University of BRISTOL

Chapter 23, Speech and Language Processing, 3rd edition draft, Jurafsky & Martin (2020).

9.5 Question Answering

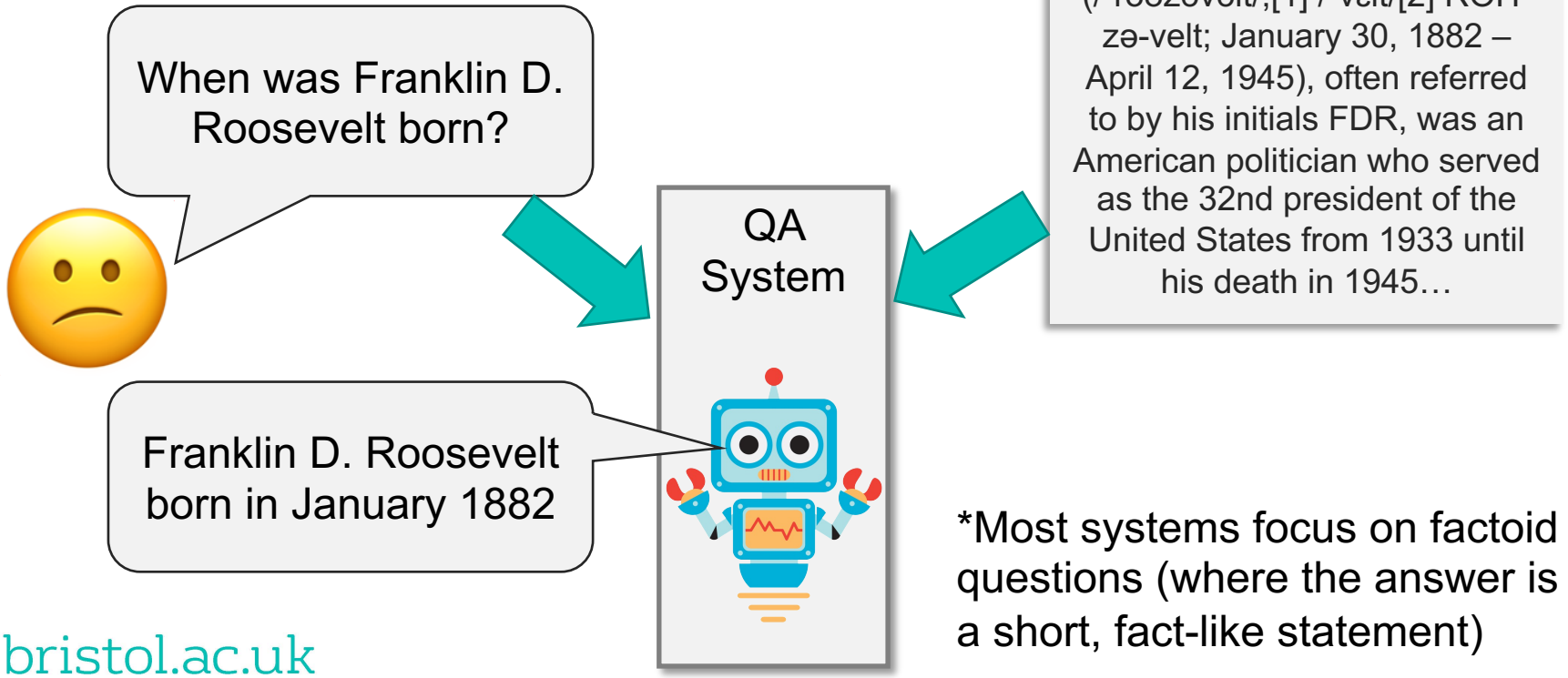
Edwin Simpson

Department of Computer Science,

University of Bristol, UK.

bristol.ac.uk

Question Answering (QA)



*Most systems focus on factoid questions (where the answer is a short, fact-like statement)

Pipeline Approach

Preprocessing

Syntax parsing

Named entity recognition

Relation extraction

Coreference resolution

Semantic parsing

Answer type detection

Candidate answer retrieval

Ranking candidate answers

A

Coreference

Chapter 21, Speech and Language Processing, 3rd edition draft, Jurafsky & Martin (2020).

- Preprocessing
- Syntax parsing
- Named entity recognition
- Relation extraction
- **Coreference resolution**
- Semantic parsing

***Victoria Chen**, CFO of Megabucks Banking, saw **her** pay jump to $2.3million…*

How much does Victoria Chen earn?

This step resolves ‘her’ to ‘Victoria Chen’.

Detect mentions: named entities, noun phrases, possessive pronouns.

Classify pairs of mentions as referring to the same entity or not.

Semantic Roles

Chapter 19, Semantic role labelling, Speech and Language Processing, 3rd edition draft, Jurafsky & Martin (2020).
Using Semantic Roles to Improve Question Answering, Shen and Lapata (2007).

- Preprocessing
- Syntax parsing
- Named entity recognitio
- Relation extraction
- Coreference resolution
- Semantic role labelling

[Lee]$_{Seller}$ sold a [text book]$_{Goods}$ to [Abby]$_{Buyer}$.

[Who]$_{Buyer}$ owns the [book]$_{Goods}$ that [Lee]$_{Seller}$ sold?

This step labels the roles of different entities in an event or situation.

This can be performed using sequence labelling.

It is a kind of shallow semantic parsing.

QA Pipeline

Preprocessing

Syntax parsing

Named entity recognition

Relation extraction

Coreference resolution

Semantic parsing

The features of candidate answers are compared with those of the question to determine the best match.

Answer type detection

Candidate answer retrieval

Ranking candidate answers

A

Chapter 19, Semantic role labelling, Speech and Language Processing, 3rd edition draft, Jurafsky & Martin (2020).

QA with a Pretrained Deep Neural Network

For each token in the passage, compute:

- Probability that it is the start of the answer, P_{start_i}
- Probability that it is the end of the answer, P_{end_i}

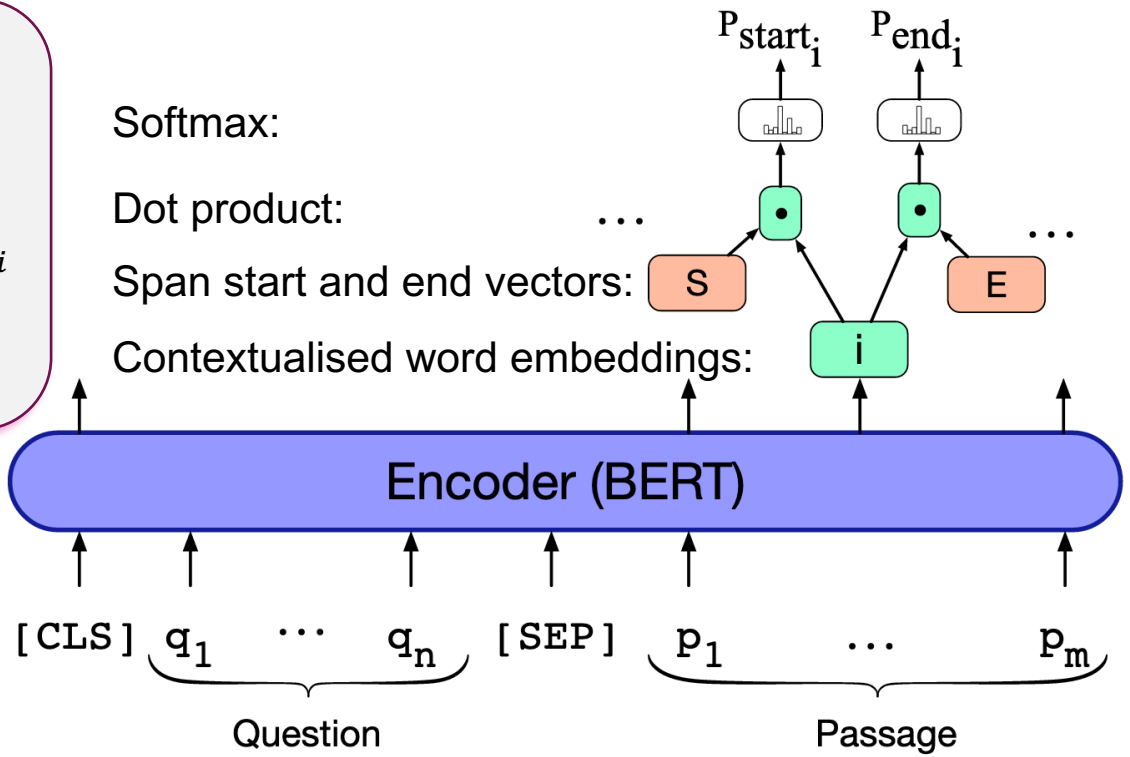


Figure 23.12, Speech and Language Processing, 3rd edition draft, Jurafsky & Martin (2020).

QA with a Pretrained Deep Neural Network

For each token in the passage, compute:

- Probability that it is the start of the answer, P_{start_i}
- Probability that it is the end of the answer, P_{end_i}

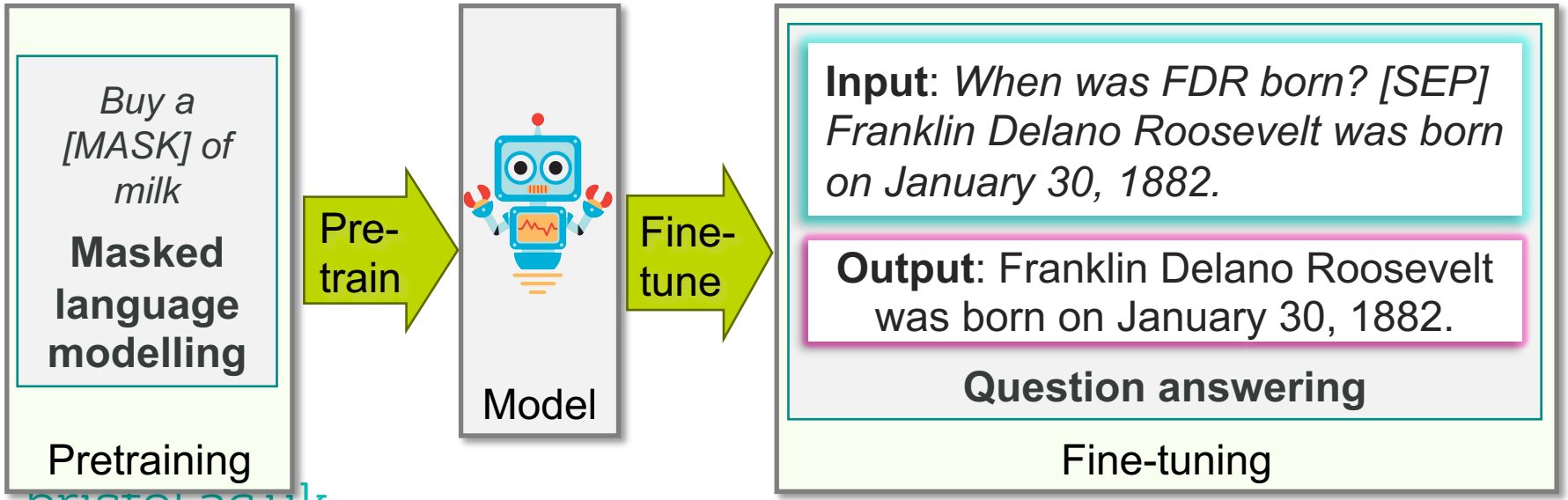
To make a prediction, compute a score for each possible span from token i to j:

- $P_{start_i} \times P_{end_j}$
- This estimates the joint probability of starting at i and ending at j
- Choose the highest-scoring span as the prediction

Section 23.2.2, Speech and Language Processing, 3rd edition draft, Jurafsky & Martin (2020).

QA with a Pretrained Deep Neural Network

- As input to the neural network for QA, we concatenate the source passage and query



Summary

- Question answering systems go beyond extracting structured data to directly answer a user's query.
- Matching questions to answers uses various features, including named entities and relations.
- Pretrained models like BERT are trained end-to-end to perform QA.
- BERT's contextualised embeddings are used by a classification layer to predict if a token is the start or end of an answer span.