Inferring Hierarchical Motifs from Execution Traces

Saba Alimadadi, Ali Mesbah, Karthik Pattabiraman

saba@northeastern.edu
All truths are easy to understand once they are discovered; the point is to discover them.

— Galileo Galilei
[Ko et al. '06], [Murphy et al. '95]
Execution Traces: Complex and Overwhelming
recurring patterns?
previously unknown
tolerant of small changes
hierarchical
Motifs of program execution

Inspired by bio-informatics

Smith-Waterman [1981]
BLAST [1990]
Most Popular Programming Language

- JavaScript: 52.5% popularity, 340,000 active repositories
- Java: 35% popularity, 255,000 active repositories
```javascript
function validateEmail () {
    // do stuff
}

function checkAddress () {
    // do more stuff
}
```
Email: <input type="text" />
Address: <input type="text" />
Occupation: <select>
  <option value="" disabled selected>Select one</option>
  <option>Student</option>
</select>
Degree: <select>
  <option value="" disabled selected>Select one</option>
  <option>Student</option>
</select>

Submit: <button type="submit" onclick="studentForm()">Submit</button>
validateEmail() checkAddress() occupation() submit()
BLAST [Altschul et al. 1990]
A
validateEmail
checkAddress
occupation
studentForm
submit
refresh
updateRegistry
informServer
successMessage
getNewList
animateChange
setDefault

B
motif1
motif2
motif3
motif4
motif5
DB Motifs

C

D
validateEmail
checkAddress
occupation
studentForm
submit

E

index.html

function validateEmail()
{
  // do stuff
}

function checkAddress()
{
  // do stuff
}

$(
  "#email"
).addEventListener("change", validateEmail, false);

$(
  ".addr"
).click(checkAddress);

$(
  ".dropdown-content"
).addEventListener("change", occupation);

function occupation()
{
  eval(this.value + ".JobForm()");
}

function avademicJobFor()
{
  var label = document.
  var degreeType = docu
  var option1 = docum
  option1.value = "Und
}

https://github.com/saltlab/sabalan
Does using **SABALAN** improve performance of developers?
Controlled Experiment

Tasks: common comprehension activities

- 14 participants
- Control: tool
- Measure: performance
Fast is fine, but **accuracy** is everything.

<table>
<thead>
<tr>
<th></th>
<th>Sabalan</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1.a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1.b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2.a</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2.b</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Accuracy (%)
Execution Traces: Complex and Overwhelming

Motifs of program execution
Inspired by bio-informatics
Smith-Waterman [1981]
BLAST [1990]

Fast is fine, but accuracy is everything.

Accuracy (%)

54 pp more accuracy