Multilingual Knowledge Graph Embeddings for Cross-lingual Knowledge Alignment

Muhao Chen¹, Yingtao Tian², Mohan Yang¹, and Carlo Zaniolo¹
University of California, Los Angeles¹
Stony Brook University²
Outline

• Background
• \textit{MTransE}—A multilingual knowledge graph embedding model
• Evaluation
• Open Challenges and Future Work
Knowledge Graphs

- Symbolic representation of entities and relations

(California, capital city, Sacramento)

**Monolingual knowledge:** triples (relation facts of entities)

**Cross-lingual knowledge:** alignment of monolingual knowledge across languages
Knowledge Graph Embeddings

- Encode entities as vectors

Semantic similarity of entities

Relational inferences as vector calculation
- France – Paris ≈ capital
- US – USD ≈ currency
- Bach – German ≈ nationality
- ...

Applications
- KG Completion
- Relation extraction from text
- Question answering

Knowledge Graph

Encode

Bach
Male
Germany
Eisenach

Semantic similarity of entities

Capture

Enable

Applications

Embeddings

Paris (0.036, -0.12, …, 0.323)
capital (0.102, 0.671, …, -0.101)
France (0.138, 0.551, …, 0.222)
...
Current KG Embedding Approaches

- Focused on embedding **monolingual** triples \((h, r, t)\)

Later approaches
- TransH [Wang et al. 2014]
- TransR [Lin et al. 2015]
- TransD [Ji et al. 2015]
- HoLE [Nickle et al. 2016]
- ComplEx [Trouillon et al. 2016]
- ...

Embedding of **monolingual** knowledge seems to be well-addressed.

What about **cross-lingual** knowledge?
Emerging challenge

• Existing works do not characterize cross-lingual knowledge
  – Entity inter-lingual links (ILLs): (ambulance --- krankenwagen)
  – Triple-wise alignment (TWA): ((State of California, capital city, Sacramento) --- (カリフォルニア, 首都, サクラメント))
  – Many KGs store such knowledge

Why important?
• Enables multilingual semantic representations
• Benefits cross-lingual NLP
  – Knowledge alignment
  – Machine translation
  – Cross-lingual Q&A
  – ...

Difficult to characterize:
• Fewer samples: Cross-lingual knowledge currently accounts for a small portion of each KB
• Larger domains: Cross-lingual knowledge applies on the entire spaces of involved languages
• Incoherence: Language-specific versions of KG are usually incoherent
• Heterogeneity: Applies to both entities and monolingual relations with inconsistent vocabularies
What does MTransE use and enable?

- Corpora: (partially-aligned) multilingual KGs
- Enabling: inferable embeddings of multilingual semantics
- Can be applied to:
  - Knowledge alignment
  - Cross-lingual Q&A
  - Multilingual chat-bots
  - ...
MTransE Model Components

• Knowledge model

\[ S_K = \sum_{(T, T') \in \delta(L_i, L_j)} S_a(T, T') \]

• Alignment model

\[ S_A = \sum_{(T, T') \in \delta(L_i, L_j)} S_a(T, T') \]

• Objective of learning
  - Minimizing \( J(\theta) = S_K + \alpha S_A \)
Different alignment techniques

- Translation vectors
  - Encoding cross-lingual transitions just like monolingual relations

- Linear Transformations
  - Transformations across embedding spaces of different languages

- Axis calibration
  - Cross-lingual counterparts have close embeddings
## Alignment Scores and Five Model Variants

- \( \text{Var}_i \) combines the \( i^{th} \) alignment model with the knowledge model

<table>
<thead>
<tr>
<th>Variant</th>
<th>Alignment Score</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{Var}_1 )</td>
<td>( S_{a_1} = |h - h'| + |t - t'| )</td>
<td>\n</td>
</tr>
<tr>
<td>( \text{Var}_2 )</td>
<td>( S_{a_2} = |h - h'| + |r - r'| + |t - t'| )</td>
<td>\n</td>
</tr>
<tr>
<td>( \text{Var}_3 )</td>
<td>( S_{a_3} = |h + v_{ij}^e - h'| + |r + v_{ij}^r - r'| + |t + v_{ij}^e - t'| ) ( v_{ij}^e = -v_{ei}^e, v_{ij}^r = -v_{ji}^r )</td>
<td>\n</td>
</tr>
<tr>
<td>( \text{Var}_4 )</td>
<td>( S_{a_4} = |M^e_{ij}h - h'| + |M^e_{ij}t - t'| ) ( M_{ij}^e \in \mathbb{R}^{k \times k}, M_{ij}^r \in \mathbb{R}^{k \times k} )</td>
<td>\n</td>
</tr>
<tr>
<td>( \text{Var}_5 )</td>
<td>( S_{a_5} = |M_{ij}^e h - h'| + |M_{ij}^r r - r'| + |M_{ij}^e t - t'| )</td>
<td>\n</td>
</tr>
</tbody>
</table>

**Axis Calibration**
- Linear Transforms

**Translation Vector**

**Linear Transforms**
Experimental Evaluation

• Cross-lingual knowledge alignment tasks
  – Entity Matching
  – Triple-wise Alignment (TWA) Verification
• Monolingual relation extraction task
• Trilingual data sets
  – Wiki-based (WK3l-15k, WK3l-120k)
  – ConceptNet-based (CN3l)
• Baselines
  – CCA [Faruqui et al. 2014] + Knowledge mode
  – OT [Xing et al. 2015] + Knowledge models

These three data sets are available at https://github.com/muhaochen/MTransE
Entity Matching

• Evaluation protocol
  – For each \((e, e')\), rank \(e'\) in the neighborhood of \(\tau(e)\)

• Training sets
  – Pairs of language-specific graphs and corresponding alignment sets

• Test data
  – Entity Inter-lingual links \(\{(e, e')\}\) (Unidirectional)

What is the German entity for the English entity “Regulation of Property”? 
**Entity Matching**

*Hits@10 on WK3l-15k*

*Hits@10 on WK3l-120k*

*Hits@10 on CN3l*

*Mean on WK3l-15k*

*Mean on CN3l*

\[\text{Var}_4 \approx \text{Var}_5 > \text{Var}_1 \approx \text{Var}_3 \approx \text{OT} > \text{Var}_2 \gg \text{CCA} > \text{LM}\]
Triple-wise Alignment Verification

We receive similar evaluation conclusions in all settings.

\[ Var_4 \approx Var_5 > Var_1 > Var_2 > Var_3 \approx OT \gg CCA > LM \]
Monolingual Relation Extraction (English, French)

- Train/Test
  - Train Sets: 90% triples and intersecting alignment sets
  - Test Sets: 10% triples
- MTransE preserves well the monolingual relations

<table>
<thead>
<tr>
<th>Axis Calibration</th>
<th>$\text{Var}_1$, $\text{Var}_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trans. Vectors</td>
<td>$\text{Var}_3$</td>
</tr>
<tr>
<td>Linear Transforms</td>
<td>$\text{Var}_4$, $\text{Var}_5$</td>
</tr>
</tbody>
</table>
Applications based on MTransE

• Multilingual Q&A

• Cross-lingual relation prediction

• Improving monolingual KG completion using multilingual correlation

• Knowledge alignment across knowledge bases
Examples of Cross-lingual Question Answering

<table>
<thead>
<tr>
<th>Query</th>
<th>Target</th>
<th>Candidates (in ascending order of rank)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Adam Lambert, genre, ?r)</td>
<td>French</td>
<td></td>
</tr>
<tr>
<td></td>
<td>German</td>
<td>popmusik, dance-pop, no wave, soul</td>
</tr>
<tr>
<td>(Ronaldinho, position, ?t)</td>
<td>French</td>
<td>milieu offensif,attaquant,quarterback, latéral gauche</td>
</tr>
<tr>
<td></td>
<td>German</td>
<td>stürmer, linker flügel, angriffsspieler, rechter flgel</td>
</tr>
<tr>
<td>(Italy, ?r, Rome)</td>
<td>French</td>
<td>capitale, plus grande ville, chef-lieu, garnison</td>
</tr>
<tr>
<td></td>
<td>German</td>
<td>hauptstadt, hauptort, verwaltungssitz, stadion</td>
</tr>
<tr>
<td>(Barack Obama, ?r, George Bush)</td>
<td>French</td>
<td>ministre-président, prédécesseur, premier ministre, président du conseil</td>
</tr>
<tr>
<td></td>
<td>German</td>
<td>vorgänger, vorgängerin, besetzung, lied</td>
</tr>
<tr>
<td>(?h, instrument, guitar)</td>
<td>French</td>
<td>Brant Bjork, Chris Garneau, David Draiman, Ian Mackaye</td>
</tr>
<tr>
<td></td>
<td>German</td>
<td>Phil Manzanera, Styles P, Tina Charles, Luke Bryan</td>
</tr>
</tbody>
</table>

**Bold-faced** ones are correct answers, *italic* ones are close answers.
Improve the embedding model

- Other forms of knowledge models and alignment models
  - Neural knowledge models such as HolE and ComplEx
  - Other alignment models such as affine transformations
  - Alignment models which consider disambiguation

- Encoding more information from multilingual KGs
  - Entity domains, class templates, entity descriptions, etc
  - Cross-lingual disambiguation

- Jointly embedding with other forms of corpora such as multilingual documents
References

• [Xing et al., 2015] Chao Xing, Dong Wang, Chao Liu, and Yiye Lin. Normalized word embedding and orthogonal transform for bilingual word translation. In NAACL HLT, pages 1006–1011, 2015.
Thank You