Perfect simulation for Bayesian Networks

Andressa Cerqueira and Florencia Leonardi

II NeuroMat Workshop: New Frontiers in Neuromathematics

São Paulo, 22 de Novembro de 2016

Probabilistic Graphical Models

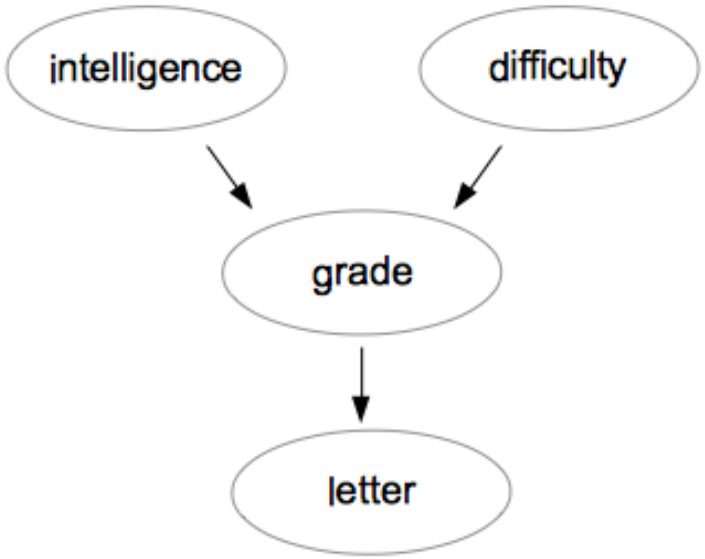
- Model: is a representation of our understanding of the world;
- Probabilistic: these models are designed to help us deal with large amounts of uncertainty;
- Graphical: the idea here is to use graphs to allow us to represent complex systems that involve a large number of variables.

Probabilistic Graphical Models - Example

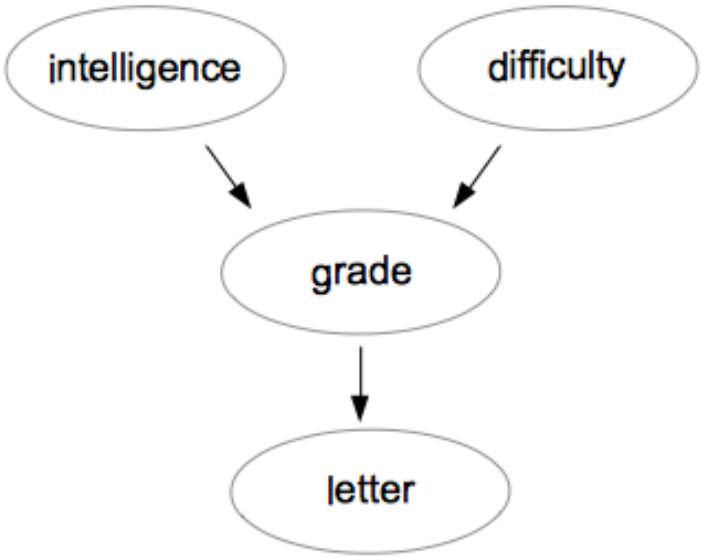
- Situation: A student who takes a course in the university;
- Variables: the intelligence of the student, the difficulty of the course, the grade of the student, the recommendation letter that the student gets

Probabilistic Graphical Models - Example

- Situation: A student who takes a course in the university;
- Variables: the intelligence of the student, the difficulty of the course, the grade of the student, the recommendation letter that the student gets

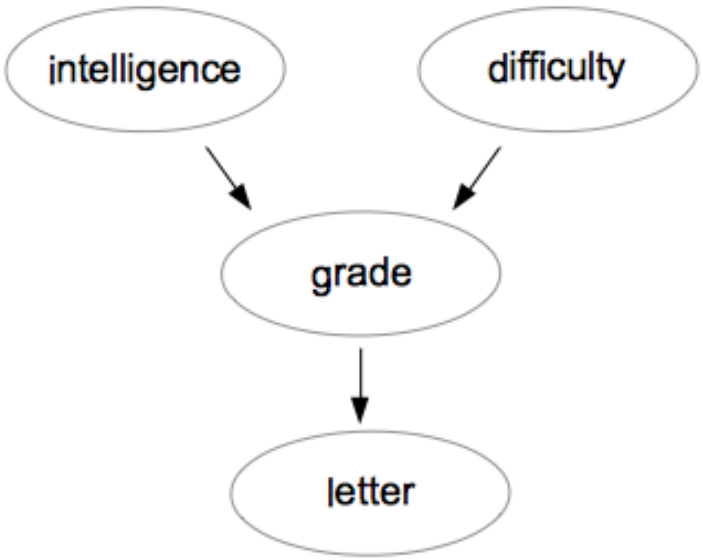


Bayesian Networks



- $\mathbb{P}(grade|intelligence, difficulty)$
- $\mathbb{P}(letter|grade)$
- letter and intelligence are conditionally independent given grade;
- letter and difficulty are conditionally independent given grade;
- intelligence and difficulty are called parents of grade;
- grade is called child of intelligence and difficulty.

Bayesian Networks

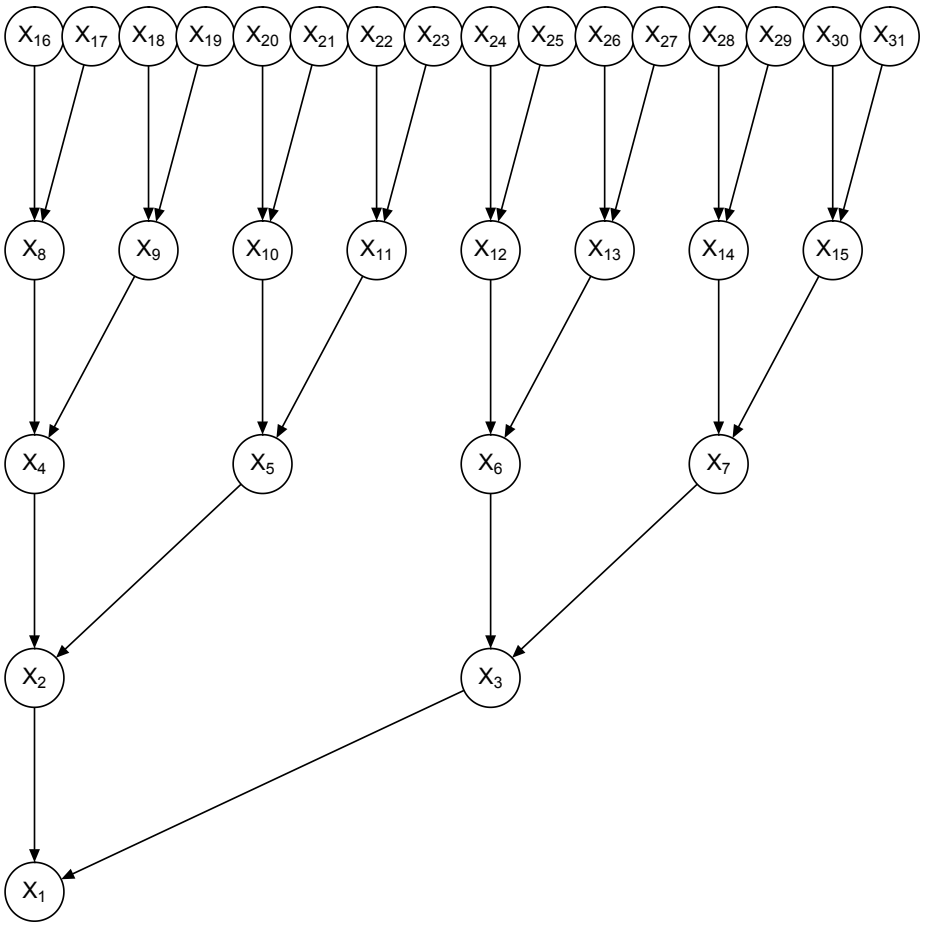


- Letter $\in \{$excelent, good, regular$\}$
- I want to sample from the variable Letter.
  How can I do it?

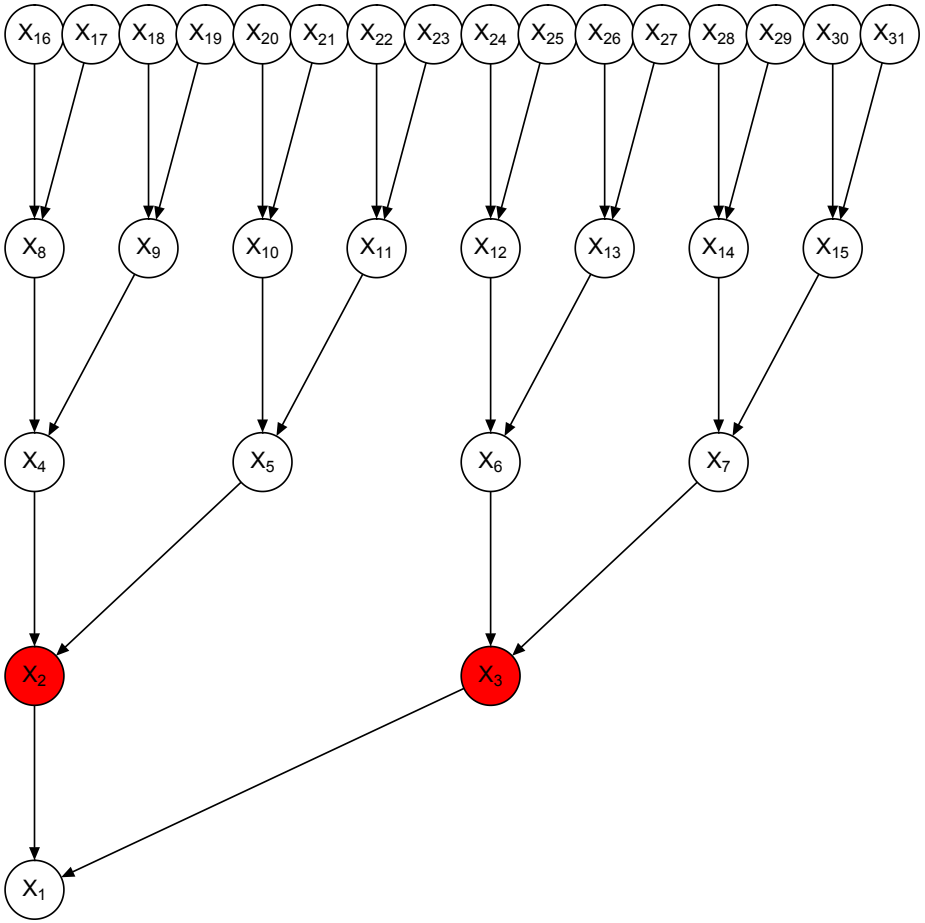
Bayesian Networks

Problem: How do we sample from a variable in a graph with thousands of vertices?

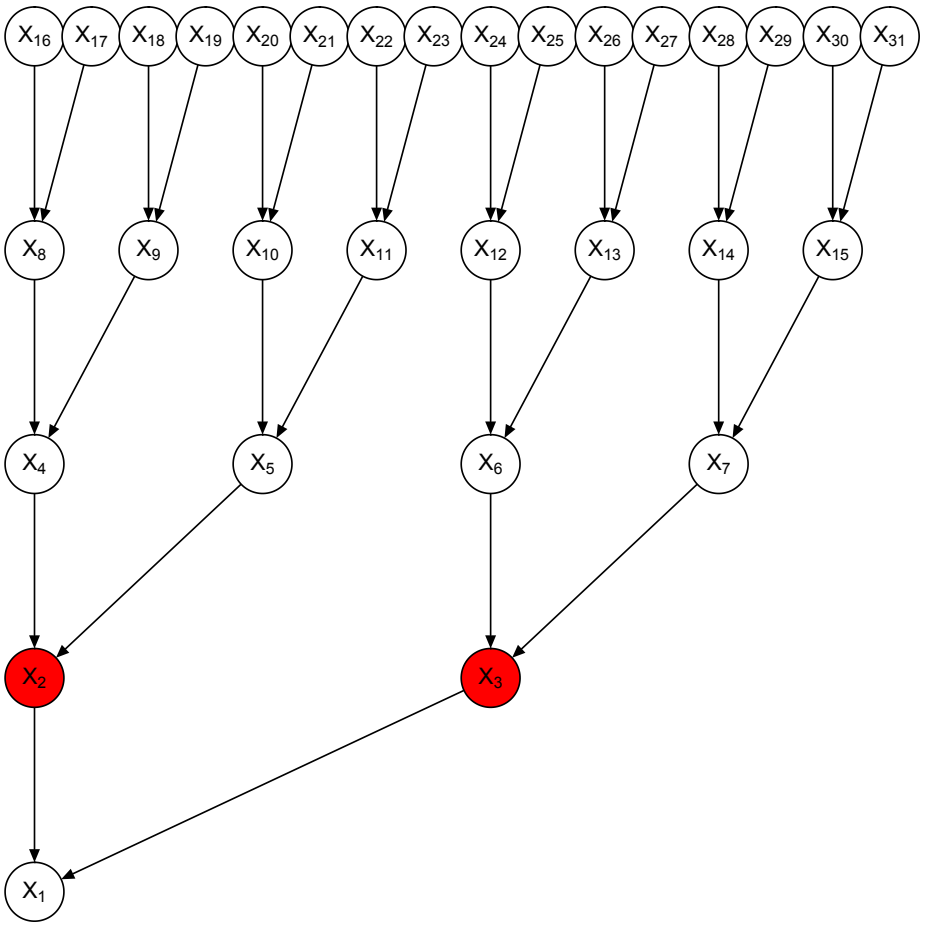
Sampling from all the vertices might be extremely demanding !!!



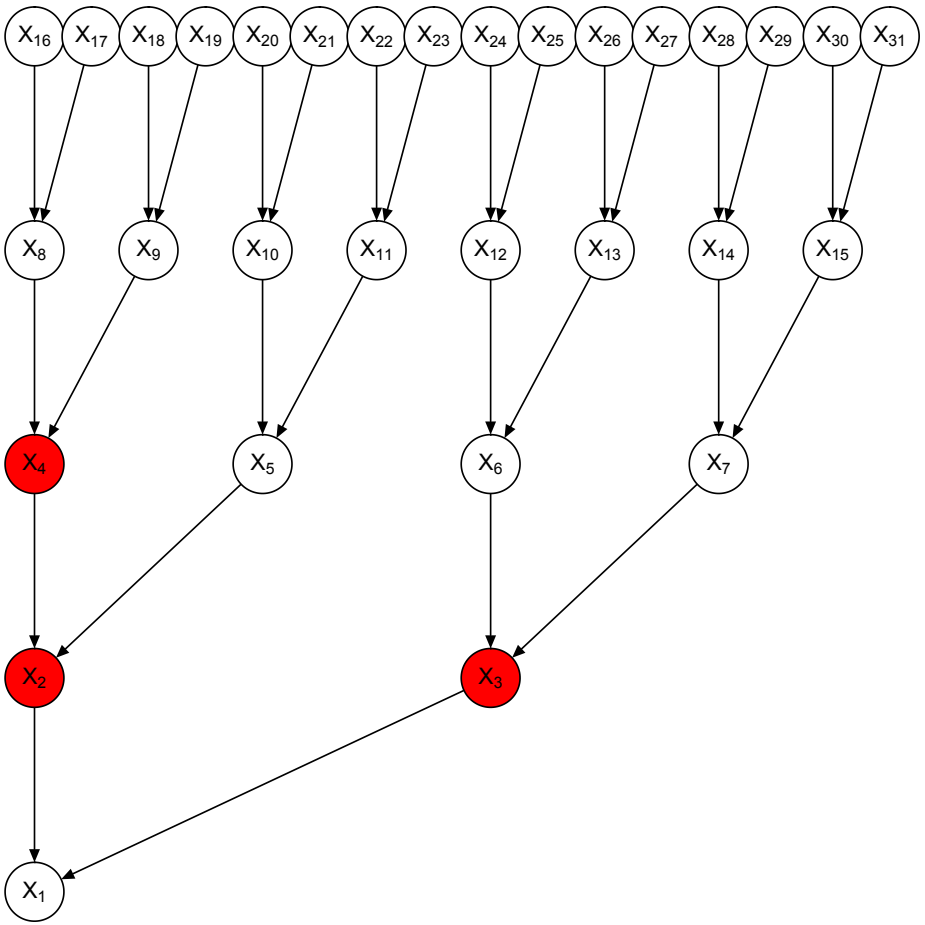
- To sample from X_1, we need to know:
- $\emptyset$, $\{X_2\}$, $\{X_2, X_3\}$



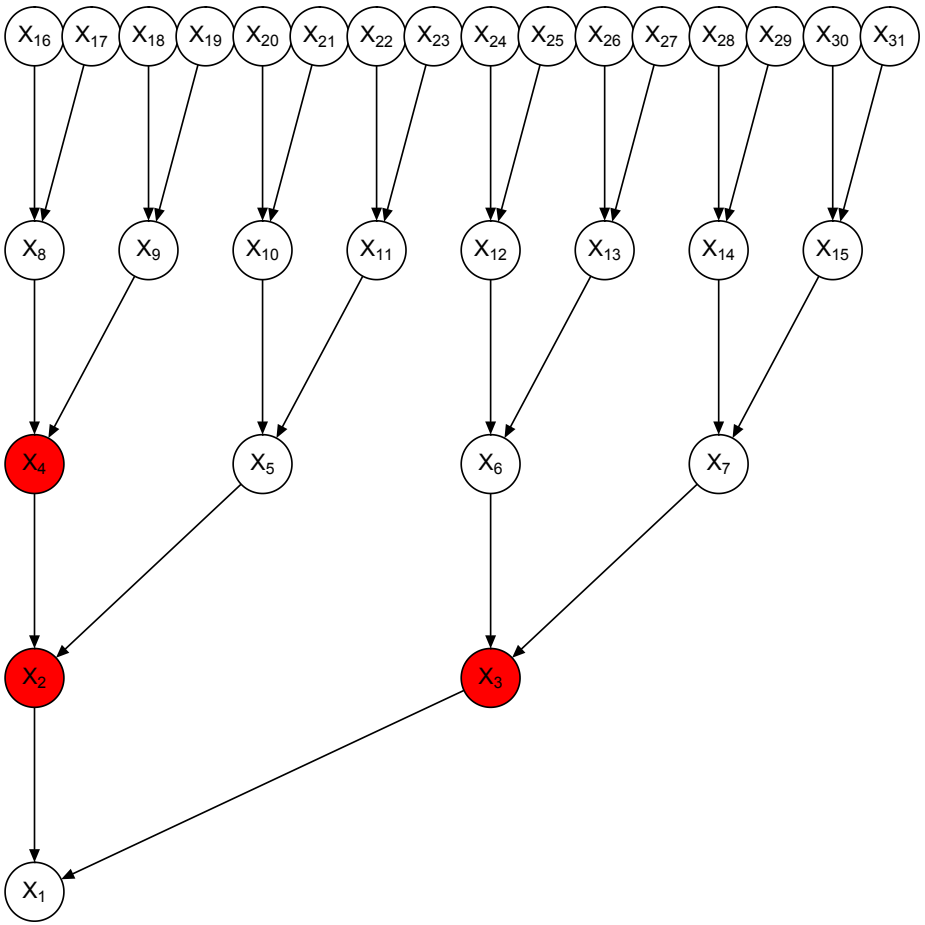
- $\emptyset$, $\{X_2\}$, $\{X_2, X_3\}$



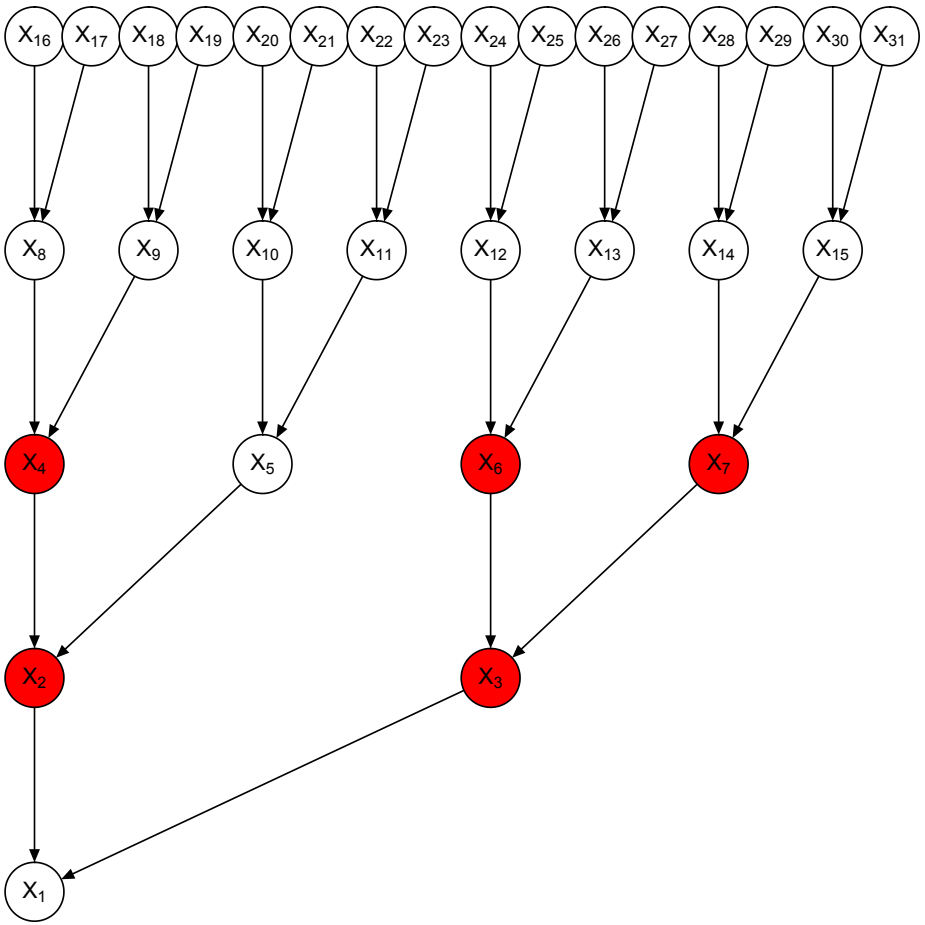
- To sample from X_2, we need to know:
- $\emptyset$, $\{X_4\}$, $\{X_4, X_5\}$



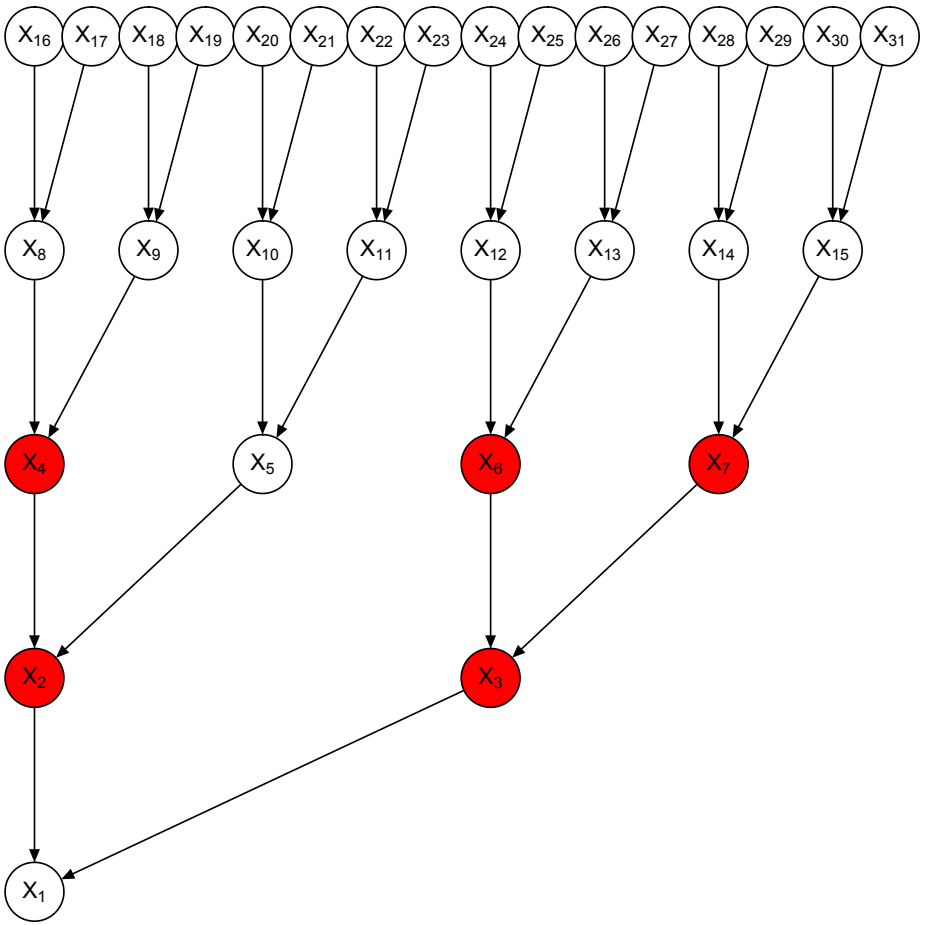
- $\emptyset$, $\{X_4\}$, $\{X_4, X_5\}$



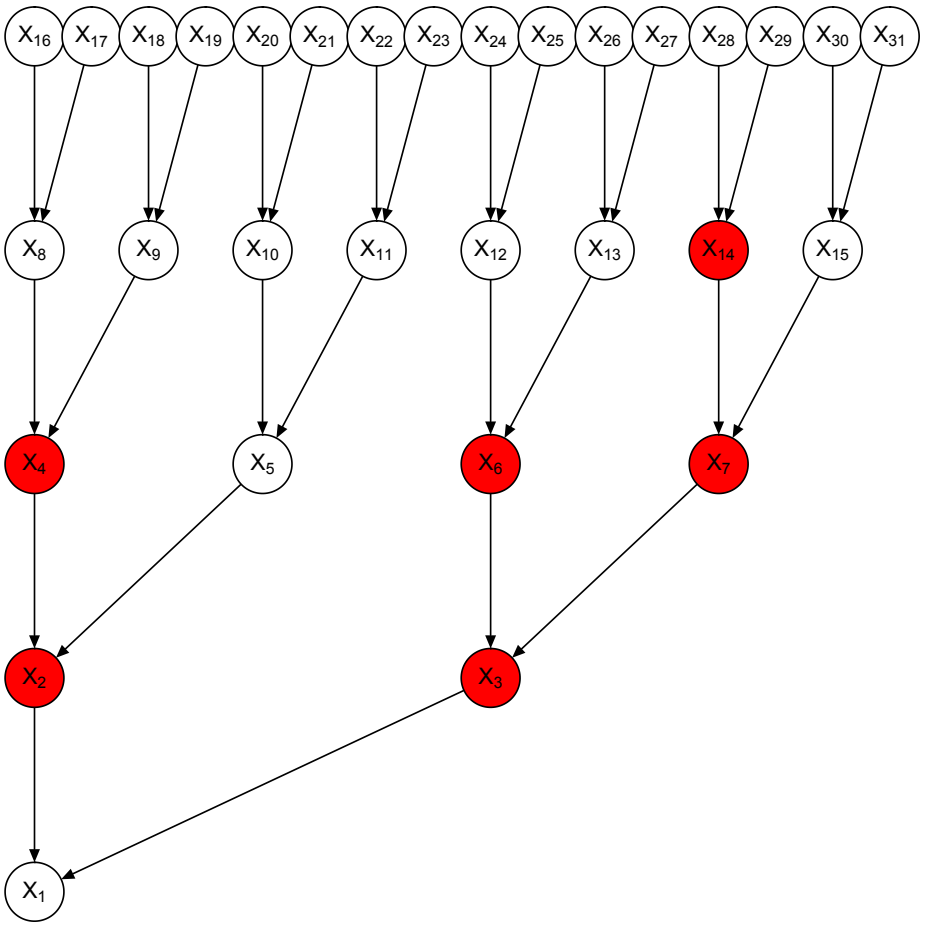
- To sample from X_3, we need to know:
- $\emptyset$, $\{X_6\}$, $\{X_6, X_7\}$



- $\emptyset$, $\{X_6\}$, $\{X_6, X_7\}$



- To sample from X_4, we need to know:
- $\emptyset$, $\{X_8\}$, $\{X_8, X_9\}$
- To sample from X_6, we need to know:
- $\emptyset$, $\{X_{12}\}$, $\{X_{12}, X_{13}\}$



- To sample from X_7, we need to know:
- $\emptyset$, $\{X_{14}\}$, $\{X_{14}, X_{15}\}$
- To sample from X_{14}, we need to know:
- $\emptyset$, $\{X_{28}\}$, $\{X_{28}, X_{29}\}$

Questions:

- How can we define the probability to select a subset of the parents of X_1?
  $\mathbb{P}(\text{to choose } \emptyset)$, $\mathbb{P}(\text{to choose} \{X_2\})$ and $\mathbb{P}(\text{to choose } \{X_2, X_3\})$
- How can we define the probability of X_1 given a subset of its parents?
  $\mathbb{P}(X_1|\emptyset)$, $\mathbb{P}(X_1|X_2)$ and $\mathbb{P}(X_1|X_2, X_3)$
- Study conditions to guarantee that the number of steps of the algorithm is sufficiently small.

References


Koller, Daphne and Friedman, Nir.
Probabilistic graphical models: principles and techniques.
MIT press.


Galves, Antonio, Löcherbach, Eva, & Orlandi, Enza. 2010.
Perfect simulation of infinite range Gibbs measures and coupling with their finite range approximations.
Journal of Statistical Physics, **138**(1-3), 476–495.


Galves, Antonio, Garcia, NL, Löcherbach, E, Orlandi, Enza, *et al.* . 2013.
Kalikow-type decomposition for multicolor infinite range particle systems.
The Annals of Applied Probability, **23**(4), 1629–1659.