Combinatorial Optimization in Computer Vision
Lecture: F. R. Schmidt and C. Domokos
Exercises: T. Möllenhoff and T. Windheuser
Winter Semester 2015/2016

Computer Vision Group
Institut für Informatik
Technische Universität München

Weekly Exercises 8

Room: 02.09.023
Tuesday, 16.12.2015, 14:15-15:45
Submission deadline: Tuesday, 16.12.2015, 11:15 , Room 02.09.023

Higher-Order Reduction Methods (4 Points)

Exercise 1 (4 Points)**.** Consider the following pseudo-Boolean function $f : \mathbb{B}^5 \to \mathbb{R}$

$$f(x_1, \ldots, x_5) = 5x_1x_2 - 7x_1x_2x_3x_4 + 2x_1x_2x_3x_5. \tag{1}$$

a) **Reduction by Substitution.** Reduce the energy (1) to an equivalent quadratic pseudo-Boolean function by substituting products of two variables by a new variable and introducing constraints.

b) **Reduction by Minimum Selection.** Apply Ishikawa's extended minimum selection reduction method to energy (1).

Solution.

a) We first substitute $x_6 = x_1x_2$ and see that $M = 1 + 5 + 7 + 2 = 15$. Using $D(x, y, z) = xy - 2xz - 2yz + 3z$ and adding the constraint $M \cdot D(x_1, x_2, x_6)$ yields

$$20x_1x_2 - 7x_3x_4x_6 + 2x_3x_5x_6 - 30x_1x_6 - 30x_2x_6 + 45x_6.$$

Applying the reduction again with $x_7 = x_3x_6$ which means adding $M \cdot D(x_3, x_6, x_7)$ yields

$$45x_6+45x_7+20x_1x_2-30x_1x_6-30x_2x_6+15x_3x_6-30x_3x_7-7x_4x_7+2x_5x_7-30x_6x_7.$$

b) We have

$$xyzt = \min_{w\in\mathbb{B}} w(-2s_1 + 3) + s_2,$$

where $s_1 = x + y + z + t$, $s_2 = xy + yz + zx + tx + ty + tz$.

Applying this yields

$$2x_1x_2x_3x_5 = \min_{x_6\in\mathbb{B}} 2(x_6(-2(x_1+x_2+x_3+x_5)+3)+x_1x_2+x_2x_3+x_3x_1+x_5x_1+x_5x_2+x_5x_3).$$

Furthermore for $a < 0$:

$$axyzt = \min_{w\in\mathbb{B}} aw(x + y + z + t - 3),$$

hence

$$-7x_1x_2x_3x_4 = \min_{x_7\in\mathbb{B}} \ -7x_7(x_1+x_2+x_3+x_4-3).$$

The final energy is given as:

$$\begin{aligned}&-7x_1x_7-7x_2x_7-7x_3x_7-7x_4x_7-21x_7-4x_1x_6-4x_2x_6-4x_3x_6\\&-4x_5x_6-6x_6+2x_1x_2+2x_2x_3+2x_1x_3+2x_1x_5+2x_2x_5+2x_3x_5+5x_1x_2.\end{aligned} \tag{2}$$

Multi-Linear Polynomial Representation (4 Points)

Exercise 2 (4 Points)**.** Let $E:\mathbb{B}^n\to\mathbb{R}, n\in\mathbb{N}$, be a pseudo-boolean function. Clearly it can be written as the multi-linear polynomial

$$E(x_1,\dots,x_n)=\sum_{\alpha\subset\{1,\dots,n\}} a_\alpha \mathbf{x}_\alpha, \tag{3}$$

where $\mathbf{x}_\alpha=\Pi_{i\in\alpha}x_i$ and $a_\alpha\in\mathbb{R}$ are the coefficients, defining E. Now let $E_\beta\in\mathbb{R},\beta\subset\{1,\dots,n\}$, be defined by $E_\beta=E(x_i=1,i\in\beta,x_j=0,j\in\{1,\dots,n\}\setminus\beta)$. For example let $n=3$, then $E_{\{1,2\}}=E(1,1,0)$ and $E_{\{2\}}=E(0,1,0)$.

a) Show that the coefficients a_α are uniquely determined by the following recursive formula:

$$a_\alpha=E_\alpha-\sum_{\beta\subsetneq\alpha}a_\beta. \tag{4}$$

b) Use the recursive formula (4) to derive the coefficients a_α of the multi-linear polynomial of E, where E is given by:

x_1	x_2	x_3	$E(x_1,x_2,x_3)$
0	0	0	2
0	0	1	-13
0	1	0	4
0	1	1	5
1	0	0	-7
1	0	1	3
1	1	0	9
1	1	1	8

Solution. a) We show the claim by induction over the size $N=|\alpha|$ of α. For $N=0$ the claim is clearly true, since

$$E_{\{\}}=E(0,\dots,0)=a_{\{\}}.$$

Also for $N=1$ the claim is true: Let $\alpha=\{i\}$, then

$$\begin{aligned}E_{\{i\}}&=E(0,\dots,0,1,0,\dots,0)=a_{\{\}}+a_{\{i\}}\\ \Rightarrow a_{\{i\}}&=E_{\{i\}}-a_{\{\}}.\end{aligned}$$

Now let Equation (4) hold some $N < n$ and let $\alpha \subset \{1, \dots, n\}$, with $|\alpha| = N+1$. Then

$$E_\alpha = \sum_{\beta \subset \alpha} a_\beta = a_\alpha + \sum_{\beta \subsetneq \alpha} a_\beta$$
$$\Rightarrow a_\alpha = E_\alpha - \sum_{\beta \subsetneq \alpha} a_\beta.$$

b)

$$E(x_1, x_2, x_3) = 2 - 9x_1 + 2x_2 - 15x_3 + 14x_1x_2 + 25x_1x_3 + 16x_2x_3 - 27x_1x_2x_3.$$

Programming (10 Points)

The goal of this exercise is to perform semantic multilabel segmentation of an image $I : \Omega \to \mathbb{R}^3$ using a contrast sensitive Potts regularizer. For the chosen dataset, the labels $\mathcal{L} = \{1, \dots, 21\}$ denote various classes such as `background`, `horse`, … The goal is to find a segmentation $x \in \mathcal{L}^n$ by minimizing the following energy:

$$E(x) = \sum_{i=1}^{n} f_i(x_i) + \sum_{i=1}^{n} \sum_{j \in \mathcal{N}(i)} f_{ij}(x_i, x_j). \tag{5}$$

The unary potentials $f_i : \mathcal{L}^n \to \mathbb{R}$ for some example images are provided as `.hdf5` files in the supplementary material `08_supp.zip` and were computed using a state-of-the-art Convolutional Neural Network[1]. The output of the network is a score for every class $s_i \in \mathbb{R}^{21}$ at each pixel $i \in \Omega$. It is stored in a pixel-first fashion, i.e. to access the cost of the l-th label at pixel (x, y) use the linear index `data[x+y*W+l*W*H]`. Compute the unary potentials as

$$f_i(x_i) = -s_i, \ \forall i \in \Omega.$$

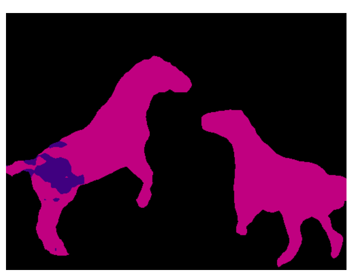
For the pairwise potentials, use the contrast sensitive Potts-model:

$$f_{ij}(x_i, x_j) = \begin{cases} 0, & \text{if } x_i = x_j, \\ g(i, j), & \text{otherwise,} \end{cases}$$

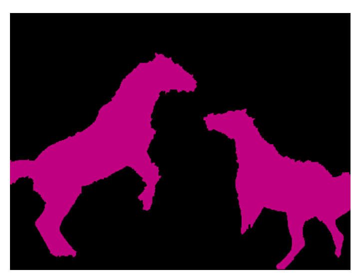
where $g(i, j) = \lambda \exp(-\gamma |I_i - I_j|)$ for some parameters $\lambda > 0$, $\gamma > 0$. Here $I_i \in \mathbb{R}^3$ denotes the color of pixel $i \in \Omega$ in the image $I : \Omega \to \mathbb{R}^3$.

Implement the model in OpenGM using `segmentation.cpp` provided in the supplementary material as a starting point. For inference, use the α-Expansion method.

Input



Unaries only



Potts

[1] Fully Convolutional Networks for Semantic Segmentation. `www.cs.berkeley.edu/~jonlong/long_shelhamer_fcn.pdf`