Lab One

Jordan Murray

Jordan.Murray1@Marist.edu

March 9, 2021

1 Problem 1.11 - Crafting a Compiler

The technique that MOSS uses to find similarities between programs is called document fingerprinting. This technique precomputes a set of hash values after avoiding whitespace and identifies for a document and then compares these hash values to other documents. By doing so, this drastically reduces the number of comparisons to be made and helps reduce computational time as well. Any documents that have a high number of matching finger prints are flagged for review.

Moss differs from other approaches for detecting possible plagiarism. One tool that is often compared to Moss is JPlag. JPlag uses a similar method of converting files to token strings, however it compares these tokens using the Greedy String Tilling algorithm. MOSS does not directly compare substrings, to make computational time faster.

1.1 References

https://yangdanny97.github.io/blog/2019/05/03/MOSS

https://www.researchgate.net/publication/2832828$_{Finding Plagiarisms_a mong_a Set_o f Programs_w ith_J Plag}$

2 Problem 3.1 - Crafting a Compiler

Statement : Main
R-Paren : (
L-Paren :)
R-Brace : {
Statement : const
Var Decl : float payment
Assign : =
Float-exp : 384.00
Var Decl : float bal
Var Decl : int month
Assign : =

Int-expr : 0
Id : bal
Assign : =
Int-expr : 15000
While : while
RParen : (
Id : bal
BoolOp : >
int-exp : 0
L-paren :)
R-Brace : {
Print : Printf
R-paren : (
String exp : "Month: %2d Balance : %10.2f
n"
Comma : ,
Id : month
Comma : ,
Id : bal
R-paren :)
Id: bal
Assign : =
Id : bal
Intop : -
Id : payment
Intop : +
float-exp: 0.015
Intop: *
Id : bal
Id: month
Assign : =
Id : month
Intop : +
int-exp : 1
L-brace : }
L brace : }

IDs might need more information returned such as a pointer so you know if you are referencing an existing value or a new id. The string token might also need more information returned because there are variables listed/ formated a specific way so that might helpful to return as well.

3 PROBLEM 1.14 - DRAGON

The advantages of using C as a target language for a compiler are it is easy to read and understand, it is easily compiled because it is a more popular language, therefore, it will be easier to understand errors when they occur, better code will be generated so fewer compiler bugs will occur. According to the Crafting a compiler text book, "if a language cannot be easily understood, then discrepancies will arise... these will in turn lead to compilers that differ in their interpretation of a program's meaning". This means that C is a good language because there are many good compilers for this language.

4 Problem 1.6.1 - Dragon

Values assigned to
W - 13
X - 11
Y - 13
Z - 11