ON BAYESIAN NETWORK INFERENCE WITH SIMPLE PROPAGATION

Cory J. Butz
butz@cs.uregina.ca
University of Regina
Canada

Jhonatan S. Oliveira
oliveira@cs.uregina.ca
University of Regina
Canada

André E. dos Santos
dossantos@cs.uregina.ca
University of Regina
Canada

Anders L. Madsen
anders@hugin.com
HUGIN EXPERT A/S
Aalborg University
Denmark
OUTLINE

• Bayesian networks
• Inference with Lazy Propagation
• Inference with Simple Propagation
• Experimental Results & Analysis
• Conclusions
A Bayesian Network (BN) consists of:

- a directed acyclic graph (DAG)
- a matching set of conditional probability tables (CPTs)

The product of the CPTs is a join probability distribution (JPD) $P(U)$
Bayesian Network Example

\[ P(U) = P(a) \cdot P(b|a) \cdot P(c|a) \cdot P(d|b,c) \cdot \ldots \cdot P(m|g,l) \]
LAZY PROPAGATION
• Madsen and Jensen (AIJ 1999)
• BN variables are clustered into nodes
• Nodes are organized as a join tree
• Each BN CPT is assigned to a join tree node
• Messages are propagated systematically
Message Construction

\[\text{message} = \sum_{N-N'} \text{Factorization at } N\]

\[M = \sum_{b,c,e,f,g,h} P(b, c) \cdot P(d = 0|b, c) \cdot P(e|d = 0) \cdot P(f|d = 0, e) \cdot P(g|e) \cdot P(h|e) \cdot P(i|d = 0, h) \cdot P(j|i) \cdot P(m|g, l)\]
DETECTING IRRELEVANT POTENTIALS

- LP constructs:
  - the domain graph $G_1$ of the factorization
  - the moralization $G_1^m$ of $G_1$
- LP tests whether the evidence separates the variables to be marginalized from the separator
- if separated, the potential is irrelevant
BUILD DOMAIN GRAPH $G_1$
BUILD MORALIZATION GRAPH $G_1^m$
TEST INDEPENDENCE FOR EACH POTENTIAL

- For $P(b,c)$, test whether evidence $d$ separates $b$ and $c$ from the separator $S = \{i,j,l,m\}$

- Thus, $P(b,c)$ is irrelevant
Determining Elimination Orderings

- LP constructs:
  - the domain graph $G_2$ of the relevant potentials
  - the moralization $G_2^m$ of $G_2$
  - obtain an elimination ordering from $G_2^m$
BUILD DOMAIN GRAPH $G_2$

\[ \mathcal{F} = \{ P(e|d = 0), P(g|e), P(h|e), P(i|d = 0, h), P(j|i), P(m|g, l) \} \]
BUILD MORALIZATION GRAPH $G_{2^m}$
• elimination ordering: $g, e, h$
NOW LP CAN BUILD THE MESSAGE

\[
M = \sum_{e,g,h} P(e|d=0) \cdot P(g|e) \cdot P(h|e) \cdot P(i|d=0, h) \cdot P(j|i) \cdot P(m|g,l)
\]

\[
= P(j|i) \cdot \sum_{h} P(i|d=0, h) \cdot \sum_{e} P(e|d=0) \cdot P(h|e) \cdot \sum_{g} P(g|e) \cdot P(m|g,l)
\]

\[
= P(j|i) \cdot P(i, m|d=0, l)
\]  (1)
SIMPLE PROPAGATION
DARWINIAN NETWORKS

- Simple Propagation arose from our work on Darwinian Networks (AI 2015)
- clever way to view CPTS

\[ P(g|e, f) \]

\[
\begin{array}{c}
\text{P}(g|e, f) \\
\Rightarrow \\
\begin{array}{c}
g \\
e \\
f
\end{array}
\end{array}
\]
MULTIPLICATION IS MERGE

○ white  +  ● black  =  ○ white
● black  +  ○ white  =  ○ white
● black  +  ● black  =  ● black
○ white  +  ○ white  =  ● black

\[ P(c|h) \cdot P(e|c, d) = P(c, e|d, h) \]
MARGINALIZATION IS
REPLICATION AND NATURAL SELECTION

\[ \sum_c P(c, e \mid d, h) = P(e \mid d, h) \]
SIMPLE PROPAGATION

SP only uses the "one in, one out" property:

a potential with one non-evidence variable in the separator and another not in the separator.
**Simple Propagation**

Evidence is $d = 0$

Variable $g$ is outside of $S$ and variables $l$ and $m$ are in $S$
SIMPLE PROPAGATION

Eliminating variable $g$ yields population $p(m|e, l)$
SIMPLE PROPAGATION

Now, variable e is out
Simple Propagation

Eliminating variable \( e \) yields \( p(h, m | d = 0, l) \)
SIMPLE PROPAGATION

Finally, variable $h$ is **out**

![Diagram showing variable $h$ being propagated through networks with relevant and irrelevant potentials.](image)
SIMPLE PROPAGATION

Eliminating variable $h$ yields population $p(i, m|d = 0, l)$

$$P(j|i) \cdot P(i, m|d = 0, l)$$  \hspace{1cm} (1)
<table>
<thead>
<tr>
<th>BN</th>
<th>Vars</th>
<th>LP</th>
<th>SP</th>
<th>Saving</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>32</td>
<td>0.06</td>
<td>0.05</td>
<td>17%</td>
</tr>
<tr>
<td>Oow</td>
<td>33</td>
<td>0.07</td>
<td>0.06</td>
<td>14%</td>
</tr>
<tr>
<td>Oow_Bas</td>
<td>33</td>
<td>0.04</td>
<td>0.03</td>
<td>25%</td>
</tr>
<tr>
<td>Mildew</td>
<td>35</td>
<td>0.05</td>
<td>0.04</td>
<td>20%</td>
</tr>
<tr>
<td>Oow_Solo</td>
<td>40</td>
<td>0.07</td>
<td>0.06</td>
<td>14%</td>
</tr>
<tr>
<td>Hkv2005</td>
<td>44</td>
<td>0.23</td>
<td>0.27</td>
<td>-17%</td>
</tr>
<tr>
<td>Barley</td>
<td>48</td>
<td>0.09</td>
<td>0.1</td>
<td>-11%</td>
</tr>
<tr>
<td>Kk</td>
<td>50</td>
<td>0.09</td>
<td>0.09</td>
<td>0%</td>
</tr>
<tr>
<td>Ship</td>
<td>50</td>
<td>0.16</td>
<td>0.17</td>
<td>-6%</td>
</tr>
<tr>
<td>Hailfinder</td>
<td>56</td>
<td>0.02</td>
<td>0.02</td>
<td>0%</td>
</tr>
<tr>
<td>Medianus</td>
<td>56</td>
<td>0.04</td>
<td>0.03</td>
<td>25%</td>
</tr>
<tr>
<td>3Nt</td>
<td>58</td>
<td>0.02</td>
<td>0.01</td>
<td>50%</td>
</tr>
<tr>
<td>Hepar_li</td>
<td>70</td>
<td>0.03</td>
<td>0.03</td>
<td>0%</td>
</tr>
<tr>
<td>Win95Pts</td>
<td>76</td>
<td>0.03</td>
<td>0.03</td>
<td>0%</td>
</tr>
<tr>
<td>System_V57</td>
<td>85</td>
<td>0.06</td>
<td>0.05</td>
<td>17%</td>
</tr>
<tr>
<td>Fwe_Model8</td>
<td>109</td>
<td>0.14</td>
<td>0.15</td>
<td>-7%</td>
</tr>
<tr>
<td>Pathfinder</td>
<td>109</td>
<td>0.12</td>
<td>0.11</td>
<td>8%</td>
</tr>
<tr>
<td>Adapt_T1</td>
<td>133</td>
<td>0.04</td>
<td>0.04</td>
<td>0%</td>
</tr>
<tr>
<td>Cc145</td>
<td>145</td>
<td>0.1</td>
<td>0.08</td>
<td>20%</td>
</tr>
<tr>
<td>Munin1</td>
<td>189</td>
<td>0.54</td>
<td>0.75</td>
<td>-39%</td>
</tr>
<tr>
<td>Andes</td>
<td>223</td>
<td>0.15</td>
<td>0.13</td>
<td>13%</td>
</tr>
<tr>
<td>Cc245</td>
<td>245</td>
<td>0.2</td>
<td>0.18</td>
<td>10%</td>
</tr>
<tr>
<td>Diabetes</td>
<td>413</td>
<td>0.34</td>
<td>0.31</td>
<td>9%</td>
</tr>
<tr>
<td>Adapt_T2</td>
<td>671</td>
<td>0.24</td>
<td>0.22</td>
<td>8%</td>
</tr>
<tr>
<td>Amirali</td>
<td>681</td>
<td>0.45</td>
<td>0.41</td>
<td>9%</td>
</tr>
<tr>
<td>Munin2</td>
<td>1003</td>
<td>0.49</td>
<td>0.45</td>
<td>8%</td>
</tr>
<tr>
<td>Munin4</td>
<td>1041</td>
<td>0.61</td>
<td>0.57</td>
<td>7%</td>
</tr>
<tr>
<td>Munin3</td>
<td>1044</td>
<td>0.66</td>
<td>0.64</td>
<td>3%</td>
</tr>
</tbody>
</table>

- Experiments conducted on optimal JTs built from real-world and benchmark BNs
- SP was faster in 18/28
- SP tied LP in 5/28
- LP was faster in 5/28
LP ANALYSIS

• Left-to-Right viewpoint

S - E

irrelevant potentials

relevant potentials
SP Analysis

• Right-to-Left viewpoint
SP Analysis

- Right-to-Left viewpoint
EXPERIMENTAL RESULTS
SP HEURISTICS

• SP is a new BN inference algorithm

• There may be more than one potential satisfying the “one in, one out” property
SP HEURISTICS

• Increasing variables in X (Inc X)
• Decreasing variables in X (Dec X)
• Increasing variables of X in S (Inc in S)
• Decreasing variables of X in S (Dec in S)
• Increasing variables in X size (Inc X Size)
• Decreasing variables in X size (Dec X Size)
• Increasing variables of X in S size (Inc in S Size)
• Decreasing variables of X in S size (Dec in S Size).
<table>
<thead>
<tr>
<th>BN</th>
<th>Vars</th>
<th>Arbitrary Order</th>
<th>Inc X</th>
<th>Dec X</th>
<th>Inc in S</th>
<th>Dec in S</th>
<th>Inc X Size</th>
<th>Dec X Size</th>
<th>Inc in S Size</th>
<th>Dec in S Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>32</td>
<td>0.05</td>
<td>0.04</td>
<td>0.04</td>
<td>0.05</td>
<td>0.05</td>
<td>0.04</td>
<td>0.05</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td>oow</td>
<td>33</td>
<td>0.06</td>
<td>0.05</td>
<td>0.06</td>
<td>0.05</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>oow_bas</td>
<td>33</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>Mildew</td>
<td>35</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>oow_solo</td>
<td>40</td>
<td>0.06</td>
<td>0.05</td>
<td>0.06</td>
<td>0.05</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>HKV</td>
<td>44</td>
<td>0.02</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>Barley</td>
<td>48</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>KK</td>
<td>50</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.07</td>
<td>0.08</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>ship</td>
<td>50</td>
<td>0.16</td>
<td>0.13</td>
<td>0.16</td>
<td>0.13</td>
<td>0.16</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
<td>0.15</td>
</tr>
<tr>
<td>hailfinder</td>
<td>56</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>medianus</td>
<td>56</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>3nt</td>
<td>58</td>
<td>0.01</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Hepar_II</td>
<td>70</td>
<td>0.24</td>
<td>0.23</td>
<td>0.32</td>
<td>0.28</td>
<td>0.31</td>
<td>0.23</td>
<td>0.32</td>
<td>0.27</td>
<td>0.31</td>
</tr>
<tr>
<td>win95pts</td>
<td>76</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>system_v57</td>
<td>85</td>
<td>0.05</td>
<td>0.05</td>
<td>0.06</td>
<td>0.05</td>
<td>0.06</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>FEW</td>
<td>109</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
<td>0.13</td>
<td>0.14</td>
</tr>
<tr>
<td>pathfinder</td>
<td>109</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
<td>0.09</td>
<td>0.10</td>
<td>0.09</td>
<td>0.09</td>
<td>0.10</td>
</tr>
<tr>
<td>Adapt_T1</td>
<td>133</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>cc145</td>
<td>145</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.07</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.08</td>
<td>0.07</td>
</tr>
<tr>
<td>Munin1</td>
<td>189</td>
<td>0.68</td>
<td>0.72</td>
<td>0.72</td>
<td>0.84</td>
<td>0.90</td>
<td>0.71</td>
<td>0.69</td>
<td>0.76</td>
<td>0.84</td>
</tr>
<tr>
<td>andes</td>
<td>223</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>cc245</td>
<td>245</td>
<td>0.17</td>
<td>0.18</td>
<td>0.18</td>
<td>0.17</td>
<td>0.18</td>
<td>0.17</td>
<td>0.18</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>Diabetes</td>
<td>413</td>
<td>0.27</td>
<td>0.28</td>
<td>0.28</td>
<td>0.28</td>
<td>0.28</td>
<td>0.28</td>
<td>0.28</td>
<td>0.28</td>
<td>0.28</td>
</tr>
<tr>
<td>Adapt_T2</td>
<td>671</td>
<td>0.19</td>
<td>0.19</td>
<td>0.19</td>
<td>0.20</td>
<td>0.20</td>
<td>0.19</td>
<td>0.20</td>
<td>0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>Amirali</td>
<td>681</td>
<td>0.40</td>
<td>0.39</td>
<td>0.44</td>
<td>0.42</td>
<td>0.40</td>
<td>0.42</td>
<td>0.44</td>
<td>0.40</td>
<td>0.41</td>
</tr>
<tr>
<td>Munin2</td>
<td>1003</td>
<td>0.44</td>
<td>0.41</td>
<td>0.42</td>
<td>0.41</td>
<td>0.43</td>
<td>0.42</td>
<td>0.44</td>
<td>0.40</td>
<td>0.39</td>
</tr>
<tr>
<td>Munin4</td>
<td>1041</td>
<td>0.51</td>
<td>0.52</td>
<td>0.51</td>
<td>0.52</td>
<td>0.52</td>
<td>0.53</td>
<td>0.51</td>
<td>0.52</td>
<td>0.51</td>
</tr>
<tr>
<td>Munin3</td>
<td>1044</td>
<td>0.53</td>
<td>0.55</td>
<td>0.56</td>
<td>0.56</td>
<td>0.56</td>
<td>0.55</td>
<td>0.56</td>
<td>0.55</td>
<td>0.55</td>
</tr>
<tr>
<td>sacso</td>
<td>2371</td>
<td>0.72</td>
<td>0.75</td>
<td>0.74</td>
<td>0.74</td>
<td>0.74</td>
<td>0.75</td>
<td>0.74</td>
<td>0.74</td>
<td>0.75</td>
</tr>
</tbody>
</table>

Tied for first: 13  16  11  17  9  12  14  11  13
Unique wins:     6  1  0  0  0  0  0  0  1
ANALYSIS

• Our experimental results suggest that SP does not require elimination orderings, provided that an optimal (or close to) join tree is built from the real-world BNs.

• It is possible that elimination orderings are needed for larger BNs or when non-optimal join trees are used, since SP’s performance degrades dramatically when applied on non-optimal join trees (Madsen et al., 2016 Canadian AI).
CONCLUSION

• SP is a new BN inference algorithm

• “one in, one out” property

• SP is faster than LP in optimal join trees

• Our heuristics were slower than choosing potentials arbitrarily