Theoretical Foundation for Model Checking Role Containment in RT

Technical Report CS-TR-2008-017

Mark Reith* Jianwei Niu William H. Winsborough
University of Texas at San Antonio
One UTSA Circle, San Antonio, Texas, USA 78249
{*mreith, niu*}*@cs.utsa.edu, wwinsborough@acm.org*

ABSTRACT

Trust management is a scalable and flexible form of access control that relies heavily on delegation techniques. While these techniques may be necessary in large or decentralized systems, stakeholders need an analysis methodology and automated tools for reasoning about who will have access to their resources today as well as in the future. When an access control policy fails to satisfy the policy author's security objectives, tools should provide information that demonstrate how and why the failure occurred. Such information is useful in that it may assist policy authors in constructing policies that satisfy security objectives, which support policy authoring and maintenance. This paper presents a collection of reduction, optimization, and verification techniques useful in determining whether security properties are satisfied by RT policies. We provide proofs of correctness as well as demonstrate the degree of effectiveness and efficiency the techniques provide through empirical evaluation. While the type of analysis problem we examine is generally intractable, we demonstrate that our reduction and optimization techniques may be able to reduce problem instances into a form that can be automatically verified.

1. INTRODUCTION

Security analysis is a critical task in the design and maintenance of access control policies. The software components that comprise critical systems often endure extensive testing if not more rigorous verification measures such as formal methods. However, policy design that fails to satisfy security properties is an equally grave hazard. Correctly configuring security systems is often very difficult. In order to demonstrate and evaluate the correctness of access control properties, it is necessary to analyze the security policy. We use the term policy to refer to mechanism configuration, and use "security property" to refer to somewhat higher-level, though still formal, security objectives.

Policy design and maintenance can benefit from automated tools and techniques for evaluating policies. Every proposed change in the policy by trusted entities may require analysis prior to committing changes. It is critical to verify that the policy as stated meets one's intended policy objectives. Closely related is impact change analysis [4] where a policy author proposes a policy change, but must demonstrate that the change will not violate security objectives before committing it.

Without analysis techniques and tools, it may be difficult to claim that a particular access control policy exhibits desired characteristics. Policy authors and stakeholders seek effective, yet usable, techniques and tools in order to demonstrate such properties [9, 10, 12, 18] as availability, *e.g.*, will Alice always have access to resource R, and safety, *e.g.*, will access to resource R always be limited to some static set of users? One particularly useful, yet expensive, type of analysis is role containment [12]. Role containment asks whether every member of one role is contained within another role, which is useful in verifying safety properties. For example, is everyone with access to the database an employee? Development of techniques to reason about role containment properties is significant as it subsumes other types of security analysis.

Writing access control policies that reflect the author's intention is a challenging task for several reasons. First, the people who are charged with constructing or maintaining security in applications and systems are not always security experts. Second, even when expertise is not an issue, manually verifying policies can be a tedious and error-prone process, particularly for large or complex policies. Third, simply knowing that a security property fails to hold in a particular policy is not sufficient. Counterexamples are policy states in which the policy fails to satisfy one or more security properties, and are valuable because they provide the policy author insight as to how the policy might fail so that it may be quickly corrected. Finally, there is a significant gap between theoretical analysis and practical techniques that needs to narrow before security analysis can be widely accepted. Part of this gap is due to a lack of tool support that leverages the theory but does not require the user to necessarily be an expert. The other part of this gap is due to a lack of evaluation of theoretical techniques. While a general technique may be inefficient, could there be usable techniques that address efficiency issues in the verification of

*The views expressed in this article are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

specific policies? We provide the reader not only a proof of correctness of our techniques, but an assessment of efficiency over a set of test cases.

We present a collection of techniques for policy reduction and optimization to improve the ability of model-checking-based techniques to analyze containment queries. We believe our reduction and optimization techniques may be able to reduce such query instances into a form that can be verified. These techniques can be automated and associated with model checking to determine if a given policy satisfies a particular security property, and always provides an example of policy failure should it exist.

The structure of this paper is as follows. Section 2 describes the RT language and the complexity of role containment analysis. Section 3 describes reductions that provide support to our verification techniques. Section 4 describes our implementation of policy model checking. Section **??** evaluates our techniques over a collection of test cases. Section 5 compares our framework with related work, and we conclude in Section 6 with our contributions and future work.

2. THE RT POLICY LANGUAGE

The role-based trust management policy language RT was designed to support highly decentralized attribute-based access control [11]. It enables resource providers to make authorization decisions about resource requesters of whom they have no prior knowledge. This is achieved by delegating authority for characterizing principals in the system to other entities that are in a better position to provide the characterization. For instance, to grant discounted service to students, a resource provider might delegate to universities the authority to identify students and delegate to accrediting boards the authority to identify universities.

A significant problem that policy authors face in this context is that of determining the extent of their exposure through delegation to untrusted or semi-trusted principals. The security analysis problem [12] in this context consists of determining whether changes made by principals that are not fully trusted could cause certain policy objectives to become violated. One example of the problem would ask whether anyone outside the organization could, because of changes made by principals outside the inner circle, gain access to the organization's sensitive data. In this section, we summarize RT and the security analysis of it.

2.1 Overview of RT Syntax & Semantics

In RT, all principals are able to define their own roles and to assign other principals to them. A role owner can do this by issuing cryptographically verifiable, role-defining statements of a few different types. To her own roles, she can add a specific principal or she can add the members of another role. In the latter case, she is delegating authority to the owner of the other role. Delegating authority to another owner can occur in two ways. First she can identify a specific principal as a delegate. Secondly, she can identify a collection of principals as delegates such that these principals are grouped by a role. Set intersection and union are also both available for role definition.

The RT language consists of two primary objects called roles and principals. A principal is an entity such as a person or software agent. Each role can be described as a set of principals and is of the form "principal.role_name". One interpretation of this role is that the principal considers the members (also principals) of this role to have an attribute denoted by the role name. For example, *Alice.friend* may be a role that contains the principals who Alice considers friends.

Type	Syntax	Description
Type I	$A.r \leftarrow D$	Simple Member
Type II	$A.r \leftarrow B.r_1$	Simple Inclusion
Type III	$A.r \leftarrow B.r_1.r_2$	Linking Inclusion
Type IV	$A.r \leftarrow B.r_1 \cap C.r_2$	Intersection Inclusion

Figure 1: RT Statements

The basic RT language consists of four types of statements as shown by Figure 1 [12]. Type I statements directly introduce individual principals to roles. For example, *Alice.friend* $\leftarrow$ *Bob* identifies Bob as a friend of Alice. A given principal must appear in a Type I statement if it is to be contained by any role. Type II statements express a form of delegation that describes the implication that if principals are in one role, then they are in another role as well. For example, the statement *Alice.friend* $\leftarrow$ *Bob.friend* describes the situation in which if a principal is a friend of Bob, then they are also a friend of Alice. Type III statements provide a mechanism to delegate to a set of principals identified by membership in a role, rather than a single principal. For example, the statement *Alice.friend* $\leftarrow$ *Bob.family.friend* says that any friend of Bob's family is also a friend of Alice. It does not imply that Alice's friends include Bob's family. Finally, Type IV statements introduce intersection such that a principal must be in the two given roles in order to be included. For example, *Alice.friend* $\leftarrow$ *Bob.friend* $\cap$ *Carl.friend* says that only those principals who are both Bob's friends and Carl's friends are introduced into the set of Alice's friends. Note that disjunction is provided through multiple statements defining the same role.

Policy classes [12] are defined based on the type of statements included in a policy. $RT[\,]$ consists of simple member and inclusion statements. $RT[\twoheadleftarrow]$ includes $RT[\,]$ and introduces linking inclusion statements. $RT[\cap]$ also includes $RT[\,]$ and introduces intersection inclusion statements. Finally, $RT[\twoheadleftarrow, \cap]$ is comprised of four types of policy statements.

For any given RT policy statement, we call the left hand side of the arrow the *defined role* and the right hand side of the arrow the *role expression*. In Type III statements, the right hand side of the arrow is called a *linked role expression*, the *base-linked role* is the role that contains the principals upon which the linking role is applied, and the *sub-linked role* is the role produced by the linking role. For example, in $A.r \leftarrow B.r_1.r_2$, $B.r_1.r_2$ is a linked role expression, $B.r_1$ is the base-linked role, and every role of the form $X.r_2$ in which X is a member of $B.r_1$ is a sub-linked role. A *policy state* $\mathcal{P}$ is a set of statements. The set of role names in $\mathcal{P}$ is given by $\mathsf{Names}(\mathcal{P})$; the set of principals in $\mathcal{P}$ is given by $\mathsf{Principals}(\mathcal{P})$; the set of roles of $\mathcal{P}$ $\mathsf{Roles}(\mathcal{P}) = \{A.r \mid A \in \mathsf{Principals}(\mathcal{P}) \wedge r \in \mathsf{Names}(\mathcal{P})\}$. The *restriction of a policy* $\mathcal{P}$ to a given set of roles τ is given by $\mathcal{P}\restriction_\tau = \{A.r \leftarrow e \in \mathcal{P} \mid A.r \in \tau\}$.

Definition 1 (Semantics of RT). *The semantics of RT policy $\mathcal{P}$ is by the least fixpoint of the following function over functions $\pi : \mathsf{Roles}(\mathcal{P}) \rightarrow \wp(\mathsf{Principals}(\mathcal{P}))$, in which $\wp$*

denotes powerset:

For all $A.r \in \mathsf{Roles}(\mathcal{P})$, $T_{\mathcal{P}}(\pi)[\![A.r]\!] = \{D \mid$
$A.r \leftarrow D \in \mathcal{P} \vee$
$(A.r \leftarrow B.r_1 \in \mathcal{P} \wedge D \in \pi[\![B.r_1]\!]) \vee$
$(A.r \leftarrow B.r_1.r_2 \in \mathcal{P} \wedge \exists Z.Z \in \pi[\![B.r_1]\!] \wedge D \in \pi[\![Z.r_2]\!]) \vee$
$(A.r \leftarrow B.r_1 \cap C.r_2 \in \mathcal{P} \wedge D \in \pi[\![B.r_1]\!] \wedge D \in \pi[\![C.r_2]\!])\}$

(In the above, Quite quotes ([\![and]\!]) delimit syntactic objects denoted by metavariables, such as A and r.) The least fixpoint of $T_{\mathcal{P}}$ *can be computed as follows:*

$$T_{\mathcal{P}} \uparrow^0 = \pi_0, \text{ in which for all } A.r \in \mathsf{Roles}(\mathcal{P}),\ \pi_0[\![A.r]\!] = \emptyset$$

$$T_{\mathcal{P}} \uparrow^{i+1} = T_{\mathcal{P}}(T_{\mathcal{P}} \uparrow^i) \qquad T_{\mathcal{P}} \uparrow^{\omega} = \bigcup_{i < \omega} T_{\mathcal{P}} \uparrow^i$$

As the number of principals and role names in $\mathcal{P}$ *is finite, this increasing sequence converges at a finite stage. We now define the semantics as the function* $[\![\cdot]\!]_{\mathcal{P}}$ *that given any role* $B.r_1$*, is defined by* $[\![B.r_1]\!]_{\mathcal{P}} = T_{\mathcal{P}} \uparrow^{\omega} [\![B.r_1]\!]\}$.

2.2 RT Policy Analysis

Policy analysis [12] as we consider it here examines whether the specified relationships between roles hold in all reachable policy states. We explain reachable policy states below. The relationships, called *queries*, are set containments and take the form $\varrho \sqsupseteq \lambda$ in which ϱ and λ are each either roles or explicit (constant) sets of principals. For instance, $X.u \sqsupseteq A.r$ holds if every member of $A.r$ is a member of $X.u$ in every reachable policy state $\mathcal{P}'$, *i.e.*, $[\![A.r]\!]_{\mathcal{P}'} \supseteq [\![X.u]\!]_{\mathcal{P}'}$. Queries of this form can be used to express many important security properties such as availability, safety, liveness and mutual exclusion. For instance, a safety property might be that everyone in the role that has access to the secret database is in the employee role.

In general, any policy state can evolve into any other policy state by having principals issue new policy statements and revoke old ones. In security analysis we ask whether queries hold in all policy states that differ from a given current policy state *only* by changes to roles outside some trusted set. Intuitively, this corresponds to the fact that we expect certain principals to cooperate with us in our goal of preserving certain desired security properties. Specifically we assume that to this end these principals agree not to add or remove statements defining certain roles that they control. Other roles are not be assumed to be managed in cooperation with our goals. This intuition leads [12] to the defined two sets of roles that are used to determine the reachable policy states, the set of *growth-restricted* roles $\mathcal{G}_{\mathcal{R}}$ and the set of *shrink-restricted* roles $\mathcal{S}_{\mathcal{R}}$. Such a pair is called a *restriction rule* and is denoted by $\mathcal{R} = (\mathcal{G}_{\mathcal{R}}, \mathcal{S}_{\mathcal{R}})$.

Growth-restricted roles ($\mathcal{G}_{\mathcal{R}}$) are not allowed to have new statements defining them added to the state. Shrink-restricted roles ($\mathcal{S}_{\mathcal{R}}$) are not allowed to have statements defining them removed. We write $\mathcal{P} \overset{*}{\mapsto}_{\mathcal{R}} \mathcal{P}'$ to indicate that $\mathcal{P}' \restriction_{\mathcal{G}_{\mathcal{R}}} \subseteq \mathcal{P}$ and $\mathcal{P}' \supseteq \mathcal{P} \restriction_{\mathcal{S}_{\mathcal{R}}}$. It is important to note that these restrictions are not actually enforced. They are simply assumptions under which the analysis is performed. Their presence enables the analysis to provide us with assurances of things like, "So long as the people I trust do not make policy changes without first running the analysis, only company employees will be able to access the secret database."

Queries of certain restricted forms can be analyzed and verified in polynomial time. These include queries in which at least one of ϱ and λ is an explicit set. They also include situations in which only Type I and Type II statements are allowed in the policy state. However, when both ϱ and λ are roles and all forms of statements are allowed, the decision problem is EXPTIME complete [20]. This is unfortunate because such properties are extremely useful. For instance, suppose we want to determine whether only employees could ever get access to a company's secret database. This can be determined efficiently if the set of employees is enumerated explicitly in the query. However, this does not consider the effect of employee turnover. By identifying employees and those users with access to the database both as roles, we can determine whether the desired property will continue to hold as new employees are added. Thus we seek techniques that can solve queries of this general form as often as possible.

DEFINITION 2 (RCPI). *An* instance of a role containment problem *(RCPI) is given by a triple* $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$. *An RCPI is said to be* satisfied *if and only if* $[\![X.u]\!]_{\mathcal{P}'} \supseteq [\![A.r]\!]_{\mathcal{P}'}$ *for each* $\mathcal{P}'$ *such that* $\mathcal{P} \overset{*}{\mapsto}_{\mathcal{R}} \mathcal{P}'$. *In this case we also say that* $\mathcal{P}$ satisfies $X.u \sqsupseteq A.r$ under $\mathcal{R}$.

3. REDUCTIONS

This section describes several reductions that transform one RCPI into another that is typically less expensive to evaluate. Our findings in Section **??** indicate that, when using our model checking technique and our platform configuation, these reductions often make the difference between being unable to evaluate an RCPI and being able to do so. We conjecture that they may also reduce the cost of applying other approaches to solve RCPI problems, such as one based one on the proof method of Sistla [20, 21].

3.1 Infinite State Space Reduction

Any analysis technique that operates by exhaustively examining reachable states must address the fact that in our analysis problem the size of reachable policy states is unbounded, and hence the state space is infinite. In this section we present a subspace of bounded size that was previously identified [12] and that has the property that, given any query, any restriction rule, and any initial state, there exists a reachable state in which the query is violated if and only if there exists a reachable state within the bounded state space in which the query is violated.

Given a policy state $\mathcal{P}$, a restriction rule $\mathcal{R}$, and a query $\mathcal{Q} = X.u \sqsupseteq A.r$, the state space that must be considered consists of all states that are reachable from $\mathcal{P}$ under $\mathcal{R}$ and that are composed of statements in $\mathcal{P}$ and in $\mathcal{N}$, in which $\mathcal{N}$ is constructed as follows: $\mathcal{N} = \{A.r \longleftarrow D \mid r \in \mathsf{Names}(\mathcal{P}) \wedge \mathsf{A}, \mathsf{D} \in \mathsf{Principals}(\mathcal{P}) \cup \mathsf{NewPrinc}(\mathcal{P}, \mathcal{Q})\}$, $\mathsf{NewPrinc}(\mathcal{P}, \mathcal{Q})$ is a set of new principals of size 2^K, $K = |\mathsf{SigRoles}(\mathcal{P}, \mathcal{Q})|$, and $\mathsf{SigRoles}(\mathcal{P}, \mathcal{Q})$ is the set $\{X.u\} \cup \{A.r_1 \mid A.r \longleftarrow A.r_1.r_2 \in \mathcal{P}\} \cup \{B_1.r_1, B_2.r_2 \mid A.r \longleftarrow B_1.r_1 \cap B_2.r_2 \in \mathcal{P}\}$. It is shown by Li *et al.* [12] that this set of reachable states has the desired property. We consider the number of new principals used here to be conservative in the sense that any fewer number of principals does not guarantee this desired property.

3.2 Unrestricted Role Reduction

We now present a result that in many cases enables us to further reduce the size of the state space that must be explored. The idea is that given an initial state $\mathcal{P}$, we can often perform the analysis using a smaller initial state and

obtain identical results. It is important to note that this is significantly different than the Lower Bound $LB(\mathcal{P})$ and Upper Bound $UB(\mathcal{P})$ programs described in [12]. The $LB(\mathcal{P})$ program removes all statements defining a role $C.r \notin \mathcal{S}_\mathcal{R}$ from $\mathcal{P}$ for the purpose of evaluating a query of the form $X.u \sqsupseteq \{B \mid B \in Principals\}$, whereas the $UB(\mathcal{P})$ program adds a special principal τ representing all principals to each growth unrestricted role for the purposes of evaluating a query of the form $\{B \mid B \in Principals\} \sqsupseteq A.r$. Our reduction program removes all statements defining a role $C.r \notin \mathcal{S}_\mathcal{R} \wedge C.r \notin \mathcal{G}_\mathcal{R}$ for the purpose of evaluating a query of the form $X.u \sqsupseteq A.r$.

We define $\equiv_\mathcal{R}$, a binary relation over policy states, by $\mathcal{P}_1 \equiv_\mathcal{R} \mathcal{P}_2$ if and only if $\mathcal{P}_1 \overset{*}{\mapsto}_\mathcal{R} \mathcal{P}_2$ and $\mathcal{P}_2 \overset{*}{\mapsto}_\mathcal{R} \mathcal{P}_1$. Note that $\equiv_\mathcal{R}$ is an equivalence relation: (1) it is reflexive, as $\mathcal{P} \overset{*}{\mapsto}_\mathcal{R} \mathcal{P}$ for all $\mathcal{P}$; (2) it is symmetric by construction; (3) it is transitive because $\overset{*}{\mapsto}_\mathcal{R}$ is transitive.

We also define the core of $\mathcal{P}$, $\mathsf{core}_\mathcal{R}(\mathcal{P})$, to be the subset of $\mathcal{P}$ consisting of those statements that define growth-restricted or shrink-restricted roles.

THEOREM 1. *Given any policy states $\mathcal{P}_1$ and $\mathcal{P}_2$ and any restriction rule $\mathcal{R}$, $\mathcal{P}_1 \equiv_\mathcal{R} \mathcal{P}_2$ if and only if $\mathsf{core}_\mathcal{R}(\mathcal{P}_1) = \mathsf{core}_\mathcal{R}(\mathcal{P}_2)$.*

PROOF. The "if" direction is straightforward: starting from any state $\mathcal{P}$, it is clear that any other state with the same core is reachable from $\mathcal{P}$. For the "only if" direction, suppose $\mathsf{core}_\mathcal{R}(\mathcal{P}_1) \neq \mathsf{core}_\mathcal{R}(\mathcal{P}_2)$. There are two cases, depending on whether the cores differ in statements that define roles that are growth-restricted or roles that are shrink-restricted. If $\mathsf{core}_\mathcal{R}(\mathcal{P}_1) \backslash \mathsf{core}_\mathcal{R}(\mathcal{P}_2)$ contains a statement defining a growth restricted role, then $\mathcal{P}_2 \overset{*}{\mapsto}_\mathcal{R} \mathcal{P}_1$ does not hold. If $\mathsf{core}_\mathcal{R}(\mathcal{P}_1) \backslash \mathsf{core}_\mathcal{R}(\mathcal{P}_2)$ contains a statement defining a shrink restricted role, then $\mathcal{P}_1 \overset{*}{\mapsto}_\mathcal{R} \mathcal{P}_2$ does not hold. □

Given this result, it is clear that the least state (under the subset ordering) that is equivalent to a given $\mathcal{P}$ is $\mathsf{core}_\mathcal{R}(\mathcal{P})$. It follows from the definition of $\equiv_\mathcal{R}$ and the transitivity of $\overset{*}{\mapsto}_\mathcal{R}$ that for all $\mathcal{P}'$, $\mathcal{P} \overset{*}{\mapsto}_\mathcal{R} \mathcal{P}'$ if and only if $\mathsf{core}_\mathcal{R}(\mathcal{P}) \overset{*}{\mapsto}_\mathcal{R} \mathcal{P}'$. Moreover, $\mathsf{core}_\mathcal{R}(\mathcal{P})$ is the least set for which this is true. Thus, given an initial state $\mathcal{P}$, it is both correct and more efficient to perform our analysis using $\mathsf{core}_\mathcal{R}\ (\mathcal{P})$ as the initial state instead.

DEFINITION 3. *Given a policy $\mathcal{P}$ and restriction rule $\mathcal{R}$, URR ($\mathcal{P}$, $\mathcal{R}$) = $\mathsf{core}_\mathcal{R}\ (\mathcal{P})$*

3.3 Cone of Influence Reduction

A given security policy $\mathcal{P}$ may contain statements that do not affect the membership of queried roles. Such extraneous statements can safely be filtered from the policy model in order to reduce the size of the problem. This reduction removes statements that are said to be outside of a role's cone of influence. This reduction is particularly significant because it has the potential to remove linked or intersection inclusion type statements that contribute to a larger number of principals in the model.

Given a set of roles, Λ, the following definition constructs a set of roles in $\mathcal{P}$. A role, ρ, is in the constructed set if the membership of some role in Λ depends in some way on the membership of ρ. This dependency is reflective of relationship established between roles via RT statements. We call this set of roles *DefRoles*.

DEFINITION 4. *Let Λ and $\mathcal{M}$ be sets of roles, and $\mathcal{P}$ be a policy. We define DefRoles($\mathcal{P}, \Lambda, \mathcal{M}$) to be the least set of roles $\mathcal{O}$ satisfying the following conditions:*

- $\Lambda \subseteq \mathcal{O}$
- $(\lambda \in \mathcal{O} \ \wedge \ \lambda \leftarrow B.r_1 \in \mathcal{P} \wedge B.r_1 \notin \mathcal{M}) \Rightarrow B.r_1 \in \mathcal{O}$
- $(\lambda \in \mathcal{O} \ \wedge \ \lambda \leftarrow B.r_1.r_2 \in \mathcal{P} \wedge D \in Principals) \Rightarrow ((B.r_1 \notin \mathcal{M} \Rightarrow B.r_1 \in \mathcal{O}) \wedge (D.r_2 \notin \mathcal{M} \Rightarrow D.r_2 \in \mathcal{O}))$
- $(\lambda \in \mathcal{O} \ \wedge \ \lambda \leftarrow B.r_1 \ \cap \ C.r_2 \in \mathcal{P}) \Rightarrow ((B.r_1 \notin \mathcal{M} \Rightarrow B.r_1 \in \mathcal{O}) \wedge (C.r_2 \notin \mathcal{M} \Rightarrow C.r_2 \in \mathcal{O}))$

Using *DefRoles*, we define the *Cone of Influence* (COI) as a policy constructed from those statements that define roles in *DefRoles*. In other words, we retain only those statements that influence the answer to an RCPI.

DEFINITION 5. *Given an RCPI $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$, we define*

$$COI(\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle) = \mathcal{P}\restriction_{DefRoles(\mathcal{P}, \{X.u\}, \mathcal{S}_\mathcal{R}) \cup DefRoles(\mathcal{P}, \{A.r\}, \mathcal{G}_\mathcal{R})}$$

THEOREM 2. *Given any RCPI $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$, let $\mathcal{P}' = COI(\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle)$. Then $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$ is satisfied if and only if $\langle \mathcal{P}', \mathcal{R}, X.u \sqsupseteq A.r \rangle$ is satisfied.*

Proof for this and subsequent reductions are given in [17].

3.4 Decomposition

It is sometimes useful to decompose an RCPI into subproblems that can be solved separately. The membership of $A.r$ is the union of $[\![e_i]\!]_\mathcal{P}$ for $i = 1 \ldots n$ in which $e_1 \ldots e_n$ enumerates $\{e \mid A.r \leftarrow e \in \mathcal{P}\}$. Thus if $A.r$ is growth and shrink restricted, then for each reachable state $\mathcal{P}'$ and each principal $E \in [\![A.r]\!]_{\mathcal{P}'}$, there is some $i \in \{1 \ldots n\}$ such that $E \in [\![e_i]\!]_{\mathcal{P}'}$. By isolationg e_i through the use of a new role $A'.r'$, we construct from a given RCPI a collection of new RCPIs such that the orginal is satisfied just in case each RCPI in the collection is satisfied. We demonstrate in Section **??** that the decomposed RCPIs can sometimes be successfully solved by our analysis tool when the original cannot.

Decompose constructs the collection of new RCPIs.

DEFINITION 6. *Given an RCPI $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$, let A' be a new principal not in $\mathcal{P}$ and r' be a new role name not in $\mathcal{P}$. We define Decompose($\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$) = $\{\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq \rho \rangle \mid A.r \leftarrow \rho \in \mathcal{P} \wedge \rho \in \mathsf{Roles}\} \cup \{\langle \mathcal{P} \cup \{A'.r' \leftarrow e\}, \langle \mathcal{G}_\mathcal{R} \cup \{A'.r'\}, \mathcal{S}_\mathcal{R} \cup \{A'.r'\} \rangle, X.u \sqsupseteq A'.r' \rangle \mid A.r \leftarrow e \in \mathcal{P} \wedge e \notin \mathsf{Roles}\}$.*

Note that each of the new policies contains at most one occurrence of $A'.r'$, and that there is no occurrence of either A' or r' in any statement body. Furthermore, when e is a role, we do not make use of the new role $A'.r'$. We now formally describe the relationship between the solution to the original RCP instance and the solutions of the new problems.

THEOREM 3. *Given an RCPI $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$, if $A.r \in \mathcal{G}_\mathcal{R} \cap \mathcal{S}_\mathcal{R}$, then $\mathcal{P}$ satisfies $X.u \sqsupseteq A.r$ under $\mathcal{R}$ if and only if $\mathcal{P}'$ satisfies $X.u \sqsupseteq \rho$ under $\mathcal{R}'$ for each $\langle \mathcal{P}', \mathcal{R}', X.u \sqsupseteq \rho \rangle \in Decompose(\mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r)$.*

3.5 Chain Reduction

One type of optimization that is often useful involves pruning the part of the policy that does not impact the membership of the queried roles. This effectively collapses a set of states into a single state such that role containment evaluation on that single state is representative of all the collapsed states. The efficiency of this is evident from not having to evaluate all reachable subsets of pruned statements. It is trivial to see the benefit of such a reduction on the initial state, but more practical technique would be one that can be applied to each reachable state. The most effective pruning would require the evaluation of the queried roles to determine which statements could be removed, however little is gained by evaluating queried role membership in order to prune the policy and then re-evaluating the queried role membership. We propose a technique for pruning a policy without evaluating the queried role membership that is guaranteed to be conservative but not necessarily the smallest pruned policy possible. This is possible by recognizing that given a reachable state, a role that is not defined by any statement must be empty in that state. Since empty roles cannot contribute any principal to the queried roles per the monotonicity of the language, other statements that depend on these empty roles may also be removed. Such an approach can be efficiently implemented as a set of constraints on the reachable statespace, thus reducing the size of the reachable statespace without incurring the cost of evaluating role membership.

DEFINITION 7. *Given any policy $\mathcal{P}$ and any restriction rule $\mathcal{R}$, we define $Chain(\mathcal{P}, \mathcal{R})$ to be the greatest set of policy statements $\mathcal{P}'$ that satisfies the following constraints:*

1. $\mathcal{P}' \subseteq \mathcal{P}$
2. $B.r \in \mathcal{G}_{\mathcal{R}} \wedge \neg\exists e.\ B.r \leftarrow e \in \mathcal{P}' \Rightarrow$
   $\neg\exists \lambda.\ \lambda \leftarrow B.r \in \mathcal{P}' \wedge$
   $\neg\exists \lambda\ \exists r_2.\ \lambda \leftarrow B.r.r_2 \in \mathcal{P}' \wedge$
   $\neg\exists \lambda\ \exists C\ \exists r_2.\ \lambda \leftarrow B.r \cap C.r_2 \in \mathcal{P}'$
   $\neg\exists \lambda\ \exists C\ \exists r_2.\ \lambda \leftarrow C.r_2 \cap B.r \in \mathcal{P}'$

where $\Rightarrow$ represents implication, $\neg$, negation, $\exists$, an existential quantification, and e is a role expression.

THEOREM 4. *Given any RCPI $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$, and any policy state $\widehat{\mathcal{P}}$ such that $\mathcal{P} \overset{*}{\mapsto}_{\mathcal{R}} \widehat{\mathcal{P}}$, let $\mathcal{P}' = Chain(\widehat{\mathcal{P}}, \mathcal{R})$. Then $[\![X.u]\!]_{\widehat{\mathcal{P}}} = [\![X.u]\!]_{\mathcal{P}'}$ and $[\![A.r]\!]_{\widehat{\mathcal{P}}} = [\![A.r]\!]_{\mathcal{P}'}$.*

Consider the example in Figure 2 where $\mathcal{P}$ is the initial policy and $\widehat{\mathcal{P}}$ is a reachable state according to $\mathcal{R}$. In removing the statement $B.r \leftarrow C.r$, it can easily be determined that $[\![B.r]\!]_{\widehat{\mathcal{P}}} = \emptyset$ since no statement defines $B.r$. Since $B.r$ is empty, then clearly the statements $X.u \leftarrow B.r$ and $A.r \leftarrow B.r$ do nothing to contribute to the membership of the queried roles, and thus can be safely removed, yielding $\mathcal{P}'$. We can then further reduce this policy state by applying the Cone of Influence reduction and removing statements that define roles not in *DefRoles*, yielding $\mathcal{P}''$. In this example, $\mathcal{P}$ would have required us to examine at least 2^9 states (9 statements defining non-shrink restricted roles), but since $\mathcal{P}'$ is the smallest representative of the set $\{\mathcal{P}' \cup \mathcal{P}''' \mid \mathcal{P}''' \subseteq \{0, 2, 5, 6, 7, 8\}\}$, we save ourselves from examining 2^6 states through this technique.

Index	Policy Statement of $\mathcal{P}$		
0	$X.u \leftarrow B.r$	6	$C.r \leftarrow E.r$
1	$X.u \leftarrow J$	7	$D.r \leftarrow F$
2	$A.r \leftarrow B.r$	8	$D.r \leftarrow G$
3	$A.r \leftarrow K$	9	$E.r \leftarrow H$
4	$B.r \leftarrow C.r$	10	$E.r \leftarrow I$
5	$C.r \leftarrow D.r$		

$\mathcal{G}_{\mathcal{R}} = \{X.u, A.r, B.r, C.r, D.r, E.r, F.r\}$, $\mathcal{S}_{\mathcal{R}} = \{E.r\}$
RCPI: $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$

Policy State	Statement Indices	Comment
$\mathcal{P}$	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10	Initial policy
$\widehat{\mathcal{P}}$	0, 1, 2, 3, 5, 6, 7, 8, 9, 10	Reachable state
$\mathcal{P}'$	1, 3, 5, 6, 7, 8, 9, 10	Apply $Chain(\widehat{\mathcal{P}}, \mathcal{R})$
$\mathcal{P}''$	1, 3, 9, 10	Apply $COI(\mathcal{P}', \mathcal{R})$

Figure 2: Chain Reduction Example

4. POLICY MODEL CHECKING

Automated techniques for verifying RT policies are necessary for two important reasons. First, manual techniques may be tedious and error prone even in relatively simple RT policies. Second, it is not difficult to succumb to the complexity of even moderately sized policies, particularly since policies devised in a decentralized environment may exhibit characteristics that interrelate several roles in complicated ways. For these reasons, automated analysis seems especially useful to policy authors and stakeholders. We specifically explore the use of model checking as an automated tool to verify if a given policy satisfies a given security specification.

We present a novel technique for model checking any given role containment problem in RT. We address the issue of cyclic dependencies by incorporating the fixpoint evaluation into the model with the intent to let the model checking tool find a counterexample as a result of this process. Such a dependency occurs when at least one role depends, directly or indirectly, on its own membership to reach the least fixpoint in membership evaluation. We also describe a collection of reductions and optimizations that assist the model checking effort by reducing the policy model into a form that may be tractable. Section 3 formally described these techniques and here we describe how they fit into the overall analysis.

We begin by providing an overview to model checking with a specific emphasis on symbolic model checking. Next, we describe our policy model and translation to the model checking tool. Finally, we discuss how reductions and optimizations are applied in our technique.

4.1 Model Checking Overview

Model checking [1, 2] is an automated verification technique that constructs a finite model of a system and exhaustively explores the state space of this model to determine if desired properties hold in all reachable states. When they do not, a counterexample will be produced to show an error trace, which can be used to fix the model or the property specification. The model is composed of state and transition relations. The state of the system is determined by the current values of all state variables, whereas the state space is the set of all possible states. The transition relation is defined by the set of next assignments which execute concurrently in a step to determine the next state of the model.

We are only interested in those states that are reachable from a given start state.

The properties we want to check are called specifications, which are expressed in temporal logic [15] formulas. Temporal logic is a language for expressing properties related to a sequence of states in terms of temporal logic operators and logic connectives (e.g., $\wedge$ and $\vee$). Temporal operators X, F, and G represent next state, some future state, and all future states, respectively. For example, Gp means that property p is always true in all possible states.

SMV [14] is a mature, symbolic model checking tool. Although conceptually similar to the aforementioned description, it operates on a set of states at a time rather than each explicit state. This is accomplished through the use of specialized data structures encoding the set of initial states, the transition relation, and the set of states representing the negated specification. The transition relation is repeatedly applied to the negated specification until a fixpoint is reached. For each application of the transition relation, if the produced set of states intersects with any of the initial states, the model fails to satisfy the specification. This approach often permits substantially larger models to be verified than the explicit approach.

4.2 Model Checking RT Policies

We now formally describe the construction of our policy model, associated finite state machine, and translation to SMV.

4.2.1 Maximal Relevant Policy Statements

We begin by constructing a set of policy statements that serve as the foundation for evaluating the input policy. The set of maximal relevant policy statements (MRPS) represents a policy state reachable from the initial policy $\mathcal{P}$ and constructed by adding a set of simple member statements that introduce an appropriate number of new principals into each non-growth restricted role. From the Statespace Reduction in Section 3.1, it is easy to see that if $\mathcal{P}$ fails to satisfy the query, then there exists some $\mathcal{P}'$ and principal E such that $E \in [\![A.r]\!]_{\mathcal{P}'}$, $E \notin [\![X.u]\!]_{\mathcal{P}'}$, and $\mathcal{P}' \subseteq MRPS(\mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r)$. Thus $MRPS(\mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r)$ is considered a maximal policy state since the set of reachable states to be evaluated are subsets of this set. Figure 3 formally describes the construction of $MRPS(\mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r)$.

Two important observations regarding the placement and composition of principals in the $MRPS(\mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r)$ should be made from Figure 3. First, while the set $ModelPrinc(\mathcal{P})$ consists of both existing and new principals, we can limit the set of existing principals to those that are directly introduced via simple member statements in $\mathcal{P}$ and those that, when combined with a linked role name in $\mathcal{P}$, produce a role found in $\mathcal{P}$. It is easy to see that all other existing principals in $\mathcal{P}$ produce roles that are representable by new principals. Let $B \notin EffectPrinc(\mathcal{P})$ and $B.r_1$ be a role that is not found in $\mathcal{P}$. If $B.r_1 \in \mathcal{G}_{\mathcal{R}}$, then $B.r_1$ must be empty since there are no statements defining it in $\mathcal{P}$. If $B.r_1 \in \mathcal{S}_{\mathcal{R}}$, then no principal is forced into $B.r_1$ since there are no statements defining it in $\mathcal{P}$. Thus $B.r_1$ is effectively unrestricted and B may serve as a new principal.

Second, model principals are only added to roles that are both non-growth restricted and in the set of $DefRoles(\mathcal{P}, A.r)$. This is easy understood from the examination of $\mathcal{P}'$, where $E \in [\![A.r]\!]_{\mathcal{P}'}$, $E \notin [\![X.u]\!]_{\mathcal{P}'}$. Clearly the membership of $X.u$ is irrelevant so long as $E \notin [\![X.u]\!]_{\mathcal{P}'}$. Thus minimizing the introduction of principals into $DefRoles(\mathcal{P}, X.u) - DefRoles(\mathcal{P}, A.r)$ serves to reduce the number of policy states to be examined while preserving $\mathcal{P}'$ if one should exist.

$$
\begin{aligned}
S &= \alpha \times \beta \times \{fix, eval\} \\
\alpha &= \{\mathcal{P}' \mid \widehat{\mathcal{P}} \restriction_{\mathcal{S}_{\mathcal{R}}} \subseteq \mathcal{P}' \wedge \mathcal{P}' \subseteq \widehat{\mathcal{P}}\} \\
\beta &= Roles(\widehat{\mathcal{P}}) \rightarrow 2^{ModelPrinc(\mathcal{P})} \\
&\quad (i.e., \beta \text{ is the set of functions from roles to sets of} \\
&\quad \text{principals.}) \\
\widehat{\mathcal{P}} &= MRPS(\mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r) \\
&Roles(\mathcal{P}) \\
&\quad = \{B.r \mid B.r \leftarrow e \in \mathcal{P} \vee \lambda \leftarrow B.r.r_1 \in \mathcal{P} \vee \\
&\quad \lambda \leftarrow B.r \cap C.r_1 \in \mathcal{P} \vee \lambda \leftarrow C.r_1 \cap B.r \in \mathcal{P}\} \cup \\
&\quad SubLinkRoles(\mathcal{P}) \\
&SubLinkRoles(\mathcal{P}) \\
&\quad = \{B.r \mid B \in ModelPrinc(\mathcal{P}) \\
&\quad \wedge\, r \in LinkedRoleNames(\mathcal{P})\} \\
\sigma_0 &= \langle \widehat{\mathcal{P}}, \pi_0 \rangle \\
\pi_0(\lambda) &= \emptyset, \text{ forall } \lambda \in Roles(\widehat{\mathcal{P}}) \\
\delta &= \{\langle\langle \mathcal{P}_1, \pi, eval\rangle, \langle \mathcal{P}_1, \pi', eval\rangle\rangle \mid \pi' = T_{\mathcal{P}_1}(\pi)\} \cup \\
&\quad \{\langle\langle \mathcal{P}_1, \pi, eval\rangle, \langle \mathcal{P}_1, \pi, fix\rangle\rangle \mid \pi = T_{\mathcal{P}_1}(\pi)\} \cup \\
&\quad \{\langle\langle \mathcal{P}_1, \pi, fix\rangle, \langle \mathcal{P}_2, \pi_0, eval\rangle\rangle \mid \widehat{\mathcal{P}} \restriction_{\mathcal{S}_{\mathcal{R}}} \subseteq \mathcal{P}_2 \wedge \mathcal{P}_2 \subseteq \widehat{\mathcal{P}}\}
\end{aligned}
$$

Figure 4: Construction of AFSM for $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$

4.2.2 Analysis Finite State Machines

The following definition shows how we use $MRPS(\mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r)$ to define the finite state machine that we model check.

DEFINITION 8 (AFSM). *Given an RCPI, $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$, the corresponding* analysis finite state machine *(AFSM) is given by the 3-tuple, $\langle S, \sigma_0, \delta \rangle$, in which S is a finite, non-empty set of states, $\sigma_0 \in S$ is the initial state, $\delta \subseteq S \times S$ is a transition relation, and these structures are constructed as shown in Figure 4.*

The AFSM state includes not only a policy state, but also a mapping that takes a role to a set of principals in that role. Because the RT language allows cyclic dependencies among roles, we found it necessary to use a sequence of state transitions to calculate the fixpoint that determines the membership of each role. This is accomplished by differentiating between what we call *fixpoint mode* (*fix*) and *evaluation mode* (*eval*). Evaluation mode signifies that the calculation of role memberships has not stabilized, so the query should not be evaluated yet; from such a state, the fixpoint evaluation should proceed and the policy state should not change. Fixpoint mode signifies that role membership under the given policy is now known and the query should be evaluated; provided the query is satisfied, from such a state, a new policy state should be chosen.

$MRPS(\mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r) = \mathcal{P} \cup \{\lambda \leftarrow e \mid \lambda \notin \mathcal{G}_\mathcal{R} \wedge \lambda \in DefRoles(\mathcal{P}, \{A.r\}) \wedge e \in ModelPrinc(\mathcal{P})\}$, where

1. $ModelPrinc(\mathcal{P}) = NewPrinc(\mathcal{P}) \cup EffectPrinc(\mathcal{P}) \cup DirectPrinc(\mathcal{P})$
2. $NewPrinc\ (\mathcal{P})$ is any set of principals of size $2^{|SigRoles(\mathcal{P})|}$ such that $NewPrinc(\mathcal{P}) \cap (Principals(\mathcal{P}) \cup Principals(\mathcal{R})) = \emptyset$
3. $SigRoles(\mathcal{P}) = \{X.u\} \cup \{B.r_1 \mid C.r \leftarrow B.r_1.r_2 \in \mathcal{P}\} \cup \{B_1.r_1, B_2.r_2 \mid C.r \leftarrow B_1.r_1 \cap B_2.r_2 \in \mathcal{P}\}$
4. $EffectPrinc(\mathcal{P}) = \{B \mid B \in Principals(\mathcal{P}) \wedge B.r \leftarrow e \in \mathcal{P} \wedge r \in LinkedRoleNames(\mathcal{P})\}$
5. $DirectPrinc(\mathcal{P}) = \{B \mid \lambda \leftarrow B \in \mathcal{P}\}$
6. $LinkedRoleNames(\mathcal{P}) = \{r_1 \mid \lambda \leftarrow B.r.r_1 \in \mathcal{P}\}$
7. $Principals(\mathcal{P}) = \{B \mid B.r \leftarrow e \in \mathcal{P} \vee \lambda \leftarrow B \in \mathcal{P} \vee \lambda \leftarrow B.r.r_1 \in \mathcal{P} \vee \lambda \leftarrow B.r \cap C.r_1 \in \mathcal{P} \vee \lambda \leftarrow C.r \cap B.r_1 \in \mathcal{P}\}$
8. $Principals(\mathcal{R}) = \{B \mid B.r \in \mathcal{G}_\mathcal{R} \vee B.r_1 \in \mathcal{S}_\mathcal{R}\}$

Figure 3: Construction of $MRPS(\mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r)$

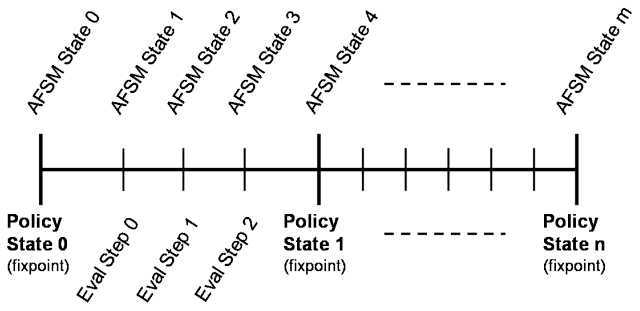

Figure 5: Policy to AFSM Mapping of States

4.2.3 Translation to SMV

Now that we have established the AFSM for a particular policy, restriction rule, and query, we can now translate the AFSM into SMV's input language. Translation consists of constructing four core components: data structure declaration, initialization, next state relations, and specification. Together these components define the set of states that will be evaluated against the specification. These states define not only a particular policy state, but also the membership of roles in the policy state as it is evaluated against the specification. Recall that the finite state machine represents not only policy states, but intermediate role membership evaluations of each policy state. In constructing it this way, the finite state machine simulates the fixpoint calculation to determine the membership of the queried roles. Only policy states that have reached the simulated fixpoint are evaluated against the specification.

Each finite state machine declares a set of data structures consisting of a bit vector representing the set of statements in the model, a bit vector per role representing the membership of that role, and a boolean variable that distinguishes two evaluation modes. The statement bit vector defines the current policy state by indicating which policy statements are included in that state. Each index into the bit vector corresponds to a particular policy statement in the maximal policy state. Statements that define shrink restricted roles and are included in the initial policy are present in all reachable states. In addition, each role is represented by a distinct bit vector where each index represents whether a principal is a member of that role. Finally, the evaluation mode boolean provides a means to evaluate policies with circular dependencies. The boolean signifies whether the current policy state is still performing a fixpoint calculation or not. Simulating fixpoint calculations is necessary for correctly handling policies with cyclic dependencies. The policy state is only allowed to change if the fix point simulation has finished for a particular state.

Initialization of these variables is a straightforward process. Each statement bit variable that defines a role that is shrink restricted is asserted, while every other statement variable is non-deterministically set as either asserted or not asserted. This constructs a set of initial policy states that are reachable from the initial input policy state. By using a set of initial policy model states rather than a single initial state, we may trade memory (larger BDD size) for speed (fewer fixpoint calculations). Role membership bits are initialized to represent empty roles, and the mode bit is always initialized to assert a fixpoint mode.

The next state relation component defines how the AFSM may change. This component is converted by SMV into a transition relation. The manner in which the model policy state is changed is determined by the fixpoint mode. If the model is in a fixpoint mode, the policy state must remain consistent while the membership of the roles is updated. This occurs until a fixpoint (none of the roles can introduce any more principals) is reached, at which point the statement bit vector is non-deterministically set to a new reachable policy state, the role membership bit vectors are re-initialized, and the mode is reset to fixpoint mode.

The specification component expresses the role containment property to be verified. Given a policy property such as $X.u \sqsupseteq A.r$, the specification verifies the model property that all model states in which the fixpoint calculation has completed implies that there exists no witness principal in the membership of $A.r$ that is not also found it the membership of $X.u$. Figure 6 describe such SMV specifications. For comparison, we include other types of RT queries represented as SMV specifications as well.

4.2.4 Translation of Policy without Cycles

Our previous work [16] remains useful in cases where the input policy does not contain cycles, and the preferred ap-

Property	RT Query	SMV LTL Specification	Notes
Availability	$A.r \sqsupseteq \{C, D\}$ Always	`assert G (fixpoint -> (Ar[0] & Ar[1]))`	C and D in A.r
Safety	$\{C, D\} \sqsupseteq A.r$ Always	`assert G (fixpoint -> (~ Ar[2]))`	E not in A.r
Containment	$A.r \sqsupseteq B.r$ Always	`assert G (fixpoint -> (Ar | Br = Ar))`	Nothing new in B.r

Figure 6: RT Queries to SMV Specifications

proach in such cases since the overhead of simulating the fixpoint role membership calculation can be removed. Specifically the state space we need to consider is limited to the number of subsets of statements from the *MRPS*. This can be easily understood from the fact that in an acyclic policy, the membership of each role needs to be evaluated at most once, and thus can be implemented as a set of constraints. This permits larger models to be verified than the approach described above. We compare the general translation against the non-cyclic translation in Section **??**.

4.3 Optimizations

In Section 3, we demonstrated the correctness of several techniques that may assist in reducing the size of the input policy and possibly reduce the complexity of the analysis. We describe these techniques below as they are applied to our analysis technique.

4.3.1 Pre-Processing

Many of the reductions from Section 3 take a policy $\mathcal{P}$, restriction rule $\mathcal{R}$, and role containment query $X.u \sqsupseteq A.r$ (an instance of RCP such as $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$) and produce either an equivalent problem or a set of problems that when combined produce the original problem. Such new problems are often easier to verify than the original, and for this reason such an effort is significant. It is important to note that these new problems can be solved using any other RT analysis technique, such as those of Sistla *et al.* [21], since the reduction does not assume anything about the manner of analysis used.

Our experience with these reductions allows us to suggest a partial ordering when applying these reductions. We have found that applying *COI* first, followed by *Decompose*, followed by *COI* to be a highly effective reduction strategy. We apply *COI* first to remove as many obvious statements as possible. This sometimes removes statements that impede the use of *Decompose*. We finish with the *COI* reduction to remove statements that no longer influence the membership of the queried roles due to the removal of other statements using another reduction. As such, applying *Chain* followed by *COI* is another effective and straightforward strategy. These partial orderings of reductions are always applied before constructing the *MRPS*.

4.3.2 Restriction Relaxation

Another interesting strategy for simplifying verification involves relaxing the restriction rule. Observe that given an instance of the RCP such as $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$, if one or more roles in $\mathcal{R}$ are relaxed (removed completely from $\mathcal{R}$) producing $\mathcal{R}'$ and $\langle \mathcal{P}, \mathcal{R}', X.u \sqsupseteq A.r \rangle$ is satisfied, then $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$ is satisfied. Clearly any statements defining a role that is not restricted can be removed by URR. This strategy, called *Restriction Relaxation*, effectively under-specifies the policy and should be applied after the pre-processing reductions but prior to construction of the *MRPS*. Based on our understanding and experience with this problem, it is our conjecture that this technique is sound but not complete.

Initial Policy $\mathcal{P}$	
$A.r \leftarrow B.r_1 \cap C.r_2$	
$B.r_1 \leftarrow D.r_3.r_4$	
$C.r_2 \leftarrow E.r_5.r_4$	
$F.r_6 \leftarrow D.r_3 \cap E.r_5$	$\mathcal{G}_\mathcal{R} = \{A.r, B.r_1, C.r_2, F.r_6, X.u\}$
$X.u \leftarrow F.r_6.r_4$	$\mathcal{S}_\mathcal{R} = \{A.r, B.r_1, C.r_2, F.r_6, X.u\}$
$X.u \leftarrow D.r_3$	
$X.u \leftarrow E.r_5$	RCPI: $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$

New Principals: P_1, P_2

New Model Statements	
$D.r_3 \leftarrow P_1$	$P_1.r_4 \leftarrow P_1$
$D.r_3 \leftarrow P_2$	$P_1.r_4 \leftarrow P_2$
$E.r_5 \leftarrow P_1$	$P_2.r_4 \leftarrow P_1$
$E.r_5 \leftarrow P_2$	$P_2.r_4 \leftarrow P_2$

a. Principal Abstraction 2

New Principals: P_1, P_2, P_3

New Statements
$D.r_3 \leftarrow P_1$
$E.r_5 \leftarrow P_2$
$P_1.r_4 \leftarrow P_3$
$P_2.r_4 \leftarrow P_3$

b. Counterexample

Figure 7: Incompleteness of Principal Abstraction

4.3.3 Principal Abstraction

In many cases the number of new principals added to the *MRPS* makes the model difficult to analyze since not only are there significantly more policy statements to consider, but the maximum membership size of each role increases. This is problematic since the state space of the AFSM depends on both the number of subsets of policy statements and the number of subsets of principals in each role. One technique for analyzing such policies is to under-approximate the *MRPS* by adding fewer than the conservative number of new principals. Recall that the conservative number of new principals was described in Section 3.1. Any violation of the query detected by the model checker is clearly valid since a counterexample is produced. However upon verification, it is indeterminant whether the result is a false positive or not. A false positive could occur in a situation where additional new principals may be necessary to expose a counterexample. We call this technique *Principal Abstraction*, and it is applied during *MRPS* construction.

Figure 7 demonstrates how the Principal Abstraction can interfere with finding all counterexamples. In this example, it is necessary to add three new principals in order to expose a counterexample. When we apply our technique and limit the number of new principals to two, we see that no subset of new model statements will produce a counterexample. In this case, the problem has to do with an insufficient number of unique sub-linked roles being created. In our experience of constructing and evaluating policies for role containment by hand, it is often the case that we need one new principal for each unique linked role expression and one new principal to serve as the witness. However, in many cases where linked role names are not shared among linked role expressions, two

new principals are sufficient to expose counterexamples.

4.3.4 Chain Reduction

Chain Reduction is applied as a set of constraints on the transition relation of the AFSM and implemented as a set of `TRANS` statements in SMV. Specifically, each constraint takes the form of a condition that forces certain policy statements to be absent if other policy statements also happen to be absent, thus forcing sets of states to collapse to a single *representative state*. The relationship between such policy statements is identified by examining how each statement defines or depends on a given role. For each role, we identify the statements that define the role (*defining statements*) as well as the statements that depend on the role (*dependency statements*). If all defining statements are absent from a given role, then we may also remove all dependency statements for that role as well. In SMV, `TRANS` statements are explicit transition relation statements that overlay the existing transition relation. By incorporating these constraints into the transition relation, only representative states are examined.

4.4 Characteristics of Difficult Problems

Clearly we cannot expect that all role containment problems will become tractable through the use of these techniques, however examination of such techniques is a significant contribution since it helps identify the characteristics of the more difficult problems. In particular, policies in which the queried roles are in a cycle and each role in the cycle is restricted in some way can be very difficult to reduce and evaluate. It is not difficult to understand that COI, URR, and Decomposition most likely will not be effective in such a case due to a significant overlap of restricted *DefRoles* of the queried roles. A second characteristic of difficult role containment problems has to do with a large number of principals introduced into the policy. This can occur in the *MRPS* through the introduction of new principals, but also in policies that happen to define restricted roles with many simple member statements. In the first case, we may be able to use principal abstraction to address part of the problem, but in the second case we will need a new abstraction technique. We suspect that such an abstraction may be possible as a result of future work. Finally, we reiterate a policy characteristic from Section 4.3.3 that suggests we may need many new principals in cases where the policy contains linked role expressions that share linked role names. For instance, if the policy were to contain the statements $A.r \leftarrow B.r_1.r_2$ and $X.u \leftarrow C.r_3.r_2$, then we may need sufficient principals to construct unique sub-linked roles using the linked role name r_2 for each linked role expression just in case this is necessary to expose a counterexample.

5. RELATED WORK

Security analysts of access control systems and policies have increasingly leveraged automated tool support to verify properties in support of security objectives. Jha et al. [10] verify such properties as authorization, availability, and shared access of the SPKI/SDSI policy language through the use of a language containment type of model checking. SPKI/SDSI is related to RT in that it can be reduced to a subset of the language, specifically the policy class $RT[\leftarrow]$, which does not support intersection inclusion statements. It is similar to our approach in that it is able to reason about policies that are unbounded in size, as well as describing specifications in a temporal logic.

Sistla et. al. [20, 21] provides a framework for reasoning about security analysis of dynamic RT policies. Of significant value is their proof of a tight EXPTIME complexity for role containment queries. Their approach focuses on using language containment to determine whether the membership of one role is contained in another role across policy states. It is unclear whether counterexamples can be constructed from this approach since no new principals are added to the evaluated policy.

Fisler et al. [4] analyzes the impact of policy changes on role-based access control (RBAC) systems using their Margrave tool. Such policies are represented as multi-terminal BDD's for efficient storage and manipulation. Unlike dynamic policy analysis which reasons about whether properties will hold despite future policy changes from unknown entities, their work focuses on specific policy changes by security administrators. They verify separation of duty properties, and they are able to produce counterexamples when a property is not satisfied.

Schaad et al. [19] also verifies separation of duty properties in RBAC systems, but uses a mature model checking tool called NuSMV. NuSMV is a model checker of the SMV family and supports additional features such as bounded model checking and step-wise exploration of the reachable state space. Their version of RBAC supports limited delegation and revocation of user permissions, though it does not support delegation of authorization as seen in modern trust management languages.

Other work [3, 5, 6, 7, 8, 13, 22, 23] also supports the use of formal tools to verify properties of security policies, however none develop a framework applicable to RT role containment problems.

6. CONCLUSIONS

Crafting and validating access control policies that reflect the author's intention is a difficult task for several reasons. First, stakeholders are generally neither security nor formal tool experts. They require an intuitive means of reasoning about the protection of their resources. Additionally, they require insightful counterexamples to describe why their policies fail to meet their expectations. Second, the complexity of large policies and intractability of certain cases makes manual security analysis an unreasonable approach. Thus tools that automate verification are clearly necessary.

Our contribution addresses these concerns through a collection of policy reduction and optimization techniques. Such techniques can be associated with automated model checking techniques, which always produce a counterexample when the policy fails to satisfy a property. This provides stakeholders the ability to verify properties without necessarily being experts in formal tools. Our contribution also includes proofs demonstrating correctness of our techniques and analysis of the policy characteristics that remain difficult to verify. We propose a more thorough examination of principal abstraction and restriction relaxation techniques and proofs on correctness as future work.

Acknowledgements

William H. Winsborough is supported in part by NSF awards CCR-0325951, CCF-0524010, and CNS-0716750,

and Texas Advanced Research Program award ARP 010115-0037-2007. Jianwei Niu is supported in part by Texas Advanced Research Program award ARP 010115-0037-2007, and University of Texas at San Antonio research award TRAC-2008.

7. REFERENCES

[1] E. M. Clarke, E. A. Emerson, and A. P. Sistla. Automatic verification of finite-state concurrent systems using temporal logic specifications. *ACM TOPLAS*, 8(2):244–263, 1986.

[2] E. M. Clarke, O. Grumberg, and D. A. Peled. *Model Checking*. The MIT Press, 1999.

[3] M. Drouineaud, M. Bortin, P. Torrini, and K. Sohr. A first step towards formal verification of security policy properties for RBAC. In *Proceedings of Fourth International Conference on Quality Software*, pages 60–67, 2004.

[4] K. Fisler, S. Krishnamurthi, L. A. Meyerovich, and M. C. Tshchantz. Verification and change-impact analysis of access-control policies. In *ICSE*, pages 196–205. ACM Press, 2005.

[5] D. Gilliam, J. Powell, and M. Bishop. Application of lightweight formal methods to software security. In *WETICE*, 2005.

[6] P. Giorgini, F. Massacci, J. Mylopoulos, and N. Zannone. Modeling security requirements through ownership, permission and delegation. In *RE*, pages 167–176. IEEE Computer Society, 2005.

[7] D. P. Guelev, M. Ryan, and P. Y. Schobbens. Model checking access control policies. In *Proceedings of the 7th Information Security Conference*, volume 3225 of *LNCS*. Springer-Verlag, 2004.

[8] F. Hansen and V. Oleshchuk. Conformance checking of RBAC policy and its implementation. In *Information Security Practice and Experience*, volume 3439 of *LNCS*, pages 144–155. Springer Berlin/Heidelberg, 2005.

[9] M. A. Harrison, W. L. Ruzzo, and J. D. Ullman. Protection in operating systems. *Commun. ACM*, 19(8):461–471, 1976.

[10] S. Jha and T. Reps. Model checking SPKI/SDSI. *Journal of Computer Security*, 12:317–353, 2004.

[11] N. Li, J. C. Mitchell, and W. H. Winsborough. Design of a role-based trust management framework. In *SSP*, pages 114–130. IEEE Computer Society Press, May 2002.

[12] N. Li, J. C. Mitchell, and W. H. Winsborough. Beyond proof-of-compliance: Security analysis in trust management. *JACM*, 52(3):474–514, 2005.

[13] M. J. May, C. A. Gunter, and I. Lee. Privacy APIs: Access control techniques to analyze and verify legal privacy policies. In *CSFW*, 2006.

[14] K. McMillan. *Symbolic Model Checking: An Approach to the State Explosion Problem*. Kluwer Academic, 1993.

[15] A. Pnueli. The temporal logic of programs. In *Proceedings of the 18th IEEE Symposium on Foundations of Computer Science*, volume 526, pages 46–67, 1977.

[16] M. Reith, J. Niu, and W. H. Winsborough. Apply model checking to security analysis in trust management. In *Proceedings of the First International Workshop on Security Technologies for Next Generation Collaborative Business Applications*, 2007.

[17] M. Reith, J. Niu, and W. H. Winsborough. Reductions and optimizations for the analysis of trust management policy. Technical Report CS-TR-2008-017, UTSA, 2008.

[18] R. Sandhu. The typed access matrix model. In *Symposium on Research in Security and Privacy*, pages 122–136. IEEE Computer Society, 1992.

[19] A. Schaad, V. Lotz, and K. Sohr. A model-checking approach to analysing organisational controls in a loan origination process. In *SACMAT '06: Proceedings of the eleventh ACM symposium on Access control models and technologies*, pages 139–149, New York, NY, USA, 2006. ACM Press.

[20] A. P. Sistla and M. Zhou. Analysis of dynamic policies. In *Proceedings of Foundations of Computer Security and Automated Reasoning for Security Protocol Analysis*, pages 233–262, 2006.

[21] A. P. Sistla and M. Zhou. Analysis of dynamic policies. *Information and Computation*, 206(2-4):185–212, 2008.

[22] N. Zhang, M. Ryan, and D. P. Guelev. Synthesising verified access control systems in XACML. In *FMSE*, pages 56–65. ACM Press, 2004.

[23] N. Zhang, M. Ryan, and D. P. Guelev. Evaluating access control policies through model checking. In *Proceedings of the 8th Information Security Conference*, volume 3650 of *LNCS*, pages 446–460. Springer-Verlag, 2005.

APPENDIX

A. REDUCTION PROOFS

This section provides proofs for the theorems presented in this paper. We begin by describing information necessary for understanding the proofs, followed by the proofs themselves.

A.1 Background

LEMMA 5 (LMW05). *Given $\mathcal{P}$ and $\mathcal{R}$, two roles $X.u$ and $A.r$ in Roles($\mathcal{P}$), if $X.u$ does not contain $A.r$, then there exists a $\mathcal{P}'$ reachable from $\mathcal{P}$ and principal E such that $E \in [\![A.r]\!]_{\mathcal{P}'}$, $E \notin [\![X.u]\!]_{\mathcal{P}'}$, $\mathcal{P}' - \mathcal{P}$ has only simple member statements, and $\mathcal{P}'$ uses only roles names in $\mathcal{P}$.*

This lemma tells us that without loss of generality, we can consider only those $\mathcal{P}'$ reachable from $\mathcal{P}$ such that $\mathcal{P}' - \mathcal{P}$ consists solely of simple member statements and uses no role names other than those occurring in $\mathcal{P}$. The intuition behind this lemma can be understood by recognizing that for any $\mathcal{P}''$ reachable from $\mathcal{P}$ where $E \in [\![A.r]\!]_{\mathcal{P}''}$ and $E \notin [\![X.u]\!]_{\mathcal{P}''}$, the set of statements $\mathcal{P}'' - \mathcal{P}$ must consist of the set of statements $\{\lambda \leftarrow e \mid \lambda \notin \mathcal{G}_{\mathcal{R}}\}$, and as such $\mathcal{P}'' - \mathcal{P}$ may be rewritten as a set of simple member statements $\{\lambda \leftarrow B \mid \lambda \notin \mathcal{G}_{\mathcal{R}} \wedge B \in [\![\lambda]\!]_{\mathcal{P}''}\}$. Let $\mathcal{P}' = \mathcal{P} \cup \{\lambda \leftarrow B \mid \lambda \notin \mathcal{G}_{\mathcal{R}} \wedge B \in [\![\lambda]\!]_{\mathcal{P}''}\}$. Clearly $\mathcal{P}'$ is semantically equivalent to $\mathcal{P}$, $\mathcal{P}' - \mathcal{P}$ consists solely of simple member statements, and $\mathcal{P}'$ is reachable from $\mathcal{P}$.

A.2 Cone of Influence Reduction

We use a variant of *DefRoles* that includes roles in $\mathcal{M}$, but does not include roles used in their definitions.

DEFINITION 9. *Let Λ and $\mathcal{M}$ be sets of roles, and $\mathcal{P}$ be a policy. We define $DefRoles^+(\mathcal{P}, \Lambda, \mathcal{M})$ to be the least set of roles $\mathcal{O}$ satisfying the following conditions:*

- $\Lambda \subseteq \mathcal{O}$
- $(\lambda \in \mathcal{O} \ \wedge \ \lambda \notin \mathcal{M} \ \wedge \ \lambda \leftarrow B.r_1 \in \mathcal{P}) \Rightarrow B.r_1 \in \mathcal{O}$
- $(\lambda \in \mathcal{O} \ \wedge \ \lambda \notin \mathcal{M} \ \wedge \ \lambda \leftarrow B.r_1.r_2 \in \mathcal{P} \wedge D \in Principals) \Rightarrow (B.r_1 \in \mathcal{O} \wedge D.r_2 \in \mathcal{O})$
- $(\lambda \in \mathcal{O} \wedge \lambda \notin \mathcal{M} \wedge \lambda \leftarrow B.r_1 \cap C.r_2 \in \mathcal{P}) \Rightarrow (B.r_1 \in \mathcal{O} \wedge C.r_2 \in \mathcal{O})$

Proof: There are two parts: "If" and "Only if".

"Only if": Let $\mathcal{P}'''$ be reachable from $\mathcal{P}'$ under $\mathcal{R}$ and let E satisfy $E \notin [\![X.u]\!]_{\mathcal{P}'''} \wedge E \in [\![A.r]\!]_{\mathcal{P}'''}$. Let $\mathcal{P}'' = \mathcal{P}''' \cup \mathcal{P}\restriction_{\mathcal{S}_\mathcal{R}}$.

We show $\mathcal{P}''$ is reachable from $\mathcal{P}$ under $\mathcal{R}$. Consider any $\lambda \leftarrow e \in \mathcal{P} - \mathcal{P}''$. We know that $\lambda \notin \mathcal{S}_\mathcal{R}$ because $\mathcal{P}\restriction_{\mathcal{S}_\mathcal{R}} \subseteq \mathcal{P}''$, so removing $\lambda \leftarrow e$ is permitted by $\mathcal{R}$. Now consider any $\lambda \leftarrow e \in \mathcal{P}'' - \mathcal{P}$. By construction of $\mathcal{P}''$ it must be that $\lambda \leftarrow e \in \mathcal{P}'''$, which is reachable from $\mathcal{P}'$. Since $\mathcal{P}' \subseteq \mathcal{P}$, it follows that $\lambda \leftarrow e \in \mathcal{P}''' - \mathcal{P}'$, which implies that $\lambda \notin \mathcal{G}'_\mathcal{R}$. Since $\mathcal{G}_\mathcal{R} \subseteq \mathcal{G}'_\mathcal{R}$, $\lambda \leftarrow e$ can be added to $\mathcal{P}$, as required.

Next we show $E \notin [\![X.u]\!]_{\mathcal{P}''}$ and $E \in [\![A.r]\!]_{\mathcal{P}''}$. The latter follows easily from $E \in [\![A.r]\!]_{\mathcal{P}'''}$ and from $\mathcal{P}''' \subseteq \mathcal{P}''$, the latter of which tells us that $[\![A.r]\!]_{\mathcal{P}'''} \subseteq [\![A.r]\!]_{\mathcal{P}''}$.

We show $E \notin [\![X.u]\!]_{\mathcal{P}''}$ by proving that for all roles $B.r' \in DefRoles(P, X.u, \mathcal{S}_\mathcal{R})$, $[\![B.r']\!]_{\mathcal{P}''} \subseteq [\![B.r']\!]_{\mathcal{P}'''}$. For this we use induction on i to show that if $m(B, r', D) \in T_{\mathcal{P}''} \uparrow^i$, then $m(B, r', D) \in T_{\mathcal{P}'''} \uparrow^j$, for some $j \in \mathcal{N}$. The base case is trivial. For the step, consider any $m(B, r', D) \in T_{\mathcal{P}''} \uparrow^{i+1}$ and fix the statement $B.r' \leftarrow e$ that is used to introduce $m(B, r', D) \in T_{\mathcal{P}''} \uparrow^{i+1}$. There are four cases based on the struture of e.

Case $e = D$**:** Because $B.r' \in DefRoles(\mathcal{P}, X.u, \mathcal{S}_\mathcal{R})$, it follows that $B.r' \leftarrow D \in \mathcal{P}'$. If $B.r' \in \mathcal{S}_\mathcal{R}$, then $B.r' \leftarrow D \in \mathcal{P}'''$ by definition f reachability. If not, then $B.r' \leftarrow D \in \mathcal{P}'''$ follws by construction of $\mathcal{P}''$. In either case we have $m(B, r', D) \in T_{\mathcal{P}'''} \uparrow^1$.

Case $e = C.r_1$**:** We obtain $B.r' \leftarrow C.r_1 \in \mathcal{P}'''$ as we did $B.r' \leftarrow D \in \mathcal{P}'''$ in the first case. It follows that $m(C, r_1, D) \in T_{\mathcal{P}''} \uparrow^i$. We show that there is a j for which $m(C, r_1, D) \in T_{\mathcal{P}'''} \uparrow^j$.

By definition of *DefRoles*, $B.r' \in DefRoles(\mathcal{P}, X.u, \mathcal{S}_\mathcal{R})$ implies either $C.r_1 \notin \mathcal{S}_\mathcal{R}$ or $C.r_1 \in DefRoles(\mathcal{P}, X.u, \mathcal{S}_\mathcal{R})$. When $C.r_1 \in DefRoles(\mathcal{P}, X.u, \mathcal{S}_\mathcal{R})$, we obtain $m(C, r_1, D) \in T_{\mathcal{P}'''} \uparrow^j$ from the induction assumption. When $C.r_1 \notin DefRoles(\mathcal{P}, X.u, \mathcal{S}_\mathcal{R})$, it follows that there is no statement defining $C.r_1$ in $\mathcal{P}'$. Thus it must be that $C.r_1 \notin \mathcal{G}_\mathcal{R}$ and that $C.r_1 \leftarrow D\mathcal{P}'' - \mathcal{P}$. (Recall that by Lemma 5 we can assume without loss of generality that $\mathcal{P}'' - \mathcal{P}$ consists of simple member statements.) Since $C.r_1 \notin \mathcal{S}_\mathcal{R}$, it follows from the construction of $\mathcal{P}''$ that $C.r_1 \leftarrow D\mathcal{P}'''$. Therefore $m(C, r_1, D) \in T_{\mathcal{P}'''} \uparrow^1$.

We now obtain $m(B, r', D) \in T_{\mathcal{P}'''} \uparrow^{j+1}$, as required.

The cases in which $e = C.r_1.r_2$ and $e = C.r_1 \cap F.r_2$ are similar to the second case above.

"If": Let $\mathcal{P}''$ be reachable from $\mathcal{P}$ and let E satisfy $E \notin [\![X.u]\!]_{\mathcal{P}''} \wedge E \in [\![A.r]\!]_{\mathcal{P}''}$.

It is convenient in this case to construct $\widehat{\mathcal{P}}$ from $\mathcal{P}''$ which is also reachable from $\mathcal{P}$ and satisfies $E \notin [\![X.u]\!]_{\widehat{\mathcal{P}}} \wedge E \in [\![A.r]\!]_{\widehat{\mathcal{P}}}$.

Let $\mathcal{P}_0 = \{F.r'' \leftarrow D \mid F.r'' \in DefRoles^+(P, A.r, \mathcal{G}_\mathcal{R}) - \mathcal{G}_\mathcal{R} \ \wedge \ D \in [\![F.r]\!]_{\mathcal{P}''}\}$. Let $\widehat{\mathcal{P}} = (\mathcal{P}'' - \mathcal{P}''\restriction_{DefRoles^+, X.u, \mathcal{S}_\mathcal{R}) - (\mathcal{S}_\mathcal{R} \cup DefRoles(P, A.r, \mathcal{G}_\mathcal{R}))}) \cup \mathcal{P}_0$.

It can be shown by induction that for all $B.r' \in DefRoles(\mathcal{P}, X.u, \mathcal{S}_\mathcal{R})$, $[\![B.r']\!]_{\widehat{\mathcal{P}}} \subseteq [\![B.r']\!]_{\mathcal{P}''}$ and that for all $B.r' \in DefRoles(\mathcal{P}, A.r, \mathcal{G}_\mathcal{R})$, $[\![B.r']\!]_{\widehat{\mathcal{P}}} = [\![B.r']\!]_{\mathcal{P}''}$. It follows from this that $E \notin [\![X.u]\!]_{\widehat{\mathcal{P}}} \wedge E \in [\![A.r]\!]_{\widehat{\mathcal{P}}}$.

Now we construct $\mathcal{P}'''$ reachable from $\mathcal{P}'$ by $\mathcal{P}'''\widehat{\mathcal{P}} \ \restriction_{M(\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle)} \cup \mathcal{P}_0$, in which $M(\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle) = DefRoles(\mathcal{P}, X.u, \mathcal{S}_\mathcal{R}) \cup DefRoles(\mathcal{P}, A.r, \mathcal{G}_\mathcal{R})$.

It can be shown that $\mathcal{P}'''$ is reachable from $\mathcal{P}'$.

It can now be shown that $[\![B.r']\!]_{\mathcal{P}'''} \subseteq [\![B.r']\!]_{\widehat{\mathcal{P}}}$ for all $B.r'$ defined in $\mathcal{P}'''$. From this we get $E \notin [\![X.u]\!]_{\mathcal{P}'''}$.

To show that $E \in [\![A.r]\!]_{\mathcal{P}'''}$, one can use induction on i to show that for all roles $B.r' \in M(\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle)$, if $m(B, r', D) \in T_{\widehat{\mathcal{P}}} \uparrow^i$ then there is a $j \in \mathcal{N}$ such that $m(B, r', D) \in T_{\mathcal{P}'''} \uparrow^j$.

While we have completed this proof, space constraints prevent our including it here.

A.3 Decomposition Reduction

Proof:

We prove entailment in both directions.

We begin by showing the "only if" part, which we do by showing its contrapositive. That is, we assume there is a $\langle \mathcal{P}', \mathcal{R}', X.u \sqsupseteq \rho \rangle \in Decompose(\mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r)$ such that in some state $\mathcal{P}''$ reachable from $\mathcal{P}'$ under $\mathcal{R}'$, there is a principal E such that $E \in [\![\rho]\!]_{\mathcal{P}''}$, but $E \notin [\![X.u]\!]_{\mathcal{P}''}$. Using this assumption, we show $\mathcal{P}$ does not satisfy $X.u \sqsupseteq A.r$ under $\mathcal{R}$. There are now four cases given by the type of statement $A.r \leftarrow \alpha \in \mathcal{P}$ that is used to construct $\langle \mathcal{P}', \mathcal{R}', X.u \sqsupseteq \rho \rangle$.

Case 1: α is a principal. In this case $\rho = A'.r'$ and $[\![\rho]\!]_{\mathcal{P}''} = \{E\}$. Given the relationship between $\mathcal{R}$ and $\mathcal{R}'$, it is not hard to see that $\mathcal{P}''' = (\mathcal{P}'' - \{A'.r' \leftarrow E\})$ is reachable from $\mathcal{P}$ under $\mathcal{R}$. By the construction in *Decompose* in this case, $A.r \leftarrow E \in \mathcal{P}$. Since $A.r \in \mathcal{S}_\mathcal{R}$, it follows that $A.r \leftarrow E \in \mathcal{P}'''$ and thus $E \in [\![A.r]\!]_{\mathcal{P}'''}$. Because $E \notin [\![X.u]\!]_{\mathcal{P}''}$, it follows by the monotonic nature of the semantics that $E \notin [\![X.u]\!]_{\mathcal{P}'''}$, completing the proof in this case.

Case 2: α is a role. In this case $\rho = \alpha$ and $E \in [\![\rho]\!]_{\mathcal{P}''}$ for some $\mathcal{P}''$ reachable from $\mathcal{P}$ under $\mathcal{R}$. By the construction in *Decompose*, in this case $E \in [\![A.r]\!]_{\mathcal{P}''}$ since $A.r \leftarrow \alpha \in \mathcal{P}''$ because, as above, $A.r \leftarrow \alpha \in \mathcal{P}$ and $A.r \in \mathcal{S}_\mathcal{R}$.

Case 3: α is a linked role of the form $B.r_1.r_2$. In this case $\rho = A'.r'$ and $E \in [\![\rho]\!]_{\mathcal{P}''}$. Given the relationship between $\mathcal{R}$ and $\mathcal{R}'$, it is not hard to see that $\mathcal{P}''' = (\mathcal{P}'' - \{A'.r' \leftarrow B.r_1.r_2\})$ is reachable from $\mathcal{P}$ under $\mathcal{R}$. By the construction in *Decompose*, $A'.r' \leftarrow \alpha$ is the only statement defining $A'.r'$ in $\mathcal{P}''$. Thus there exist some $C \in [\![B.r_1]\!]_{\mathcal{P}'''}$ such that $E \in [\![C.r_2]\!]_{\mathcal{P}'''}$. We also have $A.r \leftarrow B.r_1.r_2 \in \mathcal{P}'''$ because $A.r$ is shrink restricted so $E \in [\![A.r]\!]_{\mathcal{P}'''}$. Because $E \notin [\![X.u]\!]_{\mathcal{P}''}$, it follows by the monotonic nature of the semantics that $E \notin [\![X.u]\!]_{\mathcal{P}'''}$, completing the proof in this case.

Case 4: α is the intersection of two roles $B.r_1$ and $C.r_2$. In this case $\rho = A'.r'$ and $E \in [\![\rho]\!]_{\mathcal{P}''}$. Given the relationship between $\mathcal{R}$ and $\mathcal{R}'$, it is not hard to see that $\mathcal{P}''' = (\mathcal{P}'' - \{A'.r' \leftarrow B.r_1 \cap C.r_2\})$ is reachable from $\mathcal{P}$ under $\mathcal{R}$. By the construction in *Decompose*, $A'.r' \leftarrow \alpha$

is the only statement defining $A'.r'$ in $\mathcal{P}''$. Thus it must be the case that $E \in [\![B.r_1]\!]_{\mathcal{P}'''} \cap [\![C.r_2]\!]_{\mathcal{P}'''}$. We also have $A.r \leftarrow B.r_1 \cap C.r_2 \in \mathcal{P}'''$ because $A.r$ is shrink restricted so $E \in [\![A.r]\!]_{\mathcal{P}'''}$. Because $E \notin [\![X.u]\!]_{\mathcal{P}''}$, it follows by the monotonic nature of the semantics that $E \notin [\![X.u]\!]_{\mathcal{P}'''}$, completing the proof in this the last case.

Now we show that entailment holds in the other direction, again by showing the contrapositive. We assume $\mathcal{P}''$ is reachable from $\mathcal{P}$ under $\mathcal{R}$ and there exists some principal E such that $E \in [\![A.r]\!]_{\mathcal{P}''}$ and $E \notin [\![X.u]\!]_{\mathcal{P}''}$. Using this assumption, we show there exists at least one $\langle \mathcal{P}', \mathcal{R}', X.u \sqsupseteq \rho \rangle \in \mathit{Decompose}(\mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r)$ where $\mathcal{P}'$ does not satisfy $X.u \sqsupseteq \rho$ under $\mathcal{R}'$.

Because $E \in [\![A.r]\!]_{\mathcal{P}''}$ it must be the case that $m(A, r, E) \in T_{\mathcal{P}''} \uparrow^\omega$, which means that (at least) one of the disjuncts in Definition 1 of $T_{\mathcal{P}''}(\pi)$ must hold when π is taken to be $T_{\mathcal{P}''} \uparrow^\omega$. Recall that $T_{\mathcal{P}''} \uparrow^\omega = T_{\mathcal{P}''}(T_{\mathcal{P}''} \uparrow^\omega)$. Thus we have four cases, one for each disjunct.

Case 1: $m(A, r, E) \in T_{\mathcal{P}''} \uparrow^\omega$ is generated by $A.r \leftarrow E \in \mathcal{P}''$. Since $A.r \in \mathcal{G}_\mathcal{R}$ it follows that $A.r \leftarrow E \in \mathcal{P}$. Thus there exists $\langle \mathcal{P}', \mathcal{R}', X.u \sqsupseteq A'.r' \rangle \in \mathit{Decompose}(\mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r)$ in which $\mathcal{P}' = \mathcal{P} \cup \{A'.r' \leftarrow E\}$ and $\mathcal{R}' = \langle \mathcal{G}_\mathcal{R} \cup \{A'.r'\}, \mathcal{S}_\mathcal{R} \cup \{A'.r'\} \rangle$. Since $\mathcal{P}''$ is reachable from $\mathcal{P}$ under $\mathcal{R}$, it is easy to check that $\mathcal{P}''' = \mathcal{P}'' \cup A'.r' \leftarrow E$ is reachable from $\mathcal{P}'$ under $\mathcal{R}'$, and that $[\![A'.r']\!]_{\mathcal{P}'''} = \{E\}$. A simple induction on i shows that for all principals B_1 and B_2, and for all role names r_3, such that $B_1 \neq A'$ and $r_3 \neq r'$, $m(B_1, r_3, B_2) \in T_{\mathcal{P}''} \uparrow^i$ if and only if $m(B_1, r_3, B_2) \in T_{\mathcal{P}'''} \uparrow^i$. From this we get $E \notin [\![X.u]\!]_{\mathcal{P}'''}$ as required.

Case 2: $m(A, r, E) \in T_{\mathcal{P}''} \uparrow^\omega$ is generated by $A.r \leftarrow B.r_1 \in \mathcal{P}'' \wedge m(B, r_1, E) \in T_{\mathcal{P}''} \uparrow^\omega$. In this case we have $E \in [\![B.r_1]\!]_{\mathcal{P}''}$ and $E \notin [\![X.u]\!]_{\mathcal{P}''}$. By hypothesis $\mathcal{P}''$ is reachable from $\mathcal{P}$ under $\mathcal{R}$. Furthermore, by construction, $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq B.r_1 \rangle \in \mathit{Decompose}(\mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r)$, thus completing the proof in this case.

Case 3: $m(A, r, E) \in T_{\mathcal{P}''} \uparrow^\omega$ is generated by $A.r \leftarrow B.r_1.r_2 \in \mathcal{P}'' \wedge \exists Z.\ m(B, r_1, Z) \in T_{\mathcal{P}''} \uparrow^\omega \wedge m(Z, r_2, E) \in T_{\mathcal{P}''} \uparrow^\omega$. Fix such a Z. Since $A.r \in \mathcal{G}_\mathcal{R}$ it follows that $A.r \leftarrow B.r_1.r_2 \in \mathcal{P}$. Thus there exists $\langle \mathcal{P}', \mathcal{R}', X.u \sqsupseteq A'.r' \rangle \in \mathit{Decompose}(\mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r)$ in which $\mathcal{P}' = \mathcal{P} \cup \{A'.r' \leftarrow B.r_1.r_2\}$ and $\mathcal{R}' = \langle \mathcal{G}_\mathcal{R} \cup \{A'.r'\}, \mathcal{S}_\mathcal{R} \cup \{A'.r'\} \rangle$. Since $\mathcal{P}''$ is reachable from $\mathcal{P}$ under $\mathcal{R}$, it is easy to check that $\mathcal{P}''' = \mathcal{P}'' \cup A'.r' \leftarrow E$ is reachable from $\mathcal{P}'$ under $\mathcal{R}'$, and that $E \in [\![A'.r']\!]_{\mathcal{P}'''}$. A simple induction on i shows that for all principals B_1 and B_2, and for all role names r_3, such that $B_1 \neq A'$ and $r_3 \neq r'$, $m(B_1, r_3, B_2) \in T_{\mathcal{P}''} \uparrow^i$ if and only if $m(B_1, r_3, B_2) \in T_{\mathcal{P}'''} \uparrow^i$. From this we get $E \notin [\![X.u]\!]_{\mathcal{P}'''}$, $m(B, r_1, Z) \in T_{\mathcal{P}'''} \uparrow^\omega$, and $m(Z, r_2, E) \in T_{\mathcal{P}'''} \uparrow^\omega$. Since $A'.r' \leftarrow B.r_1.r_2 \in \mathcal{P}'''$, it follows that $m(A', r', E) \in T_{\mathcal{P}'''} \uparrow^\omega$, giving us $E \in [\![A'.r']\!]_{\mathcal{P}'''}$ as required.

Case 4: $m(A, r, E) \in T_{\mathcal{P}''} \uparrow^\omega$ is generated by $A.r \leftarrow B.r_1 \cap C.r_2 \in \mathcal{P}'' \wedge m(B, r_1, E) \in T_{\mathcal{P}''} \uparrow^\omega \wedge m(C, r_2, E) \in T_{\mathcal{P}''} \uparrow^\omega$. Since $A.r \in \mathcal{G}_\mathcal{R}$ it follows that $A.r \leftarrow B.r_1 \cap C.r_2 \in \mathcal{P}$. Thus there exists $\langle \mathcal{P}', \mathcal{R}', X.u \sqsupseteq A'.r' \rangle \in \mathit{Decompose}(\mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r)$ in which $\mathcal{P}' = \mathcal{P} \cup \{A'.r' \leftarrow B.r_1.r_2\}$ and $\mathcal{R}' = \langle \mathcal{G}_\mathcal{R} \cup \{A'.r'\}, \mathcal{S}_\mathcal{R} \cup \{A'.r'\} \rangle$. Since $\mathcal{P}''$ is reachable from $\mathcal{P}$ under $\mathcal{R}$, it is easy to check that $\mathcal{P}''' = \mathcal{P}'' \cup A'.r' \leftarrow E$ is reachable from $\mathcal{P}'$ under $\mathcal{R}'$, and that $E \in [\![A'.r']\!]_{\mathcal{P}'''}$. A simple induction on i shows that for all principals B_1 and B_2, and for all role names r_3, such that $B_1 \neq A'$ and $r_3 \neq r'$, $m(B_1, r_3, B_2) \in T_{\mathcal{P}''} \uparrow^i$ if and only if $m(B_1, r_3, B_2) \in T_{\mathcal{P}'''} \uparrow^i$. From this we get $E \notin [\![X.u]\!]_{\mathcal{P}'''}$, $m(B, r_1, E) \in T_{\mathcal{P}'''} \uparrow^\omega$, and $m(C, r_2, E) \in T_{\mathcal{P}'''} \uparrow^\omega$. Since $A'.r' \leftarrow B.r_1 \cap C.r_2 \in \mathcal{P}'''$, it follows that $m(A', r', E) \in T_{\mathcal{P}'''} \uparrow^\omega$, giving us $E \in [\![A'.r']\!]_{\mathcal{P}'''}$ as required.

Observe that *Decompose* may be applied iteratively to each new RCP and potentially constructing even more new RCP's. ■

A.4 Chain Reduction

Proof: The greatest solution to this pair of requirements can be computed iteratively as follows. Define $\mathcal{P}_0 = \mathcal{P}$ and for each $i \in \mathcal{N}$, define $\mathcal{P}_{i+1}$ to be the result of fixing some $B.r$ that satisfies the antecedent and that occurs in the body of some statement in $\mathcal{P}_i$ and defining $\mathcal{P}_{i+1}$ to be the result of removing that statement from $\mathcal{P}_i$. When no such statement remains, the result $\mathcal{P}'$ is returned.

The proof is by induction on i in the construction of $\mathcal{P}_i$. We show that for all $i \in \mathcal{N}$, $\langle \mathcal{P}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$ is satisfied if and only if $\langle \mathcal{P}_i, \mathcal{R}, X.u \sqsupseteq A.r \rangle$ is satisfied. Since $\mathcal{P}_0 = \mathcal{P}$, the basis is trivial. For the step, we assume the hypothesis for $\mathcal{P}_i$ and show that it holds for $\mathcal{P}_{i+1}$. Because we have the induction assumption, it suffices to show $\langle \mathcal{P}_i, \mathcal{R}, X.u \sqsupseteq A.r \rangle$ is satisfied if and only if $\langle \mathcal{P}_{i+1}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$ is satisfied.

The "if" part: We assume $\langle \mathcal{P}_i, \mathcal{R}, X.u \sqsupseteq A.r \rangle$ is not satisfied and show that $\langle \mathcal{P}_{i+1}, \mathcal{R}, X.u \sqsupseteq A.r \rangle$ is not satisfied. Consider any $\mathcal{P}'$ reachable from $\mathcal{P}_i$ and any E such that $E \in [\![A.r]\!]_{\mathcal{P}'}$ and $E \notin [\![X.u]\!]_{\mathcal{P}'}$. Let $\lambda \leftarrow e$ be the statement removed from $\mathcal{P}_i$ to obtain $\mathcal{P}_{i+1}$. Fix $B.r$ to be the role that appears in e and that satisfies $B.r \in \mathcal{G}_\mathcal{R} \wedge \neg \exists e'.\ B.r \leftarrow e' \in \mathcal{P}_i$. Let $\mathcal{P}'' = \mathcal{P}' - \{\lambda \leftarrow e\}$. Since $\mathcal{P}'$ is reachable from $\mathcal{P}_i$, clearly $\mathcal{P}''$ is reachable from $\mathcal{P}_{i+1}$. Since $B.r \in \mathcal{G}_\mathcal{R}$, $\forall D.\ m(B, r, D) \notin T_{\mathcal{P}_i} \uparrow^\omega$ for all $i \in \mathcal{N}$. So in no case can $\lambda \leftarrow e$ be used in the construction of any element in $T_{\mathcal{P}_i} \uparrow^\omega$. Using this observation, a simple induction on the steps of $T_{\mathcal{P}_i} \uparrow^j$ and $T_{\mathcal{P}_{i+1}} \uparrow^j$ shows that the semantics generated by $\mathcal{P}_i$ and $\mathcal{P}_{i+1}$ are identical.

For the "only if" part, we consider $\mathcal{P}''$ reachable from $\mathcal{P}_{i+1}$ and define $\mathcal{P}' = \mathcal{P}'' \cup \{\lambda \leftarrow e\}$. The proof the proceeds similarly. ■