

In Section 4.1.2 of paper: *Mining Term Association Patterns from Search Logs for Effective Query Reformulation*, there is an important and subtle typo in the formula:

$$t_C(s|w) = \frac{\exp(-D[P_C(\cdot|s)||\tilde{P}_C(\cdot|w)])}{\sum_s \exp(-D[P_C(\cdot|s)||\tilde{P}_C(\cdot|w)])}$$

The correct formula should reverse s and w .

$$t_C(s|w) = \frac{\exp(-D[P_C(\cdot|w)||\tilde{P}_C(\cdot|s)])}{\sum_s \exp(-D[P_C(\cdot|w)||\tilde{P}_C(\cdot|s)])}$$

The difference is actually significant due to asymmetry of the KL-Divergence computation.