Co-training Embeddings of Knowledge Graphs and Entity Descriptions for Cross-lingual Entity Alignment

Muhao Chen¹, Yingtao Tian², Kai-Wei Chang¹, Steven Skiena², and Carlo Zaniolo¹

¹University of California, Los Angeles
²Stony Brook University
Outline

• Background
• *KDCoE*—A multilingual knowledge graph embedding model
• Evaluation
• Future Work
Multilingual Knowledge Bases

- Symbolic representation of entities and relations in different languages
  + Accompanying literal knowledge (entity descriptions)

Monolingual knowledge: relation facts of entities (Triples)

EN triple: (Ulugh Beg, occupation, astronomer)
FR triple: (Ulugh Beg, activité, astronome)

Cross-lingual knowledge: alignment of monolingual knowledge

Inter-lingual Link (ILL): (astronomer@EN, astronome@FR)

Literal knowledge: entity descriptions

An astronomer is a scientist in the field of astronomy who concentrates their studies on a specific question or field outside of the scope of Earth...

Un astronome est un scientifique spécialisé dans l’étude de l’astronomie...
Multilingual Knowledge Graph Embeddings

• Multilingual KG Embeddings

Entities

Separate embedding spaces

Paris (0.036, -0.12, ..., 0.323)
France (0.138, 0.551, ..., 0.222)
...

Monolingual Relations

Semantic Transfer

Transformations

(Monolingual) vector algebraic operations

(Cross-lingual) transforms of embedding spaces

Applications

- Knowledge alignment
- Phrasal translation
- Causality reasoning
- Cross-lingual QA
- etc...
Existing Approaches

**MTransE** [Chen et al. 2017a; 2017b]
- Joint learning of structure encoders and an alignment model
- Alignment techniques: Linear transforms (best), vector translations, collocation (minimizing L2 distance)

**JAPE** [Sun et al. 2017]
+ Logistic-based proximity normalizer for entity attributes

**ITransE** [Zhu et al. 2017]
- self-training + cross-lingual collocation of entity embeddings

**PSG** [Yeo et al. 2018]

**Transformations+Translation** [Otani et al. 2018]

...
Critical Challenges

• Inconsistent monolingual knowledge

• Insufficient cross-lingual seed alignment

• Zero-shot scenarios

• What if some entities do not appear in the KG structure?

• Language-specific embedding spaces are highly incoherent

• Require semi-supervised cross-lingual learning

• Inducing a large portion entity alignment (e.g. 80%) based on a very small portion (20%) is extremely challenging
**KDCoE-Knowledge Graph and Entity Descriptions Co-training Embeddings**

- Embedding KG and entity descriptions for semi-supervised cross-lingual learning
- Encoding two types of knowledge
  1. Weakly-aligned KG structures
  2. Literal descriptions of entities in each language
- Iterative co-training of two model components
  1. A multilingual KG embedding model (KGEM)
  2. An entity description embedding model (DEM)
**KG Structure Embedding Model (KGEM)**

MTransE-LT [Chen et al. 2017a]

- **Knowledge model**

\[
S_K = \sum_{L \in \{L_i, L_j\}} \sum_{(h, r, t) \in G_L \land (\hat{h}, r, \hat{t}) \in G_L} \left[ f_r(h, t) - f_r(\hat{h}, \hat{t}) + \gamma \right]_+
\]

s.t. \( f_r(h, t) = \|h + r - t\|_2 \)

- **Alignment model**

\[
S_A = \sum_{(e, e') \in I(L_i, L_j)} \|M_{ij}e - e'\|_2
\]

- **Learning objective function**

\[
S_{KG} = S_K + \alpha S_A
\]

To capture monolingual KG structures in, and cross-lingual semantic transfer across separated embedding spaces.
Entity Description Embedding Model (DEM)

Siamese Attentive GRU + Pre-trained BilBOWA embeddings [Gouws et al. 2015]

Logistic loss

\[ S_D = \sum_{(e,e') \in I(L_1,L_2)} -LL_1 - LL_2 \]

\[ LL_1 = \log \sigma(d_e^T d_{e'}) + \sum_{k=1}^{k \sim U(e_k \in E_{L_i})} \mathbb{E} [\log \sigma(-d_{e_k}^T d_{e'})] \]

\[ LL_1 = \log \sigma(d_e^T d_{e'}) + \sum_{k=1}^{k \sim U(e_k \in E_{L_j})} \mathbb{E} [\log \sigma(-d_e^T d_{e_k})] \]

Stratified negative sharing [Chen et al. 2017c]

- Efficiently sharing negative samples within a batch

To collocate the embeddings of multilingual entity description counterparts

Logistic Loss + Stratified negative sharing

An astronomer is a scientist in the field of astronomy who concentrates their studies on a specific question or field outside of the scope of Earth...

Un astronome est un scientifique spécialisé dans l’étude de l’astronomie...
Iterative Co-training Process

1. Train MTransE-LT until converge
2. Propose seed alignment with high confidence using KG Embeddings
3. Train the bilingual description embedding model until converge
4. Propose seed alignment with high confidence using description embeddings

Unaligned entities
Seed alignment

Encoder

EN
FR

EN
FR
Experimental Evaluation

• WK3l-60k Dataset: Wikipedia-based trilingual KG with entity descriptions
• Knowledge alignment tasks
  1. Semi-supervised entity alignment (use around 20% seed alignment to predict the rest)
  2. Zero-shot alignment (entities do not appear in KG for training)
• Cross-lingual KG completion

<table>
<thead>
<tr>
<th>Data</th>
<th>#En</th>
<th>#Fr</th>
<th>#De</th>
<th>ILL Lang</th>
<th>#Train</th>
<th>#Valid</th>
<th>#Test</th>
<th>#Zero-shot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trippes</td>
<td>569,393</td>
<td>258,337</td>
<td>224,647</td>
<td>En-Fr</td>
<td>13,050</td>
<td>2,000</td>
<td>39,155</td>
<td>5,000</td>
</tr>
<tr>
<td>Desc.</td>
<td>67,314</td>
<td>45,842</td>
<td>43,559</td>
<td>En-De</td>
<td>12,505</td>
<td>2,000</td>
<td>41,018</td>
<td>5,632</td>
</tr>
</tbody>
</table>

Table 1: Statistics of the Wk3l60k dataset.
Entity Alignment

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

• Baselines
  – MTransE variants [Chen et al. 2017a]
  – ITransE [Zhu et al. 2017]
  – OT [Xing et al. 2015] + TransE

• Metrics
  - Hits@1, Hits@10, MRR

What is the German entity for the English entity “Regulation of Property”?
• MTransE-LT (same as KDCoE iteration 1) performs better than other baselines.
• KDCoE gradually improves the performance through each iteration of co-training, and eventually almost doubles Hit@1.
Zero-shot Entity Alignment

Induce the embeddings of unseen entities based on their descriptions (in either language)

- AttGRU + BiLbowa represents the best description representation technique.
- Within iterations of co-training, KDCoE gradually improves zero-shot alignment of entities that do not appear in the KG structure.
Preliminary Results of Cross-lingual KG Completion

A new KG completion approach based on cross-lingual knowledge transfer:

- Given a query \((h, r, ?t)\) in a less populated language version of KG (Fr, De), transfer the query to the intermediate embedding space of a well-populated version of KG (EN), then transfer the answer back.
- Preliminary results show plausibility of this new approach.
- How about ensemble models on multiple bridges of languages to co-populate few target languages?
Future Work

• Learning approaches
  - Empirical studies on other forms of KGEM
  - Ensemble models on multiple bridges to improve cross-lingual KG completion
  - Other approaches to leverage entity descriptions (e.g. weak and strong word pairs [Tissier et al. 2017])

• Applications
  - Cross-lingual semantic search of entities (based on natural language descriptions).
  - Cross-lingual Wikification.
References

Thank You