

**CAP 6515 HOMEWORK ASSIGNMENT 3**  
**DUE ON 12-08-2022**

**Note:** Any solution to an algorithm design question MUST contain the following four sections:

- (1) **Problem statement.** A clear unambiguous statement of the problem to be solved, which includes the input, the output, and the object function with the constraints.
- (2) **Algorithm description.** A clear, unambiguous description of the algorithm.
- (3) **Correctness proof.** A convincing mathematical argument that the algorithm described solves the computational problem described.
- (4) **Time analysis.** A time analysis of the algorithm, up to order, in terms of all relevant parameters.

You may use any algorithms and data structures from class.

1. RNA BASE PAIR MAXIMIZATION PROBLEM (60 PTS)

Formalize the **pseudocode** to predict an RNA secondary structure based on the base pair maximization model including the **trace back** procedure for the predicted secondary structure. Please refer to slides 18-21 in “3.2.RNAfolding.pdf”.

The secondary structure can be represented by the matching parentheses and dots to denote paired and free bases, respectively. For example:

```
GAGCCAUUAGCUCAGUUGGUAGAGCAUCUGACUUUUAUCAGAGGGUCGAAGGUUCGAGUCCUUAUGGCUCA
(((((((..<<<<.....>>>>..<<<<.....>>>>.....<<<<.....>>>>)))))).
```

2. K-EXONS SPLICED ALIGNMENT (40 PTS)

One disadvantage of the *exon chaining* formulation of the spliced alignment algorithm is that it may prefer to concatenate many short putative exons to maximize the alignment score. Modify the spliced alignment algorithm to consider the optimal alignment between a genomic sequence and an mRNA sequence with at most  $k$  exons.