Knowledge Discovery & Data Mining

— Markov Blanket & Boundary —

Instructor: Yong Zhuang

[*yong.zhuang@gvsu.edu*](mailto:yong.zhuang@gvsu.edu)

Filter Method

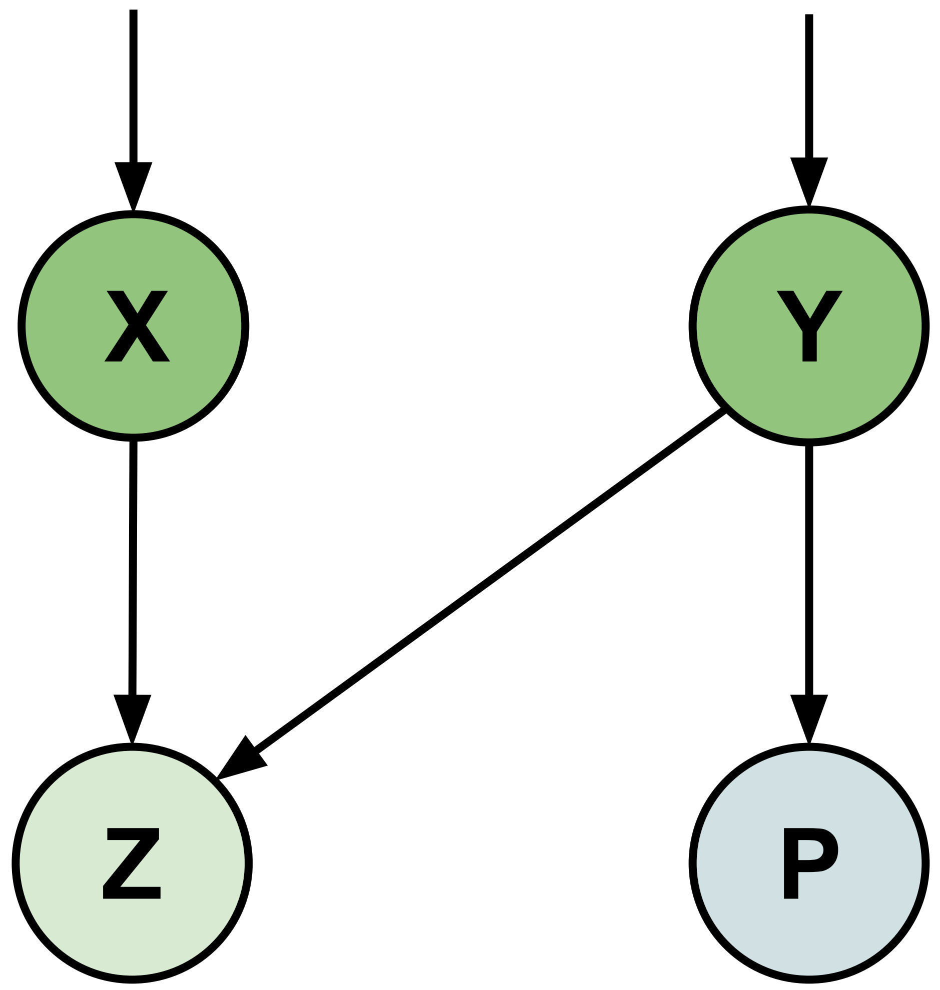
Filter methods select features from the dataset irrespective of the use of any predictive models. They are based only on general characteristics such as correlation with the feature to be predicted. Filtering methods suppress the least interesting features. Other features will be part of a classification or regression model used to classify or predict data. These methods are particularly efficient in terms of computation time and are robust to overfitting.

Outline

- Bayesian Belief Networks
- Markov Blanket & Boundary
- Filter Method: Incremental Association Markov blanket (IAMB)

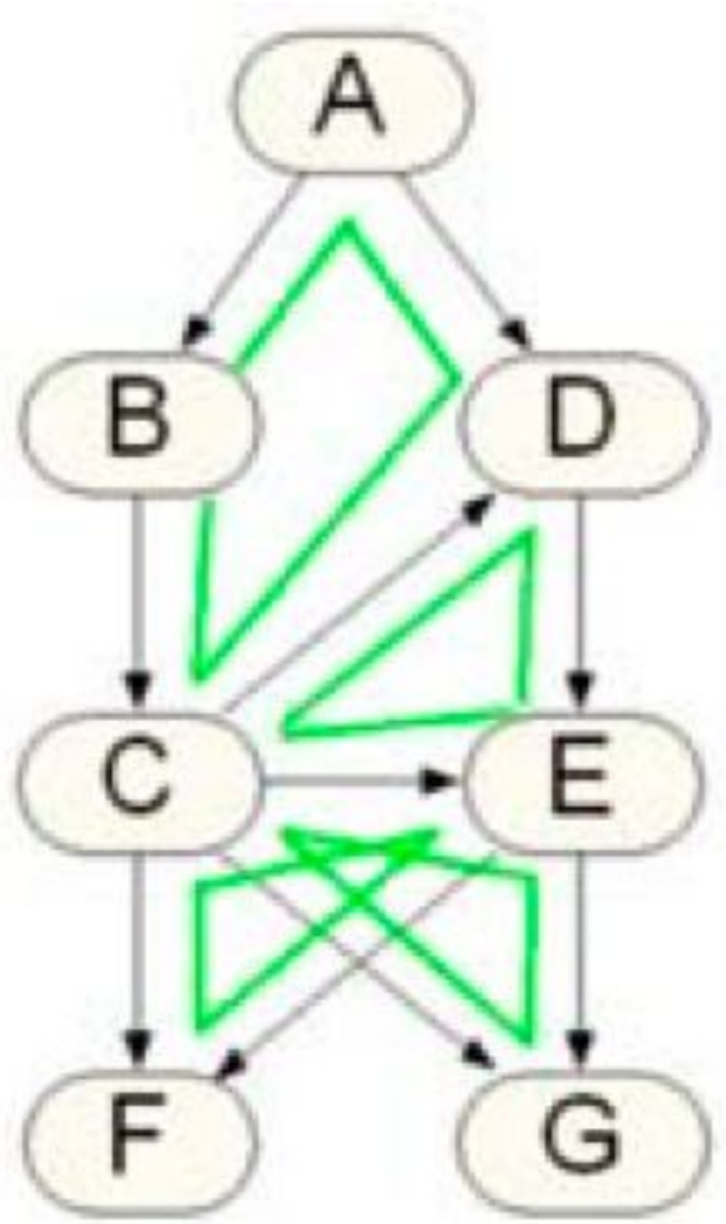
Bayesian Belief Networks

- **Bayesian belief networks** (also known as **Bayesian networks**, **probabilistic networks**): allow *class conditional independencies* between *subsets* of variables
- A (*directed acyclic*) graphical model of causal relationships
  - Represents dependency among the variables
  - Gives a specification of joint probability distribution

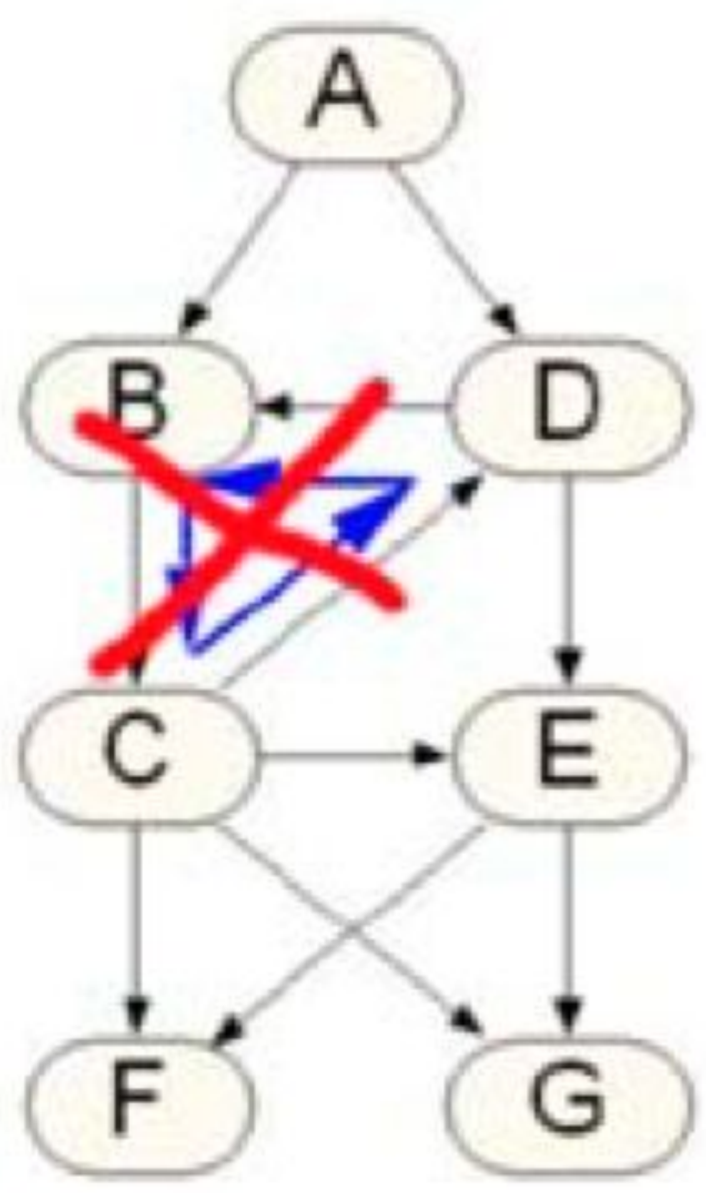


- Nodes: random variables
- Links: dependency
- X and Y are the parents of Z, and Y is the parent of P
- No dependency between Z and P
- Has no loops/cycles

Bayes Network



A valid Bayes net

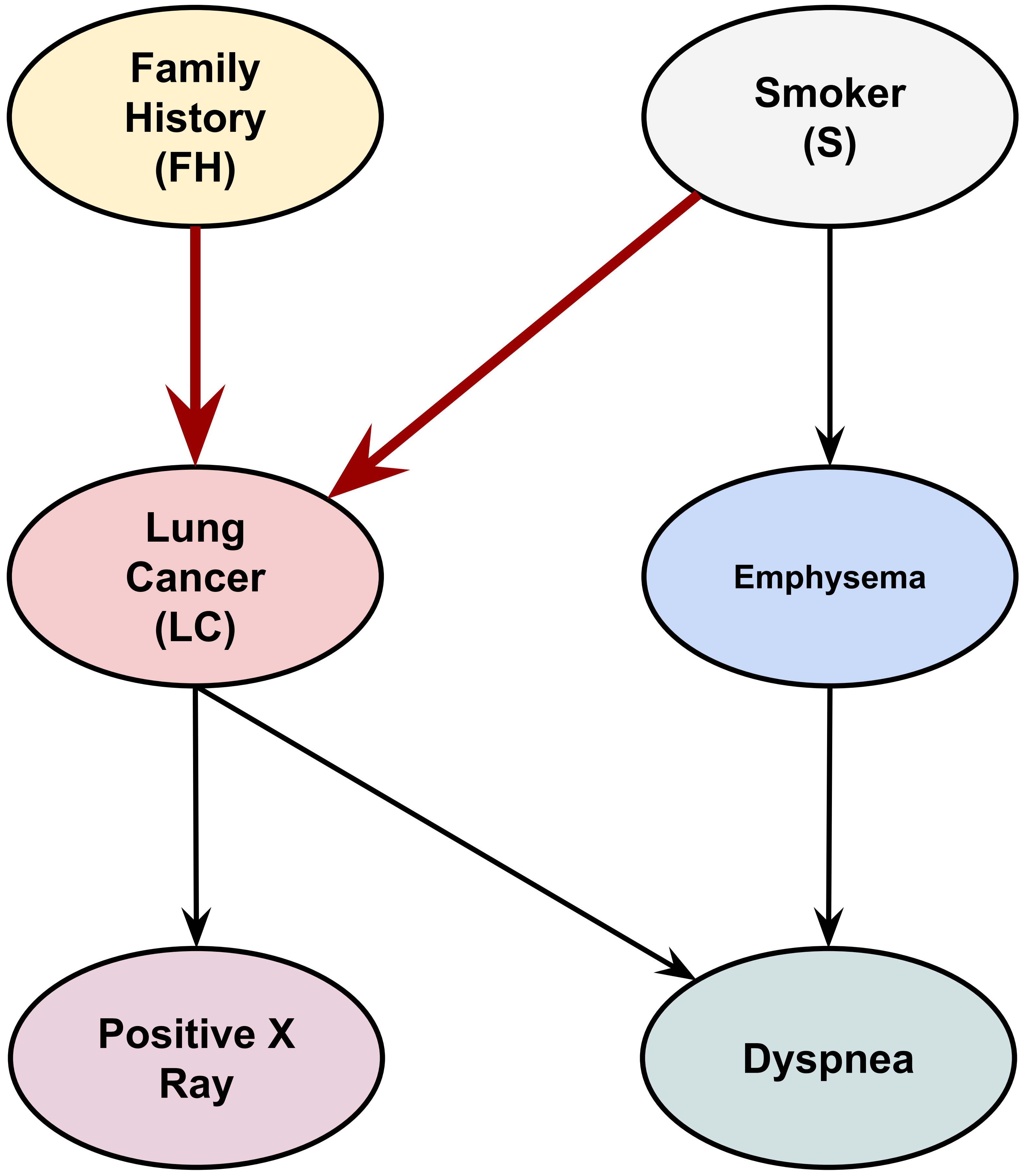


Not a Bayes net

What is Bayes Network used for?

- Diagnosis: P(cause|symptom)=?
- Prediction: P(symptom|cause)=?
- Classification: max P(class|data)
- Decision-making (given a cost function)

Bayes Network



Bayesian Belief Network

CPT: **Conditional Probability Table** for variable **Lung Cancer**

	(FH, S)	(FH, ~S)	(~FH, S)	(~FH, ~S)
LC	0.8	0.5	0.7	0.1
~LC	0.2	0.5	0.3	0.9

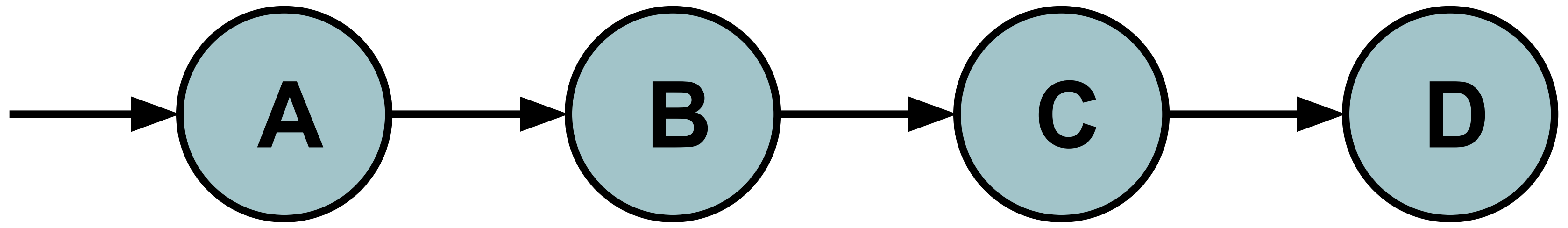
conditional probability for each possible combination of its parents

Derivation of the probability of a particular combination of values of **X**, from CPT:

$$P(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} P(x_i | \mathrm{Parents}(x_i))$$

Bayes Network

Example. Given a Bayes Network G with four Nodes A,B,C, and D as follows, what is the expression for $P(x_A, x_B, x_C, x_D)$

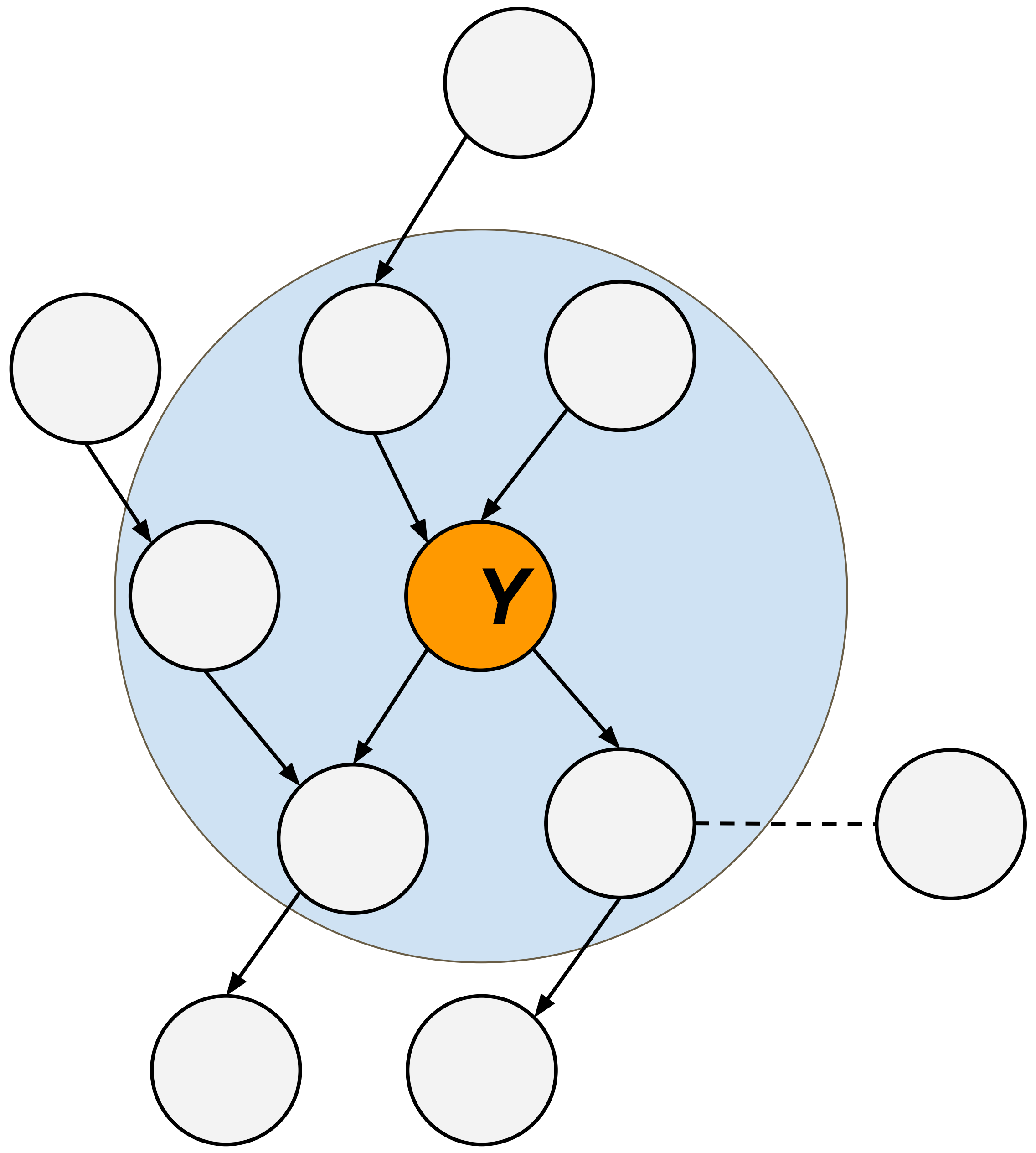


$$P(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} P(x_i|\text{Parents}(x_i))$$

Markov Boundary Theory

In faithful distributions, the Markov boundary of a node contains all the variables that shield the node from the rest of the Bayesian network. This means that the Markov boundary of a node is the only knowledge needed to predict the behavior of that node.

--- Judea Pearl, 1988.



In faithful distributions, Markov boundary corresponds to a local causal neighborhood of the that variable and consists of all its direct causes, effects, and causes of the direct effects.

--- Neapolitan, 2004; Tsamardinos and Aliferis, 2003.

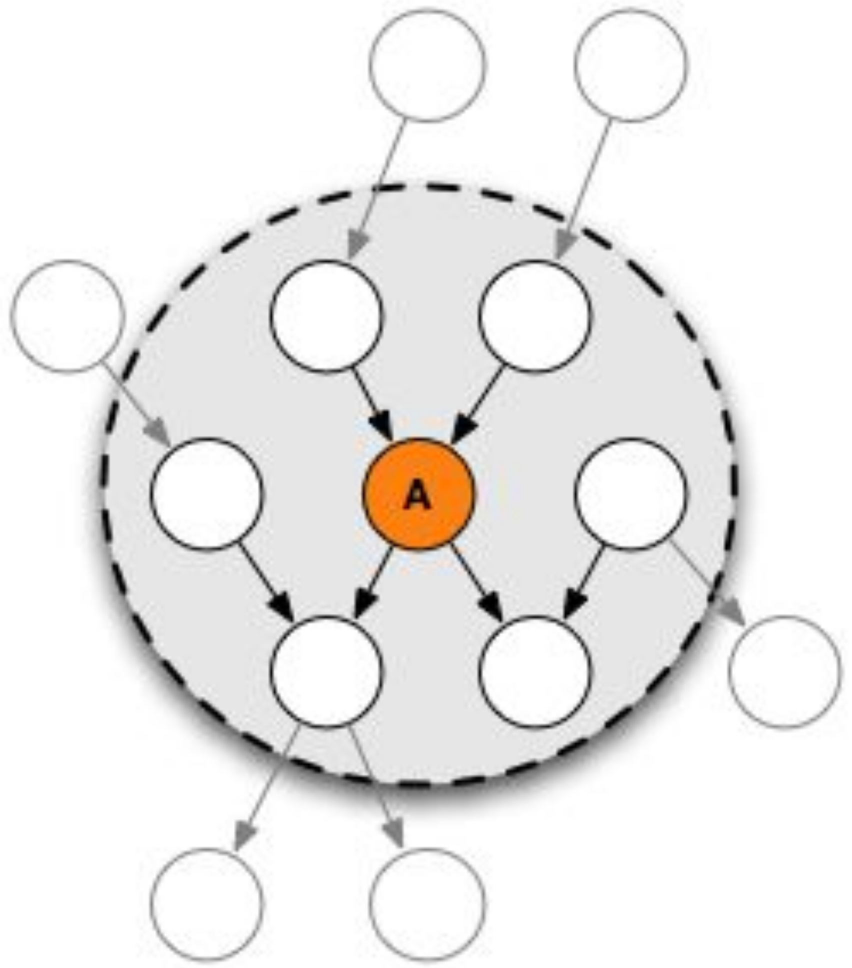
Markov Blanket & Boundary

Markov blanket: A Markov blanket $\boldsymbol{M}$ *of the response variable* $\mathrm{T} \in \boldsymbol{V}$ *in the joint probability distribution* $\mathbb{P}$ *over variables* $\boldsymbol{V}$ *is a set of variables conditioned on which all other variables are independent of* T, *that is, for every* $\mathrm{X} \in (\boldsymbol{V} \setminus \boldsymbol{M} \setminus \{\mathrm{T}\})$, $\mathrm{T} \perp \mathrm{X} \mid \boldsymbol{M}$.

Trivially, the set of all variables $\boldsymbol{V}$ excluding T is a Markov blanket of T. Also one can take a small Markov blanket and produce a larger one by adding arbitrary (predictively redundant or irrelevant) variables. Hence, only minimal Markov blankets are of interest.

Markov boundary: *If no proper subset of* $\boldsymbol{M}$ *satisfies the definition of Markov blanket of* T, *then* $\boldsymbol{M}$ *is called a Markov boundary of* T.

In a Bayesian network, the Markov boundary of node A includes its parents, children and the other parents of all of its children.



Incremental Association Markov blanket (IAMB)

<u>Algorithm</u> **<u>IAMB</u>**

<u>Input</u>: dataset $\mathbb{D}$ (a sample from distribution $\mathbb{P}$) for variables $\boldsymbol{V}$, including a response variable T.

<u>Output</u>: a Markov blanket $\boldsymbol{M}$ of T.

Phase I: Forward

Initialize $\boldsymbol{M}$ with an empty set
Initialize the set of eligible variables $\boldsymbol{E} \leftarrow \boldsymbol{V} \setminus \{T\}$
Repeat
    $Y \leftarrow \text{argmax}_{X \in \boldsymbol{E}}$ Association$(T, X \mid \boldsymbol{M})$
    $\boldsymbol{E} \leftarrow \boldsymbol{E} \setminus \{Y\}$
    If $T \not\perp Y \mid \boldsymbol{M}$ then
        $\boldsymbol{M} \leftarrow \boldsymbol{M} \cup \{Y\}$
        $\boldsymbol{E} \leftarrow \boldsymbol{V} \setminus \boldsymbol{M} \setminus \{T\}$
Until $\boldsymbol{E}$ is empty

Phase II: Backward

For each $X \in \boldsymbol{M}$
    If $T \perp X \mid (\boldsymbol{M} \setminus \{X\})$ then
        $\boldsymbol{M} \leftarrow \boldsymbol{M} \setminus \{X\}$
End
Output $\boldsymbol{M}$

Summary

- Bayesian Belief Networks
- Markov Blanket & Boundary
- Filter Method: Incremental Association Markov blanket (IAMB)