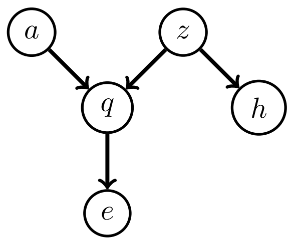
The purpose of the tutorials is twofold: First, they help you better understand the lecture material. Secondly, they provide exam preparation material. You are not expected to complete all questions before the tutorial sessions. Start early and do as many as you have time for.

Exercise 1. ***Directed graph concepts***

We here consider the directed graph below that was partly discussed in the lecture.



(a) List all trails in the graph (of maximal length)

(b) List all directed paths in the graph (of maximal length)

(c) What are the descendants of z?

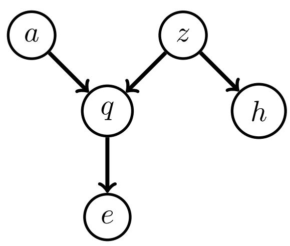
(d) What are the non-descendants of q?

(e) Which of the following orderings are topological to the graph?

- (a,z,h,q,e)
- (a,z,e,h,q)
- (z,a,q,h,e)
- (z,q,e,a,h)

Exercise 2. ***Ordered and local Markov properties, d-separation***

We continue with the investigation of the graph from Exercise 1 shown below for reference.



(a) The ordering (z, h, a, q, e) is topological to the graph. What are the independencies that follow from the ordered Markov property?

(b) What are the independencies that follow from the local Markov property?

(c) The independency relations obtained via the ordered and local Markov property include

$$a \perp\!\!\!\perp \{z, h\} \qquad q \perp\!\!\!\perp h \mid \{a, z\}$$

Verify them by d-separation

(d) Verify that $q \perp\!\!\!\perp h \mid \{a, z\}$ holds by manipulating the probability distribution induced by the graph.

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

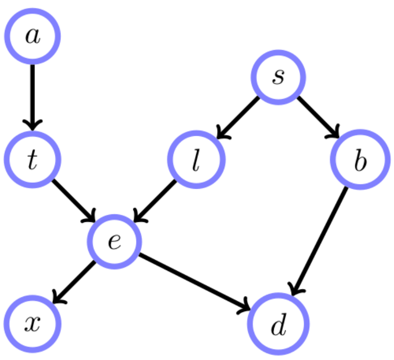
(f) Use d-separation to check whether $a \perp\!\!\!\perp h \mid e$ holds.

(g) Determine the Markov blanket of z.

(h) Assume all variables in the graph are binary. How many numbers do you need to specify, or learn from data, in order to fully specify the probability distribution?

Exercise 3. ***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. $t \perp\!\!\!\perp s \mid d$
2. $l \perp\!\!\!\perp b \mid s$
3. $a \perp\!\!\!\perp s \mid l$
4. $a \perp\!\!\!\perp s \mid l, d$

(b) Can we simplify $p(l|b, s)$ to $p(l|s)$?

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

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 $x \perp\!\!\!\perp w \mid z$.

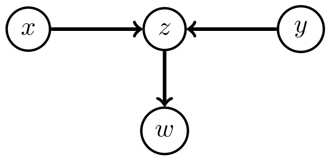
(b) We have seen that $x \perp\!\!\!\perp y|z$ is characterised by $p(x, y|z) = p(x|z)p(y|z)$ or, equivalently, by $p(x|y, z) = p(x|z)$. Show that further equivalent characterisations are

$$p(x, y, z) = p(x|z)p(y|z)p(z) \quad \text{and} \tag{1}$$

$$p(x, y, z) = a(x, z)b(y, z) \quad \text{for some non-neg. functions } \ a(x, z) \ \text{ and } \ b(x, z). \tag{2}$$

The characterisation in Equation (2) will be important for undirected graphical models.

(c) For the directed graphical model below, show that $x \perp\!\!\!\perp y$ and generally $x \not\perp\!\!\!\perp y \mid w$ without using d-separation.



The exercise shows that not only conditioning on a collider but also on a descendent of it activates the trail between x and y.