

GIT TEST

Bayes' rule:

$$\mathbb{P}(A_i|B) = \frac{\mathbb{P}(B|A_i) \cdot \mathbb{P}(A_i)}{\sum_j \mathbb{P}(B|A_j) \cdot \mathbb{P}(A_j)}$$

Binary Bayes' rule:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A) \cdot \mathbb{P}(A)}{\mathbb{P}(B)}$$

Scriptie Weikamp:

Main question: *can we find a reliable (approximate) inference method that works theoretically as well as practically?*

- To what extent do the commonly used algorithms give reliable results?
- To what extent is the BL method useful in doing inference in Bayesian networks?
Not useful

Chapter 2, section 2.1

Def. (Bayesian inference) The calculation of the *posterior distribution*:

$$\mathbb{P}(X = x|E = e).$$

Chapter 2, section 2.2

Def. (Chain rule)

$$\mathbb{P}(X, Y) = \mathbb{P}(X|Y)\mathbb{P}(Y)$$

Def. (General chain rule)

$$\mathbb{P}(X_1, \dots, X_n) = \sum_{i=1}^n \mathbb{P}(X_i|X_1, \dots, X_{n-1})$$

Def. (Conditionally independent) We say that A and B are conditionally independent given C if

$$\mathbb{P}(A|B, C) = \mathbb{P}(A|C).$$

Equivalently,

$$\mathbb{P}(A, B|C) = \mathbb{P}(A|C)\mathbb{P}(B|C).$$

We write $(A \perp\!\!\!\perp B \mid C)$

Def. (Local independencies) Local independencies, denoted by $\mathcal{I}_l(\mathcal{G})$

For each $X_i : (X_i \perp\!\!\!\perp ND(X_i) \mid Pa(X_i))$.

Def. (Factorization - chain rule for Bayesian networks)

$$\mathbb{P}(X_1, \dots, X_n) = \prod_{i=1}^n \mathbb{P}(X_i \mid Pa(X_i))$$

Def. (Factorization - chain rule for Bayesian networks)

$$\mathbb{P}(X_1, \dots, X_n) = \prod_{i=1}^n \mathbb{P}(X_i \mid Pa(X_i))$$

Chapter 3, section 3.2

Inference methods:

- exact
 - Variable elimination
 - Clique trees
 - Recursive conditioning
- approximate
 - Optimization
 - * *Loopy belief Propagation* (lack of convergence, strong dependencies)
 - Sampling inference methods
 - * forward sampling
 - * rejected sampling
 - * likelihood sampling
 - * importance sampling
 - * MCMC
 - Gibbs sampling
 - prune sampling

Def. (Factor) Each $\mathbb{P}(X_i \mid Pa(X_i))$ can be written as ϕ_i .

Def. (Factor marginalisation) Let \mathbf{X} be a set of variables and $Y \notin \mathbf{X}$ a variable. Let $\phi(\mathbf{X}, Y)$ be a factor. We define the factor marginalization of Y in ϕ , denoted $\sum_Y \phi$, to be a factor ψ over \mathbf{X} such that

$$\psi(X) = \sum_Y \phi(\mathbf{X}, Y).$$

. **(Variable Elimination)** Choosing an elimination sequence, applying factor marginalization and finally normalizing we can obtain marginal probabilities $\mathbb{P}(X_i)$, where X_i is a variable of interest in the Bayesian network.

Drawbacks: if many connections \implies large amount of factors ψ , which are products of $\phi_i \implies$ exponential growth of the CPT of ψ . (memory intensive)

Chapter 4, section 4.1

Def. (Collection of all CPT-indices of a BN)

$\mathcal{C} = \{k(i) : c_{k(i)} \text{ is a CPT-value in the } i\text{-th CPT } \mathbb{P}(X_i \mid Pa(X_i)), \text{ indexed by } k(i), i=1, \dots, n\}$

Def. (CPT-indices corresponding to state \mathbf{x})

$$\mathcal{C}_{\mathbf{x}}$$

Def. (States corresponding to a set of CPT-indices) Given a collection of CPT-indices \mathcal{C} . The set $S_{\mathcal{C}}$ of states \mathbf{x} that use only the CPT-indices in the collection \mathcal{C} are called states corresponding to the CPT-indices.

Def. (Pruning around \mathbf{x} , collection of pruned CPT-indices)

$$\begin{aligned} \mathcal{C}_{\mathbf{x},p} &= \emptyset \\ \forall k(i) \in \mathcal{C} : & \begin{cases} \text{add } k(i) \text{ to } \mathcal{C}_{\mathbf{x},p} & \text{w.p. } 1 - c_{k(i)} \\ \text{do not add } k(i) \text{ to } \mathcal{C}_{\mathbf{x},p} & \text{w.p. } c_{k(i)} \end{cases} \end{aligned}$$

Def. (Collection of non-pruned CPT-indices)

$$\begin{aligned} \mathcal{C}_{\mathbf{x},n} &:= \mathcal{C} \setminus \mathcal{C}_{\mathbf{x},p} \\ \text{Note that: } \mathcal{C}_{\mathbf{x}} &\subset \mathcal{C}_{\mathbf{x},n} \text{ and } \mathbf{x} \in S_{\mathcal{C}_{\mathbf{x},n}}. \end{aligned}$$

Def. (Uniform sampling over a set of states) Let $S_{\mathcal{C}_{\mathbf{x},n}}$ be the set of feasible states corresponding to the CPT-indices which are not pruned. We define

$$\mathcal{U}(S_{\mathcal{C}_{\mathbf{x},n}})$$

as the uniform distribution over the states in $S_{\mathcal{C}_{\mathbf{x},n}}$ and we write

$$\mathcal{U}(S_{\mathcal{C}_{\mathbf{x},n}})(y) = \frac{1}{|S_{\mathcal{C}_{\mathbf{x},n}}|}$$

for the probability of sampling state \mathbf{y} with respect to this uniform distribution.

(Remark) - *(Could be proven more formally)* With strictly positive probability we have that $\mathcal{C}_{\mathbf{x}^{(i-1)},n}$ contains all the non-zero indices in \mathcal{C} . Implying that $S_{\mathcal{C}_{\mathbf{x}^{(i-1)},n}}$ contains all feasible states of the BN.

Def. (Pruning around \mathbf{x} and \mathbf{y} , collection of pruned CPT-indices)

$$\begin{aligned} \mathcal{C}_{\{\mathbf{x},\mathbf{y}\},p} &= \emptyset \\ \forall k(i) \in \mathcal{C} : & \begin{cases} \text{add } k(i) \text{ to } \mathcal{C}_{\mathbf{x},p} \text{ when } k(i) \notin \mathcal{C}_{\mathbf{x}}, \mathcal{C}_{\mathbf{y}} & \text{w.p. } 1 - c_{k(i)} \\ \text{do not add } k(i) \text{ to } \mathcal{C}_{\mathbf{x},p} & \text{w.p. } c_{k(i)} \end{cases} \end{aligned}$$

Def. (Collection of non-pruned CPT-indices)

$$\mathcal{C}_{\{\mathbf{x},\mathbf{y}\},n} := \mathcal{C} \setminus \mathcal{C}_{\{\mathbf{x},\mathbf{y}\},p}$$

Def. (Total probability of transitioning from $\mathbf{x} \rightarrow \mathbf{y}$)

$$Q(\mathbf{x} \rightarrow \mathbf{y}) = \sum_{j=1}^K Q_j(\mathbf{x} \rightarrow \mathbf{y})$$

Chapter 4, section 4.2

Def. (Non-trivial steps of prune sampling)

1) Generating an initial state

- forward sampling: how many forward sampling walks will it take to generate one feasible solution?
 - random forward sampling
 - hybrid forward sampling
- 2) Sampling uniformly over the pruned BN, i.e. sampling from the distribution $\mathcal{U}(S_{C_{\mathbf{x},n}})$.

Def. (Literature about generating initial states)

[18] James D Park, *Using Weighted MAX-SAT engines to solve MPE*, pp. 682-687

Def. (Literature about uniformly sampling $\mathcal{U}(S_{C_{\mathbf{x},n}})$)

[28] Wei Wei, Jordan Erenrich and Bart Selman, *Towards efficient sampling: Exploiting random walk strategies*, Aaai 2004, pp. 670-676