CS 373, Spring 2009
Mock Midterm 2

Problem 1: Short Answer (12 points)

Answer "yes" or "no" to the following questions. No explanations are required. (All the strings in this problem use the same, fixed alphabet.)

(a) Is the language $\left\{ ww^R w \;\middle|\; w \in \{\mathtt{a},\mathtt{b}\}^* \right\}$ a context-free language?

(b) If L is a non-regular language over Σ^*, and h is a homomorphism, then $h(L)$ must also be non-regular. Is this statement correct?

(c) Suppose all the words in language L are no more than 1024 characters long. Then L must be regular. Is this statement correct?

(d) If L_1 and L_2 be two languages, the ***xor*** of the two languages is

$$L_1 \oplus L_2 = (L_1 \setminus L_2) \cup (L_2 \setminus L_1) = \left\{ w \;\middle|\; \begin{array}{c} w \in L_1 \text{ and } w \notin L_2 \\ \text{or} \\ w \notin L_1 \text{ and } w \in L_2 \end{array} \right\}.$$

If L_1 and L_2 are both context-free, then $L_1 \oplus L_2$ must also be context-free. Is this statement correct?

(e) If L is a language, its prefix language is

$$P(L) = \left\{ w \;\middle|\; \text{there exists } x \text{ s.t. } wx \in L \right\}.$$

If L is context-free, then the language $P(L)$ is context free. Is this statement correct?

Problem 2: Grammar design (8 points)

Let $\Sigma = \{\mathtt{a},\mathtt{b}\}$.

Let $J = \left\{ w_1\#w_2\#\ldots\#w_{n-1}\#w_n \;\middle|\; n \geq 2, w_i \in \Sigma^* \text{ for all } i, \text{ and for some } i, |w_i| = |w_{i+1}| \right\}$

In other words, an element of J is a list of at least two strings of a's and b's, separated by #'s. In the list, some pair of adjacent strings have the same length. E.g. J contains `b#aa#bbb#aba` but not `a#bbb#b`.

Give a context-free grammar whose language is J. Be sure to indicate what its start symbol is.

Problem 3: RA design (8 points)

Let

$$J = \left\{ w \in \{\mathtt{a}, \mathtt{b}\}^* \;\middle|\; \begin{array}{c} w \text{ contains only } \mathtt{a}\text{'s} \\ \text{or} \\ w \text{ has an equal number of } \mathtt{a}\text{'s and } \mathtt{b}\text{'s} \end{array} \right\}.$$

For example, J contains ϵ, aaa, and aabbba. But abbba is not in J.

Give the state diagram for a RA whose language is J. Include brief comments explaining the design of your RA, to help us understand how it works.

Problem 4: Short Answer II (8 points)

The answers to these problems should be short and not complicated.

(a) Suppose we know that the language $B = \left\{ \mathtt{a}^n \mathtt{b}^n \mathtt{c}^n \;\middle|\; n \geq 2 \right\}$ is not context-free. Let L be the language $L = \left\{ \mathtt{a}^n \mathtt{b} \mathtt{c}^n \mathtt{b} \mathtt{d}^n \;\middle|\; n \geq 0 \right\}$. Prove that L is not context-free using closure properties and the fact that B is not context-free.

(b) Define what it means for a grammar G to be in Chomsky Normal Form (CNF).

A grammar is in Chomsky normal form every rule has one of the following forms:

Problem 5: Pumping Lemma (8 points)

Suppose $\Sigma = \{\mathtt{a}, \mathtt{b}\}$ and let $L = \left\{ \mathtt{a}^n w \;\middle|\; w \in \Sigma^* \text{ and } |w| = n \right\}$. That is, L contains even-length strings whose first half contains only a's. Prove that L is not regular by filling in the missing parts of the following pumping lemma proof.

Suppose that L were regular. Let p be the constant given by the pumping lemma.

Consider the string $w_p =$

Because $w_p \in L$ and $|w_p| \geq p$, there must exist strings x, y, and z such that $w_p = xyz$, $|xy| \leq p$, $|y| > 0$, and $xy^i z \in L$ for every $i \geq 0$.

Since

is *not* in L, we have a contradiction. Therefore, L must not have been regular.

Problem 6: Formal notation (8 points)

Let $(M, main, \{(Q_m, \Sigma \cup M, \delta_m, q_0^m, F_m)\}_{m \in M})$ be an RA recognizing the language L, here L is defined over an alphabet Σ. Define the language L' to be

$$L' = \left\{ x\#y \;\middle|\; x, y \in L \right\}.$$

Describe how to construct an RA recognizing L'. (You can safely assume that $\#$ is not in $\Sigma \cup M$.)

(a) Describe the ideas behind your construction in words and/or pictures.

(b) Give the details of your construction in formal notation. That is, for the new RA recognizing L', specify the set of states and modules, the initial and final states, the details of the transition function. The input alphabet for the new machine will be $\Sigma \cup \{\#\}$.

Problem 7: Induction (8 points)

Let $\Sigma = \{\mathtt{a}, \mathtt{b}\}$. Given any string w in Σ^*, let $A(w)$ be the number of a's in w and $B(w)$ be the number of b's in w.

Suppose that grammar G has the following rules:

$$\mathsf{S} \rightarrow \mathtt{aSb} \mid \mathtt{aSS} \mid \mathtt{ab},$$

where S is the start symbol. Use induction on the derivation length to prove that $A(w) \geq B(w)$ for any string w in $L(G)$.