Machine Learning Principles and Methods	Due: April 24, 2017
Homework 1	
Instructor: Stephan Bongers, Tineke Blom	Email: s.r.bongers@uva.nl, t.blom2@uva.nl

You are allowed to discuss with your colleagues but you should write the answers in *your own words*. If you discuss with others, write down the name of your collaborators on top of the first page. No points will be deducted for collaborations. If we find similarities in solutions beyond the listed collaborations we will consider it as cheating.
We will not accept any late submissions under any circumstances. The solutions to the previous homework will be handed out in the class at the beginning of the next homework session. After this point, late submissions will be automatically graded zero.

Problem 1. (2 points)

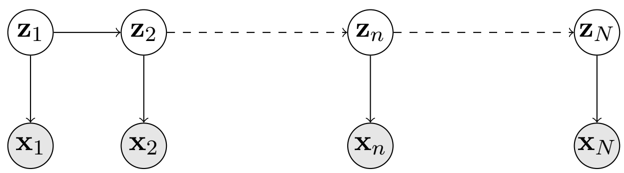

Figure 1: Markov chain of latent variables.

Given a graphical model in Figure 1. Show that:

1. $p(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}|\mathbf{x}_n, \mathbf{z}_n) = p(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}|\mathbf{z}_n)$
2. $p(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}|\mathbf{z}_{n-1}, \mathbf{z}_n) = p(\mathbf{x}_1, \ldots, \mathbf{x}_{n-1}|\mathbf{z}_{n-1})$
3. $p(\mathbf{x}_{n+1}, \ldots, \mathbf{x}_N|\mathbf{z}_n, \mathbf{z}_{n+1}) = p(\mathbf{x}_{n+1}, \ldots, \mathbf{x}_N|\mathbf{z}_{n+1})$
4. $p(\mathbf{z}_{N+1}|\mathbf{z}_N, \mathbf{X}) = p(\mathbf{z}_{N+1}|\mathbf{z}_N)$, where $\mathbf{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$

Use d-separation for the first two equalities and use the factorization properties of the graphical model for the third and fourth equalities.

Problem 2. (1.5 points)

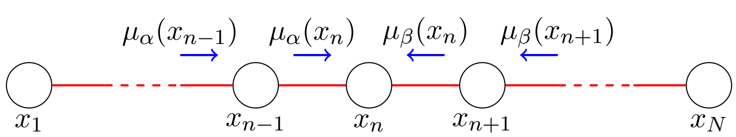

Figure 2: Chain of nodes model

Apply the sum-product algorithm (as in Bishop's section 8.4.4) to the chain of nodes model in Figure 2 and show that the results of message passing algorithm (as in Bishop's section 8.4.1) are recovered as a special case, that is

$$\begin{aligned} p(x_n) &= \frac{1}{Z}\mu_\alpha(x_n)\mu_\beta(x_n) \\ \mu_\alpha(x_n) &= \sum_{x_n-1} \psi_{n-1,n}(x_{n-1},x_n)\mu_\alpha(x_{n-1}) \\ \mu_\beta(x_n) &= \sum_{x_{n+1}} \psi_{n+1,n}(x_{n+1},x_n)\mu_\beta(x_{n+1}) \end{aligned}$$

where $\psi_{i,i+1}(x_i, x_{i+1})$ is a potential function defined over clique $\{x_i, x_{i+1}\}$.

Problem 3. (1.5 points) Consider the inference problem of evaluating $p(\mathbf{x}_n|\mathbf{x}_N)$ for the graph shown in Figure 2, for all nodes $n \in \{1, \ldots, N-1\}$. Show that the message passing algorithm can be used to solve this efficiently, and discuss which messages are modified and in what way.

Problem 4. (2 points) Run sum-product algorithm on the graph in Figure 3 with node x_3 designed as the root. Using the computed messages given in *p.409 Bishop*.

1. Show that the correct marginals are obtained for x_1 and x_3.
2. Show that the sum-product algorithm on this graph gives the correct joint distribution for x_1, x_2.

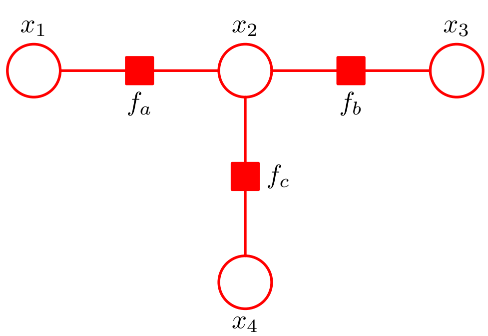

Figure 3: A simple factor graph

Problem 5. (2 points) Show that the marginal distribution for the variables $\mathbf{x}_s$ in a factor $f_s(\mathbf{x}_s)$ in a tree-structured factor graph, after running the sum-product message passing algorithm, can be written as

$$p(\mathbf{x}_s) = f_s(\mathbf{x}_s) \prod_{i \in \mathrm{ne}(f_s)} \mu_{x_i \to f_s(x_i)}$$

where $\mathrm{ne}(f_s)$ denotes the set of variable nodes that are neighbors of the factor node f_s