Faulty
Bug Localization as a Service

Gurpreet Singh

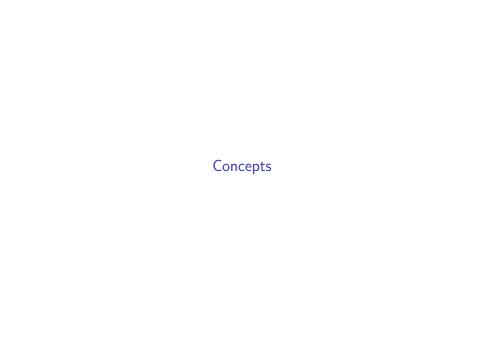2018-04-09

Background and Related Work

Preprocessing scripts

- Extracted rsf from code
- Formatted bug reports for processing
- Prepares file structure for the BugLocalization project

BugLocalization

- Used extracted bug reports and code rsf
- Produces output ranking files based on probability of containing a bug
- Written in Shell, Python and Java
- User interface was command line based with minimal input
- Only supported single operations

Concepts

RSF

- A map of relationships betweens tokens within a codebase
- A token is a keyword in the codebase (ex. method)
- Created using preprocessing scripts
- Allows us to verify tokens inside of bug reports

Bug Report

- A database entry for each bug report
- Contains a String field containing what a user wrote inside their bug report

Token Expansion

- Tokens extracted from the bug report can be expanded
- Expansion means including tokens that are referenced by the original token set

Development Objectives

Develop a complete system

- ▶ Combine scripts and java code into a easy to access complete system
- ▶ System should be able to handle new codebases and improve with new bug reports

Be able to integrate with a Github based workflow

- Fetch new issues are they are created in a repo and process them
- Push results back to the issue page so the developers can get a head start
- Allow users to authenticate using Github

Operate as a standalone service with a UI

- ▶ Work in a similar fashion to other CI tools (like travis, etc)
- ▶ Allow users control to hook into any of their repositories