

Exercise 1. *I-maps*

(a) Which of three graphs represent the same set of independencies? Explain.



Solution. To check whether the graphs are I-equivalent, we have to check the skeletons and the immoralities. All have the same skeleton, but graph 1 and graph 2 also have the same immorality. The answer is thus: graph 1 and 2 encode the same independencies.



- (b) Let p(a, z, q, e, h) = p(a)p(z)p(q|a, z)p(e|q)p(h|z). In the lecture "Expressive Power of Graphical Models", we have seen that the graph in Figure 1 is a minimal I-map for p. Assuming that the graph is a perfect map for p, determine minimal directed I-maps for the orderings
 - (a, z, h, q, e)
 - (e, h, q, z, a)

Are the two I-maps that you obtain I-equivalent?



Figure 1: Perfect map for Exercise 1, question (b).

Solution. To find a minimal I-map, we can use the same procedure we used to simplify the factorisation obtained by the chain rule:

1. Assume an ordering of the variables. Denote the ordered random variables by x_1, \ldots, x_d .

2. For each *i*, find a minimal subset of variables $\pi_i \subseteq \text{pre}_i$ such that

$$x_i \perp \{ \operatorname{pre}_i \setminus \pi_i \} \mid \pi_i$$

holds for p.

3. Construct a graph with parents $pa_i = \pi_i$.

Checking whether the independencies hold can be difficult. But here, we are given a factorisation for p from which we can construct the directed graph in Figure 1. We thus can use graphical methods to check whether an independency holds (e.g. by d-separation). Note: If the graph does not indicate that a certain independency holds, we had to generally check, however, whether it indeed does not hold for a specific distribution. If we don't, we won't obtain a minimal I-map but just an I-map. This is because the graph may not be a perfect map, and p may have independencies that are not encoded in the graph. This is why here, we made the assumption that the graph in Figure 1 is a perfect map; for perfect maps, the lack of an independence assertion by the graph (d-connection) means that the variables are indeed not independent, i.e. that they are dependent.

The ordering (a, z, h, q, e) gives rise to the same graph as in Figure 1.

For the ordering (e, h, q, z, a), we build a graph where e is the root. From Figure 1 (and the perfect map assumption), we see that $h \perp e$ does not hold. We thus set e as parent of h, see first graph in Figure 2. Then:

- We consider q: $\operatorname{pre}_q = \{e, h\}$. There is no subset π_q of pre_q on which we could condition to make q independent of $\operatorname{pre}_q \setminus \pi_q$, so that we set the parents of q in the graph to $\operatorname{pa}_q = \{e, h\}$. (Second graph in Figure 2.)
- We consider z: $\operatorname{pre}_z = \{e, h, q\}$. From the graph in Figure 1, we see that for $\pi_z = \{q, h\}$ we have $z \perp \operatorname{pre}_z \setminus \pi_z | \pi_z$. Note that $\pi_z = \{q\}$ does not work because $z \perp e, h | q$ does not hold. We thus set $\operatorname{pa}_z = \{q, h\}$. (Third graph in Figure 2.)
- We consider a: $\operatorname{pre}_a = \{e, h, q, z\}$. This is the last node in the ordering. To find the minimal set π_a for which $a \perp \operatorname{pre}_a \setminus \pi_a | \pi_a$, we can determine its Markov blanket MB(a). The Markov blanket is the set of parents (none), children (q), and co-parents of a(z) in Figure 1, so that MB(a) = $\{q, z\}$. We thus set $\operatorname{pa}_a = \{q, z\}$. (Fourth graph in Figure 2.)



Figure 2: Exercise 1, Question (b):Construction of a minimal directed I-map for the ordering (e, h, q, z, a).

Since the skeletons of the two minimal I-maps obtained with the different orderings do not match, we don't have I-equivalence. Note that the ordering (e, h, q, z, a) yields a denser graph (Figure 2) than the graph in Figure 1. While a minimal I-map, the graph does e.g. not show that $a \perp z$. Furthermore, the causal interpretation of the two graphs is different.

(c) For the collection of random variables (a, z, h, q, e) you are given the following Markov blankets for each variable:

- $MB(a) = \{q,z\}$
- $MB(z) = \{a,q,h\}$
- $MB(h) = \{z\}$
- $MB(q) = \{a, z, e\}$
- $MB(e) = \{q\}$
- (i) Draw the undirected minimal I-map.
- (ii) Indicate a Gibbs distribution that satisfies the independence relations specified by the Markov blankets.

Solution. Connecting each variable to all variables in its Markov blanket yields the desired undirected minimal I-map (see lecture slides). Note that the Markov blankets are not mutually disjoint.



After MB(a) After MB(z) After MB(q)

For positive distributions, the set of distributions that satisfy the local Markov property relative to a graph (as given by the Markov blankets) is the same as the set of Gibbs distributions that factorise according to the graph. Given the I-map, we can now easily find the Gibbs distribution

$$p(a, z, h, q, e) = \phi_1(a, z, q)\phi_2(q, e)\phi_3(z, h)$$

Note that we used the maximal clique (a, z, q).

Exercise 2. Conversion between graphs

(a) For distributions that factorises over the graph below, find the minimal undirected I-map.



Solution. To derive an undirected minimal I-map from a directed one, we have to construct the moralised graph where the "unmarried" parents are connected by a covering edge. This is because each conditional $p(x_i|pa_i)$ corresponds to a factor $\phi_i(x_i, pa_i)$ and we need to connect all variables that are arguments of the same factor with edges.

Statistically, the reason for marrying the parents is as follows: An independency $x \perp y$ {child, other nodes} does not hold in the directed graph in case of collider connections but would hold in the undirected graph if we didn't marry the parents. Hence links between the parents must be added.

It is important to add edges between *all* parents of a node. Here, $p(x_4|x_1, x_2, x_3)$ corresponds to a factor $\phi(x_4, x_1, x_2, x_3)$ so that all four variables need to be connected. Just adding edges $x_1 - x_2$ and $x_2 - x_3$ is not enough.

The moral graph, which is the requested minimal undirected I-map, is shown below.



(b) The graph below is a directed minimal I-map for the hidden Markov model. Find the corresponding undirected minimal I-map.



Solution. The graph does not contain any head-head connections. The undirected minimal I-map is thus obtained by removing all arrows from the graph.



(c) For the undirected I-map below, what is a corresponding directed minimal I-map?



Solution. We use the ordering x_1, x_2, x_3, x_4, x_5 and follow the general procedure to construct the directed minimal I-map while reading the independencies from the undirected graph:

- x_2 is not independent from x_1 so that we set $pa_2 = \{x_1\}$. See first graph in Figure 3.
- Since x_3 is connected to both x_1 and x_2 , we generally don't have $x_3 \perp x_2, x_1$. We cannot make x_3 independent from x_2 by conditioning on x_1 because there are two paths from x_3 to x_2 and x_1 only blocks the upper one. Moreover, x_1 is a neighbour of x_3 so that conditioning on x_2 does make them independent. Hence we must set $pa_3 = \{x_1, x_2\}$. See second graph in Figure 3.
- For x_4 , we see from the undirected graph, that $x_4 \perp x_1 \mid x_3, x_2$. The graph further shows that removing either x_3 or x_2 from the conditioning set is not possible and conditioning on x_1 won't make x_4 independent from x_2 or x_3 . We thus have $pa_4 =$ $\{x_2, x_3\}$. See fourth graph in Figure 3.
- The same reasoning shows that $pa_5 = \{x_3, x_4\}$. See last graph in Figure 3.

This results in the triangulated directed graph in Figure 3 on the right.



Figure 3: Answer to Exercise 2, Question (c).

To see why triangulation is necessary consider the case where we didn't have the edge between x_2 and x_3 as in Figure 4. The directed graph would then imply that $x_3 \perp x_2 \mid x_1 \mid x_2 \mid x_3 \mid x$ (check!). But this independency assertion does not hold in the undirected graph so that the graph in Figure 4 is not an I-map.



Not a directed I-map for the undirected graphical model defined by the graph in Figure 4: Question(c) of Exercise 2.

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 $p(y_1, y_2 | x_1, x_2) = p(y_1, y_2 | x_1, x_2) p(x_1) p(x_2)$ factorises as ŗ

$$p(y_1, y_2|x_1, x_2) = p(y_1|x_1)p(y_2|x_2)\phi(y_1, y_2)n(x_1, x_2)$$
(1)

with $n(x_1, x_2)$ equal to

$$n(x_1, x_2) = \left(\int p(y_1|x_1)p(y_2|x_2)\phi(y_1, y_2)\mathrm{d}y_1\mathrm{d}y_2\right)^{-1}.$$
(2)

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).



In the additional practice questions, you are asked to verify that $p(y_1, y_2, x_1, x_2)$ satisfies the following independencies

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

We here investigate whether undirected and directed graph can represent these independencies.

(a) Draw the undirected graph for $p(y_1, y_2, x_1, x_2)$ and check whether graph separation allows us to see all independencies listed above.

Solution. We write

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

as

$$p(y_1, y_2, x_1, x_2) = \phi_1(y_1, x_1)\phi_2(y_2, x_2)\phi_3(y_1, y_2)\phi_4(x_1, x_2) \quad \text{with} \quad (S.1)$$

$$\phi_1(y_1, x_1) = p(y_1 | x_1) p(x_1) \tag{S.2}$$

$$\phi_2(y_2, x_2) = p(y_2|x_2)p(x_2) \tag{S.3}$$

$$\phi_3(y_1, y_2) = \phi(y_1, y_2) \tag{S.4}$$

$$\phi_4(x_1, x_2) = n(x_1, x_2) \tag{S.5}$$

The corresponding undirected graph is as follows.



While the graph implies $x_1 \perp \!\!\!\perp y_2 \mid y_1, x_2$ and $x_2 \perp \!\!\!\perp y_1 \mid y_2, x_1$, the independency $x_1 \perp \!\!\!\perp x_2$ is not represented.

(b) Draw the directed graph for $p(y_1, y_2, x_1, x_2) = p(y_1, y_2|x_1, x_2)p(x_1)p(x_2)$ and check whether graph separation allows us to see all independencies listed above.

Solution. If we use the ordering x_1, x_2, y_1, y_2 , we obtain the graph on the left. If we use the ordering x_1, x_2, y_2, y_1 , we obtain the graph on the right.



The graphs do represent $x_1 \perp x_2$ but not $x_1 \perp y_2 \mid y_1, x_2$ and $x_2 \perp y_1 \mid y_2, x_1$. Moreover, the graphs imply a directionality between y_1 and y_2 , and a direct influence of x_1 on y_2 , and of x_2 on y_1 , in contrast to the original modelling goals.