Marginalization in Lazy Propagation

Alg. 3.1.1 [Madsen's Dissertation]

Let $\Phi = \{\phi_1, \ldots, \phi_n\}$ be a set of potentials.

If marginalization of $X$ is invoked on $\Phi$, then

1. set $\Phi_x = \{\phi \in \Phi | x \in \text{dom}(\phi)\}$

2. $\phi_x^\ast = \sum_{\phi \in \Phi_x} \prod_{x \in \Phi \setminus \Phi_x} \phi$

3. $\Phi^\ast = \{\phi_x^\ast\} \cup \Phi \setminus \Phi_x$

$\Phi^\ast$ is the set of potentials that results from
eliminating $x$ from $\Phi$.

This is Defn. 4.1 on p. 116 with different terminology.

elimination for marginalisation

$\Phi^t$ for $X^t$.

4.8 Stochastic Simulation in Bayesian Networks

$P(E = y) \propto \frac{N(E = y)}{N} = \frac{\text{number of cases in which } E = y}{\text{total number of cases}}$
\[ P(B=y) = \frac{N(B=y)}{N} = \frac{62}{100} = .62 \]

\[ P(E) \approx \left( \frac{N(E=y)}{N}, \frac{N(E=n)}{N} \right) = \left( \frac{99}{100}, \frac{1}{100} \right) = .99, .01 \]

The algorithm (p. 468 [JDZ]).

1. Let \( \langle x_1, \ldots, x_n \rangle \) be a topological ordering of the variables (e.g., \( A, B, C, D, E \)).

2. For \( j = 1 \to N \):
   a) For \( i = 1 \to n \):
      - Sample a state \( x_i \) of \( X_i \) using \( P(X_i | \text{parent}(X_i)) \), where \( \text{parent}(X_i) \) is the configuration already sampled for \( \text{parent}(X_i) \).
   b) If \( \mathbf{x} = (x_1, \ldots, x_n) \) is consistent with \( E \), then
\[ \mathbb{N}(X_k = x_k) \] = \[ \mathbb{N}(X_k = x_k) + 1 \], where

- \( x_k \) is the state that was sampled for \( X_k \).

Let \( x \), the second \( x - \) because it is inconsistent with the evidence.

3. Return

\[
P(X_k = x_k | e) = \frac{N(X_k = x_k)}{\sum_{x \in \mathbb{X}} P(x_k) P(X_k = x_k | e)}
\]

\[
P(X_k = x_k | e) = \frac{P(X_k = x_k | e)}{P(e)}
\]

The alg. described requires generating a lot of irrelevant samples when the evidence has low probability.
4.8.2 Likelihood weighting was designed to overcome this problem.