Homework 2

Stefan Schmid, stefan.schmid@fhnw.ch

December 4, 2022

1 GCN

Question 1.1

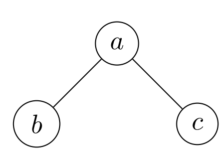
★ **Solution** ★ The two graphs are isomorphic. A bijective function f mapping from the graph on the left–hand side to the graph on the right–hand side is given by the following table

n	f(n)
1	A
2	D
3	H
4	E
5	B
6	C
7	G
8	F

It is easily verified, that this function preserves the edges and is thus a graph isomorphism. For example, the edge $(1, 5)$ is present in the graph on the RHS as $(A, B) = (f(1), f(5)$.

Question 1.2

★ **Solution** ★ We can take $V_1 = V_2$ and $E_1 = E_2$ give by



We set the initial features of all nodes to 1, i.e.

$$h_a^{(0)} = h_b^{(0)} = h_c^{(0)} = 1.$$

If we pick $v_1 \in V_1$ to be the node a and $v_2 \in V_2$ to be b, then the updated features after one iteration with $\max$ aggregation become

$$\begin{aligned} h_b^{(2)} &= \max\{h_a^{(0)}\} = 1 \\ h_a^{(2)} &= \max\{h_b^{(0)}, h_c^{(0)}\} = \max\{1, 1\} = 1 \end{aligned}$$

and with mean aggregation

$$\begin{aligned} h_b^{(2)} &= \mathrm{mean}\{h_a^{(0)}\} = 1 \\ h_a^{(2)} &= \mathrm{mean}\{h_b^{(0)}, h_c^{(0)}\} = \mathrm{mean}\{1, 1\} = \frac{1}{2}(1+1) = 1. \end{aligned}$$

But for sum aggregation we obtain

$$\begin{aligned} h_b^{(2)} &= \mathrm{sum}\{h_a^{(0)}\} = 1 \\ h_a^{(2)} &= \mathrm{sum}\{h_b^{(0)}, h_c^{(0)}\} = 1 + 1 = 2. \end{aligned}$$

Question 1.3

★ **Solution** ★ Note that for each iteration i, if $l_v^{(i)} = l_u^{(i)}$, $v, u \in V_1 \dot{\cup} V_2$, then $h_v^{(i)} = h_u^{(i)}$. This can be shown by induction on i. Clearly the assumption is true if $i = 0$ since $l_v^{(0)} = x(v) = h_v^{(0)}$ holds for any node in V_1 and V_2. Assume our claim is true for i. If $l_v^{(i+1)} = l_u^{(i+1)}$, then by the injectivity of the HASH function

$$\begin{aligned} l_v^{(i)} &= l_u^{(i)} \text{ and} \\ \{\!\{l_x^{(i)}; x \in \mathcal{N}(v)\}\!\} &= \{\!\{l_x^{(i)}; x \in \mathcal{N}(u)\}\!\}. \end{aligned} \tag{1}$$

This defines a bijection $\varphi\colon \mathcal{N}(v) \to \mathcal{N}(u)$ such that $l_x^{(i)} = l_{\varphi(x)}^{(i)}$ for any $x \in \mathcal{N}(v)$. Therefore $h_x^{(i)} = h_{\varphi(x)}^{(i)}$ by the induction hypothesis. This in turn implies

$$\{\!\{h_x^{(i)}; x \in \mathcal{N}(v)\}\!\} = \{\!\{h_{\varphi(x)}^{(i)}; x \in \mathcal{N}(v)\}\!\} = \{\!\{h_x^{(i)}; x \in \mathcal{N}(u)\}\!\}$$

because φ is a bijection between the neighborhoods of v and u. Due to the equality of multi–sets we also have $h_{\mathcal{N}(v)}^{(i+1)} = h_{\mathcal{N}(u)}^{(i+1)}$, i.e. the aggregation step yields the same result. By (1) and the induction hypothesis we also have $h_v^{(i)} = h_u^{(i)}$, so that the combine step also gives us $h_v^{(i+1)} = h_u^{(i+1)}$. This gives us a well–defined map $\Phi^{(i)}\colon l_v^{(i)} \mapsto h_v^{(i)}$.

Assume that the WL test returns true, i.e.

$$\{\{l_v^{(K)};\ v \in V_1\}\} = \{\{l_v^{(K)};\ v \in V_2\}\}.$$

Then

$$l_v^{(K)} = l_{\psi(v)}^{(K)}$$

for some bijection $\psi\colon V_1 \to V_2$, and thus by the induction above

$$h_v^{(K)} = h_{\psi(v)}^{(K)}$$

and

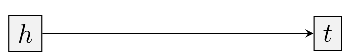
$$\{\{h_v^{(K)}; v \in V_1\}\} = \{\{h_{\psi(v)}^{(K)}; v \in V_1\}\} = \{\{h_v^{(K)}; v \in V_2\}\}$$

i.e. our algorithm decides that the graphs are isomorphic.

2 Node Embeddings with TransE

Question 2.1

★ **Solution** ★ A simple graph with two nodes and one relation between them does the trick:



The embeddings $\mathbf{h} = (1, 0)$, $\mathbf{l} = (0, 0)$, and $\mathbf{t} = (1, 0)$ minimize the loss to 0 but we can not tell from the embeddings only if $(h, l, t) \in S$ or $(t, l, h) \in S$.

Question 2.2

★ **Solution** ★ The same graph and embeddings as in Question 2 minimizes the loss to 0 because the set $S'_{(h,l,t)}$ is empty. This is due to the fact that only h or t may be corrupted and thus (t, l, h) is not a valid corrupted triple. But there are no other possibilities.

For a graph that allows the set of corrupted triples to be non–empty we can choose the graph

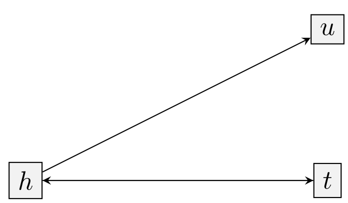
We have $S'_{(h,l,t)} = \{(u,l,t),(h,l,u)\}$ ($S = \{(h,l,t)\}$). Setting $\mathbf{h} = \mathbf{t} = \mathbf{u} = (1,0)$ and $\mathbf{l} = (0,0)$ still minimizes the loss to 0. But so would $\mathbf{h} = (1,0), \mathbf{t} = (0,1), \mathbf{u} = (-1,0)$ with $\mathbf{l} = (-1,1)$ since $d(\mathbf{h}+\mathbf{l},\mathbf{t}) = 0$ and $-d(\mathbf{h'}+\mathbf{l},\mathbf{t'})$ is always negative.

Question 2.3

★ **Solution** ★ The algorithm could take embeddings that have very small L_2 norm and would thus minimize the loss. All the embeddings would be close to 0 and therefore the distance between $\mathbf{h}+\mathbf{l}$ and $\mathbf{t}$ would be very small.

Question 2.4

★ **Solution** ★ Assume we have a graph with three nodes h, t, u, a single relation and $S = \{(h,l,t),(t,l,h),(h,l,u)\}$.



A perfect embedding is not possible because TransE can not model symmetric relations. For a perfect embedding we would need $\mathbf{h}+\mathbf{l} = \mathbf{t}$ and $\mathbf{t}+\mathbf{l} = \mathbf{h}$. Now if $\mathbf{h} = (0,0)$, then $\mathbf{t} = \mathbf{l}$ and $\mathbf{t} = -\mathbf{l}$ which implies $\mathbf{t} = \mathbf{l} = (0,0)$. But then also $\mathbf{u} = \mathbf{h}+\mathbf{l} = (0,0)$ and therefore $(u,l,h) \in S$ which is not the case. We come to the same conclusion if $\mathbf{t} = (0,0)$. So we must have $\mathbf{h}, \mathbf{t} \neq (0,0)$ and therefore $\mathbf{l} = (0,0)$. But then again $\mathbf{h} = \mathbf{t} = \mathbf{u}$ and $(u,l,h) \in S$ which is not the case.

3 Expressive Power of Knowledge Graph Embeddings

Question 3.1

★ **Solution** ★ As discussed in Question 2 TransE cannot model symmetric relations. (The proof is independent of k.)

We can model inverse relations though. If l' describes the inverse relation to l then $\mathbf{l}' = -\mathbf{l}$. Clearly, if $(h, l, t) \in S$, then $\mathbf{h} + \mathbf{l} = \mathbf{t}$ and therefore $\mathbf{t} - \mathbf{l} = \mathbf{h}$ so that $(t, l', h) \in S$.

We can also model composition of relations. If we have two relations given by l and k and a relation j that is the composition of l and k, then we can simply put $\mathbf{j} = \mathbf{k} + \mathbf{l}$. To see this, assume $(h, l, t) \in S$ and $(t, k, u) \in S$. Then $\mathbf{h} + \mathbf{l} = \mathbf{t}$ and $\mathbf{t} + \mathbf{k} = \mathbf{u}$. Therefore,

$$\mathbf{h} + \mathbf{j} = \mathbf{h} + \mathbf{l} + \mathbf{k} = \mathbf{t} + \mathbf{k} = \mathbf{u}$$

which implies $(h, j, u) \in S$.

Question 3.2

★ **Solution** ★ To simplify the solution we assume that the embeddings of entities are given by a d–dimensional complex vectors and $\mathbf{l} \in \mathbb{C}^d$ with each entry on the unit circle. This means that $\mathbf{l} = (e^{i\varphi_1}, \dots, e^{i\varphi_d})$ for some rotation angles $\varphi_1, \dots, \varphi_d$. The $\mathbf{h} \circ \mathbf{l}$ can be achieved by multiplying $\mathbf{h}$ and $\mathbf{l}$ componentwise. Let us denote the componentwise multiplication by $\mathbf{h} \odot \mathbf{l}$ for this exercise. Assume symmetry so that $\mathbf{h} \circ \mathbf{l} = \mathbf{t}$ and $\mathbf{t} \circ \mathbf{l} = \mathbf{h}$. This implies $\mathbf{t} \times \mathbf{l} \times \mathbf{l} = \mathbf{t}$ and therefore $\varphi_1, \dots, \varphi_d \in \pi\mathbb{Z}$. This only leaves finitely many possibilities for $\mathbf{l}$ and thus the number of different symmetric relations is bounded. Hence, RotatE cannot model arbitrary many relations for fixed embedding space dimension k.

However, RotatE can model inverse relations. If $(h, l, t) \in S$ and l' is the inverse relation to l, then $\mathbf{h} \times \mathbf{l} = \mathbf{t}$, $\mathbf{l} = (e^{i\varphi_1}, \dots, e^{i\varphi_d})$, and $\mathbf{t} \times \mathbf{l}' = \mathbf{h}$ with $\mathbf{l}' = (e^{-i\varphi_1}, \dots, e^{-i\varphi_d})$.

Similarly to TransE we can also model composition of relations. The proof is the same as in 3 by replacing $+$ with $\times$.

Question 3.3

★ **Solution** ★ Similar to TransE, RotatE cannot model 1–to–N relations. Pick the following graph

$$
\begin{tikzcd}
 & \boxed{u} \\
\boxed{h} \arrow[ur] \arrow[r] & \boxed{t}
\end{tikzcd}
$$

where the relation could for example be "student of". Since $(h, l, t), (h, l, u) \in S$ we have $\mathbf{h} \times \mathbf{l} = \mathbf{t}$ and $\mathbf{h} \times \mathbf{l} = \mathbf{u}$ and therefore $\mathbf{t} = \mathbf{u}$. But we have $t \neq u$.

4 Honor Code

(X) I have read and understood Stanford Honor Code before I submitted my work.
$**$ Collaboration: Write down the names & SUNetIDs of students you collaborated with on Homework 2 (None if you didn't).$**$
$**$Note: Read our website on our policy about collaboration!$**$