Midterm

CIS 756, Fall 1998

October 30, 1998

Read the whole exam first (there are 6 pages) and plan your time. You have 48 minutes to complete all three questions. There are a total of 100 points. The exam is open book and open notes, but closed neighbors. Good luck!

Name:

Student-ID:

1. Scanning (30 pts)
   Consider the following format for floating point literals. Assume a fraction part consisting of a string of one or more decimal digits with an optional decimal point (period) that can be added immediately before or after *any* of the digits. In addition, an optional exponent can be appended to the fraction part, consisting of an `e` or `E`, followed by an optional sign, followed by one or more decimal digits. A floating point literal must contain either a decimal point, or an exponent, or both.

   Examples: `.0`, `0.`, `0.1`, `.01e-01`, `01E00`

   (a) (15 pts) Draw an NFA for this construct.

   (b) (15 pts) Give a regular expression for it. You may use ϵ.

2. Context free grammars, LL parsing (30 pts)
Consider the following grammar:

$E \rightarrow E + T \mid T$
$T \rightarrow \texttt{id} \mid \texttt{id}() \mid \texttt{id}(L)$
$L \rightarrow E; L \mid E$

The nonterminals are E, T, and L. The terminals are +, `id`, (,), ;. The start symbol is E.

(a) (15 pts) Modify the grammar such that it can be parsed by an LL(1) parser.

(b) (15 pts) Show Nullable, FIRST, and FOLLOW and derive the LL(1) parse table for the modified grammar.

3. Context free grammars, LL and LR parsing (40 pts)
Consider the following simple context free grammars:

Grammar G_1	Grammar G_2
$G \rightarrow A\$$	$G \rightarrow A\$$
$A \rightarrow \epsilon$	$A \rightarrow \epsilon$
$A \rightarrow bAb$	$A \rightarrow Abb$

The start symbols are G, the nonterminals are G and A, and the terminal symbols are b and \$. Note that these grammars generate the same language: strings consisting of even numbers of b symbols (including zero of them).

(a) (15 pts) Attempt to show a shift-reduce parse of the input string *bbbb* for a parser for grammar G_1. Show the contents of the stack, the input, and the actions (in the style of Figure 3.18 on Page 58 but without the subscripts for the parse states).
Indicate any conflicts and describe why they are conflicts. Is G_1 LR(1)? Is it LR(0)?

(b) (15 pts) Attempt to show a shift-reduce parse of the input string *bbbb* for a parser for grammar G_2. Show the contents of the stack, the input, and the actions (in the style of Figure 3.18 on Page 58 but without the subscripts for the parse states).

Indicate any conflicts and describe why they are conflicts. Is G_2 LR(1)? Is it LR(0)?

(c) (5 pts) Indicate whether G_1 and G_2 are LL(1). You don't need to construct their LL(1) parse tables, but you may argue from other properties.

(d) (5 pts) Of the language classes we have discussed in class, what is the *smallest* category into which $L(G_1)$ fits? Justify your answer. [Hint: This is a trick question!]