Bioinformatics - Supervision 2

Work is due 24 hours before the supervision. Please hand in at Student Administration at the Computer Lab, or email (tab43@cam.ac.uk)

General questions
Answer the following questions in your own words:
1. What problem does CLUSTALW address, and how does it work?
2. How can we measure the quality of a multiple sequence alignment?
3. Discuss three ways to reconstruct DNA sequences if we only get to see (partly overlapping) fragments of DNA.
4. What is the "additivity" property for a matrix, and why is it useful to have?
5. Explain what microarray data is.

Algorithms
For each of the following algorithms below, write a few sentences on the following:
- The abstract problem it solves
- A practical use of the algorithm
- How the algorithm works
- The space and time complexity

1. Fitch parsimony
2. Sankoff parsimony
3. UPGMA
4. Neighbor Joining
5. K-means clustering
6. Progressive greedy K-means
7. Markov Clustering