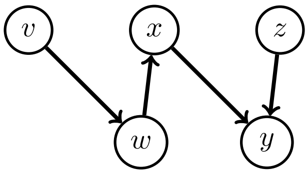
Exercises for the tutorials: 1 and 2.

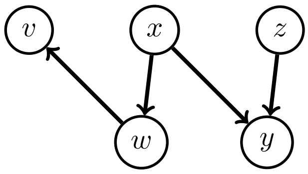
The other exercises are for self-study and exam preparation. All material is examinable unless otherwise mentioned.

Exercise 1. *I-equivalence*

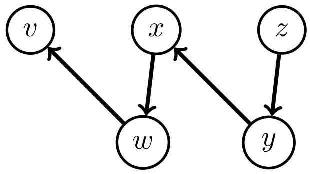
(a) Which of three graphs represent the same set of independencies? Explain.



Graph 1

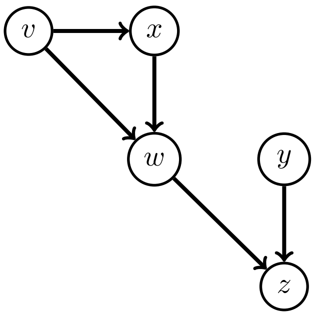


Graph 2

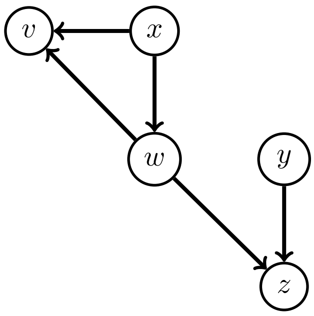


Graph 3

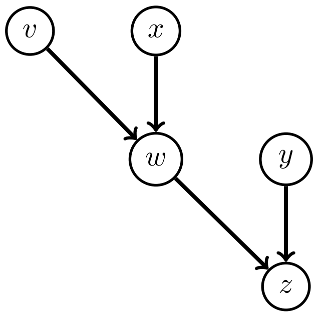
(b) Which of three graphs represent the same set of independencies? Explain.



Graph 1

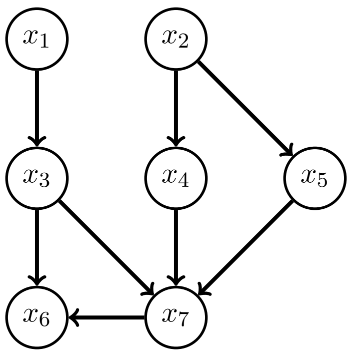


Graph 2



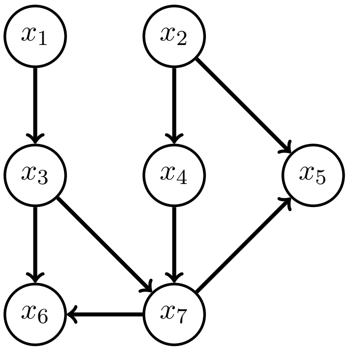
Graph 3

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

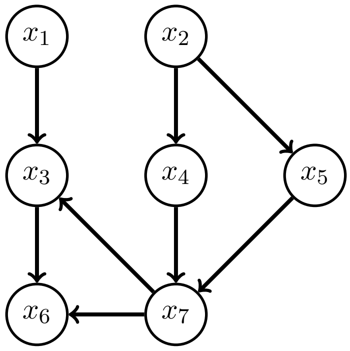


Graph 0

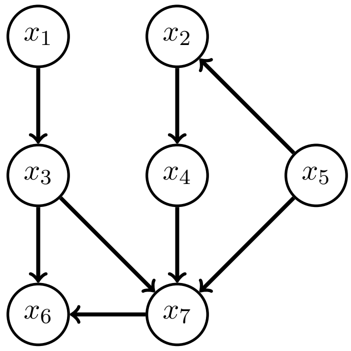
For each of the three graphs below, explain whether the graph is a perfect map, an I-map, or not an I-map for $\mathcal{U}$.



Graph 1



Graph 2



Graph 3

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?

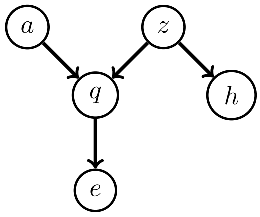


Figure 1: Perfect I-map G for Exercise 2, question (a).

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

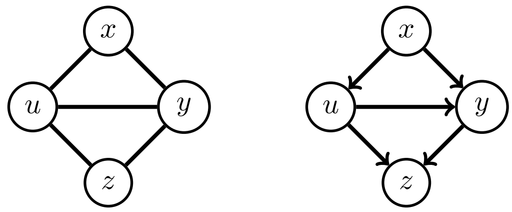
- MB(a) = {q,z}
- MB(z) = {a,q,h}
- MB(h) = {z}
- MB(q) = {a,z,e}
- MB(e) = {q}

(i) Draw the undirected minimal I-map representing the independencies.

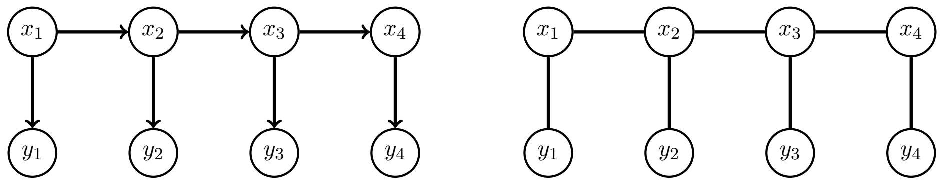
(ii) Indicate a Gibbs distribution that satisfies the independence relations specified by the Markov blankets.

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

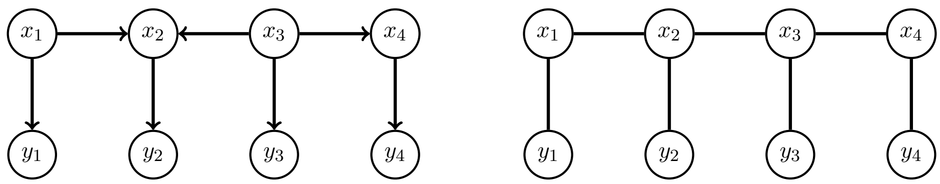
(a) Verify that the following two graphs are I-equivalent by listing and comparing the independencies that each graph implies.



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



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



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.

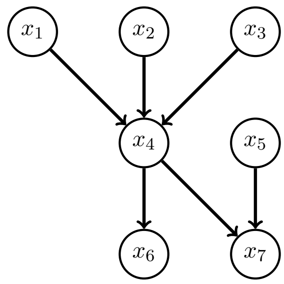
Show that this procedure leads to the following recipe to convert DAGs to undirected minimal I-maps:

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 $\mathcal{M}(G)$.

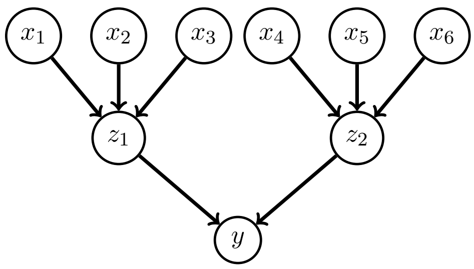
Exercise 5. ***Moralisation exercise***

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

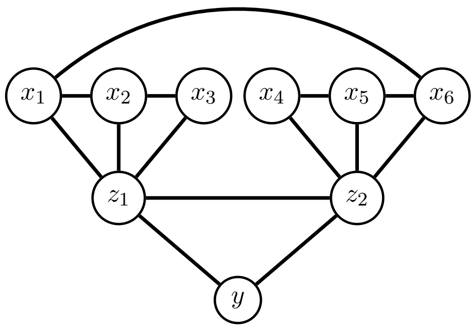


Exercise 6. ***Moralisation exercise***

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.



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:

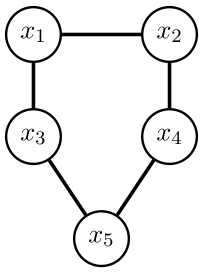
1. Choose an ordering of the random variables.
2. For all variables x_i, use H to determine a *minimal* subset π_i of the predecessors pre_i such that

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

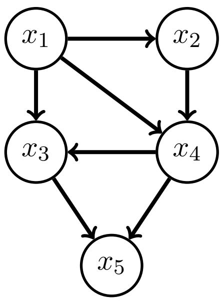
holds.
3. Construct a DAG with the π_i as parents pa_i of x_i.

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.



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

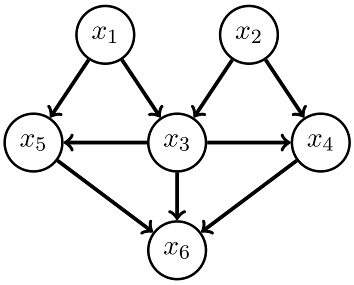


Exercise 8. ***I-maps, minimal I-maps, and I-equivalency***

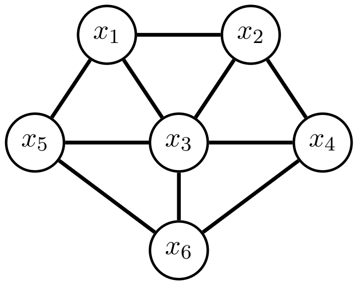
Consider the following probability density function for random variables $x_1, \ldots, x_6$.

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

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

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

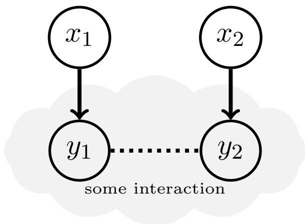
We here consider the probabilistic model $p(y_1, y_2, x_1, x_2) = p(y_1, y_2|x_1, x_2)p(x_1)p(x_2)$ where $p(y_1, y_2|x_1, x_2)$ factorises as

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

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

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

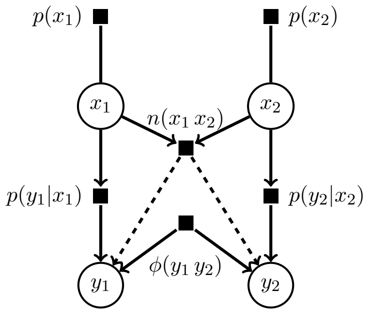
to show that $p(y_1, y_2, x_1, x_2)$ satisfies the following independencies

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

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

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

(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.