Review sheet for BNFO 236 final exam (Spring 2017)

Sequence alignment algorithms:

1. Needleman-Wunsch
    a. Compute alignment scoring matrix V and traceback matrix T
2. Smith-Waterman
    a. Compute alignment scoring matrix V and traceback matrix T
    b. Difference between this and Needleman-Wunsch
3. BLAST-like alignment
4. Running time of above algorithms as function of size of input sequences
5. Be able to manually run the above algorithms on a given input
6. Database search
7. BLOSUM62
    a. Compute alignment score with BLOSUM62 substitution scoring matrix
    b. Incorporate BLOSUM62 into Needleman-Wunsch, Smith-Waterman, and BLAST-like alignment

Phylogeny reconstruction:

1. Bipartitions to compare trees
2. Bootstrapping to evaluate statistical consistency

UPGMA phylogeny reconstruction:

1. Newick representation, tree data structure, tree bipartitions, Juke-Cantor evolutionary distance calculation, genome evolutionary distance
2. UPGMA algorithm
3. Running time of UPGMA as function of size of input distance matrix
4. Be able to manually run UPGMA on a given input
5. UPGMA variants – update distance matrix be largest or smallest number between two nodes
6. UPGMA correctness

Neighbor joining (NJ) phylogeny reconstruction:

1. NJ algorithm
2. Running time of NJ as function of size of input distance matrix
3. Be able to manually run NJ on a given input (first iteration is sufficient)
4. Difference between NJ and UPGMA

Python:

1. Functions: scalars vs lists (or dictionaries) as parameters
2. Difference in assignment between scalars and lists (or dictionaries)