Undirected Graphical Models

Machine Learning
CSx824/ECEx242
Bert Huang
Virginia Tech

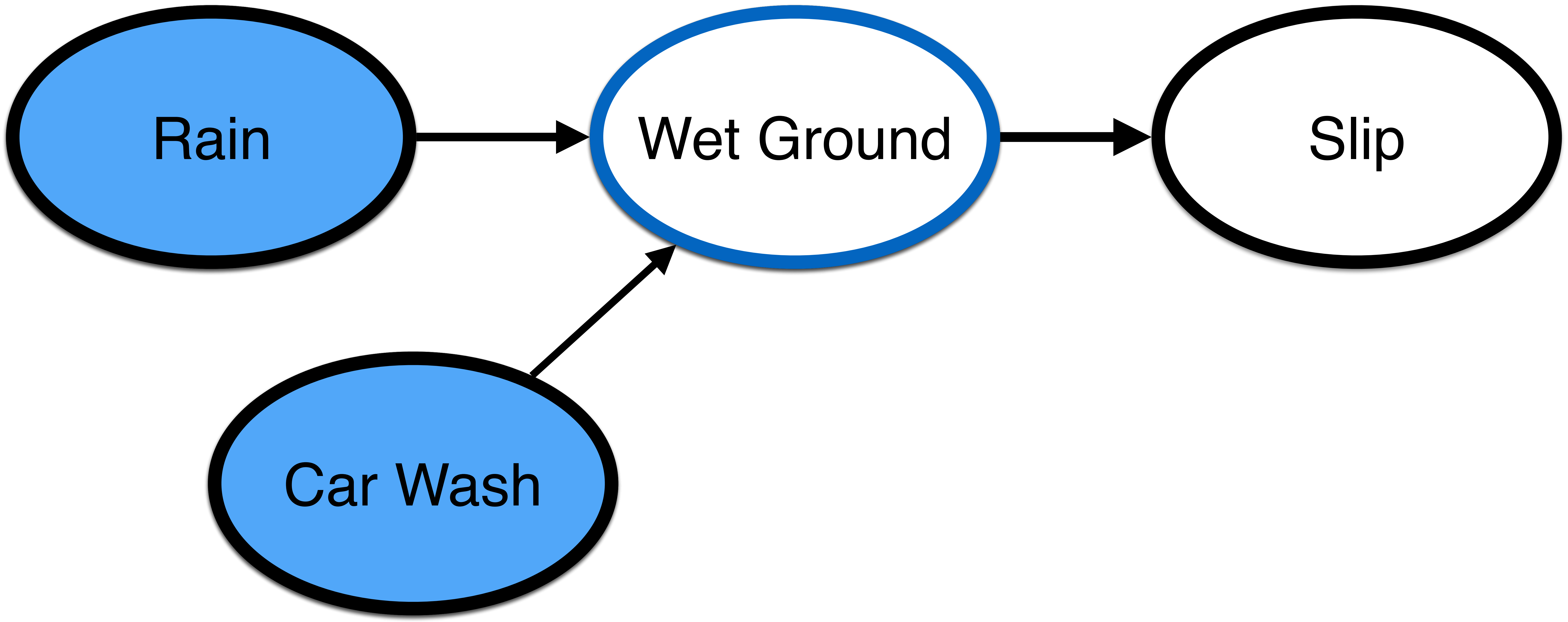
Outline

- Undirected graphical models, Markov random fields
- Independence in MRFs
- Are Bayesian networks MRFs?

Review: Bayesian Networks

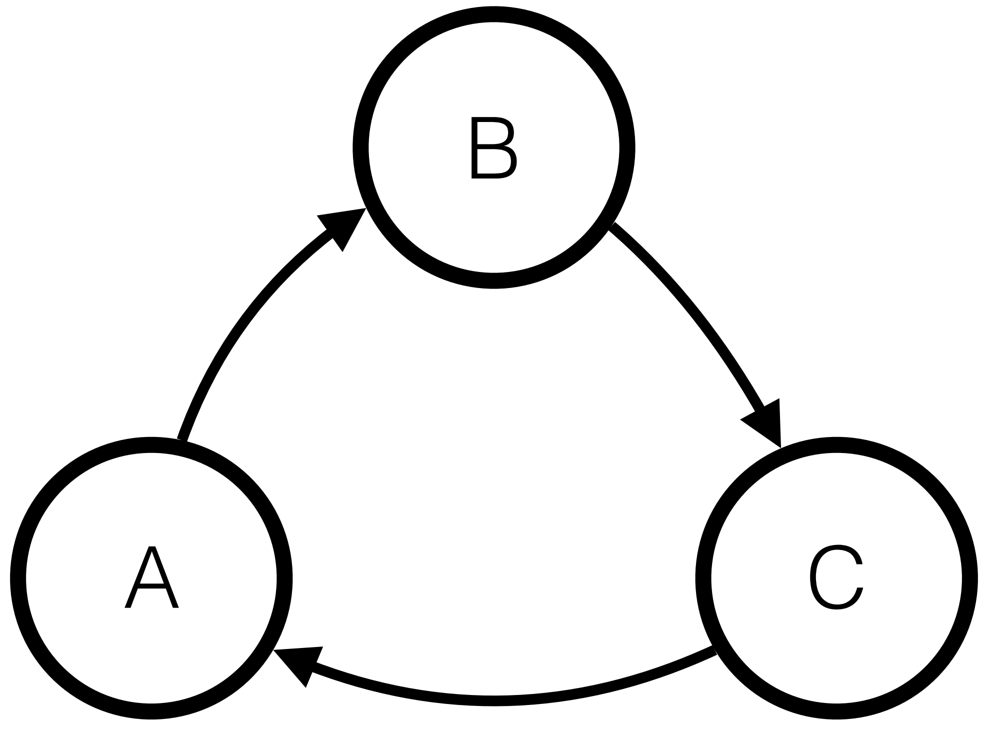
$P(R, W, S, C) = P(R)\,P(C)\,P(W \mid C, R)\,P(S \mid W)$

$P(X \mid Parents(X))$



Acyclicity of Bayes Nets

invalid Bayes net

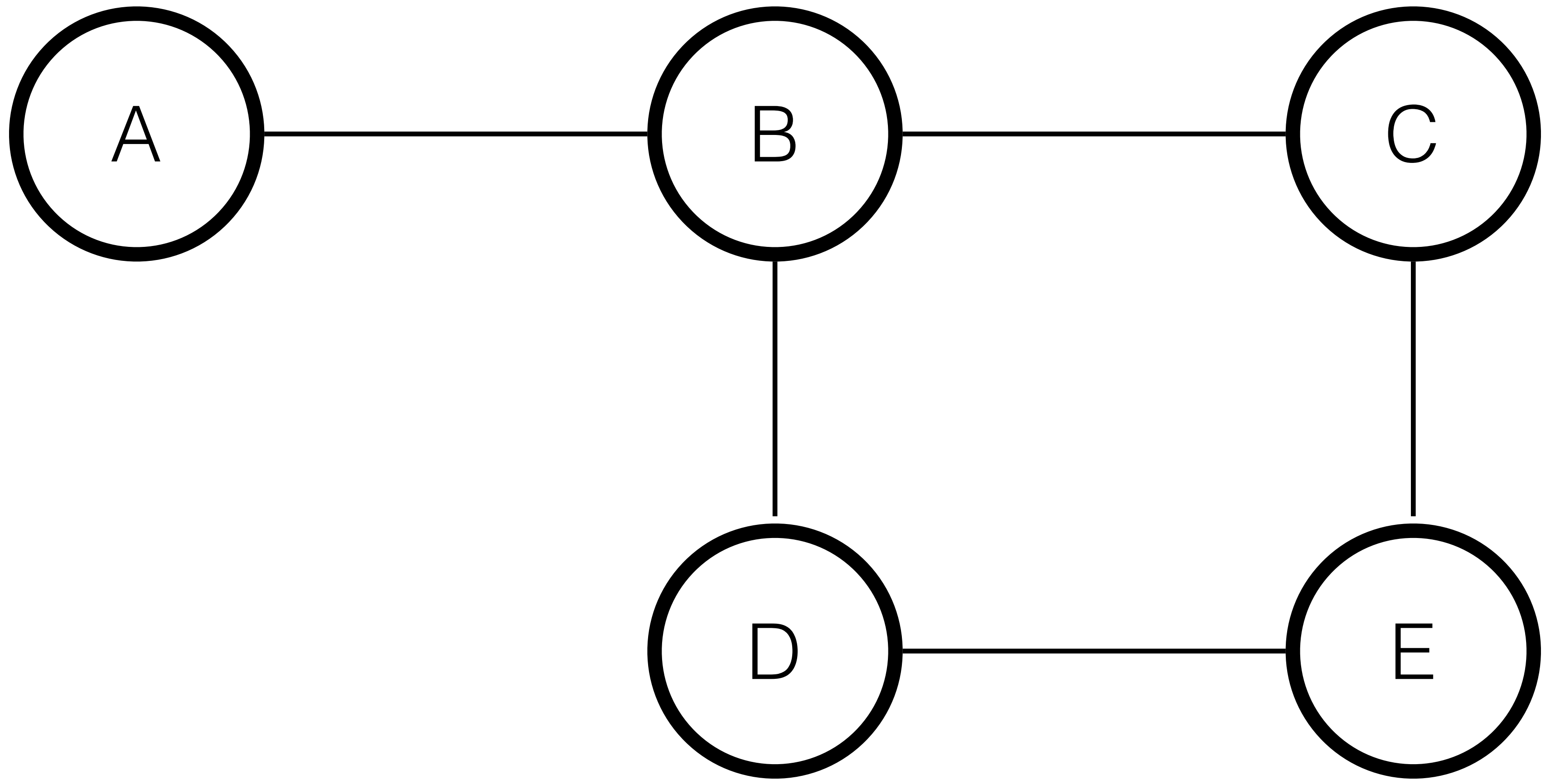


$$P(A, B, C) = P(B|A)P(C|B)P(A|C) = P(B, C|A)P(A|C) = P(A, B, C|C)?$$

Only "makes sense" if P(A) = P(B) = P(C) = 1

All meaningful Bayes nets are directed, acyclic graphs (DAGs)

Undirected Graphical Models



$$P(A, B, C, D, E) \propto \phi(A, B)\phi(B, C)\phi(B, D)\phi(C, E)\phi(D, E)$$

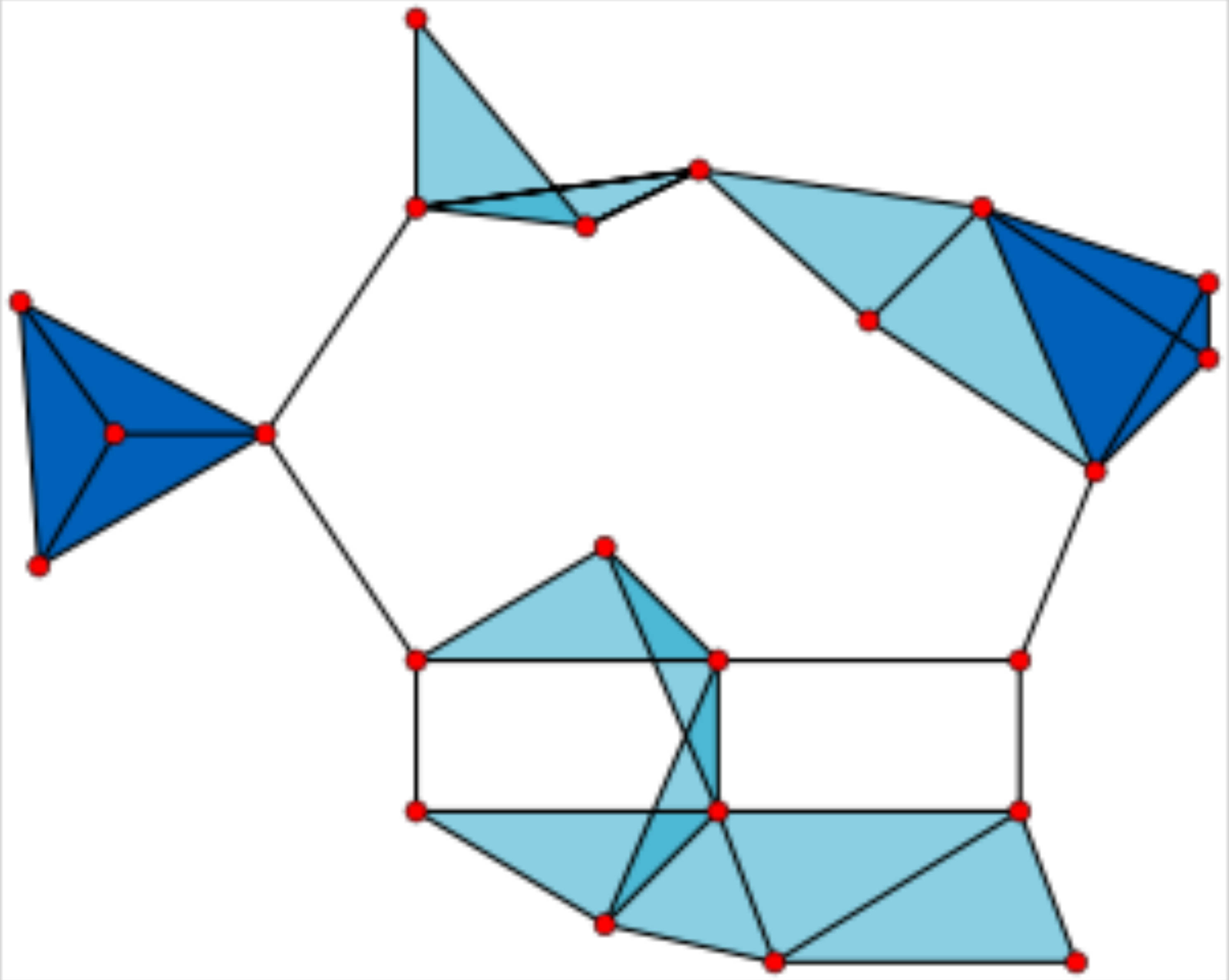
$$P(X) = \frac{1}{Z} \prod_{c \in \text{cliques}(G)} \phi_C(x_c)$$

potential functions

Clique (graph theory)

From Wikipedia, the free encyclopedia

In the mathematical area of graph theory, a **clique** (/ˈkliːk/ or /ˈklɪk/) is subset of vertices of an undirected graph, such that its induced subgraph is complete; that is, every two distinct vertices in the clique are adjacent. Cliques are one of the basic concepts of graph theory and are used in many other mathematical problems and constructions on graphs. Cliques have also been studied in computer science: the task of finding whether there is a clique of a given size in a graph (the clique problem) is NP-complete, but despite this hardness result, many algorithms for finding cliques have been studied.

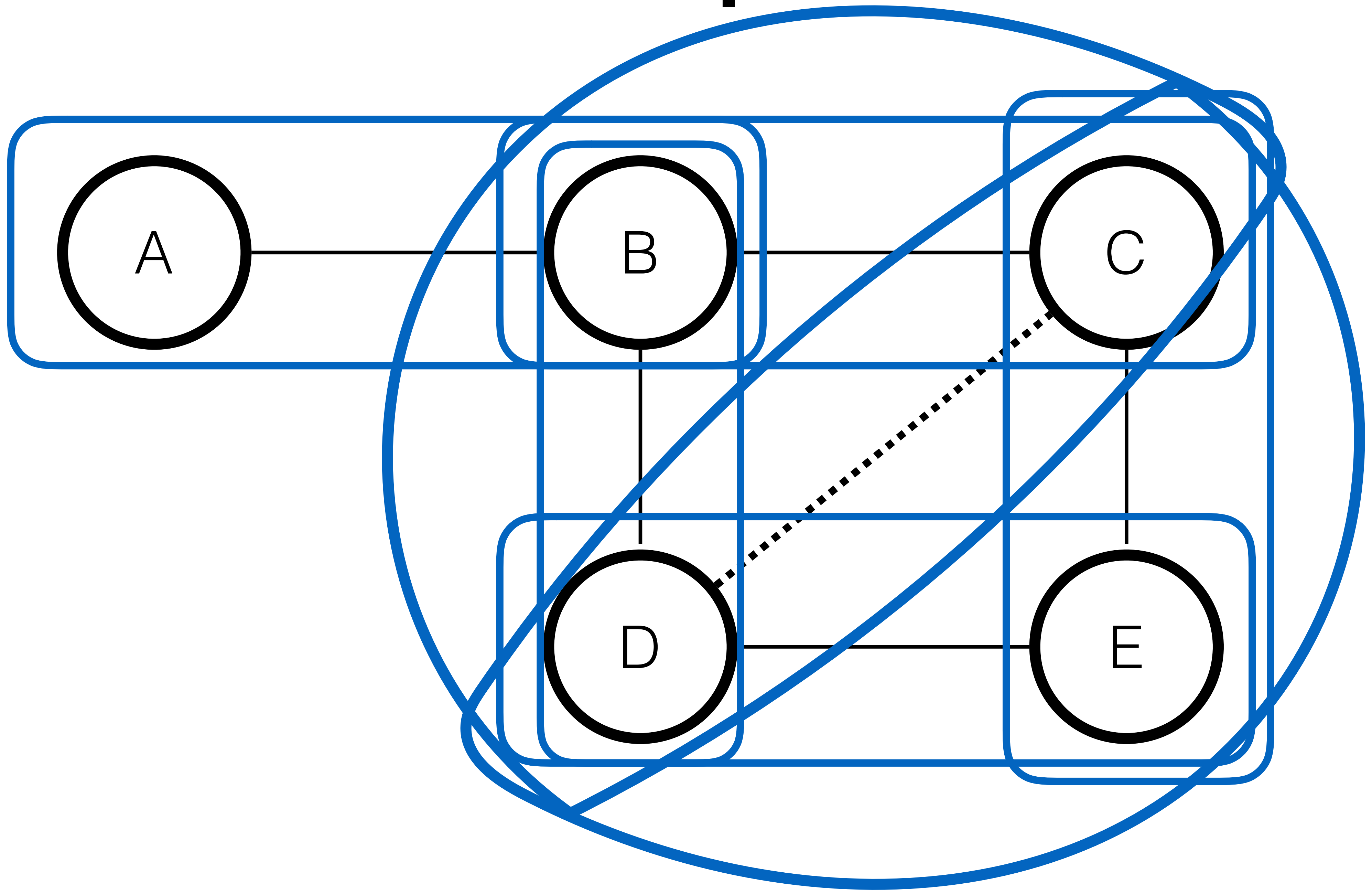


A graph with

- 23 × 1-vertex cliques (the vertices),
- 42 × 2-vertex cliques (the edges),
- 19 × 3-vertex cliques (the light and dark blue triangles), and

Although the study of complete subgraphs goes back at least to the graph-theoretic reformulation of Ramsey theory by Erdős & Szekeres (1935),[1] the term *clique* comes from Luce & Perry (1949), who used complete subgraphs in social networks to model cliques of people; that is, groups of people all of whom

Undirected Graphical Models



$$P(A, B, C, D, E) \propto \phi(A, B)\phi(B, C)\phi(B, D)\phi(C, E)\phi(D, E)$$

$$\phi(A, B)\phi(B, C, D)\phi(C, D, E)$$

$$P(X) = \frac{1}{Z} \prod_{c \in \text{cliques}(G)} \phi_C(x_c)$$

potential functions

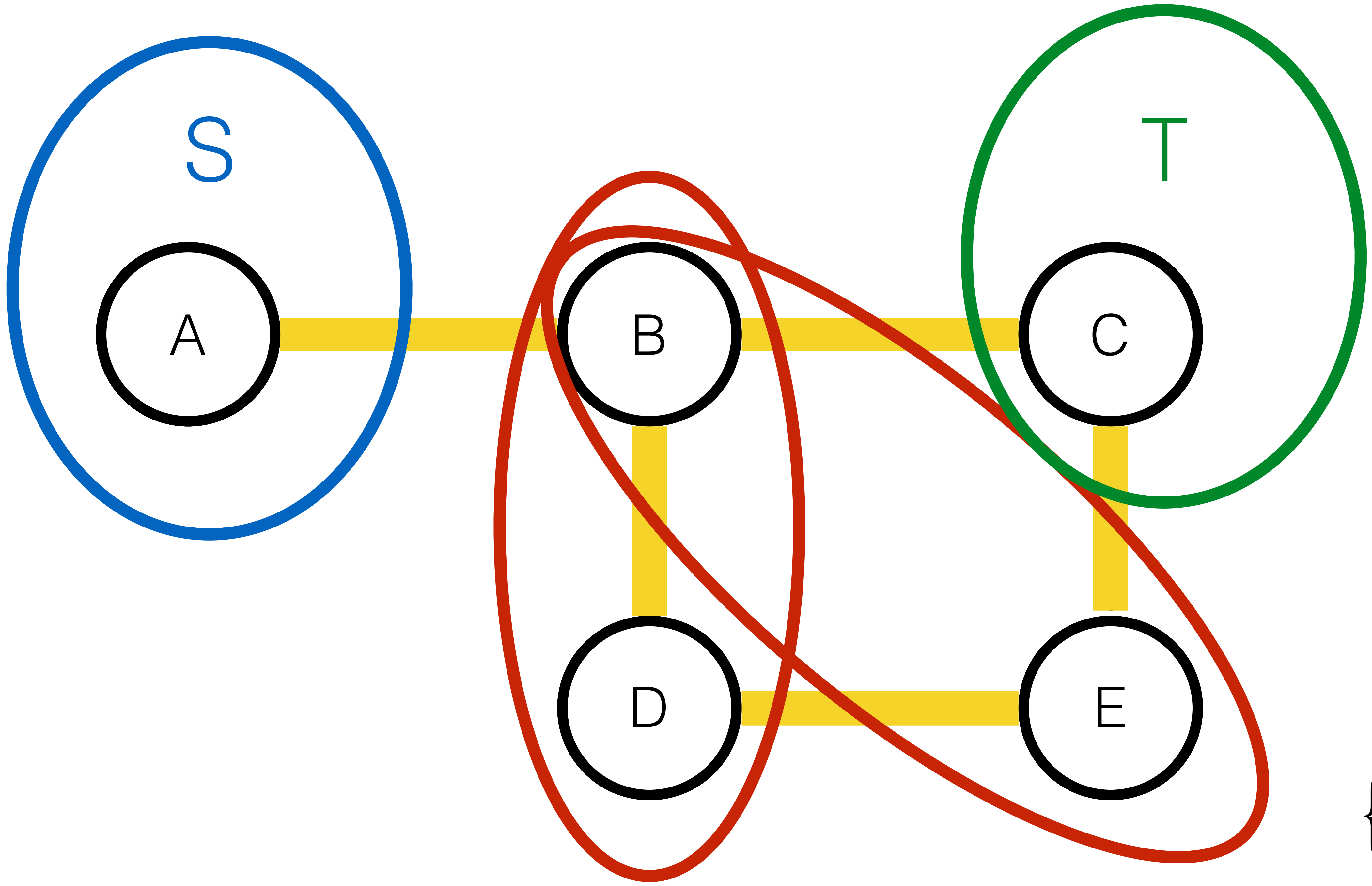
Markov Random Fields

- Any two subsets S and T of variables are conditionally independent given a **separating subset**
  - All paths between S and T must travel through the separating subset

paths:

A-B-C

A-B-D-E-C



separating subsets

{B,D}, {B,E}, {B,D,E}

Independence Corollaries

- Any two non-adjacent variables are conditionally independent given all other variables

- Any variable is conditionally independent of the other variables given its neighbors

    - Markov blanket

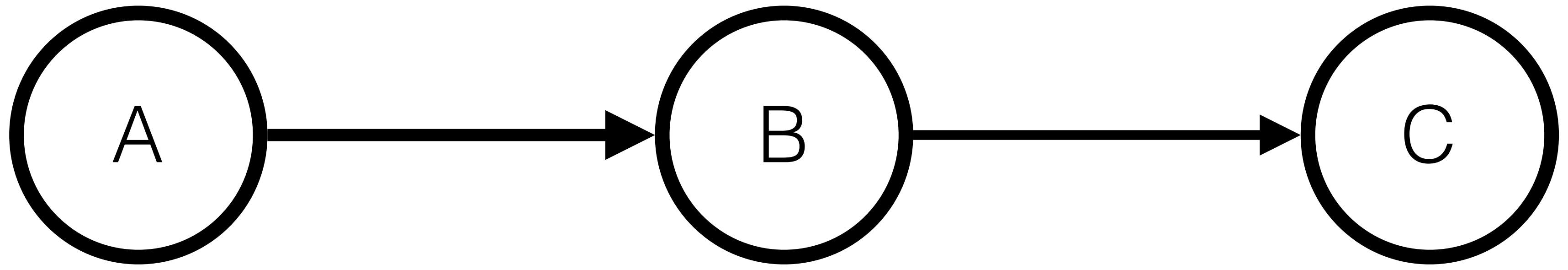
Bayesian Networks as MRFs

$$p(A, B) = p(A)p(B|A)$$

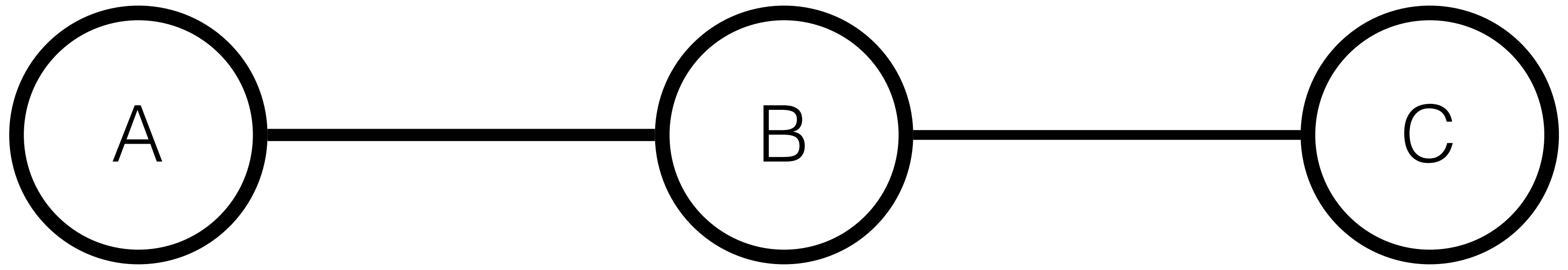
$$p(A, B) \propto \phi(A, B)$$

converting a single edge to a pairwise clique potential is easy

Bayesian Networks as MRFs



$$p(A, B, C) = p(A)p(B|A)P(C|B)$$



$$p(A, B, C) \propto \phi(A, B)\phi(B, C)$$

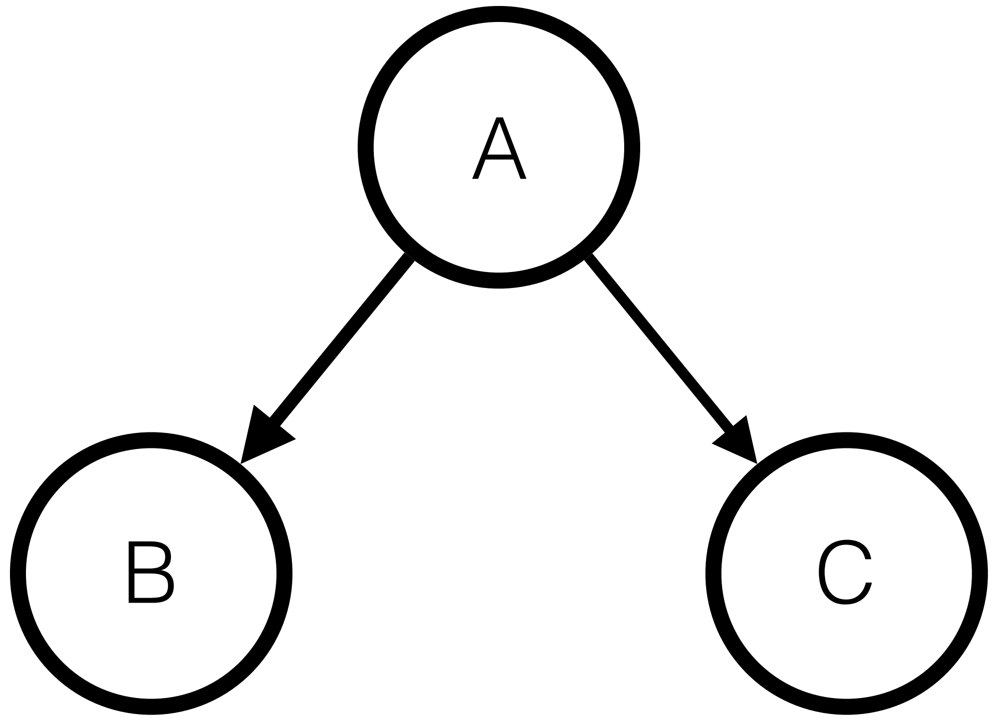
$$\phi(A, B) \leftarrow P(A)P(B|A)$$

$$\phi(B, C) \leftarrow P(C|B)$$

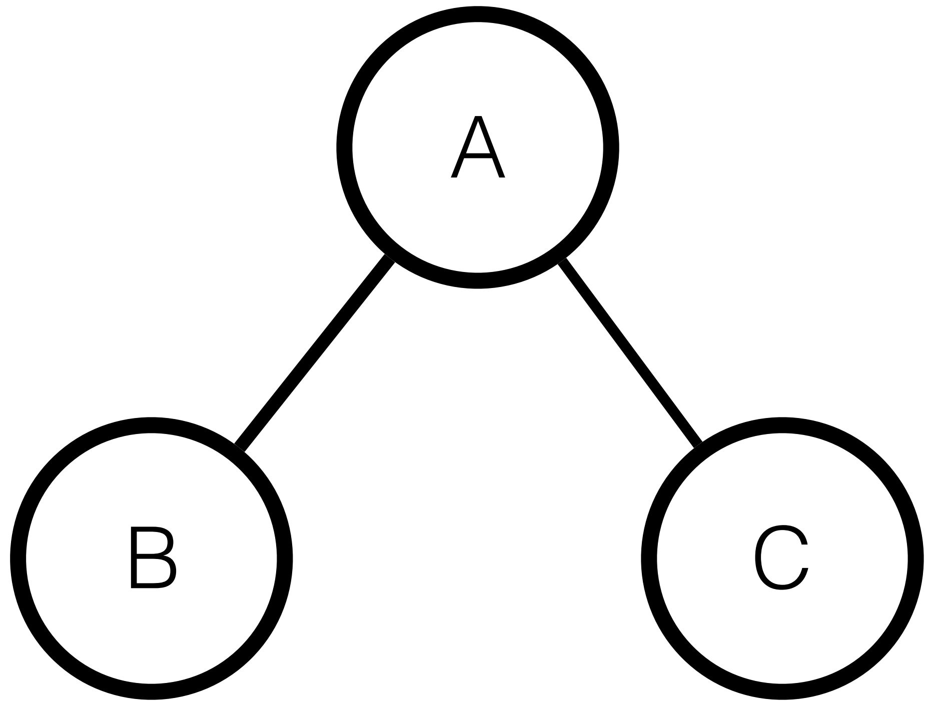
chains are easy too

parameterization is not unique

Bayesian Networks as MRFs



$$p(A, B, C) = p(A)P(B|A)P(C|A)$$



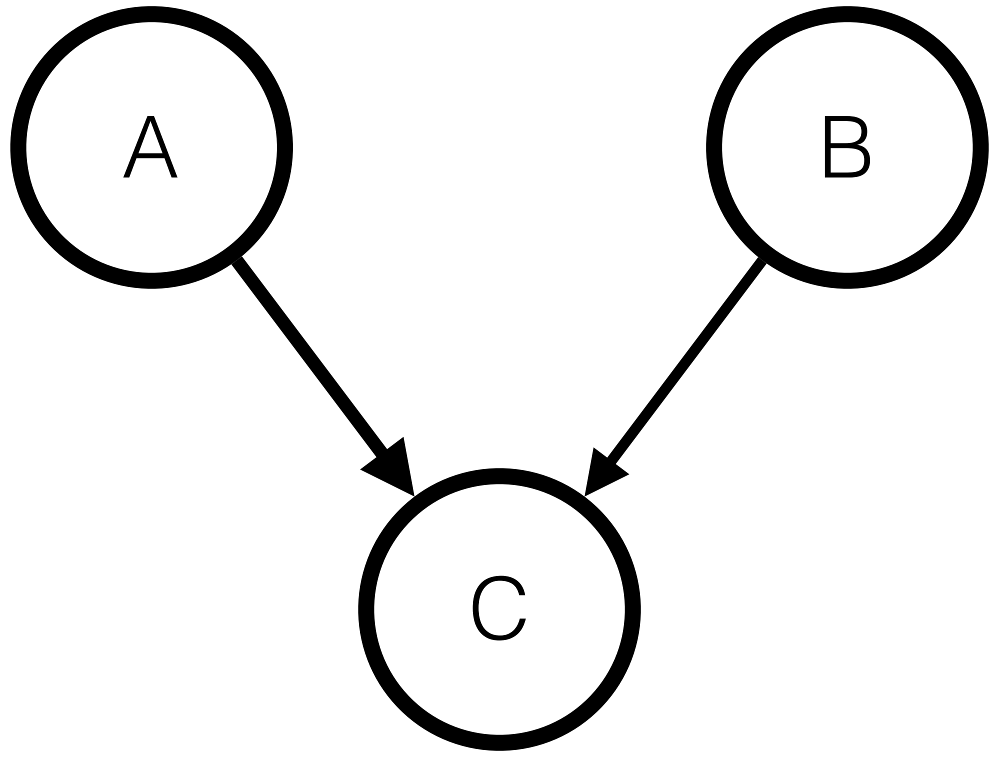
$$p(A, B, C) \propto \phi(A, B), \phi(A, C)$$

$$\phi(A, B) \leftarrow P(A)P(B|A)$$

$$\phi(A, C) \leftarrow P(C|A)$$

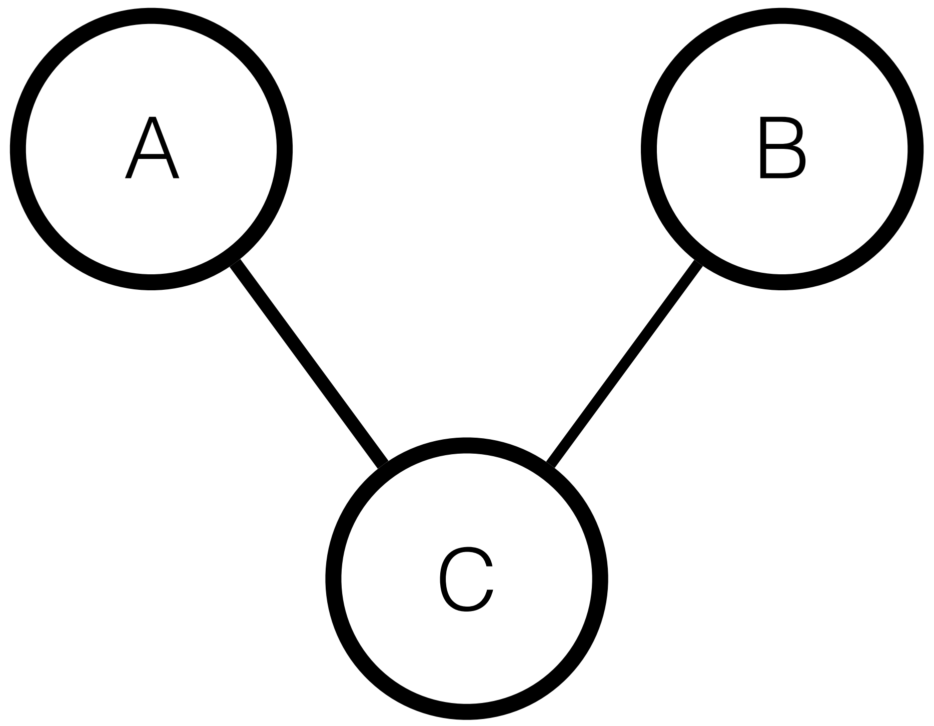
shared parents also easy

Bayesian Networks as MRFs



$$p(A, B, C) = p(A)p(B)p(C|A, B)$$

A and B are **dependent** given C



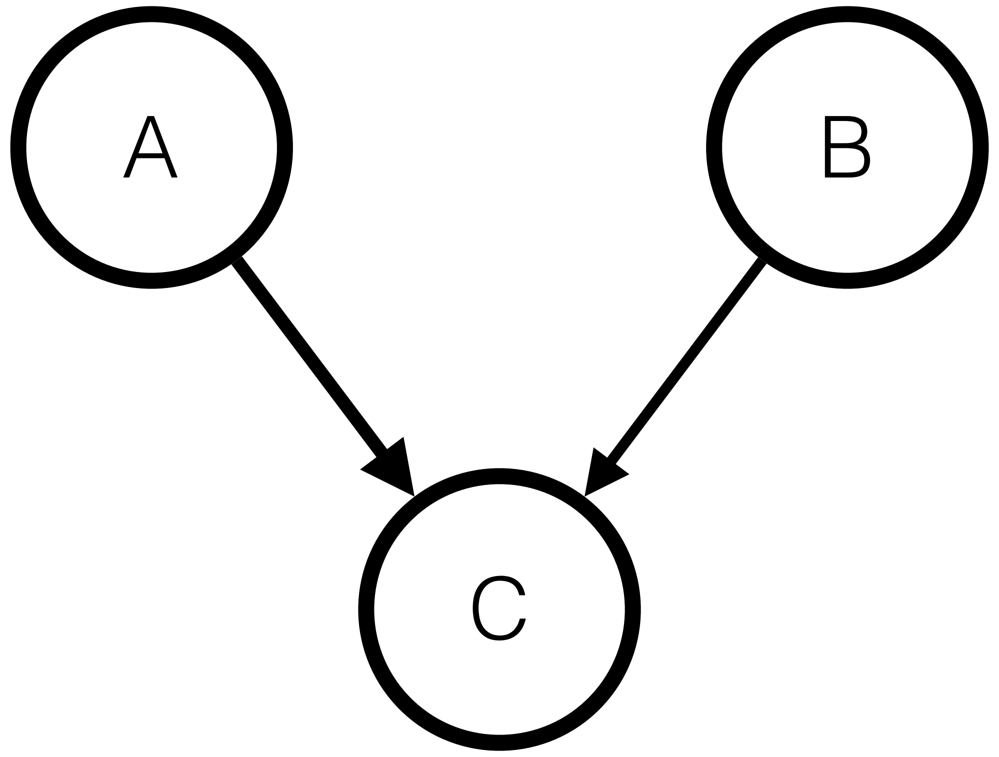
$$p(A, B, C) \propto \phi(A, C)\phi(B, C)$$

A and B are **independent** given C

can’t be correct

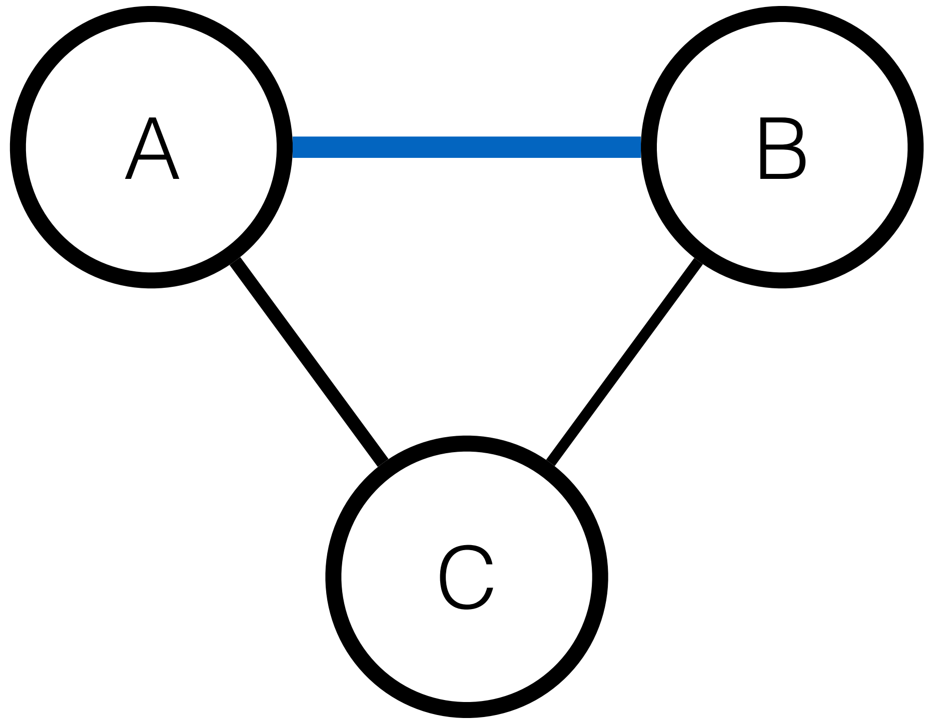
shared child

Moralizing Parents



$$p(A, B, C) = p(A)p(B)p(C|A, B)$$

A and B are **dependent** given C



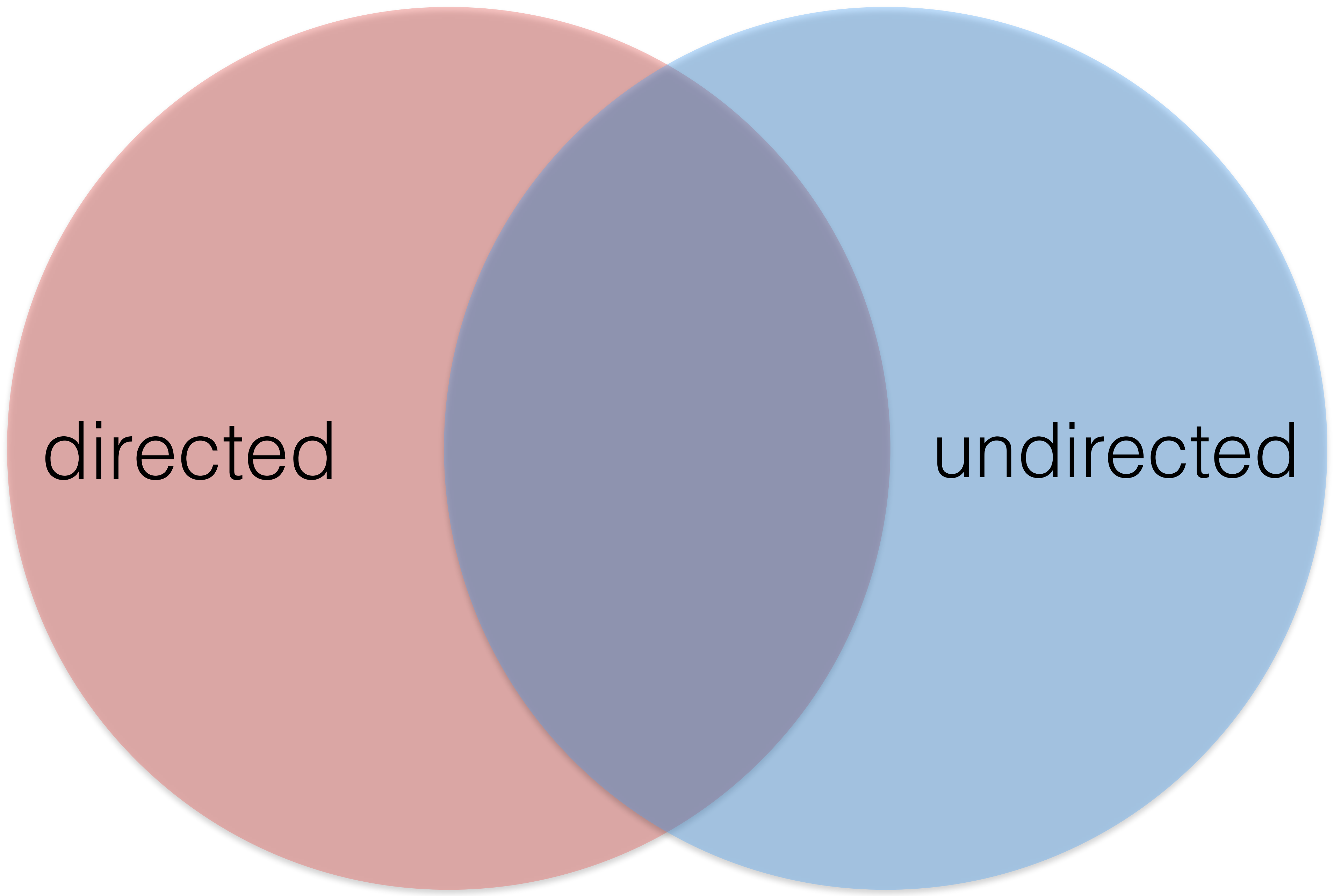
shared child

$$p(A, B, C) \propto \phi(A, C)\phi(B, C)\ \phi(A, B)$$

A and B are **independent** given C

Converting Bayes Nets to MRFs

- Moralize all co-parents
- Lose marginal independence of parents



Summary

- Undirected graphical models, Markov random fields
- Independence in MRFs
- Are Bayesian networks MRFs? No.
  - and MRFs are not Bayesian networks
- Next time: inference via belief propagation