



Pinaka: Symbolic Execution Meets Incremental Solving (Competition Contribution)

Eti Chaudhary and Saurabh Joshi

Department of Computer Science and Engineering,
Indian Institute of Technology Hyderabad, Kandi, India
`{cs17mtech11029, sbjoshi}@iith.ac.in`

Abstract. Many modern-day solvers offer functionality for incremental SAT solving, which preserves the state of the solver across invocations. This is beneficial when multiple, closely related SAT queries need to be fed to the solver. Pinaka is a symbolic execution engine which makes aggressive use of incremental SAT solving coupled with eager state infeasibility checks. It is built on top of the CProver/Symex framework. Pinaka supports both Breadth First Search and Depth First Search as state exploration strategies along with partial and full incremental modes. For SVCOMP 2019, Pinaka is configured to use partial incremental mode with Depth First Search strategy.

Keywords: Symbolic execution · Incremental solving · Software bug detection

1 Verification Technique

Pinaka extends symbolic execution with incremental solving coupled with eager infeasibility checks. A pure symbolic execution [6] engine builds a logical formula representing a potential execution path using symbolic values which may then be passed on to theorem-provers/solvers. An UNSAT outcome from the solver implies that the verification condition will not be violated along that path, whereas a SAT outcome provides a scenario leading to failure of an assertion during an execution along that path. The number of paths in a program blow-up exponentially as the number of branches increases. Pinaka, being a single-path symbolic execution engine, never merges two paths (i.e., diamonds). It employs eager infeasibility checks to avoid unnecessary exploration. Rather than making queries to the solver only when a path encounters an assertion, a query is made everytime a branch is encountered to check its feasibility. Infeasible branches are not explored further. These eager checks help Pinaka tremendously in reducing its search efforts. Pinaka is further powered by incremental solving [5] offered by

E. Chaudhary—Jury Member.

many state-of-the-art solvers such as MiniSAT [3]. Incremental Solving greatly benefits our technique by reducing the overhead encountered due to eager infeasibility checks. Pinaka has Depth First Search (DFS) and Breadth First Search (BFS) as its search strategies. It offers two different modes of operation: Partial Incremental (PI) Mode and Full Incremental (FI) Mode.

In PI mode, a single solver instance is maintained along a search path. In the event that a branch is encountered, only a partial path is encoded from the current point to the previous point from which a query was made along the current path. For example, in Fig. 1, a query would be made at b_1 . If both s_1 and s_2 are feasible, s_2 is put in a queue and the current solver instance is used to further explore the path starting from s_1 . When b_2 is encountered, only the path from s_1 to b_2 is encoded and added in the current solver instance before making a query. If both the branches at b_2 are infeasible, a new solver instance is created and a path from the initial state to another symbolic state (e.g., s_2) in the queue is encoded and the path along that symbolic state is explored further. Essentially, a new solver instance is created every-time a backtrack happens. Using BFS in PI mode is very memory consuming because for every symbolic state in the queue, a corresponding solver instance is retained. Running Pinaka with this combination is not recommended.

In FI mode, a single solver instance is retained throughout. In Fig. 1, if $b_1 \rightarrow s_1$ is a feasible branch, a new activation variable $a_{b_1 s_1}$ is created. Let $\phi_{b_1 b_2}$ be the encoding of the path from b_1 to b_2 . When b_2 is encountered, $a_{b_1 s_1} \Rightarrow \phi_{b_1 b_2}$ is added in the solver, and $a_{b_1 s_1}$ is added as an assumption to enforce the path. Since the underlying SAT solvers integrated with Pinaka do not allow *popping of a stack*, upon backtrack, $\neg a_{b_1 s_1}$ is set as an assumption to disable the constraints generated by this fragment of the path.

FI mode is beneficial when the input program does not have too many paths. Otherwise, the solver becomes quite slow over time with a large memory footprint. For a large program with too many paths, the benefit of a lower memory footprint and speed of PI mode outweighs its overhead of instantiating a new solver instance on every backtrack.

Loops are handled just like branches. Consider the program fragment given in Fig. 2. Assume that along some path where $(x_1 = 1) \wedge (y_3 = -1)$ the loop is encountered. Branch $x_1 \geq 3$ is infeasible along this path and will not be explored. Since $x_1 < 3$ is feasible, it is explored further by unrolling this iteration on-the-fly. Therefore, the path will further add $y_3 < 0$. Since it is feasible, $x_2 = x_1 + 1$ is added and feasibility of $x_2 < 3$ will be checked. After one more unrolling $x_3 < 3$

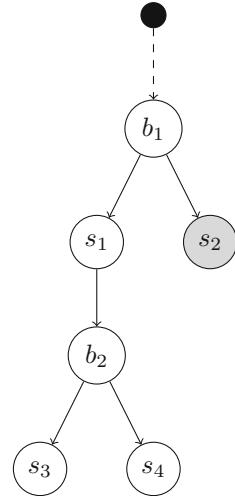


Fig. 1. Branching state in a program graph

```
while(x<3)
{
    if (y < 0 )
    { x=x+1; }
}
```

Fig. 2. Handling loops

will be found infeasible, thus guaranteeing termination of the loop along this path. Note that, along a path having $y_3 = 2$, the loop will be non-terminating for that path. In this case, Pinaka may not terminate. Function calls, including recursion, are handled in a similar fashion by inlining a call on-the-fly. Therefore, even though Pinaka provides an option of `--unwind NUM` to specify an unwinding limit, it does not mandate that a loop unwinding limit is specified. If a user-given unwinding limit is not sufficient to reach an assertion violation, it declares the program as *safe*, which may be unsound. To ensure soundness, we run it without any loop unwinding limit. For *unsafe* programs, upon encountering the first assertion violation, Pinaka terminates and reports a failure. For *safe* programs, however, Pinaka terminates only if all the paths of the program are terminating.

2 Architecture

Pinaka 0.1 is built upon the CProver [2]+ Symex [10] framework¹. Taking a *C* program as input, it makes use of CProver framework APIs to convert the input *C* program to a GOTO program. CProver APIs further come into play for pre-processing of GOTO-programs, witness generation, transformation passes such as setting the rounding-mode for floating-point operations, handling complex data types, etc. Pinaka implements PI and FI mode and eager infeasibility checks along with BFS/DFS exploration. Apart from DFS, none of those features are present in Symex [10]. Additionally, we make use of our forked version of the Symex repository in which we fix many bugs, especially for handling recursive procedures and ternary operators. As of now Pinaka only supports some MiniSAT-like solvers (i.e., Glucose [4], MapleSAT [7]) and not SMT solvers. Once a program has been verified, a verification successful/failed outcome is generated along with the appropriate witness (Fig. 3).

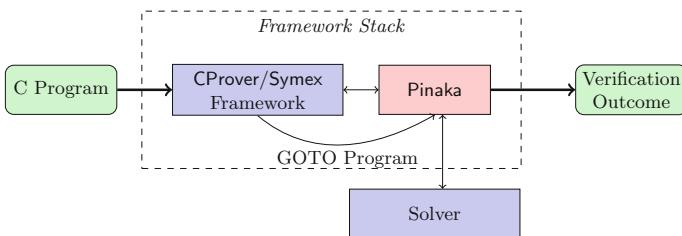


Fig. 3. Architectural overview of Pinaka

¹ Note the specific Symex version used on which Pinaka is built.

3 Strengths and Weaknesses

Modulo the soundness of the CProver/Symex framework and the back-end solver, Pinaka’s technique is sound. In addition, if CProver/Symex do not have any over-approximation in modeling the *C* constructs, then a bug reported by Pinaka would indeed be a bug.

As explained in Sect. 1, Pinaka can potentially be non-terminating if for some input value there is a non-terminating path. However, termination of verification process guarantees the termination of the underlying program (modulo the approximations introduced in modeling by CProver/Symex) *if* the program is declared *safe* by Pinaka. A notable strength of Pinaka lies in its speed. A majority of Pinaka’s verification outcomes were obtained under a mere limit of 10 s. A clear display of the same can be seen in the ReachSafety-Floats category, where Pinaka came in second with 1500 s CPU time [9], as compared to other tools that share a similar score but require 4 to 10 times more CPU time.

One major weakness of the current version Pinaka is a lack of techniques for loop invariants. Even with eager infeasibility checks and incremental solving, there is still a need for more loop-directed abstraction based approaches. Furthermore, support for handling multi-dimensional arrays is still lacking.

4 Tool Setup and Configuration

Pinaka 0.1 is available for download at <https://github.com/sbjoshi/Pinaka>. The repository contains a description of Pinaka’s working along with all the necessary configuration files required to run Pinaka SVCOMP style. All the instructions are listed in a stepwise manner in the README.md file. Although Pinaka is built on top of the CProver/Symex framework, the binary itself is sufficient and the tool does not require any additional pre-requisites. Pinaka has been tested on Ubuntu 18.04.

Pinaka 0.1 submitted for SVCOMP 2019 runs DFS in PI Mode as for SVCOMP benchmarks we found this combination the best. No loop unwinding limit was specified to retain soundness. For SVCOMP’19 [1,9] it uses Glucose-Syrup (Glucose-4.1) [4] as its solver back-end. Tool’s default search strategy, i.e., DFS may be overridden by providing `--bfs` option. Similarly, a default of FI mode may also be overridden by providing `--partial-incremental` option. Other additional options may be explored from the `--help` menu. The set of global parameters passed to the tool are: (1) `--graphml-witness`: to specify the witness file to be generated, (2) `--propertyfile`: to specify the property file, (3) `--32/--64`: to define the architecture to be used. Pinaka 0.1 participated in all ReachSafety subcategories *except* ReachSafety-Sequentialized, and also participated in NoOverflows and Termination meta-categories in SVCOMP 2019.

5 Software Project and Contributors

Pinaka is a result of very heavy code rewriting and refactoring of VerifOx [8] (developed by Saurabh Joshi) with a lot of feature additions and bug fixes. **Pinaka** is developed at Indian Institute of Technology, Hyderabad, India. It is available at <https://github.com/sbjoshi/Pinaka> under BSD License. The authors acknowledge the financial support from DST, India under SERB ECR 2017 grant.

References

1. Beyer, D.: Automatic verification of C and Java programs: SV-COMP 2019. In: Beyer, D., et al. (eds.) TACAS 2019, Part 3. LNCS, vol. 11429, pp. 133–155. Springer, Cham (2019)
2. CPROVER homepage. <http://www.cprover.org>. Accessed 10 Feb 2019
3. Een, N., Sörensson, N.: Minisat v2. 0 (beta). Solver description, SAT race 2006 (2006)
4. Glucose's homepage. <http://www.labri.fr/perso/lsimon/glucose/>. Accessed 10 Feb 2019
5. Hooker, J.N.: Solving the incremental satisfiability problem. *J. Logic Program.* **15**(1–2), 177–186 (1993)
6. King, J.C.: Symbolic execution and program testing. *Commun. ACM* **19**(7), 385–394 (1976)
7. Maplesat homepage. <https://sites.google.com/a/gsd.uwaterloo.ca/maplesat/>. Accessed 10 Feb 2019
8. Mukherjee, R., Joshi, S., Griesmayer, A., Kroening, D., Melham, T.: Equivalence checking of a floating-point unit against a high-level c model. In: FM (2016)
9. SVCOMP 2019 results. <https://sv-comp.sosy-lab.org/2019/results/results-verified/>. Accessed 4 Jan 2019
10. Symex repository. <https://github.com/diffblue/symex/tree/9b5a72cf992d29a905441f9dfa6802379546e1b7>. Accessed 10 Feb 2019

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

