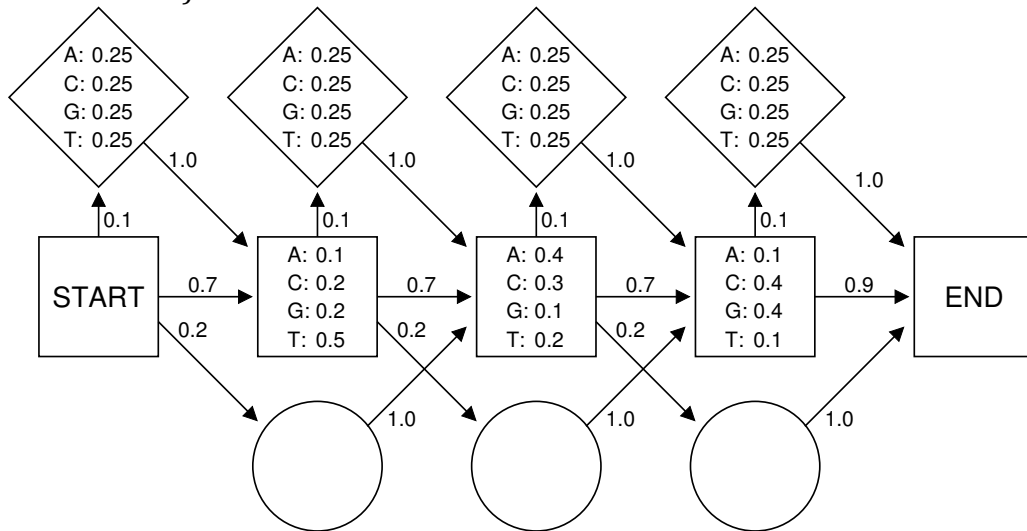# Bioinformatik für Biologen SS 2019

Complementary Homework 9

This homework will not be collected or graded, but its contents and the papers to read may be relevant in an upcoming lecture, test or assignment.

1. You have the following pHMM (all transitions with zero probability have been deleted):


- Mark all silent states and all emitting states in the HMM.
- Which are the Match, Insertion and Deletion states – what do they do?
- What is the typical length of the reflected Motif?
- What is the probability the 2nd base in the pattern gets deleted (according to the pHHM)?
- Determine (at least) two paths through the pHMM which can produce the sequence ATG.
- What is/are the motif sequence(s) that can be produced by the HMM which highest probability? (You do not need Viterbi or Forward/Backward for this.)

2. Using pHMMs for alignment:
- Explain how you would align sequences using an HMM?
- Explain why you would not place two characters of different sequences into the same alignment column, even if they are assigned to the same insert state.
- Come up with examples where you align sequences through the HMM above. Design you examples such that they use all different state types.

3. Given the following alignment:

```
T1    A  C  A  T  C  G  T  C  A  C  T  T  T
T2    A  G  A  T  C  A  T  C  C  C  T  G  T
T3    C  G  A  G  G  G  C  C  G  G  G  G  T
T4    C  A  A  T  C  T  T  C  T  C  G  G  T
T5    C  G  A  G  C  T  C  C  A  G  G  A  T
```

- Mark all variable sites.
- Mark all parsimony informative sites.
- Without assuming a tree structure, note down for each column what is the theoretically minimal number of mutations required to produce this column.

4. Use the sixth column from the alignment before and the following tree:

```
 T1     T2   T3 T4    T5
{  }   {  }{  }{  } {  }

{    }          {    }
    {    }
        {     }
```
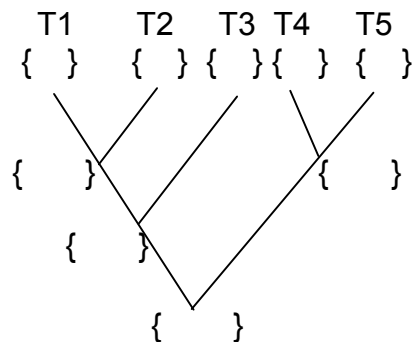
a) Fill the nucleotides into the brackets ({...}) according to the Fitch algorithm.
b) What is the parsimony score (i.e. the number of mutations needed, also called tree-length) for this column on this tree?