Clique Tree Propagation
The propagation algorithm as presented can only deal with *trees*.

Can be extended to *polytrees* (i.e., singly connected graphs with multiple parents per node).

However, it cannot handle networks that contain loops!
Main Objectives

Transform the cyclic directed graph into a secondary structure without cycles.
Find a decomposition of the underlying joint distribution.

Task

Combine nodes of the original (primary) graph structure.
These groups form the nodes of a secondary structure.
Find a transformation that yields tree structure.
Secondary Structure

We will generate an undirected graph mimicking (some of) the conditional independence statements of the cyclic directed graph.

Maximal cliques are identified and form the nodes of the secondary structure. Specify a so-called potential function for every clique such that the product of all potentials yields the initial joint distribution.

In order to propagate evidence, create a tree from the clique nodes such that the following property is satisfied:

If two cliques have some attributes in common, then these attributes have to be contained in every clique of the path connecting the two cliques. (called the running intersection property, RIP)

Justification

Tree: Unique path of evidence propagation.

RIP: Update of an attribute reaches all cliques which contain it.
**Complete Graph**

An undirected Graph $G = (V, E)$ is called complete, if every pair of (distinct) nodes is connected by an edge.

**Induced Subgraph**

Let $G = (V, E)$ be an undirected graph and $W \subseteq V$ a selection of nodes. Then, $G_W = (W, E_W)$ is called the subgraph of $G$ induced by $W$ with $E_W$ being

$$E_W = \{(u, v) \in E \mid u, v \in W\}.$$
**Complete Set, Clique**

Let $G = (V, E)$ be an undirected graph. A set $W \subseteq V$ is called *complete* iff it induces a complete subgraph. It is further called a *clique*, iff $W$ is maximal, i.e. it is not possible to add a node to $W$ without violating the completeness condition.

\[ a) \quad W \text{ is complete } \iff W \text{ induces a complete subgraph} \]

\[ b) \quad W \text{ is a clique } \iff W \text{ is complete and maximal} \]

![Graph diagram](image)

\[ C_1 = \{A, B, C, D\} \]
\[ C_2 = \{B, D, E\} \]
\[ C_3 = \{E, F\} \]

3 cliques
**Perfect Ordering**

Let $G = (V, E)$ be an undirected graph with $n$ nodes and $\alpha = (v_1, \ldots, v_n)$ a total ordering on $V$. Then, $\alpha$ is called *perfect*, if the following sets

$$\text{adj}(v_i) \cap \{v_1, \ldots, v_{i-1}\}$$

are complete, where $\text{adj}(v_i) = \{w \mid (v_i, w) \in E\}$ returns the adjacent nodes of $v_i$.

$\alpha = (A, C, D, F, E, B, H, G)$

<table>
<thead>
<tr>
<th>$i$</th>
<th>$\text{adj}(v_i)$</th>
<th>${v_1, \ldots, v_{i-1}} \cap \text{adj}(v_i)$</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>${C}$</td>
<td>$\emptyset \cap {C}$</td>
<td>$\emptyset$</td>
</tr>
<tr>
<td>2</td>
<td>${A, D, F}$</td>
<td>${A} \cap {A, D, F}$</td>
<td>${A}$</td>
</tr>
<tr>
<td>3</td>
<td>${C, B, E, F}$</td>
<td>${A, C} \cap {C, B, E, F}$</td>
<td>${C}$</td>
</tr>
<tr>
<td>4</td>
<td>${G, C, D, E, H}$</td>
<td>${A, C, D} \cap {G, C, D, E, H}$</td>
<td>${C, D}$</td>
</tr>
<tr>
<td>5</td>
<td>${B, D, F, H}$</td>
<td>${A, C, D, F} \cap {B, D, F, H}$</td>
<td>${D, F}$</td>
</tr>
<tr>
<td>6</td>
<td>${D, E}$</td>
<td>${A, C, D, F, E} \cap {D, E}$</td>
<td>${D, E}$</td>
</tr>
<tr>
<td>7</td>
<td>${F, E}$</td>
<td>${A, C, D, F, E, B} \cap {F, E}$</td>
<td>${F, E}$</td>
</tr>
<tr>
<td>8</td>
<td>${F}$</td>
<td>${A, C, D, F, E, B, H} \cap {F}$</td>
<td>${F}$</td>
</tr>
</tbody>
</table>
Running Intersection Property

Let $G = (V, E)$ be an undirected graph with $p$ cliques. An ordering of these cliques has the running intersection property (RIP), if for every $j > 1$ there exists an $i < j$ such that:

$$C_j \cap \left(C_1 \cup \ldots \cup C_{j-1}\right) \subseteq C_i$$

$$\xi = \langle C_1, C_2, C_3, C_4, C_5, C_6 \rangle$$

$x$ has running intersection property
If a node ordering $\alpha$ of an undirected graph $G = (V, E)$ is perfect and the cliques of $G$ are ordered according to the highest rank (w. r. t. $\alpha$) of the containing nodes, then this clique ordering has RIP.

<table>
<thead>
<tr>
<th>Clique</th>
<th>Rank</th>
<th>$\alpha$ C1</th>
<th>$\alpha$ C2</th>
<th>$\alpha$ C3</th>
<th>$\alpha$ C4</th>
<th>$\alpha$ C5</th>
<th>$\alpha$ C6</th>
</tr>
</thead>
<tbody>
<tr>
<td>${A, C}$</td>
<td>$\max{\alpha(A), \alpha(C)}$</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>${C, D, F}$</td>
<td>$\max{\alpha(C), \alpha(D), \alpha(F)}$</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>${D, E, F}$</td>
<td>$\max{\alpha(D), \alpha(E), \alpha(F)}$</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>${B, D, E}$</td>
<td>$\max{\alpha(B), \alpha(D), \alpha(E)}$</td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>${F, E, H}$</td>
<td>$\max{\alpha(F), \alpha(E), \alpha(H)}$</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>${F, G}$</td>
<td>$\max{\alpha(F), \alpha(G)}$</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

How to get a perfect ordering?
**Triangulated Graph**

An undirected graph is called *triangulated* if every simple loop (i.e. path with identical start and end node but with any other node occurring at most once) of length greater than 3 has a chord.

![Graphs](image)

- **not triangulated**
- **triangulated**
- **not triangulated**
- **no chord for \((A, B, E, C)\)**
Maximun Cardinality Search

Let $G = (V, E)$ be an undirected graph. An ordering according maximum cardinality search (MCS) is obtained by first assigning 1 to an arbitrary node. If $n$ numbers are assigned the node that is connected to most of the nodes already numbered gets assigned number $n + 1$. 

3 can be assigned to $D$ or $F$
6 can be assigned to $H$ or $B$
If an undirected graph is triangulated, then the ordering obtained by MCS is perfect.

To check whether a graph is triangulated is efficient to implement. The optimization problem that is related to the triangulation task is NP-hard. However, there are good heuristics.

**Moral Graph**
Let $G = (V, E)$ be a directed acyclic graph. If $u, w \in W$ are parents of $v \in V$ connect $u$ and $w$ with an (arbitrarily oriented) edge. After the removal of all edge directions the resulting graph $G_m = (V, E')$ is called the *moral graph* of $G$. 

Rudolf Kruse
Bayesian Networks
Given directed graph.

A \rightarrow C \rightarrow F \rightarrow G

B \leftarrow D \leftarrow E \leftarrow H
Join-Tree Construction (2)

- Moral graph
Join-Tree Construction (3)

- Moral graph
- Triangulated graph
Join-Tree Construction (4)

- Moral graph
- Triangulated graph
- MCS yields perfect ordering
Join-Tree Construction (5)

- Moral graph
- Triangulated graph
- MCS yields perfect ordering
- Clique order has RIP

Rudolf Kruse, Bayesian Networks
Join-Tree Construction (6)

- Moral graph
- Triangulated graph
- MCS yields perfect ordering
- Clique order has RIP
- Form a join-tree

Two cliques can be connected if they have a non-empty intersection. The generation of the tree follows the RIP. In case of a tie, connect cliques with the largest intersection. (e.g. $DBE - FED$ instead of $DBE - CFD$)

Break remaining ties arbitrarily.
**Qualitative knowledge**

Metastatic cancer is a possible cause of brain tumor, and is also an explanation for increased total serum calcium. In turn, either of these could explain a patient falling into a coma. Severe headache is also possibly associated with a brain tumor.

**Special case**

The patient suffers from heavy headache.

**Query**

Will the patient fall into coma?
Example: Choice of State Space

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Possible Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$ metastatic cancer</td>
<td>$\text{dom}(A) = {a_1, a_2}$ \quad \cdot_1 = \text{existing}</td>
</tr>
<tr>
<td>$B$ increased total serum calcium</td>
<td>$\text{dom}(B) = {b_1, b_2}$ \quad \cdot_2 = \text{notexisting}$</td>
</tr>
<tr>
<td>$C$ brain tumor</td>
<td>$\text{dom}(C) = {c_1, c_2}$</td>
</tr>
<tr>
<td>$D$ coma</td>
<td>$\text{dom}(D) = {d_1, d_2}$</td>
</tr>
<tr>
<td>$E$ severe headache</td>
<td>$\text{dom}(E) = {e_1, e_2}$</td>
</tr>
</tbody>
</table>

Exhaustive state space:

$$\Omega = \text{dom}(A) \times \text{dom}(B) \times \text{dom}(C) \times \text{dom}(D) \times \text{dom}(E)$$

Marginal and conditional probabilities are of interest for the user!
Example 3: Choice of the conditional probabilities

\[
\begin{align*}
P(e_1 | c_1) &= 0.8 & \text{headaches common, but more common if tumor present} \\
P(e_1 | c_2) &= 0.6 \\
P(d_1 | b_1, c_1) &= 0.8 \\
P(d_1 | b_1, c_2) &= 0.8 & \text{coma rare but common, if either cause is present} \\
P(d_1 | b_2, c_1) &= 0.8 \\
P(d_1 | b_2, c_2) &= 0.05 \\
P(b_1 | a_1) &= 0.8 & \text{increased calcium uncommon,} \\
P(b_1 | a_2) &= 0.2 & \text{but common consequence of metastases} \\
P(c_1 | a_1) &= 0.2 & \text{brain tumor rare, and uncommon consequence of metastases} \\
P(c_1 | a_2) &= 0.05 \\
P(a_1) &= 0.2 & \text{incidence of metastatic cancer in relevant clinic}
\end{align*}
\]
Example: Metastatic Cancer

Dependencies

Moralization/Triangulation

MCS, hyper graph
Example (2)

Quantitative knowledge:

<table>
<thead>
<tr>
<th>(a, b, c)</th>
<th>P(a, b, c)</th>
<th>(b, c, d)</th>
<th>P(b, c, d)</th>
<th>(c, e)</th>
<th>P(c, e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a_1, b_1, c_1</td>
<td>0.032</td>
<td>b_1, c_1, d_1</td>
<td>0.032</td>
<td>c_1, e_1</td>
<td>0.064</td>
</tr>
<tr>
<td>a_2, b_1, c_1</td>
<td>0.008</td>
<td>b_2, c_1, d_1</td>
<td>0.032</td>
<td>c_2, e_1</td>
<td>0.552</td>
</tr>
<tr>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>c_1, e_2</td>
<td>0.016</td>
</tr>
<tr>
<td>a_2, b_2, c_2</td>
<td>0.608</td>
<td>b_2, c_2, d_2</td>
<td>0.608</td>
<td>c_2, e_2</td>
<td>0.368</td>
</tr>
</tbody>
</table>

Decomposition:

\[
P(A, B, C, D, E) = P(A)P(B | A)P(C | A)P(D | BC)P(E | C) = \frac{P(A, B)P(B, C, D)P(C, E)}{P(BC)P(C)}
\]
Marginal distributions in the HUGIN tool.
Example (4)

Conditional marginal distributions with evidence $E = e_1$
Other Loop-Breaking Techniques

2. Stochastic Simulation (1987)
3. Tree clustering (Spiegelhalter & Lauritzen 1986)
4. Node elimination (Shachter 1986)
Applications of Bayesian Networks

- Medical Diagnosis
- Clinical Decision Support
- Complex Genetic Models
- Crime Risk Factors Analysis
- Spatial Dynamics in Geography
- Risk Management in Robotics
- Conservation of a threatened Bird
- Classification of Wines
- Student Modelling
- Sensor Validation

- An Information Retrieval System
- Reliability Analysis of Systems
- Terrorism Risk Management
- Credit-Rating of Companies
- Modelling of Mineral Potential
- Pavement and Bridge Management
- Complex Industrial Process Operation
- Probability of Default for Large Corporates
- Inference Problems in Forensic Science