ECE521 W17 Tutorial 10

Shenlong Wang and Renjie Liao

*Some of materials are credited to Jimmy Ba, Eric Sudderth, Chris Bishop

UNIVERSITY OF TORONTO

Introduction to A4

1, Graphical Models

2, Message Passing

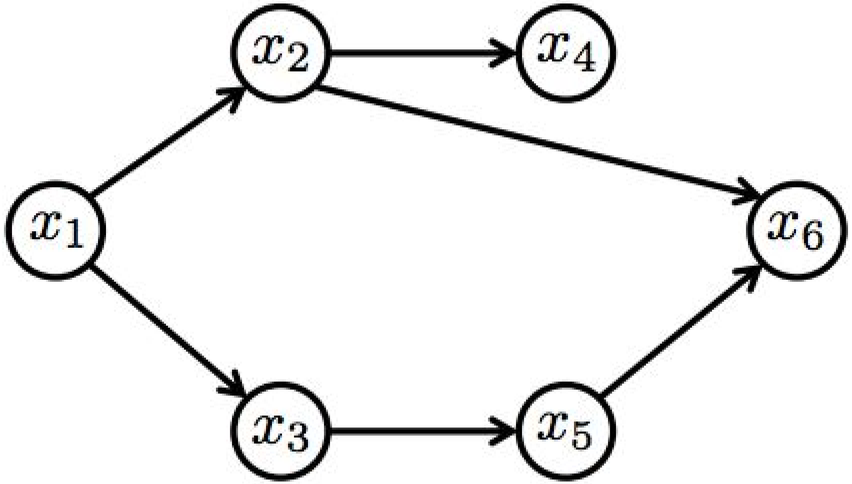
3, HMM

Introduction to A4

1, Graphical Models

$$p(x) = p(x_1)p(x_2 \mid x_1)p(x_3 \mid x_1)p(x_4 \mid x_2)p(x_5 \mid x_3)p(x_6 \mid x_2, x_5)$$

Draw Bayes-net:



Introduction to A4

Factor graph:

$$p(x) = \frac{1}{Z} \prod_{f \in \mathcal{F}} \psi_f(x_f) \qquad\qquad Z = \sum_x \prod_{f \in \mathcal{F}} \psi_f(x_f)$$

$Z > 0 \longrightarrow$ normalization constant (partition function)

$\psi_f(x_f) \geq 0 \longrightarrow$ arbitrary non-negative *potential function*

$\mathcal{F} \longrightarrow$ set of hyperedges linking subsets of nodes $f \subseteq \mathcal{V}$

$\mathcal{V} \longrightarrow$ set of N nodes or vertices, $\{1, 2, \ldots, N\}$

Introduction to A4

1, Graphical Models

$$p(x) = p(x_1)p(x_2 \mid x_1)p(x_3 \mid x_1)p(x_4 \mid x_2)p(x_5 \mid x_3)p(x_6 \mid x_2, x_5)$$

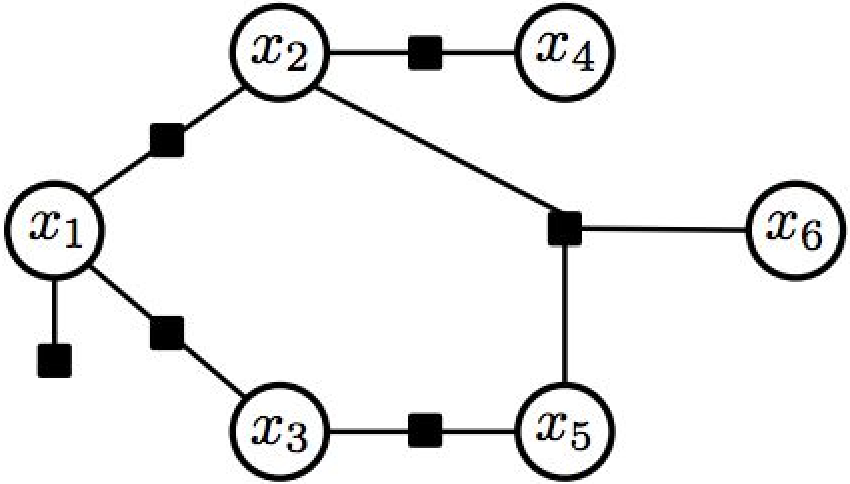
Draw factor-graph: $p(x) \propto \psi_1(x_1)\psi_2(x_2, x_1)\psi_3(x_3, x_1)\psi_4(x_4, x_2)\psi_5(x_5, x_3)\psi_6(x_6, x_2, x_5)$

Introduction to A4

1, Graphical Models

$$p(x) = p(x_1)p(x_2 \mid x_1)p(x_3 \mid x_1)p(x_4 \mid x_2)p(x_5 \mid x_3)p(x_6 \mid x_2, x_5)$$

Draw factor-graph: $p(x) \propto \psi_1(x_1)\psi_2(x_2, x_1)\psi_3(x_3, x_1)\psi_4(x_4, x_2)\psi_5(x_5, x_3)\psi_6(x_6, x_2, x_5)$

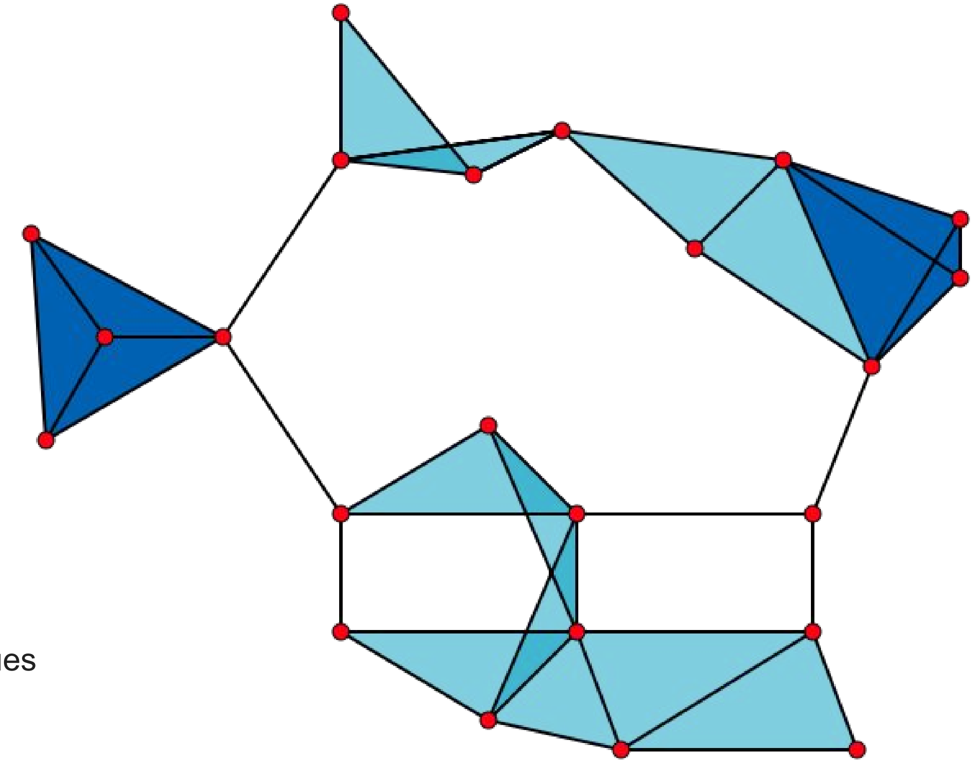


Introduction to A4

1, Graphical Models

Clique: A subset of vertices of an undirected graph such that every two distinct vertices in the clique are adjacent.

Maximal Clique: A clique that cannot be extended by including one more adjacent vertex



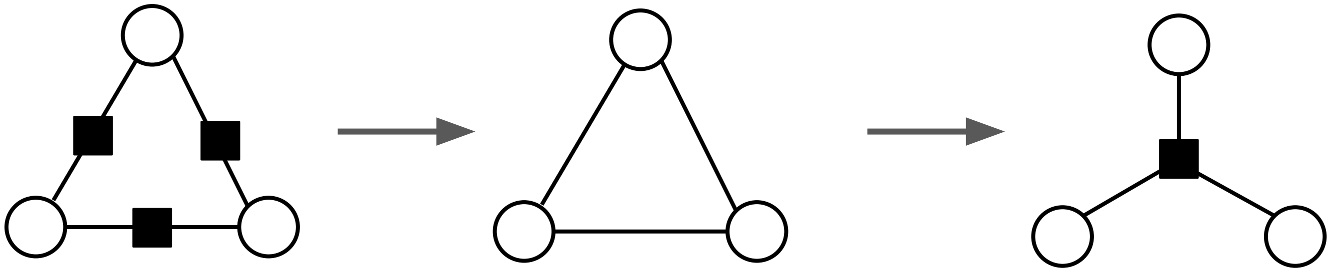
The 11 light blue 3-cliques (triangles) form maximal cliques

The 2 dark blue 4-cliques form maximal cliques

Introduction to A4

1, Graphical Models

Conversion:



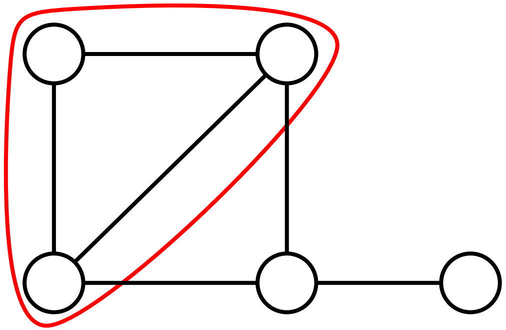
a FG expresses pair-wise factorization

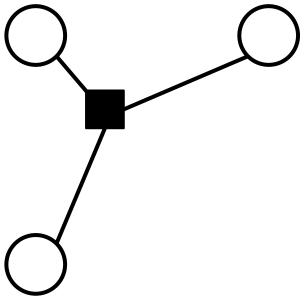
Convert the FG on the left to a MRF

Convert the MRF back to a FG, we lost the factorization property

Introduction to A4

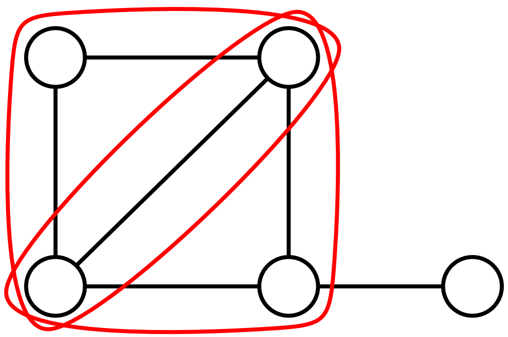
- Converting Markov Random Fields to factor graph takes the following steps:
  - Consider all the maximum cliques of the MRF
  - Create a factor node for each of the maximum cliques
  - Connect all the nodes of the maximum clique to the new factor nodes

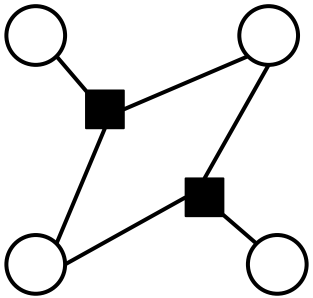




Introduction to A4

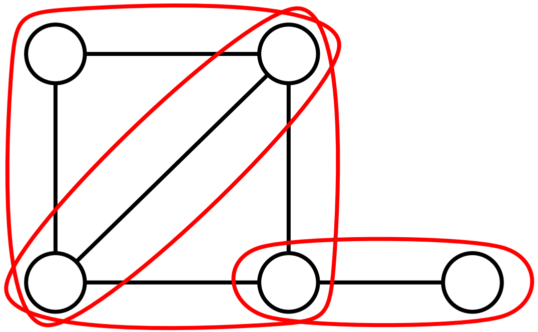
- Converting Markov Random Fields to factor graph takes the following steps:
  - Consider all the maximum cliques of the MRF
  - Create a factor node for each of the maximum cliques
  - Connect all the nodes of the maximum clique to the new factor nodes

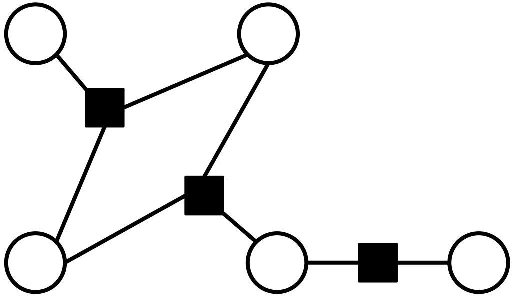




Introduction to A4

- Converting Markov Random Fields to factor graph takes the following steps:
    - Consider all the maximum cliques of the MRF
    - Create a factor node for each of the maximum cliques
    - Connect all the nodes of the maximum clique to the new factor nodes

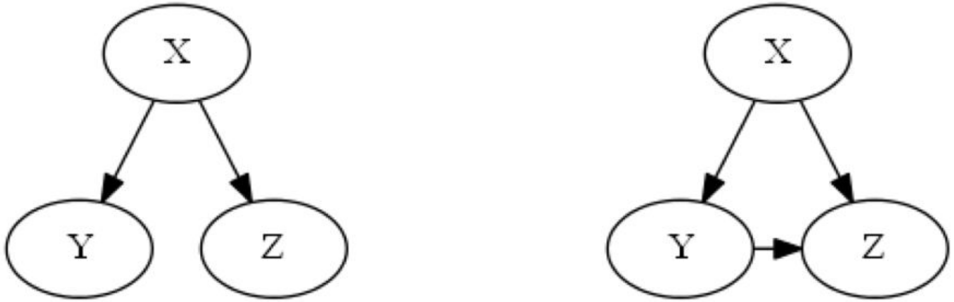




Introduction to A4

1, Graphical Models

Conditional Independence:



(a) $M_1 : Y \perp\!\!\!\perp Z | X$ **(b)** $M_2 : Y \not\perp\!\!\!\perp Z | X$

Introduction to A4

2, Message Passing

Sum-Product Algorithm

a.k.a. Belief Propagation, directed graphical models were used in expert systems, which need to infer their “beliefs” about the probabilities of various events

3, HMM

Forward-Backward Algorithm

Sum Product Algorithm

- The sum-product algorithm is used to compute probabilities for a **subset** of the variables of a joint distribution, e.g. $P(a, b, c, d)$
  - Marginal distributions, e.g. $P(b)$
  - Joint distributions of a subset of variables, e.g. $P(a,b)$
  - Conditional distributions (often the posterior distributions), e.g. $P(a,c \mid d)=P(a,c,d) / P(d)$

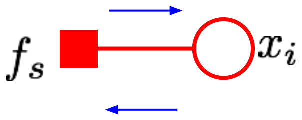
Message notations

- We use function μ to denote messages.

- On each edge of the factor graph, there are two messages traveling in opposite directions

- We use subscript to denote the origin and the destination of these messages, e.g.:

$\mu_{f_s \to x_i}(x_i)$ factor-to-variable

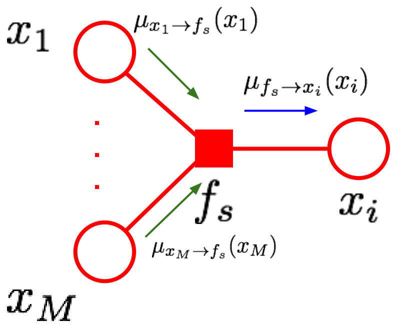


$\mu_{x_i \to f_s}(x_i)$ variable-to-factor

The sum product algorithm on factor graph

- Two rules in the sum-product algorithm:
  - Factor-to-variable messages:

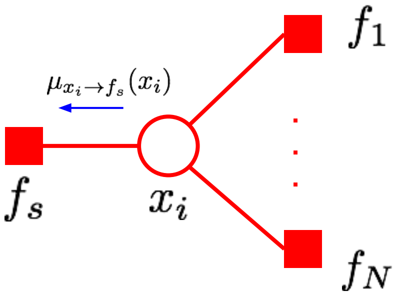
$$\mu_{f_s \to x_i}(x_i) = \sum_{Ne(f_s)\backslash x_i} f_s(x_i, x_1, \cdots, x_M) \underbrace{\prod_{x_m \in Ne(f_s)\backslash x_i} \mu_{x_m \to f_s}(x_m)}_{\text{Incoming messages}}$$



The sum product algorithm on factor graph

- Two rules in the sum-product algorithm:
    - Variable-to-factor messages:

$$\mu_{x_i \to f_s}(x_i) = \prod_{f_n \in Ne(x_i) \backslash f_s} \mu_{f_n \to x_i}(x_i)$$



The sum product algorithm on factor graph

- Two rules in the sum-product algorithm:
  - Factor-to-variable messages:

$$\mu_{f_s \to x_i}(x_i) = \sum_{Ne(f_s)\backslash x_i} f_s(x_i, x_1, \cdots, x_M) \prod_{x_m \in Ne(f_s)\backslash x_i} \mu_{x_m \to f_s}(x_m)$$

  - Variable-to-factor messages:

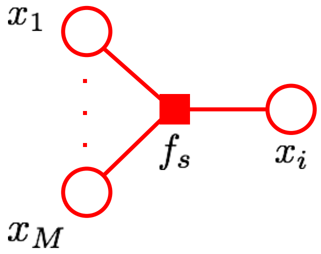
$$\mu_{x_i \to f_s}(x_i) = \prod_{f_n \in Ne(x_i)\backslash f_s} \mu_{f_n \to x_i}(x_i)$$

The sum product algorithm on factor graph

- How to start the sum-product algorithm:
  - Choose a node in the factor graph as the root node
  - Compute all the leaf-to-root messages
  - Compute all the root-to-leaf messages
- Initial conditions:
  - Starting from a factor leaf/root node, the initial factor-to-variable message is the factor itself
  - Starting from a variable leaf/root node, the initial variable-to-factor message is a vector of ones

Compute probabilities with messages

- How to convert messages to actual probabilities:



Compute unnormalized probabilities using messages (the sum-product algorithm):

unnormalized marginal probabilities:

$$g(x_i) = \prod_{f_n \in Ne(x_i)} \mu_{f_n \to x_i}(x_i)$$

normalization constant:

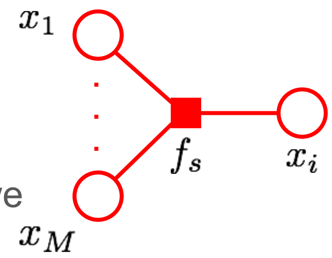
$$Z = \sum_{x_i} g(x_i)$$

marginal probabilities:

$$P(x_i) = \frac{1}{Z} g(x_i)$$

Compute probabilities with messages

- How to convert messages to actual probabilities:



Let $\{x_i\} \cup X_s \subseteq \{x_i, x_1, \cdots, x_M\}$ be a subset of the variables we wish to compute joint distribution over

Compute unnormalized probabilities using messages (the sum-product algorithm):

unnormalized joint probabilities:

$$g(x_i, X_s) = f_s(x_i, X_s) \prod_{x_m \in \{x_i, X_s\}} \prod_{f_n \in Ne(x_m) \setminus f_s} \mu_{f_n \to x_m}(x_m)$$

When all the variables in $\{x_i\} \cup X_s$ are connected to the same factor. We make the computation efficient

normalization constant:

$$Z = \sum_{x_i} g(x_i)$$

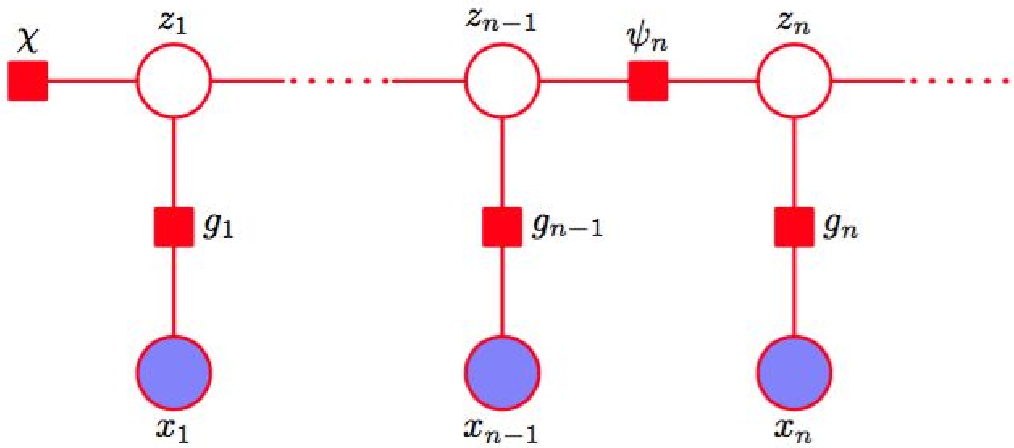
marginal probabilities:

$$P(x_i, x_1, \cdots, x_k) = \frac{1}{Z} g(x_i, x_1, \cdots, x_k)$$

Equivalence

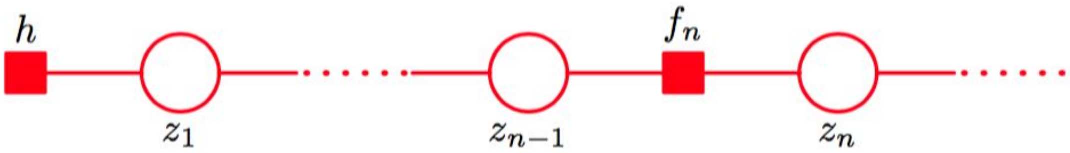
Sum-product algorithm is equivalent to forward-backward algorithm in the context of Hidden Markov Models

Factor graph of HMM:



Forward-Backward Algorithm

For the purpose of inference, we can simplify the factor graph as below:

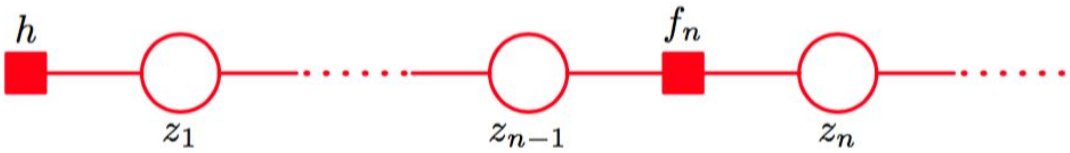


Factors are:

$$\begin{aligned} h(\mathbf{z}_1) &= p(\mathbf{z}_1)p(\mathbf{x}_1|\mathbf{z}_1) \\ f_n(\mathbf{z}_{n-1}, \mathbf{z}_n) &= p(\mathbf{z}_n|\mathbf{z}_{n-1})p(\mathbf{x}_n|\mathbf{z}_n) \end{aligned}$$

Forward-Backward Algorithm

For the purpose of inference, we can simplify the factor graph as below:



Factors are:

$$
\begin{aligned}
h(\mathbf{z}_1) &= p(\mathbf{z}_1)p(\mathbf{x}_1|\mathbf{z}_1) \\
f_n(\mathbf{z}_{n-1}, \mathbf{z}_n) &= p(\mathbf{z}_n|\mathbf{z}_{n-1})p(\mathbf{x}_n|\mathbf{z}_n)
\end{aligned}
$$

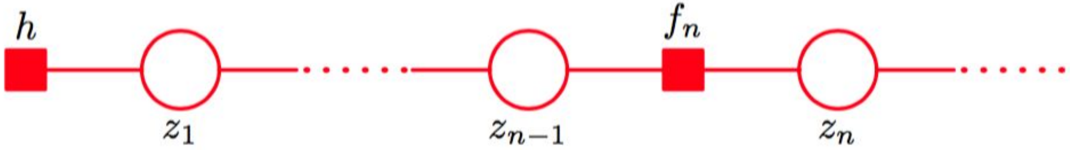
Messages are:

$$
\begin{aligned}
\underline{\mu_{\mathbf{z}_{n-1} \to f_n}(\mathbf{z}_{n-1}) = \mu_{f_{n-1} \to \mathbf{z}_{n-1}}(\mathbf{z}_{n-1})} \\
\mu_{f_n \to \mathbf{z}_n}(\mathbf{z}_n) = \sum_{\mathbf{z}_{n-1}} f_n(\mathbf{z}_{n-1}, \mathbf{z}_n)\mu_{\mathbf{z}_{n-1} \to f_n}(\mathbf{z}_{n-1})
\end{aligned}
$$

Quiz: Why?

Forward-Backward Algorithm

For the purpose of inference, we can simplify the factor graph as below:



Rewrite message:

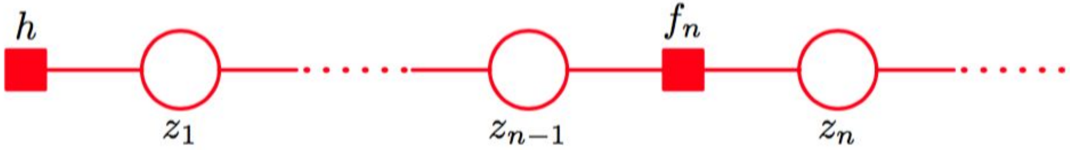
$$\mu_{\mathbf{z}_{n-1} \to f_n}(\mathbf{z}_{n-1}) = \mu_{f_{n-1} \to \mathbf{z}_{n-1}}(\mathbf{z}_{n-1})$$

$$\mu_{f_n \to \mathbf{z}_n}(\mathbf{z}_n) = \sum_{\mathbf{z}_{n-1}} f_n(\mathbf{z}_{n-1}, \mathbf{z}_n) \mu_{\mathbf{z}_{n-1} \to f_n}(\mathbf{z}_{n-1})$$

$$\boxed{\mu_{f_n \to \mathbf{z}_n}(\mathbf{z}_n) = \sum_{\mathbf{z}_{n-1}} f_n(\mathbf{z}_{n-1}, \mathbf{z}_n) \mu_{f_{n-1} \to \mathbf{z}_{n-1}}(\mathbf{z}_{n-1})}$$

Forward-Backward Algorithm

For the purpose of inference, we can simplify the factor graph as below:



Rewrite message:

$$\boxed{\mu_{f_n \to \mathbf{z}_n}(\mathbf{z}_n) = \sum_{\mathbf{z}_{n-1}} f_n(\mathbf{z}_{n-1}, \mathbf{z}_n) \mu_{f_{n-1} \to \mathbf{z}_{n-1}}(\mathbf{z}_{n-1})}$$

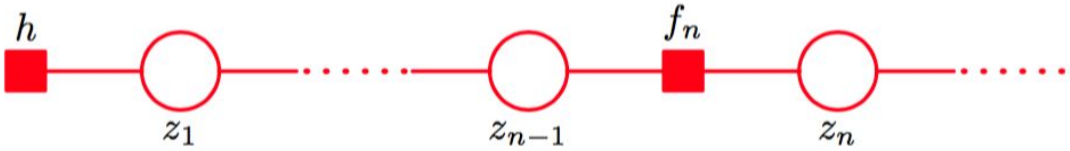
$$\alpha(\mathbf{z}_n) = \mu_{f_n \to \mathbf{z}_n}(\mathbf{z}_n)$$

Replace factor:

$$\alpha(\mathbf{z}_n) = p(\mathbf{x}_n|\mathbf{z}_n) \sum_{\mathbf{z}_{n-1}} \alpha(\mathbf{z}_{n-1}) p(\mathbf{z}_n|\mathbf{z}_{n-1})$$

Forward-Backward Algorithm

For the purpose of inference, we can simplify the factor graph as below:



Rewrite message:

$$\mu_{z_{n+1} \to f_{n+1}}(\mathbf{z}_{n+1}) = \mu_{f_{n+2} \to z_{n+1}}(\mathbf{z}_{n+1})$$

$$\mu_{f_{n+1} \to z_n}(\mathbf{z}_n) = \sum_{\mathbf{z}_{n+1}} f_{n+1}(\mathbf{z}_n, \mathbf{z}_{n+1}) \mu_{f_{n+2} \to z_{n+1}}(\mathbf{z}_{n+1})$$

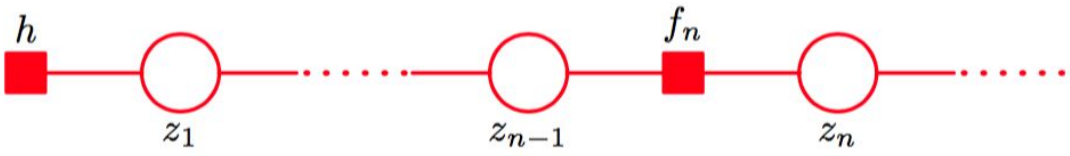
$$\beta(\mathbf{z}_n) = \mu_{f_{n+1} \to \mathbf{z}_n}(\mathbf{z}_n)$$

Replace factor:

$$\beta(\mathbf{z}_n) = \sum_{\mathbf{z}_{n+1}} \beta(\mathbf{z}_{n+1}) p(\mathbf{x}_{n+1} | \mathbf{z}_{n+1}) p(\mathbf{z}_{n+1} | \mathbf{z}_n)$$

Forward-Backward Algorithm

For the purpose of inference, we can simplify the factor graph as below:

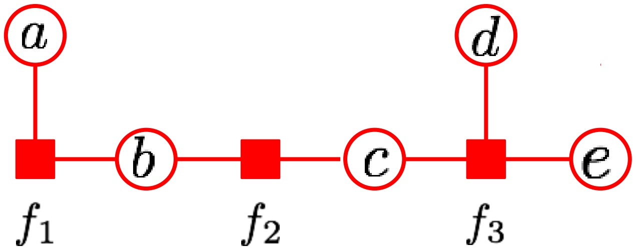


Obtain Posterior:

$$p(\mathbf{z}_n, \mathbf{X}) = \mu_{f_n \to \mathbf{z}_n}(\mathbf{z}_n)\mu_{f_{n+1} \to \mathbf{z}_n}(\mathbf{z}_n) = \alpha(\mathbf{z}_n)\beta(\mathbf{z}_n)$$

$$\gamma(\mathbf{z}_n) = \frac{p(\mathbf{z}_n, \mathbf{X})}{p(\mathbf{X})} = \frac{\alpha(\mathbf{z}_n)\beta(\mathbf{z}_n)}{p(\mathbf{X})}$$

Example of the sum product algorithm



$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

A set of binary random variables: $a, b, c, d, e \in \{0, 1\}$

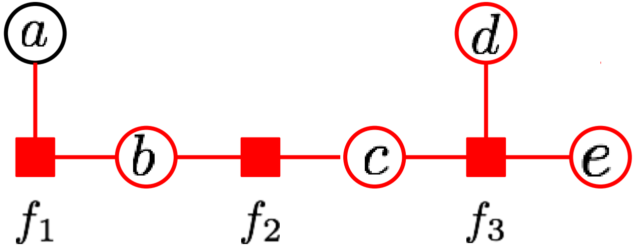
$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

The sum product algorithm on factor graph

- How to start the sum-product algorithm:
  - **Choose a node in the factor graph as the root node**
  - Compute all the leaf-to-root messages
  - Compute all the root-to-leaf messages
- Initial conditions:
  - Starting from a factor leaf/root node, the initial factor-to-variable message is the factor itself
  - Starting from a variable leaf/root node, the initial variable-to-factor message is a vector of ones

Example of the sum product algorithm



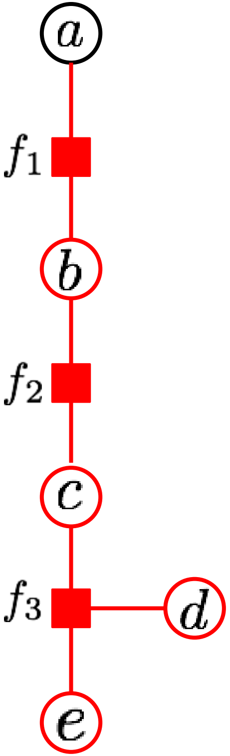
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



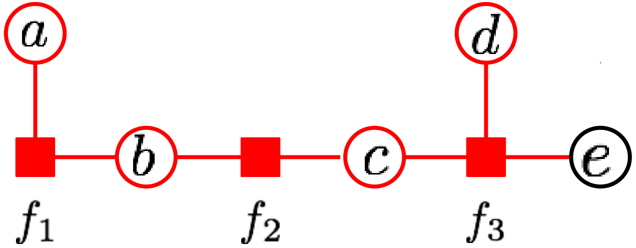
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



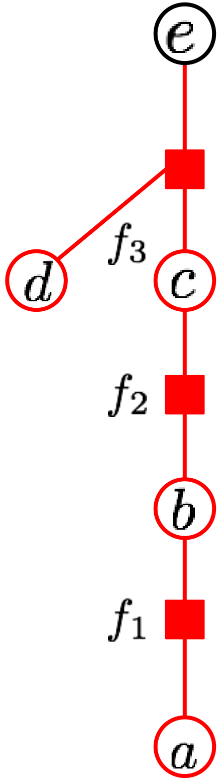
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

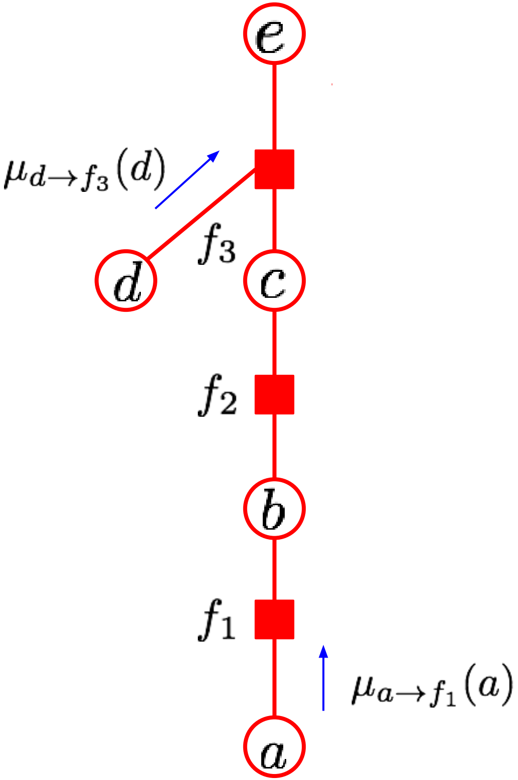
$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

The sum product algorithm on factor graph

- How to start the sum-product algorithm:
  - Choose a node in the factor graph as the root node
  - **Compute all the leaf-to-root messages**
  - Compute all the root-to-leaf messages
- Initial conditions:
  - Starting from a factor leaf/root node, the initial factor-to-variable message is the factor itself
  - Starting from a variable leaf/root node, the initial variable-to-factor message is a vector of ones

Example of the sum product algorithm



$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

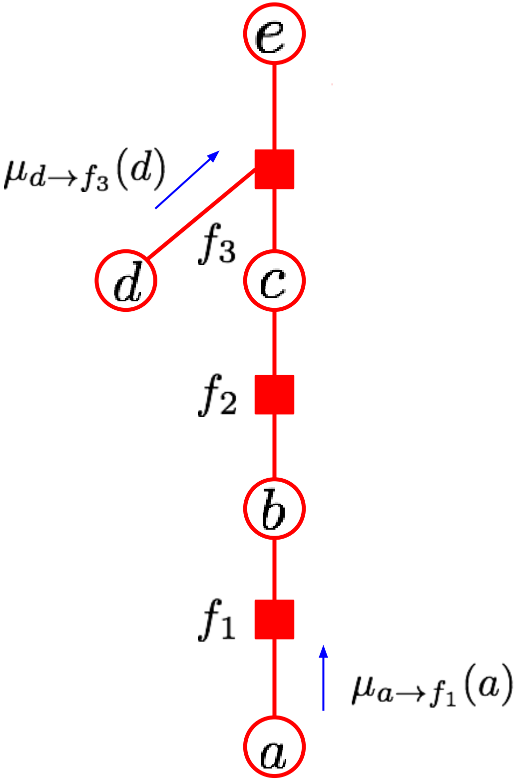
$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

The sum product algorithm on factor graph

- How to start the sum-product algorithm:
    - Choose a node in the factor graph as the root node
    - Compute all the leaf-to-root messages
    - Compute all the root-to-leaf messages
- Initial conditions:
    - Starting from a factor leaf/root node, the initial factor-to-variable message is the factor itself
    - **Starting from a variable leaf/root node, the initial variable-to-factor message is a vector of ones**

Example of the sum product algorithm



$$\mu_{a \to f_1}(a) = [1, 1]^T$$
$$\mu_{d \to f_3}(d) = [1, 1]^T$$

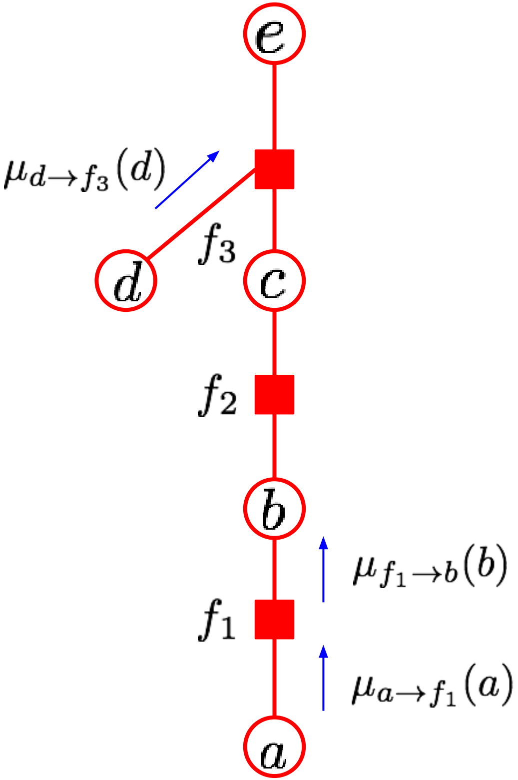
$$f_1(a, b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1, 2 \\ 1, 2 \end{bmatrix}$$

$$f_2(b, c) = \begin{bmatrix} 3, 2 \\ 1, 0.5 \end{bmatrix}$$

$$f_3(c = 0, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

$$f_3(c = 1, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$

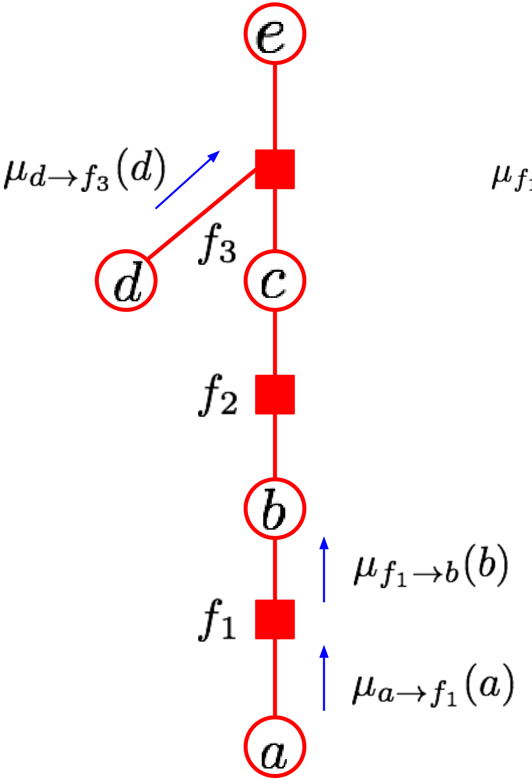
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$

$$\mu_{f_1\to b}(b) = \sum_a f_1(a,b)\mu_{a\to f_1(a)}(a)$$

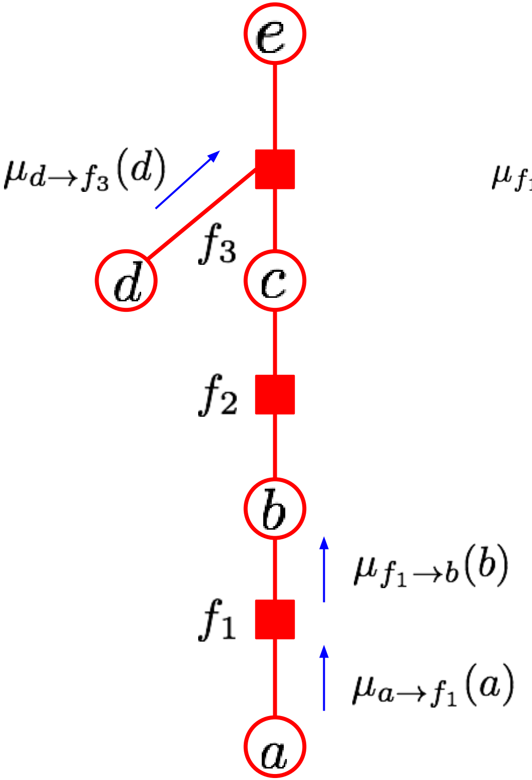
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$

$$\mu_{f_1\to b}(b) = \sum_a f_1(a,b)\mu_{a\to f_1(a)}(a)$$
$$= \begin{bmatrix}1,2\\1,2\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}3\\3\end{bmatrix}$$

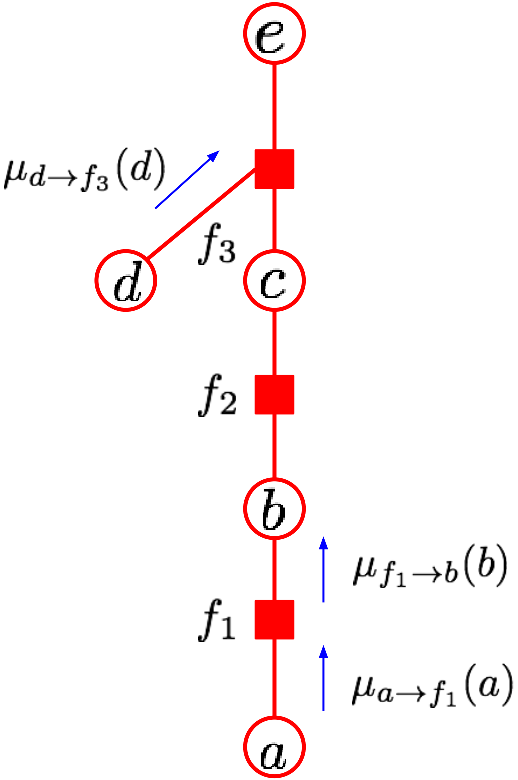
$$f_1(a,b) = \begin{bmatrix}f_1(0,0), f_1(1,0)\\ f_1(0,1), f_1(1,1)\end{bmatrix} = \begin{bmatrix}1,2\\1,2\end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix}3,2\\1,0.5\end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix}1,2\\2,1\end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix}1,2\\2,1\end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$

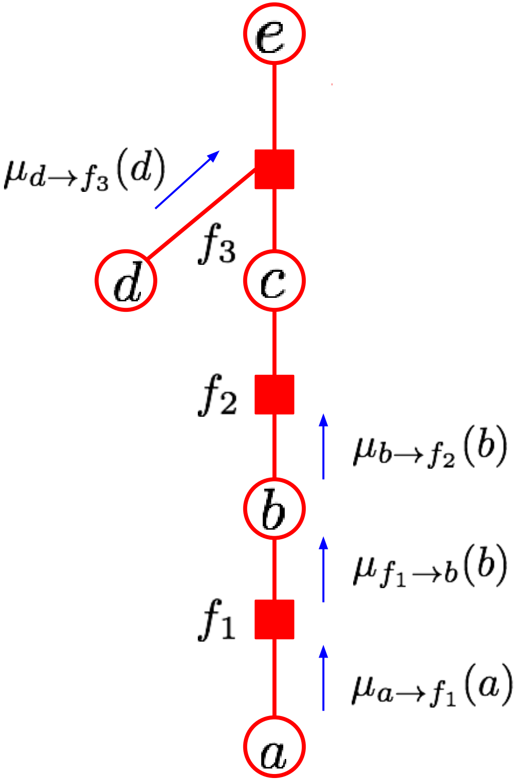
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$

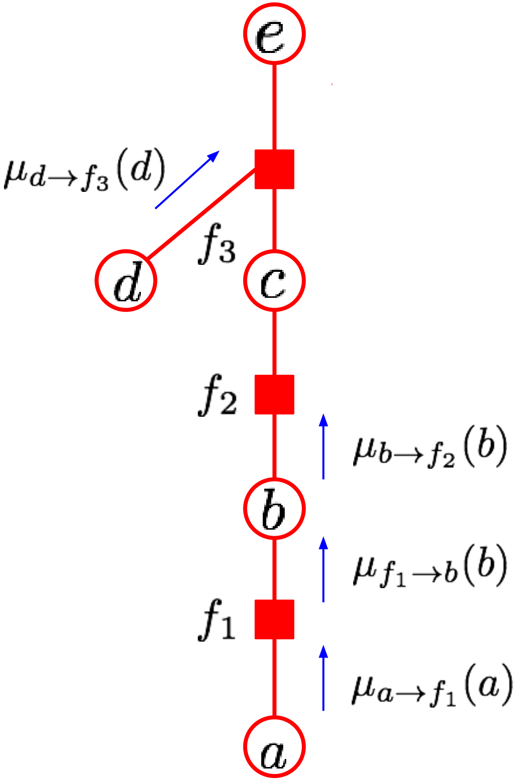
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$

$$\mu_{b\to f2}(b) = \mu_{f_1\to b}(b)$$

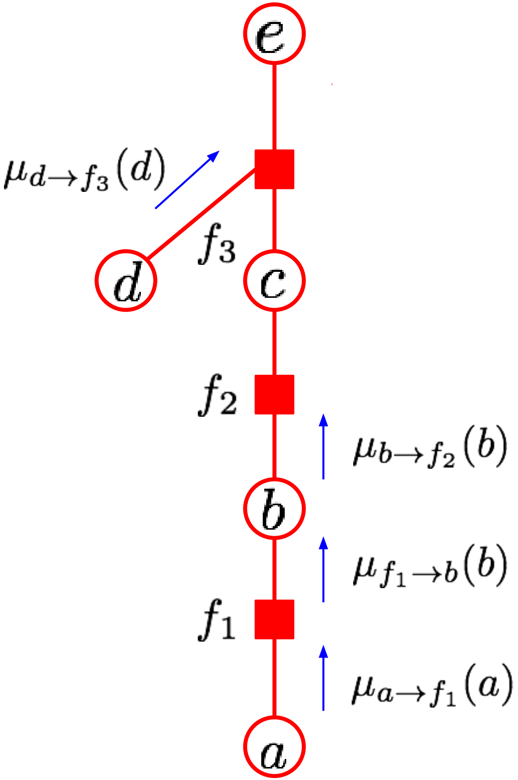
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$

$$\mu_{b\to f2}(b) = \mu_{f_1\to b}(b) = [3,3]^T$$

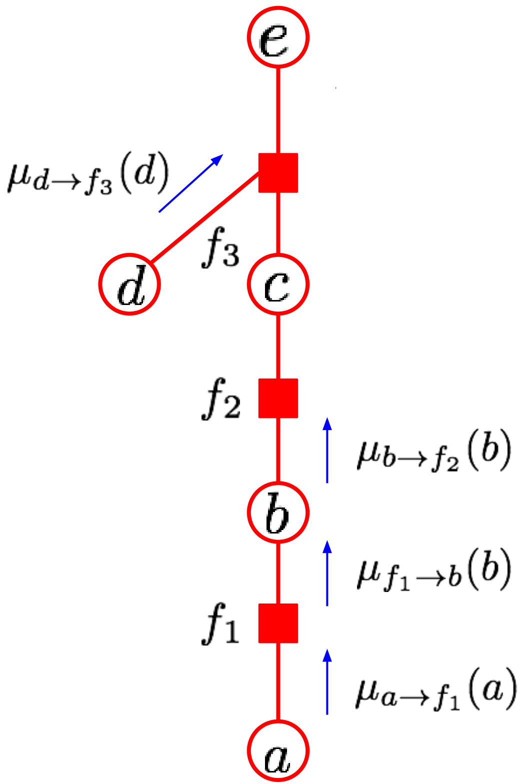
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$

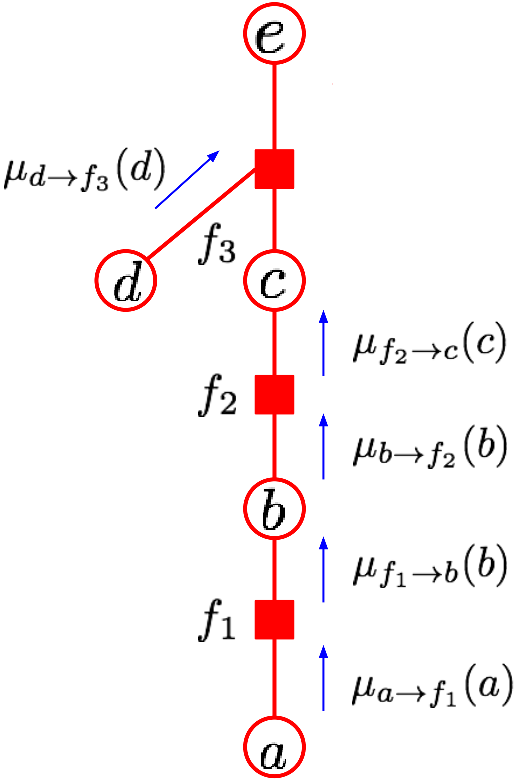
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$

$$\mu_{f_2\to c}(c) = \sum_b f_2(b,c)\mu_{b\to f_2}(b)$$
$$= \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix} \begin{bmatrix} 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 15 \\ 4.5 \end{bmatrix}$$

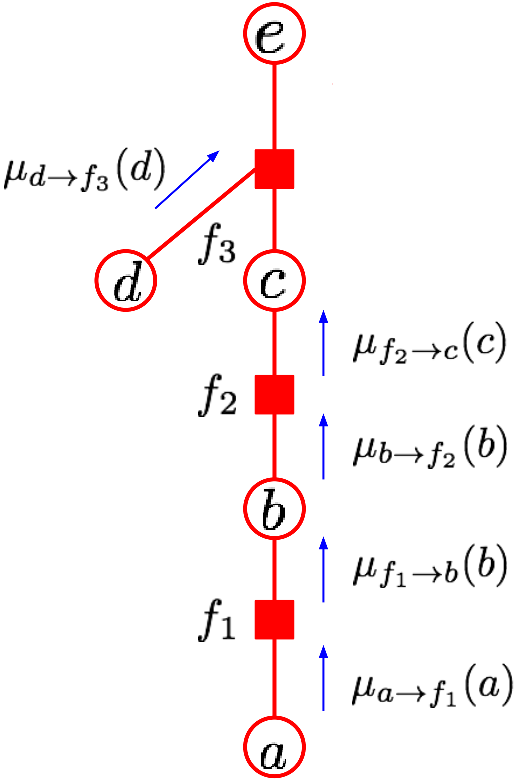
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$

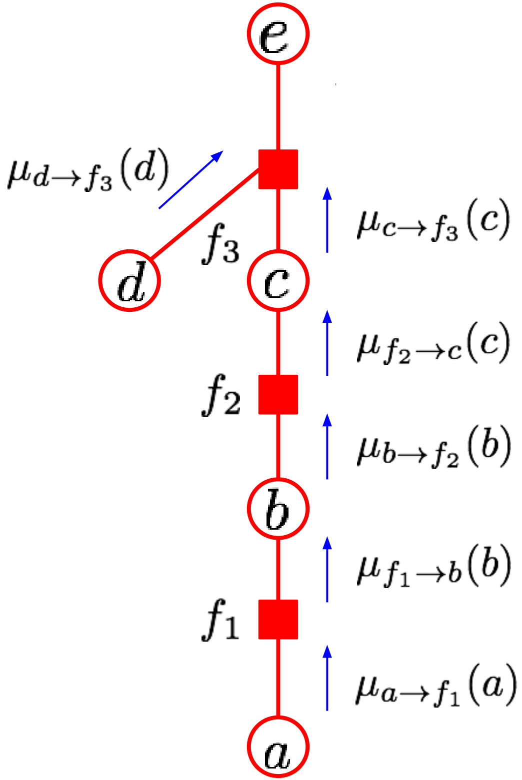
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [15,4.5]^T$$

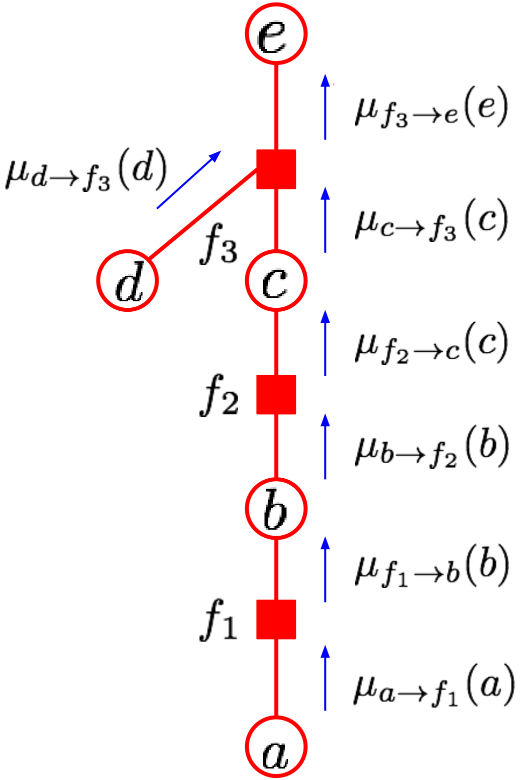
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [15,4.5]^T$$

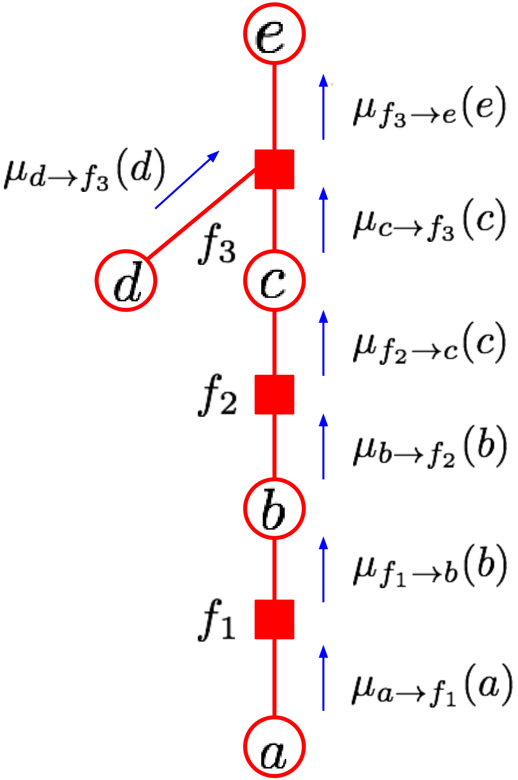
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [15,4.5]^T$$

$$\mu_{f_3\to e}(e)$$
$$=\sum_{c,d} f_3(c,d,e)\mu_{c\to f_3}(c)\mu_{d\to f_3}(d)$$
$$=15\begin{bmatrix}1,2\\2,1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}+4.5\begin{bmatrix}1,2\\2,1\end{bmatrix}\begin{bmatrix}1\\1\end{bmatrix}$$
$$=\begin{bmatrix}58.5\\58.5\end{bmatrix}$$

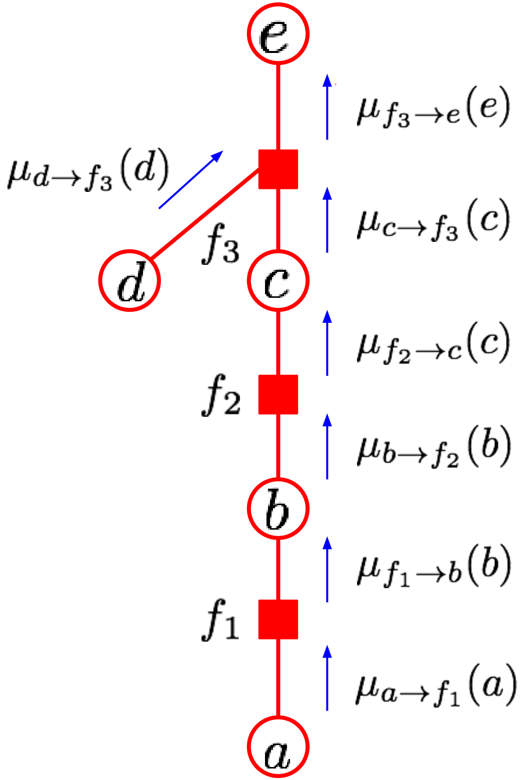
$$f_1(a,b) = \begin{bmatrix}f_1(0,0), f_1(1,0)\\f_1(0,1), f_1(1,1)\end{bmatrix} = \begin{bmatrix}1,2\\1,2\end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix}3,2\\1,0.5\end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix}1,2\\2,1\end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix}1,2\\2,1\end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a \to f_1}(a) = [1, 1]^T$$
$$\mu_{d \to f_3}(d) = [1, 1]^T$$
$$\mu_{f_1 \to b}(b) = [3, 3]^T$$
$$\mu_{b \to f2}(b) = [3, 3]^T$$
$$\mu_{f_2 \to c}(c) = [15, 4.5]^T$$
$$\mu_{c \to f_3}(c) = [15, 4.5]^T$$
$$\mu_{f_3 \to e}(e) = [58.5, 58.5]^T$$

$$f_1(a, b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1, 2 \\ 1, 2 \end{bmatrix}$$

$$f_2(b, c) = \begin{bmatrix} 3, 2 \\ 1, 0.5 \end{bmatrix}$$

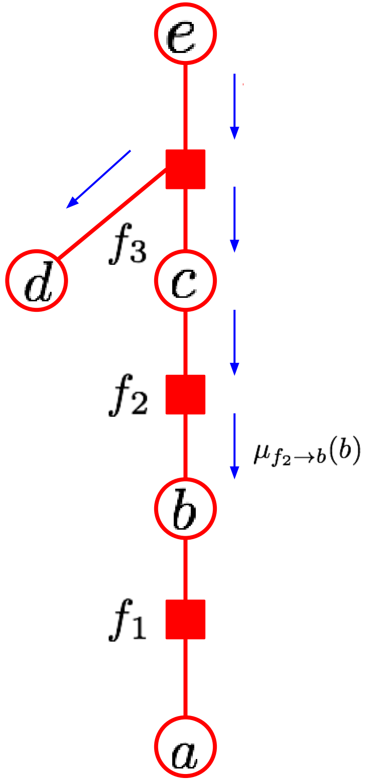
$$f_3(c = 0, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

$$f_3(c = 1, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

The sum product algorithm on factor graph

- How to start the sum-product algorithm:
  - Choose a node in the factor graph as the root node
  - Compute all the leaf-to-root messages
  - **Compute all the root-to-leaf messages**
- Initial conditions:
  - Starting from a factor leaf/root node, the initial factor-to-variable message is the factor itself
  - Starting from a variable leaf/root node, the initial variable-to-factor message is a vector of ones

Example of the sum product algorithm



$$\mu_{a \to f_1}(a) = [1, 1]^T$$
$$\mu_{d \to f_3}(d) = [1, 1]^T$$
$$\mu_{f_1 \to b}(b) = [3, 3]^T$$
$$\mu_{b \to f2}(b) = [3, 3]^T$$
$$\mu_{f_2 \to c}(c) = [15, 4.5]^T$$
$$\mu_{c \to f_3}(c) = [15, 4.5]^T$$
$$\mu_{f_3 \to e}(e) = [58.5, 58.5]^T$$
$$\mu_{e \to f_3}(e) = [1, 1]^T$$
$$\mu_{f_3 \to d}(e) = [58.5, 58.5]^T$$
$$\mu_{f_3 \to c}(c) = [6, 6]^T$$
$$\mu_{c \to f_2}(c) = [6, 6]^T$$

$$\mu_{f_2 \to b}(b) = \begin{bmatrix} 3, 2 \\ 1, 0.5 \end{bmatrix}^T \begin{bmatrix} 6 \\ 6 \end{bmatrix}$$
$$= [24, 15]^T$$

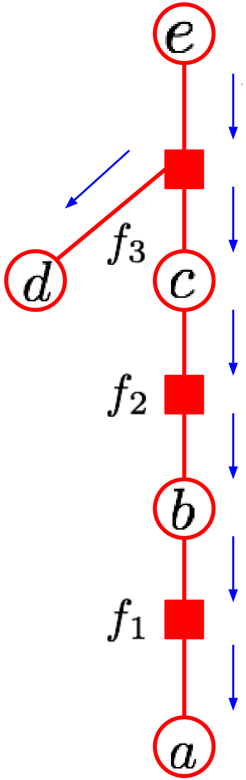
$$f_1(a, b) = \begin{bmatrix} f_1(0, 0), f_1(1, 0) \\ f_1(0, 1), f_1(1, 1) \end{bmatrix} = \begin{bmatrix} 1, 2 \\ 1, 2 \end{bmatrix}$$

$$f_2(b, c) = \begin{bmatrix} 3, 2 \\ 1, 0.5 \end{bmatrix}$$

$$f_3(c = 0, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

$$f_3(c = 1, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [15,4.5]^T$$
$$\mu_{f_3\to e}(e) = [58.5,58.5]^T$$
$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\mu_{f_3\to d}(e) = [58.5,58.5]^T$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\mu_{c\to f_2}(c) = [6,6]^T$$
$$\mu_{f_2\to b}(b) = [24,15]^T$$
$$\mu_{b\to f_1}(b) = [24,15]^T$$
$$\mu_{f_1\to a}(a) = [39,78]^T$$

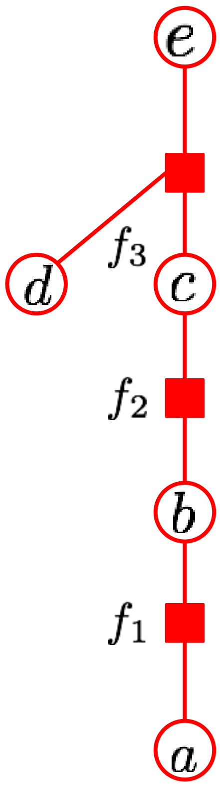
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [15,4.5]^T$$
$$\mu_{f_3\to e}(e) = [58.5,58.5]^T$$
$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\mu_{f_3\to d}(e) = [58.5,58.5]^T$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\mu_{c\to f_2}(c) = [6,6]^T$$
$$\mu_{f_2\to b}(b) = [24,15]^T$$
$$\mu_{b\to f_1}(b) = [24,15]^T$$
$$\mu_{f_1\to a}(a) = [39,78]^T$$

$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

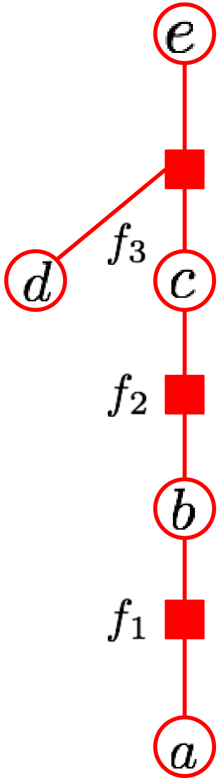
$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

How do we know we did it right?

Example of the sum product algorithm



We can verify the correctness by computing normalization constant Z

$$Z = \sum_{x_i} g(x_i)$$

$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [15,4.5]^T$$
$$\mu_{f_3\to e}(e) = [58.5,58.5]^T$$
$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\mu_{f_3\to d}(e) = [58.5,58.5]^T$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\mu_{c\to f_2}(c) = [6,6]^T$$
$$\mu_{f_2\to b}(b) = [24,15]^T$$
$$\mu_{b\to f_1}(b) = [24,15]^T$$
$$\mu_{f_1\to a}(a) = [39,78]^T$$

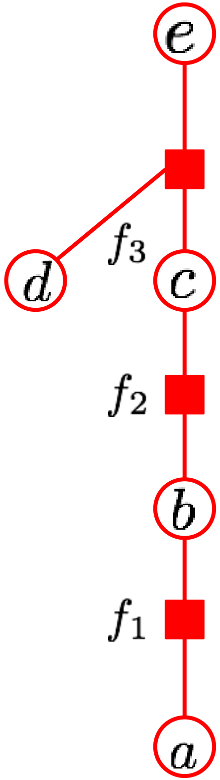
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



We can verify the correctness by computing normalization constant Z

$$Z = \sum_{x_i} g(x_i)$$

$$Z = \sum_{e} g(e) = \sum_{e} \mu_{f_3 \to e}(e) = 117$$

$$\mu_{a \to f_1}(a) = [1, 1]^T$$
$$\mu_{d \to f_3}(d) = [1, 1]^T$$
$$\mu_{f_1 \to b}(b) = [3, 3]^T$$
$$\mu_{b \to f2}(b) = [3, 3]^T$$
$$\mu_{f_2 \to c}(c) = [15, 4.5]^T$$
$$\mu_{c \to f_3}(c) = [15, 4.5]^T$$
$$\mu_{f_3 \to e}(e) = [58.5, 58.5]^T$$
$$\mu_{e \to f_3}(e) = [1, 1]^T$$
$$\mu_{f_3 \to d}(e) = [58.5, 58.5]^T$$
$$\mu_{f_3 \to c}(c) = [6, 6]^T$$
$$\mu_{c \to f_2}(c) = [6, 6]^T$$
$$\mu_{f_2 \to b}(b) = [24, 15]^T$$
$$\mu_{b \to f_1}(b) = [24, 15]^T$$
$$\mu_{f_1 \to a}(a) = [39, 78]^T$$

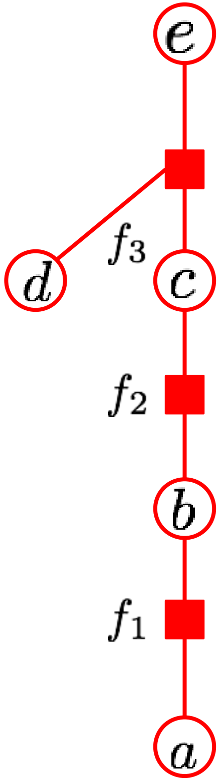
$$f_1(a, b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1, 2 \\ 1, 2 \end{bmatrix}$$

$$f_2(b, c) = \begin{bmatrix} 3, 2 \\ 1, 0.5 \end{bmatrix}$$

$$f_3(c = 0, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

$$f_3(c = 1, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

Example of the sum product algorithm



We can verify the correctness by computing normalization constant Z

$$Z = \sum_{x_i} g(x_i)$$

$$Z = \sum_e g(e) = \sum_e \mu_{f_3 \to e}(e) = 117$$

$$Z = \sum_b g(b)$$

$$= \sum_b \mu_{f_1 \to b}(b) \mu_{f_2 \to b}(b)$$

$$= 39 \times 3 = 117$$

$$\mu_{a \to f_1}(a) = [1,1]^T$$

$$\mu_{d \to f_3}(d) = [1,1]^T$$

$$\mu_{f_1 \to b}(b) = [3,3]^T$$

$$\mu_{b \to f2}(b) = [3,3]^T$$

$$\mu_{f_2 \to c}(c) = [15, 4.5]^T$$

$$\mu_{c \to f_3}(c) = [15, 4.5]^T$$

$$\mu_{f_3 \to e}(e) = [58.5, 58.5]^T$$

$$\mu_{e \to f_3}(e) = [1,1]^T$$

$$\mu_{f_3 \to d}(e) = [58.5, 58.5]^T$$

$$\mu_{f_3 \to c}(c) = [6,6]^T$$

$$\mu_{c \to f_2}(c) = [6,6]^T$$

$$\mu_{f_2 \to b}(b) = [24, 15]^T$$

$$\mu_{b \to f_1}(b) = [24, 15]^T$$

$$\mu_{f_1 \to a}(a) = [39, 78]^T$$

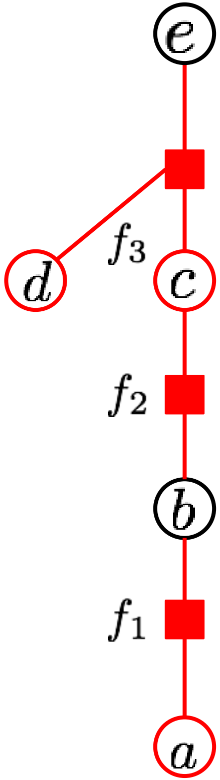
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



Naive computation:

$$Z = \sum_{a,b,c,d,e} P(a,b,c,d,e)$$

$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [15,4.5]^T$$
$$\mu_{f_3\to e}(e) = [58.5,58.5]^T$$
$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\mu_{f_3\to d}(e) = [58.5,58.5]^T$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\mu_{c\to f_2}(c) = [6,6]^T$$
$$\mu_{f_2\to b}(b) = [24,15]^T$$
$$\mu_{b\to f_1}(b) = [24,15]^T$$
$$\mu_{f_1\to a}(a) = [39,78]^T$$

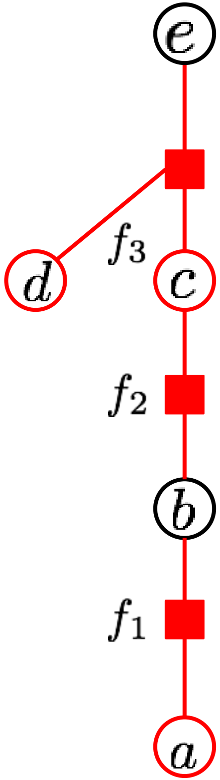
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



Naive computation:

$$Z = \sum_{a,b,c,d,e} P(a,b,c,d,e)$$

O(2^5)

$$\mu_{a \to f_1}(a) = [1,1]^T$$
$$\mu_{d \to f_3}(d) = [1,1]^T$$
$$\mu_{f_1 \to b}(b) = [3,3]^T$$
$$\mu_{b \to f2}(b) = [3,3]^T$$
$$\mu_{f_2 \to c}(c) = [15, 4.5]^T$$
$$\mu_{c \to f_3}(c) = [15, 4.5]^T$$
$$\mu_{f_3 \to e}(e) = [58.5, 58.5]^T$$
$$\mu_{e \to f_3}(e) = [1,1]^T$$
$$\mu_{f_3 \to d}(e) = [58.5, 58.5]^T$$
$$\mu_{f_3 \to c}(c) = [6,6]^T$$
$$\mu_{c \to f_2}(c) = [6,6]^T$$
$$\mu_{f_2 \to b}(b) = [24, 15]^T$$
$$\mu_{b \to f_1}(b) = [24, 15]^T$$
$$\mu_{f_1 \to a}(a) = [39, 78]^T$$

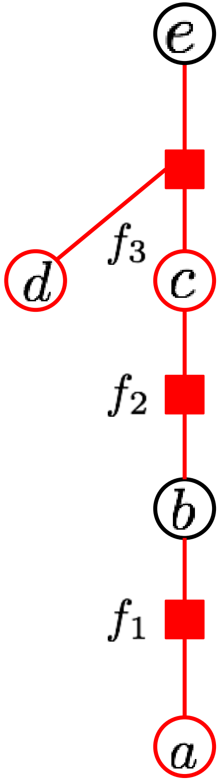
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



Naive computation:

$$Z = \sum_{a,b,c,d,e} P(a,b,c,d,e)$$

O(2^5)

Message-passing:

$$Z = \sum_e g(e)$$

O(2^3)
Dominated by the largest factor

$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [15,4.5]^T$$
$$\mu_{f_3\to e}(e) = [58.5,58.5]^T$$
$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\mu_{f_3\to d}(e) = [58.5,58.5]^T$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\mu_{c\to f_2}(c) = [6,6]^T$$
$$\mu_{f_2\to b}(b) = [24,15]^T$$
$$\mu_{b\to f_1}(b) = [24,15]^T$$
$$\mu_{f_1\to a}(a) = [39,78]^T$$

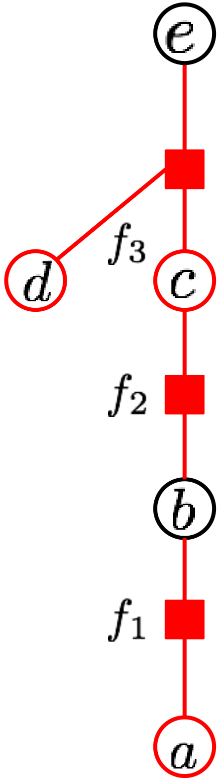
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



Naive computation:

$$Z = \sum_{a,b,c,d,e} P(a,b,c,d,e)$$

O(2^5)

Message-passing:

$$Z = \sum_{e} g(e)$$

O(2^3)
Dominated by the largest factor

$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [15,4.5]^T$$
$$\mu_{f_3\to e}(e) = [58.5,58.5]^T$$
$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\mu_{f_3\to d}(e) = [58.5,58.5]^T$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\mu_{c\to f_2}(c) = [6,6]^T$$
$$\mu_{f_2\to b}(b) = [24,15]^T$$
$$\mu_{b\to f_1}(b) = [24,15]^T$$
$$\mu_{f_1\to a}(a) = [39,78]^T$$

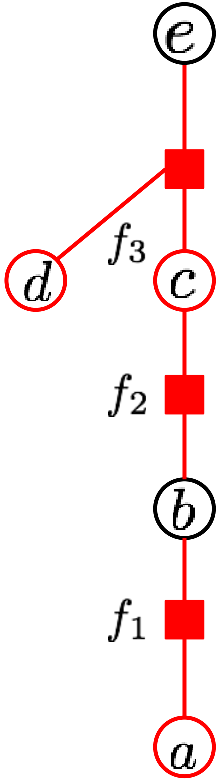
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



How do we compute P(e, b)?

$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [15,4.5]^T$$
$$\mu_{f_3\to e}(e) = [58.5,58.5]^T$$
$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\mu_{f_3\to d}(e) = [58.5,58.5]^T$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\mu_{c\to f_2}(c) = [6,6]^T$$
$$\mu_{f_2\to b}(b) = [24,15]^T$$
$$\mu_{b\to f_1}(b) = [24,15]^T$$
$$\mu_{f_1\to a}(a) = [39,78]^T$$

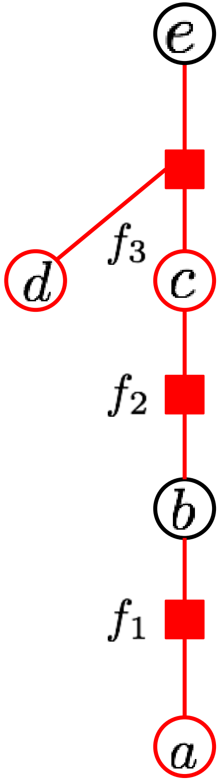
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



How do we compute P(e, b)?

We can eliminate the summation over b or e during message-passing

$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [15,4.5]^T$$
$$\mu_{f_3\to e}(e) = [58.5,58.5]^T$$
$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\mu_{f3\to d}(e) = [58.5,58.5]^T$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\mu_{c\to f_2}(c) = [6,6]^T$$
$$\mu_{f_2\to b}(b) = [24,15]^T$$
$$\mu_{b\to f_1}(b) = [24,15]^T$$
$$\mu_{f_1\to a}(a) = [39,78]^T$$

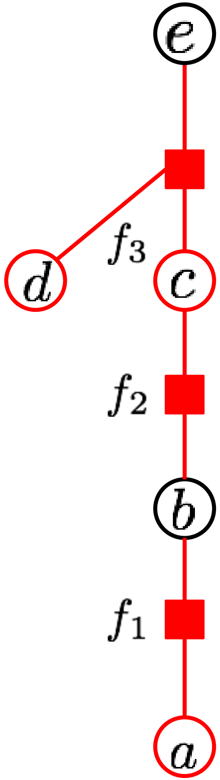
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



How do we compute P(e, b)?

We can eliminate the summation over b or e during message-passing

Assume choose e as an anchor and eliminate summation over b:

$$g(b, e) = \mu_{f_3 \to e}(b, e)$$

Or we can choose b as an anchor and eliminate summation over c:

$$g(b, e) = \mu_{f_2 \to b}(b, e)\mu_{f_1 \to b}(b)$$

$$\mu_{a \to f_1}(a) = [1, 1]^T \qquad f_1(a, b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1, 2 \\ 1, 2 \end{bmatrix}$$

$$\mu_{d \to f_3}(d) = [1, 1]^T$$

$$\mu_{f_1 \to b}(b) = [3, 3]^T$$

$$\mu_{b \to f2}(b) = [3, 3]^T \qquad f_2(b, c) = \begin{bmatrix} 3, 2 \\ 1, 0.5 \end{bmatrix}$$

$$\mu_{f_2 \to c}(c) = [15, 4.5]^T$$

$$\mu_{c \to f_3}(c) = [15, 4.5]^T \qquad f_3(c = 0, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

$$\mu_{f_3 \to e}(e) = [58.5, 58.5]^T$$

$$\mu_{e \to f_3}(e) = [1, 1]^T \qquad f_3(c = 1, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

$$\mu_{f_3 \to d}(e) = [58.5, 58.5]^T$$

$$\mu_{f_3 \to c}(c) = [6, 6]^T$$

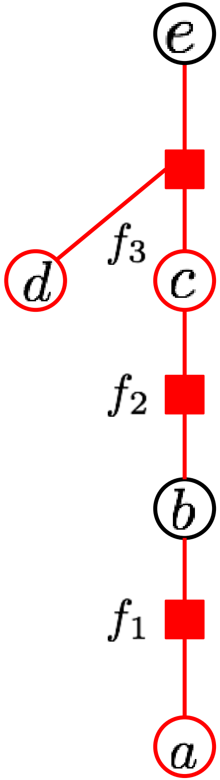
$$\mu_{c \to f_2}(c) = [6, 6]^T$$

$$\mu_{f_2 \to b}(b) = [24, 15]^T$$

$$\mu_{b \to f_1}(b) = [24, 15]^T$$

$$\mu_{f_1 \to a}(a) = [39, 78]^T$$

Example of the sum product algorithm



How do we compute P(e, b)?

Assume choose e as an anchor and eliminate summation over b:

$$g(b, e) = \mu_{f_3 \to e}(b, e)$$

Need to re-compute the messages to carry b along

$$\mu_{a \to f_1}(a) = [1, 1]^T \qquad f_1(a, b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1, 2 \\ 1, 2 \end{bmatrix}$$

$$\mu_{d \to f_3}(d) = [1, 1]^T$$

$$\mu_{f_1 \to b}(b) = [3, 3]^T$$

$$\mu_{b \to f2}(b) = [3, 3]^T \qquad f_2(b, c) = \begin{bmatrix} 3, 2 \\ 1, 0.5 \end{bmatrix}$$

$$\mu_{f_2 \to c}(c) = [15, 4.5]^T$$

$$\mu_{c \to f_3}(c) = [15, 4.5]^T \qquad f_3(c = 0, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

$$\mu_{f_3 \to e}(e) = [58.5, 58.5]^T$$

$$\mu_{e \to f_3}(e) = [1, 1]^T \qquad f_3(c = 1, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

$$\mu_{f_3 \to d}(e) = [58.5, 58.5]^T$$

$$\mu_{f_3 \to c}(c) = [6, 6]^T$$

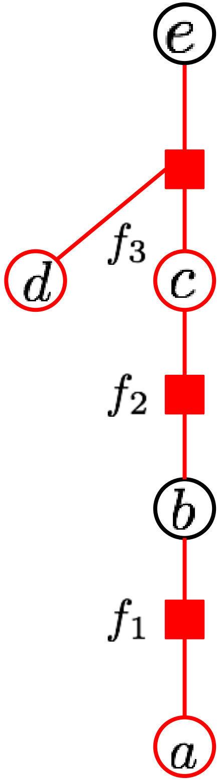
$$\mu_{c \to f_2}(c) = [6, 6]^T$$

$$\mu_{f_2 \to b}(b) = [24, 15]^T$$

$$\mu_{b \to f_1}(b) = [24, 15]^T$$

$$\mu_{f_1 \to a}(a) = [39, 78]^T$$

Example of the sum product algorithm



How do we compute P(e, b)?

Assume choose e as an anchor and eliminate summation over b:

$$g(b, e) = \mu_{f_3 \to e}(b, e)$$

Need to re-compute the messages to carry b along

$$\mu_{a \to f_1}(a) = [1, 1]^T$$
$$\mu_{d \to f_3}(d) = [1, 1]^T$$
$$\mu_{f_1 \to b}(b) = [3, 3]^T$$
$$\mu_{b \to f2}(b) = [3, 3]^T$$
$$\mu_{f_2 \to c}(c) = [15, 4.5]^T$$
$$\mu_{c \to f_3}(c) = [15, 4.5]^T$$
$$\mu_{f_3 \to e}(e) = [58.5, 58.5]^T$$
$$\mu_{e \to f_3}(e) = [1, 1]^T$$
$$\mu_{f_3 \to d}(e) = [58.5, 58.5]^T$$
$$\mu_{f_3 \to c}(c) = [6, 6]^T$$
$$\mu_{c \to f_2}(c) = [6, 6]^T$$
$$\mu_{f_2 \to b}(b) = [24, 15]^T$$
$$\mu_{b \to f_1}(b) = [24, 15]^T$$
$$\mu_{f_1 \to a}(a) = [39, 78]^T$$

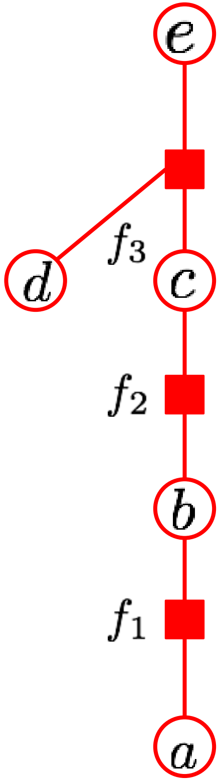
$$f_1(a, b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1, 2 \\ 1, 2 \end{bmatrix}$$

$$f_2(b, c) = \begin{bmatrix} 3, 2 \\ 1, 0.5 \end{bmatrix}$$

$$f_3(c = 0, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

$$f_3(c = 1, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

Example of the sum product algorithm



How do we compute P(e, b)?

$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$

New messages carrying b:

$$\mu_{f_2\to c}(b,c) \qquad \mu_{f_2\to c}(c) = [15, 4.5]^T$$
$$\mu_{c\to f_3}(b,c) \qquad \mu_{c\to f_3}(c) = [15, 4.5]^T$$
$$\mu_{f_3\to e}(b,e) \qquad \mu_{f_3\to e}(e) = [58.5, 58.5]^T$$

$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\mu_{f_3\to d}(e) = [58.5, 58.5]^T$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\mu_{c\to f_2}(c) = [6,6]^T$$
$$\mu_{f_2\to b}(b) = [24, 15]^T$$
$$\mu_{b\to f_1}(b) = [24, 15]^T$$
$$\mu_{f_1\to a}(a) = [39, 78]^T$$

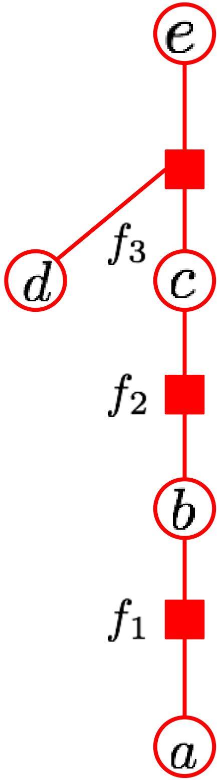
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



How do we compute P(e, b)?

New messages carrying b:

$$\mu_{f_2 \to c}(b, c)$$
$$\mu_{c \to f_3}(b, c)$$
$$\mu_{f_3 \to e}(b, e)$$

$$g(b, e) = \mu_{f_3 \to e}(b, e)$$

$$P(b, e) = \frac{1}{Z} \mu_{f_3 \to e}(b, e)$$

$$\mu_{a \to f_1}(a) = [1, 1]^T$$
$$\mu_{d \to f_3}(d) = [1, 1]^T$$
$$\mu_{f_1 \to b}(b) = [3, 3]^T$$
$$\mu_{b \to f2}(b) = [3, 3]^T$$
$$\mu_{f_2 \to c}(c) = [15, 4.5]^T$$
$$\mu_{c \to f_3}(c) = [15, 4.5]^T$$
$$\mu_{f_3 \to e}(e) = [58.5, 58.5]^T$$
$$\mu_{e \to f_3}(e) = [1, 1]^T$$
$$\mu_{f_3 \to d}(e) = [58.5, 58.5]^T$$
$$\mu_{f_3 \to c}(c) = [6, 6]^T$$
$$\mu_{c \to f_2}(c) = [6, 6]^T$$
$$\mu_{f_2 \to b}(b) = [24, 15]^T$$
$$\mu_{b \to f_1}(b) = [24, 15]^T$$
$$\mu_{f_1 \to a}(a) = [39, 78]^T$$

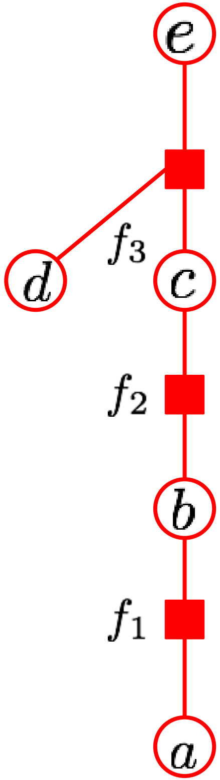
$$f_1(a, b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1, 2 \\ 1, 2 \end{bmatrix}$$

$$f_2(b, c) = \begin{bmatrix} 3, 2 \\ 1, 0.5 \end{bmatrix}$$

$$f_3(c = 0, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

$$f_3(c = 1, d, e) = \begin{bmatrix} 1, 2 \\ 2, 1 \end{bmatrix}$$

Example of the sum product algorithm



How do we compute P(e, b)?

New messages carrying b: $\mu_{f_2\to c}(b,c)$ $\mu_{c\to f_3}(b,c)$ $\mu_{f_3\to e}(b,e)$

$$\mu_{f_2\to c}(b,c) = f_2(b,c)\mu_{b\to f_2}(b) = \begin{bmatrix} 3\times 3 & 2\times 3 \\ 1\times 3 & 0.5\times 3 \end{bmatrix} = \begin{bmatrix} 9 & 6 \\ 3 & 1.5 \end{bmatrix}$$

$$\mu_{c\to f_3}(b,c) = \mu_{f_2\to c}(b,c) = \begin{bmatrix} 9 & 6 \\ 3 & 1.5 \end{bmatrix}$$

$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [15,4.5]^T$$
$$\mu_{f_3\to e}(e) = [58.5,58.5]^T$$
$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\mu_{f_3\to d}(e) = [58.5,58.5]^T$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\mu_{c\to f_2}(c) = [6,6]^T$$
$$\mu_{f_2\to b}(b) = [24,15]^T$$
$$\mu_{b\to f_1}(b) = [24,15]^T$$
$$\mu_{f_1\to a}(a) = [39,78]^T$$

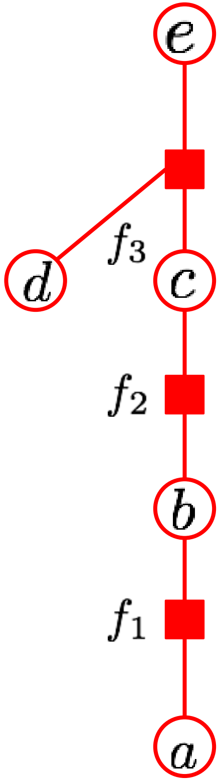
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



How do we compute P(e, b)?

New messages carrying b:

$$\mu_{f_2\to c}(b,c)$$
$$\mu_{c\to f_3}(b,c)$$
$$\mu_{f_3\to e}(b,e)$$

$$\mu_{c\to f_3}(b,c) = \mu_{f_2\to 3}(b,c) = \begin{bmatrix} 9 & 6 \\ 3 & 1.5 \end{bmatrix}$$

$$\mu_{f_3\to e}(b,e) = \sum_{c,d} f_3(c,d,e)\mu_{c\to f_3}(b,c)\mu_{d\to f_3}(d)$$

$$\mu_{f_{3\to e}}(b=0,e) = 9\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 3\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 36 \\ 36 \end{bmatrix}$$

$$\mu_{f_{3\to e}}(b=1,e) = 6\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} + 1.5\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 22.5 \\ 22.5 \end{bmatrix}$$

$$g(b,e) = \mu_{f_{3\to e}}(b,e) = \begin{bmatrix} 36 & 22.5 \\ 36 & 22.5 \end{bmatrix}$$

$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [15,4.5]^T$$
$$\mu_{f_3\to e}(e) = [58.5,58.5]^T$$
$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\mu_{f_3\to d}(e) = [58.5,58.5]^T$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\mu_{c\to f_2}(c) = [6,6]^T$$
$$\mu_{f_2\to b}(b) = [24,15]^T$$
$$\mu_{b\to f_1}(b) = [24,15]^T$$
$$\mu_{f_1\to a}(a) = [39,78]^T$$

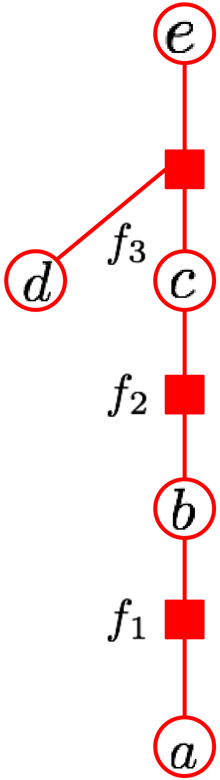
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



How do we compute P(e, b)?

$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$

$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

New messages carrying b: $\mu_{f_2\to c}(b,c)$, $\mu_{c\to f_3}(b,c)$, $\mu_{f_3\to e}(b,e)$

$$\mu_{f_2\to c}(c) = [15, 4.5]^T$$
$$\mu_{c\to f_3}(c) = [15, 4.5]^T$$
$$\mu_{f_3\to e}(e) = [58.5, 58.5]^T$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

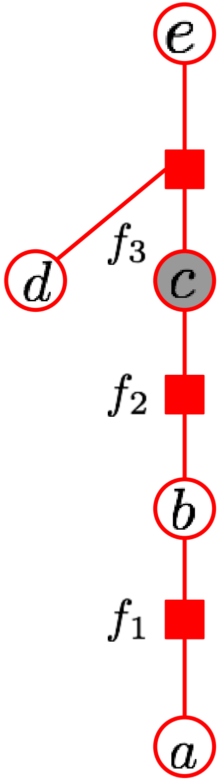
$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\mu_{f_3\to d}(e) = [58.5, 58.5]^T$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\mu_{c\to f_2}(c) = [6,6]^T$$
$$\mu_{f_2\to b}(b) = [24,15]^T$$
$$\mu_{b\to f_1}(b) = [24,15]^T$$
$$\mu_{f_1\to a}(a) = [39,78]^T$$

$$g(b,e) = \mu_{f_{3\to e}}(b,e) = \begin{bmatrix} 36 & 22.5 \\ 36 & 22.5 \end{bmatrix}$$

$$P(b,e) = \frac{1}{Z} g(b,e) = \frac{1}{117} \begin{bmatrix} 36 & 22.5 \\ 36 & 22.5 \end{bmatrix}$$

Example of the sum product algorithm



What if we observed c as c =[0,1]?

$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [15,4.5]^T$$
$$\mu_{f_3\to e}(e) = [58.5,58.5]^T$$
$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\mu_{f_3\to d}(e) = [58.5,58.5]^T$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\mu_{c\to f_2}(c) = [6,6]^T$$
$$\mu_{f_2\to b}(b) = [24,15]^T$$
$$\mu_{b\to f_1}(b) = [24,15]^T$$
$$\mu_{f_1\to a}(a) = [39,78]^T$$

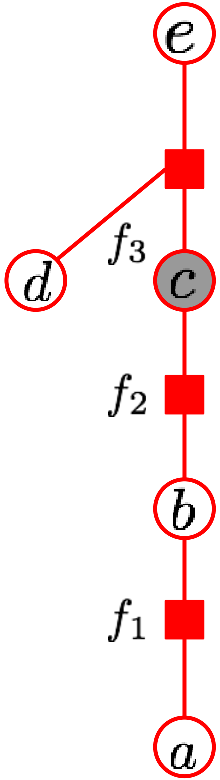
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



What if we observed c as c =[0,1]?

It changes the outgoing messages from c

$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [15,4.5]^T$$
$$\mu_{f_3\to e}(e) = [58.5,58.5]^T$$
$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\mu_{f_3\to d}(e) = [58.5,58.5]^T$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\mu_{c\to f_2}(c) = [6,6]^T$$
$$\mu_{f_2\to b}(b) = [24,15]^T$$
$$\mu_{b\to f_1}(b) = [24,15]^T$$
$$\mu_{f_1\to a}(a) = [39,78]^T$$

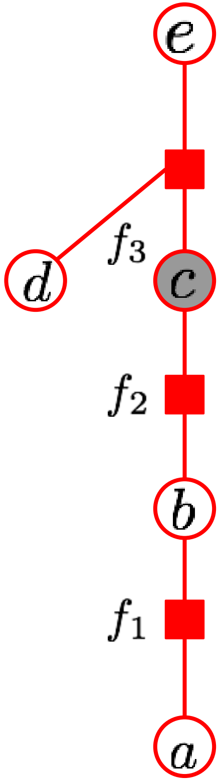
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



What if we observed c as c =[0,1]?

It changes the outgoing messages from c
And their derivative messages

$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\cancel{\mu_{c\to f_3}(c) = [15,4.5]^T}$$
$$\cancel{\mu_{f_3\to e}(e) = [58.5,58.5]^T}$$
$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\cancel{\mu_{f_3\to d}(e) = [58.5,58.5]^T}$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\cancel{\mu_{c\to f_2}(c) = [6,6]^T}$$
$$\cancel{\mu_{f_2\to b}(b) = [24,15]^T}$$
$$\cancel{\mu_{b\to f_1}(b) = [24,15]^T}$$
$$\cancel{\mu_{f_1\to a}(a) = [39,78]^T}$$

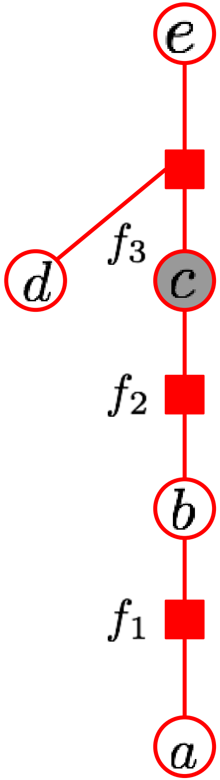
$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Example of the sum product algorithm



What if we observed c as c =[0,1]?

It changes the outgoing messages from c
And their derivative messages

$$\mu_{a\to f_1}(a) = [1,1]^T$$
$$\mu_{d\to f_3}(d) = [1,1]^T$$
$$\mu_{f_1\to b}(b) = [3,3]^T$$
$$\mu_{b\to f2}(b) = [3,3]^T$$
$$\mu_{f_2\to c}(c) = [15,4.5]^T$$
$$\mu_{c\to f_3}(c) = [0,1]^T$$
$$\cancel{\mu_{f_3\to e}(e) = [58.5,58.5]^T}$$
$$\mu_{e\to f_3}(e) = [1,1]^T$$
$$\cancel{\mu_{f_3\to d}(d) = [58.5,58.5]^T}$$
$$\mu_{f_3\to c}(c) = [6,6]^T$$
$$\mu_{c\to f_2}(c) = [0,1]^T$$
$$\cancel{\mu_{f_2\to b}(b) = [24,15]^T}$$
$$\cancel{\mu_{b\to f_1}(b) = [24,15]^T}$$
$$\cancel{\mu_{f_1\to a}(a) = [39,78]^T}$$

$$f_1(a,b) = \begin{bmatrix} f_1(0,0), f_1(1,0) \\ f_1(0,1), f_1(1,1) \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,2 \end{bmatrix}$$

$$f_2(b,c) = \begin{bmatrix} 3,2 \\ 1,0.5 \end{bmatrix}$$

$$f_3(c=0,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

$$f_3(c=1,d,e) = \begin{bmatrix} 1,2 \\ 2,1 \end{bmatrix}$$

Message-passing algorithms

- Why message-passing algorithms?

- Message-passing rules are local computation that usually has much lower computational complexity than global summation

- Computation complexity in message-passing is dominated by the factors with the largest number of neighbour nodes