CS331 Autumn 2003 Mid-term Examination

Time: 3 hours | Total marks: 75

- ***The exam is open book and notes.***
- ***Results/proofs covered in class/problem sessions/assignments may simply be cited, unless specifically asked for.***
- ***Unnecessarily lengthy solutions will be penalized.***
- ***If you need to make any assumptions, state them clearly.***
- ***Do not copy solutions from others or indulge in unfair means.***
- ***Penalty for copying/unfair practices: FR grade***

1. *[15 marks]* Let $\Sigma = \{a, b\}$. We wish to construct a DFA **(not NFA)** for recognizing substrings represented by the regular expression $\mathbf{ba}(\mathbf{ab})^*$ in a stream of a's and b's (see Fig. 1). Thus, if the DFA is fed a stream of a's and b's, it should enter one of its final state(s) whenever the last few symbols read constitue a string in $\mathbf{ba}(\mathbf{ab})^*$. For example, if $b\ \underline{a}\ b\ \underline{a}\ a\ \underline{b}\ \underline{a}\ \underline{b}\ \underline{a}\ a\ \underline{b}\ b$ is fed to the DFA, with the symbols being fed in left-to-right order, then the automaton should enter a final state after reading the 2^{nd}, 4^{th}, 6^{th}, 7^{th}, 8^{th},9^{th}, and 11^{th} symbols (shown underlined). The corresponding strings recognized are: $ba, ba, baab, ba, baabab, ba$, and $baab$, respectively.

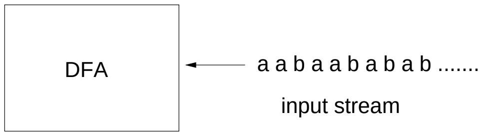

Figure 1:

Construct the above DFA.

2. *[7.5 + 7.5 marks]*

   (a) Show that $\{w \mid w \in (\mathbf{0}+\mathbf{1}+\mathbf{2})^*,\ N_0(w) \leq N_1(w) + N_2(w) \leq 2N_0(w),\ \text{ and } N_0(w), N_1(w), N_2(w) \geq 1\}$ is a CFL. Here, $N_i(w)$ represents the number of i's in w.
   (Examples of strings in the language: $100212, 00121$. *Examples of strings not in the language:* 0000122, 2210.*)*

   (b) $\{w \mid w = x.x.y,\ x \in (\mathbf{0}+\mathbf{1})^+,\ y \in (\mathbf{0}+\mathbf{1})^*\}$ is not a regular language.
   (Examples of strings in the language: $0101, 10110111$. *Examples of strings not in the language:* 0110, 011101.*)*

3. *[15 marks]* Let L be a regular language over the alphabet $\{0, 1\}$. Define $L_{\frac{1}{2}} = \{w \mid w.w \in L\}$. Prove that $L_{\frac{1}{2}}$ is regular.
   If you are outlining the construction of a DFA/NFA/regular expression for $L_{\frac{1}{2}}$, you must give at least an informal argument why the language acceoted by your DFA/NFA or that represented by your regular expression is indeed $L_{\frac{1}{2}}$.

4. *[15 marks]* Given a language L over an alphabet Σ, define $twist(L) = \{w \mid w = x.y,\ x, y \in \Sigma^*,\ x.y^R \in L\}$. Here, x^R denotes the reverse of the string x. Thus, $twist(L)$ is formed by taking strings in L, breaking them into two parts (one of which could be ϵ), reversing the second part and then concatenating them back.

   Consider the language $L(G)$ defined by the CFG $G = (\{S, A\}, \{0, 1\}, P, S)$, where P is given by $\{S \rightarrow \epsilon \mid 1A \mid 0S,\ \ A \rightarrow 0A \mid 1S\}$.

   Give a CFG for $twist(L(G)$. You must give at least an informal justification of your construction.

5. *[15 marks]* Consider the PDA shown in Fig. 2. The initial state is A and the initial symbol on the stack is Z. There are no final states. Let L denote the language accepted by the PDA *by an empty stack*.

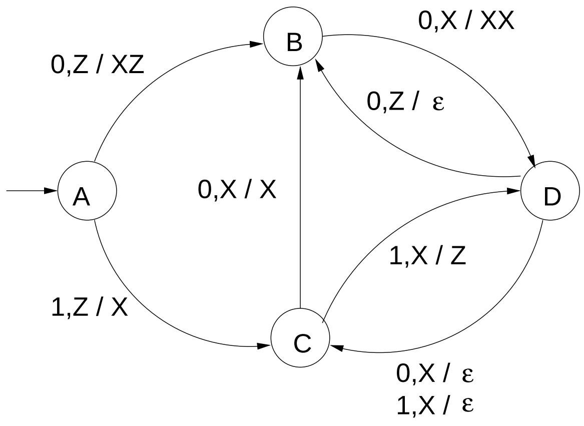

Figure 2:

Give a CFG for L in which each production rule is of the form $X \rightarrow Ya$ or $X \rightarrow a$, where $a \in \{0, 1\}$ and X and Y are (not necessarily distinct) non-terminals in the grammar. The grammar could, of course, have more than two non-terminals.