Contextual Parameter Generation for Knowledge Graph Link Prediction

George Stoica*, Otilia Stretcu*, Emmanouil Antonios Platanios*, Tom M. Mitchell, Barnabás Póczos
{gis, ostretcu, e.a.platanios, tom.mitchell, bapoczos}@cs.cmu.edu

Knowledge Graphs

KGs offer a concise way in which to store factual information.

Link Prediction

Predict missing links to questions, (Vallteri Bottas, Drives For, ?)

Existing Solutions

• Single-Hop: Infer answer directly from question
• Multi-Hop: Infer answer by traversing the KG

Additive Limitations

Existing approaches with additive interactions cannot directly represent this toy example, e0, r0, e1, r1, e2, r2, e3, r3:

\[
\begin{align*}
& h_0(e_0, r) = \phi(e_0, r) = \phi_0 e_0 + \phi_r r_0 \\
& h_1(e_1, r) = \phi(e_1, r) = \phi_1 e_1 + \phi_r r_1 \\
& h_2(e_2, r) = \phi(e_2, r) = \phi_2 e_2 + \phi_r r_2 \\
& h_3(e_3, r) = \phi(e_3, r) = \phi_3 e_3 + \phi_r r_3
\end{align*}
\]

Solutions:

\[
\begin{align*}
& e_3 = e_2 \land (\phi_0 = 0 \lor r_0 = r_1)
\end{align*}
\]

CoPER

• General: Enhances expressive power of several models
• Simple: Implemented in just 10 LoCs
• Scalable: Improves convergence by up to 28x
• Performance: outperforms state-of-the-art

Convergence Comparisons

Results

Datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th># Train</th>
<th>N_s</th>
<th>N_r</th>
<th>N_t</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kinship</td>
<td>8,544</td>
<td>104</td>
<td>25</td>
<td>6.1</td>
<td>82.9</td>
</tr>
<tr>
<td>UMLS</td>
<td>5,216</td>
<td>46</td>
<td>7.8</td>
<td>26.6</td>
<td></td>
</tr>
<tr>
<td>FB19k237</td>
<td>272,115</td>
<td>14,541</td>
<td>237</td>
<td>3.0</td>
<td>17.9</td>
</tr>
<tr>
<td>WN18RR</td>
<td>86,835</td>
<td>40,945</td>
<td>11</td>
<td>1.4</td>
<td>22</td>
</tr>
<tr>
<td>NELL-995</td>
<td>154,213</td>
<td>75,492</td>
<td>200</td>
<td>3.6</td>
<td>4.1</td>
</tr>
</tbody>
</table>

Evaluation Datasets and Summary Statistics

Design Comparisons

Convergence Ratio between CoPER-ConvE and ConvE

Additive Limitations

Existing approaches with additive interactions cannot directly represent this toy example, e0, r0, e1, r1, e2, r2, e3, r3:

\[
\begin{align*}
& h_0(e_0, r) = \phi(e_0, r) = \phi_0 e_0 + \phi_r r_0 \\
& h_1(e_1, r) = \phi(e_1, r) = \phi_1 e_1 + \phi_r r_1 \\
& h_2(e_2, r) = \phi(e_2, r) = \phi_2 e_2 + \phi_r r_2 \\
& h_3(e_3, r) = \phi(e_3, r) = \phi_3 e_3 + \phi_r r_3
\end{align*}
\]

Solutions:

\[
\begin{align*}
& e_3 = e_2 \land (\phi_0 = 0 \lor r_0 = r_1)
\end{align*}
\]

Parameter Generators

Three of many choices,
• Parameter Lookup (PL): \( \theta = W \)
• Linear Transformation\((\theta = W \cdot r)\)
• Multilayer Perceptron\((\theta = \text{MLP}(r))\)

Relation Similarities

Learnt NELL-995 Relation Palaeo Cosine Distances

Repositories

https://github.com/otiliastr/cooper