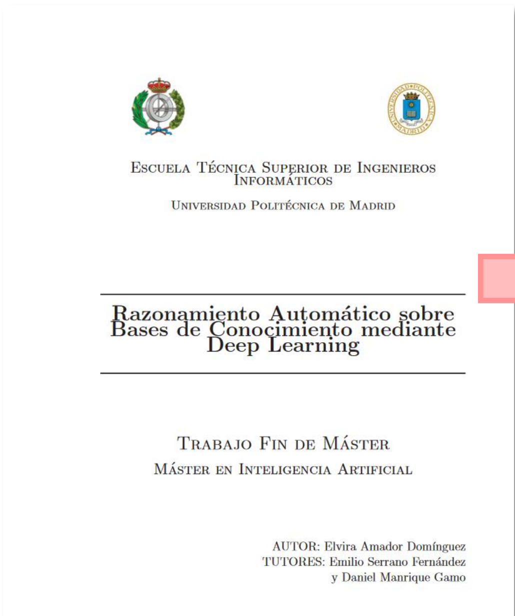
An Ontololgy-Based Deep Learning Approach for Triple Classification with Out-of-Knowledge-Base Entities

Elvira Amador-Domínguez, Ontology Engineering Group
Universidad Politécnica de Madrid, Spain
Emilio Serrano, Ontology Engineering Group, Universidad Politécnica de Madrid, Spain
Daniel Manrique, Artificial Intelligence Lab, Universidad Politécnica de Madrid, Spain
Patrick Hohenecker, Department of Computer Science, University of Oxford, Oxford
Thomas Lukasiewicz, Department of Computer Science, University of Oxford, Oxford

eamador@fi.upm.es
@elviamador

4th March 2021
Online

- This work comes from way back in time…



Escuela Técnica Superior de Ingenieros Informáticos

Universidad Politécnica de Madrid

Razonamiento Automático sobre Bases de Conocimiento mediante Deep Learning

Trabajo Fin de Máster

Máster en Inteligencia Artificial

AUTOR: Elvira Amador Domínguez
TUTORES: Emilio Serrano Fernández
y Daniel Manrique Gamo

Master Thesis
(2018)

An Ontology-Based Deep Learning Approach for Knowledge Graph Completion with Fresh Entities

Elvira Amador-Domínguez[1], Patrick Hohenecker[2], Thomas Lukasiewicz[2], Daniel Manrique[1], and Emilio Serrano[1(✉)]

[1] Department of Artificial Intelligence, Universidad Politécnica de Madrid, Madrid, Spain
{eamador,dmanrique,emilioserra}@fi.upm.es
[2] Department of Computer Science, University of Oxford, Oxford, UK
{patrick.hohenecker,thomas.lukasiewicz}@cs.ox.ac.uk

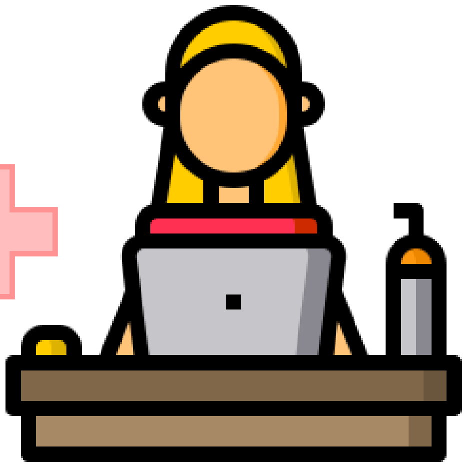
Abstract. This paper introduces a new initialization method for knowledge graph (KG) embedding that can leverage ontological information in knowledge graph completion problems, such as link classification and link prediction. Although the initialization method is general and applicable to different KG embedding approaches in the literature, such as TransE or RESCAL, this paper experiments with deep learning and specifically with the neural tensor network (NTN) model. The experimental results show that the proposed method can improve link classification for a given relation by up to 15%. In a second contribution, the proposed method allows for addressing a problem not studied in the literature and introduced here as "KG completion with fresh entities". This is the use of KG embeddings for KG completion when one or several of the entities in a triple (head, relation, tail) has not been observed in the training phase.

Keywords: Statistical relational learning · Ontological knowledge base · Knowledge Graph Embedding · Latent feature model

1 Introduction

Knowledge representation has always been one of the main challenges of Artificial Intelligence. Throughout time, several approaches have been proposed to model knowledge in a structured and comprehensive way. *Ontologies* [7] are an important approach to knowledge representation. Ontologies present a formal

DCAI Paper
(2019)



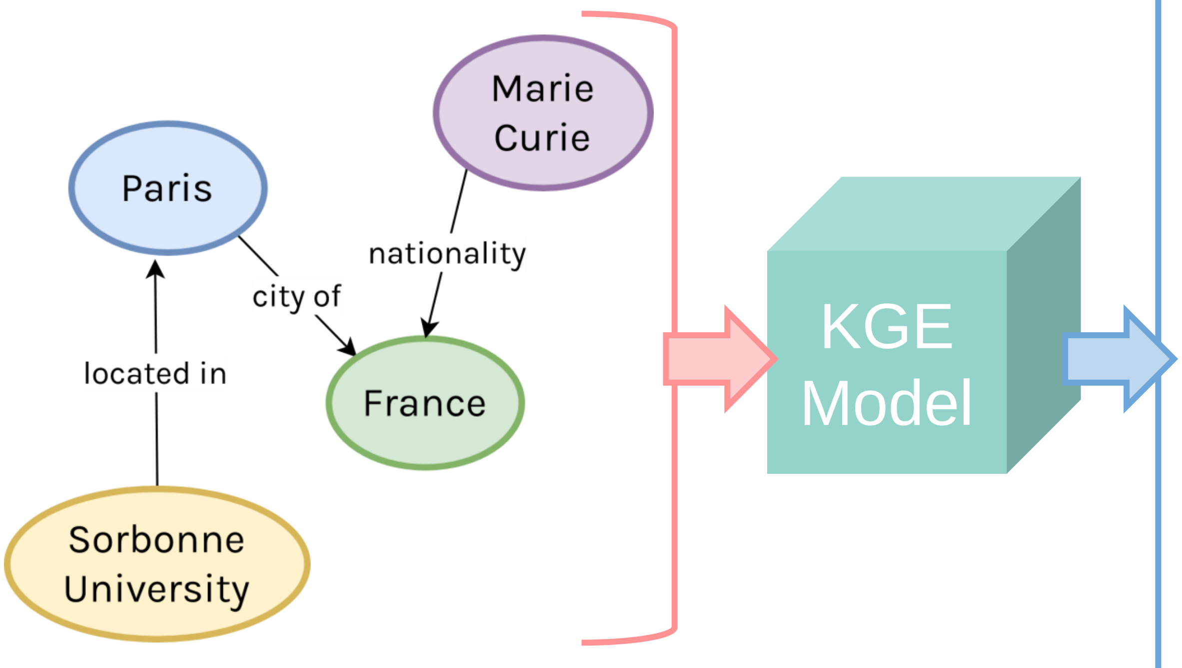
…a bit of extra work during this last year

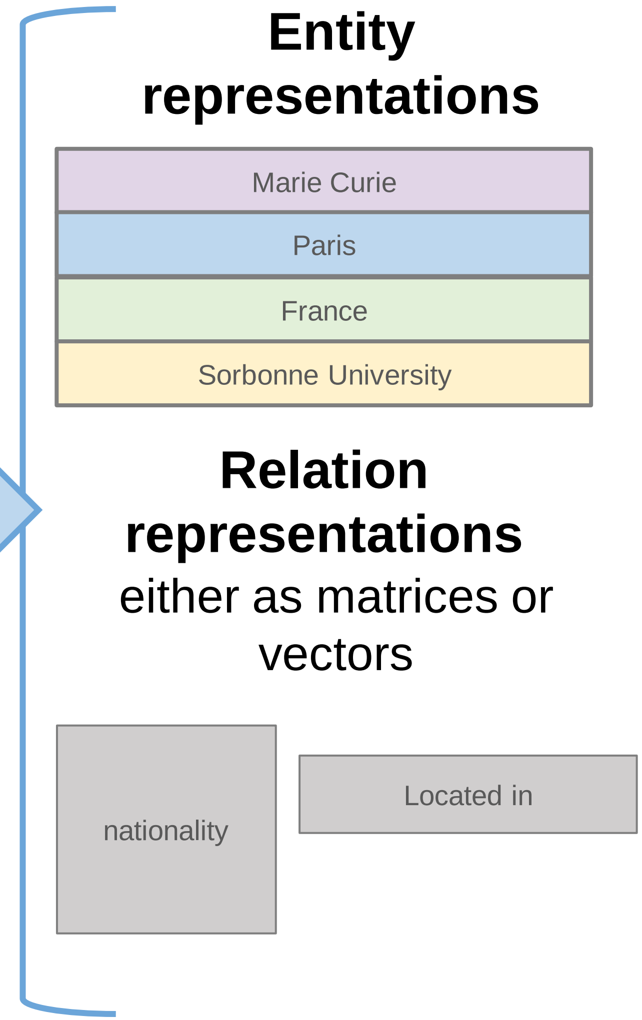
A bit of context

Knowledge graph embedding models

Proposals to enhance knowledge graph embedding models

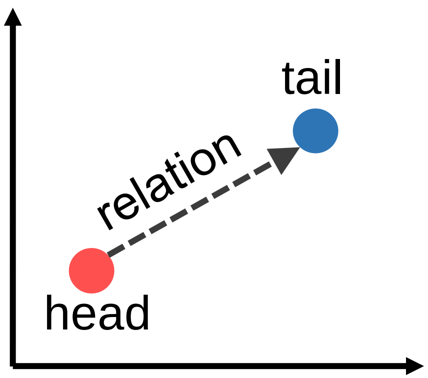
- What are Knowledge Graph Embedding (KGE) models?





Aside from arithmetically representing entities and relations, the model must also be capable of correctly computing whether a fact holds or not

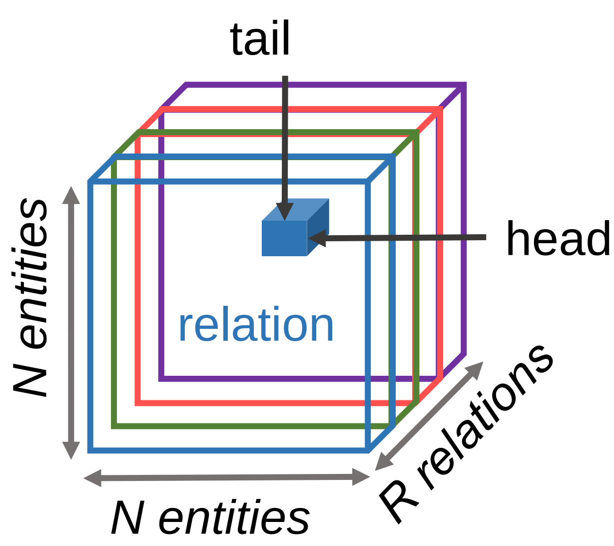
- Translation-based



$$f(h,t) = \|h + r - t\|_{1/2}$$

TransE
TransR
TransH
CrossE
HAKE…

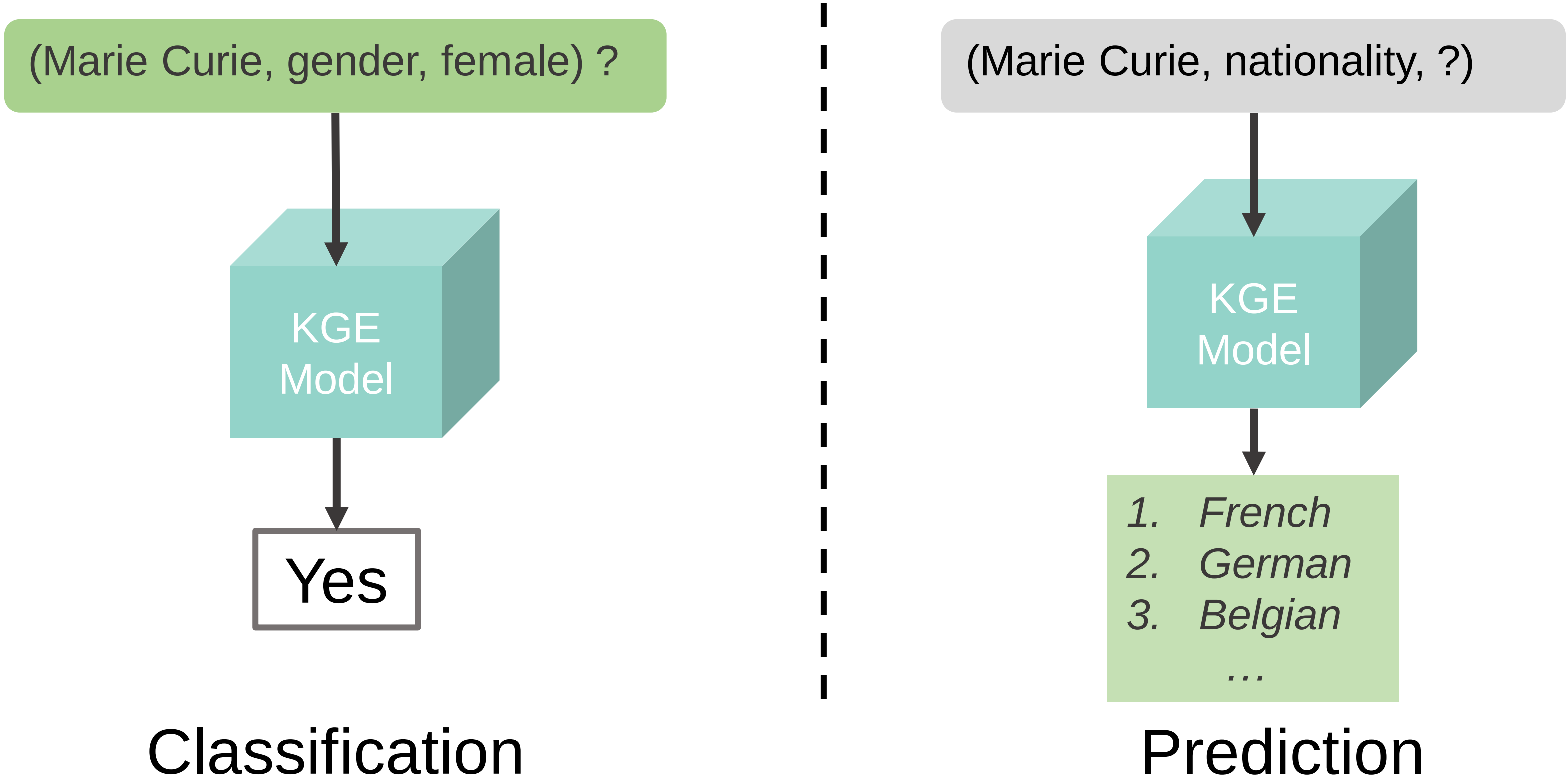
- Semantic matching



$$f(h,t) = h^T M_R t$$

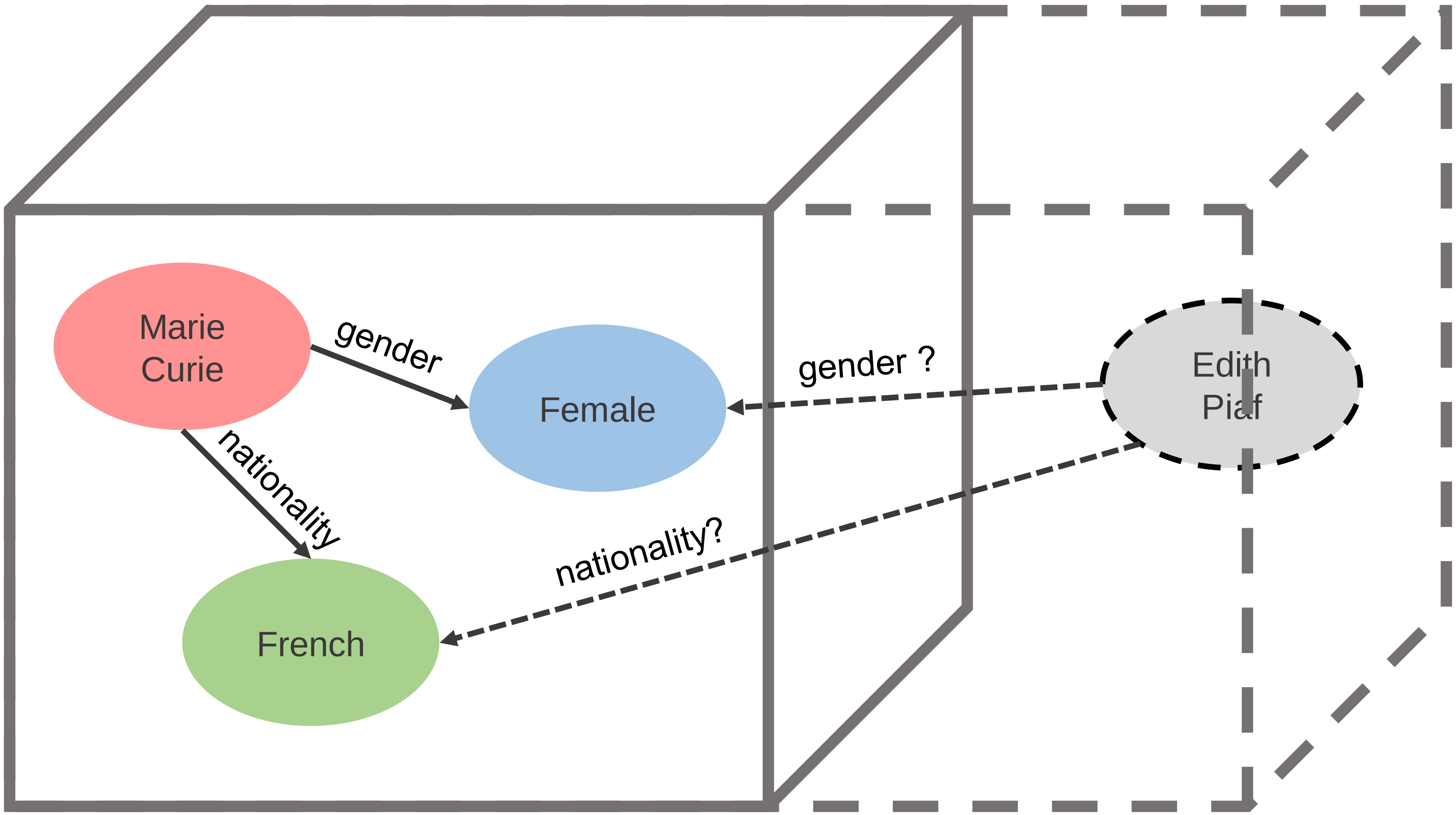
RESCAL
ComplEx
DistMult
HolE
ANALOGY
SimplE…

- Triple classification:
  - *Is this fact feasible or not?*
- Triple prediction:
  - *What is the relation that best joins these two entities?*
  - *What is the missing part of this fact?*



Virtually, no reasoning about this new entity could be performed due to:

1) Lack of existing facts in the graph
2) There is no representation attached to this entity



- There are a few existing proposals that focus on overcoming this issue…

puTransE

DKGE

Even though these approaches remove the need to retrain the model from scratch when a new entity is introduced….

⇨ Still require at least a partial retraining

⇨ Some are based on computationally expensive paradigms such as graph neural networks

⇨ Are designed for a particular model

- We asked ourselves…

1. **Where** can we act in KGE models such that OOKB entities can be introduced without changing the model design?

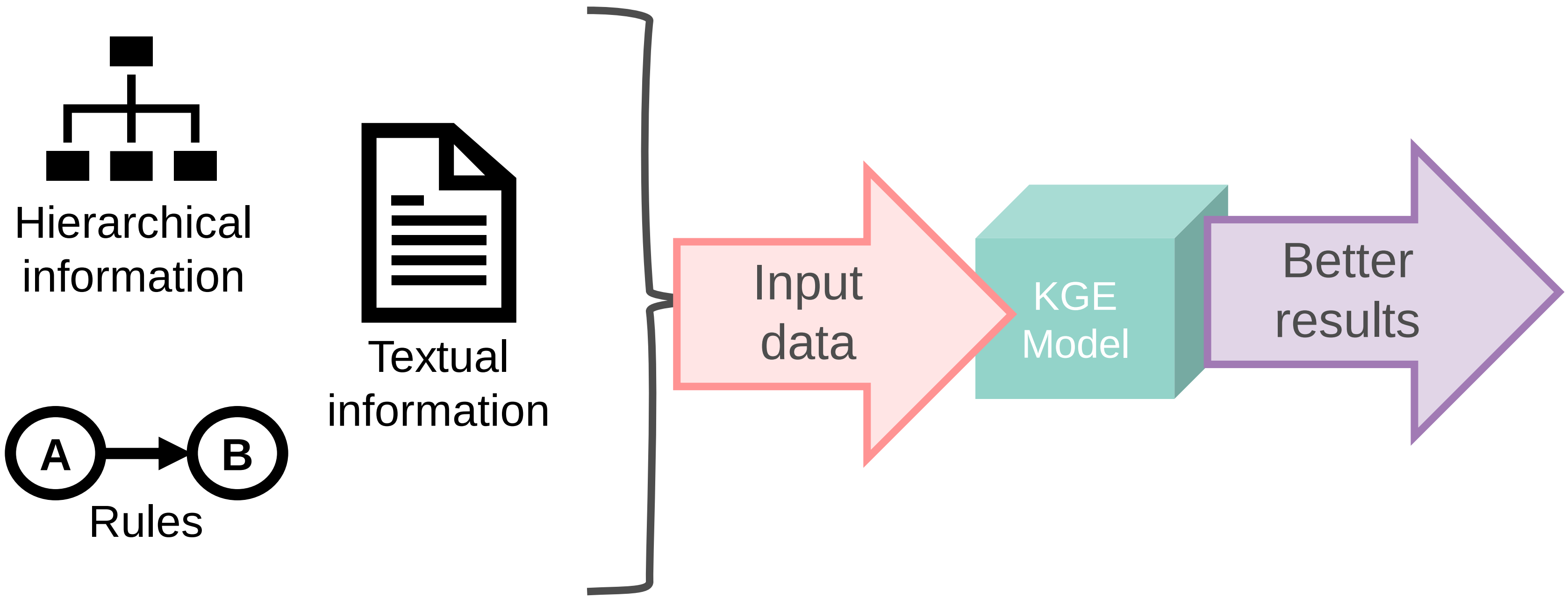
2. **What** are the common elements of both KGs and KGE models?

3. **Are there** any other domains where this issue has been already solved?

Question 1

- **Where** can we act in KGE models such that OOKB entities can be introduced without changing the model design?

⇨ Several proposals that aim to improve KG focus on the **input**



- **What** are the common elements of both KG and KGE models?

KGE Models

⇨ Entities are initialized with random values

⇨ The model 'learns' internal constraints based on the interactions between entities and relations

KGs

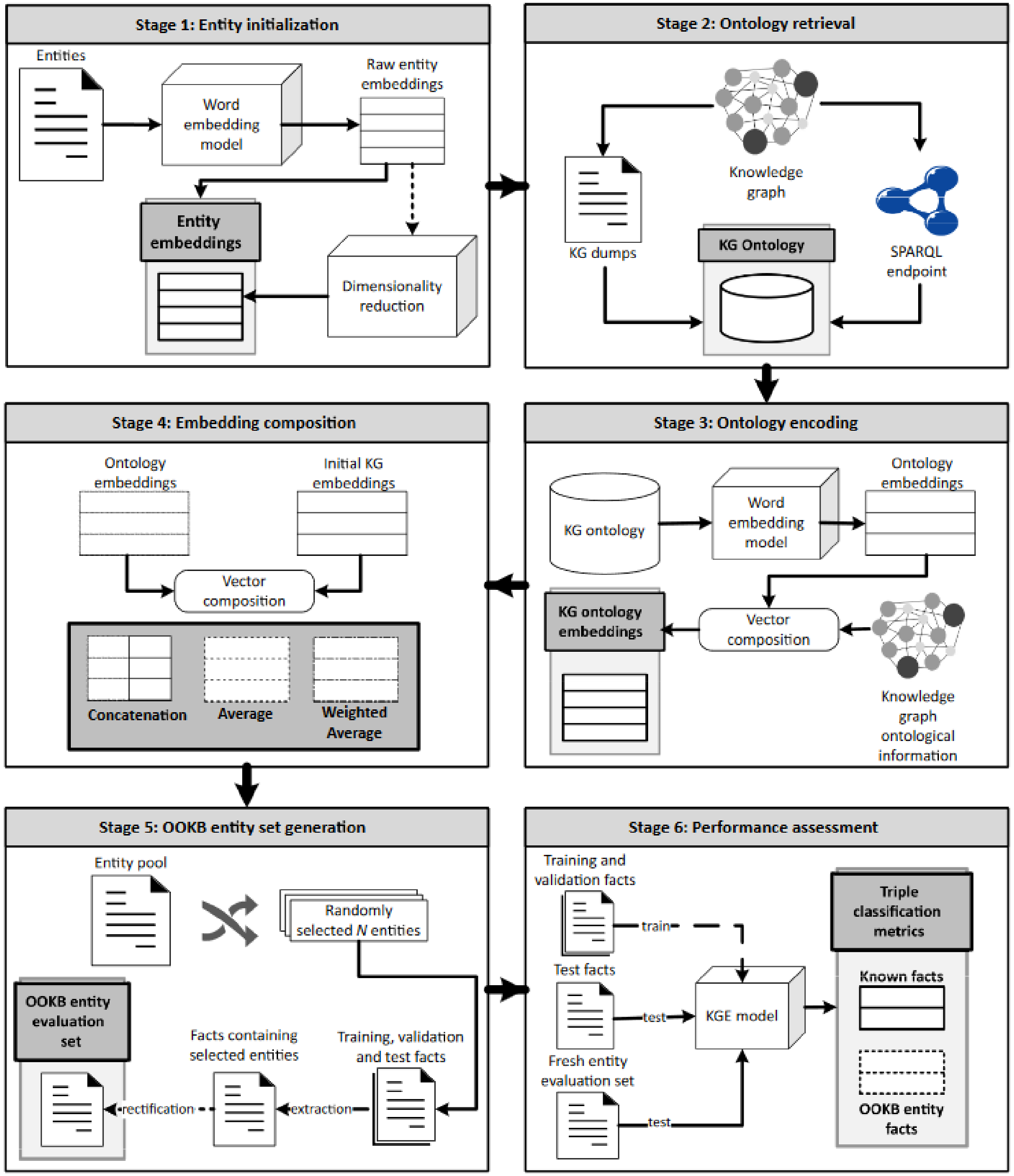
⇨ Entities in the graph are introduced under a common ontology

- **Are there** any other domains where this issue has already been solved?

  ⇨ Replace ‘entity’ with ‘word’ and you have your answer!

  ⇨ With the exception of Word2Vec, most word embedding and language representation models are capable of providing meaningful representations for unseen words

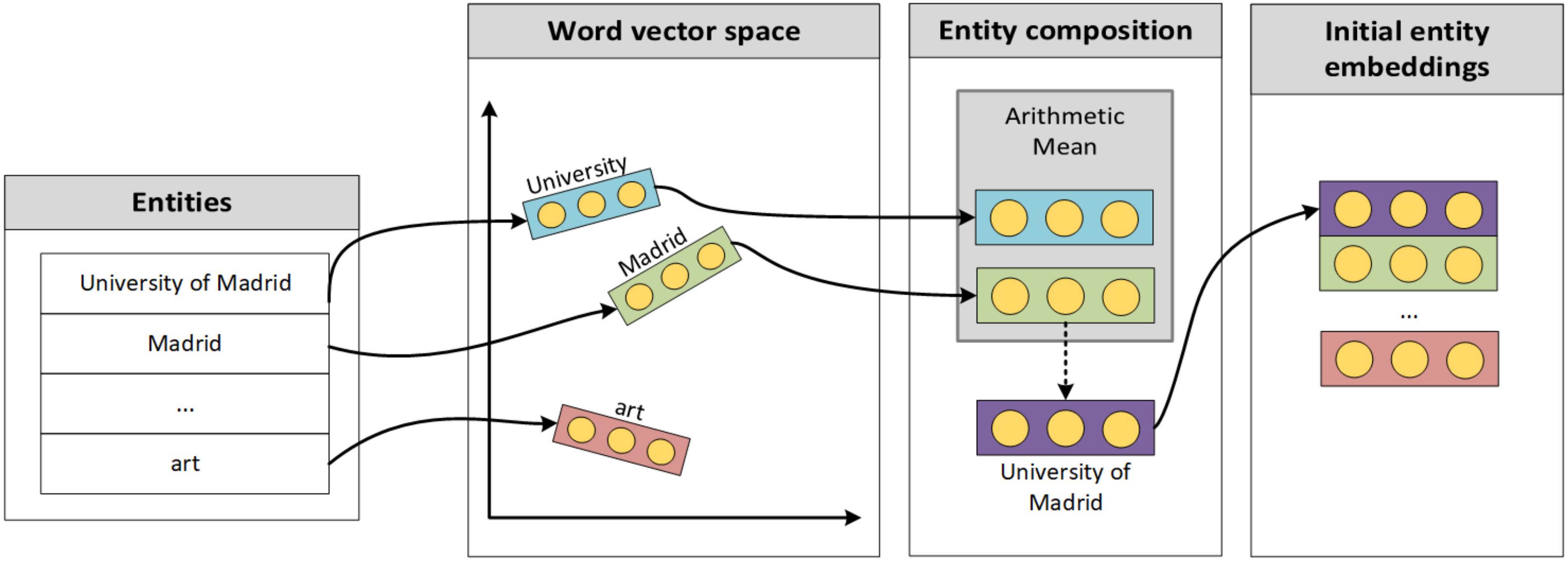
- Combine the forces of ontologies and word embeddings to create an initialization model that enables generalization and superficial reasoning over OOKB entities!



- Several word embedding approaches could be considered
  - One-hot vectors, Word2Vec, FastText, ElmO…
- Dimensionality reduction may be necessary:

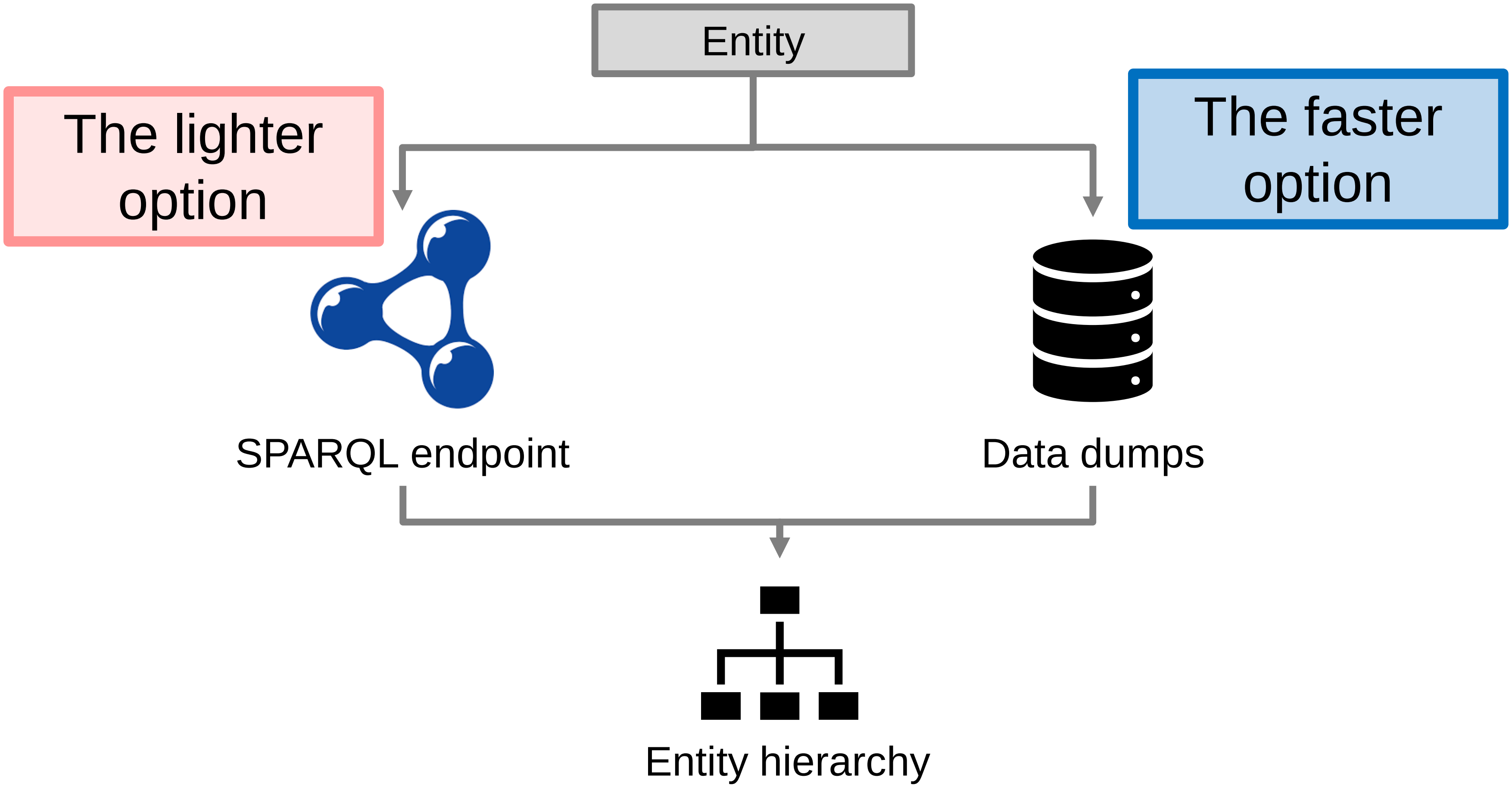
Dim 1024 > Dim 200

- Entities can be composed by one or more tokens



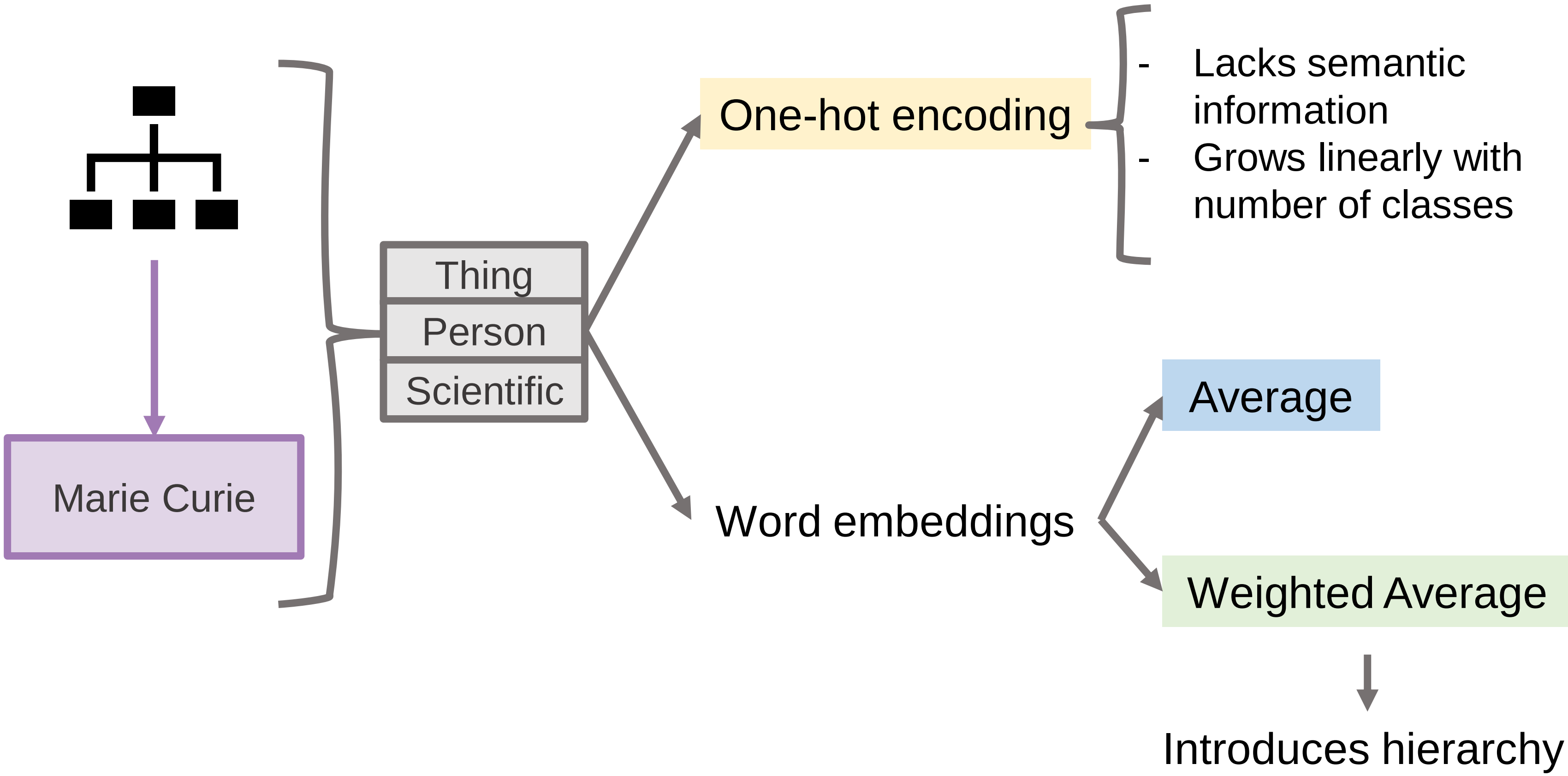
Step 2: Ontological data retrieval

- We have two choices



Step 3: Ontology encoding

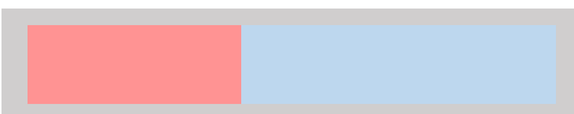
- The encoding of the ontology must be aligned with the encoding of the target
  - If entities are encoded as vectors, their ontological information **must** be a vector



- Things to consider…
  1. Dimensionality restrictions
  2. Trade-off between general and specific information

d_T = KGE model | d_O= ontological information | d_M=entity embedding

Concatenation

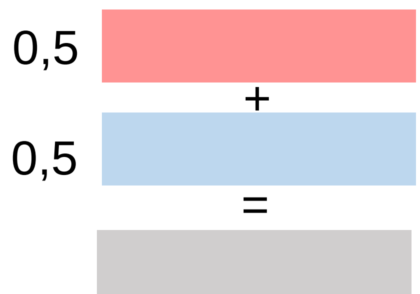


$$d_T = d_O + d_M$$

Dimensionality restriction ✗

Information trade-off ✓

Average

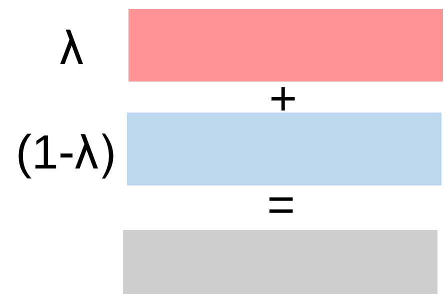


$$d_T = d_O = d_M$$

Dimensionality restriction ✓

Information trade-off ✗

Weighted Average



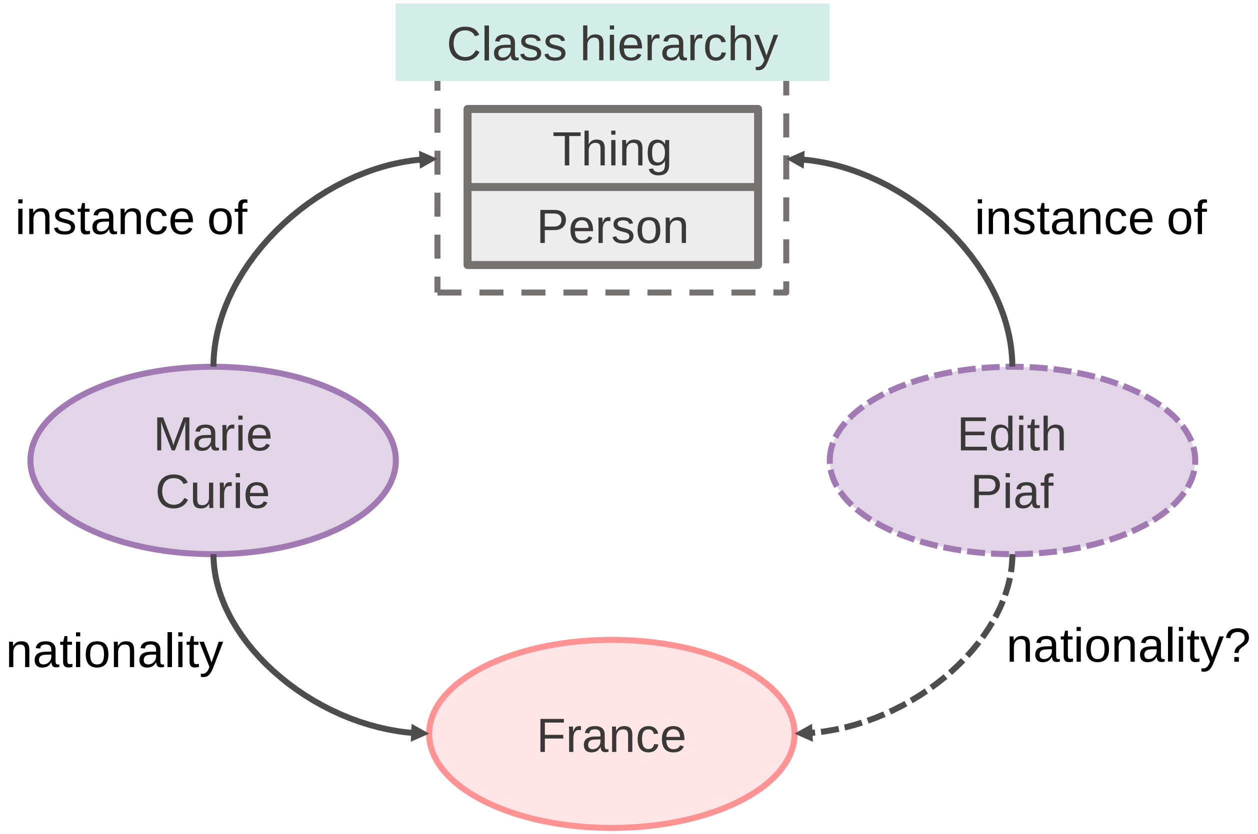
$$d_T = d_O = d_M$$

Dimensionality restriction ✓

Information trade-off ✓

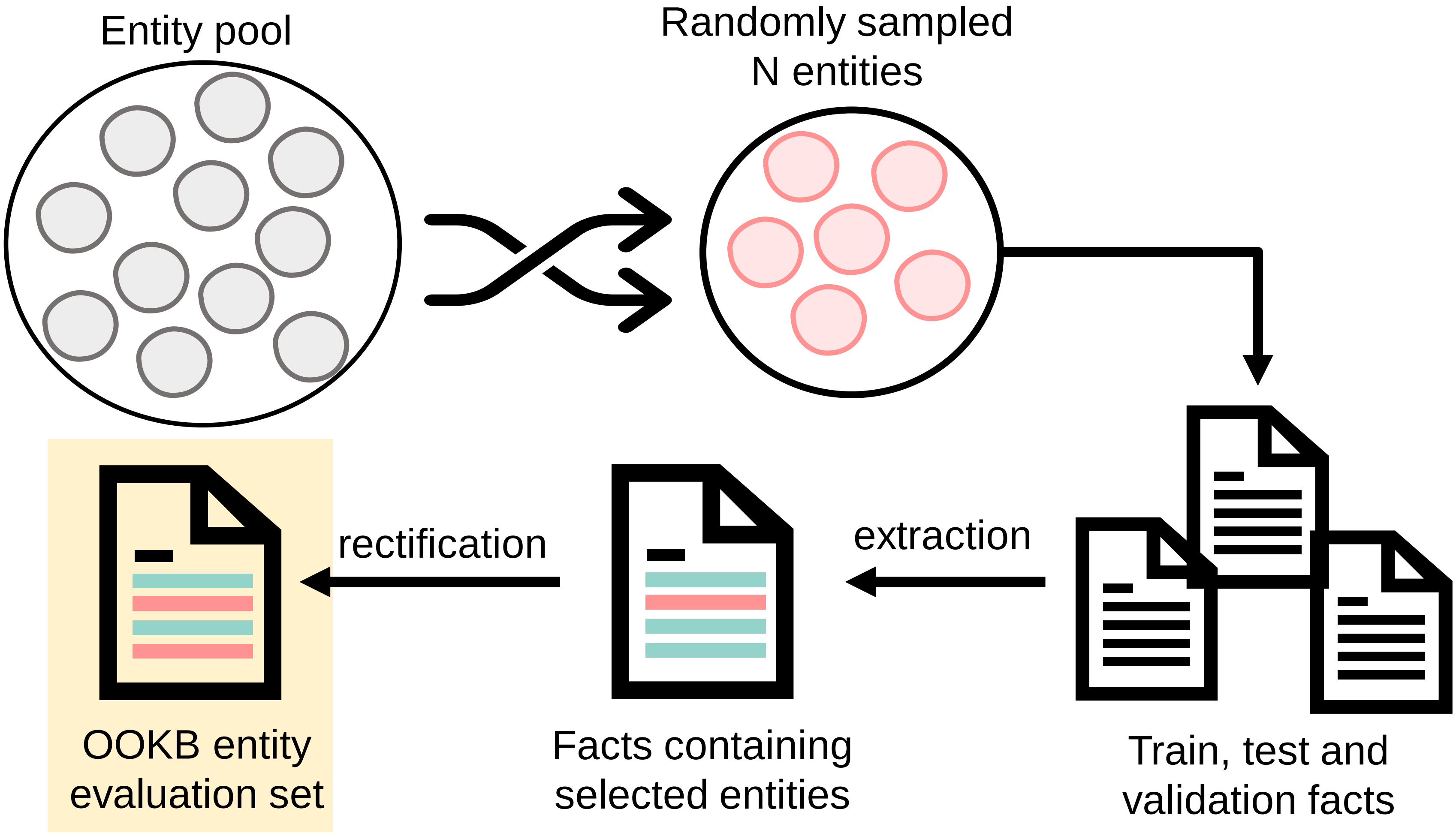
Step 5: OOKB entity set generation

- Due to the proposed initialization, unseen entities can be encoded in a superficial but reliable manner.
- OOKB entities must be initialized following the same procedure as existing entities



Step 5: OOKB entity set generation

- Out of the existing train, validation and test partitions we need to extract an additional set containing OOKB entities



- Once the model has been trained and validated with the fully-known set, we evaluate the model over the OOKB and known entity sets…

$Acc_T = accuracy\ known\ set$ $\quad$ $Acc_O = accuracy\ OOKB\ set$

$Acc_T \gg Acc_O$ ⇨ Ontological information is not been appropriately introduced

$Acc_T \approx Acc_O$ ⇨ The model is correctly introducing the ontological information and is capable of generalizing

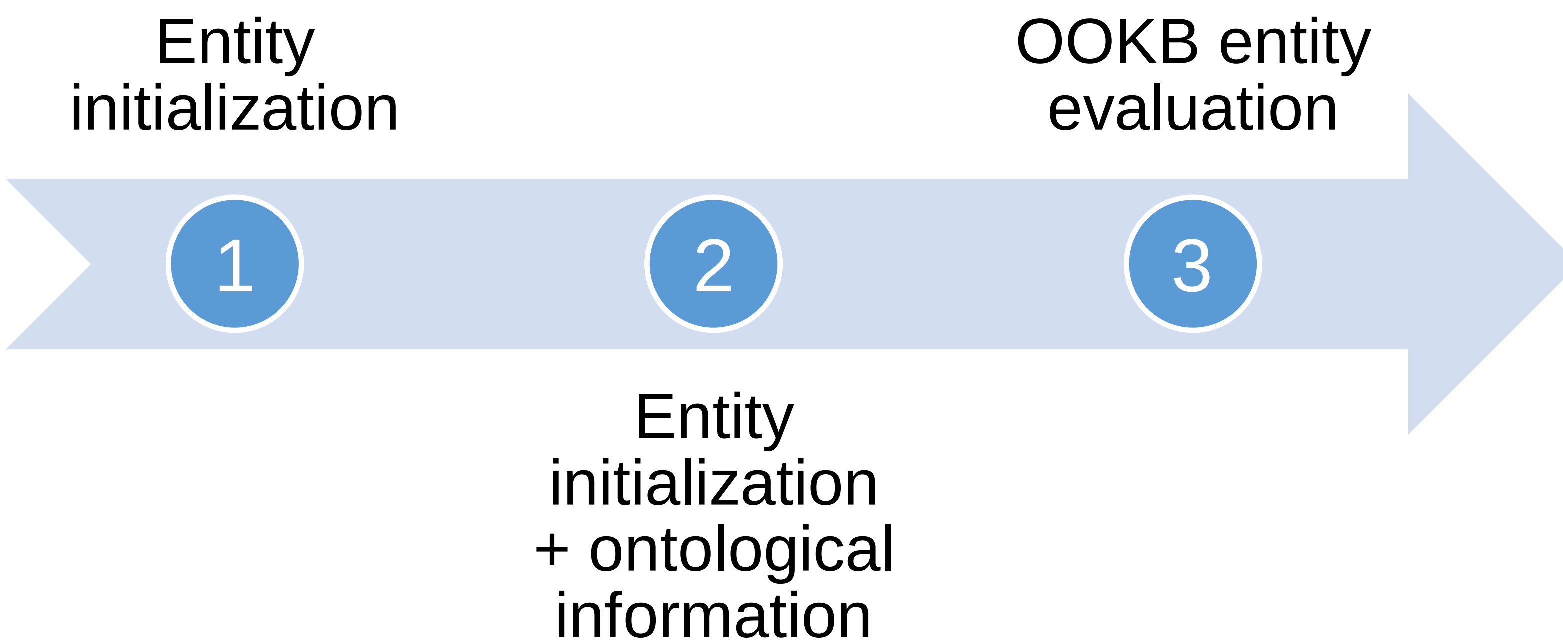
$Acc_T \ll Acc_O$ ⇨ Non-optimal parameter selection. The model is diverging from the optimal solution

- Models to evaluate: SimplE, ComplEx, DistMult, ANALOGY and TransE

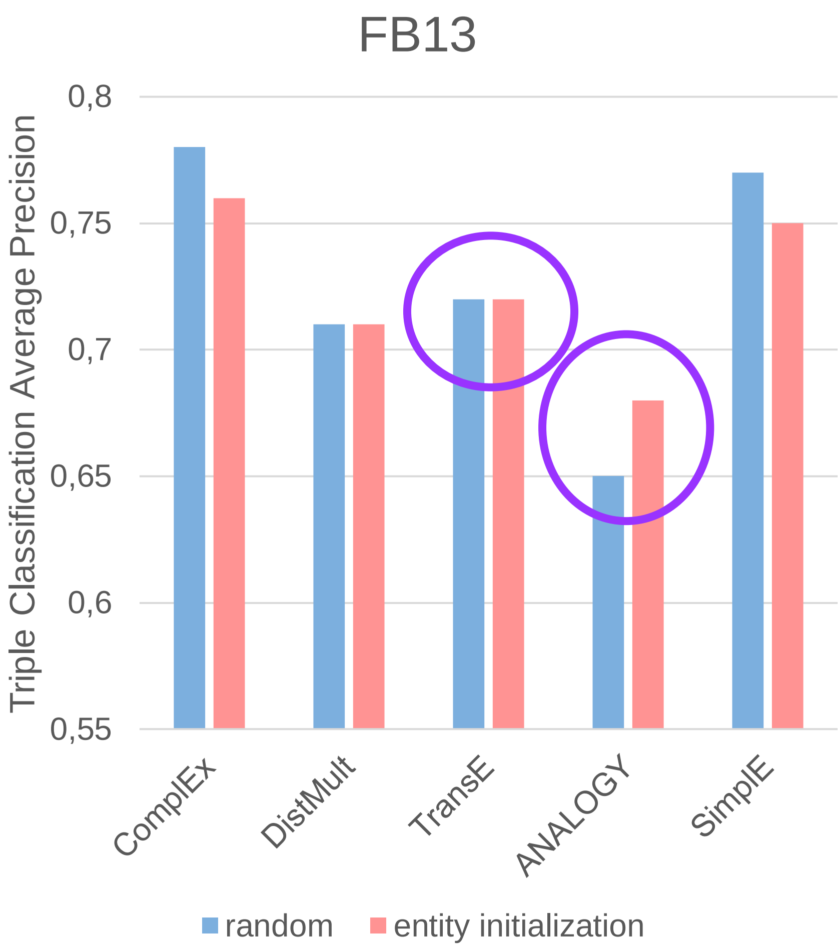
Model name	Model type	Embedding Dimensions	Scoring Function	Complexity Order
TransE [4]	Translation-based	$h, t \in \mathbb{R}^d$ $r \in \mathbb{R}^d$	$-\|h + r - t\|_{\frac{1}{2}}$	$\mathcal{O}(d)$
DistMult [37]	Semantic Matching	$h, t \in \mathbb{R}^d$ $r \in \mathbb{R}^d$	$h^T diag(r) t$	$\mathcal{O}(d)$
ComplEx [33]	Semantic Matching	$h, t \in \mathbb{C}^d$ $r \in \mathbb{C}^d$	$Re(h^T diag(r) \bar{t})$	$\mathcal{O}(d)$
ANALOGY [14]	Semantic Matching	$h, t \in \mathbb{R}^d$ $r \in \mathbb{R}^d$	$h^T M_r t$	$\mathcal{O}(d)$
SimplE [11]	Semantic Matching	$h, t \in \mathbb{R}^d$ $r, r^{-1} \in \mathbb{R}^d$	$\frac{1}{2}[(hrt^T) + (hr^{-1}t)]$	$\mathcal{O}(d)$

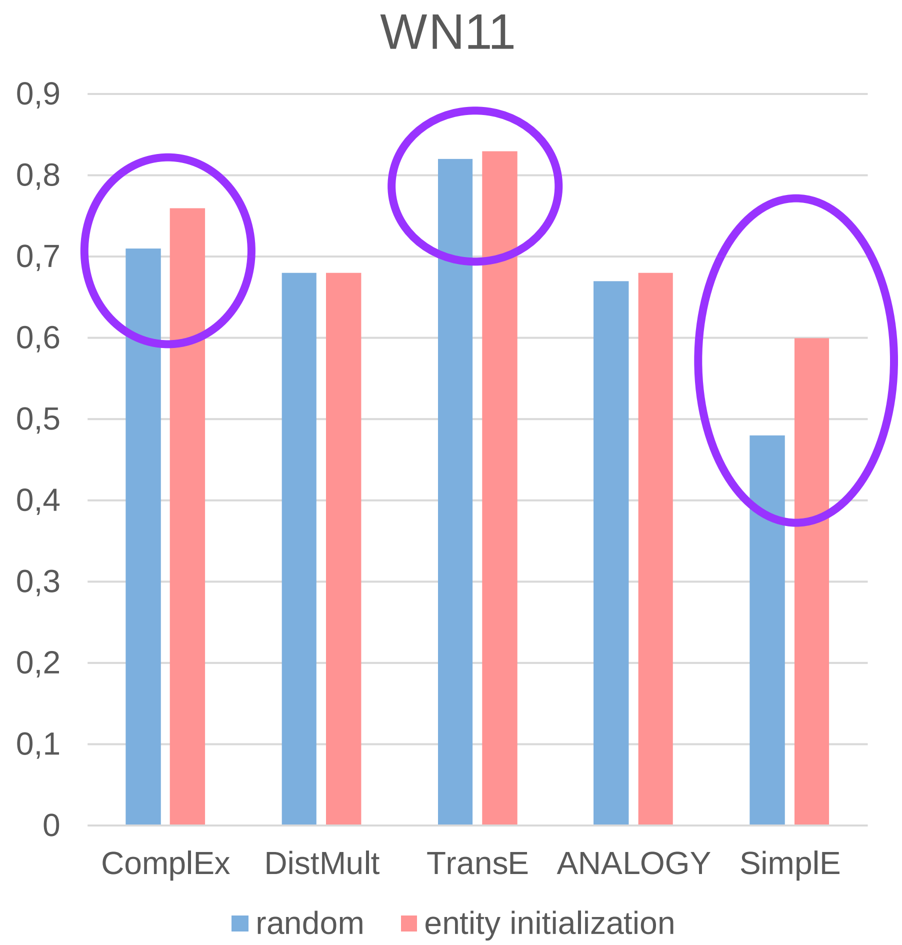
- Datasets: FB13, WN11
- Only triple classification results are reported

- Three phase experimentation

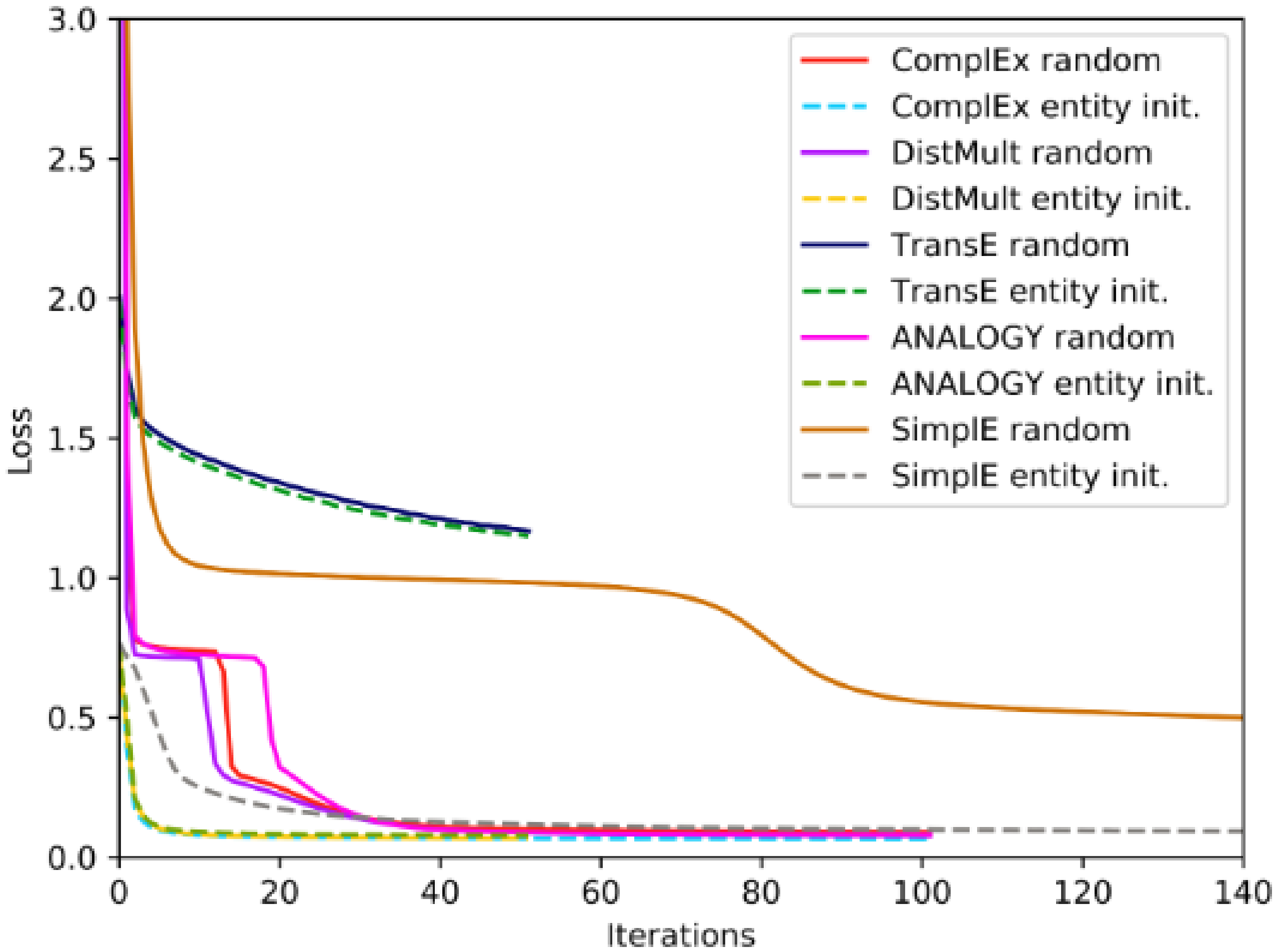


- Entities are initialized using a pretrained Word2Vec model





- Though the impact in the results may not be outstanding, entity initialization induces another interesting benefit…

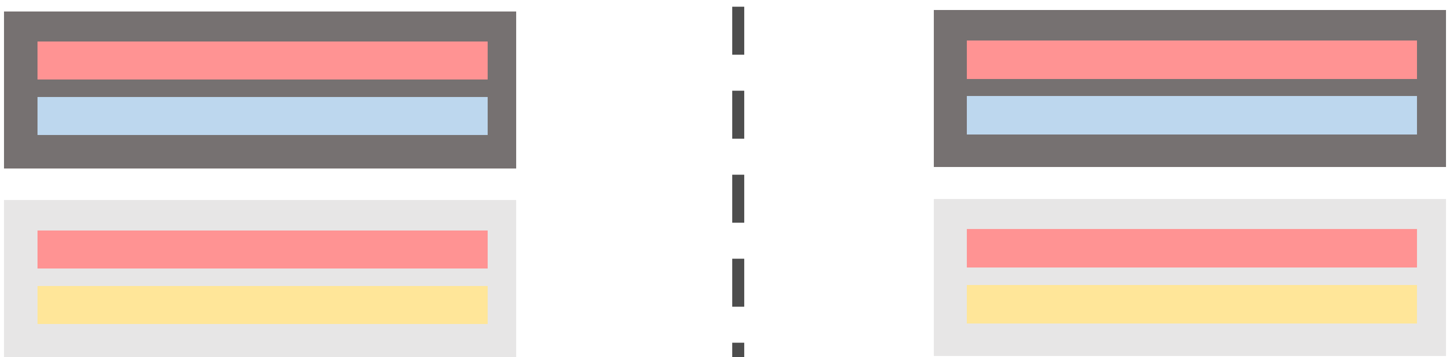


Phase 2: Ontology introduction

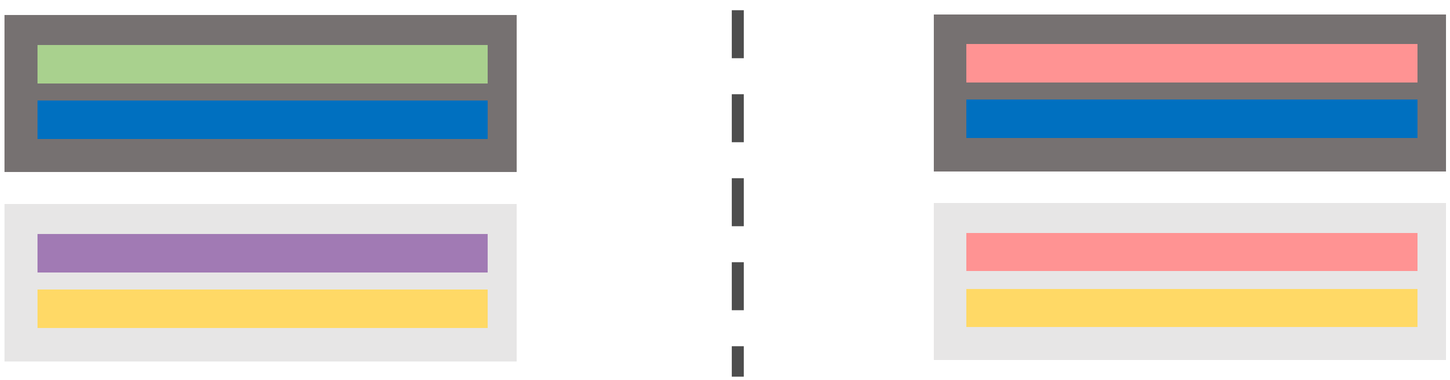
- For WN11, the homonym ontology is applied
- For FB13, two ontologies are considered: its corresponding ontology (Google KG’s) ontology and the DBpedia Ontology
- Ontology classes are encoded using the same Word2Vec model as the entities
- Ontology vectors are generated using weighted average
- **Ontological-information vectors remain static throughout training**

Non-static | Static

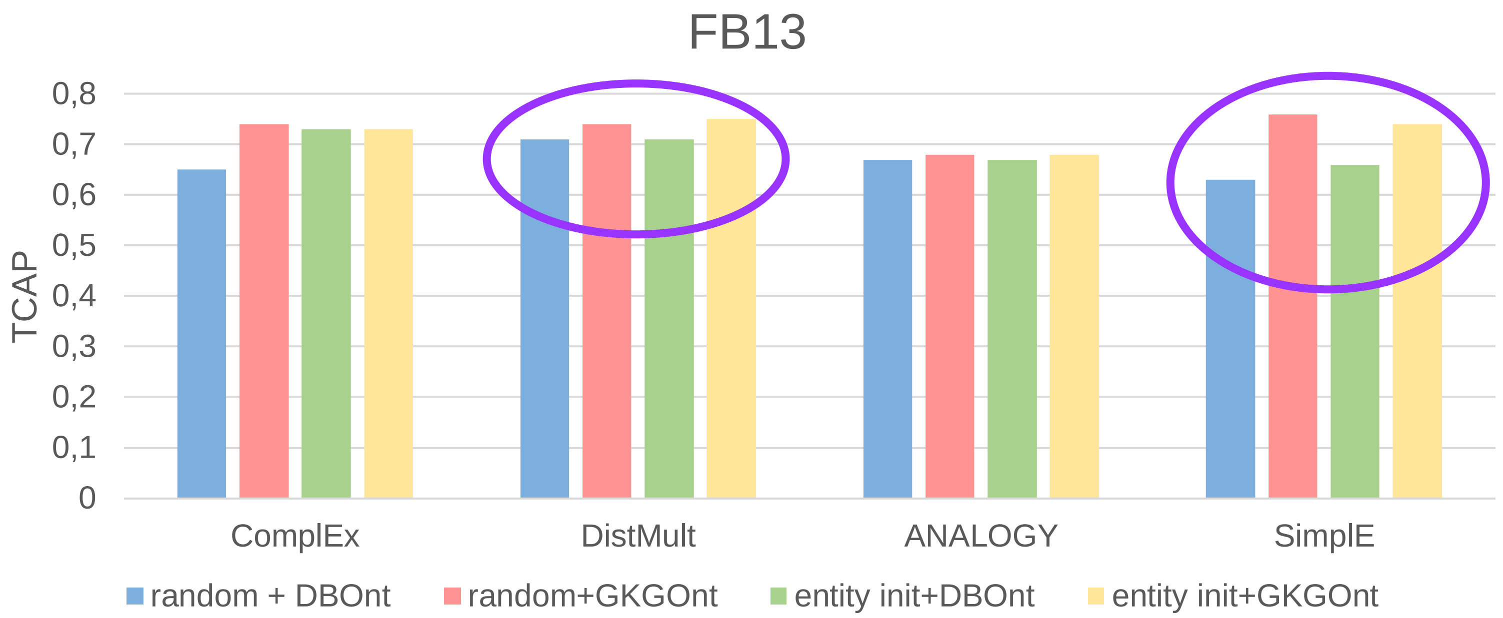
- Iteration 1

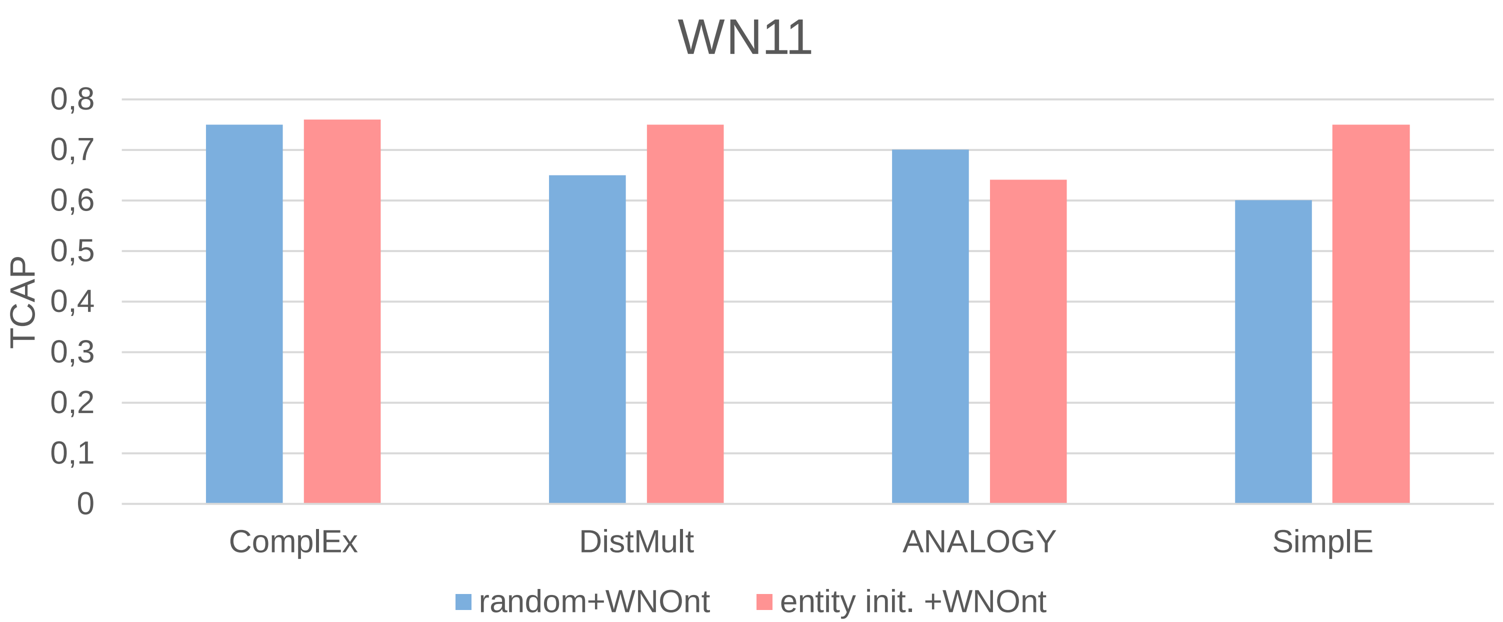


- Iteration 2

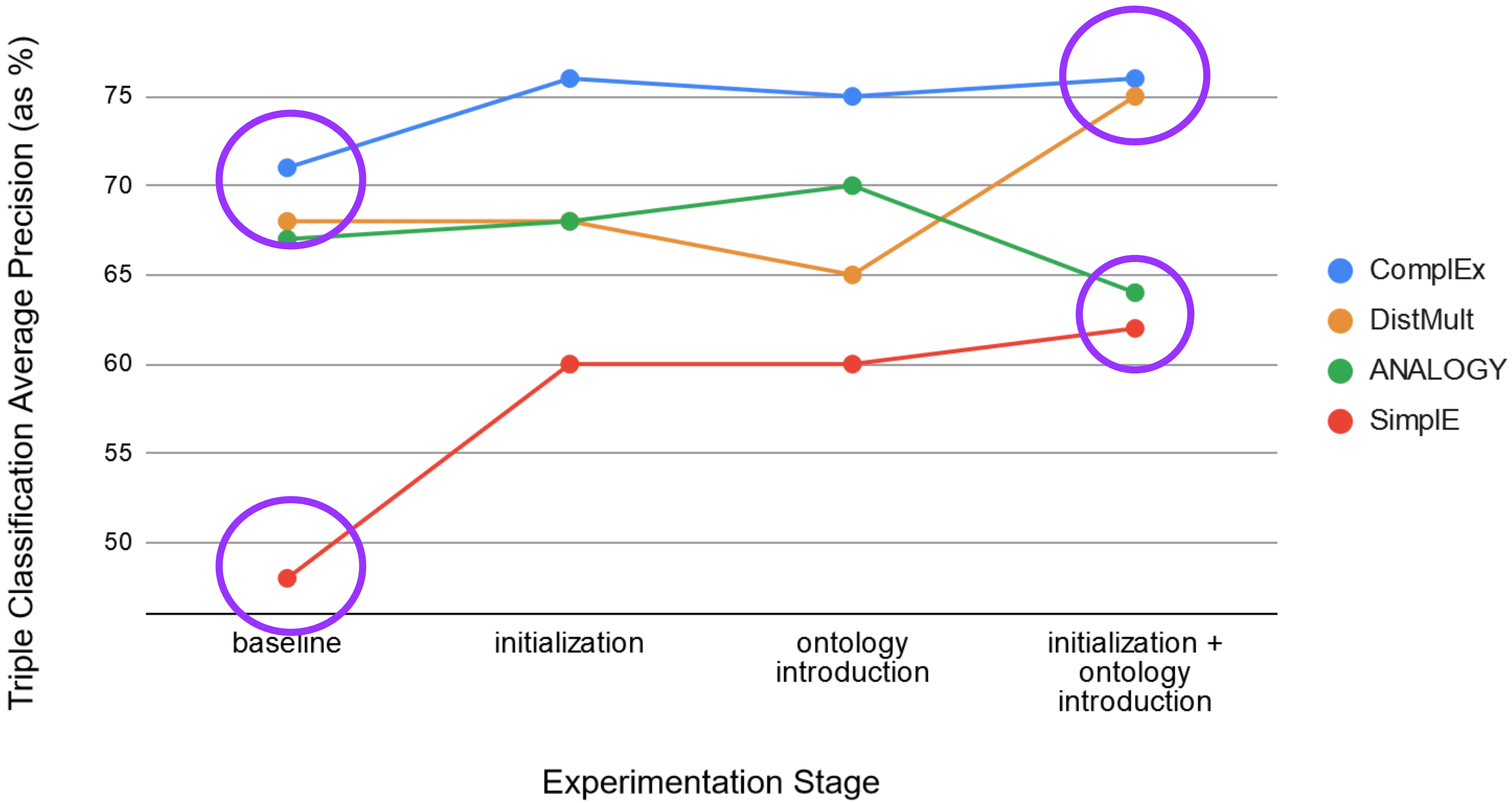


Phase 2: Ontology Introduction



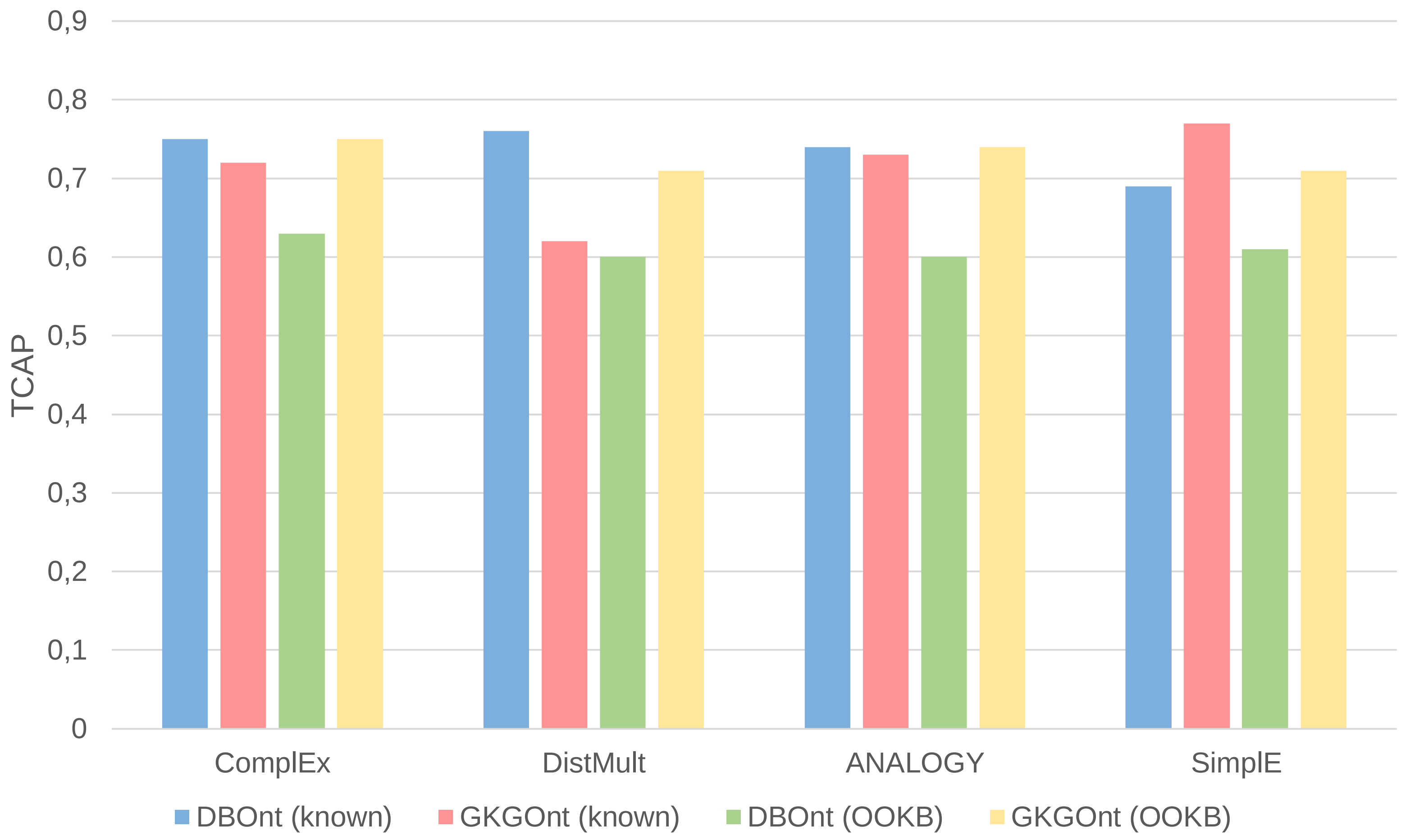


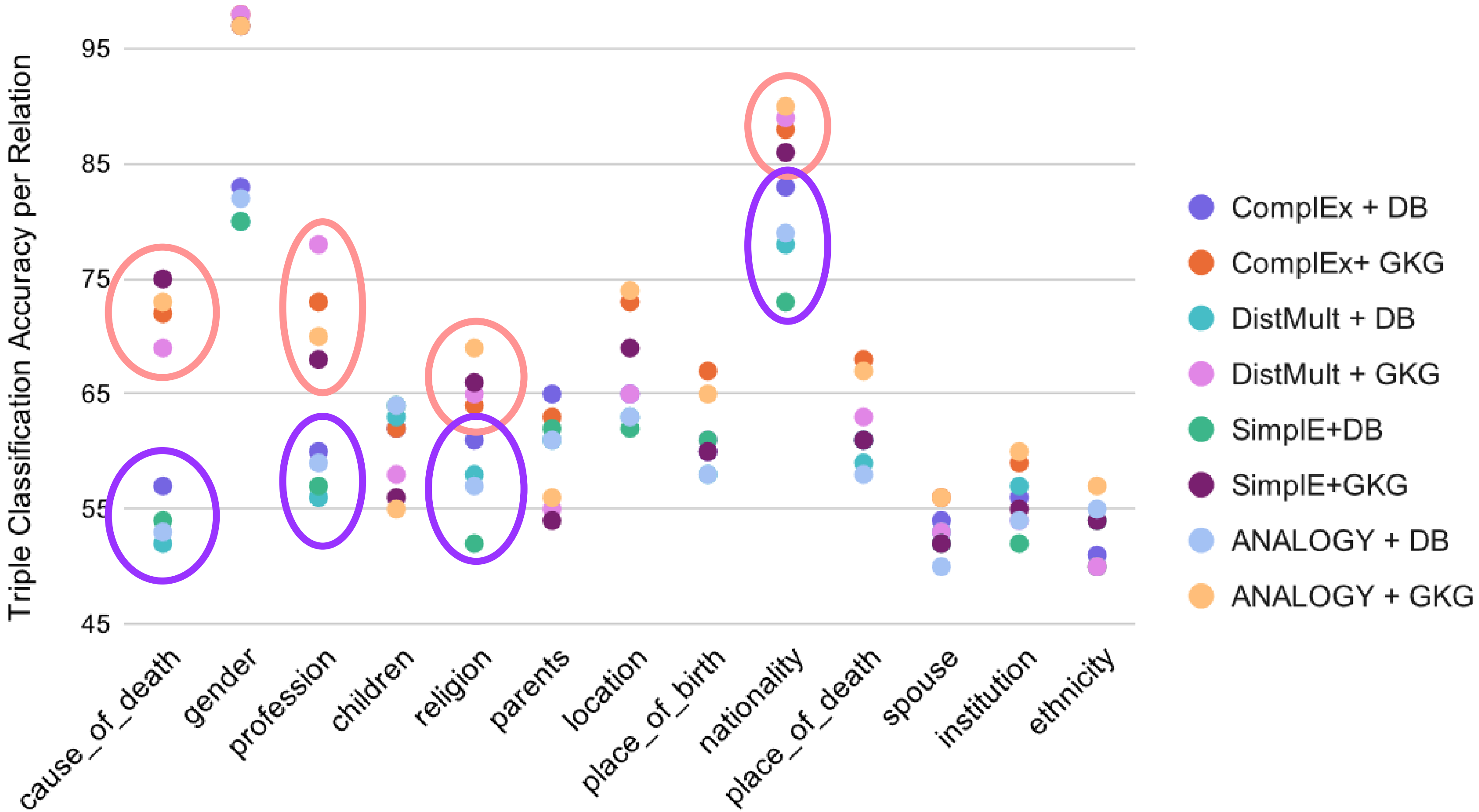
- A better view of the WN11 case



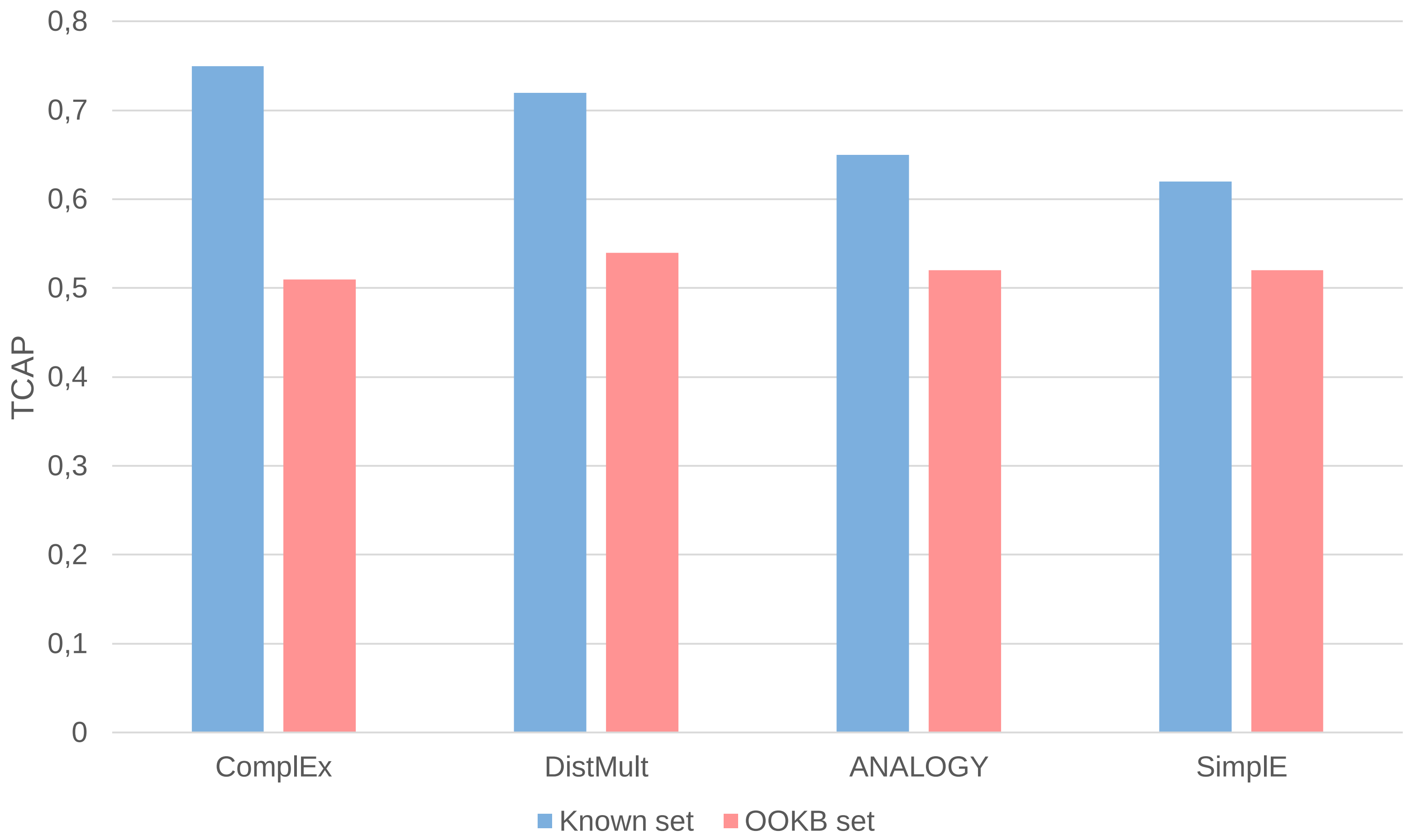
- 1500 entities were randomly selected for each dataset
- Negative samples were generated via random replacement under the closed-world assumption
- For this stage, both entity initialization and ontology information are required

- FB13

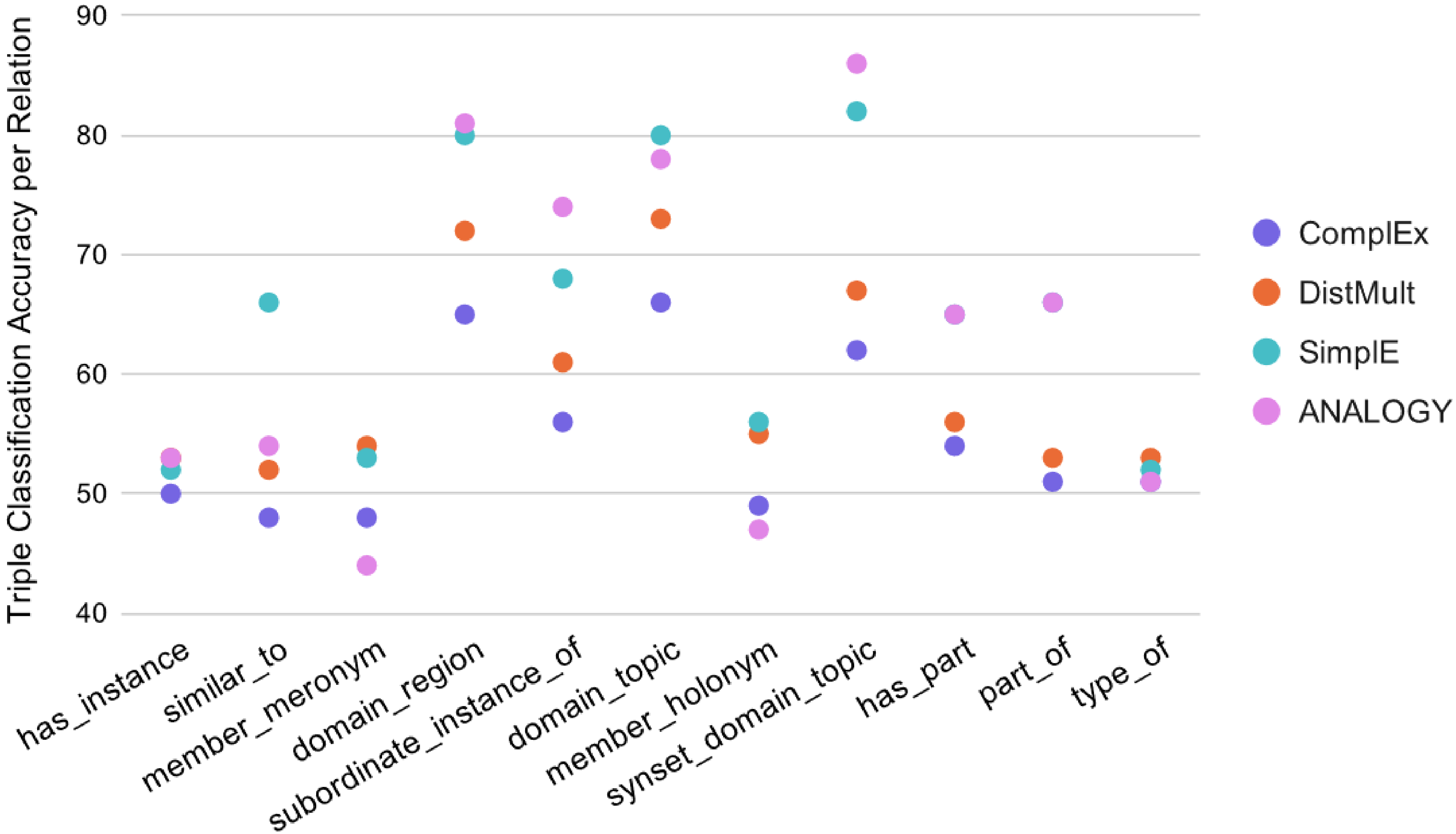

0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
TCAP
ComplEx
DistMult
ANALOGY
SimplE
DBOnt (known)
GKGOnt (known)
DBOnt (OOKB)
GKGOnt (OOKB)


Triple Classification Accuracy per Relation
95
85
75
65
55
45
cause_of_death
gender
profession
children
religion
parents
location
place_of_birth
nationality
place_of_death
spouse
institution
ethnicity
ComplEx + DB
ComplEx+ GKG
DistMult + DB
DistMult + GKG
SimplE+DB
SimplE+GKG
ANALOGY + DB
ANALOGY + GKG

- WN11


0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
TCAP
ComplEx
DistMult
ANALOGY
SimplE
Known set
OOKB set

Phase 3: OOKB entity evaluation



Conclusions

- The proposed method, despite its simplicity, can help classical KGE models to progress at least in the triple classification task in the standard benchmarks.
- By using a proper initialization, KGE models also converge much faster to their optimal solution
- The proposed initialization enables triple classification over OOKB entities while being computationally inexpensive and easy to implement alongside any KGE model.

- Evaluate this proposal to conduct transfer learning with KG embeddings
- Study the impact of employing different information sources such as rules or textual descriptions and their impact with respect to explainability
- Expand the proposal to upcoming KGE models

This work well be (hopefully soon!) published in the Information Sciences journal, yay! You can read all about this work there, but brace yourselves because it's quite a long read 😉

An Ontololgy-Based Deep Learning Approach for Triple Classification with Out-of-Knowledge-Base Entities

Elvira Amador-Domínguez, Ontology Engineering Group Universidad Politécnica de Madrid, Spain
Emilio Serrano, Ontology Engineering Group, Universidad Politécnica de Madrid, Spain
Daniel Manrique, Artificial Intelligence Lab, Universidad Politécnica de Madrid, Spain
Patrick Hohenecker, Department of Computer Science, University of Oxford, Oxford
Thomas Lukasiewicz, Department of Computer Science, University of Oxford, Oxford

eamador@fi.upm.es
@elviamador

4th March 2021
Online