From a Set of Shapes to Object Discovery

Nadia Payet and Sinisa Todorovic

Oregon State University,
Kelley Engineering Center, Corvallis, OR 97331, USA
payetn@onid.orst.edu, sinisa@eecs.oregonstate.edu

1 Learning Model Parameters: Derivation of Eq. (6)

As explained in Sec 4 under Learning, we estimate the model parameters by maximizing the acceptance rate of moving from state A to state B, defined as

$$\alpha(A \to B) = \min\left(1, \frac{q(B \to A)p(\mathcal{M} = B|G)}{q(A \to B)p(\mathcal{M} = A|G)}\right).$$

From (5) in the paper, and $(1-\rho_e^+) \leq 1$ and $(1-\rho_e^-) \leq 1$ by definition, we have:

$$\begin{aligned}\frac{q(B\to A)}{q(A\to B)} &= \frac{\prod_{e\in\text{Cut}_B^+}(1-\rho_e^+)\prod_{e\in\text{Cut}_B^-}(1-\rho_e^-)}{\prod_{e\in\text{Cut}_A^+}(1-\rho_e^+)\prod_{e\in\text{Cut}_A^-}(1-\rho_e^-)},\\ &\geq \prod_{e\in\text{Cut}_B^+}(1-\rho_e^+)\prod_{e\in\text{Cut}_B^-}(1-\rho_e^-).\end{aligned}$$

As explained in the paper, the edges in Cut_B^+ and Cut_B^- are probabilistically cut. This means that their associated likelihoods ρ_e^+ and ρ_e^- are relatively small. Therefore, for most edges in Cut_B^+ and Cut_B^- it holds that the probability of being cut is greater than the probability of being sampled, $1-\rho_e^+ \geq \rho_e^+$ and $1-\rho_e^- \geq \rho_e^-$. Therefore, we have

$$\frac{q(B\to A)}{q(A\to B)} \geq \prod_{e\in\text{Cut}_B^+}(1-\rho_e^+)\prod_{e\in\text{Cut}_B^-}(1-\rho_e^-) \geq \prod_{e\in\text{Cut}_B^+}\rho_e^+\prod_{e\in\text{Cut}_B^-}\rho_e^-.$$

Next, from (2) and (3) in the paper, we have

$$\begin{aligned}\frac{p(\mathcal{M}=B|G)}{p(\mathcal{M}=A|G)} &= \frac{p(\mathcal{M}=B)p(G|\mathcal{M}=B)}{p(\mathcal{M}=A)p(G|\mathcal{M}=A)},\\ &= \frac{e^{-w_K K_B}e^{-w_N N_B}\prod_{e\in\mathbb{E}_B^+}\rho_e^+\prod_{e\in\mathbb{E}_B^-}\rho_e^-\prod_{e\in\mathbb{E}_B^0}(1-\rho_e^+)\mathbb{1}_{l_i\neq l_j}\cdot(1-\rho_e^-)\mathbb{1}_{l_i=l_j}}{e^{-w_K K_A}e^{-w_N N_A}\prod_{e\in\mathbb{E}_A^+}\rho_e^+\prod_{e\in\mathbb{E}_A^-}\rho_e^-\prod_{e\in\mathbb{E}_A^0}(1-\rho_e^+)\mathbb{1}_{l_i\neq l_j}\cdot(1-\rho_e^-)\mathbb{1}_{l_i=l_j}},\\ &\geq \frac{e^{-w_K K_B}e^{-w_N N_B}\prod_{e\in\mathbb{E}_B^+}\rho_e^+\prod_{e\in\mathbb{E}_B^-}\rho_e^-}{e^{-w_K K_A}e^{-w_N N_A}\prod_{e\in\mathbb{E}_A^+}\rho_e^+\prod_{e\in\mathbb{E}_A^-}\rho_e^-}\prod_{e\in\mathbb{E}_B^0}(1-\rho_e^+)\mathbb{1}_{l_i\neq l_j}\cdot(1-\rho_e^-)\mathbb{1}_{l_i=l_j}.\end{aligned}$$

As explained in the paper, the edges in $\mathbb{E}_B^0$ are probabilistically cut. This means that their associated likelihoods ρ_e^+ and ρ_e^- are relatively small. Therefore, for most edges in $\mathbb{E}_B^0$ it holds that $1-\rho_e^+ \geq \rho_e^+$ and $1-\rho_e^- \geq \rho_e^-$. Therefore, we have

$$\begin{aligned}\frac{p(\mathcal{M}=B|G)}{p(\mathcal{M}=A|G)} &\geq \frac{e^{-w_K K_B} e^{-w_N N_B} \prod_{e\in\mathbb{E}_B^+}\rho_e^+ \prod_{e\in\mathbb{E}_B^-}\rho_e^-}{e^{-w_K K_A} e^{-w_N N_A} \prod_{e\in\mathbb{E}_A^+}\rho_e^+ \prod_{e\in\mathbb{E}_A^-}\rho_e^-} \prod_{e\in\mathbb{E}_B^0} \rho_e^{+\mathbb{1}_{l_i\neq l_j}} \cdot \rho_e^{-\mathbb{1}_{l_i=l_j}},\\ &= \frac{e^{-w_K K_B} e^{-w_N N_B} \prod_{e\in\mathbb{E}_B^+}\rho_e^+ \prod_{e\in\mathbb{E}_B^-}\rho_e^-}{e^{-w_K K_A} e^{-w_N N_A} \prod_{e\in\mathbb{E}_A^+}\rho_e^+ \prod_{e\in\mathbb{E}_A^-}\rho_e^-} \prod_{e\in\mathbb{E}_B^{0-}} \rho_e^+ \prod_{e\in\mathbb{E}_B^{0+}} \rho_e^- .\end{aligned}$$

From the above steps we obtain

$$\frac{q(B\to A)}{q(A\to B)}\frac{p(\mathcal{M}=B|G)}{p(\mathcal{M}=A|G)} \geq \frac{e^{-w_K K_B} e^{-w_N N_B} \prod_{e\in\mathbb{E}_B^+}\rho_e^+ \prod_{e\in\text{Cut}_B^+}\rho_e^+ \prod_{e\in\mathbb{E}_B^{0-}}\rho_e^+ \prod_{e\in\mathbb{E}_B^-}\rho_e^- \prod_{e\in\text{Cut}_B^-}\rho_e^- \prod_{e\in\mathbb{E}_B^{0+}}\rho_e^-}{e^{-w_K K_A} e^{-w_N N_A} \prod_{e\in\mathbb{E}_A^+}\rho_e^+ \prod_{e\in\mathbb{E}_A^-}\rho_e^-}.$$

Let $\tilde{\mathbb{E}}_B^+$ denote all edges in the above equation whose likelihood is $\rho+$, $\tilde{\mathbb{E}}_B^+ = \mathbb{E}_B^+ \cup \text{Cut}_B^+ \cup \mathbb{E}_B^{0-}$. Also, let $\tilde{\mathbb{E}}_B^-$ denote all edges in the above equation whose likelihood is $\rho-$, $\tilde{\mathbb{E}}_B^- = \mathbb{E}_B^- \cup \text{Cut}_B^- \cup \mathbb{E}_B^{0+}$. Then, we derive

$$\begin{aligned}\frac{q(B\to A)}{q(A\to B)}\frac{p(\mathcal{M}=B|G)}{p(\mathcal{M}=A|G)} &\geq \frac{e^{-w_K K_B} e^{-w_N N_B} \prod_{e\in\tilde{\mathbb{E}}_B^+}\rho_e^+ \prod_{e\in\tilde{\mathbb{E}}_B^-}\rho_e^-}{e^{-w_K K_A} e^{-w_N N_A} \prod_{e\in\mathbb{E}_A^+}\rho_e^+ \prod_{e\in\mathbb{E}_A^-}\rho_e^-},\\ &= \frac{e^{-w_K K_B} e^{-w_N N_B} \prod_{e\in\tilde{\mathbb{E}}_B^+} e^{-w_\delta^+\delta_e} \prod_{e\in\tilde{\mathbb{E}}_B^-} e^{-w_\delta^-(1-\delta_e)}}{e^{-w_K K_A} e^{-w_N N_A} \prod_{e\in\mathbb{E}_A^+} e^{-w_\delta^+\delta_e} \prod_{e\in\mathbb{E}_A^-} e^{-w_\delta^-(1-\delta_e)}}\end{aligned}$$

By taking the logarithm of the above equation, we derive

$$\begin{aligned}\log\left(\frac{q(B\to A)}{q(A\to B)}\frac{P(\mathcal{M}=\mathcal{B}|G)}{P(\mathcal{M}=\mathcal{A}|G)}\right) &\geq w_K(K_A-K_B)+w_N(N_A-N_B)\\ &\quad+\textstyle\sum_{e\in\mathbb{E}_A^+} w_\delta^+\delta_e - \sum_{e\in\tilde{\mathbb{E}}_B^+} w_\delta^+\delta_e\\ &\quad+\textstyle\sum_{e\in\mathbb{E}_A^-} w_\delta^-(1-\delta_e) - \sum_{e\in\tilde{\mathbb{E}}_B^-} w_\delta^-(1-\delta_e),\\ &= \boldsymbol{\phi}^{\mathrm{T}}\boldsymbol{w},\end{aligned}$$

where $\boldsymbol{w} = \left[w_K, w_N, w_\delta^+, w_\delta^-\right]^T$ and $\boldsymbol{\phi} = \begin{bmatrix} K_A - K_B \\ N_A - N_B \\ \sum_{e \in \mathbb{E}_A^+} \delta_e - \sum_{e \in \tilde{\mathbb{E}}_B^+} \delta_e \\ \sum_{e \in \mathbb{E}_A^+} (1 - \delta_e) - \sum_{e \in \tilde{\mathbb{E}}_B^+} (1 - \delta_e) \end{bmatrix}$

This is equivalent to Eq. (6) in the paper, which concludes the proof.