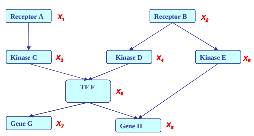
School of Computer Science
Carnegie Mellon

Bayesian & Markov Networks: A unified view

Probabilistic Graphical Models (10-708)

Lecture 3, Sep 19, 2007



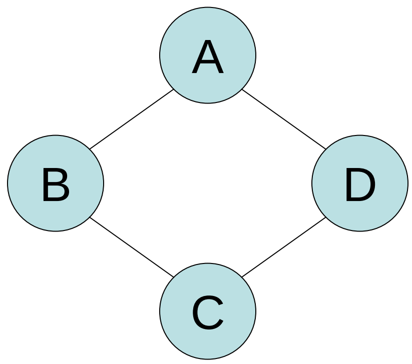
Eric Xing

Reading: KF-Chap. 5.7,5.8

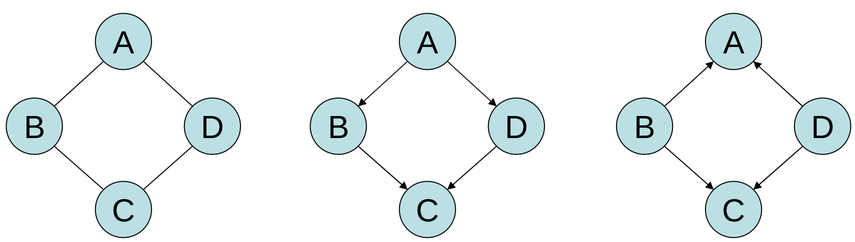
- Auditing students: please fill out forms
- Recitation:
- questions:

Question: Is there a BN that is a perfect map for a given MN?

- The "diamond" MN



Question: Is there a BN that is a perfect map for a given MN?

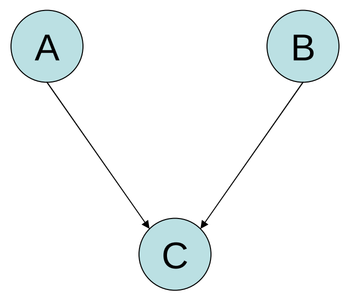


$A \perp C \mid \{B,D\}$	$A \perp C \mid \{B,D\}$	$A \perp C \mid \{B,D\}$
$B \perp D \mid \{A,C\}$	$B \perp D \mid A$	$B \perp D$

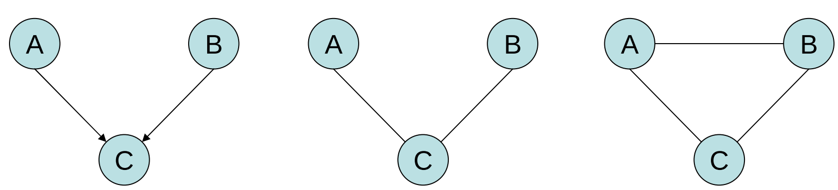
- This MN does not have a perfect I-map as BN!

Question: Is there an MN that is a perfect I-map to a given BN?

- V-structure example



Question: Is there an MN that is a perfect I-map to a given BN?



$A \perp B$

$\neg (A \perp B \mid C)$

$A \perp B \mid C$

$\neg (A \perp B)$

$\neg (A \perp B \mid C)$

$\neg (A \perp B)$

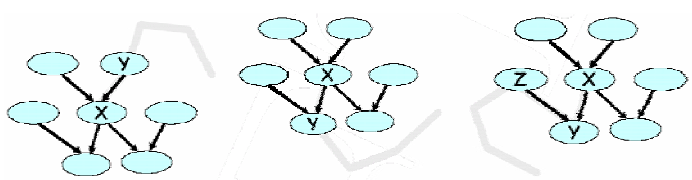
- V-structure has no equivalent in MNs!

Minimal I-maps

- Instead of attempting perfect I-maps between BNs and MNs, we can try minimal I-maps
- Recall: H is a minimal I-map for G if
  - $I(H) \subseteq I(G)$
  - Removal of a single edge in H renders it is not an I-map
- Note: If H is a minimal I-map of G, H need not necessarily satisfy all the independence relationships in G

Minimal I-maps from BNs to MNs: Markov Blanket

- Markov Blanket of X in a BN G:
  - $MB_G(X)$ is the unique minimal set U of nodes in G such that ($X \perp$ (all other nodes) | U) is guaranteed to hold for any distribution that factorizes over G
- Defn (5.7.1): $MB_G(X)$ is the set of nodes consisting of X's parents, X's children and other parents of X's children

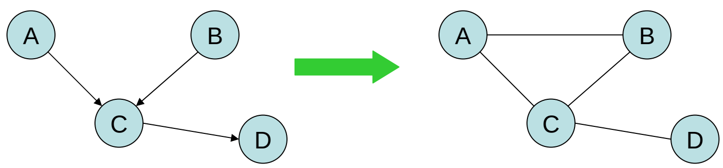


(a) (b) (c)

- Idea: The neighbors of X in H --- the minimal I-map of G --- should be $MB_G(X)$!

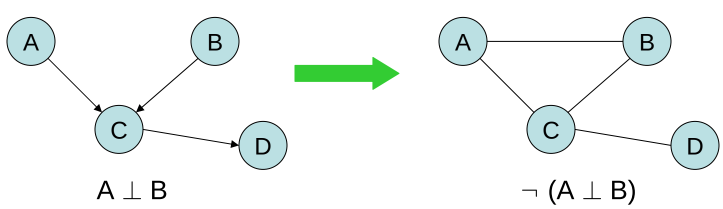
Minimal I-maps from BNs to MNs: Moral Graphs

- Defn (5.7.3): The moral graph M(G) of a BN G is an undirected graph that contains an undirected edge between X and Y if:
  - there is a directed edge between them in either direction
  - X and Y are parents of the same node
- Comment: this definition ensures $\mathrm{MB}_G(X)$ is the set of neighbors of X in M(G)



Minimal I-maps from BNs to MNs: Moral graph is the minimal I-map

- Corollary (5.7.4): The moral graph M(G) of any BN G is a minimal I-map for G
  - Moralization turns each (X, Pa(X)) into a fully connected subset
    - CPDs associated with the network can be used as clique potentials
- The moral graph loses some independence information



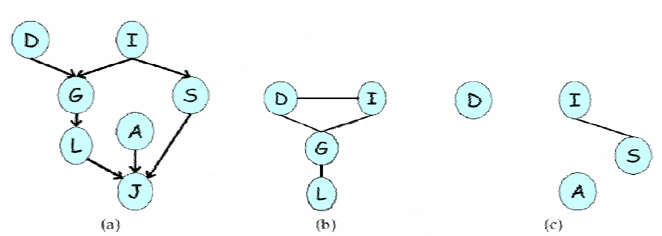
Minimal I-maps from BNs to MNs: Perfect I-maps

- Proposition (5.7.5): If the BN G is "moral", then its moralized graph M(G) is a perfect I-map of G.

- Proof sketch:
  - $I(M(G)) \subseteq I(G)$ (from before)
  - The only independence relations that are potentially lost from G to M(G) are those arising from V-structures
  - Since G has no V-structures (it is moral), no independencies are lost in M(G)

Soundness of *d*-separation

- Recall d-separation
  - Let $U = \{X, Y, Z\}$ be three disjoint sets of nodes in a BN G.
  - Let G^+ be the *ancestral graph*: the induced BN over $U \cup$ ancestors(U).
  - Then, $d\text{-}sep_G(X;Y|Z)$ iff $sep_{M(G^+)}(X;Y|Z)$



D-sep$_G$(D;I | L)

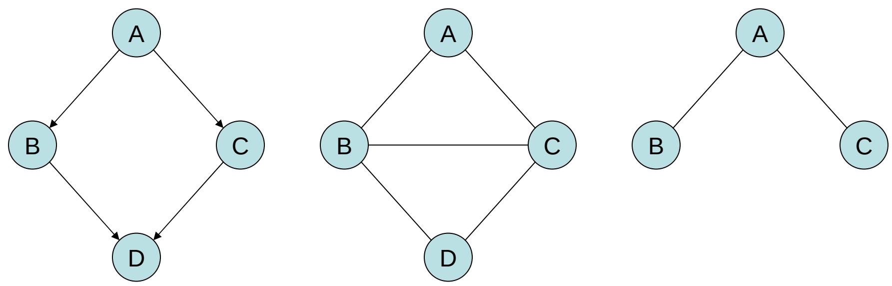
D-sep$_G$(D;I | S, A)

sep$_{M(G^+)}$(D;I | L)

Sep$_{M(G^+)}$(D;I | S,A)

Soundness of *d*-separation

- Why it works:



G: B ⊥ C | A   M(G): ¬(B ⊥ C | A)   M(G⁺): B ⊥ C | A

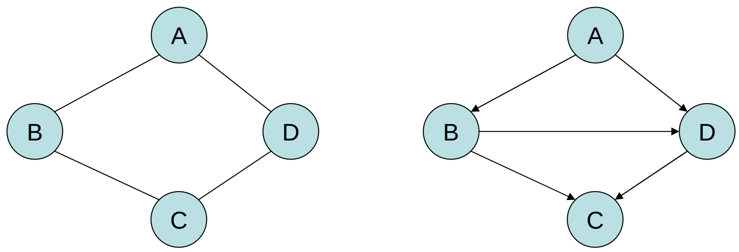
- Idea: Information *blocked* through common children in G that are not in the conditioning variables, is simulated in M(G+) by ignoring all children.

Minimal I-maps from BNs to MNs: Summary

- Moral Graph M(G) is a minimal I-map of G
- If G is moral, then M(G) is a perfect I-map of G
- D-sep$_G$(X;Y|Z) ⇔ sep$_{M(G^+)}$(X;Y|Z)
- Next: minimal I-maps from MNs to BNs ⇒

Minimal I-maps from MNs to BNs:

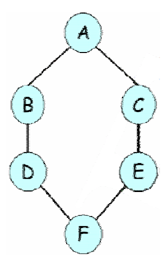
- Any BN I-map for an MN must add triangulating edges into the graph



B ⊥ D | A

Minimal I-maps from MNs to BNs: chordal graphs

- **Defn (5.7.11):**
  - Let X_1-X_2-…X_k-X_1 be a loop in a graph. A chord in a loop is an edge connecting X_i and X_j fo non-consecutive $\{X_i, X_j\}$
  - An undirected graph H is chordal if any loop X_1-X_2-…X_k-X_1 for K >= 4 has a chord



- **Defn (5.7.12):** A *directed graph* G is chordal if its *underlying undirected graph* is chordal

Minimal I-maps from MNs to BNs: triangulation

- **Thm (5.7.13):** Let H be an MN, and G be any BN minimal I-map for H. Then G can have no immoralities.
  - Intuitive reason: Immoralities introduce additional independencies that are not in the original MN
  - (cf. proof for theorem 5.7.13 in K&F)
- **Corollary (5.7.14):** Let K be *any minimal I-map* for H. Then K is necessarily chordal!
  - Because any non-triangulated loop of length at least 4 in a Bayesian network graph necessarily contains an immorality
- Process of adding edges also called *triangulation*

- **Thm (5.7.15):** Let H be a non-chordal MN. Then there is no BN G that is a perfect I-map for H.
- Proof:
  - Minimal I-map G for H is chordal
  - It must therefore have additional directed edges not present in H
  - Each additional edge eliminates some independence assumptions
  - Hence proved.

Clique trees (1)

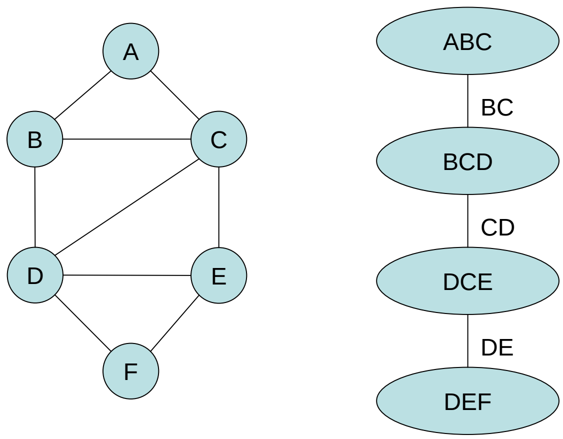
- Notation:
  - Let H be a connected undirected graph. Let $C_1,\ldots C_k$ be the set of maximal cliques in H.
  - Let T be a tree structured graph whose nodes are $C_1,\ldots C_k$.
  - Let C_i, C_j be two cliques in the tree connected by an edge. Define $S_{ij} = C_i \cap C_j$ be the sep-set between C_i and C_j
  - Let $W_{<(i,j)}$ = Variables(C_i) – Variables(S_{ij}) --- the residue set

Clique trees (2)

- A tree T is a clique tree for H if:
  - Each node corresponds to a clique in H and each maximal clique in H is a node in T
  - Each sepset $S_{i,j}$ separates $W_{<(i,j)}$ and $W_{<(j,i)}$
- Every undirected chordal graph H has a clique tree T.
  - Proof by induction (cf. Theorem 5.7.17 in K&F)
  - Example in next slide ⇒

Example

- Example chordal graph and its clique tree



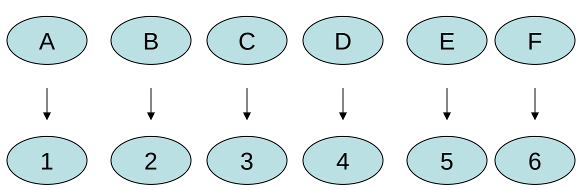
$A \perp D \mid B,C$

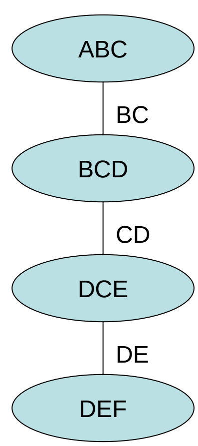
$B \perp E \mid C,D$

$C \perp F \mid D,E$

I-maps of MN as BN:

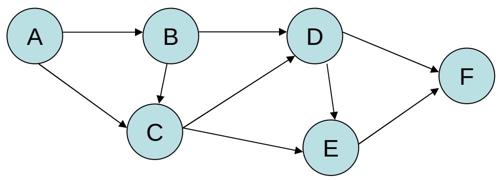
- **Thm (5.7.19):** Let H be a chordal MN. Then there exists a BN such that I(H) = I(G).
- Proof sketch:
  - Since H is an MN, it has a clique tree
  - Number the nodes consistent with clique ordering

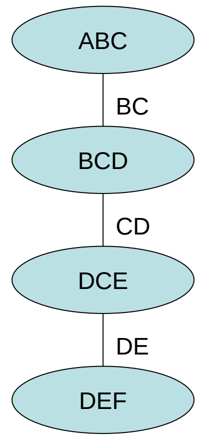




I-maps of MN as BN:

- **Thm (5.7.19):** Let H be a chordal MN. Then there exists a BN such that I(H) = I(G).
- Proof sketch (contd):
  - For each node X_i, let C_k be the first clique it occurs in.
  - Define Pa(X_i) = var{C_k} – $X_i \cap \{X_1, \ldots X_{i-1}\}$





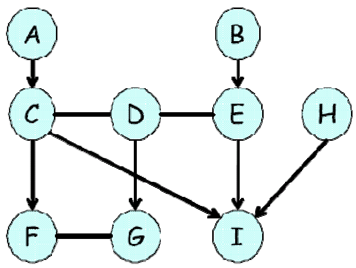
  - G and H have the same edges
  - All parents of each X_i are in the same clique node
    - ⇒ they are connected
    - ⇒ no immoralities in G

Minimal I-maps from MNs to BNs: Summary

- A minimal I-map BN of an MN is chordal
  - Obtained by triangulating the MN
- If the MN is chordal, there is a perfect BN I-map for the MN
  - Obtained from the corresponding clique-tree
- Next: Hybrids of BNs and MNs
  - Partially Directed Acyclic Graphs

Partially Directed Acyclic Graphs

- Also called chain graphs
- Nodes can be disjointly partitioned into several chain components
- An edge within the same chain component must be undirected
- An edge between two nodes in different chain components must be directed



Chain components:

{A}, {B}, {C,D,E},{F,G},{H}, {I}

Summary

- Investigated the relationship between BNs and MNs
  - They represent different families of independence assumptions
  - Chordal graphs can be represented in both
- Chain networks: superset of both BNs and MNs