Uncertainty in Artificial Intelligence - Cheat sheet

Dieter Castel & Pierre Carbonnelle

August 19, 2020

Original google docs by P. Carbonelle to be found here:
`https://docs.google.com/presentation/d/1sP-PJmo-pW4epfLQs3zoveZGceqw59pPeNAY9OM6dl8/edit#slide=id.p`

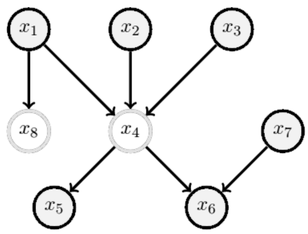
1 General Probability

Formula	Comment
$p(x,y,z) = p(x \wedge y \wedge z) = p(x,z\|y)p(y)$	Joint Probability Distribution (JPD)
$p(x \vee y) = p(x) + p(y) - p(x \wedge y)$	Disjunction for probabilities
$p(x\|y) = \frac{p(x,y)}{p(y)} = \frac{p(y\|x)p(x)}{p(y)}$	Definition of Conditional Probablity
$p(\neg x) = 1 - p(x)$	only valid for probability distributions i.e. normalized!
$p(x) = \sum_{x,z} p(x,y,z)$	marginalisation
$\sum_{y,z} p(x\|y,z) = 1$	marginalisation of CPD sums to one.
$size(p(\hat{x},\hat{y},\hat{z})) = \#dom(\hat{x}) * \#dom(\hat{y}) * \#dom(\hat{z})$	**without** independence assumptions
$p(x_1,\ldots,x_n) = p(x_1)p(x_2\|x_1)p(x_3\|x_2,x_1)\ldots p(x_k\|x_{k+1},\ldots,x_n)$	*Chainrule*

2 Distributional Independence

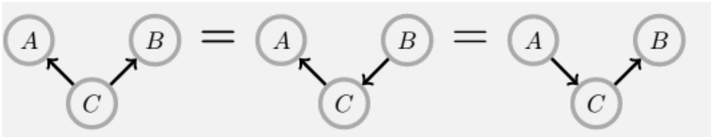
Formula	Comment
$X \perp\!\!\!\perp Y \iff \forall x \in X, y \in Y : p(x,y) = p(x)p(y)$	marginal distributional independence for variable sets X,Y
$X \perp\!\!\!\perp Y \iff \forall x \in X, y \in Y : p(x\|y) = p(x)$	marginal distributional independence with CPD def.
$X \perp\!\!\!\perp Y\|Z \iff \forall x \in X, y \in Y : p(x\|y,z) = p(x\|z)$	conditional distributional independence.

Below is the markov blanket of x_4: parents, childeren and parents of its childeren.

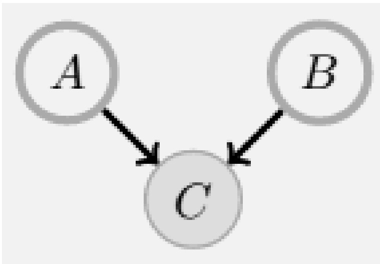
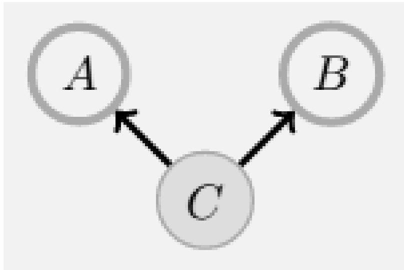


3 Graphical Independence/d-separation

In this section $\perp\!\!\!\perp$ and the like only mean GRAPHICAL independence and beware:
Graphical independence (== d-separation) $\Rightarrow$ distributional independence.
Graphical dependence (== d-connected) $\not\Rightarrow$ distributional independence.



These non-colliders are equivalent from an independence point of view.

colliders in BN	**non-colliders** in BN
	
$A \not\perp\!\!\!\perp B\|C$ or $A \top\!\!\!\top B\|C$	$A \perp\!\!\!\perp B\|C$ or $A \not\top\!\!\!\top B\|C$

3.1 Path-blocking

$$\forall p \in allpaths(x,y) : blocked(p,Z) \rightarrow isDSeparated(x,y)$$

$$\exists p \in allpaths(x,y) : infoflows(p,Z) \rightarrow isDConnected(x,y)$$

A path p is $blocked(p) \iff (1) \vee (2)$ (d-separation see p43 BRML.)

(1) COLLIDER	NON-COLLIDER (2)
	$u > v > w$
$u > v < w$	$u < v < w$
	$u < v > w$
$\exists v \in p \setminus \{x,y\}$:	$\exists v \in p \setminus \{x,y\}$:
$collider(v) \wedge collider(v) \notin Z \wedge descendants(v) \notin Z$	$noncollider(v) \wedge v \in Z$

3.2 AMDS on complete graph at once

Quickest way is with graph edits **AMDS**. For variable sets $X, Y, Z : X \perp\!\!\!\perp Y|Z$

- **Ancestral** graph (keep X,Y,Z and ancestors(X,Y,Z))

A

- **Moralize** (add edges between all parents of the same node) ($\forall v \in Amarry(parents(v))$)

M

- **Disorient** (remove arrows)

D

- **Separate** (remove all edges from nodes in Z)

S

In the final S graph all unconnected nodes are D-separated.

symmetry	**decomposition**	**weak union**	**contraction**
$A \perp\!\!\!\perp B\|C$	$A \perp\!\!\!\perp B, C$	$A \perp\!\!\!\perp B, C$	$A \perp\!\!\!\perp B \wedge A \perp\!\!\!\perp C\|B$
$\iff$	$\Downarrow$	$\Downarrow$	$\Downarrow$
$B \perp\!\!\!\perp A\|C$	$A \perp\!\!\!\perp B \wedge A \perp\!\!\!\perp C$	$A \perp\!\!\!\perp B\|C \wedge A \perp\!\!\!\perp C\|B$	$A \perp\!\!\!\perp B, D$

Graphical networks are **Markov Equivalent** $\iff$ same independencies $\iff$ same skeleton $\wedge$ same immoralities.

4 Independence Equivalencies between

5 General Inference

Formula	Comment
$p(v_{1:t}, h_{1:t}) = p(h_1) * \Pi_{i=2}^{t} p(v_t \| h_t) p(h_t \| h_{t-1})$	JPD for an HMM
$p(v_t \| h_t)(= p(v_1 \| h_1) \iff$ stationary HMM)	emission matrix
$\forall t p(v_t \| h_t) = p(v_1 \| h_1)$	emission matrix for stationary HMM
$p(h_t \| h_{t-1})$	$transmission matrix$
$X \perp\!\!\!\perp Y \| Z \iff \forall x \in X, y \in Y : p(x \| y, z) = p(x \| z)$	conditional distributional independence.

6 Hidden Markov Models (HMM)