Query Embeddings for KGs

I. Makarov & V. Pozdnyakov & D. Kiselev

BigData Academy MADE from Mail.ru Group

Graph Neural Networks and Applications

Topics

Knowledge Graphs Recap

Knowledge Graphs

Main idea

- Knowledge as graphs (linked data)
- Nodes as entities
- Labels as attributes
- Edges as relation types (heterogeneous network)

Applications

- Analytic representation of data
- Interpretable decision making
- Reasoning & QA
- Edges as relation types (heterogeneous network)

The Semantic Web

RDF representation

- r(s,p,o)= subject–predicate-object relation
- ABox representing data
- TBox representing rules (ontologies)
- rdfs:domain, rdfs:range, rdf:type, rdfs:subClassOf, rdfs:subPropertyOf
- owl:inverseOf, owl:TransitiveProperty, owl:FunctionalProperty

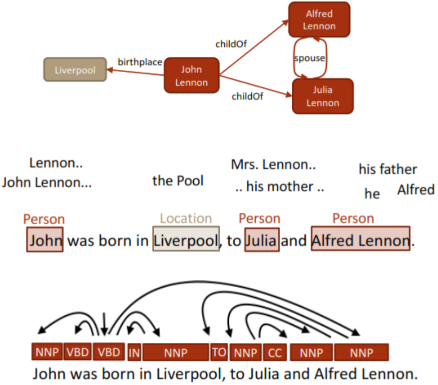
Knowledge Extraction vs. KG Construction

Problem	KE	KG
Who are entities?	NER & Coreference	Entity Linking
What are the attributes?	NER	Classification
How are they connected?	Relation extraction	Link Prediction

Table: Difference in view on knowledge mining

Knowledge Extraction

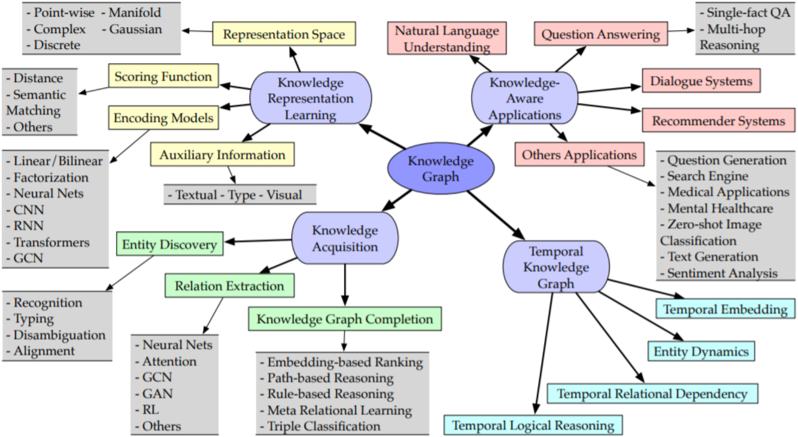
- Entity resolution, Entity linking, Relation Extraction (corpora)
- Coreference resolution (document)
- Dependency parsing, part of speech tagging, NER (sentence)



from `https://kgtutorial.github.io/`, 2018

A Survey on Knowledge Graphs: Representation, Acquisition and Applications

- Taxonomy of research and applications

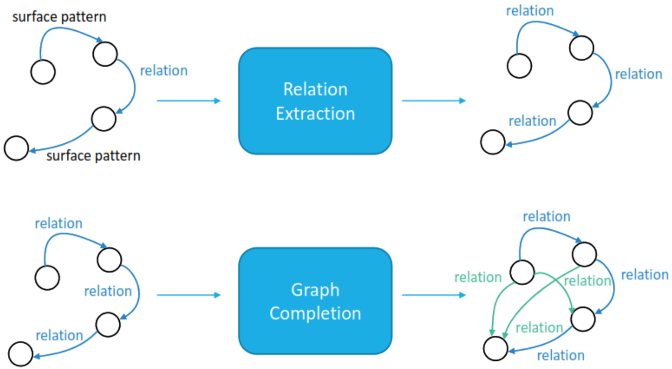


from Yu et al., 2021

Knowledge Graph Completion

Relation Extraction and KG Completion

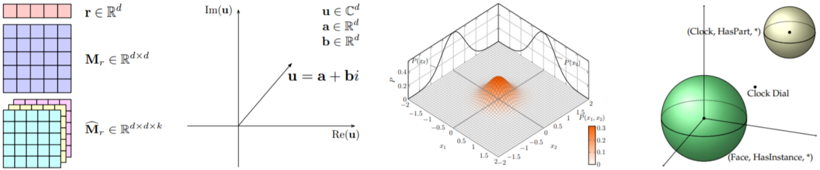
- Similar Pairs of Entities refer to similar relations (not identical)
- Similar Relations refer to paraphrases or implications
- Logical rules $\rightarrow$ Embedding space



from https://kgtutorial.github.io/, 2018

Entity and relation representations

- Different embedding spaces
- Different relation properties to capture

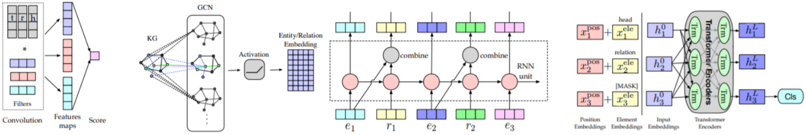


(a) Point-wise space. (b) Complex vector space. (c) Gaussian distribution. (d) Manifold space.

from Yu et al., 2021

Entity and relation encoding

- CNN uses dense layer and convolution operation
- GCN encodes knowledge graphs
- RSN encodes entity-relation sequences and with masking relationsdiscriminatively
- Transformer-based CoKE encodes triples as sequences with masked entities

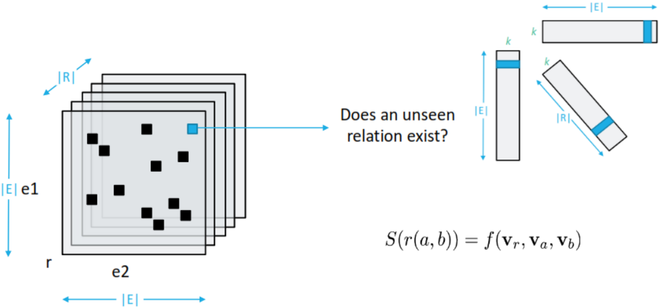


(a) CNN. (b) GCN. (c) RSN. (d) Transformer.

from Yu et al., 2021

KG Completion

Tensor Formulation of KG



$$S(r(a,b)) = f(\mathbf{v}_r, \mathbf{v}_a, \mathbf{v}_b)$$

from `https://kgtutorial.github.io/`, 2018

KG Completion

CANDECOMP/PARAFAC-Decomposition

$$S\left(r(a,b)\right) = \sum_k R_{r,k} \cdot e_{a,k} \cdot e_{b,k}$$

Tucker2 and RESCAL Decompositions

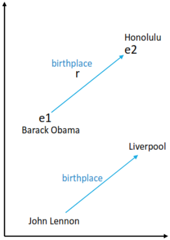
$$S\left(r(a,b)\right) = \left(\mathbf{R}_r \times \mathbf{e}_a\right) \times \mathbf{e}_b$$

Model E

$$S\left(r(a,b)\right) = \mathbf{R}_{r,1} \cdot \mathbf{e}_a + \mathbf{R}_{r,2} \cdot \mathbf{e}_b$$

Holographic Embeddings

$$S(r(a,b)) = \mathbf{R}_r \times \left(\mathbf{e}_a \star \mathbf{e}_b\right)$$



TransE

$$S\left(r(a,b)\right) = -\|\mathbf{e}_a + \mathbf{R}_r - \mathbf{e}_b\|_2^2$$

TransH

$$S\left(r(a,b)\right) = -\|\mathbf{e}_a^{\perp} + \mathbf{R}_r - \mathbf{e}_b^{\perp}\|_2^2$$

$$\mathbf{e}_a^{\perp} = \mathbf{e}_a - \mathbf{w}_r^T \mathbf{e}_a \mathbf{w}_r$$

TransR

$$S\left(r(a,b)\right) = -\|\mathbf{e}_a \mathbf{M}_r + \mathbf{R}_r - \mathbf{e}_b \mathbf{M}_r\|_2^2$$

from `https://kgtutorial.github.io/`, 2018

Scoring models II

Category	Model	Ent. embed.	Rel. embed.	Scoring Function $f_r(h,t)$
Polar coordinate	HAKE [19]	$\mathbf{h}_m, \mathbf{t}_m \in \mathbb{R}^k$ $\mathbf{h}_p, \mathbf{t}_p \in [0, 2\pi)^k$	$\mathbf{r}_m \in \mathbb{R}_+^k$ $\mathbf{r}_p, \in [0, 2\pi)^k$	$-\lVert \mathbf{h}_m \circ \mathbf{r}_m - \mathbf{t}_m \rVert_2 -$ $\lambda \lVert \sin\left(\left(\mathbf{h}_p + \mathbf{r}_p - \mathbf{t}_p \right) / 2 \right) \rVert_1$
Complex vector	ComplEx [23]	$\mathbf{h}, \mathbf{t} \in \mathbb{C}^d$	$\mathbf{r} \in \mathbb{C}^d$	$\mathrm{Re}\left(< \mathbf{r}, \mathbf{h}, \overline{\mathbf{t}} > \right) = \mathrm{Re}\left(\sum_{k=1}^{K} \mathbf{r}_k \mathbf{h}_k \overline{\mathbf{t}}_k \right)$
	RotatE [24]	$\mathbf{h}, \mathbf{t} \in \mathbb{C}^d$	$\mathbf{r} \in \mathbb{C}^d$	$\lVert \mathbf{h} \circ \mathbf{r} - \mathbf{t} \rVert$
	QuatE [25]	$\mathbf{h}, \mathbf{t} \in \mathbb{H}^d$	$\mathbf{r} \in \mathbb{H}^d$	$\mathbf{h} \otimes \frac{\mathbf{r}}{\vert\mathbf{r}\vert} \cdot \mathbf{t}$
Manifold & Group	ManifoldE [28]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$\lVert \mathcal{M}(h, r, t) - D_r^2 \rVert^2$
	TorusE [15]	$[\mathbf{h}], [\mathbf{t}] \in \mathbb{T}^n$	$[\mathbf{r}] \in \mathbb{T}^n$	$\min_{(x,y) \in ([h]+[r]) \times [t]} \lVert x - y \rVert_i$
	DihEdral [31]	$\mathbf{h}^{(l)}, \mathbf{t}^{(l)} \in \mathbb{R}^2$	$\mathbf{R}^{(l)} \in \mathbb{D}_K$	$\sum_{l=1}^{L} \mathbf{h}^{(l)\top} \mathbf{R}^{(l)} \mathbf{t}^{(l)}$
	MuRP [29]	$\mathbf{h}, \mathbf{t} \in \mathbb{B}_c^d, b_h, b_t \in \mathbb{R}$	$\mathbf{r} \in \mathbb{B}_c^d$	$-d_{\mathbb{B}}\left(\exp_{\mathbf{0}}^c \left(\mathbf{R} \log_{\mathbf{0}}^c (\mathbf{h}) \right), \mathbf{t} \oplus_c \mathbf{r} \right)^2 + b_h + b_t$
	AttH [30]	$\mathbf{h}, \mathbf{t} \in \mathbb{B}_c^d, b_h, b_t \in \mathbb{R}$	$\mathbf{r} \in \mathbb{B}_c^d$	$-d_{\mathbb{B}}^{c_r}\left(\mathrm{Att}\left(\mathbf{q}_{\mathrm{Rot}}^H, \mathbf{q}_{\mathrm{Ref}}^H; \mathbf{a}_r \right) \oplus^{c_r} \mathbf{r}_r^H, \mathbf{e}_t^H \right)^2 + b_h + b_t$
Gaussian	KG2E [26]	$\mathbf{h} \sim \mathcal{N}(\boldsymbol{\mu}_h, \boldsymbol{\Sigma}_h)$ $\mathbf{t} \sim \mathcal{N}(\boldsymbol{\mu}_t, \Sigma_t)$ $\boldsymbol{\mu}_h, \boldsymbol{\mu}_t \in \mathbb{R}^d$ $\Sigma_h, \Sigma_t \in \mathbb{R}^{d \times d}$	$\mathbf{r} \sim \mathcal{N}(\boldsymbol{\mu}_r, \Sigma_r)$ $\boldsymbol{\mu}_r \in \mathbb{R}^d, \Sigma_r \in \mathbb{R}^{d \times d}$	$\int_{x \in \mathcal{R}^{k_e}} \mathcal{N}(x; \boldsymbol{\mu}_r, \boldsymbol{\Sigma}_r) \log \frac{\mathcal{N}(x; \boldsymbol{\mu}_e, \boldsymbol{\Sigma}_e)}{\mathcal{N}(x; \boldsymbol{\mu}_r, \boldsymbol{\Sigma}_r)} dx$ $\log \int_{x \in \mathcal{R}^{k_e}} \mathcal{N}(x; \boldsymbol{\mu}_e, \boldsymbol{\Sigma}_e) \mathcal{N}(x; \boldsymbol{\mu}_r, \boldsymbol{\Sigma}_r) dx$
	TransG [27]	$\mathbf{h} \sim \mathcal{N}(\boldsymbol{\mu}_h, \boldsymbol{\sigma}_h^2 \mathbf{I})$ $\mathbf{t} \sim \mathcal{N}(\boldsymbol{\mu}_t, \Sigma_t)$ $\boldsymbol{\mu}_h, \boldsymbol{\mu}_t \in \mathbb{R}^d$	$\boldsymbol{\mu}_r^i \sim \mathcal{N}\left(\boldsymbol{\mu}_t - \boldsymbol{\mu}_h, \left(\sigma_h^2 + \sigma_t^2 \right) \mathbf{I} \right)$ $\mathbf{r} = \sum_i \pi_r^i \boldsymbol{\mu}_r^i \in \mathbb{R}^d$	$\sum_i \pi_r^i \exp\left(-\frac{\lVert \boldsymbol{\mu}_h + \boldsymbol{\mu}_r^i - \boldsymbol{\mu}_t \rVert_2^2}{\sigma_h^2 + \sigma_t^2} \right)$
Translational Distance	TransE [16]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$-\lVert \mathbf{h} + \mathbf{r} - \mathbf{t} \rVert_{1/2}$
	TransR [17]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^k, \mathbf{M}_r \in \mathbb{R}^{k \times d}$	$-\lVert \mathbf{M}_r \mathbf{h} + \mathbf{r} - \mathbf{M}_r \mathbf{t} \rVert_2^2$
	TransH [20]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r}, \mathbf{w}_r \in \mathbb{R}^d$	$-\left\lVert \left(\mathbf{h} - \mathbf{w}_r^\top \mathbf{h} \mathbf{w}_r \right) + \mathbf{r} - \left(\mathbf{t} - \mathbf{w}_r^\top \mathbf{t} \mathbf{w}_r \right) \right\rVert_2^2$
	TransA [34]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d, \mathbf{M}_r \in \mathbb{R}^{d \times d}$	$(\vert\mathbf{h} + \mathbf{r} - \mathbf{t}\vert)^\top \mathbf{W}_\mathbf{r} (\vert\mathbf{h} + \mathbf{r} - \mathbf{t}\vert)$
	TransF [35]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$(\mathbf{h} + \mathbf{r})^\top \mathbf{t} + (\mathbf{t} - \mathbf{r})^\top \mathbf{h}$
	ITransF [36]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$\left\lVert \boldsymbol{\alpha}_r^H \cdot \mathbf{D} \cdot \mathbf{h} + \mathbf{r} - \boldsymbol{\alpha}_r^T \cdot \mathbf{D} \cdot \mathbf{t} \right\rVert_\ell$
	TransAt [37]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$P_r(\sigma(\mathbf{r}_h)\mathbf{h}) + \mathbf{r} - P_r(\sigma(\mathbf{r}_t)\mathbf{t})$
	TransD [33]	$\mathbf{h}, \mathbf{t}, \mathbf{w}_h \mathbf{w}_t \in \mathbb{R}^d$	$\mathbf{r}, \mathbf{w}_r \in \mathbb{R}^k$	$-\left\lVert \left(\mathbf{w}_r \mathbf{w}_h^\top + \mathbf{I} \right) \mathbf{h} + \mathbf{r} - \left(\mathbf{w}_r \mathbf{w}_t^\top + \mathbf{I} \right) \mathbf{t} \right\rVert_2^2$
	TransM [211]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$-\theta_r \lVert \mathbf{h} + \mathbf{r} - \mathbf{t} \rVert_{1/2}$
	TranSparse [212]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^k, \mathbf{M}_r(\theta_r) \in \mathbb{R}^{k \times d}$ $\mathbf{M}_r^1(\theta_r^1), \mathbf{M}_r^2(\theta_r^2) \in \mathbb{R}^{k \times d}$	$-\lVert \mathbf{M}_r(\theta_r)\mathbf{h} + \mathbf{r} - \mathbf{M}_r(\theta_r)\mathbf{t} \rVert_{1/2}^2$ $-\lVert \mathbf{M}_r^1(\theta_r^1)\mathbf{h} + \mathbf{r} - \mathbf{M}_r^2(\theta_r^2)\mathbf{t} \rVert_{1/2}^2$

from Yu et al., 2021

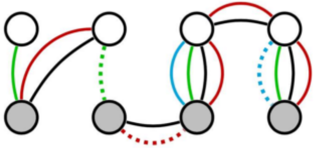
Scoring models II

- Different approaches arising from problems with relation properties preserving

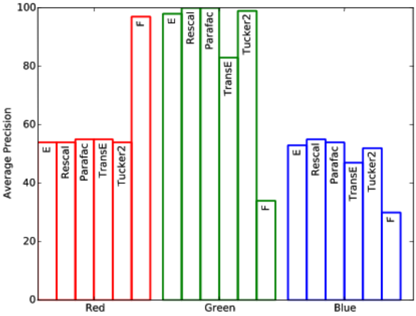
Semantic Matching	TATEC [213]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d, \mathbf{M}_r \in \mathbb{R}^{d \times d}$	$\mathbf{h}^\top \mathbf{M}_r \mathbf{t} + \mathbf{h}^\top \mathbf{r} + \mathbf{t}^\top \mathbf{r} + \mathbf{h}^\top \mathbf{D} \mathbf{t}$
	ANALOGY [22]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{M}_r \in \mathbb{R}^{d \times d}$	$\mathbf{h}^\top \mathbf{M}_r \mathbf{t}$
	CrossE [42]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$\sigma\left(\tanh\left(\mathbf{c}_r \circ \mathbf{h} + \mathbf{c}_r \circ \mathbf{h} \circ \mathbf{r} + \mathbf{b}\right) \mathbf{t}^\top\right)$
	SME [39]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$g_{\text{left}}(\mathbf{h}, \mathbf{r})^\top g_{\text{right}}(\mathbf{r}, \mathbf{t})$
	DistMult [32]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$\mathbf{h}^\top \operatorname{diag}(\mathbf{M}_r) \mathbf{t}$
	HolE [21]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$\mathbf{r}^\top (h \star t)$
	HolEx [40]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$\sum_{j=0}^{l} p\left(\mathbf{h}, \boldsymbol{r}; \mathbf{c}_j\right) \cdot \boldsymbol{t}$
	SE [8]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{M}_r^1, \mathbf{M}_r^2 \in \mathbb{R}^{d \times d}$	$-\left\\|\mathbf{M}_r^1 \mathbf{h} - \mathbf{M}_r^2 \mathbf{t}\right\\|_1$
	SimplE [48]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r}, \mathbf{r}' \in \mathbb{R}^d$	$\frac{1}{2}\left(\mathbf{h} \circ \mathbf{r} \mathbf{t} + \mathbf{t} \circ \mathbf{r}' \mathbf{t}\right)$
	RESCAL [49]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{M}_r \in \mathbb{R}^{d \times d}$	$\mathbf{h}^\top \mathbf{M}_r \mathbf{t}$
	LFM [51]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{u}_r, \mathbf{v}_r \in \mathbb{R}^p$	$\mathbf{h}^\top \sum_{i=1}^{d} \boldsymbol{\alpha}_i^r \mathbf{u}_i \mathbf{v}_i^\top \mathbf{t}$
	TuckER [52]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}_e^d$	$\mathbf{r} \in \mathbb{R}_r^d$	$\mathcal{W} \times_1 \mathbf{h} \times_2 \mathbf{r} \times_3 \mathbf{t}$
	LowFER [53]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$\left(\mathbf{S}^k \operatorname{diag}\left(\mathbf{U}^T \mathbf{h}\right) \mathbf{V}^T \mathbf{r}\right)^T \mathbf{t}$
Neural Networks	MLP [3]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$\sigma(\mathbf{w}^\top \sigma(\mathbf{W}[\mathbf{h}, \mathbf{r}, \mathbf{t}]))$
	NAM [54]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$\sigma\left(\mathbf{z}^{(L)} \cdot \mathbf{t} + \mathbf{B}^{(L+1)} \mathbf{r}\right)$
	ConvE [55]	$\mathbf{M}_h \in \mathbb{R}^{d_w \times d_h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{M}_r \in \mathbb{R}^{d_w \times d_h}$	$\sigma\left(\operatorname{vec}\left(\sigma\left(\left[\mathbf{M}_h; \mathbf{M}_r\right] * \boldsymbol{\omega}\right)\right) \mathbf{W}\right) \mathbf{t}$
	ConvKB [43]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$\operatorname{concat}\left(\sigma\left(\left[\boldsymbol{h}, \boldsymbol{r}, \boldsymbol{t}\right] * \boldsymbol{\omega}\right)\right) \cdot \mathbf{w}$
	HypER [56]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{w}_r \in \mathbb{R}^{d_r}$	$\sigma\left(\operatorname{vec}\left(\mathbf{h} * \operatorname{vec}^{-1}\left(\mathbf{w}_r \mathbf{H}\right)\right) \mathbf{W}\right) \mathbf{t}$
	SACN [44]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r} \in \mathbb{R}^d$	$g\left(\operatorname{vec}\left(\mathbf{M}\left(\mathbf{h}, \mathbf{r}\right)\right) W\right) \mathbf{t}$
	NTN [18]	$\mathbf{h}, \mathbf{t} \in \mathbb{R}^d$	$\mathbf{r}, \mathbf{b}_r \in \mathbb{R}^k, \widehat{\mathbf{M}} \in \mathbb{R}^{d \times d \times k}$ $\mathbf{M}_{r,1}, \mathbf{M}_{r,2} \in \mathbb{R}^{k \times d}$	$\mathbf{r}^\top \sigma\left(\mathbf{h}^T \widehat{\mathbf{M}} \mathbf{t} + \mathbf{M}_{r,1} \mathbf{h} + \mathbf{M}_{r,2} \mathbf{t} + \mathbf{b}_r\right)$

from Yu et al., 2021

KG Completion



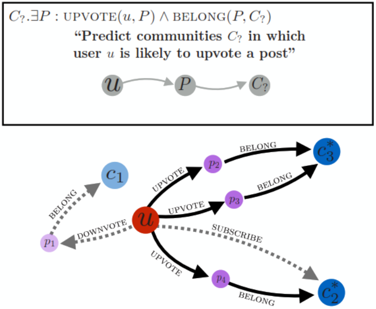
- **Red**: deterministically implied by **Black**
  - needs *pair-specific* embedding
  - Only **F** is able to generalize
- **Green**: needs to estimate entity types
  - needs *entity-specific* embedding
  - Tensor factorization generalizes, **F** doesn't
- **Blue**: implied by **Red** and **Green**
  - Nothing works much better than random

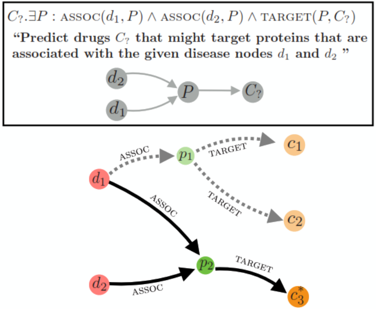


from Singh et al., 2015

Query Embedding to Knowledge Graph

- Simple conjunctive graph queries
- Example on the left shows a path query on the Reddit data with multiple answers
- Example on the right shows a more complex query with a polytree structure on the biological interaction data





from Leskovec et al., 2018

Embedding Logical Queries on Knowledge Graphs

- GQE framework takes input query q, represents this query according to its DAG (directed acyclic graph) structure
- Embedding of the query is generated based on this DAG by applying geometric operations P (projection) and I (intersection)
- Generated query embedding is used to predict the likelihood that a node satisfies the query via kNN in the embedding space

Algorithm 1: Query embedding generation

Input : Query anchor nodes $\mathcal{A}$, query variable nodes $\mathcal{B}$, query edges $\mathcal{E}_q$, a map d_q from query variables to their degree in the query DAG

Output : Query embedding $\mathbf{q}$

Q = dictionary mapping every $\mathcal{V}_i \in \mathcal{B}$ to an empty set;

for $\tau(v_i, V_j) \in \mathcal{E}_q \; : \; v_i \in \mathcal{A}$ **do**

  $Q[V_j] = Q[V_j] \cup \mathcal{P}(\mathbf{z}_{v_i}, \tau)$

while $|Q.key\_set| > 0$ **do**

  A = empty dictionary;

  for $V_i \in Q.key\_set \; : \; |Q[V_i]| = d_q(V_i)$ **do**

    $A[V_i] = \mathcal{I}(Q[V_i])$;

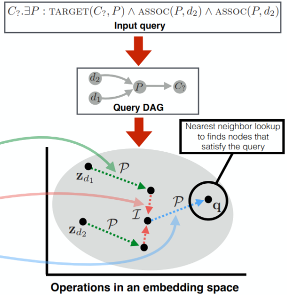
    delete $Q[V_i]$;

  for $V_i \in A.key\_set$ **do**

    for $\tau(V_j, V_k) \in \mathcal{E}_q \; : \; V_j = V_i$ **do**

      $Q[V_k] = Q[V_k] \cup \mathcal{P}(A[V_i], \tau)$;

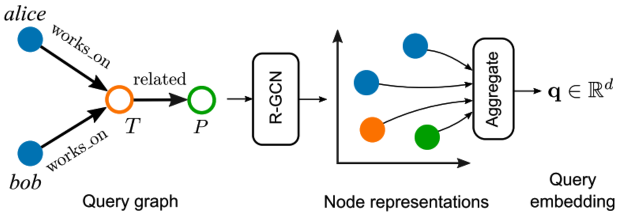
return $A[V_?]$;



from Leskovec et al., 2018

Message Passing Query Embedding

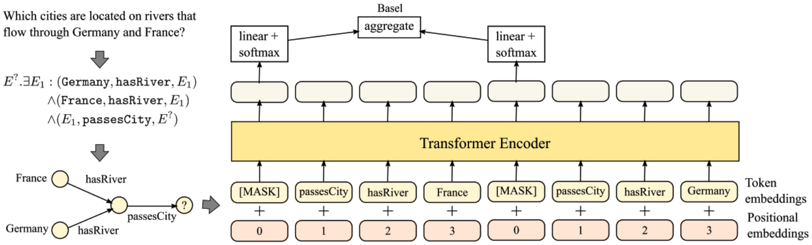
- Message Passing Query Embedding takes as input a query graph and outputs a query embedding.
- Features for each node in the query graph are embeddings of entities in the KG, or type embeddings.
- GNN propagates information across the graph, and an aggregation function yields the query embedding.



from Cochez et al., 2020

Answering Complex Queries in Knowledge Graphs with Bidirectional Sequence Encoders

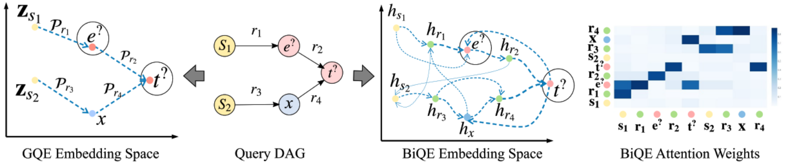
- Use of Transformers of sentences from KG
- Bidirectional Query Embedding (BIQE) embeds conjunctive queries with models based on bi-directional attention mechanisms.



from Niepert et al., 2020

Answering Complex Queries in Knowledge Graphs with Bidirectional Sequence Encoders

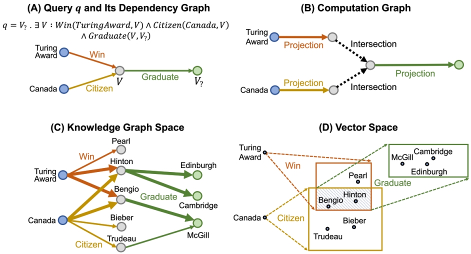
- Query embedding in GQE (left) vs. BIQE (right).
- For GQE, when computing the intersection of e or t, only the previous query context is considered and not the future.
- For BIQE, every element can attend to every other element of the query.



from Niepert et al., 2020

Query2Box: Reasoning over knowledge graphs in vector space using box embeddings

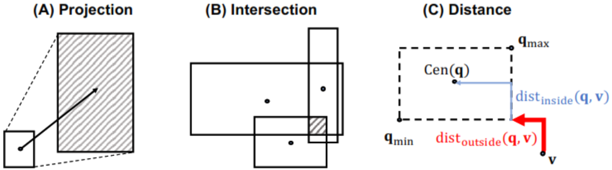
- Conjunctive query is represented with a dependency graph
- Computation graph specifies the reasoning procedure
- Knowledge graph, where green nodes denote answers to the query
- In QUERY2BOX, nodes of the KG are embedded
- Query embedding is computed according to the computation graph as a sequence of box operations



from Leskovec et al., 2020

Query2Box: Reasoning over knowledge graphs in vector space using box embeddings

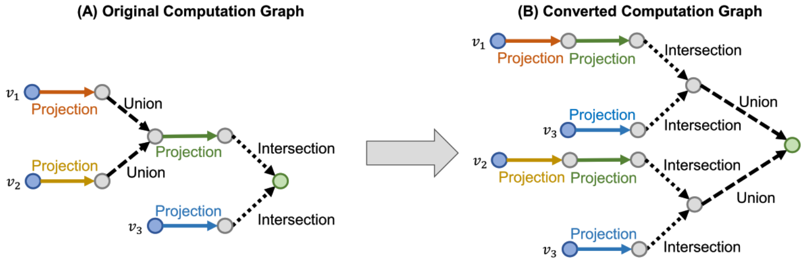
- Projection generates a larger box with a translated center
- Intersection generates a smaller box lying inside the given set of boxes
- Distance distbox is the weighted sum of dist-outside and dist-inside with lesser weight.



from Leskovec et al., 2020

Query2Box: Reasoning over knowledge graphs in vector space using box embeddings

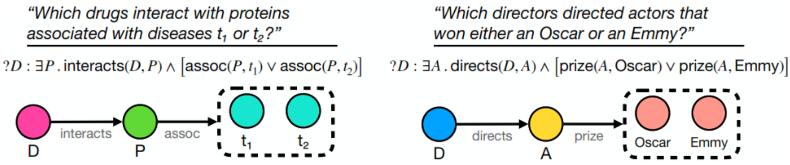
- Balancing query depth with disjunctive normal form transformation
- Better generalizability for queries with many variables



from Leskovec et al., 2020

Complex query answering with neural link predictors

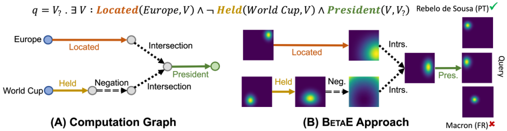
- Consequent link prediction and query probabilistic scoring over combinatorial optimized set of candidates
- Łukasiewicz loic



from Cochez et al., 2020

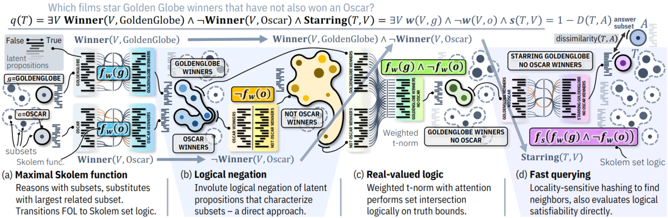
Beta Embeddings for Multi-Hop Logical Reasoning in Knowledge Graphs

- BetaE answers first-order logic queries modeling each node of the computation graph as a Beta distribution over the entity embedding space
- Each edge of the computation graph transforms the distribution via a projection, negation, or intersection probabilistic operation.
- Negation $N(\alpha, \beta) \to N(1/\alpha, 1/\beta)$, disjunction via De Moran's laws



from Leskovec et al., 2020

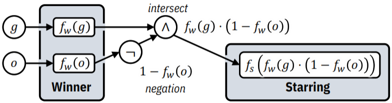
- Real-valued logic on latent propositions (latents), an array of truth bounds that describes any subset of entities.
- Learned Skolem function maps latents of singleton Oscar to latents of maximal subset of Oscar winners, and similarly for GoldenGlobe
- Complement of a subset is logical negation of latents
- Intersection of subsets is logical conjunction of latents
- q(T) = 1 - D(T; A) measures logic satisfiability of candidate answer followed by KNN.

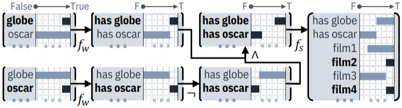


from Gray et al., 2021

Logic Embeddings for Complex Query Answering

- Computation graph for t-norm for intersect.
- Nodes are truth vectors that identify entity subsets.
- Embedding query logic reduces to a simple vectorized calculation.
- Propositions identify features, relations substitute propositions, negation flips truths, intersect retains common features.
- The final query embedding is closer to propositions of film2/4 (answer set) than film1/3.

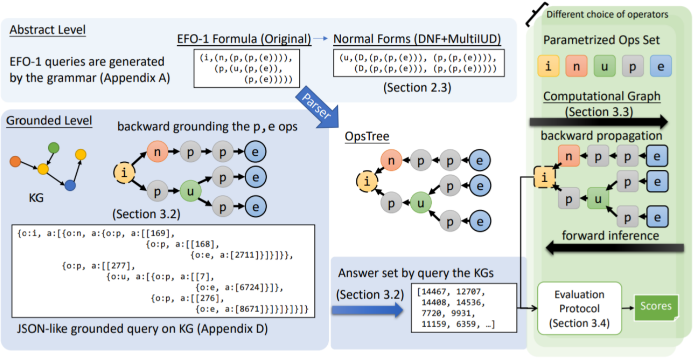




from Gray et al., 2021

Benchmarking the Combinatorial Generalizability of Complex Query Answering on Knowledge Graphs

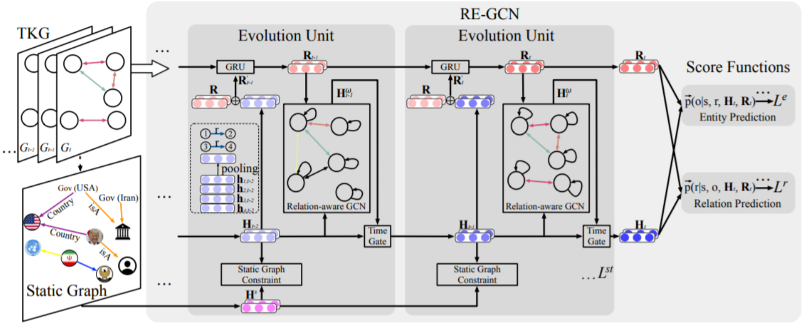
- Benchmarking querygeneration for training and validation
- Different tree-based representations of queries lead to different model performance



from Song et al., 2021

Temporal Knowledge Graph Reasoning Based on Evolutional Representation Learning

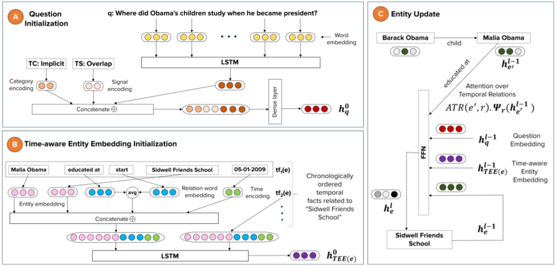
- Temporal KG (TKG) predicts facts via Recurrent Evolution network based on Graph Convolution Network (RE-GCN)
- Temporal component is implemented via autogressive GRU module over structural snapshot-based embeddings



from Cheng et al., 2021

Complex Temporal Question Answering on Knowledge Graphs

- Exaqt processed temporal intents in questions over KGs
- First, it computes question-relevant compact subgraphs within the KG, and judiciously enhances them with pertinent temporal facts, using Group Steiner Trees and fine-tuned BERT models.
- Secondly, relational graph convolutional networks (R-GCNs) with temporal attention and dynamic entity embeddings.



from Weikum et al., 2021

References

- Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, and Oksana Yakhnenko. Translating embeddings for modeling multi-relational data. In Advances in neural information processing systems, 2787–2795. 2013.
- Bishan Yang, Wen-tau Yih, Xiaodong He, Jianfeng Gao, and Li Deng. Embedding entities and relations for learning and inference in knowledge bases. arXiv preprint, 2014.
- Théo Trouillon, Johannes Welbl, Sebastian Riedel, Éric Gaussier, and Guillaume Bouchard. Complex embeddings for simple link prediction. In International Conference on Machine Learning, 2071–2080. 2016.
- Maximilian Nickel, Lorenzo Rosasco, Tomaso A Poggio, et al. Holographic embeddings of knowledge graphs. In AAAI, 1955–1961. 2016.
- Tim Dettmers, Pasquale Minervini, Pontus Stenetorp, and Sebastian Riedel. Convolutional 2d knowledge graph embeddings. In Procs of AAAI. 2018.

References

- Stanovsky, Gabriel, Julian Michael, Luke Zettlemoyer, and Ido Dagan. "Supervised open information extraction." In Proceedings of the 2018 NACL, pp. 885-895. 2018.
- Xiong, Wenhan, Mo Yu, Shiyu Chang, Xiaoxiao Guo, and William Yang Wang. "One-shot relational learning for knowledge graphs." arXiv preprint arXiv:1808.09040. 2018.
- Bosselut, Antoine, Hannah Rashkin, Maarten Sap, Chaitanya Malaviya, Asli Celikyilmaz, and Yejin Choi. "COMET: Commonsense transformers for automatic knowledge graph construction." arXiv preprint arXiv:1906.05317. 2019.
- Liu, Zhibin, Zheng-Yu Niu, Hua Wu, and Haifeng Wang. "Knowledge aware conversation generation with explainable reasoning over augmented graphs." arXiv preprint arXiv:1903.10245. 2019.
- Al-Moslmi, Tareq, Marc Gallofré Ocaña, Andreas L. Opdahl, and Csaba Veres. "Named entity extraction for knowledge graphs: A literature overview." IEEE Access 8, p. 32862-32881. 2020.

References

- Qian, Wei, Cong Fu, Yu Zhu, Deng Cai, and Xiaofei He. "Translating Embeddings for Knowledge Graph Completion with Relation Attention Mechanism." In IJCAI, pp. 4286-4292. 2018.
- Kallumadi, Surya, and William H. Hsu. "Interactive Recommendations by Combining User-Item Preferences with Linked Open Data." In Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization, pp. 121-125. 2018.
- Wang, Hongwei, Fuzheng Zhang, Jialin Wang, Miao Zhao, Wenjie Li, Xing Xie, and Minyi Guo. "Ripplenet: Propagating user preferences on the knowledge graph for recommender systems." In Proceedings of the 27th ACM International Conference on Information and Knowledge Management, pp. 417-426. 2018.
- Dai Quoc Nguyen, Tu Dinh Nguyen, Dat Quoc Nguyen, and Dinh Phung. A Novel Embedding Model for Knowledge Base Completion Based on Convolutional Neural Network. Proc. of NACL, 327–333. 2018.

References

- Ji, S., Pan, S., Cambria, E., Marttinen, P. and Philip, S.Y., 2021. A survey on knowledge graphs: Representation, acquisition, and applications. IEEE Transactions on Neural Networks and Learning Systems.
- Zhang, J., Chen, B., Zhang, L., Ke, X. and Ding, H., 2021. Neural, symbolic and neural-symbolic reasoning on knowledge graphs. AI Open, 2, pp.14-35.
- Hamilton, W.L., Bajaj, P., Zitnik, M., Jurafsky, D. and Leskovec, J., 2018. Embedding logical queries on knowledge graphs. arXiv preprint arXiv:1806.01445.
- Ren, H., Hu, W. and Leskovec, J., 2020. Query2box: Reasoning over knowledge graphs in vector space using box embeddings. arXiv preprint arXiv:2002.05969.

References

- Arakelyan, E., Daza, D., Minervini, P. and Cochez, M., 2020. Complex query answering with neural link predictors. arXiv preprint arXiv:2011.03459.
- Ren, H. and Leskovec, J., 2020. Beta embeddings for multi-hop logical reasoning in knowledge graphs. arXiv preprint arXiv:2010.11465.
- Bakel, R.V., Aleksiev, T., Daza, D., Alivanistos, D. and Cochez, M., 2020, September. Approximate Knowledge Graph Query Answering: From Ranking to Binary Classification. In International Workshop on Graph Structures for Knowledge Representation and Reasoning (pp. 107-124). Springer, Cham.
- Daza, D. and Cochez, M., 2020. Message passing query embedding. arXiv preprint arXiv:2002.02406.
- Kotnis, B., Lawrence, C. and Niepert, M., 2020. Answering complex queries in knowledge graphs with bidirectional sequence encoders. CoRR, abs/2004.02596.

References

- Luus, F., Sen, P., Kapanipathi, P., Riegel, R., Makondo, N., Lebese, T. and Gray, A., 2021. Logic Embeddings for Complex Query Answering. arXiv preprint arXiv:2103.00418.
- `https://github.com/snap-stanford/KGReasoning`
- Alivanistos, D., Berrendorf, M., Cochez, M. and Galkin, M., 2021. Query Embedding on Hyper-relational Knowledge Graphs. arXiv preprint arXiv:2106.08166.
- Wang, Z., Yin, H. and Song, Y., 2021. Benchmarking the Combinatorial Generalizability of Complex Query Answering on Knowledge Graphs. arXiv preprint arXiv:2109.08925.
- Sap, M., Le Bras, R., Allaway, E., Bhagavatula, C., Lourie, N., Rashkin, H., Roof, B., Smith, N.A. and Choi, Y., 2019, July. Atomic: An atlas of machine commonsense for if-then reasoning. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, No. 01, pp. 3027-3035).

References

- Li, Z., Jin, X., Li, W., Guan, S., Guo, J., Shen, H., Wang, Y. and Cheng, X., 2021. Temporal Knowledge Graph Reasoning Based on Evolutional Representation Learning. arXiv preprint arXiv:2104.10353.
- Jia, Z., Pramanik, S., Saha Roy, R. and Weikum, G., 2021, October. Complex Temporal Question Answering on Knowledge Graphs. In Proceedings of the 30th ACM International Conference on Information & Knowledge Management (pp. 792-802).