Join Graphs vs. Region Graphs?

Andrew Gelfand
Questions

- Can all Join Graphs (JGs) be represented as Region Graphs (RGs)?
- Can all RGs be represented as JGs?
- To what extent do we gain when we apply propagation on RG versus JG?
A Join Graph is a triple \(<JG, \rho, \psi>\) where \(JG = (V, E)\) is a graph and \(\rho\) and \(\psi\) are label functions that associate with each vertex \(v \in V\) the sets \(\rho(v) \subseteq X\) and \(\psi(v) \subseteq F\) such that:

- For each factor \(f \in F\) there is exactly one vertex \(v \in V\) such that \(f \in \psi(v)\) and \(\text{scope}[f] \subseteq \rho(v)\)

- (Connectedness) For each variable \(X_i \in X\), the set \(\{v \in V | X_i \in \rho(v)\}\) induces a connected subgraph of JG

Note: This definition is different than that of Darwiche, who insists that a path connect each JG vertex containing some variable \(X_i\).
Valid Join Graphs

- From the definition, \( JG_1 \) is a valid Join Graph
  - Even though it is not arc-minimal

Grid Markov Network

\[
p(x) = \frac{1}{Z} f_a(x_1, x_2) \cdots f_l(x_6, x_9)
\]

Join Graph \( JG_1 \)
Reparameterization

Let $\beta_{v}(x_{v})$ denote belief at vertex $v \in V$ in JG

Let $\mu_{e}(x_{e})$ denote belief on edge $e \in E$ in JG

A JG decomposition defines a reparameterization $p(x) = q_{JG}(x) = \frac{1}{Z_{JG}} \prod_{v \in V} \beta_{v}(x_{v}) \prod_{e \in E} \mu_{e}(x_{e})$

For $JG_{1}$ this is:

$q_{JG_{1}}(x) \propto \frac{\beta(x_{1},x_{2},x_{4},x_{5})\beta(x_{2},x_{3},x_{5},x_{6})\beta(x_{4},x_{5},x_{7},x_{8})\beta(x_{5},x_{6},x_{8},x_{9})}{\mu(x_{2},x_{5})\mu(x_{4},x_{5})\mu(x_{5},x_{6})\mu(x_{5},x_{8})}$
JG₁ reparameterization

- The reparameterization of JG₁ is poor

\[ q_{JG₁}(x) \propto \frac{\beta(x₁,x₂,x₄,x₅)\beta(x₂,x₃,x₅,x₆)\beta(x₄,x₅,x₇,x₈)\beta(x₅,x₆,x₈,x₉)}{\mu(x₄,x₅)\mu(x₂,x₅)\mu(x₅,x₈)\mu(x₅,x₆)} \]

\[ q_{JG₁}(x) \propto q(x₁,x₂|x₄,x₅)q(x₃,x₆|x₂,x₅)q(x₄,x₇|x₅,x₈)q(x₈,x₉|x₅,x₆) \]

\[ \uparrow \uparrow \uparrow \uparrow \uparrow \uparrow \uparrow \uparrow \]

- Not a joint distribution over \( x₁...x₉ \)

- Consider arc-minimal JG₂

\[ q_{JG₂}(x) \propto \frac{\beta(x₁,x₂,x₄,x₅)\beta(x₂,x₃,x₅,x₆)\beta(x₄,x₅,x₇,x₈)\beta(x₅,x₆,x₈,x₉)}{\mu(x₄,x₅)\mu(x₂,x₅)\mu(x₅,x₈)\mu(x₆)} \]

\[ q_{JG₂}(x) \propto q(x₁,x₂|x₄,x₅)q(x₃,x₆|x₂,x₅)q(x₄,x₇|x₅,x₈)q(x₅,x₈,x₉|x₆) \]

\[ \uparrow \uparrow \uparrow \uparrow \uparrow \uparrow \uparrow \uparrow \]
Join Graphs as Region Graphs

- $JG_1$ can be represented as a Region Graph

Outer Regions

<table>
<thead>
<tr>
<th>Regions</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1, x_2, x_4, x_5$</td>
<td>$f_a, f_c, f_g, f_h$</td>
</tr>
</tbody>
</table>

Inner Regions

<table>
<thead>
<tr>
<th>Regions</th>
<th>Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_4, x_5$</td>
<td></td>
</tr>
<tr>
<td>$x_2, x_5$</td>
<td></td>
</tr>
<tr>
<td>$x_5, x_8$</td>
<td></td>
</tr>
<tr>
<td>$x_5, x_6$</td>
<td></td>
</tr>
</tbody>
</table>
Join Graphs as Region Graphs

- JG₁ can be represented as a Region Graph

Outer Regions
- \( x_1, x_2, x_4, x_5 \) \( f_a, f_c, f_g, f_h \)  
  \( \kappa_{R1} = 1 \)
- \( x_4, x_5 \)  
  \( \kappa_{R2} = -1 \)

Inner Regions
- \( x_2, x_5 \)  
  \( \kappa_{R3} = -1 \)
- \( x_5, x_8 \)  
  \( \kappa_{R4} = -1 \)
- \( x_5, x_6 \)  
  \( \kappa_{R5} = -1 \)

- Each region is assigned a counting number \( \kappa_R \)
- Defines RG approximation:  
  \[ q_{RG}(\mathbf{x}) \propto \prod_{r \in R} \beta_r(\mathbf{x}_r)^{\kappa_r} \]
Join Graphs as Region Graphs

- Extending JG$_1$

Outer Regions:

$x_1, x_2, x_4, x_5$
$f_a, f_c, f_g, f_h$

$x_4, x_5, x_7, x_8$
$f_j, f_k, f_e$

$x_2, x_3, x_5, x_6$
$f_b, f_i, f_d$

$x_5, x_6, x_8, x_9$
$f_l, f_f$

Inner Regions:

$x_4, x_5$

$x_2, x_5$

$x_5, x_8$

$x_5, x_6$

$x_5$

$q_{RG}(x) \propto \frac{\beta(x_1, x_2, x_4, x_5)\beta(x_2, x_3, x_5, x_6)\beta(x_4, x_5, x_7, x_8)\beta(x_5, x_6, x_8, x_9)\beta(x_5)}{\mu(x_4, x_5)\mu(x_2, x_5)\mu(x_5, x_8)\mu(x_5, x_6)}$
Questions

- Can all JGs be represented as RGs?
  - Yes. Take nodes as outer regions; edges as inner regions

- Can all RGs be represented as JGs?
  - No. No way to represent the distribution \( q_{RG}(x) \) as a join graph

- To what extent we gain when we apply propagation on RG instead of JG?
  - Region graphs permit broader class of reparameterizations than Join Graphs and thus allow a richer set of approximations
Extending Join Graphs

- Can propagation on JGs be made equivalent to propagation on RGs?
  - Yes...but must think about JGs in new way

- 3 important elements in constructing JGs
  1. Choice of clusters
  2. Messages sent between clusters
  3. How messages are computed
Cluster Choice

- Very important, but very little guidance
  - Mini-bucket schematic provides guidance
  - Can choose by searching for cycle basis of MN

All 7 elementary cycles of $K_4$
Cluster Choice...

- **Cycle basis**: A set of cycles \( B = \{C_1, \ldots, C_\mu\} \) such that for every cycle \( C \) of \( G \), there is a unique subset \( B_C \) of \( B \), such that the edges appearing an odd number of times in \( B_C \) comprise \( C \).

**All 7 elementary cycles of \( K_4 \)**

1. \( \{A, B, C\} = C_1 \)
2. \( \{A, B, D\} = C_2 \)
3. \( \{A, C, D\} = C_3 \)
4. \( \{B, C, D\} = C_4 \)
5. \( \{A, B\} = C_5 \)
6. \( \{A, C\} = C_6 \)
7. \( \{B, C\} = C_7 \)

**A particular cycle basis of \( K_4 \)**

- \( \{a\} = C_1 \)
- \( \{b\} = C_7 \)
- \( \{a, b\} = C_4 \)
- \( \{a, c\} = C_5 \)
- \( \{b, c\} = C_2 \)
- \( \{a, b, c\} = C_6 \)
Messages Between Clusters

- Consider 3x3 grid and following bucket tree

\[ X_5: f_c(x_4,x_5) f_d(x_5,x_6) f_h(x_2,x_5) f_k(x_5,x_8) \]

\[ m(x_4,x_6) \text{ or } m(x_4,x_2) \text{ or } m(x_2,x_4,x_6) \text{ or } m(x_2,x_4,x_8) \]

\[ X_4: f_g(x_1,x_4) f_j(x_4,x_7) \]

Why not pass multiple messages between clusters? For a graph, what are message updates? What set of approximations does this permit?
Message Computation

- Given previous bucket tree, how do we compute message $m(x_4, x_6)$?

- $X_5$: $f_c(x_4, x_5) f_d(x_5, x_6) f_h(x_2, x_5) f_k(x_5, x_8)$

- $m(x_4, x_6)$

- $X_4$: $f_g(x_1, x_4) f_j(x_4, x_7)$

$$m_{\text{partial}}(x_4, x_6) = \sum_{x_5} f_c(x_4, x_5) f_d(x_5, x_6)$$

$$m_{\text{full}}(x_4, x_6) = \sum_{x_2, x_5, x_8} f_c(x_4, x_5) f_d(x_5, x_6) f_h(x_2, x_5) f_k(x_5, x_8)$$