Probabilistic reasoning via Weighted Model Counting

Knowledge and Data Representation Project

Mario Tapia Montero (ID 2081407)

February 23, 2024

Contents

Bayesian Networks

- Probabilistic graphical models representing relationships between variables
- Arrows show dependencies, conditional probability tables (CPTs) quantify strength

Bayesian Networks

Definition

A Bayesian Network on a set of random variables $\mathbf{X} = \{X_1, ..., X_n\}$ is a pair $\mathcal{B} = (G, P_r)$ composed of a directed acyclic graph $G = ([n], E)$ (where $[n] = \{1, ..., n\}$) and P_r specifies the conditional probabilities

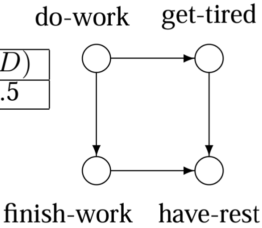
$$P_r\left(X_i = x_i | \mathbf{X}_{par(i)} = \mathbf{x}_{par(i)}\right)$$

for every $X_i \in \mathbf{X}$. $\mathcal{B}$ uniquely define the joint distribution on $\mathbf{X}$

$$P_r(\mathbf{X} = \mathbf{x}) = \prod_{i=1}^{n} P_r\left(X_i = x_i | \mathbf{X}_{par(i)} = \mathbf{x}_{par(i)}\right)$$

Bayesian Networks

- The work-rest BN (Sang, Beame, and Kautz 2005):



$p(D)$
0.5

D	$p(G)$
True	0.7
False	0.2

D	$p(F)$
True	0.6
False	0.1

F	G	$p(H)$
True	True	1
True	False	0.5
False	True	0.4
False	False	0

Probabilistic Inference

- Calculating probabilities of variables given evidence (decision-making, prediction, understanding relationships)
- **Marginal**: Probability of a single variable being in a particular state (e.g. $P(H = h)$)
- **Conditional**: Probability of one variable given the state of another (e.g. $P(G = g|D = d)$)
- **Joint**: Probability of a combination of values for multiple variables (e.g. $P(D = d, F = f, G = g, H = h)$)

Using WMC for probabilistic inference

- Problem: computing the probability of a query given an evidence in a Bayesian Network
- Translation from Bayesian network to weighted model counting

Using WMC for probabilistic inference

1. For every node p (*state variables*) with $k > 0$ parents introduce 2^k new propositional variables (*chance variables*) $p_{\mathbf{b}}$ for $\mathbf{b} \in \{0, 1\}^k$
2. To each of the introduced variables, associate the weight specified in the corresponding line of the CPT, i.e., $w(p_{\mathbf{b}}) = P_r(p = 1|par(p) = \mathbf{b})$
3. Set $w(\neg p_{\mathbf{b}})$ to $1 - w(p_{\mathbf{b}})$, and the weights of all the other literals to 1
4. For every $p_{\mathbf{b}}$ (Sang et al., 2005):
    - Define two lists: $L_1 = (\neg p_{\mathbf{b}}, p)$, and $L_2 = (p_{\mathbf{b}}, \neg p)$
    - For every $par(p)$:
        - If $par(p) = 1$: append $\neg par(p)$ to L_1 and L_2
        - Else: append $par(p)$ to L_1 and L_2
    - Append L_1 and L_2 to the initialized CNF() formula

Using WMC for probabilistic inference

5. Let $\Phi_{\mathcal{B}}$ be the conjunction of the formulas defined in step 4
6. Let $w_{\mathcal{B}}$ be the weight function given by

$$w_{\mathcal{B}}(\mathcal{I}) = \begin{cases} \prod_{\mathcal{I} \models p_{\mathbf{b}}} w(p_{\mathbf{b}}) \prod_{\mathcal{I} \not\models p_{\mathbf{b}}} w(\neg p_{\mathbf{b}}) & \text{if } \mathcal{I} \models \Phi_{\mathcal{B}} \\ 0 & \text{Otherwise} \end{cases}$$

7. Check that the weight of all models of $\Phi_{\mathcal{B}}$ is 1, namely

$$\sum_{\mathcal{I} \models \Phi_{\mathcal{B}}} w(\mathcal{I}) = 1$$

8. Answer the conditional probability query $P_r(\phi|\psi)$ by using

$$P_r(\phi|\psi) = \frac{\mathrm{WMC}(\Phi_{\mathcal{B}} \wedge \phi \wedge \psi, w_{\mathcal{B}})}{\mathrm{WMC}(\Phi_{\mathcal{B}} \wedge \psi, w_{\mathcal{B}})}$$

Encoding BN in WMC: an example

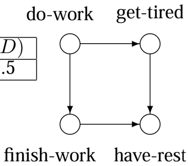
- **State variables**:

  $F(1), G(1), H(1), \neg F(1)...$

- **Chance variables**:

  $D(0.5)$,
  $f_0(0.1), f_1(0.6)$,
  $g_0(0.2), g_1(0.7)$,
  $h_{00}(0), h_{01}(0.4), h_{10}(0.5), h_{11}(1)...$

$p(D)$
0.5



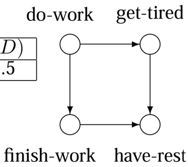
D	$p(G)$
True	0.7
False	0.2

D	$p(F)$
True	0.6
False	0.1

F	G	$p(H)$
True	True	1
True	False	0.5
False	True	0.4
False	False	0

Encoding BN in WMC: an example

- **Clauses per node**:
  - get-tired:
    $\{\{\neg D, \neg g_1, G\}, \{\neg D, g_1, \neg G\}, \{D, \neg g_0, G\}, \{D, g_0, \neg G\}\}$
  - finish-work:
    $\{\{\neg D, \neg f_1, F\}, \{\neg D, f_1, \neg F\}, \{D, \neg f_0, F\}, \{D, f_0, \neg F\}\}$
  - have-rest:
    $\{\{\neg F, \neg G, h_{11}, \neg H\}, \{\neg F, \neg G, \neg h_{11}, H\}, \{\neg F, G, h_{10}, \neg H\}, \{\neg F, G, \neg h_{10}, H\}, \{F, \neg G, h_{01}, \neg H\}, \{F, \neg G, \neg h_{01}, H\}, \{F, G, h_{00}, \neg H\}, \{F, G, \neg h_{00}, H\}\}$



$p(D)$
0.5

D	$p(G)$
True	0.7
False	0.2

D	$p(F)$
True	0.6
False	0.1

F	G	$p(H)$
True	True	1
True	False	0.5
False	True	0.4
False	False	0

Setup

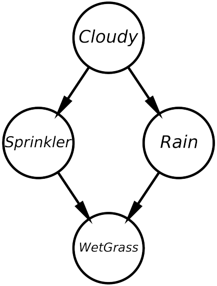
- The algorithm was implemented in **Python**, using the **bnlearn** package, and **Minisat22**

Implementation

- Input: Query and evidence
- Output: Probability via WMC
- Functions from the bnlearn package are used for validation
- The implementation was tested in **two** datasets

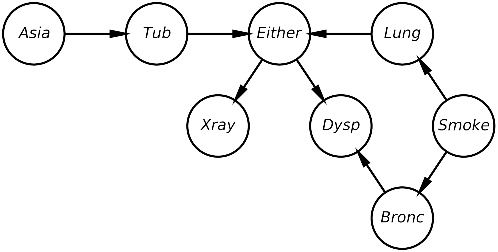
Implementation

- **Sprinkler**
  - 4 binary variables
  - 1,000 instances
  - The dataset represents a Bayesian network of four binary variables: *Cloudy*, *Sprinkler*, *Rain*, and *Wet Grass*. The variable Cloudy influences Sprinkler and Rain, which in turn affect Wet Grass.



Implementation

- **Asia**
  - 8 binary variables
  - 10,000 instances
  - The dataset represents a Bayesian network of eight binary variables: *Dyspnoea*, *Tuberculosis*, *Lung Cancer*, *Bronchitis*, *Visit to Asia*, *Smoking*, *Chest X-ray*, and *Tuberculosis versus Lung Cancer/Bronchitis*. It models the relationships between lung diseases and risk factors



Demo

Conclusion

- **Advantage**:
  - Flexible for various relationships and constraints
- **Limitation**:
  - Computational cost can grow exponentially with network size

Thanks for your attention