We have now seen two types of graphical models for representing joint probability distributions:

- **Bayesian Networks**
- **Markov Networks**

Can we convert from one type to the other?

**Bayesian Networks and Markov Networks**

- Can a Markov network represent any Bayesian network? **No!**
- Can’t use a Markov network to represent a distribution corresponding a Bayesian network v-structure

Bayesian networks can represent independence constraints that Markov networks cannot

Markov networks can represent independence constraints that Bayesian networks cannot
Bayesian Networks to Markov Networks

BNs to MNs

To convert a Bayesian Network $B$ to a Markov Network $H$, we can think of $B$ as a Gibbs distribution with:

- Each factor $\phi_{X_i}$ in $H$ corresponds to each conditional probability table $P(X_i|\text{Parents}(X_i))$ in $B$.
- The scope of factor $\phi_{X_i}$ is $X_i \cup \text{Parents}(X_i)$.
- The partition function $Z = 1$.

BNs to MNs

How do we create an undirected graph that is an I-map for the distribution $P_B$ represented by the BN?

Need to moralize each factor:

- Add an edge between $X_i$ and its parents.
- Add an edge between all the parents of $X_i$. 
BNs to MNs

The moral graph $M[G]$ of a Bayesian network structure $G$ over $X$ is the undirected graph over $X$ that contains an undirected edge between $X$ and $Y$ if:
- There is a directed edge between them in $G$ (in either direction)
- Or $X$ and $Y$ are both parents of the same node

Let $G$ be any Bayesian network graph. The moralized graph $M[G]$ is a minimal I-map for $G$.
[See proof of Proposition 4.8 in textbook]

BNs to MNs

A Bayesian network $G$ is moral if it contains no immoralities i.e. for any pair of variables $X, Y$ that share a child, there is a covering edge between $X$ and $Y$.

• The addition of the moralizing edges to the MN $H$ leads to the loss of independence information implied by the graph structure

• Only happens if moralization adds new edges into the graph
BNs to MNs

- If the directed graph $G$ is moral, then its moralized graph $M[G]$ is a perfect map of $G$.
- Unfortunately, very few directed graphs are moral

Markov Networks to Bayesian Networks

MNs to BNs

- Much harder (conceptually and computationally) to find a Bayesian network that is a minimal I-map for a Markov network
- Resulting Bayesian network might have (many) more edges than the Markov network

Example: Find a Bayesian network I-map for the given Markov Network

Build-Minimal-I-Map (from Section 3.4.1):

To find a minimal I-map for a distribution $P$:
- Pick a variable ordering
- For each variable $X_i$ in the ordering:
  - Find some minimal subset $U$ of $\{X_1, ..., X_{i-1}\}$ to be $X_i$’s parents in $G$ such that $(X_i \perp \{X_1, ..., X_{i-1}\} \mid U)$

Note: $(X_i \perp \emptyset \mid \{X_1, ..., X_{i-1}\})$ is trivially true
MNs to BNs
Example: Find a Bayesian network I-map for the given Markov Network
Ordering: A,B,C,D,E,F

(C not ⊥ B | A), so must add B as parent of C

And so on...
Notice the edges added which result in a set of triangles.
The graph is chordal (defn to follow)
MNs to BNs

- Let $X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_4$ be a loop in a graph
- A chord in the loop is an edge connecting $X_i$ and $X_j$ for two nonconsecutive nodes $X_i, X_j$.
- An undirected graph $H$ is said to be chordal if any loop $X_i \rightarrow X_2 \rightarrow \ldots \rightarrow X_k \rightarrow X_l$ for $K \geq 4$ has a chord.

Theorem 4.10: Let $H$ be a Markov network structure, and let $G$ be any Bayesian network minimal I-map for $H$. Then $G$ can have no immoralities. [Proof omitted]

MNs to BNs

Any nontriangulated loop of length at least 4 in a Bayesian network graph necessarily contains an immorality.

Corollary 4.3: Let $H$ be a Markov network structure, and let $G$ be any minimal I-map for $H$. Then $G$ is necessarily chordal.

- Turning a MN to a BN requires triangulation: adding enough edges to a graph to make it chordal.
- Leads to the loss of independence information
- When converting from BN→MN, the (moralizing) edges added are in some sense implicitly there (ie. each factor in the BN involves a node and its parents)
- When converting from MN→BN, we can introduce a large number of edges (via triangulation) which results in very large cliques.
Let $H$ be a chordal Markov network. Then there is a Bayesian network $G$ which is a perfect map for $H$ i.e. $I(H) = I(G)$

We will introduce concepts to help us sketch out a proof for this.

**Basic idea:**
- Show that any connected chordal Markov network $H$ can be decomposed into a clique tree (to be defined)
- The clique tree encodes independencies in $H$
- These independencies can be represented in a Bayesian network $G$
Chordal Graphs

Markov Network $H$

Clique Tree $T$

- Define the Sepset $S_{1,2} = C_1 \cap C_2$ (Note: the two cliques must be connected by an edge)
- Let $W_{<1,2} = \{A\}$ be all of the variables that appear in any clique on the $C_1$ side of the edge
- Let $W_{<1,2} = \{D\}$ be all of the variables that appear in any clique on the $C_2$ side of the edge

We say that 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 $W_{<j,i}$ in $H$
- Note: $W_{<i,j} \perp W_{<j,i} | S_{i,j}$

**Theorem 4.12:** Every undirected chordal graph $H$ has a clique tree $T$. [See book for proof by induction]

Chordal Graphs

How do you form a clique tree from bigger Markov networks?

Theorem 4.13: Let $H$ be a chordal Markov network. Then there is a Bayesian network $G$ such that $I(H) = I(G)$.

The process [see full proof in book]:
- Build clique tree $T$ from $H$
- Call Build-Minimal-I-Map to get Bayesian network $G$
- $G$ (moralized Bayesian network) has the same edges as $H$ due to sepset property $W_{<i,j} \perp W_{<j,i} | S_{i,j}$