Probabilistic Models and Machine Learning

No. 7

Graphical Models

Hui Jiang

Department of Electrical Engineering and Computer Science
York University, Toronto, Canada

Outline

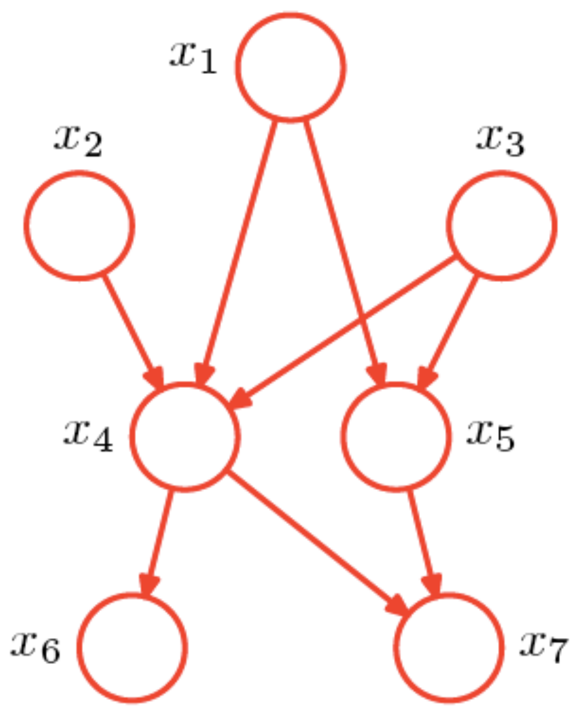
- Graphical Model: concepts
- Graphical Model: types
  - Directed Graphical Models (aka. Bayesian Networks)
  - Indirected Graphical Models (aka. Markov Random Fields)
- Exact Inference
  - Example: a chain model
  - Sum-product algorithm
  - Max-sum algorithm
- Approximate Inference
  - Loopy Belief Propagation
  - Variational Inference
  - Expectation Propagation
  - Monte Carlo Sampling

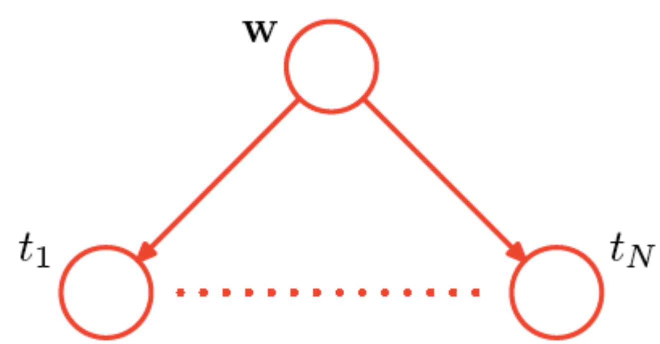
Graphical Model

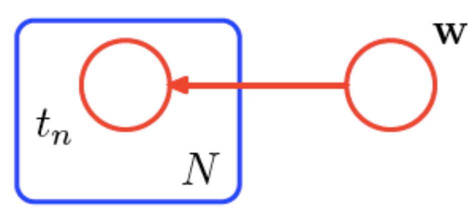
- **Use a graph to represent joint distributions of random variables**
  - **Nodes —> random variables (RV)**
  - **Linking —> dependency among RVs**
- **Graphical Models may imply conditional independence among RVs.**

Bayesian Networks (1)

- **Use a directed graph to represent joint distributions of random variables**
  - **Nodes —> random variables (RV)**
  - **Linking —> conditional distribution of children given the parents**

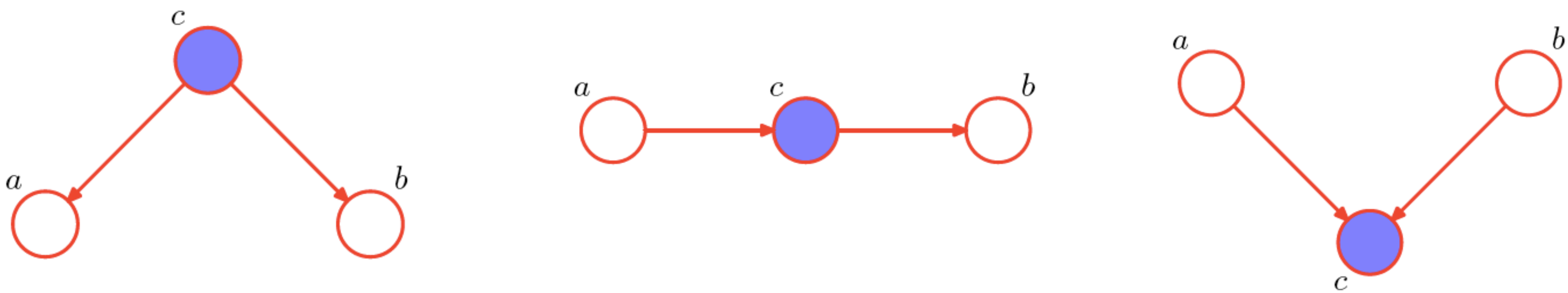




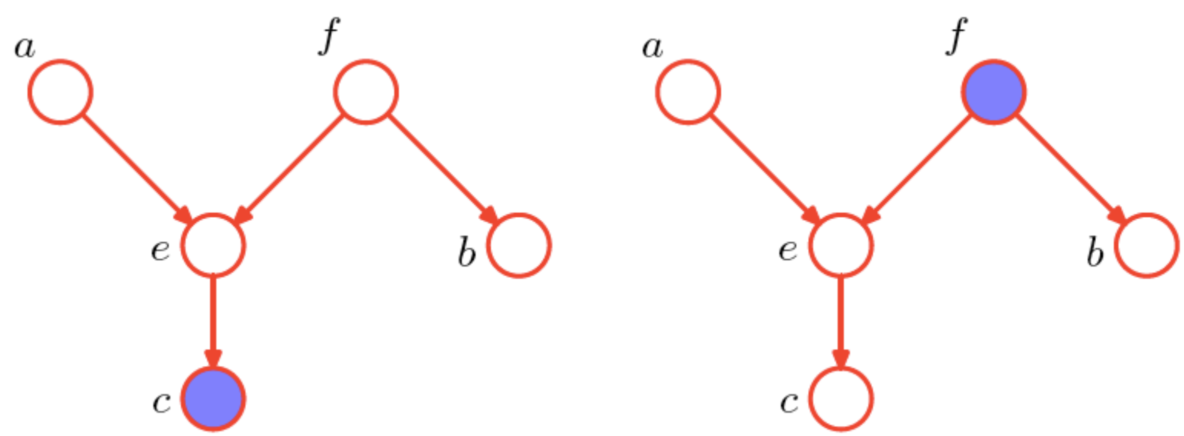


Bayesian Networks (2)

- **Conditional independence in Bayesian Networks**
  - **tail-to-tail, head-to-tail, head-to-head (explain-away)**

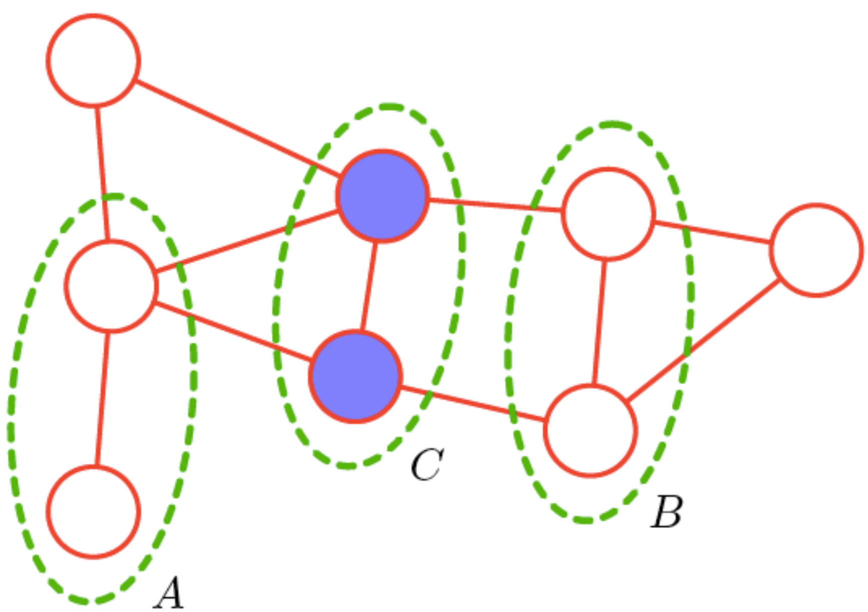


  - **D-separation rule**



Markov Random Fields (1)

- **Use an undirected graph to represent joint distributions of random variables**
  - **Nodes —> random variables (RV)**
  - **Linking —> conditional dependency**
- **Conditional independence == simple graph separation**



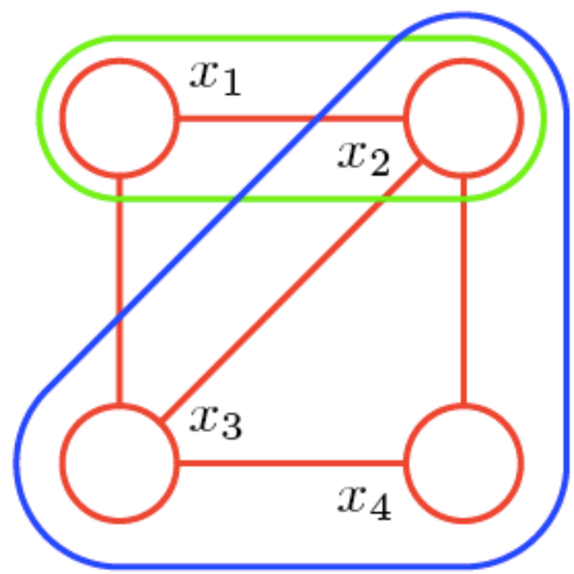
Markov Random Fields (2)

- How to form the joint probability distribution?
  - Potential functions: defined over maximal cliques

$$p(\mathbf{x}) = \frac{1}{Z}\prod_C \psi_C(\mathbf{x}_C).$$

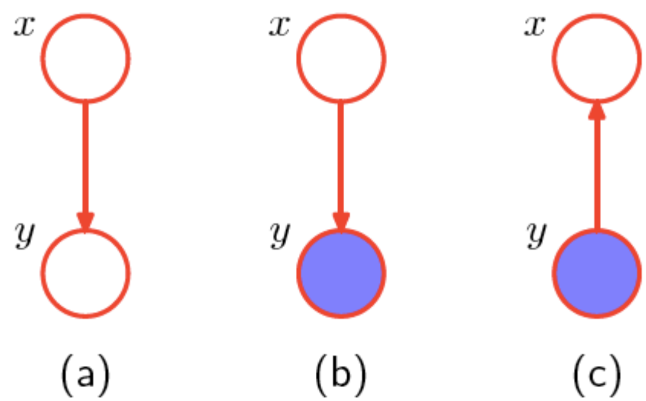
  - Partition function: normalization constant

$$Z = \sum_{\mathbf{x}}\prod_C \psi_C(\mathbf{x}_C)$$

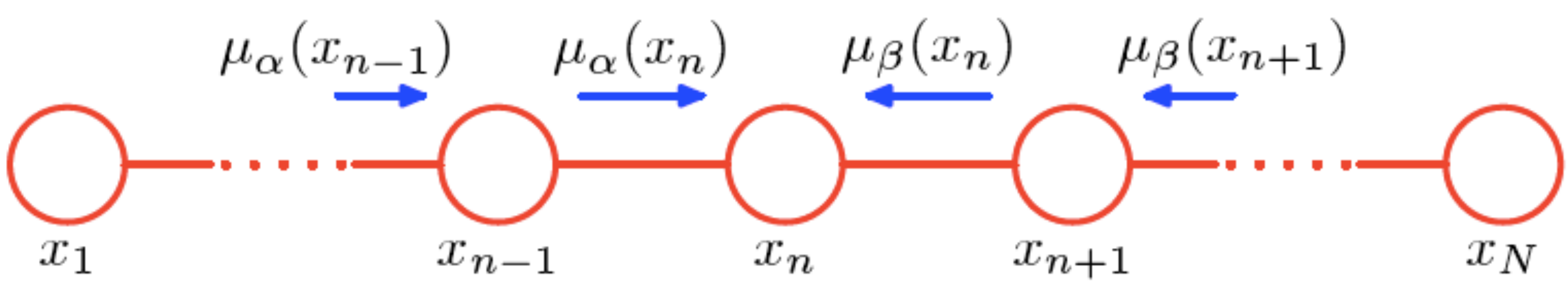


Exact Inference In Graphical Models

- **What is Inference?**
- **Message Propagation for Inference**
  - **Two nodes —> Bayes' theorem**



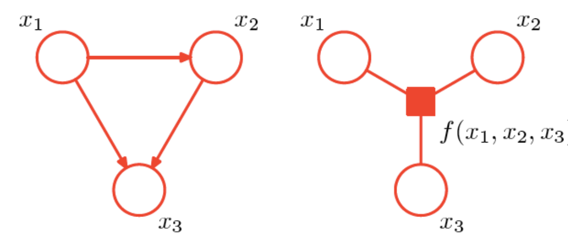
  - **Inference on a chain**



Exact Inference In Graphical Models

- **Tree-structured Graphical models**
  - **Sum-Product (Max-sum) algorithm**
- **General Graphs**
  - **Junction tree algorithms**
  - **Computationally expensive**

Sum-Product algorithm (1)

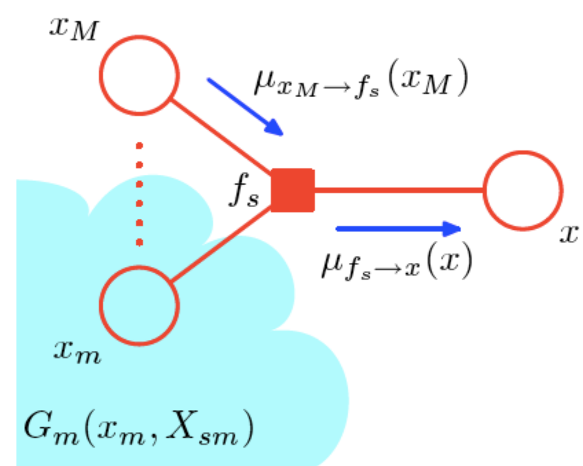


- **Form factor graphs**
- **Message propagation**
  - **factor node —> variable node**

$$\mu_{f_s \to x}(x) = \sum_{x_1} \cdots \sum_{x_M} f_s(x, x_1, \ldots, x_M) \prod_{m \in \mathrm{ne}(f_s) \setminus x} \mu_{x_m \to f_s}(x_m)$$

  - **variable node —> factor node**

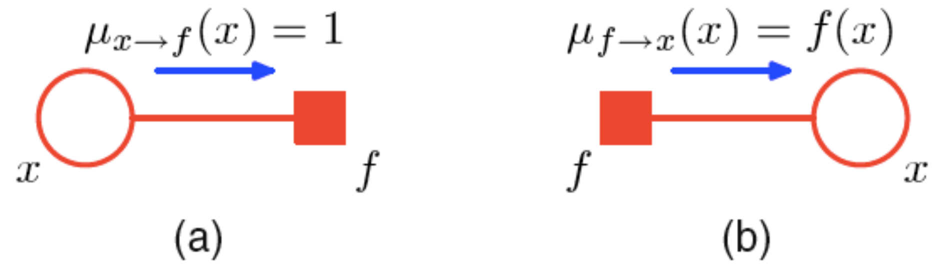
$$\mu_{x_m \to f_s}(x_m) = \prod_{l \in \mathrm{ne}(x_m) \setminus f_s} \mu_{f_l \to x_m}(x_m)$$

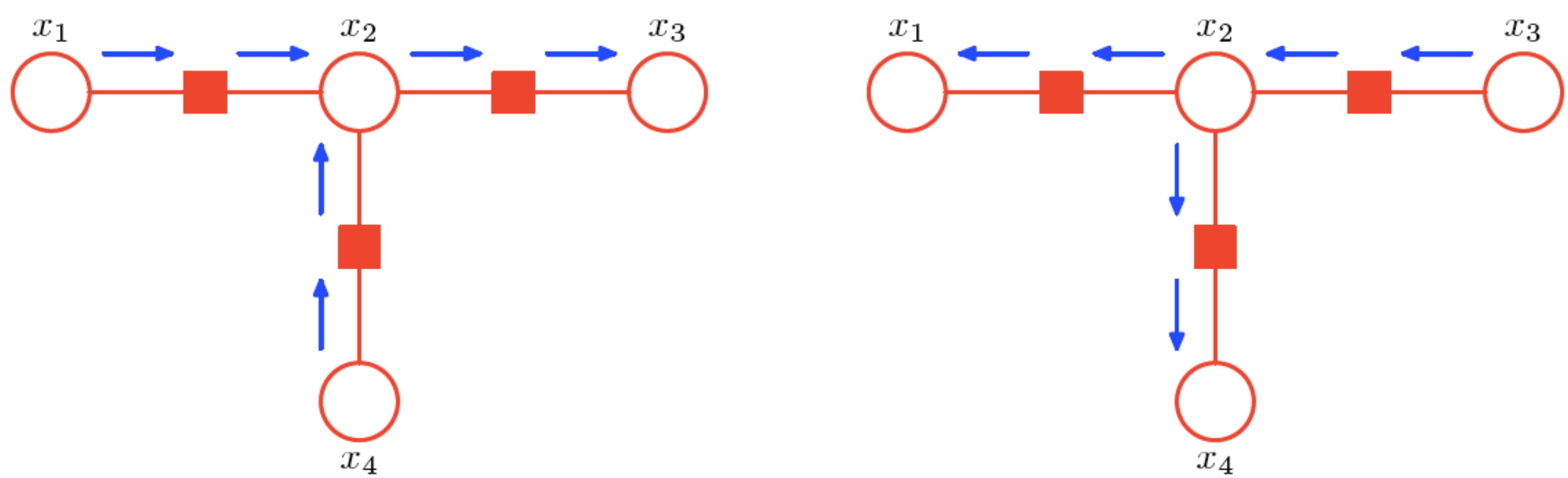


- **Two-pass propagation**
  - **from leafs to root**
  - **from root to leafs**

Sum-Product algorithm (2)

- **Two-pass propagation over the whole graph with initial messages**
  - **from leafs to root**
  - **from root to leafs**





Approximate Inference Methods

- **Loopy Belief Propagation**
- **Variational Inference**
- **Expectation Propagation**
- **Mento Carlo Sampling**

Variational Inference

- Approximate posterior distributions by the variational distributions
- Variational Distributions:
  - Factorized approximation based on mean field theory
- Variational message passing
  - updating is done via a local calculation on the graph
  - Applicable to large networks
- Example: Variational Bayes of Gaussian mixture models

Mento Carlo Sampling

- **Estimate p(x2, x5, x6, x7 | x1, x3, x4)**