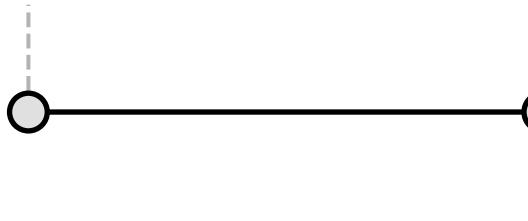


# Evolution of Sequence Modeling: From N-grams to Transformers

• Breakthrough

## N-grams

*Fixed context  
Count-based*



*Phrase-based  
Alignment models*

## RNNs

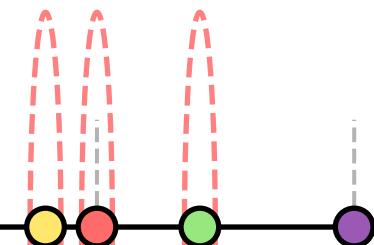
*Sequential  
Fixed output*



## Statistical MT

## Attention

*Dynamic context  
Bahdanau et al.*



*Variable I/O  
Encoder-Decoder  
Vaswani et al.*

## Seq2SeqTransformers

## GPT/BERT

*Pre-trained  
Transfer learning*

