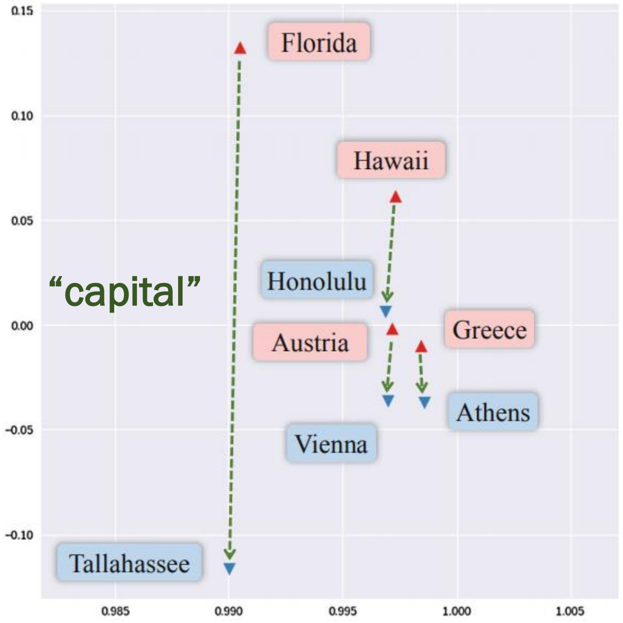
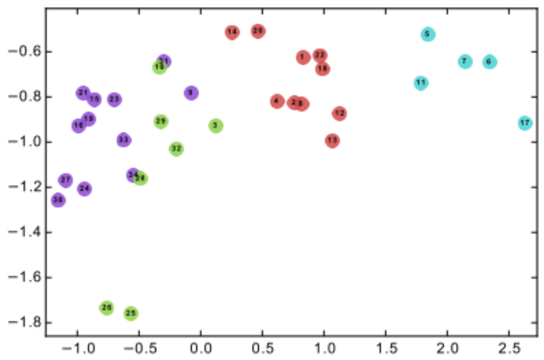
Part II: Link Prediction

Zequn Sun

Knowledge Graph Representation Learning

- Embed the discrete symbolic representations of KGs into continuous vector space.
- Why representation learning?
  - Good features are essential for successful machine learning.
  - Mitigate symbolic heterogeneities.
  - Build a unified semantic space serving knowledge-driven applications.
- KG representation learning **vs.** network embedding

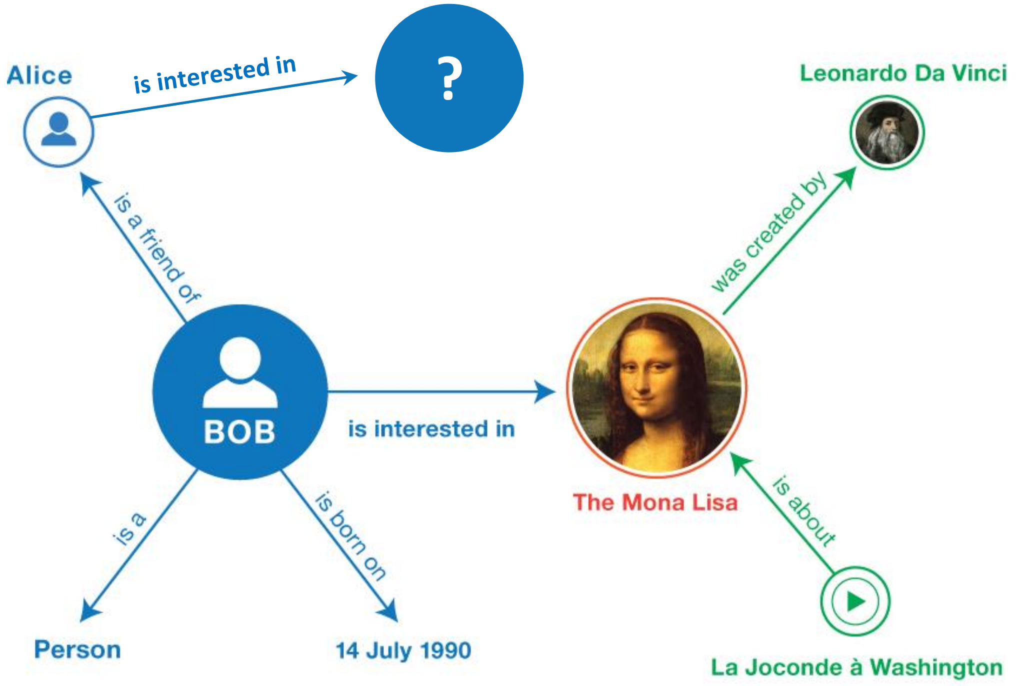




relational structures vs. topological structures

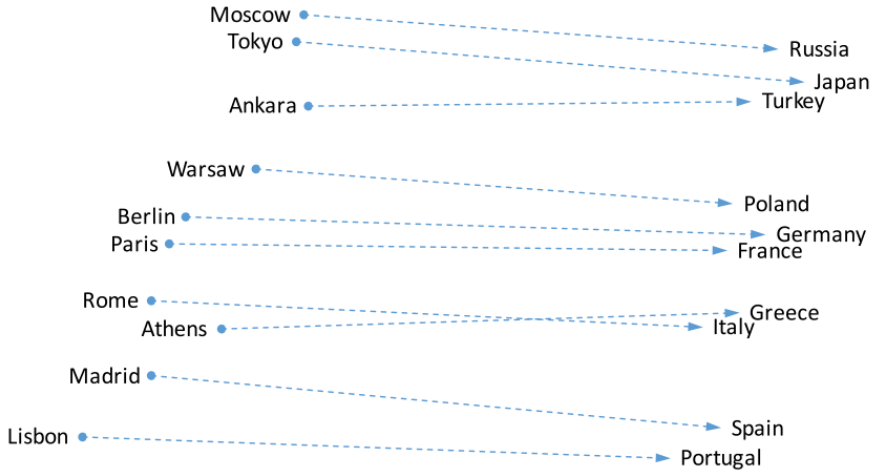
Link Prediction

- Infer the missing relation triples in a KG.
- Input
  - query (head entity, *relation*, ?) or
  - query (?, *relation*, tail entity)
- Scoring function
  - Measure the plausibility of relation triples.
  - The learning objective is to differentiate between positive triples and negatives.
- Representative models
  - Translation-based models & semantic matching models
  - Deep models
  - Non-Euclidean models

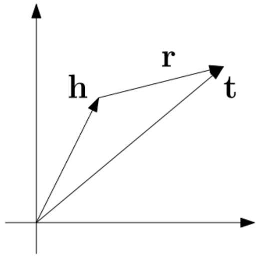


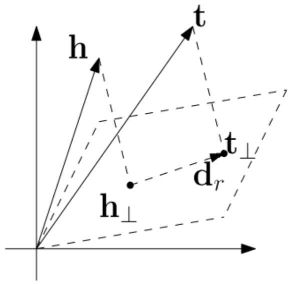
Translation-based Models: TransE/TransH/TransR

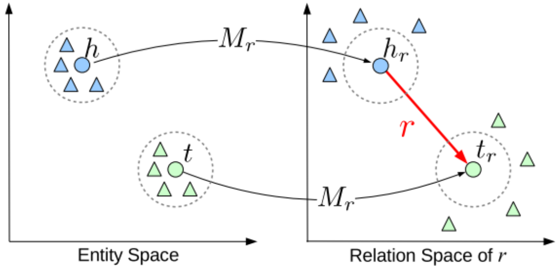
- Translation-based KG embedding methods interpret a relation as a translation vector from its head entity to tail entity.
  - Head entity vector – tail entity vector ≈ *relation* vector
- Inspired by word embeddings
  - Russia – Moscow ≈ *capital*
  - France – Paris ≈ *capital*



- Where to translate relation embeddings?
  - TransE (Bordes et al., NIPS-2013): in the joint vector space
  - TransH (Wang et al., AAAI-2014): on the relation-specific hyperplane
  - TransR (Lin et al., AAAI-2015): in the relation-specific space

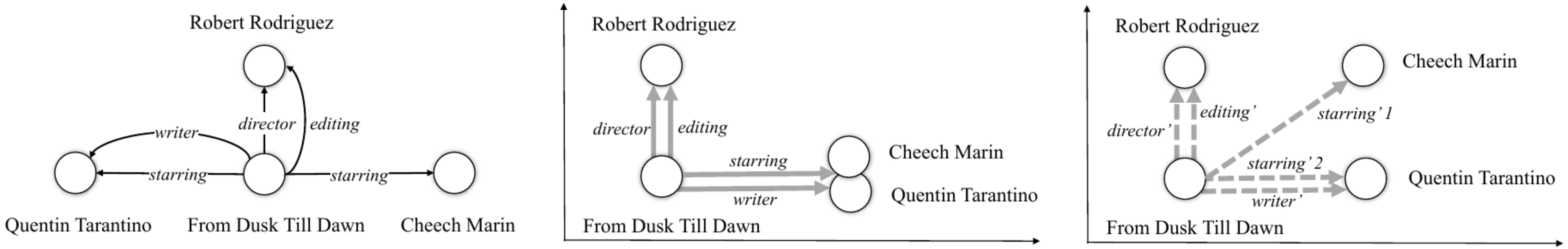






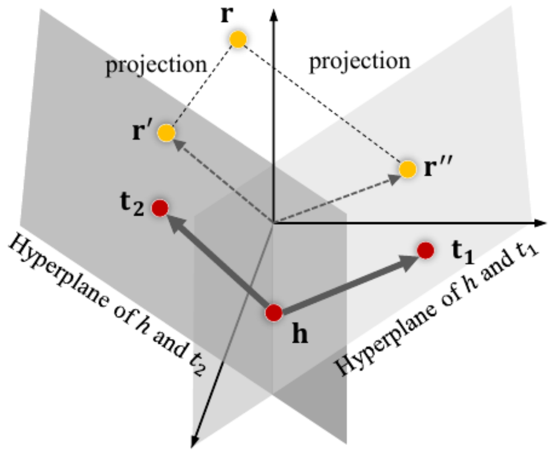
Translation-based Models: TransEdge (Sun et al., ISWC-2019)

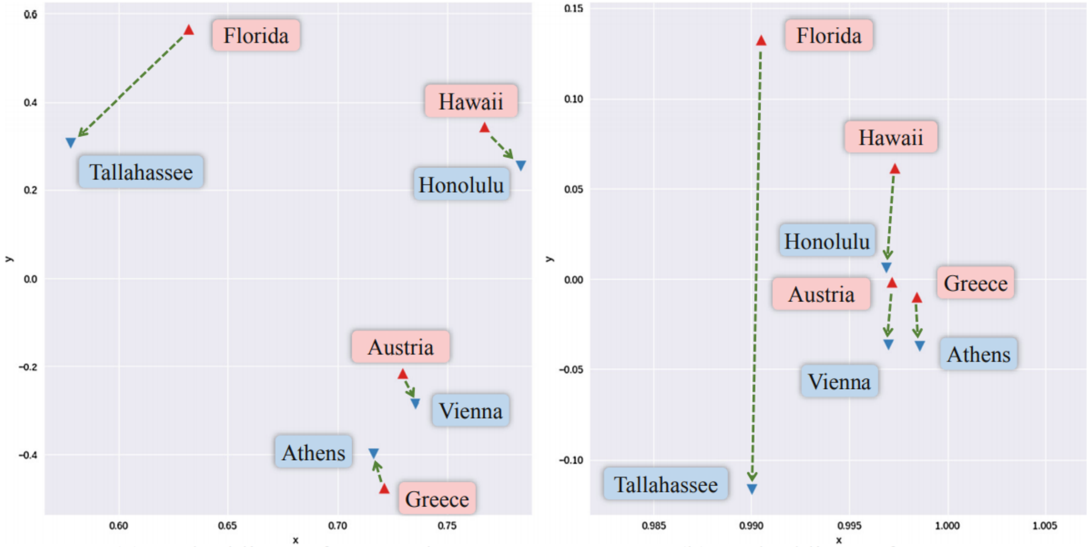
- Contextualize relation representations in terms of specific head-tail entity pairs.



(a) graph-structured relational facts (b) relation-level translation (c) edge-level translation

- Relation-contextualized translation

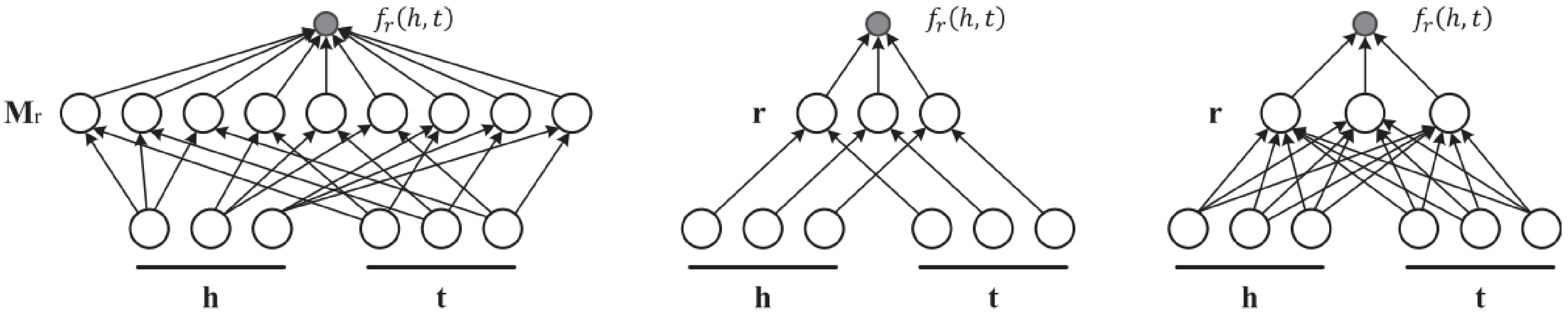




(a) Embeddings of TransEdge (b) Embeddings of MTransE

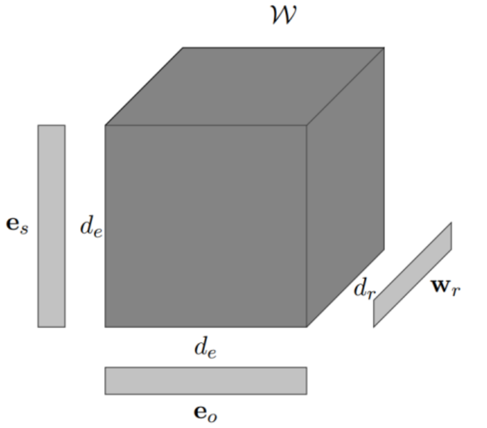
Semantic Matching Models: RESCAL/DistMult/HolE/TuckER

- Semantic matching KG embedding models exploit similarity-based functions to measure the plausibility of relation triples.
  - Bilinear embeddings: RESCAL (Nickel et al., ICML-2011), DistMult (Yang et al., ICLR-2015)
  - Holographic embeddings: HOLE (Nickel et al., AAAI-2016)



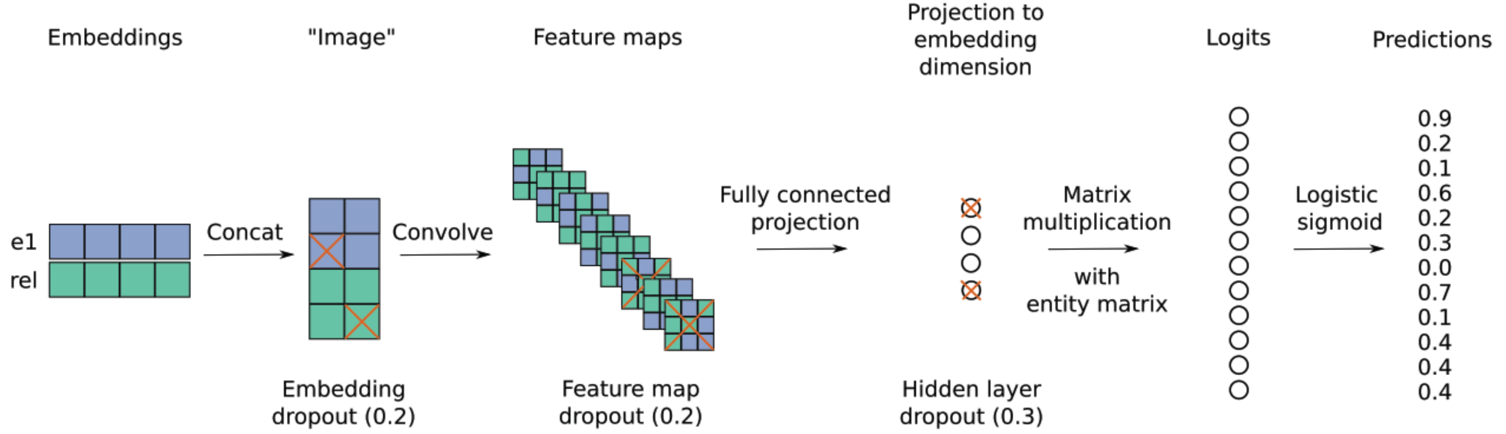
(a) RESCAL. (b) DistMult. (c) HolE.

  - Tensor decomposition embeddings: TuckER (Balazevic et al., EMNLP-2019)
    - Binary tensor representation of a KG

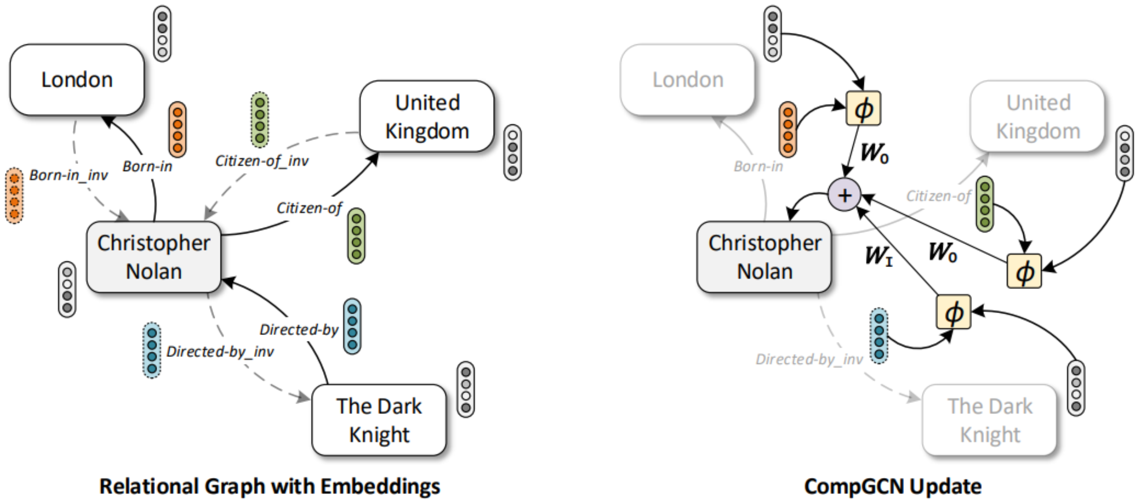


Deep Models: ConvE/CompGCN

- Convolutional neural networks: ConvE (Dettmers et al., AAAI-2018)
  - Model the interactions between entities and relations by convolutional operations over 2D shaped embeddings.



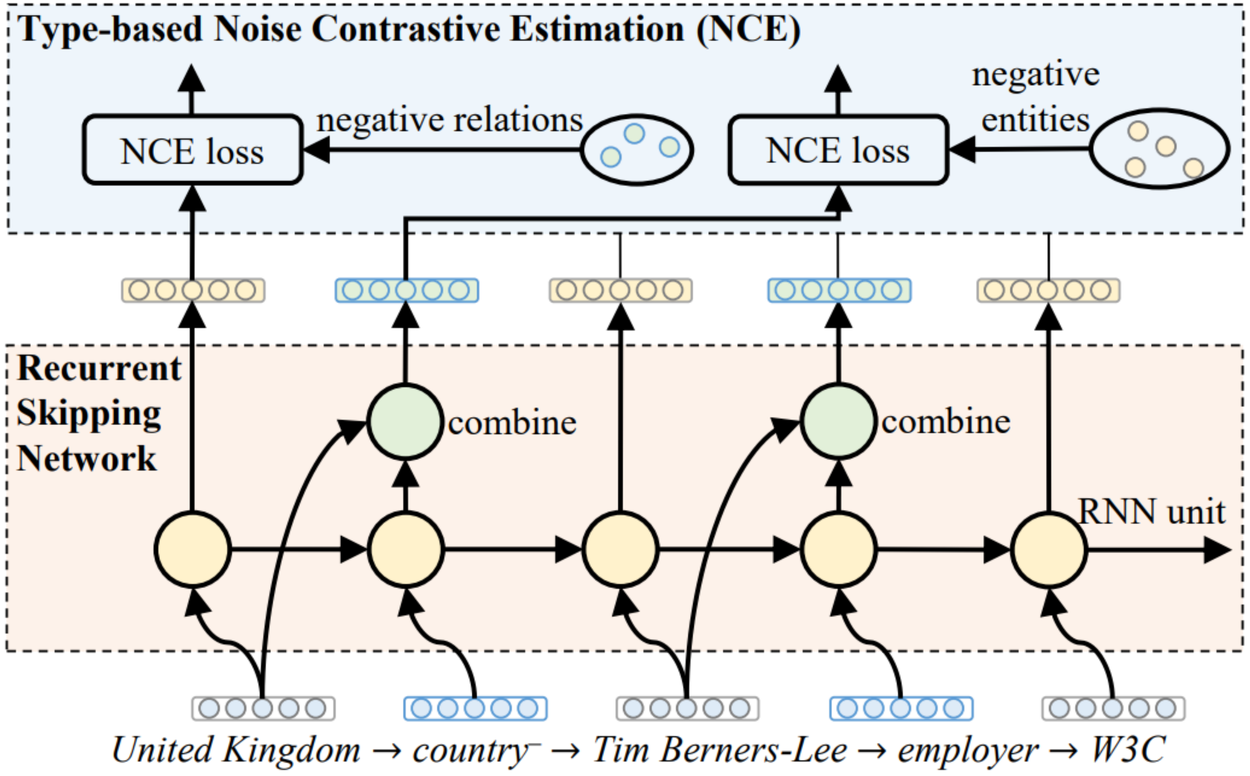
- Relational GNN: CompGCN (Vashishth et al., ICLR-2020)
  - Use entity-relation composition operations for neighborhood aggregation.



Deep Models: RSN (Guo et al., ICML-2019)

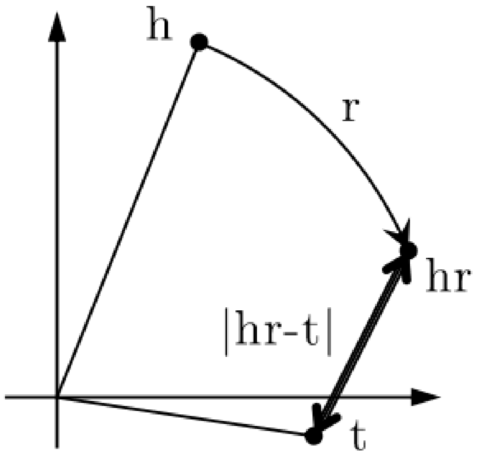
- Recurrent skipping network
  - Entities and relations appear alternately in a path.
  - A path consists of many relational triples as basic units.
  - RNNs overlook element types and local units of triples.

  - Type-based NCE
    - Entity prediction
    - Relation prediction
  - Entities participate in predicting
    - the subsequent relations
    - the relational object entities

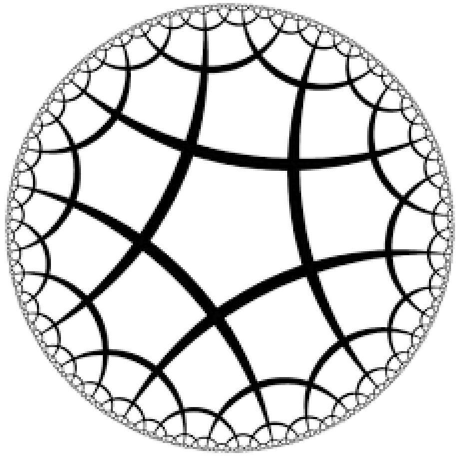


Non-Euclidean Methods: RotatE/ATTH

- Complex embeddings: RotatE (Sun et al., ICLR-2019)
  - Define each relation as a rotation from the head entity to the tail entity in the complex vector space.
  - Propose self-adversarial negative sampling for embedding learning.



- Hyperbolic embeddings: ATTH (Chami et al., ACL-2020)
  - The amount of space covered by hyperbolic geometry grows exponentially with the radius.
  - Capture hierarchical and logical patterns at low dimensions.
  - Lift existing embedding techniques (e.g., relation translation and rotation) into hyperbolic space.



Link Prediction Datasets

- FB15K and WN18 (Bordes et al., NIPS-2013)
  - Subsets of Freebase and WordNet, respectively
  - Suffer from the test data “leakage” issue due to reverse triples.
    - (A Room With A View, *film/directed_by*, James Ivory)
    - (James Ivory, *director/film*, A Room With A View)

Dataset	#entities	#relations	#train	#valid	#test
FB15k	14,951	1,345	483,142	50,000	59,071
FB15k-237	14,541	237	272,115	17,535	20,046
WN18	40,943	18	141,442	5,000	5,000
WN18RR	40,943	11	86,835	3,034	3,134

- Current benchmark (reverse triples removed, more challenging)
  - FB15K-237 (Toutanova and Chen, CVSC-2015)
  - WN18RR (Dettmers et al., AAAI-2018)
- Realistic Re-evaluation of Link Prediction (Akrami et al., SIGMOD-2020)
  - Existing embedding models would have been biased toward learning trivial patterns for link prediction.
  - **Link prediction is still a difficult task without truly effective automated solution.**

End of Part II