Exact inference in Bayesian networks

Chapter 14, Section 4

Inference tasks

Simple queries: compute posterior marginal $\mathbf{P}(X_i|\mathbf{E}=\mathbf{e})$
e.g., $P(Burglar|JohnCalls = true, MaryCalls = true)$
or shorter, $P(B|j, m)$

Conjunctive queries: $\mathbf{P}(X_i, X_j|\mathbf{E}=\mathbf{e}) = \mathbf{P}(X_i|\mathbf{E}=\mathbf{e})\mathbf{P}(X_j|X_i, \mathbf{E}=\mathbf{e})$

Optimal decisions: decision networks include utility information;
probabilistic inference required for $P(outcome|action, evidence)$

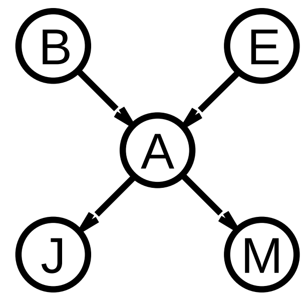
Value of information: which evidence to seek next?

Sensitivity analysis: which probability values are most critical?

Explanation: why do I need a new starter motor?

Inference by enumeration

Slightly intelligent way to sum out variables from the joint without actually constructing its explicit representation



Simple query on the burglary network:

$$\begin{aligned}
&\mathbf{P}(B|j,m) \\
&\quad = \mathbf{P}(B,j,m)/P(j,m) \\
&\quad = \alpha\mathbf{P}(B,j,m) \\
&\quad = \alpha \textstyle\sum_e \sum_a \mathbf{P}(B,e,a,j,m)
\end{aligned}$$

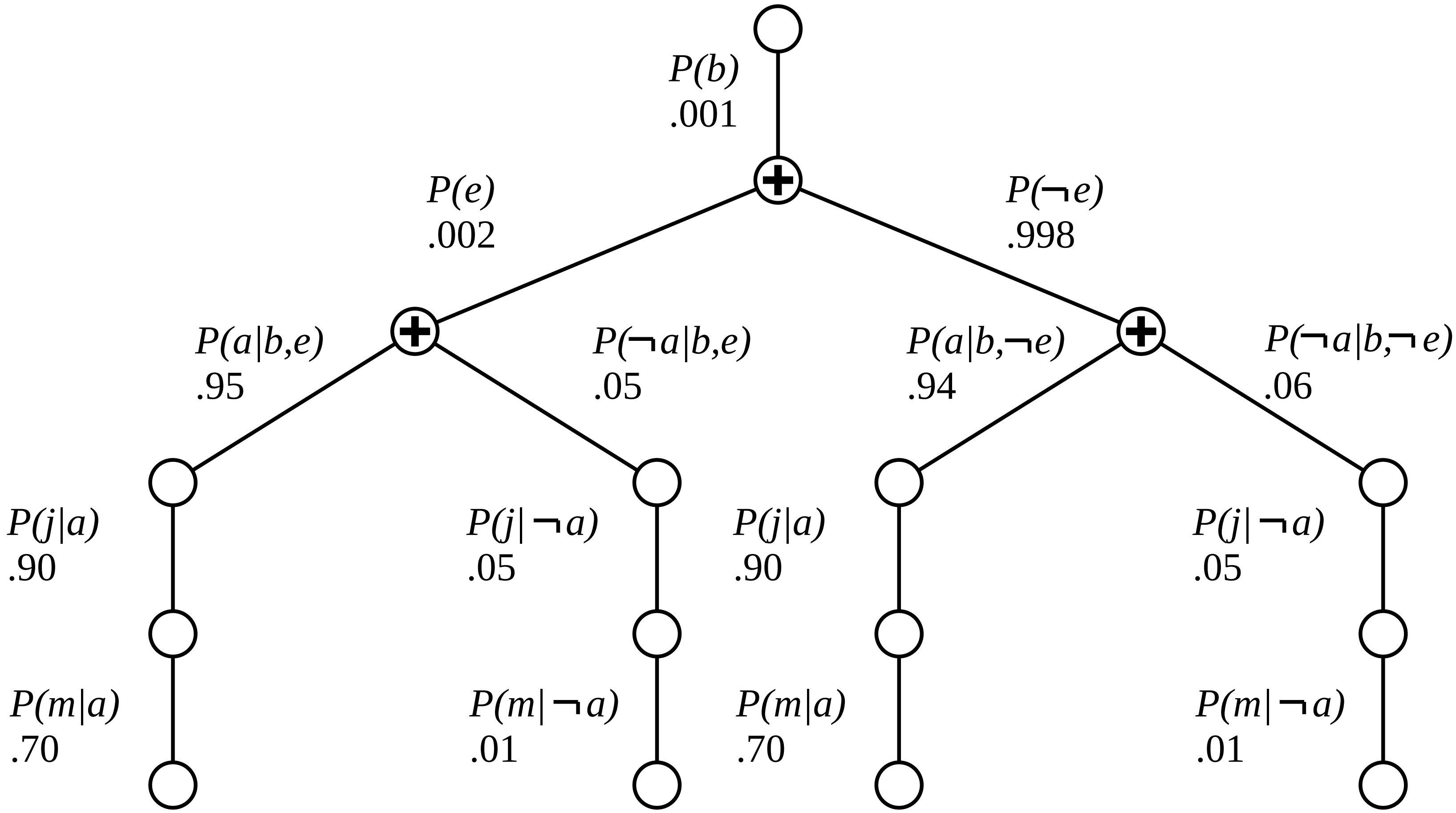
(where e and a are the hidden variables)

Rewrite full joint entries using product of CPT entries:

$$\begin{aligned}
&\mathbf{P}(B|j,m) \\
&\quad = \alpha \textstyle\sum_e \sum_a \mathbf{P}(B)P(e)\mathbf{P}(a|B,e)P(j|a)P(m|a) \\
&\quad = \alpha\, \mathbf{P}(B) \textstyle\sum_e P(e) \sum_a \mathbf{P}(a|B,e)P(j|a)P(m|a)
\end{aligned}$$

Recursive depth-first enumeration: $O(n)$ space, $O(d^n)$ time

Evaluation tree



Enumeration is inefficient: repeated computation
e.g., computes $P(j|a)P(m|a)$ for each value of e

Inference by variable elimination

Variable elimination: carry out summations right-to-left,
storing intermediate results (factors) to avoid recomputation

$$\begin{aligned}\mathbf{P}(B|j,m) &= \alpha\, \mathbf{P}(B) \textstyle\sum_e P(e) \sum_a \mathbf{P}(a|B,e)\, P(j|a)\, P(m|a) \\ &= \alpha\, \mathbf{f}_1(B) \textstyle\sum_e \mathbf{f}_2(E) \sum_a \mathbf{f}_3(A,B,E)\, \mathbf{f}_4(A)\, \mathbf{f}_5(A)\end{aligned}$$

(where $\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_4, \mathbf{f}_5,$ are 2-element vectors, and $\mathbf{f}_3$ is a $2 \times 2 \times 2$ matrix)

Sum out A to get the 2×2 matrix $\mathbf{f}_6$, and then E to get the 2-vector $\mathbf{f}_7$:

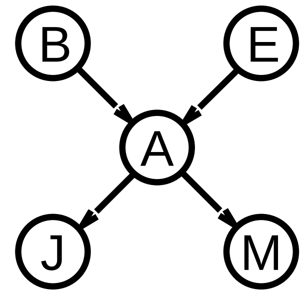
$$\begin{aligned}\mathbf{f}_6(B,E) &= \textstyle\sum_a \mathbf{f}_3(A,B,E) \times \mathbf{f}_4(A) \times \mathbf{f}_5(A) \\ &= \mathbf{f}_3(a,B,E) \times \mathbf{f}_4(a) \times \mathbf{f}_5(a) + \mathbf{f}_3(\neg a,B,E) \times \mathbf{f}_4(\neg a) \times \mathbf{f}_5(\neg a) \\ \mathbf{f}_7(B) &= \textstyle\sum_e \mathbf{f}_2(E) \times \mathbf{f}_6(B,E) \;=\; \mathbf{f}_2(e) \times \mathbf{f}_6(B,e) + \mathbf{f}_2(\neg e) \times \mathbf{f}_6(B,\neg e)\end{aligned}$$

Finally, we get this:

$$\mathbf{P}(B|j,m) = \alpha\, \mathbf{f}_1(B) \times \mathbf{f}_7(B)$$

Irrelevant variables

Consider the query $P(JohnCalls|Burglary = true)$



$$P(J|b) = \alpha P(b) \sum_e P(e) \sum_a P(a|b,e) P(J|a) \sum_m P(m|a)$$

Sum over m is identically 1; M is **irrelevant** to the query

Theorem: Y is irrelevant unless $Y \in Ancestors(\{X\} \cup \mathbf{E})$

Here, $X = JohnCalls$, $\mathbf{E} = \{Burglary\}$, and
$Ancestors(\{X\} \cup \mathbf{E}) = \{Alarm, Earthquake\}$
so $MaryCalls$ is irrelevant