THE UNIVERSITY OF MELBOURNE

Comp90042 Workshop Week 7

Table of Content

Contextual representation

Language is complex, one word can have multiple meanings

Contextual representation

Language is complex, one word can have multiple meanings

He sat on the **bank** *of the river*

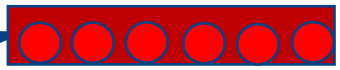
He cashed a check at the **bank**

Contextual representation

Language is complex, one word can have multiple meanings

He sat on the **bank** *of the river*

He cashed a check at the **bank**



Word embedding vector for **bank**

Word embeddings: All "**bank**" with different meanings share the same vector

Contextual representation

Language is complex, one word can have multiple meanings

He sat on the **bank** *of the river*

He cashed a check at the **bank**

How to determine the meaning of a word?

Contextual representation

Language is complex, one word can have multiple meanings

He sat on the **bank** *of the river*

He cashed a check at the **bank**

How to determine the meaning of a word?

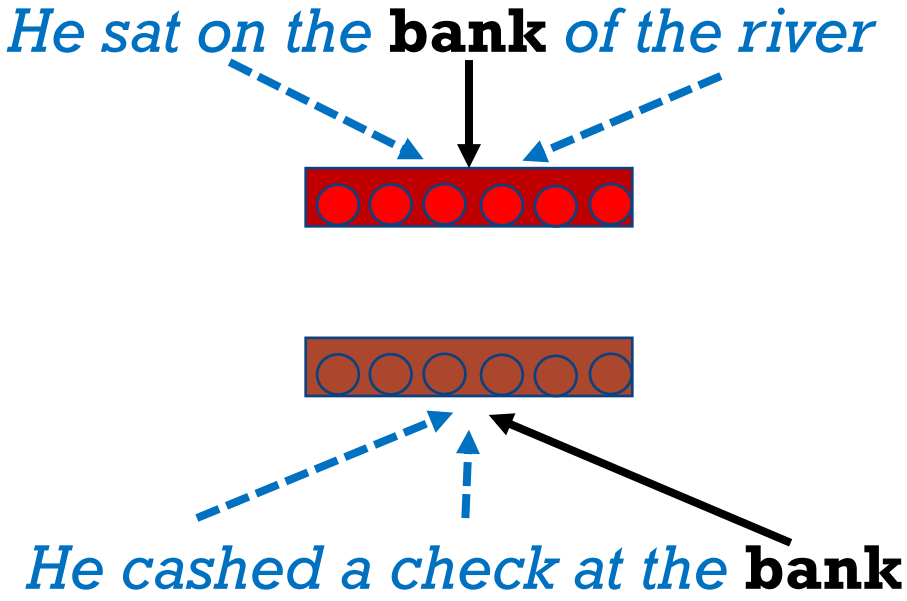
Context !

Contextual representation

Contextual representation captures the different senses or nuances of the word depending on the context.

Contextual representation

Contextual representation captures the different senses or nuances of the word depending on the context.



RNN

He cashed a check at the **bank**

Bank[1]:…a financial institution that accepts deposits and channels the money into lending activities

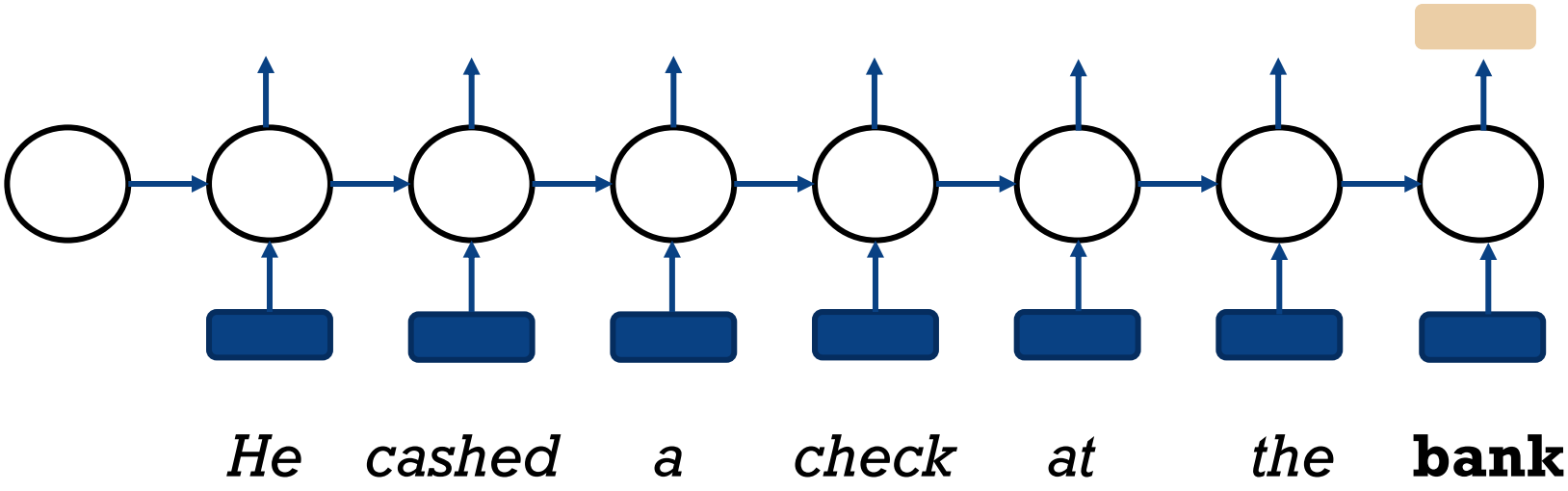
Bank[2]: sloping land (especially the slope beside a body of water)

Classification: determine which meaning of **bank** is in the sentence.

RNN

Bank[1]:…a financial institution that accepts deposits and channels the money into lending activities

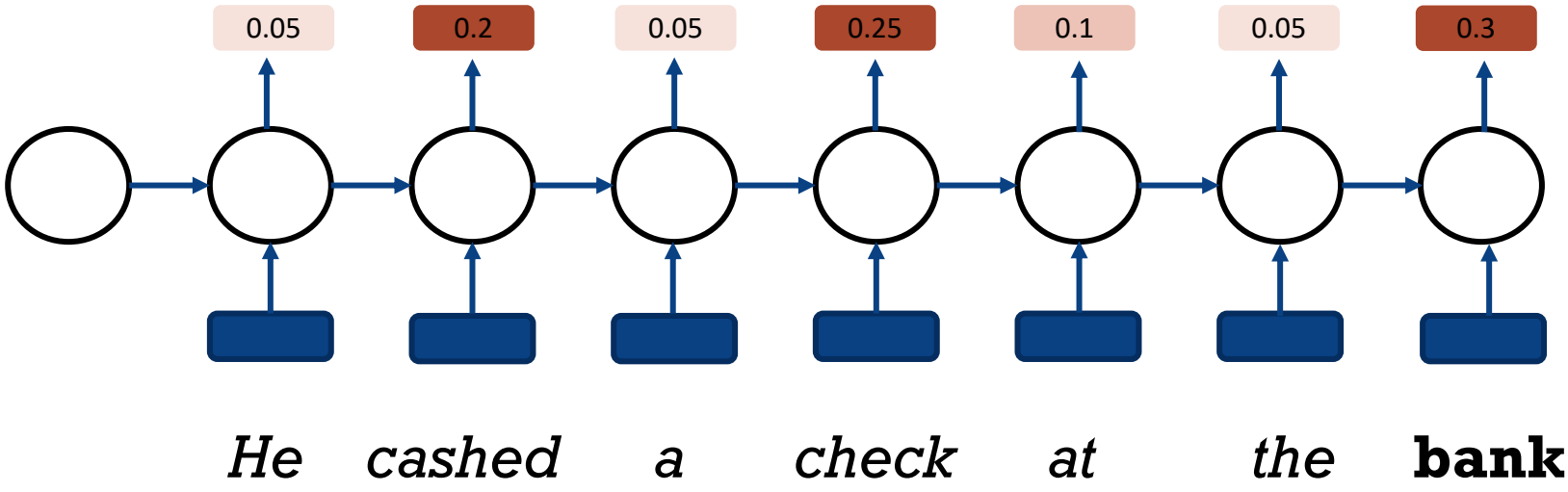
Bank[2]: sloping land (especially the slope beside a body of water)



RNN + Attention

Bank[1]:…a financial institution that accepts deposits and channels the money into lending activities

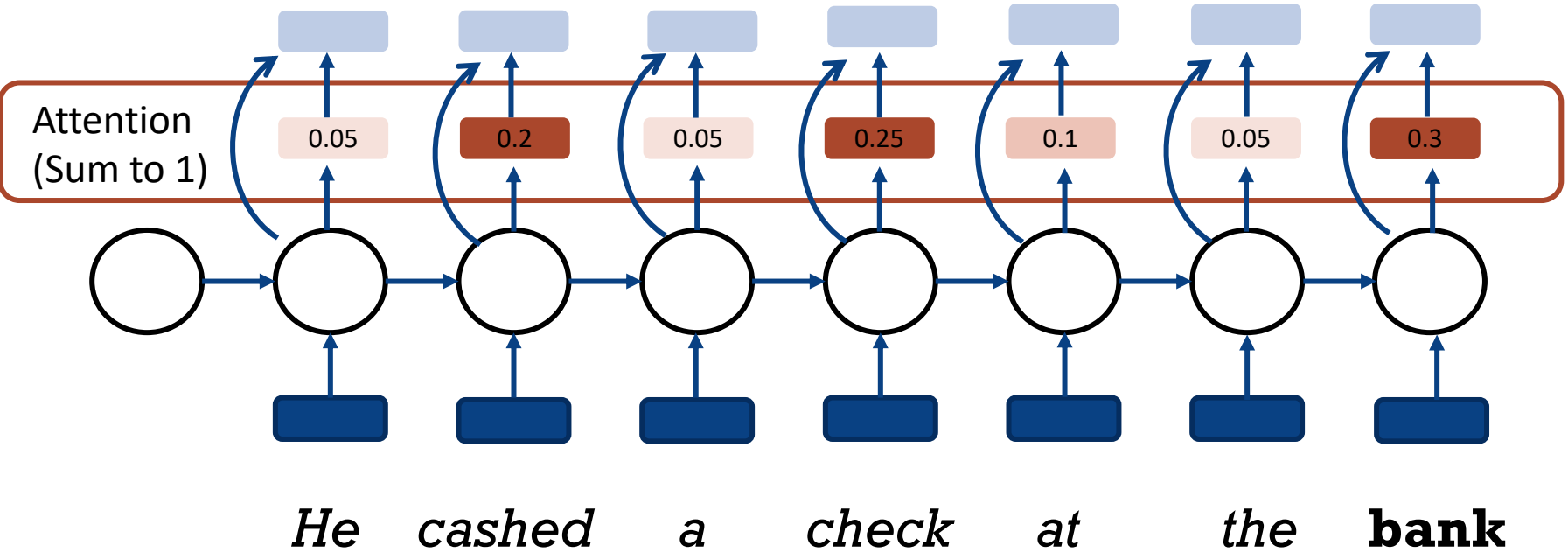
Bank[2]: sloping land (especially the slope beside a body of water)



RNN + Attention

Bank[1]:…a financial institution that accepts deposits and channels the money into lending activities

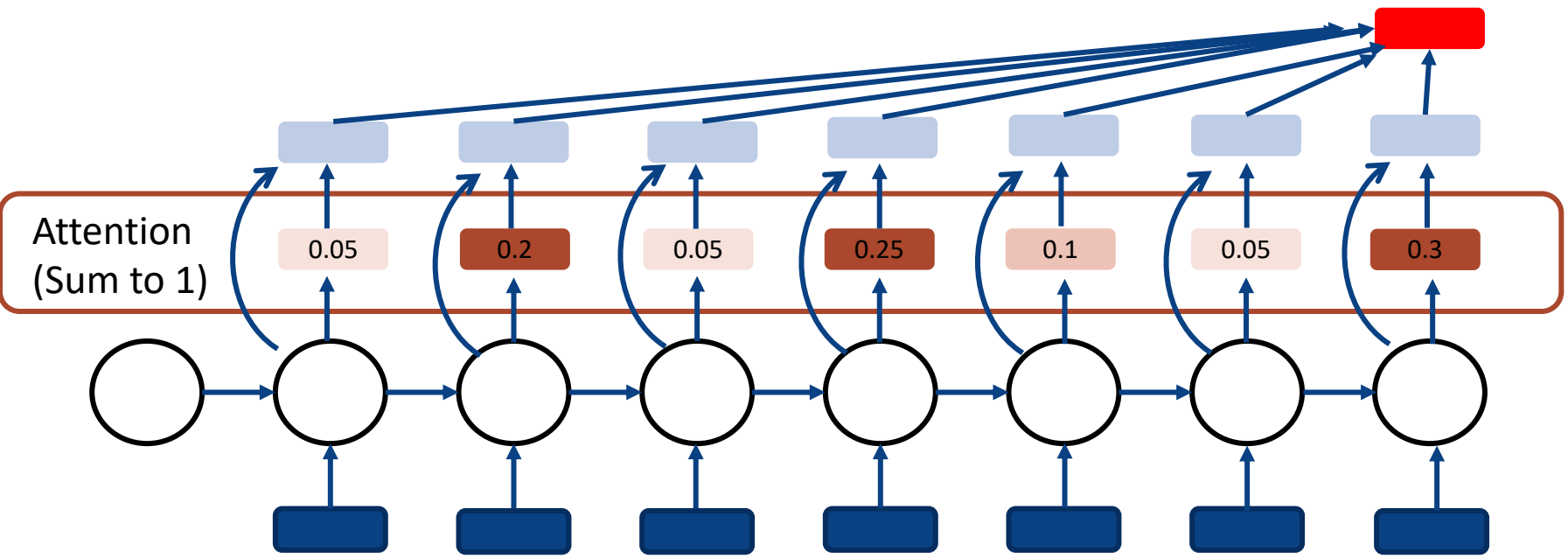
Bank[2]: sloping land (especially the slope beside a body of water)



RNN + Attention

Bank[1]:…a financial institution that accepts deposits and channels the money into lending activities

Bank[2]: sloping land (especially the slope beside a body of water)



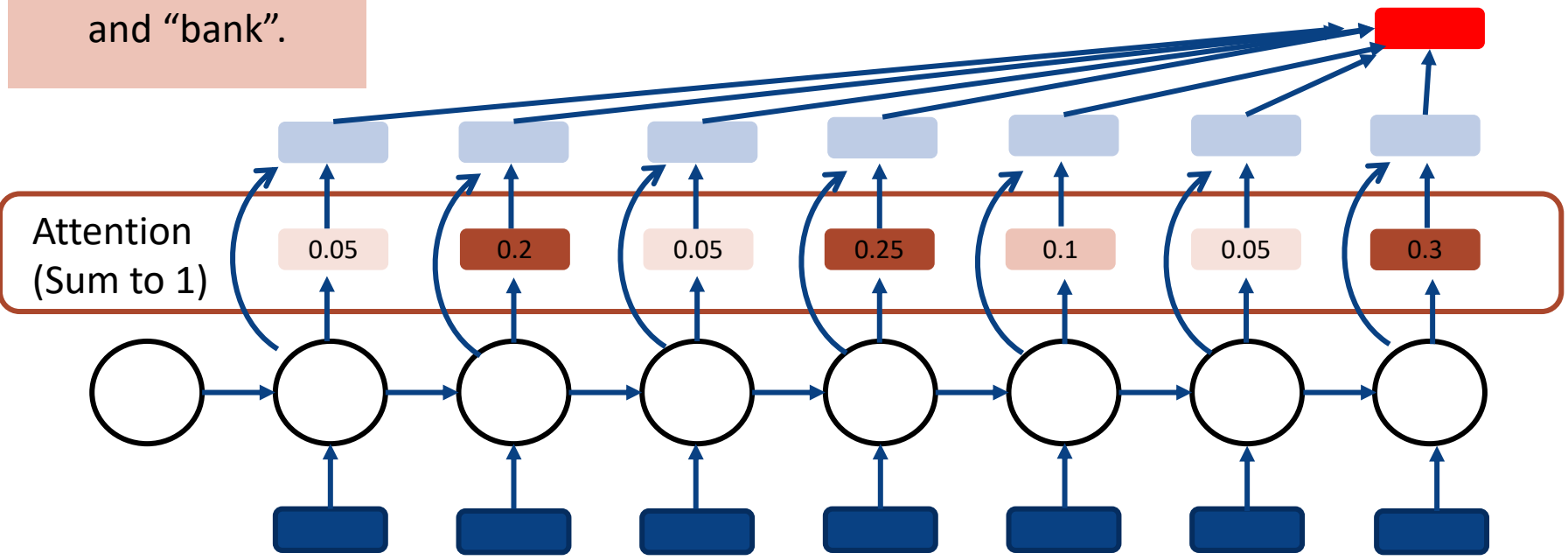
He cashed a check at the **bank**

RNN + Attention

To predict the meaning of **bank**, we need to pay more attention to "chased", "check" and "bank".

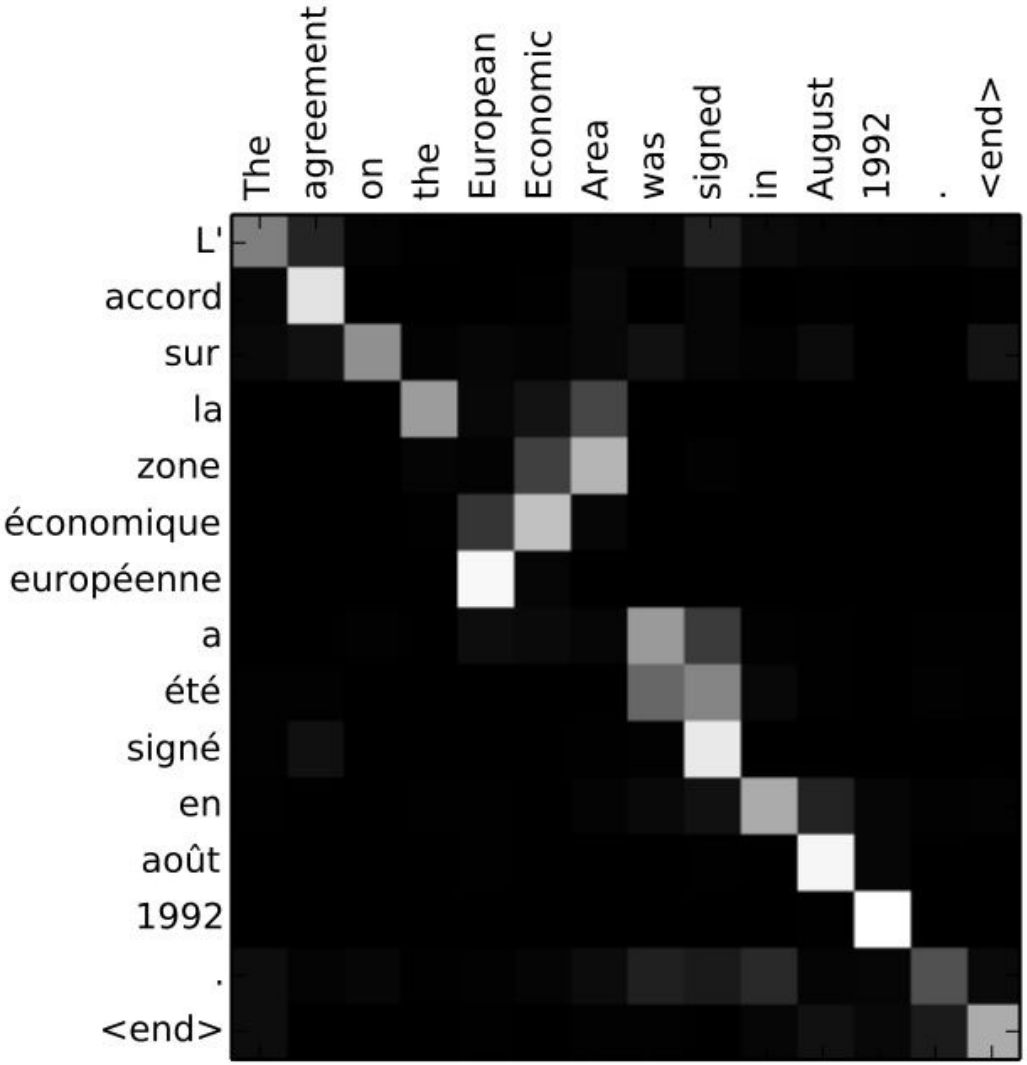
Bank[1]:...a financial institution that accepts deposits and channels the money into lending activities

Bank[2]: sloping land (especially the slope beside a body of water)



He cashed a check at the **bank**

Attention


The
agreement
on
the
European
Economic
Area
was
signed
in
August
1992
.
<end>
L'
accord
sur
la
zone
économique
européenne
a
été
signé
en
août
1992
.
<end>

Transformer

How does a transformer captures dependencies between words?

Transformer uses attention to capture dependencies between words.

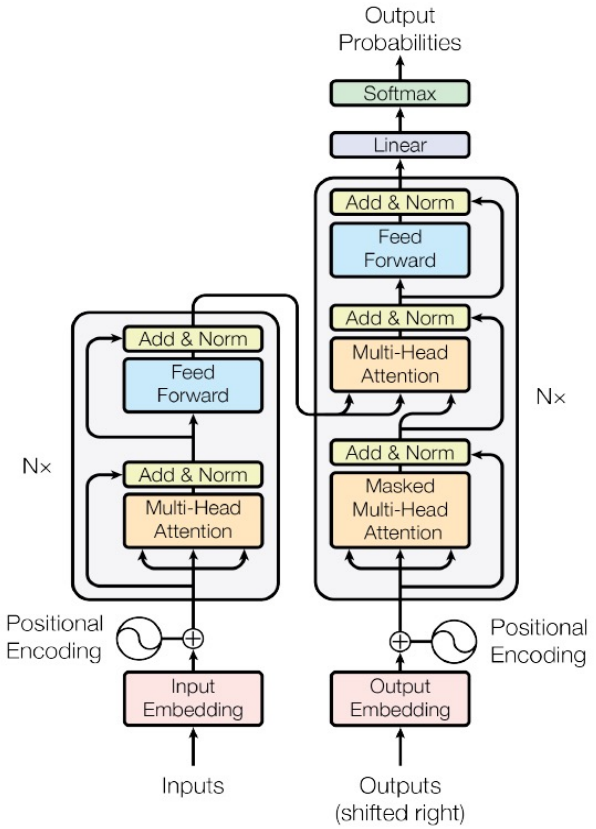


Figure 1: The Transformer - model architecture.

Figure from 'Attention is all your need' *by Vaswani et al.*

Self-Attention

$$\text{Attention}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{softmax}(\frac{\mathbf{Q}\mathbf{K}^\top}{\sqrt{n}})\mathbf{V}$$

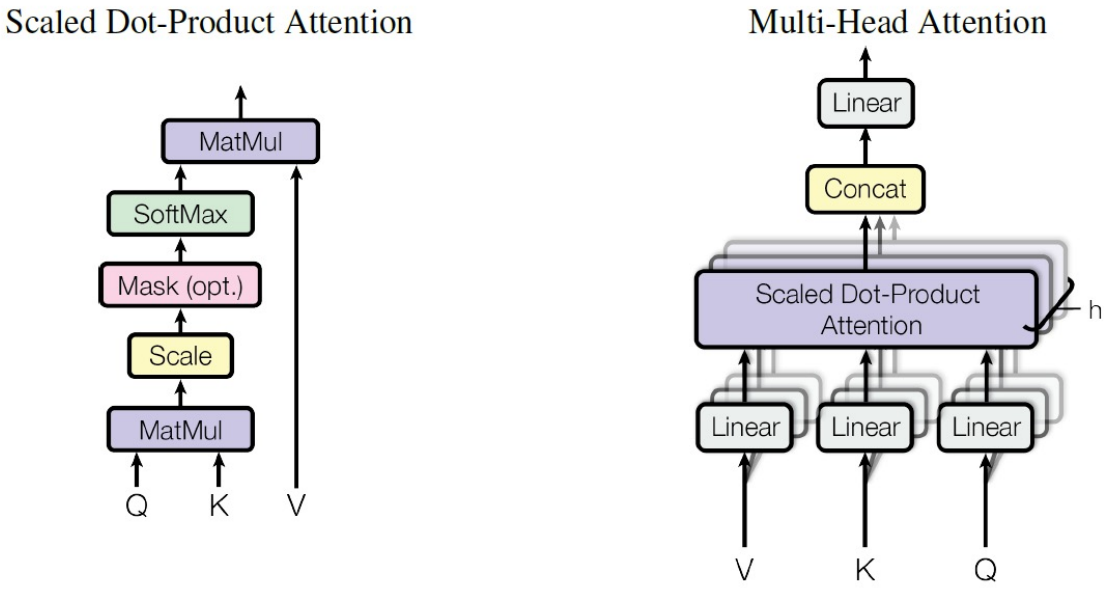
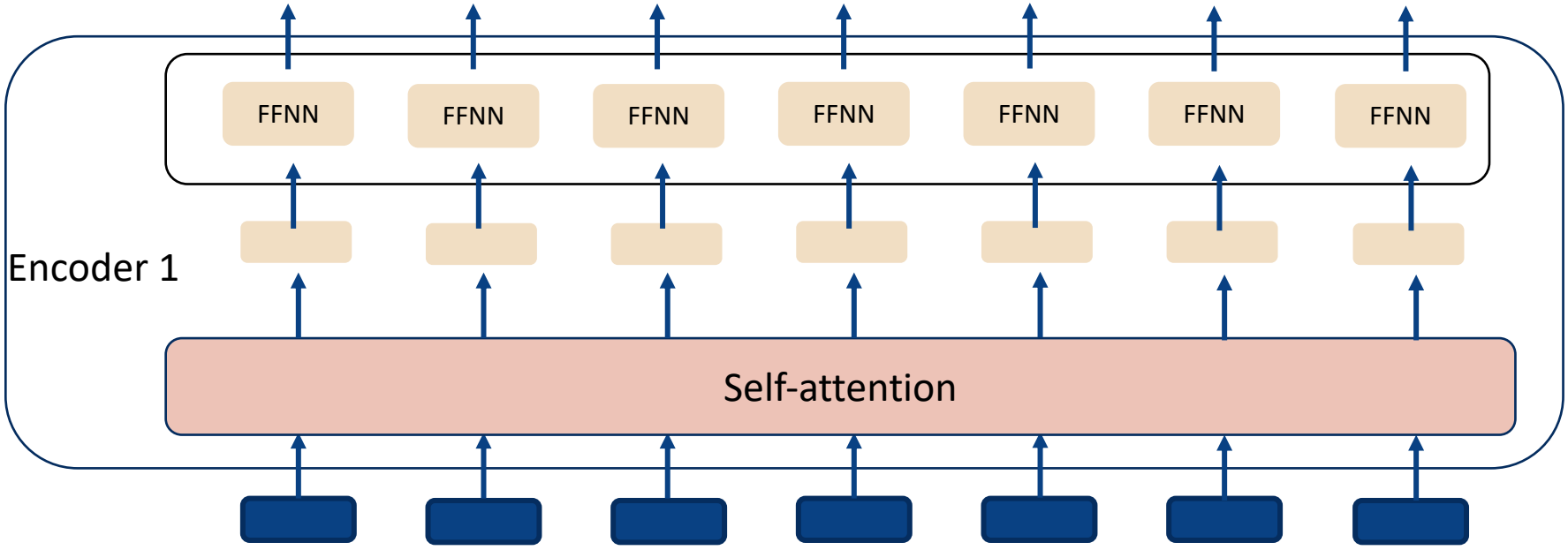


Figure 2: (left) Scaled Dot-Product Attention. (right) Multi-Head Attention consists of several attention layers running in parallel.

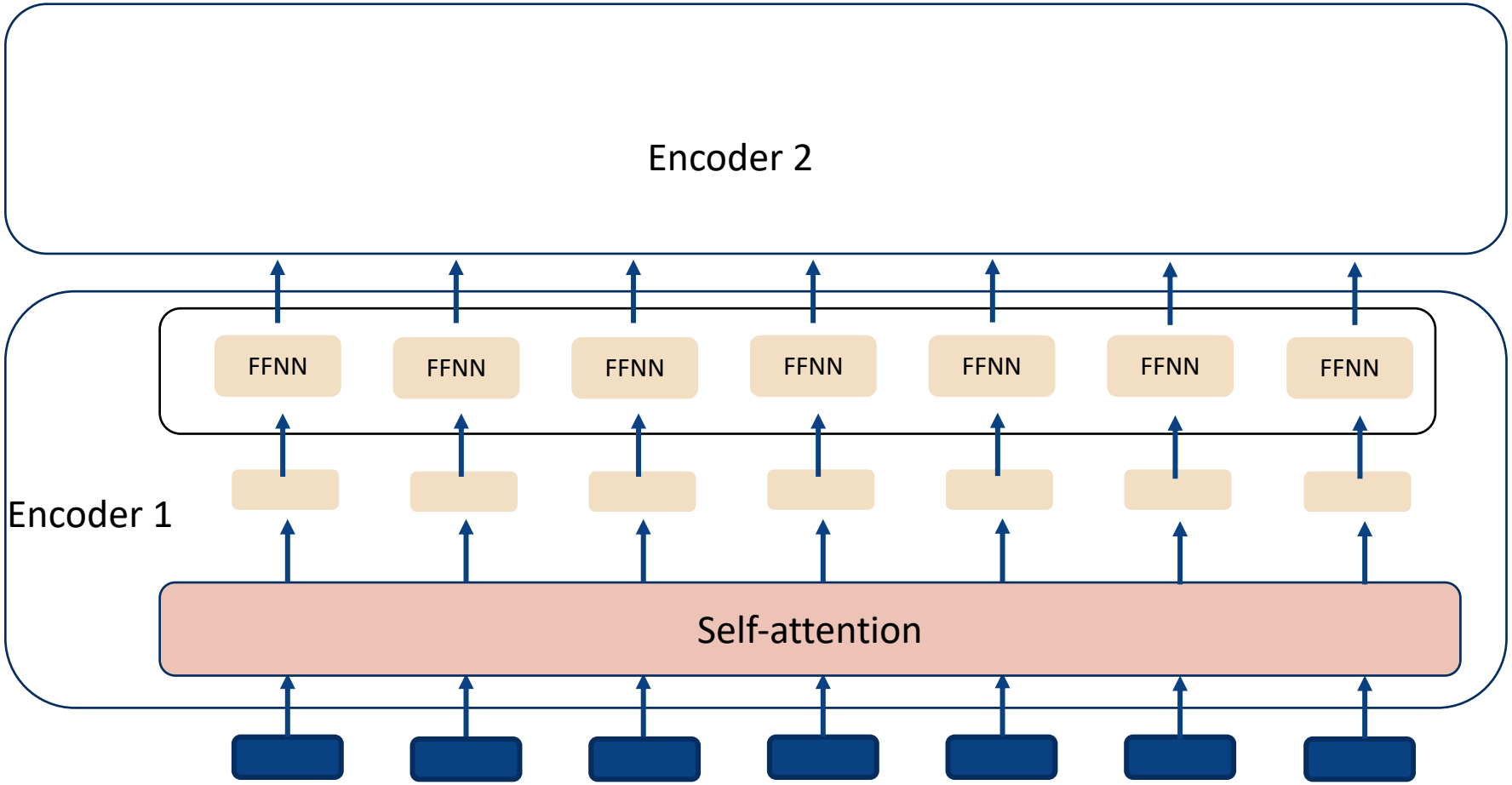
Figure from 'Attention is all your need' *by Vaswani et al.*

Transformer

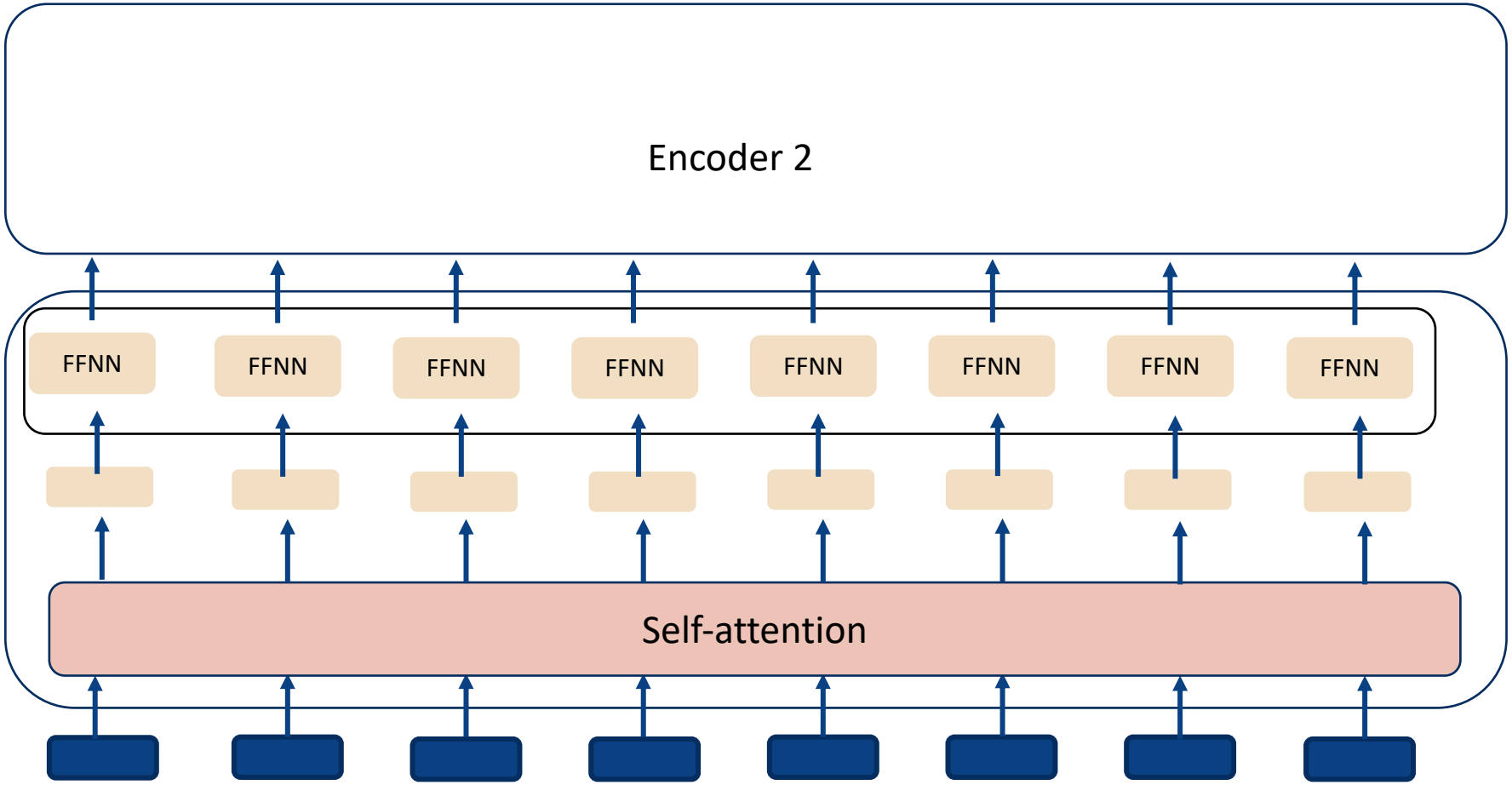


He cashed a check at the **bank**

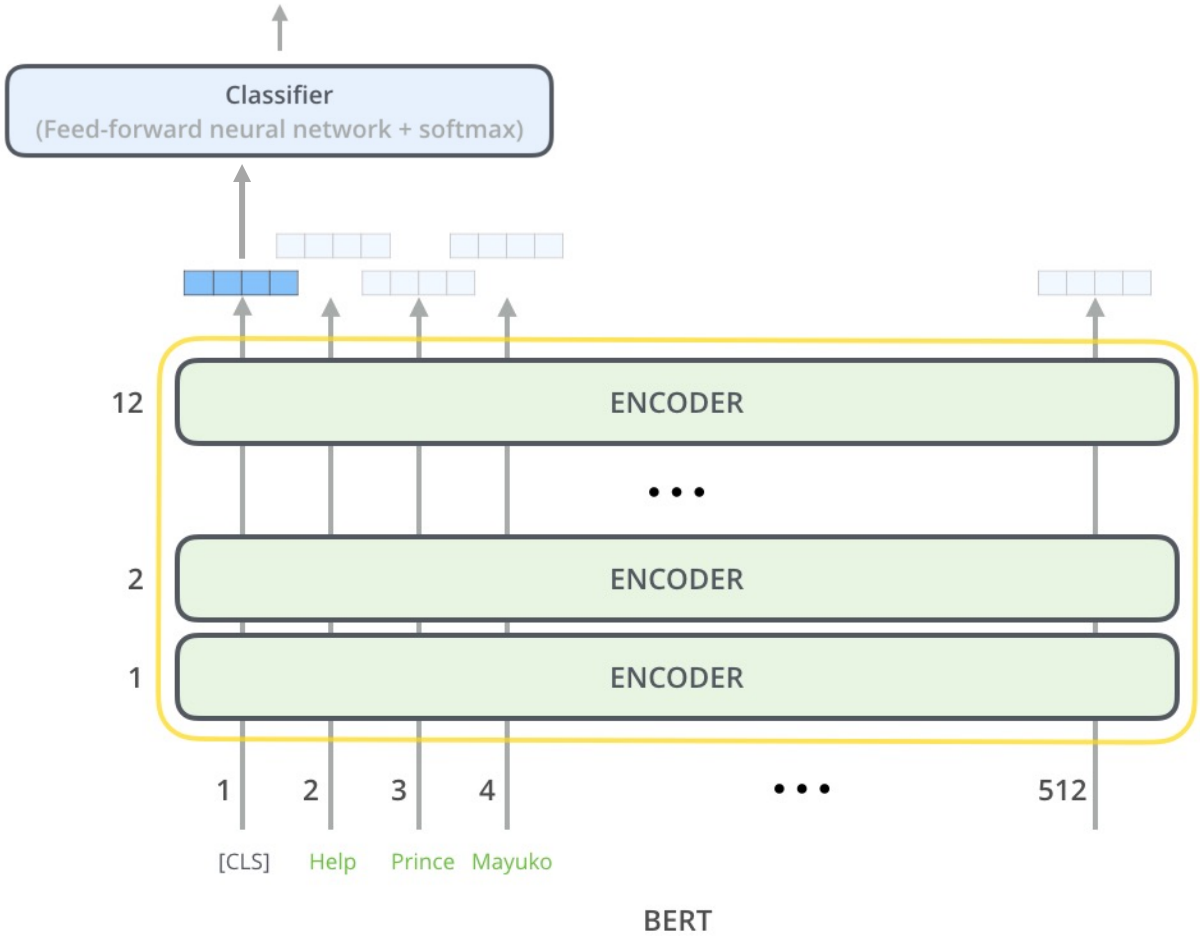
Transformer



Transformer



Bert


Classifier
(Feed-forward neural network + softmax)
12
ENCODER
2
ENCODER
1
ENCODER
1
2
3
4
512
[CLS]
Help
Prince
Mayuko
BERT

Transformer v.s. RNN

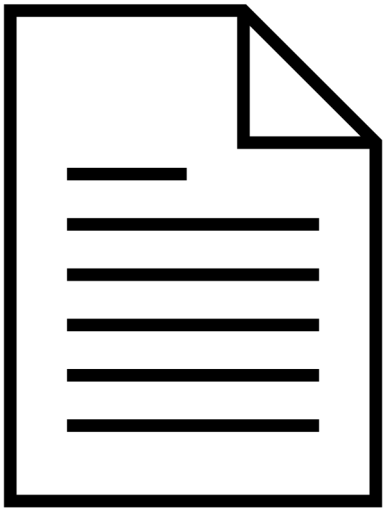
What advantages does transformer have compared to RNN?

- RNN: rely on sequential processing, current computation need to wait for previous computations has done.
- Transformer: the contextual embeddings of a target word is independent to other target words

- Significantly more parallelization, easy to scale to very large data

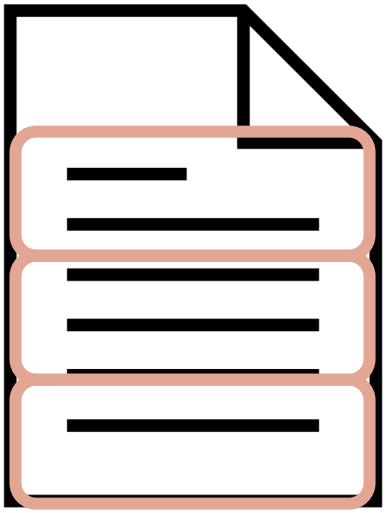
Discourse

What is discourse segmentation? What do the segments consist of, and what are some methods we can use to find them?



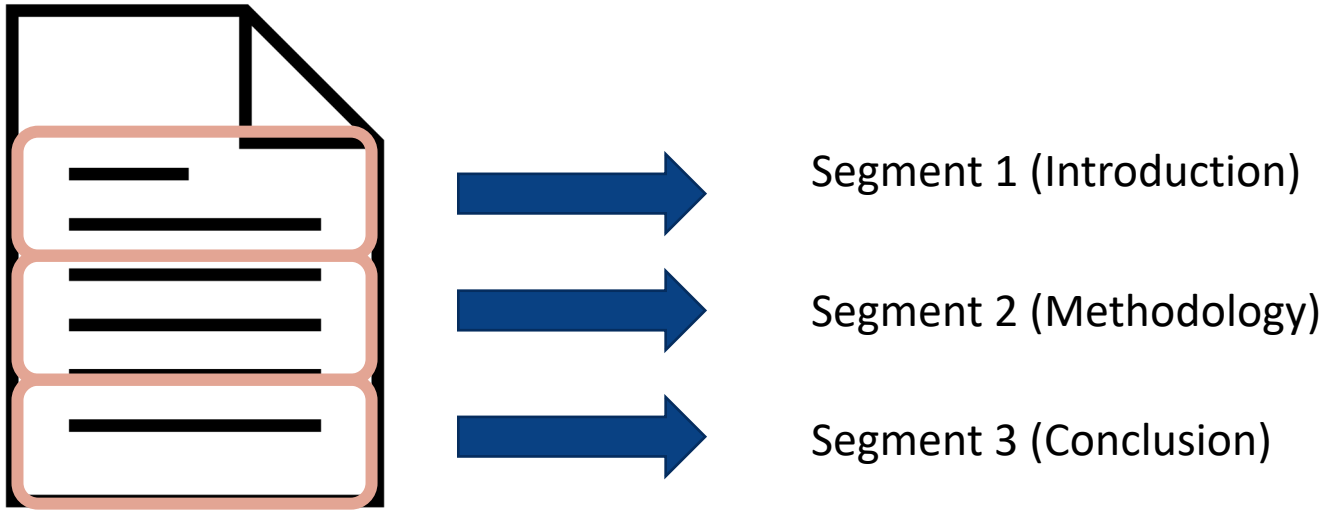
Discourse

What is discourse segmentation? What do the segments consist of, and what are some methods we can use to find them?



Discourse

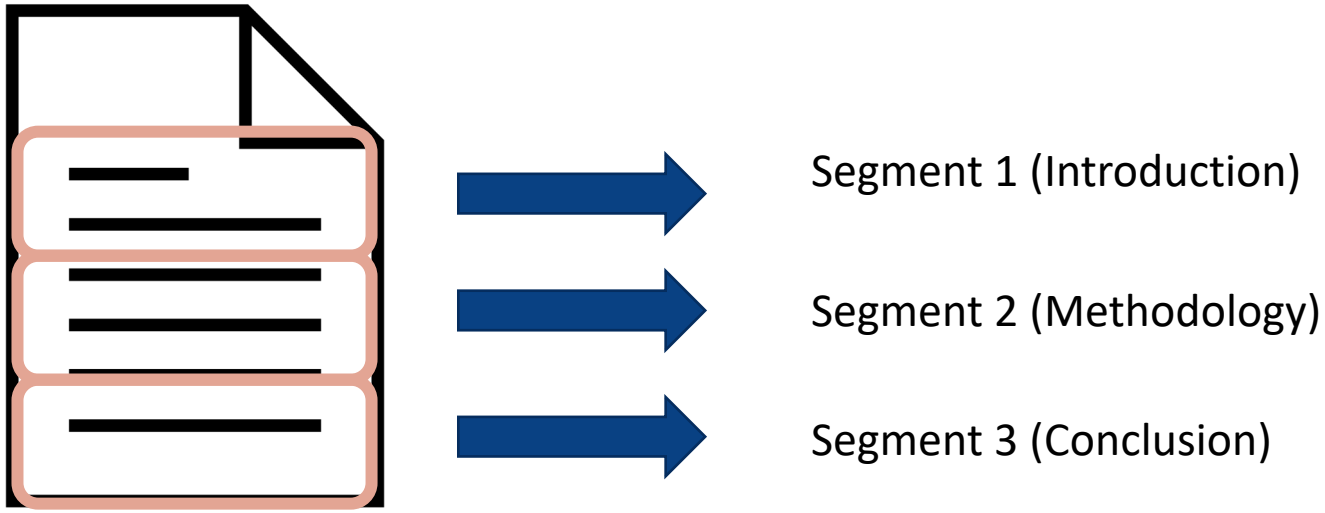
What is discourse segmentation? What do the segments consist of, and what are some methods we can use to find them?



Discourse

What is discourse segmentation? What do the segments consist of, and what are some methods we can use to find them?

divide up a text into discrete, cohesive units based on sentences



Discourse

What is discourse segmentation? What do the segments consist of, and what are some methods we can use to find them?

- Rule-based methods
- Unsupervised methods

find sentences with lexical overlap

Discourse

What is discourse segmentation? What do the segments consist of, and what are some methods we can use to find them?

- Rule-based methods
- Unsupervised methods

find sentences with lexical overlap

- Supervised methods

  (Lecture 12-page 10)

A classifier to predict paragraph boundaries

Anaphor

What is an anaphor?

Mary gave John a cat for his birthday. She is generous. He was surprised. He is fluffy.

Anaphor

What is an anaphor?

Mary gave John a cat for his birthday. She is generous. He was surprised. He is fluffy.

Anaphor

What is an anaphor?

Mary gave John a cat for his birthday. She is generous. He was surprised. He is fluffy.

An anaphor is a linguistic expression that refers back to one or more elements in the text.

Anaphor

What is anaphora resolution and why is it difficult?

Mary gave John a cat for his birthday. She is generous. He was surprised. He is fluffy.

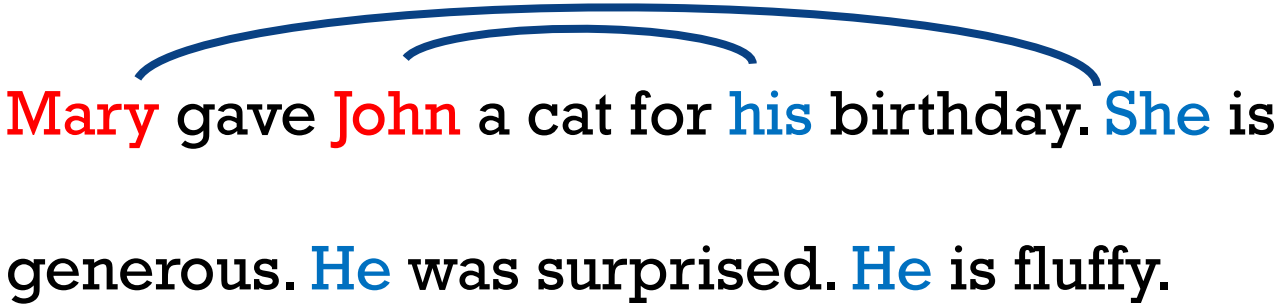
Anaphor

What is anaphora resolution and why is it difficult?

Mary gave John a cat for his birthday. She is generous. He was surprised. He is fluffy.

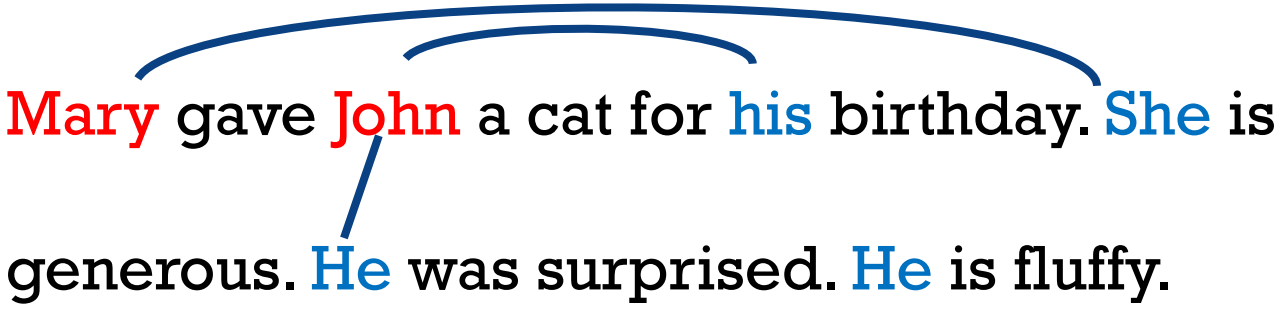
Anaphor

What is anaphora resolution and why is it difficult?



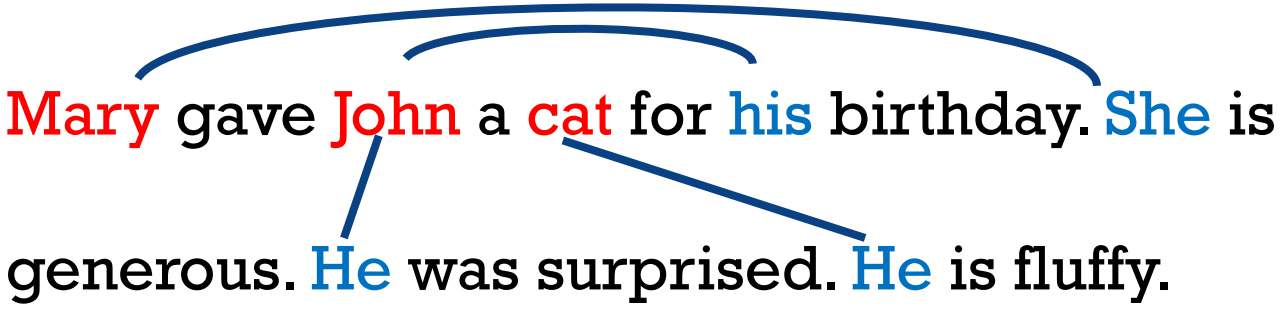
Anaphor

What is anaphora resolution and why is it difficult?



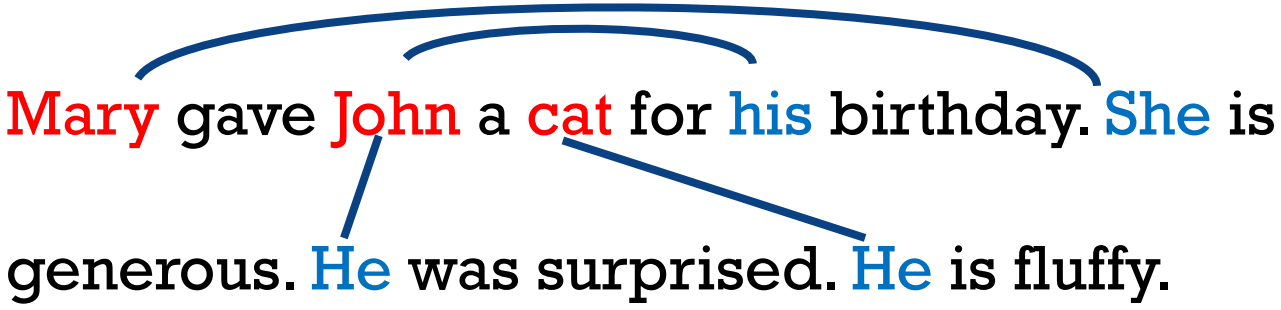
Anaphor

What is anaphora resolution and why is it difficult?



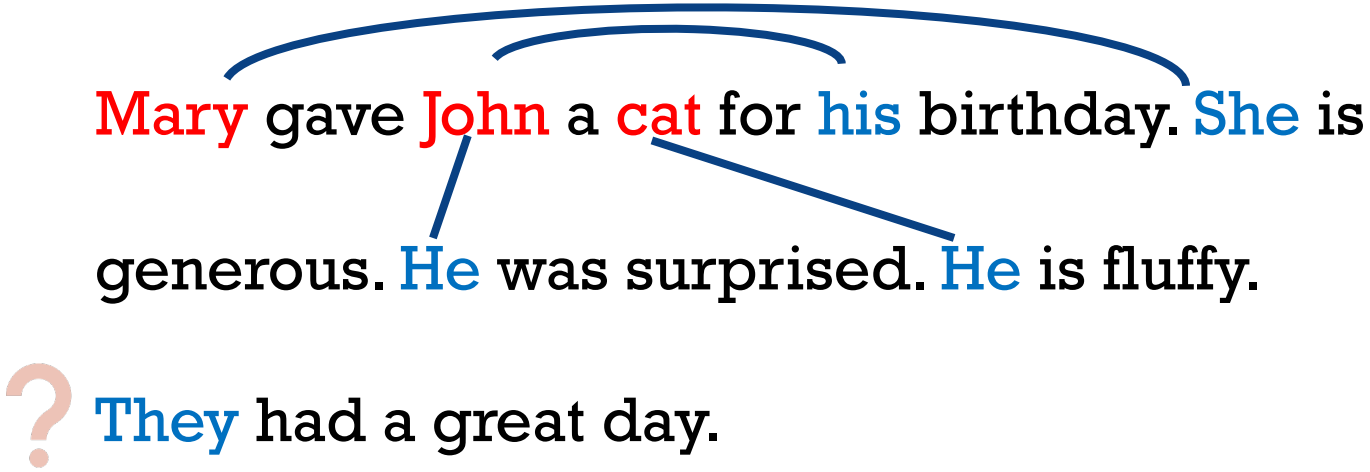
Anaphor

What is anaphora resolution and why is it difficult?



Anaphor

What is anaphora resolution and why is it difficult?



Anaphor

What are some useful heuristics (or features) to help resolve anaphora?

- Recency heuristic:

Bob enjoyed playing football with John,

he had a great day.

Anaphor

What are some useful heuristics (or features) to help resolve anaphora?

- Recency heuristic:

Anaphor

What are some useful heuristics (or features) to help resolve anaphora?

- Refer to “center”:

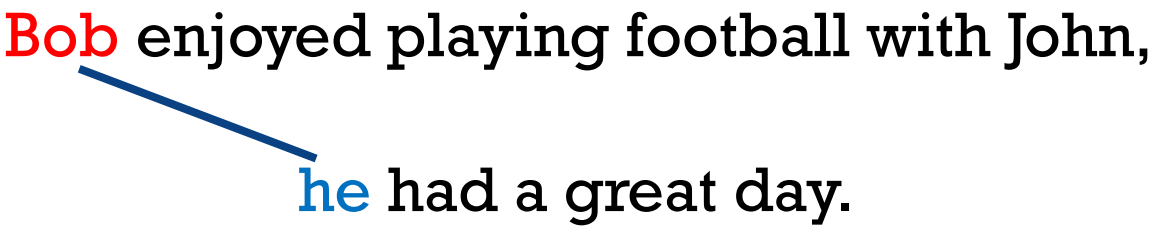
Bob enjoyed playing football with John,

he had a great day.

Anaphor

What are some useful heuristics (or features) to help resolve anaphora?

- Refer to “center”:



Programming -- BERT

Steps:

1. go to: https://colab.research.google.com/

2. Sign up or login to a Google account.

3. File > Upload Notebook