Improving Knowledge Graph Embedding Using Simple Constraints: Supplymentary Material

Boyang Ding[1,2], **Quan Wang**[1,2,3*], **Bin Wang**[1,2], **Li Guo**[1,2]
[1]Institute of Information Engineering, Chinese Academy of Sciences
[2]School of Cyber Security, University of Chinese Academy of Sciences
[3]State Key Laboratory of Information Security, Chinese Academy of Sciences
{dingboyang,wangquan,wangbin,guoli}@iie.ac.cn

Sufficient Condition for Eq. (3)

Given the non-negativity constraints of Eq. (2), a sufficient condition for Eq. (3) to hold, is to further impose the strict entailment constraints of Eq. (4). In fact, given the constraints of Eq. (2) and Eq. (4), we will always have

$$\begin{aligned}\phi(e_i, r_p, e_j) &= \langle \text{Re}(\mathbf{e}_i), \text{Re}(\mathbf{r}_p), \text{Re}(\mathbf{e}_j)\rangle \\ &\quad + \langle \text{Im}(\mathbf{e}_i), \text{Re}(\mathbf{r}_p), \text{Im}(\mathbf{e}_j)\rangle \\ &\quad + \langle \text{Re}(\mathbf{e}_i), \text{Im}(\mathbf{r}_p), \text{Im}(\mathbf{e}_j)\rangle \\ &\quad - \langle \text{Im}(\mathbf{e}_i), \text{Im}(\mathbf{r}_p), \text{Re}(\mathbf{e}_j)\rangle \\ &\leq \langle \text{Re}(\mathbf{e}_i), \text{Re}(\mathbf{r}_q), \text{Re}(\mathbf{e}_j)\rangle \\ &\quad + \langle \text{Im}(\mathbf{e}_i), \text{Re}(\mathbf{r}_q), \text{Im}(\mathbf{e}_j)\rangle \\ &\quad + \langle \text{Re}(\mathbf{e}_i), \text{Im}(\mathbf{r}_q), \text{Im}(\mathbf{e}_j)\rangle \\ &\quad - \langle \text{Im}(\mathbf{e}_i), \text{Im}(\mathbf{r}_q), \text{Re}(\mathbf{e}_j)\rangle \\ &= \phi(e_i, r_q, e_j)\end{aligned}$$

for any two entities $e_i, e_j \in \mathcal{E}$, *i.e.*, Eq. (3). Here, the first two terms of the inequality hold because $\text{Re}(\mathbf{r}_p) \leq \text{Re}(\mathbf{r}_q)$, and the last two terms because $\text{Im}(\mathbf{r}_p) = \text{Im}(\mathbf{r}_q)$, given the condition that $\text{Re}(\mathbf{e})$, $\text{Im}(\mathbf{e}) \geq \mathbf{0}$ for every $e \in \mathcal{E}$.

Equivalence between Eq. (7) and Eq. (8)

We first rewrite the constraints of the optimization Eq. (7). Specifically, the two constraints

$$\boldsymbol{\alpha} \geq \lambda\big(\text{Re}(\mathbf{r}_p) - \text{Re}(\mathbf{r}_q)\big), \quad \boldsymbol{\alpha} \geq \mathbf{0}$$

can be rewritten as a single one, *i.e.*,

$$\boldsymbol{\alpha} \geq \lambda\big[\text{Re}(\mathbf{r}_p) - \text{Re}(\mathbf{r}_q)\big]_+,$$

where $[\mathbf{x}]_+ = \max(\mathbf{0}, \mathbf{x})$ with $\max(\cdot,\cdot)$ being an entry-wise operator. Similarly, the two constraints

$$\boldsymbol{\beta} \geq \lambda\big(\text{Im}(\mathbf{r}_p) - \text{Im}(\mathbf{r}_q)\big)^2, \quad \boldsymbol{\beta} \geq \mathbf{0}$$

degenerate to a single one, *i.e.*,

$$\boldsymbol{\beta} \geq \lambda\big(\text{Im}(\mathbf{r}_p) - \text{Im}(\mathbf{r}_q)\big)^2.$$

As the objective function of Eq. (7) has to minimize $\mathbf{1}^\top(\boldsymbol{\alpha}+\boldsymbol{\beta})$ over all possible $\boldsymbol{\alpha}, \boldsymbol{\beta}$, an optimal value for this term will be

$$\lambda\mathbf{1}^\top\big[\text{Re}(\mathbf{r}_p) - \text{Re}(\mathbf{r}_q)\big]_+ + \lambda\mathbf{1}^\top\big(\text{Im}(\mathbf{r}_p) - \text{Im}(\mathbf{r}_q)\big)^2.$$

Plugging this back into the objective function and removing the degenerated constraints, we will obtain the optimization of Eq. (8).

Properties of Equivalence, Inversion, and Ordinary Entailment

For ordinary entailment $r_p \rightarrow r_q$ (neither equivalence nor inversion), the constraints of Eq. (4) directly suggest

$$\text{Re}(\mathbf{r}_p) \leq \text{Re}(\mathbf{r}_q), \quad \text{Im}(\mathbf{r}_p) = \text{Im}(\mathbf{r}_q).$$

For equivalence $r_p \leftrightarrow r_q$ ($r_p \rightarrow r_q$ and $r_q \rightarrow r_p$), we ought to have

$$\begin{aligned}\text{Re}(\mathbf{r}_p) \leq \text{Re}(\mathbf{r}_q), \quad \text{Im}(\mathbf{r}_p) = \text{Im}(\mathbf{r}_q), \\ \text{Re}(\mathbf{r}_q) \leq \text{Re}(\mathbf{r}_p), \quad \text{Im}(\mathbf{r}_q) = \text{Im}(\mathbf{r}_p),\end{aligned}$$

which imply $\mathbf{r}_p = \mathbf{r}_q$. Since

$$\begin{aligned}\phi(e_i, r_k, e_j) &= \text{Re}(\langle \mathbf{e}_i, \mathbf{r}_k, \bar{\mathbf{e}}_j\rangle) \\ &= \text{Re}(\langle \mathbf{e}_j, \bar{\mathbf{r}}_k, \bar{\mathbf{e}}_i\rangle) \\ &\triangleq \phi(e_j, r_k^{-1}, e_i)\end{aligned}$$

for any $e_i, e_j \in \mathcal{E}$ and $r_k \in \mathcal{R}$, we could represent the inverse of relation r_k (*i.e.* r_k^{-1}) as the conjugate of $\mathbf{r}_k$ (*i.e.* $\bar{\mathbf{r}}_k$). Then for inversion $r_p \leftrightarrow r_q^{-1}$, we ought to have $\mathbf{r}_p = \bar{\mathbf{r}}_q$.

*Corresponding author: Quan Wang.

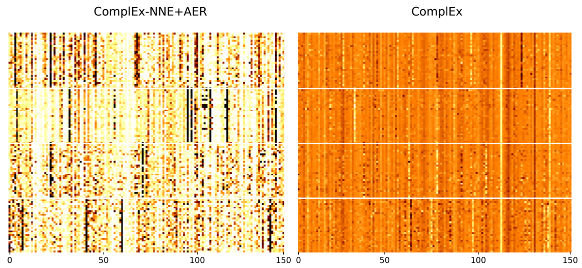

Figure 1: Visualization of imaginary components of entity representations (rows) learned by ComplEx-NNE+AER (left) and ComplEx (right). From top to bottom, entities belong to `reptile`, `wine_ region`, `species`, `programming_language` in turn. Values range from 0 (white) via 0.5 (orange) to 1 (black). Best viewed in color.

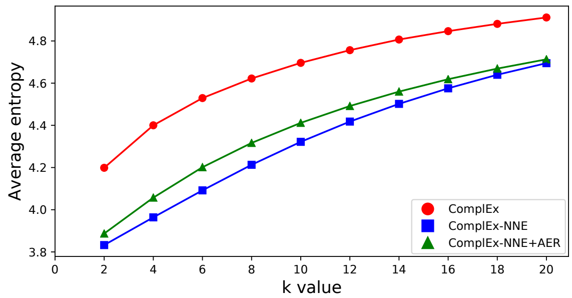

Figure 2: Average entropy over all dimensions of imaginary components of entity representations learned by ComplEx (circles), ComplEx-NNE (squares), and ComplEx-NNE+AER (triangles) as K varies.

Analyses on Imaginary Components of Entity Representations

We conduct the same analyses on imaginary components of entity representations as those conducted on real ones (§ 4.3). Figure 1 visualizes imaginary components of entity representations learned by ComplEx and ComplEx-NNE+AER, with the optimal configurations determined by link prediction. Figure 2 shows average entropy along imaginary components of entity representations learned by ComplEx, ComplEx-NNE, and ComplEx-NNE +AER.