Universal Representation Learning of Knowledge Bases by Jointly Embedding Instances and Ontological Concepts

Junheng Hao, Muhao Chen, Wenchao Yu, Yizhou Sun, Wei Wang
University of California, Los Angeles

Outline

- **Background: Knowledge Graphs and Embeddings**
- Formulation: Two-view Knowledge Graphs
- JOIE Modeling: Cross-view & Intra-view
- Experimental Results
- Conclusion & Future Work

Knowledge graphs (KGs) Are Everywhere

UCLA ScAi

General-purpose KGs

Baidu 知识图谱

Google KNOWLEDGE GRAPH

DBpedia

WIKIDATA

yago select knowledge

NELL

Freebase

Bio & Medical KGs

MeSH

neXtprot

STRING

GENEONTOLOGY Unifying Biology

PubMed

Product Graphs & E-commerce

product graph

ebay

Alibaba.com

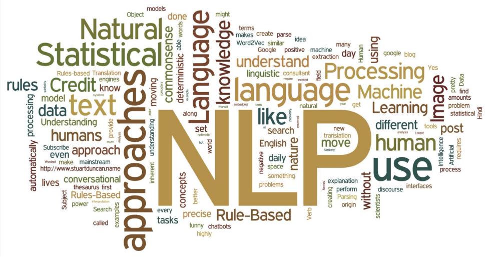
Common-sense KGs & NLP

WORDNET

ConceptNet An open, multilingual knowledge graph

Knowledge Graphs Are Foundational

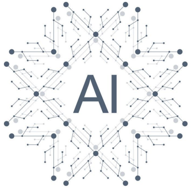
- Foundational to knowledge-driven AI systems
- Enable many downstream applications (NLP tasks, QA systems, etc)



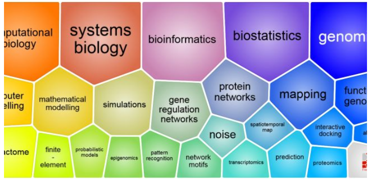
Natural Language Processing

QA & Dialogue systems

Knowledge Graphs

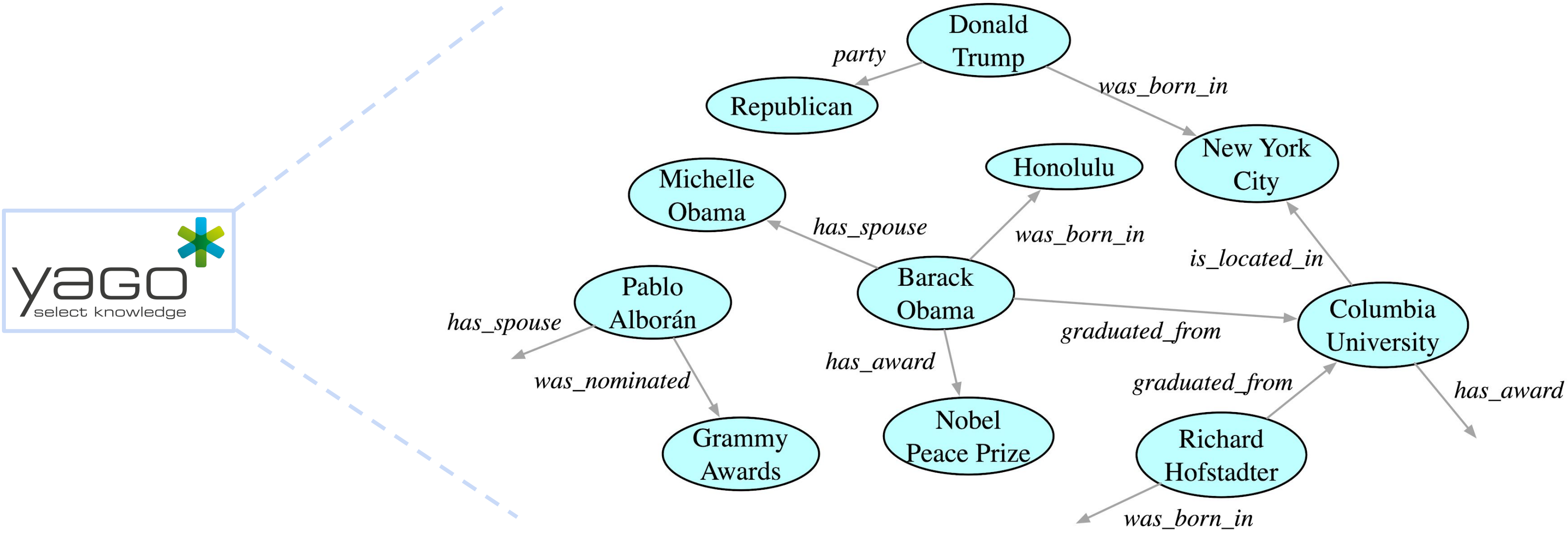


Recommendation Systems



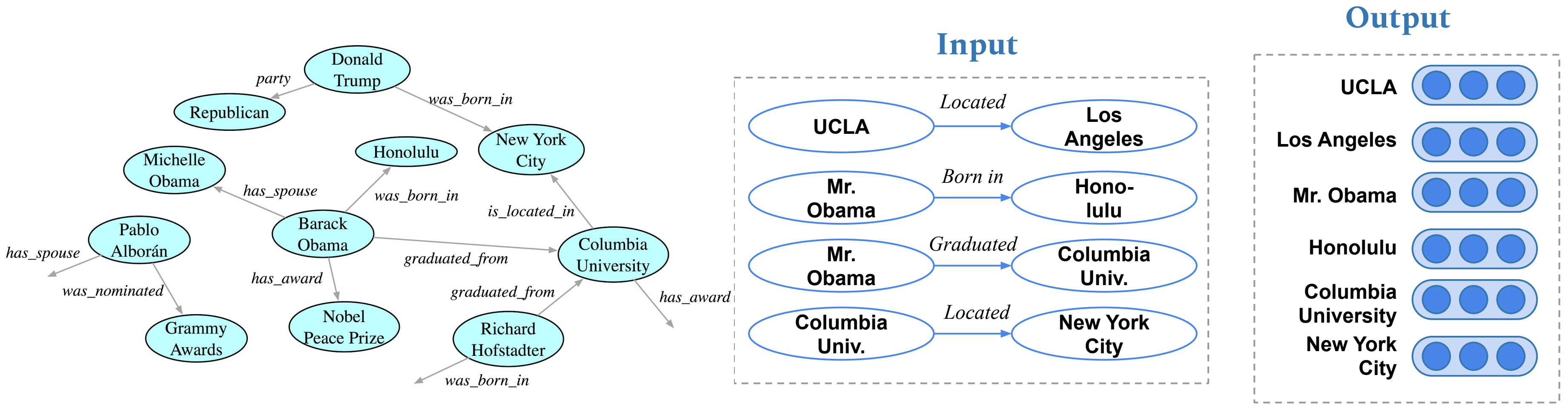
Computational Biology

KG Example From YAGO



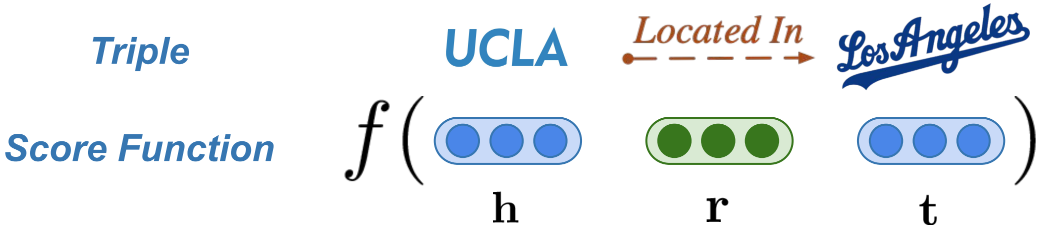
KG Embedding From Triples

- Knowledge graph embeddings represent entities and relations as latent vectors or matrices and support effective relation learning and inference.
- **Input:** Relation facts (triples)
- **Output:** Embedding representations of objects and relations



Learning KG Embeddings

- Key of existing KG embedding methods: Triple score function

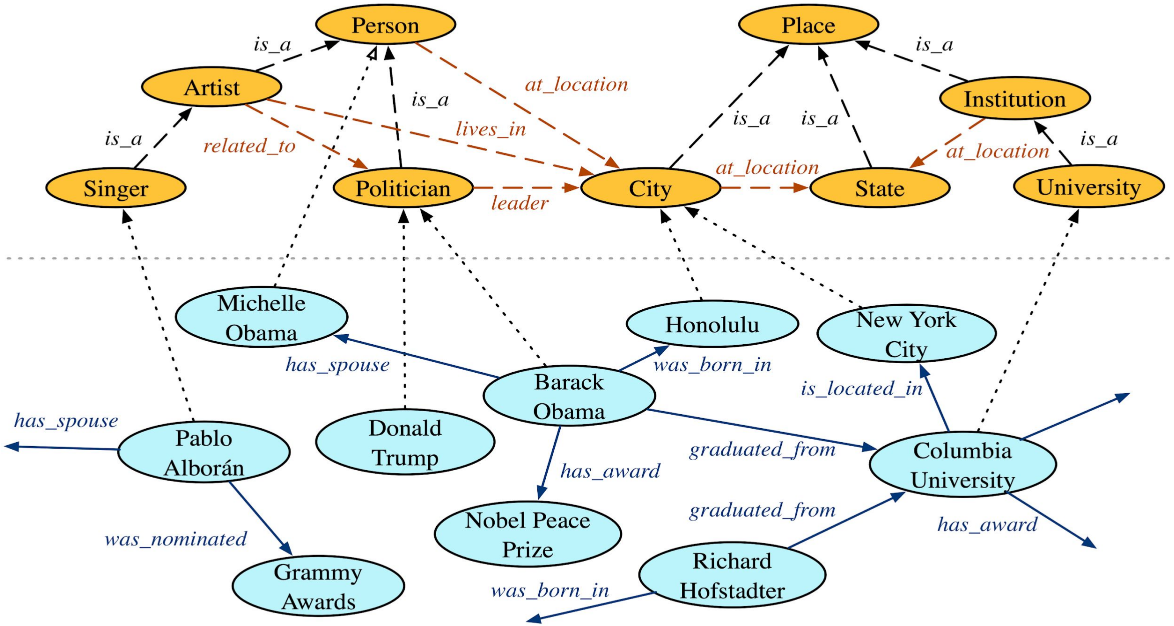


- Previous research employ various arithmetic methods to capture observed relations of entities in a single KG (for example, translational distance or similarity)

Model	Score Function	Embeddings
TransE (Bordes et al., 2013)	$-\|\|\mathbf{h}+\mathbf{r}-\mathbf{t}\|\|$	$\mathbf{h},\mathbf{r},\mathbf{t}\in\mathbb{R}^k$
TransX	$-\|\|g_{r,1}(\mathbf{h})+\mathbf{r}-g_{r,2}(\mathbf{t})\|\|$	$\mathbf{h},\mathbf{r},\mathbf{t}\in\mathbb{R}^k$
DistMult (Yang et al., 2014)	$(\mathbf{h}\circ\mathbf{t})\cdot\mathbf{r}$	$\mathbf{h},\mathbf{r},\mathbf{t}\in\mathbb{R}^k$
HolE (Nickel et al., 2016)	$(\mathbf{h}\star\mathbf{t})\cdot\mathbf{r}$	$\mathbf{h},\mathbf{r},\mathbf{t}\in\mathbb{R}^k$
ComplEx (Trouillon et al., 2016)	$\mathrm{Re}\langle\mathbf{r},\mathbf{h},\bar{\mathbf{t}}\rangle$	$\mathbf{h},\mathbf{r},\mathbf{t}\in\mathbb{C}^k$
ConvE (Dettmers et al., 2017)	$\langle\sigma(\mathrm{vec}(\sigma([\mathbf{r},\mathbf{h}]*\Omega))\mathbf{W}),\mathbf{t}\rangle$	$\mathbf{h},\mathbf{r},\mathbf{t}\in\mathbb{R}^k$
RotatE (Sun et al., 2017)	$-\|\|\mathbf{h}\circ\mathbf{r}-\mathbf{t}\|\|^2$	$\mathbf{h},\mathbf{r},\mathbf{t}\in\mathbb{C}^k$, $\|r_i\|=1$

Drawbacks & Limitation

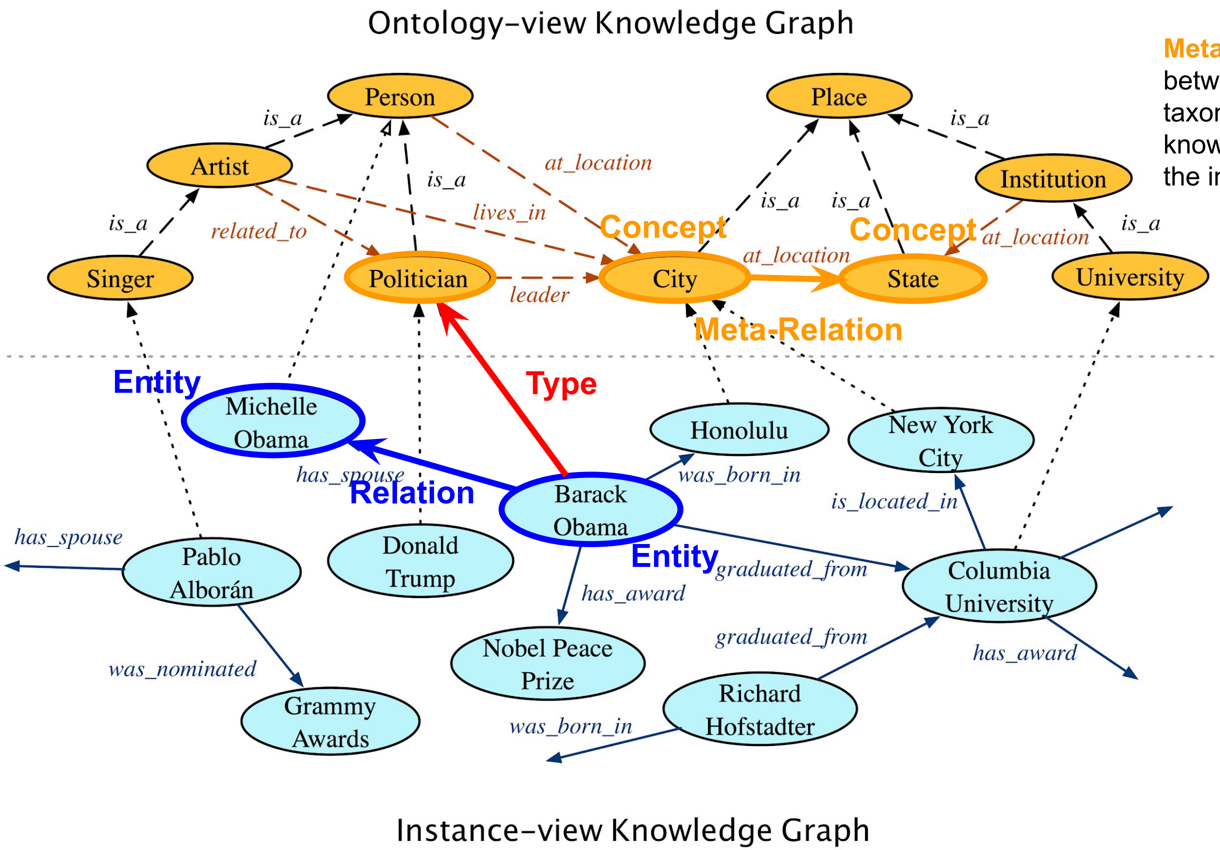
- Most existing approaches embed instance-level knowledge.
- KGs have both specific instances and general ontological concepts.



Outline

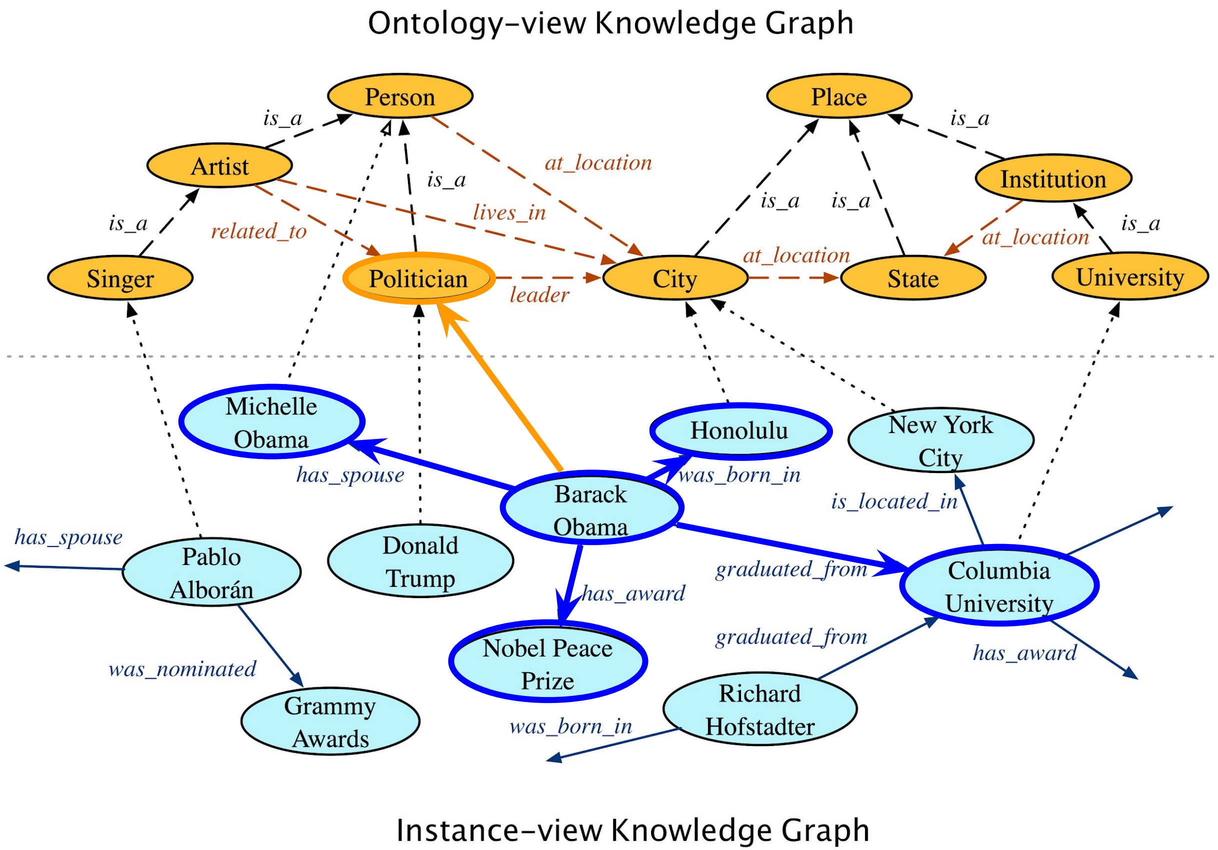
- Background: Knowledge Graphs and Embeddings
- **Formulation: Two-view Knowledge Graphs**
- JOIE Modeling: Cross-view & Intra-view
- Experimental Results
- Conclusion & Future Work

Two-view KG: More than a instance view



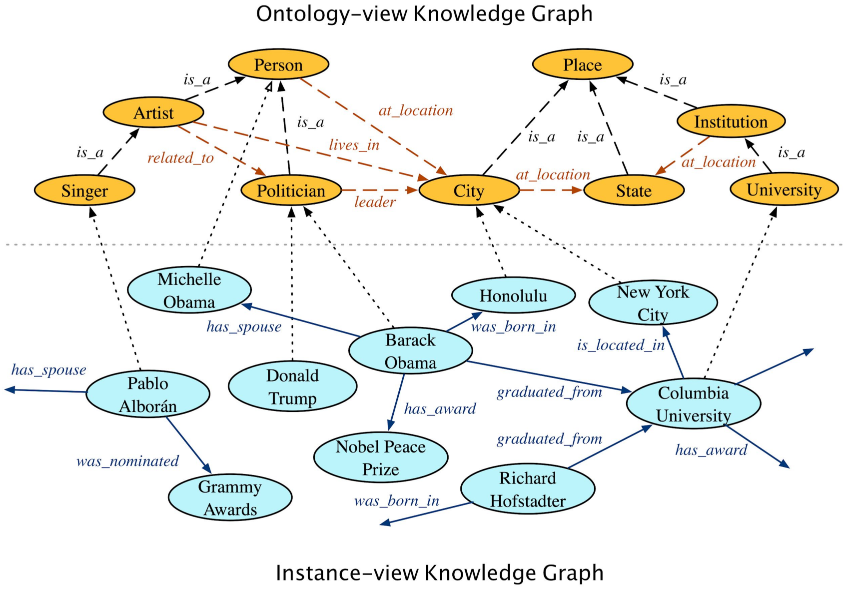
Meta-relation: Relations between concepts, including taxonomy, common-sense knowledge, which differ from the instance view.

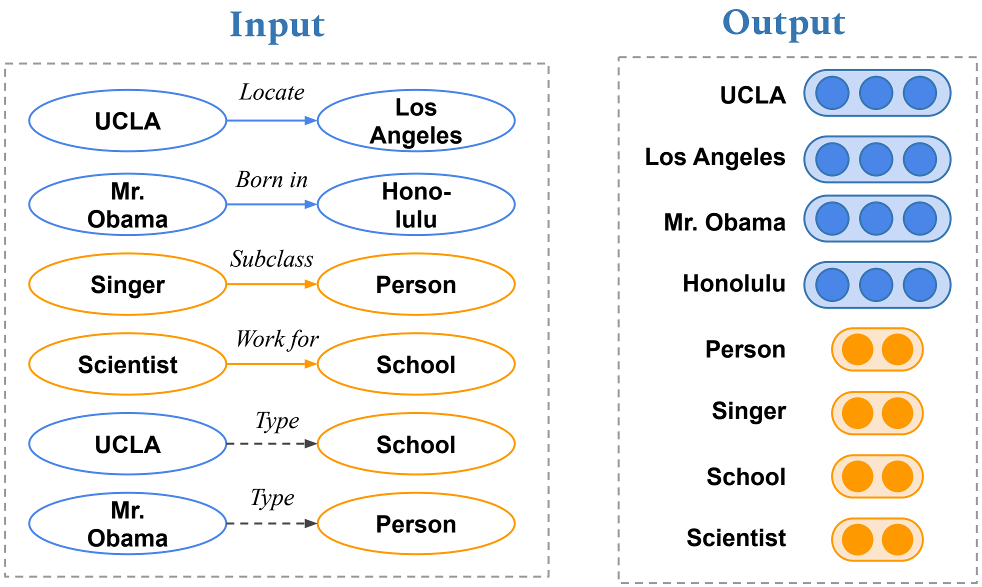
Two-view KG: More than just a set of triples



Problem Formulation

- **Input:** Instance-view KG triples, ontology-view KG triples, cross-view type links
- **Output:** Embeddings of entities, concepts, relations and meta-relations





Why Two-view KG Embeddings?

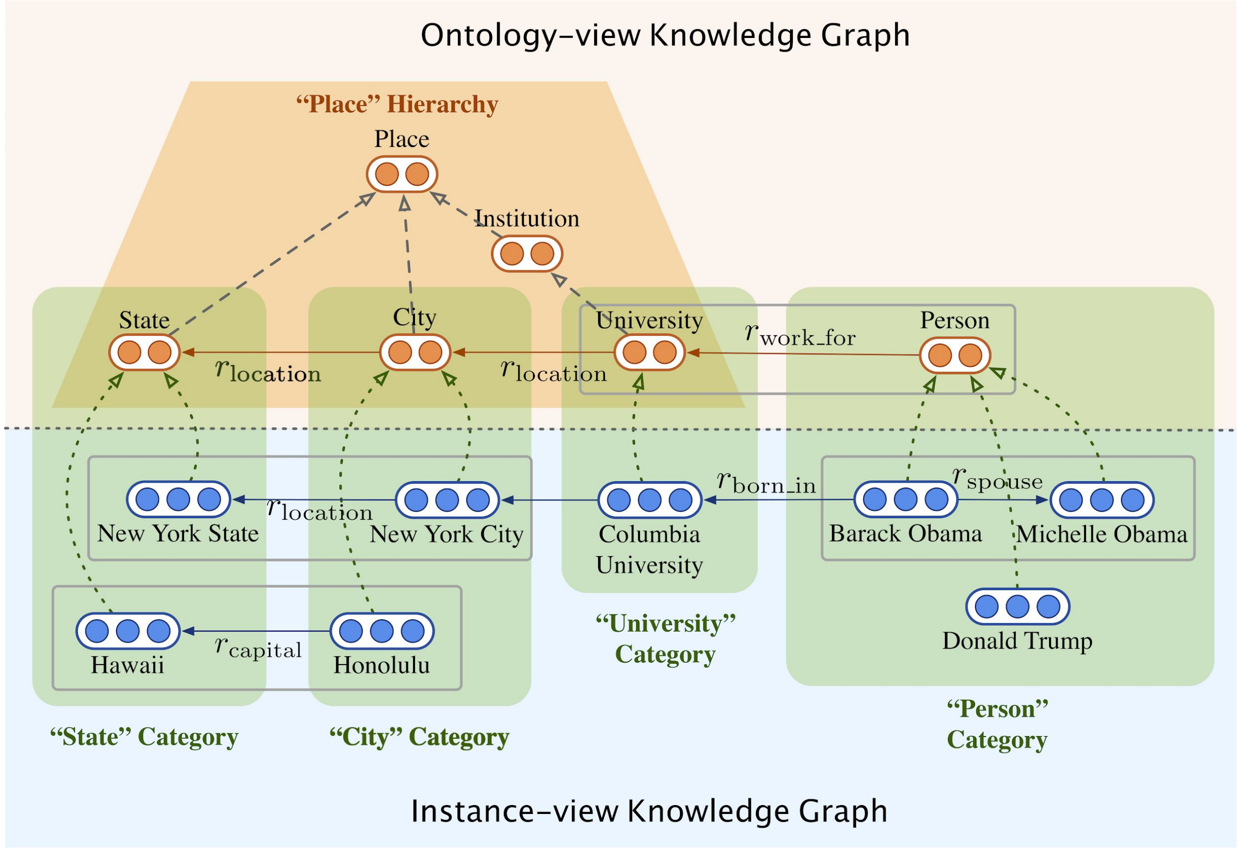
- Many existing KGs, such as YAGO and DBpedia, have constructed two views.
- Two views represent different levels of abstraction for relational knowledge, and can be used to enhance each other.
- Embeddings of a two-view KG provide more natural and clearer knowledge organization and curation, and are in line with human cognition.

Outline

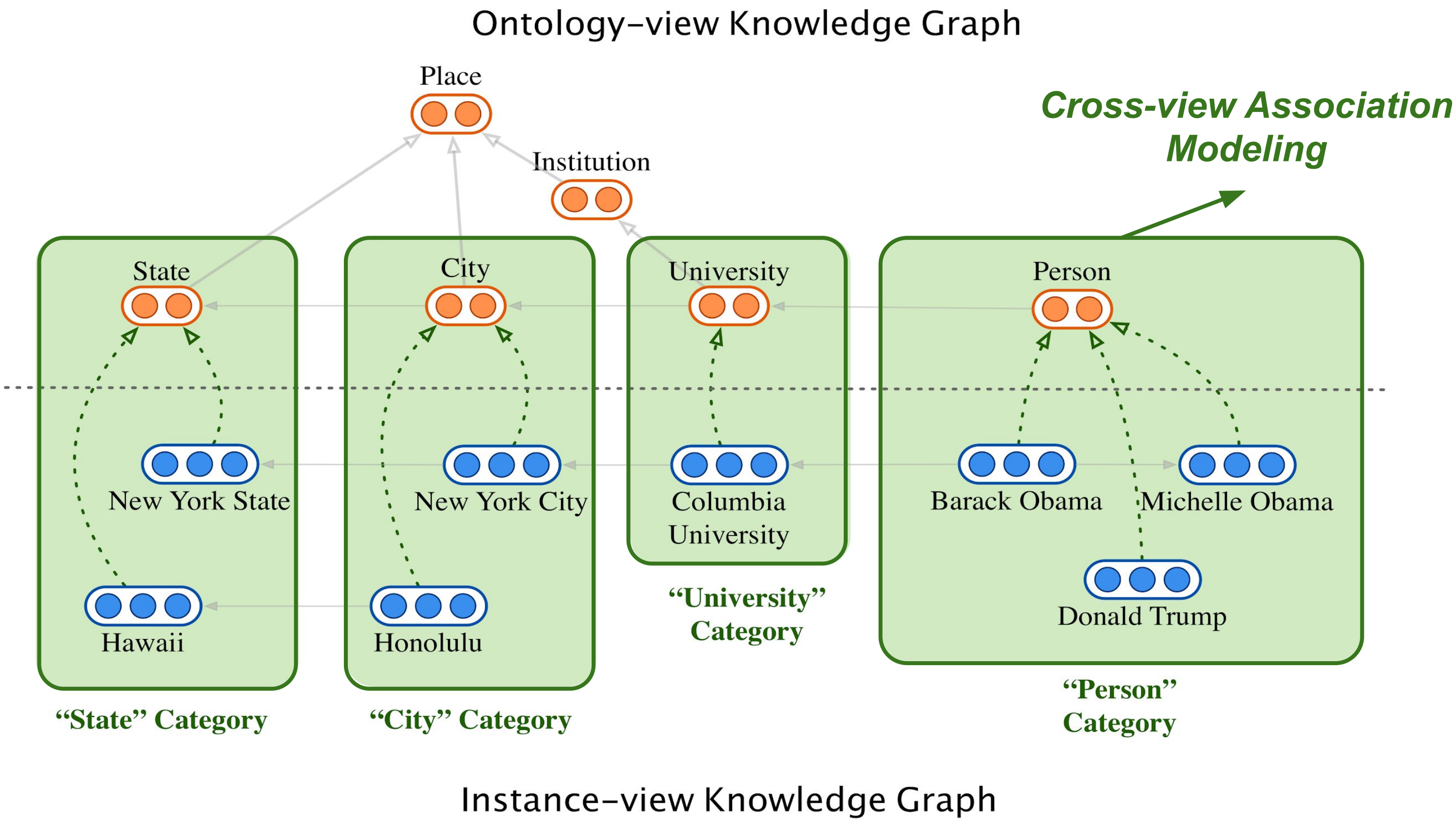
- Background: Knowledge Graphs and Embeddings
- Formulation: Two-view Knowledge Graphs
- **JOIE Modeling: Cross-view & Intra-view**
- Experimental Results & Case Study
- Conclusion & Future Work

JOIE: Modeling

- Cross-view Association model
- Intra-view model

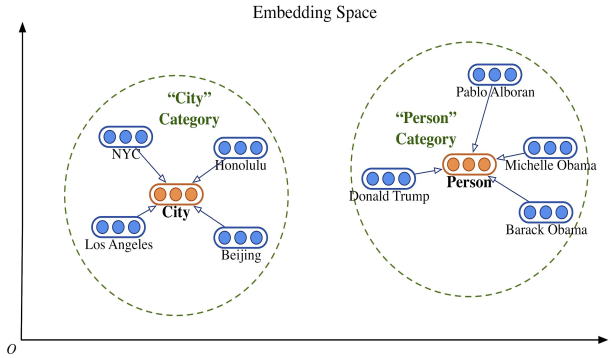


JOIE: Cross-view Association Model



JOIE: Cross-view Model

- **Goal:** capture associations between the entities **e** and corresponding concepts **c**



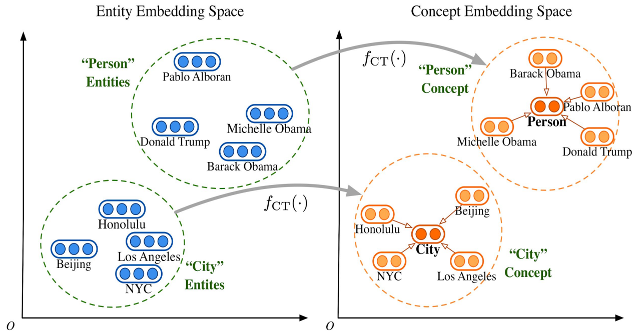
- **Cross-view Grouping (CG)**

$$J_{\text{Cross}}^{\text{CG}} = \frac{1}{|\mathcal{S}|} \sum_{(e,c)\in\mathcal{S}} \left[||\mathbf{c} - \mathbf{e}||_2 - \gamma^{\text{CG}} \right]_+$$

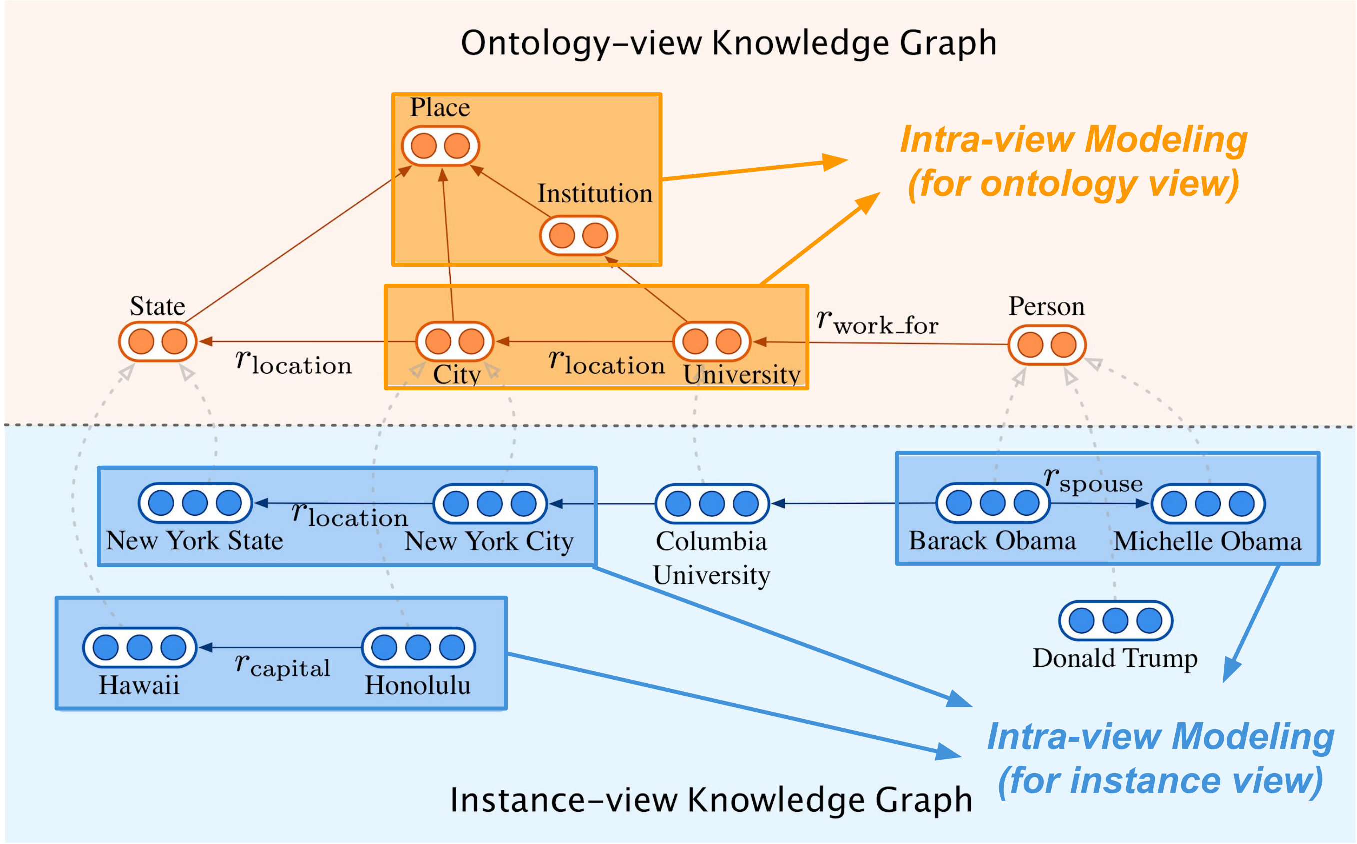
- **Cross-view Transformation (CT)**

$$f_{\text{CT}}(\mathbf{e}) = \sigma(\mathbf{W}_{\text{ct}} \cdot \mathbf{e} + \mathbf{b}_{\text{ct}})$$

$$J_{\text{Cross}}^{\text{CT}} = \frac{1}{|\mathcal{S}|} \sum_{\substack{(e,c)\in\mathcal{S} \\ \wedge (e,c')\notin\mathcal{S}}} \left[\gamma^{\text{CT}} + ||\mathbf{c} - f_{\text{CT}}(\mathbf{e})||_2 - ||\mathbf{c}' - f_{\text{CT}}(\mathbf{e})||_2 \right]_+$$



JOIE: Intra-view Model



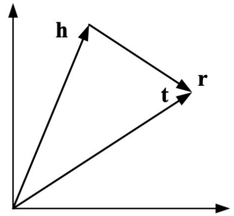
JOIE: Intra-view Model for Instance View

- Goal: To embed the relational structures in the instance view of the KB
- Apply any KG embedding techniques on instance view
  - Three representatives: TransE, DistMult, and HolE

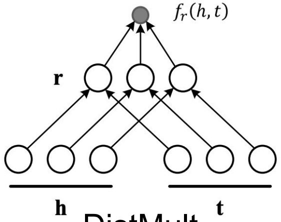
$$f_{\text{TransE}}(\mathbf{h},\mathbf{r},\mathbf{t}) = -||\mathbf{h}+\mathbf{r}-\mathbf{t}||_2$$
$$f_{\text{Mult}}(\mathbf{h},\mathbf{r},\mathbf{t}) = (\mathbf{h}\circ\mathbf{t})\cdot\mathbf{r}$$
$$f_{\text{HolE}}(\mathbf{h},\mathbf{r},\mathbf{t}) = (\mathbf{h}\star\mathbf{t})\cdot\mathbf{r}$$

- Training on marginal ranking loss

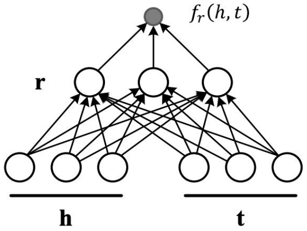
$$J_{\text{Intra}}^{\mathcal{G}} = \frac{1}{|\mathcal{G}|}\sum_{\substack{(h,r,t)\in\mathcal{G}\\ \wedge(h',r,t')\notin\mathcal{G}}}\left[\gamma^{\mathcal{G}} + f(\mathbf{h}',\mathbf{r},\mathbf{t}') - f(\mathbf{h},\mathbf{r},\mathbf{t})\right]_+$$



TransE



DistMult



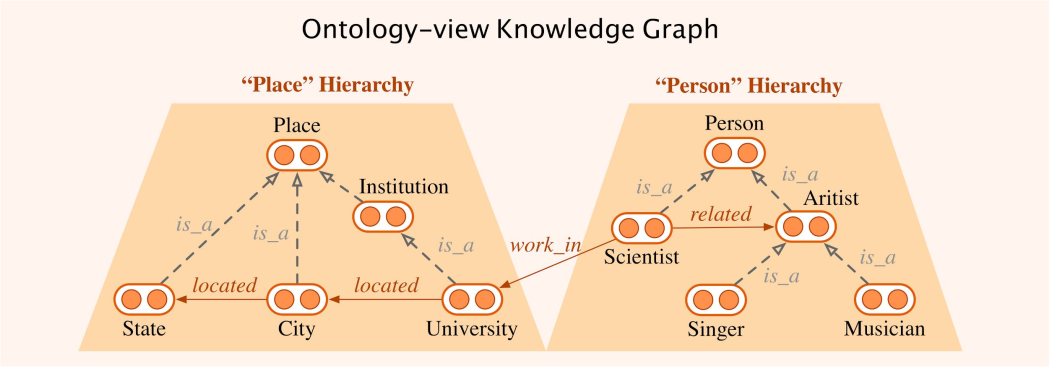
HolE

JOIE: Intra-view Model for Ontology View

- We can still follow the same techniques as the instance view. $J_{\text{Intra}} = J_{\text{Intra}}^{\mathcal{G}_I} + \alpha_1 \cdot J_{\text{Intra}}^{\mathcal{G}_O}$
- However, the hierarchical structure of the ontology-view represents critical semantics, with special meta relations such as "*is_a*" and "*subclass*".

c_l : Scientist c_h :Person

$$g_{\text{HA}}(\mathbf{c}_h) = \sigma(\mathbf{W}_{\text{HA}} \cdot \mathbf{c}_l + \mathbf{b}_{\text{HA}})$$



- Similar to CT model, we model such hierarchical structures in,

$$J_{\text{Intra}}^{\text{HA}} = \frac{1}{|\mathcal{T}|} \sum_{\substack{(c_l, c_h) \in \mathcal{T} \\ \wedge (c_l, c_h') \notin \mathcal{T}}} \left[\gamma^{\text{HA}} + ||\mathbf{c}_h - g(\mathbf{c}_l)||_2 - ||\mathbf{c_h}' - g(\mathbf{c_l})||_2\right]_+$$

JOIE: Summary & Model Highlights

- Two model components: Cross-view model and intra-view model
- Cross-view association model $\Rightarrow J_{\text{Cross}}$
  - Categorical grouping (CG)
  - Categorical transformation (CT)
- Intra-view model $\Rightarrow J_{\text{Intra}}$
  - Can apply any KG embedding on each view
  - Hierarchical-aware modeling on ontological view specifically for taxonomy meta relations
- Joint training on cross-view loss and intra-view loss

$$J = J_{\text{Intra}} + \omega \cdot J_{\text{Cross}}$$

Outline

- Background: Knowledge Graphs and Embeddings
- Formulation: Two-view Knowledge Graphs
- JOIE Modeling: Cross-view & Intra-view
- **Experimental Results & Case Study**
- Conclusion & Future Work

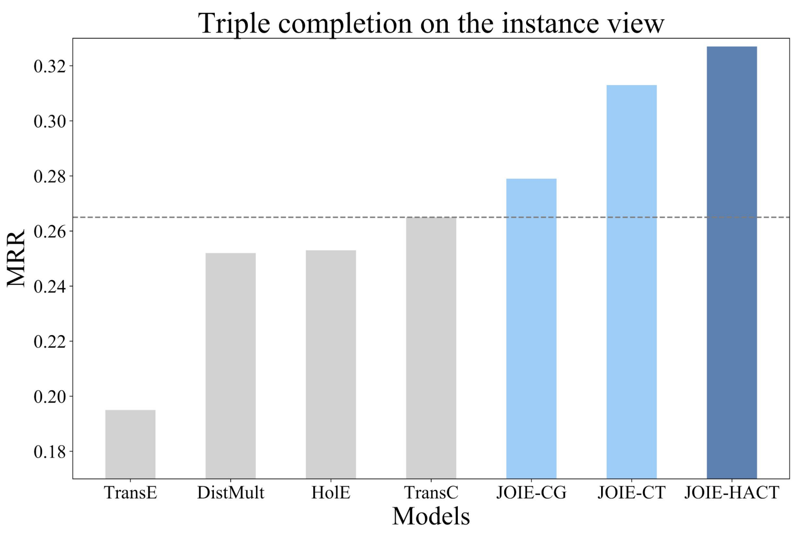
Experiment Setup

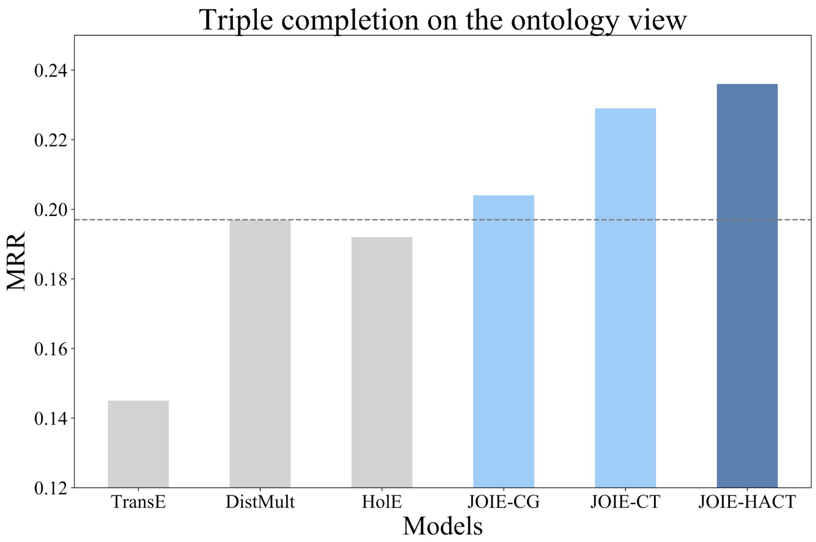
- Datasets: YAGO26K-906 (from YAGO) and DB111K-184 (from DBpedia)
- Tasks: **Triple completion** and **entity typing**
- Evaluation metrics
  - Triple completion: MRR, Hit@*K* score (*K*=1,3,10)
  - Entity typing: Accuracy (Hit@1), Hit@3 Score
- Baselines: TransE, DistMult, HolE, TransC, MTransE

Dataset	Instance Graph $\mathcal{G}_I$			Ontology Graph $\mathcal{G}_O$			Type Links $\mathcal{S}$
	#Entities	#Relations	#Triples	#Concepts	#Meta-relations	#Triples	
YAGO26K-906	26,078	34	390,738	906	30	8,962	9,962
DB111K-174	111,762	305	863,643	174	20	763	99,748

Task 1: Triple Completion

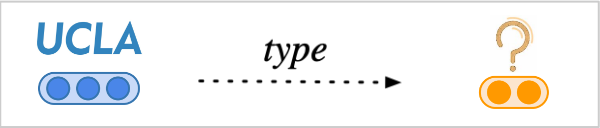
- Given the head and predicate of a triple, what is the most likely tail (answer)?

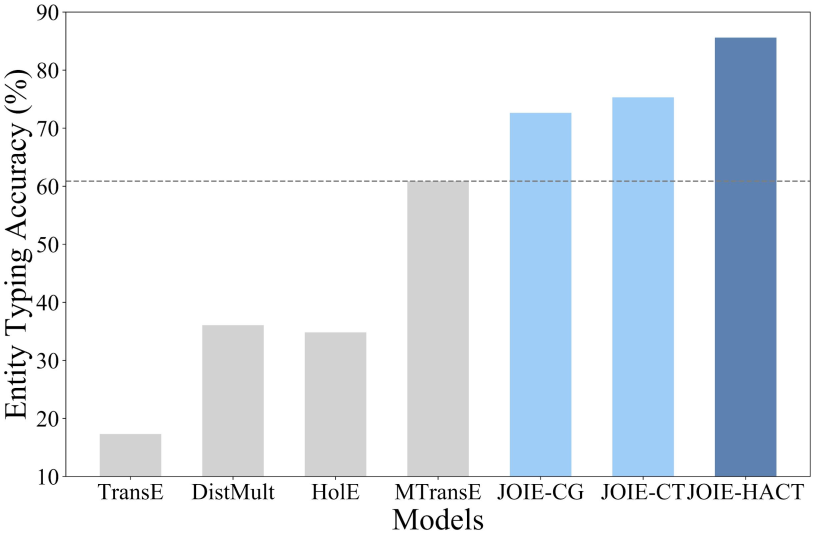




Task 2: Entity Typing

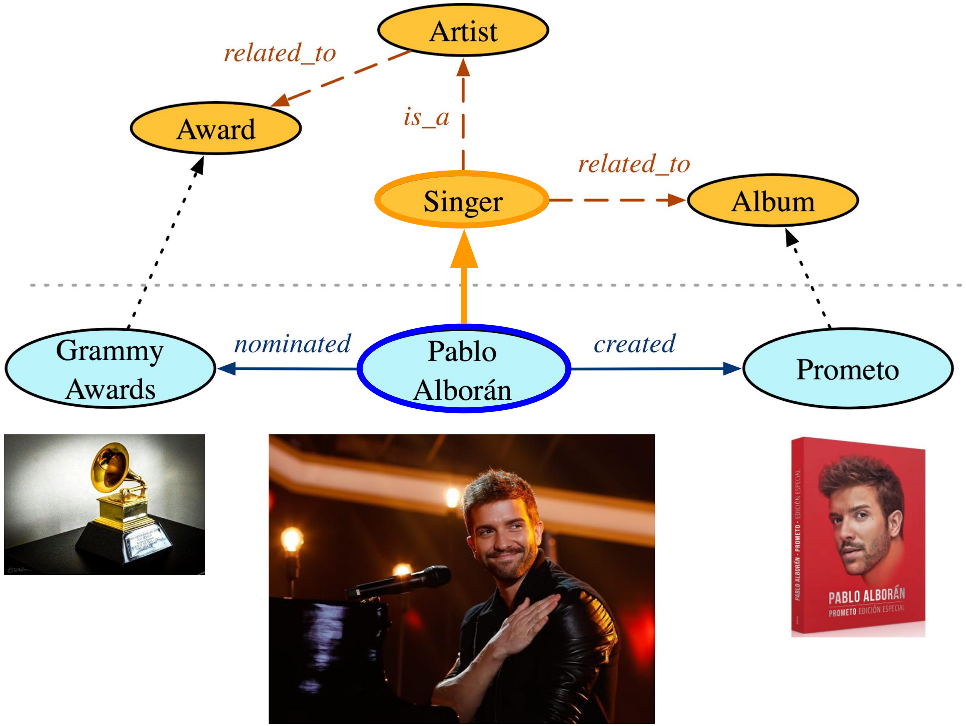
- Given an entity without a known type, what is the most likely type (concept) that it associates with?





Type inference on 30% of all entities on YAGO.

Task 2+: Long-tail Entity Typing



Example of long-tail entity typing

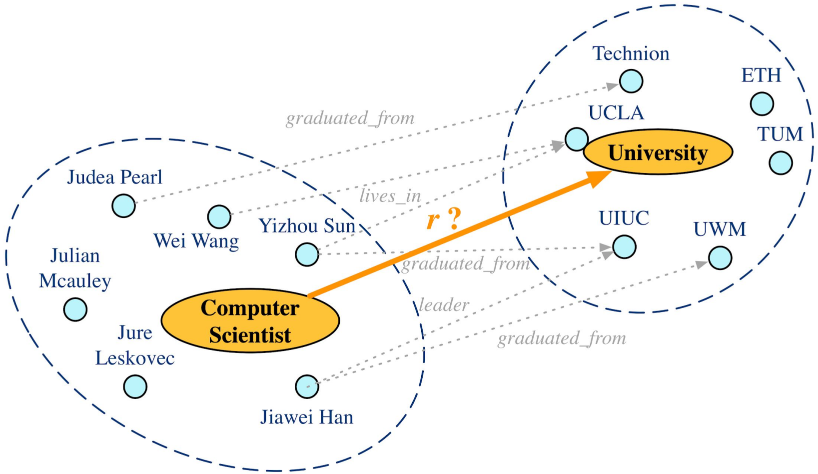
Entity	Model	Top 3 Concept Prediction
Laurence Fishburne	DistMult	football team, club, team
	MTransE	writer, **person**, artist
	JOIE	**person**, artist, philosopher
Warangal City	DistMult	country, village,**city**
	MTransE	administrative region, **city**, settlement
	JOIE	**city**, town, country
Royal Victor -ian Order	DistMult	person, writer, administrative region
	MTransE	election, award, **order**
	JOIE	award, **order**, election

Entity typing accuracy on long-tail entities

Datasets	YAGO26K-906		
Metrics	MRR	Acc.	Hit@3
DistMult	0.156	10.89	25.33
MTransE	0.526	46.45	67.25
JOIE-TransE-CG	0.708	59.97	79.80
JOIE-TransE-CT	0.737	62.05	82.60
JOIE-HATransE-CT	**0.802**	**69.66**	**87.75**

Task 3: Ontology Population

→ JOIE can help enhance the quality of ontology view and make it more complete and informative by populating the instance-level knowledge.



Examples of ontology population

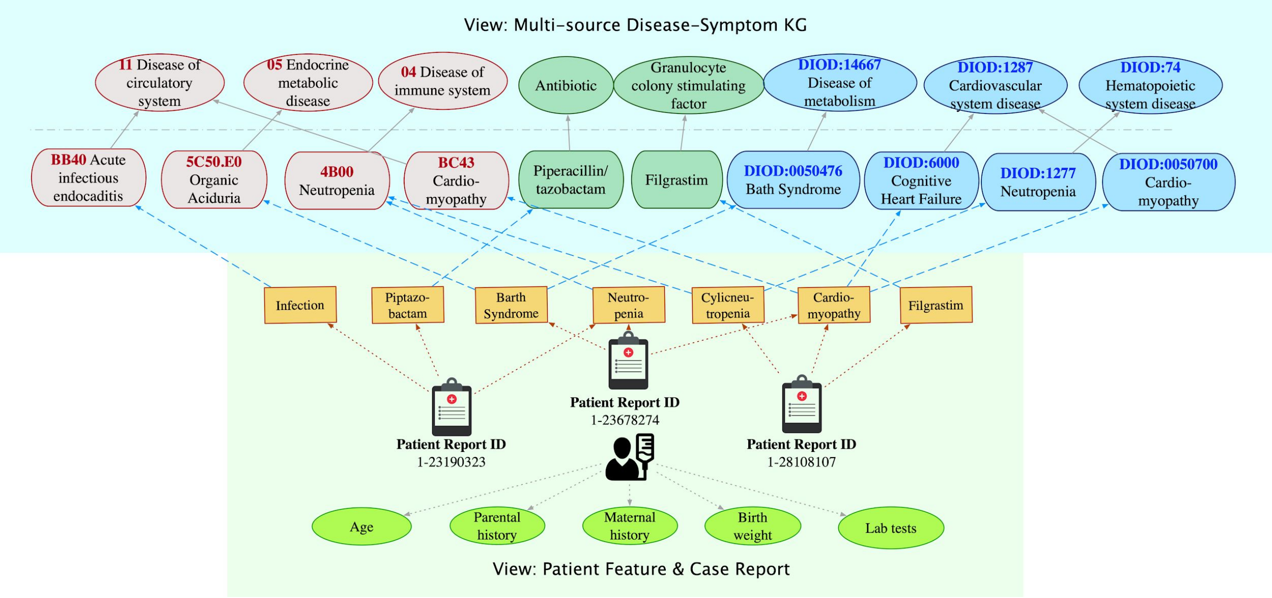
Query	Top 3 Populated Triples with distances
(scientist, ?r, university)	scientist, ***graduated from***, university (0.499) scientist, ***isLeaderOf***, university (1.082) scientist, *isKnownFor*, university (1.098)
(boxer, ?r, club)	boxer, ***playsFor***, club (1.467) boxer, *isAffiliatedTo*, club (1.474) boxer, ***worksAt***, club (1.479)
(scientist, ?r, scientist)	scientist, ***doctoralAdvisor***, scientist (0.204) scientist, ***doctoralStudent***, scientist (0.221) scientist, ***relative***, scientist (0.228)

Outline

- Background: Knowledge Graphs and Embeddings
- Formulation: Two-view Knowledge Graphs
- JOIE Modeling: Cross-view & Intra-view
- Experimental Results & Case Study
- **Conclusion & Future Work**

Conclusion & Future Work

- Joint learning on the instance and ontology views improves the KG embeddings.
- Incorporating ontologies in KGs is beneficial.
- Two-view KG modeling can be applied in a wide selection of interdisciplinary applications.
  - Disease-symptom with multiple medical KGs for automated patient case report analysis.



*Join us in **Bio-KARE** community!*

Sign up with your email:

Scan this QR Code or visit:
Bio-KARE.org/KDD2019

National Center for Biomedical
Knowledge Architecture
& Relationship Enrichment

Dr. Wei Wang (University of California, Los Angeles)
Dr. Peipei Ping (University of California, Los Angeles)
Dr. Cathy Wu (University of Delaware)
Dr. Jiawei Han (University of Illinois, Urbana-Champaign)

Do you have biomedicine's next TOP model?

Validate your model with Bio-KARE

Join the Bio-KARE community

of researchers, developers, and clinicians today

Opportunity to revolutionize model development and empower your research.
Sign up with your email:

Scan this QR Code or visit: Bio-KARE.org/KDD2019

UCLA Samueli Computer Science

ScAi
SCALABLE ANALYTICS INSTITUTE

Thank you!

Q & A

Join us in ***Bio-KARE*** *community!*

Sign up with your email:

Scan this QR Code or visit:
Bio-KARE.org/KDD2019