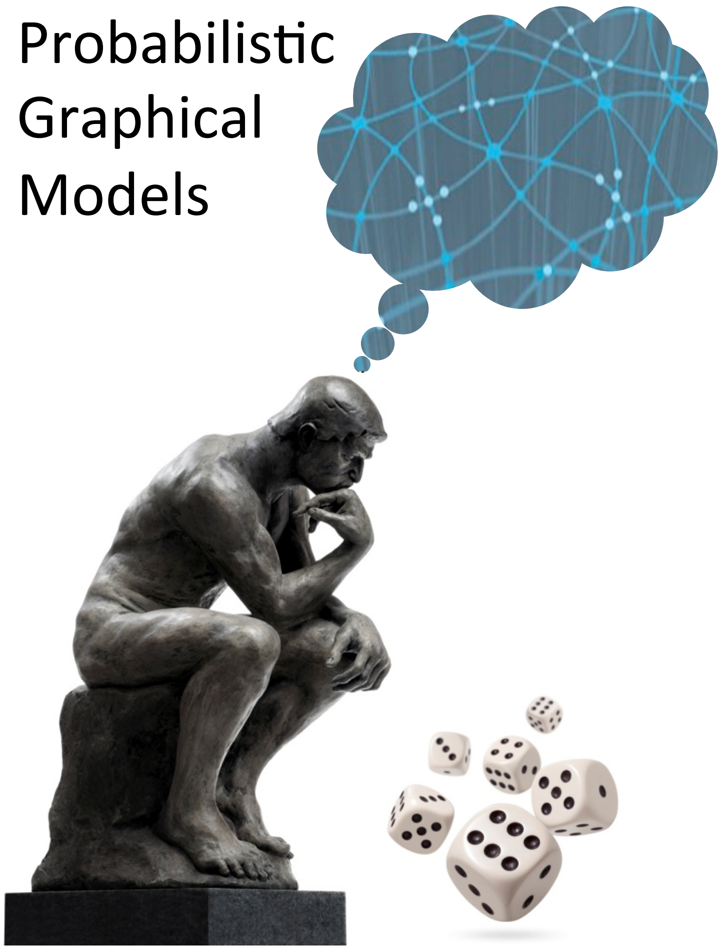
Probabilistic
Graphical
Models



Inference

Overview

Conditional Probability Queries

Conditional Probability Queries

- Evidence: $\mathbf{E} = \boldsymbol{e}$
- Query: a subset of variables $\mathbf{Y}$
- Task: compute $P(\mathbf{Y} \mid \mathbf{E}=\boldsymbol{e})$

- Applications
  - Medical/fault diagnosis
  - Pedigree analysis

NP-Hardness

worst case
general case

The following are all NP-hard

- Given a PGM P_Φ, a variable X and a value $x \in Val(X)$, compute $P_\Phi(X=x)$
  - Or even decide if $P_\Phi(X=x) > 0$
- Let $\varepsilon < 0.5$. Given a PGM P_Φ, a variable X and a value $x \in Val(X)$, and observation $\mathbf{e} \in Val(\mathbf{E})$, find a number p that has $|P_\Phi(X=x|\mathbf{E}=\mathbf{e}) - p| < \varepsilon$

Sum-Product

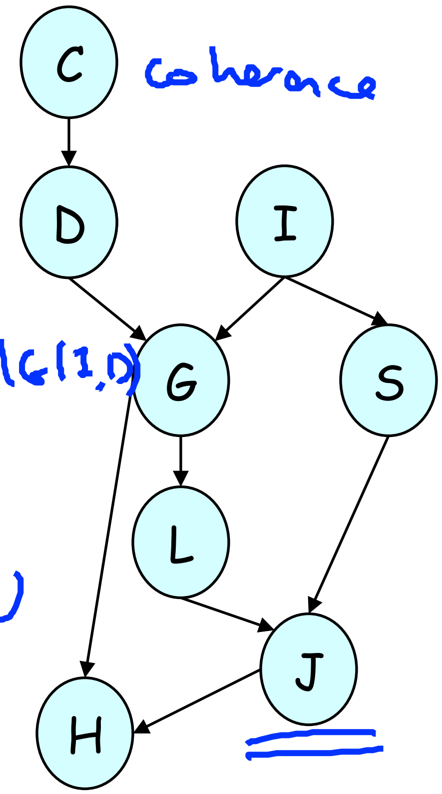
P(J)

$$\sum_{C,D,I,G,S,L,H} \phi_C(C)\phi_D(C,D)\phi_I(I)\phi_G(G,I,D)$$

$$\underbrace{\phi_S(S,I)\phi_L(L,G)\phi_J(J,L,S)\phi_H(H,G,J)}_{\text{joint dist. chain rule}}$$

P(C) coherence

P(G|I,D)



Daphne Koller

Sum-Product

$\tilde{P}(D) =$

$$\cancel{\frac{1}{Z}} \sum_{A,B,C} \underbrace{\phi_1(A,B)\phi_2(B,C)\phi_3(C,D)\phi_4(A,D)}_{\tilde{P}(A,B,C,D)}$$

$P(D) = \frac{1}{Z} \tilde{P}(D)$

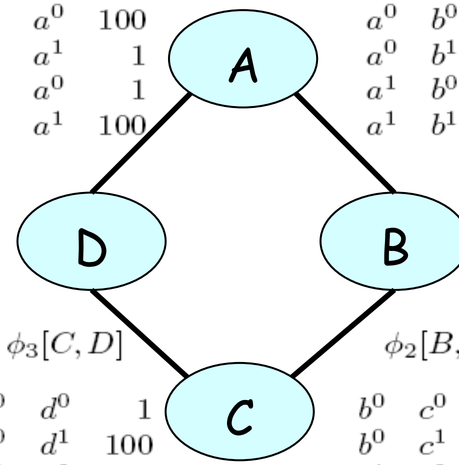
renormalization

$\phi_4[D, A]$

d^0	a^0	100
d^0	a^1	1
d^1	a^0	1
d^1	a^1	100

$\phi_1[A, B]$

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



$\phi_3[C, D]$

c^0	d^0	1
c^0	d^1	100
c^1	d^0	100
c^1	d^1	1

$\phi_2[B, C]$

b^0	c^0	100
b^0	c^1	1
b^1	c^0	1
b^1	c^1	100

Daphne Koller

Evidence: Reduced Factors

$$P(\mathbf{Y} \mid \mathbf{E} = \mathbf{e}) = \frac{P(\mathbf{Y}, \mathbf{E} = \mathbf{e})}{P(\mathbf{E} = \mathbf{e})}$$

$\boxed{W = \{X_1, \ldots, X_n\} - Y - E}$

query

evidence

$$P(\mathbf{Y}, \mathbf{E} = \mathbf{e}) = \sum_{\mathbf{W}} P(\mathbf{Y}, \mathbf{W}, \mathbf{E} = \mathbf{e})$$

$$= \sum_{\mathbf{W}} \frac{1}{Z} \prod_k \phi_k(\mathbf{D}_k, \mathbf{E} = \mathbf{e})$$

$$\propto \sum_{\mathbf{W}} \cancel{\frac{1}{Z}} \prod_k \phi'_k(\mathbf{D}'_k)$$

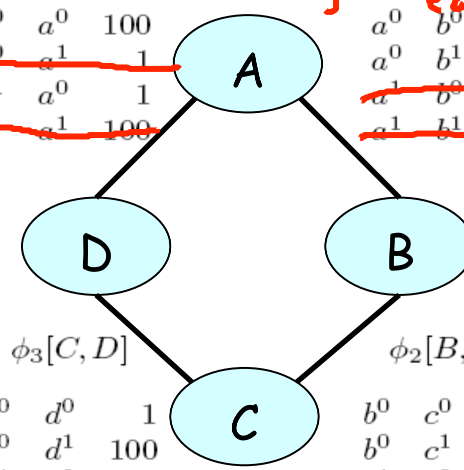
renormalize

$\phi_4[D, A]$

d^0	a^0	100
~~d^0~~	~~a^1~~	~~1~~
d^1	a^0	1
~~d^1~~	~~a^1~~	~~100~~

$\phi_1[A, B]$

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



$\phi_3[C, D]$

c^0	d^0	1
c^0	d^1	100
c^1	d^0	100
c^1	d^1	1

$\phi_2[B, C]$

b^0	c^0	100
b^0	c^1	1
b^1	c^0	1
b^1	c^1	100

$A = a^0$

Daphne Koller

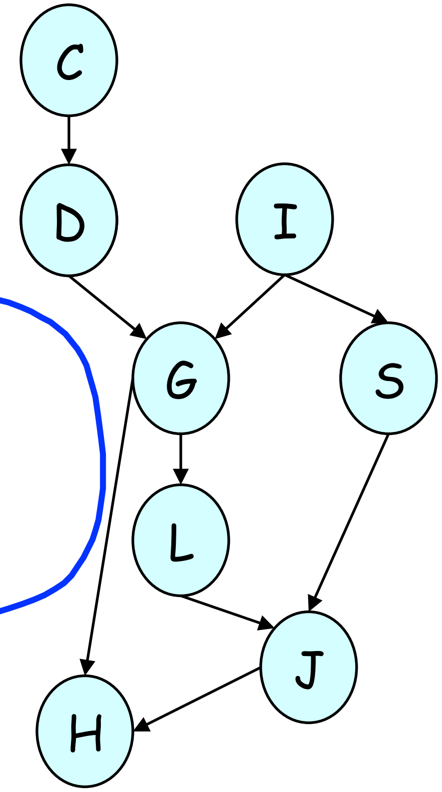
Evidence: Reduced Factors

P(J | I=i, H=h)

P(J,I=i, H=h) =

$$\sum_{C,D,I,G,S,L,H} \phi_C(C)\phi_D(C,D)\phi_I(\overset{i}{\cancel{I}})\phi_G(G,\overset{i}{\cancel{I}},D)\phi_S(S,\overset{i}{\cancel{I}})\phi_L(L,G)\phi_J(J,L,S)\phi_H(\overset{h}{\cancel{H}},G,J)$$

renormalize



Sum-Product

$$P(\boldsymbol{Y} \mid \boldsymbol{E} = \boldsymbol{e}) = \frac{P(\boldsymbol{Y}, \boldsymbol{E} = \boldsymbol{e})}{P(\boldsymbol{E} = \boldsymbol{e})}$$

reduced

$$P(Y, \boldsymbol{E} = \boldsymbol{e}) = \sum_{\boldsymbol{W}} \frac{1}{Z} \prod_k \phi'_k(\boldsymbol{D}'_k)$$ numerator

query

$$P(\boldsymbol{E} = \boldsymbol{e}) = \sum_{\boldsymbol{Y}} \sum_{\boldsymbol{W}} \frac{1}{Z} \prod_k \phi'_k(\boldsymbol{D}'_k)$$ denominator

Compute $\sum_{\boldsymbol{W}} \prod_k \phi'_k(\boldsymbol{D}'_k)$ and renormalize

normalizing

Algorithms: Conditional Probability

- Push summations into factor product
  - Variable elimination (dynamic programming)
- Message passing over a graph
  - Belief propagation
  - Variational approximations

exact
approx

- Random sampling instantiations
  - Markov chain Monte Carlo (MCMC)
  - Importance sampling

approx

Daphne Koller

Summary

- Conditional probability queries of subset of variables given evidence on others
- Summing over factor product
- Evidence simply reduces the factors
- Many exact and approximate algorithms