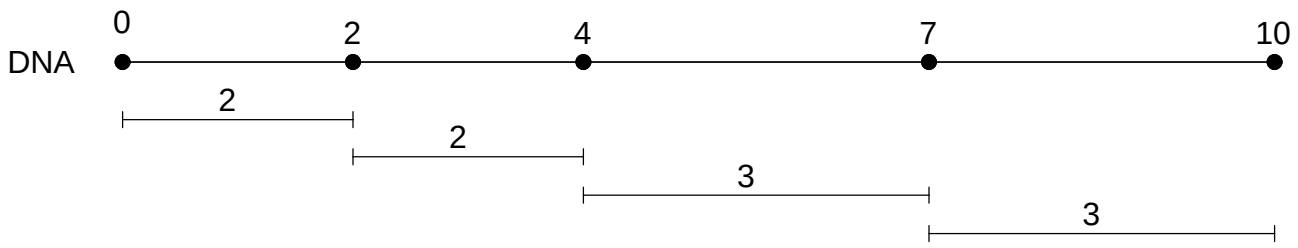
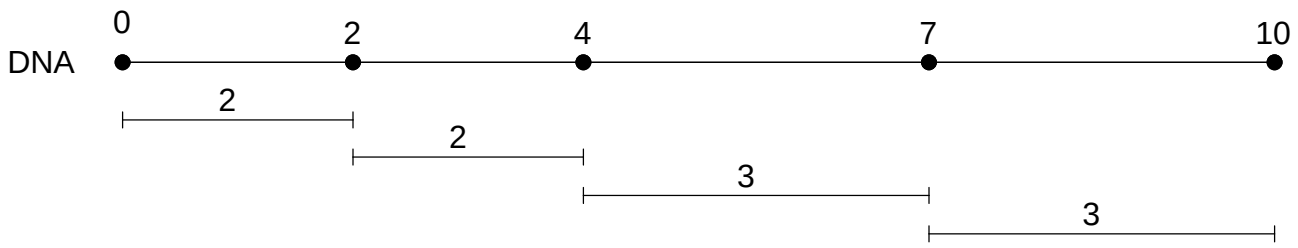
9/6/07 Name: _____

1) Consider the following DNA sequence with restriction sites marked. The complete digest is as shown.



a) The *partial digest* would generate fragments of lengths between all restriction sites. Complete the following diagram showing the partial digest.



b) A *multiset* is a set that allows duplicates. What would be the multiset for the lengths of DNA fragments for the above partial digest?

2) If $X = \{ x_1, x_2, x_3, \ldots , x_n \}$ is a set of n points on a line segment in increasing order, the ΔX denotes the multiset of all pairwise distances between points in X.

a) For $X = \{ 0, 2, 4, 7, 10 \}$ as in the above DNA, what is ΔX?

b) If the set X is stored in an array X[1..n], write a nested loops to generate ΔX.

c) Develop a formula between $|X|$ and $|\Delta X|$.

3) Consider the following brute-force algorithm for solving the partial digest problem (PDP).

```
1. BruteForcePDP(L, n):
2.     M <- maximum element in L
3.     for every set of n – 2 integers 0 < x_2 < … x_{n-1} < M
4.         X <- {0,x_2,…,x_{n-1},M}
5.         Form DX from X
6.         if DX = L
7.             return X
8.     output "no solution"
```

a) In line 4, why does the set X contain 0 and M?

b) To analyze the algorithm to determine its O( ) notation, draw a "search tree" of all sets of n-2 integer.

4) How could we be smarter about generating the possible sets of X? Hint: Use more information about the specific problem instance.