Enhancing Reuse of Constraint Solutions to Improve Symbolic Execution

Xiangyang Jia (Wuhan University)
Carlo Ghezzi (Politecnico di Milano)
Shi Ying (Wuhan University)
Outline

❖ Motivation
❖ Logical Basis of our Approach
❖ GreenTrie Framework
  ❖ Constraint Reduction
  ❖ Constraint Storing
  ❖ Constraint Querying
❖ Evaluation
❖ Conclusion and Future Work
Motivation

❖ **Symbolic Execution (SE)**
  ❖ A well-known program analysis technique, mainly used for test-case generation and bug finding.

❖ **Constraint Solving**
  ❖ The most time-consuming work in SE
  ❖ Optimization approaches:
    ❖ Irrelevant constraint elimination
    ❖ Caching and reuse
Motivation

Aggregated data over 73 applications

- **Base**
- **Irrelevant Constraint Elimination**
- **Caching**
- **Irrelevant Constraint Elimination + Caching**

[From Shauvik Roy Choudhary’s Slides]
Motivation

- Reuse of Constraint Solutions

Equivalence based approach (Green)

$x > 0$ is equivalent to $y > 0$
$x + 1 > 0 \land x \leq 1$ is equivalent to $y < 2 \land y \geq 0$ (if $x$, $y$ are integers)
Motivation

❖ Reuse of Constraint Solutions

Equivalence based approach (Green)

Subset/superset based approach (KLEE)

If $A \land B \land C$ is satisfiable, then $A \land B$ is satisfiable
If $A \land B \land C$ is unsatisfiable, then $A \land B \land C \land D$ is unsatisfiable
Motivation

❖ Reuse of Constraint Solutions

- Equivalence based approach (Green)
- Subset/superset based approach (KLEE)

If $x > 0$ is satisfiable, can we prove $x > -1$ satisfiable?
If $x < 0 \land x > 1$ is unsatisfiable, can we prove $x < -1 \land x > 2$ unsatisfiable?
Motivation

- Reuse of Constraint Solutions

If \( x > 0 \) is satisfiable, can we prove \( x > -1 \) satisfiable?
If \( x < 0 \land x > 1 \) is unsatisfiable, can we prove \( x < -1 \land x > 2 \) unsatisfiable?
Logical Basis of our Approach

Implication and Satisfiability

Providing $C_1 \rightarrow C_2$

- if $C_1$ is satisfiable, $C_2$ is satisfiable
- if $C_2$ is unsatisfiable, $C_1$ is unsatisfiable

It looks easy to apply it to constraint reuse! However, there is a problem: Implication checking with SAT/SMT solver is even more expensive than only solving the single constraint itself.
**Logical Basis of our Approach**

- **The subset/superset (KLEE)**
  - \( \{c_1, c_2\} \subseteq \{c_1, c_2, c_3\} \) means \( c_1 \land c_2 \land c_3 \rightarrow c_1 \land c_2 \)

- **Logical subset/superset**
  - Given two constraint sets \( X, Y \), if \( \forall a \in X \exists b \in Y (b \rightarrow a) \), then \( X \) is a logical subset of \( Y \), and \( Y \) is a logical superset of \( X \)
  - E.g.: \( X = \{m \neq 0, m > -1, m < 2\} \), \( Y = \{m > 1, m < 2\} \)
  - It is easy to prove that \((m > 1 \land m < 2) \rightarrow (m \neq 0 \land m > -1 \land m < 2)\)

---

**the subset/superset is a specific case of logical subset/superset**

Logical subset/superset checks more implication cases!
- the two sets might have totally different atomic constraints
- the length of logical superset may be shorter than its subset
Logical Basis of our Approach

- Implication checking rules for atomic constraints

\[
\begin{align*}
(R1) \quad & \frac{C \rightarrow C}{C} \\
(R2) \quad & \frac{n \neq n'}{P + n = 0 \rightarrow P + n' \neq 0} \\
(R3) \quad & \frac{n \geq n'}{P + n = 0 \rightarrow P + n' \leq 0} \\
(R4) \quad & \frac{n \leq n'}{P + n = 0 \rightarrow P + n' \geq 0} \\
(R5) \quad & \frac{n > n'}{P + n \leq 0 \rightarrow P + n' \neq 0} \\
(R6) \quad & \frac{n > n'}{P + n \leq 0 \rightarrow P + n' \leq 0} \\
(R7) \quad & \frac{n < n'}{P + n \geq 0 \rightarrow P + n' \neq 0} \\
(R8) \quad & \frac{n < n'}{P + n \geq 0 \rightarrow P + n' \geq 0}
\end{align*}
\]

P: non-constant prefix, n: constant number
E.g. \(x+y+3\geq 0\) has a non-constant prefix \(x+y\) and a constant number 3
GreenTrie Framework

- Architecture of GreenTrie

Two separated stores for SAT and UNSAT constraints
GreenTrie Framework

- **Architecture of GreenTrie**

  - **Symbolic Executor**

  GreenTrie

  Known: $C_1 \land C_2 \land \ldots \land C_n$  
  Unknown: $C_0$

  - **Slicing**

  - **Canonization**

  - **Reusing**

  - **Translating**

  - **Storing**

  - **Querying**

  - **Reduction**

  - **L-Trie**

  Satisfiable Constraint Store (SCS)

  - Logical Index

  - Constraint Trie

  Unsatisfiable Constraint Store (UCS)

  - Logical Index

  - Constraint Trie

  - **A constraint trie with a logical index**

  - **Constraint Solver:** CVC3, Z3, Yices, Choco...
GreenTrie Framework

- **Architecture of GreenTrie**

  - **Symbolic Executor**

  - **Known:** $C_1 \land C_2 \land \ldots \land C_n$  
    - **Slicing**
    - **Canonization**
      - **sliced:** $C_0 \land C_1 \land \ldots \land C_{n-1}$
      - **reduced:** $C' \land C'_m$
    - **Canontized:** $C'_1 \land C_0 \land \ldots \land C'_n$
  
  - **Unknown:** $C_0$
    - **Translating**
    - **Reusing**
      - if result is null
        - **Storing**
          - **put($C' \land C'_m$, SAT)**

- **L-Trie**
  - **Satisfiable Constraint Store (SCS)**
    - **Logical Index**
    - **Constraint Trie**
  - **Unsatisfiable Constraint Store (UCS)**
    - **Logical Index**
    - **Constraint Trie**

- **Implication Partial Order Graph (IPOG)**
  - Contains all atomic sub-constraints appearing in its associated constraint trie.

- **Storing the Constraints**
  - Everytime a constraint is solved (or it is proved to be unsatisfiable), SCS (respectively UCS) remove redundant sub-constraints for better matching.

**Figure 1:** The overview architecture of GreenTrie
GreenTrie Framework

• Architecture of GreenTrie

Symbolic Executor

Slicing

Canonization

Reusing

Translating

Reduction

Querying

Storing

Known: C₁^C₂...Cₙ

Unknown: C₀

sliced: C₀^C₁^Cₙ₋₁

canonized: C'₁^C₀^C'ₙ₋₁

reduced: C'₁^C'm

get(C'₁^C'm)

result

if result is null

put(C'₁^C'm;SAT)

Logical Index

Constraint Trie

Unsatisfiable Constraint Store (UCS)

Logical Index

Constraint Trie

Query reusable constraints through logical subset/superset checking
GreenTrie Framework

**Architecture of GreenTrie**

1. We discard all exceptional points that are outside the overlapping interval; in the example, the value of $E$ becomes $\{0, 4\}$.
2. If one endpoint of the overlapping interval $A$ (or $B$) belongs to $E$, we (repeatedly) change its value and eliminate $A$ (or $B$) from $E$ at the same time. In the example after this step the interval becomes $[3, 3]$ and the new value of $E$ is $\{0\}$.
3. If the overlapping interval is empty then the constraint is unsatisfiable and we report a conflict; otherwise we translate $[A, B]$ and $E$ into a constraint in normal form. In the example, the final result of our reduction is $x + y + 3 \geq 0$ and $x + y - 3 \leq 0$. 

**Constraint Storing**

L-Trie provides a different storage scheme that replaces the Redis store of Green:

- Unlike Redis, which stores the strings representing constraints and solutions as key-value pairs, L-Trie splits constraints into sub-constraint sets, and stores them into tries, in order to support logical subset and superset queries based on Theorem 1.
- L-Trie stores unsatisfiable and satisfiable constraints into separate areas: the **Unsatisfiable Constraint Store (UCS)** and the **Satisfiable Constraint Store (SCS)** respectively. The two areas are organized differently to efficiently support logical subset querying and logical superset querying, which pose different requirements.
- L-Trie maintains a logical index for each of the two tries, to support efficient check of the implication relations. The logical index is represented as an implication partial order graph (IPOG), whose nodes contain references to nodes in the trie.

**Constraint Trie**

The constraint trie is designed to store a sub-constraint set of solved constraints. The sub-constraint set is sorted in lexicographic order based on string comparison, to guarantee that sub-constraints with same non-constant prefix are kept close to each other. The labels of the constraint trie record the sub-constraints. The leaf nodes indicate the end of the constraint and are annotated with the solution (the solution is null for the leaves of the UCS trie). As shown in Fig. 2, the leaf node C2 corresponds to a constraint $v_0 + 5 > 0 \land v_0 + v_1 < 0$, which has a solution $\{v_0 : 0, v_1 : 1\}$, and its sub-constraints $v_0 + 5 > 0$ and $v_0 + v_1 < 0$, are annotated as edge labels in the path.

**Implication Partial Order Graph (IPOG)**

IPOG is a graph that contains all the atomic sub-constraints appearing in its associated constraint trie, and arranges them as a graph based on the partial order defined by the implication relation. With this graph, given a constraint $C$, we can query the sub-constraints which imply $C$, as well as the sub-constraints which $C$ implies, as we will see later. This is useful to improve the efficiency of implication checking in logical subset and superset querying. IPOG nodes are labeled by a sub-constraint and have references to all trie nodes whose input edge is labeled with exactly this sub-constraint. Through these references, it is possible to trace all the occurrences of a given sub-constraint.

**Storing the Constraints**

Everytime a constraint is solved (or it is proved to be unsatisfiable), SCS (respectively UCS) stores the solving result into stores.
Constraint Reduction

- Constraint Reduction
  - target: remove redundant sub-constraints
  - idea: interval computation-based constraint reduction

Example

\[ x + y + 3 \geq 0 \land x + y + 5 \geq 0 \land x + y - 4 \leq 0 \land x + y \neq 0 \land x + y + 6 \neq 0 \land x + y - 4 \neq 0 \]

compute: \([-3, \infty) \cap [-5, \infty) \cap (-\infty, 4] - \{0, -6, 4\} = [-3, 4) - \{0\}\]

reduced: \(x + y + 3 \geq 0 \land x + y - 4 < 0 \land x + y \neq 0\)
Constraint Storing

C3 represents a constraint \( V_0 + 5 \geq 0 \land V_1 + (-1) \leq 0 \), which has a solution \( \{v0:0, v1:-5\} \).
v₀ + 5 ≥ 0 is implied by v₀ + (-3) = 0 and v₀ + (-4) = 0.

v₀ + 5 ≥ 0 has one occurrence in the trie, therefore it has a reference to the successive trie node.
Constraint Querying

- Implication Set (IS) and Reverse Implication Set (RIS)

Example
Constraint: $v_0 \geq 0$
IS_{v_0 \geq 0}: \{v_0 + 5 \geq 0\}
RIS_{v_0 \geq 0}: \{v_0 + (-3) = 0, v_0 + (-4) = 0\}
Constraint Querying

❖ Logical Superset Checking Algorithm
❖ Find a path in trie, so that every sub-constraint in target constraint is implied by at least one constraint on this path

Example
Target: $v_0 \neq 0 \land v_0 + (-1) \neq 0 \land v_1 + (-2) \leq 0$
RIS$_{v_1 + (-2) \leq 0}$: $\{v_1 + (-1) \leq 0\}$
So, we got two candidate paths to check!

Start from these two nodes!
Constraint Querying

Logical Superset Checking Algorithm

Example
Target: \( v_0 \neq 0 \land v_0 + (-1) \neq 0 \land v_1 + (-2) \leq 0 \)

\( \text{RIS}_{v_0 \neq 1} : \{v_0 + (-3) = 0, v_0 + (-4) = 0\} \)

\( v_0 + 5 \geq 0 \) is not in the RIS, the trie root is reached, so this path doesn’t match!
Constraint Querying

Logical Superset Checking Algorithm

Example

Target: $v_0 \neq 0 \land v_0 + (-1) \neq 0 \land v_1 + (-2) \leq 0$

$\text{RIS}_{v_0 \neq 1} : \{v_0 + (-3) = 0, v_0 + (-4) = 0\}$

$v_0 + (-3) \geq 0$ is in the RIS, go on to check next sub-constraint of target!
Constraint Querying

Logical Superset Checking Algorithm

Example
Target: \( v_0 \neq 0 \land v_0 +(-1)\neq 0 \land v_1 +(-2)\leq 0 \)
\[ \text{RIS}_{v_0 \neq 0} = \{ v_0 +(-3)=0, v_0 +(-4)=0 \} \]

\( v_0 +(-3) \geq 0 \) is also in the RIS of \( v_0 \neq 0 \), now, every sub-constraint in target is implied by one constraint on this path. C4 is the reusable constraint!
Constraint Querying

Logical Subset Checking Algorithm

Target: \(v_0 + (-1) \geq 0 \land v_0 + 3 \neq 0 \land v_0 + 4 \leq 0\)

Union of ISs of the sub-constraints: \(\{v_0 \geq 0\} \cup \{v_0 + 2 \leq 0, v_0 + 1 \leq 0\}\)

\(\text{IS}_{\text{union}} = \{v_0 \geq 0, v_0 + 2 \leq 0, v_0 + 1 \leq 0\}\)

We will find a trie path, so that all its sub-constraints on the path exists in \(\text{IS}_{\text{union}}\).
Constraint Querying

- Logical Subset Checking Algorithm

**Target:** \( v_0 +(-1)>=0 \land v_0+3!= 0 \land v_0+4<= 0 \)

\( IS_{union} = \{v_0 >=0, v_0+2<= 0, v_0+1<= 0\} \)
Constraint Querying

Logical Subset Checking Algorithm

**Target:** $v_0 + (-1) \geq 0 \land v_0 + 3 \neq 0 \land v_0 + 4 \leq 0$

**IS**$_{\text{union}} = \{ v_0 \geq 0, v_0 + 2 \leq 0, v_0 + 1 \leq 0 \}$

We found two paths, so the target constraint is unsatisfiable.
Evaluation

❖ Research Question

❖ Does GreenTrie achieve better reuse and save more time than other approaches (Green, KLEE)?

❖ Benchmarks

❖ 6 programs from Green (Willem Visser’s FSE’12 paper)
❖ 1 program from Guowei Yang’s ISSTA 2012 paper.

❖ Experiment scenarios

❖ (1) reuse in a single run of the program
❖ (2) reuse across runs of different versions of the same program
❖ (3) reuse across different programs
Evaluation

❖ Experiment setup
  ❖ PC with a 2.5GHz Intel processor with 4 cores and 4Gb of memory
  ❖ We implemented GreenTrie by extending Green
  ❖ We implemented KLEE’s subset/superset checking approach, and also integrated it into Green as an extension.
  ❖ Symbolic executor: Symbolic Pathfinder (SPF)
  ❖ Constraint Solver: Z3
Reusing code across runs as a way to improve the SE process.

## Evaluation

### Reuse in a Single Run

#### Table 1: Experimental results of reuse in single run

<table>
<thead>
<tr>
<th>Program</th>
<th>$n_0$</th>
<th>$n_1$</th>
<th>$n_2$</th>
<th>$n_3$</th>
<th>$R'$</th>
<th>$R''$</th>
<th>$t_0$(ms)</th>
<th>$t_1$(ms)</th>
<th>$t_2$(ms)</th>
<th>$t_3$(ms)</th>
<th>$T'$</th>
<th>$T''$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trityp</td>
<td>32</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>0.00%</td>
<td>0.00%</td>
<td>1040</td>
<td>915</td>
<td>922</td>
<td>995</td>
<td>-8.74%</td>
<td>-7.92%</td>
</tr>
<tr>
<td>Euclid</td>
<td>642</td>
<td>552</td>
<td>464</td>
<td>464</td>
<td>15.94%</td>
<td>0.00%</td>
<td>5105</td>
<td>6503</td>
<td>7274</td>
<td>6311</td>
<td>2.95%</td>
<td>13.24%</td>
</tr>
<tr>
<td>TCAS</td>
<td>680</td>
<td>41</td>
<td>20</td>
<td>14</td>
<td>65.85%</td>
<td>30.00%</td>
<td>12742</td>
<td>3356</td>
<td>2182</td>
<td>2165</td>
<td>35.49%</td>
<td>0.78%</td>
</tr>
<tr>
<td>TreeMap1</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>0.00%</td>
<td>0.00%</td>
<td>871</td>
<td>942</td>
<td>947</td>
<td>882</td>
<td>6.37%</td>
<td>6.86%</td>
</tr>
<tr>
<td>TreeMap2</td>
<td>148</td>
<td>148</td>
<td>140</td>
<td>140</td>
<td>5.41%</td>
<td>0.00%</td>
<td>2918</td>
<td>2542</td>
<td>2851</td>
<td>2606</td>
<td>-2.52%</td>
<td>8.59%</td>
</tr>
<tr>
<td>TreeMap3</td>
<td>1080</td>
<td>956</td>
<td>833</td>
<td>806</td>
<td>15.69%</td>
<td>3.24%</td>
<td>21849</td>
<td>10729</td>
<td>9871</td>
<td>9871</td>
<td>8.00%</td>
<td>16.41%</td>
</tr>
<tr>
<td>BinTree1</td>
<td>84</td>
<td>41</td>
<td>25</td>
<td>25</td>
<td>39.02%</td>
<td>0.00%</td>
<td>1476</td>
<td>1103</td>
<td>1092</td>
<td>1027</td>
<td>6.89%</td>
<td>5.95%</td>
</tr>
<tr>
<td>BinTree2</td>
<td>472</td>
<td>238</td>
<td>133</td>
<td>118</td>
<td>50.42%</td>
<td>11.28%</td>
<td>4322</td>
<td>3648</td>
<td>3156</td>
<td>2872</td>
<td>21.27%</td>
<td>9.00%</td>
</tr>
<tr>
<td>BinTree3</td>
<td>3252</td>
<td>1654</td>
<td>939</td>
<td>873</td>
<td>47.22%</td>
<td>7.03%</td>
<td>36581</td>
<td>17197</td>
<td>14764</td>
<td>12041</td>
<td>29.98%</td>
<td>18.44%</td>
</tr>
<tr>
<td>BinomialHeap1</td>
<td>448</td>
<td>32</td>
<td>23</td>
<td>19</td>
<td>40.63%</td>
<td>17.39%</td>
<td>3637</td>
<td>2137</td>
<td>2046</td>
<td>2017</td>
<td>5.62%</td>
<td>1.42%</td>
</tr>
<tr>
<td>BinomialHeap2</td>
<td>3184</td>
<td>190</td>
<td>85</td>
<td>68</td>
<td>64.21%</td>
<td>20.00%</td>
<td>27165</td>
<td>7653</td>
<td>6442</td>
<td>6071</td>
<td>20.67%</td>
<td>5.76%</td>
</tr>
<tr>
<td>BinomialHeap3</td>
<td>23320</td>
<td>988</td>
<td>337</td>
<td>288</td>
<td>70.85%</td>
<td>14.54%</td>
<td>249224</td>
<td>28549</td>
<td>31892</td>
<td>21392</td>
<td>25.07%</td>
<td>32.92%</td>
</tr>
<tr>
<td>MerArbiter</td>
<td>60648</td>
<td>21</td>
<td>15</td>
<td>13</td>
<td>38.10%</td>
<td>13.33%</td>
<td>&gt;10min</td>
<td>304726</td>
<td>290854</td>
<td>272813</td>
<td>10.47%</td>
<td>6.20%</td>
</tr>
</tbody>
</table>

$n_i$: the number of invocations to solver  
$t_i$: running time for symbolic execution  
i=0: SE without reuse  
i=1: SE with Green  
i=2: SE with KLEE’s approach  
i=3: SE with GreenTrie  
Reuse improvement ratio: $R' = (n_1 - n_3) / n_1$  
Reuse improvement ratio: $R'' = (n_2 - n_3) / n_2$  
Time improvement ratio: $T' = (t_1 - t_3) / t_1$  
Time improvement ratio: $T'' = (t_2 - t_3) / t_2$
Evaluation

❖ Reuse in a Single Run

Table 1: Experimental results of reuse in single run

<table>
<thead>
<tr>
<th>Program</th>
<th>n0</th>
<th>n1</th>
<th>n2</th>
<th>n3</th>
<th>R'</th>
<th>R''</th>
<th>t0(ms)</th>
<th>t1(ms)</th>
<th>t2(ms)</th>
<th>t3(ms)</th>
<th>T'</th>
<th>T''</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trityp</td>
<td>32</td>
<td>28</td>
<td>28</td>
<td>28</td>
<td>0.00%</td>
<td>0.00%</td>
<td>1040</td>
<td>915</td>
<td>922</td>
<td>995</td>
<td>-8.74%</td>
<td>-7.92%</td>
</tr>
<tr>
<td>Euclid</td>
<td>642</td>
<td>552</td>
<td>464</td>
<td>464</td>
<td>15.94%</td>
<td>0.00%</td>
<td>5105</td>
<td>6503</td>
<td>7274</td>
<td>6311</td>
<td>2.95%</td>
<td>13.24%</td>
</tr>
<tr>
<td>TCAS</td>
<td>680</td>
<td>41</td>
<td>20</td>
<td>28</td>
<td>65.85%</td>
<td>30.00%</td>
<td>12742</td>
<td>3356</td>
<td>2182</td>
<td>2165</td>
<td>35.49%</td>
<td>0.78%</td>
</tr>
<tr>
<td>TreeMap1</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>0.00%</td>
<td>0.00%</td>
<td>871</td>
<td>942</td>
<td>947</td>
<td>882</td>
<td>6.37%</td>
<td>6.86%</td>
</tr>
<tr>
<td>TreeMap2</td>
<td>148</td>
<td>148</td>
<td>140</td>
<td>140</td>
<td>5.41%</td>
<td>0.00%</td>
<td>2918</td>
<td>2542</td>
<td>2851</td>
<td>2606</td>
<td>-2.52%</td>
<td>8.59%</td>
</tr>
<tr>
<td>TreeMap3</td>
<td>1080</td>
<td>956</td>
<td>833</td>
<td>806</td>
<td>15.69%</td>
<td>3.24%</td>
<td>21849</td>
<td>10729</td>
<td>11809</td>
<td>9871</td>
<td>8.00%</td>
<td>16.41%</td>
</tr>
<tr>
<td>BinTree1</td>
<td>84</td>
<td>41</td>
<td>25</td>
<td>25</td>
<td>39.02%</td>
<td>0.00%</td>
<td>1476</td>
<td>1103</td>
<td>1092</td>
<td>1027</td>
<td>6.89%</td>
<td>5.95%</td>
</tr>
<tr>
<td>BinTree2</td>
<td>472</td>
<td>238</td>
<td>133</td>
<td>118</td>
<td>50.42%</td>
<td>11.28%</td>
<td>4322</td>
<td>3648</td>
<td>3156</td>
<td>2872</td>
<td>21.27%</td>
<td>9.00%</td>
</tr>
<tr>
<td>BinTree3</td>
<td>3252</td>
<td>1654</td>
<td>939</td>
<td>873</td>
<td>47.22%</td>
<td>7.03%</td>
<td>36581</td>
<td>17197</td>
<td>14764</td>
<td>12041</td>
<td>29.98%</td>
<td>18.44%</td>
</tr>
<tr>
<td>BinomialHeap1</td>
<td>448</td>
<td>32</td>
<td>23</td>
<td>19</td>
<td>40.63%</td>
<td>17.39%</td>
<td>3637</td>
<td>2137</td>
<td>2046</td>
<td>2017</td>
<td>5.62%</td>
<td>1.42%</td>
</tr>
<tr>
<td>BinomialHeap2</td>
<td>3184</td>
<td>190</td>
<td>85</td>
<td>68</td>
<td>64.21%</td>
<td>20.00%</td>
<td>27165</td>
<td>7653</td>
<td>6442</td>
<td>6071</td>
<td>20.67%</td>
<td>5.76%</td>
</tr>
<tr>
<td>BinomialHeap3</td>
<td>23320</td>
<td>988</td>
<td>337</td>
<td>288</td>
<td>70.85%</td>
<td>14.54%</td>
<td>249224</td>
<td>28549</td>
<td>31892</td>
<td>21392</td>
<td>25.07%</td>
<td>32.92%</td>
</tr>
<tr>
<td>MerArbiter</td>
<td>60648</td>
<td>21</td>
<td>15</td>
<td>13</td>
<td>38.10%</td>
<td>13.33%</td>
<td>&gt;10min</td>
<td>304726</td>
<td>290854</td>
<td>272813</td>
<td>10.47%</td>
<td>6.20%</td>
</tr>
<tr>
<td>total/average</td>
<td>94014</td>
<td>4913</td>
<td>3066</td>
<td>2880</td>
<td>41.38%</td>
<td>6.07%</td>
<td>/ 390000</td>
<td>374012</td>
<td>341063</td>
<td>12.55%</td>
<td>9.35%</td>
<td></td>
</tr>
</tbody>
</table>

GreenTrie gets better reuse ratio and saves more time when the scale of execution increases.
Evaluation

- Reuse across Runs

Table 2: Experimental results of reuse across runs (program Euclid)

<table>
<thead>
<tr>
<th>Changes</th>
<th>$n_0$</th>
<th>$n_1$</th>
<th>$n_2$</th>
<th>$n_3$</th>
<th>$R'$</th>
<th>$R''$</th>
<th>$t_1$(ms)</th>
<th>$t_2$(ms)</th>
<th>$t_3$(ms)</th>
<th>$T'$</th>
<th>$T''$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADD#1</td>
<td>492</td>
<td>432</td>
<td>5</td>
<td>3</td>
<td>99.54%</td>
<td>60.00%</td>
<td>3896</td>
<td>1375</td>
<td>1329</td>
<td>65.89%</td>
<td>3.35%</td>
</tr>
<tr>
<td>ADD#2</td>
<td>438</td>
<td>331</td>
<td>216</td>
<td>216</td>
<td>34.74%</td>
<td>0.00%</td>
<td>2830</td>
<td>3275</td>
<td>2284</td>
<td>19.29%</td>
<td>30.26%</td>
</tr>
<tr>
<td>ADD#3</td>
<td>220</td>
<td>170</td>
<td>32</td>
<td>2</td>
<td>98.82%</td>
<td>93.75%</td>
<td>1382</td>
<td>972</td>
<td>552</td>
<td>60.06%</td>
<td>43.21%</td>
</tr>
<tr>
<td>DEL#1</td>
<td>438</td>
<td>322</td>
<td>156</td>
<td>126</td>
<td>60.87%</td>
<td>19.23%</td>
<td>3428</td>
<td>2670</td>
<td>2171</td>
<td>36.67%</td>
<td>18.69%</td>
</tr>
<tr>
<td>DEL#2</td>
<td>492</td>
<td>426</td>
<td>350</td>
<td>134</td>
<td>68.54%</td>
<td>61.71%</td>
<td>3777</td>
<td>4483</td>
<td>2046</td>
<td>45.83%</td>
<td>54.36%</td>
</tr>
<tr>
<td>DEL#3</td>
<td>642</td>
<td>552</td>
<td>112</td>
<td>111</td>
<td>79.89%</td>
<td>0.89 %</td>
<td>4649</td>
<td>2560</td>
<td>2049</td>
<td>55.93%</td>
<td>19.96%</td>
</tr>
<tr>
<td>MOD#1</td>
<td>642</td>
<td>552</td>
<td>464</td>
<td>463</td>
<td>16.12%</td>
<td>0.22%</td>
<td>4851</td>
<td>6899</td>
<td>4400</td>
<td>9.30%</td>
<td>36.22%</td>
</tr>
<tr>
<td>MOD#2</td>
<td>642</td>
<td>552</td>
<td>464</td>
<td>462</td>
<td>16.30%</td>
<td>0.43%</td>
<td>4765</td>
<td>7094</td>
<td>4351</td>
<td>8.69%</td>
<td>38.67%</td>
</tr>
<tr>
<td>MOD#3</td>
<td>642</td>
<td>551</td>
<td>442</td>
<td>433</td>
<td>21.42%</td>
<td>2.04%</td>
<td>4505</td>
<td>7481</td>
<td>4240</td>
<td>5.88%</td>
<td>43.32%</td>
</tr>
<tr>
<td>total/average</td>
<td>4648</td>
<td>3888</td>
<td>2241</td>
<td>1949</td>
<td>49.87%</td>
<td>13.03%</td>
<td>34083</td>
<td>36809</td>
<td>23422</td>
<td>31.28%</td>
<td>36.37%</td>
</tr>
</tbody>
</table>

Table 4: Experimental results of reuse across runs (program BinTree)

<table>
<thead>
<tr>
<th>Changes</th>
<th>$n_0$</th>
<th>$n_1$</th>
<th>$n_2$</th>
<th>$n_3$</th>
<th>$R'$</th>
<th>$R''$</th>
<th>$t_1$(ms)</th>
<th>$t_2$(ms)</th>
<th>$t_3$(ms)</th>
<th>$T'$</th>
<th>$T''$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADD#1</td>
<td>5930</td>
<td>1689</td>
<td>803</td>
<td>746</td>
<td>55.83%</td>
<td>7.10%</td>
<td>17978</td>
<td>20355</td>
<td>11889</td>
<td>33.87%</td>
<td>41.59%</td>
</tr>
<tr>
<td>ADD#2</td>
<td>13358</td>
<td>3938</td>
<td>2618</td>
<td>2556</td>
<td>35.09%</td>
<td>2.37%</td>
<td>35382</td>
<td>105190</td>
<td>32465</td>
<td>8.24%</td>
<td>69.14%</td>
</tr>
<tr>
<td>ADD#3</td>
<td>15602</td>
<td>540</td>
<td>0</td>
<td>0</td>
<td>100.00%</td>
<td>0/0</td>
<td>18106</td>
<td>61586</td>
<td>17180</td>
<td>5.11%</td>
<td>72.10%</td>
</tr>
<tr>
<td>DEL#1</td>
<td>13358</td>
<td>3149</td>
<td>2216</td>
<td>2185</td>
<td>30.61%</td>
<td>1.40%</td>
<td>32134</td>
<td>126488</td>
<td>31002</td>
<td>3.52%</td>
<td>75.49%</td>
</tr>
<tr>
<td>DEL#2</td>
<td>5930</td>
<td>1154</td>
<td>599</td>
<td>0</td>
<td>100.00%</td>
<td>100.00%</td>
<td>13565</td>
<td>44789</td>
<td>10932</td>
<td>19.41%</td>
<td>75.59%</td>
</tr>
<tr>
<td>DEL#3</td>
<td>3252</td>
<td>1682</td>
<td>0</td>
<td>0</td>
<td>100.00%</td>
<td>0/0</td>
<td>12945</td>
<td>11482</td>
<td>4505</td>
<td>65.20%</td>
<td>60.76%</td>
</tr>
<tr>
<td>MOD#1</td>
<td>3252</td>
<td>1682</td>
<td>1080</td>
<td>1002</td>
<td>40.43%</td>
<td>7.22%</td>
<td>14533</td>
<td>16297</td>
<td>10628</td>
<td>26.97%</td>
<td>34.79%</td>
</tr>
<tr>
<td>MOD#2</td>
<td>3252</td>
<td>1680</td>
<td>716</td>
<td>632</td>
<td>62.38%</td>
<td>11.73%</td>
<td>14147</td>
<td>13784</td>
<td>7953</td>
<td>43.78%</td>
<td>42.30%</td>
</tr>
<tr>
<td>MOD#3</td>
<td>8310</td>
<td>2377</td>
<td>1068</td>
<td>964</td>
<td>59.44%</td>
<td>9.74%</td>
<td>22772</td>
<td>32889</td>
<td>14593</td>
<td>35.92%</td>
<td>55.63%</td>
</tr>
<tr>
<td>total/average</td>
<td>72244</td>
<td>17891</td>
<td>9100</td>
<td>8085</td>
<td>54.81%</td>
<td>11.15%</td>
<td>181582</td>
<td>432860</td>
<td>141147</td>
<td>22.27%</td>
<td>67.39%</td>
</tr>
</tbody>
</table>

GreenTrie gets better reuse ratio and saves more time than both Green and KLEE’s approach.
## Evaluation

- **Reuse across Runs**

<table>
<thead>
<tr>
<th>Changes</th>
<th>$n_0$</th>
<th>$n_1$</th>
<th>$n_2$</th>
<th>$n_3$</th>
<th>$R'$</th>
<th>$R''$</th>
<th>$t_1$(ms)</th>
<th>$t_2$(ms)</th>
<th>$t_3$(ms)</th>
<th>$T'$</th>
<th>$T''$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADD#1</td>
<td>5930</td>
<td>1689</td>
<td>803</td>
<td>746</td>
<td>55.83%</td>
<td>7.10%</td>
<td>17978</td>
<td>20355</td>
<td>11889</td>
<td>33.87%</td>
<td>41.59%</td>
</tr>
<tr>
<td>ADD#2</td>
<td>13358</td>
<td>3938</td>
<td>2618</td>
<td>2556</td>
<td>35.09%</td>
<td>2.37%</td>
<td>35382</td>
<td>105190</td>
<td>32465</td>
<td>8.24%</td>
<td>69.14%</td>
</tr>
<tr>
<td>ADD#3</td>
<td>15602</td>
<td>540</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100.00%</td>
<td>18106</td>
<td>61586</td>
<td>17180</td>
<td>5.11%</td>
<td>72.10%</td>
</tr>
<tr>
<td>DEL#1</td>
<td>13358</td>
<td>3149</td>
<td>2216</td>
<td>2185</td>
<td>30.61%</td>
<td>1.40%</td>
<td>32134</td>
<td>126488</td>
<td>31002</td>
<td>3.52%</td>
<td>75.49%</td>
</tr>
<tr>
<td>DEL#2</td>
<td>5930</td>
<td>1154</td>
<td>599</td>
<td>0</td>
<td>100.00%</td>
<td>100.00%</td>
<td>13565</td>
<td>44789</td>
<td>10932</td>
<td>19.41%</td>
<td>75.59%</td>
</tr>
<tr>
<td>DEL#3</td>
<td>3252</td>
<td>1682</td>
<td>0</td>
<td>0</td>
<td>100.00%</td>
<td>0/0</td>
<td>12945</td>
<td>11482</td>
<td>4505</td>
<td>65.20%</td>
<td>60.76%</td>
</tr>
<tr>
<td>MOD#1</td>
<td>3252</td>
<td>1682</td>
<td>1080</td>
<td>1002</td>
<td>40.43%</td>
<td>7.22%</td>
<td>14553</td>
<td>16297</td>
<td>10628</td>
<td>26.97%</td>
<td>34.79%</td>
</tr>
<tr>
<td>MOD#2</td>
<td>3252</td>
<td>1680</td>
<td>716</td>
<td>632</td>
<td>62.38%</td>
<td>11.73%</td>
<td>14147</td>
<td>13784</td>
<td>7953</td>
<td>43.78%</td>
<td>42.30%</td>
</tr>
<tr>
<td>MOD#3</td>
<td>8310</td>
<td>2377</td>
<td>1068</td>
<td>964</td>
<td>59.44%</td>
<td>9.74%</td>
<td>22772</td>
<td>32889</td>
<td>14593</td>
<td>35.92%</td>
<td>55.63%</td>
</tr>
<tr>
<td>total/average</td>
<td>72244</td>
<td>17891</td>
<td>9100</td>
<td>8085</td>
<td>54.81%</td>
<td>11.15%</td>
<td>181582</td>
<td>432860</td>
<td>141147</td>
<td>22.27%</td>
<td>67.39%</td>
</tr>
</tbody>
</table>

Table 4: Experimental results of reuse across runs (program TCAS)

Table 3: Experimental results of reuse across runs (program Euclid)

- **3421 constraints in store**

- **Running time increases dramatically in KLEE’s approach**

- **GreenTrie gains better scalability than KLEE’s approach**
Evaluation

reuse across programs

Numbers of reused constraints for Green, KLEE approach and GreenTrie

<table>
<thead>
<tr>
<th>Program</th>
<th>Trityp</th>
<th>Euclid</th>
<th>TCAS</th>
<th>TreeMap</th>
<th>BinTree</th>
<th>BinomialHeap</th>
<th>Merarbiter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trityp</td>
<td>/</td>
<td>0, 0, 3</td>
<td>0, 0, 3</td>
<td>4, 4</td>
<td>0, 2, 2</td>
<td>0, 6, 7</td>
<td>0, 0, 1</td>
</tr>
<tr>
<td>Euclid</td>
<td>0, 0, 1</td>
<td>/</td>
<td>2, 5, 5</td>
<td>0, 0</td>
<td>0, 3, 4</td>
<td>0, 2, 2</td>
<td>0, 0, 2</td>
</tr>
<tr>
<td>TCAS</td>
<td>0, 0, 2</td>
<td>2, 2, 2</td>
<td>/</td>
<td>0, 0, 0</td>
<td>0, 2, 3</td>
<td>0, 3, 4</td>
<td>0, 3, 4</td>
</tr>
<tr>
<td>TreeMap</td>
<td>0, 0, 0</td>
<td>0, 0, 0</td>
<td>0, 0, 0</td>
<td>/</td>
<td>256, 449, 470</td>
<td>0, 0, 0</td>
<td>0, 0, 0</td>
</tr>
<tr>
<td>BinTree</td>
<td>0, 0, 0</td>
<td>0, 0, 0</td>
<td>0, 0, 0</td>
<td>256, 449, 470</td>
<td>1, 11, 10</td>
<td>/</td>
<td>0, 0, 0</td>
</tr>
<tr>
<td>BinomialHeap</td>
<td>2, 2, 5</td>
<td>2, 2, 5</td>
<td>2, 8, 6</td>
<td>0, 2, 3</td>
<td>1, 11, 10</td>
<td>/</td>
<td>0, 0, 0</td>
</tr>
<tr>
<td>Merarbiter</td>
<td>0, 1, 2</td>
<td>0, 2</td>
<td>0, 3, 0</td>
<td>0, 0, 0</td>
<td>0, 0, 0</td>
<td>0, 0, 0</td>
<td>/</td>
</tr>
</tbody>
</table>

GreenTrie achieves more inter-programs reuse than Green.
In some cases, GreenTrie has a little less reuse than KLEE’s approach.
The reason is that some constraints, which reuse the solution both across programs and in same program in GreenTrie, can only reuse constraints across programs in KLEE. Such constraints are counted for KLEE but not counted for GreenTrie.
Conclusion and Future Work

❖ Contributions

❖ Logical basis of implication-based reuse
❖ Trie-based store indexed with implication partial order graph
❖ Efficient logical subset/superset checking algorithms

❖ Future works

❖ Support more kinds of constraints other than linear integer constraints
❖ Reuse constraints which contains summaries
❖ Improve scalability for large-scale programs
Apologize

- I am sorry that I cannot answer your question face to face.
- If you have any question, please contact me with this email:

  jxy@whu.edu.cn