Tutorial 3

1.1

$A \perp C \mid \{B, D\}$ can be modelled with figure 1, because they are d-separated.

$B \perp D \mid \{A, C\}$ cannot be modelled, because C is a collider and conditioning on C d-connects B and D. They would be independent given only A without C: $B \perp D \mid \{A\}$.

1.2

a)

A clique is a fully connected subgraph, i.e. a set of nodes, where every node is connected to every other node.

b)

In G1: {A,C}, {C,B}

In G2: {A,D}, {A,B}, {C,D}, {C,B}, {D,B}, {A,D,B}, {C,D,B}

1.3

a) Independencies in G3 given x4 and x5

$\{x1, x2, x3\} \perp \{x6, x7, x8\} \mid \{x4, x5\}$

b) Markov Blanket for x2 and x6 in G3

In a MRF the Markov Blanket are the direct neighbours in the graph.

MB for x2: {x1, x3, x5}

MB for x6: {x4, x5, x7, x8}

MB for {x2, x6}: {x1, x3, x5, x4, x5, x7, x8}

1.4

p(a|c) = p(a,c) / p(c)

p(a,c) = Z^-1 ∑b φ(a,b) φ(b,c)

p(c) as in exercise defined.

1.5

a)

p(d|c) = p(d,c) | p(c)

$$p(d,c) = Z^{-1}\sum_a\sum_b\sum_e \Phi(a,c)\Phi(b,c)\Phi(c,d)\Phi(d,e)$$

$$\mu_a(c) = \sum_a \Phi(a,c)$$

$$\mu_b(c) = \sum_b \Phi(b,c)$$

$$\mu_e(d) = \sum_e \Phi(d,e)$$

$$p(d,c) = Z^{-1}\Phi(c,d)\mu_a(c)\mu_b(c)\mu_e(d)$$

$$p(c) = Z^{-1}\sum_a\sum_b\sum_d\sum_e \Phi(a,c)\Phi(b,c)\Phi(c,d)\Phi(d,e)$$

$$\mu_d(c) = \sum_d \Phi(c,d)\mu_e(d)$$

$$p(c) = Z^{-1}\mu_d(c)\mu_a(c)\mu_b(c)$$

$$p(d|c) = \frac{Z^{-1}\Phi(c,d)\mu_a(c)\mu_b(c)\mu_e(d)}{Z^{-1}\mu_d(c)\mu_a(c)\mu_b(c)}$$

$$p(d|c) = \frac{\Phi(c,d)\mu_e(d)}{\mu_d(c)}$$

b)

p(a|d) = p(a,d) / p(d)

$$p(a,d) = Z^{-1}\sum_b\sum_c\sum_e \Phi(a,c)\Phi(b,c)\Phi(c,d)\Phi(d,e)$$

$$\mu_e(d) = \sum_e \Phi(d,e)$$

$$\mu_b(c) = \sum_b \Phi(b,c)$$

$$p(a,d) = Z^{-1} \sum_c \Phi(a,c)\Phi(c,d)\mu_b(c)\mu_e(d)$$

$$p(d) = Z^{-1} \sum_a \sum_b \sum_c \sum_e \Phi(a,c)\Phi(b,c)\Phi(c,d)\Phi(d,e)$$

$$\mu_a(c) = \sum_a \Phi(a,c)$$

$$\mu_c(d) = \sum_c \Phi(c,d)\mu_b(c)\mu_a(c)$$

$$p(d) = Z^{-1}\mu_c(d)\mu_e(d)$$

$$p(a|d) = \frac{Z^{-1}\sum_c \Phi(a,c)\Phi(c,d)\mu_b(c)\mu_e(d)}{Z^{-1}\mu_c(d)\mu_e(d)}$$

$$p(a|d) = \frac{\sum_c \Phi(a,c)\Phi(c,d)\mu_b(c)}{\mu_c(d)}$$

2.1

$$p(x,y) = Z^{-1} \sum_{\mathrm{i}} \sum_j \Phi(x_i, y_i)\Phi(x_i, x_j)$$

$$\mathrm{p(x,y)} = \mathrm{Z}^{-1} exp\left(ln\left(\sum_i exp(-\tau|\mathrm{x_i} - \mathrm{y_i}|) \sum_{i,j} exp(-\lambda[x_i \neq x_j]) \right)\right)$$

$$p(x,y) = Z^{-1} exp\left(\sum_i (-\tau|x_i - y_i|) + \sum_{i,j} (-\lambda[x_i \neq x_j]) \right)$$

$$p(x,y) = Z^{-1} exp\left(-\left(\sum_i \tau|x_i - y_i| + \sum_{i,j} \lambda[x_i \neq x_j] \right)\right)$$

$$p(x,y) = Z^{-1} exp(-E(x,y))$$

2.2

b)

According to MSE I got the best result for

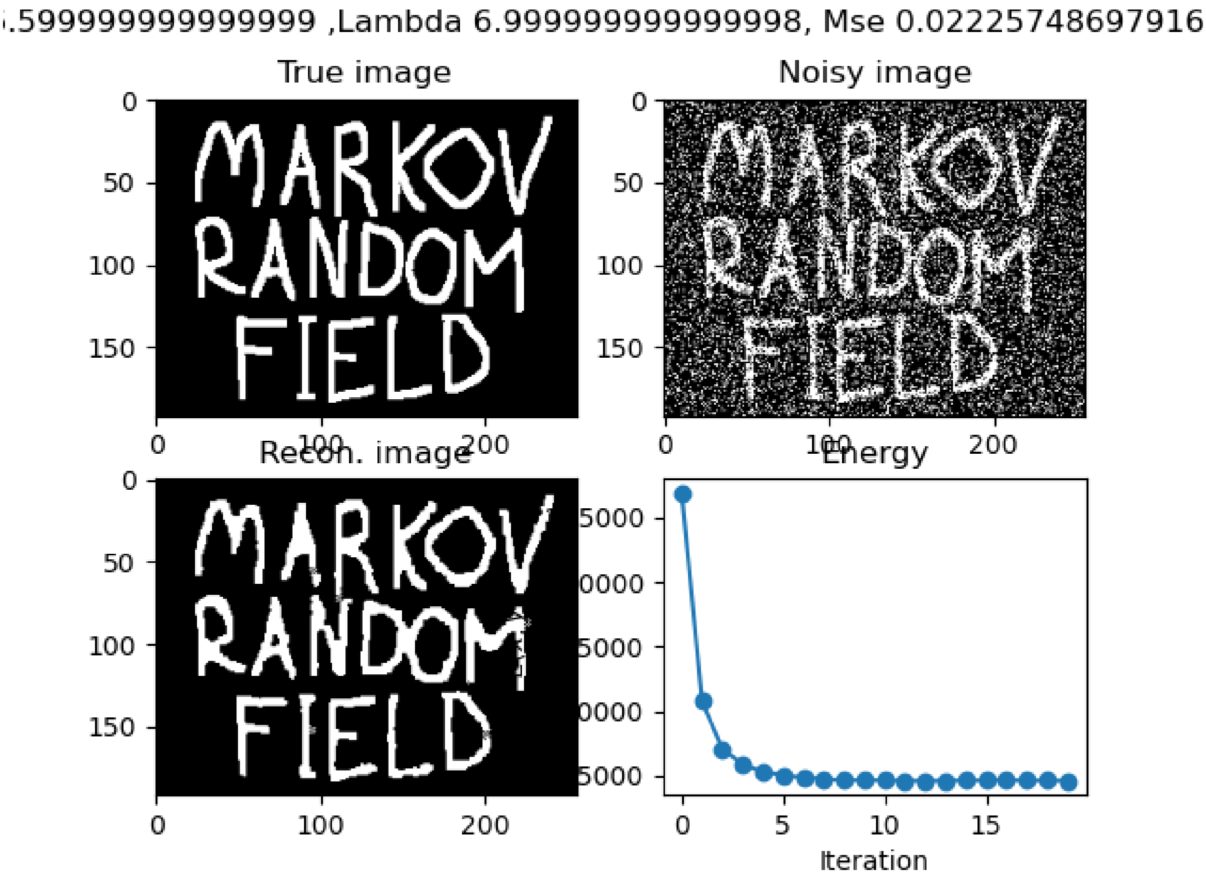
Tau=6.6

Lambda=7

Iterations=20

(Grid search with 0.2 steps for both parameters.)

Worst MSE was at ~0.199, which was basically the noisy image.



c)

Tau affects the belief how correct the observation is, thus governs the influence strength of the observation on the final reconstruction result.

Lambda governs the smoothness of the reconstructed image by determining how strongly the neighbouring nodes affect each other.

A high tau and low lambda value can result in a still noisy image.
While a low tau and high lambda value results in an oversmoothed / smudged image.
One has to find a good trade-off.

Remarks:

I would like to know how to come up with the specified factor functions in the denoising example. I feel that I have very theoretical knowledge about the topic and out-in-the-wild adapting these algorithms, designing functions, relationships between variables, entire graphs describing phenomena might be problematic.

I have learned that MRF and BN display different independencies.

I have learned Message passing and Sum Product Algorithm given the problem statement. The image denoising example was interesting.