Knowledge Graph

Zhiyuan Liu

liuzy@tsinghua.edu.cn

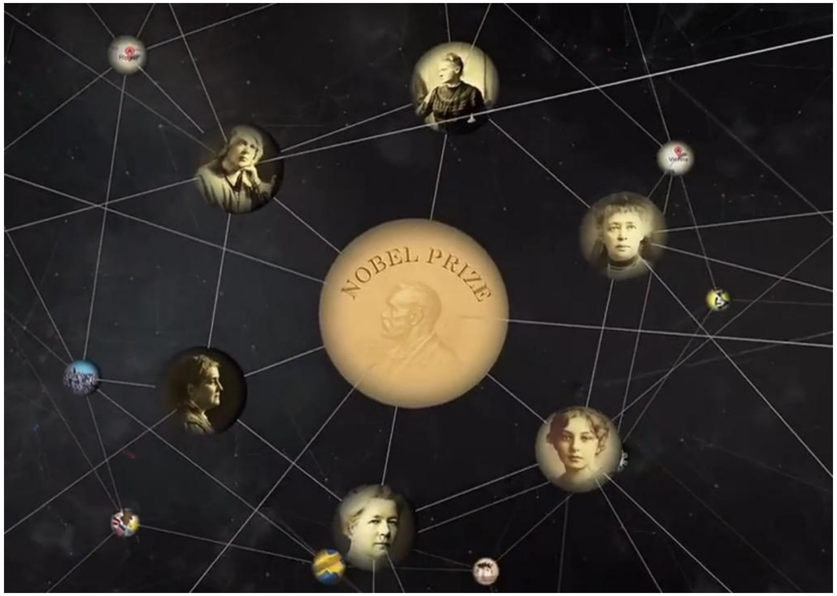
THUNLP

What Is Knowledge Graph?

Introduction

- The Knowledge Graph provides structured and detailed information about the topic in addition to a list of links to other sites

--Wikipedia



From Data to Knowledge

Data: unprocessed signals or facts

Information: interpreted and meaning attached data

Knowledge: understanding of the selected information

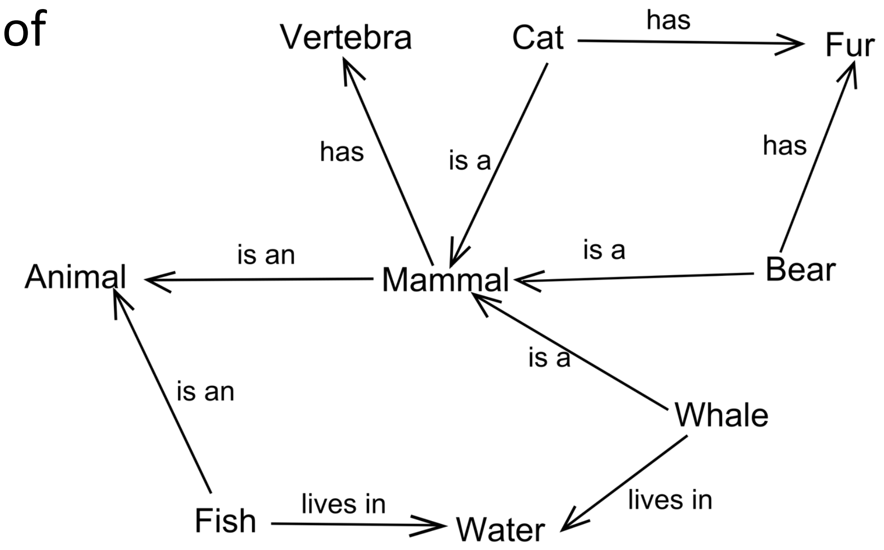
History of KG

- Knowledge to Data
  - Knowledge representation
  - Knowledge organization and storage

- Data to Knowledge
  - Acquire knowledge from data
  - Knowledgeable data representation

History of KG: Knowledge to Data

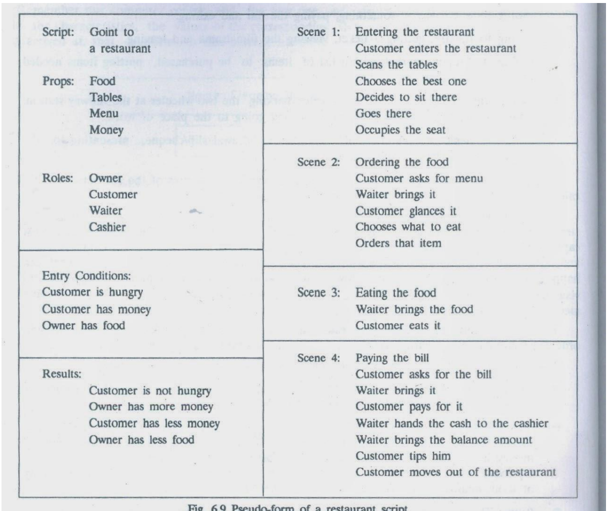
- Symbolic Logic
  - Automated reasoning
  - General Problem Solver: separation of knowledge and solver
- Semantic Network
  - Graphical representation of knowledge
  - Interlingua for machine translation



History of KG: Knowledge to Data

- Expert Systems
  - Large amount of complex data need to be managed
  - Combine logic and data
  - Examples: Japanese 5th Generation Project, MYCIN, PROLOG
- Knowledge Engineering
  - Imitate human expertise with a knowledge-based system
  - Frames and Scripts

Slot	Value	Type
ALEX	_	(This Frame)
NAME	Alex	(key value)
ISA	Boy	(parent frame)
SEX	Male	(inheritance value)
AGE	IF-NEEDED: Subtract(current,BIRTHDATE);	(procedural attachment)
HOME	100 Main St.	(instance value)
BIRTHDATE	8/4/2000	(instance value)
FAVORITE_FOOD	Spaghetti	(instance value)
CLIMBS	Trees	(instance value)
BODY_TYPE	Wiry	(instance value)
NUM_LEGS	1	(exception)


Script: Goint to a restaurant

Props: Food, Tables, Menu, Money

Roles: Owner, Customer, Waiter, Cashier

Entry Conditions:
Customer is hungry
Customer has money
Owner has food

Results:
Customer is not hungry
Owner has more money
Customer has less money
Owner has less food

Scene 1: Entering the restaurant
Customer enters the restaurant
Scans the tables
Chooses the best one
Decides to sit there
Goes there
Occupies the seat

Scene 2: Ordering the food
Customer asks for menu
Waiter brings it
Customer glances it
Chooses what to eat
Orders that item

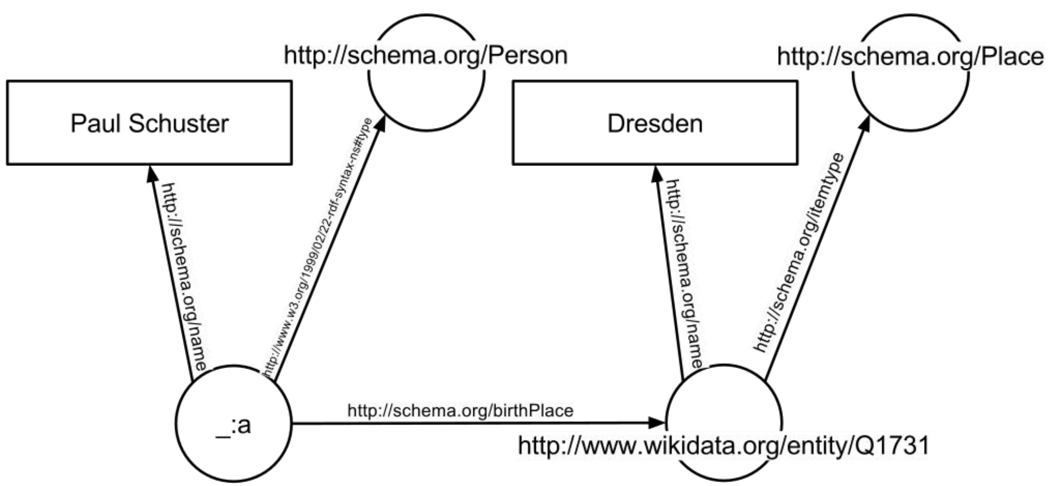
Scene 3: Eating the food
Waiter brings the food
Customer eats it

Scene 4: Paying the bill
Customer asks for the bill
Waiter brings it
Customer pays for it
Waiter hands the cash to the cashier
Waiter brings the balance amount
Customer tips him
Customer moves out of the restaurant

Fig. 6.9 Pseudo-form of a restaurant script

History of KG: Data to Knowledge

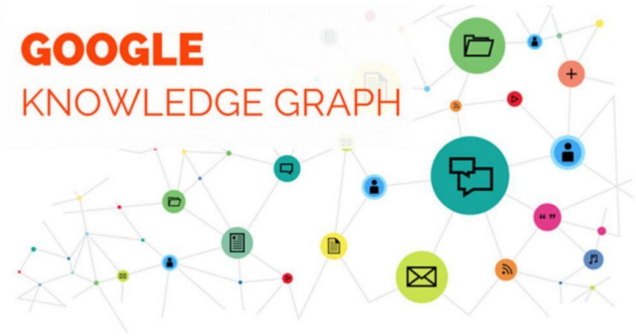
- The Web 1.0: World Wide Web
  - Global information infrastructure
  - Semi-structured data: XML, HTML, RDF
- Semantic Web
  - Semantic understanding of web resources
  - Knowledge representation: Ontology



Paul Schuster was born in Dresden

History of KG: Data to Knowledge

- The Web 2.0: Collective Intelligence
  - Open and distributed data
  - Collective knowledge engineering: Wikipedia
- Google Knowledge Graph
  - Extract and describe factual information from web
  - Things, not strings



Today's KG:

- Open Knowledge Graphs
  - DBpedia
  - YAGO
  - Freebase
  - Wikidata
- Enterprise Knowledge Graphs
  - Google knowledge graph
  - Amazon, Uber
  - Facebook, LinkedIn

Milestone

DBpedia

Wikipedia

yago select knowledge

Cyc

WordNet

HowNet

High-Quality Data Source,
5 million concepts,
muilti-lingual,
rich structural data:
Infobox, table,
list, category…

freebase

GOOGLE KNOWLEDGE GRAPH

1985

1990

2005-2010

Structural Knowledge

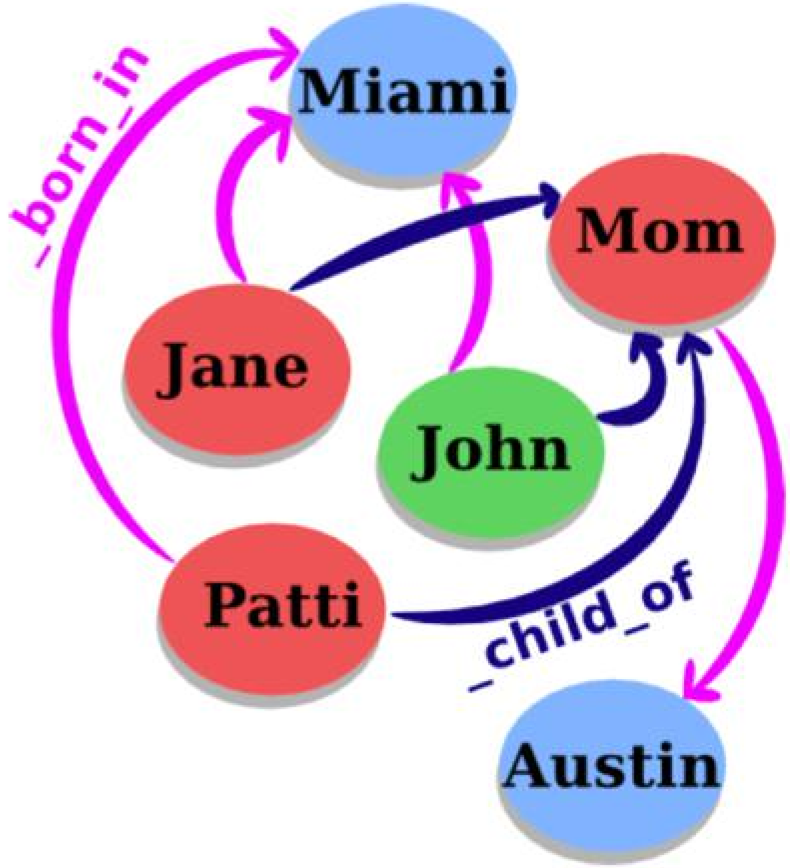
write

(***William Shakespeare***, book/author/works_written, ***Romeo and Juliet***)

head entity **relation** **tail entity**

Entities and Relations

- Knowledge structured as graph
  - Each node = an entity
  - Each edge = a relation
- Fact: (head, relation, tail)
  - head = subject entity
  - relation = relation type
  - tail = object entity



KG Features

- Graph-structured
  - A knowledge graph is organized as a graph
  - Relational, flexible and learnable
- Semantics
  - The meaning of the data is encoded alongside the data in the graph
  - Closer to natural language, human-readable

Knowledge Graph Examples

https://www.wikidata.org/

https://wordnet.princeton.edu/

https://github.com/yago-naga/yago3

https://wiki.dbpedia.org/

KG Application: Question Answering

WolframAlpha computational knowledge engine

Enter what you want to calculate or know about:

how big is China

Examples Random

Assuming "how big" is international data | Use as referring to socioeconomic data or referring to species or referring to administrative divisions instead

Assuming total area | Use population instead

Input interpretation:

China	total area

Result: Show non-metric

9.597×10^6 km^2 (square kilometers) (world rank: 4th)

Unit conversions:

9.597×10^{12} m^2 (square meters)

3.705 million mi^2 (square miles)

1.033×10^{14} ft^2 (square feet)

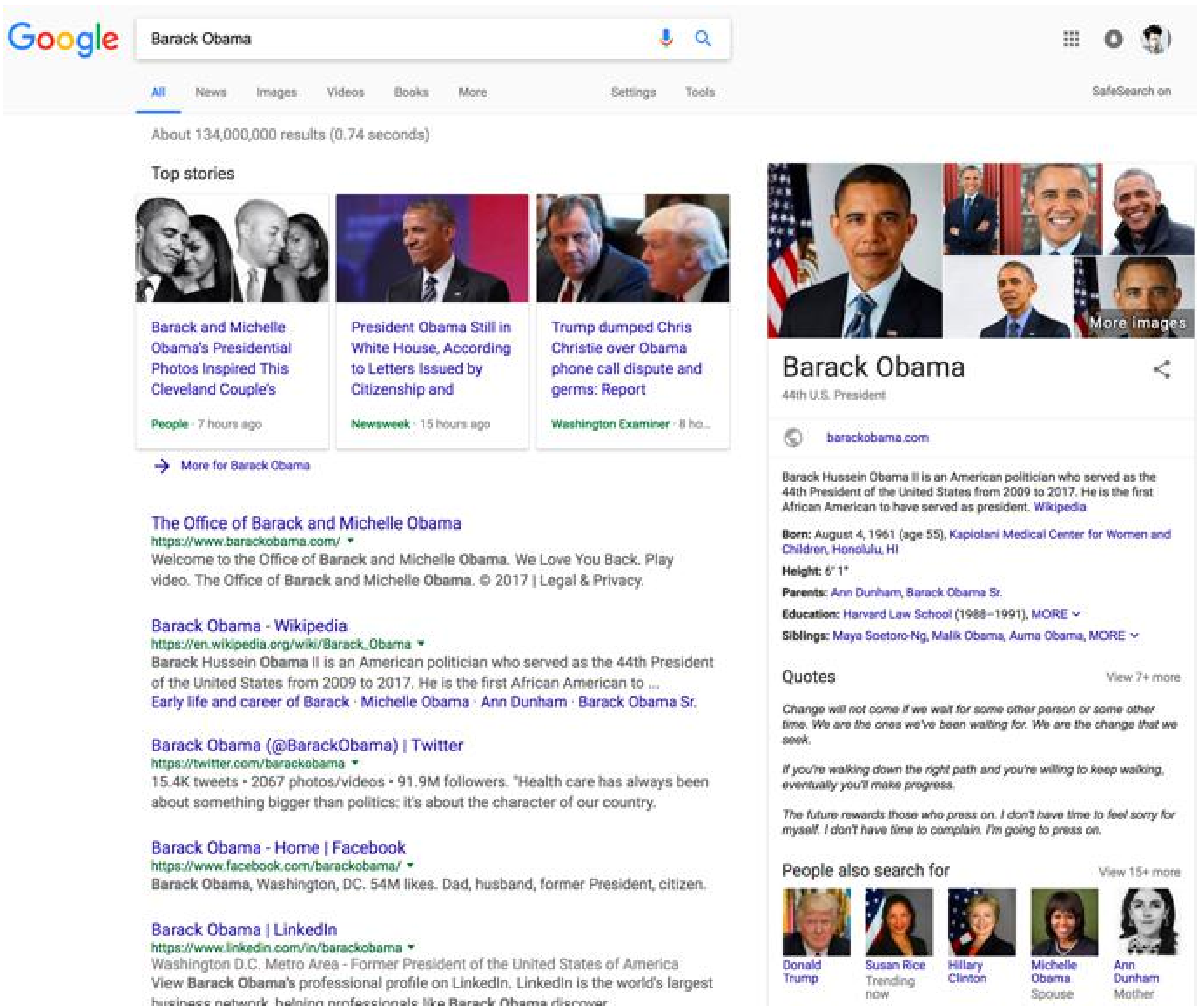
Comparisons as area:

≈ 0.96 × total area of Canada (9.98467×10^6 km^2)

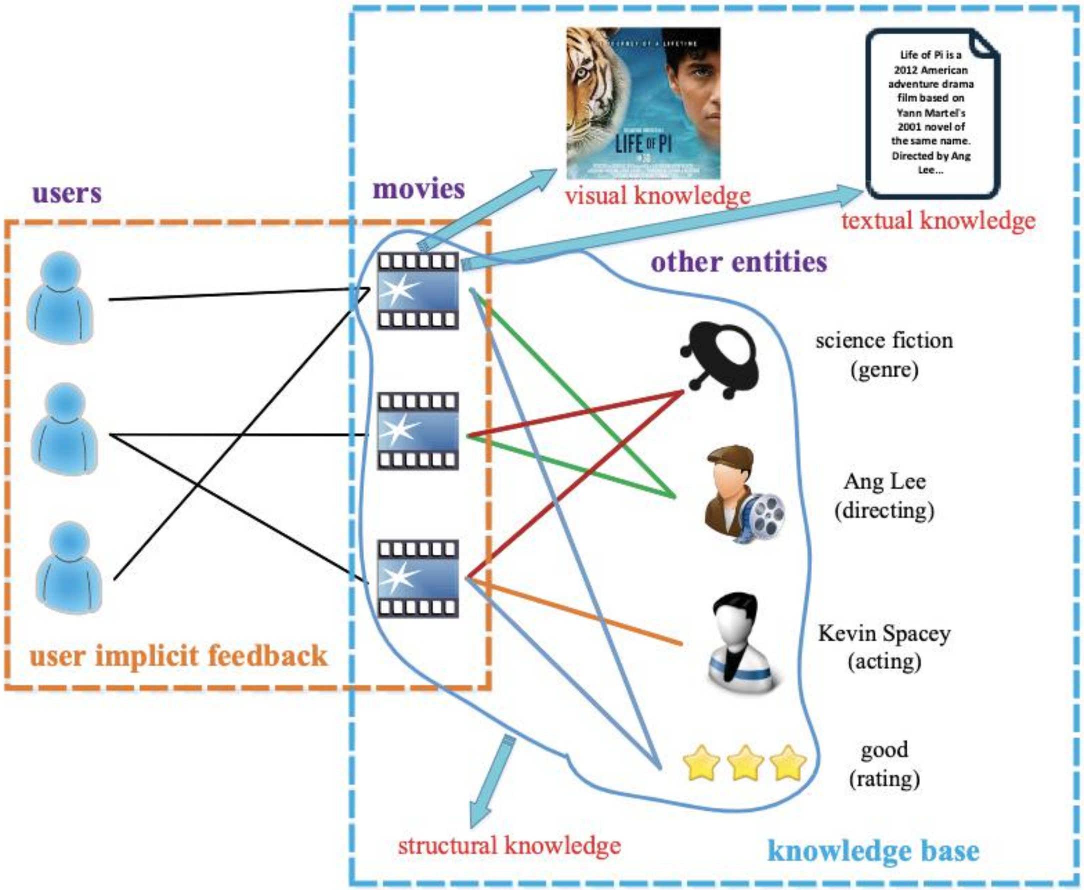
≈ 0.996 × total area of the United States (9.63142×10^6 km^2)

≈ largest extent of the Roman Empire (≈ 9 Mm^2)

KG Application: Search Engine



KG Application: Recommendation Systems



Knowledge Representation Learning

- Motivations and Aims
- Basic Models
  - Translational distance
  - Semantic matching
- Advanced Models
  - Modeling Complex Relations
  - Fusion of External Information and KG
  - Knowledge Graph Reasoning

Knowledge Representation Learning

- Motivations and Aims
- Basic Models
  - Translational distance
  - Semantic matching
- Advanced Models
  - Modeling Complex Relations
  - Fusion of External Information and KG
  - Knowledge Graph Reasoning

ML = Representation + Objective + Optimization

Yoshua Bengio. Deep Learning of Representations. AAAI 2013 Tutorial.

RL for NLP / KG

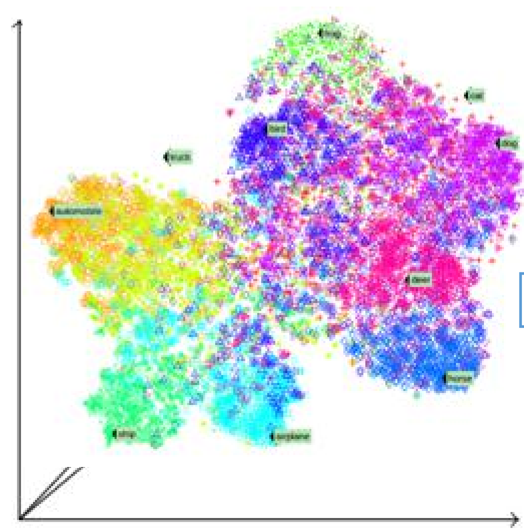
- Alleviate sparsity issues in large-scale NLP
- Enable knowledge transfer across domains and objects

Knowledge

Document

Sentence

Word



Unified Semantic Space

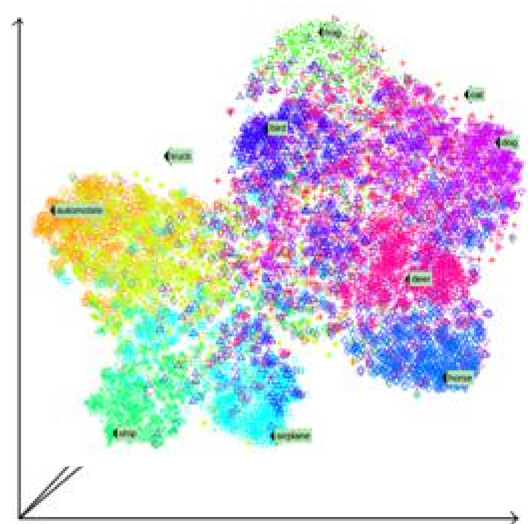
Semantic Parsing

Syntactic Parsing

Lexical Analysis

KRL

- Typical representations for KG
    - Symbolic triples (RDF)
    - Cannot efficiently measure semantic relations of entities
- How: Encode KGs into low-dimensional vector spaces



Knowledge Representation Learning

- Motivations and Aims
- Basic Models
  - Translational distance
  - Semantic matching
- Advanced Models
  - Modeling Complex Relations
  - Fusion of External Information and KG
  - Knowledge Graph Reasoning

KRL: Scoring function

- How to measure the plausibility of knowledge embeddings?
- Translational distance
  - Inspired by word2vec

$$C(king) - C(man) \approx C(queen) - C(woman)$$

$$\Downarrow$$

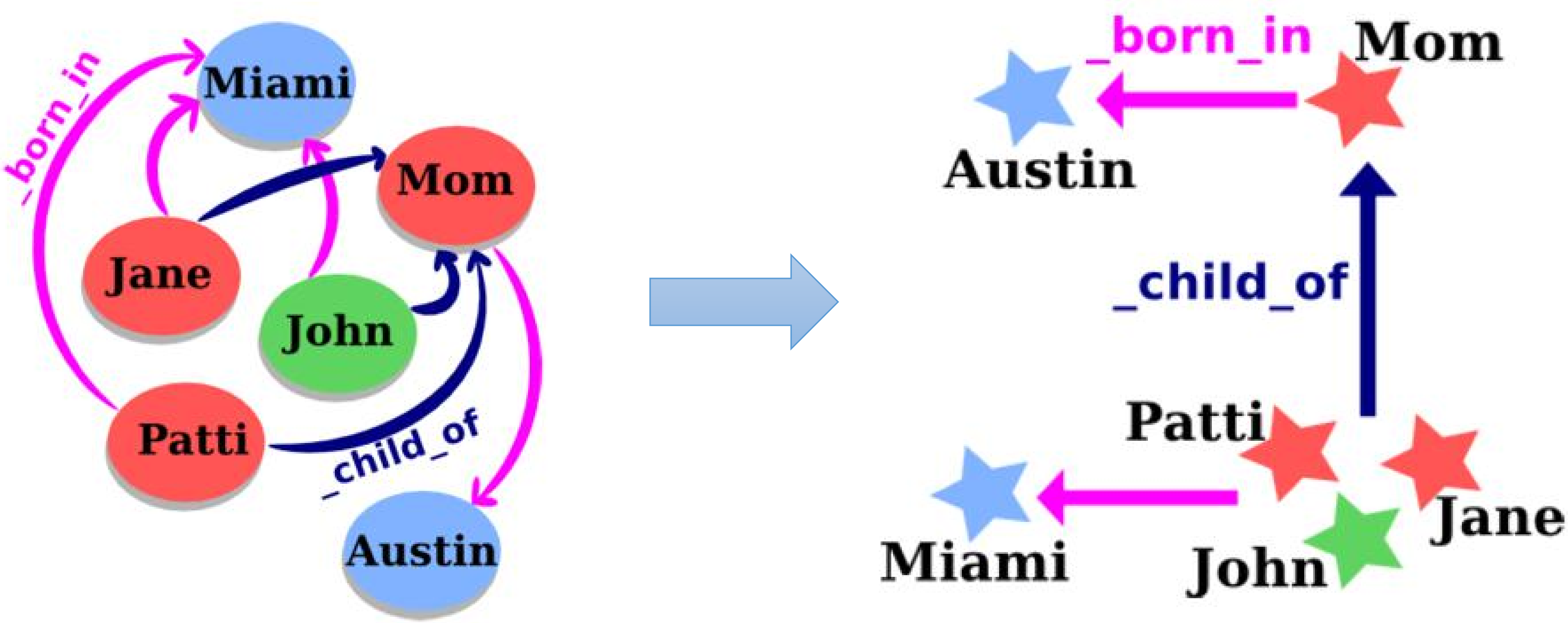
$$C(China) - C(Beijing) \approx C(France) - C(Paris)$$
$$\approx C(CapitalOf)$$

- Semantic matching
  - Transform head entity near the tail in the representation space

$$h^{\mathrm{T}} M_r \approx t^{\mathrm{T}}$$

Translational: TransE

- For each triple (head, relation, tail), relation serves as a translation from head to tail

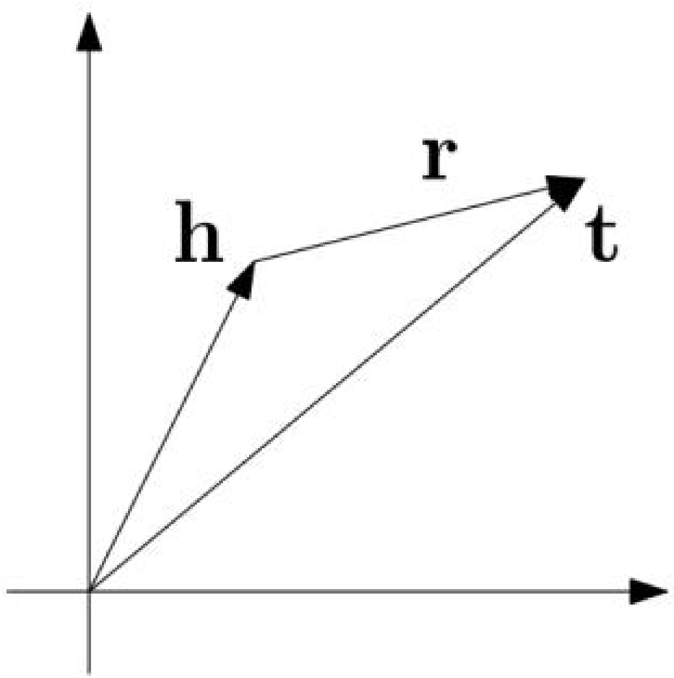


Learning objective: **h + r = t**

Translating embeddings for modeling multi-relational data. Bordes, et al. NIPS. 2013.

TransE

- For each triple (head, relation, tail), relation serves as a translation from head to tail



Learning objective: **h + r = t**

Objective Function

- Energy Function
  - For correct (h, r, t)， requires **h** + **r** = **t**

$$f(h,\ r,\ t) = \|\boldsymbol{h} + \boldsymbol{r} - \boldsymbol{t}\|$$

Objective Function

Energy Function $$f(h,\ r,\ t) = \|\boldsymbol{h} + \boldsymbol{r} - \boldsymbol{t}\|$$

Objective Function $$\sum_{(h,r,t)\in\Delta}\ \sum_{(h',r,t')\in\Delta'} [\gamma + f(h,r,t) - f(h',r,t')]_+$$

where $$[x]_+ = \max(0,x)$$

w.r.t $$\|\boldsymbol{h}\| \leq 1,\ \|\boldsymbol{t}\| \leq \boldsymbol{1}$$

Δ triple sets in KG

Δ' negative triple sets not in KG

Entity Prediction

WALL-E _has_genre ?

h + r = ?

Entity Prediction

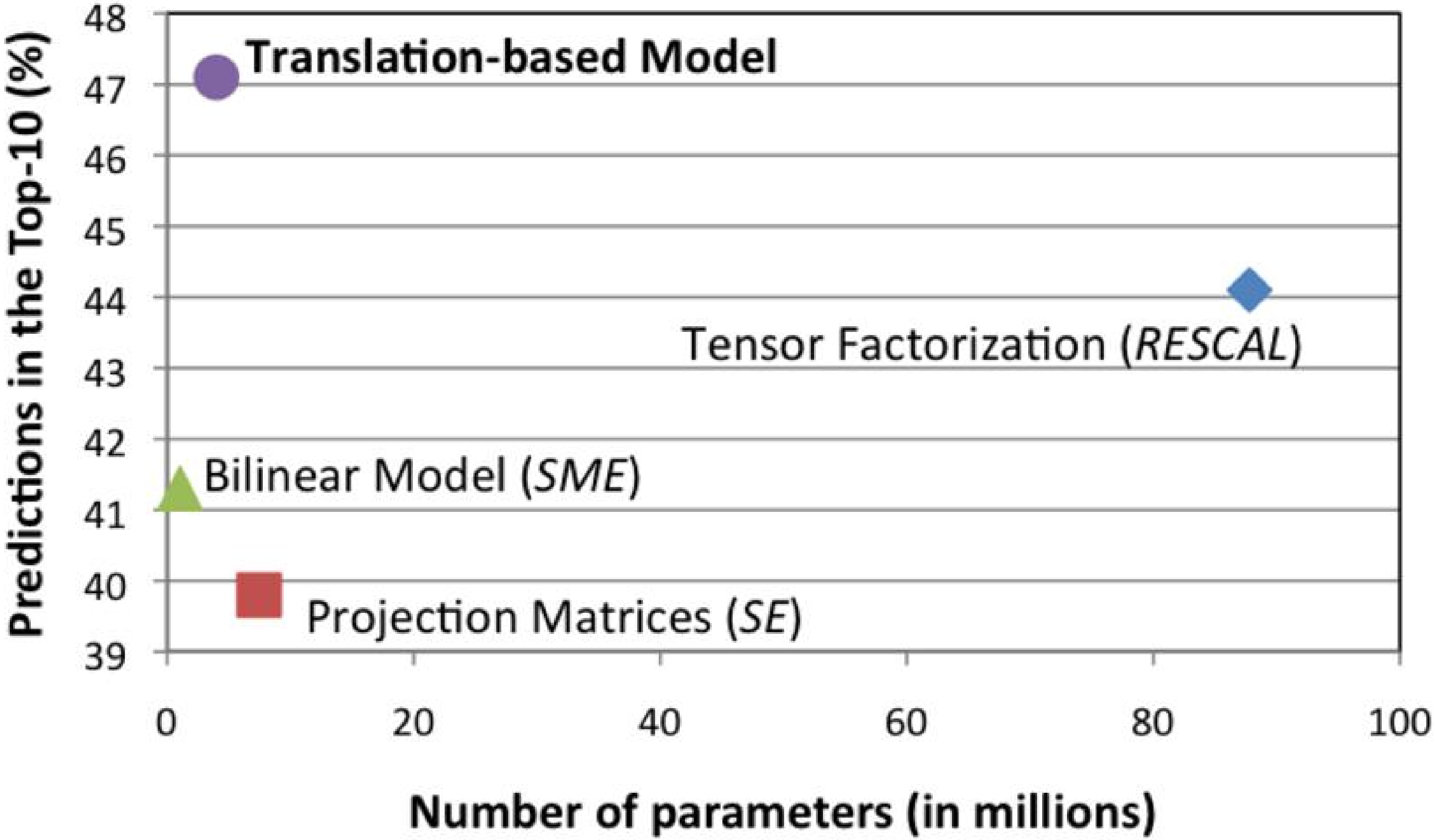
WALL-E _has_genre

Animation
Computer animation
Comedy film
Adventure film
Science Fiction
Fantasy
Stop motion
Satire
Drama
Connecting

h + r = t

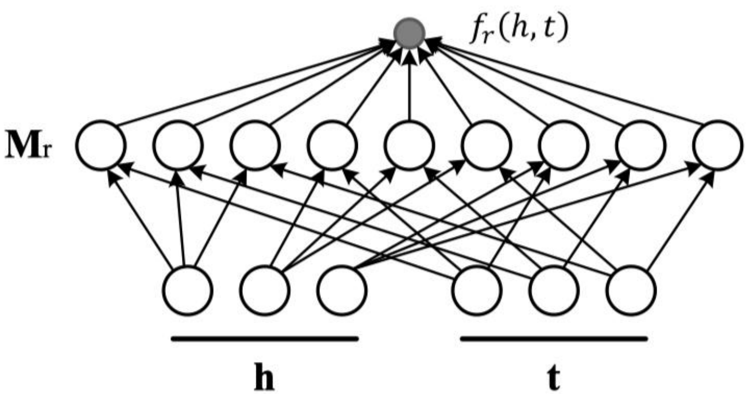
Prediction Performance

- Dataset: Freebase15K

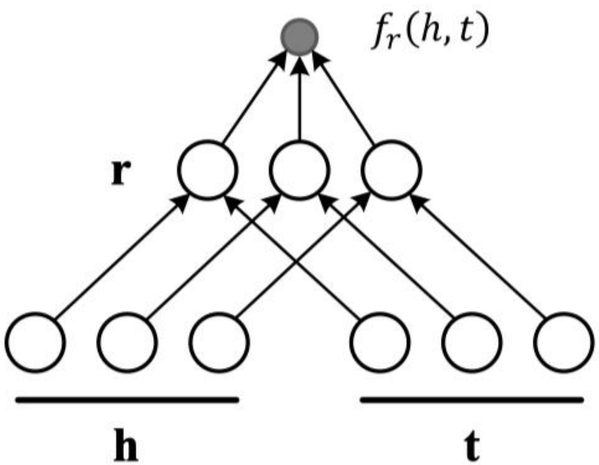


Semantic Matching

- **RESCAL**: Bilinear scoring function
- **DistMult**: Elementwise multiplication scoring function



$$f_r(h,t) = \mathbf{h}^\top \mathbf{M}_r \mathbf{t}$$



$$f_r(h,t) = \mathbf{h}^\top \mathrm{diag}(\mathbf{r})\mathbf{t}$$

A three-way model for collective learning on multi-relational data. Nickel, et al. ICML. 2011.
Embedding entities and relations for learning and inference in knowledge bases. Yang, et al. ICLR. 2015.

Key Challenges in KRL

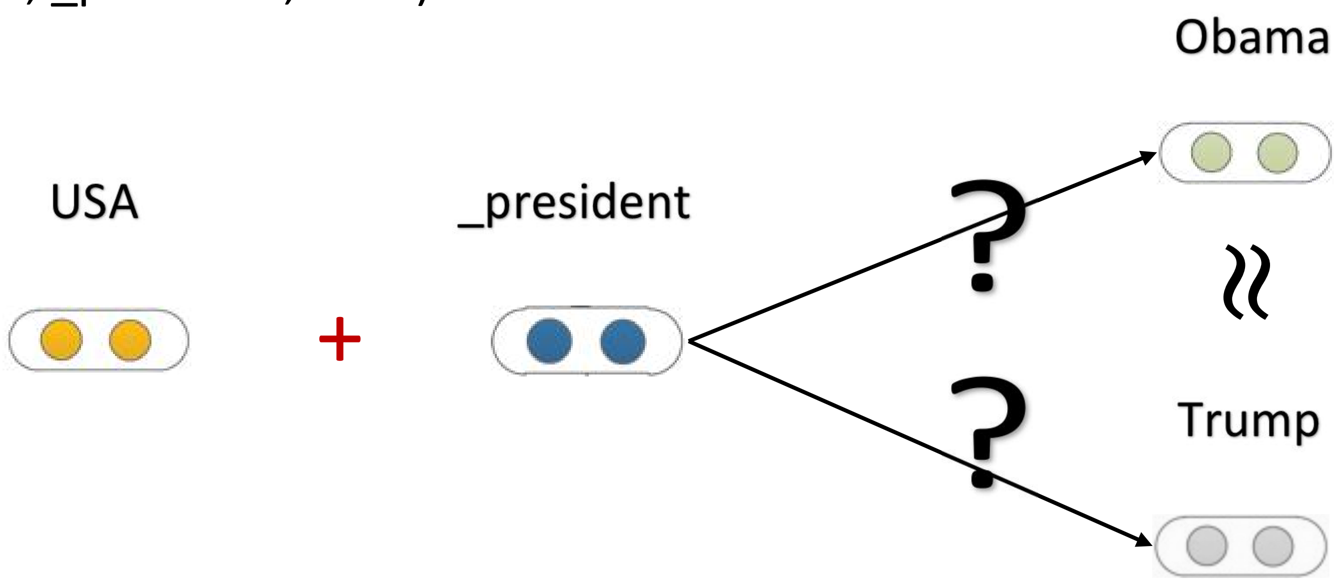
- Modeling Complex Relations
- Fusion of External Information and KG
- Knowledge Graph Reasoning

Key Challenges in KRL

- Modeling Complex Relations
- Fusion of External Information and KG
- Knowledge Graph Reasoning

Complex Relations

- 1-to-N, N-to-1, N-to-N relations
  - (USA, _president, Obama)
  - (USA, _president, Bush)



Key Challenges in KRL

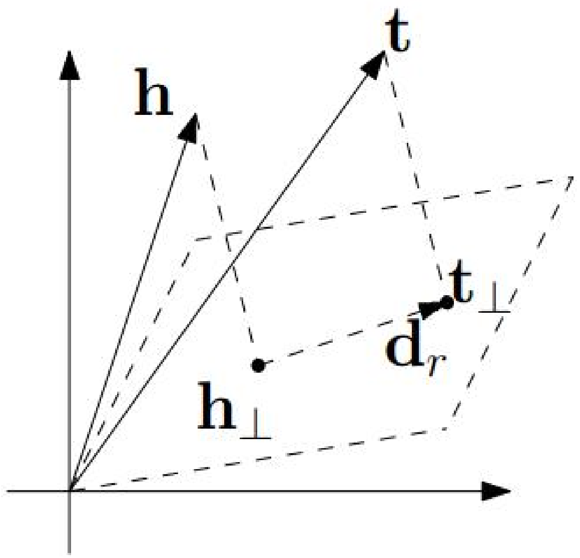
- Modeling Complex Relations
  - Projection
  - Embedding Space
    - Complex Space
  - Encoding Models
    - MLP
    - CNN
    - Transformer
    - GNN

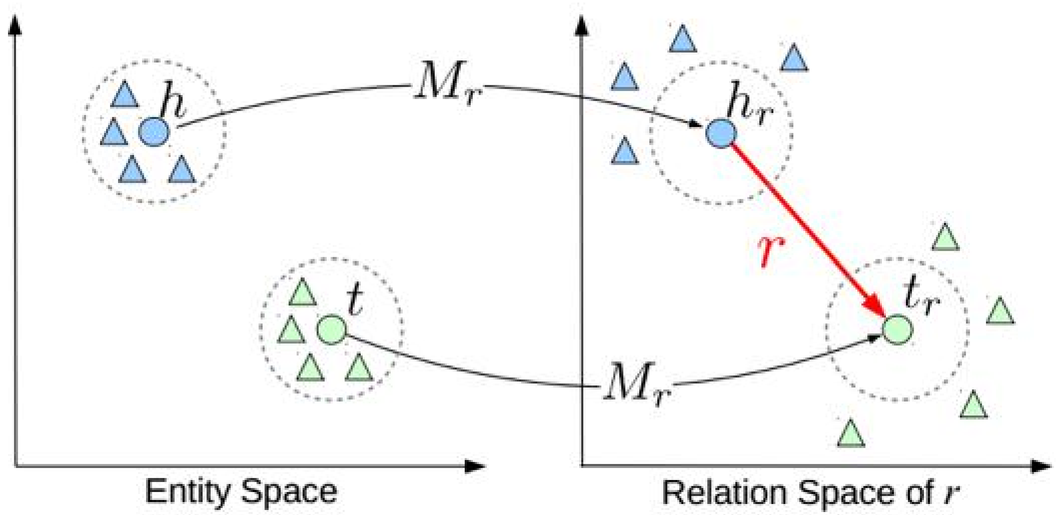
Key Challenges in KRL

- Modeling Complex Relations
  - Projection
  - Embedding Space
    - Complex Space
  - Encoding Models
    - MLP
    - CNN
    - Transformer
    - GNN

Projection

- Build relation-specific entity embeddings





TransH

$$f_r(h,t) = -\left\|\left(\mathbf{h} - \mathbf{w}_r^\top \mathbf{h} \mathbf{w}_r\right) + \mathbf{r} - \left(\mathbf{t} - \mathbf{w}_r^\top \mathbf{t} \mathbf{w}_r\right)\right\|_2^2$$

TransR

$$\mathbf{h}_r = \mathbf{h}\mathbf{M}_r, \quad \mathbf{t}_r = \mathbf{t}\mathbf{M}_r$$
$$f_r(h,t) = |\mathbf{h}_r + \mathbf{r} - \mathbf{t}_r|_{L1/L2}$$

Knowledge graph embedding by translating on hyperplanes. Wang, et al. AAAI. 2014.
Learning entity and relation embeddings for knowledge graph completion. Lin, et al. AAAI. 2015.

Examples

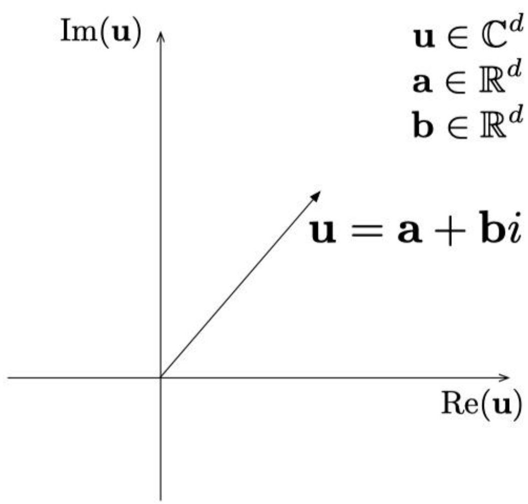
Head	Titanic		
Relation	**/film/film/genre**		
Model	TransE	TransH	TransR
1	War_film	Drama	Costume_drama
2	Period_piece	Romance_Film	Drama
3	Drama	Costume_drama	Romance_Film
4	History	Film_adaptation	Period_piece
5	Biography	Period_piece	Epic_film
6	Film_adaptation	Adventure_Film	Adventure_Film
7	Adventure_Film	LGBT	LGBT
8	Action_Film	Existentialism	Film_adaptation
9	Political_drama	Epic_film	Existentialism
10	Costume_drama	War_film	War_film

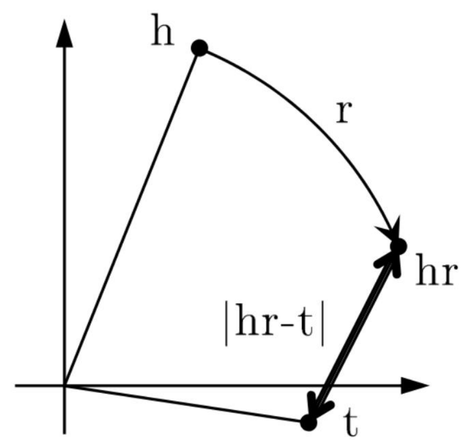
Key Challenges in KRL

- Modeling Complex Relations
  - Projection
  - Embedding Space
    - Complex Space
  - Encoding Models
    - MLP
    - CNN
    - Transformer
    - GNN

Complex Space

- **ComplEx**: Represent entities and relations in a complex space instead of real-value space
- **RotatE**: Model each relation as a rotation from the head entity to the tail entity in the complex space





$$f_r(h,t) = \mathrm{Re}\left(\mathbf{h}^\top \operatorname{diag}(\mathbf{r})\bar{\mathbf{t}}\right) = \mathrm{Re}\left(\sum_{i=0}^{d-1} [\mathbf{r}]_i \cdot [\mathbf{h}]_i \cdot [\bar{\mathbf{t}}]_i\right)$$

$$f_r(h,t) = |\mathbf{h} \circ \mathbf{r} - \mathbf{t}|_{L1/L2}$$

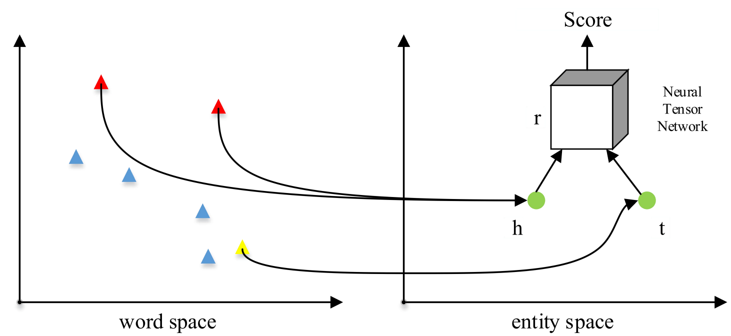
Complex embeddings for simple link prediction. Trouillon, et al. ICML. 2016.
RotatE: Knowledge graph embedding by relational rotation in complex space. Sun, et al. ICLR. 2019.

Key Challenges in KRL

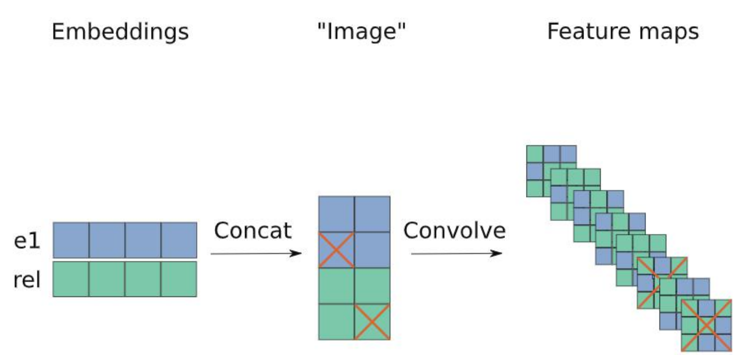
- Modeling Complex Relations
  - Projection
  - Embedding Space
    - Complex Space
  - Encoding Models
    - MLP
    - CNN
    - Transformer
    - GNN

MLP & CNN

- **NTN**: Model KGs with a Neural Tensor Network and represent entities via word vectors



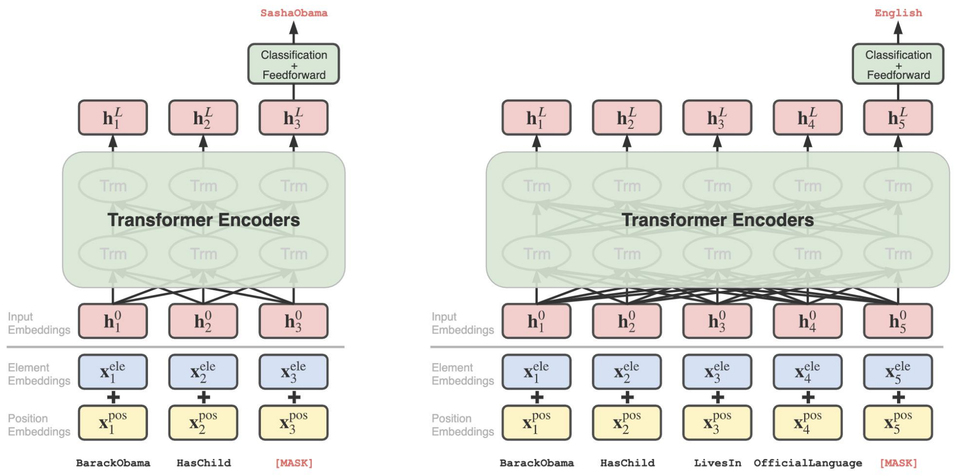
- **ConvE**: Use 2D convolution to model the interactions between entities and relations

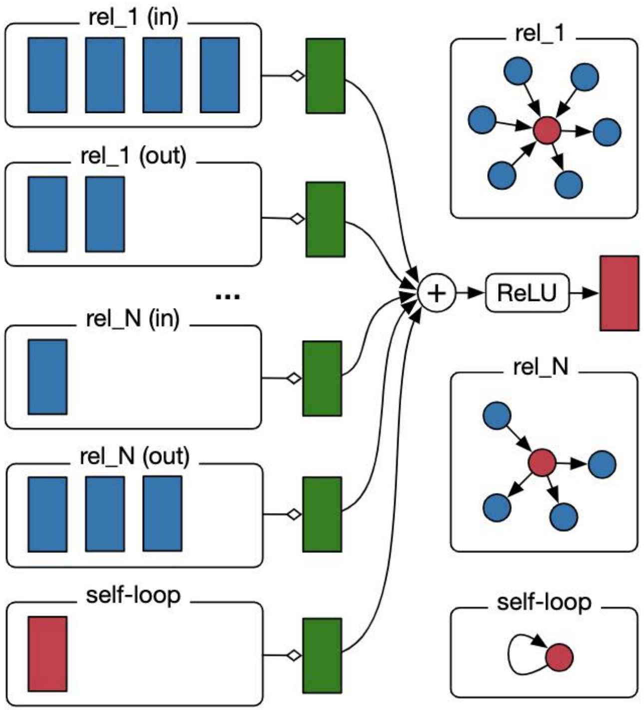


Reasoning with neural tensor networks for knowledge base completion. Socher, et al. NIPS. 2013.
Convolutional 2d knowledge graph embeddings. Dettmers, et al. AAAI. 2018.

Transformer & GNN

- **CoKE**: Employ Transformer encoders to utilize contextual information
- **R-GCN**: Relation-specific GNN

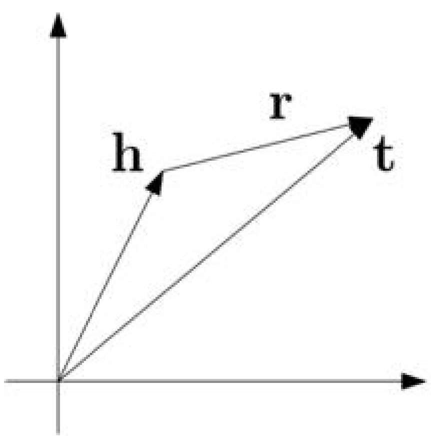


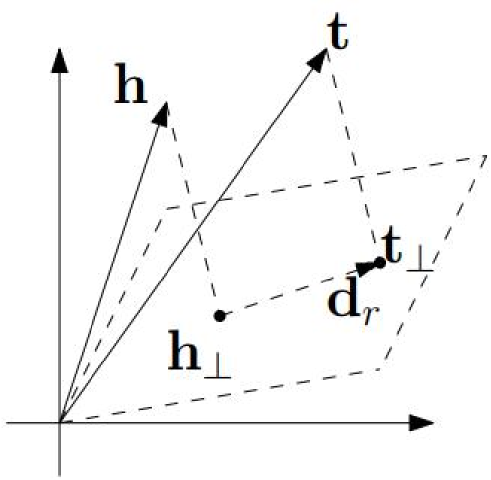


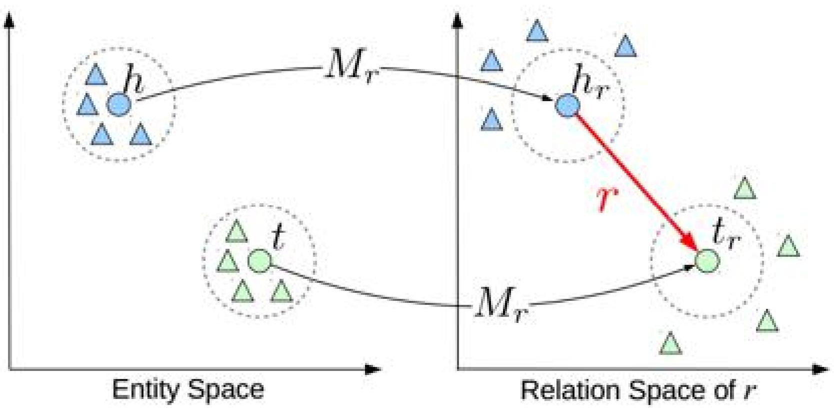
CoKE: Contextualized knowledge graph embedding. Wang, et al. Arxiv. 2019.
Modeling relational data with graph convolutional networks. Schlichtkrull, et al. ESWC. 2018.

Summary

- Basic models are too simple to handle complex relations well
  - 1-N, N-1, N-N
  - TransH, TransR, ComplEx, RotatE, NTN, ConvE, CoKE, R-GCN





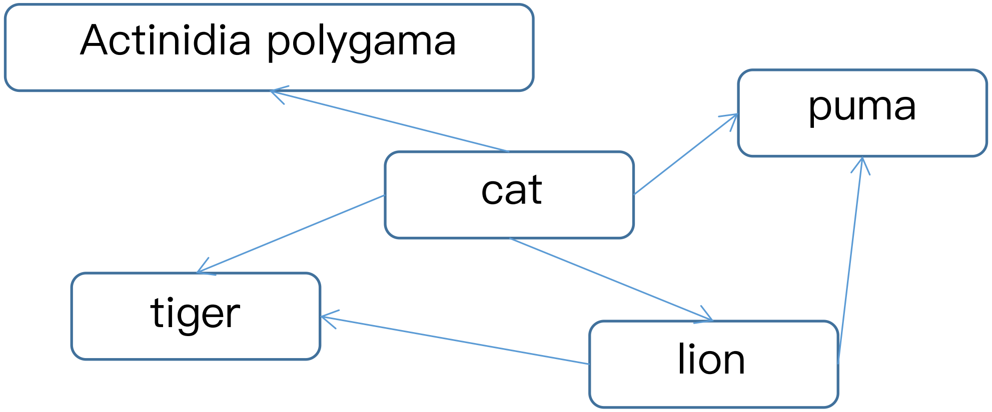


Key Challenges in KRL

- Modeling Complex Relations
- Fusion of External Information and KG
- Knowledge Graph Reasoning

Any Other Information We Can Use?

Complicated World



The domestic cat (Latin: Felis catus) or the feral cat (Latin: Felis silvestris catus) is a small, typically furry, carnivorous mammal. They are often called house cats when kept as indoor pets or simply cats when there is no need to distinguish them from other felids and felines

Three Types of Information

- Text Information
  - Construct representation with the corresponding entity descriptions, entity text
- Structure Information
  - Construct representation from the corresponding entity hierarchical structure
- Image Information
  - Construct representation from the corresponding entity images

Entity Descriptions

- KG contains rich information besides network structure

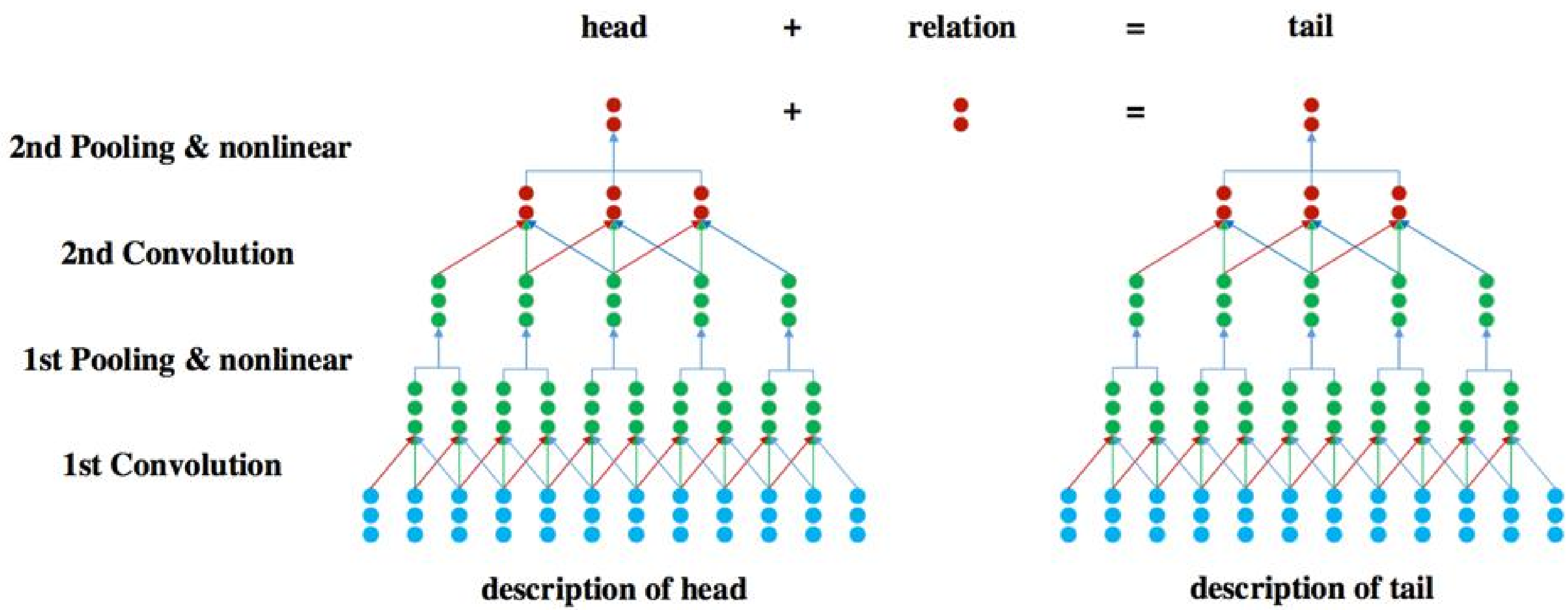
(***William Shakespeare***, book/author/works_written, ***Romeo and Juliet***)

William Shakespeare was an English poet, playwright, and actor, widely regarded as the greatest writer in the English language and the world's pre-eminent dramatist. ...

Romeo and Juliet is a tragedy written by William Shakespeare early in his career about two young star-crossed lovers whose deaths ultimately reconcile their feuding families. ...

DKRL

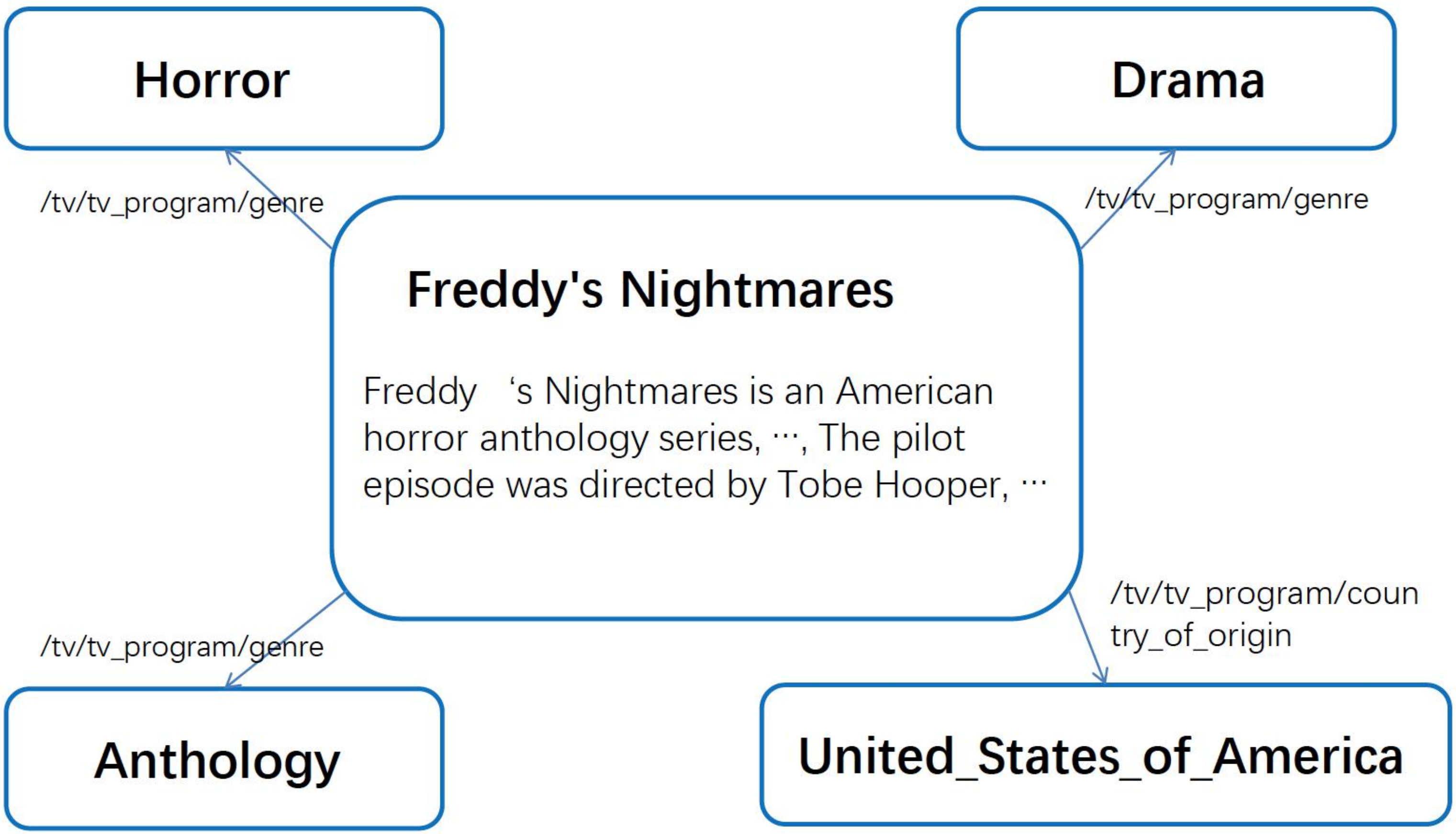
- Enhance entity representations with descriptions
- Model descriptions with CNN



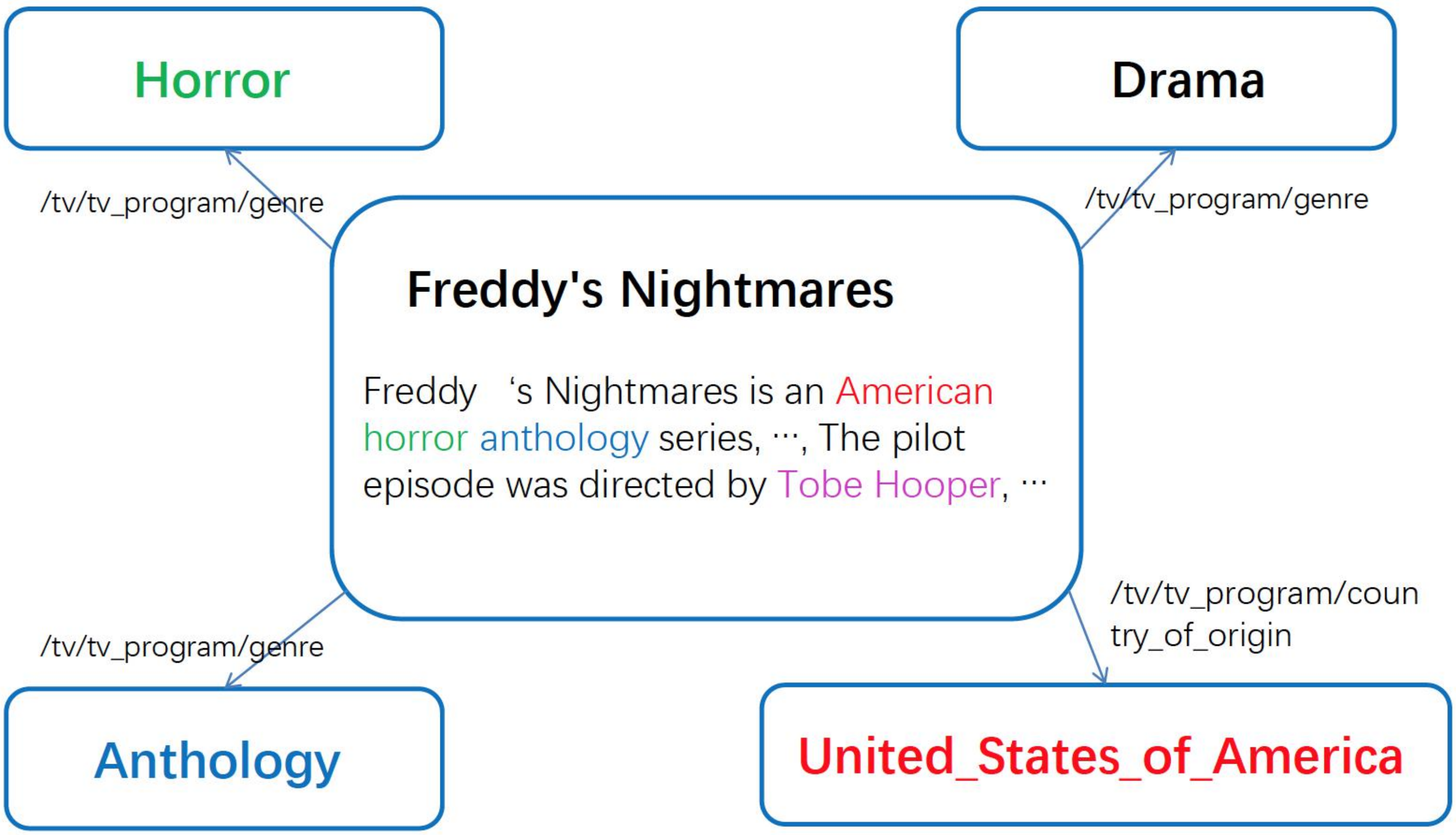
$$f_r(h,t) = -\|\mathbf{h}_s + \mathbf{r} - \mathbf{t}_s\|_1 - \|\mathbf{h}_d + \mathbf{r} - \mathbf{t}_d\|_1 - \|\mathbf{h}_s + \mathbf{r} - \mathbf{t}_d\|_1 - \|\mathbf{h}_d + \mathbf{r} - \mathbf{t}_s\|_1$$

Representation Learning of Knowledge Graphs with Entity Descriptions. Xie, et al. AAAI. 2016.

Examples


Horror
Drama
/tv/tv_program/genre
/tv/tv_program/genre
Freddy's Nightmares
Freddy 's Nightmares is an American horror anthology series, ···, The pilot episode was directed by Tobe Hooper, ···
/tv/tv_program/genre
/tv/tv_program/country_of_origin
Anthology
United_States_of_America

Examples



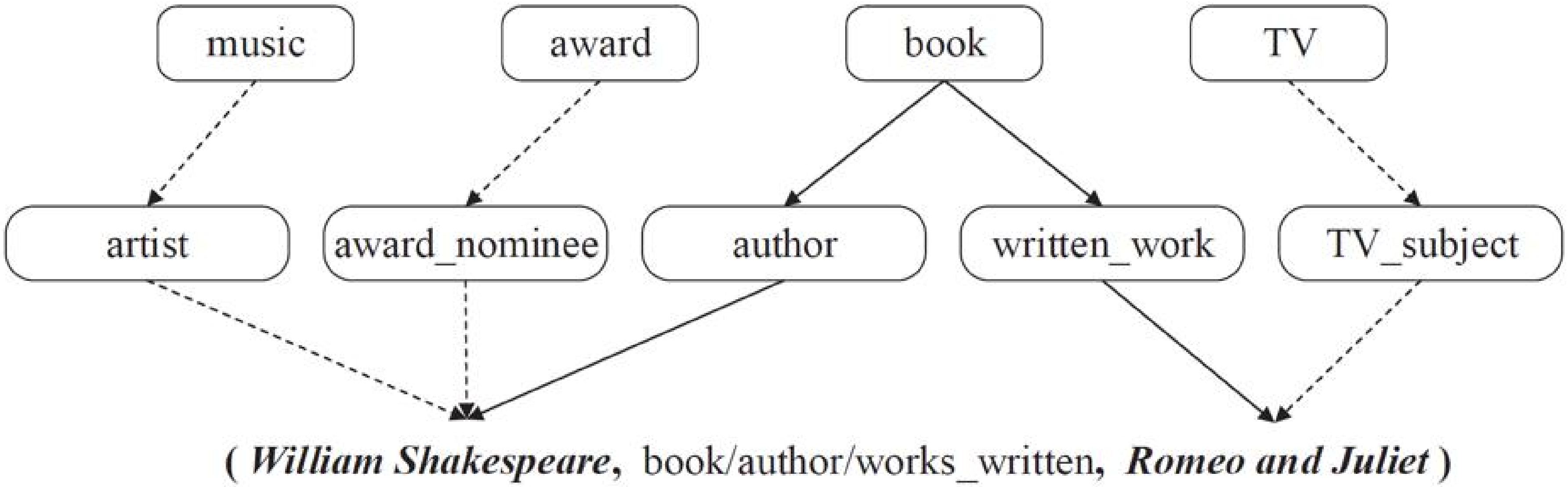
Horror
Drama
/tv/tv_program/genre
/tv/tv_program/genre
Freddy's Nightmares
Freddy 's Nightmares is an American horror anthology series, ⋯, The pilot episode was directed by Tobe Hooper, ⋯
/tv/tv_program/genre
/tv/tv_program/country_of_origin
Anthology
United_States_of_America

Three Types of Information

- Text Information
  - Construct representation with the corresponding entity descriptions, entity text
- Structure Information
  - Construct representation from the corresponding entity hierarchical structure
- Image Information
  - Construct representation from the corresponding entity images

TKRL

- Entities possessing multiple types should have various representations in different scenarios



Representation Learning of Knowledge Graphs with Hierarchical Types. Xie, et al. IJCAI. 2016.

Enhanced Energy Function

- Represent both head and tail under type-specific projection

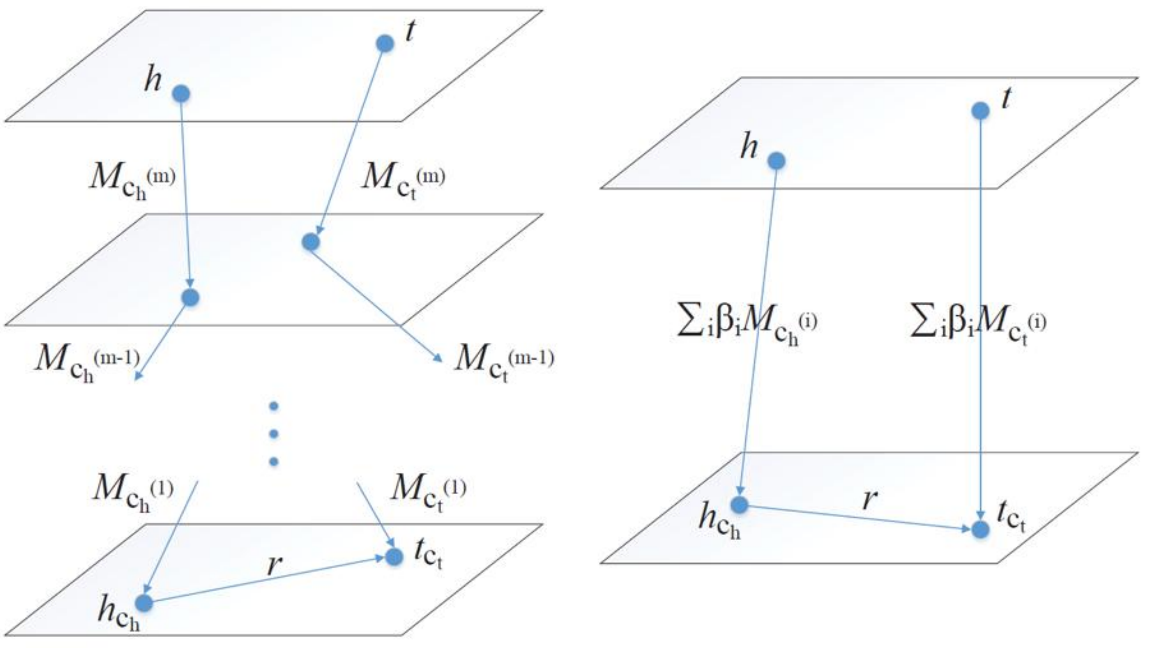
$$E(h, r, t) = ||\mathbf{h} + \mathbf{r} - \mathbf{t}||$$

$$\Downarrow$$

$$E(h, r, t) = ||\mathbf{M}_{rh}\mathbf{h} + \mathbf{r} - \mathbf{M}_{rt}\mathbf{t}||$$

Hierarchical Type Encoders

- Recursive Hierarchy Encoder (RHE)
- Weighted Hierarchy Encoder (WHE)



(a) RHE (b) WHE

Results on Long-Tail Distribution

- Hierarchical types help the performance on those long-tail entities
- Prior type knowledge is critical for sparse data

Relation Frequency	Test Number	Hits@10 for Entity (%)			Hits@1 for Relation (%)		
		TransE	TransR	TKRL (WHE)	TransE	TransR	TKRL (WHE)
$f_r <= 10$	1,444	28.0	32.4 (+4.4)	38.1 (**+10.1**)	13.2	17.0 (+3.8)	21.5 (**+8.3**)
$f_r <= 100$	4,763	49.9	54.5 (+4.6)	57.9 (**+8.0**)	45.7	50.5 (+4.8)	54.3 (**+8.6**)
$f_r <= 1000$	18,296	66.1	69.1 (+3.0)	71.6 (**+5.5**)	70.9	75.4 (+4.5)	77.8 (**+6.9**)
total	62,374	61.9	67.2 (+5.3)	69.2 (**+7.3**)	80.4	88.8 (+8.4)	89.7 (**+9.3**)

KR-EAR

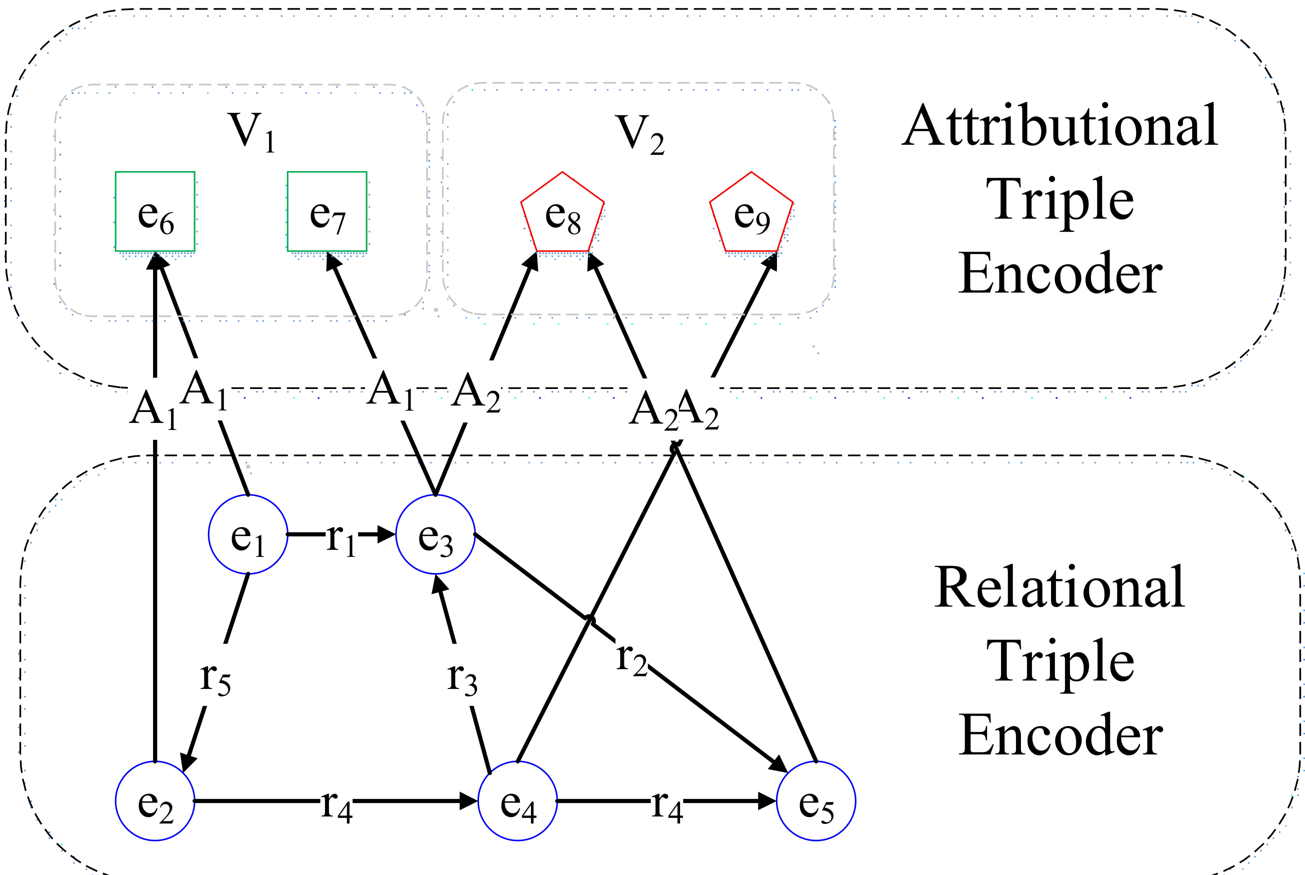
- Relational Facts in Knowledge Graph
  - Entity Attributes
  - Relations between Entities

Relation Type	Relation	E_t	E_h
Attributes	nationality	1.05	1,551.90
	gender	1.00	637,333.33
	ethnicity	1.12	41.52
	religion	1.09	107.40
Relations	parents	1.58	1.67
	capital	1.29	1.42
	author	1.02	2.17
	founder	1.37	1.31

Knowledge Representation Learning with Entities, Attributes and Relations. Lin, et al. IJCAI. 2016.

KR-EAR

- Relational Triple Encoder
- Attributional Triple Encoder

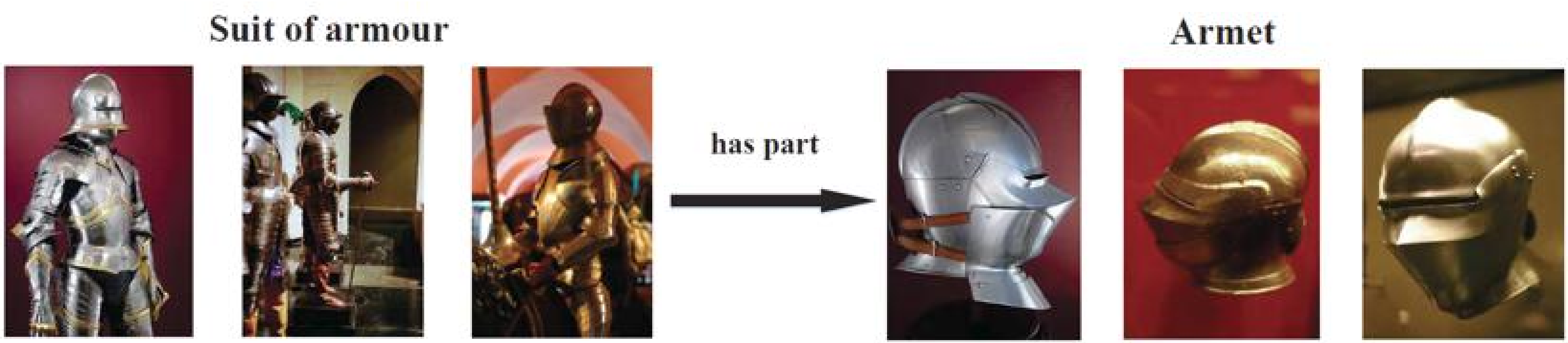


Three Types of Information

- Text Information
  - Construct representation with the corresponding entity descriptions, entity text
- Structure Information
  - Construct representation from the corresponding entity hierarchical structure
- Image Information
  - Construct representation from the corresponding entity images

Visual Information in Images

- Images provide significant visual information that intuitively describes the appearances and behaviors of entities

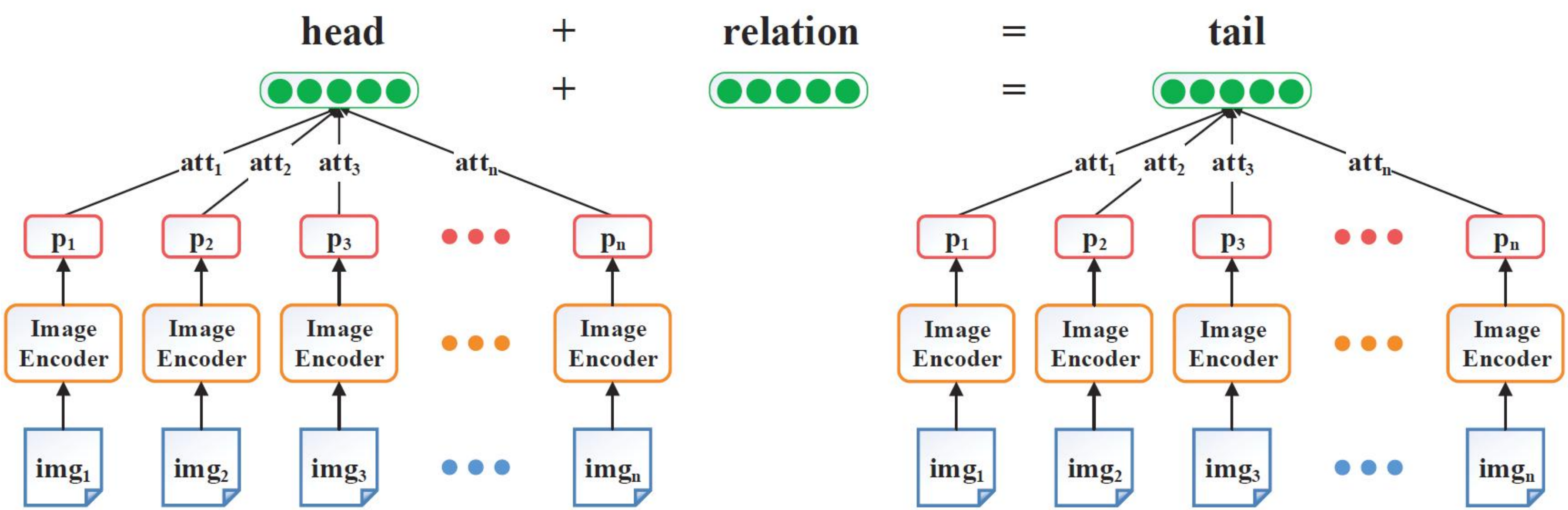


(*Suit of armour*, has_part, *Armet*)

Image-embodied Knowledge Representation Learning. Xie, et al. IJCAI. 2017.

IKRL

- Image encoder extracts informative features
- Attention module combines visual representations

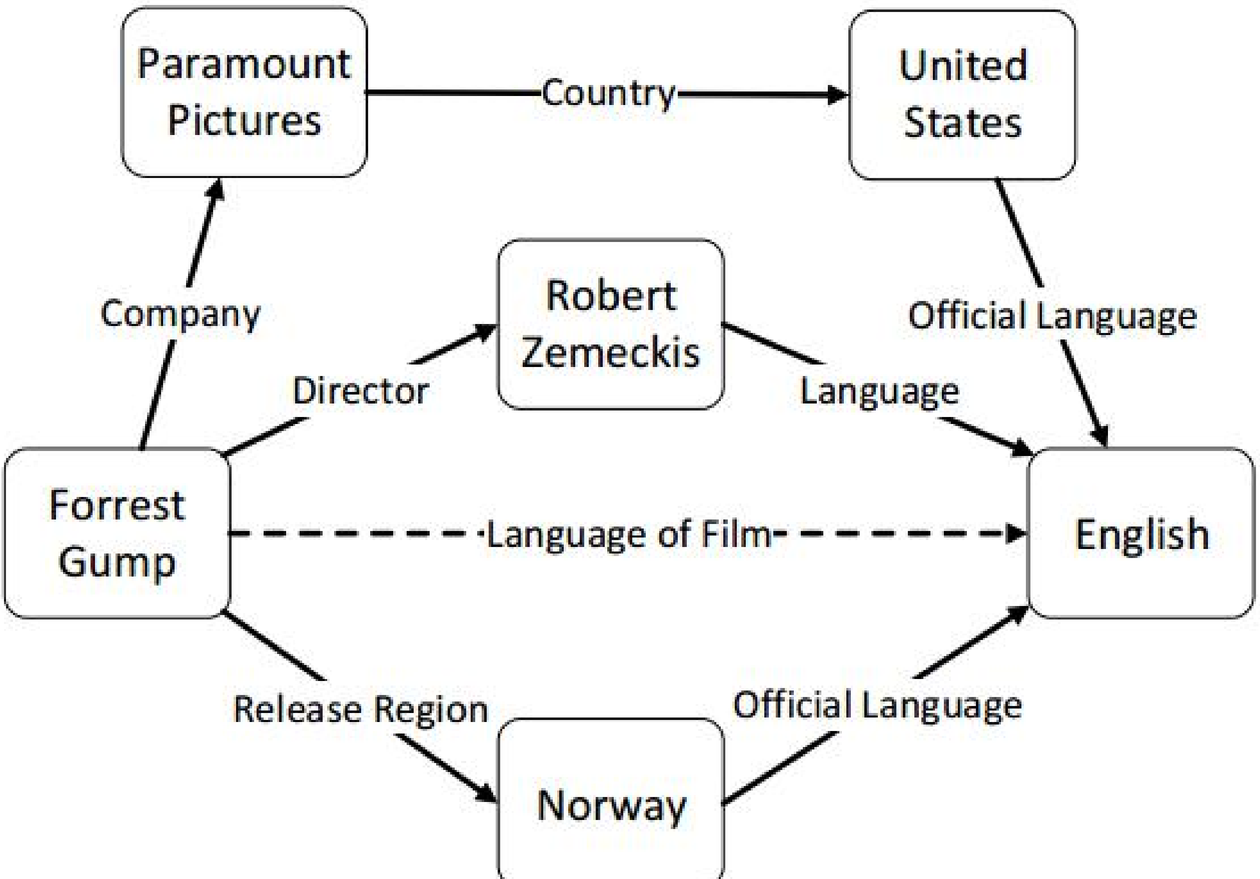


Key Challenges in KRL

- Modeling Complex Relations
- Fusion of External Information and KG
- Knowledge Graph Reasoning

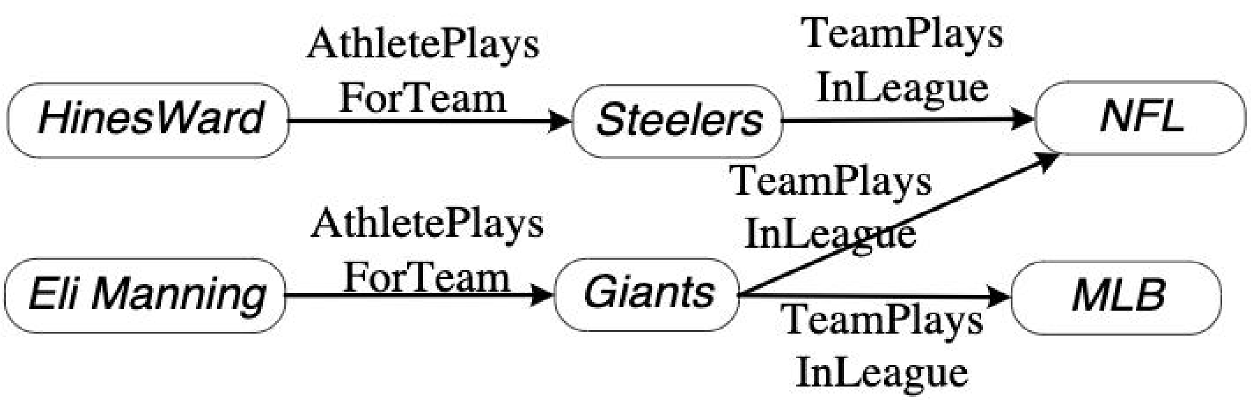
Relational Paths

- Complex inference patterns between relations



Path-based Methods

- Path-Ranking Algorithm
  - Run random walk with restarts to derive many paths
  - Use supervised training to rank paths
- Interpretable but not scalable



Random Walk Inference and Learning in A Large Scale Knowledge Base. Lao, et al. EMNLP. 2011.

Path-based TransE

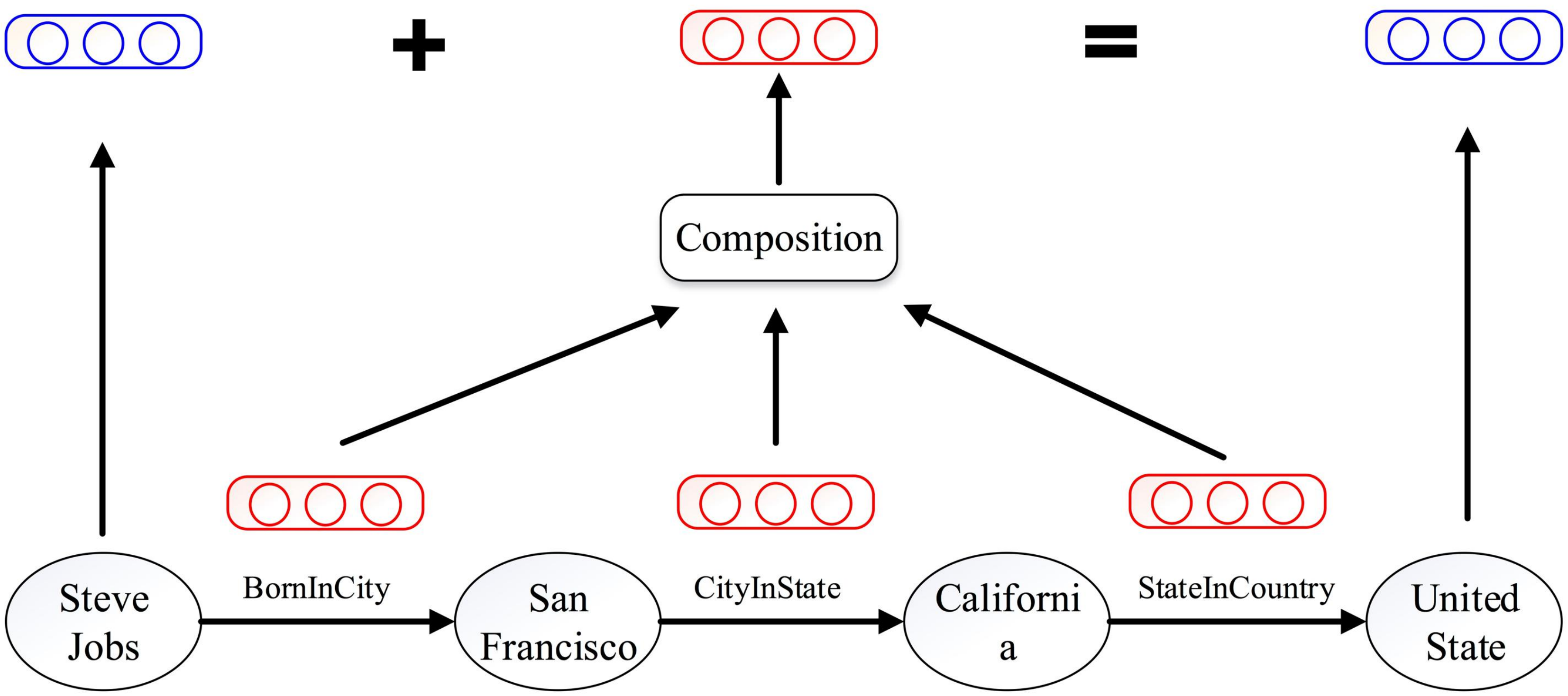
- Combine path-based and embedding-based model

	TransE	PTransE
KB	$h \xrightarrow{r} t$	$h \xrightarrow{r_1} e_1 \xrightarrow{r_2} t$
Triples	(h, r, t)	(h, r_1, e_1) (e_1, r_2, t) $(h, r_1 \circ r_2, t)$
Objectives	$\mathbf{h} + \mathbf{r} = \mathbf{t}$	$\mathbf{h} + \mathbf{r}_1 = \mathbf{e}_1$ $\mathbf{e}_1 + \mathbf{r}_2 = \mathbf{t}$ $\mathbf{h} + (\mathbf{r}_1 \circ \mathbf{r}_2) = \mathbf{t}$

Modeling Relation Paths for Representation Learning of Knowledge Bases. Lin, et al. EMNLP. 2015.

Path-based TransE

- Semantic composition: add, multiply, RNN



Logic Rules

- Markov Logic Networks
  - Composition Rules

$$\forall x, y, z \in E, \mathbf{v}_{(x,r_i,y)} \wedge \mathbf{v}_{(y,r_j,z)} \Rightarrow \mathbf{v}_{(x,r_k,z)}$$

  - Inverse Rules

$$\forall x, y \in E, \mathbf{v}_{(x,r_i,y)} \Rightarrow \mathbf{v}_{(y,r_j,x)}$$

  - Symmetric Rules

$$\forall x, y \in E, \mathbf{v}_{(x,r,y)} \Rightarrow \mathbf{v}_{(y,r,x)}$$

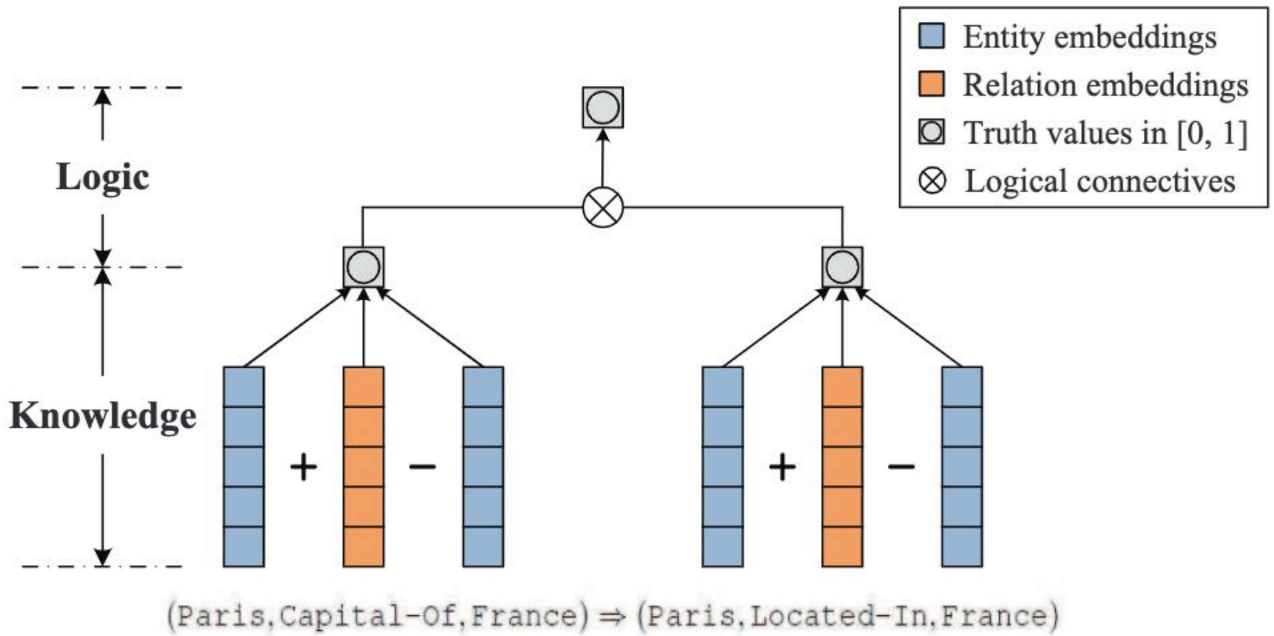
  - Subrelation Rules

$$\forall x, y \in E, \mathbf{v}(x, r_i, y) \Rightarrow \mathbf{v}(x, r_j, y)$$

- Combine rules and embeddings

KALE

- Represent and model facts and rules in a unified framework
  - Compose constituent triple scores for rule scores

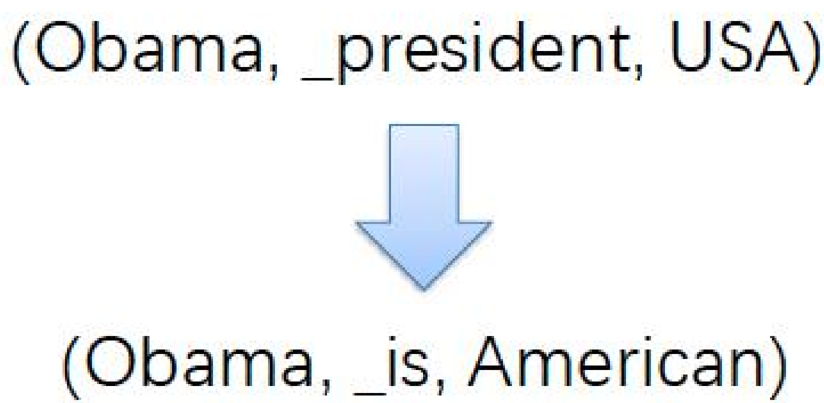


Jointly Embedding Knowledge Graphs and Logical Rules. Guo, et al. EMNLP. 2016.

Summary

- Relational paths and logic rules contain rich inference patterns about knowledge

- More complex inference patterns should be taken into consideration



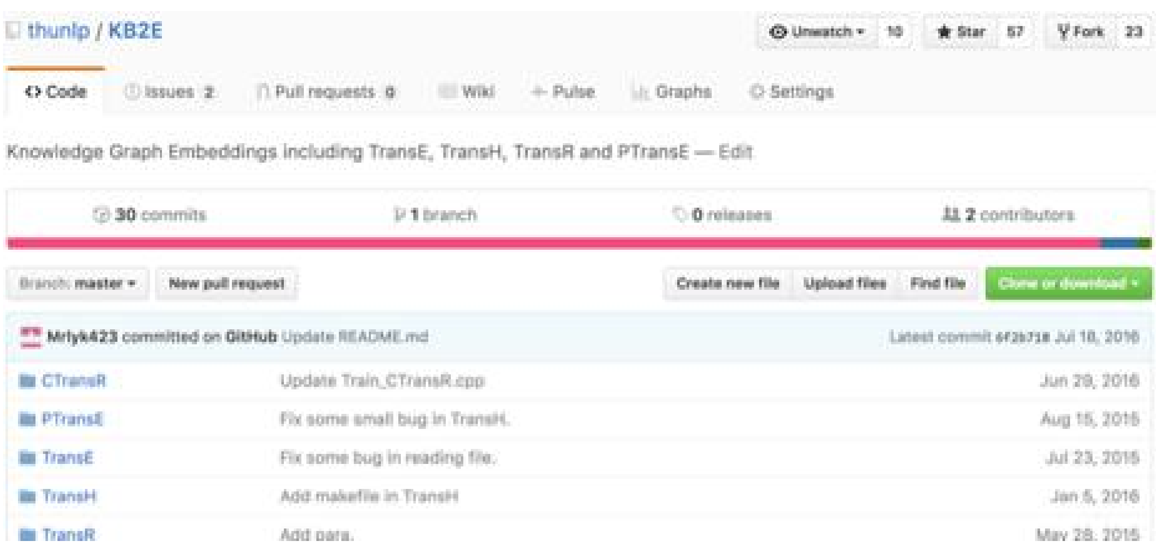
Other Key Challenges in KRL

- Online and fast learning for large-scale KGs
- Large-scale KGs are sparse, existing models cannot learn good representation for infrequent entities and relations
- Learning orders of triples are important for fast KRL
  - Curriculum Learning

Open Source Codes

https://github.com/thunlp

- OpenKE: TransE、TransH、TransR、PTransE

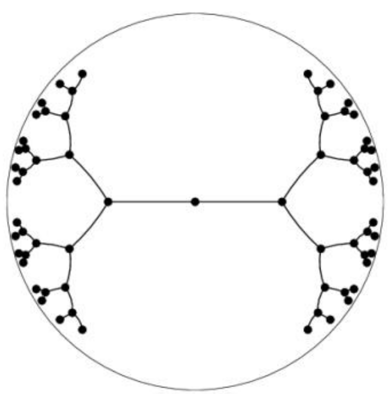


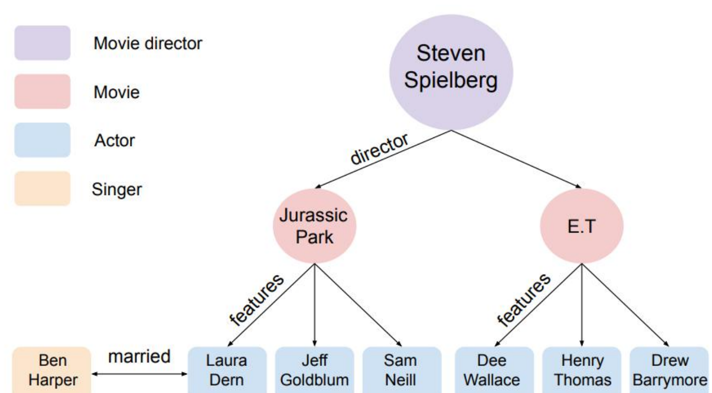
Frontiers of KRL

THUNLP

Hyperbolic Embeddings

- Hyperbolic space
  - Suitable for hierarchical data using few dimensions
  - Hierarchies are common in KGs
- ATTH
  - Relation-specific curvatures
  - parameterize hyperbolic isometries to capture relations' logical patterns





Low-Dimensional Hyperbolic Knowledge Graph Embeddings. Chami, et al. ACL. 2020.

Revisiting KRL Models

- KRL models differ along
  - Model architecture
  - Training strategies
  - Hyperparameter optimization
- What's their impact?

Publication	Model	Loss	Training	Regularizer	Optimizer	Reciprocal
Nickel et al. (2011)	**RESCAL**	MSE	Full	**L2**	ALS	No
Bordes et al. (2013)	**TransE**	**MR**	**NegSamp**	**Normalization**	**SGD**	No
Yang et al. (2015)	**DistMult**	MR	NegSamp	**Weighted** L2	**Adagrad**	No
Trouillon et al. (2016)	**ComplEx**	**BCE**	NegSamp	Weighted L2	Adagrad	No
Kadlec et al. (2017)	DistMult	**CE**	NegSamp	Weighted L2	**Adam**	No
Dettmers et al. (2018)	**ConvE**	BCE	**KvsAll**	**Dropout**	Adam	**Yes**
Lacroix et al. (2018)	ComplEx	CE	**1vsAll**	Weighted **L3**	Adagrad	Yes

MSE = mean squared error, MR = margin ranking, BCE = binary cross entropy, CE = cross entropy

You Can Teach An Old Dog New Tricks! On Training Knowledge Graph Embeddings. Ruffinelli, et al. ICLR. 2020.

Revisiting KRL Models

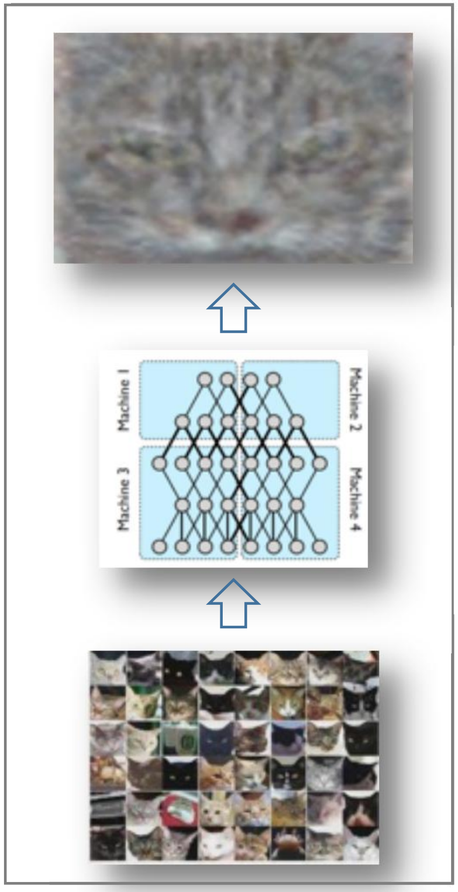
- Experimental study
  - Search in a large configuration space
  - Training strategies are important

		FB15K-237		WNRR	
		MRR	Hits@10	MRR	Hits@10
First	RESCAL (Wang et al., 2019)	27.0	42.7	42.0	44.7
	TransE (Nguyen et al., 2018)	29.4	46.5	22.6	50.1
	DistMult (Dettmers et al., 2018)	24.1	41.9	43.0	49.0
	ComplEx (Dettmers et al., 2018)	24.7	42.8	44.0	51.0
	ConvE (Dettmers et al., 2018)	32.5	50.1	43.0	52.0
Ours	RESCAL	**35.7**	**54.1**	46.7	51.7
	TransE	31.3	49.7	22.8	52.0
	DistMult	34.3	53.1	45.2	53.1
	ComplEx	34.8	53.6	**47.5**	**54.7**
	ConvE	33.9	52.1	44.2	50.4
Recent	TuckER (Balazevic et al., 2019)	**35.8**	**54.4**	47.0	52.6
	RotatE (Sun et al., 2019a)	33.8	53.3	**47.6**	**57.1**
	SACN (Shang et al., 2019)	35.0	54.0	47.0	54.4
Large	DistMult (Salehi et al., 2018)	35.7	54.8	45.5	54.4
	ComplEx-N3 (Lacroix et al., 2018)	**37.0**	**56.0**	**49.0**	**58.0**

You Can Teach An Old Dog New Tricks! On Training Knowledge Graph Embeddings. Ruffinelli, et al. ICLR. 2020.

Take Home Message

- KRL is a promising approach to construction and application of KGs
- KRL is still a rising research area, there are many open problems
- Learn from the generalization & abstraction abilities of human
  - Zero / One shot learning



How to grow a mind: statistics, structure, and abstraction. Joshua B. Tenenbaum, et al. Science. 2011.
Human-level concept learning through probabilistic program induction. Lake, et al. Science. 2015.

Outlook

- DL & KG will bring a revolution to NLP

- Knowledge-Guided Language Understanding
  - Sentence: Question Answering, Conversation
  - Document: Summarization, Reading Comprehension

- Knowledge-Guided Language Generation
  - Legal, Patent, Finance, Science

Reading Material

a. Introduction to KG

· Towards a Definition of Knowledge Graphs. Lisa Ehrlinger, Wolfram Wöß [link]

· KG Definition & History Wiki [link]

· Semantic Network [link]

b. Knowledge Representation Learning & Reasoning

· KRL paper list [link]

· Knowledge Representation Learning: A Review. (In Chinese) Zhiyuan Liu, Maosong Sun, Yankai Lin, Ruobing Xie. Computer Research and Development 2016. [link]

· A Review of Relational Machine Learning for Knowledge Graphs. Maximilian Nickel, Kevin Murphy, Volker Tresp, Evgeniy Gabrilovich. 2016. [link]

· Knowledge Graph Embedding: A Survey of Approaches and Applications. Quan Wang, Zhendong Mao, Bin Wang, Li Guo. TKDE 2017. [link]

· OpenKE [link]

· KG Reasoning paper list [link] & PPT [link]

For reading material recommendation of this course, please refer to our github.

Q&A

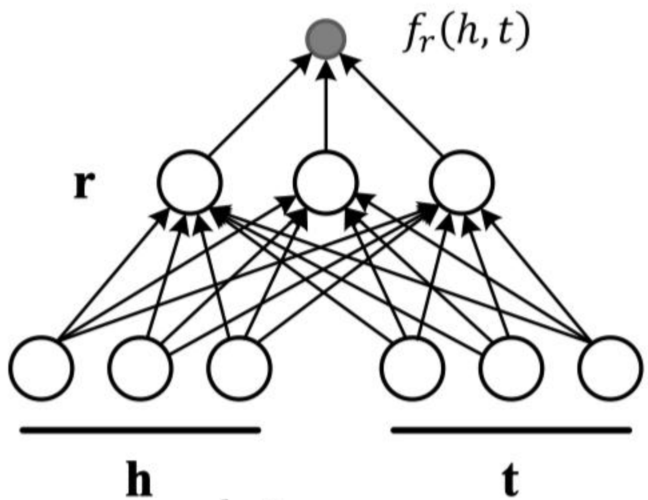
THUNLP

Appendix

THUNLP

Semantic: HolE

- Scoring function: circular correlation
  - Combine the expressive power of RESCAL with the efficiency and simplicity of DistMult



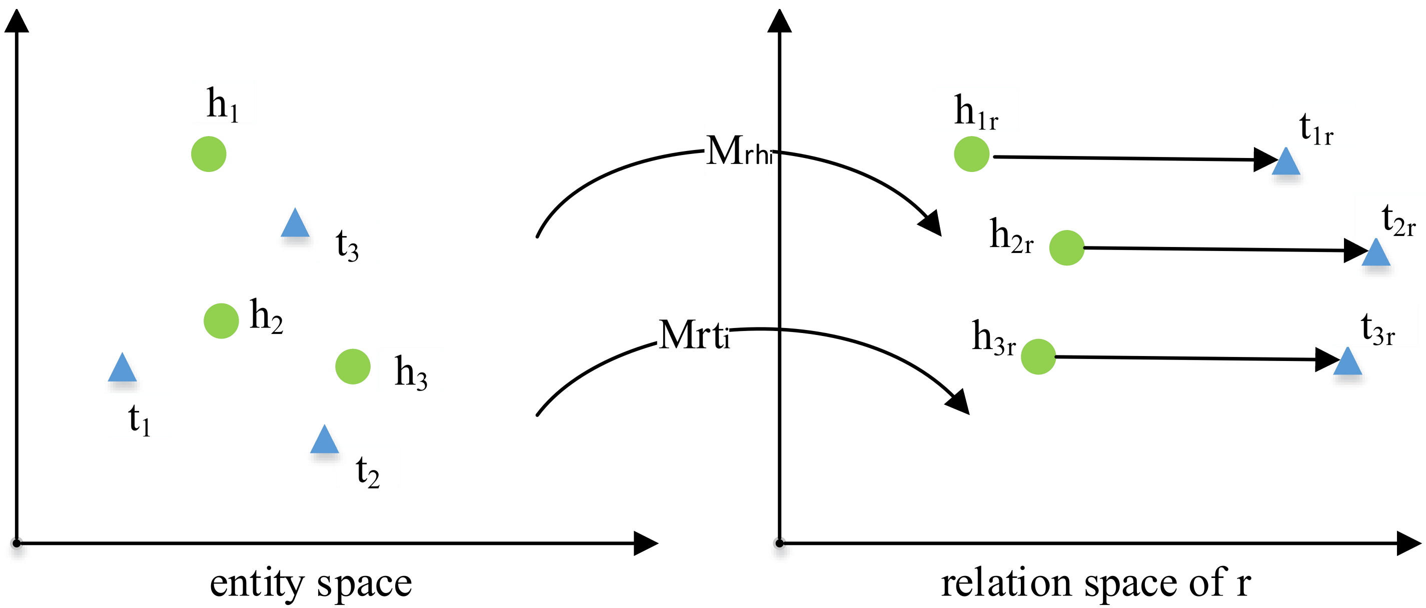
$$[\mathbf{h} \star \mathbf{t}]_i = \sum_{k=0}^{d-1} [\mathbf{h}]_k \cdot [\mathbf{t}]_{(k+i) \bmod d}.$$

$$f_r(h,t) = \mathbf{r}^\top (\mathbf{h} \star \mathbf{t})$$

Holographic embeddings of knowledge graphs. Nickel, et al. AAAI. 2016.

TransD

- Build relation-specific entity embeddings with projection matrices related not only to relation but also head/tail entities



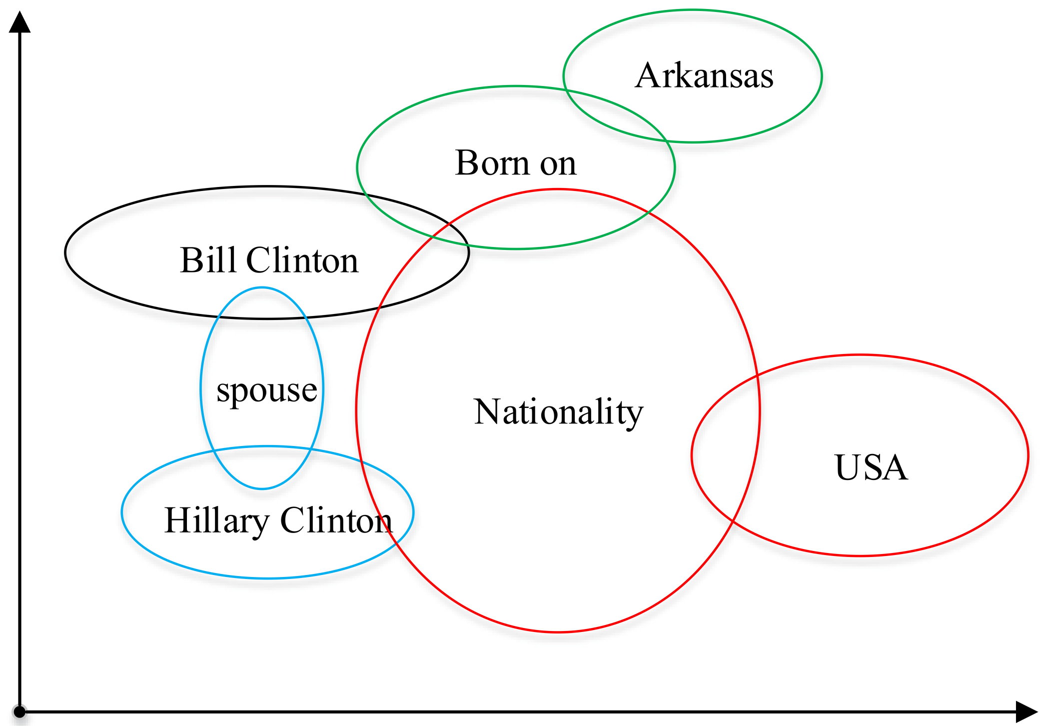
$$\mathbf{M}_r^1 = \mathbf{w}_r \mathbf{w}_h^\top + \mathbf{I}, \quad \mathbf{M}_r^2 = \mathbf{w}_r \mathbf{w}_t^\top + \mathbf{I}$$

$$\mathbf{h}_\perp = \mathbf{M}_r^1 \mathbf{h}, \quad \mathbf{t}_\perp = \mathbf{M}_r^2 \mathbf{t}$$

Knowledge graph embedding via dynamic mapping matrix. Ji, et al. ACL-IJCNLP. 2015.

KG2E

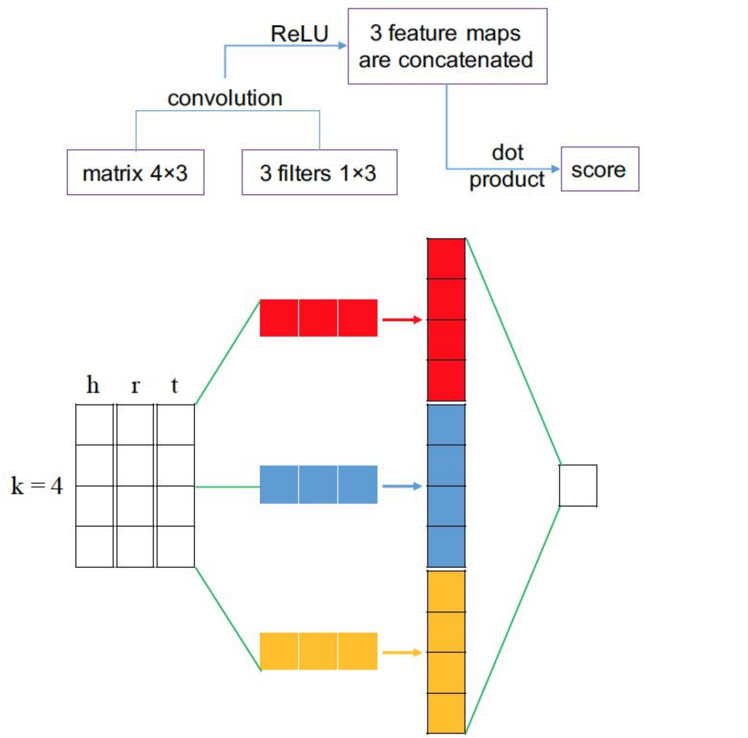
- Model relations and entities with Gaussian Distribution
  - Specially consider the (un)certainties of entities and relations



Learning to represent knowledge graphs with gaussian embedding. He, et al. CIKM. 2015.

ConvKB

- Adopt CNNs for encoding the concatenation of entities and relations

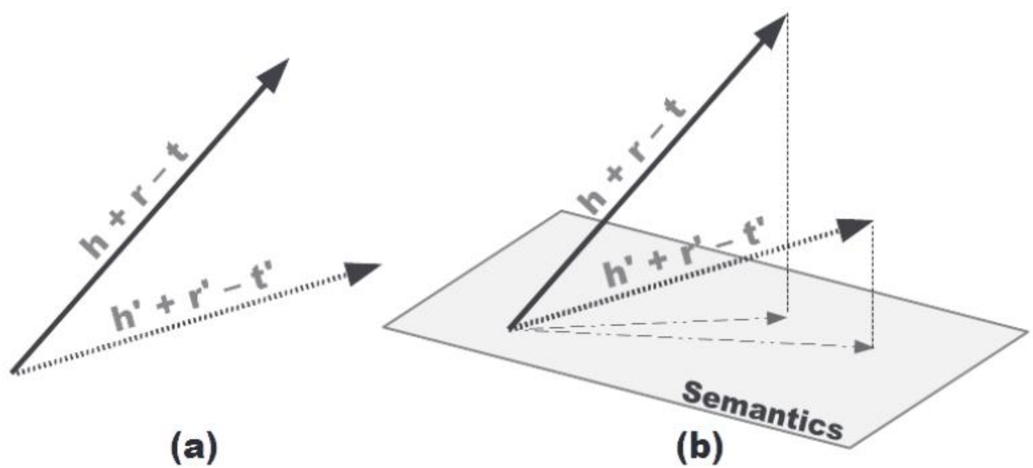


$$f_r(h,t) = \text{concat}\left(\sigma\left([\boldsymbol{h},\boldsymbol{r},\boldsymbol{t}] * \boldsymbol{\omega}\right)\right) \cdot \mathbf{w}.$$

A novel embedding model for knowledge base completion based on convolutional neural network. Nguyen, et al. NAACL. 2018.

SSP

- Project loss vectors to the semantic hyperplane
- Model the strong correlations between triples and textual descriptions

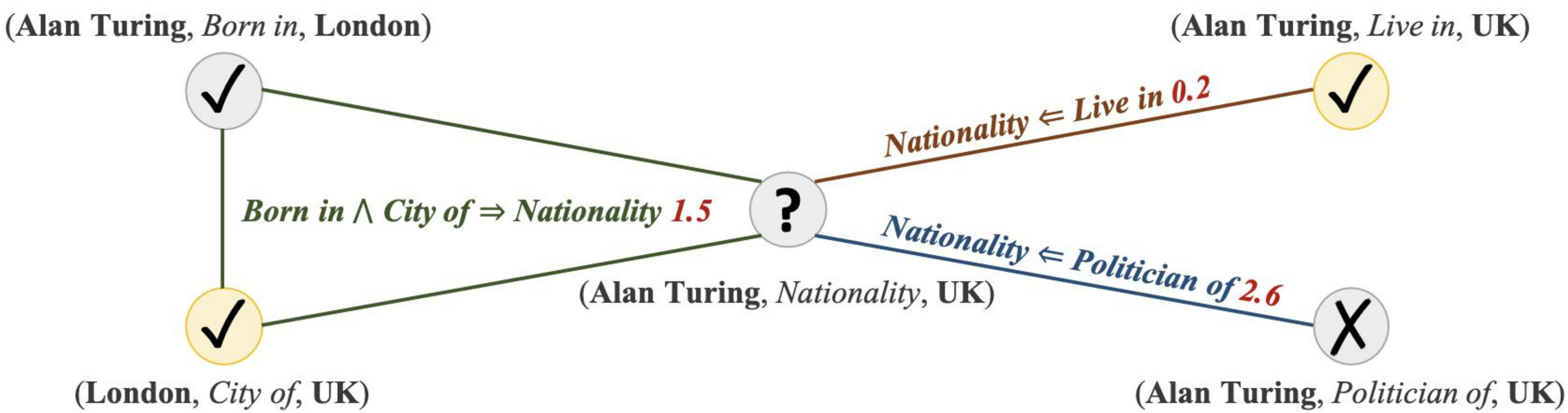


$$\mathcal{S}(\mathbf{s_h}, \mathbf{s_t}) = \frac{\mathbf{s_h} + \mathbf{s_t}}{\|\mathbf{s_h} + \mathbf{s_t}\|_2^2}$$

SSP: Semantic space projection for knowledge graph embedding with text descriptions. Xiao, et al. AAAI. 2017.

pLogicNet

- Combine MLN and KRL
  - Variational EM algorithm
  - E-step: Infer missing triples using logic rules
  - M-step: Update the weights of rules



Probabilistic Logic Neural Networks for Reasoning. Qu, et al. NIPS. 2019.