Flexible and Creative Chinese Poetry Generation Using Neural Memory

Jiyuan Zhang, Feng Yang, Dong Wang, Yang Wang, Andrew Abel, Shiyue Zhang, Andi Zhang

Center for Speech and Language Technologies (CSLT), RIIT, Tsinghua University, China
http://cslt.riit.tsinghua.edu.cn

Introduction

(1) A 5-char quatrain(Rhythm & tone):

乐游原
Climbing the Paradise Mound
向晚意不适，(*ZZPZ)
As I was not in a good mood this evening round,
驱车登古原。(PPPZP)
I went by cart to climb the Ancient Paradise Mound.
夕阳无限好，(*PPZZ)
It is now nearing dusk,
只是近黄昏。(*ZZPP)
When the setting sun is infinitely fine, which is a must.

(2) A problem about traditional neural network(fluent and trivial, a lack of innovation):

For example:

竹	春雨
竹林小立松风雨	雨声细雪春初月
一点青山不可怜	一点青山不可怜
我爱清溪无数曲	天上晴阴无数事
绿阴未到水边船	东风送客又经年

(3) Our solution:
A memory-augmented neural network.

The two aspects of the effect of our proposed memory-augmented network:

--improve the innovation of poems
--generate poem with different sytle

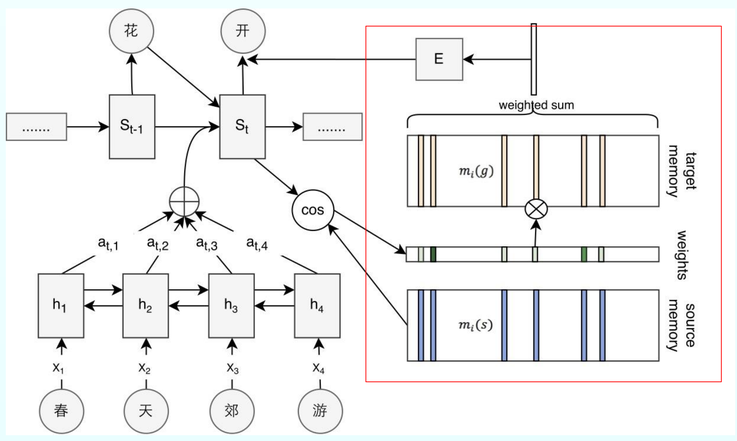
The memory-augmented neural netwok

It contains two components, the neural model component on the left, and the memory component on the right in the picture below.

Focus on the memory part in the red table:

- Source memory
  - $m_i(s) = f_d(x_{j-1}, s_{j-1}, 0)$
- Target memory
  - $m_i(g) = x_j$
- Weight:the memory elements are selected according to their fit to the present decoder status s_t, choose cosine distance to measure the fitting degree.
- The output of memory:

$$v_t = \sum_{i=1}^{K} \cos(s_t, m_i(s)) m_i(g))$$



The output of memory-augemneted neural network:

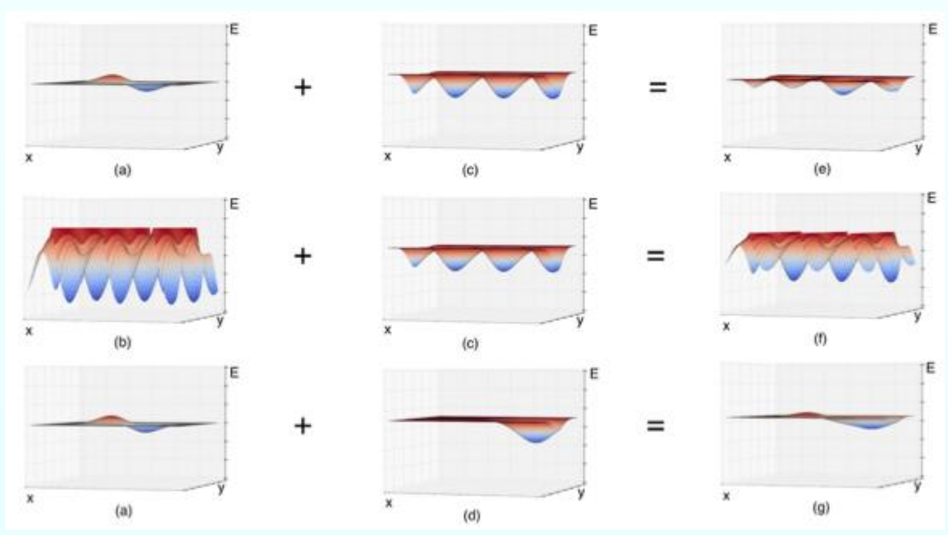
- $z_t = \sigma(s_t W + \beta v_t E)$

The β is not better than the manually-selected one.

The analysis of memory mechanism

Three scenarios where adding a memory may contribute:

- Promote innovation in an one-iteration neural network
- Regularize the innovation in an over-fitted neural model
- Generation of poems of different styles



Energy surface of Neural model | Energy surface of Memory | Energy surface of Combined Model

Experiment(innovation)

- Dataset
  - 500 quatrains randomly selected from our training corpus
- Two configurations
  - one is with a one-iteration model (C_1) and the other is with an overfitted model (C_∞).

	Preference Ratio				
	Compliance	Fluency	Theme Consistency	Aesthetic Innovation	Scenario Consistency
C_1 vs C_∞	0.59:0.41	0.68:0.32	0.70:0.30	0.68:0.32	0.69:0.31
C_1 vs C_1+Mem	0.41:0.59	0.36:0.64	0.37:0.63	0.33:0.67	0.43:0.57
C_∞ vs C_∞+Mem	0.40:0.60	0.26:0.74	0.32:0.68	0.30:0.70	0.36:0.64
C_1 vs C_∞+Mem	0.43:0.57	0.58:0.42	0.59:0.41	0.50:0.50	0.59:0.41

Experiment(style-transfer)

- Dataset
  - contains 300 quatrains with clear styles, including 100 pastoral, 100 battlefield and 100 romantic quatrains.
- General topic

	Probability			
Model	Pastoral	Battlefield	Romantic	Unclear
C_1 (Baseline)	0.09	0.04	0.18	0.69
C_1 + Pastoral Mem	0.94	0.00	0.06	0.00
C_1 + Battlefield Mem	0.05	0.93	0.00	0.02
C_1 + Romantic Mem	0.17	0.00	0.61	0.22

Method	Compliance	Fluency	Aesthetic Innovation	Scenario Consistence
C_1 (baseline)	4.10	3.01	2.53	2.94
C_1 + Pastoral Mem	4.07	3.00	3.07	3.17
C_1 + Battlefield Mem	3.82	2.63	2.60	2.95
C_1 + Romantic Mem	4.00	2.78	2.59	3.00
C_1 + All Mem	3.95	2.80	2.74	3.05

Conclusions

- The memory can encourage creative generation for regularly-trained models.
- The memory can encourage rule-compliance for overfitted models.
- The memory can modify the style of the generated poems in a flexible way.