

Decomposition Techniques for Policy Refinement

Robert Craven	Jorge Lobo	Emil Lupu	Alessandra Russo	Morris Sloman
Dept. of Computing	IBM T.J. Watson	Dept. of Computing	Dept. of Computing	Dept. of Computing
Imperial College London	New York	Imperial College London	Imperial College London	Imperial College London
SW7 2AZ		SW7 2AZ	SW7 2AZ	SW7 2AZ

Abstract—The automation of policy refinement, whilst promising great benefits for policy-based management, has hitherto received relatively little treatment in the literature, with few concrete approaches emerging. In this paper we present initial steps towards a framework for automated distributed policy refinement for both obligation and authorization policies. We present examples drawn from military scenarios, describe details of our formalism and methods for action decomposition, and discuss directions for future research.

I. INTRODUCTION

The automated refinement of high level goals and policies into implementable management and security policies would help non-programmers specifying policies in situations such as service management, pervasive applications and for health-care services. Refinement involves generating enforceable and implementable policies guaranteed to achieve and preserve the high-level security and system management goals from which the policies are derived. Some progress has been made, but as yet there is no comprehensive, accepted solution. This paper presents our ideas on the refinement of authorization and obligation policies, with a preliminary view of a formal framework, algorithms and representations.

We view the process of policy refinement as comprising three aspects: decomposition, operationalization and distribution. In *policy decomposition*—the main focus of this paper—policies expressed at higher levels of abstraction are mapped into lower-level policies; successive mappings move closer to concepts that are directly implementable. The mapping is achieved using policy-independent *refinement rules*, defined within the scope of an application-specific system model. The paper describes the syntax of the refinement rules, how they relate to and are constrained by the system model, and the way they are applied to authorization and obligation policies.

Operationalization associates abstract policy classes with specific subjects responsible for initiating actions and specific targets on which the actions are performed, obtained from the system model. Required action parameters may also depend on specific targets. The refinement process may itself need to be *distributed*, e.g. to organizations part of a collaboration: parts of the system model may be distributed, or there may be concerns about the confidentiality of policies within organizations. In a rich, distributed policy scenario, we foresee there being several phases of policy decomposition, followed by operationalization, followed by distribution.

There is a need to interleave policy refinement with analysis

to ascertain that the refined specification achieves the requirements and is consistent with system properties and limitations, as well as with existing policies. Parallel, distributed refinement could lead to policies which conflict with each other as the refinement context will not be the same across all entities. Consequently, in the current paper we build on our previous work on policy analysis [1], [2], using the same representation language, with a view to the future integration of refinement and analysis in a single framework. The nature of the interleaving itself is left for future work.

II. SYSTEM MODELS

A system model defines the scope for specification of policies: the agents (human or automated) that can be subjects and targets, services and devices on which actions can be performed, detailed specification of the action parameters (c.f. interface specification) as well as the organizational relationships of all these entities. We use UML class diagrams to define a system model as they are widely-used, well-supported and perspicuous. Subjects and targets of policies are objects of classes specified in the class diagram, with associations being used to constrain the objects to satisfy specified properties. The main relationships that refinement exploits are generalization, aggregation, association and composition, so we currently restrict the UML to these relationships between classes.

The examples used in this paper are taken from a more detailed military policy scenario [3], involving platoons, divisions, sensor networks, and the submission of daily activity reports across the organizational structure. Figure 1 shows a fragment of the UML system model for that scenario which we will use in this paper. This example contains instances of class specialization and generalization, named associations, and aggregation (there is no composition).

Our representation language is a subset of first-order logic. We use the Event Calculus [4] (EC) to describe the state of the system, and to express the conditions under which a policy applies; we therefore show how the UML used to represent domains can be translated into the EC. The variant of the EC we use is presented in full in [2]. $holdsAt(\cdot, \cdot)$ is used to represent the changing properties of a system's state, with the first argument taking a *fluent* (a property whose value changes over time) and the second a time at which the system has the property expressed by the fluent. Thus, $holdsAt(obj(s, backupServer), 4)$ could represent that the object s exists as a backup server, at time 4.

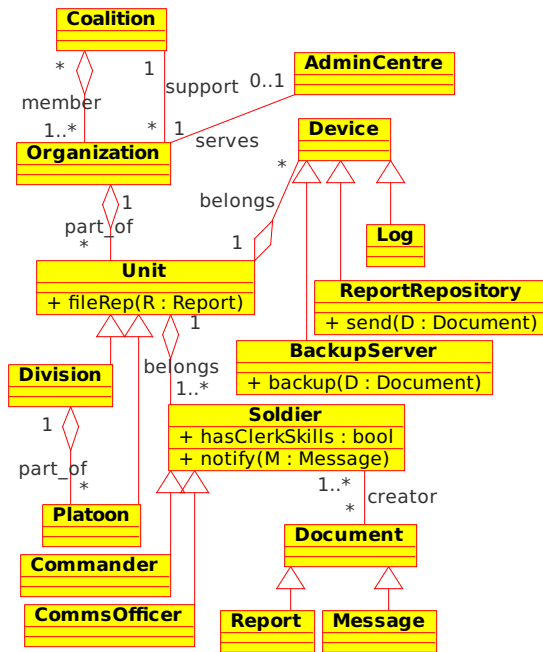


Fig. 1. UML structure for fragment of domain

There is no need to use the EC for static features of the system model, as they can be encoded directly in first-order logic, using the *class* predicate:

$$\text{class}(\text{org}) \quad \text{class}(\text{device}) \quad \text{class}(\text{adminCentre})$$

To express the relationship between classes, we use *isa* for class specialization; *assType* for association; *compType* for composition; and *aggregType* for aggregation. The relevant relationships from Figure 1 would be represented according to the following model:

$$\begin{aligned} \text{isa}(\text{division}, \text{unit}) & \quad \text{aggregType}(\text{org}, \text{part_of}, \text{cltn}) \\ \text{isa}(\text{reportRepos}, \text{device}) & \quad \text{assType}(\text{org}, \text{supports}, \text{cltn}) \end{aligned}$$

Finally, *classOp* facts are included which describe the operations possible on instances of a given class, and the types of the arguments these operations take. For instance:

$$\begin{aligned} \text{classOp}(\text{unit}, \text{fileRep}, 1, [\text{report}]) \\ \text{classOp}(\text{backupServer}, \text{backup}, 1, [\text{document}]) \end{aligned}$$

The first argument is the class name; the second, the operation name; the third, the arity; the fourth, a list of argument-types. To aid reasoning over the structure of our domain models, we introduce the following transitive version of the *isa* predicate:

$$\begin{aligned} \text{isa_trans}(X, Y) & \leftarrow \text{isa}(X, Y). \\ \text{isa_trans}(X, Z) & \leftarrow \text{isa}(X, Y), \text{isa_trans}(Y, Z). \end{aligned} \quad (1)$$

If there were a special kind of report repository called a ‘mobile report repository’, for instance, with an associated fact $\text{isa}(\text{mobileRepRep}, \text{reportRepos})$ in our knowledge base, then we could conclude that this too was a device, using the transitivity of the *isa* relationship.

Specific instances of the above classes are needed for policy operationalization. For instance, there may be both UK and US organizations in a coalition; the US may have two divisions, *eastDivision* and *westDivision*, with devices, soldiers, commanders, and so on. In our example, all divisions and platoons have a commander, a communications officer, a log, report repository, and backup server. We consider this set-up to be stable over timescales of policy specification although our approach would allow for changes. Facts about instances of classes existing at a time T are represented, in the EC, as:

$$\begin{aligned} \text{holdsAt}(\text{obj}(\text{ita}, \text{cltn}), T) \\ \text{holdsAt}(\text{obj}(\text{us}, \text{org}), T) \\ \text{holdsAt}(\text{obj}(\text{eastPlatOneComOf}, \text{commsOfficer}), T) \end{aligned}$$

We define a recursive clause for the *obj* fluent, using a new *obj_trans* fluent similar to the *isa* relationship above.

$$\begin{aligned} \text{holdsAt}(\text{obj_trans}(O, C), T) & \leftarrow \text{holdsAt}(\text{obj}(O, C), T). \\ \text{holdsAt}(\text{obj_trans}(O, C), T) & \leftarrow \\ & \text{holdsAt}(\text{obj}(O, C'), T), \text{isa_trans}(C', C). \end{aligned} \quad (2)$$

This allows us to reason, for instance, that each platoon communications officer is also a soldier—required for an authorization policy giving access to a certain document to all soldiers in the Eastern division is to be correctly applied.

Relationships between instances of classes are depicted as follows (we show a representative sample).

$$\begin{aligned} \text{holdsAt}(\text{aggreg}(\text{uk}, \text{member}, \text{ita}), T) \\ \text{holdsAt}(\text{aggreg}(\text{eastDivision}, \text{part_of}, \text{us}), T) \\ \text{holdsAt}(\text{ass}(\text{usAdminCentre}, \text{serves}, \text{us}), T) \end{aligned}$$

Definition 1 A *class definition* \mathcal{C} is a set of ground (containing no variables) instances of the predicates *class*, *isa*, *assType*, *aggregType*, *compType*, *classOp* together with the definition (1) of the *isa_trans* predicate, if that does not include cycles in the *isa* relation. We require that

- if $\text{isa}(c, c') \in \mathcal{C}$, then $\text{class}(c), \text{class}(c') \in \mathcal{C}$;
- if $\text{assType}(c, a, c') \in \mathcal{C}$, then $\text{class}(c), \text{class}(c') \in \mathcal{C}$;
- if $\text{aggregType}(c, a, c') \in \mathcal{C}$, then $\text{class}(c), \text{class}(c') \in \mathcal{C}$;
- if $\text{compType}(c, a, c') \in \mathcal{C}$, then $\text{class}(c), \text{class}(c') \in \mathcal{C}$;
- if $\text{classOp}(c, o, n, l) \in \mathcal{C}$, then $\text{class}(c) \in \mathcal{C}$, and for all $c' \in l$, $\text{class}(c') \in \mathcal{C}$, and l has length n . \lrcorner

In this way, the definition of classes, and their associations and operations is self-contained. The *isa* relationship must hold between classes explicitly stated to exist; the type of associations, aggregations and compositions must be between known classes; and operations must be on known classes.

Definition 2 An *instance definition at t* , where t is a time variable or ground member of the *Time* sort, is a set of ground instances of *holdsAt*, all of whose second arguments are t , and whose fluents are instances of *obj*, *ass*, *aggreg* or *comp*. \lrcorner

The following defines the proper relationship between class and instance definitions. (\models represents standard first-order semantic entailment.)

Definition 3 Let \mathcal{C} be a class definition, and \mathcal{I} an instance definition at time t . \mathcal{I} is *correct with respect to* \mathcal{C} if:

- If $\text{holdsAt}(\text{obj}(o, c), t) \in \mathcal{I}$ then $\text{class}(c) \in \mathcal{C}$
- If $\text{holdsAt}(\text{obj}(o, c), t), \text{holdsAt}(\text{obj}(o, c'), t) \in \mathcal{I}$ then $\mathcal{C} \models \text{isa_trans}(c, c') \vee \text{isa_trans}(c', c)$
- If $\text{holdsAt}(\text{ass}(o, a, o'), t) \in \mathcal{I}$, there are c, c' such that

$$\mathcal{C} \cup \mathcal{I} \models \text{holdsAt}(\text{obj_trans}(o, c), t) \wedge \text{holdsAt}(\text{obj_trans}(o', c'), t) \wedge \text{assType}(c, a, c')$$

- If $\text{holdsAt}(\text{aggreg}(o, a, o'), t) \in \mathcal{I}$, there are c, c' with

$$\mathcal{C} \cup \mathcal{I} \models \text{holdsAt}(\text{obj_trans}(o, c), t) \wedge \text{holdsAt}(\text{obj_trans}(o', c'), t) \wedge \text{aggregType}(c, a, c')$$

- If $\text{holdsAt}(\text{comp}(o, a, o'), t) \in \mathcal{I}$, there are c, c' such that

$$\mathcal{C} \cup \mathcal{I} \models \text{holdsAt}(\text{obj_trans}(o, c), t) \wedge \text{holdsAt}(\text{obj_trans}(o', c'), t) \wedge \text{compType}(c, a, c') \perp$$

That is, in turn: if an object is said to belong to a given class, the class must be stated to exist; if an object belongs to two different classes, one must be a subclass of the other; if an object is associated with, aggregated to form, or composed of, another object, then the two objects must belong to classes between which the association, aggregation, or composition is possible.

In Section IV we introduce *refinement rules* to define the relationship between objects and actions seen at a higher level of abstraction, to their lower-level details, contents and implementations. These rules' forms will be constrained by the UML class structure of the domain to which they apply.

III. POLICY LANGUAGE

The first stage of policy refinement is to translate the most abstract representation of a policy into a formal language upon which automated refinement techniques can work. Accordingly, we assume that policies are specified, at the highest level of abstraction, in a structured natural language, having a constrained lexicon and syntax designed for policy expression; automated translation could then compile this into the abstract, logical language we describe below. Work on structured natural languages for policies is known [5].

We present a brief summary of the language we use to represent policies. This language is abstract, in the sense that it is not meant to serve as an implementable policy language, deployed into policy decision points; rather, it is intended to serve as a generic formal language into and out of which multiple policy languages can be translated. The base language is that of constraint logic programs—normal logic programs with constraints—and we use the constraints to order the times at which the conditions of the policy must be true. Policies are expressed as rules, with the head an instance of one of the predicates *permitted*, *denied* or *obligation*, and the body of the rule representing circumstances in which the policy applies. A fuller version of the language for authorization policies is in [1]. In the following, a *constraint* c is given by:

$$c ::= s_1 = s_2 \mid s_1 < s_2 \mid s_1 \leq s_2 \\ s ::= n \mid v \mid s_1 + s_2 \mid s_1 - s_2$$

$n \in \mathbb{R}^+ \cup \{0\}$, v is a variable. A *condition* is a literal of *holdsAt*, *happens*, *permitted*, *denied*, *obligation*, *do* or *req*.

The language is divided into two parts: one for the policies and one for the system the policies are used to control. There are bridges between these parts that model the role of a Policy Enforcement Point. Authorization policies take the form

$$[\text{permitted/denied}](\text{Sub}, \text{Tar}, \text{Act}, T) \leftarrow (3) \\ L_1, \dots, L_m, C_1, \dots, C_n.$$

where the L_i are conditions in the sense defined above, and the C_i are constraints, also as above. Where the conditions L_i are true, subject to the constraints C_i , then the *Sub* is *permitted* (*denied* permission) to perform *Act* on *Tar* at time T . For example, consider the domain we sketched in Section II. A policy that [*Red Cross devices are permitted to access logs of coalition platoons*] might be formalized as

$$\text{permitted}(\text{Sub}, \text{Tar}, \text{read}, T) \leftarrow (4) \\ \text{holdsAt}(\text{obj}(\text{Sub}, \text{device}), T), \\ \text{holdsAt}(\text{ass}(\text{Sub}, \text{owner}, \text{redCross}), T), \\ \text{holdsAt}(\text{obj}(\text{Tar}, \text{log}), T), \\ \text{holdsAt}(\text{ass}(\text{Tar}, \text{owner}, C), T), \\ \text{holdsAt}(\text{obj}(C, \text{org}), T), \\ \text{holdsAt}(\text{ass}(C, \text{member}, \text{cltn}), T).$$

In the conditions, the subject *Sub* is constrained to be a device owned by the Red Cross; the target must be a *log* belonging to a nation which is a member of the coalition. There are no constraints over T , so the permission applies at all times.

Obligations require an action to be performed within a certain period; thus, in addition to the time T at which the obligation holds, reference is needed to the times T_s and T_e which are the limits of the period in which the action should be performed. Our obligation policies have the form

$$\text{obligation}(\text{Sub}, \text{Tar}, \text{Act}, T_s, T_e, T) \leftarrow (5) \\ L_1, \dots, L_m, C_1, \dots, C_n.$$

The L_i and C_i are as before—conditions and constraints. These obligations have a more general form than event-condition-action (ECA) rules, as found in many systems, such as Ponder [6]. The form considered here is more common in higher-level specifications, allowing for actions performed by humans and more complex management actions (e.g. that schedule their own actions). ECA rules are a particular form of implementing these more general obligations; our refinement process can refine the obligations into the ECAs that implement them. This happens in one of two ways: either an obligation to an action *read* (for instance) is directly refined into an ECA rule to do the *read*; or else the obligation is interpreted more along the lines found in work on norm-governed systems, and the refinement is into ECA rules notifying the user of the obligation, and taking appropriate action when the obligation is fulfilled or violated.

As an example of an obligation policy, consider [Platoon communications officers must file daily activity reports to the division their platoon belongs to between 2000h and 2200h]. This can be formalized as follows:

$$\begin{aligned} \text{obligation}(\text{Sub}, \text{Tar}, \text{fileRep}(R), 2000h, 2100h, T) \leftarrow & \quad (6) \\ & \text{holdsAt}(\text{obj}(\text{Sub}, \text{commsOfficer}), T), \\ & \text{holdsAt}(\text{ass}(\text{Sub}, \text{belongs}, P), T), \\ & \text{holdsAt}(\text{obj}(P, \text{platoon}), T), \\ & \text{holdsAt}(\text{obj}(\text{Tar}, \text{division}), T), \\ & \text{holdsAt}(\text{ass}(P, \text{part_of}, \text{Tar}), T), \\ & \text{holdsAt}(\text{obj}(R, \text{report}), T), \\ & \text{holdsAt}(\text{ass}(R, \text{reportType}, \text{dailySummary}), T), \\ & 2000h < T \leq 2200h. \end{aligned}$$

In addition to obligation policies of the form above, the syntax allows for the specification of obligations defined for classes of objects, where any member of the class (as supposed to every member of the class as above) can fulfill the obligation. These obligations are not discussed further in the current paper.

An obligation on a subject to perform an action on a target is fulfilled when the action is done. An obligation is violated when the time at which it could have been fulfilled runs out; for instance, *fulfilled* is defined as follows:

$$\begin{aligned} \text{fulfilled}(\text{Sub}, \text{Tar}, \text{Act}, T_s, T_e, T_{do}, T) \leftarrow & \quad (7) \\ & \text{do}(\text{Sub}, \text{Tar}, \text{Act}, T_{do}), \\ & \text{obligation}(\text{Sub}, \text{Tar}, \text{Act}, T_s, T_e, T_{do}), \\ & T_s \leq T_{do} < T_e, T_{do} \leq T. \end{aligned}$$

$\text{obligation}(\text{Sub}, \text{Tar}, \text{Act}, T_s, T_e, T)$ must be defined such that the constraints in its body imply $T \leq T_e$; however, the obligation may only be fulfilled at a time T_{do} before T_e .

IV. REFINEMENT RULES

In defining the management and security policies of a system, it ought only to be necessary to specify the policies at the highest level of abstraction, and then use refinement techniques to map those higher-level policies, in successive phases, to the lowest level. Policies that refer generically to sensors as targets of configuration actions, for instance, may first be mapped to types of sensor (audio, or video, say). The policies derived for audio sensors may then be mapped onto policies for different models of audio sensors, with model-specific configuration operations. Finally, policies controlling configuration operations intended for a given model of audio sensor may be mapped into policies concerning messages sent to a certain port of all audio sensors of the relevant model known to be currently attached to a sensor network.

To achieve this gradual refinement, we use rules to relate subjects, targets and actions described abstractly, or described *en masse*, to the lower level. The basic components of such rules are known as *conditioned actions*—expressions which represent the performance of an action on a target by a subject in certain contexts, and where the contexts place constraints on the types of the subjects, targets and actions and the relationships between them.

Definition 4 Let \mathcal{L} be a policy representation language.¹ A *conditioned action* for \mathcal{L} has the form

$$(\text{Sub}, \text{Tar}, \text{Act}) : C_1, \dots, C_n.$$

where each C_i has one of the forms $\text{obj}(O, C)$, $\text{ass}(O, A, O')$, $\text{aggreg}(O, A, O')$ or $\text{comp}(O, A, O')$. In each case, C must be a ground class name from the domain, and A must be a ground association, aggregation, or composition name from the domain. Any variable occurring in $\text{Sub}, \text{Tar}, \text{Act}$ must also occur in one of the C_i ; this constrains the refinement rule to apply only to objects of classes present in the UML diagram.□

For example, a conditioned action

$$\begin{aligned} (\text{Sub}, \text{Tar}, \text{read}(\text{locDev}, \text{wQuad})) : & \quad (8) \\ & \text{obj}(\text{Sub}, \text{device}), \text{ass}(\text{Sub}, \text{owner}, X), \\ & \text{obj}(\text{Tar}, \text{locationServer}), \text{ass}(\text{Tar}, \text{owner}, X). \end{aligned}$$

represents the performance of an action of reading information on the location of devices in the West Quadrant by devices which belong to the same organization as the location server from which the information is read. Such conditioned actions form the basic components of refinement rules.

Definition 5 Let \mathcal{L} be a policy representation language, and \mathcal{C} a class definition. A *refinement rule* is an expression

$$\mathcal{C} \Rightarrow C_1 \text{ then } \dots \text{ then } C_n \quad (9)$$

where C, C_1, \dots, C_n are conditioned actions, and $i \geq 1$. C is called the *top* of such a rule, and the C_i are *bottoms*.

More precisely, where R has the form

$$\begin{aligned} (\text{Sub}, \text{Tar}, \text{Act}) : \text{Conds} \Rightarrow & \quad (10) \\ (\text{Sub}_1, \text{Tar}_1, \text{Act}_1) : \text{Conds}_1 & \\ \text{then } \dots & \\ \text{then } (\text{Sub}_n, \text{Tar}_n, \text{Act}_n) : \text{Conds}_n & \end{aligned}$$

- 1) There must be $\text{obj}(\text{Sub}, C)$ and $\text{obj}(\text{Tar}, C')$ in the conditions of \mathcal{C} , such that for some N and L

$$\begin{aligned} \mathcal{C} \models & \text{class}(C) \wedge \text{class}(C') \wedge \text{classOp}(C'' \text{Act}, N, L) \\ & \wedge \text{isa_trans}(C', C'') \end{aligned}$$

and further, in each Conds_i , $1 \leq i \leq n$:

2. there is an expression $\text{obj}(\text{Sub}_i, C_i)$ (resp. $\text{obj}(\text{Tar}_i, C_i)$) in Conds_i , but not Conds , with $\mathcal{C} \models \text{isa_trans}(C_i, C)$ and $\text{obj}(\text{Sub}, C)$ (resp. $\text{obj}(\text{Tar}, C)$) is in Conds —or $\text{Sub}_i = \text{Sub}$ (resp. $\text{Tar}_i = \text{Tar}$);
3. there is $\text{aggreg}(\text{Sub}_i, A, \text{Sub})$ (resp. $\text{aggreg}(\text{Tar}_i, A, \text{Tar})$) in Conds_i which is not in Conds , and such that $\mathcal{C} \models \text{aggregType}(C_i, A, C)$, $\text{Conds} \models \text{obj}(\text{Sub}, C)$ and $\text{Conds}_i \models \text{obj}(\text{Sub}_i, C_i)$ —or else $\text{Sub}_i = \text{Sub}$ (resp. $\text{Tar}_i = \text{Tar}$);
4. there is $\text{comp}(\text{Sub}_i, A, \text{Sub})$ (resp. $\text{comp}(\text{Tar}_i, A, \text{Tar})$) in Conds_i but not in Conds , with $\mathcal{C} \models \text{compType}(C_i, A, C)$,

¹See [2]. \mathcal{L} defines the language used to specify both policies and the domains they operate on.

$Conds \models obj(Sub, C)$ and $Conds_i \models obj(Sub_i, C_i)$ —or $Sub_i = Sub$ (resp. $Tar_i = Tar$).

In addition, for each i such that $1 \leq i \leq n$

5. $\mathcal{C} \cup Conds \cup Conds_i \models classOp(Tar_i, Act_i, N_i, L_i)$ for some N_i, L_i . \square

The meaning of such rules is that the succession of actions referred to in C_1, \dots, C_n correspond to a refinement of the higher-level action in C , with **then** intended to denote temporal sequence. The constraints 1–5 ensure that our refinement rules respect the UML system model.

Specifically, constraint 1 ensures that the subjects and targets of the high-level conditioned action appearing in the refinement rule must be objects of the domain, as defined by the UML; this constraint also forces the high-level action to be one of the operations the target supports. Constraints 2–4 ensure that, as we refine, we move to more specific subjects and targets; for each C_i that constitutes a lower-level action, either the subject or the target must be more specific, and neither should be *more* general or abstract. For instance, if the subject of the top of a refinement rule is defined to be a platoon, then the subject of the action into which we refine could be a soldier belonging to that platoon (this would be a case of aggregation). Or, the target might be constrained, in the top of the refinement rule, to be a sensor—a possible target of the bottom of the refinement rule could be an audio sensor, the sensor class being a generalization of audio sensors. Finally, constraint 5 ensures that the refinement rules are well-formed—the actions Act_i of the lower-level parts of the rules must be operations it is possible to perform on the targets Tar_i .

As an illustration, consider an action of filing a daily activity report within the US military. We suppose the workflow for filing reports is predefined centrally within the US army, and the procedure must be followed throughout the organization. Seen at a more abstract perspective, soldiers file reports to units (not necessarily their own). However, for a soldier to file a report, a series of actions specified at a more concrete level must occur. He must first send the report to the report repository of the unit to which the report is to be filed. He must then backup the report to that unit’s backup server. When this has been done, the soldier must notify the communications officer of the unit that the upload has been completed. The following refinement rule represents this picture:

$$\begin{aligned}
 & (Sub, Tar, fileRep(R)) : & (11) \\
 & \quad obj(Sub, soldier), obj(Tar, unit) \\
 & \quad \Downarrow \\
 & (Sub, Tar_1, send(R)) : \\
 & \quad obj(Tar_1, reportRepos), aggreg(Tar_1, belongs, Tar) \\
 & \text{then } (Sub, Tar_2, backup(R)) : \\
 & \quad obj(Tar_2, backupServer), ass(Tar_2, belongs, Tar) \\
 & \text{then } (Sub, Tar_3, notify(upload(R))) : \\
 & \quad obj(Tar_3, commsOfficer), ass(Tar_3, belongs, Tar)
 \end{aligned}$$

Note that the same variable Sub occurs as subject of all actions, both in the top and bottom of the rule: the same soldier must

send the report, then back it up and make the notification. However, the target varies: from a report repository, to a backup server, to a communications officer.

In [7] we used concepts from data integration to formulate rules and algorithms for the refinement of targets and actions, for the same broader goal of policy refinement. Data integration concerns the way in which heterogeneous databases are mapped to each other so their vocabularies (schemata) are inter-translatable. Refinement rules as we define them here are a generalization of that previous work.

V. POLICY DECOMPOSITION

In this section, we examine in detail the decomposition stage of policy refinement. This itself divides into two parts. *Matching* verifies whether a refinement rule is applicable to a policy; and *decomposition* proper performs the refinement.

A. Matching

Recall policy (6). In order to refine this using rule (11), variables are first renamed so as no longer to be shared between policy and refinement rule. Then, an attempt is made to unify the head of the top of the refinement rule—in our example, $(Sub, Tar, fileRep(R))$ —with the corresponding features of the head of the policy. Let variables in the policy (6) which are shared with the refinement rule be given a mark $'$ to separate them; the subject, target and action can be matched using the most general unifier (m.g.u.) $\theta = \{Sub/Sub', Tar/Tar', R/R'\}$.

Next, a test is made to determine whether the constraints on the action in the body of the top of the refinement rule—in our example, $obj(Sub', soldier)$ and $obj(Tar', unit)$ with θ applied—are consistent with the conditions on the subject, target, and action found in the policy. The purpose of this consistency check is to ensure that there is an overlap between the circumstances in which the refinement rule and policy apply. The consistency check is made as follows. Let $Conds$ be the set of conditions in the top of the refinement rule, and $Conds_P$ be the set of *holdsAt* conditions in the body of the policy, with θ applied. Let $Conds_P^*\theta$ be the result of stripping the *holdsAt*(\cdot, \cdot) predicates from the members of $Conds_P\theta$, leaving the fluents. In our case, we have

$$\begin{aligned}
 Conds\theta &= \{obj(Sub', soldier), obj(Tar', unit)\} \\
 Conds_P^*\theta &= \{obj(Sub', commsOfficer), aggreg(Sub', belongs, U_s), \\
 & \quad obj(U_s, platoon), obj(Tar', division), \\
 & \quad aggreg(U_s, part_of, Tar'), obj(R', report), \\
 & \quad ass(R', reportType, dailySummary)\}
 \end{aligned}$$

We now check $Conds\theta$ and $Conds_P^*\theta$ for logical consistency, with respect to the background UML formalization. (The details are given below, in Definition 6.) If this phase succeeds, then the refinement rule is known to be applicable to the policy.

Definition 6 Let \mathcal{L} be a policy representation language, \mathcal{C} a class description. Given an obligation policy P

$$obligation(Sub', Tar', Act', T_s, T_e, T) \leftarrow Conds_P, Cons_P.$$

or an authorization policy

$[permitted/denied](Sub', Tar', Act', T) \leftarrow Conds_P, Cons_P.$

where $Conds_P$ are literals and $Cons_P$ are constraints, and a refinement rule R of the form (10), whose variables have been renamed to be different, we say that P and R *head-match with* θ if there is an m.g.u. θ such that $(Sub, Tar, Act)\theta = (Sub', \theta', Act')\theta$. Further, if P and R head-match on θ , let $Conds_P^*$ be defined as

$$\{F \mid \exists T(\text{holdsAt}(F, T) \in Conds_P, F \text{ is } obj, \text{ ass, aggreg or comp})\} \\ \cup \{\neg F \mid \exists T(\mathbf{not\ holdsAt}(F, T) \in Conds_P, F \text{ is an } obj, \text{ ass, aggreg or comp fluent})\}$$

Let \mathcal{X} be $\mathcal{C} \cup Conds_P^* \cup Conds$. R is said to *match* P on θ if

$$\forall o[obj(o, c), obj(o, c') \in \mathcal{X} \rightarrow \\ \models \mathcal{X} \rightarrow (isa_trans(c, c') \leftrightarrow \neg isa_trans(c', c))] \quad \square$$

If a refinement rule matches a policy on some θ according to Definition 6, then the rule can be applied to the policy during policy refinement.

B. Policy Decomposition

After a successful matching process, the policy is decomposed. Part of the conditions defining the scope of the refined policies comes from the original policy, part from the top of the refinement rule, and part from the conditioned actions in the bottom of the refinement rule. The circumstances in which the refined policy applies will be the intersection of these three.

Given conditioned actions C_1, \dots, C_n in the bottom of the refinement rule, n copies of $P\theta$ are made, with the subject, target and action of $P\theta$ replaced by those of each $C_i\theta$. Conditions from the top of the refinement rule that are not implied by the conditions of the original policy are added to each of these copies. Conditions from each C_i are added to the policy. Conditions are added deriving from each conditioned action C_i that the actions at the head of each $C_j, j < i$, have been performed, and in the right order.

Returning to our running example, let *PolicyConditions* be:

$$\text{holdsAt}(obj(Sub', \text{commsOfficer}), T), \\ \text{holdsAt}(aggreg(Sub', \text{belongs}, U_s), T), \\ \text{holdsAt}(obj(U_s, \text{platoon}), T), \\ \text{holdsAt}(obj(Tar', \text{division}), T), \\ \text{holdsAt}(ass(U_s, \text{part_of}, Tar'), T), \\ \text{holdsAt}(obj(R', \text{report}), T), \\ \text{holdsAt}(ass(R', \text{reportType}, \text{dailySummary}), T),$$

These are the literals that will be inherited from the higher-level policy. The results for our example are shown below (the source of the various conditions in the bodies of the policies has been marked to their left). The policy stemming from the first conditioned action, C_1 , is:

$$\text{obligation}(Sub', Tar_1, \text{send}(R'), 2000h, 2200h, T) \leftarrow \quad (12) \\ C_1 \left[\begin{array}{l} \text{holdsAt}(obj(Tar_1, \text{reportRepos}), T), \\ \text{holdsAt}(aggreg(Tar_1, \text{belongs}, Tar'), T), \end{array} \right. \\ \text{policy} \left[\begin{array}{l} \text{PolicyConditions}, \\ 2000h < T \leq 2200h. \end{array} \right.$$

The policy coming from the second conditioned action, C_2 :

$$\text{obligation}(Sub', Tar_2, \text{backup}(R'), T_1, 2200h, T) \leftarrow \quad (13) \\ C_2 \left[\begin{array}{l} \text{holdsAt}(obj(Tar_2, \text{backupServer}), T), \\ \text{holdsAt}(aggreg(Tar_2, \text{belongs}, Tar'), T), \end{array} \right. \\ \text{policy} \left[\begin{array}{l} \text{PolicyConditions}, \\ T_1 < T < 2200h, \end{array} \right. \\ C_1 \left[\text{fulfilled}(Sub', Tar_1, \text{send}(R'), 2000h, 2200h, T_1, T). \right.$$

Finally, the policy derived from C_3 is:

$$\text{obligation}(Sub', Tar_3, \text{notify}(\text{upload}(R')), T_2, 2200h, T) \leftarrow \quad (14) \\ \text{from } C_3 \left[\begin{array}{l} \text{holdsAt}(obj(Tar_3, \text{commsOfficer}), T), \\ \text{holdsAt}(aggreg(Tar_3, \text{belongs}, Tar'), T), \end{array} \right. \\ \text{policy} \left[\begin{array}{l} \text{PolicyConditions}, \\ 2000h < T_1 < 2200h, \\ T_1 < T_2 < 2200h, \end{array} \right. \\ C_1 \left[\text{fulfilled}(Sub', Tar_1, \text{send}(R'), 2000h, 2200h, T_1, T), \right. \\ C_2 \left[\text{fulfilled}(Sub', Tar_2, \text{backup}(R'), T_1, 2200h, T_2, T). \right.$$

Some description of what has happened here is appropriate. We focus on the third derived policy. The action is determined by the conditioned action

$$(Sub, Tar_3, \text{notify}(\text{upload}(R))) : \\ obj(Tar_3, \text{commsOfficer}) \\ ass(Tar_3, \text{belongs}, Tar)$$

from the rule (11). The first set of conditions in the derived policy's body is inherited from the specific conditioned action in the bottom of the action-refinement rule: these define the nature of the target of the derived policy (it is this target, and the action in the head of the policy, that are altered from the high-level policy). Next, there are conditions come from the high-level policy: these define the subject which must perform the lower level *notify*(\cdot) action, the kind of report referred to, and the relationship of the subject to platoons and divisions.

There are two additional sets of conditions in (14). These require that, for the obligation policy to hold, the obligations derived from other conditioned actions in the action decomposition rule must have been fulfilled. This is a result of the structure of the refinement rule: one should only notify of an upload, once the report has been sent and then the report has been backed up. In general, given a refinement rule $C \Rightarrow C_1 \mathbf{then} \dots \mathbf{then} C_n$, the ordering of actions is crucial. An obligation to do the action of C is an obligation to do the actions of the C_1, \dots, C_n in order, and so C_i must be performed only if C_j , for $j < i$, have been correctly executed. To return to our example policy, the times between which the action that the policy requires must be performed, are T_2 and $2200h$. $2200h$ is the final time from the original high-level policy; T_2 is the time at which the obligation derived from the previous conditioned action C_2 was fulfilled.

The refinement of positive authorization policies (with *permitted* in their heads) is very similar. The only difference concerns the temporal constraints on when the derived policies hold. In an obligation policy, an interval is given, within which the action must be performed; when the policy is refined, we

have interpreted this to mean that the entire sequence of lower-level actions must be performed in the interval, in the correct order. But with authorization policies, in general, a reference is made in the head of the policy to a single time at which the action is authorized. We therefore adopt the convention that it is the *final* action in the sequence that is authorized at this time. The high-level policy can be seen as authorizing the production of certain effects (those which the high-level action achieves), and these effects are only guaranteed to be achieved when the entire sequence of actions as been performed.

For negative authorization policies (with *denied* in their head), a slightly different tactic is taken. The denial of a high-level action is the denial of permission to execute a sequence, but one might interpret this to mean, either the denial of permission to perform *every* action in the sequence, or denial to perform only *some* of these actions, selected according to some appropriate criterion. We take the latter approach, and have chosen in the current work to disallow the *final* action of the sequence, as the one that achieves the sequence's completion. However, there is no reason why this choice of the way in which to refine negative authorization policies should not be made available to the user. Note that the final action of a sequence is denied, in this model, only if the previous actions of the sequence have already been performed; this qualification is achieved by conditions in the body of the derived policy.

Definition 7 Let P be either an obligation or authorization policy, of any of the forms

$$\begin{aligned} & \text{obligation}(Sub', Tar', Act', T_s, T_e, T) \leftarrow Conds_P, Cons_P, \\ & [\text{permitted/denied}](Sub', Tar', Act', T) \leftarrow Conds_P, Cons_P. \end{aligned}$$

Let R be a refinement rule $C \Rightarrow C_1 \text{ then } \dots \text{ then } C_n$. C has the form $(Sub, Tar, Act) : Conds$, and each C_i has the form $(Sub_i, Tar_i, Act_i) : Conds_i$. Assume that R matches P with m.g.u. θ . Recall the meaning of $Conds_P^*$ from Definition 6. Matching ensures that $Conds$ and $Conds_P^*$ are consistent with respect to the UML class definitions. Let $holdsAt(Conds)$ be defined as the set $\{holdsAt(F, T) \mid F \in Conds\}$ and let $holdsAt^*(Conds)$ is defined to be

$$\{holdsAt(F, T) \mid F \in Conds \text{ and } Conds_P \cup C \not\models holdsAt(F, T)\}$$

(i) if P is an obligation policy, then the refinement of P w.r.t. R is the set of policies

$$\begin{aligned} [& \text{obligation}(Sub_i, Tar_i, Act_i, T_i, T_e, T) \leftarrow \\ & \quad Conds_P, holdsAt^*(Conds), \\ & \quad Fulf_1, \dots, Fulf_{i-1}.] \theta \end{aligned} \quad (15)$$

Where $Fulf_j$, for $1 \leq j < i$ is

$$\text{fulfilled}(Sub_j, Tar_j, Act_j, T_{j-1}, T_j, T_e, T). \quad (16)$$

In the case where $j = 1$, the condition T_{j-1} is replaced by T_s . The $Fulf_j$ expressions ensure that the previous actions in the series of those required have been correctly performed.

(ii) if P is a positive authorization policy, then the refinement of P w.r.t. R is the set of policies

$$\begin{aligned} [& \text{permitted}(Sub_i, Tar_i, Act_i, T) \leftarrow \\ & \quad Conds_P, holdsAt^*(Conds), \\ & \quad Done_1, \dots, Done_{i-1}.] \theta \end{aligned} \quad (17)$$

$Done_j$, for $1 \leq j < i$ is

$$\begin{aligned} & \text{do}(Sub_j, Tar_j, Act_j, T_j), \\ & \quad T_{j-1} < T_j < T_j, \\ & \quad holdsAt^*(C_j). \end{aligned}$$

Where $j = 1$, there is no temporal constraint; where $j = i - 1$, the constraint is $T_{i-1} < T$.

(iii) finally, if P is a negative authorization policy, then the refinement of P with respect to R is

$$\begin{aligned} [& \text{denied}(Sub_n, Tar_n, Act_n, T_n) \leftarrow \\ & \quad Conds_P, \\ & \quad holdsAt^*(Conds), \\ & \quad Done_1, \dots, Done_{n-1}.] \theta \end{aligned} \quad (18)$$

The definition of *Done* is the same as for *permitted*. \lrcorner

This definition formalizes the preceding discussion.

C. Multiple refinement rules

In a given scenario, be true that many different refinement rules match a given policy. This could be the case when there are a number of heterogenous devices, which implement an action specified at a higher level of abstraction in different ways. For example, if a generic *configure* action applies to devices designated as *sensors*, this might be refined to an *audio sensor's* having its sampling rate set, and to an *image sensor's* having its resolution set. There would be two different refinement rules, for each type of sensor, but both may match the same policy, and a single policy should be refined using both simultaneously, so the policy may be implemented on the PDPs applying both to audio sensors and image sensors. This ability of a single higher-level policy to be refined in multiple different ways is one of the advantages of our approach.

VI. OPERATIONALIZATION AND DISTRIBUTION

Operationalization in software engineering is the process of assigning specific resources for the performance of goals. In the context of policy refinement, we understand it as the selection of named entities for the execution of policies.

In a centralized refinement model, we would expect the following arrangements. There would be a set of policy refinement rules of the form given by (5), together with a class definition \mathcal{C} , and an *instance repository* which stores information about the instances of classes known to exist, and their relationships (instance repositories can be represented as instance definitions, as in Section II). A policy to be refined enters the refinement component, and the refinement rules are iteratively applied to it—more than one refinement rule may, of course, apply to a given policy. When policies have reached the point that no more refinement rules are applicable to them, then their form is tested against a subset of the language which

is known to describe concretely implementable components of the domain: class names of actual rather than abstract devices, and the lowest level of concrete description of the properties of those devices. If the policies are expressed in this “implementable” subset of the policy representation language, then they may be operationalized.

Authorization policies grant or deny the permission to perform an action on a target; their operationalization accordingly means grounding the targets referred to in the heads of rules. If the target is not currently ground, then those literals in the body of the authorization policy which include the variable appearing in the head of the rule as target are collected, together with literals transitively sharing variables with them. These are then passed, as a query, to the database holding the current instance repository. Answers return supply a set of ground values for the target of the policy; it is to those targets that the policy is distributed, with the ground target replacing the target variable in the rule.

In the case of obligation policies, the last stage of refinement transforms the policy into ECA-type policies which (i) will notify subjects of the obligation of its imminence, and (ii) take any necessary action if the obligation is fulfilled or violated. Operationalization here means the selection of the appropriate policy monitors to carry out these notifications and responses.

In a fully distributed refinement scenario, partially-refined policies may be distributed to other refinement-centres, with local knowledge of devices and their refinement rules. Phases of decomposition, operationalization and distribution are then made locally.

VII. RELATED WORK

Policy refinement remains one of the most ambitious goals in policy-based security management because it aims to automate the realization of high-level requirements in executable implementations. Early studies (e.g. [8]) highlighted some of the specific issues and have attempted to address them in a very restricted way. However, there has been renewed interest in this problem in recent years as demonstrated by the panel discussion at IM 2007 and several recent studies [9], [10], [11], [12] which address subsets of the problem such as goal decomposition for refinement or transforming specifications using predefined mappings. [13] considers a planning approach to refining change requests into implementable tasks, but does not consider policies. In [14] a goal is essentially a utility function which can be optimised to select a set of parameters for policies used to configure a sensor network. Our approach is much more general than these rather specialised forms of policy refinement.

VIII. CONCLUSION

In this paper, we have described the process of action decomposition in a policy refinement framework we are developing. Future work must address the formalization of operationalization and distribution, as well as the full integration of distributed policy refinement with policy analysis.

We are testing the approach for specifying policies relating to inter-organizational collaborations and the use of sensor networks. We intend to develop tools to help non-technical users refine their goals into policies.

Acknowledgment Research was sponsored by the U.S. Army Research Laboratory and the U.K. Ministry of Defence under Agreement W911NF-06-3-0001. The views and conclusions in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Army Research Laboratory, the U.S. Government, the U.K. Ministry of Defence or the U.K. Government.

REFERENCES

- [1] R. Craven, J. Lobo, J. Ma, A. Russo, E. Lupu, A. Bandara, S. Calo, and M. Sloman, “Expressive policy analysis with enhanced system dynamicity,” in *ASIACCS*. ACM, 2009, pp. 239–250.
- [2] R. Craven, E. Lupu, J. Lobo, A. Bandara, S. Calo, J. Ma, A. Russo, and M. Sloman, “An expressive policy analysis framework with enhanced system dynamicity,” Technical Report, Department of Computing, Imperial College London, 2008.
- [3] R. Craven, J. Lobo, E. Lupu, J. Ma, A. Russo, and M. Sloman, “Distributed policy scenario,” ITA Technical Report, 2010.
- [4] R. A. Kowalski and M. J. Sergot, “A logic-based calculus of events,” *New Generation Comput.*, vol. 4, no. 1, pp. 67–95, 1986.
- [5] C. Brodie, C.-M. Karat, and J. Karat, “An empirical study of natural language parsing of privacy policy rules using the sparkle policy workbench,” in *SOUPS*, ser. ACM International Conference Proceeding Series, L. F. Cranor, Ed., vol. 149. ACM, 2006, pp. 8–19.
- [6] N. Damianou, N. Dulay, E. Lupu, and M. Sloman, “The ponder policy specification language,” in *POLICY*, ser. LNCS, M. Sloman, J. Lobo, and E. Lupu, Eds., vol. 1995. Springer, 2001, pp. 18–38.
- [7] R. Craven, J. Lobo, E. Lupu, A. Russo, and M. Sloman, “Security policy refinement using data integration: a position paper,” in *SafeConfig '09: Proceedings of the 2nd ACM workshop on Assurable and usable security configuration*. New York, NY, USA: ACM, 2009, pp. 25–28.
- [8] M. Abadi and L. Lamport, “The existence of refinement mappings,” *Theor. Comput. Sci.*, vol. 82, no. 2, pp. 253–284, 1991.
- [9] A. K. Bandara, E. Lupu, J. D. Moffett, and A. Russo, “A goal-based approach to policy refinement,” in *POLICY*. IEEE Computer Society, 2004, pp. 229–239.
- [10] S. Davy, B. Jennings, and J. Strassner, “Conflict prevention via model-driven policy refinement,” in *DSOM*, ser. Lecture Notes in Computer Science, R. State, S. van der Meer, D. O’Sullivan, and T. Pfeifer, Eds., vol. 4269. Springer, 2006, pp. 209–220.
- [11] J. Rubio-Loyola, J. Serrat, M. Charalambides, P. Flegkas, and G. Pavlou, “A functional solution for goal-oriented policy refinement,” in *POLICY*. IEEE Computer Society, 2006, pp. 133–144.
- [12] M. Beigi, S. B. Calo, and D. C. Verma, “Policy transformation techniques in policy-based systems management,” in *POLICY*. IEEE Computer Society, 2004, pp. 13–22.
- [13] D. Trastour, R. Fink, and F. Liu, “ChangeRefinery: Assisted Refinement of High-Level IT Change Requests,” *Policies for Distributed Systems and Networks, IEEE International Symposium on*, vol. 0, pp. 68–75, 2009.
- [14] G. Campbell and K. Turner, “Goals and policies for sensor network management,” *Sensor Technologies and Applications, International Conference on*, vol. 0, pp. 354–359, 2008.