EE / CprE / SE 491 – sdmay20-43
Analysis of GitLab projects using BOA
Week 9 Report

11/13/2019 - 11/19/2019
Client: Simanta Mitra
Faculty Advisor: Simanta Mitra

Team Members
Diego Realpe - Team Leader | Test Engineer
Adrian Hamill - Test Engineer
Megan Miller - Test Engineer | Website Manager
Benjamin Carland - Test Engineer
Yi-Hsien Tan - Test Engineer | Organizer | Website Manager

Weekly Overview
This week The team worked on more complex stuff regarding the usage of BOA, how queries can be custom made and how it can act as a very versatile language with many uses. The BOA team learned and presented to Dr. Mitra about Data structures present in BOA (Arrays, Queues, Maps) and how they can be used to collect and sort analysis data. The Data Pipeline team decided to discard the Scraping solution they had been working on as discussed last time and found a new program for migrating repos from GitLab to GitHub while preserving metadata information.

Past Week Accomplishment
Diego Realpe:
- General team organization and coordination
- Found a new solution for the GitLab migration problem, we decided to no longer work with GitLab and the Scraper repo
- New repo pulls down from GitLab and Uploads all the metadata onto GitHub, from there it can be easily adapted to the existing solution of parsing GitHub's data into Hadoop SequenceFiles.

Adrian Hamill:
- Started development on a script to utilize a new migration tool we've found which migrates issues and commits between gitlab and github.
- Got parts of the new pull down, push up solution to work locally. This solution seems to be the right one for our problem

- Additional perks that the commit data such as its author are not linked to the GitLab account, instead they are linked to their email so the transition from one place to the other conserves the iastate mail address they used for their commits

Megan Miller:
- Worked with team to write Boa Presentation 3 - Level 2. It breaks down the different parts of a boa script (using examples), providing a better understanding of the structure which will help majorly with writing scripts.

Benjamin Carland:
- Covered the Data structures that Boa allows you to implement in the presentation as well as the specific functions that they can carry out and why you might use them in a query. Covered syntax of more complex queries as well.

Yi-Hsien Tan
- Contributed to Boa Presentation 3 - Level 2. Brainstorming more ideas for code quality testing. Decided to come up with ideas instead of queries for this semester. Will discuss with our client about which ideas to be used before creating queries.

Individual Contribution

Team Member	Contribution	Weekly Hours	Total Hours
Diego Realpe	Research on pull down/push up solution, helped script writing	10	62
Adrian Hamill	Last attempts with Scraper, started new solution script	11	70
Megan Miller	Boa presentation & script components	8	68
Benjamin Carland	Boa Presentation and Looked into applications of FindBugs more	10	65
Yi-Hsien Tan	Boa presentation & code quality testing ideas	8	70

Pending Issue

Research about evaluating code quality to come up with good ideas for boa queries. Dr. Mitra asked about functions and aggregators. We are not completely solid on these so further research might be necessary.

Plans for Coming Week

- Data Pipeline: Adrian Hamil & Diego Realpe
  - Fix the reason why the links in the JSON metadata are returning errors
  - Make the pull down push up solution work on 309 repos and not only on personal ones
  - Figure out the best way to manipulate and pass the JSON object. Throw out junk we don't need from it
  - Still can't figure out how to handle when repositories aren't public, this needs to be addressed too

Boa language expert: Ben Carland & Megan Miller & Yi-Hsien Tan
  - Learn how to write custom functions in boa; start learning R programming language
  - Come up with more ideas for code quality testing.