



# 句子的表示

CS2916 大语言模型

飲水思源 愛國榮校

<https://plms.ai/teaching/index.html>



# 什么是句子的表示?

一个句子  
(a sequence of  
tokens)



合成函数



0.1  
0.2  
0.5  
0.1  
0.5



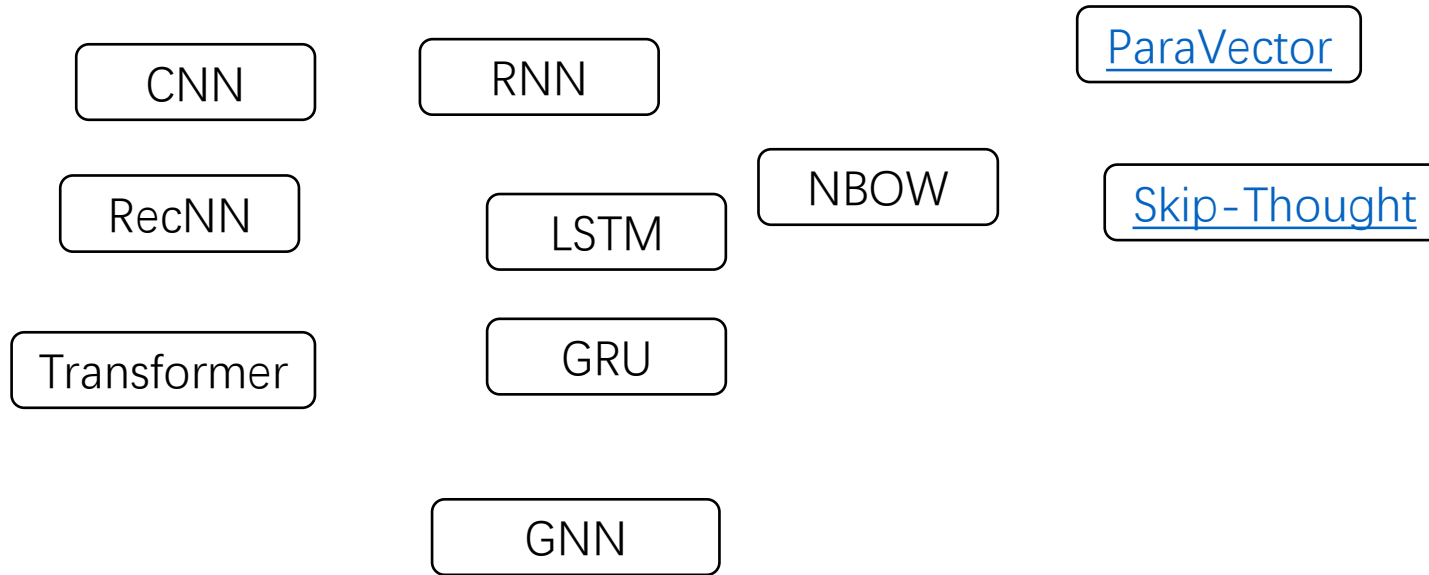
# 为什么我们需要句子的表示?

---

- 非常基础的一步：
  - 情感分类
  - 语义匹配
  - 文本摘要
  - 机器翻译



# 怎么学习句子的表示?





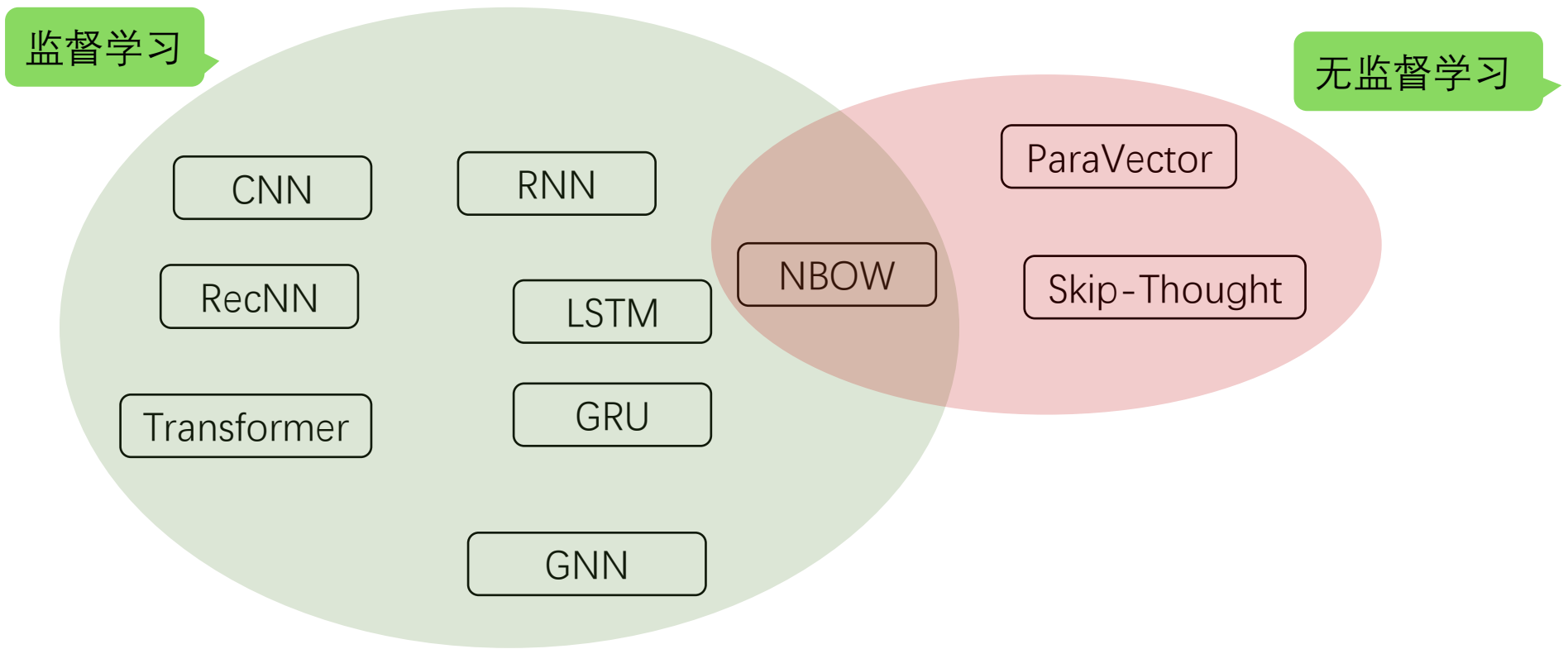
# 怎么学习句子的表示?

---

- 监督学习
- 无监督学习



# 怎么学习句子的表示?





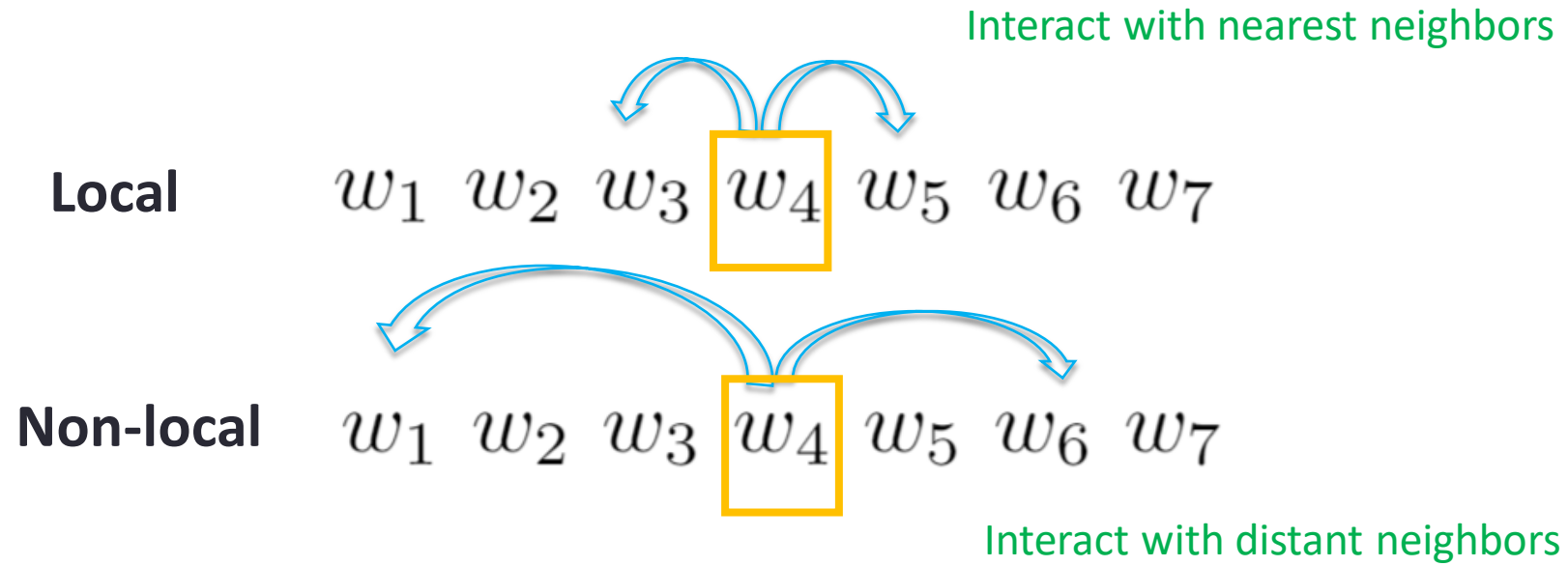
# 句子表示中的 “Structural Bias”

---

- Structural Bias: 一组整合到模型设计中的先验知识
  - 网络的连接方式
  - 网络的拓扑结构



# 网络的连接方式





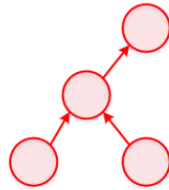


# 网络的拓扑结构

**Sequential**

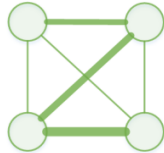


**Tree**



(syntactic parsing tree)

**Graph**



(entity relation)



# 统一的角度理解不同句子表示模型

连接方式

Local

Non-local

拓扑结构

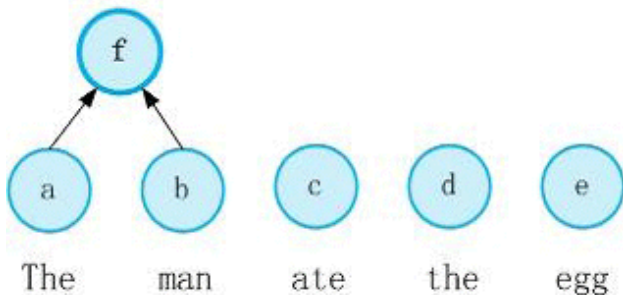
Seq.

Tree

Graph



The



RNN

LSTM

GRU

CNN



# 统一的角度理解不同句子表示模型

连接方式

Local

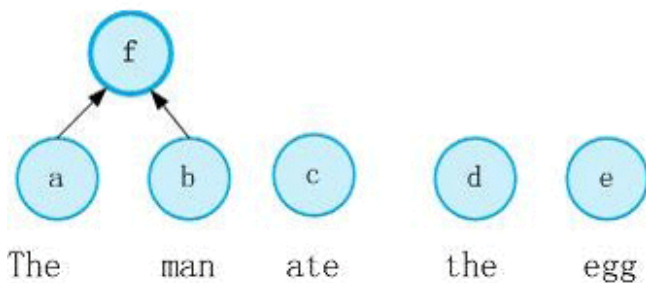
Non-local

拓扑结构

Seq.

Tree

Graph



Recursive NN

TreeLSTM

TreeCNN



# 统一的角度理解不同句子表示模型

连接方式

Local

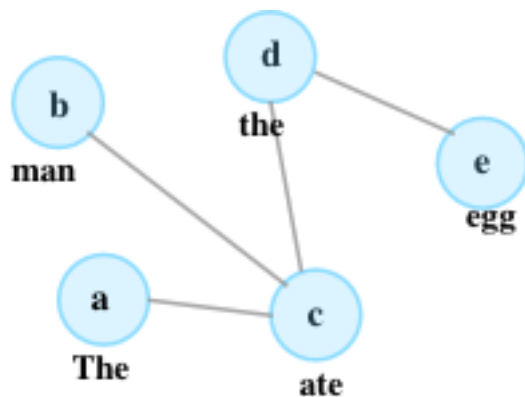
Non-local

拓扑结构

Seq.

Tree

Graph



Graph Neural Nets



# 统一的角度理解不同句子表示模型

连接方式

Local

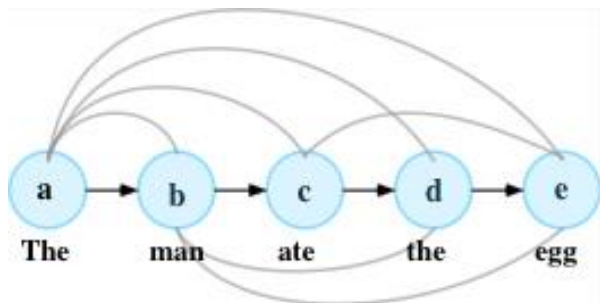
Non-local

拓扑结构

Seq.

Tree

Graph



Attention LSTM



# 统一的角度理解不同句子表示模型

连接方式

Local

Non-local

拓扑结构

Seq.

Tree

Graph

?



# 统一的角度理解不同句子表示模型

连接方式

Local

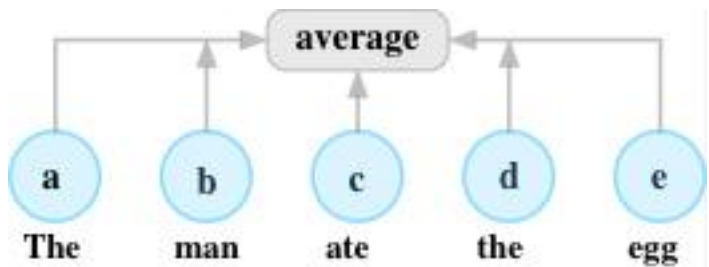
Non-local

拓扑结构

Seq.

Tree

Graph



NBOW



# 统一的角度理解不同句子表示模型

连接方式

Local

Non-local

拓扑结构

Seq.

Tree

Graph

?





# 统一的角度理解不同句子表示模型

连接方式

Local

Non-local

拓扑结构

Seq.

Tree

Graph

# Transformer