Problem 1

(a) Let $\mathsf{x} = \mathsf{z} + \mathsf{w}_1$ and $\mathsf{y} = \mathsf{z} + \mathsf{w}_2$. From the directed graphical model in Figure 1, we have $\mathsf{x} \not\perp\!\!\!\perp \mathsf{y}$ while $\mathsf{x} \perp\!\!\!\perp \mathsf{y}|\mathsf{z}$.

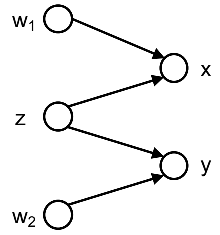

Figure 1: Directed graphical model for 1(a).

(b) "$\Longrightarrow$":

Since

$$\begin{aligned} p_{\mathsf{x},\mathsf{y},\mathsf{z}}(x,y,z) &= p_{\mathsf{x},\mathsf{y}|\mathsf{z}}(x,y|z)p_{\mathsf{z}}(z) \\ &= p_{\mathsf{x}|\mathsf{z}}(x|z)p_{\mathsf{y}|\mathsf{z}}(y|z)p_{\mathsf{z}}(z), \end{aligned}$$

we can assign $h(x,z) = p_{\mathsf{x}|\mathsf{z}}(x|z)$ and $g(y,z) = p_{\mathsf{y}|\mathsf{z}}(x|z)p_{\mathsf{z}}(z)$.

"$\Longleftarrow$":

By marginalizing x and y,

$$\begin{aligned} p_{\mathsf{x}|\mathsf{z}}(x|z) &= \sum_y p_{\mathsf{x},\mathsf{y}|\mathsf{z}}(x,y|z) \\ &= \sum_y \frac{p_{\mathsf{x},\mathsf{y},\mathsf{z}}(x,y,z)}{p_{\mathsf{z}}(z)} \\ &= \frac{h(x,z)}{p_{\mathsf{z}}(z)} \sum_y g(y,z), \end{aligned}$$

$$\begin{aligned} p_{\mathsf{y}|\mathsf{z}}(y|z) &= \sum_x p_{\mathsf{x},\mathsf{y}|\mathsf{z}}(x,y|z) \\ &= \sum_x \frac{p_{\mathsf{x},\mathsf{y},\mathsf{z}}(x,y,z)}{p_{\mathsf{z}}(z)} \\ &= \frac{g(y,z)}{p_{\mathsf{z}}(z)} \sum_x h(x,z). \end{aligned}$$

Since $\sum_x p_{\mathsf{x}|\mathsf{z}}(x|z) = 1$,

$$\frac{\sum_x h(x,z) \sum_y g(y,z)}{p_{\mathsf{z}}(z)} = 1.$$

Thus,

$$\begin{aligned} p_{x|z}(x|z)p_{y|z}(y|z) &= \frac{h(x,z)}{p_z(z)} \sum_y g(y,z) \frac{g(y,z)}{p_z(z)} \sum_x h(x,z) \\ &= \frac{\sum_x h(x,z) \sum_y g(y,z)}{p_z(z)} \frac{h(x,z)g(y,z)}{p_z(z)} \\ &= \frac{h(x,z)g(y,z)}{p_z(z)} \\ &= \frac{p_{x,y,z}(x,y,z)}{p_z(z)} \\ &= p_{x,y|z}(x,y|z). \end{aligned}$$

□

Problem 2

(a) Since $\mathcal{S}$ is a finite set, for convenience and without loss of generality, we can write $\mathcal{S} = \{1, 2, ..., N\}$, where $N = |\mathcal{S}|$.

In preparation, we compute $\gamma(1) = \mu(1)$, $\gamma(2) = \mu(1) + \mu(2)$, $\ldots$, $\gamma(N) = \mu(1) + \mu(2) + \cdots + \mu(N) = 1$. Let $\gamma(0) = 0$. This preparation step takes $O(|\mathcal{S}|)$.

Notice

$$\begin{aligned} \mathbb{P}\Big(0 \le UN \le \gamma(1)\Big) &= \mu(1), \\ \mathbb{P}\Big(\gamma(1) < UN \le \gamma(2)\Big) &= \mu(2), \\ &\cdots \\ \mathbb{P}\Big(\gamma(i-1) < UN \le \gamma(i)\Big) &= \mu(i), \\ &\cdots \\ \mathbb{P}\Big(\gamma(N-1) < UN \le 1\Big) &= \mu(N). \end{aligned}$$

Thus, we sample $U \sim \text{unif}([0,1])$, check which interval UN falls in (i.e., sandwiched by which γ interval). The $\gamma(i)$ interval UN falls in corresponds to the index i of element in $\mathcal{S}$ to output. This takes $O(|\mathcal{S}|)$ to check, as there are N such intervals.

(b) Notice

$$\begin{aligned} p(x_1, x_2, \cdots, x_n) &= p(x_1|x_2, \cdots, x_n)\; p(x_2|x_3, \cdots, x_n)\; \cdots p(x_{n-1}|x_n)\; p(x_n) \\ &= \frac{p(x_1, x_2, \cdots, x_n)}{p(x_2, x_3, \cdots, x_n)} \frac{p(x_2, x_3, \cdots, x_n)}{p(x_3, x_4, \cdots, x_n)} \cdots \frac{p(x_{n-1})}{p(x_n)} p(x_n). \end{aligned}$$

Now, from the Black-box A we have access to arbitrary marginals, which means we can compute each of the nominators and denominators above. Thus, we can compute each term of the above telescoping factorization. This involves $2n - 1 = O(n)$ calls of Black-box A.

For each of the terms, we can then use (a) to sample and the computation time is now $O(n|\mathcal{X}|)$.

(c)

(d)

Problem 3

(a)(i) From the given values of $p(x_1, x_2, x_3, x_4)$, we see

$$\begin{aligned} p(1,0,0,0) &= p(0,0,0,1) = 1/8, \\ p(1,1,0,0) &= p(0,0,1,1) = 1/8, \\ p(1,1,1,0) &= p(0,1,1,1) = 1/8, \\ p(0,0,0,0) &= p(0,0,0,0) = 1/8, \\ p(1,1,1,1) &= p(1,1,1,1) = 1/8, \end{aligned}$$

and all other combinations are all with 0 probability. Thus, $p(x_1, x_2, x_3, x_4) = (x_4, x_3, x_2, x_1)$. □

(ii) Notice that

$$\begin{aligned} (x_2, x_4) = (0,0) &\Longrightarrow x_3 = 0, \\ (x_2, x_4) = (1,0) &\Longrightarrow x_1 = 1, \\ (x_2, x_4) = (0,1) &\Longrightarrow x_1 = 0, \\ (x_2, x_4) = (1,1) &\Longrightarrow x_3 = 1, \end{aligned}$$

while the other value in x_1 or x_3 is undetermined. Thus, this proves $x_1 \perp\!\!\!\perp x_3 | x_2, x_4$. □

(b) Assume Hammersley-Clifford factorization exists, then

$$\begin{aligned} p(0,0,0,0) &= \phi_{12}(0,0)\ \phi_{23}(0,0)\ \phi_{34}(0,0)\ \phi_{41}(0,0) = 1/8, \\ p(0,0,1,0) &= \phi_{12}(0,0)\ \phi_{23}(0,1)\ \phi_{34}(1,0)\ \phi_{41}(0,0) = 0, \\ p(0,0,1,1) &= \phi_{12}(0,0)\ \phi_{23}(0,1)\ \phi_{34}(1,1)\ \phi_{41}(0,0) = 1/8, \\ p(1,1,1,0) &= \phi_{12}(1,1)\ \phi_{23}(1,1)\ \phi_{34}(1,0)\ \phi_{41}(0,1) = 1/8. \end{aligned}$$

Here, $p(0,0,1,0) = 0$ means one of the terms in $\phi_{12}(0,0)\ \phi_{23}(0,1)\ \phi_{34}(1,0)\ \phi_{41}(0,0)$ needs to equal 0.

However, $\phi_{12}(0,0) \neq 0$ because otherwise $p(0,0,0,0)$ would have been 0; $\phi_{23}(0,1) \neq 0$ because otherwise $p(0,0,1,1)$ would have been 0; $\phi_{34}(1,0) \neq 0$ because otherwise $p(1,1,1,0)$ would have been 0; $\phi_{41}(0,0) \neq 0$ because otherwise $p(0,0,1,1)$ again would have been 0. This contradicts with the assumption. □

Problem 4

(a) The diagram of the undirected graphical model is shown in Figure 2. We write the potential as

$$p_{\mathsf{x},\mathsf{t}}(x,t) = \phi_{\mathsf{x},\mathsf{t}}(x,t) = \mathbb{1}_{x=1}\, p_{\mathsf{t}}(t)\, c_t + \mathbb{1}_{x=0}\, p_{\mathsf{t}}(t)\, (1 - c_t).$$

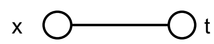

Figure 2: Undirected graphical model for 4(a).

(b) The diagram of the undirected graphical model is shown in Figure 3. The potential function between x and t, $\phi_{\mathsf{x},\mathsf{t}}(x,t)$, is the same as (a). We write the potential between t and each y_i as

$$\phi_{\mathsf{t},\mathsf{y}_i}(t, y_i) = p_{\mathsf{t}}(t)\, \exp\left[\sum_{k=1}^{K} \mathbb{1}_{t=k}\, \theta_i^k\, y_i\right].$$

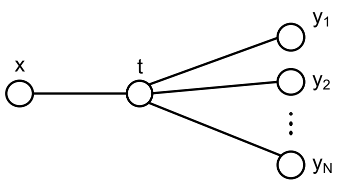

Figure 3: Undirected graphical model for 4(b).