Product-Aware Answer Generation in E-Commerce Question-Answering

Advisor:JIA-LING, KOH

Source:WSDM ’19

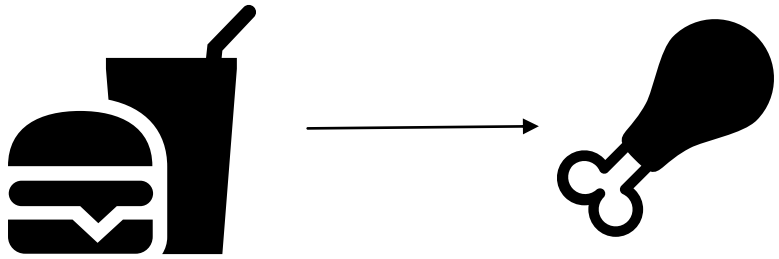
SPEAKER:rui ze Fang

DATE: 2020/12/08

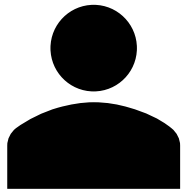
Outline

- Introduction
- Method: PAAG
- Experiment
- Conclusion

Introduction



Question:
這間速食店的炸雞好吃嗎？



Answer:
這間速食店的炸雞很好吃並且有多好吃...

Review

Introduction

- most of existing reading comprehension solutions only extract text spans from contextual passages
- It’s extremely expensive to label large amounts of explicit text spans

  (S-Net)
- Traditional loss function calculation in reading comprehension tends to generate meaningless answers such as “I don’t know”.(RAGE)

Problem Define

- Question : X^q=$\{x_1^q, x_2^q, \ldots, x_{T_q}^q\}$
- Review : X^r=$\{x_1^r, x_2^r, \ldots, x_{T_r}^r\}$
- Attribute : A =$\{(a_1^k, a_1^v), (a_2^k, a_2^v), \ldots, (a_{T_a}^k, a_{T_a}^v)\}$

  , a_i^k name of i-th attribute , a_i^v attribute content

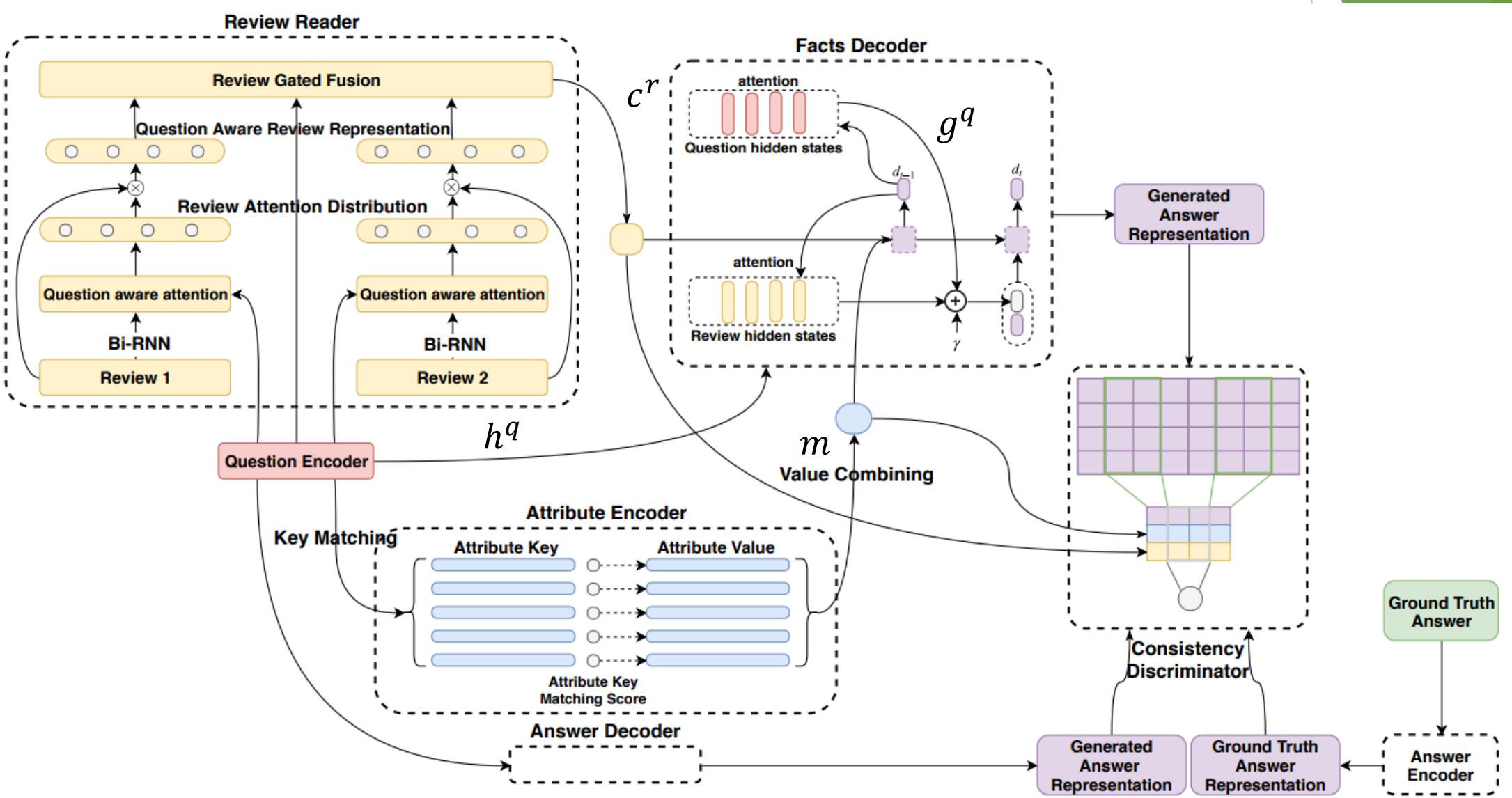
- $\hat{Y}$ = $\{\widehat{y_1}, \widehat{y_2}, \ldots, \widehat{y_{T_y}}\}$, answer of generate
- Target: maximize P(Y| X^q, X^r,A) = $\prod_{t=1}^{T_y} P(y_t|X^q, X^r, A)$

  Y = $\{y_1, y_2, \ldots, y_{T_y}\}$ ground truth answer

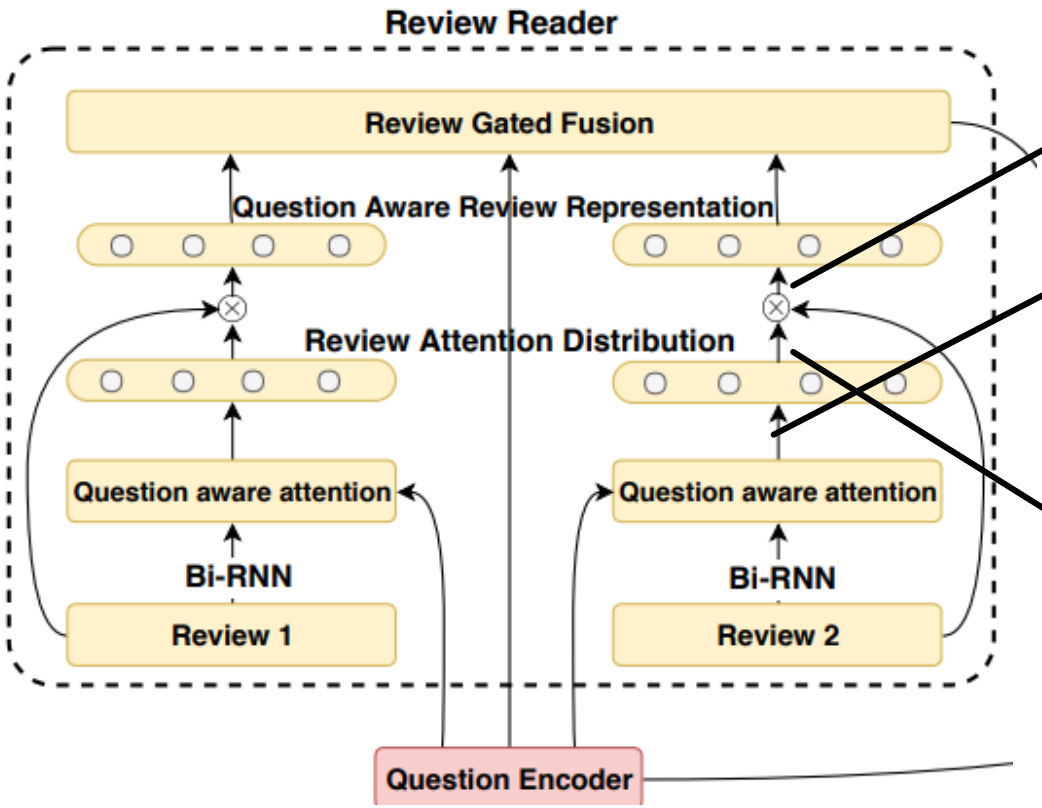
Outline

- Introduction
- Method: PAAG
- Experiment
- Conclusion

Method


Review Reader
Review Gated Fusion
Question Aware Review Representation
Review Attention Distribution
Question aware attention
Question aware attention
Bi-RNN
Bi-RNN
Review 1
Review 2
Question Encoder
h^q
c^r
Facts Decoder
attention
Question hidden states
g^q
attention
Review hidden states
γ
Generated Answer Representation
m
Value Combining
Attribute Encoder
Key Matching
Attribute Key
Attribute Value
Attribute Key Matching Score
Answer Decoder
Consistency Discriminator
Ground Truth Answer
Generated Answer Representation
Ground Truth Answer Representation
Answer Encoder

Method-
Review reader



Value

$$c_i^r = \sum_{t=1}^{T_r^i} \alpha_{i,t} h_{i,t}^r,$$

Query

Key

$$s_{i,j}^k = v^\intercal \tanh(W_q h_k^q + W_r h_{i,j}^r),$$

$$s_{i,j} = \max(s_{i,j}^1, s_{i,j}^2, \ldots, s_{i,j}^{T_q}),$$

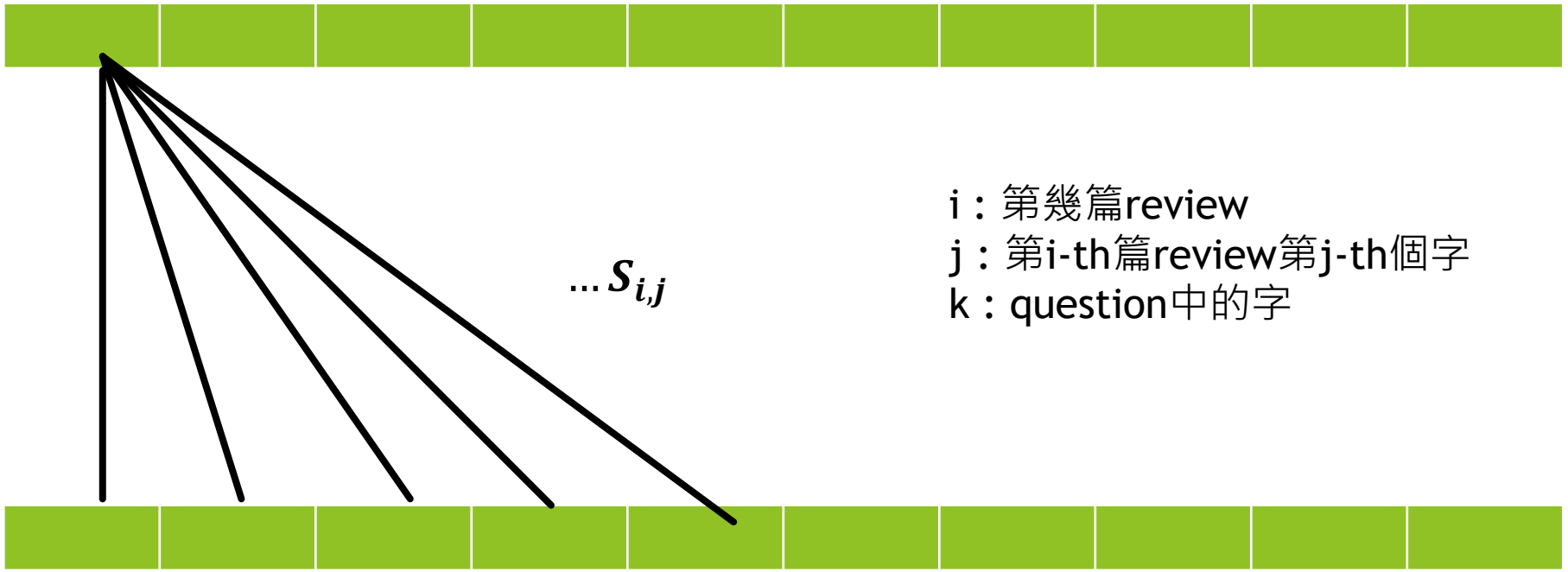
$$\alpha_{i,j} = \exp(s_{i,j}) / \sum_{t=1}^{T_r^i} \exp(s_{i,t}),$$

$$h_t^q = \text{Bi-RNN}_q(e(x_t^q), h_{t-1}^q),$$

$$h_{i,t}^r = \text{Bi-RNN}_r(e(x_{i,t}^r), h_{i,t-1}^r).$$

Attention

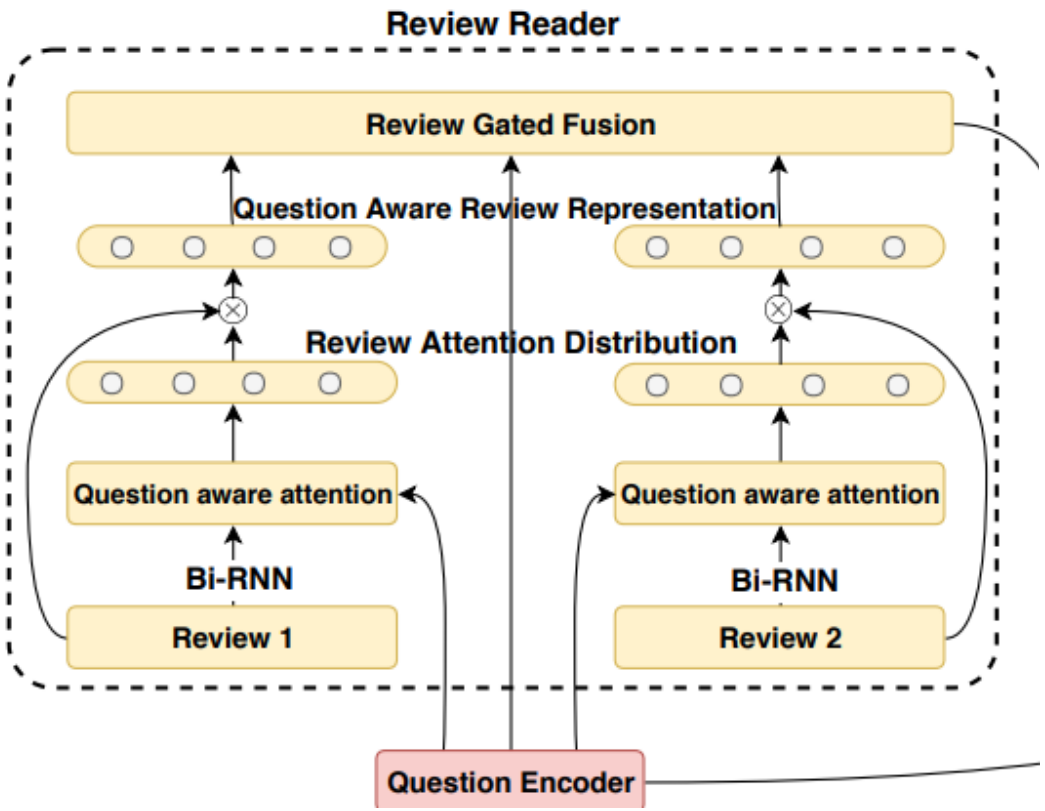
Review ($h_{i,j}^{q}$)



i：第幾篇review
j：第i-th篇review第j-th個字
k：question中的字

Question(h_{k}^{q})

Method-
Review reader

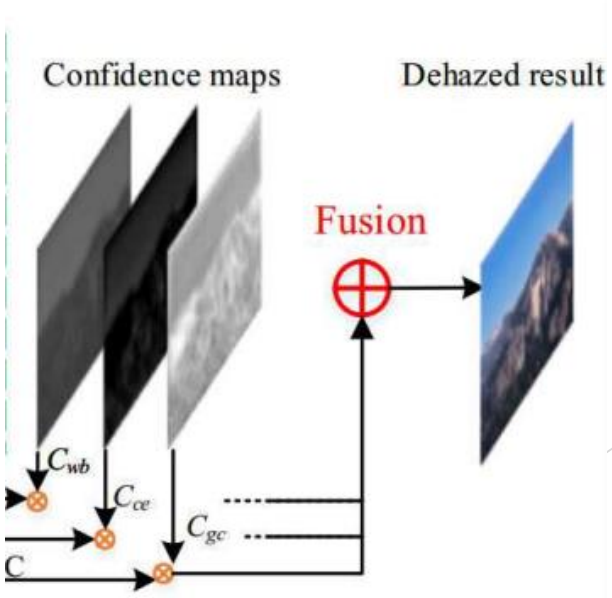


$$u_i = c_i^r W_f h_{T_q}^q,$$

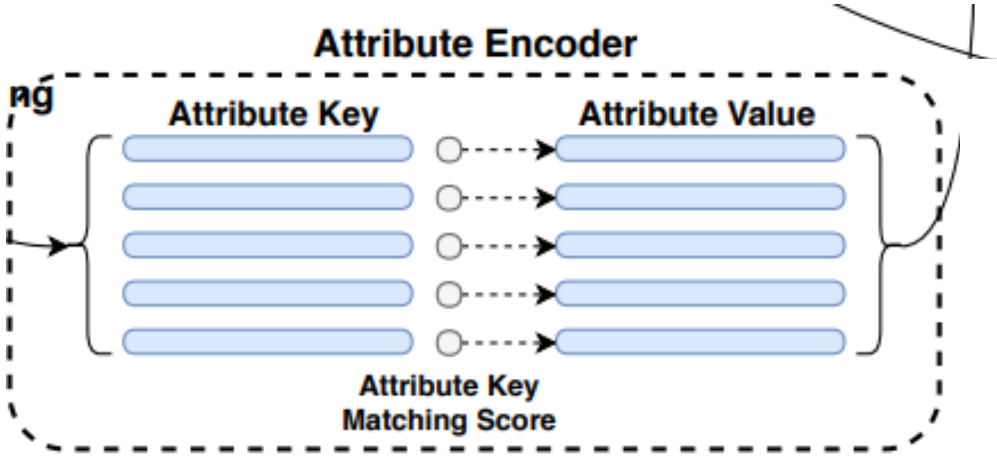
$$u_i' = \exp(u_i) / \sum_{t=1}^{T_r} \exp(u_t),$$

$$c^r = \sum_{i=1}^{T_r} u_i' c_i^r.$$

***Gated Fusion**



Method-
Attributes encoder



價格 —— 50元

口感 —— 柴

外觀 —— 都是雞胸

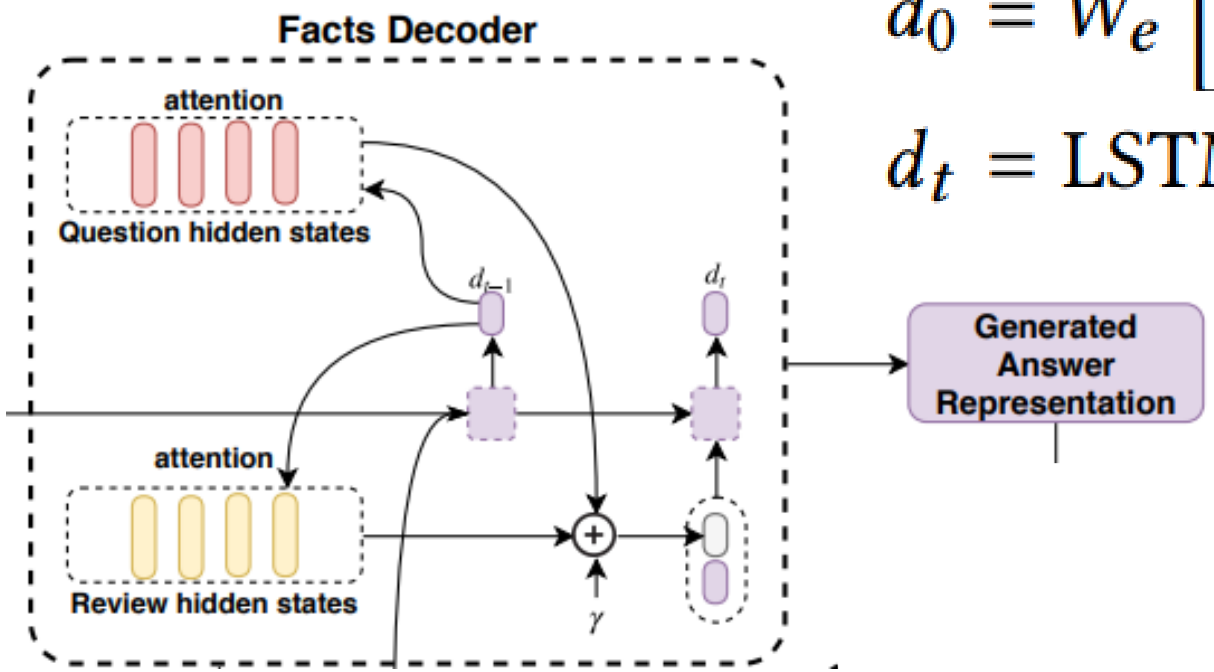
Key matching

$$P(a_i|X^q) = \frac{\exp(h^q_{T_q} W_a e(a^k_i))}{\sum_{t=1}^{T_a} \exp(h^q_{T_q} W_a e(a^k_t))},$$

Value combination

$$m = \sum_{i=1}^{T_a} P(a_i|X^q) e(a^v_i).$$

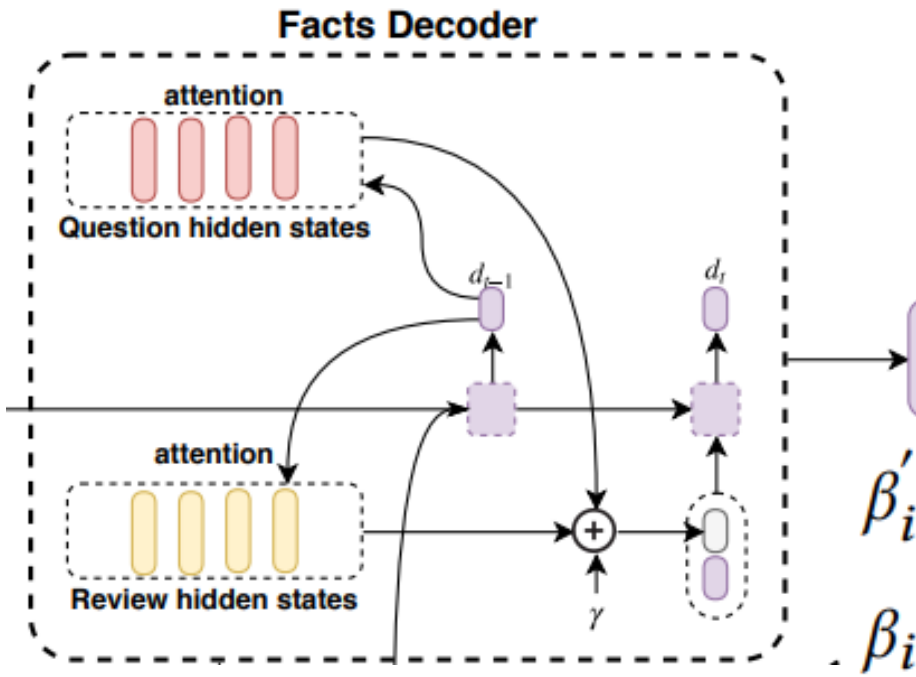
Method-
Facts decoder



$$d_0 = W_e \left[m; h^q_{T_q}; c^r \right] + b_e,$$

$$d_t = \mathrm{LSTM}\left(d_{t-1}, [g_{t-1}; e(y_{t-1})]\right)$$

Method-
Facts decoder



$$\gamma = \sigma\left(W_g d_t + b_g\right)$$

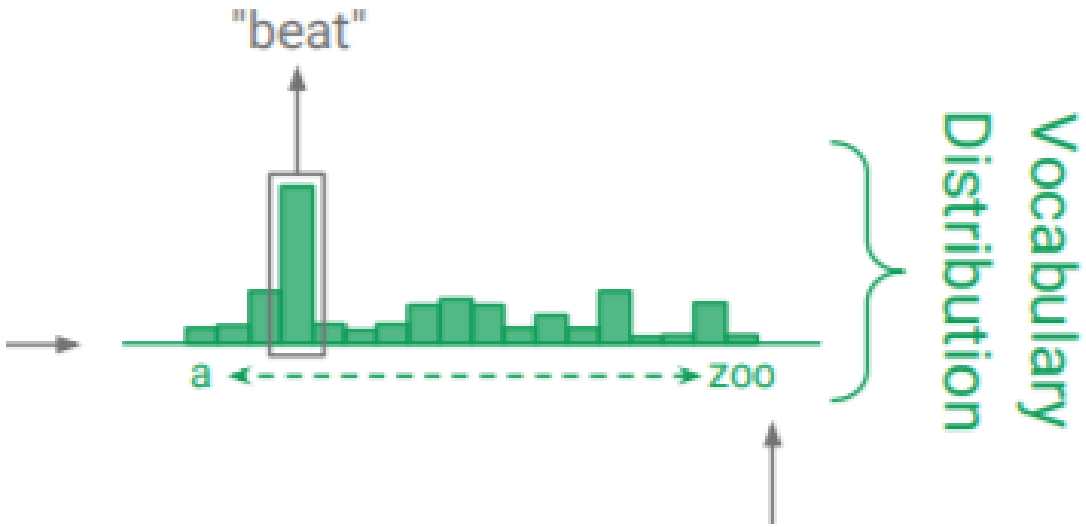
$$g_t = \left[\gamma g_t^r ; (1-\gamma)\, g_t^q\right]$$

$$\beta'_{i,t} = z^{\mathsf{T}} \tanh\left(W_s h_i^* + W_d d_t\right),$$

$$\beta_{i,t} = \exp\left(\beta'_{i,t}\right) / \sum_{j=1}^{T^q} \exp\left(\beta'_{j,t}\right)$$

$$g_t^* = \sum_{i=1}^{T} \beta_{i,t} h_i^*,$$

Method-
Facts decoder

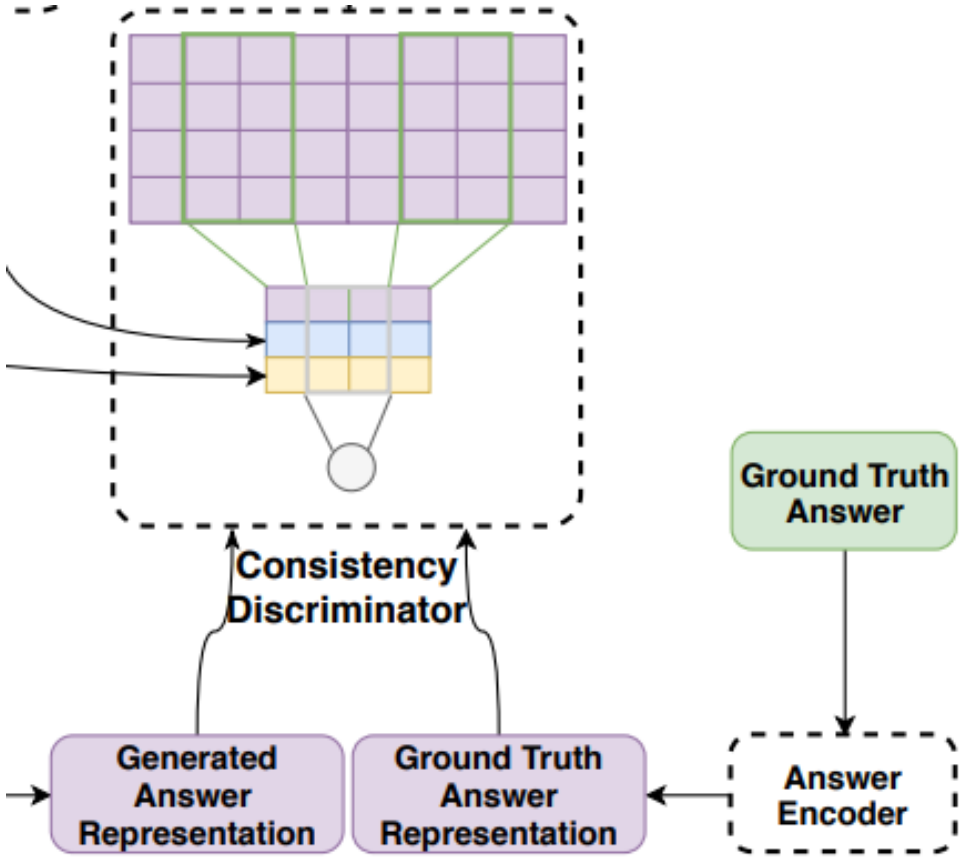


$$d_t^o = (W_o[d_t; g_t] + b_o),$$

$$P_v = \text{softmax}\left(W_v d_t^o + b_v\right)$$

$$loss_g = -1/T_y \sum_{t=1}^{T_y} \log P_v(y_t).$$

Method-
Consistency discriminator



$$\tilde{d}_t^g = LSTM(y_t, \tilde{d}_t^g)$$

$$d_t^g = W_z * \tilde{d}_t^g + b_z$$

$$n_t^* = \text{relu}\left(d_t^* \otimes W_c + b_c\right)$$

[x,x,x,x,x,x,x,x]
[y,y,y,y,y,y,y,y]
[z,z,z,z,z,z,z,z]

n_t^*

Method-
Consistency discriminator

$$D(d_t^*) = W_h \text{relu}\left(N^* + m + c^r\right) + b_h$$

$$d_t' = \epsilon d_t^o + (1-\epsilon) d_t^g$$

$$loss_d = \frac{1}{T_y}\sum_t^{T_y} D(d_t^f) + D(d_t^o) - D(d_t^g) + \lambda\left(\left\|\nabla_{d_t'} D(d_t')\right\|_2 - 1\right)^2$$

加到前面的loss function

Outline

- Introduction
- Method: PAAG
- Experiment
- Conclusion

Experiment

reviews	衣服质量很好，就是我比较瘦，这个S 号对于我来说还是比较肥。很适合孕妇穿。(The quality of the clothes is very good. Because I am thin, the S size is still quite fat for me. It is suitable for pregnant women to wear.)
	颜色漂亮，宽松舒服，就是线头有点多，竟然还有口袋，方便实用！(The clothes are beautiful in color, comfortable to wear, and the thread is a bit more. This dress has a pocket and it is convenient and practical!)
	不轻的雪纺连衣裙，很有垂坠感，觉得夏天春天都适合穿，很衬肤色白呢，胸前会有透，不过不影响穿着(This chiffon dress feels heavy when worn. I think summer and spring are suitable for wearing. Putting on this dress will bring out my skin white. But this dress will be transparent on the chest, but it will not affect the wear.)
attributes	裙长: 中裙||裙型: A 字裙||材质: 涤纶||袖长: 七分袖||领型: 其它||袖型: 其它||上市时间: 2018 春季||版型: A 型
question	怀孕五个多月能穿吗(I have been pregnant for more than five months, can I wear it?)
reference	能穿到生都没问题的(You can wear it until your child is born)
S2SA	可以的，我儿子三个月就穿了(I can wear it, my son wore it when he was three months old.)
RAGF	可以啊，我的就是五个月(I can wear it, my pregnancy is five months.)
PAAG	可以的，我就是(I can wear it, I am pregnant.)

Experiment

	BLEU	BLEU1	BLEU2	BLEU3	BLEU4
Text generation methods					
S2SA	1.6186	15.4754	3.1437	0.8267	0.1706
S2SAR	1.7549	15.1708	3.2156	0.9078	0.2142
SNet	0.9550	13.7029	2.5374	0.4007	0.0597
QS	1.6848	15.4961	2.9508	0.8315	0.2119
PAAG	**2.0189**▲	**16.2232**▲	**3.5711**▲	**1.0290**▲	**0.2787**▲
Sentence extraction methods					
BM25	0.4125	6.9630	0.7097	0.1333	0.0439
TF-IDF	0.2548	5.5480	0.5127	0.0779	0.0190

Experiment

	Fluency		Consistency	
	mean	variance	mean	variance
Text generation methods				
S2SA	2.22	0.3	1.62	0.29
S2SAR	2.405	0.365	1.82	0.39
SNet	1.93	0.36	1.355	0.225
QS	2.335	0.285	1.725	0.355
PAAG	2.865▲	0.105	2.205▲	0.445
Sentence extraction methods				
BM25	2.70	0.24	1.45	0.29
TF-IDF	2.48	0.38	1.14	0.12

Experiment

	BLEU	BLEU1	BLEU2	BLEU3	BLEU4
RAGF	1.7931	15.7213	3.3705	0.9385	0.2079
RAGFD	1.8597	15.9021	3.4160	0.9409	0.2340
RAGFWD	1.9389	16.1755	**3.5986**	0.9865	0.2461
PAAG	**2.0189**	**16.2232**	3.5711	**1.0290**	**0.2787**

Acronym	Gloss
RAGF	**R**eview reader + **A**ttributes encoder + **G**ated **F**usion
RAGFD	RAGF + consistency **D**iscriminator
RAGFWD	RAGF + **W**asserstein consistency **D**iscriminator
PAAG	RAGFWD + Gradient Penalty

Experiment

	PAAG		S2SA		S2SAR	
	BLEU1	BLEU2	BLEU1	BLEU2	BLEU1	BLEU2
Jewelry	**19.53**	**6.35**	17.65	4.26	18.74	4.69
Mattress	**18.89**	4.14	16.35	3.00	17.52	**5.57**
Clothing	18.18	**5.17**	**18.39**	4.98	18.36	4.68
Kitchenware	**18.00**	**4.31**	15.23	3.19	17.15	4.09
Power and Handtools	**16.34**	**3.98**	13.73	3.20	15.60	3.22
Skin Care	18.01	**4.57**	15.39	3.55	**18.33**	4.40
Gardening	13.67	**2.30**	11.86	1.52	**15.74**	**2.30**
Baby	**18.22**	**4.51**	16.95	3.71	17.27	3.75
Automotive Accessories	17.46	**3.43**	15.49	3.14	**17.86**	3.00
Gift	**19.25**	3.93	17.23	3.06	18.39	**4.24**

Outline

- Introduction
- Method: PAAG
- Experiment
- Conclusion

Conclusion

- PAAG is better than state-of-the-art baselines in terms of metric-based evaluations and human evaluations.
- we have verified the effectiveness of different product categories using PAAG.