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

Muhao Chen[1], Yingtao Tian[2], Kai-Wei Chang[1], Steven Skiena[2], and Carlo Zaniolo[1]
[1]University of California, Los Angeles
[2]Stony 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)

Cross-lingual knowledge: alignment of monolingual knowledge

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

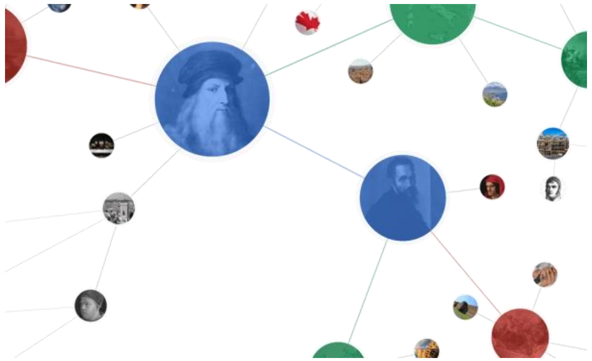
EN triple: (***Ulugh Beg***, *occupation*, ***astronomer***)

FR triple: (***Ulugh Beg***, *activité*, ***astronome***)

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...

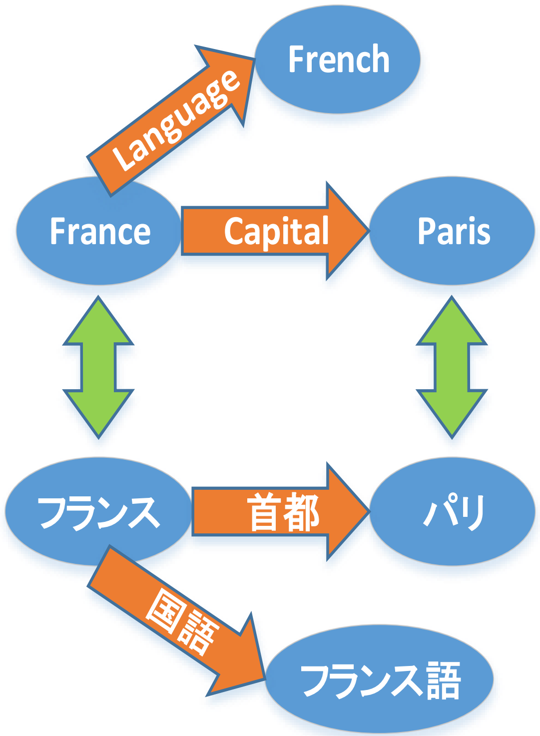
Un astronome est un scientifique spécialisé dans l'étude de l'astronomie...

Literal knowledge: entity descriptions



Multilingual Knowledge Graph Embeddings

- **Multilingual KG Embeddings**

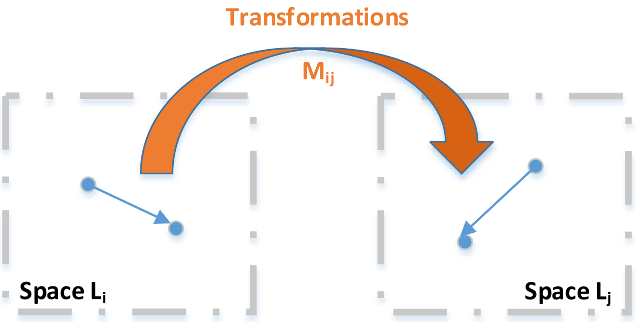


Entities

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

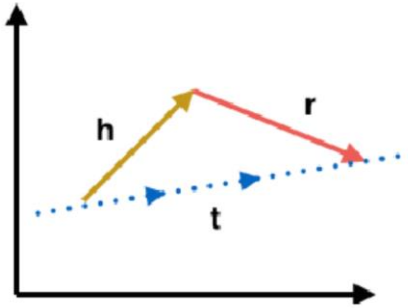
Semantic Transfer

Separated embedding spaces



(Cross-lingual) transforms of embedding spaces

Monolingual Relations



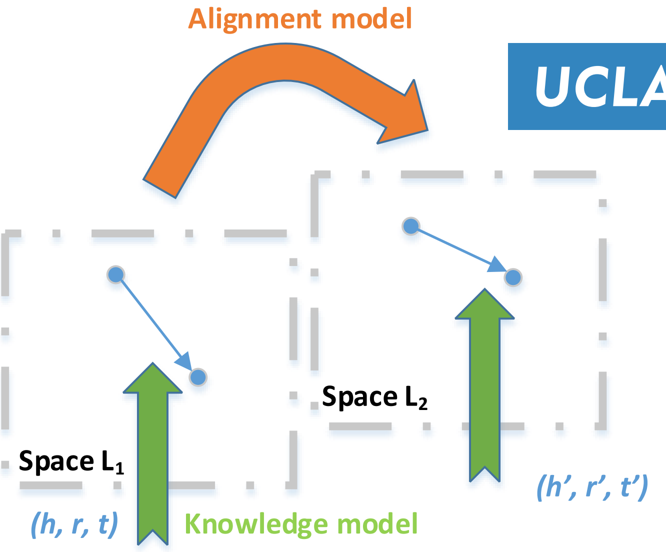
(Monolingual) vector algebraic operations

- **Applications**
  - Knowledge alignment
  - Phrasal translation
  - Causality reasoning
  - Cross-lingual QA
  - etc..

Existing Approaches

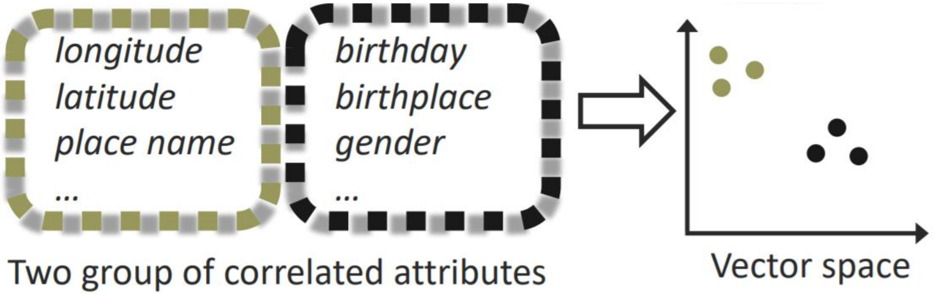
MTransE [Chen et al. 2017a; 2017b]

- Joint learning of structure encoders and an alignment model
- Alignment techniques: Linear transforms (best), vector translations, collocation (minimizing L2 distance)



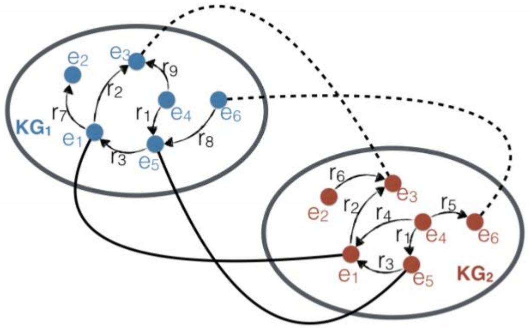
JAPE [Sun et al. 2017]

+ 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

- **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**

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

KDCoE-**K**nowledge Graph and **E**ntity **D**escriptions **Co**-training **E**mbeddings

UCLA

- 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]

TransE encoders for each langauge

- *Knowledge model*

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

s.t. $f_r(h,t) = \|\mathbf{h} + \mathbf{r} - \mathbf{t}\|_2$

Linear transformation induced from cross-lingual seed aligment

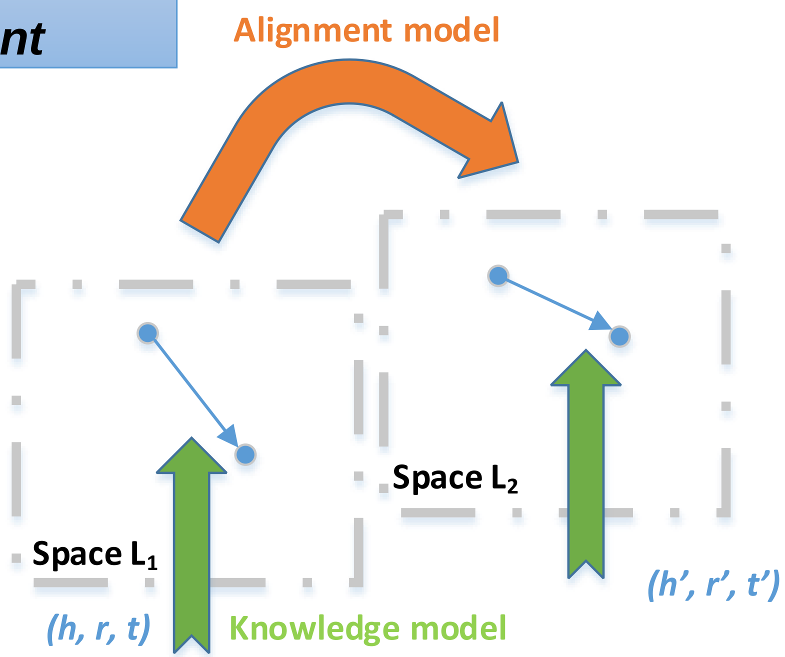
- *Alignment model*

$$S_A = \sum_{(e,e') \in I(L_i, L_j)} \left\| \mathbf{M}_{ij} \mathbf{e} - \mathbf{e}' \right\|_2$$

- *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]

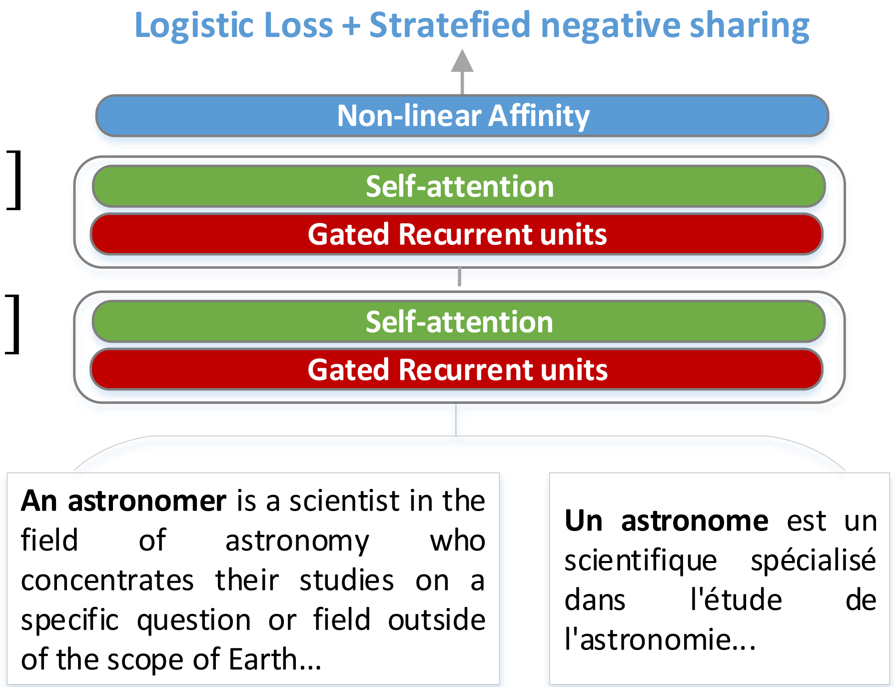
To **collocate** the embeddings of **multilingual entity description counterparts**

Logistic loss

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

$$LL_1 = \log\sigma(\mathbf{d}_e^\top \mathbf{d}_{e'}) + \sum_{k=1}^{|B_d|} \mathbb{E}_{e_k\sim \mathrm{U}(e_k\in E_{L_i})}[\log\sigma(-\mathbf{d}_{e_k}^\top \mathbf{d}_{e'})]$$

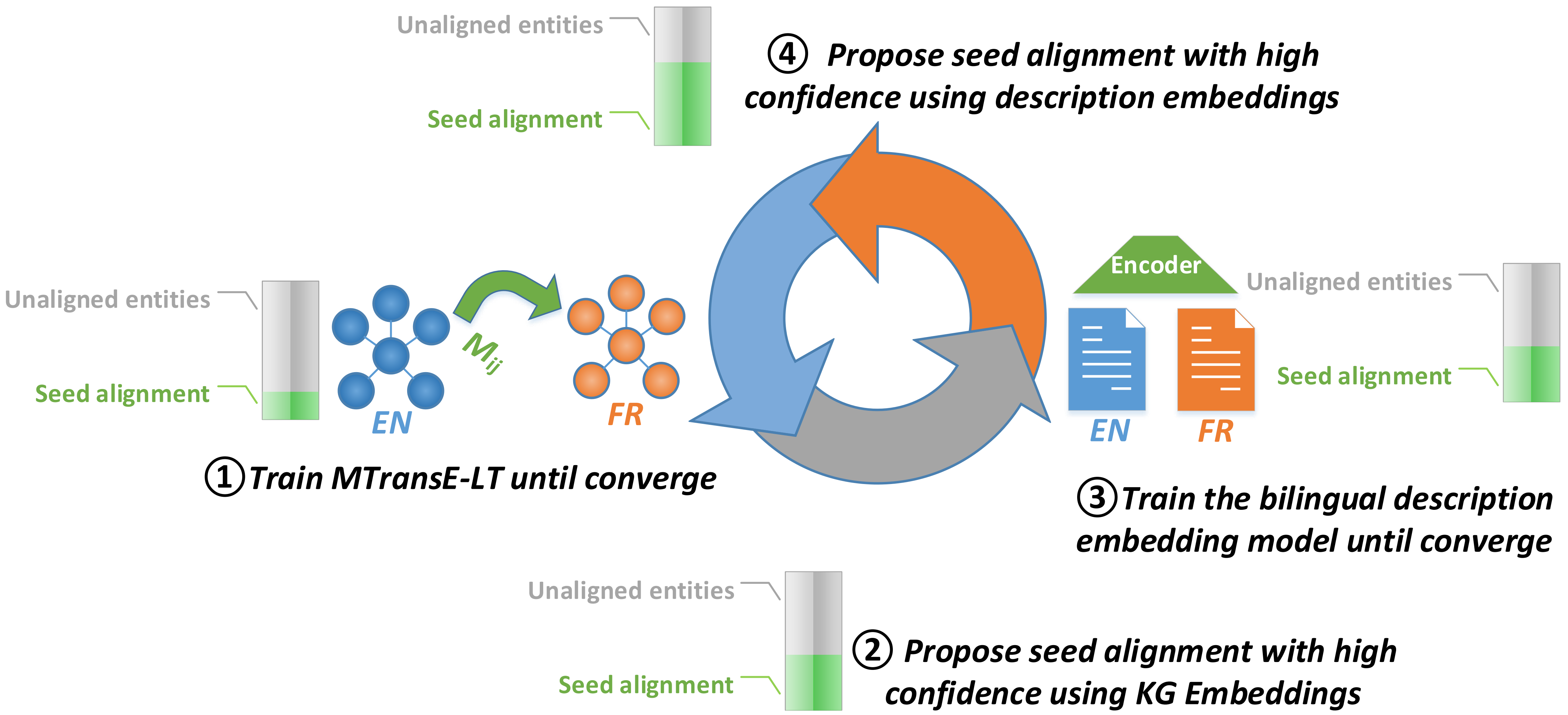
$$LL_1 = \log\sigma(\mathbf{d}_e^\top \mathbf{d}_{e'}) + \sum_{k=1}^{|B_d|} \mathbb{E}_{e_k\sim \mathrm{U}(e_k\in E_{L_j})}[\log\sigma(-\mathbf{d}_{e}^\top \mathbf{d}_{e_k})]$$



Stratified negative sharing [Chen et al. 2017c]

- Efficiently sharing negative samples within a batch

Iterative Co-training Process



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

Data	#En	#Fr	#De	ILL Lang	#Train	#Valid	#Test	#Zero-shot
Triples	569,393	258,337	224,647	En-Fr	13,050	2,000	39,155	5,000
Desc.	67,314	45,842	43,559	En-De	12,505	2,000	41,018	5,632

Table 1: Statistics of the Wk3l60k dataset.

Entity Alignment

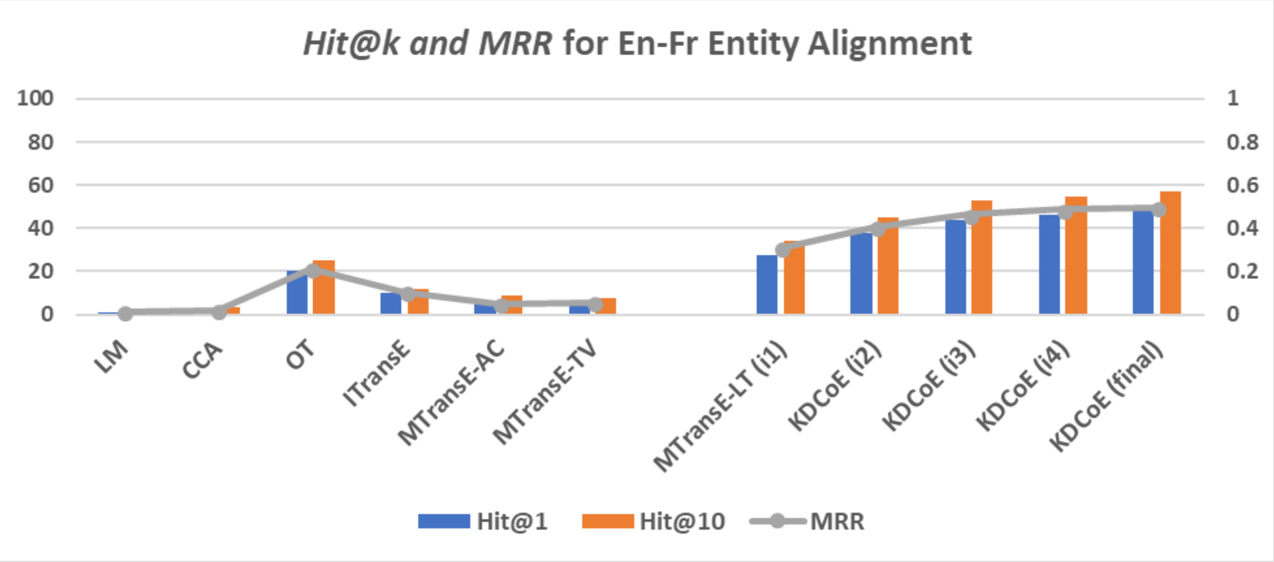
What is the German entity for the English entity "Regulation of Property"?

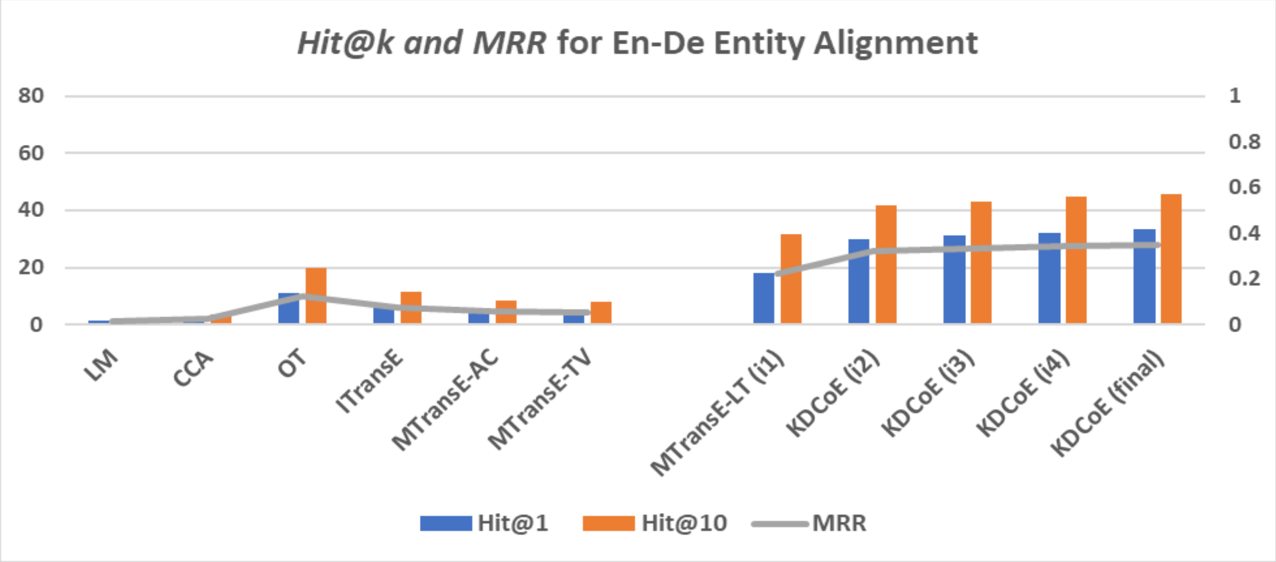
UCLA

- Evaluation protocol
  – For each (e, e'), rank $\boldsymbol{e'}$ in the neighborhood of $\tau(\boldsymbol{e})$
- Baselines
  – MTransE variants [Chen et al. 2017a]
  – ITransE [Zhu et al. 2017]
  – LM [Mikolov et al. 2013] + TransE
  – CCA [Faruqui et al. 2014] + TransE
  – OT [Xing et al. 2015] + TransE
- Metrics
  - Hits@1, Hits@10, MRR

Entity Alignment

UCLA

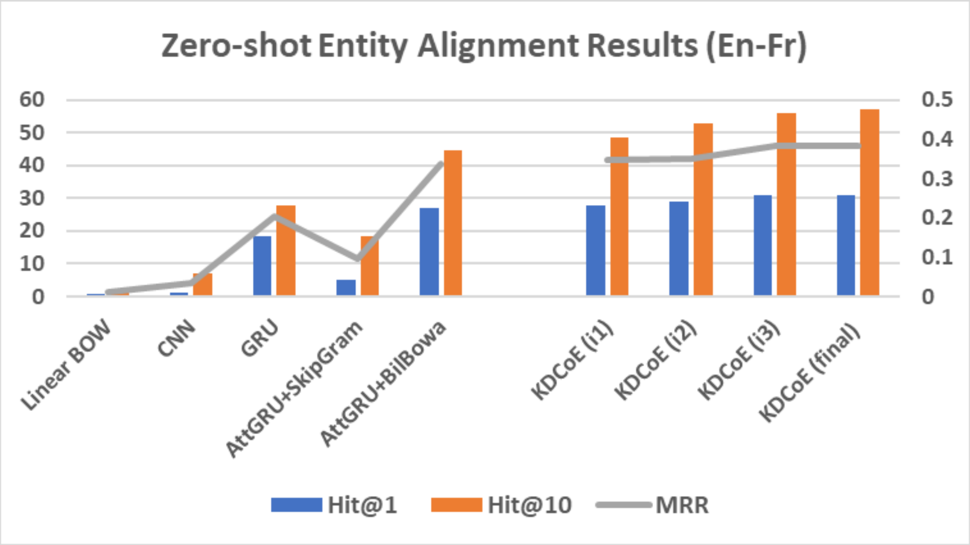


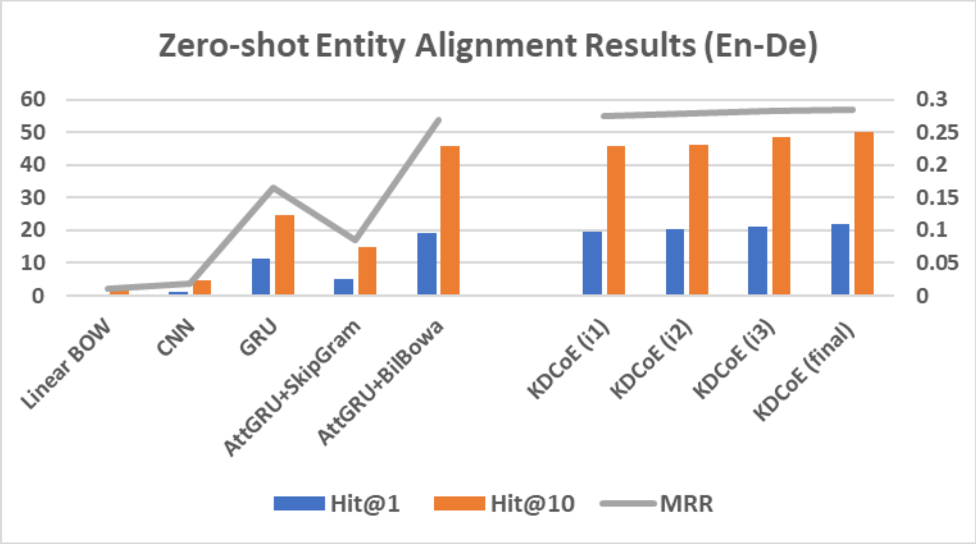


- MTransE-LT (same as KDCoE iteration 1) performs better than other baselines.
- KDCoE gradually improves the performance through each iteration of co-training, and eventually almost doubles Hit@1.

Zero-shot Entity Alignment

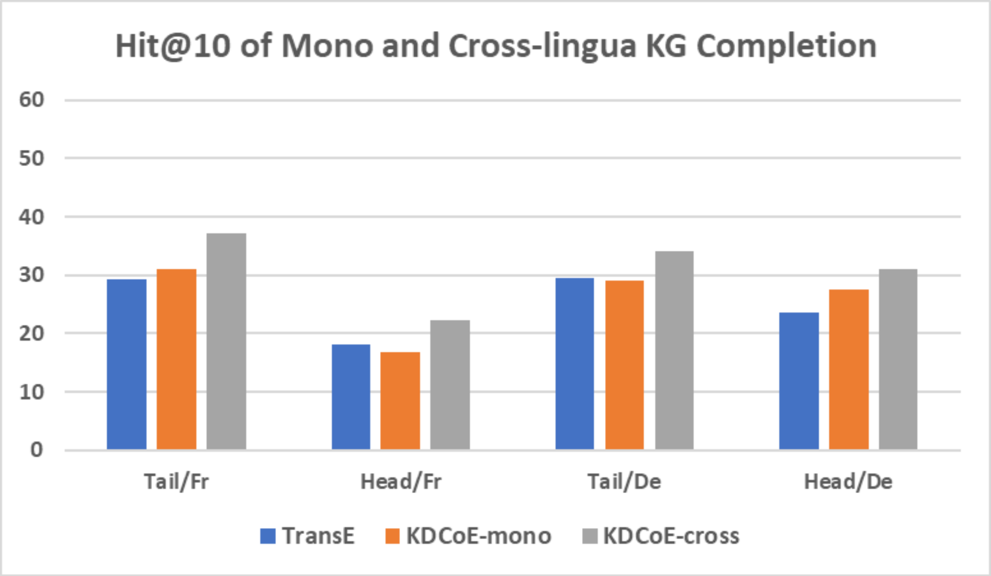
Induce the embeddings of unseen entities based on their descriptions (in either language)

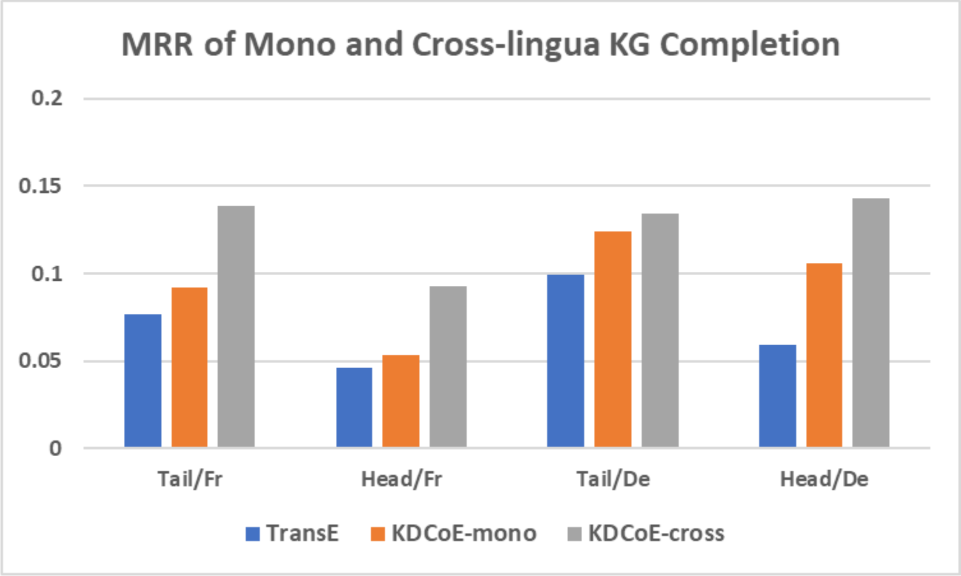




- AttGRU + BilBowa represents the best description representation technique.
- Within iterations of co-training, KDCoE gradually improves zero-shot alignment of entities that do not appear in the KG structure.

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

UCLA

- 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.

UCLA

References

1. [Tissier et al. 2017] Tissier, Julien, et al. "Dict2vec: Learning Word Embeddings using Lexical Dictionaries." EMNLP. 2017.
2. [Chen et al. 2017a] Chen, Muhao, et al. "Multilingual knowledge graph embeddings for cross-lingual knowledge alignment." IJCAI. 2017.
3. [Chen et al. 2017b] Chen, Muhao, et al. "Multi-graph Affinity Embeddings for Multilingual Knowledge Graphs." AKBC. 2017
4. [Chen et al. 2017c] Chen, Ting, et al. "On Sampling Strategies for Neural Network-based Collaborative Filtering,". KDD. 2017
5. [Mikolov et al. 2013] Mikolov, Tomas, et al. "Exploiting similarities among languages for machine translation. CoRR, 2013.". CoRR. 2013.
6. [Faruqui et al. 2014] Faruqui, Manaal, et al. "Improving vector space word representations using multilingual correlation." EACL, 2014.
7. [Xing et al. 2015] Xing, Chao, et al. "Normalized word embedding and orthogonal transform for bilingual word translation." NAACL, 2015.
8. [Zhu et al. 2017] Zhu, Hao, et al. "Iterative entity alignment via knowledge embeddings." IJCAI, 2017.
9. [Gouws et al. 2015] Gouws, Stephan, et al. "Bilbowa: Fast bilingual distributed representations without word alignments." ICML, 2015.
10. [Sun et al. 2017] Zequn Sun, et al. "Cross-lingual entity alignment via joint attribute-preserving embedding." ISWC, 2017.

Thank You