## Exercise 1. I-maps

(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.

(b) Let $p(a, z, q, e, h)=p(a) p(z) p(q \mid a, z) p(e \mid q) p(h \mid z)$. 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 $\pi_{i} \subseteq \operatorname{pre}_{i}$ such that

$$
x_{i} \Perp\left\{\operatorname{pre}_{i} \backslash \pi_{i}\right\} \mid \pi_{i}
$$

holds for $p$.
3. Construct a graph with parents $\mathrm{pa}_{i}=\pi_{i}$.

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 $\pi_{q}$ of $\operatorname{pre}_{q}$ on which we could condition to make $q$ independent of $\operatorname{pre}_{q} \backslash \pi_{q}$, so that we set the parents of $q$ in the graph to $\mathrm{pa}_{q}=\{e, h\}$. (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 $z \Perp \operatorname{pre}_{z} \backslash \pi_{z} \mid \pi_{z}$. Note that $\pi_{z}=\{q\}$ does not work because $z \Perp e, h \mid q$ does not hold. We thus set $\mathrm{pa}_{z}=\{q, h\}$. (Third graph in Figure 2.)
- We consider $a:$ pre $_{a}=\{e, h, q, z\}$. This is the last node in the ordering. To find the minimal set $\pi_{a}$ for which $a \Perp \operatorname{pre}_{a} \backslash \pi_{a} \mid \pi_{a}$, we can determine its Markov blanket $\operatorname{MB}(a)$. The Markov blanket is the set of parents (none), children ( $q$ ), and co-parents of $a(z)$ in Figure 1, so that $\operatorname{MB}(a)=\{q, z\}$. We thus set pa $a_{a}=\{q, z\}$.(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:

- $M B(a)=\{q, z\}$
- $M B(z)=\{a, q, h\}$
- $M B(h)=\{z\}$
- $M B(q)=\{a, z, e\}$
- $M B(e)=\{q\}$
(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.


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 $p\left(x_{i} \mid \mathrm{pa}_{i}\right)$ corresponds to a factor $\phi_{i}\left(x_{i}, \mathrm{pa}_{i}\right)$ 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 $x \Perp$ $y \mid\{c h i l d$, 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, $p\left(x_{4} \mid x_{1}, x_{2}, x_{3}\right)$ corresponds to a factor $\phi\left(x_{4}, x_{1}, x_{2}, x_{3}\right)$ 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 $\mathrm{pa}_{2}=\left\{x_{1}\right\}$. 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 $\mathrm{pa}_{3}=\left\{x_{1}, x_{2}\right\}$. 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 $\mathrm{pa}_{4}=$ $\left\{x_{2}, x_{3}\right\}$. See fourth graph in Figure 3.
- The same reasoning shows that $\mathrm{pa}_{5}=\left\{x_{3}, x_{4}\right\}$. 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 $x_{3} \Perp x_{2} \mid x_{1}$ (check!). But this independency assertion does not hold in the undirected graph so that the graph in Figure 4 is not an I-map.


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

## Exercise 3. Limits of directed and undirected graphical models

We here consider the probabilistic model $p\left(y_{1}, y_{2}, x_{1}, x_{2}\right)=p\left(y_{1}, y_{2} \mid x_{1}, x_{2}\right) p\left(x_{1}\right) p\left(x_{2}\right)$ where $p\left(y_{1}, y_{2} \mid x_{1}, x_{2}\right)$ factorises as

$$
\begin{equation*}
p\left(y_{1}, y_{2} \mid x_{1}, x_{2}\right)=p\left(y_{1} \mid x_{1}\right) p\left(y_{2} \mid x_{2}\right) \phi\left(y_{1}, y_{2}\right) n\left(x_{1}, x_{2}\right) \tag{1}
\end{equation*}
$$

with $n\left(x_{1}, x_{2}\right)$ equal to

$$
\begin{equation*}
n\left(x_{1}, x_{2}\right)=\left(\int p\left(y_{1} \mid x_{1}\right) p\left(y_{2} \mid x_{2}\right) \phi\left(y_{1}, y_{2}\right) \mathrm{d} y_{1} \mathrm{~d} y_{2}\right)^{-1} . \tag{2}
\end{equation*}
$$

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 $p\left(y_{1}, y_{2}, x_{1}, x_{2}\right)$ satisfies the following independencies

$$
x_{1} \Perp x_{2} \quad x_{1} \Perp y_{2}\left|y_{1}, x_{2} \quad x_{2} \Perp y_{1}\right| y_{2}, x_{1}
$$

We here investigate whether undirected and directed graph can represent these independencies.
(a) Draw the undirected graph for $p\left(y_{1}, y_{2}, x_{1}, x_{2}\right)$ and check whether graph separation allows us to see all independencies listed above.

Solution. We write

$$
p\left(y_{1}, y_{2}, x_{1}, x_{2}\right)=p\left(y_{1} \mid x_{1}\right) p\left(y_{2} \mid x_{2}\right) \phi\left(y_{1}, y_{2}\right) n\left(x_{1}, x_{2}\right) p\left(x_{1}\right) p\left(x_{2}\right)
$$

as

$$
\begin{align*}
p\left(y_{1}, y_{2}, x_{1}, x_{2}\right) & =\phi_{1}\left(y_{1}, x_{1}\right) \phi_{2}\left(y_{2}, x_{2}\right) \phi_{3}\left(y_{1}, y_{2}\right) \phi_{4}\left(x_{1}, x_{2}\right)  \tag{S.1}\\
\phi_{1}\left(y_{1}, x_{1}\right) & =p\left(y_{1} \mid x_{1}\right) p\left(x_{1}\right)  \tag{S.2}\\
\phi_{2}\left(y_{2}, x_{2}\right) & =p\left(y_{2} \mid x_{2}\right) p\left(x_{2}\right)  \tag{S.3}\\
\phi_{3}\left(y_{1}, y_{2}\right) & =\phi\left(y_{1}, y_{2}\right)  \tag{S.4}\\
\phi_{4}\left(x_{1}, x_{2}\right) & =n\left(x_{1}, x_{2}\right) \tag{S.5}
\end{align*}
$$

The corresponding undirected graph is as follows.


While the graph implies $x_{1} \Perp y_{2} \mid y_{1}, x_{2}$ and $x_{2} \Perp y_{1} \mid y_{2}, x_{1}$, the independency $x_{1} \Perp x_{2}$ is not represented.
(b) Draw the directed graph for $p\left(y_{1}, y_{2}, x_{1}, x_{2}\right)=p\left(y_{1}, y_{2} \mid x_{1}, x_{2}\right) p\left(x_{1}\right) p\left(x_{2}\right)$ 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.

