Exact Inference: Variable Elimination

Vibhav Gogate

THE UNIVERSITY OF TEXAS AT DALLAS
Erik Jonsson School of Engineering and Computer Science

Sum-product Inference Problems

- Given a PGM containing factors $\phi_1, \ldots, \phi_m$ and an assignment $\mathbf{E} = \mathbf{e}$ to a subset of its variables (evidence), find
  1. $\Pr(\mathbf{E} = \mathbf{e})$: Probability of evidence
  2. $\Pr(X | \mathbf{E} = \mathbf{e})$: Conditional distribution at a variable X given evidence (marginals)

By definition, these are sum-product tasks and can be reduced to computing partition function of a Markov network. Why?

$$Z = \sum_{\mathbf{x}} \prod_{i=1}^{m} \phi_i(\mathbf{x}_i)$$

where $\mathbf{x}_i$ is the projection of $\mathbf{x}$ on the variables involved in ϕ_i.

Naive sum-product algorithm

$$Z = \sum_{\mathbf{x}} \prod_{i=1}^{m} \phi_i(\mathbf{x}_i)$$

- Initialize $Z = 0$
- Iterate over all possible assignments $\mathbf{x}$
  - Compute the value of each assignment $value(\mathbf{x}) = \prod_{i=1}^{m} \phi_i(\mathbf{x}_i)$ (`Product`)
  - $Z = Z + value(\mathbf{x})$ (`Sum`)
- **Return** Z

Product Operation

$$\phi(A, B, C) = \phi(A, B) \times \phi(B, C)$$

a^1	b^1	0.5
a^1	b^2	0.8
a^2	b^1	0.1
a^2	b^2	0
a^3	b^1	0.3
a^3	b^2	0.9

b^1	c^1	0.5
b^1	c^2	0.7
b^2	c^1	0.1
b^2	c^2	0.2

a^1	b^1	c^1	$0.5 \cdot 0.5 = 0.25$
a^1	b^1	c^2	$0.5 \cdot 0.7 = 0.35$
a^1	b^2	c^1	$0.8 \cdot 0.1 = 0.08$
a^1	b^2	c^2	$0.8 \cdot 0.2 = 0.16$
a^2	b^1	c^1	$0.1 \cdot 0.5 = 0.05$
a^2	b^1	c^2	$0.1 \cdot 0.7 = 0.07$
a^2	b^2	c^1	$0 \cdot 0.1 = 0$
a^2	b^2	c^2	$0 \cdot 0.2 = 0$
a^3	b^1	c^1	$0.3 \cdot 0.5 = 0.15$
a^3	b^1	c^2	$0.3 \cdot 0.7 = 0.21$
a^3	b^2	c^1	$0.9 \cdot 0.1 = 0.09$
a^3	b^2	c^2	$0.9 \cdot 0.2 = 0.18$

Sum-out Operation

$$\sum_{a \in A} \phi(A, B) = \psi(B)$$

$\sum_{a \in A}$ $\phi(A, B)$

a^0	b^0	30
a^0	b^1	5
a^1	b^0	1
a^1	b^1	10

$=$ $\psi(B)$

b^0	
b^1	

Idea: Distribute Sums over Products

Example:

$\phi(A, B)$

a^0	b^0	30 (w_0)
a^0	b^1	5 (w_1)
a^1	b^0	1 (w_2)
a^1	b^1	10 (w_3)

$\phi(B, C)$

b^0	c^0	100 (w_4)
b^0	c^1	1 (w_5)
b^1	c^0	1 (w_6)
b^1	c^1	100 (w_7)

$$\begin{aligned} Z &= \sum_{b \in B} \sum_{a \in A} \sum_{c \in C} \phi(a, b)\phi(b, c) && (1) \\ &= \sum_{b \in B} \sum_{a \in A} \phi(a, b) \left\{ \sum_{c \in C} \phi(b, c) \right\} && (2) \\ &= \sum_{b \in B} \psi(b) \left\{ \sum_{a \in A} \phi(a, b) \right\} && (3) \\ &= \sum_{b \in B} \psi(b)\chi(b) && (4) \end{aligned}$$

Idea: Distribute Sums over Products

Example:

$\phi(A, B)$

a^0	b^0	30 (w_0)
a^0	b^1	5 (w_1)
a^1	b^0	1 (w_2)
a^1	b^1	10 (w_3)

$\phi(B, C)$

b^0	c^0	100 (w_4)
b^0	c^1	1 (w_5)
b^1	c^0	1 (w_6)
b^1	c^1	100 (w_7)

$$
\begin{aligned}
Z &= w_0 w_4 + w_0 w_5 + w_1 w_6 + w_1 w_7 \\
&\quad + w_2 w_4 + w_2 w_5 + w_3 w_6 + w_3 w_7 \\
&= w_0(w_4 + w_5) + w_1(w_6 + w_7) \\
&\quad + w_2(w_4 + w_5) + w_3(w_6 + w_7) \\
&= (w_0 + w_2)(w_4 + w_5) \\
&\quad + (w_1 + w_3)(w_6 + w_7)
\end{aligned}
$$

Bucket Elimination

Intuitive data-structure for performing variable elimination!

- Let $X_1, \ldots, X_n$ be an ordering of the variables and let Φ denote the current set (database) of functions
- Associate each variable with a bucket
- Process the buckets from top to bottom $X_1, \ldots, X_n$
  - Put all functions in Φ that mention X_i in the bucket of X_i and remove them from Φ
  - Compute the product of all functions in the bucket of X_i. Let us call the resulting function ψ
  - Sum-out X_i from ψ. Call the new function χ
  - Add χ to Φ

Important: At the end, Φ will contain one or more functions with one entry. The partition function equals the product of these entries.

Bucket Elimination: Example and Animation