

Math 127, **Homework # 3**

due: Tuesday, February 27

- (1) Write a short essay (\leq one page) about the software **PHYLIP**. Who developed **PHYLIP** and why? What does it do and who uses it? Illustrate the use of **PHYLIP** for one example chosen by you.
- (2) Consider the following two DNA sequences of length ten:

$$\sigma^1 = \text{ACGGCTTAGG} \quad \text{and} \quad \sigma^2 = \text{CGAGTCTATG}.$$

How many alignments do these two sequences have? Find all alignments of these two sequences that maximize the number of matches.

- (3) The sequences in Problem (2) define a polynomial $f_{\sigma^1\sigma^2}$ as in equation (2.15) on page 56. Let $h_{\sigma^1\sigma^2}$ be the polynomial in three unknowns m, x, s which is obtained from $f_{\sigma^1\sigma^2}$ by specializing $\theta_{uu} = m$, $\theta_{uv} = x$, $\theta_{u-} = \theta_{-u} = s$ for any two distinct nucleotides u, v , and $\theta'_{kl} = 1$ for any $k, j \in \{H, I, D\}$. Compute the Newton polytope of $h_{\sigma^1\sigma^2}$.
- (4) Explain the meaning of Figure 7.3 on page 204 in your own words.
- (5) Let $n = 6$ and determine the metric d which gives the pairwise distances (in miles) among the six cities *Atlanta*, *Chicago*, *Houston*, *Miami*, *New York* and *San Francisco*. Build a phylogenetic tree on these “taxa” from the data d by applying the Neighbor-Joining Algorithm (pages 73/74).
- (6) The following 5×5 -matrix D has two unknown parameters x and y :

$$D = \begin{pmatrix} 0 & 4 & 10 & 8 & 7 \\ 4 & 0 & 12 & 10 & 9 \\ 10 & 12 & 0 & x & 7 \\ 8 & 10 & x & 0 & y \\ 7 & 9 & 7 & y & 0 \end{pmatrix}$$

Draw the set of all points (x, y) in the plane such that D is a metric. In your diagram, mark all points (x, y) such that D is a tree metric.