

## Lecture 16: Dynamic Memory Network

2018 年 8 月 5 日

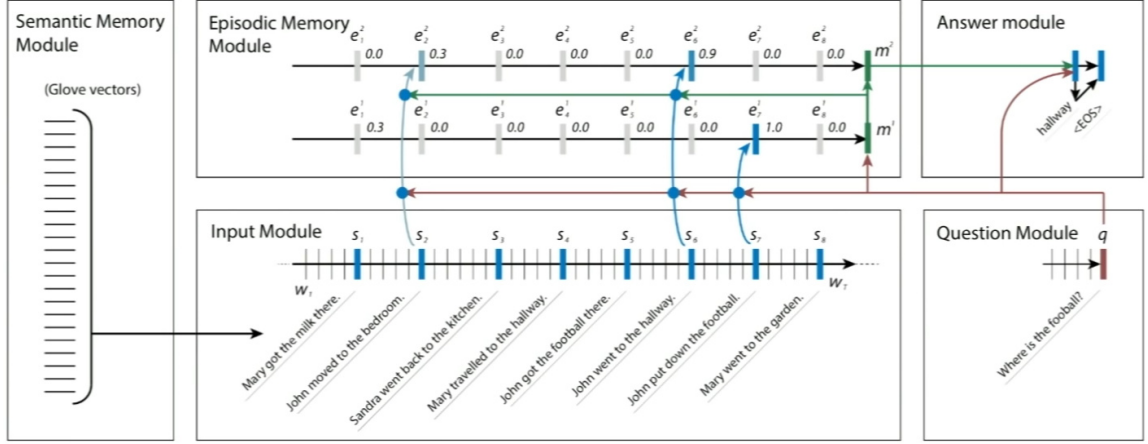
# 1 Introduction

其实 NLP 的许多问题都可以被转化为 QA 问题，比如情感分析，词性标注和机器翻译，非常厉害。

I: Mary walked to the bathroom.	I: Jane has a baby in Dresden.
I: Sandra went to the garden.	Q: What are the named entities?
I: Daniel went back to the garden.	A: Jane - person, Dresden - location
I: Sandra took the milk there.	I: Jane has a baby in Dresden.
Q: Where is the milk?	Q: What are the POS tags?
A: garden	A: NNP VBZ DT NN IN NNP .
I: Everybody is happy.	I: I think this model is incredible
Q: What's the sentiment?	Q: In French?
A: positive	A: Je pense que ce modèle est incroyable.

然而当时所有的 QA 模型都是针对具体的问题，首先没有人探讨过如何生成一个 General 的 model 去学习所有的任务；其次，Fully Joint multitask learning 共享所有的参数（所有任务参数设置相同），这非常困难。

## 2 Dynamic Memory network



大体思路：带着问题和阅读第一次的记忆，阅读第二次，不断加深记忆。

### 2.1 Input

输入 Embedding, encoder 就是传统的 GRU，文章以句子分割获得  $T_C$  个 fact representations  $c_t$ ，问题获得最后一个 hidden state  $q$ 。

### 2.2 Episodic Memory

先上公式，其中  $g$  是 Attention weight：

$$h_t^i = g_t^i GRU(c_t, h_{t-1}^i) + (1 - g_t^i) h_{t-1}^i$$

$$e_i = h_{T_C}^i$$

这一层我称之为 Attention Gated GRU，通过 Attention weight 控制当前 hidden state 来自上一步 GRU 和原文的比例。其中 Attention weight 取决于 candidate fact  $c_t$ ，a previous memory  $m^{i-1}$ ，and the question  $q$ 。

$$g_t^i = G(c_t, m^{i-1}, q)$$

$$G(c, m, q) = \text{softmax}(W^{(2)} \tanh(W^{(1)} z(c, m, q) + b^{(1)}) + b^{(2)})$$

$$z(c, m, q) = [c, m, q, c \circ q, c \circ m, |c - q|, |c - m|, c^T W^{(b)} q, c^T W^{(b)} m]$$

$z$  向量是一个超长拼接向量，经过一层 perceptron layer，一层线性层，然后 softmax 获得对不同部分关注权重。

### 2.3 Answer layer

简单的 GRU， $a_0$  初始化为最后一个 pass 的 memory  $a_0 = m^{T_M}$ 。

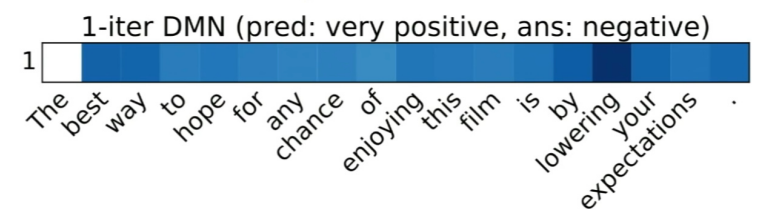
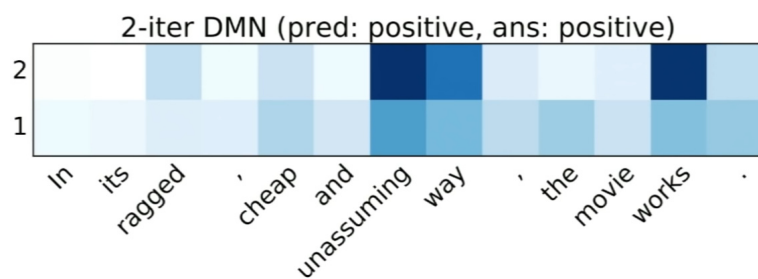
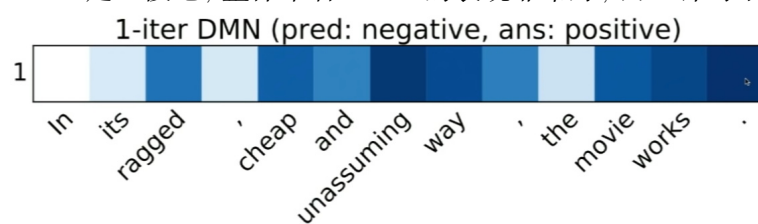
$$a_t = GRU([y_{t-1}, q], a_{t-1})$$

$$y_t = \text{softmax}(W^{(a)} a_t)$$

### 3 Performance

Max passes	task 3 three-facts	task 7 count	task 8 lists/sets	sentiment (fine grain)
0 pass	0	48.8	33.6	50.0
1 pass	0	48.8	54.0	51.5
2 pass	16.7	49.1	55.6	<b>52.1</b>
3 pass	64.7	83.4	83.4	50.1
5 pass	<b>95.2</b>	<b>96.9</b>	<b>96.5</b>	N/A

task3 是三段论，整体来看 DMN 的表现非常好，而且针对不同任务需选择不同次数的 pass。



不光文字可以使用 DMN，图像也可以，通过 CNN encode 图片作为输入，训练模型后可  
可视化 Attention part。



What is the main color on the bus ?



Answer: **blue**



What type of trees are in the background ?



Answer: **pine**



How many pink flags are there ?



Answer: **2**



Is this in the wild ?



Answer: **no**