Extensible Access Control with Authorization Contracts

Scott Moore
Harvard University (USA)
sdmoore@fas.harvard.edu

Christos Dimoulas
Harvard University (USA)
chrdimo@seas.harvard.edu

Robert Bruce Findler
Northwestern University (USA)
robbie@eecs.northwestern.edu

Matthew Flatt
University of Utah (USA)
mflatt@cs.utah.edu

Stephen Chong
Harvard University (USA)
chong@seas.harvard.edu

Abstract
Existing programming language access control frameworks do not meet the needs of all software components. We propose an expressive framework for implementing access control monitors for components. The basis of the framework is a novel concept: the authority environment. An authority environment associates rights with an execution context. The building blocks of access control monitors in our framework are authorization contracts: software contracts that manage authority environments. We demonstrate the expressiveness of our framework by implementing a diverse set of existing access control mechanisms and writing custom access control monitors for three realistic case studies.

Categories and Subject Descriptors D.3.1 [PROGRAMMING LANGUAGES]: Formal Definitions and Theory—Semantics; D.2.4 [SOFTWARE ENGINEERING]: Software/Program Verification—Programming by contract

Keywords access control; contracts; authorization logic

1. Introduction
An access control monitor mediates requests to call sensitive operations and allows each call if and only if the request possesses the necessary rights to call the operation. Broadly speaking, when an access control mechanism is presented with a call to a sensitive operation, it must be able to answer two questions. First, which rights are required for the call? And second, which rights does the request possess? The design of an access control mechanism specifies, implicitly or explicitly, the answers to these questions.

For example, Unix file permissions describe which users are allowed to call which operations on a file. The access control mechanism uses file permissions to determine what rights are necessary to call different sensitive operations. Each Unix process executes on behalf of a specific user, and a request to call an operation possesses the same rights as the user of the process that issues the request. Thus, file permissions answer the first question, and the rights of the user associated with a process answer the second question. Importantly, Unix associates users and processes in two different ways. By default, a new process runs on behalf of the same user as the process that spawned it. But a process can run on behalf of a different user if it runs a setuid executable, the operating system launches a new process to run the executable and associates the new process with the user that owns the executable, rather than the user that invoked it. Hence, this feature creates services that provide restricted access to resources that an invoking user could not otherwise access.

Similar to operating systems, software components also need access control mechanisms to prevent unauthorized clients from calling sensitive operations while allowing authorized ones to do so. Thus, when responding to a request to call a sensitive operation, access control mechanisms for components must be able to answer the same two questions: which rights are necessary for the call and which rights the request possesses.

However, access control needs of components vary, and it is impossible to choose a single answer to these questions that satisfies all component authors. To make things worse, access control mechanisms for general purpose programming languages have made design choices that are not suitable for all application domains and are typically mutually incompatible. For example, Java stack inspection \cite{40} determines the rights associated with a call site by walking the stack from the current stack frame. In contrast, object-capability languages (e.g., E \cite{25} and Caja \cite{26}) determine rights by the
The lexical structure of the program: a code may call operations on exactly those resources that are reachable from variables in the code’s text.

In this paper we propose a new, extensible access control framework that allows component authors to design access control monitors that suit their needs. The framework supports the design and implementation of many different novel and existing access control monitors for software components. Moreover, because different monitors are implemented using a common framework, different software components within the same application can use different access control mechanisms.

The framework builds on a novel concept: the authority environment. Just as each execution context has a variable environment that maps variable identifiers to values, each execution context has an authority environment that associates the context with its rights to call operations. The rights that a call to a sensitive operation possesses are those possessed by the authority environment of the call’s execution context.

By analogy with dynamic and lexical scoping of variable environments, we identify two ways in which an execution context can receive authority:
1. dynamically, by inheriting the authority environment of the surrounding execution context, and
2. lexically, by capturing the authority environment of the execution context where it is defined.

Returning to the Unix file system example, a process receives authority dynamically when it inherits the user of the process that launched it. A process receives authority “lexically” when it runs a setuid executable.

Based on the correspondence with variable scoping, we define a framework for designing access control monitors as sets of monitor actions that manipulate authority environments (§2). We implement our framework as a library for Racket [16] without changes to the language’s runtime. We use higher-order contracts [15] to specify where an access control monitor should interpose on a program and how it should manage authority environments. Contracts are executable specifications attached to software components that support separation of concerns by removing defensive checks from code implementing functionality [22,24]. In the same way, our authorization contracts separate the task of access control from the program’s functionality.

The design of this framework presents four major contributions:
1. the introduction of authority environments as a unifying concept for access control mechanisms (§2),
2. the introduction of context contracts to check and enforce properties of execution contexts (§3.1),
3. a novel authorization logic for representing and querying authority in authority environments (§3.2), and
4. authorization contracts that specialize context contracts for managing authority environments and enforcing access control policies expressed in the logic (§3.3).

We have used the framework to implement diverse access control mechanisms: discretionary access control, stack inspection, history-based access control, and object-capabilities (§4). We demonstrate the practicality of our approach with three realistic case studies (§5).

2. Authority Environments
In this section, we introduce authority environments as a unifying concept for access control. First, we review the differences between lexical and dynamic scoping (§2.1). Then we describe the connection between lexical and dynamic scoping and access control (§2.2) and show how we can use scoping in the design of a framework for writing access control monitors (§2.3). Throughout, we use small examples in the Racket programming language [16].

2.1 Lexical and Dynamic Scoping
The scope of a variable binding is the spatial and temporal part of the program in which it is visible. A common way to categorize strategies for assigning scopes to bindings is as either lexical or dynamic. Earlier work distinguishes between the scope of a binding, which describes where the binding is visible in the program text, and the extent of a binding, which describes when the binding is visible during execution. Dynamic scope often refers to bindings that have dynamic extent and “indefinite” scope. Here, we use dynamic scope to refer to bindings that have dynamic extent and lexical scope, also called “fluid” scope [17,18,19].

Under lexical scoping, a variable refers to the binding from its closest binder in the textual structure of the program. For example, in the Racket expression below, the variable x in function f refers to the binding in the outer-most let statement. The evaluation of this expression returns 0 since the inner-most let statement has no effect on the value x binds within f.

\[
\begin{align*}
\text{let } ([x \ 0]) \\
\text{let } ([f \ (\lambda \ () \ x)]) \\
\text{let } ([x \ 42]) \\
(f))
\end{align*}
\]

In a programming language with fluid scoping, programmers can instead associate a binding with the dynamic extent of an expression. That binding is visible to any code that runs in the dynamic extent of the expression. For example, the following Racket expression defines a new fluidly-scoped variable x with default value 0. The parameterize expression binds x to the value 42 in the dynamic extent of its body. The variable x in the body of f refers to the most recent binding rather than the closest one in the program text. Since f is invoked within the parameterize expression, the program evaluates to 42 instead of 0.

\[
\begin{align*}
\text{let } ([x \ (\text{make-parameter} \ 0)]) \\
\text{let } ([f \ (\lambda \ () \ (x))]) \\
\text{parameterize } ([x \ 42]) \\
(f))
\end{align*}
\]
Fluid scoping is a useful programming construct because it allows the context of an expression to communicate with its callees without explicitly threading arguments through the program. For example, a library function for printing may offer a parameter that determines the standard output file. Instead of threading that file as an argument through every function call leading to the printf routine, a client program can instead set the parameter once and all calls to printf in the body of the program see the client-specified file.

### 2.2 Scoping for Access Control

This ability to pass contextual information from an execution context to an eventual callee closely matches the problem of correctly determining the authority of a request to call a sensitive operation. To demonstrate this relationship, consider the design of a web application with multiple users. A key component of this application is a login function that authenticates users and executes code on their behalf:

```lisp
(define login user guess onSuccess)
(if (check-password? user guess) 
  (run-as-user user onSuccess) 
  (error "Wrong password!"))
```

This login function takes three arguments: the user attempting to authenticate, the password guess, and a callback on Success to invoke with the user’s rights if the password is correct. After checking the password, the login function changes the state of the program to indicate that the current user is now user and then calls onSuccess.

The body of onSuccess may attempt to access sensitive resources. For example, it may try to update a user’s profile. To avoid an unauthorized update, the update-profile function checks whether the current user has sufficient rights:

```lisp
(define update-profile profileUser text)
(if (can-update? currentUser profileUser) ...
  (error "Unauthorized!"))
```

Function can-update? compares the current user with the user who owns the profile to determine whether the update is authorized. This code thus implicitly uses the authority of its context, i.e., the current user, in much the same way that code accesses the dynamically scoped bindings from its context. By managing authority as an implicit context in this way, we can avoid modifying the code between the decision to run a computation with particular authority and the call to the sensitive operation. This has two advantages. First, threading authority explicitly through the program reduces extensibility, since third party code would need to be aware of and correctly handle authority explicitly. Second, if the code is untrustworthy, it might attempt to subvert the access control checks that protect the sensitive operation by fabricating its own authority.

Another requirement of the security of this application is that only code running with the authority of the main loop is allowed to switch users. According to the Principle of Least Privilege [31], we should further limit the code that is allowed to switch users to just the login function, and switch to an unprivileged user for the rest the program. Crucially, the body of the login function must still use the authority that was in its environment when it was created, i.e., the authority of the main loop. In a sense, for login, we wish to close over the authority of the main loop, in the same way that closures capture lexically scoped bindings.

To achieve this, we build on the analogy between scoping and access control and introduce the concept of an authority environment. An authority environment associates rights with an execution context, just as a variable environment associates bindings with an execution context. Just like variable environments, authority environments can be captured and associated with code, updated, and extended with new bindings for the dynamic extent of a computation. In this application, the authority environment of an execution context records the user on whose behalf the code executes. Section 2.3 shows how authority environments help enforce access control in our running example, including how to create a secure login function. Section 3 generalizes authority environments so that we can express a wide variety of access control mechanisms.

### 2.3 From Access Control to Authorization Contracts

Using the concept of an authority environment, we build an access control monitor that manipulates and inspects the authority environments of the example web application. The monitor consists of actions that describe how events in the execution of the application interact with its authority environment. We describe our framework for defining monitors in detail in Section 3. Here we explain only the features relevant to the example.

Figure 1 shows our example monitor. The monitor specifies three actions: setuid/c, chuser/c, and checkuser/c. Each

```lisp
(define-monitor users
  (monitor-interface
    setuid/c chuser/c checkuser/c)
  (action
    [chuser/c (user)
      #:on-create (do-create)
      #:on-apply (do-apply
        #:check (> @ current-principal user user)
        #:set-principal user))
    [checkuser/c (user)
      #:on-create (do-create)
      #:on-apply (do-apply
        #:check (> @ current-principal user user))
    [setuid/c
      #:on-create (do-create)
      #:on-apply (do-apply
        #:set-principal closure-principal))])
```

---

*Figure 1. Defining a simple access control monitor*
action defines a higher-order function contract [15]. When one of these contracts is attached to a function, the contract captures the current authority environment and associates it with the function. When the function is called, the contract has access to both the authority environment at the call site and the authority environment that it has captured. The monitor configures each action-contract with two hooks: #:on-create and #:on-apply. By changing these hooks, monitor designers can implement actions that implement different forms of “lexically” and dynamically scoped authority environments.

Action chuser/c is parameterized with an argument that identifies the user whose authority should be used during the execution of the body of a contracted function. The #:on-apply hook for chuser/c ignores the authority it has closed over and sets the active principal to user for the dynamic extent of the body of the contracted function, but only if the #:check holds, that is, the current-principal has authority over user user. Otherwise, it raises a security violation as a contract violation. Monitor action checkuser/c is also parameterized with a user. Upon a call of its contracted function, it checks that the current-principal has authority over user. If the check succeeds, the action does not change the authority environment. If the check fails, the action raises a security violation. The final monitor action, setuid/c, creates an authority closure: calling a function with this contract changes the current principal in the authority environment to the closed-over principal closure-principal for the dynamic extent of the function’s body.

Using the monitor, we can now reimplement the web application. First, we can replace code that defensively performs authorization checks with contracts that enforce authority requirements:

```
(define/contract
  ((update-profile someUser text)
   ->a ([user principal?] [text string?])
    #:auth (user) (checkuser/c user) any))
...
```

This revised implementation of update-profile uses the Racket form `define/contract` to attach a contract to the update-profile function. This contract is a `dependent contract` [15] for a function. It says that update-profile takes two arguments: user, which must be a principal?, and text, which must be a string?. The keyword argument #:auth (user)(checkuser/c user) says that the authorization contract for this function depends on the user argument and attaches the contract (checkuser/c user) to the function. Finally, the range of this contract is any, making no requirements on the return values. The definition of the function can now elide the authorization check.

We also revise the implementation of login:

```
(define/contract
  (login user guess onSuccess)
  (->a ([user principal?] [guess string?] [onSuccess (user) (chuser/c user)])
   #:auth () setuid/c any)
  (if (check-password? user guess)
      (onSuccess)
      (error "Wrong password")))
```

The keyword argument #:auth () setuid/c attaches the setuid/c action to the login function, capturing the authority of the program context where it is created. This allows the login function to execute with the captured authority and thus allows the application to switch to a less privileged principal without losing the ability to safely authenticate as a different user. In the original implementation, login uses a hand-rolled function run-as-user to confine onSuccess within the authority of user. In the revised code, login can invoke onSuccess directly. The contract on the onSuccess argument attaches the action (chuser/c user) to the function. This ensures that any call to onSuccess has the correct authority.

Rewriting this application to use authorization contracts makes the authorization requirements of each function clear, while also simplifying its implementation by removing authorization code that was spread throughout the program.

### 3. A Framework for Access Control

In this section, we present the general design of our framework with a formal model. First, we show how we extend existing higher-order function contracts to context contracts that check and modify information about their execution context (§3.1). Context contracts are expressive enough to enforce a wide range of properties. However, this flexibility makes it difficult to use them to implement and reason about access control. To free users from this burden, our framework provides a specialized interface for defining authorization contracts. The interface simplifies the definition of context contracts for access control in two ways. First, it specifies a common representation for authorization environments (§3.2). At the core of this representation is a novel authorization logic that describes how authority captured in a closure may be used. Second, it defines combinators for building authorization contracts (§3.3). Authorization contracts are specializations of context contracts that use the authorization logic to succinctly describe how they manage authority environments.

#### 3.1 A Contract System with Context Contracts

We model higher-order contracts and context contracts as extensions to an applied lambda calculus with modules and parameters, which implement dynamic binding. Figure 2 describes the syntax of our model. Figure 3 gives the reduction semantics of the model. Evaluation contexts that are not presented in Figure 2 are standard and enforce call-by-value, left-to-right evaluation. The type system and the definition of

[1] In Racket, a keyword argument is a (possibly optional) argument passed by keyword instead of position. A keyword is a symbol starting with #:.
evaluation contexts are given in Appendix A. Though we omit the extension for clarity, we have also extended the model with first class continuations, following Takikawa et al. [37], because our implementation language supports them and they interact in interesting ways with our framework.

The semantics of common language features is standard. Below, we explain the semantics of modules, parameters, higher-order contracts, and context contracts.

3.1.1 Modules

A program $p$ is a sequence of modules followed by a top-level expression. A module simultaneously defines a collection of values owned by a single component and a set of contracts for those values. Each module

module $\ell$ exports $x_{e_1}, \ldots$ where $y_1 = e_1, \ldots$

has a label $\ell$ and defines a set of values $y_1, \ldots$. Values within the module are visible only to subsequent modules in the program if they are exported. An export declaration $x_{e_1}$ with $x_{e_2}$ binds $x_{v_1}$ in the rest of the program to the value defined as $x_{e_1}$ in the where clause, but only after attaching to it the contract defined as $x_{e_2}$ in the where clause. Meta-function import, shown in Figure 3, substitutes occurrences of $x_{e_1}$ in the rest of the program with monitored values $^\ell $mon$^j_k(c, v)$ that enforce the contract.

3.1.2 Higher-order contracts

Term $^\ell $mon$^j_k(c, v)$ attaches contract $c$ to value $v$ and monitors whether uses of $v$ satisfy the contract. Labels $j$, $k$, and $\ell$, identify, respectively, the module that imposed the contract, the module that provided the value, and the module that is the client of the value. The top-level expression of a program is identified by the distinguished label $\ell_0$. These labels are used to assign blame when a contract is violated. The simplest contract is a flat contract $\flat(c, v)$ that takes a predicate $v$ as an argument. Flat contracts can be applied only to values of base types Int, Bool, and Unit. When the contract is attached to a value, the predicate is applied to the value. If the predicate returns true, the value passes to its context. Otherwise, the contract system stops the program and raises an error blaming the provider of the value.

Contract $c_d : \tau \rightarrow (c_r) c_r$ is a higher-order function contract. It specifies a contract $c_d$ for the domain of the function and a contract $c_r$ for the range of the function. In addition, it specifies a context contract $c_c$ for the contract. Context contracts, novel to this work, are higher-order contracts that enforce restrictions on the execution context of function calls. They are described in detail below. Attaching contract $c_d : \tau \rightarrow (c_r) c_r$ to a function returns a new value that enforces contracts on the argument and results of the function and applies the context contract $c_c$. A contract $c_d : \tau \rightarrow (c_r) c_r$ and $\ell$ parameterize $p$ for the function $c_r$ to a function returns a new value that enforces contracts on the argument and results of the function and applies the context contract $c_c$. A contract $c_d : \tau \rightarrow (c_r) c_r$ and $\ell$ parameterize $p$ for the function $c_r$ to a function returns a new value that enforces contracts on the argument and results of the function and applies the context contract $c_c$.

Contract $c_d : \tau \rightarrow (c_r) p$ is an indy-dependent contract for higher-order functions. This contract corresponds to the $\rightarrow a$ contracts from Section 2. The contract is dependent since the contract uses the argument of a contracted function to choose a context contract for the function. Applying a function $v$ with this contract has four steps. First, the argument is wrapped with the contract for the domain, $c_d$. Second, the wrapped argument is passed to the function $\lambda x : \tau. c_r$ to construct a context contract. Third, the resulting context contract is attached to $v$, which is applied to the wrapped argument. Finally, the contract for the range, $c_r$, is attached to the result of the application.

3.1.3 Parameters

Parameters are first-class values that can be used to access and install dynamic bindings. Parameters implement fluid scope because access to their dynamic bindings is controlled lexically by access to the parameter itself. The expression $\text{make-parameter } e$ creates a new parameter $p(r)$ with default value the result of $e$, where $r$ is a fresh tag uniquely identifying the parameter. The default value is recorded in the store $\sigma$. Term $\text{parameterize } p(r) = \sigma e_1 e_2$ installs the result of $e_1$ as the new value of the parameter $p(r)$ for the dynamic extent of $e_2$. Accessing the value of a parameter

---

2 An indy-dependent contract is a dependent function contract that uses the “indy” strategy for blame assignment [8].
Figure 3. Reduction semantics
with term \( \ell(p(r)) \) returns the value of the closest enclosing parameterize for \( p(r) \) in the current evaluation context. If there is no such term, it returns the current value for \( r \) in the store. Similarly, term \( p(r) := v \) mutates the current binding for the parameter, updating the parameter associated with either the closest enclosing parameterize for \( p(r) \) in the current evaluation context or the value for \( r \) in the store, if there is no such expression.

The parameter contract \( \text{param/c}(c) \) is a higher-order contract that restricts uses of a parameter. A contracted parameter \( v \) reduces to a proxy \( \ell(\text{param/p}^k(c,v)) \) that records the labels of the contract, provider, and client modules and intercepts uses of the parameter to enforce that values bound to the parameter meet contract \( c \).

### 3.1.4 Context contracts

To track properties of execution contexts, context contracts use parameters to install and access relevant state. A context contract interposes on programs at two key times: when the contract is attached to a function and when the contracted function is called. First, the parameter contract specifies a guard function \( \ell(r) \) that records the labels of the contract, provider, and client modules and intercepts uses of the parameter to enforce that values bound to the parameter meet contract \( c \).

A context contract

\[
\text{ctx/c}(v_c, (v_g, \Rightarrow v_{p_c} \leftarrow v_{v_a}), \ldots), \quad v_a, (v_{g_a}, \Rightarrow v_{p_a} \leftarrow v_{v_a}), \ldots)
\]

has four parts:

1. \( v_c \), a predicate that checks whether the context is appropriate when the contract is attached,
2. \((v_g, \Rightarrow v_{p_c} \leftarrow v_{v_a}), \ldots\) a list of guarded parameterizations, described below, to close over when the contract is attached,
3. \( v_a \), a predicate that checks whether the context is appropriate when the contract function is called, and
4. \((v_{g_a}, \Rightarrow v_{p_a} \leftarrow v_{v_a}), \ldots\) a list of guarded parameterizations to be installed around the body of the contracted function if the contract check succeeds.

The first two parts are evaluated when the contract is attached to a value. First, the predicate \( v_c \) is executed to allow the contract to check the current context. If the predicate returns \#t, a contract error is raised blaming the client of the contract. Otherwise, each guarded parameterization \((v_g, \Rightarrow v_{p_c} \leftarrow v_{v_a})\) from part 2 is evaluated in turn. Each guarded parameterization specifies a guard function \( v_g \), a parameter \( v_{p_c} \), and a value function \( v_{v_a} \). If invoking the guard thunk \( v_g \) returns \#t, the corresponding value thunk \( v_{v_a} \) is executed to produce a new value. This value is “closed over” and re-installed for parameter \( v_{p_c} \) when the contracted function is applied. The predicate \( v_a \) and the remaining parameterizations are recorded in a proxy \( \ell(\text{ctx/p}^k(v_a, (v_g, \Rightarrow v_{p_a} \leftarrow v_{v_a}), \ldots)) \).

The proxy enforces additional checks and parameterizations when the contracted function is called. First, the param-
the interposition is local to individual components, an access control monitor can be installed around a component without a global enforcement mechanism or the cooperation of other components.

Put differently, complete monitoring guarantees that contracts can enforce the same set of properties as reference monitors: an arbitrary prefix-closed property of a sequence of events. For contracts, these events are the attachment of contracts to values and the use of contracted values. In contrast, for inline reference monitors built with aspects, this set of events is determined by the point-cuts selected by the policy. In either case, the programmer must correctly identify relevant events and specify the policy, but can assume the policy is enforced.

The formal definition and proof of complete monitoring for our contract system can be found in the accompanying technical report [28].

### 3.2 Representing Authority

In principle, a programmer can use context contracts to enforce arbitrary properties of execution contexts such as access control, but in practice this requires the careful design of an appropriate representation of the relevant information as an environment, i.e., a set of parameters. In particular, for access control this requires a representation of the authority of an execution context.

The authority of an execution context describes the rights it has to perform sensitive operations. In different access control mechanisms, the form and organization of these rights varies. For example, in a web application, a session executes on behalf of a particular user whose rights may change over time in accordance with the access control policies attached to the application’s resources. In the Java stack inspection framework, rights are sets of “permissions” possessed by activation records that can be queried with the checkPermission operation.

A common way to describe the structure of authority in an access control system involves a mapping from subjects (users, processes, or security domains) to access rights for objects (resources that require the protection of the access control system) [21]. In practice, subjects and objects may comprise the same entities, so we refer to both as security principals (or, simply, principals).

To build a framework that supports many different access control mechanisms, we need a general way to express and reason about principals, the authority of principals, and how principals delegate and restrict their authority. For this purpose, we use an authorization logic [2] based on the Flow-Limited Authorization Model (FLAM) [5]. We briefly describe our logic, highlighting where it differs from FLAM. In Section 3.3 we employ this logic to represent authority as a set of parameters that are managed by specialized context contracts, dubbed authorization contracts.

Figure 6 presents the syntax of our logic. We assume an enumerable set of primitive principals \( \mathcal{P} \). Primitive principals represent program entities that possess rights, such as users, modules, or activation records. We assume a most trusted principal \( \top \) and a least trusted principal \( \bot \). For principals \( p \) and \( q \), the conjunctive principal \( p \land q \) is a principal with the authority of both \( p \) and \( q \). Similarly, the disjunctive principal \( p \lor q \) has the authority of either \( p \) or \( q \).

If principal \( p \) trusts principal \( q \), we write \( q \triangleright p \), and say that \( q \) acts for \( p \). The acts-for relation is reflexive and transitive, and induces a lattice structure over the set of principals, with conjunction as join, disjunction as meet, and \( \top \) and \( \bot \) as the top and bottom elements of the lattice.

Principals may assert the existence of trust relationships. A delegation \( p \trianglerighteq q \triangleright r \) means that principal \( r \) asserts that \( p \) acts for \( q \) (or, equivalently, that \( q \) delegates its authority to \( p \)). Of course, whether a principal \( s \) believes the assertion depends on whether \( s \) trusts principal \( r \). (We differ from FLAM in that we describe only the integrity of delegations, not their confidentiality.)

Judgment \( D \vdash r \trianglerighteq p \trianglerighteq q \) denotes that given the set of delegations \( D \), principal \( r \) believes that principal \( p \) acts for principal \( q \). Intuitively, \( r \) believes that \( p \) acts for \( q \) if that trust relationship can be derived using only delegations asserted by principals that \( r \) trusts.

Figure 7 presents the inference rules for the judgment \( D \vdash r \trianglerighteq p \trianglerighteq q \). Rules \( \text{BOT}, \text{TOP}, \text{REFL}, \text{TRANS}, \text{CONJ-LEFT}, \text{CONJ-RIGHT}, \text{DISJ-LEFT}, \) and \( \text{DISJ-RIGHT} \) are standard and provide the underlying lattice structure for the acts-for relation. Rule \( \text{DEL} \) captures the intuition that principal \( r \) trusts only delegations asserted by principals that it trusts, that is delegations \( p \trianglerighteq q \otimes s \) where \( D \vdash r \trianglerighteq s \trianglerighteq r \).

We have three additional principal constructors. Principal \( p \trianglerighteq \alpha \) is the projection of the authority of principal \( p \) on dimension \( \alpha \). We use projections to limit or attenuate the

---

3FLAM considers basis projections and ownership projections. The projections we use here are more general and have less structure than either. In addition to preserving the acts-for lattice, the only structure we impose is that projections are commutative: \( p \trianglerighteq \alpha \trianglerighteq \beta = p \trianglerighteq \beta \trianglerighteq \alpha \).
authority of a principal, and to identify access rights. For example, \( p \mapsto \text{files} \) may refer to principal \( p \) ’s authority restricted to \( p \) ’s rights to access the file system. Similarly, principal \( p \mapsto \text{obj} \mapsto \text{invoke} \) (equivalently \( p \mapsto \text{invoke} \mapsto \text{obj} \)) might refer to the right to invoke a particular object belonging to principal \( p \). Principal \( p \) can grant this right to another principal \( q \) by asserting a delegation: \( q \mapsto p \mapsto \text{obj} \mapsto \text{invoke} \).

We leave projection dimensions underspecified, and access control mechanisms can define their own dimensions. For any projection dimension \( \alpha \), principal \( p \) acts for principal \( p \mapsto \alpha \), as captured in Rule Proj. Typically the converse does not hold, and so \( p \mapsto \alpha \) has strictly less authority than \( p \).

Novel to this work, we introduce closure principals \( \overline{D} p \) and \( \overline{D} q \). Given a set of delegations \( D \) and principal \( p \), the left-closure principal \( \overline{D} p \) represents \( p \) with all of the trust relationships derivable from \( D \) where \( p \) delegates its authority to other principals. The right-closure principal \( \overline{D} q \) represents \( p \) with all of the trust relationships derivable from \( D \) where \( p \) acts for other principals. In our framework, delegations may change over time. Closure principals are useful because they allow us to capture trust relationships as they exist at particular moments in time. In particular, closure principals are a principled mechanism to describe how authority captured by a context contract should be combined with the current authority environment based on which parts of the closed over authority environment are trusted by principals in the current authority environment.

Rule Closure-Left shows that \( D : r \vdash p \geq \overline{D} q \) holds when there is some principal \( s \) such that at the time of closure creation (i.e., with delegation set \( D' \)), \( s \) believed that \( p \) acted for \( q \) (premise \( D' : s \vdash p \geq q \)), and moreover, right now (i.e., with delegation set \( D \)) principal \( r \) trusts principal \( s \) (premise \( D : r \vdash \overline{D} q \)). Typically, \( s \) and \( r \) are the same principal, meaning that \( r \) at-time-\( D \) trusts the decisions made by \( r \) at-time-\( D' \). Rule Closure-Right is similar and \( D : r \vdash \overline{D} q \geq \overline{D} p \) holds when there is some principal \( s \) such that at the time the closure was taken \( s \) believed that \( p \) acted for \( q \) (premise \( D' : s \vdash p \geq q \)), and principal \( r \) trusts \( s \) at-time-\( D' \) (premise \( D : r \vdash \overline{D} s \geq r \)).

To query whether a particular set of delegations satisfies an acts-for relation, we use a proof search algorithm adapted from FLAM [5]. We give examples of using delegations to implement different authorization mechanisms in Section 4.

Based on this logic, we represent an authority environment as:

1. a principal, who is responsible for the current execution context, and
2. a delegation set, which records the current trust relationships between principals.

The latter has two sub-parts: a global, mutable delegation set, and a set of delegations that are in place only for a currently executing context.

### 3.3 Authorization Contracts

Using authority environments, we can now introduce authorization contracts. Authorization contracts specialize context contracts in two ways. First, they prevent interference from untrustworthy code by using parameters that the rest of the program does not have access to. Second, they use a high-level representation of authority environments rather than directly manipulating parameters. Authorization contracts provide a structured way to describe how the underlying context contracts should manipulate authority environments.

Authorization contracts are defined as monitor actions using the define-monitor form (9). In this section, we model a pared-down version of define-monitor as an extension to the language model from Section 3.1. Figure 8 displays the syntax of the extension. The extension introduces new types, constructors, and operations for principals, delegations, and delegations sets, including an expression that evaluates an acts-for judgment: \( e ; e \vdash e \geq e \). The define-monitor form corresponds to the monitor \( (a \ldots) \) form that can appear in the where clause of a module definition in the extended model. Each \( a \) in monitor \( (a \ldots) \) is an action specification. An action specification \( \text{action } x (y : \tau, \ldots) \) (ce,ae) has a name \( (x) \), a set of arguments \( (y : \tau, \ldots) \), and two terms \( (ce \text{ and } ae) \) that define the action’s #:on-create and #:on-apply hooks.

The term for the #:on-create hook has the form

\[ \text{check: cee add: cee remove: cee setl-principal: cee cee closure-principal: cee cee closure-delegations: cee } \]

and specifies what the authorization contract should do when the contract is applied to a value. In particular it describes

<table>
<thead>
<tr>
<th>Bot</th>
<th>Top</th>
<th>Proj</th>
<th>CONJ-LEFT</th>
<th>CONJ-RIGHT</th>
</tr>
</thead>
<tbody>
<tr>
<td>( D : r \vdash p \geq 1 )</td>
<td>( D : r \vdash p \geq p )</td>
<td>( D : r \vdash p \geq p )</td>
<td>( D : r \vdash p_k \geq q ) ( k \in {1,2} )</td>
<td>( D : r \vdash q_k \geq q_2 )</td>
</tr>
<tr>
<td>DISJ-LEFT</td>
<td>( D : r \vdash p_1 \geq q ) ( D : r \vdash p_2 \geq q )</td>
<td>( D : r \vdash p \geq q )</td>
<td>( D : r \vdash p \geq p_1 \wedge p_2 \geq q )</td>
<td>( D : r \vdash p_1 \wedge p_2 \geq q )</td>
</tr>
<tr>
<td>CLOSURE-LEFT</td>
<td>( D' : s \vdash p \geq q )</td>
<td>( D' : s \vdash p \geq q )</td>
<td>( D : r \vdash p \geq q )</td>
<td>( D : r \vdash p \geq q )</td>
</tr>
<tr>
<td>CLOSURE-RIGHT</td>
<td>( D' : s \vdash p \geq q )</td>
<td>( D : r \vdash p \geq q )</td>
<td>( D : r \vdash p \geq q )</td>
<td>( D : r \vdash p \geq q )</td>
</tr>
<tr>
<td>REFLECTED</td>
<td>( D : r \vdash p \geq q )</td>
<td>( D : r \vdash p \geq q )</td>
<td>( D : r \vdash p \geq q )</td>
<td>( D : r \vdash p \geq q )</td>
</tr>
</tbody>
</table>

| TRANSFORMED   | \( D : r \vdash p \geq q \)  | \( D : r \vdash p \geq q \)  | \( D : r \vdash p \geq q \)  | \( D : r \vdash p \geq q \)  |

Figure 7. Inference rules for judgment \( D : l \vdash p \geq q \)
how to modify each part of the authority environment. Its field \texttt{check} accepts an acts-for judgment. If this judgement does not hold, a contract error is raised blaming the client of the contract. Field \texttt{add} accepts a set of delegations to add to the global delegation set. Field \texttt{remove} accepts a set of delegations to remove from the global delegation set, if present. Field \texttt{set!-principal} changes the current principal to the given principal. Fields \texttt{closure-principal} and \texttt{closure-delegations} accept a principal and a set of delegations, respectively, and record the principal and delegations for use upon a call to the contracted function. Terms in each of these six fields can access the pieces of the current authority environment using \texttt{current-principal} and \texttt{current-delegations}.

The second term corresponds to the \#on-apply hook, which specifies what the authorization contract should do upon a call of the contracted function. It has the form

\[
\]

Similar to a \texttt{ce} term, it allows the configuration of the contract’s behavior. As before, \texttt{check} accepts an acts-for judgment and raises a contract error if it does not hold. Likewise, fields \texttt{add} and \texttt{remove} mutate the global delegation set, and field \texttt{set!-principal} changes the current principal. The \texttt{scope} field accepts a set of delegations, but this set is installed only for the dynamic extent of the current function call, rather than added to the global delegation set. The \texttt{set-principal?} field requires a boolean value. If that value is \#false, the current authorization environment is extended with a principal for the dynamic extent of the function call. This \(1\) allows changing the principal visible within the extent of the function call and \(2\) prevents contracts that change the principal during the extent of the function call from modifying the principal of the enclosing context. In addition to accessing the current principal and delegations from the authority environment, the seven fields of an \texttt{ae} term can access the principal and delegations closed over by the contract with terms \texttt{closure-principal} and \texttt{closure-delegations}.

To give a detailed semantics for \texttt{monitor} terms, we use a compilation function that replaces \texttt{monitor} expressions with terms that explicitly construct context contracts. The compilation uses five parameters: one each for the current principal, global delegation set, and scoped delegation set, plus a pair to record the closed-over principal and delegations. Each \texttt{monitor} term generates a fresh set of parameters, preventing separately defined monitors from interfering with each other. Each \texttt{action} term compiles to a single context contract that closes over the fresh parameters. The full compilation function is listed in Appendix \[A\] along with typing judgments for authorization contracts and the semantics of expressions that operate on principals, dimensions, delegations, and delegation sets.

The hooks for defining actions are sufficiently flexible to implement a variety of access control mechanisms (§4). Here, we briefly describe some of the ways programmers can configure authorization contracts.

**Mutable Authority** Many access control mechanisms have a global policy that changes over time. For example, in discretionary access control, users can grant or revoke access to their resources. We can implement this with a contract that adds or removes (global) delegations.

**Dynamically-scoped Authority** An authority closure can inherit the authority environment from its calling context by ignoring the authority environment it closes over.

**“Lexically”-scoped Authority** An authority closure can isolate itself from the authority of its calling context by replacing the authority environment at a call site with the authority that it closes over.

Moreover, different access control mechanisms may require authorization contracts that blend these different strategies. For example, implementing \texttt{isetuid}-like authority closures requires capturing the principal but not the delegations the closures close over. Otherwise, updates to the global, mutable discretionary access control policy would be forgotten when a \texttt{setuid} function runs.

\[\text{Figure 8. Syntax extensions for authorization contracts.}\]
(define-monitor monitor-name
  (monitor-interface
    action-name ... extra-name ...)
  (monitor-syntax-interface
    syntax-name ...)
  (action ; definitions of monitor actions
    [action-name (action-var ...)
      #:on-create on-create-hook
      #:on-apply on-apply-hook]
    ...
  )
  (extra ; additional monitor abstractions
    (define extra-name extra-body)
    ...
  )
  (syntax ; syntactic monitor abstractions
    (define-syntax syntax-name syntax-body)
    ...
  )
))

Figure 9. The define-monitor form

4.1 The define-monitor Form

Figure 9 shows the complete syntax of define-monitor. It has two sections in addition to the action section we have seen before: extra and syntax. The first defines extra functions and contracts that the programmer wants to include in the interface of a monitor. These are usually contracts that combine two or more actions together or contracts that fix the arguments of an action. The syntax section defines macros that serve as syntactic abstractions over the monitor’s interface, for example, to automate the placement of authority contracts when defining a function. We give examples of definitions in the extra and syntax sections later. The monitor-interface and monitor-syntax-interface clauses specify which elements are available to users of the monitor. After defining a monitor monitor-name, a client can instantiate it with (run monitor-name). This creates a fresh monitor, i.e., one with a fresh authority environment and contracts.

The most complicated part of defining a monitor is writing the two hooks for each monitor action. To facilitate this, we provide two functions, do-create and do-apply, that simplify this process. Each function has optional keyword arguments corresponding to the fields of an action form in Section 3.3. The functions provide default values for any argument not specified. Thus, the programmer need only specify the results of the hooks they care about. For instance, the default value for the argument with keyword #:check seen in Figure 1 is an acts-for relation that is always true.

4.2 A Stack Inspection Monitor

In stack inspection [40], code obtains permissions based on static properties such as the package it belongs to. At runtime, code can choose to enable its static permissions making them eligible for satisfying an access control check. Before a sensitive operation, stack inspection checks for the presence of a particular permission by walking the run-time call stack until a frame from code that has enabled the permission is found. To prevent luring attacks [40], stack inspection additionally requires that all execution contexts between the enabled permission and the authorization check have the required static permission. Despite this protection, untrusted code may be able to influence the program even if its frames are no longer on the stack. As a result, modern adaptations of stack inspection provide additional support for capturing the permissions of the stack at some point in an execution and reinstating them for a later check.

Implementations of stack inspection provide the following primitives: checkPermission, which checks that a frame on the stack has the required permission enabled and that all intervening frames have the required static permission; doPrivileged, which enables the static permissions of the current code for its dynamic extent, possibly using captured permissions instead of the current permissions; and getContext, which captures the permissions of the stack at some point in execution. In addition, the implementation must provide a mechanism to associate static permissions with code.

To realize stack inspection using authorization contracts, a monitor must provide (1) actions that implement these primitives and (2) a way to grant static permissions to code. In our monitor, the actions for (1) are check-permission/c, do-privileged/c, and context/c. To track which permissions are held by code on the stack, we use the authority environment to grant permissions to individual frames, each represented by a distinct principal. Each stack frame has three projections that are used to manage its authority. The static projection indicates the permissions granted to the code statically. The enable projection has the authority of the permissions enabled for this frame. The active projection represents permissions that would satisfy a privilege check. We say a principal has a particular permission if it acts for the corresponding projection of the ⊤ principal.

We use one additional monitor action, privileged/c, to indicate the static permissions a piece of code possesses and to enforce that a stack frame’s active projection acts for exactly those permissions for which checkPermission should succeed. Action privileged/c takes a list of permissions (each of which is a projection of the ⊤ principal). On an #:on-apply event, it creates a new principal callee to represent the new stack frame and adds delegations initializing these projections for the dynamic extent of the function:
These delegations give callee the specified static permissions (by asserting that the callee’s static dimension acts for the conjunctive principle permissions), assert that the new frame inherits the active permissions from the previous frame, and require that the callee has both static and enabled permissions to make them active.

Tracking the authority of each frame in this way makes walking the stack unnecessary. Action check-permission/c only checks that the active projection of the current principal acts for all of the requested permissions.

Action do-privileged/c enables the current frame’s static permissions by adding a delegation from the frame’s static projection to its enable projection for the dynamic extent of the wrapped function.

Action context/c is used to capture the permissions of the current stack for future permission checks. It captures the current authorization environment when it is attached to a function. When it is invoked, it installs the same set of delegations as privileged/c, except that the first delegation that grants static permissions gets replaced with a delegation that derives permissions from the active permissions of the captured frame at the time they were captured:

\[
\begin{align*}
\geq & (\vdash \text{callee static}) \\
\geq & (\vdash \text{callee enable}) \\
& (\vdash \text{current-principal active}) \\
& \text{current-principal} \\
\geq & (\vdash \text{callee active}) \\
& (\forall (\vdash \text{callee enable}) (\vdash \text{callee static})) \\
& \text{callee}. \\
\end{align*}
\]

The right-closure principal on the right hand side of this delegation acts for all of the principals that closure-principal acted for when the closure was created.

The monitor must also provide (2) a way to grant static permissions to code. Because Racket does not have class-loading facilities that would allow permissions to be granted to code at load-time, we use macros to attach authorization contracts to code that should have static permissions. In particular, the monitor provides a new definition form define/rights in its syntax section. This form works like the define form, but takes two additional arguments: a set of permissions and a contract to apply to the definition. It defines a function wrapped with the given contract and a privileged/c contract. In addition, the macro define/rights coerces any function arguments or free-variables appearing in the body of the function to authority closures by applying an additional contract unprivileged/c, which is defined in the extra section of the monitor. Action unprivileged/c switches to the \( \bot \) principal for the dynamic extent of the closure it wraps, preventing any check-permission/c actions from succeeding. Thus, these contracts prevent functions that were not defined with define/rights from using code that requires permissions.

Figure 10 shows an example program using the stack inspection monitor. There are three functions defined using define/rights. Two of these functions are trusted to access the filesystem: read-file and read-privileged. However, read-file should not be used directly, so it checks that the filesystem permission has been enabled by one of its callers. Function read-privileged enables the filesystem permission, but only calls read-file if the file is safe to read. Function malicious does not have the filesystem permission but attempts to read “/etc/passwd” anyway, so invoking this function results in a contract violation. The contract violation says that the stack frame corresponding to the call to read-file does not have the necessary permission filesys.

5. Case studies

To evaluate the use of our framework in practical applications, we developed three case studies. The first adds simple authorization contracts to the implementation of a card game to ensure that player’s moves affect only the parts of the game state they control. The second secures a plugin interface of the DrRacket development environment and demonstrates how the flexibility of the framework can support complex security mechanisms. The third, which mirrors the example from Section 2.2, replaces authorization checks in a web application with authorization contracts.

We evaluated the performance of our framework on each case study. The experiments were conducted on a MacBook Pro with a 2.6 GHz Intel Core i5 and 16GB of RAM running Mac OS X 10.11 and Racket 6.4.0.9. In the first two case studies, authorization contracts have significant impact on the performance of the benchmarks. However, both case studies are worst case scenarios: they have no existing code
implementing access control (and so we are strictly adding functionality), and after adding contracts, they invoke many access control checks (tens of thousands in the case of the card game) while performing cheap operations. Moreover, in the DrRacket case study, the absolute overhead for each benchmark due to authorization contracts is less than 45ms, but the relative overhead is high since the baseline running time is less than 15ms. The third case study replaces existing access control checks with authorization contracts, with negligible impact on performance. Our implementation is a prototype, and we anticipate that optimizations in the implementation of our contracts can further reduce their overhead.

**Preventing Cheating in a Card Game** We have used authorization contracts to enforce a security policy for a functional implementation of the card game Dominion[4]. The exact rules of Dominion do not matter for our purpose, except that each player collects cards in a local deck and attempts to outscore the rest of the players by playing cards from their deck. During each turn, players can play cards from their deck to either purchase additional cards or attack other players, forcing them to discard some of their cards.

In this implementation, each player is a program that runs in its own process and responds automatically to messages from a central broker. The broker maintains the shared inventory of cards and a mirror of each player’s local deck. Players perform moves by sending messages to the broker describing the move.

To perform a move, the player sends a message to the broker identifying a card to play. In response, the broker updates its copy of the game state to reflect the move and, if the move involves an attack on another player, informs the other player of the attack. The other player then has an opportunity to defend by choosing which card to discard and the broker again updates the game state.

The broker represents the local deck of each player as an immutable record `player` and the state of the game as an immutable structure `game` that holds a list of `player` records. The first element in this list corresponds to the player who makes the next move. The broker is implemented as a core function that delegates to two functions: `move` and `defend`. Both functions perform functional updates to the relevant structures.

We enforce the policy that the broker only updates the current player’s deck or a defending player’s deck. The monitor that enforces this policy specifies three authorization contracts: `deprivilege/c`, which sets the principal for the dynamic extent of a function to `⊥`; `switch-player/c name`, which sets the principal for the dynamic extent of a function to the player with name `name`; and `check-player/c name`, which checks before calling a function if the current principal is the player with name `name`.

---

[4] The implementation is part of the teaching material of a long running undergraduate Functional Programming course.
can access or modify similar parts of DrRacket. For instance, methods that write to the clipboard (e.g. cut and copy) belong to the same group while methods that change how DrRacket displays content (e.g. set-max-width and set-line-spacing) belong to a second group. Each group has a corresponding privilege that is required to invoke the group’s methods. For example, the privileges ReadClipboard and ChangeEditorView grant access to the methods mentioned above. Methods can belong to multiple groups. Access control checks around each method verify that the authority of a calling execution context has the necessary privileges.

In addition to methods that require specific privileges to invoke, text% has sensitive methods that should be invoked only by another method of the text% object. For example, the on-delete method should never be invoked directly as its correctness depends on DrRacket’s state. Instead, key bindings should invoke the delete method that subsequently calls on-delete. To support this use case, we require an additional privilege to call on-delete that is granted during the dynamic extent of delete.

The stack-inspection-like access control mechanism we have described so far is not sufficient. Some methods of text% install callbacks that are triggered by subsequent events. For example, add-undo registers a callback that runs when the user wishes to undo the action of a key binding. This callback should not run with the authority of its calling context, but instead should use the privileges of the action that created it. To achieve this, we create authority closures around any callbacks registered by an action.

Our monitor represents each privilege as a unique principal and represents sets of principals as conjunctions and disjunctions of principals. It defines three actions: check/c, enable/c, and closure/c. Upon an #:on-apply event, the first action consumes principal perms and checks if the current principal has permissions that imply perms. Then the action switches the current principal to a principal that only has permissions perms. When a function wrapped with enable/c is applied, it switches the current principal to a principal that has the same permissions as the current principal augmented with perms. The #:on-create event of the third action creates an authority closure. When the authority closure is applied, it installs the closed-over principal.

We use the actions of the monitor to define an authorization contract for the keybinding interface:

\[
\text{(check/c ChangeEditorView). In essence, text/c defines a security policy for the editor.}
\]

To assess the monitor’s performance, we ran a series of 30 benchmarks, adapted from DrRacket’s test suite, that simulate a sequence of keystrokes that trigger built-in key bindings. We ran these benchmarks with the monitor off and on. When the monitor is on, the prototype grants the minimum set of privileges necessary for each key binding. For each benchmark, we measured the time required to retrieve and execute each key binding. Our measurements show that the authority monitor increases median response time by 3–7× and increases response time at the 99th percentile by 3–5×. However, for an IDE, a response time fast enough for interactive use is more important. Our prototype achieves this goal with a maximum response time of 53ms.

**Authentication in a Web Application** The Racket package system allows users to discover and install packages from a public index service. Individual users can add new packages or update old ones by logging into the index service web application, which is implemented using the Racket web-server. Requests to add or modify packages are issued to the application as asynchronous http requests. The baseline implementation of the application uses macros to authenticate the user and perform any required access control checks before processing the request. For example, the jsonp/pkg/modify endpoint authenticates the current user and checks that they are an author of the package they are attempting to modify. This approach to access control is brittle, since it requires that the checks included for each endpoint accurately capture the privileges required when processing the response.

Using authorization contracts, we are able to separate the tasks of authentication and authorization in the index service web application. Rather than performing a different set of access control checks for each endpoint, all endpoints now simply invoke an authenticate function that checks whether the current session is valid and which user is logged in, then invokes a procedure to process the request, like the login function from Section 2.2. The access control policies for sensitive operations like updating a package are enforced by adding authorization contracts that implement the necessary checks to the web application’s data model. There are two types of checks: (is-author/c pkg) which checks that the logged in user is an author of package pkg, and is-curator/c, which checks whether the logged in user has “curator” status, which allows them to tag packages with information about their quality.

To evaluate the new implementation’s performance, we measured the latency of 1,000 repeated requests to modify a package record. Replacing inline checks with authorization contracts has minimal impact on performance. Median latency was 283ms for the baseline implementation versus 281ms with authorization contracts. At the 99th percentile, using authorization contracts latency was 338ms versus 330ms with the baseline implementation.
6. Related Work

The connection between scoping and access control has been implicit in prior work on security in programming languages but has never been a central concept for extensible access control. Morris’s seminal paper “Protection in Programming Languages” [29] describes how lexical scope can be used to create security abstractions within a program. More recently, the object-capability paradigm has embraced lexical scope as an organizing security principle [25]. Wallach and Felten [39] note that “in some ways, [stack inspection] resembles dynamic variables (where free variables are resolved from the caller’s environment rather than from the environment in which the function is defined).” Phung et al. [30] use dynamic and lexical scoping to associate principals with executing code in order to correctly enforce security policies on programs that mix JavaScript and ActionScript code.

Inlined Reference Monitors An alternative approach to language-level access control is inlined reference monitoring. Reference monitors observe the actions taken by a system and intercede to prevent violations of a security policy [3]. They can enforce a large class of policies [32]. Inlined reference monitoring (IRM) weaves the implementation of a reference monitor into the program being monitored [12]. Many implementations of inlined reference monitoring rely on aspects to identify security relevant actions during program execution [6,7,12,14,20]. Policies supported by these tools typically focus on access patterns for sensitive resources. While policies supported by our framework can be encoded this way, as in Erlingsson and Schneider’s IRM implementation of Java stack inspection [13], policies where the authority of code depends on application state require duplicating code. A further disadvantage of IRMs is that they require a global transformation of the program to inline the security monitor. Because authorization contracts are applied at component boundaries, our framework requires only local modifications.

Authorization Logics Authorization logics give a formal language to express access control policies [11]. Authorization logics have been used to understand existing access control mechanisms, including Java stack inspection [39], Aura [19] and Fine [36] implement access control using proof-carrying authentication, where proofs of formulas in an authorization logic are used as capabilities [4]. Our access control logic is inspired by the Flow-Limited Authorization Model [5], which uses projections to describe attenuated authority without requiring additional constructs such as roles or groups.

Contracts for Security Previous work has used contracts to enforce limited access control policies. Moore et al. [27] use contracts to constrain the use of capabilities in a secure shell scripting language. Dimoulas et al. [9] use contracts to control the flow of capabilities between components in object-capability languages. Heidegger et al. [18] use contracts to specify which fields of an object may be accessed by a component. However, each of these systems is specialized to enforce a specific type of access control policy. Disney et al. [10] introduce temporal higher-order contracts that enforce that sequences of function calls and returns match a specification. Schollier et al.’s computational contracts [33] can enforce a wide range of trace properties on programs. Unlike authorization contracts and temporal higher-order contracts, computational contracts use aspects to interpose on program events. Both of these systems support arbitrarily powerful monitors, but like inlined reference monitoring, provide limited support for writing complex access control policies like stack inspection or discretionary access control.

Scoped Aspects for Security Dutchyn et al. [11] enforce simple access control policies with lexically and dynamically-scoped aspects. With additional aspect scoping mechanisms, Toledo et al. [38] encode full Java stack inspection. While aspects can enforce a wide range of access control mechanisms, authorization contracts offer linguistic support for implementing diverse (and customized) access control mechanisms with ease. Doing the same with aspects, if possible, requires brittle and complex encodings.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant Numbers CCF-1438271, CNS-1524052, and CCF-1526324. This research is also supported by the Air Force Research Laboratory and a Google Faculty Research Award.

References


A. Details of model

A.1 Evaluation contexts

\[
E ::= [i] \mid E e \mid v E \mid \text{let } x = E \text{ in } e \mid \text{if } E \text{ then } e \text{ else } e \mid E \oplus e \mid v \oplus E \mid E \leq e \mid v \leq E
\]

- make-parameter \( E \mid ?E \mid E := e \mid p(r) := E \)
- parameterize \( E = c \mid p \mid E \)
- parameterize \( p(r) = E \)
- parameterize \( p(r) = v \)

\[
\Gamma; \Sigma \vdash (E \to (\lambda x : \tau. e)) \mid (c : \tau \to (E \to c) \mid e : \tau \to (c \to (E \to e), \ldots, e(e \to e), \ldots))
\]

\[
\text{ctx} / \text{ctc} / \text{guard} / \text{mon} / \text{check} / \text{module}
\]

A.2 Typing judgments

\[
\Sigma \vdash p : \tau
\]

\[
\begin{align*}
\varnothing; \Sigma & \vdash m_1 : \Gamma_1 & \ldots & \ldots & \ldots & \ldots & \ldots & \Gamma_n ; \Sigma \vdash e : \tau \\
\Gamma_1 ; \ldots & \ldots & \ldots & \ldots & \Gamma_n ; \Sigma & \vdash e_1 : \tau
\end{align*}
\]

\[
\Sigma \vdash x \text{ with } x_1, \ldots, x_m \text{ where } x_1 = e_1, \ldots, x_m = e_m \vdash E
\]

\[
\begin{align*}
\varnothing; \Sigma & \vdash e_1 : \tau_1 & \ldots & \ldots & \ldots & \ldots & \ldots & \Gamma_n ; \Sigma \vdash e_n : \tau_n \\
\forall 1 \leq i < n', \exists 1 \leq n'' \leq n, x_i & \equiv y_i \land \tau_i \neq \tau \land \tau_i \neq \tau_c t c t & \land \tau_i \neq \tau_c t c t & \\
\Gamma & \equiv \{ y_1 : \tau_1, \ldots, y_n : \tau_n \} \mid (x_i, \ldots, x_{i'}) \\
\Gamma & \vdash \text{module } e \text{ exports } x_1 \text{ with } x_1, \ldots, x_m \text{ where } y_1 = e_1, \ldots, y_n = e_n \vdash E
\end{align*}
\]
A.3 Typing judgments for authorization contract extensions

\[ \Gamma; \Sigma \vdash e_1 : \text{Unit} \to \text{Bool} \quad \Gamma; \Sigma \vdash e_2 : \text{Unit} \to \text{Bool} \]
\[ \Gamma; \Sigma \vdash e_{p_1} : \text{Unit} \to \text{Bool} \quad \Gamma; \Sigma \vdash e_{c_1} : \text{Unit} \to \tau_{e_1} \]
\[ \Gamma; \Sigma \vdash e_{g_1} : \text{Unit} \to \text{Bool} \quad \Gamma; \Sigma \vdash e_{p_2} : \tau_{a_1} \quad \Gamma; \Sigma \vdash e_{a_1} : \text{Unit} \to \tau_{a_1} \]
\[ \Gamma; \Sigma \vdash \text{ctx}(e_1, (e_{p_1} \Rightarrow e_{p_2} \leftarrow e_{a_1}), \ldots, e_{a_1} \Rightarrow e_{a_1}) : \text{ctx} \quad \Gamma; \Sigma \vdash \text{param}(p)(e, e) : \tau \quad \Gamma; \Sigma \vdash e : \tau \quad \Gamma; \Sigma \vdash e : \tau \]

\[ \begin{align*}
\Gamma; \Sigma \vdash \text{guard}(e_g, v_p, v_c, e_f) : \tau_d \to \tau_r \\
\Gamma; \Sigma \vdash \text{install}(p)(v_p, e_c, e_f) : \tau_d \to \tau_r \\
\end{align*} \]
A.4 Compiling authorization contract extensions

\[
\begin{align*}
\Gamma; \Sigma \vdash \text{cee} : \tau \\
\Gamma; \Sigma \vdash \text{ae} : \tau \\
\Gamma; \Sigma \vdash \text{let} x = \text{ae} \text{ in} \, \text{ae} : \tau \\
\Gamma; \Sigma \vdash \text{cee} : \tau \\
\Gamma; \Sigma \vdash \text{let} x = \text{cee} \text{ in} \, \text{cee} : \tau \\
\Gamma; \Sigma \vdash \text{current-principal} : \text{Prin} \\
\Gamma; \Sigma \vdash \text{current-delegations} : \text{DelSet} \\
\Gamma; \Sigma \vdash \text{closure-principal} : \text{Prin} \\
\Gamma; \Sigma \vdash \text{closure-delegations} : \text{DelSet}
\end{align*}
\]

\[
\begin{align*}
\text{compile} & \left[ \text{module } \ell \text{ exports } x_1 \text{ with } x_{c_1}, \ldots \text{ where } y_1 = e_2, \ldots; p \right] = \\
& \text{compile-monitor} \left[ \text{module } \ell \text{ exports } x_1 \text{ with } x_{c_1}, \ldots \text{ where } y_1 = e_2, \ldots; \text{compile}[p] \right] \\
\text{compile} & = e \\
\text{compile-monitor} & \left[ \text{module } \ell \right. \\
& \text{exports } x_1 \text{ with } x_{c_1}, \ldots \\
& \text{where } y_1 = e_1, \ldots, y_n = e_n, \\
& \text{monitor} \left( \text{action} \, x_{a_1} \, (y_{a_1} : \tau_{y_{a_1}}, \ldots) \, (\text{ce}_1, \text{ae}_1), \ldots, \text{action} \, x_{a_n} \, (y_{a_n} : \tau_{y_{a_n}}, \ldots) \, (\text{ce}_n, \text{ae}_n) \right), \\
& y_{n+1} = e_{n+1}, \ldots, y_m = e_m \\
\text{where } p, d, s, cp, cd, curp, and curd & \text{ are fresh} \\
\text{compile-action} & \left[ \text{action} \, x \, (y : \tau_y, \ldots) \, (\text{ce}_a, \text{ae}_a) \right] = \\
& \lambda y : \tau_y, \ldots \\
& \text{ctx/c} \left( \text{compile-cecheck} \, [\text{ce}, p, d, s, \text{curp}, \text{curd}] \right), \\
& \text{compile-ce} \left( [\text{ce}, \text{cp}, \text{curp}, \text{curd}] \right), \\
& \text{compile-ce} \left( [\text{ce}, \text{cp}, \text{curp}, \text{curd}] \right), \\
& \text{compile-ce} \left( [\text{ae}, p, d, s, \text{curp}, \text{curd}] \right), \\
& \text{compile-ce} \left( [\text{ae}, s, \text{cp}, \text{curp}, \text{curd}] \right), \\
& \text{compile-ce} \left( [\text{ae}, p, \text{cp}, \text{curp}, \text{curd}] \right)
\end{align*}
\]

\[
\begin{align*}
\text{compile-cecheck} & \left[ \text{check} : e_1 \, \text{add} : e_2 \, \text{remove} : e_3 \, \text{set}-!-\text{principal} : e_4 \, \text{closure-principal} : e_5 \, \text{closure-delegations} : e_6, p, d, s, \text{curp}, \text{curd} \right] = \\
& \lambda s : \text{Unit} \\
& \text{let} \_ = \text{curp} := ?p \text{ in} \\
& \text{let} \_ = \text{curd} := ?d \cup ?s \text{ in} \\
& \text{let} p_1 \geq p_2 \Rightarrow p_1 = \text{compile-ce}\left[ e_1, \text{curp}, \text{curd} \right] \text{ in} \\
& \text{if } \text{curd} \Rightarrow p_1 \geq p_2 \text{ then} \\
& \text{let} \text{add} = \text{compile-ce}\left[ e_2, \text{curp}, \text{curd} \right] \text{ in} \\
& \text{let} \text{remove} = \text{compile-ce}\left[ e_3, \text{curp}, \text{curd} \right] \text{ in} \\
& \text{let} \text{setprin} = \text{compile-ce}\left[ e_4, \text{curp}, \text{curd} \right] \text{ in} \\
& \text{let} \_ = d := (\text{?d} \cup \text{add})/\text{remove} \text{ in} \\
& \text{let} \_ = p := \text{setprin} \text{ in} \\
& \# ! \\
& \text{else} \# ! \\
\text{compile-cecheck} & \left[ \text{check} : e_1 \, \text{add} : e_2 \, \text{remove} : e_3 \, \text{set}-!-\text{principal} : e_4 \, \text{closure-principal} : e_5 \, \text{closure-delegations} : e_6, p, d, s, \text{curp}, \text{curd} \right] = \\
& \left( \lambda s : \text{Unit} \Rightarrow \\
& \text{let} \_ = \text{cp} \leftarrow \lambda s : \text{Unit} \Rightarrow \text{compile-ce}\left[ e_5, \text{curp}, \text{curd} \right] \text{ in} \\
& \text{let} \_ = \text{cd} \leftarrow \lambda s : \text{Unit} \Rightarrow \text{compile-ce}\left[ e_6, \text{curp}, \text{curd} \right] \text{ in} \\
& \left( \lambda s : \text{Unit} \Rightarrow \\
& \text{let} \_ = \text{cp} \leftarrow \lambda s : \text{Unit} \Rightarrow \text{compile-ce}\left[ e_5, \text{curp}, \text{curd} \right] \text{ in} \\
& \text{let} \_ = \text{cd} \leftarrow \lambda s : \text{Unit} \Rightarrow \text{compile-ce}\left[ e_6, \text{curp}, \text{curd} \right] \text{ in} \\
& \left( \lambda s : \text{Unit} \Rightarrowight) \right)
\end{align*}
\]
A.6 Reduction semantics for authorization contract extensions

\[\text{compile-cee}(\text{check: } e_1 \text{ add: } e_2 \text{ remove: } e_3 \text{ scope: } e_4 \text{ set-principal?: } e_5 \text{ principal: } e_6 \text{ set!-principal: } e_7, p, d, s, c, p_c, c_p, c_d, c_{p_c}, c_{p_d}) = \lambda _\_ : \text{Unit.} \]

\[
\begin{align*}
\text{let } _\_ &= \text{curp} := p \text{ in} \\
\text{let } _\_ &= \text{curd} := d \cup ?s \text{ in} \\
\text{let } p_1 \geq p_2 \circ p_3 &= \text{compile-aceg}[e_1, \text{curp, curd, cp, cd}] \text{ in} \\
\text{if } (\text{curd}) : p_3 \vdash p_1 \geq p_2 \text{ then} \\
\text{let } add &= \text{compile-aceg}[e_2, \text{curp, curd, cp, cd}] \text{ in} \\
\text{let } remove &= \text{compile-aceg}[e_3, \text{curp, curd, cp, cd}] \text{ in} \\
\text{let } setprin &= \text{compile-aceg}[e_7, \text{curp, curd, cp, cd}] \text{ in} \\
\text{let } _\_ &= d : = (?d \cup add)/\text{remove} \text{ in} \\
\text{let } _\_ &= p := setprin \text{ in} \\
\# \text{ if} \\
\end{align*}
\]

\[\text{compile-aee}_p(\text{check: } e_1 \text{ add: } e_2 \text{ remove: } e_3 \text{ scope: } e_4 \text{ set-principal?: } e_5 \text{ principal: } e_6 \text{ set!-principal: } e_7, p, c, p_c, c_p, c_d) =
\]

\[\text{compile-aee}_p(\text{check: } e_1 \text{ add: } e_2 \text{ remove: } e_3 \text{ scope: } e_4 \text{ set-principal?: } e_5 \text{ principal: } e_6 \text{ set!-principal: } e_7, s, c, p_c, c_p, c_d) =
\]

\[\text{compile-acee}(\text{let } x = c_{ee1} \text{ in } c_{ee2}, \text{curp, curd}) = \text{let } x = \text{compile-acee}[c_{ee1}, \text{curp, curd}] \]

\[\text{compile-acee}(\text{let } x = a_{ee1} \text{ in } a_{ee2}, \text{curp, curd, cp, cd}) = \text{let } x = \text{compile-acee}[a_{ee1}, \text{curp, curd, cp, cd}] \]

\[\text{compile-acee}(\text{current-principal, curp, curd}) = \text{?curp} \]

\[\text{compile-acee}(\text{current-principal, curp, curd}) = \text{?curp} \]

\[\text{compile-acee}(\text{current-delegations, curp, curd}) = \text{?curd} \]

\[\text{compile-acee}(\text{current-delegations, curp, curd}) = \text{?curd} \]

\[\text{compile-acee}(\text{closure-principal, curp, curd}) = \text{?cp} \]

\[\text{compile-acee}(\text{closure-delegations, curp, curd}) = \text{?cd} \]

\[\text{compile-acee}(e, \text{curp, curd}) = e \]

\[\text{compile-acee}(e, \text{curp, curd}) = e \]

A.5 Authorization contract extension evaluation contexts

\[E ::= \ldots \mid E \vdash e \mid v \vdash e \mid E ; e \vdash e \mid v \vdash E \mid v \vdash v \vdash E \]

\[\mid E \vdash e \circ E \mid v \vdash E \circ e \mid v \vdash v \vdash E \mid \{E\} \mid E \cup e \mid v \cup E \mid E \backslash e \mid v \backslash E \]

(A.6) Reduction semantics for authorization contract extensions

\[\langle E[_{\text{new-principal}}], \sigma \rangle \rightarrow \langle E[p], \sigma \rangle \text{ where } p \text{ is fresh} \]

\[\langle E[_{\text{new-dimension}}], \sigma \rangle \rightarrow \langle E[d], \sigma \rangle \text{ where } d \text{ is fresh} \]

\[\langle E[\{p_1 \geq p_1 \circ p_1, \ldots ; p_n \geq p_n\}], \sigma \rangle \rightarrow \langle E[#t], \sigma \rangle \text{ if } \{p_1 \geq p_1 \circ p_1, \ldots ; p_n \geq p_n\} \]

\[\langle E[\{p_1 \geq p_1 \circ p_1, \ldots ; p_n \geq p_n\}], \sigma \rangle \rightarrow \langle E[#t], \sigma \rangle \text{ if } \{p_1 \geq p_1 \circ p_1, \ldots ; p_n \geq p_n\} \]

\[\langle E[\{v_{11}, \ldots ; v_{1n}\} \cup \{v_{21}, \ldots ; v_{2m}\}], \sigma \rangle \rightarrow \langle E[v_{31}, \ldots ; v_{3k}], \sigma \rangle \text{ where } \{v_{31}, \ldots ; v_{3k}\} = \{v_{11}, \ldots ; v_{1n}\} \cup \{v_{21}, \ldots ; v_{2m}\} \]

\[\langle E[\{v_{11}, \ldots ; v_{1n}\} \backslash \{v_{21}, \ldots ; v_{2m}\}], \sigma \rangle \rightarrow \langle E[v_{31}, \ldots ; v_{3k}], \sigma \rangle \text{ where } \{v_{31}, \ldots ; v_{3k}\} = \{v_{11}, \ldots ; v_{1n}\} \backslash \{v_{21}, \ldots ; v_{2m}\} \]

\[\langle E[\{v \circ f(v)\}], \sigma \rangle \rightarrow \langle E[v], \sigma \rangle \]

\[\langle E[\{v_1, v_2, \ldots ; v_f(v)\}], \sigma \rangle \rightarrow \langle E[\{v_2, \ldots ; v_f(v_1)\}], \sigma \rangle \]

\[\langle E[\text{let } x_a \geq x_1 \circ x = v_a \geq v_1 \circ v_r \text{ in } e], \sigma \rangle \rightarrow \langle E[v_1, v_r, v_a / x_1, x, x_a], e], \sigma \rangle \]