CS 373, Spring 2009
Exam 2, 7-9pm, April 2, 2009

INSTRUCTIONS (read carefully)

- Fill in your name, netid, and discussion section time below. Also write your netid on the other pages (in case they get separated).

NAME:

NETID: | **DISC:**

- There are 7 problems, on pages numbered 1 through 8. Make sure you have a complete exam.
- The point value of each problem is indicated next to the problem, and in the table below.
- Points may be deducted for solutions which are correct but excessively complicated, hard to understand, or poorly explained.
- The exam is designed for slightly over one hour, but you have the full two hours to finish it.
- It is wise to skim all problems and point values first, to best plan your time.
- This is a closed book exam. No notes of any kind are allowed. Do all work in the space provided, using the backs of sheets if necessary. See the proctor if you need more paper.
- Please bring any apparent bugs to the attention of the proctors.
- After the midterm is over, discuss its contents with other CS 273 students **only** after verifying that they have also taken the exam (e.g. they aren't about to take the conflict exam).

Problem	Possible	Score
1	12	
2	8	
3	8	
4	8	
5	8	
6	8	
7	8	
Total	60	

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 $\left\{ \mathtt{a}^i \mathtt{b}^{i+j} \mathtt{c}^j \;\middle|\; i, j \geq 0 \right\}$ a context-free language?

Yes: ☐ **No**: ☐

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

Yes: ☐ **No**: ☐

(c) Suppose L is a language over $\{\mathtt{0}\}^*$, such that if $\mathtt{0}^i$ is in L then $\mathtt{0}^{i+2}$ is in L. Then L must be regular. Is this statement correct?

Yes: ☐ **No**: ☐

(d) If L_1 and L_2 be two languages. If L_1 and L_2 are both context-free, then $L_1 \setminus L_2$ must also be context-free. Is this statement correct?

Yes: ☐ **No**: ☐

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

$$S(L) = \left\{ w \;\middle|\; \text{there exists } x \in L \text{ s.t. } w \text{ is a subsequence of } x \right\}.$$

A word w is a ***subsequence*** of x, if one can delete $|w| - |x|$ characters from x and get w. For example, $w = \mathtt{abc}$ is a subsequence of $x = \mathtt{cbacbacba}$ since $\mathtt{cb\underline{a}c\underline{b}a\underline{c}ba}$.

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

Yes: ☐ **No**: ☐

(f) Consider a ***parallel recursive automata*** $\mathtt{M}$ – it is made out of two recursive automatas $\mathtt{N}_1$ and $\mathtt{N}_2$ and it accepts a word w if both $\mathtt{N}_1$ and $\mathtt{N}_2$ accepts w. The parallel recursive automata $\mathtt{M}$ might accept a language that is not context-free. Is this statement correct?

Yes: ☐ **No**: ☐

Problem 2: Grammar design (8 points)

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

$$J = \left\{ w \;\middle|\; \#_{\mathtt{a}}(w) = \#_{\mathtt{b}}(w) \text{ or } \#_{\mathtt{b}}(w) = \#_{\mathtt{c}}(w) \right\},$$

where $\#_z(w)$ is the number of appearances of the character z in w. For example, the word $x = \mathtt{baccacbbcb} \in \mathsf{L}(J)$ since $\#_{\mathtt{a}}(x) = 2$, $\#_{\mathtt{b}}(x) = 4$, and $\#_{\mathtt{c}}(x) = 4$. Similarly, the word $y = \mathtt{abbccc} \notin \mathsf{L}(J)$ since $\#_{\mathtt{a}}(y) = 1$, $\#_{\mathtt{b}}(y) = 2$, and $\#_{\mathtt{c}}(y) = 3$.

Give a context-free grammar whose language is J. Be sure to indicate what its start symbol is. (Hint: First provide a CFG for the easier language $K = \left\{ w \in \{\mathtt{a}, \mathtt{b}\}^* \;\middle|\; \#_{\mathtt{a}}(w) = \#_{\mathtt{b}}(w) \right\}$ and modify it into the desired grammar.)

Problem 3: RA design (8 points)

Let

$$J = \left\{ w \in \{\mathtt{a},\mathtt{b},\mathtt{c},\mathtt{d}\}^* \;\middle|\; \begin{array}{l} w = \mathtt{a}^i\mathtt{b}^i\mathtt{c}^n\mathtt{d}^n \text{ or} \\ w = \mathtt{c}^n\mathtt{a}^i\mathtt{b}^i\mathtt{d}^n \\ \text{ for some } i,n \geq 0 \end{array} \right\}.$$

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

$$L = \left\{ w \in \{\mathtt{a}, \mathtt{b}, \mathtt{c}, \mathtt{d}, \mathtt{e}, \mathtt{f}\}^* \;\middle|\; \#_{\mathtt{a}}(w) = \#_{\mathtt{d}}(w) = \#_{\mathtt{f}}(w) \right\},$$

where $\#_z(w)$ is the number of appearances of the character z in w. Prove that L is not context-free using closure properties and the fact that B is not context-free.

(b) Let w be a word of length n in the language generated by a grammar $\mathcal{G}$ which is in Chomsky Normal Form (CNF). First, how many internal nodes does the parse tree for w using $\mathcal{G}$ has?

Secondly, assume that $\mathcal{G}$ has k variables, and $n = |w| > 2^k$. Is the language $\mathsf{L}(\mathcal{G})$ finite or not? Justify your answer.

Problem 5: Not regular (8 points)

Suppose $\Sigma = \{\mathtt{a}, \mathtt{b}\}$ and let $L = \left\{ ww \;\middle|\; w \in \Sigma^* \right\}$. That is, a word in L is a concatenation of two copies of some string. Provide a direct (and short) proof that L is not regular (a proof not using the pumping lemma is preferable).

Problem 6: Finite or not. (8 points)

Let $\mathcal{G}$ be a context-free grammar over $\Sigma = \{\mathtt{a}, \mathtt{b}\}$, and let m be some fixed number.

(i) Prove that $L_{\geq m} = \left\{ w \in \mathsf{L}(\mathcal{G}) \;\middle|\; |w| \geq m \right\}$ is a context free language.

(This is a very short proof of this – your answer should be at most 40 words. Longer answers would get no points.)

(ii) Assume that (i) holds, and furthermore, one can compute the grammar of $L_{\geq m}$ given L and m. Describe an algorithm that decides whether the language L is finite or not. Prove that your algorithm works.

Problem 7: Proof. (8 points)

Consider the grammar

$$\Longrightarrow \quad \mathsf{S} \rightarrow \mathtt{a}\mathsf{S}\mathtt{b} \mid \mathsf{T}, \qquad \mathsf{T} \rightarrow \mathtt{c}\mathsf{T} \mid \epsilon.$$

Prove that the language of this grammar is

$$L = \left\{ \mathtt{a}^n \mathtt{c}^k \mathtt{b}^n \;\middle|\; n, k \geq 0 \right\}.$$

Your proof must be formal, correct and ***short***.

Hint: Prove (by induction, naturally) exactly what are the words that can be derived by a tree with i internal nodes, with T as the root (i.e., these are the strings that can be derived from T in i derivation steps). Next, prove a similar claim about all the words that can be generated by a parse tree with i internal nodes, having S as the root.