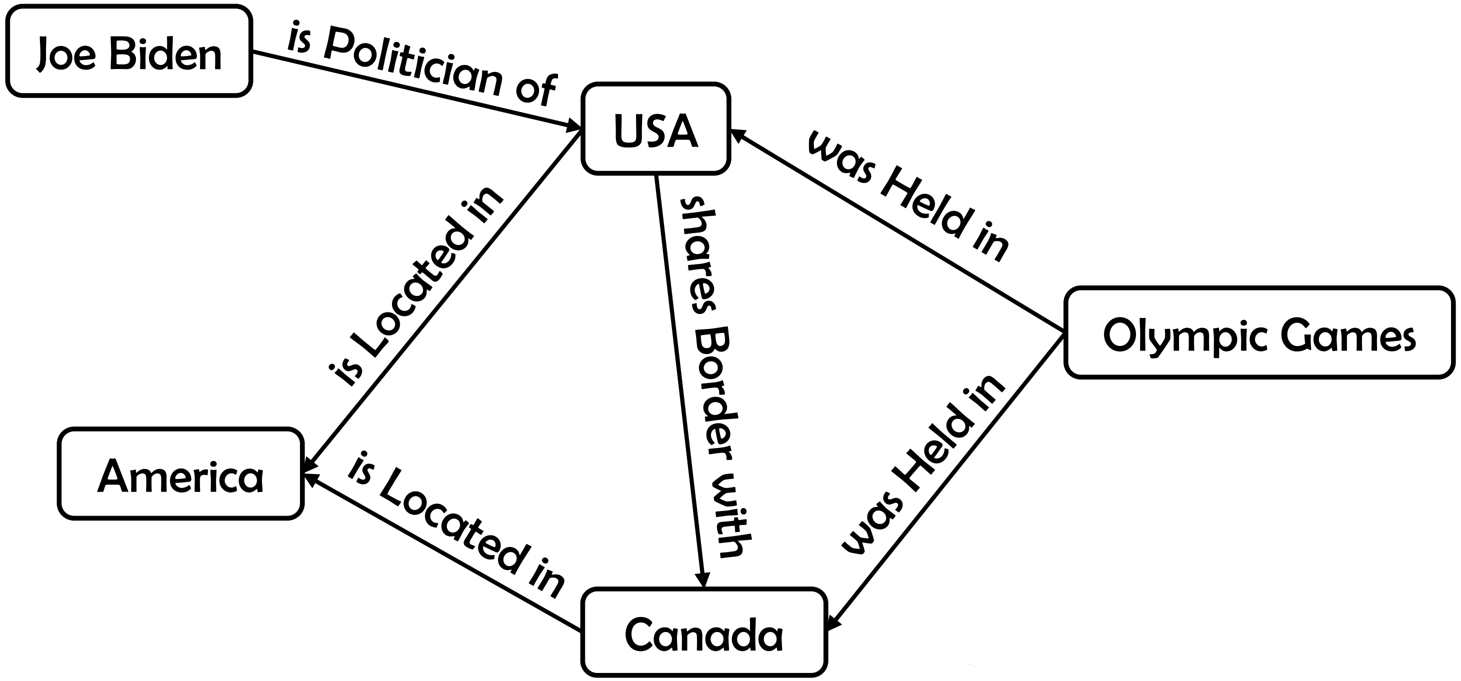
Knowledge Graph Embedding via Metagraph Learning

SIGIR 2021 Short Paper

Chanyoung Chung and Joyce Jiyoung Whang
KAIST School of Computing

Knowledge Graphs

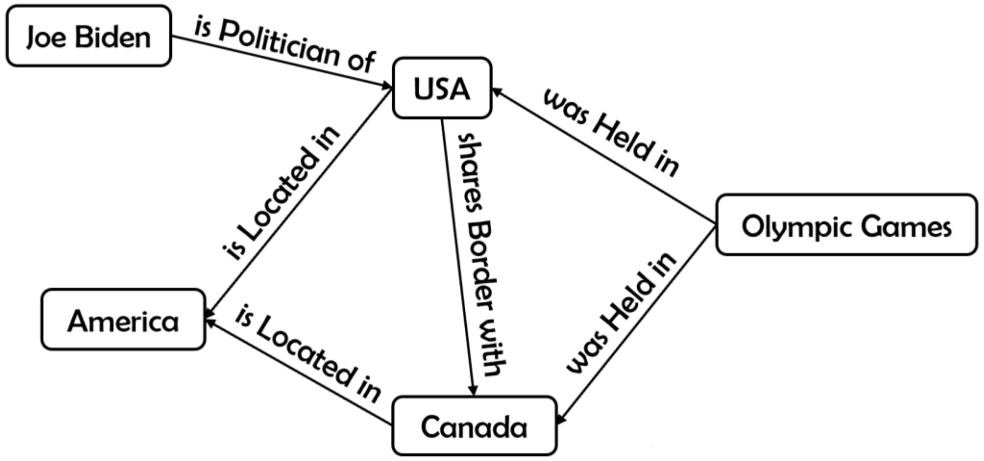
- Human knowledge as a directed graph
  - Each fact is represented as a triplet (head entity, relation, tail entity)

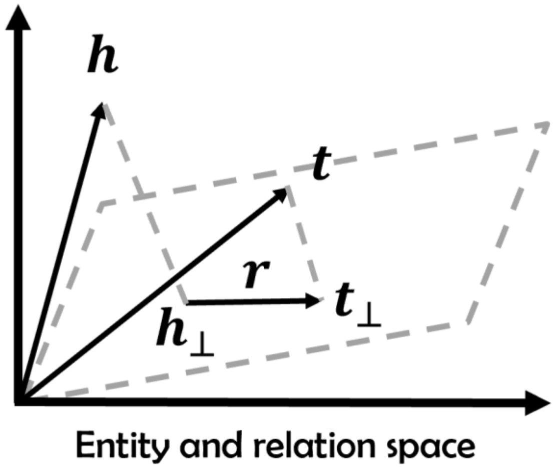


(Joe Biden, is Politician of, USA)
(USA, is Located in, America)
(Canada, is Located in, America)
(USA, shares Border with, Canada)
(Olympic Games, was Held in, USA)
(Olympic Games, was Held in, Canada)

Knowledge Graph Embedding

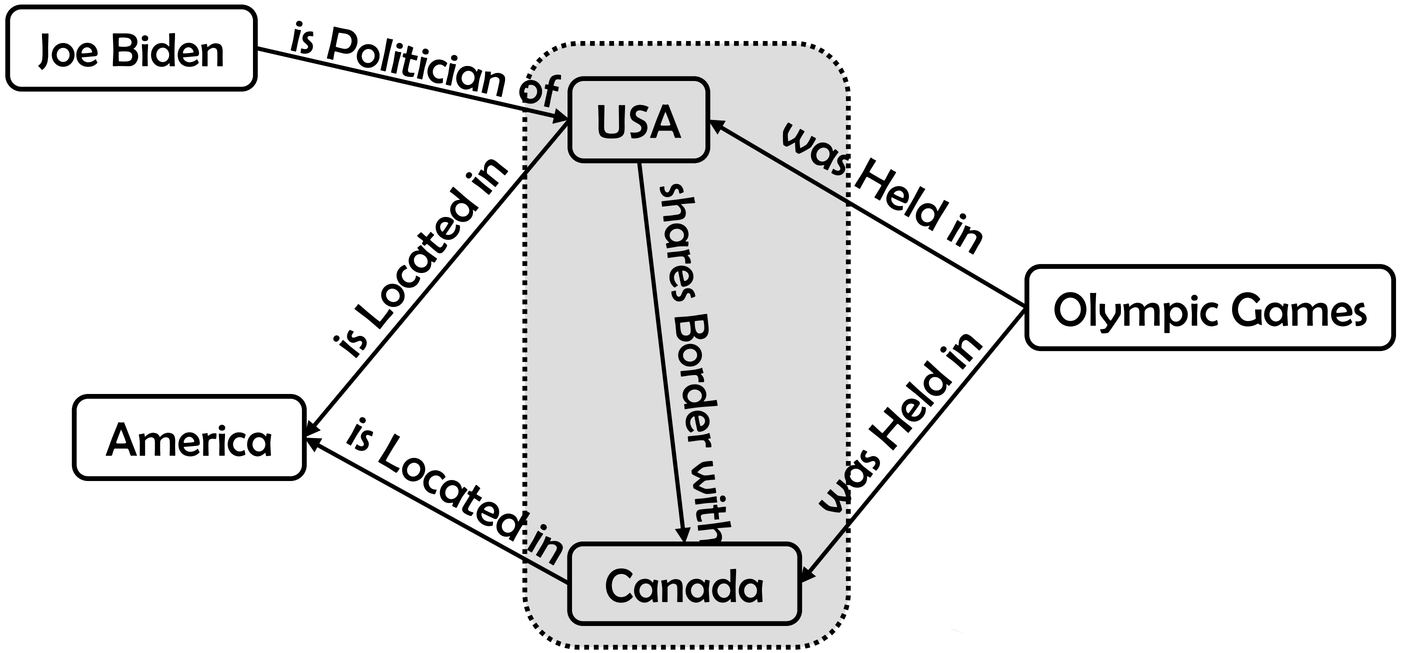
- Representation Learning Technique
  - Projects entities and relations into a continuous feature space.
  - Applicable to solving diverse problems such as link prediction.
- Knowledge Graph Embedding Methods
  - Translational Distance Models (e.g., TransE, TransH, TransR)
  - Semantic Matching Models (e.g., DistMult, RotatE)





Intuition

- Semantic closeness can be inferred by the structural similarity between entities.
  - If two entities share the same tail entity or the same head entity with the same relation, they might belong to the same semantic category.

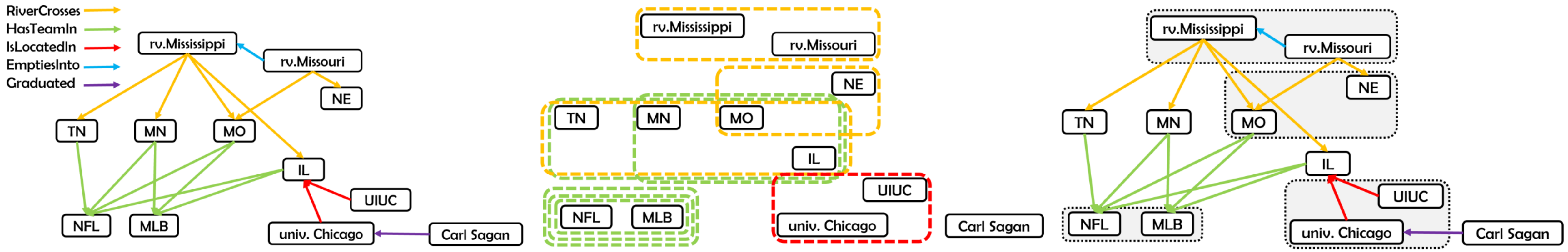


'USA' and 'Canada' are structurally similar.

- They share the same tail entity 'America' with the relation 'is Located in'.
- They share the same head entity 'Olympic Games' with the relation 'was Held in'.

Overview

- Hypergraph representation of a knowledge graph
  - Connect a set of semantically close entities by looking at the structure of a knowledge graph.
- Group similar entities by hypergraph clustering.
  - Based on the newly defined affinity metric, perform hypergraph clustering.



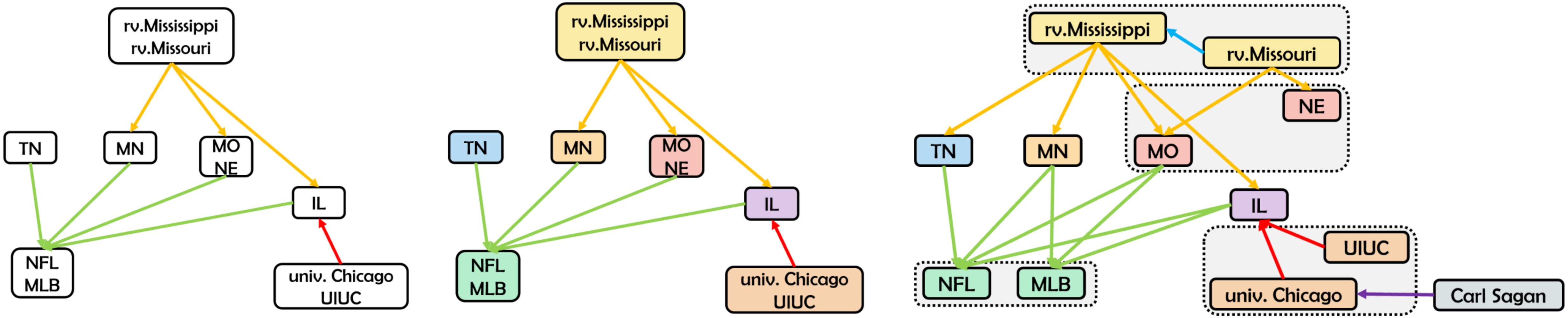
Knowledge Graph

Hypergraph Representation

Grouping Entities

Overview

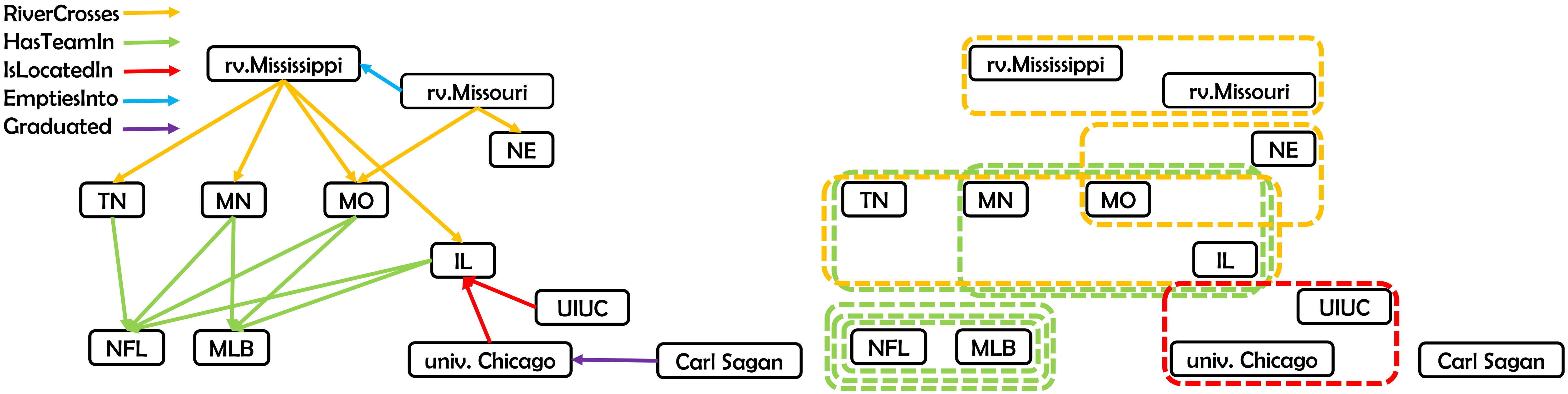
- Metagraph embedding
  - Form super-entities by merging entities in the same group and learn embeddings.
- Pre-train a knowledge graph embedding model
  - Initialize corresponding entities and relations with the learned representations.



Metagraph Metagraph Embedding Pre-trained Knowledge Graph Embedding

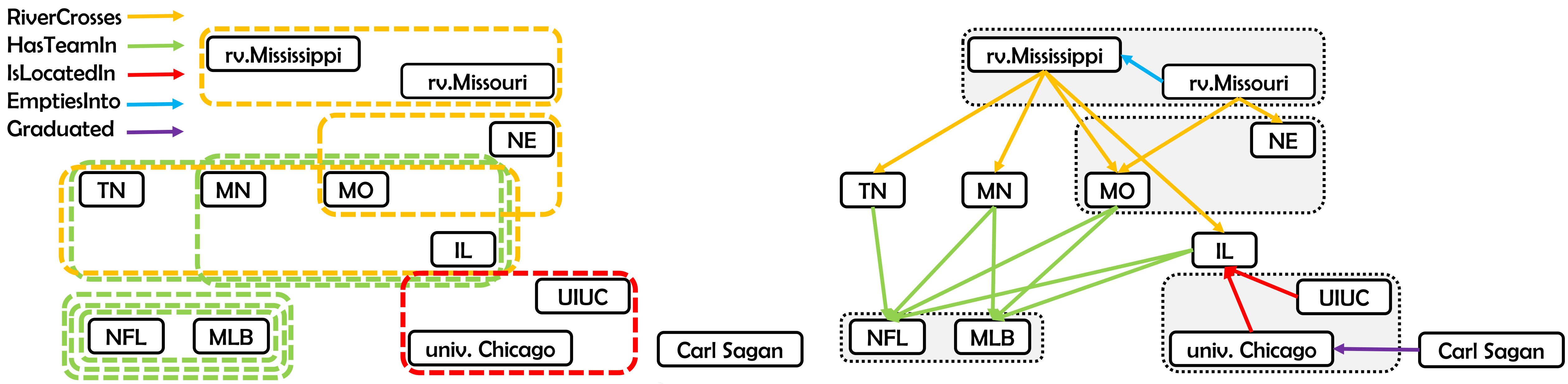
Hypergraph Representation

- Connect a set of entities that share the same head (tail) entity with the same relation.
  - A hyperedge can connect an arbitrary number of nodes.
- Define a new affinity score between two entities → reflects the structural similarity
  - Affinity between entities i & j: $a_{ij} = \sum_{l \in \mathcal{L}} 1/{d_l}^2$ where $\mathcal{L}$ is the set of hyperedges that contain both i and j, and d_l indicates the number of entities included in the hyperedge l.



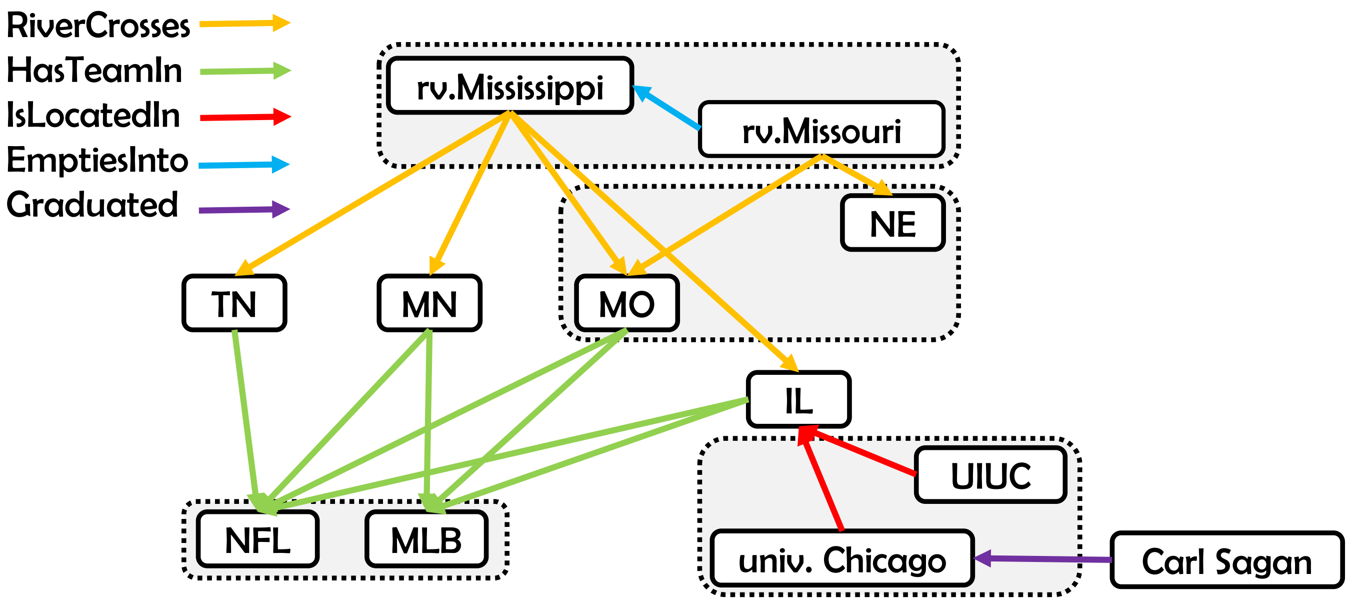
Grouping Entities

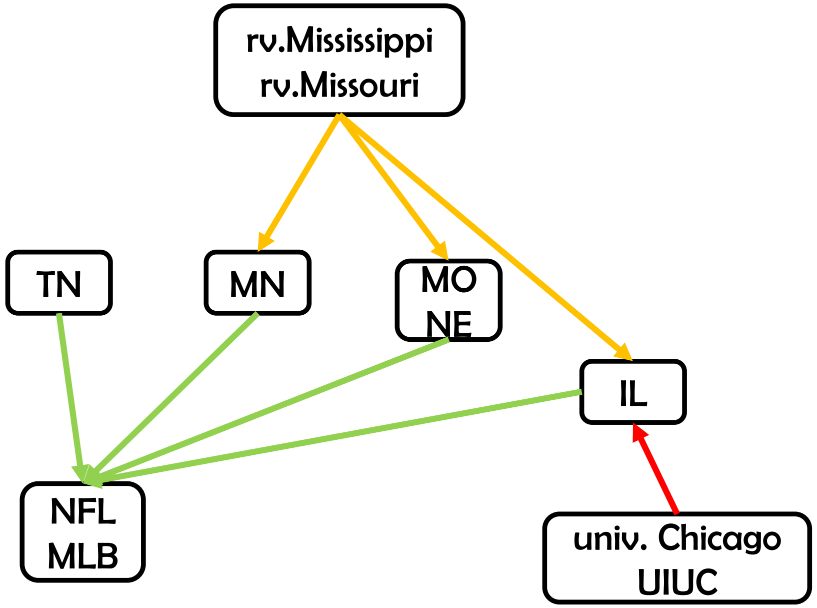
- Normalize the affinity scores with respect to individual entities.
- Perform hypergraph clustering with the normalized affinity scores to group entities.
  - We use an agglomerative hierarchical clustering.



Metagraph of a Knowledge Graph

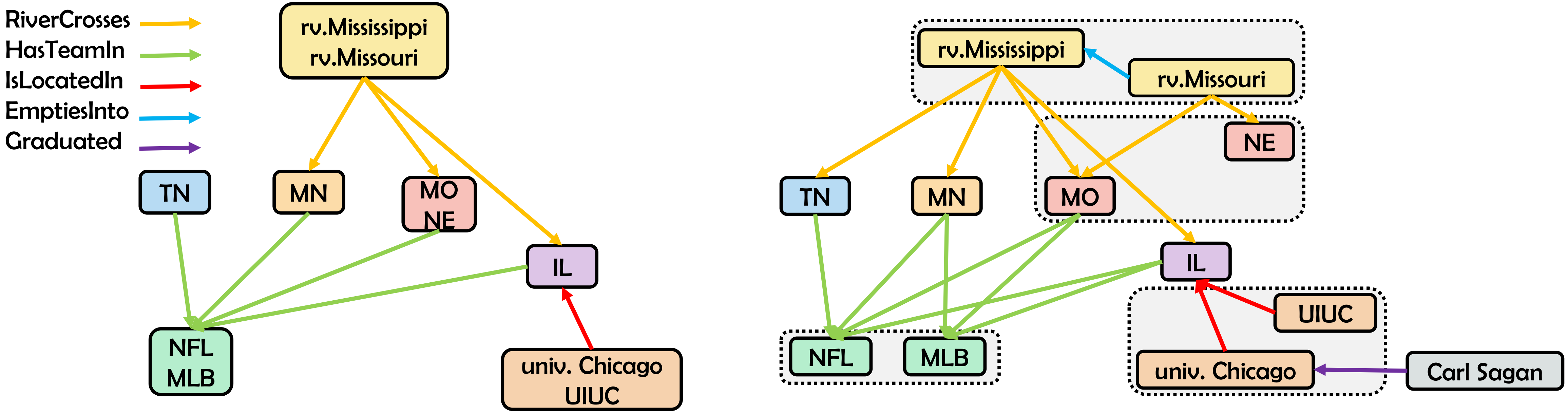
- Merge entities in the same group to form a super-entity.
- Metagraph → preserves a core structure of the given knowledge graph.
  - Within-group triplets are dropped.
  - Between-group triplets are probabilistically dropped.





Pre-training of Knowledge Graph Embedding

- Run a knowledge graph embedding method on the metagraph.
- Initialize the corresponding entities and relations in the original knowledge graph with the learned representations on the metagraph.
  - Entities in the same super-entity are initialized with the same representations.



Top Similar Entities from NELL-995

- Target entity: emotion_thankfulness
  - emotion_graditude, emotion_admiration, emotion_happiness, emotion_joy, emotion_deep_love, emotion_jealousy, emotion_thanks
- Target entity: software_microsoft_word
  - software_internet_explorer, software_microsoft_frontpage, software_microsoft_powerpoint, software_notepad, software_autocad
- Target entity: sport_american_football
  - sport_ski, sport_scout, sport_skiing, sport_golf, sport_judo

Datasets

- We cluster the entities so that the size of the metagraph becomes about half of the original knowledge graph in terms of the number of triplets in the train set.
  - We perform clustering only using the train set.

		Train set			Validation set			Test set		
		$\|V\|$	$\|R\|$	$\|E\|$	$\|V\|$	$\|R\|$	$\|E\|$	$\|V\|$	$\|R\|$	$\|E\|$
FB15K	Original KG	14,951	1,345	483,142	13,292	916	50,000	13,584	961	59,071
	Metagraph	10,387	1,142	224,824	8,377	686	21,778	8,611	703	25,433
NELL-995	Original KG	74,432	200	149,678	765	12	543	3,747	12	3,992
	Metagraph	48,693	200	67,588	382	12	261	2,387	12	2,118
WN18	Original KG	40,943	18	141,442	7,802	18	7,802	7,845	18	5,000
	Metagraph	27,250	18	63,423	3,695	18	3,695	3,689	18	2,227

Experimental Results

- Link Prediction Results
  - A positive gain indicates that our method is better than the original model.

		MR (↓)	MRR (↑)	Hit@10 (↑)	Total Gain (↑)
FB15K	TransE	89.0	0.596	0.733	
	meta-TransE	75.0	0.551	0.798	
	Gain (↑)	15.8%	-7.5%	8.8%	17.1%
	DistMult	106.4	0.414	0.644	
	meta-DistMult	143.7	0.541	0.786	
	Gain (↑)	-35.0%	30.8%	21.9%	17.7%
	RotatE	34.4	0.691	0.869	
	meta-RotatE	33.6	0.690	0.871	
	Gain (↑)	2.1%	-0.1%	0.2%	2.2%

Experimental Results

- Link Prediction Results
  - A positive gain indicates that our method is better than the original model.

		MR (↓)	MRR (↑)	Hit@10 (↑)	Total Gain (↑)
NELL-995	TransE	7202.4	0.278	0.477	
	meta-TransE	6507.5	0.287	0.434	
	Gain (↑)	9.6%	3.3%	-9.1%	3.8%
	DistMult	10312.7	0.298	0.388	
	meta-DistMult	8046.0	0.288	0.397	
	Gain (↑)	22.0%	-3.6%	2.3%	20.7%
	RotatE	9243.9	0.350	0.428	
	meta-RotatE	8618.7	0.352	0.435	
	Gain (↑)	6.8%	0.7%	1.8%	9.3%

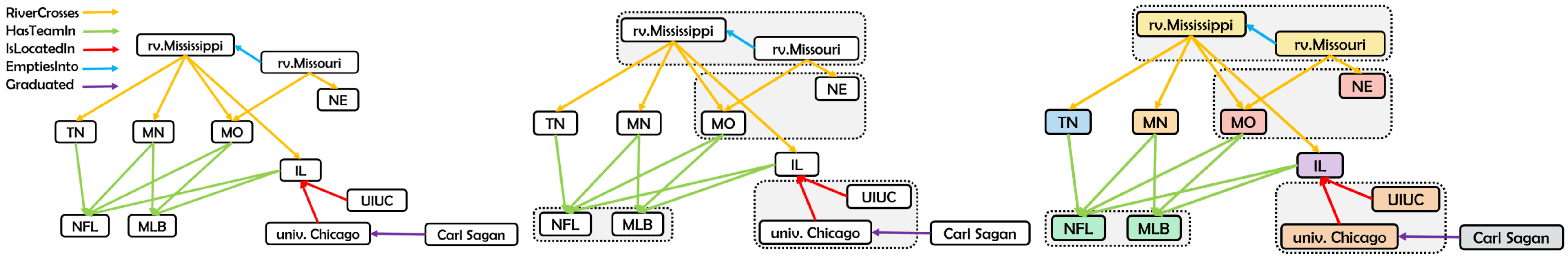
Experimental Results

- Link Prediction Results
  - A positive gain indicates that our method is better than the original model.

		MR (↓)	MRR (↑)	Hit@10 (↑)	Total Gain (↑)
WN18	TransE	210.4	0.521	0.943	
	meta-TransE	185.9	0.535	0.949	
	Gain (↑)	11.6%	2.7%	0.7%	15.0%
	DistMult	301.1	0.320	0.550	
	meta-DistMult	289.1	0.463	0.732	
	Gain (↑)	4.0%	44.7%	33.2%	81.9%
	RotatE	76.681	0.661	0.882	
	meta-RotatE	73.718	0.655	0.884	
	Gain (↑)	3.9%	-0.9%	0.2%	3.2%

Summary

- By identifying a set of structurally similar entities, group semantically close entities.
- Construct the metagraph of a knowledge graph → a pre-training method
- Empirically, our pre-training method improves the link prediction performance.
- Future work: overlapping clustering and multi-level clustering



Big Data Intelligence Lab @ KAIST

bdi-lab.kaist.ac.kr