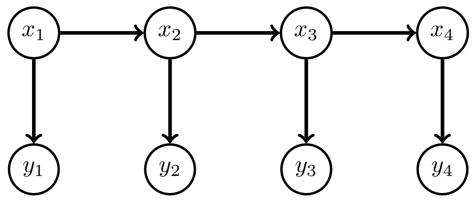
Exercises for the tutorials: 1 and 2.

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

Exercise 1. *Conversion to factor graphs*

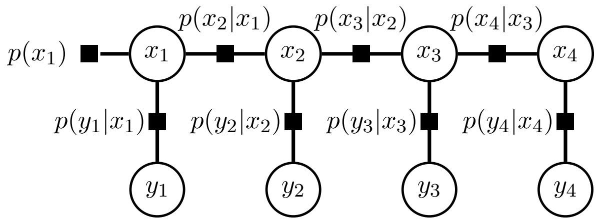
(a) Draw an undirected factor graph for the directed graphical model defined by the graph below.



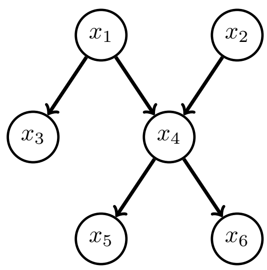
Solution. The graph specifies probabilistic models that factorise as

$$p(x_1, \ldots, x_4, y_1, \ldots, y_4) = p(x_1)p(y_1|x_1) \prod_{i=2}^{4} p(y_i|x_i)p(x_i|x_{i-1})$$

It is the graph for a hidden Markov model. The corresponding factor graph is shown 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).

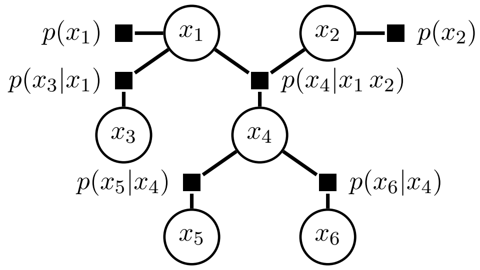


Solution. For the factor graph, we note that the directed graph defines the following class of probabilistic models

$$p(x_1, \ldots x_6) = p(x_1)p(x_2)p(x_3|x_1)p(x_4|x_1, x_2)p(x_5|x_4)p(x_6|x_4)$$

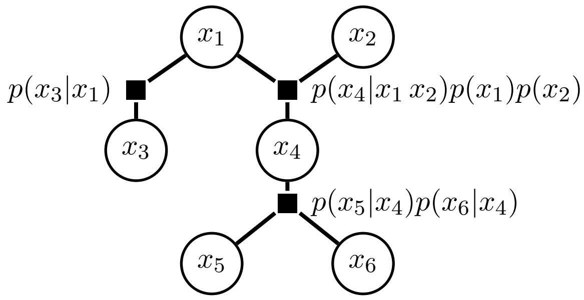
This gives the factor graph on right in the figure below.

©Michael U. Gutmann and Chris Williams, UoE, 2018-23 CC BY 4.0



Note:

- One may choose to group some factors together in order to obtain a factor graph with a particular structure (see factor graph below)



Exercise 2. ***Variable elimination for the "car start" network***

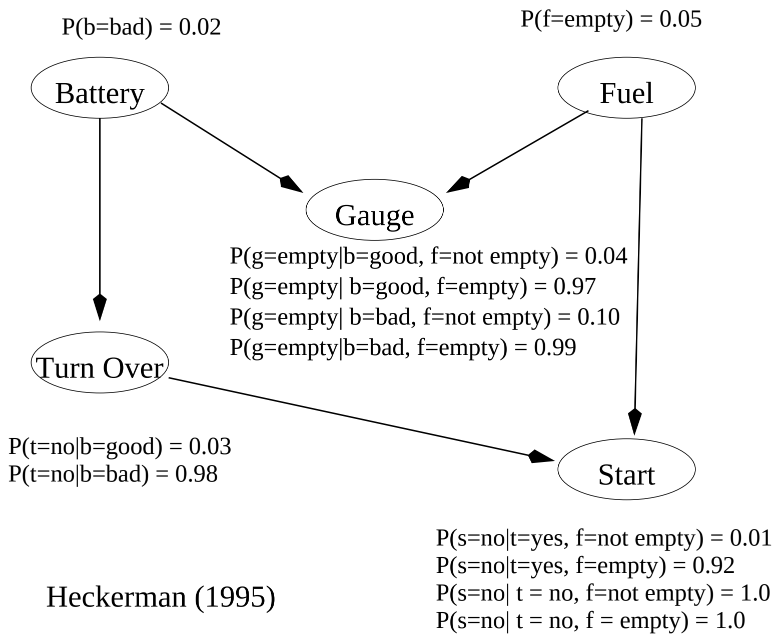


Figure 1: Simple network describing if a car's engine starts, due to Heckerman (1995).

Consider the belief network given in Figure 1, which gives the probability of a car starting (or not) (s), depending on the state of the battery (b), if it has fuel (f), the state of the fuel gauge (g), and if the car "turns over" (t).

©Michael U. Gutmann and Chris Williams, UoE, 2018-23 CC BY 4.0

Below we code the variables as 0/1, with:
$b = 0 \equiv b = bad$, *and* $b = 1 \equiv b = good$,
$f = 0 \equiv f = empty$, *and* $f = 1 \equiv f = not\ empty$,
$g = 0 \equiv g = empty$, *and* $g = 1 \equiv g = not\ empty$,
$t = 0 \equiv t = no$, *and* $t = 1 \equiv t = yes$,
$s = 0 \equiv s = no$, *and* $s = 1 \equiv s = yes$.

We have that $p(b = 0) = 0.02$, *and* $p(f = 0) = 0.05$. *The other CPTs are given by*

b	f	$p(g = 0\|b, f)$
0	0	0.99
0	1	0.10
1	0	0.97
1	1	0.04

b	$p(t = 0\|b)$
0	0.98
1	0.03

t	f	$p(s = 0\|t, f)$
0	0	1.0
0	1	1.0
1	0	0.92
1	1	0.01

(a) *From the graphical model we can read off*

$$p(b, f, g, t, s) = p(b)p(f)p(g|b, f)p(t|b)p(s|t, f). \tag{1}$$

Now consider the potential representation

$$p(b, f, g, t, s) \overset{def}{=} \phi_{bt}(b, t)\phi_{stf}(s, t, f)\phi_{gbf}(g, b, f).$$

Assign the factors in eq. 1 to the potentials in a valid way.

Solution. Here we assume the assignment $\phi_{bt}(b, t) = p(b)p(t|b)$, $\phi_{stf}(s, t, f) = p(f)p(s|t, f)$ and $\phi_{gbf}(g, b, f) = p(g|b, f)$. (Other valid assigments of the $p(b)$ and $p(f)$ factors to potentials are also possible.)

(b) *We wish to compute* $p(b, f|s = 0)$ *(i.e., the car does not start). This conditioning has the effect of turning*

$$p(b, f, g, t, s) = \phi_{bt}(b, t)\phi_{stf}(s, t, f)\phi_{gbf}(g, b, f)$$

into

$$p(b, f, g, t|s = 0) \propto \phi_{bt}(b, t)\phi_{tf}^{s=0}(t, f)\phi_{gbf}(g, b, f).$$

Identify the modification needed to the factor(s) in the potential ϕ_{stf} *to turn it into* $\phi_{tf}^{s=0}$.

Solution. To condition on $s = 0$, we have $\phi_{tf}^{s=0} = p(f)p(s = 0|t, f)$

(c) *First eliminate* g *to obtain a potential representation for* $p(b, f, t|s = 0)$.

Solution. We could compute a table with values of $\phi_{gbf}(g = 0, b, f)$ and $\phi_{gbf}(g = 1, b, f)$ for all for combinations of b and f values, but in our assignment of potentials we have chosen $\phi_{gbf}(g, b, f) = p(g|b, f)$. Hence summing out g will give the value of 1 for all values of b and f.

$$p(b, f, t|s = 0) \propto \phi_{bt}(b, t)\phi_{tf}^{s=0}(t, f).$$

(d) *Now eliminate* t *to obtain a potential representation for* $p(b, f|s = 0)$. *Evaluate this for all four possible combinations of* b *and* f. *Normalize this to obtain the posterior probability of all four configurations of* b *and* f.

©Michael U. Gutmann and Chris Williams, UoE, 2018-23 CC BY 4.0

Solution. We wish to sum out the variable t in the potential representation

$$p(b, f, t|s = 0) \propto \phi_{bt}(b, t)\phi_{tf}^{s=0}(t, f).$$

There is no shortcut for this, we need to compute the potentials $\phi_{bt}(b, t)$ and $\phi_{tf}^{s=0}(t, f)$ as tables, and sum out t to obtain $\tilde{\phi}_{bf}^{s=0}$. To obtain the posterior probability we normalize by summing over the 4 configurations to obtain
$Z = \tilde{\phi}_{bf}^{s=0}(b = 0, f = 0) + \tilde{\phi}_{bf}^{s=0}(b = 0, f = 1) + \tilde{\phi}_{bf}^{s=0}(b = 1, f = 0) + \tilde{\phi}_{bf}^{s=0}(b = 1, f = 1)$.

b	t	$p(b)$	$p(t\|b)$	ϕ_{bt}
0	0	0.02	0.98	0.0196
0	1	0.02	1-0.98	0.0004
1	0	1-0.02	0.03	0.0294
1	1	1-0.02	1-0.03	0.9506

t	f	$p(f)$	$p(s = 0\|t, f))$	$\phi_{tf}^{s=0}$
0	0	0.05	1.0	0.05
0	1	1-0.05	1.0	0.95
1	0	0.05	0.92	0.046
1	1	1-0.05	0.01	0.0095

b	f	$\phi_{b,t=0} * \phi_{t=0,f}^{s=0}$	$\phi_{b,t=1} * \phi_{t=1,f}^{s=0}$	$\tilde{\phi}_{bf}^{s=0}$	$\tilde{\phi}_{bf}^{s=0}/Z$
0	0	0.0196*0.05	0.0004*0.046	0.0009984	0.0098
0	1	0.0196*0.95	0.0004*0.0095	0.0186	0.1830
1	0	0.0294*0.05	0.9506*0.046	0.0452	0.4441
1	1	0.0294*0.95	0.9506*0.0095	0.0370	0.3631

where $Z = 0.0009984 + 0.0186 + 0.0452 + 0.0370 = 0.1018$ (working to 4 dec. pl.).

(e) *Wellman and Henrion (1993)*[1] *state:*

> *"Explaining away" is a common pattern of reasoning in which the confirmation of one cause of an observed or believed event reduces the need to invoke alternative causes.*

Think about the four posterior probabilites obtained above. Given that the car does not start, do these broadly meet with your expectations? Do you observe "explaining away"?

Solution. The MAP configuration is for $b = 1,\ f = 0$, i.e. the fuel is empty but the battery is good.

Perhaps surprisingly the second most probable configuration is $b = 1,\ f = 1$, i.e. the battery is good and the fuel is not empty. Although the probability that $t = 0$ is only 0.03 if $b = 1$, the fact that $p(s = 0|t = 0, f = 1) = 1.0$ means that there is signficant mass in the posterior on $b = 1,\ f = 1$.

The configuration $b = 0,\ f = 1$ also has high posterior probability, as the bad battery will give rise to a low probability of the car turning over, and hence starting.

It is notable that the posterior probability of a "double failure" $b = 0,\ f = 0$ is less than 1%. Thus we do observe explaining away—setting either $b = 0$ (with $f = 1$) or $f = 0$ (with $b = 1$) is enough to explain the failure of the car to start, and the very improbable double failure is not supported.

(f) *In general we write computer programs to do inference in such models. See* `https://github.com/vsimkus/pmr2023-pgm-demo` *for a demo for this question prepared by the PMR TA Vaidotas Simkus using the* `pgmpy` *Python package. You can run the notebook on Google Colab directly via the link* `http://colab.research.google.com/github/vsimkus/pmr2023-pgm-demo`.

[1] IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(3) pp 287–292 (1993).

©Michael U. Gutmann and Chris Williams, UoE, 2018-23 CC BY 4.0

Exercise 3. ***Limits of directed and undirected graphical models***

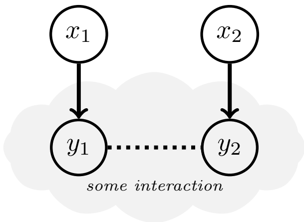
We here consider the probabilistic model $p(y_1, y_2, x_1, x_2) = p(y_1, y_2|x_1, x_2)p(x_1)p(x_2)$ *where* $p(y_1, y_2|x_1, x_2)$ *factorises as*

$$p(y_1, y_2|x_1, x_2) = p(y_1|x_1)p(y_2|x_2)\phi(y_1, y_2)n(x_1, x_2) \tag{2}$$

with $n(x_1, x_2)$ *equal to*

$$n(x_1, x_2) = \left(\int p(y_1|x_1)p(y_2|x_2)\phi(y_1, y_2)\mathrm{d}y_1\mathrm{d}y_2\right)^{-1}. \tag{3}$$

In the lecture "Factor Graphs", we used the model to illustrate the setup where x_1 *and* x_2 *are two independent inputs that each control the interacting variables* y_1 *and* y_2 *(see graph below).*



(a) *Use the basic characterisations of statistical independence*

$$u \perp\!\!\!\perp v|z \Longleftrightarrow p(u, v|z) = p(u|z)p(v|z) \tag{4}$$

$$u \perp\!\!\!\perp v|z \Longleftrightarrow p(u, v|z) = a(u, z)b(v, z) \qquad (a(u, z) \geq 0, b(v, z) \geq 0) \tag{5}$$

to show that $p(y_1, y_2, x_1, x_2)$ *satisfies the following independencies*

$$x_1 \perp\!\!\!\perp x_2 \qquad\qquad x_1 \perp\!\!\!\perp y_2 \mid y_1, x_2 \qquad\qquad x_2 \perp\!\!\!\perp y_1 \mid y_2, x_1$$

Solution. The pdf/pmf is

$$p(y_1, y_2, x_1, x_2) = p(y_1|x_1)p(y_2|x_2)\phi(y_1, y_2)n(x_1, x_2)p(x_1)p(x_2)$$

For $\mathbf{x_1} \perp\!\!\!\perp \mathbf{x_2}$
We compute $p(x_1, x_2)$ as

$$\begin{aligned}
p(x_1, x_2) &= \int p(y_1, y_2, x_1, x_2)\mathrm{d}y_1\mathrm{d}y_2 && \text{(S.1)}\\
&= \int p(y_1|x_1)p(y_2|x_2)\phi(y_1, y_2)n(x_1, x_2)p(x_1)p(x_2)\mathrm{d}y_1\mathrm{d}y_2 && \text{(S.2)}\\
&= n(x_1, x_2)p(x_1)p(x_2)\int p(y_1|x_1)p(y_2|x_2)\phi(y_1, y_2)\mathrm{d}y_1\mathrm{d}y_2 && \text{(S.3)}\\
&\overset{(3)}{=} n(x_1, x_2)p(x_1)p(x_2)\frac{1}{n(x_1, x_2)} && \text{(S.4)}\\
&= p(x_1)p(x_2). && \text{(S.5)}
\end{aligned}$$

Since $p(x_1)$ and $p(x_2)$ are the univariate marginals of x_1 and x_2, respectively, it follows from (4) that $x_1 \perp\!\!\!\perp x_2$.

For $\mathbf{x_1} \perp\!\!\!\perp \mathbf{y_2} \mid \mathbf{y_1}, \mathbf{x_2}$
We rewrite $p(y_1, y_2, x_1, x_2)$ as

$$\begin{aligned}
p(y_1, y_2, x_1, x_2) &= p(y_1|x_1)p(y_2|x_2)\phi(y_1, y_2)n(x_1, x_2)p(x_1)p(x_2) && \text{(S.6)}\\
&= [p(y_1|x_1)p(x_1)n(x_1, x_2)]\,[p(y_2|x_2)\phi(y_1, y_2)p(x_2)] && \text{(S.7)}\\
&= \phi_A(x_1, y_1, x_2)\phi_B(y_2, y_1, x_2) && \text{(S.8)}
\end{aligned}$$

©Michael U. Gutmann and Chris Williams, UoE, 2018-23 CC BY 4.0

With (5), we have that $x_1 \perp\!\!\!\perp y_2 \mid y_1, x_2$. Note that $p(x_2)$ can be associated either with ϕ_A or with ϕ_B.

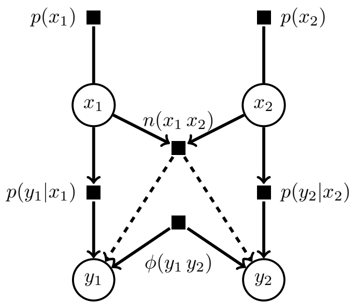
For $\mathbf{x_2} \perp\!\!\!\perp \mathbf{y_1} \mid \mathbf{y_2}, \mathbf{x_1}$

We use here the same approach as for $x_1 \perp\!\!\!\perp y_2 \mid y_1, x_2$. (By symmetry considerations, we could immediately see that the relation holds but let us write it out for clarity). We rewrite $p(y_1, y_2, x_1, x_2)$ as

$$\begin{aligned} p(y_1, y_2, x_1, x_2) &= p(y_1|x_1)p(y_2|x_2)\phi(y_1, y_2)n(x_1, x_2)p(x_1)p(x_2) && \text{(S.9)} \\ &= [p(y_2|x_2)n(x_1, x_2)p(x_2)p(x_1))]\,[p(y_1|x_1)\phi(y_1, y_2)]) && \text{(S.10)} \\ &= \tilde{\phi}_A(x_2, x_1, y_2)\tilde{\phi}_B(y_1, y_2, x_1) && \text{(S.11)} \end{aligned}$$

With (5), we have that $x_2 \perp\!\!\!\perp y_1 \mid y_2, x_1$.

(b) (optional, not examinable) *The following factor graph represents $p(y_1, y_2, x_1, x_2)$:*



Use the separation rules for factor graphs to verify that we can find all independence relations. The separation rules are (see Barber, section 4.4.1, or the original paper by Brendan Frey: https://arxiv.org/abs/1212.2486):

"If all paths are blocked, the variables are conditionally independent. A path is blocked if one or more of the following conditions is satisfied:

1. *One of the variables in the path is in the conditioning set.*
2. *One of the variables or factors in the path has two incoming edges that are part of the path (variable or factor collider), and neither the variable or factor nor any of its descendants are in the conditioning set."*

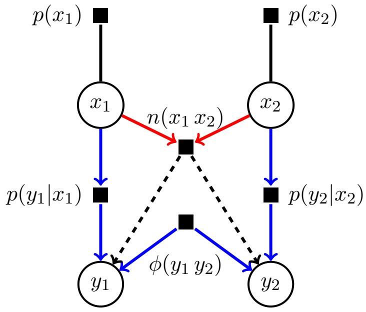
Remarks:

- *"one or more of the following" should best be read as "one of the following".*
- *"incoming edges" means directed incoming edges*
- *the descendants of a variable or factor node are all the variables that you can reach by following a path (containing directed or directed edges, but for directed edges, all directions have to be consistent)*
- *In the graph we have dashed directed edges: they do count when you determine the descendants but they do not contribute to paths. For example, y_1 is a descendant of the $n(x_1, x_2)$ factor node but $x_1 - n - y_2$ is not a path.*

Solution. $\mathbf{x_1} \perp\!\!\!\perp \mathbf{x_2}$

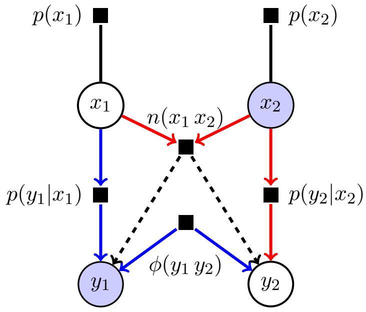
There are two paths from x_1 to x_2 marked with red and blue below:

©Michael U. Gutmann and Chris Williams, UoE, 2018-23 CC BY 4.0



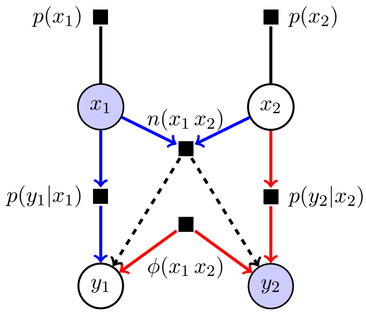
Both the blue and red path are blocked by condition 2.

$\mathbf{x_1 \perp\!\!\!\perp y_2 \mid y_1, x_2}$
There are two paths from x_1 to y_2 marked with red and blue below:



The observed variables are marked in blue. For the red path, the observed x_2 blocks the path (condition 1). Note that the $n(x_1, x_2)$ node would be open by condition 2. The blue path is blocked by condition 1 too. In directed graphical models, the y_1 node would be open, but here while condition 2 does not apply, condition 1 still applies (note the *one or more of ...* in the separation rules), so that the path is blocked.

$\mathbf{x_2 \perp\!\!\!\perp y_1 \mid y_2, x_1}$
There are two paths from x_2 to y_1 marked with red and blue below:



The same reasoning as before yields the result.

Finally note that x_1 and x_2 are not independent given y_1 or y_2 because the upper path through $n(x_1, x_2)$ is not blocked whenever y_1 or y_2 are observed (condition 2).

Credit: this example is discussed in the original paper by B. Frey (Figure 6).

©Michael U. Gutmann and Chris Williams, UoE, 2018-23 CC BY 4.0