AI 研习社 · 大讲堂

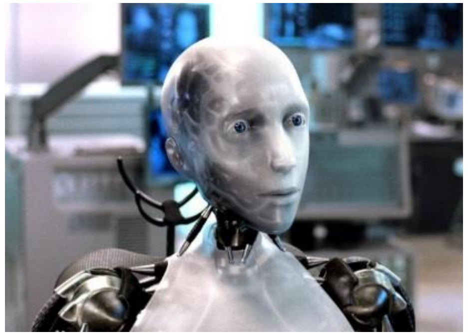
“九歌”——基于深度学习的中国古典诗歌自动生成系统

清华大学 计算机系
人工智能研究院
自然语言处理与社会人文计算实验室

矣晓沅

- 任务背景及“九歌”作诗系统简介
- 基于显著性上下文机制的诗歌生成
- 基于工作记忆模型的诗歌生成
- 基于互信息的无监督风格诗歌生成

任务背景及“九歌”作诗系统简介



- 机器翻译
- 专家系统
- 对话系统
- ⋮
- 诗歌生成

任务背景及“九歌”作诗系统简介

- 娱乐应用
- 诗词教育
- 文学研究
- 启发其他类型文本的生成
- …

任务背景及“九歌”作诗系统简介

- 娱乐应用
- 诗词教育
- 文学研究
- 启发其他类型文本的生成
- …

自动分析、理解和生成文学性文本的理想切入点！

任务背景及“九歌”作诗系统简介

第一阶段:
基于规则和模板的方法

- ASPERA system (Gerv´as, 2001)
- Haiku system (Wu et al., 2009).

……

任务背景及“九歌”作诗系统简介

第一阶段:
基于规则和模板的方法

- ASPERA system (Gerv´as, 2001)
- Haiku system (Wu et al., 2009).

……

第二阶段:
统计机器学习方法

- Genetic Algorithm (Manurung, 2003; Levy, 2001)
- Summarization Method (Yan et al., 2013)
- SMT (Jiang and Zhou, 2008; He et al., 2012)
- ……

任务背景及“九歌”作诗系统简介

第一阶段:
基于规则和模板的方法

- ASPERA system (Gerv´as, 2001)
- Haiku system (Wu et al., 2009).

……

第二阶段:
统计机器学习方法

- Genetic Algorithm (Manurung, 2003; Levy, 2001)
- Summarization Method (Yan et al., 2013)
- SMT (Jiang and Zhou, 2008; He et al., 2012)
- ……

第三阶段:
神经网络方法

- RNNPG (Zhang and Lapata, 2014)
- Polish (Yan, 2016)
- Planning (Wang et al., 2016)
- ……

任务背景及“九歌”作诗系统简介

九歌是清华大学自然语言处理与社会人文计算实验室研发的自动诗歌生成系统。九歌研究团队由孙茂松教授带领。

系统采用最新的深度学习技术，结合多个为诗歌生成专门设计的不同模型，基于超过30万首的诗歌进行训练学习。九歌系统能够产生集句诗、近体诗、藏头诗、宋词等不同体裁的诗歌。该系统在生成诗歌的质量上有显著提升。作为一款融合现代技术和中国古典文化的有趣应用，九歌在推动自然语言处理技术发展，弘扬中华优秀传统文化等方面都有所帮助。

任务背景及“九歌”作诗系统简介

- 绝句生成。系统根据用户输入的一个关键词生成五言或七言绝句。
- 藏头绝句生成。用户给定1~4个藏头字，用户生成对应的五言或七言藏头绝句。
- 风格绝句生成。用户给定一个关键词，系统生成指定风格(如边塞、闺怨、山水田园等)的绝句。
- 集句诗生成。用户给定一句首句，系统自动计算匹配剩余诗句，形成一首集句诗。
- 宋词生成。根据用户输入的多个关键词，生成指定词牌的宋词，已支持词牌超过20个。

任务背景及“九歌”作诗系统简介

- 2017年4月，九歌测试版上线；
- 2017年8月，九歌参加央视一套黄金时间大型科学挑战类节目《机智过人》，与三位当代优秀诗人同台竞技，比拼诗词创作能力；当期节目网络播放量达约1000万次；
- 2017年9月10日，九歌正式版上线；
- 2017年9月， 九歌受邀参加首届中国北京国际语言文化博览会；
- 2017年11月，九歌受邀参加《机智过人》人工智能年度盛典，与著名演员张凯丽同台表演小品；
- 2017年12月，九歌受邀在第七届吴文俊人工智能科学技术奖颁奖晚会进行作诗展示；
- 2017年12月，九歌参加央视三套“收获2017”节目录制；
- 2018年5月， 九歌受邀参加央视“五月的鲜花”五四青年晚会录制；
- 九歌获全国计算语言学学术会议（CCL）2017 最佳系统展示奖；
- 自2017年9月上线至今，九歌累计访问量达到100万次。

任务背景及“九歌”作诗系统简介

导师：
孙茂松教授

杨成
2014 级博士生

陈慧敏
2015级博士生

矣晓沅 2016级硕士生

梁健楠 2016级硕士生

郭志芃 2017级硕士生

参与系统开发的本科生：李文浩、杨宗瀚、魏钧宇

任务背景及“九歌”作诗系统简介

https://jiuge.thunlp.cn

基于显著性上下文机制的诗歌生成

Chinese Poetry Generation with a Salient-Clue Mechanism

Xiaoyuan Yi, Ruoyu Li, Maosong Sun

In CoNLL 2018

基于显著性上下文机制的诗歌生成

- Innovation (Zhang et al., 2017)
- Rhythmic Constraints (Ghazvininejad et al., 2016)
- Keywords Insertion (Wang et al., 2016)

…

基于显著性上下文机制的诗歌生成

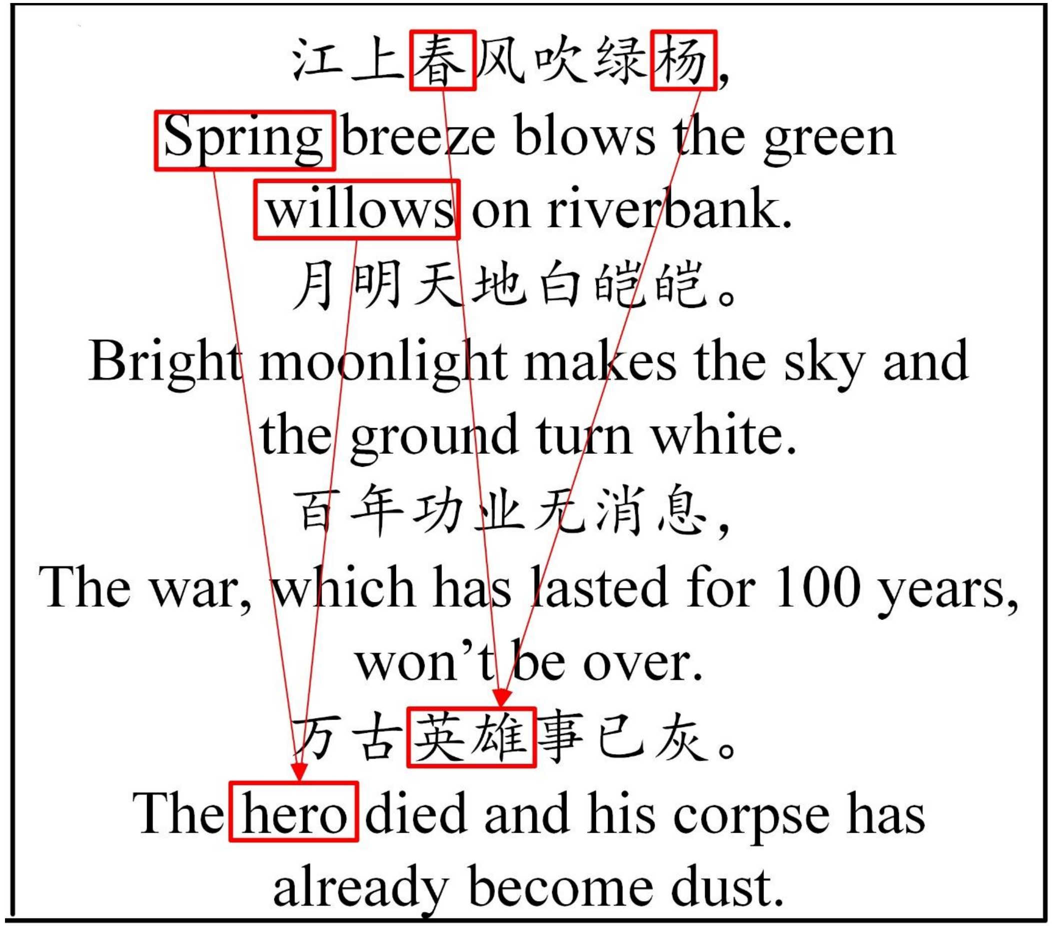
- Innovation (Zhang et al., 2017)
- Rhythmic Constraints (Ghazvininejad et al., 2016)
- Keywords Insertion (Wang et al., 2016)

…

Context Coherence？

AI 研习社 · 大讲堂

基于显著性上下文机制的诗歌生成

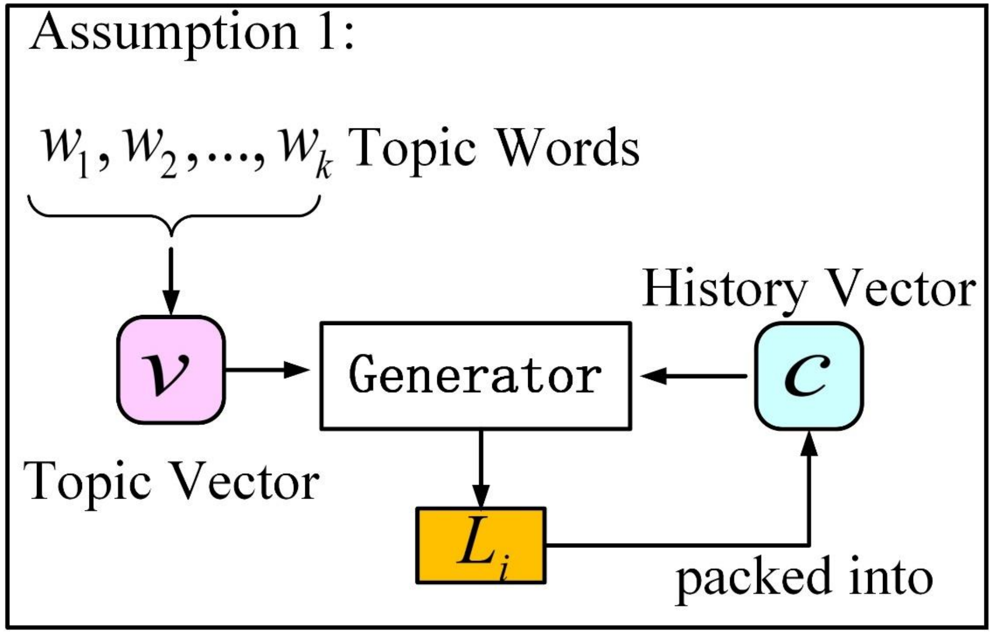


A poem generated by (Yan, 2016). The input keyword is *chun feng* (spring breeze). Red boxes and arrows show the inconsistency.

AI 研习社 · 大讲堂

基于显著性上下文机制的诗歌生成

(1) *Single* history vector (Zhang and Lapata 2014; Yan,2016)



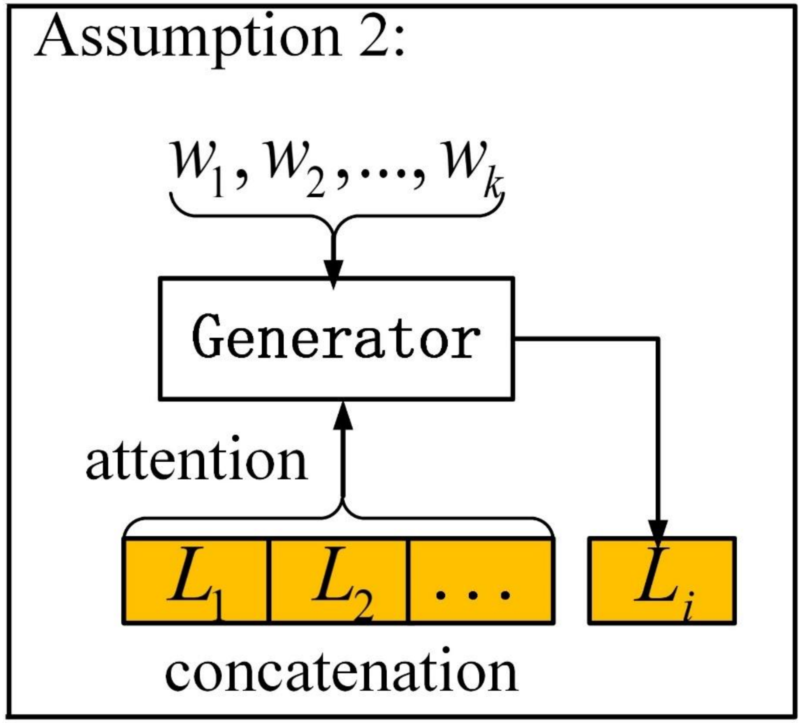
A graphical illustration of assumption 1.

- Insufficient capacity for maintaining the full history.
- Informative words and noises are mixed (e.g., stop words).
- Indiscriminate and inefficient use of the context.

Packing Full Context (PFC)

基于显著性上下文机制的诗歌生成

(2) The ability of exploring *unlimited* history. (Wang et al., 2016; Zhang et al., 2017)



A graphical illustration of assumption 2.

Keyword	The Preceding Text	Current Line
床	—	床前明月光
霜	床前明月光	疑是地上霜
明月	床前明月光; 疑是地上霜	举头望明月
故乡	床前明月光; 疑是地上霜; 举头望明月	低头思故乡

Training triples extracted from a quatrain in (Wang et al., 2016)

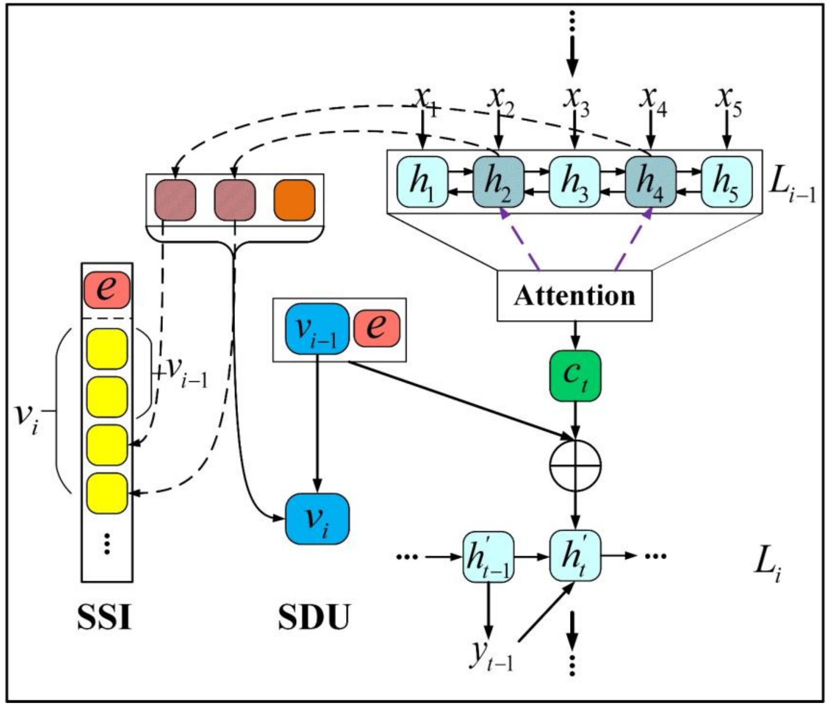
Too long input/output sequences

The performance of seq2seq model degrades, even with an attention mechanism.

N Lines to one line (nLto1L)

AI 研习社 · 大讲堂

基于显著性上下文机制的诗歌生成



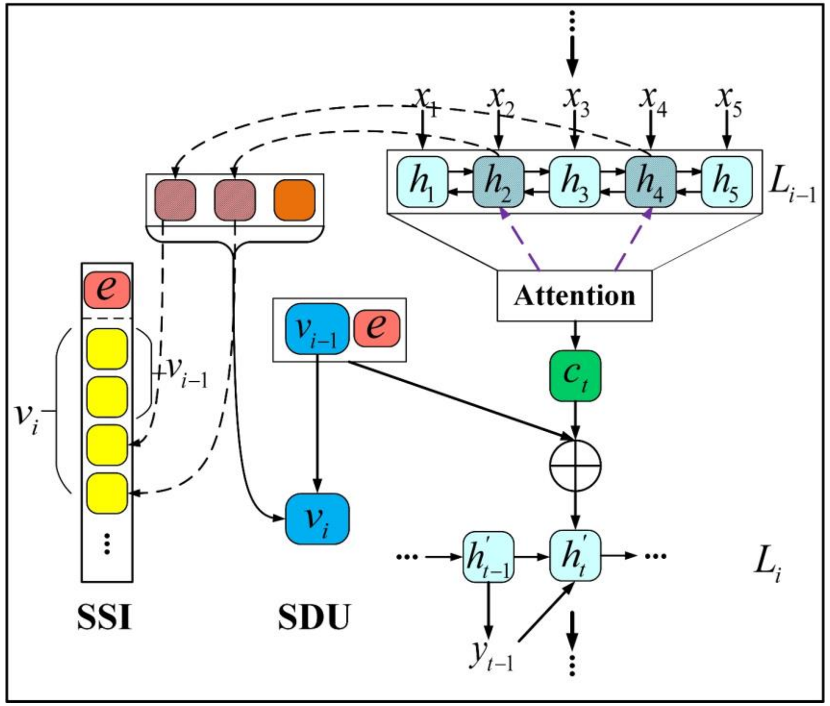
A graphical illustration of the proposed Salient-Clue mechanism. v_i is the salient-clue vector and e is the extension vector. We design two strategies for updating the salient clue. SDU: v_i is kept at the same size; SSI: the size of v_i increases during the generation process.

Design Philosophy: ignore the uninformative parts (e.g., stop words) and use some salient characters in context to represent the full context and form a salient clue, which is used to guide the generation process.

1. For each generated line: selects salient and informative characters to form the salient-clue.
2. When generating each line: utilizes the salient-clue.

AI 研习社 · 大讲堂

基于显著性上下文机制的诗歌生成



A graphical illustration of the proposed Salient-Clue mechanism. v_i is the salient-clue vector and e is the extension vector. We design two strategies for updating the salient clue. SDU: v_i is kept at the same size; SSI: the size of v_i increases during the generation process.

Basic framework: Bidirectional LSTM Encoder-Decoder with attention mechanism (Bahdanau et al., 2015)

$$h'_t = LSTM(h'_{t-1}, emb(y_{t-1}), c_t), \quad (1)$$

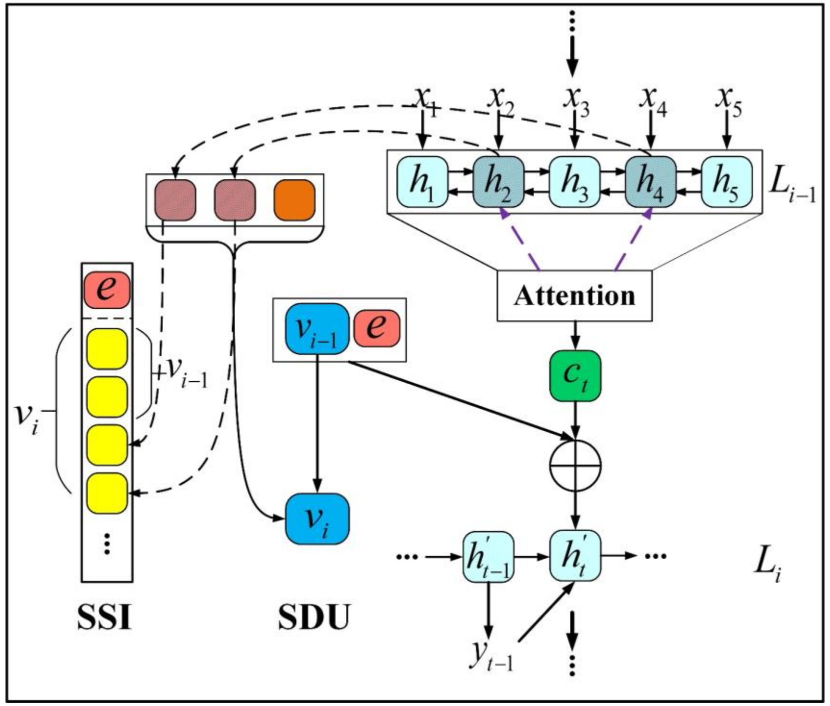
$$p(y_t|y_{1:t-1}, L_{1:i-1}) = g(h'_t, emb(y_{t-1}), c_t, v), \quad (2)$$

$$r_j = [(\mathbf{w}_{out} * A) \odot \mathbf{w}_{in}]_j, \quad (3)$$

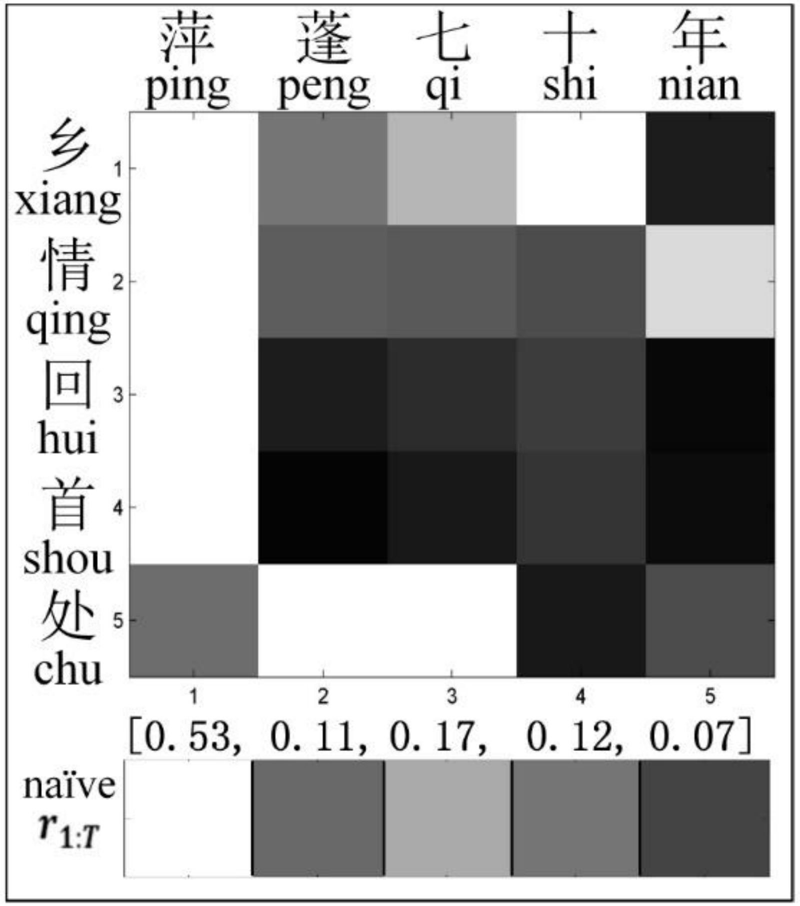
the attention alignment matrix in the attention mechanism

AI 研习社 · 大讲堂

基于显著性上下文机制的诗歌生成

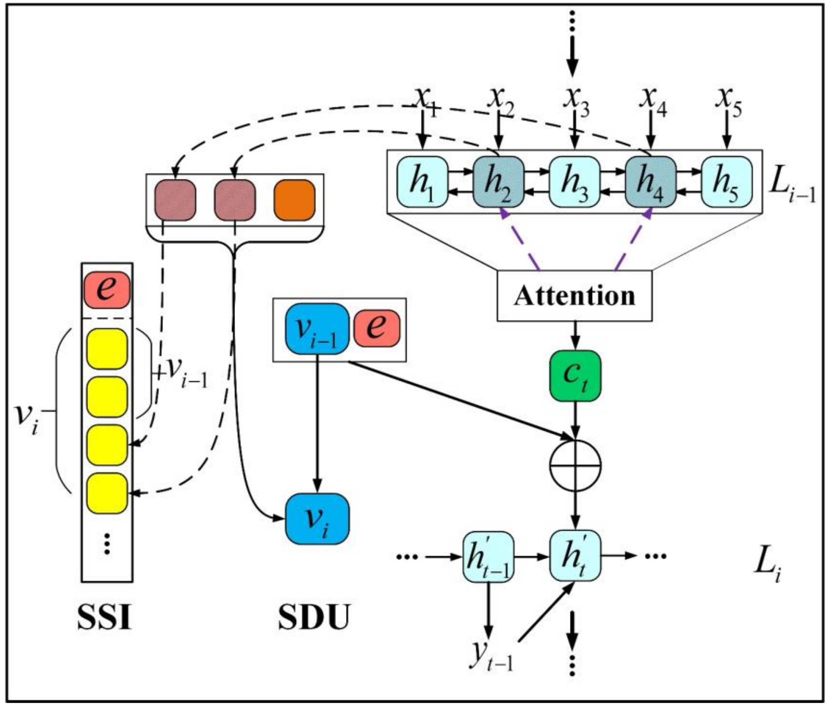


A graphical illustration of the proposed Salient-Clue mechanism. v_i is the salient-clue vector and e is the extension vector. We design two strategies for updating the salient clue. SDU: v_i is kept at the same size; SSI: the size of v_i increases during the generation process.

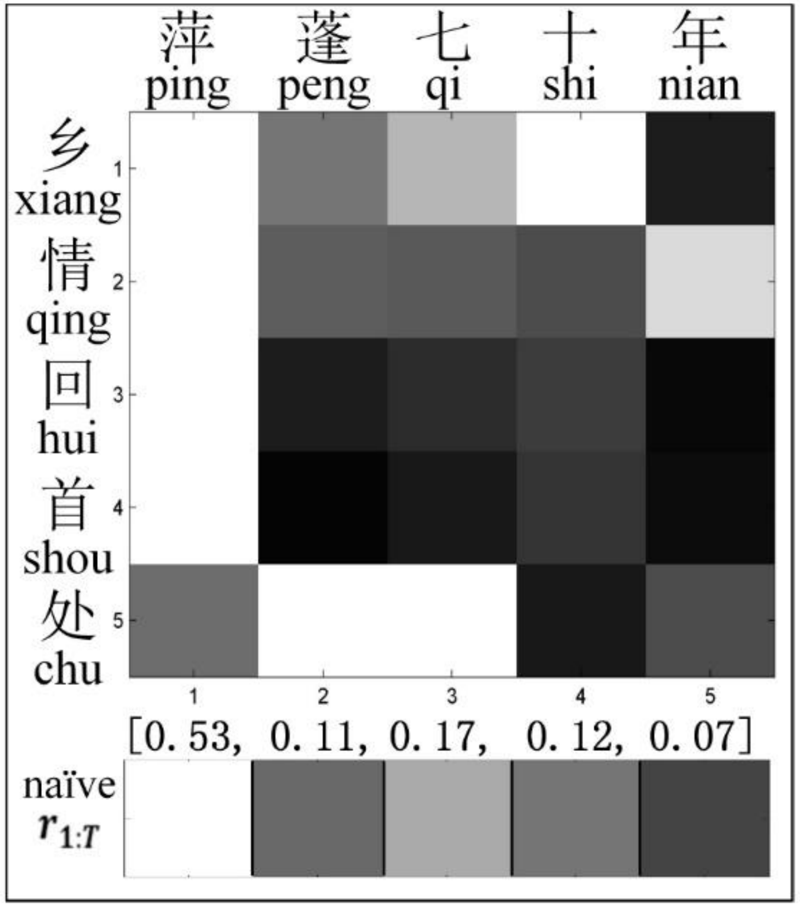


An example of calculating the saliency score of each character (in the x-axis) from the attention matrix (0:black, 1:white), in the naive Salient-Clue. The scores are normalized to interval [0,1] here.

基于显著性上下文机制的诗歌生成

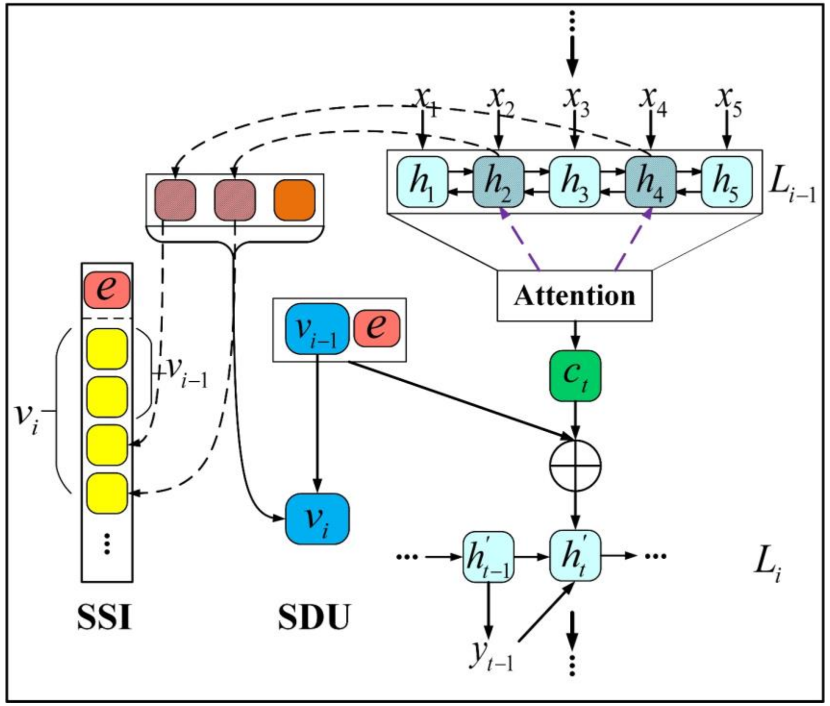


A graphical illustration of the proposed Salient-Clue mechanism. v_i is the salient-clue vector and e is the extension vector. We design two strategies for updating the salient clue. SDU: v_i is kept at the same size; SSI: the size of v_i increases during the generation process.



An example of calculating the saliency score of each character (in the x-axis) from the attention matrix (0:black, 1:white), in the naive Salient-Clue. The scores are normalized to interval [0,1] here.

基于显著性上下文机制的诗歌生成



A graphical illustration of the proposed Salient-Clue mechanism. v_i is the salient-clue vector and e is the extension vector. We design two strategies for updating the salient clue. SDU: v_i is kept at the same size; SSI: the size of v_i increases during the generation process.

Algorithm 1 Saliency Selection Algorithm

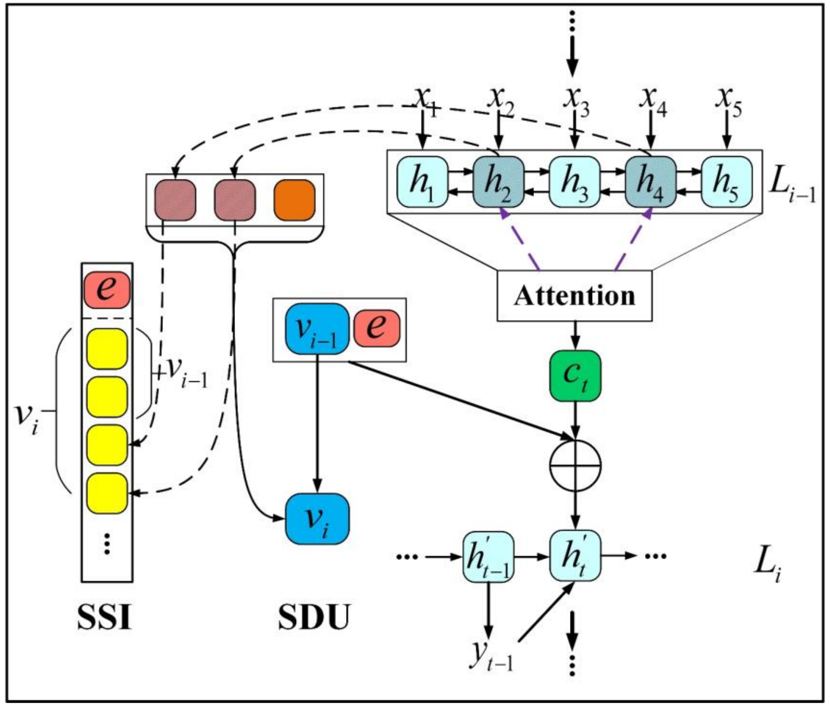
Inputs: The saliency scores of characters in the preceding line, $r_{1:T}$; K;

Outputs: The number of finally selected salient characters, N; The indices of selected characters in the preceding line, $m_{1:N}$;

Calculate the mean value of $r_{1:T}$, avg;
Calculate the standard deviation of $r_{1:T}$, std;
Get sorted indices $i_{1:T}$ in descending order of $r_{1:T}$;
$k = 1$; $val = avg + 0.5 * std$;
while $(r_{i_k} \geq val)$ $\quad and \quad$ $(k \leq K)$ **do**
  $m_k = i_k$; $val = val * 0.618$ (the golden ratio); $k = k + 1$;
end while
$N = k - 1$;
return $N, m_{1:N}$;

$$N, m_{1:N} = SSal(r_{1:T}, K), \quad (4)$$

基于显著性上下文机制的诗歌生成



A graphical illustration of the proposed Salient-Clue mechanism. v_i is the salient-clue vector and e is the extension vector. We design two strategies for updating the salient clue. SDU: v_i is kept at the same size; SSI: the size of v_i increases during the generation process.

Saliency Dynamic Update (SDU) V.S. PFC

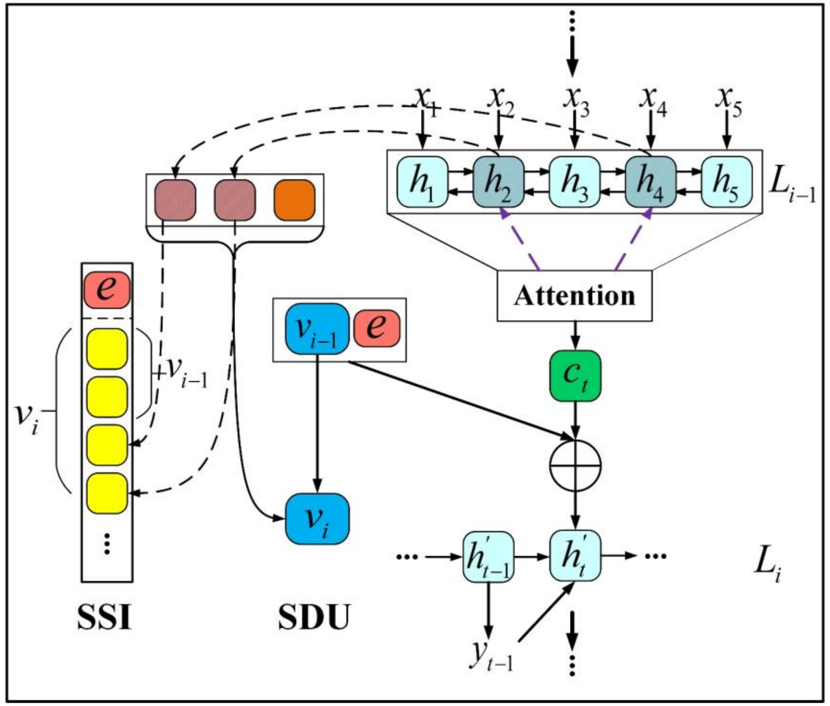
$$s = \frac{\sum_{k=1}^{N} r_{m_k} * h_{m_k}}{\sum_{k'}^{N} r_{m_{k'}}}, \tag{5}$$

$$v_i = \sigma(v_{i-1}, s), v_0 = \vec{0}, \tag{6}$$

Saliency Sensitive Identity (SSI) V.S nLto1L

$$v_i = [v_{i-1}; h_{m_1}; ...; h_{m_N}], \tag{7}$$

基于显著性上下文机制的诗歌生成



A graphical illustration of the proposed Salient-Clue mechanism. v_i is the salient-clue vector and e is the extension vector. We design two strategies for updating the salient clue. SDU: v_i is kept at the same size; SSI: the size of v_i increases during the generation process.

Extensions of Salient-Clue

$$p(y_t|y_{1:t-1}, L_{1:i-1}) = g(h_t', emb(y_{t-1}), c_t, [v_{i-1}; e]). \quad (8)$$

- Intent Salient-Clue

Feed the keyword into Encoder, then vector e is calculated by a non-linear transformation of the average of their hidden states.

- Style Salient-Clue

Simply use a style embedding as the vector e. Use LDA to train the whole corpus. For three main styles, Pastoral, Battlefield and Romantic, find the corresponding topics manually. Then all poems are labeled by LDA inference.

AI 研习社 · 大讲堂

基于显著性上下文机制的诗歌生成

Experimental Results

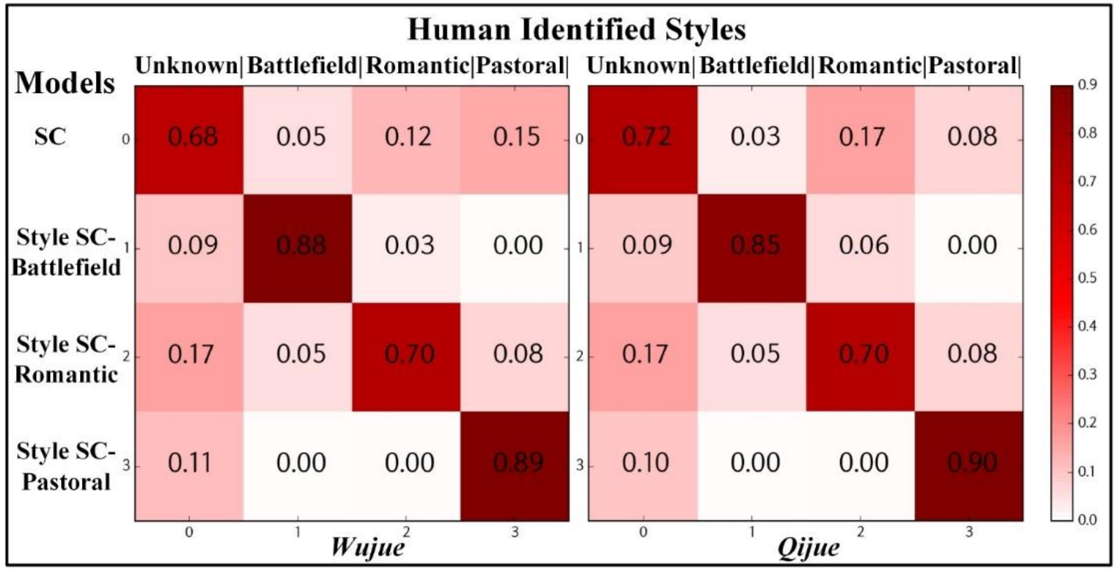
	Models	*Wujue*	*Qijue*
Different Models	Planning	0.460	0.554
	iPoet	0.502	0.591
	seq2seqPG	0.466	0.620
	SC	**0.532**	**0.669**
Different Strategies of SC	naive-TopK-SDU	0.442	0.608
	naive-SSal-SDU	0.471	0.610
	tfidf-SSal-SDU	**0.533**	0.648
	tfidf-SSal-SSI	0.530	0.667
	tfidf-SSal-SSI-intent	0.532	**0.669**

Table 1: BLEU evaluation results. The scores are calculated by the multi-bleu.perl script.

Models	**Fluency**		**Coherence**		**Meaningfulness**		**Poeticness**		**Entirety**	
	Wujue	*Qijue*	*Wujue*	*Qijue*	*Wujue*	*Qijue*	*Wujue*	*Qijue*	*Wujue*	*Qijue*
Planning	2.56	2.84	2.50	2.64	2.49	2.64	2.59	2.88	2.39	2.66
iPoet	3.13	3.45	2.89	2.91	2.60	2.80	2.79	3.05	2.54	2.85
seq2seqPG	3.54	3.65	3.31	3.16	3.15	3.01	3.26	3.29	3.06	3.08
SC	4.01^{**}	4.04^{**}	3.85^{**}	3.86^{**}	3.55^{**}	3.63^{**}	$\mathbf{3.74}^{**}$	$\mathbf{3.69}^{*}$	3.63^{**}	$\mathbf{3.70}^{**}$
Style-SC	$\mathbf{4.03}^{**}$	$\mathbf{4.16}^{**}$	$\mathbf{3.90}^{**}$	$\mathbf{4.01}^{**}$	$\mathbf{3.68}^{**}$	$\mathbf{3.75}^{**}$	3.61^{*}	3.68^{*}	$\mathbf{3.65}^{**}$	$\mathbf{3.70}^{**}$
Human	4.09	4.43	3.90	4.33^{+}	3.94	4.35^{++}	3.83	4.24^{++}	3.81	4.24^{++}

Table 2: Human evaluation results. Diacritics * ($p < 0.05$) and ** ($p < 0.01$) indicates SC models significantly outperform the three baselines; + ($p < 0.05$) and ++ ($p < 0.01$) indicates Human is significantly better than all the five models. The Intraclass Correlation Coefficient of the four groups of scores is 0.596, which indicates an acceptable inter-annotator agreement.

基于显著性上下文机制的诗歌生成



Style control evaluation results. The values are ratios that generated poems are identified as different styles by human experts.

Models	***Wujue***	***Qijue***
Random	0.271	0.247
tf-idf	0.417	0.378
naive-TopK SC	0.347	0.415
naive-SSal SC	0.431	0.441
tfidf-SSal SC	**0.525**	**0.461**

Table 3: Saliency selection results. Random: randomly select K characters for three times and use the average Jaccard values. tf-idf: directly select K characters in terms of tf-idf, without SC.

基于显著性上下文机制的诗歌生成

濛濛烟雨暗江天，一片青山水自妍。为问故人何处所，夕阳依旧绿杨边。
Misty rain darkens the sky.
The green mountains and the rivers are so beautiful.
Where is my old friend?
I know, he is standing beside green willows at sunset.
(a)

茫茫烟雨不胜愁，千古英雄泪眼流。一代兴亡今已矣，百年遗恨付沧洲。
There is endless sadness in the misty rain.
Tears at last welled from the hero's eyes.
The war has ended now,
I leave for my hometown with a century-old hate.
(b)

濛濛烟雨湿衣裳，满地春风十里香。一曲琵琶无处觅，隔帘明月是何郎。
The misty rain wets my dress.
Spring wind brings the fragrance of flowers.
Playing pi-pa (a kind of Chinese musical instrument),
the boy out of the curtain, I want to know who you are.
(c)

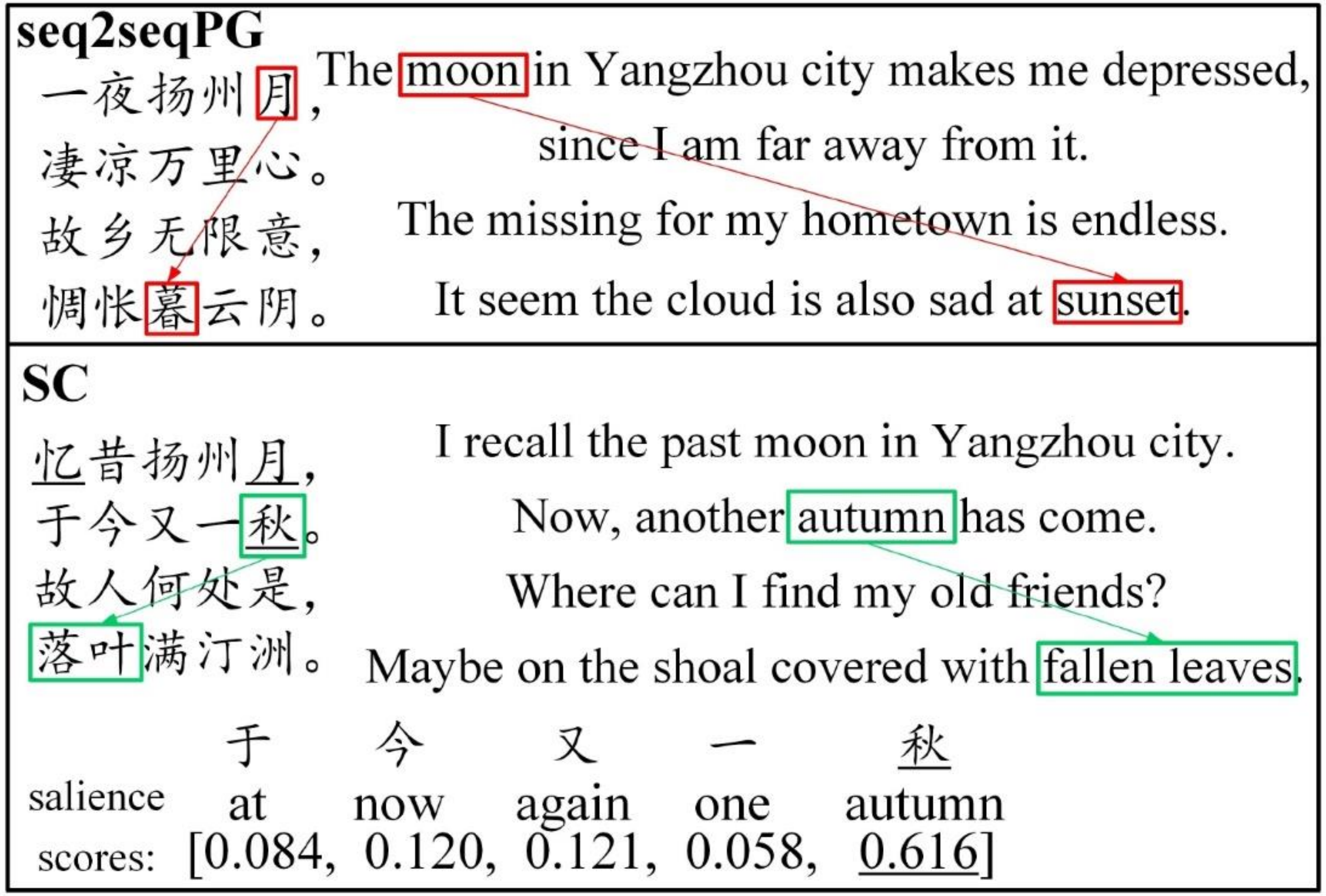
濛濛烟雨一番新，野水平畴绿涨津。何处渔郎归去晚，数家沽酒卖鱼人。
The misty rain refreshes the world.
Beside the cropland, river rises up to the wharf.
At sunset, where is the fisherman going?
The fish shop and the bar!
(d)

Four *Qijue*s generated with the same keyword “烟雨” (misty rain) as input. (a) non-style by SC. (b) Battlefield style by Style-SC. (c) Romantic style by Style-SC. (d) Pastoral style by Style-SC.

AI 研习社 · 大讲堂

基于显著性上下文机制的诗歌生成

Case Study



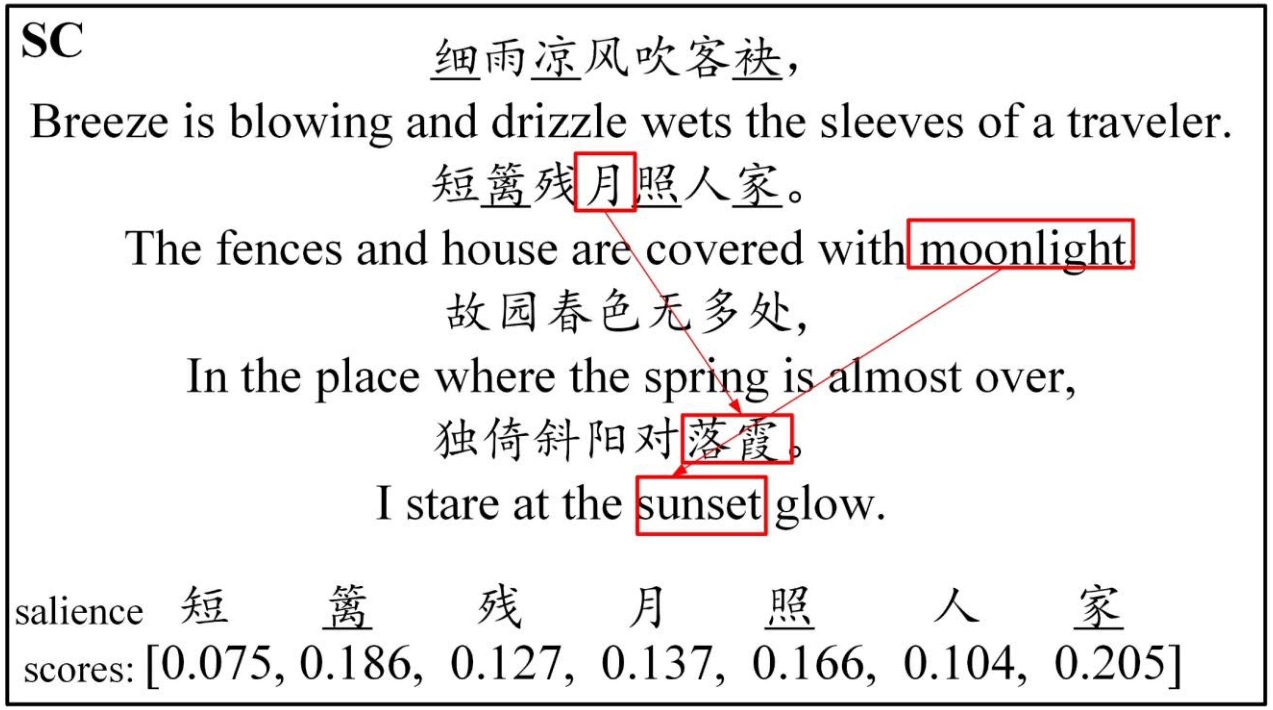
Two *Wujue*s generated with the same input. Green boxes and arrows show consistencies, and the red ones show inconsistencies. Automatically selected charcters by SC are underlined.

AI 研习社 · 大讲堂

基于显著性上下文机制的诗歌生成

Weakness

1. Inflexible selection method

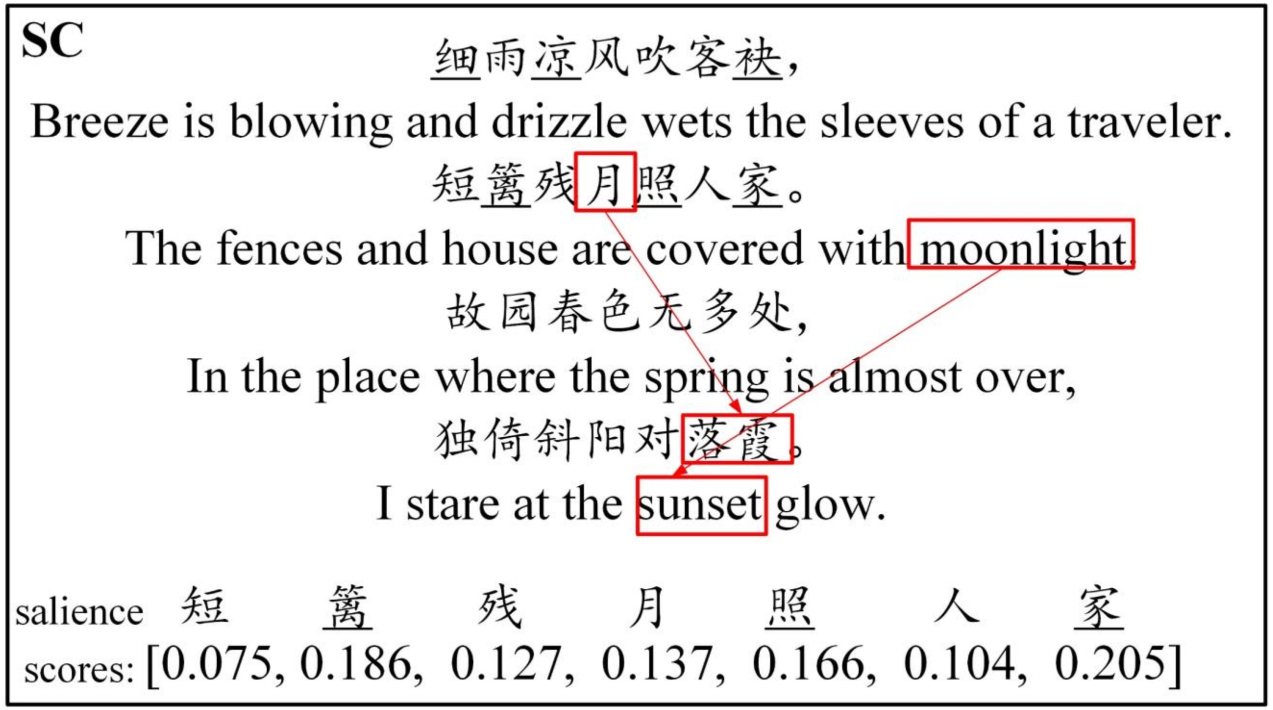


A negative example. A *Qijue* generated by our SC model. Red box and arrow show the inconsistency.

基于显著性上下文机制的诗歌生成

Weakness

1. Inflexible selection method



A negative example. A *Qijue* generated by our SC model. Red box and arrow show the inconsistency.

2. Supervised learning for style transfer

- Expensive labeled data
- Losing some fluency and Poeticness

基于工作记忆模型的诗歌生成

Chinese Poetry Generation with a Working Memory Model

Xiaoyuan Yi, Maosong Sun, Ruoyu Li, Zonghan Yang

In IJCAI 2018

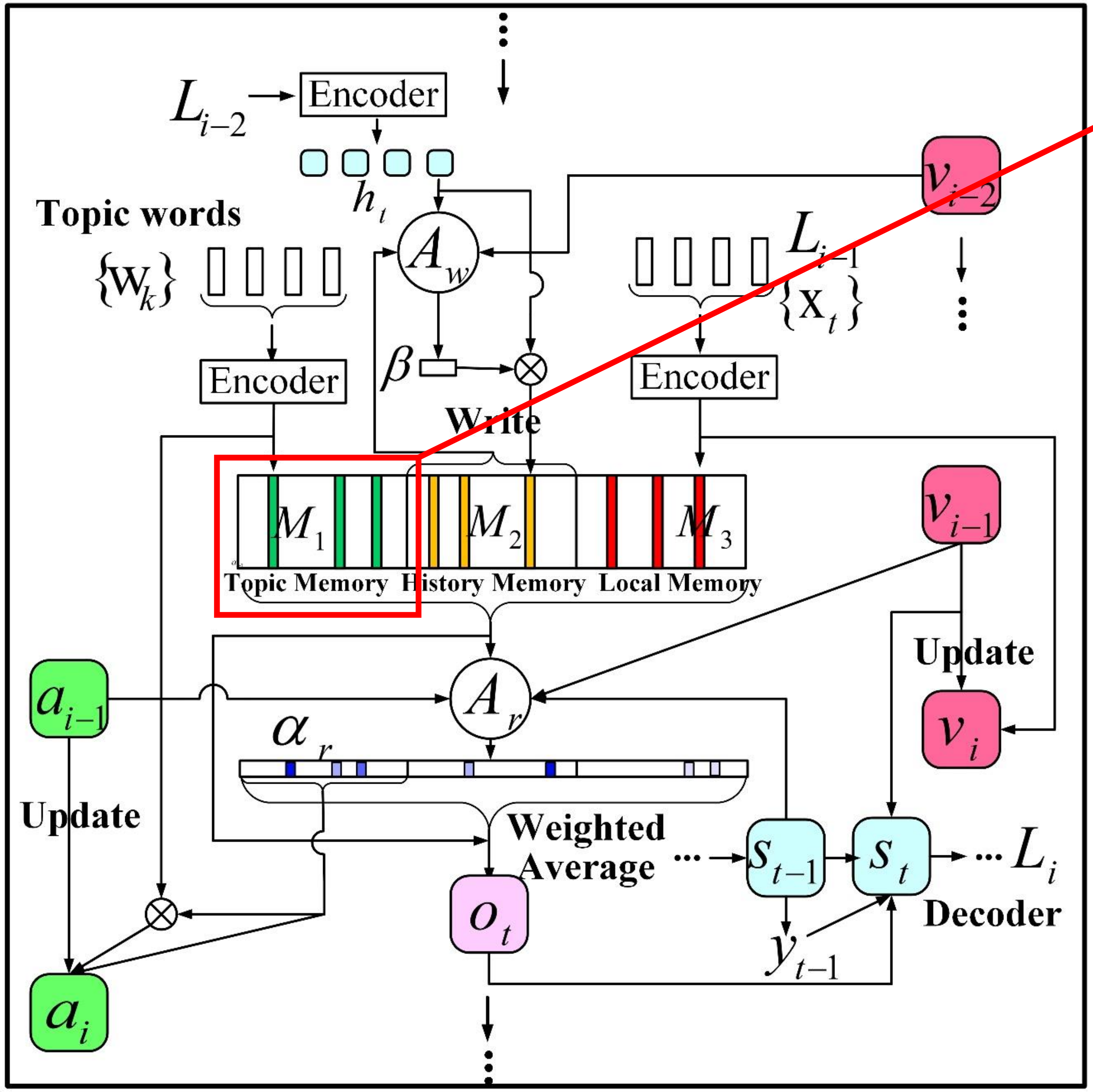
基于工作记忆模型的诗歌生成

Coherence in Psycholinguistics

Coherence is achieved if the reader can connect the incoming sentence to the content in working memory and to the major messages and points of the text [Sanders *et al.*, 2001].

The working memory is a system with a *limited capacity* that is responsible for *holding information* available for reasoning, decision-making and behavior [Priti and Miyake, 1999].

基于工作记忆模型的诗歌生成



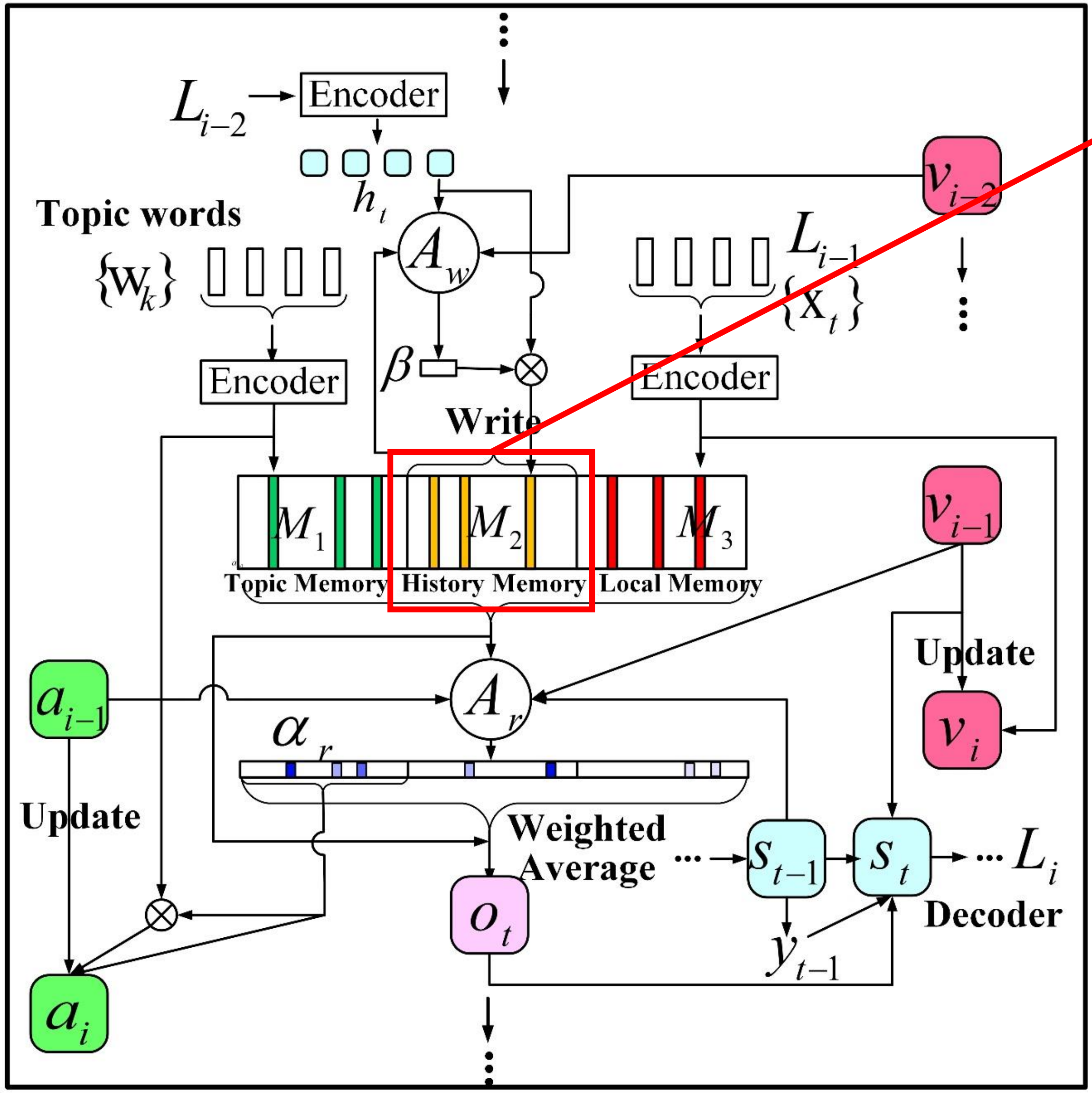
Topic Memory M_1

Each topic word is maintained in the topic memory explicitly and independently.

Major messages and points of the text !

Flexible order and form of the topic expression!

基于工作记忆模型的诗歌生成



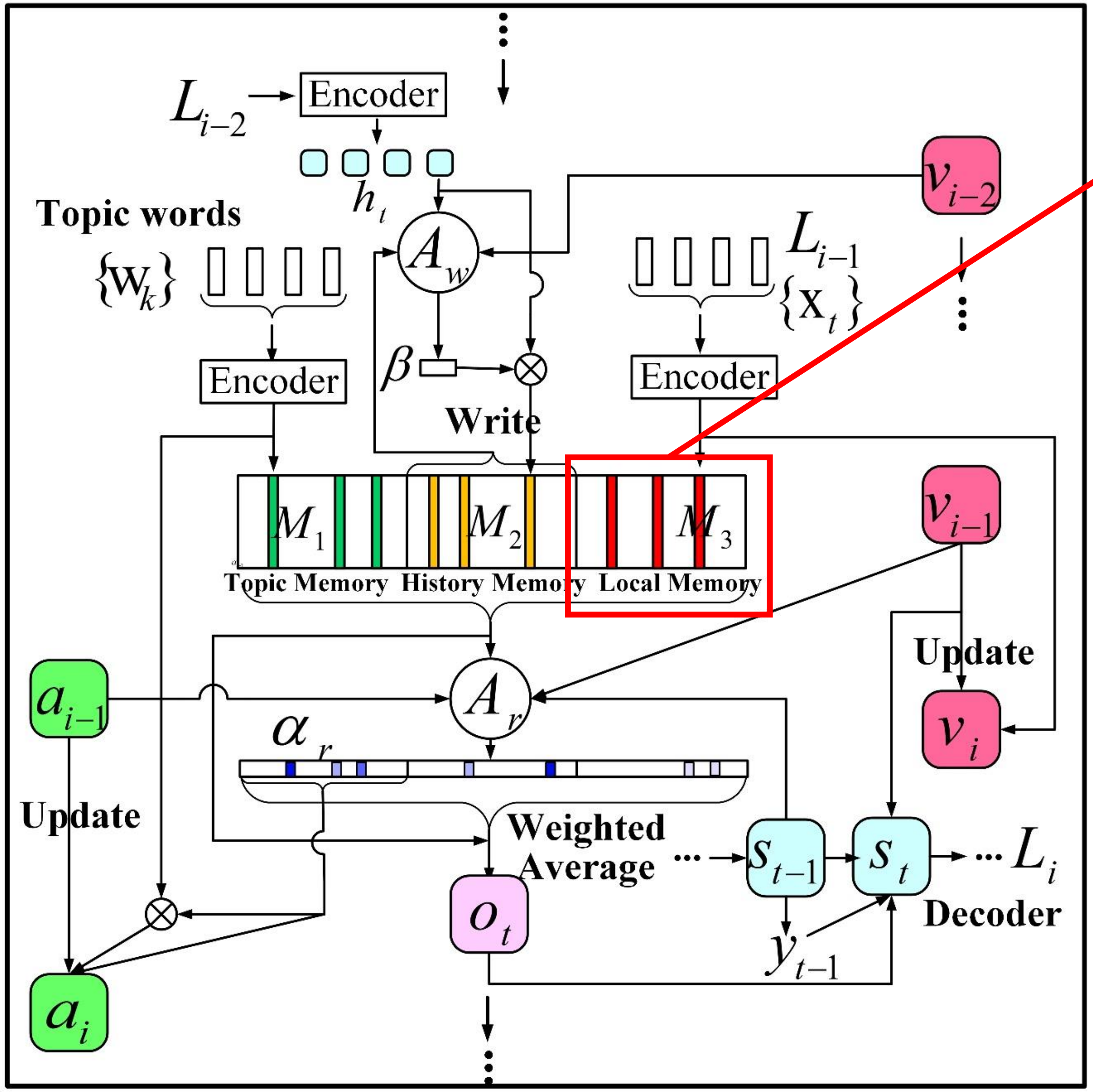
History Memory M_2

1. For each generated line: selects salient and informative characters to write into it.
2. When generating each line: reads most relevant from it.

Multiple, independent but limited memory slots→ enough capacity!

Always keep a coherent information flow in the history memory!

基于工作记忆模型的诗歌生成



Local Memory M_3

Save the previously generated line.

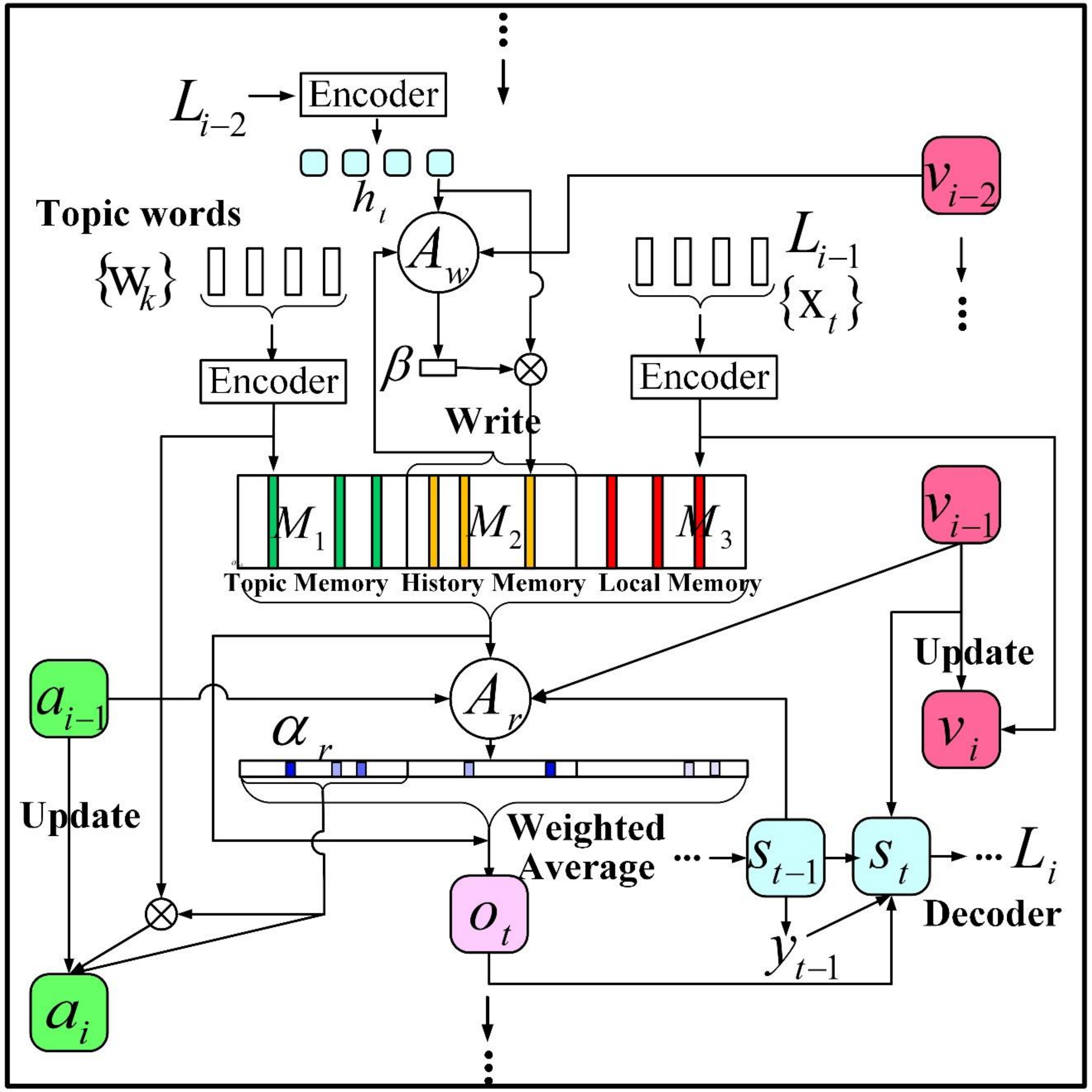
Strong semantic associations between two adjacent lines in Chinese poetry！

白日依山尽，

黄河入海流。

The local memory provides full short-distance history.

基于工作记忆模型的诗歌生成



Decoder:

$$s_t = GRU(s_{t-1}, [e(y_{t-1}); o_t; g_t; v_{i-1}]),$$
$$p(y_t | y_{1:t-1}, L_{1:i-1}, w_{1:K_1}) = softmax(W s_t),$$

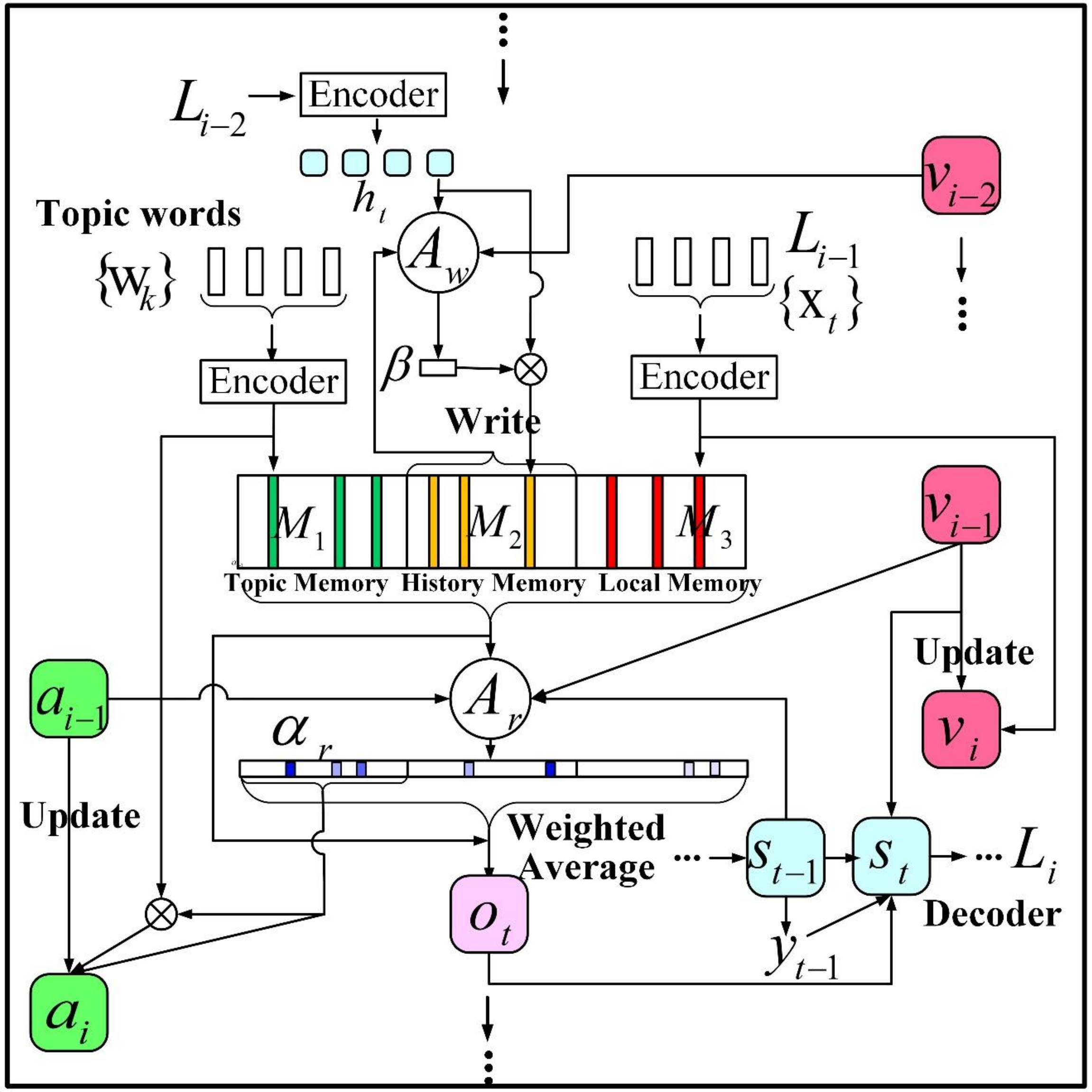
Global trace vector:

$$v_i = \sigma(v_{i-1}, \frac{1}{T_{enc}} \sum_{t=1}^{T_{enc}} h_t), v_0 = \mathbf{0}.$$

Addressing function: $\alpha = A(\tilde{M}, q)$

$$z_k = b^T \sigma(\tilde{M}[k], q),$$
$$\alpha[k] = softmax(z_k),$$

基于工作记忆模型的诗歌生成



Memory reading: $M = [M_1; M_2; M_3]$

$$\alpha_r = A_r(M, [s_{t-1}; v_{i-1}]),$$

$$o_t = \sum_{k=1}^{K} \alpha_r[k] * M[k],$$

Jointly read the three memory modules!

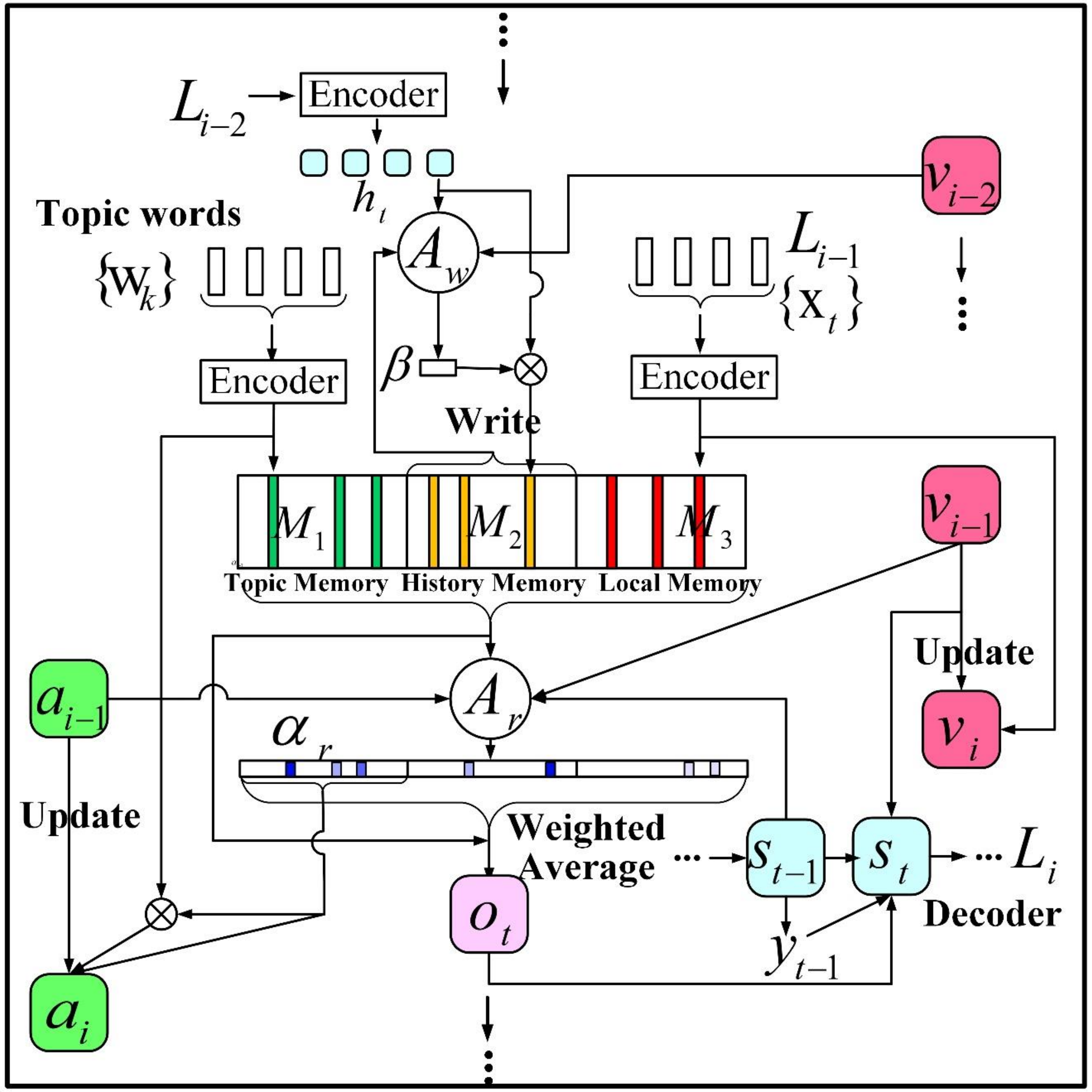
Memory writing:

$$\alpha_w = A_w(\tilde{M}_2, [h_t; v_{i-1}]),$$

$$\beta[k] = I(k = \arg\max_j \alpha_w[j]),$$

$$\tilde{M}_2[k] \leftarrow (1 - \beta[k]) * \tilde{M}_2[k] + \beta[k] * h_t,$$

基于工作记忆模型的诗歌生成



Genre embedding

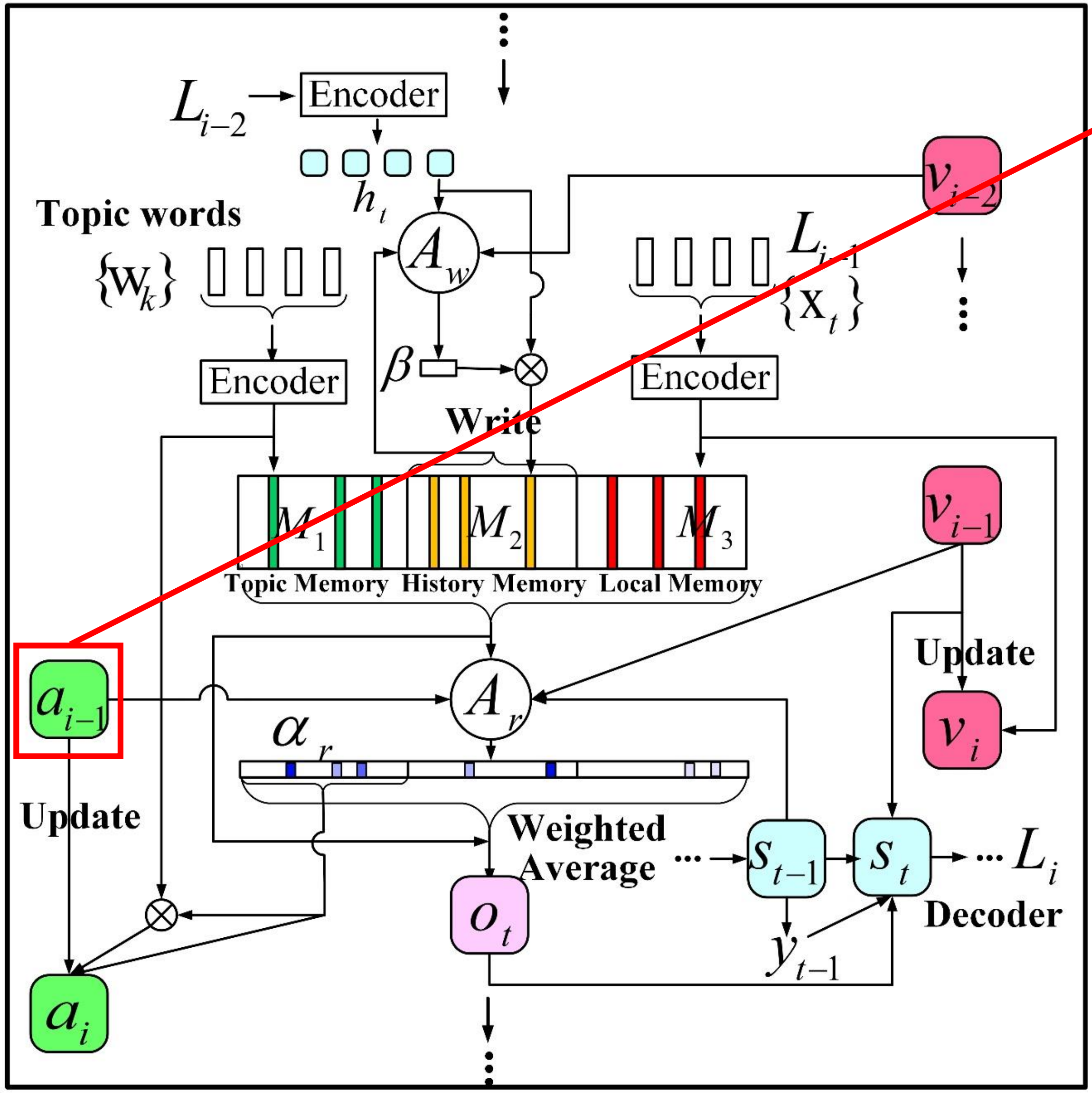
$$s_t = GRU(s_{t-1}, [e(y_{t-1}); o_t; \boxed{g_t}; v_{i-1}]),$$
$$p(y_t|y_{1:t-1}, L_{1:i-1}, w_{1:K_1}) = softmax(W s_t),$$

The genre embedding vector

- The length of each line.
- The phonological category of each character

Transparent to poetry genre!

基于工作记忆模型的诗歌生成



Topic Trace mechanism

Record the usage of topics in a more explicit way!

- Maintain the content of used topics
- Explicitly records the times of reading each topic

Improve topic expression ratio! (>70%)

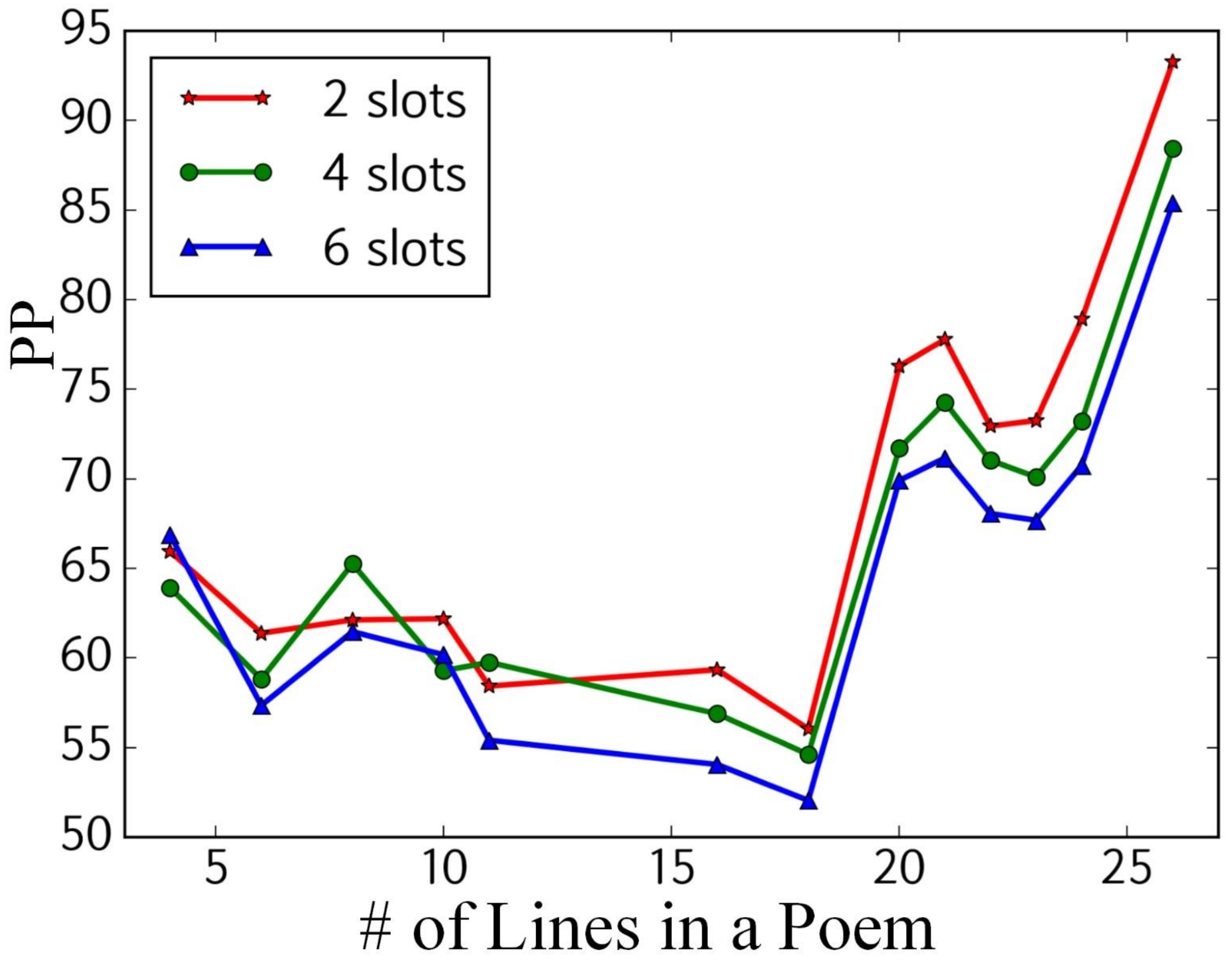
基于工作记忆模型的诗歌生成

Experiments

	Models	Fluency	Meaning	Coherence	Relevance	Aesthetics
Quatrains	Planning	2.28	2.13	2.18	2.50	2.31
	iPoet	2.54	2.28	2.27	2.13	2.45
	FCPG	2.36	2.15	2.15	2.65	2.28
	WM	**3.57****	**3.45****	**3.55****	**3.77****	**3.47****
	Human	3.62	3.52	3.59	3.78	3.58
Iambics	iambicGen	2.48	2.73	2.78	2.36	3.08
	WM	**3.39****	**3.69****	**3.77****	**3.87****	**3.87****
	Human	4.04	4.10++	4.13++	4.03	4.09
Lyrics	lyricGen	1.70	1.65	1.81	2.24	1.99
	WM	**2.63****	**2.49****	**2.46****	**2.53**	**2.66****
	Human	3.43++	3.20++	3.41++	3.34++	3.26++

Table 4: Human evaluation results. Diacritic ** ($p < 0.01$) indicates WM significantly outperforms baselines; ++ ($p < 0.01$) indicates Human is significantly better than all models. The Intraclass Correlation Coefficient of the four groups of scores is 0.5, which indicates an acceptable inter-annotator agreement.

基于工作记忆模型的诗歌生成



On iambics, perplexity over different numbers of lines in a poem

基于工作记忆模型的诗歌生成

Case Study

柳丝不绾系春愁，
Uncoiled willow twigs tie the spring sorrow.
梦里思量竟日留。
Missing from the dream stays all the day.
别后情怀似水流。
After bidding farewell to you, my feelings flow like water.
恨难酬，
With regret hard to release,
锦字红笺泪眼浮。
when reading the letter from you, I burst into tears.

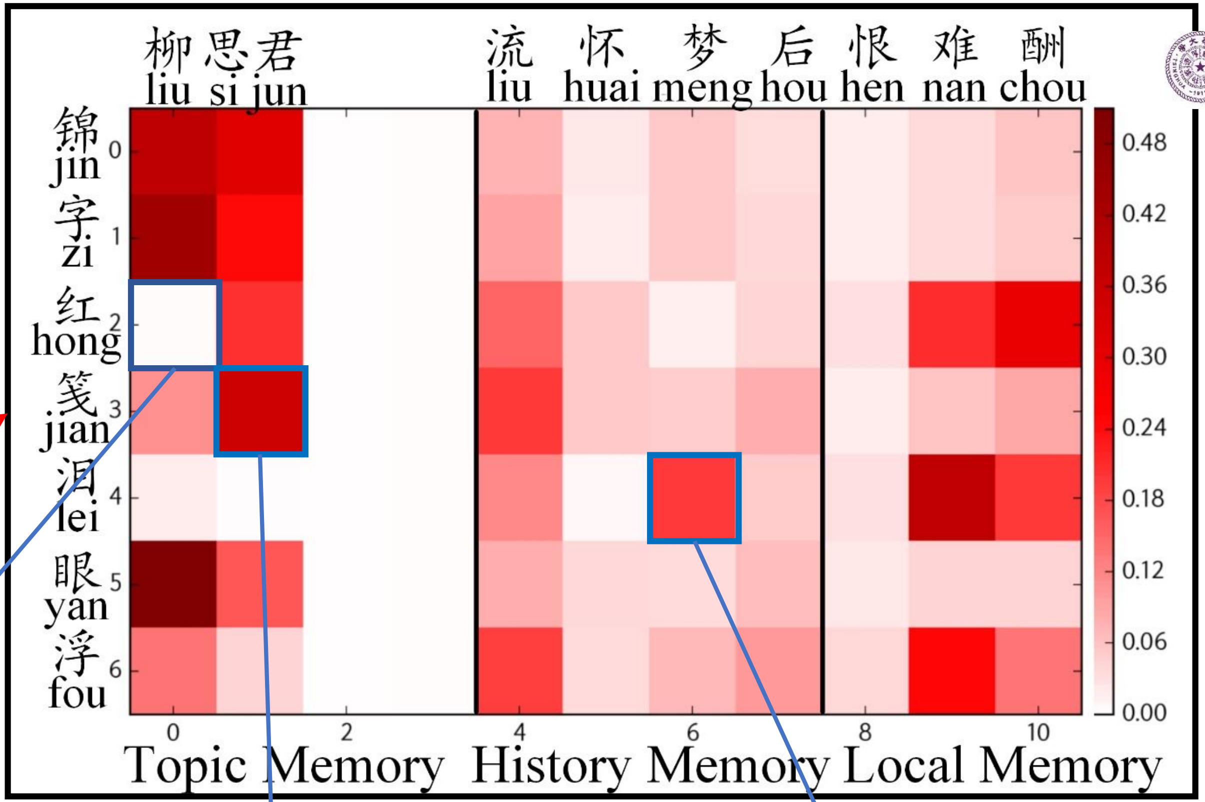
An iambic generated by our model, taking 柳 (willow) and 思君 (missing you) as input words.

基于工作记忆模型的诗歌生成

Case Study

柳丝不绾系春愁，
Uncoiled willow twigs tie the spring sorrow.
梦里思量竟日留。
Missing from the dream stays all the day.
别后情怀似水流。
After bidding farewell to you, my feelings flow like water.
恨难酬，
With regret hard to release,
锦字红笺泪眼浮。
when reading the letter from you, I burst into tears.

An iambic generated by our model, taking 柳 (willow) and 思君 (missing you) as input words.



The visualization of memory (in the x-axis) reading probabilities, α_r, when generating the last line (in the y-axis) of the iambic.

柳, willow ~~→~~ 红,red　　思君, missing you → 笺, letter　　梦, dream → 泪, tears

基于工作记忆模型的诗歌生成

柳丝不绾系春愁，
Uncoiled willow twigs tie the spring sorrow.
梦里思量竟日留。
Missing from the dream stays all the day.
别后情怀似水流。
After bidding farewell to you, my feelings flow like water.
恨难酬，
With regret hard to release,
锦字红笺泪眼浮。
when reading the letter from you, I burst into tears.

An iambic generated by our model, taking liu (willow) and sijun (missing you) as input words.

flexible form!

基于工作记忆模型的诗歌生成

柳丝不绾系春愁，
Uncoiled willow twigs tie the spring sorrow.
梦里思量竟日留。
Missing from the dream stays all the day.
别后情怀似水流。
After bidding farewell to you, my feelings flow like water.
恨难酬，
With regret hard to release,
锦字红笺泪眼浮。
when reading the letter from you, I burst into tears.

An iambic generated by our model, taking liu (willow) and sijun (missing you) as input words.

flexible form!

柳丝无力挽春愁，
Willow twigs could not tie the spring sorrow.
燕子归来恨未休。
The returned swallows arouse my endless regret.
记得当年锦绣楼，
I still remember, at that time, in the splendid mansion,
为君留，
I stay for you.
别后相思泪满眸。
After bidding farewell to you, I miss you with my eyes full of tears.

Force the keyword, sijun (missing you) not to appear in the first three lines.

flexible order!

基于互信息的无监督风格诗歌生成

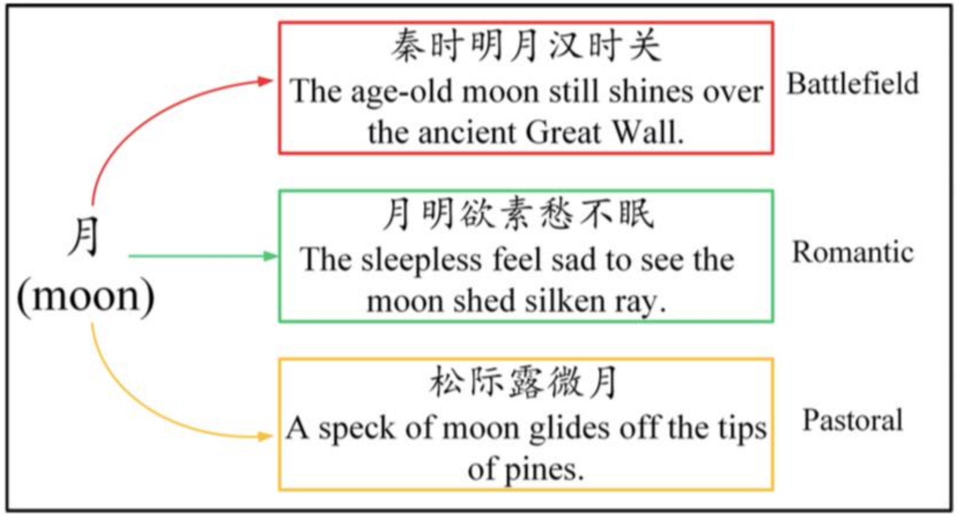
Stylistic Chinese Poetry Generation via Unsupervised Style Disentanglement

Cheng Yang, Maosong Sun, Xiaoyuan Yi, Wenhao Li

In EMNLP 2018

基于互信息的无监督风格诗歌生成

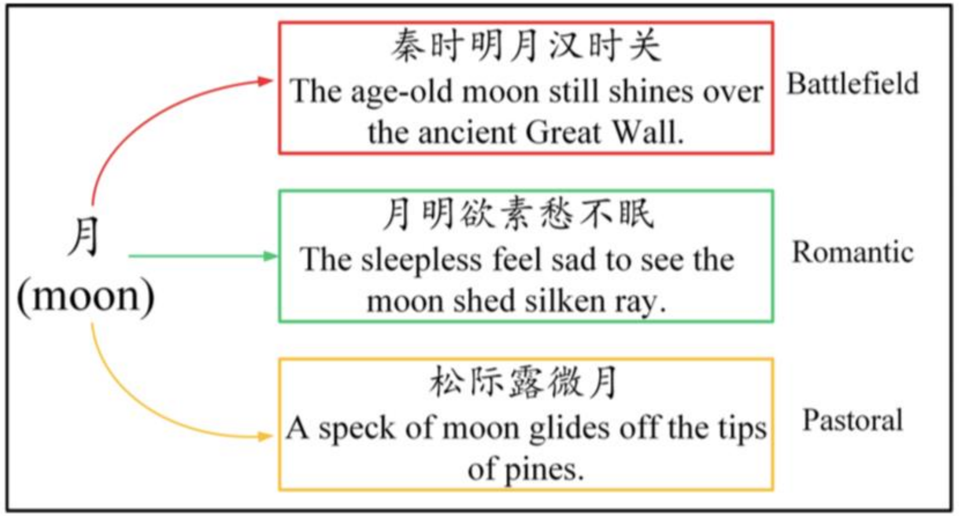
Motivation



An example of poems in diverse styles under the same keyword.

基于互信息的无监督风格诗歌生成

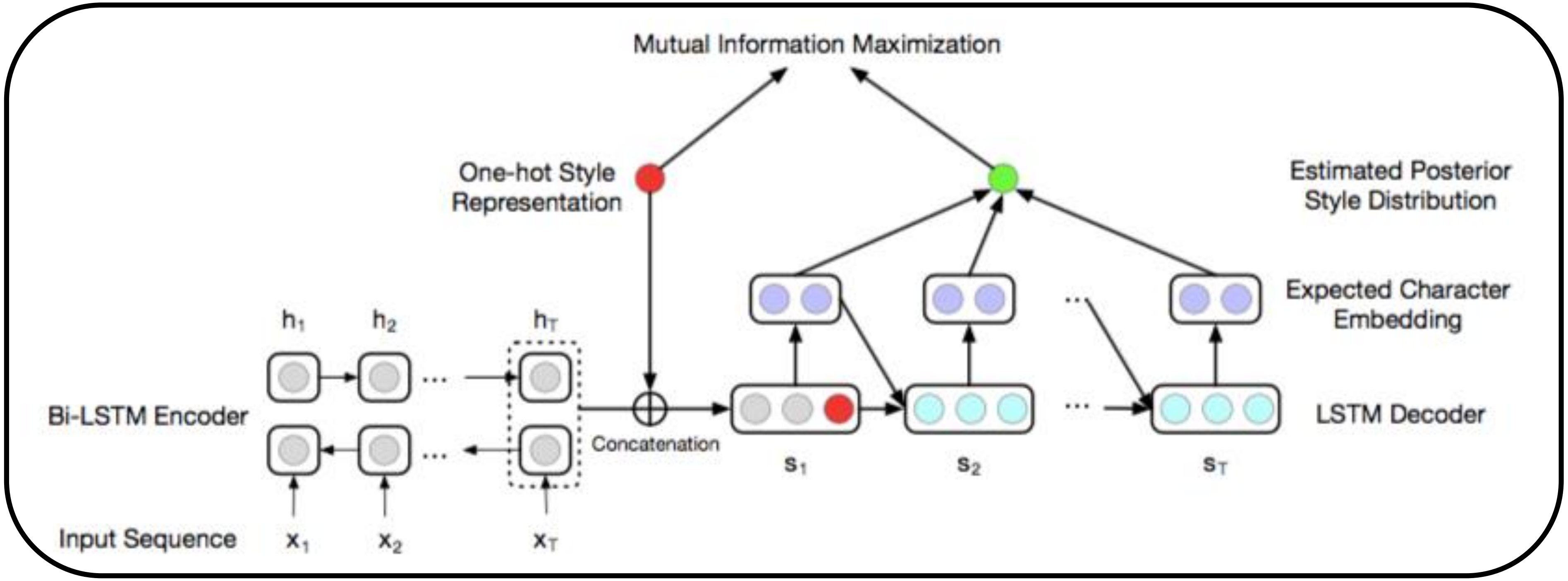
Motivation



An example of poems in diverse styles under the same keyword.

- Our model should be able to generate multiple outputs given the same input.
- The outputs should be diverse and different in style from each other.
- The method should be unsupervised since there is no explicit label or topic labeling for poems in the corpus.
- There should be no loss in other criteria, e.g, fluency.

基于互信息的无监督风格诗歌生成



An overview of style disentanglement by mutual information maximization.

基于互信息的无监督风格诗歌生成

Basic Framework

- Encoder

$$\overrightarrow{h_i} = LSTM_{forward}(\overrightarrow{h_{i-1}}, e(x_i)),$$

$$\overleftarrow{h_i} = LSTM_{backward}(\overleftarrow{h_{i-1}}, e(x_{T-i+1})),$$

$$h_i = [\overrightarrow{h_i}, \overleftarrow{h_{T-i+1}}]$$

- Decoder

$$s_i = LSTM_{decoder}(s_{i-1}, [e(y_{i-1}), a_i]),$$

$$a_i = attention(s_{i-1}, h_{1:T}).$$

$$p(y_i|y_1y_2\ldots y_{i-1}, X) = g(y_i|s_i),$$

$$s_0 = h_T$$

基于互信息的无监督风格诗歌生成

Basic Framework

- Decoder

$$s_i = LSTM_{decoder}(s_{i-1}, [e(y_{i-1}), a_i]),$$

$$p(y_i|y_1 y_2 \ldots y_{i-1}, X) = g(y_i|s_i),$$

$$s_0 = h_T$$

$$s_0 = [onehot(k);\ h_T]$$

Our model takes two arguments as input:
input sentence s_{input} and
style id $k \in 1,2, \ldots, K$.

基于互信息的无监督风格诗歌生成

Add a regularization term to force a strong dependency relationship between the input style id and generated sentence!

Given two random variables X and Y, the mutual information I(X,Y) measures "the amount of information" obtained about one random variable given another one.

$$I(X,Y) = \int_Y \int_X p(X,Y) \log \frac{p(X,Y)}{p(X)p(Y)} dX dY.$$

基于互信息的无监督风格诗歌生成

Assume that the input style id is a uniformly distributed random variable Sty and $\Pr(Sty) = \frac{1}{K}, k = 1,2, \ldots, K$

Maximize the mutual information between the style distribution $\Pr(Sty)$ and the generated sentence distribution $\Pr(X;Y)$given input sentence X.

基于互信息的无监督风格诗歌生成

$$
\begin{aligned}
& I(\Pr(Sty), \Pr(Y;X)) \\
= & \sum_{k=1}^{K} \Pr(Sty=k) \int_{Y|k;X} \log \frac{\Pr(Y, Sty=k; X)}{\Pr(Sty=k)\Pr(Y;X)} dY \\
= & \sum_{k=1}^{K} \Pr(Sty=k) \int_{Y|k;X} \log \frac{\Pr(Y, Sty=k; X)}{\Pr(Y;X)} dY \\
& - \sum_{k=1}^{K} \Pr(Sty=k) \log \Pr(Sty=k) \\
= & \sum_{k=1}^{K} \Pr(Sty=k) \int_{Y|k;X} \log \Pr(Sty=k|Y) dY + \log K \\
= & \int_{Y;X} \sum_{k=1}^{K} \boxed{\Pr(Sty=k|Y)} \log P(Sty=k|Y) dY + \log K.
\end{aligned}
$$

基于互信息的无监督风格诗歌生成

$$
\begin{aligned}
&I(\Pr(Sty), \Pr(Y;X)) \\
&= \sum_{k=1}^{K} \Pr(Sty=k) \int_{Y|k;X} \log \frac{\Pr(Y, Sty=k; X)}{\Pr(Sty=k)\Pr(Y;X)} dY \\
&= \sum_{k=1}^{K} \Pr(Sty=k) \int_{Y|k;X} \log \frac{\Pr(Y, Sty=k; X)}{\Pr(Y;X)} dY \\
&\quad - \sum_{k=1}^{K} \Pr(Sty=k) \log \Pr(Sty=k) \\
&= \sum_{k=1}^{K} \Pr(Sty=k) \int_{Y|k;X} \log \Pr(Sty=k|Y) dY + \log K \\
&= \int_{Y;X} \sum_{k=1}^{K} \boxed{\Pr(Sty=k|Y)} \log P(Sty=k|Y) dY + \log K.
\end{aligned}
$$

Variational Lower Bound

$$
\begin{aligned}
&I(\Pr(Sty), \Pr(Y;X)) - \log K \\
&= \int_{Y;X} \sum_{k=1}^{K} \Pr(Sty=k|Y) \log \Pr(Sty=k|Y) dY \\
&= \int_{Y;X} \sum_{k=1}^{K} \Pr(Sty=k|Y) \log Q(Sty=k|Y) dY \\
&\quad + \int_{Y;X} \sum_{k=1}^{K} \Pr(Sty=k|Y) \log \frac{\Pr(Sty=k|Y)}{Q(Sty=k|Y)} dY \\
&= \int_{Y;X} \sum_{k=1}^{K} \Pr(Sty=k|Y) \log Q(Sty=k|Y) dY \\
&\quad + \int_{Y;X} KL(\Pr(Sty|Y), Q(Sty|Y)) dY \\
&\geq \int_{Y;X} \sum_{k=1}^{K} \Pr(Sty=k|Y) \log Q(Sty=k|Y) dY \\
&= \sum_{k=1}^{K} \Pr(Sty=k) \int_{Y|k;X} \log Q(Sty=k|Y) dY.
\end{aligned}
$$

基于互信息的无监督风格诗歌生成

$$
\begin{aligned}
& I(\Pr(Sty), \Pr(Y;X)) \\
&= \sum_{k=1}^{K} \Pr(Sty=k) \int_{Y|k;X} \log \frac{\Pr(Y, Sty=k; X)}{\Pr(Sty=k)\Pr(Y;X)} dY \\
&= \sum_{k=1}^{K} \Pr(Sty=k) \int_{Y|k;X} \log \frac{\Pr(Y, Sty=k; X)}{\Pr(Y;X)} dY \\
&\quad - \sum_{k=1}^{K} \Pr(Sty=k) \log \Pr(Sty=k) \\
&= \sum_{k=1}^{K} \Pr(Sty=k) \int_{Y|k;X} \log \Pr(Sty=k|Y) dY + \log K \\
&= \int_{Y;X} \sum_{k=1}^{K} \boxed{\Pr(Sty=k|Y)} \log P(Sty=k|Y) dY + \log K.
\end{aligned}
$$

Variational Lower Bound

$$
\begin{aligned}
& I(\Pr(Sty), \Pr(Y;X)) - \log K \\
&= \int_{Y;X} \sum_{k=1}^{K} \Pr(Sty=k|Y) \log \Pr(Sty=k|Y) dY \\
&= \int_{Y;X} \sum_{k=1}^{K} \Pr(Sty=k|Y) \log Q(Sty=k|Y) dY \\
&\quad + \int_{Y;X} \sum_{k=1}^{K} \Pr(Sty=k|Y) \log \frac{\Pr(Sty=k|Y)}{Q(Sty=k|Y)} dY \\
&= \int_{Y;X} \sum_{k=1}^{K} \Pr(Sty=k|Y) \log Q(Sty=k|Y) dY \\
&\quad + \int_{Y;X} KL(\Pr(Sty|Y), Q(Sty|Y)) dY \\
&\geq \int_{Y;X} \sum_{k=1}^{K} \Pr(Sty=k|Y) \log Q(Sty=k|Y) dY \\
&= \boxed{\sum_{k=1}^{K} \Pr(Sty=k) \int_{Y|k;X} \log Q(Sty=k|Y) dY.}
\end{aligned}
$$

基于互信息的无监督风格诗歌生成

Maximize $$\sum_{k=1}^{K} \Pr(Sty = k) \int_{Y|k;X} \log Q(Sty = k|Y) dY.$$

$$Q(Sty|Y) = softmax(W \cdot \frac{1}{T} \sum_{i=1}^{T} e(y_i)),$$

基于互信息的无监督风格诗歌生成

Maximize $$\sum_{k=1}^{K} \Pr(Sty = k) \int_{Y|k;X} \log Q(Sty = k|Y) dY.$$

$$Q(Sty|Y) = softmax(W \cdot \frac{1}{T} \sum_{i=1}^{T} e(y_i)),$$

Problem: Impossible to integrate over all possible sequences.

AI 研习社 · 大讲堂

基于互信息的无监督风格诗歌生成

Maximize $$\sum_{k=1}^{K} \Pr(Sty = k) \int_{Y|k;X} \log Q(Sty = k|Y) dY.$$

$$Q(Sty|Y) = softmax(W \cdot \frac{1}{T} \sum_{i=1}^{T} e(y_i)),$$

Problem: Impossible to integrate over all possible sequences.

Solution: We only generate an expected embedding sequence and suppose $Y|k;X$ has one hundred percent probability generating this one.

基于互信息的无监督风格诗歌生成

Expected Character Embedding (Kočciský et al., 2016)

$$expect(i; k, X) = \sum_{c \in V} g(c|s_i)e(c),$$

$$s_{i+1} = LSTM_{decoder}(s_i, [expect(i; k, X), a_{i+1}]).$$

基于互信息的无监督风格诗歌生成

Expected Character Embedding (Kočciský et al., 2016)

$$expect(i;k,X) = \sum_{c \in V} g(c|s_i)e(c),$$

$$s_{i+1} = LSTM_{decoder}(s_i, [expect(i;k,X), a_{i+1}]).$$

$$\sum_{k=1}^{K} \Pr(Sty = k) \int_{Y|k;X} \log Q(Sty = k|Y)dY. \approx \frac{1}{K} \sum_{k=1}^{K} \log\{softmax(W \cdot \frac{1}{T} \sum_{i=1}^{T} expect(i;k,X))[k]\},$$

$$= \mathcal{L}_{reg}$$

AI 研习社 · 大讲堂

基于互信息的无监督风格诗歌生成

Maximize

$$Train(X,Y) = \sum_{i=1}^{T} \log p(y_i | y_1 y_2 \ldots y_{i-1}, X) + \lambda \mathcal{L}_{reg},$$

基于互信息的无监督风格诗歌生成

Maximize

$$Train(X,Y) = \boxed{\sum_{i=1}^{T} \log p(y_i|y_1 y_2 \dots y_{i-1}, X)} + \lambda \mathcal{L}_{reg},$$

style irrelevant generation likelihood,
which computed by setting one-hot
style representation to an all-zero vector.

基于互信息的无监督风格诗歌生成

Maximize

$$Train(X,Y) = \boxed{\sum_{i=1}^{T} \log p(y_i|y_1 y_2 \ldots y_{i-1}, X)} + \lambda \mathcal{L}_{reg},$$

style irrelevant generation likelihood,
which computed by setting one-hot
style representation to an all-zero vector.

Ensures that the decoder can generate
fluent and coherent outputs.

基于互信息的无监督风格诗歌生成

Maximize

$$Train(X,Y) = \boxed{\sum_{i=1}^{T} \log p(y_i | y_1 y_2 \ldots y_{i-1}, X)} + \boxed{\lambda \mathcal{L}_{reg},}$$

style irrelevant generation likelihood, which computed by setting one-hot style representation to an all-zero vector.

Style regularization

Ensures that the decoder can generate fluent and coherent outputs.

基于互信息的无监督风格诗歌生成

Maximize

$$Train(X,Y) = \boxed{\sum_{i=1}^{T} \log p(y_i | y_1 y_2 \dots y_{i-1}, X)} + \boxed{\lambda \mathcal{L}_{reg},}$$

style irrelevant generation likelihood, which computed by setting one-hot style representation to an all-zero vector.

Ensures that the decoder can generate fluent and coherent outputs.

Style regularization

Guarantees the style-specific output has a strong dependency on the one-hot style representation input.

基于互信息的无监督风格诗歌生成

Maximize

$$Train(X, Y) = \boxed{\sum_{i=1}^{T} \log p(y_i | y_1 y_2 \ldots y_{i-1}, X)} + \boxed{\lambda \mathcal{L}_{reg},}$$

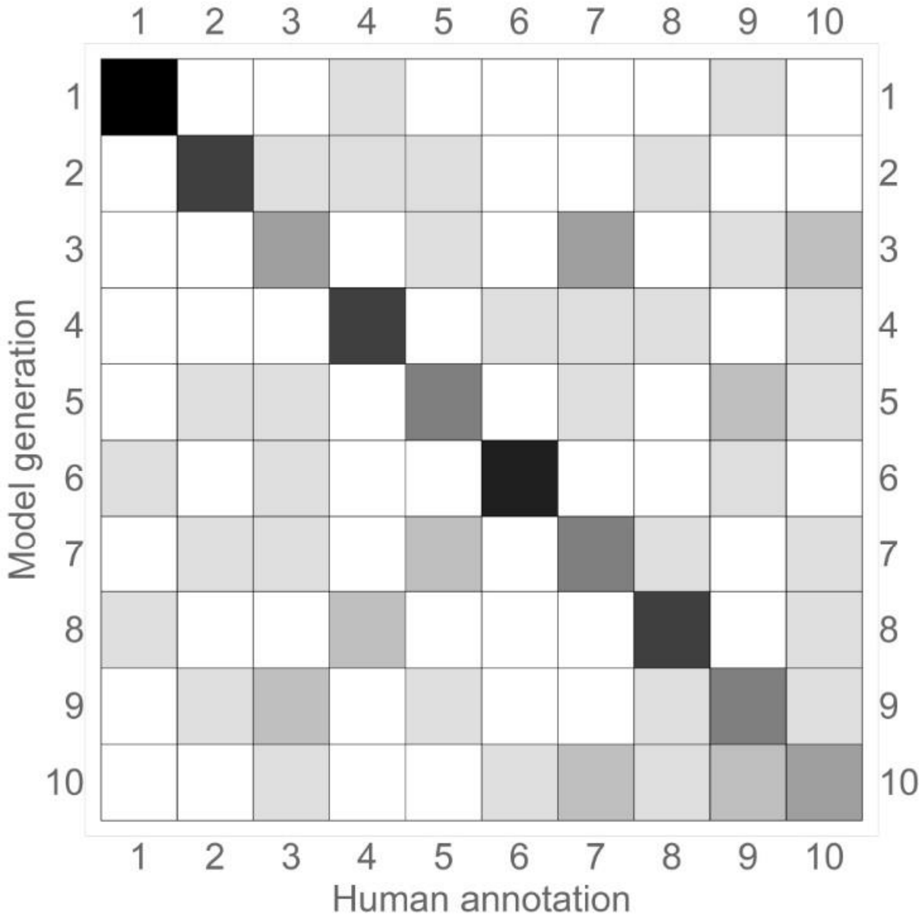
our model is not task-specific: the regularization term can be added to any other basic models conveniently for diverse or stylistic generations.

AI 研习社 · 大讲堂

基于互信息的无监督风格诗歌生成

Style id	Keywords
1	loneliness, melancholy
2	the portrait of landscape
3	sorrow during roaming
4	hermit, rural scenes
5	grand scenery, regrets about old events
6	sorrow during drinking
7	emotions towards life experience
8	the portrait of hazy sceneries
9	reminiscence, homesickness
10	sadness about seasons

Representative keywords for poems generated by each learned style.



Experimental results on style recognition. Each row represents the human annotation of corresponding style generations. The diagonal blocks are correct classifications. Darker block indicates higher probability.

基于互信息的无监督风格诗歌生成

浊酒一杯聊酩酊，
After a cup of unstrained wine,
I have been a little drunk
白云千里断鸿濛。
I saw the cloud split the sky apart.
马蹄踏破青山路，
On horseback, I pass through every road across the mountain,
惆怅斜阳落日红。
but can only watch the red sun falling down with sorrow.

(a) Style 1: "loneliness, melancholy"

浊酒一杯聊酩酊，
After a cup of unstrained wine,
I have been a little drunk
扁舟何处问渔樵。
With a narrow boat, where could I find the hermits?
行人莫讶归来晚，
Friends, don't be surprised that I come back so late,
万里春风到海潮。
I have seen the great tide and the grand spring breeze.

(b) Style 4: "hermit, rural scenes"

浊酒一杯聊酩酊，
After a cup of unstrained wine,
I have been a little drunk
浮云何处觅仙踪。
I wonder on which cloud I can see the presence of the gods.
迢迢十二峰头月，
The moon above the mount seems farther and farther.
漠漠千山暮霭浓。
The mist among the hill becomes thicker and thicker.

(c) Style 8: "the portrait of hazy sceneries"

谢谢！

- 九歌系统在线页面: https://jiuge.thunlp.cn
- THUNLP 实验室主页: http://nlp.csai.tsinghua.edu.cn

Mail: mtmoonyi@gmail.com

AI 研习社 · 大讲堂

Reference

- Pablo Gerv´as. 2001. *An Expert System for the Composition of Formal Spanish Poetry*. Springer London.
- Xiaofeng Wu, Naoko Tosa, and Ryohei Nakatsu. 2009. New hitch haiku: An interactive renku poem composition supporting tool applied for sightseeing navigation system. In *Proceedings of the 8th International Conference on Entertainment Computing*, pages 191–196, Paris, France.
- Hisar Maruli Manurung. 2003. *An evolutionary algorithm approach to poetry generation*. Ph.D. thesis, University of Edinburgh.
- Robert P. Levy. 2001. A computational model of poetic creativity with neural network as measure of adaptive fitness. In *Proceedings of the ICCBR-01 Workshop on Creative Systems*.
- Rui Yan, Han Jiang, Mirella Lapata, Shou-De Lin, Xueqiang Lv, and Xiaoming Li. 2013. I, poet:automatic chinese poetry composition through a generative summarization framework under constrained optimization. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence*, pages 2197–2203, Beijing, China.
- Long Jiang and Ming Zhou. 2008. Generating Chinese couplets using a statistical mt approach. In *Proceedings of the 22nd International Conference on Computational Linguistics*, pages 377–384, Manchester, UK.
- Jing He, Ming Zhou, and Long Jiang. 2012. Generating chinese classical poems with statistical machine translation models. In *Proceedings of the 26th AAAI Conference on Artificial Intelligence*, pages 1650–1656, Toronto, Canada.

Reference

- Xingxing Zhang and Mirella Lapata. 2014. Chinese poetry generation with recurrent neural networks. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pages 670–680, Doha, Qatar.
- Rui Yan. 2016. i,poet:automatic poetry composition through recurrent neural networks with iterative polishing schema. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, pages 2238–2244, New York, USA.
- Zhe Wang, Wei He, Hua Wu nad Haiyang Wu, Wei Li, Haifeng Wang, and Enhong Chen. 2016. Chinese poetry generation with planning based neural network. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics:Technical Papers*, pages 1051–1060, Osaka, Japan.
- Jiyuan Zhang, Yang Feng, Dong Wang, Yang Wang, Andrew Abel, Shiyue Zhang, and Andi Zhang. 2017 Flexible and creative chinese poetry generation using neural memory. *In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, pages 1364–1373. Association for Computational Linguistics.
- Marjan Ghazvininejad, Xing Shi, Yejin Choi, and Kevin Knight. Generating topical poetry. *In Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1183–1191, Texas, USA, 2016.
- Dzmitry Bahdanau, KyungHyun Cho, and Yoshua Bengio. 2015. Neural machine translation by jointly learning to align and translate. *In Proceedings of the 2015 International Conference on Learning Representations*, San Diego, CA.

AI 研习社 · 大讲堂

Reference

- Tomáš Kočciský, Gábor Melis, Edward Grefenstette, Chris Dyer, Ling Wang, Phil Blunsom, and Karl Moritz Hermann. 2016. Semantic parsing with semi-supervised sequential autoencoders. *In Conference on Empirical Methods in Natural Language Processing*, pages 1078–1087.