LECTURE 25: QUESTION ANSWERING

Mehmet Can Yavuz, PhD.

Adapted from Julia Hockenmaier, NLP S2023 - course material
https://courses.grainger.illinois.edu/cs447/sp2023/

KNOWLEDGE-BASED QA

KNOWLEDGE-BASED QA

- **Paradigm 1: Graph-based QA:**
- Assumes you have a knowledge based of “facts” (facts = RDF triplets: predicate with two arguments, can also be expressed as a knowledge graph):
- Ada Lovelace birth-year 1815
- Claude Shannon birth-year 1916
- William Shakespeare author Hamlet
- … … ….
- [data sets: SimpleQuestions, FreebaseQA, WebQA etc.]
- When was Ada Lovelace born?
- Who was born in 1815?
- …

KNOWLEDGE-BASED QA

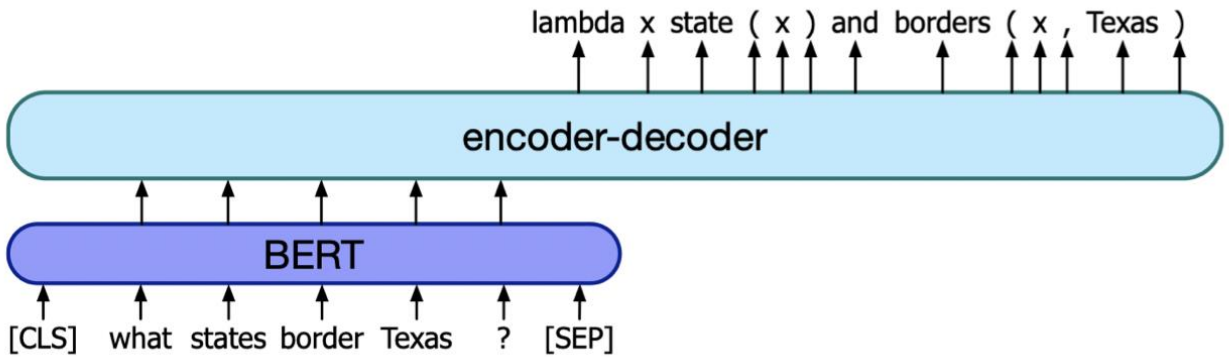
- **Paradigm 2: QA by semantic parsing**
  - Assumes facts are given in a database.
  - QA can be done by translating query to DB query

Question	Logical form
What states border Texas?	$\lambda x.\text{state}(x) \wedge \text{borders}(x, \text{texas})$
What is the largest state?	$\text{argmax}(\lambda x.\text{state}(x), \lambda x.\text{size}(x))$
I'd like to book a flight from San Diego to Toronto	SELECT DISTINCT f1.flight_id FROM flight f1, airport_service a1, city c1, airport_service a2, city c2 WHERE f1.from_airport=a1.airport_code AND a1.city_code=c1.city_code AND c1.city_name= 'san diego' AND f1.to_airport=a2.airport_code AND a2.city_code=c2.city_code AND c2.city_name= 'toronto'
How many people survived the sinking of the Titanic?	(count (!fb:event.disaster.survivors fb:en.sinking_of_the_titanic))
How many yards longer was Johnson's longest touchdown compared to his shortest touchdown of the first quarter?	ARITHMETIC diff(SELECT num(ARGMAX(SELECT)) SELECT num(ARGMIN(FILTER(SELECT))))

SEMANTIC PARSING FOR QA

Earlier approaches: often grammar-based
(e.g. based on CCG)

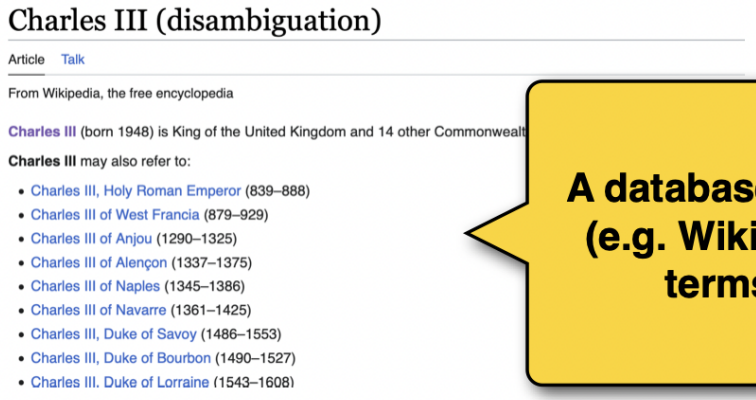
Current approaches: seq2seq based models:



ENTITY LINKING

Map mentions of entities in text to the corresponding entry in an ontology (e.g. based on Wikipedia pages):

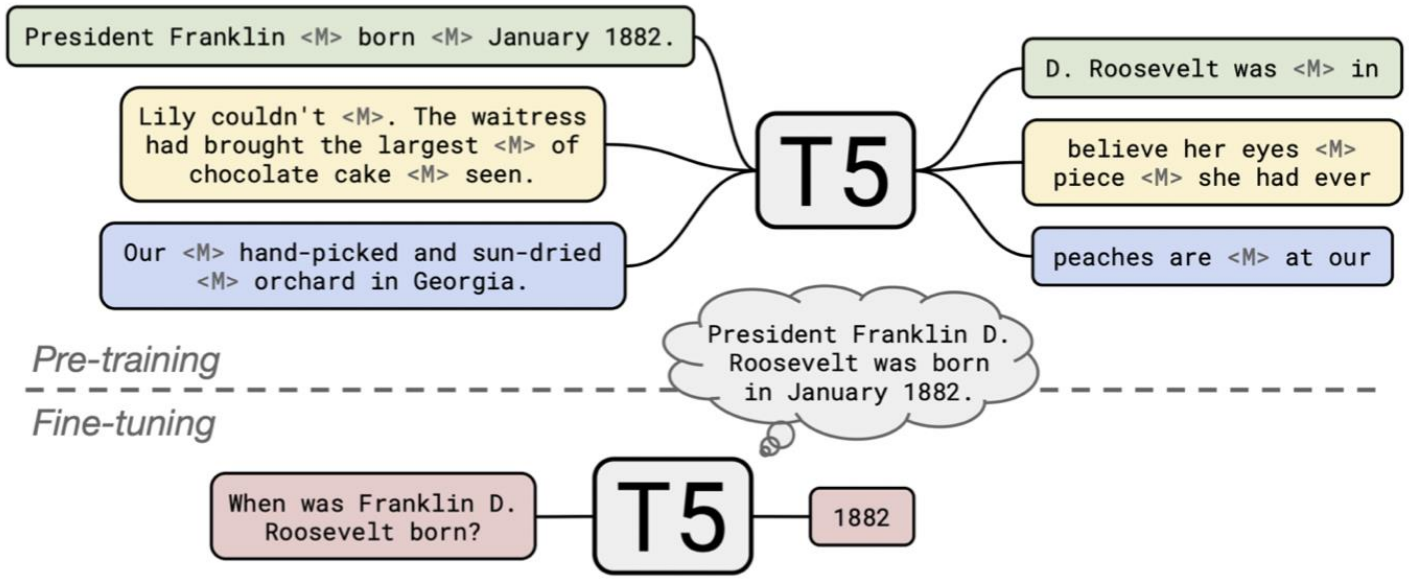
- **Mention detection**: which spans are entity mentions?
- **Mention disambiguation**: which entry does a mention refer to?



A database obtained from Wikipedia (e.g. Wikidata) would use different terms for each Charles III.

This is useful for **knowledge-based QA** (and other applications)

USING LLMS (E.G. T5) FOR QA



Pre-training: masked language modeling
(fill in masked out word)
Fine-tuning: predict answer (for QA datasets)

OTHER QA TASKS

MORE RECENT DEVELOPMENTS IN QA

QA is a very active area of research

Retrieval-based QA is often seen as too simplistic (especially when billed as “reading comprehension”)

More recent developments include datasets whose answers require several steps of reasoning (multi-hop QA), as well answers that require commonsense knowledge.

Visual QA: answer questions about an image.

SCIENCE EXAMS AS TESTBED FOR QA

Task: Answer **multiple choice questions** from 8th-grade science exams

- 1. Which equipment will best separate a mixture of iron filings and black pepper?
- (1) magnet (2) filter paper (3) triple-beam balance (4) voltmeter

This requires a lot of **background knowledge** that has to be acquired from somewhere (e.g. textbooks), and reasoning capabilities

https://allenai.org/content/docs/Aristo_Milestone.pdf