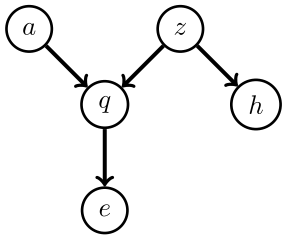
Exercise 1. ***More on ordered and local Markov properties, d-separation***

We continue with the investigation of the graph below

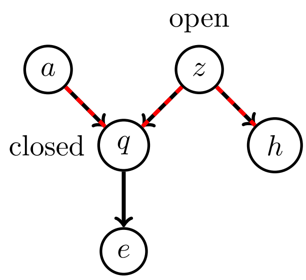


(a) *Why can the ordered or local Markov property not be used to check whether $a \perp\!\!\!\perp h \mid e$ may hold?*

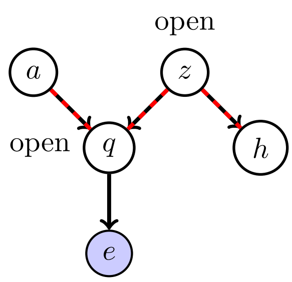
Solution. The independencies that follow from the ordered or local Markov property require conditioning on parent sets. However, e is not a parent of any node so that the above independence assertion cannot be checked via the ordered or local Markov property.

(b) *Use d-separation to check whether $a \perp\!\!\!\perp h \mid e$ holds.*

Solution. The trail from a to h is shown below in red together with the default states of the nodes along the trail.



Conditioning on e opens the q node since q in a collider configuration on the path.



The trail from a to h is thus active, which means that the relationship does not hold because $a \not\perp\!\!\!\perp h \mid e$ for some distributions that factorise over the graph.

(c) *The independency relations obtained via the ordered and local Markov property include $a \perp\!\!\!\perp \{z, h\}$. Verify the independency using d-separation.*

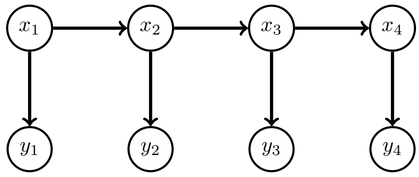
Solution. All paths from a to z or h pass through the node q that forms a head-head connection along that trail. Since neither q nor its descendant e is part of the conditioning set, the trail is blocked and the independence relation follows.

(d) *Determine the Markov blanket of z.*

Solution. The Markov blanket is given by the parents, children, and co-parents. Hence: $\mathrm{MB}(z) = \{a, q, h\}$.

Exercise 2. *Hidden Markov models*

This exercise is about directed graphical models that are specified by the following DAG:



These models are called "hidden" Markov models because we typically assume to only observe the y_i and not the x_i that follow a Markov model.

(a) *Show that all probabilistic models specified by the DAG factorise as*

$$p(x_1, y_1, x_2, y_2, \dots, x_4, y_4) = p(x_1)p(y_1|x_1)p(x_2|x_1)p(y_2|x_2)p(x_3|x_2)p(y_3|x_3)p(x_4|x_3)p(y_4|x_4)$$

Solution. From the definition of directed graphical models it follows that

$$p(x_1, y_1, x_2, y_2, \dots, x_4, y_4) = \prod_{i=1}^{4} p(x_i|\mathrm{pa}(x_i)) \prod_{i=1}^{4} p(y_i|\mathrm{pa}(y_i)).$$

The result is then obtained by noting that the parent of y_i is given by x_i for all i, and that the parent of x_i is x_{i-1} for $i = 2, 3, 4$ and that x_1 does not have a parent ($\mathrm{pa}(x_1) = \varnothing$).

(b) *Derive the independencies implied by the ordered Markov property with the topological ordering $(x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4)$*

Solution.

$$y_i \perp\!\!\!\perp x_1, y_1, \dots, x_{i-1}, y_{i-1} \mid x_i \qquad x_i \perp\!\!\!\perp x_1, y_1, \dots, x_{i-2}, y_{i-2}, y_{i-1} \mid x_{i-1}$$

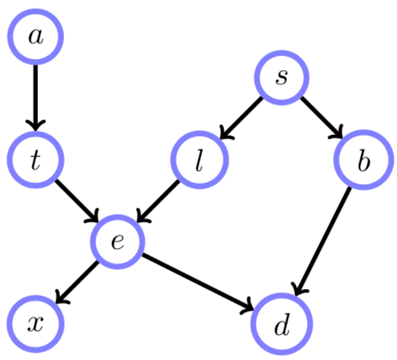
(c) *Derive the independencies implied by the ordered Markov property with the topological ordering $(x_1, x_2, \dots, x_4, y_1, \dots, y_4)$.*

Solution. For the x_i, we use that for $i \geq 2$: $\text{pre}(x_i) = \{x_1, \ldots, x_{i-1}\}$ and $\text{pa}(x_i) = x_{i-1}$. For the y_i, we use that $\text{pre}(y_1) = \{x_1, \ldots, x_4\}$, that $\text{pre}(y_i) = \{x_1, \ldots, x_4, y_1, \ldots, y_{i-1}\}$ for $i > 1$, and that $\text{pa}(y_i) = x_i$. The ordered Markov property then gives:

$$
\begin{aligned}
x_3 &\perp\!\!\!\perp x_1 \mid x_2 & x_4 &\perp\!\!\!\perp \{x_1, x_2\} \mid x_3 \\
y_1 &\perp\!\!\!\perp \{x_2, x_3, x_4\} \mid x_1 & y_2 &\perp\!\!\!\perp \{x_1, x_3, x_4, y_1\} \mid x_2 \\
y_3 &\perp\!\!\!\perp \{x_1, x_2, x_4, y_1, y_2\} \mid x_3 & y_4 &\perp\!\!\!\perp \{x_1, x_2, x_3, y_1, y_2, y_3\} \mid x_4
\end{aligned}
$$

Exercise 3. *More on the chest clinic (based on Barber's exercise 3.3)*

The directed graphical model in Figure 1 is the "Asia" example of Lauritzen and Spiegelhalter (1988). It concerns the diagnosis of lung disease (T=tuberculosis or L=lung cancer). In this model, a visit to some place in A=Asia is thought to increase the probability of tuberculosis.



$x =$ Positive X-ray
$d =$ Dyspnea (Shortness of breath)
$e =$ Either Tuberculosis or Lung Cancer
$t =$ Tuberculosis
$l =$ Lung Cancer
$b =$ Bronchitis
$a =$ Visited Asia
$s =$ Smoker

Figure 1: Graphical model for Exercise 3 (Barber Figure 3.15).

(a) *Explain which of the following independence relationships hold for all distributions that factorise over the graph.*

1. $a \perp\!\!\!\perp s \mid l$

Solution.

- There are two trails from a to s: (a, t, e, l, s) and (a, t, e, d, b, s)
- The trail (a, t, e, l, s) features a collider node e that blocks the trail (the trail is also blocked by l).
- The trail (a, t, e, d, b, s) is blocked by the collider node d.
- All trails are blocked so that the independence relation holds.

2. $a \perp\!\!\!\perp s \mid l, d$

Solution.

- There are two trails from a to s: (a, t, e, l, s) and (a, t, e, d, b, s)
- The trail (a, t, e, l, s) features a collider node e that is opened by the conditioning variable d but the l node is closed by the conditioning variable l: the trail is blocked

- The trail (a, t, e, d, b, s) features a collider node d that is opened by conditioning on d. On this trail, e is not in a head-head (collider) configuration) so that all nodes are open and the trail active.
- Hence, the independence relation does generally not hold.

(b) *Let g be a (deterministic) function of x and t. Is the expected value $\mathbb{E}[g(x,t) \mid l, b]$ equal to $\mathbb{E}[g(x,t) \mid l]$?*

Solution. The question boils down to checking whether $x, t \perp\!\!\!\perp b \mid l$. For the independence relation to hold, all trails from both x and t to b need to be blocked by l.

- For x, we have the trails (x, e, l, s, b) and (x, e, d, b)
- Trail (x, e, l, s, b) is blocked by l
- Trail (x, e, d, b) is blocked by the collider configuration of node d.
- For t, we have the trails (t, e, l, s, b) and (t, e, d, b)
- Trail (t, e, l, s, b) is blocked by l.
- Trail (t, e, d, b) is blocked by the collider configuration of node d.

As all trails are blocked we have $x, t \perp\!\!\!\perp b \mid l$ and $\mathbb{E}[g(x,t) \mid l, b] = \mathbb{E}[g(x,t) \mid l]$.

Exercise 4. *Independencies*

This exercise is on further properties and characterisations of statistical independence.

(a) *Without using d-separation, show that $x \perp\!\!\!\perp \{y, w\} \mid z$ implies that $x \perp\!\!\!\perp y \mid z$ and $x \perp\!\!\!\perp w \mid z$.*
Hint: use the definition of statistical independence in terms of the factorisation of pmfs/pdfs.

Solution. We consider the joint distribution $p(x, y, w|z)$. By assumption

$$p(x, y, w|z) = p(x|z)p(y, w|z) \tag{S.1}$$

We have to show that $x \perp\!\!\!\perp y|z$ and $x \perp\!\!\!\perp w|z$. For simplicity, we assume that the variables are discrete valued. If not, replace the sum below with an integral.

To show that $x \perp\!\!\!\perp y|z$, we marginalise $p(x, y, w|z)$ over w to obtain

$$p(x, y|z) = \sum_w p(x, y, w|z) \tag{S.2}$$

$$= \sum_w p(x|z)p(y, w|z) \tag{S.3}$$

$$= p(x|z) \sum_w p(y, w|z) \tag{S.4}$$

Since $\sum_w p(y, w|z)$ is the marginal $p(y|z)$, we have

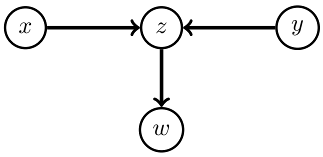
$$p(x, y|z) = p(x|z)p(y|z), \tag{S.5}$$

which means that $x \perp\!\!\!\perp y|z$.

To show that $x \perp\!\!\!\perp w|z$, we similarly marginalise $p(x, y, w|z)$ over y to obtain $p(x, w|z) = p(x|z)p(w|z)$, which means that $x \perp\!\!\!\perp w|z$.

(b) *For the directed graphical model below, show that the following two statements hold without using d-separation:*

$$x \perp\!\!\!\perp y \quad \textit{and} \tag{1}$$
$$x \not\perp\!\!\!\perp y \mid w \tag{2}$$



The exercise shows that not only conditioning on a collider node but also on one of its descendents activates the trail between x and y. You can use the result that $x \perp\!\!\!\perp y|w \Leftrightarrow p(x, y, w) = a(x, w)b(y, w)$ for some non-negative functions $a(x, w)$ and $b(y, w)$.

Solution. The graphical model corresponds to the factorisation

$$p(x, y, z, w) = p(x)p(y)p(z|x, y)p(w|z).$$

For the marginal $p(x, y)$ we have to sum (integrate) over all (z, w)

$$\begin{aligned}
p(x, y) &= \sum_{z,w} p(x, y, z, w) && \text{(S.6)}\\
&= \sum_{z,w} p(x)p(y)p(z|x, y)p(w|z) && \text{(S.7)}\\
&= p(x)p(y) \sum_{z,w} p(z|x, y)p(w|z) && \text{(S.8)}\\
&= p(x)p(y) \underbrace{\sum_{z} p(z|x, y)}_{1} \underbrace{\sum_{w} p(w|z)}_{1} && \text{(S.9)}\\
&= p(x)p(y) && \text{(S.10)}
\end{aligned}$$

Since $p(x, y) = p(x)p(y)$ we have $x \perp\!\!\!\perp y$.
For $x \not\perp\!\!\!\perp y|w$, compute $p(x, y, w)$ and use the result $x \perp\!\!\!\perp y|w \Leftrightarrow p(x, y, w) = a(x, w)b(y, w)$.

$$\begin{aligned}
p(x, y, w) &= \sum_{z} p(x, y, z, w) && \text{(S.11)}\\
&= \sum_{z} p(x)p(y)p(z|x, y)p(w|z) && \text{(S.12)}\\
&= p(x) \underbrace{p(y) \sum_{z} p(z|x, y)p(w|z)}_{k(x,y,w)} && \text{(S.13)}
\end{aligned}$$

Since $p(x, y, w)$ cannot be factorised as $a(x, w)b(y, w)$, the relation $x \perp\!\!\!\perp y|w$ cannot generally hold.