Comparing Verification Condition Generation with Symbolic Execution

... with an Introduction to Syxc

Z3 SIG Meeting
2nd to 4th November 2011, Cambridge

Malte Schwerhoff, ETH Zürich
Joint work with Yannis Kassios, Peter Müller
Outline

1. Background
2. Verifiers
3. Experiment
Background
Chalice

- Microsoft research language by Rustan Leino and others
- User provided specifications
  - Pre- and postconditions, loop invariants
  - *Pure functions* for information hiding (getters) and specification reasons
  - *Abstract predicates* for information hiding and to specify recursive structures
- Reasoning about side effects with *implicit dynamic frames*
- Fork-join concurrency with static dead-lock detection
- Shared data due to *fractional permissions*, monitors and locks
- Asynchronous communication via message passing channels
Verification condition generation vs. symbolic execution

→ Query prover once per method with all available information

→ Query prover often with limited information
Verifiers

(well, mainly one)
VCG-based Chalice verifier: Chalice

- Encodes a Chalice program in the *intermediate verification language* Boogie
- Heaps and permissions are encoded in Boogie as updateable maps
- Boogie verifier computes weakest preconditions and uses Z3 to discharge proof obligations
- User functions are axiomatised, i.e. for all heaps, permissions maps, and arguments, a function application is equivalent to evaluating its body w.r.t. the given heap, permissions and arguments
- Chalice is implemented in Scala
class Cell {
    var x: int
    
    method set(y: int)
        requires acc(x)
        ensures acc(x) && get() == y
        { x := y }

    function get(): int
        requires rd(x)
        { x }
}
Chalice: Example verification

```java
class Cell {
    var x: int

    method set(y: int)
        requires acc(x)
        ensures acc(x) && get() == y
        { x := y }

    function get(): int
        requires rd(x)
        { x }
}

method client(c: Cell, y: int)
    requires acc(c.x)
    ensures acc(c.x) && c.get() == y + 1
    { // Π(c,x)=100 ⇒ (Π(c,x)=100 ∧ ∀ u • (Π(c,x)=100 ∧ Cell.get(Η[(c,x)↦u],Π,c)=y
        // ⇒ (Π(c,x)=100 ∧ Cell.get(Η[(c,x)↦Η(c,x)+1][(c,x)↦u],Π,c)=y+1)))

        // Π(c,x)=100 ∧ (∀ u • (Π(c,x)=100 ∧ Cell.get(Η[(c,x)↦u],Π,c)=y
        // ⇒ (Π(c,x)=100 ∧ Cell.get(Η[(c,x)↦Η(c,x)+1][(c,x)↦u],Π,c)=y+1)))
        call c.set(y)

        // Π(c,x)=100 ∧ Cell.get(Η[(c,x)↦Η(c,x)+1],Π,c)=y+1
c.x := c.x + 1

        // Π(c,x)=100 ∧ Cell.get(Η,Π,c)=y+1
    }
```

- Weakest precondition is computed bottom-up, starting with the user-provided postcondition
- Weakest precondition must be implied by the user-provided precondition
SE-based Chalice verifier: Syxc

- VeriCool-style symbolic execution with a *symbolic state* comprising:
  - A *store* $\gamma$ mapping local variables to symbolic values (*terms*)
  - Symbolic *heaps* $h$ and $g$ (old heap) consisting of *heap chunks*, i.e. quadruples
    \[ r.f \mapsto v \# p \]
    representing
    “field $r.f$ has the value $v$ and we have $p$ permissions to the field”
  - *Path conditions* $\pi$ with assumptions such as $v > 0$, $r \neq \text{null}$
- Syxc is written in Scala
• Store and heap are maintained by Syxc, i.e. as Scala data structures, path conditions are managed by Z3’s push-pop scopes:
  - Knowledge is partitioned, Z3 does not know about heaps and permissions
  - Idea: Relieving Z3 from reasoning about the heap increases performance

• Functions are not axiomatised
  - Equality between function application and body evaluation is added when a function application is evaluated
  - Necessary, since heap and path conditions are separated, but functions are still framed by their heap footprint
Syxc: Example verification

```java
class Cell {
    var x: int
    method set(y: int)
        requires acc(x)
        ensures acc(x) && get() == y
    {
        x := y
    }

    method client(c: Cell, y: int)
        requires acc(c.x)
    {
        // γ: {c ↦ tc, y ↦ ty}
        // h: {tc.x ↦ tx # 100}
        // π: {}
        call c.set(y)
        // h: {tc.x ↦ tx' # 100}
        // π: {Cell.get(tc, tx') == ty ∧ Cell.get(tc, tx') == tx'}
        c.x := c.x + 1
        // h: {tc.x ↦ tx' + 1 # 100}
        assert acc(c.x) && c.get() == y + 1
        // π: {Cell.get(tc, tx') == ty ∧ Cell.get(tc, tx') == tx, // ∧ Cell.get(tc, tx' + 1) == tx' + 1}
    }
}
```

- Symbolic execution proceeds top-down
- Z3 is invoked with the current π to discharge proof obligations arising from preconditions and assertions
· Separation between heap and path conditions has consequences

```java
method client(c1: Cell, c2: Cell, b: bool)
  requires rd(c1.x) && rd(c2.x)
  { 
    if (c1 == c2) {
      assert c1.x == c2.x
    }
  }
```
Syxc: State separation consequence I

- Separation between heap and path conditions has consequences

```java
method client(c1: Cell, c2: Cell, b: bool)
    requires rd(c1.x) && rd(c2.x)
{
    if (c1 == c2) {
        assert c1.x == c2.x
    }
}
```

```
exec
if (c1 == c2) {
    assert c1.x == c2.x
}
```
Syxc: State separation consequence I

- Separation between heap and path conditions has consequences

```java
method client(c1: Cell, c2: Cell, b: bool)
    requires rd(c1.x) && rd(c2.x)
{
    if (c1 == c2) {
        assert c1.x == c2.x
        // Would fail naively because
        // tx1 = tx2 is unknown to Z3
    }
}
```

- Solution: *Heap compression*, that is, modifying the heap according to the current path conditions
Syxc: State separation consequence I

- Separation between heap and path conditions has consequences

```java
method client(c1: Cell, c2: Cell, b: bool)
    requires rd(c1.x) && rd(c2.x)
{
    if (c1 == c2) {
        assert c1.x == c2.x
            // Holds
    }
}
```

- For all pairs \( t, t' \) of receivers and all fields \( f \) in \( h \), invoke Z3 to check if \( t = t' \); if so, merge chunks

- Expensive operation: \( \mathcal{O}(n^2) \)

- Worse, newly added equalities require iterative proceeding: \( \mathcal{O}(n^3) \)
Syxc: State separation consequence II

- Separation between heap and path conditions has consequences

```java
method client(c1: Cell, c2: Cell)
requires acc(c1.x, 50) && acc(c2.x, 60)
{
    assert c1 != c2
}
```
Syxc: State separation consequence II

- Separation between heap and path conditions has consequences

```java
method client(c1: Cell, c2: Cell)
    requires acc(c1.x, 50) && acc(c2.x, 60)
{
    assert c1 != c2
    // Would fail naively because Z3 does not know about permissions (at most 100)
}
```

- Solution: Compute *object disjointness* based on field names and permissions:
  - For all pairs of chunks with receivers \( t, t' \) and the same field id \( f \), invoke Z3 to check if the combined permissions exceed 100
  - If so, add assumption \( t \neq t' \) to the path conditions
  - Expensive operation: \( \mathcal{O}(n^2) \)
Syxc: State separation consequence II

- Separation between heap and path conditions has consequences

```java
method client(c1: Cell, c2: Cell)
  requires acc(c1.x, 50) && acc(c2.x, 60)
{
  assert c1 != c2
    // Holds
}
```

- Solution: Compute *object disjointness* based on field names and permissions:
  - For all pairs of chunks with receivers \( t, t' \) and the same field id \( f \), invoke Z3 to check if the combined permissions exceed 100
  - If so, add assumption \( t \neq t' \) to the path conditions
  - Expensive operation: \( \mathcal{O}(n^2) \)
Syxc: Branching

- Symbolic execution branches on
  - If-then-else statements
  - If-then-else expressions (ternary operator) and implications
  - → possible explosion of the number of paths

```java
method client(c: Cell, b: bool)
  requires b ==> acc(c.x)
  { ... }
```

- Two execution paths with different heaps (and path conditions)
  1. Heap \( h \) contains chunk \( tc.x \mapsto tx \# 100 \)
  2. Heap \( h \) does not contain such a chunk

- Possible optimisation: *conditional chunks*
  - Single path with chunk \( tc.x \mapsto tx \# (tb ?100 : 0) \)
  - Prototypical implementation looked promising, but no systematic benchmarks have been run
Syxc: Branching

```java
method client(b: bool, i: int)
    requires b => i > 0
{ ... }
```

- Pure expression, i.e. it has no effect on the heap
  - Caused heavy branching in an impl. of Peterson’s Algorithm (4x slower)
- Since the expression is pure
  - Encode it as a Z3 implication \( tb \Rightarrow ti > 0 \)
  - Add it to \( \pi \), continue without branching
- Decreased verification time significantly; interpretations
  - Z3 is heavily optimised and branches faster than Syxc?
  - Additional knowledge allows Z3 to prune branches? (our assumption)
Experiment
Setup

- Benchmarked 22 tests from the Chalice test suite
- Tests existed already before Syxc had been developed
- Tests exercise main the Chalice features such as fractional permissions, objects, threads, locks, and message passing

- Syxc uses Chalice (as a library) to parse input into its AST
- Z3 3.1, SMTLib2 frontend via std-io, interaction is logged in a file
- We ensured that Syxc and Chalice use Z3 in comparable settings, for example by using nearly identical Z3 configurations

- Averaging statistics over ten runs per tool and test case:
  - Verify each test case and record the runtime
  - Run Z3 on the created SMTLib2 logfile to get Z3 statistics
## Results

<table>
<thead>
<tr>
<th>File</th>
<th>LOC</th>
<th>Meth.</th>
<th>Time (s)</th>
<th>QI</th>
<th>Conflicts</th>
<th>Time (s)</th>
<th>QI</th>
<th>Conflicts</th>
<th>Time (%)</th>
<th>QI (%)</th>
<th>Conflicts (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVLTree.iterative</td>
<td>212</td>
<td>3</td>
<td>2.24</td>
<td>231</td>
<td>177</td>
<td>9.45</td>
<td>146238</td>
<td>1288</td>
<td>23.70</td>
<td>0.16</td>
<td>13.72</td>
</tr>
<tr>
<td>AVLTree.nokeys</td>
<td>572</td>
<td>19</td>
<td>34.21</td>
<td>2357</td>
<td>4304</td>
<td>154.03</td>
<td>2541875</td>
<td>19593</td>
<td>22.21</td>
<td>0.09</td>
<td>21.97</td>
</tr>
<tr>
<td>Cell</td>
<td>131</td>
<td>11</td>
<td>1.78</td>
<td>233</td>
<td>40</td>
<td>3.23</td>
<td>9599</td>
<td>526</td>
<td>55.11</td>
<td>2.43</td>
<td>7.60</td>
</tr>
<tr>
<td>CopyLessMessagePassing</td>
<td>54</td>
<td>3</td>
<td>1.46</td>
<td>218</td>
<td>35</td>
<td>2.72</td>
<td>4769</td>
<td>210</td>
<td>53.68</td>
<td>4.57</td>
<td>16.67</td>
</tr>
<tr>
<td>CopyLessMessagePassing-w/a</td>
<td>62</td>
<td>3</td>
<td>1.71</td>
<td>751</td>
<td>75</td>
<td>2.76</td>
<td>4014</td>
<td>181</td>
<td>61.96</td>
<td>18.71</td>
<td>41.44</td>
</tr>
<tr>
<td>CopyLessMessagePassing-w/a2</td>
<td>66</td>
<td>3</td>
<td>1.66</td>
<td>792</td>
<td>78</td>
<td>2.92</td>
<td>10542</td>
<td>325</td>
<td>56.85</td>
<td>7.51</td>
<td>24.00</td>
</tr>
<tr>
<td>Dining-philosophers</td>
<td>74</td>
<td>3</td>
<td>1.62</td>
<td>1181</td>
<td>128</td>
<td>3.09</td>
<td>17144</td>
<td>416</td>
<td>52.43</td>
<td>6.89</td>
<td>30.82</td>
</tr>
<tr>
<td>Iterator</td>
<td>123</td>
<td>6</td>
<td>3.65</td>
<td>436</td>
<td>210</td>
<td>3.27</td>
<td>7577</td>
<td>316</td>
<td>111.62</td>
<td>5.75</td>
<td>66.39</td>
</tr>
<tr>
<td>Iterator2</td>
<td>111</td>
<td>6</td>
<td>1.69</td>
<td>174</td>
<td>41</td>
<td>3.17</td>
<td>12867</td>
<td>304</td>
<td>53.31</td>
<td>1.35</td>
<td>13.49</td>
</tr>
<tr>
<td>LoopLockChange</td>
<td>69</td>
<td>5</td>
<td>1.26</td>
<td>220</td>
<td>42</td>
<td>2.60</td>
<td>2624</td>
<td>207</td>
<td>48.46</td>
<td>8.38</td>
<td>20.29</td>
</tr>
<tr>
<td>OwickiGries</td>
<td>31</td>
<td>2</td>
<td>1.21</td>
<td>121</td>
<td>41</td>
<td>2.61</td>
<td>5696</td>
<td>206</td>
<td>46.36</td>
<td>2.12</td>
<td>19.90</td>
</tr>
<tr>
<td>PetersonsAlgorithm</td>
<td>70</td>
<td>3</td>
<td>2.12</td>
<td>1196</td>
<td>278</td>
<td>7.36</td>
<td>160574</td>
<td>1500</td>
<td>28.80</td>
<td>0.74</td>
<td>18.53</td>
</tr>
<tr>
<td>ProdConsChannel</td>
<td>78</td>
<td>6</td>
<td>1.43</td>
<td>282</td>
<td>117</td>
<td>2.78</td>
<td>4640</td>
<td>316</td>
<td>51.44</td>
<td>6.08</td>
<td>37.03</td>
</tr>
<tr>
<td>Producer-consumer</td>
<td>171</td>
<td>11</td>
<td>1.91</td>
<td>466</td>
<td>147</td>
<td>3.84</td>
<td>17667</td>
<td>598</td>
<td>49.74</td>
<td>2.64</td>
<td>24.58</td>
</tr>
<tr>
<td>Quantifiers</td>
<td>28</td>
<td>1</td>
<td>1.02</td>
<td>16</td>
<td>4</td>
<td>2.23</td>
<td>196</td>
<td>12</td>
<td>45.74</td>
<td>8.16</td>
<td>33.33</td>
</tr>
<tr>
<td>RockBand</td>
<td>109</td>
<td>13</td>
<td>1.30</td>
<td>79</td>
<td>14</td>
<td>2.92</td>
<td>9129</td>
<td>341</td>
<td>44.52</td>
<td>0.87</td>
<td>4.11</td>
</tr>
<tr>
<td>Sieve</td>
<td>56</td>
<td>4</td>
<td>1.42</td>
<td>308</td>
<td>93</td>
<td>2.74</td>
<td>7596</td>
<td>246</td>
<td>51.82</td>
<td>4.05</td>
<td>37.80</td>
</tr>
<tr>
<td>Swap</td>
<td>19</td>
<td>2</td>
<td>0.98</td>
<td>10</td>
<td>2</td>
<td>2.20</td>
<td>261</td>
<td>20</td>
<td>44.55</td>
<td>3.83</td>
<td>10.00</td>
</tr>
<tr>
<td>Prog1</td>
<td>33</td>
<td>4</td>
<td>1.07</td>
<td>40</td>
<td>27</td>
<td>2.32</td>
<td>1101</td>
<td>106</td>
<td>46.12</td>
<td>3.63</td>
<td>25.47</td>
</tr>
<tr>
<td>Prog2</td>
<td>52</td>
<td>7</td>
<td>1.11</td>
<td>22</td>
<td>17</td>
<td>2.31</td>
<td>258</td>
<td>28</td>
<td>48.05</td>
<td>8.53</td>
<td>60.71</td>
</tr>
<tr>
<td>Prog3</td>
<td>163</td>
<td>16</td>
<td>1.58</td>
<td>627</td>
<td>144</td>
<td>3.53</td>
<td>21632</td>
<td>640</td>
<td>44.76</td>
<td>2.90</td>
<td>22.50</td>
</tr>
<tr>
<td>Prog4</td>
<td>19</td>
<td>1</td>
<td>1.07</td>
<td>76</td>
<td>22</td>
<td>2.32</td>
<td>877</td>
<td>61</td>
<td>15.12</td>
<td>1.47</td>
<td>17.77</td>
</tr>
</tbody>
</table>

LOC = Lines of code  
Meth. = Methods  
Time = Verification time (sec)  
QI = Quantifier instantiations  
Conflicts ≅ Explored search space
### Results

<table>
<thead>
<tr>
<th>Name</th>
<th>LOC</th>
<th>Meth.</th>
<th>Time (s)</th>
<th>QI</th>
<th>Conflicts</th>
<th>Time (s)</th>
<th>QI</th>
<th>Conflicts</th>
<th>Time (%)</th>
<th>QI (%)</th>
<th>Conflicts (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVLTree.iterative</td>
<td>212</td>
<td>3</td>
<td>2.24</td>
<td>231</td>
<td>177</td>
<td>9.45</td>
<td>146238</td>
<td>1288</td>
<td>23.70</td>
<td>0.16</td>
<td>13.72</td>
</tr>
<tr>
<td>AVLTree.nokeys</td>
<td>572</td>
<td>19</td>
<td>34.21</td>
<td>2357</td>
<td>4304</td>
<td>154.03</td>
<td>2541875</td>
<td>19593</td>
<td>22.21</td>
<td>0.09</td>
<td>21.97</td>
</tr>
<tr>
<td>Cell</td>
<td>131</td>
<td>11</td>
<td>1.78</td>
<td>233</td>
<td>40</td>
<td>3.23</td>
<td>9599</td>
<td>526</td>
<td>55.11</td>
<td>2.43</td>
<td>7.60</td>
</tr>
<tr>
<td>CopyLessMessagePassing</td>
<td>54</td>
<td>3</td>
<td>1.46</td>
<td>218</td>
<td>35</td>
<td>2.72</td>
<td>4769</td>
<td>210</td>
<td>53.68</td>
<td>4.57</td>
<td>16.67</td>
</tr>
<tr>
<td>CopyLessMessagePassing-w/a</td>
<td>62</td>
<td>3</td>
<td>1.71</td>
<td>751</td>
<td>75</td>
<td>2.76</td>
<td>4014</td>
<td>181</td>
<td>61.96</td>
<td>18.71</td>
<td>41.44</td>
</tr>
<tr>
<td>CopyLessMessagePassing-w/a2</td>
<td>66</td>
<td>3</td>
<td>1.66</td>
<td>792</td>
<td>78</td>
<td>2.92</td>
<td>10542</td>
<td>325</td>
<td>56.85</td>
<td>7.51</td>
<td>24.00</td>
</tr>
<tr>
<td>Dining-philosophers</td>
<td>74</td>
<td>3</td>
<td>1.62</td>
<td>1181</td>
<td>128</td>
<td>3.09</td>
<td>17144</td>
<td>416</td>
<td>52.43</td>
<td>6.89</td>
<td>30.82</td>
</tr>
<tr>
<td>Iterator</td>
<td>123</td>
<td>6</td>
<td>3.65</td>
<td>436</td>
<td>210</td>
<td>3.27</td>
<td>7577</td>
<td>316</td>
<td>111.62</td>
<td>5.75</td>
<td>66.39</td>
</tr>
<tr>
<td>Iterator2</td>
<td>111</td>
<td>6</td>
<td>1.69</td>
<td>174</td>
<td>41</td>
<td>3.17</td>
<td>12867</td>
<td>304</td>
<td>53.31</td>
<td>1.35</td>
<td>13.49</td>
</tr>
<tr>
<td>LoopLockChange</td>
<td>69</td>
<td>5</td>
<td>1.26</td>
<td>220</td>
<td>42</td>
<td>2.60</td>
<td>2624</td>
<td>207</td>
<td>48.46</td>
<td>8.38</td>
<td>20.29</td>
</tr>
<tr>
<td>OwickiGries</td>
<td>31</td>
<td>2</td>
<td>1.21</td>
<td>121</td>
<td>41</td>
<td>2.61</td>
<td>5696</td>
<td>206</td>
<td>46.36</td>
<td>2.12</td>
<td>19.90</td>
</tr>
<tr>
<td>PetersonsAlgorithm</td>
<td>70</td>
<td>3</td>
<td>2.12</td>
<td>1196</td>
<td>278</td>
<td>7.36</td>
<td>160574</td>
<td>1500</td>
<td>28.80</td>
<td>0.74</td>
<td>18.53</td>
</tr>
<tr>
<td>ProdConsChannel</td>
<td>78</td>
<td>6</td>
<td>1.43</td>
<td>282</td>
<td>117</td>
<td>2.78</td>
<td>4640</td>
<td>316</td>
<td>51.44</td>
<td>6.08</td>
<td>37.03</td>
</tr>
<tr>
<td>Producer-consumer</td>
<td>171</td>
<td>11</td>
<td>1.91</td>
<td>466</td>
<td>147</td>
<td>3.84</td>
<td>17667</td>
<td>598</td>
<td>49.74</td>
<td>2.64</td>
<td>24.58</td>
</tr>
<tr>
<td>Quantifiers</td>
<td>28</td>
<td>1</td>
<td>1.02</td>
<td>16</td>
<td>4</td>
<td>2.23</td>
<td>196</td>
<td>12</td>
<td>45.74</td>
<td>8.16</td>
<td>33.33</td>
</tr>
<tr>
<td>RockBand</td>
<td>109</td>
<td>13</td>
<td>1.30</td>
<td>79</td>
<td>14</td>
<td>2.92</td>
<td>9129</td>
<td>341</td>
<td>44.52</td>
<td>0.87</td>
<td>4.11</td>
</tr>
<tr>
<td>Sieve</td>
<td>56</td>
<td>4</td>
<td>1.42</td>
<td>308</td>
<td>93</td>
<td>2.74</td>
<td>7596</td>
<td>246</td>
<td>51.82</td>
<td>4.05</td>
<td>37.80</td>
</tr>
<tr>
<td>Swap</td>
<td>19</td>
<td>2</td>
<td>0.98</td>
<td>10</td>
<td>2</td>
<td>2.20</td>
<td>261</td>
<td>20</td>
<td>44.55</td>
<td>3.83</td>
<td>10.00</td>
</tr>
<tr>
<td>Prog1</td>
<td>33</td>
<td>4</td>
<td>1.07</td>
<td>40</td>
<td>27</td>
<td>2.32</td>
<td>1101</td>
<td>106</td>
<td>46.12</td>
<td>3.63</td>
<td>25.47</td>
</tr>
<tr>
<td>Prog2</td>
<td>52</td>
<td>7</td>
<td>1.11</td>
<td>22</td>
<td>17</td>
<td>2.31</td>
<td>258</td>
<td>28</td>
<td>48.05</td>
<td>8.53</td>
<td>60.71</td>
</tr>
<tr>
<td>Prog3</td>
<td>163</td>
<td>16</td>
<td>1.58</td>
<td>627</td>
<td>144</td>
<td>3.53</td>
<td>21632</td>
<td>640</td>
<td>44.76</td>
<td>2.90</td>
<td>22.50</td>
</tr>
<tr>
<td>Prog4</td>
<td>19</td>
<td>1</td>
<td>1.07</td>
<td>76</td>
<td>22</td>
<td>2.32</td>
<td>877</td>
<td>61</td>
<td>46.12</td>
<td>8.67</td>
<td>36.07</td>
</tr>
</tbody>
</table>

- **Longest method (iterative); longest example (nokeys)**
- **Best runtime (24%) and quantifier instantiation (0.2%)**
## Results

<table>
<thead>
<tr>
<th>File</th>
<th>LOC</th>
<th>Meth.</th>
<th>Time (s)</th>
<th>QI</th>
<th>Conflicts</th>
<th>Time (s)</th>
<th>QI</th>
<th>Conflicts</th>
<th>Time (%)</th>
<th>QI (%)</th>
<th>Conflicts (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVLTree.iterative</td>
<td>212</td>
<td>3</td>
<td>2.24</td>
<td>231</td>
<td>177</td>
<td>9.45</td>
<td>146238</td>
<td>1288</td>
<td>23.70</td>
<td>0.16</td>
<td>13.72</td>
</tr>
<tr>
<td>AVLTree.nokeys</td>
<td>572</td>
<td>19</td>
<td>34.21</td>
<td>2357</td>
<td>4304</td>
<td>154.03</td>
<td>2541875</td>
<td>19593</td>
<td>22.21</td>
<td>0.09</td>
<td>21.97</td>
</tr>
<tr>
<td>Cell</td>
<td>131</td>
<td>11</td>
<td>1.78</td>
<td>233</td>
<td>40</td>
<td>3.23</td>
<td>9599</td>
<td>526</td>
<td>55.11</td>
<td>2.43</td>
<td>7.60</td>
</tr>
<tr>
<td>CopyLessMessagePassing</td>
<td>54</td>
<td>3</td>
<td>1.46</td>
<td>218</td>
<td>35</td>
<td>2.72</td>
<td>4769</td>
<td>210</td>
<td>53.68</td>
<td>4.57</td>
<td>16.67</td>
</tr>
<tr>
<td>CopyLessMessagePassing-w/a</td>
<td>62</td>
<td>3</td>
<td>1.71</td>
<td>751</td>
<td>75</td>
<td>2.76</td>
<td>4014</td>
<td>181</td>
<td>61.96</td>
<td>18.71</td>
<td>41.44</td>
</tr>
<tr>
<td>CopyLessMessagePassing-w/a2</td>
<td>66</td>
<td>3</td>
<td>1.66</td>
<td>792</td>
<td>78</td>
<td>2.92</td>
<td>10542</td>
<td>325</td>
<td>56.85</td>
<td>7.51</td>
<td>24.00</td>
</tr>
<tr>
<td>Dining-philosophers</td>
<td>74</td>
<td>3</td>
<td>1.62</td>
<td>1181</td>
<td>128</td>
<td>3.09</td>
<td>17144</td>
<td>416</td>
<td>52.43</td>
<td>6.89</td>
<td>30.82</td>
</tr>
<tr>
<td>Iterator</td>
<td>123</td>
<td>6</td>
<td>3.65</td>
<td>436</td>
<td>210</td>
<td>3.27</td>
<td>7577</td>
<td>316</td>
<td>111.62</td>
<td>5.75</td>
<td>66.39</td>
</tr>
<tr>
<td>Iterator2</td>
<td>111</td>
<td>6</td>
<td>1.69</td>
<td>174</td>
<td>41</td>
<td>3.17</td>
<td>12867</td>
<td>304</td>
<td>53.31</td>
<td>1.35</td>
<td>13.49</td>
</tr>
<tr>
<td>LoopLockChange</td>
<td>69</td>
<td>5</td>
<td>1.26</td>
<td>220</td>
<td>42</td>
<td>2.60</td>
<td>2624</td>
<td>207</td>
<td>48.46</td>
<td>8.38</td>
<td>20.29</td>
</tr>
<tr>
<td>OwickiGries</td>
<td>31</td>
<td>2</td>
<td>1.21</td>
<td>121</td>
<td>41</td>
<td>2.61</td>
<td>5696</td>
<td>206</td>
<td>46.36</td>
<td>2.12</td>
<td>19.90</td>
</tr>
<tr>
<td>PetersonsAlgorithm</td>
<td>70</td>
<td>3</td>
<td>2.12</td>
<td>1196</td>
<td>278</td>
<td>7.36</td>
<td>160574</td>
<td>1500</td>
<td>28.80</td>
<td>0.74</td>
<td>18.53</td>
</tr>
<tr>
<td>ProdConsChannel</td>
<td>78</td>
<td>6</td>
<td>1.43</td>
<td>282</td>
<td>117</td>
<td>2.78</td>
<td>4640</td>
<td>316</td>
<td>51.44</td>
<td>6.08</td>
<td>37.03</td>
</tr>
<tr>
<td>Producer-consumer</td>
<td>171</td>
<td>11</td>
<td>1.91</td>
<td>466</td>
<td>147</td>
<td>3.84</td>
<td>17667</td>
<td>598</td>
<td>49.74</td>
<td>2.64</td>
<td>24.58</td>
</tr>
<tr>
<td>Quantifiers</td>
<td>28</td>
<td>1</td>
<td>1.02</td>
<td>16</td>
<td>4</td>
<td>2.23</td>
<td>196</td>
<td>12</td>
<td>45.74</td>
<td>8.16</td>
<td>33.33</td>
</tr>
<tr>
<td>RockBand</td>
<td>109</td>
<td>13</td>
<td>1.30</td>
<td>79</td>
<td>14</td>
<td>2.92</td>
<td>9129</td>
<td>341</td>
<td>44.52</td>
<td>0.87</td>
<td>4.11</td>
</tr>
<tr>
<td>Sieve</td>
<td>56</td>
<td>4</td>
<td>1.42</td>
<td>308</td>
<td>93</td>
<td>2.74</td>
<td>7596</td>
<td>246</td>
<td>51.82</td>
<td>4.05</td>
<td>37.80</td>
</tr>
<tr>
<td>Swap</td>
<td>19</td>
<td>2</td>
<td>0.98</td>
<td>10</td>
<td>2</td>
<td>2.20</td>
<td>261</td>
<td>20</td>
<td>44.55</td>
<td>3.83</td>
<td>10.00</td>
</tr>
<tr>
<td>Prog1</td>
<td>33</td>
<td>4</td>
<td>1.07</td>
<td>40</td>
<td>27</td>
<td>2.32</td>
<td>1101</td>
<td>106</td>
<td>46.12</td>
<td>3.63</td>
<td>25.47</td>
</tr>
<tr>
<td>Prog2</td>
<td>52</td>
<td>7</td>
<td>1.11</td>
<td>22</td>
<td>17</td>
<td>2.31</td>
<td>258</td>
<td>28</td>
<td>48.05</td>
<td>8.53</td>
<td>60.71</td>
</tr>
<tr>
<td>Prog3</td>
<td>163</td>
<td>16</td>
<td>1.58</td>
<td>627</td>
<td>144</td>
<td>3.53</td>
<td>21632</td>
<td>640</td>
<td>44.76</td>
<td>2.90</td>
<td>22.50</td>
</tr>
<tr>
<td>Prog4</td>
<td>19</td>
<td>1</td>
<td>1.07</td>
<td>76</td>
<td>22</td>
<td>2.32</td>
<td>877</td>
<td>61</td>
<td>46.12</td>
<td>8.67</td>
<td>36.07</td>
</tr>
</tbody>
</table>

- Similar behaviour for test case PetersonsAlgorithm: Runtime speedup above average, very low QI ratio
### Results

<table>
<thead>
<tr>
<th>File</th>
<th>LOC</th>
<th>Meth.</th>
<th>Time (s)</th>
<th>QI</th>
<th>Conflicts</th>
<th>Time (s)</th>
<th>QI</th>
<th>Conflicts</th>
<th>Time (%)</th>
<th>QI (%)</th>
<th>Conflicts (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVLTree.iterative</td>
<td>212</td>
<td>3</td>
<td>2.24</td>
<td>231</td>
<td>177</td>
<td>9.45</td>
<td>146238</td>
<td>1288</td>
<td>23.70</td>
<td>0.16</td>
<td>13.72</td>
</tr>
<tr>
<td>AVLTree.nokeys</td>
<td>572</td>
<td>19</td>
<td>34.21</td>
<td>2357</td>
<td>4304</td>
<td>154.03</td>
<td>2541875</td>
<td>19593</td>
<td>22.21</td>
<td>0.09</td>
<td>21.97</td>
</tr>
<tr>
<td>Cell</td>
<td>131</td>
<td>11</td>
<td>1.78</td>
<td>233</td>
<td>40</td>
<td>3.23</td>
<td>9599</td>
<td>526</td>
<td>55.11</td>
<td>2.43</td>
<td>7.60</td>
</tr>
<tr>
<td>CopyLessMessagePassing</td>
<td>54</td>
<td>3</td>
<td>1.46</td>
<td>218</td>
<td>35</td>
<td>2.72</td>
<td>4769</td>
<td>210</td>
<td>53.68</td>
<td>4.57</td>
<td>16.67</td>
</tr>
<tr>
<td>CopyLessMessagePassing-w/a</td>
<td>62</td>
<td>3</td>
<td>1.71</td>
<td>751</td>
<td>75</td>
<td>2.76</td>
<td>4014</td>
<td>181</td>
<td>61.96</td>
<td>18.71</td>
<td>41.44</td>
</tr>
<tr>
<td>CopyLessMessagePassing-w/a2</td>
<td>66</td>
<td>3</td>
<td>1.66</td>
<td>792</td>
<td>78</td>
<td>2.92</td>
<td>10542</td>
<td>325</td>
<td>56.85</td>
<td>7.51</td>
<td>24.00</td>
</tr>
<tr>
<td>Dining:philosophers</td>
<td>74</td>
<td>7</td>
<td>1.62</td>
<td>1181</td>
<td>128</td>
<td>3.00</td>
<td>17144</td>
<td>416</td>
<td>52.43</td>
<td>6.89</td>
<td>30.82</td>
</tr>
<tr>
<td>Iterator</td>
<td>123</td>
<td>6</td>
<td>3.65</td>
<td>436</td>
<td>210</td>
<td>3.27</td>
<td>7577</td>
<td>316</td>
<td>111.62</td>
<td>5.75</td>
<td>66.39</td>
</tr>
<tr>
<td>Iterator2</td>
<td>111</td>
<td>6</td>
<td>1.69</td>
<td>174</td>
<td>41</td>
<td>3.17</td>
<td>12867</td>
<td>304</td>
<td>53.31</td>
<td>1.35</td>
<td>13.49</td>
</tr>
<tr>
<td>LoopLockChange</td>
<td>69</td>
<td>5</td>
<td>1.26</td>
<td>220</td>
<td>42</td>
<td>2.60</td>
<td>2624</td>
<td>207</td>
<td>48.46</td>
<td>8.38</td>
<td>20.29</td>
</tr>
<tr>
<td>OwickiGries</td>
<td>31</td>
<td>2</td>
<td>1.21</td>
<td>121</td>
<td>41</td>
<td>2.61</td>
<td>5696</td>
<td>206</td>
<td>46.36</td>
<td>2.12</td>
<td>19.90</td>
</tr>
<tr>
<td>PetersonsAlgorithm</td>
<td>70</td>
<td>3</td>
<td>2.12</td>
<td>1196</td>
<td>278</td>
<td>7.36</td>
<td>160574</td>
<td>1500</td>
<td>28.80</td>
<td>0.74</td>
<td>18.53</td>
</tr>
<tr>
<td>ProdConsChannel</td>
<td>78</td>
<td>6</td>
<td>1.43</td>
<td>282</td>
<td>117</td>
<td>2.78</td>
<td>4640</td>
<td>316</td>
<td>51.44</td>
<td>6.08</td>
<td>37.03</td>
</tr>
<tr>
<td>Producer-consumer</td>
<td>171</td>
<td>11</td>
<td>1.91</td>
<td>466</td>
<td>147</td>
<td>3.84</td>
<td>17667</td>
<td>598</td>
<td>49.74</td>
<td>2.64</td>
<td>24.58</td>
</tr>
<tr>
<td>Quantifiers</td>
<td>28</td>
<td>1</td>
<td>1.02</td>
<td>16</td>
<td>4</td>
<td>2.23</td>
<td>196</td>
<td>12</td>
<td>45.74</td>
<td>8.16</td>
<td>33.33</td>
</tr>
<tr>
<td>RockBand</td>
<td>109</td>
<td>13</td>
<td>1.30</td>
<td>79</td>
<td>14</td>
<td>2.92</td>
<td>9129</td>
<td>341</td>
<td>44.52</td>
<td>0.87</td>
<td>4.11</td>
</tr>
<tr>
<td>Sieve</td>
<td>56</td>
<td>4</td>
<td>1.42</td>
<td>308</td>
<td>93</td>
<td>2.74</td>
<td>7596</td>
<td>246</td>
<td>51.82</td>
<td>4.05</td>
<td>37.80</td>
</tr>
<tr>
<td>Swap</td>
<td>19</td>
<td>2</td>
<td>0.98</td>
<td>10</td>
<td>2</td>
<td>2.20</td>
<td>261</td>
<td>20</td>
<td>44.55</td>
<td>3.83</td>
<td>10.00</td>
</tr>
<tr>
<td>Prog1</td>
<td>33</td>
<td>4</td>
<td>1.07</td>
<td>40</td>
<td>27</td>
<td>2.32</td>
<td>1101</td>
<td>106</td>
<td>46.12</td>
<td>3.63</td>
<td>25.47</td>
</tr>
<tr>
<td>Prog2</td>
<td>52</td>
<td>7</td>
<td>1.11</td>
<td>22</td>
<td>17</td>
<td>2.31</td>
<td>258</td>
<td>28</td>
<td>48.05</td>
<td>8.53</td>
<td>60.71</td>
</tr>
<tr>
<td>Prog3</td>
<td>163</td>
<td>16</td>
<td>1.58</td>
<td>627</td>
<td>144</td>
<td>3.53</td>
<td>21632</td>
<td>640</td>
<td>44.76</td>
<td>2.90</td>
<td>22.50</td>
</tr>
<tr>
<td>Prog4</td>
<td>19</td>
<td>1</td>
<td>1.07</td>
<td>76</td>
<td>22</td>
<td>2.32</td>
<td>877</td>
<td>61</td>
<td>46.12</td>
<td>8.67</td>
<td>36.07</td>
</tr>
</tbody>
</table>

- **Conflicts are less conclusive:**
  - **Iterator:** conflicts above average, average QI, negative speedup
  - **Prog2:** conflicts above average, QI even higher, average speedup
Conclusions

- Test suite is still rather small, results have to be taken with caution!
- Syxc requires in average only 5% quantifier instantiations → indicates that SE is more predictive than VCG
- Syxc causes in average only 25% conflicts, but with greater variation → indicates that SE-based verification is more focused
- We assume that the number of QI is significantly smaller due to state separation into heap and path conditions
  - Less information for the prover, less non-goal-directed work
  - Separation of labour between verifier and prover seems beneficial
- However, limited information might not always be beneficial, as hinted at by the branching problem
Future work

- Investigate outliers
- Evaluate different heap compression strategies
- Implement new permission model
- Implement and benchmark conditional chunks
- Parallelisation: Methods, branches, others?
- Responsiveness: Cache symbolic state to re-execute as little as possible?
- Much more difficult: a debugger on top of Syxc
  - How to present symbolic state?
  - Which information is relevant and helpful, which isn’t?
  - Evaluate different approaches
Questions?
Backup Slides
class Cell {
    var x: int

    predicate P { acc(x) }

    function get(): int
        requires rd(P)
    { unfolding rd(P) in x }

    method set(y: int)
        requires P
    ensures P && get() == y
    {
        unfold P
        x := y
        fold P
    }

    method clone() returns (c: Cell)
        requires rd(P)
    ensures rd(P) && c.P && c.get() == get()
    {
        c := new Cell
        fold c.P
        call c.set(get())
    }
}

class Client {
    method fails(c: Cell)
        requires c.P
    {
        fork tk1 := c.set(1)
        fork tk2 := c.set(2) // ERR
    }

    method succeeds(c: Cell)
        requires c.P
    {
        fork tk1 := c.clone()
        fork tk2 := c.clone()
    }
}
Syxc: Branching

```
method client(c: Cell, b: bool)
  requires b => acc(c.x)
{ ... }
```

- Possible optimisation: *conditional chunks*
  - Single path with chunk \( tc.x \mapsto tx \# (tb \oplus 100 : 0) \)
  - Disadvantage: Now every field dereferencing requires a Z3 invocation in order to check if we have non-zero permissions
- Prototypical implementation looked promising
  - verification time of massively branching programs dropped significantly
  - verification time of other programs increased slightly
Tools

- Latest Chalice version uses a new permission model not yet supported by Syxc, hence we had to use a slightly outdated Chalice version
- Syxc uses Chalice (as a library) to parse input into an AST
- Recent Boogie version; limited to one error per Chalice method
- Z3 3.1, smtlib2 frontend via std-io, interaction is logged in a file
- Syxc uses nearly the same Z3 settings as Chalice does, except
  - Syxc requires Z3 to respond to every command, not only to `check-sat`
  - Syxc uses global declarations, not scope-local ones
- Other differences:
  - Syxc encodes snapshots, references and lists as Z3 integers → might increase the number of quantifier instantiations
  - Syxc uses Boogie’s sequence axiomatisation, but they range over integers only, whereas Boogie’s are polymorphic → might increase the workload for Z3
### Results

<table>
<thead>
<tr>
<th>Name</th>
<th>LOC</th>
<th>Cell</th>
<th>Syxc</th>
<th>Chalice</th>
<th>Syxc rel. to Chalice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell</td>
<td>60</td>
<td>1.12</td>
<td>1.24</td>
<td>1.46</td>
<td>1.66</td>
</tr>
<tr>
<td>LoopLockChange</td>
<td>60</td>
<td>1.12</td>
<td>1.24</td>
<td>1.46</td>
<td>1.66</td>
</tr>
<tr>
<td>ProdConsChannel</td>
<td>60</td>
<td>1.12</td>
<td>1.24</td>
<td>1.46</td>
<td>1.66</td>
</tr>
<tr>
<td>Prog1</td>
<td>572</td>
<td>1.12</td>
<td>1.24</td>
<td>1.46</td>
<td>1.66</td>
</tr>
<tr>
<td>Prog2</td>
<td>60</td>
<td>1.12</td>
<td>1.24</td>
<td>1.46</td>
<td>1.66</td>
</tr>
<tr>
<td>Prog3</td>
<td>60</td>
<td>1.12</td>
<td>1.24</td>
<td>1.46</td>
<td>1.66</td>
</tr>
<tr>
<td>Prog4</td>
<td>60</td>
<td>1.12</td>
<td>1.24</td>
<td>1.46</td>
<td>1.66</td>
</tr>
<tr>
<td>AVLTree.iterative</td>
<td>256</td>
<td>2.24</td>
<td>2.41</td>
<td>2.51</td>
<td>2.61</td>
</tr>
<tr>
<td>AVLTree.nokeys</td>
<td>256</td>
<td>2.24</td>
<td>2.41</td>
<td>2.51</td>
<td>2.61</td>
</tr>
<tr>
<td>Cell</td>
<td>131</td>
<td>1.78</td>
<td>1.46</td>
<td>2.92</td>
<td>3.27</td>
</tr>
<tr>
<td>CopyLessMessagePassing</td>
<td>171</td>
<td>1.78</td>
<td>1.46</td>
<td>2.92</td>
<td>3.27</td>
</tr>
<tr>
<td>CopyLessMessagePassing-w/a</td>
<td>171</td>
<td>1.78</td>
<td>1.46</td>
<td>2.92</td>
<td>3.27</td>
</tr>
<tr>
<td>CopyLessMessagePassing-w/o</td>
<td>171</td>
<td>1.78</td>
<td>1.46</td>
<td>2.92</td>
<td>3.27</td>
</tr>
<tr>
<td>Dining-philosophers</td>
<td>3</td>
<td>1.62</td>
<td>1.51</td>
<td>1.91</td>
<td>2.23</td>
</tr>
<tr>
<td>Iterator</td>
<td>123</td>
<td>3.65</td>
<td>1.37</td>
<td>3.17</td>
<td>3.65</td>
</tr>
<tr>
<td>Iterator2</td>
<td>123</td>
<td>3.65</td>
<td>1.37</td>
<td>3.17</td>
<td>3.65</td>
</tr>
<tr>
<td>Prog1</td>
<td>60</td>
<td>1.26</td>
<td>1.69</td>
<td>1.69</td>
<td>1.69</td>
</tr>
<tr>
<td>OwickiGries</td>
<td>60</td>
<td>1.26</td>
<td>1.69</td>
<td>1.69</td>
<td>1.69</td>
</tr>
<tr>
<td>PetersonsAlgorithm</td>
<td>60</td>
<td>1.26</td>
<td>1.69</td>
<td>1.69</td>
<td>1.69</td>
</tr>
<tr>
<td>ProdConsChannel</td>
<td>60</td>
<td>1.26</td>
<td>1.69</td>
<td>1.69</td>
<td>1.69</td>
</tr>
<tr>
<td>Producer-consumer</td>
<td>60</td>
<td>1.26</td>
<td>1.69</td>
<td>1.69</td>
<td>1.69</td>
</tr>
<tr>
<td>Quantifiers</td>
<td>28</td>
<td>1.30</td>
<td>1.83</td>
<td>1.83</td>
<td>1.83</td>
</tr>
<tr>
<td>RockBand</td>
<td>109</td>
<td>1.30</td>
<td>1.83</td>
<td>1.83</td>
<td>1.83</td>
</tr>
<tr>
<td>Sieve</td>
<td>56</td>
<td>1.42</td>
<td>3.84</td>
<td>3.84</td>
<td>3.84</td>
</tr>
<tr>
<td>Swap</td>
<td>19</td>
<td>0.98</td>
<td>1.07</td>
<td>1.07</td>
<td>1.07</td>
</tr>
<tr>
<td>Prog1</td>
<td>33</td>
<td>1.07</td>
<td>1.07</td>
<td>1.07</td>
<td>1.07</td>
</tr>
<tr>
<td>Prog2</td>
<td>33</td>
<td>1.07</td>
<td>1.07</td>
<td>1.07</td>
<td>1.07</td>
</tr>
<tr>
<td>Prog3</td>
<td>163</td>
<td>1.58</td>
<td>2.23</td>
<td>2.23</td>
<td>2.23</td>
</tr>
<tr>
<td>Prog4</td>
<td>19</td>
<td>1.07</td>
<td>2.23</td>
<td>2.23</td>
<td>2.23</td>
</tr>
</tbody>
</table>

1st block: tests expected to fail

2nd block: tests expected to succeed

- **Cell**
  - LOC: 60
  - Time (s): 1.12
  - QI: 13
  - Conflicts: 1
  - Expected to fail

- **AVLTree.iterative**
  - LOC: 256
  - Time (s): 2.24
  - QI: 231
  - Conflicts: 177
  - Expected to fail

- **AVLTree.nokeys**
  - LOC: 256
  - Time (s): 34.21
  - QI: 2357
  - Conflicts: 4304
  - Expected to fail

- **Cell**
  - LOC: 131
  - Time (s): 1.78
  - QI: 233
  - Conflicts: 40
  - Expected to fail

- **OwickiGries**
  - LOC: 60
  - Time (s): 1.26
  - QI: 220
  - Conflicts: 42
  - Expected to fail

- **PetersonsAlgorithm**
  - LOC: 70
  - Time (s): 2.12
  - QI: 1196
  - Conflicts: 278
  - Expected to fail

- **ProdConsChannel**
  - LOC: 78
  - Time (s): 1.43
  - QI: 282
  - Conflicts: 117
  - Expected to fail

- **Producer-consumer**
  - LOC: 171
  - Time (s): 1.91
  - QI: 466
  - Conflicts: 147
  - Expected to fail

- **Quantifiers**
  - LOC: 28
  - Time (s): 1.02
  - QI: 16
  - Conflicts: 4
  - Expected to fail

- **RockBand**
  - LOC: 109
  - Time (s): 1.30
  - QI: 79
  - Conflicts: 14
  - Expected to fail

- **Sieve**
  - LOC: 56
  - Time (s): 1.42
  - QI: 308
  - Conflicts: 93
  - Expected to fail

- **Swap**
  - LOC: 19
  - Time (s): 0.98
  - QI: 10
  - Conflicts: 2
  - Expected to fail

- **Prog1**
  - LOC: 33
  - Time (s): 1.07
  - QI: 40
  - Conflicts: 27
  - Expected to fail

- **Prog2**
  - LOC: 52
  - Time (s): 1.11
  - QI: 22
  - Conflicts: 17
  - Expected to fail

- **Prog3**
  - LOC: 163
  - Time (s): 1.58
  - QI: 627
  - Conflicts: 144
  - Expected to fail

- **Prog4**
  - LOC: 19
  - Time (s): 1.07
  - QI: 76
  - Conflicts: 22
  - Expected to fail