Lecture

Knowledge-based Systems

Lecture 08: Common Sense Knowledge and Factual Knowledge

Dr. Mohsen Mesgar

Universität Duisburg-Essen

Recall ...

(Artificial) intelligence: the ability acquire knowledge and adopt that knowledge to new environment and tasks to achieve goals.

Knowledge base (KB) is the core of KB systems

Two types of KBs (based on knowledge representations)

Symbolic

Connectionist

Evaluating KBs using semantic relatedness tasks

Semantic relations between words

Multilingual Knowledge bases

Any other open questions?

In this lecture, you learn about …

- Commonsense knowledge
- Factual knowledge
- How to capture these types of knowledge

Today

Universität Duisburg Essen
Offen im Denken

Commonsense Knowledge

Motivational questions

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

- Who is taller, Prince William or his baby son Prince George?
- Can you make a salad out of a polyester shirt?
- If you stick a pin into a carrot, does it make a hole in the carrot or in the pin?

How do we know all the answers?

Commonsense Knowledge

- knowledge which is assumed that every ordinary human knows about
  - the basic level of practical **knowledge** and **reasoning**
  - concerning everyday **situations** and **events**
  - that are commonly shared among most people

- Example:
  - if you meet a person the first time, you know for sure that this person has parents without ever having talked to or about her.
  - This is because you know a person could not exist if she had no parents
- Example, it’s ok to keep the closet door open, but it’s not ok to keep the fridge door open, as the food inside might go bad.

Why is common sense knowledge important?

- Essential for humans to live and interact with each other in a reasonable and safe way.
- Essential for AI to understand and interact with human

Commonsense in Children

Implicit Message

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

- Commonsense makes a system (human or machine) to understand hidden information provided by the speaker or writer.
  - For example sarcasm, etc.

- Anyone who has seen the unforgettable horse's head scene in The Godfather immediately realizes what's going on

https://www.youtube.com/watch?v=VC1_tdnZq1A

- It's not just that it's unusual to see a severed horse head, it's clear that Tom Hagen is sending Jack Woltz a message – if I can decapitate your horse, I can decapitate you; cooperate, or else.

Winograd Schema - Contributions

- Small binary questions that any English speaking individual should be able to answer
- Questions involve thinking and cannot be answered even with access to a large corpus
- Need to disambiguate pronouns based on commonsense
- PROS:
  - Easy to understand questions for humans
  - No trickery or evasiveness can work
  - Can be evaluated quickly and automatically

Penn Engineering

Taken from: https://www.seas.upenn.edu/~cis700dr/Spring19/presentation_slides/Winograd.pptx

Winograd Schema – Example 1

I poured water from the bottle into the cup until **it** was **full**. What was **full**?

I poured water from the bottle into the cup until **it** was **empty**. What was **empty**?

Penn Engineering

Taken from: https://www.seas.upenn.edu/~cis700dr/Spring19/presentation_slides/Winograd.pptx

Winograd Schema – Example 1

I poured water from the bottle into the cup until **it** was **full**. What was **full**? Cup

I poured water from the bottle into the cup until **it** was **empty**. What was **empty**? Bottle

Penn Engineering

Taken from: https://www.seas.upenn.edu/~cis700dr/Spring19/presentation_slides/Winograd.pptx

Winograd Schema – Example II

Paul tried to call George on the phone but **he** wasn't **successful**. Who was not **successful**?

Paul tried to call George on the phone but **he** wasn't **available**. Who was not **available**?

Penn Engineering

Taken from: https://www.seas.upenn.edu/~cis700dr/Spring19/presentation_slides/Winograd.pptx

Winograd Schema – Example II

Paul tried to call George on the phone but **he** wasn't **successful**. Who was not **successful**? Paul

Paul tried to call George on the phone but **he** wasn't **available**. Who was not **available**?
George

Penn Engineering

Taken from: https://www.seas.upenn.edu/~cis700dr/Spring19/presentation_slides/Winograd.pptx

Winograd Schema - Properties

- Two similar parties (Male/Female/Object)
- Pronoun used that could refer to either
- Question asked that involves disambiguation of the pronoun
- Sentence has a special word which when replaced by an alternate word changes the answer to the question
- Special and alternate word need not be opposites

Penn Engineering

Taken from: https://www.seas.upenn.edu/~cis700dr/Spring19/presentation_slides/Winograd.pptx

What do we need to solve WS?

- Lexical knowledge – Language models? Distributional representations?
- Structural understanding – Syntax? Parsing?
- Background knowledge – Knowledge bases?

Penn Engineering

Taken from: https://www.seas.upenn.edu/~cis700dr/Spring19/presentation_slides/Winograd.pptx

What do we need to solve WS?

- Lexical knowledge – Language models? Distributional representations?
- Structural understanding – Syntax? Parsing?
- Background knowledge – Knowledge bases?

John asked **George** a question, but he was reluctant to [**answer**/**repeat**] it. Who was reluctant to [**answer**/**repeat**] it?

George asked **John** a question, but he was reluctant to [**answer**/**repeat**] it. Who was reluctant to [**answer**/**repeat**] it?

- Meaning of the words or the sentence
- Knowledge that a question needs to be answered
- Sentence structure to decide who asked the question and who should answer it

Penn Engineering

Taken from: https://www.seas.upenn.edu/~cis700dr/Spring19/presentation_slides/Winograd.pptx

Solving the WS Problems

- "Resolving Complex Cases of Definite Pronouns" (Rahman & Ng; 2012)
  - Rank the two candidate antecedents for the target pronoun
  - Features based on complex linguistic rules are used
  - http://www.hlt.utdallas.edu/~vince/papers/emnlp12.pdf
- "Solving hard coreference problems" (Peng & Kashabi & Roth; 2015)
  - WS as coreference resolution or clustering
  - http://cogcomp.org/page/publication_view/762
- "The Winograd Schema Challenge and Reasoning about Correlation" (Bailey et al; 2015)
  - Determine correlation between different phrases in the sentences after substituting the correct answer
  - https://www.cs.utexas.edu/~vl/papers/wsc_SS7.pdf

Penn Engineering

Taken from: https://www.seas.upenn.edu/~cis700dr/Spring19/presentation_slides/Winograd.pptx

Solving the WS Problems

- "Combing Context and Commonsense Knowledge Through Neural Networks for Solving Winograd Schema Problems" (Liu et al; 2017)
  - Learn word representations based on current context and knowledge base
  - Predict the correct antecedent for the pronoun either based on semantic similarity or using a neural network
  - https://www.aaai.org/ocs/index.php/SSS/SSS17/paper/viewFile/15392/14552
- "Addressing the Winograd Schema Challenge as a Sequence Ranking Task" (Opitz & Frank; 2018)
  - Replace pronoun by antecedent and generate a plausibility score for each sentence using a neural network
  - http://aclweb.org/anthology/W18-4105

Penn Engineering

Taken from: https://www.seas.upenn.edu/~cis700dr/Spring19/presentation_slides/Winograd.pptx

Winograd schema in other languages

- Chinese - https://cs.nyu.edu/davise/papers/WinogradSchemas/WSChinese.html
- French - http://www.llf.cnrs.fr/winograd-fr
- Japanese - http://arakilab.media.eng.hokudai.ac.jp/~kabura/collection_ja.html

Penn Engineering

Taken from: https://www.seas.upenn.edu/~cis700dr/Spring19/presentation_slides/Winograd.pptx

Challenges with WS

- Coming up with questions!!!!
  - Answers can be too obvious because of properties of the two parties
    - The women stopped taking the pills because they were [**pregnant/carcinogenic**]. Which individuals were [**pregnant/carcinogenic**] ?
    - Women can't be carcinogenic
    - Pills can't be pregnant
    - Such information can be learnt through a large corpus of text via co-occurrence statistics easily and are excluded form WS datasets

Penn Engineering

Taken from: https://www.seas.upenn.edu/~cis700dr/Spring19/presentation_slides/Winograd.pptx

Challenges with WS

- Coming up with questions!!!!
  - Phrases need not be Google-proof
    - Many astronomers are engaged in the search for distant galaxies. They are spread all over the [**earth/universe**]. What are spread all over the [**earth/universe**] ?
    - "Galaxies are spread all over the universe" may be found on the web
      - Direct fact but the system would have to find it and use it
    - "There are many galaxies found throughout the universe" may be found
      - System would have to understand that it means the same as the question
    - Even though they could still be difficult, such questions are excluded in WS datasets

Penn Engineering

Taken from: https://www.seas.upenn.edu/~cis700dr/Spring19/presentation_slides/Winograd.pptx

Challenges with WS

- Coming up with questions!!!!
  - Lack of knowledge in human subjects on whom questions are tested
    - The large ball crashed through the table because it is made up of [**steel/Styrofoam**]. What is made of [steel/Styrofoam]?
    - Some person might not know what is Styrofoam so such questions are excluded
    - Though a system might understand these if they appeared in some large corpus it used

Penn Engineering

Taken from: https://www.seas.upenn.edu/~cis700dr/Spring19/presentation_slides/Winograd.pptx

Summary of Winograd Schema

- Winograd Schema as an alternative test
  - Binary questions involving pronoun disambiguation
  - No deception
  - Easy to evaluate
  - Difficult to come up with questions

Penn Engineering

Taken from: https://www.seas.upenn.edu/~cis700dr/Spring19/presentation_slides/Winograd.pptx

Today

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Factual Knowledge

What is Factual Knowledge?

- Having basic information about a particular subject or discipline.
- Include the terminology and the specific details or elements of a subject.
- Serves as basic building blocks to understand the larger relationships among important information that define a subject.

Examples

- Within an adventure game, students may learn that China is in Asia,
- From the instructions of a game, students may learn that start-up disk refers to a 3.5-inch square disk that fits in a disk drive.
- Capital of Germany.

Factual Knowledge Vs. Common-Sense Knowledge

Factual Knowledge

- Who won the FIFA world cup in 2018?
- Has Tesla Motor released automated car in 2017?

Vs

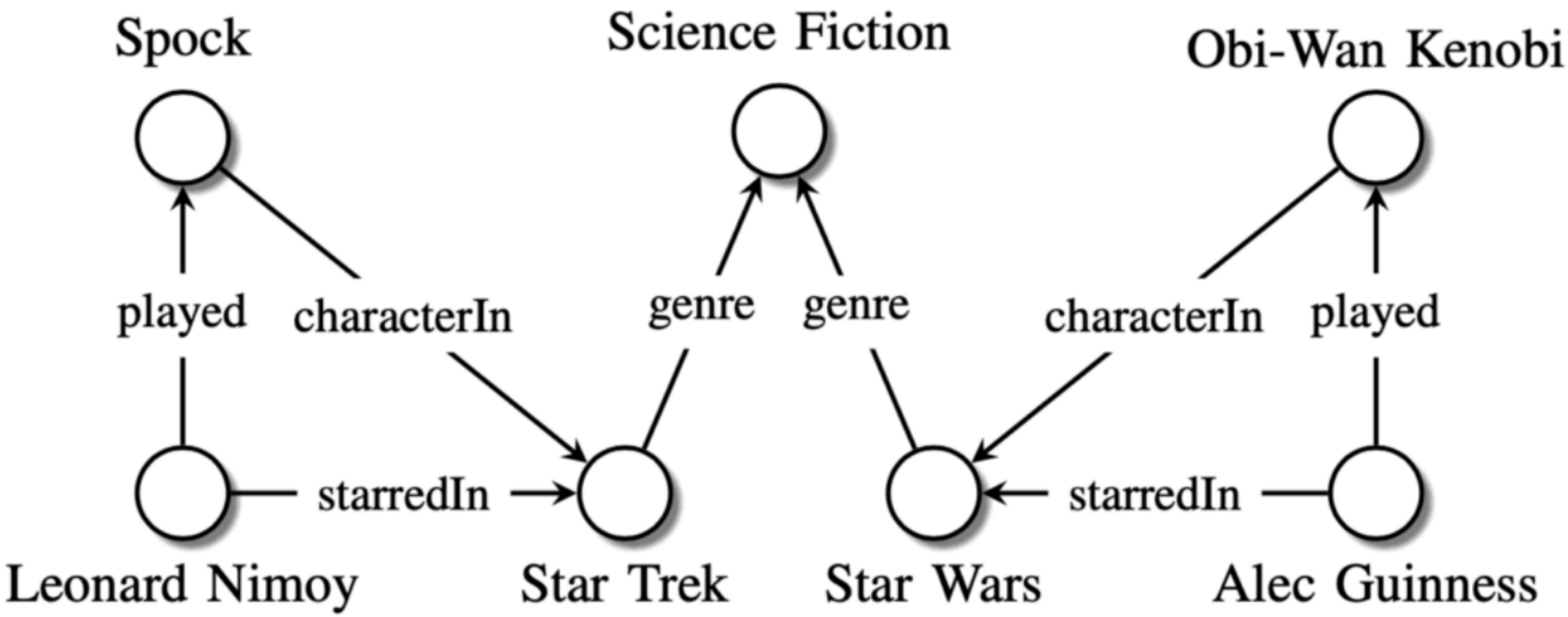
Common sense Knowledge

- Who is taller, Prince William or his baby son Prince George?
- Can you make a salad out of a polyester shirt? I
- If you stick a pin into a carrot, does it make a hole in the carrot or in the pin?

Knowledge Graphs

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken



Entities and Relationships

- Knowledge Graphs (KGs) are a way of structuring information in graph form, by representing **entities** (eg: people, places, objects) as nodes, and **relationships** between entities (eg: being married to, being located in) as edges.

Example : (New York City, type, City) states that the entity New York City has type City.

Facts in KGs

- **Facts** are typically represented as “SPO” triples: *(Subject, Predicate, Object)*.

Example:

(New York City, country, United States)

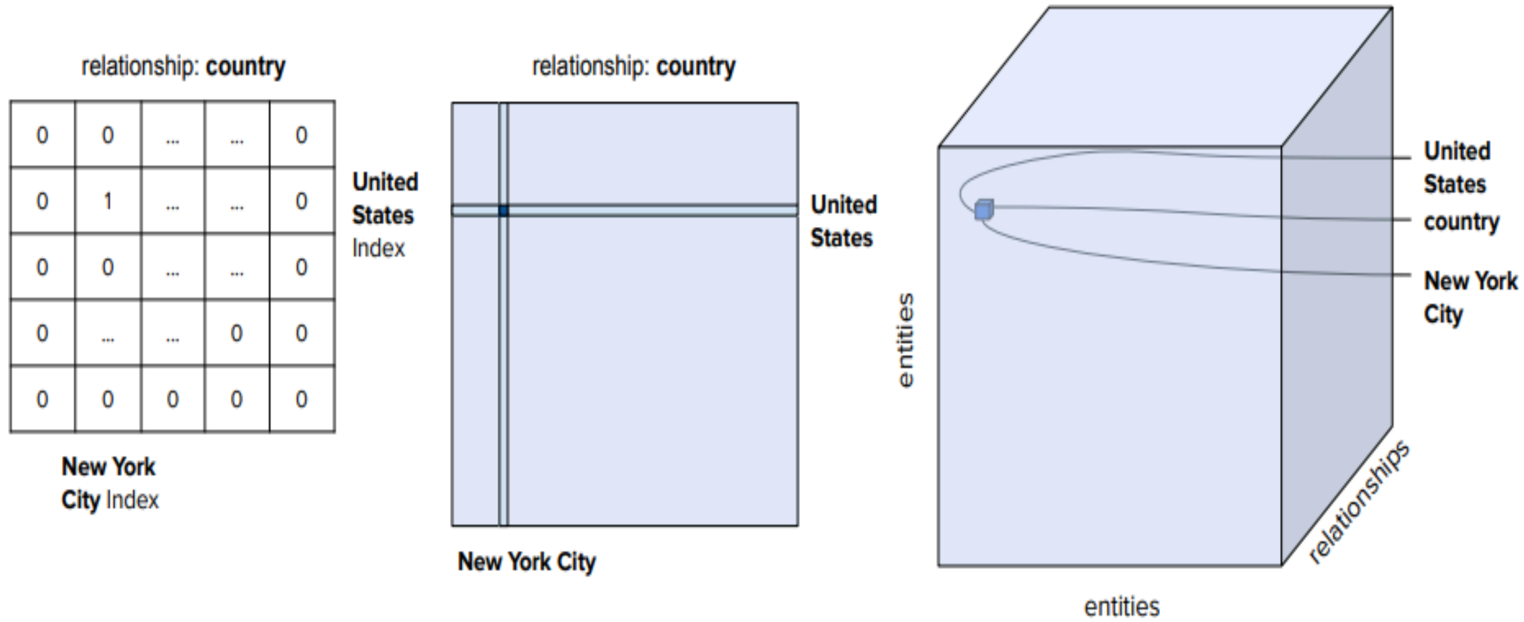
- Adding new facts can be done with the use of logical inference.

(Washington D.C., capital, United States) ➔ (Washington D.C., country, United States).

logic rule: $\forall X,Y$: capital(X,Y) $\Rightarrow$ country(X,Y)).

KGs in 3D Space

Simple way to visualize a knowledge graph is considering it as a 3-order adjacency tensor



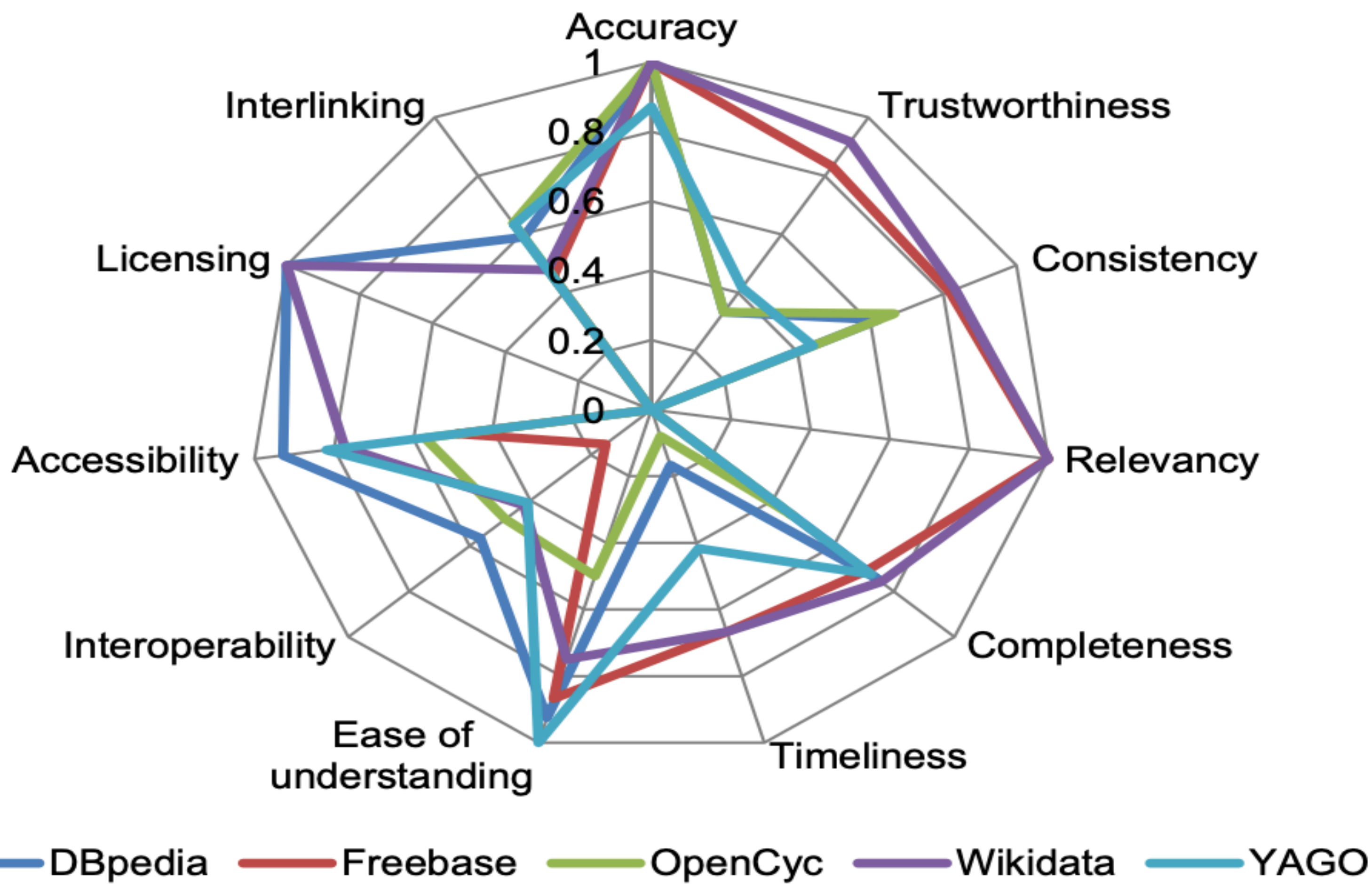
Ways of building Knowledge Graphs

- Manually, by experts or volunteers
- By automatically extracting them from semi-structured text (eg: Wikipedia infoboxes)
- By automatically extracting them from unstructured text (using natural language processing techniques)

KG Datasets

	DBpedia	Freebase	OpenCyc	Wikidata	YAGO
Number of triples	411 885 960	3 124 791 156	2 412 520	748 530 833	1 001 461 792
Number of classes	736	53 092	116 822	302 280	569 751
Number of relations	2819	70 902	18 028	1874	106
No. of unique predicates	60 231	784 977	165	4839	88 736
Number of entities	4 298 433	49 947 799	41 029	18 697 897	5 130 031
Number of instances	20 764 283	115 880 761	242 383	142 213 806	12 291 250
Avg. number of entities per class	5840.3	940.8	0.35	61.9	9.0
No. of unique subjects	31 391 413	125 144 313	261 097	142 278 154	331 806 927
No. of unique non-literals in object position	83 284 634	189 466 866	423 432	101 745 685	17 438 196
No. of unique literals in object position	161 398 382	1 782 723 759	1 081 818	308 144 682	682 313 508

Key features

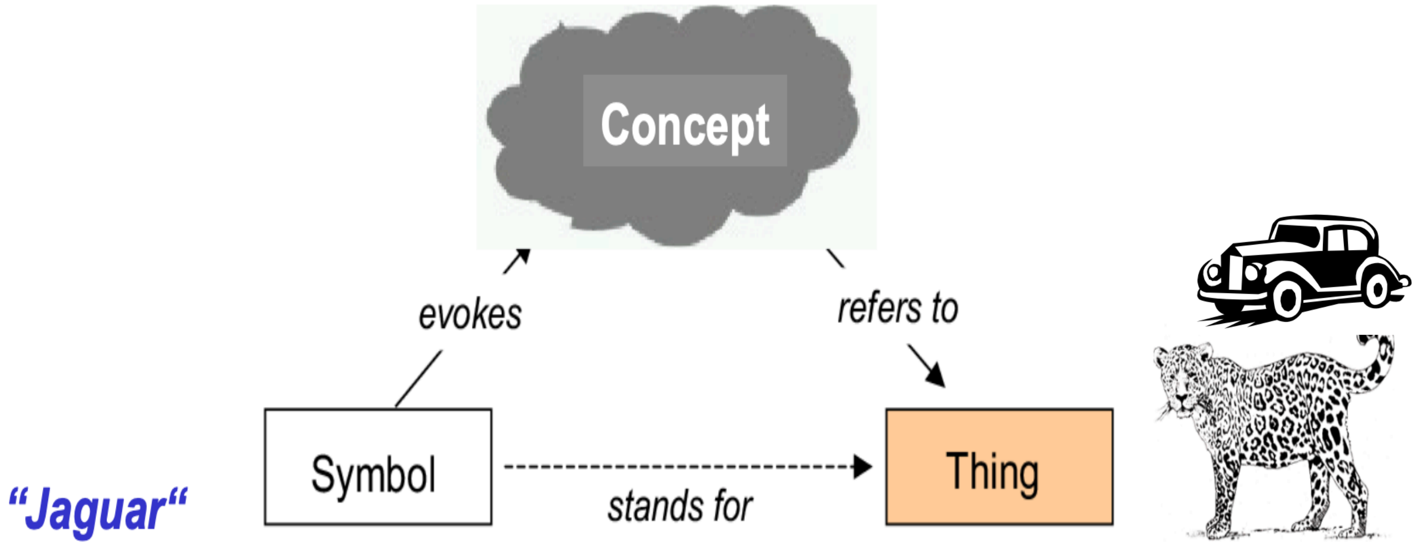


Ontology

UNIVERSITÄT DUISBURG ESSEN
Offen im Denken

- Describes the common words, concepts and relationships between concepts used to describe and represent an area of knowledge.

- Humans require words (or at least symbols) to communicate efficiently. The mapping of words to things is indirect. We do it by creating concepts that refer to things.



Ontology Vs. Knowledge Graphs

ontology + data = knowledge graph

Ontologies

are *generalized* data models, meaning that they only model *general* types of things that share certain properties, but don't include information about *specific* individuals

Example:

Book → has author → Author

Knowledge Graph

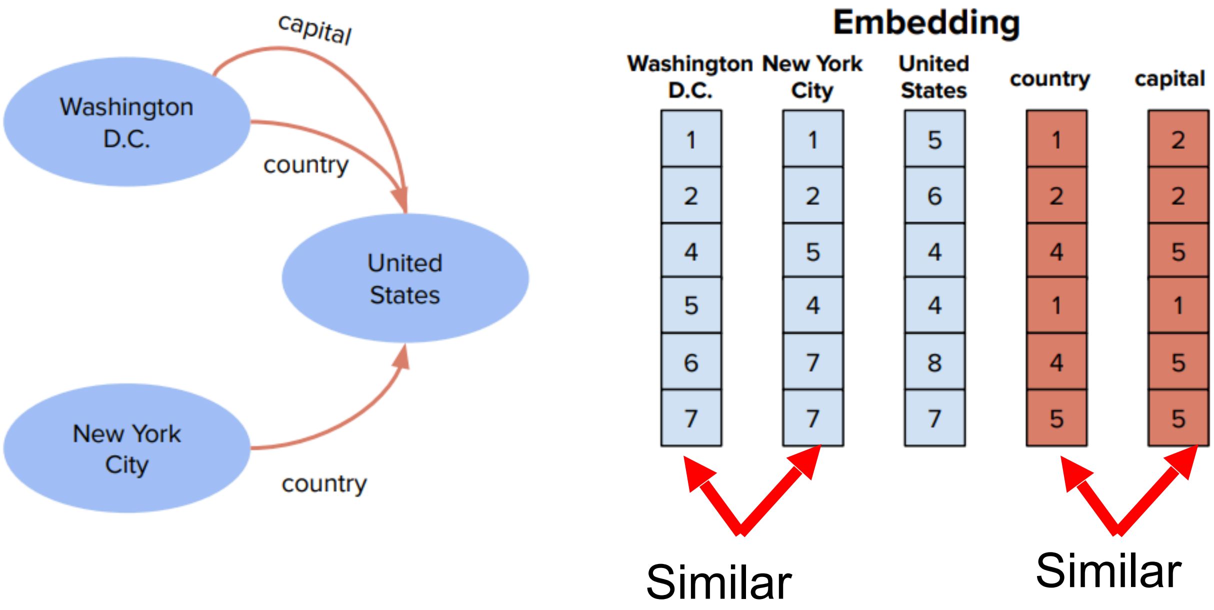
Use ontology as a framework

Example:

"To kill a Mocking Bird" → has author –> "Harper Lee"

Knowledge Graph Embeddings (KGE)

- Generation of vector representations of the elements that form a knowledge graph



Knowledge Graph Embeddings (KGE)

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

- Embeddings are generated in such a way to capture latent properties of the semantics in the knowledge graph:

Similar entities and similar relationships will be represented with similar vectors (as shown in the figure of previous slide)

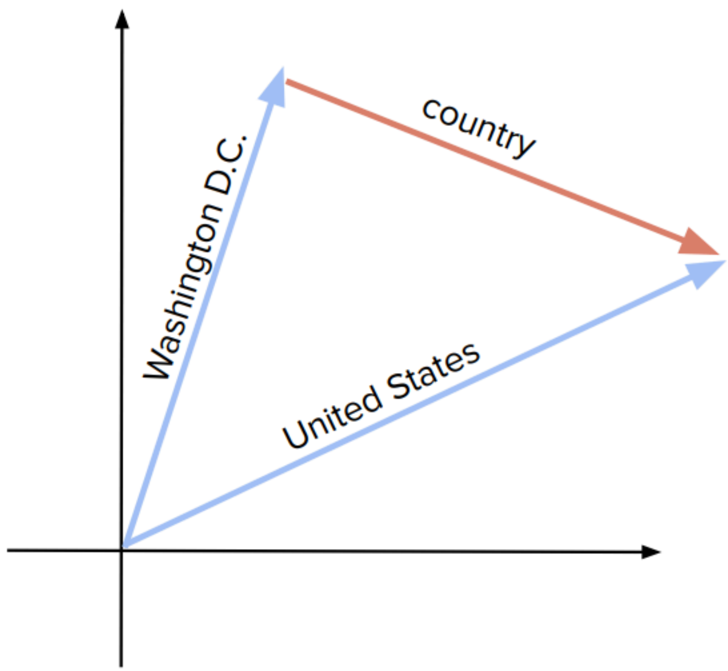
- KG embeddings can be used for link prediction, since they show interesting predictive abilities and are not directly constrained by logical rules

Example: we might not know the meaning of the entity New York City, but it can be inferred from its topic by looking at closest entities in the geometric space (i.e. Washington D.C. and United States).

TranE (variant of KGE)

- Trained on a large scale data set with 1M entities, 25k relationships and more than 17M training samples.

- The driving idea is that the sum of the subject vector with the predicate vector should generate the vector representation of the object as output



Vec(Washington D.C.) + Vec(country) ≈ Vec(United States)
subject predicate Object

Understanding of larger relationships

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

Having basic information (factual knowledge), complex relationships can be understood.

Neural Knowledge Language Model (NKLM)

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

- Conventional Language Models, limited in their ability in dealing with factual knowledge because these are usually represented by named entities such as person names, place names, years, etc. Such as:

Kanye West, a famous <unknown> and the husband of <unknown>, released his latest album <unknown> in <unknown>.

Solution:

Neural Knowledge Language Model (NKLM) combines symbolic knowledge provided by the knowledge graph with language models.

Summary

- Factual knowledge is the accumulation of facts that develop gradually for a period of time.
- Commonsense knowledge is expected to know by every human being (including children) where as factual knowledge may be segregated.
- Ontology provide the frameworks for knowledge graphs.
- Latent representation of knowledge graphs are resulted into knowledge embeddings.
-

References

UNIVERSITÄT DUISBURG ESSEN

Offen im Denken

- **Mandatory**
  - Check Moodle
- **Optional**

- A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko. Translating embeddings for modeling multi-relational data. In NIPS, pages 2787--2795, 2013.
- Federico Bianchi and Gaetano Rossiello and Luca Costabello and Matteo Palmonari and Pasquale Minervini, Knowledge Graph Embeddings and Explainable AI, 2020
- Michael Färber and Achim Rettinger, Which Knowledge Graph Is Best for Me?, 2018
- https://towardsdatascience.com/an-introduction-to-knowledge-graphs-841bbc0e796e
- https://homes.cs.washington.edu/~msap/acl2020-commonsense/

Thank You