

Instructions: This is an open book, open note, 110 minute exam. You are not allowed to use your homework assignments or your project. There are 5 questions worth a total of 100 points. You must answer all of the questions. Put your name on all pieces of paper that you turn in. Show all of your work, write legibly and good luck.

1) (40 points) You are to add a new language element to Classcal and enhance your compiler appropriately. The language should support a multi-way conditional statement. We'll call it a "switch" statement.

Syntactically a switch looks like the following:

```
switch <expr> is
  case <literal>..
```

The cases can consist of ranges of literals, a single literal value is considered a trivial range, and a statement that is to be evaluated if the expressions value lies within the range.

Semantically, this is equivalent to a series of nested if-then-else's with conditions that correspond to expression/literal comparisons. Thus, when a case's statement executes the next statement to be executed is the one following the switch statement itself. The literal ranges for a case statement must be disjoint. There is a required default case that handles any values not matched by an explicit case. There are obviously type consistency requirements for literals and expressions.

Describe in detail the changes you would make to your compiler to:

- a) Enhance the scanner to handle the new lexical elements.
- b) Enhance the parser to handle the new syntactic elements.
- c) Enhance the translation schemes to handle the new semantic elements, including defining any new AST nodes necessary.
- d) Enhance the code generator to handle the new AST nodes.

If you need any assumptions, e.g., about string comparisons, just state them explicitly and proceed.

2) (10 points) It has been said that compilers are one area of computer science where theory has been successfully applied. Justify this statement by describing at least three ways that theory has been applied. Be as detailed as possible.

3) (15 points) Consider the following program:

```
program drawIRs;
var b : boolean;
    s : string;

procedure flip(var b : boolean);
begin
    b := not b
end { flip } ;

procedure conditionalPrint(x : integer);
    var l : integer;
begin
    l := x;
    if b then begin
        writeString(s);
        while l > 0 do begin
            writeInteger(l);
            l := l - 1
        end { while }
    end { if };      {<----- Point A }
    writeLn
end { conditionalPrint };

begin
    s := "The values are ";
    b := FALSE;
    conditionalPrint(7);
    flip(b);
    conditionalPrint(7)
end { drawIRs }.
```

Based on the internal representations discussed in class:

- Draw a picture of the symbol table at **Point A** in the parse.
- Draw a picture of the AST for this program.
- Draw a picture of the CFG for the conditionalPrint procedure.

Include attributes and values for nodes/edges where appropriate.

4) (20 points) We know that all regular languages are also context-free. We know that there are context-free grammars that are not LL(1). This question is about the relationship between regular languages and LL(1) grammars. In class we discussed right linear grammars, also called regular grammars. They have a pretty simple form with rules that look like:

$$\begin{array}{ccc} \langle A \rangle & \rightarrow & a \langle B \rangle \\ & & | \quad a \end{array}$$

For the regular expression $a; (b;c)^*; a$

- a) Give a regular grammar for this expression
- b) Compute FIRST, FOLLOW for this grammar
- c) Is it in LL(1) form? If not, can you convert it to LL(1) form?
- e) If yes, do so and build the predictive parse table.

5) (15 points) Classcal doesn't allow for user defined types. If we were to allow users to define their own types we have two options for judging type equivalence : by type name or by type structure. Assume a type constructor, e.g., typedef, has been added to Classcal.

- a) Give a program that is correct under structural equivalence but not under name equivalence.
- b) Sketch out how you would implement user defined types under both approaches.
- c) Which one is simpler for the compiler writer?