Matthew Musich

Lab 1

1.11)

MOSS uses a few techniques to find similarities between text files. MOSS does this by using a few different algorithms that can determine similarities based on a percentage. Overall the main technique that is done is called Fingerprinting via k-grams. K-grams are a consistent amount of characters in the files that are repeating, that are then hashed and compared using an algorithm that compares the hashes with other files. This can be done with a Karp-Rabin algorithm or and All-to-all technique. All of this leads into the winnowing algorithm, which takes fingerprinting and compares it alongside positions within a document. This is all done while avoiding whitespace and comments within code. MOSS is very different from other plagiarism checkers because code isn't something that can be easily detectable like an exact copy from something like the English language. Where comparing a sequence of words can be done just by checking characters. Code is something that can be abstract, yet have the same function as something that might be "worded" differently.

3.1)

```
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;
	}
}
```

```
(program main)
(symbol leftParen)
(symbol RightParen)
(symbol leftbracket)
(symbol leftParen)
(symbol leftParen)
(type const)
(type float)
(id payment)
(op equals)
(decimal 384.00)
(symbol semicolon)
```

(type float)
(id bal)
(symbol semicolon)
(type int)
(id month)
(op equals)
(num 0)
(symbol semicolon)
(id bal)
(op equals)
(num 15000)
(symbol semicolon)
(statement while)
(symbol leftParen)
(id bal)
(op greaterthan)
(num 0)
(symbol RightParen)
(symbol leftbracket)
(statement printf)
(symbol leftParen)
(symbol leftQuote)
(string “Month: %2d Balance: %10.2f\n”)\
(symbol comma)
(id month)
(symbol comma)
(id bal)
(symbol RightParen)
(symbol semicolon)
(id bal)
(op equal)
(id bal)
(op subtract)
(id payment)
(op add)
(decimal 0.015)
(op times)
(id bal)
(symbol semicolon)
(id month)
(op equals)
(id month)
(op plus)

(num 1)
(symbol semicolon)
(symbol rightbracket)
(symbol rightbracket)

Exercise 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?

C is a good target language for a compiler because it is easily ran in many different environments and OS's. It is a relatively low level language that has many compilers that are very well made which allows it to be fairly well optimised. Lastly it is a language that is hardware independent since the compilers that work with it generally take care of any hardware issues.

Exercise 1.6.1 : For the block-structured C code of Fig. 1.13(a), indicate the values assigned to w, x, y, and z.

```
int w, x, y, z;
int i = 4; int j = 5;
{           int j = 7;
            i = 6;
            w = i + j;
          }
x = i + j;
{           int i = 8;
           y = i + j;
           }
z = i + j;
```

w = 13
x = 11
y = 13
z = 11