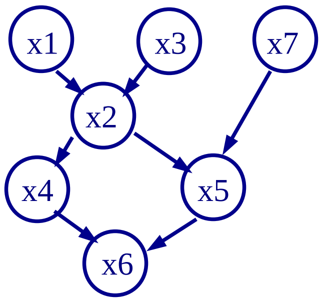
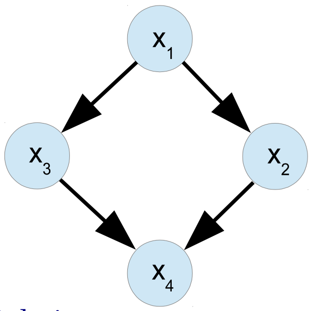
ML2, Summer 2020,
Exercise 2, Supervised structured learning

2.1. The joint probability in the variables x1,..., x7 shall be given as
p(x1, x2, x3, x4, x5, x6, x7) = p(x1)p(x3)p(x7)p(x2|x1,x3)p(x5|x7, x2)p(x4|x2)p(x6|x5, x4).
Draw the directed graphical model (DGM) for this joint probability!

Solution:



2.2. Write down the joint probability for the DGM given in the pictures:

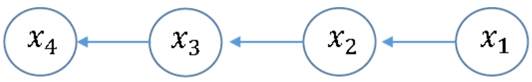


Solution:
p(x1)p(x3|x1)p(x2|x1)p(x4|x2,x3)

p(x1)p(x2|x1)p(x3|x2)p(x4|x3)

2.3. Max marginal inference in chains

A Bus goes from your home to the university where you leave it at the 4th stop.
With certain probabilities the bus arrives on time a the stops x_i. Calculate these probabilities given the directed graphical model (DGM):



Algorithm for marginal computation ("Sum-Product Message Passing")
1. Compute Messages from right to left ⇐
2.. Read out all marginals

and the following (conditional) probability tables $p(x_i|x_{i-1})$ (x=0 means late, x=1 on time):

x1	
0	0,29
1	0,71

x2|x1 x2	x1 = 0	x1 = 1
0	0,39	0,2
1	0,61	0,8

x3|x2 x3	x2 = 0	x2 = 1
0	0,77	0,51
1	0,23	0,49

x4|x3 x4	x3 = 0	x3 = 1
0	0,11	0,66
1	0,89	0,34

(equivalently: What is the maximum marginals solution of this probability distribution?)

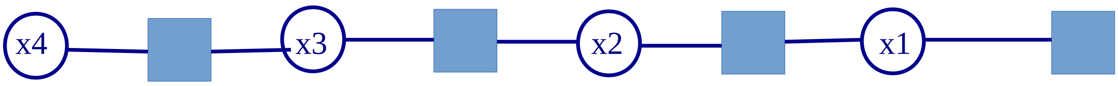
Solution:
p(x1)p(x2|x1)p(x3|x2)p(x4|x3)

$$p(x_4)=\sum_{x_3} p(x_4|x_3)\,p(x_3)=\sum_{x_3} p(x_4|x_3)\sum_{x_2} p(x_3|x_2)\sum_{x_1} p(x_2|x_1)\,p(x_1)$$

$$p(x_1=1)=0.71$$
$$p(x_2=1)=p(x_2=1|x_1=0)\,p(x_1=0)+p(x_2=1|x_1=1)\,p(x_1=1)$$
$$=0.61\cdot 0.29+0.71\cdot 0.8=0.745$$
$$p(x_2=0)=0.255$$
$$p(x_3=1)=0.424$$
$$p(x_4=1)=0.657$$

2.4. Factor graphs

For the chain in 2.2: Draw the faktor graph and write down the formula that corresponds to this factor graph. The conditional probabilities from the directed graphical model are inserted as factors here.
Solution:



$$p(x_{1,}x_{2,}x_{3,}x_4)=\psi(x_1)\,\psi(x_1,x_2)\,\psi(x_2,x_3)\,\psi(x_3,x_4)$$ (already normalized)

2.5. Let a factor be

$$\psi(x_1,x_2)=\begin{Bmatrix} x_1+x_2 \mid 0\leqslant x_1,x_2\leqslant 2 \\ 0 \quad \mid x_1,x_2<0;\, x_1,x_2>2 \end{Bmatrix}.$$ There are no more factors.

Which distribution $p(x_{1,}x_2)$ follows from that factor?

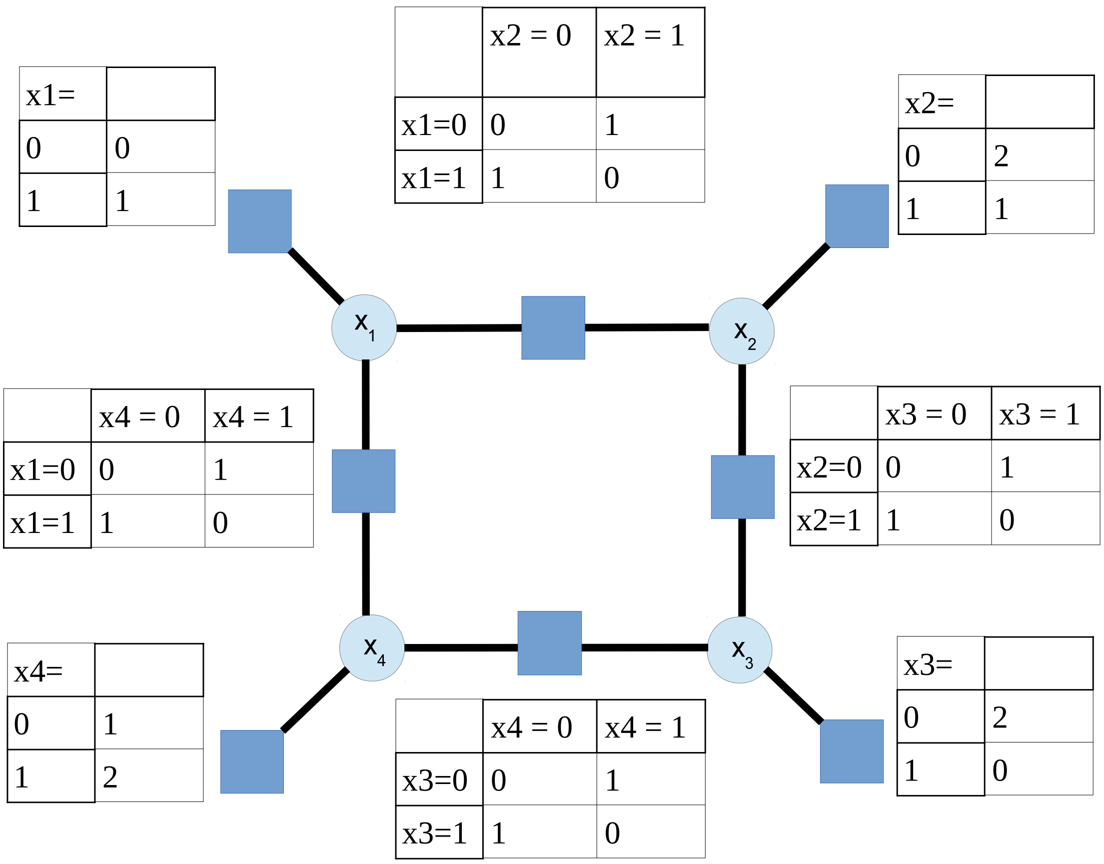
Solution:

`Normalization is needed:`

$$f=\int_0^2\int_0^2 (x_1+x_2)\,dx_1 dx_2=8$$

$$p(x_1,x_2)=\begin{Bmatrix} \frac{x_1+x_2}{8} \mid 0\leqslant x_1,x_2\leqslant 2 \\ 0 \quad \mid x_1,x_2<0;\, x_1,x_2>2 \end{Bmatrix}$$

2.6. Given is the following energy function
(of a Gibbs-Boltzmann-Distribution):



x1=	
0	0
1	1

	x2 = 0	x2 = 1
x1=0	0	1
x1=1	1	0

x2=	
0	2
1	1

	x4 = 0	x4 = 1
x1=0	0	1
x1=1	1	0

	x3 = 0	x3 = 1
x2=0	0	1
x2=1	1	0

x4=	
0	1
1	2

	x4 = 0	x4 = 1
x3=0	0	1
x3=1	1	0

x3=	
0	2
1	0

Calculate the MAP solution with ICM.

Solution: ICM for this task is explained in this video. As mentioned at the end of the video: start with another initialization and find out if you get the same solution.