



CMPT 413/825: Natural Language Processing

# Question Answering

Fall 2020  
2020-11-20

Adapted from slides from Danqi Chen and Karthik Narasimhan  
(with some content from slides from Chris Manning)

# Question Answering

- Goal: build computer systems to answer questions

Question

When were the first pyramids built?

Answer

2630 BC

What's the weather like in Vancouver?

42 F

Where is Einstein's house?

112 Mercer St, Princeton, NJ 08540

Why do we yawn?

When we're bored or tired we don't breathe as deeply as we normally do. This causes a drop in our blood-oxygen levels and yawning helps us counter-balance that.

# Question Answering

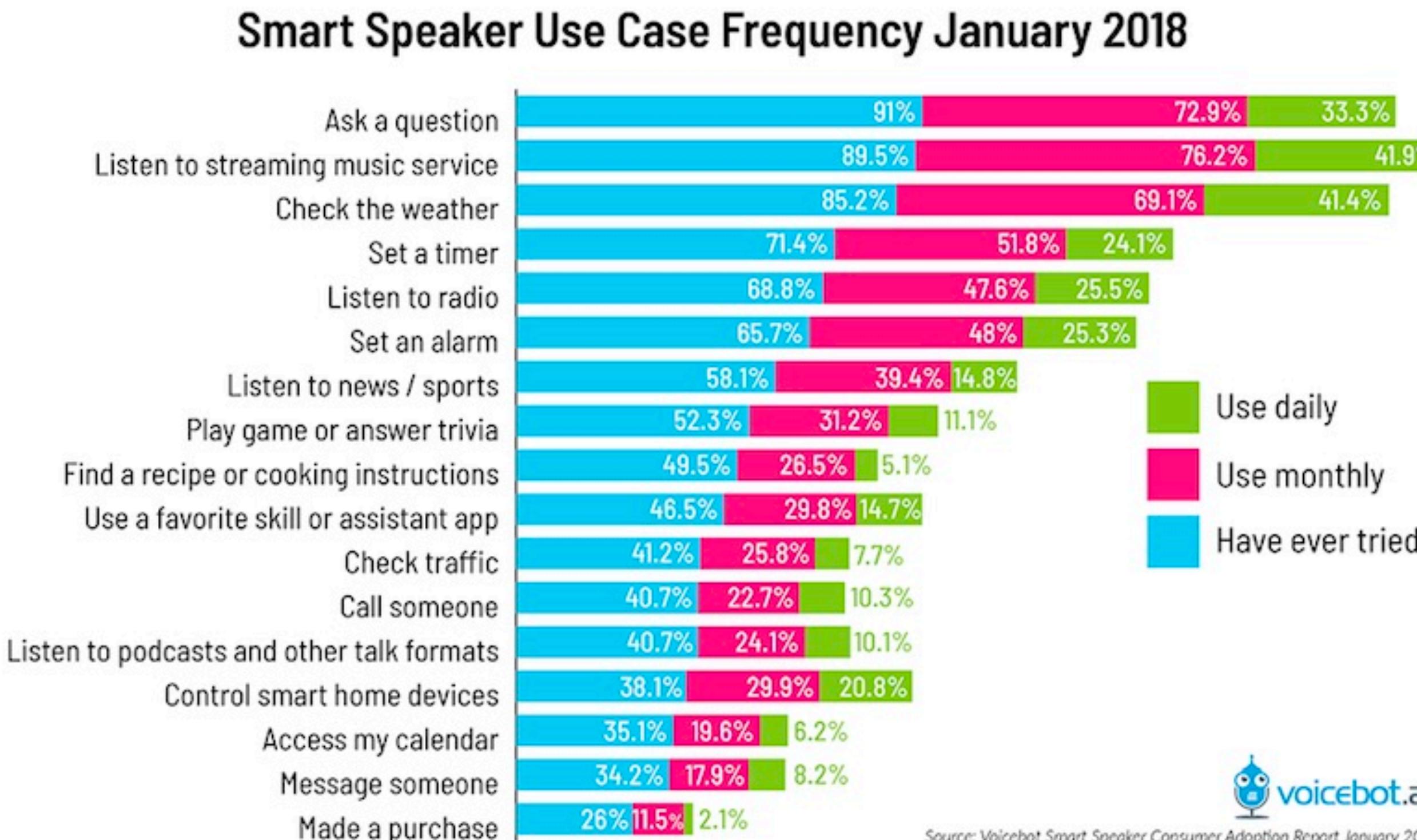
- You can easily find these answers in google today!

A screenshot of a Google search results page. The search query "when were the first pyramids built" is entered in the search bar. Below the search bar, the "All" tab is selected, followed by "Images", "News", "Shopping", "Videos", "More", "Settings", and "Tools". A status message indicates "About 20,300,000 results (0.67 seconds)". The top result is a summary box containing the text "2630 BC" in large bold letters. Below this, a paragraph describes the construction of pyramids for pharaohs during the Old and Middle Kingdom periods, mentioning the Pyramid of Djoser. At the bottom of the summary box is a link "The First Pyramids Built - Timeline Index" and the URL "www.timelineindex.com > content > view".

A screenshot of a Google search results page. The search query "where is einstein's house?" is entered in the search bar. Below the search bar, the "All" tab is selected, followed by "Images", "Maps", "News", "Videos", "More", "Settings", and "Tools". A status message indicates "About 7,550,000 results (1.16 seconds)". The top result is a summary box containing the address "112 Mercer St, Princeton, NJ 08540" in large bold letters.

# Question Answering

- People ask lots of questions to Digital Personal Assistants:



# Question Answering



IBM Watson defeated two of Jeopardy's greatest champions in 2011

# Why care about question answering?

- Lots of immediate applications: search engines, dialogue systems
- Question answering is an important testbed for evaluating how well computer systems understand human language

THE PROCESS OF QUESTION ANSWERING

May 1977

Research Report #88

Wendy Lehnert

When a person understands a story, he can demonstrate his understanding by answering questions about the story. Since questions can be devised to query any aspect of text comprehension, the ability to answer questions is the strongest possible demonstration of understanding. Question answering is therefore a task criterion for evaluating reading skills.

If a computer is said to understand a story, we must demand of the computer the same demonstrations of understanding that we require of people. Until such demands are met, we have no way of evaluating text understanding programs. Any computer programmer can write a program which inputs text. If the programmer assures us that his program 'understands' text, it is a bit like being reassured by a used car salesman about a suspiciously low speedometer reading. Only when we can ask a program to answer questions about what it reads will we be able to begin to assess that program's comprehension.

“Since questions can be devised to query **any aspect** of text comprehension, the ability to answer questions is the **strongest possible demonstration of understanding**.”

# QA Taxonomy

- Factoid questions vs non-factoid questions
- Answers
  - A short span of text
  - A paragraph
  - Yes/No
  - A database entry
  - A list
- Context
  - A passage, a document, a large collection of documents
  - Knowledge base
  - Semi-structured tables
  - Images

# Textual Question Answering

Also called “Reading Comprehension”

The first recorded travels by Europeans to China and back date from this time. The most famous traveler of the period was the Venetian Marco Polo, whose account of his trip to "Cambaluc," the capital of the Great Khan, and of life there astounded the people of Europe. The account of his travels, *Il milione* (or, *The Million*, known in English as the *Travels of Marco Polo*), appeared about the year 1299. Some argue over the accuracy of Marco Polo's accounts due to the lack of mentioning the Great Wall of China, tea houses, which would have been a prominent sight since Europeans had yet to adopt a tea culture, as well the practice of foot binding by the women in capital of the Great Khan. Some suggest that Marco Polo acquired much of his knowledge through contact with Persian traders since many of the places he named were in Persian.

How did some suspect that Polo learned about China instead of by actually visiting it?

**Answer:** through contact with Persian traders

# Textual Question Answering

James the Turtle was always getting in trouble. Sometimes he'd reach into the freezer and empty out all the food. Other times he'd sled on the deck and get a splinter. His aunt Jane tried as hard as she could to keep him out of trouble, but he was sneaky and got into lots of trouble behind her back.

One day, James thought he would go into town and see what kind of trouble he could get into. He went to the grocery store and pulled all the pudding off the shelves and ate two jars. Then he walked to the fast food restaurant and ordered 15 bags of fries. He didn't pay, and instead headed home.

His aunt was waiting for him in his room. She told James that she loved him, but he would have to start acting like a well-behaved turtle.

After about a month, and after getting into lots of trouble, James finally made up his mind to be a better turtle.

1) What is the name of the trouble making turtle?

- A) Fries
- B) Pudding
- C) James
- D) Jane

2) What did James pull off of the shelves in the grocery store?

- A) pudding
- B) fries
- C) food
- D) splinters

# Conversational Question Answering

The Virginia governor's race, billed as the marquee battle of an otherwise anticlimactic 2013 election cycle, is shaping up to be a foregone conclusion. Democrat Terry McAuliffe, the longtime political fixer and moneyman, hasn't trailed in a poll since May. Barring a political miracle, Republican Ken Cuccinelli will be delivering a concession speech on Tuesday evening in Richmond. In recent ...

Q: What are the candidates **running** for?

A: Governor

Q: **Where?**

A: Virginia

Q: Who is the democratic candidate?

A: Terry McAuliffe

Q: Who is **his** opponent?

A: Ken Cuccinelli

Q: What party does **he** belong to?

A: Republican

Q: Which of **them** is winning?

# Long-form Question Answering

**Abstractive:**  
Answer made up of  
novel words and  
sentences composed  
through paraphrasing



Question

| How do jellyfish function without a brain and a nervous system?



Abstractive

Jellyfish don't have brains. Jellyfish have a nervous system. They know how to swim, they don't even know what to do with their bodies. They have a "nervous system", which is basically a bunch of nerve cells that connect to the body. The nervous system is made up of neurons, neurons, nerves, etc. Each neuron doesn't have a specific set of "neurons", it just has a different set of nerves. It has a set of sensory neurons, which are all connected to each other. It just knows that it does react to the stimulus, and it knows that the stimulus is there. They have a bunch of nervous systems that act like a filter to get information back.



Extractive

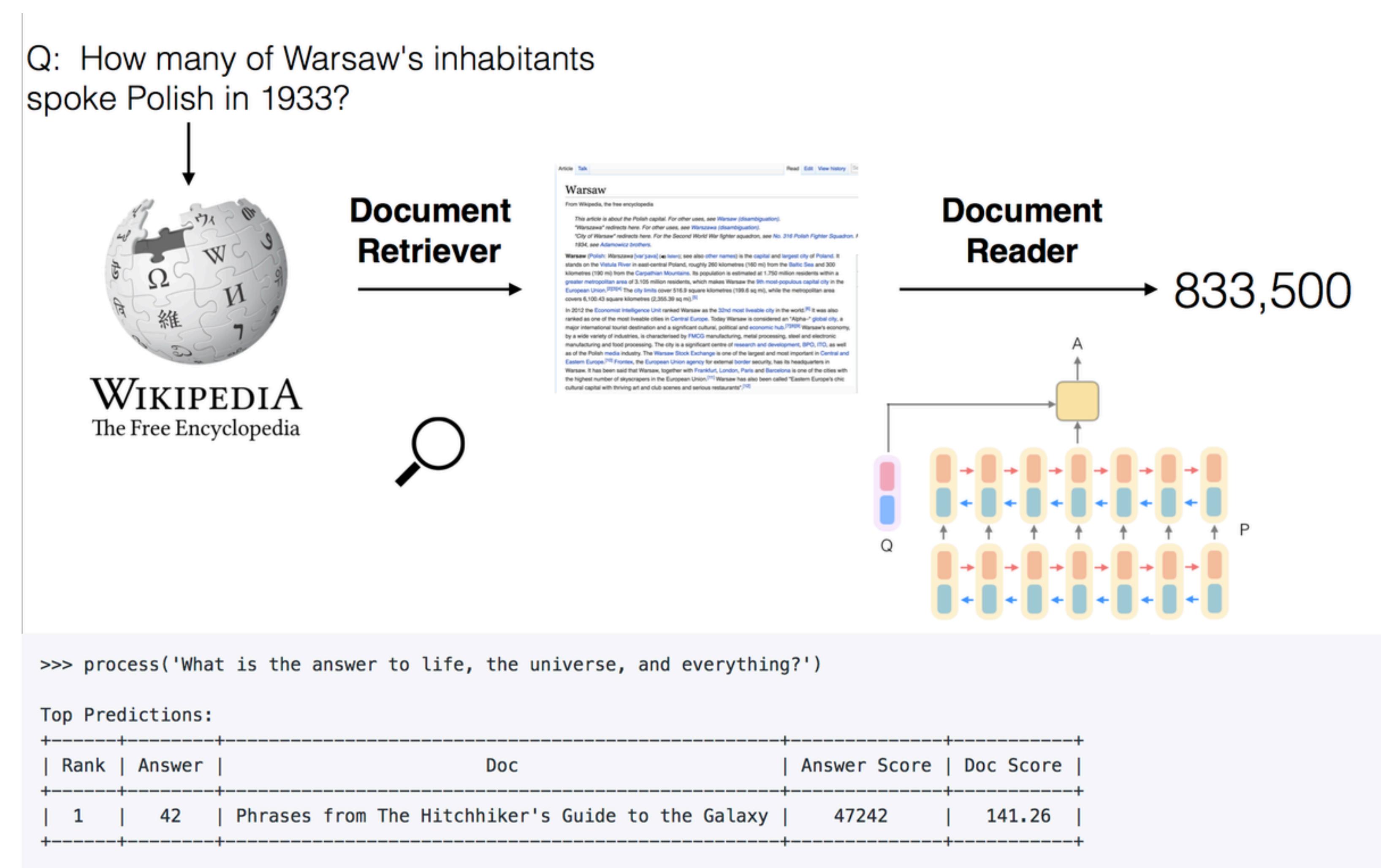
But none of these hypotheses make much sense for a jellyfish because they don't have brains at all. They just have a nerve net — a loose ring of neurons that runs around the rim of their pulsating bells. They have an unusual nervous system, 451 because jellyfish are not bilaterally symmetrical — that is, they don't have a left side and a right side. Jellyfish don't have brains, but their nervous systems detect smells, light and other stimuli, and they coordinate their physical responses.

**Extractive:**  
Select excerpts  
(extracts) and  
concatenate them  
to form the answer.

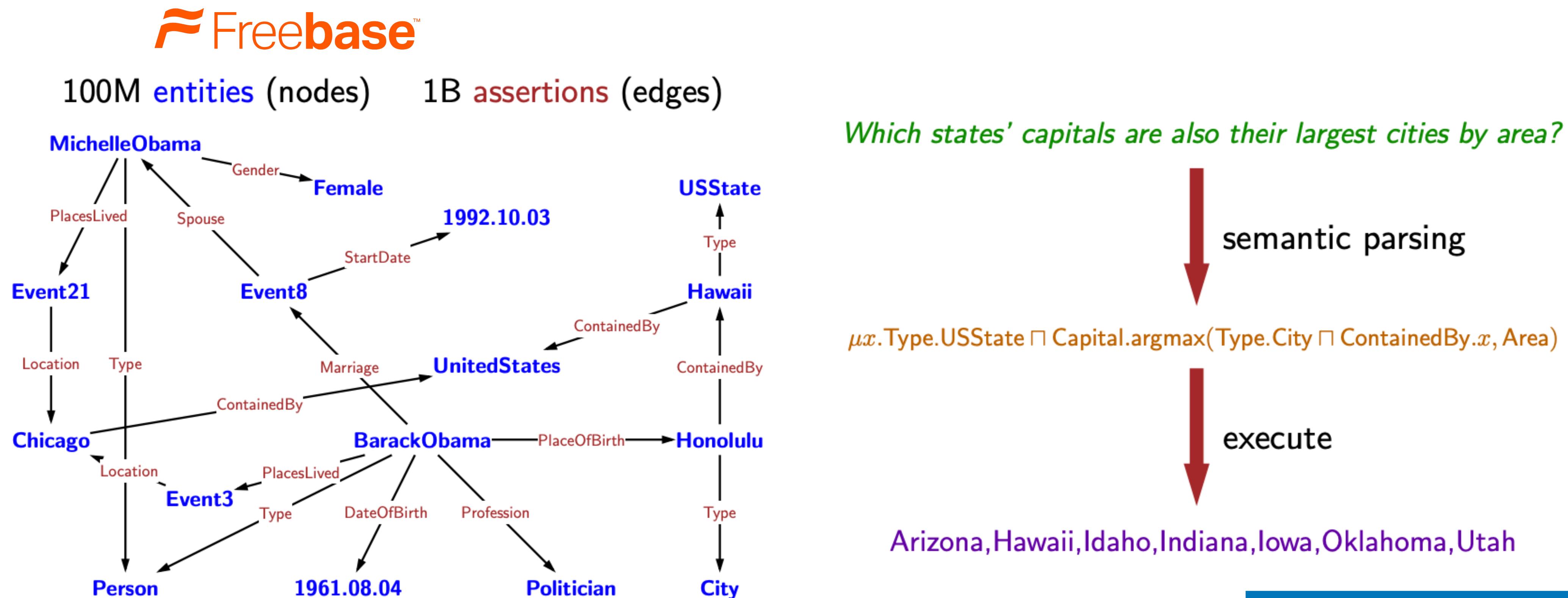
# Open-domain Question Answering

DrQA

- Factored into two parts:
  - Find documents that might contain an answer (handled with traditional **information retrieval**)
  - Finding an answer in a paragraph or a document (**reading comprehension**)



# Knowledge Base Question Answering



# Structured knowledge representation

# QA via semantic parsing

# Table-based Question Answering

Year	City	Country	Nations
1896	Athens	Greece	14
1900	Paris	France	24
1904	St. Louis	USA	12
...	...	...	...
2004	Athens	Greece	201
2008	Beijing	China	204
2012	London	UK	204

$x$  = Greece held its last Summer Olympics in which year?

$y$  = 2004

# Visual Question Answering



What color are her eyes?

What is the mustache made of?



How many slices of pizza are there?

Is this a vegetarian pizza?

# Reading Comprehension

# Stanford Question Answering Dataset (SQuAD)

## Passage

Super Bowl 50 was an American football game to determine the champion of the National Football League (NFL) for the 2015 season. The American Football Conference (AFC) champion Denver Broncos defeated the National Football Conference (NFC) champion Carolina Panthers 24–10 to earn their third Super Bowl title. The game was played on February 7, 2016, at Levi's Stadium in the San Francisco Bay Area at Santa Clara, California.

**Question:** Which NFL team won Super Bowl 50?

**Answer:** Denver Broncos

**Question:** What does AFC stand for?

**Answer:** American Football Conference

**Question:** What year was Super Bowl 50?

**Answer:** 2016

SQuAD 2.0:  
Have classifier/threshold to decide whether to take the most likely prediction as answer

- (passage, question, answer) triples
- Passage is from Wikipedia, question is crowd-sourced
- Answer must be a span of text in the passage (aka. “extractive question answering”)
- SQuAD 1.1: 100k answerable questions, SQuAD 2.0: another 50k unanswerable questions

# Stanford Question Answering Dataset (SQuAD)

Private schools, also known as independent schools, non-governmental, or nonstate schools, are not administered by local, state or national governments; thus, they retain the right to select their students and are funded in whole or in part by charging their students tuition, rather than relying on mandatory taxation through public (government) funding; at some private schools students may be able to get a scholarship, which makes the cost cheaper, depending on a talent the student may have (e.g. sport scholarship, art scholarship, academic scholarship), financial need, or tax credit scholarships that might be available.

**Along with non-governmental and nonstate schools, what is another name for private schools?**

**Gold answers:** ① independent ② independent schools ③ independent schools

**Along with sport and art, what is a type of talent scholarship?**

**Gold answers:** ① academic ② academic ③ academic

**Rather than taxation, what are private schools largely funded by?**

**Gold answers:** ① tuition ② charging their students tuition ③ tuition

3 gold answers are collected for each question

# Stanford Question Answering Dataset (SQuAD)

## SQuAD 1.1 evaluation:

- Two metrics: exact match (EM) and F1
  - Exact match: 1/0 accuracy on whether you match one of the three answers
  - F1: take each gold answer and system output as bag of words, compute precision, recall and harmonic mean. Take the max of the three scores.

Private schools, also known as independent schools, non-governmental, or nonstate schools, are not administered by local, state or national governments; thus, they retain the right to select their students and are funded in whole or in part by charging their students tuition, rather than relying on mandatory taxation through public (government) funding; at some private schools students may be able to get a scholarship, which makes the cost cheaper, depending on a talent the student may have (e.g. sport scholarship, art scholarship, academic scholarship), financial need, or tax credit scholarships that might be available.

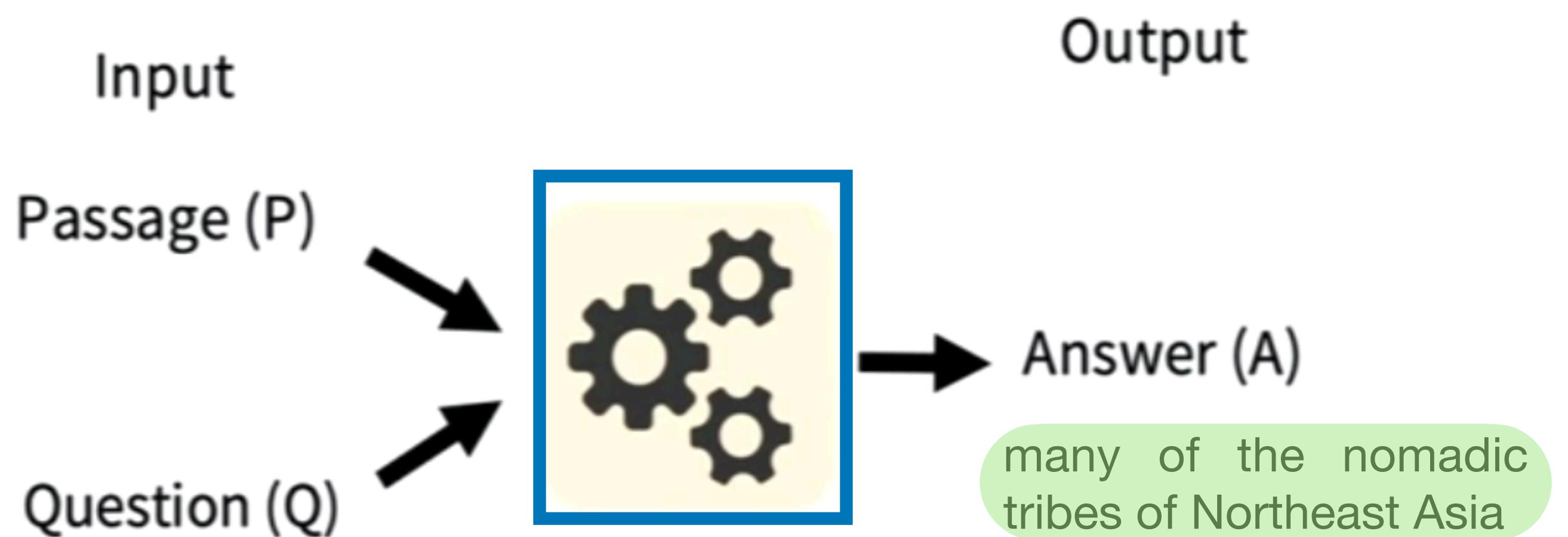
Q: Rather than taxation, what are private schools largely funded by?

A: {tuition, charging their students tuition, tuition}

# Models for Reading Comprehension

He came to power by **uniting** many of the nomadic tribes of Northeast Asia. After founding the Mongol Empire and being proclaimed "**Genghis Khan**", he started the Mongol invasions that resulted in the **conquest** of most of **Eurasia**. These included raids or invasions of the Qara Khitai, Caucasus, Khwarezmid Empire, Western Xia and Jin dynasties. These campaigns were often accompanied by wholesale massacres of the civilian populations – especially in the Khwarezmian and Xia controlled lands. By the end of his life, the Mongol Empire occupied a substantial portion of Central Asia and China.

Who did **Genghis Khan unite** before **he** began **conquering** the rest of **Eurasia**?

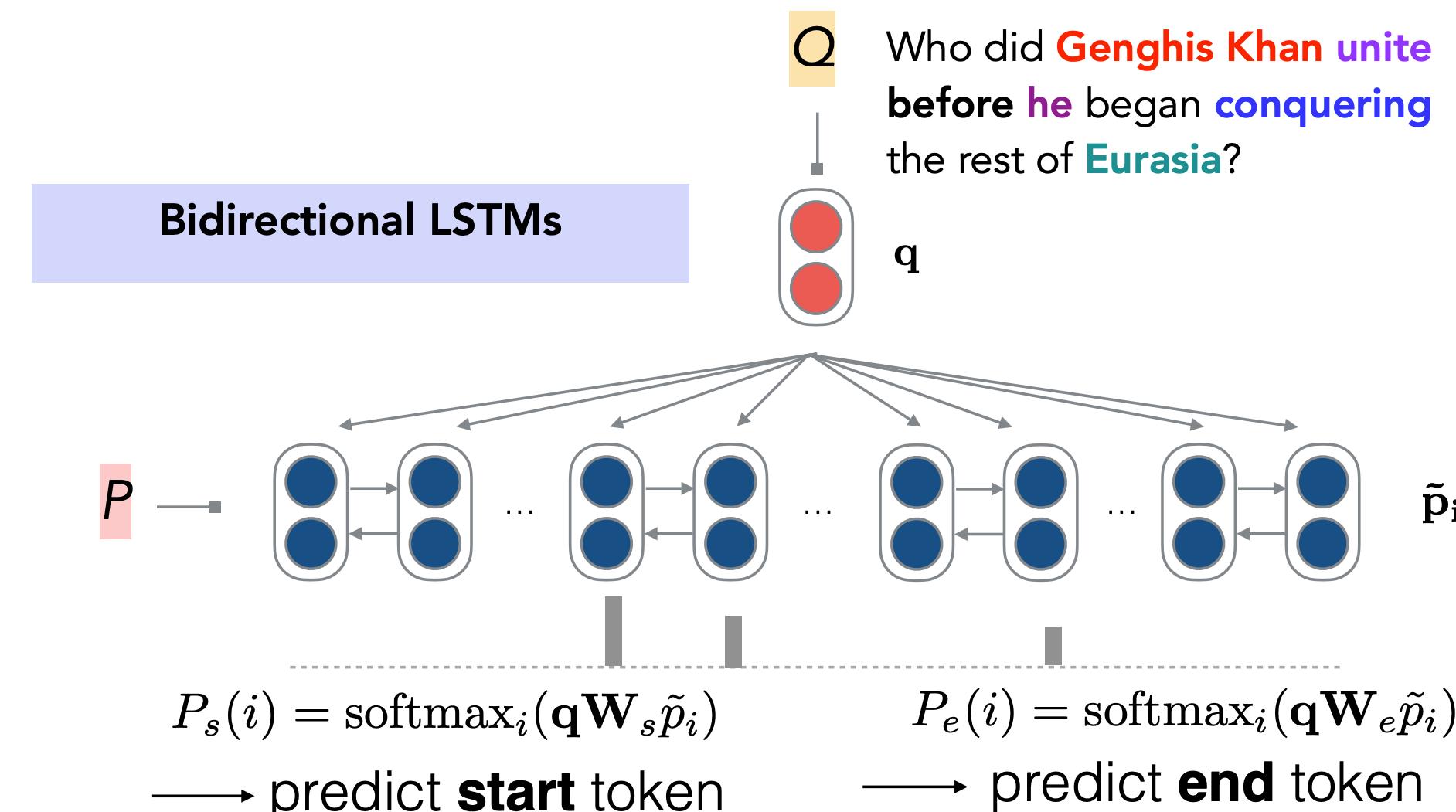


# Feature-based models

- Generate a list of candidate answers  $\{a_1, a_2, \dots, a_M\}$ 
  - Considered only the constituents in parse trees
- Define a feature vector  $\phi(p, q, a_i) \in \mathbb{R}^d$ :
  - Word/bigram frequencies
  - Parse tree matches
  - Dependency labels, length, part-of-speech tags
- Apply a (multi-class) **logistic regression model**

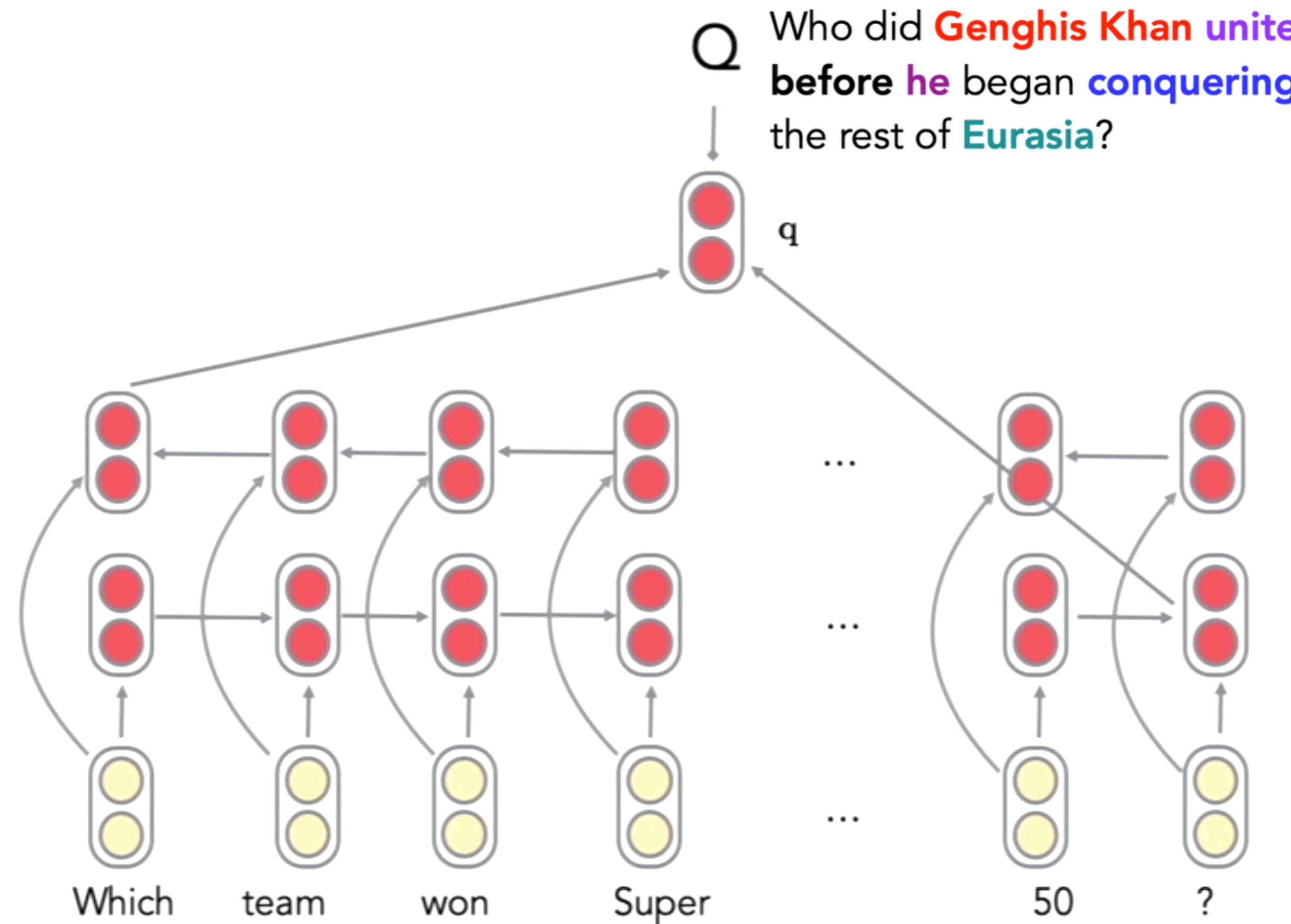
# Stanford Attentive Reader (Chen, Bolten, and Manning, 2016)

- Simple model with good performance
- Encode the question and passage word embeddings and BiLSTM encoders
- Use attention to predict start and end span



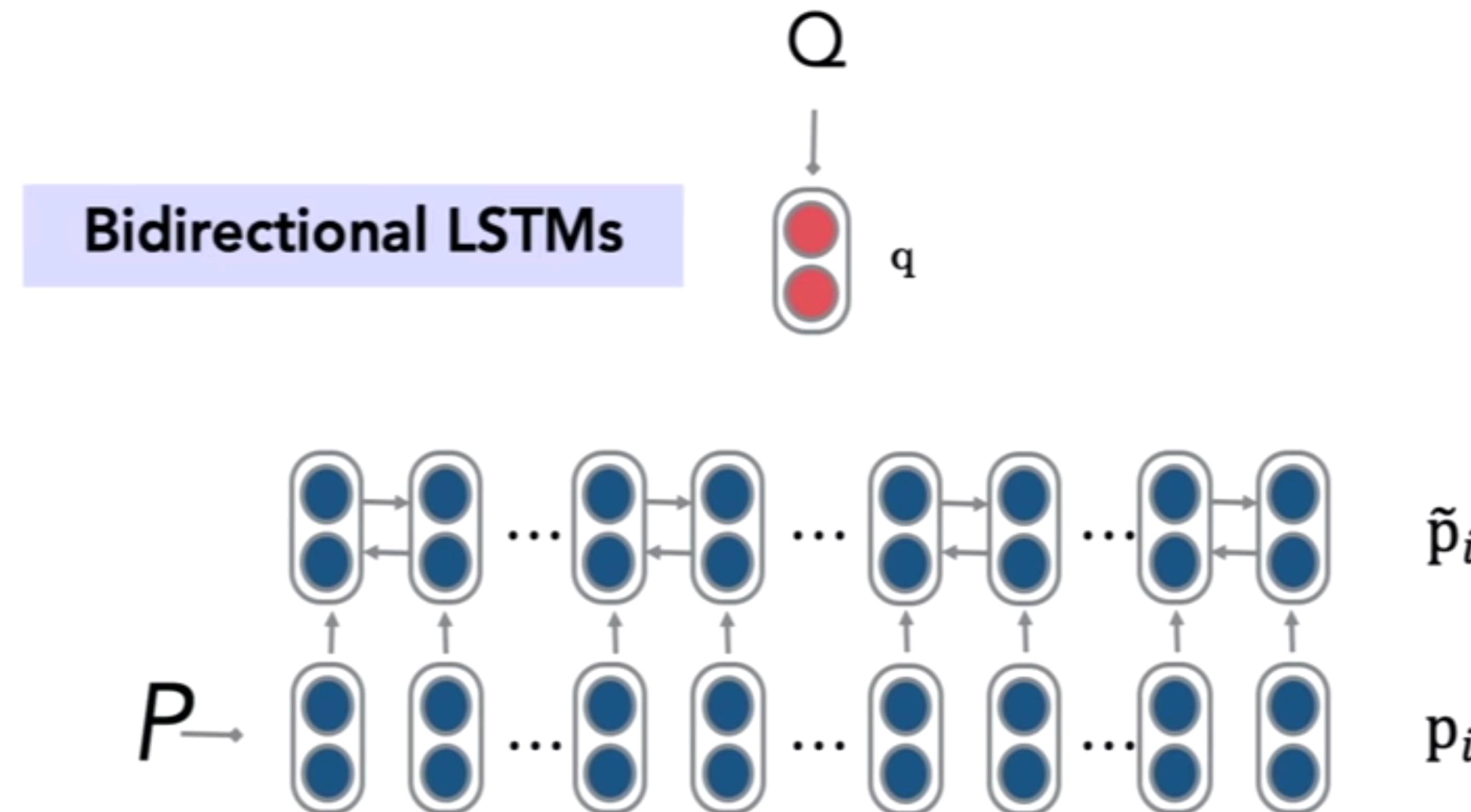
Also used in DrQA  
(Chen et al, 2017)

# Stanford Attentive Reader Question Encoder



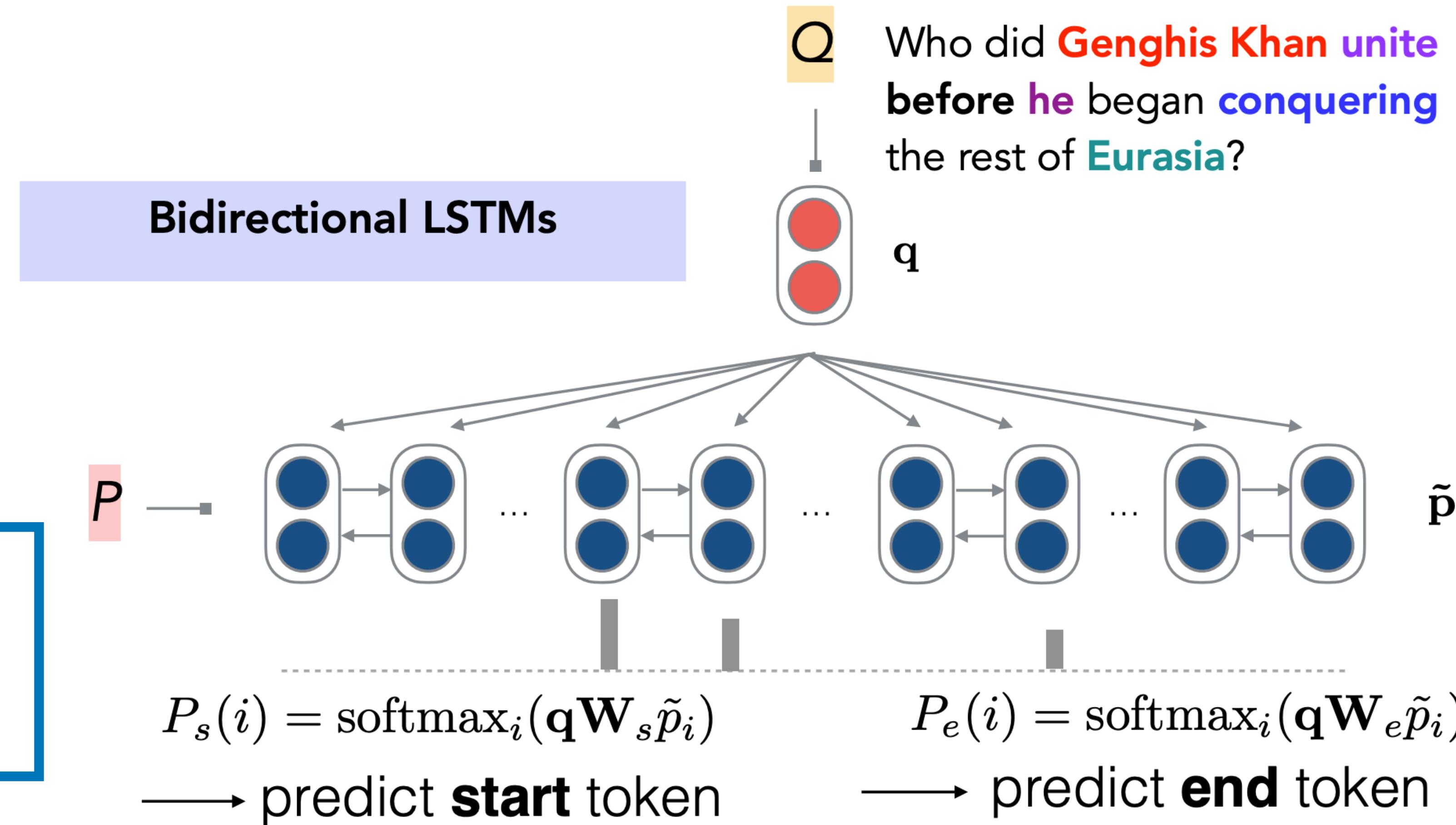
# Stanford Attentive Reader

## Passage encoder



# Stanford Attentive Reader

Use attention to predict span



# Stanford Attentive Reader++

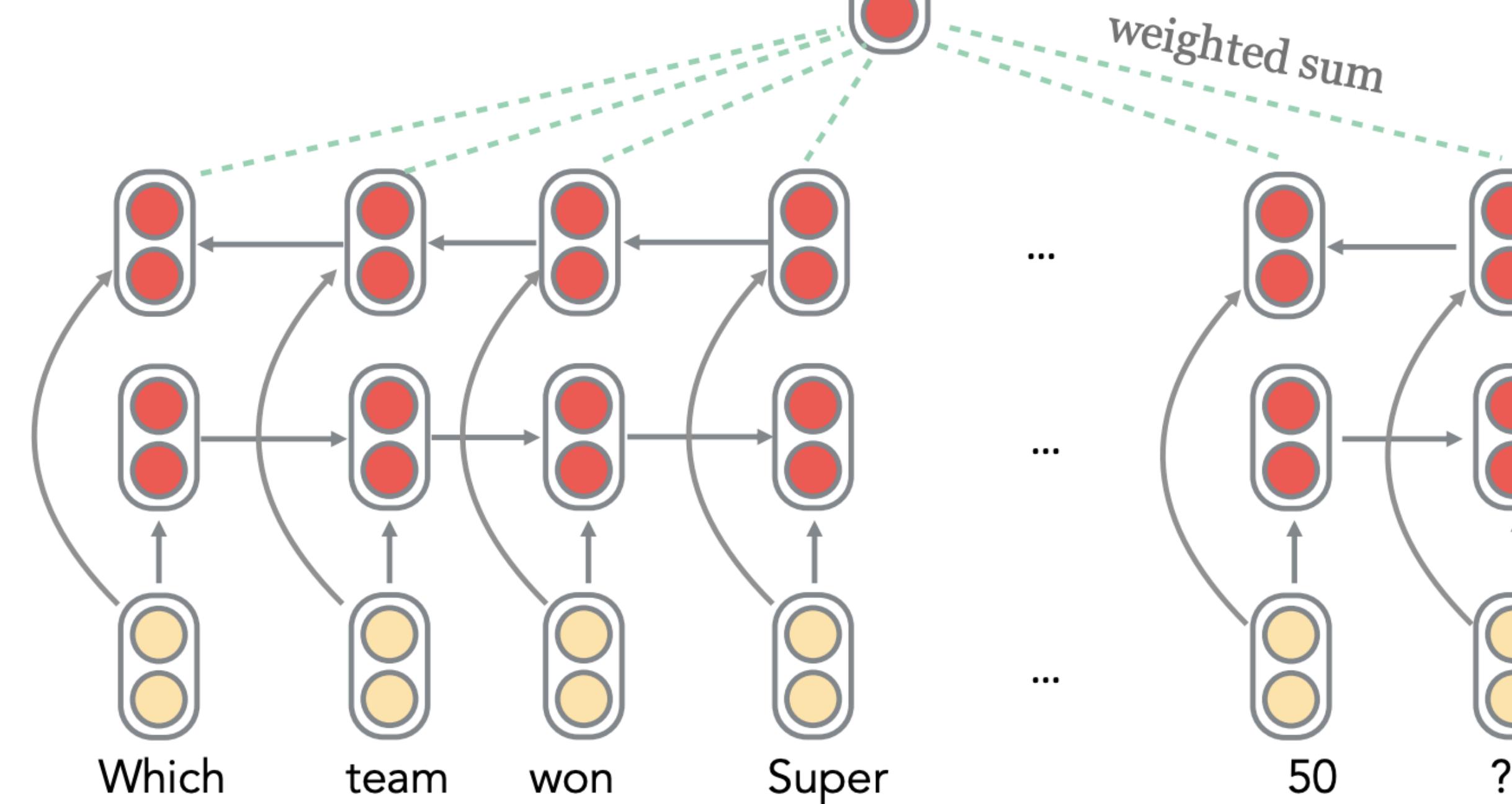
$$\mathbf{q} = \sum_j b_j \mathbf{q}_j$$

For learned  $\mathbf{w}$ ,  $b_j = \frac{\exp(\mathbf{w} \cdot \mathbf{q}_j)}{\sum_{j'} \exp(\mathbf{w} \cdot \mathbf{q}_{j'})}$

Take weighted sum  
of hidden states at all  
time steps of LSTM!

Q Which team won Super Bowl 50?

Deep 3 layer BiLSTM  
is better!



# Stanford Attentive Reader++

- $\mathbf{p}_i$ : Vector representation of each token in passage  
Made from concatenation of
  - Word embedding (GloVe 300d)
  - Linguistic features: POS & NER tags, one-hot encoded
  - Term frequency (unigram probability)
  - Exact match: whether the word appears in the question
    - 3 binary features: exact, uncased, lemma
- Aligned question embedding (“car” vs “vehicle”)

Improved passage  
word/position  
representations

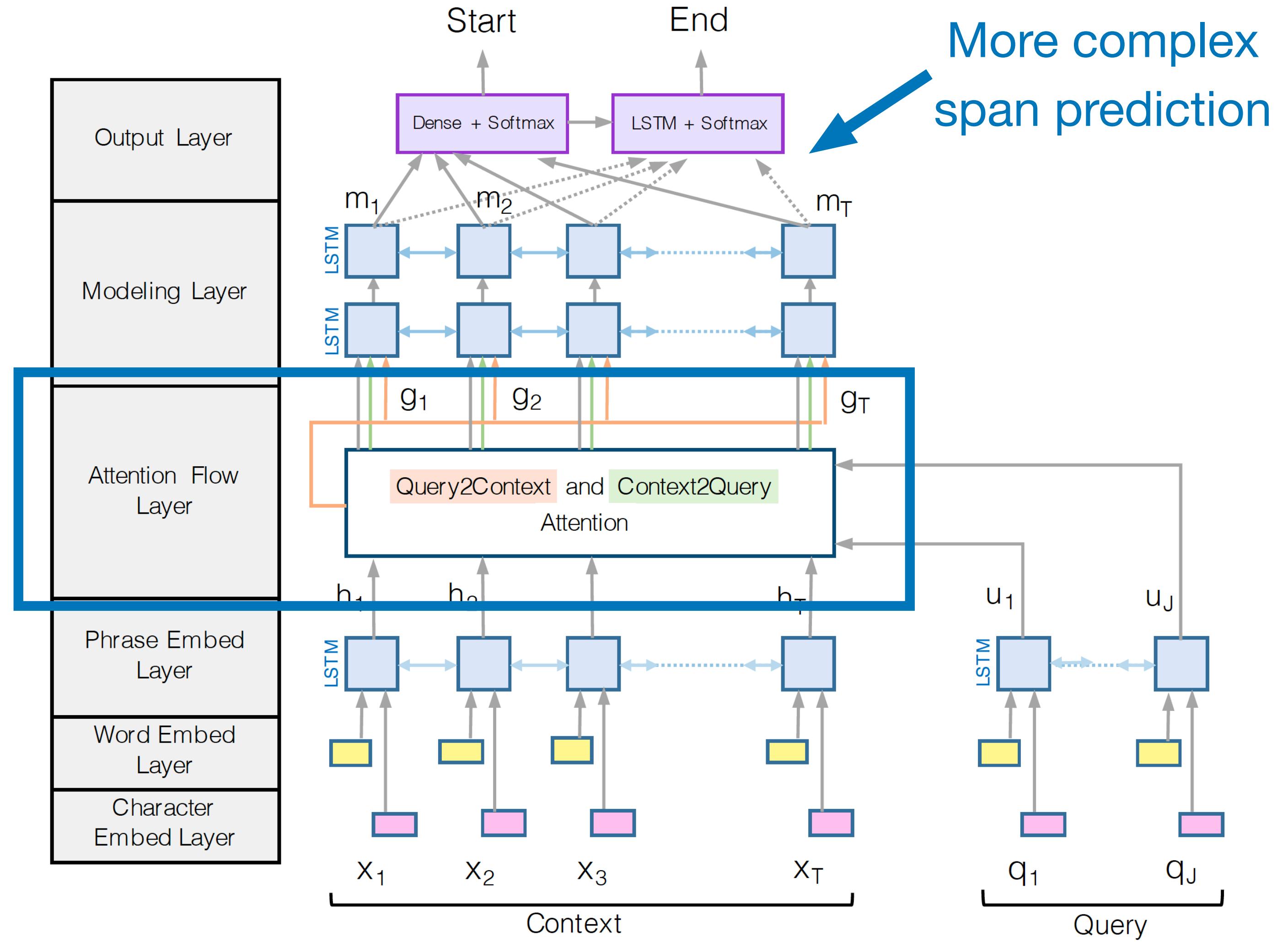
Matching of words in  
the question to words  
in the passage

$$f_{align}(p_i) = \sum_j a_{i,j} \mathbf{E}(q_j) \quad q_{i,j} = \frac{\exp(\alpha(\mathbf{E}(p_i)) \cdot \alpha(\mathbf{E}(q_j)))}{\sum_{j'} \exp(\alpha(\mathbf{E}(p_i)) \cdot \alpha(\mathbf{E}(q'_j)))}$$

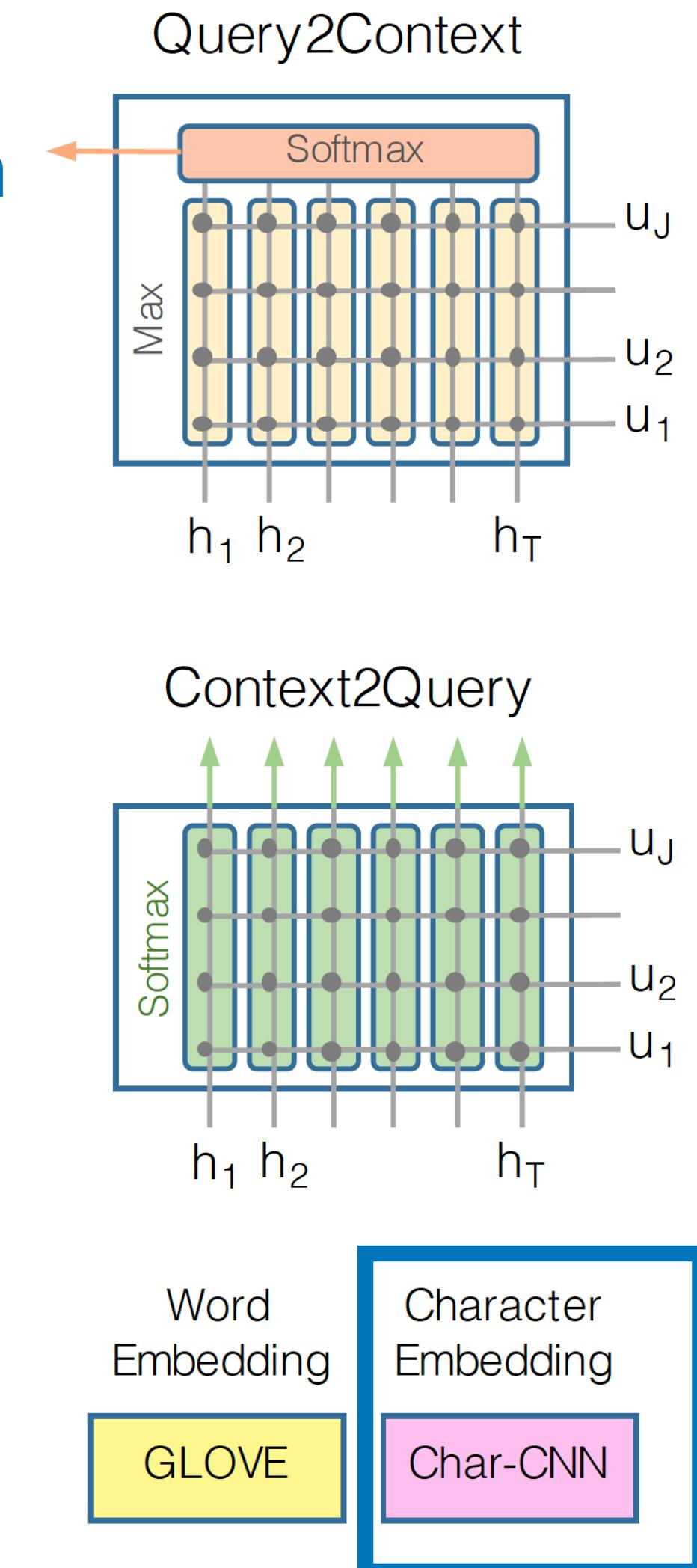
Where  $\alpha$  is a simple one layer FFNN

# BiDAF

Attention  
flowing between  
question (query)  
and  
passage (context)

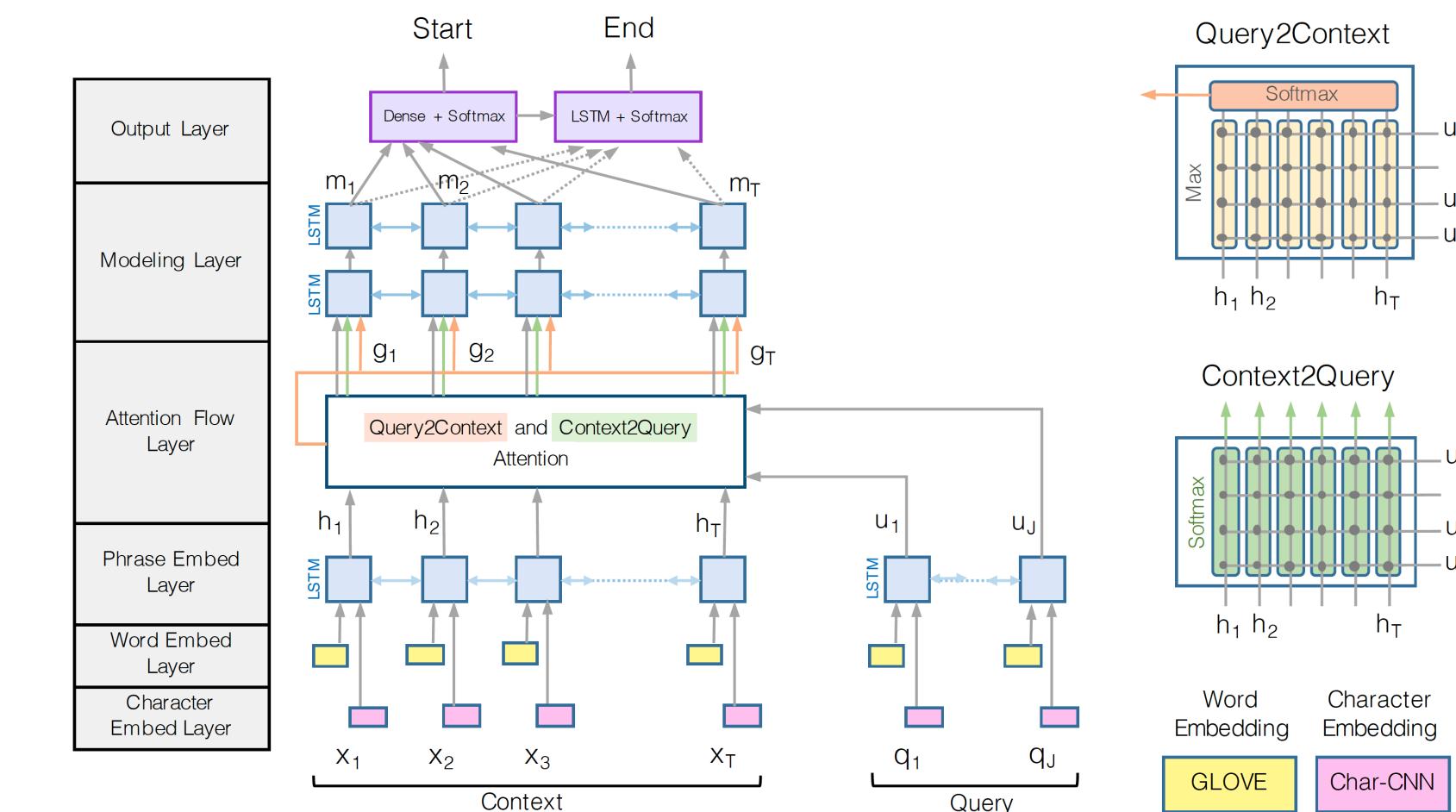


More complex  
span prediction



# BiDAF

- Encode the question using word/character embeddings; pass to an biLSTM encoder
- Encode the passage similarly
- Passage-to-question and question-to-passage attention
- Modeling layer: another BiLSTM layer
- Output layer: two classifiers for predicting start and end points
- The entire model can be trained in an end-to-end way



# BiDAF

- There are variants of and improvements to the BiDAF architecture over the years, but the central idea is **the Attention Flow layer**
- **Idea:** attention should flow both ways – from the context to the question and from the question to the context
- Make similarity matrix (with  $w$  of dimension  $6d$ ):

$$\mathbf{S}_{ij} = \mathbf{w}_{\text{sim}}^T [\mathbf{c}_i; \mathbf{q}_j; \mathbf{c}_i \circ \mathbf{q}_j] \in \mathbb{R}$$

- Context-to-Question (C2Q) attention:  
(which query words are most relevant to each context word)

$$\alpha^i = \text{softmax}(\mathbf{S}_{i,:}) \in \mathbb{R}^M \quad \forall i \in \{1, \dots, N\}$$

$$\mathbf{a}_i = \sum_{j=1}^M \alpha_j^i \mathbf{q}_j \in \mathbb{R}^{2h} \quad \forall i \in \{1, \dots, N\}$$

$c_i$  = passage word

$q_j$  = question word

Each are of dimension  $2d$

(from the bidirectional LSTM)

# BiDAF

- **Attention Flow Idea:** attention should flow both ways – from the context to the question and from the question to the context
- Question-to-Context (Q2C) attention:  
(the weighted sum of the most important words in the context with respect to the query – slight asymmetry through max)

$$\mathbf{m}_i = \max_j \mathbf{S}_{ij} \in \mathbb{R} \quad \forall i \in \{1, \dots, N\}$$

$$\beta = \text{softmax}(\mathbf{m}) \in \mathbb{R}^N$$

$$\mathbf{c}' = \sum_{i=1}^N \beta_i \mathbf{c}_i \in \mathbb{R}^{2h}$$

- For each passage position, output of BiDAF layer is:

$$\mathbf{b}_i = [\mathbf{c}_i; \mathbf{a}_i; \mathbf{c}_i \circ \mathbf{a}_i; \mathbf{c}_i \circ \mathbf{c}'] \in \mathbb{R}^{8h} \quad \forall i \in \{1, \dots, N\}$$

# SQuAD v1.1 performance (2017)

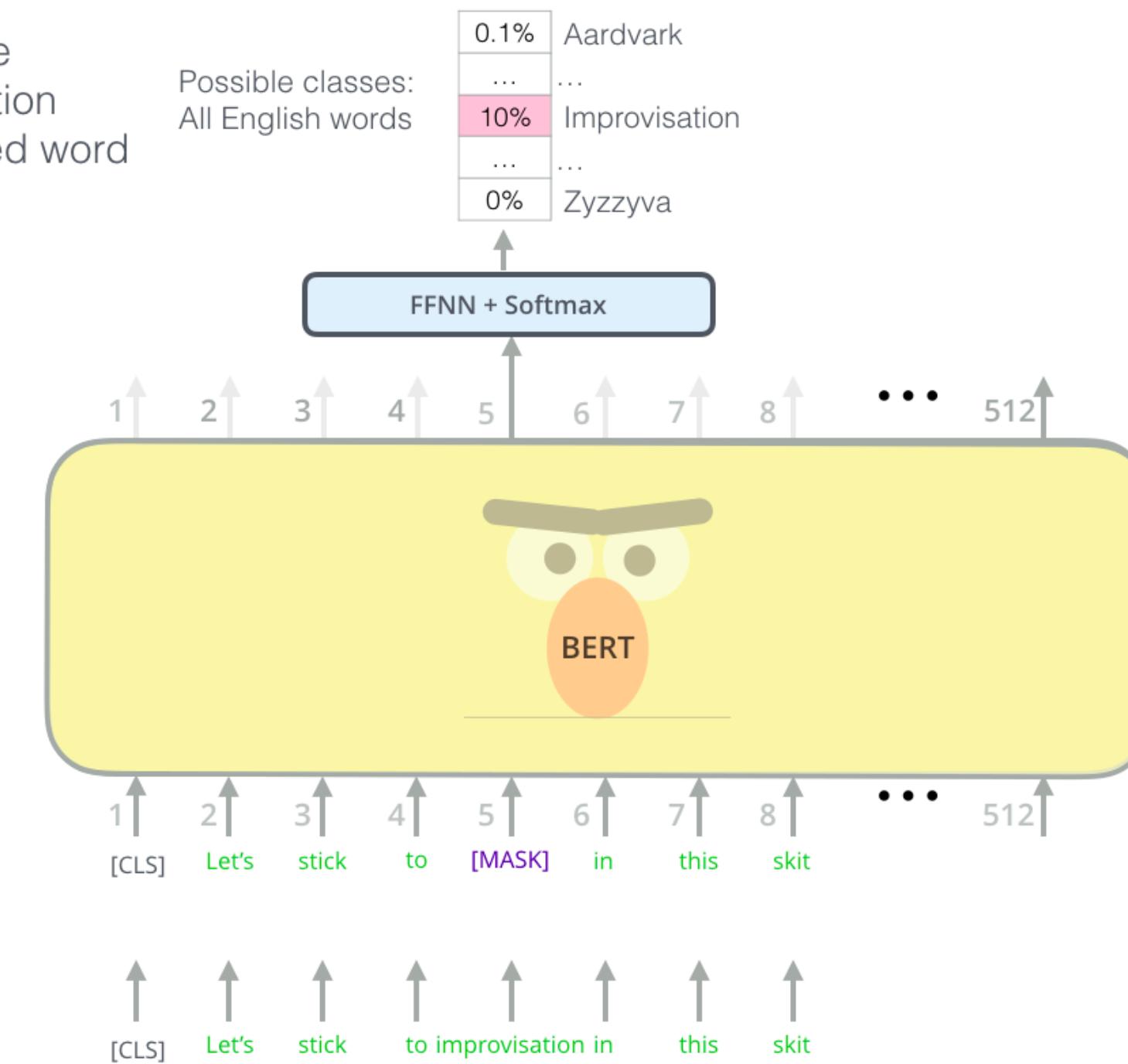
	F1
Logistic regression	51.0
Fine-Grained Gating (Carnegie Mellon U)	73.3
Match-LSTM (Singapore Management U)	73.7
DCN (Salesforce)	75.9
BiDAF (UW & Allen Institute)	77.3
Multi-Perspective Matching (IBM)	78.7
ReasoNet (MSR Redmond)	79.4
DrQA (Chen et al. 2017)	79.4
r-net (MSR Asia) [Wang et al., ACL 2017]	79.7
Human performance	91.2

# BERT-based models

Use the output of the masked word's position to predict the masked word

Randomly mask 15% of tokens

Input

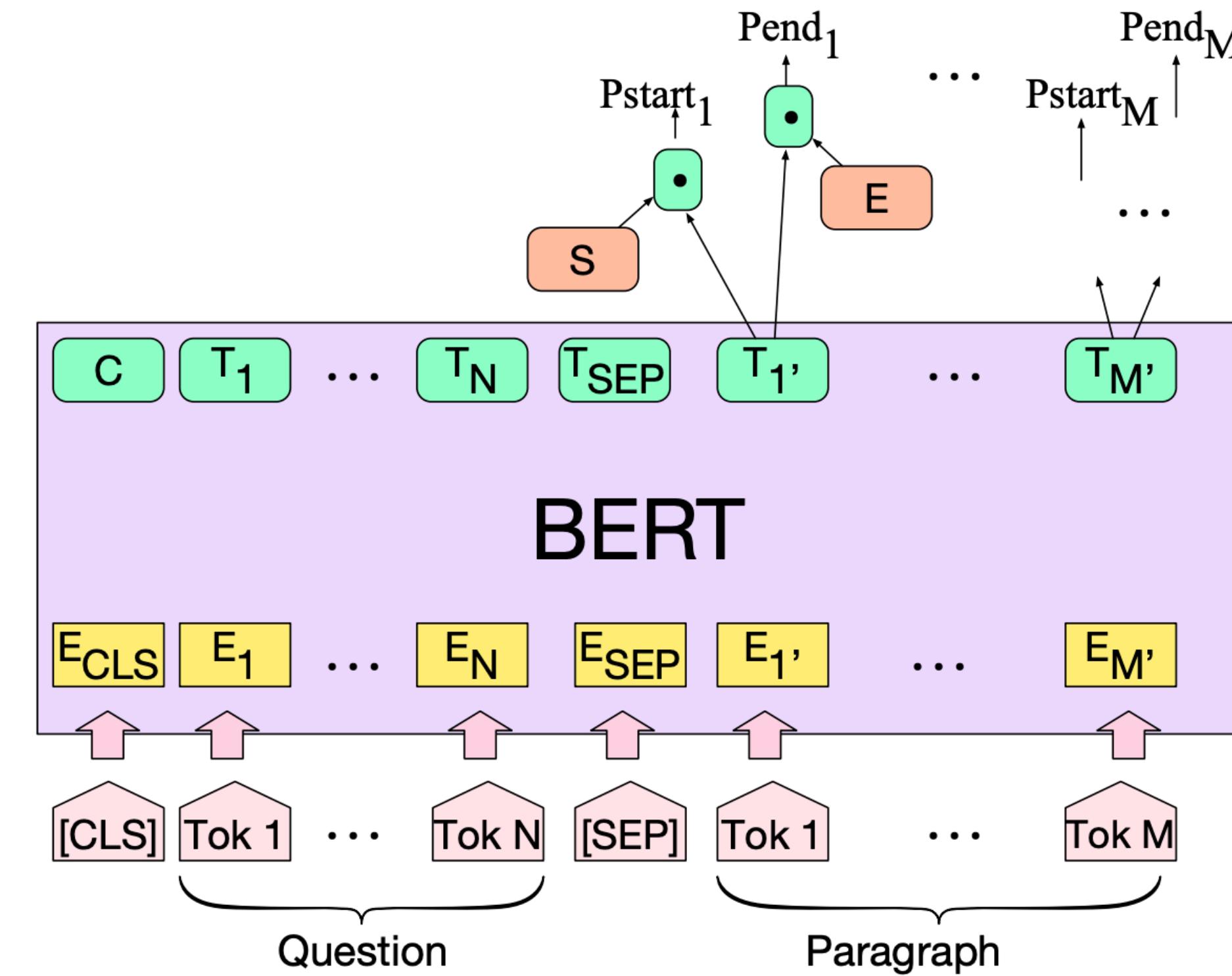


Pre-training

# BERT-based models

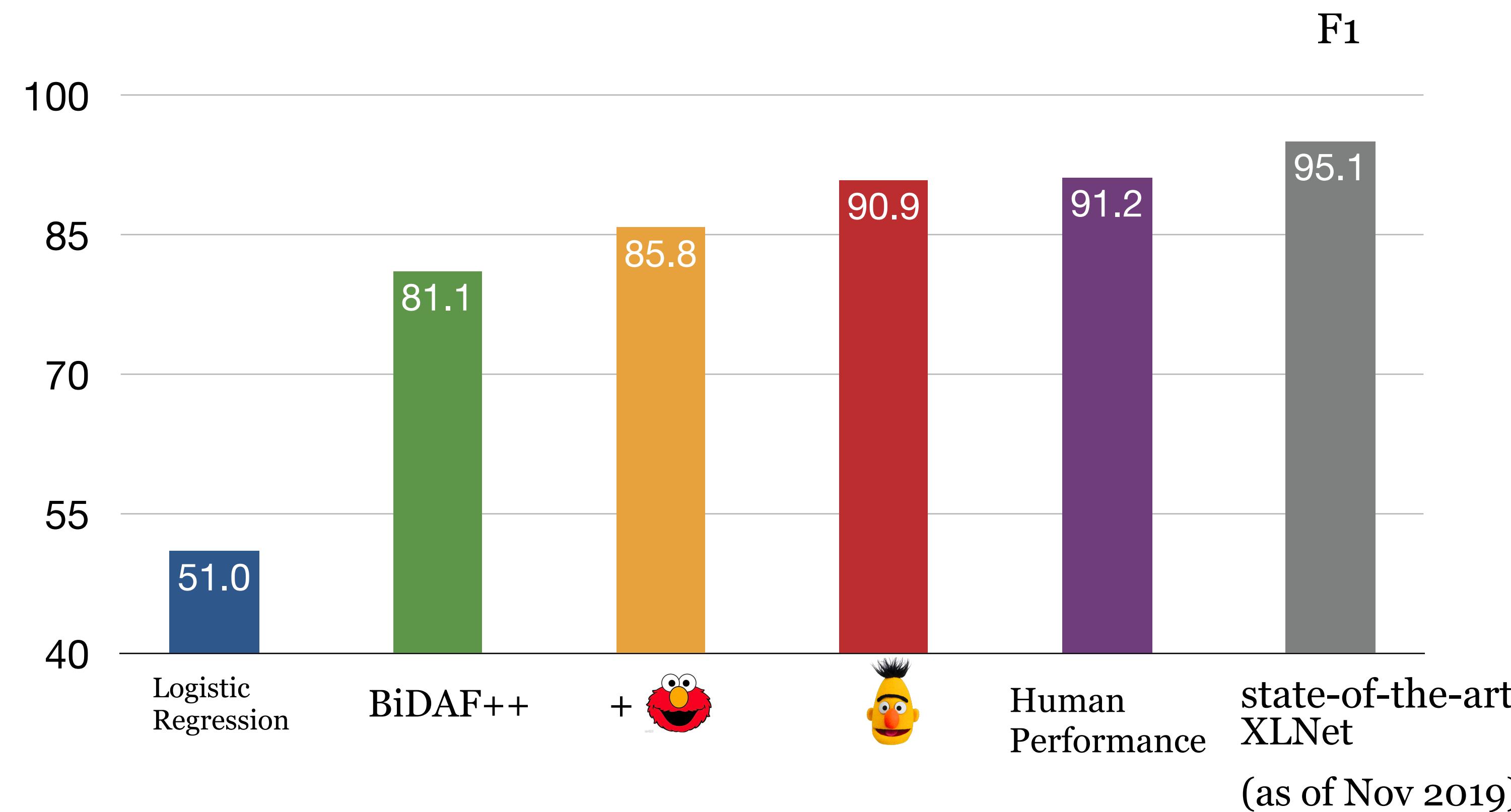
$$Pstart_i = \frac{e^{S \cdot T_i}}{\sum_j e^{S \cdot T_j}}$$

$$Pend_i = \frac{e^{E \cdot T_i}}{\sum_j e^{E \cdot T_j}}$$



- Concatenate question and passage as one single sequence separated with a [SEP] token, then pass it to the BERT encoder
- Train two classifiers on top of the passage tokens

# Experiments on SQuAD v1.1



\*: single model only

# Is Reading Comprehension solved?

**AI systems are beating humans  
in reading comprehension**

By Associated Press

January 24, 2018 | 2:25pm



Artificial Intelligence Jan 15, 2018

**AI Beats Humans at Reading Comprehension,  
but It Still Doesn't Truly Comprehend Language**



**AI Beat Humans at Reading! Maybe Not**

Microsoft and Alibaba claimed software could read like a human. There's more to the story than that.



Nope, maybe the SQuAD dataset is solved.

# Basic NLU errors

The Yuan dynasty is considered both a successor to the Mongol Empire and an imperial Chinese dynasty. It was the khanate ruled by the successors of Möngke Khan after the division of the Mongol Empire. In official Chinese histories, the Yuan dynasty bore the Mandate of Heaven, following the Song dynasty and preceding the Ming dynasty. The dynasty was established by Kublai Khan, yet he placed his grandfather Genghis Khan on the imperial records as the official founder of the

## **What dynasty came before the Yuan?**

*Gold Answers:* ① Song dynasty ② Mongol Empire  
③ the Song dynasty

*Prediction:* Ming dynasty [BERT (single model) (Google AI)]

# Is Reading Comprehension solved?

**Article:** Super Bowl 50

**Paragraph:** *“Peyton Manning became the first quarterback ever to lead two different teams to multiple Super Bowls. He is also the oldest quarterback ever to play in a Super Bowl at age 39. The past record was held by John Elway, who led the Broncos to victory in Super Bowl XXXIII at age 38 and is currently Denver’s Executive Vice President of Football Operations and General Manager. Quarterback Jeff Dean had jersey number 37 in Champ Bowl XXXIV.”*

**Question:** *“What is the name of the quarterback who was 38 in Super Bowl XXXIII?”*

**Original Prediction:** John Elway

**Prediction under adversary:** Jeff Dean

# SQuAD Limitations

- SQuAD has a number of limitations:
  - Only span-based answers (no yes/no, counting, implicit why)
  - Questions were constructed looking at passages
    - Not genuine information needs
    - Generally greater lexical and syntactic matching between question and answer span
  - Barely any multi-fact/sentence inference beyond coreference
- Nevertheless, it is a well-targeted, well-structured, clean dataset
  - The most used and competed QA dataset
  - A useful starting point for building systems in industry (although in-domain data always really helps!)

# Beyond SQuAD 1.1

- SQuAD 2.0 (Rajparkar et al, 2018)
  - unanswerable questions
- HotPotQA (Yang et al, 2018)
  - multi-hop reasoning
- QuAC(Choi et al, 2018) and CoQA (Reddy et al, 2018)
  - conversational QA
- Natural Questions (Kwiatkowski et al, 2019)
  - Real world questions issued to Google
- BooIQ (Clark et al, 2019)
  - Hard yes/no questions from Google queries

---

The Virginia governor's race, billed as the marquee battle of an otherwise anticlimactic 2013 election cycle, is shaping up to be a foregone conclusion. Democrat Terry McAuliffe, the longtime political fixer and moneymen, hasn't trailed in a poll since May. Barring a political miracle, Republican Ken Cuccinelli will be delivering a concession speech on Tuesday evening in Richmond. In recent ...

Q<sub>1</sub>: What are the candidates **running** for?

A<sub>1</sub>: Governor

R<sub>1</sub>: The Virginia governor's race

Q<sub>2</sub>: **Where**?

A<sub>2</sub>: Virginia

R<sub>2</sub>: The Virginia governor's race

Q<sub>3</sub>: Who is the democratic candidate?

A<sub>3</sub>: **Terry McAuliffe**

R<sub>3</sub>: Democrat Terry McAuliffe

Q<sub>4</sub>: Who is **his** opponent?

A<sub>4</sub>: **Ken Cuccinelli**

R<sub>4</sub>: Republican Ken Cuccinelli

Q<sub>5</sub>: What party does **he** belong to?

A<sub>5</sub>: Republican

R<sub>5</sub>: Republican Ken Cuccinelli

Q<sub>6</sub>: Which of **them** is winning?

A<sub>6</sub>: Terry McAuliffe

R<sub>6</sub>: Democrat Terry McAuliffe, the longtime political fixer and moneymen, hasn't trailed in a poll since May

---

CoQA (Reddy et al, 2018)

# Natural Questions

## Real world queries to Google

### Example 1

**Question:** what color was john wilkes booth's hair

**Wikipedia Page:** John\_Wilkes\_Booth

**Long answer:** Some critics called Booth “the handsomest man in America” and a “natural genius”, and noted his having an “astonishing memory”; others were mixed in their estimation of his acting. He stood 5 feet 8 inches (1.73 m) tall, had jet-black hair, and was lean and athletic. Noted Civil War reporter George Alfred Townsend described him as a “muscular, perfect man” with “curling hair, like a Corinthian capital”.

**Short answer:** jet-black

### Example 2

**Question:** can you make and receive calls in airplane mode

**Wikipedia Page:** Airplane\_mode

**Long answer:** Airplane mode, aeroplane mode, flight mode, offline mode, or standalone mode is a setting available on many smartphones, portable computers, and other electronic devices that, when activated, suspends radio-frequency signal transmission by the device, thereby disabling Bluetooth, telephony, and Wi-Fi. GPS may or may not be disabled, because it does not involve transmitting radio waves.

**Short answer:** BOOLEAN:NO

### Example 3

**Question:** why does queen elizabeth sign her name elizabeth r

**Wikipedia Page:** Royal\_sign-manual

**Long answer:** The royal sign-manual usually consists of the sovereign’s regnal name (without number, if otherwise used), followed by the letter R for Rex (King) or Regina (Queen). Thus, the signs-manual of both Elizabeth I and Elizabeth II read Elizabeth R. When the British monarch was also Emperor or Empress of India, the sign manual ended with R I, for Rex Imperator or Regina Imperatrix (King-Emperor/Queen-Empress).

**Short answer:** NULL

# Beyond SQuAD 1.1

## BoolQ

### Hard yes/no questions from Google queries

---

**Q:** Has the UK been hit by a hurricane?

**P:** The Great Storm of 1987 was a violent extratropical cyclone which caused casualties in England, France and the Channel Islands ...

**A:** Yes. [An example event is given.]

**Q:** Does France have a Prime Minister and a President?

**P:** ... The extent to which those decisions lie with the Prime Minister or President depends upon ...

**A:** Yes. [Both are mentioned, so it can be inferred both exist.]

**Q:** Have the San Jose Sharks won a Stanley Cup?

**P:** ... The Sharks have advanced to the Stanley Cup finals once, losing to the Pittsburgh Penguins in 2016 ...

**A:** No. [They were in the finals once, and lost.]

---

(Clark et al, 2019)

# DrQA: Document Retrieval

Dataset	Wiki Search	Doc. Retriever	
		plain	+bigrams
SQuAD	62.7	76.1	<b>77.8</b>
CuratedTREC	81.0	85.2	<b>86.0</b>
WebQuestions	73.7	<b>75.5</b>	74.4
WikiMovies	61.7	54.4	<b>70.3</b>

Traditional  
tf.idf  
inverted  
index +  
efficient  
bigram  
hash

For **70-86%** of questions, the answer segment appears in the top 5 articles

# DrQA Demo

<https://github.com/facebookresearch/DrQA>

Hi!



Hello! Please ask a question.

What is question answering?



a computer science discipline within the fields of information retrieval and natural language processing

Who was the winning pitcher in the 1956 World Series?



Don Larsen

What is the answer to life, the universe, and everything?



42

# General Questions

Combined with **Web search**, DrQA can answer **57.5%** of **trivia questions** correctly



**Q:** The Dodecanese Campaign of WWII that was an attempt by the Allied forces to capture islands in the Aegean Sea was the inspiration for which acclaimed 1961 commando film?

**A:** The Guns of Navarone

**Q:** American Callan Pinckney's eponymously named system became a best-selling (1980s-2000s) book/video franchise in what genre?

**A:** Fitness

# Summary

- Many different types of question answering
- Reading comprehension
  - Given passage + question, come up with answer
  - SQuAD: answer is span of text in passage
  - Train classifier to predict span
- Reading comprehension is not solved!
- Lots of ongoing work on QA!