Markov Logic Networks for Knowledge Base Completion: A Theoretical Analysis Under the MCAR Assumption (Appendix)

Ondřej Kuželka
Department of CS
CTU in Prague
Prague, Czech Republic

Jesse Davis
Department of CS
KU Leuven
Leuven, Belgium

A APPENDIX

A.1 PROOF OF THEOREM 7

Here we prove Theorem 7. The proof follows the steps of the proof of Theorem 5; there are just some additional details that are arguably not necessary for understanding the main ideas, which is why we deferred it to appendix.

Proof of Theorem 7. We first redefine the random variable $\langle \mathbf{w}, Q(\Phi, \omega) \rangle$ as a function of independent Bernoulli random variables $B_1, \dots, B_{|\omega^*|}$ satisfying $P[B_i = 0] = \delta$, where δ is the subsampling rate from Equation 6 (main paper). We suppose that there is some (arbitrary) ordering of the atoms in $\omega^* = \{a_1, \dots, a_{|\omega^*|}\}$ so that we could uniquely identify each B_i with an atom a_i in ω^*. Then we define a function $g : \{0,1\}^{|\omega^*|} \to 2^{\omega^*}$ as: $g(b_1, \dots, b_{\omega^*}) \mapsto \{a_i \in \omega^* | b_i = 1\}$. Finally we define $Q_{\mathbf{w},\Phi}(b_1, \dots, b_{\omega^*}) \triangleq \langle \mathbf{w}, Q(\Phi, g(b_1, \dots, b_{|\omega^*|})) \rangle$. It is easy to see that $\langle \mathbf{w}, Q(\Phi, \omega) \rangle$ and $Q_{\mathbf{w},\Phi}(B_1, \dots, B_{|\omega^*|})$ have the same distribution. We also assume w.l.o.g. that ω^* contains only relations that also appear in Φ (since the rest of the relations in ω^* do not influence the values $Q(\Phi, \omega)$). We denote by $\mathcal{R}_\Phi \subseteq \mathcal{R}$ the set of relations present in Φ.

From McDiarmid's inequality [1] we have

$$P[|Q_{\mathbf{w},\Phi}(B_1, \dots, B_{|\omega^*|}) - \mathbb{E}[Q_{\mathbf{w},\Phi}]| \geq \varepsilon] \leq 2 \cdot \exp\left(\frac{-2\varepsilon^2}{\sum_{j=1}^{|\omega^*|} c_j^2} \right) \quad (1)$$

provided that $|Q_\alpha(B_1, \dots, B_j, \dots, B_{\omega^*}) - Q_\alpha(B_1, \dots, B'_j \dots, B_{\omega^*})| \leq c_j$ holds for every j and every value of B_j and B'_j.

It follows from Lemma 1 that we can set $c_j := \sum_{k=1}^m \|\mathbf{w}\| \cdot |\alpha_k| \cdot |\Delta|^{-A_j}$, where w_k is the k-th component of the weight vector $\mathbf{w}$ and A_j is the arity of the atom a_j, in (1).

Let us split ω^* into disjoint subsets ω_1^*, ω_2^*, $\dots$, ω_M^* where each ω_i^* contains all atoms from ω^* with exactly i unique constants. Then we can write

$$\sum_{j=1}^{|\omega^*|} c_j^2 = \sum_{i=1}^{|\omega^*|} \left(\|\mathbf{w}\| \cdot \sum_{i=1}^m |\alpha_i| \cdot |\Delta|^{-A_i} \right)^2 = |\omega_1^*| \cdot \left(\frac{\|\mathbf{w}\| \cdot \sum_{i=1}^m |\alpha_i|}{|\Delta|} \right)^2 + \dots + |\omega_M^*| \cdot \left(\frac{\|\mathbf{w}\| \cdot \sum_{i=1}^m |\alpha_i|}{|\Delta|^M} \right)^2. \quad (2)$$

We can also bound every $|\omega_i^*|$ as $|\omega_i^*| \leq i^{M-1} \cdot |\mathcal{R}_\alpha| \cdot |\Delta|^i$. By substituting this into (2) and assuming that $|\Delta| \geq (M+1)^M$, we obtain (we omit here the detailed algebraic manipulations which are the same as in the proof of Theorem 5)

$$\sum_{j=1}^{|\omega^*|} c_j^2 \leq |\mathcal{R}_\Phi| \cdot \|\mathbf{w}\|^2 \cdot \left(\sum_{i=1}^m |\alpha_i| \right)^2 \cdot \left(\frac{1}{|\Delta|} + \frac{2^{M-1}}{|\Delta|} + \dots + \frac{M^{M-1}}{|\Delta|^M} \right) \leq 2 \cdot \frac{|\mathcal{R}_\Phi| \cdot \|\mathbf{w}\|^2 \cdot (\sum_{i=1}^m |\alpha_i|)^2}{|\Delta|}.$$

Finally, plugging this into (1) finishes the proof. □

References

[1] Colin McDiarmid. On the method of bounded differences. *Surveys in combinatorics*, 141(1):148–188, 1989.