Homework 4 - Theory Questions

1)

There are two different probabilities for A and G at node 6 because it is connected to an internal and an external branch. Since A is determined at the tip, it influences the probability of node 6 quite strongly.

2)

Everything downward (towards the tips) of the tree could be kept, but everything upward (toward the root) would need to be recalculated. This is because the Felsenstein pruning algorithm is a bottom-up approach.

3)

UPGMA	**Maximum Likelihood**
Uses arithmetic mean	Uses statistical approach to maximize likelihood
Uses a distance matrix	Doesn't use a distance matrix
Only makes rooted trees	Can make rooted or unrooted trees
Assumes that there is an equal rate of evolution for all sequences, molecular clock hypothesis	Does not assume equal rate of evolution between sequences

4) If we have sequences of length 4, then we'll have 3 nodes. The recursions must be computed on each site, resulting in 12 calculations.

Now if we do it for sequences of length 5, we will have 4 nodes, resulting in 20 calculations.

20/12=1.666… Therefore the number of calculations increase by 66%.

5)

By the definition of time reversibility this would affect the likelihood. This is because, a stochastic process is said to be time reversible if its probabilistic structure is unaffected by reversing the direction of time.