ECE 6504: Advanced Topics in Machine Learning

Probabilistic Graphical Models and Large-Scale Learning

Topics:

- Bayes Nets: Representation/Semantics
  - v-structures
  - Probabilistic influence, Active Trails

Readings: Barber 3.3; KF 3.3.1-3.3.2

Dhruv Batra

Virginia Tech

Plan for today

- Notation Clarification
- Errata #1: Number of parameters in disease network
- Errata #2: Car start v-structure example
- Bayesian Networks
  - Probabilistic influence & active trails
  - d-separation
  - General (conditional) independence assumptions in a BN

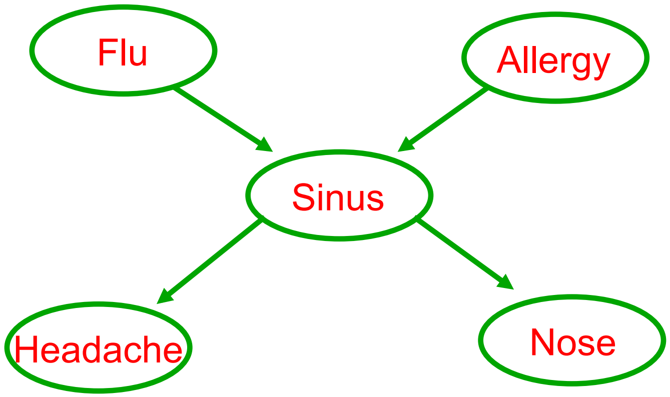
A general Bayes net

- Set of random variables
- Directed acyclic graph
  - Encodes independence assumptions
- CPTs
  - Conditional Probability Tables
- Joint distribution:

$$P(X_1, \ldots, X_n) = \prod_{i=1}^{n} P\left(X_i \mid \mathbf{Pa}_{X_i}\right)$$

Factorized distributions

- Given
  - Random vars $X_1,\ldots,X_n$
  - *P* distribution over vars
  - BN structure *G* over same vars



- *P* factorizes according to *G* if

$$P(X_1,\ldots,X_n) = \prod_{i=1}^{n} P\left(X_i \mid \mathbf{Pa}_{X_i}\right)$$

How many parameters in a BN?

- Discrete variables $X_1, \ldots, X_n$

- Graph
  - Defines parents of X_i, $\mathbf{Pa}_{X_i}$

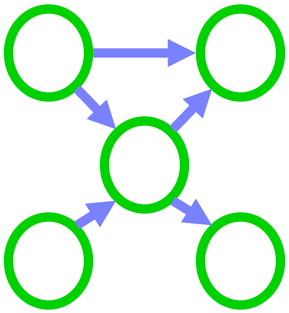
- CPTs – $P(X_i | \mathbf{Pa}_{X_i})$

Independencies in Problem

World, Data, reality:

True distribution *P* contains independence assertions

BN:



Graph *G* encodes local independence assumptions

Bayes Nets

- BN encode (conditional) independence assumptions.
  - I(G) = {X indep of Y given Z}

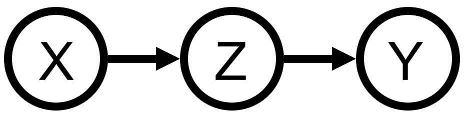
- Which ones?
- And how can we easily read them?

Local Structures

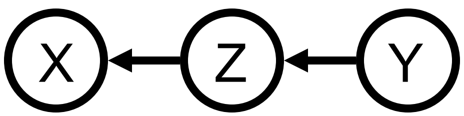
- What’s the smallest Bayes Net?

Local Structures

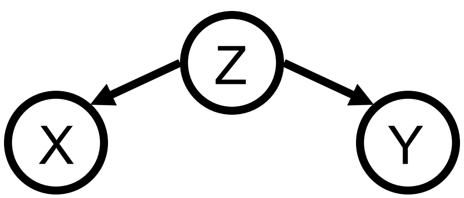
Indirect causal effect:



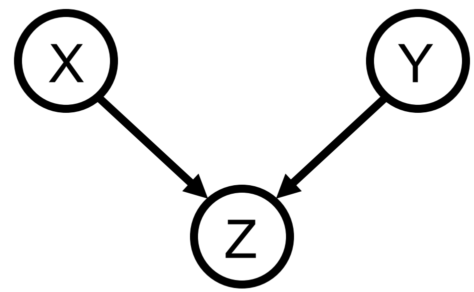
Indirect evidential effect:



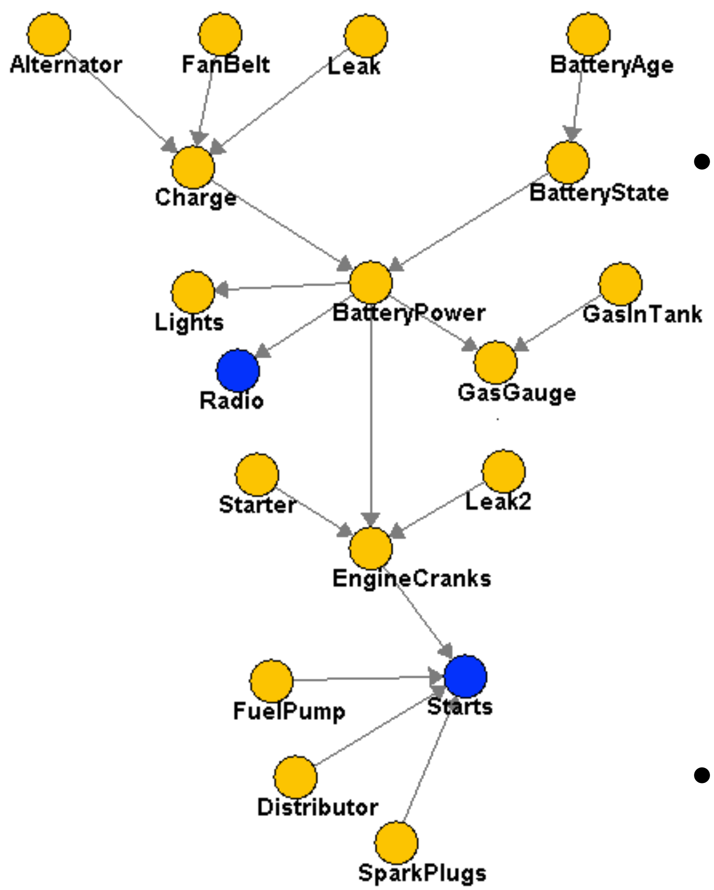
Common cause:



Common effect:



Car starts BN



- 18 binary attributes
- Inference
  - P(BatteryAge|Starts=f)
- 2^{18} terms, why so fast?

Bayes Ball Rules

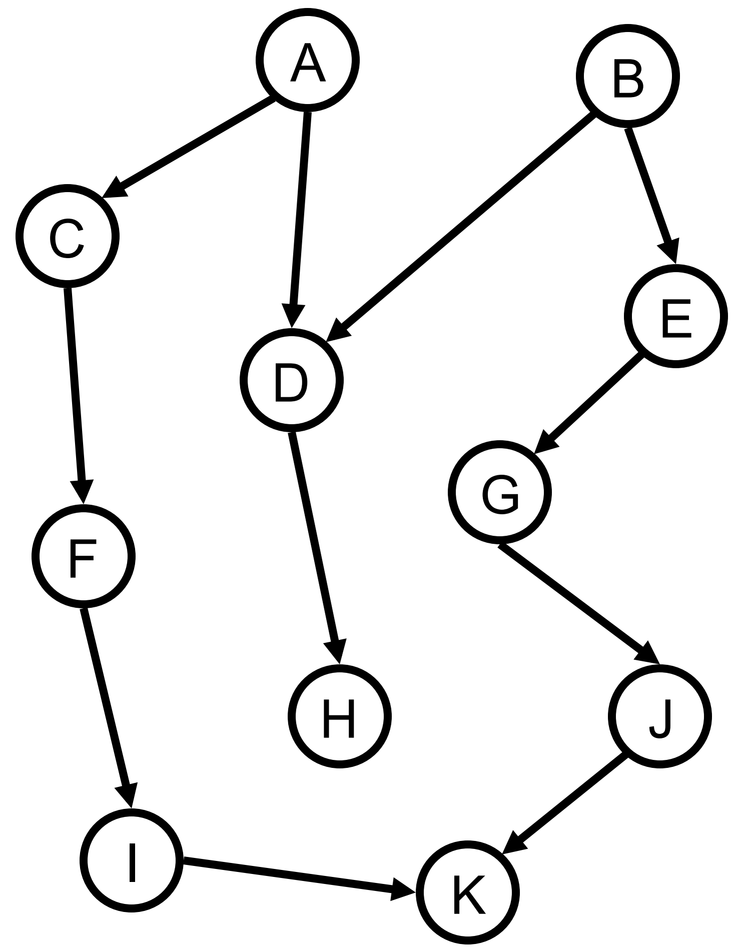
- Flow of information
  - on board

Active trails formalized

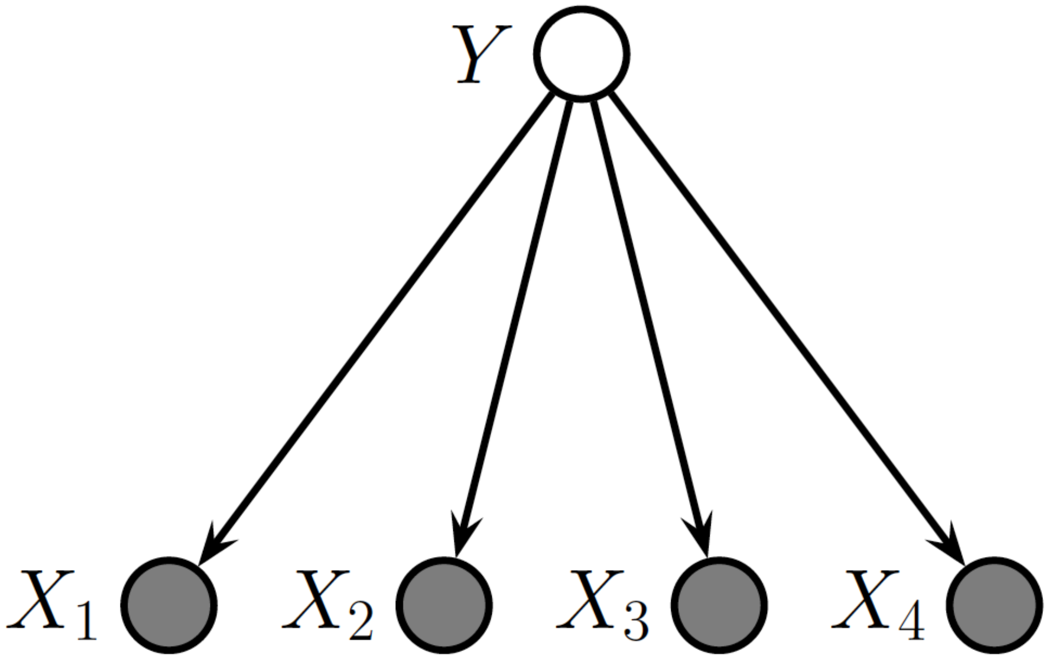
- Let variables $\boldsymbol{O} \subseteq \{X_1,\ldots,X_n\}$ be observed
- A path $X_1 - X_2 - \cdot\cdot\cdot - X_k$ is an **active trail** if for each consecutive triplet:
  - $X_{i-1} \rightarrow X_i \rightarrow X_{i+1}$, and X_i is **not observed** ($X_i \notin \boldsymbol{O}$)
  - $X_{i-1} \leftarrow X_i \leftarrow X_{i+1}$, and X_i is **not observed** ($X_i \notin \boldsymbol{O}$)
  - $X_{i-1} \leftarrow X_i \rightarrow X_{i+1}$, and X_i is **not observed** ($X_i \notin \boldsymbol{O}$)
  - $X_{i-1} \rightarrow X_i \leftarrow X_{i+1}$, and X_i **is observed** ($X_i \in \boldsymbol{O}$), or one of its descendents **is observed**

Active trails and Independence

- **Theorem**: Variables $\mathbf{X_i}$ **and** $\mathbf{X_j}$ are independent given $\boldsymbol{Z}$ if
  - **no active trail** between X_i and X_j when variables $\boldsymbol{Z} \subseteq \{X_1,\ldots,X_n\}$ are observed



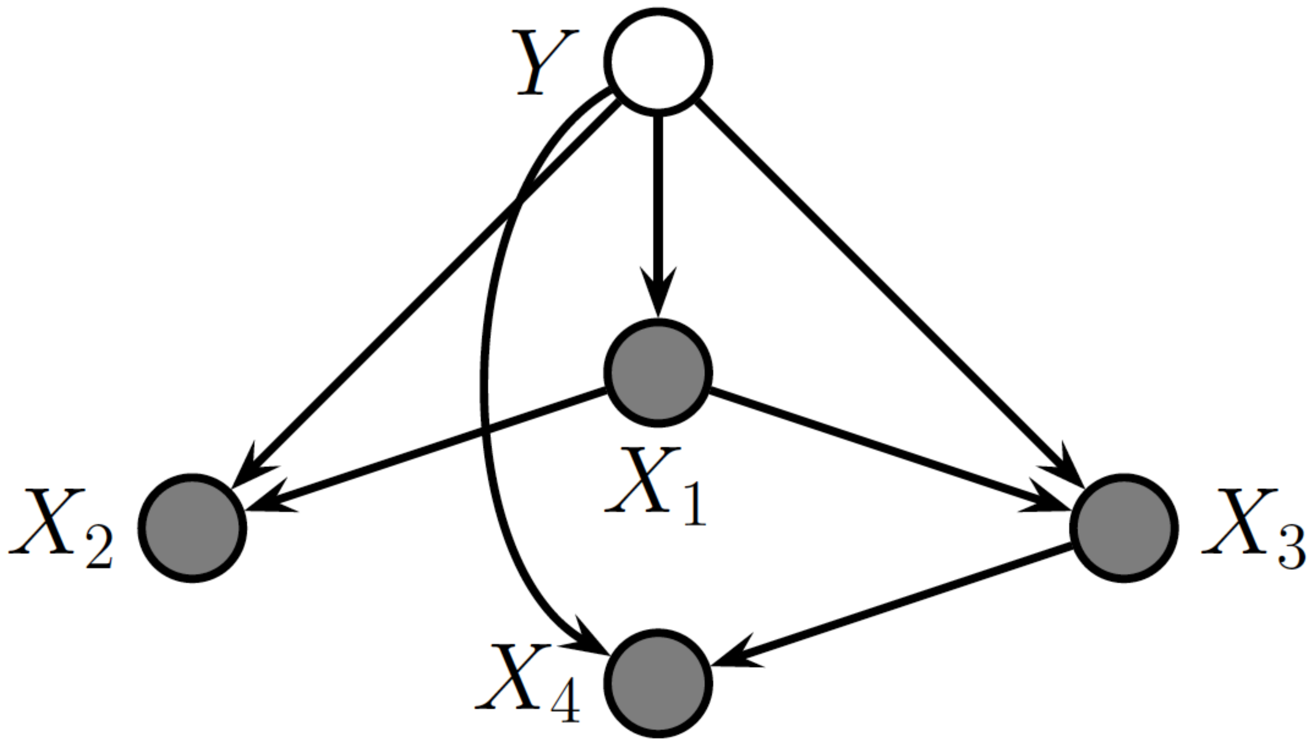
Name That Model



Naïve Bayes:

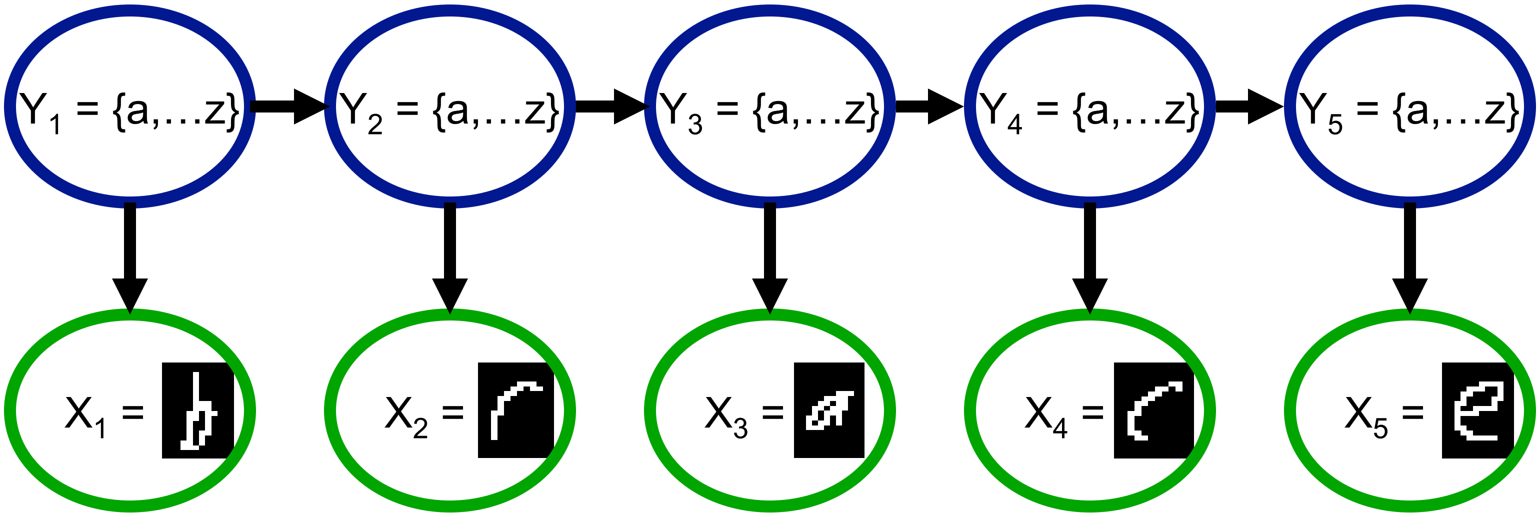
$$p(y, \mathbf{x}) = p(y) \prod_{j=1}^{D} p(x_j | y)$$

Name That Model



Tree-Augmented Naïve Bayes (TAN)

Name That Model



Hidden Markov Model (HMM)

Figure Credit: Carlos Guestrin