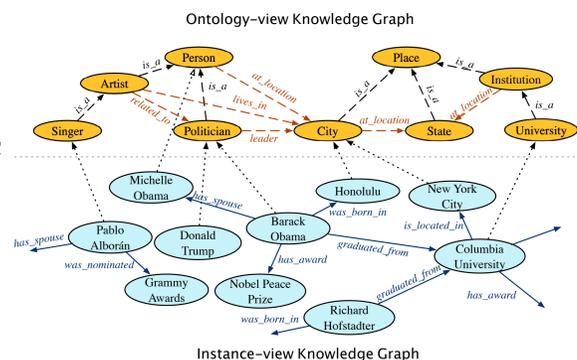


Junheng Hao, Muhao Chen, Wenchao Yu, Yizhou Sun and Wei Wang
Computer Science Department, University of California, Los Angeles
Email: jhao@cs.ucla.edu | Website: <https://www.haojunheng.com/project/joie-kdd/>

EMBEDDING TWO-VIEW KNOWLEDGE GRAPH

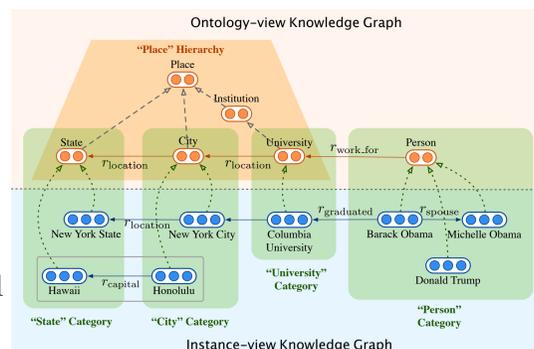
- Instance View**
 - Relational triplets between specific entities
- Ontology View**
 - Triplets of abstract concepts connected by semantic meta-relations
 - Hierarchical structure of the ontology view (e.g. "subclass" meta-relations between concepts)
- Cross-view Type links (alignment)**
 - Connecting concepts and instances (entity types)



JOIE: MODEL OVERVIEW

- Embedding-based method**
 - Learning latent representations for all objects in KG
- Cross-view Association Models**
 - Cross-view Grouping (CG)
 - Cross-view Transformation (CT)
- Intra-view Models**
 - Default intra-view models (TransE, DistMult, HoIE)
 - Hierarchy-aware modeling for the ontology view
- Joint training on intra-view model (for both views) and cross-view model.**

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



CROSS-VIEW MODELS

Goal: capture many-to-one association between the entities e and the corresponding concept c

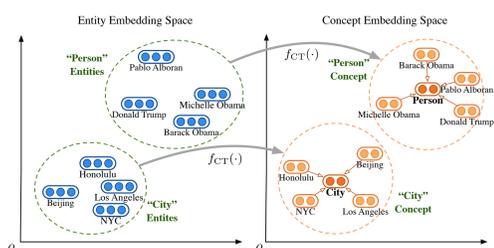
Cross-view Grouping

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

Cross-view Transformation

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

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



INTRA-VIEW MODELS

Goal: Embed the relational structures in each view

Techniques: TransE, DistMult, HoIE

$$f_{\text{TransE}}(\mathbf{h}, \mathbf{r}, \mathbf{t}) = -||\mathbf{h} + \mathbf{r} - \mathbf{t}||_2$$

$$f_{\text{Mult}}(\mathbf{h}, \mathbf{r}, \mathbf{t}) = (\mathbf{h} \circ \mathbf{t}) \cdot \mathbf{r}$$

$$f_{\text{HoIE}}(\mathbf{h}, \mathbf{r}, \mathbf{t}) = (\mathbf{h} \star \mathbf{t}) \cdot \mathbf{r}$$

Triple Loss

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

Hierarchy-Aware model for the ontology-view

⇒ For example, given, c_l : singer and c_h : person:

$$g_{\text{HA}}(c_h) = \sigma(\mathbf{W}_{\text{HA}} \cdot c_l + \mathbf{b}_{\text{HA}})$$

$$J_{\text{Intra}}^{\text{HA}} = \frac{1}{|\mathcal{T}|} \sum_{(c_l, c_h) \in \mathcal{T}} [\gamma^{\text{HA}} + ||c_h - g(c_l)||_2 - ||c_h' - g(c_l)||_2]_+$$

Total training loss for Intra-view model

$$J_{\text{Intra}} = J_{\text{Intra}}^{\mathcal{G}} + \alpha_1 \cdot J_{\text{Intra}}^{\mathcal{G}_O} + \alpha_2 \cdot J_{\text{Intra}}^{\text{HA}}$$

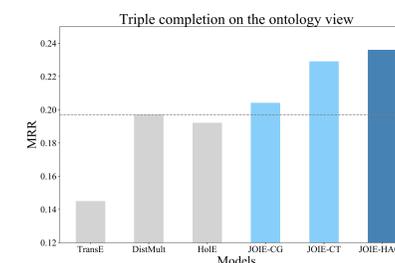
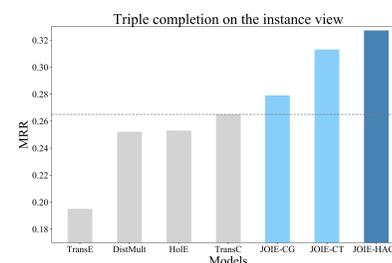
RESULTS: TRIPLE COMPLETION & ENTITY TYPING

Dataset YAGO26K-906 (from YAGO) and DB11K-184 (from DBpedia)

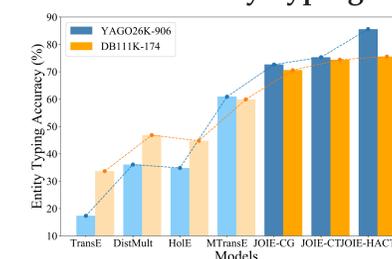
Dataset	\mathcal{G}_I Ent.	\mathcal{G}_I Rel.	\mathcal{G}_I Tri.	\mathcal{G}_O Con.	\mathcal{G}_O MRel.	\mathcal{G}_I Tri.	Typelinks
YAGO26K-906	26,078	34	390,738	906	30	8,962	9,962
DB11K-174	111,762	305	863,643	174	20	763	99,748

Evaluation Triple completion (link prediction) and entity typing

Task 1: Triple completion on both views (YAGO)

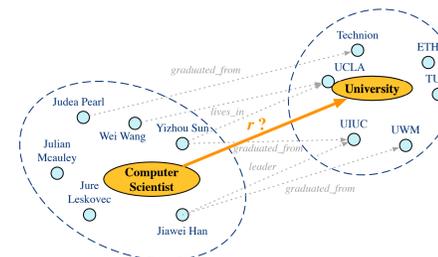


Task 2: Entity Typing



JOIE: APPLICATIONS

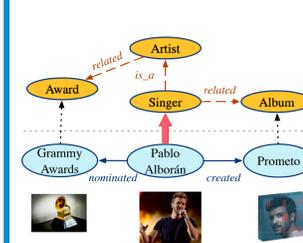
Ontology Population



Examples of populated ontology

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

Long-tail Entity Typing



Examples of long-tail entity typing

Entity	Model	Top 3 Predictions	Long-tail entity typing accuracy	
			Datasets	YAGO
Laurence Fishburne	DistMult	football, club, team	10.89	16.48
	MTransE	writer, person, artist	46.45	46.67
	JOIE	person, artist, writer	59.97	64.45
Warangal City	DistMult	country, village, city	62.05	66.35
	MTransE	region, city, settlement	62.05	66.35
	JOIE	city, town, country	69.66	67.34

ACKNOWLEDGEMENT

Supported by NIH R01GM115833, U01HG008488, U54GM114833 NSF DBI-1565137, DGE-1829071, NSF III-1705169, NSF Career Award 1741634 and Amazon Research Award.



Paper Link

