Tracking Origins
Andreas Zeller
Today’s Topics

• Exploring History
• Dynamic Slicing
• Leveraging Origins
Exploring the Past

A typical debugging session looks like this:

1. Set a breakpoint
2. Start program, reaching breakpoint
3. Step, Step, Step, …
4. Oops! I’ve gone too far!
Omniscient Debugging

Thread.new(<DemoRunnable_2>, "Sorter") -> <Sorter_2>
<Sorter_2>.start() -> void
<Sorter_2>.sort(16, 19) -> void
<Sorter_2>.sort(16, 19) -> 1885
<Sorter_2>.sort(16, 17) -> void
<Sorter_2>.sort(16, 18) -> void
sort -> void
<Sorter_2>.join() -> void

Stack
<DemoRunnable_0>.run()
Demo_0.sort(0, 19)
Demo_0.sort(11, 19)
Demo_0.sort(16, 19)

Locals
start 16
end 19
i --
j --
tmp --
middle --
newEnd --

TTY Output
A badMethod threw: java.lang.NullPointerException: Bad met
Starting QuickSort: 20
-- Done sorting --
-2 0 1
-1 0 1
-2 237 2
-3 240 3
-4 480 4
-5 486 5
-6 492 6
-7 729 7
-8 735 8
-9 978 9

From last: -437 stamps, -0.026secs  First Line in: <Demo_0>.sort(16, 19) -> void
How does it work?

- ODB records a trace of the entire execution history
- Slows down programs by a factor of 10
- Records about 100 MB/s
- Now available in commercial tools
Dynamic Slicing

- Static slices apply to all program runs:
  - General + reusable, but imprecise
- A dynamic slice applies to a single run:
  - Specific and precise
Static Slicing

- Given a statement B, the backward slice contains all statements that could influence the read variables or execution of B
- Formally:
  \[ S^B(B) = \{ A | A \rightarrow^* B \} \]
1 n = read();
2 a = read();
3 x = 1;
4 b = a + x;
5 a = a + 1;
6 i = 1;
7 s = 0;
8 while (i <= n) {
9     if (b > 0)
10        if (a > 1)
11           x = 2;
12     s = s + x;
13     i = i + 1;
14 }
15 write(s);

Static slice for (s, 15)  Dynamic slice for (s, 15)
1 n = read();
2 a = read();
3 x = 1;
4 b = a + x;
5 a = a + 1;
6 i = 1;
7 s = 0;
8 while (i <= n) {
9     if (b > 0)
10         if (a > 1)
11             x = 2;
12     s = s + x;
13     i = i + 1;
14 }
15 write(s);

1. Obtain a trace of the execution
2. Get the variables that are read and written
3. Assign an empty slice to each written variable
4. Compute the slices from start to end:

\[
\text{DynSlice}(w) = \bigcup_i (\text{DynSlice}(r_i) \cup \{\text{line}(r_i)\})
\]
<table>
<thead>
<tr>
<th>Trace</th>
<th>Write</th>
<th>Read</th>
<th>Dynamic Slice</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>1 n = read();</code></td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>2 a = read();</code></td>
<td>a</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>3 x = 1;</code></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>4 b = a + x;</code></td>
<td>b</td>
<td>a, x</td>
<td>2, 3</td>
</tr>
<tr>
<td><code>5 a = a + 1;</code></td>
<td>a</td>
<td>a</td>
<td>2</td>
</tr>
<tr>
<td><code>6 i = 1;</code></td>
<td>i</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>7 s = 0;</code></td>
<td>s</td>
<td></td>
<td></td>
</tr>
<tr>
<td><code>8 while (i &lt;= n) {</code></td>
<td>p8</td>
<td>i, n</td>
<td>6, 1</td>
</tr>
<tr>
<td><code>9    if (b &gt; 0)</code></td>
<td>p9</td>
<td>b, p8</td>
<td>4, 2, 3, 8, 6, 1</td>
</tr>
<tr>
<td><code>10   if (a &gt; 1)</code></td>
<td>p10</td>
<td>a, p9</td>
<td>5, 2, 9, 4, 2, 3, 8, 6, 1</td>
</tr>
<tr>
<td><code>12    s = s + x;</code></td>
<td>s</td>
<td>s, x, p8</td>
<td>7, 3, 8, 6, 1</td>
</tr>
<tr>
<td><code>13    i = i + 1;</code></td>
<td>i</td>
<td>i, p8</td>
<td>8, 6, 1</td>
</tr>
<tr>
<td><code>8 while (i &lt;= n) {</code></td>
<td>p8</td>
<td>i, n</td>
<td>13, 8, 6, 1</td>
</tr>
<tr>
<td><code>9    if (b &gt; 0)</code></td>
<td>p9</td>
<td>b, p8</td>
<td>4, 2, 3, 13, 8, 6, 1</td>
</tr>
<tr>
<td><code>10   if (a &gt; 1)</code></td>
<td>p10</td>
<td>a, p9</td>
<td>5, 2, 9, 4, 2, 3, 13, 8, 6, 1</td>
</tr>
<tr>
<td><code>12    s = s + x;</code></td>
<td>s</td>
<td>s, x, p8</td>
<td>12, 7, 3, 6, 8, 1, 13</td>
</tr>
<tr>
<td><code>13    i = i + 1;</code></td>
<td>i</td>
<td>i, p8</td>
<td>13, 8, 6, 1</td>
</tr>
<tr>
<td><code>8 while (i &lt;= n) {</code></td>
<td>p8</td>
<td>i, n</td>
<td>13, 8, 6, 1</td>
</tr>
<tr>
<td><code>15 write(s);</code></td>
<td>o15</td>
<td>s</td>
<td>12, 7, 3, 6, 8, 1, 13</td>
</tr>
</tbody>
</table>

\[
\text{DynSlice}(w) = \bigcup_i (\text{DynSlice}(r_i) \cup \{\text{line}(r_i)\})
\]
<table>
<thead>
<tr>
<th>Trace</th>
<th>Write</th>
<th>Read</th>
<th>Dynamic Slice</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>x = 1;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>b = a + x;</td>
<td>b</td>
<td>a, x</td>
</tr>
<tr>
<td>5</td>
<td>a = a + 1;</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>6</td>
<td>i = 1;</td>
<td>i</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>s = 0;</td>
<td>s</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>while (i &lt;= n) {</td>
<td>p8</td>
<td>i, n</td>
</tr>
<tr>
<td>9</td>
<td>if (b &gt; 0)</td>
<td>p9</td>
<td>b, p8</td>
</tr>
<tr>
<td>10</td>
<td>if (a &gt; 1)</td>
<td>p10</td>
<td>a, p9</td>
</tr>
<tr>
<td>12</td>
<td>s = s + x;</td>
<td>s</td>
<td>s, x, p8</td>
</tr>
<tr>
<td>13</td>
<td>i = i + 1;</td>
<td>i</td>
<td>i, p8</td>
</tr>
<tr>
<td>8</td>
<td>while (i &lt;= n) {</td>
<td>p8</td>
<td>i, n</td>
</tr>
<tr>
<td>9</td>
<td>if (b &gt; 0)</td>
<td>p9</td>
<td>b, p8</td>
</tr>
<tr>
<td>10</td>
<td>if (a &gt; 1)</td>
<td>p10</td>
<td>a, p9</td>
</tr>
<tr>
<td>12</td>
<td>s = s + x;</td>
<td>s</td>
<td>s, x, p8</td>
</tr>
<tr>
<td>13</td>
<td>i = i + 1;</td>
<td>i</td>
<td>i, p8</td>
</tr>
<tr>
<td>8</td>
<td>while (i &lt;= n) {</td>
<td>p8</td>
<td>i, n</td>
</tr>
<tr>
<td>15</td>
<td>write(s);</td>
<td>o15</td>
<td>s</td>
</tr>
<tr>
<td>Trace</td>
<td>Write</td>
<td>Read</td>
<td>Dynamic Slice</td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>------</td>
<td>----------------</td>
</tr>
<tr>
<td>1 n = read();</td>
<td>n</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 a = read();</td>
<td>a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 x = 1;</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 b = a + x;</td>
<td>b</td>
<td>a, x</td>
<td>2, 3</td>
</tr>
<tr>
<td>5 a = a + 1;</td>
<td>a</td>
<td>a</td>
<td>2</td>
</tr>
<tr>
<td>6 i = 1;</td>
<td>i</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 s = 0;</td>
<td>s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 while (i &lt;= n) {</td>
<td>p8</td>
<td>i, n</td>
<td>6, 1</td>
</tr>
<tr>
<td>9 if (b &gt; 0)</td>
<td>p9</td>
<td>b, p8</td>
<td>4, 2, 3, 8, 6, 1</td>
</tr>
<tr>
<td>10 if (a &gt; 1)</td>
<td>p10</td>
<td>a, p9</td>
<td>5, 2, 9, 4, 2, 3, 8, 6, 1</td>
</tr>
<tr>
<td>12 s = s + x;</td>
<td>s</td>
<td>s, x, p8</td>
<td>7, 3, 8, 6, 1</td>
</tr>
<tr>
<td>13 i = i + 1;</td>
<td>i</td>
<td>i, p8</td>
<td>8, 6, 1</td>
</tr>
<tr>
<td>8 while (i &lt;= n) {</td>
<td>p8</td>
<td>i, n</td>
<td>13, 8, 6, 1</td>
</tr>
<tr>
<td>9 if (b &gt; 0)</td>
<td>p9</td>
<td>b, p8</td>
<td>4, 2, 3, 13, 8, 6, 1</td>
</tr>
<tr>
<td>10 if (a &gt; 1)</td>
<td>p10</td>
<td>a, p9</td>
<td>5, 2, 9, 4, 2, 3, 13, 8, 6, 1</td>
</tr>
<tr>
<td>12 s = s + x;</td>
<td>s</td>
<td>s, x, p8</td>
<td>12, 7, 3, 6, 8, 1, 8, 13</td>
</tr>
<tr>
<td>13 i = i + 1;</td>
<td>i</td>
<td>i, p8</td>
<td>13, 8, 6, 1</td>
</tr>
<tr>
<td>8 while (i &lt;= n) {</td>
<td>p8</td>
<td>i, n</td>
<td>13, 8, 6, 1</td>
</tr>
<tr>
<td>15 write(s);</td>
<td>o15</td>
<td>s</td>
<td>12, 7, 3, 6, 8, 1, 13</td>
</tr>
</tbody>
</table>
n = read();
a = read();
x = 1;
b = a + x;
a = a + 1;
i = 1;
s = 0;
while (i <= n) {
    if (b > 0)
        if (a > 1)
            x = 2;
    s = s + x;
    i = i + 1;
}
write(s);

1 n = read(); // n = 2
2 a = read(); // a = 0
3 x = 1;
4 b = a + x;
5 a = a + 1;
6 i = 1;
7 s = 0;
8 while (i <= n) {
9     if (b > 0)
10         if (a > 1)
11             x = 2;
12     s = s + x;
13     i = i + 1;
14 }
15 write(s);

Static slice for (s, 15)  Dynamic slice for (s, 15)
Discussion

• Dynamic slices are much more precise than static slices (applied to the one run, that is)

• From some variable, a backward slice encompasses on average
  • 30% of the entire program (static slice)
  • 5% of the executed program (dynamic slice)

• Overhead as in omniscient debugging
Question: Why didn’t Pac resize 0.5?

Answer:
One or more of these actions prevented Pac resize 0.5 from happening. Try following the arrows and checking each action to find out what went wrong.
Question: Why didn't Pac resize 0.5?

Answer:
One or more of these actions prevented Pac resize 0.5 from happening. Try following the arrows and checking each action to find out what went wrong.
questions are chosen from a hierarchical menu

code related to the question is highlighted

Question: Why didn't Pac resize 0.5?

Answer:
One or more of these actions prevented Pac resize 0.5 from happening. Try following the arrows and checking each action to find out what went wrong.
“Why did” questions

• Take the dynamic slice of the variable
• Follow at most two dependencies
• If programmer wants, follow dependencies transitively
```
1 n = read(); // n = 2
2 a = read(); // a = 0
3 x = 1;
4 b = a + x;
5 a = a + 1;
6 i = 1;
7 s = 0;
8 while (i <= n) {
9     if (b > 0)
10        if (a > 1)
11            x = 2;
12     s = s + x;
13     i = i + 1;
14 }
15 write(s);
```

""Why did s = 2 in Line 15?"

""Because s = 1 and i = 2""
“Why didn’t” questions

• Follow back control dependencies to closest controlling statement(s)

• Do a “why did” question on each

• Again, follow at most two dependencies
“Why didn’t x = 2 in Line 11?”

```
1 n = read(); // n = 2
2 a = read(); // a = 0
3 x = 1;
4 b = a + x;
5 a = a + 1;
6 i = 1;
7 s = 0;
8 while (i <= n) {
9     if (b > 0)
10        if (a > 1)
11           x = 2;
12     s = s + x;
13     i = i + 1;
14 }
15 write(s);
```

“Because a = 1 and b = 1”
Discussion

The WHYLINE combines

• omniscient debugging
• static slicing
• dynamic slicing

in an attractive package, showcasing the state of the art in interactive debugging
Tracking Infections

1. Start with the infected value as seen in the failure
2. Follow back the dependencies
3. Observe and judge origins – are they sane?
4. If some origin is infected, repeat at Step 2
5. All origins are sane? Here’s the infection site!
Concepts

★ Omniscient debugging allows for simple exploration of the entire execution history

★ Dynamic slicing tells the origin of a value

★ To track down an infection, follow dependencies and observe origins, repeating the process for infected origins