



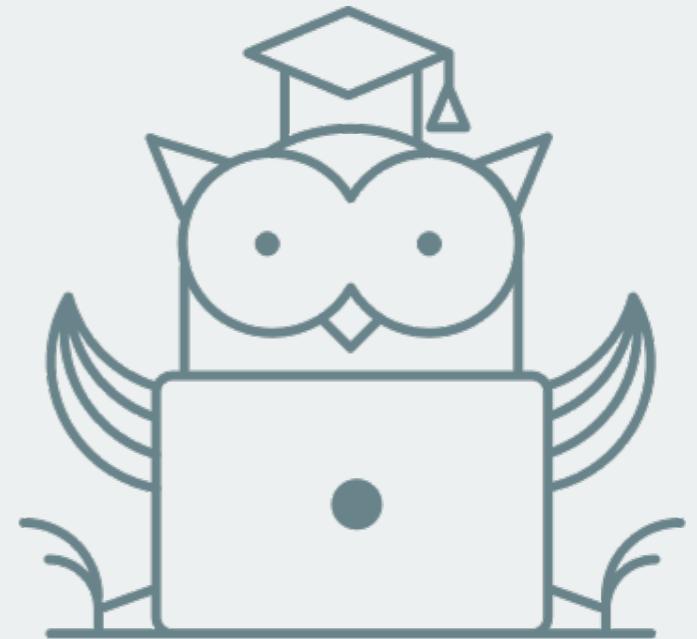
ОНЛАЙН-ОБРАЗОВАНИЕ

# Внимание

Порой во что-то нужно всмотреться, чтобы понять ценность.

Т/с «Шерлок»

Артур Кадурин  
Insilico Medicine

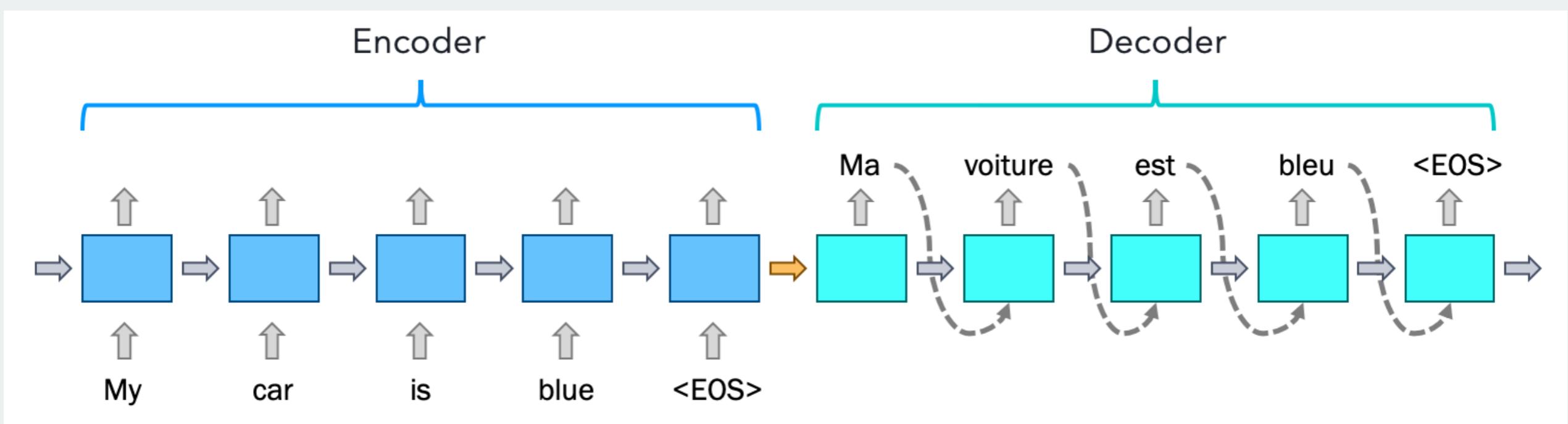


1. Внимание
2. Самолюбование
3. Титры
4. Сцена после титров

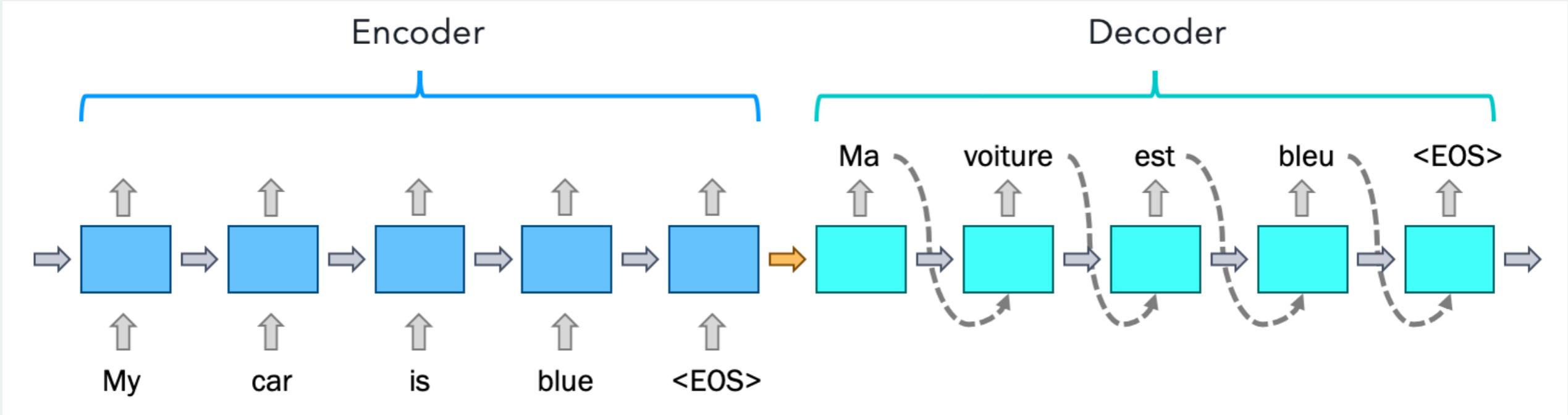


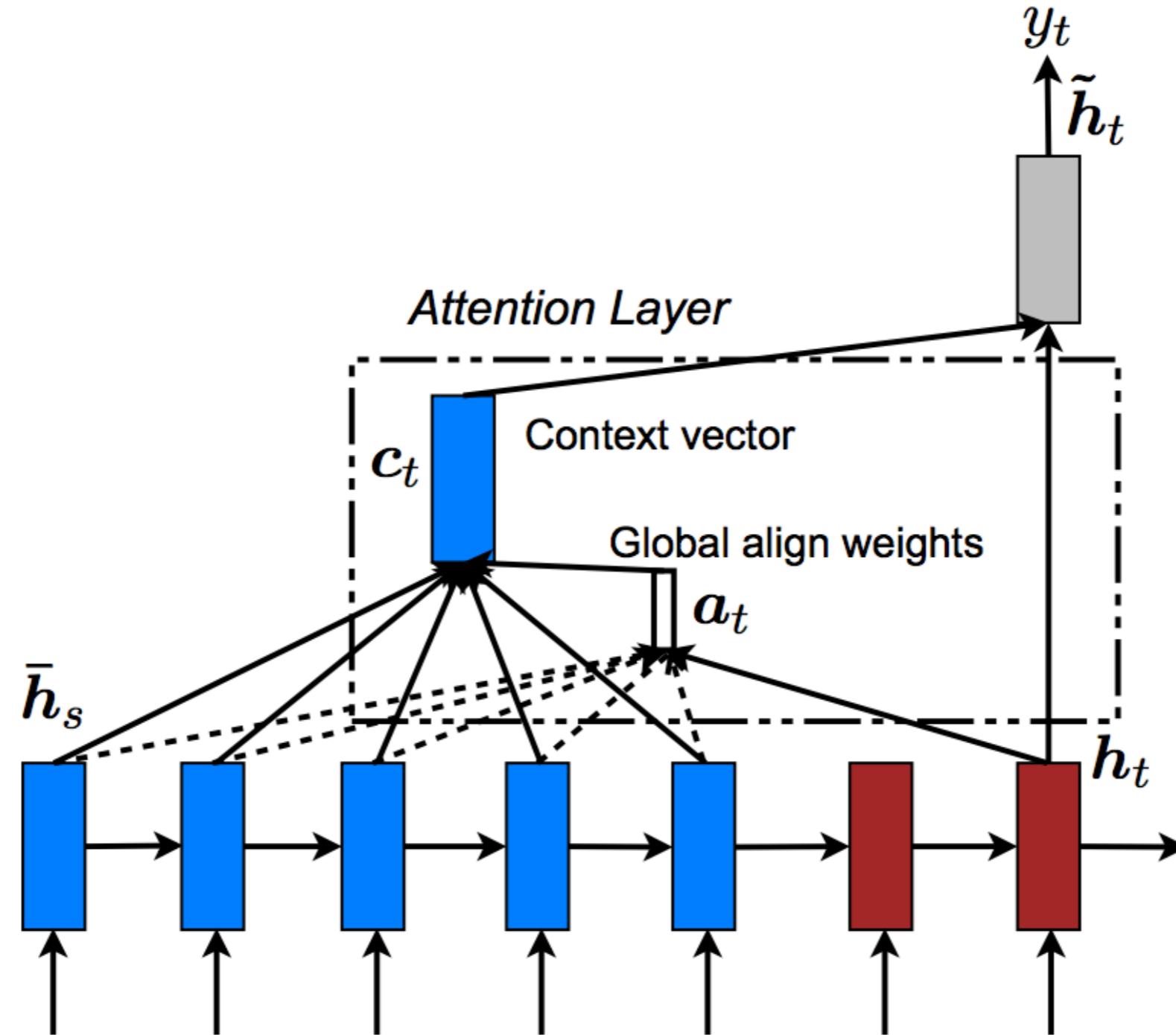
Идея внимания в нейронных сетях, как это часто бывает, пришла из нейробиологии. Известно, что человек в каждый момент времени фокусируется на какой-то небольшой части поступающих сигналов и практически игнорирует все остальное.

В нейронных сетях механизм внимания был предложен в публикации посвященной машинному переводу.

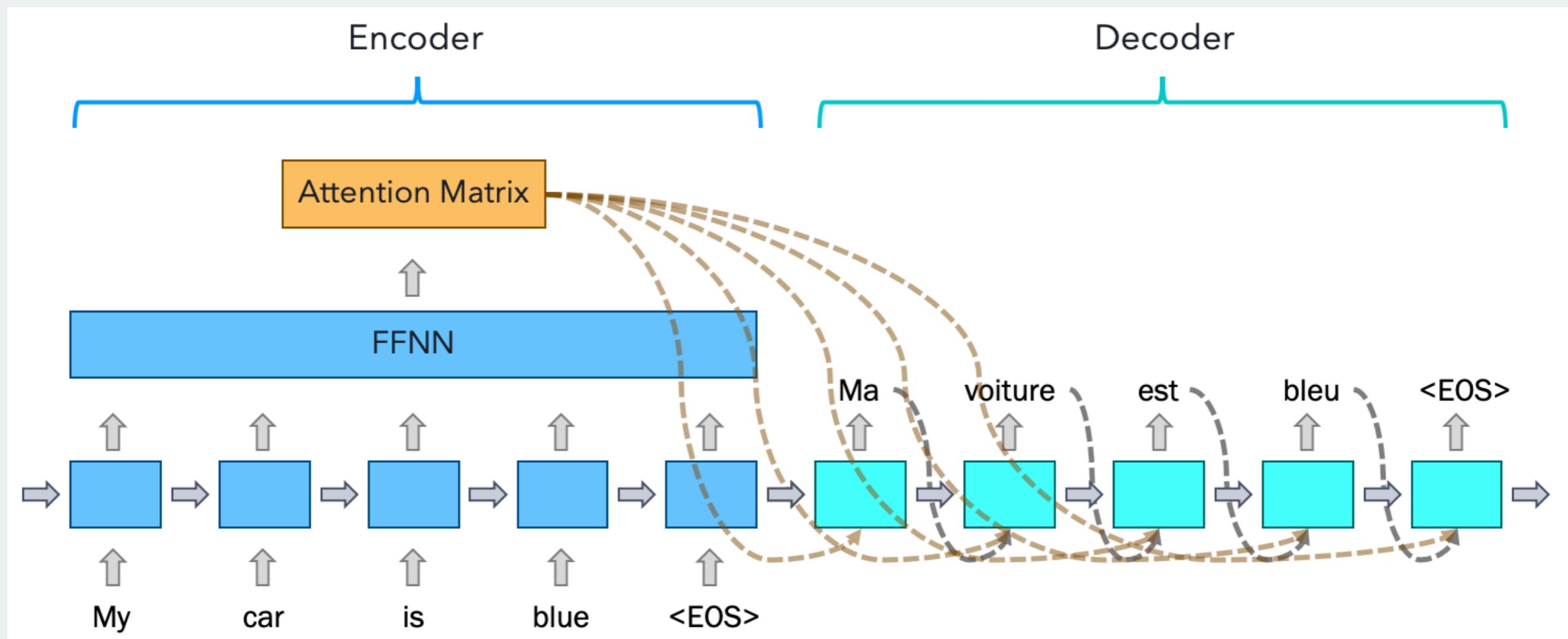


В seq2seq моделях вся информация из энкодера в декодер передается через единственный вектор посередине — глобальное представление входной последовательности. Но при декодировании хочется «смотреть» на конкретные детали.





В seq2seq моделях вся информация из энкодера в декодер передается через единственный вектор посередине — глобальное представление входной последовательности.



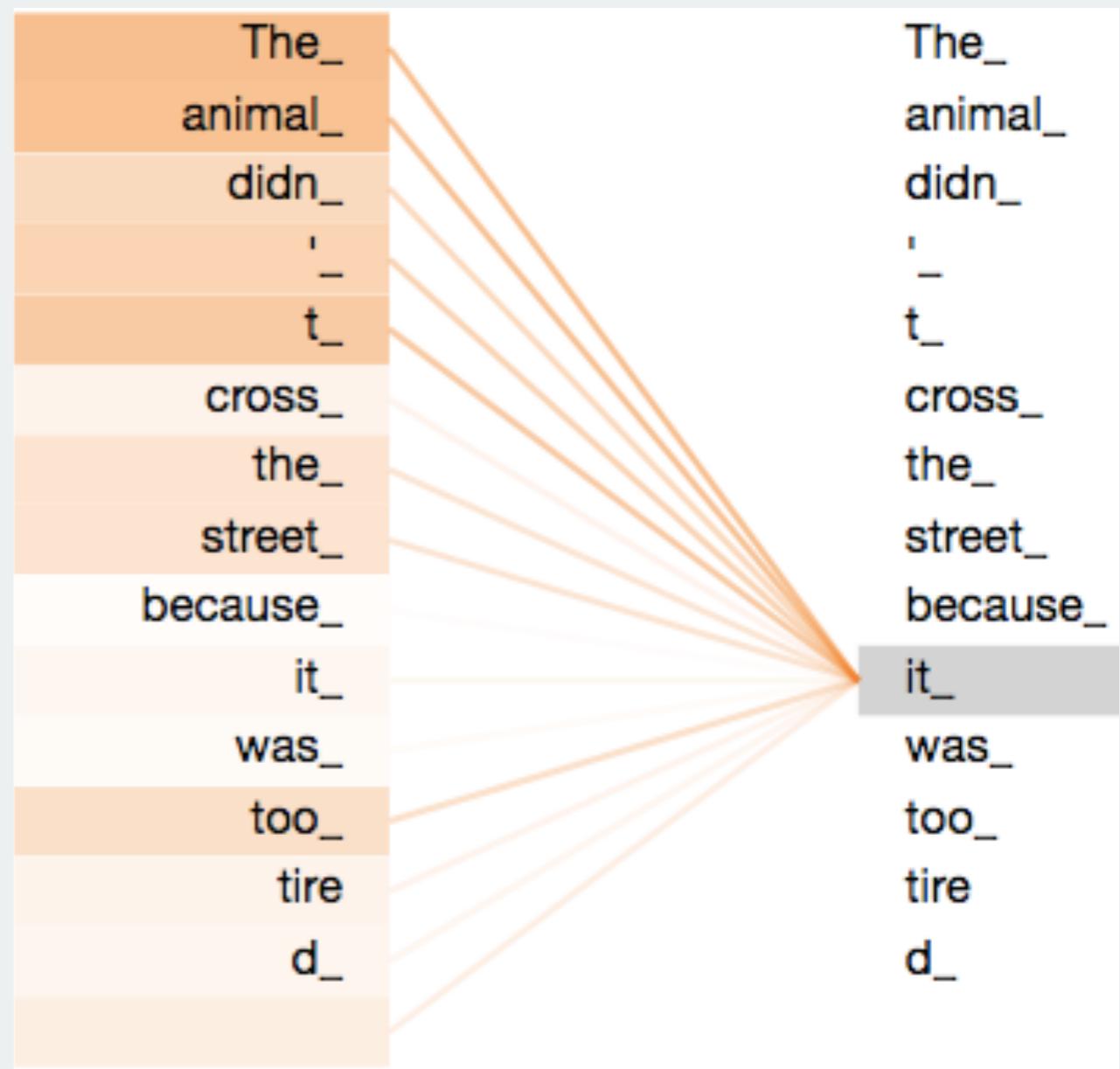
1. Внимание
2. Самолюбование
3. Титры
4. Сцена после титров



# Self-attention



Внимание может быть направлено «внутрь» себя. В моделях с самовниманием при вычислении очередного эмбеддинга для конкретного элемента последовательности учитывается его «похожесть» на остальные элементы.



# Self-attention



Для того чтобы задать  
самовнимание вводят 3  
дополнительных набора  
параметров:

$W^Q$  — Query,

$W^K$  — Key,

$W^V$  — Value,

где  $X$  — это последовательность  
выходов рекуррентной ячейки.

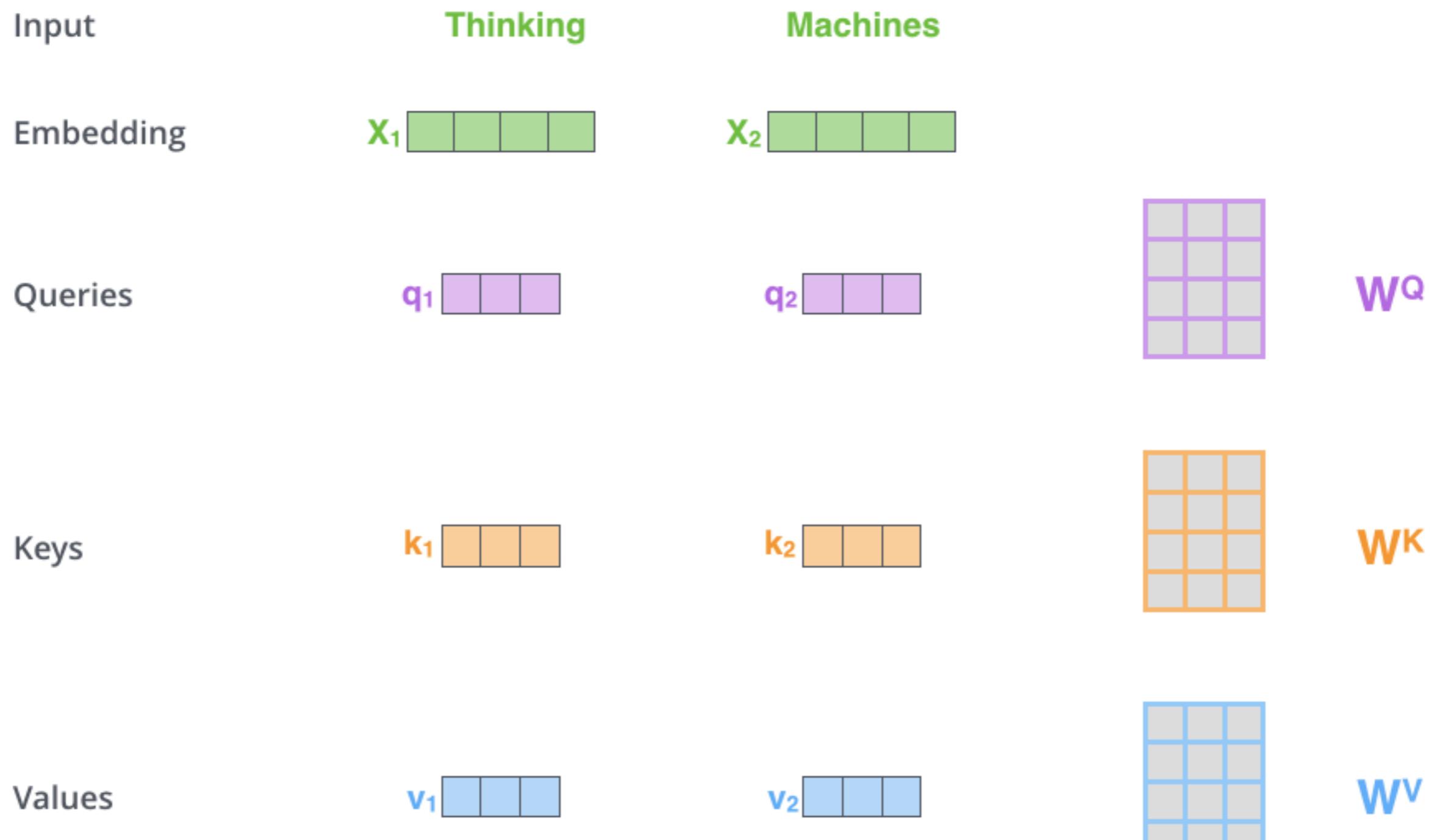
$$X \times W^Q = Q$$

$$X \times W^K = K$$

$$X \times W^V = V$$



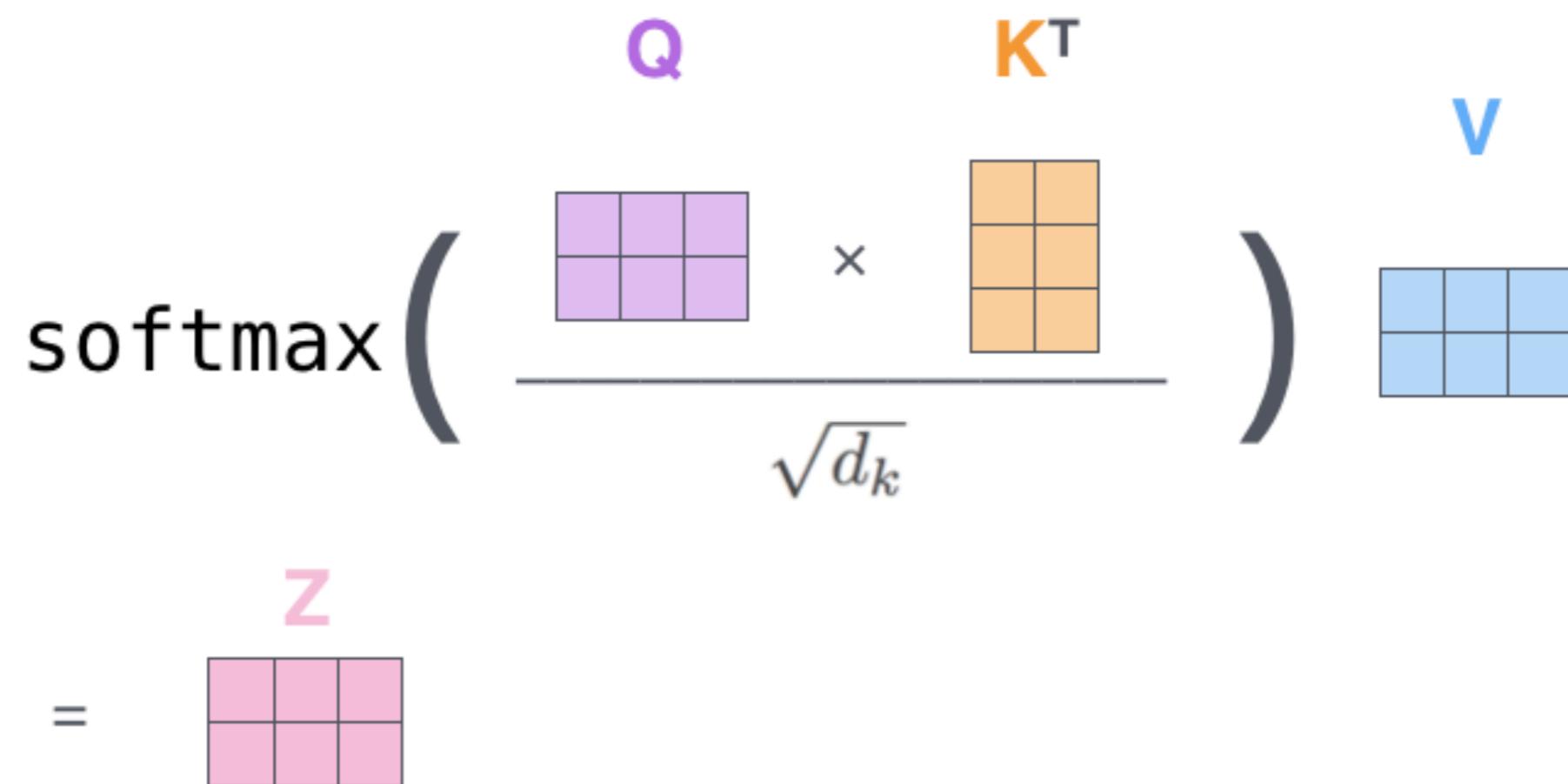
# Self-attention



# Self-attention

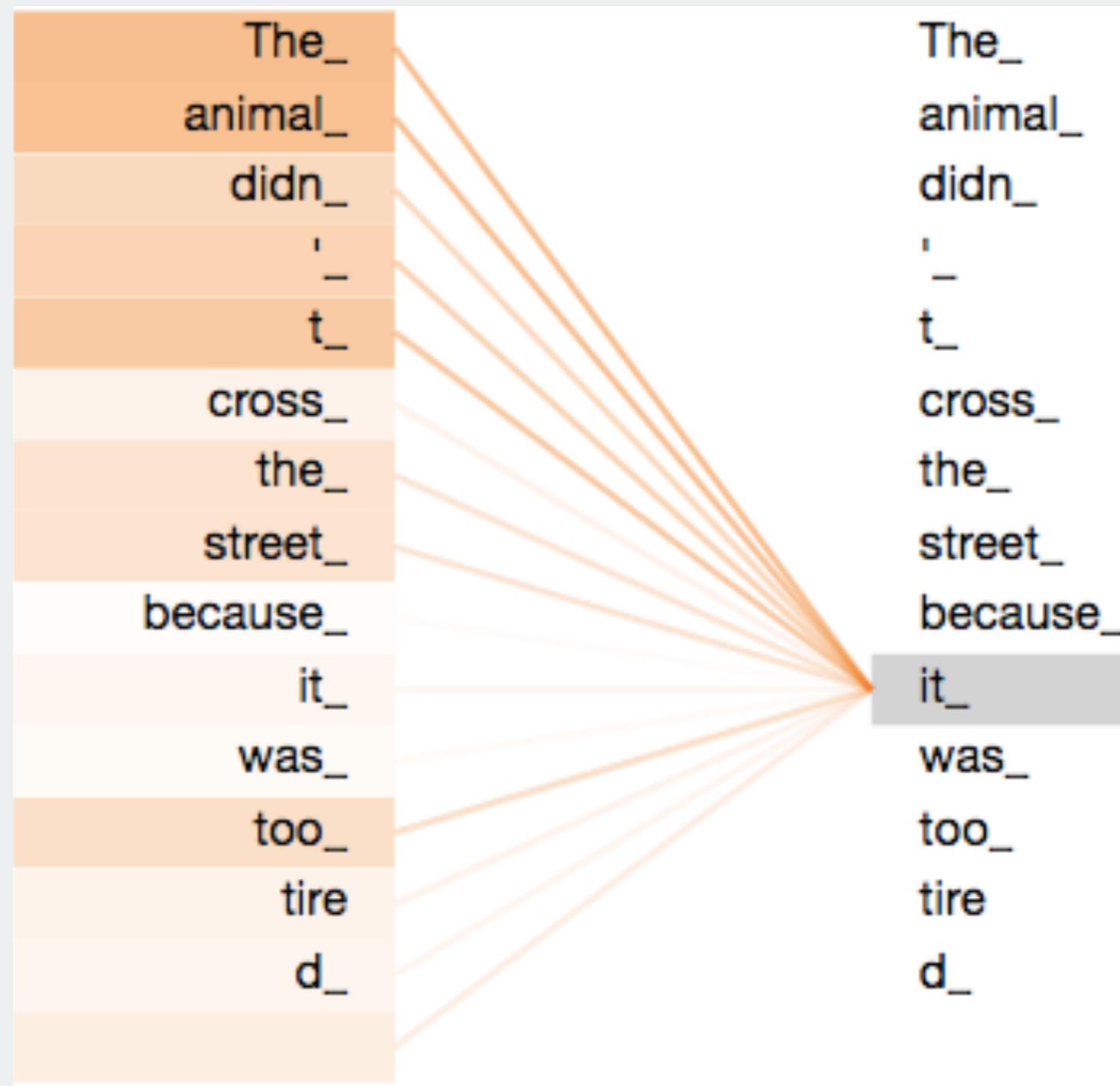


Вектор запроса( $Q$ ) для каждого из  $X$  умножаются на все векторы ключей( $K$ ). В результате этого получается матрица размером  $d_k \times d_k$ , где  $d_k$  — это длина  $K$



# Self-attention

O T U S



1. Внимание
2. Самолюбование
- 3. Титры**
4. Сцена после титров



# Captioning



В задачах описания изображений механизм внимания используется наиболее естественным образом. В процессе генерации текста, рекуррентная сеть «сматривает» на соответствующие части изображения.



The man at bat readies to swing at the pitch while the umpire looks on.



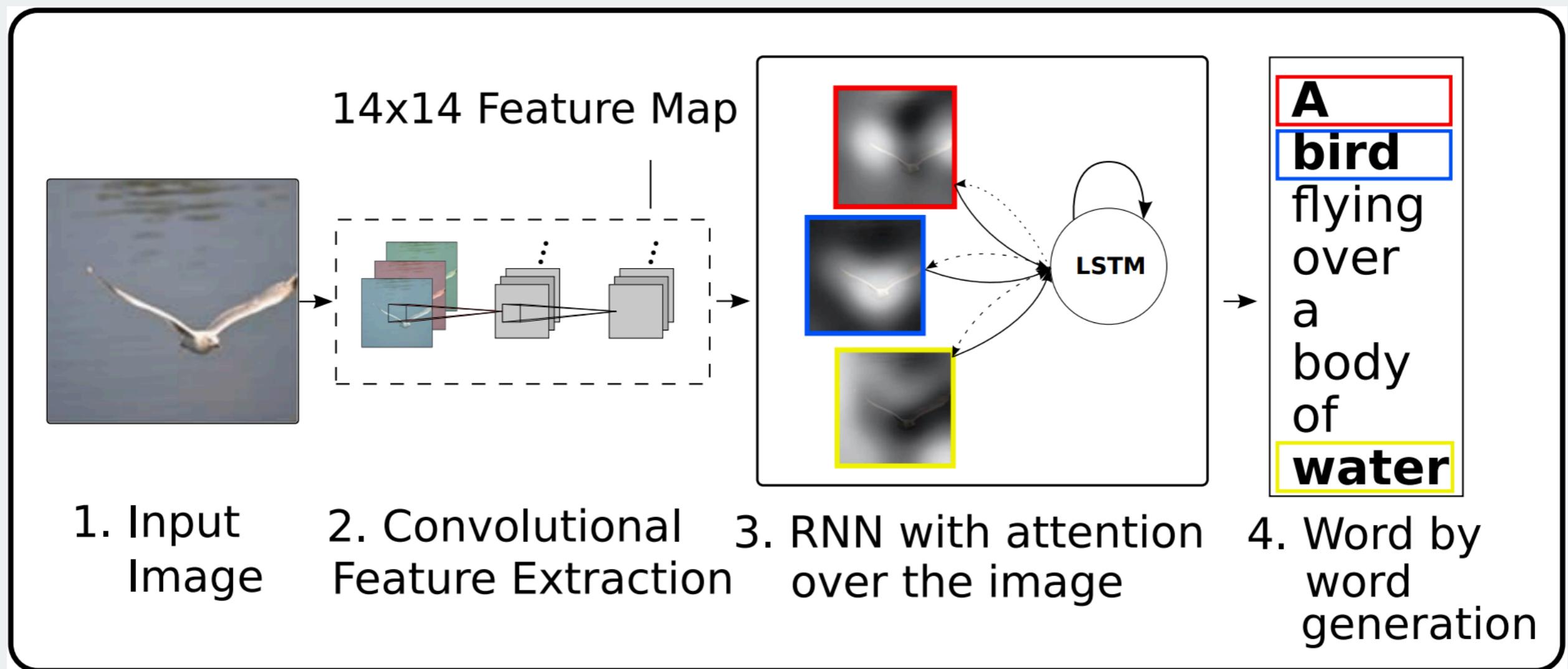
A large bus sitting next to a very tall building.



# Captioning



Каждый из «пикселей» выходного слоя сверточной сети соответствует региону исходного изображения. Если их взвесить вниманием и просуммировать, то LSTM сфокусируется на важном.



# Captioning



В статье “Show, Attend and Tell” 2015г. продемонстрированы SOTA результаты на тот момент.



A woman is throwing a frisbee in a park.



A dog is standing on a hardwood floor.



A stop sign is on a road with a mountain in the background.



A little girl sitting on a bed with a teddy bear.



A group of people sitting on a boat in the water.



A giraffe standing in a forest with trees in the background.



# Captioning

O T U S

Но без файлов не обошлось.



A large white bird standing in a forest.



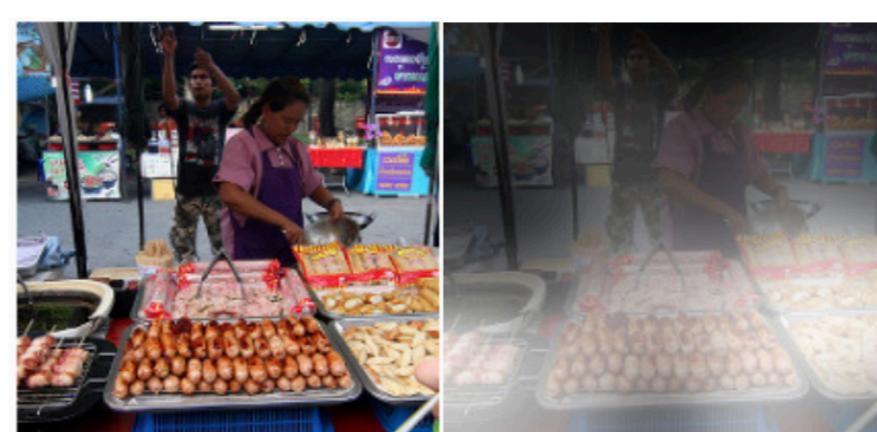
A woman holding a clock in her hand.



A man wearing a hat and  
a hat on a skateboard.



A person is standing on a beach  
with a surfboard.



A woman is sitting at a table  
with a large pizza.



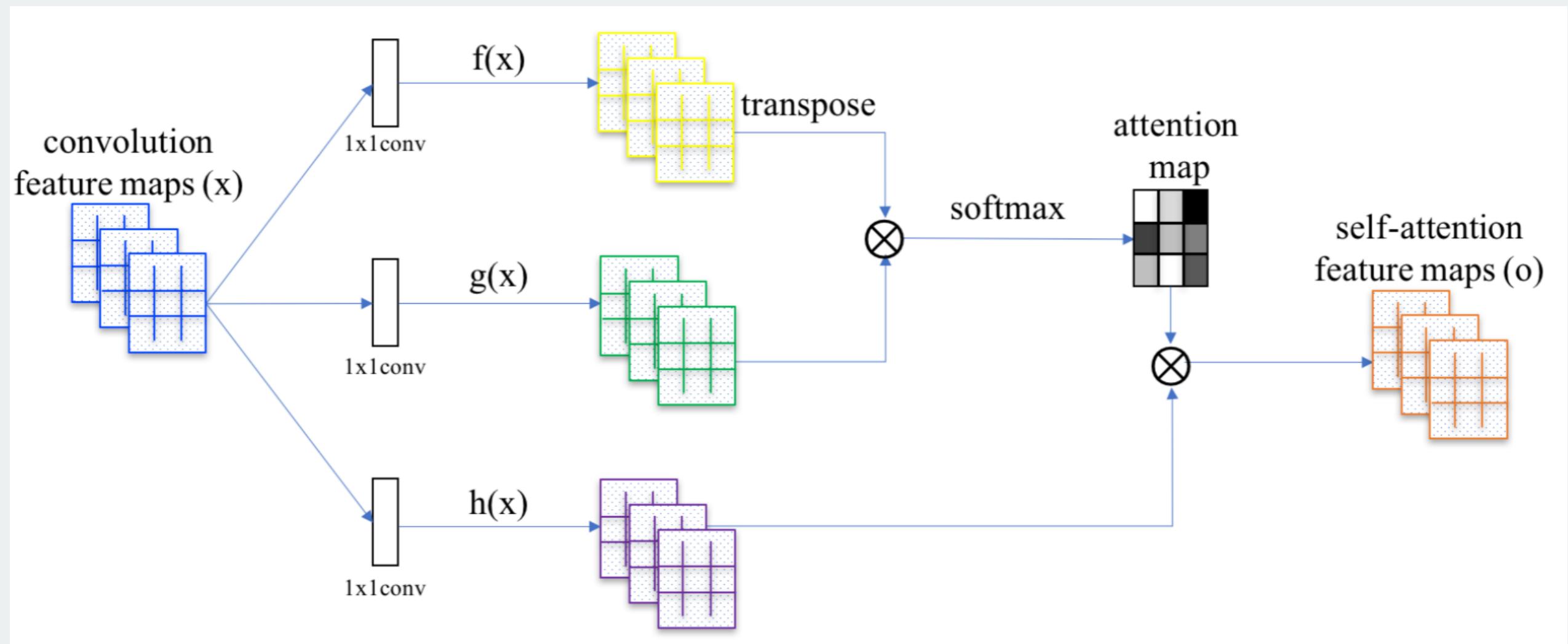
A man is talking on his cell phone  
while another man watches.



# Self-attention CNN

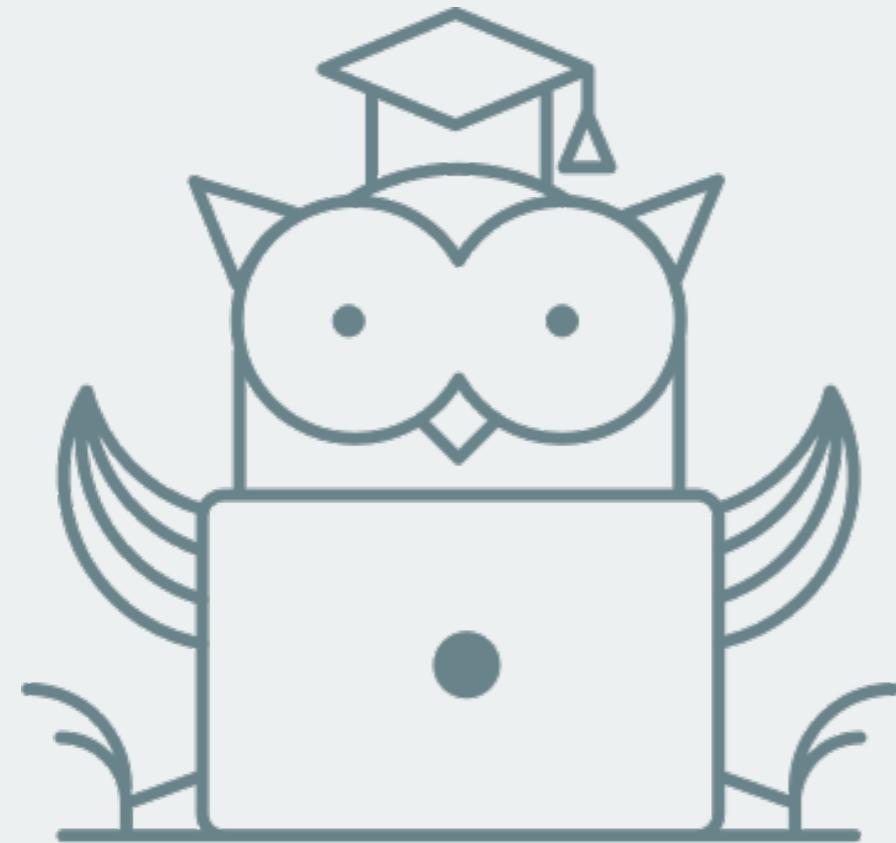


Как и в рекуррентных сетях, к сверточным сетям тоже применим self-attention. Например, такой подход может помочь при генерации изображения.



1. Внимание
2. Самолюбование
3. Титры
- 4. Сцена после титров**





Спасибо  
за внимание!