Inferring CIs from BNs

We start by enforcing a set of Local-CIs in a BN: $x_i \perp\!\!\!\perp ND(x_i)|Pa(x_i)$
Using axioms of probabilities, we can infer several others.

Axioms of CIs:

Let X, Y, Z, W be sets of variables.

- Decomposition: If $X \perp\!\!\!\perp \{Y, Z\} \implies X \perp\!\!\!\perp Y, X \perp\!\!\!\perp Z$
  - But the reverse is not necessarily true. Let x, y, z be binary variables.

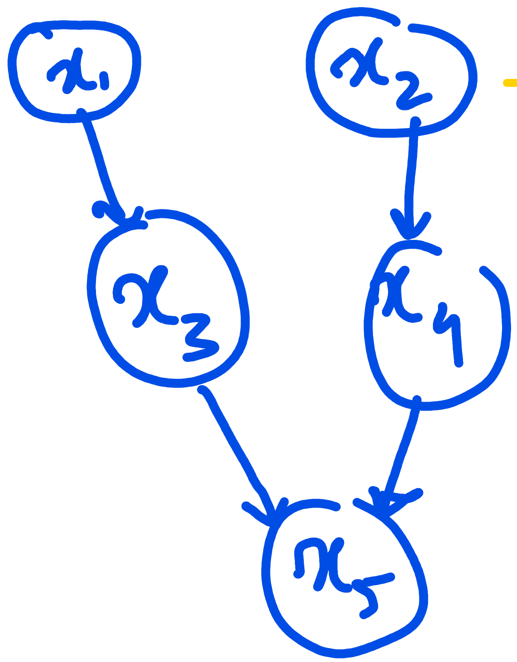
$$
\begin{aligned}
x &= y \bigoplus z, \quad y \perp\!\!\!\perp z \\
p(y) &= 1/2, P(z) = 1/2 \\
\text{Then} \quad P(x|y) &= P(x), P(x|z) = P(x), \textit{but} \;\; P(x|y,z) \neq P(x)
\end{aligned}
$$

Inferring CIs from BNs

- Weak Union: If $X \perp\!\!\!\perp \{Y, Z\} \implies X \perp\!\!\!\perp Y|Z$
  - But, in general, two independent variables can become dependent in the presence of a third variable. That is, if $X \perp\!\!\!\perp Y$ and $Z \not\perp\!\!\!\perp \{X, Y\}$ then it is not necessarily true that $X \perp\!\!\!\perp Y|Z$.

- Contraction: If $X \perp\!\!\!\perp Y|Z, X \perp\!\!\!\perp Z, \implies X \perp\!\!\!\perp \{Y, Z\}$

Examples of CIs that hold in BN but not covered by local-CI



$x_3 \perp\!\!\!\perp x_4 \mid x_2$

$x_3 \perp\!\!\!\perp x_2$

(contraction)

$\Rightarrow \boxed{x_3 \perp\!\!\!\perp x_4}$

holds in BN but not part of local-CI.

Global CIs in a BN

Three sets of variables X, Y, Z. If Z d-separates X from Y in BN then, $X \perp\!\!\!\perp Y | Z$.

In a directed graph H, Z d-separates X from Y if all paths P from any X to Y is blocked by Z.

A path P is blocked by Z when

$x_1 \rightleftarrows x_2 \rightarrow x_i \rightarrow x_{i+1} \rightleftarrows \cdots y$

x_i

$x \rightleftarrows x_2 \cdots \rightarrow z \rightarrow x_{i+1} \rightleftarrows \cdots - y$

1. $x_1 \to x_2 \to \ldots x_k$ and $x_i \in Z$
2. $x_1 \leftarrow x_2 \leftarrow \ldots x_k$ and $x_i \in Z$
3. $x_1 \ldots \leftarrow x_i \to \ldots x_k$ and $x_i \in Z$
4. $x_1 \ldots \to x_i \leftarrow \ldots x_k$ and $x_i \notin Z$ and $Desc(x_i) \notin Z$

Theorem

The d-separation test identifies the complete set of conditional independencies that hold in all distributions that conform to a given Bayesian network.

Global CIs Examples

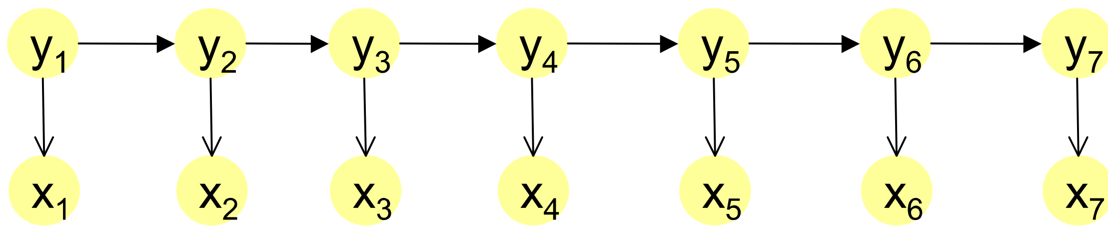
Global CIs and Local-CIs

In a BN, the set of CIs combined with the axioms of probability can be used to derive the Global-CIs.

Proof is long but easy to understand. Sketch of a proof available in the supplementary.

Popular Bayesian networks

- Hidden Markov Models: speech recognition, information extraction



  - State variables: discrete phoneme, entity tag
  - Observation variables: continuous (speech waveform), discrete (Word)
- Kalman Filters: State variables: continuous
  - Discussed later
- Topic models for text data
  1. Principled mechanism to categorize multi-labeled text documents while incorporating priors in a flexible generative framework
  2. Application: news tracking
- QMR (Quick Medical Reference) system
- PRMs: Probabilistic relational networks:

Can All Distributions be Represented as BNs?

$x \perp\!\!\!\perp y, x \perp\!\!\!\perp z, y \perp\!\!\!\perp z, x \not\perp\!\!\!\perp \{x, y\}$