Today

- Overview of class project topics
- Overview of the scientific method
- Paper discussion
    - Open discussion
      - Why do we need statistical tests?
      - Why is it particularly important for randomized algorithms?
      - What’s wrong with the current state of the art?
      - Significance and effect size
      - Parametric vs. non-parametric statistics
      - Internal, external, and construct validity

Class projects: overview

Logistics
- 4 students per project group
- 2 presentations (early feedback and final presentation)
  - Informal 10-15min + questions

High-level topics
- UI Design
- Static analysis
- Empirical studies
- Development of features for two real-world projects:
  - Major: mutation analysis framework
  - Defects4J: database and framework for empirical studies

Class projects: UI Design (multiple groups)

Goal: Visualization of program mutations and mutation analysis results (e.g., standalone, IDE plug-in, html, ...)

- A binary search implementation.
  - public class BinarySearch implements Iterator,unnable
  
  - public int find (int [] data, final int key)
  
  - if (key < data [0]) return -1;
  
  - for (int i = 0; i < data. length; i++)
  
  - if (key == data [i]) return i;
  
  - return -1;

- Code snippet for mutation analysis.
Class projects: UI Design (multiple groups)

**Goal:** Visualization of program mutations and mutation analysis results (e.g., standalone, IDE plug-in, html, ...)

How to visualize mutants and their results?

Class projects: Static analysis

**Goal:** Implement a data-flow analysis (e.g., in Soot), in particular a dependency analysis.

Class projects: Empirical studies (multiple groups)

**Goal for project 1:** Repeat a published experiment

**Goal for project 2:** Perform statistical analyses, in particular cross validation, on a given data set

Class projects: Development (multiple groups)

**Goal for Major:**
- Develop a mutation analyzer (standalone, Ant plug-in, Maven plug-in, IDE plug-in)

**Goals for Defects4J:**
- Develop a bug-mining infrastructure
- Develop a semi-automated review system for patches
The scientific method

Question ➔ Observations ➔ Hypothesis ➔ Predictions

The scientific method

Question ➔ Observations ➔ Hypothesis ➔ Predictions

Experiment ➔ Predictions
The scientific method

Question \rightarrow Observations \rightarrow Hypothesis

Experiment \rightarrow Predictions

Falsifiable

Data collection and analysis!

Repeatable

When do we stop and what are the next steps?
The scientific method: common mistakes

Data collection → Observations → Hypothesis → Predictions → Data Analysis

"If you torture the data long enough, it will confess." (Ronald Harry Coase)

Paper discussion


- Significance and effect size
- Parametric vs. non-parametric statistics
- Internal, external, and construct validity
- Open discussion:
  - Why do we need statistical tests?
  - Why is this important for randomized algorithms?
  - What's wrong with the current state of the art?