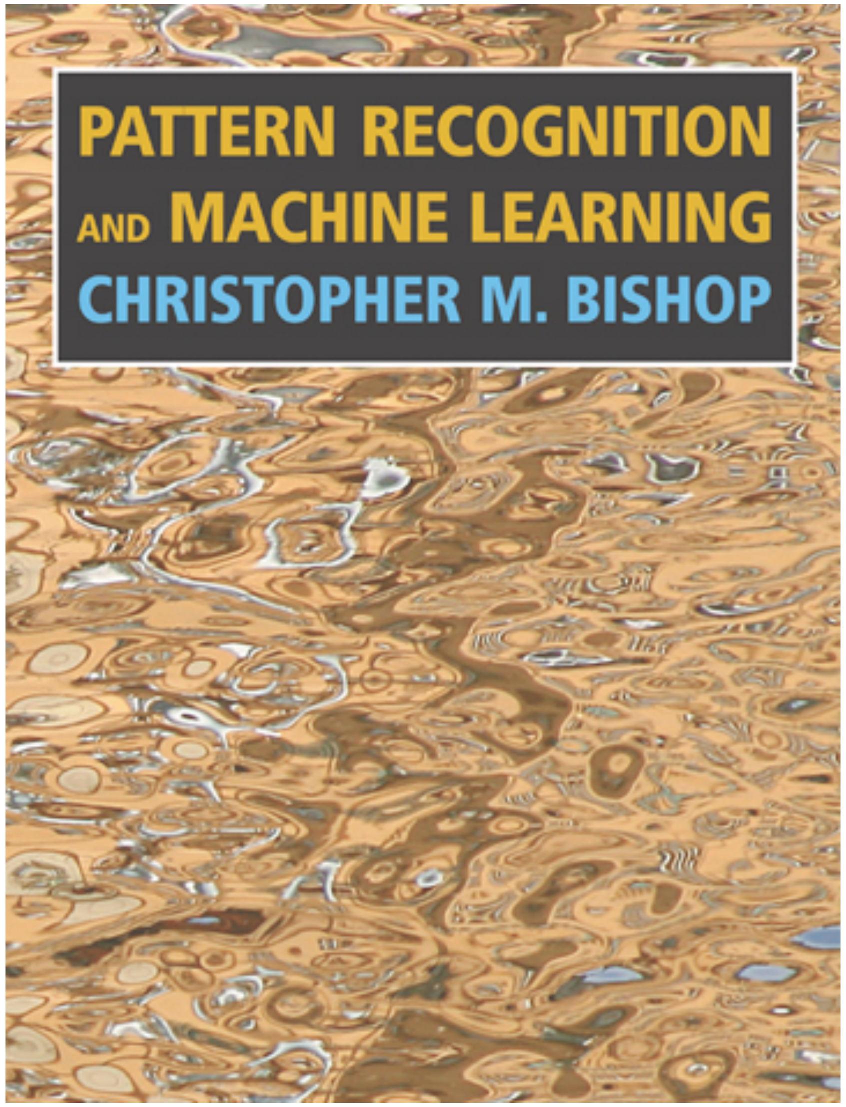
15-381
ARTIFICIAL INTELLIGENCE
LECTURE 9: INFERENCE ON BAYESIAN NETWORKS

Fall 2010

PATTERN RECOGNITION AND MACHINE LEARNING



http://research.microsoft.com/en-us/um/people/cmbishop/prml/

REVIEW: PROBABILITY

Marginal Probability

Joint Probability

SUM RULE

$$p(X) = \sum_Y p(X,Y)$$

PRODUCT RULE

$$p(X,Y) = p(Y|X)p(X)$$

Conditional Probability

Likelihood

Prior Probability

BAYES' THEOREM

$$p(X|Y) = \frac{p(Y|X)p(X)}{p(Y)}$$

Posterior Probability

Evidence

REVIEW: BAYES NETS

$$p(\mathbf{x}) = \prod_{k=1}^{K} p(x_k|\mathrm{pa}_k)$$

REVIEW: BAYES NETS

- GRAPHICAL MODEL

$$p(\mathbf{x}) = \prod_{k=1}^{K} p(x_k|\mathrm{pa}_k)$$

REVIEW: BAYES NETS

- GRAPHICAL MODEL
  - EACH NODE DENOTES A RANDOM VARIABLE

$$p(\mathbf{x}) = \prod_{k=1}^{K} p(x_k|\mathrm{pa}_k)$$

REVIEW: BAYES NETS

- GRAPHICAL MODEL
  - EACH NODE DENOTES A RANDOM VARIABLE
  - DIRECTED ACYCLIC GRAPH

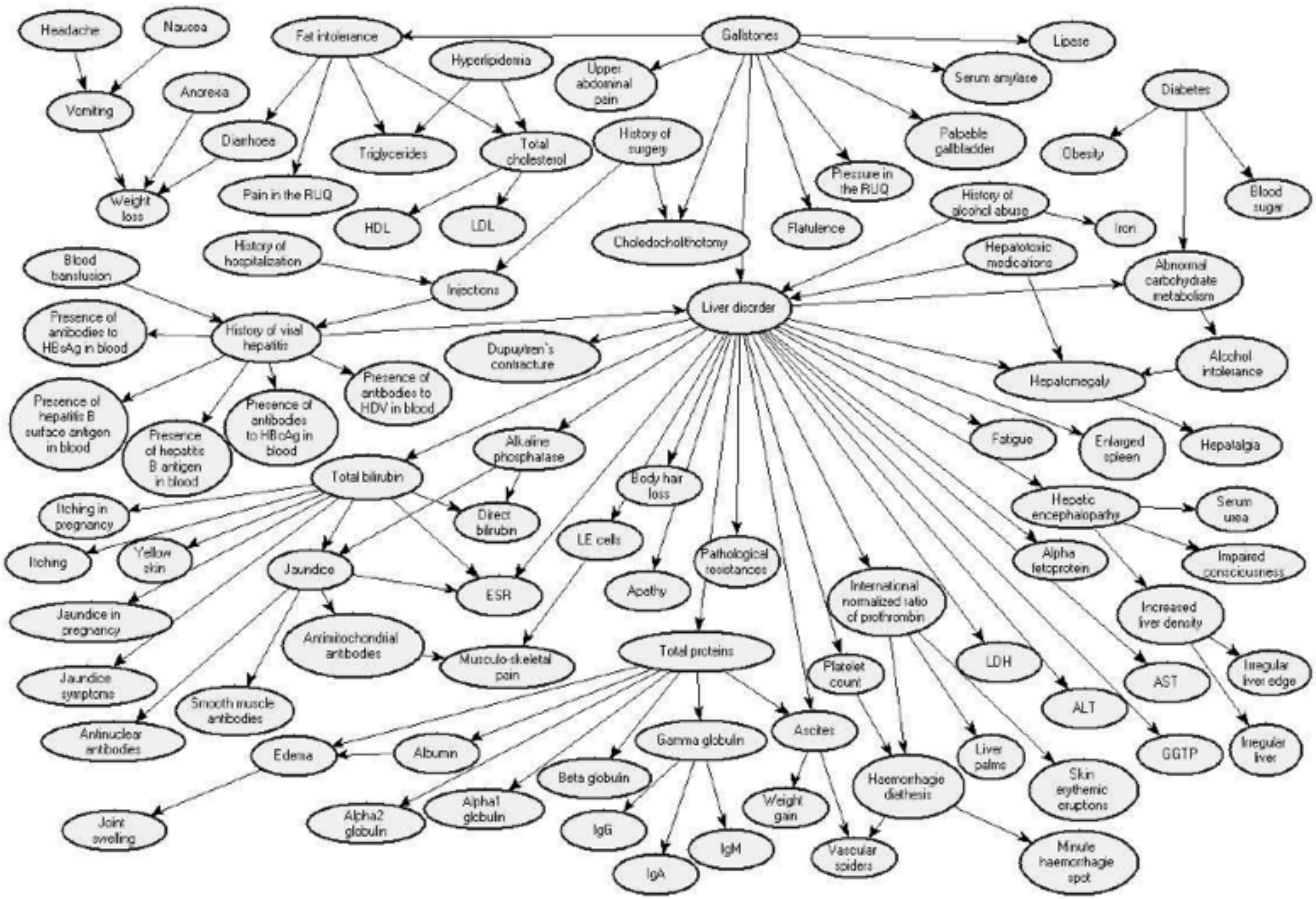
$$p(\mathbf{x}) = \prod_{k=1}^{K} p(x_k|\mathrm{pa}_k)$$

REVIEW: BAYES NETS

- GRAPHICAL MODEL
  - EACH NODE DENOTES A RANDOM VARIABLE
  - DIRECTED ACYCLIC GRAPH
  - CONDITIONAL PROBABILITY TABLES

$$p(\mathbf{x}) = \prod_{k=1}^{K} p(x_k|\mathrm{pa}_k)$$

Diagnosis of Liver Disorders



Agnieszka Onisko, Marek J. Druzdzel and Hanna Wasyluk. [A Bayesian network model for diagnosis of liver disorders](). In *Proceedings of the Eleventh Conference on Biocybernetics and Biomedical Engineering*, pages 842-846, Warsaw, Poland, December 2-4, 1999.

CONDITIONAL INDEPENDENCE

$$p(a|b, c) = p(a|c)$$

CONDITIONAL INDEPENDENCE

$$p(a|b,c) = p(a|c)$$

$$\begin{aligned} p(a,b|c) &= p(a|b,c)p(b|c) \\ &= p(a|c)p(b|c) \end{aligned}$$

CONDITIONAL INDEPENDENCE

$$p(a|b,c) = p(a|c)$$

$$\begin{aligned} p(a,b|c) &= p(a|b,c)p(b|c) \\ &= p(a|c)p(b|c) \end{aligned}$$

$$a \perp b|c$$

CONDITIONAL INDEPENDENCE

$$p(a|b,c) = p(a|c)$$

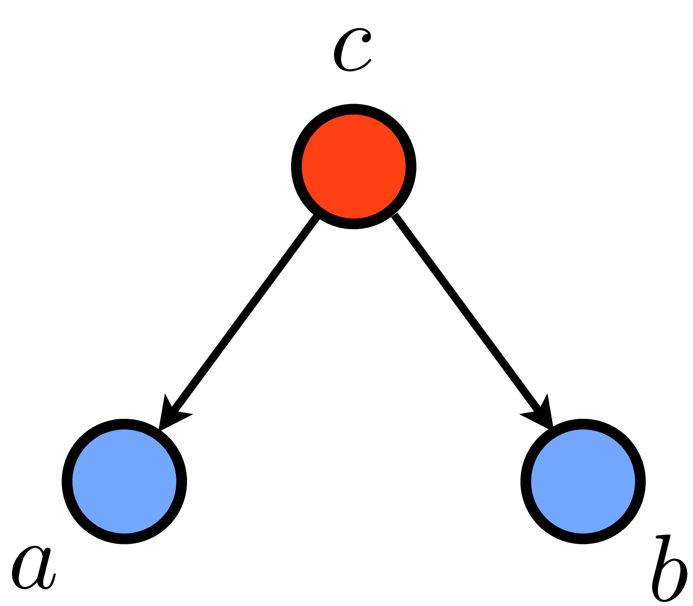
$$\begin{aligned} p(a,b|c) &= p(a|b,c)p(b|c) \\ &= p(a|c)p(b|c) \end{aligned}$$

$$a \perp b|c$$

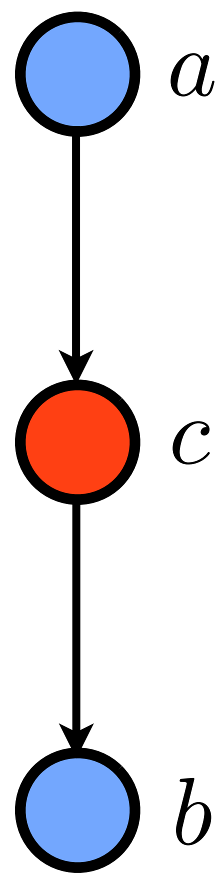
"*a* and *b* are independent given *c*"

THREE GRAPHS

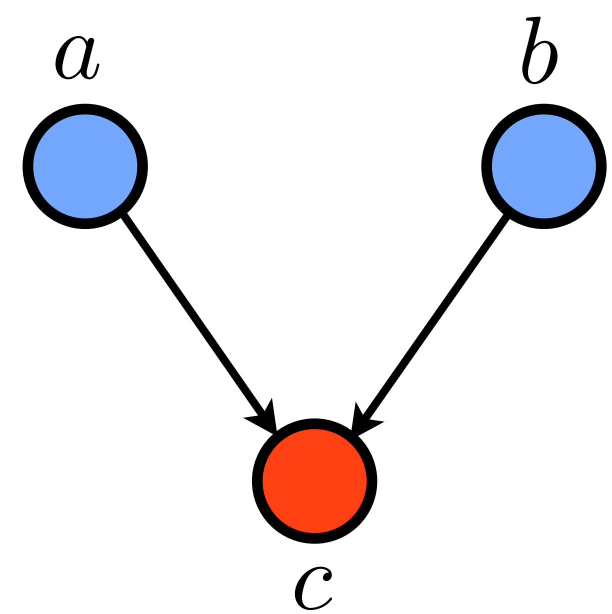
COMMON CAUSE



CAUSAL CHAIN



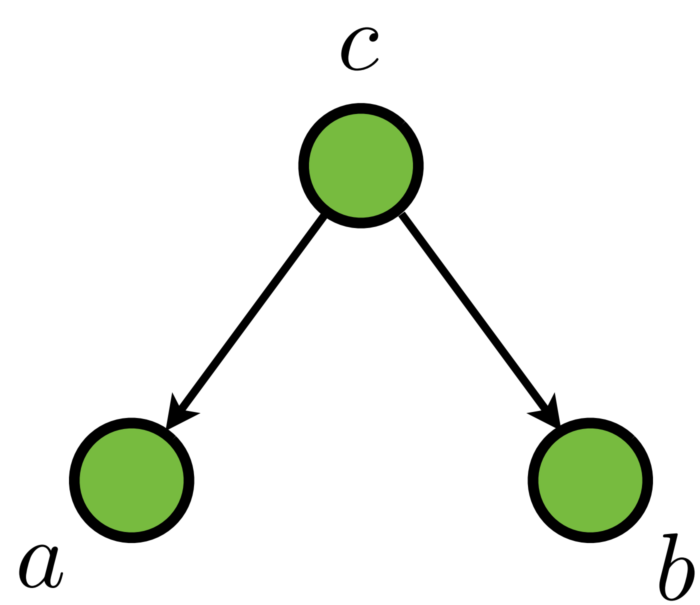
COMMON EFFECT



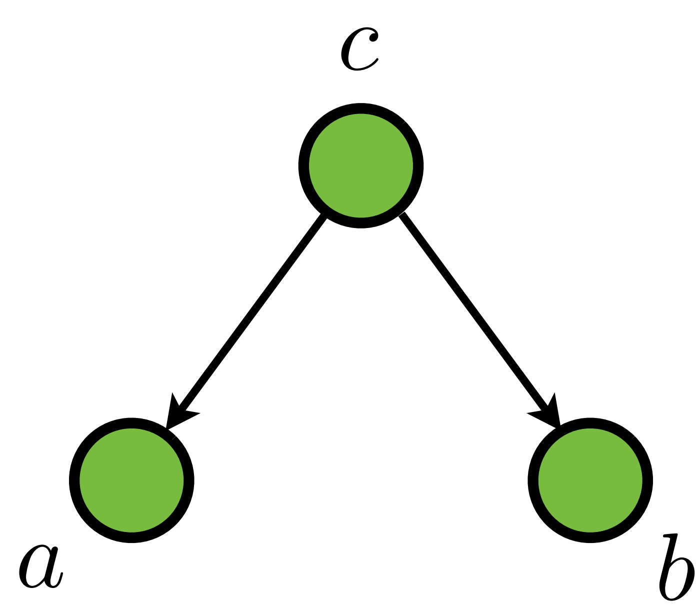
$$a \perp b | \emptyset \;\Leftrightarrow\; p(a, b) = p(a)p(b)$$

$$a \perp b | c \;\Leftrightarrow\; p(a, b|c) = p(a|c)p(b|c)$$

COMMON CAUSE

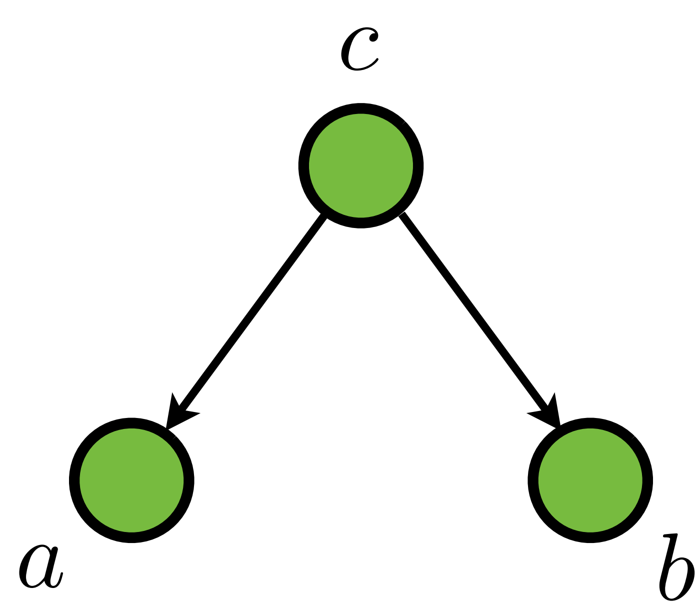


COMMON CAUSE



$$p(a, b, c) = p(a|c)p(b|c)p(c)$$

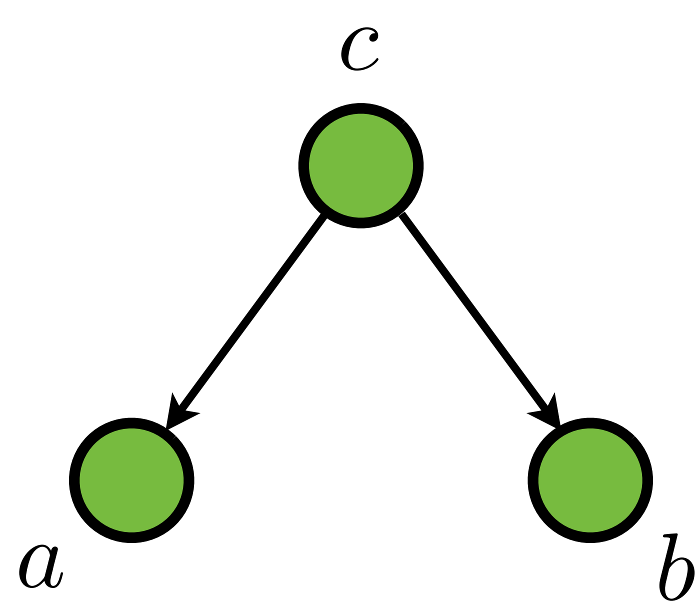
COMMON CAUSE



$$p(a, b, c) = p(a|c)p(b|c)p(c)$$

$$p(a, b)$$

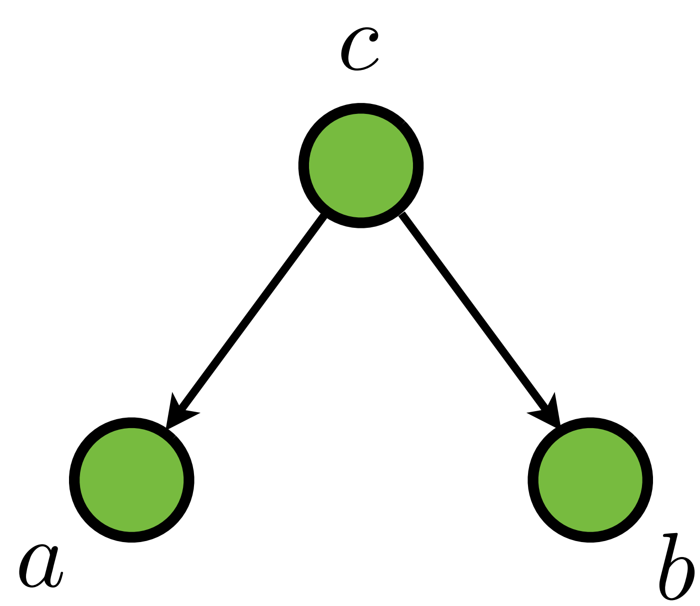
COMMON CAUSE



$$p(a,b,c) = p(a|c)p(b|c)p(c)$$

$$p(a,b) = \sum_c p(a|c)p(b|c)p(c)$$

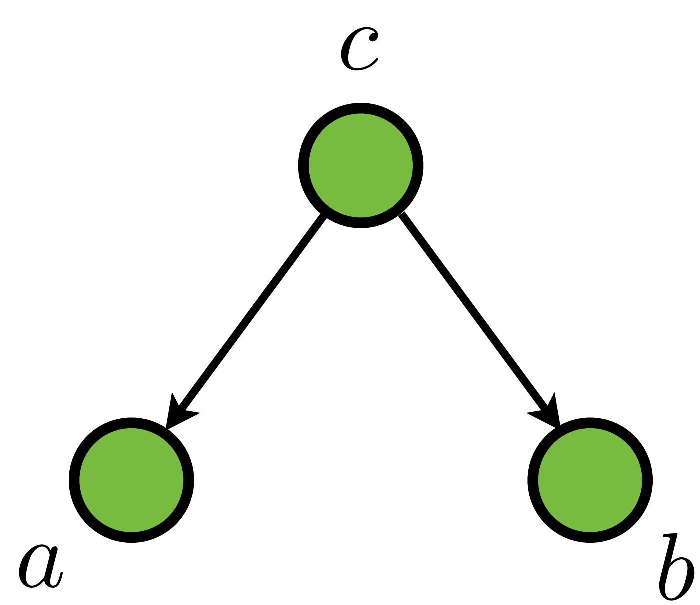
COMMON CAUSE



$$p(a,b,c) = p(a|c)p(b|c)p(c)$$

$$\begin{aligned} p(a,b) &= \sum_c p(a|c)p(b|c)p(c) \\ &\neq p(a)p(b) \end{aligned}$$

COMMON CAUSE



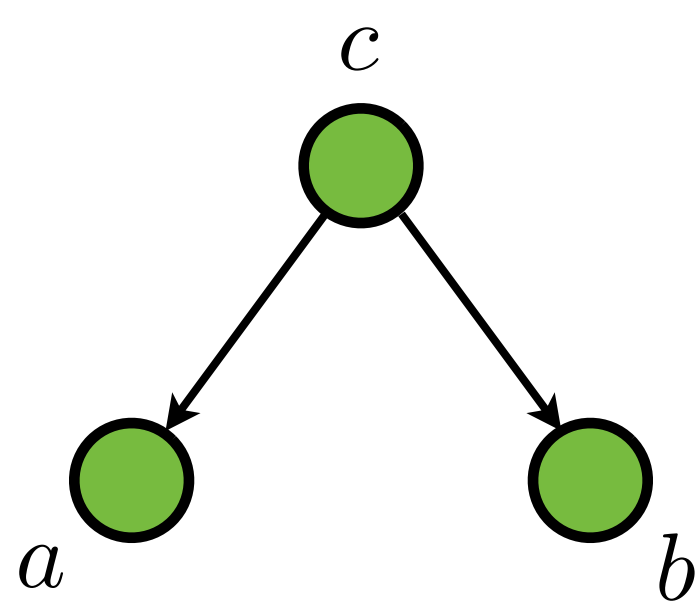
$$p(a, b, c) = p(a|c)p(b|c)p(c)$$

$$p(a, b) = \sum_c p(a|c)p(b|c)p(c)$$

$$\neq p(a)p(b)$$

$$a \not\perp b|\emptyset$$

COMMON CAUSE



$$p(a,b,c) = p(a|c)p(b|c)p(c)$$

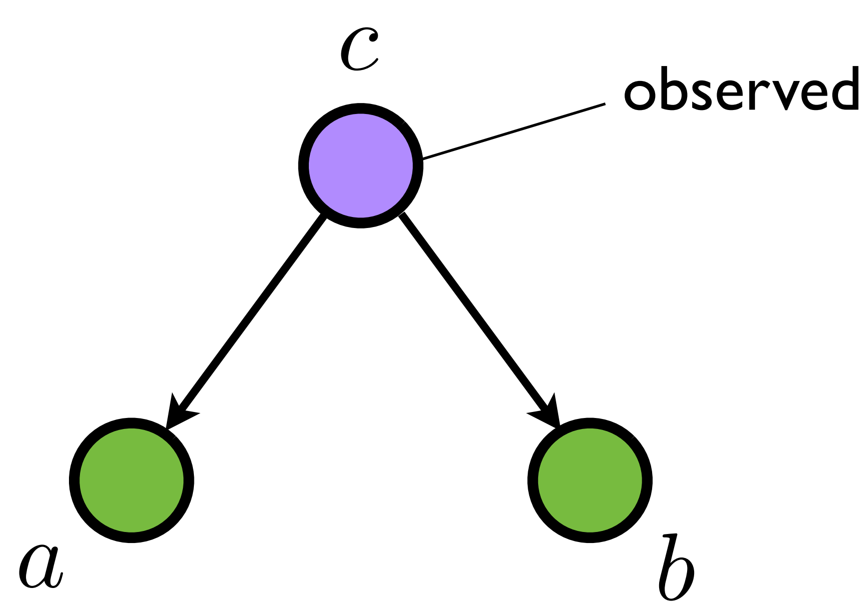
$$p(a,b) = \sum_c p(a|c)p(b|c)p(c)$$

$$\neq p(a)p(b)$$

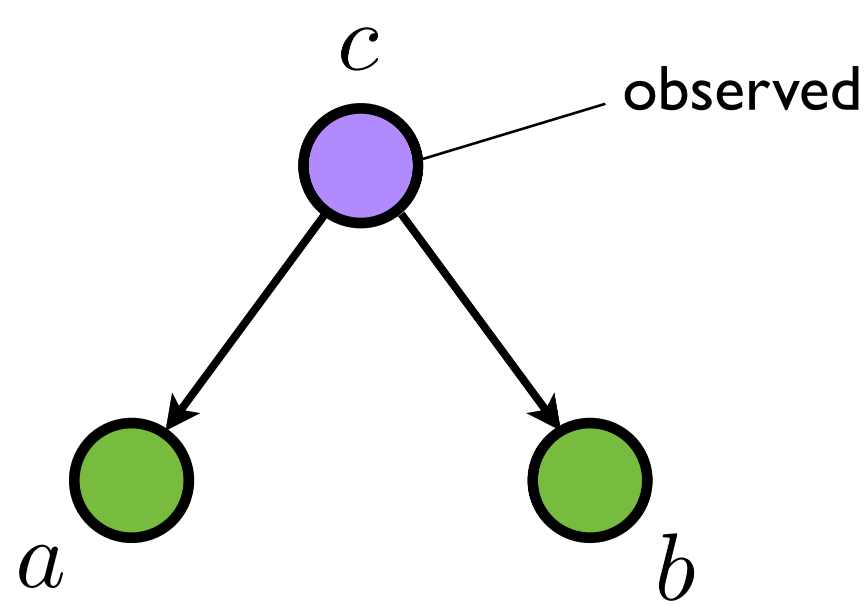
$$a \not\perp b | \emptyset$$

a **and** ***b*** **are <u>not</u> independent**

COMMON CAUSE

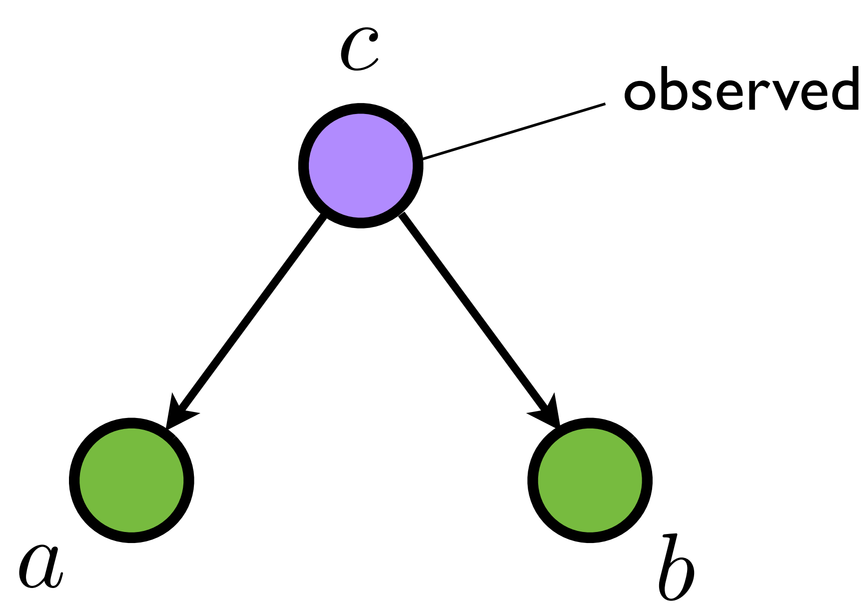


COMMON CAUSE



$$p(a, b|c) = \frac{p(a, b, c)}{p(c)}$$

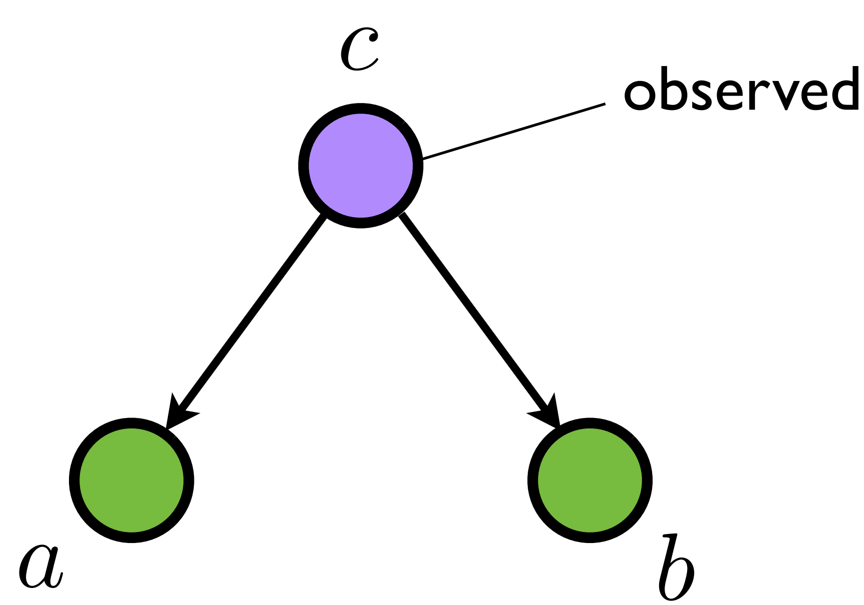
COMMON CAUSE



$$p(a,b|c) = \frac{p(a,b,c)}{p(c)}$$

$$= \frac{p(a|c)p(b|c)p(c)}{p(c)} = p(a|c)p(b|c)$$

COMMON CAUSE

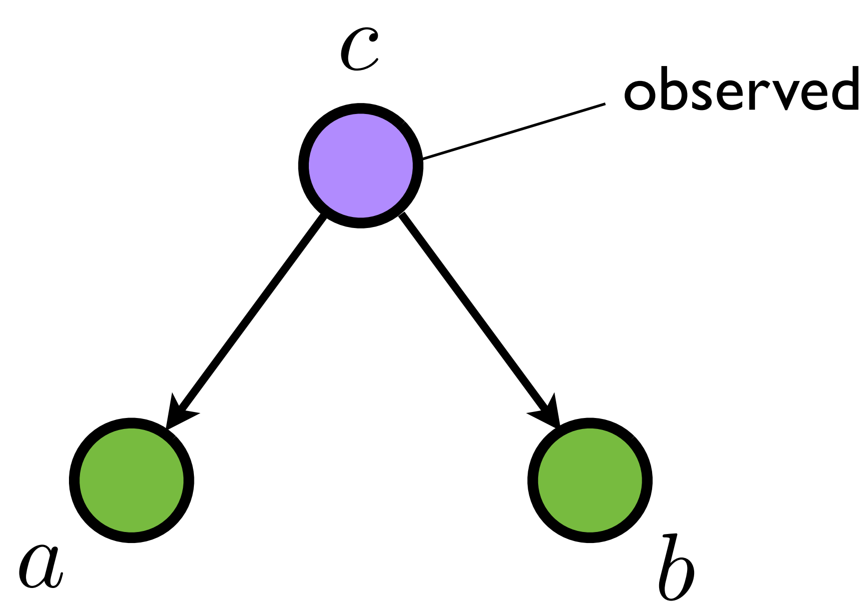


$$p(a,b|c) = \frac{p(a,b,c)}{p(c)}$$

$$= \frac{p(a|c)p(b|c)p(c)}{p(c)} = p(a|c)p(b|c)$$

$$a \perp b|c$$

COMMON CAUSE



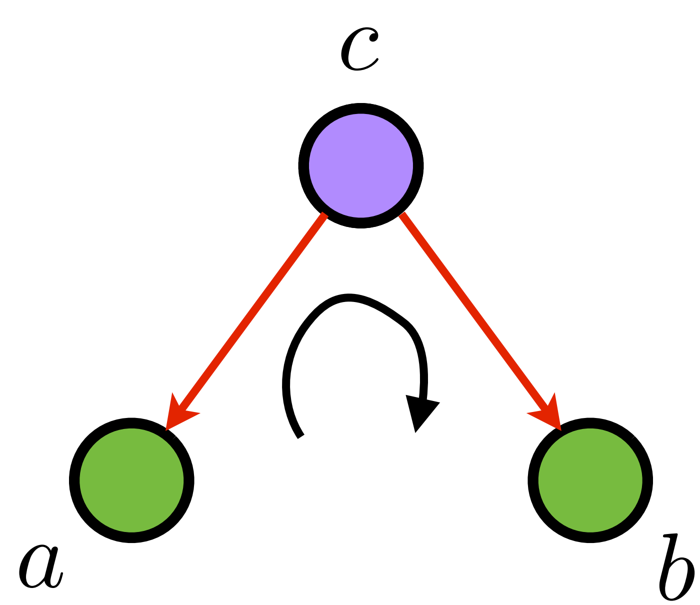
$$p(a, b|c) = \frac{p(a, b, c)}{p(c)}$$

$$= \frac{p(a|c)p(b|c)p(c)}{p(c)} = p(a|c)p(b|c)$$

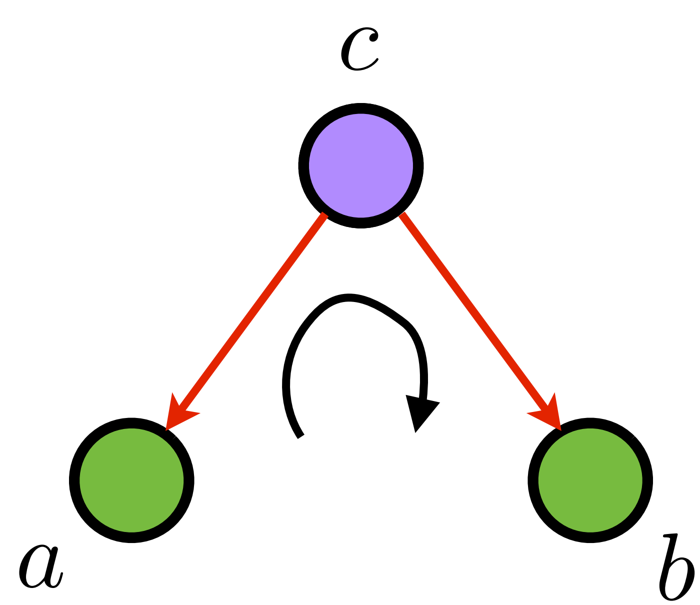
$$a \perp b|c$$

a is conditionally independent of *b* given *c*

TAIL-TO-TAIL

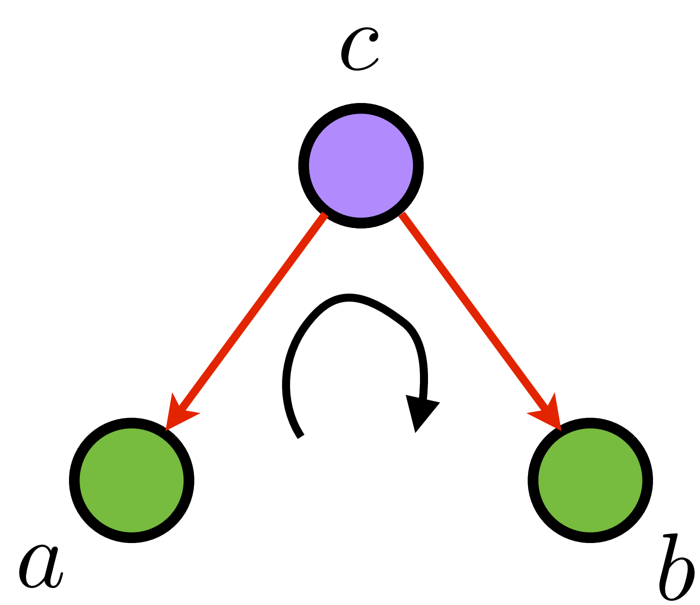


TAIL-TO-TAIL



Node c is *tail-to-tail*

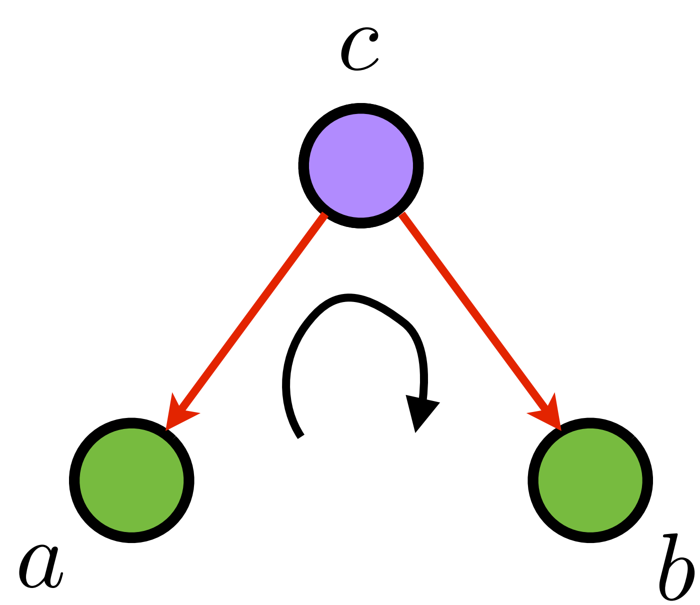
TAIL-TO-TAIL



Node c is *tail-to-tail*

Path exists from a to b through c

TAIL-TO-TAIL

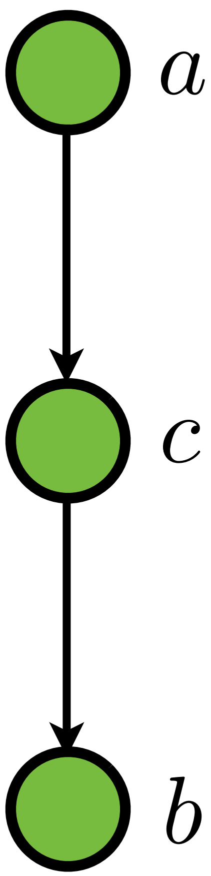


Node *c* is *tail-to-tail*

Path exists from *a* to *b* through *c*

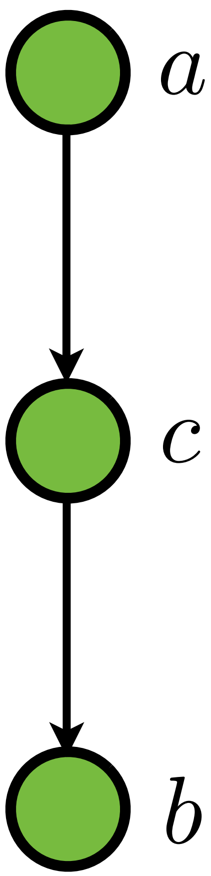
When *c* is observed, it ‘blocks’ the path from *a* to *b*

CAUSAL CHAIN



$$p(a, b, c) = p(a)p(c|a)p(b|c)$$

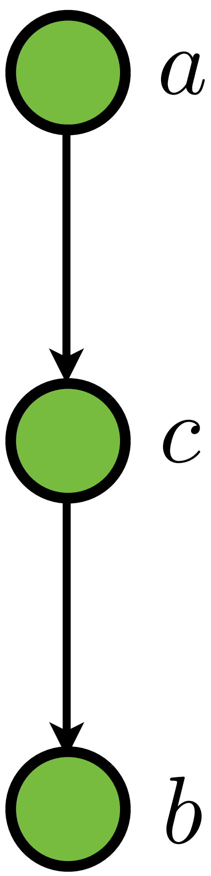
CAUSAL CHAIN



$$p(a, b, c) = p(a)p(c|a)p(b|c)$$

$$p(a, b)$$

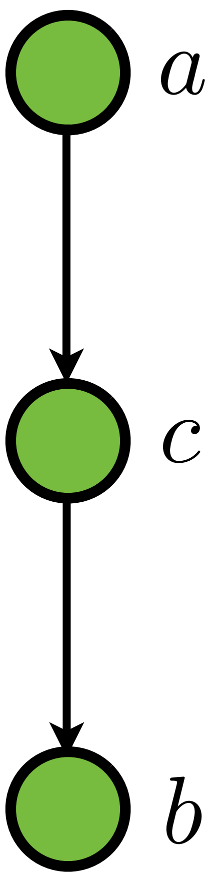
CAUSAL CHAIN



$$p(a,b,c) = p(a)p(c|a)p(b|c)$$

$$p(a,b) = \sum_c p(a)p(c|a)p(b|a) = p(a)\sum_c p(c|a)p(b|a)$$

CAUSAL CHAIN

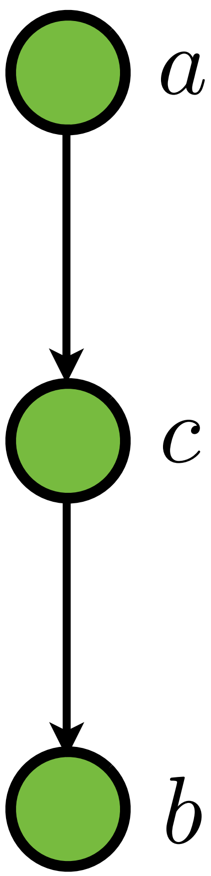


$$p(a,b,c) = p(a)p(c|a)p(b|c)$$

$$p(a,b) = \sum_c p(a)p(c|a)p(b|a) = p(a)\sum_c p(c|a)p(b|a)$$

$$= p(a)p(b|a)$$

CAUSAL CHAIN



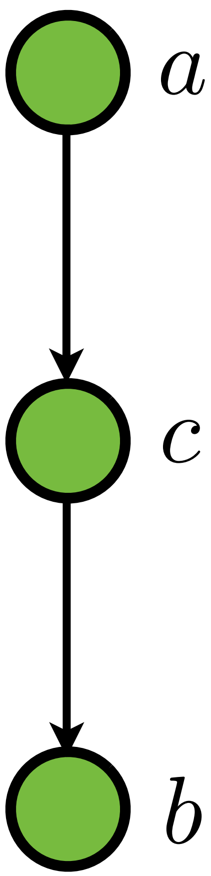
$$p(a,b,c) = p(a)p(c|a)p(b|c)$$

$$p(a,b) = \sum_c p(a)p(c|a)p(b|a) = p(a)\sum_c p(c|a)p(b|a)$$

$$= p(a)p(b|a)$$

$$a \not\perp\!\!\!\perp b|\emptyset$$

CAUSAL CHAIN



$$p(a, b, c) = p(a)p(c|a)p(b|c)$$

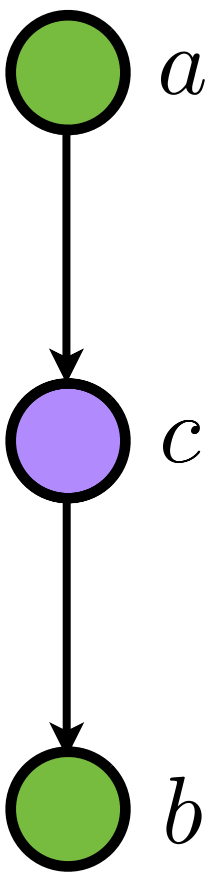
$$p(a, b) = \sum_c p(a)p(c|a)p(b|a) = p(a) \sum_c p(c|a)p(b|a)$$

$$= p(a)p(b|a)$$

$$a \not\perp b | \emptyset$$

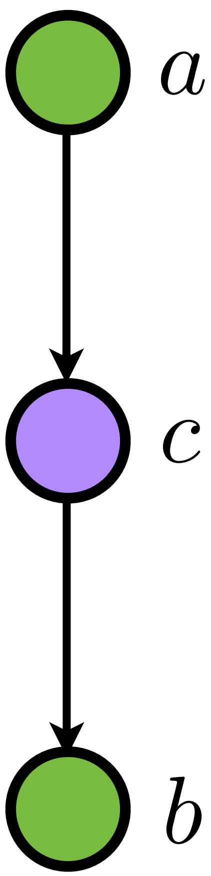
a and *b* are <u>not</u> independent

CAUSAL CHAIN



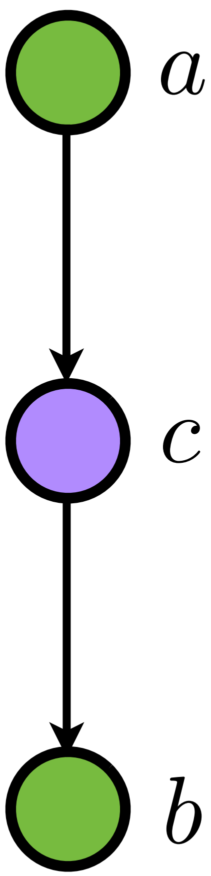
$$p(a, b|c) = \frac{p(a)p(c|a)p(b|c)}{p(c)}$$

CAUSAL CHAIN



$$p(a, b|c) = \frac{p(a, b, c)}{p(c)} = \frac{p(a)p(c|a)p(b|c)}{p(c)}$$

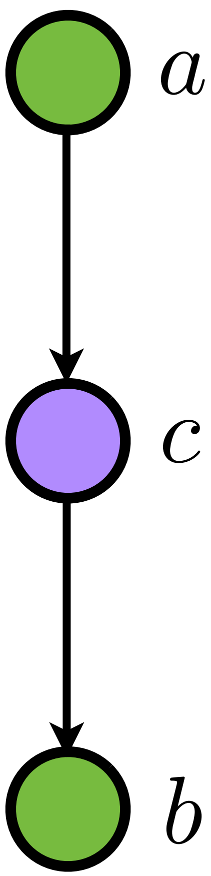
CAUSAL CHAIN



$$p(a,b|c) = \frac{p(a,b,c)}{p(c)} = \frac{p(a)p(c|a)p(b|c)}{p(c)}$$

$$= p(a|c)p(b|c)$$

CAUSAL CHAIN



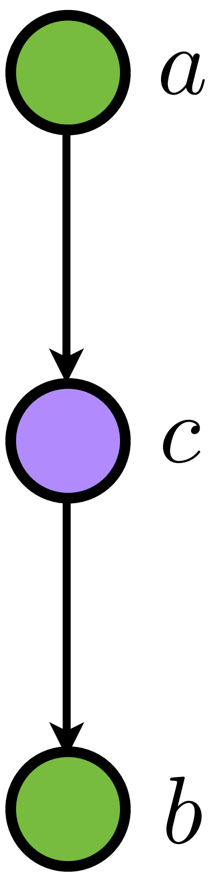
$$p(a,b|c) = \frac{p(a,b,c)}{p(c)} = \frac{p(a)p(c|a)p(b|c)}{p(c)}$$

$$= p(a|c)p(b|c)$$

$$a \perp b|c$$

Thursday, September 23, 2010

CAUSAL CHAIN



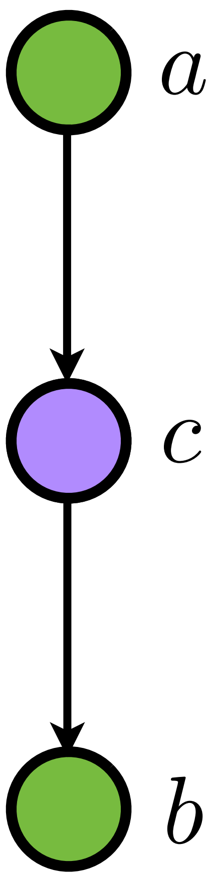
$$p(a,b|c) = \frac{p(a,b,c)}{p(c)} = \frac{p(a)p(c|a)p(b|c)}{p(c)}$$

$$= p(a|c)p(b|c)$$

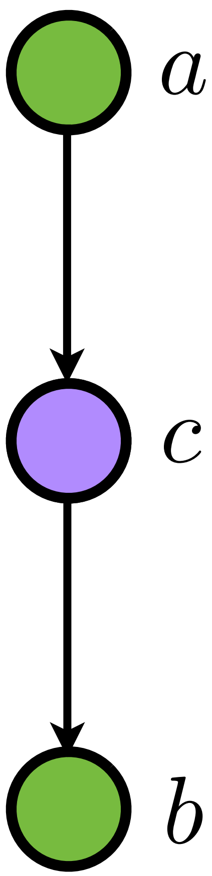
$$a \perp b|c$$

a is conditionally independent of *b* given *c*

HEAD-TO-TAIL

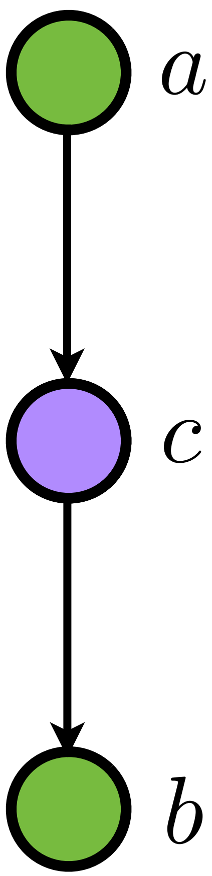


HEAD-TO-TAIL



Node c is *head-to-tail*

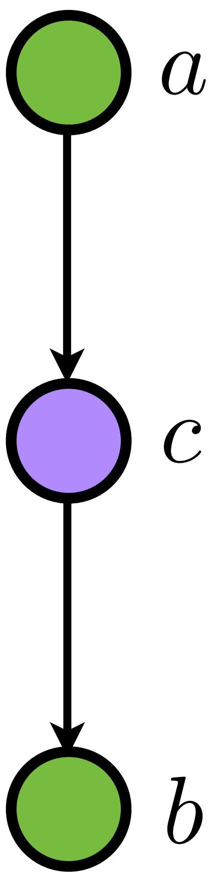
HEAD-TO-TAIL



Node *c* is *head-to-tail*

Path exists from *a* to *b* through *c*

HEAD-TO-TAIL

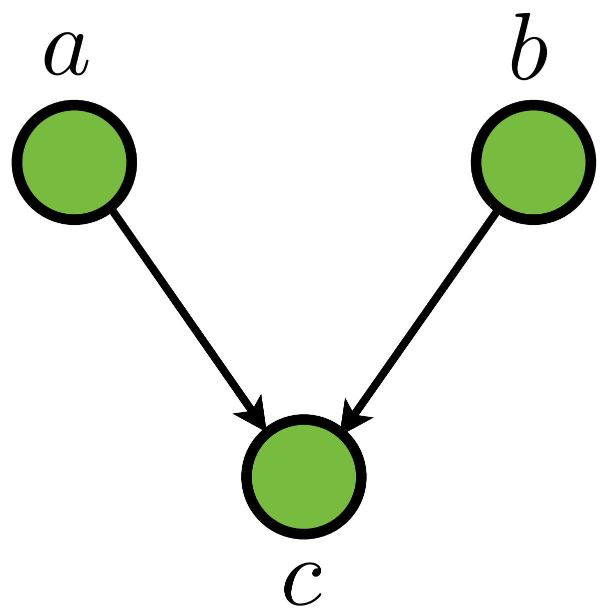


Node *c* is *head-to-tail*

Path exists from *a* to *b* through *c*

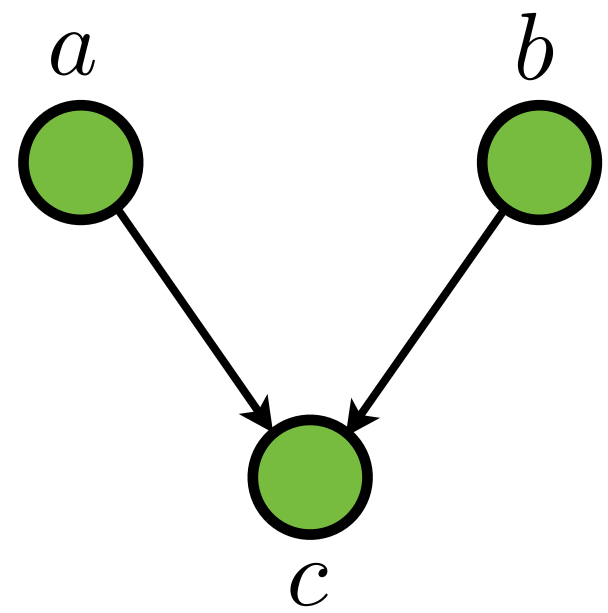
When *c* is observed, it 'blocks' the path from *a* to *b*

COMMON EFFECT



$$p(a, b, c) = p(a)p(b)p(c|a, b)$$

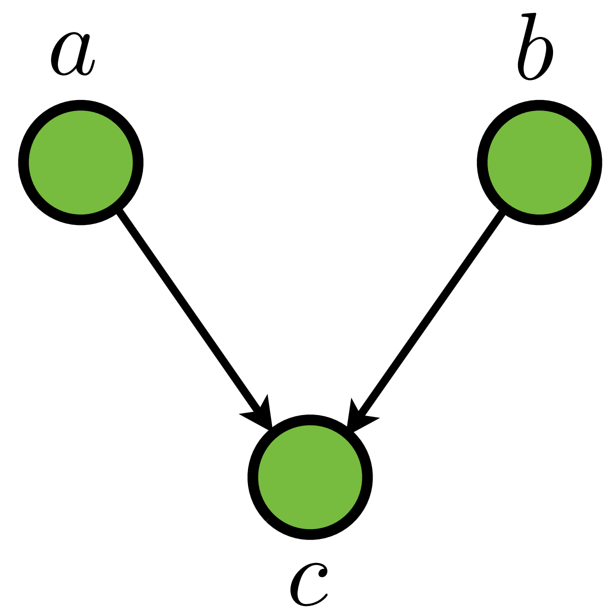
COMMON EFFECT



$$p(a,b,c) = p(a)p(b)p(c|a,b)$$

$$p(a,b) = \sum_c p(a)p(b)p(c|a,b)$$

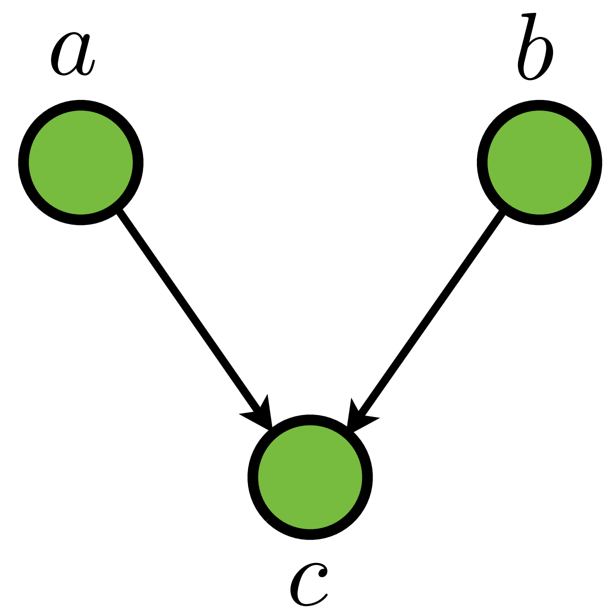
COMMON EFFECT



$$p(a,b,c) = p(a)p(b)p(c|a,b)$$

$$p(a,b) = \sum_c p(a)p(b)p(c|a,b) = p(a)p(b)\sum_c p(c|a,b)$$

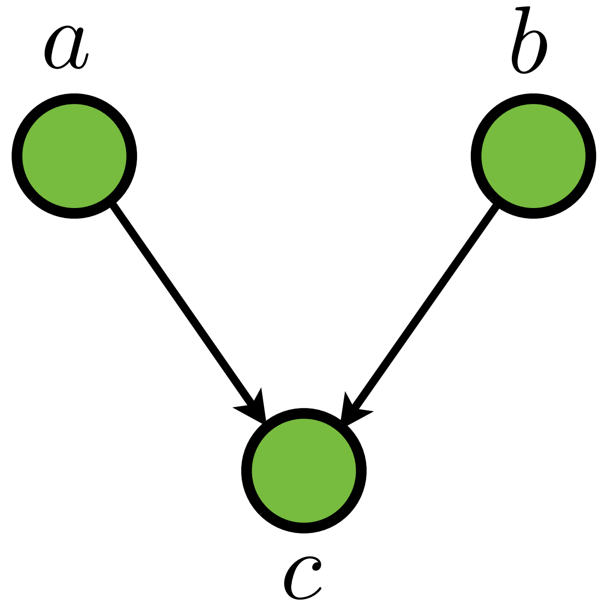
COMMON EFFECT



$$p(a,b,c) = p(a)p(b)p(c|a,b)$$

$$p(a,b) = \sum_c p(a)p(b)p(c|a,b) = p(a)p(b)\sum_c p(c|a,b)$$
$$= p(a)p(b)$$

COMMON EFFECT



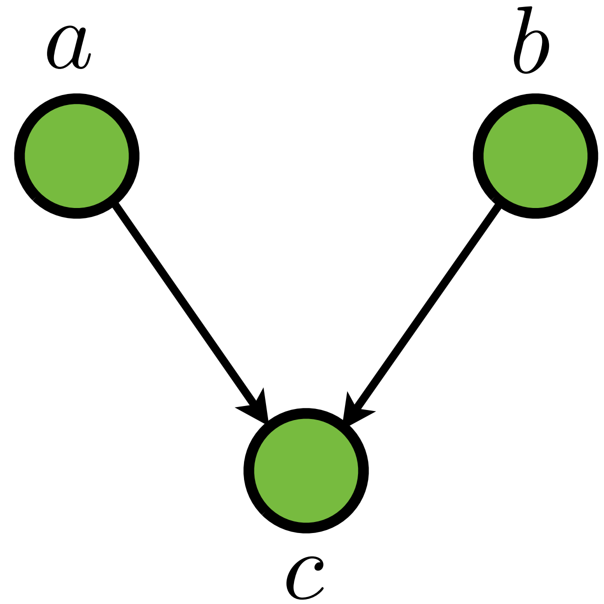
$$p(a, b, c) = p(a)p(b)p(c|a, b)$$

$$p(a, b) = \sum_c p(a)p(b)p(c|a, b) = p(a)p(b) \sum_c p(c|a, b)$$

$$= p(a)p(b)$$

$$a \perp b|\emptyset$$

COMMON EFFECT



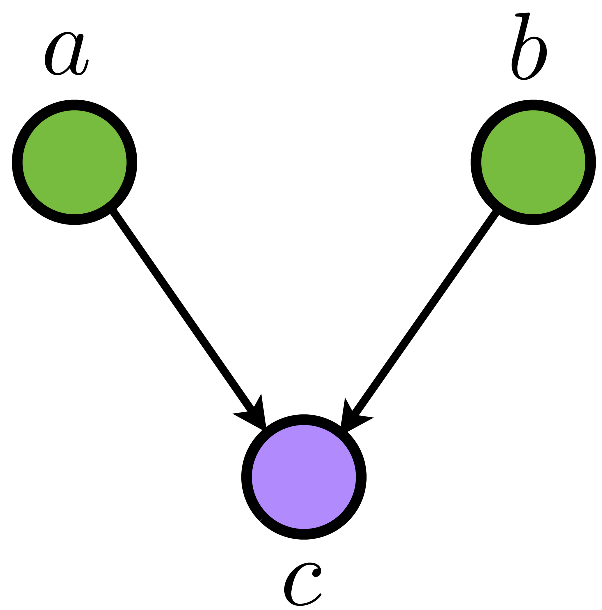
$$p(a,b,c) = p(a)p(b)p(c|a,b)$$

$$p(a,b) = \sum_c p(a)p(b)p(c|a,b) = p(a)p(b)\sum_c p(c|a,b)$$
$$= p(a)p(b)$$

$$a \perp b|\emptyset$$

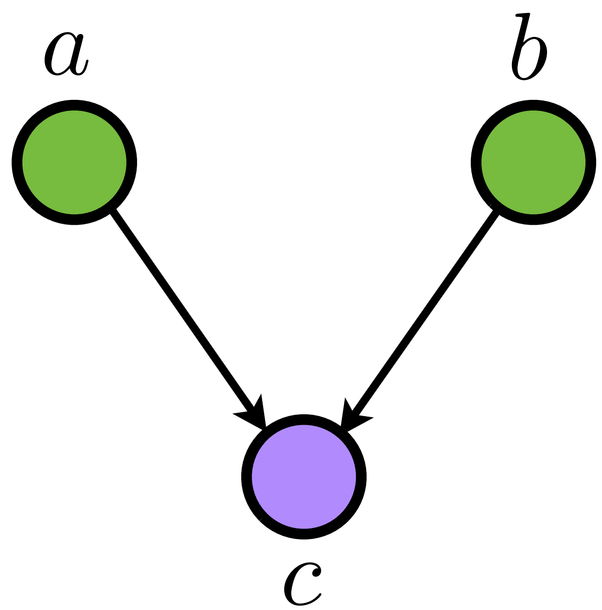
a **and** ***b*** **are independent**

COMMON EFFECT



$$p(a, b|c) = \frac{p(a, b, c)}{p(c)}$$

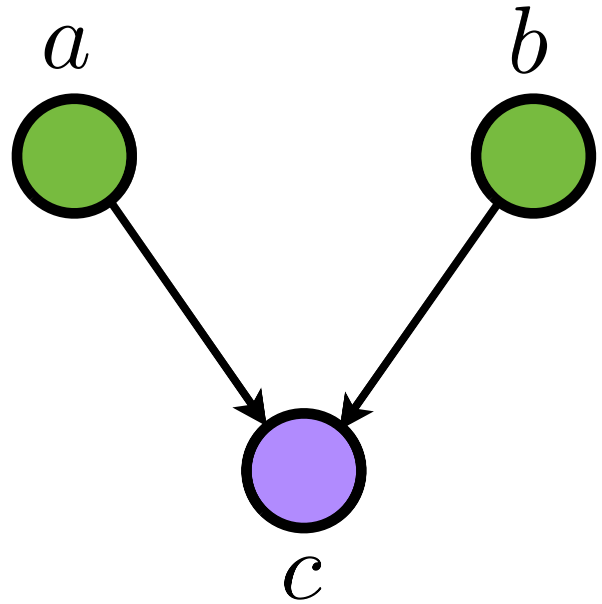
COMMON EFFECT



$$p(a, b|c) = \frac{p(a, b, c)}{p(c)}$$

$$= \frac{p(a)p(b)p(c|a, b)}{p(c)}$$

COMMON EFFECT

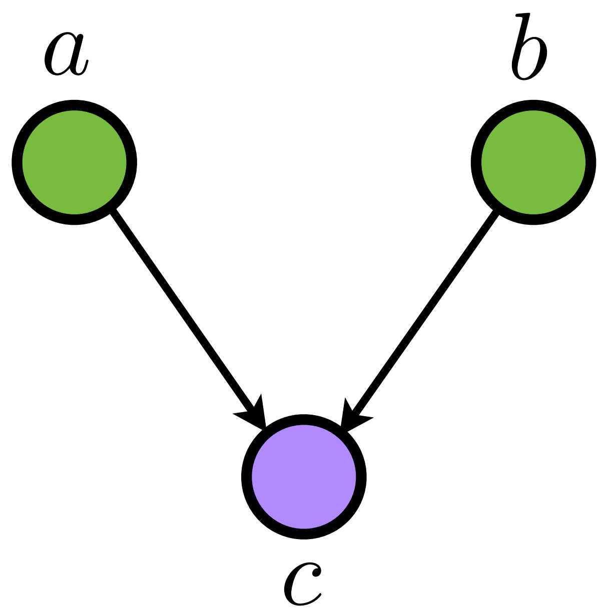


$$p(a, b|c) = \frac{p(a, b, c)}{p(c)}$$

$$= \frac{p(a)p(b)p(c|a, b)}{p(c)}$$

$$\neq p(a|c)p(b|c)$$

COMMON EFFECT



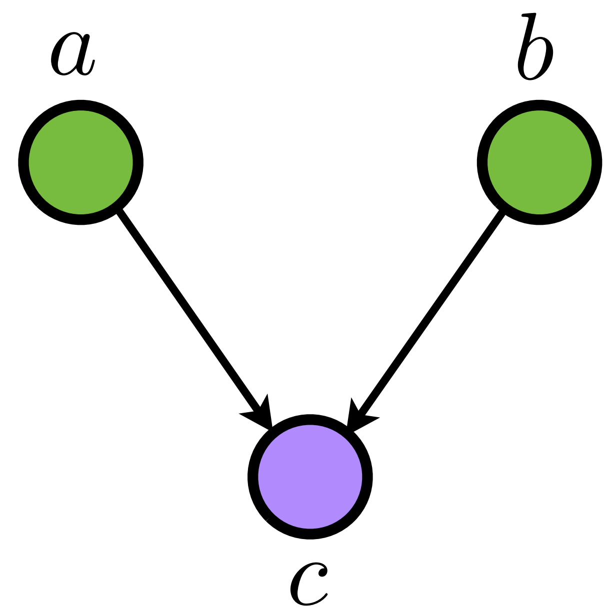
$$p(a,b|c) = \frac{p(a,b,c)}{p(c)}$$

$$= \frac{p(a)p(b)p(c|a,b)}{p(c)}$$

$$\neq p(a|c)p(b|c)$$

$$a \not\perp b|c$$

COMMON EFFECT



$$p(a, b|c) = \frac{p(a, b, c)}{p(c)}$$

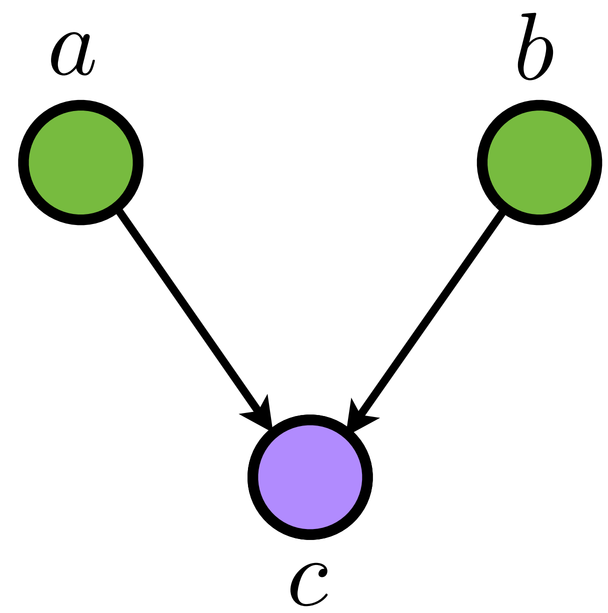
$$= \frac{p(a)p(b)p(c|a, b)}{p(c)}$$

$$\neq p(a|c)p(b|c)$$

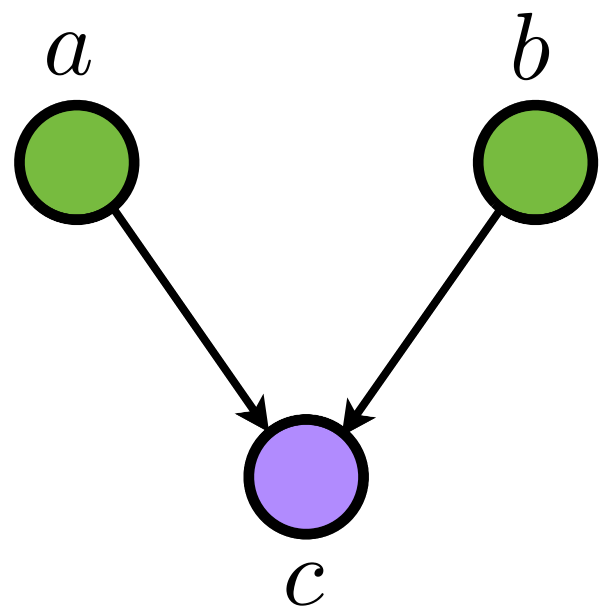
$$a \not\perp b|c$$

a and *b* not conditionally independent given *c*

COMMON EFFECT

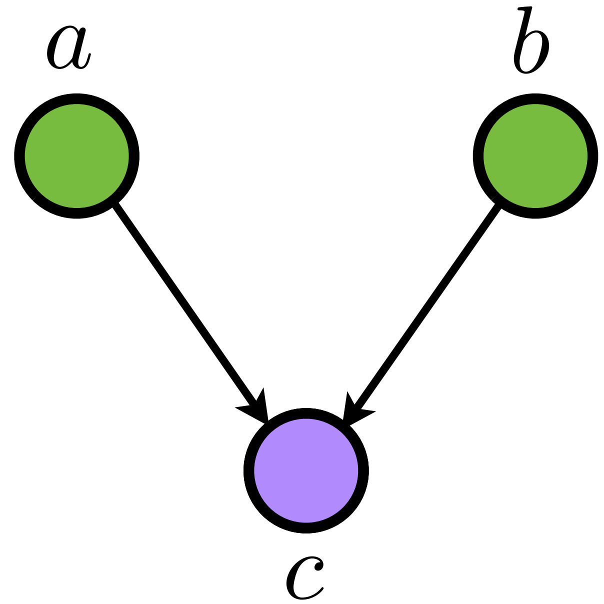


COMMON EFFECT



Node c is *head-to-head* wrt path

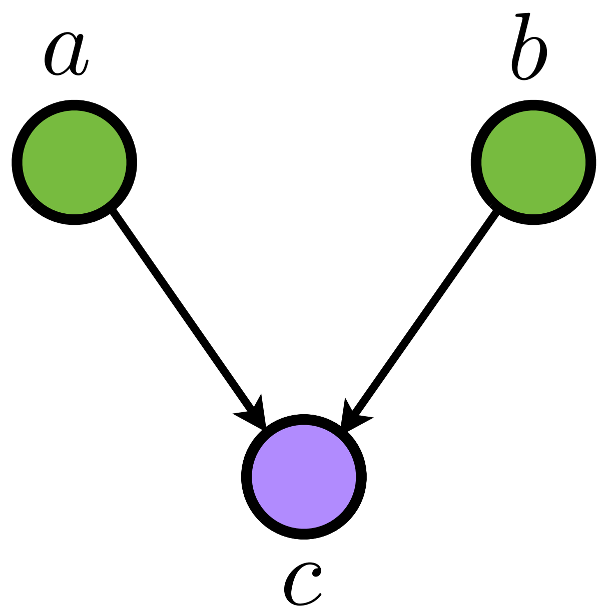
COMMON EFFECT



Node *c* is *head-to-head* wrt path

c ‘blocks’ path from *a* to *b*

COMMON EFFECT

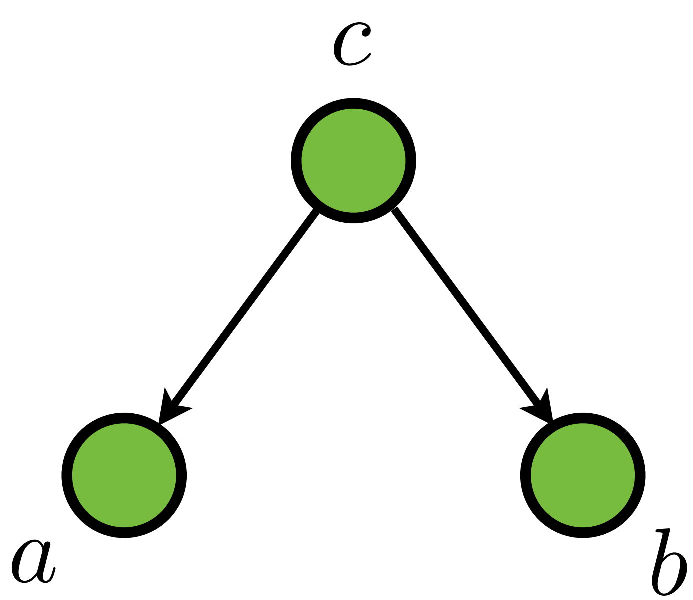


Node *c* is *head-to-head* wrt path

c 'blocks' path from *a* to *b*

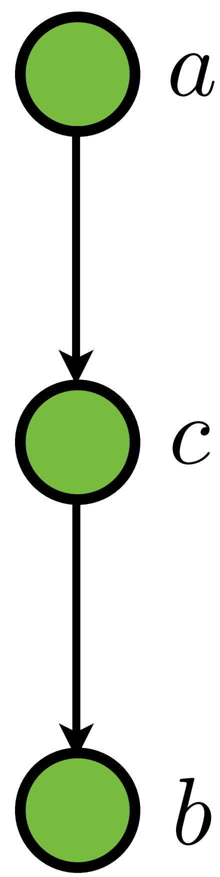
When *c* is observed, it unblocks the path from *a* to *b*

SUMMARY



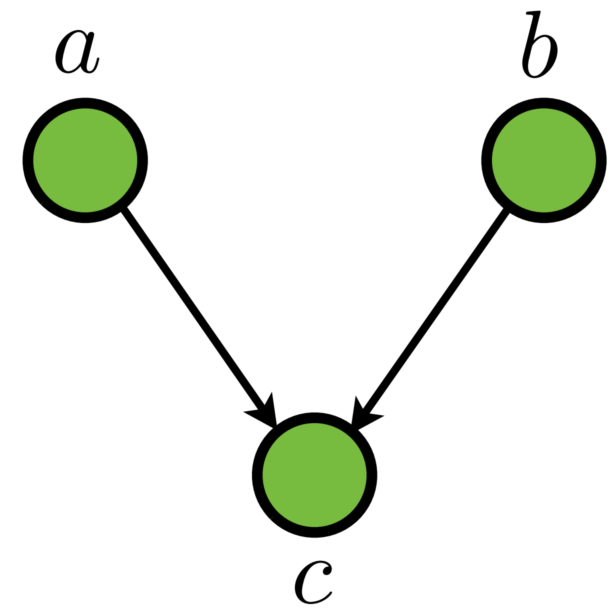
$$a \not\perp b | \emptyset$$

$$a \perp b | c$$



$$a \not\perp b | \emptyset$$

$$a \perp b | c$$

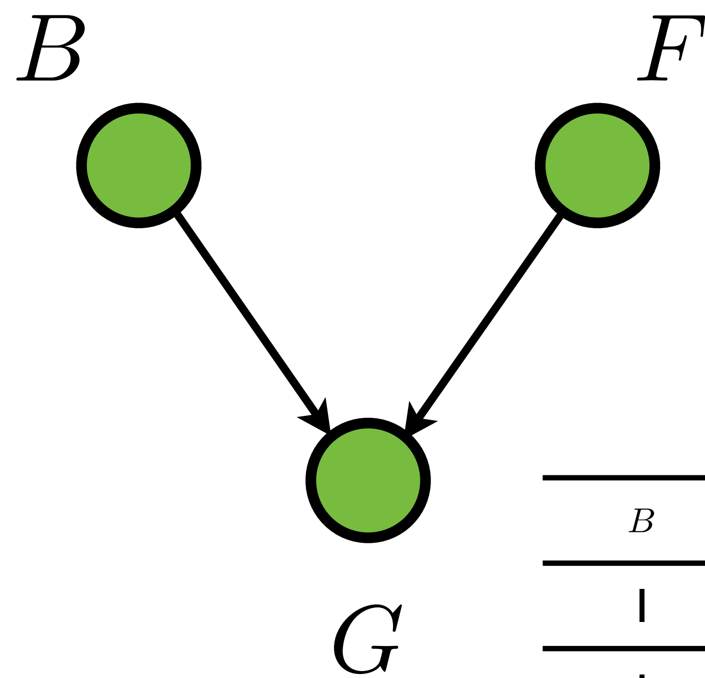


$$a \perp b | \emptyset$$

$$a \not\perp b | c$$

HEAD-TO-HEAD REVISITED

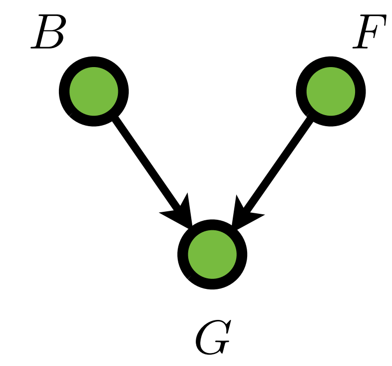
$p(B = 1) = 0.9$



$p(F = 1) = 0.9$

B	F	$p(G\|B, F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

BATTERY-FUEL-GAUGE



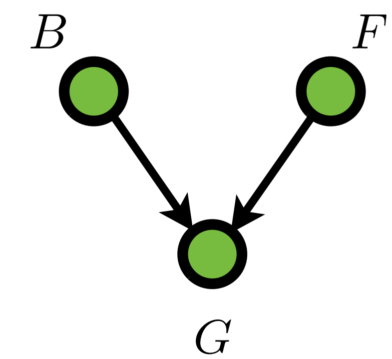
prior

$$p(F = 0) = 0.1$$

$$p(B = 1) = 0.9$$
$$p(F = 1) = 0.9$$

B	F	$p(G\|B,F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

BATTERY-FUEL-GAUGE



$$p(B=1)=0.9$$
$$p(F=1)=0.9$$

B	F	$p(G\|B,F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

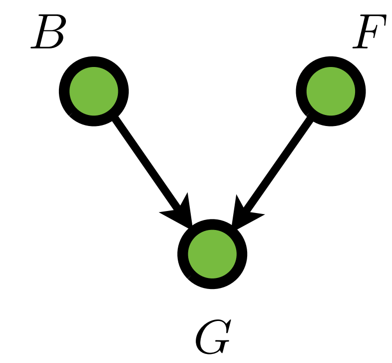
prior

$$p(F=0)=0.1$$

$$p(F=0|G=0)$$

posterior

BATTERY-FUEL-GAUGE



$$p(B = 1) = 0.9$$
$$p(F = 1) = 0.9$$

B	F	$p(G\|B, F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

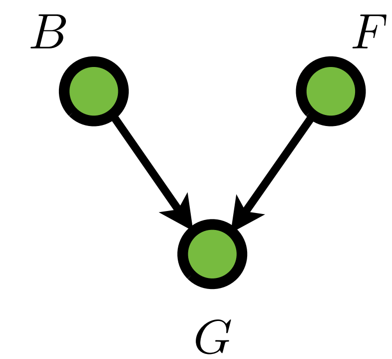
prior

$$p(F = 0) = 0.1$$

$$p(F = 0|G = 0) = \frac{p(G = 0|F = 0)p(F = 0)}{p(G = 0)}$$

posterior

BATTERY-FUEL-GAUGE



$$p(B = 1) = 0.9$$
$$p(F = 1) = 0.9$$

B	F	$p(G \vert B, F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

prior

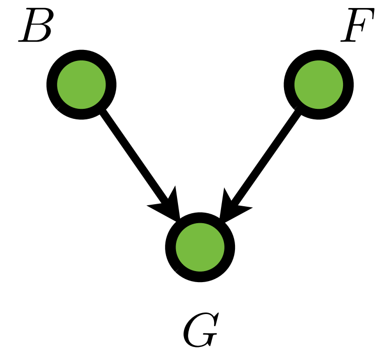
$$p(F = 0) = 0.1$$

$$p(F = 0|G = 0) = \frac{p(G = 0|F = 0)p(F = 0)}{p(G = 0)}$$

posterior

$$p(G = 0) = \sum_{B \in \{0,1\}} \sum_{F \in \{0,1\}} p(G = 0, B, F)$$

BATTERY-FUEL-GAUGE



prior

$$p(F = 0) = 0.1$$

$$p(F = 0|G = 0) = \frac{p(G = 0|F = 0)p(F = 0)}{p(G = 0)}$$

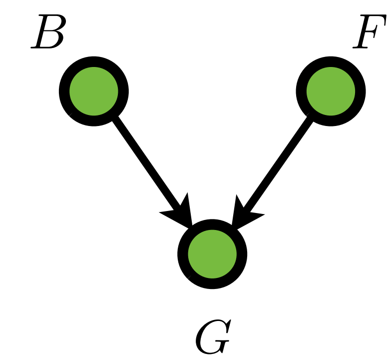
posterior

$$p(G = 0) = \sum_{B\in\{0,1\}} \sum_{F\in\{0,1\}} p(G = 0, B, F)$$

$$= \sum_{B\in\{0,1\}} \sum_{F\in\{0,1\}} p(G = 0|B, F)p(B)p(F)$$

$$p(B = 1) = 0.9$$
$$p(F = 1) = 0.9$$

B	F	$p(G\|B, F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1



BATTERY-FUEL-GAUGE

$$p(B=1)=0.9$$
$$p(F=1)=0.9$$

B	F	$p(G\|B,F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

prior

$$p(F=0)=0.1$$

$$p(F=0|G=0) = \frac{p(G=0|F=0)p(F=0)}{p(G=0)}$$

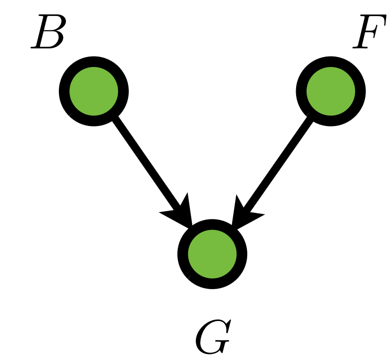
posterior

$$p(G=0) = \sum_{B\in\{0,1\}} \sum_{F\in\{0,1\}} p(G=0,B,F)$$

$$= \sum_{B\in\{0,1\}} \sum_{F\in\{0,1\}} p(G=0|B,F)p(B)p(F)$$

$$= 0.315$$

BATTERY-FUEL-GAUGE



$$p(B=1) = 0.9$$
$$p(F=1) = 0.9$$

B	F	$p(G \mid B, F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

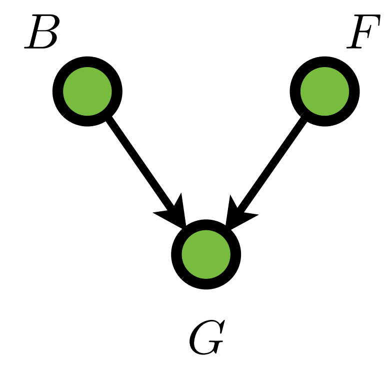
prior

$$p(F=0) = 0.1$$

$$p(F=0 \mid G=0) = \frac{p(G=0 \mid F=0)\,p(F=0)}{0.315}$$

posterior

BATTERY-FUEL-GAUGE



$$p(B=1) = 0.9$$
$$p(F=1) = 0.9$$

B	F	$p(G \mid B, F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

prior

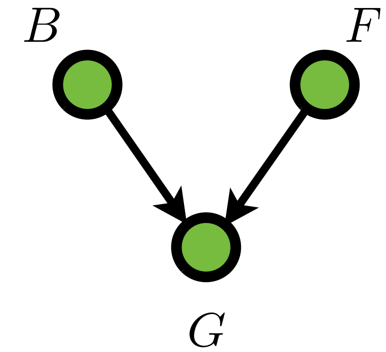
$$p(F=0) = 0.1$$

$$p(F=0 \mid G=0) = \frac{p(G=0 \mid F=0)p(F=0)}{0.315}$$

posterior

$$p(G=0 \mid F=0) = \sum_{B \in \{0,1\}} p(G=0 \mid B, F=0)p(B)$$

BATTERY-FUEL-GAUGE



$$p(B=1)=0.9$$
$$p(F=1)=0.9$$

B	F	$p(G\|B,F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

prior

$$p(F=0)=0.1$$

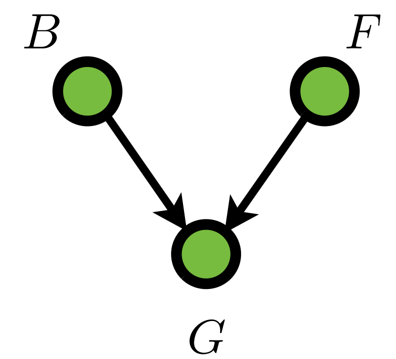
$$p(F=0|G=0) = \frac{p(G=0|F=0)p(F=0)}{0.315}$$

posterior

$$p(G=0|F=0) = \sum_{B\in\{0,1\}} p(G=0|B,F=0)p(B)$$

$$=0.81$$

BATTERY-FUEL-GAUGE



$$p(B=1)=0.9$$
$$p(F=1)=0.9$$

B	F	$p(G\|B,F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

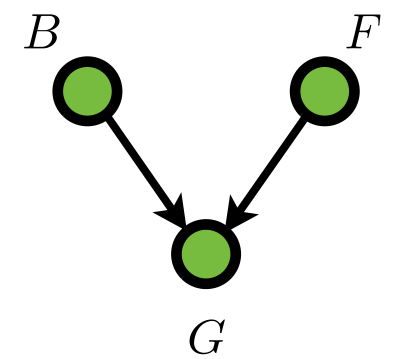
prior

$$p(F=0)=0.1$$

$$p(F=0|G=0) = \frac{0.81 \cdot p(F=0)}{0.315}$$

posterior

BATTERY-FUEL-GAUGE



$$p(B = 1) = 0.9$$
$$p(F = 1) = 0.9$$

B	F	$p(G\|B,F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

prior

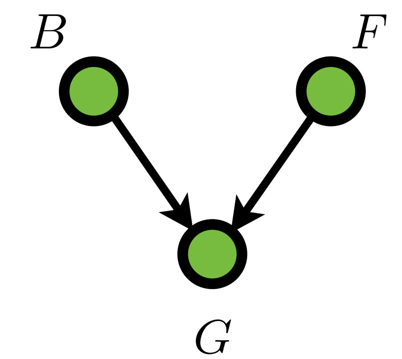
$$p(F = 0) = 0.1$$

$$p(F = 0|G = 0) = \frac{0.81 \cdot p(F = 0)}{0.315}$$

posterior

$$= \frac{0.81 \times 0.1}{0.315}$$

BATTERY-FUEL-GAUGE



$$p(B=1)=0.9$$
$$p(F=1)=0.9$$

B	F	$p(G\|B,F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

prior

$$p(F=0)=0.1$$

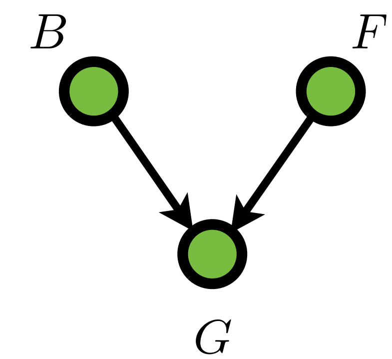
$$p(F=0|G=0) = \frac{0.81 \cdot p(F=0)}{0.315}$$

posterior

$$= \frac{0.81 \times 0.1}{0.315}$$

$$= 0.257$$

BATTERY-FUEL-GAUGE



$$p(B=1) = 0.9$$
$$p(F=1) = 0.9$$

B	F	$p(G\|B,F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

prior

$$p(F=0) = 0.1$$

$$p(F=0|G=0) = \frac{0.81 \cdot p(F=0)}{0.315}$$

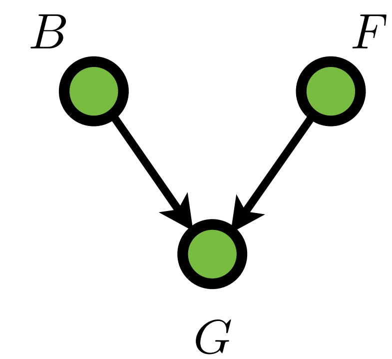
posterior

$$= \frac{0.81 \times 0.1}{0.315}$$

$$= 0.257$$

$$p(F=0|G=0) > p(F=0)$$

BATTERY-FUEL-GAUGE



$$p(B=1)=0.9$$
$$p(F=1)=0.9$$

B	F	$p(G\|B,F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

prior

$$p(F=0)=0.1$$

$$p(F=0|G=0)=\frac{0.81\cdot p(F=0)}{0.315}$$

posterior

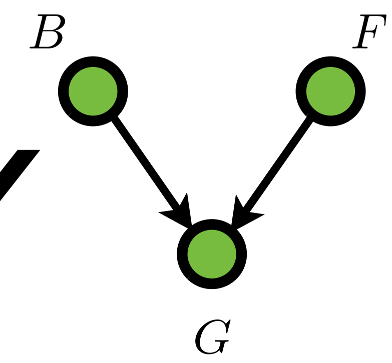
$$=\frac{0.81\times 0.1}{0.315}$$

$$=0.257$$

$$p(F=0|G=0)>p(F=0)$$

posterior prior

EXPLAINING AWAY



prior

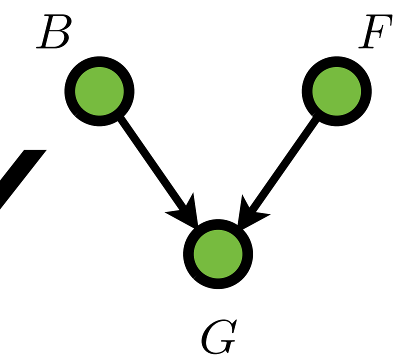
$$p(F = 0) = 0.1$$

$$p(F = 0 | G = 0, B = 0) = 0.111$$

$$p(B = 1) = 0.9$$
$$p(F = 1) = 0.9$$

B	F	$p(G\|B, F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

EXPLAINING AWAY



prior

$$p(F=0)=0.1$$

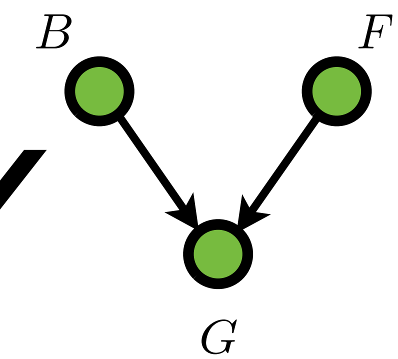
$$p(F=0|G=0,B=0)=0.111$$

$$p(B=1)=0.9$$
$$p(F=1)=0.9$$

B	F	$p(G\|B,F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

$$p(F=0|G=0)>p(F=0|G=0,B=0)$$

EXPLAINING AWAY



prior

$$p(F = 0) = 0.1$$

$$p(F = 0|G = 0, B = 0) = 0.111$$

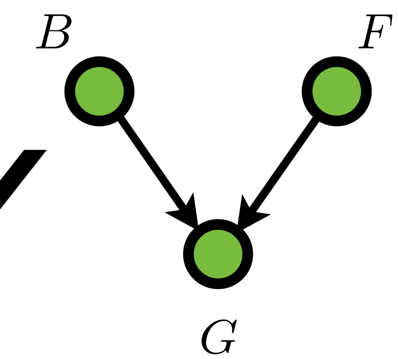
$$p(B = 1) = 0.9$$
$$p(F = 1) = 0.9$$

B	F	$p(G\|B, F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

$$p(F = 0|G = 0) > p(F = 0|G = 0, B = 0)$$

$$p(F = 0|G = 0) > p(F = 0)$$

EXPLAINING AWAY



$$p(B=1)=0.9$$
$$p(F=1)=0.9$$

B	F	$p(G\|B,F)$
1	1	0.8
1	0	0.2
0	1	0.2
0	0	0.1

prior

$$p(F=0)=0.1$$

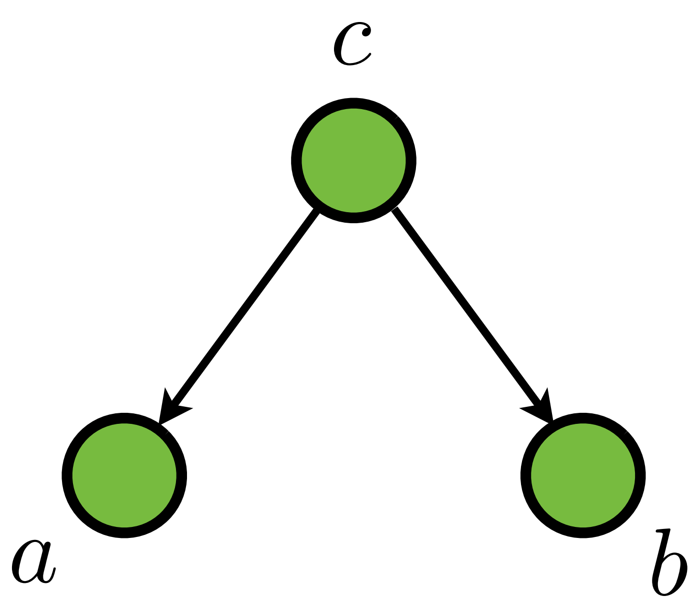
$$p(F=0|G=0,B=0)=0.111$$

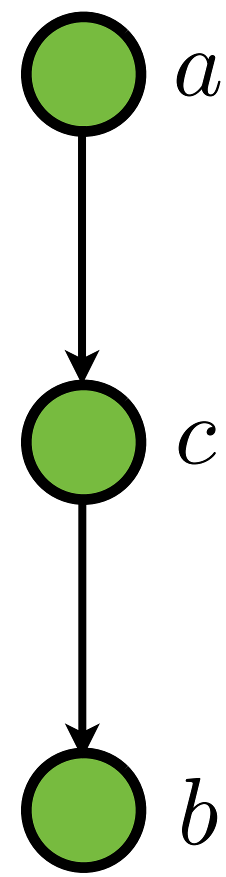
$$p(F=0|G=0)>p(F=0|G=0,B=0)$$

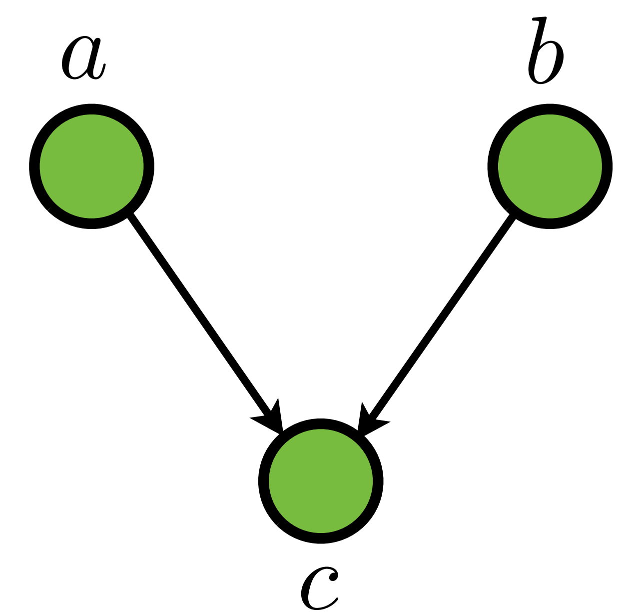
$$p(F=0|G=0)>p(F=0)$$

F and *B* become dependent as a result of observing *G*

SUMMARY







$a \not\perp b|\emptyset$

$a \perp b|c$

$a \not\perp b|\emptyset$

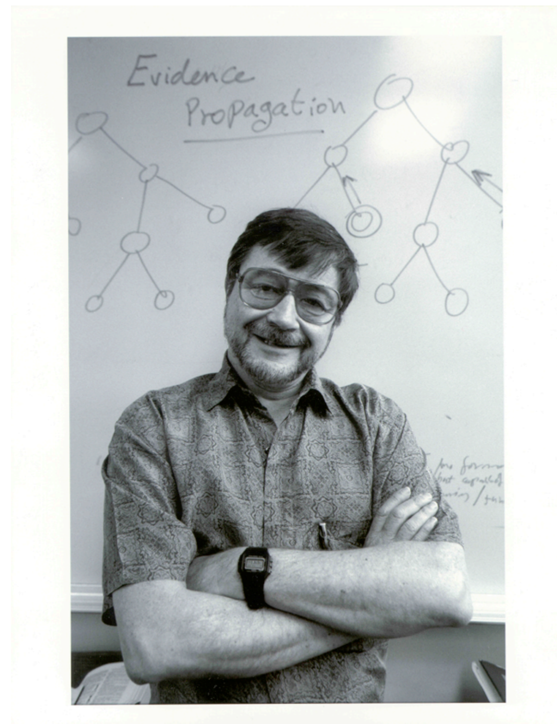
$a \perp b|c$

$a \perp b|\emptyset$

$a \not\perp b|c$

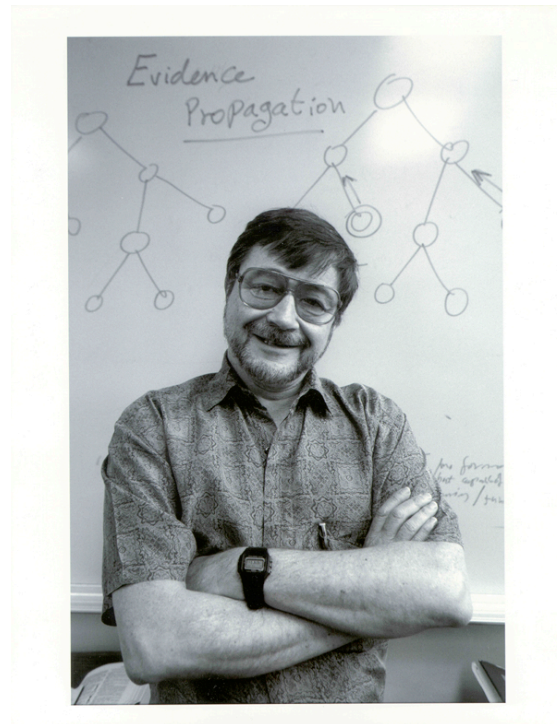
D-SEPARATION

JUDEA PEARL



D-SEPARATION

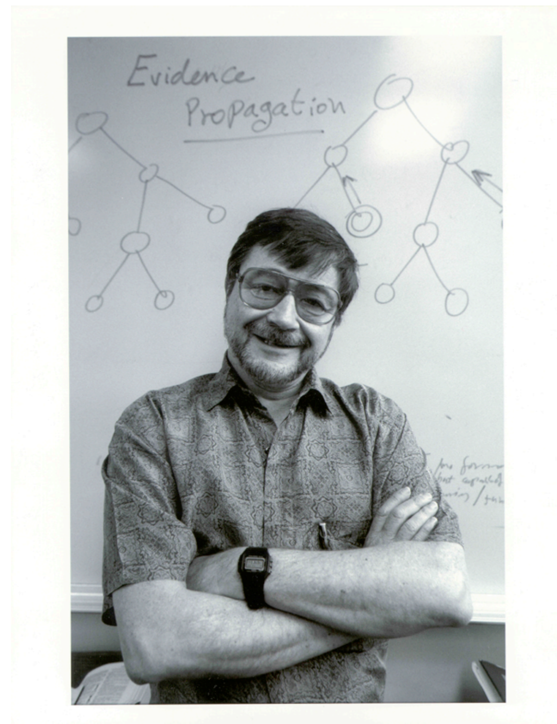
JUDEA PEARL



- DIRECTION-DEPENDENT SEPARATION

D-SEPARATION

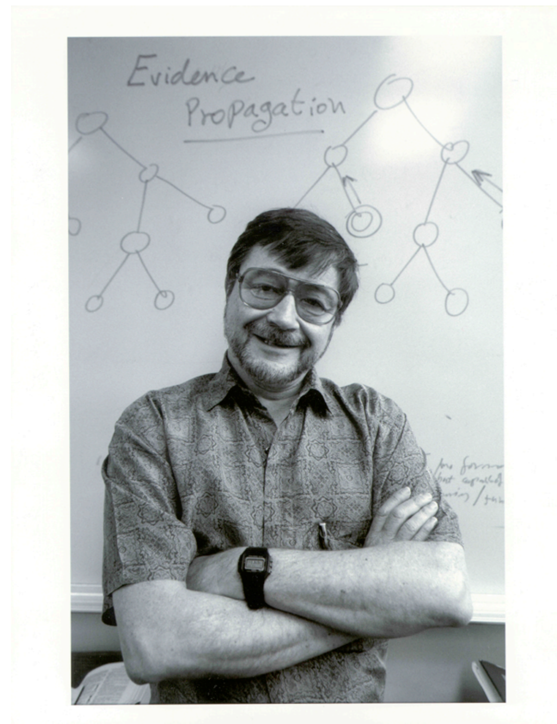
JUDEA PEARL



- DIRECTION-DEPENDENT SEPARATION
- A TECHNIQUE TO DETERMINING CONDITIONAL INDEPENDENCE PROPERTIES FROM GRAPHICAL MODELS

D-SEPARATION

JUDEA PEARL



- DIRECTION-DEPENDENT SEPARATION
- A TECHNIQUE TO DETERMINING CONDITIONAL INDEPENDENCE PROPERTIES FROM GRAPHICAL MODELS
- "IS THE SET OF VARIABLES **A** CONDITIONALLY INDEPENDENT OF THE SET **B** GIVEN THE SET **C**?"

D-SEPARATION

Task: Determine if $A \perp B|C$

D-SEPARATION

Task: Determine if $A \perp B | C$

Consider all possible paths from any node in A to any node in B

D-SEPARATION

Task: Determine if $A \perp B|C$

Consider all possible paths from any node in A to any node in B

A path is blocked if it includes a node where either

D-SEPARATION

Task: Determine if $A \perp B | C$

Consider all possible paths from any node in A to any node in B

A path is blocked if it includes a node where either

the arrows on the path meet head-to-tail or tail-to-tail at the node, and the node is in C

D-SEPARATION

Task: Determine if $A \perp B|C$

Consider all possible paths from any node in A to any node in B

A path is blocked if it includes a node where either

the arrows on the path meet head-to-tail or tail-to-tail at the node, and the node is in C

the arrows on the path meet head-to-head and neither the node nor any of its descendants is in C

D-SEPARATION

Task: Determine if $A \perp B|C$

Consider all possible paths from any node in A to any node in B

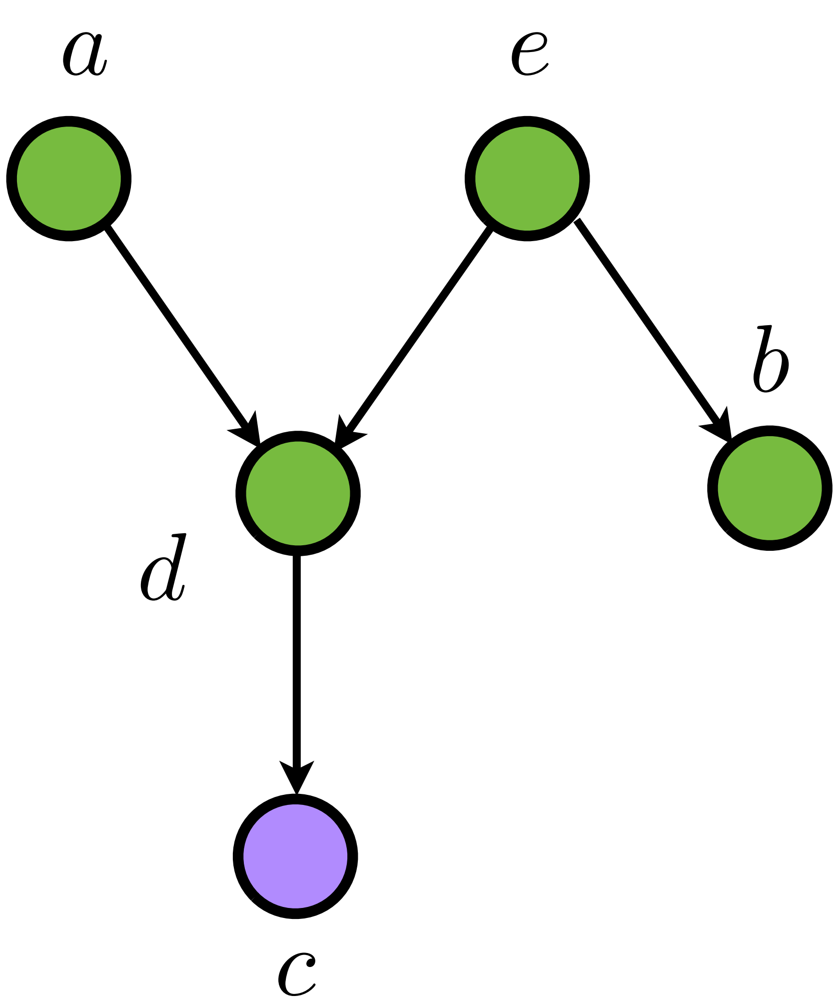
A path is blocked if it includes a node where either

the arrows on the path meet head-to-tail or tail-to-tail at the node, and the node is in C

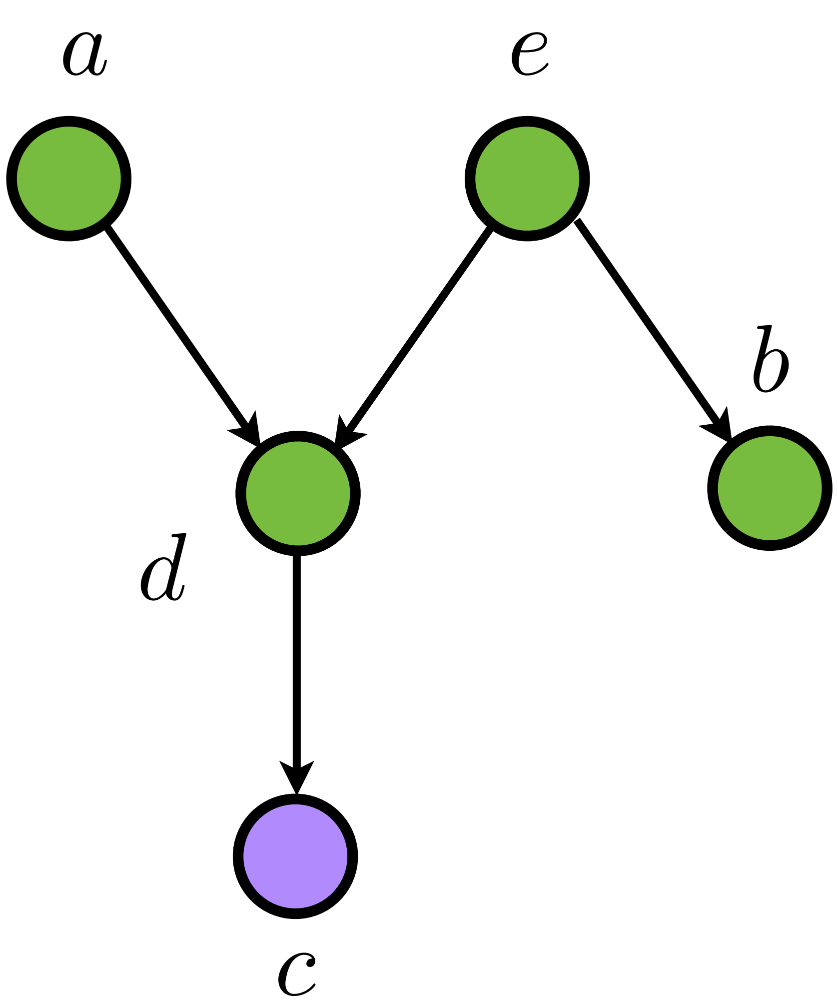
the arrows on the path meet head-to-head and neither the node nor any of its descendants is in C

If all paths are blocked then A is d-separated from B by C, then

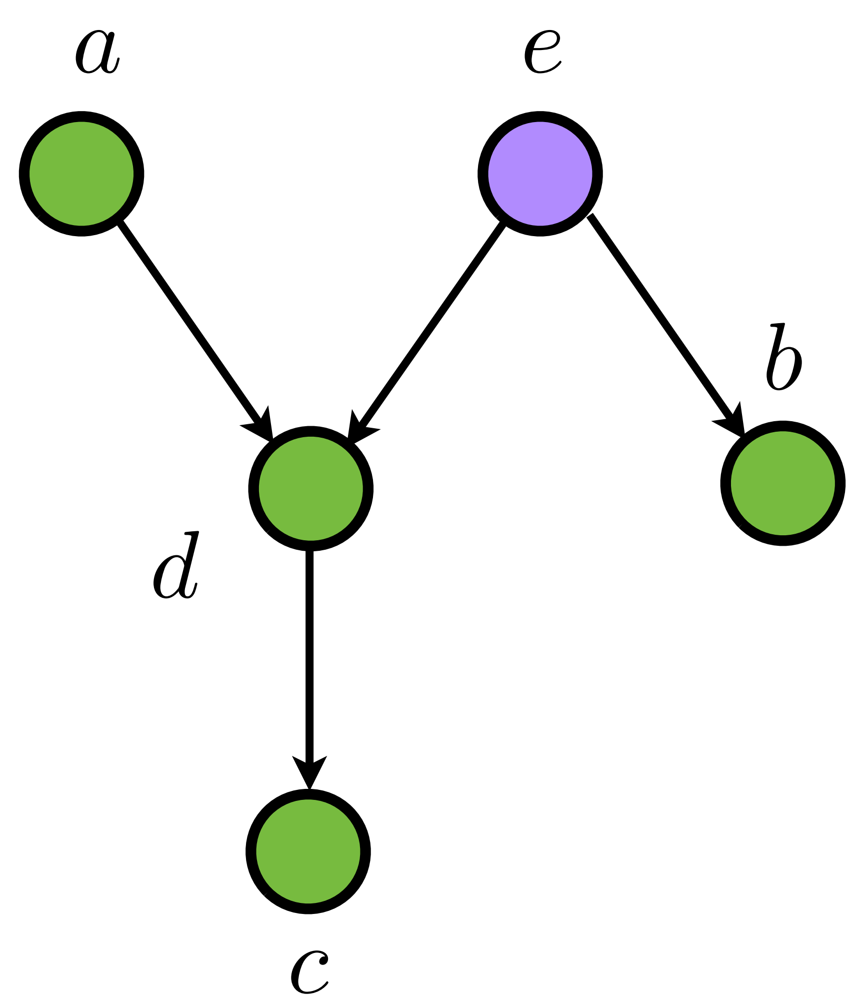
$$A \perp B|C$$



$$a \perp b | c \ ?$$



$a \perp b|c$?



$a \perp b|e$?

INFERENCE ON BAYESIAN NETWORKS

INFERENCE ON BAYESIAN NETWORKS

JOHN

INFERENCE ON BAYESIAN NETWORKS

JOHN

$$p(B|J = \texttt{TRUE}) = ?$$

INFERENCE ON BAYESIAN NETWORKS

$p(J|A = \texttt{TRUE}) = 0.9$

JOHN

$p(B|J = \texttt{TRUE}) = ?$

INFERENCE ON BAYESIAN NETWORKS

$$p(J|A = \texttt{TRUE}) = 0.9$$

$$p(B = \texttt{TRUE}) = 0.001$$

$$p(J|A = \texttt{TRUE}) = 0.9$$

$$p(J|A = \texttt{FALSE}) = 0.05$$

JOHN

$$p(B|J = \texttt{TRUE}) = ?$$

INFERENCE ON BAYESIAN NETWORKS

$$p(J|A = \texttt{TRUE}) = 0.9$$

$$p(B = \texttt{TRUE}) = 0.001$$

$$p(J|A = \texttt{TRUE}) = 0.9$$

$$p(J|A = \texttt{FALSE}) = 0.05$$

JOHN

$$p(B|J = \texttt{TRUE}) = ?$$

In 1000 days:

INFERENCE ON BAYESIAN NETWORKS

$$p(J|A = \texttt{TRUE}) = 0.9$$

$$p(B = \texttt{TRUE}) = 0.001$$

$$p(J|A = \texttt{TRUE}) = 0.9$$

$$p(J|A = \texttt{FALSE}) = 0.05$$

JOHN

$$p(B|J = \texttt{TRUE}) = ?$$

In 1000 days:

There will be 1 burglary

INFERENCE ON BAYESIAN NETWORKS

$$p(J|A = \texttt{TRUE}) = 0.9$$

$$p(B = \texttt{TRUE}) = 0.001$$

$$p(J|A = \texttt{TRUE}) = 0.9$$

$$p(J|A = \texttt{FALSE}) = 0.05$$

JOHN

$$p(B|J = \texttt{TRUE}) = ?$$

In 1000 days:

There will be 1 burglary

John will call 50 times!

INFERENCE ON BAYESIAN NETWORKS

$$p(J|A = \texttt{TRUE}) = 0.9$$

$$p(B = \texttt{TRUE}) = 0.001$$

$$p(J|A = \texttt{TRUE}) = 0.9$$

$$p(J|A = \texttt{FALSE}) = 0.05$$

JOHN

$$p(B|J = \texttt{TRUE}) = ?$$

In 1000 days:

There will be 1 burglary

John will call 50 times!

$$p(B|J = \texttt{TRUE}) \simeq 0.2$$

INFERENCE

INFERENCE

- WHEN SOME VARIABLES ARE OBSERVED WHAT CAN WE SAY ABOUT THE UNOBSERVED VARIABLES?

INFERENCE

- WHEN SOME VARIABLES ARE OBSERVED WHAT CAN WE SAY ABOUT THE UNOBSERVED VARIABLES?
- DIRECTED GRAPH TO SPECIFY MODEL

INFERENCE

- WHEN SOME VARIABLES ARE OBSERVED WHAT CAN WE SAY ABOUT THE UNOBSERVED VARIABLES?
- DIRECTED GRAPH TO SPECIFY MODEL
- FACTOR GRAPH FOR INFERENCE AND LEARNING

INFERENCE

- WHEN SOME VARIABLES ARE OBSERVED WHAT CAN WE SAY ABOUT THE UNOBSERVED VARIABLES?
- DIRECTED GRAPH TO SPECIFY MODEL
- FACTOR GRAPH FOR INFERENCE AND LEARNING

$$ab + ac = a(b + c)$$

FACTOR GRAPH

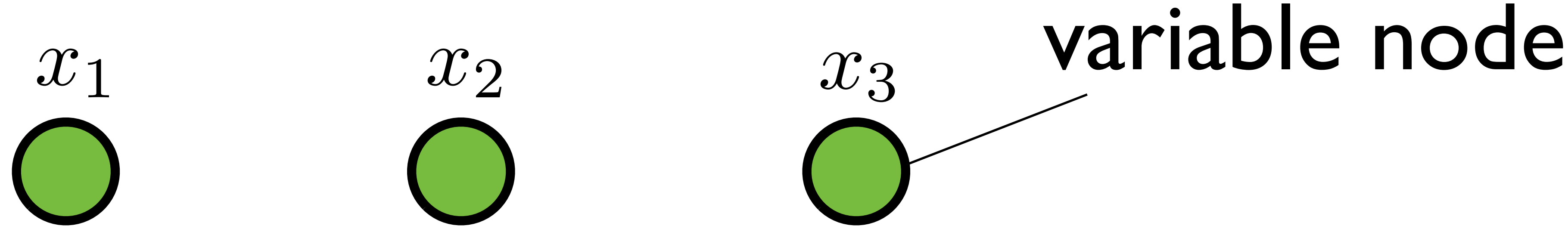
$$p(\mathbf{x}) = \prod_s f_s(\mathbf{x}_s)$$

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

FACTOR GRAPH

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

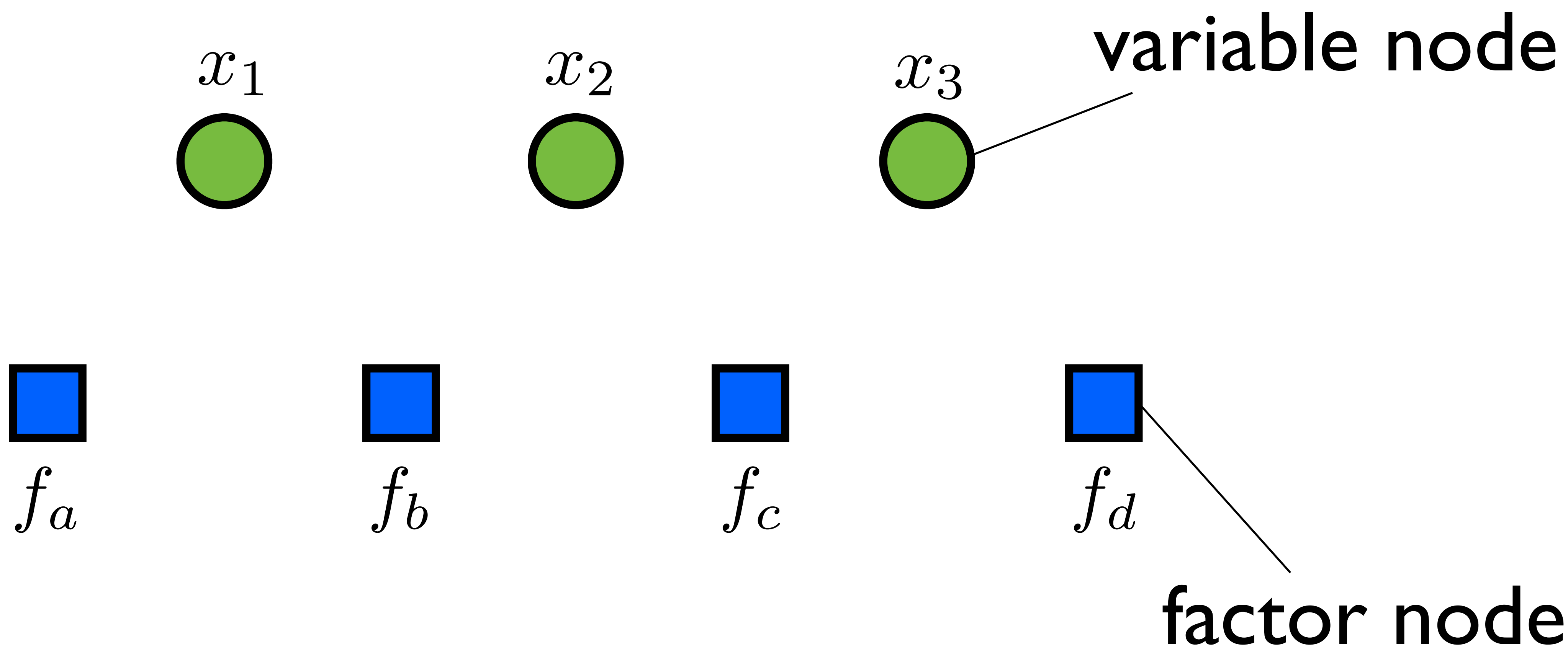
$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_1, x_2) f_c(x_2, x_3) f_d(x_3)$$



FACTOR GRAPH

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

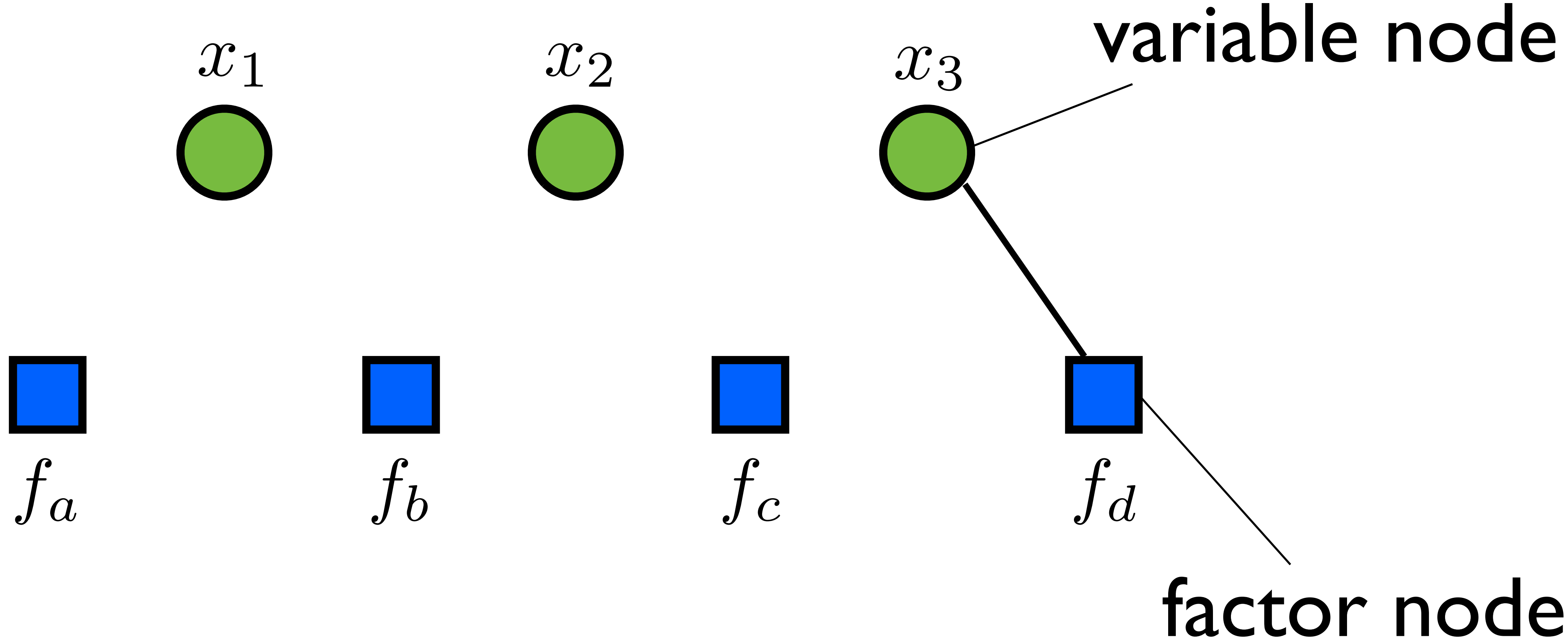
$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_1, x_2) f_c(x_2, x_3) f_d(x_3)$$



FACTOR GRAPH

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

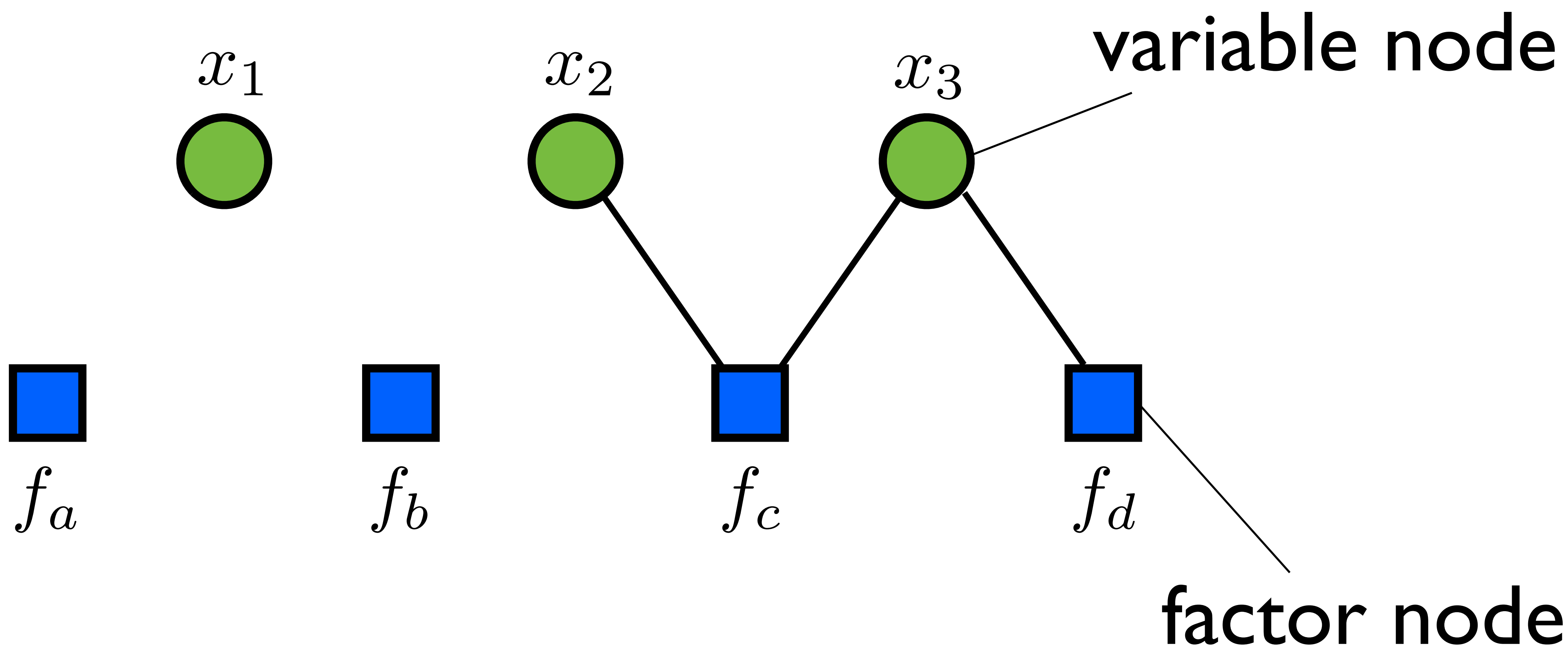
$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_1, x_2) f_c(x_2, x_3) f_d(x_3)$$



FACTOR GRAPH

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

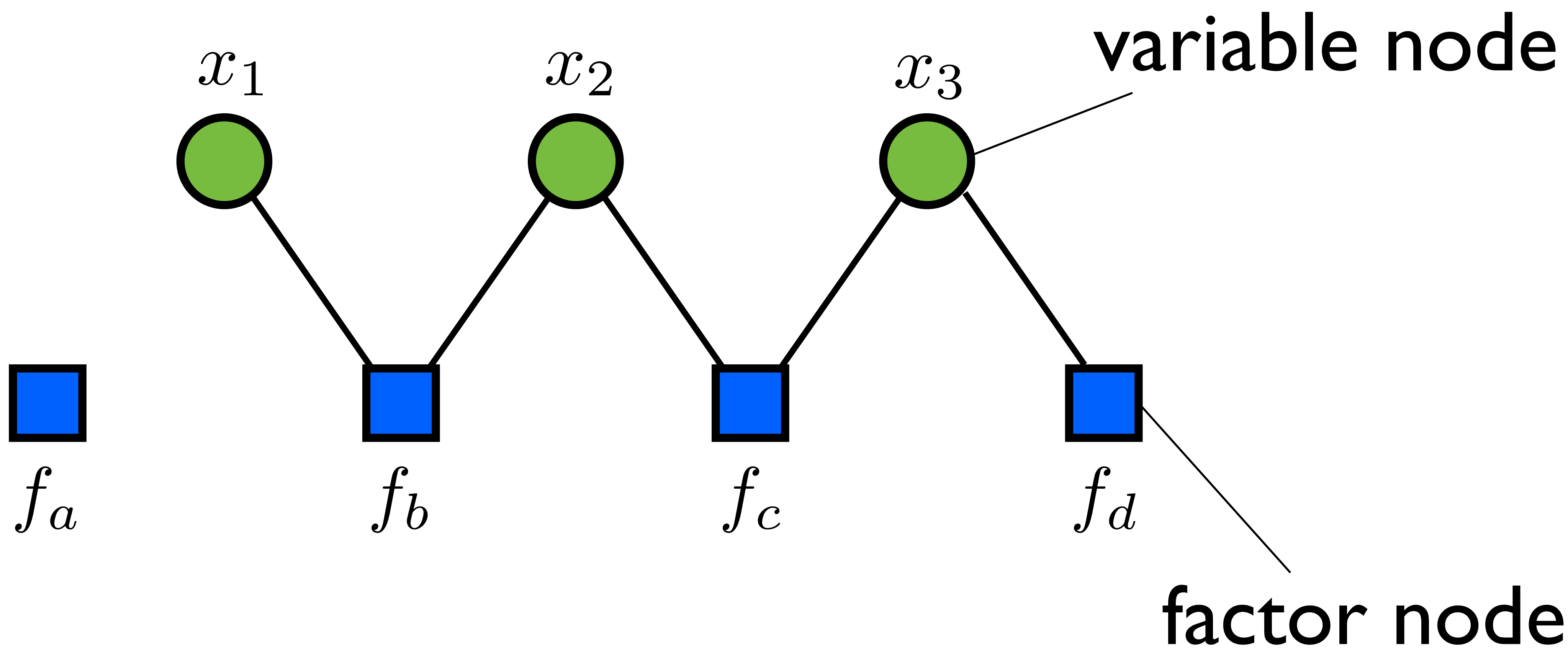
$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_1, x_2) f_c(x_2, x_3) f_d(x_3)$$



FACTOR GRAPH

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

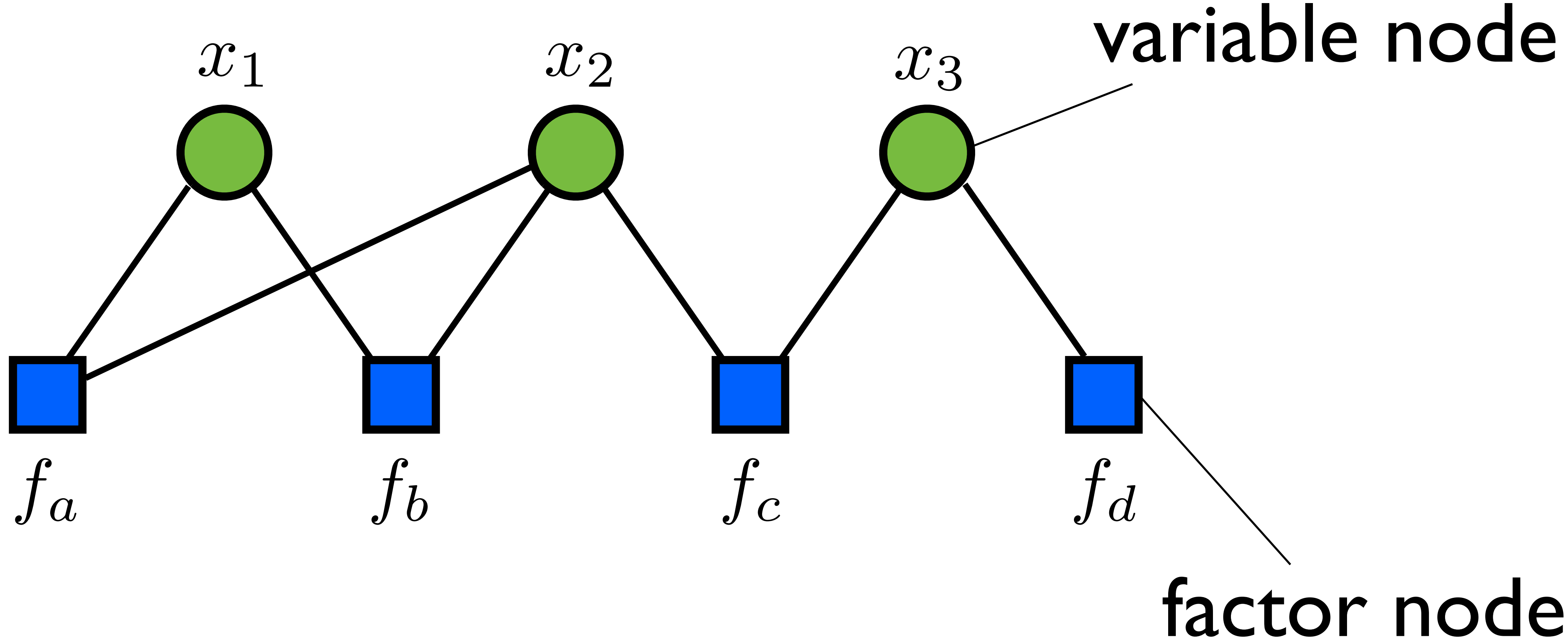
$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_1, x_2) f_c(x_2, x_3) f_d(x_3)$$



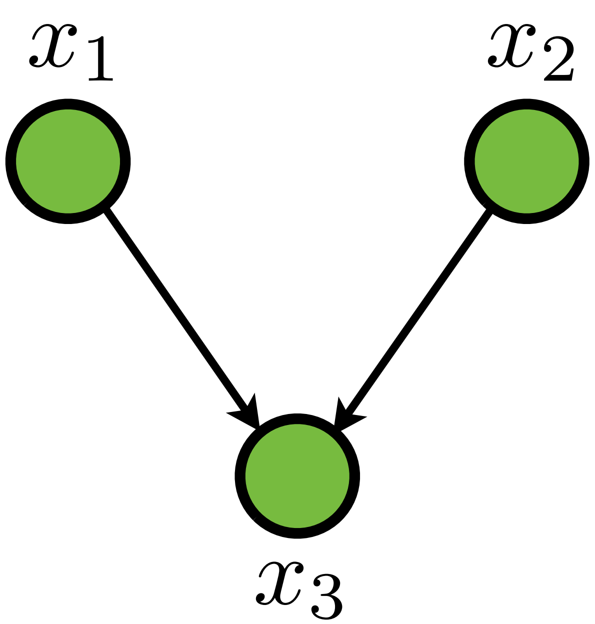
FACTOR GRAPH

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

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

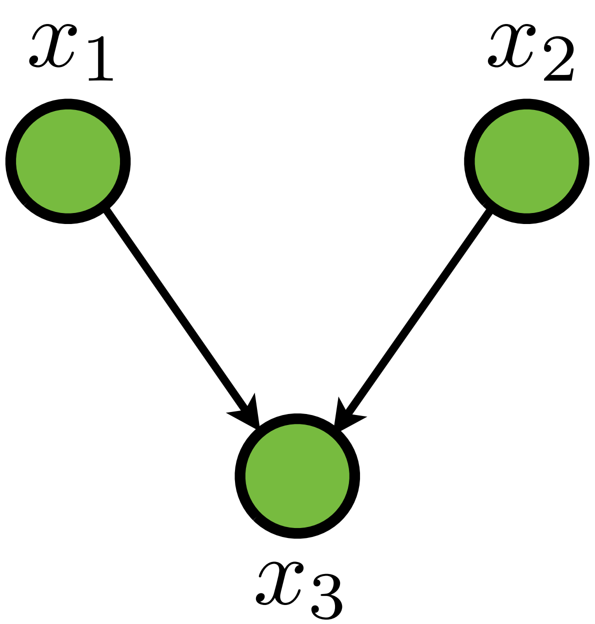


FACTOR GRAPH



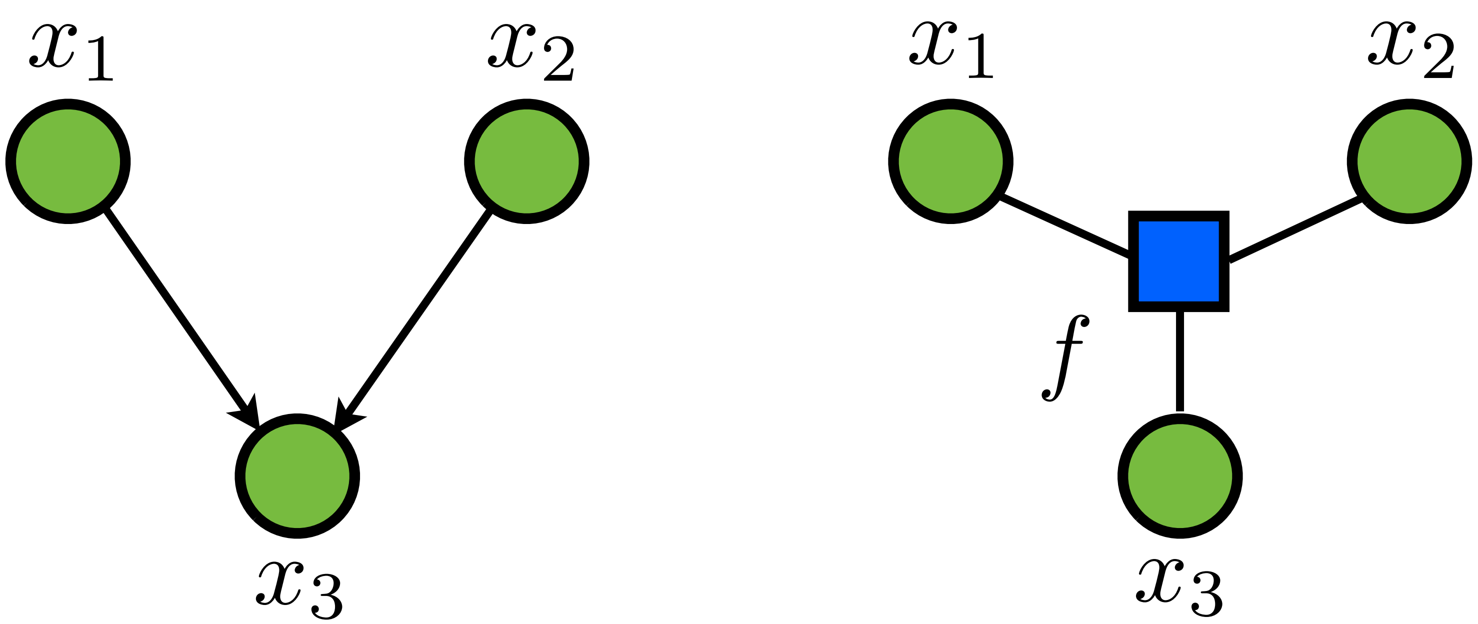
$p(\mathbf{x})$

FACTOR GRAPH



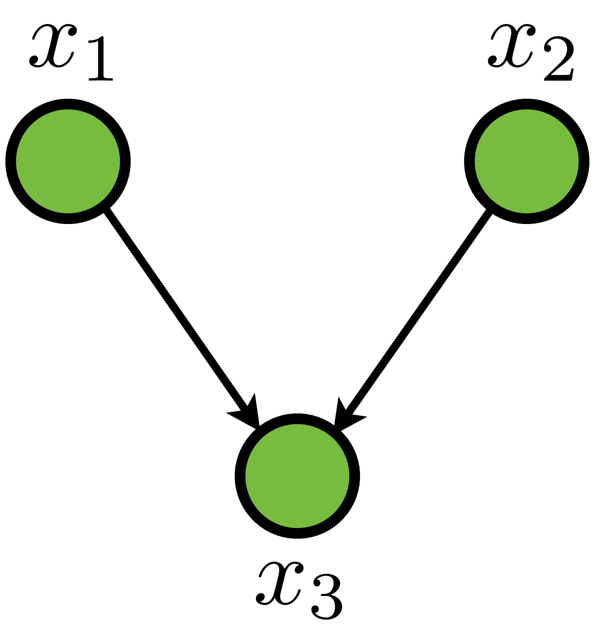
$$p(\mathbf{x}) = \frac{p(x_1, x_2, x_3)}{f}$$

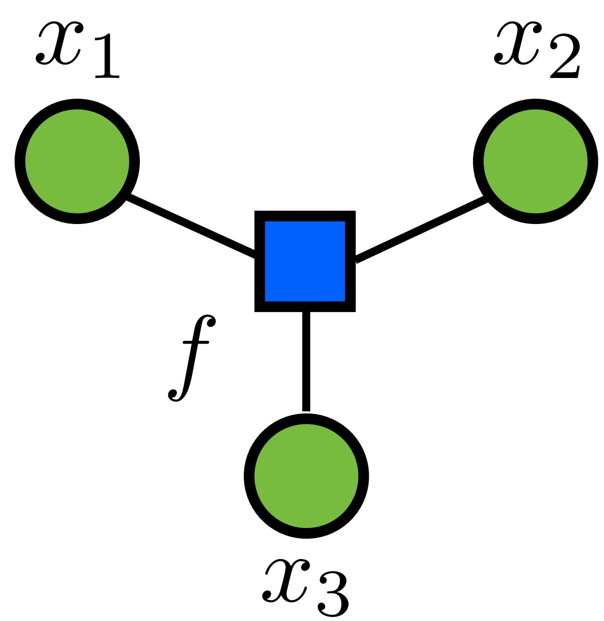
FACTOR GRAPH



$$p(\mathbf{x}) = \frac{p(x_1, x_2, x_3)}{f}$$

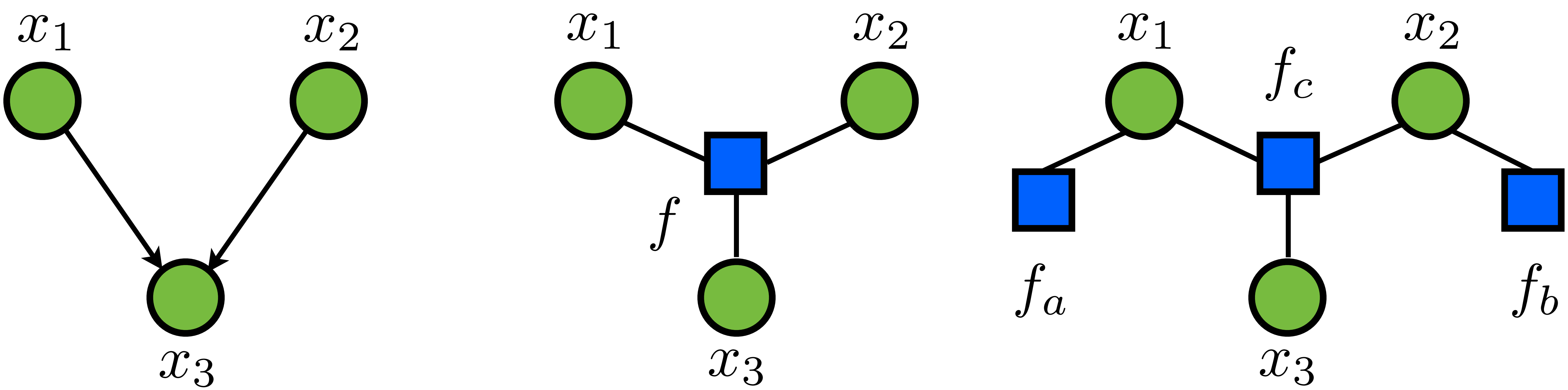
FACTOR GRAPH





$$p(\mathbf{x}) = \underbrace{p(x_1, x_2, x_3)}_{f} = \underbrace{p(x_1)}_{f_a}\underbrace{p(x_2)}_{f_b}\underbrace{p(x_3|x_1, x_2)}_{f_c}$$

FACTOR GRAPH



$$p(\mathbf{x}) = \underbrace{p(x_1, x_2, x_3)}_{f} = \underbrace{p(x_1)}_{f_a}\underbrace{p(x_2)}_{f_b}\underbrace{p(x_3|x_1, x_2)}_{f_c}$$

SUM-PRODUCT ALGORITHM

SUM-PRODUCT ALGORITHM

- TREE OR *POLYTREE*

SUM-PRODUCT ALGORITHM

- TREE OR *POLYTREE*
- BELIEF PROPAGATION: SPECIAL CASE

SUM-PRODUCT ALGORITHM

- TREE OR *POLYTREE*
- BELIEF PROPAGATION: SPECIAL CASE
- GOAL:

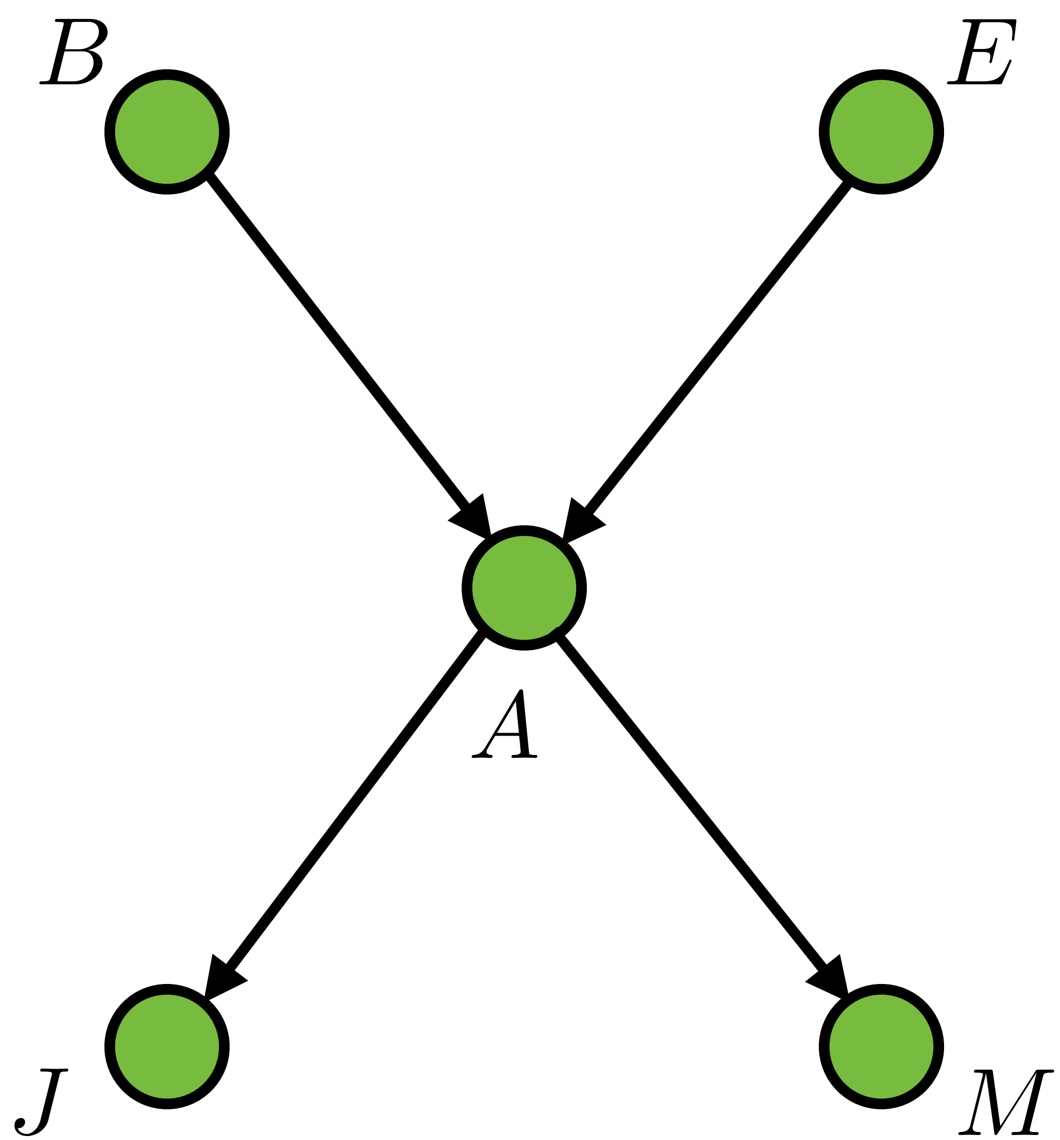
SUM-PRODUCT ALGORITHM

- TREE OR *POLYTREE*
- BELIEF PROPAGATION: SPECIAL CASE
- GOAL:
  - OBTAIN EFFICIENT EXACT ALGORITHM FOR FINDING MARGINALS

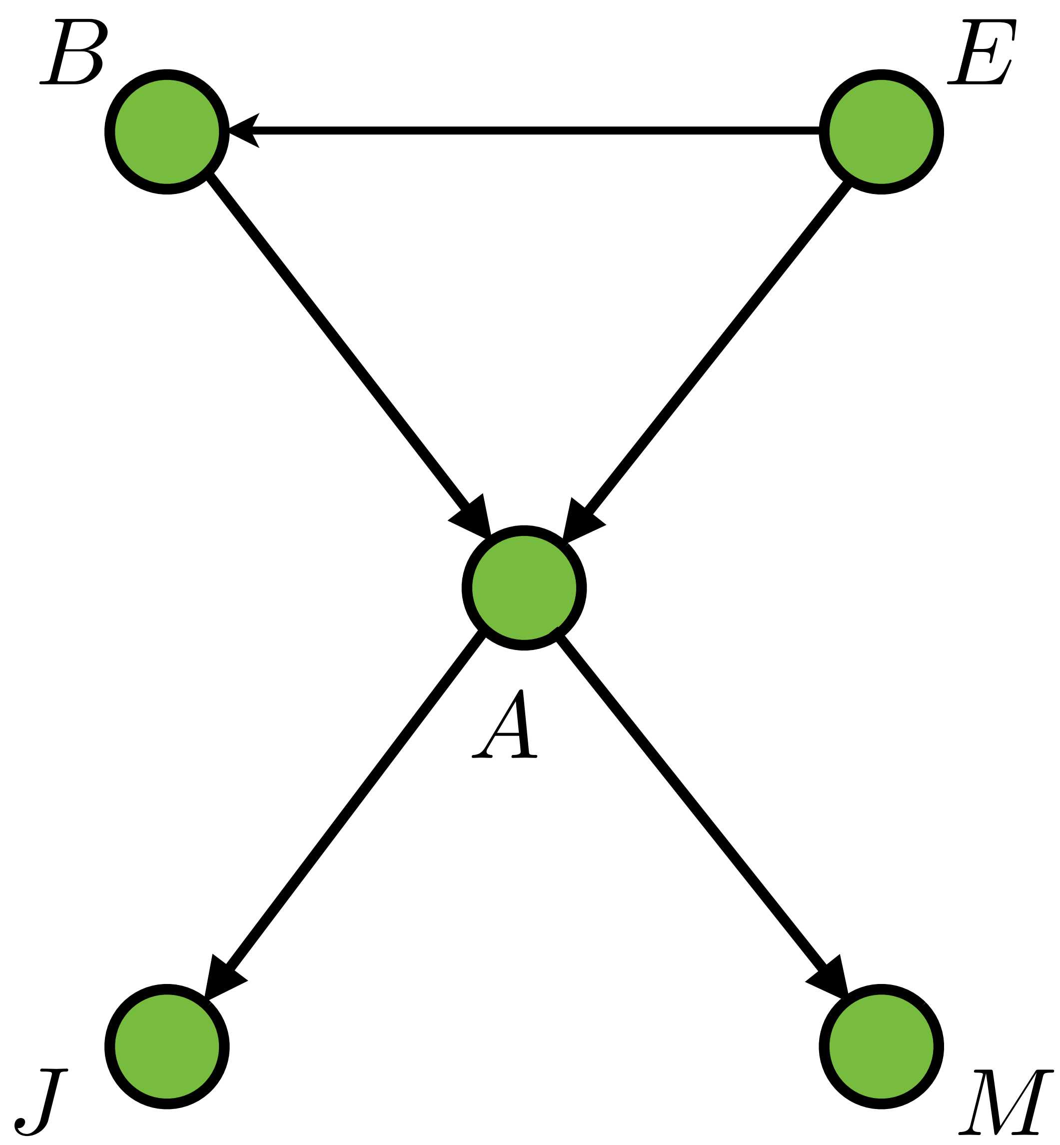
SUM-PRODUCT ALGORITHM

- TREE OR *POLYTREE*
- BELIEF PROPAGATION: SPECIAL CASE
- GOAL:
  - OBTAIN EFFICIENT EXACT ALGORITHM FOR FINDING MARGINALS
  - ALLOW COMPUTATIONS TO BE SHARED EFFICIENTLY

POLYTREE

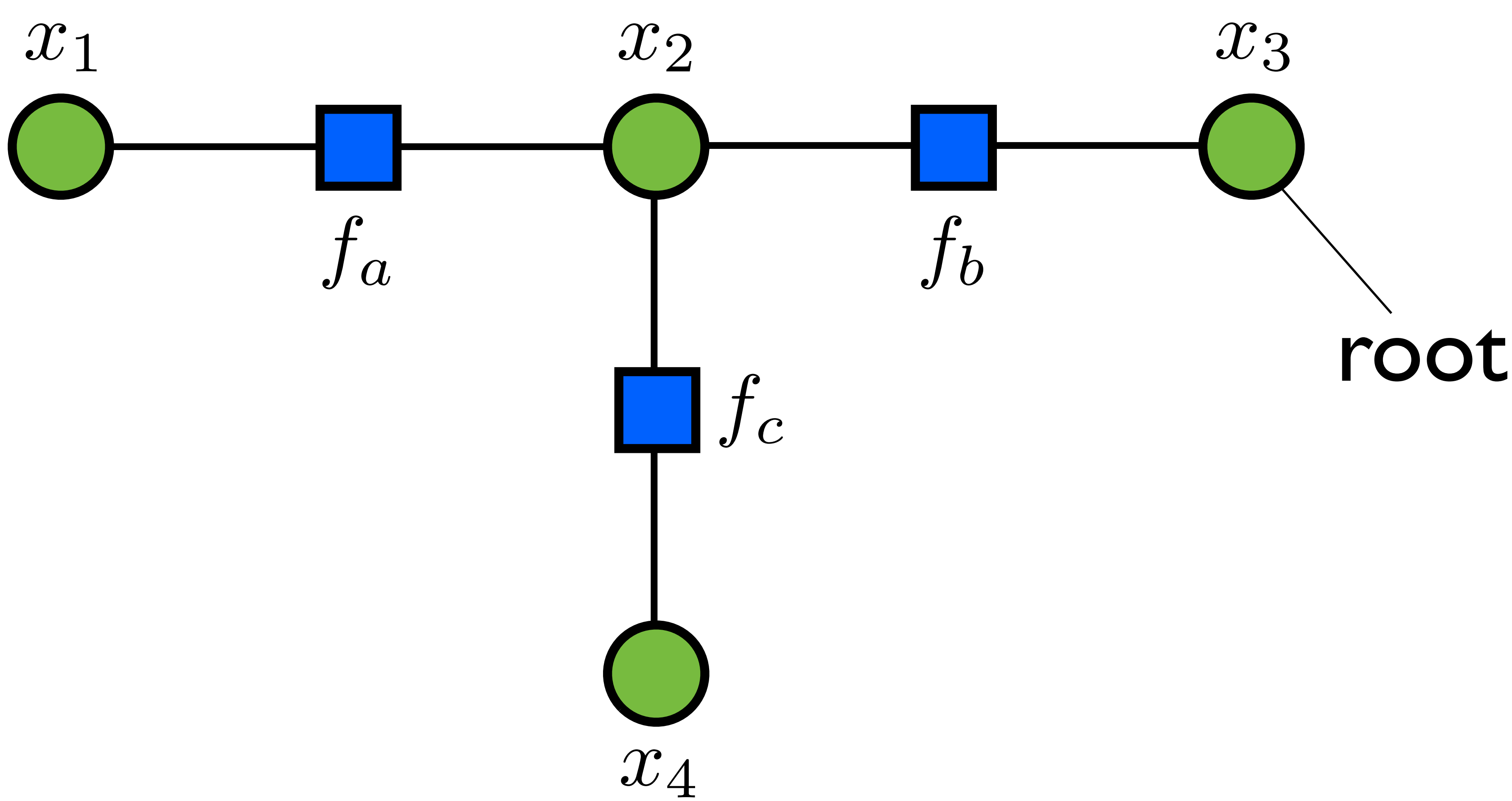


~~POLYTREE~~ DAG



SUM-PRODUCT ALGORITHM

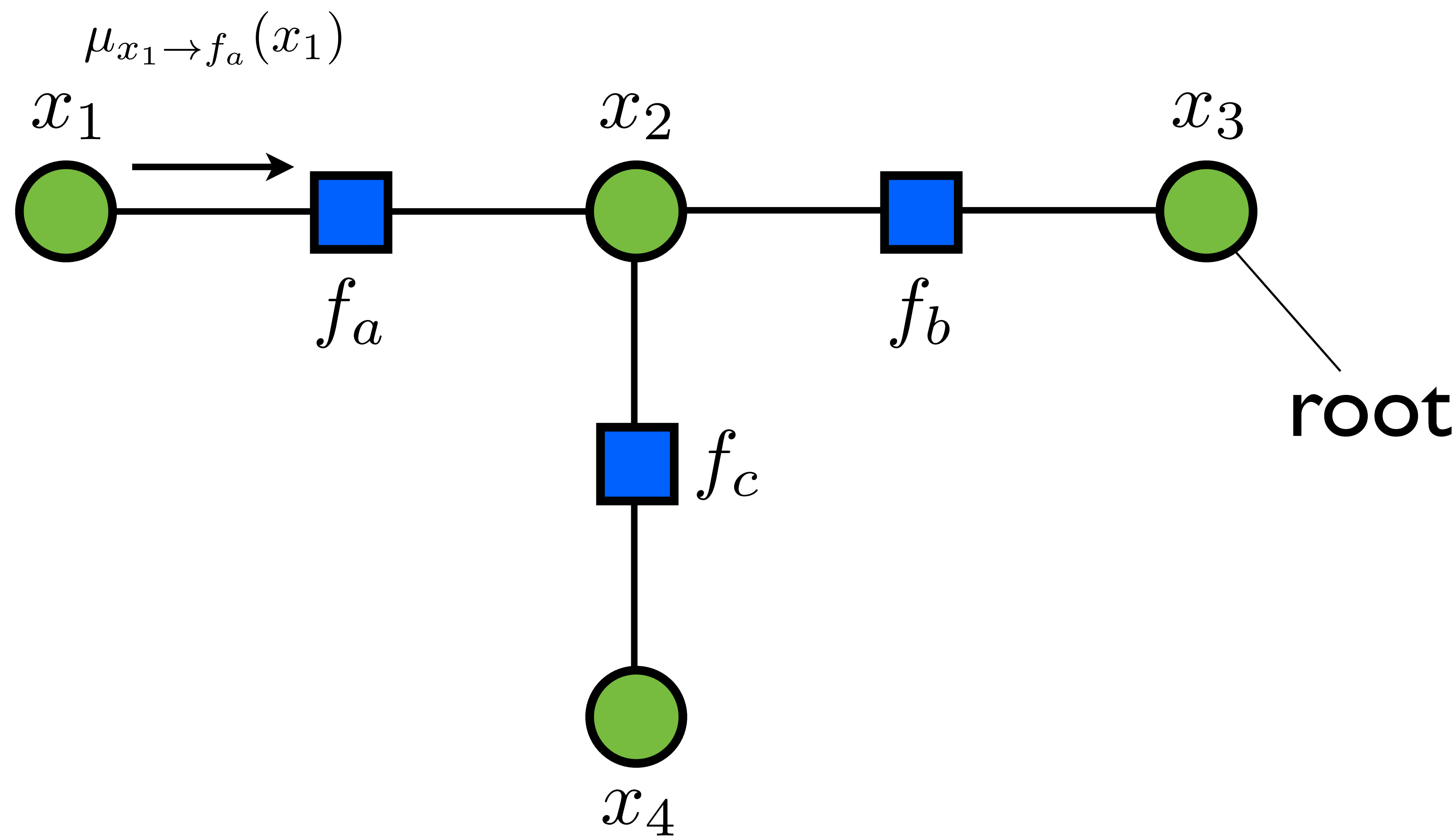
EXAMPLE



$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4)$$

SUM-PRODUCT ALGORITHM

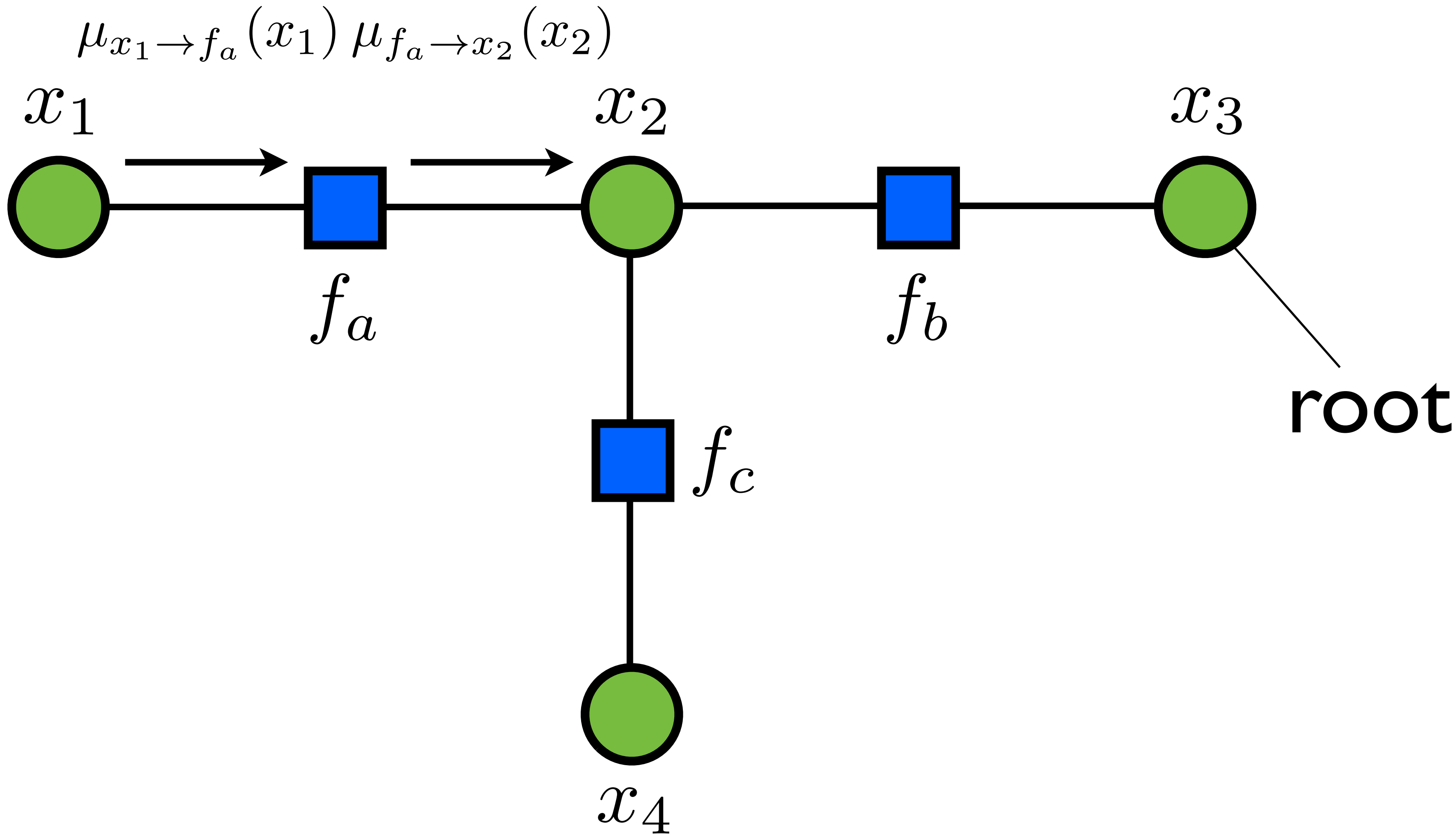
EXAMPLE



$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4)$$

SUM-PRODUCT ALGORITHM

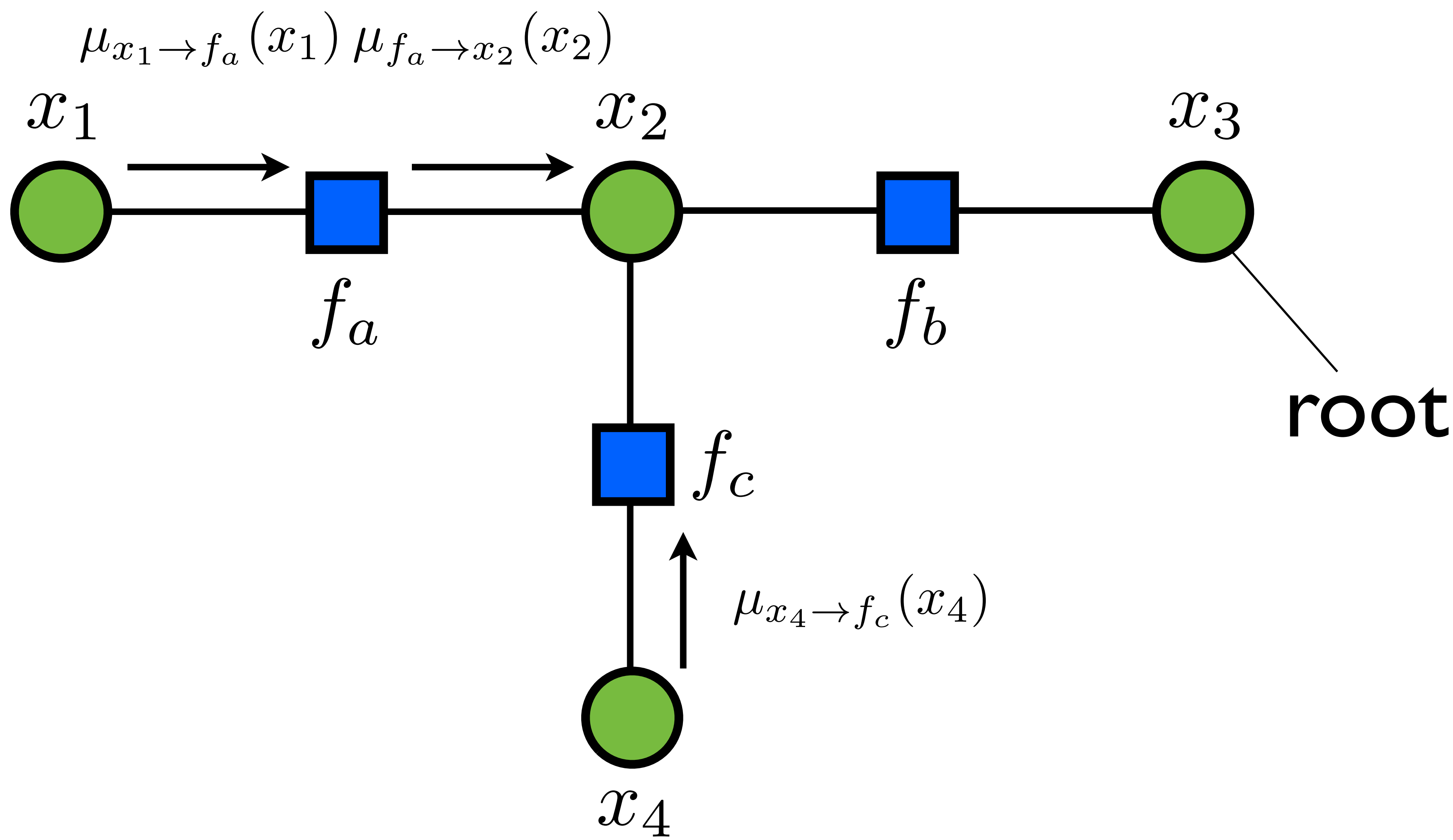
EXAMPLE



$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4)$$

SUM-PRODUCT ALGORITHM

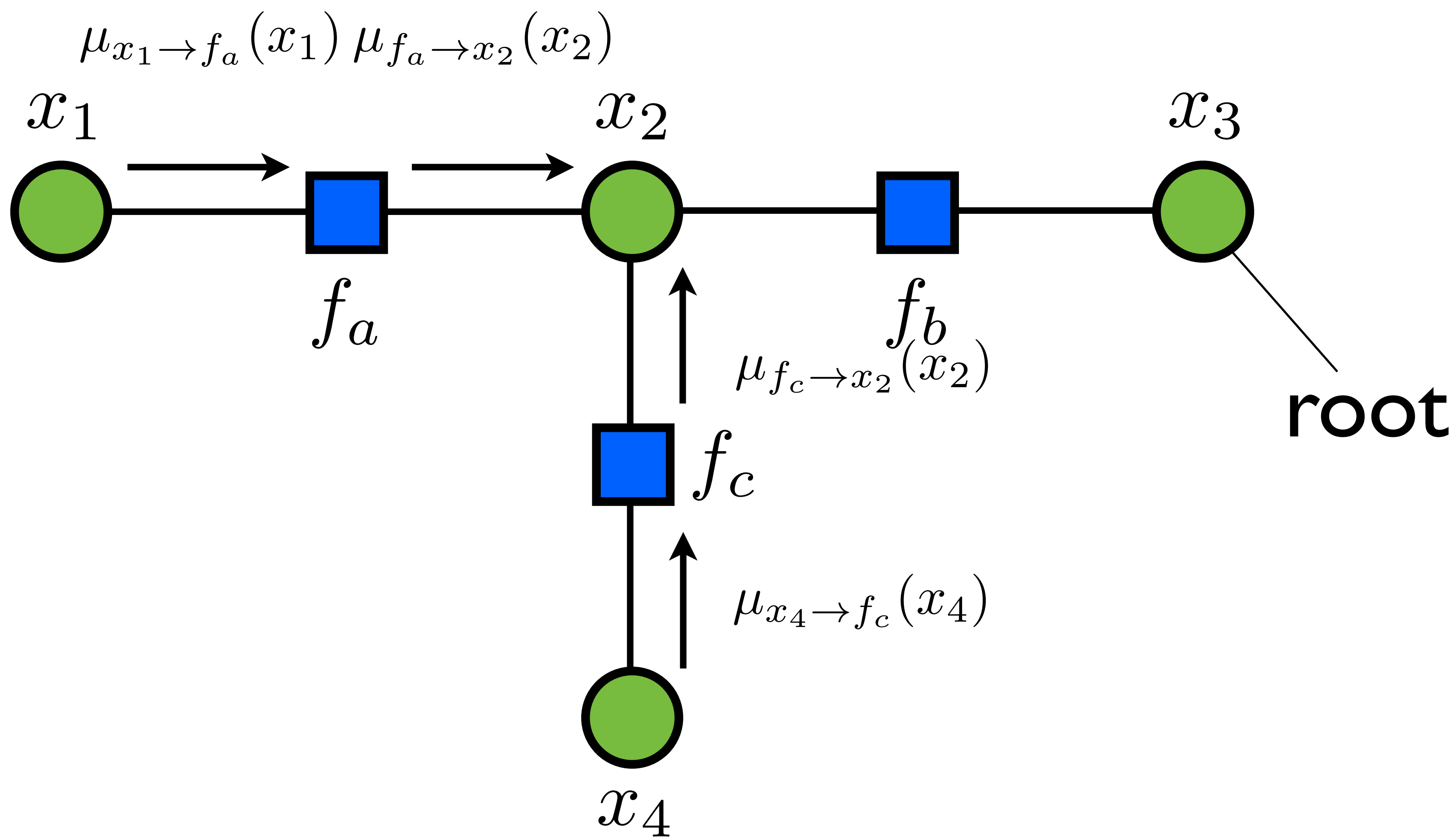
EXAMPLE



$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4)$$

SUM-PRODUCT ALGORITHM

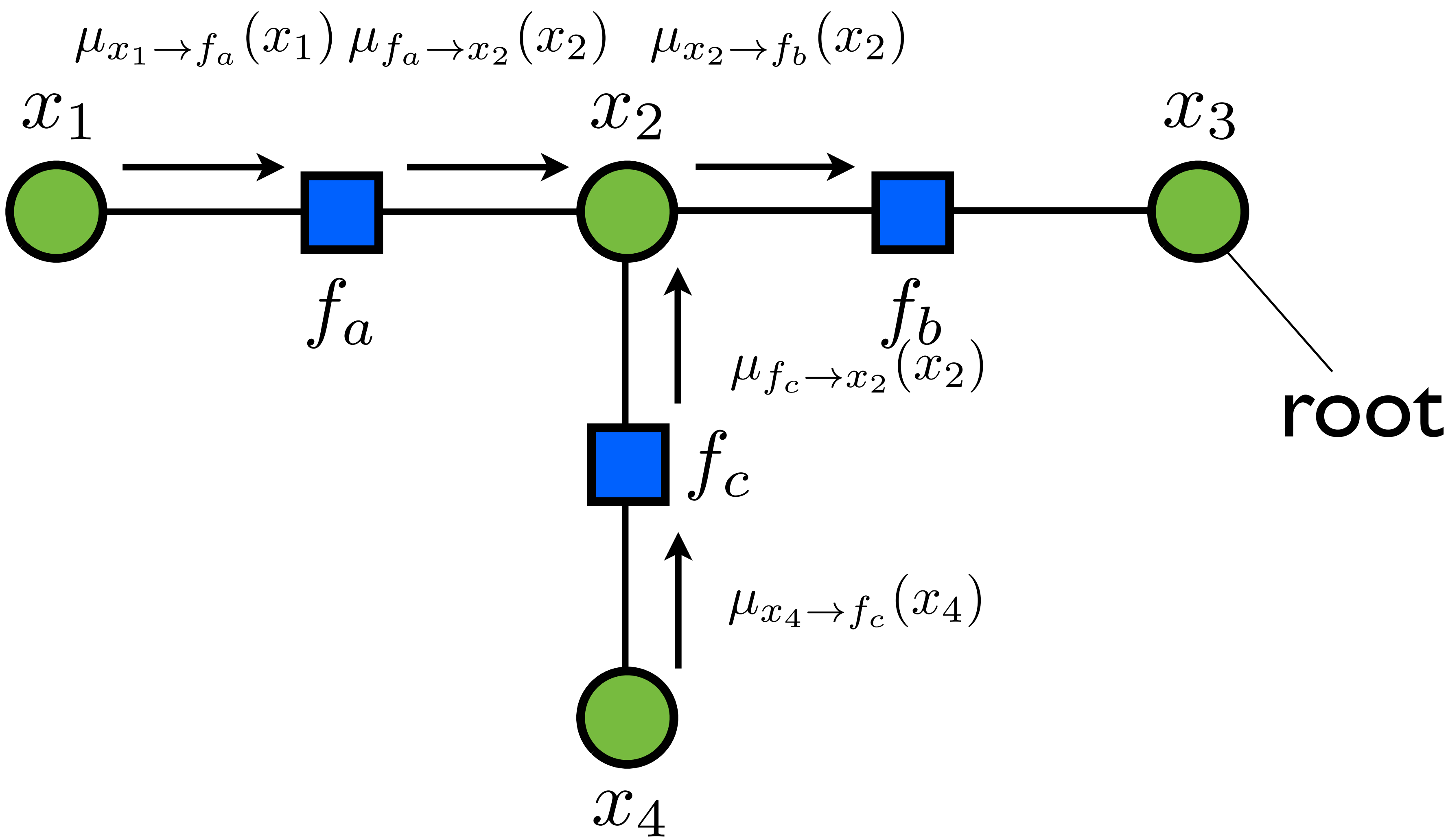
EXAMPLE



$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4)$$

SUM-PRODUCT ALGORITHM

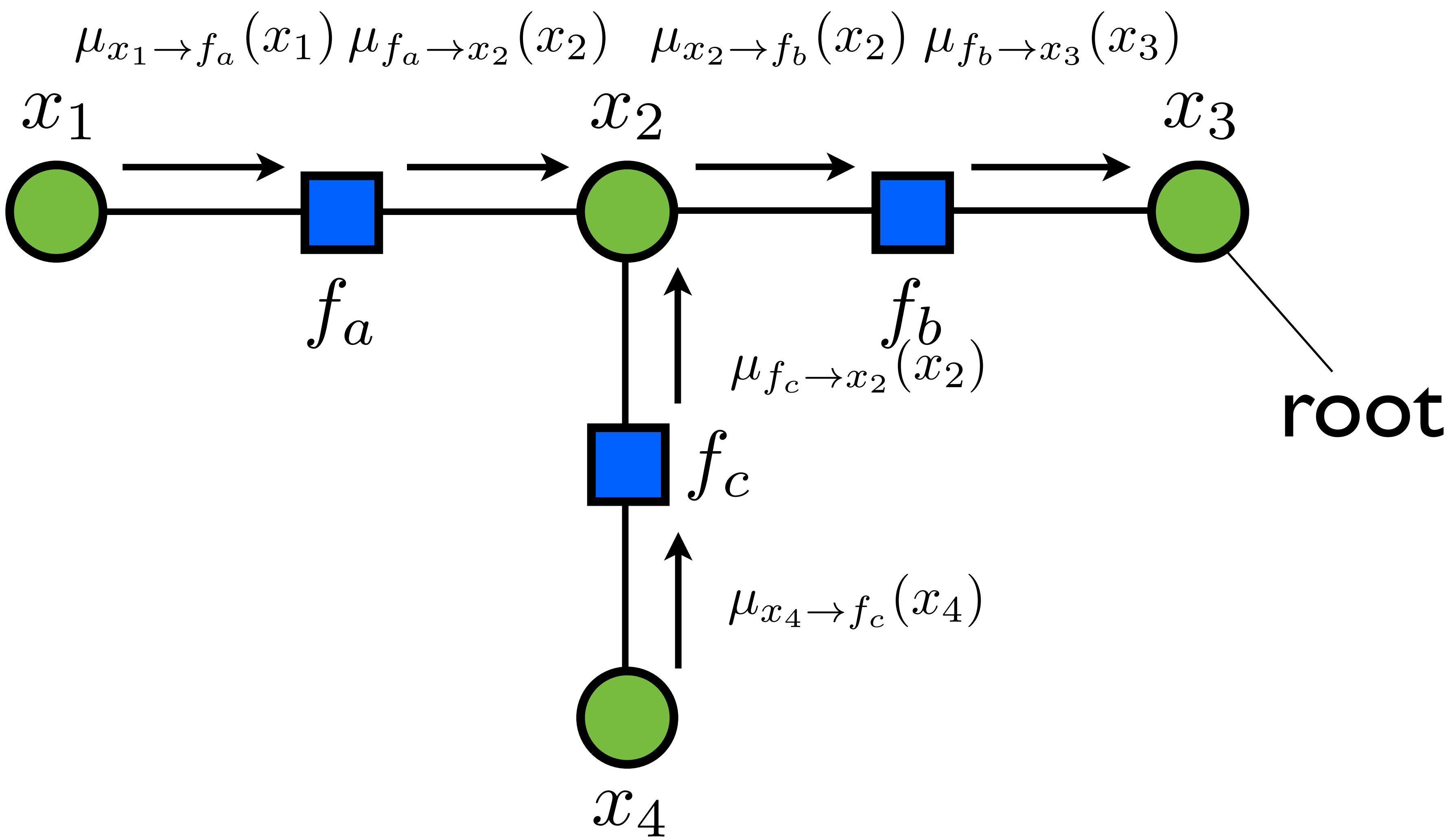
EXAMPLE



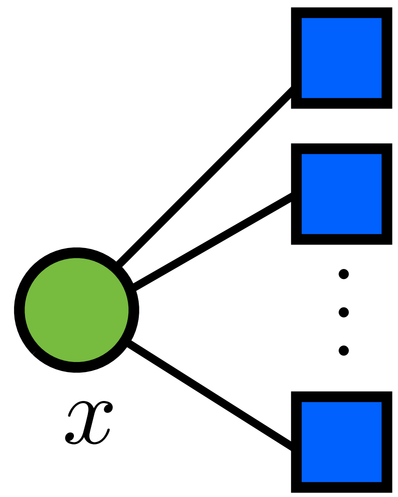
$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4)$$

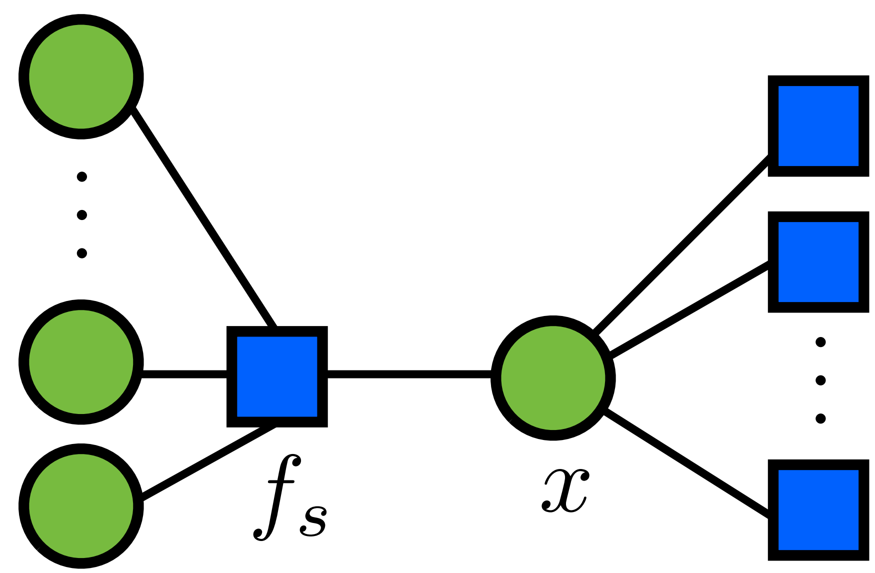
SUM-PRODUCT ALGORITHM

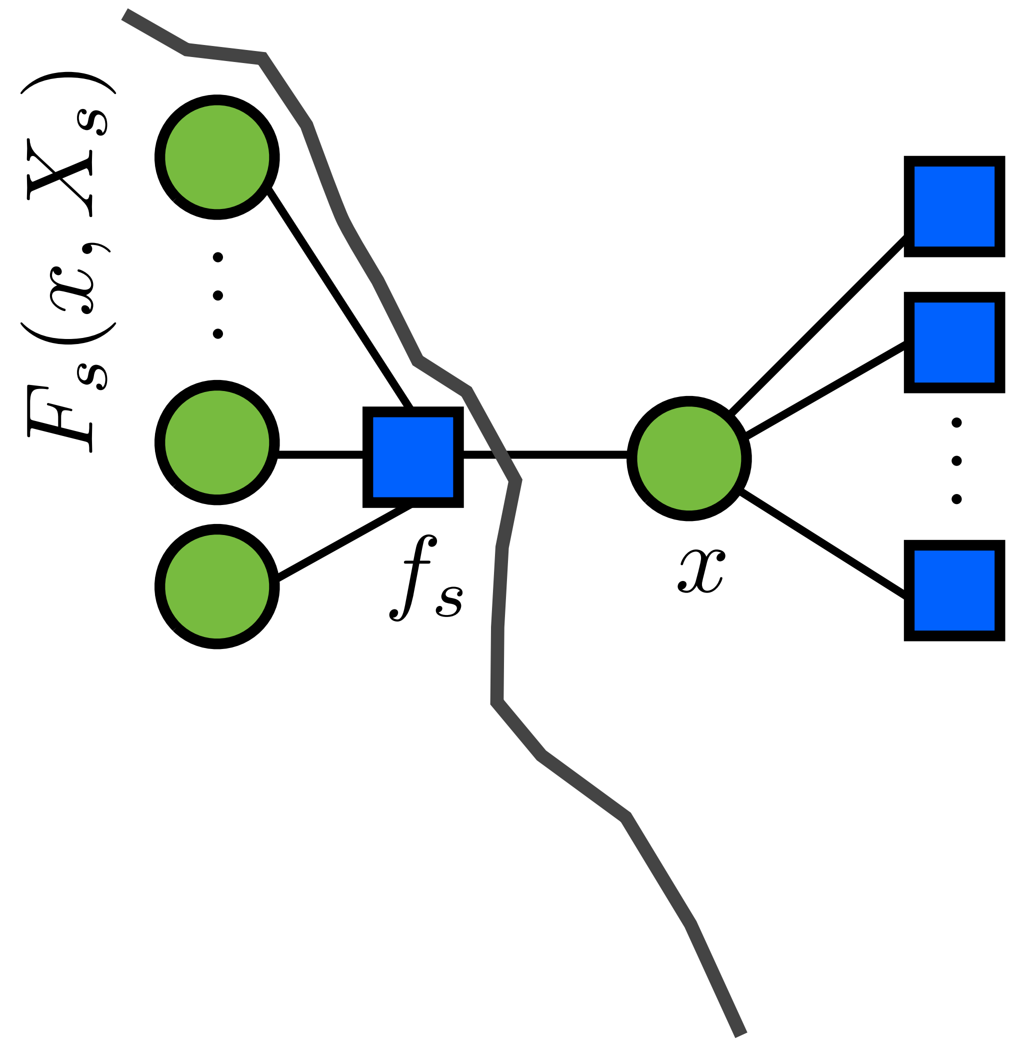
EXAMPLE

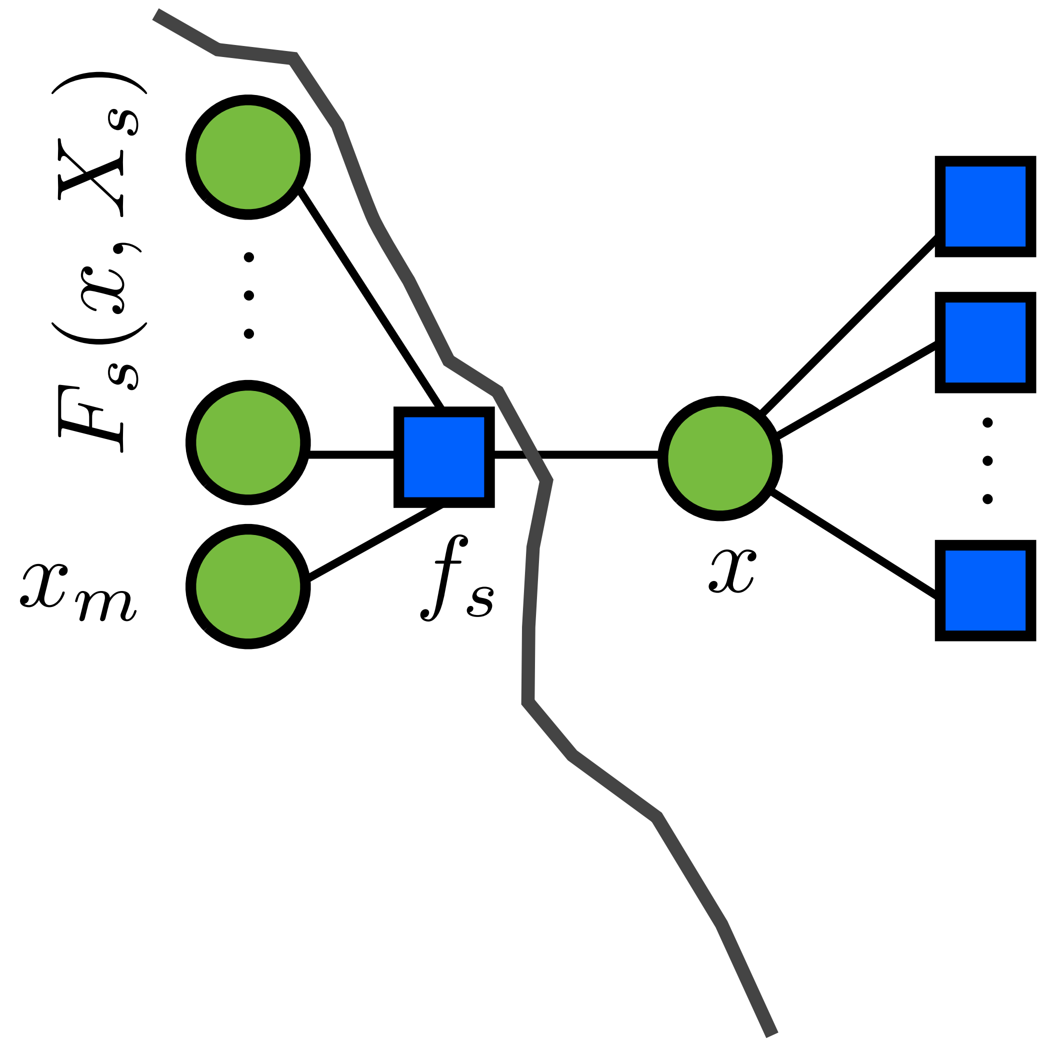


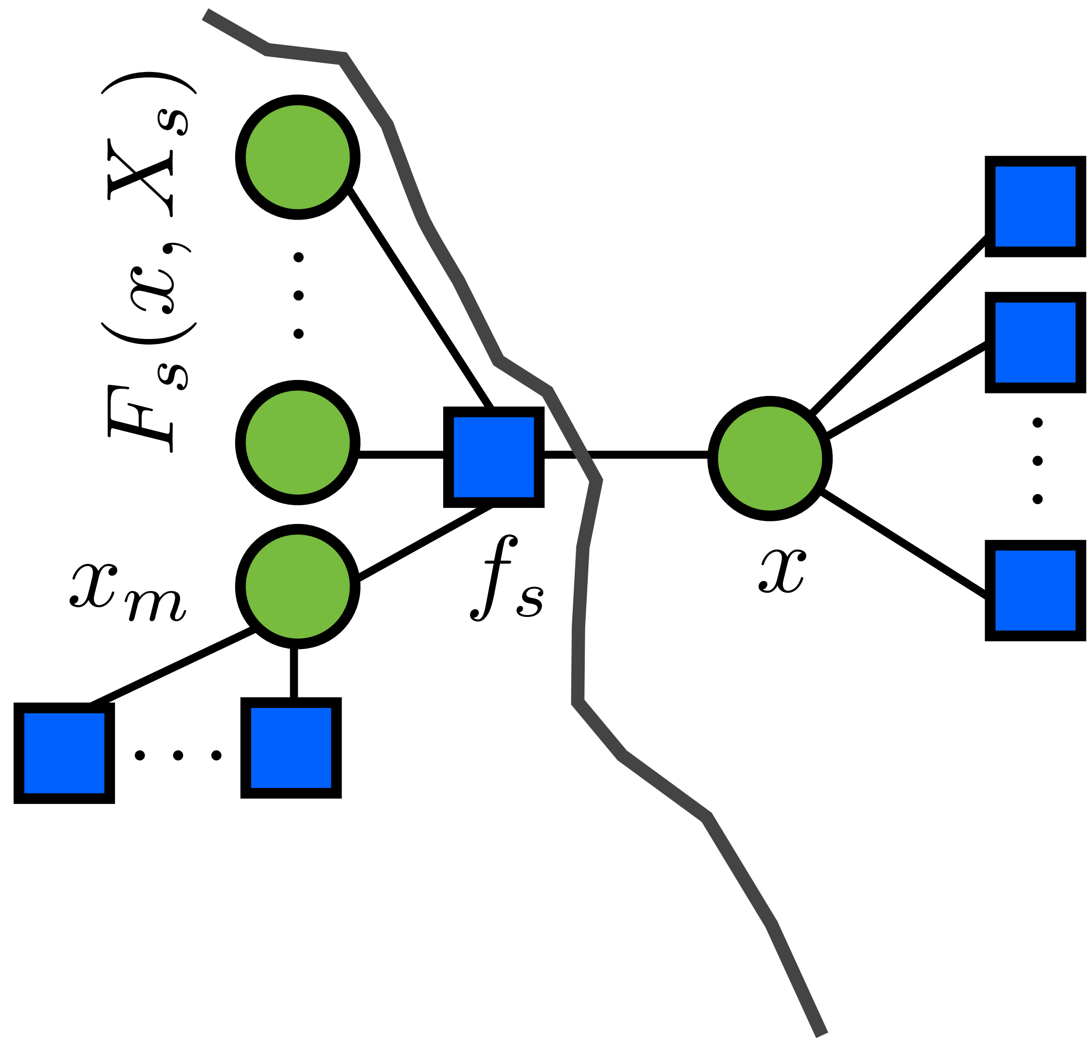
$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4)$$

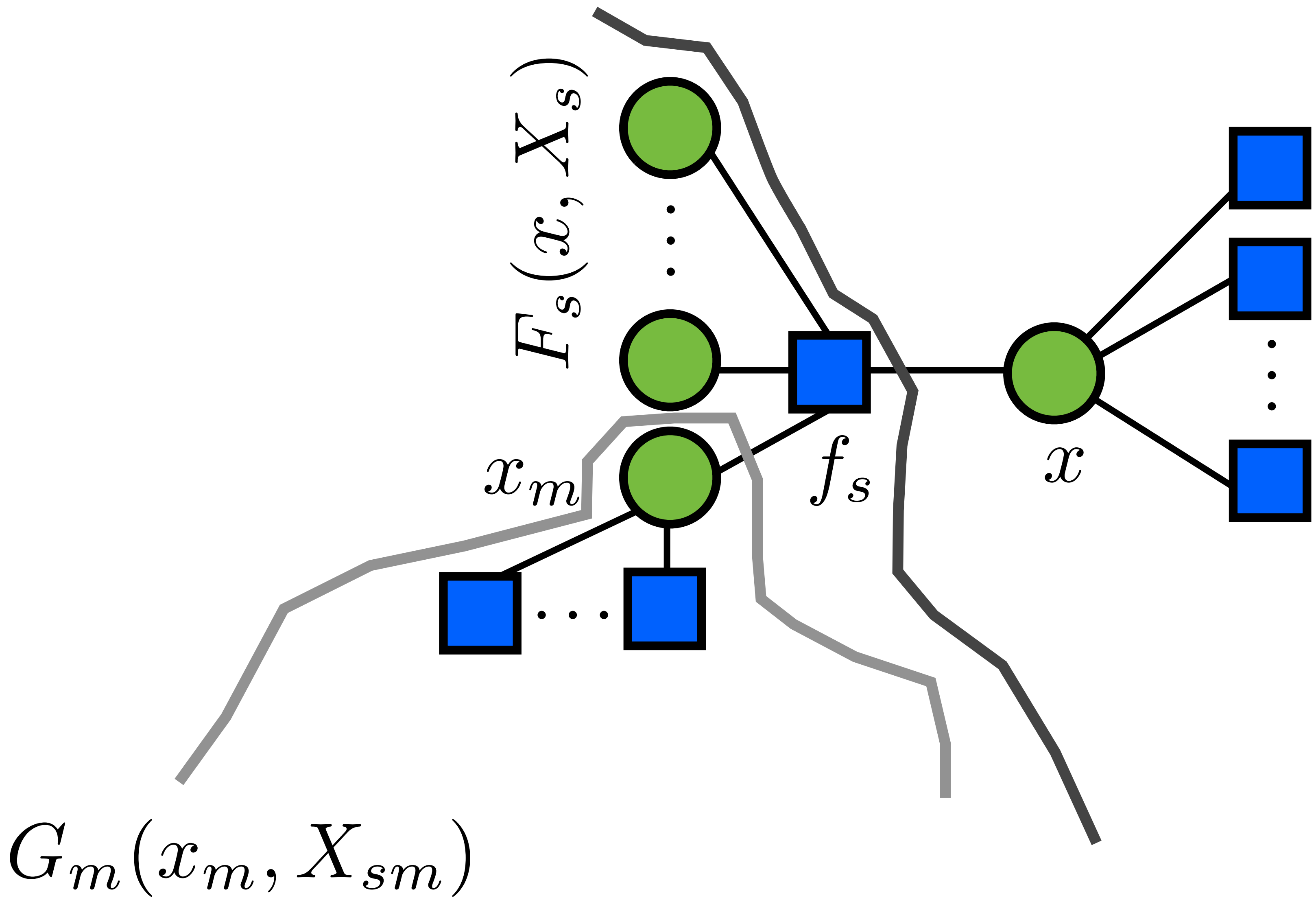

x

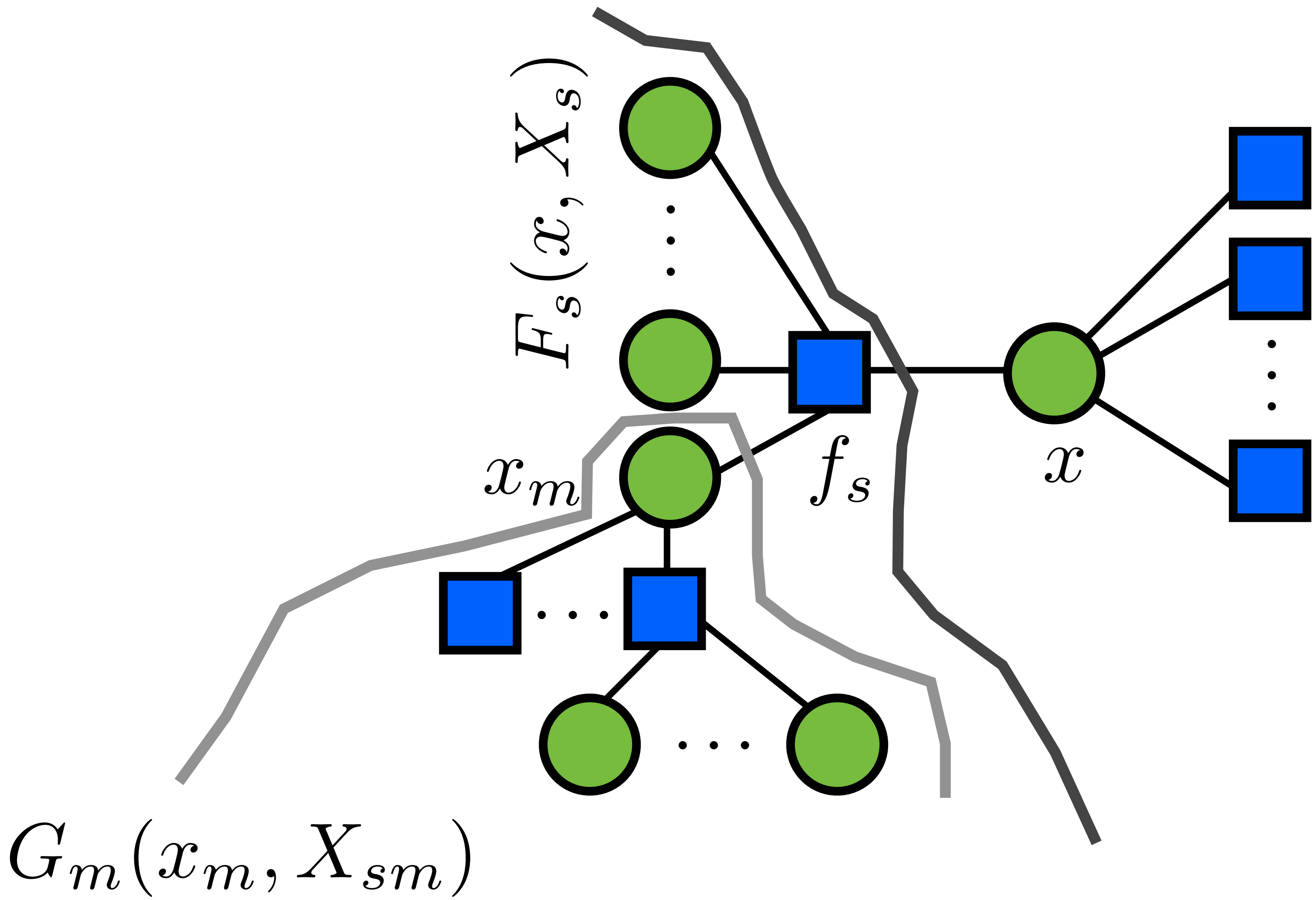

f_s
x


$F_s(x, X_s)$
f_s
x

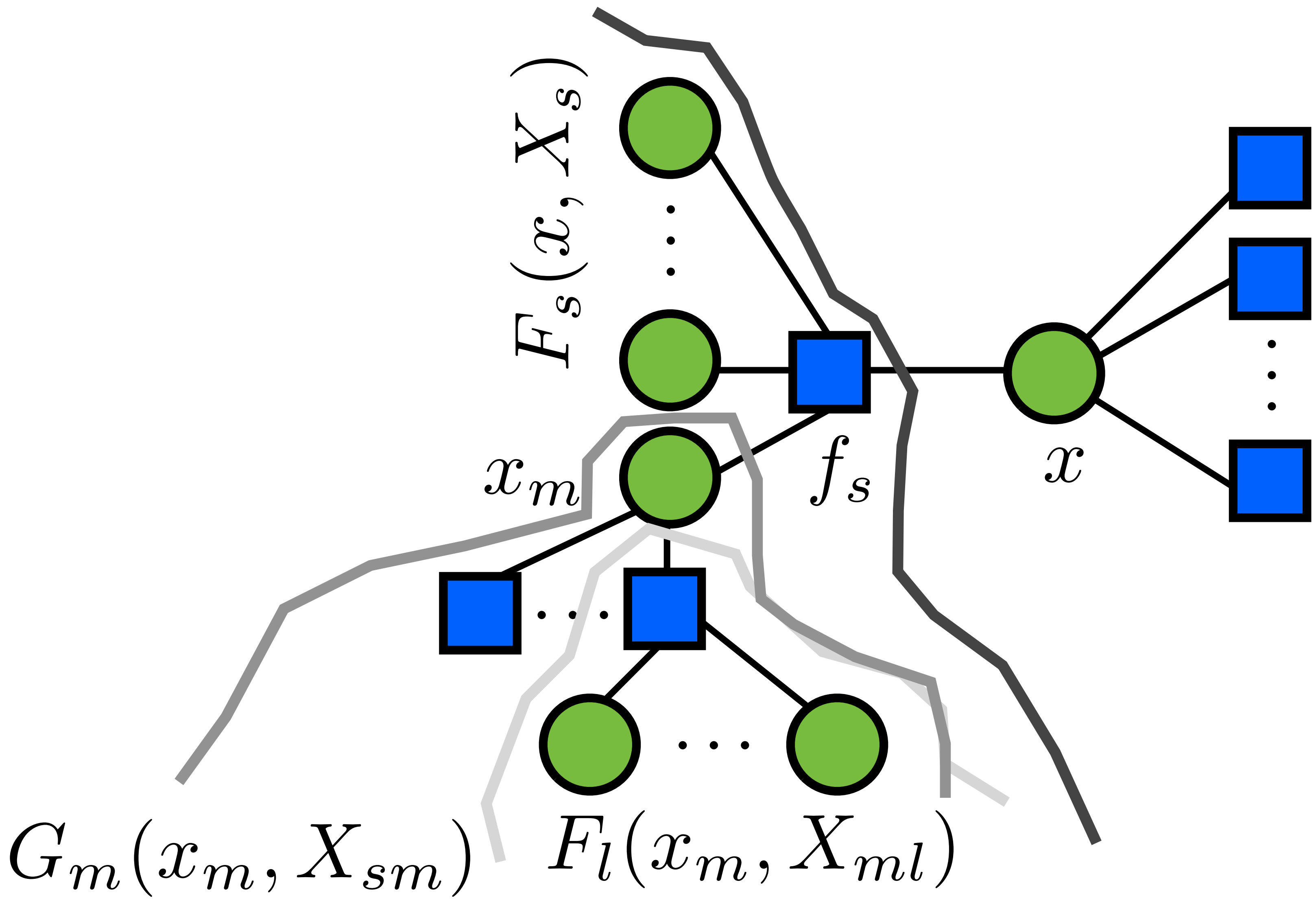

$F_s(x, X_s)$
x_m
f_s
x


$F_s(x, X_s)$
x_m
f_s
x


$F_s(x, X_s)$
x_m
f_s
x
$G_m(x_m, X_{sm})$



$F_s(x, X_s)$
x_m
f_s
x
$G_m(x_m, X_{sm})$


$F_s(x, X_s)$
x_m
f_s
x
$G_m(x_m, X_{sm})$
$F_l(x_m, X_{ml})$

SUM-PRODUCT ALGORITHM

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

SUM-PRODUCT ALGORITHM

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

$$\mathbf{x} = \{x, x_1, x_2, \cdots, x_N\}$$

SUM-PRODUCT ALGORITHM

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

$$\mathbf{x} = \{x, x_1, x_2, \cdots, x_N\}$$

$\mathbf{x}\setminus x$: all of **x** except *x*

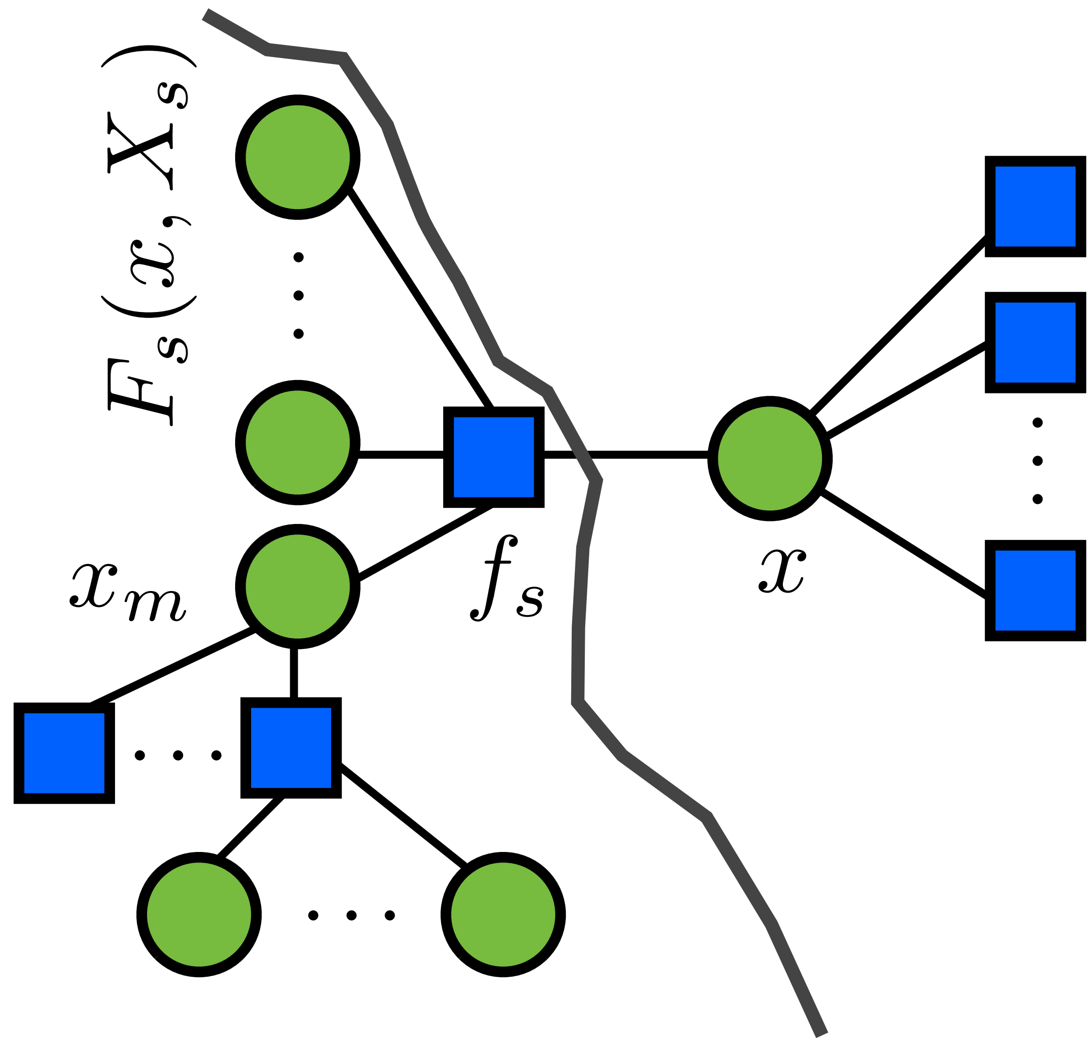
SUM-PRODUCT ALGORITHM

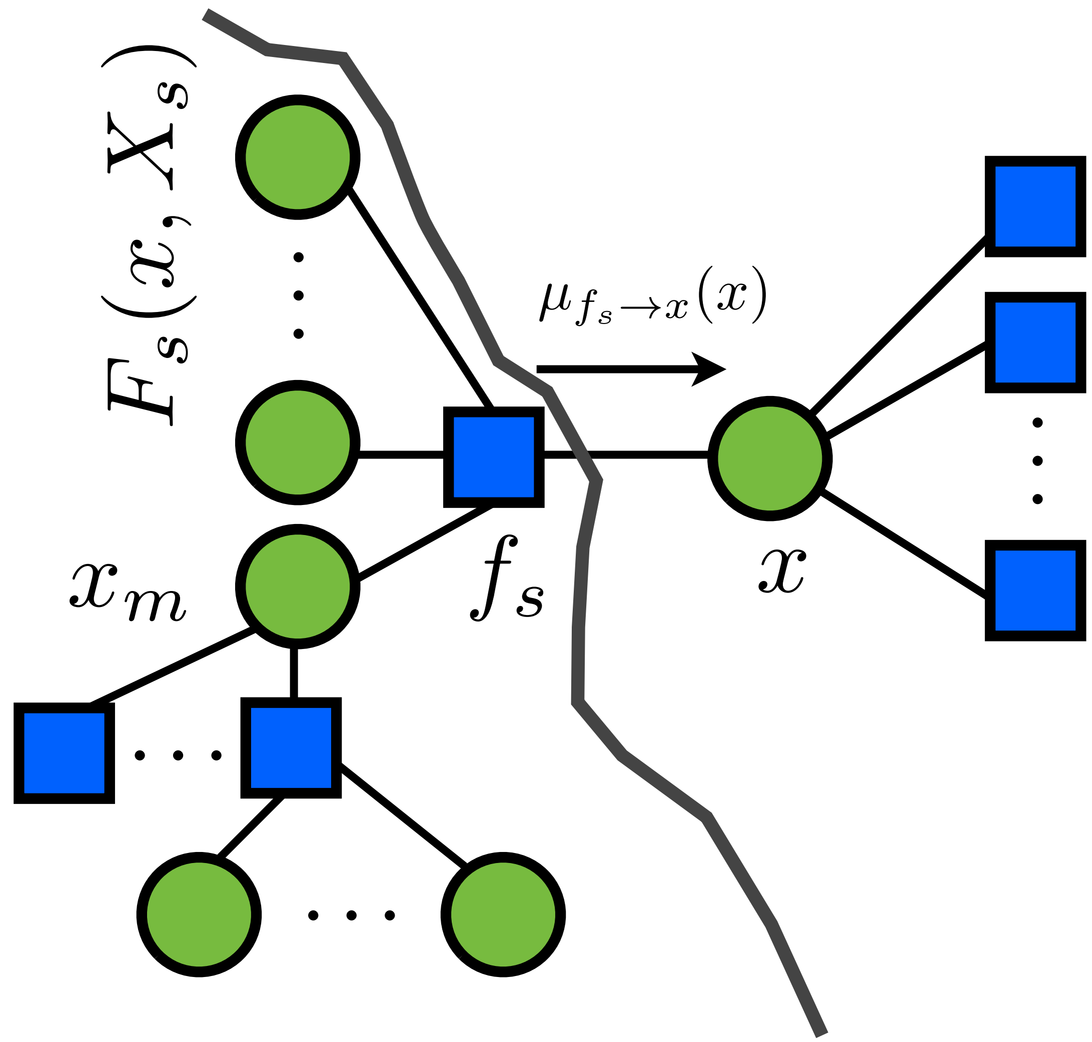
$$p(x) = \sum_{\mathbf{x}\backslash x} p(\mathbf{x})$$

$$\mathbf{x} = \{x, x_1, x_2, \cdots, x_N\}$$

$\mathbf{x}\backslash x$: all of **x** except *x*

$$p(x) = \sum_{x_1} \sum_{x_2} p(x, x_1, x_2) = \sum_{\mathbf{x}\backslash x} p(x, x_1, x_2)$$


$F_s(x, X_s)$
x_m
f_s
x


$F_s(x, X_s)$
$\mu_{f_s \to x}(x)$
x_m
f_s
x

SUM-PRODUCT ALGORITHM

FACTOR TO VARIABLE

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

SUM-PRODUCT ALGORITHM

FACTOR TO VARIABLE

$$p(x) = \sum_{\mathbf{x} \setminus x} p(\mathbf{x}) \qquad p(\mathbf{x}) = \prod_{s \in \mathrm{ne}(x)} F_s(x, X_s)$$

SUM-PRODUCT ALGORITHM

FACTOR TO VARIABLE

$$p(x) = \sum_{\mathbf{x} \setminus x} p(\mathbf{x}) \qquad p(\mathbf{x}) = \prod_{s \in \mathrm{ne}(x)} F_s(x, X_s)$$

$\mathrm{ne}(x)$: neighbors of *x*

SUM-PRODUCT ALGORITHM

FACTOR TO VARIABLE

$$p(x) = \sum_{\mathbf{x} \setminus x} p(\mathbf{x}) \qquad p(\mathbf{x}) = \prod_{s \in \mathrm{ne}(x)} F_s(x, X_s)$$

$\mathrm{ne}(x)$: neighbors of *x*

$s \in \mathrm{ne}(x)$: a particular neighbor of *x*

SUM-PRODUCT ALGORITHM

FACTOR TO VARIABLE

$$p(x) = \sum_{\mathbf{x} \setminus x} p(\mathbf{x}) \qquad p(\mathbf{x}) = \prod_{s \in \mathrm{ne}(x)} F_s(x, X_s)$$

$\mathrm{ne}(x)$: neighbors of *x*

$s \in \mathrm{ne}(x)$: a particular neighbor of *x*

$$p(x) = \prod_{s \in \mathrm{ne}(x)} \Big[\sum_{X_s} F_s(x, X_s) \Big]$$

SUM-PRODUCT ALGORITHM

FACTOR TO VARIABLE

$$p(x) = \sum_{\mathbf{x}\backslash x} p(\mathbf{x}) \qquad p(\mathbf{x}) = \prod_{s \in \mathrm{ne}(x)} F_s(x, X_s)$$

$\mathrm{ne}(x)$: **neighbors of *x***

$s \in \mathrm{ne}(x)$: **a particular neighbor of *x***

$$p(x) = \prod_{s \in \mathrm{ne}(x)} \Big[\sum_{X_s} F_s(x, X_s) \Big]$$

$$= \prod_{s \in \mathrm{ne}(x)} \mu_{f_s \to x}(x)$$

SUM-PRODUCT ALGORITHM

FACTOR TO VARIABLE

$$p(x) = \sum_{\mathbf{x}\setminus x} p(\mathbf{x}) \qquad p(\mathbf{x}) = \prod_{s\in \mathrm{ne}(x)} F_s(x, X_s)$$

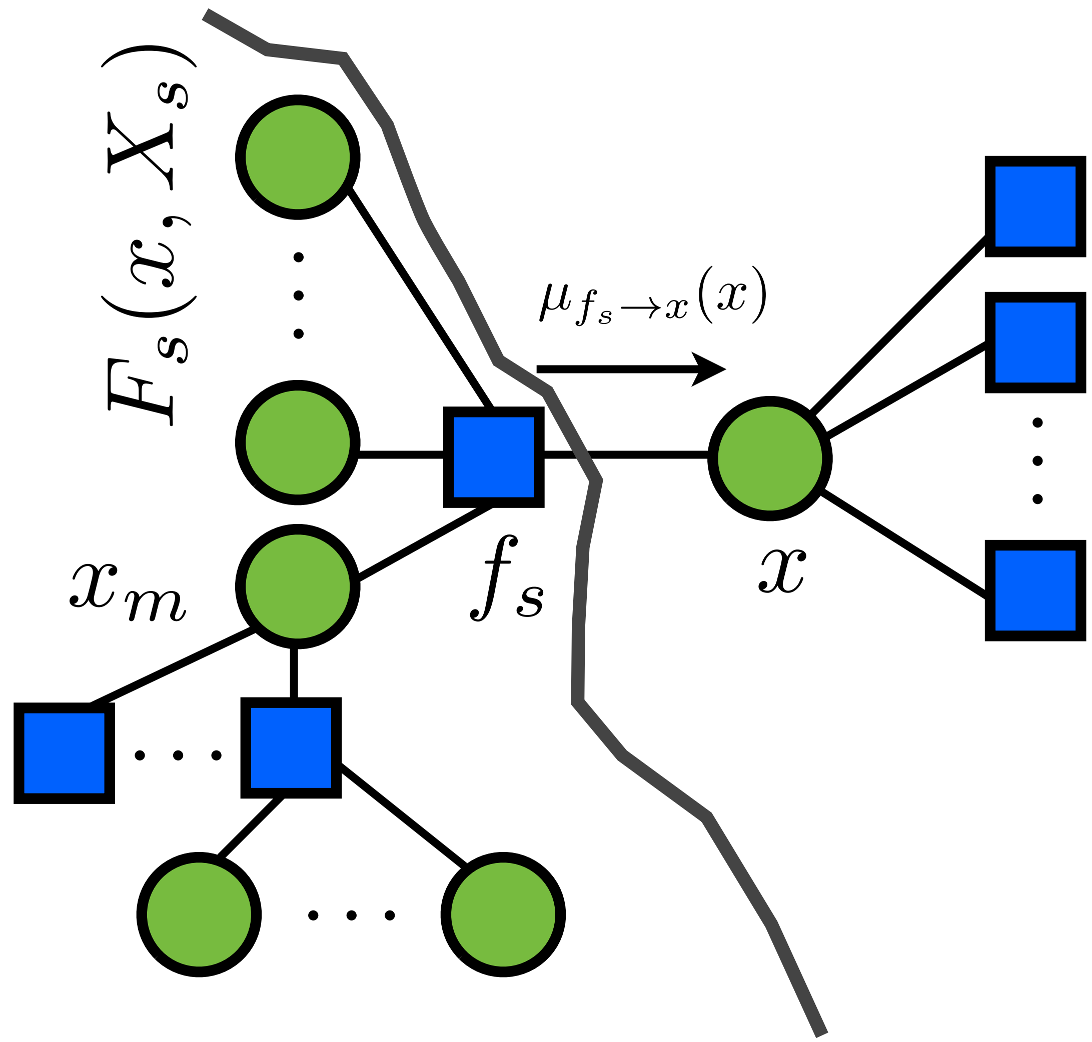
$\mathrm{ne}(x)$: neighbors of *x*

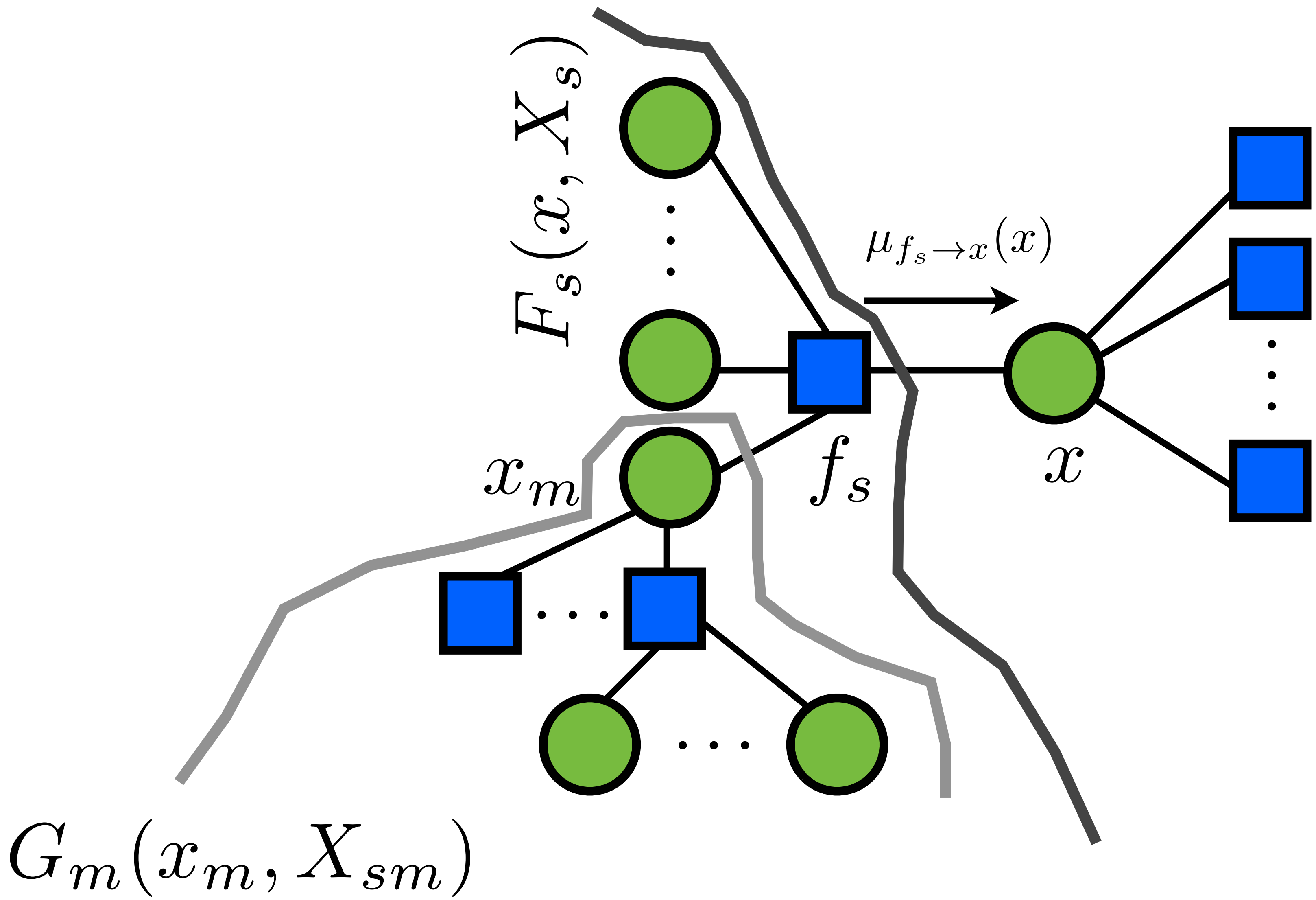
$s \in \mathrm{ne}(x)$: a particular neighbor of *x*

$$p(x) = \prod_{s\in \mathrm{ne}(x)} \left[\sum_{X_s} F_s(x, X_s) \right]$$

$$= \prod_{s\in \mathrm{ne}(x)} \mu_{f_s \to x}(x)$$

$$\boxed{\mu_{f_s \to x}(x) \equiv \sum_{X_s} F_s(x, X_s)}$$


$F_s(x, X_s)$
$\mu_{f_s \to x}(x)$
x_m
f_s
x


$F_s(x, X_s)$
$\mu_{f_s \to x}(x)$
x_m
f_s
x
$G_m(x_m, X_{sm})$

SUM-PRODUCT ALGORITHM

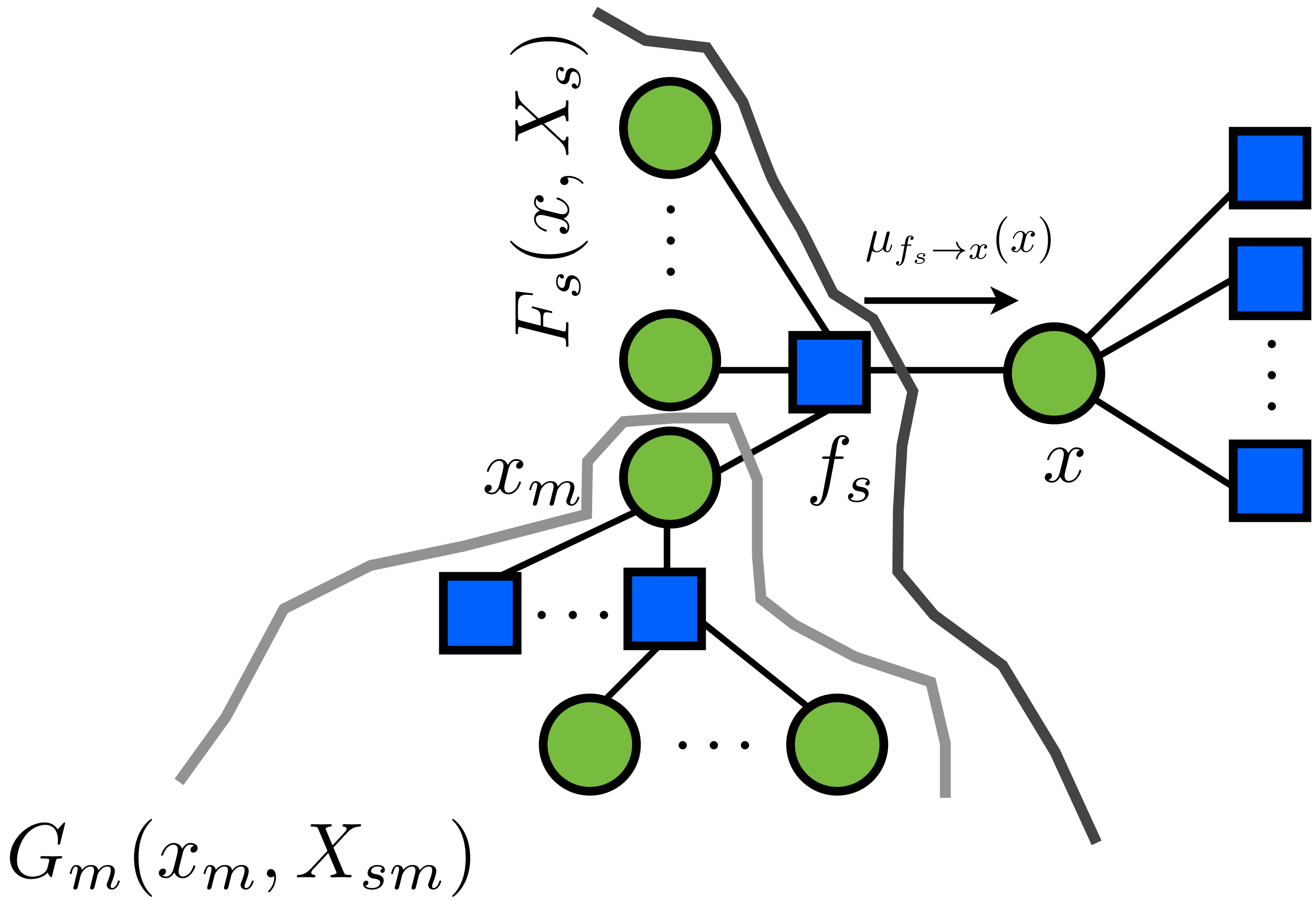
$$\mu_{f_s \to x}(x) \equiv \sum_{X_s} F_s(x, X_s)$$

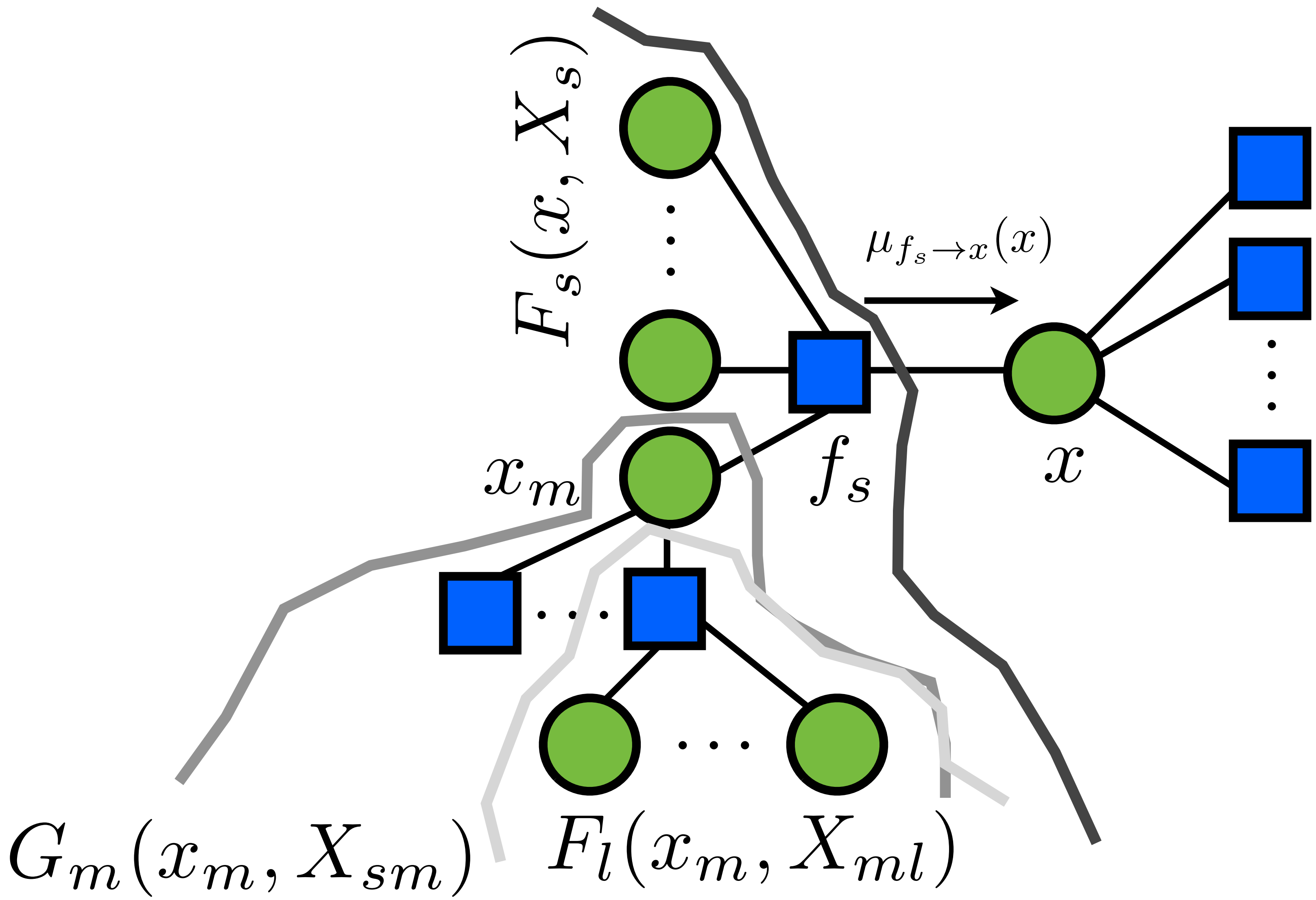
$F_s(x, X_s)$

SUM-PRODUCT ALGORITHM

$$\mu_{f_s \to x}(x) \equiv \sum_{X_s} F_s(x, X_s)$$

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


$F_s(x, X_s)$
$\mu_{f_s \to x}(x)$
x_m
f_s
x
$G_m(x_m, X_{sm})$


$F_s(x, X_s)$
$\mu_{f_s \to x}(x)$
x_m
f_s
x
$G_m(x_m, X_{sm})$
$F_l(x_m, X_{ml})$

SUM-PRODUCT ALGORITHM

VARIABLE TO FACTOR

$$\mu_{f_s \to x}(x) \equiv \sum_{X_s} F_s(x, X_s)$$

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

SUM-PRODUCT ALGORITHM

VARIABLE TO FACTOR

$$\mu_{f_s \to x}(x) \equiv \sum_{X_s} F_s(x, X_s)$$

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

$$\mu_{f_s \to x} = \sum_{X} f_s(x, x_1, \cdots, x_M) \prod_{m \in \mathrm{ne}(f_s) \setminus x} \left[\sum_{X_{sm}} G_m(x_m, X_{sm}) \right]$$

SUM-PRODUCT ALGORITHM

VARIABLE TO FACTOR

$$\mu_{f_s \to x}(x) \equiv \sum_{X_s} F_s(x, X_s)$$

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

$$\mu_{f_s \to x} = \sum_X f_s(x, x_1, \cdots, x_M) \prod_{m \in \mathrm{ne}(f_s) \setminus x} \left[\sum_{X_{sm}} G_m(x_m, X_{sm}) \right]$$

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

SUM-PRODUCT ALGORITHM

VARIABLE TO FACTOR

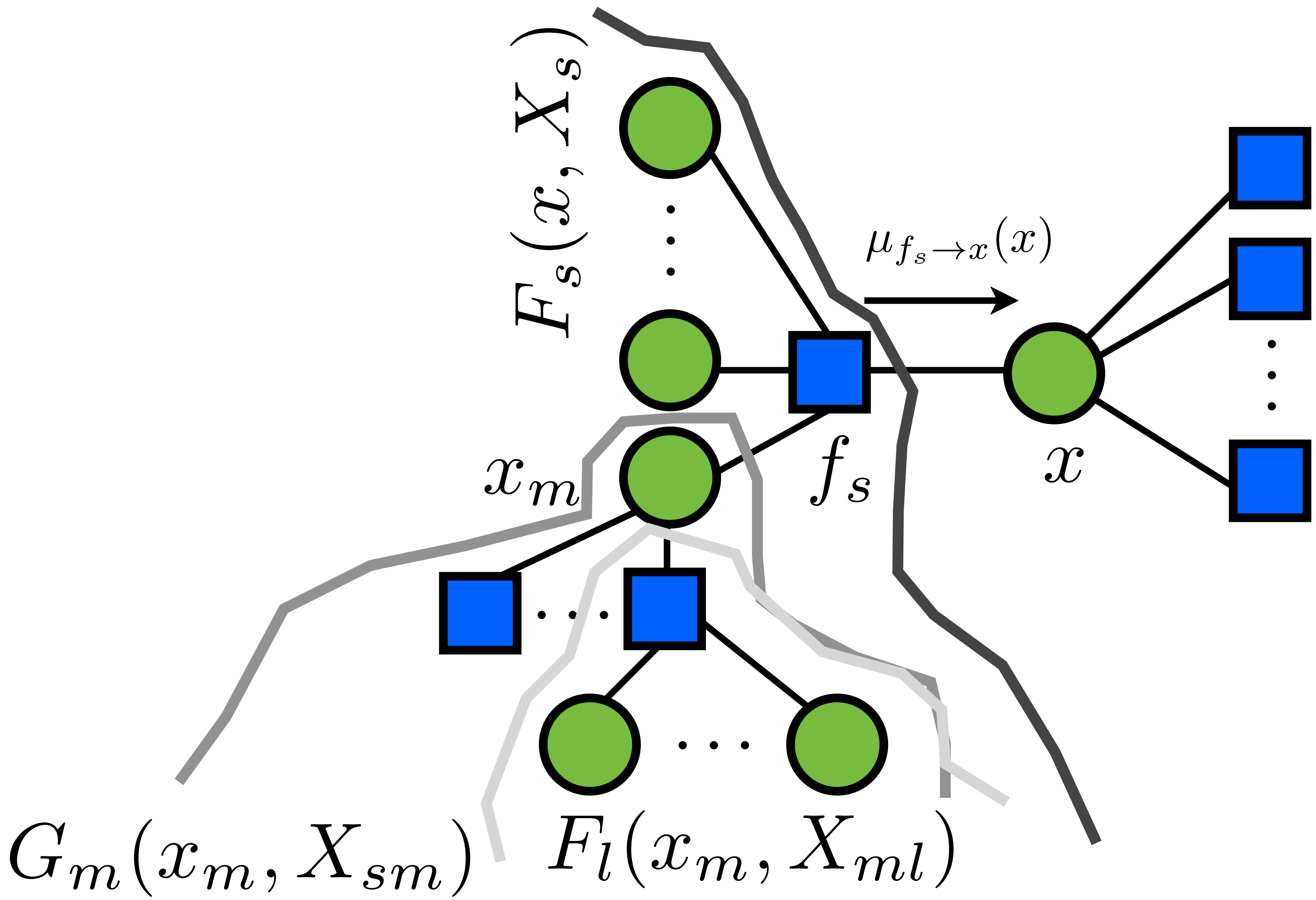
$$\mu_{f_s \to x}(x) \equiv \sum_{X_s} F_s(x, X_s)$$

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

$$\mu_{f_s \to x} = \sum_{X} f_s(x, x_1, \cdots, x_M) \prod_{m \in \mathrm{ne}(f_s) \backslash x} \left[\sum_{X_{sm}} G_m(x_m, X_{sm}) \right]$$

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

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


$F_s(x, X_s)$
$\mu_{f_s \to x}(x)$
x_m
f_s
x
$G_m(x_m, X_{sm})$
$F_l(x_m, X_{ml})$

SUM-PRODUCT ALGORITHM

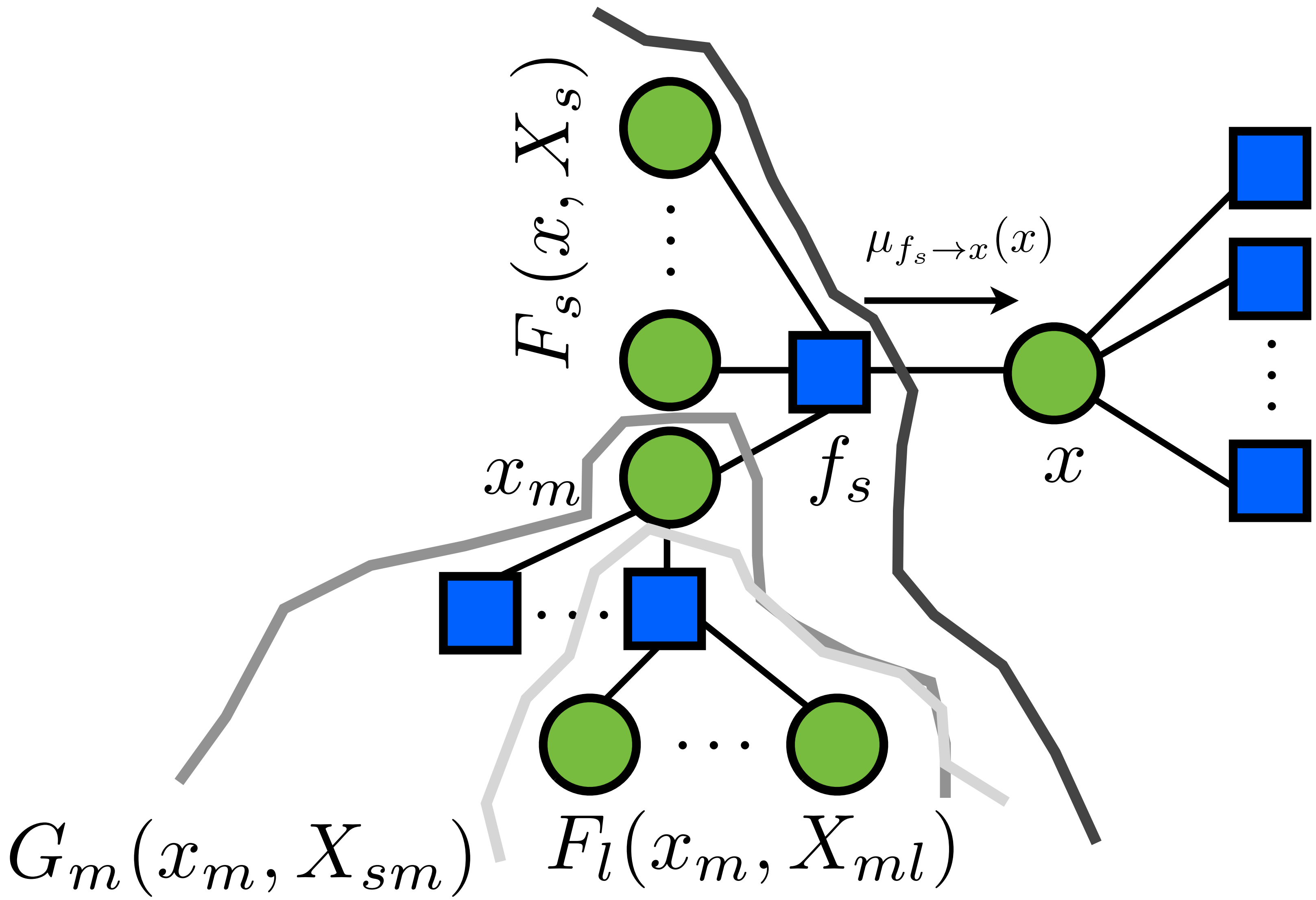
$$\mu_{f_s \to x}(x) \equiv \sum_{X_s} F_s(x, X_s)$$

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

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

$$\mu_{x_m \to f_s(x_m)} = \prod_{l = \mathrm{ne} \in (x_m) \setminus f_s} \left[\sum_{X_{ml}} F_l(x_m, X_{ml}) \right]$$

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



$F_s(x, X_s)$
$\mu_{f_s \to x}(x)$
x_m
f_s
x
$G_m(x_m, X_{sm})$
$F_l(x_m, X_{ml})$

SUM-PRODUCT ALGORITHM

LEAF NODES

$$\mu_{f_s \to x}(x) \equiv \sum_{X_s} F_s(x, X_s)$$

SUM-PRODUCT ALGORITHM

LEAF NODES

$$\mu_{f_s \to x}(x) \equiv \sum_{X_s} F_s(x, X_s)$$

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

SUM-PRODUCT ALGORITHM

LEAF NODES

$$\mu_{f_s \to x}(x) \equiv \sum_{X_s} F_s(x, X_s)$$

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

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

SUM-PRODUCT ALGORITHM

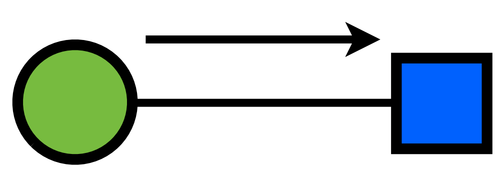
LEAF NODES

$$\mu_{f_s \to x}(x) \equiv \sum_{X_s} F_s(x, X_s)$$

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

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

$$\mu_{x \to f}(x) = 1$$



SUM-PRODUCT ALGORITHM

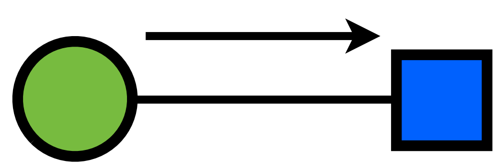
LEAF NODES

$$\mu_{f_s \to x}(x) \equiv \sum_{X_s} F_s(x, X_s)$$

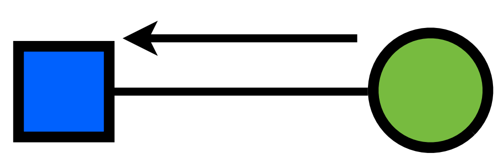
$$\mu_{f_s \to x} = \sum_{X} f_s(x, x_1, \cdots, x_M) \prod_{m \in \mathrm{ne}(f_s) \setminus x} \mu_{x_m \to f_s}(x_m)$$

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

$$\mu_{x \to f}(x) = 1$$



$$\mu_{f \to x}(x) = f(x)$$



SUM-PRODUCT ALGORITHM

ALGORITHM

Task: Evaluate p(*x*)

View *x* as root of factor graph

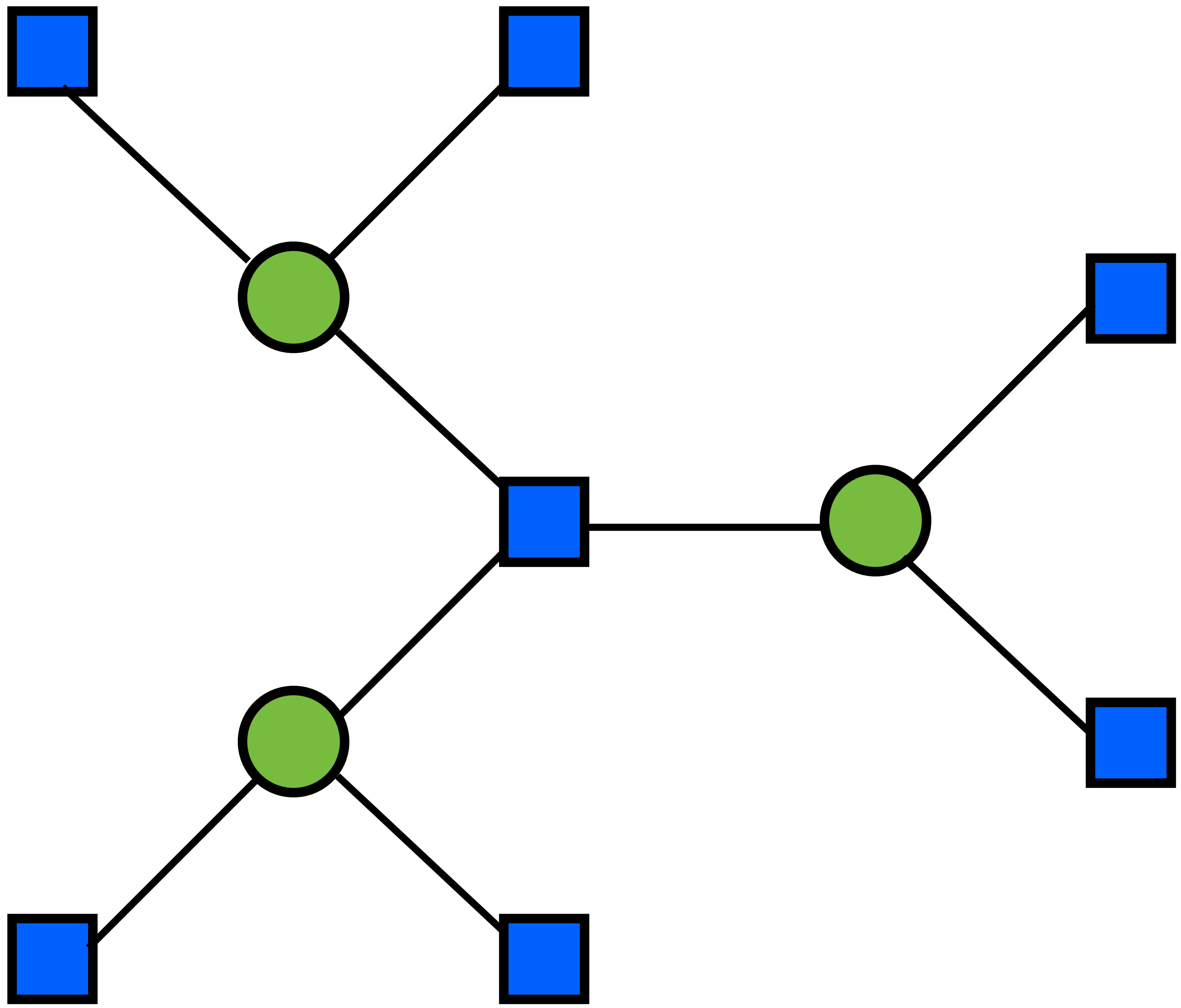
Initiate message at the leaves of the graph

Recursively pass messages until root has received message from all neighbors

Evaluate the marginal

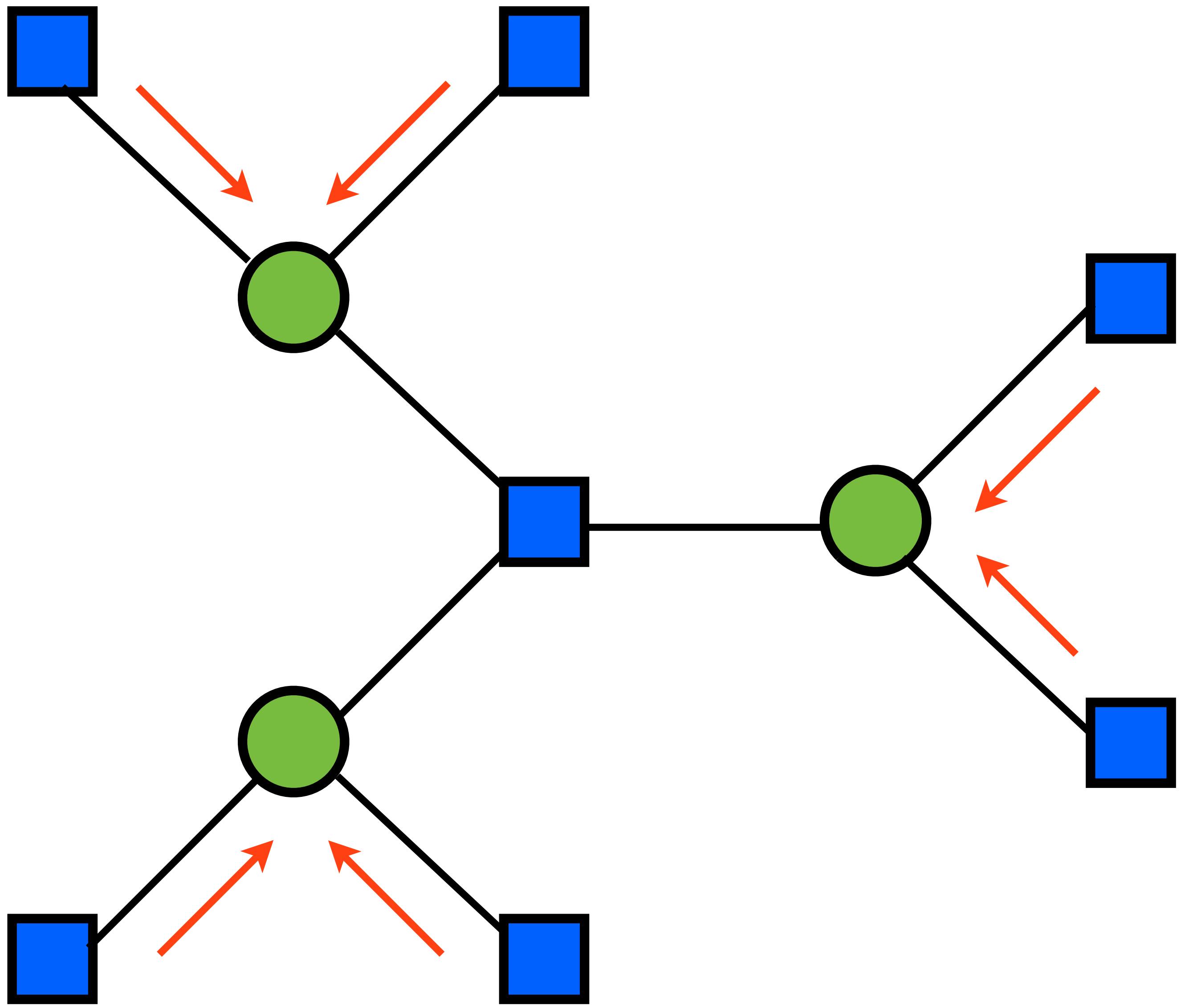
SUM-PRODUCT ALGORITHM

MARGINAL FOR EVERY NODE



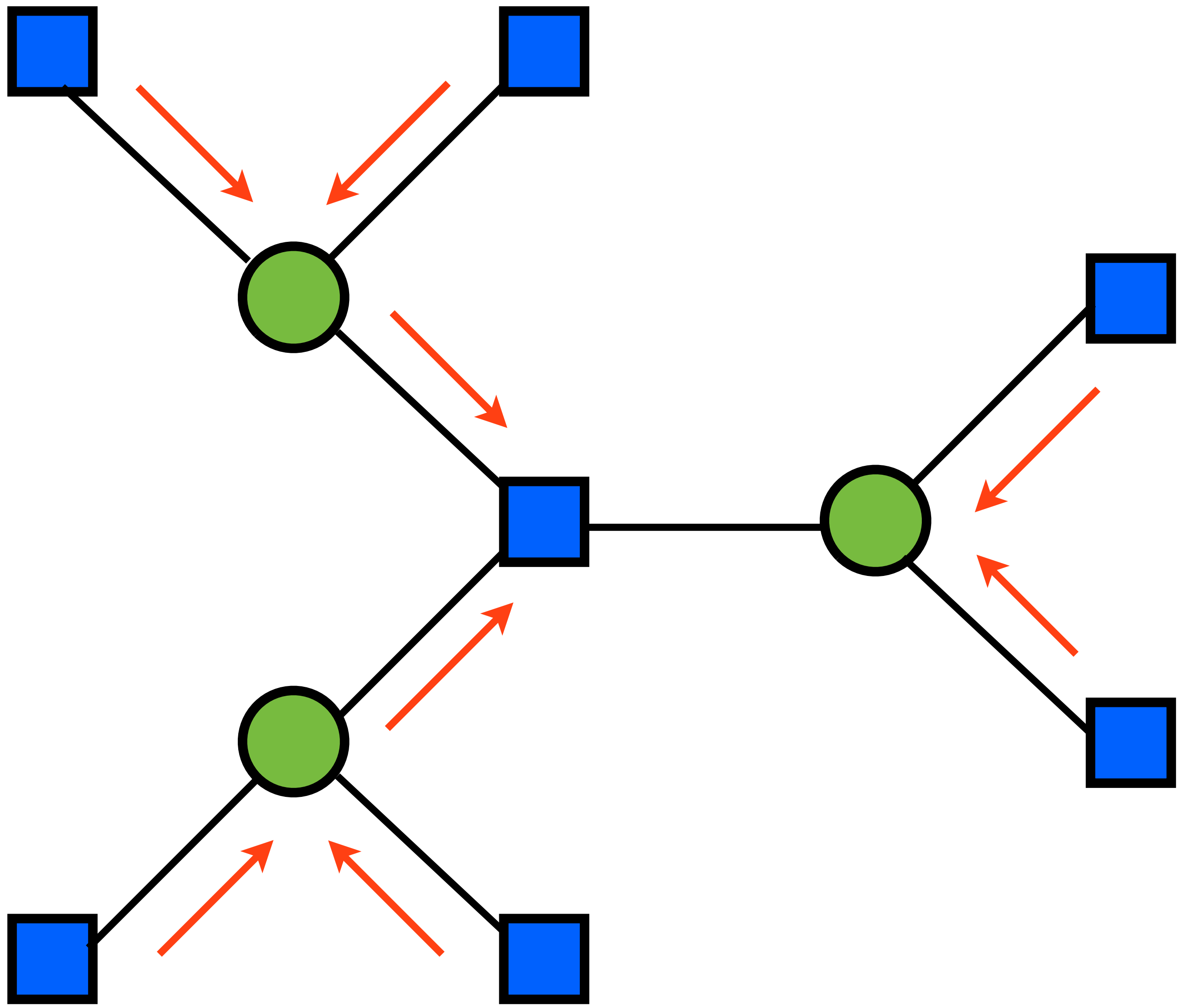
SUM-PRODUCT ALGORITHM

MARGINAL FOR EVERY NODE



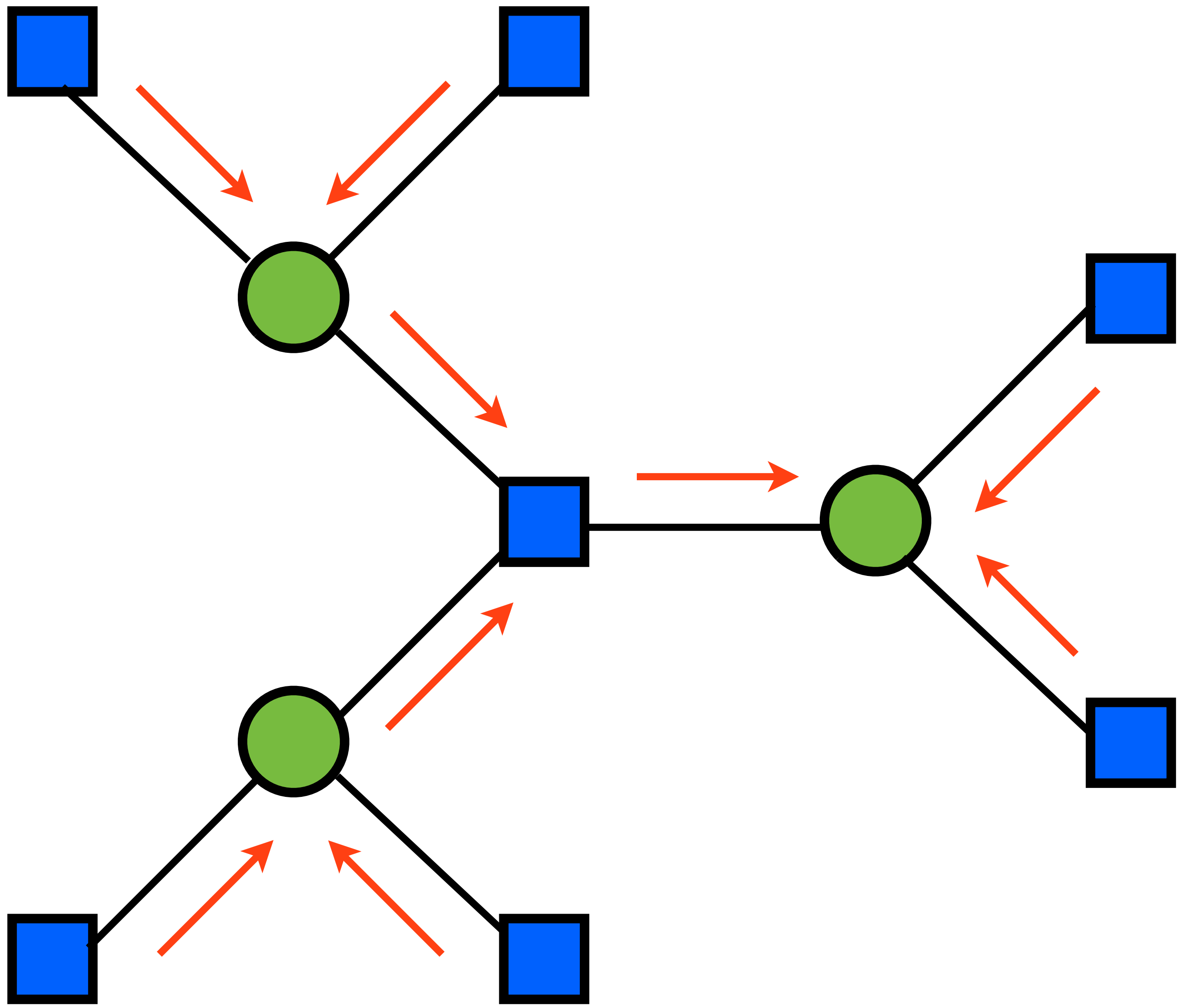
SUM-PRODUCT ALGORITHM

MARGINAL FOR EVERY NODE



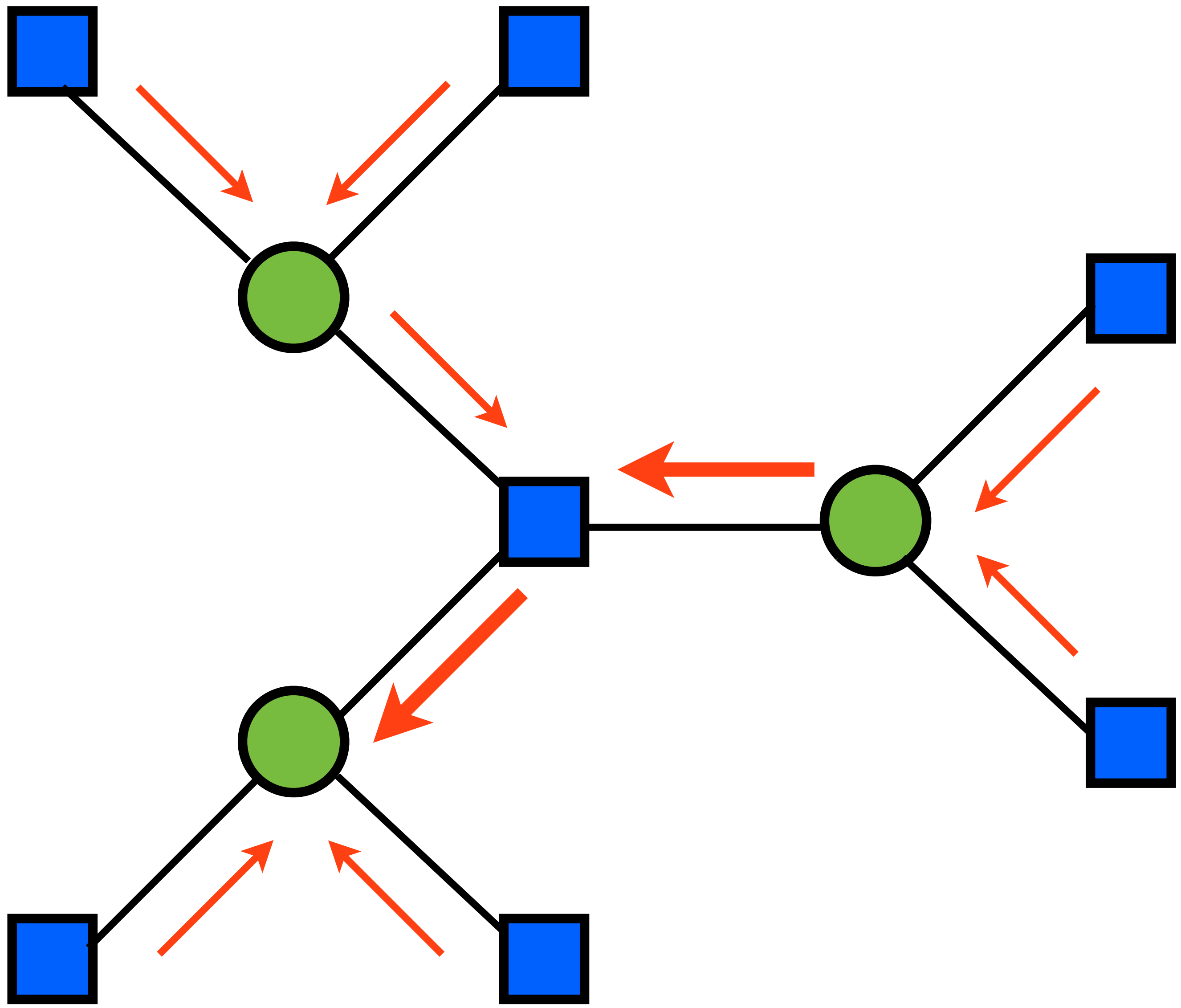
SUM-PRODUCT ALGORITHM

MARGINAL FOR EVERY NODE



SUM-PRODUCT ALGORITHM

MARGINAL FOR EVERY NODE



SUM-PRODUCT ALGORITHM

Task: Efficiently compute marginals for all *x*

Pick any *x* as root of factor graph

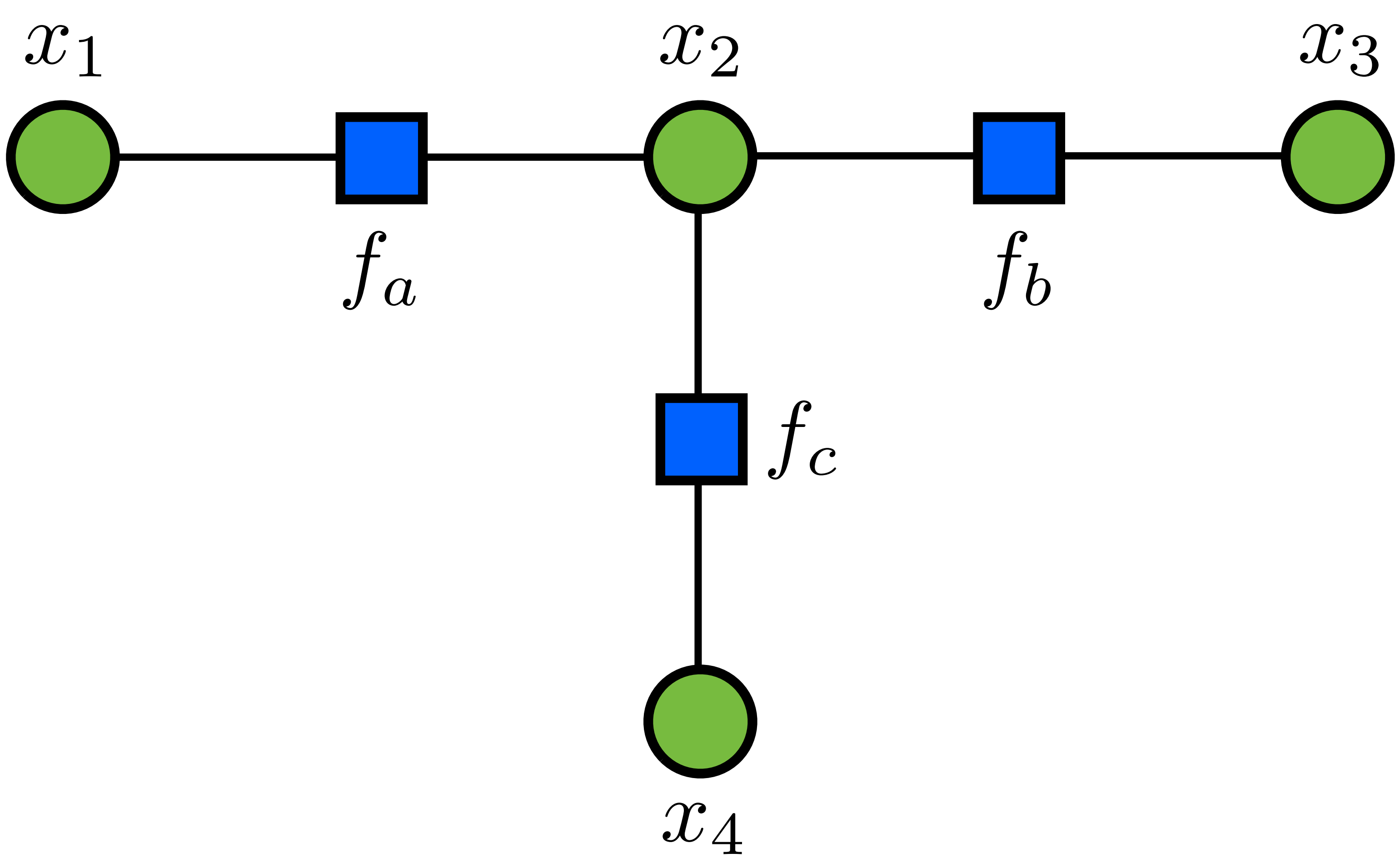
Send message from leaves to root

Send message back from root to leaves

Calculate marginal distribution for all *x*

SUM-PRODUCT ALGORITHM

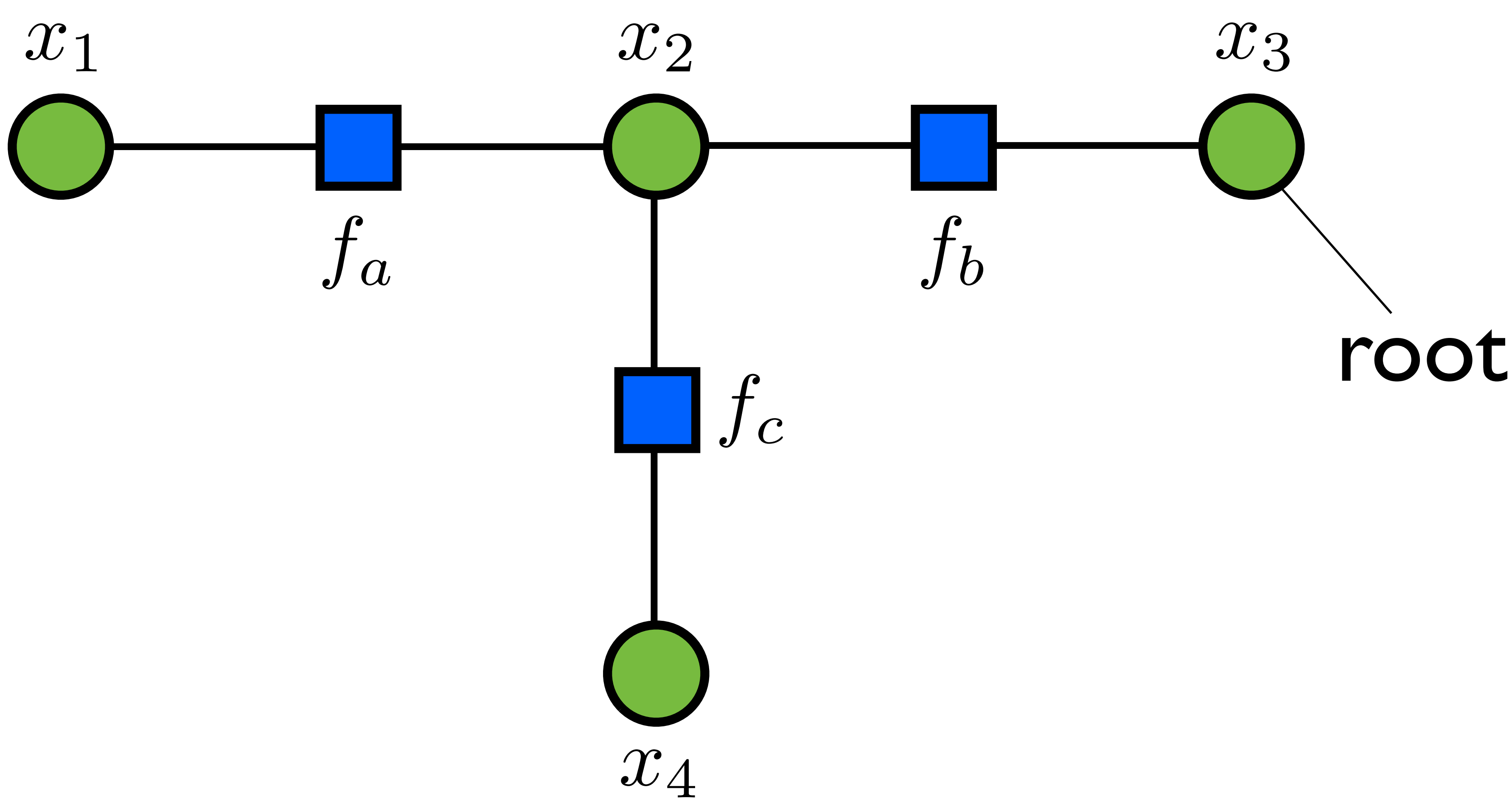
EXAMPLE



$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4)$$

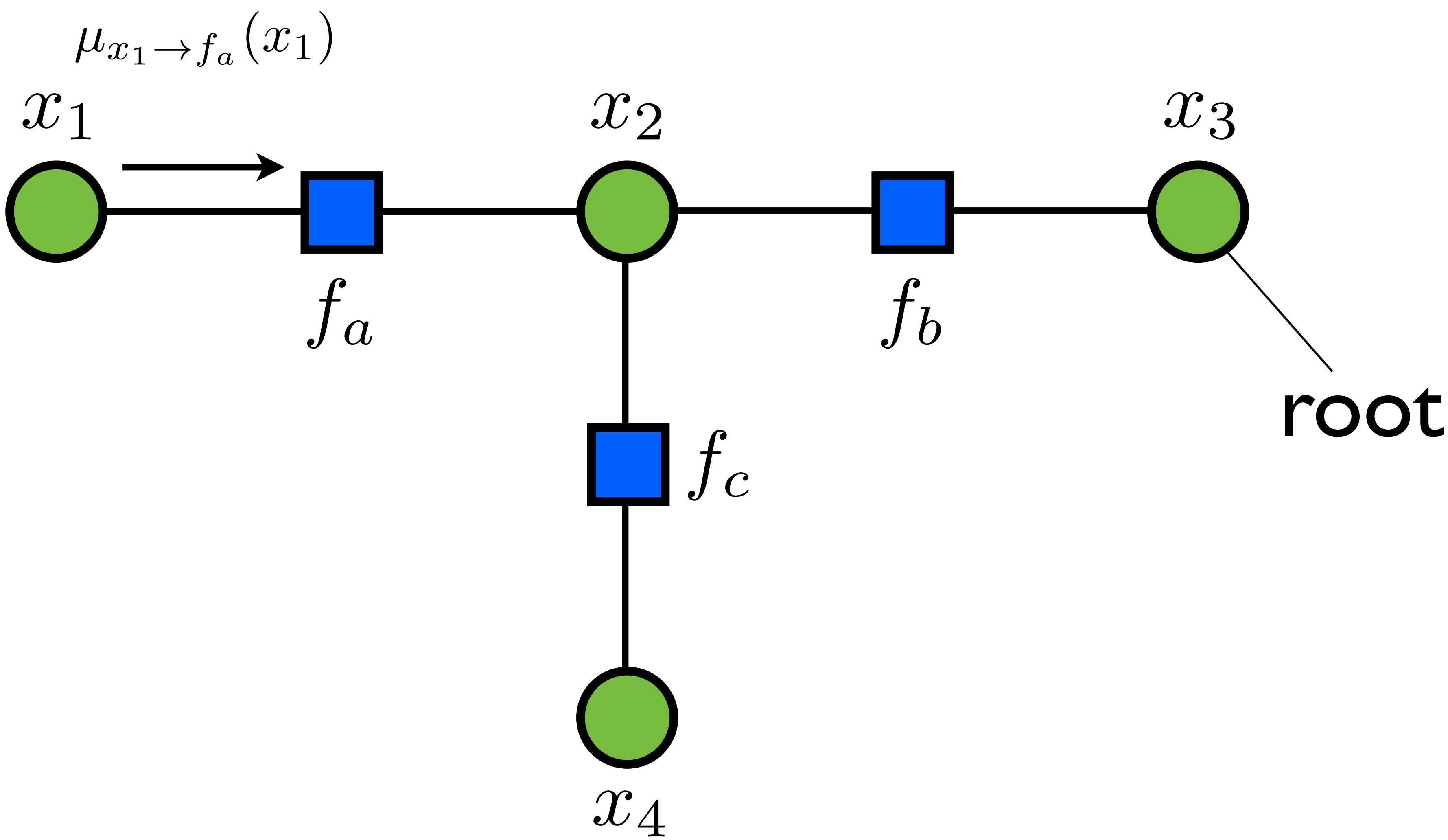
SUM-PRODUCT ALGORITHM

EXAMPLE



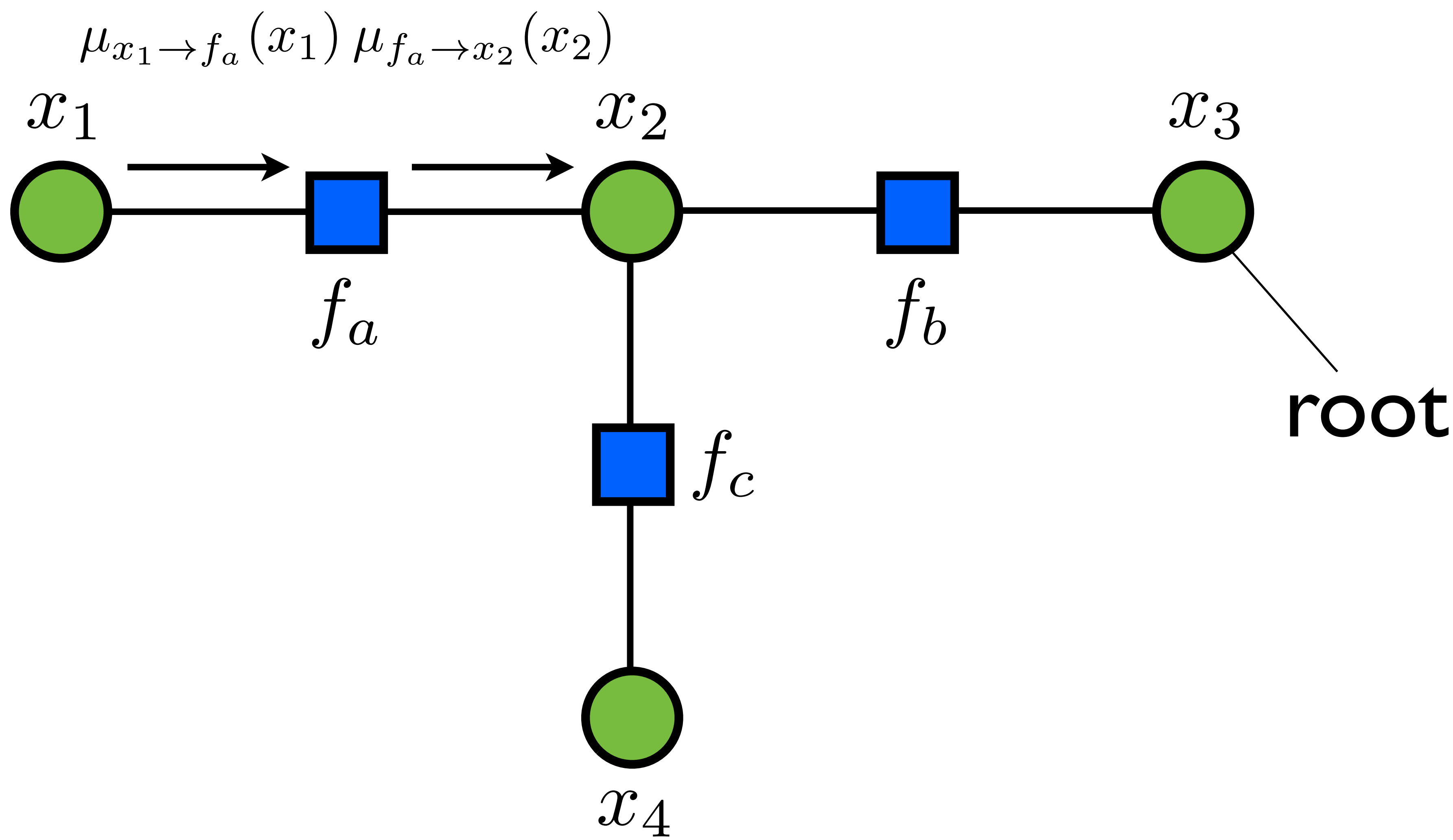
SUM-PRODUCT ALGORITHM

EXAMPLE



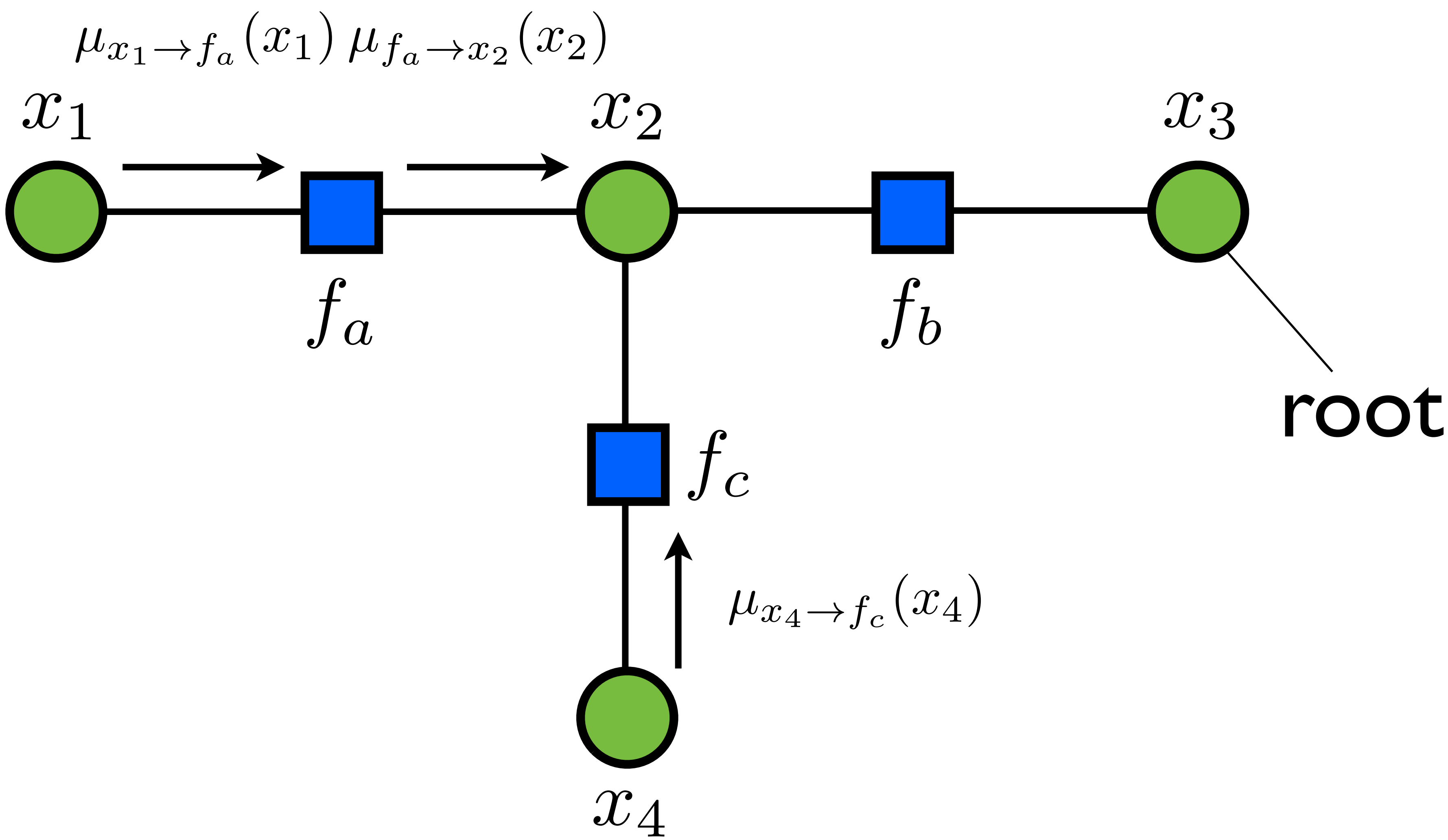
SUM-PRODUCT ALGORITHM

EXAMPLE



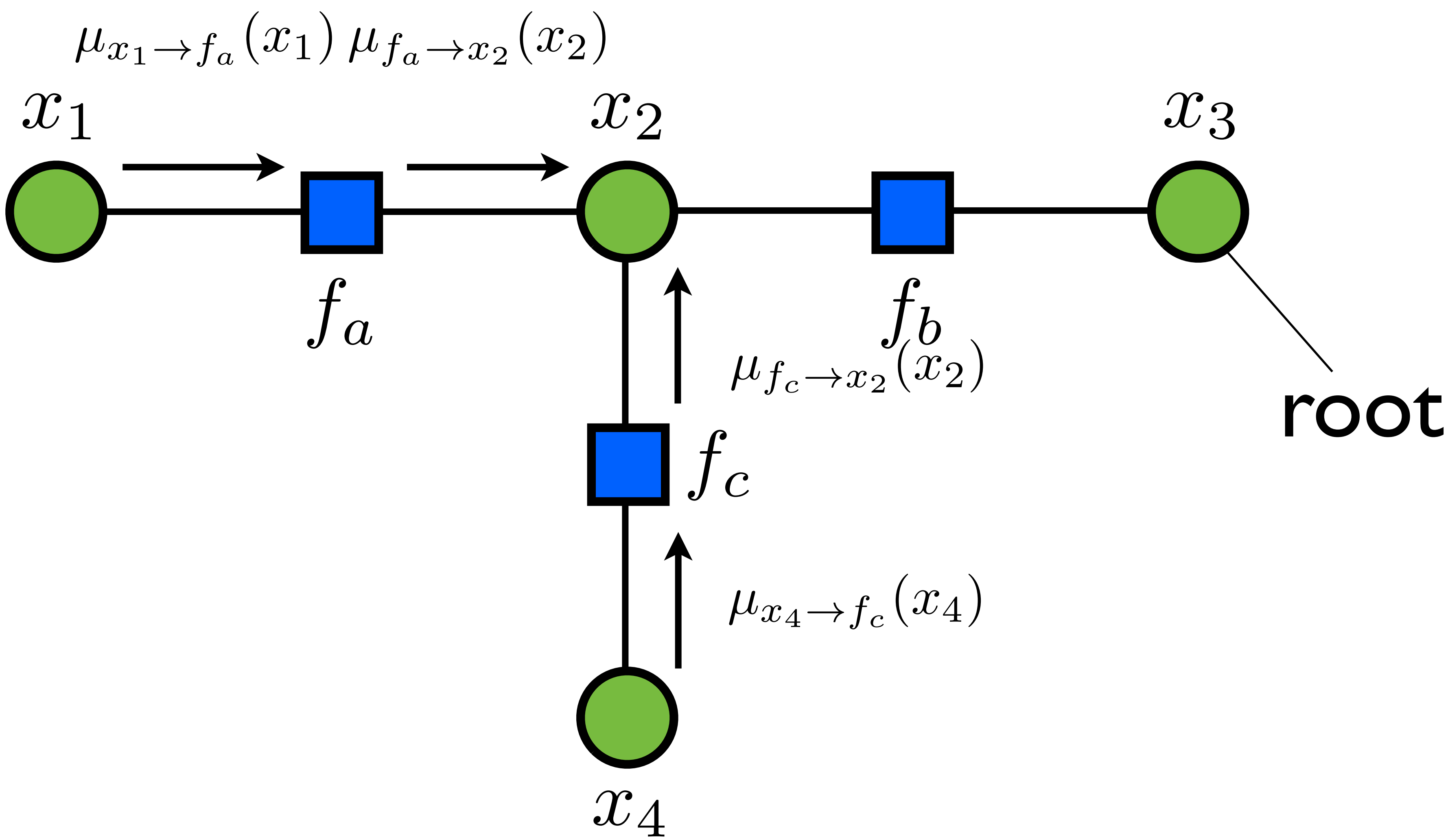
SUM-PRODUCT ALGORITHM

EXAMPLE



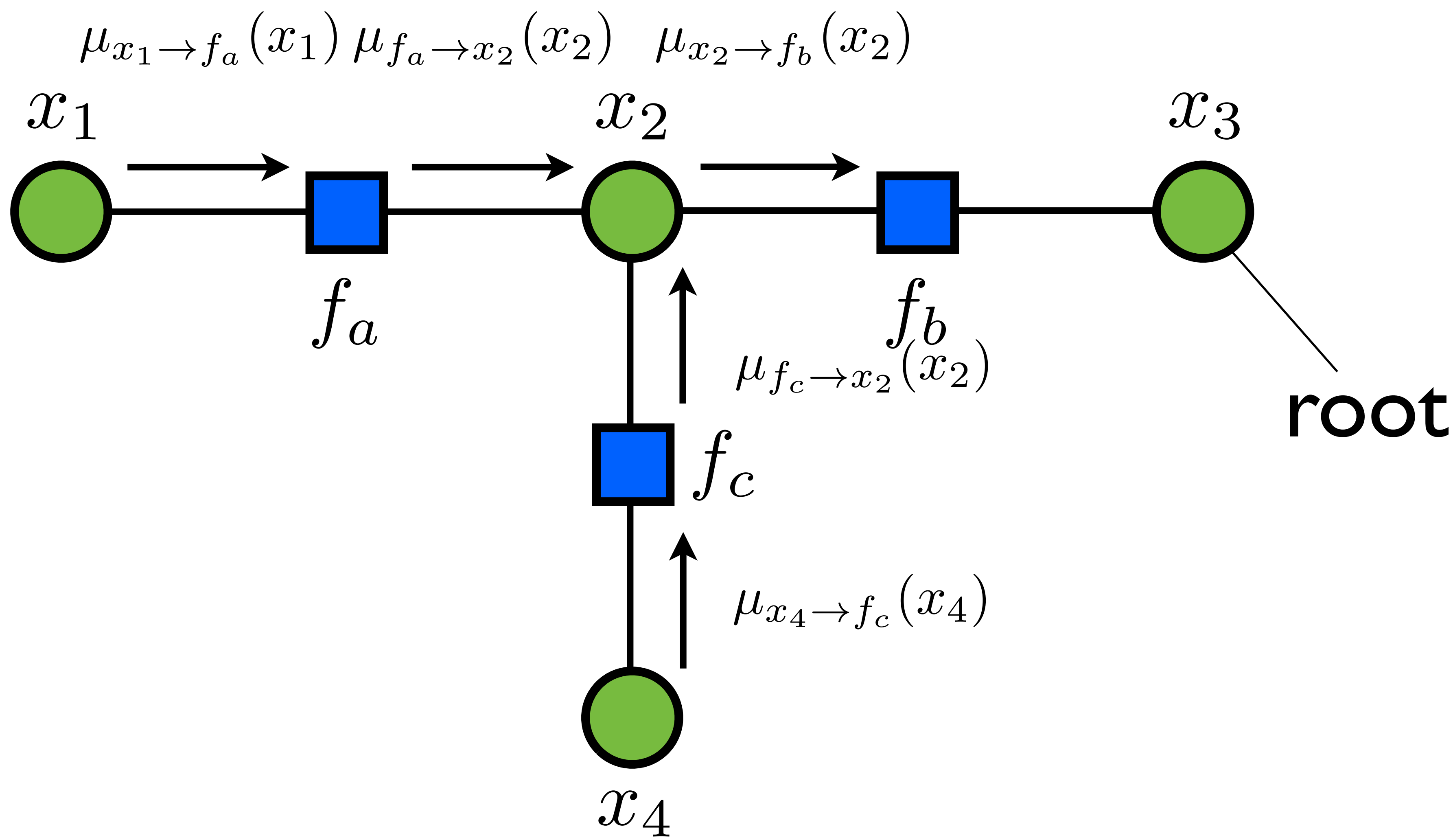
SUM-PRODUCT ALGORITHM

EXAMPLE



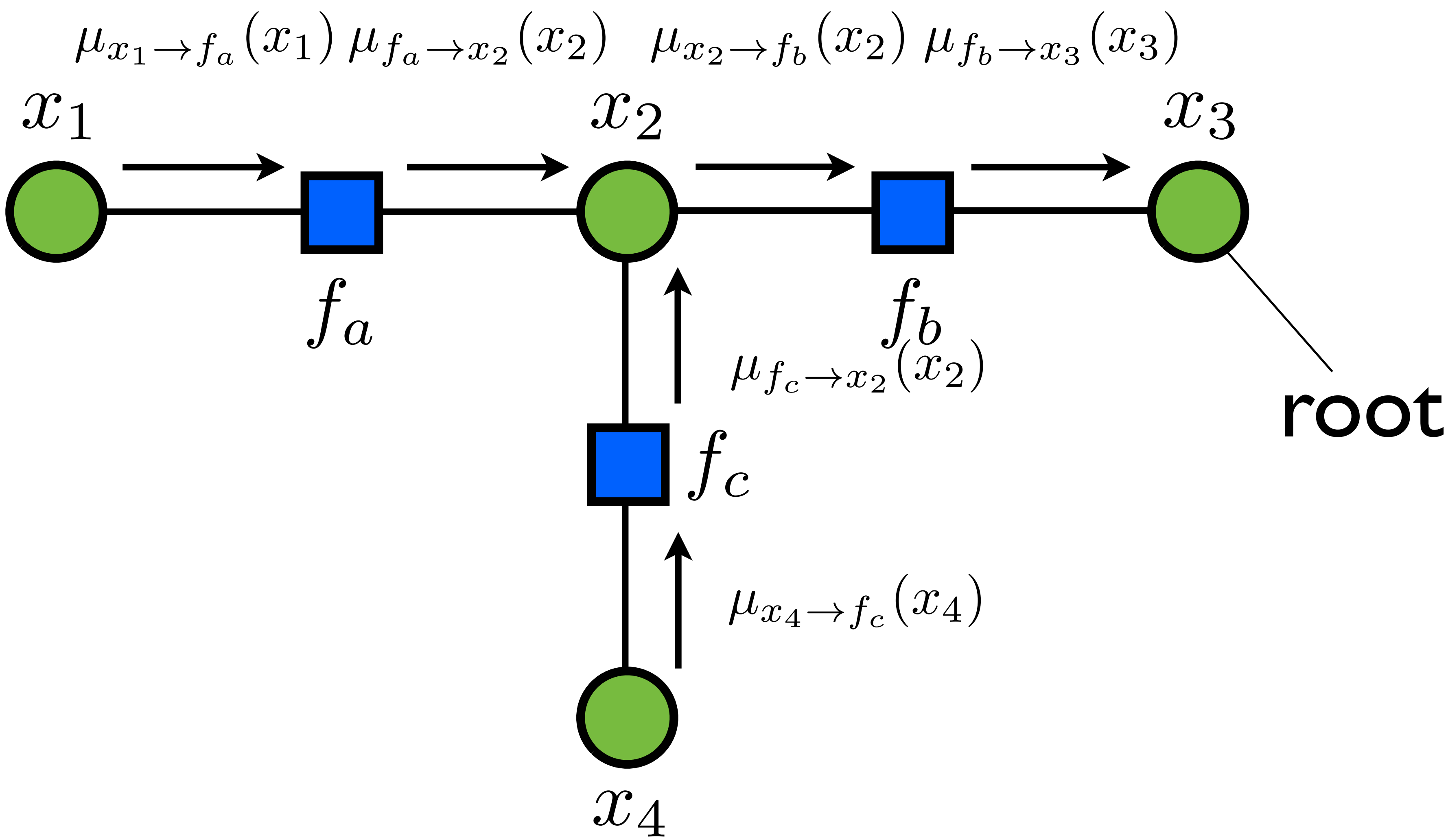
SUM-PRODUCT ALGORITHM

EXAMPLE

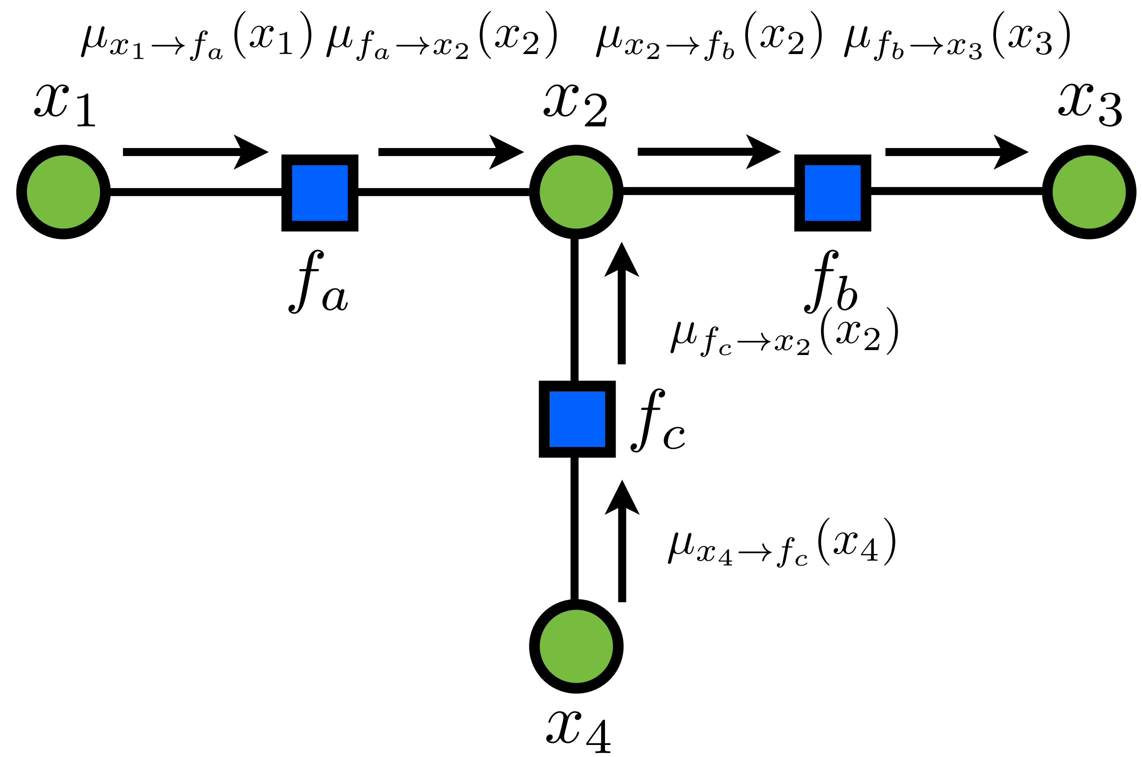


SUM-PRODUCT ALGORITHM

EXAMPLE

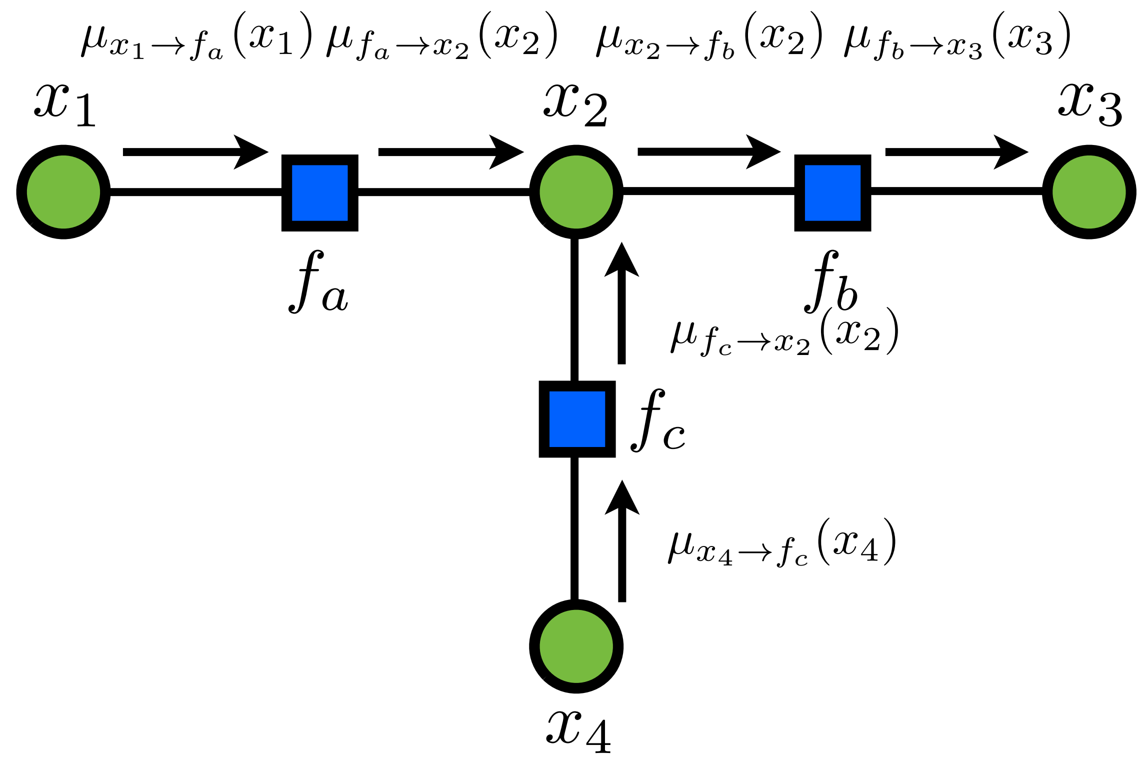


FORWARD MESSAGES



FORWARD MESSAGES

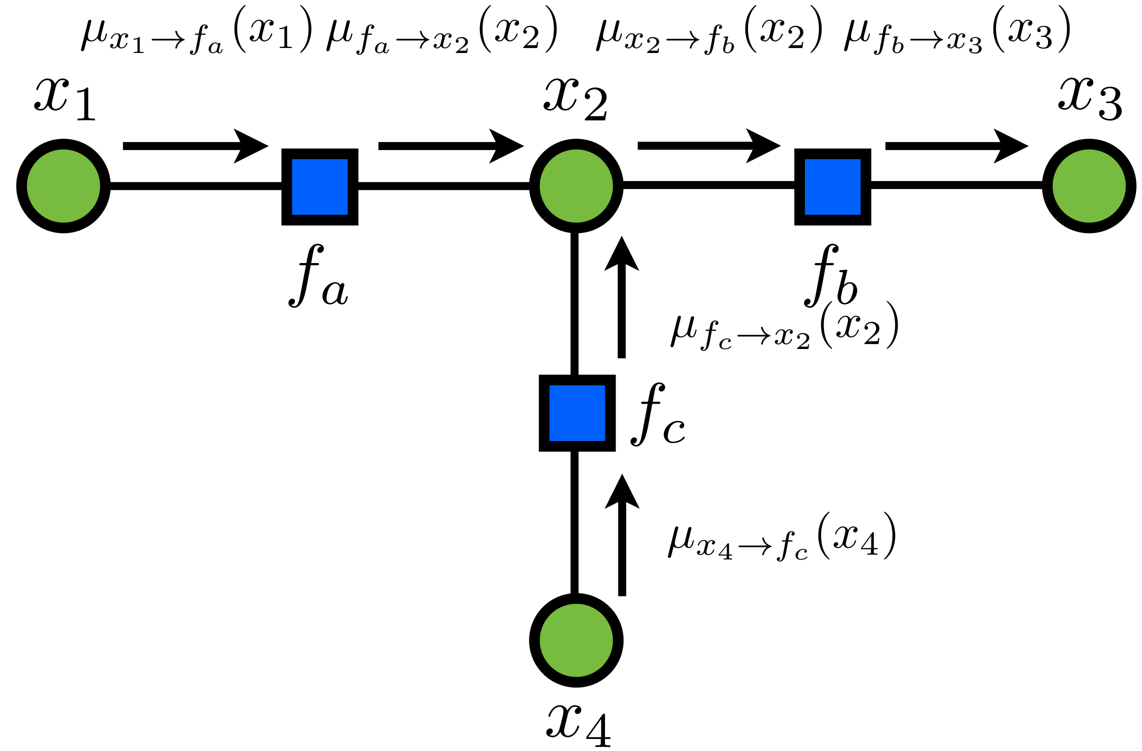
$$\mu_{x_1 \to f_a}(x_1) = 1$$



FORWARD MESSAGES

$$\mu_{x_1 \to f_a}(x_1) = 1$$

$$\mu_{f_a \to x_2}(x_2) = \sum_{x_1} f_a(x_1, x_2)$$

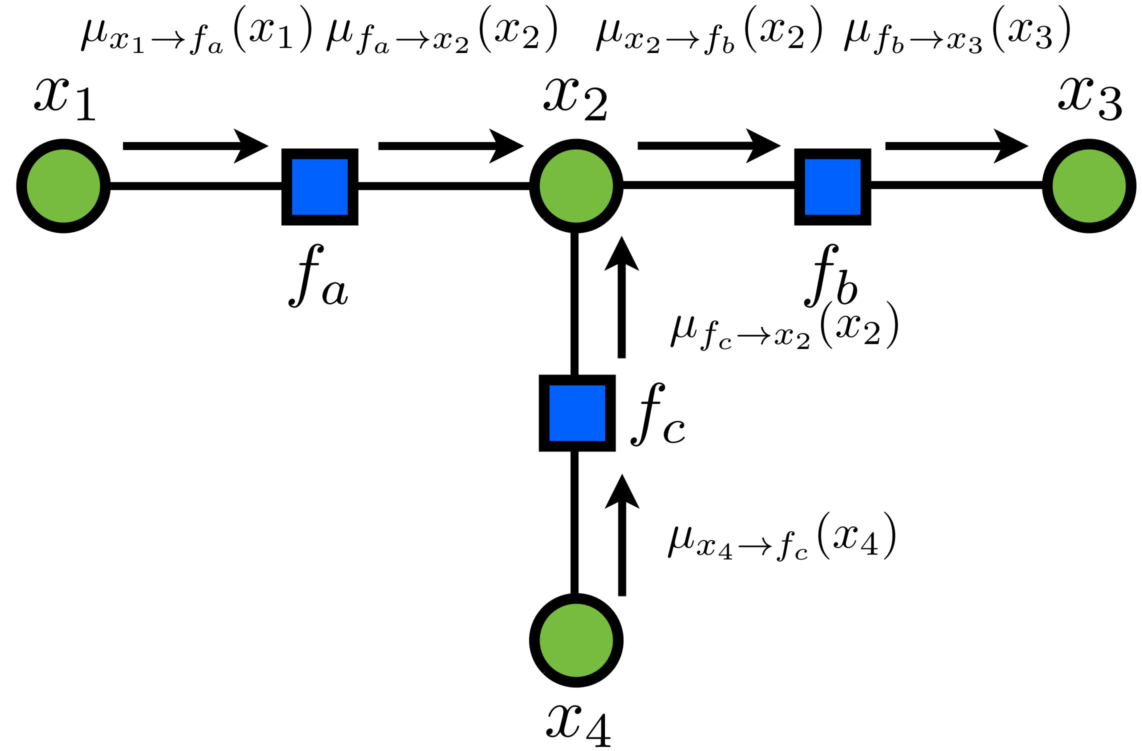


FORWARD MESSAGES

$$\mu_{x_1 \to f_a}(x_1) = 1$$

$$\mu_{f_a \to x_2}(x_2) = \sum_{x_1} f_a(x_1, x_2)$$

$$\mu_{x_4 \to f_c}(x_4) = 1$$



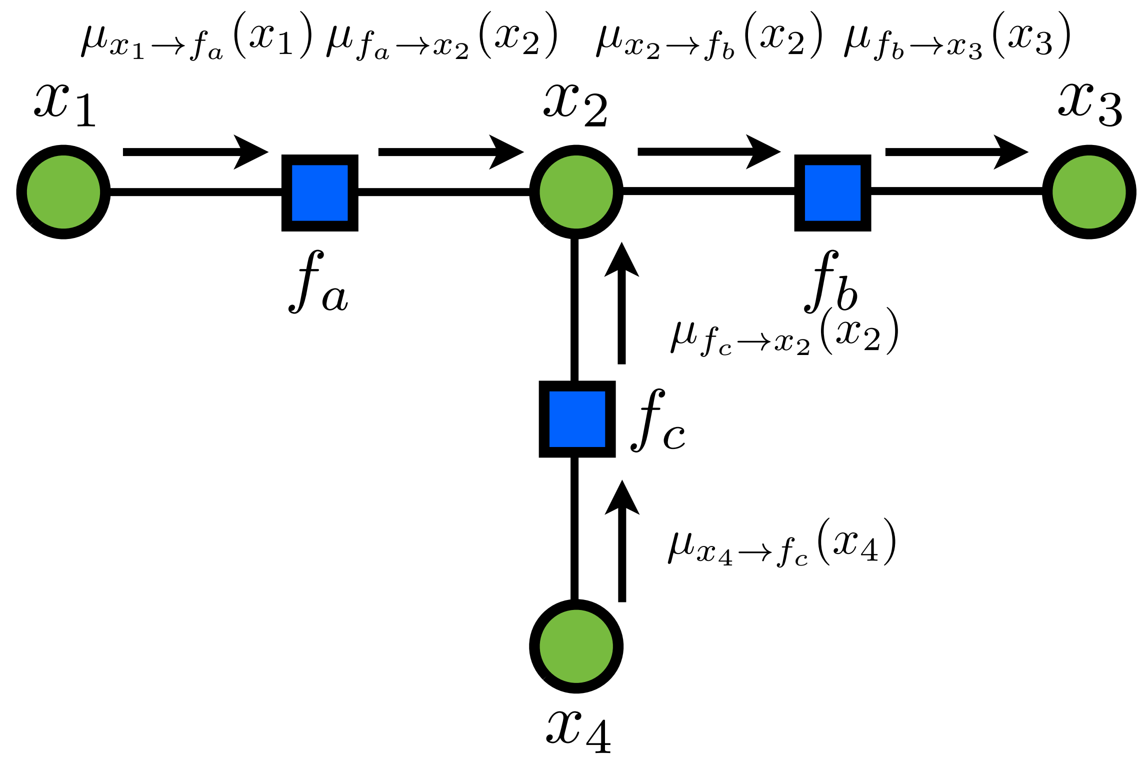
FORWARD MESSAGES

$$\mu_{x_1 \to f_a}(x_1) = 1$$

$$\mu_{f_a \to x_2}(x_2) = \sum_{x_1} f_a(x_1, x_2)$$

$$\mu_{x_4 \to f_c}(x_4) = 1$$

$$\mu_{f_c \to x_2}(x_2) = \sum_{x_1} f_c(x_2, x_4)$$



FORWARD MESSAGES

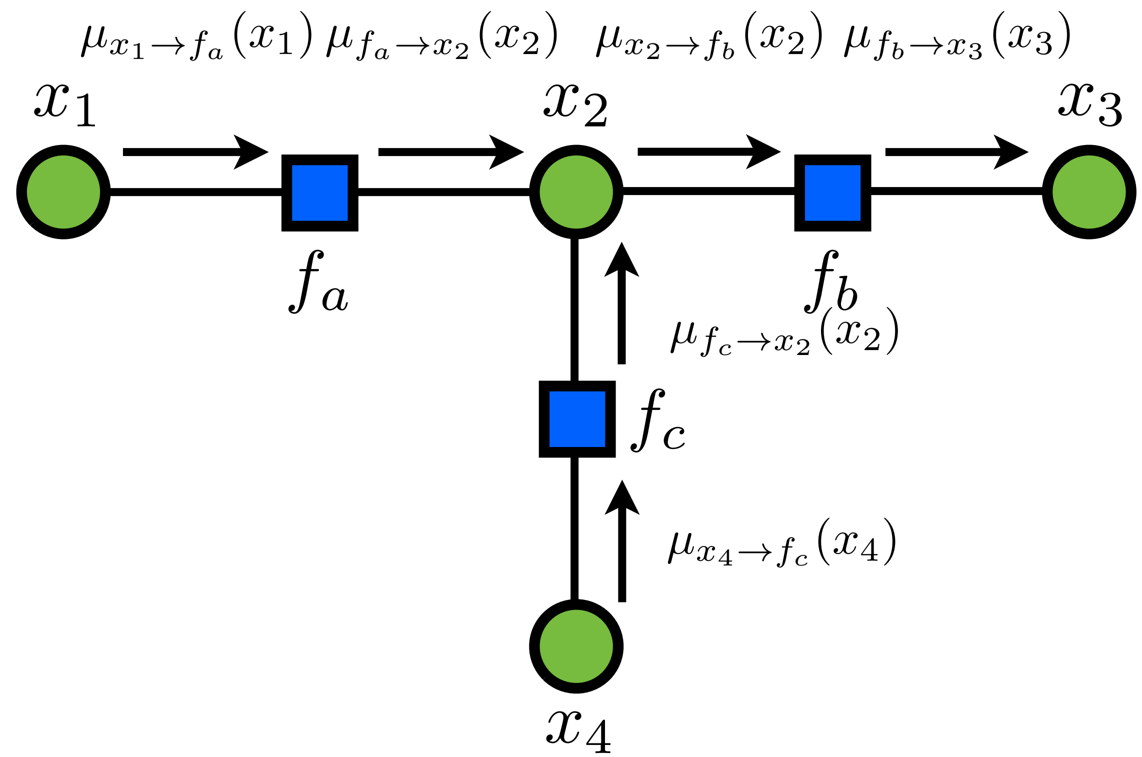
$$\mu_{x_1 \to f_a}(x_1) = 1$$

$$\mu_{f_a \to x_2}(x_2) = \sum_{x_1} f_a(x_1, x_2)$$

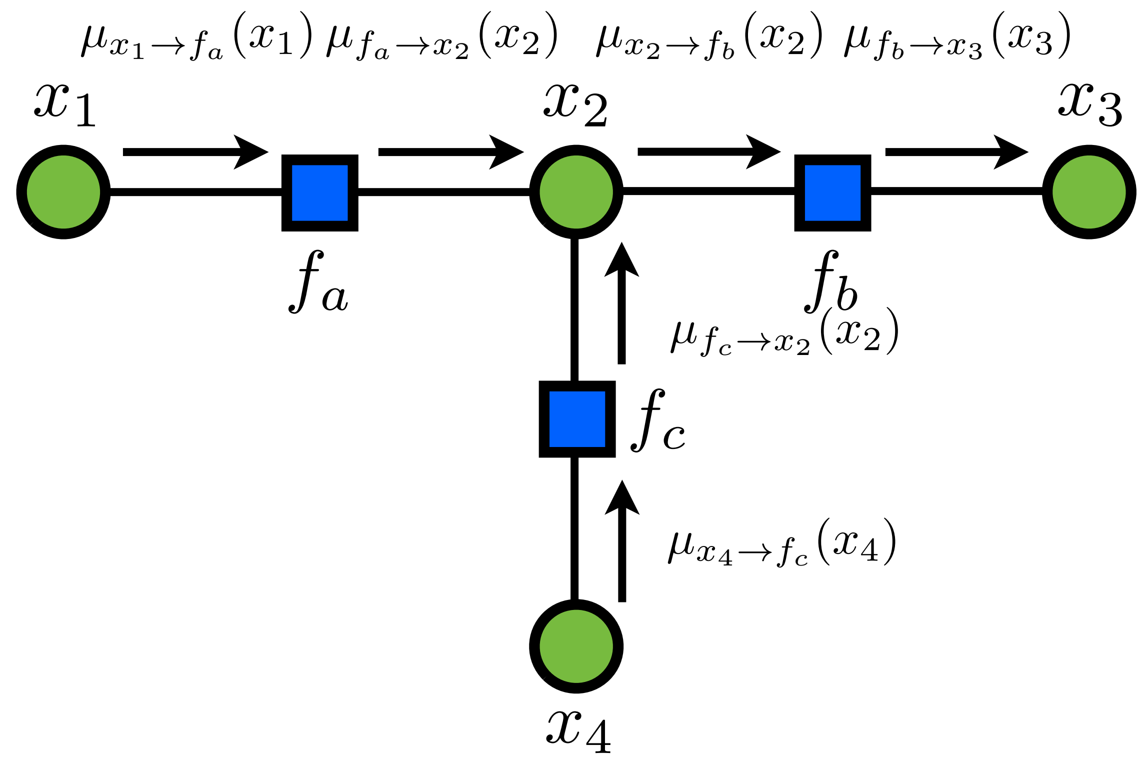
$$\mu_{x_4 \to f_c}(x_4) = 1$$

$$\mu_{f_c \to x_2}(x_2) = \sum_{x_1} f_c(x_2, x_4)$$

$$\mu_{x_2 \to f_b}(x_2) = \mu_{f_a \to x_2}(x_2)\mu_{f_c \to x_2}(x_2)$$



FORWARD MESSAGES



$$\mu_{x_1 \to f_a}(x_1) = 1$$

$$\mu_{f_a \to x_2}(x_2) = \sum_{x_1} f_a(x_1, x_2)$$

$$\mu_{x_4 \to f_c}(x_4) = 1$$

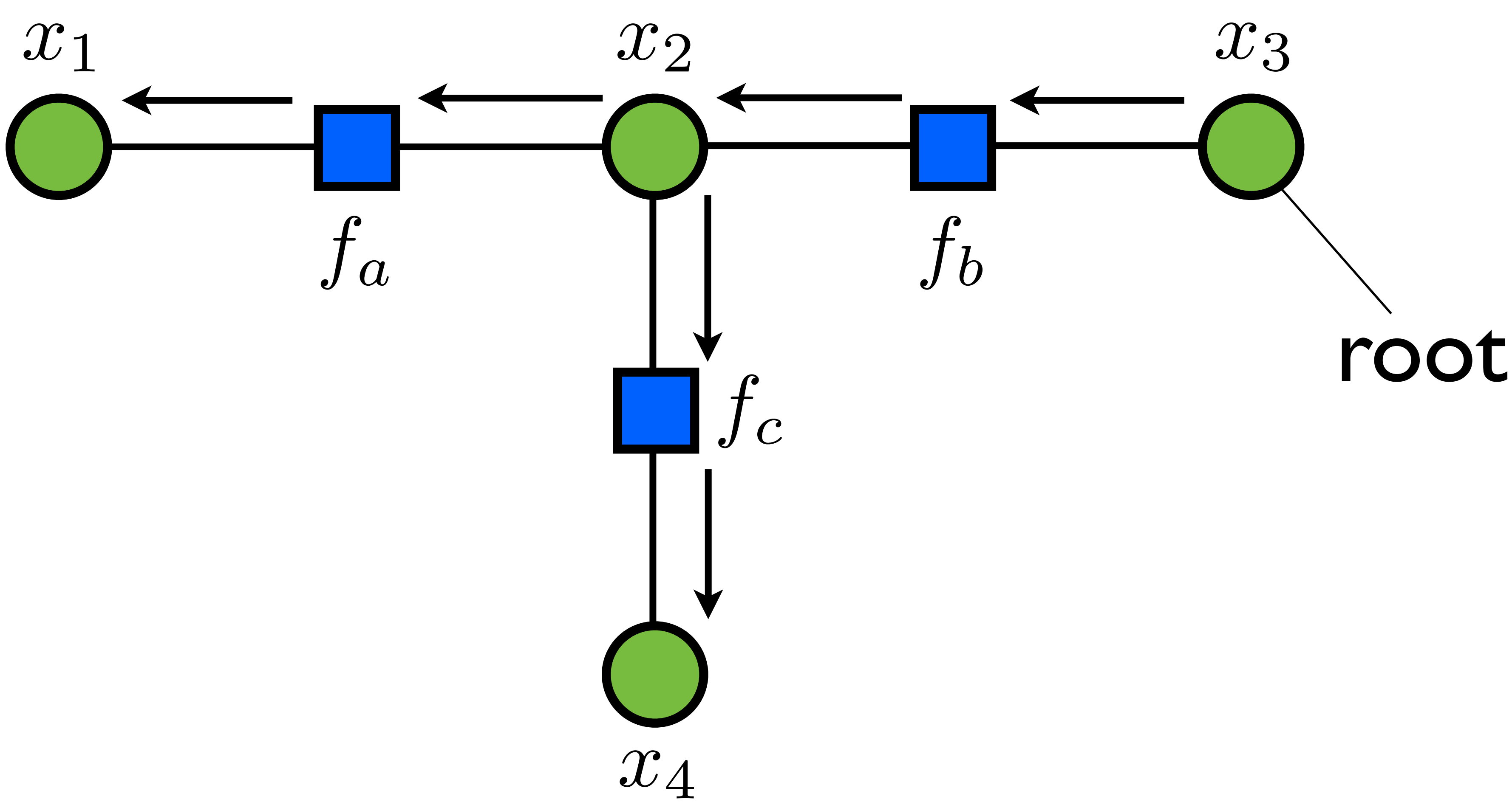
$$\mu_{f_c \to x_2}(x_2) = \sum_{x_1} f_c(x_2, x_4)$$

$$\mu_{x_2 \to f_b}(x_2) = \mu_{f_a \to x_2}(x_2)\mu_{f_c \to x_2}(x_2)$$

$$\mu_{f_b \to x_3}(x_3) = \sum_{x_2} f_b(x_2, x_3)\mu_{x_2 \to f_b}(x_2)$$

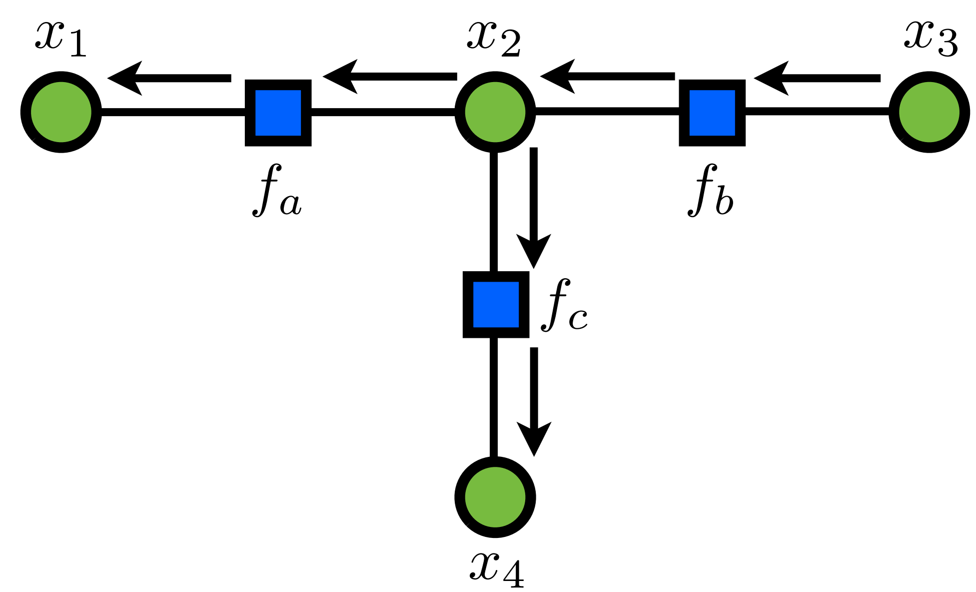
SUM-PRODUCT ALGORITHM

EXAMPLE



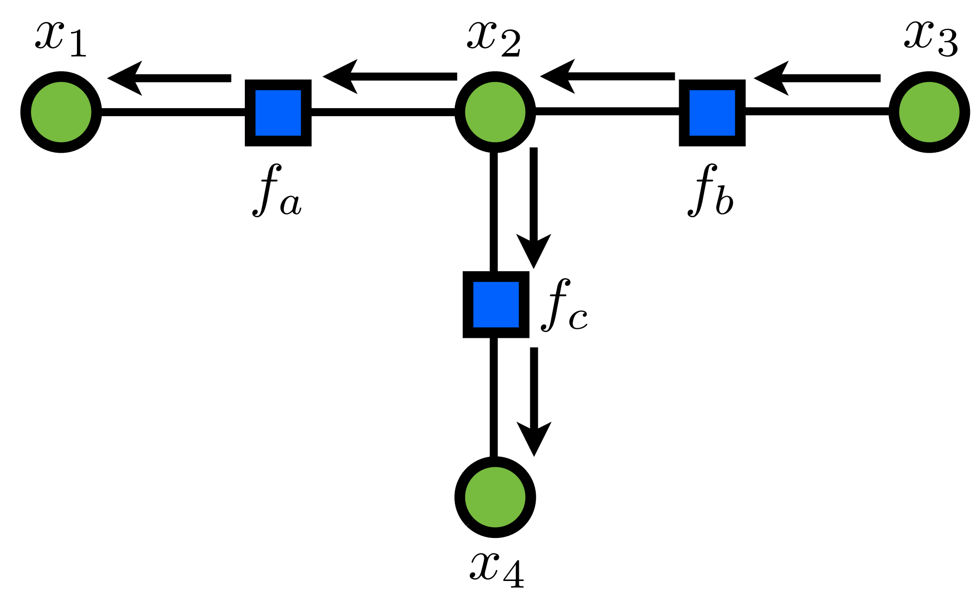
$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4)$$

BACKWARD MESSAGES



BACKWARD MESSAGES

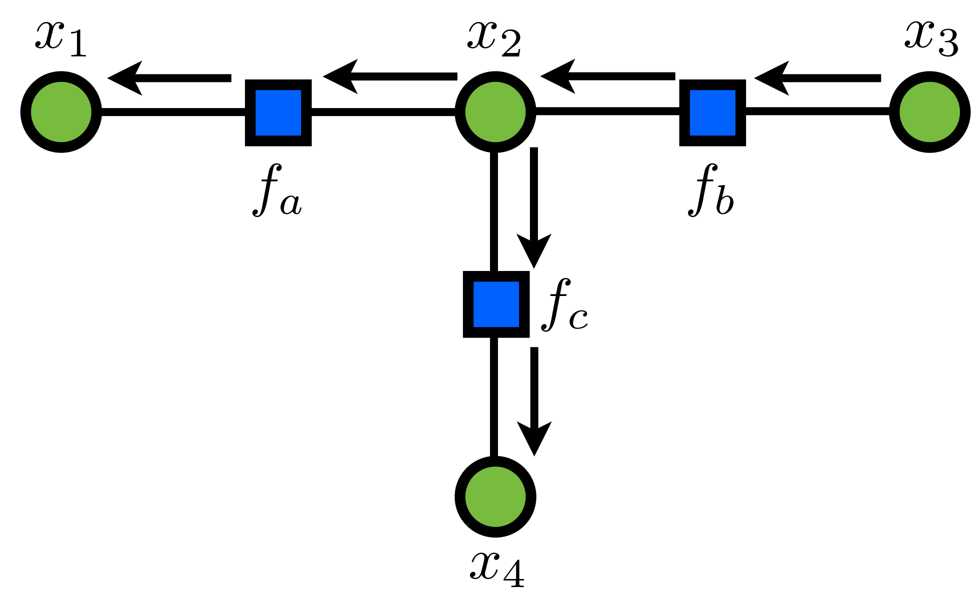
$$\mu_{x_3 \to f_b}(x_3) = 1$$



BACKWARD MESSAGES

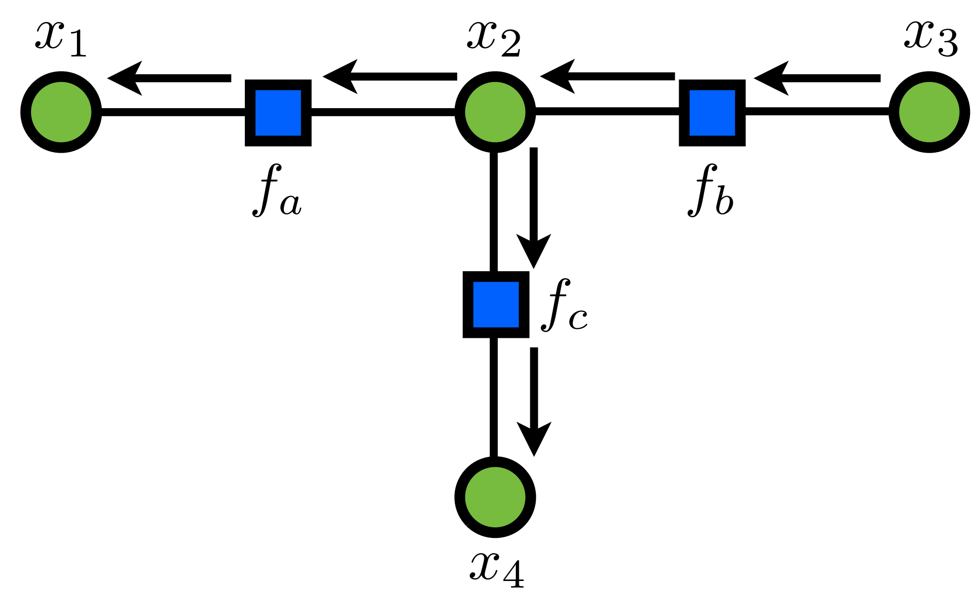
$$\mu_{x_3 \to f_b}(x_3) = 1$$

$$\mu_{f_b \to x_2}(x_2) = \sum_{x_3} f_b(x_2, x_3)$$



BACKWARD MESSAGES

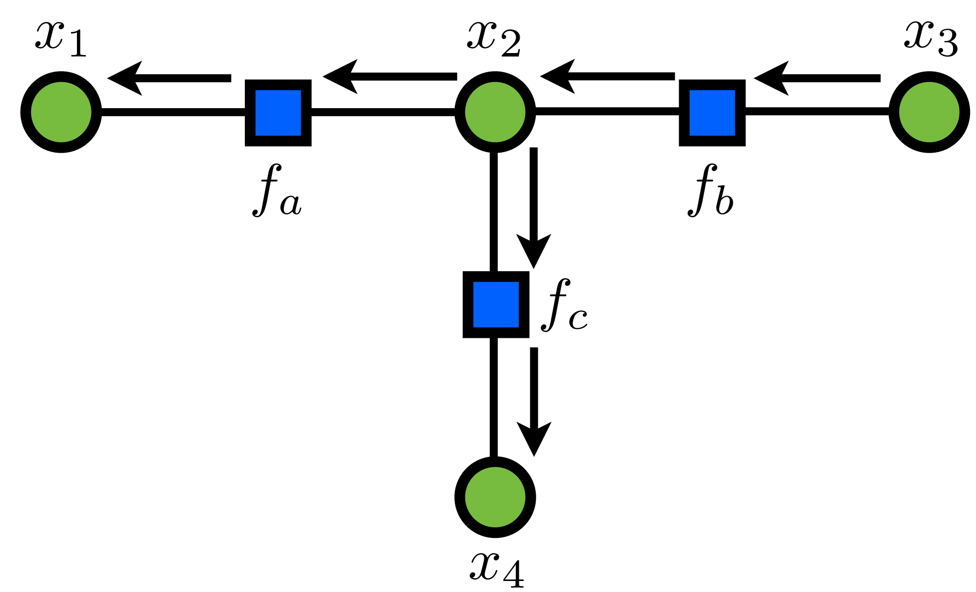
$$\mu_{x_3 \to f_b}(x_3) = 1$$



$$\mu_{f_b \to x_2}(x_2) = \sum_{x_3} f_b(x_2, x_3)$$

$$\mu_{x_2 \to f_a}(x_2) = \mu_{fb \to x_2}(x_2)\mu_{fc \to x_2}(x_2)$$

BACKWARD MESSAGES



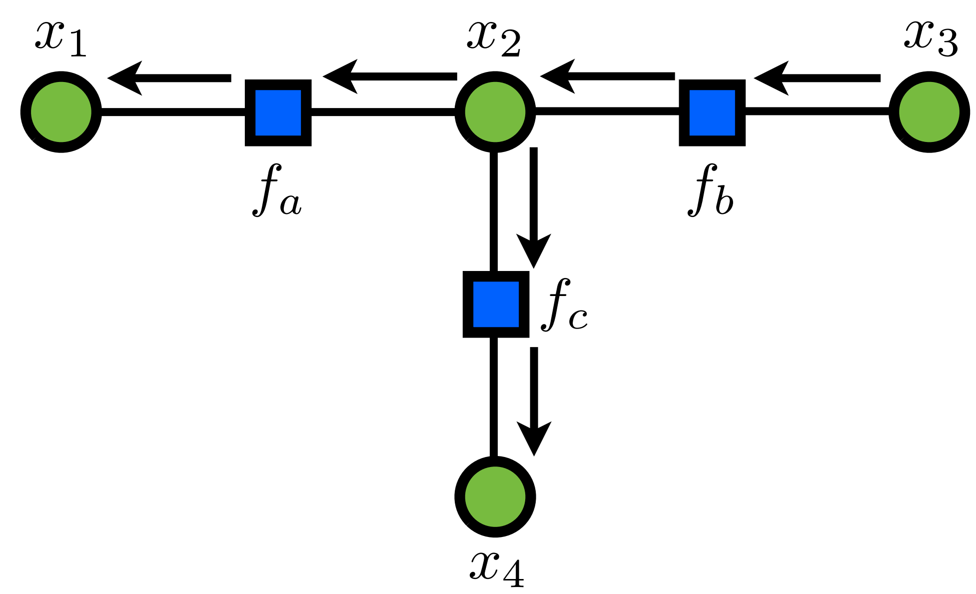
$$\mu_{x_3 \to f_b}(x_3) = 1$$

$$\mu_{f_b \to x_2}(x_2) = \sum_{x_3} f_b(x_2, x_3)$$

$$\mu_{x_2 \to f_a}(x_2) = \mu_{fb \to x_2}(x_2)\mu_{fc \to x_2}(x_2)$$

$$\mu_{f_a \to x_1}(x_1) = \sum_{x_2} f_a(x_1, x_2)\mu_{x_2 \to f_a}(x_2)$$

BACKWARD MESSAGES



$$\mu_{x_3 \to f_b}(x_3) = 1$$

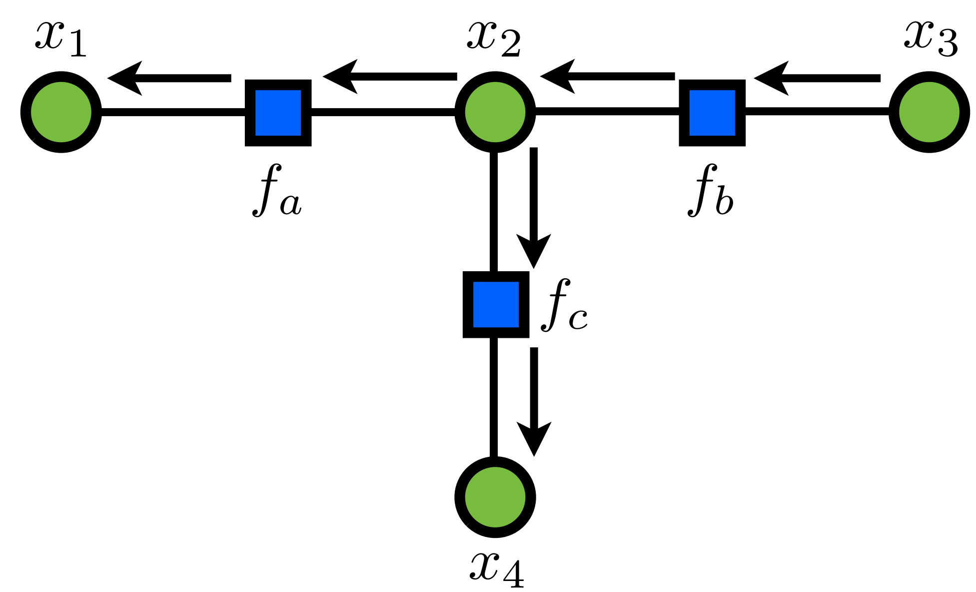
$$\mu_{f_b \to x_2}(x_2) = \sum_{x_3} f_b(x_2, x_3)$$

$$\mu_{x_2 \to f_a}(x_2) = \mu_{fb \to x_2}(x_2)\mu_{fc \to x_2}(x_2)$$

$$\mu_{f_a \to x_1}(x_1) = \sum_{x_2} f_a(x_1, x_2)\mu_{x_2 \to f_a}(x_2)$$

$$\mu_{x_2 \to f_c}(x_2) = \mu_{f_a \to x_2}(x_2)\mu_{f_b \to x_2}(x_2)$$

BACKWARD MESSAGES



$$\mu_{x_3 \to f_b}(x_3) = 1$$

$$\mu_{f_b \to x_2}(x_2) = \sum_{x_3} f_b(x_2, x_3)$$

$$\mu_{x_2 \to f_a}(x_2) = \mu_{fb \to x_2}(x_2)\mu_{fc \to x_2}(x_2)$$

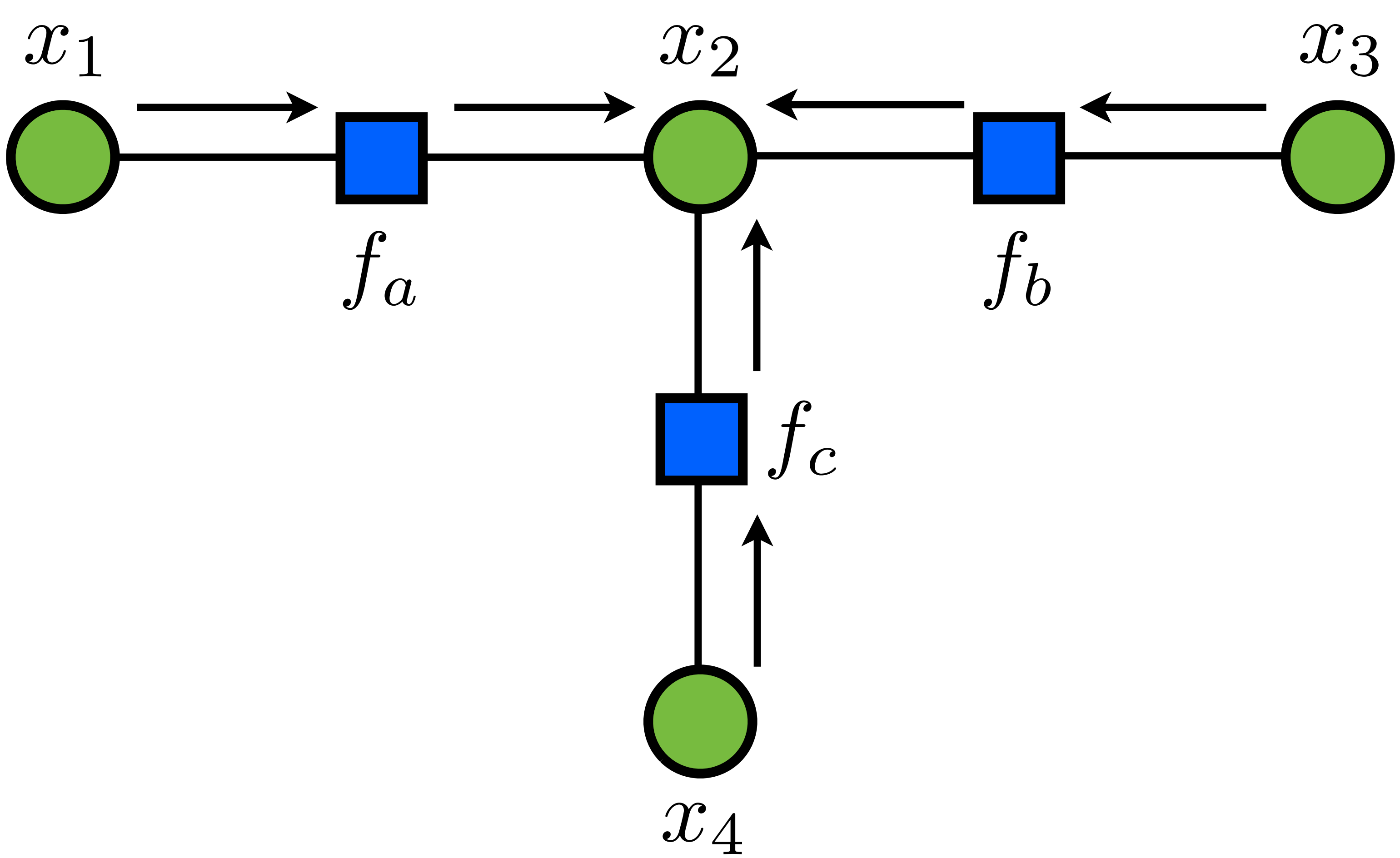
$$\mu_{f_a \to x_1}(x_1) = \sum_{x_2} f_a(x_1, x_2)\mu_{x_2 \to f_a}(x_2)$$

$$\mu_{x_2 \to f_c}(x_2) = \mu_{f_a \to x_2}(x_2)\mu_{f_b \to x_2}(x_2)$$

$$\mu_{f_c \to x_4}(x_4) = \sum_{x_2} f_c(x_2, x_4)\mu_{x_2 \to f_c}(x_2)$$

SUM-PRODUCT ALGORITHM

EXAMPLE



$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4)$$

SUM-PRODUCT ALGORITHM

EXAMPLE

$$p(x_2) = \mu_{f_a \to x_2}(x_2)\mu_{f_b \to x_2}(x_2)\mu_{f_c \to x_2}(x_2)$$

$$p(\mathbf{x}) = f_a(x_1, x_2)f_b(x_2, x_3)f_c(x_2, x_4)$$

SUM-PRODUCT ALGORITHM

EXAMPLE

$$p(x_2) = \mu_{f_a \to x_2}(x_2)\mu_{f_b \to x_2}(x_2)\mu_{f_c \to x_2}(x_2)$$

$$= \Big[\sum_{x_1} f_a(x_1, x_2)\Big]\Big[\sum_{x_3} f_b(x_2, x_3)\Big]\Big[\sum_{x_4} f_c(x_2, x_4)\Big]$$

$$p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4)$$

SUM-PRODUCT ALGORITHM

EXAMPLE

$$
\begin{aligned}
p(x_2) &= \mu_{f_a \to x_2}(x_2)\mu_{f_b \to x_2}(x_2)\mu_{f_c \to x_2}(x_2) \\
&= \left[\sum_{x_1} f_a(x_1, x_2)\right]\left[\sum_{x_3} f_b(x_2, x_3)\right]\left[\sum_{x_4} f_c(x_2, x_4)\right] \\
&= \sum_{x_1}\sum_{x_3}\sum_{x_4} f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4)
\end{aligned}
$$

$$
p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4)
$$

SUM-PRODUCT ALGORITHM

EXAMPLE

$$
\begin{aligned}
p(x_2) &= \mu_{f_a \to x_2}(x_2)\mu_{f_b \to x_2}(x_2)\mu_{f_c \to x_2}(x_2) \\
&= \Big[\sum_{x_1} f_a(x_1, x_2)\Big]\Big[\sum_{x_3} f_b(x_2, x_3)\Big]\Big[\sum_{x_4} f_c(x_2, x_4)\Big] \\
&= \sum_{x_1}\sum_{x_3}\sum_{x_4} f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4) \\
&= \sum_{x_1}\sum_{x_3}\sum_{x_4} p(\mathbf{x})
\end{aligned}
$$

$$
p(\mathbf{x}) = f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4)
$$

SUMMARY

- D-SEPARATION
- FACTOR GRAPHS
- SUM-PRODUCT ALGORITHM
- FURTHER READING:
  - MARKOV RANDOM FIELDS
  - MAX-PRODUCT ALGORITHM
  - LOOPY BELIEF PROPAGATION