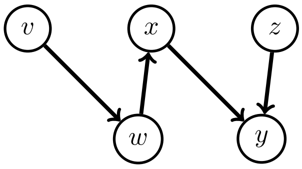
Exercises for the tutorials: 1 and 2.

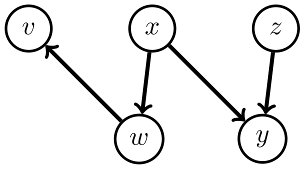
The other exercises are for self-study and exam preparation. All material is examinable unless otherwise mentioned.

Exercise 1. *I-equivalence*

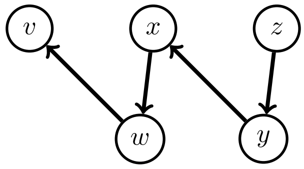
(a) *Which of three graphs represent the same set of independencies? Explain.*



Graph 1

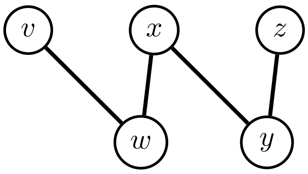


Graph 2

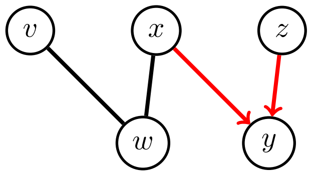


Graph 3

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.

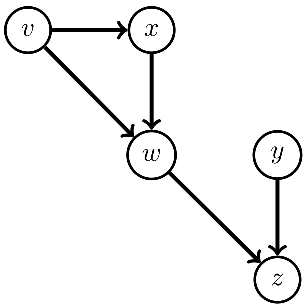


skeleton

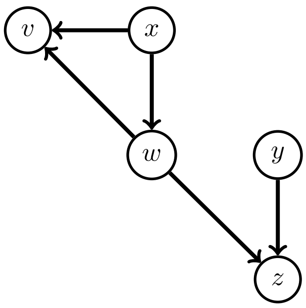


immorality

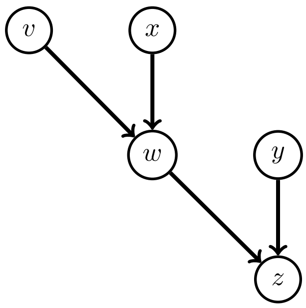
(b) *Which of three graphs represent the same set of independencies? Explain.*



Graph 1

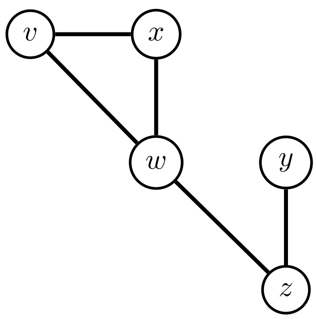


Graph 2

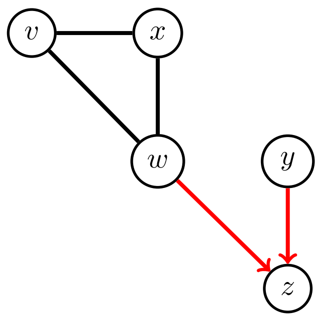


Graph 3

Solution. The skeleton of graph 3 is different from the skeleton of graphs 1 and 2, so that graph 3 cannot be I-equivalent to graph 1 or 2, and we do not need to further check the immoralities for graph 3. Graph 1 and 2 have the same skeleton, and they also have the same immorality. Hence, graph 1 and 2 are I-equivalent. Note that node w in graph 1 is in a collider configuration along trail $v - w - x$ but it is not an immorality because its parents are connected (covering edge); equivalently for node v in graph 2.

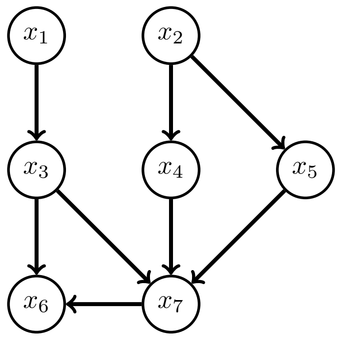


skeleton



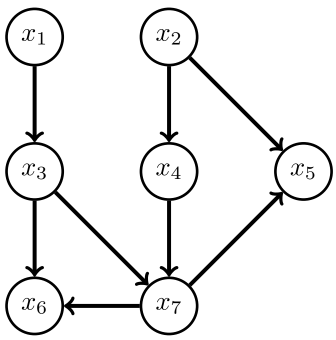
immorality

(c) *Assume the graph below is a perfect map for a set of independencies $\mathcal{U}$.*

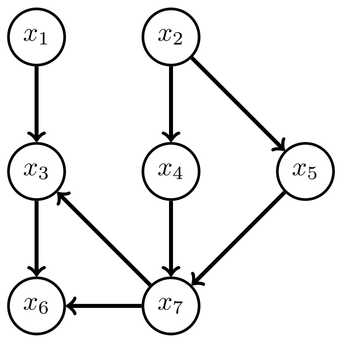


Graph 0

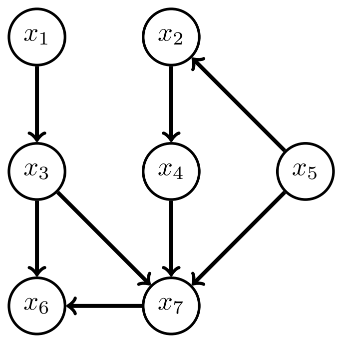
For each of the three graphs below, explain whether the graph is a perfect map, an I-map, or not an I-map for $\mathcal{U}$.



Graph 1



Graph 2



Graph 3

Solution.

- Graph 1 has an immorality $x_2 \to x_5 \leftarrow x_7$ which graph 0 does not have. The graph is thus not I-equivalent to graph 0 and can thus not be a perfect map. Moreover, graph 1 asserts that $x_2 \perp\!\!\!\perp x_7 | x_4$ which is not case for graph 0. Since graph 0 is a perfect map for $\mathcal{U}$, graph 1 asserts an independency that does not hold for $\mathcal{U}$ and can thus not be an I-map for $\mathcal{U}$.
- Graph 2 has an immorality $x_1 \to x_3 \leftarrow x_7$ which graph 0 does not have. Graph 2 thus asserts that $x_1 \perp\!\!\!\perp x_7$, which is not the case for graph 0. Hence, for the same reason as for graph 1, graph 2 is not an I-map for $\mathcal{U}$.

- Graph 3 has the same skeleton and set of immoralities as graph 0. It is thus I-equivalent to graph 0, and hence also a perfect map.

Exercise 2. *Minimal I-maps*

(a) *Assume that the graph G in Figure 1 is a perfect I-map for $p(a, z, q, e, h)$. Determine the minimal directed I-map using the ordering (e, h, q, z, a). Is the obtained graph I-equivalent to G?*

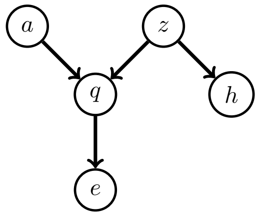


Figure 1: Perfect I-map G for Exercise 2, question (a).

Solution. To find a minimal I-map, we can use the procedure that we used to simplify the chain rule and visualise the obtained factorisation as a DAG. Since we are given a perfect I-map G for p, we can use the graph to check whether p satisfies a certain independency. This gives the following recipe:

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

$$x_i \perp\!\!\!\perp \{\text{pre}_i \setminus \pi_i\} \mid \pi_i$$

   is in $\mathcal{I}(G)$ (only works if G is a perfect I-map for $\mathcal{I}(p)$)
3. Construct a graph with parents $\text{pa}_i = \pi_i$.

Note: For I-maps G that are not perfect, if the graph does not indicate that a certain independency holds, we have to check that the independency indeed does not hold for p. If we don't, we won't obtain a minimal I-map but just an I-map for $\mathcal{I}(p)$. This is because p may have independencies that are not encoded in the graph G.

Given 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 $h \perp\!\!\!\perp e$ does not hold. We thus set e as parent of h, see first graph in Figure 2. Then:

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

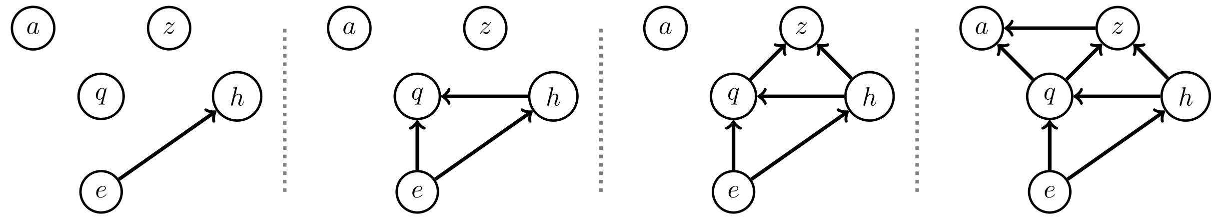


Figure 2: Exercise 2, Question (a):Construction of a minimal directed I-map for the ordering (e, h, q, z, a).

Since the skeleton in the obtained minimal I-map is different from the skeleton of G, we do not have I-equivalence. Note that the ordering (e, h, q, z, a) yields a denser graph (Figure 2) than the graph in Figure 1. Whilst a minimal I-map, the graph does e.g. not show that $a \perp\!\!\!\perp z$. Furthermore, the causal interpretation of the two graphs is different.

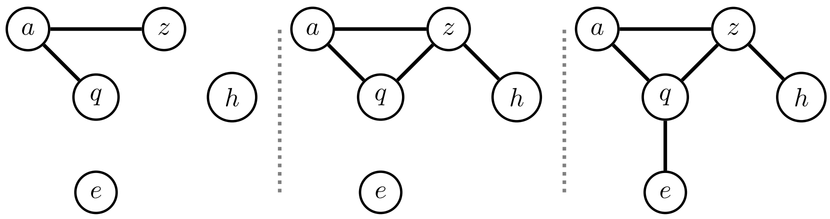
(b) For the collection of random variables (a, z, h, q, e) you are given the following Markov blankets for each variable:

- *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 representing the independencies.

(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)

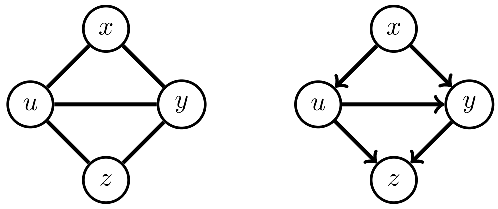
For positive distributions, the set of distributions that satisfy the local Markov property relative to a graph (as given by the Markov blankets) is the same as the set of Gibbs distributions that factorise according to the graph. Given the I-map, we can now easily find the Gibbs distribution

$$p(a, z, h, q, e) = \frac{1}{Z}\phi_1(a, z, q)\phi_2(q, e)\phi_3(z, h),$$

where the ϕ_i must take positive values on their domain. Note that we used the maximal clique (a, z, q).

Exercise 3. *I-equivalence between directed and undirected graphs*

(a) *Verify that the following two graphs are I-equivalent by listing and comparing the independencies that each graph implies.*



Solution. First, note that both graphs share the same skeleton and the only reason that they are not fully connected is the missing edge between x and z.

For the DAG, there is also only one ordering that is topological to the graph: x, u, y, z. The missing edge between x and y corresponds to the only independency encoded by the graph: $z \perp\!\!\!\perp \text{pre}_z \setminus \text{pa}_z | \text{pa}_z$, i.e.

$$z \perp\!\!\!\perp x | u, y.$$

This is the same independency that we get from the directed local Markov property.

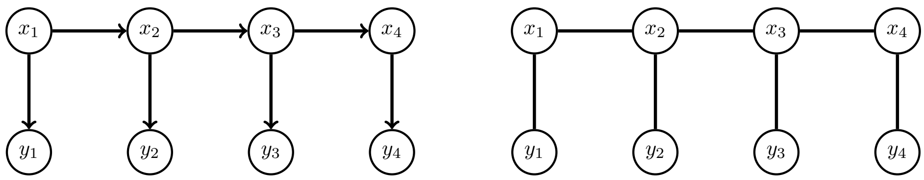
For the undirected graph,

$$z \perp\!\!\!\perp x | u, y$$

holds because u, y block all paths between z and x. All variables but z and x are connected to each other, so that no further independency can hold.

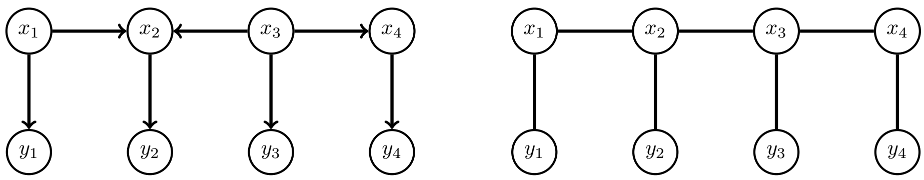
Hence both graphs only encode $z \perp\!\!\!\perp x | u, y$ and they are thus I-equivalent.

(b) *Are the following two graphs, which are directed and undirected hidden Markov models, I-equivalent?*



Solution. The skeleton of the two graphs is the same and there are no immoralities. Hence, the two graphs are I-equivalent.

(c) *Are the following two graphs I-equivalent?*



Solution. The two graphs are not I-equivalent because $x_1 - x_2 - x_3$ forms an immorality. Hence, the undirected graph encodes $x_1 \perp\!\!\!\perp x_3 | x_2$ which is not represented in the directed graph. On the other hand, the directed graph asserts $x_1 \perp\!\!\!\perp x_3$ which is not represented in the undirected graph.

Exercise 4. ***Moralisation: Converting DAGs to undirected minimal I-maps***

In the lecture, we had the following recipe to construct undirected minimal I-maps for $\mathcal{I}(p)$:

- *Determine the Markov blanket for each variable x_i*
- *Construct a graph where the neighbours of x_i are given by its Markov blanket.*

We can adapt the recipe to construct an undirected minimal I-map for the independencies $\mathcal{I}(G)$ encoded by a DAG G. What we need to do is to use G to read out the Markov blankets for the variables x_i rather than determining the Markov blankets from the distribution p.

Show that this procedure leads to the following recipe to convert DAGs to undirected minimal I-maps:

1. *For all immoralities in the graph: add edges between* all *parents of the collider node.*
2. *Make all edges in the graph undirected.*

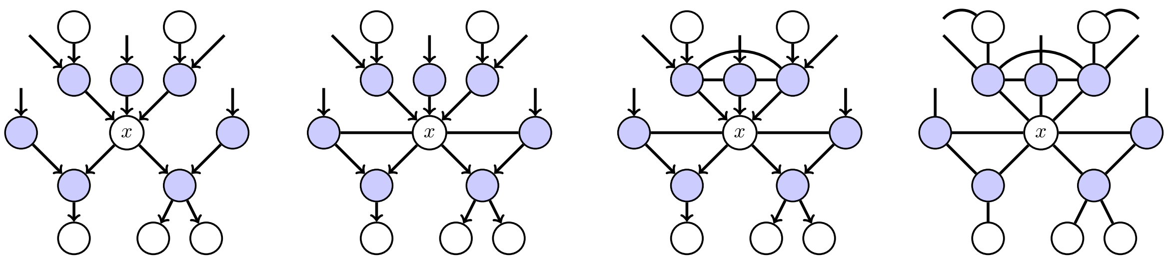
The first step is sometimes called "moralisation" because we "marry" all the parents in the graph that are not already directly connected by an edge. The resulting undirected graph is called the moral graph of G, sometimes denoted by $\mathcal{M}(G)$.

Solution. The Markov blanket of a variable x is the set of its parents, children, and co-parents, as shown in the graph below in sub-figure (a). The parents and children are connected to x in the directed graph, but the co-parents are not directly connected to x. Hence, according to "Construct a graph where the neighbours of x_i are its Markov blanket.", we need to introduce edges between x and all its co-parents. This gives the intermediate graph in sub-figure (b).

Now, considering the top-left parent of x, we see that for that node, the Markov blanket includes the other parents of x. This means that we need to connect all parents of x, which gives the graph in sub-figure (c). This is sometimes called "marrying" the parents of x. Continuing in this way, we see that we need to "marry" all parents in the graph that are not already married.

Finally, we need to make all edges in the graph undirected, which gives sub-figure (d).

A simpler approach is to note that the DAG specifies the factorisation $p(\mathbf{x}) = \prod_i p(x_i|\text{pa}_i)$. We can consider each conditional $p(x_i|\text{pa}_i)$ to be a factor $\phi_i(x_i, \text{pa}_i)$ so that we obtain the Gibbs distribution $p(\mathbf{x}) = \prod_i \phi_i(x_i|\text{pa}_i)$. Visualising the distribution by connecting all variables in the same factor $\phi_i(x_i|\text{pa}_i)$ leads to the "marriage" of all parents of x_i. This corresponds to the first step in the recipe because x_i is in a collider configuration with respect to the parent nodes. Not all parents form an immorality but this does here not matter because those that do not form an immorality are already connected by a covering edge in the first place.

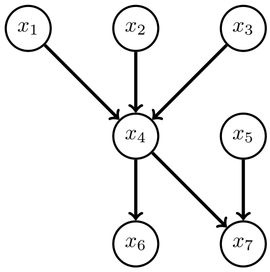


(a) DAG (b) Intermediate step 1 (c) Intermediate step 2 (d) Undirected graph

Figure 3: Answer to Exercise 4: Illustrating the moralisation process

Exercise 5. ***Moralisation exercise***

For the DAG G below find the minimal undirected I-map for $\mathcal{I}(G)$.

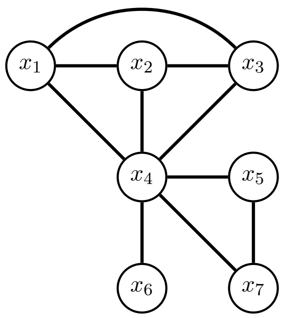


Solution. To derive an undirected minimal I-map from a directed one, we have to construct the moralised graph where the "unmarried" parents are connected by a covering edge. This is because each conditional $p(x_i|\text{pa}_i)$ corresponds to a factor $\phi_i(x_i, \text{pa}_i)$ and we need to connect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

Statistically, the reason for marrying the parents is as follows: An independency $x \perp\!\!\!\perp y|\{\text{child, other nodes}\}$ 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.

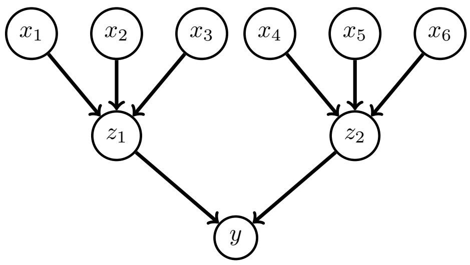
It is important to add edges between *all* parents of a node. Here, $p(x_4|x_1, x_2, x_3)$ corresponds to a factor $\phi(x_4, x_1, x_2, x_3)$ so that all four variables need to be connected. Just adding edges $x_1 - x_2$ and $x_2 - x_3$ would not be enough.

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

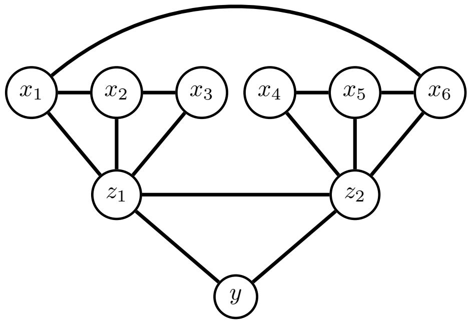


Exercise 6. ***Moralisation exercise***

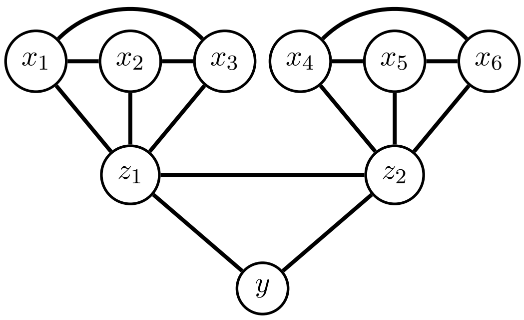
Consider the DAG G:



A friend claims that the undirected graph below is the moral graph $\mathcal{M}(G)$ of G. Is your friend correct? If not, state which edges needed to be removed or added, and explain, in terms of represented independencies, why the changes are necessary for the graph to become the moral graph of G.



Solution. The moral graph $\mathcal{M}(G)$ is an undirected minimal I-map of the independencies represented by G. Following the procedure of connecting "unmarried" parents of colliders, we obtain the following moral graph of G:



We can thus see that the friend's undirected graph is not the moral graph of G.

The edge between x_1 and x_6 can be removed. This is because for G, we have e.g. the independencies $x_1 \perp\!\!\!\perp x_6|z_1$, $x_1 \perp\!\!\!\perp x_6|z_2$, $x_1 \perp\!\!\!\perp x_6|z_1, z_2$ which is not represented by the drawn undirected graph.

We need to add edges between x_1 and x_3, and between x_4 and x_6. Otherwise, the undirected graph makes the wrong independency assertion that $x_1 \perp\!\!\!\perp x_3|x_2, z_1$ (and equivalent for x_4 and x_6).

Exercise 7. *Triangulation: Converting undirected graphs to directed minimal I-maps*

In Exercise 4 we adapted a recipe for constructing undirected minimal I-maps for $\mathcal{I}(p)$ to the case of $\mathcal{I}(G)$, where G is a DAG. The key difference was that we used the graph G to determine independencies rather than the distribution p.

We can similarly adapt the recipe for constructing a directed minimal I-map for $\mathcal{I}(p)$ to build a directed minimal I-map for $\mathcal{I}(H)$, where H is an undirected graph:

1. *Choose an ordering of the random variables.*

2. *For all variables x_i, use H to determine a* minimal *subset π_i of the predecessors* pre_i *such that*

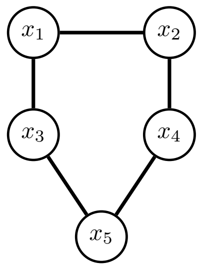
$$x_i \perp\!\!\!\perp (\mathrm{pre}_i \setminus \pi_i) \mid \pi_i$$

holds.

3. *Construct a DAG with the π_i as parents pa_i of x_i.*

Remarks: (1) Directed minimal I-maps obtained with different orderings are generally not I-equivalent. (2) The directed minimal I-maps obtained with the above method are always chordal graphs. Chordal graphs are graphs where the longest trail without shortcuts is a triangle (https://en.wikipedia.org/wiki/Chordal_graph). They are thus also called triangulated graphs. We obtain chordal graphs because if we had trails without shortcuts that involved more than 3 nodes, we would necessarily have an immorality in the graph. But immoralities encode independencies that an undirected graph cannot represent, which would make the DAG not an I-map for $\mathcal{I}(H)$ any more.

(a) *Let H be the undirected graph below. Determine the directed minimal I-map for $\mathcal{I}(H)$ with the variable ordering x_1, x_2, x_3, x_4, x_5.*



Solution. We use the ordering x_1, x_2, x_3, x_4, x_5 and follow the conversion procedure:

- x_2 is not independent from x_1 so that we set $\mathrm{pa}_2 = \{x_1\}$. See first graph in Figure 4.
- Since x_3 is connected to both x_1 and x_2, we don't have $x_3 \perp\!\!\!\perp x_2, x_1$. 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 $\mathrm{pa}_3 = \{x_1, x_2\}$. See second graph in Figure 4.
- For x_4, we see from the undirected graph, that $x_4 \perp\!\!\!\perp x_1 \mid x_3, x_2$. 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 $\mathrm{pa}_4 = \{x_2, x_3\}$. See fourth graph in Figure 4.
- The same reasoning shows that $\mathrm{pa}_5 = \{x_3, x_4\}$. See last graph in Figure 4.

This results in the triangulated directed graph in Figure 4 on the right.

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 5. The directed graph would then imply that $x_3 \perp\!\!\!\perp x_2 \mid x_1$ (check!). But this independency assertion does not hold in the undirected graph so that the graph in Figure 5 is not an I-map.

(b) *For the undirected graph from question (a) above, which variable ordering yields the directed minimal I-map below?*

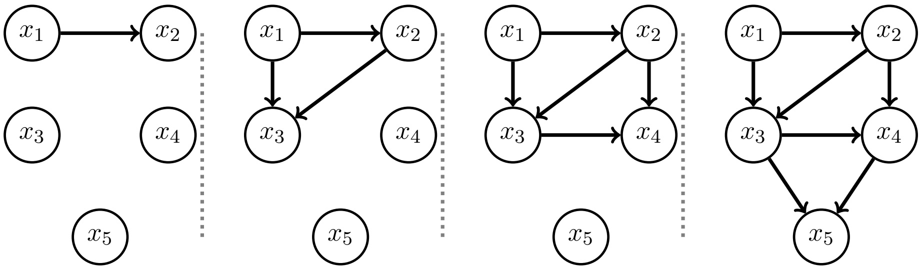


Figure 4: . Answer to Exercise 7, Question (a).

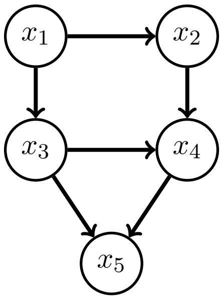
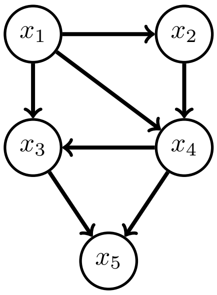


Figure 5: Not a directed I-map for the undirected graphical model defined by the graph in Exercise 7, Question (a).



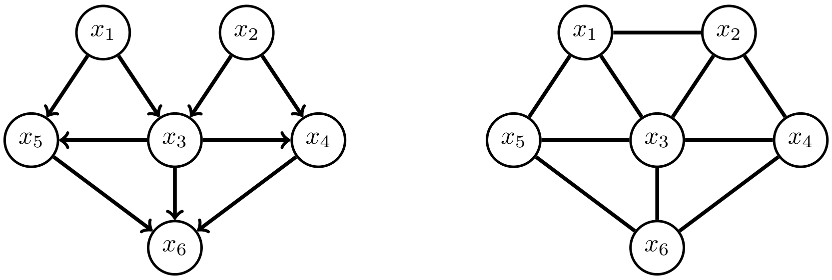
Solution. x_1 is the root of the DAG, so it comes first. Next in the ordering are the children of x_1: x_2, x_3, x_4. Since x_3 is a child of x_4, and x_4 a child of x_2, we must have x_1, x_2, x_4, x_3. Furthermore, x_3 must come before x_5 in the ordering since x_5 is a child of x_3, hence the ordering used must have been: x_1, x_2, x_4, x_3, x_5.

Exercise 8. *I-maps, minimal I-maps, and I-equivalency*

Consider the following probability density function for random variables $x_1, \ldots, x_6$.

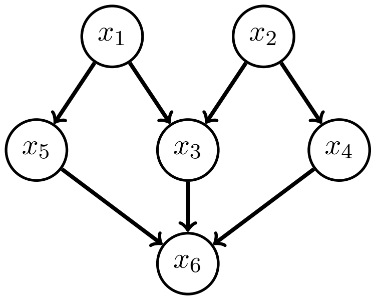
$$p_a(x_1, \ldots, x_6) = p(x_1)p(x_2)p(x_3|x_1, x_2)p(x_4|x_2)p(x_5|x_1)p(x_6|x_3, x_4, x_5)$$

For each of the two graphs below, explain whether it is a minimal I-map, not a minimal I-map but still an I-map, or not an I-map for the independencies that hold for p_a.



graph 1 *graph 2*

Solution. The pdf can be visualised as the following directed graph, which is a minimal I-map for it.



Graph 1 defines distributions that factorise as

$$p_b(\mathbf{x}) = p(x_1)p(x_2)p(x_3|x_1, x_2)p(x_4|x_2, x_3)p(x_5|x_1, x_3)p(x_6|x_3, x_4, x_5). \tag{S.1}$$

Comparing with $p_a(x_1, \ldots, x_6)$, we see that only the conditionals $p(x_4|x_2, x_3)$ and $p(x_5|x_1, x_3)$ are different. Specifically, their conditioning set includes x_3, which means that Graph 1 encodes fewer independencies than what $p_a(x_1, \ldots, x_6)$ satisfies. In particular $x_4 \perp\!\!\!\perp x_3|x_2$ and $x_5 \perp\!\!\!\perp x_3|x_1$ are not represented in the graph. This means that we could remove x_3 from the conditioning sets, or equivalently remove the edges $x_3 \to x_4$ and $x_3 \to x_5$ from the graph without introducing independence assertions that do not hold for p_a. This means graph 1 is an I-map but not a minimal I-map.

Graph 2 is not an I-map. To be an undirected minimal I-map, we had to connect variables x_5 and x_4 that are parents of x_6. Graph 2 wrongly claims that $x_5 \perp\!\!\!\perp x_4 \mid x_1, x_3, x_6$.

Exercise 9. *Limits of directed and undirected graphical models*

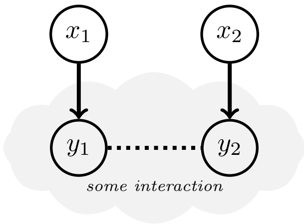
We here consider the probabilistic model $p(y_1, y_2, x_1, x_2) = p(y_1, y_2|x_1, x_2)p(x_1)p(x_2)$ *where* $p(y_1, y_2|x_1, x_2)$ *factorises as*

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

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

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

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).*



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

$$u \perp\!\!\!\perp v|z \Longleftrightarrow p(u,v|z) = p(u|z)p(v|z) \tag{3}$$
$$u \perp\!\!\!\perp v|z \Longleftrightarrow p(u,v|z) = a(u,z)b(v,z) \qquad (a(u,z) \geq 0, b(v,z) \geq 0) \tag{4}$$

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

$$\begin{aligned}
p(x_1, x_2) &= \int p(y_1, y_2, x_1, x_2)\mathrm{d}y_1\mathrm{d}y_2 && \text{(S.2)}\\
&= \int p(y_1|x_1)p(y_2|x_2)\phi(y_1, y_2)n(x_1, x_2)p(x_1)p(x_2)\mathrm{d}y_1\mathrm{d}y_2 && \text{(S.3)}\\
&= n(x_1, x_2)p(x_1)p(x_2) \int p(y_1|x_1)p(y_2|x_2)\phi(y_1, y_2)\mathrm{d}y_1\mathrm{d}y_2 && \text{(S.4)}\\
&\overset{(2)}{=} n(x_1, x_2)p(x_1)p(x_2)\frac{1}{n(x_1, x_2)} && \text{(S.5)}\\
&= p(x_1)p(x_2). && \text{(S.6)}
\end{aligned}$$

Since $p(x_1)$ and $p(x_2)$ are the univariate marginals of x_1 and x_2, respectively, it follows from (3) 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

$$\begin{aligned}
p(y_1, y_2, x_1, x_2) &= p(y_1|x_1)p(y_2|x_2)\phi(y_1, y_2)n(x_1, x_2)p(x_1)p(x_2) && \text{(S.7)}\\
&= [p(y_1|x_1)p(x_1)n(x_1, x_2)]\,[p(y_2|x_2)\phi(y_1, y_2)p(x_2)] && \text{(S.8)}\\
&= \phi_A(x_1, y_1, x_2)\phi_B(y_2, y_1, x_2) && \text{(S.9)}
\end{aligned}$$

With (4), 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

$$\begin{aligned}
p(y_1, y_2, x_1, x_2) &= p(y_1|x_1)p(y_2|x_2)\phi(y_1, y_2)n(x_1, x_2)p(x_1)p(x_2) && \text{(S.10)}\\
&= [p(y_2|x_2)n(x_1, x_2)p(x_2)p(x_1))]\,[p(y_1|x_1)\phi(y_1, y_2)]) && \text{(S.11)}\\
&= \tilde{\phi}_A(x_2, x_1, y_2)\tilde{\phi}_B(y_1, y_2, x_1) && \text{(S.12)}
\end{aligned}$$

With (4), we have that $x_2 \perp\!\!\!\perp y_1 \mid y_2, x_1$.

(b) *Is there an undirected perfect map for the independencies satisfied by $p(y_1, y_2, x_1, x_2)$?*

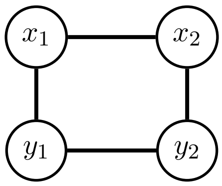
Solution. We write

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

as a Gibbs distribution

$$p(y_1, y_2, x_1, x_2) = \phi_1(y_1, x_1)\phi_2(y_2, x_2)\phi_3(y_1, y_2)\phi_4(x_1, x_2) \qquad \text{with} \tag{S.13}$$
$$\phi_1(y_1, x_1) = p(y_1|x_1)p(x_1) \tag{S.14}$$
$$\phi_2(y_2, x_2) = p(y_2|x_2)p(x_2) \tag{S.15}$$
$$\phi_3(y_1, y_2) = \phi(y_1, y_2) \tag{S.16}$$
$$\phi_4(x_1, x_2) = n(x_1, x_2). \tag{S.17}$$

Visualising it as an undirected graph gives an I-map:



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. Hence the graph is not a perfect map. Note further that removing any edge would result in a graph that is not an I-map for $\mathcal{I}(p)$ anymore. Hence the graph is a minimal I-map for $\mathcal{I}(p)$ but that we cannot obtain a perfect I-map.

(c) *Is there a directed perfect map for the independencies satisfied by $p(y_1, y_2, x_1, x_2)$?*

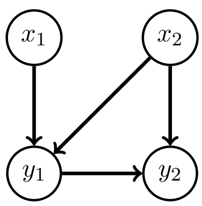
Solution. We construct directed minimal I-maps for $p(y_1, y_2, x_1, x_2) = p(y_1, y_2|x_1, x_2)p(x_1)p(x_2)$ for different orderings as explained in the lecture. We will see that they do not represent all independencies in $\mathcal{I}(p)$ and hence that they are not perfect I-maps.

To guarantee unconditional independence of x_1 and x_2, the two variables must come first in the orderings (either x_1 and then x_2 or the other way around).

If we use the ordering x_1, x_2, y_1, y_2, and that

- $x_1 \perp\!\!\!\perp x_2$
- $y_2 \perp\!\!\!\perp x_1 | y_1, x_2$, which is $y_2 \perp\!\!\!\perp \text{pre}(y_2) \setminus \pi | \pi$ for $\pi = (y_1, x_2)$

are in $\mathcal{I}(p)$, we obtain the following directed minimal I-map:

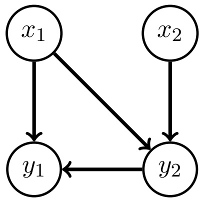


The graphs misses $x_2 \perp\!\!\!\perp y_1 \mid y_2, x_1$.

If we use the ordering x_1, x_2, y_2, y_1, and that

- $x_1 \perp\!\!\!\perp x_2$
- $y_1 \perp\!\!\!\perp x_2 | x_1, y_2$, which is $y_1 \perp\!\!\!\perp \mathrm{pre}(y_1) \setminus \pi | \pi$ for $\pi = (x_1, y_2)$

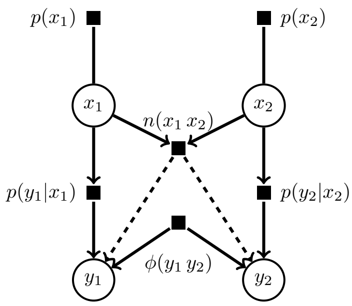
are in $\mathcal{I}(p)$, we obtain the following directed minimal I-map:



The graph misses $x_1 \perp\!\!\!\perp y_2 \mid y_1, x_2$.

Moreover, the graphs imply a directionality between y_1 and y_2, or a direct influence of x_1 on y_2, or of x_2 on y_1, in contrast to the original modelling goals.

(d) (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:

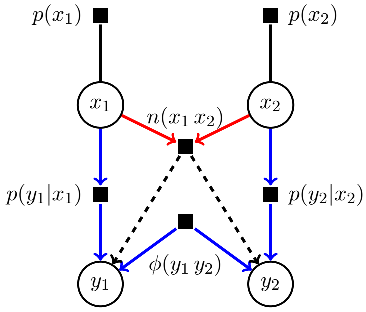
1. *One of the variables in the path is in the conditioning set.*
2. *One of the variables or factors in the path has two incoming edges that are part of the path (variable or factor collider), and neither the variable or factor nor any of its descendants are in the conditioning set."*

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

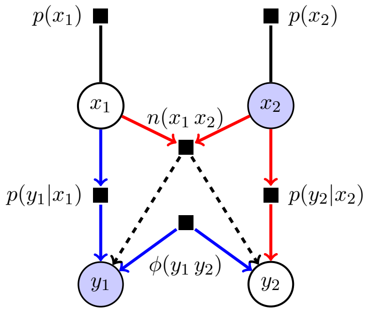
There are two paths from x_1 to x_2 marked with red and blue below:



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

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

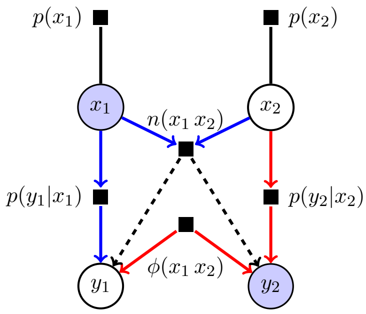
There are two paths from x_1 to y_2 marked with red and blue below:



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).