

# Algorithms and Data Structures for Bioinformatics( BI-GY 7453)

## Assignment - 1

### Criteria for evaluation :

- Commenting of code
- The code file you have submitted needs to be executable and satisfy all the test cases.
- A clear and comprehensive explanation of your problem-solving approach.

### Question 1: ( 5 marks )

Find the time complexity of the following pseudo-code :

#### 1.1. Code Snippet 1 ( 2 marks )

```
1  int a = 0 , b = 0;
2  for ( i = 0; i < N ; i ++ ) {

3    a = a + rand () ;
4  }
5  for ( j = 0; j < M ; j ++ ) {

6    b = b + rand () ;
7  }
```

#### 1.2. Code Snippet 2 ( 2 marks)

```
1  int l , t , k = 0;
2  for ( l = n / 2; l <= n ; l ++ ) {

3    for ( t = 2; t <= n ; t = t * 2 ) {

4      k = k + n / 2;
5    }
6  }
```

#### 1.3. Code Snippet 3 ( 1 mark )

```
1  int a = 0 , i = N ;
```

```
2 while ( i > 0) {  
3   a += i ;  
4   i /= 2;  
  
5 }
```

### **Question 2 : ( 5 marks )**

Given an array of numbers, find the GCD of the array elements (Hint: Use the Euclidean algorithm discussed in the lecture)

#### **Example Input :**

arr [ ] = {1, 2, 3}

**Output:** 1

#### **Example Explanation :**

The GCD of three or more numbers equals the product of the prime factors common to all the numbers.

### **Question 3 : (10 marks)**

In the context of genetic sequencing, you are given the alignment of two DNA sequences, where one is the reference sequence (REF) and the other is a mutated sequence (MUT). Your task is to design a simple algorithm to detect and report the mutations present in the MUT sequence compared to the REF sequence.

#### **Example:**

- REF Sequence: ATCGATCGA

- MUT Sequence: ATCAATCGT

#### **Mutations detected:**

1. Substitution at position 4: G  $\rightarrow$  A

2. Substitution at position 9: A  $\rightarrow$  T

#### **Guidelines:**

- Present a clear algorithmic approach. ( 4 marks)
- Provide a function in Python to implement the mutation detection algorithm. ( 5 marks)
- Discuss the time complexity of your algorithm. ( 1 mark)