**Syllabus, Chapter 3:**

# Independences and Graphical Representations

## A qualitative notion of independence

**Observation**:

People are capable of making statements about independences among variables without having to perform numerical calculations.

**Conclusion**:

In human reasoning behaviour, the qualitative notion of independence is more fundamental than the quantitative notion of independence.

# The (probabilistic) independence relation of a joint distribution

**Definition**: Let $\boldsymbol{V}$ be a set of random variables and let $\Pr$ be a joint probability distribution on $\boldsymbol{V}$.

The independence relation $I_{\Pr}$ of $\Pr$ is a set $I_{\Pr} \subseteq \mathcal{P}(\boldsymbol{V}) \times \mathcal{P}(\boldsymbol{V}) \times \mathcal{P}(\boldsymbol{V})$, defined for all $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z} \subseteq \boldsymbol{V}$ by

$$(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y}) \in I_{\Pr} \text{ if and only if } \Pr(\boldsymbol{X} \mid \boldsymbol{Y} \wedge \boldsymbol{Z}) = \Pr(\boldsymbol{X} \mid \boldsymbol{Z})$$

**Remarks**:

- $(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y}) \in I_{\Pr}$ will be written as $I_{\Pr}(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y})$;
  $(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y}) \notin I_{\Pr}$ will be written as $\neg I_{\Pr}(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y})$;
- a statement $I_{\Pr}(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y})$ is called an independence statement for the joint distribution $\Pr$.

# Properties of $I_{\mathrm{Pr}}$: symmetry

**Lemma**: $I_{\mathrm{Pr}}(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y})$ if and only if $I_{\mathrm{Pr}}(\boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{X})$

**Proof**: [NB This is an exercise from the Syllabus. In proofs you should also explain your steps, like done in the lectures.]

$$
\begin{aligned}
I_{\mathrm{Pr}}(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y}) &\iff \Pr(\boldsymbol{X} \mid \boldsymbol{Y} \wedge \boldsymbol{Z}) = \Pr(\boldsymbol{X} \mid \boldsymbol{Z}) \\
&\iff \frac{\Pr(\boldsymbol{X} \wedge \boldsymbol{Y} \wedge \boldsymbol{Z})}{\Pr(\boldsymbol{Y} \wedge \boldsymbol{Z})} = \frac{\Pr(\boldsymbol{X} \wedge \boldsymbol{Z})}{\Pr(\boldsymbol{Z})} \\
&\iff \frac{\Pr(\boldsymbol{X} \wedge \boldsymbol{Y} \wedge \boldsymbol{Z})}{\Pr(\boldsymbol{X} \wedge \boldsymbol{Z})} = \frac{\Pr(\boldsymbol{Y} \wedge \boldsymbol{Z})}{\Pr(\boldsymbol{Z})} \\
&\iff \Pr(\boldsymbol{Y} \mid \boldsymbol{X} \wedge \boldsymbol{Z}) = \Pr(\boldsymbol{Y} \mid \boldsymbol{Z}) \\
&\iff I_{\mathrm{Pr}}(\boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{X})
\end{aligned}
$$

■

# Properties of $I_{\mathrm{Pr}}$: decomposition

**Lemma**: $I_{\mathrm{Pr}}(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y} \cup \boldsymbol{W}) \Rightarrow I_{\mathrm{Pr}}(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y}) \ \wedge \ I_{\mathrm{Pr}}(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{W})$

**Proof**: (sketch) (Note: for $\boldsymbol{U} = \boldsymbol{Y} \cup \boldsymbol{W}$, $c_{\boldsymbol{U}} = c_{\boldsymbol{Y}} \wedge c_{\boldsymbol{W}}$!)
Suppose that

$\mathrm{Pr}(\boldsymbol{X} \mid \boldsymbol{Y} \wedge \boldsymbol{W} \wedge \boldsymbol{Z}) = \mathrm{Pr}(\boldsymbol{X} \mid \boldsymbol{Z})$. Then, by definition,

$$\mathrm{Pr}(\boldsymbol{X} \wedge \boldsymbol{Y} \wedge \boldsymbol{W} \wedge \boldsymbol{Z}) = \mathrm{Pr}(\boldsymbol{Y} \wedge \boldsymbol{W} \wedge \boldsymbol{Z}) \cdot \frac{\mathrm{Pr}(\boldsymbol{X} \wedge \boldsymbol{Z})}{\mathrm{Pr}(\boldsymbol{Z})}$$

For $\mathrm{Pr}(\boldsymbol{X} \mid \boldsymbol{Y} \wedge \boldsymbol{Z})$ we find that

$$\begin{aligned}
\mathrm{Pr}(\boldsymbol{X} \mid \boldsymbol{Y} \wedge \boldsymbol{Z}) &= \frac{\mathrm{Pr}(\boldsymbol{X} \wedge \boldsymbol{Y} \wedge \boldsymbol{Z})}{\mathrm{Pr}(\boldsymbol{Y} \wedge \boldsymbol{Z})} \\
&= \frac{\sum_{c_{\boldsymbol{W}}} \mathrm{Pr}(\boldsymbol{X} \wedge \boldsymbol{Y} \wedge \boldsymbol{Z} \wedge c_{\boldsymbol{W}})}{\mathrm{Pr}(\boldsymbol{Y} \wedge \boldsymbol{Z})} \\
&= \frac{\mathrm{Pr}(\boldsymbol{X} \wedge \boldsymbol{Z})}{\mathrm{Pr}(\boldsymbol{Z})} = \mathrm{Pr}(\boldsymbol{X} \mid \boldsymbol{Z}) \qquad \blacksquare
\end{aligned}$$

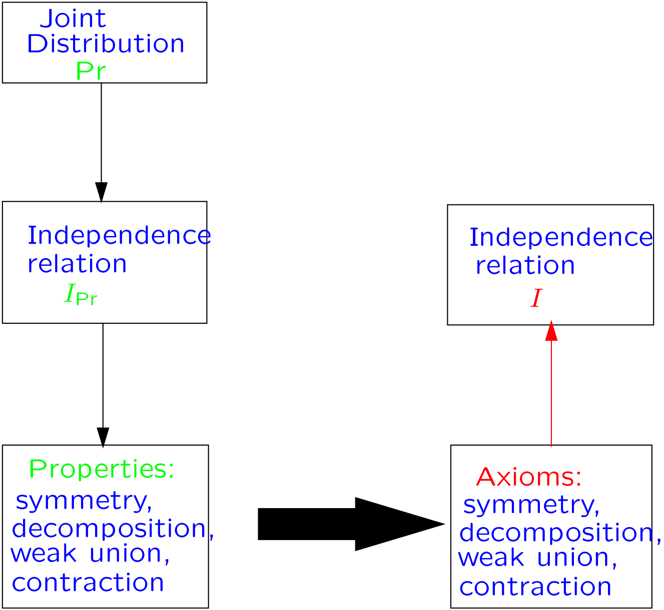
## Properties of $I_{\mathrm{Pr}}$: weak union, contraction

**Lemma**:

- if $I_{\mathrm{Pr}}(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y} \cup \boldsymbol{W})$ then $I_{\mathrm{Pr}}(\boldsymbol{X}, \boldsymbol{Z} \cup \boldsymbol{W}, \boldsymbol{Y})$ (weak union);
- if $I_{\mathrm{Pr}}(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{W})$ and $I_{\mathrm{Pr}}(\boldsymbol{X}, \boldsymbol{Z} \cup \boldsymbol{W}, \boldsymbol{Y})$ then $I_{\mathrm{Pr}}(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y} \cup \boldsymbol{W})$ (contraction)
- (for strictly positive $\mathrm{Pr}$ also the intersection property holds; see syllabus)

**Proof**: left as exercise 3.1.

# The definition of the independence relation



# The (qualitative) independence relation $I$

**Definition**:

Let $V$ be a set of random variables and let $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{W} \subseteq \boldsymbol{V}$.

An independence relation $I$ on $\boldsymbol{V}$ is a ternary relation $I \subseteq \mathcal{P}(\boldsymbol{V}) \times \mathcal{P}(\boldsymbol{V}) \times \mathcal{P}(\boldsymbol{V})$ that satisfies the following properties:

- if $I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y})$ then $I(\boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{X})$;
- if $I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y} \cup \boldsymbol{W})$ then $I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y})$ and $I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{W})$;
- if $I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y} \cup \boldsymbol{W})$ then $I(\boldsymbol{X}, \boldsymbol{Z} \cup \boldsymbol{W}, \boldsymbol{Y})$;
- if $I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{W})$ and $I(\boldsymbol{X}, \boldsymbol{Z} \cup \boldsymbol{W}, \boldsymbol{Y})$ then $I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y} \cup \boldsymbol{W})$.

The first property is called the symmetry *axiom*; the second is called the decomposition axiom; the third is referred to as the weak union axiom; the last one is called contraction.

# An example

**Lemma**:

Let $I$ be an independence relation on a set of random variables $V$. We have that for all $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{W} \subseteq \boldsymbol{V}$:

$$\text{if } I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y}) \text{ and } I(\boldsymbol{X} \cup \boldsymbol{Z}, \boldsymbol{Y}, \boldsymbol{W}) \text{ then } I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{W})$$

**Proof**:

We observe that

$$I(\boldsymbol{X} \cup \boldsymbol{Z}, \boldsymbol{Y}, \boldsymbol{W}) \Rightarrow_{\text{symm}} I(\boldsymbol{W}, \boldsymbol{Y}, \boldsymbol{X} \cup \boldsymbol{Z}) \Rightarrow_{\text{weakunion}}$$

$$\Rightarrow I(\boldsymbol{W}, \boldsymbol{Y} \cup \boldsymbol{Z}, \boldsymbol{X}) \Rightarrow_{\text{symm}} I(\boldsymbol{X}, \boldsymbol{Y} \cup \boldsymbol{Z}, \boldsymbol{W})$$

From $I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y})$, $I(\boldsymbol{X}, \boldsymbol{Y} \cup \boldsymbol{Z}, \boldsymbol{W})$ and the contraction axiom we have that $I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{W} \cup \boldsymbol{Y})$; decomposition now gives $I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{W})$. ■

## Representing independences

Different ways exist of representing an independence relation:

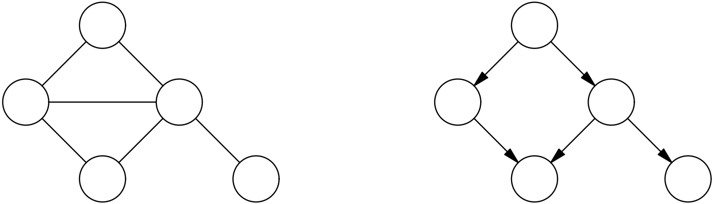
- explicitly list all independence statements of the relation;
- explicitly list only the independence statements of a suitable subset of the relation (the ‘basis’) — all other statements are implicitly represented by means of the axioms;
- code the independence relation in a graph;
- …

# An example

Consider $\boldsymbol{V} = \{V_1, V_2, V_3, V_4\}$ and independence relation $I$ on $\boldsymbol{V}$:

| | | |
|---|---|---|
| $I(\{V_1\}, \emptyset, \{V_4\})$ | $I(\{V_2\}, \emptyset, \{V_1\})$ | $I(\{V_4\}, \{V_1\}, \{V_2\})$ |
| $I(\{V_2\}, \emptyset, \{V_4\})$ | $I(\{V_1, V_4\}, \emptyset, \{V_2\})$ | $I(\{V_4\}, \{V_1\}, \{V_3\})$ |
| $I(\{V_3\}, \emptyset, \{V_4\})$ | $I(\{V_2, V_4\}, \emptyset, \{V_1\})$ | $I(\{V_4\}, \{V_1\}, \{V_2, V_3\})$ |
| $I(\{V_4\}, \emptyset, \{V_1\})$ | $I(\{V_2\}, \emptyset, \{V_1, V_4\})$ | $I(\{V_1\}, \{V_2\}, \{V_4\})$ |
| $I(\{V_4\}, \emptyset, \{V_2\})$ | $I(\{V_1\}, \emptyset, \{V_2, V_4\})$ | $I(\{V_3\}, \{V_2\}, \{V_4\})$ |
| $I(\{V_4\}, \emptyset, \{V_3\})$ | $I(\{V_2\}, \{V_1\}, \{V_4\})$ | $I(\{V_1, V_3\}, \{V_2\}, \{V_4\})$ |
| $I(\{V_1, V_2\}, \emptyset, \{V_4\})$ | $I(\{V_3\}, \{V_1\}, \{V_4\})$ | $I(\{V_4\}, \{V_2\}, \{V_1\})$ |
| $I(\{V_1, V_3\}, \emptyset, \{V_4\})$ | $I(\{V_2, V_3\}, \{V_1\}, \{V_4\})$ | $I(\{V_4\}, \{V_2\}, \{V_3\})$ |
| $I(\{V_2, V_3\}, \emptyset, \{V_4\})$ | $I(\{V_4\}, \{V_1, V_2\}, \{V_3\})$ | $I(\{V_4\}, \{V_2\}, \{V_1, V_3\})$ |
| $I(\{V_4\}, \emptyset, \{V_1, V_2\})$ | $I(\{V_2\}, \{V_1, V_3\}, \{V_4\})$ | $I(\{V_1\}, \{V_3\}, \{V_4\})$ |
| $I(\{V_4\}, \emptyset, \{V_1, V_3\})$ | $I(\{V_4\}, \{V_1, V_3\}, \{V_2\})$ | $I(\{V_2\}, \{V_3\}, \{V_4\})$ |
| $I(\{V_4\}, \emptyset, \{V_2, V_3\})$ | $I(\{V_1\}, \{V_2, V_3\}, \{V_4\})$ | $I(\{V_1, V_2\}, \{V_3\}, \{V_4\})$ |
| $I(\{V_1, V_2, V_3\}, \emptyset, \{V_4\})$ | $I(\{V_4\}, \{V_2, V_3\}, \{V_1\})$ | $I(\{V_1\}, \{V_4\}, \{V_2\})$ |
| $I(\{V_4\}, \emptyset, \{V_1, V_2, V_3\})$ | $I(\{V_4\}, \{V_3\}, \{V_1, V_2\})$ | $I(\{V_2\}, \{V_4\}, \{V_1\})$ |
| $I(\{V_1\}, \emptyset, \{V_2\})$ | $I(\{V_4\}, \{V_3\}, \{V_1\})$ | $I(\{V_3\}, \{V_1, V_2\}, \{V_4\})$ |

# Coding an independence relation with a graph

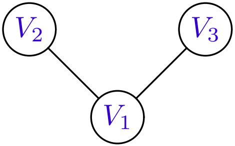


The global idea of coding an independence relation $I$ on $\boldsymbol{V}$ in graph $G$:

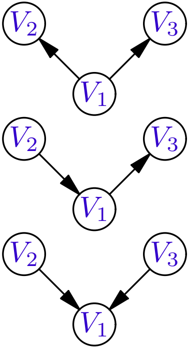
- each variable $V_i \in \boldsymbol{V}$ corresponds to a node $V_i \in \boldsymbol{V}_G$;
- (combinations of) edges/arcs define a graphical notion of (d-)separation;
- there exists a mapping between (d-)separation and relation $I$

# Graphs as a language

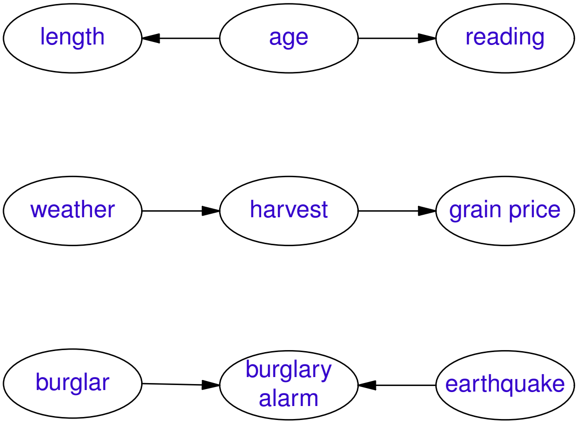
The directed graph provides a more expressive language than the undirected graph:




vs.



# Example reasoning patterns




- do arcs capture causality? not necessarily
- do reasoning patterns really differ? last differs from first two
- what if we drop the direction of the arcs? last one incomplete

# “Intercausal” interaction

Consider an experiment with two coins and a bell: the bell sounds iff the two coins have the same outcome after a toss.

Consider: variable $C_1$: the outcome of tossing coin one;
variable $C_2$: the outcome of tossing coin two;
variable $B$: whether or not the bell sounds;
independence relation $I$ for this experiment.

We have, among others, that

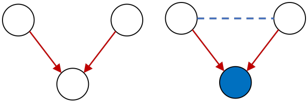
$$
\begin{array}{ll}
I(\{C_1\}, \emptyset, \{C_2\}) & \neg I(\{C_1\}, \{B\}, \{C_2\}) \\
I(\{C_1\}, \emptyset, \{B\}) & \neg I(\{C_1\}, \{C_2\}, \{B\}) \\
I(\{C_2\}, \emptyset, \{B\}) & \neg I(\{C_2\}, \{C_1\}, \{B\})
\end{array}
$$

This independence relation is an example of an independence relation with an induced dependency.

# Directed versus undirected graph language

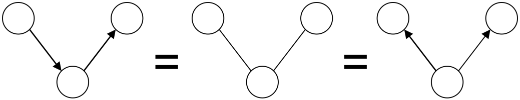
Arc directions encode the possibility of induced dependencies:

- a head-to-head node captures the possible occurrence of an "intercausal" interaction between its parents:



Note:

An acyclic(!) directed graph without head-to-head nodes encodes the same independences as its underlying undirected graph:



# Undirected graphs:
## separation and (in)dependence

The lectures and the examination focus on directed graphical models. However, the concept of separation in undirected graphs is easier than the concept of d-separation in directed graphs, so you may want to study undirected graphs first. Both course syllabus and these slides (with accompanying video) provide the necessary information and exercises for that.

Skip to directed graphs

## The separation criterion: introduction

**Definition**:
Let $G = (\boldsymbol{V}_G, \boldsymbol{E}_G)$ be an undirected graph with edges $\boldsymbol{E}_G$ and nodes $\boldsymbol{V}_G = \{V_1, \ldots, V_n\}$, $n > 1$.

Let $s$ be a path in $G$ from a node $V_i$ to a node $V_j$.

The path $s$ is blocked by a set of nodes $\boldsymbol{Z} \subseteq \boldsymbol{V}_G$, if at least one node from $\boldsymbol{Z}$ is *on* the path $s$.

If $s$ is not blocked by $\boldsymbol{Z}$, the path is called active given $\boldsymbol{Z}$.

# The separation criterion

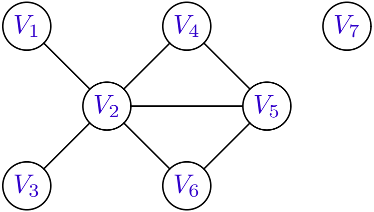
**Definition**:
Let $G = (\boldsymbol{V}_G, \boldsymbol{E}_G)$ be an undirected graph. Let $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z} \subseteq \boldsymbol{V}_G$ be sets of nodes in $G$.

The set $\boldsymbol{Z}$ separates the set $\boldsymbol{X}$ from $\boldsymbol{Y}$ in $G$— Notation: $\langle \boldsymbol{X} \mid \boldsymbol{Z} \mid \boldsymbol{Y} \rangle_G$— if *every simple* path in $G$ from a node in $\boldsymbol{X}$ to a node in $\boldsymbol{Y}$ is *blocked* by $\boldsymbol{Z}$.

**Remarks**:

- the above notion is known as the separation criterion for undirected graphs;
- if there is no path between the nodes $\boldsymbol{X}$ and $\boldsymbol{Y}$ in a graph $G$, then $\langle \boldsymbol{X} \mid \emptyset \mid \boldsymbol{Y} \rangle_G$.

# An example




Which of the following separation statements are valid?

a) $\langle \{V_1\} \mid \{V_2\} \mid \{V_3, V_6\} \rangle_G$

b) $\langle \{V_4\} \mid \{V_2, V_5\} \mid \{V_6\} \rangle_G$

c) $\langle \{V_4\} \mid \{V_1, V_2, V_5\} \mid \{V_6\} \rangle_G$

d) $\langle \{V_1\} \mid \{V_4\} \mid \{V_5\} \rangle_G$

e) $\langle \{V_1, V_5, V_6\} \mid \emptyset \mid \{V_7\} \rangle_G$

f) $\langle \{V_2\} \mid \{V_5\} \mid \{V_7\} \rangle_G$

g) $\langle \{V_1\} \mid \{V_5\} \mid \{V_2\} \rangle_G$

Answer: all except d) and g)

## Independence relations and undirected graphs

**Definition**: Let $I$ be an independence relation on a set of random variables $\boldsymbol{V}$. Let $G = (\boldsymbol{V}_G, \boldsymbol{E}_G)$ be an undirected graph with $\boldsymbol{V}_G = \boldsymbol{V}$.

- graph $G$ is called a dependency map (D-map) for $I$ if for all $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z} \subseteq \boldsymbol{V}$ we have:

  if $I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y})$ then $\langle \boldsymbol{X} \mid \boldsymbol{Z} \mid \boldsymbol{Y} \rangle_G$;

- graph $G$ is called an independency map (I-map) for $I$ if for all $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z} \subseteq \boldsymbol{V}$ we have:

  if $\langle \boldsymbol{X} \mid \boldsymbol{Z} \mid \boldsymbol{Y} \rangle_G$ then $I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y})$;

- graph $G$ is called a perfect map (P-map) for $I$ if $G$ is both a dependency map and an independency map for $I$.

## undirected D-maps: what can they tell?

Let $I$ be an independence relation and $G$ an undirected graph.

Consider a D-map for $I$, then

$V_1$ and $V_2$ neighbours[2] $\Longrightarrow$ $V_1$, $V_2$ dependent

$V_1$ and $V_2$ non-neighbours $\Longrightarrow$ ?? $V_1$ and $V_2$ can be:

$dependent$,

$independent$, $or$

$conditionally\ independent$

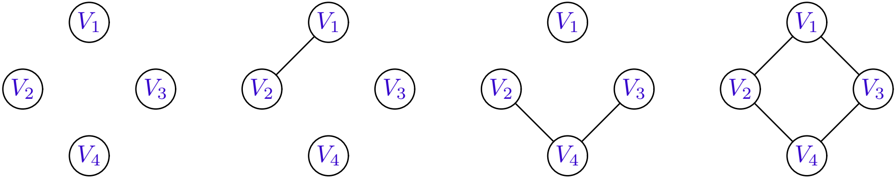
[2] i.e. directly connected by an edge

# An example

Consider the independence relation $I$ on $\boldsymbol{V} = \{V_1, \ldots, V_4\}$, defined by

$$I(\{V_1\}, \{V_2, V_3\}, \{V_4\}) \text{ and } I(\{V_2\}, \{V_1, V_4\}, \{V_3\})$$

Which of the following undirected graphs are examples of D-maps for $I$ ?




See Exercise 3.5 (Syllabus) Answer: all

## Undirected I-maps: what can they tell?

Let $I$ be an independence relation and $G$ an undirected graph.

Consider an I-map for $I$, then

$V_1$ and $V_2$ non-neighbours $\Longrightarrow$ $V_1$, $V_2$ (condit.) independent

$V_1$ and $V_2$ neighbours $\Longrightarrow$ ?? $V_1$ and $V_2$ can be

$dependent,$

$independent,\ or$

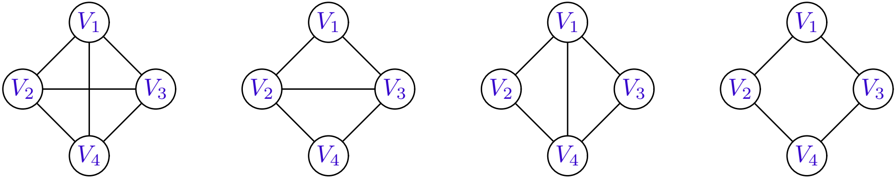
$conditionally\ independent$

# An example

Consider the independence relation $I$ on $\boldsymbol{V} = \{V_1, \ldots, V_4\}$, defined by

$$I(\{V_1\}, \{V_2, V_3\}, \{V_4\}) \text{ and } I(\{V_2\}, \{V_1, V_4\}, \{V_3\})$$

Which of the following undirected graphs are examples of I-maps for $I$ ?




(See Exercise 3.5) Answer: all

# Properties of $I$

Let $I$ be an independence relation on a set of random variables $\boldsymbol{V}$.

**Lemma**:
Every independence relation $I$ has an undirected D-map.

**Proof**:
The undirected graph $G = (\boldsymbol{V}, \emptyset)$ is a D-map for $I$. ■

**Lemma**:
Every independence relation $I$ has an undirected I-map.

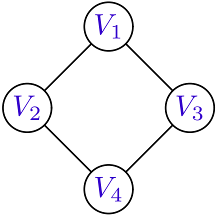
**Proof**:
The undirected graph $G' = (\boldsymbol{V}, \boldsymbol{V} \times \boldsymbol{V})$ is an I-map for $I$. ■

# An example

Consider the independence relation $I$ on $\boldsymbol{V} = \{V_1, \ldots, V_4\}$, defined by

$$I(\{V_1\}, \{V_2, V_3\}, \{V_4\}) \text{ and } I(\{V_2\}, \{V_1, V_4\}, \{V_3\})$$

The following undirected graph is a perfect map for $I$:




Is this P-map for $I$ unique ?

Does every $I$ have a P-map ?

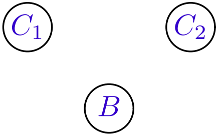
Answer: yes

Answer: no, see next slide

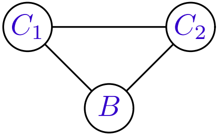
# An example

Reconstider the experiment with the two coins and the bell.

- the following graph is the only D-map for the independence relation $I$ of this experiment:




- the following graph is the only I-map for $I$:




- $I$ does not have a perfect map !

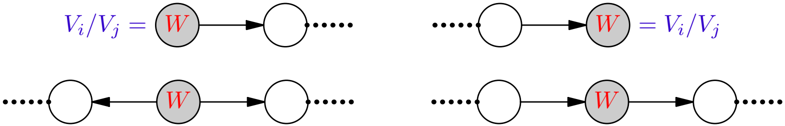
# Directed acyclic graphs:
## d-separation and (in)dependence

# The d-separation criterion: introduction

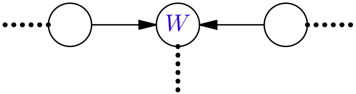
**Definition**: Let $G = (\boldsymbol{V}_G, \boldsymbol{A}_G)$ be an acyclic directed graph (DAG), and let $s$ be a chain in $G$ between $V_i$ and $V_j \in \boldsymbol{V}_G$.

Chain $s$ is blocked (or: in-active) by a set $\boldsymbol{Z} \subseteq \boldsymbol{V}_G$ if $s$ contains a node $W$ for which one of the following holds:

- $W \in \boldsymbol{Z}$ and $W$ has at most one incoming arc *on* chain $s$:

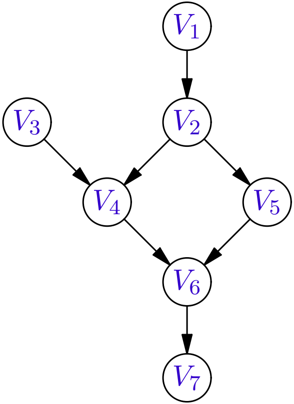



- $\boldsymbol{\sigma}^*(W) \cap \boldsymbol{Z} = \emptyset$ and $W$ has two incoming arcs *on* chain $s$:



# An example

Consider the following DAG and some of its chains:




1) $V_4, V_2, V_5$ from $V_4$ to $V_5$
2) $V_1, V_2, V_5, V_6, V_7$ from $V_1$ to $V_7$
3) $V_3, V_4, V_6, V_5$ from $V_3$ to $V_5$
4) $V_2, V_4$ from $V_2$ to $V_4$

Which chains are blocked by which of the following sets?

$\emptyset, \{V_2\}, \{V_5\}, \{V_2, V_5\},$
$\{V_4\}, \{V_6\}, \{V_4, V_6\}, \{V_7\}$

Answers: $\emptyset$:3; $\{V_2\}$: all; $\{V_5\}$: 1,2,3; $\{V_2, V_5\}$: all; $\{V_4\}$: 1,3,4; $\{V_6\}$: ~~none~~ 2; $\{V_4, V_6\}$: ~~1,3,4~~ all; $\{V_7\}$: 2

# The d-separation criterion

**Definition**:
Let $G = (V_G, A_G)$ be an acyclic directed graph. Let $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z} \subseteq \boldsymbol{V}_G$ be sets of nodes in $G$.

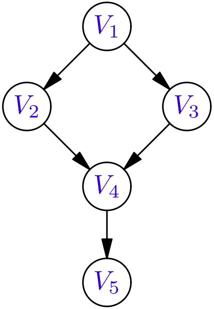
The set $\boldsymbol{Z}$ d-separates $\boldsymbol{X}$ from $\boldsymbol{Y}$ in $G$—notation: $\langle \boldsymbol{X} \mid \boldsymbol{Z} \mid \boldsymbol{Y} \rangle_G^d$ —if *every simple chain* in $G$ from a node in $\boldsymbol{X}$ to a node in $\boldsymbol{Y}$ is blocked by $\boldsymbol{Z}$.

**Remarks**:

- The above notion is known as the d-separation criterion;
- $\langle \boldsymbol{X} \mid \emptyset \mid \boldsymbol{Y} \rangle_G^d$ indicates that all chains between $\boldsymbol{X}$ and $\boldsymbol{Y}$, if any, contain a head-to-head node;
- if $\boldsymbol{X}$ and $\boldsymbol{Y}$ are not d-separated by $\boldsymbol{Z}$, we say that they are d-connected given $\boldsymbol{Z}$.

# An example

Consider the following DAG and d-separation statements:




a) $\langle \{V_1\} \mid \{V_2, V_3\} \mid \{V_5\} \rangle_G^d$

b) $\langle \{V_1\} \mid \{V_4\} \mid \{V_5\} \rangle_G^d$

c) $\langle \{V_2\} \mid \{V_1\} \mid \{V_3\} \rangle_G^d$

d) $\langle \{V_2\} \mid \{V_1, V_5\} \mid \{V_3\} \rangle_G^d$

e) $\langle \{V_2\} \mid \emptyset \mid \{V_3\} \rangle_G^d$

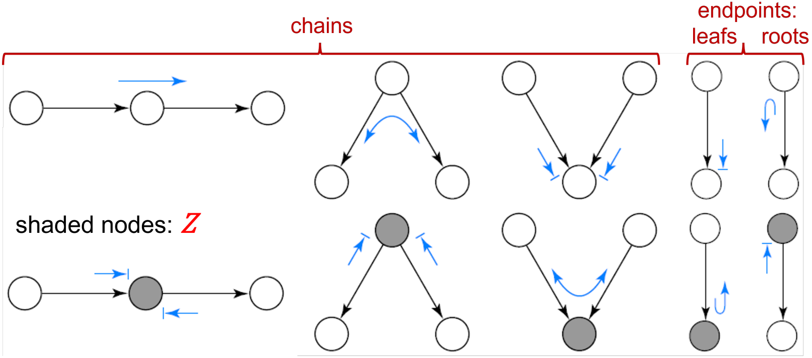
f) $\langle \{V_1\} \mid \{V_3, V_4\} \mid \{V_2\} \rangle_G^d$

Which d-separation statements are valid in the graph ?

Answer: a,b,c

# Bayes-Ball for determining d-separation

$\langle \boldsymbol{X} \mid \boldsymbol{Z} \mid \boldsymbol{Y} \rangle_G^d$ ? Drop bouncing balls at $\boldsymbol{X}$, which bounce from node to node along chains, using the 10 rules of Bayes-ball:




- balls quite a chain at a stop →|
- any node reached (visited) by a ball is on an active chain
- $\boldsymbol{Y}$ consists of all non-visited nodes

# Independence relations and directed graphs

**Definition**:
Let $I$ be an independence relation on a set of random variables $\boldsymbol{V}$.
Let $G = (\boldsymbol{V}_G, \boldsymbol{A}_G)$ be an acyclic directed graph with $\boldsymbol{V}_G = \boldsymbol{V}$.

- graph $G$ is called a dependency map ( D-map ) for $I$ if for every $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z} \subseteq \boldsymbol{V}$ we have that:

  if $I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y})$ then $\langle \boldsymbol{X} \mid \boldsymbol{Z} \mid \boldsymbol{Y} \rangle_G^d$;

- graph $G$ is called an independency map ( I-map ) for $I$ if for every $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z} \subseteq \boldsymbol{V}$ we have that:

  if $\langle \boldsymbol{X} \mid \boldsymbol{Z} \mid \boldsymbol{Y} \rangle_G^d$ then $I(\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{Y})$;

- graph $G$ is called a perfect map ( P-map ) for $I$ if $G$ is both a D-map and an I-map for $I$.

## Directed D-maps: what can they tell?

Let $I$ be an independence relation and $G$ a DAG.

Consider a D-map for $I$, then

$V_1$ and $V_2$ neighbours[3] $\Longrightarrow$ $V_1$, $V_2$ dependent

$V_1$ and $V_2$ non-neighbours $\Longrightarrow$ ?? $V_1$ and $V_2$ can be:

$dependent$,

$independent$,

$conditionally\ dependent,\ or$

$conditionally\ independent$

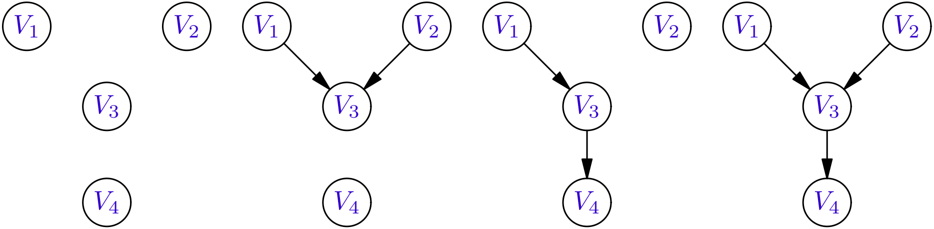
[3] i.e. directly connected by an arc

# An example

Consider the independence relation $I$ on $\boldsymbol{V} = \{V_1, \ldots, V_4\}$ defined by

$$I(\{V_1\}, \emptyset, \{V_2\}) \text{ and } I(\{V_1, V_2\}, \{V_3\}, \{V_4\})$$

Which of the following DAGs are D-maps for $I$ ?




See Exercise 3.8 (Syllabus) Answer: all

## Directed I-maps: what can they tell ?

Let $I$ be an independence relation and $G$ a DAG.

Consider an I-map for $I$, then

$V_1$ and $V_2$ non-neighbours $\Longrightarrow$ $V_1$, $V_2$ (cond.) independent,
or possibly(!) induced:
conditionally dependent

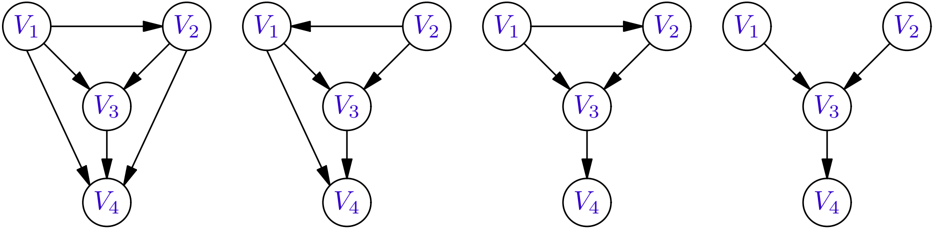
$V_1$ and $V_2$ neighbours $\Longrightarrow$ ?? $V_1$ and $V_2$ can be:
$dependent,$
$independent,$
$conditionally\ dependent,\ or$
$conditionally\ independent$

# An example

Consider the independence relation $I$ on $\boldsymbol{V} = \{V_1, \ldots, V_4\}$ defined by

$$I(\{V_1\}, \emptyset, \{V_2\}) \text{ and } I(\{V_1, V_2\}, \{V_3\}, \{V_4\})$$

Which of the following DAGs are I-maps for $I$ ?



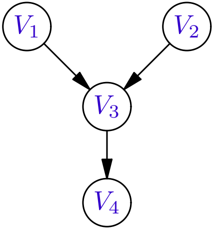

(See Exercise 3.8) Answer: all

# An example

Consider the independence relation $I$ on $\boldsymbol{V} = \{V_1, \ldots, V_4\}$ defined by

$$I(\{V_1\}, \emptyset, \{V_2\}) \text{ and } I(\{V_1, V_2\}, \{V_3\}, \{V_4\})$$

The following DAG is a perfect map for $I$:



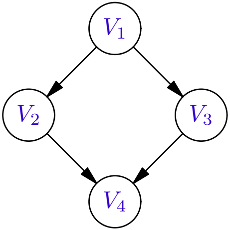

Is this P-map for $I$ unique ? yes

## An example

Consider the independence relation $I$ on $\boldsymbol{V} = \{V_1, \ldots, V_4\}$ defined by

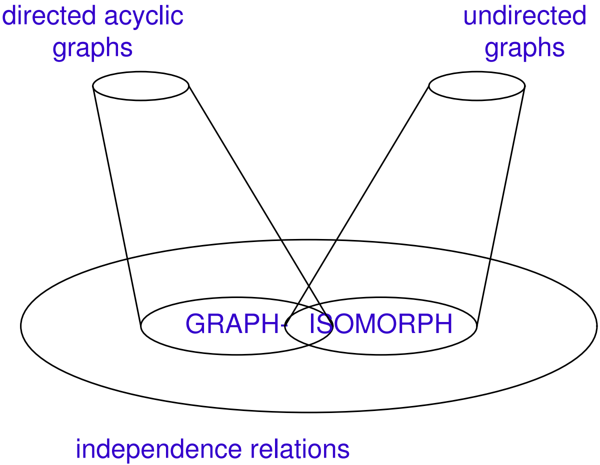
$$I(\{V_1\}, \{V_2, V_3\}, \{V_4\}) \text{ and } I(\{V_2\}, \{V_1, V_4\}, \{V_3\})$$

The relation $I$ has an undirected P-map, but does not have a directed P-map. Consider for example the following DAG $G$:




In graph $G$ we have that $\langle \{V_1\} \mid \{V_2, V_3\} \mid \{V_4\} \rangle_G^d$, but also that $\langle \{V_2\} \mid \{V_1\} \mid \{V_3\} \rangle_G^d$ !

# Independence relations and their graphical representation



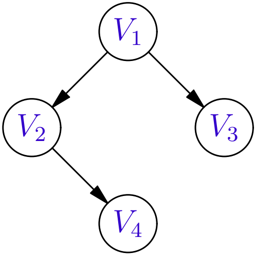

(Graph-isomorph: independence relation with perfect map.)

# An I-map or a D-map ?

Reconsider the independence relation $I$ on $\boldsymbol{V} = \{V_1, \ldots, V_4\}$ defined by

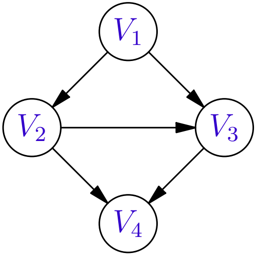
$$I(\{V_1\}, \{V_2, V_3\}, \{V_4\}) \text{ and } I(\{V_2\}, \{V_1, V_4\}, \{V_3\})$$

Compare the following two representations of independence relation $I$:




a D-map

and




an I-map

## A minimal I-map

**Definition**: Let $I$ be an independence relation on a set of random variables $\boldsymbol{V}$. Let $G = (\boldsymbol{V}_G, \boldsymbol{A}_G)$ be a graph with $\boldsymbol{V}_G = \boldsymbol{V}$.

The graph $G$ is called a minimal I-map for $I$ if the following conditions hold:

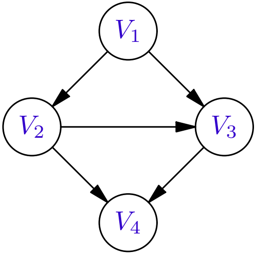
- $G$ is an I-map for $I$, and
- no proper subgraph of $G$ is an I-map for $I$.

## An example

Consider the independence relation $I$ on $\boldsymbol{V} = \{V_1, \ldots, V_4\}$ defined by

$$I(\{V_1\}, \{V_2, V_3\}, \{V_4\}) \text{ and } I(\{V_2\}, \{V_1, V_4\}, \{V_3\})$$

The following DAG is a minimal I-map for $I$:




Is this minimal I-map for $I$ unique ? no

# Directed or undirected ? (I)

Directed and undirected I-maps are related.

**Definition**: The moral graph of a DAG $G = (\boldsymbol{V}_G, \boldsymbol{A}_G)$ is the undirected graph obtained as follows:

- for each $V_k \in \boldsymbol{V}_G$ add an edge between each pair of unconnected parents $V_i, V_j \in \boldsymbol{\rho}_G(V_k)$;
- drop the directions of all arcs.

**Definition**: A graph is triangulated or chordal if any loop of length $\geq 4$ contains a shortcut.

**Proposition**: Let $I$ be an independence relation over $\boldsymbol{V}$. Consider graphs $G = (\boldsymbol{V}_G, \boldsymbol{A}_G)$ and $G' = (\boldsymbol{V}, \boldsymbol{E}_{G'})$. Then,

$$G \text{ is an I-map for } I \quad \underset{\text{triangulation+add direction}}{\overset{\text{moralisation+drop direction}}{\begin{array}{c}\Longrightarrow \\ \Longleftarrow\end{array}}} \quad G' \text{ is an I-map for } I$$

## Directed or undirected ? (II)

Consider the following properties (partly proven later):

- Let $G$ be a directed acyclic graph. Then $G$ is a directed I-map of $I_{\Pr} \iff \Pr$ can be written as

$$\Pr(\boldsymbol{V}) = \prod_{V_i} \Pr(V_i \mid \boldsymbol{\rho}_G(V_i))$$

- Let $G$ be an undirected graph. Then $G$ is an undirected I-map of $I_{\Pr} \iff$[4] $\Pr$ can be written as

$$\Pr(\boldsymbol{V}) = K \cdot \prod_{\boldsymbol{C}_i} \Phi(\boldsymbol{C}_i) \quad \longleftarrow$$

what's the meaning of these clique potentials?!?

for some normalisation factor $K$.

[4] $\Longrightarrow$ requires $\Pr$ to be strictly positive