CS 273: Intro to Theory of Computation, Spring 2008
Problem Set 9
Due Tuesday, March 25th, 10am

This homework contains five problems. Please submit each on a **separate sheet of paper**. Turn in your homework at Elaine Wilson's office (3229 Siebel).

1. CHOMSKY NORMAL FORM.

   (a) Remove nullable variables from the following grammar (with start symbol S):

   $$\begin{aligned} \mathsf{S} &\rightarrow \mathtt{a}\mathsf{T}\mathtt{a} \mid \mathtt{b}\mathsf{B}\mathtt{b} \mid \mathsf{BB} \\ \mathsf{T} &\rightarrow \mathsf{C} \\ \mathsf{B} &\rightarrow \mathsf{S} \mid \mathsf{T} \\ \mathsf{C} &\rightarrow \mathsf{S} \mid \epsilon \end{aligned}$$

   (b) This grammar (with start symbol S) has no nullable variables. Generate its Chomsky normal form.

   $$\begin{aligned} \mathsf{S} &\rightarrow \mathsf{TSB} \\ \mathsf{T} &\rightarrow \mathtt{a}\mathsf{TS} \mid \mathtt{a} \\ \mathsf{B} &\rightarrow \mathsf{S}\mathtt{b}\mathsf{S} \mid \mathsf{T} \mid b \end{aligned}$$

2. CONTEXT FREE LANGUAGE CLOSURE PROPERTIES.

   We know (example 2.37 in Sipser) that the following language is not context free:

   $$C = \left\{ \mathtt{a}^i \mathtt{b}^j \mathtt{c}^k \,\middle|\, 0 \leq i \leq j \leq k \right\}.$$

   Using closure properties of context-free languages, and the fact that C is not context-free, prove that the following languages are not context free:

   (a) $J = \left\{ \mathtt{a}^i \mathtt{b}^j \mathtt{c}^{k-1} \,\middle|\, 1 \leq i \leq j \leq k \right\}$

   (b) $K = \left\{ \mathtt{a}^i \mathtt{c}^j \mathtt{b}^k \mathtt{c}^n \,\middle|\, 0 \leq i \leq j \leq k, 0 \leq n \right\}$

3. Grammar-based induction.

Let G be the grammar with start symbol S, terminal alphabet $\Sigma = \{a, b\}$ and the following rules:

$S \rightarrow aX \mid Y.$

$X \rightarrow aS \mid a.$

$Y \rightarrow bbY \mid aa.$

Claim version 1. *If w is a string in $L(G)$, then w contains an even number of* a*'s.*

It's actually easier to prove the following, stronger and more explicit claim:

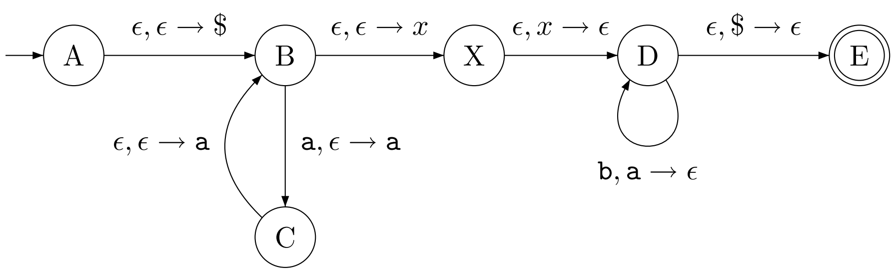
Claim version 2. *For any n, show that if a string $w \in \Sigma^*$ can be derived from either* S *or* Y *in n steps, then w contains an even number of* a*'s.*

(a.) The original claim involved strings in $L(G)$, i.e. strings that can be derived from the start symbol S.
Why did we extend the claim to include derivations starting with the variable Y? Why didn't we extend it even further to include derivations starting with the variable X?

(b.) Prove version 2 of the claim using strong induction on the derivation length n.

4. PDA to CFG conversion.

Consider the following PDA.



Recall the proof that a PDA is a context free language on page 119–120 of Sipser (and in the notes for Lecture 14). For the above PDA.

(a.) Generate rules defined by the first bullet point of the proof on page 120.

(b.) What is the start variable?

(c.) How many rules are generated by the second bullet point. Explain why your answer is correct.

5. Give me that old time PDA, its good enough for me, it was good enough for my father.

Tony had just released into the market a new model of PDA called Blu-PDA (Sushiba released a competing PDA product called HD-PDA, but thats really a topic for a different exercise).

Instead of a stack, like the good old PDA, the new Blu-PDA has a queue. You can push/pop characters from both sides of the queue (thus, the Blu-PDA can see both characters stored in the front and back of the queue when making a transition decision [and the current input character of course]). Since Tony is targeting this product to the amateur market, they decided to limit the queue size to 25 (if the queue size exceeds 25, then it stops and reject the input). Tony claims that the new Blu-PDA is a major breakthrough and a considerably stronger computer than a PDA.

(Of course, if the Blu-PDA does strange things like reading characters from an empty queue, or popping characters from an empty queue, then it immediately rejects the input.)

(a) So, given a Blu-PDA, is it equivalent to a PDA, DFA, or is it stronger than both?

(b) Explain clearly why your answer is correct.

(c) (5 point bonus) Prove your answer in detail.