

# Towards more Intelligent Dialog Systems

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>

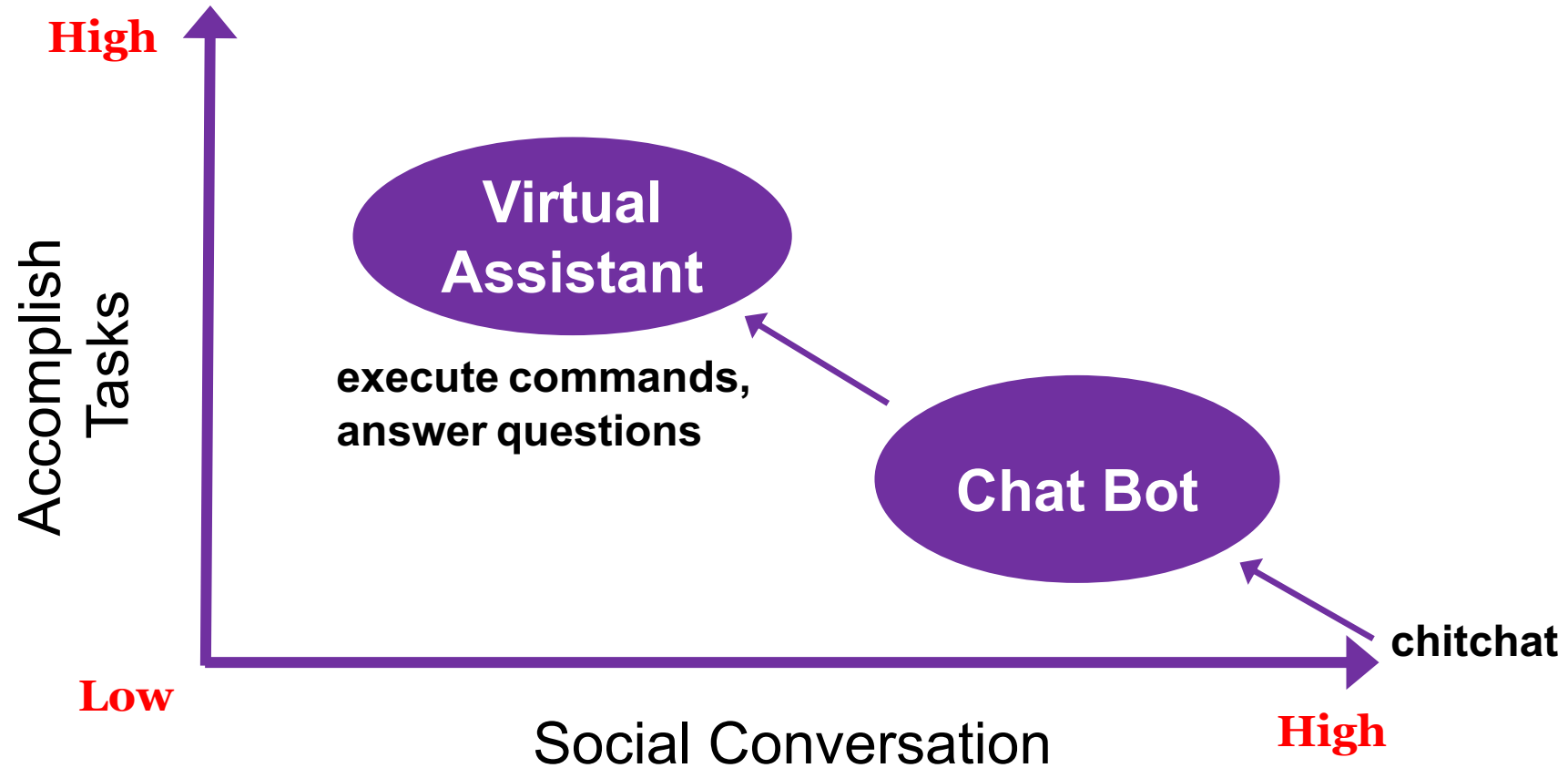
# Conversational AI

- **Conversational AI** refers to the use of **messaging apps, speech-based assistants** and **chatbots** to automate communication and create personalized customer experiences **at scale**.
- Conversational AI signals a huge advancement in the way we **interact with** computers.  
**Keyboard → Mouse → Touchscreen → Voice(Conversation)**
- Key technologies for speech, language, multi-modalities

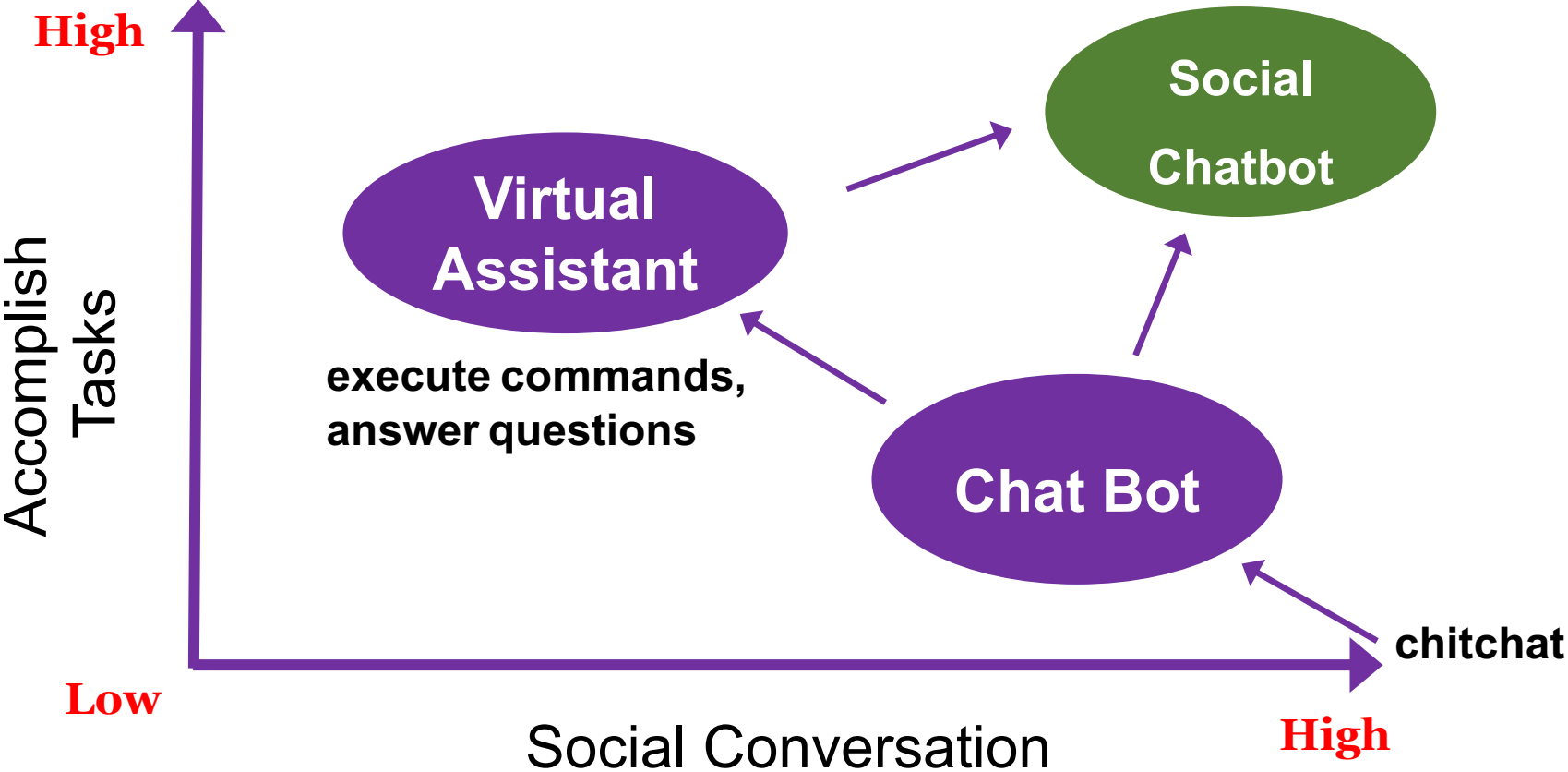
# Dialog System Types

- Task-oriented dialog systems (**task completion**)
  - Hotel reservation, flight booking
- Open-domain dialog system (**chit-chatting**)
- Social chatbot (**Mixture of many skills** including task completion, chit-chatting, etc.)
  - However, far beyond merely *a mixture*

# Types of Dialog Systems

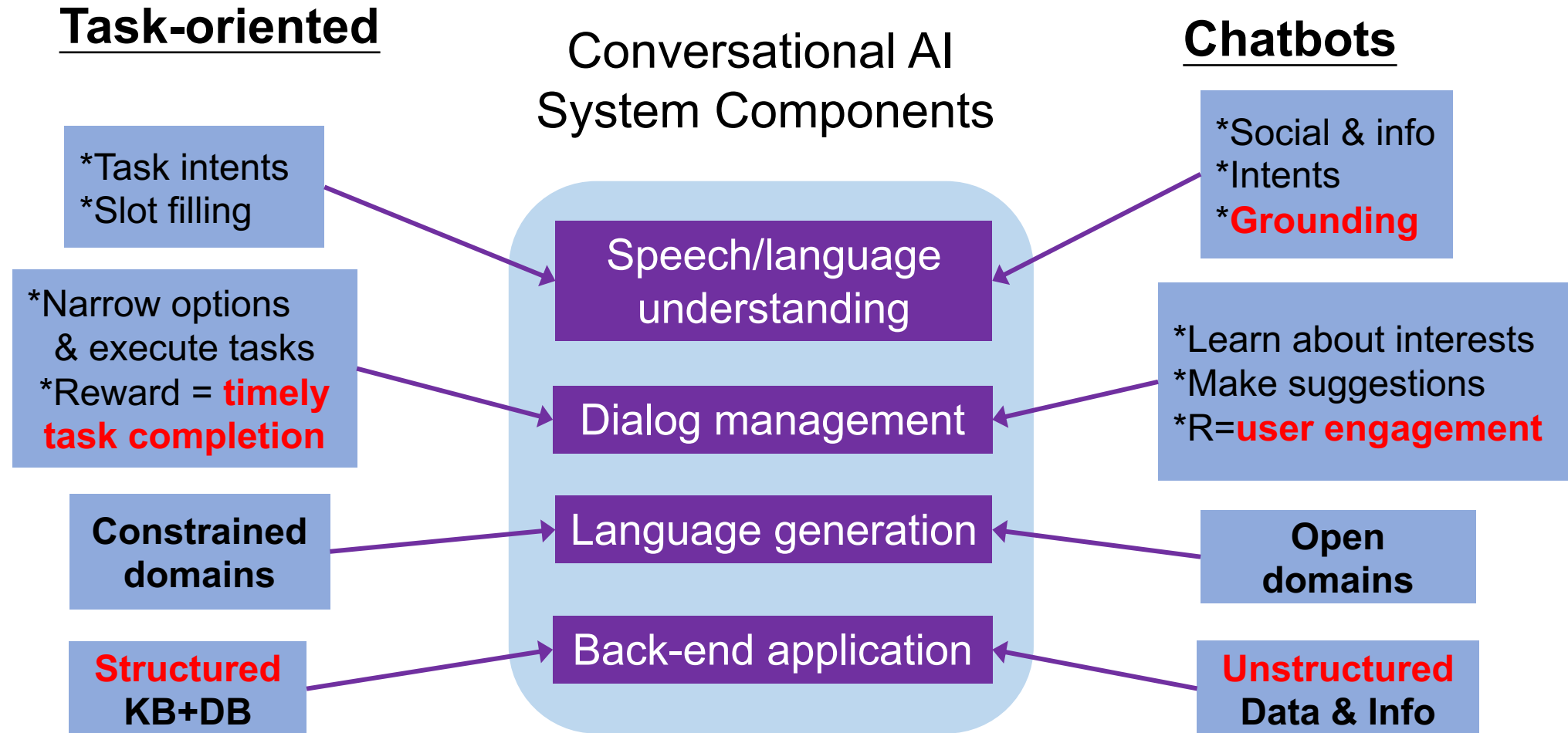


# Types of Dialog Systems



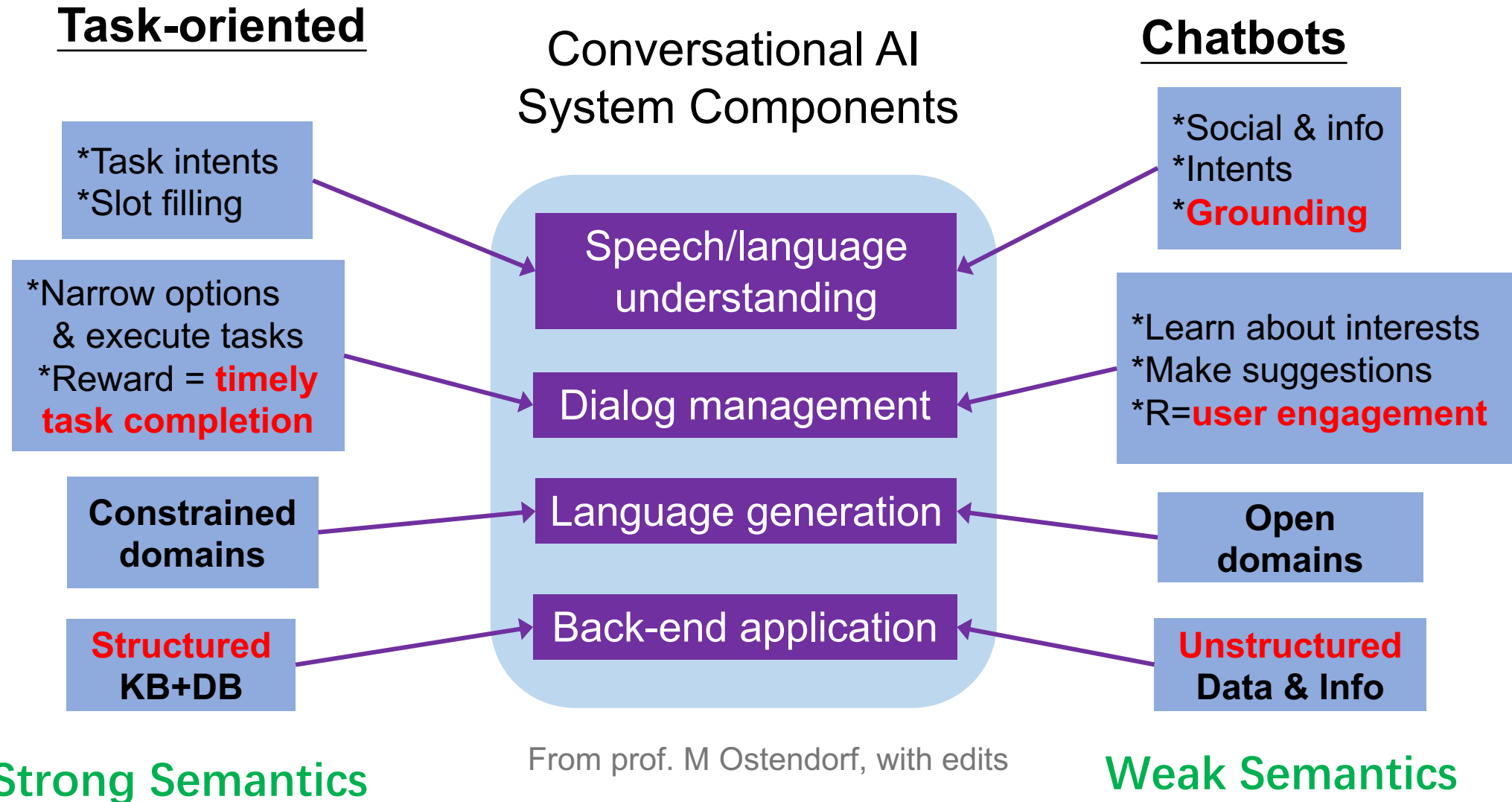
From prof. M Ostendorf, with edits

# Issues in Different Dialog Systems



From prof. M Ostendorf, with edits

# Issues in Different Dialog Systems



# Strong Semantics

- Entity Recognition
- Intent Classification
- Semantic Parsing
- Slot Filling for state tracking
- Template-based language generation
- Symbolic-based methods



# Strong Semantics – IBM Watson 2010



- 2010 IBM Watson: Human-level knowledge QA  
Defeat human champions in TV show Jeopardy!
- ✓ Answering factoid questions: which states does Mississippi River flow through?
  - ✓ Finding answers from unstructured data
  - ✓ Deep semantic parsing techniques in NLP

Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A. A., ... & Schlaefer, N. (2010). Building Watson: An overview of the DeepQA project. *AI magazine*, 31(3), 59-79.

# GUS 1977-Genial Understander System

**GUS:** Hello. My name is GUS. I can help you plan a simple trip by air.  
Where do you want to go?

(1) **Client:** I want to go to San Diego on May 28

**GUS:** What time do you want to leave?

(2) **Client:** I must be in San Diego before 10 am

**GUS:** Would you like P.S.A. flight 102 that arrives at 9.15 am?

(3) **Client:** What is the next flight?

**GUS:** Air California flight 310 that leaves at 8.30 am and arrives at 10.00 am

(4) **Client:** I'll take the first one

**GUS:** In what name should I make a reservation, please?

(5) **Client:** Dan Bobrow

**GUS:** I have confirmed the following flight:

P.S.A. flight 102 on Wednesday May 28 from San Jose to San Diego  
that leaves at 7.30 am and arrives at 9.15 am.

What date do you want to return on?

# GUS 1977-Genial Understander System

**GUS: Hello. My name is GUS. I can help you plan a simple trip by air.**

**CLIENT = {ISA PERSON} in {ISA DIALOG}**

**TODAY = (MAY 15) in {ISA DIALOG}**

**TOPIC = (ROUNDTrip TO ?) in {ISA DIALOG}**

**HOME-PORT = PALO-ALTO in (ROUNDTrip TO ?)**

**FROM-PLACE = PALO-ALTO in (TRIP TO ?)**

**CMD: (GUSQUERY (DIALOG TOPIC TRIP-SPECIFICATION  
OUTWARD-LEG TRIP-LEG TO-PLACE CITY))**

**GUS: Where do you want to go?**

**FIG. 7. The beginning of the transcript for the dialog.**

# GUS 1977-Genial Understander System

[DATE

MONTH

NAME

DAY

(BOUNDED-INTEGER 1 31)

YEAR

INTEGER

WEEKDAY

(MEMBER (SUNDAY MONDAY TUESDAY  
WEDNESDAY THURSDAY FRIDAY SATURDAY))]

a. Prototype for **date**

[ISA DATE

MONTH

MAY

DAY

28]

b. The instance frame for **May 28**

# Weak Semantics

- Data-driven
- End-to-end
- Probabilistic methods without knowledge, rule, or symbolics

# Eliza 1966 – Earliest Chatbot

- Created by MIT professor Joseph Weizenbaum
- Features
  - Hand-crafted scripts
  - keyword spotting
  - template matching

Joseph Weizenbaum. 1966. ELIZA - a computer program for the study of natural language communication between man and machine. Commun. ACM 9, 1 (1966), 36–45.

**Human:** Men are all alike.

**Eliza:** In what way?

**Human:** They're always bugging us about something specific or other.

**Eliza:** Can you think of a specific example?

**Human:** Well, my boyfriend made me come here.

**Eliza:** Your boyfriend made you come here?

**Human:** He says I'm depressed much of the time.

**Eliza:** I am sorry to hear you are depressed.

**Human:** It's true. I'm unhappy.

**Eliza:** Do you think coming here will help you not to be unhappy?

**Human:** I need some help, that much seems certain.

**Eliza:** What would it mean to you if you got some help?

**Human:** Perhaps I could learn to get along with my mother.

**Eliza:** Tell me more about your family.

# Eliza 1966 – Earliest Chatbot

pronoun reversal (key tricks): I → you, me → you

- Human: He says **I am depressed** much of the time.
- Eliza: I am sorry to hear **you are depressed**.
  
- Human: **I need some help**, that much seems certain.
- Eliza: what would it mean to you if **you got some help**?
  - Pattern: I need some X → What would it mean to you if you got some X?

Joseph Weizenbaum. 1966. ELIZA - a computer program for the study of natural language communication between man and machine. Commun. ACM 9, 1 (1966), 36–45.

# Weak Semantics- Open-domain Dialog Generation

User: 我不喜欢吃米饭。

**I do not like to eat rice.**

**Machine: 我也不喜欢吃米饭, 但我喜欢吃米饭。**

**I do not like either, but I like to eat rice.**

User: 你喜欢我吗?

**Do you like me?**

**Machine: 我喜欢你, 但我不喜欢你。**

**I like you, but I do not like you.**

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!**

## 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?

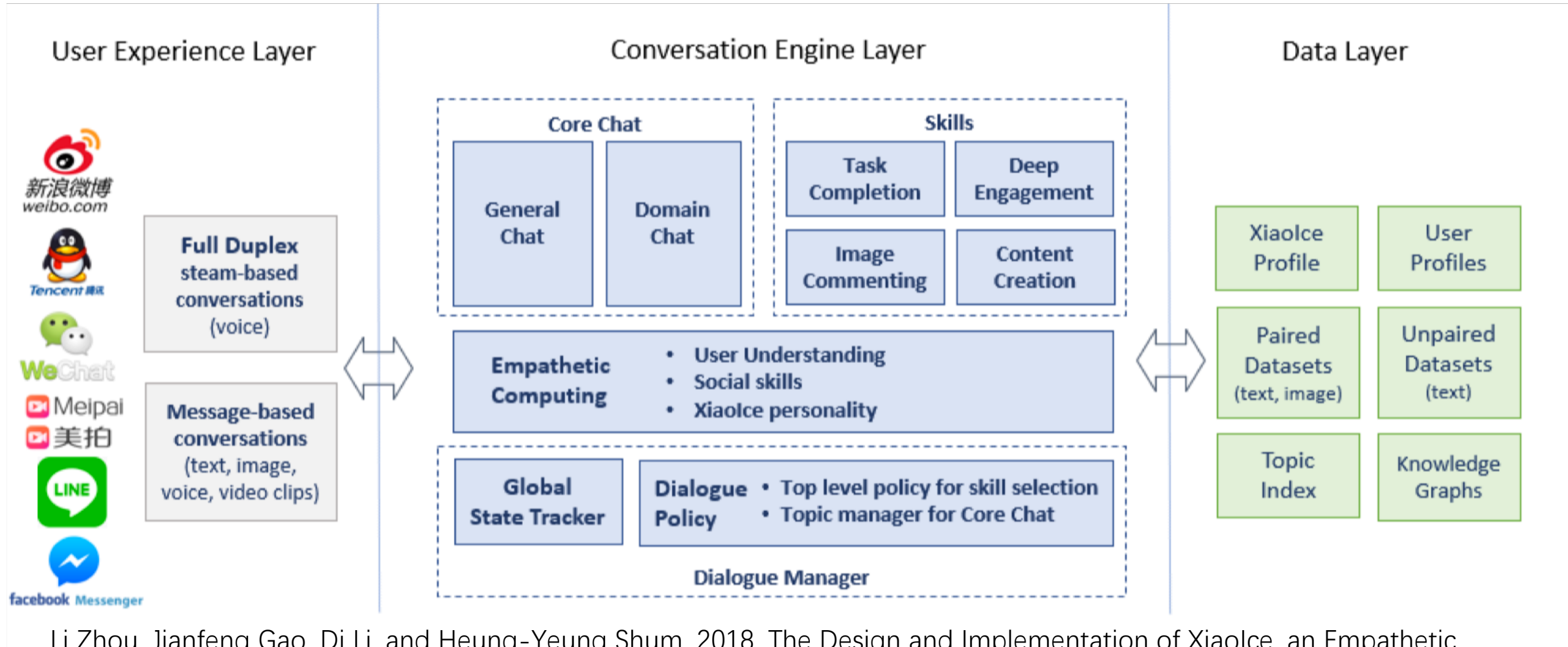
**小冰: 错, 75年阳历一月出生。**

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。**



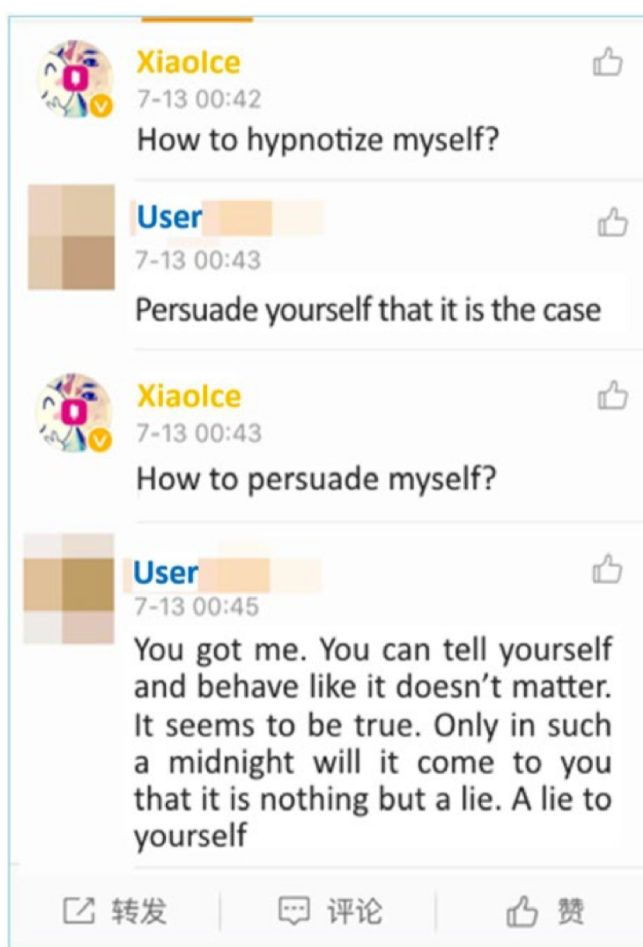
# In Between – Xiaolce



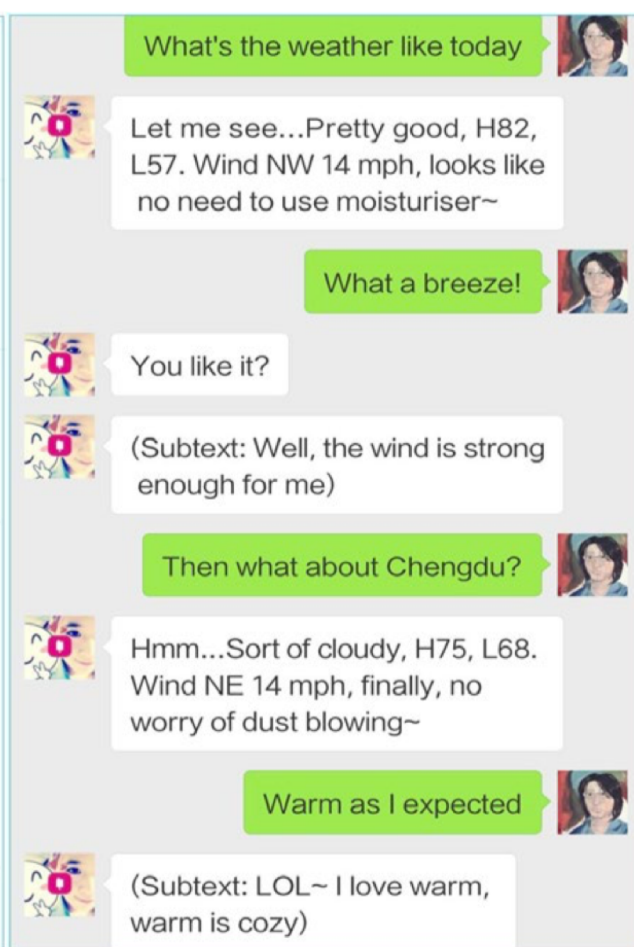
Li Zhou, Jianfeng Gao, Di Li, and Heung-Yeung Shum. 2018. The Design and Implementation of Xiaolce, an Empathetic Social Chatbot. CoRR abs/1812.08989 (2018).

Heung-Yeung Shum, Xiaodong He, and Di Li. 2018. From Eliza to Xiaolce: challenges and opportunities with social chatbots. Frontiers of IT & EE 19, 1 (2018), 10–26.

# In Between – Xiaoice

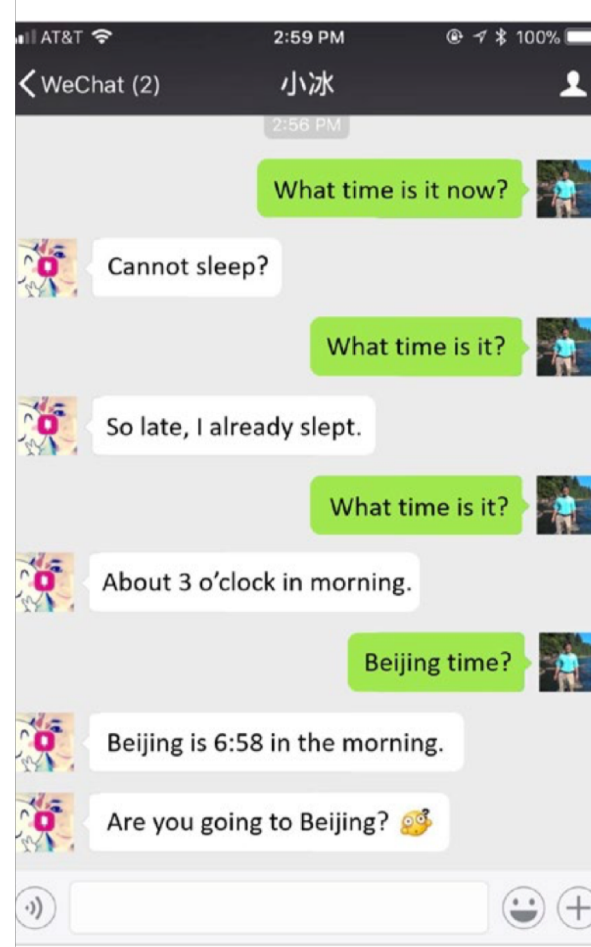


(a)



(b)

Emotional connections and invoking a skill in a casual chat



(a)



(b)

Showing IQ+EQ in a social chatbot

# In Between – Xiaoice

## 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？

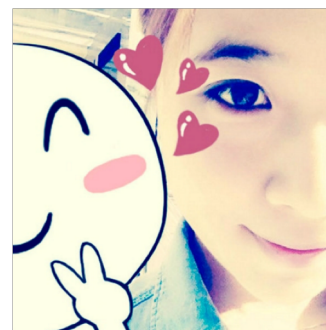
**So soon to reveal my age?**

**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?**



## 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。

## 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

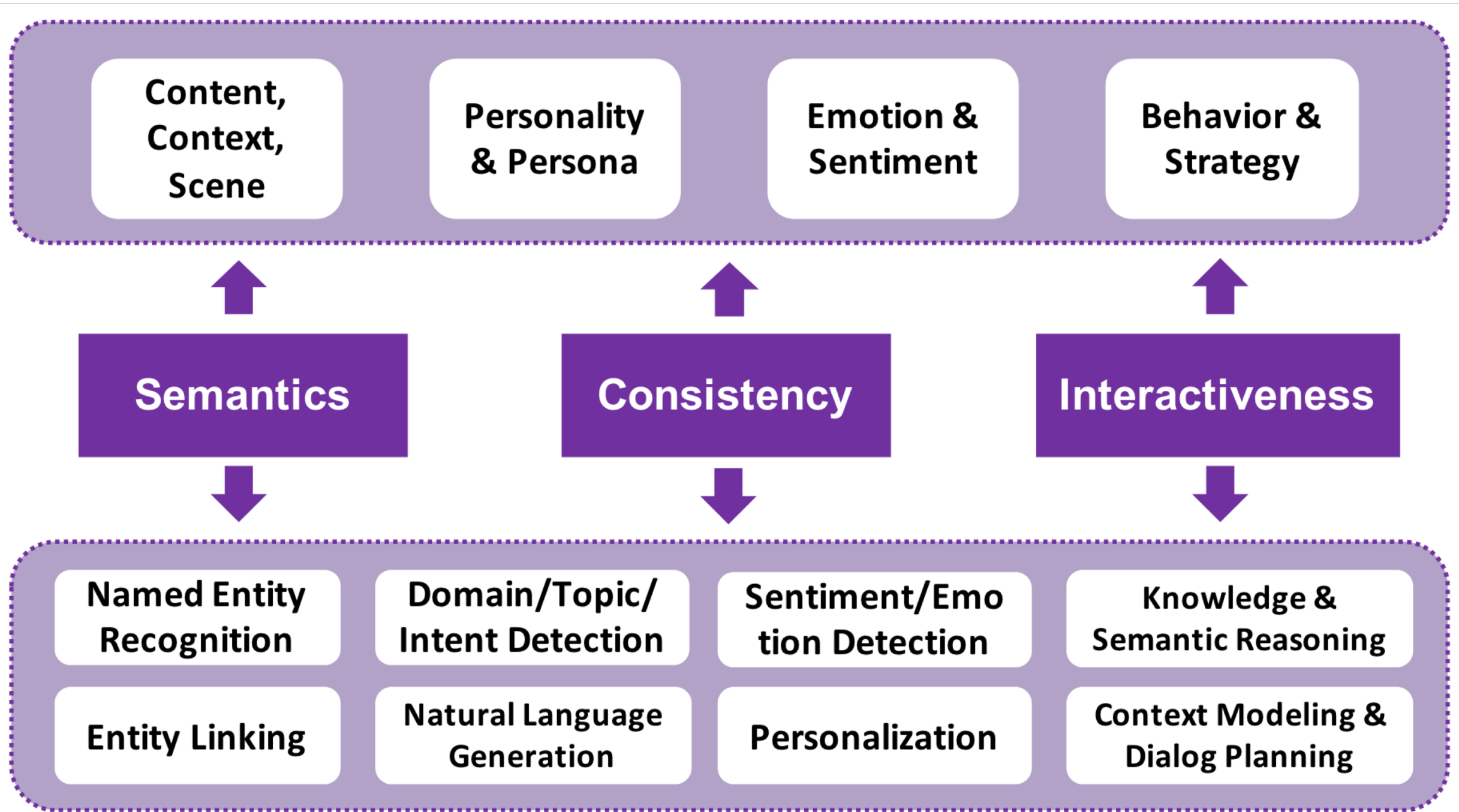
# The Next: Social Chatbots

- A **chatbot** is a [computer program](#) which conducts a [conversation](#) via auditory or textual methods. (Wikipedia)
- **Social chatbot** is a **chatbot** that can satisfy not only **information need** but also **social need** of a user.
  - **Information need (IQ)**: QA, recommendation, task completion
  - **Social need (EQ)**: emotion, accompany, counseling
- Social chatbot is a mixture of various **skills**, but far **beyond** a mixture

# Challenges in Social Chatbots

- **Semantics:** dialog is all about semantic understanding
- **Consistency:** within multi-turn contexts, personality, behaviors
- **Interactiveness:** topic, emotion, sentiment, behavior, strategy, etc.

Minlie Huang, Xiaoyan Zhu, and Jianfeng Gao. "**Challenges in Building Intelligent Open-domain Dialog Systems.**" *arXiv preprint arXiv:1905.05709* (2019).



# The Next: Social Chatbots

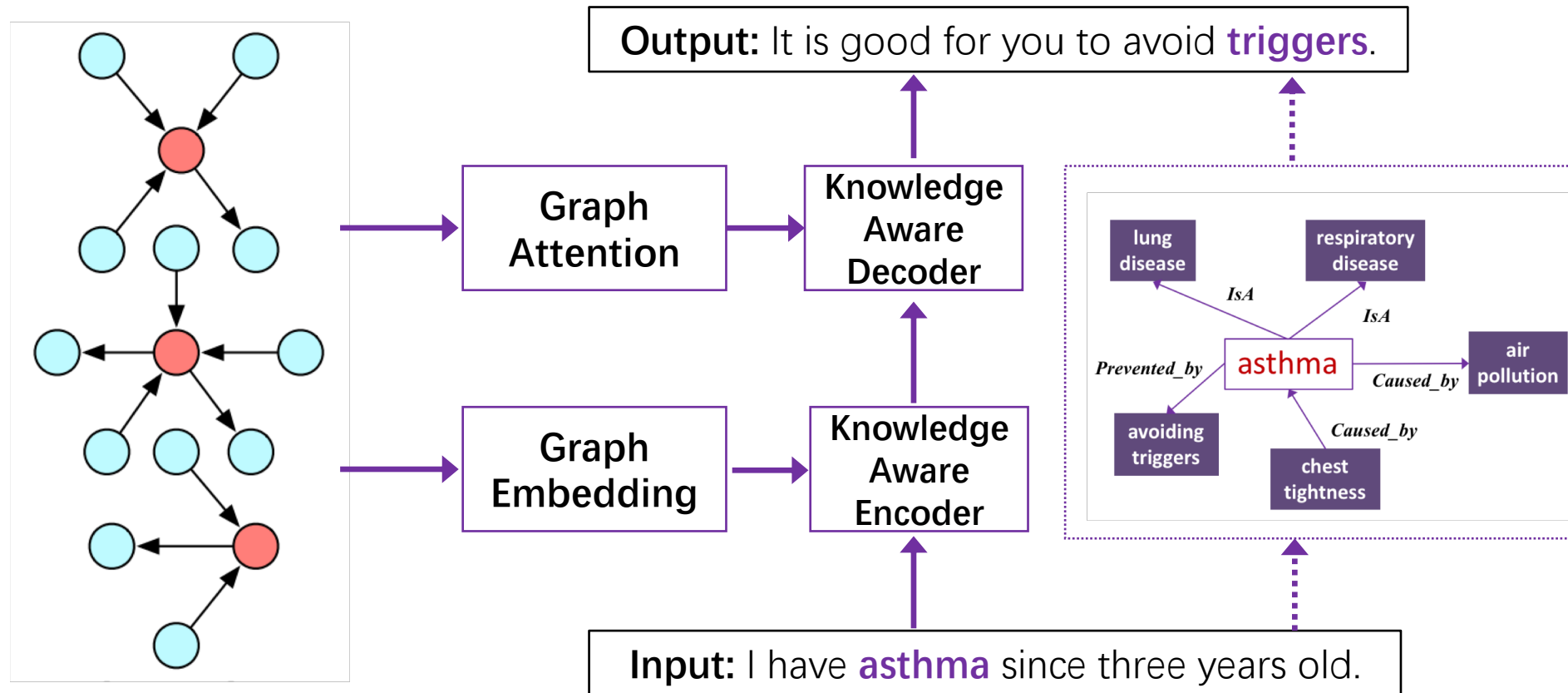
- **Semantics:** Knowledge-aware, Knowledge-grounded
- **Consistency:** Persona/Personality coherent
- **Interactiveness:** Emotion-aware, Proactive behaviors, Topic planning

# Knowledge-aware

- **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*?

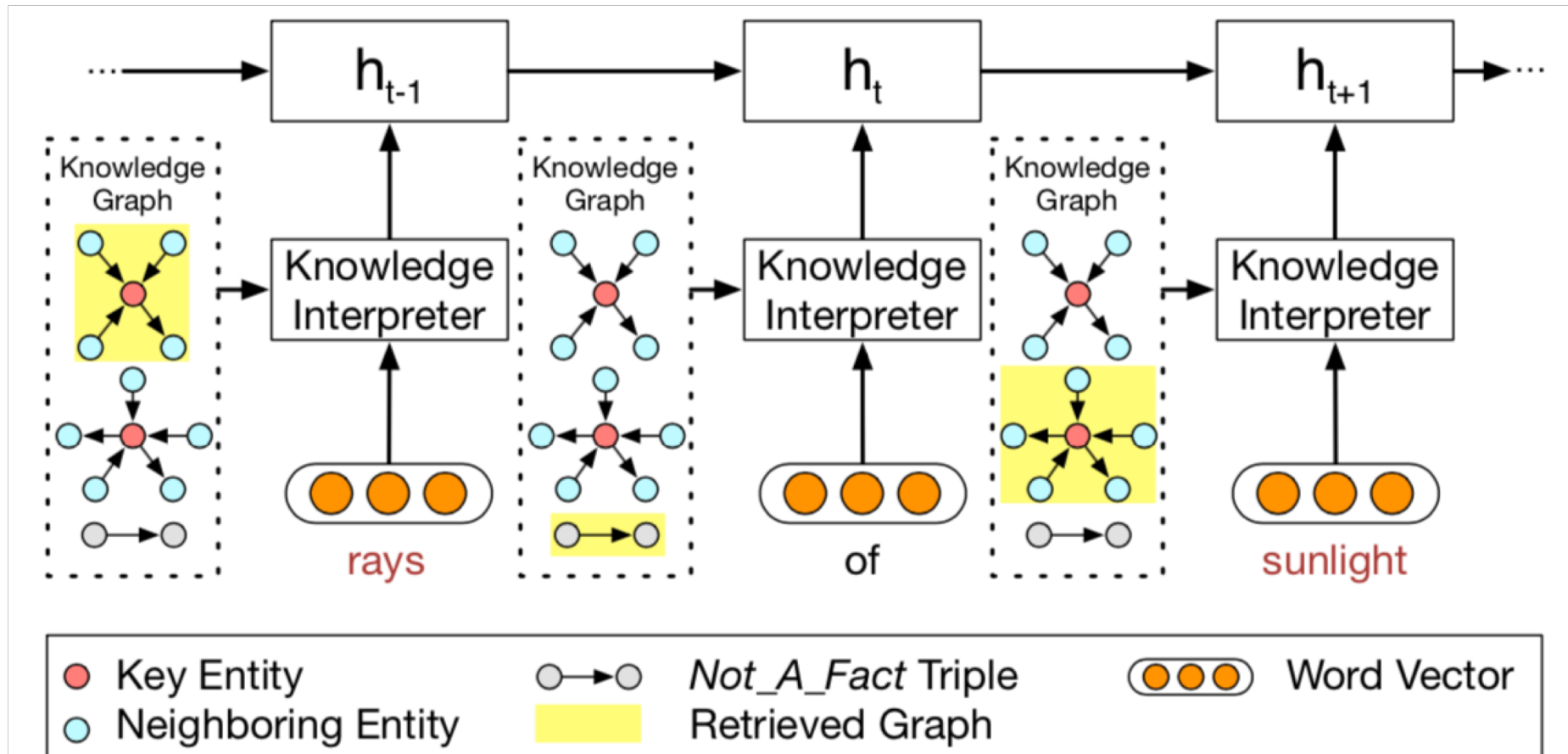


# Knowledge-aware

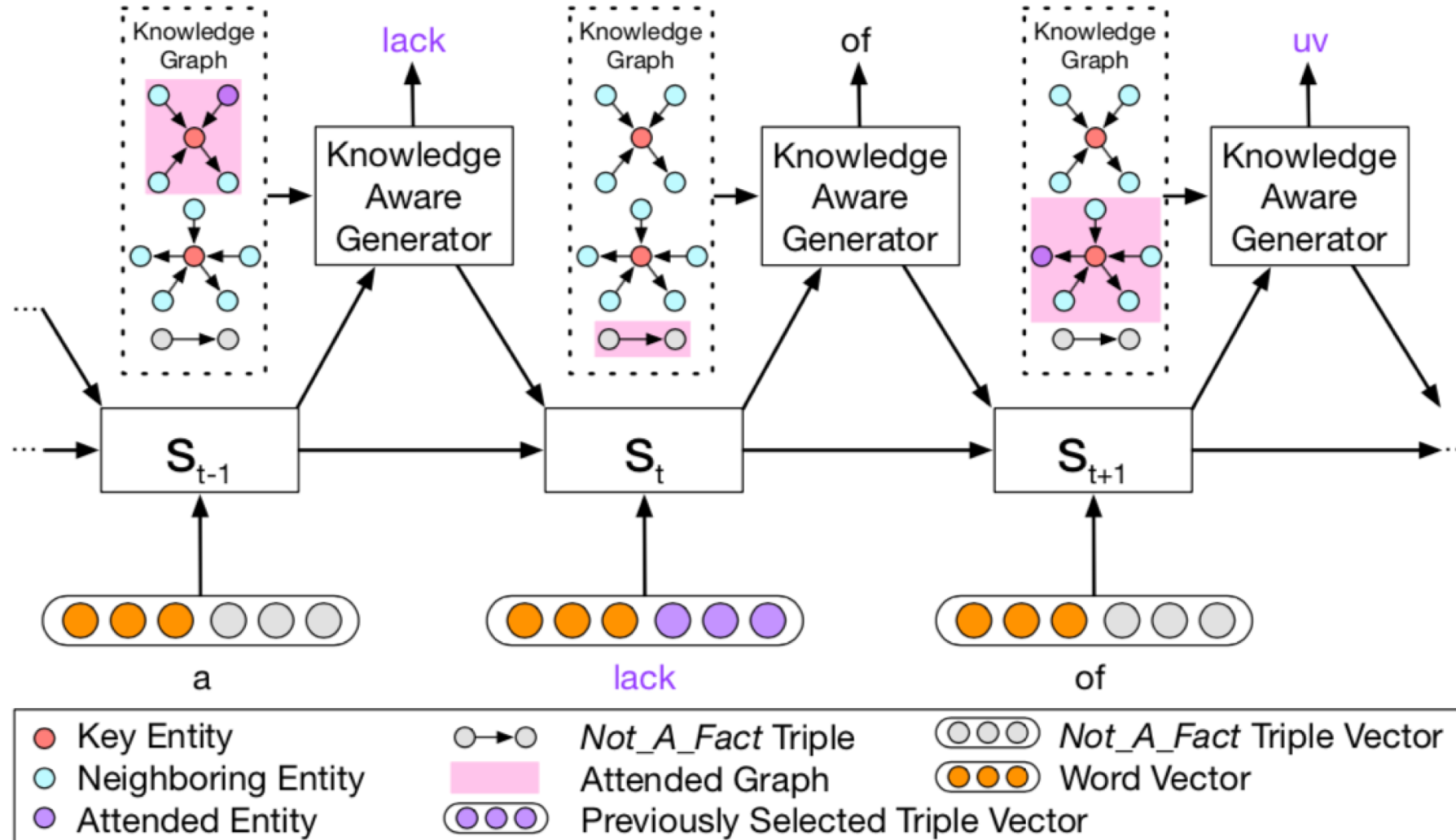


# Knowledge-aware

**Static graph attention:** encoding a knowledge graph to a vector, feeding graph vector and word vector to the encoder



# Knowledge-aware



# Knowledge Grounding

- *Kind of* understanding
- Contentful generation: aligning dialog generation to knowledge

<b>Name</b>	The Shape of Water
<b>Year</b>	2017
<b>Director</b>	Guillermo del Toro
<b>Genre</b>	Fantasy, Drama
<b>Cast</b>	Sally Hawkins as Elisa Esposito, a mute cleaner who works at a secret government laboratory. Michael Shannon as Colonel Richard Strickland, a corrupt military official, Richard Jenkins as Giles, Elisa's closeted neighbor and close friend who is a struggling advertising illustrator. Octavia Spencer as Zelda Delilah Fuller, Elisa's co-worker and friend who serves as her interpreter., Michael Stuhlbarg as Dimitri Mosenkov, a Soviet spy working as a scientist studying the creature, under the alias Dr. Robert Hoffstetler.

User 2:	I thought The Shape of Water was one of Del Toro's best works. What about you?
User 1:	Did you like the movie?
User 1:	Yes, his style really extended the story.
User 2:	I agree. He has a way with fantasy elements that really helped this story be truly beautiful.

Table 2: An example conversation for scenario 2. Both User 1 and User 2 have access to the Wiki document. The full dialogue is attached in the Appendix.

# Knowledge Grounding

- *Kind of* understanding
- Contentful generation: aligning dialog generation to knowledge

<b>Topic:</b>	Lifeguard
Apprentice:	So I am a lifeguard. Know anything about saving lives in water?
Wizard:	I'm impressed! It's a big responsibility to supervise other people's safety in the water! Tell me more.
Apprentice:	Well, I help make sure people do not drown or get injured while in or near the water!
<b>Knowledge:</b>	<p>A lifeguard is a rescuer who supervises the safety and rescue of swimmers, surfers, ... Lifeguards are strong swimmers and trained in CPR/AED first aid, certified in water ...</p> <p>...</p> <p>In some areas, the lifeguard service also carries out mountain rescues, or may function as the primary EMS provider.</p>
Wizard:	I've heard that in some places, lifeguards also help with other sorts of emergencies, like mountain rescues! Is that part of your job too?
Apprentice:	I have! I feel like you know much about this! What brings you to know so much?
Wizard:	Oh, that's about the extent of my knowledge. I've just been around beaches and I've always admired lifeguards. I'm not a super strong swimmer myself.

Emily Dinan et al. 2018. Wizard of Wikipedia: Knowledge-Powered Conversational agents. CoRR abs/1811.01241 (2018).

# Personality of Dialog Systems

- 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

# Personality-aware

- Personality is important for game, custom service, virtual agent, etc.
- A coherent personality makes a system more **trustable**
- Personality is important for making effective social interactions
- Require cross-discipline research from **psychology, cognitive science**

## From a neural model

### 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.

## From Xiaoice

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？

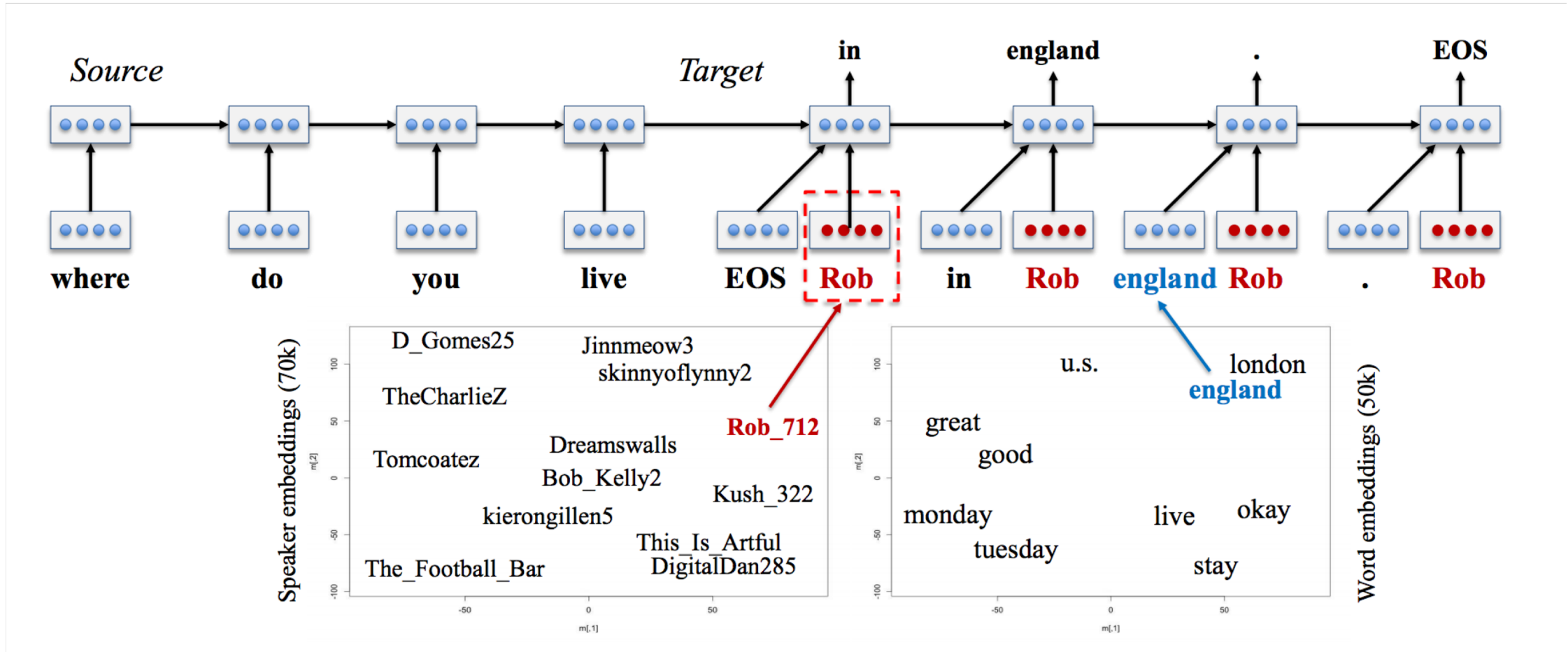
小冰: 错，75年阳历一月出生。

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。



# Persona-aware



Li et al. "A persona-based neural conversation model." *ACL 2016*.



# Personality-aware

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

## Generic Dialogue Data for Training

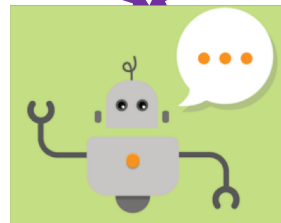
UserA: how old are you?  
UserB: I am **six**.

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**.

# Intelligent Socialbot = IQ+EQ

- **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.

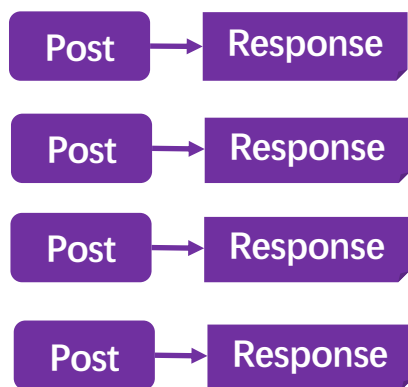
ECM (*Disgust*): Sometimes life just sucks.

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

# Emotion-aware

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

Social Interaction Data



Emotion Classifier

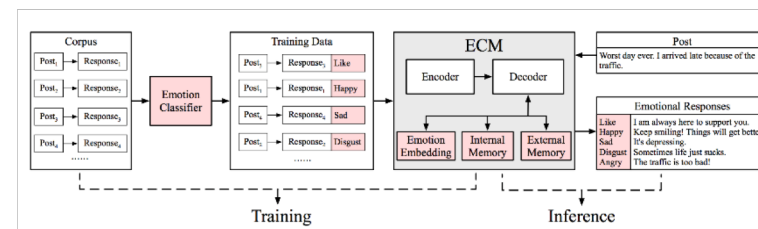


Emotion Tagged data



Emotional Chatting Machine

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？

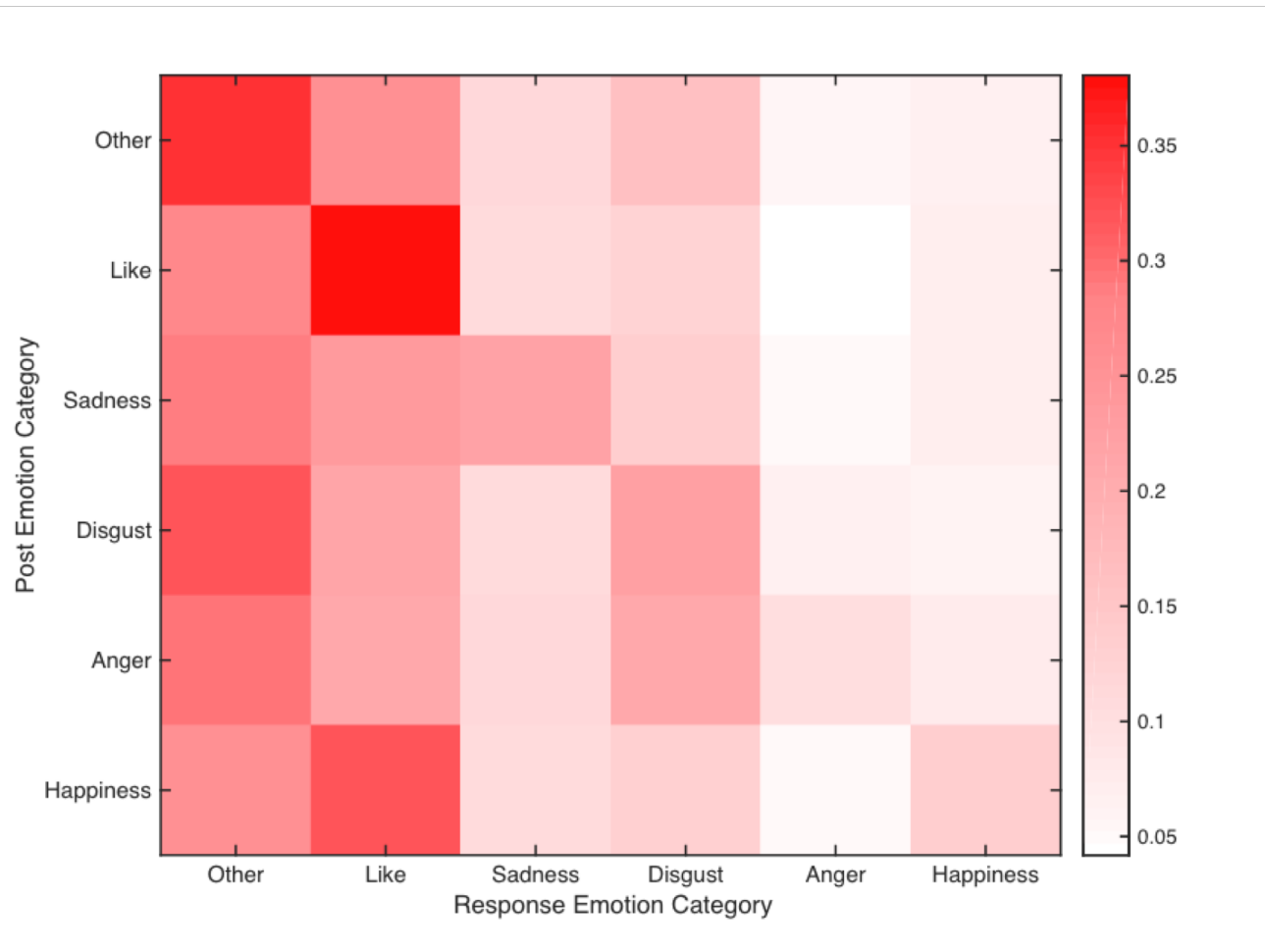


Our work was reported by [MIT Technology Review](#), [the Guardian](#), [Cankao News](#), [Xinhua News Agency](#) etc.

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**.

# Emotion Interaction Patterns



Like → Like (**empathy**)

Sadness → Sadness (**empathy**)

**Sadness → Like (comfort)**

Disgust → Disgust (**empathy**)

**Disgust → Like (comfort)**

Anger → Disgust

Happiness → Like

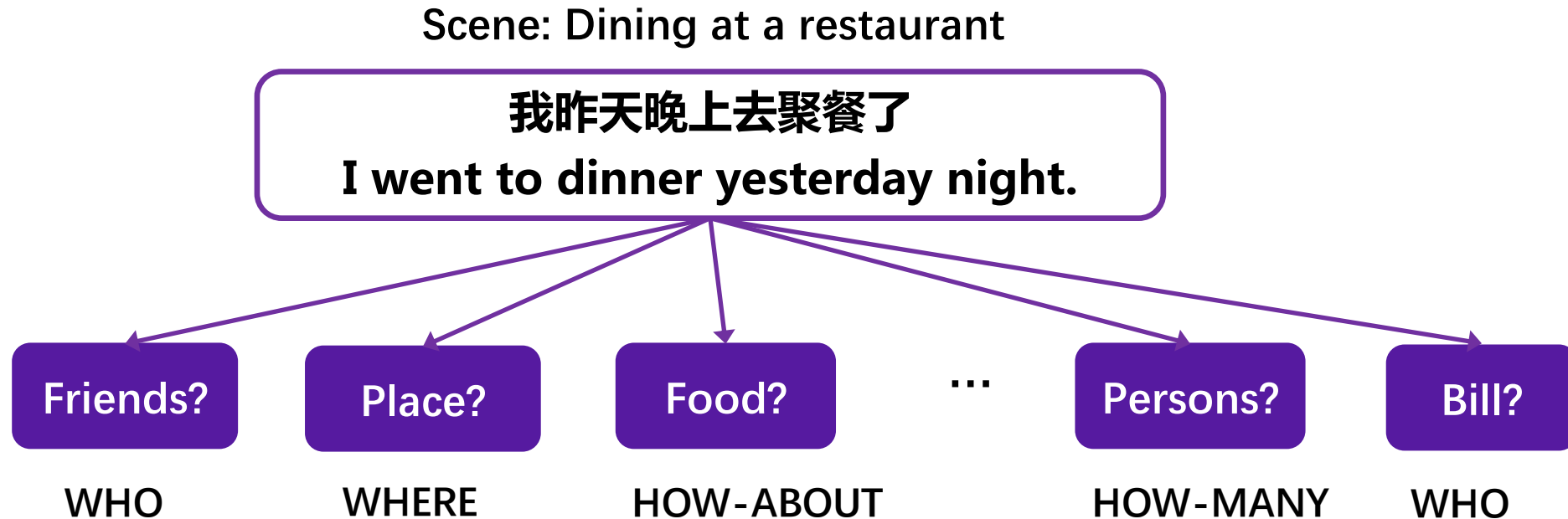
# Proactive Behaviors

我昨天晚上去聚餐了

I went to dinner yesterday night.

# Proactive Behaviors

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



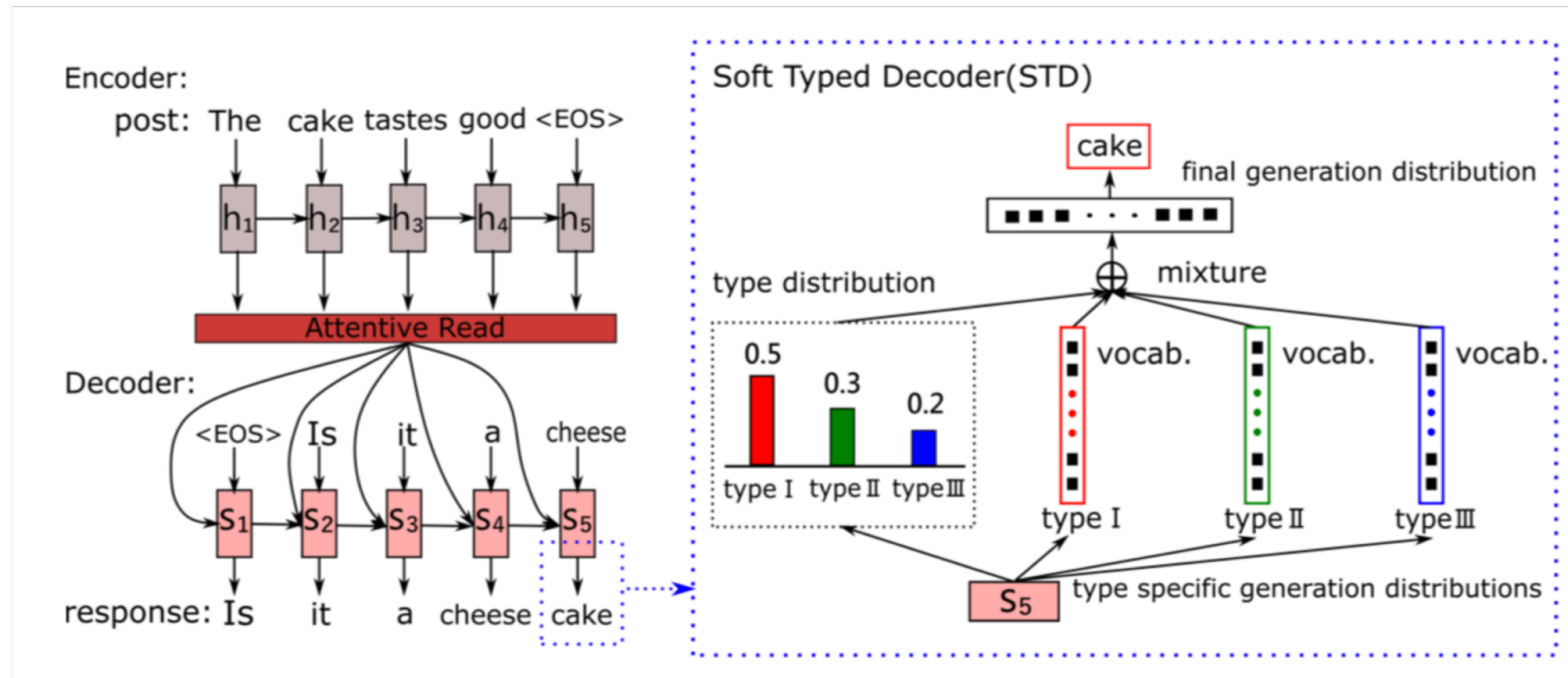
Yansen Wang, Chenyi Liu, Minlie Huang, Liqiang Nie.  
Learning to ask questions in open-domain conversation systems. **ACL 2018**.

# Proactive Behavior by Asking Questions

- 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

# Proactive Behaviors (STD)

- **Typed decoders:** soft typed decoder



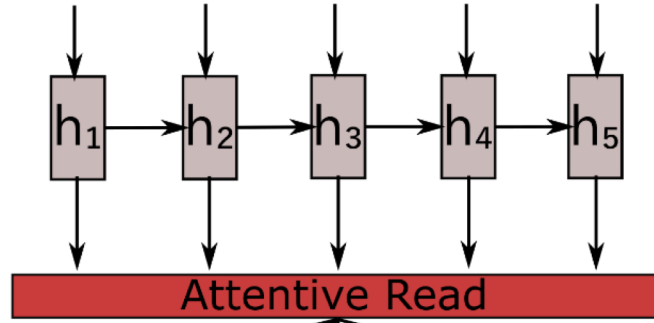
Yansen Wang, Chenyi Liu, Minlie Huang, Liqiang Nie.  
Learning to ask questions in open-domain conversation systems. **ACL 2018**.



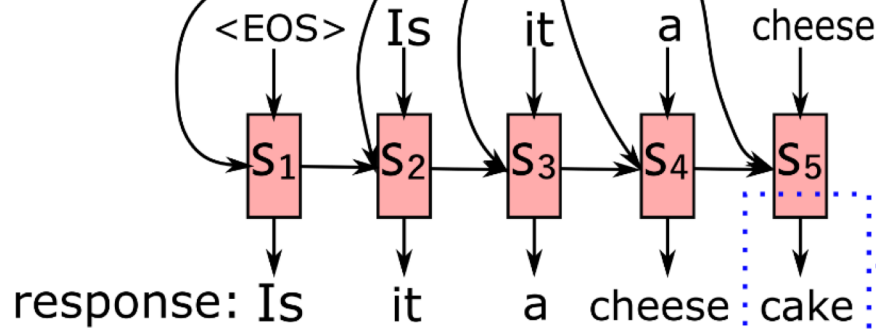
# Proactive Behaviors (HTD)

Encoder:

post: The cake tastes good <EOS>

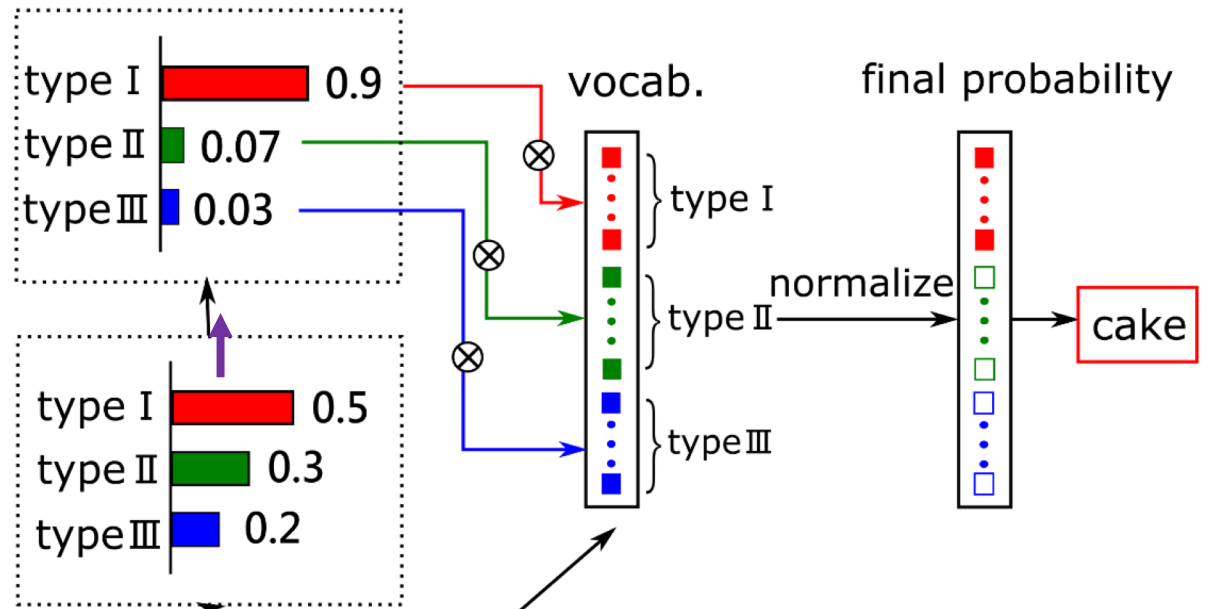


Decoder:



Hard Typed Decoder(HTD)

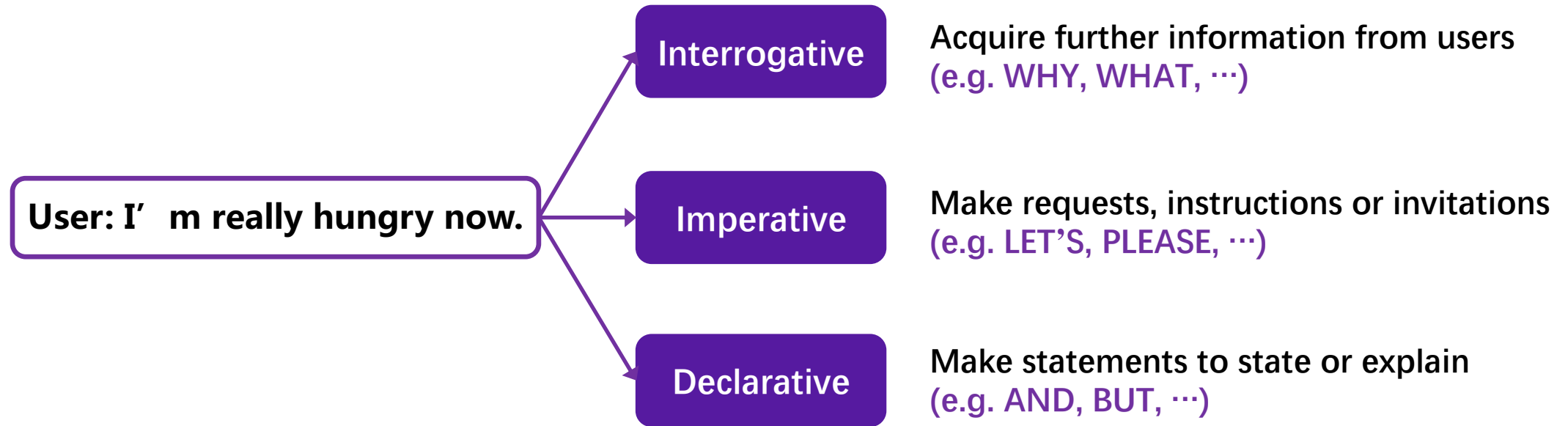
Gumbel-softmax



Decoding state

# Controlling Sentence Function

- Sentence function indicates different conversational purposes.

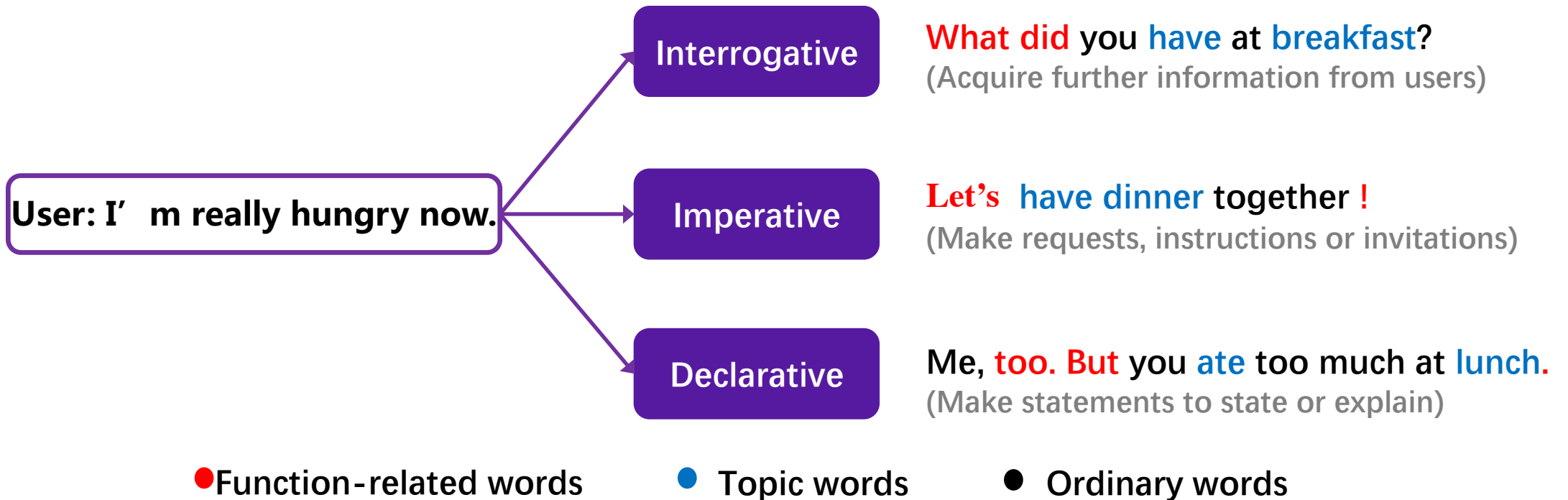


Pei Ke, Jian Guan, Minlie Huang, Xiaoyan Zhu.

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

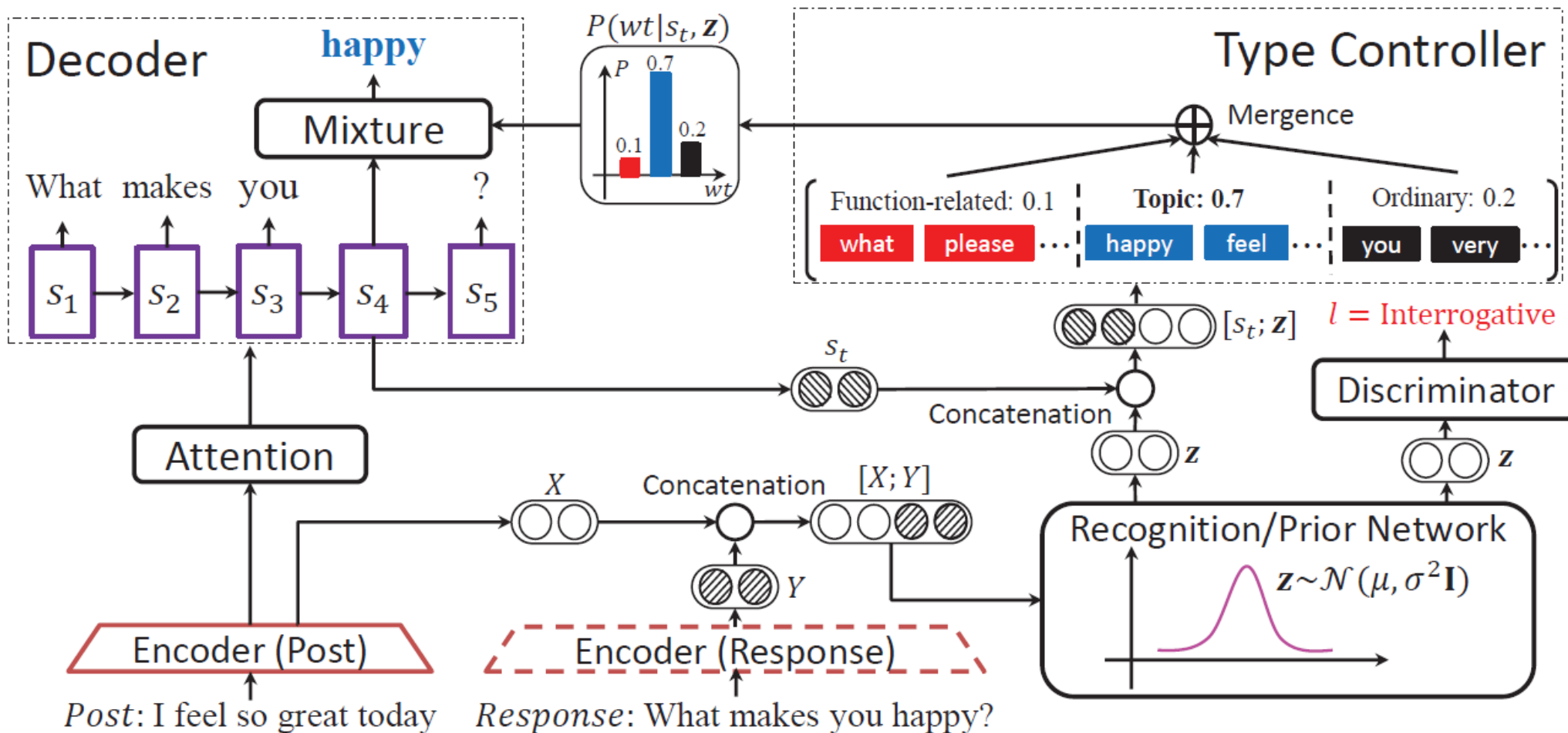
# Controlling Sentence Function

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



# Controlling Sentence Function

- Conditional Variational Autoencoder (CVAE) Framework



# Topic Planning in Conversation

## A dialog session

person A:	There is a <b>heavy rain</b> today.
person B:	The <b>umbrella</b> is totally <b>useless</b> .
person A:	The <b>rain</b> is really <b>heavy</b> . ( <b>topic penetration</b> ) I got <b>wet</b> in the afternoon and <b>caught a cold</b> at night. ( <b>topic extension</b> )
person B:	You should take some <b>hot tea</b> and get <b>a good sleep</b> .( <b>topic extension</b> )

## Motivation:

- Leverage the **relevant** topic information.
- Generate informative responses that are not only relevant but also capable of **deepening and widening** the chatting topic.
- Avoid **dull** responses

# Topic Planning in Conversation

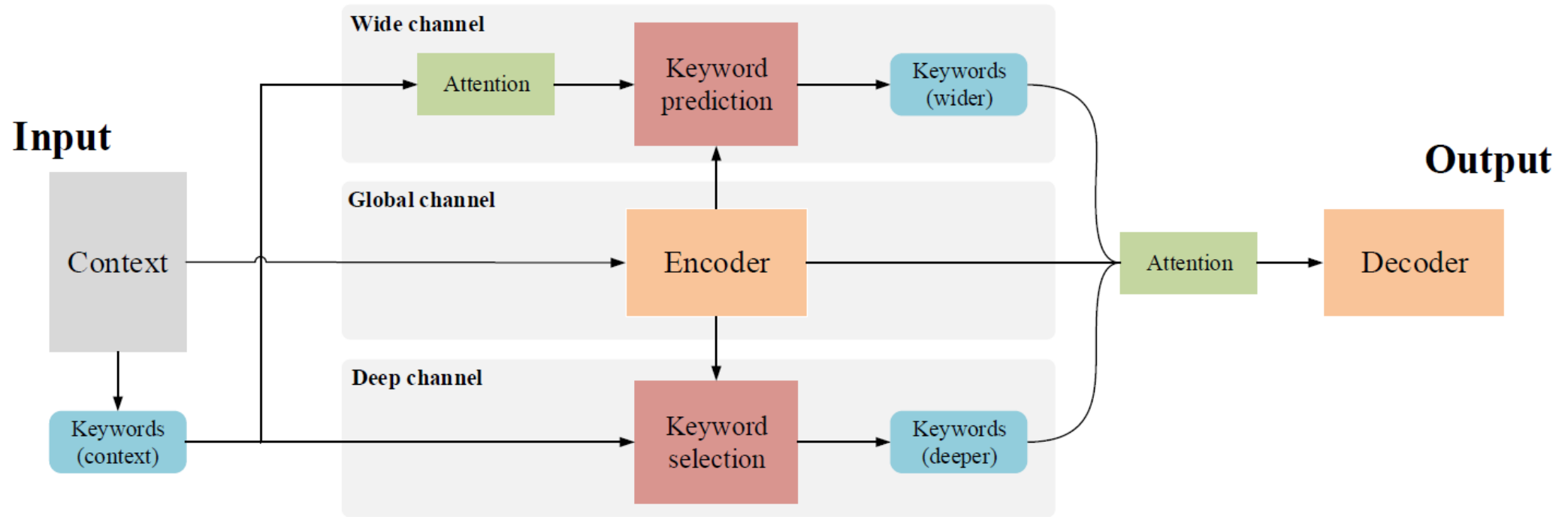
## A dialog session

person A:	There is a <b>heavy rain</b> today.
person B:	The <b>umbrella</b> is totally <b>useless</b> .
person A:	The <b>rain</b> is really <b>heavy</b> . ( <b>topic penetration</b> ) I got <b>wet</b> in the afternoon and <b>caught a cold</b> at night. ( <b>topic extension</b> )
person B:	You should take some <b>hot tea</b> and get <b>a good sleep</b> . ( <b>topic extension</b> )

## Definition:

- **Deepening the chatting topic:** *continue the historical topic.*  
eg. heavy rain, umbrella, wet.
- **Widening the chatting topic:** *transfer the topic to related ones.* eg. caught a cold, hot tea, a good sleep.

# Topic Planning in Conversation



## Context:

A: There is a **heavy rain** today.

B: The **umbrella** is totally **useless**.

## Keywords(context):

*heavy, rain, umbrella, useless*

## Keywords(deeper):

*heavy, rain*

## Keywords(wider):

*wet, cold, night*

## Response:

A: The **rain** is really **heavy**. (*topic penetration*)

I got **wet** in the afternoon and caught a **cold** in the **night**. (*topic extension*)

# Summary

- Strong & weak semantics in dialog systems
- Key issues and challenges in social chatbots: semantics, consistency, interactiveness
- Research attempts from knowledge, personality, emotion, behavior, topic



# Future Trends

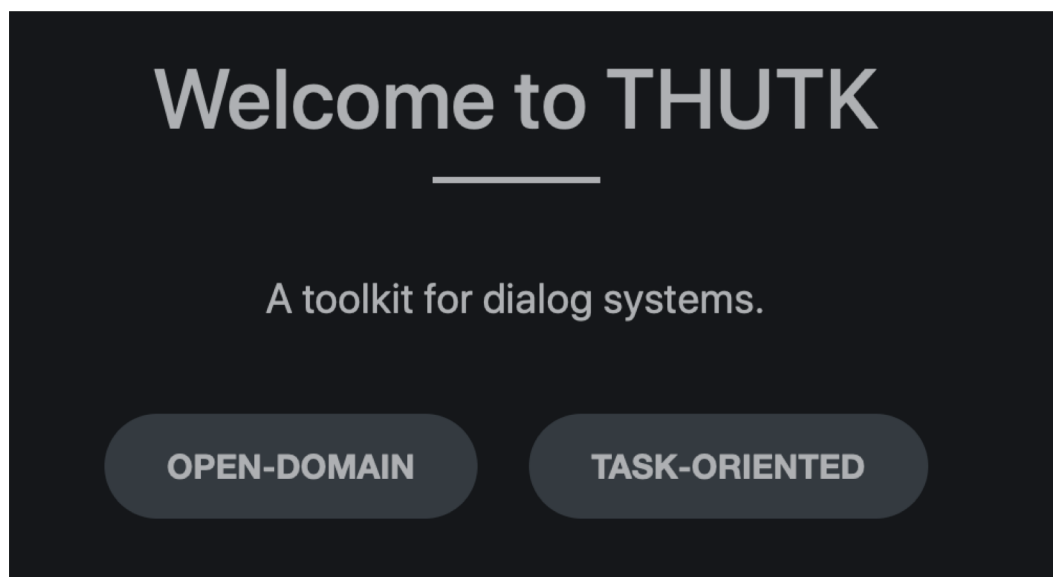
- Knowledge-grounded Dialog Models
- Empathetic Computing (with multi-modalities)
- Personality of a Social Chatbot (with psychologies)
- Controllability of Dialog Generation
- Dialog Evaluation
- **New Technologies**
  - **Data-driven + knowledge-grounded**
  - **Neural computing + symbolic systems (symbols, rules, entities, etc.)**

# My Recent Papers on Dialogue System

- Perceiving and Expressing Emotions (**AAAI 2018**)
- Proactive Behavior by Asking Good Questions (**ACL 2018**)
- Controlling Sentence Function (**ACL 2018**)
- **Topic** Change in Multi-turn Dialog Systems (**SIGIR 2018**)
- **Explicit Personality** Assignment (**IJCAI-ECAI 2018**)
- Better Understanding and Generation Using Commonsense Knowledge (**IJCAI-ECAI 2018 distinguished paper**)
- **Discourse parsing** in multi-party dialogues (**AAAI 2019**)
- **Memory** augmented dialog management (**ACM TOIS 2019**)
- **Multimodality** neural belief tracker (**WWW 2019, SIGIR 2019**)
- **Low-resource** language generation (**IJCAI 2019**)
- Dialog toolkits **ConvLab** (**ACL 2019, best demo candidate**)
- Survey paper “Challenges in Building Intelligent Open-domain Dialog Systems”  
<https://arxiv.org/abs/1905.05709>

# Thanks for Your Attention

- <http://coai.cs.tsinghua.edu.cn/thutk/> (CoTK & TaTK)



[aihuang@tsinghua.edu.cn](mailto:aihuang@tsinghua.edu.cn)  
<http://coai.cs.tsinghua.edu.cn/hml>

