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

Exercise 1. *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?*

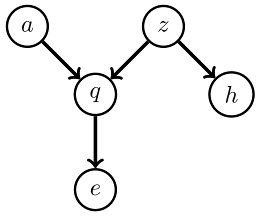


Figure 1: Perfect I-map G for Exercise 1, question (a).

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.)

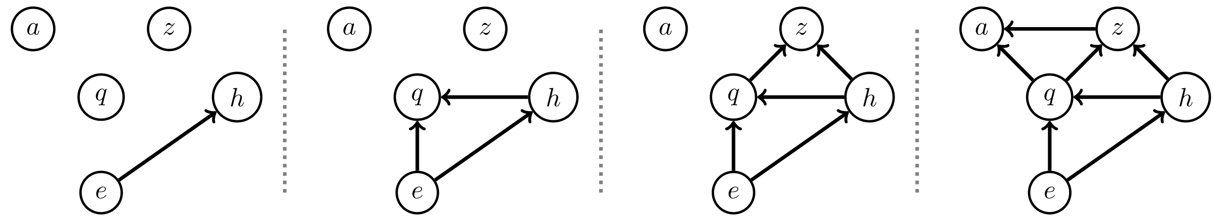


Figure 2: Exercise 1, Question (a):Construction of a minimal directed I-map for the ordering (e, h, q, z, a).

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.

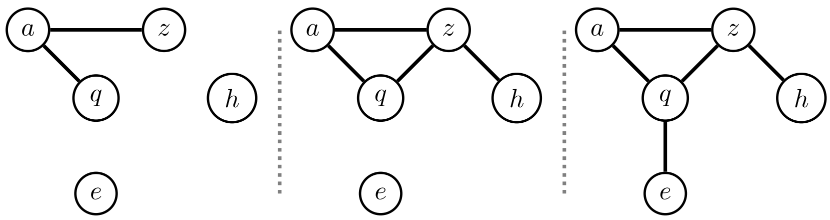
(b) *For the collection of random variables (a, z, h, q, e) you are given the following Markov blankets for each variable:*

- *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)

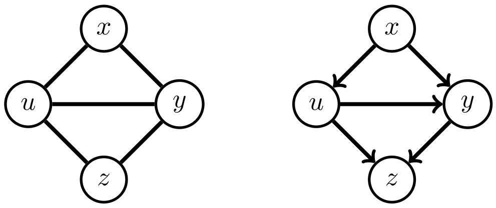
For positive distributions, the set of distributions that satisfy the local Markov property relative to a graph (as given by the Markov blankets) is the same as the set of Gibbs distributions that factorise according to the graph. Given the I-map, we can now easily find the Gibbs distribution

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

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

Exercise 2. ***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.

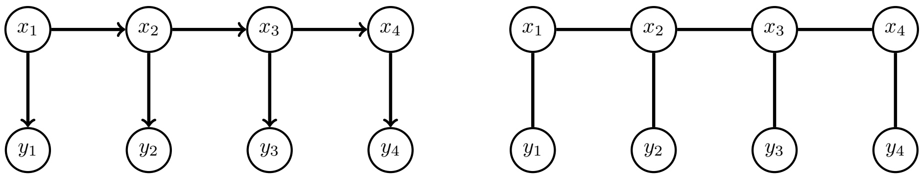
For the undirected graph,

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

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.

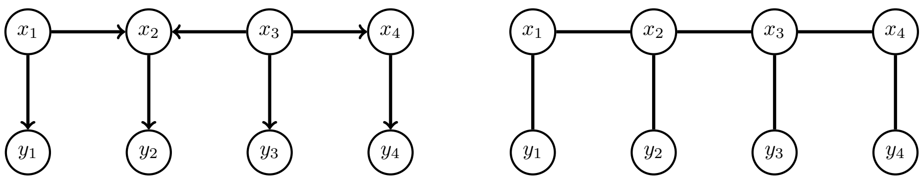
Hence both graphs only encode $z \perp\!\!\!\perp x | u, y$ and they are thus I-equivalent.

(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.

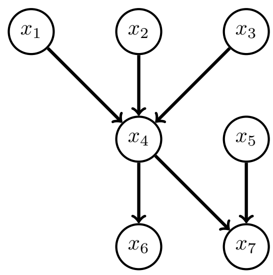
(c) *Are the following two graphs I-equivalent?*



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 3. ***Moralisation exercise***

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

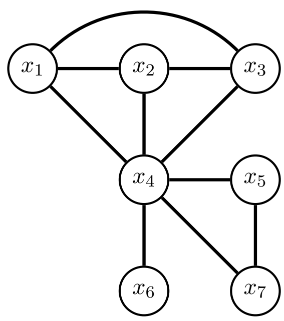


Solution. To derive an undirected minimal I-map from a directed one, we have to construct the moralised graph where the "unmarried" parents are connected by a covering edge. This is because each conditional $p(x_i|\text{pa}_i)$ corresponds to a factor $\phi_i(x_i, \text{pa}_i)$ and we need to connect 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 4. ***Triangulation: Converting undirected graphs to directed minimal I-maps***

In the lecture, we have seen a recipe for constructing directed minimal I-maps for $\mathcal{I}(p)$. We here adapt it to build a directed minimal I-map for $\mathcal{I}(H)$, where H is an undirected graph. The difference to the procedure in the lecture is that we here use the graph H to determine independencies rather than the distribution p.

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*

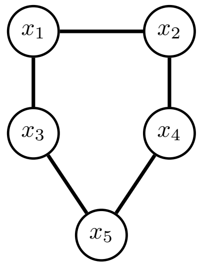
$$x_i \perp\!\!\!\perp (\text{pre}_i \setminus \pi_i) \mid \pi_i$$

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.*



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 $\text{pa}_2 = \{x_1\}$. See first graph in Figure 3.
- 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 $\text{pa}_3 = \{x_1, x_2\}$. See second graph in Figure 3.
- For x_4, we see from the undirected graph, that $x_4 \perp\!\!\!\perp x_1 \mid x_3, x_2$. 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 $\text{pa}_4 = \{x_2, x_3\}$. See fourth graph in Figure 3.
- The same reasoning shows that $\text{pa}_5 = \{x_3, x_4\}$. See last graph in Figure 3.

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

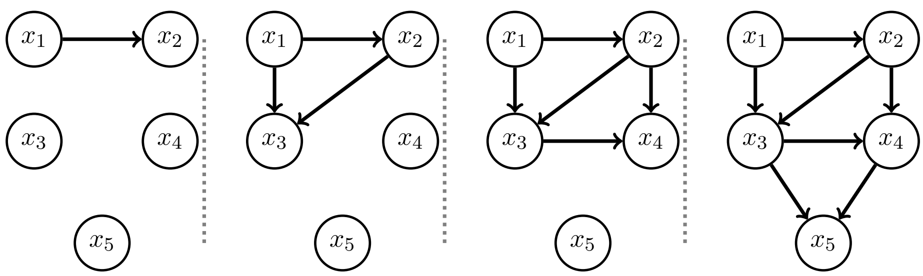


Figure 3: . Answer to Exercise 4, Question (a).

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 $x_3 \perp\!\!\!\perp x_2 \mid x_1$

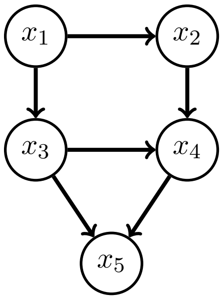
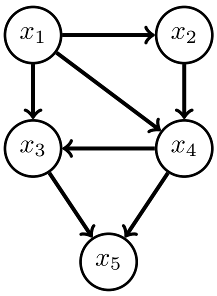


Figure 4: Not a directed I-map for the undirected graphical model defined by the graph in Exercise 4, Question (a).

(check!). But this independency assertion does not hold in the undirected graph so that the graph in Figure 4 is not an I-map.

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



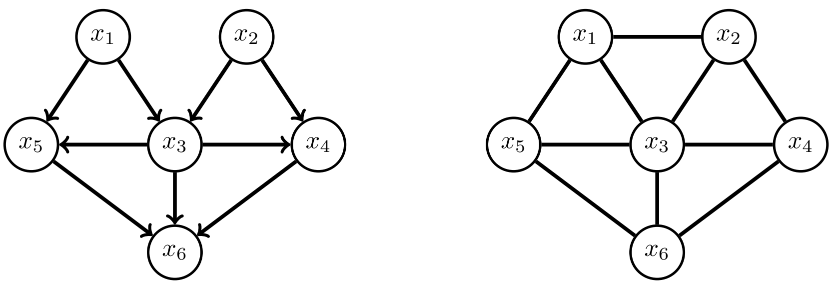
Solution. x_1 is the root of the DAG, so it comes first. Next in the ordering are the children of x_1: x_2, x_3, x_4. Since x_3 is a child of x_4, and x_4 a child of x_2, we must have x_1, x_2, x_4, x_3. Furthermore, x_3 must come before x_5 in the ordering since x_5 is a child of x_3, hence the ordering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

Exercise 5. *I-maps, minimal I-maps, and I-equivalency*

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

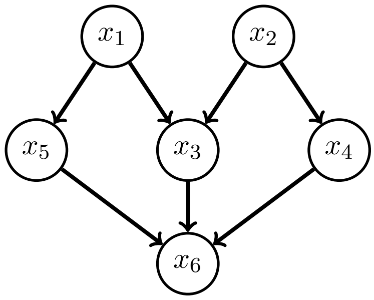
$$p_a(x_1, \ldots, x_6) = p(x_1)p(x_2)p(x_3|x_1, x_2)p(x_4|x_2)p(x_5|x_1)p(x_6|x_3, x_4, x_5)$$

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 $x_5 \perp\!\!\!\perp x_4 \mid x_1, x_3, x_6$.