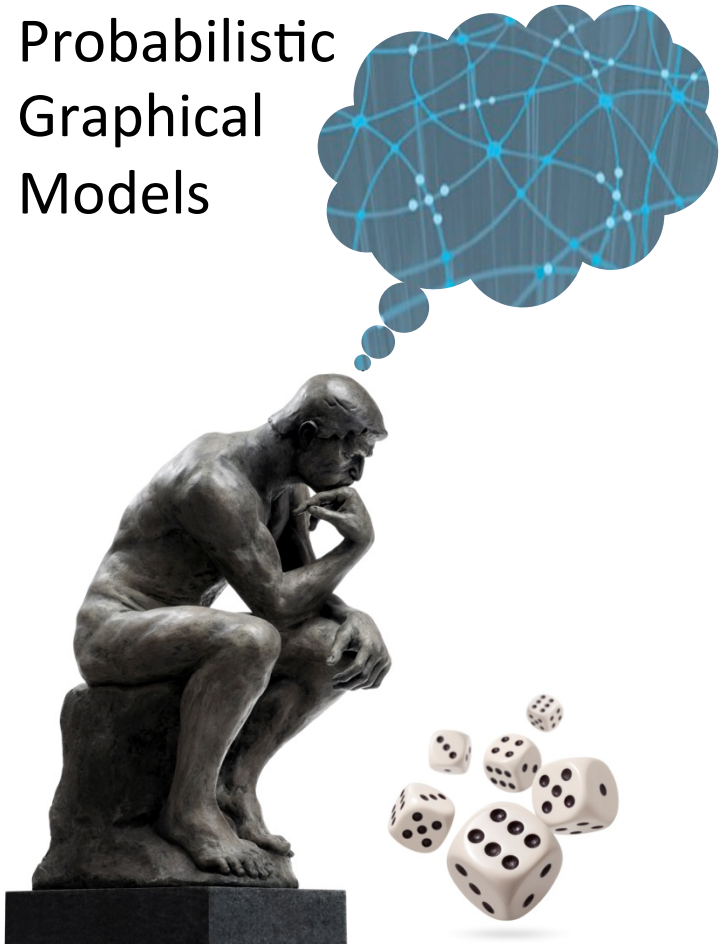
Probabilistic Graphical Models



Inference

Overview

MAP Inference

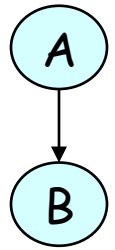
Maximum a Posteriori (MAP)

- Evidence: $\mathbf{E}=\boldsymbol{e}$
- Query: all other variables $\mathbf{Y}$ ($\mathbf{Y}=\{X_1,\ldots,X_n\}-\mathbf{E}$)
- Task: compute $MAP(\mathbf{Y}|\mathbf{E}=\boldsymbol{e}) = \text{argmax}_{\mathbf{y}} P(\mathbf{Y}=\mathbf{y} \mid \mathbf{E}=\boldsymbol{e})$
  - Note: there may be more than one possible solution

- Applications
  - Message decoding: most likely transmitted message
  - Image segmentation: most likely segmentation

MAP ≠ Max over Marginals

- Joint distribution
  - $P(a^0, b^0) = 0.04$
  - $P(a^0, b^1) = 0.36$
  - $P(a^1, b^0) = 0.3$
  - $P(a^1, b^1) = 0.3$



P(A)

I	
a^0	a^1
0.4	0.6

P(B|A)

	B	
A	B^0	B^1
a^0	0.1	0.9
a^1	0.5	0.5

Daphne Koller

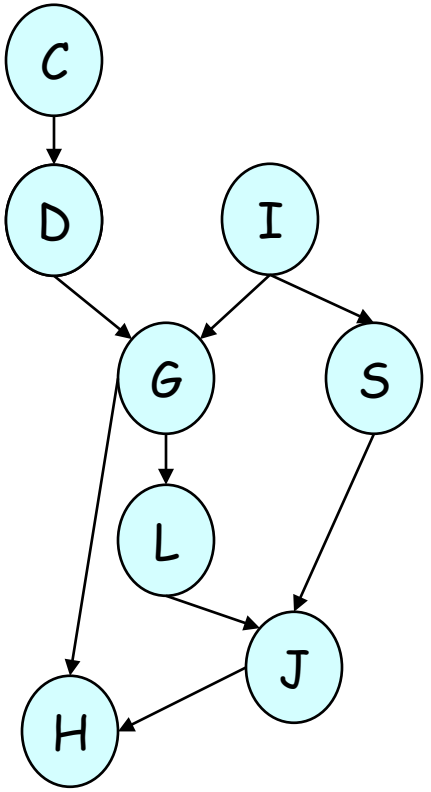
NP-Hardness

The following are NP-hard

- Given a PGM P_Φ, find a joint assignment $\mathbf{x}$ with highest probability $P_\Phi(\mathbf{x})$
- Given a PGM P_Φ and a probability p, decide if there is an assignment $\mathbf{x}$ such that $P_\Phi(\mathbf{x}) > p$

Max-Product

$$\text{argmax}_{C,D,I,G,S,G,L,J,H} \; \phi_C(C)\phi_D(C,D)\phi_I(I)\phi_G(G,I,D)$$
$$\phi_S(S,I)\phi_L(L,G)\phi_J(J,L,S)\phi_H(H,G,J)$$



Max-Product

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

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

constant wrt $\boldsymbol{Y}$

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

partition — constant wrt $\boldsymbol{Y}$ — reduced relative to $\boldsymbol{E} = \boldsymbol{e}$

$$\propto \prod_k \phi'_k(\boldsymbol{D}'_k)$$

$$\text{argmax}_{\boldsymbol{Y}} P(\boldsymbol{Y} \mid \boldsymbol{E} = \boldsymbol{e}) = \text{argmax}_{\boldsymbol{Y}} \prod_k \phi'_k(\boldsymbol{D}'_k)$$

Algorithms: MAP

optimization

- Push maximization into factor product
  - Variable elimination — Max product
- Message passing over a graph
  - Max-product belief propagation
- Using methods for integer programming
- For some networks: graph-cut methods
- Combinatorial search

Summary

- MAP: single coherent assignment of highest probability
  - Not the same as maximizing individual marginal probabilities
- Maxing over factor product
- Combinatorial optimization problem
- Many exact and approximate algorithms