TEXT GENERATION WITH VARIATION MODELS

大组会

于孟萱

2019.10.14

LONG TEXT GENERATION

- **Sentiment-Controllable Chinese Poetry Generation (IJCAI 2019)**
- **Long and Diverse Text Generation with Planning-based Hierarchical Variational Model (EMNLP 2019)**
- **Syntax-Infused Variational Auto-encoder for Text Generation (ACL 2019)**

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION

Sentiment-Controllable Chinese Poetry Generation

Huimin Chen[1,2,3,4*], **Xiaoyuan Yi**[1,2,3,4*], **Maosong Sun**[1,2,3,4†], **Wenhao Li**[1,2,3,4], **Cheng Yang**[1,2,3,4], **Zhipeng Guo**[1,2,3,4]
[1]Department of Computer Science and Technology, Tsinghua University
[2]Institute for Artificial Intelligence, Tsinghua University
[3]Beijing National Research Center for Information Science and Technology, Tsinghua University
[4]State Key Lab on Intelligent Technology and Systems, Tsinghua University
{chm15, yi-xy16, cheng-ya14, liwh16, gzp17}@mails.tsinghua.edu.cn, sms@tsinghua.edu.cn

IJCAI 2019

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION

Achieved significant improvements in fluency and coherence but neglected to generate sentiment-controllable poetry.

- Sentiment collapse: the generated poems tend to be neutral and meaningless descriptions)
- Sentiment bias: the generated poems tend to express negative sentiment (may hurts the semantics and diversity)

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION

Contributions:

- Build a fine-grained manually-labelled sentimental Chinese Poetry Corpus
- Utilise a Semi-supervised Variational AutoEncoder for generating sentimental poetry

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION

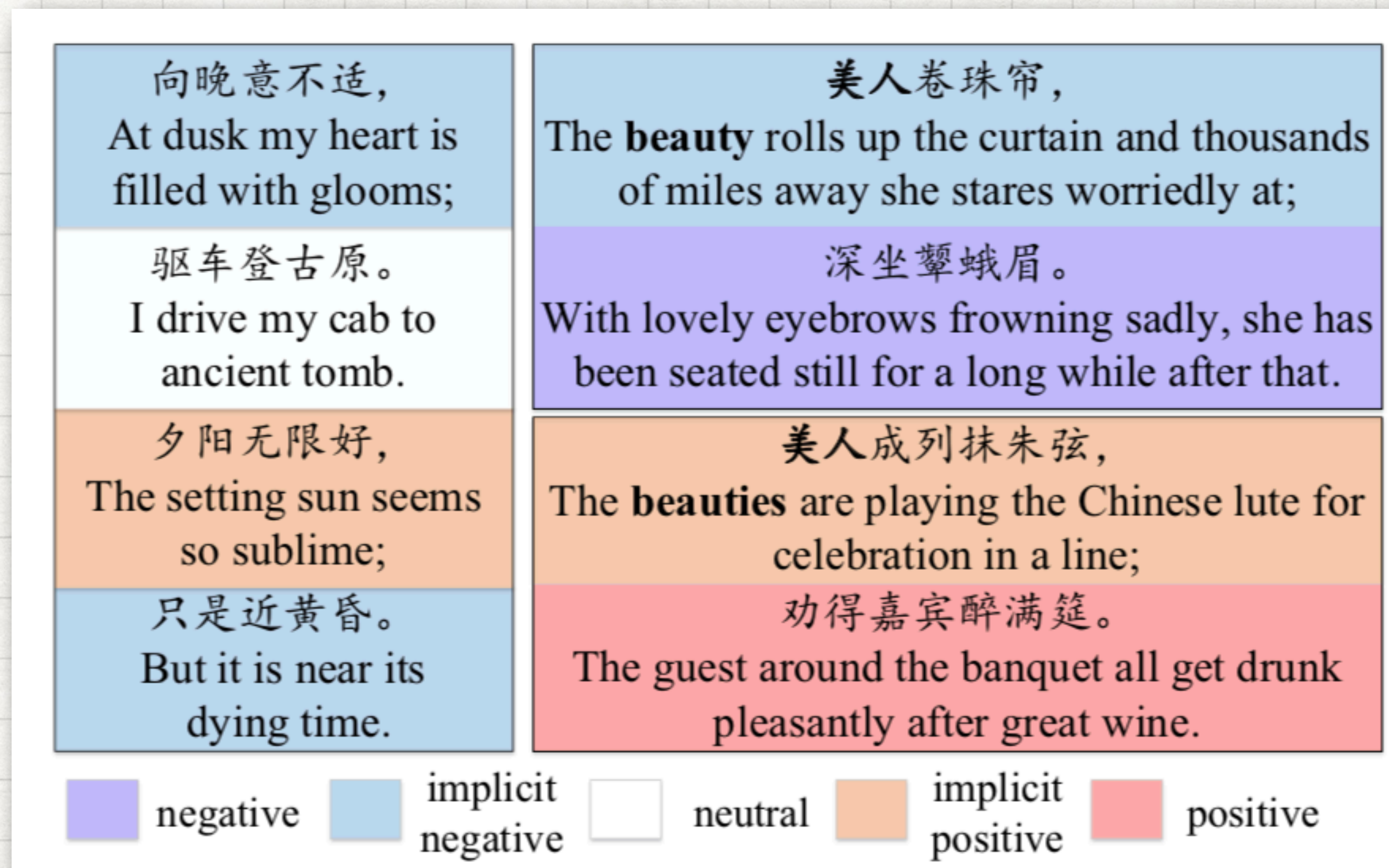
Fine-grained Sentimental Poetry Corpus

- Amount: 5,000 Chinese quatrains.

Annotate each poem and each line

into 5 classes,

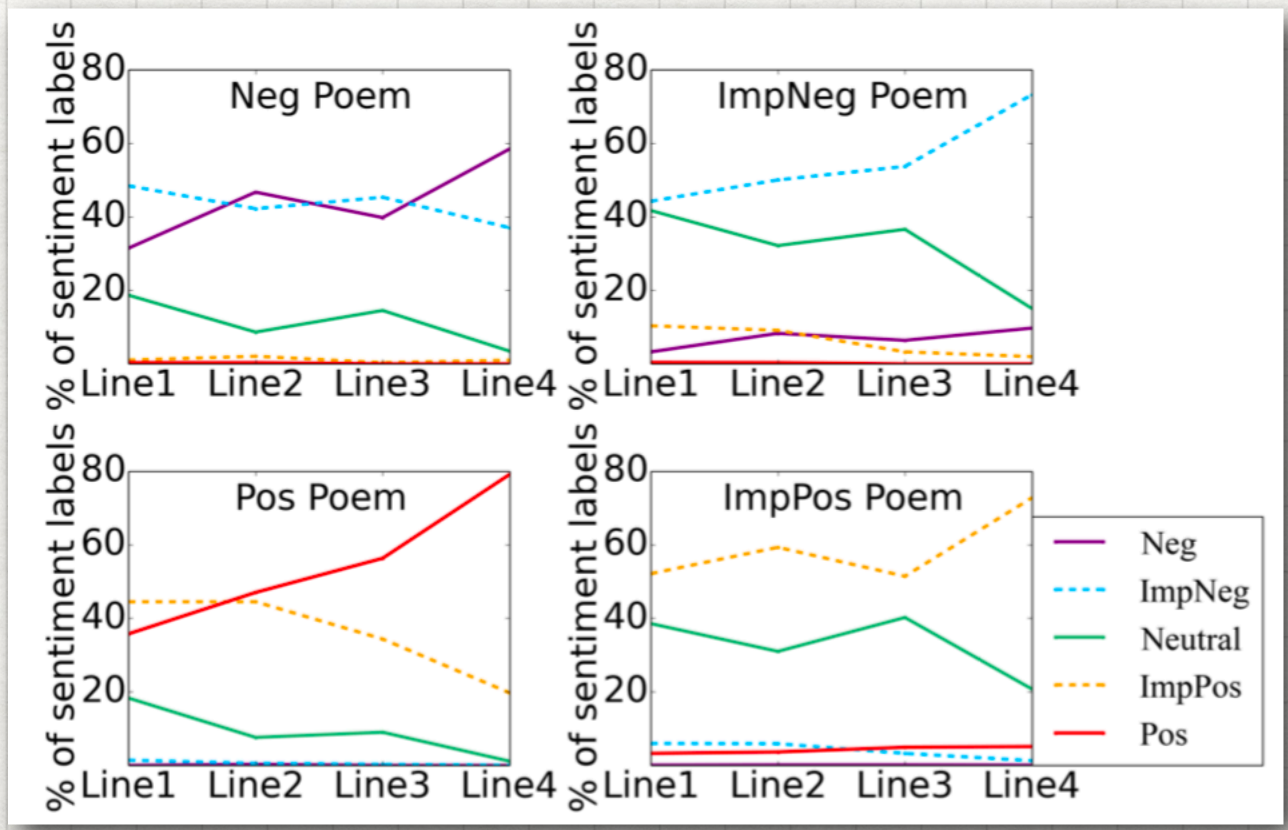
- negative
- implicit negative
- neutral
- implicit positive
- positive



SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION

Fine-grained Sentimental Poetry Corpus

Granularity	#Neg.	#Implicit Neg.	#Neutral	#Implicit Pos.	#Pos.
Whole Poem	289	1,467	1,328	1,561	355
Line1	143	1,023	2,337	1,310	187
Line2	268	1,138	1,936	1,423	235
Line3	212	1,107	2,320	1,083	278
Line4	315	1,317	1,650	1,357	361



Model both the sentiment of the <u>whole poetry</u> and the sentiment transition patterns <u>across different lines</u>.

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION

Also utilised 146,835 unlabelled poetry

For each poem, extracted three keywords using TEXT RANK

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION

Semi-supervised Sentiment-Controllable Poetry Generation Model (SCPG)

Problem formulation:

- x: poem with n lines {x1 , x2 , ..., xn }
- w: keyword which represents the main topic of x,
- y: holistic sentiment of x
- {y1 , y2 , ..., yn } as the sentiments expressed in each line,
- pl(x, w, y, y1:n) the distributions over labelled dataset
- pu(x, w) the distributions over unlabelled datasets

With the keyword w, aim to generate poems not only holding the holistic sentiment y for the whole poem but also expressing the sentiment yi for each line xi

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION

SCPG -

- Holistic Sentiment Control Module
- Temporal Sentiment Control Module

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION

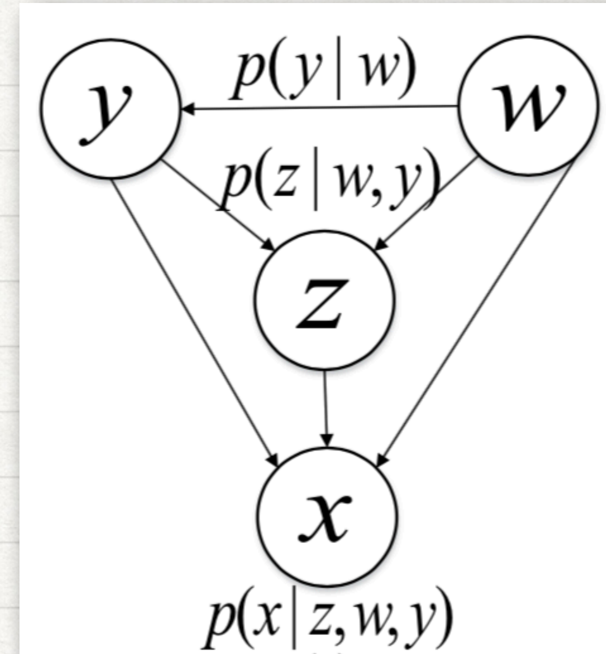
Goal: to learn a conditional joint distribution p(x, y, z | w), where z is the latent variable

- Holistic Sentiment Control Module

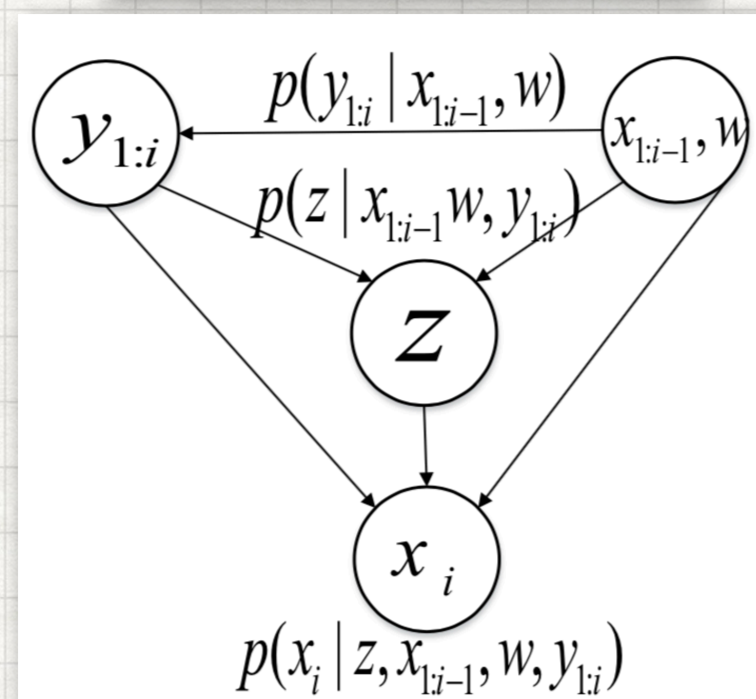
  p(x, y, z | w) = p(x | z, w, y) * p(z | y, w) * p(y | w)

- Temporal Sentiment Control Module

$$\log\, p(x_{1:n}, y_{1:n}|w) = \log\, p(x_1, y_1|w) + \sum_{i=2}^{n} \log\, p(x_i, y_i|x_{1:i-1}, y_{1:i-1}, w).$$

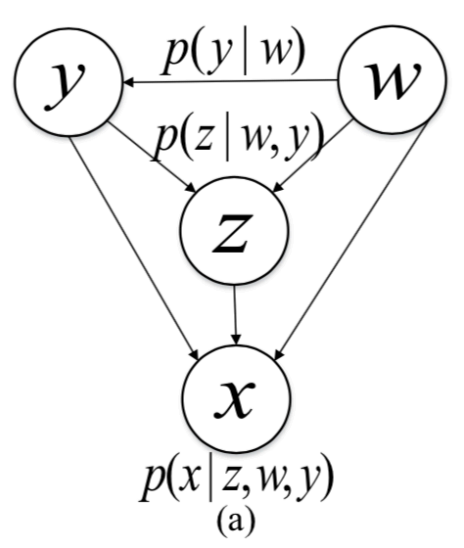


(a)



(b)

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION



SCPG - Holistic Sentiment Control Module

- Labelled data

Reconstruction Loss

$$\begin{aligned}\log\ p(x,y|w) &\geq \boxed{\mathbb{E}_{q(z|x,w,y)}[\log\ p(x|z,w,y)]} \\ &\boxed{-KL[q(z|x,w,y)||p(z|w,y)]} + \boxed{\log\ p(y|w)} \\ &= -\mathcal{L}(x,y,w)\end{aligned}$$

KL Loss

Classifier Loss

- Unlabelled data

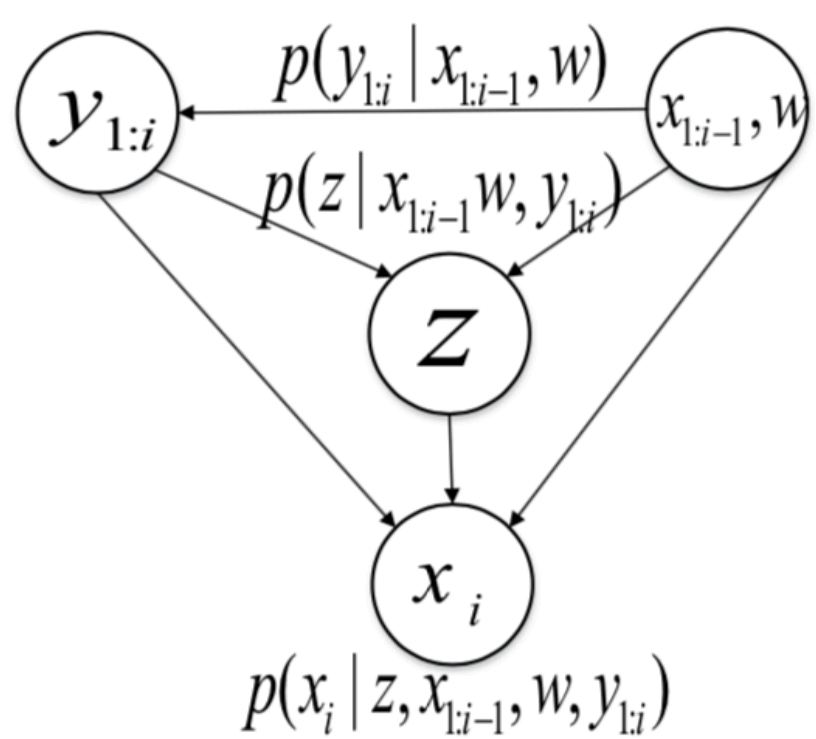
$$\begin{aligned}\log p(x|w) &= \iint q(y,z|x,w) \log\ p(x|w) dy dz \\ &\geq \mathbb{E}_{q(y|x,w)}[-\mathcal{L}(x,y,w) \boxed{-\log\ q(y|x,w)]} \\ &= -\mathcal{U}(x,w).\end{aligned}$$

classifier: simultaneously trained to sample sentiments for unlabelled poems

IJCAI 2019

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION

SCPG - Temporal Sentiment Control Module



(b)

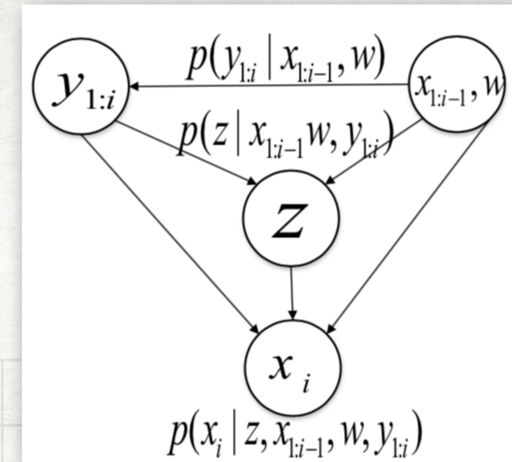
- Each line xi is generated by the decoder p(xi | z, x1:i−1, w, y1:i)

  (conditioned on previous content sequence and sentiment sequence)

- Each yi is predicted by a time sequence classifier p(yi | x1:i−1, w, y1:i−1)

- The content sequence and the sentiment sequence are interactively generated

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION



(b)

SCPG - Temporal Sentiment Control Module

- Labelled data

$$\begin{aligned}\log p(x,y|w) &\geq \mathbb{E}_{q(z|x,w,y)}[\log p(x|z,w,y)] \\ &- KL[q(z|x,w,y)||p(z|w,y)] + \log p(y|w) \\ &= -\mathcal{L}(x,y,w)\end{aligned}$$

Condition: x1:i−1, y1:i−1, w

log p(xi, yi | x1:i−1, y1:i−1, w)

−L(x1:i, y1:i, w)

- Unlabelled data

$$-\mathcal{U}(x_{1:i},w) = \frac{1}{M}\sum_{k=1}^{M}[-\mathcal{L}(x_{1:i},y_{1:i},w) - \boxed{\log q(y_{1:i}|x_{1:i},w)]}$$

where $q(y_{1:i}|x_{1:i},w)$ can be factorized as:

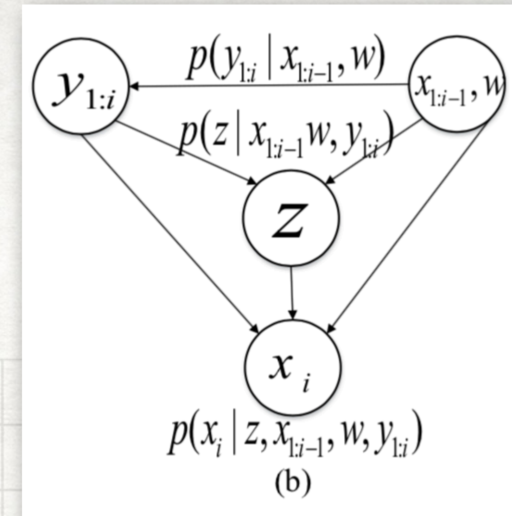
$$q(y_1|x_{1:i},w)q(y_2|x_{1:i},w,y_1)\ldots q(y_i|x_{1:i},w,y_{1:i-1})$$

Assume future content has no influence on current or past sentiment

time sequence classifier: predicting each yi

IJCAI 2019

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION



SCPG - Temporal Sentiment Control Module

- Unlabelled data

time sequence classifier: predicting each yi

$$-\mathcal{U}(x_{1:i}, w) = \frac{1}{M}\sum_{k=1}^{M}[-\mathcal{L}(x_{1:i}, y_{1:i}, w) - \boxed{\log q(y_{1:i}|x_{1:i}, w)}]$$

time sequence classifier:

$$c_i = h_1(c_{i-1}, x_i), c_0 = f_1(w)$$
$$m_i = h_2(m_{i-1}, c_{i-1}, y_{i-1}), m_0 = f_2(w), y_0 = y$$
$$q(y_i|x_{1:i}, w, y_{1:i-1}) = softmax(m_i),$$

- C: context
- Y: sentiment
- M: hidden state

IJCAI 2019

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION

Other Details for the SCPG:

Model Structure: CVAE-Conditional Variational AutoEnocder

Encoder / Decoder: Bidirectional GRU (initial decoder state to $s0 = f(z, y, w)$,

where f is a non-linear layer, to involve the sentiment and keyword.)

Latent variable z: a isotropic Gaussian distribution

When training, z is sampled from $q(z|x, w, y)$ and when testing z is sampled from

$p(z|w, y)$.

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION

Experiments:

Models	Flu.	Coh.	Mea.	Poe.	Ove.
Models without sentiment controlling module					
WM	3.09	2.86	2.8	2.79	2.71
CVAE	2.50	2.41	2.33	2.39	2.24
MRL	3.20	2.96	2.88	**2.95**	**2.86**
Models with sentiment controlling module					
SBasic	2.21	2.09	1.99	1.96	1.87
SCPG-H	3.07	2.82	2.81	2.72	2.65
SCPG-T	**3.23**	2.93	2.88	2.87	2.78
SCPG-HT	**3.23**	**3.04**	**2.92**	2.83	2.78
GT	4.03	4.13	4.00	3.90	3.83

Human evaluation

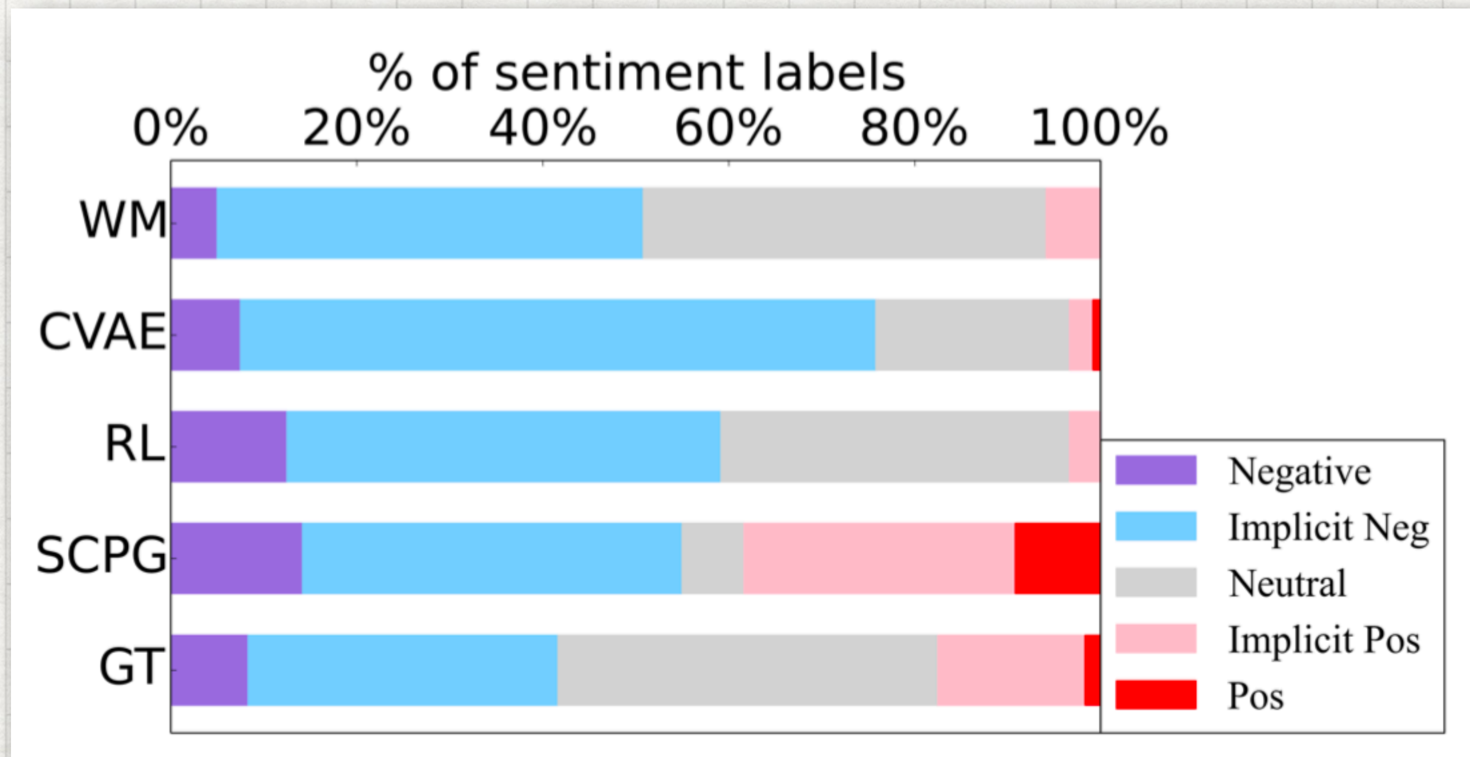
Models	Jaccard Similarity
WM	3.3%
CVAE	1.8%
MRL	**1.2%**
SCPG-HT	1.5%
GT	0.05%

Semantic Diversity

Baseline:

- WM - working memory model
- CVAE - conditional variational autoen- coder
- MRL - reinforcement learning framework

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION



Sentiment Diversity

Models	5 Classes	3 Classes
Holistic Sentiment Accuracy		
SBasic	0.412	0.633
SCPG-HT	0.461	0.733
Line Sentiment Accuracy		
SCPG-T	0.379	0.550
SCPG-HT	0.441	0.637

sentiment control accuracy

SENTIMENT-CONTROLLABLE CHINESE POETRY GENERATION

泠泠流水声， The water releasing sound is flowing beside;	泠泠玉立水中央， In the center of a pool flowing with the sound I can see her gracefully standing;
窈窕小窗明。 Through the small window the thin moonlight is shed inside.	翠羽啁嘈入曲廊。 At the corridor the bird with the beautiful feather I can hear of its chanting.
夜静无人语， I have no one to talk with at the lonely and tranquil night;	万里相逢真乐事， Have been thousands of miles apart, our meeting makes me really pleasant and thrilled;
心惊白发生。 Frighten and sorrow in mind, my hair has gradually become white.	四时佳气蔼嘉祥。 In every minute and every corner is the auspicious and joyful atmosphere filled.

negative | implicit negative | neutral | implicit positive | positive

Poems generated by holistic sentiments "implicit neg" and "pos" respectively given the same keyword "sound of water".

LONG AND DIVERSE TEXT GENERATION WITH PLANNING-BASED HIERARCHICAL VARIATIONAL MODEL

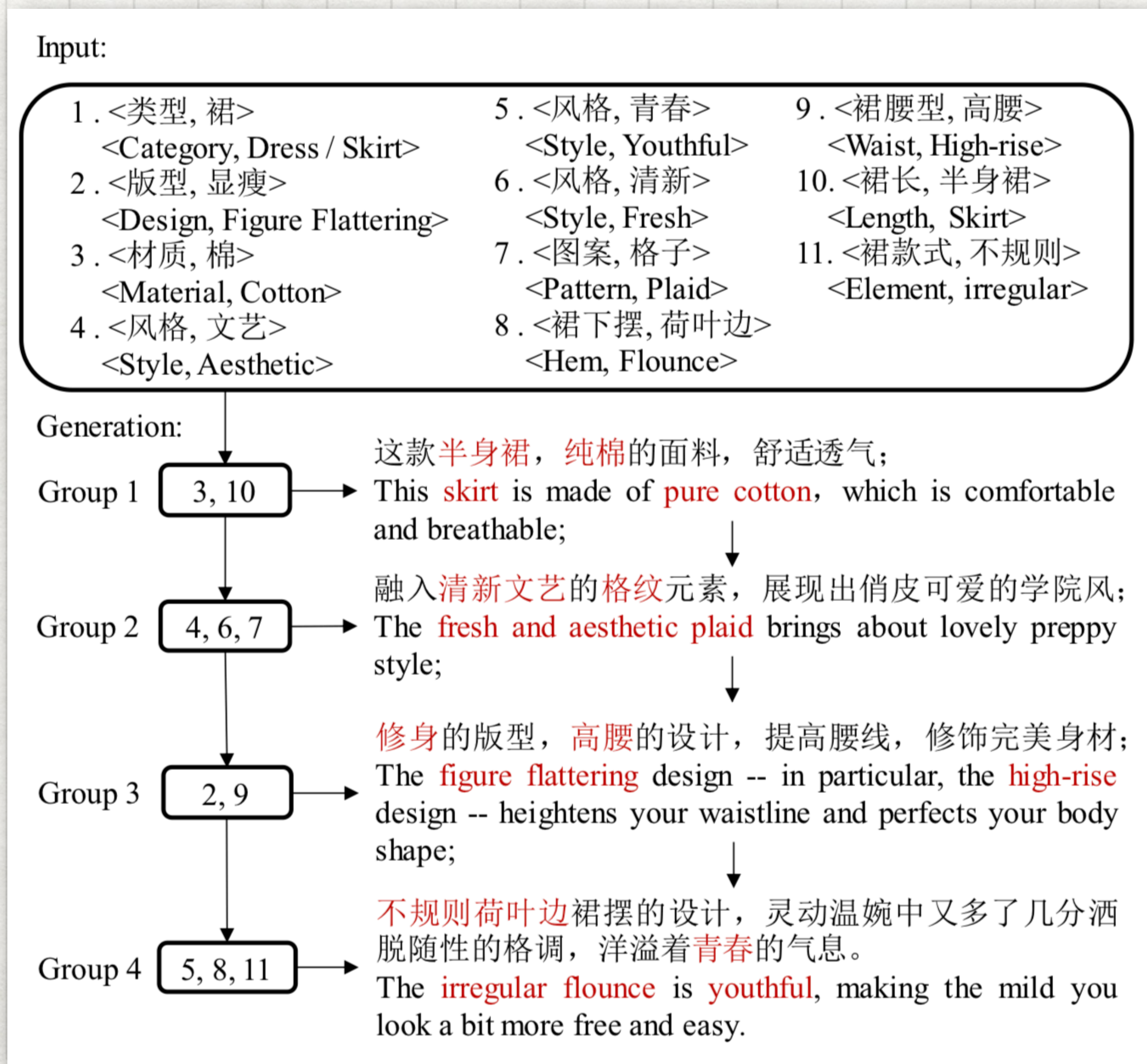
Long and Diverse Text Generation with Planning-based Hierarchical Variational Model

Zhihong Shao[1], Minlie Huang[1]*, Jiangtao Wen[1], Wenfei Xu[2], Xiaoyan Zhu[1]
[1] Institute for Artificial Intelligence, State Key Lab of Intelligent Technology and Systems
[1] Beijing National Research Center for Information Science and Technology
[1] Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
[2] Baozun, Shanghai, China
szh19@mails.tsinghua.edu.cn, aihuang@tsinghua.edu.cn
jtwen@tsinghua.edu.cn, xuwenfeilittle@gmail.com, zxy-dcs@tsinghua.edu.cn

EMNLP 2019

LONG AND DIVERSE TEXT GENERATION WITH PLANNING-BASED HIERARCHICAL VARIATIONAL MODEL

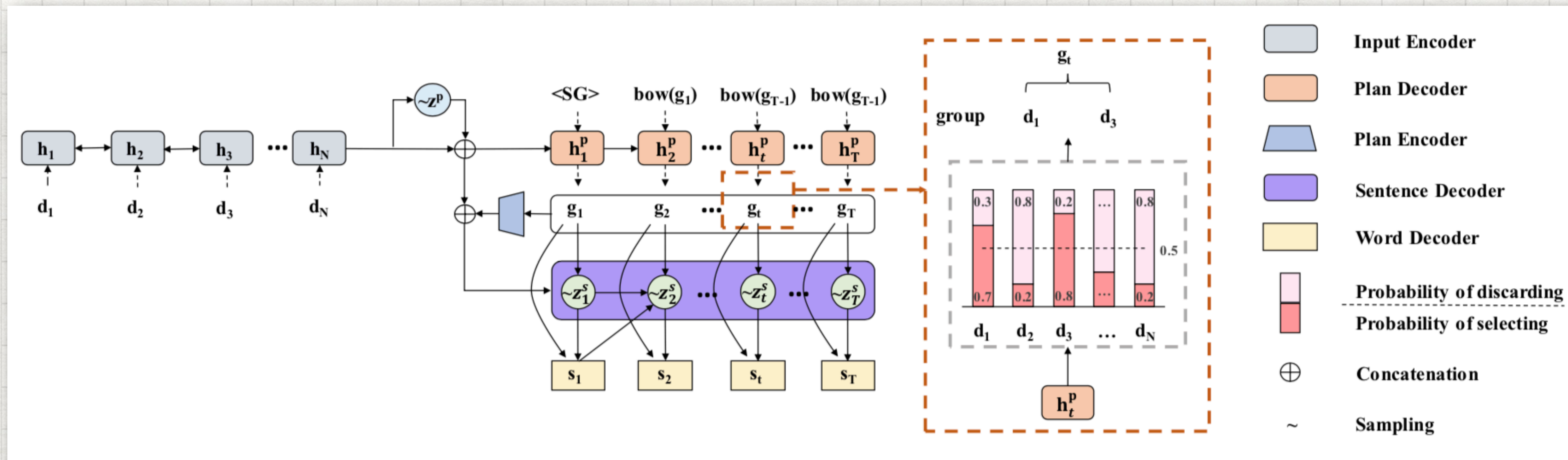
Aiming to boost up the <u>coherency</u> and <u>diversity</u> issue of Data-to-text generation tasks.
Proposed a Planning-based Hierarchical Variational Model (PHVM)



1. Encode a list of input attribute-value pairs
2. Generate a sequence of groups, each of which is a subset of input items
3. Each sentence is then generated conditioned on the corresponding group and its previous generated sentences.

EMNLP 2019

LONG AND DIVERSE TEXT GENERATION WITH PLANNING-BASED HIERARCHICAL VARIATIONAL MODEL



- Given input data x = {d1, d2, ..., dN }, generate a long and diverse text y = s1 s2 ...sT that covers x as much as possible.
- Advertising text generation task: x consists of specifications about a product where each di is an attribute-value pair < ai , vi >.
- Recipe text generation task: x is an ingredient list where each di is an ingredient

EMNLP 2019

LONG AND DIVERSE TEXT GENERATION WITH PLANNING-BASED HIERARCHICAL VARIATIONAL MODEL

Input		
1 . <类型, 裙> <Category, Dress / Skirt> 2 . <版型, 显瘦> <Design, Figure Flattering> 3 . <材质, 蕾丝> <Material, Lace> 4 . <颜色, 黑色> <Color, Black>	5 . <风格, 简约> <Style, Minimalist> 6 . <风格, 性感> <Style, Sexy> 7 . <裙型, A字> <Shape, A-line> 8 . <裙长, 长裙> <Length, Long>	9 . <裙袖长, 七分袖> <Sleeve Length, Three-quarter Sleeve> 10. <裙领型, 圆领> <Collar, Round> 11. <裙款式, 拼接> <Element, Stitching>

Methods	Missing Pairs	texts
Checklist	11	这款黑色蕾丝长裙，简约的圆领设计，修饰颈部曲线，性感迷人。 This dress with black lace, the minimalist round collar flatters the curve of your neck which is sexy and attractive. 七分袖设计，修饰手臂曲线，更显纤细修长。 The three-quarter sleeves flatter the curves of your arms, making your arms look slender. A字版型，遮肉显瘦，不挑身材，适合各种身材。

Advertising Text Generation

Title:	Drunken apple/pumpkin pie	
Ingredients:	2 eggs 3/4 c brown sugar ; firmly packed 1 9-inch pie shell ; unbaked 1 c cooked pumpkin - mashed and drained 1 c applesauce ; thick & chunky 1 tb flour 1/2 ts salt	1 ts each cinnamon and ginger 1/4 ts nutmeg 1/8 ts each allspice and cloves 1 1/2 c half-and-half or vaporated milk 1 ts vanilla 1 c pecan halves 2 tb rum
Truth	In a bowl, beat together the eggs and sugar until light. Mix in the pumpkin, applesauce, flour, salt, cinnamon, ginger, nutmeg, allspice, cloves, half-and-half and vanilla; Blend thoroughly. Pour into pie shell. Arrange pecan halves over top of filling. And bake 30 to 35 minutes longer , or until filling is firm and a knife inserted 1 " from the edge comes out clean. Cool on a wire rack. At serving time, warm rum in a small container suitable for pouring. Light the rum with a match and pour immediately while flaming over the pie.	

Recipe Generation

LONG AND DIVERSE TEXT GENERATION WITH PLANNING-BASED HIERARCHICAL VARIATIONAL MODEL

II. Plan the inputs for the sentences

Use a plan GRU decoder to calculate

$argmax_{g_t} P(g_t | g_{<t}, x, z^p)$

- At time step t, the plan decoder makes a binary prediction for each input item

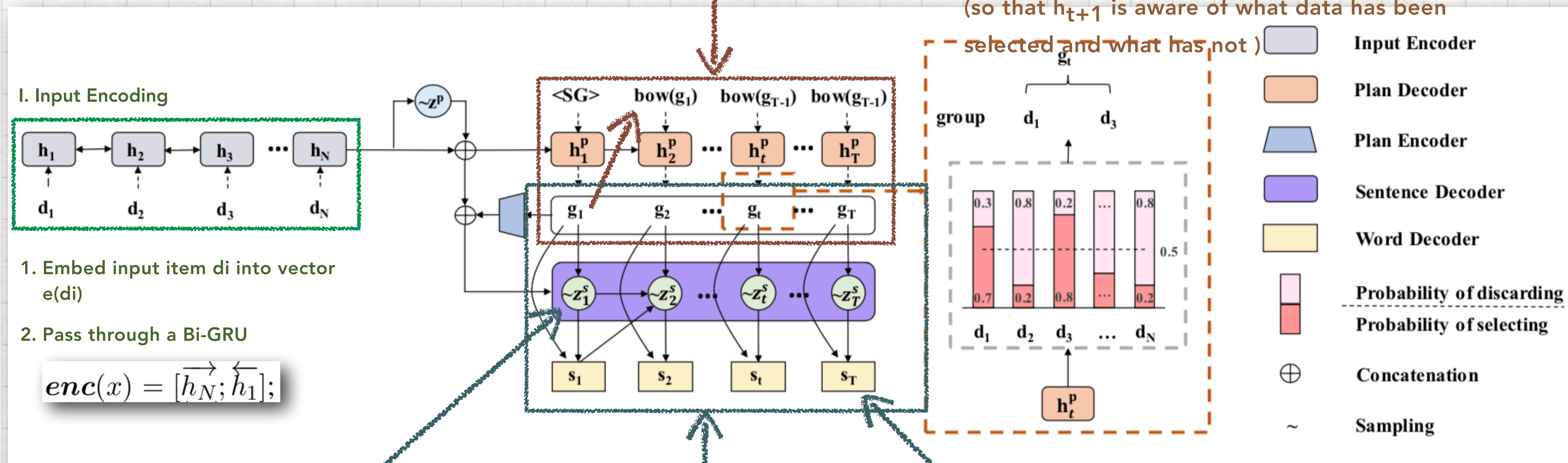
$P(d_i \in g_t) = \boldsymbol{\sigma}(v_p^T \boldsymbol{tanh}(W_p[h_i; h_t^p] + b_p))$

- If p() > 0.5 => select the input
- Previous generated gt-1 is feed to the decoder (so that h_{t+1} is aware of what data has been selected and what has not)

I. Input Encoding

1. Embed input item di into vector e(di)
2. Pass through a Bi-GRU

$\boldsymbol{enc}(x) = [\overrightarrow{h_N}; \overleftarrow{h_1}];$



- Sentence Level Generation

$h_t^s = \boldsymbol{GRU}_s([z_{t-1}^s; h_{t-1}^w], h_{t-1}^s)$

III. Hierarchical Generation Process

$c = \{x, z^p\}$

$y = argmax_y P(y|g, c)$

- Word Level Generation

$w_k^t \sim P(w_k^t | w_{<k}^t, s_{<t}, z_t^s, g, c)$

EMNLP 2019

LONG AND DIVERSE TEXT GENERATION WITH PLANNING-BASED HIERARCHICAL VARIATIONAL MODEL

Advertising Text Generation:

Models	BLEU (%)	Coverage (%)	Length	Distinct-4 (%)	Repetition-4 (%)
Checklist	**4.17**	84.52**	83.61**	21.95**	46.40**
CVAE	4.02	77.65**	80.96**	41.69**	36.58**
Pointer-S2S	3.88**	85.97**	74.88**	18.16**	36.78**
Link-S2S	3.90**	70.49**	**95.65**	16.64**	59.83**
PHVM (ours)	2.85**	**87.05**	89.20**	**72.87**	**3.90**
w/o z^p	3.07**	84.74**	91.97**	70.51**	4.19
w/o z_t^s	3.38**	84.89**	75.28**	42.32**	20.88**

Recipe Generation:

Models	BLEU (%)	Coverage (%)	Length	Distinct-4 (%)	Repetition-4 (%)
Checklist §	3.0	67.9	N/A	N/A	N/A
Checklist	2.6**	66.9*	67.59	30.67**	39.1**
CVAE	4.6	63.0**	57.49**	52.53**	38.7**
Pointer-S2S	4.3	70.4**	59.18**	30.72**	36.4**
Link-S2S	1.9**	53.8**	40.34**	24.93**	31.6**
PHVM (ours)	**4.6**	**73.2**	**70.92**	**67.86**	**17.3**

EMNLP 2019

LONG AND DIVERSE TEXT GENERATION WITH PLANNING-BASED HIERARCHICAL VARIATIONAL MODEL

Manual Evaluation

Models	Grammaticality			κ	Coherence			κ
	Win (%)	Lose (%)	Tie (%)		Win (%)	Lose (%)	Tie (%)	
PHVM vs. Checklist	59.0**	23.5	17.5	0.484	54.5*	42.5	3.0	0.425
PHVM vs. CVAE	69.5**	13.5	17.0	0.534	60.0**	37.0	3.0	0.426
PHVM vs. Pointer-S2S	76.5**	17.0	6.5	0.544	56.5**	39.0	4.5	0.414
PHVM vs. Link-S2S	66.0**	28.5	5.5	0.462	62.5**	31.5	6.0	0.415

SYNTAX-INFUSED VARIATIONAL AUTO-ENCODER FOR TEXT GENERATION

Syntax-Infused Variational Autoencoder for Text Generation

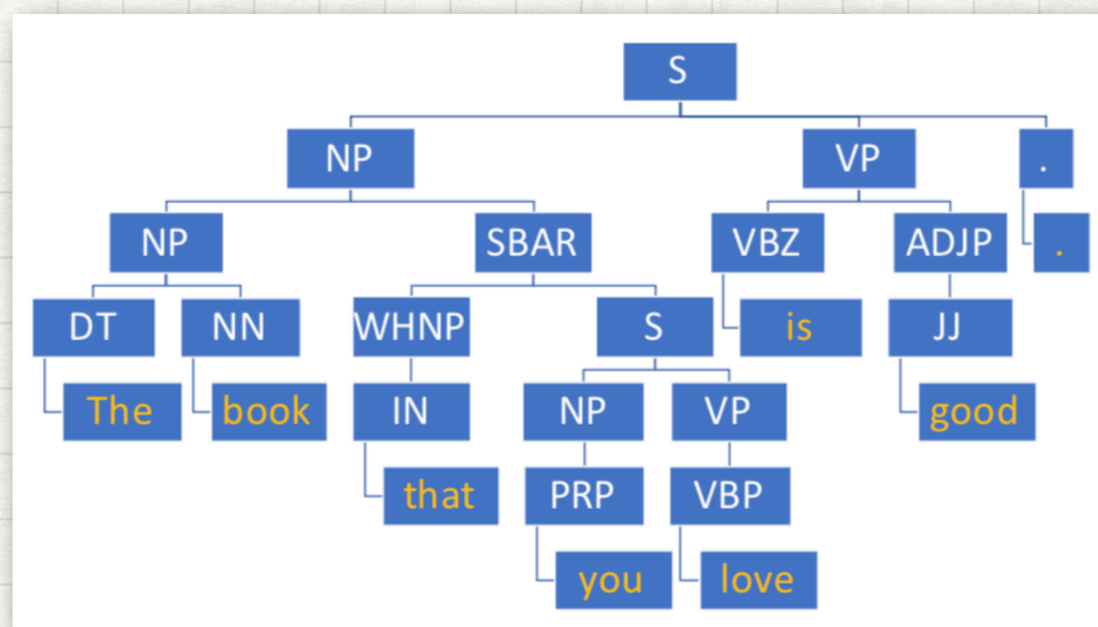
Xinyuan Zhang[1,*], **Yi Yang**[2,*], **Siyang Yuan**[1], **Dinghan Shen**[1], **Lawrence Carin**[1]
[1]Duke University
[2]ASAPP Inc.
xy.zhang@duke.edu, yyang@asapp.com

ACL 2019

SYNTAX-INFUSED VARIATIONAL AUTO-ENCODER FOR TEXT GENERATION

Motivation: VAE generation models tends to generate ungrammatical sentences

Utilise augmented data by introducing the syntactic tree, to enrich the latent representation and make the generated sentences more grammatical and fluent.



To simplify the encoding and decoding processes, remove leaf nodes and linearise the bracketed parse structure into a syntactic tree sequence

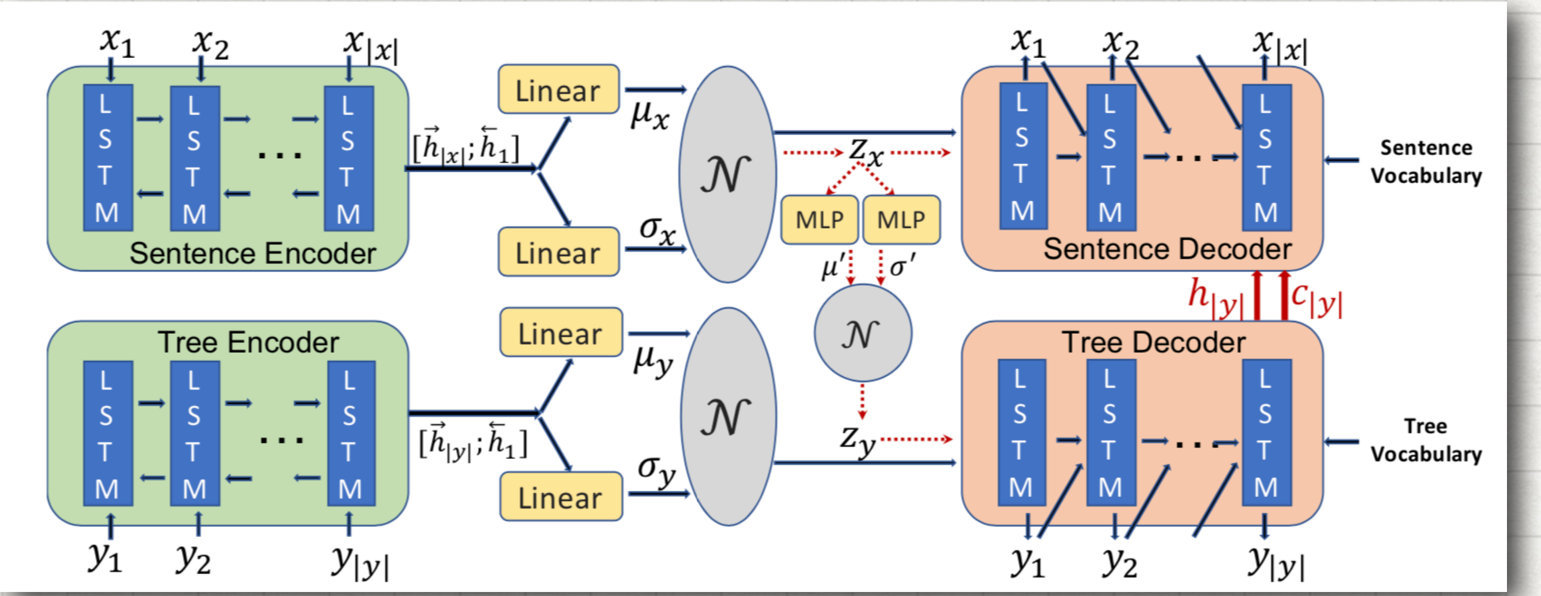
(S(NP(NP(DT)(NN))(SBAR(WHNP(IN))(S(NP(PRP))(VP(VBP)))))(VP(VBZ)(ADJP(JJ)))(.)).

ACL 2019

SYNTAX-INFUSED VARIATIONAL AUTO-ENCODER FOR TEXT GENERATION

Syntax-infused VAE model

Problem Formulation:



Given a sentence x and syntactic tree y, the goal is to jointly encode x and y into latent representations $z_x \in R^d$ and $z_y \in R^d$, and then decode them jointly from the two latent spaces

Becuz current VAE-based language models cannot accommodate two separate latent spaces for z_x and z_y.

To incorporate x, y, z_x, and z_y in one VAE framework, the objective needs to be redesigned to optimize the log joint likelihood log p(x, y).

ACL 2019

SYNTAX-INFUSED VARIATIONAL AUTO-ENCODER FOR TEXT GENERATION

SIVAE-c

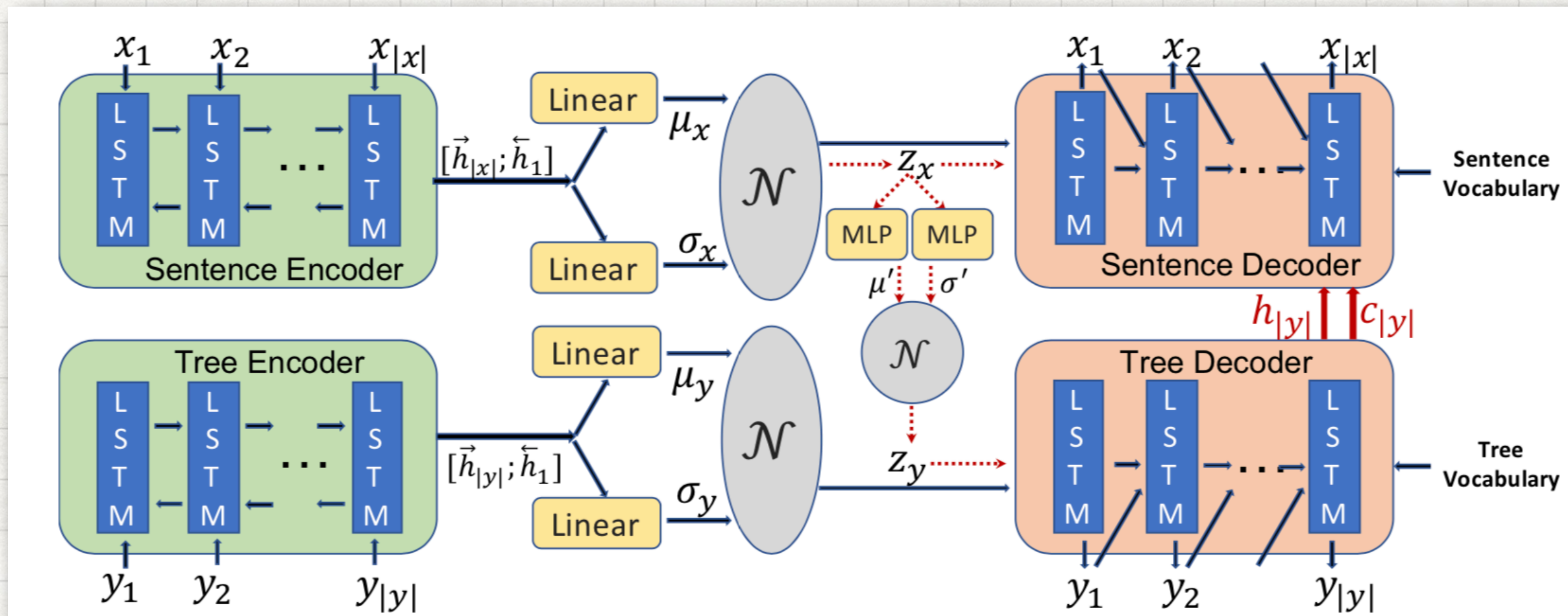
directly capturing the dependencies between z_x and z_y, presumes that semantic information should influence syntax structure

SIVAE-i

generates sentences and syntactic trees assuming the priors $p(z_x)$ and $p(z_y)$ are independent.

SYNTAX-INFUSED VARIATIONAL AUTO-ENCODER FOR TEXT GENERATION

SIVAE-c: Modeling Syntax-Semantics Dependencies



$$p(\boldsymbol{x}, \boldsymbol{y}) = \int_{d_{\boldsymbol{z}_x}} \int_{d_{\boldsymbol{z}_y}} p(\boldsymbol{x}|\boldsymbol{y}, \boldsymbol{z}_x) p(\boldsymbol{y}|\boldsymbol{z}_x, \boldsymbol{z}_y) \cdot p(\boldsymbol{z}_y|\boldsymbol{z}_x) p(\boldsymbol{z}_x) d_{\boldsymbol{z}_y} d_{\boldsymbol{z}_x},$$

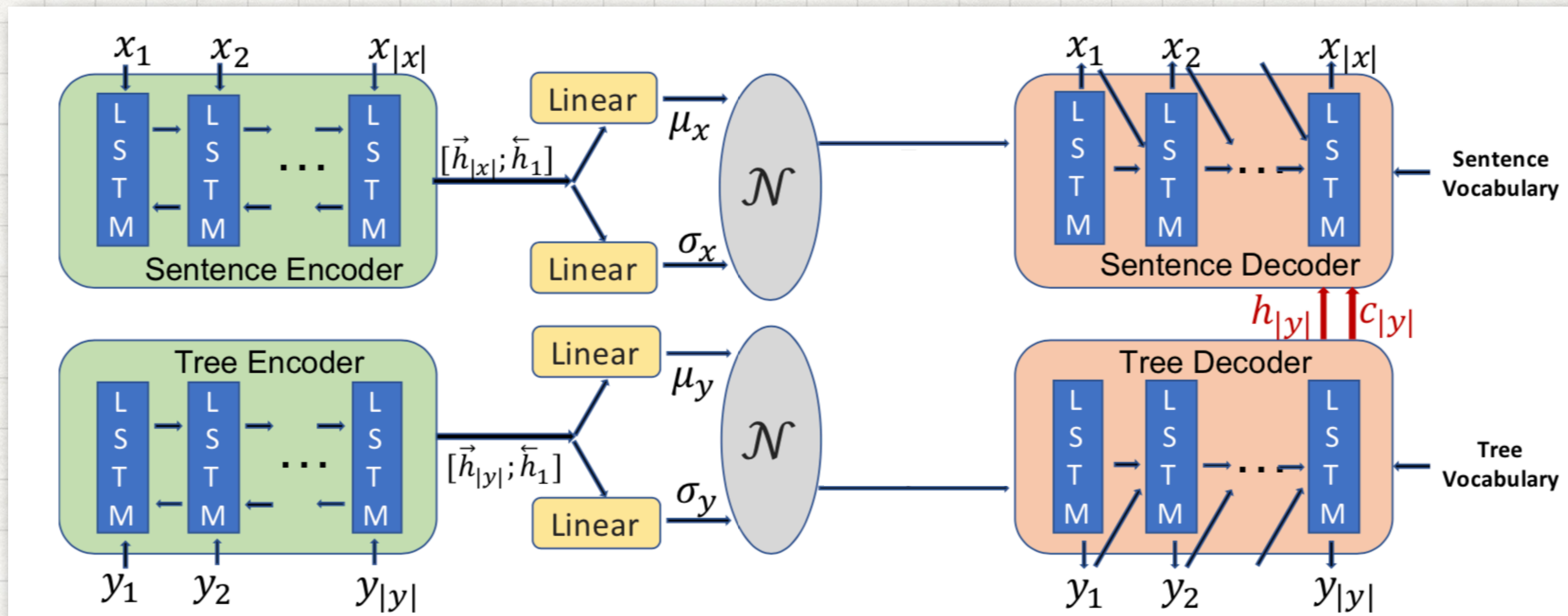
Loss:

$$\log p(\boldsymbol{x}, \boldsymbol{y}) \geq \mathcal{L}(\boldsymbol{x}, \boldsymbol{y}; \theta, \phi, \psi) = \quad (2)$$
$$\boxed{\mathbb{E}_{q_\phi(\boldsymbol{z}_x|\boldsymbol{x})} \log p_\theta(\boldsymbol{x}|\boldsymbol{y}, \boldsymbol{z}_x) - \mathrm{KL}[q_\phi(\boldsymbol{z}_x|\boldsymbol{x}) || p(\boldsymbol{z}_x)]}$$
$$\boxed{+ \mathbb{E}_{q_\phi(\boldsymbol{z}_y|\boldsymbol{y}, \boldsymbol{z}_x)} \log p_\theta(\boldsymbol{y}|\boldsymbol{z}_y) - \mathrm{KL}[q_\phi(\boldsymbol{z}_y|\boldsymbol{y}, \boldsymbol{z}_x) || p_\psi(\boldsymbol{z}_y|\boldsymbol{z}_x)],}$$

ACL 2019

SYNTAX-INFUSED VARIATIONAL AUTO-ENCODER FOR TEXT GENERATION

SIVAE-i:



$$\log p(\boldsymbol{x}, \boldsymbol{y}) \geq \mathcal{L}(\boldsymbol{x}, \boldsymbol{y}; \theta, \phi) = \tag{9}$$
$$\mathbb{E}_{q_\phi(\boldsymbol{z}_x|\boldsymbol{x})} \log p_\theta(\boldsymbol{x}|\boldsymbol{y}, \boldsymbol{z}_x) - \mathrm{KL}[q_\phi(\boldsymbol{z}_x|\boldsymbol{x}) \| p(\boldsymbol{z}_x)]$$
$$+ \mathbb{E}_{q_\phi(\boldsymbol{z}_y|\boldsymbol{y})} \log p_\theta(\boldsymbol{y}|\boldsymbol{z}_y) - \mathrm{KL}[q_\phi(\boldsymbol{z}_y|\boldsymbol{y}) \| p(\boldsymbol{z}_y)].$$

ACL 2019

SYNTAX-INFUSED VARIATIONAL AUTO-ENCODER FOR TEXT GENERATION

This independence assumption manifests syntactically-controlled sentence generation as it allows to **alter the syntactic structure**, desirable for related tasks like **paraphrase generation**.

1) Encode the original sentence to z_x;
2) Select and encode a syntactic template into z_y;
3) Generate the reconstructed syntactic sequence y from $p(y|z_y)$;
4) Generate the paraphrase of the original sentence that conforms to y from $p(x|y, z_x)$.

SYNTAX-INFUSED VARIATIONAL AUTO-ENCODER FOR TEXT GENERATION

Model	PTB					
	Standard			Inputless		
	PPL	NLL	KL	PPL	NLL	KL
KN5	145	132	-	593	169	-
LSTM-LM	110	124	-	520	165	-
VAE	112	125	2	317	153	13
SIVAE-c	98(**1.6**)	121(**53**)	5(0.5)	286(**2.4**)	150(**99**)	17(1.3)
SIVAE-i	**90**(1.7)	**119**(60)	9(1.0)	**261**(2.6)	**147**(108)	24(2.5)

Decoder purely relies on the latent representations without any use of prior words

Model	wiki90M					
	Standard			Inputless		
	PPL	NLL	KL	PPL	NLL	KL
KN5	141	141	-	588	182	-
LSTM-LM	105	133	-	521	179	-
VAE	106	133	5	308	164	22
SIVAE-c	94(**1.6**)	130(**56**)	12(1.0)	278(**2.3**)	161(**99**)	29(2.4)
SIVAE-i	**89**(1.7)	**128**(63)	16(1.9)	**256**(2.4)	**158**(104)	36(5.1)

ACL 2019

SYNTAX-INFUSED VARIATIONAL AUTO-ENCODER FOR TEXT GENERATION

Targeted Syntactic Evaluation: to examine whether the proposed methods improve th grammar of generated sentences

Subject-verb agreement (SVA): Third-person present English verbs need to agree with the number of their subjects.

For example, *simple SVA*:

(a). The author laughs.

(b). *The author laugh.

Reflexive anaphoras (RA): A reflective pronoun such as *himself* needs to agree in number (and gender) with its antecedent.

For example, *simple RA*:

(a). The senators embarrassed themselves.

(b). *The senators embarrassed herself.

Negative polarity items (NPI): Words like *any* and *ever* that can only be used in the scope of negation are negative polarity items.

For example, *simple NPI*:

(a). No students have ever lived here.

(b). *Most students have ever lived here.

Model	SVA		RA		NPI	
	S	C	S	C	S	C
Humans	0.96	0.85	0.96	0.87	0.98	0.81
KN5	0.79	0.50	0.50	0.50	0.50	0.50
LSTM-LM	0.94	0.56	0.83	0.55	0.50	0.50
VAE	0.94	0.57	0.84	0.57	0.51	0.50
SIVAE-c	**0.97**	**0.75**	**0.89**	**0.64**	**0.57**	**0.52**
SIVAE-i	0.95	0.71	0.88	0.63	0.56	0.52

S: simple; C:complex

ACL 2019

SYNTAX-INFUSED VARIATIONAL AUTO-ENCODER FOR TEXT GENERATION

Paraphrase Generation

Template	Paraphrase
original	the discovery of dinosaurs has long been accompanied by a legend .
(SBARQ (NP) (VP) (,) (SQ) (?))	the discovery of dinosaurs has been a legend , is it ?
(S (") (NP) (VP) (") (NP) (VP) (.))	" the discovery of dinosaurs is a legend " he said .
(S (VP) (,) (NP) (.))	having been accompanied , the unk lengend .
original	in 1987 a clock tower and a fountain were erected at council unk monument .
(S (PP) (PP) (NP) (VP) (.))	in 1987 at council a fountain was erected .
(S (VP) (NP) (CC) (NP) (PP) (.))	build a clock and a fountain at council unk unk .
(S (NP) (;) (S) (PP) (.))	a clock p ; he shops everything on the fountain at unk unk .

Table 4: Examples of syntactically controlled paraphrases generated by SIVAE-i. We show two successful and one failed (in blue) generations with different templates for each input sentence.

Ori	the new york times has been one of the best selling newspapers in america .
Gen1	the new york times also has been used as american best selling newspaper .
Gen2	the new york times also has been used as a " unk " that sells in america .
Gen3	the new york times also has been used as the best " unk " selling in america .

Table 5: An example of paraphrases generated by SIVAE-c.

Model	PTB			wiki90M		
	Rele	Read	Div	Rele	Read	Div
VAE	2.63	3.07	2.77	3.03	3.20	2.60
SIVAE-c	2.93	3.47	2.80	3.27	3.67	2.73
SIVAE-i	3.00	3.30	3.13	3.37	3.53	3.20

Table 6: Human evaluation results on Relevance, Readability, and Diversity of generated paraphrases.

ACL 2019

Thanks for Listening !