Expressive Power of Graphical Models

Chris Williams
(based on slides by Michael U. Gutmann)

Probabilistic Modelling and Reasoning (INFR11134)
School of Informatics, The University of Edinburgh

Spring Semester 2024

Recap

- Need for efficient representation of probabilistic models
  - Restrict the number of directly interacting variables by making independence assumptions
  - Restrict the form of interaction by making parametric family assumptions
- DAGs and undirected graphs to represent independencies and factorisations
- Equivalences between independencies (Markov properties) and factorisation
- Rules for reading independencies from the graph that hold for all distributions that factorise over the graph

Program

1. Graphs as independency maps (I-maps)
2. Equivalence of I-maps (I-equivalence)

Program

1. Graphs as independency maps (I-maps)
   - I-maps
   - Perfect maps
   - Minimal I-maps
   - Strengths and weaknesses of directed and undirected graphs
2. Equivalence of I-maps (I-equivalence)

I-map

- We have seen that graphs represent independencies. We say that they are independency maps (I-maps).
- *Definition*: Let $\mathcal{U}$ be a set of independencies that random variables $\mathbf{x} = (x_1, \ldots x_d)$ satisfy. A DAG or undirected graph K with nodes x_i is said to be an independency map (I-map) for $\mathcal{U}$ if the independencies $\mathcal{I}(K)$ asserted by the graph are part of $\mathcal{U}$:
$$\mathcal{I}(K) \subseteq \mathcal{U}$$
- $\mathcal{I}(K)$ may miss some independencies that hold in $\mathcal{U}$.
- An I-map is a "directed I-map" if K is a DAG, and an "undirected I-map" if K is an undirected graph.

I-map

The set of independencies $\mathcal{U}$ can be specified in different ways. For example:

- as a list of independencies, e.g.

$$\mathcal{U} = \{x_1 \perp\!\!\!\perp x_2\}$$

- as the independencies implied by another graph K_0

$$\mathcal{U} = \mathcal{I}(K_0)$$

- denoting by $\mathcal{I}(p)$ all the independencies satisfied by a specific distribution p, we can have

$$\mathcal{U} = \mathcal{I}(p)$$

I-maps and factorisation

- We have previously found that all independencies asserted by the graph K hold for all p that factorise over K.
- Hence, if p factorises over K, we have

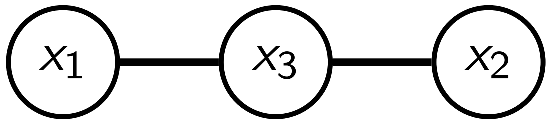
$$\mathcal{I}(K) \subseteq \mathcal{I}(p)$$

  and K is an I-map for $\mathcal{I}(p)$
- But we do not have guarantees that $\mathcal{I}(K)$ equals $\mathcal{I}(p)$ since, as we have seen, $\mathcal{I}(K)$ may miss some independencies that hold for p.

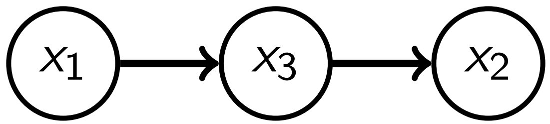
Examples of I-maps

Consider $\mathcal{U} = \{x_1 \perp\!\!\!\perp x_2,\ x_1 \perp\!\!\!\perp x_2 | x_3,\ x_2 \perp\!\!\!\perp x_3,\ x_2 \perp\!\!\!\perp x_3 | x_1\}$

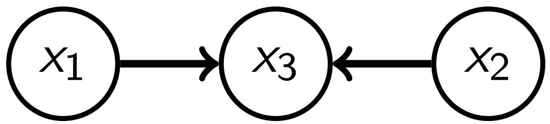
- $\mathcal{I}(H) = \{x_1 \perp\!\!\!\perp x_2 | x_3\} \subset \mathcal{U}$



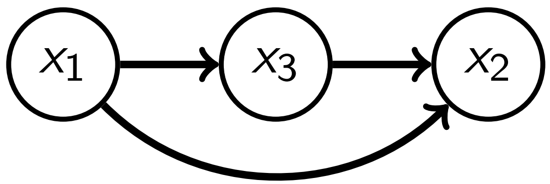
- $\mathcal{I}(G) = \{x_1 \perp\!\!\!\perp x_2 | x_3\} \subset \mathcal{U}$



- $\mathcal{I}(G) = \{x_1 \perp\!\!\!\perp x_2\} \subset \mathcal{U}$



- $\mathcal{I}(G) = \varnothing \subset \mathcal{U}$



Remarks

- I-maps are not unique.
- Different I-maps may make the same independence assertions. (discussed later under I-equivalence)
- Criterion for an I-map is that the independence assertions made by the graph are true. I-maps are not concerned with the number of independence assertions made.
- I-maps of $\mathcal{U}$ are allowed to "miss" some independencies in $\mathcal{U}$.
- The complete graph does not make any assertions. Empty set is trivially a subset of any $\mathcal{U}$, so that the complete graph is trivially an I-map.

Perfect maps

- *Definition*: K is said to be a perfect I-map (or P-map) for $\mathcal{U}$ if $\mathcal{I}(K) = \mathcal{U}$.
- Let K be a DAG or an undirected graph. For what set $\mathcal{U}$ of independencies is a graph K a perfect map?
- We have seen that:
  if X are Y and not (d-)separated by Z *then* $X \not\perp\!\!\!\perp Y|Z$ for some p that factorises over K (some $\equiv$ not all)
- Contrapositive: (Reminder: $A \Rightarrow B \Leftrightarrow \bar{B} \Rightarrow \bar{A}$)

  if $X \perp\!\!\!\perp Y|Z$ for all p that factorise over K *then* X and Y are (d-)separated by Z
- Denote by $\mathcal{P}_K$ the set of all p that factorise over K. We thus have:

$$\left[\bigcap_{p \in \mathcal{P}_K} \mathcal{I}(p)\right] \subseteq \mathcal{I}(K)$$

Perfect maps and factorisation

- Since for all individual p we have $\mathcal{I}(K) \subseteq \mathcal{I}(p)$, it follows that

$$\left[\bigcap_{p \in \mathcal{P}_K} \mathcal{I}(p)\right] \subseteq \mathcal{I}(K) \subseteq \left[\bigcap_{p \in \mathcal{P}_K} \mathcal{I}(p)\right]$$

  and hence that

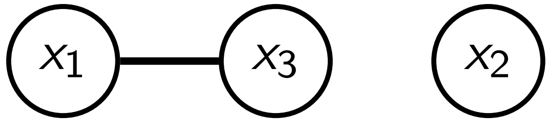
$$\mathcal{I}(K) = \bigcap_{p \in \mathcal{P}_K} \mathcal{I}(p)$$

- In plain English: K is a perfect map for the independencies that hold for all p that factorise over the graph.
- This result is not very surprising. It just says that K is a perfect map for the graphical models (set of distributions) that were defined by K in the first place!

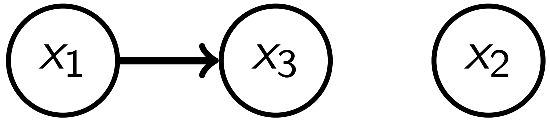
Examples of P-maps

Consider again $\mathcal{U} = \{x_1 \perp\!\!\!\perp x_2,\ x_1 \perp\!\!\!\perp x_2 | x_3,\ x_2 \perp\!\!\!\perp x_3,\ x_2 \perp\!\!\!\perp x_3 | x_1\}$

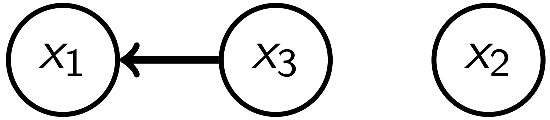
- $\mathcal{I}(H) = \mathcal{U}$



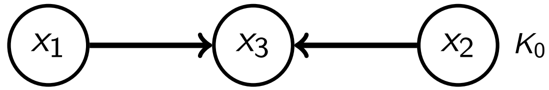
- $\mathcal{I}(G) = \mathcal{U}$



- $\mathcal{I}(G) = \mathcal{U}$

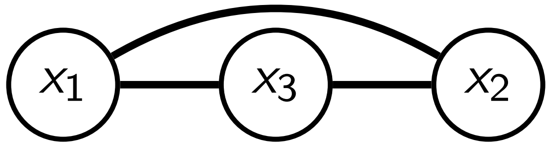


Collider does not have an undirected P-map

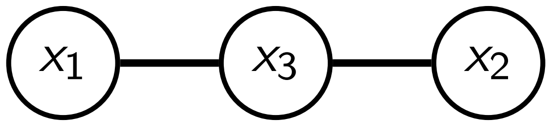


Consider the independencies represented by the collider K_0.

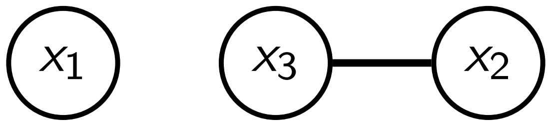
- Let $\mathcal{U} = \mathcal{I}(K_o) = \{x_1 \perp\!\!\!\perp x_2\}$
- I-map for $\mathcal{U}$: $\mathcal{I}(H) = \{\}$



- Not an I-map for $\mathcal{U}$: graph wrongly asserts $x_1 \perp\!\!\!\perp x_2 \mid x_3$



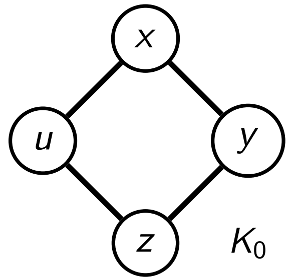
- Not an I-map for $\mathcal{U}$: graph wrongly asserts $x_1 \perp\!\!\!\perp x_3$



- Going through all undirected graphs shows that there is no undirected perfect I-map for $\mathcal{U}$.

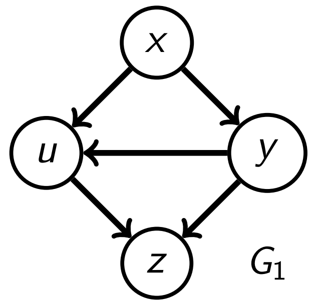
Diamond does not have a directed P-map

Consider the independencies represented by the diamond configuration K_0.



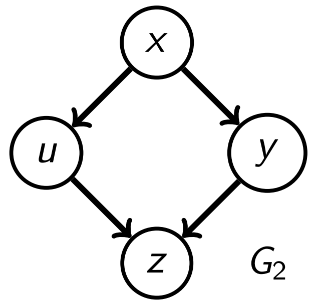
K_0

- Let $\mathcal{U} = \mathcal{I}(K_0) = \{x \perp\!\!\!\perp z | u, y;\ u \perp\!\!\!\perp y | x, z\}$
- G_1 is an I-map for $\mathcal{U}$:
  $\mathcal{I}(G_1) = \{x \perp\!\!\!\perp z | u, y\} \subset \mathcal{U}$



G_1

- G_2 is not an I-map for $\mathcal{U}$:
  graph wrongly asserts $u \perp\!\!\!\perp y | x$
- Going through all DAGs shows that there is no directed perfect I-map for $\mathcal{U}$.



G_2

Minimal I-maps

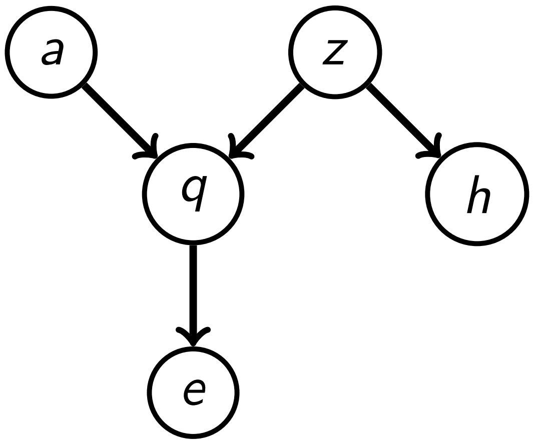
- Directed or undirected perfect maps may not always exist.
- On the other hand, criterion for a graph to be an I-map is weak (full graph is an I-map!).
- Compromise: Let us "sparsify" I-maps so that they become more useful.
- *Definition*: A minimal I-map is an I-map such that if you remove an edge (more independencies), the resulting graph is not an I-map any more.
- Note: A perfect map for $\mathcal{U}$ is also a minimal I-map for $\mathcal{U}$
  (being perfect is a stronger requirement than being minimal)

Our previous visualisations of $p(\mathbf{x})$ are minimal I-maps

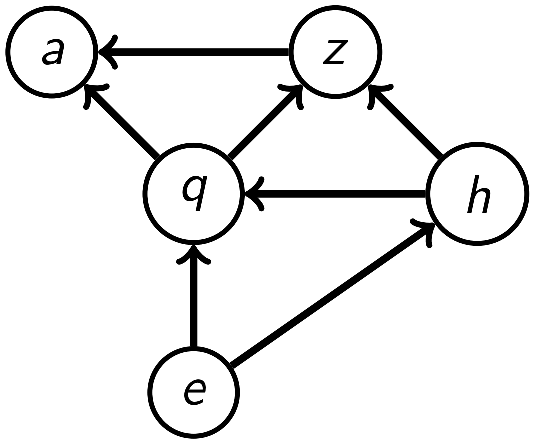
- To visualise $p(\mathbf{x})$ as a DAG:
  - Ordering + independencies $x_i \perp\!\!\!\perp (\mathrm{pre}_i \setminus \pi_i) \mid \pi_i$ that $p(\mathbf{x})$ satisfies, where π_i is a *minimal* subset of the predecessors
  - Construct a graph with the π_i as parents pa_i
  - Gives a minimal I-map of $\mathcal{I}(p)$ because the π_i are the *minimal* subsets.
- To visualise $p(\mathbf{x})$ as an undirected graph:
  - Determine the Markov blanket for each variable x_i
  - Construct a graph where the neighbours of x_i are its Markov blanket.
  - Gives a minimal I-map of $\mathcal{I}(p)$ because the Markov blanket is the *minimal* set of variables that makes the x_i independent from the remaining variables.

Directed minimal I-maps are not unique

Consider p with perfect I-map G_1. Use G_1 to determine $x_i \perp\!\!\!\perp (\mathrm{pre}_i \setminus \pi_i) \mid \pi_i$ for a given ordering of the variables.



Graph G_1



Minimal I-map G_2 for ordering (e, h, q, z, a), see exercises

- Directed (minimal) I-maps are not unique.
- Here: $\mathcal{I}(G_2) \subset \mathcal{I}(G_1) = \mathcal{I}(p)$.
- The minimal directed I-maps from different orderings may not represent the same independencies. (they are not I-equivalent)

Pros/cons of directed and undirected graphs

- Some independencies are more easily represented with DAGs, others with undirected graphs.
- Both directed and undirected graphical models have strengths and weaknesses.
- Undirected graphs are suitable when interactions are symmetrical and when there is no natural ordering of the variables, but they cannot represent "explaining away" phenomena (colliders).
- DAGs are suitable when we have an idea of the data generating process (e.g. what is "causing" what), but they may force directionality where there is none.
- It is possible to combine the individual strengths with mixed/partially directed graphs (see e.g. Barber, Section 4.3; Lauritzen, Section 3.2.3, not examinable).

Program

1. Graphs as independency maps (I-maps)

2. Equivalence of I-maps (I-equivalence)
   - I-equivalence for DAGs: check the skeletons and the immoralities
   - I-equivalence for undirected graphs: check the skeletons
   - I-equivalence between directed and undirected graphs

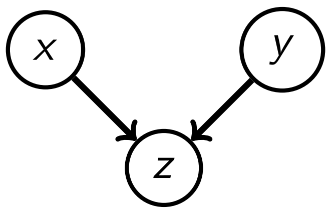
I-equivalence for DAGs

- How do we determine whether two DAGs make the same independence assertions (that they are "I-equivalent")?
- From d-separation: what matters is
  - which node is connected to which irrespective of direction (skeleton)
  - the set of collider (head-to-head) connections

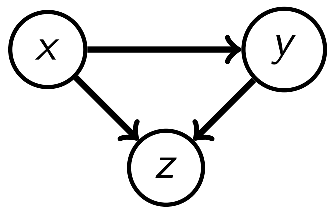
Connection	$p(x, y)$	$p(x, y \mid z)$
$x \rightarrow z \rightarrow y$	$x \not\perp\!\!\!\perp y$	$x \perp\!\!\!\perp y \mid z$
$x \leftarrow z \rightarrow y$	$x \not\perp\!\!\!\perp y$	$x \perp\!\!\!\perp y \mid z$
$x \rightarrow z \leftarrow y$	$x \perp\!\!\!\perp y$	$x \not\perp\!\!\!\perp y \mid z$

I-equivalence for DAGs

- The situation $x \perp\!\!\!\perp y$ and $x \not\perp\!\!\!\perp y \mid z$ can only happen if we have colliders without a "covering edge" $x \to y$ or $x \leftarrow y$, that is when parents of the collider node are not directly connected.
- Colliders without a covering edge are called "immoralities".
- Theorem: For two DAGs G_1 and G_2:
  G_1 and G_2 are I-equivalent $\iff$ G_1 and G_2 have the same skeleton and the same set of immoralities.
  (for a proof, see e.g. Theorem 4.4, Koski and Noble, 2009; not examinable)



$x \perp\!\!\!\perp y$ and $x \not\perp\!\!\!\perp y \mid z$
Collider w/o covering edge

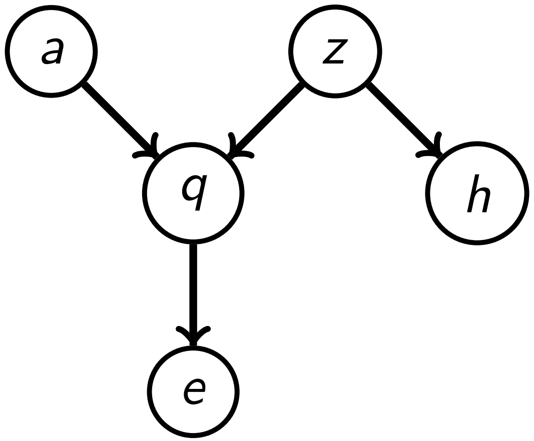


$x \not\perp\!\!\!\perp y$ and $x \not\perp\!\!\!\perp y \mid z$
Collider with covering edge

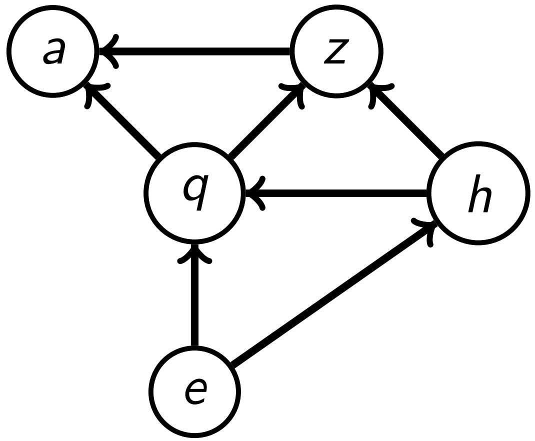
Example

Not I-equivalent because of skeleton mismatch:

G_1:



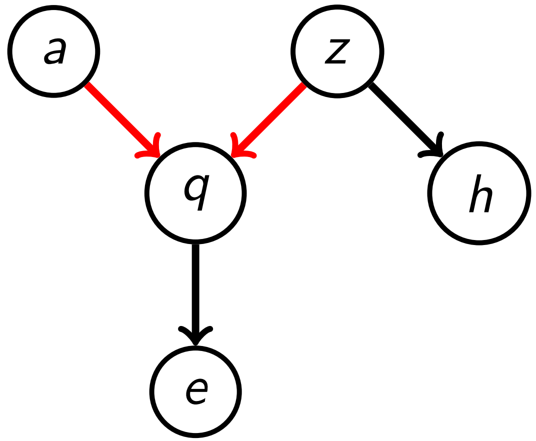
G_2:



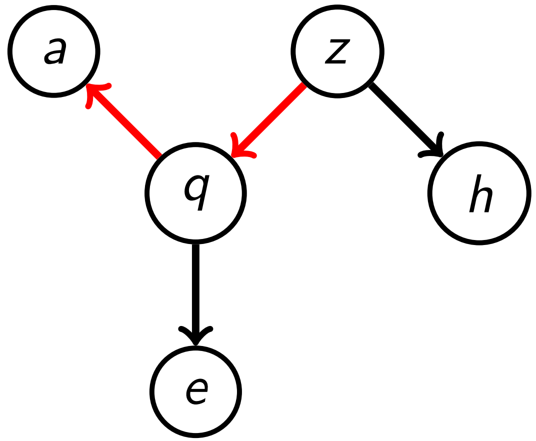
Example

Not I-equivalent because of immoralities mismatch:

G_1:



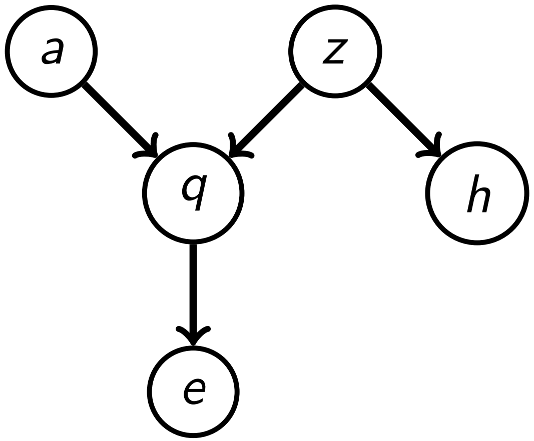
G_2:



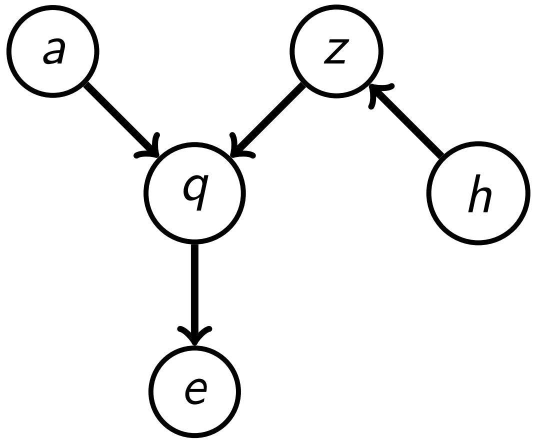
Example

I-equivalent (same skeleton, same immoralities):

G_1:

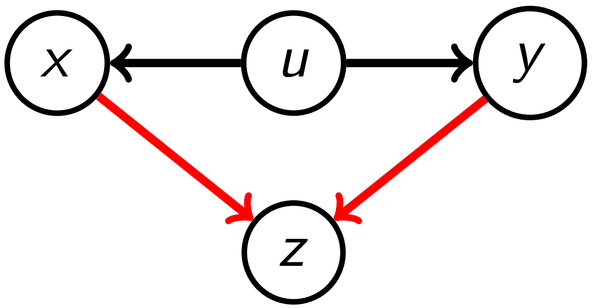


G_2:

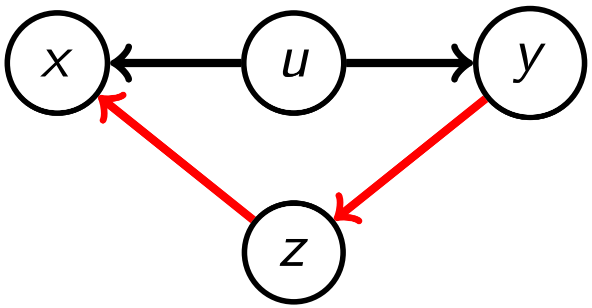


Example

Not I-equivalent (immoralities mismatch)



$x \perp\!\!\!\perp y \mid u$ and $x \not\!\perp\!\!\!\perp y \mid u, z$
Immorality: collider w/o covering edge



$x \not\!\perp\!\!\!\perp y \mid u$ and $x \perp\!\!\!\perp y \mid u, z$
Not an immorality

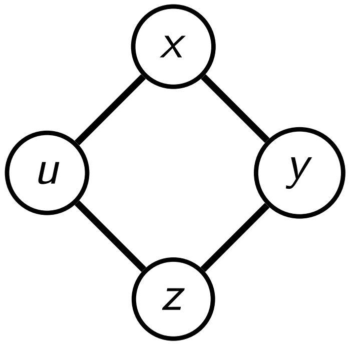
I-equivalence for undirected graphs

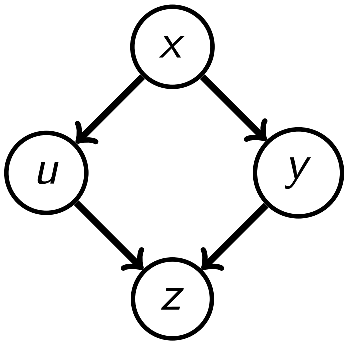
- Different undirected graphs make different independence assertions.
- I-equivalent if their skeleton is the same.

I-equivalence between directed and undirected graphs

Recall the example about non-existence of P-maps:

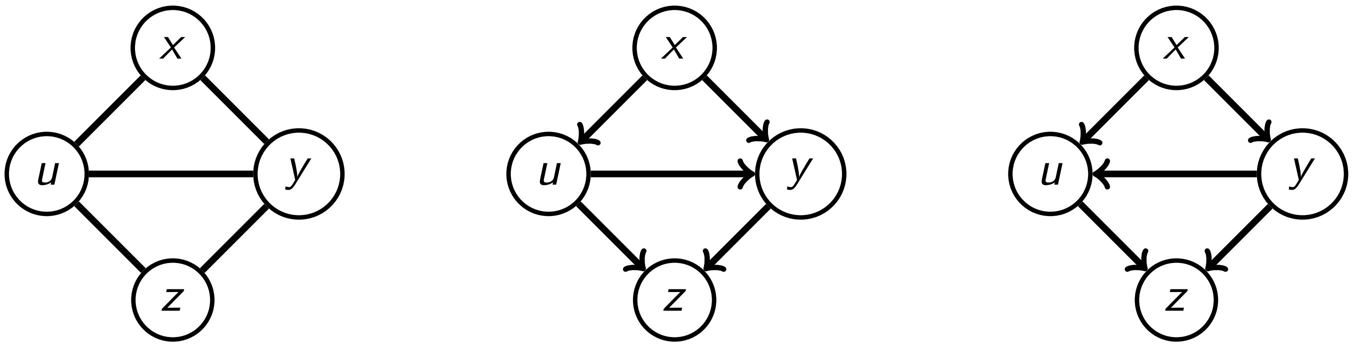
- Immoralities (colliders without a covering edge) allow DAGs to represent independencies that cannot be represented with undirected graphs (e.g. $x \perp\!\!\!\perp y$ without enforcing $x \perp\!\!\!\perp y|z$)
- Diamond configurations (where the loop has length > 3) allow undirected graphs to represent independencies that DAGs cannot represent.
- Connection between the two: Turning a diamond configuration into a DAG introduces an immorality.





I-equivalence between directed and undirected graphs

- For DAGs without immoralities, only the skeleton is relevant for I-equivalence. Since the orientation of the arrows does not matter, we can just replace them with undirected edges to obtain an I-equivalent undirected graph.
- Relatedly, for chordal/triangulated undirected graphs (where the longest loop without shortcuts is a triangle), introducing arrows does not lead to immoralities (there is always a covering edge!) and obtained DAGs are I-equivalent to the undirected graph.
- Example of I-equivalent graphs:



(note the covering edge between u and y)

Program recap

1. Graphs as independency maps (I-maps)
   - I-maps
   - Perfect maps
   - Minimal I-maps
   - Strengths and weaknesses of directed and undirected graphs

2. Equivalence of I-maps (I-equivalence)
   - I-equivalence for DAGs: check the skeletons and the immoralities
   - I-equivalence for undirected graphs: check the skeletons
   - I-equivalence between directed and undirected graphs

Further Reading

- There is some material on I-maps in Bishop (2006) sec. 8.3.4, but with less detail than here
- Koller and Friedman (2009) cover I-maps (sec. 3.2.3.1 for DGMs, sec. 4.3.3 for UGMs), minimal and perfect I-maps (secs. 3.4.1, 3.4.2) and I-equivalance (sec. 3.3.4).

But following carefully the material (by MG) in the slides and tutorials is sufficient for PMR

Credits

These slides are modified from ones produced by Michael Gutmann, made available under Creative Commons licence CC BY 4.0.