

Q1. Basically this means that  $L$  does not have  $aaa$  as a substring.

$Q = \{q_0, q_1, q_2, q_3\}$ .  $\Sigma = \{a, b\}$ .  $F = \{q_0, q_1, q_2\}$ .

$\delta(q_i, a) = q_{i+1}$ , if  $i \leq 2$ .

$\delta(q_i, b) = q_0$ , if  $i \leq 2$ .

$\delta(q_3, a) = \delta(q_3, b) = q_3$ .

Q2. The statement is false.

Let  $R = a$ ,  $S = b$ . Then,  $bb$  is in  $Lang((R + \epsilon)(S + \epsilon)^*)$  but not in  $Lang((R + \epsilon)^*(S + \epsilon))$ .

Q3.

$S \rightarrow aaSb$

$S \rightarrow abTb \mid T \mid X$

$T \rightarrow bbTb \mid Y$

$X \rightarrow aXa \mid Y$

$Y \rightarrow bYa \mid c$

Intuitively,  $S$  first matches off the (pair of)  $a$ 's to the left of  $c$  with the  $b$ 's to the right of  $c$ . Then it goes to  $T$  ( $abTb$ ) or  $X$  depending on whether the number of  $a$ 's to the left of  $c$  is less or at least double the number of  $b$ 's to the right of  $c$ .

$T$  matches the (pair of)  $b$ 's to the left of  $c$  with the  $b$ 's to the right of  $c$ .

$X$  matches the  $a$ 's to the left of  $c$  with the  $a$ 's to the right of  $c$ .

$Y$  matches the  $b$ 's to the left of  $c$  with the  $a$ 's to the right of  $c$ .

They are done in appropriate order to maintain the order of  $a$ 's and  $b$ 's.

Q4. The grammar is ambiguous because  $bba$  has the following two different leftmost derivation.

$S \Rightarrow bbSa \Rightarrow bbRa \Rightarrow bbTa \Rightarrow bba$

$S \Rightarrow R \Rightarrow bRa \Rightarrow bTa \Rightarrow bbTa \Rightarrow bba$

Basically, the above grammar generates the language  $\{b^i a^j : i \geq j\}$ . The unambiguous grammar for it is:

$S \rightarrow bSa \mid T$

$T \rightarrow bT \mid \epsilon$

It is unambiguous, as any  $a$  is only generated by using the production  $S \rightarrow bSa$ . After using  $S \rightarrow T$ , only  $b$ 's are generated, one by one.

Q5. True. Let  $X = \{w : aba \text{ is a substring of } w\}$ . Note that  $X$  is regular. Now,  $Lang(G)$  is good iff  $Lang(G) \cap \bar{X}$  is empty. This is used in the following algorithm to decide the question in the claim.

(0): If  $G$  is not a context free grammar, output false.

(1): Now assume  $G$  is a CFG. Convert it to a NPDA  $P$  such that  $Lang(G) = Lang(P)$  as done in class.

(2): Let  $A$  be a DFA for accepting  $(a + b)^* aba (a + b)^*$ .

(3): Let  $B$  be a DFA for accepting the complement of  $Lang(A)$ .

(4): Let  $P'$  be a NPDA for accepting  $Lang(P) \cap Lang(B)$  as done in class.

(5): Let  $G'$  be a grammar for  $Lang(P')$  as done in class.

(6): Check if  $Lang(G') = \emptyset$  as done in class.

(7): Output yes, iff the above check is true.

Q6. False.

Let  $L = \{ba^i ba^j ba^t : M_i \text{ accepts } w_j \text{ in } \leq t \text{ time steps}\}$ .

Note that  $L$  is recursive as shown in tutorials.

Now suppose  $L' = \text{substring}(L)$  is recursive. Then,  $L' \cap ba^*ba^*b$  must also be recursive. But This means  $\{ba^i ba^j b : M_i \text{ accepts } w_j\}$  is recursive, a contradiction to result done in class.

Thus, Claim is false.

Q7.

Given  $M_i$ , design  $M_{f(i)}$  such that:

$M_{f(i)}$  accepts  $(ab)^j$  for all  $j$ .

Furthermore,  $M_{f(i)}$  accepts  $aa$  iff  $\text{Lang}(M_i)$  is not empty.

Then,  $a^{f(i)} \in L_7$  iff  $M_i \in L_e$  (where  $L_e = \{M : \text{Lang}(M) = \emptyset\}$ ). Thus,  $L_e \leq_m L_7$ . As  $L_e$  is not r.e., we have that  $L_7$  is not r.e.

Q8. To see that the problem is in NP, consider the following: Guess a sequence of results of the games and check if each member of the audience has their prediction true for at least one game.

To see NP-hardness, we do a reduction from 3-SAT. Suppose an instance of 3-SAT is given, which has  $n$  variables and  $m$  clauses,  $C_1, C_2, \dots, C_m$ .

Then, form an instance of the problem in Q8 as follows. There are  $n$  games and  $m$  members of the audience. If  $x_j \in C_i$ , then  $i$ -th member of the audience predicts that  $A$  wins game  $j$ . If  $\neg x_j \in C_i$ , then  $i$ -th member of the audience predicts that  $B$  wins game  $j$ .

Now, if 3-SAT formula is satisfiable using truth function  $t(x_j)$ , then we have match fixer have  $A$  win game  $j$  iff  $t(x_j)$  is true. It is easy to verify that, for  $1 \leq i \leq m$ , the  $i$ -th member of the audience has at least one prediction true since clause  $C_i$  is satisfied. On the other hand, if match fixer can have the results of the games such that each member of the audience has at least one prediction true, then let  $t(x_j) = \text{true}$  iff  $A$  wins game  $j$ . Now,  $C_i$  is satisfied as at least one of the predictions of the  $i$ -th member of the audience is true.