

## Markov Chains and HMMs

1. Markov chain is used to generate a binary string. Initial probability is given by  $P(0) = 0.5$  and  $P(1) = 0.5$  and the transition probabilities are

$p$	0	1
0	1/5	4/5
1	1/4	3/4

- Find the probability that string 00000 is generated in the first 5 steps.
  - Find the probability that string 110011 is generated in the first 6 steps.
2. Markov chain with the end state is used to generate a binary string. Initial probability is given by  $P(0) = 0.5$  and  $P(1) = 0.5$ . Transition probabilities are given in the table.

$p$	$B$	0	1	$E$
$B$	0	0.5	0.5	0
0	0	0.2	0.7	0.1
1	0	0.5	0.2	0.3
$E$	0	0	0	1

- Find the probability that the string starts with 01.
  - Find the probability of generating the string 110011.
3. Consider a Markov Chain with 6 non-silent states. How many paths of states of length  $k$  (including the begin and the end state) are there?
  4. There are two types of coins: Type 1 - fair coin, Type 2 - probability of heads is 0.25. One switches from Type 1 to Type 2 with probability 0.2 and from Type 2 to Type 1 with probability 0.4. One starts with Type 1 with probability 0.6 and ends the game when using Type 1 with probability 0.1. One starts with Type 2 with probability 0.4 and ends the game when using Type 2 with probability 0.3. Describe the HMM model and find the most probable path for  $HHHTT$ .
  5. Consider two types of coins: Type 1 - fair coin, Type 2 - probability of heads is 1/4. One switches from Type 1 to Type 2 with probability 0.2, from Type 2 to Type 1 with probability 0.4 and one stops playing (no matter what type is used) with probability 0.1.

- Describe the HMM model.
- Use Viterbi's algorithm to find probability of the most probable path for string HHTHTT.
- Use the forward algorithm to find  $P(HHTH)$ .

6. Consider the CpG HMM.

- In the + model the transition probability from C to G is 0.2, transition probability from C to C is 0.2, transition probability from G to C is 0.2, transition probability from G to G is 0.1.
- In the - model the transition probability from C to G is 0.05, transition probability from C to C is 0.1, transition probability from G to C is 0.1, transition probability from G to G is 0.1.
- Probability of switching from a particular state in the + model to a particular state in the - model is 0.02.
- Probability of switching from a particular state in the - to a particular state in the + model is 0.03.
- Transition probability from the begin state to a state in the + model is 0.1 and to a state in the - model is 0.15.
- Transition probability of going to the end state from any state (but the begin state) is 0.02.
- Note that some other transition probabilities are not given as they are not need for the problem.

Find the most probable path for generating CCGGCG.

7. Describe an efficient algorithm (better in English than in the pseudo-code) which in a Markov Chain computes the probability that the length of a generated string is equal to  $l$ . Estimate its running time.