

1. This problem provides some experience with the Chomsky hierarchy (regular, context-free, context-sensitive, and unrestricted grammars). For each of the following languages (denoted  $L$ ), provide a grammar  $G$  occupying the most restrictive level of the hierarchy possible.

(a)  $\Sigma = \{a, b\}$ ,  $L(G) = \{a^n b^n \mid n \in \mathbb{N}\}$ .

(b)  $\Sigma = \{a, b\}$ ,  $L(G) = \{\alpha \in \Sigma^* \mid \alpha \text{ is a palindrome}\}$ .

(c)  $\Sigma = \{a, b\}$ ,  $L(G) = \{a^n b^m \mid n \in \mathbb{N} \text{ odd}, m \in \mathbb{N} \text{ even}\}$ .

(d)  $\Sigma = \{a, b, c\}$ ,  $L(G) = \{a^n b^n c^n \mid n \in \mathbb{N}\}$ .

2. Transform the following context-free grammars into Chomsky normal form.

(a) 
$$\begin{array}{lll} S & \rightarrow & aA \mid bA \mid aE \mid bE \\ A & \rightarrow & S \\ E & \rightarrow & \epsilon \end{array}$$

(b) 
$$\begin{array}{lll} S & \rightarrow & ab \mid A \\ A & \rightarrow & bAaA \mid b \end{array}$$

3. In this problem, we study some properties of the Nussinov algorithm for RNA folding. Recall that, given an RNA sequence  $R = r_1 \dots, r_L$ ,  $B(i, j)$  is the maximal number of base-pairings for the substring  $r_i \dots r_j$ , and satisfies the recursion,

$$B(i, j) = \max \begin{cases} B(i+1, j-1) + \sigma(r_i, r_j), \\ B(i+1, j), \\ B(i, j-1), \\ \max_{i < k < j-1} \{B(i, k) + B(k+1, j)\}, \end{cases}$$

for  $i+1 < j$ , with base cases  $B(i, i) = 0$  and  $B(i, i+1) = 0$ . Here  $\sigma(r_i, r_j) = 1$  if  $r_i$  and  $r_j$  can base pair, 0 otherwise.

- (a) (modified from Durbin, question 10.3) What is the minimum length of a hairpin loop in the algorithm above? Modify the Nussinov folding algorithm so that hairpin loops have a minimum length of  $h$ .
- (b) How can you modify the Nussinov folding algorithm to handle *circular* RNA?