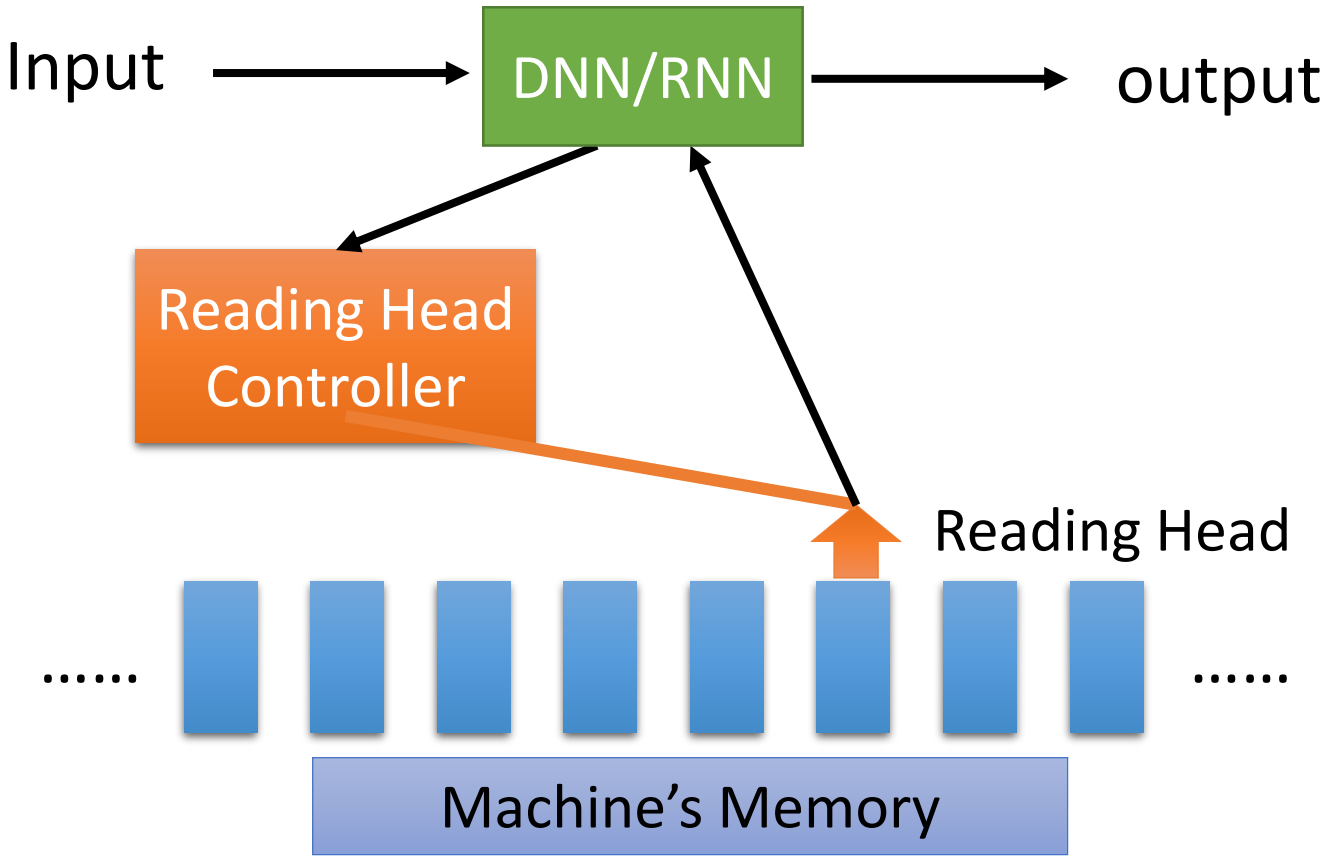
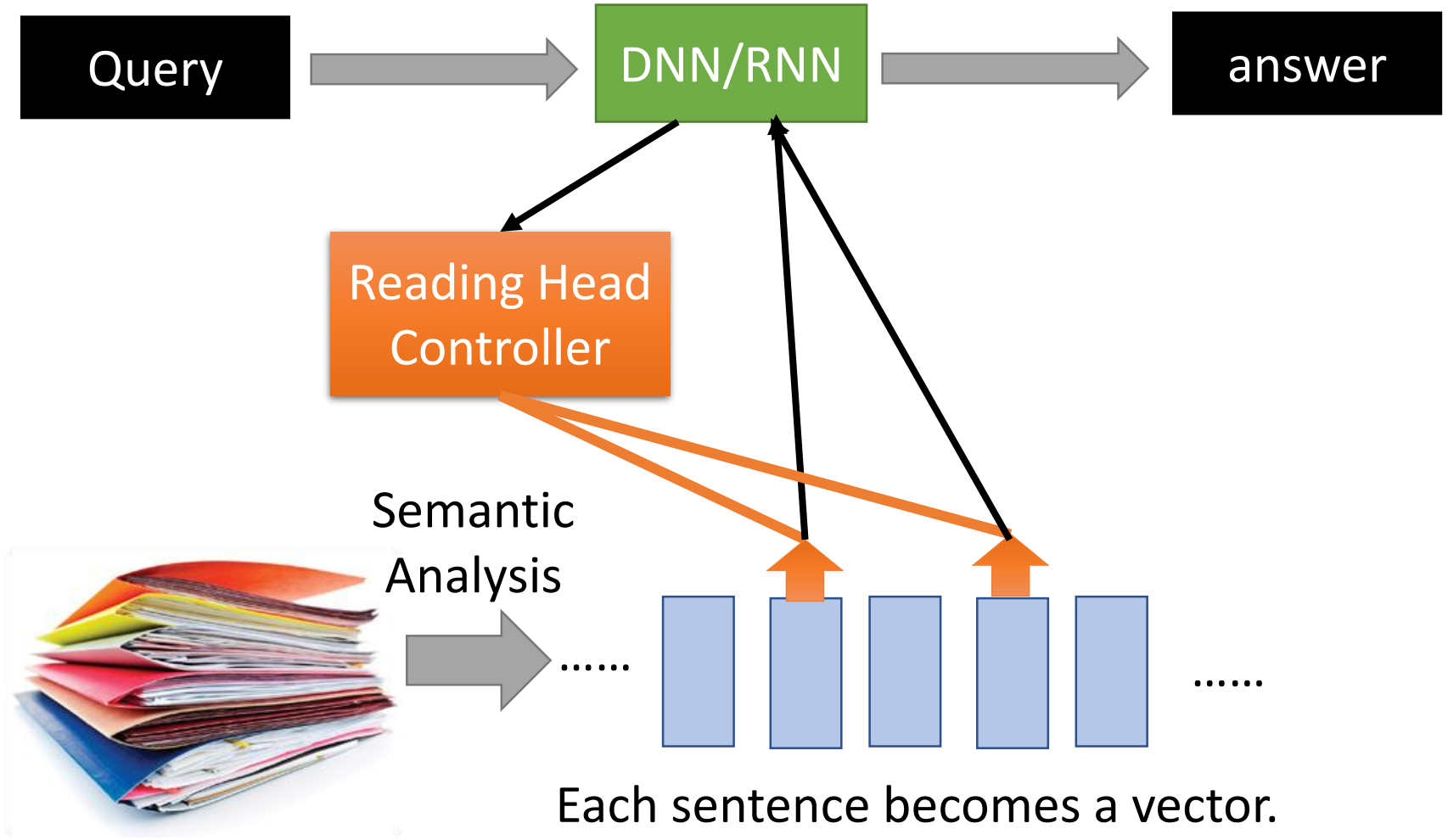
Attention-based Model

External Memory

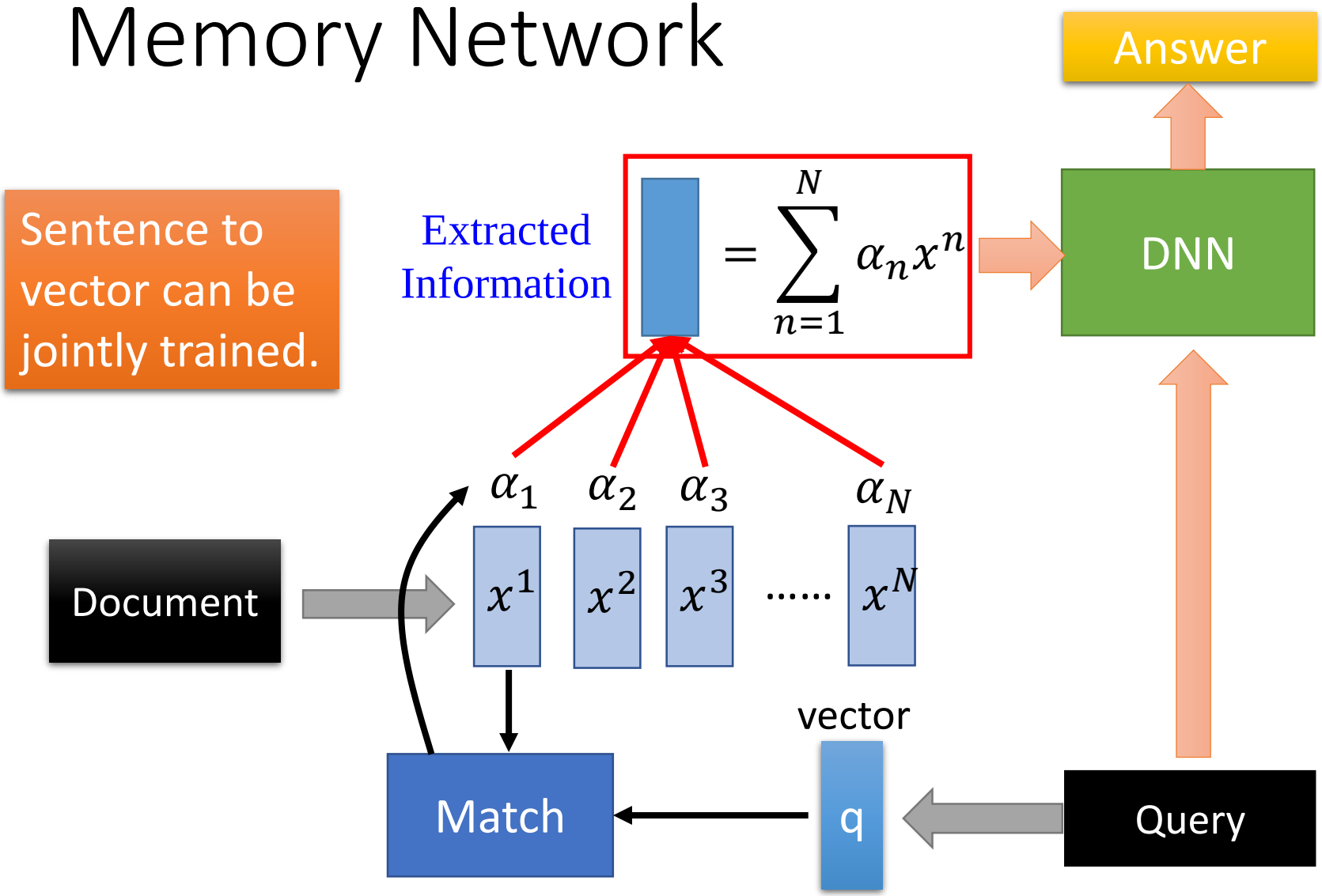


Ref:
http://speech.ee.ntu.edu.tw/~tlkagk/courses/MLDS_2015_2/Lecture/Attain%20(v3).ecm.mp4/index.html

Reading Comprehension

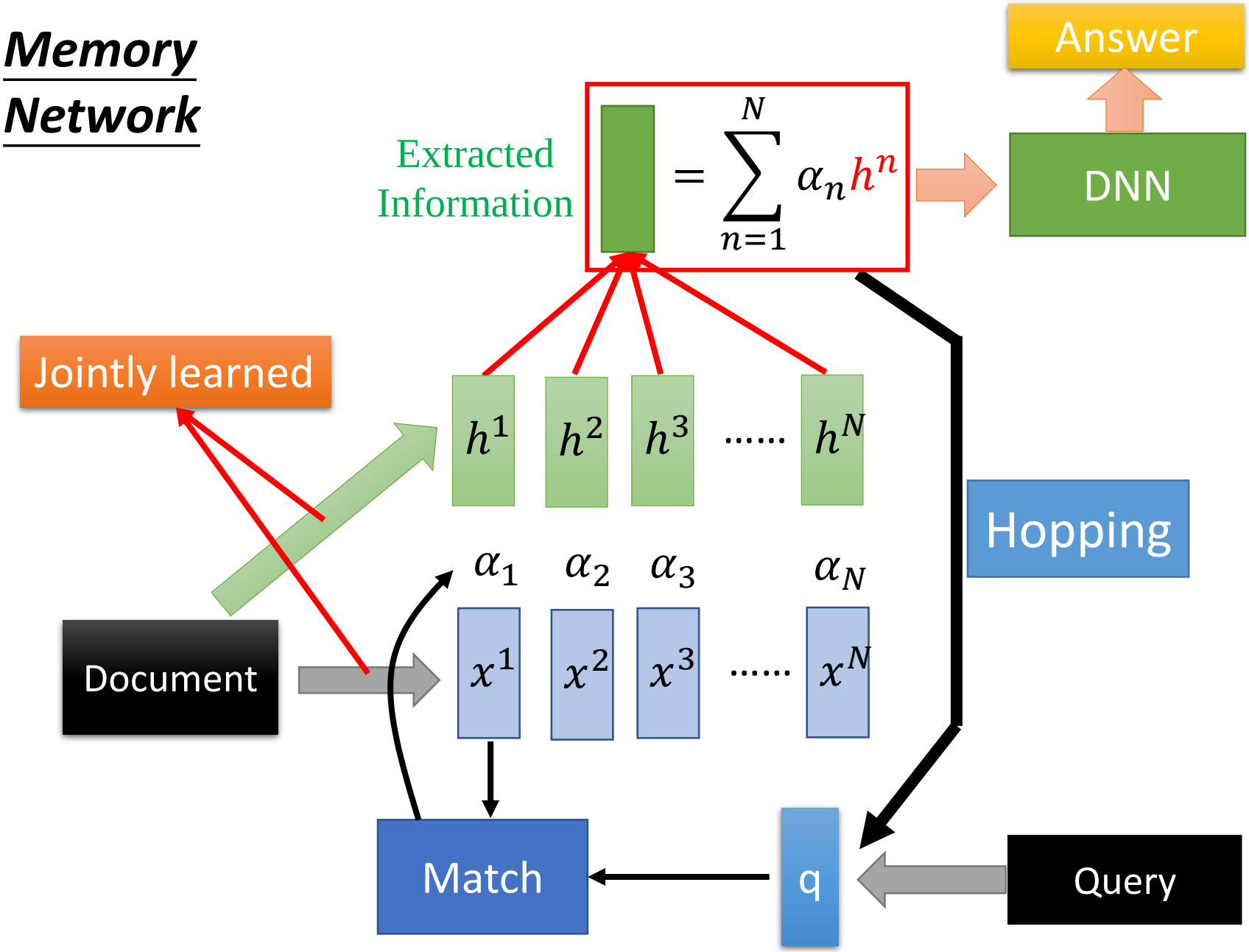


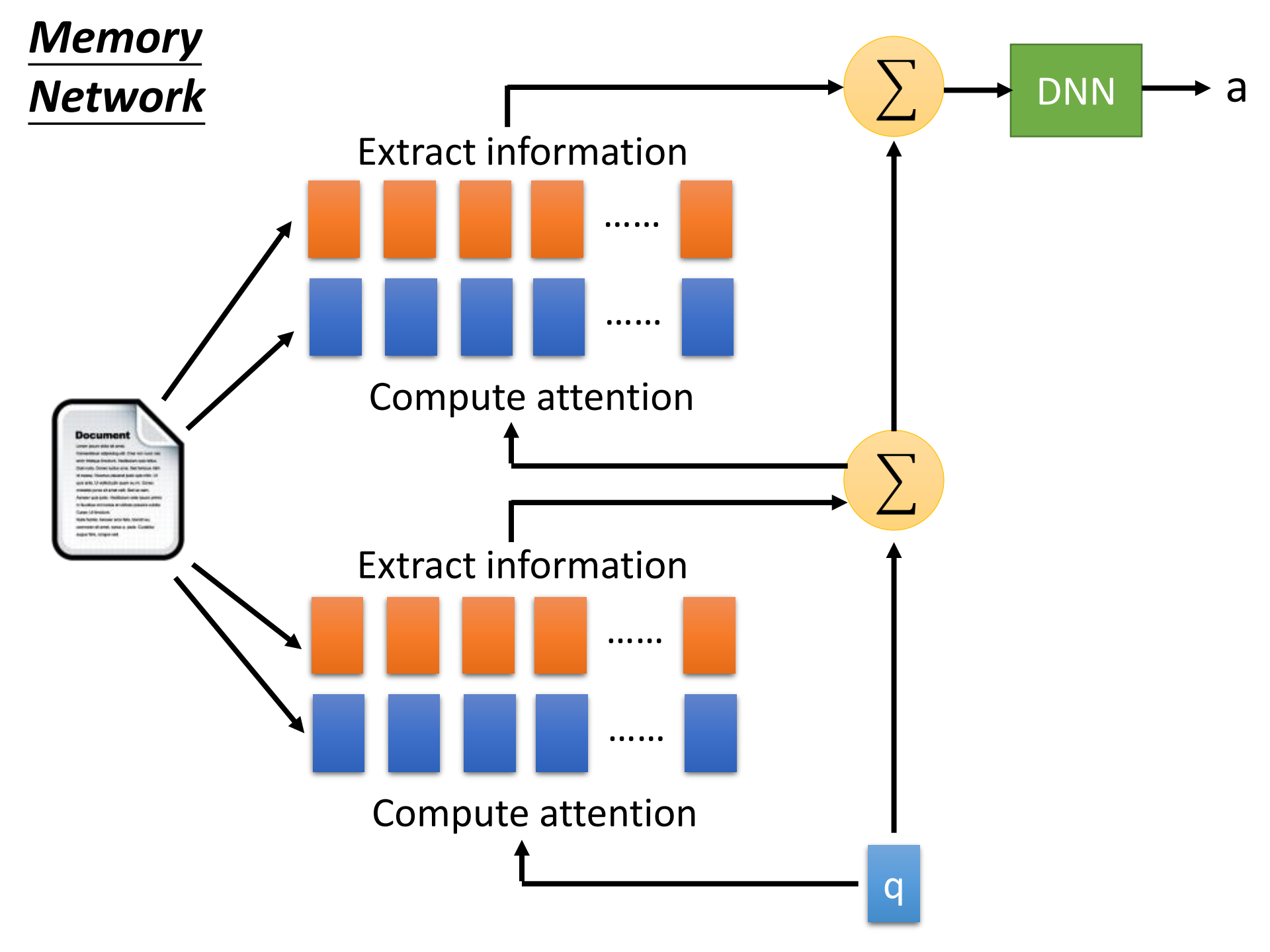
Memory Network



Sainbayar Sukhbaatar, Arthur Szlam, Jason Weston, Rob Fergus, "End-To-End Memory Networks", NIPS, 2015

Memory Network




Memory Network
Extract information
Compute attention
Extract information
Compute attention
Document
DNN
a
q

Multiple-hop

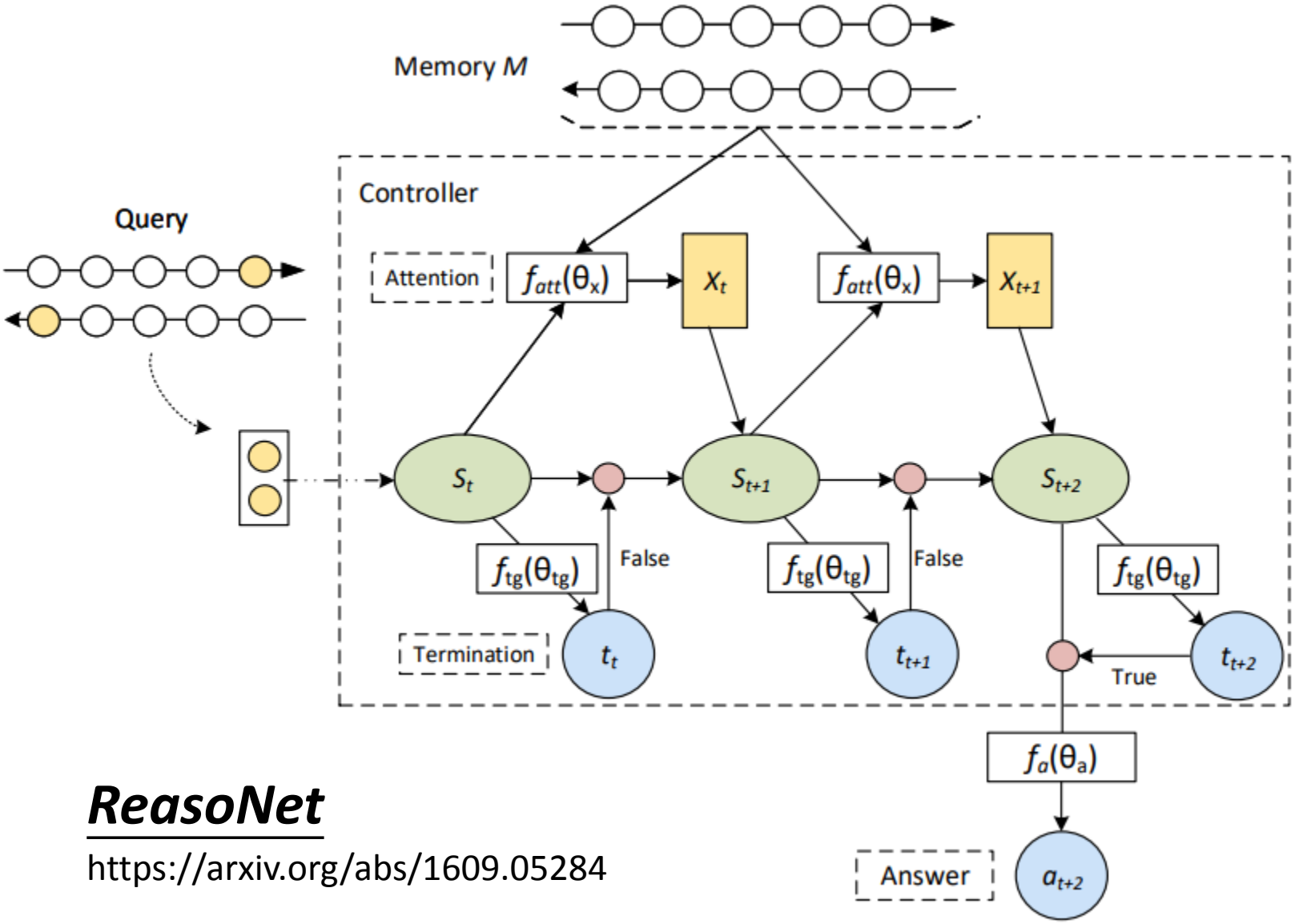
- End-To-End Memory Networks. S. Sukhbaatar, A. Szlam, J. Weston, R. Fergus. NIPS, 2015.

The position of reading head:

Story (16: basic induction)	Support	Hop 1	Hop 2	Hop 3
Brian is a frog.	yes	0.00	0.98	0.00
Lily is gray.		0.07	0.00	0.00
Brian is yellow.	yes	0.07	0.00	1.00
Julius is green.		0.06	0.00	0.00
Greg is a frog.	yes	0.76	0.02	0.00
What color is Greg? Answer: yellow	**Prediction: yellow**			

Keras has example:
https://github.com/fchollet/keras/blob/master/examples/babi_memnn.py

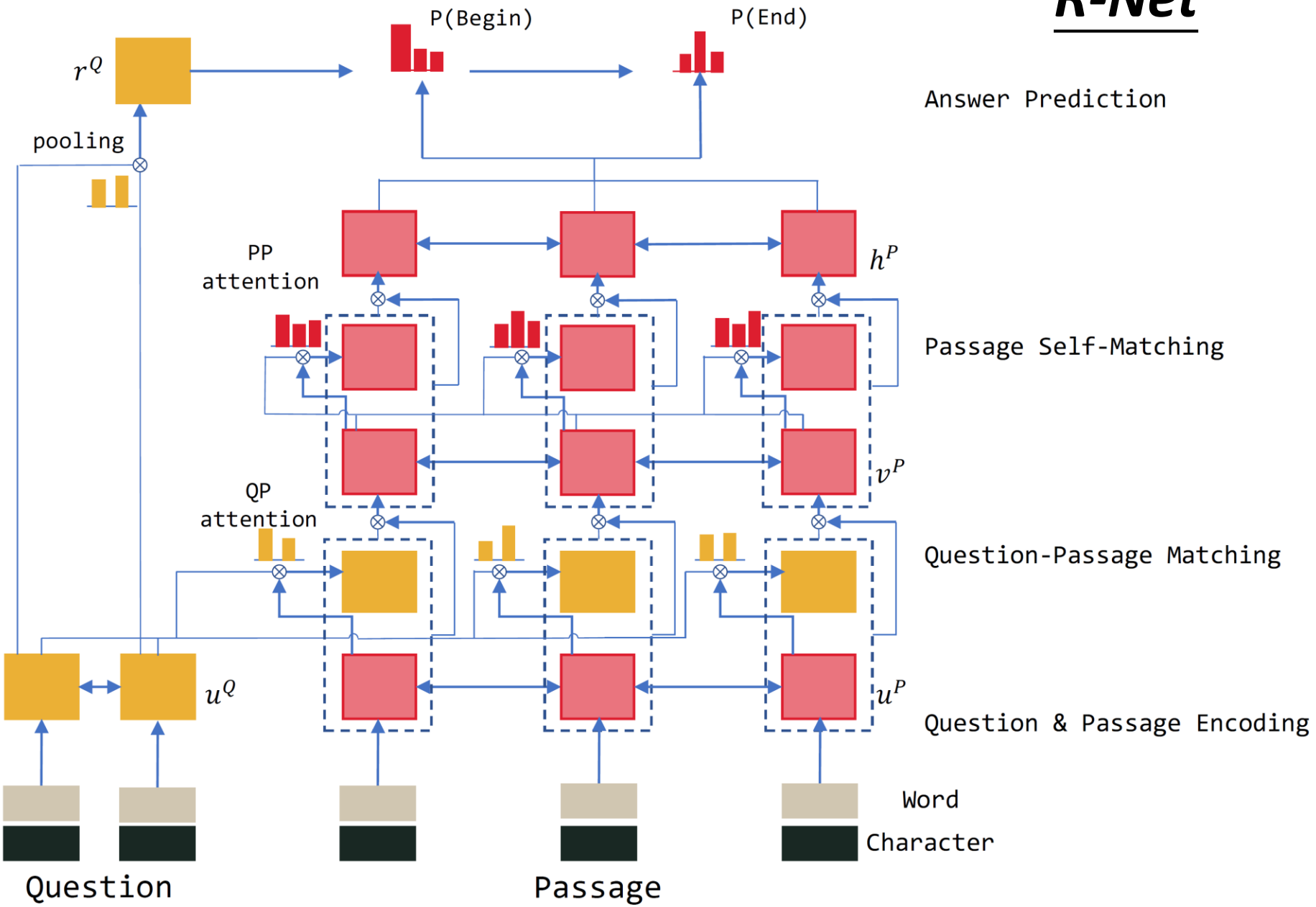
Multiple-hop

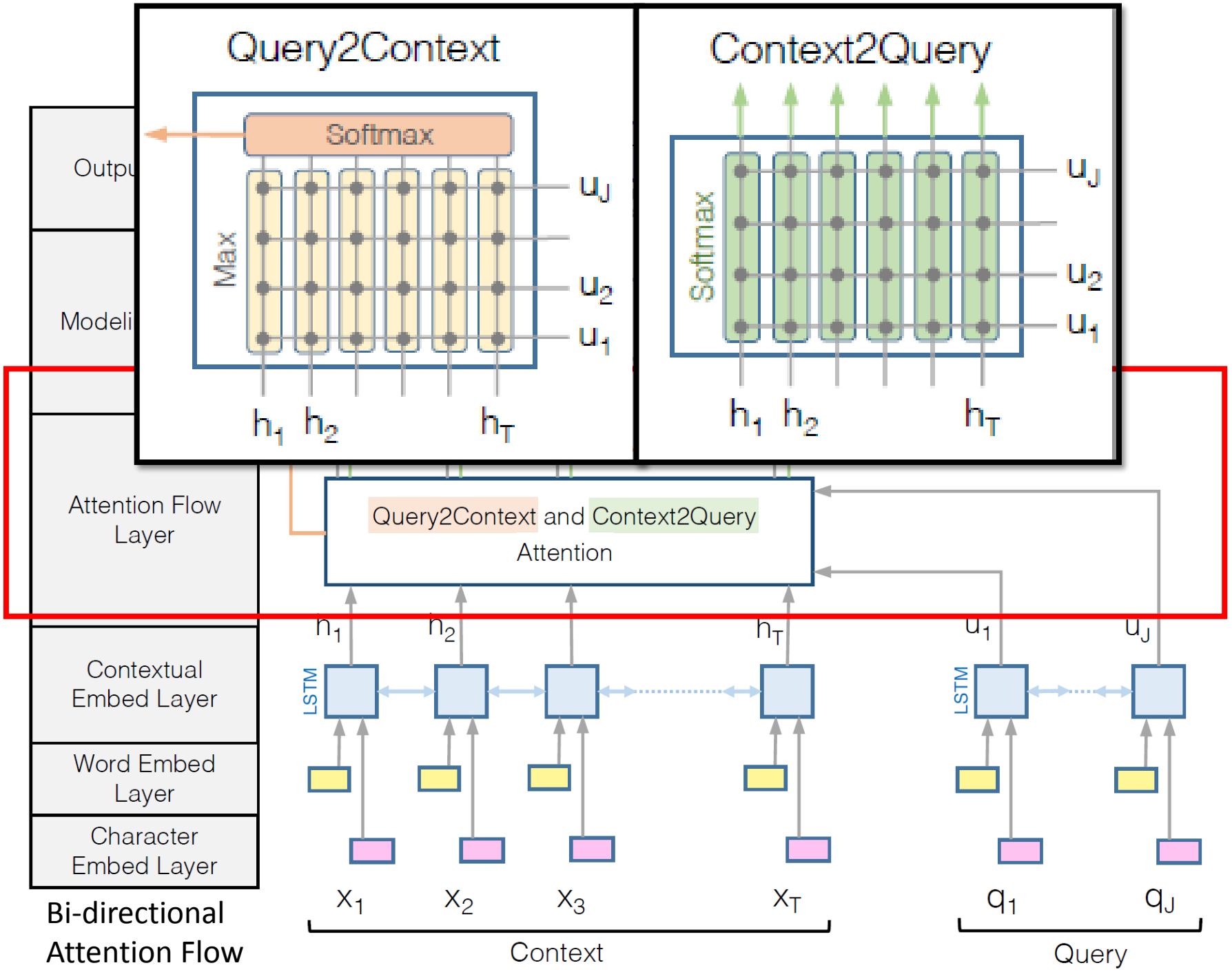


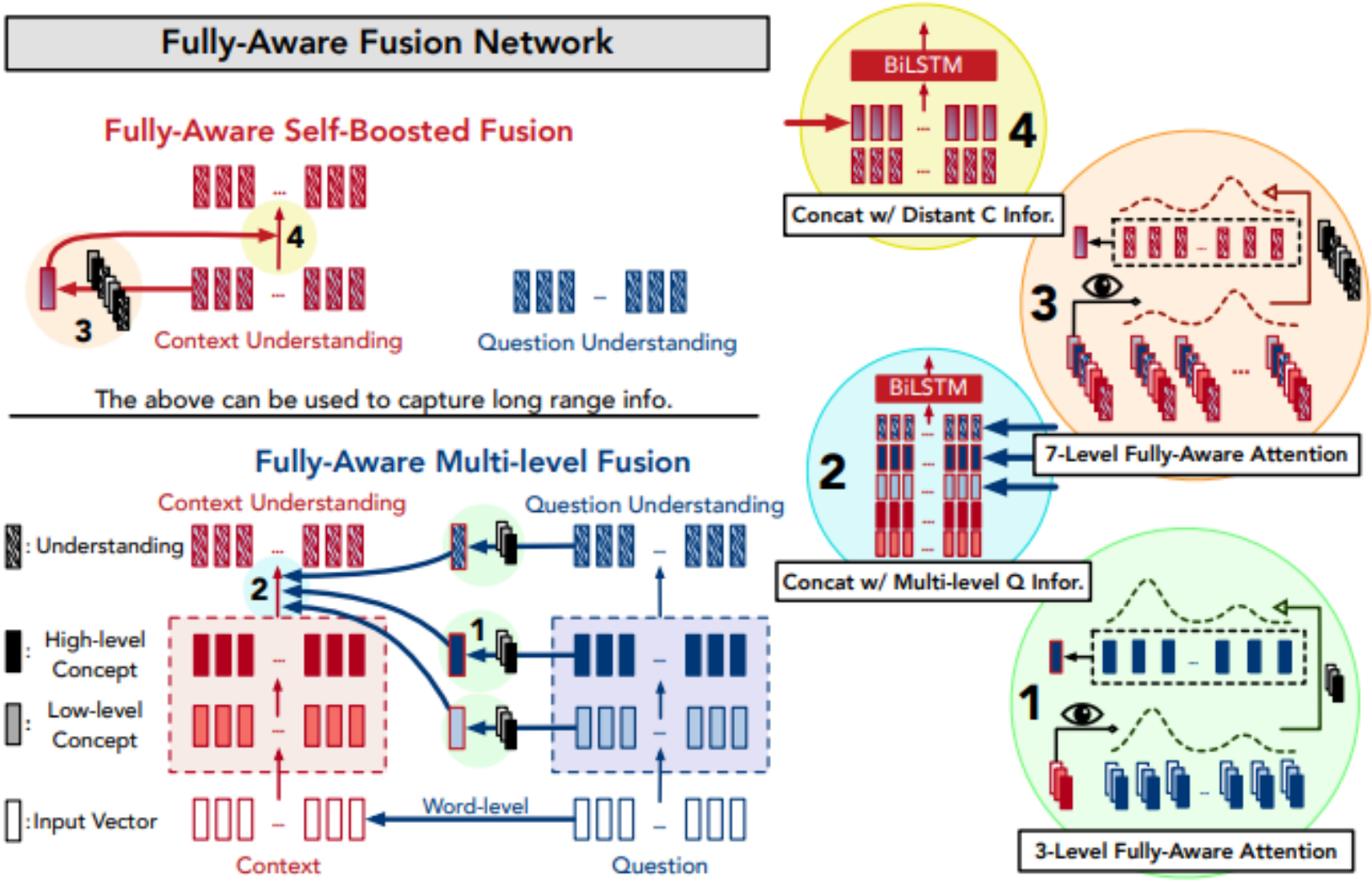
ReasoNet

https://arxiv.org/abs/1609.05284

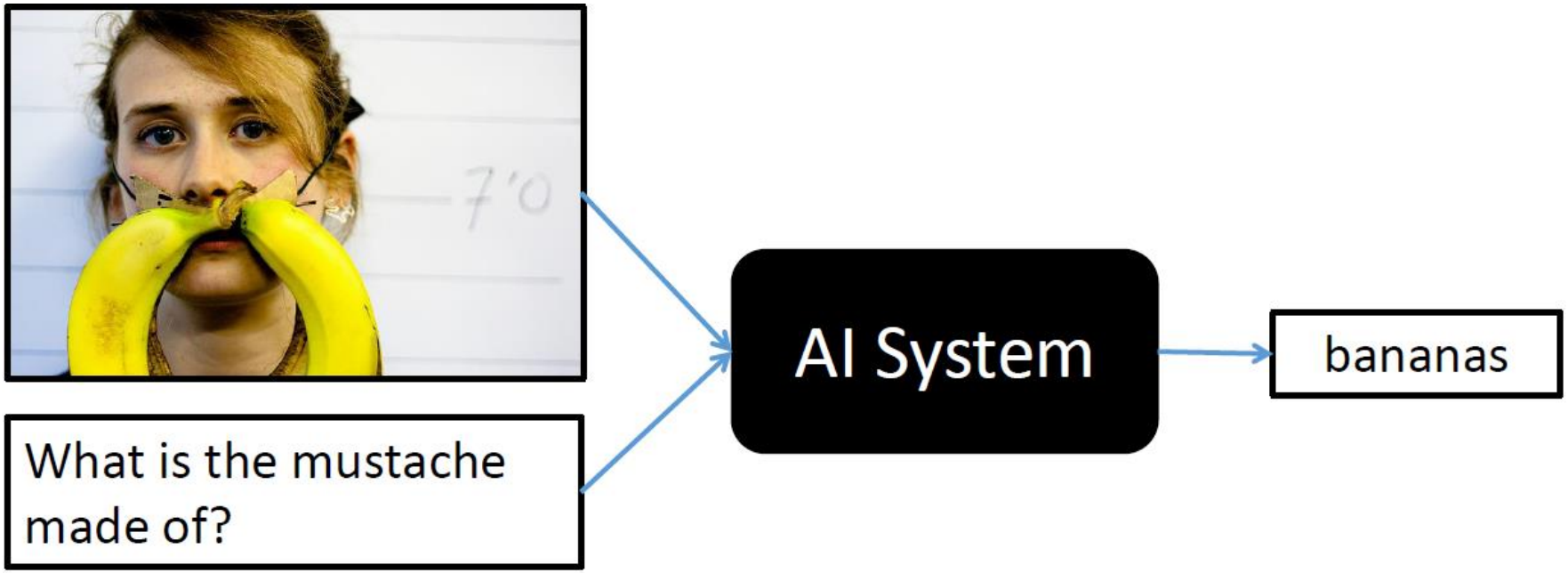
R-Net


P(Begin)
P(End)
r^Q
Answer Prediction
pooling
PP attention
h^P
Passage Self-Matching
v^P
QP attention
Question-Passage Matching
u^Q
u^P
Question & Passage Encoding
Word
Character
Question
Passage


Query2Context
Softmax
Max
u_J
u_2
u_1
h_1 h_2 h_T
Context2Query
Softmax
u_J
u_2
u_1
h_1 h_2 h_T
Outpu
Modeli
Attention Flow Layer
Query2Context and Context2Query Attention
Contextual Embed Layer
LSTM
h_1 h_2 h_T
u_1 u_J
Word Embed Layer
Character Embed Layer
x_1 x_2 x_3 x_T
q_1 q_J
Context
Query
Bi-directional Attention Flow

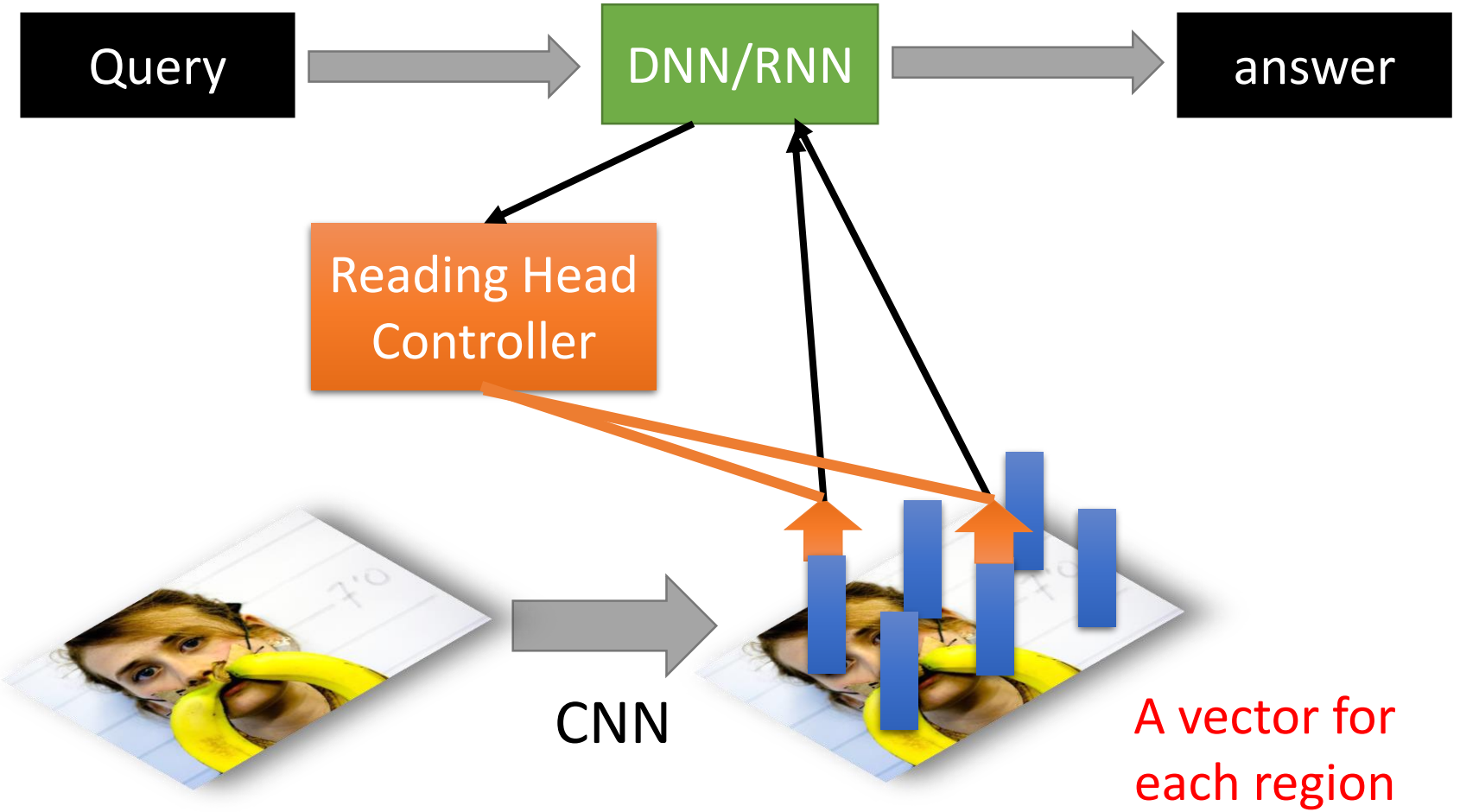

Fully-Aware Fusion Network
Fully-Aware Self-Boosted Fusion
4
3
Context Understanding
Question Understanding
The above can be used to capture long range info.
Fully-Aware Multi-level Fusion
Context Understanding
Question Understanding
: Understanding
2
1
: High-level Concept
: Low-level Concept
: Input Vector
Word-level
Context
Question
BiLSTM
4
Concat w/ Distant C Infor.
3
7-Level Fully-Aware Attention
BiLSTM
2
Concat w/ Multi-level Q Infor.
1
3-Level Fully-Aware Attention

Visual Question Answering



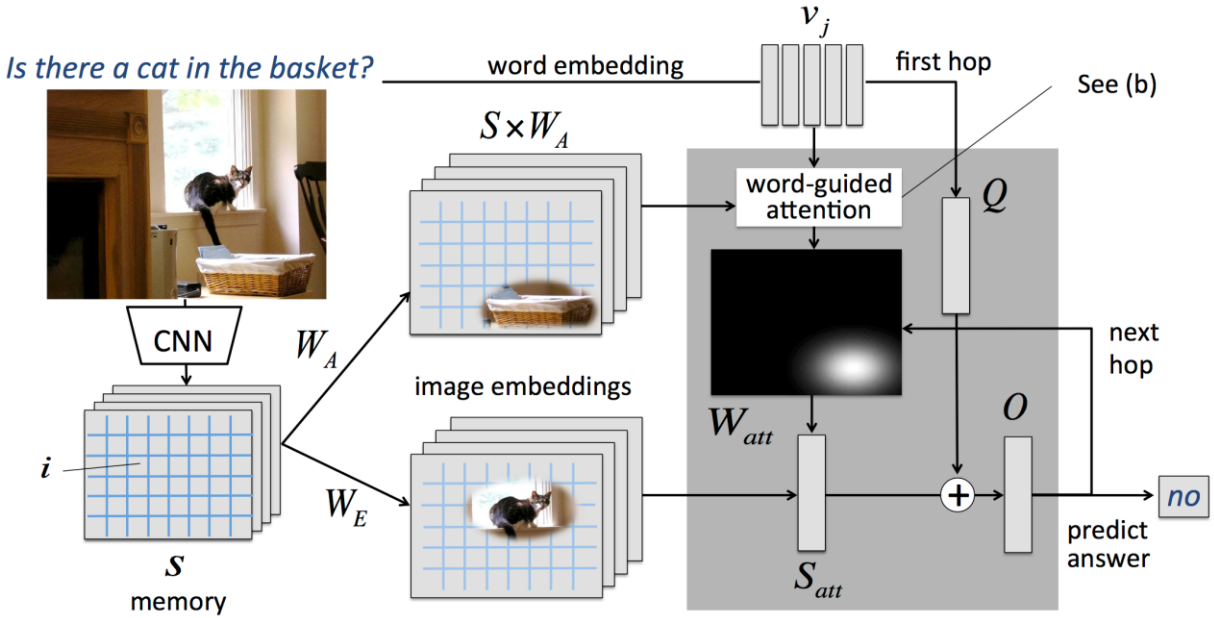
source: http://visualqa.org/

Visual Question Answering

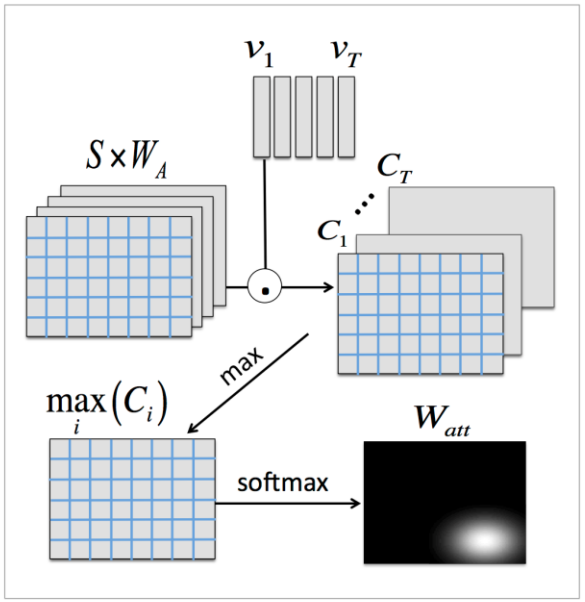


Visual Question Answering

- Huijuan Xu, Kate Saenko. Ask, Attend and Answer: Exploring Question-Guided Spatial Attention for Visual Question Answering. arXiv Pre-Print, 2015



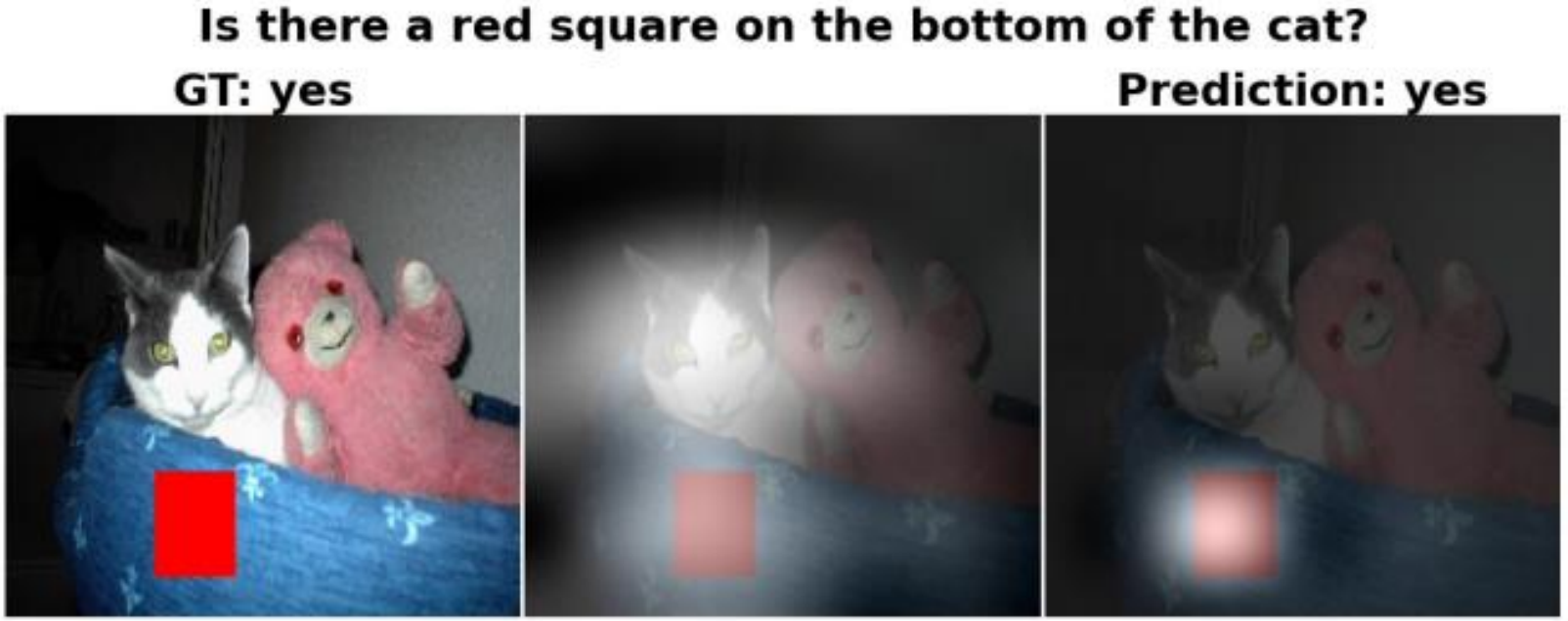
(a) Overview



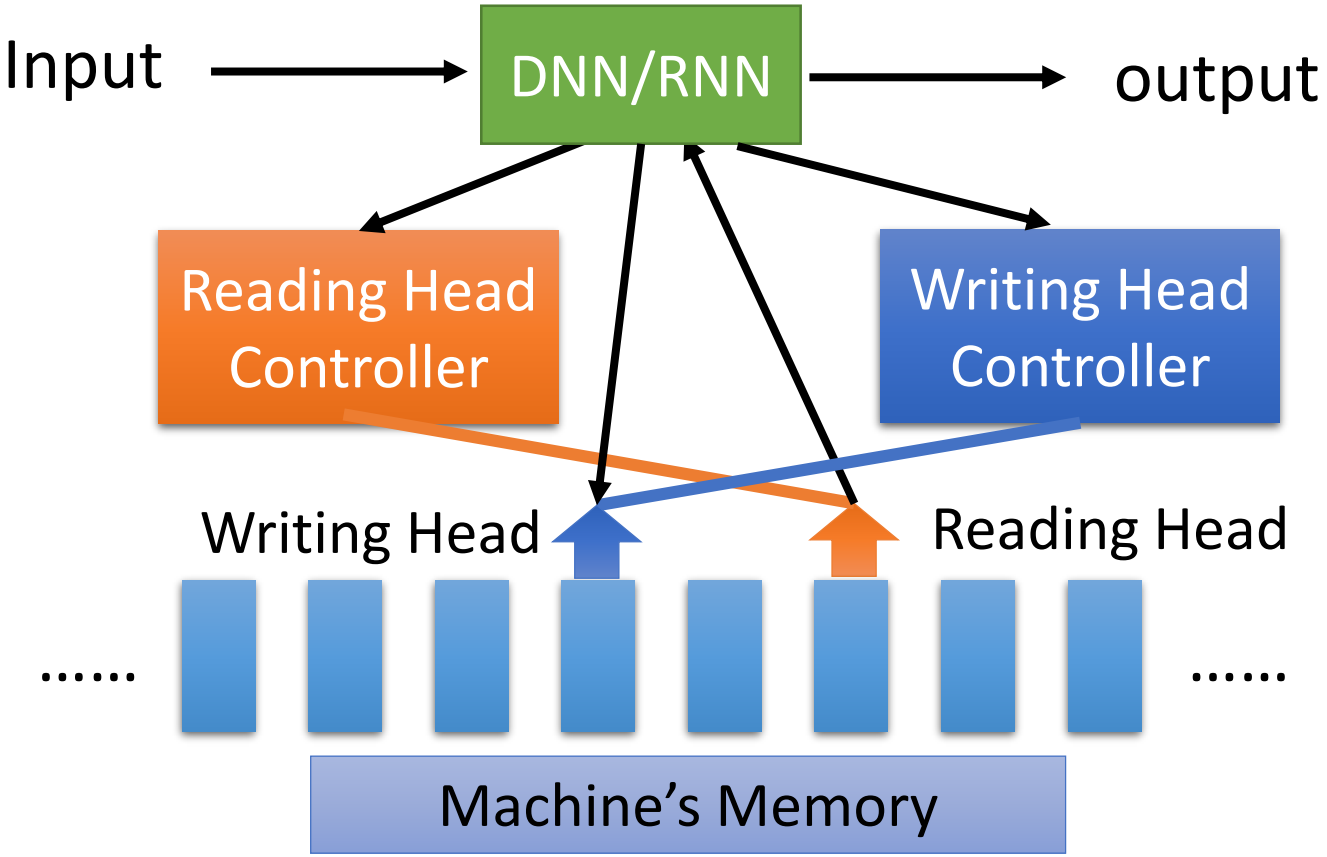
(b) Word-guided attention

Visual Question Answering

- Huijuan Xu, Kate Saenko. Ask, Attend and Answer: Exploring Question-Guided Spatial Attention for Visual Question Answering. arXiv Pre-Print, 2015

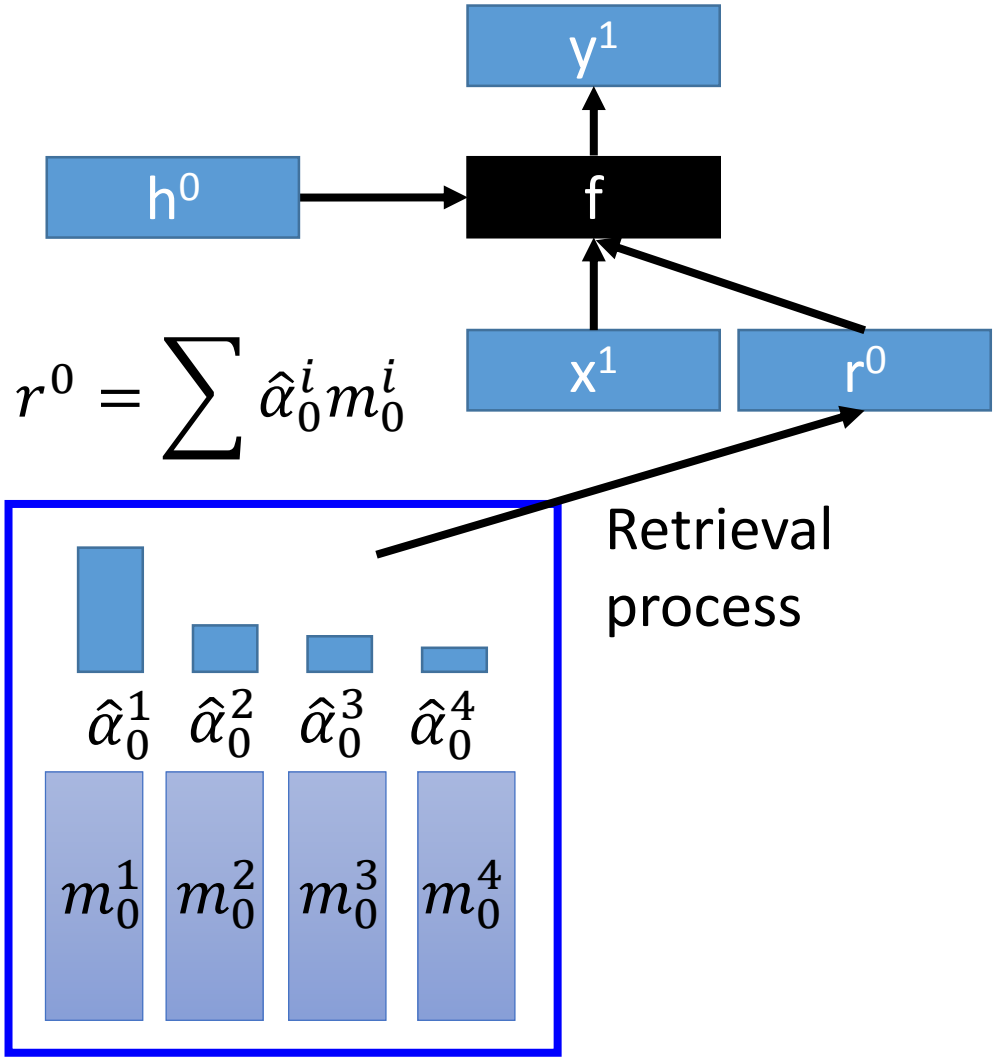


External Memory v2

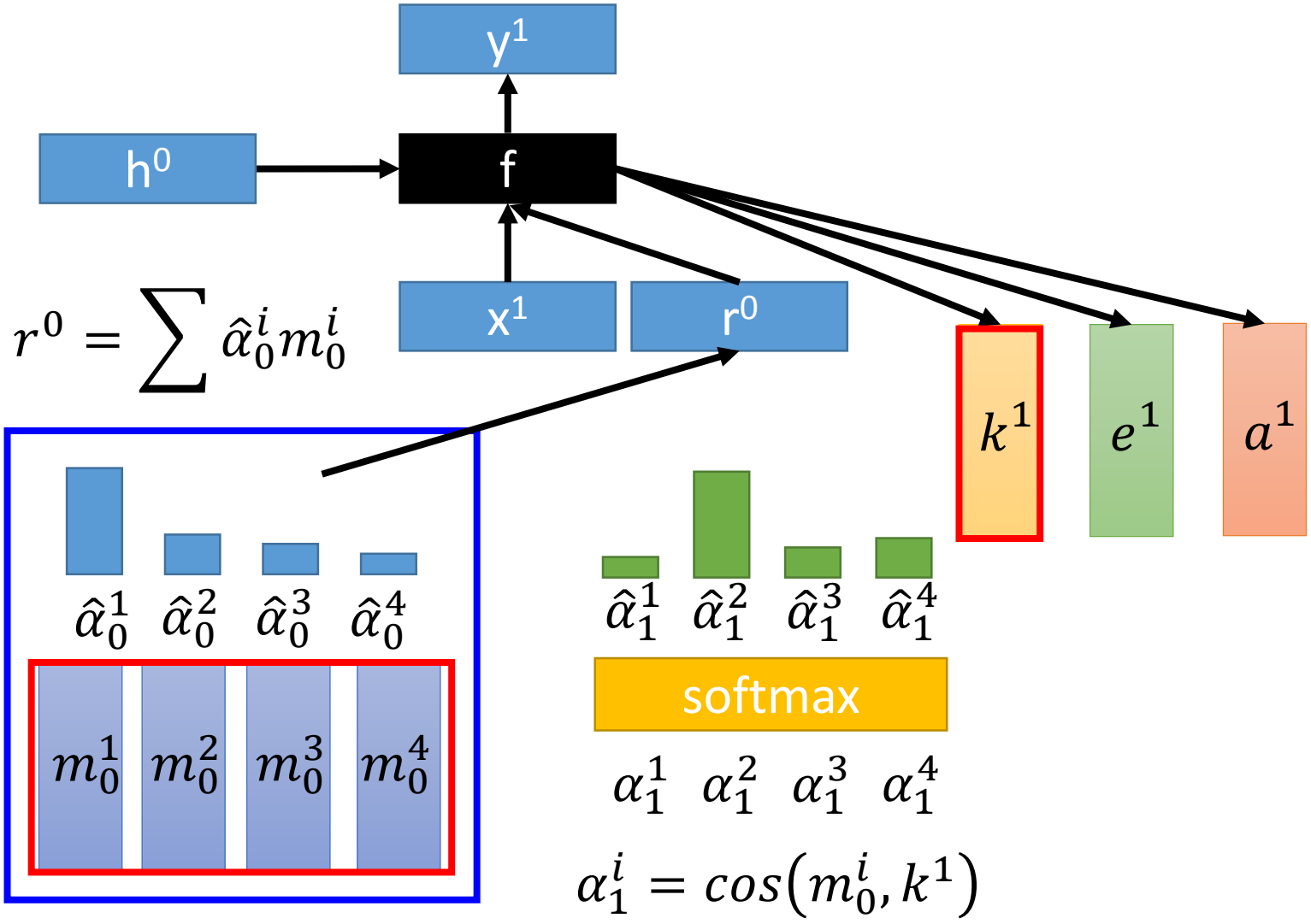


Neural Turing Machine

Neural Turing Machine



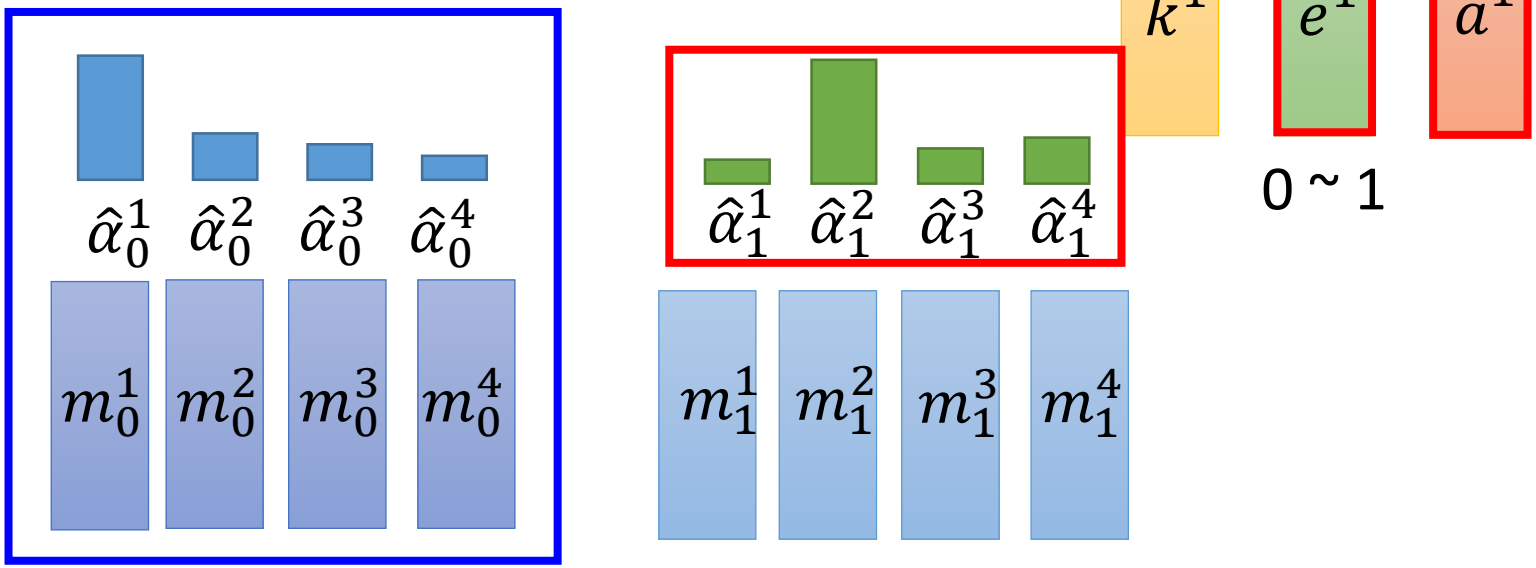
Neural Turing Machine



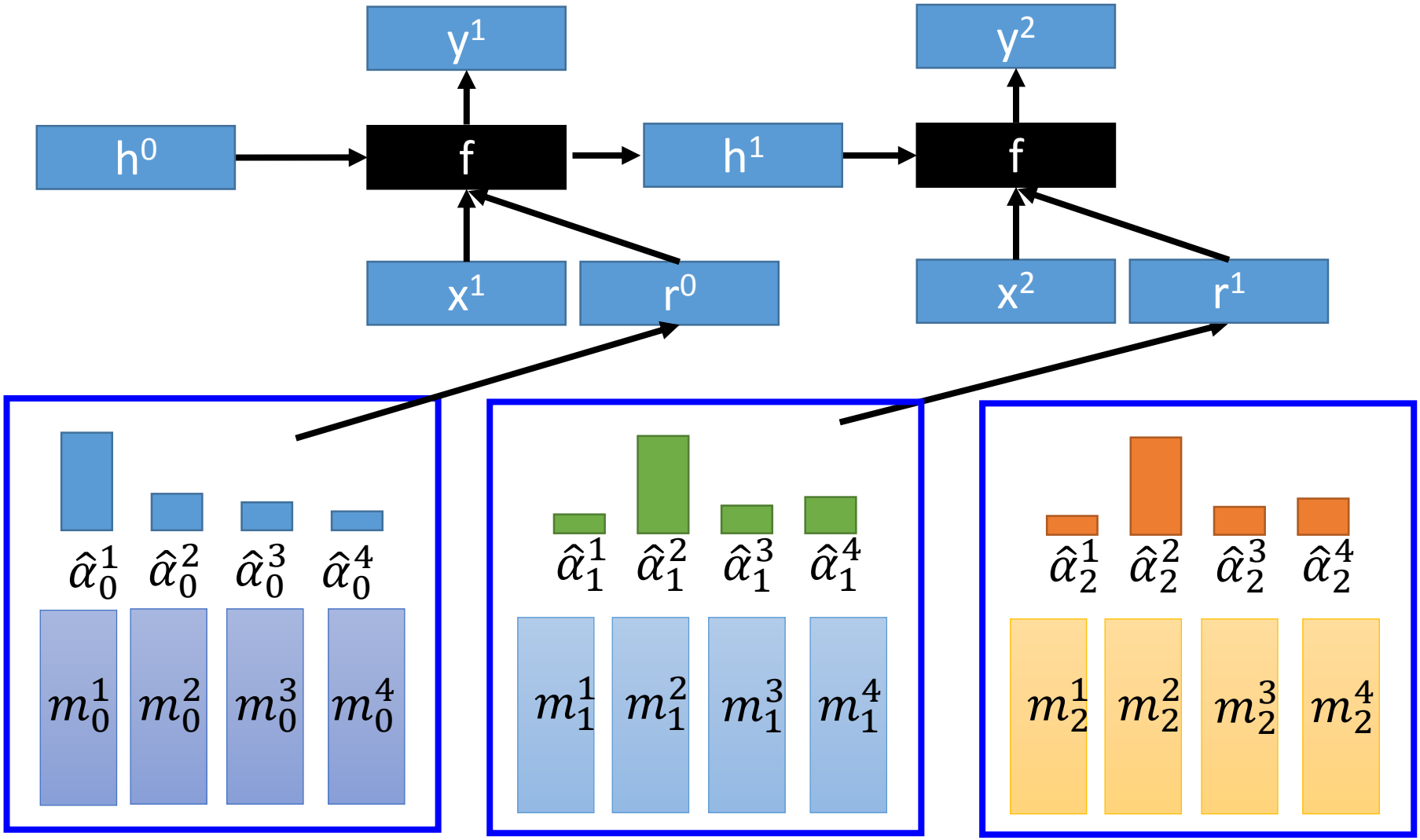
Neural Turing Machine

$$m_1^i = m_0^i - \hat{\alpha}_1^i \, e^1 \odot m_0^i + \hat{\alpha}_1^i \, a^1$$

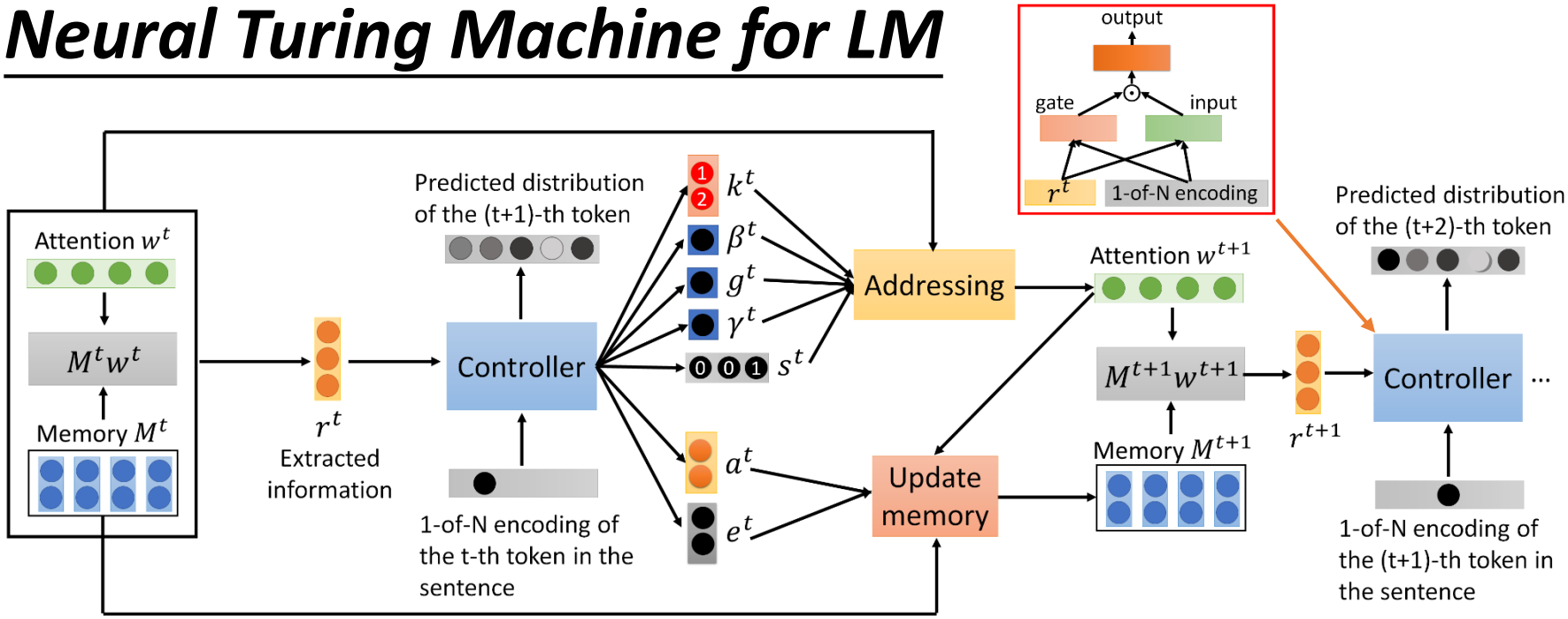
(element-wise)

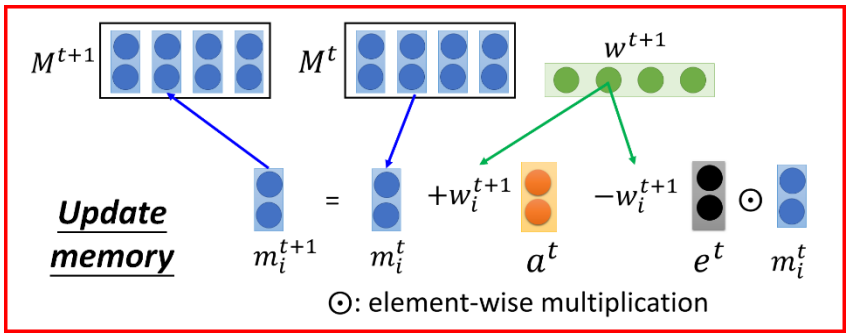


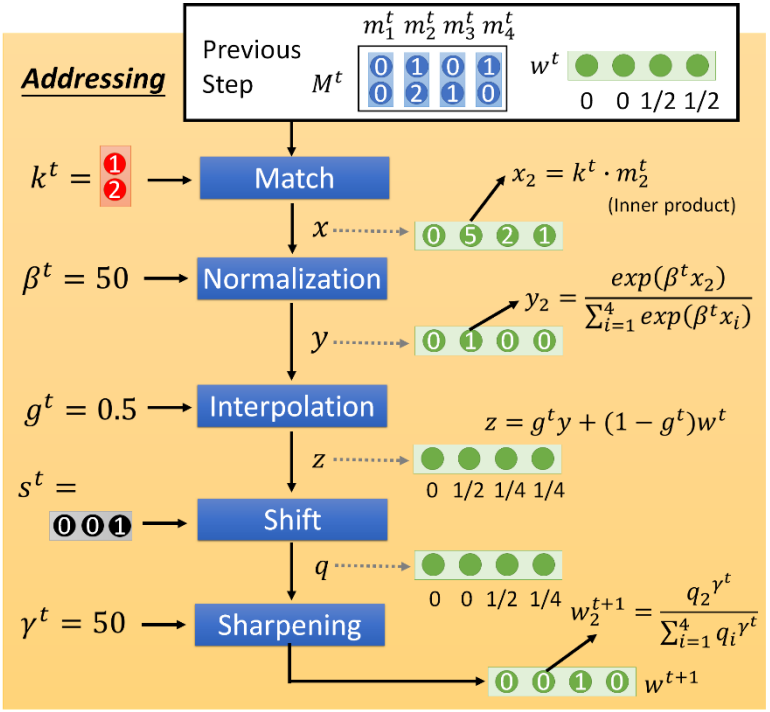
Neural Turing Machine



Neural Turing Machine for LM







Wei-Jen Ko, Bo-Hsiang Tseng, Hung-yi Lee, "Recurrent Neural Network based Language Modeling with Controllable External Memory", ICASSP, 2017

Armand Joulin, Tomas Mikolov, Inferring Algorithmic Patterns with Stack-Augmented Recurrent Nets, arXiv Pre-Print, 2015

Stack RNN