transformation that replaces each element from $P(k)$ in a formula with its permuted value.

Example 5. Let π be the following permutation:

$$\pi := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

Let sv be a state variable, then:

$$M_\pi(3 \neq 1 \to \text{sv}[3] \neq \text{sv}[1]) = (1 \neq 2 \to \text{sv}[1] \neq \text{sv}[2])$$

2.5 Indistinguishable Elements

We have loosely stipulated that T must have "indistinguishable" elements. In this section, we make this assumption precise.

Assumption 1 (P Has Indistinguishable Elements)**.** Let $j, k \in \mathbb{N}$ such that $j \geq k$ and F be a quantified sentence in PNF. Let $s \in \text{States}(j)$ such that $s \models F(k)$. If π is a permutation then it is also the case that $M_\pi(s) \models F(k)$.

3 Helper Lemmas

Lemma 1. Let $j, k \in \mathbb{N}$ such that $j \geq k$, $s \in \text{States}(j)$, and F be a universally quantified formula. Then:

$$(s \models F(k)) \leftrightarrow (\forall f \in \text{Gr}(F, k), s \models f)$$

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

Now suppose that $\forall f \in \text{Gr}(F, k), s \models f$. Suppose, for the sake of contradiction, that $s \not\models F(k)$. 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, $\phi(x_1, ..., x_m) \in \text{Gr}(F, k)$ which, by our original assumption, implies $\neg s$. Hence we have both s and $\neg s$ and we have reached a contradiction. $\square$

Lemma 2 (Gr Members Sat)**.** Let F be a formula and $k \in \mathbb{N}$ be given. Then:

$$\forall g \in \text{Gr}(F, k), F(k) \models g$$

Proof. Suppose that $g \in \text{Gr}(F, k)$, then there exists a formula f such that $g = f[\mathbf{V} \mapsto P(k)]$ and $F \in par\_formula(f)$. We will prove the claim by induction on the number of free variables that f has.

Base case: If f has just one free variable, then either $F(k) = \forall x \in P(k), f(x)$ or $F(k) = \exists x \in P(k), f(x)$. In both cases:

$$F(k) \models f[\mathbf{V} \mapsto P(k)] = g$$

Now suppose that the claim holds for $1, ..., i$ free variables. If f has $i+1$ free variables, then either $F(k) = \forall x_{i+1} \in P(k), Q_i x_i \in P(k), ..., Q_1 x_1 \in P(k), f(x_{i+1}, x_i, ..., x_1)$ or $F(k) = \exists x_{i+1} \in P(k), Q_i x_i \in P(k), ..., Q_1 x_1 \in P(k), f(x_{i+1}, x_i, ..., x_1)$. By the inductive hypothesis, $Q_i x_i \in P(k), ..., Q_1 x_1 \in P(k), f(x_{i+1}, x_i, .$

Ah shoot. We can't use induction on num free vars, we'll have to use structural induction. $\square$

Lemma 3 (Gr Closed Under Permutation)**.** Let g be a ground formula, F be a quantified formula, and $k \in \mathbb{N}$ be given. Let $\pi : P(k) \to P(k)$ be a permutation, then:

$$(g \in \text{Gr}(F, k)) \leftrightarrow (M_\pi(g) \in \text{Gr}(F, k))$$

Proof. Suppose that $g \in \text{Gr}(F, k)$, then there exists a formula f such that $g = f[\mathbf{V} \mapsto P(k)]$ and $F \in par\_formula(f)$. However:

$$M_\pi(g) = M_\pi(f[\mathbf{V} \mapsto P(k)]) = f[\mathbf{V} \mapsto \pi(P(k))]$$

The other direction is straightforward if we realize that π^{-1} is also a permutation. $\square$

Lemma 4 (Minimum Gr)**.** Let F be a formula, f be a ground formula, and $k \in \mathbb{N}$ be given. Suppose that $\text{Elems}(f) \subseteq P(j)$ where $j \leq k$, then:

$$f \in \text{Gr}(F, k) \to f \in \text{Gr}(F, j)$$

Proof. I will need a better definition for Gr to prove this one. For now, proof by obviousness. □

Lemma 5. Let $k \in \mathbb{N}$, and $s \in \text{States}(k)$ such that $s \models \Phi(k)$. Then for $j \leq k$, it is also the case that $s \models \Phi(j)$.

Proof. Let k and $j \leq k$ be given and suppose that $s \models \Phi(k)$. By Lemma 1, $\forall f \in \text{Gr}(F, k), s \models \Phi(k)$. Now observe that $\text{Gr}(\Phi, j) \subseteq \text{Gr}(\Phi, k)$ due to the fact that Φ is a universally quantified PNF formula. Thus it is also the case that $\forall f \in \text{Gr}(F, j), s \models \Phi(j)$, and then the result follows from Lemma 1. □

4 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 6 (M-N Initiation)**.** Suppose that $\Phi(m)$ is an inductive invariant for $T(m)$, then $I(k) \to \Phi(k)$ for all $k > m$.

Proof. Coming soon.

□

Lemma 7 (M-N Consecution)**.** 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 inductive for any $k > m+n$.

Proof. Assume that $[\Phi \wedge \Delta \to \Phi'](m+n)$ is valid. Let $k > m+n$ be given and $s \in \text{States}(k)$ such that $s \models \Phi(k)$. Let $\delta \in \text{Gr}(\Delta, k)$, i.e. δ is a ground "transition". Let $t \in \text{States}(k)$ such that $t' \models (s \wedge \delta)$, that is, t' is an arbitrary "next" state of s. Finally, let $f' \in \text{Gr}(\Phi', k)$ be arbitrary, then, by Lemma 1 and the fact that Φ' is in PNF and universally quantified, it suffices to show that $t' \models f'$.

Next, we will construct a permutation π as follows: let $x_1, ..., x_j$ be the distinct elements of $P(k)$ used in δ and f'. We know that $j \leq m+n$ because Δ quantifies over n variables while Φ quantifies over m variables. Then:

$$\pi := \begin{pmatrix} x_1 & x_2 & \dots & x_j \\ 1 & 2 & \dots & j \end{pmatrix}$$

And hence by construction, $\text{Elems}(M_\pi(\delta)) \subseteq P(j)$ and $\text{Elems}(M_\pi(f')) \subseteq P(j)$. First, we immediately see that $M_\pi(\delta) \in \text{Gr}(\Delta, k)$ and $M_\pi(f') \in \text{Gr}(\Phi', k)$ by Lemma 3. Next, we further notice:

$$\text{Elems}(M_\pi(\delta)) \subseteq P(j) \subseteq P(m+n)$$

and

$$\text{Elems}(M_\pi(f')) \subseteq P(j) \subseteq P(m+n)$$

And thus by Lemma 4, we see that $M_\pi(\delta) \in \text{Gr}(\Delta, m+n)$ and $M_\pi(f') \in \text{Gr}(\Phi', m+n)$.

Now because $s \models \Phi(k)$, we see that $s \models \Phi(m+n)$ by Lemma 5, and furthermore $M_\pi(s) \models \Phi(m+n)$ by Assumption 1. Notice:

$$M_\pi(t' \models (s \wedge \delta)) \leftrightarrow M_\pi(t' \to (s \wedge \delta)) \leftrightarrow (M_\pi(t') \to M_\pi(s \wedge \delta)) \leftrightarrow M_\pi(t') \models M_\pi(s \wedge \delta)$$

Now:

$$M_\pi(t') \models M_\pi(s \wedge \delta) = M_\pi(s) \wedge M_\pi(\delta) \models [\Phi(m+n) \wedge \Delta(m+n)] = [\Phi \wedge \Delta](m+n)$$

Thus $M_\pi(t') \models [\Phi \wedge \Delta](m+n)$, which in turn implies $M_\pi(t') \models \Phi'(m+n)$ by our initial assumption.

Informal: Notice that $\text{Elems}(M_\pi(\delta)) \subseteq P(m+n)$, and hence the elements of the set $P(k) - P(m+n)$ have the same constraints in t' as they do in s; this means $M_\pi(t') \models (\Phi(k) - \Phi(m+n))'$ (for an abuse of notation). This further implies that $M_\pi(t') \models \Phi(k)$, and hence by Assumption 1, it follows that $t' \models \Phi(k)'$. In particular, by Lemma 1, $t' \models f'$.

□

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. □

5 Case Studies

In this section we visit several (more coming soon) distributed protocols that are parameterized by a single sort and satisfy Assumption 1.

5.1 Peterson's Mutex Protocol

Peterson's Mutex Protocol can be encoded with a transition relation Δ in PNF that quantifies over two variables. A sample inductive invariant candidate is given in [1] that quantifies of two variables and works for $|P| = 2$:

```
Phi == \A p,q \in ProcSet :
   /\ pc[p] \in {"a3","a4","cs"} => flag[p]
   /\ (p#q /\ pc[p] = "cs" /\ pc[q] = "a4") => turn = p
   /\ (p # q) => ~(pc[p] = "cs" /\ pc[q] = "cs")
```

However, by the M-N Theorem, we must show that Φ is an inductive invariant for the cases when $|P| = 1, ..., 4$. In fact, we easily see that $\Phi(3)$ fails to be inductive in the following counter example:

```
/\ turn = 1                /\ turn = 2
/\ pc[1] = "cs"            /\ pc[1] = "cs"
/\ pc[2] = "a4"    ->      /\ pc[2] = "a4"
/\ pc[3] = "a3"            /\ pc[3] = "a4"
/\ a3(3,2)
```

References

[1] Parametric Peterson's Mutex Protocol. https://github.com/iandardik/iinf/blob/master/ii_cutoff/mn_thm/PetersonParametric.tla, 2022.

[2] Mordechai Ben-Ari. *Mathematical Logic for Computer Science.* Springer Publishing Company, Incorporated, 3rd edition, 2012.

[3] 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.