Human-Oriented Robotics
Prof. Dr. Kai Arras, Social Robotics Lab
Lab instructors: Timm Linder, Luigi Palmieri, Billy Okal

Winter term 2014/2015
University of Freiburg
Department of Computer Science

Exercise 3: Basics of Probabilistic Reasoning – Solutions –

Exercise 3.1: Probabilistic Graphical Models

a) The resulting probabilistic graphical model is the one in Figure 1

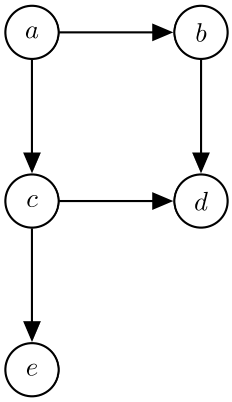

Figure 1: Probabilistic Graphical Model a

b) Figure a :

$$p(\mathbf{x}) = p(x_1)p(x_2)p(x_3)p(x_4)p(x_5|x_1, x_2, x_3, x_4)p(x_6|x_2, x_3)p(x_7|x_1, x_3, x_4)p(x_8|x_5, x_6)p(x_9|x_5, x_6, x_7)$$

Figure b :

$$p(\mathbf{x}) = p(x_1)p(x_2)p(x_3)p(x_4|x_1, x_2, x_3)p(x_5|x_2)p(x_6|x_4, x_5)p(x_7|x_5)p(x_8|x_4, x_5)p(x_9|x_6, x_8)p(x_{10}|x_7, x_8)$$

Exercise 3.2: Markov Chains

a) The Model 1 and 4 are not Markov Chains. Number 2 is a second order Markov Chain. Number 3 is a first order Markov Chain.

Exercise 3.3: State-space models The full distribution is given by;

$$\begin{aligned} p(s_1, \dots, s_k, a_1, \dots, a_k, r_1, \dots, r_k) = & \quad p(s_1)p(s_2|s_1, a_1)p(s_3|s_2, a_2) \cdots p(s_k|s_{k-1}, a_{k-1}) \\ & \cdot \quad p(a_1|s_1)p(a_2|s_2) \cdots p(a_k|s_k) \\ & \cdot \quad p(r_1|s_1, a_1) \cdots p(r_k|s_k, a_k) \\ = & \quad p(s_1) \prod_{i=2}^{K} p(s_i|s_{i-1}, a_{i-1}) \prod_{i=1}^{K} p(a_i|s_i) \prod_{i=1}^{K} p(r_i|s_i, a_i) \end{aligned}$$

Exercise 3.4: Joint Distribution

First of all, the sum in

$$\sum_{x_1} p(\mathbf{x}) = \sum_{x_1} p(x_1, \ldots, x_K)$$

means that we marginalize out x_1 (i.e. we sum up over all possible outcomes of x_1). In the continuous case, this would be an integral over $\mathrm{d}x_1$.

We want to show that:

$$\sum_{x_1} ... \sum_{x_K} p(\mathbf{x}) = 1$$

$$\sum_{x_1} ... \sum_{x_K} p(\mathbf{x}) = \sum_{x_1} ... \sum_{x_K} \prod_{k=1}^{K} p(x_k|\mathrm{pa}_k) = 1$$

We assume that the nodes in the graph have been numbered such that x_1 is the root node and no arrows lead from a higher numbered node to a lower numbered node (i.e. nodes with lower index are conditionally independent from nodes with higher index). We can then marginalize over the nodes in reverse order, starting with x_K. In the following product, we can extract the K-th factor:

$$\prod_{k=1}^{K} p(x_k|pa_k) = p(x_K|\mathrm{pa}_K) \prod_{k=1}^{K-1} p(x_k|\mathrm{pa}_k)$$

which, when inserting this back into the full formula, leads to

$$\sum_{x_1} ... \sum_{x_K} \prod_{k=1}^{K} p(x_k|pa_k) = \sum_{x_1} ... \sum_{x_K} p(x_K|\mathrm{pa}_K) \prod_{k=1}^{K-1} p(x_k|\mathrm{pa}_k).$$

Since none of the other variables depend on x_K and each of the conditional distributions is assumed to be correctly normalized, i.e.

$$\sum_{x_K} p(x_K|\mathrm{pa}_K) = 1,$$

we can factor out the K-th factor from the sum and then cancel out the remaining sum term in square brackets (because it is equal to 1):

$$\sum_{x_1} ... \sum_{x_{k-1}} \left[\sum_{x_K} p(x_K|\mathrm{pa}_K) \right] \prod_{k=1}^{K-1} p(x_k|\mathrm{pa}_k) = \sum_{x_1} ... \sum_{x_K - 1} \prod_{k=1}^{K-1} p(x_k|\mathrm{pa}_k)$$

Repeating this process $K-2$ times, we are left with one final marginalization step leading to

$$\sum_{x_1} p(x_1|\emptyset) = 1 \quad \Leftrightarrow \quad 1 = 1.$$