Markov Random Fields (MRFs)
Markov random fields (MRFs)

- Undirected graph on variables
- Each variable is independent of all other variables, given its neighbors
 Cliques and maximal cliques

Clique, not maximal
Not a clique

Maximal clique

Maximal clique
The distribution for an MRF

\[ P(x_1, \ldots, x_N) = \alpha \prod_i \Psi_i(x_{C_i}) \]

- \( i \) = index of maximal clique
  - \( C_i \) = set of indices of variables
  - \( x_{C_i} \) = set of variables
- \( \Psi_i(x_{C_i}) \) is a “potential” (“local function”)
- \( \alpha \) is a normalizing constant
Burglar MRF

1 maximal clique: $Q_1 = \{e, b, a\}$

Clique potential: $\phi_1(e, b, a)$

Distribution: $P(e, b, a) = \alpha \phi_1(e, b, a)$

Are $e$ and $b$ independent? CAN’T TELL!
Occlusion model

\[ P(z, m, f, b) = P(b)P(f)\left(\prod_{i=1}^{K} P(m_i \mid f)\right)\left(\prod_{i=1}^{K} P(z_i \mid m_i, f, b)\right). \]

\[ P(z, m, f, b) = \pi_b \pi_f \left(\prod_{i=1}^{K} \alpha_{fi}^m (1 - \alpha_{fi})^{1-m_i} N(z_i; \mu_{fi}, \psi_{fi})^m_i N(z_i; \mu_{bi}, \psi_{bi})^{1-m_i}\right). \]
Line processes
(Geman and Geman 1984)

Maximal clique

Patterns with *high* $\phi$

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Patterns with *low* $\phi$

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Factor Graphs
Factor Graphs: Unification

- Factorization
- Independencies
- Inference

1 simple algorithm

B. J. FREY
Factor Graphs

- Bipartite: Variable nodes, function nodes
- Each function node is associated with a function that depends on the neighboring variables
- Joint distribution proportional to product of functions

\[ P(x, y, z) = P(x)P(y)P(z|x, y) \]

\[ P(x, y, z) = P(x)P(y)f(x, z)g(y, z) \]

\[ P(x, y, z) = f(x, y)g(y, z)h(z, x) \]
Occlusion model

\[ P(z, m, f, b) = P(b)P(f)\left(\prod_{i=1}^{K} P(m_i \mid f)\right)\left(\prod_{i=1}^{K} P(z_i \mid m_i, f, b)\right). \]

\[ P(z, m, f, b) = \pi_b \pi_f \left(\prod_{i=1}^{K} \alpha_{fi}^m (1 - \alpha_{fi})^{1-m_i} N(z_i; \mu_{fi}, \psi_{fi})^{m_i} N(z_i; \mu_{bi}, \psi_{bi})^{1-m_i}\right). \]

B. J. FREY
Converting a BN to a FG

Pinch operation:

Retain directed edges to indicate conditional probabilities:

Directed Factor Graph
• So FG >= BN, but are FGs a strict superset?

• Yes.
Occlusion model

BN to FG
Converting an MRF to a FG

Maximal cliques

Function nodes
• So FG >= MRF, but are FGs a strict superset?

• Yes.
Occlusion model

\[ P(z, m, f, b) = \pi_b \pi_f \left( \prod_{i=1}^{K} \alpha_{fi}^{m_i} (1 - \alpha_{fi})^{1-m_i} N(z_i; \mu_{fi}, \psi_{fi})^{m_i} N(z_i; \mu_{bi}, \psi_{bi})^{1-m_i} \right). \]

B. J. FREY
FG for line process

Clique potential
Distributed constraint models
(coding, satisfiability, learning algorithms)

Constr 1  Constr 2  ...

\[ x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7 \quad x_8 \quad x_9 \]

\[ P = \text{product of many potentials on small subsets} \]

Well-suited to factor graph representation

BN: Directed edges don’t have an interpretation

MRF: When the number of constraints is large, the MRF is fully-connected (there is one maximal clique) – inefficient