Chapter 2:

Preliminaries
Random variables

Let $V = \{V_1, \ldots, V_n\}$, $n \geq 1$, be a set of random variables. Each variable $V_i \in V$ can take on one of $m \geq 2$ values; for now we consider 2-valued variables:

- $V_i = true$, denoted by $v_i$;
- $V_i = false$, denoted by $\neg v_i$ (or by $\overline{v_i}$).

The set $V$ spans a Boolean Algebra of logical propositions $\mathcal{V}$:

- $T(\text{rue}), F(\text{alse}) \in \mathcal{V}$;
- for all variables $V_i \in V$ we have that $v_i \in \mathcal{V}$;
- for all $x \in \mathcal{V}$ we have that $\neg x \in \mathcal{V}$;
- for all $x, y \in \mathcal{V}$ we have that $x \land y \in \mathcal{V}$ and $x \lor y \in \mathcal{V}$.

The elements of $\mathcal{V}$ obey the usual rules of propositional logic.
The joint probability distribution

Definition:

Let $\mathcal{V}$ be the Boolean Algebra of propositions spanned by a set of random variables $\mathcal{V}$. Let $\Pr : \mathcal{V} \to [0, 1]$ be a function such that

- Pr is **positive**: for each $x \in \mathcal{V}$ we have that $\Pr(x) \geq 0$ and, more specifically, $\Pr(\mathsf{F}) = 0$;
- Pr is **normed**: $\Pr(\mathsf{T}) = 1$;
- Pr is **additive**: we have, for each $x, y \in \mathcal{V}$ with $x \land y \equiv \mathsf{F}$, that $\Pr(x \lor y) = \Pr(x) + \Pr(y)$.

The function $\Pr$ is a joint probability distribution on $\mathcal{V}$; the function value $\Pr(x)$ is the probability of $x$. 

Independence of propositions

**Definition:** Let $\mathcal{V}$ be the Boolean Algebra of propositions spanned by a set of random variables $V$. Let $\Pr$ be a joint probability distribution on $V$.

Two propositions $x, y \in \mathcal{V}$ are called **independent** in $\Pr$ if

$$\Pr(x \wedge y) = \Pr(x) \cdot \Pr(y)$$

The propositions $x, y \in \mathcal{V}$ are called **conditionally independent** given the proposition $z \in \mathcal{V}$ if we have that

$$\Pr(x \wedge y | z) = \Pr(x | z) \cdot \Pr(y | z)$$
The two notions of independence (1)

• Consider two propositions $x, y \in \mathcal{V}$ such that $x$ and $y$ are independent $^1$:

Can $z \in \mathcal{V}$ exist such that $x$ and $y$ are dependent given $z$?

• Yes:

$^1$The square has area 1, representing the total probability mass.
The two notions of independence (2)

- Consider two propositions $x, y \in \mathcal{V}$ such that $x$ and $y$ are dependent:

  \[ x \quad y \]

  Can $z \in \mathcal{V}$ exist such that $x$ and $y$ are conditionally independent given $z$?

- Yes:

  \[ x \quad y \quad z \]
Configurations

Let $V$ be a set of random variables and let $W \subseteq V$.

- a configuration $c_W$ of $W$ is a conjunction of value assignments to the variables from $W$;
- convention: $c_\emptyset = T$;
- $w$ is used to denote a specific configuration of $W$.
- $W$ also indicates all possible configurations to the set $W$ (notation abuse!): $W$ is then considered to be a template that can be filled in with any configuration $c_W$.

Example: Let $W = \{V_1, V_3, V_7\}$. $W = V_1 \land V_3 \land V_7$ denotes a configuration template: filling in values for $V_i$ results in proper propositions/configurations. Some configurations $c_W$ of $W$ are:

\[
V_1 = true \land V_3 = true \land V_7 = false \\
v_1 \land \neg v_3 \land v_7 \\
\neg v_1 \land v_3 \land \neg v_7
\]
Conventions and notation

In the remainder of this course, for distributions on $\Pr(V)$:

- rather than talking about propositions $x \in \mathcal{V}$ spanned by $\mathcal{V}$
- we refer to configurations $c_{\mathcal{V}}$ of $\mathcal{V}$

<table>
<thead>
<tr>
<th>Variables/templates (capital)</th>
<th>Set (bold faced)</th>
<th>Singleton</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{V}$</td>
<td>$\mathcal{V}$</td>
<td>$\mathcal{V}$</td>
</tr>
<tr>
<td>$c_{\mathcal{V}}, v$</td>
<td>$c_{\mathcal{V}}, v$</td>
<td></td>
</tr>
</tbody>
</table>

- conjunctions are often left implicit: e.g. $v_1 \land v_2$ denotes $v_1 \land v_2$;
- note the following differences (!)

probabilities: $\Pr(c_{\mathcal{V}}), \Pr(c_{\mathcal{V}}), \Pr(v), \Pr(v), \Pr(v \mid c_E)$
distributions: $\Pr(\mathcal{V}), \Pr(\mathcal{V}), \Pr(\mathcal{V} \mid e)$
distribution sets: $\Pr(\mathcal{V} \mid \mathcal{E}), \Pr(\mathcal{V} \mid \mathcal{E})$
Independence of variables

**Definition:** Let $V$ be a set of random variables and let $X, Y, Z \subseteq V$. Let $\Pr$ be a joint distribution on $V$.

The set of variables $X$ is called **conditionally independent of** the set $Y$ given the set $Z$ in $\Pr$, if we have that

$$\Pr(X | Y \land Z) = \Pr(X | Z)$$

**Remarks:**

- the expression $\Pr(X | Y \land Z) = \Pr(X | Z)$ represents that $\Pr(c_X | c_Y \land c_Z) = \Pr(c_X | c_Z)$ holds for all configurations $c_X$, $c_Y$ and $c_Z$ of $X$, $Y$ and $Z$;
- $\Pr(X | Y \land Z) = \Pr(X | Z) \Rightarrow 
  \Pr(X \land Y | Z) = \Pr(X | Z) \cdot \Pr(Y | Z)$ (what about $\Leftarrow$).
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 $V$ be a set of random variables and let $\Pr$ be a joint probability distribution on $V$.

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

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

Remarks:

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


Properties of $I_{Pr}$: symmetry

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

**Proof:**

\[ I_{Pr}(X, Z, Y) \iff \Pr(X \mid Y \land Z) = \Pr(X \mid Z) \]

\[ \iff \frac{\Pr(X \land Y \land Z)}{\Pr(Y \land Z)} = \frac{\Pr(X \land Z)}{\Pr(Z)} \]

\[ \iff \frac{\Pr(X \land Y \land Z)}{\Pr(X \land Z)} = \frac{\Pr(Y \land Z)}{\Pr(Z)} \]

\[ \iff \Pr(Y \mid X \land Z) = \Pr(Y \mid Z) \]

\[ \iff I_{Pr}(Y, Z, X) \]
Properties of $I_{Pr}$: decomposition

Lemma: $I_{Pr}(X, Z, Y \cup W) \Rightarrow I_{Pr}(X, Z, Y) \land I_{Pr}(X, Z, W)$

Proof: (sketch) (Note: $c_{Y \cup W} = c_Y \land c_W$) Suppose that

$Pr(X \mid Y \land W \land Z) = Pr(X \mid Z)$. Then, by definition,

$$Pr(X \land Y \land W \land Z) = Pr(Y \land W \land Z) \cdot \frac{Pr(X \land Z)}{Pr(Z)}$$

For $Pr(X \mid Y \land Z)$ we find that

$$Pr(X \mid Y \land Z) = \frac{Pr(X \land Y \land Z)}{Pr(Y \land Z)}$$

$$= \sum_{c_W} \frac{Pr(X \land Y \land Z \land c_W)}{Pr(Y \land Z)}$$

$$= \frac{Pr(X \land Z)}{Pr(Z)} = Pr(X \mid Z) \quad \blacksquare$$
Properties of $I_{Pr}$: weak union, contraction

Lemma:

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

Proof: left as exercise 3.1.

What about $\Leftarrow$?
The definition of the independence relation

Joint Distribution $Pr$

Independence relation $I_{Pr}$

Properties: symmetry, decomposition, weak union, contraction

Independence relation $I$

Axioms: symmetry, decomposition, weak union, contraction
The (qualitative) independence relation $I$

**Definition:**
Let $V$ be a set of random variables and let $X, Y, Z, W \subseteq V$.

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

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

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.
Lemma:
Let $I$ be an independence relation on a set of random variables $V$. We have that

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

for all $X, Y, Z, W \subseteq V$.

Proof:
We observe that

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

$$\implies I(W, Y \cup Z, X) \implies_{\text{symm}} I(X, Y \cup Z, W)$$

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

\[\blacksquare\]
Representing independences

Different ways exist of representing an independence relation:

• all independence statements of the relation are explicitly stated;
• only the independence statements of a suitable subset of the relation are explicitly stated — all other statements are implicitly represented by means of the axioms;
• the independence relation is coded in a graph;
• . . .
An example

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

\[
\begin{align*}
I(\{V_1\}, \emptyset, \{V_4\}) & \quad I(\{V_2\}, \emptyset, \{V_4\}) & \quad I(\{V_4\}, \emptyset, \{V_1\}) \\
I(\{V_2\}, \emptyset, \{V_4\}) & \quad I(\{V_1, V_4\}, \emptyset, \{V_2\}) & \quad I(\{V_4\}, \{V_1\}, \{V_2\}) \\
I(\{V_3\}, \emptyset, \{V_4\}) & \quad I(\{V_2, V_4\}, \emptyset, \{V_1\}) & \quad I(\{V_4\}, \{V_1\}, \{V_3\}) \\
I(\{V_4\}, \emptyset, \{V_1\}) & \quad I(\{V_2\}, \emptyset, \{V_1, V_4\}) & \quad I(\{V_4\}, \{V_1\}, \{V_2, V_3\}) \\
I(\{V_4\}, \emptyset, \{V_2\}) & \quad I(\{V_1\}, \emptyset, \{V_2, V_4\}) & \quad I(\{V_1\}, \{V_2\}, \{V_4\}) \\
I(\{V_4\}, \emptyset, \{V_3\}) & \quad I(\{V_2\}, \{V_1\}, \{V_4\}) & \quad I(\{V_3\}, \{V_2\}, \{V_4\}) \\
I(\{V_4\}, \emptyset, \{V_3\}) & \quad I(\{V_2\}, \{V_1\}, \{V_4\}) & \quad I(\{V_4\}, \{V_2\}, \{V_1\}) \\
I(\{V_1, V_2\}, \emptyset, \{V_4\}) & \quad I(\{V_3\}, \{V_1\}, \{V_4\}) & \quad I(\{V_4\}, \{V_2\}, \{V_1\}) \\
I(\{V_1, V_3\}, \emptyset, \{V_4\}) & \quad I(\{V_2, V_3\}, \{V_1\}, \{V_4\}) & \quad I(\{V_4\}, \{V_2\}, \{V_3\}) \\
I(\{V_2, V_3\}, \emptyset, \{V_4\}) & \quad I(\{V_4\}, \{V_1, V_2\}, \{V_3\}) & \quad I(\{V_4\}, \{V_2\}, \{V_1, V_3\}) \\
I(\{V_4\}, \emptyset, \{V_1, V_2\}) & \quad I(\{V_2\}, \{V_1, V_3\}, \{V_4\}) & \quad I(\{V_1\}, \{V_3\}, \{V_4\}) \\
I(\{V_4\}, \emptyset, \{V_1, V_3\}) & \quad I(\{V_4\}, \{V_1, V_3\}, \{V_2\}) & \quad I(\{V_2\}, \{V_3\}, \{V_4\}) \\
I(\{V_4\}, \emptyset, \{V_2, V_3\}) & \quad I(\{V_1\}, \{V_2, V_3\}, \{V_4\}) & \quad I(\{V_1, V_2\}, \{V_3\}, \{V_4\}) \\
I(\{V_4\}, \emptyset, \{V_2, V_3\}) & \quad I(\{V_4\}, \{V_2, V_3\}, \{V_1\}) & \quad I(\{V_1\}, \{V_4\}, \{V_2\}) \\
I(\{V_4\}, \emptyset, \{V_2, V_3\}) & \quad I(\{V_4\}, \{V_3\}, \{V_1, V_2\}) & \quad I(\{V_2\}, \{V_4\}, \{V_1\}) \\
I(\{V_4\}, \emptyset, \{V_1, V_2, V_3\}) & \quad I(\{V_4\}, \{V_3\}, \{V_1\}) & \quad I(\{V_3\}, \{V_1, V_2\}, \{V_4\}) \\
I(\{V_4\}, \emptyset, \{V_1, V_2, V_3\}) & \quad I(\{V_4\}, \{V_3\}, \{V_1\}) & \quad I(\{V_3\}, \{V_1, V_2\}, \{V_4\})
\end{align*}
\]
The representation of an independence relation in an undirected graph

Consider an independence relation $I$ and an undirected graph:

The global idea is:

- represent each variable $V_i$ by a node $V_i$ in the graph, and v.v.;
- code the independence statements of $I$ by means of missing edges.
The separation criterion: introduction

Definition:
Let $G = (V_G, E_G)$ be an undirected graph with edges $E_G$ and nodes $V_G = \{V_1, \ldots, V_n\}$, $n > 1$.

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 $Z \subseteq V_G$, if at least one node from $Z$ is on the path $s$.

If $s$ is not blocked by $Z$, the path is called active given $Z$. 
The separation criterion

Definition:
Let $G = (V_G, E_G)$ be an undirected graph. Let $X, Y, Z \subseteq V_G$ be sets of nodes in $G$.

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

Remarks:

- the above notion is known as the separation criterion for undirected graphs;
- if there is no path between the nodes $X$ and $Y$ in a graph $G$, then $\langle X \mid \emptyset \mid Y \rangle_G$. 
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)  $\langle \{V_1\} \mid \{V_5\} \mid \{V_2\} \rangle_G$
Independence relations and undirected graphs

Definition: Let $I$ be an independence relation on a set of random variables $V$. Let $G = (V_G, E_G)$ be an undirected graph with $V_G = V$.

- **graph $G$ is called a dependency map** (D-map) for $I$ if for all $X, Y, Z \subseteq V$ we have:
  
  \[ \text{if } I(X, Z, Y) \text{ then } \langle X \mid Z \mid Y \rangle_G; \]

- **graph $G$ is called an independency map** (I-map) for $I$ if for all $X, Y, Z \subseteq V$ we have:
  
  \[ \text{if } \langle X \mid Z \mid Y \rangle_G \text{ then } I(X, Z, Y); \]

- **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 do 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 $\implies V_1, V_2$ dependent

$\neg \langle \{V_1\} \mid Z \mid \{V_2\} \rangle_G$ $\neg I(\{V_1\}, Z, \{V_2\})$

$V_1$ and $V_2$ non-neighbours $\implies$ ??

$\langle \{V_1\} \mid Z \mid \{V_2\} \rangle_G$ $\text{dependent}$

$\text{independent}$

$\text{conditionally independent}$

Note: statements hold for all $Z \subseteq V_G \setminus (\{V_1\} \cup \{V_2\})$!
An example

Consider the independence relation $I$ on $V = \{V_1, \ldots, V_4\}$, 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 $I$?
Undirected I-maps: what do 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 $\implies V_1, V_2$ (cond.) independent

\[ \langle \{V_1\} \mid Z \mid \{V_2\} \rangle_G \]

$I(\{V_1\}, Z, \{V_2\})$

$V_1$ and $V_2$ neighbours $\implies ??$

\[ \neg \langle \{V_1\} \mid Z \mid \{V_2\} \rangle_G \]

dependent

independent

conditionally independent

Note: statements hold for all $Z \subseteq V_G \setminus (\{V_1\} \cup \{V_2\})$!
An example

Consider the independence relation $I$ on $V = \{V_1, \ldots, V_4\}$, 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 $I$?
Properties of $I$

Let $I$ be an independence relation on a set of random variables $V$.

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

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

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

**Proof:**
The undirected graph $G' = (V, V \times V)$ is an I-map for $I$. ■
Consider the independence relation $I$ on $\mathbf{V} = \{V_1, \ldots, V_4\}$, 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 $I$: 

![Graph with nodes $V_1$, $V_2$, $V_3$, $V_4$ connected in a square]

Is this P-map for $I$ unique? Does every $I$ have a P-map?
An example

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

\[
I(\{C_1\}, \emptyset, \{C_2\}) \quad \neg I(\{C_1\}, \{B\}, \{C_2\})
\]
\[
I(\{C_1\}, \emptyset, \{B\}) \quad \neg I(\{C_1\}, \{C_2\}, \{B\})
\]
\[
I(\{C_2\}, \emptyset, \{B\}) \quad \neg I(\{C_2\}, \{C_1\}, \{B\})
\]

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

Reconsider the experiment with the two coins and the bell.

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

• the following graph is an I-map for $I$:

• Does $I$ have a perfect map?
The representation of an independence relation in a directed graph

Consider an independence relation $I$ and a directed graph $G$:

The global idea is:

- represent each variable $V_i$ of $I$ by a node $V_i$ in $G$, and v.v.;
- code the independence statements of $I$ by means of missing arcs in the graph;
- use the direction of the arcs to represent induced dependencies.
The formalism of the directed graph is more expressive than the formalism of the undirected graph:
Causality?

Consider the following examples:

- length → age → reading
- weather → harvest → grain price
- burglar → alarm → earthquake
Introduction, continued

We aim to represent the following (in)dependences with directed graphs:

- \( I(\{V_2\}, \emptyset, \{V_3\}) \) and \( \neg I(\{V_2\}, \{V_1\}, \{V_3\}) \):

- \( I(\{V_2\}, \{V_1\}, \{V_3\}) \) and \( \neg I(\{V_2\}, \emptyset, \{V_3\}) \):

- \( I(\{V_2\}, \{V_1\}, \{V_3\}) \) and \( \neg I(\{V_2\}, \emptyset, \{V_3\}) \):
The d-separation criterion: introduction

**Definition:** Let \( G = (V_G, A_G) \) be an acyclic directed graph (DAG), and let \( s \) be a chain in \( G \) between \( V_i \) and \( V_j \in V_G \).

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

- \( W \in Z \) and \( W \) has at most one incoming arc on chain \( s \):
  \[
  V_i/V_j = \begin{array}{c}
  \circlearrowleft \\
  \circlearrowright \\
  V_i/V_j = W
  \end{array}
  \]
- \( \sigma^*(W) \cap Z = \emptyset \) and \( W \) has two incoming arcs on chain \( s \):
  \[
  \begin{array}{c}
  \circlearrowleft \\
  \circlearrowright \\
  W
  \end{array}
  \]
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 of these chains is blocked by which of the following sets?

$\{V_2\}, \{V_5\}, \{V_2, V_5\}, \{V_4\}, \{V_6\}, \{V_4, V_6\}$
The d-separation criterion

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

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

**Remarks:**

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

Consider the following DAG and d-separation statements:

\[ a) \quad \langle \{V_1\} \mid \{V_2, V_3\} \mid \{V_5\}\rangle^d_G \]
\[ b) \quad \langle \{V_1\} \mid \{V_4\} \mid \{V_5\}\rangle^d_G \]
\[ c) \quad \langle \{V_2\} \mid \{V_1\} \mid \{V_3\}\rangle^d_G \]
\[ d) \quad \langle \{V_2\} \mid \{V_1, V_5\} \mid \{V_3\}\rangle^d_G \]
\[ e) \quad \langle \{V_2\} \mid \emptyset \mid \{V_3\}\rangle^d_G \]
\[ f) \quad \langle \{V_1\} \mid \{V_3, V_4\} \mid \{V_2\}\rangle^d_G \]

Which d-separation statements are valid in the graph?
Bayes-Ball for determining d-separation

Determine if \( \langle X \mid Z \mid Y \rangle^d_G \) by dropping bouncing balls at \( X \) and following the 10 rules of Bayes-ball:

- \( Z \) is shaded
- a chain is active until a ball travelling along it meets a stop
- any node visited by a Bayes ball cannot be in \( Y \)
Independence relations and directed graphs

**Definition:**
Let $I$ be an independence relation on a set of random variables $V$. Let $G = (V_G, A_G)$ be an acyclic directed graph with $V_G = V$.

- the graph $G$ is called a (directed) dependency map (D-map) for $I$ if for every $X, Y, Z \subseteq V$ we have that:
  
  if $I(X, Z, Y)$ then $\langle X | Z | Y \rangle^d_G$;

- the graph $G$ is called a (directed) independency map (I-map) for $I$ if for every $X, Y, Z \subseteq V$ we have that:
  
  if $\langle X | Z | Y \rangle^d_G$ then $I(X, Z, Y)$;

- the graph $G$ is called a (directed) perfect map (P-map) for $I$ if $G$ is both a dependency map and an independency map for $I$. 
Directed D-maps: what do 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 $\implies V_1, V_2$ dependent

$\neg \langle \{V_1\} \mid Z \mid \{V_2\} \rangle_G \implies \neg I(\{V_1\}, Z, \{V_2\})$

$V_1$ and $V_2$ non-neighbours $\implies$ ??

$\langle \{V_1\} \mid Z \mid \{V_2\} \rangle_G$

dependent

independent

conditionally dependent ($Z = \emptyset$)

conditionally independent ($Z \neq \emptyset$)

Note: statements hold for all $Z \subseteq V_G \setminus (\{V_1\} \cup \{V_2\})$!
An example

Consider the independence relation $I$ on $V = \{V_1, \ldots, V_4\}$ 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 $I$?
Directed I-maps

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

Consider an I-map for $I$, then

$V_1$ and $V_2$ non-neighbours $\implies V_1, V_2$ (cond.) independent, or cond. dependent (= induced)

$$\langle \{V_1\} \mid Z \mid \{V_2\} \rangle_G \quad I(\{V_1\}, Z, \{V_2\})$$

$V_1$ and $V_2$ neighbours $\implies ??$

$$\neg \langle \{V_1\} \mid Z \mid \{V_2\} \rangle_G \quad \text{dependent}
\text{independent}
\text{conditionally dependent}
\text{conditionally independent}$$

Note: statements hold for all $Z \subseteq V_G \setminus (\{V_1\} \cup \{V_2\})$!
Consider the independence relation $I$ on $V = \{V_1, \ldots, V_4\}$ 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 $I$?
Consider the independence relation $I$ on $V = \{V_1, \ldots, V_4\}$ 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?
An example

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

\[
I(\{V_1\}, \{V_2, V_3\}, \{V_4\}) \quad \text{and} \quad I(\{V_2\}, \{V_1, V_4\}, \{V_3\})
\]

The relation $I$ does not have a directed perfect map. Consider for example the following DAG $G$:

In graph $G$ we have that $\langle \{V_1\} \mid \{V_2, V_3\} \mid \{V_4\} \rangle^d_G$, but also that $\langle \{V_2\} \mid \{V_1\} \mid \{V_3\} \rangle^d_G$!
Independence relations and their graphical representation

Directed acyclic graphs

Undirected graphs

Independence relations

(Graph-isomorph: independence relation with perfect map.)
An I-map or a D-map?

Reconsider the independence relation $I$ on $V = \{V_1, \ldots, V_4\}$ 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
Recall what we were looking for...

- Compact representation of independence relation of $\Pr$;
- Factorise joint more efficiently than with chain rule $\rightarrow$ store (conditional) distributions involving less variables:

\[
\Pr(V) = \Pr(V_n | V_{n-1} \land \ldots \land V_1) \cdot \ldots \cdot \Pr(V_2 | V_1) \cdot \Pr(V_1)
\] (chain rule)

\[
= \ldots
\]

\[
= \ldots
\]

\[
= \Pr(V_n) \cdot \ldots \cdot \Pr(V_2) \cdot \Pr(V_1)
\] (assuming mutual independence among all $V_i$)

- $\Pr(X \land Y) = \Pr(X) \cdot \Pr(Y)$ is mathematically correct only if $X$ is truly independent of $Y$
A minimal I-map

Definition: Let $I$ be an independence relation on a set of random variables $V$. Let $G = (V_G, A_G)$ be a graph with $V_G = V$.

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 $V = \{V_1, \ldots, V_4\}$ 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?
Directed and undirected I-maps are related.

**Definition:** The moral graph of a DAG $G = (V_G, A_G)$ is the undirected graph obtained as follows:

- for each $V_k \in V_G$ add an edge between each pair of unconnected parents $V_i, V_j \in \rho_G(V_k)$;
- 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 $V$. Consider graphs $G = (V_G, A_G)$ and $G' = (V, E_{G'})$. Then,

\[
G \text{ is an I-map for } I \quad \iff \quad \text{moralisation+drop direction} \quad \iff \quad G' \text{ is an I-map for } I \quad \iff \quad \text{triangulation+add direction}
\]
Directed or undirected? (II)

Consider independence relation $I_{Pr}$ over $V$ and graph $G$ with $V = V_G$. Consider the following properties (partly proven later):

- Let $G$ be a DAG. Then $G$ is a minimal directed $I$-map of $I_{Pr}$ if and only if $Pr$ factorises as
  \[ Pr(V) = \prod_{V_i} Pr(V_i \mid \rho_G(V_i)) \]

- Let $G$ be an undirected graph. Then $G$ is an undirected $I$-map of $I_{Pr}$ if and only if $Pr$ can be written as
  \[ Pr(V) = K \cdot \prod_{C_i} \Phi(C_i) \]
  what’s the meaning of these clique potentials?!?
  for some normalisation factor $K$. 