

## Exercise 2

*Nir Raviv 200683548**Roi Tabach 203022983**Andrey Leshenko 322026527**nirraviv@mail.tau.ac.il**roi.tabach@gmail.com**andrey.leshenko@gmail.com*

## Q1

a

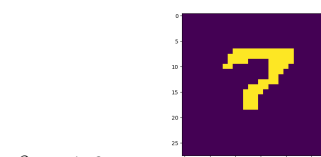
b

c

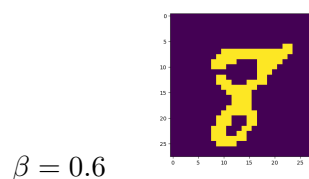
The code for this question can be found in [ANDREY TO THE RESCUE].

As recommended we worked in log space, normalized messages, and used the given skeleton. The interesting parts of the code are found in Vertex class: functions compute message and get belief.

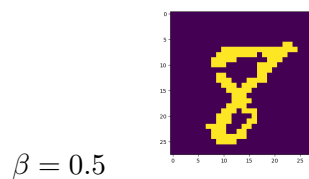
We will plot what happens when  $\beta$  is changed,  $\alpha = 1$  constantly:



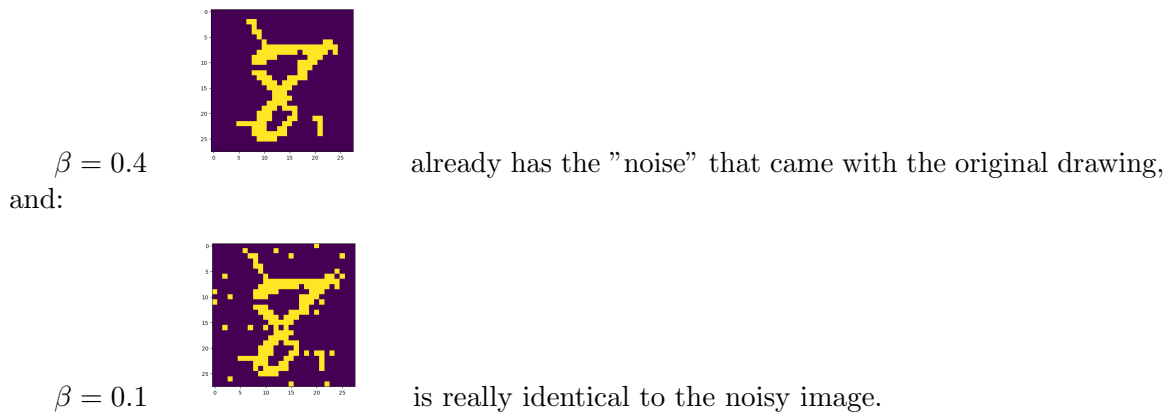
is super smooth, the image loses its meaning.



is great.



is a bit too loyal to the original image.



## Q2

We need to show that the tree-width of a 2D grid graph of size  $(m, m)$  is at most  $m$ . According to [1] if a graph has tree-width of at most  $k$  then  $k + 1$  omniscient cops can catch a robber on  $G$ .

On our 2D grid graph,  $G$ , we will show that we can use a strategy to win the cops and robber game with  $m + 1$  cops.

Denote the vertices of  $G$  by  $v_{ij}$ . First place  $m$  cops on  $v_{i1}$  for every  $j$  (i.e. the "first column" of  $G$ ). Then for  $j = 1, \dots, m$  add a cop to  $v_{j2}$  and remove the cop from  $v_{j1}$  until all the "second column" vertices are occupied by cops. Continue this process until all vertices  $v_{im}$  are occupied by cops.

We can see that at any iteration the cops separate the later columns from the earlier columns so the robber cannot reach earlier columns and will be captured eventually.

QED.

## Q3

We now consider the sum-product message update on a tree graph, with the messages updated simultaneously, not recursively. so:

$$m_{ij}^{t+1}(x_j) = \sum_{x_i} \phi_{ij}(x_i, x_j) \cdot \prod_{k \in N(j) \setminus i} m_{ki}^t(x_i) \quad (1)$$

We need to prove that at iteration  $t = n$  this converges to the true marginals.

We know from class that the recursive algorithm converges to the true marginals, so we will only prove the following fact: While updating using the simultaneous equations, for an edge  $e = i \rightarrow j$ , the message computed with the new ("parallel") algorithm converges to the message computed the recursive algorithm, after at most  $d$  steps, where  $d$  is defined here as the depth of the sub-tree of  $i$  as a descendant of  $j$ .

We will use induction on the depth. For the base case, with depth 0, we have to initialize with the uniform distribution, and the result is trivial.

Now we assume for  $d_j \leq t$ , and prove for  $d_j = t + 1$ . We start with (1), and from the assumption we can replace each  $m_{ki}^t(x_i)$  with the recursive counterpart  $m_{ki}(x_i)$ , to get

$$m_{ij}^{t+1}(x_j) = \sum_{x_i} \phi_{ij}(x_i, x_j) \cdot \prod_{k \in N(j) \setminus i} m_{ki}(x_i)$$

$$m_{ij}^{t+1}(x_j) = m_{ij}(x_j)$$

QED.

## Q4

We consider the pairwise MRF:  $p(x_1, \dots, x_n) \propto \prod_{ij \in E} \phi_{ij}(x_i, x_j)$ . And define the following “pseudo-marginals”:

$$\mu_i(x_i) \propto \prod_{k \in N(i)} m_{ki}(x_i) \quad (2)$$

$$\mu_{ij}(x_i, x_j) \propto \phi_{ij}(x_i, x_j) \prod_{k \in N(i) \setminus j} m_{ki}(x_i) \prod_{k \in N(j) \setminus i} m_{kj}(x_j) \quad (3)$$

**a**

We will show that the above pseudo-marginals are indeed correct for tree graphs. For the next part, we define  $ST_j$  to be the sub-tree defined by node  $j$  when node  $i$  is the root of the tree.  $E_{ST_j}$  will be the edges of this sub-tree. Claim:

$$m_{ji}(x_i) = \sum_{(x_k \mid k \in ST_j)} \left( \phi_{ij}(x_i, x_j) \prod_{ab \in E_{ST_j}} \phi_{ab}(x_a, x_b) \right) \quad (4)$$

We will prove this claim using induction on the depth of  $ST_j$ . In the base case  $j$  is a leaf,  $ST_j = j$  and  $E_{ST_j} = \emptyset$ . We then have:

$$m_{ji}(x_i) = \sum_{x_k} \phi_{ij}(x_i, x_j) = \sum_{(x_k \mid k \in ST_j)} \left( \phi_{ij}(x_i, x_j) \prod_{ab \in E_{ST_j}} \phi_{ab}(x_a, x_b) \right) \quad (5)$$

For the induction step we use the fact that  $j$ -s children have lower subtrees than  $j$ , so the induction hypothesis holds for them:

$$m_{ji}(x_i) = \sum_{x_j} \phi_{ij}(x_i, x_j) \prod_{k \in N(j) \setminus i} m_{kj}(x_j) \quad (6)$$

$$= \sum_{x_j} \phi_{ij}(x_i, x_j) \prod_{k \in N(j) \setminus i} \left[ \sum_{(x_l \mid l \in ST_k)} \left( \phi_{jk}(x_j, x_k) \prod_{ab \in E_{ST_k}} \phi_{ab}(x_a, x_b) \right) \right] \quad (7)$$

The outer product is made out of terms, each of which applies to a single subtree. Each term is a sum of that goes over all possible assignment for this subtree. Because the different subtrees are disjoint, using the distributivity property we can rearrange the equation and have the an outer sum that goes over all assignments for the union of all subtrees (which is  $ST_j$ ), and for each assignment we have a product of its  $\phi$ -s. After rearrangement we are left with:

$$= \sum_{(x_l \mid l \in ST_j)} \left( \phi_{ij}(x_i, x_j) \prod_{ab \in E_{ST_j}} \phi_{ab}(x_a, x_b) \right) \quad (8)$$

Which proves the claim. We will now use the claim to substitute it into the definitions of our pseudo-marginals:

$$\mu_i(x_i) \propto \prod_{k \in N(i)} m_{ki}(x_i) = \prod_{k \in N(i)} \left[ \sum_{(x_k \mid k \in ST_j)} \left( \phi_{ik}(x_i, x_k) \prod_{ab \in E_{ST_j}} \phi_{ab}(x_a, x_b) \right) \right] \quad (9)$$

If we consider  $i$  to be the root of the tree, the summation covers all sub-trees defined by  $i$ -s children, therefore it covers all nodes except  $i$ . The inner product covers all edges inside the sub-trees defined by  $i$ -s children, and the outer product also contains all the edges that connect  $i$  to its children, therefore the the product covers all edges in the tree. We can now rewrite the equation as follows:

$$\mu_i(x_i) \propto \sum_{\{x_1, \dots, x_n\} \setminus x_i} \left( \prod_{ab \in E} \phi_{ab}(x_a, x_b) \right) \propto \sum_{\{x_1, \dots, x_n\} \setminus x_i} p(x_1, \dots, x_n) \quad (10)$$

And this shows that  $\mu_i$  is indeed the correct marginal for trees. Similarly, we will substitute into  $\mu_{ij}(x_i, x_j)$ . The resulting equation is very similar to the previous one, but takes up twice as much space. Each message  $m_{ji}$  transforms into its expanded form, and then we are left with a sum of products. Finally we get:

$$\mu_{ij}(x_i, x_j) \propto \sum_{\{x_1, \dots, x_n\} \setminus x_i, x_j} \left( \prod_{ab \in E} \phi_{ab}(x_a, x_b) \right) \propto \sum_{\{x_1, \dots, x_n\} \setminus x_i, x_j} p(x_1, \dots, x_n) \quad (11)$$

Which is the correct marginal. Our pseudo-marginals are indeed the correct marginals for MRF tree graphs.

**b**

We will show that for any graph  $E$  (possible non-tree) it holds that:

$$p(x_1, \dots, x_n) \propto \prod_i \mu_i(x_i) \prod_{ij \in E} \frac{\mu_{ij}(x_i, x_j)}{\mu_i(x_i) \mu_j(x_j)} \quad (12)$$

We will substitute the definitions of the marginals into the above equation, and then see that all the messages cancel out and only the terms we need are left:

$$\prod_i \mu_i(x_i) \prod_{ij \in E} \frac{\mu_{ij}(x_i, x_j)}{\mu_i(x_i) \mu_j(x_j)} = \prod_i \mu_i(x_i) \prod_{ij \in E} \frac{\phi_{ij}(x_i, x_j) \prod_{k \in N(i) \setminus j} m_{ki}(x_i) \prod_{k \in N(j) \setminus i} m_{kj}(x_j)}{\prod_{k \in N(i)} m_{ki}(x_i) \prod_{k \in N(j)} m_{kj}(x_j)} \quad (13)$$

The large fraction has many similarities between its numerator and denominator. The differences are the  $\phi_{ij}$  term in the numerator, and the fact that the product in the numerator skips over  $m_{ji}(x_i)$  and  $m_{ij}(x_j)$ . After simplification:

$$= \prod_i \mu_i(x_i) \prod_{ij \in E} \frac{\phi_{ij}(x_i, x_j)}{m_{ji}(x_i) m_{ij}(x_j)} = \left( \prod_i \prod_{k \in N(i)} m_{ki}(x_i) \right) \prod_{ij \in E} \frac{\phi_{ij}(x_i, x_j)}{m_{ji}(x_i) m_{ij}(x_j)} \quad (14)$$

Going over all vertices and then over all inbound edges for each vertex is the same as going over all edges in both directions. Therefore:

$$= \left( \prod_{ij \in E} m_{ji}(x_i) m_{ij}(x_j) \right) \prod_{ij \in E} \frac{\phi_{ij}(x_i, x_j)}{m_{ji}(x_i) m_{ij}(x_j)} = \prod_{ij \in E} \phi_{ij}(x_i, x_j) \propto p(x_1, \dots, x_n) \quad (15)$$

QED.

**c**

We will show that at a fixed point of LBP we have:  $\sum_{x_j} \mu_{ij}(x_i, x_j) = \mu_i(x_i)$ . First, change the marginals into their definitions:

$$\sum_{x_j} \mu_{ij}(x_i, x_j) = \sum_{x_j} \phi_{ij}(x_i, x_j) \prod_{k \in N(i) \setminus j} m_{ki}(x_i) \prod_{k \in N(j) \setminus i} m_{kj}(x_j) = \quad (16)$$

$$= \left( \prod_{k \in N(i) \setminus j} m_{ki}(x_i) \right) \left[ \sum_{x_j} \phi_{ij}(x_i, x_j) \prod_{k \in N(j) \setminus i} m_{kj}(x_j) \right] = \quad (17)$$

The term inside the square brackets is exactly the definition of  $m_{ji}$ . Because we are in a fixed point of LBP, all values are stable and the messages are consistent with their initial

definition (which might not be the case if the algorithm is still running).  $m_{ji}$  is up to date with all values in the graph and therefore:

$$= \left( \prod_{k \in N(i) \setminus j} m_{ki}(x_i) \right) [m_{ji}] = \prod_{k \in N(i)} m_{ki}(x_i) = \mu_i(x_i) \quad (18)$$

We have shown that at fixed points of LBP:  $\sum_{x_j} \mu_{ij}(x_i, x_j) = \mu_i(x_i)$ . QED.

## References

### References

- [1] P.D. Seymour and R. Thomas. Graph searching and a min-max theorem for tree-width. *Journal of Combinatorial Theory, Series B*, 58(1):22 – 33, 1993. [2-2](#)