Mining Parametric Specifications

Choonghwan Lee  Feng Chen  Grigore Roșu  
Department of Computer Science  University of Illinois at Urbana-Champaign  
{clee83, _, grosu}@illinois.edu

ABSTRACT
Specifications carrying formal parameters that are bound to concrete data at runtime can effectively and elegantly capture multi-object behaviors or protocols. Unfortunately, parametric specifications are not easy to formulate by non-experts and, consequently, are rarely available. This paper presents a general approach for mining parametric specifications from program executions, based on a strict separation of concerns: (1) a trace slicer first extracts sets of independent interactions from parametric execution traces; and (2) the resulting non-parametric trace slices are then passed to any conventional non-parametric property learner. The presented technique has been implemented in jMiner, which has been used to automatically mine many meaningful and non-trivial parametric properties of OpenJDK 6.

Categories and Subject Descriptors
F.3.1 [Logics and Meanings of Programs]: Specifying and Verifying and Reasoning about Programs

General Terms
Verification, Algorithms

Keywords
Parametric specifications, dynamic analysis

1. INTRODUCTION

Formal specifications define behaviors that systems or parts of systems should obey. A parametric specification is a formal specification that carries parameters that are bound to concrete object instances at runtime. As an example, Figure 1 shows a finite state automaton (FSA) that describes a parametric specification involving two Java classes, Collection and Iterator. Each edge represents an event, such as calling a method. (init) represents calling the constructor of Collection and update calling a method that changes the contents of the Collection object specified as parameter c; add, remove and clear can be such methods. iterator represents creating an Iterator object for a Collection object and has two parameters: the underlying Collection object c and the created Iterator object i. hasNext and next represent invocations on methods hasNext and next of Iterator, respectively. The next method returns the next element in the iteration and the hasNext method checks if the iteration has more elements (i.e., hasNext checks if next is safe). The specification in Figure 1 states the following safety property: if an iterator i is created for a collection c, the contents of c should not be changed while i is being used; indeed, once the FSA enters state 4, no method calls of the iterator are allowed (state 4 does not accept hasNext or next). A violation of this property results in a runtime exception. Figure 1 also shows a typical usage pattern of Iterator: every call to next is guarded by a call to hasNext.

The use of parameters is crucial in order to properly distinguish among different object interactions. Indeed, if one omits the parameters of the specification in Figure 1, then one would wrongly report a protocol violation when, e.g., two consecutive next calls are observed on two distinct iterators. Current parametric specification monitoring approaches maintain a clean separation between different object interactions, each interaction being observed by a distinct monitor; if an event is relevant to several monitors, e.g., an update event of a collection c which is relevant to all iterators i over c, then that event is dispatched to all interested monitors [5]. This way, monitors are not perturbed by irrelevant events or by unfortunate interaction interleavings. The major objective of this paper is to present a technique that achieves the same degree of separation between different object interactions, but for mining instead of monitoring.

Numerous specification mining approaches have been proposed, e.g., [2, 15, 20, 9, 1, 14, 17, 18, 8, 3, 13, 21, 7, 10] among many others. Parametric specifications are much more challenging to mine than non-parametric ones, mainly due to the complexity of handling parameters. Indeed, as discussed in Section 7, there are no approaches that provide a satisfactory solution for mining parametric specifications.
in their full generality and invulnerability to "unfortunate" interaction interleaving. For example, no approach that we are aware of is able to mine the specification in Figure 1 in the presence of arbitrary interaction interleavings.

**Contributions.** In this paper we present an effective and generic dynamic approach for mining parametric specifications. We strongly separate the tasks of parameter handling and of specification learning, which has two major benefits: (1) parameter handling makes no assumption on the types of specifications to mine, resulting in a generic parametric mining framework; and (2) learning is not affected by parameters or perturbed by interaction interleavings, reducing the effort to re-use existing algorithms (or to develop new ones) and increasing the overall precision of mining. The presented approach has been implemented in and extensively evaluated with jmMiner. The major tasks performed by jmMiner are: run the observed program and collect a parametric execution trace; slice the obtained parametric trace into a set of independent object interactions, called trace slices; pass the resulting trace slices to any learner; finally, put everything together into one or more parametric specifications. When used on packages that provide unit tests, jmMiner provides a completely automated parametric specification mining solution: building upon the hypothesis that unit tests are deviced to stress interactions among objects and methods that are likely to obey some protocol, jmMiner learns the events of interest by first running and observing the existing unit tests. This worked well in practice and, indeed, jmMiner has been used to automatically mine many properties of OpenJDK 6, like the one in Figure 1.

The rest of this paper is organized as follows. Section 2 highlights our overall approach to mining parametric specifications. Section 3 explains our novel trace slicing technique. Section 4 describes the other parts of our framework in detail. Sections 5 and 6 discuss experiments and limitations. Section 7 discusses related work. Section 8 concludes.

## 2. APPROACH OVERVIEW

Here we give a high-level overview of our mining approach.

**Definition 1.** *(Event specification)* We write meth-ods as \( m(T_1, T_2, \ldots, T_{n}) \), where \( m \) is the method name, \( T_i \) is the receiver type, \( T_e \) is the return type, and \( T_{p1}, \ldots, T_{pn} \) are the types of its parameters; for uniformity, we call each \( T_i, T_r, T_{p1}, \ldots, T_{pn} \) a method parameter. If \( M \) is a set of methods, let \( X_M \) be all the method parameters of reference type for all methods in \( M \). An event specification is a pair \((M, X)\), where \( M \) is a set of methods and \( X \subseteq X_M \).

For example, consider a set of methods \( M = \{\text{iterator.hasNext(Iterator, boolean)}, \text{iterator.next(Iterator, Object)}\} \). Collection iterator \((\text{Collection, Iterator, Object})\) because boolean is not a reference type. \((M, X_M)\) forms an event specification, but one may prefer event specification \((M, X)\), where \( X \) drops Object from \( X_M \). An event specification defines a set of related methods and their parameters that would likely obey a meaningful parametric specification.

Parametric specifications are specifications adding parameters to properties specified using any (non-parametric) trace formalism [4, 5]. Here we only consider regular patterns, definable through FSA; we only used FSA learners so far.

**Definition 2.** A parametric specification is a tuple \((S, M, X, s_0, \delta, F)\), where \((M, X)\) is an event specification and \((S, M, s_0, \delta, F)\) is an FSA, where \( S \) is the set of states, \( s_0 \) is the initial state, \( \delta : S \times M \rightarrow S \) is the transition (partial) function, and \( F \subseteq S \) is the set of final states. Figure 1 shows one such parametric specification. As a notational convenience, we write methods in \( M \) with their parameters in \( X \) (and drop the other parameters in \( X_M - X \)).

Our parametric specification mining approach consists of two stages, also giving the architecture of jmMiner, as depicted in Figure 2: event specification mining (Section 4.1) and parametric specification mining (Sections 3 and 4.2). The former yields a set of event specifications. The latter mines a parametric specification for each event specification.

Providing precise event specifications is inconvenient, since it requires expert knowledge about the target software. A set of overly diverse methods would result in a too complex specification likely to be application-specific, while a set of too few methods would result in a too simple specification covering only part of the usage pattern. Our approach is to automatically learn event specifications from unit tests, when available, based on the hypothesis that unit tests check the behavior of tightly interacting objects and thus the methods involved in such interactions likely obey some specification.

The parametric specification mining stage takes an event specification and various program execution traces as input, and yields a parametric specification as output. To obtain program execution traces, we wrote a Java Virtual Machine Tool Interface (JVMTI)[11] agent that logs information about the invoked methods and their arguments and attached it to the JVM. JVMTI provides a convenient means to access the call stack and to uniquely identify objects. Our JVMTI agent records all method invocations from all threads in chronological order, so that interactions that span over multiple threads can be recognized. The events and parameters which are not relevant for the given event specification are then filtered out. Figure 3 shows a resulting execution trace fragment when the event specification is \({\{\text{Collection.add(Collection, Iterator, Iterator.hasNext(Iterator, Iterator.hasNext))}, \text{Collection.iterator}\}}\). Each line corresponds to a method invocation and contains a method name together with the paired actual type and object identifier for each method parameter. Note that our execution trace filtering takes into account Java’s implicit upcasting. For example, \{Arraylist.add is not
Figure 3: Fragment of an execution trace.

filtered out because the method add is defined by ArrayList’s superclass Collection and is included in the event specification.

As explained in Section 1, the event parameters in execution traces play a crucial role. Indeed, if one dropped the parameters then one would see just one large and likely unreal (and thus unsuitable for mining) interaction that interleaves many different object interactions. For example, in the trace in Figure 3, events 8 and 9 are unrelated since their Iterator objects are different. However, a specification learner making abstraction of parameters would infer a faulty specification allowing two consecutive events.

To infer accurate specifications no matter how interactions interleave, independent interactions need to be considered separately. To this end, our approach introduces trace slicing. Informally, an interaction is a series of events that manipulate a set of objects. For example, two interactions exist in Figure 3: [add, iterator, hasNext, next, hasNext, add] (events 1, 2, 3, 4, 5 and 9) and [iterator, hasNext, next] (events 6, 7 and 8) on ArrayList. Our trace slicer identifies such interactions and passes them separately to the specification learner.

A specification learner finally infers a parametric specification from the set of separate interactions. The learner need not be aware of parameters and, consequently, one can reuse existing learners and/or easily develop new ones. If an FSA learner is employed, it would infer the pattern of alternation between hasNext and next from the above two separate interactions. The inferred FSA is finally annotated with parameters, which produces a parametric specification.

3. SLICING TRACES

In this section, we first define trace slicing, a process that dispatches events (in the given execution trace) to trace slices corresponding to different parameter bindings, according to the given event specification. We then introduce the concept of complete and connected parameter instances in order to remove trace slices that are meaningless and thus can generate noise in the process of mining specifications.

We show that a parametric trace can comprise, in the worst-case scenario, exponentially many trace slices (corresponding to complete and connected parameter instances).

3.1 Parametric Trace Slicing

Our terminology used in this section is borrowed from [5].

Definition 3. Let \( \mathcal{E} \) be a set of non-parametric events, called base events or simply events. An \( \mathcal{E} \)-trace, or trace, is a finite sequence of events in \( \mathcal{E} \), i.e., an element in \( \mathcal{E}^* \). We write \( e \in w \) when event \( e \in \mathcal{E} \) appears in trace \( w \in \mathcal{E}^* \).

Let \( [A \rightarrow B] \) and \( [A \rightarrow B] \) denote the sets of total and respectively partial functions from \( A \) to \( B \). What follows extends the definition above to parametric events and traces.

Definition 4. (Parametric events and traces) Let \( X \) be a set of parameters and let \( V_X \) be a set of corresponding parameter values. If \( \mathcal{E} \) is a set of base events (Def. 3), then let \( \mathcal{E}(X) \) denote the set of corresponding parametric events \( e(\theta) \), where \( e \) is a base event in \( \mathcal{E} \) and \( \theta \) is a partial function in \( X \rightarrow V_X \). Let \( \text{Dom}(\theta) \) be \( \{ x \in X \mid \theta(x) \text{ defined} \} \) and \( \perp \in X \rightarrow V_X \) be the map undefined everywhere; i.e., \( \text{Dom}(\perp) = \emptyset \). Partial maps in \( X \rightarrow V_X \) may also be called parameter instances or parameter bindings. A parametric trace is a trace with events in \( \mathcal{E}(X) \), that is, a word in \( \mathcal{E}(X)^* \).

In Figure 3, \( \mathcal{E} = \{ \text{add}, \text{iterator}, \text{hasNext}, \text{next} \} \), \( X = \{ \text{Collection}, \text{iterator} \} \) and \( V_X = \{ 158, 119, \ldots \} \). add and \( \{ \text{ArrayList} : 158 \} \) are the base event and, resp., the parameter binding of the first parametric event in the parametric trace shown in Figure 3.

Definition 5. We say that \( \theta \) is less informative than \( \theta' \), written \( \theta' \sqsubseteq \theta \), iff for any \( x \in X \), if \( \theta'(x) \) is defined then \( \theta(x) \) is also defined and \( \theta'(x) = \theta(x) \).

For example, \( \{ \text{ArrayList} : 158 \} \) \( \sqsubseteq \) \( \{ \text{ArrayList} : 158, \text{AbstractList} : 119 \} \).\]

Definition 6. (Trace slicing) Given parametric trace \( \tau \in \mathcal{E}(X)^* \) and partial function \( \theta : X \rightarrow V_X \), let the \( \theta \)-trace slice \( \tau_{\theta} \) of \( \tau \) be the non-parametric trace in \( \mathcal{E}^* \) defined as:

\[
\begin{align*}
\text{let } \ell_\theta &= \text{empty trace/word}, \\
\ell_\theta(x) &= \text{empty trace/word when } \theta(x) \text{ undefined, and} \\
\{ \tau(e(\theta)) \} &= \text{tuple } \{ (\tau(x))_{\theta(x)} \text{ when } \theta(x) \text{ defined} \} \quad (\tau(e(\perp))_{\perp})
\end{align*}
\]

A trace slice \( \tau_{\theta} \) first filters out all the parametric events that are irrelevant to the parameter instance \( \theta \). For example, when the given parameter instance is \( \{ \text{ArrayList} : 158, \text{AbstractList} : 119 \} \) in Figure 3, the resulting trace slice does not contain event 6. Similarly, events 7 and 8 are also filtered out. Event 1, which binds only ArrayList, is added to the trace slice corresponding to \( \{ \text{ArrayList} : 158, \text{AbstractList} : 119 \} \) because \( \{ \text{ArrayList} : 158 \} \) \( \sqsubseteq \) \( \{ \text{ArrayList} : 158, \text{AbstractList} : 119 \} \).

A trace slice \( \tau_{\theta} \) also forgets the parameter bindings of parametric events. As a result, a trace slice is non-parametric and merely a list of base events. For example, the trace slice corresponding to \( \{ \text{ArrayList} : 158, \text{AbstractList} : 119 \} \) is \[ \{ \text{add}, \text{iterator}, \text{hasNext}, \text{next} \} \]. Dropping parameter information enables the parametric specification mining stage to use any learners as long as they take as input a set of strings, where a string is a list of base events.

Although the intuition is clear, developing efficient and correct trace slicing algorithms is non-trivial. First, traversing the trace more than once is undesirable due to efficiency concerns. Second, an event may contain an incomplete binding of the given set of parameters. For example, for the trace in Figure 3, if we choose \( \{ \text{Collection}, \text{iterator} \} \) as the set of parameters, an add event contains only a Collection parameter, leaving the iterator parameter unbound. Furthermore, an event may belong to multiple trace slices because its parameter instance can be less informative than many other parameter instances introduced by the trace. For example, if the trace in Figure 3 contained another event iterator, the first event would also belong to the trace slice of \( \{ \text{ArrayList} : 158, \text{AbstractList} : 219 \} \).

3.2 Complete and Connected Instances

Since a trace slice corresponds to a parameter instance, selecting an appropriate set of parameter instances is crucial in order to prevent meaningless trace slices (such trace slices may result in learning inaccurate specifications). To select only appropriate parameter instances, we use two criteria: completeness and connectedness.
Figure 4: Part of an execution trace used for mining the specification in Figure 10.

A parameter instance is complete if \( \text{Dom}(\theta) = X \), where \( X \) is the set of parameters in the given event specification. Incomplete parameter instances are considered inappropriate, and trace slices corresponding to those parameter instances are suppressed; i.e., \( r_\theta \) is suppressed if \( \text{Dom}(\theta) \neq X \). For example, if \( X = \{\text{Collection}, \text{Iterator}\} \) in Figure 3, the trace slice corresponding to \( (\text{AbstractList}:119) \) is suppressed. The trace slice for this incomplete parameter instance is indeed meaningless because it does not represent an interaction between a collection and an iterator.

For some cases, no event provides a complete parameter instance. One such example is illustrated in Figure 4, when \( X = \{\text{Socket}, \text{SocketInputStream}, \text{SocketOutputStream}\} \). The first four events are part of one interaction (\( \text{Socket}:262, \text{SocketInput-Stream}:227 \)), but there is no event that provides the complete parameter instance. Parameter instances from multiple events therefore need to be combined.

**Definition 7.** Two parameter instances \( \theta \) and \( \theta' \) are compatible iff for any \( x \in \text{Dom}(\theta) \cap \text{Dom}(\theta') \), \( \theta(x) = \theta'(x) \). We can combine compatible parameter instances \( \theta \) and \( \theta' \), written \( \theta \sqcup \theta' \):

\[
\theta \sqcup \theta' = \begin{cases} 
\theta(x) & \text{when } \theta(x) \text{ is defined} \\
\theta'(x) & \text{when } \theta'(x) \text{ is defined} \\
\text{undefined} & \text{otherwise}
\end{cases}
\]

For example, \( \langle \text{Socket}:262, \text{SocketInputStream}:227 \rangle \) is compatible with \( \langle \text{Socket}:262, \text{SocketOutputStream}:288 \rangle \) but is not compatible with \( \langle \text{Socket}:562, \text{SocketOutputStream}:588 \rangle \). Two parameter instances disagreeing on any parameter are incompatible, and thus cannot be combined. \( \langle \text{Socket}:262, \text{SocketInputStream}:227 \rangle \sqcup \langle \text{Socket}:562, \text{SocketOutputStream}:288 \rangle \) is \( \langle \text{Socket}, \text{SocketInputStream}:227, \text{SocketOutputStream}:288 \rangle \), obtaining a complete instance when \( X = \{\text{Socket}, \text{SocketInputStream}, \text{SocketOutputStream}\} \).

Combining multiple parameter instances is therefore necessary for achieving a complete parameter instance. However, if done blindly, it may introduce spurious parameter instances. For example, \( \langle \text{Socket}:562, \text{SocketInputStream}:227, \text{SocketOutputStream}:588 \rangle \) obtained by combining \( \langle \text{Socket}:562, \text{SocketOutputStream}:588 \rangle \) and \( \langle \text{Socket}:262, \text{SocketOutputStream}:227 \rangle \) in Figure 4, is spurious because an interaction involving all three objects does not exist in the given trace. Trace slices corresponding to spurious parameter instances are noise for mining. To filter out such spurious parameter instances, we introduce the concept of connected parameter instances.

**Definition 8.** If \( \tau \in \mathcal{E}(X)^* \), we define \( \tau \)-connectedness of parameter instance \( \theta \) as follows: 1) if \( e(\theta) \in \tau \) then \( \theta \) is \( \tau \)-connected; and 2) if \( \theta_1, \theta_2 \) are \( \tau \)-connected, compatible, and \( \theta_1 \cap \theta_2 \neq \emptyset \), then \( \theta_1 \cup \theta_2 \) is also \( \tau \)-connected.

Therefore, a parameter instance is \( \tau \)-connected iff it is formed by combining parameter instances of events in \( \tau \) that share parameter bindings. For example, \( \langle \text{Socket}, \text{SocketInput-Stream}:227, \text{SocketOutputStream}:288 \rangle \) is \( \tau \)-connected in Figure 4 because of events 2 and 3, but \( \langle \text{Socket}, \text{SocketInput-Stream}:227, \text{SocketOutputStream}:588 \rangle \) is not. In cases where there is no ambiguity, we will say connected instead of \( \tau \)-connected.

Connectedness is motivated by the following observation: in most cases we are interested in mining specifications for a set of interacting objects; if two objects appear in the same event, then they interact with each other. Therefore, all the objects contained in a connected parameter instance directly or indirectly interact with one another. Experiments using our technique, discussed in Section 5, show that passing only the trace slices corresponding to connected parameter instances to the specification learner (and discarding the other trace slices) effectively removes noise in the mining process, resulting in accurate specifications.

Computing all possible connected parameter instances is hard. One should not mistakenly think that this problem reduces to computing the ordinary connected components of a graph, where a vertex represents a parameter instance and an edge exists iff the two associated parameter instances \( (\theta_1 \text{ and } \theta_2) \) are compatible and \( \theta_1 \cap \theta_2 \neq \emptyset \). Figure 5 shows one such graph. \( P, Q, R \) and 5 are parameters, and \( p_0, q_0, r_1, r_2 \), and \( s_0 \) are parameter values. The graph-connected component in Figure 5 correctly suggests that \( (P:p_0, Q:q_0) \cup \{Q:q_0, R:r_1\} \cup \{R:r_2, S:s_0\} \) is \( \tau \)-connected. However, it also suggests that \( (P:p_0, Q:q_0) \cup \{Q:q_0, R:r_2\} \cup \{R:r_1, S:s_0\} \) is \( \tau \)-connected, which is wrong. Indeed, computing the graph-connected components does not take into consideration the compatibility between parameter instances, while computing the \( \tau \)-connected parameter instances must. For example, \( \langle Q:q_0, R:r_1 \rangle \) and \( \langle Q:q_0, R:r_2 \rangle \) are incompatible, but the standard graph-connected component fails to recognize it.

### 3.3 Complexity of Trace Slicing

In what follows we calculate the worst-case complexity of the trace slicing problem in terms of the number of trace slices as a function of the total length \( n \) of the original parameter trace and the size of \( X \), the set of parameters. More precisely, we show that there are approximately \(^1\left(\frac{2^n}{n!}\right)^m \) trace slices in the worst case when \( m \geq 1 \), \( m+1 \) is the size of \( X \). Note that if \( |X| = 1 \) then we have at most \( n \) trace slices and they are easy to compute. However, if \( |X| = \frac{2}{3} + 1 \) then we have \( 2^\frac{2}{3} \) trace slices, showing that the addition of conflicting edges (like in Figure 5) makes the graph-connected component problem harder. The maximum of \( \left(\frac{2^n}{n!}\right)^m \) is actually reached when \( m = \frac{2}{3} \), in which case it becomes \( e^{2^\frac{2}{3}} \).

Suppose that \( X = \{P_0, P_1, \ldots, P_m\} \) for some \( m > 0 \) and that \( \tau = e_1(\theta_1) e_2(\theta_2) \ldots e_n(\theta_n) \). The worst case is when any two events have at least one common parameter value, so that \( \theta \cap \theta' \neq \emptyset \) for any two parameter instances \( \theta \) and \( \theta' \) such that \( e(\theta), e'(\theta') \in \tau \); we can achieve that with minimal resources, by designating a parameter instance \( P_0:p_0 \) and assuming that that is common to all events. Each event may be in conflict with a certain number of other events. For example, suppose that \( e_1(\theta_1) \) is in conflict with \( a_1 - 1 \) events on parameter \( P_1 \), where \( a_1 > 0 \). The other \( a_1 - 1 \) events are also in conflict with each other, so we have a “cluster” of \( a_1 \) events which are in conflict with each other on parameter \( P_1 \). The worst case is when the conflicting \( a_1 \) events are in

---

\(^1\)We do an approximate analysis, making abstraction of the fact that \( m \) may not divide \( n \), etc.
conflict with no other event and when, for each trace slice corresponding to the remaining events, each of them yields a new trace slice. Thus, assuming that the remaining events generate s trace slices, we have a × s slices in total. We can iterate the argument above and obtain a1 × a2 × ... × am trace slices when we split the n events of \( \tau \) into clusters of \( a_1, a_2, \ldots, a_m \) events with \( a_1 + a_2 + \cdots + a_m = n \), each cluster containing those events conflicting on precisely one of the parameters \( P_1, P_2, \ldots, P_m \), respectively. Note that this is not only an over-approximation. It can actually happen, as shown in Figure 6. The product is maximized when \( a_1 = a_2 = \cdots = a_m = \frac{n}{m} \), in which case it becomes \( (\frac{n}{m})^m \).

Therefore, assuming that \( X \) is fixed a priori, as it is usually the case, we can only have a polynomial (in the length of the original parametric trace) number of trace slices. If \( X \) is not fixed, then one can actually fabricate an absolute worst-case scenario, which maximizes \( (\frac{n}{m})^m \). This case occurs when \( m = \frac{n}{\log n} \), in which case the number of trace slices is exponential: \( e^\log n \). Although it is likely in practice that the size of \( X \) is co-related to the length of the trace, it is instructive to have a clear understanding of the worst-case complexity of the problem that we are attempting to solve.

### 3.4 Slicing Algorithm

As discussed in Section 3.3, the number of trace slices is \( (\frac{n}{m})^m \) in the worst case. Since all trace slices can be distinct, this number gives a lower bound for all trace slicing algorithms. This lower bound is hard to achieve, though, since computing complete and connected parameter instances may require several operations of instance combination. For example, \( \left\langle P_{0;0}, P_{1;1}, P_{2;2}, \ldots, P_{m;m} \right\rangle \) in Figure 6 can be obtained only after at least \( m \) instance combination operations: \((\left\langle P_{0;0}, P_{1;1} \right\rangle \cup \left\langle P_{0;0}, P_{2;2} \right\rangle \cup \left\langle P_{0;0}, P_{3;3} \right\rangle \cup \cdots \cup \left\langle P_{0;0}, P_{m;m} \right\rangle)\). Furthermore, a trace slicing algorithm needs to search for compatible parameter instances in order to create combined parameter instances.

Our trace slicing algorithm in Figure 7, called SLICER, traverses the given trace only once and avoids the construction of meaningless trace slices. SLICER has two stages: (1) it first processes the entire parametric trace, event by event, constructing intermediate results \( \Delta \); and (2) then it constructs the set of trace slices \( \Psi \), each corresponding to a complete and connected parameter instance.

During the first stage, SLICER stores in \( \Delta \) intermediate trace slices only for parameter instances that actually occur in observed events. Neither combined parameter instances nor trace slices for them are created at this stage. The second stage, \textsc{ConstructConnected}, constructs \( \Omega \) holding all possible connected parameter instances by combining compatible parameter instances in the loop at lines 2–3. For each complete and connected parameter instance, its corresponding trace slice is finally constructed at lines 4–6. \( \Gamma \) collects all intermediate trace slices corresponding to \( \theta \)-subinstances. \textsc{MergeTraces} is essentially the merge function of merge sort, using the position of events in the trace for comparison (events in trace slices are listed chronologically).

**Theorem 1.** After running SLICER on \( \tau \in \mathcal{E}(X) \),

1. \( \Psi(\theta) \) is defined iff \( \theta \) is \( \tau \)-connected and \( \text{Dom}(\theta) = X \).
2. If \( \Psi(\theta) \) is defined, then \( \Psi(\theta) = \tau_{\theta} \).

This theorem tells all trace slices can be retrieved from \( \Psi \).

We next analyze the complexity of SLICER. It first calls HandleEvent \( n \) times, and, assuming that a self-balancing binary search tree is used for \( \Delta \), the complexity of \textsc{HandleEvent} is \( O(\log n) \). The loop at lines 2–3 in \textsc{ConstructConnected} can pick \( \theta_1 \) and \( \theta_2 \) from \( \Omega \times \Omega \), and each iteration takes \( O(m) \) time for checking the compatibility and combining the two parameter instances. There are \( |\Omega| \) iterations of the loop at lines 4–6, with each iteration taking \( O(\log n) \) time. The running time of the entire algorithm is thus \( O(n \log n + |\Omega|^2 \cdot m + |\Omega| \cdot m) = O(n \log n + |\Omega|^2 \cdot m) \). Since the algorithm creates all possible connected parameter instances, \( |\Omega| \) can be calculated as follows: the number of connected parameters with \( |\text{Dom}(\theta)| = i + 1 \) is \( \binom{m}{i} \cdot (\frac{n}{m})^i \) because we can choose \( i \) parameters and there are \( \frac{n}{m} \) parameter values for each parameter. Thus, we have \( |\Omega| = \sum_{i=0}^{m} \binom{m}{i} \cdot (\frac{n}{m})^i = (\frac{n}{m} + 1)^m \), and the time complexity of SLICER is \( O(n \log n + (\frac{n}{m} + 1)^m \cdot m) = O((\frac{n}{m} + 1)^m \cdot m) \). For the space complexity, it needs to maintain \( O(|\Omega|) \) connected parameter instances of length \( O(m) \) during trace slicing. It also needs space for \( (\frac{n}{m})^m \) trace slices of size \( m \) as illustrated in Figure 6. Therefore, the space complexity is \( O((\frac{n}{m} + 1)^m \cdot m + (\frac{n}{m})^m \cdot m) = O((\frac{n}{m} + 1)^m \cdot m) \). SLICER iterates through all possible connected parameter instances in the loop at lines 2–3 in \textsc{ConstructConnected}. In our experiments, this step was comparatively the most expensive w.r.t. performance, so we have investigated several possibilities to optimize it. We next describe two of our optimizations which bring considerable performance benefits. Instead of blindly picking a pair of parameter instances from \( \Omega \) and combining them, our implementation proceeds in a bottom-up manner. At the first step, it picks two parameter instances \( (\theta_1 \text{ and } \theta_2) \) such that \( |\text{Dom}(\theta_1)| = |\text{Dom}(\theta_2)| = N \), and creates \( \theta_1 \cup \theta_2 \), if necessary. After handling all parameter instances with \( N \) parameter bindings, it picks parameter instances with \( N+1 \) parameter bindings, and so on and so forth until \( N \) reaches the size of \( X \), the set of parameters. This way, a parameter instance is
considered for compatibility within only a limited window, reducing the number of iterations.

Our second optimization is to group parameter instances so that all parameter instances in the same group bind exactly the same parameters. Grouping also reduces the number of iterations at lines 2-3 in `CONSTRUCTCONNECTED`. For example, if \( (p_1, q_1) \) is chosen as \( \theta_1 \), all parameter instances that belong to the group corresponding to \( (R, S) \) will be excluded from the list of candidates for \( \theta_2 \) because any parameter instance in this group would result in \( \theta_1 \cap \theta_2 = \emptyset \).

4. LEARNING IN JMINER

Here we discuss how the trace slicing technique in Section 3 is incorporated in our JMiner parametric mining approach, by means of discussing JMiner’s event specification learner (from unit tests) as well as its specification learner (based on a refinement of the off-the-shelf PFSA learner[16]).

4.1 Mining Event Specifications

The event specification learner dynamically infers a set of event specifications from the target software. It takes as input the source code, a target package name, and unit tests (all for the target software). Providing these inputs is easy and requires no expert knowledge. For example, in order to mine specifications in the java.util package of OpenJDK 6, what the user needs to provide is: the source code of OpenJDK 6, the target package name java.util, and the unit tests for java.util. Here is, e.g., a typical unit test in OpenJDK 6:

```java
1. import java.util.*;
2. public class CheckForComodification {
3.   private static final int LENGTH = 10;
4.   public static void main(String[] args) throws Exception {
5.     List<Integer> list = new ArrayList<Integer>;
6.     for (int i = 0; i < LENGTH; i++) list.add(i);
7.     try {
8.         if (i == LENGTH - 2) list.remove(i);
9.     } catch(ConcurrentModificationException e) { return; }
10.    throw new RuntimeException
11.    } }
```

This unit test case is written to test if a concurrent modification of a Collection object is detected and a runtime exception is raised. As the Java compiler translates the for-each loop at lines 7-8 into `ArrayList.iterator(AbstractListStr, hasNext and AbstractListStr, next)`, an execution of this test case will reveal an interaction between `ArrayList` and `AbstractListStr`.

An advantage of using unit tests for mining event specifications is that tightly interacting objects are well isolated. For example, the unit test case above considers only one issue, namely detecting a concurrent modification. This isolation avoids tangential relationships among issues, which are usually application-specific and thus not likely to obey a generic specification. For example, an interaction on a `FileReader` object and an interaction on a `FileWriter` object can be related in an application through a `File` object, because both `FileReader` and `FileWriter` objects can be constructed with the same `File` object. Although somewhat related, it is expected that reading and writing are two separate issues and thus specifications involving both are unnecessary. Such tangential relationships rarely occur in unit tests. Moreover, users can obtain unit tests for free as many software packages are shipped with them. Unit tests are well maintained as they are frequently run and failed cases are promptly addressed.

Learning related methods and parameters. We first trace unit test executions using a JVTMI agent like the one in Section 2, but one which also logs the thread identifier and the depth of the call stack in front of each event (each thread has its stack and every method invocation is logged). Here is, e.g., part of the trace logged by the unit test above:

```
1. 1 2 ArrayList.<init>(ArrayList:689)
2. 1 2 ArrayList.add(ArrayList:689, Integer:830)
3. 1 2 ArrayList.ensureCapacity(ArrayList:689)
4. 1 2 AbstractList.iterator(ArrayList:689, AbstractListStr:950)
5. 1 2 AbstractListStr.hasNext(AbstractListStr:950)
6. 1 3 AbstractListStr.size(AbstractListStr:689)
```

One can infer that add invoked ensureCapacity; and (init), add, iterator, hasNext and next were invoked in order by the same method (not shown here), since they have the same thread identifier and depth of the call stack. Although a trace from a unit test case usually contains relatively few events, it may still be too large to precisely infer related methods.

We next analyze the execution trace in order to remove irrelevant events. Many heuristics are possible here; we prefer to use two heuristics which were also used in [19, 15] and appear to work well: keep only events corresponding to methods (1) which are defined in the user-specified target package, and (2) which are directly invoked by methods of the class that declares the main entry of a test case. Our rationale for using these heuristics is that unit tests are rarely written for interactions among multiple packages, and that a unit test consists of one core class for performing the actual test and other utility classes for supporting the core class.

All the relevant events are then partitioned into groups of related events. Two events are directly related if they share at least one common argument, and related if they are connected through a sequence of directly related events. For example, (init) (event 1) and add (event 2) are directly related to `ArrayList:689`, and (init) (event 1) and next (event 7) are related through iterator (event 4).

An event specification is then created for each partition. For each object used as a receiver in a partition, its type is generalized to the least specific type that specifies all methods involving that object. Then, the least specific type and all the involved methods (after their declaring types are generalized to the least specific types) are added as a parameter and, respectively, methods. For the trace above, e.g., a partition including `ArrayList:689` and `AbstractListStr:950` is first generated. Then, `ArrayList` is generalized to `AbstractList` because `AbstractList` specifies all involved methods. As a result, a parameter `AbstractList` and the generalized methods (`AbstractList.(init)`, `AbstractList.add` and `AbstractList.iterator`) are added. Similarly, `AbstractListStr:950` adds a parameter `Iterator` and two methods (`Iterator.hasNext` and `Iterator.next`) These two parameters and five methods form an event specification.

Filtering out generics parameters. Generics may yield undesirable event specifications. Consider, for example, the following execution trace:

```
1. ArrayList.<init>(ArrayList:62)
4. AbstractListStr.hasNext(AbstractListStr:519)
5. AbstractListStr.next(AbstractListStr:519, Locale:57)
6. ArrayList.<init>(ArrayList:1279)
```

The above event specification learner would identify `[hasNext, (init), add]` (events 4, 5, 6 and 7) as the interaction on `ArrayList:1279`, `AbstractListStr:519`, `Locale:57` (`ArrayList` and `AbstractListStr` are subclasses of `AbstractList` and `Iterator`, respectively). This interaction is spurious as it is about conec-
4.2 Learning Parametric Specifications

\textit{jMiner} takes the event specifications learned from unit tests as explained above and passes them to its core component, the trace slicer. As discussed in Sections 2 and 3, the trace slicer takes an event specification and an execution trace produced by a program that exercises the target package, and produces a set of trace slices, each corresponding to an observed interaction. The generated trace slices are then passed to any non-parametric specification learner. We included one such specification learner in \textit{jMiner}, which we discuss below. Our learner consists of two components: an off-the-shelf PFSA learner and an FSA refiner.

\textbf{PFSA learner}. A PFSA is an FSA where each transition is labelled with how often the transition occurs. A PFSA learner takes a set of strings as input and infers a PFSA. Several PFSA learning approaches have been proposed [16]; we here adopt the sk-strings algorithm [16], since it performs well at inferring small FSAs. It first constructs an FSA that precisely accepts the input set of strings. Each transition is then annotated with a frequency, saying how many times that transition was observed. The sk-strings algorithm then generalizes by merging states which are sk-equivalent: two states are sk-equivalent iff corresponding sets of bounded strings (ones that are frequently generated from each of the two states) are matched. As a result of this approximation, two states can be merged even when they are not strictly equivalent, making it possible for the inferred FSA to accept not only the input strings but also other “similar” strings. The reader is referred to [16] for more details. After running the sk-strings algorithm, the learner drops the frequency information, yielding an ordinary FSA. Figure 8 shows the FSA inferred by the sk-strings algorithm for our Collection-\textit{iterator} example (Figure 1 shows the desired specification).

\textbf{FSA refiner}. Although PFSA learner’s approximations are generally desirable in many application domains, the resulting FSAs turned out to often be overly general in our domain of mining specifications from execution traces, in that the mined FSAs accept a large number of undesirable traces. For example, the trace \{\langle init\rangle, \text{iterator}, \text{hasNext}, \text{next}\} is accepted by the FSA in Figure 8, but it is impossible to occur in any program execution (only one \text{iterator} event can be observed in any interaction between a collection and an \text{iterator}). To prevent over-generalized specifications, our specification learner implemented in \textit{jMiner} refines the FSA inferred by the sk-strings algorithm.

The overall goal of our refiner is to eliminate transitions caused by over-generalization while keeping desirably generalized transitions. An obvious step for avoiding over generalization is to remove all the transitions that are never taken by any of the trace slices provided as the input of the specification learner. For example, the \text{iterator} transition from state 4 to state 2 in Figure 8 is never taken (the same \text{iterator} object cannot be created twice), so it can be safely removed. However, that is not enough: the resulting FSA still accepts infeasible interactions containing multiple \text{iterators}, e.g., \{\langle init\rangle, iterator, hasNext, update, iterator, hasNext\}.

The fundamental problem is that \textit{jMiner} learners do not take the context into account when merging states, while the context is important in programming languages. For example, the two transitions from states 3 and 4 to state 1 in Figure 8 were obtained by the PFSA learner by merging two contextually different states. Modifying a PFSA learner to take the context into account seems hard. Instead, our approach is to “reverse engineer” the FSA generated by the PFSA learner to partly take the context into account. More precisely, our refiner first expands the inferred FSA as follows: if a state \(s\) has \(n\) incoming edges from the other states, then \(s\) is replaced by \(n\) corresponding states \((s_1, s_2, \ldots, s_n)\) in order to differentiate the originating state. Figure 9 shows the expanded FSA of the one in Figure 8. States \(1_2\) and \(1_3\) indicate that an iterator has been already used, whereas state \(1_4\) indicates that no iterators have been used.

Using the expanded FSA, the refiner removes all the transitions that are never taken by the given trace slices (i.e., caused by over-generalization), in our case \(1_2 \rightarrow 2_1, 1_3 \rightarrow 2_1\), and \(4_1 \rightarrow 2_2\). Then the refiner finally eliminates the states with no incoming transitions and merges states that have the same outgoing transitions. For example, state \(2_2\) is eliminated, and states \(1_2, 1_3\) and states \(3_1, 3_2\) are merged, respectively. The resulting FSA for the one in Figure 9 is shown in Figure 1. The reader is referred to [6] for more details.

5. EVALUATION OF \textit{jMiner}

We applied \textit{jMiner} to mining specifications in four packages of OpenJDK 6: java.\text{util}, java.\text{io}, java.\text{lang} and java.\text{net}. OpenJDK 6 contains various unit tests and is well docu-
mented, allowing us to validate the mined specifications. Selection of the smallest number of events (event specification mining) is easy because the unit tests are well structured.

We used execution traces obtained from the DaCapo benchmark suite 9.12 and the Apache JAMES Server 2.3.1 as training sets for specification mining. We used DaCapo (which contains 14 programs and a harness to execute each program) for java.util, java.io and java.lang, and Apache JAMES (which contains several test cases for whole-system checking) for java.net. Table 1 shows information on the traces used in our experiments. Two kinds of traces were used: traces for mining event specifications and traces for mining parametric specifications. We limited the execution time for each program in DaCapo to one hour.

Table 1: Traces used in the experiments.

<table>
<thead>
<tr>
<th>Target package</th>
<th>Event Specifications</th>
<th>Parametric Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># files</td>
<td># events</td>
</tr>
<tr>
<td>java.io</td>
<td>382</td>
<td>28,835,588</td>
</tr>
<tr>
<td>java.lang</td>
<td>372</td>
<td>41,784,568</td>
</tr>
<tr>
<td>java.util</td>
<td>370</td>
<td>65,854,349</td>
</tr>
<tr>
<td>java.net</td>
<td>221</td>
<td>9,429,744</td>
</tr>
</tbody>
</table>

Table 3: Times (minutes; Windows, 3GHz, 1GB).

Figure 10 shows a specification of a client-side stream socket. The constructor of Socket connects a new socket to the peer specified by its arguments. Then, getInputStream(), getOutputStream(), and accept() are used to connect to the remote socket.

ServerSocket. Figure 11 shows an inferred specification for the server-side socket. After a ServerSocket object s is instantiated, accept listens for a connection and accepts it, returning a new socket e. getInputStream and getOutputStream return an InputStream object i and an OutputStream object o respectively, which can be used for data transmission. After these operations, close can be invoked to close the connection. This behavior spans over multiple threads in most cases because multiple clients can connect to the same port represented by a single ServerSocket object, and a server needs to handle them simultaneously. The trace slices in our experiments indeed involved two threads: the data transfer was processed in a separate thread. If each thread’s trace was considered separately like in [15], the specification could not be mined.

Figure 11: ServerSocket specification mined with jMiner.
Knowing that \texttt{add}, \texttt{remove}, \texttt{update} and \texttt{clear} are similar (all these methods update a \texttt{Collection}), we merged the three methods into one hypothetical method, which we called \texttt{update}.

\begin{figure}[h]
\centering
\begin{tikzpicture}

\t\node (init) at (0,0) {0};
\t\node (r) at (1,0) {1};
\t\node (close) at (2,0) {2};
\t\node (read) at (1,-1) {<init>(s)};
\t\node (close2) at (1,-2) {close(r)};
\t\node (in) at (1,-1.7) {read(r)};
\t\node (out) at (1,-2.6) {0 1 2};
\t\node (close2) at (1,-2.9) {close(r)};

\t\path[->] (init) edge (r)
\t\node (close) at (2,0) {2};
\t\node (read) at (1,-1) {<init>(s)};
\t\node (close2) at (1,-2) {close(r)};
\t\node (in) at (1,-1.7) {read(r)};
\t\node (out) at (1,-2.6) {0 1 2};
\t\node (close2) at (1,-2.9) {close(r)};
\end{tikzpicture}
\caption{Reader specification mined using JMiner.}
\end{figure}

\textit{Reader, Writer.} Figure 12 shows a specification of a \texttt{Reader} object, stating that \texttt{read} can be repeatedly called before close. It does not enforce the invocation of close, similarly to the Socket specification above. JMiner also mined a similar specification for \texttt{Writer}. These specifications are simple, but can detect the wrong invocation of \texttt{read or write} after close.

\section{Limitations}

We have identified, during our experiments, a few limitations of our approach. First, the learning process is limited to the observed behaviors. This is an inherent limitation of all dynamic approaches. For example, the specifications in Figures 10 and 11 wrongly enforce the order between \texttt{getInputStream} and \texttt{getOutputStream} because this was consistently observed in the training set. Also, in Figure 13 one may expect that the specification should state that \texttt{nextToken} is guarded by \texttt{hasMoreTokens}. Surprisingly, the inferred one allows the invocation of the two methods in an arbitrary order since it was actually observed that Xalan, an application in DaCapo, calls \texttt{nextToken} without calling \texttt{hasMoreTokens}.

\begin{figure}[h]
\centering
\begin{tikzpicture}

\t\node (init) at (0,0) {0};
\t\node (r) at (1,0) {1};
\t\node (close) at (2,0) {2};
\t\node (in) at (1,-1) {<init>(s)};
\t\node (close2) at (1,-2) {close(r)};
\t\node (read) at (1,-1.7) {read(r)};
\t\node (out) at (1,-2.6) {0 1 2};
\t\node (close2) at (1,-2.9) {close(r)};
\end{tikzpicture}
\caption{StringTokenizer spec. mined using JMiner.}
\end{figure}

After inspecting the source code of Xalan, we could see that the pattern is not defective because Xalan first retrieves the number of tokens by calling \texttt{countTokens} and then consecutively calls \texttt{nextToken} as many times as specified by \texttt{countTokens}. Due to \texttt{countTokens}, a specification on \texttt{StringTokenizer} cannot be stricter than Figure 13. Considering \texttt{countTokens} as well does not improve the specification because our technique cannot infer that the return value of this method indicates the number of allowed \texttt{nextToken} calls. This limitation is inherent to all FSA-based approaches: FSA cannot count.

\section{Related Work}

Ammons et al. [2] propose a technique for mining specifications from execution traces and user-provided input: functions of interest, attributes for those functions, and a scenario seed. It extracts a set of API usage scenarios from execution traces and then passes it to a PFSA learner. Providing attributes requires in-depth knowledge (one should understand the side-effect of each function); the user should imagine a hypothetical object corresponding to a scenario, and should mark a parameter as \texttt{define} or, respectively, as \texttt{use} if the parameter changes or depends upon the state of the object. Scenarios are identified starting from the seed event, searching the execution trace along \texttt{define-use} chain. Having explicit seed events and using the chain reduce the search space, but it may result in failing to recognize complete interactions. For example, assume the following trace:

1. \texttt{ArrayList.add(Collection:1)}
2. \texttt{AbstractList.iterator(Collection:1, Iterator:2)}
3. \texttt{AbstractList.iterator(Collection:1, Iterator:3)}
4. \texttt{AbstractList.iterator(Collection:1, Iterator:4)}
5. \texttt{ArrayList.remove(Collection:1)}
6. \texttt{ArrayList.add(Collection:1)}

Also suppose that the seed event is \texttt{iterator} (it is the only event that connects \texttt{Collection} and \texttt{Iterator}) and that \texttt{add} \texttt{defines} \texttt{Collection}, \texttt{iterator defines} \texttt{iterator} and \texttt{uses} \texttt{Collection}, and \texttt{hasNext uses} \texttt{Iterator}. For example, event 2 depends on event 1 \texttt{defines} \texttt{Collection:1} and event 2 \texttt{uses} \texttt{Collection:1}. From these inputs, two scenarios will be extracted: 1 2 3 and 1 4 5. None of them are complete with regards to the interaction between \texttt{Collection} and \texttt{Iterator}: none includes event 6 because \texttt{add} does not \texttt{use} \texttt{Collection} and, consequently, event 6 cannot be reached along \texttt{define-use} chain in either of the scenarios. As a result, the inferred FSA will lack the transition from \texttt{hasNext} to \texttt{add}, preventing any update after using an \texttt{iterator}. Marking \texttt{add} as both \texttt{define} and \texttt{use} is not proper either; it will add both events 4 and 6 to the scenario of events 1, 2 and 3, which is wrong because event 6 does not belong to the same interaction. No matter how attributes are adjusted, their technique cannot infer the comprehensive and correct FSA shown in Figure 1 in situations where JMiner can.

Pradel and Gross [15] propose a dynamic mining technique based on collecting from execution traces a list of related receiver-method pairs up to a user-specified level of nested method calls. Unlike ours, their technique does not consider individual interactions separately. Therefore, it may merge individual interactions and thus infer inaccurate specifications. For example, if the execution trace in Figure 3 is observed within a method, their technique will not consider the two interactions separately and, consequently, infer a faulty specification that allows consecutive calls to \texttt{next}. Moreover, it cannot infer a specification that spans over multiple threads, since it creates a separate trace for each thread; e.g., the specification in Figure 11 cannot be mined. Furthermore, it may fail to mine specifications from distantly related events if the value of the level of nested method calls is too small. If, on the other hand, the value is too large, it may produce specifications that include too many methods and would likely be application-specific.

Yang et al. [20] propose a technique to find all pairs of methods that satisfy the predefined particular pattern (ab)c* from execution traces. Although their chaining heuristic composes somewhat more complex patterns, such as (abc)* (by connecting related specifications into a chain), it cannot infer useful and more complex patterns like in Figure 1. Gabel and Su [9] extend [20]; their work considers an additional predefined pattern (ab\(c^*\)). It then combines instances of these basic patterns, generating complex patterns. Unlike our approach, it neglects parameters; thus, it may infer noisy specifications from sequences of irrelevant events that happen to match the predefined patterns.

Dallmeier et al. [7] also present a technique for mining FSAs from execution traces. A state in the FSA inferred by their work represents the results of inspector methods that observe the internal state, such as is\texttt{Empty} and \texttt{hasNext}, whereas a state in our approach is abstract (e.g., “before using an \texttt{Iterator}”). Associating each state with inspectors can help users to easily understand the specification, but it fails to capture implicit states such as “an \texttt{iterator} for a \texttt{collection is being used}” because no methods in \texttt{Collection} can...
observe it. In our approach, the sequence of method calls can be imperfect, which is the reason why their work considers only one object and is essentially non-parametric.

Lorenzoli et al. [14] propose an advanced algorithm to mine extended finite state machines (EFSMs), i.e., FSMs extended with state constraints. Their approach proposes no trace slicing technique, which is at the core of our approach; instead, they assume that the training traces are already given as separate interactions with constraints. In principle, one could use state constraints to encode parameter instances, but considering the huge number of parameter instances encountered in our experiments, we believe that would be impractical. We have only looked at mining parametric FSAs; it may, however, be beneficial to plug their learner into jMiner and use the latter for learning EFSMs from trace slices produced by our algorithm in Section 3.

Henkel et al. [10] present a dynamic mining technique which is specific to container classes. Their technique actively constructs various operations (by invoking methods of a container), observes the state of the container, and then infers relations among distinct operations, such as state equivalence. In their technique, parameters are predefined and interactions on them are not considered; parameters are inserted into a container and solely used to determine the state of the enclosing container. In contrast, our technique passively observes interactions on parameters occurring in existing programs and then infers the FSA by generalizing all the observed interactions.

Acharya et al. [1] propose a static technique that generates a set of traces along possible execution paths directly from the source code, and then produces an API usage pattern from it. Since it mines partial orders, the resulting specifications cannot describe loops; thus, it cannot mine the specifications shown in this paper. Zhong et al. [21] also present a static mining technique for sequential patterns from open source repositories. Unlike our approach, their tool does not consider individual interactions separately. For example, if there are multiple distinct interactions on Collection in a method, their tool can extract a faulty method call sequence. Since their tool inlines multiple methods, the probability that a method call sequence consists of multiple interactions on Collection is high, which makes this approach improper to mine specifications of frequently used classes.

Chen and Rosu [5] introduce trace slicing for monitoring. There is an inherent duality between parametric specification monitoring and parametric specification mining: they both rely on a parametric trace slicing process for identifying interactions, followed either by monitoring the resulting trace slices against given specifications in the first case, or by inferring the specifications that best explain the observed trace slices in the second case. However, a blind use of off-the-shelf trace slicing techniques for monitoring leads to noisy and inefficient trace slicers for mining. It is allowed for trace slicers for monitoring to generate trace slices corresponding to incomplete or unconnected parameter instances because such trace slices can be ignored by the underlying monitor. In the context of mining, however, such trace slices would result in faulty specifications.

8. CONCLUSION

This paper presented a generic and automated approach to mine accurate parametric specifications from execution traces with minimal effort. An event specification learner automatically infers event specifications from unit tests; a trace slicer then identifies independent interactions, allowing one to apply various learning techniques that do not handle parameters. An automaton refiner makes the inferred specifications more accurate by eliminating spurious transitions. Experiments indicate that the technique is effective.

9. REFERENCES