Homework 3

JD Gallego Posada, *University of Amsterdam* 24/04/2017

Problem 1

Lemma
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $A, B, C \in \mathcal{F}$ be arbitrary events. Then, $\mathbb{P}(A|B \cap C) = \mathbb{P}(A|C)$ iff $A \perp\!\!\!\perp B|C$.
Proof

$$\mathbb{P}(A|B \cap C) = \frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(B \cap C)} = \frac{\mathbb{P}(A \cap B|C)\,\mathbb{P}(C)}{\mathbb{P}(B \cap C)} \overset{A \perp\!\!\!\perp B|C}{=} \frac{\mathbb{P}(A|C)\,\mathbb{P}(B|C)\,\mathbb{P}(C)}{\mathbb{P}(B \cap C)} = \mathbb{P}(A|C). \quad \blacksquare$$

1. All paths from $\mathbf{x}_i$ to $\mathbf{x}_n$ are blocked by $\mathbf{z}_n$ - case **a** - for $i < n \Rightarrow \mathbf{x}_1, \dots, \mathbf{x}_{n-1} \perp\!\!\!\perp \mathbf{x}_n|\mathbf{z}_n$.
2. All paths from $\mathbf{x}_i$ to $\mathbf{z}_n$ are blocked by $\mathbf{z}_{n-1}$ - case **a** - for $i < n \Rightarrow \mathbf{x}_1, \dots, \mathbf{x}_{n-1} \perp\!\!\!\perp \mathbf{z}_n|\mathbf{z}_{n-1}$.
3.

$$\begin{aligned} p(\mathbf{x}_{n+1}, \dots, \mathbf{x}_N, \mathbf{Z}) &= \sum_{x_1} \cdots \sum_{x_n} p(\mathbf{X}, \mathbf{Z}) = p(\mathbf{z}_1) \prod_{k=2}^{N} p(\mathbf{z}_k|\mathbf{z}_{k-1}) \sum_{x_1} \cdots \sum_{x_n} \prod_{m=1}^{N} p(\mathbf{x}_m|\mathbf{z}_m) \\ &= p(\mathbf{z}_1) \prod_{k=2}^{N} p(\mathbf{z}_k|\mathbf{z}_{k-1}) \prod_{m=n+1}^{N} p(\mathbf{x}_m|\mathbf{z}_m). \end{aligned}$$

$$\begin{aligned} \frac{p(\mathbf{x}_{n+1}, \dots, \mathbf{x}_N, \mathbf{z}_n, \mathbf{z}_{n+1})}{p(\mathbf{z}_{n+1})} &= \frac{1}{p(\mathbf{z}_{n+1})} \sum_{z_1} \cdots \sum_{z_{n-1}} \sum_{z_{n+2}} \cdots \sum_{z_N} p(\mathbf{z}_1) \prod_{k=2}^{N} p(\mathbf{z}_k|\mathbf{z}_{k-1}) \prod_{m=n+1}^{N} p(\mathbf{x}_m|\mathbf{z}_m) \\ &= \frac{1}{p(\mathbf{z}_{n+1})} p(\mathbf{z}_n) p(\mathbf{z}_{n+1}|\mathbf{z}_n) p(\mathbf{x}_{n+1}|\mathbf{z}_{n+1}) p(\mathbf{x}_{n+2}, \dots, \mathbf{x}_N|\mathbf{z}_{n+1}) \\ &= p(\mathbf{z}_n|\mathbf{z}_{n+1}) p(\mathbf{x}_{n+1}, \dots, \mathbf{x}_N|\mathbf{z}_{n+1}). \end{aligned}$$

4.

$$p(\mathbf{z}_{N+1}, \mathbf{X}|\mathbf{z}_N) = \frac{p(\mathbf{z}_{N+1}, \mathbf{X}, \mathbf{z}_N)}{p(\mathbf{z}_N)} = \frac{\sum_{\mathbf{z}_1} \cdots \sum_{\mathbf{z}_{N-1}} p(\mathbf{X}, \mathbf{Z})}{p(\mathbf{z}_N)} p(\mathbf{z}_{N+1}|\mathbf{z}_N) = p(\mathbf{X}|\mathbf{z}_N) p(\mathbf{z}_{N+1}|\mathbf{z}_N).$$

Problem 2

Note that the maximal cliques are of the form $\{x_i, x_{i+1}\}$. Thus, we can transform the given MRF into a FG where the corresponding potential function for the mentioned clique is $\psi_{i,i+1}(x_i, x_{i+1})$. Thus, let x_n be the root and start from the leaves:

$$\begin{aligned} \mu_{x_1 \to \psi_{1,2}}(x_1) &= 1 \\ \mu_{\psi_{1,2} \to x_2}(x_2) &= \sum_{x_1} \psi_{1,2}(x_1, x_2) \cdot 1 \\ \mu_{x_2 \to \psi_{2,3}}(x_2) &= \mu_{\psi_{1,2} \to x_2}(x_2) \\ &\vdots \end{aligned}$$

$$\begin{aligned} \mu_{x_N \to \psi_{N-1,N}}(x_N) &= 1 \\ \mu_{\psi_{N-1,N} \to x_{N-1}}(x_{N-1}) &= \sum_{x_N} \psi_{N-1,N}(x_{N-1}, x_N) \cdot 1 \\ \mu_{x_{N-1} \to \psi_{N-2,N-1}}(x_{N-1}) &= \mu_{\psi_{N-1,N} \to x_{N-1}}(x_{N-1}) \\ &\vdots \end{aligned}$$

Define,

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

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

Therefore, $\tilde{p}(x_n) = \mu_\alpha(x_n)\mu_\beta(x_n) \Rightarrow p(x_n) = \frac{1}{Z}\mu_\alpha(x_n)\mu_\beta(x_n)$.

Problem 3

If we observe $\mathbf{x}_N$ to have a value $\hat{\mathbf{x}}_N$, this has the same effect as multiplying the joint distribution by $\mathbb{1}\left[\mathbf{x}_N = \hat{\mathbf{x}}_N\right]$. Since $\tilde{p}(\mathbf{x})$ is a product over the potential functions, we can collect this factor into the factor $\psi_{N-1,N}$.

Summations over $\mathbf{x}_N$ then contain only one term in which $\mathbf{x}_N = \hat{\mathbf{x}}_N$. Note that for calculating $p(\mathbf{x}_n|\mathbf{x}_N)$ we can apply the same message passing algorithm and renormalize by $p(\mathbf{x}_N = \hat{\mathbf{x}}_N)$.

Problem 4

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

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

$$\begin{aligned}
\tilde{p}(x_1, x_2) &= \sum_{\mathbf{x}\backslash\{x_1, x2\}} \tilde{p}(\mathbf{x}) = \sum_{x_3}\sum_{x_4} f_a(x_1, x_2) f_b(x_2, x_3) f_c(x_2, x_4) \\
&= f_a(x_1, x_2)\left(\sum_{x_3} f_b(x_2, x_3)\right)\left(\sum_{x_4} f_c(x_2, x_4)\right) \\
&= f_a(x_1, x_2)\mu_{f_b \to x_2}(x_2)\mu_{f_c \to x_2}(x_2) \\
&= f_a(x_1, x_2)\mu_{x_2 \to f_a}(x_2)(1) \\
&= f_a(x_1, x_2) \prod_{i \in ne(f_a)} \mu_{x_i \to f_a}(x_i).
\end{aligned}$$

Problem 5

Following Bishop's notation in section 8.4,

$$\begin{aligned}
p(\mathbf{x}_s) &= f_s(\mathbf{x}_s) \prod_{i \in ne(f_s)} \mu_{x_i \to f_s}(x_i) \\
&= f_s(\mathbf{x}_s) \prod_{i \in ne(f_s)} \prod_{l \in ne(x_i)\backslash f_s} \mu_{f_l \to x_i}(x_i) \\
&= f_s(\mathbf{x}_s) \prod_{i \in ne(f_s)} \prod_{l \in ne(x_i)\backslash f_s} \sum_{X_l} F_l(x_i, X_l) \\
&= \sum_{\mathbf{x}\backslash\mathbf{x}_s} f_s(\mathbf{x}_s) \prod_{i \in ne(f_s)} \prod_{l \in ne(x_i)\backslash f_s} F_l(x_i, X_l) \\
&= \sum_{\mathbf{x}\backslash\mathbf{x}_s} p(\mathbf{x}).
\end{aligned}$$