# Test-Driven Development (TDD)

#### **TDD:** Description

 Developer writes an automated test case prior to writing the product code implementing the feature under test

Most frequently used as an enhancement to unitlevel testing

May also be used with acceptance testing

Green

Red

Refactor

Add a test for yet-to-bebuild functionality. This test will fail (test suite becomes red).

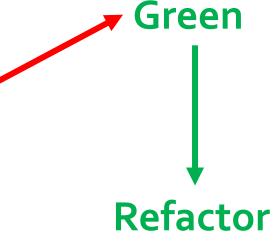
Green

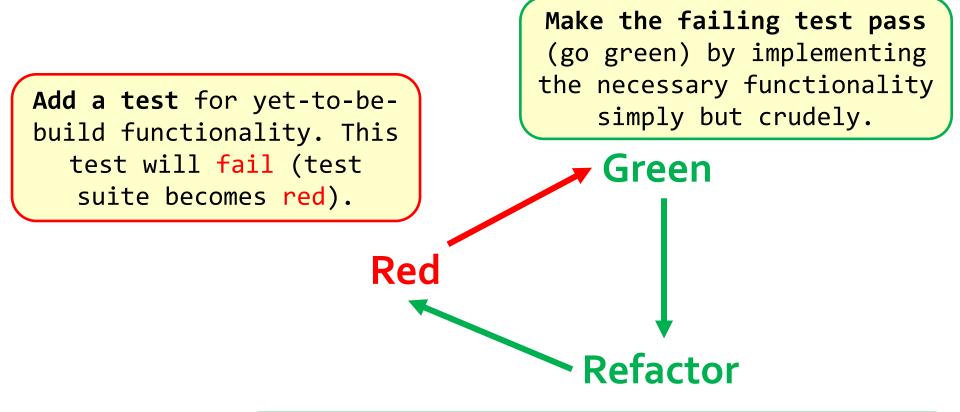
Red

Refactor

Red

Add a test for yet-to-bebuild functionality. This test will fail (test suite becomes red). Make the failing test pass (go green) by implementing the necessary functionality simply but crudely.





Use refactoring to ensure the overall code
base is as clean and well-designed as possible
for currently-implemented functionality.

# TDD: Measuring Defect Removal Effectiveness

 Direct Measurement: No known method (because TDD may prevent as well as remove defects)

#### Indirect Measurement:

- Measure the reduction in down-stream defect density in a large code population before/after introducing TDD
- ■Published Results: 20..91% (Williams, Brown, et al.)

#### TDD: Why it Works (Potential Explanation)

- Thinking about test cases before writing the code helps incrementally design the software
  - Writing test cases before code requires writing stubs/mocks that interact with other objects through interfaces designed incrementally
- •TDD encourages more thorough testing (which remove more defects). Evidence: TDD seems to increase:
  - LinesOfTestCode / LinesOfProductCode ratio and
  - test coverage

#### TDD: Advantages

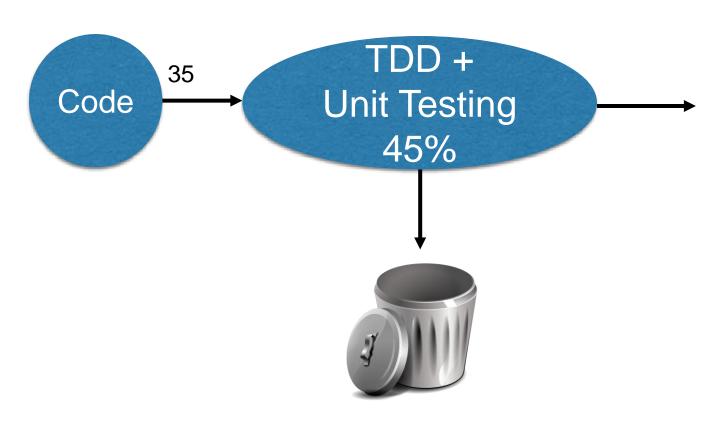
•May reduce time to find a defect because problem is known to lie within just a few lines of new code

#### TDD: Disadvantages

- Test and product code are written by the same developer (or pair), effectively limiting their thoroughness to their skill
  - (similar to a Code Review)
- Increase code churn:
  - Throw away (or refactor) tests when code is refactored

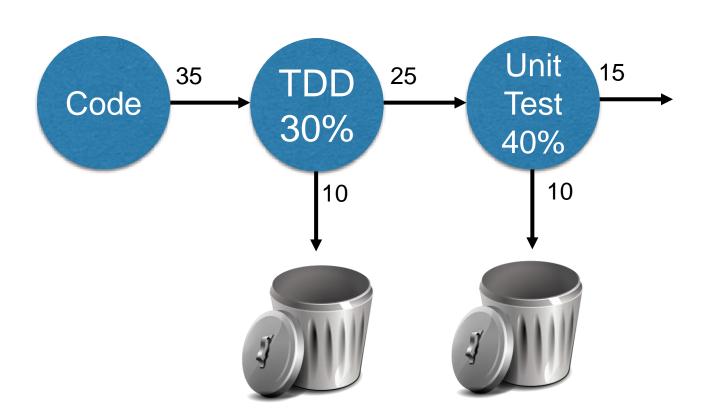
#### TDD: Modeling

Version 1: Model TDD as a more effective unit-level testing activity



#### TDD: Modeling

Version 2: Model TDD as a distinct defect removal activity with an effectiveness of about 30% preceding a traditional unit-level test



# Static Analysis

"Smart people make dumb mistakes"

### Static Analysis: Description

 Automated review of source code — essentially free (i.e., cheap to run)

"Static" because code is not executed

Static Analysis Tools available for most important languages

- Many tools have user-editable rules that configure what errors they will flag
  - •e.g., "I'm only interested in null pointer exceptions"

### Static Analysis: Example Tools

- **Java:** FindBugs, Checkstyle, ThreadSafe, Coverity
- •JavaScript: Google's Closure, JSLint, JSHint
- •Objective-C: Clang (in Xcode)
- **C#:** FxCop, Coverity
- •C/C++: Coverity, PRQA QA C, QA C++, SLAM Project
- ■PHP: RIPS
- **Python:** Pylint
- Special Purpose: Fortify

#### Sample Code for Review

```
// try to get the service
Object obj = ServiceManager.getService(FileOpenerService.class, myFinder);

// Preferred service is not found, use the only one available instead
if ((obj == null) && (!myFinder.equals(finders[0])))
   obj = ServiceManager.getService(FileOpenerService.class, finders[0]);

// Open the file!

((FileOpenerService)obj).openFile(fileName, view);
```

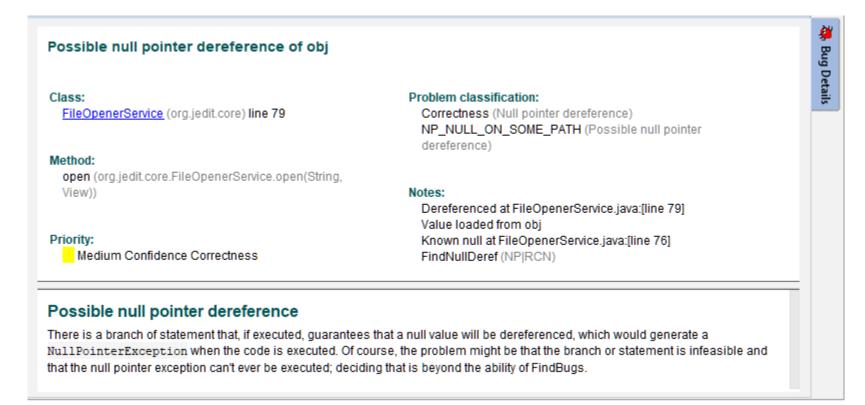
#### Sample Code for Review

```
// try to get the service
Object obj = ServiceManager.getService(FileOpenerService.class, myFinder);

// Preferred service is not found, use the only one available instead
if ((obj == null) && (!myFinder.equals(finders[0])))
   obj = ServiceManager.getService(FileOpenerService.class, finders[0]);

// Open the file!

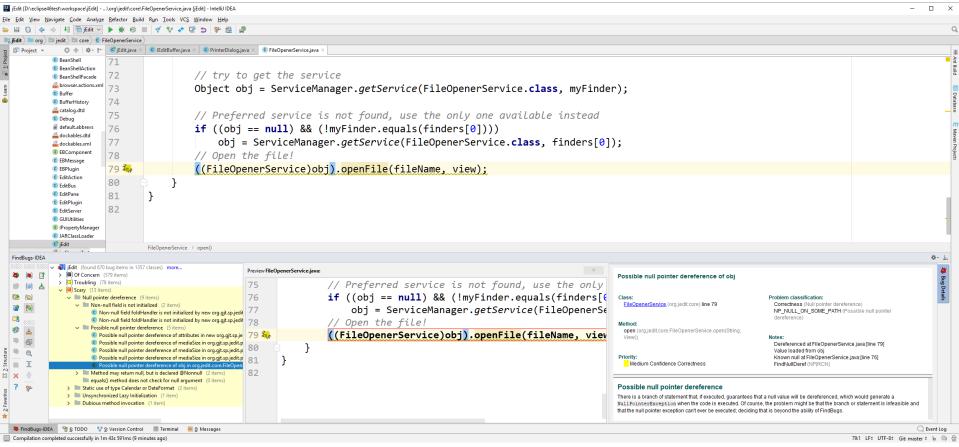
((FileOpenerService)obj).openFile(fileName, view);
```



# Example Error found by FindBugs 🔌



"Possible null pointer dereference of obj"

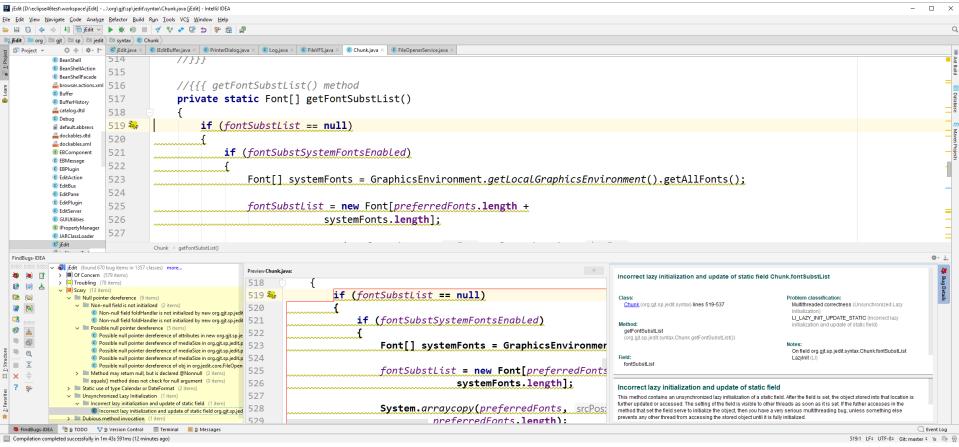


http://findbugs.sourceforge.net/ (IntelliJ IDEA Plugin)

# Example Error found by FindBugs 🔌



"Incorrect lazy initialization and update of static filed Chunk.fontSubstList"



http://findbugs.sourceforge.net/ (IntelliJ IDEA Plugin)

# Sonarcloud example

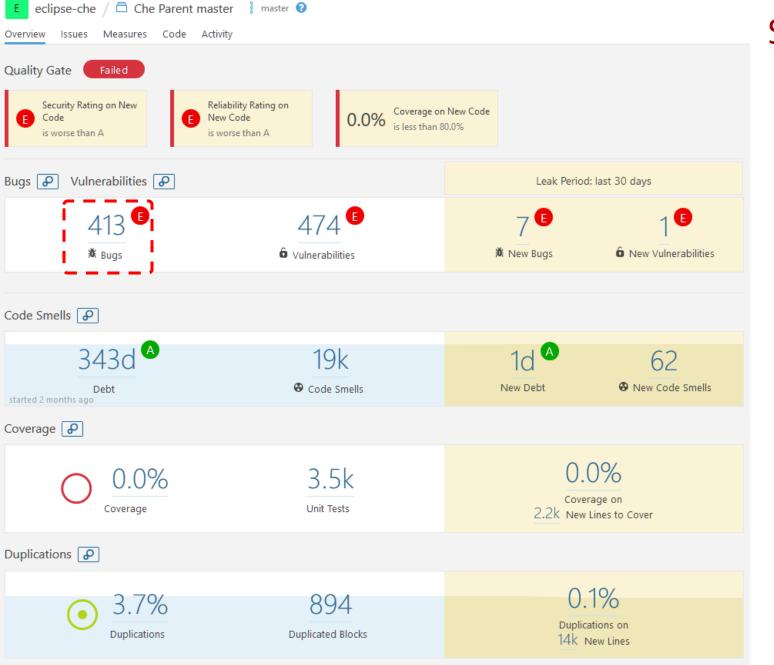


■ README.md

#### **Eclipse Che - Eclipse Next-Generation IDE**



https://www.eclipse.org/che/. Next-generation Eclipse platform, developer workspace serve workspaces that include their dependencies including embedded containerized runtimes, V makes workspaces distributed, collaborative, and portable to run anywhere on a desktop o

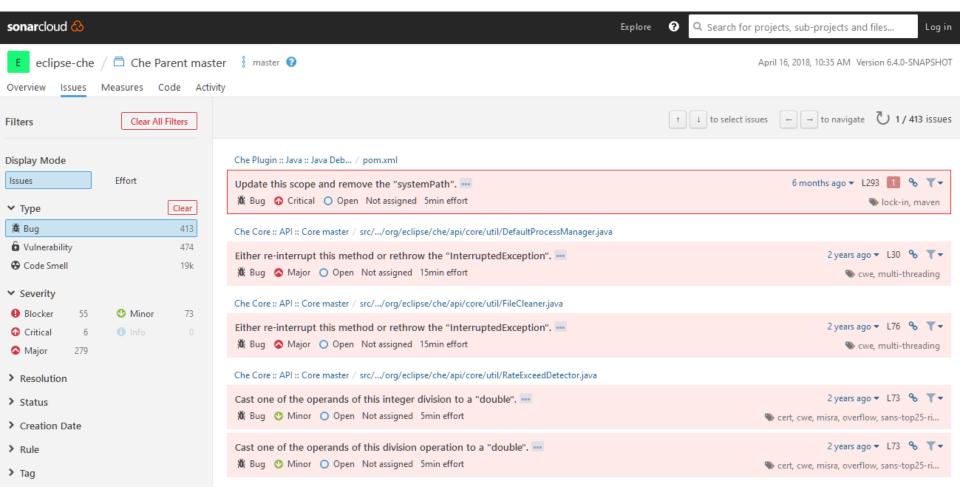


Sonarcloud example



# Sonarcloud example





#### https://sonarcloud.io/dashboard?id=org.eclipse.che%3Ache-parent%3Amaster

#### Static Analysis

- •Run static analysis before integration on a sprint/master branch
  - nobody wants to break the build for a defect the static analysis tool would have caught!
  - No developers wants to review code containing bugs that static analysis tools would have caught
    - Developer time is far more important and costly (\$\$\$)

## Static Analysis: Measuring Defect Removal Effectiveness

Direct Measurement: Just count #defects repaired

Remember... finding a defect doesn't remove it

# Static Analysis: Activities Removing the Same Defects

 Compare to manual code review but much, much cheaper

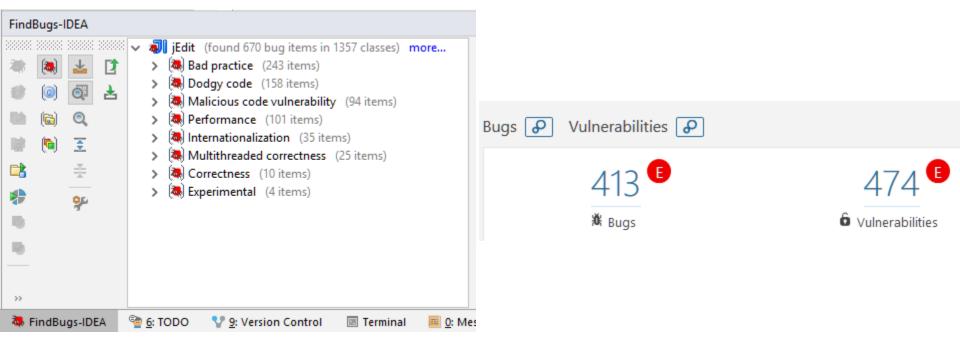
#### Static Analysis: Advantages

Ridiculously simple and easy to use

- High-end tools can find cross-class problems, e.g.:
  - dereferencing a null pointer
  - memory leaks
  - security vulnerabilities
  - division by zero, etc.
- that even code reviews may overlook

### Static Analysis: Disadvantages

- Some tools can report noise\* rather than customervisible defects
  - Developers need to configure the rules
- \*large number of "errors/warning/code smells" that aren't important enough to fix by developers



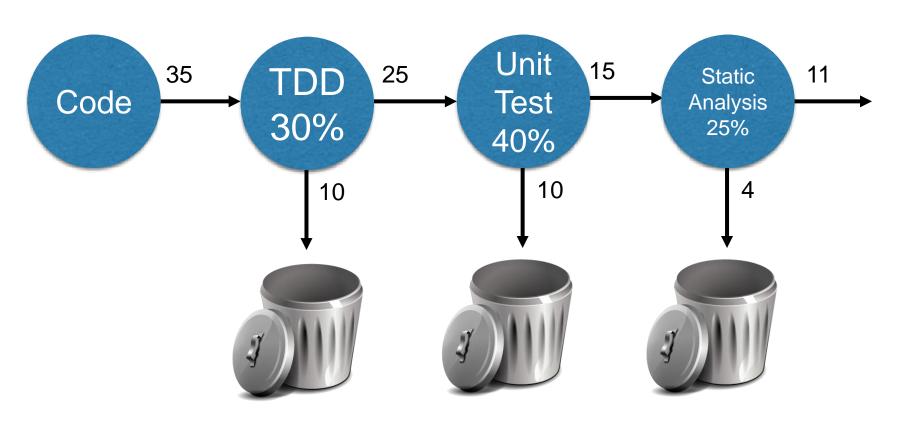
### Static Analysis: Modeling

 In real-world development, Static Analysis probably should follow TDD and Unit-Level Testing

 And should always proceed a manual Code Review (to avoid wasting the Code Review team's time on what can be automated)

Effectiveness claims vary widely (15..90%) depending upon the definition of a "defect"

# Modeling Static Analysis



# Code Reviews

#### Code Reviews

#### Includes:

- Formal Code Inspections
- Code Reviews
- Peer Reviews
- Code Walkthroughs any examination of the source code to find defects

#### Code Review

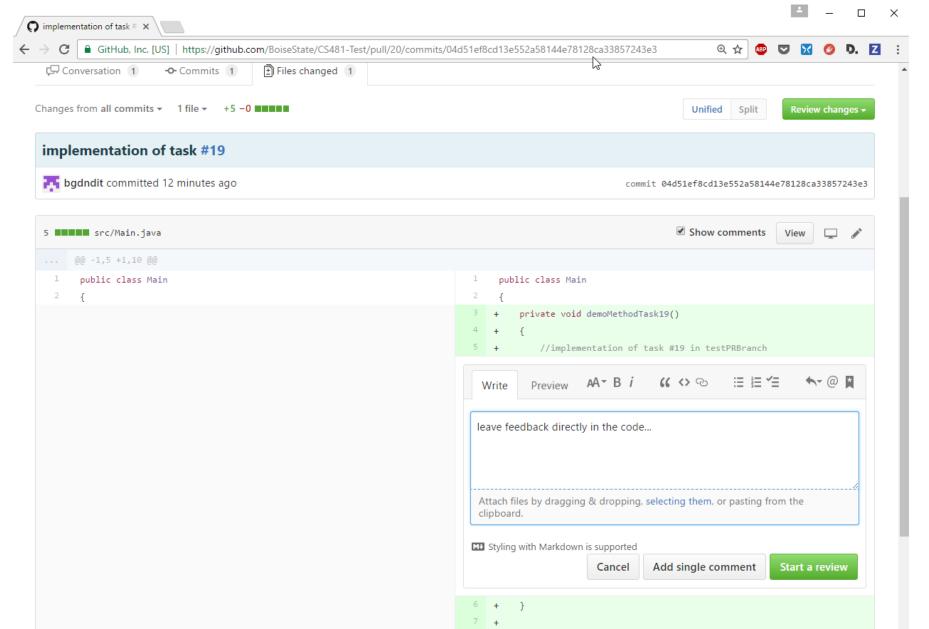
Code reviews place a burden on other Developers

You likely want to clean-up your code as much as possible before sending it out for review by others

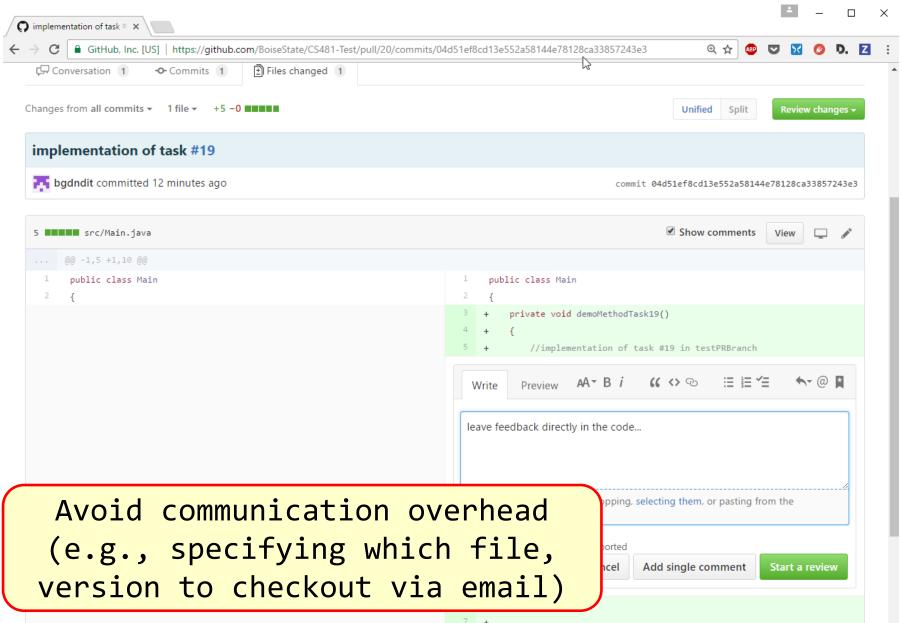
#### Where to Perform Code Reviews?

- •And you likely want to have it reviewed before you push it to the remote repository and share its defects with your Team by breaking the build (or worse!)
  - ■That means... you probably want to perform code reviews on either your local repository...
  - Or on a feature/story branch if you're using them (as your exposure is limited to just those few other developers contributing to Tasks on the same story)
  - Pull-Requests are great for code reviews

# Code Review within Pull Request Example



### Code Review within Pull Request Example



### Code Review: Why does it work?

 "Another set of eyes" will often see something, especially integration problems, you miss

- The act of explaining how your code works sometimes reveals its shortcomings to you
- •Group Synergy: The result of a group can exceed the sum of its individuals working alone.
  - Somehow, the interactions of the group prompt its individuals to higher performance than possible when working alone.

# Code Review: Many Variations

•Informal Code Review, Peer Review: Development team meets and discusses the source code

- Walkthrough: Author leads the development team through the code, explaining how it works
- Tool Assisted: e.g., GitHub Pull-Requests Workflow, Gerrit

# Code Review: Many Variations

Informal Code Review, Peer Review: Development team meets and discusses the source code

- Walkthrough: Author leads the development team through the code, explaining how it works
- Tool Assisted: e.g., GitHub Pull-Requests Workflow, Gerrit

```
gerrit / gerrit-server/src/main/java/com/google/gerrit/server/change/PatchSetinserter.java
    private PatchSet patchSet;
                                                                                                        private PatchSet patchSet;
   private ChangeMessage changeMessage;
                                                                                                        private ChangeMessage changeMessage;
    private SshInfo sshInfo;
                                                                                                        private SshInfo sshInfo;
                                                                                                        private ValidatePolicy validatePolicy = ValidatePolicy.GERRIT;
    private ValidatePolicy validatePolicy = ValidatePolicy.GERRIT;
    private boolean draft;
                                                                                                        private boolean draft;
                                                                                                        private boolean runHooks = true;
    private boolean runHooks;
                                                                                                                        Why do you move this out of the constructor? Initially I assumed this... Jan 28 2:55 PM
                                                                                                                        Because it would be identical between the two constructors, so it sa... Jan 28 3:19 PM
   private boolean sendMail;
                                                                                                        private boolean sendMail = true;
   private Account. Id uploader;
                                                                                                        private Account. Id uploader;
   private BatchRefUpdate batchRefUpdate;
                                                                                                        private BatchRefUpdate batchRefUpdate;
                                                                                                        public PatchSetInserter(ChangeHooks hooks,
    public PatchSetInserter(ChangeHooks hooks,
```

# Code Reviews: Measuring Defect Removal Effectiveness

- Direct Measurement:
  - Count the number of repaired bugs
- ■The effectiveness of the formal inspection process can achieve 65%\*

■The informal reviews used in CS471/CS481 perform at the 25..30% level

<sup>\* (</sup>Jones, "Measuring Defect Potentials and Defect Removal Effectiveness" 2008)

# Code Reviews: Advantages

Cost to repair defects found by Code Reviews is usually low because defect reports often cite a particular method or even line-of-code where the problems lie

 Lightweight code reviews can be cost effective in terms of #defectsRemoved / dollar

Shared code ownership

# Code Reviews: Disadvantages

- Some studies report that as many as 75% of the defects found by Code Reviews are "code evolution" (health) problems, not customer-visible functionality
  - If your team is not repairing these problems, then Code Reviews are wasting their time

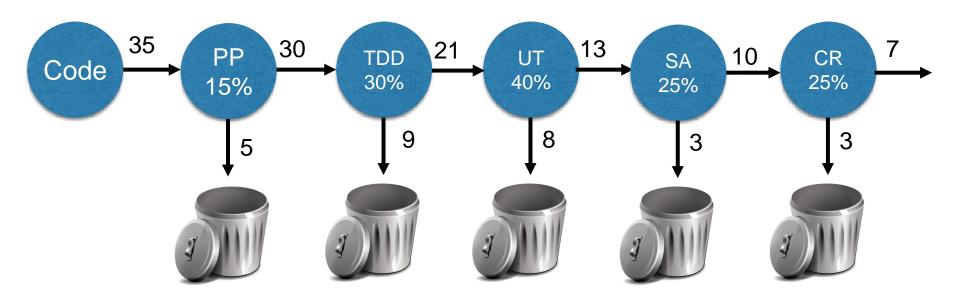
 The formal heavyweight approaches may be expensive in terms of #defectsRemoved / dollar

# Code Reviews: Modeling

 Code Reviews should usually be placed following Static Analysis in your defect removal pipeline

■In CS471/CS481, you can expect 25% effectiveness

# Modeling Everything Discussed So-Far



# Exercise: Code Review (Whiteboard only)

```
01: public class Triangle {
02:
     int a, b, c; //The lengths of the three sides
03: Triangle(int a, int b, int c) {
04: a = a;
05: b = b;
06: c = c;
07:
   }
08:
   //Two equal sides
09:
   boolean isIsosceles() {
10:
       if (a==b) return true;
    if (b==c) return true;
11:
12:
    return false;
13:
    }
14:
    //Three equal sides
15:
    boolean isEquilateral() {
16:
       if ((a==b) \&\& (b==c)) return true;
17:
       return false;
18:
19:
    //Three unequal sides
20:
   boolean isScalene() {
21:
       if ((a!=b) \&\& (b!=c)) return true;
22:
   return false;
23: }
24:}
```

# "Upstream"\* Defect Removal Activities

- Pair Programming
- Test-Driven Development
- •Unit-Level Testing
- Static Analysis
- Code Reviews

# "Downstream"\* Defect Removal Activities

- Integration Testing
- Regression Testing
- System-Level Testing
  - Acceptance Testing
  - Beta Testing

# Integration Testing

# Integration-Level Testing: Description

- Verifies that a subset of the units work well together
- Exercises sub-systems (larger than a unit, smaller than the entire system):
  - package(s)
  - two or more classes working together

Preferably initiated by an automated build system

•May use test dummies / fakes / stubs / spyes / mocks / doubles for major sub-systems (e.g., database)

# Unit Test vs. Integration Test

# Unit Test vs. Integration Test



See animated version at:

https://media.giphy.com/media/307rbPDRHIHwbmcOBy/giphy.gif

# Integration-Level Testing Example

- Module A: "Retrieve Image from Database"
  - should have Unit Tests

- Module B: "Display Image"
  - should have Unit Tests

- Integration tests (Module A + Module B):
  - "Will the code be able to display data loaded from a database?"

# Integration-Level Testing: Effectiveness

■1996Jones: 25..40%

■Kan\*: 36%

Values may be different for modern projects

■CS471 expectation: <20%

<sup>\*</sup> Kan, Stephen. Metrics and Models in Software Quality Engineering 2nd Edition. Addison Wesley. 2003.

# Integration-Level Testing: Why do it?

- Cheapest downstream activity focused on integration defects
- Is sometimes the last fully-automated (i.e., cheaper than manual tests) defect removal activity
  - Cost (\$) is low ⇒ can be executed on every build
  - Helps to discover if you broke-the-build
- Most thorough regression testing activity if you don't have automated acceptance tests executing on every build

# Regression Testing

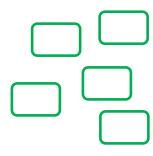
### Regression Testing

- Re-runs tests that previously passed
- Looking for code changes that introduced new defects into previously working source code
  - Did my code change break your perfectly working code?
- Can be performed at many levels
  - unit
  - •integration (preferred)
  - system
- Can use both white and black-box approaches
- Works well if automated in Continuous Integration

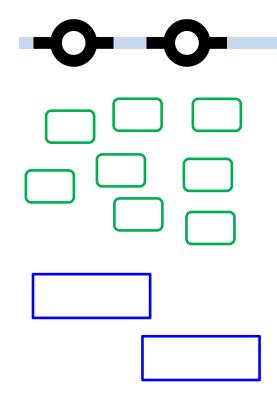
Throughout development, team adds Unit-, Integration- and System-level tests

# Legend Unit-level Test Integration-level Test System-level Test Commit





- Unit-level Test
- Integration-level Test
- System-level Test
- Commit

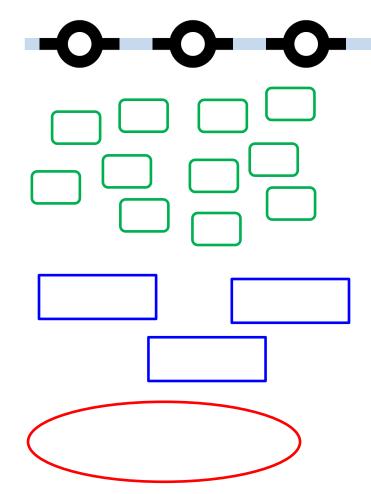


Unit-level Test

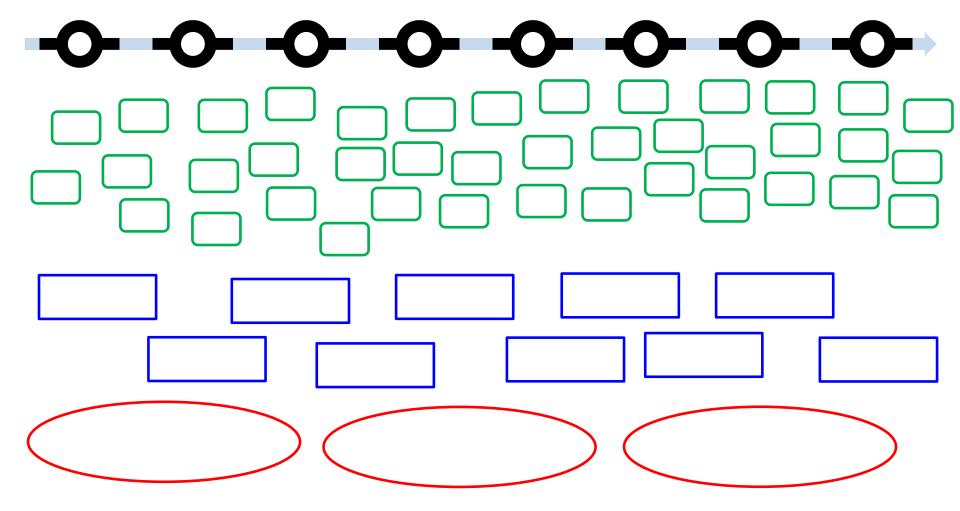
Integration-level Test

System-level Test

• Commit



- Unit-level Test
- Integration-level Test
- System-level Test
- -Commit

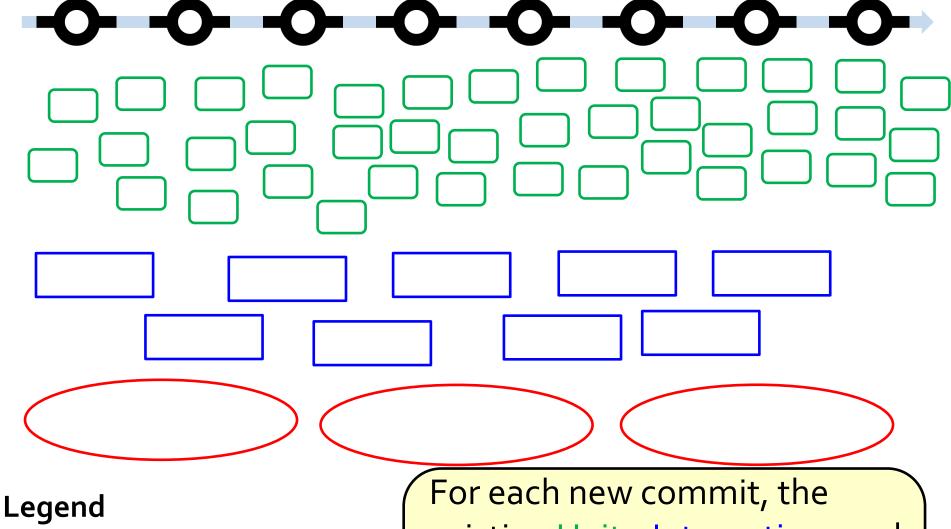


Unit-level Test

Integration-level Test

System-level Test

**O**- Commit



Unit-level Test

Integration-level Test

System-level Test

Commit

existing Unit-, Integration- and System-level tests become regression tests (that ideally should be run automatically)

# System Testing

# System-Level Testing

- Exercising the fully-assembled product
  - Example: Acceptance Testing, Beta Testing

Verifies that all the individual units work together

## White Box Approach

- Exercises the internal structure of the source code as opposed to its external functionality
  - aka Structural Testing

Example: Code Review

•NB: Testers must have access to the source code

# Black Box Approach

Testers don't have access to the source code

- Exercises the product without examining its internal structure
  - Focus on input / output

- Example:
  - Acceptance Testing
  - Stress Test
  - Any exercise of an interface's functionality without examination of its underlying source code

# Functional Testing Approach

 A form of System-Level Black Box testing of a product's functionality

- Yes, there is non-functional testing:
  - performance testing
  - usability testing
  - A/B testing
  - load/stress testing
  - reliability testing
  - etc.

# Example A/B testing

#### Compare variation A with B



# Acceptance Testing

## **Acceptance Testing**

System-level, black-box test of the entire product

- Often dominated by a functional approach
  - But may include usability, reliability, stress & performance
- "Acceptance" refers to the customer's "Acceptance Criteria" (i.e., how will the customer know if a User Story is "Done")

# Manual vs. Automated Testing

#### Manual Testing

- Test procedure executed by a human (e.g., keyboardand-mouse)
- Preferably following a written script to promote repeatability

#### Automated Testing

- Computer executed testing
  - NB: automation refers to the execution of the test, not the analysis and repair of the underlying problem
- Executable script (shell or special-purpose test tool or harness)

### Acceptance Testing: Manual / Automated?

Acceptance Testing is often performed on a GUI

•Manual GUI tests may not yield reproducible results without a written "test procedure" (a script for humans)

- Automation can be a significant investment
  - https://en.wikipedia.org/wiki/List\_of\_GUI\_testing\_tools
  - http://blog.dreamcss.com/tools/gui-testing-tools/
  - Automation works best when you need to run the same test over and over without changing the GUI

# Selenium – Automating Web Testing

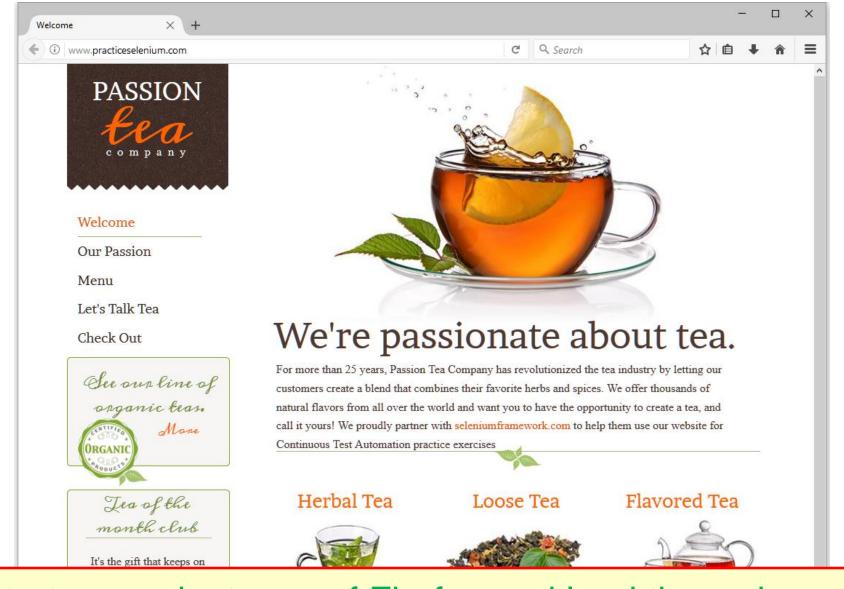
Selenium automates browsers

- Facilitates testing of web applications automatically
- Selenium "is like jUnit for testing GUI and Web applications"
  - write your tests in: Java, C#, Ruby, Python, JS (Node)
  - run your tests in: Firefox, Chrome, IE, Safari, etc.

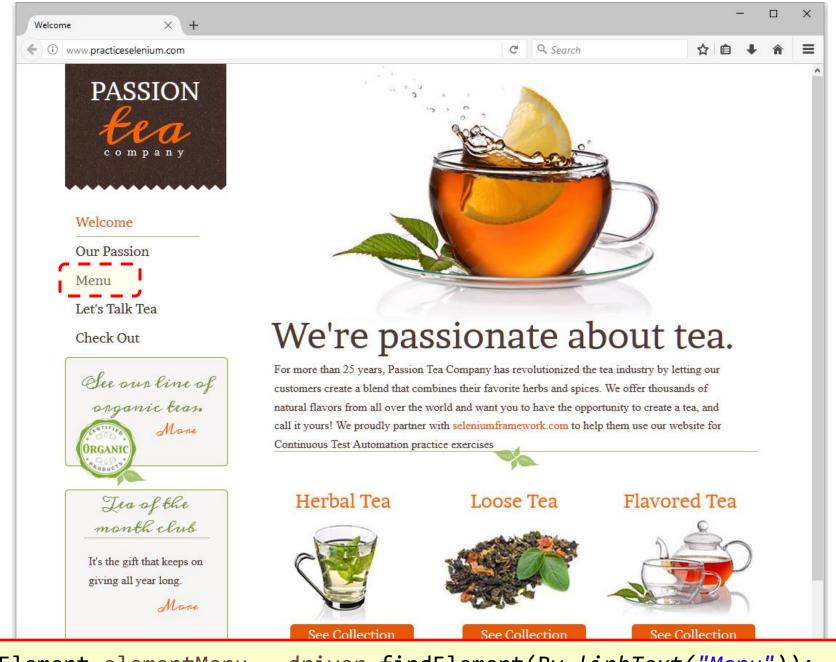
# Sample Java code using Selenium WebDriver that automates executing a web page in Firefox

```
WebDriver driver = new FirefoxDriver();
driver.get("http://www.practiceselenium.com/");
System.out.println("Successfully loaded the website " + "http://www.practiceselenium.com/");
Thread.sleep(2000);
WebElement elementMenu = driver.findElement(By.linkText("Menu"));
elementMenu.click();
Thread.sleep(2000);
WebElement elementCheckout = driver.findElement(By.linkText("Check Out"));
elementCheckout.click();
Thread.sleep(2000);
WebElement elementEmail = driver.findElement(By.id("email"));
elementEmail.sendKeys("johndoe@someemail.com");
Thread.sleep(100);
WebElement elementName = driver.findElement(By.id("name"));
elementName.sendKeys("John Doe");
//...
Select selectionCardType = new Select(driver.findElement(By.id("card type")));
selectionCardType.selectByVisibleText("Mastercard");
//...
WebElement elementPlaceOrder =
driver.findElement(By.xpath("/html/body/div/div/div[1]/div/div[1]/div/div/form/div/button"));
elementPlaceOrder.click();
Thread.sleep(2000);
System.out.println("TODO: Assert if checkout operation was successful");
driver.quit();
```

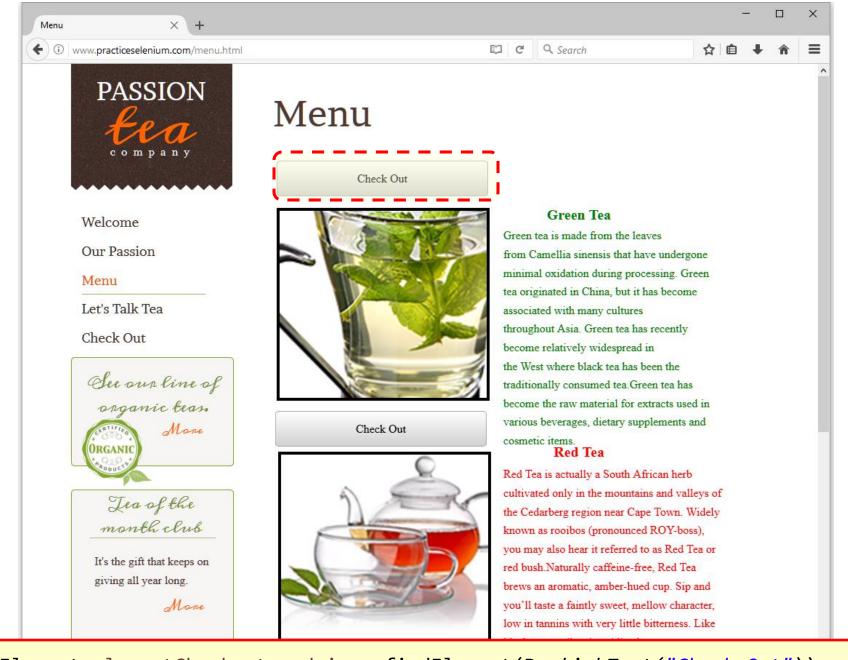
```
WebDriver driver = new FirefoxDriver();
driver.get("http://www.practiceselenium.com/");
System.out.println("Successfully loaded the website " + "http://www.practiceselenium.com/");
Thread.sleep(2000);
WebElement elementMenu = driver.findElement(By.linkText("Menu"));
elementMenu.click();
               Load a page and execute a checkout process
Thread.sleep(2000);
WebElement elementCheckout = driver.findElement(By.linkText("Check Out")
elementCheckout.click();
Thread.sleep(2000);
WebElement elementEmail = driver.findElement(By.)
elementEmail.sendKeys("johndoe@someemail.co"
Thread.sleep(100);
WebElement elementName
elementName.sendKeys
//...
Select selecti
selectionCardTy
                       egyVisibleText("Mastercard");
//...
WebElement elementPlaceOrder =
driver.findElement(By.xpath("/html/body/div/div/div[1]/div/div[1]/div/div/form/div/button"));
elementPlaceOrder.click();
Thread.sleep(2000);
System.out.println("TODO: Assert if checkout operation was successful");
driver.quit();
```



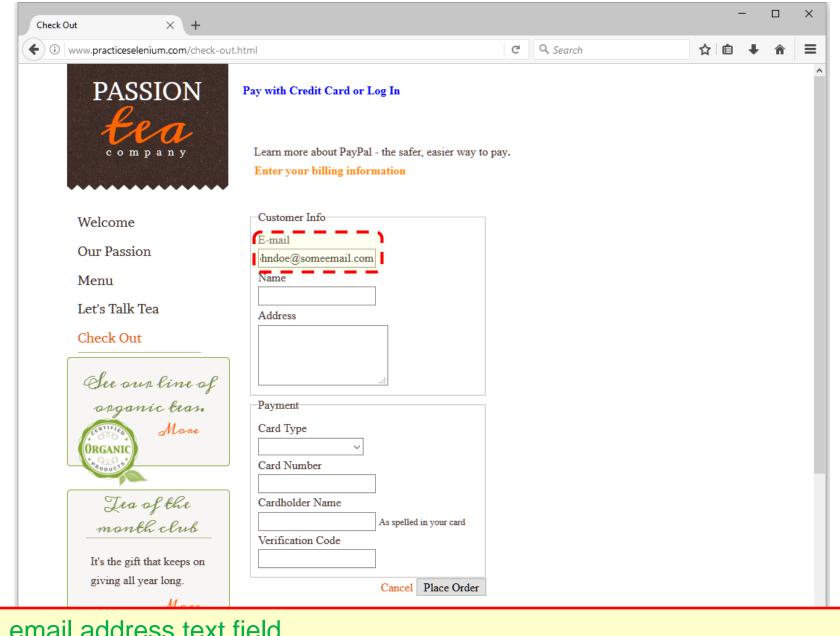
//start a new instance of Firefox and load the web page
WebDriver driver = new FirefoxDriver();
driver.get("http://www.practiceselenium.com/");



WebElement elementMenu = driver.findElement(By.linkText("Menu"));
elementMenu.click();

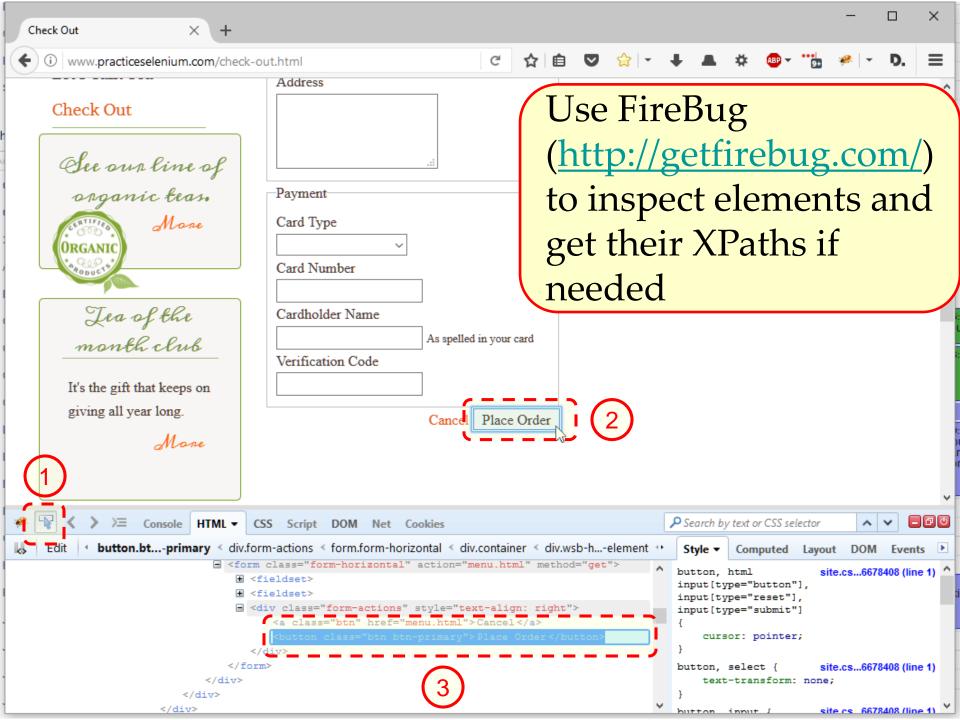


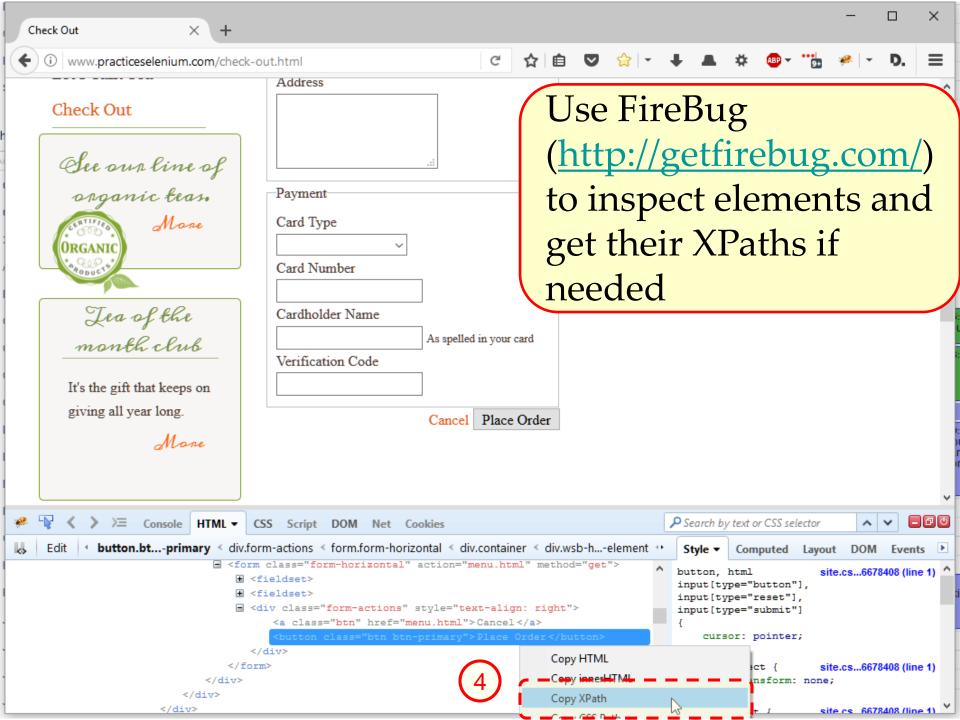
WebElement elementCheckout = driver.findElement(By.linkText("Check Out"));
elementCheckout.click();

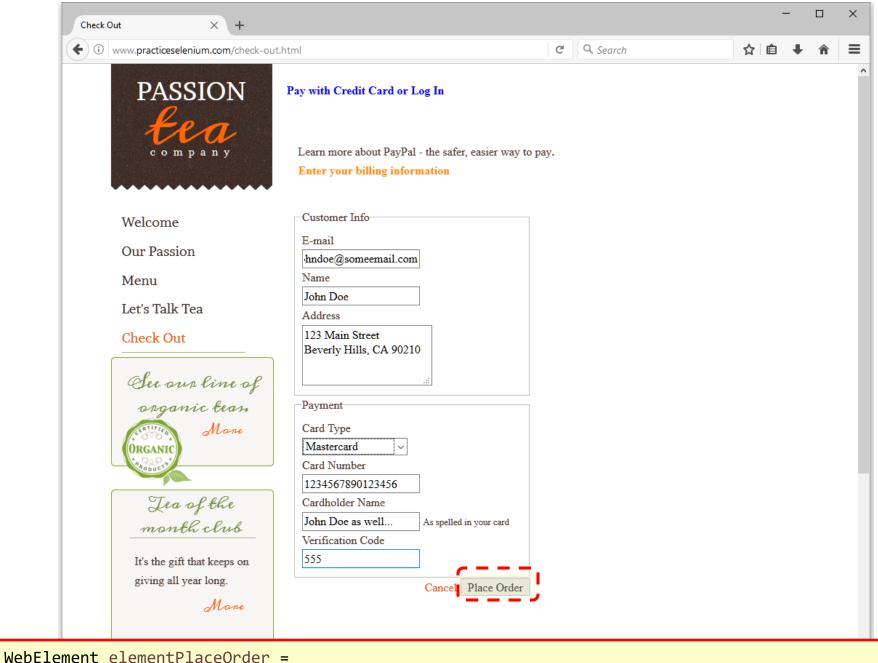


#### //fill email address text field

WebElement elementEmail = driver.findElement(By.id("email")); elementEmail.sendKeys("johndoe@someemail.com");







driver.findElement(By.xpath("/html/body/div/div[1]/div/div[1]/div/div/form/div/button"));
elementPlaceOrder.click();

#### Selenium Resources

- http://docs.seleniumhq.org/
  - Official Selenium page where you can download the Selenium WebDriver from

- http://toolsqa.com/selenium-tutorial/
  - very detailed and easy to follow tutorial for getting up to speed running with Selenium in Eclipse to test web apps.

#### Selenium Resources

- http://www.seleniumframework.com/demo-sites/
  - provides a list of websites to practices your Selenium tests

- <u>http://www.techbeamers.com/websites-to-practice-selenium-webdriver-online/</u>
  - provides a list of websites to practices your Selenium tests

## Marathon – Automating GUI Testing

- Cross-platform GUI test automation framework for:
  - Java/Swing
  - ■Java/FX
  - Web applications
- Very similar in functionality/usage with Selenium

Allows to test Java GUI apps

#### Marathon Resources

- <u>https://marathontesting.com/</u>
  - Official Marathon page
- https://confengine.com/selenium-conf 2016/proposal/2532/java-swing-java-fx-application-testing-using-selenium-webdriver
  - •1 hour video tutorial to get started
    - sample code used in video (can be easily imported as an Eclipse project) and tested:
      - https://github.com/jalian-systems/marathon-demo

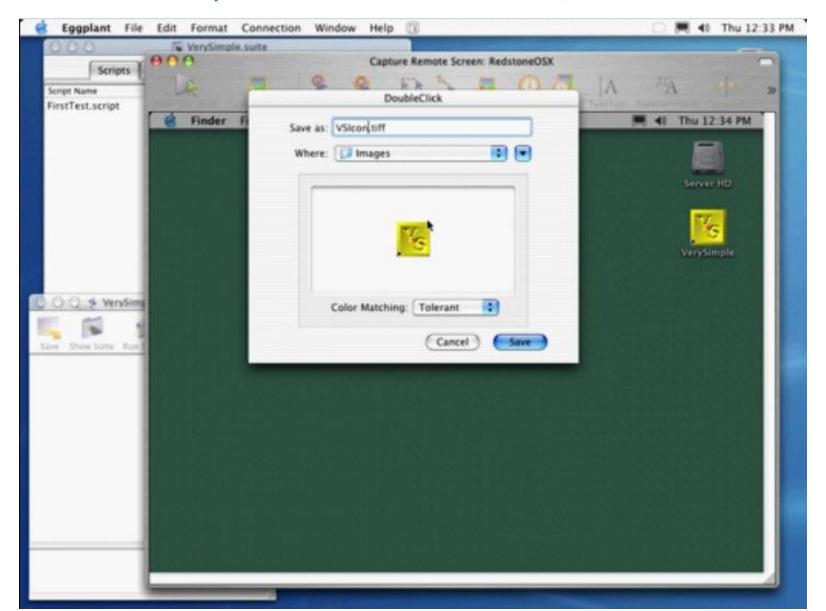
# Eggplant – Automating Web/GUI Testing

Cross-platform GUI test automation framework for any application

- It uses image analysis/recognition to identify:
  - which button to press
  - where to input text
  - where to scroll
  - etc.

https://www.testplant.com/eggplant/testing-tools/

# Short video demoing Eggplant functionality <a href="https://vimeo.com/27490840">https://vimeo.com/27490840</a>



# Acceptance Testing: Effectiveness

**1**996Jones: **25..50**%

■2003Kan (IBM): **57%** 

■CS481 indirect measurement: 40%

### **Everybody Does Acceptance Testing**

- It's likely the most commonly used defect removal activity
  - It is mandatory to do it

- It finds defects that customers actually care about!
  - ■However...

#### Acceptance Testing (AT): Warnings

- By itself, AT is inadequate
  - unless your business can tolerate shipping half your defects to customers

- Management will sometimes react to quality problems by throwing money at AT
  - The cost of removing the next defect with AT rises dramatically (and reaches a saturation point)
  - Solution: complement AT with another defect removal activity