Lab One

Alexis Dionne

alexis.dionne1@Marist.edu

February 8, 2019

1 Crafting a Compiler

Chapter 1 Question 11

The Measure Of Software Similarity (MOSS) [SWA03] tool can detect similarity of programs written in a variety of modern programming languages. Its main application has been in detecting similarity of programs submitted in computer science classes, where such similarity may indicate plagiarism (students, beware!). In theory, detecting equivalence of two programs is undecidable, but MOSS does a very good job of ïňĄnding similarity in spite of that limitation. Investigate the techniques MOSS uses to ïňĄnd similarity. How does MOSS diïňĂer from other approaches for detecting possible plagiarism?

MOSS is the preferred method for professors to use when comparing source code, and it was developed by a couple of professors at Stanford University. MOSS is white space insensitive, and it uses k-gram hashing to determine the similarity level of strings. The professors used this to develop "Winnowing" which finds matches at least as long as the "guaranteed threshold" entered by the user and ignores strings that are shorter than the "noise threshold". The program ignores words like "the" and compares large strings of text to determine if there are only small changes like a variable name. It however is not very smart when it comes to comparing differences in logical structures. For example, if the original uses a for loop and the suspect uses a while loop, it will go mostly undetected.

Chapter 3 Question 1

Assume the following text is presented to a C scanner:

```
main() {
        const float payment = 384.00;
        float bal;
        int month = 0;
        bal= 15000;
        while (bal>0){
                printf("Month: %2d Balance:      %10.2f\n", month, bal);
                bal=bal-payment+0.015* bal;
                month=month+1;
        }
}
```

What token sequence is produced? For which tokens must extra information be returned in addition to the token code?

id open paren close paren open bracket
type(constant) id
type id
id operation integer
while open paren id operation integer close paren open bracket
id
id operation id operation id operation decimal operation id
id operation id operation integer
close bracket
close bracket

2 Dragon

1.1.4

A compiler that translates a high-level language into another high-level language is called a *source-to-source* translator. What advantages are there to using C as a target language for a compiler?
The C language has a large number of compilers available for almost any hardware.

1.6.1

For the block-structured C code of Fig. 1.13(a), indicate the values assigned to *w, x, y, and z.*
w = 13
x = 9
y = 13
z = 9