A Survey of
Symbolic Executions Techniques

Hallet Adrien Sens Loan

October 14, 2018

Abstract

1 Introduction

1.1 A definition

The first occurences of symbolic execution described the then-new method as a middle ground [9] between the two most-used method of its time. On one hand, program testing (*e.g.: unit testing*) can not always detect a fault in a program and producing a correct test sample and proving that it indeed is correct is not that easy. On the other hand, program proving can indeed ensure that a program is correct from its entry point to the result but it heavily relys on the proof definitions by the programmer and the formal definition of the problem.
Nowadays, symbolic execution is both described as (part of) the core of many modern techniques to software testing [13] and an effective way to create tests suites with extensive coverage. [2]

1.2 The concept

The idea behind symbolic execution is to test an algorithm with *symbolic values* rather than concrete values. So instead of using unit testing where a variable is set to a (usually random) value, the symbolic execution maintains a formula that contains all the possible values for the code to reach a particular point in the program. This formula is updated every time the program reaches a branching point. In figure 1, we show an example from [14] of a symbolic execution. Notice how it produces constraints over the variables to explore the algorithm's branching tree.

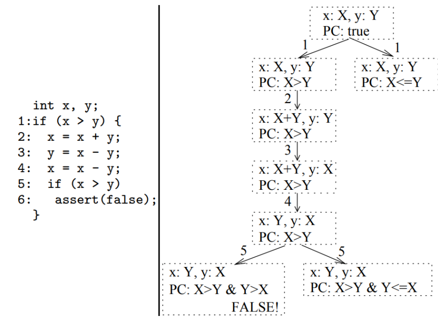

Figure 1: Swapping two integers and its symbolic execution tree

2 History

We find the first papers on symbolic execution around 1975 [9]. Early methods proposed simple structures to hold conditions with a SAT-solver, the support of simple data types and were focused on algorithm testing (instead of large programs). While papers continue to grow on the subject, it really is around 2005 that symbolic execution really becomes a large thing with more and more frameworks and tools for software verification. The last ten years have seen more papers on the subject[1] (around 200.000) than the three decades following its introduction (half that much). Among the supporters of the method, we can identify two clear research poles, China [3,16,17] and the United States of America, with an heavy participation of Microsoft [6] which was pooling heavy resources in software and OS reliability, and the NASA [12,14,15].

3 Method

3.1 Useful concepts

Algorithms can be modeled as graphs where nodes are basic blocks (*i.e.: a part, one or multiple instructions with a single entry and exit point*) and edges are the branches (issued from conditional statements). Def-use pairs use the same concept, although they base their graph on the *definition* and *usage* of a variable. With the *branches* and *def-use pairs*, we can model an algorithm's behavior to follow the values of its variables and determine the *execution path*.

[1] Data from Google Scholar

3.2 Basis

Symbolic execution works on those concepts by updating an internal list of symbols. The execution generates a new symbol for each introduced variable in an algorithm [9]. The symbolic execution runs over the algorithm's statements and builds the symbolic values when it encounters a branching point. The symbolically executed algorithm creates a *state* [14] containing the symbolics values, a counter identifying the next line to be executed and a *path condition*. This path condition is a simple boolean formula over the symbols, it creates a constraint for the algorithm to reach the current state of the program (this also allows to check for unreachable paths in programs [1]). The path condition allows to recreate the execution up to its state. The states are stored in a *symbolic execution tree* with the states as nodes and the transitions as edges (see figure 1).

3.3 Problems

3.3.1 State-Space Explosion

Symbolic Execution cannot be that perfect and hosts its bundle of problems that reduce either the confidence in or the performances of the concept. We have seen symbolic execution tree in 3.2. Small algorithms can use such methods but actual programs need to be tested in *integration*. In large environments, the tree's branching factor will produce too many nodes (*state-space explosion*) for the performances to stay relevant, sometimes creating infinite loops in the graph [11]. Reducing the state space or pruning them is not enough. To improve the performances, we can depth-first-search the graph but it does not prevent infinite loops until we add a max depth (as KLEE or EXE do). Pruning heuristics (*e.g.:def-use pairs distance*) can be used to reduce the tree's branching factor, we may randomly select a path (emphi.e: the path condition), weighting the shallowest nodes to avoid deadloops. The random technique is exploited by a lot of *fuzzing techniques* [3] which uses the injection of random values to detect program's faults. Another solution lies in the *concolic execution* (see 4.1).

3.3.2 Modeling the memory

Another technical difficulty lies in the *memory model*. As said before, a symbol α represents a variable `a` with a value. In the programming world, it means there is a pointer `a*` that stores the address to an allocated memory block. The initial approach [9] proposed a *fully symbolic memory*. The

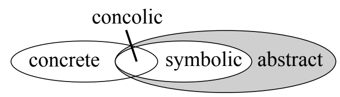

Figure 2: Concrete and abstract execution machine models

symbols are stored in plain states holding their path condition with either a duplication of the states depending on their memory status called *state forking* where every possible execution is forked from the main branch (*e.g.: accessing an array of size 10 with a variable i that depends on the context will fork 10 states from the main memory state, each one with i from 0 to 9*). Fully symbolic memory is therfore slowed by an increased *state-space explosion*. To reduce memory usage, some tools may represent an artificial memory space [10], which reduces the size of the general space by allowing more memory to each path condition, although doing so will void some execution paths, leaving potential bugs out of the symbolic execution. An italian approach [5] proposes the use of a *symbolic* adress which holds the condition to which address the variable points to, which drastically reduces the amount of memory states.

Instead of fully modeling the memory, there also is the *abstract symbol table* [17] which records tuples [variable, address, symbolic value]. This method has the advantage of supporting complex data types (some fully memory models cannot express structures) and memory aliasing (instead of creating a new variable copied from another in older methods).

4 Variants

4.1 Concolic execution

The name "concolic" is a portmanteau of the words "concrete" and "symbolic", the idea of this testing method is to mix symbolic execution alongside concrete ones.

Concolic execution approaches

This technique concept was first introduced on 2005 [7] (more details on section 5.1). Since then the idea was further extended and combined with other testing techniques.

However, the general principle has been explored with different angles.

4.1.1 Dynamic Symbolic Execution

Dynamic Symbolic Execution (DSE) also kwown as *dynamic test generation* [7] is a popular approach of concolic execution. Its main feature is to have the concrete execution drive the symbolic execution.

This method need to add a new store in order to save the concrete execution information.
We first choose an arbitrary value as input for our parameters. Then it executes the program concretely and symbolically at the same time updating both stores and the path constraints. Whenever the concrete execution takes a branch, the symbolic execution is directed toward the same branch and the constraints extracted from the branch condition are added to the current set of path constraints.
In order to explore different paths, the path conditions given by one or more branches can be negated and the solver invoked to find a satisfying assignment for the new constraints.
We can repeat this process as many time as we want to achieve the desired coverage.

Notice that it exists different strategies on the choice of the branch to negate, this crucial heuristic choice depends of the tool.

Downside : Imperfect symbolic execution

False Negative Missed path. For example, when another function from the one tested is not symbolically tracked but its result is needed to explore a particular path.

Path Divergence In some situations the engine can't guess that no input can provoke an error. In other word, whenever an actual execution path does not match the program path predicted by symbolic execution for a given input vector. For example, assert on a negative value of an absolute value due to the untracked side effect of the `abs()` function.

According to [8] they calculated a divergence rates of over 60 %

4.1.2 Selective Symbolic Execution

TODO [Loan] : Lire article [4] section 2

5 Tools and languages

Many tools exist for symbolic execution, Wikipedia mention 22 of them. Another source claiming to "curate a list of awesome symbolic execution resources including tools" mention 35 different tools spread over 10 different languages.

5.1 *DART* : Directed Automated Random Testing

DART is presented as a tool for automatically testing software using concolic testing method. It was introduced in 2005 making it the first the first tool to be created using concolic techniques and more specifically dynamic symbolic execution techniques (see section 4.1.1).

5.1.1 Methodology

DART combines three main techniques [7] in order to automate the process of testing for a particular software :

1. An automated extraction of the interface of a program with its external environment using static source-code parsing
2. An automatic generation of a test driver for this interface that performs random testing to simulate the most general environment the program can operate in
3. A dynamic analysis of how program behaves under random testing and automatic generation of new test inputs to direct systemically the execution along alternatives program paths

DART chooses the *depth-first* strategy whenever it has to negate a branch.

```
Function foo(int x, int y):
    if x != y then
        if 2 * x == x + 10 then
            ERROR;
        end
    end
    return SUCCESS;
```

5.1.2 Example

Let consider the following program :

This function is defective as it may lead to an error statement for some value of x and y.

DART start by guessing values for both x and y for instance 269167349 and 889801541. With this values the function return successfully, during the execution two predicates were formed created by the `if` statements, in our case the path constraint at the end is : $\langle x_0 \neq y_0, 2 \times x_0 \neq x_0 + 10\rangle$ with x_0 and y_0 both beings *symbolic variables*.

While we maintain this predicates, all path will lead to the same end. So in order to force the program through a potential different outcome we change one of the predicate and look at the result. If we negate the last predicate we have the following path constraint : $\langle x_0 \neq y_0, 2 \times x_0 = x_0 + 10\rangle$ in which $x_0 = 10$ and $y_0 = 889801541$ is a solution. Using this values as inputs the program end up into the `ERROR` as wanted.

5.1.3 Key strength/originality

The main strength of DART is that testing can be performed completely automatically on any program that compiles – there is no need to write any test driver or harness code.

During testing, DART detects standard errors such as program crashes, assertion violations, and non-termination.

DART provides an attractive alternative approach to static analyzers, because it is based on high-precision dynamic analysis instead, while being fully automated as static analysis. The main advantage of DART over static analysis is that every execution leading to an error that is found by DART is guaranteed to be sound. Two areas where we expect DART to compete especially well against static analyzers are the detection of interprocedural bugs and of bugs that arise through the use of library functions.

DART is overall complementary to static analysis since it has its own limitations, namely the computational expense of running tests and the sometimes limited effectiveness of dynamic test generation to improve over random testing.

6 Conclusions

References

[1] Romain Aissat, Frédéric Voisin, and Burkhart Wolff. Infeasible paths elimination by symbolic execution techniques. In Jasmin Christian Blanchette and Stephan Merz, editors, *Interactive Theorem Proving*, pages 36–51, Cham, 2016. Springer International Publishing.

[2] Cristian Cadar and Koushik Sen. Symbolic execution for software testing: Three decades later. 56:82–90, 02 2013.

[3] Chen Chen, Baojiang Cui, Jinxin Ma, Runpu Wu, Jianchao Guo, and Wenqian Liu. A systematic review of fuzzing techniques. *Computers and Security*, 75:118 – 137, 2018.

[4] Vitaly Chipounov, Volodymyr Kuznetsov, and George Candea. The s2e platform: Design, implementation, and applications. *ACM Transactions on Computer Systems (TOCS)*, 30(1):2, 2012.

[5] Emilio Coppa, Daniele Cono D'Elia, and Camil Demetrescu. Rethinking pointer reasoning in symbolic execution. In *Proceedings of the 32Nd IEEE/ACM International Conference on Automated Software Engineering*, ASE 2017, pages 613–618, Piscataway, NJ, USA, 2017. IEEE Press.

[6] Manuel Costa, Miguel Castro, Lidong Zhou, Lintao Zhang, and Marcus Peinado. Bouncer: Securing software by blocking bad input. pages 117–130, Stevenson, Washington, USA, October 2007. Association for Computing Machinery, Inc.

[7] Patrice Godefroid, Nils Klarlund, and Koushik Sen. Dart: directed automated random testing. In *ACM Sigplan Notices*, volume 40, pages 213–223. ACM, 2005.

[8] Patrice Godefroid, Michael Y. Levin, and David A. Molnar. Automated whitebox fuzz testing. In *NDSS*, 2008.

[9] J.C. King. A new approach to program testing. 10:228–233, 06 1975.

[10] M. Papadakis and N. Malevris. Automatic mutation test case generation via dynamic symbolic execution. In *2010 IEEE 21st International Symposium on Software Reliability Engineering*, pages 121–130, Nov 2010.

[11] E. J. Schwartz, T. Avgerinos, and D. Brumley. All you ever wanted to know about dynamic taint analysis and forward symbolic execution (but might have been afraid to ask). In *2010 IEEE Symposium on Security and Privacy*, pages 317–331, May 2010.

[12] S. Shen, S. Ramesh, S. Shinde, A. Roychoudhury, and P. Saxena. Neuro-Symbolic Execution: The Feasibility of an Inductive Approach to Symbolic Execution. *ArXiv e-prints*, July 2018.

[13] David Trabish, Andrea Mattavelli, Noam Rinetzky, and Cristian Cadar. Chopped symbolic execution. In *International Conference on Software Engineering (ICSE 2018)*, 5 2018.

[14] Willem Visser, Corina S. Păsăreanu, and Sarfraz Khurshid. Test input generation with java pathfinder. In *Proceedings of the 2004 ACM SIGSOFT International Symposium on Software Testing and Analysis*, ISSTA '04, pages 97–107, New York, NY, USA, 2004. ACM.

[15] Guowei Yang, Suzette Person, Neha Rungta, and Sarfraz Khurshid. Directed incremental symbolic execution. 2014.

[16] Rui Zhang and Cynthia Sturton. A recursive strategy for symbolic execution to find exploits in hardware designs. pages 1–9, 06 2018.

[17] Dai Ziying, Xiaoguang Mao, Ma Xiaodong, and Wang Rui. A memory model for symbolic execution, 01 2010.