PRINT FAMILY NAME:__________________________

Examination 1
COT 6315 - Formal Languages and Computation Theory
24 May 1999

Instructions:

- **Read all these instructions; failure to follow instructions will result in loss of points.**
- Do not start the examination until instructed to do so.
- This test is closed book, closed notes, but you are permitted one 8.5" by 11" crib sheet (two sides).
- **Answer all questions from Part A (first 4 questions) and then any two questions from Part B (last 3 questions) below.**
- Show your work.
- **Start the answer to each question on a new page** (i.e., do NOT put the answer to more than one question on the same page).
- **Assemble your answers in numerical order of the questions when you submit them.**
- **Leave a 1" square in the upper left corner for a staple.**
- Be sure to include your name on your answer sheets.
- You have 120 minutes to complete this examination.
- **Read and sign the following statement.** You may write this on your exam and sign it there if you wish to take the exam questions home with you today. Do not discuss this exam with anyone in this course who has not yet taken this exam.

On my honor, I have neither given nor received unauthorized aid on this examination, and I will not discuss the contents of this examination with any student who has not yet taken this examination.

Signed:

Part I

Answer **all four (4)** questions.

1. Given regular expressions r and s, for each of the pairs of regular expressions given below, determine whether the languages they represent are the same or not. If so, prove it. If not, then demonstrate it, showing non-inclusion in all directions possible (e.g., if $L_1 \not\subseteq L_2$, then provide w in $L_1 \setminus L_2$; if $L_2 \not\subseteq L_2$ also, then provide w' in $L_2 \setminus L_1$).

   (a) (3) $r^*s(r+s)^*, \quad (r+s)^*sr^*$

   (b) (3) $(rs+r)^*r, \quad r(sr+r)^*$

   (c) (2) $r^* + s^*, \quad rs^* + sr^* + s^*r + (r^*s)^*$

   (d) (2) What must be true of the language represented by a regular expression that does not include Kleene closure ($*$)? Why?

2. For each of the languages below, classify it as regular, context free, or neither, and prove your answer (e.g., if L is context free but not regular, you must prove it is not regular but is context free). Proof sketches of constructions and for pumping lemma application are OK.

   (a) (5) $L_1 = \{w_1\#w_2 \mid w_1, w_2 \in \{a,b\}^*,\ \#_a(w_1) = \#_a(w_2) \text{ and } \#_b(w_1) = \#_b(w_2)\}$

   (b) (5) $L_2 = \{wxw^R \mid w, x \in \{a,b\}^+\}$

   (c) (5) $L_3 = \{0^i1^j \mid i < j\}$

   (d) (5) $L_4 = \{x \in \{[a,b,c]\}^* \mid a,b,c \in \{0,1\},\ x = x_1x_2...x_n,\ x_i = [a_i,b_i,c_i]$, and interpreting $a = a_1a_2...a_n$, $b = b_1b_2...b_n$ and $c = c_1c_2...c_n$ as binary numbers with x_1 the most significant bits, $c = a + b\}$
   ($e.g.$, $[0,1,1][1,0,1][0,0,1][1,1,1][1,1,0] \in L_4$ since $01011_2 + 10011_2 = 11110_2$, while $[1,0,0][0,1,1] \not\in L_4$ since 1 $01_2 \neq 01_2$.)

3. (10) Recall that the Pumping Lemma for CFLs gives in its proof a value for the pumping constant for a given CFL represented by CFG $G = \ < V, T, R, S >$ as $p_G = b^{|V|+2}$, where $b = max\{|u| \mid A \rightarrow u \in R\}$ is the length of the longest right hand side over all productions in G, and $|V|$ is the size of the variable set of G. Provide and argue the correctness of a tighter upper bound on p_G. What is the bound you obtain for a CFG in CNF?

4. (10) Give a formal PDA recognizing the following language L and prove it accepts exactly the words in L. A state diagram is sufficiently formal as a description.

$$L = \{a^ib^jc^k \mid i = j \text{ or } i = k \text{ or } j = k\}.$$

Part II

Answer **any two (2)** questions.

5. (25) Prove that for any PDA $P = < Q, \Sigma, \Gamma, \delta, q_0, F >$, there exists another PDA $P' = < Q', \Sigma, \Gamma', \delta', q_0', F' >$, such that $L(P) = L(P')$ and

   (a) $|F'| = 1$,

   (b) P' only accepts with an empty stack

   (c) δ' only contains push and pop moves (i.e., if $(q, t) \in \delta'(p, a, t')$ then either $t = \epsilon$ or $t' = \epsilon$ but not both)

   In other words, the stack size always increases or decreases, and if it increases, it does so without regard to the top of stack symbol.

6. (25) Prove that it suffices to reason about Regular Languages with a binary alphabet by proving the following. For any DFA $M =< Q, \Sigma, \delta, q_0, F >$, with $|\Sigma| > 2$, there exists another DFA $M' =< Q', \Sigma', \delta', q_0', F' >$, with $|\Sigma'| = \{0, 1\}$, and a mapping of the symbols of Σ into bit strings, $f : \Sigma \to \{0, 1\}^n$ where $n = \lceil log_2(|\Sigma|) \rceil$, with extensions $\tilde{f} : \Sigma^* \to \{0, 1\}^*$ defined on strings to be $\tilde{f}(a_1 a_2 \ldots a_n) = f(a_1) f(a_2) \ldots f(a_n)$ and $\hat{f}$ defined on languages to be $\hat{f}(L) = \{\tilde{f}(w) | w \in L\}$, such that $\hat{f}(L(M)) = L(M')$. In other words, M' accepts $\tilde{f}(w)$ iff M accepts w.

7. (25) Consider the equation $X = rX + s$, where r and s are regular expressions (taken here to be the languages they represent), X is a language (or its regular expression), rX is the concatenation of the languages represented by r and X, and $+$ is union. Assuming $\epsilon \notin r$, what is X? Show X is unique. What if $\epsilon \in r$?