Lehrstuhl für STEUERUNGS- UND REGELUNGSTECHNIK Technische Universität München Prof. Dongheui Lee	MACHINE LEARNING IN ROBOTICS Exercises: Markov processes & HMM

Exercise 1

Let us call $\boldsymbol{P}$ the transition matrix of an irreducible Markov chain with finitely many states.

a) Show that $\boldsymbol{Q} = \frac{1}{2}(\boldsymbol{I} + \boldsymbol{P})$, where $\boldsymbol{I}$ stands for the identity matrix, is the transition matrix of an irreducible and aperiodic Markov chain.

b) Show that $\boldsymbol{P}$ and $\boldsymbol{Q} = \frac{1}{2}(\boldsymbol{I} + \boldsymbol{P})$ have the same stationary distributions.

c) Discuss, physically, how the two chains are related.

Exercise 2

A protocol for data transmission shall be analysed using a Markov chain with 3 states. The probability for the transition from $state1$ (*check interface for incoming data*) to $state2$ (*check address*) is 0.1. The address is correct with probability 0.4. In this case, there is a transition to $state3$ (*message received*). Otherwise, the system returns to $state1$. If a message was received and there is no further message (probability 0.7), the system leaves $state3$ and enters in the $state1$. If there is a further message, it enters in the $state2$.

a) Specify the matrix of transition probabilities.

b) Draw the corresponding Markov chain.

c) What is the probability for the system to be in $state1$?

Exercise 3

The EM algorithm finds parameters $\boldsymbol{\theta}$ which maximize

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{i-1}) = \mathbb{E}_{\boldsymbol{Z}}[\ln p(\boldsymbol{X}, \boldsymbol{Z}|\boldsymbol{\theta})|\boldsymbol{X}, \boldsymbol{\theta}^{i-1}]$$

Similarly, the Baum-Welch algorithm finds the model parameters λ which maximize

$$Q(\lambda|\lambda^{i-1}) = \mathbb{E}_{\mathcal{Q}}[\log p(\mathcal{O}, \mathcal{Q}|\lambda)|\mathcal{O}, \lambda^{i-1}] = \sum_{\forall \mathcal{Q}} \log\left[p(\mathcal{O}, \mathcal{Q}|\lambda)\right] p(\mathcal{Q}|\mathcal{O}, \lambda^{i-1})$$

where $\mathcal{O} = [o_1, o_2, \ldots, o_T]$.

Show that the maximization of $Q(\lambda|\lambda^{i-1})$, with the constraint $\sum_{i=1}^{N} \hat{\pi}_i = 1$, leads to the update equation

$$\hat{\pi}_i = \gamma_1(i) = \sum_{j=1}^{N} \xi_1(i,j) = \sum_{j=1}^{N} p(q_1 = s_i, q_2 = s_j|\mathcal{O}, \lambda^{i-1}) = p(q_1 = s_i|\mathcal{O}, \lambda^{i-1})$$

Exercise 4

A Hidden Markov Model with 2 states $\{s_1 = H, s_2 = C\}$ and 3 possible observations based on the number of observed sizes $\{small, medium, large\}$ is given with transition probability matrix:

$$A = \begin{pmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{pmatrix}$$

and observation matrix:

$$B = \begin{pmatrix} 0.1 & 0.4 & 0.5 \\ 0.7 & 0.2 & 0.1 \end{pmatrix}$$

The prior probabilities of states are $\pi_1 = 0.6, \pi_2 = 0.4$.

a) Compute the probability of the state sequence $\{q_1 = H \; q_2 = H \; q_3 = C \; q_4 = C\}$.

b) Compute the probability $P(o_1 = S \; o_2 = M \; o_3 = S \; o_4 = L \mid q_1 = H \; q_2 = H \; q_3 = C \; q_4 = C)$.

Exercise 5

Two HMMs with different structure are shown in Fig. 1.

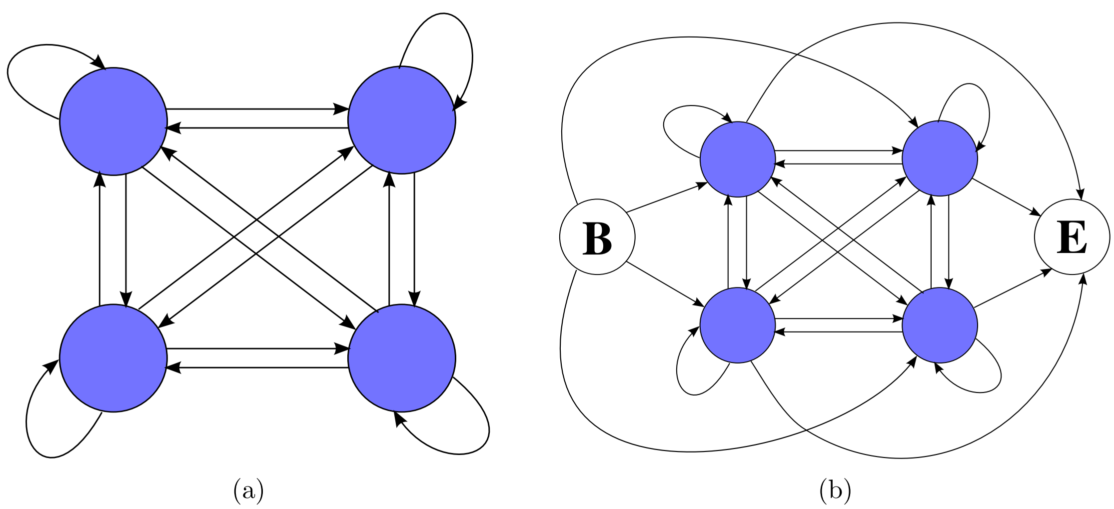

Figure 1: HMMs with different structures.

a) Consider the HMM in Fig. 1(a). Show that the sum of the probability of all possible state sequences $\mathcal{Q} = [q_1, \ldots, q_L]$ of length L is equal to 1.

b) Assume that the HMM has a begin (B) and an end (E) state, as in Fig. 1(b). The end state has probability ε. Show that the sum of the probability over all state sequences $\mathcal{Q} = [q_1, \ldots, q_L]$ of length L (and properly terminating by making a transition to the end state) is $p = \varepsilon(1 - \varepsilon)^{L-1}$. Use this result to show that the sum of the probability over all possible state sequences of any length is 1.

Hint: Use the result $\sum_{i=0}^{\infty} x^i = 1/(1 - x)$ for $0 < x < 1$.