

Name: KEY
 Bioinformatics Quiz #1 [100 pts]
 02/03/2010

1)

What are three differences between Smith-Waterman and Needleman-Wunsch [6 pts]?

- | | |
|----------------------|---------------------------|
| SW | NW |
| - LOCAL | - GLOBAL |
| - NO NEGATIVE VALUES | - NEGATIVE VALUES ALLOWED |

2) - START BACKTRACKING AT HIGHEST SCORE - START BACKTRACKING AT LOWER RIGHT CORNER

Use the appropriate dynamic programming algorithm to LOCALLY align the sequences below. Use +2 for a match, -2 for a mismatch, and -1 for a gap.

Show ALL optimal alignments and give the corresponding scores. [50 points]

		C	G	G	A	T
	0	0	0	0	0	0
C	0	2	1	0	0	0
G	0	1	4	3	2	1
A	0	0	3	2	5	4
T	0	0	2	1	4	7
A	0	0	1	0	3	6

```

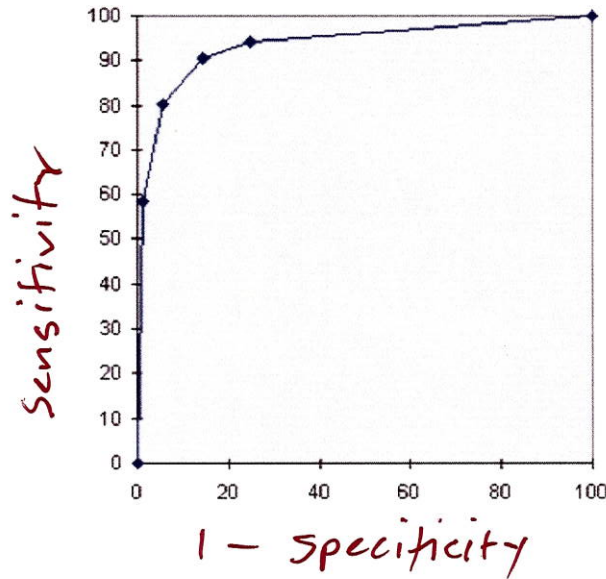
C - G A T
|   | | |
C G G A T
  
```

Score = 7

```

C G - A T
| | | | |
C G G A T
  
```

3. In the following ROC (Receiver Operating Curve) label the axes and also define each axis in terms of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) [16 pts].



ACTUAL

TP	FP
FN	TN

PREDICTION

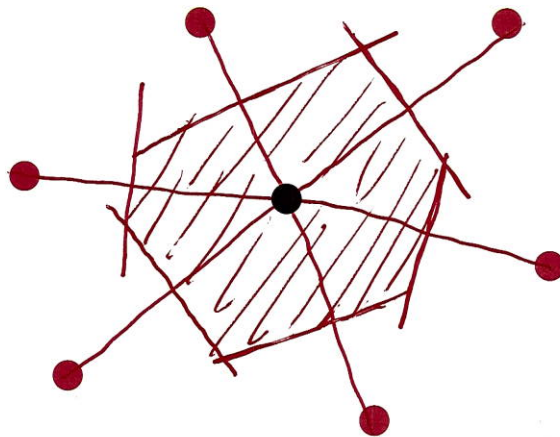
$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Specificity} = \frac{TN}{FP + TN}$$

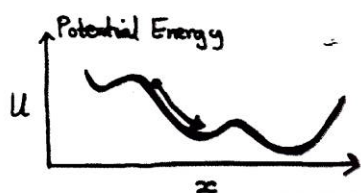
4. Which statistical distribution is used to assess the statistical significance of sequence alignments and why [6 pts]?

- Extreme value distribution
- Max (independent random variables)

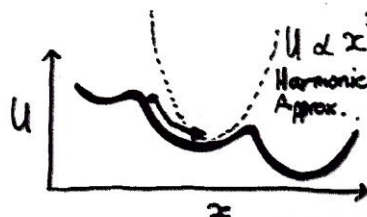
5. Draw the Voronoi volume for the central atom (black circle) in the figure below [10 pts]:



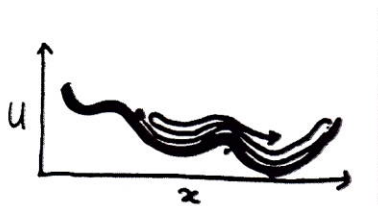
6. Label the four panels below in the red boxes [12 pts]:



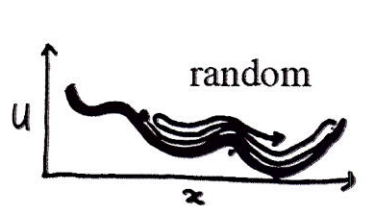
MINIMIZATION



NORMAL MODE ANALYSIS



MOLECULAR DYNAMICS



MONTE CARLO