Enhancing Symbolic Execution by Machine Learning Based Solver Selection

Sheng-Han Wen*, Wei-Loon Mow*, Wei-Ning Chen*, Chien-Yuan Wang* and Hsu-Chun Hsiao*
*Department of Computer Science and Information Engineering, National Taiwan University, Taiwan
†Research Center for IT Innovation, Academia Sinica, Taiwan
{r05922009, b03902110, b01902093, b04902084, hchsiao}@csie.ntu.edu.tw

Abstract—Constraint solving creates a serious performance bottleneck in symbolic execution. Examining a plethora of SMT solvers with diverse capabilities, we address the following research questions: How can the performance of symbolic execution improve if it can pick a priori the best solver for a given path constraint? How can such a prediction oracle be practically implemented? In this work, we first define the solver selection problem in symbolic execution and its evaluation metrics, and perform a preliminary study to gauge potential performance improvement through solver selection. We then present the design and implementation of Path Constraint Classifier (PCC), a machine learning based meta-solver that aims to reduce overall constraint solving latency by dynamically selecting a solver per query. Using machine learning seems straightforward, yet surprisingly underexplored; one main technical challenge is how to avoid excessive overhead introduced by feature extraction. We address this challenge by taking advantage of the structural characteristics of symbolic execution. Our experiments confirm that the overall solver time can be reduced by 10.3% in the KLEE dataset and 46% in the benchmark dataset, while the solver prediction procedure only accounts for 2% to 10% of overall solving time.

I. INTRODUCTION

Symbolic execution is an automated program analysis technique for software testing and vulnerability discovery. Symbolic execution systematically explores possible program execution paths, represents each path as a set of constraints, and sends the constraints to a solver for satisfiability checks and input generation.

Despite being a key enabler of symbolic execution, constraint solving is a computationally intractable problem in theory and remains a serious performance bottleneck in practice. Palikareva and Cadar [27] showed that solvers can use up to 99% of total time in symbolic execution. The performance may further deteriorate when advanced symbolic execution techniques, such as state merging, increase the constraint complexity. Kuznetsov et al. [21] showed that merging multiple states may end up increasing the overall solving time as the size of the symbolic path constraints increases.

Because state-of-the-art SMT solvers address intractability via various optimizations, they have different capabilities and solving times. No existing solver can consistently outperform the others, as shown in our preliminary study in Section III-A.

To take advantage of multiple solvers with diverse capabilities, several symbolic execution engines support multiple solvers in parallel. For example, KLEE [10] and JDART [25] embed the pySMT framework [15] and the JConstraints interface, respectively, as a unified interface for solvers. In addition, because solver selection is a type of algorithm selection problem, several portfolio-solving techniques have been applied to select solvers [19], [34]. However, because the time spent on constraint processing and prediction was not included in their evaluation, it remains unclear how effective these solutions are when applied to improve the overall performance of symbolic execution.

In this work, we define the solver selection problem in symbolic execution: Given a set of solvers and a sequence of path constraints generated by a symbolic execution engine, we would like to have a prediction oracle that selects a solver for each path constraint, such that the overall solving and prediction time is minimized. An interesting difference between the generic solver selection problem and the one in symbolic execution is the path constraints collected by the symbolic execution engine have high similarity with each other. This is because when the symbolic execution engine explores to a certain state with path constraint \( \{c_1, c_2, \ldots, c_n\} \), the engine can choose to explore the next state with constraint \( c_{n+1} \) or its negation \( \neg c_{n+1} \). Choosing either will share a set of similar constraints except the last one.

To approximate this optimization problem, we propose Path Constraint Classifier (PCC), a system that automatically selects a solver by utilizing machine learning techniques. PCC predicts which solver can perform well when receiving a path constraint. The use of machine learning seems straightforward, yet is surprisingly underexplored, with one main technical challenge being how to avoid excessive overhead introduced by machine learning. We address this challenge by taking advantage of the domain knowledge of symbolic execution. Specifically, to accelerate feature extraction, we identify a set of features that can be quickly extracted from path constraints collected during symbolic execution, and propose using a Constraint Feature Cache Table (CFCT) to drastically reduce the feature extraction time by leveraging the structural characteristics of symbolic execution.

Overall, when taking the extra solver prediction and feature extraction overhead into consideration, our system performs better than individual SMT solvers in the KLEE and benchmark datasets: Our system achieves 1.12x and 1.85x speedup in the KLEE and benchmark datasets, respectively, and has the highest solve rate in these two datasets. Particularly, thanks to CFCT, the feature extraction procedure speeds up 28x to 788x...
in the SE dataset and 11x in the benchmark dataset.

Because PCC can be considered as a meta-solver, it can be integrated with other optimization techniques (e.g., constraint simplification, query reduction, and improved solvers) to further improve the scalability of constraint solving in symbolic execution.

The contributions of this paper are listed below:

- We compare the performance of different solvers over three constraint datasets and show the amount of time reduction one might gain by employing solver selection.
- We define the solver selection problem in symbolic execution and identify the fundamental distinctions between it and the generic solver selection problem.
- We design and implement a system called Path Constraint Classifier based on machine learning techniques, with domain-specific optimizations to accelerate the extraction of path constraint features.
- We discuss several future directions to further improve PCC, including the use of backup solvers, logic selection, delayed solving, and dynamic timeout.

The source code of PCC is available on GitHub\(^1\).

II. BACKGROUND AND RELATED WORK

Symbolic execution is a well-known software testing technique. Given a program, a symbolic execution engine systematically explores it and generates path constraints that represent program states using symbolic variables. During exploration, the symbolic execution engine heavily relies on Satisfiability Modulo Theories (SMT) solvers, such as Z3, for two important tasks: (1) checking the satisfiability of a path constraint; (2) obtaining concrete input values that can be used to reach the corresponding state of a path constraint.

Able to generate concrete inputs for software crashes, symbolic execution is an effective method of software testing and has been applied to small-scale industry applications and for academic usage recently. However, when it comes to large-scale applications, many scalability issues arise, limiting the practicability of symbolic execution. One of the performance bottlenecks is constraint solving. An experiment [27] about the time spent on constraint solving shows that the constraint solving time without any optimization could account for up to 90% of the total time of symbolic execution.

Several methods have been proposed to speed up constraint solving by reducing the number of constraints [10], [17], reusing proofs [1], [2], [10], [20], [32], simplifying complex constraints (e.g., nonlinear constraints) [16], [29], or searching for approximate solutions to complex constraints [7], [13], [22], [31]. In addition, specialized constraint solvers are proposed to better support specific operations or library functions. For example, Z3-str [35] treats strings as a primitive type to efficiently solve strings-related theories. CORAL [31] utilizes meta-heuristic search algorithms to support theories with transcendental functions such as trigonometric and logarithmic functions. Our solution can be integrated with these optimization techniques to further improve the scalability of constraint solving in symbolic execution.

As solvers have diverse capabilities, several proposals explore the idea of querying different solvers for different types of constraints. KLEE [10] supports a multi-solver framework using pySMT. However, querying all the supported solvers in parallel is resource-consuming. Hence, our work aims to predict the best solver for a given path constraint and only queries this solver for improved performance. Pasareanu et al. [28] split a path constraint into simple and complex parts. The solutions of the simple part will be reused to help simplify the complex one. This technique has been improved by Hybrid-KLEE [23], which proposes to divide a path constraint into linear and nonlinear parts. The solutions of the linear parts will be treated as initial seeds of a local search algorithm to solve the complex parts. Instead of having pre-defined classifications, our work proposes an automated constraint classification procedure based on machine learning.

SATzilla [34] considers SAT solver selection as an algorithm selection problem and adopts existing portfolio-solving techniques. Healy et al. [19] apply similar techniques to the SMT solving domain. Our work focuses on how to apply portfolio-solving in the symbolic execution domain, and proposes domain-specific optimizations to accelerate the extraction of path constraint features.

III. SOLVER SELECTION PROBLEM IN SYMBOLIC EXECUTION

A. Preliminary Study

We first conduct an experiment to show that no solver can consistently outperform the others and simple classification based on constraint logic is insufficient. This motivates the development of a more sophisticated solver selection method in symbolic execution.

a) Data: We use the SMT-LIB standard [5], [6] as a common format to express constraints. SMT-LIB is well documented and is compatible with most of the current symbolic execution engines and SMT solvers. We obtain constraints with various logic types from the SMT-COMP benchmark [3]. A logic in SMT-LIB can be viewed as a depiction of the expression language, including the theory (e.g., linear or nonlinear arithmetic), the primitive types (e.g., integers, reals or bitvector) and corresponding operations, and also the limitations between the conjunction of primitives. We also use constraint data collected by running two symbolic execution engines, KLEE and angr (the SE dataset). See Section V-A for details.

b) Solvers: We choose five popular SMT solvers from SMT-COMP, which are Z3 [12], MathSAT [11], CVC4 [4], Yices [14], and Boolector [8]. For each constraint, we query each of the five solvers and identify the best solver that uses the least solving time.

c) Results: TABLE I and TABLE II show that even though Boolector seems to be the best solver for the SE dataset, Yices requires the least solving time to solve all the data. We can conclude that Boolector might struggle to solve particular constraints, which increases the total solving time. On the other hand, though Yices is the fastest solver for 5%-37% of constraints, it is stable (i.e., few timeouts and errors) and uses the least total solving time.

\(^1\)https://github.com/csenslab-PCC/PathConstraintClassifier
techniques to predict the performance of each solver when solving a given path constraint, and selects the solver with the best predicted performance.

We implement PCC with an embedded solver interaction interface so that the selected solver can be queried directly according to the classification result. Therefore, one can also view PCC as a special kind of SMT solver that can be flexibly integrated with other symbolic execution engines.

Figure 1 presents an overview of PCC’s internal architecture and a flowchart of how a path constraint is classified to a solver. PCC is composed of three major components: Feature Extractor, Classification Module and Solver Interaction Interface. We also propose a special component called Constraint Feature Cache Table in Feature Extractor to further improve the performance of PCC. The main classification workflow is as follows:

- **Feature Extractor**: Given a new incoming path constraint, the Feature Extractor will generate features for the path constraint in cooperation with Constraint Feature Cache Table, which stores all the generated features in a table in order to speed up the generation of identical features, and pass the features to Classification Module.

- **Classification Module**: The classification module will perform a classification algorithm to determine the most suitable solver for a given path constraint, and then query the chosen solver through the Solver Interaction Interface.

TABLE III shows the ratio of the best solvers and the breakdowns based on logic type of the SMT-COMP benchmark dataset. The result confirms that no solver can consistently outperform the others and simple classification based on constraint logic is insufficient. Thus, it is possible to improve overall performance via careful solver selection. In addition, while logic type is an important factor affecting solvers’ performance, there is still no clear winner in the logic types commonly seen in symbolic execution (e.g., QF_ABV) [9]. Hence, instead of relying on a single solver, we are motivated to combine the strength of multiple solvers.

**B. Problem Definition**

We now define the solver selection problem in symbolic execution. Compared to the generic solver selection problem, this formulation takes into account the time overhead of the selection procedure and the sequence of path constraints is generated by a symbolic execution engine.

**Solver Selection Problem in Symbolic Execution.** Given a sequence of $M$ path constraints $PC = \{pc_1, pc_2, ..., pc_M\}$ and a set of $N$ constraint solvers $S = \{s_1, s_2, ..., s_N\}$, we formulate the solver selection problem in symbolic execution as finding $s_i \in S$ such that $T(S, PC)$, which is the total solving time including solver selection time and solver solving time, is minimized.

$$T(S, PC) = \sum_{i=1}^{M} (T_{solve}(s_i, pc_i) + T_{determine}(pc_i))$$ (1)

where $T_{solve}$ stands for the time spent on solving path constraint $pc$ using solver $s$, and $T_{determine}$ stands for the time spent on determining the best SMT solver for a given path constraint.

Note that this formulation does not include one-time costs such as initialization or training time in ML.

**IV. PCC DESIGN AND IMPLEMENTATION**

To address the solver selection problem in symbolic execution, we propose a new symbolic execution component: **Path Constraint Classifier (PCC)**. PCC leverages machine learning techniques to predict the performance of each solver when solving a given path constraint, and selects the solver with the best predicted performance.

We implement PCC with an embedded solver interaction interface so that the selected solver can be queried directly according to the classification result. Therefore, one can also view PCC as a special kind of SMT solver that can be flexibly integrated with other symbolic execution engines.

**IV. PCC DESIGN AND IMPLEMENTATION**

To address the solver selection problem in symbolic execution, we propose a new symbolic execution component: **Path Constraint Classifier (PCC)**. PCC leverages machine learning techniques to predict the performance of each solver when solving a given path constraint, and selects the solver with the best predicted performance.

We implement PCC with an embedded solver interaction interface so that the selected solver can be queried directly according to the classification result. Therefore, one can also view PCC as a special kind of SMT solver that can be flexibly integrated with other symbolic execution engines.
Fig. 1: The architecture of Path Constraint Classifier

- **Solver Interaction Interface**: When the chosen solver finishes its task, the interaction interface will help transfer and output the result to the upper layer component.

In the following subsection, we describe the design and implementation of Path Constraint Classifier in details, including the features, extraction method and optimization, and finally, the classification algorithm.

**A. Feature Extraction**

Recall that in order to evaluate the performance of different SMT solvers, we use the SMT-LIB standard as a common expression of a path constraint. Therefore, to extract the features of a given path constraint, we first express it in the SMT-LIB form. This conversion is supported by most symbolic execution engines. Next, the path constraint is parsed into a syntax tree, with every internal node representing an operator and every leaf representing an operand (i.e., symbol or constant). An example of a syntax tree is shown in Figure 2.

**B. Feature Selection**

To capture the characteristics of a path constraint, we select 119 features for solver prediction. These features can be roughly classified into four categories:

- **Syntax Node Statistics Features** are the statistics of syntax node properties, including the appearance of arithmetic operators such as `add`, `minus`, `multiply` and `divide`; bitvector operators such as `bv-add`, `bv-minus`, `bv-and` and `bv-or`; quantifiers `forall` and `exists`. There are about 62 features in this category.
- **Tree Structure Features** contain the structure properties of a tree, including the depth of a tree, the number of leaves and nodes.
- **Variable Features** are the features related to variables, such as the number of variables and unique variables, and also the number of variable clauses such as `AND` and `OR`.
- **Logic Features** is a 50-dimension vector composed of only Boolean values, which indicates the types of logic the path constraint belongs to.

The majority of the chosen features are related to **constraint size**, as a larger constraint tends to cause higher burden on a solver. In addition, these features can be easily extracted, reducing overhead imposed on the solver selection procedure. An example of an easily extracted feature is the depth feature, as all we have to do is traverse all the syntax sub-trees and find the one with the biggest depth value, where the time complexity is $O(\#\text{of subtrees})$.

**C. Constraint Feature Cache Table**

In symbolic execution, every path constraint of a state and its child states share a large proportion of identical constraints. More precisely, the only difference will be the last constraint of its child state. Due to this, we are likely to waste excessive time extracting the same features and thus slow down the performance of symbolic execution.

To accelerate the feature extraction procedure, we design a special component called **Constraint Feature Cache Table (CFCT)**. The idea is based on hash-consing [18]. Before extracting features from a syntax tree, we first query whether the features of a given syntax tree have already been stored in CFCT. If so, we directly take the stored features. Otherwise, we extract the required properties from its root and apply the same procedure to all the syntax sub-trees, then store it in CFCT for future use. To show the performance of CFCT, we conduct an experiment to compare extraction times with and without the usage of CFCT, as shown in Section V-B.

**D. Classification Model**

We use **Deep Neural Network (DNN)** with nine dense layers as our classification model, which is a lightweight machine learning model whose classification procedure takes negligible time. Given the features of a path constraint, the model will predict the probability for every SMT solver being the most suitable solver to solve it. This probability indicates which solver is least likely to result in losses among all the other solvers. For example, if the probability of the `z3` solver is 1.0, then `z3` will result in the least amount of loss when solving a given path constraint. The loss function can be flexibly defined depending on the optimization objective.

We design two versions of DNN model with different loss functions and heuristics in mind.

- **DNN-Alpha**: In the first version of the model, our heuristic is very simple: if we can accurately predict the fastest solver, then the total solving time could be minimized. Hence, in this model, we only use the fastest solver as the answer to
each problem and train the model with categorical cross entropy loss.

- **DNN-Beta**: In the second version of the model, we try to take solvability and solver performance into consideration. After all, in symbolic execution, choosing a solver that fails to solve a path constraint is worse than choosing a slow one that successfully solves it. Hence, we give penalty to those that fail to solve due to timeout or error, and assign a loss value based on the time of a successful solve:

\[
\begin{align*}
\text{Loss}(\text{Solver}) &= \begin{cases} 
A \cdot \text{Norm}(T_{\text{Solver}}), & \text{if solved} \\
B, & \text{if timeout} \\
C, & \text{if error}
\end{cases} \\
\text{Norm}(T) &= \frac{T - T_{\text{min}}}{T_{\text{max}} - T_{\text{min}}}
\end{align*}
\]

\(\text{Norm}(\cdot)\) is a function that normalizes the solving time \(T_{\text{Solver}}\) of Solver according to \(T_{\text{max}}\) and \(T_{\text{min}}\), which are the maximal and minimal values of the solving time collected from all the solvers that successfully solve this problem. \(A, B\) and \(C\) are adjustable parameters to evaluate the outcome of a solving procedure, which can be determined by one’s preference for each solving outcome. For example, suppose our preference order for solving outcomes are \{Successfully Solve\} > \{Timeout\} > \{Error\}, which means “it is better for a solver to Successfully Solve a problem than to Timeout”. Therefore, according to our preference, a possible assignment for \(A, B\) and \(C\) could be 50, 100 and 200.

V. Evaluation

A. Data Collection

We use two datasets for experiments: \(\text{SE}\) and \(\text{benchmark}\). The \(\text{SE}\) dataset is collected from two symbolic execution engines: angr [30] and KLEE [10]. Specifically, we perform symbolic execution on eight target binaries from GNU Coreutils 8.29 and dump all the path constraint queries into a .smt2 file. To increase diversity, we also collect the \(\text{benchmark dataset}\) from SMT-COMP [3], which contains constraints of several different logics type stored in the .smt2 file format.

The benchmark dataset and the \(\text{SE}\) dataset contain 8,199 and 59,073 files, respectively. The \(\text{SE}\) dataset contains 39,776 files generated by KLEE and 19,297 files generated by angr.

B. Constraint Feature Cache

To evaluate the effectiveness of using the constraint feature cache table, we compare the extraction times between PCC with and without the usage of CFCT.

For each extraction procedure, we set the time limit to 20 seconds and record the time spent on extraction as well as the number of extracted syntax nodes. The result is shown in Figure 3 (more details in Table IV). The result in the \text{RANDOM} column is an average of 10 iterations of 500 randomly sampled .smt2 files from the benchmark dataset.

In this experiment, we use the \(\text{SE}\) dataset collected from KLEE. According to the experiment results, using CFCT greatly speeds up the extraction procedure for both types of data. Additionally, the \(\text{SE}\) dataset enjoys a higher speedup ratio because the \(\text{SE}\) dataset contains a higher percentage of duplicate constraints than the benchmark dataset.

The speedup factor for binary \text{pathchk} (i.e., 788x) is much higher than other binaries. A possible explanation for this could be that some of the path constraints in \text{pathchk} contain a huge number of constraints, but most are identical. Consequently, a large amount of time can be saved by using CFCT, demonstrating yet again that CFCT can effectively improve the efficiency of feature extraction.

For the benchmark dataset, where all data have relatively low dependency compared to the \(\text{SE}\) dataset, the usage of CFCT results in 11x speedup. This shows that CFCT can greatly speed up the feature extraction procedure for all kinds of datasets.

C. Solving Performance

Training Data Generation. We use five SMT solvers to solve all the constraints with a given timeout (100 seconds in the current experiment) and record the corresponding outcome, including the solving time, accuracy of the answer, and whether an error or timeout event occurs. As for model training, we split our data into training and testing datasets as follows:

- **Benchmark dataset**: We randomly choose 60% of them to be our training dataset, and the remainder is our testing dataset.
- **SE dataset**: We divide the binaries into two groups, one for generating training data and the other reserved for testing.

The reason for splitting the \(\text{SE}\) dataset in this manner is to reduce the chance of information leaks during the training phase. Since a path constraint could share a large proportion of identical constraints with another path constraint in the same binary, randomly splitting the \(\text{SE}\) dataset could mean that the training dataset contains knowledge about the testing dataset, thus making prediction results unreliable.

Experiment Settings. We train our models and compare the prediction results of them directly using an SMT solver. Five solvers are evaluated: cvc4, z3, msat, btor, and yices. In addition, the \text{best} solver represents the ideal selection that always correctly predicts the fastest solver, with zero prediction overhead, for solving a given path constraint.
**Table IV:** Feature extracting time with and without the usage of feature cache

**Experiment Results.** The experiment results are shown in Figure 4 (more details in TABLE V). Based on the results, the performance of our DNN models is the closest to the best solver with respect to both the solve rate and the total time in the KLEE and benchmark datasets.

In the KLEE dataset, almost all of the SMT solvers have a 100% solve rate. Our DNN models perform better than other SMT solvers in terms of total time, even with additional feature extraction and prediction overhead.

In the benchmark dataset, the DNN-beta model has the highest solve rate compared to other solvers. Despite the fact that btor requires the least total time, it can only solve about 20% of the data, which may be intolerable as this could prevent the symbolic execution engine from exploring further paths if most of the constraints are unsolved. z3 seems to be the best solver among the five SMT solvers, because z3 can solve almost 98% of data within a reasonable time. However, our DNN-beta results in about 2x speedup and can solve almost 99% of data.

In the angr dataset, our DNN models do not perform well but the performance is still tolerable compared with other solvers like btor, cvc4 and z3, which result in 1.8x to 5x performance speedup. We believe that increasing the size of the angr dataset can improve the performance of our DNN model.

Notice that in the SE dataset, the performance of DNN-alpha is better than DNN-beta. This is because all the solvers we tested can solve all symbolic data, and the factors of timeout and error do not matter much. Thus, choosing the fastest solver will be the best decision. On the other hand, in the benchmark dataset, the number of timeouts and errors becomes an important factor for solver selection. DNN-beta can make better decisions in choosing the fastest and solvable solver than choosing the fastest but unsolvable solver.

**VI. CONCLUSION AND FUTURE WORK**

In this work, we present Path Constraint Classifier (PCC), a new component in symbolic execution engines, which aims to improve the efficiency of constraint solving by predicting the fastest solver for a given path constraint. We first conduct a preliminary study on modern SMT solvers to demonstrate the need for solver selection. Next, we define the solver selection problem in symbolic execution and propose a machine learning based solution with several optimizations to tackle this problem. PCC transforms path constraints into a syntax tree to extract features, and uses Deep Neural Network (DNN) with two self-designed loss functions to predict the best solver in different scenarios. We also propose constraint feature cache to greatly reduce the overhead in feature extraction. Finally, we evaluate the solving performance of PCC and demonstrate that it can achieve better performance than individual SMT solvers, thus improving the efficiency of constraint solving.

Several interesting research directions remain for future exploration. As our solution is based on machine learning, one future work is to perform feature engineering and model tuning (e.g., a better loss function) to increase the robustness of prediction results while reducing the overhead cost of the determination procedure. To improve the model, we would also like to collect more symbolic data (e.g., path constraints in the
<table>
<thead>
<tr>
<th>Solver</th>
<th>Solve Rate</th>
<th>Number of Timeout</th>
<th>Number of Error</th>
<th>Solving Time</th>
<th>Timeout Time</th>
<th>Error Time</th>
<th>Feature Extraction Time</th>
<th>Predict Time</th>
<th>Total Time</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>KLEE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>best</td>
<td>100.00%</td>
<td>0</td>
<td>0</td>
<td>137.70</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>137.70</td>
</tr>
<tr>
<td>DNN-alpha</td>
<td>100.00%</td>
<td>0</td>
<td>0</td>
<td>150.09</td>
<td>0.00</td>
<td>0.00</td>
<td>14.54</td>
<td>2.28</td>
<td>166.91</td>
</tr>
<tr>
<td>DNN-beta</td>
<td>100.00%</td>
<td>0</td>
<td>0</td>
<td>150.61</td>
<td>0.00</td>
<td>0.00</td>
<td>14.54</td>
<td>2.23</td>
<td>167.38</td>
</tr>
<tr>
<td>yices</td>
<td>100.00%</td>
<td>0</td>
<td>0</td>
<td>186.18</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>186.18</td>
</tr>
<tr>
<td>btor</td>
<td>100.00%</td>
<td>0</td>
<td>0</td>
<td>226.26</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>226.26</td>
</tr>
<tr>
<td>msat</td>
<td>100.00%</td>
<td>0</td>
<td>0</td>
<td>444.24</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>444.24</td>
</tr>
<tr>
<td>z3</td>
<td>100.00%</td>
<td>0</td>
<td>0</td>
<td>559.31</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>559.31</td>
</tr>
<tr>
<td>cvc4</td>
<td>99.99%</td>
<td>1</td>
<td>0</td>
<td>1526.81</td>
<td>100.01</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>1626.82</td>
</tr>
<tr>
<td><strong>ANGR</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>best</td>
<td>100.00%</td>
<td>0</td>
<td>0</td>
<td>117.17</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>117.17</td>
</tr>
<tr>
<td>yices</td>
<td>100.00%</td>
<td>0</td>
<td>0</td>
<td>174.37</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>174.37</td>
</tr>
<tr>
<td>msat</td>
<td>100.00%</td>
<td>0</td>
<td>0</td>
<td>318.17</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>318.17</td>
</tr>
<tr>
<td>DNN-alpha</td>
<td>100.00%</td>
<td>0</td>
<td>0</td>
<td>318.47</td>
<td>0.00</td>
<td>0.00</td>
<td>5.10</td>
<td>1.14</td>
<td>324.71</td>
</tr>
<tr>
<td>DNN-beta</td>
<td>99.99%</td>
<td>1</td>
<td>0</td>
<td>324.96</td>
<td>102.02</td>
<td>0.00</td>
<td>5.10</td>
<td>1.11</td>
<td>433.19</td>
</tr>
<tr>
<td>btor</td>
<td>99.98%</td>
<td>2</td>
<td>0</td>
<td>557.09</td>
<td>204.04</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>761.13</td>
</tr>
<tr>
<td>cvc4</td>
<td>100.00%</td>
<td>0</td>
<td>0</td>
<td>1101.88</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>1101.88</td>
</tr>
<tr>
<td>z3</td>
<td>100.00%</td>
<td>0</td>
<td>0</td>
<td>1638.07</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>1638.07</td>
</tr>
<tr>
<td><strong>BENCHMARK</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>btor</td>
<td>18.49%</td>
<td>(97.62%)</td>
<td>1</td>
<td>26001</td>
<td>335.23</td>
<td>102.41</td>
<td>1.54</td>
<td>-</td>
<td>439.18</td>
</tr>
<tr>
<td>best</td>
<td>100.00%</td>
<td>0</td>
<td>0</td>
<td>883.04</td>
<td>0.00</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
<td>883.04</td>
</tr>
<tr>
<td>DNN-beta</td>
<td>98.65%</td>
<td>30</td>
<td>12</td>
<td>1445.6</td>
<td>3102.72</td>
<td>0.21</td>
<td>101.88</td>
<td>0.36</td>
<td>4650.77</td>
</tr>
<tr>
<td>DNN-alpha</td>
<td>97.87%</td>
<td>34</td>
<td>34</td>
<td>1505.95</td>
<td>3442.35</td>
<td>80.10</td>
<td>101.88</td>
<td>0.36</td>
<td>5130.64</td>
</tr>
<tr>
<td>msat</td>
<td>30.15%</td>
<td>(94.94%)</td>
<td>34</td>
<td>3265.41</td>
<td>3470.19</td>
<td>4.98</td>
<td>-</td>
<td>-</td>
<td>6740.58</td>
</tr>
<tr>
<td>yices</td>
<td>57.69%</td>
<td>(94.73%)</td>
<td>68</td>
<td>1241.03</td>
<td>6801.32</td>
<td>6.23</td>
<td>-</td>
<td>-</td>
<td>8048.58</td>
</tr>
<tr>
<td>z3</td>
<td>97.96%</td>
<td>62</td>
<td>3</td>
<td>2146.23</td>
<td>6386.44</td>
<td>79.81</td>
<td>-</td>
<td>-</td>
<td>8612.48</td>
</tr>
<tr>
<td>cvc4</td>
<td>93.61%</td>
<td>98</td>
<td>106</td>
<td>4364.52</td>
<td>9811.21</td>
<td>3.22</td>
<td>-</td>
<td>-</td>
<td>14178.95</td>
</tr>
</tbody>
</table>

TABLE V: The performance of different SMT solvers and our PCC models at benchmark, angr and KLEE dataset

Another interesting direction is examining how to combine PCC with other solver optimization techniques, including constraint simplification, query reduction, and improved solvers, and evaluating the performance. It is also important to note that when using PCC, the symbolic execution engines may frequently switch between different solvers. This might cause frequent context switching that affects the overall performance of symbolic execution.

Finally, existing symbolic execution engines mostly treat solvers as a blackbox. It would be interesting to see whether and how a symbolic execution engine could dynamically determine how to encode a path condition [21] such that the constraints can be solved quickly by solvers.

ACKNOWLEDGMENT

This work was supported in part by the Ministry of Science and Technology of Taiwan under grant MOST 108-2636-E-002-008 and by Institute for Information Industry under grant 106-EC-17-D-11-1502.

*The bracket is the solve rate of supported logic.
†The solver specializes on specific types of logic, which may lead to a high error rate because unsupported logic is counted as an error in our experiment.

REFERENCES


M. Gario, A. Micheli, and F. B. Kessler, “Pysmt: a solver-agnostic B. Dutertre, “Yices 2.2,” in
C. Cadar, D. Dunbar, and D. Engler, “Klee: Unassisted and automatic
X. Jia, C. Ghezzi, and S. Ying, “Enhancing reuse of constraint