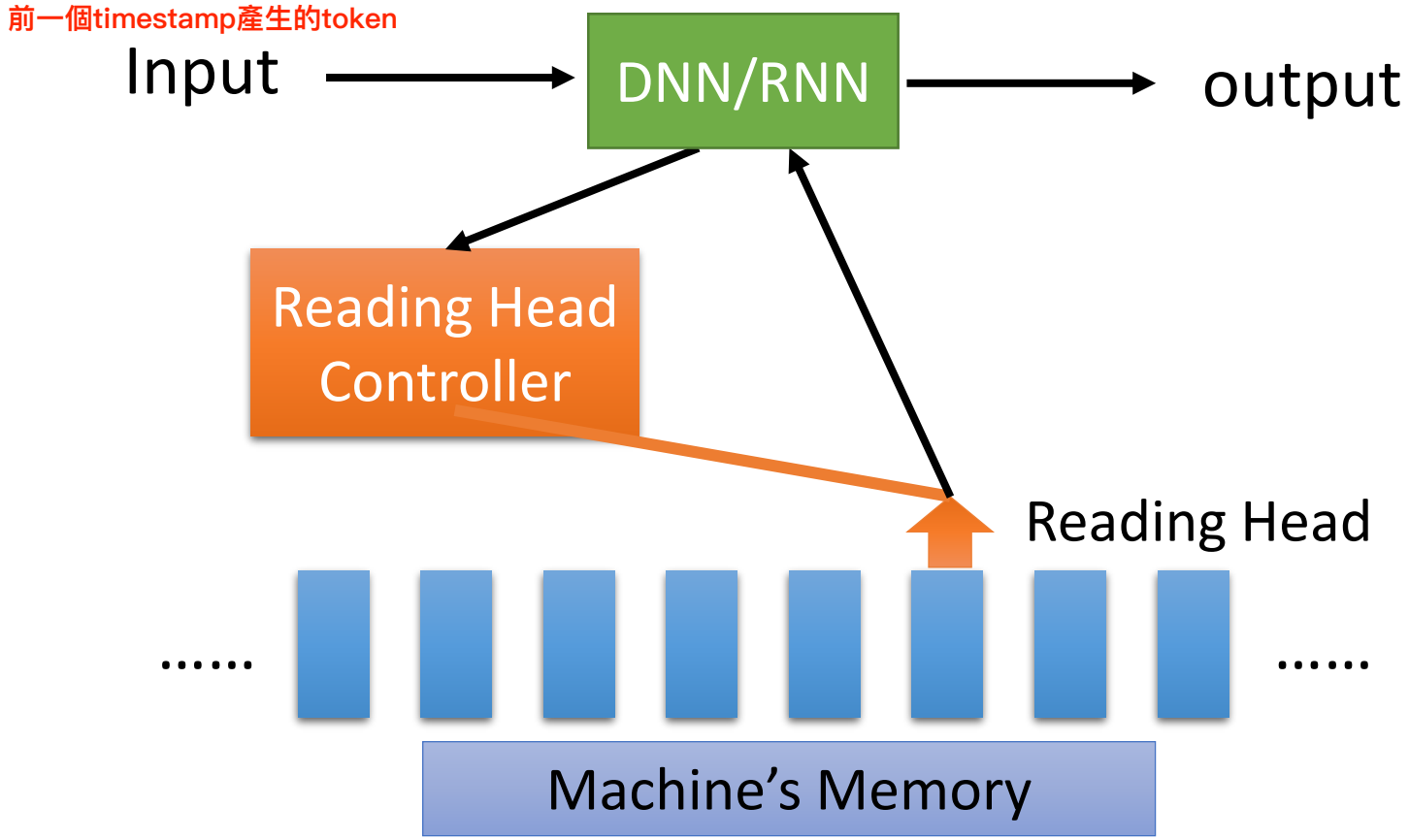
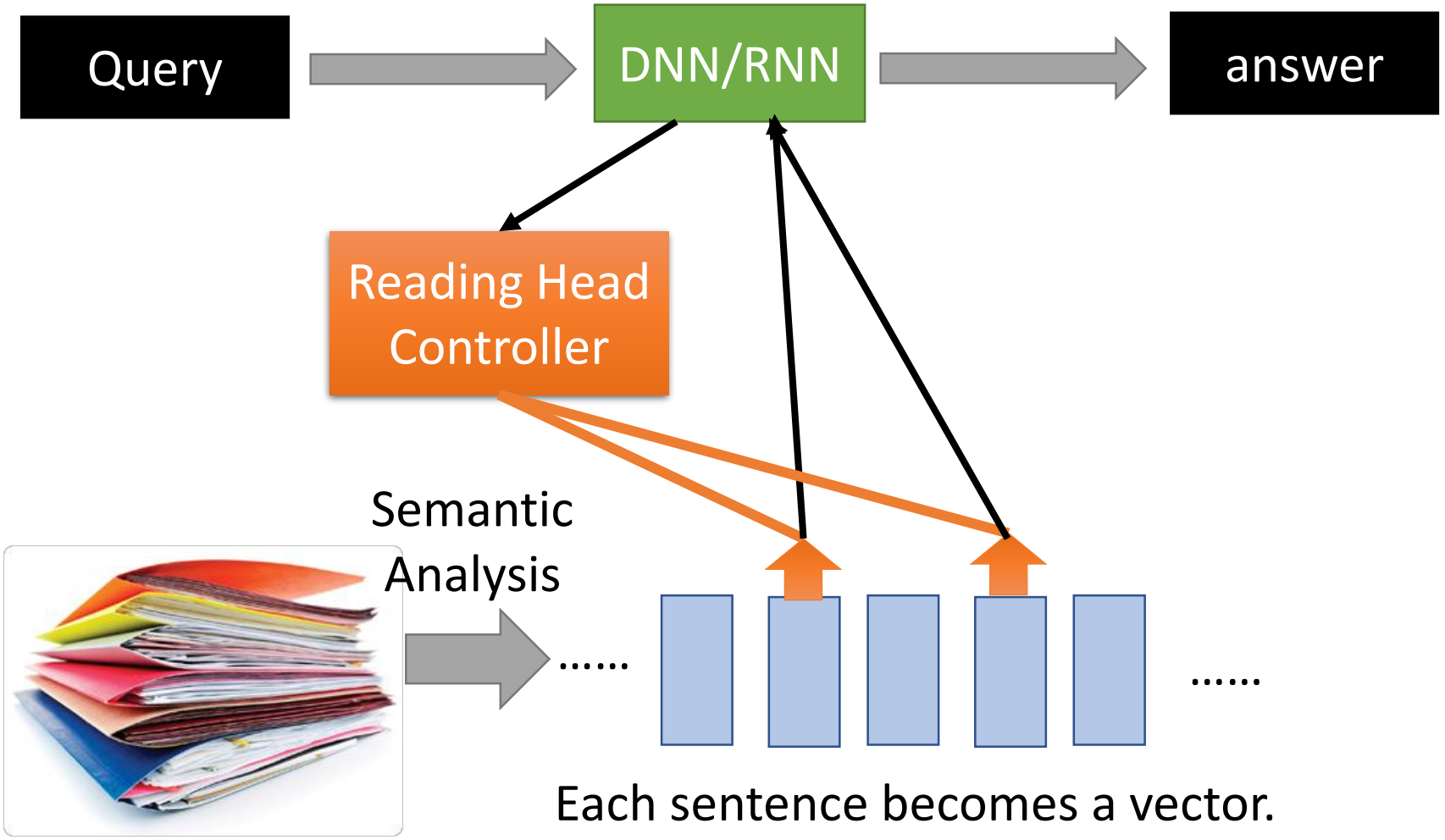
Attention-based Model

External Memory



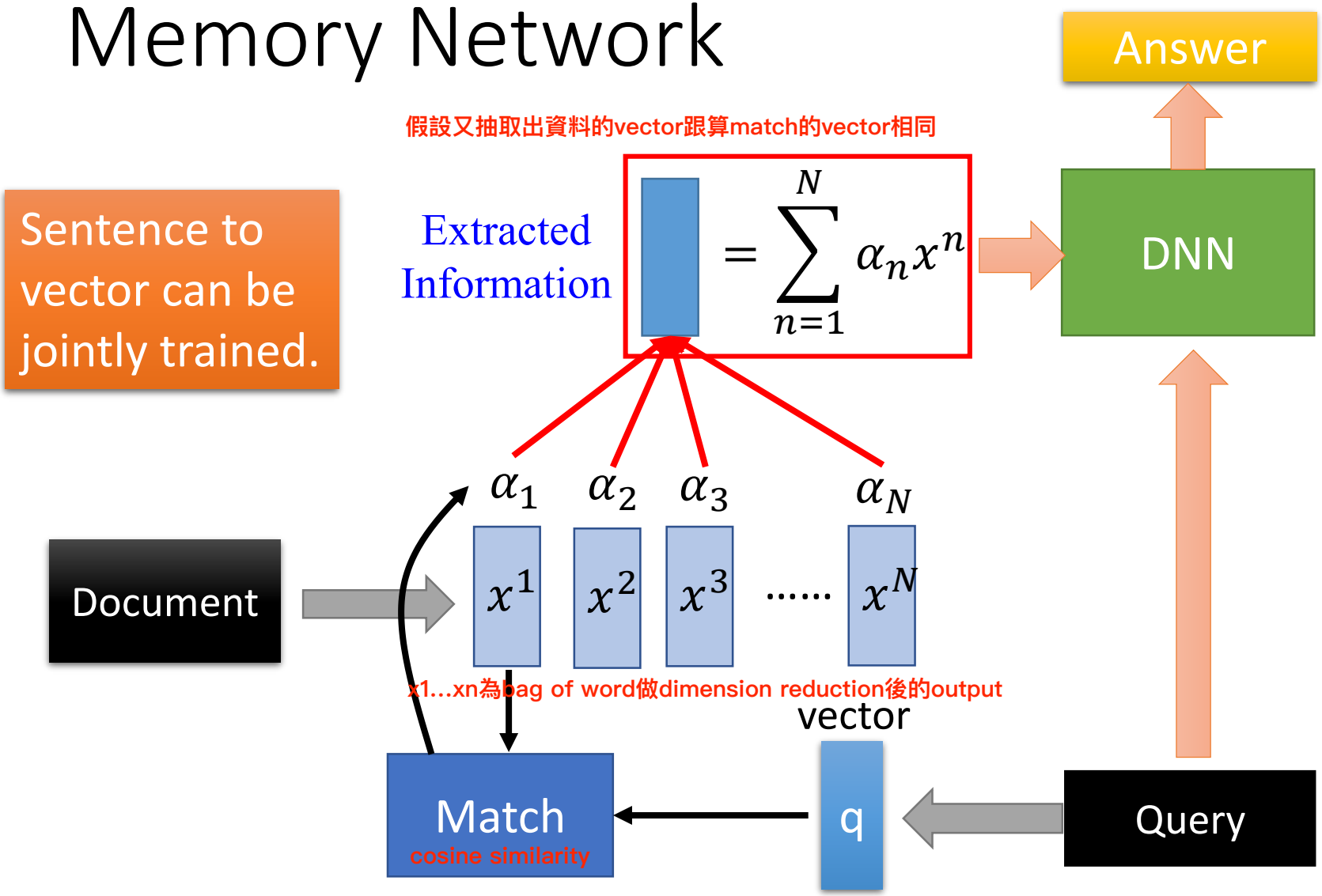
Ref:
http://speech.ee.ntu.edu.tw/~tlkagk/courses/MLDS_2015_2/Lecture/Attain%20(v3).ecm.mp4/index.html

Reading Comprehension



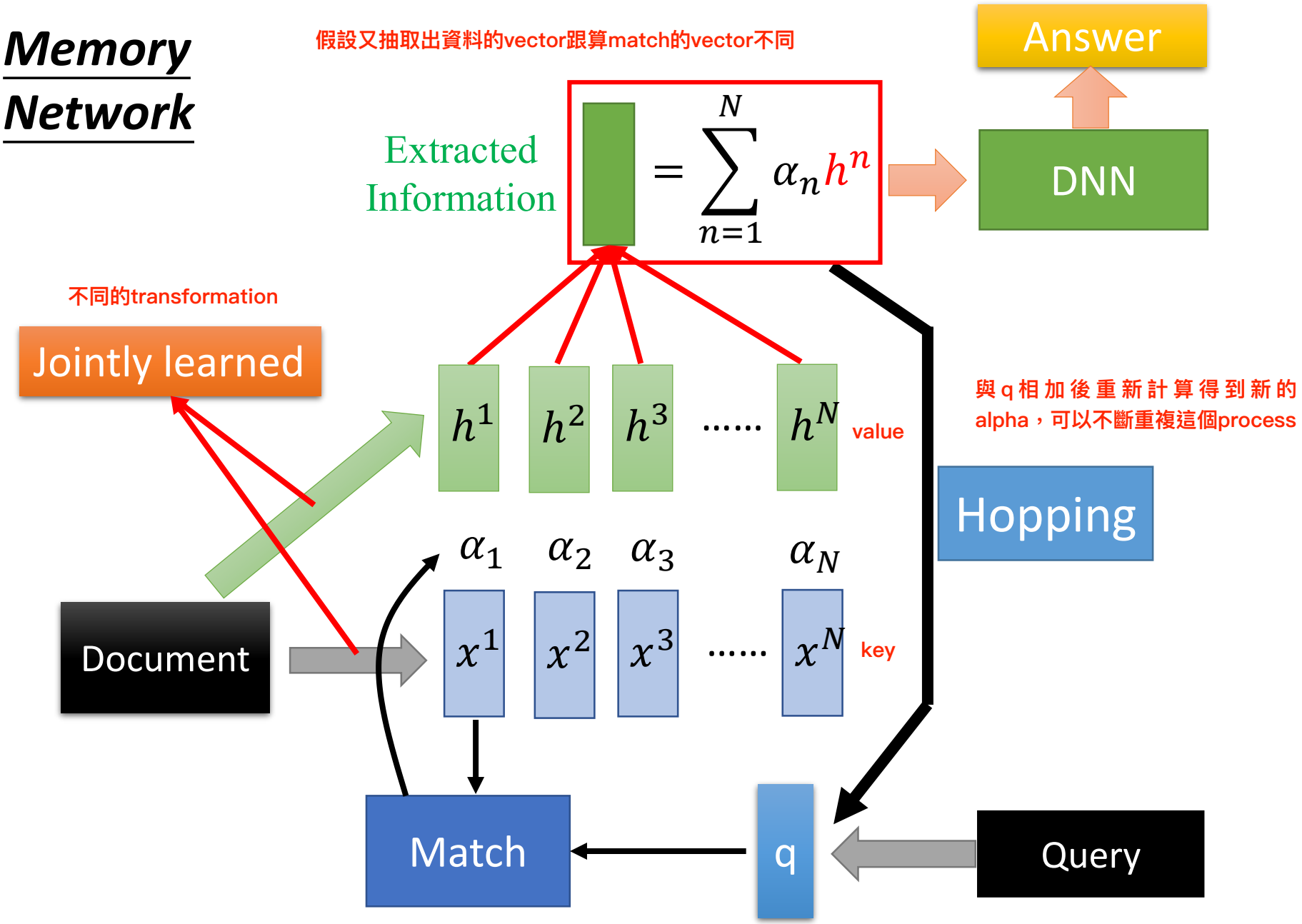
SQUAD – attention based model completion like Kaggle

Memory Network



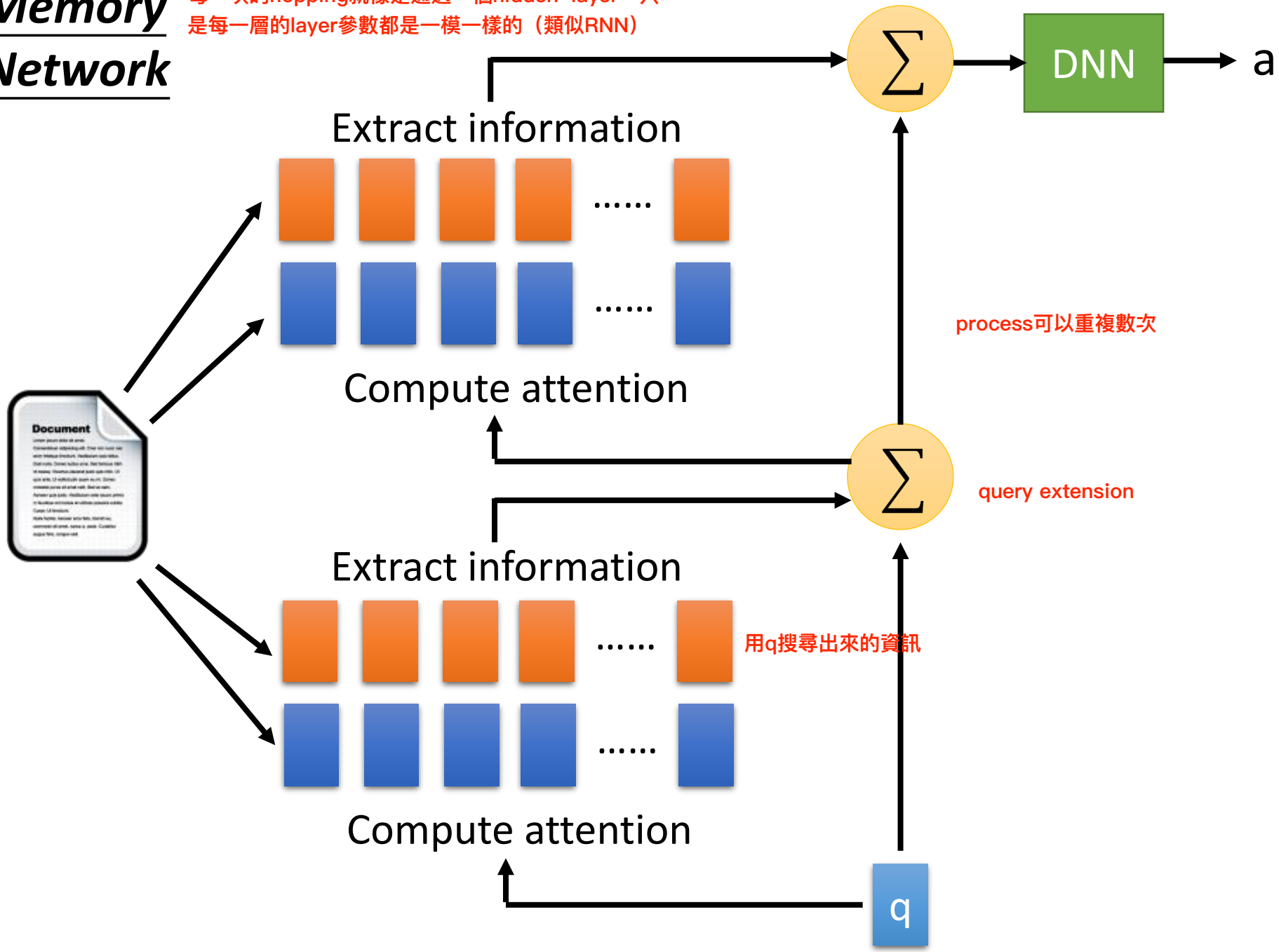
Sainbayar Sukhbaatar, Arthur Szlam, Jason Weston, Rob Fergus, “End-To-End Memory Networks”, NIPS, 2015

Memory Network



Memory Network

每一次的hopping就像是通過一個hidden layer，只是每一層的layer參數都是一模一樣的（類似RNN）



新的paper

R-Net (by Microsoft)

Di-directional Attention Flow

Fully-Aware Fusion Network

Multiple-hop

ReasoNet，要做幾次hop是machine自己決定的

https://arxiv.org/abs/1609.05284

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

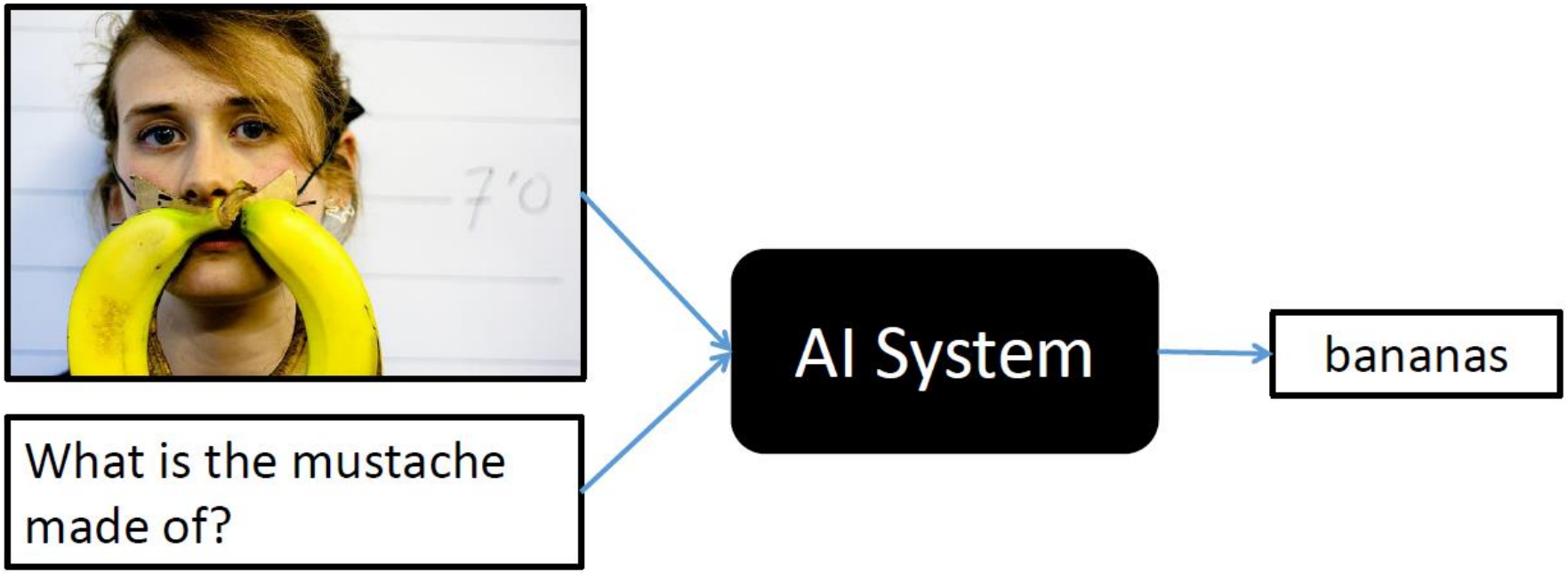
The position of reading head:

attention weight

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**			

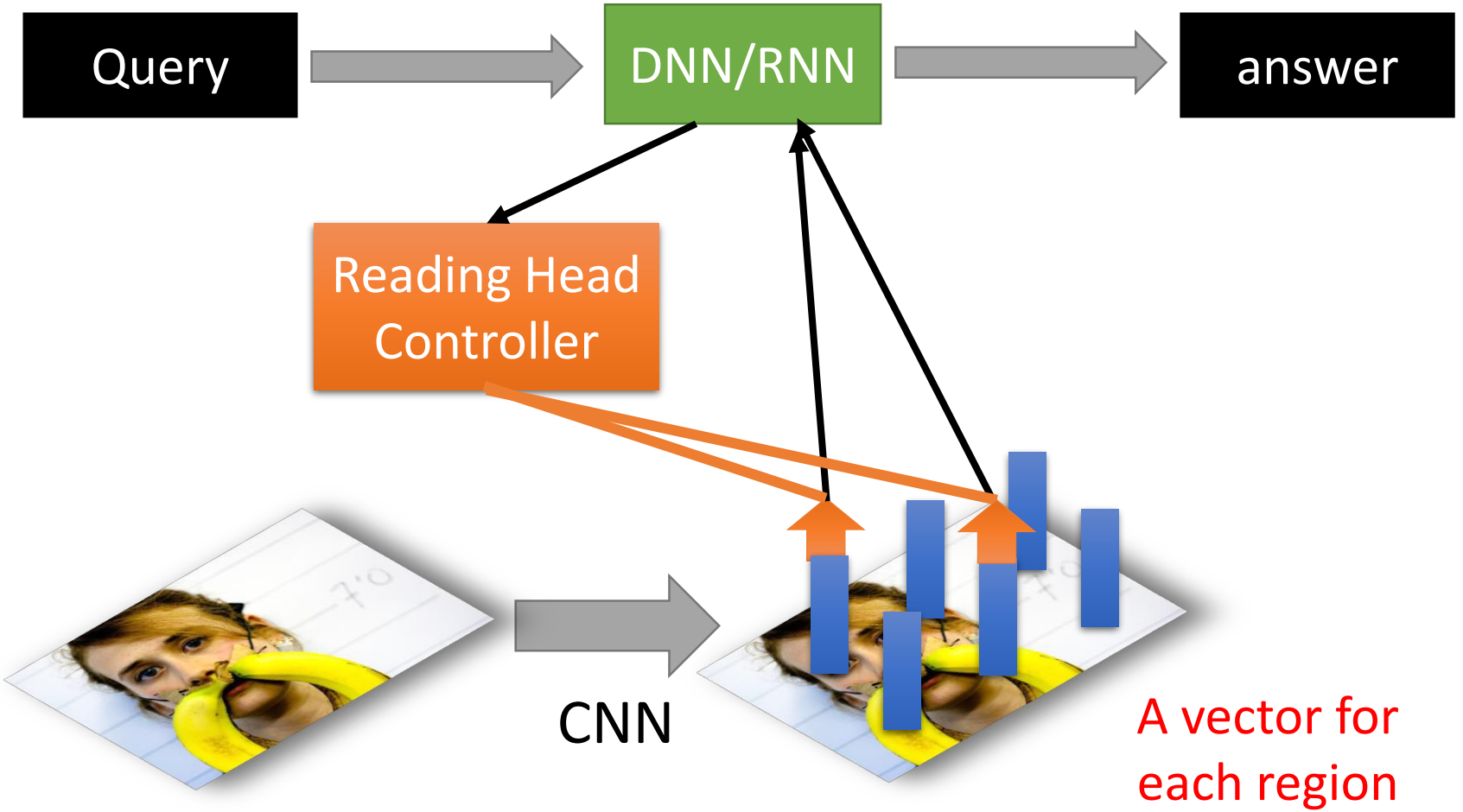
Keras has example:
https://github.com/fchollet/keras/blob/master/examples/babi_memnn.py

Visual Question Answering



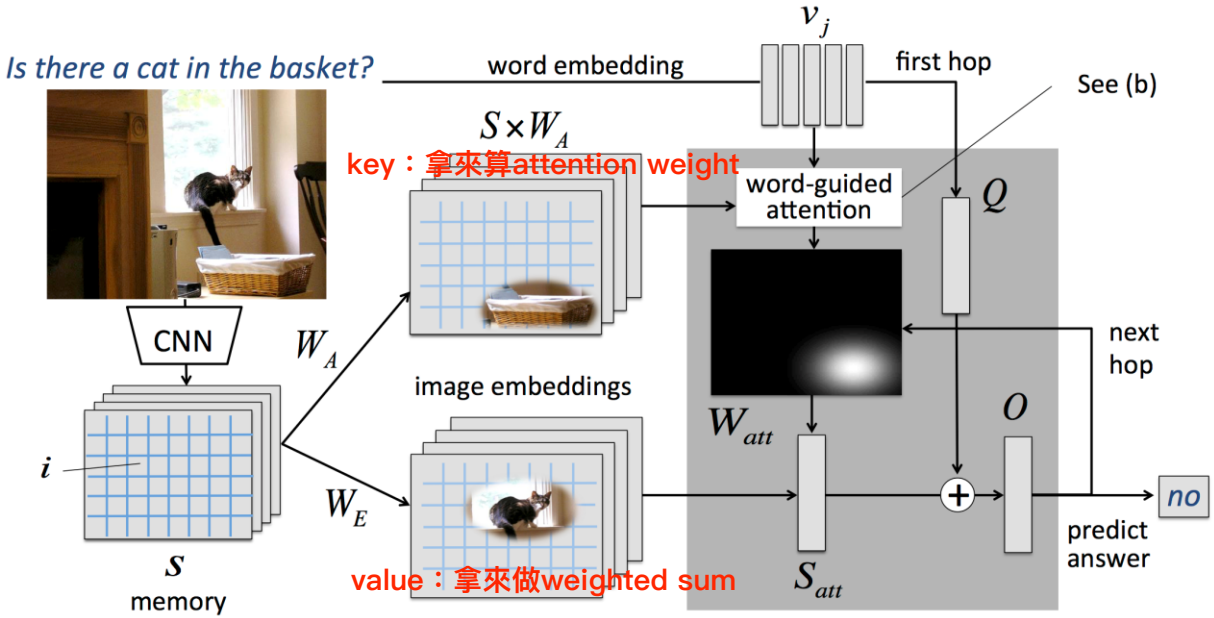
source: http://visualqa.org/

Visual Question Answering

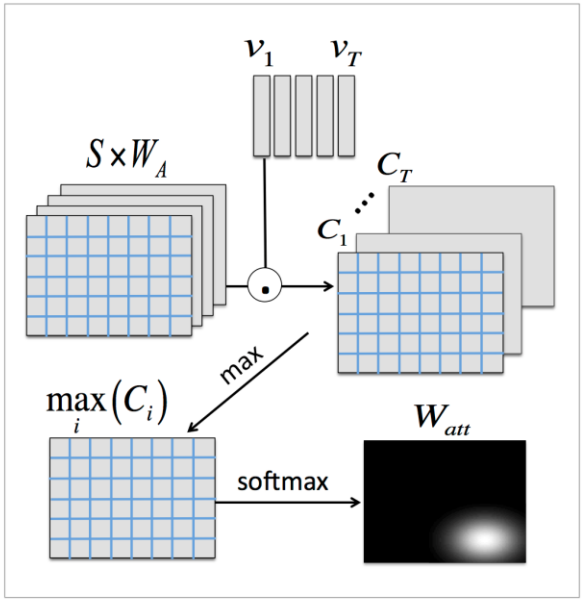


Visual Question Answering

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



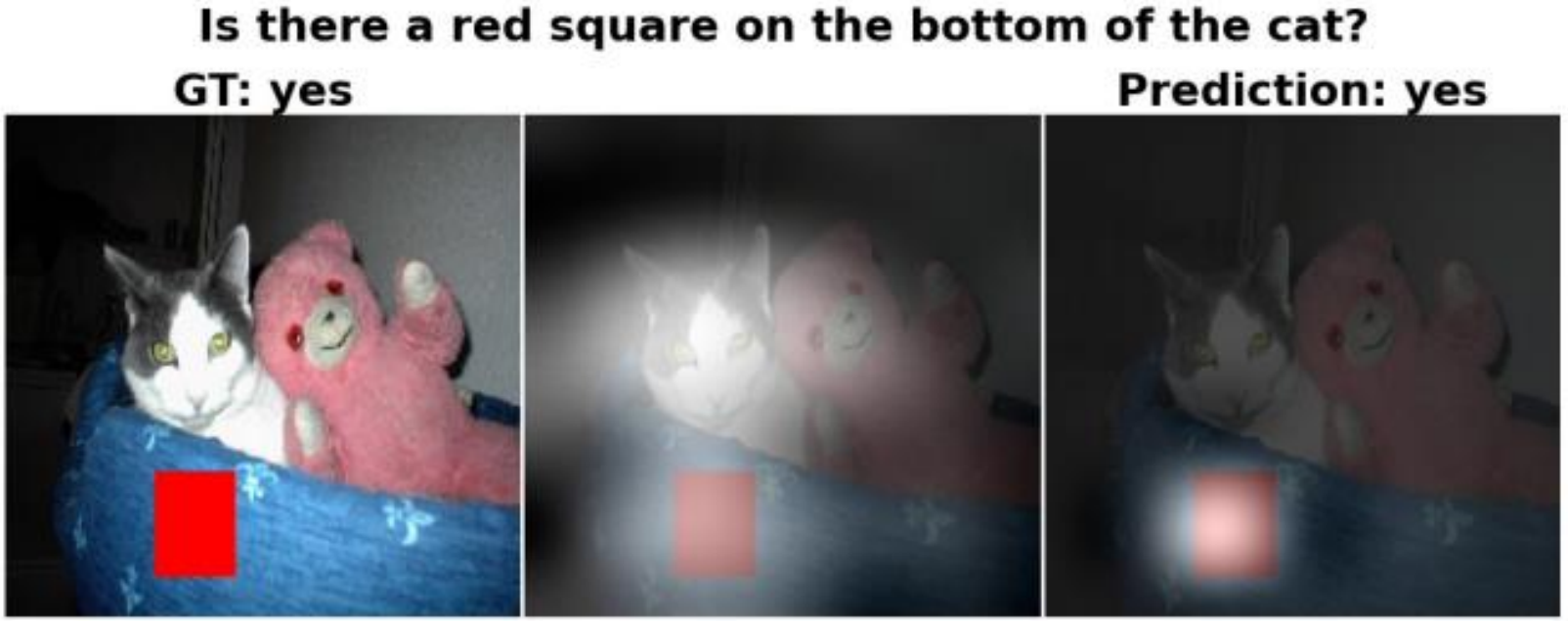
(a) Overview



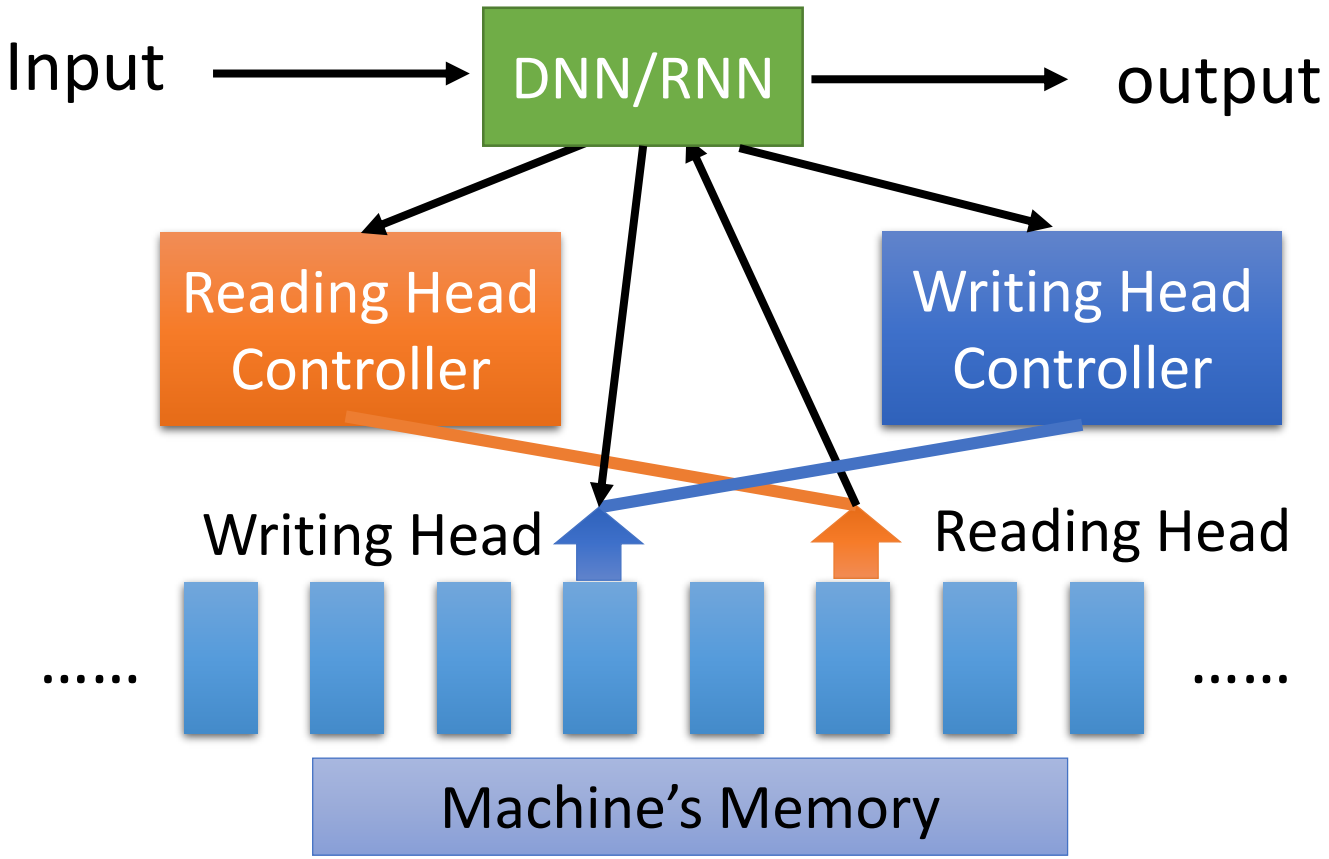
(b) Word-guided attention

Visual Question Answering

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



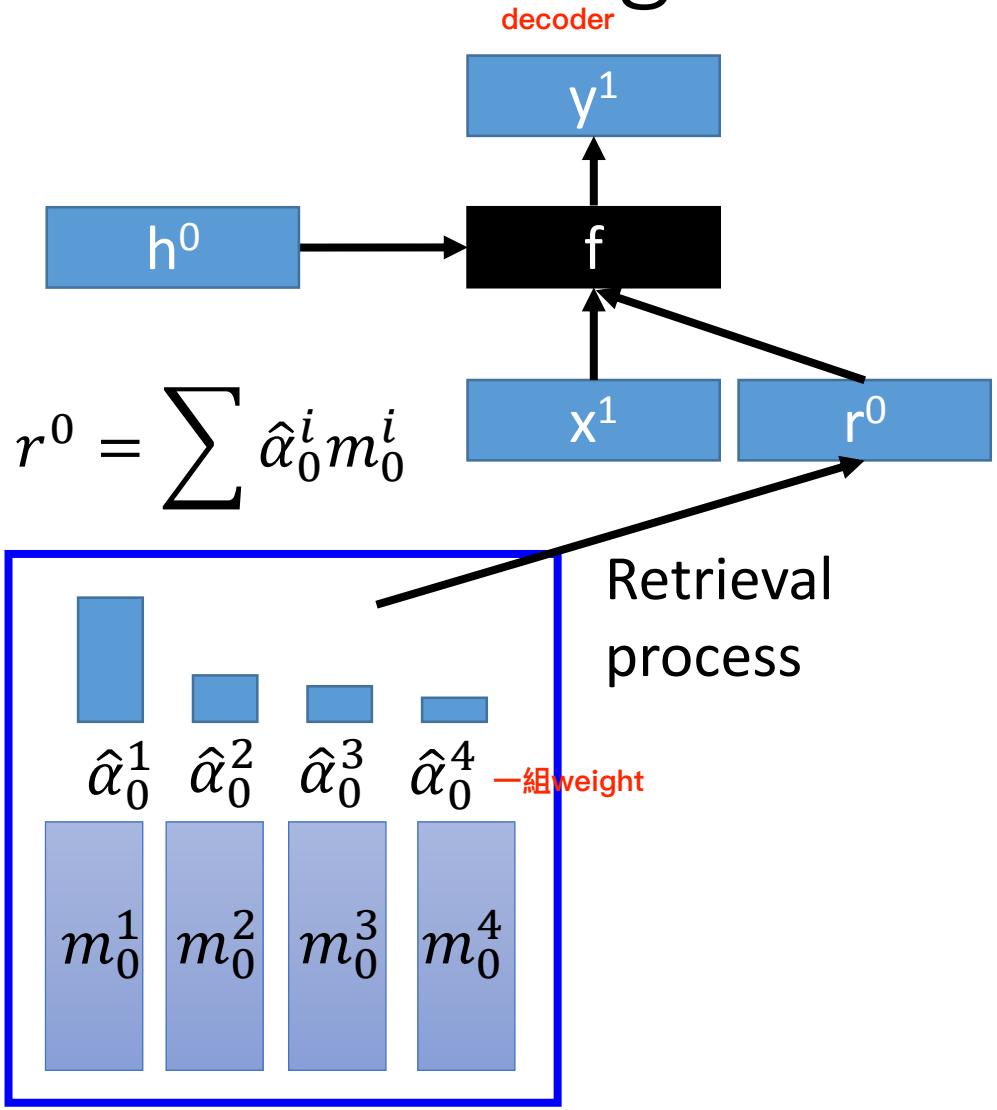
External Memory v2



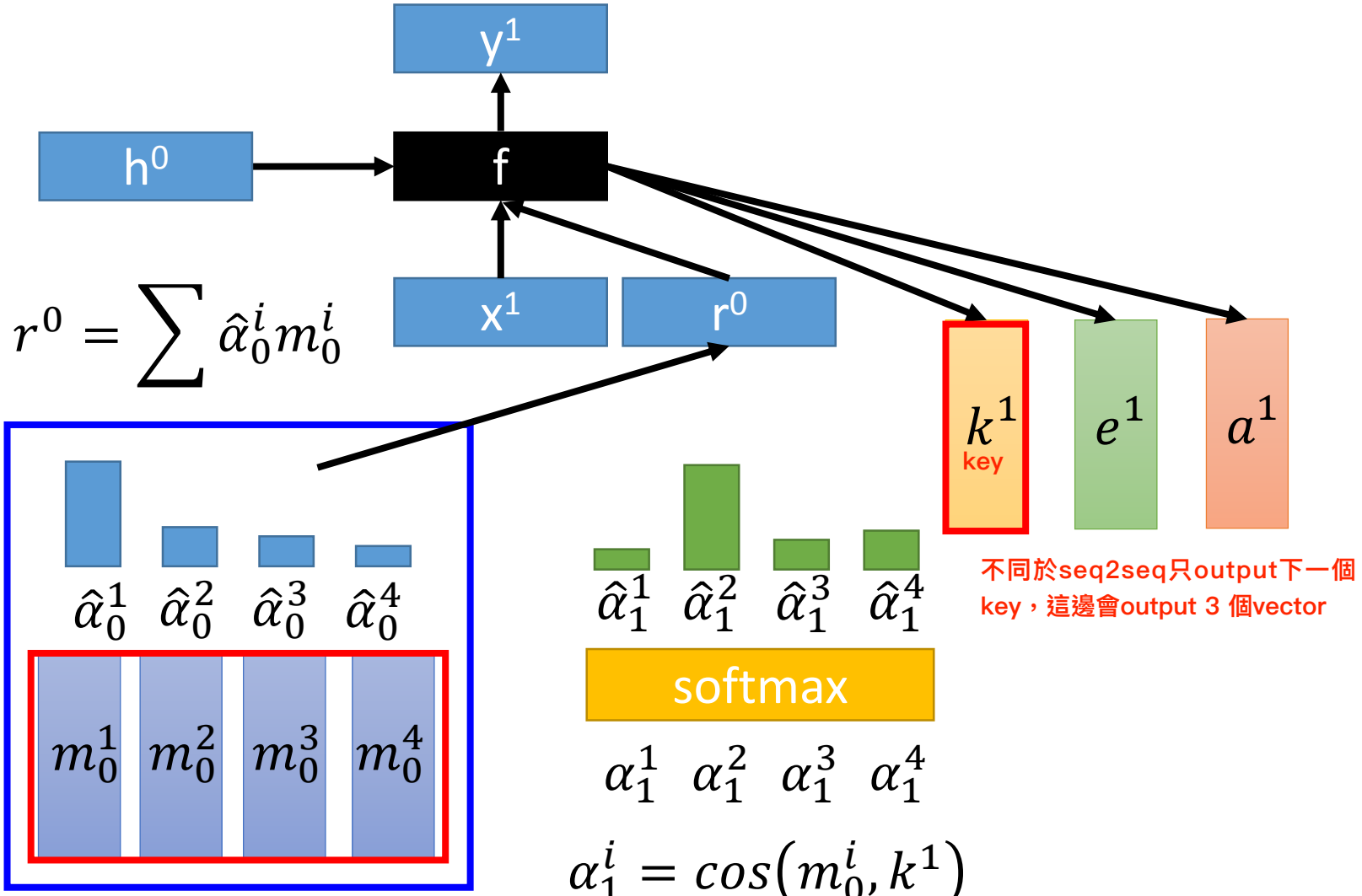
Neural Turing Machine

不只去讀machine mem.還可以去改寫

Neural Turing Machine



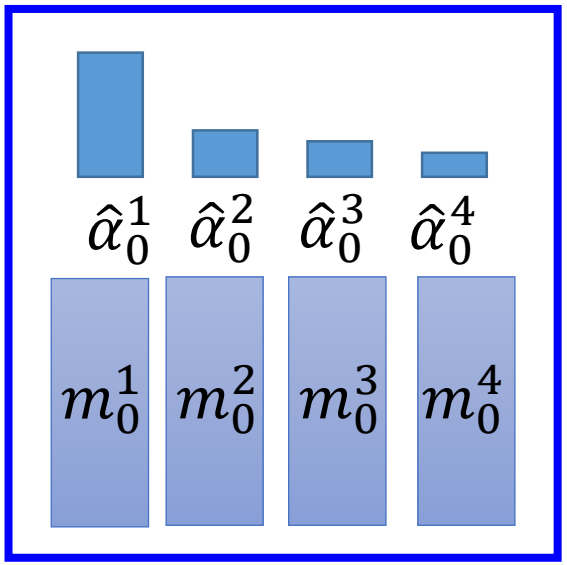
Neural Turing Machine

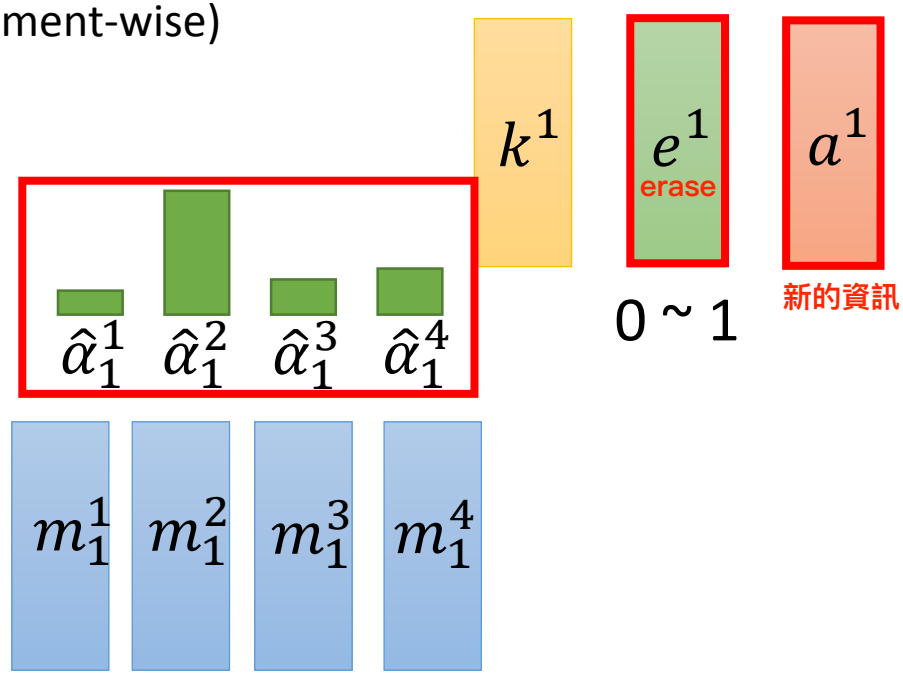


Neural Turing Machine

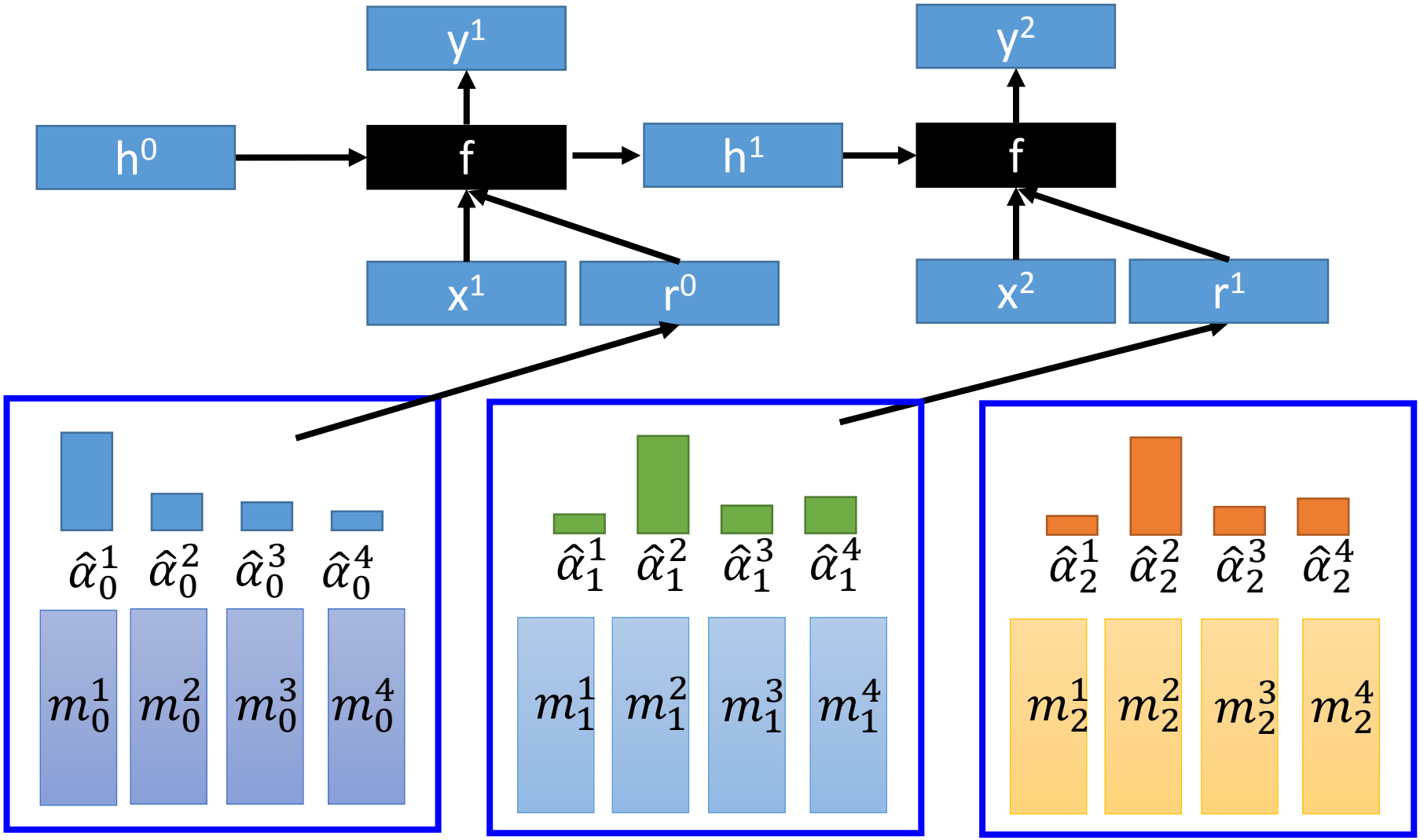
$$m_1^i = m_0^i - \hat{\alpha}_1^i \, e^1 \odot m_0^i + \hat{\alpha}_1^i \, a^1$$

(element-wise)

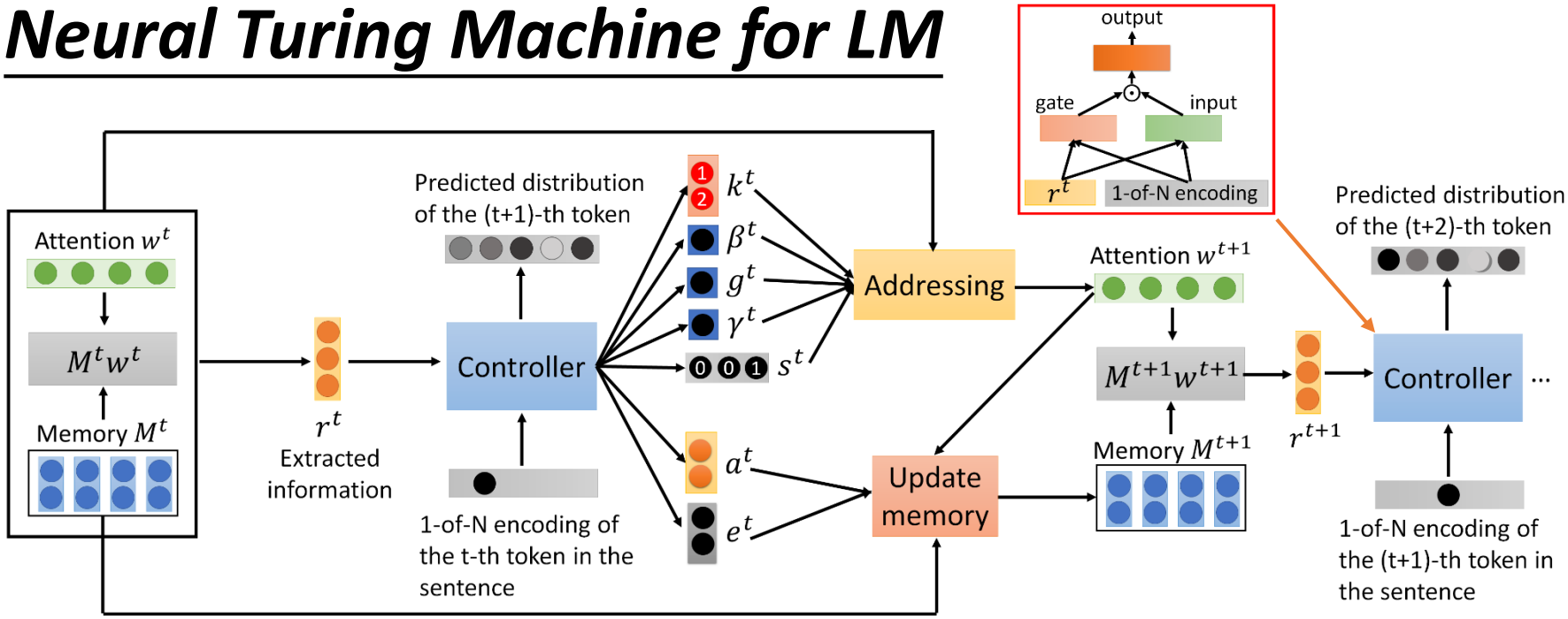


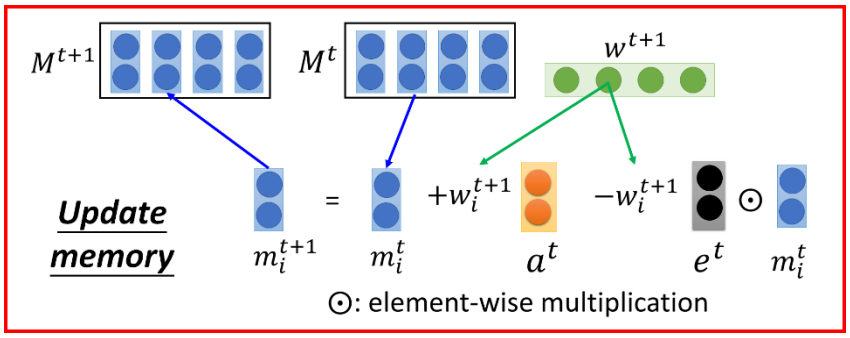


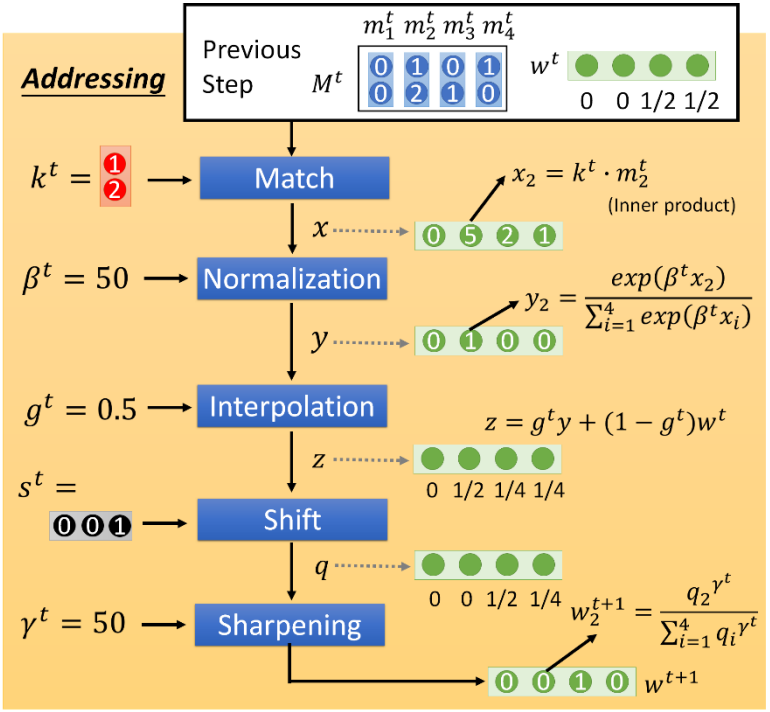
Neural Turing Machine



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