Bayesian Network Reasoning with Gibbs Sampling

Todd W. Neller
Gettysburg College
Outline

• Prerequisites
• Objectives
• What is provided
• Example exercise
• Possible uses and extensions
Prerequisites

- Audience: Undergraduate Intro AI students
- Prerequisite knowledge:
  - DAGs
  - Axioms of probability
  - Conditional probabilities
  - Conditional probability tables (CPTs)
  - Competent programmers (pseudocode → code)
Objectives

• Students will
  – Implement the core of the **Gibbs Sampling** algorithm, a Markov Chain Monte Carlo (MCMC) algorithms for Bayesian Network (BN) reasoning
  – Empirically observe important concepts of reasoning, e.g. **conditional independence, diagnostic inference, causal, intercausal, and mixed inference, D-separation**
  – Experience both **strengths** and **weaknesses** of this MCMC approach, gaining an appreciation for its place in our algorithmic toolset.
What is Provided

• Brief introduction to BN reasoning
  – Recommendations for supplementary readings
• Java software to parse a simple grammar for describing BN structure, CPTs, and evidence and build a helpful BN data structure
• Exercises to guide a student to an empirical understanding of core BN reasoning concepts
Example Exercise – Intercausal Inference (Explaining Away)

\[ P(a) = \{0.20\} \]
\[ P(b|a) = \{0.20, 0.80\} \]
\[ P(c|a) = \{0.05, 0.20\} \]
\[ P(d|b,c) = \{0.05, 0.80, 0.80, 0.80\} \]
\[ P(e|c) = \{0.60, 0.80\} \]

After iteration 1000000:
Variable, Average Conditional Probability, Fraction True
a, 0.4248869857096668, 0.424857
b, 0.8001614923088834, 0.800057
c, 0.20001273834624336, 0.199574
e, 0.6399147999906855, 0.640206

BNGibbsSampler
Example Exercise – Intercausal Inference (Explaining Away) +c

After iteration 1000000:
Variable, Average Conditional Probability, Fraction True
a, 0.49929080000000346, 0.49861
b, 0.49916600000000317, 0.498817
e, 0.800000000000106631, 0.799731

P(a) = {0.20}
P(b|a) = {0.20, 0.80}
P(c|a) = {0.05, 0.20}
P(d|b,c) = {0.05, 0.80, 0.80, 0.80}
P(e|c) = {0.60, 0.80}
evidence
d
c

BNGibbsSampler
Example Exercise – Intercausal Inference (Explaining Away) +b

After iteration 1000000:
Variable, Average Conditional Probability, Fraction True
a, 0.49982754285114167, 0.5
c, 0.1249383398139281, 0.124496
e, 0.6248991999897608, 0.62497
Example Exercise – Intercausal Inference (Explaining Away) Summary

• Given evidence of $d$ ($P(d) = 1$):

<table>
<thead>
<tr>
<th>Evidence</th>
<th>+d</th>
<th>+d +c</th>
<th>+d +b</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P(b)$</td>
<td>.8</td>
<td>.5</td>
<td>1</td>
</tr>
<tr>
<td>$P(c)$</td>
<td>.2</td>
<td></td>
<td>.125</td>
</tr>
</tbody>
</table>

![Diagram of the intercausal inference network with probabilities for each node:

- $A$: $P(A) = 0.2$
- $B$: $P(B) = 0.2$ (FALSE), $P(B) = 0.8$ (TRUE)
- $C$: $P(C) = 0.05$ (FALSE), $P(C) = 0.2$ (TRUE)
- $D$: $P(D) = 0.05$ (FALSE FALSE), $P(D) = 0.8$ (FALSE TRUE), $P(D) = 0.8$ (TRUE FALSE), $P(D) = 0.8$ (TRUE TRUE)
- $E$: $P(E) = 0.6$ (FALSE), $P(E) = 0.8$ (TRUE)]
Possible Uses and Extensions

• Complete exercises emphasize core Gibbs Sampling implementation and experiential learning of BN reasoning concepts

• Solutions are available for instructors
  – Could provide Gibbs sampling implementation and focus simply on empirical observations of BN reasoning concepts

• Ease of BN specification allows easy assignment of additional BN exercises with various foci:
  – Representation focus
  – Estimation qualitative/quantitative focus
  – Estimation accuracy focus
Conclusion

• This assignment provides:
  – Easy parsing of simple BN specification
  – Auto construction of BN data structure with helper methods to aid focus on the core implementation of Gibbs Sampling
  – Basic exercises to efficiently and experientially teach about BN reasoning
  – Possibility of simplification, extension, various foci

• [http://modelai.gettysburg.edu/2017/mc2](http://modelai.gettysburg.edu/2017/mc2)