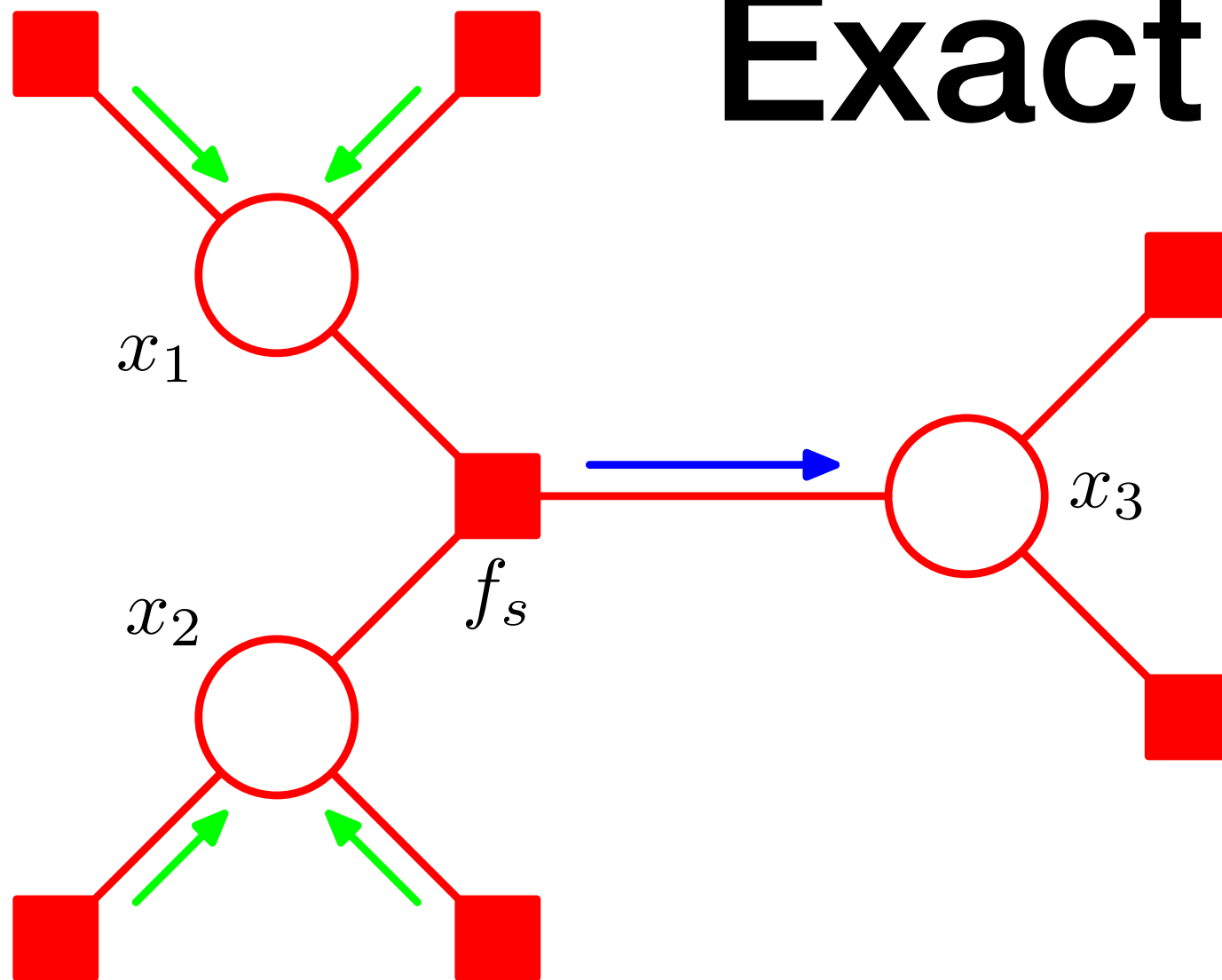


# Exact Inference



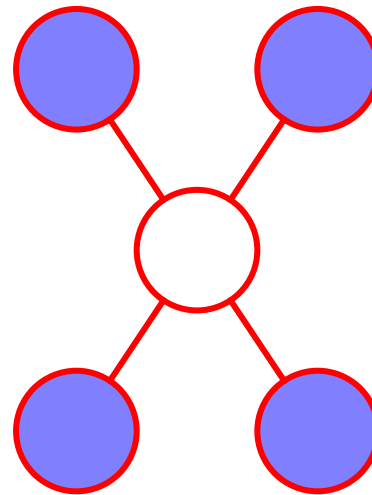
Factor Graphs through Max-Sum Algorithm

Figures from Bishop PRML Sec. 8.3/8.4

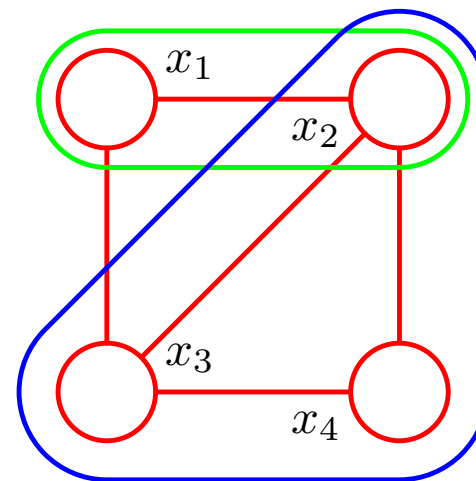
# Building Blocks

UGMs, Cliques, Factor Graphs

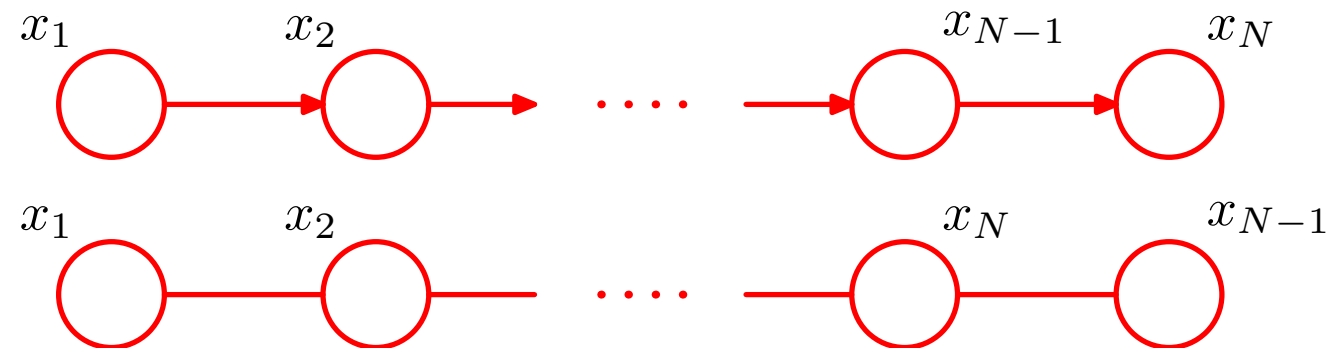
## Markov Random Fields / UGMs



**Parameterization: maximal cliques**



## Example: Equivalent DGM and UGM

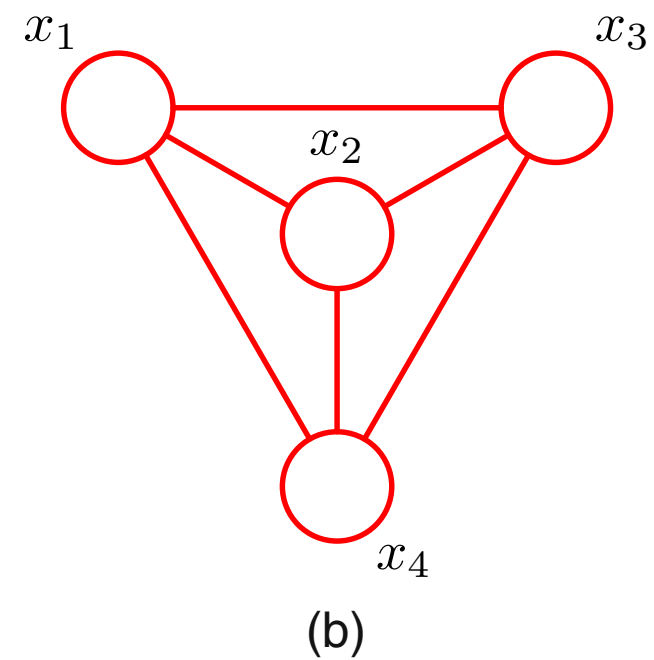
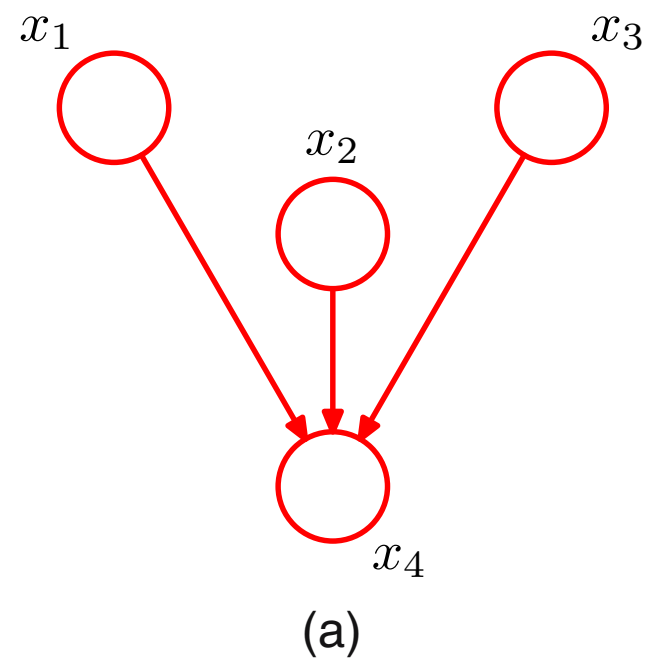


$$p(\mathbf{x}) = p(x_1)p(x_2|x_1)p(x_3|x_2) \cdots p(x_N|x_{N-1}).$$

$$p(\mathbf{x}) = \frac{1}{Z} \psi_{1,2}(x_1, x_2) \psi_{2,3}(x_2, x_3) \cdots \psi_{N-1,N}(x_{N-1}, x_N)$$

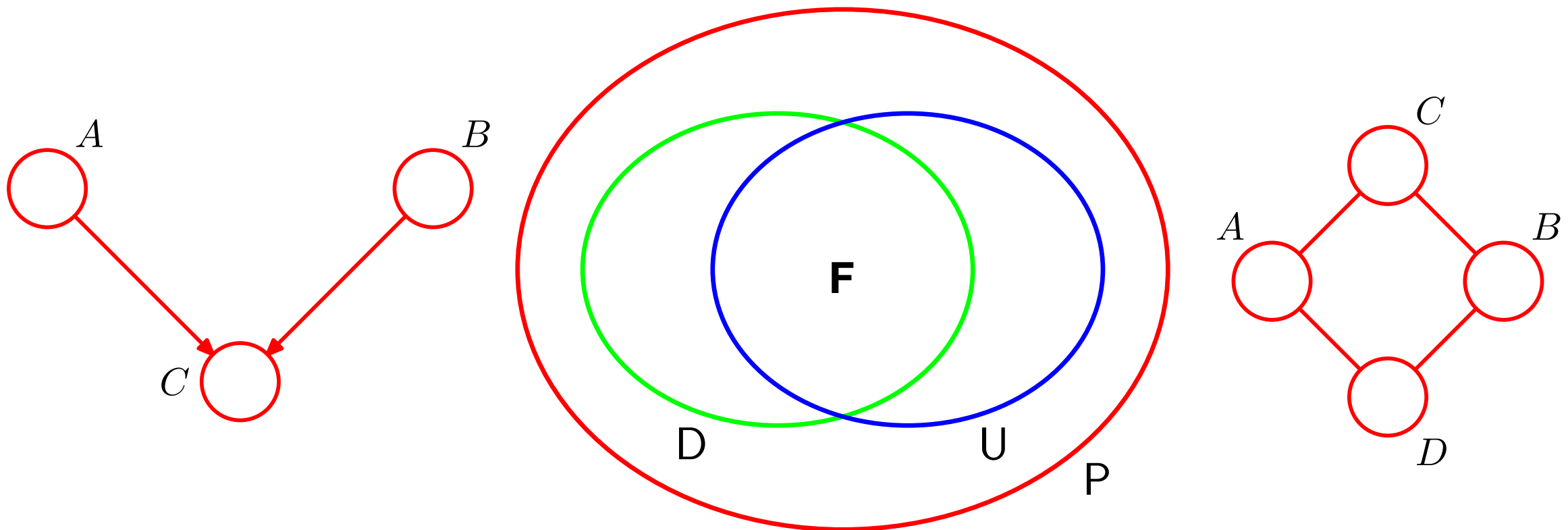
$$\begin{aligned} \psi_{1,2}(x_1, x_2) &= p(x_1)p(x_2|x_1) \\ \psi_{2,3}(x_2, x_3) &= p(x_3|x_2) \\ &\vdots \\ \psi_{N-1,N}(x_{N-1}, x_N) &= p(x_N|x_{N-1}) \end{aligned}$$

## Conversion: “Moralization” (Marry the Parents of Every Child)



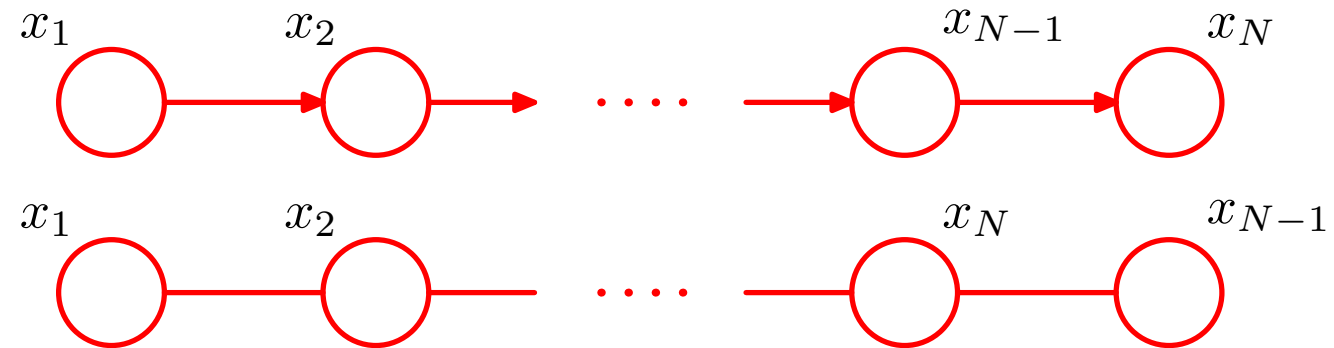
$$p(\mathbf{x}) = p(x_1)p(x_2)p(x_3)p(x_4|x_1, x_2, x_3)$$

## DGMs and UGMs represent distinct distributions



# **Motivations: Exact Inference in a Chain**

## Query probability of a configuration for node $X_n$ : $p(\mathbf{x}_n)$

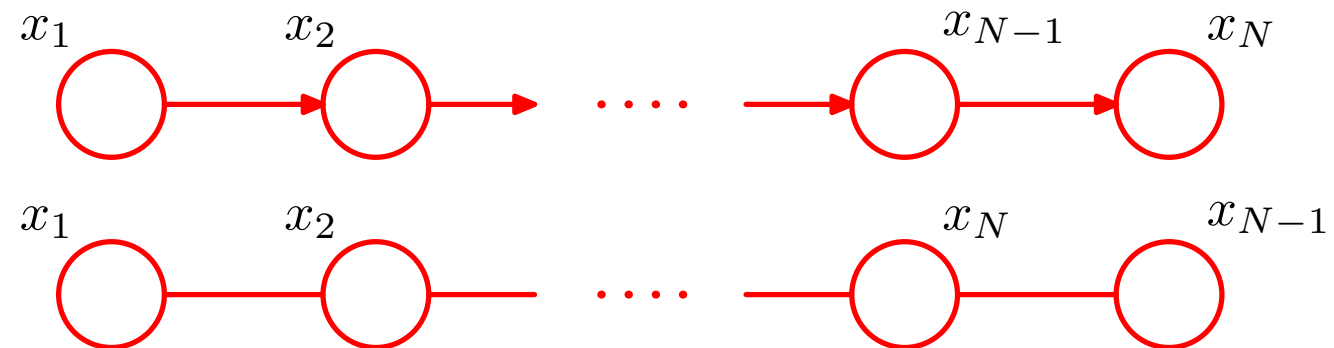


$$p(\mathbf{x}) = \frac{1}{Z} \psi_{1,2}(x_1, x_2) \psi_{2,3}(x_2, x_3) \cdots \psi_{N-1,N}(x_{N-1}, x_N)$$

$$p(x_n) = \sum_{x_1} \cdots \sum_{x_{n-1}} \sum_{x_{n+1}} \cdots \sum_{x_N} p(\mathbf{x}).$$



**Query probability of a configuration for node  $X_n$ :  $p(\mathbf{x}_n)$**



$$p(\mathbf{x}) = \frac{1}{Z} \psi_{1,2}(x_1, x_2) \psi_{2,3}(x_2, x_3) \cdots \psi_{N-1,N}(x_{N-1}, x_N)$$

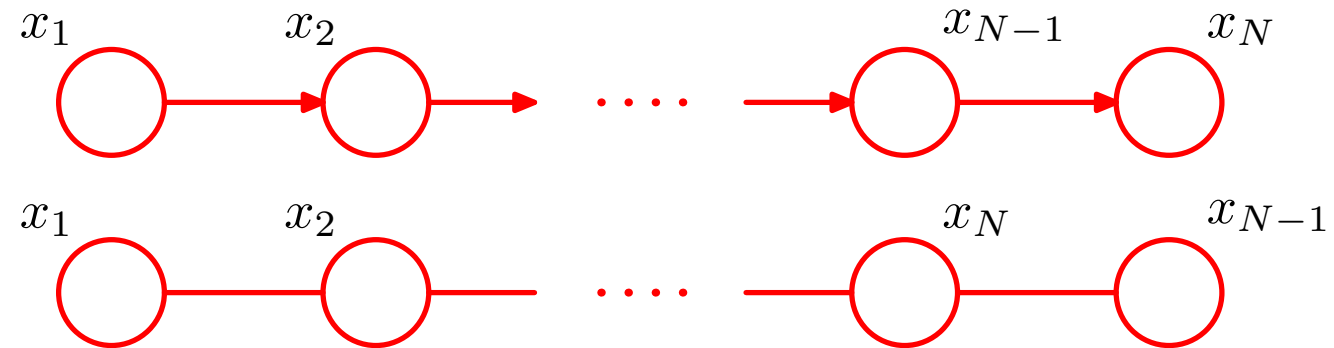
$$p(x_n) = \sum_{x_1} \cdots \sum_{x_{n-1}} \sum_{x_{n+1}} \cdots \sum_{x_N} p(\mathbf{x}).$$

**Naively:**

**N variables, K states per variable: computation complexity?**

$$p(x_n) = \sum_{x_1} \cdots \sum_{x_{n-1}} \sum_{x_{n+1}} \cdots \sum_{x_N} p(\mathbf{x}).$$

**Query probability of a configuration for node  $X_n$ :  $p(\mathbf{x}_n)$**



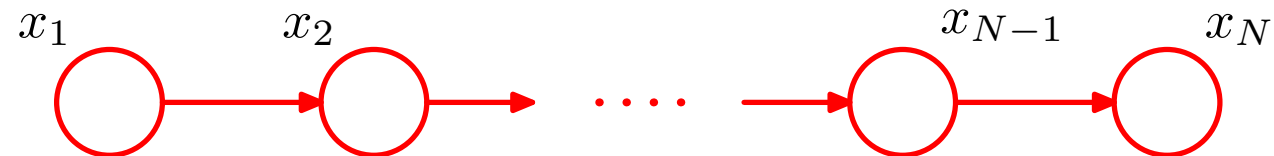
$$p(\mathbf{x}) = \frac{1}{Z} \psi_{1,2}(x_1, x_2) \psi_{2,3}(x_2, x_3) \cdots \psi_{N-1,N}(x_{N-1}, x_N)$$

$$p(x_n) = \sum_{x_1} \cdots \sum_{x_{n-1}} \sum_{x_{n+1}} \cdots \sum_{x_N} p(\mathbf{x}).$$

**We ignored the conditional independence! Notice for  $x_n$ :**

$$\sum_{x_N} \psi_{N-1,N}(x_{N-1}, x_N)$$

**Be clever about order of computation:**

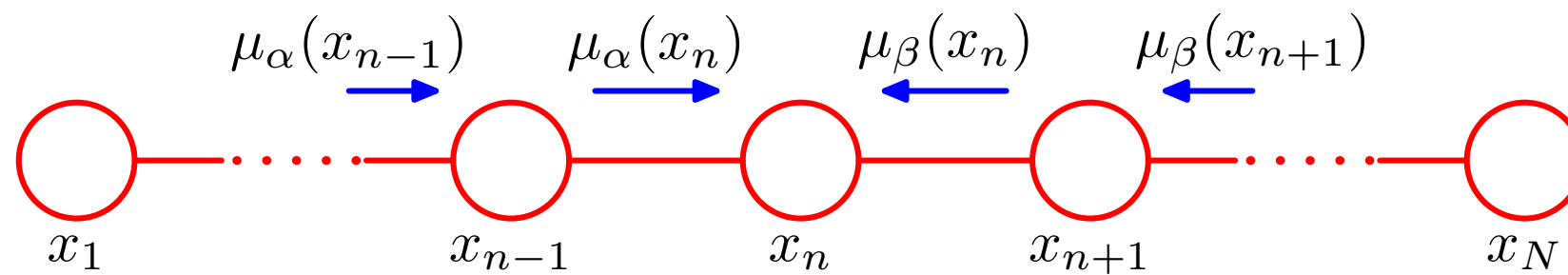


$$p(x_n) = \frac{1}{Z}$$

$$\underbrace{\left[ \sum_{x_{n-1}} \psi_{n-1,n}(x_{n-1}, x_n) \cdots \left[ \sum_{x_2} \psi_{2,3}(x_2, x_3) \left[ \sum_{x_1} \psi_{1,2}(x_1, x_2) \right] \right] \cdots \right]}_{\mu_\alpha(x_n)}
 \underbrace{\left[ \sum_{x_{n+1}} \psi_{n,n+1}(x_n, x_{n+1}) \cdots \left[ \sum_{x_N} \psi_{N-1,N}(x_{N-1}, x_N) \right] \cdots \right]}_{\mu_\beta(x_n)} . \quad (8.52)$$

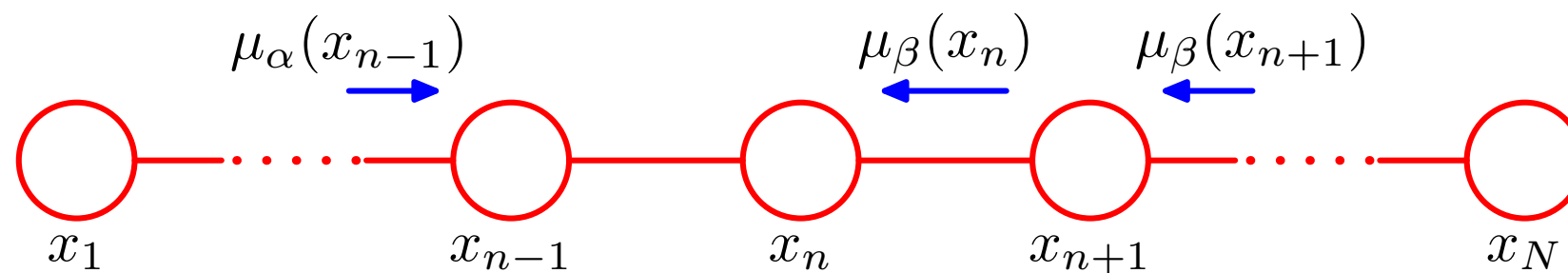
$$p(x_n) = \frac{1}{Z} \mu_\alpha(x_n) \mu_\beta(x_n)$$

**Be clever about order of computation:**



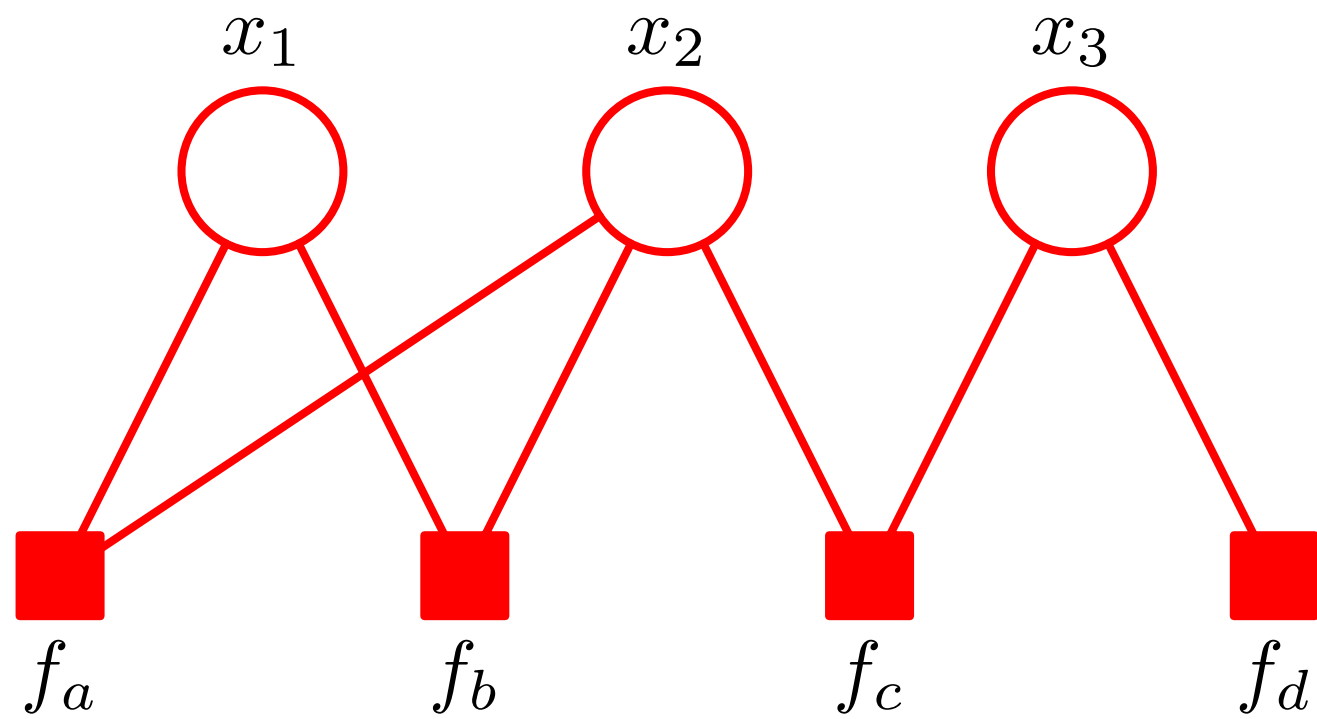
$$p(x_n) = \frac{1}{Z} \mu_\alpha(x_n) \mu_\beta(x_n)$$

**We get joint marginals over variables, too:**



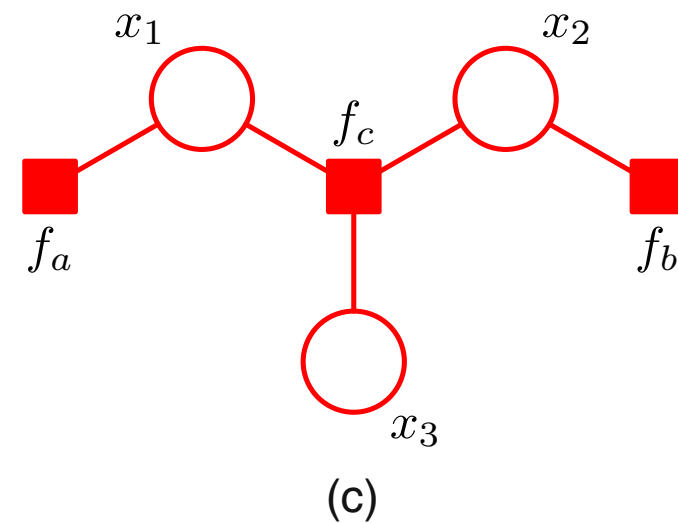
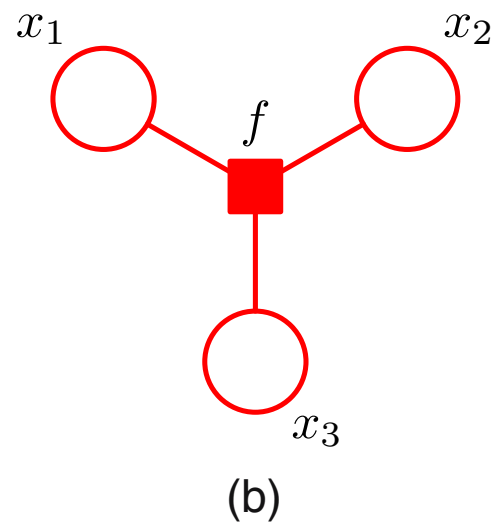
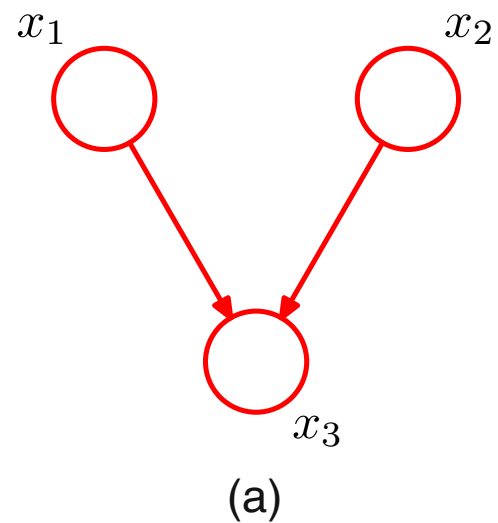
$$p(x_{n-1}, x_n) = \frac{1}{Z} \mu_\alpha(x_{n-1}) \psi_{n-1,n}(x_{n-1}, x_n) \mu_\beta(x_n).$$

# Factor Graph Review



$$p(\mathbf{x}) = \prod_s f(\mathbf{x}_s)$$

$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_1, x_2) f_c(x_2, x_3) f_d(x_3)$$

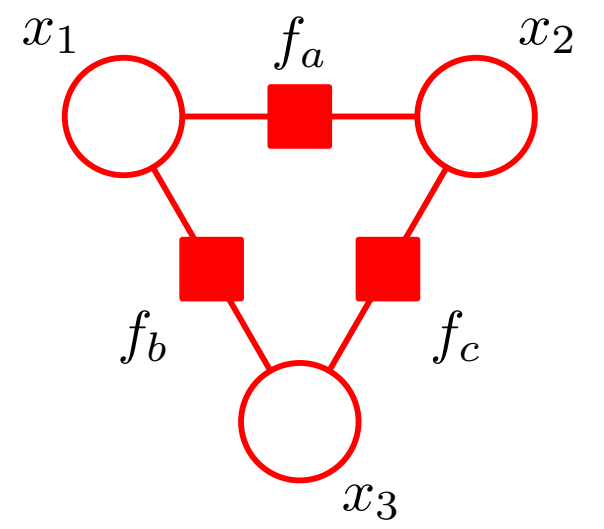
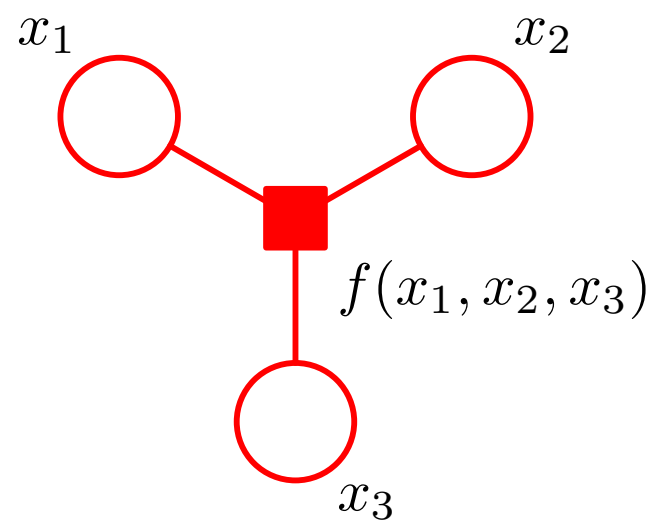
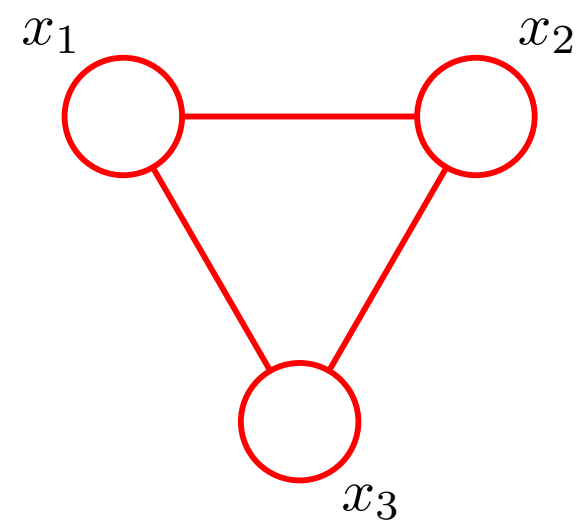


$$(a) \quad p(x_1)p(x_2)p(x_3|x_1, x_2)$$

$$(b) \quad f(x_1, x_2, x_3) = p(x_1)p(x_2)p(x_3|x_1, x_2)$$

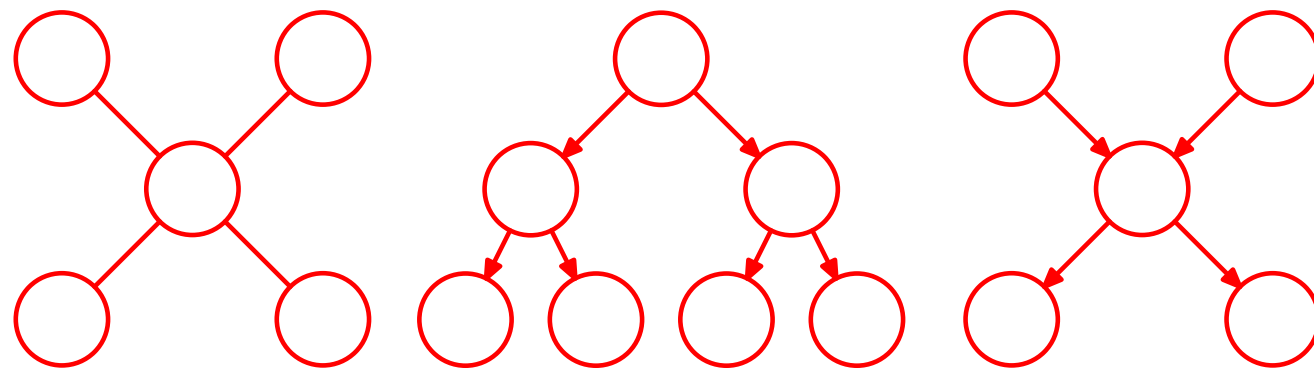
$$(c) \quad f_a(x_1) = p(x_1), f_b(x_2) = p(x_2), f_c(x_1, x_2, x_3) = p(x_3|x_1, x_2)$$



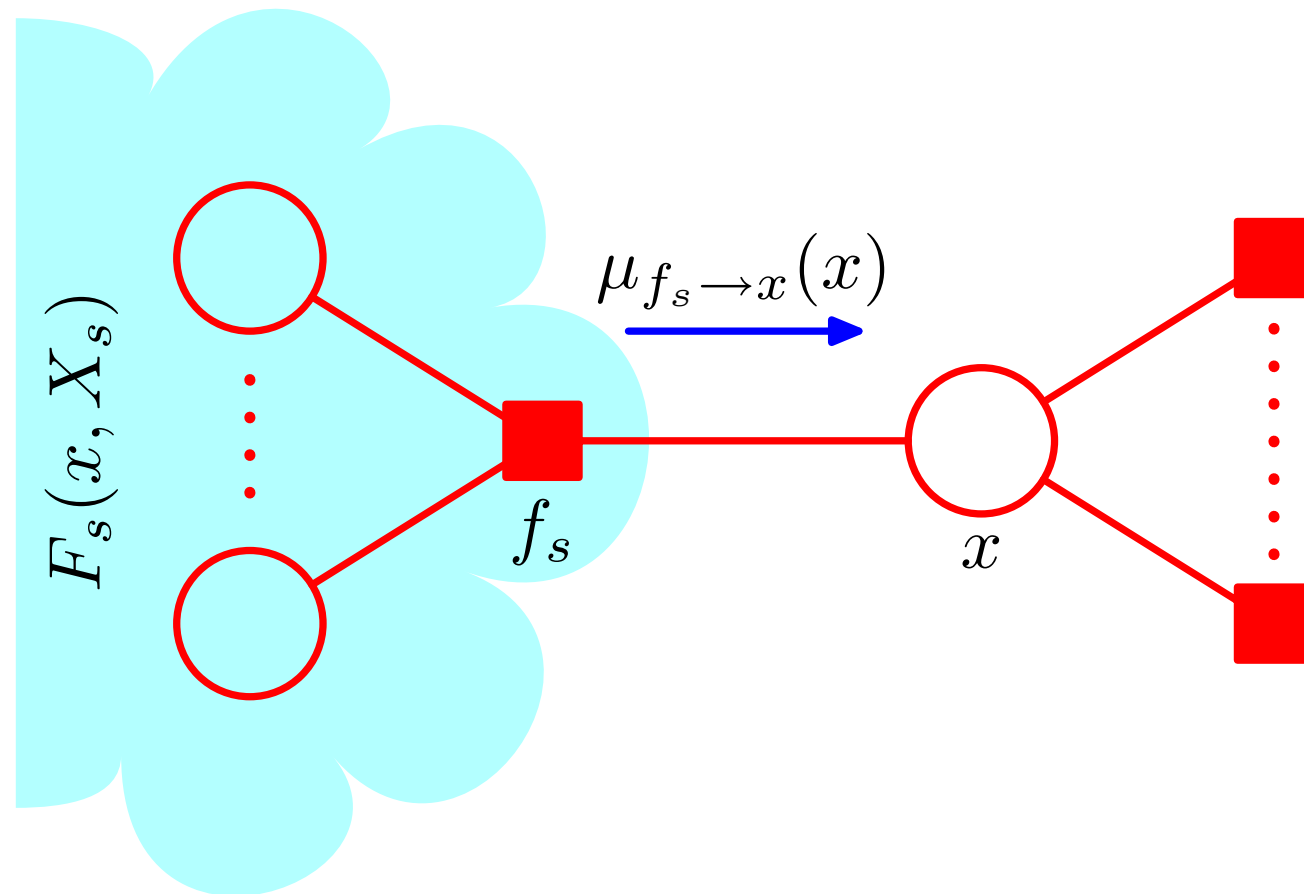


# Sum-Product Algorithm

Generalize Exact Inference in Chains  
to Tree-Structured PGMs



## Problem setup: notation



$$p(x) = \sum_{\mathbf{x} \setminus x} p(\mathbf{x})$$

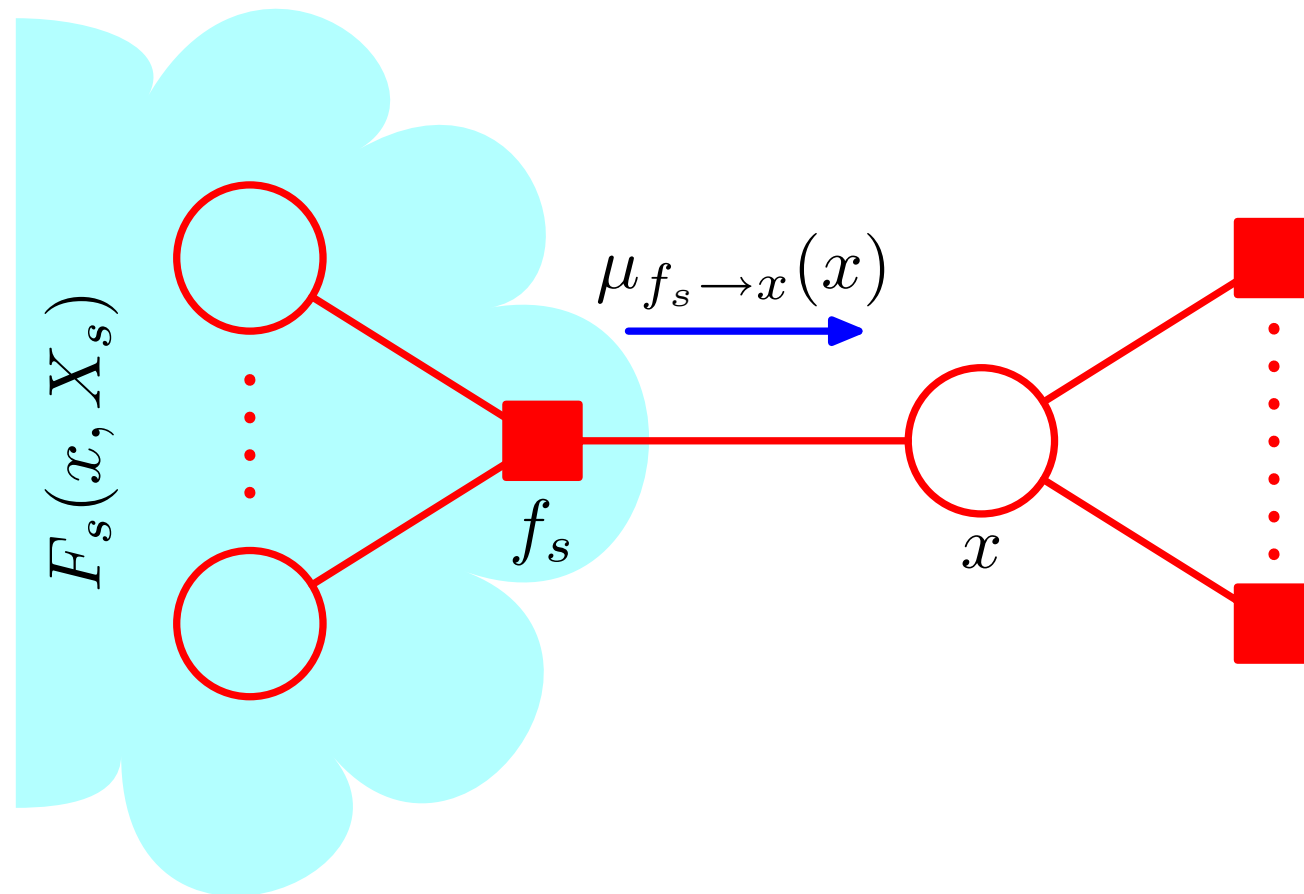
$$p(\mathbf{x}) = \prod_{s \in \text{ne}(x)} F_s(x, X_s)$$

**ne(x)**: set of factor nodes that are neighbours of  $x$

**X<sub>s</sub>**: set of all variables in the subtree connected to the variable node  $x$  via factor node  $f_s$

**F<sub>s</sub>(x, X<sub>s</sub>)**: the product of all the factors in the group associated with factor  $f_s$

## Problem setup: notation



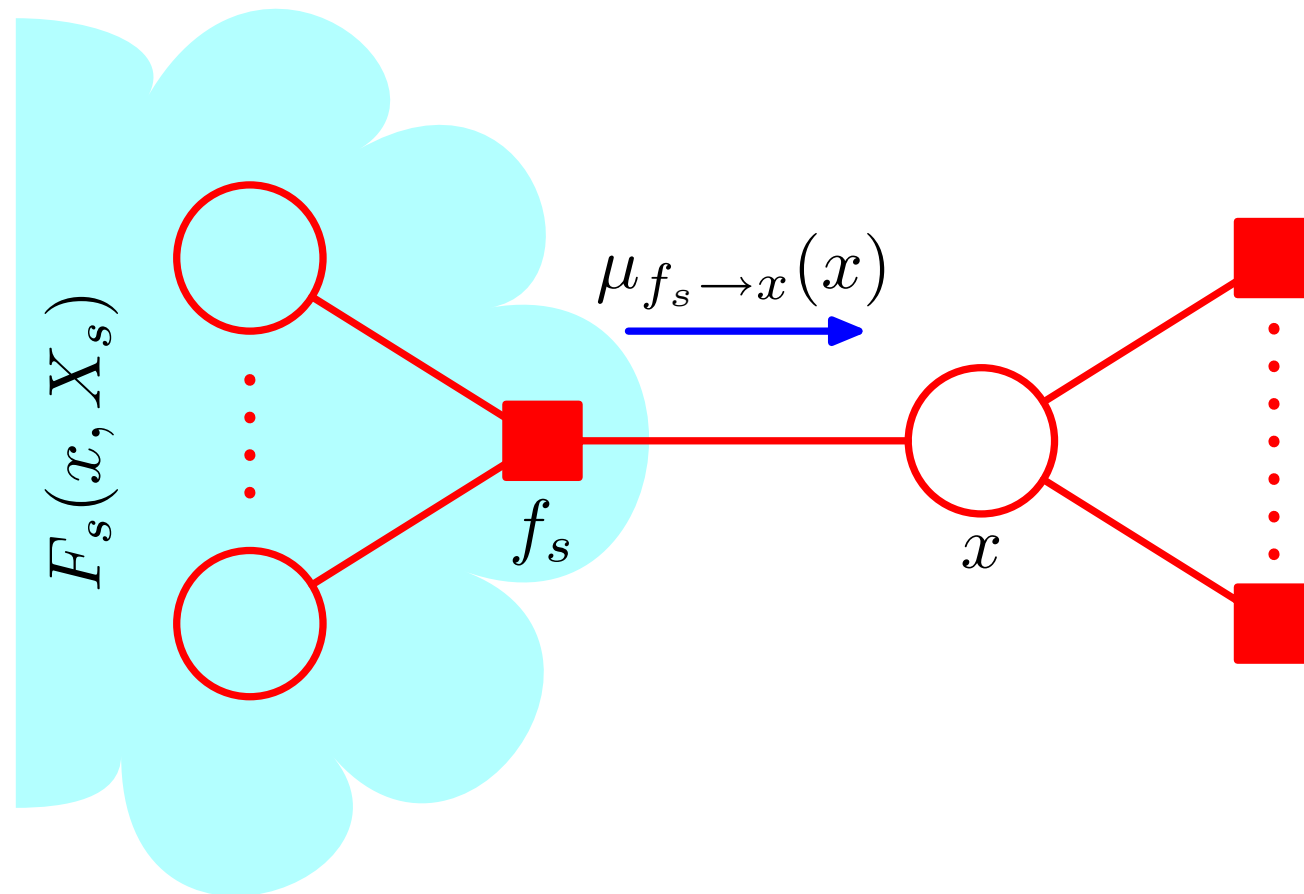
$$\begin{aligned}
 p(x) &= \prod_{s \in \text{ne}(x)} \left[ \sum_{X_s} F_s(x, X_s) \right] \\
 &= \prod_{s \in \text{ne}(x)} \mu_{f_s \rightarrow x}(x).
 \end{aligned}$$

**ne(x)**: set of factor nodes that are neighbours of  $x$

**X<sub>s</sub>**: set of all variables in the subtree connected to the variable node  $x$  via factor node  $f_s$

**F<sub>s</sub>(x, X<sub>s</sub>)**: the product of all the factors in the group associated with factor  $f_s$

## Problem setup: notation



$$\begin{aligned} p(x) &= \prod_{s \in \text{ne}(x)} \left[ \sum_{X_s} F_s(x, X_s) \right] \\ &= \prod_{s \in \text{ne}(x)} \mu_{f_s \rightarrow x}(x). \end{aligned}$$

**We evaluate the marginal  $p(x)$  as product of messages from surrounding factors!**

## Factor messages: decomposition

$$\begin{aligned} p(x) &= \prod_{s \in \text{ne}(x)} \left[ \sum_{X_s} F_s(x, X_s) \right] \\ &= \prod_{s \in \text{ne}(x)} \boxed{\mu_{f_s \rightarrow x}(x)}. \end{aligned}$$

Each factor is itself described by a factor sub-graph, so we can decompose:

$$F_s(x, X_s) = f_s(x, x_1, \dots, x_M) G_1(x_1, X_{s1}) \dots G_M(x_M, X_{sM})$$

**(Each variable associated with  $f_s$  is  $\{x, x_1, \dots, x_M\}$ )**

Rewriting the factor-to-variable message:

$$\boxed{\mu_{f_s \rightarrow x}(x)} = \sum_{x_1} \dots \sum_{x_M} f_s(x, x_1, \dots, x_M) \prod_{m \in \text{ne}(f_s) \setminus x} \left[ \sum_{X_{xm}} G_m(x_m, X_{sm}) \right]$$

## Factor messages: decomposition

$$\begin{aligned} p(x) &= \prod_{s \in \text{ne}(x)} \left[ \sum_{X_s} F_s(x, X_s) \right] \\ &= \prod_{s \in \text{ne}(x)} \mu_{f_s \rightarrow x}(x). \end{aligned}$$

Each factor is itself described by a factor sub-graph, so we can decompose:

$$F_s(x, X_s) = f_s(x, x_1, \dots, x_M) G_1(x_1, X_{s1}) \dots G_M(x_M, X_{sM})$$

**(Each variable associated with  $f_s$  is  $\{x, x_1, \dots, x_M\}$ )**

Rewriting the factor-to-variable message:

$$\begin{aligned} \mu_{f_s \rightarrow x}(x) &= \sum_{x_1} \dots \sum_{x_M} f_s(x, x_1, \dots, x_M) \prod_{m \in \text{ne}(f_s) \setminus x} \left[ \sum_{X_{xm}} G_m(x_m, X_{sm}) \right] \\ &= \sum_{x_1} \dots \sum_{x_M} f_s(x, x_1, \dots, x_M) \prod_{m \in \text{ne}(f_s) \setminus x} \boxed{\mu_{x_m \rightarrow f_s}(x_m)} \quad (8.66) \end{aligned}$$

## Factor-to-variable messages: decomposition

$$\begin{aligned} p(x) &= \prod_{s \in \text{ne}(x)} \left[ \sum_{X_s} F_s(x, X_s) \right] \\ &= \prod_{s \in \text{ne}(x)} \mu_{f_s \rightarrow x}(x). \end{aligned}$$

Each factor is itself described by a factor sub-graph, so we can decompose:

$$F_s(x, X_s) = f_s(x, x_1, \dots, x_M) G_1(x_1, X_{s1}) \dots G_M(x_M, X_{sM})$$

(Each variable associated with  $f_s$  is  $\{x, x_1, \dots, x_M\}$ )

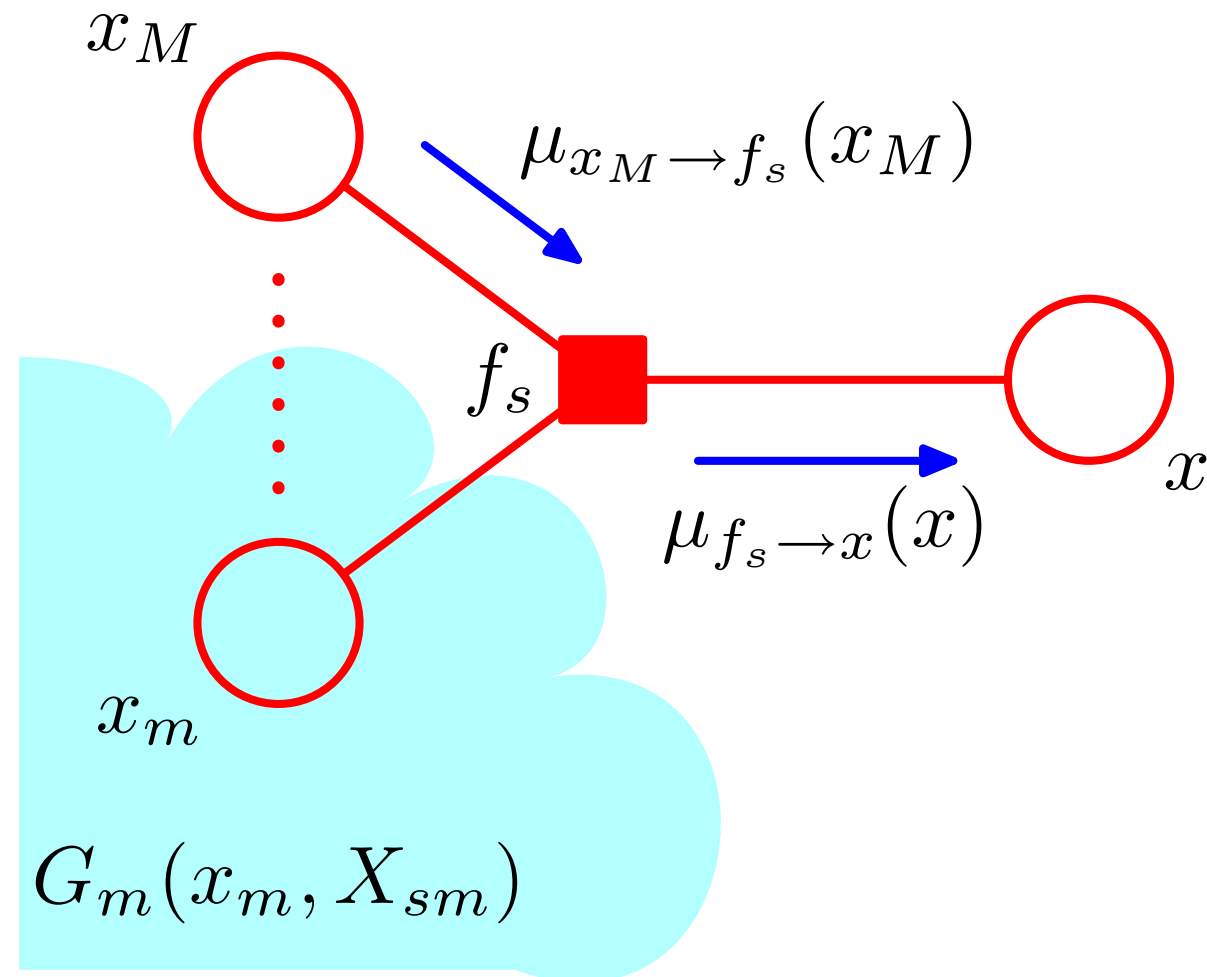
Rewriting the factor-to-variable message:

$$\begin{aligned} \mu_{f_s \rightarrow x}(x) &= \sum_{x_1} \dots \sum_{x_M} f_s(x, x_1, \dots, x_M) \prod_{m \in \text{ne}(f_s) \setminus x} \left[ \sum_{X_{sm}} G_m(x_m, X_{sm}) \right] \\ &= \sum_{x_1} \dots \sum_{x_M} f_s(x, x_1, \dots, x_M) \prod_{m \in \text{ne}(f_s) \setminus x} \boxed{\mu_{x_m \rightarrow f_s}(x_m)} \quad (8.66) \end{aligned}$$

$$\mu_{x_m \rightarrow f_s}(x_m) \equiv \sum_{X_{sm}} G_m(x_m, X_{sm})$$

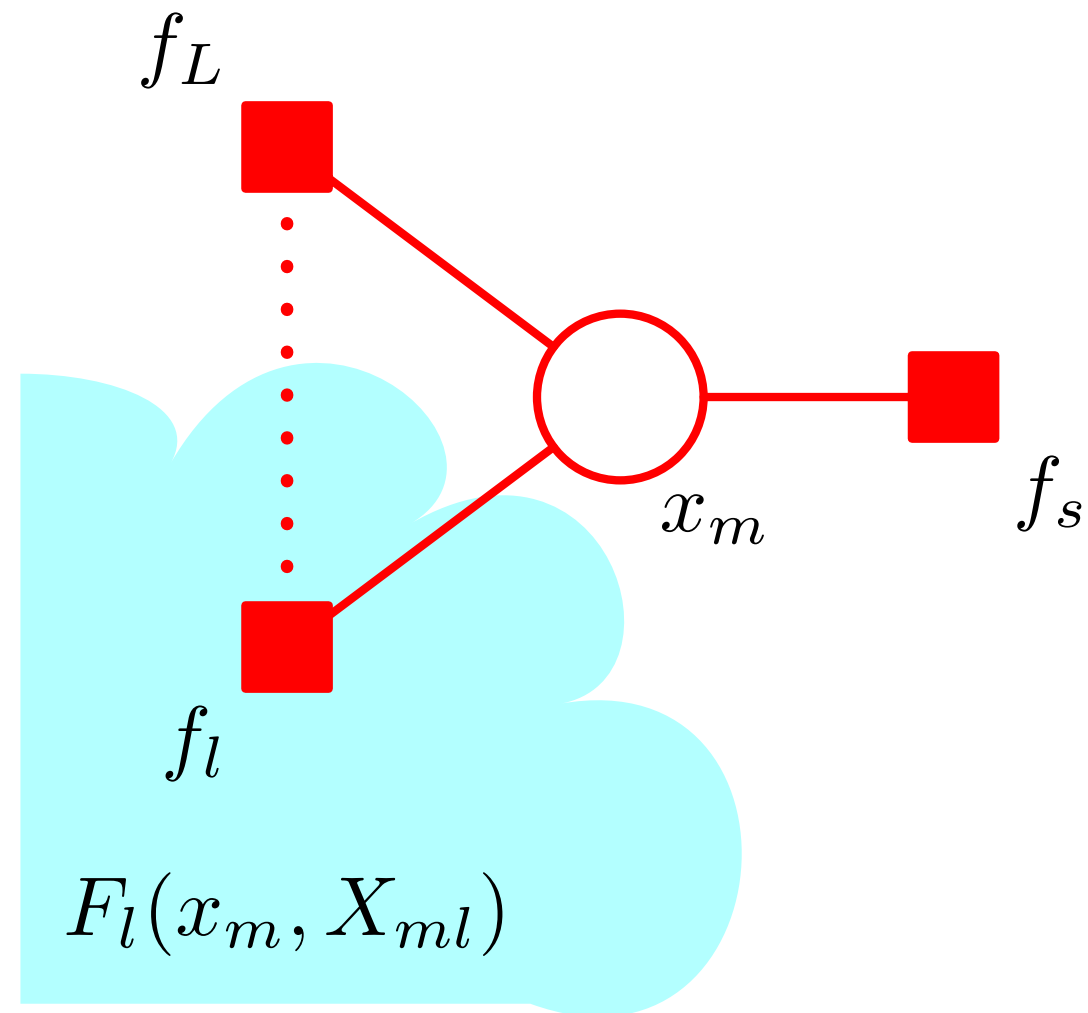


## Variable-to-factor messages: decomposition



$$\mu_{x_m \rightarrow f_s}(x_m) \equiv \sum_{X_{sm}} G_m(x_m, X_{sm})$$

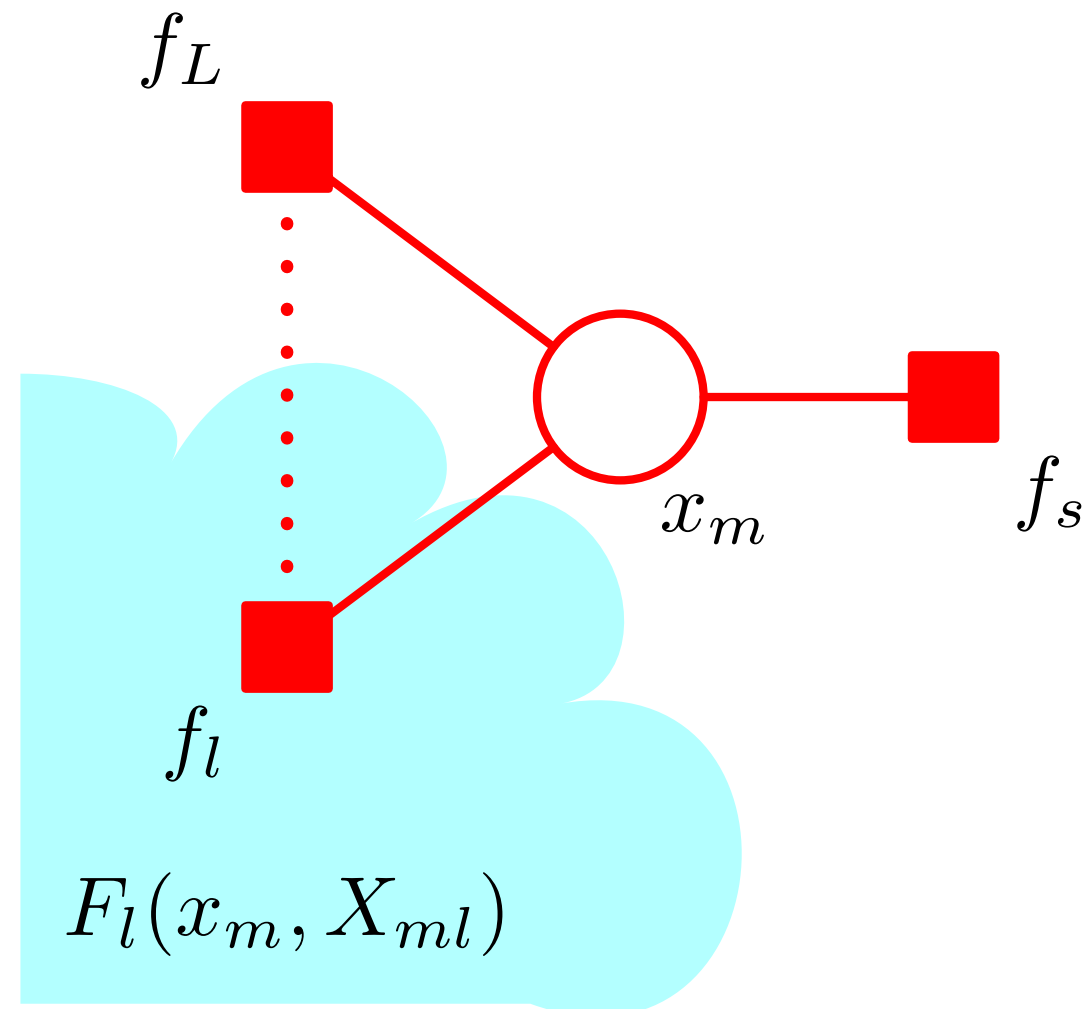
## Factor-to-variable messages: one step back towards the leaves



$$F_s(x, X_s) = f_s(x, x_1, \dots, x_M) G_1(x_1, X_{s1}) \dots G_M(x_M, X_{sM})$$

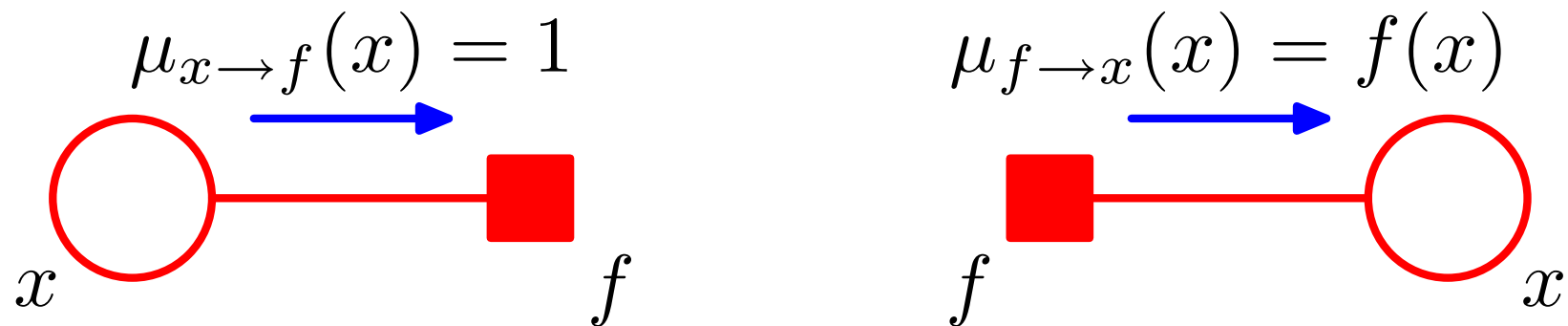
$$G_m(x_m, X_{sm}) = \prod_{l \in \text{ne}(x_m) \setminus f_s} F_l(x_m, X_{ml})$$

## Factor-to-variable messages: one step back towards the leaves



$$\begin{aligned}\mu_{x_m \rightarrow f_s}(x_m) &= \prod_{l \in \text{ne}(x_m) \setminus f_s} \left[ \sum_{X_{ml}} F_l(x_m, X_{ml}) \right] \\ &= \prod_{l \in \text{ne}(x_m) \setminus f_s} \mu_{f_l \rightarrow x_m}(x_m)\end{aligned}$$

## Sum-Product Initialization at Leaves



## Sum-Product: Marginal distribution over $x$

$$\begin{aligned} p(x) &= \prod_{s \in \text{ne}(x)} \left[ \sum_{X_s} F_s(x, X_s) \right] \\ &= \prod_{s \in \text{ne}(x)} \mu_{f_s \rightarrow x}(x). \end{aligned}$$

**See Bishop p. 409 for a fully worked, simple example!**