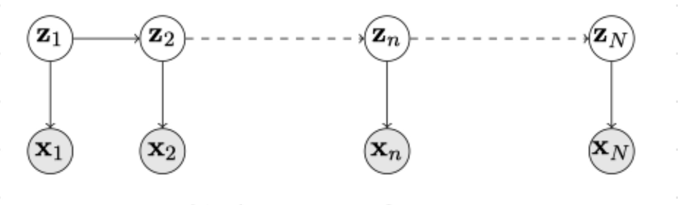
ML2 Home Assignment 4

id: 12179078

Volodymyr Medentsiy

Problem 1

Given BN with $X = \{x_1, \dots, x_N\}$, $Z = \{z_1, \dots, z_N\}$.

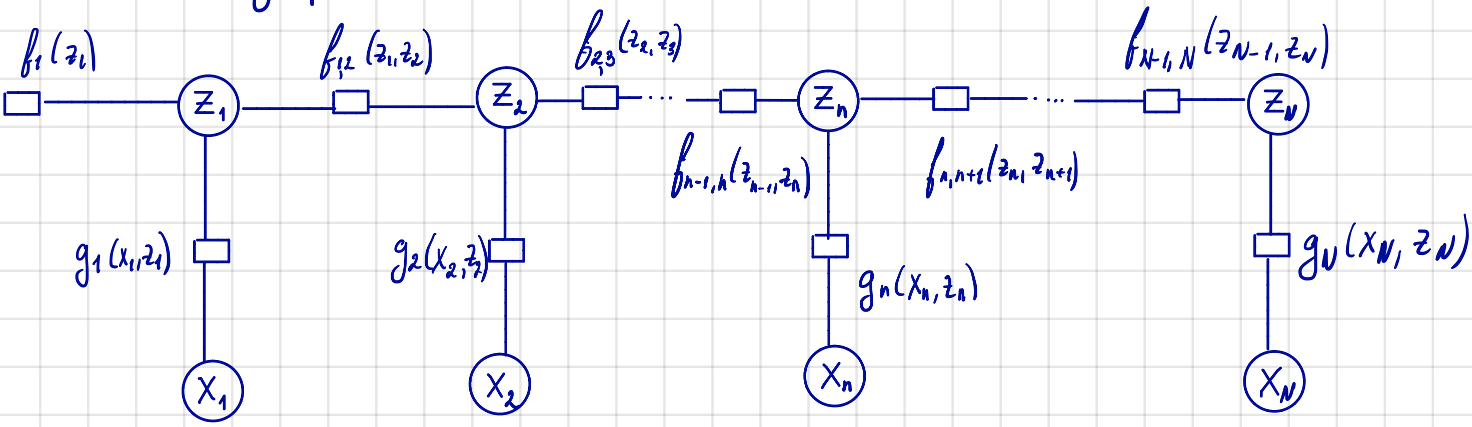


1. The factorized joint distribution $p(Z, X)$:

$$p(Z, X) = p(z_1) \cdot p(x_1 | z_1) \cdot p(z_2 | z_1)\, p(x_2 | z_2) \cdots p(z_N | z_{N-1}) \cdot p(x_N | z_N)$$

$$= p(z_1)\, p(x_1 | z_1) \prod_{i=2}^{N} p(z_i | z_{i-1})\, p(x_i | z_i)$$

2. Factor graph:



where $\{x_1, \dots, x_N, z_1, \dots, z_N\}$ – variables and $\{f_1(z_1), f_{12}(z_1, z_2), \dots, f_{N-1,N}(z_{N-1}, z_N),$ $g_1(x_1, z_1), \dots, g_N(x_N, z_N)\}$ – factors.

3.

$$p(X,Z) = \frac{1}{Z} f_1(z_1)\, f_{1,2}(z_1,z_2) \cdots f_{N-1,N}(z_{N-1},z_N) \cdot g_1(x_1,z_1) \cdots g_N(x_N,z_N) =$$

$$= \frac{1}{Z} f_1(z_1) \prod_{i=1}^{N-1} f_{i,i+1}(z_i,z_{i+1}) \cdot \prod_{i=1}^{N} g_i(x_i,z_i)$$

Because factor graph was derived from a directed graph $Z=1$ and

$f_1(z_1) = p(z_1)$, $f_{i,i+1}(z_i, z_{i+1}) = p(z_{i+1} | z_i)$ $\forall\, i = \overline{1, N-1}$ and

$g_i(x_i, z_i) = p(x_i | z_i)$ $\forall\, i = \overline{1, N}$

4. $$p(z_n | X) = \frac{p(X|z_n)\, p(z_n)}{p(X)} = \frac{\alpha(z_n)\,\beta(z_n)}{p(X)}$$

for $n=1$: $$p(z_1|X) = \frac{p(X|z_1)\,p(z_1)}{p(X)} = \left[x_1 \perp\!\!\!\perp x_j \,|\, z_1 \;\; \forall j > 1, \text{ because } 1 \text{ and } x_j \text{ are d-separated } z_1 \right] =$$

$$= \frac{p(x_1|z_1) \cdot p(x_2, \dots, x_N | z_1)\; p(z_1)}{p(X)}$$

for $n=2$:

$$p(z_2|X) = \frac{p(X|z_2)\,p(z_2)}{p(X)} = \left[\begin{array}{l} x_1 \perp\!\!\!\perp x_2 \,|\, z_2 \\ x_1 \perp\!\!\!\perp x_j \,|\, z_2 \\ x_2 \perp\!\!\!\perp x_j \,|\, z_2 \end{array} \;\; \forall j > 2 \right]$$

$$= \frac{p(x_1|z_2)\,p(x_2|z_2)\; p(x_3, \dots, x_N | z_2)\; p(z_2)}{p(X)}$$

So in general case: $$p(z_n|X) = \frac{p(z_n)}{p(X)}\, p(x_n|z_n)\, p(\{x_1, \dots, x_{n-1}\}|z_n) \cdot$$

$$\cdot\, p(\{x_{n+1}, \dots, x_N\}|z_n).$$

$$\alpha(z_n) = p(x_n|z_n)\, p(\{x_1, \dots, x_{n-1}\}|z_n) \cdot p(z_n) = p(x_1, \dots x_n, z_n)$$

$$\alpha(z_n) = \sum_{z_{n-1}} p(x_1, \dots, x_n, z_n | z_{n-1}) p(z_{n-1}) = \left[\{x_1, \dots, x_{n-1}\} \perp\!\!\!\perp \{x_n, z_n\} | z_{n-1} \right] =$$

$$= \sum_{z_{n-1}} p(x_1, \dots, x_{n-1} | z_{n-1}) p(x_n, z_{n-1}) p(z_{n-1}) = \sum_{z_{n-1}} p(x_1, \dots, x_{n-1}, z_{n-1}) p(x_n, z_n | z_{n-1})$$

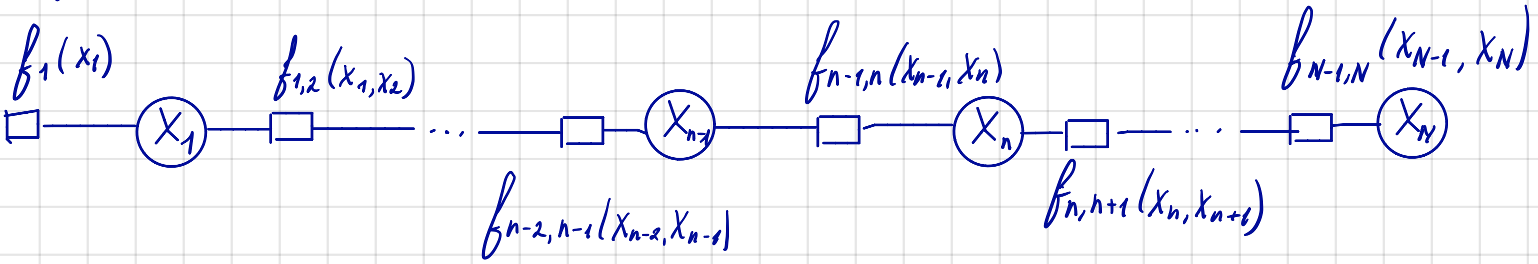
$$= \sum_{z_{n-1}} \alpha(z_{n-1}) p(x_n, z_n | z_{n-1})$$

$$\beta(z_n) = p(x_{n+1}, \dots x_N | z_n) = \sum_{z_{n+1}} p(x_{n+1}, \dots x_N, z_{n+1} | z_n) = \sum_{z_{n+1}} \frac{p(x_{n+1}, \dots x_N, z_n | z_{n+1}) p(z_{n+1})}{p(z_n)} =$$

$$= \left[\{x_{n+1}, z_n\} \perp\!\!\!\perp \{x_{n+2}, \dots x_N\} | z_{n+1} \right] = \sum_{z_{n+1}} \frac{p(x_{n+1}, z_n, z_{n+1}) \cdot p(x_{n+2} \dots x_N | z_{n+1})}{p(z_n)} = \sum_{z_{n+1}} \beta(z_{n+1}) p(x_{n+1}, z_{n+1} / z_n)$$

Problem 2

Let us draw a corresponding factor graph to apply sum-product algorithm:



with $f_{i-1,i}(x_{i-1}, x_i) = \psi_{i-1,i}(x_{i-1}, x_i) \quad \forall i = \overline{2, N}$

$f_1(x_1) = 1.$

$$\tilde{p}(x_n) = \prod_{s \in ne(x)} \mu_{f_s \to x_n}(x_n) = \mu_{f_{n-1,n} \to x_n}(x_n) \cdot \mu_{f_{n,n+1} \to x_n}(x_n) \quad \forall n = \overline{2, N-1}$$

$\tilde{p}(x_1) = \mu_{f_{1,2} \to x_1}(x_1)$ and $\tilde{p}(x_N) = \mu_{f_{N-1,N} \to x_N}(x_N).$

Lets run sum-product algorithm to find $\mu_{f_s \to x_n}(x_n)$. I choose x_N as a root node and x_1 as a leaf. So starting with a leaf node:

$$\mu_{x_1 \to f_{1,2}}(x_1) = 1$$

$$\mu_{f_{1,2} \to x_2}(x_2) = \sum_{x_1} f_1(x_1, x_2)$$

$$\mu_{x_2 \to f_{2,3}}(x_2) = \mu_{f_{1,2} \to x_2}(x_2)$$

$$\mu_{f_{2,3} \to x_3}(x_3) = \sum_{x_2} f_{2,3}(x_2, x_3) \cdot \mu_{x_2 \to f_{2,3}}(x_2).$$

so $\forall i = \overline{2, N-1}$: $\mu_{x_i \to f_{i,i+1}}(x_i) = \mu_{f_{i-1,i} \to x_i}(x_i)$

$$\mu_{f_{i,i+1} \to x_{i+1}}(x_{i+1}) = \sum_{x_i} f_{i,i+1}(x_i, x_{i+1}) \cdot \mu_{x_i \to f_{i,i+1}}(x_i) =$$

$$= \sum_{x_i} f_{i,i+1}(x_i, x_{i+1}) \cdot \mu_{f_{i-1,i} \to x_i}(x_i).$$

Now we can propagate messages from the root node to the leaf:

$$\mu_{x_N \to f_{N-1,N}}(x_N) = 1$$

$$\mu_{f_{N-1,N} \to x_{N-1}}(x_{N-1}) = \sum_{x_N} f_{N-1,N}(x_{N-1}, x_N)$$

$$\mu_{x_{N-1} \to f_{N-2,N-1}}(x_{N-1}) = \mu_{f_{N-1,N} \to x_{N-1}}(x_{N-1})$$

$$\mu_{f_{N-2,N-1} \to x_{N-2}}(x_{N-2}) = \sum_{x_{N-1}} f_{N-2,N-1}(x_{N-2}, x_{N-1}) \cdot \mu_{x_{N-1} \to f_{N-2,N-1}}(x_{N-1})$$

So general formula for messages propagated in the backward direction:

$$\mu_{x_i \to f_{i-1,i}}(x_i) = \mu_{f_{i,i+1} \to x_i}(x_i) \quad \forall i = \overline{1, N-1}$$

$$\mu_{f_{i,i+1} \to x_i}(x_i) = \sum_{x_{i+1}} f_{i,i+1}(x_i, x_{i+1}) \mu_{x_{i+1} \to f_{i,i+1}}(x_{i+1}) =$$

$$= \sum_{x_{i+1}} f_{i,i+1}(x_i, x_{i+1}) \mu_{f_{i+1,i+2} \to x_{i+1}}(x_{i+1})$$

Using the notation from the Figure 2:

$$\mu_{f_{n-1,n} \to x_n}(x_n) = \mu_\alpha(x_n) = \sum_{x_{n-1}} \psi_{n-1,n}(x_{n-1}, x_n) \mu_\alpha(x_{n-1})$$

$$\mu_{f_{n,n+1} \to x_n}(x_n) = \mu_\beta(x_n) = \sum_{x_{n+1}} \psi_{n+1,n}(x_{n+1}, x_n) \mu_\beta(x_{n+1})$$

So we have shown that $\tilde{p}(x_n) = \mu_\alpha(x_n) \cdot \mu_\beta(x_n)$, because we worked with an undirected graph $p(x_n) = \frac{1}{Z} \tilde{p}(x_n)$ where

Z is a normalizing coefficient.

PROBLEM 3

Considering that $p(x_n | x_N) \propto p(x_n, x_N)$ to evaluate $p(x_n | x_N)$ we can find $p(x_n, x_N)$.

$$p(x_n, x_N) = \sum_{X \setminus x_n, x_N} p(X) = \frac{1}{Z} \sum_{x_{n-1}} \psi_{n-1,n}(x_{n-1}, x_n) \cdot$$

$$\sum_{x_{n-2}} \psi_{n-2,n-1}(x_{n-2}, x_{n-1}) \cdot \ldots \sum_{x_1} \psi_{1,2}(x_1, x_2) \cdot \sum_{x_{n+1}} \psi_{n,n+1}(x_n, x_{n+1})$$

$$\sum_{x_{n+2}} \psi_{n+1,n+2}(x_{n+1}, x_{n+2}) \cdot \ldots \cdot \sum_{x_{N-1}} \psi_{N-2,N-1}(x_{N-2}, x_{N-1}) \psi_{N-1,N}(x_{N-1}, x_N) =$$

$$= \frac{1}{Z} \mu_\alpha(x_n) \cdot \tilde{\mu}_\beta(x_n, x_N).$$

with $\tilde{\mu}_\beta(x_n, x_N) =$

$$= \sum_{x_{n+1}} \psi_{n,n+1}(x_n, x_{n+1}) \ldots \cdot \sum_{x_{N-1}} \psi_{N-2,N-1}(x_{N-2}, x_{N-1}) \psi_{N-1,N}(x_{N-1}, x_N) =$$

$$= \sum_{x_{n+1}} \psi_{n,n+1}(x_n, x_{n+1}) \tilde{\mu}_\beta(x_{n+1}, x_N)$$

and $\tilde{\mu}_\beta(x_{N-1}, x_N) = \sum_{x_{N-1}} \psi_{N-2,N-1}(x_{N-2}, x_{N-1}) \psi_{N-1,N}(x_{N-1}, x_N)$

$$\mu_\alpha(x_n) = \sum_{x_{n-1}} \psi_{n-1,n}(x_{n-1}, x_n) \cdot \mu_\alpha(x_{n-1}).$$

μ_α does not change while μ_β should be adjusted.

Problem 4

Using Bishop's equations 8.61, 8.62, 8.63, 8.64, 8.66, 8.68, 8.69:

$$p(x_s) = \sum_{X \setminus x_s} f_s(x_s) \prod_{i \in ne(f_s)} \prod_{j \in ne(x_i) \setminus f_s} F_j(x_i, X_{ij}) =$$

$$= f_s(x_s) \prod_{i \in ne\, f(s)} \sum_{X \setminus x_s} \prod_{j \in ne(x_i) \setminus f_s} F_j(x_i, X_{ij}) =$$

$$= f_s(x_s) \prod_{i \in ne\, f(s)} \sum_{X \setminus x_s} G_i(x_i, X_{si}) =$$

$$= f_s(x_s) \prod_{i \in ne\, f(s)} \mu_{x_i \to f_s}(x_i)$$