

MS2a, Exercises Week 8


Rune Lyngsø

December 9, 2009

A Hidden Markov Model Use

- a. Consider a hidden Markov model emitting sequences over the alphabet $\{A, C, G, T\}$. The model has two states, 0 and 1, that are equiprobable start states. Transition probabilities are 0.75 for remaining in a state and 0.25 for switching to the other state. In state 0 A and G are emitted with probability 0.45 while C and T are emitted with probability 0.05. In state 1 A and G are emitted with probability 0.05 while C and T are emitted with probability 0.45. What is the probability of observing the sequence ACTG? Observe that we do not have an end state providing explicit termination, so the model will not model a sequence length distribution. Rather, for every sequence length it models a distribution over sequence content.
- b. What is the most likely sequence of hidden states, and how probable is it?
- c. What is the most likely hidden state at position 2, summing over all possible paths, and how probable is it?

B Hidden Markov Model Design

- a. The occasionally dishonest casino is a standard HMM example. The casino usually plays with a fair six-sided die, but occasionally switching to a biased die that has 50% chance of coming up  (apparently the casino wins when this happens), with the other five rolls being equiprobable. Describe a HMM modelling the die rolls at this casino, where the expected length of a fair die run is 50 and the expected length of a biased die run is 10.
- b. Construct a HMM that generates the sequence A^i , i.e. the sequence of i As, with probability 2^{-i} for $i \geq 1$, if possible. Otherwise argue it is not possible.

- c. Construct a HMM that generates the sequence $A^i C^j$ with probability 2^{-i-j} for $i, j \geq 1$, if possible. Otherwise argue that it is not possible.
- d. Construct a HMM that generates the sequence $A^i C^i$ with probability 2^{-i} for $i \geq 1$, if possible. Otherwise argue that it is not possible.