Secondary use of data in EHR systems

Fan Yang∗ Chris Hankin† Flemming Nielson ‡ Hanne Riis Nielsen§

December 21, 2013

Abstract

We show how to use aspect-oriented programming to separate security and trust issues from the logical design of mobile, distributed systems. The main challenge is how to enforce various types of security policies, in particular predictive access control policies - policies based on the future behavior of a program. A novel feature of our approach is that advice is able to analyze the future use of data. We consider a number of different security policies, concerning both primary and secondary use of data, some of which can only be enforced by analysis of process continuations.

1 Introduction

Whilst there is broad agreement that security and other non-functional properties should be designed into systems ab initio it is also recognized that, as society becomes more IT-savvy, our expectations about security and privacy evolve. This is usually followed by changes in regulation in the form of standards and legislation. Thus, whilst we would still argue that security should feature in the initial design of a system, there is merit in separating out security and other non-functional properties so that they can be updated without disturbing the functional aspects of the system.

This paper focuses on designing a language for specifying policies for access control and explicit flow of information. The traditional approach to enforcing such security policies is to use a reference monitor [1] that dynamically tracks the execution of the program; it makes appropriate checks on each basic operation being performed, either blocking the operation or allowing it to proceed. In concrete systems this is implemented as part of the operating system or as part of the interpreter for the language at hand (e.g. the Java byte code interpreter); in both cases as part of the trusted computing base. Sometimes it is found to be more cost effective to systematically modify the code so as to explicitly perform the checks that the

∗DTU Informatics, Technical University of Denmark, Lyngby, Denmark
†ISST, Imperial College London, London United Kingdom
‡DTU Informatics, Technical University of Denmark, Lyngby, Denmark
§DTU Informatics, Technical University of Denmark, Lyngby, Denmark
reference monitor would otherwise have imposed \[2\]. In any case, even small modification in the security policies may involve substantial changes in the code or the underlying system.

The notion of *aspect-oriented programming* \[3, 4\] is an interesting approach to separation of concerns. The enforcement of security policies is an obvious candidate for such separation of concerns, e.g., because the security policy can be implemented by more skilled or more trusted programmers, or indeed because security considerations can be retrofitted by (re)defining *advice* to suit the (new) security policy. The detailed definition of the advice will then make decisions about how to possibly modify the operation being trapped. This calls for a modified language (like AspectJ \[3\] for Java) that supports the use of aspects and where a notion of trapping operations and applying advice has been incorporated. It is possible to systematically modify the code so as to explicitly perform the operations that the advice would have imposed (e.g. \[3\]).

In many cases the aspect-oriented approach provides a more flexible way for dealing with modifications in security policies \[5, 6, 7, 8, 9\] than the use of reference monitors. It facilitates the use of frameworks for security policies that may be well suited to the task at hand but that are perhaps not of general applicability and therefore not appropriate for incorporating into a reference monitor.

**Outline of the paper.** In this paper we are primarily interested in the modelling of mobile, distributed systems. The work is based on the coordination language KLAIM \[10\] (reviewed in Section 2) that facilitates distribution of data, mobility of code and handling of dynamically evolving, open systems. The main contribution of the paper is the design of AspectKE, an aspect-oriented extension of KLAIM that facilitates the trapping of actions (presented in Section 3) as well as processes (presented in Section 5), which can enforce traditional access control policies (presented in Section 4), as well as predictive access control policies (presented in Section 6)—i.e., security policies based on the future behavior of a program.

To evaluate our language design we shall throughout the paper illustrate its features using a running example based on a health information system for a care facility for the elderly people in New South Wales, Australia \[11\]. In Section 4 we show how to use AspectKE to enforce basic *primary use of data policies*, that is, policies concerned with the right to access data. Here we consider the three classical access control models, namely discretionary access control, mandatory access control and role-based access control \[1\]. Furthermore, we illustrate how multiple security policies can be integrated into existing systems thereby allowing policies to be refined at later stages in the system development. In Section 6 we show how *secondary use of data* policies can be modelled; these policies are concerned with how data is used once it has been obtained \[12\]. They can be enforced by using predictive access control policy enforcement mechanism offered by AspectKE. Here we exploit the ability to analyze not only the behavior of remotely executed processes but also the future use of data. To the best of our knowledge few, if any, proposals have ever used aspect oriented programming to tackle secondary use of data policies and provide a predictive access control policy enforcement mechanism.

Finally, in Section 7 we present related work and we conclude in Section 8 with a discussion of
the experience gathered from a proof-of-concept implementation \cite{13} and outline some future work.

2 Background: KLAIM

AspectKE is an extension of the KLAIM \textit{(Kernel Language for Agents Interaction and Mobility)} coordination language \cite{10} with support for aspect oriented programming. In this section we will review the fragment of KLAIM that will be used for AspectKE in the following section.

KLAIM is a language specifically designed to program distributed systems consisting of several mobile components that interact through multiple distributed tuple spaces (databases). KLAIM uses a Linda-like generative communication model \cite{14} but, instead of using Linda’s global shared tuple space (shared database), KLAIM associates a local tuple space with each node of a net. Each node may also have processes associated with it; the KLAIM computing primitives allow programmers to distribute and retrieve data and processes to and from locations (nodes) of a net, evaluate processes at remote locations and introduce new locations to the net.

2.1 Syntax of KLAIM

The syntax of a fragment of KLAIM is displayed in Table \ref{table1}.

A net (in \textbf{Net}) is a parallel composition of located processes and/or located tuples. For simplicity, components of tuples can be location constants only.\footnote{Compared with the original KLAIM, we do not allow processes to be components of tuples.} We use the notation $\overrightarrow{l}$ to represent a sequence of location constants and $\epsilon$ is used to represent the empty sequence. Nets must be closed: all variables must be in the scope of a defining occurrence (indicated by an exclamation mark).

A process (in \textbf{Proc}) can be a parallel composition of processes, a guarded sum of action prefixed processes, or a replicated process (indicated by the $\ast$ operator). We write $0$ for a
nullary sum, \( a.P \) for a unary sum, and \( a_1.P_1 + a_2.P_2 \) for a binary sum.

An action (in \texttt{Act}) operates on locations, tuples and processes: a tuple can be output to, input from (read and delete the source) and read from (read and keep the source) a location; processes can be spawned at a location; new locations can also be created. The actual operation performed by an action is called a \textit{capability} (in \texttt{Cap}) – this is a key concept when formalizing uses of data later. We do not distinguish real locations and data: all of them are called locations (in \texttt{Loc}) in our setting, which can be location constants \( l \), defining (i.e. binding) occurrences of location variables \( !u \) (where the scope is the entire process to the right of the occurrence), and use of location variables \( u \).

**Well-Formedness of Locations and Actions.** We do not allow multiple defining occurrences of the same variable in an action. We also prohibit bound variables and free variables from sharing any name in a single action. Thus we disallow \texttt{in}(!\(u, !u\)@\(l\)) as well as \texttt{in}(!\(u, u\)@\(l\)).

### 2.2 Semantics of KLAIM

Informally the meaning of a KLAIM program is as follows:

1. a node is selected for the next step of execution
2. if the process at the node is a choice, then one of the enabled choices is chosen non-deterministically and executed as described in the following four steps
3. if the prefix of the process is an output action, the output is performed
4. if the prefix of the process is an input (either destructive or non-destructive), the input action is enabled if there is a matching tuple at the target location, and the input is performed and appropriate variables are bound in the remainder of the process
5. if the prefix is an eval, the process is spawned at the target location
6. if the prefix is a newloc, the network is dynamically extended with a new location and the continuation process is given the address of that location
7. then return to Step 1

Notice that we do not need to deal with parallelism and replication within nodes because, at the cost of having duplicate addresses in the network, these can be lifted to the net level.

### 2.3 Running Example

Health Care Information Systems are gradually becoming prevalent and indispensable to our society. An electronic health record (EHR), part of a system’s database, stores a patient’s data and is created, developed, and maintained by the health care providers.
To illustrate the use of KLAIM, we now introduce a typical EHR system, which is inspired by [11], and the scenario presented here is used throughout the paper.

The EHR database (EHDB) stores all patient healthcare records and we assume that there are two types of data recorded for each patient: medical records (MedicalRecord) and private notes (PrivateNote). Medical records are entries created by doctors and so are the private notes; however the latter are of a more confidential nature. Also we distinguish between past records (Past) that have been entered into the EHR system previously and recent records (Recent) that have been created since the patient was admitted to the hospital. We therefore assume that the EHR database contains tuples with the following five fields:

| Field          | Description                                           |
|----------------|-------------------------------------------------------|
| patient        | The name of the patient                               |
| recordtype     | The type of record: MedicalRecord or PrivateNote      |
| author         | The author of the record                              |
| createdtime    | The time of creation of the record: Recent or Past    |
| subject        | The record’s content                                  |

For example ⟨Alice, MedicalRecord, DrSmith, Recent, text⟩ is a recent medical record of Alice, created by DrSmith and it has content text.

Doctors and nurses, as well as the patient, can access a patient’s record. We model these actors as locations in a network; the process at the location represents the actions of the individual and the data is the individual’s local “knowledge”. As an example the following process expresses that DrSmith reads one of the Past medical records for Alice created by DrHansen before she was admitted to this hospital, writes some of the information in her own note (in location DrSmith) and then creates a new medical record for the patient:

\[
\text{DrSmith} :: \quad \text{read}(\text{Alice}, \text{MedicalRecord}, \text{DrHansen}, \text{Past}, !\text{content})@\text{EHDB}.
\]

\[
\quad \text{out}(\text{Alice, content})@\text{DrSmith}.
\]

\[
\quad \text{out}(\text{Alice, MedicalRecord, DrSmith, Recent, newtext})@\text{EHDB}
\]

Here DrSmith will first consult location EHDB and read a five-tuple whose first four components are Alice, MedicalRecord, DrHansen, and Past respectively and the corresponding fifth component is assigned to variable content. The second action will write the content read at the first action to the location associated with DrSmith. The final construct will write a new five-tuple to location EHDB for this patient whose last three components indicate that the author is DrSmith, it is a Recent medical record and the content is newtext.

To illustrate the semantics of KLAIM let us consider the following net, consisting of locations EHDB and DrSmith:

\[
\text{EHDB} :: \langle \text{Alice, MedicalRecord, DrHansen, Past, alicetext} \rangle
\]

\[
\quad || \quad \text{EHDB} :: \langle \text{Bob, PrivateNote, DrJensen, Recent, bobtext} \rangle
\]

\[
\quad || \quad \text{DrSmith} :: \quad \text{read}(\text{Alice, MedicalRecord, DrHansen, Past, !content})@\text{EHDB}.
\]

\[
\quad \quad \text{out}(\text{Alice, content})@\text{DrSmith}.
\]

\[
\quad \quad \text{out}(\text{Alice, MedicalRecord, DrSmith, Recent, newtext})@\text{EHDB}
\]
The execution may proceed as follows:

\[
\text{EHDB} :: \langle \text{Alice}, \text{MedicalRecord}, \text{DrHansen}, \text{Past}, \text{alicetext} \rangle
\]

\[
\text{DrSmith} :: \text{read}(\text{Alice}, \text{MedicalRecord}, \text{DrHansen}, \text{Past}, \text{!content})@\text{EHDB}.
\]

\[
\text{out}(\text{Alice}, \text{content})@\text{DrSmith}.
\]

\[
\text{out}(\text{Alice}, \text{MedicalRecord}, \text{DrSmith}, \text{Recent}, \text{newtext})@\text{EHDB}
\]

\[
\rightarrow
\text{EHDB} :: \langle \text{Alice}, \text{MedicalRecord}, \text{DrHansen}, \text{Past}, \text{alicetext} \rangle
\]

\[
\text{DrSmith} :: \text{out}(\text{Alice}, \text{alicetext})@\text{DrSmith}.
\]

\[
\text{out}(\text{Alice}, \text{MedicalRecord}, \text{DrSmith}, \text{Recent}, \text{newtext})@\text{EHDB}
\]

\[
\rightarrow
\text{EHDB} :: \langle \text{Alice}, \text{MedicalRecord}, \text{DrHansen}, \text{Past}, \text{alicetext} \rangle
\]

\[
\text{DrSmith} :: \text{out}(\text{Alice}, \text{alicetext})@\text{DrSmith}.
\]

\[
\text{DrSmith} :: \langle \text{Alice}, \text{alicetext} \rangle
\]

DrSmith first reads the tuple \( \langle \text{Alice}, \text{MedicalRecord}, \text{DrHansen}, \text{Past}, \text{alicetext} \rangle \) from \( \text{EHDB} \); the binding of the variable \text{content} is reflected in the continuation of the process. In the second step \text{DrSmith} outputs a tuple that consists of \text{Alice} together with a bound \text{content} (alicetext) to her own tuple space. In the final step, a new tuple that represents a new medical record is written to location \( \text{EHDB} \).

### 3 AspectKE: Trapping Actions

We now show how to integrate aspects into KLAIM by presenting the basic features of AspectKE, with a focus on how aspects trap actions in a KLAIM program. We consider a global set of aspects.

#### 3.1 Syntax

The Syntax of AspectKE is given by Tables 2 and 1 (the KLAIM syntax).

Table 2 introduces a system \( S \) (in System) that consists of a net \( N \) and a sequence of global aspect declarations \( \overrightarrow{asp} \). An aspect declaration (in Asp) takes the form \( A[cut] \triangleq body \): \( A \) is the aspect name, and \( body \) (in Advice) is the advice to the trapped action. Each action (the Act in Table 1) is a potential join point that can be intercepted by AspectKE’s pointcut (in Cut).
Moreover, _ is introduced as a don’t-care parameter in the cut version of actions, and in the test primitive of conditional expressions (BExp). It can match any type of location used in the program. Note in the cut, the occurrence of !_u and u have different meaning from those of KLAIM; a plain variable in a pointcut can only match an actual location and banged (!) variables in the pointcut can only match against binding occurrences of variables, while the don’t-care (_) can match both in the join point.

Each aspect gives a unique run-time suggestion (either break or proceed) which may depend on the evaluation of a conditional expression. The suggestion break suppresses the trapped action whilst proceed allows the trapped action to be executed. In case of multiple aspects that trap an action, break takes precedence over proceed. The primitive test(ℓ_t0)@ℓ evaluates to tt if a tuple exists in the tuple space of ℓ which matches ℓ_t0. Besides basic boolean expressions, condition cond also includes bounded existential quantification and universal quantification – this allows simple queries to the databases occurring in the nets.

In contrast to other aspect languages, the condition is part of the advice instead of being part of the pointcut (being evaluated before intercepting a join point). Evaluating the condition after intercepting a join point allows a more natural modelling of security policies.

### Well-formedness of Cuts.
In addition to the well-formedness conditions for KLAIM, we require that the variables in a cut are pairwise distinct. We shall also impose that aspects are closed: any free variable in the body is defined in the cut. Additionally, when !_u is used in a cut pattern, u should not be used in conditions except in the context of set expressions.

**Example 1** To illustrate how aspects can be composed in AspectKE that work with the KLAIM
program, the following simple aspect gives advice to the running example in section 2.3.

\[ A_{1}^{\text{out}}[\text{user :: out}(\_\_\_\_\_ \_ data)@\text{DrSmith}] \]
\[ \triangleq \text{case}(\text{data} = \text{alicetext}) \]
\[ \text{break; proceed} \]

The aspect traps an \texttt{out} action of processes running at location \texttt{DrSmith} that attempt to send a tuple with two fields. If the actual value of the second field is equal to \texttt{alicetext}, the aspect will break the execution of the action and its continuation process. Otherwise, the action continues.

3.2 Semantics

The base semantics is that of KLAIM (Section 2) but now, before executing an action (all actions in a KLAIM program are potential join points), we check to see if any aspect applies to the action and combine the advice of all applicable aspects. Each advice is either that the action be allowed to proceed or not. We resolve possible conflicts by ensuring that any aspect that disallows an action has priority. Aspects are applied in definition order but, because aspects can only allow or disallow the join point to proceed, the order is actually immaterial.

\textbf{Example 2} Suppose we have a system that contains the same net as in running example of Section 2.3 and aspect \( A_{1}^{\text{out}} \) in Example 1:

\[ A_{1}^{\text{out}}[\text{user :: out}(\_\_\_\_\_ \_ data)@\text{DrSmith}] \]
\[ \triangleq \text{case}(\text{data} = \text{alicetext}) \]
\[ \text{break; proceed} \]

\begin{align*}
\text{let} & \quad \text{in} \\
\text{EHDB ::}< & \quad \text{Alice, MedicalRecord, DrHansen, Past, alicetext}\rangle \\
\mid \mid & \quad \text{EHDB ::}< & \quad \text{Bob, PrivateNote, DrJensen, Recent, bobtext}\rangle \\
\mid \mid & \quad \text{DrSmith ::}< & \quad \text{read}(\text{Alice, MedicalRecord, DrHansen, Past, !content})@\text{EHDB}. \quad \text{out}(\text{Alice, content})@\text{DrSmith}. \quad \text{out}(\text{Alice, MedicalRecord, DrSmith, Recent, newtext})@\text{EHDB}
\end{align*}

and some steps of execution (omitting the aspect definition):
Aspect $A^\text{out}_1$ does not trap the read action, thus the read action executes and binds content with alicetext. But $A^\text{out}_1$ traps the first out action, and the result substitution is

$$[\text{DrSmith}/\text{user}, \text{alicetext}/\text{data}]$$

and the case condition \(\text{data} = \text{ Alicetext}\) evaluates to \(\text{tt}\), thus the aspect breaks the execution of this action and its continuation process.

4 Worked Examples: Advice for Access Control Models

To evaluate the expressiveness of the language and show its language features, we now show how AspectKE can be used to enforce access control policies by utilizing three well-known access control models, namely discretionary access control (Section 4.1), mandatory access control (Section 4.2) and role-based access control (Section 4.3), and how AspectKE can introduce new aspects for retrofitting new policies to existing systems (Section 4.4).

Since patient confidentiality is an important issue in the health care industry it is imperative that EHRs are protected [15]. To help achieve this goal, governments define many types of security policies, encapsulated in various acts and guides (e.g. [16, 17]). Throughout the paper, we will enforce several security policies for the EHR system that was introduced in Section 2.3 and this shows different features of the language.

The first is a primary use of data policy inspired by [11] which regulates the basic access control concerning the read and write rights owned by doctors and nurses:

Doctors can read all patients’ medical records and private notes, while nurses can read all patients’ medical records but cannot read any private notes. Medical records and private notes can only be created by doctors.

For simplicity, here we restrict ourselves to only focussing on read, in and out actions, while eval and newloc actions will be discussed further when enforcing other security policies.
4.1 Discretionary Access Control

We will show how to enforce the above policy with discretionary access control (DAC), which is a type of access control as a means of restricting access to objects based on the identity of subjects and/or the groups to which they belong[18]. We do so by using an access control matrix containing triples \((s, o, c)\) identifying which subjects \(s\) can perform which operations \(c\) on which objects \(o\). If we use the KLAIM programming model, we should equip the semantics of KLAIM with a reference monitor that consults the access control matrix when an action is executed to check if the action is permitted. In AspectKE we can directly use aspects to elegantly inline the reference monitor to enforce this discretionary access control policy.

**Example 3** The access control matrix is stored in location DAC, which contains tuples: \((user, recordtype, capability)\). For example, if DrSmith is a doctor and NsOlsen is a nurse, then DAC might contain the following tuples:

- \((\text{DrSmith}, \text{MedicalRecord}, \text{read})\)
- \((\text{DrSmith}, \text{PrivateNote}, \text{read})\)
- \((\text{DrSmith}, \text{MedicalRecord}, \text{out})\)
- \((\text{DrSmith}, \text{PrivateNote}, \text{out})\)
- \((\text{NsOlsen}, \text{MedicalRecord}, \text{read})\)

We also assume that the location DAC can only be modified by privileged users, thus doctors and nurses cannot perform any in and out action on it. This can be enforced by other aspects but we omit them here.

The following aspect declarations will impose the desired requirements.

\[
A^{\text{read}}_{\text{p1A1}}[\text{user} :: \text{read}(\_\_\_\_, \_\_\_\_, \text{read}) \@ \text{DAC}] = \text{case}(\text{test}(\text{user}, \_\_\_\_, \text{read}) \@ \text{DAC}) \\text{proceed; break}
\]

\[
A^{\text{in}}_{\text{p1A2}}[\text{user} :: \text{in}(\_\_\_\_, \_\_\_\_, \text{in}) \@ \text{DAC}] = \text{case}(\text{test}(\text{user}, \_\_\_\_, \text{in}) \@ \text{DAC}) \\text{proceed; break}
\]

\[
A^{\text{out}}_{\text{p1A3}}[\text{user} :: \text{out}(\_\_\_\_, \_\_\_\_, \text{out}) \@ \text{DAC}] = \text{case}(\text{test}(\text{user}, \_\_\_\_, \text{out}) \@ \text{DAC}) \\text{proceed; break}
\]

Aspects \(A^{\text{read}}_{\text{p1A1}}, A^{\text{in}}_{\text{p1A2}}, \) and \(A^{\text{out}}_{\text{p1A3}}\) enforce the above policy by using DAC, where the access rights for each user are actually described. Note that the second field of the tuple operated by these cut actions is recordtype, which trap an action that clearly specifies a concrete record type.

Consider the following KLAIM program that is a variant of the running example in Section 2.3 (in that the user is nurse NsOlsen instead of doctor DrSmith) and is equipped with the above four aspects:

\[
\text{NsOlsen} :: \text{read}(\text{Alice, MedicalRecord, DrHansen, Past, !content}) \@ \text{EHDB}.
\text{out}(\text{Alice, content}) \@ \text{NsOlsen}.
\text{out}(\text{Alice, MedicalRecord, NsOlsen, Recent, newtext}) \@ \text{EHDB}
\]
The first read action will be trapped by aspect $A_{p1a1}^{\text{read}}$, and the resulting substitution is

$$[\text{NsOlsen}/user, \text{MedicalRecord}/\text{recordtype}]$$

and the condition $\text{test}(\text{NsOlsen}, \text{MedicalRecord}, \text{read})@\text{DAC}$ is evaluated. Since NsOlsen has the appropriate right according to DAC we proceed and perform this read action thereby giving rise to the binding of content to alicetext.

The second action will not be trapped by any of the aspects, so it will simply be performed and the tuple $(\text{Alice, alicetext})$ is output to location NsOlsen.

The last action will be trapped by aspect $A_{p1a3}^{\text{out}}$ and after the substitution we evaluate the condition $\text{test}(\text{NsOlsen}, \text{MedicalRecord}, \text{out})@\text{DAC}$ which is evaluated to $\text{ff}$ and thus we break the execution.

However, the KLAIM program can also execute read or in actions without specifying the record type, e.g., using $!\text{recordtype}$ instead of $\text{recordtype}$, users can thus get a record as follows:

\[ \text{NsOlsen :: read}(\text{Alice}, !\text{recordtype}, \text{DrHansen}, \text{Past}, !\text{content})@\text{EHDB} \]

where a successful input action can retrieve any type of EHR record.

None of the above aspects can trap these input actions, thus we have to enforce additional aspects so that the above input actions will not bypass our aspects and consequently break the policy. The simple aspects forbid any attempts to read or in (read and then remove) EHR records without specifying the record type:

$$A_{p1a1}^{\text{read}} [\text{user} :: \text{read}(\_, !\text{recordtype}, \_, \_\_, \_)@\text{EHDB}] \triangleq \text{break}$$

$$A_{p1a3}^{\text{in}} [\text{user} :: \text{in}(\_, !\text{recordtype}, \_, \_\_, \_\_)@\text{EHDB}] \triangleq \text{break}$$

One may wonder why not build the above two aspects on top of aspects $A_{p1a1}^{\text{read}}$ and $A_{p1a2}^{\text{in}}$ by directly replacing $\text{recordtype}$ with $!\text{recordtype}$ in their pointcut, respectively. The reason is that these aspects will not be well-formed: when trapping actions, $\text{recordtype}$ binds with a variable, which cannot be used in a test condition such as $\text{test}(\text{user, recordtype, read})@\text{DAC}$.

### 4.2 Mandatory Access Control

In this subsection we will show how to enforce the above policies by using mandatory access control (MAC), which is a means of restricting access to objects based on the sensitivity (as represented by a label) of the information contained in the objects and the formal authorization (i.e., clearance) of subjects[15]. Before enforcing the above policy, we first impose a comparable classical MAC policy - the Bell-LaPadula security policy[1] based on a mandatory access control model. Later we enforce the above policy as a variant of the Bell-LaPadula policy. In the presentation, security levels are assigned to subjects (as clearances) and objects (as labels).
**Example 4** In this scenario, we just need two security levels, and may assign security levels to subjects as follows: doctors have level high and nurses have level low; similarly we may assign objects as follows: private notes have level high and medical records have level low.

To model this policy we need to introduce a location MAC that stores tuples of the form: \( (\text{user}, \text{securitylevel}) \) and \( (\text{recordtype}, \text{securitylevel}) \). Continuing Example 3, we create the location MAC with the tuples:

- \( (\text{DrSmith}, \text{high}) \)
- \( (\text{NsOlsen}, \text{low}) \)
- \( (\text{PrivateNote}, \text{high}) \)
- \( (\text{MedicalRecord}, \text{low}) \)

As before we also assume that the location MAC can only be modified by privileged users.

Firstly, we enforce the Bell-LaPadula security policy [1] to illustrate that AspectKE can enforce a well-known mandatory access control policy. Then we will enforce our example policy, with small modifications based on the aspects that enforce Bell-LaPadula policy.

If we enforce the Bell-LaPadula security policy, the first part of the policy states that a subject is allowed to read or input data from any object provided that the subject’s security level dominates that of the object. In our case, this guarantees no read up: that is, low subjects (nurses) cannot read high objects (private notes) but can only read low objects (medical records); however, high subjects (doctors) can access both kinds of records.

The no read up part of the policy can be enforced by aspects as follows:

\[
A^{\text{read}}_{\text{pl1}}[\text{user} :: \text{read}(\_ , \_ , \_ , \_ )@\text{EHDB}] \\
\triangleq \text{case}(\lnot (\text{test(user,low)}@\text{MAC} \land \text{test(recordtype,high)}@\text{MAC})) \\
\text{proceed}; \\
\text{break}
\]

The second part of the policy (a simplified form of Bell-LaPadula star property [2]) states that a subject can write to any object provided that the security level of the object dominates that of the subject (no write down). In our case high subjects (doctors) cannot write low objects (medical records) but low subjects (nurses) can write to both kinds of records.

The no write down of the policy can be enforced by the aspect below:

\[
A^{\text{out}}_{\text{pl1}}[\text{user} :: \text{out}(\_ , \_ , \_ , \_ )@\text{EHDB}] \\
\triangleq \text{case}(\lnot (\text{test(user,high)}@\text{MAC} \land \text{test(recordtype,low)}@\text{MAC})) \\
\text{proceed}; \\
\text{break}
\]

Additionally, we have an aspect for the read action to prevent users from reading records without specifying the record type, and an aspect for the in action to prevent users from reading and deleting records:

\[
A^{\text{read}}_{\text{pl1}}[\text{user} :: \text{read}(\_ , \_ , \_ , \_ )@\text{EHDB}] \\
\triangleq \text{break}
\]

\[
A^{\text{in}}_{\text{pl1}}[\text{user} :: \text{in}(\_ , \_ , \_ , \_ )@\text{EHDB}] \\
\triangleq \text{break}
\]
These aspects correctly enforce our policy about reading patient records. However, the no write down policy is not quite right for our example, instead we depart from the Bell-LaPadula policy and define:

\[ A_{p1_{L2}}^{\text{out}} \{ \text{user} :: \text{out}(\text{recordtype, } \text{high}) @ EHDB \} \]

\[ \triangleq \text{case}(\text{test}(\text{user}, \text{high}) @ \text{MAC}) \]

\[ \text{proceed;} \]

\[ \text{break} \]

This aspect allows doctors to write any kind of record.

The aspect \( A_{p1_{L2}}^{\text{out}} \) together with \( A_{p1_{L3}}^{\text{read}}, A_{p1_{L3}}^{\text{read}}, A_{p1_{L4}}^{\text{in}} \) reflect a mandatory access control model which satisfies our policy. In this case we only allow high users (doctors) to write patient records. Hence nurse NsOlsen in Example 4 cannot execute the third action as it will be blocked by \( A_{p1_{L2}}^{\text{out}} \), which would be allowed with \( A_{p1_{L2}}^{\text{out}} \) from the Bell-LaPadula security policy.

### 4.3 Role-Based Access Control

Role-based access control (RBAC) [14] is another access control mechanism which allows the central administration of security policies and is often more flexible and elegant for modelling security policies. The simplest model in the RBAC family is \( RBAC_0 \), where there are three sets of entities called user, role, and permission. A user can be assigned multiple roles (role assignment) and a role can have multiple permissions (permission assignment) to corresponding operations. In addition, the user can initiate a session during which the user activates some subset of roles that he or she has been assigned. A user can execute an operation only if the user’s active roles have the permission to perform that operation.

**Example 5** To implement the security policy for patient records, we use a model that does not differentiate a user’s assigned role and active role (we assume that the assigned roles of all users are activated by default), so we only need location RDB with tuples (user, role):

- \( \langle \text{DrSmith, Doctor} \rangle \)
- \( \langle \text{NsOlsen, Nurse} \rangle \)

For permission assignment we also need a location to describe each role’s permission. This can be done by storing tuples (role, object, capability) at PDB:

- \( \langle \text{Doctor, MedicalRecord, read} \rangle \)
- \( \langle \text{Doctor, PrivateNote, read} \rangle \)
- \( \langle \text{Doctor, MedicalRecord, out} \rangle \)
- \( \langle \text{Doctor, PrivateNote, out} \rangle \)
- \( \langle \text{Nurse, MedicalRecord, read} \rangle \)

Once more we assume that the locations RDB and PDB can only be modified by privileged users.
The following aspects then implement the required policy:

\[
A_{\text{plc1}}^{\text{read}}[\text{user} :: \text{read}(_, \text{recordtype}, _, _, _) @ \text{EHDB}] \\
\triangleq \text{case}(\exists \text{role} \in \{ \text{Doctor}, \text{Nurse} \} : \\
(\text{test}(\text{user}, \text{role}) @ \text{RDB} \land \text{test}(\text{role}, \text{recordtype}, \text{read}) @ \text{PDB})) \\
\text{proceed}; \\
\text{break}
\]

\[
A_{\text{plc1}}^{\text{in}}[\text{user} :: \text{in}(_, \text{recordtype}, _, _, _) @ \text{EHDB}] \\
\triangleq \text{case}(\exists \text{role} \in \{ \text{Doctor}, \text{Nurse} \} : \\
(\text{test}(\text{user}, \text{role}) @ \text{RDB} \land \text{test}(\text{role}, \text{recordtype}, \text{in}) @ \text{PDB})) \\
\text{proceed}; \\
\text{break}
\]

\[
A_{\text{plc1}}^{\text{out}}[\text{user} :: \text{out}(_, \text{recordtype}, _, _, _) @ \text{EHDB}] \\
\triangleq \text{case}(\exists \text{role} \in \{ \text{Doctor}, \text{Nurse} \} : \\
(\text{test}(\text{user}, \text{role}) @ \text{RDB} \land \text{test}(\text{role}, \text{recordtype}, \text{out}) @ \text{PDB})) \\
\text{proceed}; \\
\text{break}
\]

These three aspects are useful for interrupting the execution when a user attempts to operate on EHR records with a concrete record type, which essentially relies on the tuples from RDB and PDB. They also show the benefit of admitting quantifiers into the conditional expressions.

Similar to the previous subsections, we have to enforce additional aspects for capturing user attempts to access EHR records without specifying the record type.

\[
A_{\text{plc1}}^{\text{read}}[\text{user} :: \text{read}(_, !\text{recordtype}, _, _, _) @ \text{EHDB}] \\
\triangleq \text{break}
\]

\[
A_{\text{plc1}}^{\text{in}}[\text{user} :: \text{in}(_, !\text{recordtype}, _, _, _) @ \text{EHDB}] \\
\triangleq \text{break}
\]

In the following sections we will use the role-based access control mechanism since it is best suited for enforcing security policies in a large organization.

### 4.4 Advice for Retrofitting Policies

Now we will show how aspects in AspectKE can retrofit new security policies into an exiting system that is being developed/updated or has already been deployed. Concretely, when a new functionality has been introduced to the existing system we will show how we enforce new policies to cater for the new requirements (Section 4.4.1); when a policy has been proposed to refine existing policies, we will show how we enforce policies based on the existing functionality of the system (Section 4.4.2); on the other hand, sometimes additional functionality has to be introduced to the system to implement aspects which refine existing policies (Section 4.4.3).

Indeed, the possibility to refine, renew and retrofit security policies into an existing/evolving system is very important for IT systems. Taking the EHR system as an example, as the public debate about security standards (especially for privacy) evolves, governments have to modify the corresponding acts and guides. As a consequence, the security policy for an EHR system will undergo frequent change \textsuperscript{[20]}. Moreover, the IT system itself will always be enhanced
by new functionalities, which means that new policies need to be enforced. The National IT Strategy for the Danish Health Care Service states that “it is also important to acknowledge the fact that IT is not just implemented once and for all”[21].

4.4.1 Enforcing Security Policy for New Functionality

AspectKE can be used for enforcing new security policies when a new functionality has been introduced at any phase of the system development. The running example in Section 2.3 introduces the functionality of the EHR system as regards reading and writing from and to the EHR database (EHDB). Now we introduce another (new) function to the EHR system that enables a manager to add a patient, or delete information from the system.

In our programming model, each patient is represented by a location. A manager can add a new patient to the system as follows:

\[ \text{MgDavis :: newloc}(!\text{patient}).\text{out} (\text{patient}, \text{Patient})@\text{RDB} \]

First a new location for the patient is created, then it is registered in the role database RDB by the out action. To delete a user one can simply perform an in action to the location RDB.

We shall now show how to enforce the following security policy regarding the manager role at the hospital[11], for this new added functionality.

\[ \text{In the hospital, only managers are allowed to add a user to, or delete from the system.} \]

**Example 6** The following aspects will enforce the above security policy.

\[ A_{\text{newloc}}^{\text{p2}}[\text{user :: newloc}(\_)] \triangleq \text{case(test}(\text{user}, \text{Manager})@\text{RDB} \land \text{test}(\text{Manager}, \text{Location, newloc}@\text{PDB}) \]

\[ \quad \text{proceed;} \]

\[ \quad \text{break} \]

\[ A_{\text{out}}^{\text{p2}}[\text{user :: out}(\_\_)]@\text{RDB} \]

\[ \triangleq \text{case(test}(\text{user}, \text{Manager})@\text{RDB} \land \text{test}(\text{Manager}, \text{RDB, out}@\text{PDB}) \]

\[ \quad \text{proceed;} \]

\[ \quad \text{break} \]

\[ A_{\text{in}}^{\text{p2}}[\text{user :: in}(\_\_)]@\text{RDB} \]

\[ \triangleq \text{case(test}(\text{user}, \text{Manager})@\text{RDB} \land \text{test}(\text{Manager}, \text{RDB, in}@\text{PDB}) \]

\[ \quad \text{proceed;} \]

\[ \quad \text{break} \]

These aspects are composed in a way that is similar to the Example 5. Before using them we need to put the tuple (MgDavis, Manager) into RDB and the tuples (Manager, Location, newloc), (Manager, RDB, out), and (Manager, RDB, in) into PDB. Finally, once again we assume that PDB can only be modified by privileged users.

Example 6 shows that AspectKE can enforce a policy for new functionality (code), no matter when this functionality is developed, potentially in the entire life cycle of the EHR system.
This is because the underlying aspect-oriented mechanism allows new aspects to intercept all join points (including a join point of new code) and give appropriate advice. Example 6 also shows how to give advice on an action (newloc) that creates new node in a net. Note that Examples 5 and 6 use the role-based access control model which shows that the tuples introduced above at PDB are good at expressing permissions that only directly rely on roles and objects. However, some permission assignments are more complex and therefore we shall also use logical formulae to express permission assignments in the following sections.

4.4.2 Refining Security Policy with Existing Functionality

AspectKE can be used to refine an existing security policy at any phase of the system development. The following is a policy which can be considered as a refinement of the previous policy (enforced by Example 3-5) to protect patients’ privacy. This policy is also inspired by [11]:

Private notes can only be viewed on the basis of the doctor-patient confidentiality – doctors cannot view private notes that were not created by themselves; nurses can only view recent medical records which were created after the patient has been admitted.

Traditional programming paradigms normally necessitate modifying existing code to enforce this extra policy, while the aspect approach simply requires additional aspects.

Example 7 The first part of the policy can be expressed by the aspects shown below. These two aspects will prevent a doctor from reading a private note written by another doctor or reading a private note without clearly specifying the author of the note.

\[
A_{p31}^{\text{read}}[\text{user :: read}(\cdot, \text{PrivateNote}, \text{author}, \cdot, \cdot)@EHDB] \\
\triangleq \begin{cases} 
\text{case}(\text{test}(\text{user}, \text{Doctor}@RDB \land (\text{user} = \text{author})) \\
\text{break} \\
\text{case}(\neg \text{test}(\text{user}, \text{Doctor}@RDB)) \\
\text{break}
\end{cases}
\]

Note that the second case of \(A_{p31}^{\text{read}}\) and \(A_{p32}^{\text{read}}\) allow any users registered in RDB except doctors to read a private note, which includes users taking roles as nurses. Allowing nurses to read a private note is not problematic, as these two aspects only reflect the intention of this policy. Preventing nurses to read a private note is enforced in the previous policy.

These two aspects are supposed to work together with those aspects defined in the Examples 3 to 5. For example if a nurse tries to read a private note, these aspects will proceed the execution but aspects in Examples 3-5 will suggest break. Another example is if a doctor tries to read a private note written by other doctors, aspects in Examples 3-5 will allow the action.
to proceed whilst these aspects will break the execution. Since break takes precedence over proceed, the final decision suggested from this advice will be to block the execution in both cases.

The second part of the policy says that nurses can only read recent medical records which can be expressed by the following aspects where, once again, the aspects should be considered in conjunction with those of Examples 3 to 4. These two aspects will prevent a nurse from reading a past medical record or reading a medical record without specifying the created time.

\[
A_{\text{read}}^{p3_1} \triangleq \begin{cases} 
\text{case} & (\text{test(user, Nurse)} \land \text{createdtime} = \text{Recent}) \\
\text{proceed; break} \\
\text{case} & (\lnot \text{test(user, Nurse)}) \\
\text{proceed;} \\
\text{break} 
\end{cases}
\]

Note that no new functionality is required to enforce this policy, as these aspects only rely on the existing program (i.e., node RDB).

4.4.3 Refining Security Policy with New Functionality

Sometimes refining an existing security policy is necessary. We have to introduce additional functionality to support the implementation of the aspects for policy refinement. Now we consider the following security policy that restricts access to the database to certain locations in the hospital [22]:

A nurse can only read medical records of the patients who are in the wards located on the nurse’s working floor. Furthermore, the nurse can only access medical records through the computers that are located on that specific floor. But in the emergency room, a nurse does not have this restriction.

Example 8 To express this policy as aspects we shall assume that the current location database CLDB records every user’s current location information (indicating that they are using computers at that location), and that the assigned location database ALDB stores every user’s assigned room (e.g. for nurses this is their working floors and rooms which they are responsible for; for patients this is the floors and wards that they are on). These two databases store tuples with the same fields \(\langle \text{user}, \text{floor}, \text{room} \rangle\) and can only be modified by privileged users.

The set Floor contains the actual floors of the hospital (e.g. f1,f2). The set Room contains the actual rooms of the hospital and includes two types of rooms: ordinary wards (e.g. w1,w2) and the special room EmergencyRoom.
The appropriate advice can now be expressed as follows:

\[
A_{\text{read}}^{p4_1}[\text{user} :: \text{read}(\text{patient}, \text{MedicalRecord}, _, _, _)@\text{EHDB}]
\]
\[
\triangleq \text{case(test}(\text{user}, \text{Nurse})@\text{RDB}\land
\exists \text{floor} \in \text{Floor} : \text{test}(\text{user, floor, EmergencyRoom})@\text{CLDB})
\]
\[
\text{proceed};
\]
\[
\text{case(test}(\text{user, Nurse})@\text{RDB}\land
\exists \text{floor} \in \text{Floor} \exists \text{room} \in \text{Room} : ( \text{test}(\text{user, floor, room})@\text{CLDB}\land
\text{test}(\text{user, floor, room})@\text{ALDB}\land
\text{test}(\text{patient, floor, room})@\text{ALDB}))
\]
\[
\text{proceed};
\]
\[
\text{case}(\neg \text{test}(\text{user, Nurse})@\text{RDB})
\]
\[
\text{proceed};
\]
\[
\text{break}
\]
\[
A_{\text{read}}^{p4_2}[\text{user} :: \text{read}(!\text{patient}, \text{MedicalRecord}, _, _, _)@\text{EHDB}]
\]
\[
\triangleq \text{case(test}(\text{user}, \text{Nurse})@\text{RDB}\land
\exists \text{floor} \in \text{Floor} : \text{test}(\text{user, floor, EmergencyRoom})@\text{CLDB})
\]
\[
\text{proceed};
\]
\[
\text{case}(\neg \text{test}(\text{user, Nurse})@\text{RDB})
\]
\[
\text{proceed};
\]
\[
\text{break}
\]

In \(A_{\text{read}}^{p4_1}\), the first case caters for the situation where a nurse is working in the emergency room; the second case allows the read action when a nurse is trying to access the medical record for a patient who is on the same ward/floor as where the nurse is currently at and assigned to work, the third case allows a user who is not a nurse to perform the read action. The aspect \(A_{\text{read}}^{p4_2}\) is similar to the aspect \(A_{\text{read}}^{p4_1}\) except that it does not contain the second case. This is because reading a medical record without specifying the name of the patient is not acceptable for a nurse who is not working in the emergency room.

As in Example 7, these aspects will work together with all previously introduced security policies. Moreover, these aspects are both in the spirit of role-based access control, and they demonstrate that when composing aspects in AspectKE for larger and more complex security policies of an organization, role-based access control can be very efficacious, as has already been observed in the literature [19].

Note new functionality has to be introduced to enforce this policy such as the newly introduced nodes for databases (e.g., CLDB, ALDB). This new functionality can be developed as part of the main logic of the EHR system, or can be merely developed and maintained to enforce security policies. We observe that although we might need to extend the functionality of the system’s main logic to enforce certain security policies, the policies themselves are still described in aspects (even though they directly or indirectly rely on new nodes/processes), which are still separated out of the main logic.
5 AspectKE: Trapping Processes

5.1 Motivation

Classical reference monitors are incapable of enforcing security policies based on the future behavior of programs, rather they rely only on information gathered by monitoring execution steps \[23\], and perform history-based dynamic checks. However, security policies concerned with information flow that cannot be implemented correctly without a security check of the overall behavior of the program exist.

For example, in a software system, remote evaluation involves the transmission of programs from a client to a server for subsequent execution at the server. However, as the programs transmitted might perform unintended operations at the server side, a security check is usually needed. A typical example of this is Java applets which can be transferred to a remote system and executed by the Java Virtual Machine (JVM). Since the unknown applets are not always written by trusted users, the JVM has certain mechanisms for ensuring that the applet will not be able to do malicious actions, e.g., the bytecode verifier \[24\] and sandbox mechanism \[25\].

As another example, in the EHR domain, rather than enforcing the primary use of data policies for direct patient care domain as shown in the previous sections, there is a trend to define and enforce secondary use of data policies. Here data is used outside of direct health care delivery that includes activities such as analysis, research, public health etc, even though it still lacks a robust infrastructure of policies and is surrounded with complex ethical, political, technical and social issues \[12\]. Compared with primary use of data, whose focus is on regulating “someone has some rights to access some data”, it focuses on defining "how the data can be used after it is released to someone". The enforcement of such policies requires security checks in the form of inspection of the flow of data.

In general, program analysis techniques concern computing reliable, approximate information about the dynamic behavior of programs \[26\], and the derivation of useful information by simulating execution of all possible paths of the executing program. For example, type systems can be used to enforce various kinds of information flow as well as access control security policies (like Jif \[27\] for Java).

In this section, we extend the AspectKE language presented in Section \[3\] which enables AspectKE not only to trap the action, but also to trap a process that is to be executed in the future. The process can be a process that is to be sent to a remote site, or a process continuation of a trapped action. This enables us to perform simple forms of program analyses, called behavior analysis, that syntactically inspect the trapped processes, i.e., actions in a new thread to be executed (at local/remote sites) or the remaining actions in the current thread. In the following section, we will show how to use simple behavior analysis techniques to enforce various security policies that require checking of the future behavior of a program, the so-called predictive access control policies, and that detect and prevent execution of the potential malicious operations at the earliest stage.
5.2 Extended Syntax and Semantics

| cut ∈ Cut | cut ::= · · · | ℓ :: ca.X |
| cond ∈ BExp | cond ::= · · · | c ∈ set | ℓ ∈ set | set = ∅ |
| set ∈ Set | set ::= · · · | {c} | Act(X) | Loc_c(X) | FV(X) | FV_c(X) | LC(X) | LC_c(X) | LVar_* |

Table 3: AspectKE Syntax - Aspects for Trapping Processes

Table 3 shows the extended syntax of AspectKE.

The pointcut (Cut) in Table 2 has been extended with ℓ :: ca.X, which not only binds the action, but also binds the program continuation after the trapped action to a variable X.

Table 3 extends BExp and set in Table 2, which can be used for defining properties that require syntactic analysis of the processes to be executed (usualy the continuation of the trapped action bound by X). The set-forming behavior analysis functions Act, Loc_c, LC, LC_c, FV, FV_c will be explained in the following sections when needed, but we have collected their definitions in Table 4 where fv, bv, lc, cap and loc are the obvious extraction functions for free variables, bound variables, location constants, capabilities and target locations, respectively.

Note these functions expose different data-flow information of processes bound by process variables, and can be used to enforce predictive access control policies, namely access control policies that depend on the future behavior of a program.

**Example 9** To illustrate how aspects in the extended AspectKE trap a KLAIM program and extract its properties, the following aspect gives advice to the running example in section 2.3.

\[
A_{read}^{user :: \text{read}(\_\_\_\_\_\_)}@\text{DrSmith}.X = \begin{aligned}
\text{case} & (\text{out} \in \text{Act}(X)) \\
& \quad \text{break}; \\
& \quad \text{proceed}
\end{aligned}
\]

This aspect traps a read action of processes running at location DrSmith, when reading a tuple with five fields. The process continuation of the trapped action will be recorded in variable X. Function Act returns all actions in processes represented by X. If the actual process bound by X contains any out actions, the aspect will break the execution of the action and its continuation process. Otherwise, the action continues.

Suppose we have a system that contains the same net as in running example of Section 2.3 and aspect \( A_{read}^{user :: \text{read}(\_\_\_\_\_\_)}@\text{DrSmith}.X \):

\[
\begin{aligned}
\text{let} & \quad \text{break}; \\
& \quad \text{proceed}
\end{aligned}
\]

\( \text{in} \)
Act(\(P_1 | P_2\)) = Act(\(P_1\)) \cup Act(\(P_2\))
Act(\(\Sigma a_i P_1\)) = \bigcup\{|\cap(a_i)\} \cup Act(\(P_1\))
Act(\(* P\)) = Act(\(P\))

Loc_c(\(P_1 | P_2\)) = Loc_c(\(P_1\)) \cup Loc_c(\(P_2\))
Loc_c(\(\Sigma a_i P_1\)) = \bigcup\{|\loc(a_i)\} \cup Loc_c(\(P_1\))
Loc_c(\(* P\)) = Loc_c(\(P\))

LC(\(P_1 | P_2\)) = LC(\(P_1\)) \cup LC(\(P_2\))
LC(\(\Sigma a_i P_1\)) = \bigcup\{|\loc(a_i)\} \cup LC(\(P_1\))
LC(\(* P\)) = LC(\(P\))

FV(\(P_1 | P_2\)) = FV(\(P_1\)) \cup FV(\(P_2\))
FV(\(\Sigma a_i P_1\)) = \bigcup\{|\fv(a_i)\} \cup (FV(\(P_1\)) \setminus bv(a_i))
FV(\(* P\)) = FV(\(P\))

Table 4: Behavior Analysis Functions

\text{EHDB} :: (\text{Alice, MedicalRecord, DrHansen, Past, alicetext})
\quad \mid \quad \text{EHDB} :: (\text{Bob, PrivateNote, DrJensen, Recent, bobtext})
\quad \mid \quad \text{DrSmith} :: \text{read}(\text{Alice, MedicalRecord, DrHansen, Past, !content}@\text{EHDB}).
\quad \text{out}(\text{Alice, content}@\text{DrSmith}).
\quad \text{out}(\text{Alice, MedicalRecord, DrSmith, Recent, newtext}@\text{EHDB})

and some steps of execution (omitting the aspect definition):

\begin{align*}
\text{EHDB} :: (\text{Alice, MedicalRecord, DrHansen, Past, alicetext})
\quad \mid \quad \text{EHDB} :: (\text{Bob, PrivateNote, DrJensen, Recent, bobtext})
\quad \mid \quad \text{DrSmith} :: \text{read}(\text{Alice, MedicalRecord, DrHansen, Past, !content}@\text{EHDB}).
\quad \text{out}(\text{Alice, content}@\text{DrSmith}).
\quad \text{out}(\text{Alice, MedicalRecord, DrSmith, Recent, newtext}@\text{EHDB})
\rightarrow \quad \text{EHDB} :: (\text{Alice, MedicalRecord, DrHansen, Past, alicetext})
\quad \mid \quad \text{EHDB} :: (\text{Bob, PrivateNote, DrJensen, Recent, bobtext})
\quad \mid \quad \text{DrSmith} :: \text{0}
\end{align*}

Aspect A_2^{\text{read}} traps the read action, whose result substitution is
\[\text{[DrSmith/\text{user,}}\]
\text{out}(\text{Alice, content}@\text{DrSmith}.out(\text{Alice, MedicalRecord, DrSmith, Recent, newtext}@\text{EHDB} / X)]

Here Act(\(P\)) \subseteq \{\text{out, in, read, eval, newloc}\} is the set of capabilities performed by the process
P (see Table 4). In this case, Act returns set \{\text{out}\}, the case condition evaluates to \text{tt}, thus the aspect breaks the execution of this action and its continuation process.

6 Worked Examples: Advice for Usage Control

In this section, we now show how to use the extended AspectKE to enforce security policies that require behavior analyses of processes to be executed in the future, and we clarify the meaning of set-forming behavior analysis functions (from \text{Set} in Table 3 and 4) through examples – enforcing several predictive access control EHR security polices to the target EHR system presented in Section 2.3. In Section 6.1, we show how to enforce policies by utilizing the set-forming functions that check properties of remotely evaluating processes. In Section 6.2 we show how to enforce policies by checking properties of the continuation process at the current thread.

6.1 Remote Evaluation

Using the action \text{eval}, AspectKE can easily express a process’s remote evaluation. Moreover, using behavior analysis, AspectKE can check the content of the transmitted process by composing various aspects that embody different security considerations. This gives us a flexible way of controlling the use of remote processes. We will enforce a security policy in a distributed mobile environment that has to consider both direct and indirect access to tuple spaces, and in the latter case AspectKE shows great usefulness.

Consider a policy concerning removal of data from the system [11]:

Doctors, nurses and patients are not allowed to delete records from the database – only the administrator of the database can do that.

Thus, in terms of AspectKE, only the administrator is allowed to perform an \text{in} action on the EHR database.

Example 10 In section 4.1-4.3, we introduced aspects for restricting the \text{in} actions to access the EHR database when enforcing a basic access control policy. Here we shall slightly update them to reflect the new policy. As we prefer to use the role-based access control model, we show how to update the relevant aspects in Section 4.3.

For aspect \text{A}^{\text{read}}_{\text{plCS}}, we need to add role \text{Administrator} to role set, while updating tuple spaces \text{RDB} with tuple \langle \text{AdWalker, Administrator} \rangle, and \text{PDB} with tuples \langle \text{Administrator, MedicalRecord, in} \rangle and \langle \text{Administrator, PrivateNote, in} \rangle.

In aspect \text{A}^{\text{in}}_{\text{plCS}}, we have forbidden all users to delete records from the EHR database. Now
we relax this requirement and replace it with the following aspect.

\[ A_{\text{p5}}^{\text{in}}[\text{user :: in(\_!, recordtype, \_!, \_!)} @ \text{EHDB}] \]
\[ \triangleq \text{case(test(user, Administrator)} @ \text{RDB} \]
\[ \text{proceed; break} \]

This breaks direct attempts to perform \text{in} actions, only actions by the administrator are allowed.

This advice only deals with direct attempts to delete data; we also have to cater for processes like

\[ \text{NsOlsen :: eval(in(patient, !recordtype, !author, !createdtime, !subject) @ EHDB)} \]
\[ @ \text{AdWalker} \]

where anyone (e.g., \text{NsOlsen}) who spawns an arbitrary process on the administrator \text{AdWalker}’s node can behave as an administrator and delete the records.

This behavior can be captured by an aspect for \text{eval} actions that targets the \text{AdWalker} location, without using any behavior analysis functions.

\[ A_{\text{p5A}}^{\text{eval}}[\text{user :: eval(Y)} @ \text{AdWalker.X}] \]
\[ \triangleq \text{case(test(user, Administrator)} @ \text{RDB} \]
\[ \text{proceed; break} \]

However, this aspect is too restrictive as it disallows the possibility for other users to perform well-behaved actions on behalf of an administrator (e.g., \text{out}, \text{read} etc.).

Using behavior analysis functions, we are able to check the process in advance so that less restrictive policies can be enforced. For example, the following aspect prevents remotely spawned \text{in} actions on \text{AdWalker}, but allows other types of action.

\[ A_{\text{p5B}}^{\text{eval}}[\text{user :: eval(Y)} @ \text{AdWalker.X}] \]
\[ \triangleq \text{case(test(user, Administrator)} @ \text{RDB} \]
\[ \text{proceed; break} \]
\[ \text{case(test(user, \_)} @ \text{RDB} \land \lnot (\text{in} \in \text{Act(Y)}) \]
\[ \text{proceed; break} \]

Here \text{Act(P)} is the set of capabilities performed by the process \text{P} (see Table 4). It follows that any \text{in} actions within \text{Y} will be trapped. There is no restriction for actions other than \text{in} actions, so remote code can still perform actions like \text{out} and \text{read}.

We may want to be more liberal and allow \text{in} actions on locations distinct from \text{EHDB}. To do so we introduce \text{Loc}_{\text{in}}(\text{P}) to be the set of locations \ell, where \text{in(\_!)} @ \ell occur in \text{P}; note that this set may include location constants as well as location variables (see Table 4).
Rather than aspects $A_{p5A}^{eval}$ and $A_{p5B}^{eval}$, we could use the aspect:

$$A_{p5C}^{eval}[\text{user} :: \text{eval}(Y)@\text{AdWalker}.X]$$

\begin{align*}
\triangleq & \text{case(test(user, Administrator)@RDB)} \\
& \text{proceed;}
\end{align*}

\begin{align*}
\text{case(test(user, __)@RDB} & \wedge (LVar \cup \{EHDB\}) \cap \text{Locin}(Y) = \emptyset) \\
& \text{proceed;}
\end{align*}

\begin{align*}
\text{break}
\end{align*}

where $LVar_*$ is the set of all location variables in the program. This aspect allows users to perform an in action on behalf of an administrator when the target locations will never be EHDB, which is the least restrictive aspect for enforcing the same policy.

Example 10 shows that whilst security policies may be very simple to enforce superficially, execution of remotely spawned code might easily invalidate policies which appear to be reasonable. Using aspects, we are able to elegantly update the enforcement of the security policy to cater for this. Furthermore, the examples also illustrate AspectKE’s capability of checking the remote code before it is executed, which gives the users greater flexibility to enforce a less restrictive but more precise policy.

One might wonder whether combining the use of newloc and eval actions will break the above security policies. Consider the following example:

$$\text{NsOlsen} :: \text{newloc}(!u).\text{out}(u, \text{Administrator}@\text{RDB}).$$

$$\text{eval}(\text{in}(\text{patient, !recordtype, !author, !createdtime, !subject}@\text{EHDB}@u)@u$$

Here $\text{NsOlsen}$ tries to create a new location, register it to RDB, then execute the in action from the new location. This attempt will not work: if policy $A_{p2}^{newloc}$ and $A_{p2}^{out}$ from Example 6 are still enforced, $\text{NsOlsen}$ is neither able to create any new location nor update RDB since only a Manager can do that.

In fact, all aspects defined in this paper will directly break any action executed at a location that has not been registered in RDB, and this includes all attempts to use newloc and eval to bypass the security policies as shown above, as only the manager can update RDB through aspects defined in Example 6.

6.2 Using Program Continuations

Now we show how AspectKE can trap the continuation process and use behavior analysis functions to get the future behavior of the executing process, which enables the advice to control and avoid executing the malicious processes as early as possible.

As we have mentioned before, our society is moving in the direction of trying to exploit patient healthcare records that already exist for new purposes (known as secondary use of data). At the Canadian Institute of Health Research [28] it is stated that “health research based on the secondary use of data contributes to our present level of understanding of the causes, patterns of expression and natural history of diseases.” This raises new challenges for developing an
For example, researchers who are making secondary use of data should not be able to access the identity of patients. Therefore we want to prevent them from executing a process like

\[
\text{RsMiller :: } \text{read}(\text{patient}, \text{recordtype}, \text{author}, \text{createdtime}, \text{subject})@\text{EHDB}
\]

which explicitly specifies the patient whose records the researcher is interested in reading. We can easily compose aspects that forbid the researcher from performing such actions (but allow nurses, doctors, etc...) that are similar to those aspects that have been shown in Section 4.

In order for the researcher to blindly get a patient’s healthcare record, the researcher may perform a process such as

\[
\text{RsMiller :: } \text{read}(\text{!patient}, \text{recordtype}, \text{author}, \text{createdtime}, \text{subject})@\text{EHDB}
\]

The difference between the \textit{read} action in this program and those in the previous examples is the use of \texttt{!} in front of \texttt{patient}, i.e., the researcher does not specify which patient’s healthcare record is to be read. However, after the researcher has obtained the healthcare record, s/he can still \textit{use} the patient’s identity. We might want to prevent the researcher from executing a process like

\[
\text{RsMiller :: } \text{read}(\text{!patient}, \text{MedicalRecord}, \text{author}, \text{createdtime}, \text{subject})@\text{EHDB}, \text{out}(\text{patient, subject})@\text{Publication}
\]

whereas it would be acceptable to execute the process

\[
\text{RsMiller :: } \text{read}(\text{!patient}, \text{MedicalRecord}, \text{author}, \text{createdtime}, \text{subject})@\text{EHDB}, \text{out}(\text{subject})@\text{Publication}
\]

since the second program will not \textit{use} the identity of the patient whose record has been selected.

The following policy is extensively discussed and accepted around the world and is specified directly or indirectly in a number of codes (e.g. [17, 28, 29, 12]):

\begin{quote}
 Researchers should not read and use the patient’s healthcare records of an EHR system in a way that might potentially expose the identity of the patient.
\end{quote}

\textbf{Example 11} The following aspects, which replace the aspects \(A_{p1c1}^{\text{read}}\) and \(A_{p1c4}^{\text{read}}\) in Example 5, enforce this policy, which enforce both the basic access control policies for doctors and nurses, and policies for the researchers regarding their rights to read EHR records. It is necessary to revise the aspects because our previous development was only for primary use of data.

\[
A_{p61}^{\text{read}[\text{user :: read(patient, recordtype, _ _ _ _ _ _)}@\text{EHDB} . X]}
\triangleq \text{case}\{\exists \text{role } \in \{\text{Doctor, Nurse}\} : \\
(\text{test(user, role)}@\text{RDB} \land \text{test(role, recordtype, read)}@\text{PDB})
\}
\text{proceed;}
\text{break}
A_{p59}^{\text{read}[\text{user :: read(patient, !recordtype, _ _ _ _ _ _)}@\text{EHDB} . X]}
\triangleq \text{break}
\]
Aspect $A^{\text{read}}_{p63}$ is replaced/divided by aspects $A^{\text{read}}_{p61}$ and $A^{\text{read}}_{p64}$, according to whether a patient name is clearly specified or not. Similarly, $A^{\text{read}}_{p1c4}$ is replaced/divided by aspects $A^{\text{read}}_{p62}$ and $A^{\text{read}}_{p64}$. In both cases, additional conditions regarding researchers are only added to aspects when a patient name is not specified, i.e., $A^{\text{read}}_{p63}$ and $A^{\text{read}}_{p64}$. In these two aspects, the behavior analysis function $FV$ is used, which returns the set of free variables of $P$ (see Table 4).

In our case the aspect $A^{\text{read}}_{p63}$ will bind $X$ with the out actions of the above two programs: in program (1) with out$(\text{patient}, \text{subject})$@Publication, and in program (2) with out$(\text{subject})$@Publication. Thus $\neg(\text{patient} \in FV(X))$ would be evaluated to ff for the first case and to tt for the second case. Using this extra information from the behavior analysis function, the advice can be based on some properties of the future execution of the continuation process (in particular, whether or not patient will ever be used). And the suggestions from aspect $A^{\text{read}}_{p63}$ would be break for the first case and proceed for the second one. Note at the point that the access control decision has been made, $\text{patient}$ in the join point read action is still a defining occurrence variable and thus does not bind with any actual location yet, and behavior analysis merely analyses the future behavior of process based on its static information.

The above aspect is too restrictive since it forbids the execution of meaningful read actions as well. As one of the case studies performed by the Canadian Government [28] indicates:

In practice, the researchers might need to do some data linkage operations between different databases.

For example, we may allow the researcher to extract several records for the same patient from different databases and put that information together as in the process

$$\text{RsMiller :: read}(!\text{patient}, \text{MedicalRecord}, !\text{author}, !\text{createdtime}, !\text{subject1})@\text{EHDB}.$$  
$$\text{read}(!\text{patient}, \text{MedicalRecord}, !\text{author}, !\text{createdtime}, !\text{subject2})@\text{EHDB2}.$$  
$$\text{out}(!\text{subject1}, \text{subject2})@\text{Publication}$$

where we introduce another EHR database EHDB2 that is located at another hospital.

Now there are two databases so we need to consider which policies are suitable for each of them, respectively. For illustration purposes, we might simply demand that the second database has the same security policy as the original one. Thus the second read action would be denied, since the original policy prohibits a researcher from reading a specific patient’s
healthcare record. Because of this, establishing data linkage in the direct manner of program (3) will never work.

To make the data linkage work we can restrict the researcher’s access to the data through a trusted location (EHDB for example) by remote evaluation, and let the trusted location perform the actual data linkage actions. In this way the policy will allow the second read action whenever it is executed from the trusted location:

\[
\text{RsMiller} :: \quad \text{eval} \left( \text{read}(!\text{patient, MedicalRecord, !author, !createdtime, !subject1}) \oplus \text{EHDB.} \right.
\]
\[
\text{read}(\text{patient, MedicalRecord, !author, !createdtime, !subject2}) \oplus \text{EHDB2.}
\]
\[
\text{out}(\text{subject1, subject2}) \oplus \text{Publication}
\]
\[
\text{EHDB}
\]

As demonstrated in the last subsection, if remote evaluation is allowed, we need to pay closer attention to the overall security of the system. For example, in the event that the researcher attempts to get a doctor to link databases and output the private information as in

\[
\text{RsMiller} :: \quad \text{eval} \left( \text{read}(!\text{patient, MedicalRecord, !author, !createdtime, !subject1}) \oplus \text{EHDB.} \right.
\]
\[
\text{read}(\text{patient, MedicalRecord, !author, !createdtime, !subject2}) \oplus \text{EHDB2.}
\]
\[
\text{out}(\text{patient, subject1, subject2}) \oplus \text{Publication}
\]
\[
\text{DrSmith}
\]

or in the event that the researcher attempts to obtain the records of a patient whose name is selected from his own tuple space either before evaluating a process at an EHR database or during the evaluation procedure.

\[
\text{RsMiller} :: \quad \text{read}(!\text{patient}) \oplus \text{RsMiller.}
\]
\[
\text{eval} \left( \text{read}(\text{patient, MedicalRecord, !author, !createdtime, !subject1}) \oplus \text{EHDB.} \right.
\]
\[
\text{read}(\text{patient, MedicalRecord, !author, !createdtime, !subject2}) \oplus \text{EHDB2.}
\]
\[
\text{out}(\text{patient, subject1, subject2}) \oplus \text{Publication}
\]
\[
\text{DrSmith}
\]

\[
\text{RsMiller} :: \quad \text{eval} \left( \text{read}(!\text{patient}) \oplus \text{RsMiller.} \right.
\]
\[
\text{read}(\text{patient, MedicalRecord, !author, !createdtime, !subject1}) \oplus \text{EHDB.}
\]
\[
\text{read}(\text{patient, MedicalRecord, !author, !createdtime, !subject2}) \oplus \text{EHDB2.}
\]
\[
\text{out}(\text{patient, subject1, subject2}) \oplus \text{Publication}
\]
\[
\text{DrSmith}
\]

we need an aspect that inspects the actions performed by a researcher performing an eval action on all the EHR databases or locations other than the EHR databases.
Example 12 This motivates the aspect:

\[
A_{\text{eval}}^{\text{pt}}[\text{user} :: \text{eval}(Y)@[\text{dest}.X]] \\
= \text{case}(\text{test}(\text{user}, \text{Researcher})@[\text{RDB} \land \text{test}(\text{dest}, \text{DataBase})@[\text{RDB}]) \land \\
\forall x \in \text{LC}_{\text{read}}(Y) : \neg(\text{test}(x, \text{Patient})@[\text{RDB}]) \land \\
\forall y \in \text{Loc}_{\text{read}}(Y) : (\text{test}(y, \text{DataBase})@[\text{RDB}]) \}
\]

proceed;

\[
\text{case}(\text{test}(\text{user}, \text{Researcher})@[\text{RDB} \land \\
(\neg\text{test}(\text{dest}, \text{DataBase})@[\text{RDB} \land \text{test}(\text{dest}, \_)[@\text{RDB}]) \land \\
\forall x \in \text{Loc}_{\text{out}}(Y) : (\text{test}(x, \text{DataBase}) \lor x \in \{\text{dest}\}) )
\]

proceed;

\[
\text{case}(\neg\text{test}(\text{user}, \text{Researcher})@[\text{RDB} \land \text{test}(\text{user}, \_))
\]

proceed;

break

The first case of the aspect ensures that a researcher can directly evaluate processes in the EHR databases if s/he is not able to get the name of the patient whose record s/he is trying to obtain either before evaluation or during evaluation (by reading a patient name from a location other than EHR databases). Note that LC\text{read}(P) is a set of location constants in read actions of process P (defined in Table 7). The second case guarantees that when sending a process to other remote locations, the process only contains out actions that are performed on the EHR databases (trusted locations) or the tuple space associated with that remote location.

In summary, the examples in this section show the versatility of AspectKE when dealing with remote evaluation and future execution paths of processes. We illustrated the usefulness of each behavior analysis function when enforcing access control policies that requires the future behavior of process. When selecting the appropriate behavior analysis functions, it is possible to enforce less restrictive policies and avoid the execution of malicious processes as early as possible. More importantly, in a highly complex and privacy related computing environment, with policies that are changed frequently such as the EHR system, enforcing various access control and data usage policies through security aspects shows the potential of being a very flexible and elegant approach.

7 Related Work

7.1 Policy Enforcement Mechanisms and Aspect Oriented Programming

There are various techniques for enforcing security policies, the most traditional one is a reference monitor that observes software execution and dynamically mediates all access to objects by subjects. Instead of mixing the monitoring code in the target system, Inlined Reference Monitors (IRM) [2] use a load-time, trusted program rewriter to insert security code into a target application, resulting in a self-monitoring application that performs security checks as it executes. There are many IRM systems implemented by various program rewritters (e.g. [30, 31, 32, 33]), ensuring that different types of application will obey their corresponding security policies.
Independently, the aspect-oriented programming (AOP) paradigm [3] has emerged and has served as another effective mechanism for tackling the same issue. Indeed security is naturally identified as one kind of crosscutting concern that aspect-oriented programming was designed to deal with: instead of using a rewriter to inject monitoring code, security policies are coded through aspects which are invoked when the target program executes certain actions. There is a close connection between AOP and IRM. Hamlen and Jones [34] propose an aspect-oriented security policy specification language SPoX for enforcement by IRMs which establish a formal connection between AOP and IRMs. JavaMOP [35] implements IRM by using AspectJ aspects as the instrumentation mechanism. AspectKE takes the AOP approach to internalize the reference monitor for enforcing security policy to tuple space systems, and directly encodes security concerns in aspects.

Most research focuses on the class of security policies that can be enforced by monitoring execution of a target system [23] and hence are enforceable by traditional reference monitors. AspectKE allows us to perform a behavior analysis on future execution of the target system, giving us the capability of enforcing policies that go beyond reference monitors. In [36], the authors outline several promising methods for enforcing security policies: IRM, type systems and certifying compilers. The authors also argue that synergies among these approaches will achieve remarkable results. We believe our approach – aspects with behavior analysis – is comparable as an alternative to IRM with Type systems.

Several AOP languages can identify the data-flow and control-flow between join points, which can serve as powerful policy enforcement mechanisms. AspectJ’s cflow [33] captures the control flow between join points. Dataflow pointcut [37] identifies join points based on the dataflow of information. Tracematches [33] can give advice based on the execution history of computation. However, these systems can only refer to the past and current events, in contrast AspectKE can refer to future events. A few AOP languages propose mechanisms so that aspects can be triggered by control flow of a program in the future, e.g., pcflow [39] and transcut [40], however, they lack support for providing dataflow information in the future as AspectKE does. Some advanced AOP languages (e.g., [41, 42]) offer ways of referring to the future behavior of a program in the aspect, which in theory can be used for specifying security policies that depend on the future control-flow and data-flow. However, as they usually lack formal semantics and also normally only offer access to low level (e.g., bytecode-level) information of a program, this makes it hard to understand and develop appropriate underlying analyses for enforcing security policies. Moreover, they lack high level abstraction for presenting the analysis results as AspectKE’s behavior analysis functions do, which makes it non-trivial to use the results for composing security policies. The formal semantics of AspectKE clarifies the way of developing useful behavior (program) analyses and presenting the analysis results through appropriate language abstraction, and formally pave a way of integrating program analysis techniques into the policy specification and enforcement procedure.

7.2 Security With Aspect-Oriented Programming

There are many papers that explore the AOP mechanism to enforce security policies. One line of work directly or indirectly use the popular Java-based general purpose AOP languages
like AspectJ [3], Hyper/J [43], CaesarJ [44] to express and enforce security policies.

In [45] the authors present general guidelines for how to compose access control aspects in AspectJ, while in [46] an enforcement of application-specific policies in an access control service is implemented in CaesarJ. Phung and Sands [8] identify classes of reference monitor-style policies that can be defined and enforced by AspectJ and present a method to realize some history-dependent security policies which cannot be naturally expressed in AspectJ. Ramachandran et al. [47] discuss using AspectJ for implementing multilevel security and demonstrate how aspects, in comparison to traditional programming, can guarantee better security assurance. Similarly, AspectKE can be used to enforce a wide range of security policies but focuses on access control and explicit flow of information.

Oliveira et al. [48] use their own rewrite-based system to express access control policies and then map them into an AspectJ program; in [49], availability requirements are expressed in a formal model that combines deontic and temporal logics, and are then translated into availability aspects in AspectJ. One advantage of these approaches is that policies can be formalized through security oriented languages which are more suitable for security considerations than general purpose languages, another advantage is that some policy languages have a formal semantics which enables formal verification. AspectKE is designed with security in mind and has a formal semantics which will enable us to reason about the AspectKE’s policies in future work.

Even though these well-known AOP languages are industrial strength and can be readily used for a policy enforcement mechanism, they have their limitations: it is very hard to apply these AOP languages to the types of systems we have been studying. For example, the languages are designed for programs that run on a local machine, and they do not naturally support pointcuts for a distributed system. They also lack a pointcut mechanism to capture the future execution of a program, and thus are unable to enforce predictive access control policies. These issues are explicitly addressed by AspectKE. Particularly, we are the first to report how to use aspect oriented programming techniques to enforce secondary use of data policies that are becoming increasingly important in large IT systems.

As we have done, some researchers design their own special purpose aspect languages or systems to study security enforcement mechanisms. For example, HarmlessAML [5] is an aspect-oriented extension of Standard ML, and has a type system that guarantees well-typed harmless advice does not interfere with mainline logic computation. AspectKE is an aspect-oriented extension of a tuple space system, with a different research focus, aim at enforcing access control policies to a distributed computing model and studying how properties obtained from the behavior analyses can be used to specify access control policies. As mentioned earlier, in [37] the dataflow pointcut is proposed which specifies where aspects should be applied based on the origins of values in the past execution, which is useful in a situation where flow of information is important. AspectKE tackles similar issues, but checks the flow of information in the future.
7.3 Distribution with Aspect-Oriented Programming

Much work has been done regarding how to deploy and weave aspects for distributed systems: some work is relevant for language design of distributed AOP with explicit distribution [50, 51, 52], other work explores the implementation of AOP middleware to support distributed AOP [53, 54, 55]. AspectKE is closer to the language design of distributed AOP which naturally follows the KLAIM programming model and uses remote pointcuts [50] that identify join points in a program running on a different location. However, AspectKE does not aim at enhancing the flexibility of mechanisms to deploy, instantiate and execute distributed aspects, e.g., support advice execution over remote hosts, as AWED [51] and ReflexD [52] have achieved, rather it focuses on integrating analysis components for reasoning about the local or mobile processes to support advanced access control in a distributed setting. Compared with these languages, AspectKE provides a well defined security enforcement mechanism to tuple space systems that supports process mobility.

AO4BPEL [56] is an aspect extension of the process-oriented composition language BPEL, which was originally designed for composing Web Services. Work in [57, 58] discusses different principles of using AOP to implement coordination systems (in AspectJ), but that are not related to security. Recently, another variant of AspectK, AspectKB [59] has been proposed; it uses Belnap Logic to deal with conflicts when distributed advices are composed in a coordination environment.

7.4 Security in Coordination Languages and Security Policy Languages

AspectKE extends KLAIM, first presented in [10], later evolved into the KLAIM family (reviewed in [60]), including cKlaim, OpenKlaim, HotKlaim, OKlaim and X-Klaim etc. The prototype language of KLAIM is Klava [61], implemented in Java and has proved to be suitable for programming many distributed applications involving code mobility and our AspectKE* prototype language [13] is built on top of it. Regarding policy enforcement of these languages, some authors use control and data flow analyses that are written in the Flow Logic approach (e.g. [62, 63]), others use type systems (e.g. [64]), and [65] combine these two lines of work. They can be used to enforce very advanced security policies, however, all of them require the user to explicitly annotate policies in the main code (e.g. attach policies to each location), while our approach avoids this by specifying them inside the aspects, thus achieving a better separation of concerns.

Secure shared date-space coordination languages can be classified into two categories with regard to the underlying access control mechanisms [66]: the entity-driven approach (additional information, associated to resources such as tuple spaces, tuples and single data fields, list the entities which are allowed to access the resources) e.g., Secure Lime [67] and KLAIM [10]; and the knowledge-driven approach (resources are decorated with additional information and the processes can access the resources only in the event that they prove to keep the knowledge of such additional information) e.g., SecOS [68] and SecSpaces [69]. Our approach is suitable for expressing access control policies that fit both an entity-driven approach and a knowledge-driven approach, as the additional information is essentially expressed in aspects.
and is directly embedded in neither resources nor processes. Moreover, this additional information is not limited to the past and current facts used in the previous work, e.g., password [67], locks [68] or partitions [69], but also facts about the future, e.g., how particular data will be used, which is useful for enforcing predictive access control policies that only AspectKE can enforce.

Binder [70] and Cassandra [20] are very powerful logic-based security policy languages, which are both based on the datalog logic-programming language. In [20], there is a substantial case study performed using Cassandra that is based on the UK National Health Service procurement exercise. In AspectKE, our security policies are mainly expressed based on logical formulae and predicates in the aspects. The difference is AspectKE also provides predicates that can also describe future events. There are other prominent policy languages like Protune [71], Rei [72], Ponder [73], and KeyNote [74], which can express basic access control policies very well. Only Ponder and Rei can express usage control through obligation policies but, unlike our approach, neither language can enforce them and has to trust that the party receiving the data uses it in proper ways [75].

8 Conclusion

8.1 Our Experience of a Proof-of-Concept Implementation

We have designed and implemented a proof-of-concept programming language AspectKE* [13], based on the core concepts of AspectKE, which can be compiled and executed under a Java environment for building secure distributed systems and is freely distributed[2]. The runtime system of AspectKE* is built on top of Klava [61]. Klava, a Java package that implements the core concept of KLAIM, can be used to program tuple space operations for building distributed systems.

AspectKE* is generally implemented following the semantics of AspectKE, however, it takes a different approach regarding the implementation strategy of the behavior analyses. When aspects need the results of behavior analysis functions, AspectKE needs to perform the corresponding behavior analyses at runtime for each step of program execution, which is not practical enough due to the potential large runtime overhead. On the other hand, AspectKE* uses a two-staged implementation strategy that gathers fundamental static analysis information at the process’s load-time (the analyses are performed on the bytecode representation of the processes), and evaluates the program analysis predicates and functions (similar to the behavior analysis functions) at runtime by combining the results of the load-time analysis and runtime information, which efficiently implements the conceptual model of AspectKE.

We implemented a secure tuple space based EHR workflow system in AspectKE*, and all aspects presented in the previous sections are implemented, except for aspect $A_{eval}^p$, as the behavior analysis functions $LC$ and $LC_e$ are not currently supported by AspectKE*. Besides the EHR system, we also built a secure tuple space based chat application in AspectKE*. In

[^2]: http://www.graco.c.u-tokyo.ac.jp/ppp/projects/aspectklava.en
we show how to enforce access control policies which require analysis of future behavior of a process onto a chat system that contains untrusted components. These two applications demonstrate that the AspectKE model can be efficiently implemented and executed in real world settings.

8.2 Final Words and Future work

We have presented AspectKE, an aspect-oriented extension of the coordination language KLAIM [76]. This has provided a concrete vehicle for presenting our approach; the distributed tuple spaces provide a natural model of the kind of system that motivated our work. However, the approach could equally well be applied to more classical process algebraic languages; the join points in this case being read and write accesses to channels.

Compared with our previous work AspectK [77], AspectKE empowers aspects to trap not only the matched actions but also the process continuation and the processes being evaluated remotely. AspectKE also provides various behavior analysis functions and enables us to reason about the future execution of processes, which improves the capability of standard reference monitors that normally only deal with history based security policies. To achieve this, we simplify AspectK so that actions before and after proceed or break are not allowed. If we had allowed these actions, a safe behavior analysis would be very difficult to achieve, since the processes to be executed might execute more actions (inserted by aspects at runtime) than planned. This is an interesting direction for the future work and will require more powerful program analyses than the behavior analyses of the present paper.

We evaluated the expressiveness of AspectKE by investigating its policy enforcement capability through examples in an EHR setting. We have demonstrated that AspectKE can enforce discretionary access control, mandatory access control as well as role-based access control. Furthermore, AspectKE can elegantly retrofit new policies to an existing system with minimum effort at any phase during the system development cycle. We have also shown that in a distributed and mobile system, the information flow is very hard to control. AspectKE can enforce a range of predictive access control policies to cater for this issue, through composing aspects that check remote evaluation and the program continuation. We enforce both primary and secondary use of data, which shows that AspectKE is suitable to cater for old as well as new challenges in such a complex distributed computing environment, where security and privacy are of great importance. We also briefly report the experience of our proof-of-concept implementation of AspectKE, namely AspectKE*, and its usage for enforcing the security policies presented in this paper to an EHR workflow system and enforcing other security policies to a chat application.

The examples presented in this paper were chosen so as to illustrate the different characteristics of AspectKE, but they were also chosen so as to constitute a complete set of access control policies. Here we conjecture that our aspects indeed enforce the complete set of access control policies, but in order to validate our conjecture we need to apply a formal validation method to it.

There are other challenging secondary use of data policies which require control not only of
direct flows but also of indirect flows [78]: e.g., after storing specific data into a doctor’s own tuple space, the doctor should not allow them to be transferred into a researcher’s tuple spaces by indirectly passing through another location. In this case, checking all the parallel executing processes with existing security aspects that rely on more advanced static analyses [26] might be needed (e.g., to predict all possible data that can be stored in a certain tuple space [62, 63]).

We find that the combination of aspects with behavior and static analysis techniques shows great potential for serving as a flexible and powerful mechanism for policy enforcement, and as a promising method of building security and trust in a distributed and mobile environment.

Acknowledgement

This project was partially funded by the Danish Strategic Research Council (project 2106-06-0028) “Aspects of Security for Citizens”. We would like to thank Hidehiko Masuhara and Tomoyuki Aotani for working with us in AspectKE*, and for their comments on the paper.

References

[1] D. Gollmann, Computer Security, 1st Edition, John Wiley & Sons, 1999.
[2] Ú. Erlingsson, F. B. Schneider, IRM enforcement of Java stack inspection, in: IEEE Symposium on Security and Privacy, 2000, pp. 246–255.
[3] G. Kiczales, E. Hilsdale, J. Hugunin, M. Kersten, J. Palm, W. G. Griswold, An overview of AspectJ, in: Proceedings of the 15th European Conference on Object-Oriented Programming, ECOOP’01, Vol. 2072 of LNCS, Springer, 2001, pp. 327–353.
[4] G. Kiczales, J. Lamping, A. Mendhekar, C. Maeda, C. V. Lopes, J.-M. Loingtier, J. Irwin, Aspect-oriented programming, in: Proceedings of the European Conference on Object-Oriented Programming, ECOOP’97, Springer, 1997, pp. 220–242.
[5] D. S. Dantas, Analyzing security advice in functional aspect-oriented programming languages, Ph.D. thesis (2007).
[6] S. Gao, Y. Deng, H. Yu, X. He, K. Beznosov, K. Cooper, Applying aspect-orientation in designing security systems: A case study, in: Proceedings of the 16th International Conference on Software Engineering & Knowledge Engineering, SEKE’04, 2004, pp. 360–365.
[7] G. Georg, I. Ray, R. B. France, Using aspects to design a secure system, in: Proceedings of the 8th International Conference on Engineering of Complex Computer Systems, ICECCS’02), IEEE Computer Society, 2002, pp. 117–.
[8] P. H. Phung, D. Sands, Security policy enforcement in the OSGi framework using aspect-oriented programming, in: Proceedings of the 32nd Annual IEEE International Computer
Software and Applications Conference, COMPSAC’08, IEEE Computer Society, 2008, pp. 1076–1082.

[9] B. D. Win, W. Joosen, F. Piessens, Developing secure applications through aspect-oriented programming, in: Aspect-Oriented Software Development, Addison-Wesley, 2002, pp. 633–650.

[10] R. De Nicola, G. L. Ferrari, R. Pugliese, KLAIM: A kernel language for agents interaction and mobility, IEEE Transactions on Software Engineering 24 (5) (1998) 315–330.

[11] M. Evered, S. Bögeholz, A case study in access control requirements for a health information system, in: Proceedings of the second workshop on Australasian information security, Data Mining and Web Intelligence, and Software Internationalisation, ACSW Frontiers’04, Australian Computer Society, Inc., 2004, pp. 53–61.

[12] C. Safran, M. Bloomrosen, W. Hammond, S. Labkoff, S. Markel-Fox, P. Tang, D. Detmer, Toward a National Framework for the Secondary Use of Health Data: An American Medical Informatics Association White Paper, Journal of the American Medical Informatics Association 14 (1) (2007) 1–9.

[13] F. Yang, H. Masuhara, T. Aotani, F. Nielson, H. R. Nielson, AspectKE*:security aspects with program analysis for distributed systems, in: Proceedings of the 9th Workshop on Aspects, Components, and Patterns for Infrastructure Software, ACP4IS’10, Universitt Potsdam, 2010, pp. 27–31.

[14] D. Gelernter, Generative communication in Linda, ACM Trans. Program. Lang. Syst. 7 (1) (1985) 80–112.

[15] K. T. Win, A review of security of electronic health records, Electronic Health Records: security, safety and archiving 34 (1) (2005) 13–18.

[16] National Health Service, UK, Integrated Care Records service: Output Based Specification version 2, located at http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuid (2003).

[17] Department of Health, UK, NHS Code of Practice–Confidentiality, located at http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuid (2003).

[18] L. Qiu, Y. Zhang, F. Wang, M. Kyung, H. Mahajan, Trusted computer system evaluation criteria, National Computer Security Center.

[19] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, C. E. Youman, Role-based access control models, IEEE Computer 29 (2) (1996) 38–47.

[20] M. Y. Becker, P. Sewell, Cassandra: Flexible trust management, applied to electronic health records, in: Proceedings of the 17th IEEE Computer Security Foundations Workshop, CSFW’04, IEEE Computer Society, 2004, pp. 139–154.
[21] National Board of Health, National IT Strategy 2003-2007 for the Danish Health Care Service (2003).

[22] K. Beznosov, Requirements for access control: US healthcare domain, in: Proceedings of the third ACM workshop on Role-based access control, RBAC’98, ACM, 1998, p. 43.

[23] F. B. Schneider, Enforceable security policies, ACM Transactions on Information and System Security (TISSEC) 3 (1) (2000) 30–50.

[24] X. Leroy, Java bytecode verification: an overview, in: Computer Aided Verification, Springer, 2001, pp. 265–285.

[25] S. Oaks, Java security, O’Reilly & Associates, Inc. Sebastopol, CA, USA (1998) 456.

[26] F. Nielson, H. R. Nielson, C. Hankin, Principles of Program Analysis, 2nd Edition, Springer, Berlin, Germany, 2005.

[27] A. Myers, N. Nystrom, L. Zheng, S. Zdancewic, Jif: Java information flow, Software release. Located at http://www.cs.cornell.edu/jif, July 2003.

[28] Canadian Institutes of Health Research, Secondary Use of Personal Information in Health Research: Case Studies (2002).

[29] Danish Data Protection Agency, The Act on Processing of Personal Data, located at http://www.datatilsynet.dk/english/the-act-on-processing-of-personal-data/ (2000).

[30] Úlfar Erlingsson, F. B. Schneider, SASI enforcement of security policies: a retrospective, in: Proceedings of the workshop on New security paradigms, NSPW’99, ACM, 2000, pp. 87–95.

[31] D. Evans, A. Twyman, Flexible policy-directed code safety, in: IEEE Symposium on Security and Privacy, 1999, pp. 32–45.

[32] L. Bauer, J. Ligatti, D. Walker, Composing security policies with polymer, in: Proceedings of the 2005 ACM SIGPLAN conference on Programming language design and implementation, PLDI’05, ACM, 2005, pp. 305–314.

[33] K. W. Hamlen, Security policy enforcement by automated program-rewriting, Ph.D. thesis, adviser-Morrisett,, Greg (2006).

[34] K. W. Hamlen, M. Jones, Aspect-oriented in-lined reference monitors, in: Proceedings of the third ACM SIGPLAN workshop on Programming languages and analysis for security, PLAS’08, ACM, 2008, pp. 11–20.

[35] F. Chen, Java-MOP: A monitoring oriented programming environment for Java, in: Proceedings of the 11th International Conference on Tools and Algorithms for the Construction and Analysis of Systems, TACAS’05, Springer, 2005, pp. 546–550.

[36] F. B. Schneider, J. G. Morrisett, R. Harper, A language-based approach to security, in: Informatics - 10 Years Back. 10 Years Ahead, Vol. 2000 of LNCS, Springer, 2001, pp. 86–101.
[37] H. Masuhara, K. Kawauchi, Dataflow pointcut in aspect-oriented programming, in: Proceedings of the 1st Asian Symposium on Programming Languages and Systems, APLAS’03, Vol. 2895 of LNCS, Springer, 2003, pp. 105–121.

[38] C. Allan, P. Avgustinov, A. Christensen, L. Hendren, S. Kuzins, O. Lhoták, O. de Moor, D. Sereni, G. Sittampalam, J. Tibble, Adding trace matching with free variables to AspectJ, in: Proceedings of the 20th annual ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications, OOPSLA’05, ACM, 2005, p. 364.

[39] G. Kiczales, The fun has just begun, Keynote AOSD.

[40] H. Sadat-Mohtasham, H. Hoover, Transactional pointcuts: designation reification and advice of interrelated join points, in: Proceedings of the 8th international conference on Generative programming and component engineering, GPCE’09, ACM, 2009, pp. 35–44.

[41] T. Aotani, H. Masuhara, SCoPE: an AspectJ compiler for supporting user-defined analysis-based pointcuts, in: Proceedings of the 6th international conference on Aspect-oriented software development, AOSD’07, ACM, 2007, pp. 161–172.

[42] S. Chiba, K. Nakagawa, Josh: an open AspectJ-like language, in: Proceedings of the 3rd International Conference on Aspect-Oriented Software Development, AOSD’04, ACM, 2004, pp. 102–111.

[43] H. Ossher, P. Tarr, Hyper/J: multi-dimensional separation of concerns for Java, in: Proceedings of the 22nd international conference on Software engineering, ICSE’00, ACM, 2000, pp. 734–737.

[44] I. Aracic, V. Gasiumas, M. Mezini, K. Ostermann, An overview of CaesarJ, in: Transactions on Aspect-Oriented Software Development I, Vol. 3880 of LNCS, Springer, 2006, pp. 135–173.

[45] B. D. Win, B. Vanhaute, B. D. Decker, Security through aspect-oriented programming, in: Proceedings of the IFIP TC11 WG11.4 First Annual Working Conference on Network Security, Kluwer, 2001, pp. 125–138.

[46] T. Verhanneman, F. Piessens, B. De Win, E. Truyen, W. Joosen, Implementing a modular access control service to support application-specific policies in caesarJ, in: Proceedings of the 1st workshop on Aspect oriented middleware development, AOMD’05, ACM, 2005.

[47] R. Ramachandran, D. J. Pearce, I. Welch, AspectJ for multilevel security, in: Proceedings of the 5th Workshop on Aspects, Components, and Patterns for Infrastructure Software, ACP4IS’06, 2006, pp. 1–5.

[48] A. S. de Oliveira, E. K. Wang, C. Kirchner, H. Kirchner, Weaving rewrite-based access control policies, in: Proceedings of the 2007 ACM workshop on Formal methods in security engineering, FMSE’07, ACM, 2007, pp. 71–80.

[49] F. Cuppens, N. Cuppens-Boulahia, T. Ramard, Availability enforcement by obligations and aspects identification, in: Proceedings of the The 1st International Conference on
Availability, Reliability and Security, ARES’06, IEEE Computer Society, 2006, pp. 229–239.

[50] M. Nishizawa, S. Chiba, M. Tatsubori, Remote pointcut: a language construct for distributed AOP, in: Proceedings of the 3rd International Conference on Aspect-Oriented Software Development, AOSD’04, ACM, 2004, pp. 7–15.

[51] L. D. Benavides Navarro, M. Südholt, W. Vanderperren, B. De Fraine, D. Suvée, Explicitly distributed AOP using AWED, Research Report RR-5882, INRIA (2006). URL http://hal.inria.fr/inria-00071386/en/

[52] É. Tanter, R. Toledo, A versatile kernel for distributed AOP, in: Distributed Applications and Interoperable Systems, Springer, 2006, pp. 316–331.

[53] R. Pawlak, L. Seinturier, L. Duchien, G. Florin, F. Legond-Aubry, L. Martelli, JAC: an aspect-based distributed dynamic framework, Software Practice and Experience 34 (12) (2004) 1119–1148.

[54] B. Lagaisse, W. Joosen, True and transparent distributed composition of aspect-components, in: Proceedings of the ACM/IFIP/USENIX 2006 International Conference on Middleware, Middleware’06, Springer, 2006, pp. 42–61.

[55] E. Truyen, N. Janssens, F. Sanen, W. Joosen, Support for distributed adaptations in aspect-oriented middleware, in: Proceedings of the 7th international conference on Aspect-oriented software development, AOSD’08, ACM, 2008, pp. 120–131.

[56] A. Charfi, M. Mezini, Aspect-oriented web service composition with AO4BPEL, in: Proceedings of the European Conference on Web Services, ECOWS’04, Vol. 3250 of LNCS, Springer, 2004, pp. 168–182.

[57] S. Capizzi, R. Solmi, G. Zavattaro, From endogenous to exogenous coordination using aspect-oriented programming, in: Proceedings of the 6th International Conference on Coordination Models and Languages, COORDINATION’04, Vol. 2949 of LNCS, Springer, 2004, pp. 105–118.

[58] A. W. Colman, J. Han, Coordination systems in role-based adaptive software, in: Proceedings of the 7th International Conference on Coordination Models and Languages, COORDINATION’05, Vol. 3454 of LNCS, Springer, 2005, pp. 63–78.

[59] C. Hankin, F. Nielson, H. R. Nielson, Advice from belnap policies, in: Proceedings of the 22nd IEEE Computer Security Foundations Symposium, CSF’09, IEEE Computer Society, 2009, pp. 234–247.

[60] L. Bettini, V. Bono, R. De Nicola, G. L. Ferrari, D. Gorla, M. Loreti, E. Moggi, R. Pugliese, E. Tuosto, B. Venneri, The klaim project: Theory and practice, in: Proceedings of the Global Computing. Programming Environments, Languages, Security, and Analysis of Systems, IST/FET International Workshop, GC’03, Vol. 2874 of LNCS, Springer, 2003, pp. 88–150.

[61] L. Bettini, R. De Nicola, R. Pugliese, Klava: a Java package for distributed and mobile applications, Software-Practice and Experience 32 (14) (2002) 1365–1394.
[62] R. R. Hansen, C. W. Probst, F. Nielson, Sandboxing in myKlaim, in: Proceedings of the The First International Conference on Availability, Reliability and Security, ARES’06, IEEE Computer Society, 2006, pp. 174–181.

[63] R. R. Hansen, F. Nielson, H. R. Nielson, C. W. Probst, Static validation of licence conformance policies, in: Proceedings of the The Third International Conference on Availability, Reliability and Security, ARES’08, IEEE Computer Society, 2008, pp. 1104–1111.

[64] R. De Nicola, G. L. Ferrari, R. Pugliese, B. Venneri, Types for access control, Theoretical Computer Science 240 (1) (2000) 215–254.

[65] R. De Nicola, D. Gorla, R. R. Hansen, F. Nielson, H. R. Nielson, C. W. Probst, R. Pugliese, From flow logic to static type systems for coordination languages, in: Proceedings of the 10th International Conference on Coordination Models and Languages, COORDINATION’08, Vol. 5052 of LNCS, Springer, 2008, pp. 100–116.

[66] R. Focardi, R. Lucchi, G. Zavattaro, Secure shared data-space coordination languages: a process algebraic survey, Science of Computer Programming 63 (1) (2006) 3–15.

[67] R. Handorean, G. Roman, Secure sharing of tuple spaces in ad hoc settings, Electronic Notes in Theoretical Computer Science 85 (3) (2003) 122–141.

[68] J. Vitek, C. Bryce, M. Oriol, Coordinating processes with secure spaces, Science of Computer Programming 46 (1-2) (2003) 163–193.

[69] R. Gorrieri, R. Lucchi, G. Zavattaro, Supporting secure coordination in SecSpaces, Fundamenta Informaticae 73 (4) (2006) 479–506.

[70] J. DeTreville, Binder, a logic-based security language, in: IEEE Symposium on Security and Privacy, 2002, pp. 105–113.

[71] P. A. Bonatti, D. Olmedilla, Driving and monitoring provisional trust negotiation with metapolicies, in: Proceedings of the 6th IEEE International Workshop on Policies for Distributed Systems and Networks, POLICY’05, IEEE Computer Society, 2005, pp. 14–23.

[72] L. Kagal, T. W. Finin, A. Joshi, A policy based approach to security for the semantic web, in: Proceedings of the Second International Semantic Web Conference, ISWC’03, Vol. 2870 of LNCS, Springer, 2003, pp. 402–418.

[73] N. Damianou, N. Dulay, E. Lupu, M. Sloman, A language for specifying security and management policies for distributed systems, Tech. rep., London: Department of Computing, Imperial College (2000).

[74] M. Blaze, J. Feigenbaum, J. Ioannidis, A. Keromytis, The keynote trust-management system version 2 (1999).

[75] C. Duma, A. Herzog, N. Shahmehri, Privacy in the semantic web: What policy languages have to offer, in: Proceedings of the 8th IEEE International Workshop on Policies for Distributed Systems and Networks, POLICY’07, IEEE Computer Society, 2007, pp. 109–118.
[76] L. Bettini, R. De Nicola, R. Pugliese, G. L. Ferrari, Interactive mobile agents in X-Klaim, in: Proceedings of the 7th Workshop on Enabling Technologies Infrastructure for Collaborative Enterprises, WETICE’98, 1998, pp. 110–117.

[77] C. Hankin, F. Nielson, H. R. Nielson, F. Yang, Advice for coordination, in: Proceedings of the 10th International Conference on Coordination Models and Languages, COORDINATION’08, Vol. 5052 of LNCS, Springer, 2008, pp. 153–168.

[78] A. Sabelfeld, A. C. Myers, Language-based information-flow security, IEEE Journal on Selected Areas in Communications 21 (1) (2003) 5–19.