Selenium Study

Laurent Christophe
lachrist@vub.ac.be
Test Automation

Rapid release cycle

Development → Integration → GUI → Test Automaton

SUT
Test Automation

Test Automaton → Firefox → Google Servers

- Fetch
- Fill
- Click
- Assert
- Quit

GET

GET
Fragility Problem

Test Automaton → SUT → Test Automaton → SUT' → Test Automaton' → SUT'
A Recognized Problem


Querying SourceForge with BOA

Boa Source Code

```
out: output collection of string;
proj: Project = input;
check: bool = false;
visit(proj, visitor {
    before ast: ASTRoot -> {
        exists (i: int ; match(".*selenium.*", lowercase(ast.imports[i]))) {
            check = true;
        }
    }
});
if (check) {
    out << proj.project_url;
}
```

Input Dataset

2013 May

http://boa.cs.iastate.edu/boa/
Querying SourceForge with BOA

<table>
<thead>
<tr>
<th>Output for Job 2829</th>
</tr>
</thead>
<tbody>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/somap">http://sourceforge.net/projects/somap</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/1portal">http://sourceforge.net/projects/1portal</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/japsurveylib">http://sourceforge.net/projects/japsurveylib</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/bitprime">http://sourceforge.net/projects/bitprime</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/junitbook">http://sourceforge.net/projects/junitbook</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/htmlunit">http://sourceforge.net/projects/htmlunit</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/flashrecruit">http://sourceforge.net/projects/flashrecruit</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/specrunner">http://sourceforge.net/projects/specrunner</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/atarrabi">http://sourceforge.net/projects/atarrabi</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/itracker">http://sourceforge.net/projects/itracker</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/actioncenters">http://sourceforge.net/projects/actioncenters</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/jameleon">http://sourceforge.net/projects/jameleon</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/adempiere">http://sourceforge.net/projects/adempiere</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/iwii-organizer">http://sourceforge.net/projects/iwii-organizer</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/officefloor">http://sourceforge.net/projects/officefloor</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/xrace">http://sourceforge.net/projects/xrace</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/cruisereport">http://sourceforge.net/projects/cruisereport</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/storytestiq">http://sourceforge.net/projects/storytestiq</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/wonder">http://sourceforge.net/projects/wonder</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/simpleyql">http://sourceforge.net/projects/simpleyql</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/wicketpagetest">http://sourceforge.net/projects/wicketpagetest</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/sitools2">http://sourceforge.net/projects/sitools2</a></td>
</tr>
<tr>
<td>out[]= <a href="http://sourceforge.net/projects/pcollecta">http://sourceforge.net/projects/pcollecta</a></td>
</tr>
</tbody>
</table>

... 105 Results
Querying GitHub
Querying GitHub

Wow

Much results
(120,000)

Max 100 pages

10 Hits / Pages
Querying GitHub

Result slicing (< 1000)

Advanced search criterion

import openqa.selenium.size:100..200

We've found 265 code results

Java

1. import org.openqa.selenium.WebDriver;
2. import org.openqa.selenium.support.ui.Wait;
3. import org.openqa.selenium.support.ui.WebDriverWait;

3. import org.openqa.selenium.support.ui.WebDriverWait;
4. public class WaiterOfPageLoad {
5. }

Languages
Java 198
Groovy 51
Scala 10
Querying GitHub

GitHub

Web Scraping

API Calls

Cloning

SQL DB

Extending

Filtering

GIT Repos
Using GitHub API

Ivan Žužak

... Finally, while you're free to use the API as much as the rate limit permits, please do so at a reasonable pace. In other words, don't spin up a hundred EC instances and start making API requests in parallel -- that kind of applications are hard to tell apart from malicious attacks. Making a few requests in parallel should be okay, and will earn you a Good Citizen badge. ...

Ivan Žužak

... When you make a request to the Repository Statistics API, you will often get a 202 status code back, which means that we've started a job to compute the statistics data and that you should come back later to pick it up [1]. However, there might be situations where this data is never computed (e.g. due to repository corruption). In such cases, those endpoints will keep returning a 202 status code every time you come back to pick the data up. ...
Scrapping GitHub

Stop, Demo Time