University of Toronto
Department of Electrical and Computer Engineering

Final Examination

CSC467 - Compilers and Interpreters

Fall 2010

Duration: 2.5 hours

Print your name neatly in the space provided below.

Name:	
ID Number:	

Grade:

Question	Mark
1	
2	
3	
4	
5	
Total	

1. **(10 marks):** Clearly mark T (True) or F (False) for the following statements:

- T / F) Any regular expression can be recognized by a program generated by YACC or BISON.
- T / F) A software implementation of a deterministic finite automaton (DFA) with a large transition table is slow on a modern microprocessor.
- T / F) Lexical analysis occupies only a negligible fraction of total runtime in a compiler.
- T / F) YACC produces a top-down parser, and a recursive descent parser is a bottom up parser.
- T / F) Recursive descent parser requires a left-recursive grammar.
- T / F) A program conforming to a context free grammar cannot be recognized by a NFA.
- T / F) Symbol tables are used for syntax checking.
- T / F) An abstract syntax tree is usually used for code optimization.
- T / F) A basic block is a collection of any instructions convenient for program analysis.
- T / F) A program that passes semantic checking is guaranteed to be correct at runtime.

2. **(10 marks):** Consider the following program that computes Fibonacci numbers.

```
int fib( int n ) {
     if( n == 1 || n == 2 ) return 1;
L
     return fib( n-1) * fib( n-2 );
     }
```

Draw the snapshot of runtime stack frame for a call of fib(4), before fib(1) is returned.

3. **(30 marks):** Consider the following assembly code, where t1-t8 are temporaries, and u, v, w, x, y, z, are memory addresses.

```
load t1 [u]         ; load from address u to register r1
load t2 [v]
add t3, t1, t2      ; t3 = t1 + t2
sub t4, t1, t2      ; t4 = t1 – t2
store t3 [w]
store t4 [x]
load t5, [y]
load t6, [z]
add t7, t5, t6
add t8, t5, t4
```

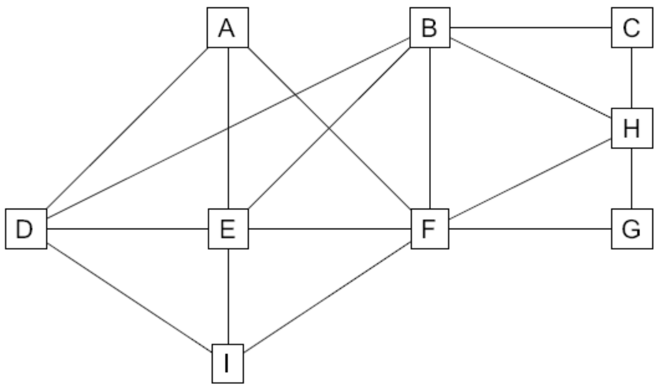
a) Assuming the program is run on a VLIW processor with unlimited number of hardware resources, and the latency of each instruction is 1 cycle. What is the minimal number of cycles the program can complete? Substantiate your answer by showing a schedule of the instructions.

b) Repeat a), assuming there could be at most 1 memory instruction executed per cycle.

c) Repeat a), assuming that by program analysis, it can be proved that

- $y \neq w$, $y \neq x$;
- $z \neq w$, $z \neq x$.

4. **(30 marks):** The figure below is a interference graph of nine values (A through *I*) for register allocation.



a) Determine the coloring order of nodes using the algorithm taught in class. Substantiate your answer by drawing the remainder graph in each step, as well as the criterion you use in each step.

b) Given the coloring order developed in a), determine the register assignment. Substantiate your answer by drawing the remainder graph in each step, and state how you choose the color of the node in each step.

c) Can this graph be colored with at most **three** colors? Is yes, give an color assignment. If not, explain why.

5. **(20 marks):** Many software programs have security vulnerabilities where user input can directly manipulate critical data. One method to prevent such vulnerabilities is called **taint analysis**, which determines whether a data value is **tainted**, in other words, dependent, possibly transitively, on the user input value. For the sake of simplicity, consider a simple language consisting of only the following kinds of statements. In addition, each variable is assigned only once.

- x = input ();
- x = <constant>;
- x1 = x2 + x3;
- if(x) goto L;

a) find the set of tained values in the following code.

(1) a = 10;
(2) b = input ();
(3) c = b + a;
(4) d = c + 10;
(5) e = 10;
(6) f = c + c;

b) Devise an algorithm to find the set of all tainted values in a program of the above language. Carefully state your assumptions and intermediate representation of the program. Also carefully state how you initialize the values of data structures used in your analysis.