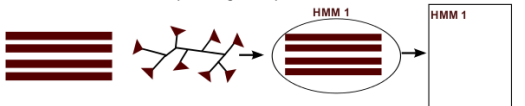
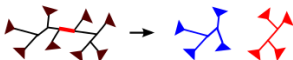


Generating ensemble of HMMs

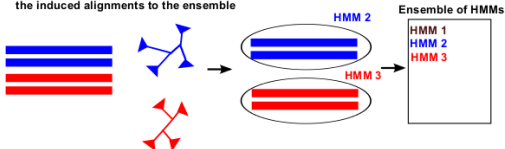
1. Build an HMM on the entire MSA and add it to the ensemble.
If the MSA has more than 10 sequences, go to step 2.



2. Partition the tree on the **centroid edge** that separates the tree into two approximately equally sized subtrees



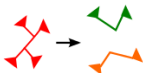
3. Find the alignments induced by each of the subtrees and add the HMMs built on the induced alignments to the ensemble



4. Repeat the decomposition steps 2 and 3 on each subtree with more than 10 leaves that meets either one of the two following criteria:

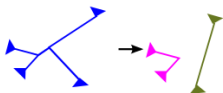
A) Contains more leaves than 10% of the total number of sequences in of the original input alignment

Example: Subtree contains 45 leaves and original alignment contained 100 sequences



B) The mean pairwise sequence identity of the induced MSA is below 0.4

Example: The induced MSA has a mean pairwise sequence identity of 0.31



5. Once none of the subtrees meet any of the decomposition criteria, return the ensemble of HMMs produced by this iterative process

Ensemble of HMMs

HMM 1
HMM 2
HMM 3
HMM 4
HMM 5
HMM 6
HMM 7