

Modular Access Control via Strategic Rewriting

Daniel J. Dougherty¹, Claude Kirchner², H el ene Kirchner², Anderson Santana de Oliveira²

¹ Worcester Polytechnic Institute

² INRIA & LORIA*

Abstract. Security policies, in particular access control, are fundamental elements of computer security. We address the problem of authoring and analyzing policies in a modular way using techniques developed in the field of term rewriting, focusing especially on the use of rewriting strategies. Term rewriting supports a formalization of access control with a clear declarative semantics based on equational logic and an operational semantics guided by strategies. Well-established term rewriting techniques allow us to check properties of policies such as the absence of conflicts and the property of always returning a decision. A rich language for expressing rewriting strategies is used to define a theory of modular construction of policies, in which we can better understand the preservation of properties of policies under composition. The robustness of the approach is illustrated on the composition operators of XACML.

1 Introduction

Access control is at the heart of computer security. It has grown beyond mediating operating-system interactions between users and files and now plays a central role in web-based systems, privacy policies, and business rules. Accompanying these expanded applications of access control, our conception of the mechanism of authorization now goes beyond the classical model [28] of access-control matrices, and we view access control decisions as the embodiment of a set of rules. We call such a set of rules an access-control *policy*. Although monitoring and enforcement mechanisms are important aspects of the study of access control, the size and complexity of the systems being treated mean that the policies themselves are interesting software artifacts in their own right. They are sensitive to complex conditions on the policy environment, which represents the data that a program respecting the policy manipulates, such as attributes of subjects and resources and relations among these. They are not easy to get right.

In light of these considerations, it is now typical in large or complex systems to disentangle policy from application code. Policies are written in domain-specific, typically declarative languages, such as the industrial standard XACML [34], and reasoning about the correctness of policies is a subtle matter. It is common wisdom that a key to designing, reasoning about, and maintaining a large system is modularity, with corresponding attention to the mechanisms by which the models in a system interact.

* LORIA: UMR 7503 CNRS-INPL-INRIA-Nancy2-UHP; Nancy, France.

In this paper, we are interested in the question of building access-control policies in a modular fashion, and taking some initial steps towards a theory of how parts of a policy interact.

We propose *term rewriting* [6, 2] as a formalism for representing access control policies. Rewriting is a well-established paradigm whose applications include theoretical foundations for functional programming languages and theorem provers. It is flexible and expressive enough to capture a wide range of policy frameworks arising in practice and indeed it is a universal model of computation. It has a clean declarative semantics, based on equational logic. There is an active research community supporting efficient implementations and tools for reasoning about properties such as termination and confluence of rewrite systems. One can view rewrite systems as an intermediate language for policies; our thesis in this paper is that some of the more interesting aspects of reasoning about policies are profitably viewed in this context.

Indeed, rewriting is not a single formalism but rather a family of variations on a robust paradigm of directed equality. It is easy to see that simple term rewriting can capture policies such as Unix file-permissions rules, the richer setting of conditional rewriting is as rich as the language of Datalog explored by several authors (notably in trust-management research), and—as sketched below—core XACML policies can be captured by rewriting under strategy.

To give a flavor of how term rewriting can capture policy rules, we may consider the following rules, adapted from the XACML specification [34]: - A person, identified by his or her social security number, may read any record for which he or she is the designated patient:

$$req(patient(x), read, record(x)) \rightarrow permit.$$

Here *patient* names the function from patient numbers to patients as Subjects and *record* is a function from patient numbers to health records as Resources, while *read* is a constant symbol, of sort Actions.

The variable x is implicitly universally quantified, so that the rewriting above captures the generality of the access rule; and the repetition of the variable as a parameter has the effect of enforcing the binding between the patient and his record.

- An administrator shall not be permitted to write to medical elements of a patient record:

$$req(admin(x), write, record(y)) \rightarrow deny.$$

Here *any* administrator, named perhaps by his employee number, is denied write access to *any* health record: note the use of distinct variables in the rule. Also note the use of explicit *deny* as a decision. It is crucially important to modularity of policies that *deny* is not treated as the negation of *permit*: this will be further illustrated in the body of the paper.

- It is straightforward to capture certain notions of authorization hierarchy. For example, to say that subject s_2 inherits from subject s_1 all access rights involving resource r , it suffices to have the following rule in a policy:

$$req(s_1, x, r) \rightarrow req(s_2, x, r)$$

Here x is a variable ranging over actions. Note that this rule is a refinement of the type of inheritance typically incorporated into a Role-based Access Control Model (in

which one role may inherit all privileges from another, uniformly across all actions and resources).

In a large organization, there are many classes of “Subjects” with different needs for access to an immense variety of “Resources”. For example, in a hospital there are rules governing the access of patients to their health records, their financial records, and the like, while at the same time, there are rules for employee access to these same records as well as to resources quite different from health records. Meanwhile, other entities such as insurance carriers are subject to yet another set of rules for access to these data and more. The different constituencies (patients, staff, insurers) are almost certainly going to have somewhat different—even competing—requirements on their use of the data and place different emphases on the security goals (confidentiality, availability, integrity) of policies. It is natural to imagine that the sets of rules describing these various modes of access should not be authored and maintained in a single monolithic policy. In this setting, *the theory of composition of policies* becomes crucially important.

As a very simple example, imagine that rules for patient data access and rules for staff data access are composed in separate policy documents, \wp_p and \wp_s respectively. What should we say about the decision of \wp_p in the context of a request by an administrator to write a health record? Assuming \wp_p will not explicitly compute a decision (permit or deny) upon such a request, we must uniformly assume a *default* decision, perhaps default-deny, for all requests not handled directly. But this immediately leads to the conclusion that composing policies is something more subtle than taking their union. Consider by contrast a request by an administrator to read the next-of-kin information for a patient. A default-deny by \wp_p for this request would mean that when \wp_p was combined with \wp_s , which may explicitly compute a *permit* for this request, the resulting logical theory, taken in a naive sense, would be contradictory.

So at the very least, one must make a distinction between a policy decision which is computed in a “direct” way from the policy rules, and one taken as a default. The situation is even more interesting if two modules of a policy compute contradictory decisions: if the policy is to be coherent in practice there must be a principled way to combine the modules, a mechanism that lends itself to clean design and supports analysis and verification. The combination method we explore in this paper is that of *rewriting strategies*.

The remaining sections of this paper are organized as follows: we recall in Section 2 the main notions on rewrite rules and strategies used in this paper. In Section 3 we give the definition, suitable properties and examples of an access control policies expressed in the rewrite-based framework. We formalize policy composition in Section 4, its suitable properties, and we illustrate our approach on the composition operators of XACML. We discuss related and further works in Section 5.

2 Background

Basic definitions on term rewriting can be found in [6, 2]. Let us recall those which are used in the following. A many-sorted signature $(\mathcal{S}, \mathcal{F})$, or \mathcal{F} for short, is a set of sorts \mathcal{S} and a set of function symbols \mathcal{F} . Each $f \in \mathcal{F}$ has a profile $f : S_1 \times \dots \times S_n \rightarrow S$,

where $S_1, \dots, S_n, S \in \mathcal{S}$, and is associated to a natural number by the arity function ($\text{ar} : \mathcal{F} \rightarrow \mathbb{N}$). When $\text{ar}(f) = 0$, the function symbol f is called a constant.

$\mathcal{T}(\mathcal{F}, \mathcal{X})$ is the set of well-sorted terms built from a given finite set \mathcal{F} of function symbols and a denumerable set \mathcal{X} of variables. The set of variables occurring in a term t is denoted by $\mathcal{V}\text{ar}(t)$. If $\mathcal{V}\text{ar}(t)$ is empty, t is called a *ground term* and $\mathcal{T}(\mathcal{F})$ is the set of ground terms. For $f \in \mathcal{F}$, $f(\mathcal{T}(\mathcal{F}), \dots, \mathcal{T}(\mathcal{F}))$ denotes the set of ground terms with f as top symbol.

A *substitution* σ is an assignment from \mathcal{X} to $\mathcal{T}(\mathcal{F}, \mathcal{X})$, with a finite domain $\{x_1, \dots, x_k\}$ and written $\sigma = \{x_1 \mapsto t_1, \dots, x_k \mapsto t_k\}$.

A rewrite rule is an ordered pair of terms, denoted as $l \rightarrow r$, $l, r \in \mathcal{T}(\mathcal{F}, \mathcal{X})$, where l is not a variable and $\mathcal{V}\text{ar}(r) \subseteq \mathcal{V}\text{ar}(l)$ such that l and r belong to a same sort. The terms l and r are respectively called the left-hand side and the right-hand side of the rule. A rewrite system R on $\mathcal{T}(\mathcal{F}, \mathcal{X})$ is a (finite or infinite) set of rewrite rules. Rules can be labeled to easily distinguish among them. A rewrite rule $l \rightarrow r$ is a *collapsing rule* if r is a variable. It is a *duplicating rule* if there exists a variable that has more occurrences in r than in l . A function symbol which is not the top symbol of any rule in R is called a *constructor*. Other symbols are called *defined functions*.

Given a rewrite system R , a term t rewrites to a term t' , which is denoted $t \rightarrow_R t'$ if there exists a rewrite rule $l \rightarrow r$ of R , a position ω in t , a substitution σ , satisfying $t|_{\omega} = \sigma(l)$, such that $t' = t[\sigma(r)]_{\omega}$.

A rewriting derivation of the rewrite system R is any sequence of rewriting steps $t_1 \rightarrow_R t_2 \rightarrow_R \dots$. The *source* of such a derivation is t_1 . When the derivation is finite, its last term is called its *target*. R induces a derivability relation $\xrightarrow{*}_R$ on terms: $t \xrightarrow{*}_R t'$ if there exists a rewriting derivation from t to t' . If the derivation contains at least one step, it is denoted by $\xrightarrow{+}_R$. A rewrite system is *terminating* (or *strongly normalizing*) if all rewriting derivations are finite. A term t is *R-normalized* (or in *R-normal form*) when the empty derivation is the only one with source t ; a derivation is *normalizing* when its target is *R-normalized*. A rewrite system R is *weakly terminating* if every term t is the source of a normalizing derivation. It is *confluent* if for all terms t, u, v , $t \xrightarrow{*}_R u$ and $t \xrightarrow{*}_R v$ implies $u \xrightarrow{*}_R s$ and $v \xrightarrow{*}_R s$, for some s . When it is clear from the context, we may omit the index R .

The notion of strategy used in this paper is fundamental in rewriting, and we give here a general presentation of the main ideas. We adopt a general definition, slightly different from the one used in [6]: a *rewrite strategy* ζ for the rewrite system R is a subset of the set of all derivations of R . The *application of a strategy* ζ on a term t is denoted $[\zeta](t)$ and defined as the set of all targets t' of the derivations of source t in ζ . We denote $[\zeta](t)_{\downarrow}$ its subset that contains only the targets in normal form. The *domain* of a strategy is the set of terms that are source of a derivation in ζ . When no derivation in ζ has source t , we say that the strategy application on t fails. The result of the application of a failing strategy on a term t is the empty set.

In this paper, we will consider only strategies that are stable by concatenation (i.e. $t \xrightarrow{*}_R t' \in \zeta$ and $t' \xrightarrow{*}_R t'' \in \zeta$ implies $t \xrightarrow{*}_R t' \xrightarrow{*}_R t'' \in \zeta$). A strategy could be described by enumerating all its elements or more suitably by a *strategy language*. From elementary strategies expressions directly issued from a rewrite system R , more elaborated strategies expressions are built like in ELAN [24], Stratego [39], Tom [3, 33]

or more recently MAUDE [30]. The semantics of such a language is naturally described in the rewriting calculus [9, 10]. We describe below the main elements of the strategy language of interest in this paper.

Given a rewrite system R over $\mathcal{T}(\mathcal{F}, \mathcal{X})$, rewrite rules in R are elementary or atomic strategies. For instance, if a and b are constants, the application of the rewrite rule $a \rightarrow b$ to the term a is denoted $[a \rightarrow b](a)$ and evaluates to $\{b\}$.

A strategy expression ζ may take arguments ζ_1, \dots, ζ_n , and the resulting expression is expressed functionally: $\zeta(\zeta_1, \dots, \zeta_n)$. Notice that this is consistent with the notation $\zeta(R)$ as soon as the definition of ζ does not depend on its arguments order. When it is clear from the context, we identify the strategy expression and the strategy (i.e. the set of derivations it represents). In a consistent way, the application of a strategy expression to a term is defined as the application of the strategy it represents.

A simple strategy is the sequential application of two rules. It is described by the concatenation operator “seq”. For instance $[\text{seq}(l_1 \rightarrow r_1, l_2 \rightarrow r_2)](t)$ denotes $[l_2 \rightarrow r_2]([l_1 \rightarrow r_1](t))$. This strategy operator extends naturally to multiple arguments:

$$[\text{seq}(\zeta_1, \dots, \zeta_n)](t) = [\zeta_n]([\zeta_{n-1}]([\dots[\zeta_1](t)]))$$

Identity and failure are strategies easy to understand:

$$[\text{id}](t) = \{t\} \quad [\text{fail}](t) = \emptyset$$

The strategy computing all derivations obtained by application of a rewrite system R is called `universal`; it takes as argument the set of rules under consideration:

$$[\text{universal}(R)](t) = \{t' \mid t \xrightarrow{*}_R t'\}$$

For instance, we have:

$$\begin{aligned} [\text{universal}(a \rightarrow a)](a) &= \{a\} \\ [\text{universal}(f(x) \rightarrow f(f(x)))](f(a)) &= \{f(a), f(f(a)), f(f(f(a))), \dots\} \end{aligned}$$

One can successively try to apply several strategies using the `choice` operator: its first argument is applied if it does not fail, otherwise the second one is applied (and may fail too).

$$\begin{aligned} [\text{choice}(\zeta_1, \zeta_2)](t) &= [\zeta_1](t) \quad \text{if } [\zeta_1](t) \neq \emptyset \\ [\text{choice}(\zeta_1, \zeta_2)](t) &= [\zeta_2](t) \quad \text{if } [\zeta_1](t) = \emptyset \end{aligned}$$

Clearly `choice` is associative and therefore its syntax is extended to be applicable to a list of strategies:

$$\text{choice}(\zeta_1, \zeta_2, \dots, \zeta_n) = \text{choice}(\zeta_1, \text{choice}(\zeta_2, \dots, \zeta_n))$$

Other strategies allow controlling the application of rules over sub-terms of a term. The strategy `one` must succeed on at least one of the sub-terms of a term. On the other hand, `all` application must succeed on each sub-term, otherwise, the result is failure:

$$\begin{aligned} [\text{one}(\zeta)](f(t_1, \dots, t_n)) &= f(t_1, \dots, [\zeta](t_i), \dots, t_n), \text{ if } [\zeta](t_i) \neq \emptyset \\ [\text{all}(\zeta)](f(t_1, \dots, t_n)) &= f([\zeta](t_1), \dots, [\zeta](t_n)), \text{ if } \forall i \in \{1, \dots, n\}, [\zeta](t_i) \neq \emptyset \end{aligned}$$

Using the above set of operators, we can define recursive ones which iterate the application of a strategy to a term, for example:

$$\text{try}(\zeta) = \text{choice}(\zeta, \text{id}) \quad \text{repeat}(\zeta) = \text{try}(\text{seq}(\zeta, \text{repeat}(\zeta)))$$

It is worth noticing that `try` and `repeat` never fail. Other high-level strategies implement term traversal and normalization on terms and are well-known in the rewrite system literature:

$$\begin{aligned} \text{topDown}(\zeta) &= \text{seq}(\zeta, \text{all}(\text{topDown}(\zeta))) \\ \text{bottomUp}(\zeta) &= \text{seq}(\text{all}(\text{bottomUp}(\zeta)), \zeta) \\ \text{OnceTopDown}(\zeta) &= \text{choice}(\zeta, \text{one}(\text{OnceTopDown}(\zeta))) \\ \text{OnceBottomUp}(\zeta) &= \text{choice}(\text{one}(\text{OnceBottomUp}(\zeta)), \zeta) \\ \text{innermost}(\zeta) &= \text{repeat}(\text{onceBottomUp}(\zeta)) \\ \text{outermost}(\zeta) &= \text{repeat}(\text{onceTopDown}(\zeta)) \end{aligned}$$

Example 1. Some examples of strategy application are:

$$\begin{aligned} [\text{universal}(a \rightarrow b, a \rightarrow c)](a) &= \{a, b, c\} \\ [\text{choice}(a \rightarrow b, a \rightarrow c)](a) &= \{b\} \\ [\text{choice}(a \rightarrow c, a \rightarrow b)](b) &= \emptyset \\ [\text{try}(b \rightarrow c)](a) &= \{a\} \\ [\text{repeat}(\text{choice}(b \rightarrow c, a \rightarrow b))](a) &= \{c\} \end{aligned}$$

3 Rewrite-Based Policies

Recent developments in access control are aimed to express various constraints on the environment where policies run, in order to capture real world requirements from policy authors, such as time, location, and any other condition involving attributes of principals and objects.

In this context, it is important to embark expressive computational power in the definition of policies. As the notion of rule is quite natural in the context of policy specifications, we propose here a quite general definition of access control.

In our model, rewrite rules transform input terms representing access requests into access decision terms. In order to take the raw computational power of term rewriting and to enhance the agility of the policy specification language, we use strategies to explicitly control the rule application. We define rewrite-based policies as follows, where Q stands for queries (or requests) and D for decisions.

Definition 1 (Security Policy). *An access control security policy, \wp , is a 5-tuple $(\mathcal{F}, D, R, Q, \zeta)$ such that:*

1. \mathcal{F} is a signature;
2. D is a non-empty set of closed terms: $D \subseteq \mathcal{T}(\mathcal{F})$;
3. R is a set of rewrite rules over $\mathcal{T}(\mathcal{F}, \mathcal{X})$;
4. Q is a set of terms from $\mathcal{T}(\mathcal{F})$: $Q \subseteq \mathcal{T}(\mathcal{F})$;
5. ζ is a rewrite strategy for R .

Let us explain the main design choices made in this definition.

- First we consider that the policy specification and its environment are described as terms built over the signature \mathcal{F} . The set of possible decisions to be taken by the policy is denoted by D . Indeed, D is often a set of constants and the two main constants in D are usually *permit* and *deny*. But since it is crucial to model also policies that do not directly take decision, it can be useful to have a constant *not applicable* that simply expresses the fact that the current policy in the current context cannot decide about the access. Moreover, the result returned by a policy could be more elaborated than just a constant and can be in general a term containing further information like the time and duration the access is granted. What is significant is not treating the failure to derive a permission as a denial. In contrast to [20] for example, in which this later design is followed, we can treat explicitly decisions such as *deny* and *not applicable*. This is a crucial advantage for merging rules, since in purely logic-based works, there is no way to handle in the theory what happens when a policy which derives *deny* for a request q is merged with another which then derives *permit* explicitly, for the same q .
- The rewrite system R describes the behavior of the policy as well as some necessary computations which explain how its environment evolves. The role of the strategy is to point derivations of R whose interest is to produce decisions.
- The requests are a subset of terms. They typically express questions of the form: should a certain entity be granted access to a resource given the current configuration of the policy environment.
- The last component is the strategy which allows one to finely specify the evaluation order of the policy rules.

One of the main nice consequences of this approach, in addition to its expressivity, which we illustrate on the examples below, is that it allows us to take advantage of all the results obtained by the rewriting community since the last thirty years. Amongst such results, we investigate in this section confluence, termination and sufficient completeness.

Example 2. This example is taken from the NetFilter how-to³. Suppose an Internet user wants to set his firewall to block any traffic coming from the exterior to the local network. Since the interface associated to Internet connections is usually `ppp0`, a simple method is to reject all new packets coming from this source. In order to demonstrate the fact that it might be convenient for a policy to contain rules beyond those which *directly* compute decisions, we also give some additional rules which allow two different local computers to share the same external IP address: for each outgoing packet whose origin is a local machine, its head is rewritten to a single address.

- Let the policy signature be:

<i>pckt</i>	:	<i>Address</i> × <i>Address</i> × <i>State</i>	→	<i>Packet</i>
<i>filter</i>	:	<i>Packet</i>	→	<i>Decision</i>
<i>new, established</i>	:		→	<i>State</i>
<i>drop, accept</i>	:		→	<i>Decision</i>
<i>eth0, ppp0</i>	:		→	<i>Address</i>

³ <http://www.netfilter.org>

- The set of constant symbols representing decisions is $D = \{accept, drop\}$
- Consider R as the following rules, where $src, dst : Address$, and $s : State$ are variables:

$$\begin{array}{ll}
filter(pkt(src, dst, established)) & \rightarrow accept \\
filter(pkt(eth0, dst, new)) & \rightarrow accept \\
filter(pkt(ppp0, dst, new)) & \rightarrow drop \\
pkt(10.1.1.1, ppp0, s) & \rightarrow pkt(123.123.1.1, ppp0, s) \\
pkt(10.1.1.2, ppp0, s) & \rightarrow pkt(123.123.1.1, ppp0, s)
\end{array}$$

- The set Q contains ground terms with top symbol $filter$.
- A possible strategy for this policy, among others that guarantee a normalization process, is $\zeta = innermost(R)$.

This defines a security policy as all conditions of Definition 1 are satisfied.

Example 3. As already suggested in the introduction, we can model a policy for a clinical system (this example is adapted from the XACML specification [34], and first presented in the rewrite-based formalism in [12]).

- The policy signature \mathcal{F} contains the following symbols:

$$\begin{array}{lll}
accs & : Request \times Condition & \rightarrow Decision \\
req & : Subject \times Action \times Object & \rightarrow Request \\
read, write & : & \rightarrow Action \\
permit, deny, na & : & \rightarrow Decision \\
patient, phy & : Number & \rightarrow Subject \\
admin, per & : Number & \rightarrow Subject \\
record & : Number & \rightarrow Object \\
guard & : Subject \times Subject & \rightarrow Condition, \\
respPhy & : Subject \times Subject & \rightarrow Condition \\
urgency & : & \rightarrow Condition
\end{array}$$

- The set of decisions is $D = \{permit, deny, na\}$.
- R is the following set of rules, where variables are $x, y : Number$; $r : Object$; $c : Condition$:

$$\begin{array}{ll}
accs(req(patient(x), read, record(x)), c) & \rightarrow permit \\
accs(req(per(x), read, record(y)), guard(per(x), patient(y))) & \rightarrow permit \\
accs(req(phy(x), read, record(y)), respPhy(phy(x), patient(y))) & \rightarrow permit \\
accs(req(phy(x), write, record(y)), respPhy(phy(x), patient(y))) & \rightarrow permit \\
accs(req(admin(x), read, r), c) & \rightarrow deny \\
accs(req(admin(x), write, r), c) & \rightarrow deny.
\end{array}$$

In the order of appearance, these rules state that: a patient can read his own record, the guardian of a person can read the record for that person, the responsible physician of a patient can read or write data for his record, the last two rules deny any access of administrators to records.

- The set of requests Q is the set of all terms of the form $accs(\mathcal{T}(\mathcal{F}), \mathcal{T}(\mathcal{F}))$.
- One can adopt the strategy $\zeta = \text{choice}(R, accs(q, c) \rightarrow na)$, which introduces a default rule for this policy, where $q : Request$. The terms in Q which are not reduced by the rules from R will be rewritten into na . The example presented here is a security policy according to Definition 1.

The policy illustrated in this example has some desirable properties; for example the evaluation of a request is guaranteed to return a unique result, as will be demonstrated shortly.

A security policy is consistent if it computes at most one access decision:

Definition 2 (Consistency). A security policy $\wp = (\mathcal{F}, D, R, Q, \zeta)$ is consistent if for every query $q \in Q$, ζ applied to q returns at most one result: $\forall q \in Q$, the cardinality of $[\zeta](q) \cap D$ is less than or equal to 1.

This means that for every query evaluation, a deterministic result is computed by the application of ζ on the terms of Q . In the case where the strategy leads to a derivation that does not terminate on q , the cardinality of $[\zeta](q)$ is 0, the policy is still considered as consistent.

Example 4. Consider the following policy:

$$\begin{aligned} \wp_1 = (\mathcal{F}_1 &= \{g : Decision \times Decision \rightarrow Decision, permit, deny : Decision\}, \\ D_1 &= \{permit, deny\}, \\ R_1 &= \{g(x, y) \rightarrow x, g(x, y) \rightarrow y\}, \\ Q_1 &= g(\mathcal{T}(\mathcal{F}), \mathcal{T}(\mathcal{F})), \\ \zeta_1 &= \text{universal}(R)) \end{aligned}$$

Then \wp_1 is a security policy under the conditions expressed in Definition 1, but it clearly fails to be consistent, since $[\zeta_1](g(permit, deny)) = \{permit, deny\}$.

Since we assume strategies to be closed by concatenation, confluence under strategy can be simply expressed:

Definition 3 (Confluence under strategy). A rewrite system R is confluent under a strategy ζ when $\forall u, v_1, v_2 \in \mathcal{T}(\mathcal{F}, \mathcal{X})$ such that $\{v_1, v_2\} \subseteq [\zeta](u)$ then $[\zeta](v_1) \cap [\zeta](v_2) \neq \emptyset$.

If we consider the `universal` strategy, the above definition reduces to the usual one of confluence. Therefore:

Proposition 1. The policy $(\mathcal{F}, D, R, Q, \text{universal})$ is consistent as soon as R is confluent on $\mathcal{T}(\mathcal{F}, \mathcal{X})$.

A second fundamental property is termination:

Definition 4 (Termination). A security policy $\wp = (\mathcal{F}, D, R, Q, \zeta)$ is terminating if for every $q \in Q$, all derivations of source q in ζ are finite.

A fundamental property is that terminating and confluent term-rewriting systems evaluate any term to an unique normal form.

Example 5. The policy from Example 3 is terminating and confluent, which can be easily checked by analyzing the rules in R . This guarantees that the evaluation of any request will return a unique decision.

Example 6. Consider the policy:

$$\begin{aligned} \wp_2 = (\mathcal{F}_2 &= \{a : \text{Decision}, \text{permit} : \text{Decision}, \text{deny} : \text{Decision}\}, \\ D_2 &= \{\text{permit}, \text{deny}\}, \\ R_2 &= \{a \rightarrow a, a \rightarrow \text{deny}\}, \\ Q_2 &= \{a\}, \\ \zeta_2 &= \text{universal}(R)) \end{aligned}$$

\wp_2 is a security policy. In contrast to the previous example, this policy is consistent (since the corresponding rewrite relation is confluent), but it is not terminating.

Some simple sufficient conditions allows us to apply termination results from rewrite theory:

Proposition 2. *A policy $(\mathcal{F}, D, R, Q, \zeta)$ terminates provided that all derivations in ζ are finite or if R is strongly terminating (i.e. all derivations in $\text{universal}(R)$ are finite).*

To ensure strong termination, classical quite powerful termination tools can be used like recursive path orderings [13] or dependency pairs [1]. Termination allows one to localize confluence check following Newmann's lemma and this can be made operational via the completion algorithm [25]. Therefore we inherit sufficient conditions for confluence and termination of policies using the `universal` strategy. Since in general we may use the finer notion of termination and confluence under strategies, this opens new research questions to establish sufficient conditions also for rich strategies.

Another expected property of a policy strategy is that it is able to evaluate every incoming request into at least one decision, following its strategy. This is expressed through the decision completeness property:

Definition 5 (Decision Completeness). *A security policy $\wp = (\mathcal{F}, D, R, Q, \zeta)$ is decision complete if $\forall q \in Q, [\zeta](q) \neq \emptyset$ and $[\zeta](q)_{\downarrow} \subseteq D$.*

This definition is close to the definition of sufficient completeness of a rewrite system, which states that every ground term evaluates to a term exclusively built with constructors and possibly variables [11, 23]. Several algorithms have been developed to check sufficient completeness or to complete a set of patterns to ensure this property [8].

Proposition 3. *A policy $(\mathcal{F}, D, R, \mathcal{T}(\mathcal{F}), \text{universal}(\mathcal{R}))$ is decision complete provided that D is the set of terms built from constructors only and that R is terminating and sufficiently complete.*

An alternative set of conditions ensuring completeness is that R is weakly terminating and that an innermost strategy is used, as shown in [17].

4 Policy Composition

Let us now focus on the problem of *combining* policies in a modular way, relying on the long history of research in combining rewrite systems. This combination consists in taking the union of signatures and rules of the two policy components, choosing the sets of requests and decisions, and building a strategy for the combination of the two strategies in each component of the composition. However, combining naively access-control policies can result in inconsistent or non-terminating policies and we show how syntactic conditions and strategies may help to keep these suitable properties for the composition of two policies. Based on the example of XACML policy combinators, we explore the idea of a rich combination language for policies based on rewriting strategies.

Definition 6 (Policy Composition). *A composition of the two policies $\wp_i = (\mathcal{F}_i, D_i, R_i, Q_i, \zeta_i)$ ($i = 1, 2$) is any policy $\wp = (\mathcal{F}, D, R, Q, \zeta)$, where:*

1. $\mathcal{F}_1 \cup \mathcal{F}_2 \subseteq \mathcal{F}$;
2. $D_1 \cup D_2 \subseteq D \subseteq \mathcal{T}(\mathcal{F})$;
3. $R_1 \cup R_2 \subseteq R$;
4. $Q_1 \cup Q_2 \subseteq Q \subseteq \mathcal{T}(\mathcal{F})$;
5. ζ is a rewrite strategy for R .

Observe that when combining policies it may be convenient to introduce symbols not occurring in the original policies. The set of requests for the combined policy contains terms of the form determined by its sub-policies, but may also contain any additional well-formed closed terms that can be constructed from the combined policy signature. For example, suppose that $\mathcal{F}_1 = \{0, f\}$, $Q_1 = f(\mathcal{T}(\mathcal{F}_1))$ and $\mathcal{F}_2 = \{g\}$, $Q_2 = g(\mathcal{T}(\mathcal{F}_2))$, then a valid request would be $g(f(0))$.

The combination strategy is in charge of defining how the composed policy rewrites request terms. It may or not be built in a modular way by composing ζ_1 and ζ_2 . It often can be expressed as a functional composition of component strategies.

Example 7. Based on the policy from Example 3, let us show how we can extend it with additional rules. Consider the access control rules R' below:

$$\begin{array}{ll} \mathit{auth}(\mathit{req}(\mathit{phy}(x), \mathit{write}, r), \mathit{urgency}) & \rightarrow \mathit{permit} \\ \mathit{auth}(q, c) & \rightarrow \mathit{na} \end{array}$$

A strategy $\zeta' = \mathit{choice}(R', R, \mathit{auth}(q, c) \rightarrow \mathit{na})$ extends the previous policy by enforcing the rule for urgency cases first, and at the same time does not interfere with the decisions generated by the previous set of rules. In the case rule application fails for any of these cases, the combined policy delivers the *not-applicable* decision. This is a direct consequence of the semantics of the `choice` strategy.

The next example illustrates that much care must be taken in composing two policies.

Example 8. Consider the policies \wp_1 , from Example 4, and the policy \wp_3 below.

$$\begin{aligned} \wp_3 = (\mathcal{F}_3 = \{ & f : Decision \times Decision \times Decision \rightarrow Decision, \\ & permit : Decision, deny : Decision \} \\ D_3 = \{ & permit, deny \} \\ R_3 = \{ & f(permit, deny, x) \rightarrow f(x, x, x), \\ & f(deny, permit, x) \rightarrow f(x, x, x), \\ & f(x, x, x) \rightarrow x \}, \\ Q_3 = & f(\mathcal{T}(\mathcal{F}_2), \mathcal{T}(\mathcal{F}_2), \mathcal{T}(\mathcal{F}_2)), \\ \zeta_3 = & \text{universal}(R_2)) \end{aligned}$$

The composition \wp of \wp_1 and \wp_2 can be defined in a straightforward way as $\wp =$:

$$(\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_3, D = D_1 = D_3, R = R_1 \cup R_3, Q = \mathcal{T}(\mathcal{F}_1 \cup \mathcal{F}_3), \zeta = \text{universal}(R))$$

These two policies are clearly terminating and share only symbols *permit* and *deny*. It is therefore quite intuitive to believe that their composition will be also terminating. But this is false since the following request has an infinite derivation: $f(g(permit, deny), g(permit.deny), g(permit, deny)) \rightarrow f(permit, g(permit, deny), g(permit, deny)) \rightarrow f(permit, deny, g(permit, deny)) \rightarrow f(g(permit, deny), g(permit.deny), g(permit, deny)) \dots$

A property of rewrite systems is said to be *modular* if it is preserved under composition of systems. Many modularity results for confluence and termination of rewrite systems have been produced and the interested reader can refer for instance to [35] for a survey. Confluence and termination are in general not modular properties for rewrite systems. In the context of rewrite systems on disjoint signatures, confluence is modular, while termination is not [37]. However, adding syntactic conditions on rewrite rules or on the existence of a simplification ordering, allows getting positive results. Relying on the results of the rewrite system community [38, 36, 31, 18, 26], we can state the following useful results about composition of security policies.

Proposition 4. *Let us consider two policies $\wp_i = (\mathcal{F}_i, D_i, R_i, Q_i, \text{universal}(R_i))$ ($i = 1, 2$) such that \mathcal{F}_1 and \mathcal{F}_2 are disjoint and their composition $\wp = (\mathcal{F}_1 \cup \mathcal{F}_2, D_1 \cup D_2, R_1 \cup R_2, \mathcal{T}(\mathcal{F}_1 \cup \mathcal{F}_2), \text{universal}(R))$. If \wp_1 and \wp_2 are consistent, then \wp is consistent. If \wp_1 and \wp_2 are terminating, then so is \wp , provided:*

1. *neither R_1 nor R_2 contain collapsing rules, or*
2. *neither R_1 nor R_2 contain duplicating rules, or*
3. *R_1 or R_2 contains neither collapsing rules nor duplicating rules, or*
4. *termination of R_1 and of R_2 are proved by a simplification ordering.*

Relaxing the disjointness assumption of signatures in the previous results led to consider constructor-sharing systems [27], composable systems [32] or hierarchical combinations of rewrite systems generalizing the previous ones by allowing a certain sharing of defined symbols [14].

The interest of rewriting strategies appears again in their composition. For instance, in contrast to termination, innermost termination has a nice modular behavior, for disjoint unions, constructor-sharing systems, composable systems and for certain hierarchical combinations. We can take advantage of such results about innermost termination [19] to state the following result:

Proposition 5. *Let us consider two policies $\wp_i = (\mathcal{F}_i, D_i, R_i, Q_i, \text{innermost}(R_i))$ ($i = 1, 2$) such that \mathcal{F}_1 and \mathcal{F}_2 are disjoint or share only constructors, and \wp be their composition $(\mathcal{F}_1 \cup \mathcal{F}_2, D_1 \cup D_2, R_1 \cup R_2, \mathcal{T}(\mathcal{F}_1 \cup \mathcal{F}_2), \text{innermost}(R_1 \cup R_2))$. Then \wp is terminating as soon as \wp_1 and \wp_2 are.*

Example 9. Let us consider again the policies \wp_1 , from Example 4 and \wp_3 , from Example 8, but now with different strategies $\zeta'_1 = \text{innermost}(R_1)$ and $\zeta'_3 = \text{innermost}(R_3)$. Their combination $\wp = (\mathcal{F}_1 \cup \mathcal{F}_3, D, R_1 \cup R_3, \mathcal{T}(\mathcal{F}_1 \cup \mathcal{F}_3), \text{innermost}(R))$ is terminating according to Proposition 5.

Semantics of XACML Policy Combiners. In this paragraph, we show that the rewriting approach can capture the behavior of the access-control language XACML. Using rewriting and strategies it is easy to simulate the basic evaluation of XACML, computing *permit* or *deny* on a given request via a single policy. We do not present the translation from XACML to rewrite systems here due to the lack of space; details can be found in [15]. Instead, we show how the XACML policy *combiners* can be captured. The main combiners are listed below.

- *permit-overrides*: whenever *one* of the policies answers to a request with a *granting* decision, the final authorization for the composed policy will be granted. The policy will generate a *denial* only in the case at least one of the sub-policies denies the request, and all others return *not-applicable*.
- *deny-overrides*: this combiner has a similar semantics to *permit-overrides*, with the difference that denials take precedence.
- *first-applicable*: the decision produced by the combined policy corresponds to the authorization determined by the first sub-policy that does not fail, and whose decision is different from *not-applicable*.

To simulate *permit-overrides*, consider the following set of rules R_{po} .

$$\begin{aligned}
po(\text{permit}, y) &\rightarrow \text{permit} \\
po(x, \text{permit}) &\rightarrow \text{permit} \\
po(\text{deny}, na) &\rightarrow \text{deny} \\
po(na, \text{deny}) &\rightarrow \text{deny} \\
po(na, na) &\rightarrow na
\end{aligned}$$

We add an additional rule, R_{wrap} , whose purpose is to wrap any incoming request with the po function:

$$q(x) \rightarrow po(q(x), q(x))$$

In order to encode the *permit-overrides* combiner over two sub-policies, we can write the following strategy:

$$[\zeta_{po}](q) = \text{seq}([R_{wrap}](q), \text{onceBottomUp}(\zeta_1), \text{onceBottomUp}(\zeta_2), R_{po})$$

It means that the global policy intercepts requests before they are evaluated by the sub-policies. This builds a new term (of the form $po(t, t)$), containing two subterms, which

will be separately evaluated in a bottom-up fashion by the component sub-policy strategies ζ_1 and ζ_2 , then the reductions with R_{po} occur in the top position of the wrapped term, generating a final decision.

Therefore, given two policies $\wp_i = (\mathcal{F}_i, \{permit, deny, na\}, R_i, Q_i, \zeta_i)(i = 1, 2)$ the *permit-overrides* combiner is defined as

1. $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2 \cup \{po : Decision \times Decision \rightarrow Decision, q : Request \rightarrow Decision\}$;
2. $D = \{permit, deny, na\}$;
3. $R = R_1 \cup R_2 \cup R_{po} \cup R_{wrap}$;
4. $Q = \{q(t) \mid t \in Q_1 \cup Q_2\}$;
5. $\zeta = \zeta_{po}$.

Clearly, *deny-overrides* can be simulated in an analogous way. In order to simulate the *first-applicable* combining algorithm, we need only to construct the strategy that schedules the rules in the order they appear in the policy file.

5 Related and further work

There are numerous proposals for languages in which to express access-control policies, many of them “logic-based”. References [16, 20, 21, 22] represent a small sample of those.

With respect to policy composition, a number of works have a close relationship with the formalization introduced here. Bonatti et al. [7] address the composition problem through an algebra of composition operators that is able to define policy templates, among other operations; Wijesekera and Jajodia [40] take a similar approach. The operator definitions can be adapted to several languages and situations, since their definition is orthogonal to the underlying authorization language. On the other hand, we have shown how to deliver fine-grained control over the rule interaction of sub-policies.

Another alternative for composing access control policies is implemented by the Polymer system [5], which proposes rather classical operators on policies (conjunction, precedence, etc), and that allows reusing the policy objects, modifying them by executing additional actions, in order to specialize or enforce the policy.

A completely different approach to composition is taken by Lee et al in [29], based on the non-monotonic properties of defeasible logics. Here a single operator is proposed, which takes into account a precedence relation among the policies. We advocate that this kind of composition can also be achieved using rewriting strategies, which can readily define priorities on the rules.

Future work. We are using the Tom system to prototype and study the behavior of rewrite-based policies. Tom can also support the compilation process of our access control policies into JAVA classes, through the formal island framework [4]. The concepts presented in this paper provide a formal basis for reasoning about policies. This opens the way to the application of automated analysis tools to proving properties of policies.

References

- [1] T. Arts and J. Giesl. Termination of term rewriting using dependency pairs. *Theoretical Computer Science*, 236:133–178, 2000.
- [2] F. Baader and T. Nipkow. *Term Rewriting and all That*. Cambridge University Press, 1998.
- [3] E. Balland, P. Brauner, R. Kopetz, P.-E. Moreau, and A. Reilles. Tom Manual. LORIA, Nancy (France), version 2.4 edition, October 2006.
- [4] E. Balland, C. Kirchner, and P.-E. Moreau. Formal islands. In M. Johnson and V. Vene, editors, *AMAST*, volume 4019 of *Lecture Notes in Computer Science*, pages 51–65. Springer, 2006.
- [5] L. Bauer, J. Ligatti, and D. Walker. Composing security policies with polymer. In V. Sarkar and M. W. Hall, editors, *PLDI*, pages 305–314. ACM, 2005.
- [6] M. Bezem, J. W. Klop, and R. de Vrijer, editors. *Term Rewriting Systems*. Cambridge University Press, 2002.
- [7] P. A. Bonatti, S. D. C. di Vimercati, and P. Samarati. An algebra for composing access control policies. *ACM Trans. Inf. Syst. Secur.*, 5(1):1–35, 2002.
- [8] A. Bouhoula. Spike: a system for sufficient completeness and parameterized inductive proofs. In A. Bundy, editor, *CADE*, volume 814 of *Lecture Notes in Computer Science*, pages 836–840. Springer, 1994.
- [9] H. Cirstea and C. Kirchner. The rewriting calculus — Part I and II. *Logic Journal of the Interest Group in Pure and Applied Logics*, 9:427–498, May 2001.
- [10] H. Cirstea, C. Kirchner, L. Liquori, and B. Wack. Rewrite strategies in the rewriting calculus. In B. Gramlich and S. Lucas, editors, *Electronic Notes in Theoretical Computer Science*, volume 86. Elsevier, 2003.
- [11] H. Comon. Sufficient completeness, term rewriting systems and "anti-unification". In J. H. Siekmann, editor, *CADE*, volume 230 of *Lecture Notes in Computer Science*, pages 128–140. Springer, 1986.
- [12] A. S. de Oliveira. Rewriting-based access control policies. In M. Fernandez and C. Kirchner, editors, *Proceedings of the 1st International Workshop on Security and Rewriting Techniques - SecRet'06*, June 2006.
- [13] N. Dershowitz. Termination of rewriting. *Journal of Symbolic Computation*, 3(1 & 2):69–116, 1987.
- [14] N. Dershowitz. Hierarchical termination. In *Proceedings 4th International Workshop on Conditional Term Rewriting Systems, Jerusalem (Israel)*, volume 968 of *Lecture Notes in Computer Science*, pages 89–105. Springer-Verlag, 1994.
- [15] D. J. Dougherty. Core XACML and term-rewriting systems. Technical Report WPI-CS-TR-07-07, Worcester Polytechnic Institute, 2007.
- [16] D. J. Dougherty, K. Fisler, and S. Krishnamurthi. Specifying and reasoning about dynamic access-control policies. In U. Furbach and N. Shankar, editors, *IJCAR*, volume 4130 of *Lecture Notes in Computer Science*, pages 632–646. Springer, 2006.
- [17] I. Gnaedig and H. Kirchner. Computing constructor forms with non terminating rewrite programs. In A. Bossi and M. J. Maher, editors, *PPDP*, pages 121–132. ACM, 2006.
- [18] B. Gramlich. Generalized sufficient conditions for modular termination of rewriting. In H. Kirchner and G. Levi, editors, *Proceedings of the 3rd Algebraic and Logic Programming Conference*, volume 632 of *Lecture Notes in Computer Science*, pages 53–68. Springer-Verlag, September 1992.
- [19] B. Gramlich. On proving termination by innermost termination. In H. Ganzinger, editor, *Proceedings 7th Conference on Rewriting Techniques and Applications, New Brunswick (New Jersey, USA)*, volume 1103 of *Lecture Notes in Computer Science*, pages 93–107. Springer-Verlag, July 1996.

- [20] J. Y. Halpern and V. Weissman. Using first-order logic to reason about policies. In *CSFW*, pages 187–201. IEEE Computer Society, 2003.
- [21] S. Jajodia, P. Samarati, M. L. Sapino, and V. S. Subrahmanian. Flexible support for multiple access control policies. *ACM Trans. Database Syst.*, 26(2):214–260, 2001.
- [22] A. Kalam, R. Baida, P. Balbiani, S. Benferhat, F. Cuppens, Y. Deswarte, A. Mieke, C. Saurel, and G. Trouessin. Organization based access control. *Policies for Distributed Systems and Networks, 2003. Proceedings. POLICY 2003. IEEE 4th International Workshop on*, pages 120–131, 2003.
- [23] D. Kapur, P. Narendran, D. J. Rosenkrantz, and H. Zhang. Sufficient-completeness, ground-reducibility and their complexity. *Acta Inf.*, 28(4):311–350, 1991.
- [24] C. Kirchner, H. Kirchner, and M. Vittek. Designing clp using computational systems. In P. V. Hentenryck and S. Saraswat, editors, *Principles and Practice of Constraint Programming*, chapter 8, pages 133–160. MIT press, 1995.
- [25] D. E. Knuth and P. B. Bendix. Simple word problems in universal algebras. In J. Leech, editor, *Computational Problems in Abstract Algebra*, pages 263–297. Pergamon Press, Oxford, 1970.
- [26] M. Kurihara and A. Ohuchi. Modularity of simple termination of term rewriting systems. *Journal of IPS Japan*, 31(5):633–642, 1990.
- [27] M. Kurihara and A. Ohuchi. Modularity of simple termination of term rewriting systems with shared constructors. *Theor. Comput. Sci.*, 103(2):273–282, 1992.
- [28] B. Lampson. Protection. *ACM Operating Systems Review. Vol.*, 8:18–24, 1974.
- [29] A. J. Lee, J. P. Boyer, L. Olson, and C. A. Gunter. Defeasible security policy composition for web services. In M. Winslett, A. D. Gordon, and D. Sands, editors, *FMSE*, pages 45–54. ACM, 2006.
- [30] N. Martí-Oliet, J. Meseguer, and A. Verdejo. Towards a strategy language for maude. *Electr. Notes Theor. Comput. Sci.*, 117:417–441, 2005.
- [31] A. Middeldorp. A sufficient condition for the termination of the direct sum of term rewriting systems. In *Proceedings 4th IEEE Symposium on Logic in Computer Science, Pacific Grove*, pages 396–401, 1989.
- [32] A. Middeldorp and Y. Toyama. Completeness of combinations of constructor systems. In *Proceedings 4th Conference on Rewriting Techniques and Applications, Como (Italy)*, 1991. also Report CS-R9058, CWI, 1990.
- [33] P.-E. Moreau, C. Ringeissen, and M. Vittek. A pattern matching compiler for multiple target languages. In G. Hedin, editor, *CC*, volume 2622 of *Lecture Notes in Computer Science*, pages 61–76. Springer, 2003.
- [34] T. Moses. eXtensible Access Control Markup Language (XACML) version 2.0. Technical report, OASIS, 2005.
- [35] E. Ohlebusch. *Advanced Topics in Term Rewriting*. Springer, 2002.
- [36] M. Rusinowitch. On termination of the direct sum of term rewriting systems. *Information Processing Letters*, 26(2):65–70, 1987.
- [37] Y. Toyama. Counterexamples to termination for the direct sum of term rewriting systems. Technical report, NTT Electrical Communications Laboratories Japan, 1987.
- [38] Y. Toyama. On the Church-Rosser property for the direct sum of term rewriting systems. *Journal of the ACM*, 34(1):128–143, January 1987.
- [39] E. Visser. Stratego: A language for program transformation based on rewriting strategies. System description of Stratego 0.5. In A. Middeldorp, editor, *Rewriting Techniques and Applications (RTA'01)*, volume 2051 of *Lecture Notes in Computer Science*, pages 357–361. Springer-Verlag, May 2001.
- [40] D. Wijesekera and S. Jajodia. A propositional policy algebra for access control. *ACM Trans. Inf. Syst. Secur.*, 6(2):286–325, 2003.