Java Ranger: Static Regions for Efficient Symbolic Execution of Java

***Abstract*—Merging related execution paths is a powerful technique for reducing path explosion in symbolic execution. One approach, introduced and dubbed "veritesting" by Avgerinos et al., works by statically translating a bounded control flow region into a single formula. This approach is a convenient way to achieve path merging as a modification to a pre-existing single-path symbolic execution engine. Avgerinos et al. evaluated their approach in a symbolic execution tool for binary code, but different design considerations apply when building tools for other languages. In this paper we explore the best way to use a veritesting approach in the symbolic execution of Java.**

Because Java code typically contains many small dynamically dispatched methods, it is important to include them in veritesting regions; we introduce a *higher-order* veritesting technique to do so modularly. Java's typed memory structure is very different from a binary, but we show how the idea of static single assignment (SSA) form can be applied to object references to statically account for aliasing. More formally, we describe our veritesting algorithms as syntax-directed transformations of a structured intermediate representation, which highlights their logical structure. We have implemented our algorithms in Java Ranger, an extension to the widely used Symbolic Pathfinder tool for Java bytecode. Our empirical evaluation shows that veritesting greatly reduces the search space of Java symbolic execution benchmarks, while our expanded capabilities provided a significant further improvement.

I. Introduction

Symbolic execution is a popular analysis technique that performs non-standard execution of a program: data operations generate formulas over inputs, and the branch constrains along an execution path are combined into a predicate. Originally developed in the 1970s [1], [2], symbolic execution is a convenient building block for program analysis, since arbitrary query predicates can be combined with the logical program representation, and solutions to these constraints are program inputs illustrating the queried behavior. Some of the many application of symbolic execution include test generation [3], [4], equivalence checking [5], [6], vulnerability finding [7], [8], and protocol correctness checking [9]. Symbolic execution tools are available for many languages, including CREST [10] for C source code, KLEE [11] for C/C++ via LLVM, JDart [12] and Symbolic PathFinder [13] for Java, and S2E [14], FuzzBALL [15], and angr [8] for binary code. More here...explain the 'ecosystem' - tools for different languages: KLEE, FuzzBall, Java Symbolic Pathfinder, ...

Although symbolic analysis is a very popular technique, scalability is a substantial challenge for symbolic execution. Dynamic state merging [16] provides one way to alleviate scalability challenges by opportunistically merging dynamic symbolic executors, which can be performed on paths Add std. cite or on environments FM paper from 2014 on Javascript?. Other techniques include CEGAR/subsumption Add references from ASE 2017 paper: More Effective Interpolations in Software Model Checking.

Veritesting [17] is a different recently proposed technique that can dramatically improve the performance of symbolic execution. Rather than explicitly merge paths or check subsumption relationships, Veritesting simply encodes a local region of a program containing branches as a disjunctive region for symbolic analysis. If any path within the region meets an exit point, then the disjunctive formula is satisfiable. This often allows many paths to be collapsed into a single path involving the region. In previous work [17], bounded static code regions have been shown to find more bugs, and achieve more node and path coverage, when implemented at the X86 binary level for compiled C programs. This provides motivation for investigating integration of introducing static regions with symbolic execution at the Java bytecode level.

```
1  // x = ArrayList of symbolic integers with
2  // concrete length
3  for (int i = 0; i < x.size(); i++) {
4    // Begin region for static unrolling
5    if (x.get(i) < 0) sum += -1;
6    else if (x.get(i) > 0) sum += 1;
7    // End region for static unrolling
8  }
9  if (sum < 0) System.out.println("neg");
10 else if (sum > 0) System.out.println("pos");
11 else System.out.println("bug");
```

Listing 1: An example to loop through a symbolic array with three execution paths through the loop body

We present an example demonstrating the potential benefit of integrating static code regions with SPF in Listing 1. The example checks if positive or negative integers occur more frequently in the list x, and it contains a bug if x contains an equal number of positive and negative integers. The three-way branch on lines 5, 6 causes the total number of execution paths required to cover the *for* loop to be 3^{len}. However, this three-way branch can be combined into a multi-path region and represented as a disjunctive predicate. We present such predicates in SMT2 notation in Listing 2 assuming x to contain two symbolic integers named *x0* and *x1* (*len* equals 2). The updates to *sum* in the two loop iterations are captured by *sum0* and *sum1*. Using such predicates to represent the three-way branch on lines 5, 6 of Listing 1 allows us to have only one execution path through the loop body. Figure 1 shows a comparison of the number of execution paths explored to find

the bug on line 11 of Listing 1. The exponential speed-up from our predicates, representing a multi-path region, allows us to find the bug using just three test cases.

Unfortunately, as originally proposed, Veritesting would be unable to create a static region for this loop because it involves non-local control jumps (the calls to the `get` methods). This is not an impediment for compiled C code, as the C compiler will usually automatically inline the code for short methods such as `get`. However, Java has an *open world* assumption, and most methods are *dynamically dispatched*, meaning that the code to be run is not certain until resolved at runtime, so the compiler is unable to perform these optimizations.

In Java, programs often consist of many small methods that are dynamically dispatched, leading to poor performance for näive implementations of bounded static regions. Thus, to be successful, we must be able to inject the static regions associated with the calls into the dispatching region. We call such regions *higher order* as they require a region as an argument and can return a region that may need to be further interpreted. Given support for such regions, we can make analysis of programs such as 1 trivial for large loop depths. In our experiments, we demonstrate 100x speedups on several models (in general, the more paths contained within a program, the larger the speedup) over the unmodified Java SPF tool using this approach.

```
1  ; one variable per array entry
2  (declare-fun x0 () (_ BitVec 32))
3  (declare-fun x1 () (_ BitVec 32))
4  ; a variable to represent 'sum'
5  (declare-fun sum () (_ BitVec 32))
6  ; one 'sum' variable per loop iteration
7  (declare-fun sum0 () (_ BitVec 32))
8  (declare-fun sum1 () (_ BitVec 32))
9  ; unrolled lines 5, 6 in Listing 1
10 (assert
11   (or (and (= x0 #x00000000) (= sum0 #x00000000))
12     (or (and (bvsgt x0 #x00000000) (= sum0 #x00000001))
13         (and (bvslt x0 #x00000000) (= sum0 #xffffffff)))
     ))
14 ; second iteration of unrolling lines 5, 6
15 (assert
16   (or (and (= x1 #x10000000) (= sum1 #x10000000))
17     (or (and (bvsgt x1 #x10000000) (= sum1 #x10000001))
18         (and (bvslt x1 #x10000000) (= sum1 #xffffffff)))
     ))
19 ; merge function for 'sum' variable
20 (assert (= sum (bvadd sum0 sum1)))
21 ; branch on line 9 of Listing 1
22 (assert (bvslt sum #x00000000))
```

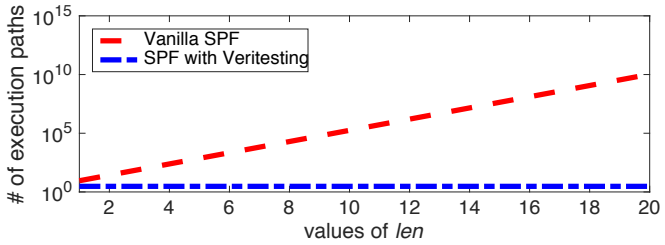
Listing 2: SMT2 representation of multi-path execution in Listing 1 using *len* = 2

II. Preliminaries

KEEP?

In this section, we describe the different technologies that are used to construct Java Ranger.

Fig. 1: Comparing number of execution paths from Listing 1 using vanilla SPF and SPF with static unrolling



A. SPF

Symbolic PathFinder (SPF) [13] combines symbolic execution with model checking and constraint solving for test case generation. In this tool, programs are executed on symbolic inputs representing multiple concrete inputs. Values of variables are represented as numeric constraints, generated from analysis of the code structure. These constraints are then solved to generate test inputs guaranteed to reach that part of code. Essentially SPF performs symbolic execution for Java programs at the bytecode level. Symbolic PathFinder uses the analysis engine of the NASA Ames JPF model checking tool (i.e. jpf-core) [18].

B. WALA and SSA Form

In order to find the static regions that we compile into disjunctive predicates, we use the open-source WALA library []. WALA can construct several different intermediate representations from Java bytecode including single-static assignment (SSA) form []. The SSA form is particularly attractive for constructing static regions as (for loopless code) there is a unique variable for each assignment, affording a straightforward translation into a predicate representing the region.

We chose WALA specifically over competing solutions such as Soot [] because it maintains the bytecode offset of statements and the stack locations of local variables used in the SSA form. This allows us to easily interface with the existing SPF code, by binding (through equalities) the symbolic or concrete values stored on the stack for inputs and also to store the variables corresponding to final values of computed variables back onto the stack.

III. Challenges

While the performance improvement demonstrated on the code in Listing 1 is impressive, it is perhaps not representative of most Java code. Java conventions encourage the use of indirection when accessing class fields using non-static *get* and *set* methods, as well as liberal use of exceptions. Unlike C compilers, which assume a "closed world" and often inline simple functions into the body of calling methods to improve performance, the Java compiler must assume an "open world" in which a class may be used in a new context, so inlining of non-final methods is unsafe.

Previous approaches to veritesting exit static code regions when indirect calls to functions or non-local jumps are made. In this section, we explore how the structure of Java programs reducees the performance of a näive veritesting approach.

A. Exit Points

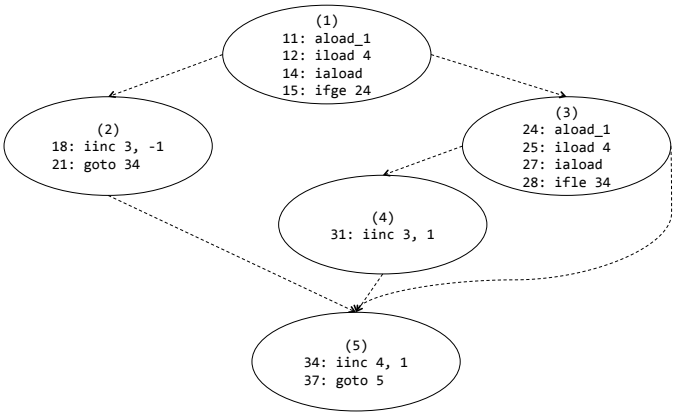
Should we re-run this analysis with WALA for consistency's sake? Integrating veritesting with SPF requires that we can represent a region of a Java program as a disjunctive formula with multiple exit points. Each exit point describes a possible distinct continuation of the current path after the static code region completes execution. Avgerinos et al. [17] defined exit points as unresolved jumps, function boundaries, and system calls. These exit points are nodes in a control-flow graph which represent non-local control flow, and therefore, need to be explored using plain dynamic symbolic execution. In the context of Java bytecode, we find such non-local control flow in five ways listed as follows: (1) return statements, (2) exceptions, (3) virtual function invocations (*invokevirtual*, *invokeinterface*), (4) reflection, (5) native calls.

The primary benefit of implementing veritesting comes from its conversion of branches into disjunctions in a multi-path region. But this benefit exists only when the number of different exit points from the disjunctive formula is less than the number of execution paths through the region in the first place. For example, all execution paths in the first three-way branch in Listing 1 joined together on line 7, causing the three-way branch to have a single exit point. Therefore, it is crucial for us to study the number of exit points for each of our statically-analyzed regions vs. the number of branches within the region.

These five types of exit points create the kind of non-local control flow which formed the frontier of the visible control-flow graph created as a result of *CFGReduce* step by Avgerinos et al. However, many of these exit points are used pervasively by Java developers. For example, the Visitor design pattern is used extensively by the ASM framework [19], Soot [20] and makes use of Java's dynamic dispatch mechanism. Running into exit points too often causes our statically-analyzed regions to be small and our performance gain from having fewer branches to be lost.

We investigated the occurrence of these exit points by creating a Soot-based static analysis of six large open-source projects written in Java. Software faults from these six projects are maintained in the Defects4J [21] repository. We used Soot to create a control-flow graph for every method body in every class file in these six projects. For each control-flow graph, we used nodes corresponding to conditional jump bytecodes as a starting point of our analysis. We measured the number of instructions encountered when traversing down each side of the branch until we get to the immediate post dominator [22] of our starting point. If there were no exit points encountered on any side of the branch, we considered this region as a pure multi-path region and calculated its size in bytecode instructions. Finally, we allowed up to five nested branches and calculated the number of bytecode instructions from the

Fig. 2: Control-flow graph with `else return;` added as line 7 of Listing 1



	#classes	if-return	if-invokevirtual	if-throw	region size
chart	679	8.44	27.47	4.33	13.59
closure	1339	7.35	22.1	9.5	11.66
lang	170	6.70	11.64	7.09	9.60
math	1104	18.27	56.61	9.56	27.06
mockito	382	6.02	12.51	8.05	13.57
time	209	7.79	13.10	7.08	8.10

TABLE I: Soot-based analysis for number of bytecode instructions between starting and exit points

earliest, as well as, the latest starting point in our control-flow graph traversal to an exit point.

As an example, we modify the three-way branch in Listing 1 by adding a `else return;` to get the snippet shown in Listing 3.

```
1 for (int i = 0; i < len; i++) {
2   if (x[i] < 0) sum += -1;
3   else if (x[i] > 0) sum += 1;
4   else return;
5 }
```

Listing 3: Listing 1 modified to have a return statement in the three-way branch

This results in bytecode that produces the control-flow graph shown in Figure 2. Nodes are numbered 1 through 6 with node (1) being the starting point and node (6) being the immediate post-dominator of node (1). Nodes contain bytecode instructions along with instruction offsets. The added *return* statement creates an exit point, which causes the three-way branch in Listing 3 to contain two exit points. Counting the number of instructions after the `ifge` instruction in node (1) through node (2) to node (6) gives us two instructions. The same count going through nodes (3) and (4) to (6) gives us 6 instructions, and through nodes (3) to (5) gives us 4 instructions. The presence of the *return* statement in node (5) prevents this region from being a pure multi-path region.

We report our results from a Soot-based static analysis in Tables I and II. Table I shows the average size and Table II shows the number of times each such count was reported. The *if-return*, *if-invokevirtual*, *if-throw* columns in Table I report the average number of instructions observed between

	if-return	if-invokevirtual	if-throw	region count
chart	1712	7760	521	6627
closure	3853	7466	138	9258
lang	3602	1589	539	2065
math	2219	5582	662	15375
mockito	372	572	15	574
time	1202	984	204	1421

TABLE II: Number of occurences in Soot-based static analysis

any *if* opcode-containing bytecode instruction and a *return*, *invokevirtual* or *invokeinterface*, *throw* opcode-containing bytecode instruction. These same columns in Table II report the number of times we observed an occurence of one of *return*, *invokevirtual*, *invokeinterface*, *throw* opcode-containing bytecode instructions before reaching the immediate post-dominator of the starting *if* node on any side of the branch. Tables I and II show that while we discover thousands of regions which do not contain any exit points, these regions are small. They also show that early *return* instructions are another often used construct in Java. We believe that creating multi-path regions for these no-exit-point cases alone would provide a significant performance boost to SPF. Table II shows *invokevirtual* or *invokeinterface* instructions are encountered far more often than *return* or *throw* instructions. This can be explained by the pervasive use of runtime polymorphism and exception handling by Java developers. Instead of using *invokevirtual* and *invokeinterface* instructions as exit points, if we can continue our predicate construction for multi-path regions beyond them, we would almost double the number of multi-path regions.

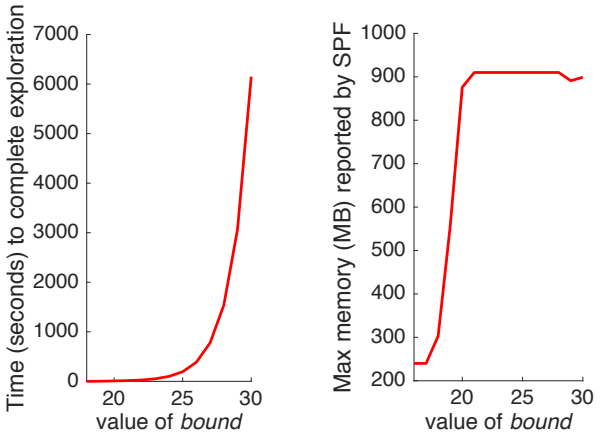
B. Shared Expressions

Veritesting causes regions of code to be executed using static symbolic execution. Symbolic formulas representing the static symbolic execution are then gathered at the exit points of the region and added to the path expression and symbolic store of dynamic symbolic execution. This causes large disjunctive formulas to be substituted and reused multiple times, necessitating the use of techniques like hash consing [23], or its variants such as maximally-shared graphs [24], or using expression caching [25]. To evaluate reuse of structurally equivalent expressions in SPF, consider the code shown in Listing 4.

```
1 // x is symbolic, bound is concrete
2 public void testSharing(int x, int bound) {
3   for(int i=0; i < bound; i++) x = x + x;
4   if ( x < -50 || x > 50) ...;
5   else ...
6 }
```

Listing 4: An example with an increasing formula size with every loop iteration

The function *testSharing* adds the value of x to itself in every loop iteration on line 3 of Listing 4. The number of loop iterations is controlled by a user-supplied value for *bound*. On line 4, the code branches on the value of the value of x. We symbolically executed the *testSharing* method with x set to be symbolic and *bound* set to be a concrete value. We set the minimum and maximum symbolic integer values to be -100 and 100 respectively. We increased the value of *bound* from 1 to 30 and recorded the time taken for complete path coverage. Figure 3 shows the trend seen in running time and memory usage for increasing values of *bound*. Figure 3a shows that the running time remains constant until the value for *bound* is 18, and then starts to rise exponentially. Figure 3b shows that at a value of 17 for *bound*, the memory usage starts to rise from 240 MB and has reached 910 MB when *bound* equals 21. We also observed that the number of expression objects undergoes a linear increase with the value of *bound*. These three observations lead us to the hypothesis that while SPF does share expression objects internally, the traversal of such expressions breaks the sharing and causes an increase in time and memory.



(a) Time required for covering paths in Listing 4

(b) Maximum memory usage of SPF for covering paths in Listing 4

Fig. 3: Time and memory usage of Listing 4 when increasing *bound*

C. Complex Expressions

During exploration, SPF creates conjunctions of expressions and adds them to its *PathCondition* to determine satisfiability of paths. These expressions are allowed to have a *Comparator* (a comparison operator such as !=) as the top-level operator; however, comparison and Boolean operators are not allowed in sub-expressions. Thus, the current set of SPF expressions is insufficiently expressive to represent the disjunctive formulas required for multi-path regions.

IV. Implementing Veritesting for Java

In order to implement veritesting for SPF, we are leveraging existing tools and framework features within SPF. The trickiest implementation aspects involve determining the bounds of static code regions over Java bytecodes and the mechanism for switching from "standard" symbolic execution using the SPF *SymbolicListener* class to one that can evaluate these multi-path code regions. We briefly describe our prototype implementations for these aspects.

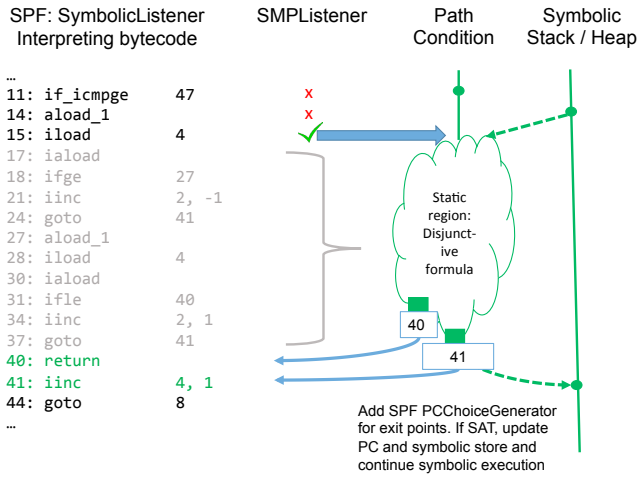
A. WALA-based analysis for veritesting

Veritesting requires static construction of predicates of a multi-path region which represent changes to the path expression of the dynamic symbolic executor. It also requires construction of a control-flow graph of method bodies from Java bytecode and finding exit points of the region, which in turn requires creation of a control-flow graph of the region. Implementing veritesting is made simpler by using a static single assignment (SSA) [26] representation of the multi-path region. Using an SSA form allows us to use the ϕ-expressions created by the SSA form and translate them into points at the end of the veritesting region where updates to system state along different paths in the region can be merged. MWW: Vaibhav please update to describe WALA WALA [] is a static analysis framework for Java programs that has both these features, with ExceptionalUnitGraph [27] and the Shimple IR [28]. For simple regions with only one exit point, like the one presented in Listing 1, we were able to use Soot to automate static construction of the predicate representing an update to the expression. For doing this, we used nodes with more than one successor as the starting point, found the immediate post-dominator of the starting point, and traversed the control-flow graph on all sides of such branches. During such a traversal, we constructed predicates representing the multi-path region, similar to the ones presented in Listing 2. As explained in Section III-A, including virtual function invocations in the construction of our predicates amplifies the benefits of veritesting even further. We plan to automate this inclusion in the future using Soot. Providing SPF with updates to be made to its symbolic store also requires Soot to maintain stack location information for variables. We plan to automate SPF's symbolic store updates using Soot in the future.

B. Integrating Veritesting with Symbolic PathFinder

Integration of veritesting requires changing Symbolic PathFinder so that it can use a Soot-based analysis. We present the sequence of actions SPF must take to implement veritesting in Figure 4. The bytecode used in Figure 4 was obtained by compiling source code shown in Listing 3 with its corresponding control-flow graph shown in Figure 2. This integration assumes our prior Soot-based analyis provides SPF with a table that maps instruction offsets (representing the start of a veritesting region) to a set containing (1) the multi-path region predicate to be added to the path expression as a conjunction, (2) symbolic store updates, (3) exit points, (4) the expression to branch on to one of the exit points. Using SPF's listener mechanism, we add a listener (named *SMPListener* in Figure 4) which listens for instructions that are starting points for a Symbolic Multi-Path (SMP) region. On finding such an instruction, *SMPListener* (1) updates the path expression, which may involve using the symbolic stack and/or heap, (2) updates the symbolic stack and heap, (3) creates a branch (using SPF's *PCChoiceGenerator*) to jump to one of the exit points (which are instructions at offsets 40 and 41 in Figure 4). SPF will then continue plain symbolic execution.

Fig. 4: Veritesting with Symbolic PathFinder



Thus, veritesting causes SPF to explore fewer branches. For example, SPF only explores a two-way branch in Figure 4.

C. Representation of Static Regions

MWW: - we should provide an AST of the constraint language MWW: - we should describe the different types of "holes" MWW: - we should describe how regions with multiple exit points are managed (currently we do not handle these regions, correct?

V. Experiment

We would like to measure the performance of Java Ranger against the baseline of single-path exploration using Java Symbolic PathFinder. This can be examined in several dimensions: the wall-clock time of the solving process, the number of paths explored. In addition, we would also like to gather metrics about the regions themselves, in order to better understand where static regions are effective.

Therefore, we investigate the following research questions:

- RQ1: How much do higher-order static regions (HOSR) improve the performance of symbolic execution?
- RQ2: How much do HOSRs reduce the number of paths explored?
- RQ3: How do HOSRs affect the number and expense of calls to the SMT solver?
- RQ4: How does the size of computed HOSRs affect the performance of the approach?

For each question, we examine three different configurations of Java Ranger: a version that creates simple regions (one branch) only (Java Ranger-SR), one that creates complex regions with multiple branches but no non-local jumps (Java Ranger-CR), and one that operates over higher-order complex regions containing non-local jumps (Java Ranger-CR+HO). This allows us to examine (at a coarse level) how each feature impacts the experimental results.

TABLE III: Performance of Java Ranger vs. Java Symbolic Pathfinder

Program	SPF	JR CR+HO	JR SR	JR CR
Program1	0.005	2.335	0.192	0.355
Program2	0.014	13.297	0.589	1.473

TABLE IV: Number of Paths for Java Rangervs. Java Symbolic Pathfinder

Program	SPF	JR CR+HO	JR SR	JR CR
Program1	XXX	YYY	ZZZ	AAA
Program2	XXX	YYY	ZZZ	AAA

A. Experimental Setup

Information here about benchmark models and machine configuration

The benchmarks should be a superset of at least one previous paper, and better yet, multiple papers.

VI. Results

In this section, we examine experimental results from the perspective of each research question.

A. Performance

We consider the performance of Java Ranger against Java Symbolic Pathfinder in Table III.

B. Number of Paths

We consider the number of paths explored by Java Ranger against Java Symbolic Pathfinder in Table IV.

C. Number of SMT Calls

We consider the number of SMT calls explored by Java Ranger against Java Symbolic Pathfinder in Table V. Note that this does not directly correspond to the number of paths, because static regions tend to make more of the variables symbolic, leading to larger numbers of solver calls per path.

D. Region Size

We consider the region size produced by each configuration in Table VI

TABLE V: Number and Aggregate Time of SMT Solver Calls for Java Ranger vs. Java Symbolic Pathfinder

Program	SPF	JR CR+HO	JR SR	JR CR
Program1	XXX (YY)	YYY (YY)	ZZZ (YY)	AAA (YY)
Program2	XXX (YY)	YYY (YY)	ZZZ (YY)	AAA (YY)

TABLE VI: Size of Regions for Java Ranger vs. Java Symbolic Pathfinder

Program	SPF	JR CR+HO	JR SR	JR CR
Program1	XXX	YYY	ZZZ	AAA
Program2	XXX	YYY	ZZZ	AAA

VII. Discussion

- Small regions cause performance problems, especially when they make previously concrete information symbolic. This can lead to many more solver calls, even when the number of paths is reduced.
- In some models, it is possible to reduce the number of paths to one. The static region approach essentially constructs a unrolled version of the program, similar to what tools like CBMC construct. This can only happen on relatively static models that do not have a lot of object construction leading to multiple dispatch paths. HOSRs are more flexible for these situations and allow specialization depending on dispatch type, which we believe will lead to better performance for highly-dynamic models.
- In general, the solver time does not rise dramatically for disjunctive paths. Since (in the limit) we reduce the number of paths exponentially by removing branches, we can perform relatively expensive analyses as preprocessing steps and at instantiation if we are able to instantiate a static region, and still end up with much better performance.
- Other lessons?

VIII. Related Work

The original idea for veritesting was presented by Avgerinos et al. [17]. They implemented veritesting on top of MAYHEM [29], a system for finding bugs at the X86 binary level which uses symbolic execution. Their implementation demonstrated dramatic performance improvements and allowed them to find more bugs, and have better node and path coverage. Veritesting has also been integrated with another binary level symbolic execution engine named `angr` [8]. Veritesting was added to `angr` with similar goals of statically and selectively merging paths to mitigate path explosion. However, path merging from veritesting integration with `angr` caused complex expressions to be introduced which overloaded the constraint solver. Using the Green [25] solver may alleviate such problems when implementing veritesting with SPF. Another technique named *MultiSE* for merging symbolic execution states incrementally was proposed by Sen et al. [30]. MultiSE computes a set of guarded symbolic expressions for every assignment and does not require identification of points where previously forked dynamic symbolic executors need to be merged. MultiSE complements predicate construction for multi-path regions beyond standard exit points (such as *invokevirtual*, *invokeinterface*, *return* statements). Combining both techniques, while a substantial implementation effort, has the potential to amplify the benefits from both techniques.

Finding which reflective method call is being used, or handling dynamic class loading are known problems for static analysis tools. TamiFlex [31] provides an answer that is sound with respect to a set of previously seen program runs. Integrating veritesting runs into similar problems, and using techniques from TamiFlex would allow static predicate construction beyond exit points caused by reflection or dynamic class loading.

IX. Conclusion

Future work:

Extensions:

1) Simplification of static regions at instantiation time using constant propagation, code specialization, etc.
2) Adding support for parallel Java programs with static regions
3) Adaptation of metric-based test-case generation for static regions: statement, branch, MCDC, observable metrics.
4) Parallelization of the analysis process (similar to Staats work in 2011)
5) Adding support for bypass of complex expressions: *symcrete* execution involving regions.
6) Integration with the Green constraint solver

Other improvements:

1) Simplification of the Symbolic PathFinder constraint mechanism
2) Interpolation-based path subsumption checks (c.f.: "More Effective Interpolations in Software Model Checking" - ASE 2017")

References

[1] J. C. King, "Symbolic execution and program testing," *Commun. ACM*, vol. 19, no. 7, pp. 385–394, 1976. [Online]. Available: http://doi.acm.org/10.1145/360248.360252

[2] L. A. Clarke, "A system to generate test data and symbolically execute programs," *IEEE Trans. Software Eng.*, vol. 2, no. 3, pp. 215–222, 1976. [Online]. Available: https://doi.org/10.1109/TSE.1976.233817

[3] P. Godefroid, N. Klarlund, and K. Sen, "Dart: Directed automated random testing," in *Proceedings of the 2005 ACM SIGPLAN Conference on Programming Language Design and Implementation*, ser. PLDI '05. New York, NY, USA: ACM, 2005, pp. 213–223. [Online]. Available: http://doi.acm.org/10.1145/1065010.1065036

[4] K. Sen, D. Marinov, and G. Agha, "Cute: A concolic unit testing engine for c," in *Proceedings of the 10th European Software Engineering Conference Held Jointly with 13th ACM SIGSOFT International Symposium on Foundations of Software Engineering*, ser. ESEC/FSE-13. New York, NY, USA: ACM, 2005, pp. 263–272. [Online]. Available: http://doi.acm.org/10.1145/1081706.1081750

[5] D. A. Ramos and D. R. Engler, "Practical, low-effort equivalence verification of real code," in *Proceedings of the 23rd International Conference on Computer Aided Verification*, ser. CAV'11. Berlin, Heidelberg: Springer-Verlag, 2011, pp. 669–685. [Online]. Available: http://dl.acm.org/citation.cfm?id=2032305.2032360

[6] V. Sharma, K. Hietala, and S. McCamant, "Finding Substitutable Binary Code By Synthesizing Adaptors," in *11th IEEE Conference on Software Testing, Validation and Verification (ICST)*, Apr. 2018.

[7] N. Stephens, J. Grosen, C. Salls, A. Dutcher, R. Wang, J. Corbetta, Y. Shoshitaishvili, C. Kruegel, and G. Vigna, "Driller: Augmenting fuzzing through selective symbolic execution." in *NDSS*, vol. 16, 2016, pp. 1–16.

[8] Y. Shoshitaishvili, R. Wang, C. Salls, N. Stephens, M. Polino, A. Dutcher, J. Grosen, S. Feng, C. Hauser, C. Kruegel, and G. Vigna, "Sok: (state of) the art of war: Offensive techniques in binary analysis," in *2016 IEEE Symposium on Security and Privacy (SP)*, May 2016, pp. 138–157.

[9] W. Sun, L. Xu, and S. Elbaum, "Improving the cost-effectiveness of symbolic testing techniques for transport protocol implementations under packet dynamics," in *Proceedings of the 26th ACM SIGSOFT International Symposium on Software Testing and Analysis*, ser. ISSTA 2017. New York, NY, USA: ACM, 2017, pp. 79–89. [Online]. Available: http://doi.acm.org/10.1145/3092703.3092706

[10] J. Burnim and K. Sen, "Heuristics for scalable dynamic test generation," in *23rd IEEE/ACM International Conference on Automated Software Engineering (ASE)*, 2008, pp. 443–446. [Online]. Available: https://doi.org/10.1109/ASE.2008.69

[11] C. Cadar, D. Dunbar, and D. R. Engler, "KLEE: unassisted and automatic generation of high-coverage tests for complex systems programs," in *8th USENIX Symposium on Operating Systems Design and Implementation (OSDI)s*, 2008, pp. 209–224. [Online]. Available: http://www.usenix.org/events/osdi08/tech/full_papers/cadar/cadar.pdf

[12] K. Luckow, M. Dimjašević, D. Giannakopoulou, F. Howar, M. Isberner, T. Kahsai, Z. Rakamarić, and V. Raman, "JDart: A dynamic symbolic analysis framework," in *Proceedings of the 22nd International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS)*, ser. Lecture Notes in Computer Science, M. Chechik and J.-F. Raskin, Eds., vol. 9636. Springer, 2016, pp. 442–459.

[13] C. S. Păsăreanu, W. Visser, D. Bushnell, J. Geldenhuys, P. Mehlitz, and N. Rungta, "Symbolic pathfinder: integrating symbolic execution with model checking for java bytecode analysis," *Automated Software Engineering*, vol. 20, no. 3, pp. 391–425, Sep 2013. [Online]. Available: https://doi.org/10.1007/s10515-013-0122-2

[14] V. Chipounov, V. Kuznetsov, and G. Candea, "The S2E platform: Design, implementation, and applications," *ACM Trans. Comput. Syst.*, vol. 30, no. 1, pp. 2:1–2:49, 2012. [Online]. Available: http://doi.acm.org/10.1145/2110356.2110358

[15] D. Babic, L. Martignoni, S. McCamant, and D. Song, "Statically-directed dynamic automated test generation," in *Proceedings of the 20th International Symposium on Software Testing and Analysis (ISSTA)*, 2011, pp. 12–22. [Online]. Available: http://doi.acm.org/10.1145/2001420.2001423

[16] V. Kuznetsov, J. Kinder, S. Bucur, and G. Candea, "Efficient state merging in symbolic execution," in *Proceedings of the 33rd ACM SIGPLAN Conference on Programming Language Design and Implementation*, ser. PLDI '12. New York, NY, USA: ACM, 2012, pp. 193–204.

[17] T. Avgerinos, A. Rebert, S. K. Cha, and D. Brumley, "Enhancing symbolic execution with veritesting," in *Proceedings of the 36th International Conference on Software Engineering*, ser. ICSE 2014. New York, NY, USA: ACM, 2014, pp. 1083–1094. [Online]. Available: http://doi.acm.org/10.1145/2568225.2568293

[18] W. Visser, K. Havelund, G. Brat, S. Park, and F. Lerda, "Model checking programs," *Automated Software Engineering*, vol. 10, no. 2, pp. 203–232, Apr 2003. [Online]. Available: https://doi.org/10.1023/A:1022920129859

[19] E. Bruneton, R. Lenglet, and T. Coupaye, "Asm: a code manipulation tool to implement adaptable systems," *Adaptable and extensible component systems*, vol. 30, no. 19, 2002.

[20] R. Vallée-Rai, P. Co, E. Gagnon, L. Hendren, P. Lam, and V. Sundaresan, "Soot - a java bytecode optimization framework," in *Proceedings of the 1999 Conference of the Centre for Advanced Studies on Collaborative Research*, ser. CASCON '99. IBM Press, 1999, pp. 13–24.

[21] R. Just, D. Jalali, L. Inozemtseva, M. D. Ernst, R. Holmes, and G. Fraser, "Are mutants a valid substitute for real faults in software testing?" in *Proceedings of the 22Nd ACM SIGSOFT International Symposium on Foundations of Software Engineering*, ser. FSE 2014. New York, NY, USA: ACM, 2014, pp. 654–665. [Online]. Available: http://doi.acm.org/10.1145/2635868.2635929

[22] A. V. Aho, R. Sethi, and J. D. Ullman, *Compilers: principles, techniques, and tools*. Addison-wesley Reading, 2007, vol. 2.

[23] E. Goto, "Monocopy and associative algorithms in an extended lisp," TR 74-03, University of Tokyo, Tech. Rep., 1974.

[24] D. Babic, "Exploiting structure for scalable software verification," Ph.D. dissertation, PhD thesis, University of British Columbia, Vancouver, Canada, 2008.

[25] W. Visser, J. Geldenhuys, and M. B. Dwyer, "Green: Reducing, reusing and recycling constraints in program analysis," in *Proceedings of the ACM SIGSOFT 20th International Symposium on the Foundations of Software Engineering*, ser. FSE '12. New York, NY, USA: ACM, 2012, pp. 58:1–58:11. [Online]. Available: http://doi.acm.org/10.1145/2393596.2393665

[26] G. Bilardi and K. Pingali, "The static single assignment form and its computation," Tech. Rep., 1999.

[27] M. U. Sable Research Group, "Exceptional Unit Graph (Soot API)," https://www.sable.mcgill.ca/soot/doc/soot/toolkits/graph/ExceptionalUnitGraph.html, 2017.

[28] ——, "A Brief Overview Of Shimple," https://github.com/Sable/soot/wiki/A-brief-overview-of-Shimple, 2017.

[29] S. K. Cha, T. Avgerinos, A. Rebert, and D. Brumley, "Unleashing mayhem on binary code," in *2012 IEEE Symposium on Security and Privacy*, May 2012, pp. 380–394.

[30] K. Sen, G. Necula, L. Gong, and W. Choi, "Multise: Multi-path symbolic execution using value summaries," in *Proceedings of the 2015 10th Joint Meeting on Foundations of Software Engineering*, ser. ESEC/FSE 2015. New York, NY, USA: ACM, 2015, pp. 842–853. [Online]. Available: http://doi.acm.org/10.1145/2786805.2786830

[31] E. Bodden, A. Sewe, J. Sinschek, H. Oueslati, and M. Mezini, "Taming reflection: Aiding static analysis in the presence of reflection and custom class loaders," in *Proceedings of the 33rd International Conference on Software Engineering*, ser. ICSE '11. New York, NY, USA: ACM, 2011, pp. 241–250.