Probabilistic
Graphical
Models

Inference

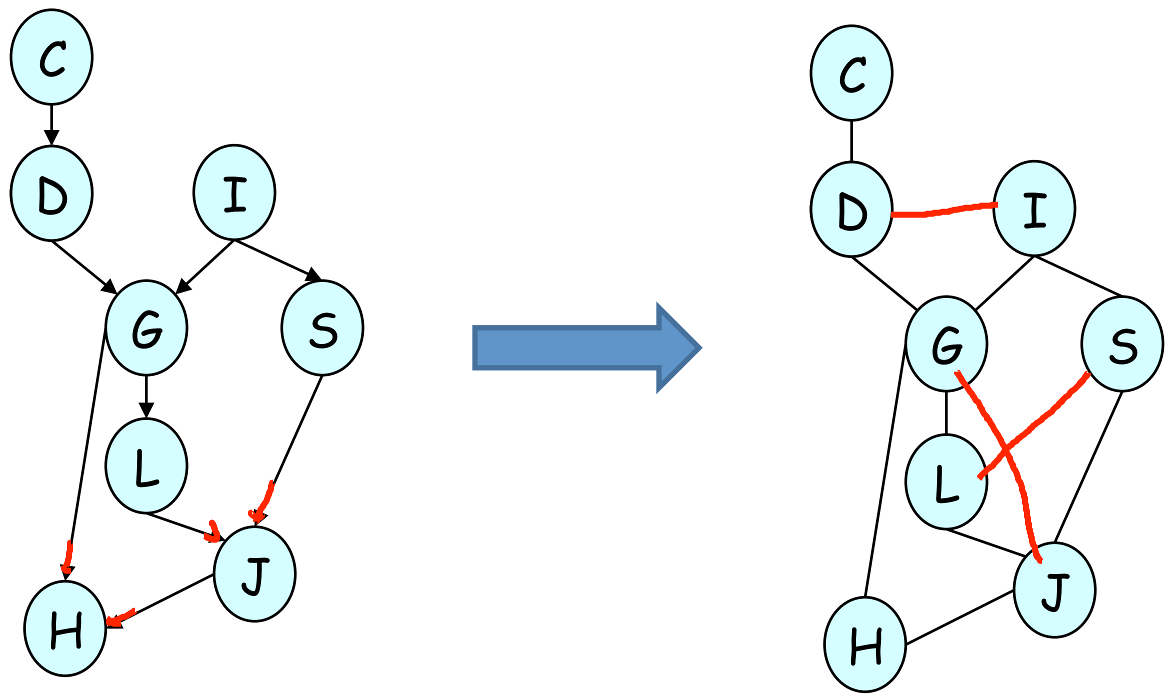
Variable Elimination

Graph-Based Perspective

Initial Graph

$$\phi_J(J,L,S)\phi_L(L,G)\phi_S(S,I)\phi_G(G,I,D)\phi_H(H,G,J)\phi_I(I)\phi_D(C,D)\phi_C(C)$$

moralization



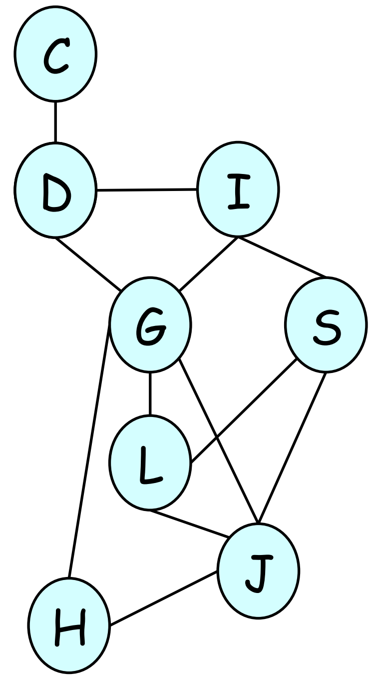
Daphne Koller

Elimination as Graph Operation

$$\phi_J(J,L,S)\phi_L(L,G)\phi_S(S,I)\phi_G(G,I,D)\phi_H(H,G,J)\phi_I(I)\phi_D(C,D)\phi_C(C)$$

- Eliminate: C

$$\tau_1(D) = \sum_C \phi_C(C)\phi_D(C,D)$$



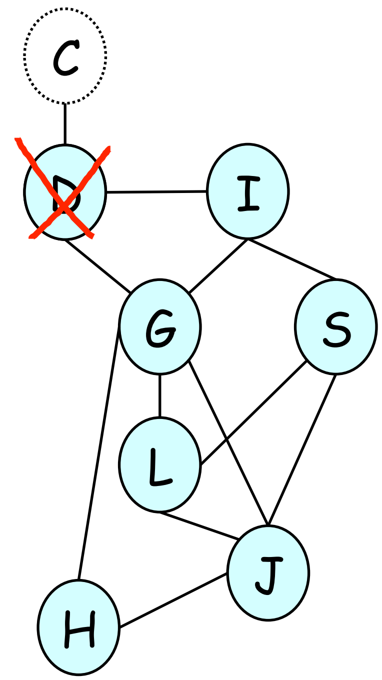
Induced Markov network for the current set of factors

Elimination as Graph Operation

$$\phi_J(J,L,S)\phi_L(L,G)\phi_S(S,I)\phi_G(G,I,D)\phi_H(H,G,J)\phi_I(I)\tau_1(D)$$

- Eliminate: D

$$\tau_2(G,I) = \sum_D \phi_G(G,I,D)\tau_1(D)$$



Induced Markov network for the current set of factors

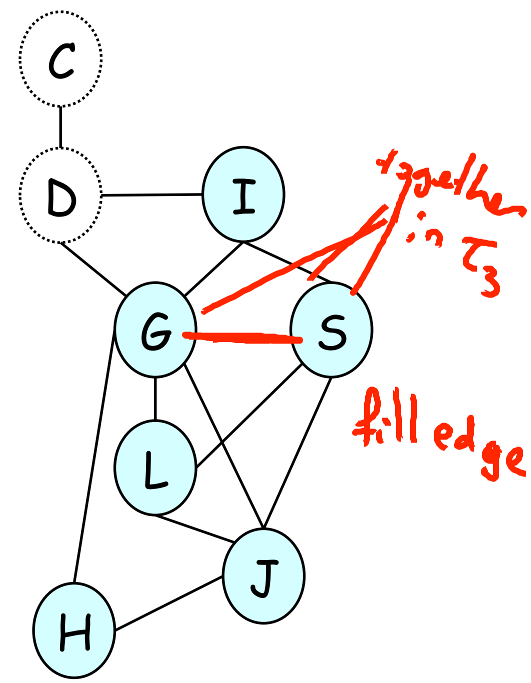
Elimination as Graph Operation

$$\phi_J(J,L,S)\phi_L(L,G)\phi_S(S,I)\phi_I(I)\phi_H(H,G,J)\tau_2(G,I)$$

- Eliminate: I

$$\tau_3(S,G) = \sum_I \phi_S(S,I)\phi_I(I)\tau_2(G,I)$$

all variables connected to I become connected directly



Induced Markov network for the current set of factors

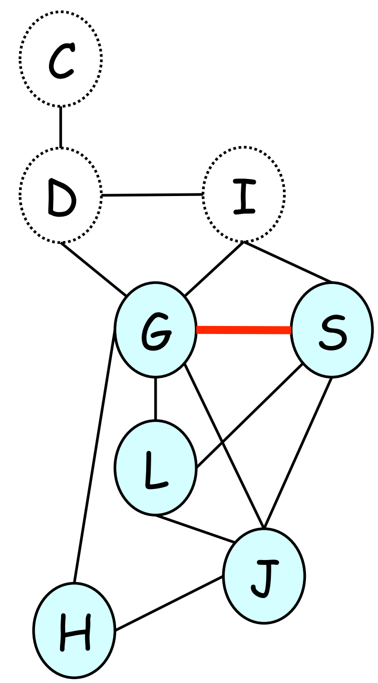
Daphne Koller

Elimination as Graph Operation

$$\phi_J(J,L,S)\phi_L(L,G)\phi_H(H,G,J)\tau_3(S,G)$$

- Eliminate: H

$$\tau_4(G,J)=\sum_H \phi_H(H,G,J)$$

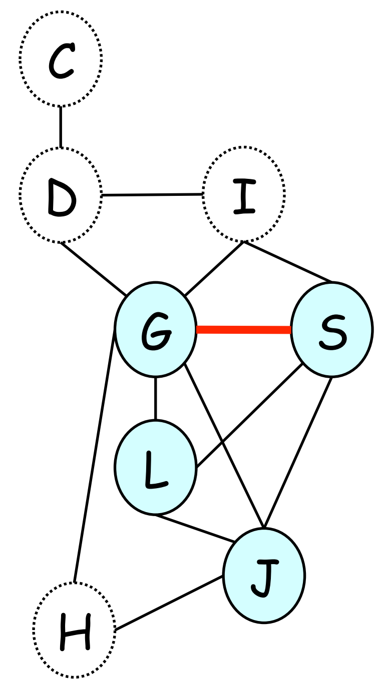


Induced Markov network for the current set of factors

Elimination as Graph Operation

$$\phi_J(J,L,S)\phi_L(L,G)\tau_3(S,G)\tau_4(G,J)$$

- Eliminate: G



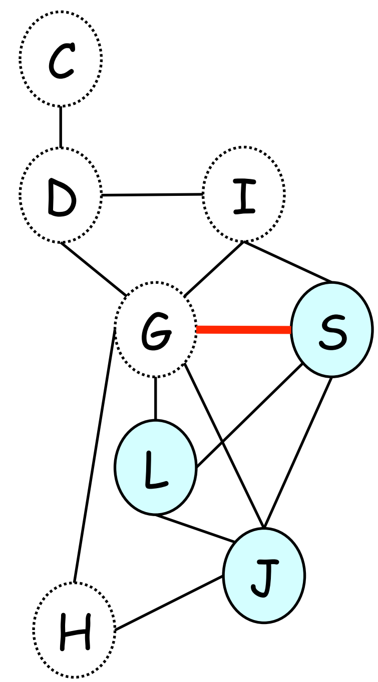
$$\tau_5(L,J) = \sum_G \phi_L(L,G)\tau_3(S,G)\tau_4(G,J)$$

Induced Markov network for the current set of factors

Elimination as Graph Operation

$$\phi_J(J,L,S)\tau_5(L,J)$$

- Eliminate: L,S

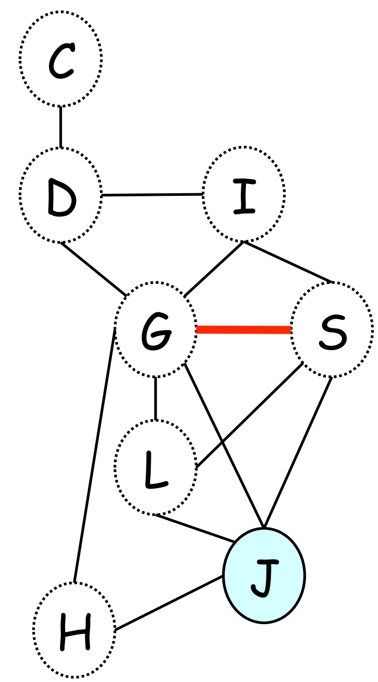


Induced Markov network for the current set of factors

Elimination as Graph Operation

$$\phi_J(J,L,S)\tau_5(L,J)$$

- Eliminate: L,S

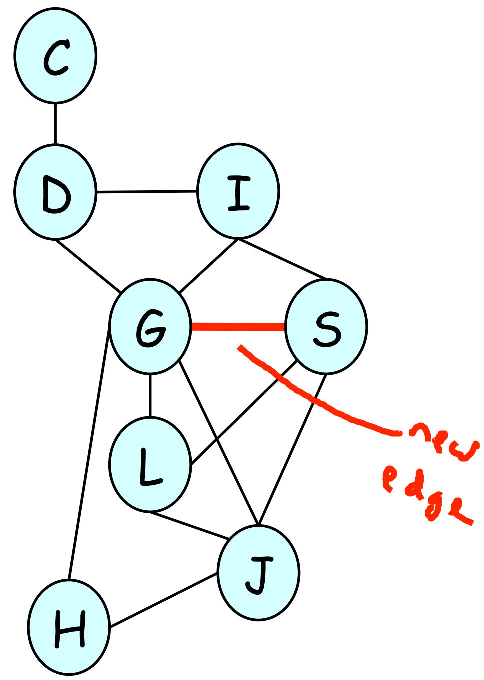


Induced Markov network for the current set of factors

Daphne Koller

Induced Graph

- The induced graph $I_{\Phi,\alpha}$ over factors Φ and ordering α:
  - Undirected graph
  - X_i and X_j are connected if they appeared in the same factor in a run of the VE algorithm using α as the ordering



Cliques in the Induced Graph

maximal fully connected subgraph

- **Theorem:** Every factor produced during VE is a clique in the induced graph

$$\tau_1(D) = \sum_C \phi_C(C)\phi_D(C,D)$$

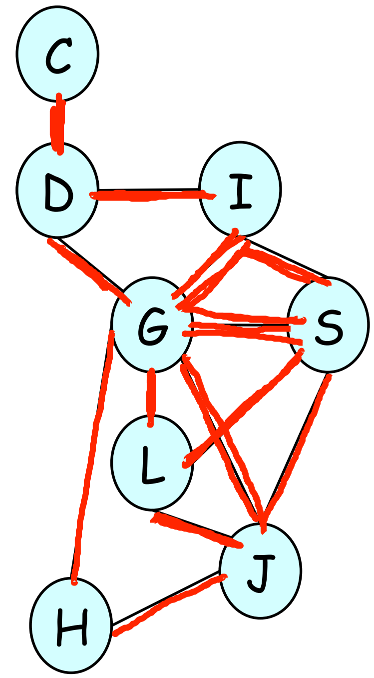
$$\tau_2(G,I) = \sum_D \phi_G(G,I,D)\tau_1(D)$$

$$\tau_3(S,G) = \sum_I \phi_S(S,I)\phi_I(I)\tau_2(G,I)$$

$$\tau_4(G,J) = \sum_H \phi_H(H,G,J)$$

$$\tau_5(L,J) = \sum_G \phi_L(L,G)\tau_3(S,G)\tau_4(G,J)$$

$$\tau_6 = \sum_{L,S} \phi_J(J,L,S)\tau_5(L,J)$$



Daphne Koller

Cliques in the Induced Graph

- **Theorem:** Every (maximal) clique in the induced graph is a factor produced during VE

$$\tau_1(D) = \sum_C \phi_C(C)\phi_D(C,D)$$

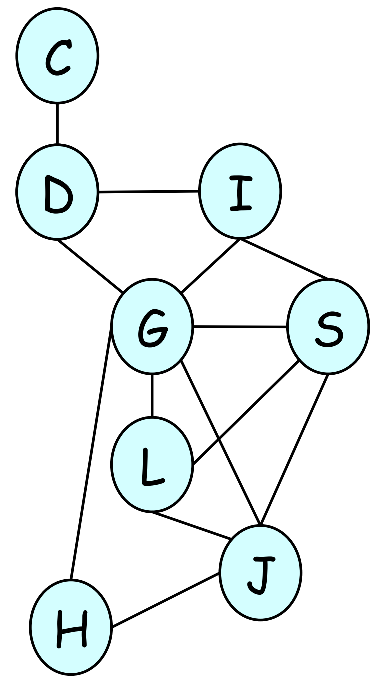
$$\tau_2(G,I) = \sum_D \phi_G(G,I,D)\tau_1(D)$$

$$\tau_3(S,G) = \sum_I \phi_S(S,I)\phi_I(I)\tau_2(G,I)$$

$$\tau_4(G,J) = \sum_H \phi_H(H,G,J)$$

$$\tau_5(L,J) = \sum_G \phi_L(L,G)\tau_3(S,G)\tau_4(G,J)$$

$$\tau_6 = \sum_{L,S} \phi_J(J,L,S)\tau_5(L,J)$$



Daphne Koller

Cliques in the Induced Graph

- **Theorem:** Every (maximal) clique in the induced graph is a factor produced during VE

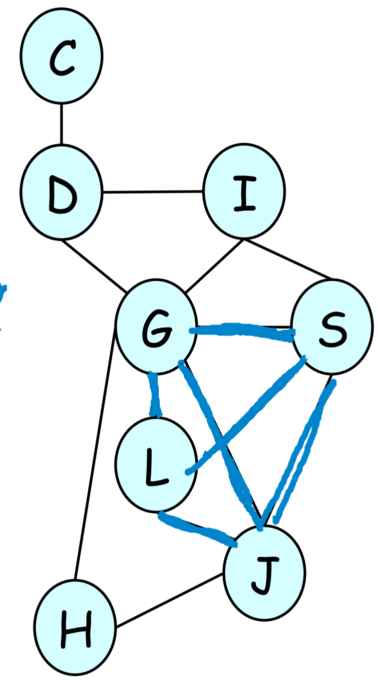
Consider a max clique —
some variable is first to be eliminated

once a variable is eliminated, no new neighbors are added to it

⇒ when eliminated it already had all the clique members as neighbors

⇒ participated in factor with all these other variables

⇒ when multiplied together, we have a factor over all of them

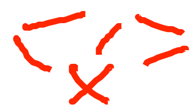


Induced Width

- The width of an induced graph is the number of nodes in the largest clique in the graph minus 1
- Minimal induced width of a graph K is $\min_\alpha(\text{width}(I_{K,\alpha}))$
- Provides a lower bound on best performance of VE to a model factorizing over K

Summary

- Variable elimination can be viewed as transformations on undirected graph
  - Elimination connects all node's current neighbors



- Cliques in resulting induced graph directly correspond to algorithm's complexity