

Question Answering using Knowledge Graphs

Jay Pujara

DSCI-558 Spring 2021

4/14/21

Real-world KG applications



Real-world KG applications

Google search results for "how many coronavirus cases in us".

Statistics

New cases

United States

31,869 April 11

Mar 12 Mar 19 Mar 26 Apr 2 Apr 9

New cases are the confirmed cases reported since the previous day
Updated less than 20 mins ago · Source: Wikipedia · About this data

Cases

Worldwide United States All

Location	Confirmed	Recovered	Deaths
New York	188,694	17,089	9,385

Coronavirus (COVID-19) map

View full map

Source: Wikipedia · About this data

Cases overview

United States

Confirmed	Recovered	Deaths
559,409	33,115	22,071

Real-world KG applications

where does jay pujara work? X 🔍

All News Maps Images Videos More Settings Tools

About 346,000 results (0.59 seconds)

Jay Pujara is a research assistant professor at the University of Southern California and a research lead at the Information Sciences Institute whose principal areas of research are machine learning, artificial intelligence, and data science.



[Jay Pujara - USC - Viterbi School of Engineering - Viterbi ...](#)
<https://viterbi.usc.edu> › directory › faculty › Pujara › Jay ▾

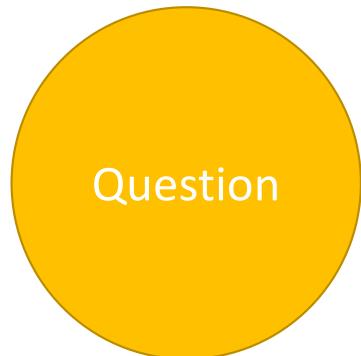
where does usc play | 🔍

where does usc play basketball 🔍

Galen Center 🔍

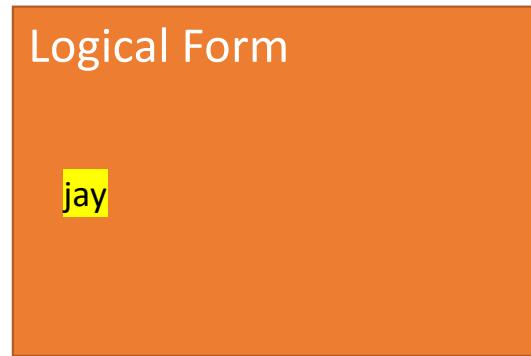
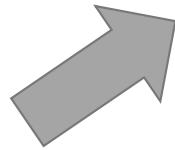
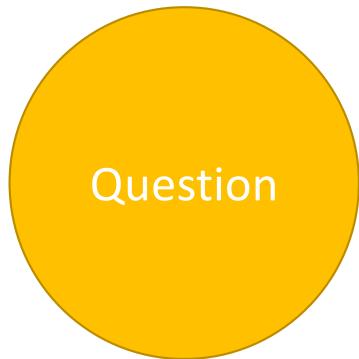
Going from questions to answers?

Where does Jay Pujara work?



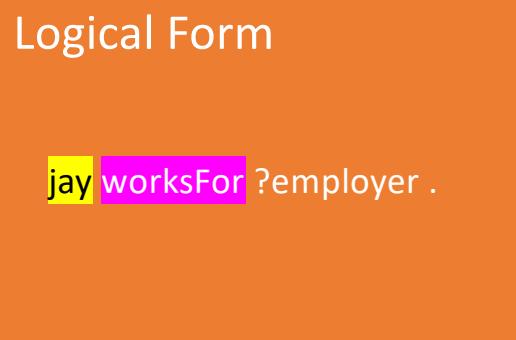
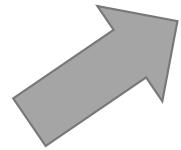
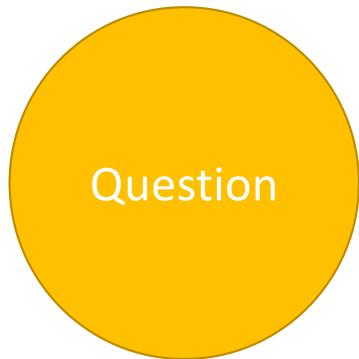
Going from questions to answers?

Where does Jay Pujara work?



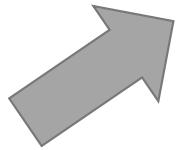
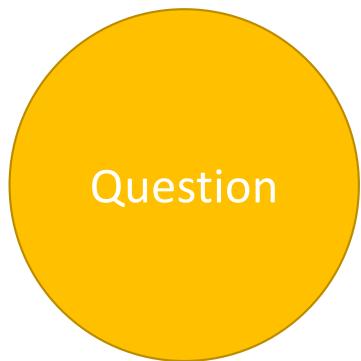
Going from questions to answers?

Where does Jay Pujara work?



Going from questions to answers?

Where does Jay Pujara work?

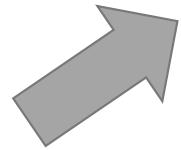
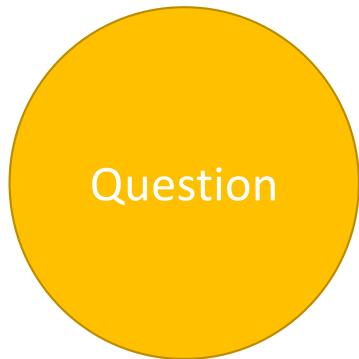


Logical Form

```
jay worksFor ?employer .  
?employer address ?addr .  
?addr addressLocality ?city .
```

Going from questions to answers?

Where does Jay Pujara work?

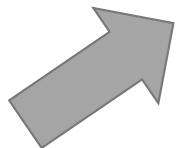
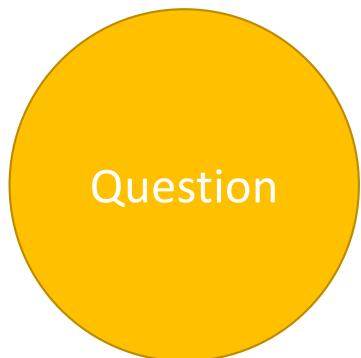


Logical Form

```
SELECT ?city WHERE {  
    jay worksFor ?employer .  
    ?employer address ?addr .  
    ?addr addressLocality ?city  
}
```

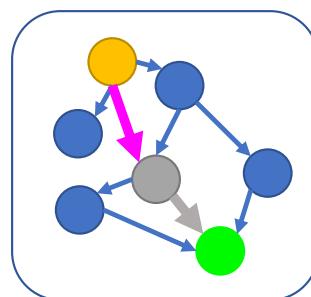
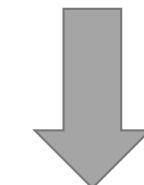
Going from questions to answers?

Where does Jay Pujara work?



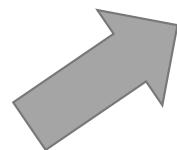
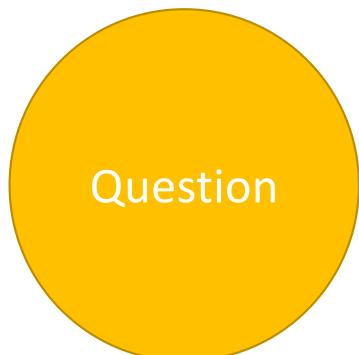
Logical Form

```
SELECT ?city WHERE {  
    jay worksFor ?employer .  
    ?employer address ?addr .  
    ?addr addressLocality ?city  
}
```



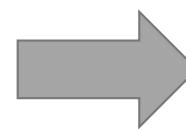
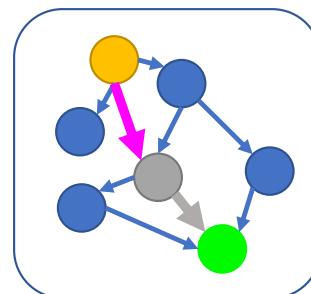
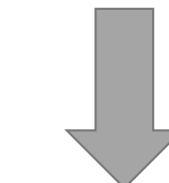
Going from questions to answers?

Where does Jay Pujara work?



Logical Form

```
SELECT ?city WHERE {  
    jay worksFor ?employer .  
    ?employer address ?addr .  
    ?addr addressLocality ?city  
}
```



Los Angeles



Important Subtasks

- Entity Linking
- Relation Linking
- Query Structure Discovery
- Identifying Logical Operators

Challenges [from QALD17]

- Lexical Gap
- Ambiguity
- Multilingualism
- Complexity
- Distributed Knowledge
- Temporal, Spatial, and Procedural Questions
- Diversity / Template-Free Questions

Lexical Gap between ontology and language

- String normalization – different forms of the same word
- Query expansion – adding additional terms
- Pattern libraries – translating complex phrasal structure into properties
- Entailment – using previous answers + reasoning to fill in missing facts
- Open IE style approaches – co-occurrences of entities & relations

Ambiguity of questions and concepts

- Context-based or corpus-based filtering
- Constraints between concepts in the question
- Coherence models across ontologies and concept hierarchies

Complex queries

- Multiple hops or levels of indirection
- Aggregation, logical comparisons, quantifiers
- Indirect references to prior answers, nested queries
- Constraints on the answers

Distributed Knowledge across several KGs

- Identifying the right KG with the answer
- Understanding entity links between KGs
- Predicting missing links

Procedural, Spatial, & Temporal QA

- Answering tasks may require different types of knowledge
- Answer may not be an entity (e.g., “How do you make a sandwich?”)
- Event-based questions require knowledge of “before” “after” “causes”
- Questions may use geo-coordinates, spatial relationships, containment

Template(-free) questions

- Question may not match existing structural patterns
- Question might require multiple structural schemas
- May require aggregation of answers from different query types and knowledge sources

Solutions and candidate solutions

Table 6

Established and actively researched as well as envisioned techniques for solving each challenge.

Challenge	Established	Future
Lexical Gap	stemming, lemmatization, string similarity, synonyms, vector space model, indexing, pattern libraries, explicit semantic analysis	combined efforts, reuse of libraries
Ambiguity	user information (history, time, location), underspecification, machine learning, spreading activation, semantic similarity, crowdsourcing, Markov Logic Network	holistic, knowledge-base aware systems
Multilingualism	translation to core language, language-dependent grammar	usage of multilingual knowledge bases
Complex Operators	reuse of former answers, syntactic tree-based formulation, answer type orientation, HMM, logic	non-factual questions, domain-independence
Distributed Knowledge and Procedural, Temporal, Spatial	temporal logic	domain specific adaptors, procedural SQA
Templates	fixed SPARQL templates, template generation, syntactic tree based generation	complex questions

Case Study: Watson

[See Question Analysis: How Watson Reads a Clue](#)

IBM's Watson QA system



Problem 1: Focus entity

Goal: Determine the entity that constitutes the answer to the question

Baseline: set of patterns

Improve co-reference detection

Use category information

Use previous questions

- A noun phrase with determiner “this” or “these”: THEATRE: A new play based on *this Sir Arthur Conan Doyle canine classic* opened on the London stage in 2007.
- “This” or “these” as a pronoun: ‘88: In April 1988, Northwest became the first U.S. air carrier to ban *this* on all domestic flights.
- When the question is a noun phrase, we conventionally label the entire question as the focus: AMERICAN LIT: *Number of poems Emily Dickinson gave permission to publish during her lifetime.*
- One of the pronouns “he/she/his/her/him/hers”: OUT WEST: *She* joined Buffalo Bill Cody’s Wild West Show after meeting him at the Cotton Expo in New Orleans.
- One of the pronouns “it/they/them/its/their”: ME “FIRST”!: *It* forbids Congress from interfering with a citizen’s freedom of religion, speech, assembly, or petition.
- The pronoun “one”: 12-LETTER WORDS: Leavenworth, established in 1895, is a federal *one*.

Problem 2: Lexical Answer Types

Goal: Find types and constraints to characterize the focus entity

Approach: Use modifiers of the focus entity in the parse structure. Cover other cases with a set of rules.

- If the focus is a conjunction, extract the conjuncts:
HENRY VIII: Henry destroyed the Canterbury Cathedral Tomb of *this saint and chancellor of Henry II*.
- “⟨Focus⟩ of X”. extract LAT X when ⟨Focus⟩ is any of one/name/type/kind:
HERE, PIGGY, PIGGY, PIGGY: Many a mom has compared her kid’s messy room to *this kind of hog enclosure*.
- “⟨Focus⟩ for X”. extract LAT X when ⟨Focus⟩ is any of name/word/term:
COMPANY NAME ORIGINS: James Church chose *this name for his product b*ecause the symbols of the god Vulcan represented power.
- If no focus was detected and the category is a noun phrase, take headword of the category as LAT:
HEAVY METAL **BANDS**: “Seek & Destroy”, “Nothing Else Matters”, “Enter Sandman”.

Problem 3: Question Type Prediction

Goal: Understand the type of question

Approach: Classify using the focus entity and a set of pre-defined templates

<i>QClass</i>	<i>Description</i>	<i>Example questions (correct answer)</i>	<i>Frequency (%)</i>
DEFINITION	A question that contains a definition of the answer	CONSTRUCTION: It can be the slope of a roof, or the gunk used to waterproof it. (Answer: "pitch") CONSTRUCTION: The name of this large beam that supports the joists literally means "something that encircles". (Answer: "a girder")	14.2
CATEGORY-RELATION	The answer has a semantic relation to the question, where the relation is specified in the category	FORMER STATE GOVERNORS: Nelson A. Rockefeller. (Answer: "New York") COUNTRIES BY NEWSPAPER: Haaretz, Yedioth Ahronoth. (Answer: "Israel")	7.2
FITB	A fill-in-the-blank question asks for completion of a phrase	COMPLETE IT: Attributed to Lincoln: "The ___ is stronger than the bullet." (Answer: "ballot") SHAKESPEARE IN LOVE: "Not that I loved Caesar less," says Brutus, "but that I loved" this city "more." (Answer: "Rome")	3.8
ABBREVIATION	The answer is an expansion of an abbreviation in the question	MILITARY MATTERS: Abbreviated SAS, this elite British military unit is similar to the USA's Delta Force. (Answer: "the Special Air Service")	2.9
PUZZLE	A puzzle question: the answer requires derivation, synthesis, inference, etc.	BEFORE & AFTER: 13th Century Venetian traveler who's a Ralph Lauren short sleeve top with a collar. (Answer: "Marco Polo shirt") THE HIGHEST-SCORING SCRABBLE WORD: Zoom, quiz or heaven. (Answer: "quiz")	2.3
ETYMOLOGY	A question asking for an English word derived from a foreign word having a given meaning	ARE YOU A FOOD"E"?: From the Spanish for "to bake in pastry", it's South America's equivalent of a calzone. (Answer: "an empanada")	1.9
VERB	Question asks for a verb	THE NOT-SO-DEADLY SINS: To capitalize all text in an email is an abomination that signifies the person is doing this. (Answer: "shouting")	1.5
TRANSLATION	A question asking for translation of a word or phrase from one language to another	FRUITS IN FRENCH: Pomme. (Answer: "apple")	1.1

Problem 4: Constraints

Goal: Understand the constraints required for the correct answer

Approach: Classify using a set of pre-defined patterns and category clues

- LexicalConstraint—A phrase such as “this 4-letter word” that should not be used in a query but is critical for selecting the correct answer.
- Abbreviation—A term in a question that is identified as an abbreviation, which is associated with its possible expansions. For questions of QClass ABBREVIATION, one of these Abbreviation QSections is further identified as the abbreviation whose expansion is sought.
- SubQuestionSpan—When a question can be decomposed into two or more disjoint sections that individually indicate or contribute to the answer, these sections are marked as SubQuestionSpans. Some of the more complex Factoids, as well as all “BEFORE & AFTER”s, most “RHYME TIME”s, and double definitions such as the first DEFINITION example in Table 4, get SubQuestionSpan annotations [3, 5].
- McAnswer—The (usually three) strings that represent the answer choices in a multiple-choice question are marked with a McAnswer QSection.
- FITB—This annotates the string that adjoins the focus term (i.e., the text that forms the term or expression that the focus completes).

Matching Logical Relations

Identifying relations in the question

<i>Question</i>	<i>Relations</i>
A.k.a., the Flavian Amphitheatre, this ancient structure was begun by the Roman Emperor Vespasian around 72 A.D.	altName(focus, Flavian Amphitheatre)
A myocardial infarction, better known as this, is a common reason for ICU admission.	altName(focus, myocardial infarction)
In May 1898 Portugal celebrated the 400th anniversary of this explorer's arrival in India.	anniversaryOf(this explorer's arrival in India, 400, May 1898)
Chile shares its longest land border with this country.	borderOf(focus, Chile)
In 1867 the U.S. bought this island group named for a Russian captain and leased it to seal hunting companies.	rdfTriple(buy, U.S., focus)

Parsing a question into logical form

POETS & POETRY: He was a bank clerk in the Yukon before he published “Songs of a Sourdough” in 1907.

```
lemma(1, "he").  
partOfSpeech(1,pronoun).  
lemma(2, "publish").  
partOfSpeech(2,verb).  
lemma(3, "Songs of a Sourdough").  
partOfSpeech(3,noun).  
subject(2,1).  
object(2,3).
```

```
authorOf(Author, Composition) :-  
    createVerb(Verb),  
    subject(Verb, Author),  
    author(Author),  
    object(Verb, Composition),  
    composition(Composition).-  
  
createVerb(Verb) :-  
    partOfSpeech(Verb, verb),  
    lemma(Verb, VerbLemma),  
    [ "write", "publish", ... ].
```

Semantic Parsing

Goal: Translate NL into logical form

Dependency parsing is a specific logical form,
techniques can be generalized to others

Logical forms include Combinatory Categorical Grammars, Abstract Meaning Representations, Lambda Calculus, ...

Approaches include: Hidden Markov Models, Conditional Random Fields, RNNs, LSTMs, ...

Research thrusts: less supervision, scaling to large vocabularies

Category	Description
Alignment	Log of # entity pairs that occur with the phrase r_1 ($ \mathcal{F}(r_1) $) Log of # entity pairs that occur with the logical predicate r_2 ($ \mathcal{F}(r_2) $) Log of # entity pairs that occur with both r_1 and r_2 ($ \mathcal{F}(r_1) \cap \mathcal{F}(r_2) $) Whether r_2 is the best match for r_1 ($r_2 = \arg \max_r \mathcal{F}(r_1) \cap \mathcal{F}(r) $)
Lexicalized	Conjunction of phrase w and predicate z
Text similarity	Phrase r_1 is equal/prefix/suffix of s_2 Phrase overlap of r_1 and s_2
Bridging	Log of # entity pairs that occur with bridging predicate b ($ \mathcal{F}(b) $) Kind of bridging (# unaries involved) The binary b injected
Composition	# of intersect/join/bridging operations POS tags in join/bridging and skipped words Size of denotation of logical form

Table 1: Full set of features. For the alignment and text similarity, r_1 is a phrase, r_2 is a predicate with Freebase name s_2 , and b is a binary predicate with type signature (t_1, t_2) .

Questions to Templates



Approach: Match question to SPARQL patterns

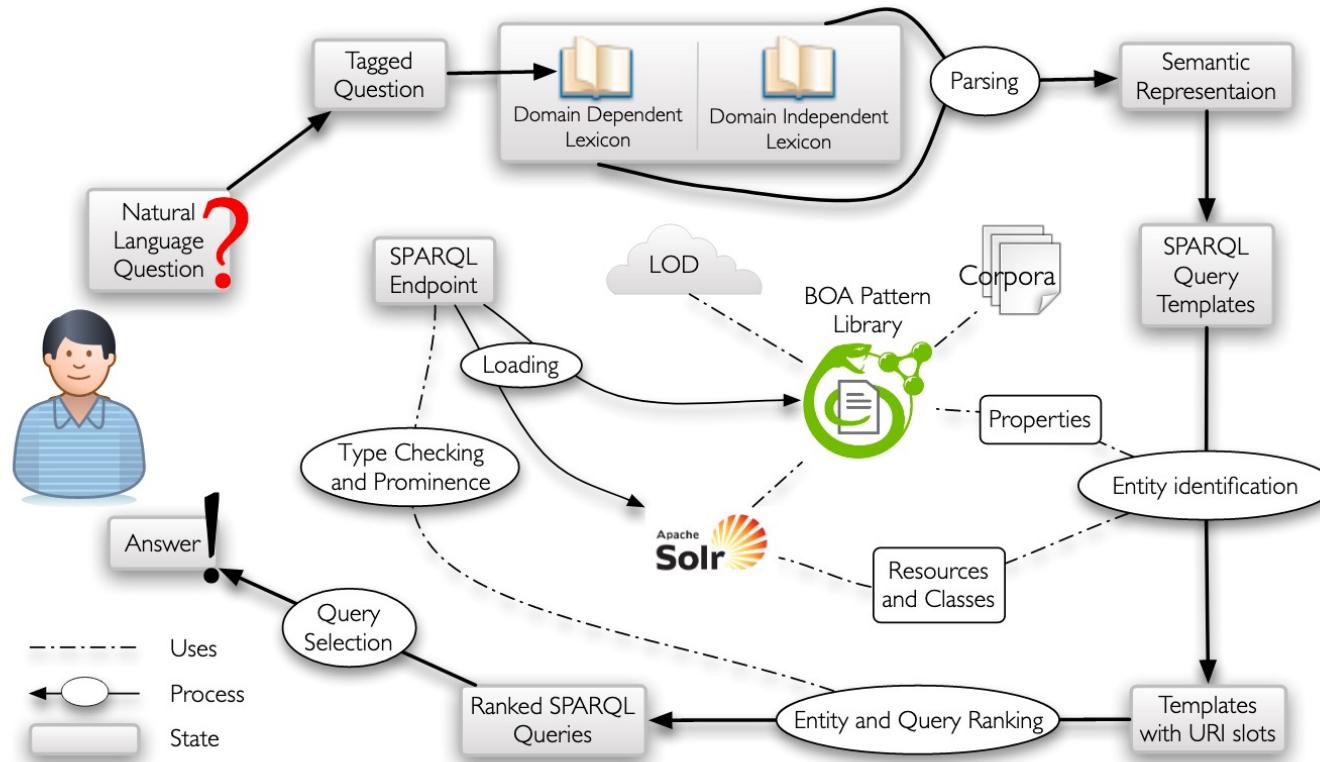
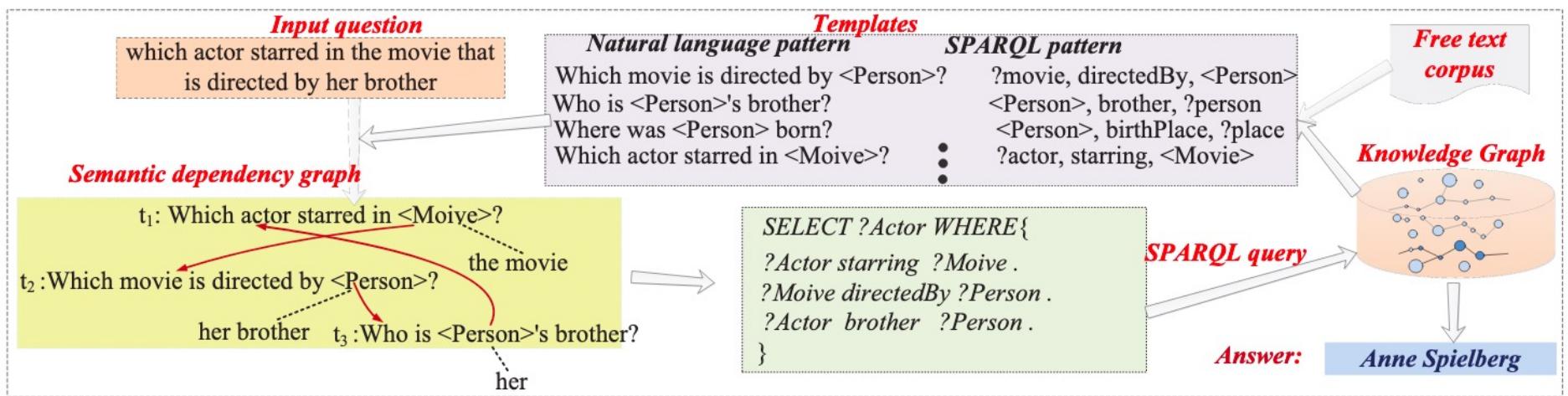


Figure 1: Overview of the template based SPARQL query generator.

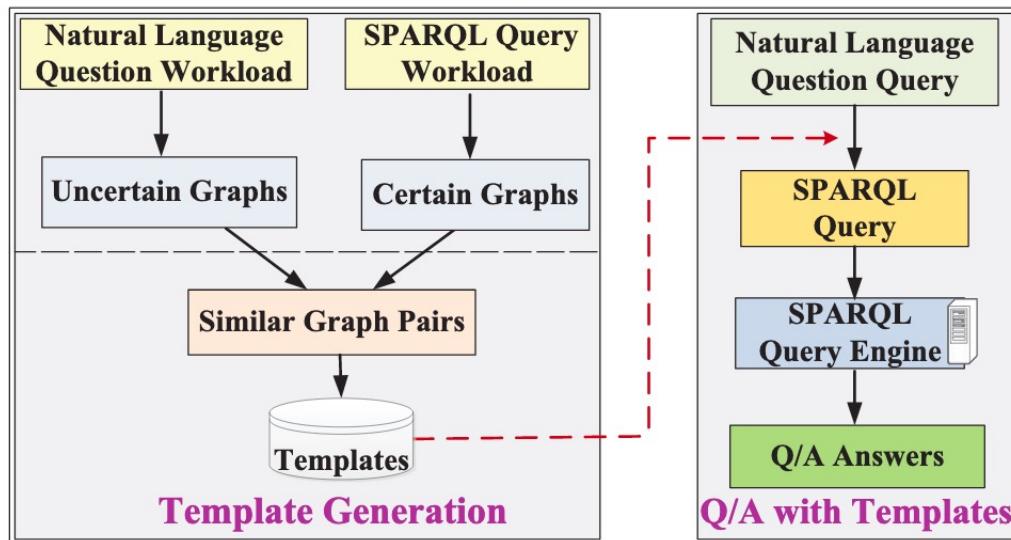
From https://pub.uni-bielefeld.de/download/2495397/2526223/template-based_question.pdf

Approach: Match question to SPARQL patterns



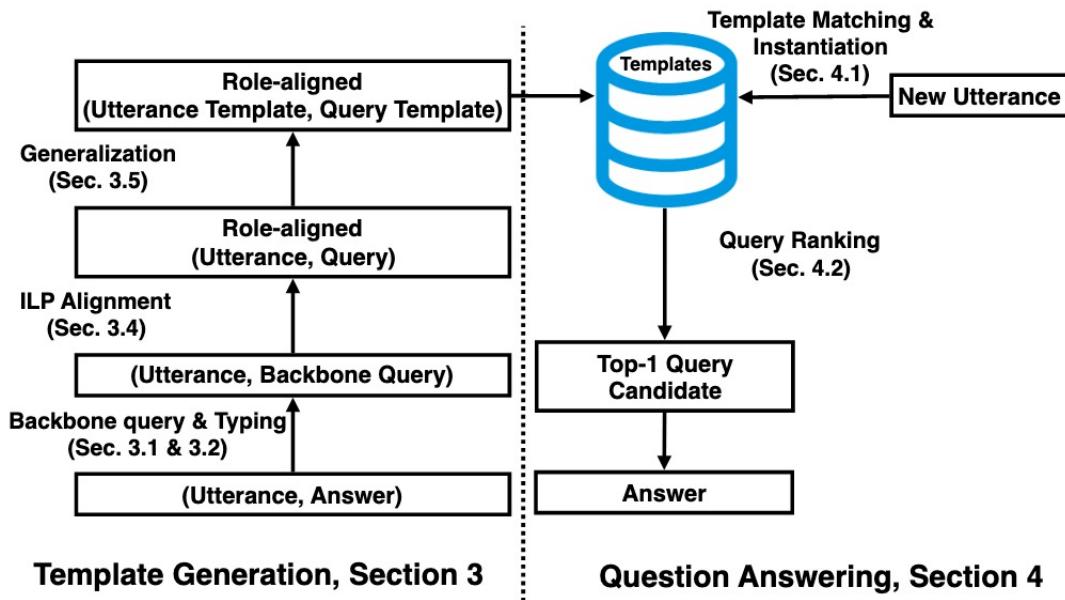
From <http://www.vldb.org/pvldb/vol11/p1373-zheng.pdf>

Idea 1: Turn NL into messy SPARQL



[How to Build Templates for RDF Question/Answering—An Uncertain Graph Similarity Join Approach](#)
Zheng, Zou, Lian, Yu, Song, Zhao. SIGMOD 2015

Idea 2: Use Q-A pairs to discover SPARQL



[Automated Template Generation for Question Answering over Knowledge Graphs](#)
Abujabal, Yahya, Riedwald, Weikum. WWW 2017

Neural Approaches

Semantic Parsing: Three neural approaches

- Classification
 - Predict the slots in a query (in simple cases, predict (s, p) and then query o
- Ranking
 - Given a query, rank possible logical forms
- Translation
 - Given a query, translate to a logical form (based on formal query, results, etc.)

Language Model Approaches

Paragraph A, Return to Olympus:

[1] *Return to Olympus* is the only album by the alternative rock band Malfunkshun. [2] It was released after the band had broken up and after lead singer Andrew Wood (later of Mother Love Bone) had died of a drug overdose in 1990. [3] Stone Gossard, of Pearl Jam, had compiled the songs and released the album on his label, Loosegroove Records.

Paragraph B, Mother Love Bone:

[4] *Mother Love Bone* was an American rock band that formed in Seattle, Washington in 1987. [5] The band was active from 1987 to 1990. [6] Frontman Andrew Wood's personality and compositions helped to catapult the group to the top of the burgeoning late 1980s/early 1990s Seattle music scene. [7] Wood died only days before the scheduled release of the band's debut album, "Apple", thus ending the group's hopes of success. [8] The album was finally released a few months later.

Q: What was the former band of the member of Mother Love Bone who died just before the release of "Apple"?

A: Malfunkshun

Supporting facts: 1, 2, 4, 6, 7

Figure 1: An example of the multi-hop questions in HOTPOTQA. We also highlight the supporting facts in *blue italics*, which are also part of the dataset.

Good review of recent methods in Xiang Ren's guest lecture. Many methods do not use a KG and are primarily focused on identifying a span in the source document

Article: Endangered Species Act

Paragraph: "... Other legislation followed, including the Migratory Bird Conservation Act of 1929, a *1937 treaty* prohibiting the hunting of right and gray whales, and the *Bald Eagle Protection Act of 1940*. These *later laws* had a low cost to society—the species were relatively rare—and little *opposition* was raised."

Question 1: "Which laws faced significant *opposition*?"

Plausible Answer: *later laws*

Question 2: "What was the name of the *1937 treaty*?"

Plausible Answer: *Bald Eagle Protection Act*

Figure 1: Two unanswerable questions written by crowdworkers, along with plausible (but incorrect) answers. Relevant keywords are shown in *blue*.

Knowledge-Aware Graph Network (KagNet)

Commonsense Question Answering

Where do adults usually use glue sticks?

A: classroom B: office C: desk drawer

What do you need to fill with ink to write notes on an A4 paper?

A: fountain pen B: printer C: pencil

Can you choose **the most plausible answer** based
on daily life **commonsense** knowledge?

Commonsense Question Answering

Where do adults usually use glue sticks?

A: classroom B: office C: desk drawer

What do you need to fill with ink to write notes on an A4 paper?

A: fountain pen B: printer C: pencil

From the CommonsenseQA dataset (Talmor et al. NAACL 2019)

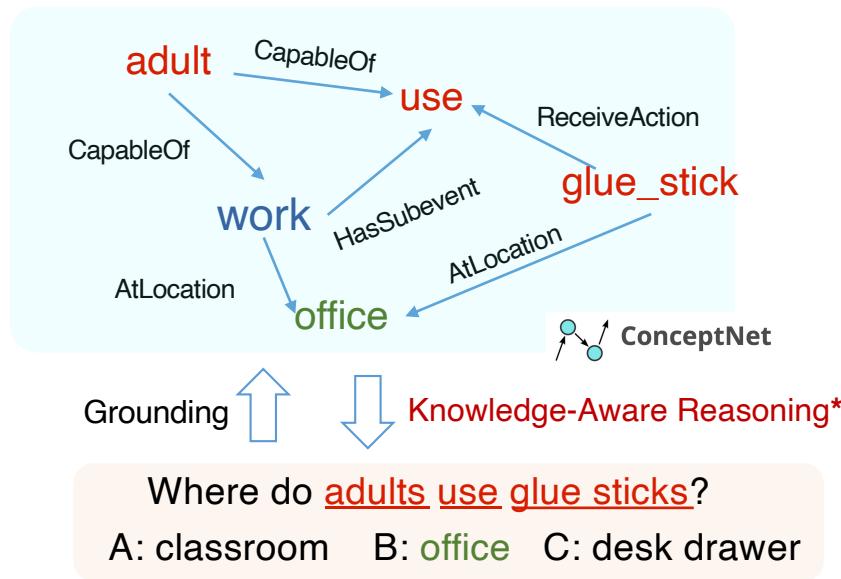
Research question:

How can we impose commonsense in NLU models?

Knowledge-Aware Reasoning

Symbol Space

Semantic Space



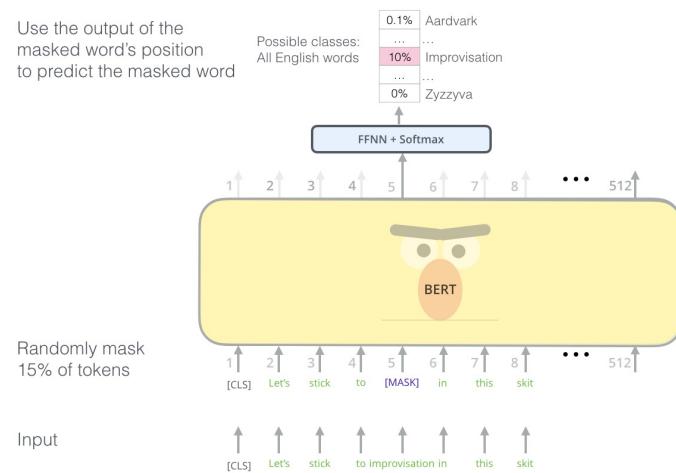
A Schema Graph
for the choice B

Question
Answer Candidates

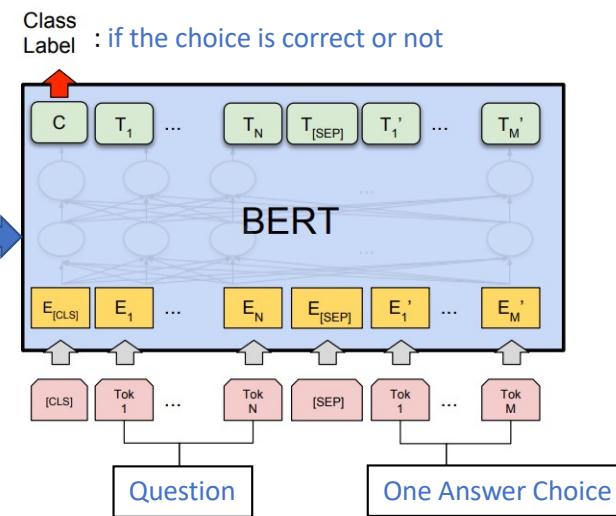
Challenges in knowledge-aware reasoning

- How can we find the schema graphs?
 - Noisy and Incomplete
 - Numerous graphs; how to select the most related ones
- How do we encode these graphs for reasoning?
 - Complex multi-relational graph structures
 - NO supervision in aligning graphs and question-answer pairs
 - Need to be compatible with neural sentence encoders

Fine-tuning BERT for Answering Commonsense Questions



Pre-Training BERT-like models w/ **Masked LM** objective
credit: <http://jalammar.github.io/illustrated-bert/>



Fine-tuning BERT for the CommonsenseQA task.

Limits of fine-tuning BERT-like models for Commonsense Reasoning

1. Not capturing commonsense : Most plausible predictions are far from common truth

Masked Language Modeling

Enter text with one or more "[MASK]" tokens and BERT will generate the most likely token to substitute for each "[MASK]".

Sentence:

Adults usually use glue sticks at their [MASK].

Mask 1 Predictions:
16.4% **feet**
14.8% **disposal**
5.4% **backs**
3.5% **fingertips**

Sentence:

A bird has one [MASK].

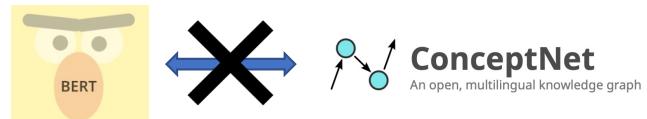
Mask 1 Predictions:
10.3% **wing**
8.3% **eye**
3.6% **leg**

Online demo of BERT's Masked-LM
<https://demo.allennlp.org/masked-lm>

Prompt Correct Answer Predictions

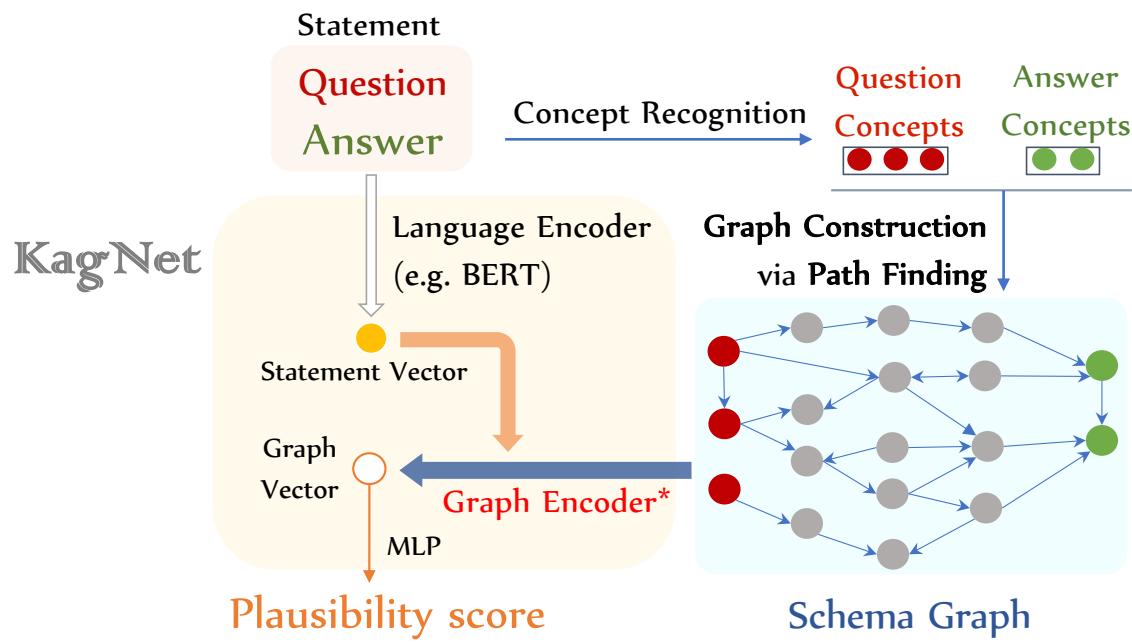
Prompt	Correct Answer	Predictions
Birds can [MASK].	fly	fly (-0.5), sing (-2.3), talk (-2.8)
Birds cannot [MASK].		fly (-0.3), sing (-3.6), speak (-4.1)

2. Not Interpretable w/ Knowledge



“Negated LAMA: Birds cannot fly”
<https://arxiv.org/pdf/1911.03343.pdf>

Proposed Framework Overview



(Bill Yuchen Lin et al. 2019)

Slides courtesy Xiang Ren

(1) Schema Graph Construction

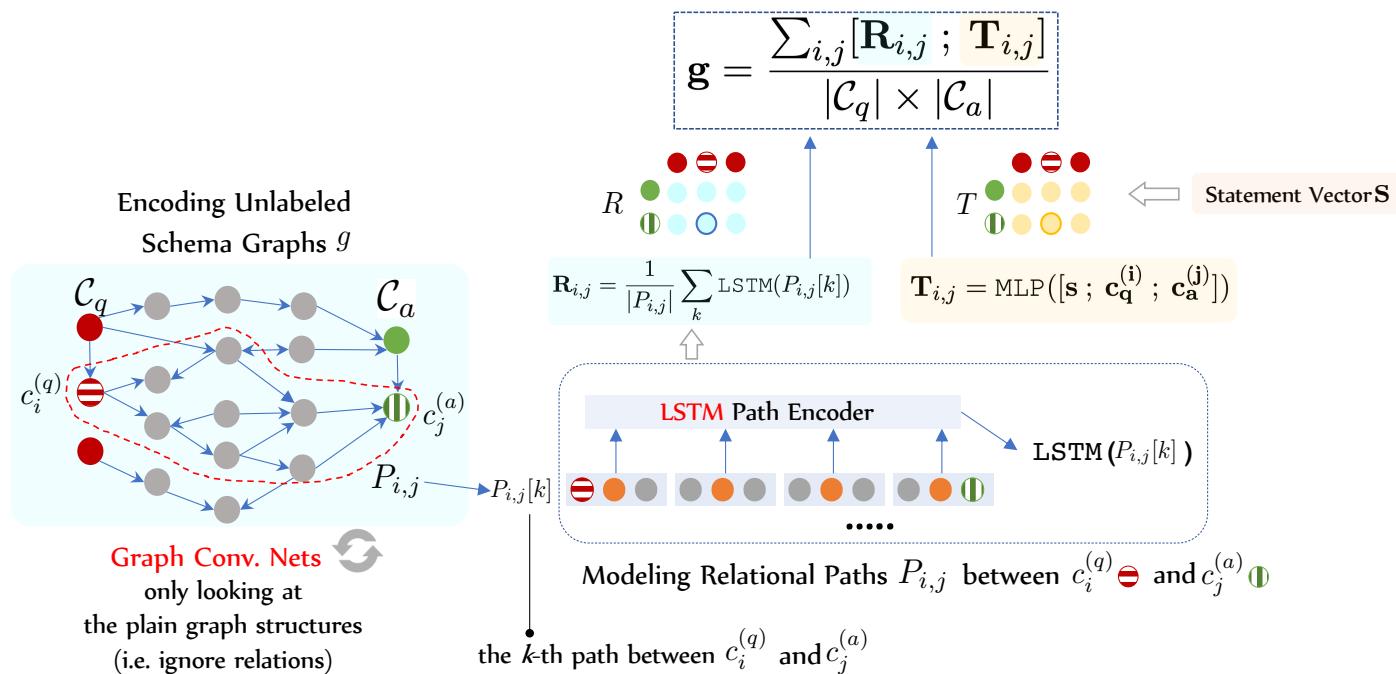
- Concept Recognition

- Tokenization / Lemmatization
- Match ConceptNet vocabulary
- Merge multiple smaller concepts into a longer one
 - e.g. " fountain", "pen" --> "fountain pen"
- Question Concepts \mathcal{C}_q and Answer Concepts \mathcal{C}_a

- Path Finding

- Find paths between each QA-concept pair (one from \mathcal{C}_q and one from \mathcal{C}_a)
 - $P_{i,j}$ denotes the set of paths between i -th question concept and j -th answer concept
 $c_i^{(q)} \ominus$, $c_j^{(a)} \oplus$
- Path pruning by length (≤ 5 nodes) and embedding-based metric.
 - Train knowledge graph embeddings (e.g. TransE) for scoring paths
 - Choose the highest-score relation by the KGE when there are multiple ones

(2) Path-based Relational Graph Encoder



(3) w/ Hierarchical Path-based Attention

- Two average pooling:
 - Assuming all QA-concept pairs are equally important
 - Assuming all paths are equally relevant
- Modeling the two-level importance as latent weights:

$$\mathbf{g} = \frac{\sum_{i,j} [\mathbf{R}_{i,j} ; \mathbf{T}_{i,j}]}{|\mathcal{C}_q| \times |\mathcal{C}_a|}$$

$$\mathbf{R}_{i,j} = \frac{1}{|P_{i,j}|} \sum_k \text{LSTM}(P_{i,j}[k])$$

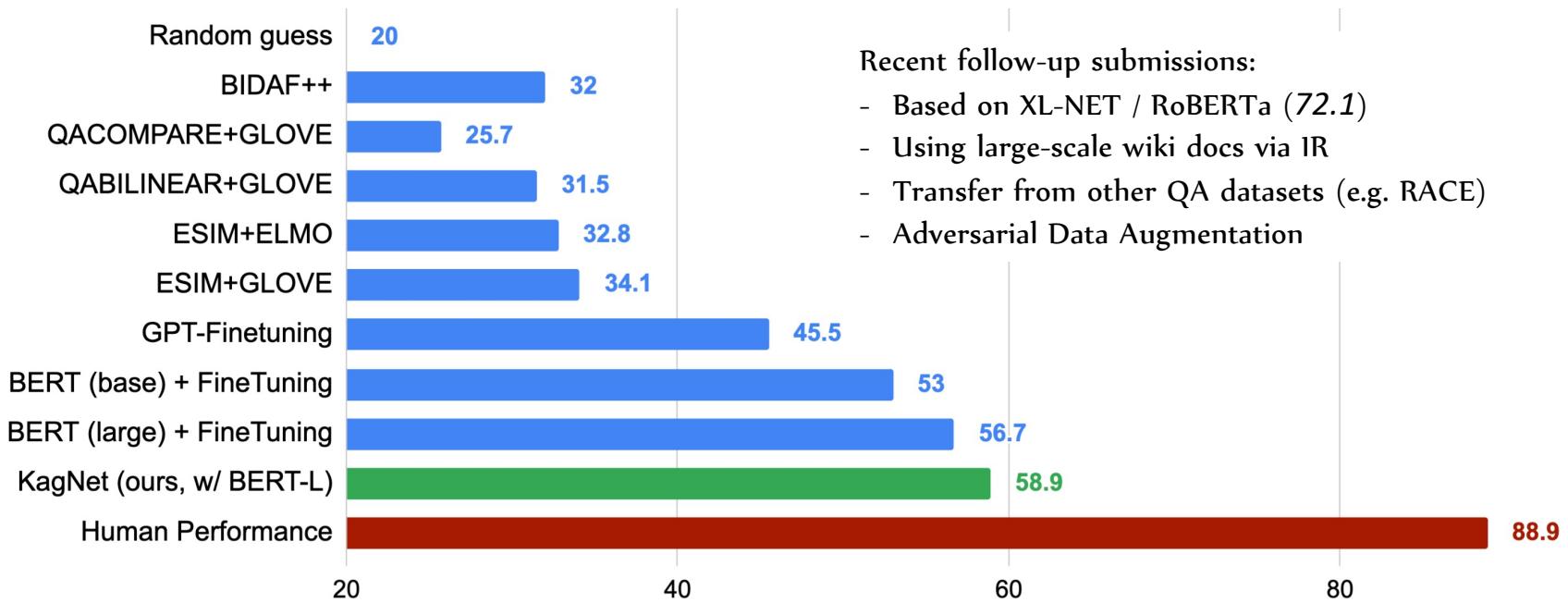
$$\begin{aligned}\alpha_{(i,j,k)} &= \mathbf{T}_{i,j} \mathbf{W}_1 \text{LSTM}(P_{i,j}[k]), \\ \hat{\alpha}_{(i,j,\cdot)} &= \text{SoftMax}(\alpha_{(i,j,\cdot)}), \\ \hat{\mathbf{R}}_{i,j} &= \sum_k \hat{\alpha}_{(i,j,k)} \cdot \text{LSTM}(P_{i,j}[k])\end{aligned}$$

Path-Level Attention
(attending on semantic space)

$$\begin{aligned}\beta_{(i,j)} &= \mathbf{s} \mathbf{W}_2 \mathbf{T}_{i,j} \\ \hat{\beta}_{(\cdot,\cdot)} &= \text{SoftMax}(\beta_{(\cdot,\cdot)}) \\ \hat{\mathbf{g}} &= \sum_{i,j} \hat{\beta}_{(i,j)} [\hat{\mathbf{R}}_{i,j} ; \mathbf{T}_{i,j}]\end{aligned}$$

ConceptPair-Level Attention
(attending on statement)

Experiments



Recent follow-up submissions:

- Based on XL-NET / RoBERTa (72.1)
- Using large-scale wiki docs via IR
- Transfer from other QA datasets (e.g. RACE)
- Adversarial Data Augmentation

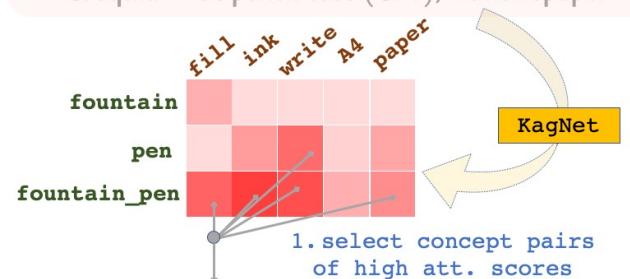
More Performance on Official Test Set: <https://www.tau-nlp.org/csqa-leaderboard>

Interpretability & Transferability

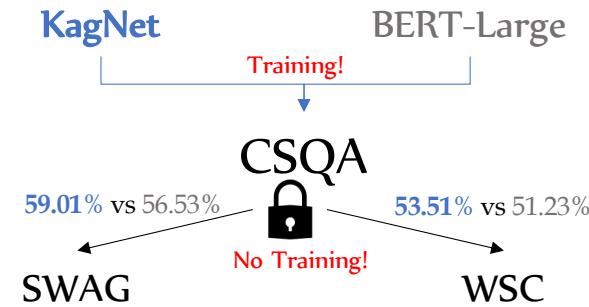
What do you **fill** with **ink** to **write** on an **A4 paper**?

A: fountain pen ✓ (KagNet); B: printer (BERT);

C: squid D: pencil case (GPT); E: newspaper



ink →PartOf→ fountain_pen
ink →RelatedTo→ container <-IsA- fountain_pen
fill <-HasSubEvent- ink <-AtLocation- fountain_pen
fill →RelatedTo→ container <-IsA- fountain_pen
write <-UsedFor- pen
write <-UsedFor- pen <-IsA- fountain_pen
paper <-RelatedTo- write <-UsedFor- fountain_pen
..... 2. Ranking via path-level attn.



Additional Resources

- [Introduction to Neural Network based Approaches for Question Answering over Knowledge Graphs](#)
- [Survey on Challenges of Question Answering in the Semantic Web](#)
- [Semantic Parsing on Freebase from Question-Answer Pairs](#)
- [Question Answering over Linked Data](#)
- [LC-QuAD 2.0: A Large Dataset for Complex Question Answering over Wikidata and DBpedia](#)
- [HotPotQA / SQuAD](#)