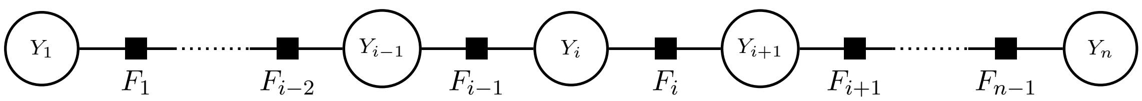
Weekly Exercise 9

Dr. Csaba Domokos and Lingni Ma
Technische Universität München, Computer Vision Group
June 14, 2016 (submission deadline: June 21, 2016)

Probabilistic Inference (8 Points)

Exercise 1 (**Inference on chains**, 2 Points). Consider the following factor graph, which is a chain:



The joint distribution can be written in the form

$$p(\mathbf{y}) = \frac{1}{Z}\psi_{F_1}(y_1, y_2)\psi_{F_2}(y_2, y_3) \cdot \ldots \cdot \psi_{F_{n-1}}(y_{n-1}, y_n) ,$$

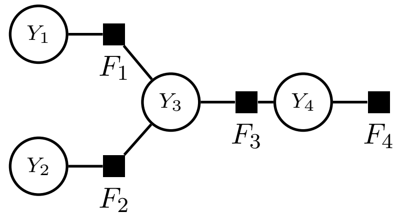
where $Z = \sum_{\mathbf{y}} \prod_{i=1}^{n-1} \psi_{F_i}(y_i, y_{i+1})$ denotes the partition function. Show that the *marginal distribution* $p(y_i)$ decomposes into the product of two factors:

$$p(y_i) = \frac{1}{Z} r_{F_{i-1}\to Y_i}(y_i) r_{F_i \to Y_i}(y_i) .$$

Solution. Using the definition of $p(y_i)$ we have:

$$\begin{aligned}
p(y_i) &= \sum_{y_1} \cdots \sum_{y_{i-1}} \sum_{y_{i+1}} \cdots \sum_{y_n} p(\mathbf{y}) \\
&= \sum_{y_1} \cdots \sum_{y_{i-1}} \sum_{y_{i+1}} \cdots \sum_{y_n} \frac{1}{Z}\psi_{F_1}(y_1, y_2)\psi_{F_2}(y_2, y_3) \cdot \ldots \cdot \psi_{F_{n-1}}(y_{n-1}, y_n) \\
&= \frac{1}{Z}\sum_{y_1} \cdots \sum_{y_{i-1}} \sum_{y_{i+1}} \cdots \sum_{y_n} \psi_{F_1}(y_1, y_2)\psi_{F_2}(y_2, y_3) \cdot \ldots \cdot \psi_{F_{n-1}}(y_{n-1}, y_n) \\
&= \frac{1}{Z}\left[\sum_{y_{i-1}} \psi_{F_{i-1}}(y_{i-1}, y_i) \ldots \left[\sum_{y_2} \psi_{F_2}(y_2, y_3)\left[\sum_{y_1} \psi_{F_1}(y_1, y_2)\right]\right] \ldots \right] \\
&\quad \cdot \left[\sum_{y_{i+1}} \psi_{F_i}(y_i, y_{i+1}) \ldots \left[\sum_{y_n} \psi_{F_{n-1}}(y_{n-1}, y_n)\right] \ldots \right] \\
&= \frac{1}{Z} r_{F_{i-1}\to Y_i}(y_i) r_{F_i \to Y_i}(y_i).
\end{aligned}$$

Exercise 2 (**Sum-product and max-sum**, 6 Points)**.** Consider the following factor graph,



The potential functions are defined for $k \in \{1, 2, 3\}$

$$\psi_{F_k}(y_i, y_j) = \exp\left(-(|y_i - y_j| + (c_k - y_i)^2)\right) \text{ , where } c_1 = 0, c_2 = 1, c_3 = 1 \text{ , and}$$
$$\psi_{F_4}(y_4) = \exp(-(2 - y_4)^2) \ .$$

Assume Y_4 as the root node and $\mathbf{y} \in \mathcal{L}^4 = \{0, 1, 2\}^4$.

a) Perform the `sum-product algorithm` in order to achieve *probabilistic inference* of the model expressed by the graph above. Show the intermediate steps in details.

b) Perform the `max-sum algorithm` to achieve *MAP inference* of the model expressed by the graph above. Show the intermediate steps in details.

Solution. Instead of the potential functions, we use energy functions, defined as

$$E_k(y_i, y_j) = |y_i - y_j| + (c_k - y_i)^2 \text{ , where } c_1 = 0, c_2 = 1, c_3 = 1 \text{ , and}$$
$$E_4(y_4) = (2 - y_4)^2 \ .$$

a) Let us consider Y_4 as the root node. Starting with the leaf nodes, we then have the following *leaf–to–root* messages:

$$q_{Y_1 \to F_1}(y_1) = 1 \qquad \text{for } \forall y_1 \in \mathcal{L} \tag{1}$$
$$r_{F_1 \to Y_3}(y_3) = \sum_{y_1 \in \mathcal{L}} \exp(-E_{F_1}(y_1, y_3)) \cdot q_{Y_1 \to F_1}(y_1) \tag{2}$$
$$q_{Y_2 \to F_2}(y_2) = 1 \qquad \text{for } \forall y_2 \in \mathcal{L} \tag{3}$$
$$r_{F_2 \to Y_3}(y_3) = \sum_{y_2 \in \mathcal{L}} \exp(-E_{F_2}(y_2, y_3)) \cdot q_{Y_2 \to F_2}(y_2) \tag{4}$$
$$q_{Y_3 \to F_3}(y_3) = r_{F_1 \to Y_3}(y_3) \cdot r_{F_2 \to Y_3}(y_3) \tag{5}$$
$$r_{F_3 \to Y_4}(y_4) = \sum_{y_3 \in \mathcal{L}} \exp(-E_{F_3}(y_3, y_4)) \cdot q_{Y_3 \to F_3}(y_3) \tag{6}$$
$$r_{F_4 \to Y_4}(y_4) = \exp(-E_{F_4}(y_4)) \ . \tag{7}$$

Starting with the root node, we then have the following *root–to–leaf* messages:

$$q_{Y_4 \to F_4}(y_4) = r_{F_3 \to Y_4}(y_4) \tag{8}$$

$$q_{Y_4 \to F_3}(y_4) = r_{F_4 \to Y_4}(y_4) \tag{9}$$

$$r_{F_3 \to Y_3}(y_3) = \sum_{y_4 \in \mathcal{L}} \exp(-E_{F_3}(y_3, y_4)) \cdot q_{Y_4 \to F_3}(y_4) \tag{10}$$

$$q_{Y_3 \to F_2}(y_3) = r_{F_1 \to Y_3}(y_3) \cdot r_{F_3 \to Y_3}(y_3) \tag{11}$$

$$r_{F_2 \to Y_2}(y_2) = \sum_{y_3 \in \mathcal{L}} \exp(-E_{F_2}(y_2, y_3)) \cdot q_{Y_3 \to F_2}(y_3) \tag{12}$$

$$q_{Y_3 \to F_1}(y_3) = r_{F_2 \to Y_3}(y_3) \cdot r_{F_3 \to Y_3}(y_3) \tag{13}$$

$$r_{F_1 \to Y_1}(y_1) = \sum_{y_2 \in \mathcal{L}} \exp(-E_{F_1}(y_1, y_2)) \cdot q_{Y_3 \to F_1}(y_3) \ . \tag{14}$$

Easy calculations show that the messages are given as:

	0	1	2
$q_{Y_1 \to F_1}$	1	1	1
$r_{F_1 \to Y_3}$	1.5032	0.63855	0.027532
$q_{Y_2 \to F_2}$	1	1	1
$r_{F_2 \to Y_3}$	0.78555	1.2707	0.78555
$q_{Y_3 \to F_3}$	1.1808	0.81139	0.021628
$r_{F_3 \to Y_4}$	0.73398	0.97412	0.36524
$r_{F_4 \to Y_4}$	0.018316	0.36788	1
$q_{Y_4 \to F_4}$	0.73398	0.97412	0.36524
$q_{Y_4 \to F_4}$	0.018316	0.36788	1
$r_{F_3 \to Y_3}$	0.10631	0.7425	0.41858
$q_{Y_3 \to F_3}$	0.15981	0.47412	0.011524
$r_{F_2 \to Y_2}$	0.12353	0.53715	0.076361
$q_{Y_3 \to F_1}$	0.083513	0.94347	0.32881
$r_{F_1 \to Y_1}$	0.21201	0.38002	0.14501

The partition function Z can be calculated as:

$$Z = \sum_{y_1 \in \mathcal{L}} r_{F_1 \to Y_1}(y_1) \tag{15}$$

$$= \sum_{y_2 \in \mathcal{L}} r_{F_2 \to Y_2}(y_2) \tag{16}$$

$$= \sum_{y_3 \in \mathcal{L}} r_{F_1 \to Y_3}(y_3) \cdot r_{F_2 \to Y_3}(y_3) \cdot r_{F_3 \to Y_3}(y_3) \tag{17}$$

$$= \sum_{y_4 \in \mathcal{L}} r_{F_3 \to Y_4}(y_4) \cdot r_{F_4 \to Y_4}(y_4) = 0.737 \ . \tag{18}$$

b) Let us consider Y_4 as the root node. Starting with the leaf nodes, we then have the following *leaf–to–root* messages:

$$q_{Y_1 \to F_1}(y_1) = 0 \qquad \text{for } \forall y_1 \in \mathcal{L} \tag{19}$$
$$r_{F_1 \to Y_3}(y_3) = \max_{y_1 \in \mathcal{L}} \{-E_{F_1}(y_1, y_3) + q_{Y_1 \to F_1}(y_1)\} \tag{20}$$
$$q_{Y_2 \to F_2}(y_2) = 0 \qquad \text{for } \forall y_2 \in \mathcal{L} \tag{21}$$
$$r_{F_2 \to Y_3}(y_3) = \max_{y_2 \in \mathcal{L}} \{-E_{F_2}(y_2, y_3) + q_{Y_2 \to F_2}(y_2)\} \tag{22}$$
$$q_{Y_3 \to F_3}(y_3) = r_{F_1 \to Y_3}(y_3) + r_{F_2 \to Y_3}(y_3) \tag{23}$$
$$r_{F_3 \to Y_4}(y_4) = \max_{y_3 \in \mathcal{L}} \{-E_{F_3}(y_3, y_4) + q_{Y_3 \to F_3}(y_3)\} \tag{24}$$
$$r_{F_4 \to Y_4}(y_4) = -E_{F_4}(y_4) \; . \tag{25}$$

Easy calculations show that the messages are given as:

	0	1	2
$q_{Y_1 \to F_1}$	0	0	0
$r_{F_1 \to Y_3}$	0	-1	-2
$q_{Y_2 \to F_2}$	0	0	0
$r_{F_2 \to Y_3}$	-1	0	-1
$q_{Y_3 \to F_3}$	-1	-1	-3
$r_{F_3 \to Y_4}$	-2	-1	-4
$r_{F_4 \to Y_4}$	-4	-1	0

Hence, the maximizing energy is given as:

$$E(\mathbf{y}^*) = \max_{y_4 \in \mathcal{L}} \{r_{F_3 \to Y_4}(y_4) + r_{F_4 \to Y_4}(y_4)\} = -2 \; .$$

To find a maximizing configuration, we have the following sequence of updates:

$$y_4^* \in \text{argmax}_{y_4 \in \mathcal{L}} \{r_{F_3 \to Y_4}(y_4) + r_{F_4 \to Y_4}(y_4)\} = \{1\} \tag{26}$$
$$y_3^* \in \text{argmax}_{y_4 \in \mathcal{L}} \{-E_{F_3}(y_3, 1) + r_{F_2 \to Y_3}(y_3) + r_{F_1 \to Y_3}(y_3)\} = \{1\} \tag{27}$$
$$y_2^* \in \text{argmax}_{y_2 \in \mathcal{L}} \{-E_{F_2}(y_2, 1)\} = \{1\} \tag{28}$$
$$y_1^* \in \text{argmax}_{y_1 \in \mathcal{L}} \{-E_{F_1}(y_1, 1)\} = \{0, 1\}. \tag{29}$$

Thus, two global maximizers are given by:

$$\mathbf{y}^* \in \{(0,1,1,1), (1,1,1,1)\} \; .$$