Possible queries

- Inference
  \[ P(F = t \mid H = t) \]

- Most probable explanation
  \[ P(F = t, d = a, s = s, N = n \mid H = t) \]

- Active data collection
  do you have a runny nose?
General probabilistic inference

- Query: \( P(X \mid e) \)
  - Notation: capital letter \( X \) \( \rightarrow \) answer for assignment
  - Defn. of cond. prob.

- Using Bayes rule:
  \[ P(X \mid e) = \frac{P(X, e)}{P(e)} \]
  - Constant doesn't depend on \( X \)

- Normalization:
  \[ P(X \mid e) \propto P(X, e) \]

Marginalization of \( S \)

\[
\begin{align*}
P(F=t, N=t) &= P(F) P(S|F) P(N|S) \\
P(F=t, N=t) &= P(F=t, S=t, N=t) + P(F=t, S=f, N=t) \\
&= \sum_S P(F=t, S, N=t) \\
&= \sum_S P(F=t) P(S|F=t) P(N=t|S)
\end{align*}
\]
Probabilistic inference example

Inference seems exponential in number of variables!
Actually, inference in graphical models is NP-hard 😞

Fast probabilistic inference example – Variable elimination

(Potential for) Exponential reduction in computation!
Understanding variable elimination –
Exploiting distributivity

\[ a(b + c) = ab + ac \]

\[ P(F, N=\text{t}) = \sum_s P(F) P(S=\text{t}|F) P(N=\text{t}|S=\text{t}) \]

\[ + P(F) P(S=\text{f}|F) P(N=\text{t}|S=\text{f}) \]

\[ = P(F) \sum_s P(S=\text{t}|F) P(N=\text{t}|S=\text{t}) \]

\[ + P(S=\text{f}|F) P(N=\text{t}|S=\text{f}) \]

\[ = P(F) \sum_s P(S=\text{t}|F) P(N=\text{t}|S) \]

Understanding variable elimination –
Order can make a HUGE difference

\[ P(F, N=\text{t}) = \sum_{a,b} P(F) P(a) P(G(F,a)) P(h,b) P(w=\text{t}) \]

\[ = \sum_{a,b} P(F) P(a) \sum S P(G(F,a), P(h,b), P(w=\text{t})) \]

\[ = \sum_{a,b} P(F) P(a) \sum S g_1(F,a) \phi_{\text{values}} \]

\[ = \sum_{a,b} P(F) P(a) \sum S \prod_{i=1}^{n} P(x_i, c) \]

\[ \text{OR} \]

\[ P(x_1) = \sum_{x_2, \ldots, x_n} P(c) P(x_1 | c) \prod_{i=1}^{n} P(x_i | c) \]

\[ \text{OR} \]

\[ \text{OR} \]

\[ = \sum_{S = \{ x_1, \ldots, x_n \}} P(c) P(x_1 | c) \prod_{i=1}^{n} P(x_i | c) \]

\[ \text{OR} \]

\[ = \sum_{S = \{ x_1, \ldots, x_n \}} P(c) P(x_1 | c) \prod_{i=1}^{n} P(x_i | c) \]

\[ \text{OR} \]

\[ = \sum_{S = \{ x_1, \ldots, x_n \}} P(c) P(x_1 | c) \prod_{i=1}^{n} P(x_i | c) \]

©2005-2009 Carlos Guestrin
Understanding variable elimination – Another example

Variable elimination algorithm

1. Given a BN and a query $P(X|e) \neq P(X,e)$
2. Instantiate evidence $e$, e.g., $S, Y, N \neq t$
3. Choose an ordering on variables, e.g., $X_1, \ldots, X_n$
   - For $i = 1$ to $n$, if $X_i \notin \{X,e\}$
     - Collect factors $f_1, \ldots, f_k$ that include $X_i$
   - Generate a new factor by eliminating $X_i$ from these factors

$$g = \sum \prod_{j=1}^{k} f_j$$

- Variable $X_i$ has been eliminated!
- Normalize $P(X,e)$ to obtain $P(X|e)$

$P(t) = \sum_{X} P(x,e)$ \& normalization

$P(y | N=t) \triangleq \sum_{S,h} P(s) P(h | s) P(N=t | s) P(c | N=t) \sum_{S,h} P(h | s) P(y | h, c) \frac{g_1(s, y, c)}{P(y | s, c)}$
Complexity of variable elimination – (Poly)-tree graphs

Variable elimination order:
- Start from “leaves” up
- Find topological order, eliminate variables in reverse order

Linear in number of variables!!! (versus exponential)

Complexity of variable elimination – Graphs with loops

if I eliminate: every variable takes on k values

Exponential in number of variables in largest factor generated
Complexity of variable elimination – Tree-width

Moralize graph:
Connect parents into a clique and remove edge directions

Size of largest clique of triangulation of

Complexity of VE elimination:
(“Only”) exponential in tree-width
(Loosely speaking Tree-width - 1 is lower bounded by “maximum minimum node cut”)

Example: Large tree-width with small number of parents

Compromise representation \(\Rightarrow\) Easy inference ️
Choosing an elimination order

- Choosing best order is NP-complete
  - Reduction from MAX-Clique
- Many good heuristics (some with guarantees)
- Ultimately, can’t beat NP-hardness of inference
  - Even optimal order can lead to exponential variable elimination computation
- In practice
  - Variable elimination often very effective
  - Many (many many) approximate inference approaches available when variable elimination too expensive