Link Prediction in Knowledge Graphs Using Graph Embedding

July 7, 2020

Emil Bækdahl

Department of Computer Science

Aalborg University

AALBORG UNIVERSITY

DENMARK

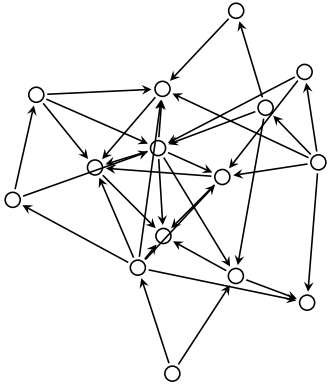
Quick Note

Agenda

Link Prediction
The Problem Setting

- True knowledge

$$\mathcal{G} = (V, E).$$



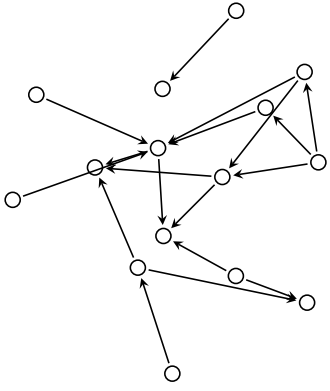
Link Prediction
The Problem Setting

- True knowledge

$$\mathcal{G} = (V, E).$$

- Observed knowledge

$$\mathcal{G}' = (V, E') \text{ where } E' \subset E.$$



Link Prediction
The Problem Setting

- True knowledge

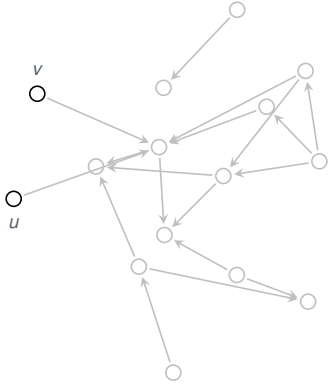
$$\mathcal{G} = (V, E).$$

- Observed knowledge

$$\mathcal{G}' = (V, E') \text{ where } E' \subset E.$$

- Given $(u, v) \in V \times V$, is it true that

$$(u, v) \in E.$$

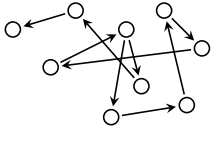


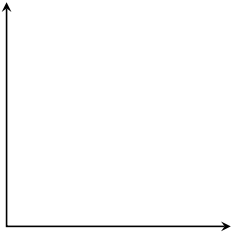
Link Prediction as an Embedding Problem

- An *embedding* is a representation of one mathematical structure in the terms of another, e.g. *graphs* represented as *vectors*.

Link Prediction as an Embedding Problem

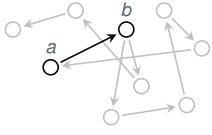
Learning Embeddings

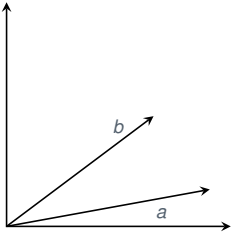




Link Prediction as an Embedding Problem

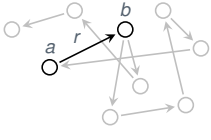
Learning Embeddings

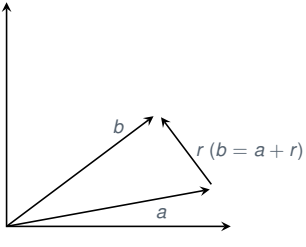




Link Prediction as an Embedding Problem

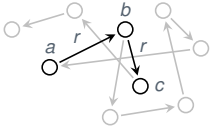
Learning Embeddings

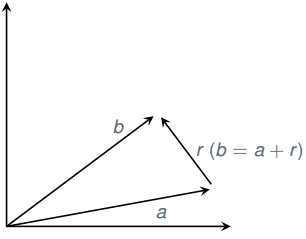




Link Prediction as an Embedding Problem

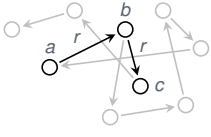
Learning Embeddings

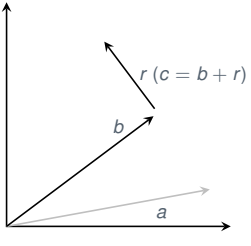




Link Prediction as an Embedding Problem

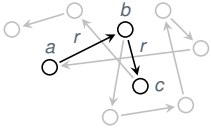
Learning Embeddings

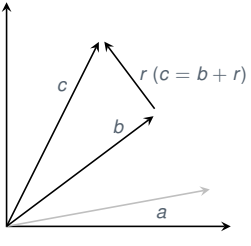




Link Prediction as an Embedding Problem

Learning Embeddings

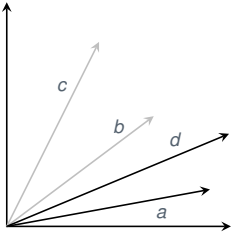




Link Prediction as an Embedding Problem

Inferring with Embeddings

- Are *d* and *a* related with *r*?



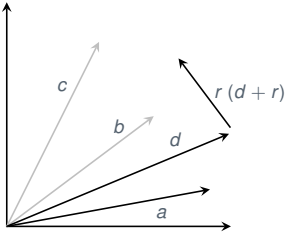
Link Prediction as an Embedding Problem

Inferring with Embeddings

- Are d and a related with r?
- We use the distance

$$(d + r) - a$$

as plausibility measure.



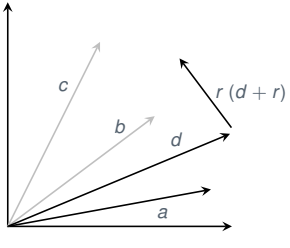
Link Prediction as an Embedding Problem

Inferring with Embeddings

- Are d and a related with r? Probably not!
- We use the distance

$$(d + r) - a$$

as plausibility measure.



Link Prediction as an Embedding Problem

Embedding Semantic Hierarchies

- Hierarchy-Aware Knowledge Graph Embedding (HAKE)

Link Prediction as an Embedding Problem

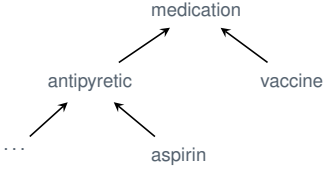
Embedding Semantic Hierarchies

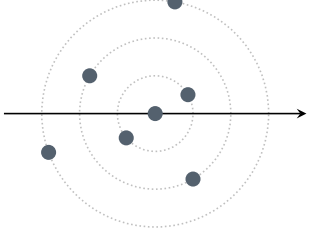
- Hierarchy-Aware Knowledge Graph Embedding (HAKE)
- Knowledge naturally exhibits hierarchies.

Link Prediction as an Embedding Problem

Embedding Semantic Hierarchies

- Hierarchy-Aware Knowledge Graph Embedding (HAKE)
- Knowledge naturally exhibits hierarchies.
- Polar coordinates help encode entities on different hierarchical levels.

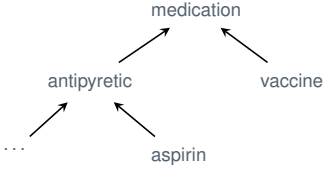


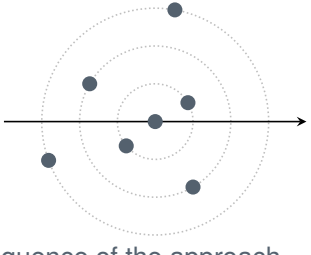


Link Prediction as an Embedding Problem

Embedding Semantic Hierarchies

- Hierarchy-Aware Knowledge Graph Embedding (HAKE)
- Knowledge naturally exhibits hierarchies.
- Polar coordinates help encode entities on different hierarchical levels.





- Hierarchies are not explicitly embedded but a consequence of the approach.

Experiment Results

- Replication: 0.172% to -11.7% deviation.

Experiment Results

- Replication: 0.172% to -11.7% deviation.
- Domain-specific data: close to 50% of replication performance (at best).

Experiment Results
Visualising Embeddings

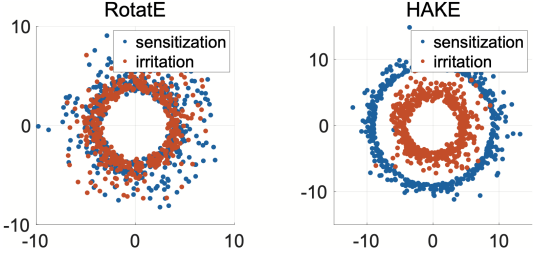


Figure: HAKE Paper

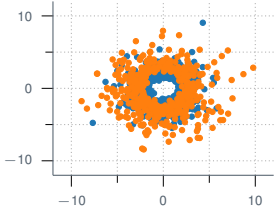


Figure: Domain-Specific

Improving Link Prediction
Why (Not) Embeddings?

- Graph embeddings are fast and efficient for inferring.

Improving Link Prediction
Why (Not) Embeddings?

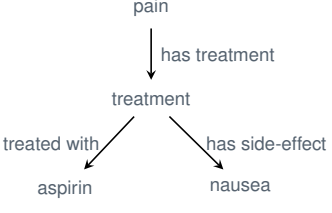
- Graph embeddings are fast and efficient for inferring.
- Flawed evaluation methods due to reverse and Cartesian product relations.

Improving Link Prediction
Why (Not) Embeddings?

- Graph embeddings are fast and efficient for inferring.
- Flawed evaluation methods due to reverse and Cartesian product relations.
- Transparency, explainabilty, and interesting graph properties are lost.

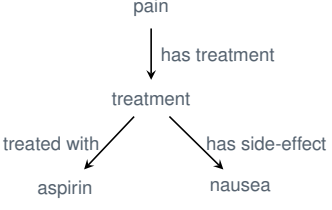
Improving Link Prediction
Then What?

- Knowledge base representations may be inappropriate (Cartesian product relations).



Improving Link Prediction
Then What?

- Knowledge base representations may be inappropriate (Cartesian product relations).



- Similarity measures may give better results, transparency, and explainabilty.

Thank you!

AALBORG UNIVERSITY
DENMARK