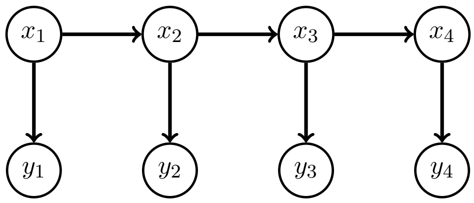
Exercises for the tutorials: 2(a-c) and 4(a-b).

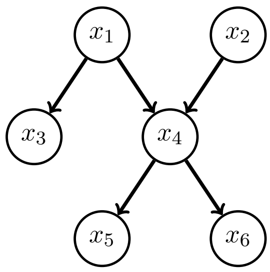
The other exercises are for self-study and exam preparation. All material is examinable unless otherwise mentioned.

Exercise 1. *Conversion to factor graphs*

(a) Draw an undirected factor graph for the directed graphical model defined by the graph below.

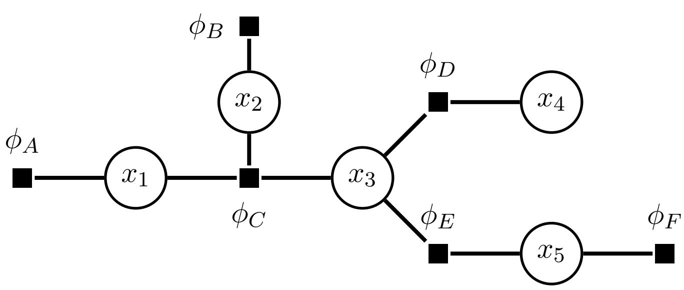


(b) Draw an undirected factor graph for directed graphical models defined by the graph below (this kind of graph is called a polytree: there are no loops but a node may have more than one parent).



Exercise 2. *Sum-product message passing*

We here re-consider the factor tree from the lecture on exact inference.



Let all variables be binary, $x_i \in \{0, 1\}$, and the factors be defined as follows:

x_1	ϕ_A
0	2
1	4

x_2	ϕ_B
0	4
1	4

x_1	x_2	x_3	ϕ_C
0	0	0	4
1	0	0	2
0	1	0	2
1	1	0	6
0	0	1	2
1	0	1	6
0	1	1	6
1	1	1	4

x_3	x_4	ϕ_D
0	0	8
1	0	2
0	1	2
1	1	6

x_3	x_5	ϕ_E
0	0	3
1	0	6
0	1	6
1	1	3

x_5	ϕ_F
0	1
1	8

(a) Mark the graph with arrows indicating all messages that need to be computed for the computation of $p(x_1)$.

(b) Compute the messages that you have identified.

Assuming that the computation of the messages is scheduled according to a common clock, group the messages together so that all messages in the same group can be computed in parallel during a clock cycle.

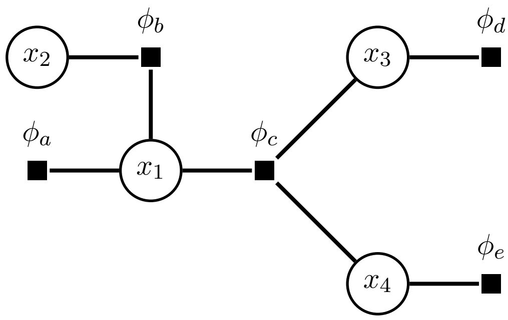
(c) What is $p(x_1 = 1)$?

(d) Draw the factor graph corresponding to $p(x_1, x_3, x_4, x_5 | x_2 = 1)$ and provide the numerical values for all factors.

(e) Compute $p(x_1 = 1 | x_2 = 1)$, re-using messages that you have already computed for the evaluation of $p(x_1 = 1)$.

Exercise 3. ***Sum-product message passing***

The following factor graph represents a Gibbs distribution over four binary variables $x_i \in \{0, 1\}$.



The factors ϕ_a, ϕ_b, ϕ_d are defined as follows:

x_1	ϕ_a
0	2
1	1

x_1	x_2	ϕ_b
0	0	5
1	0	2
0	1	2
1	1	6

x_3	ϕ_d
0	1
1	2

and $\phi_c(x_1, x_3, x_4) = 1$ if $x_1 = x_3 = x_4$, and is zero otherwise.

For all questions below, justify your answer:

(a) Compute the values of $\mu_{x_2 \to \phi_b}(x_2)$ for $x_2 = 0$ and $x_2 = 1$.

(b) Assume the message $\mu_{x_4 \to \phi_c}(x_4)$ equals

$$\mu_{x_4 \to \phi_c}(x_4) = \begin{cases} 1 & \text{if } x_4 = 0 \\ 3 & \text{if } x_4 = 1 \end{cases}$$

Compute the values of $\phi_e(x_4)$ for $x_4 = 0$ and $x_4 = 1$.

(c) Compute the values of $\mu_{\phi_c \to x_1}(x_1)$ for $x_1 = 0$ and $x_1 = 1$.

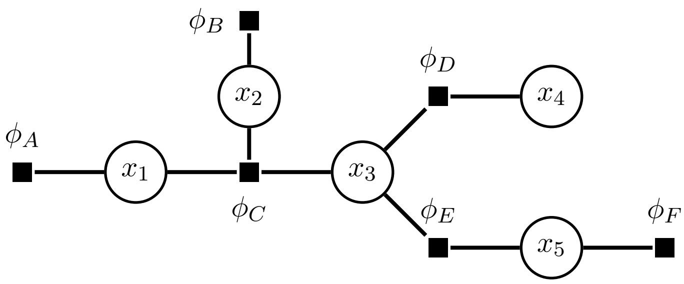
(d) The message $\mu_{\phi_b \to x_1}(x_1)$ equals

$$\mu_{\phi_b \to x_1}(x_1) = \begin{cases} 7 & \text{if } x_1 = 0 \\ 8 & \text{if } x_1 = 1 \end{cases}$$

What is the probability that $x_1 = 1$, i.e. $p(x_1 = 1)$?

Exercise 4. ***Max-sum message passing***

We here compute most probable states for the factor graph and factors below.



Let all variables be binary, $x_i \in \{0, 1\}$, and the factors be defined as follows:

x_1	ϕ_A
0	2
1	4

x_2	ϕ_B
0	4
1	4

x_1	x_2	x_3	ϕ_C
0	0	0	4
1	0	0	2
0	1	0	2
1	1	0	6
0	0	1	2
1	0	1	6
0	1	1	6
1	1	1	4

x_3	x_4	ϕ_D
0	0	8
1	0	2
0	1	2
1	1	6

x_3	x_5	ϕ_E
0	0	3
1	0	6
0	1	6
1	1	3

x_5	ϕ_F
0	1
1	8

(a) Will we need to compute the normalising constant Z to determine $\text{argmax}_{\mathbf{x}}\, p(x_1, \ldots, x_5)$?

(b) Compute $\text{argmax}_{x_1,x_2,x_3}\, p(x_1, x_2, x_3|x_4 = 0, x_5 = 0)$ via max-sum message passing.

(c) Compute $\text{argmax}_{x_1,\ldots,x_5}\, p(x_1, \ldots, x_5)$ via max-sum message passing with x_1 as root.

(d) Compute $\text{argmax}_{x_1,\ldots,x_5}\, p(x_1, \ldots, x_5)$ via max-sum message passing with x_3 as root.

Exercise 5. ***Inference for two linearly dependent Gaussian RVs***

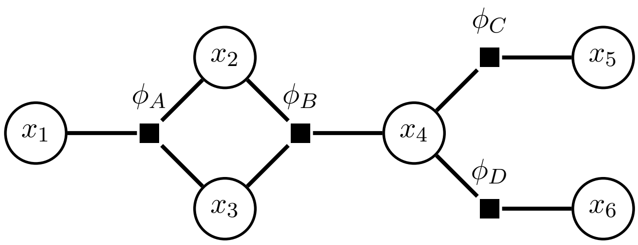
You have prior knowledge that an unknown variable $X \sim N(0, 1)$. You can make an observation of the variable Y which is related to X by $Y = wX + \mu_y + \epsilon$, with w and μ constants, and $\epsilon \sim N(0, \sigma^2)$ independent of X. The graphical model is $X \to Y$. (The factor w might arise e.g. because you want to measure X in centimeters, but your ruler is in inches. μ_y may arise due to an offset between the origins of the coordinates in the X and Y spaces.)

(a) Show that $\mathbb{E}[Y] = \mu_y$.

(b) You now want to make inferences for X given the observation $Y = y$. The conditional distribution $p(X = x|Y = y)$ is Gaussian. Compute its posterior mean and variance. HINT: You cause the conditioning formula for Gaussians given in the slides, or manipulate the expressions for the Gaussians $p(x)$ and $p(y|x)$ directly.

Exercise 6. ***Choice of elimination order in factor graphs***

Consider the following factor graph, which contains a loop:



Let all variables be binary, $x_i \in \{0, 1\}$, and the factors be defined as follows:

x_1	x_2	x_3	ϕ_A
0	0	0	4
1	0	0	2
0	1	0	2
1	1	0	6
0	0	1	2
1	0	1	6
0	1	1	6
1	1	1	4

x_2	x_3	x_4	ϕ_B
0	0	0	2
1	0	0	2
0	1	0	4
1	1	0	2
0	0	1	6
1	0	1	8
0	1	1	4
1	1	1	2

x_4	x_5	ϕ_C
0	0	8
1	0	2
0	1	2
1	1	6

x_4	x_6	ϕ_D
0	0	3
1	0	6
0	1	6
1	1	3

(a) Draw the factor graph corresponding to $p(x_2, x_3, x_4, x_5 \mid x_1 = 0, x_6 = 1)$ and give the tables defining the new factors $\phi_A^{x_1=0}(x_2, x_3)$ and $\phi_D^{x_6=1}(x_4)$ that you obtain.

(b) Find $p(x_2 \mid x_1 = 0, x_6 = 1)$ using the elimination ordering (x_4, x_5, x_3):

(i) Draw the graph for $p(x_2, x_3, x_5 \mid x_1 = 0, x_6 = 1)$ by marginalising x_4
Compute the table for the new factor $\tilde{\phi}_4(x_2, x_3, x_5)$

(ii) Draw the graph for $p(x_2, x_3 \mid x_1 = 0, x_6 = 1)$ by marginalising x_5
Compute the table for the new factor $\tilde{\phi}_{45}(x_2, x_3)$

(iii) Draw the graph for $p(x_2 \mid x_1 = 0, x_6 = 1)$ by marginalising x_3
Compute the table for the new factor $\tilde{\phi}_{453}(x_2)$

(c) Now determine $p(x_2 \mid x_1 = 0, x_6 = 1)$ with the elimination ordering (x_5, x_4, x_3):

(i) Draw the graph for $p(x_2, x_3, x_4, \mid x_1 = 0, x_6 = 1)$ by marginalising x_5
Compute the table for the new factor $\tilde{\phi}_5(x_4)$

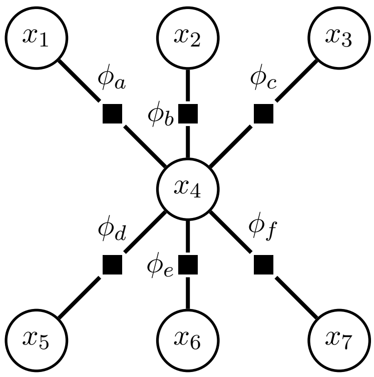
(ii) Draw the graph for $p(x_2, x_3 \mid x_1 = 0, x_6 = 1)$ by marginalising x_4
Compute the table for the new factor $\tilde{\phi}_{54}(x_2, x_3)$

(iii) Draw the graph for $p(x_2 \mid x_1 = 0, x_6 = 1)$ by marginalising x_3
Compute the table for the new factor $\tilde{\phi}_{543}(x_2)$

(d) Which variable ordering, (x_4, x_5, x_3) or (x_5, x_4, x_3) do you prefer?

Exercise 7. ***Choice of elimination order in factor graphs***

We would like to compute the marginal $p(x_1)$ by variable elimination for a joint pmf represented by the following factor graph. All variables x_i can take K different values.



(a) A friend proposes the elimination order $x_4, x_5, x_6, x_7, x_3, x_2$, i.e. to do x_4 first and x_2 last. Explain why this is computationally inefficient.

(b) Propose an elimination ordering that achieves $O(K^2)$ computational cost per variable elimination and explain why it does so.

