Probabilistic Graphical Models

Provides us with a framework for inference on probability distributions

Graphical model (informally): Nodes represent RVs
edges represent influence

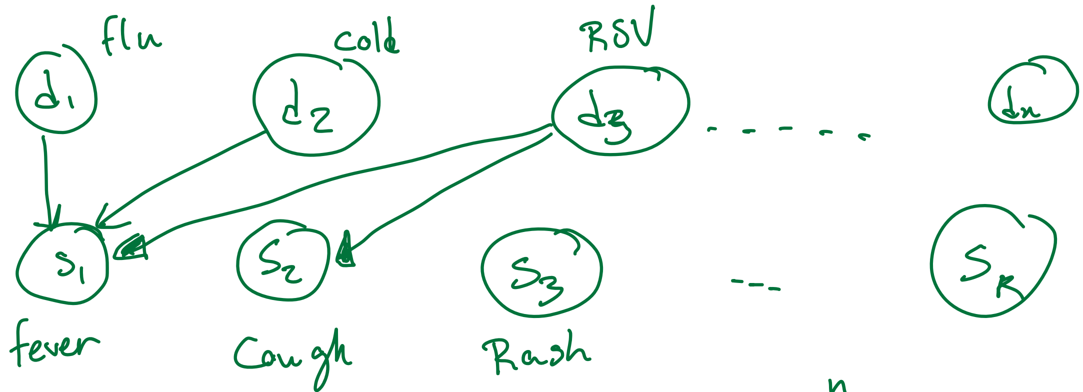
Less dense graphs will correspond to probability distributions with structure that can allow for efficient inference.

Inference tasks
Draw conclusions about distribution

Example 1:

Diseases [unobserved]

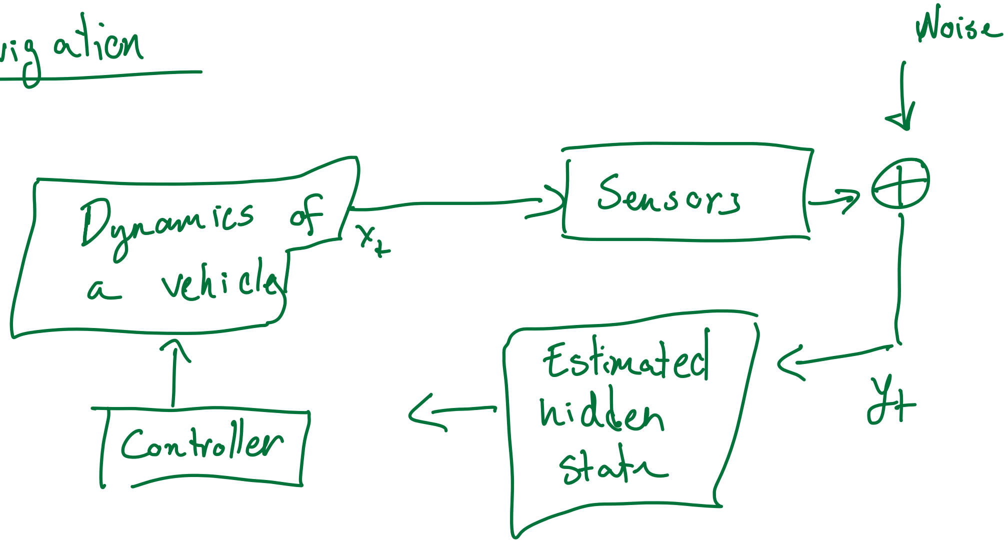
Symptoms [observed]

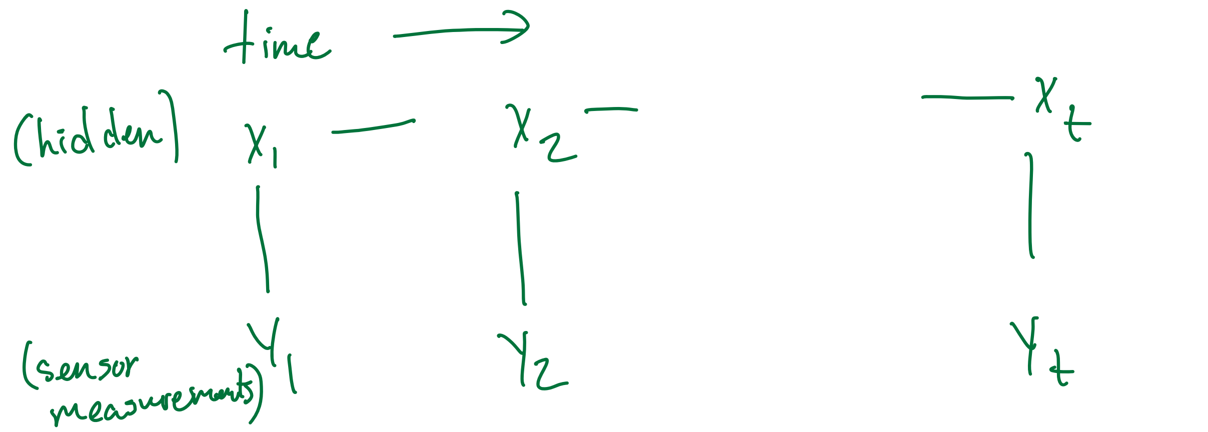


A patient can be represented as $d \in \{0,1\}^n$
Also as $f_i \in \{0,1\}^k$

Structure of this Bayesian network will allow efficient inference of, e.g. $\Pr[d_3 = 1 \mid s = (01000)]$

Navigation





Inference Given a transcript $y_1 \dots y_t$

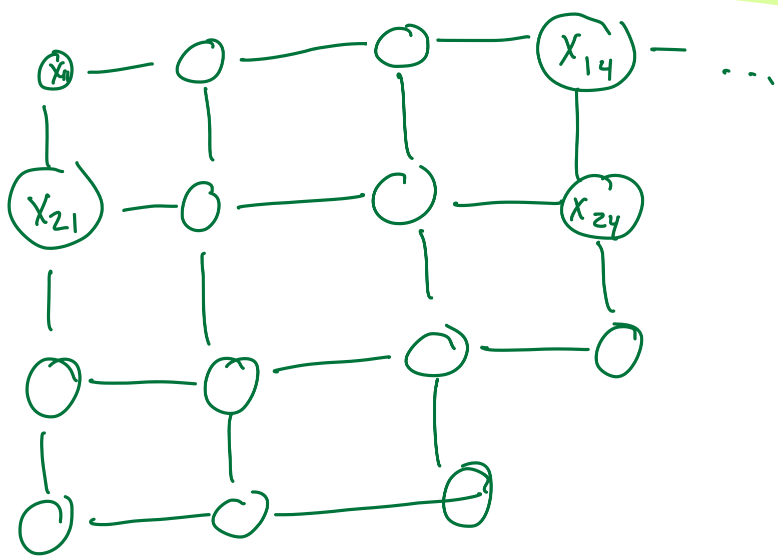
$$\Pr_y[x_t \mid y_1 \dots y_t]$$

Image superresolution

Markov Random field

Given a low-dimensional set of pixels, fill in the other pixels

$$\Pr[X \mid x_{11}, x_{14}, x_{21}, x_{24}]$$



General theme

[We know] Probability distribution over $X = [x_1, \dots x_n]$

$x_i \in \mathcal{X}$

$|\mathcal{X}| < \infty$

We observe $Y = [Y_1, \dots Y_k]$

[We want] inference

Examples

- Find the most probable realization

$$\underset{x}{\operatorname{argmax}} \; Pr_Y(x)$$

- Marginals

$$Pr_Y[X_1]$$

- Sample $\quad x \sim P_Y(x)$

Hard? $\quad n \gg 1$, $k \gg 1$ $\quad \tilde{O}(|\mathcal{X}|^{n \times k})$

If, for example $\quad x_i \perp x_j \mid Y \quad \forall i, j$

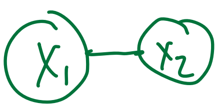
then $P_Y(x)$, representable using $O(|\mathcal{X}| \cdot n)$

PGMs

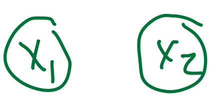
Graph Separation

$\uparrow$

Graph $\longrightarrow$ Factorizations $\longrightarrow$ (Conditional) Independence



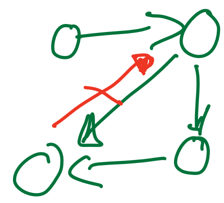
$P[x_1, x_2]$



$Pr[x_1] \, Pr[x_2]$ $\qquad x_1 \perp x_2$

★ Directed Graphical Models

Def: Directed, acyclic graph $G = (V, E)$



- Parent set of $i \in V$ $\quad \pi(i) = \{ j \in V \mid (j,i) \in E \}$

A directed graphical model on $G = (V, E) \to$ a DAG is a family of distributions that factorizes as

$$\Pr[x_1, \dots, x_n] \triangleq \prod_{i=1}^{n} f_i(x_i, x_{\pi(i)}) \qquad (\exists f_i)$$

Claim

$$f_i(x_i, x_{\pi(i)}) = P[x_i \mid x_{\pi(i)}]$$