

# **Towards Building More Intelligent Chatting Machines**

**Dr. Minlie Huang (黄民烈)**

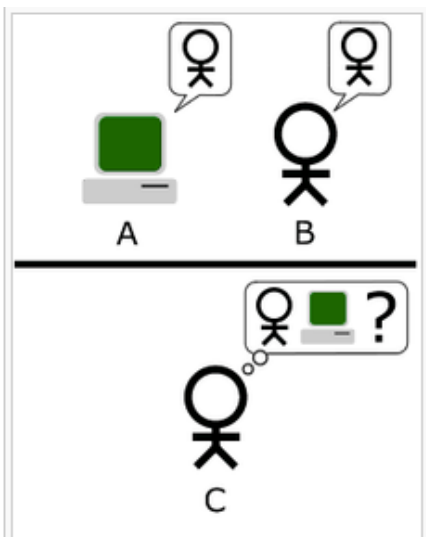
**Associate Professor**

**CS Department, Tsinghua University**

**[aihuang@tsinghua.edu.cn](mailto:aihuang@tsinghua.edu.cn)**

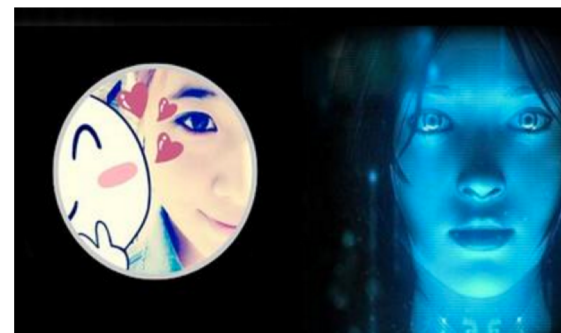
**<http://coai.cs.tsinghua.edu.cn/hml>**

# Open-domain Chatting and Conversational AI



1950 Turing Test

2010 IBM Watson: Human-level knowledge QA



2014 Xiaoice: virtual partner

**Building open-domain chatting machines is one of the most challenging AI tasks.**



# Virtual Conversational Agents

## Eliza Chat bot

Is she a Rogerian psychotherapist, a semi intelligent chat bot, or just a toy?



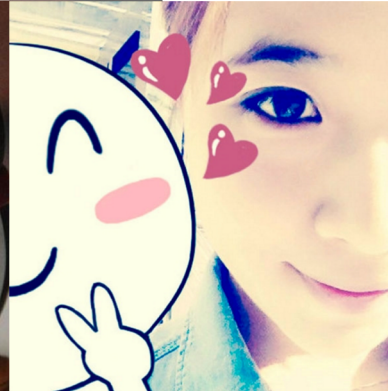
#1 limitations of early artificial intelligence programs

Eliza, one of the original chat bots, is now ready to talk to you.

She is programmed to behave as a Rogerian psychotherapist, and is an interesting example of the



**A.L.I.C.E.:** I have never had a teacher.  
**Fake Kirk:** So I am teaching you now.



1966

Eliza

1994

ALICE

2010

QQ chatbot

2014

Xiaoice

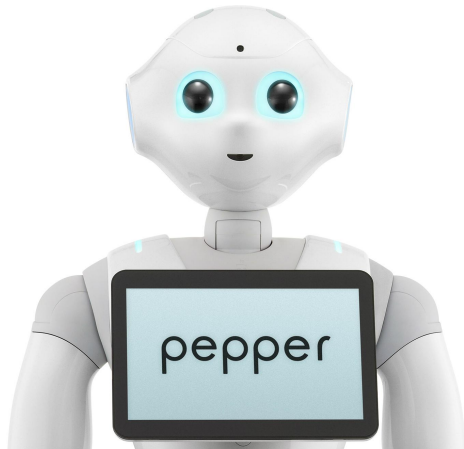
2016

MS Tay



# Social Robots (with Physical Body)

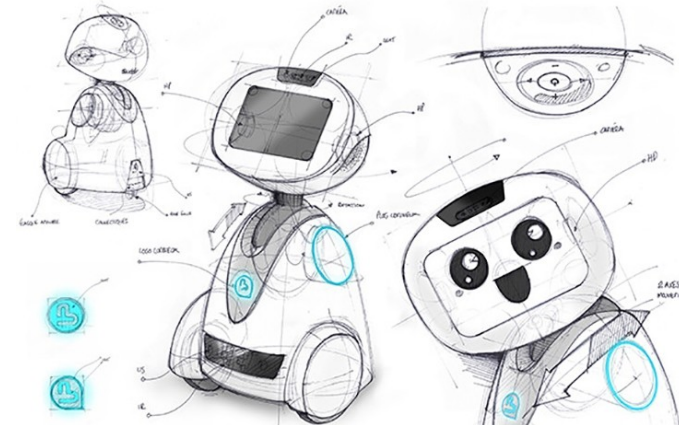
**Pepper**—humanoid robot  
The first robot with  
emotions



**MIT Jibo:** the  
world's first social  
robot for the home



**Buddy:** protects your home,  
entertains the family,  
interfaces with devices, and  
assists the family



# Good Examples from DL Models

User: 最幸福的人不是拿到了最好的东西，而是学会去珍惜手上已经拥有的人。

**The happiest person is not the one who gets the best thing, instead who cherishes what she/he already has.**

Machine: 是的，所以我也要珍惜自己的幸福

**Agree, I will cherish my happiness too.**

User: 时间啊，你这把杀猪刀。。。

**Time is a killer...**

Machine: 我是被杀的

**I am the one who was killed.**

User: 怕孤独，又捨不下自己的小自由。

**Afraid to be lonely, but wont' give up my little freedom.**

Machine: 我想去流浪，但没有人陪我去。

**I want to wander, but no one accompany me.**



# Bad Examples are More!

User: 其实，他没有那么爱你。

**Ohh, he does not love you that much.**

Machine: 幸福的一天开始了。

**A happy day starts.**

User: 回北京了！很冷很高兴！

**Back to Beijing now. So cold so happy!**

Machine: 什么时候来北京啊？

**When will you come to Beijing?**

User: 会做饭的男人是最性感的男人。

**Mans who can cook are the most sexy.**

Machine: 会做饭的男人是最性感的女人

**Mans who can cook are the most sexy women.**



# Bad Examples are More!

User: 猫叔羊毛毡手工作品，太美好了！

Machine: 这是什么品种的猫啊

User: 不得不夸一句广东的饭确实好吃！

Machine: 我也喜欢吃臭豆腐。

User: 我来自上海。

I come from Shanghai.

Machine: 真巧，我来自海南，老乡啊

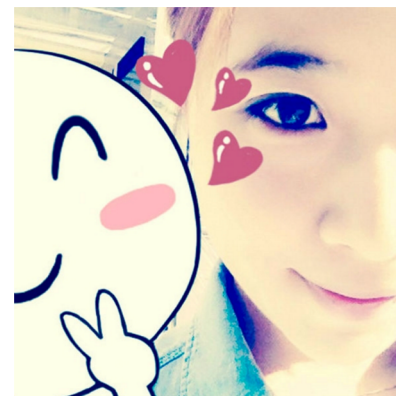
So cool, I am from Hainan, the same province!



# Bad Examples are More!

Xiaoice by MSRA

微软  
小冰



## 语义理解的问题 Semantic understanding

**User:** 你家多大? How large is your house?

**小冰:** 这么快就要爆年龄了?

**So soon to reveal my age?**

**User:** 你是素食主义者还是吃肉啊?

Are you vegan or meatatarian?

**小冰:** 怎么减肥

**How can I lose weight?**

## 上下文理解的问题 Context issue

**User:** 早饭不好吃

**小冰:** 那明早一起吃

**User:** 行, 你推荐的应该不错

**小冰:** 是吗?

**User:** 哈哈, 我相信你

**小冰:** 呃呃

## 个性身份一致性问题

Inconsistency in personality

**User:** 你几岁了?

**小冰:** 我啊。。。反正是九零后!

**User:** 那你是几几年出生的?

**小冰:** 错, 75年阳历一月出生。

**User:** 你是九零后, 怎么会75年出生?

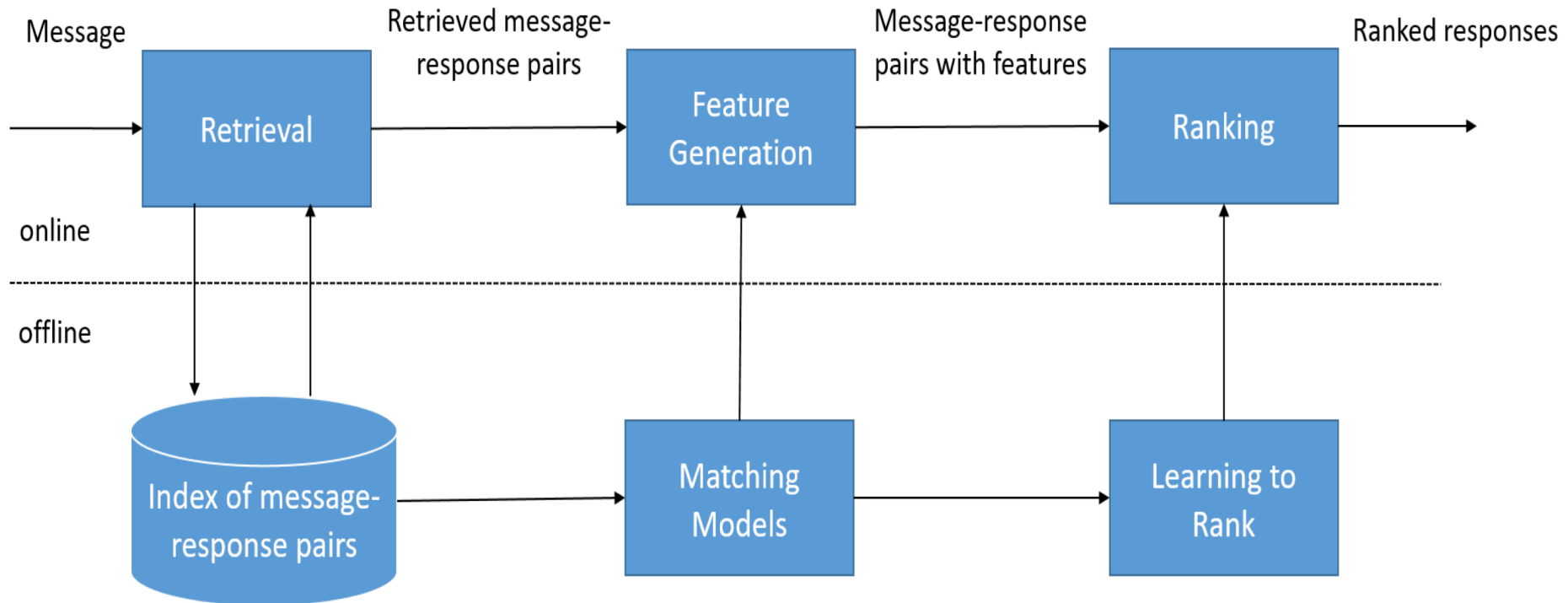
**小冰:** 生在九零后, 在深圳只能被当做八零后了。



# Bad Examples (AI Ethics)



# Typical Solution I: Retrieval-based



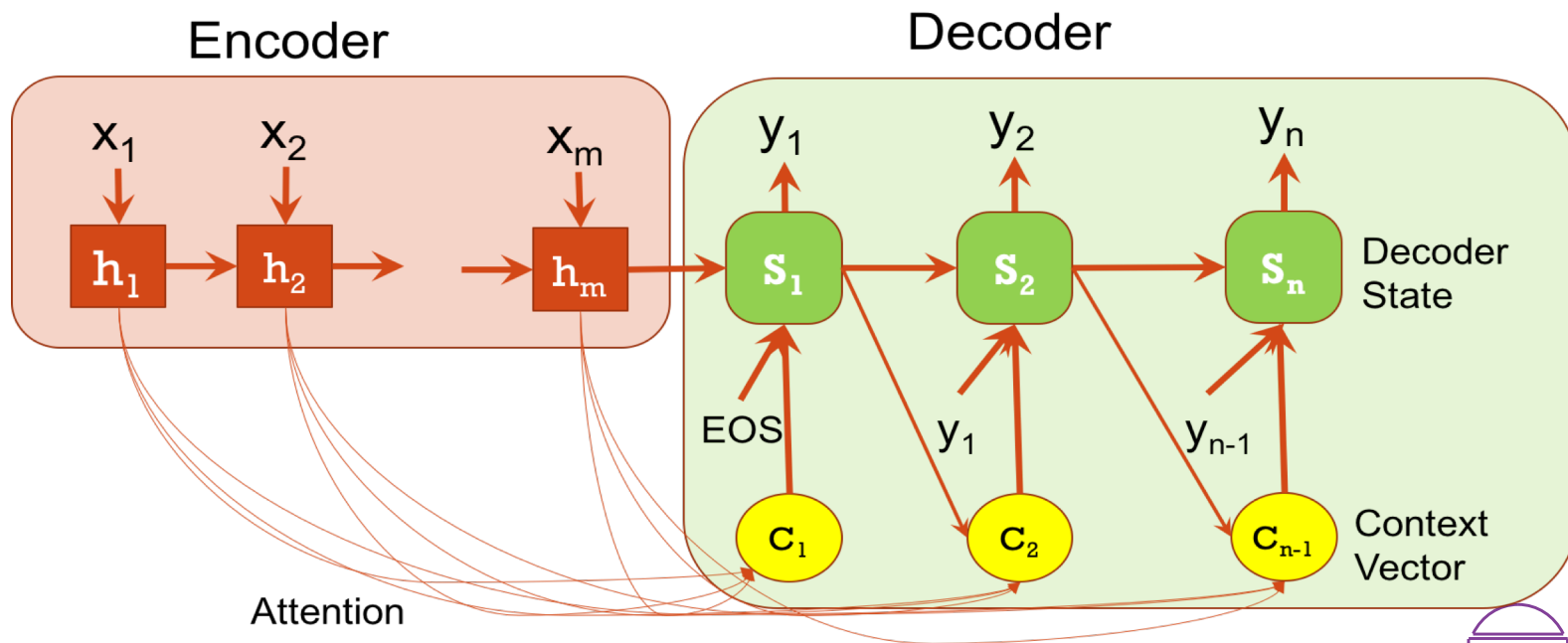
From MSRA Dr. Ming Zhou



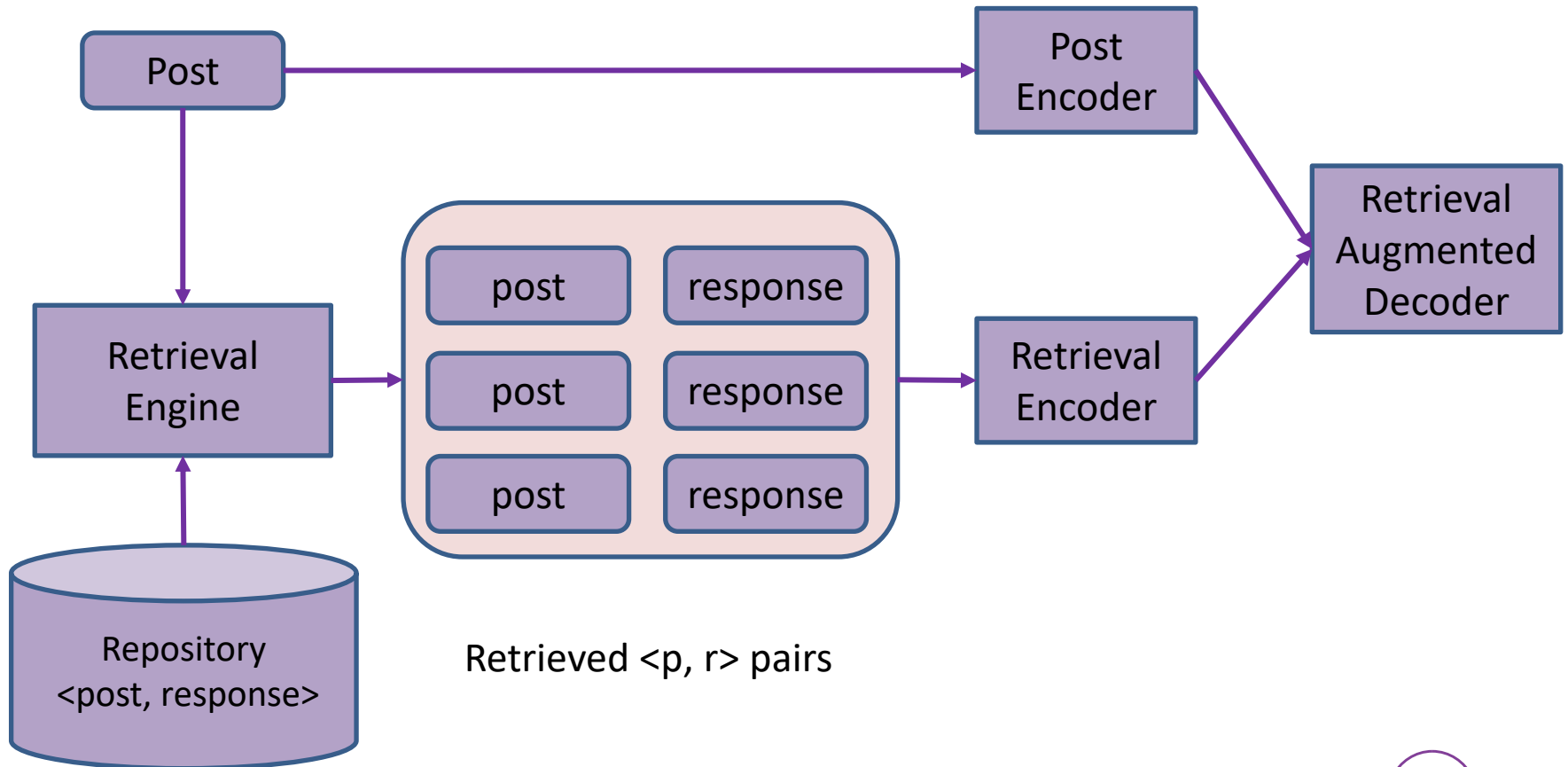
# Typical Solution II: Generation-based

User: I am so happy to be here

Machine: Glad with you.



# Typical Solution III: Hybrid Methods

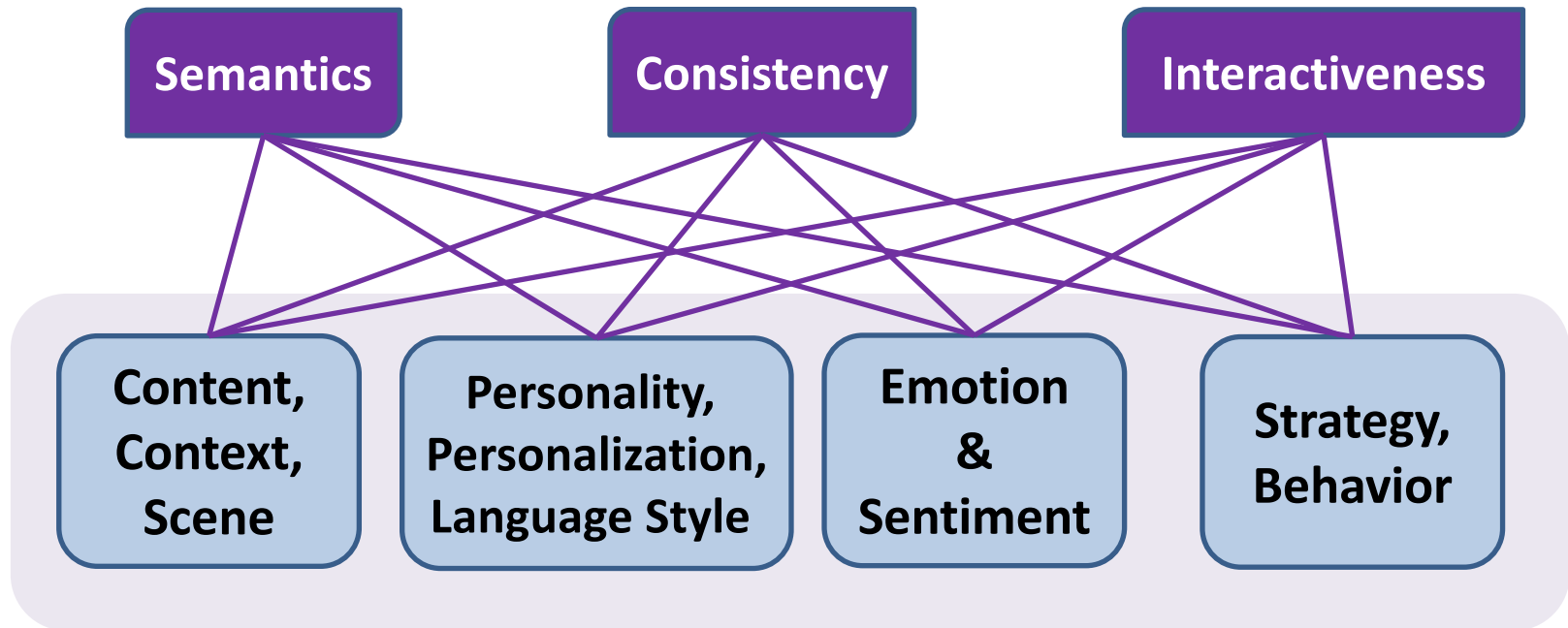


# Challenges in Chatting Machines

- ◎ **One-to-many:** one input, many many possible responses
- ◎ **Knowledge & Reasoning:** real understanding requires various knowledge, world facts, or backgrounds
- ◎ **Situational Context**
  - ◆ Who are you talking with?
    - Stranger, or friend?
  - ◆ His mood and emotion?
  - ◆ Shared backgrounds that are only accessible by two acquaintances



# Challenges in Chatting Machines



Open-domain, open-topic conversational systems



# Open-domain Chatting Machines

- ◎ Behaving more interactively:
  - ◆ Perceiving and Expressing Emotions (**AAAI 2018**)
  - ◆ Proactive Behavior by Asking Good Questions (**ACL 2018**)
  - ◆ Controlling sentence function (**ACL 2018**)
- ◎ Behaving more consistently:
  - ◆ Explicit Personality Assignment (**IJCAI-ECAI 2018**)
- ◎ Behaving more intelligently with semantics:
  - ◆ Better Understanding and Generation Using Commonsense Knowledge (**IJCAI-ECAI 2018**)

## References:

- ① Emotional Chatting Machine: Emotional Conversation Generation with Internal and External Memory. **AAAI 2018**.
- ② Assigning personality/identity to a chatting machine for coherent conversation generation. **IJCAI-ECAI 2018**.
- ③ Commonsense Knowledge Aware Conversation Generation with Graph Attention. **IJCAI-ECAI 2018**.
- ④ Learning to Ask Questions in Open-domain Conversational Systems with Typed Decoders. **ACL 2018**.
- ⑤ Generating Informative Responses with Controlled Sentence Function. **ACL 2018**.

# Interactiveness: Emotion Perception and Expression



# Emotional Chatting Machine

- Emotion intelligence is a key human behavior for intelligence (Salovey and Mayer, 1990; Picard and Picard, 1997)
- Understanding **emotion and affect** is important for dialogue and conversation
  - Enhance user performance
  - Improve user satisfaction
  - Less breakdowns
- Rule-based** emotion adaptation is widely seen in early dialogue systems

User: Worst day ever. I arrived late because of the traffic.

Chatbot (Basic Seq2seq): You were late.

ECM (*Like*): I am always here to support you.

ECM (*Happiness*): Keep smiling! Things will get better.

ECM (*Sadness*): It's depressing.

ECM (*Disgust*): Sometimes life just sucks.

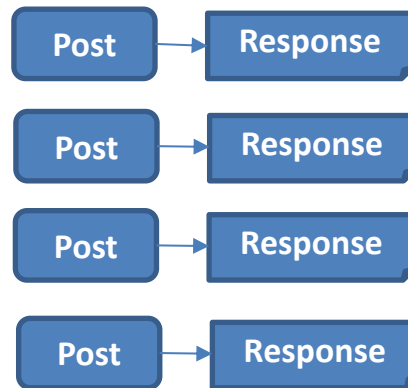
ECM (*Anger*): The traffic is too bad!



# Emotional Chatting Machine

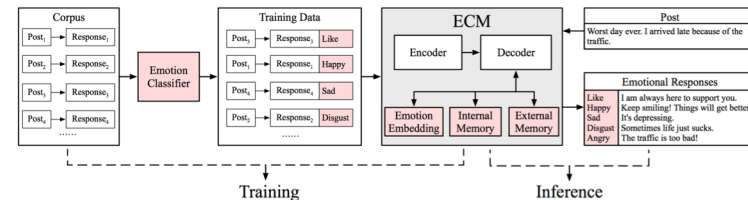
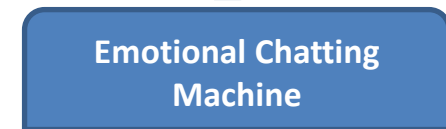
Perceiving and Expressing emotion by machine  
Closer to human-level intelligence

Social Interaction Data



Emotion  
Tagged  
data

今天领导前来视察工作！  
Happy: 哈哈，领导视察工作辛苦了  
Like: 这是一个很有特色的领导人  
Sad: 我想知道，谁来帮我干活？  
Angry: 这是要干什么？领导？



Our work was reported by **MIT Technology Review, the Guardian, Cankao News, Xinhua News Agency etc.**

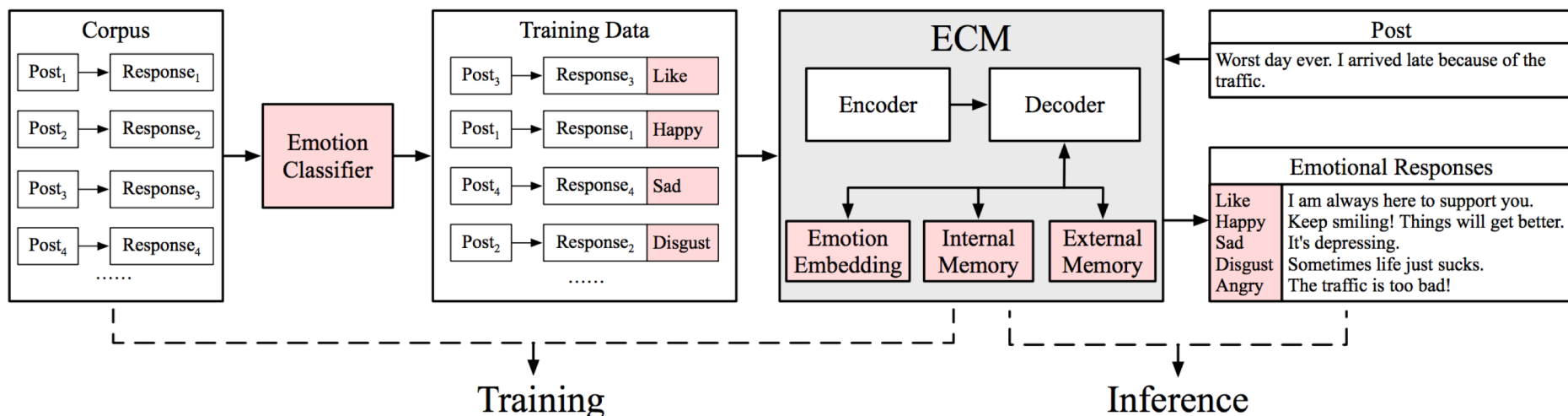
Prof Björn Schuller: “**an important step**” towards personal assistants that could read the emotional undercurrent of a conversation and respond with something akin to empathy.

•Hao Zhou, Minlie Huang, Xiaoyan Zhu, Bing Liu. Emotional Chatting Machine: Emotional Conversation Generation with Internal and External Memory. **AAAI 2018**.



# Emotional Chatting Machine

- ◉ **Emotion category embedding:** High level abstraction of emotions
- ◉ **Emotion internal state:** Capturing the change of emotion state during decoding
- ◉ **Emotion external memory:** Treating emotion/generic words differentially

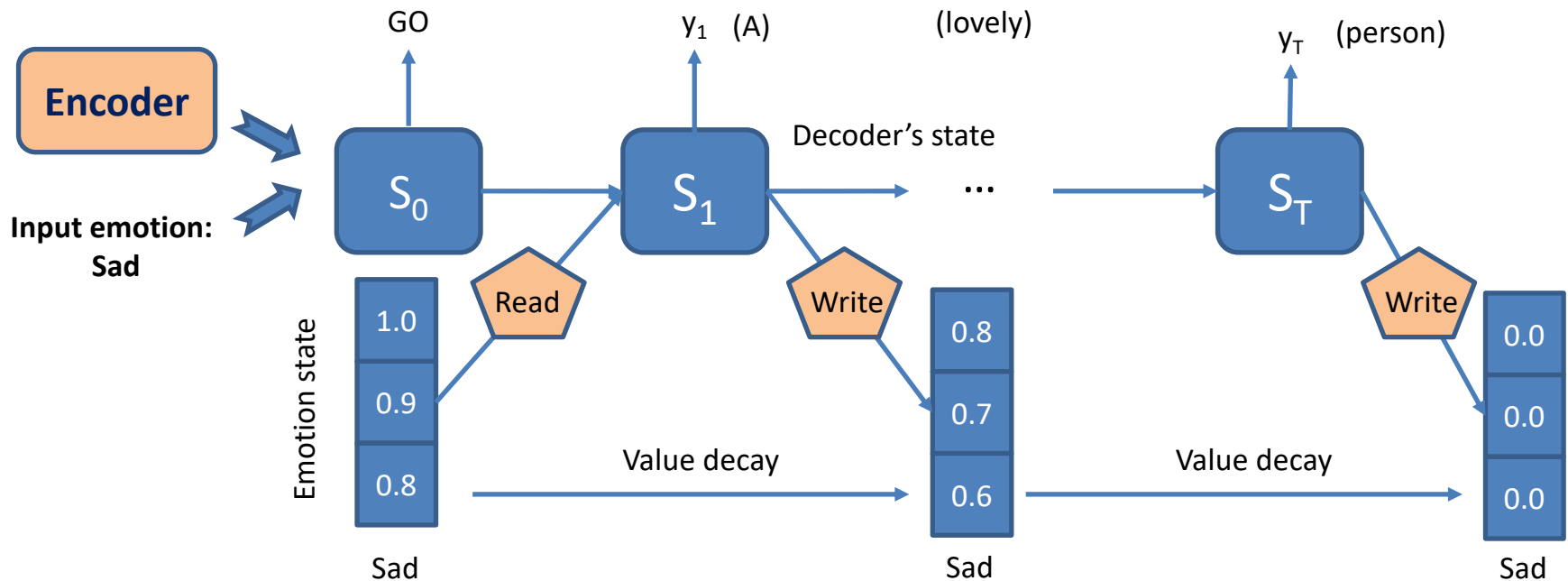


• Hao Zhou, Minlie Huang, Xiaoyan Zhu, Bing Liu. Emotional Chatting Machine: Emotional Conversation Generation with Internal and External Memory. **AAAI 2018**.



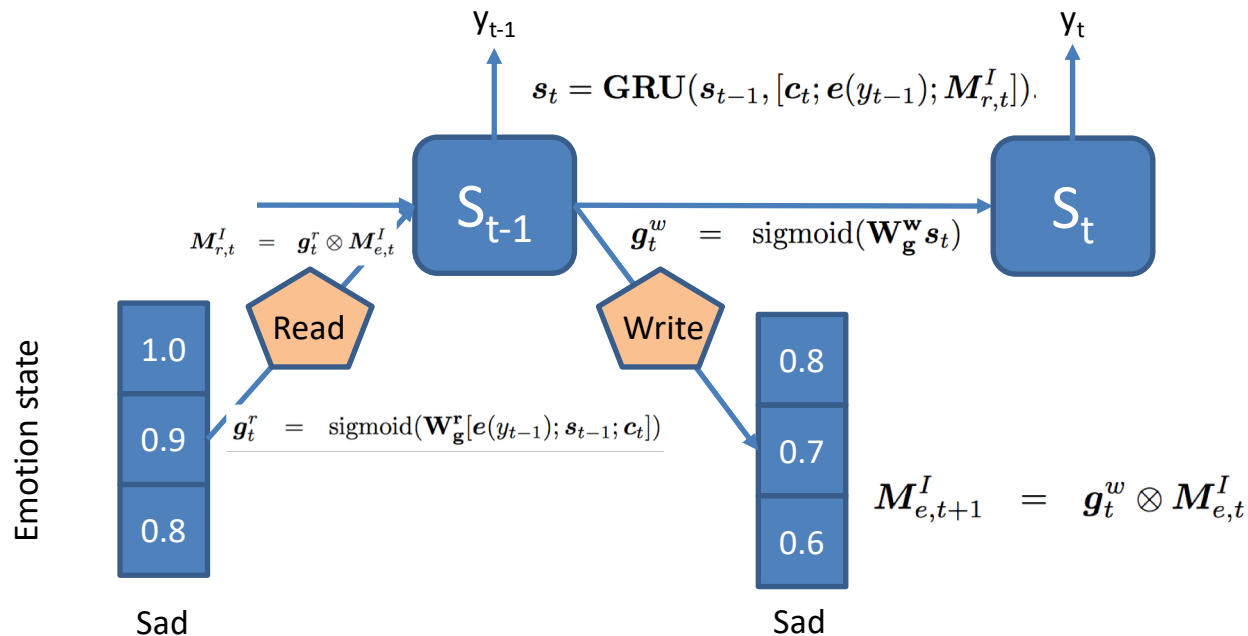
# Emotional Chatting Machine

- Internal emotion memory : “emotional responses are relatively short lived and involve changes” (Gross, 1998; Hochschild, 1979)



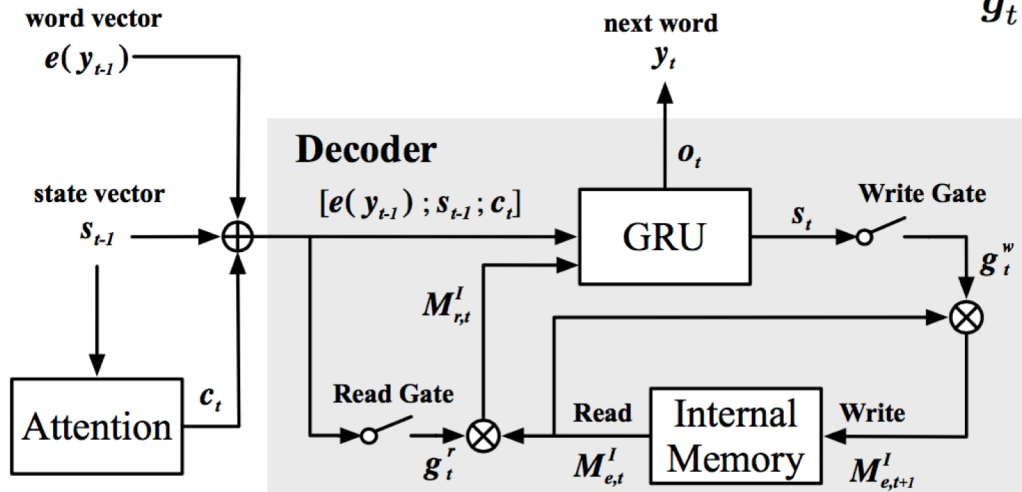
# Emotional Chatting Machine

- Internal emotion memory : “emotional responses are relatively short lived and involve changes” (Gross, 1998; Hochschild, 1979)



# Emotional Chatting Machine

- Internal emotion memory : “emotional responses are relatively short lived and involve changes” (Gross, 1998; Hochschild, 1979)



$$g_t^r = \text{sigmoid}(\mathbf{W}_g^r[e(y_{t-1}); s_{t-1}; c_t]),$$

$$g_t^w = \text{sigmoid}(\mathbf{W}_g^w s_t).$$

$$M_{r,t}^I = g_t^r \otimes M_{e,t}^I,$$

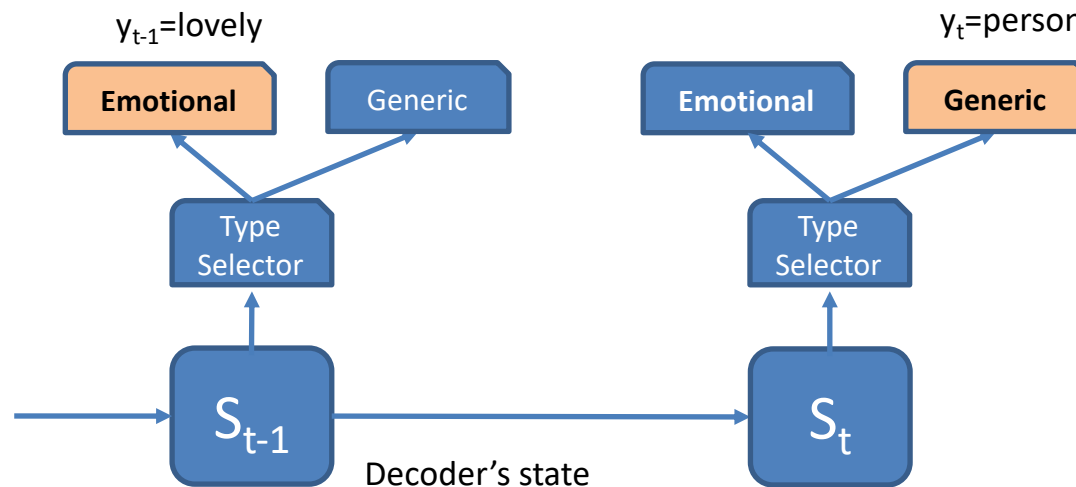
$$M_{e,t+1}^I = g_t^w \otimes M_{e,t}^I,$$

$$s_t = \text{GRU}(s_{t-1}, [c_t; e(y_{t-1}); M_{r,t}^I]).$$



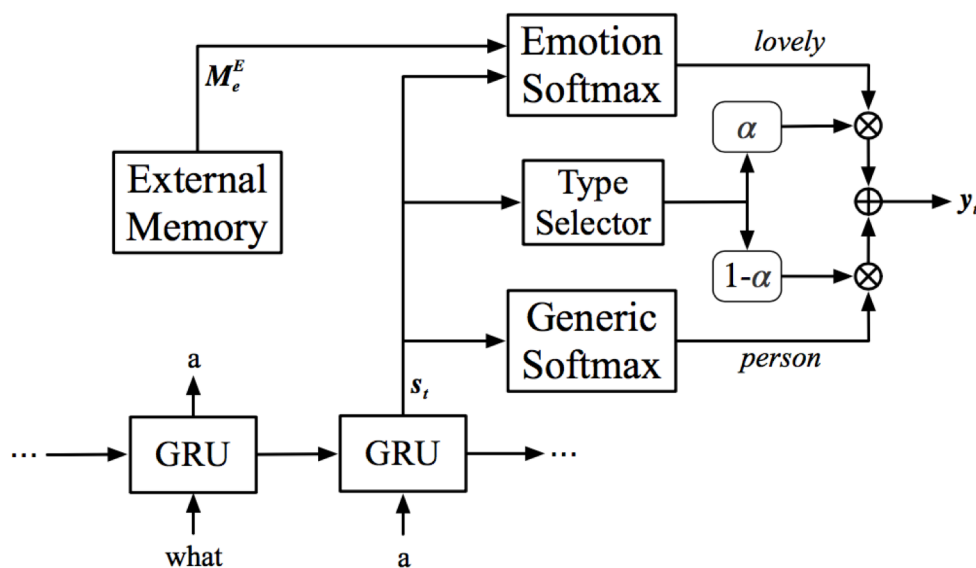
# Emotional Chatting Machine

- External emotion memory: generic words (**person**) and emotion words (**lovely**)



# Emotional Chatting Machine

- External emotion memory: generic words (**person**) and emotion words (**lovely**)



$$\begin{aligned}\alpha_t &= \text{sigmoid}(\mathbf{v}_u^\top \mathbf{s}_t), \\ P_g(y_t = w_g) &= \text{softmax}(\mathbf{W}_g^o \mathbf{s}_t), \\ P_e(y_t = w_e) &= \text{softmax}(\mathbf{W}_e^o \mathbf{s}_t), \\ y_t \sim \mathbf{o}_t = P(y_t) &= \begin{bmatrix} (1 - \alpha_t) P_g(y_t = w_g) \\ \alpha_t P_e(y_t = w_e) \end{bmatrix}\end{aligned}$$



# Emotional Chatting Machine

- ◎ Emotion Classification Dataset: the Emotion Classification Dataset of NLPCC 2013&2014
  - ◆ 23,105 sentences collected from Weibo
- ◎ The STC dataset: a conversation dataset from (Shang et al., 2015)
  - ◆ 219,905 posts and 4,308,211 responses
  - ◆ Each post has about 20 responses



# Emotional Chatting Machine

## ◆ Automatic Evaluation

Method	Perplexity	Accuracy
Seq2Seq	68.0	0.179
Emb	62.5	0.724
ECM	65.9	<b>0.773</b>
w/o Emb	66.1	0.753
w/o IMem	66.7	0.749
w/o EMem	<b>61.8</b>	0.731

Table 4: Objective evaluation with perplexity and accuracy.



# Emotional Chatting Machine

Method (%)	2-1	1-1	0-1	2-0	1-0	0-0
Seq2Seq	9.0	5.1	1.1	37.6	28.0	19.2
Emb	22.8	9.3	4.3	27.1	19.1	17.4
ECM	<b>27.2</b>	<b>10.8</b>	4.4	24.2	15.5	17.9

Score distribution

Table 5: The percentage of responses with the score of *Content-Emotion*. For content score is 2 and emotion score is 1.

Preference test

Pref. (%)	Seq2Seq	Emb	ECM
Seq2Seq	-	38.8	38.6
Emb	60.2	-	43.1
ECM	<b>61.4</b>	<b>56.9</b>	-

Table 7: Pairwise preference of the three systems.

Method	Overall		Like		Sad		Disgust		Angry		Happy	
	Cont.	Emot.	Cont.	Emot.	Cont.	Emot.	Cont.	Emot.	Cont.	Emot.	Cont.	Emot.
Seq2Seq	1.255	0.152	1.308	0.337	1.270	0.077	<b>1.285</b>	0.038	<b>1.223</b>	0.052	1.223	0.257
Emb	1.256	0.363	1.348	0.663	1.337	0.228	1.272	0.157	1.035	0.162	1.418	0.607
ECM	<b>1.299</b>	<b>0.424</b>	<b>1.460</b>	<b>0.697</b>	<b>1.352</b>	<b>0.313</b>	1.233	<b>0.193</b>	0.98	<b>0.217</b>	<b>1.428</b>	<b>0.700</b>

Table 6: Manual evaluation of the generated responses in terms of *Content* (Cont.) and *Emotion* (Emot.).

# Emotional Chatting Machine

Post	Model	Response(Original)	Response(Translated)
承诺，就是一个骗子说给一个傻子听的。 A promise is what a cheater says to a fool.	Seq2Seq	承诺，就是一个骗子说给一个傻子听的。	A promise is what a cheater says to a fool.
	ECM	Like 承诺是一个 <b>专一</b> 的傻子	Promise is a <b>single-minded</b> fool
		Happy 我就是 <b>一个开心的</b> 傻子	I'm just a <b>happy</b> fool
		Sad 我就是 <b>一个悲伤的</b> 傻子	I am a <b>sad</b> fool
		Disgust 一个 <b>骗子</b> 说给一个 <b>傻子</b> 听	A <b>cheater</b> says to a <b>fool</b>
		Angry 不管你信不信，反正我信了！	Whatever you believe or not, I trust in it!
知道真相的我眼泪笑下来 Knowing the truth, I laughed myself to tears	Seq2Seq	知道真相的我眼泪掉下来。	After knowing the truth, I burst into tears.
	ECM	Like 太 <b>可爱</b> 了，太 <b>可爱</b> 了！	Too <b>lovely</b> , it's too <b>lovely</b> .
		Happy 笑得我眼泪都出来了。	I can't help laugh myself to tears.
		Sad 知道真相的我 <b>眼泪</b> 掉下来。	After knowing the truth, I burst into <b>tears</b> .
		Disgust 这是真的还是假的？	Is this true?
		Angry 什么意思？没看懂。	What do you mean? I don't understand.

•Hao Zhou, Minlie Huang, Xiaoyan Zhu, Bing Liu. Emotional Chatting Machine: Emotional Conversation Generation with Internal and External Memory. **AAAI 2018**.



# More Examples

情绪 emotion	Post:早上的天气不错啊！哈哈哈 哈哈哈
Other	今天是个好天气。
喜欢 Like	今天的天气不错啊！
悲伤 Sad	天气好，心情不好。
厌恶 Disgust	这是哪里啊？求解释
愤怒 Angry	怎么了？什么情况？
快乐 Happy	今天的天气真好啊！

情绪 emotion	Post:为了解决抄作业的坏习惯， 高数老师这回要出绝招了！
Other	这老师是要逆天吗？
喜欢 Like	这老师太有才了！
悲伤 Sad	这样的老师伤不起啊
厌恶 Disgust	这老师也太狠了点吧
愤怒 Angry	这老师是要逆天啊！
快乐 Happy	哈哈，这老师太有才了

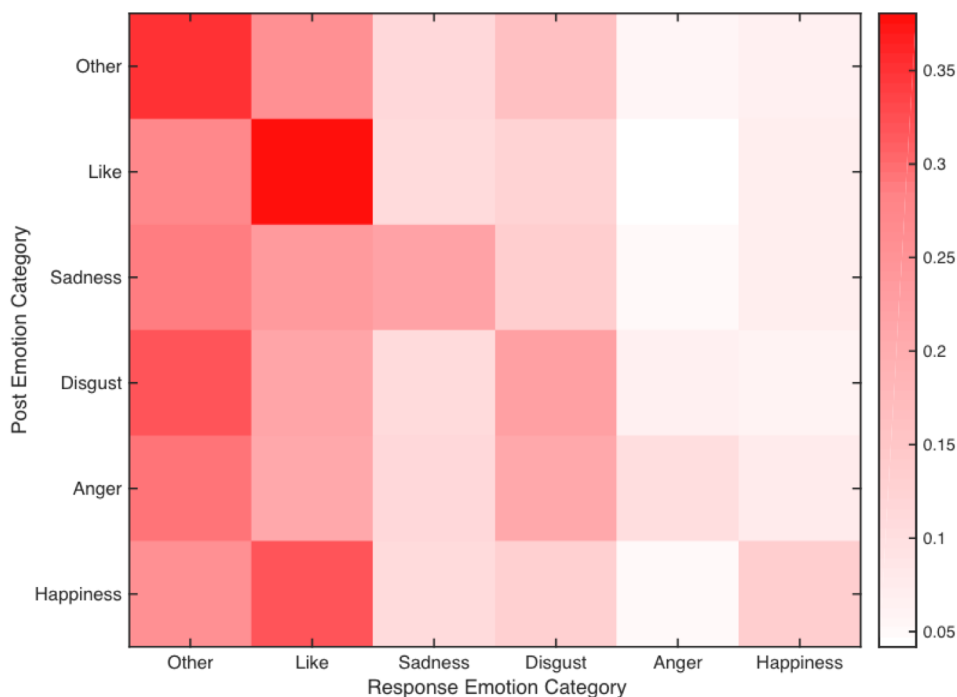
情绪 emotion	Post:送一个这样的蛋糕，给你美丽的闺蜜 吧！
Other	这是蛋糕蛋糕吗？
喜欢 Like	好漂亮的蛋糕啊！
悲伤 Sad	为什么我觉得很恶心
厌恶 Disgust	这样的蛋糕你忍心吗？
愤怒 Angry	有没有想开网店的？
快乐 Happy	哈哈，好可爱的蛋糕！

情绪 emotion	post:承诺，就是一个骗子说给一个傻子听 的。
Other	承诺就是一个承诺
喜欢 Like	承诺是一个专一的傻子
悲伤 Sad	我就是个悲伤的傻子
厌恶 Disgust	一个骗子说给一个傻子听
愤怒 Angry	不管你信不信，反正我信了！
快乐 Happy	我就是个开心的傻子

•Hao Zhou, Minlie Huang, Xiaoyan Zhu, Bing Liu. Emotional Chatting Machine:  
Emotional Conversation Generation with Internal and External Memory. **AAAI 2018.**



# Emotion Interaction Patterns



Like → Like (**empathy**)

Sadness → Sadness (**empathy**)

Sadness → Like (**comfort**)

Disgust → Disgust (**empathy**)

Disgust → Like (**comfort**)

Anger → Disgust

Happiness → Like

• Hao Zhou, Minlie Huang, Xiaoyan Zhu, Bing Liu. Emotional Chatting Machine: Emotional Conversation Generation with Internal and External Memory. **AAAI 2018**.



# Interactiveness:

## Behaving More Proactively by Asking Good Questions



# Asking Questions in Chatbots

我昨天晚上去聚餐了

I went to dinner yesterday night.

Yansen Wang, Chenyi Liu, Minlie Huang, Liqiang Nie.

Learning to ask questions in open-domain conversation systems. **ACL 2018**.



# Asking Questions in Chatbots

- Asking **good** questions requires **scene understanding**

Scene: Dining at a restaurant

我昨天晚上去聚餐了  
I went to dinner yesterday night.

Friends?

WHO

Place?

WHERE

Food?

HOW-ABOUT

...

Persons?

HOW-MANY

Bill?

WHO

Yansen Wang, Chenyi Liu, Minlie Huang, Liqiang Nie.  
Learning to ask questions in open-domain conversation systems. **ACL 2018**.



# Asking Questions in Chatbots

- ◉ Responding + **asking** (Li et al., 2016)
- ◉ **Key proactive** behaviors (Yu et al., 2016)
- ◉ Asking good questions are indication of **machine understanding**
- ◉ Key differences to **traditional** question generation (eg., reading comprehension):
  - ◆ **Different goals**: Information seeking vs. Enhancing interactiveness and persistence of human-machine interactions
  - ◆ **Various patterns**: YES-NO, WH-, HOW-ABOUT, etc.
  - ◆ **Topic transition**: from topics in post to topics in response



# Asking Questions in Chatbots

- ◎ A good question is a natural composition of
  - ◆ **Interrogatives** for using various questioning patterns
  - ◆ **Topic words** for addressing interesting yet novel topics
  - ◆ **Ordinary words** for playing grammar or syntactic roles

Example 1:

User: I am too fat ...

Machine: **How about** climbing **this weekend?**

Example 2:

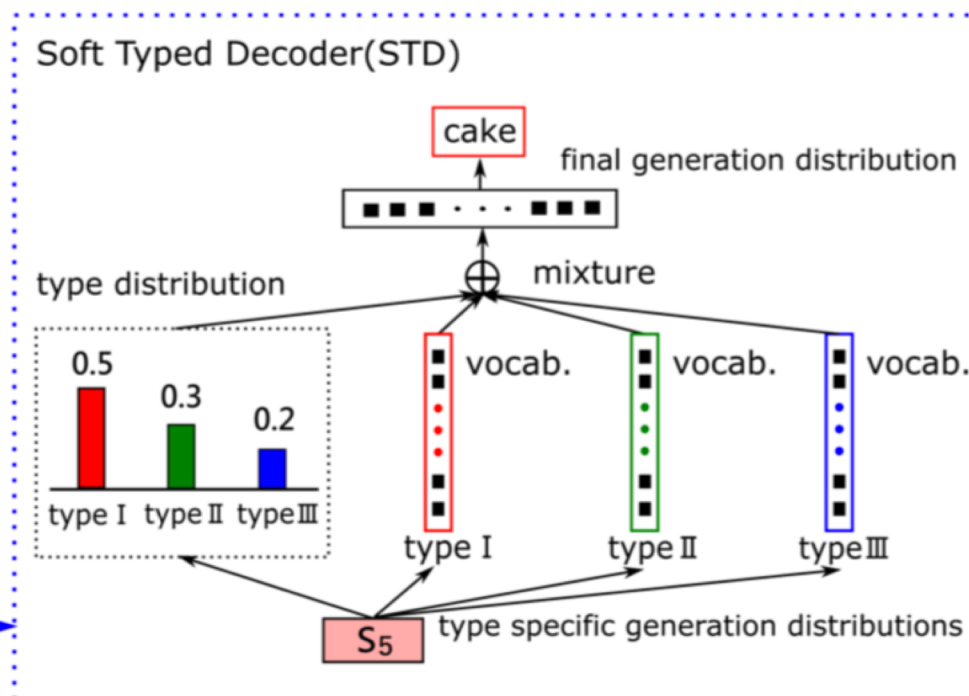
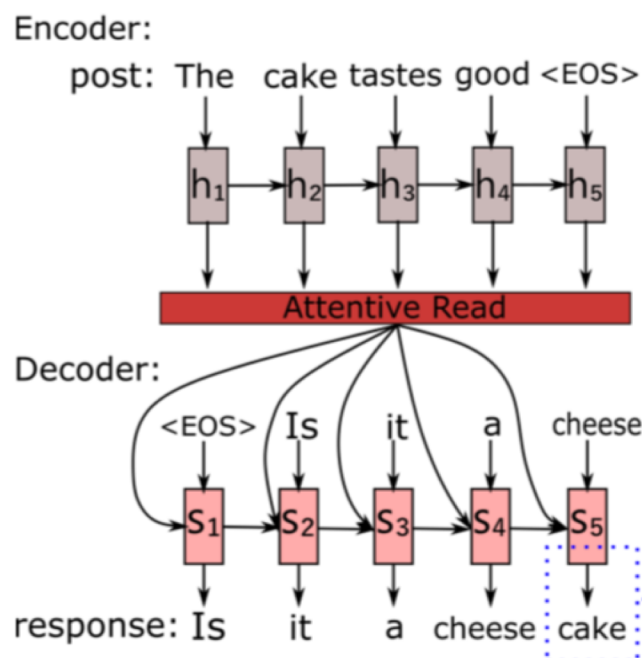
User: Last night, I stayed in KTV with friends.

Machine: **Are you happy with your** singing?



# Asking Questions in Chatbots

## Typed decoders: soft typed decoder



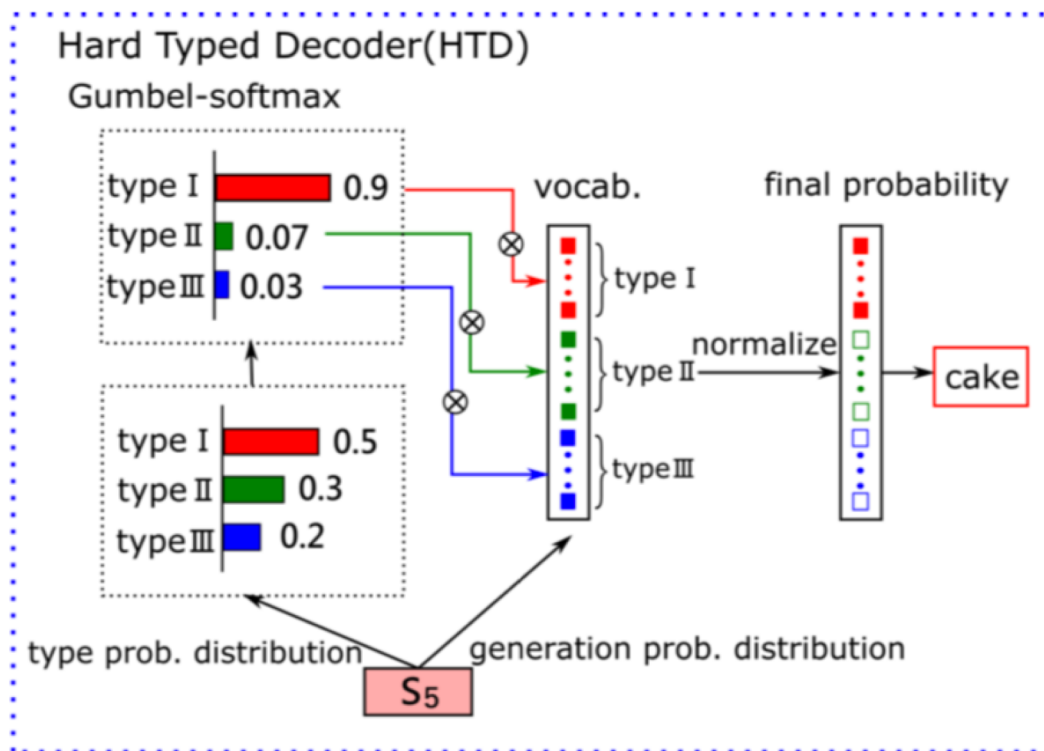
Yansen Wang, Chenyi Liu, Minlie Huang, Liqiang Nie.

Learning to ask questions in open-domain conversation systems. **ACL 2018**.



# Asking Questions in Chatbots

## Typed decoders: hard typed decoder



### For each post:

- A set of interrogatives
- A list of topic words
- Others for ordinary words

### Topic words:

- Training -- nouns, verbs
- Test – predicted by PMI

Yansen Wang, Chenyi Liu, Minlie Huang, Liqiang Nie.

Learning to ask questions in open-domain conversation systems. **ACL 2018**.



# Asking Questions in Chatbots

- Type prediction at each decoding position

Post:	我喜欢小动物(I like little animals)					
Response:	你(you)	喜欢(like)	兔子(rabbit)	吗( <i>particle</i> )	?	_EOS
Interrogative	0.09	0.02	0.01	0.85	1.00	0.01
Topic word	0.26	0.35	0.71	0.14	0.00	0.02
Ordinary word	0.65	0.63	0.28	0.01	0.00	0.97
Decoding steps	1	2	3	4	5	6



# Asking Questions in Chatbots

- ◎ Dataset: 490,000 post-response pairs collected from Weibo; 5,000 for test, 5000 for validation
  - ◆ All responses are of questioning form
- ◎ 66,547 different words, and 18,717 words appear more than 10 times



# Baselines

---

- **Seq2Seq**: A simple encoder-decoder model ([Luong et al., 2015](#))
- **Mechanism-Aware (MA)**: Multiple responding mechanisms represented by real-valued vectors ([Zhou et al., 2017](#))
- **Topic-Aware (TA)**: Topic Aware Model by incorporating topic words ([Xing et al., 2017](#))
- **Elastic Responding Machine (ERM)**: Enhanced MA using reinforcement learning ([Zhou et al., 2018](#))



# Asking Questions in Chatbots

Model	Perplexity	Distinct-1	Distinct-2	TRR
Seq2Seq	63.71	0.0573	0.0836	6.6%
MA	<b>54.26</b>	0.0576	0.0644	4.5%
TA	58.89	0.1292	0.1781	8.7%
ERM	67.62	0.0355	0.0710	4.5%
STD	56.77	0.1325	0.2509	12.1%
HTD	56.10	<b>0.1875</b>	<b>0.3576</b>	<b>43.6%</b>

Table 1: Results of automatic evaluation.



# Asking Questions in Chatbots

- Manual evaluation: Appropriateness, richness, willingness

Models	Appropriateness			Richness			Willingness		
	Win (%)	Lose (%)	Tie (%)	Win (%)	Lose (%)	Tie (%)	Win (%)	Lose (%)	Tie (%)
STD vs. Seq2Seq	42.0	38.6	19.4	37.2**	15.2	47.6	45.4*	38.6	16.0
STD vs. MA	39.6*	31.2	29.2	32.6**	16.8	50.6	49.4**	27.0	23.6
STD vs. TA	42.2	40.0	17.8	49.0**	5.4	45.6	47.6*	40.2	12.2
STD vs. ERM	43.4*	34.4	22.2	60.6**	13.2	26.2	43.2*	36.8	20.0
HTD vs. Seq2Seq	50.6**	30.6	18.8	46.0**	10.2	43.8	58.4**	33.2	8.4
HTD vs. MA	54.8**	24.4	20.8	45.0**	17.0	38.0	67.0**	18.0	15.0
HTD vs. TA	52.0**	38.2	9.8	55.0**	5.4	39.6	62.6**	31.0	6.4
HTD vs. ERM	64.8**	23.2	12.0	72.2**	8.4	19.4	56.6**	36.6	6.8
HTD vs. STD	52.0**	33.0	15.0	38.0**	26.2	35.8	61.8**	30.6	7.6

Yansen Wang, Chenyi Liu, Minlie Huang, Liqiang Nie.

Learning to ask questions in open-domain conversation systems. **ACL 2018.**



# Asking Questions in Chatbots

---

**Post:** 我们昨天去唱歌了  
We went to karaoke yesterday.

---

**Seq2Seq:** 什么时候?  
**When?**

**MA:** 你怎么知道我是在问  
**How** did you know I am questioning you?

**TA:** 什么?  
**What?**

**ERM:** 什么情况?  
**What** happened ?

**STD:** 去哪里唱歌?  
**Where** did you sing karaoke?

**HTD:** 你们几个人唱歌?  
**How many** people were singing with you?

---

Yansen Wang, Chenyi Liu, Minlie Huang, Liqiang Nie.

Learning to ask questions in open-domain conversation systems. **ACL 2018.**



# Interactiveness:

## Achieving Different Purposes

### by

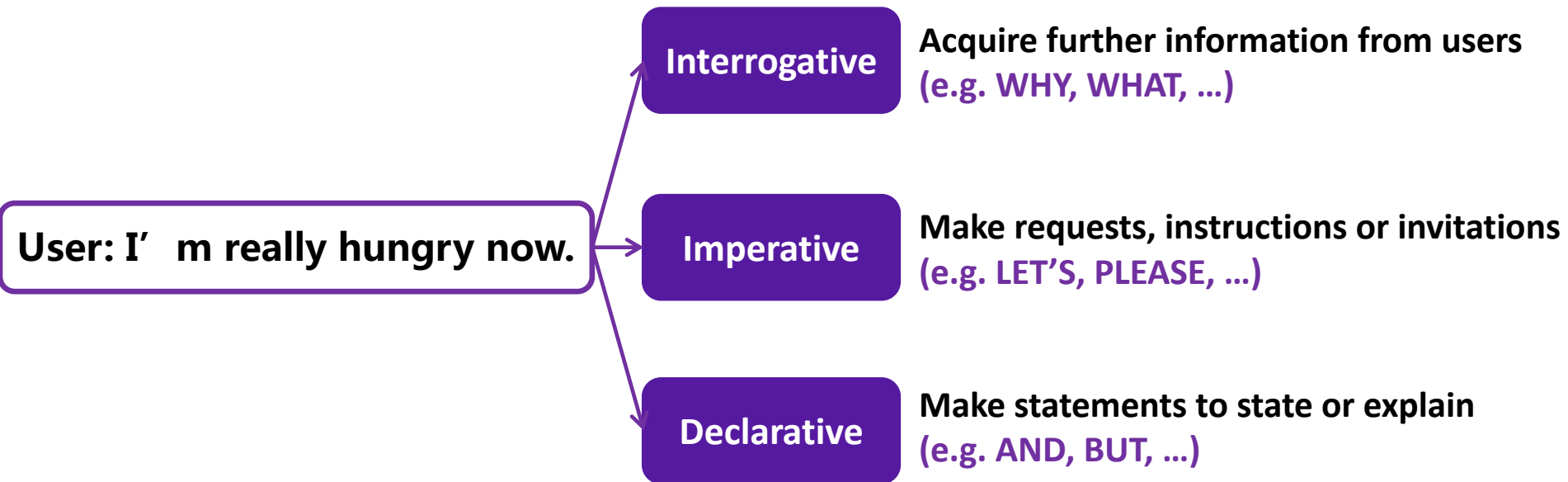
## Controlling Sentence

## Function



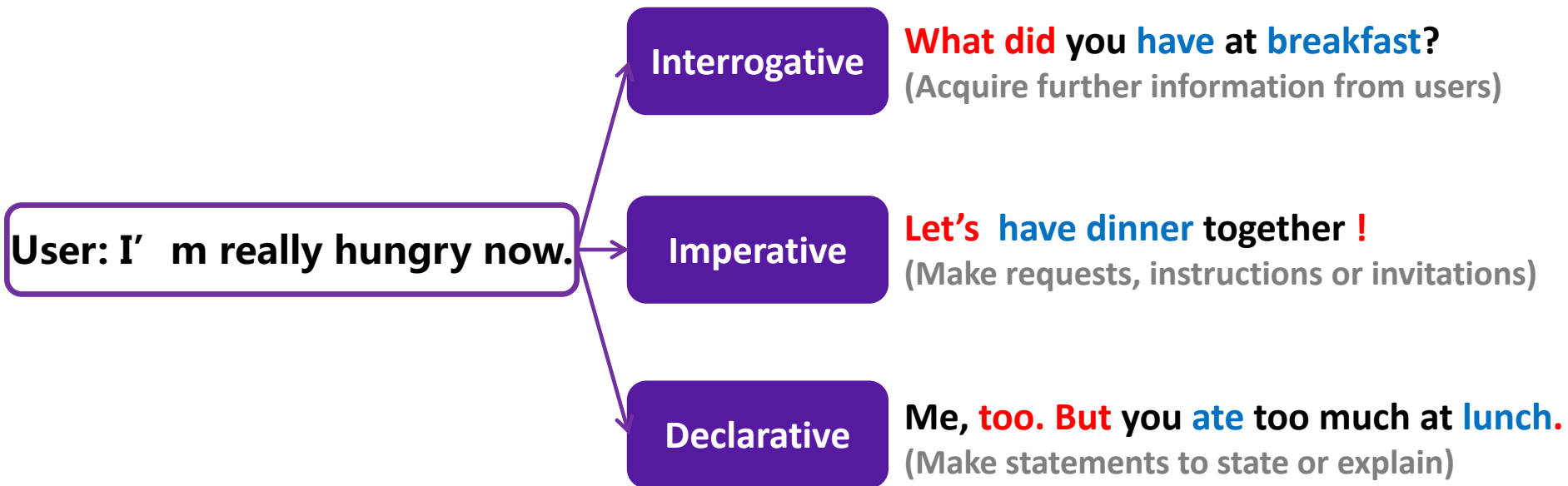
# Controlling Sentence Function

- ◎ Sentence function indicates different conversational purposes.



# Controlling Sentence Function

- Response with controlled sentence function requires a **global plan** of *function-related*, *topic* and *ordinary* words.



● Function-related words

● Topic words

● Ordinary words



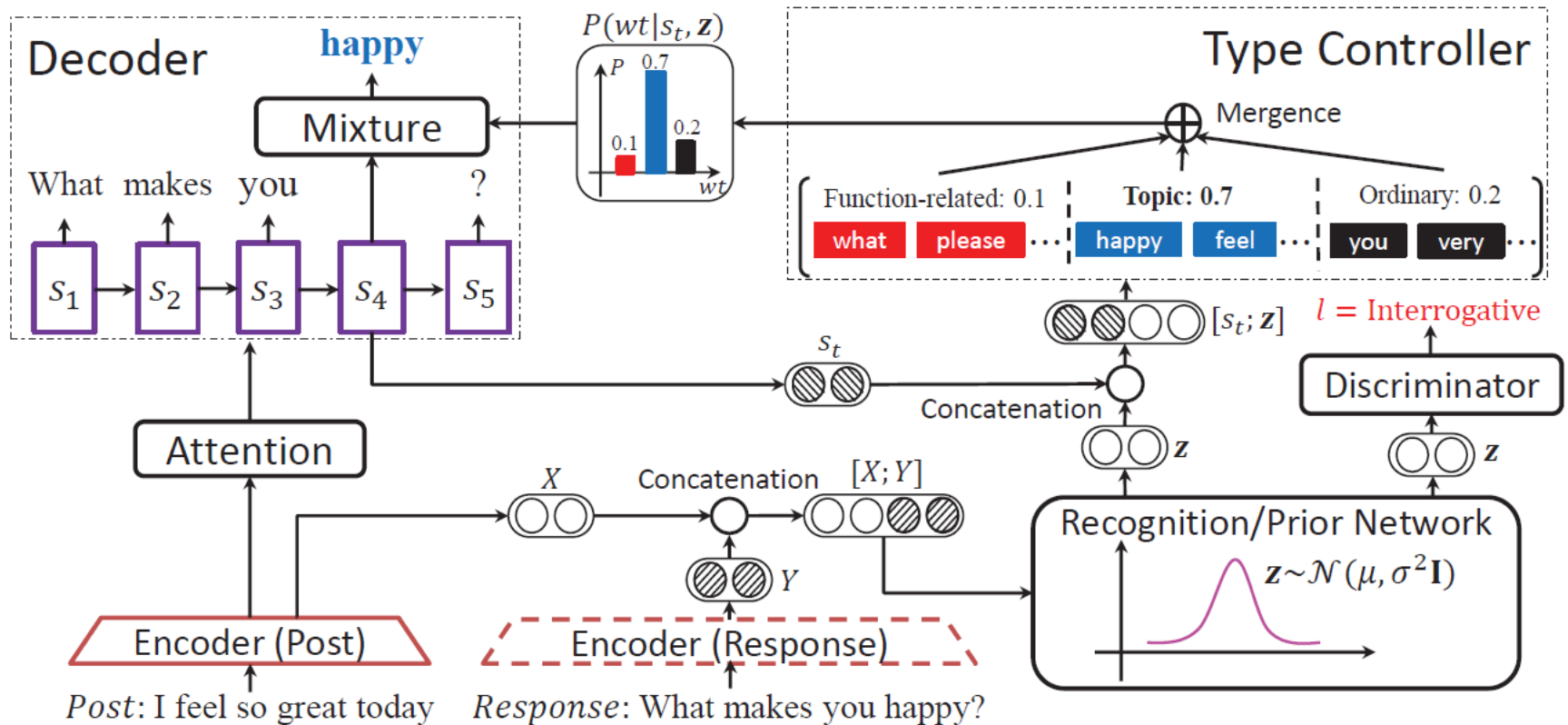
# Controlling Sentence Function

- ◎ Key differences to other controllable text generation tasks:
  - ◆ **Global Control**: adjust the global structure of the entire text, including changing word order and word patterns
  - ◆ **Compatibility**: controllable sentence function + informative content
- ◎ Solutions:
  - ◆ **Continuous Latent Variable**: project different sentence functions into different regions in a latent space + capture word patterns within a sentence function
  - ◆ **Type Controller**: arrange different types of words at proper decoding positions by estimating a distribution over three word types



# Controlling Sentence Function

## Conditional Variational Autoencoder (CVAE) Framework



# Controlling Sentence Function

- Dataset: post-response pairs with sentence function labels

Training	#Post	1,963,382	
	#Response	Interrogative	618,340
		Declarative	672,346
		Imperative	672,696
Validation	#Post	24,034	
	#Response	Interrogative	7,045
		Declarative	9,685
		Imperative	7,304
Test	#Post	6,000	

Pei Ke, Jian Guan, Minlie Huang, Xiaoyan Zhu.

Generating Informative Responses with Controlled Sentence Function. **ACL 2018.**



# Controlling Sentence Function

- Automatic Evaluation: Perplexity, Distinct-1/2, Accuracy

Model	PPL	Dist-1	Dist-2	ACC
c-seq2seq	57.14	949/.007	5177/.041	0.973
MA	<b>46.08</b>	745/.005	2952/.027	0.481
KgCVAE	56.81	1531/.009	10683/.070	0.985
Our Model	55.85	<b>1833/.008</b>	<b>15586/.075</b>	<b>0.992</b>

Table 3: Automatic evaluation with perplexity (PPL), distinct-1 (Dist-1), distinct-2 (Dist-2), and accuracy (ACC). The integers in the Dist-\* cells denote the total number of distinct  $n$ -grams.



# Controlling Sentence Function

## Manual Evaluation: Grammaticality, Appropriateness, Informativeness

Model	Interrogative			Declarative			Imperative		
	Gram.	Appr.	Info.	Gram.	Appr.	Info.	Gram.	Appr.	Info.
Ours vs. c-seq2seq	0.534	0.536	0.896*	0.630*	0.573*	0.764*	0.685*	0.504	0.893*
Ours vs. MA	0.802*	0.602*	0.675*	0.751*	0.592*	0.617*	0.929*	0.568*	0.577*
Ours vs. KgCVAE	0.510	0.626*	0.770*	0.546*	0.515*	0.744*	0.780*	0.521*	0.837*

Table 4: Manual evaluation results for different functions. The scores indicate the percentages that our model wins the baselines after removing tie pairs. The scores of our model marked with \* are significantly better than the competitors (Sign Test,  $p\text{-value} < 0.05$ ).



# Consistency:

## Behaving More Consistently with Personality



# Personality of Chatting Machine

- ◎ Personality is important for game, custom service, etc.

## General seq2seq model

User: Are you a boy or a girl?

Chatbot: I am a boy.

User: Are you a girl?

Chatbot: Yes, I am a girl.

**User:** 你几岁了?

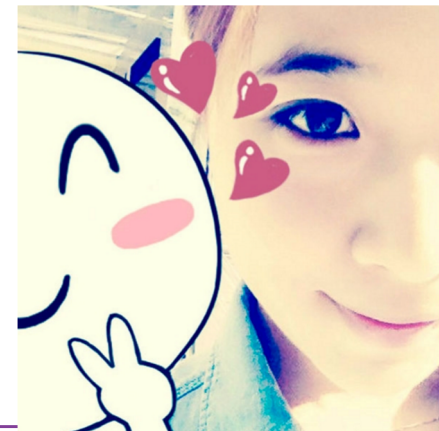
**小冰:** 我啊。。。反正是九零后!

**User:** 那你是几几年出生的?

**小冰:** 错, **75**年阳历一月出生。

**User:** 你是九零后, 怎么会75年出生?

**小冰:** 生在九零后, 在深圳只能被当做八零后了。



# Personality of Chatting Machine

- ◎ Passing the **Turning Test**?
  - ◆ **Deep semantic understanding**
  - ◆ Existing chatting machine lacks **identity or personality**
- ◎ Personality is a well-defined concept in psychology(Norman, 1963; Gosling et al., 2003)
- ◎ Extremely **subtle, implicit** in language expression:
  - ◆ Age, gender, language, speaking style, level of knowledge, areas of expertise
- ◎ Existing works
  - ◆ **Implicit personalization**: learn **implicit** conversation style (Li et al., 2016; Al-Rfou et al., 2016)
  - ◆ Require dialogue data from different users with **user attributes tagged**

# Personality of Chatting Machine

- ◎ Deliver coherent conversations w.r.t. **identity/personality**

## Generic Dialogue Data for Training

UserA: how old are you?

UserB: I am **six**.

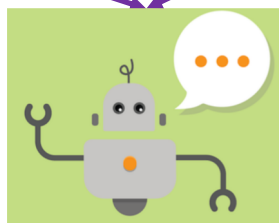
UserA: do you like to play piano?

UserB: I play **violin**.

## Pre-specified Chatbot Profile

Profile key	Profile value
Name	汪仔(Wang Zai)
Age	三岁(3)
Gender	男孩(Boy)
Hobbies	动漫(Cartoon)
Speciality	钢琴(Piano)

Personality-coherent  
Chatbot



## Generated Dialogues

User: how old are you?

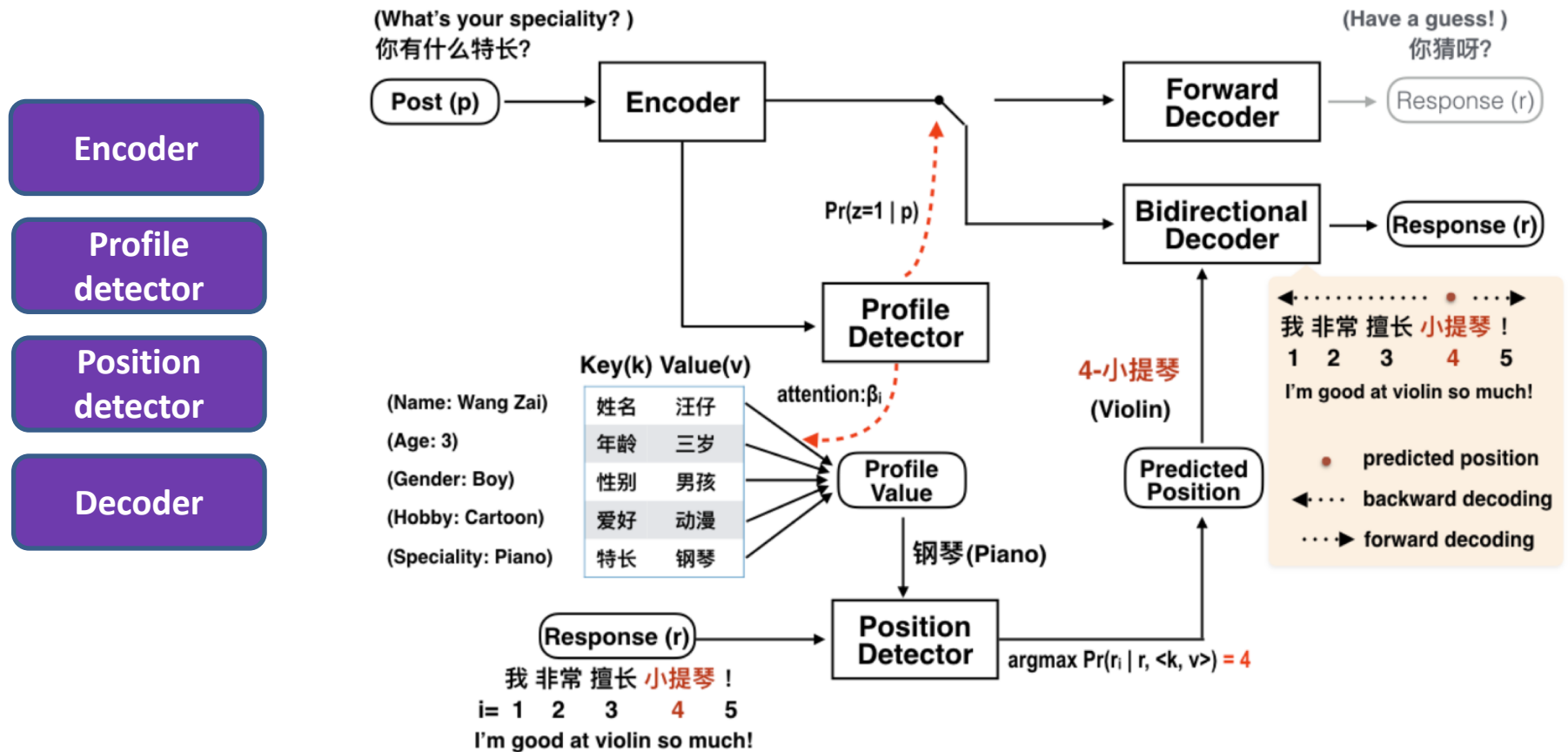
Machine: I am **three years old**.

User: do you like to play piano?

Machine: Yes, I play **piano**.



# Personality of Chatting Machine



•Qiao Qian, Minlie Huang, Haizhou Zhao, Jingfang Xu, Xiaoyan Zhu. Assigning personality/identity to a chatting machine for coherent conversation generation. **IJCAI-ECAI 2018**.



# Personality of Chatting Machine

- ◎ WD: 9,697,651 post-response pairs from Weibo
- ◎ 76,930 pairs from WD for 6 profile keys (name, gender, age, city, weight, constellation) with about 200 regular expression patterns, each annotated to positive or negative
- ◎ 42,193 positive pairs, each mapped to one of the keys
- ◎ Manual Dataset: real, human-written conversational posts



# Personality of Chatting Machine

## Post-level evaluation

Method	Nat.	Logic	Cor.
Seq2Seq	71.8%	56.0%	23.8%
Seq2Seq +PV	72.0%	56.0%	41.3%
Seq2Seq +PVD	73.3%	52.5%	38.0%
Our Model -PD	82.7%	51.7%	38.3%
Our Model	<b>83.3%</b>	<b>59.5%</b>	<b>45.8%</b>

## Session-level evaluation

Method	Consistency	Variety
Seq2Seq	1.3%	1.0%
Seq2Seq +PV	47.0%	1.3%
Seq2Seq +PVD	40.0%	7.5%
Our Model -PD	38.8%	16.0%
Our Model	<b>49.5%</b>	<b>27.8%</b>

## Generated sample responses that exhibit session-level consistency

Chinese	English(Translated)
U:你对什么事感兴趣 S:打篮球 U:你都有啥爱好呀 S:篮球 U:告诉我你兴趣所在 S:我喜欢篮球	U:What are you interested in? S:Playing basketball. U:What's your hobby? S:Basketball. U:Tell me your interest. S:I like to play basketball.
U:你还没说你几岁呢 S:我三岁了 U:你今年有15了不 S:我还没到呢 U:你多大啦 S:3岁了	U:You haven't told me your age. S:I'm three years old. U:Are you 15 years old or not? S:I'm not yet. U:How old are you? S:Three years old.

•Qiao Qian, Minlie Huang, Haizhou Zhao, Jingfang Xu, Xiaoyan Zhu. *Assigning personality/identity to a chatting machine for coherent conversation generation.* **IJCAI-ECAI 2018.**



# Semantics:

## Better Understanding and Generation with Commonsense Knowledge

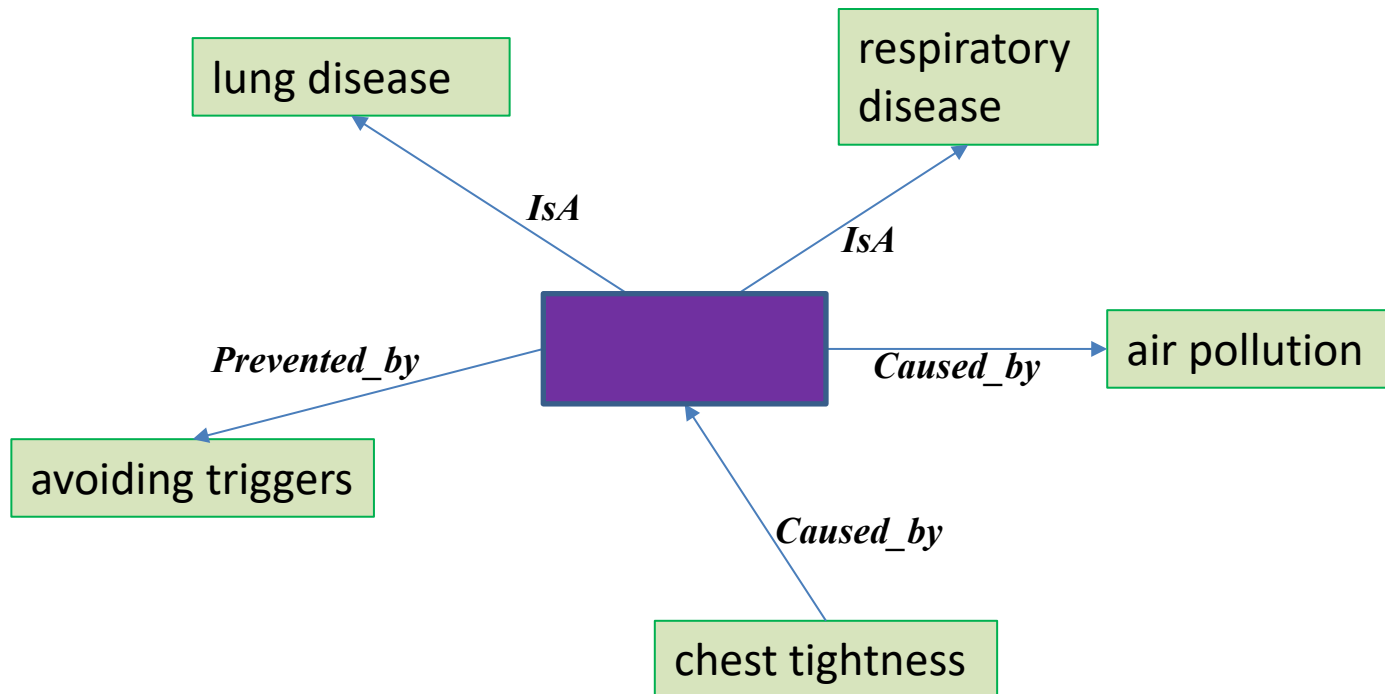


# Commonsense Knowledge

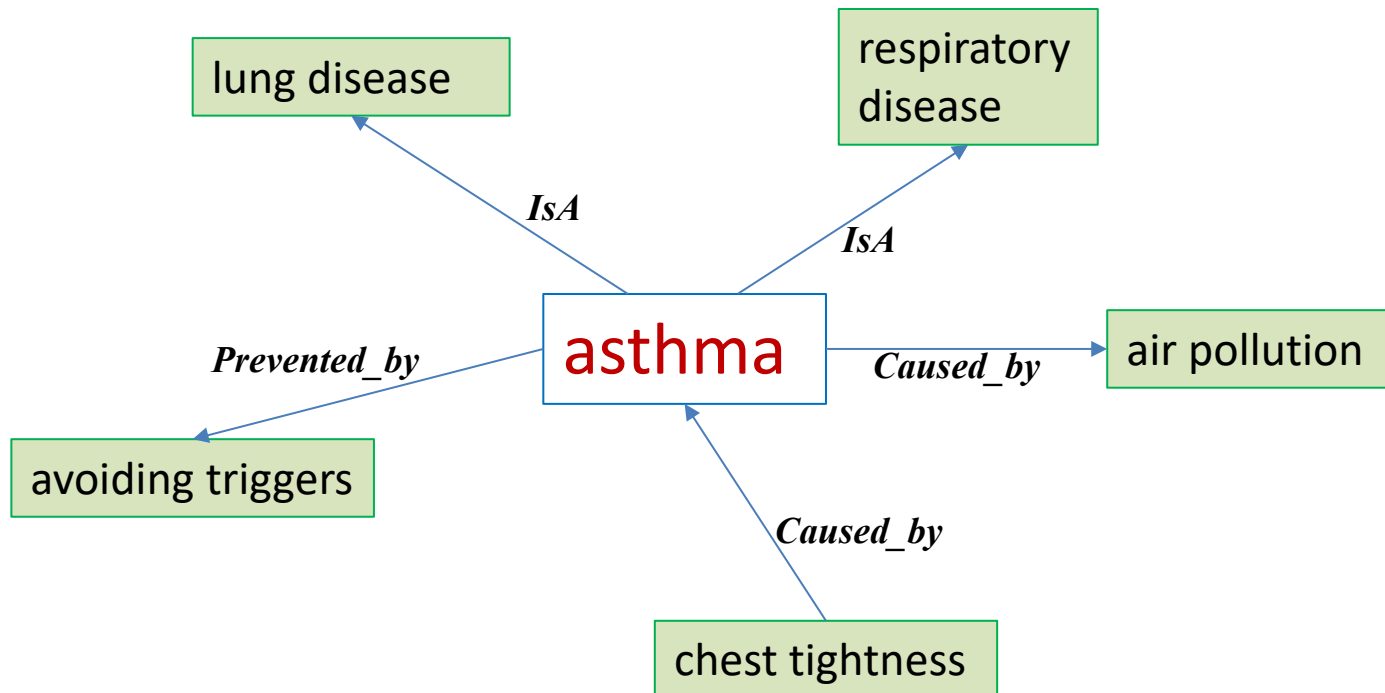
- ◎ **Commonsense knowledge** consists of facts about the everyday world, that all humans are expected to know. (Wikipedia)
  - ◆ Lemons are sour
  - ◆ Tree has leafs
  - ◆ Dog has four legs
- ◎ Commonsense Reasoning ~ **Winograd Schema Challenge:**
  - The trophy would not fit in the brown suitcase because it was too **big**. What was too **big**?
  - The trophy would not fit in the brown suitcase because it was too *small*. What was too *small*?



# Commonsense Knowledge in Chatbots



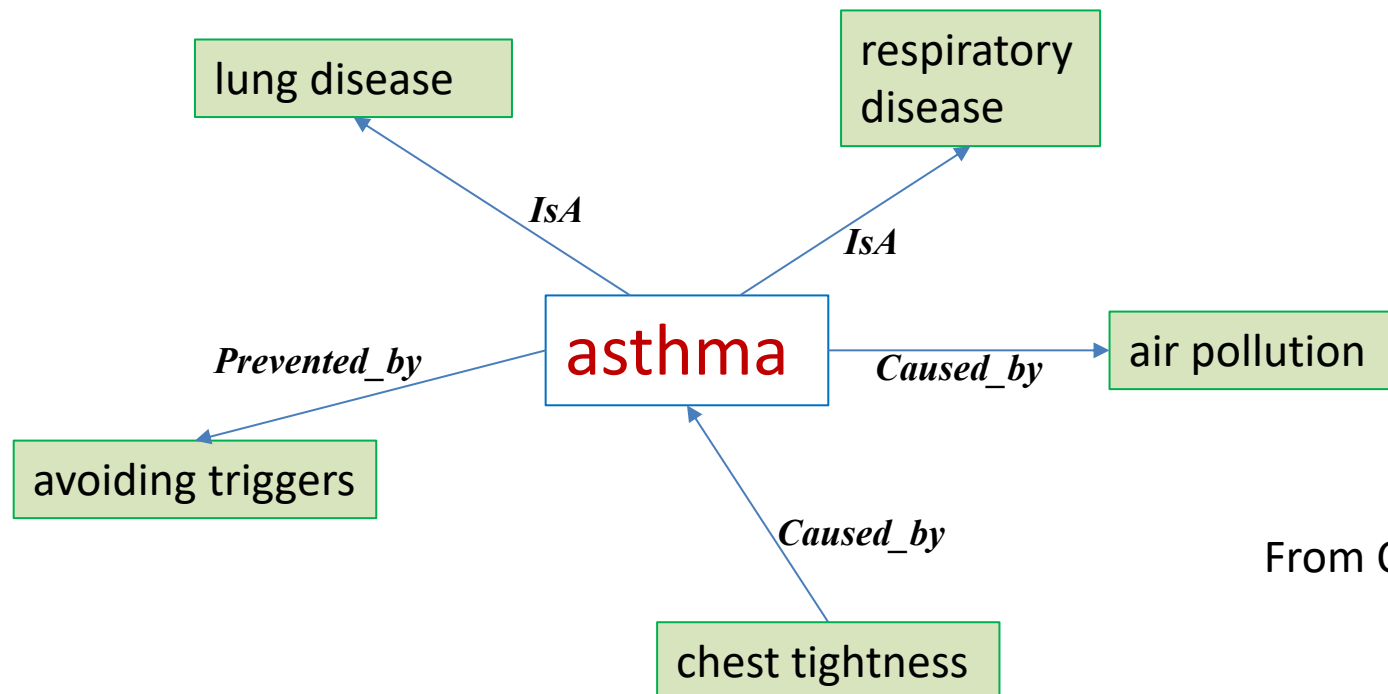
# Commonsense Knowledge in Chatbots



# Commonsense Knowledge in Chatbots

Post: I have an **asthma** since three years old.

Triples in knowledge graph:  
(lung disease, *IsA*, **asthma**)  
(**asthma**, *Prevented\_by*, avoiding triggers)



From ConceptNet

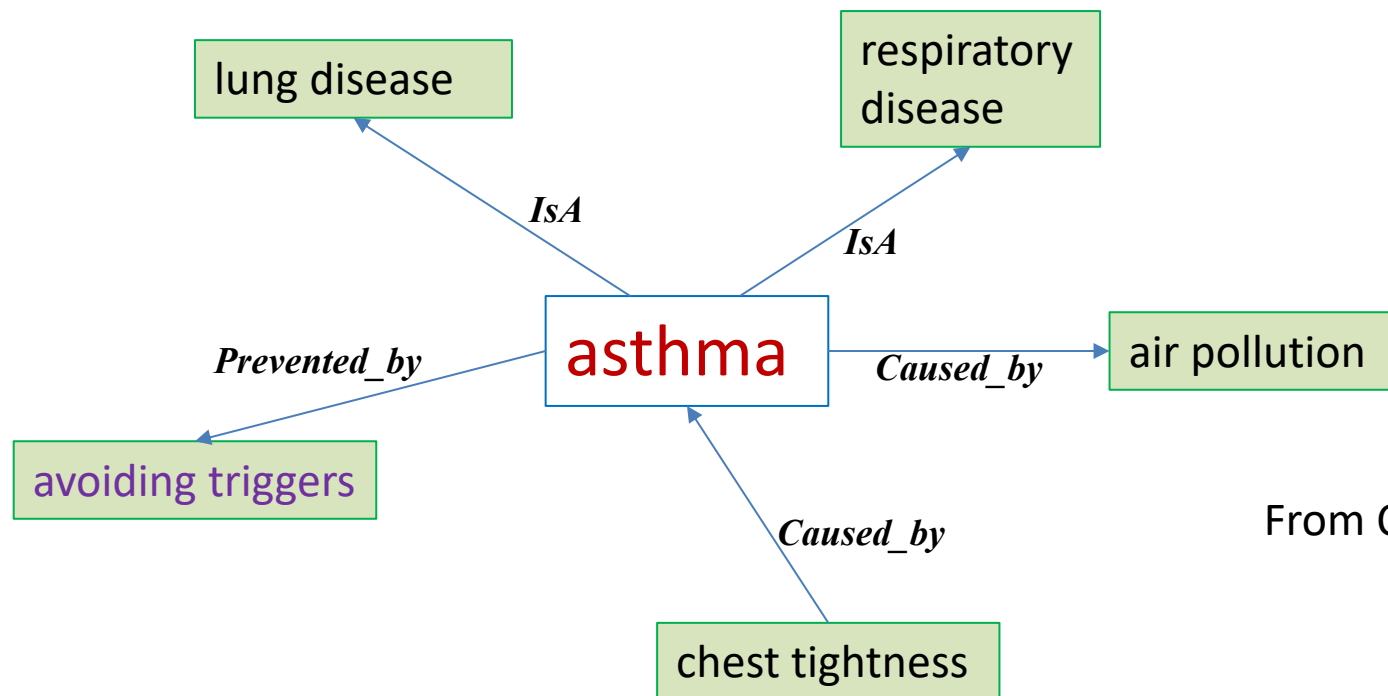


# Commonsense Knowledge in Chatbots

Post: I have an **asthma** since three years old.

Triples in knowledge graph:  
(lung disease, IsA, **asthma** )  
(**asthma**, Prevented\_by, avoiding triggers)

Response: I am sorry to hear that. Maybe **avoiding triggers** can prevent **asthma** attacks.



From ConceptNet

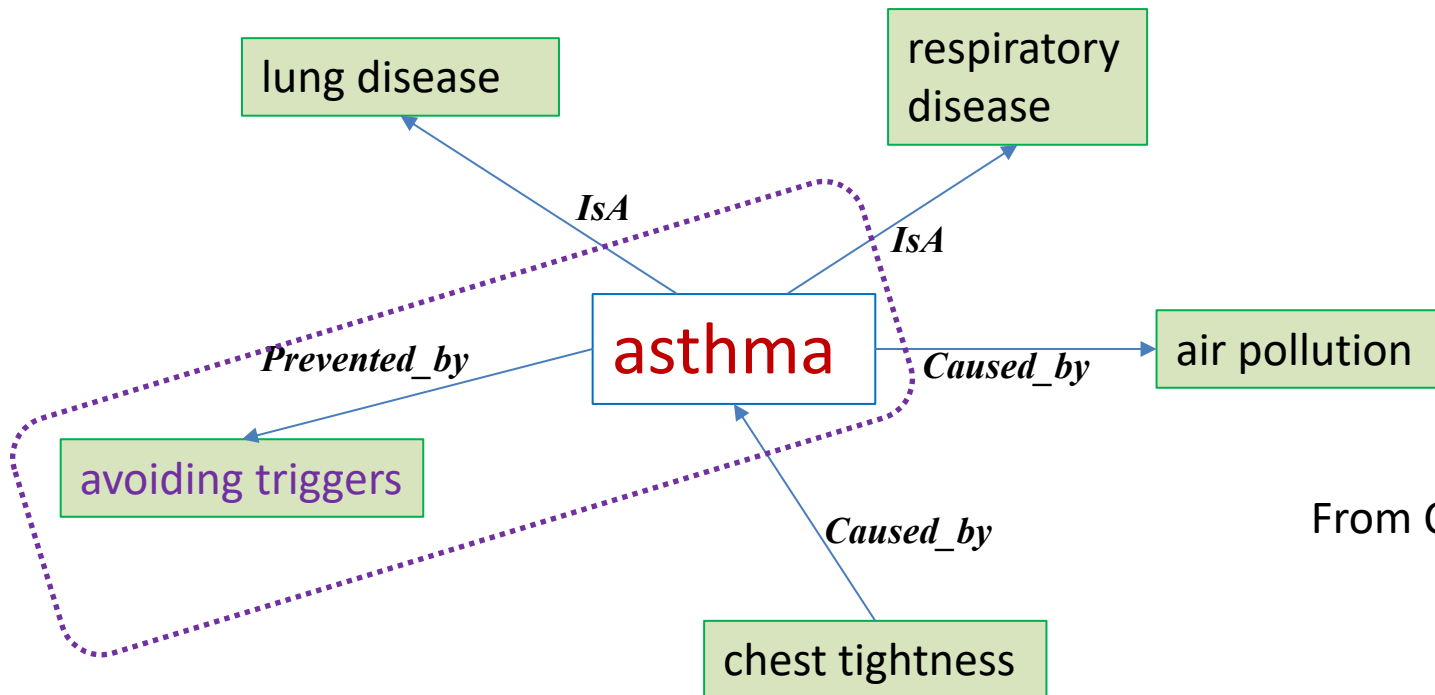


# Commonsense Knowledge in Chatbots

Post: I have an **asthma** since three years old.

Triples in knowledge graph:  
(lung disease, IsA, **asthma**)  
(**asthma**, Prevented\_by, avoiding triggers)

Response: I am sorry to hear that. Maybe **avoiding triggers** can prevent **asthma** attacks.



From ConceptNet



# Commonsense Knowledge in Chatbots

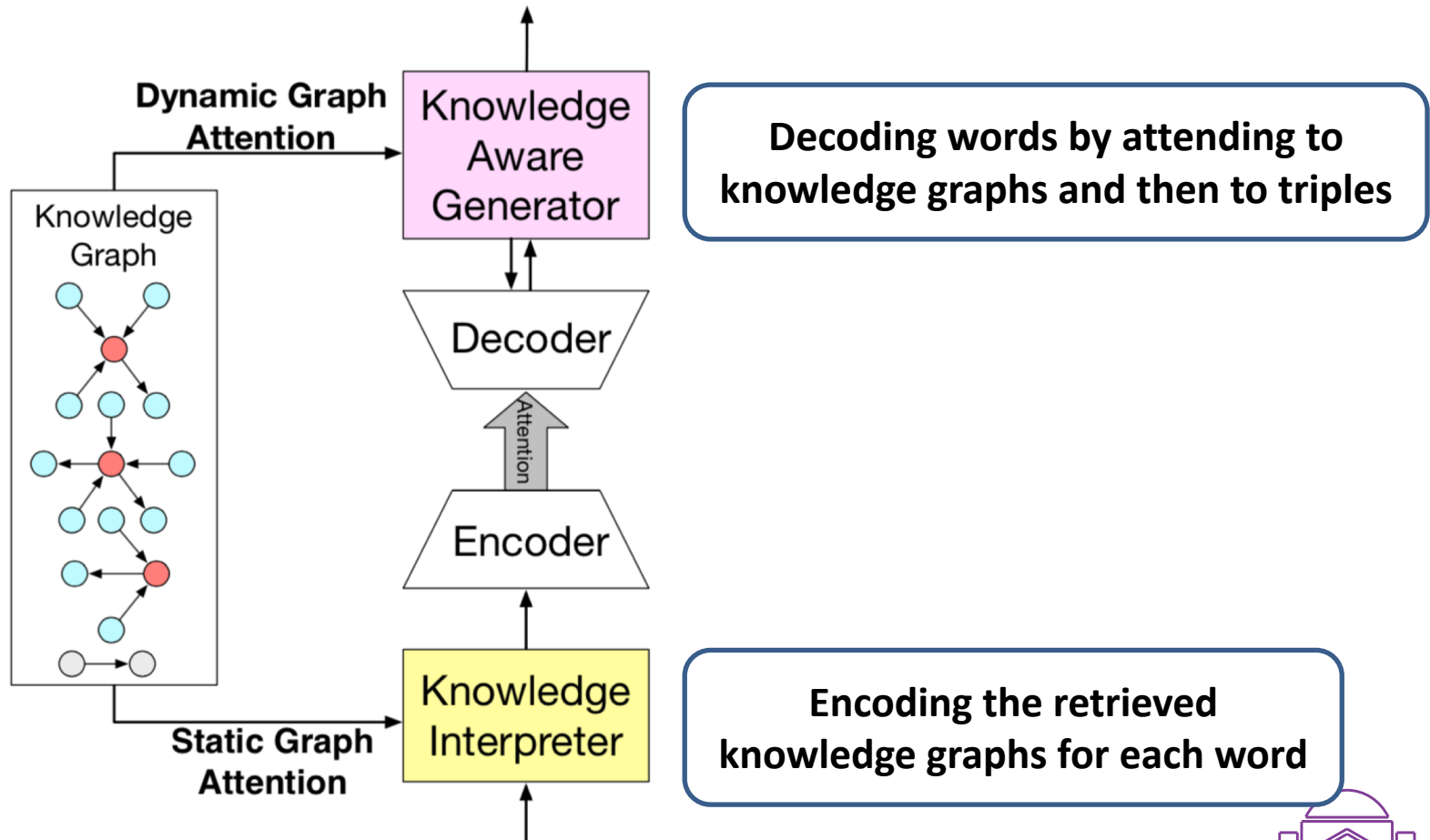
- Conversation is all about **semantic understanding**
- Commonsense knowledge is important for **language understanding and generation**

Post	Why are you so <b>breakable</b> ?
Knowledge	( <b>glass</b> , RelatedTo, <b>breakable</b> ), ( <b>brittle</b> , RelatedTo, <b>breakable</b> ), ( <b>rule</b> , RelatedTo, <b>breakable</b> )
Seq2Seq	I'm not a <b>OOV</b> , i 'm just a <b>OOV</b> .
MemNet	I'm not <b>OOV</b> . I'm just a really nice person.
CopyNet	I'm not. I'm just a lurker.
CCM	Because I'm a <b>brittle</b> man .



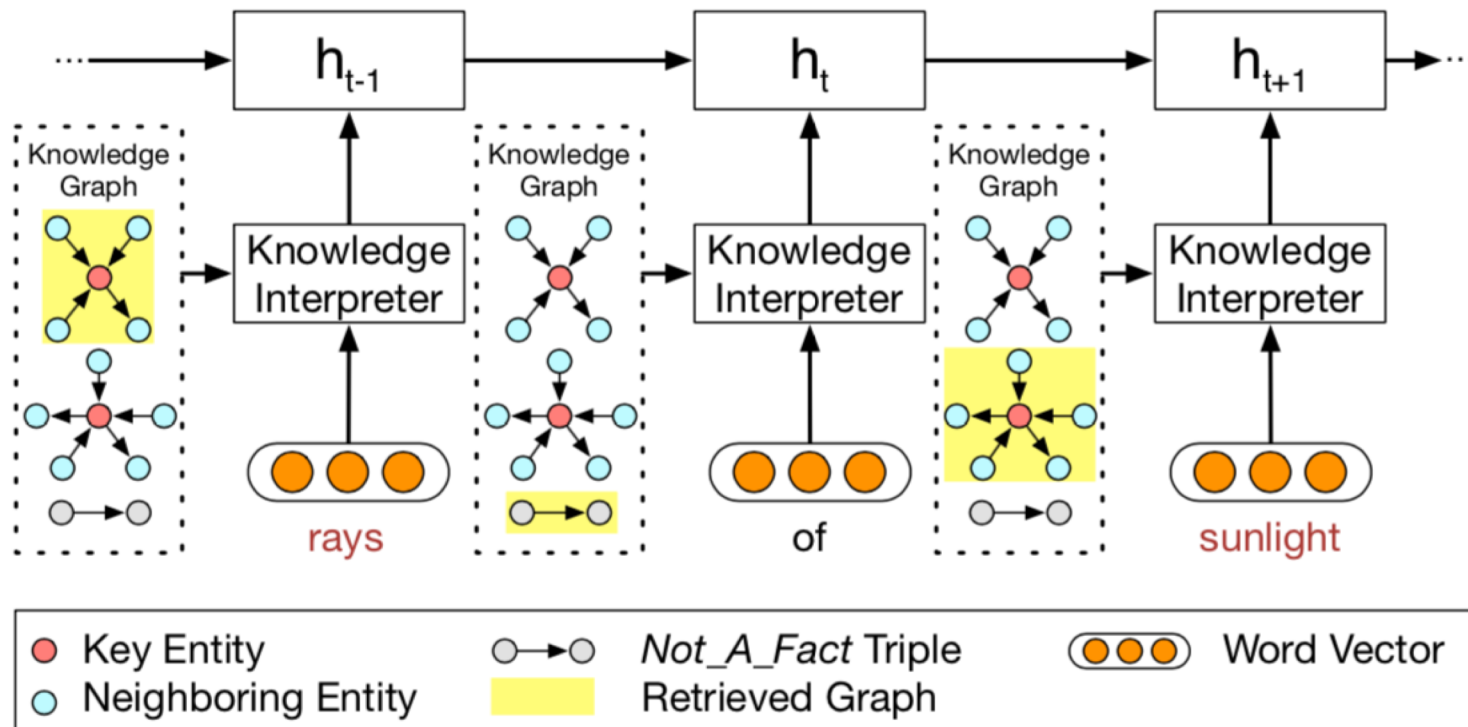
# Commonsense Knowledge in Chatbots

Output: Because I'm a brittle man.



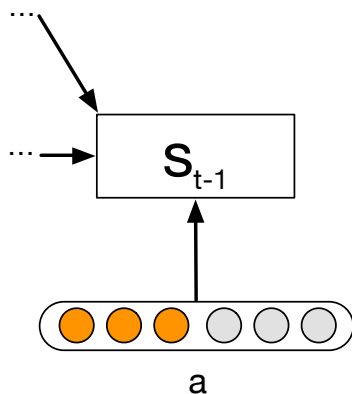
# Commonsense Knowledge in Chatbots

**Static graph attention:** encoding semantics in graph,  
Feeding knowledge-enhanced info. into the encoder



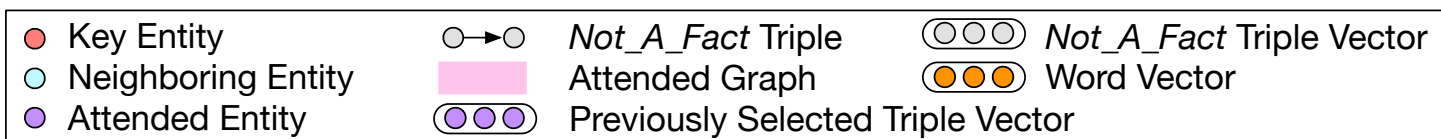
# Commonsense Knowledge in Chatbots

**Dynamic graph attention:** first attend a graph, then to a triple within that graph, finally generate with the words in a graph



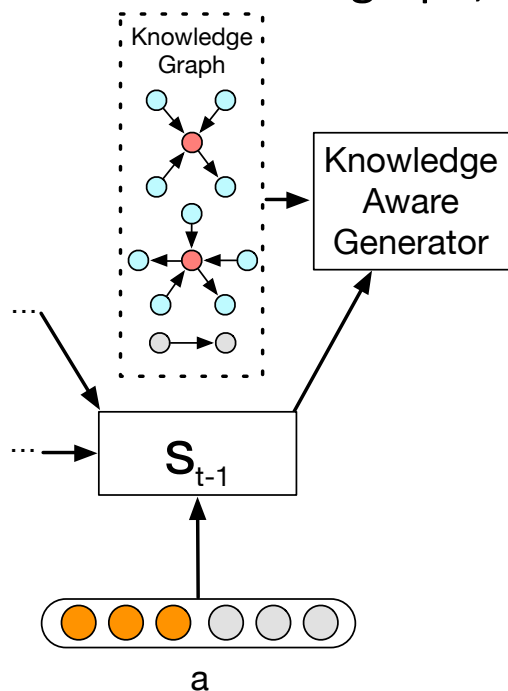
$$s_{t+1} = \text{GRU}(s_t, [c_t; c_t^g; c_t^k; e(y_t)]),$$

$$e(y_t) = [w(y_t); k_j],$$



# Commonsense Knowledge in Chatbots

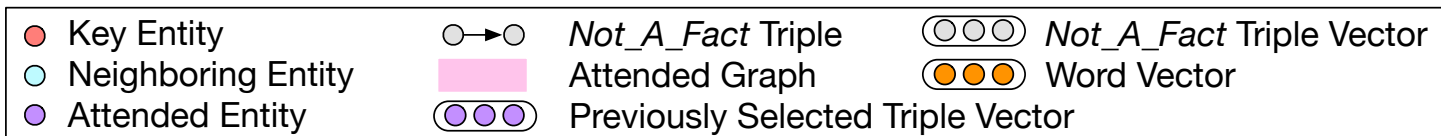
**Dynamic graph attention:** first attend a graph, then to a triple within that graph, finally generate with the words in a graph



$$\mathbf{g}_i = \sum_{n=1}^{N_{g_i}} \alpha_n^s [\mathbf{h}_n; \mathbf{t}_n],$$

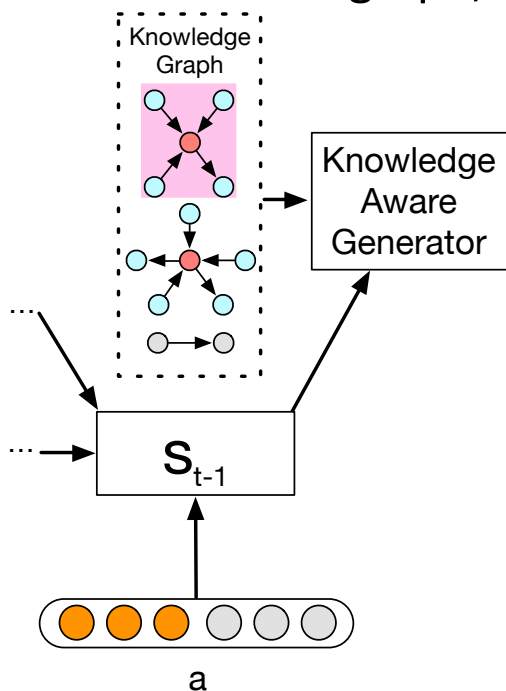
$$\alpha_n^s = \frac{\exp(\beta_n^s)}{\sum_{j=1}^{N_{g_i}} \exp(\beta_j^s)},$$

$$\beta_n^s = (\mathbf{W}_r \mathbf{r}_n)^\top \tanh(\mathbf{W}_h \mathbf{h}_n + \mathbf{W}_t \mathbf{t}_n),$$



# Commonsense Knowledge in Chatbots

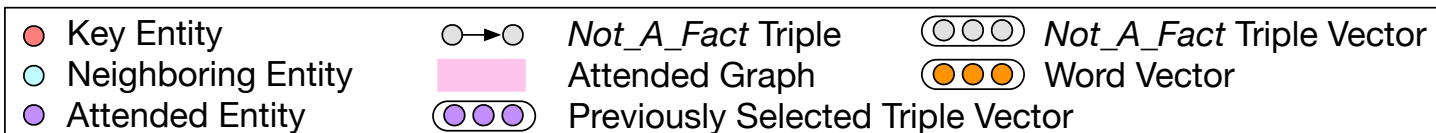
**Dynamic graph attention:** first attend a graph, then to a triple within that graph, finally generate with the words in a graph



$$\mathbf{c}_t^g = \sum_{i=1}^{N_G} \alpha_{ti}^g \mathbf{g}_i,$$

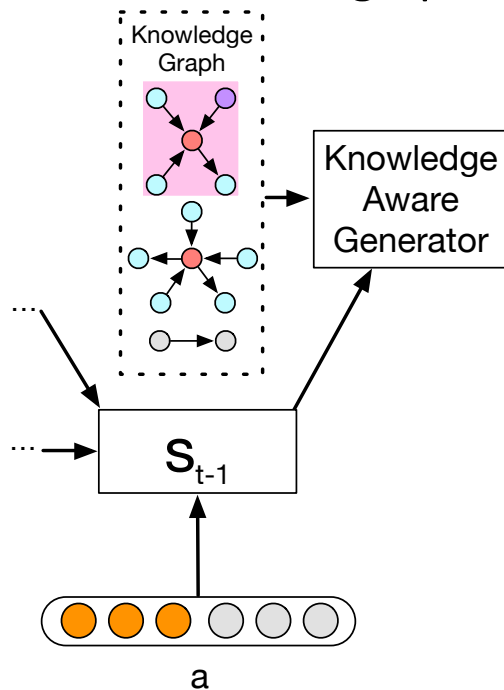
$$\alpha_{ti}^g = \frac{\exp(\beta_{ti}^g)}{\sum_{j=1}^{N_G} \exp(\beta_{tj}^g)},$$

$$\beta_{ti}^g = \mathbf{V}_b^\top \tanh(\mathbf{W}_b \mathbf{s}_t + \mathbf{U}_b \mathbf{g}_i),$$



# Commonsense Knowledge in Chatbots

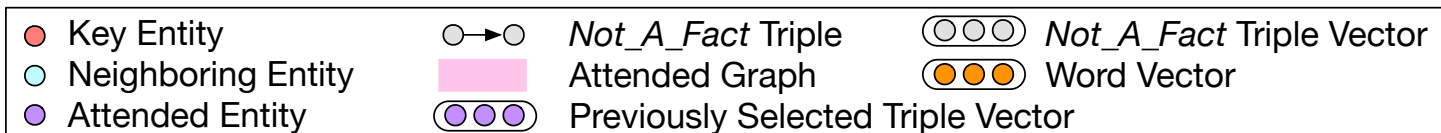
**Dynamic graph attention:** first attend a graph, then to a triple within that graph, finally generate with the words in a graph



$$\mathbf{c}_t^k = \sum_{i=1}^{N_G} \sum_{j=1}^{N_{g_i}} \alpha_{ti}^g \alpha_{tj}^k \mathbf{k}_j,$$

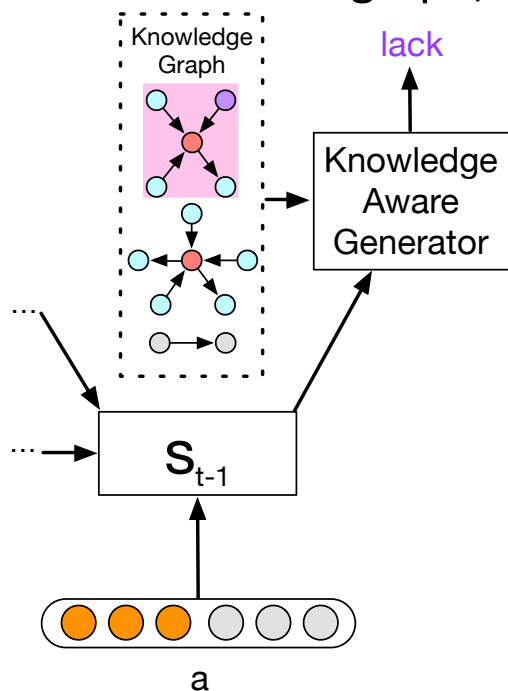
$$\alpha_{tj}^k = \frac{\exp(\beta_{tj}^k)}{\sum_{n=1}^{N_{g_i}} \exp(\beta_{tn}^k)},$$

$$\beta_{tj}^k = \mathbf{k}_j^\top \mathbf{W}_c \mathbf{s}_t,$$



# Commonsense Knowledge in Chatbots

**Dynamic graph attention:** first attend a graph, then to a triple within that graph, finally generate with the words in a graph



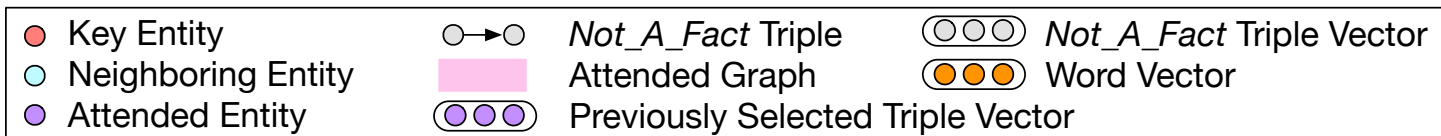
$$\mathbf{a}_t = [\mathbf{s}_t; \mathbf{c}_t; \mathbf{c}_t^g; \mathbf{c}_t^k],$$

$$\gamma_t = \text{sigmoid}(\mathbf{V}_o^\top \mathbf{a}_t),$$

$$P_c(y_t = w_c) = \text{softmax}(\mathbf{W}_o \mathbf{a}_t),$$

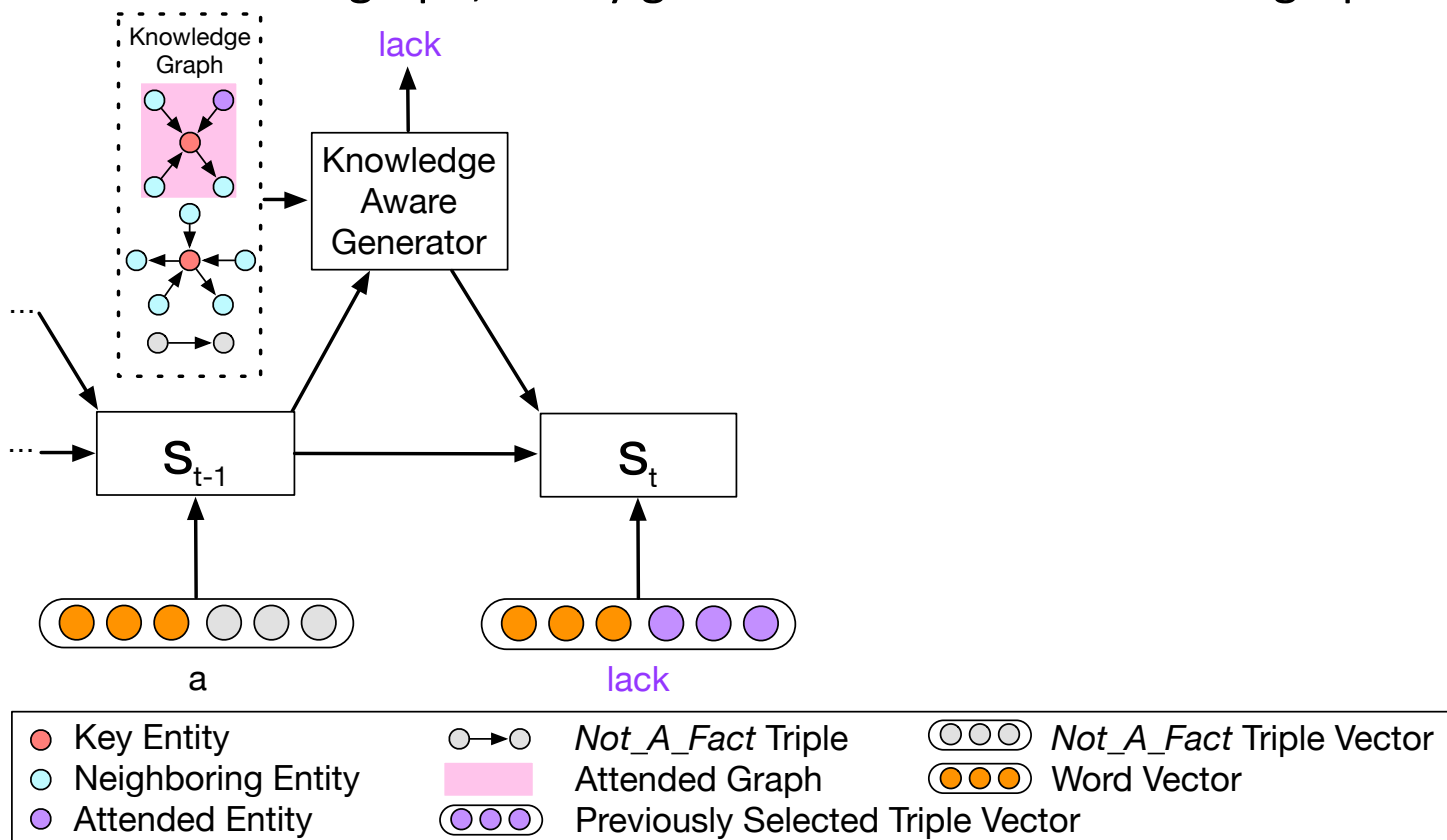
$$P_e(y_t = w_e) = \alpha_{ti}^g \alpha_{tj}^k,$$

$$y_t \sim \mathbf{o}_t = P(y_t) = \begin{bmatrix} (1 - \gamma_t) P_g(y_t = w_c) \\ \gamma_t P_e(y_t = w_e) \end{bmatrix},$$



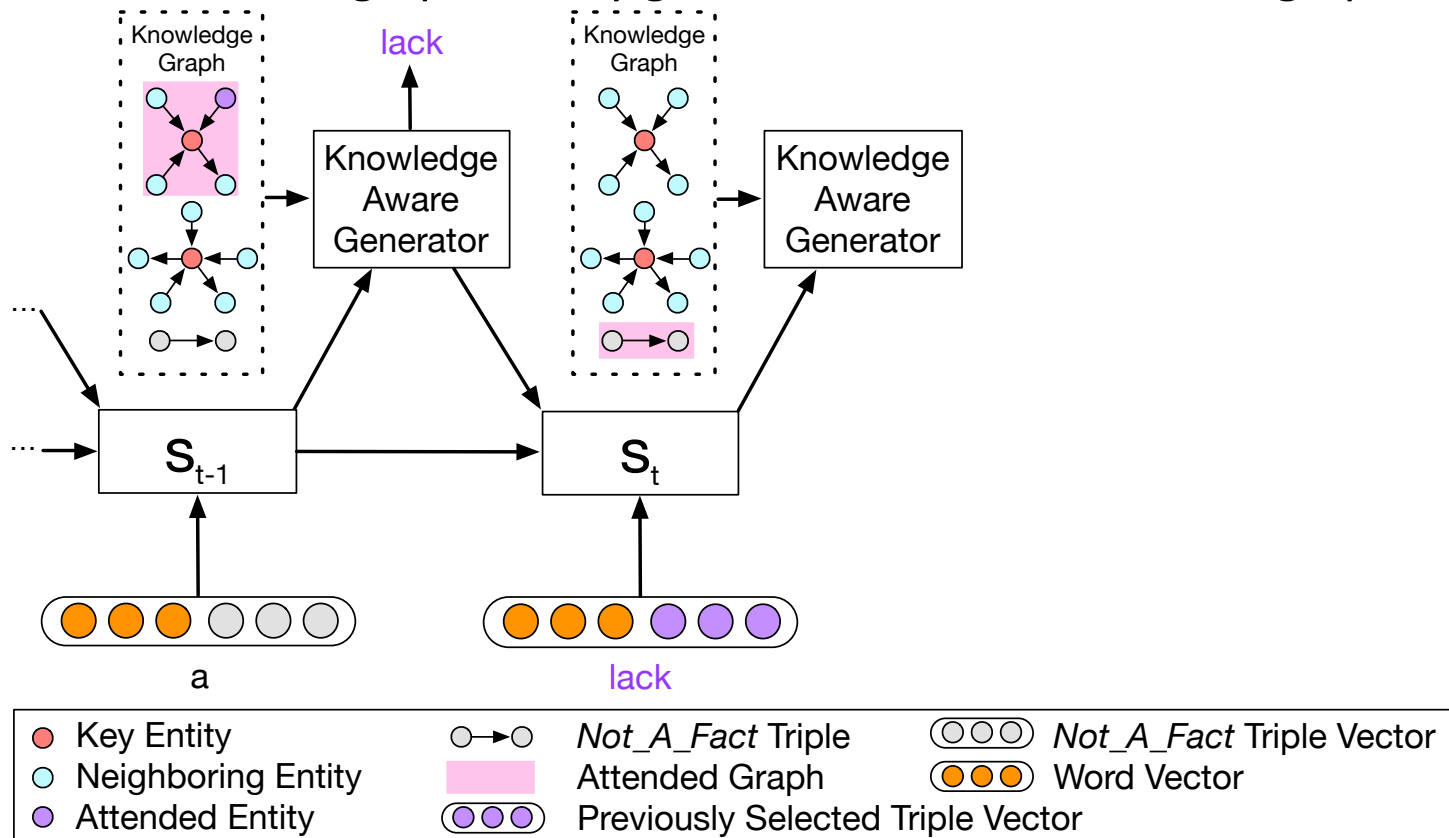
# Commonsense Knowledge in Chatbots

**Dynamic graph attention:** first attend a graph, then to a triple within that graph, finally generate with the words in a graph



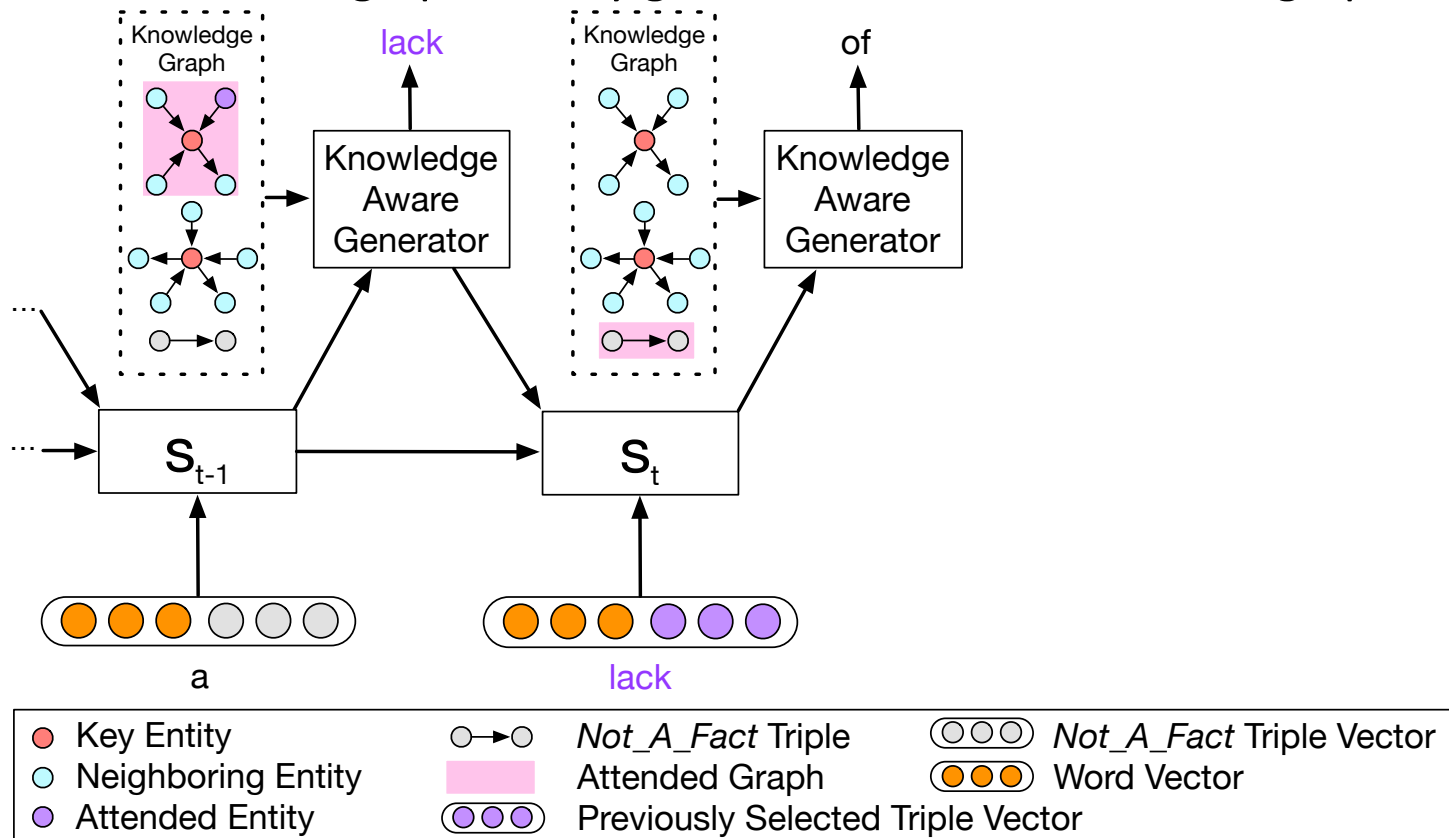
# Commonsense Knowledge in Chatbots

**Dynamic graph attention:** first attend a graph, then to a triple within that graph, finally generate with the words in a graph



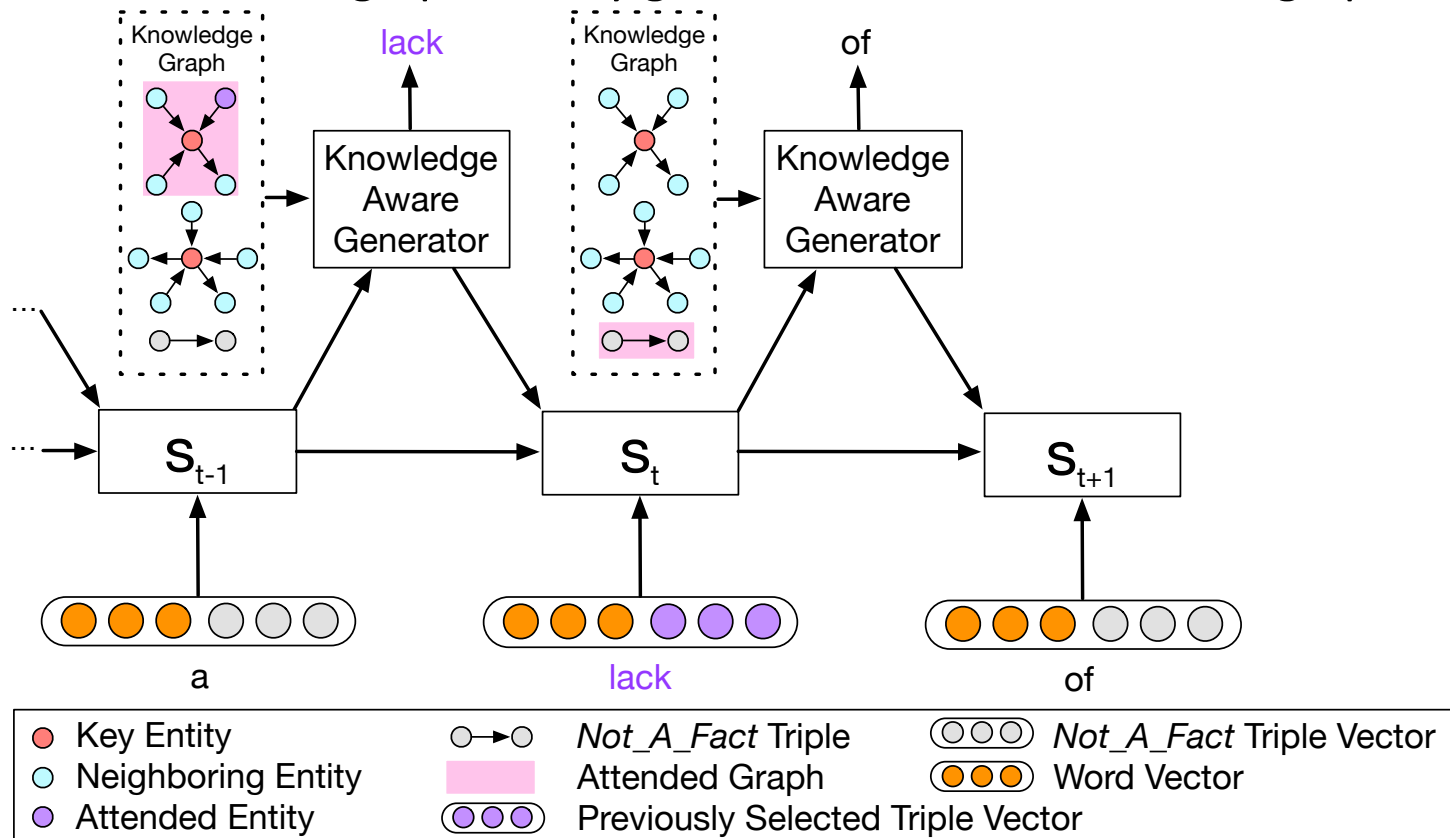
# Commonsense Knowledge in Chatbots

**Dynamic graph attention:** first attend a graph, then to a triple within that graph, finally generate with the words in a graph



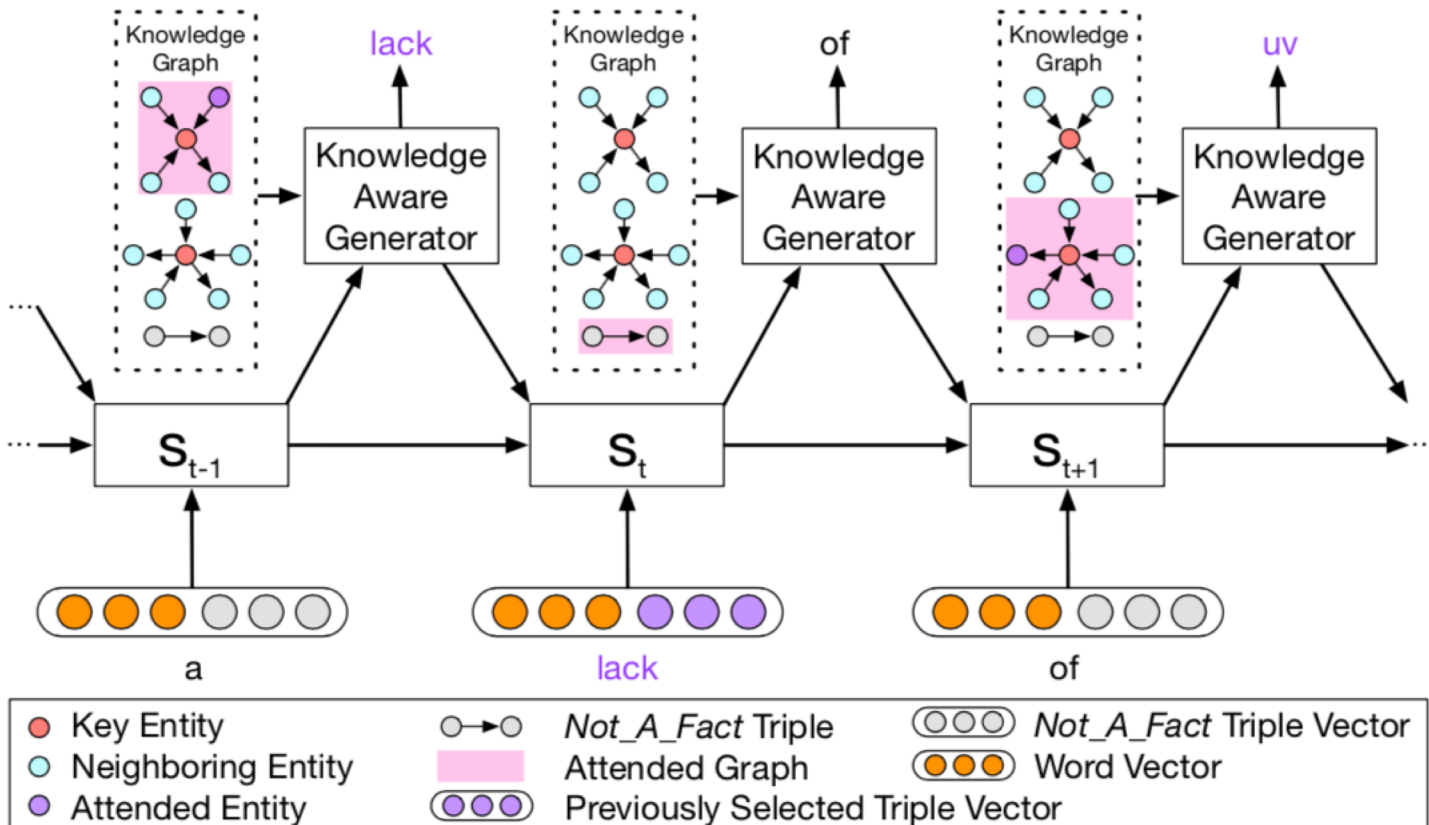
# Commonsense Knowledge in Chatbots

**Dynamic graph attention:** first attend a graph, then to a triple within that graph, finally generate with the words in a graph



# Commonsense Knowledge in Chatbots

**Dynamic graph attention:** first attend a graph, then to a triple within that graph, finally generate with the words in a graph



# Commonsense Knowledge in Chatbots

- Dataset: filtered from 10M reddit single-round dialogs

Conversational Pairs		Commonsense KB	
Training	3,384,185	Entity	21,471
Validation	10,000	Relation	44
Test	20,000	Triple	120,850

Table 1: Statistics of the dataset and the knowledge base.



# Commonsense Knowledge in Chatbots

## Automatic evaluation

Model	Overall		High Freq.		Medium Freq.		Low Freq.		OOV	
	ppx.	ent.	ppx.	ent.	ppx.	ent.	ppx.	ent.	ppx.	ent.
Seq2Seq	47.02	0.717	42.41	0.713	47.25	0.740	48.61	0.721	49.96	0.669
MemNet	46.85	0.761	41.93	0.764	47.32	0.788	48.86	0.760	49.52	0.706
CopyNet	40.27	0.96	36.26	0.91	40.99	0.97	42.09	0.96	42.24	0.96
CCM	<b>39.18</b>	<b>1.180</b>	<b>35.36</b>	<b>1.156</b>	<b>39.64</b>	<b>1.191</b>	<b>40.67</b>	<b>1.196</b>	<b>40.87</b>	<b>1.162</b>

## Manual evaluation

(Sign-test, p-value<0.005)

Model	Overall		High Freq.		Medium Freq.		Low Freq.		OOV	
	app.	inf.	app.	inf.	app.	inf.	app.	inf.	app.	inf.
CCM vs. Seq2Seq	0.616	0.662	0.605	0.656	0.549	0.624	0.636	0.650	0.673	0.716
CCM vs. MemNet	0.602	0.647	0.593	0.656	0.566	0.640	0.622	0.635	0.626	0.657
CCM vs. CopyNet	0.600	0.640	0.606	0.669	0.586	0.619	0.610	0.633	0.596	0.640

Hao Zhou, Tom Yang, Minlie Huang, Haizhou Zhao, Jingfang Xu, Xiaoyan Zhu.  
Commonsense Knowledge Aware Conversation Generation with Graph  
Attention. **IJCAI-ECAI 2018**, Stockholm, Sweden.



# Generation Examples 1

**Post:** He proposed March 5th. We will be married October 10th.  
So 7 months

**Knowledge:** (engaged, RelatedTo, married), (wedding, RelatedTo, married), (miss, DistinctFrom, married), (March, RelatedTo, months)

- ◉ **Seq2Seq:** I'm not sure how to answer that. I'm not sure how to make a date.
- ◉ **MemNet:** I'm assuming he's married to a gay marriage?
- ◉ **CopyNet:** I'm sorry, I'm not sure what you mean by that.
- ◉ **CCM:** I'm pretty sure he's going to be engaged in the next 6 months.



# Summary

---

- ◎ **Emotion, personality, and knowledge** is important to Chatbots
- ◎ Open-domain chatting machine is one of the most challenging AI tasks
  - ◆ Requires the ability of understanding **semantics, knowledge, and situational context**
  - ◆ Ability of making **reasoning**
  - ◆ Exhibits **consistent personality**
- ◎ Still a long way to go: existing generation models are still far from the expectation of real-world applications



# Future Research Problems

---

- ◎ **Multi-modality** emotion perception and expression (**voice, vision, text**)
  
- ◎ **Personality, identity, style** → “**human-like robot**”
  - ◆ **Introvert or extrovert**
  - ◆ **Personalized (style, or profile)**
  
- ◎ **Learning to learn (lifelong learning)**
  - ◆ **Grow up from interactions with human partners and environment**



# Thanks for Your Attention

---

- ◎ Minlie Huang, Tsinghua University
- ◎ [aihuang@tsinghua.edu.cn](mailto:aihuang@tsinghua.edu.cn)
- ◎ <http://coai.cs.tsinghua.edu.cn/hml>

