



前沿技术综述

CCL 2018

# 人机对话技术前沿及动态

张伟男

哈尔滨工业大学

社会计算与信息检索研究中心

2018.10.21



# 人机对话系统四个主要功能

	聊天 Chit-Chat	知识问答 Knowledge	任务执行 Task	推荐 Recommendation
目的	闲聊	知识获取	完成任务或动作	信息推荐
领域	开放域	开放域	特定域（垂类）	特定域
话轮数 评价	话轮越多越好	越少越好	越少越好	越少越好
应用	娱乐、情感陪护、 营销沟通	客服、教育	虚拟个人助理	个性化推荐
典型 系统	小冰	Watson、 Wolfram Alpha	Siri、Cortana、 Allo、度秘、灵犀	Quartz/ 今日头条

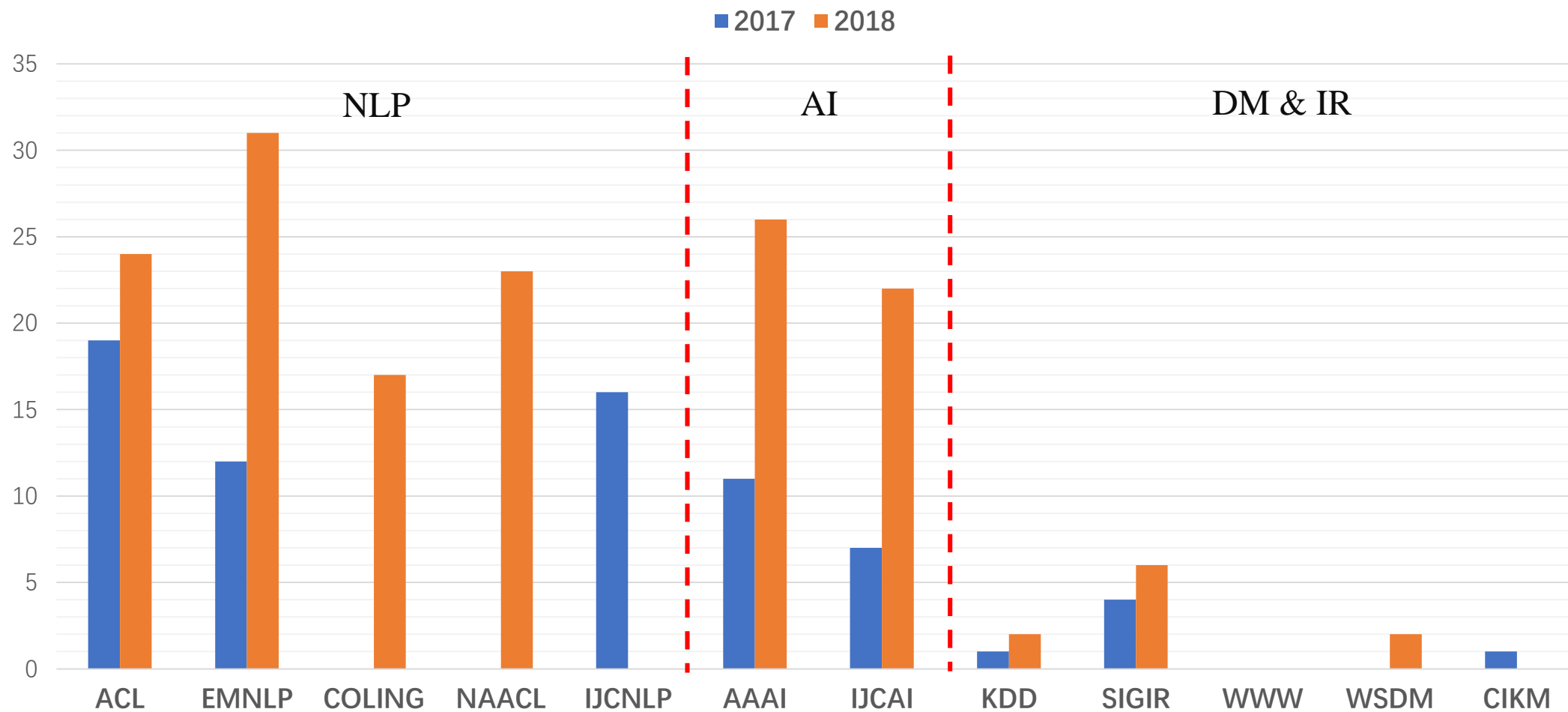
# 大纲

- 热点及趋势分析
- 技术及任务介绍
- 数据及评测汇总

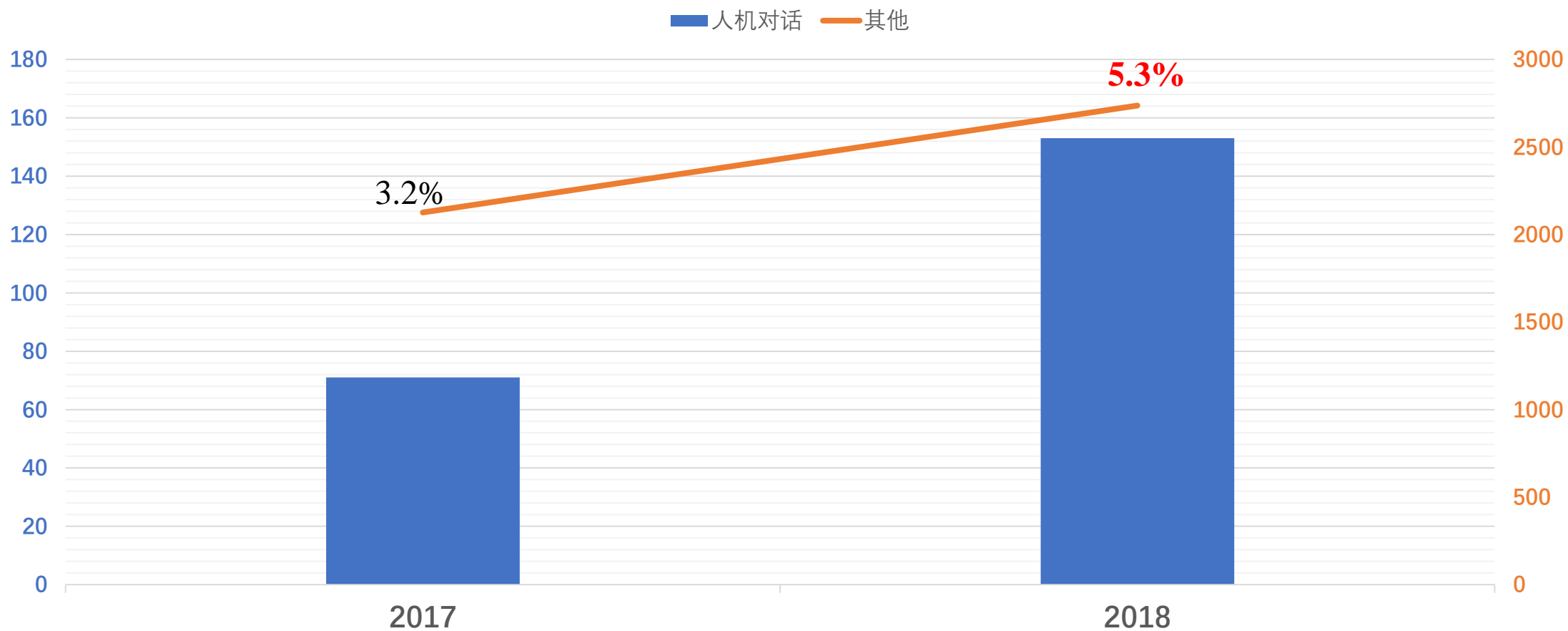
# 大纲

- 热点及趋势分析
- 技术及任务介绍
- 数据及评测汇总

# 人机对话会议论文分布



# 人机对话会议论文数量及比例









# 大纲

- 热点及趋势分析
- 技术及任务介绍
- 数据及评测汇总

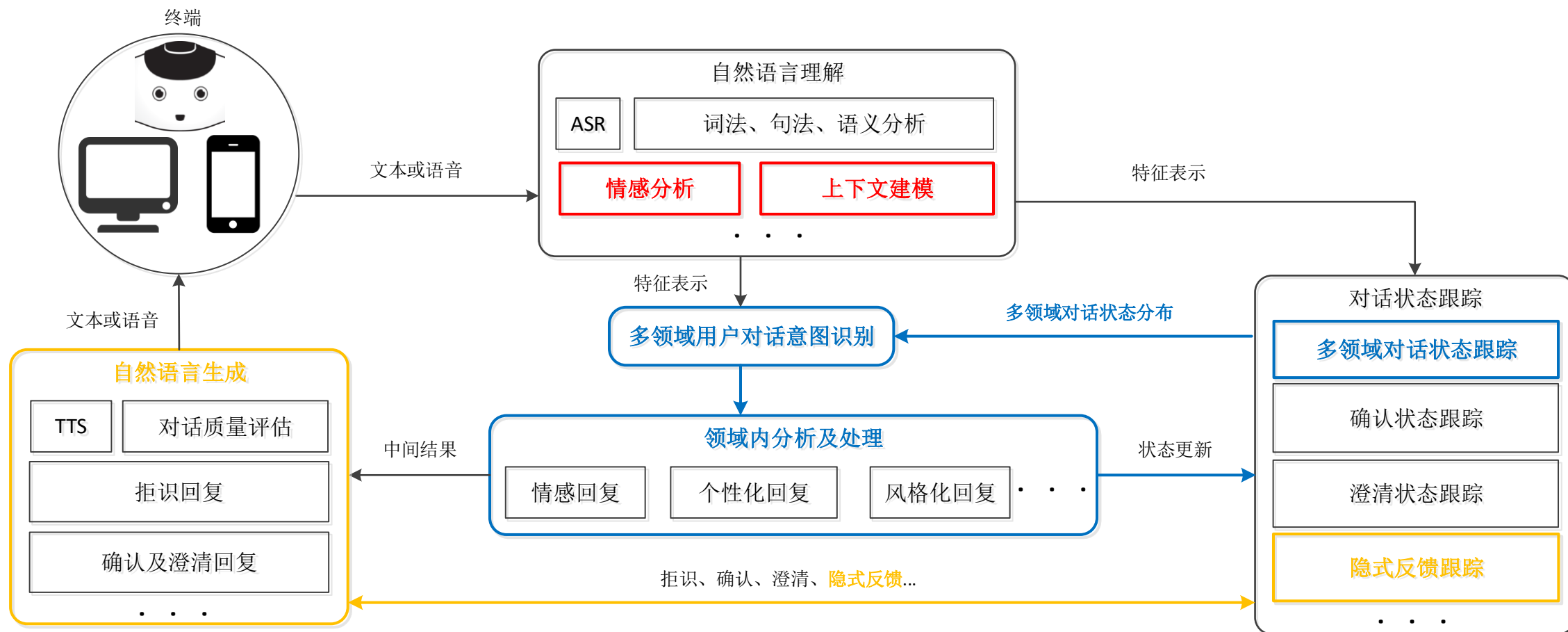
# 人机对话系统功能及相关技术

	自然语言理解	对话管理	自然语言生成
聊天	情感及意图识别	上下文序列建模及 候选回复评分	开放域聊天回复
问答	问句分析及分类	文本的检索及 知识库匹配	文本片段或知识库实体
任务	意图分类及语义槽填充	对话状态跟踪及 策略学习	确认、澄清、完成等
推荐	主题及兴趣识别	用户兴趣匹配及 推荐内容排序	推荐内容

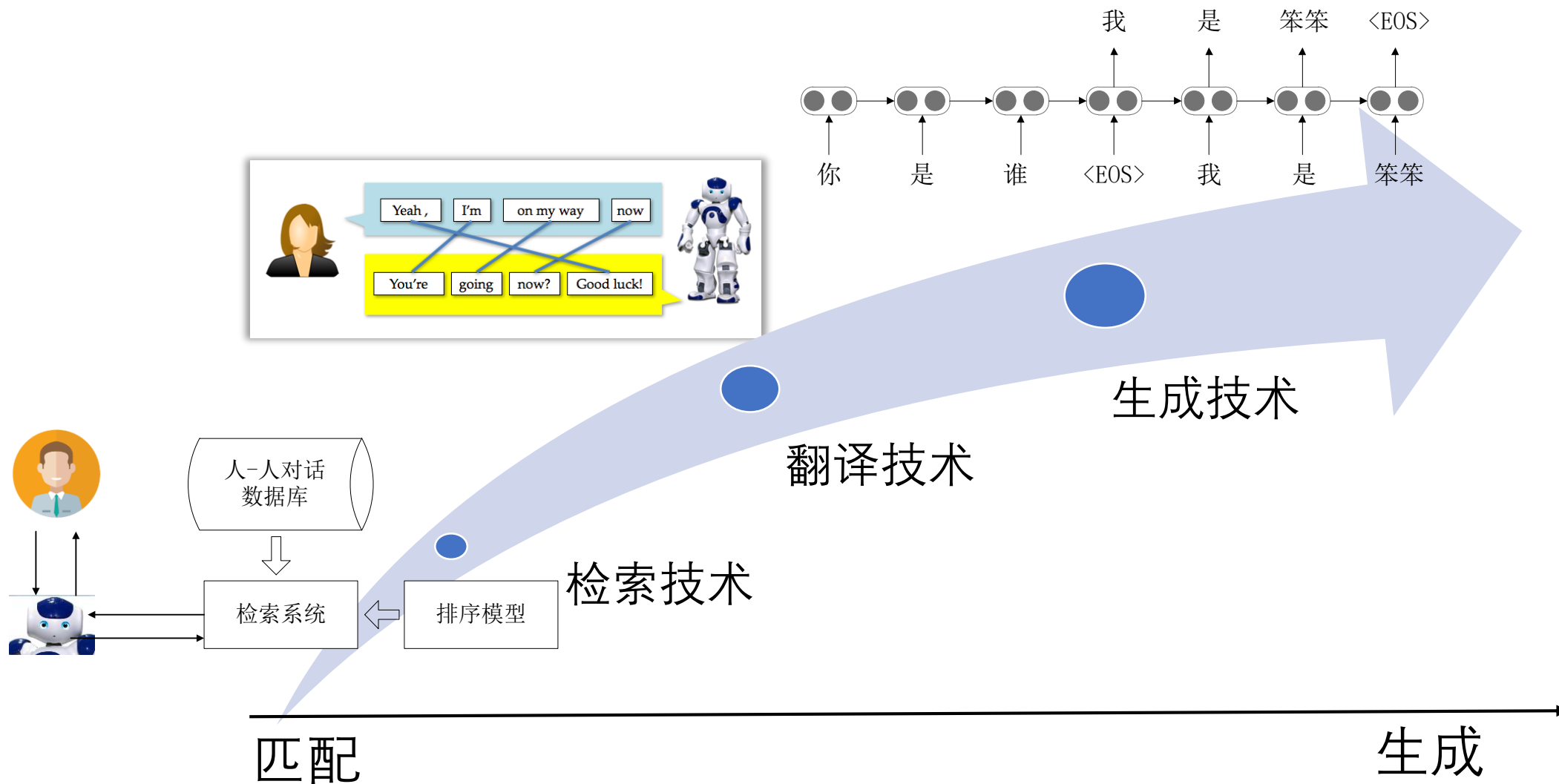
# 人机对话系统功能及相关技术

	自然语言理解	对话管理	自然语言生成
聊天	情感及意图识别	上下文序列建模及候选回复评分	开放域聊天回复
问答	问句分析及分类	文本的检索及知识库匹配	文本片段或知识库实体
任务	意图分类及语义槽填充	对话状态跟踪及策略学习	确认、澄清、完成等
推荐	主题及兴趣识别	用户兴趣匹配及推荐内容排序	推荐内容

# 开放域对话(Open-domain Conversation, 聊天)系统构成



# 聊天技术发展过程



# 聊天技术存在的问题

- M: 你今年多大了?
  - R: *我不知道*
  - M: 你怎么不知道?
  - R: *我也是*
  - M: 呵呵好吧
  - R: *好的*
- M: 我考试得了满分
  - R: 太棒了
  - M: 我该怎么告诉妈妈?
  - R: *快去给她个惊喜*
- M: 我考试没及格
  - R: 悲剧了
  - M: 我该怎么告诉妈妈?
  - R: *快去给她个惊喜*

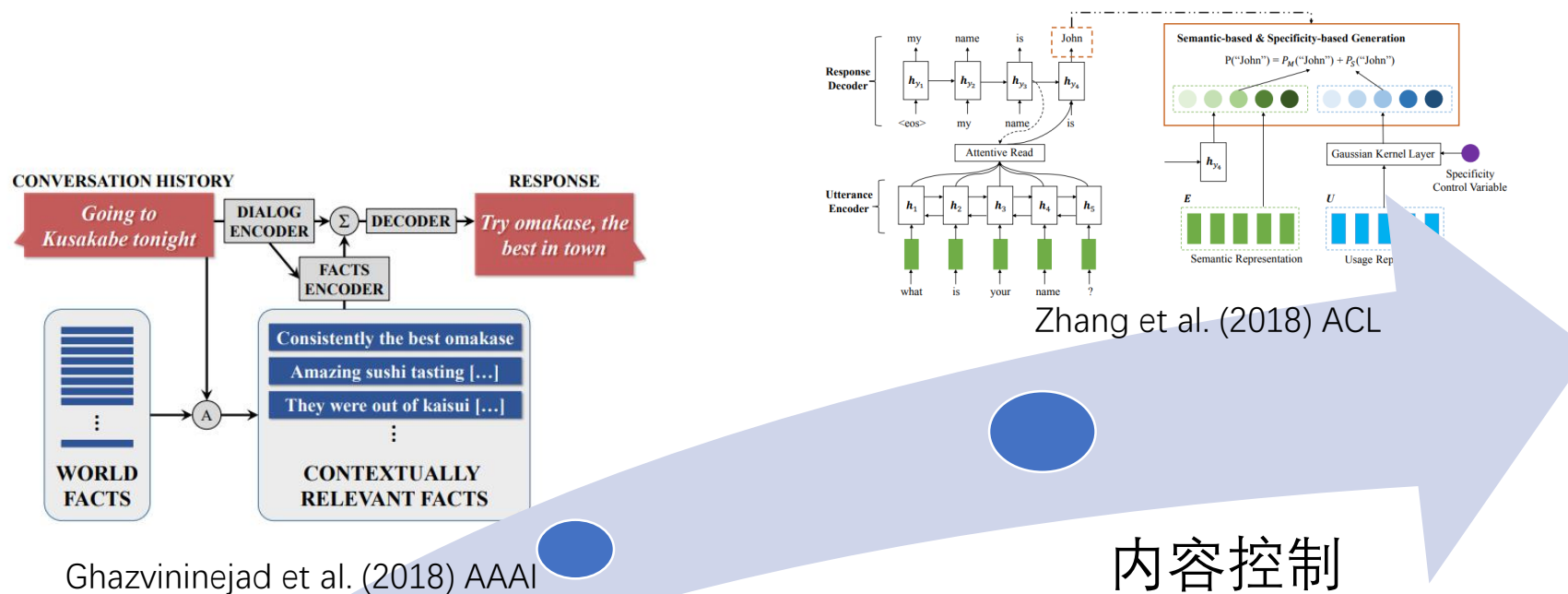
# 聊天技术存在的问题

- M: 你今年多大了?
- R: *我不知道*
- M: 你怎么不知道?
- R: *我也是*
- M: 呵呵好吧
- R: *好的*
- M: 我考试得了满分
- R: 太棒了
- M: 我该怎么告诉妈妈?
- R: *快去给她个惊喜*
- M: 我考试没及格
- R: 我的天!
- M: 我该怎么告诉妈妈?
- R: *快去给她个惊喜*

万能回复  
相关性差

局部学习  
上下文无关

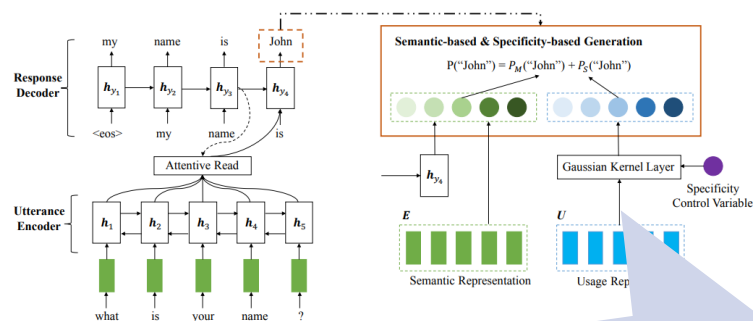
# 单轮聊天生成技术的发展过程



知识嵌入

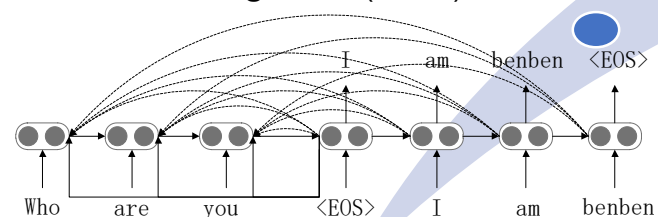
内容控制

Zhang et al. (2018) ACL



Shang et al. (2015) ACL

序列到序列

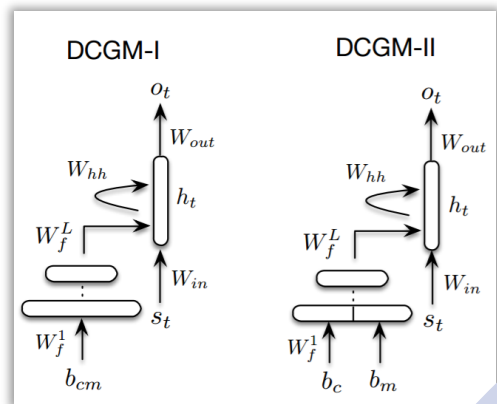


模型优化

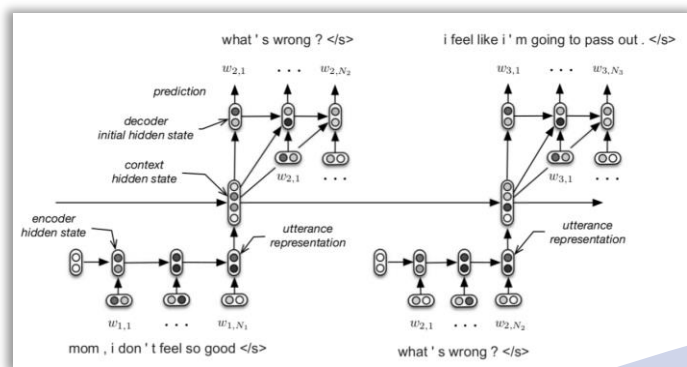
内容优化



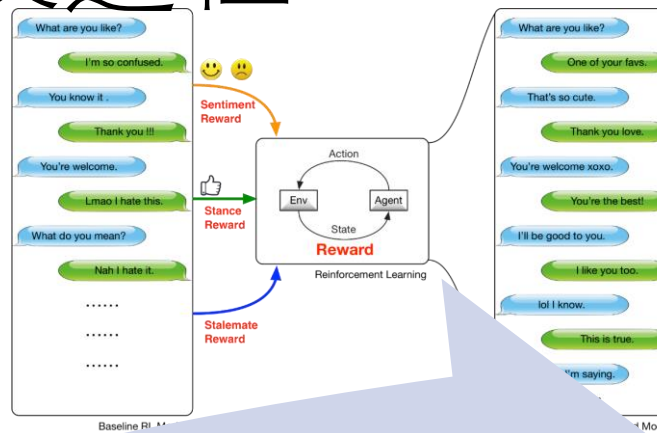
# 多轮聊天生成技术的发展过程



Sordoni et al. (2015) NAACL



Serban et al. (2017) AAAI



Zhang et al. (2018) AAAI

重排序模型

层次化模型

强化学习模型

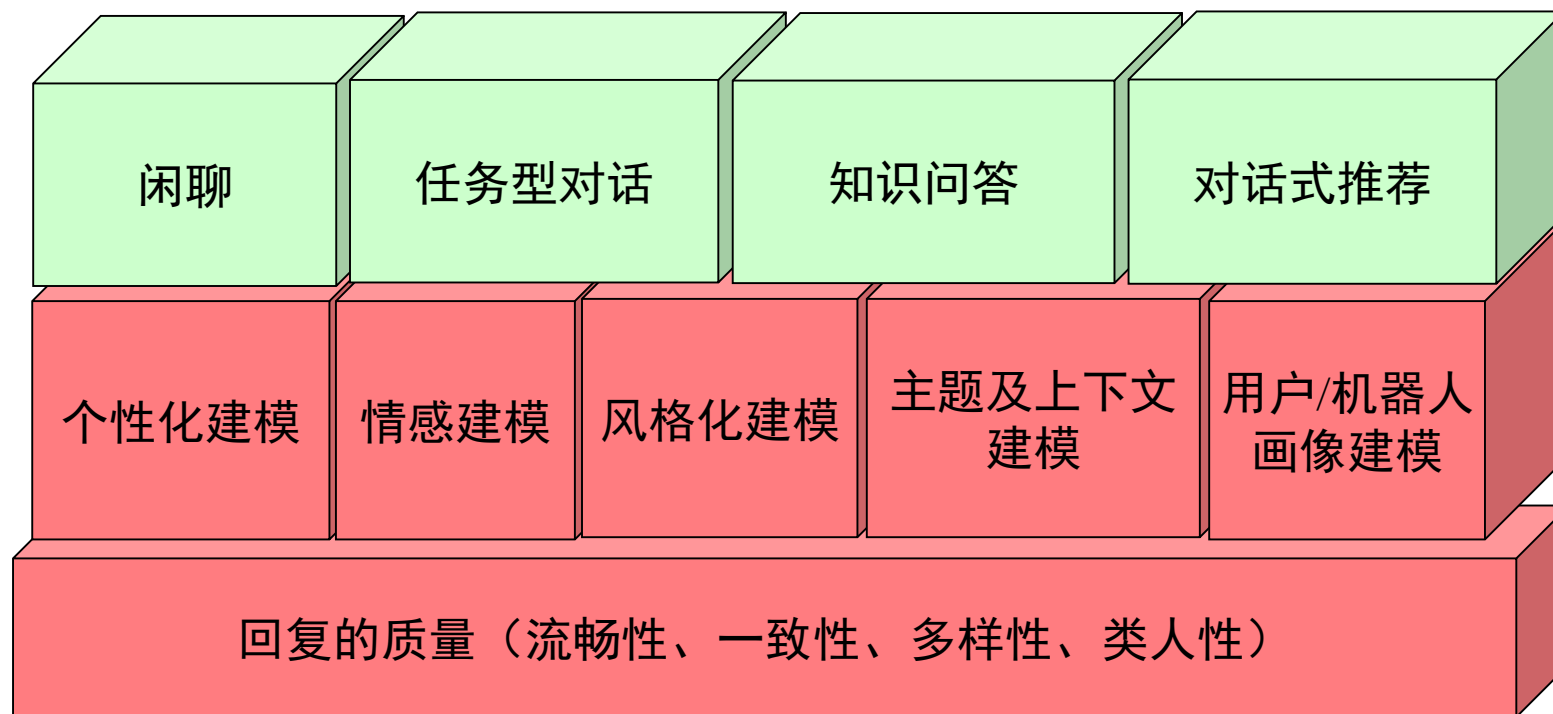
深度学习

强化学习

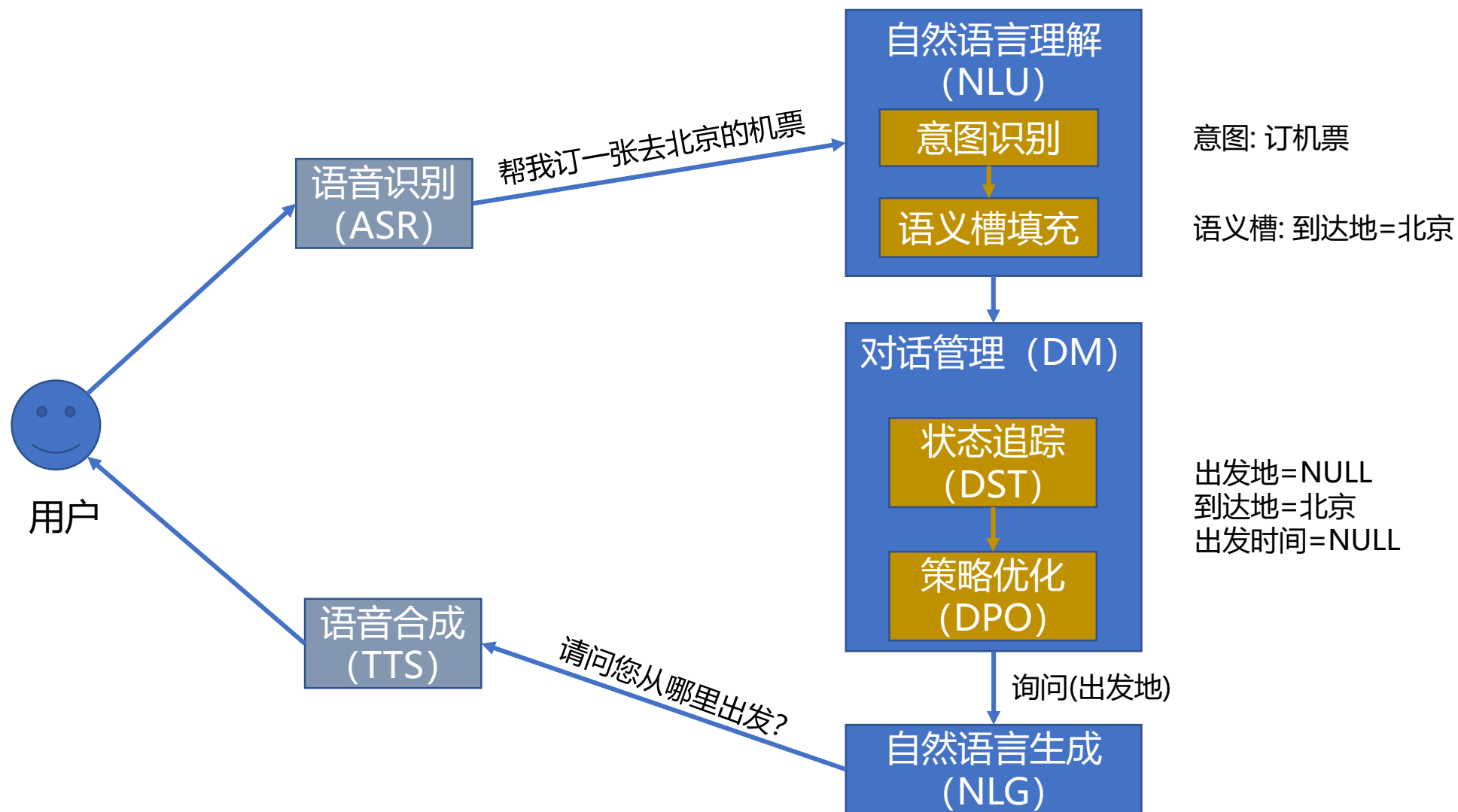
# 聊天技术的最新进展



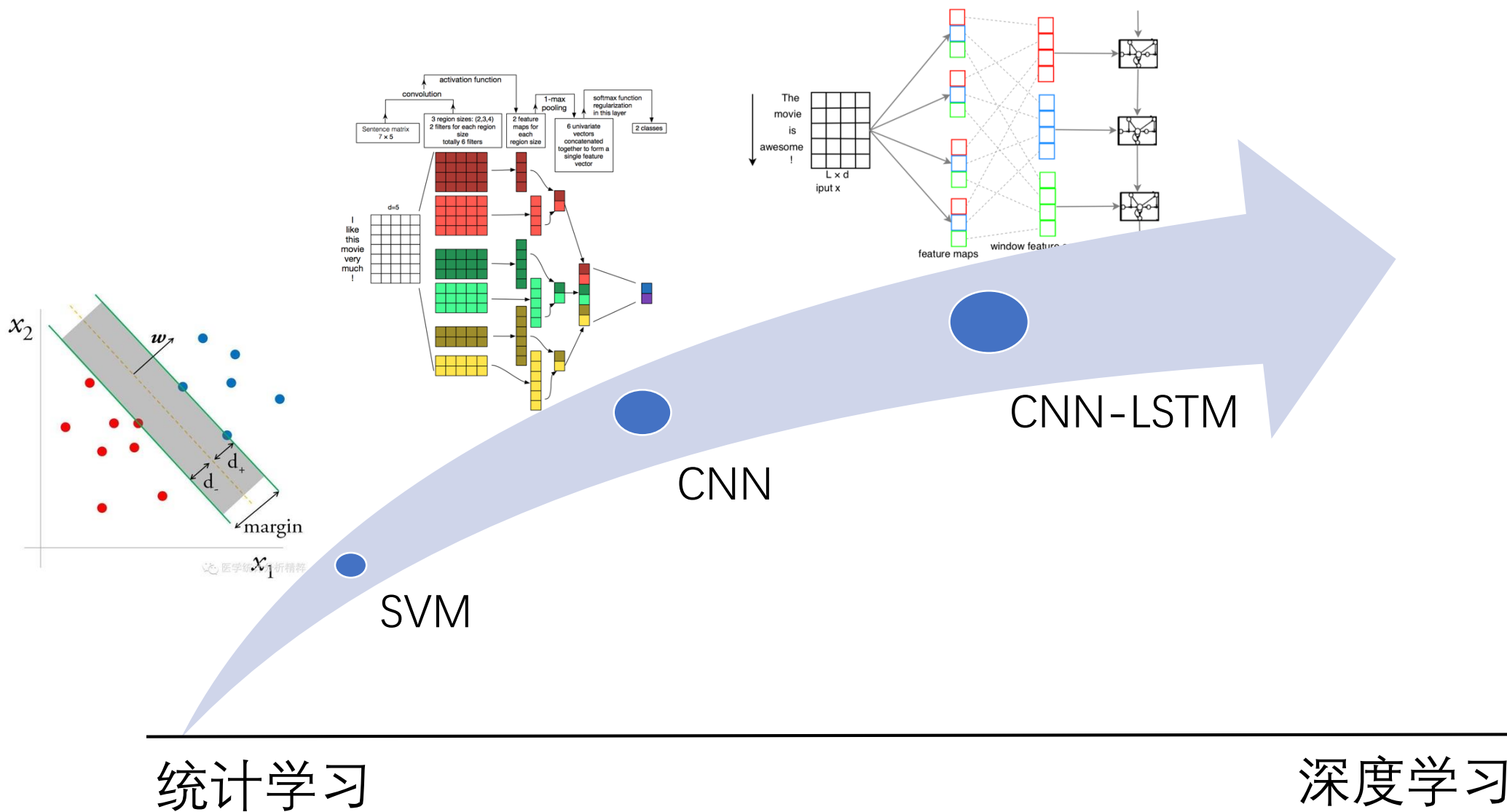
# 聊天技术的发展趋势



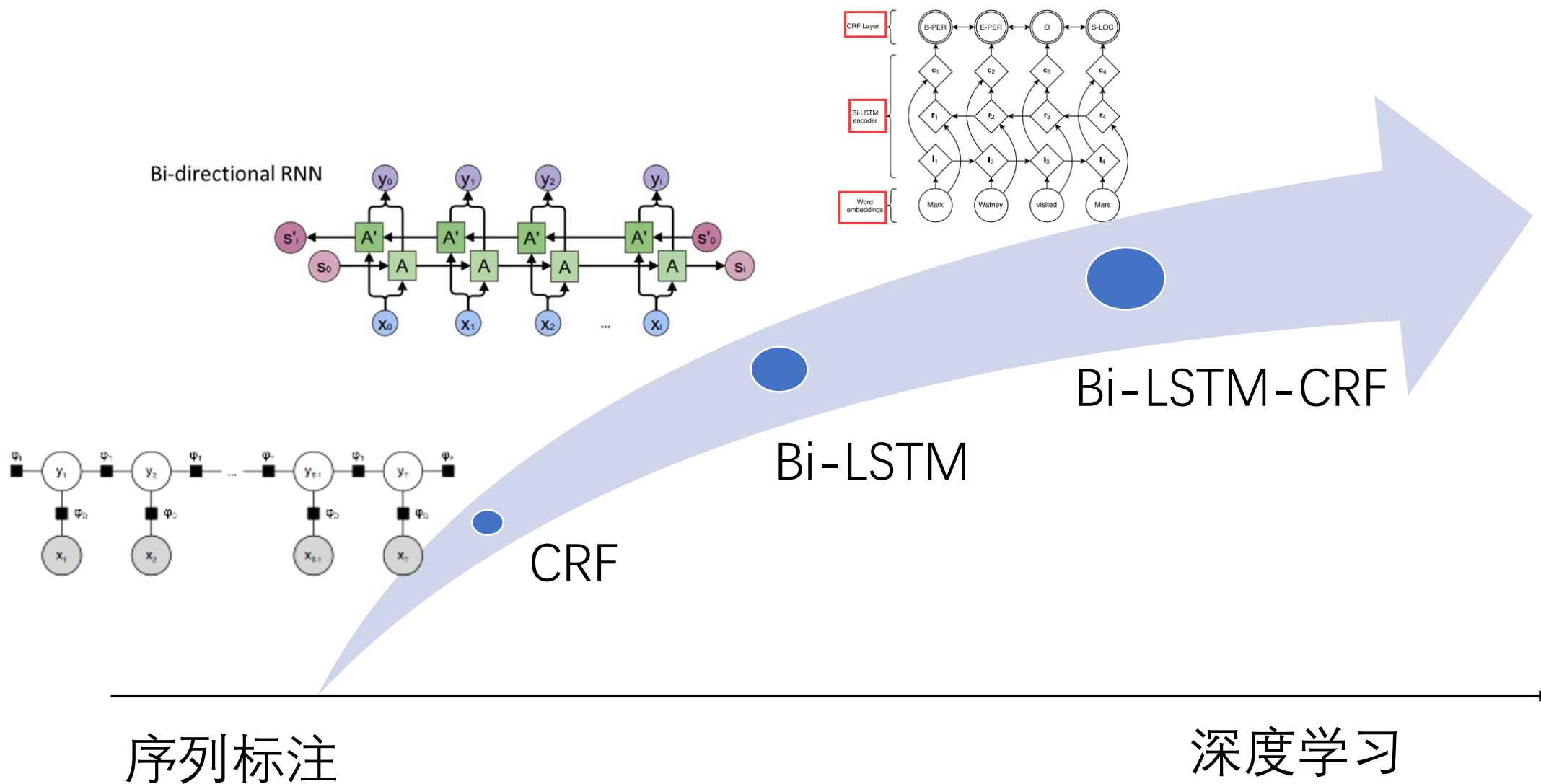
# 任务型对话系统构成



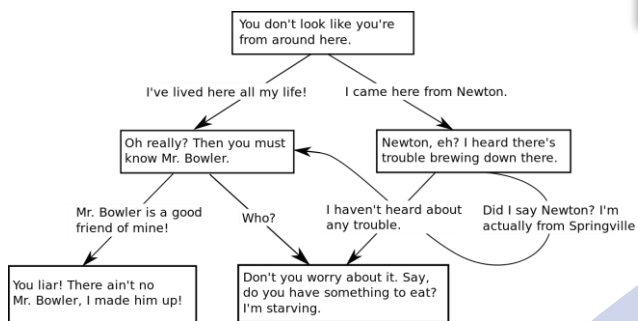
# 意图识别技术发展过程



# 语义槽填充技术的发展过程

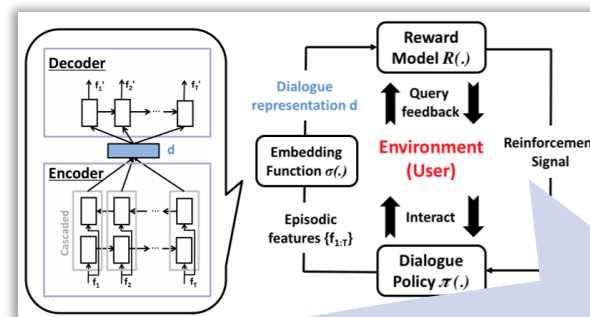


# 对话管理技术的发展过程



```

System: East Pittsburg Bus Schedules. Say a bus route, like 28X, or say I'm not sure.
hello(), request(route), example(route=28x), example(route=dont_know)
User: 61A
SLU: 0.77 inform(route=61a)
      0.12 inform(route=61)
      0.01 inform(route=61d)
System: Okay, 61A. To change, say go back. Where are you leaving from?
impl-conf(route=61a), example(act=goback), request(from)
User: Downtown
SLU: 0.59 inform(from.desc=downtown)
      0.10 inform(from.desc=from downtown)
System: Okay, downtown. You can always say go back. And where are you going to?
impl-conf(from.desc=downtown), example(act=goback), request(to)
  
```



基于规则

基于规则

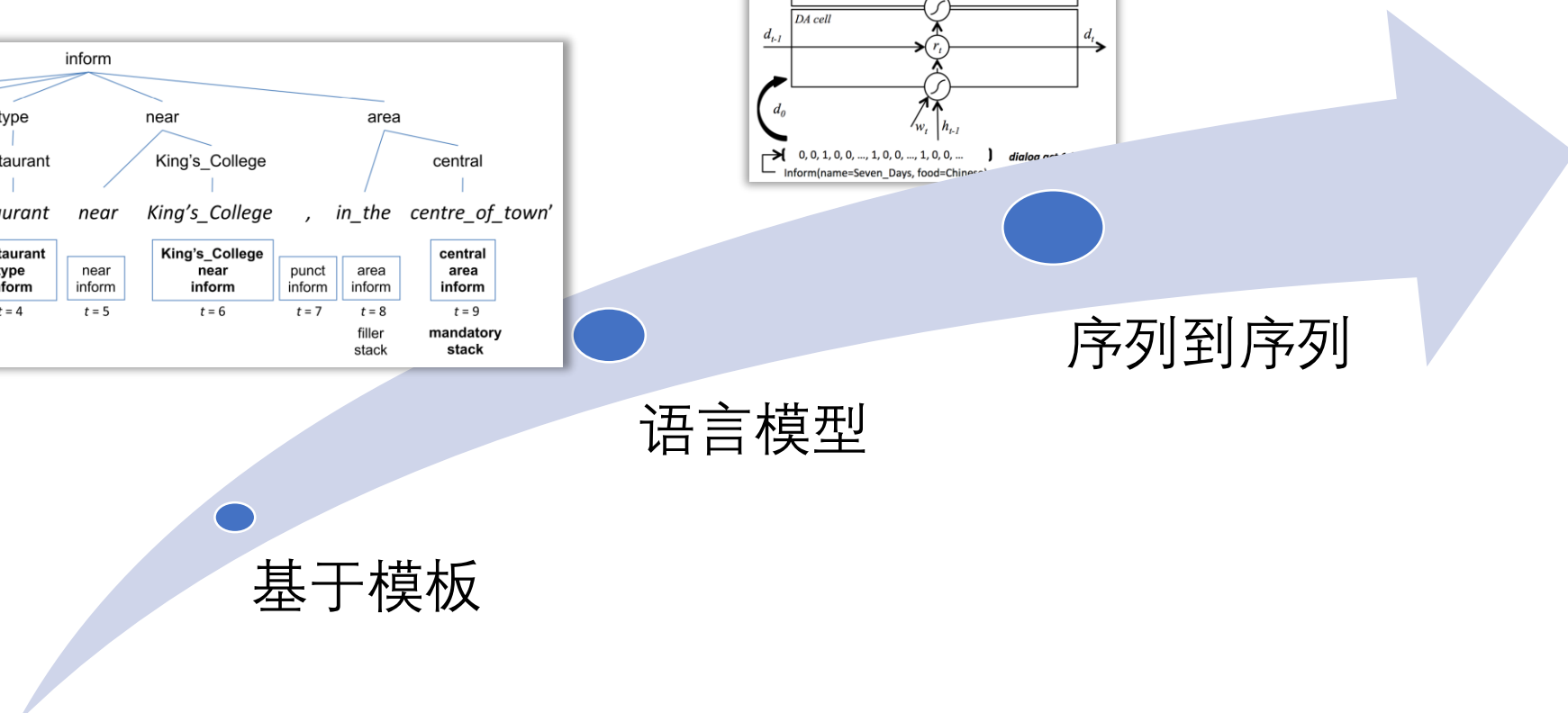
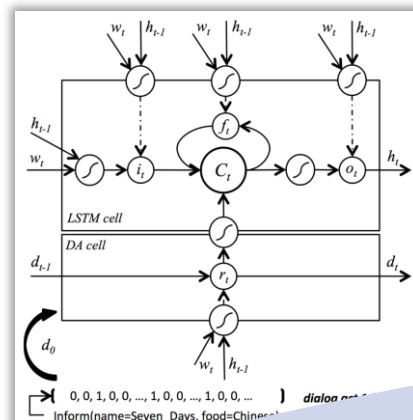
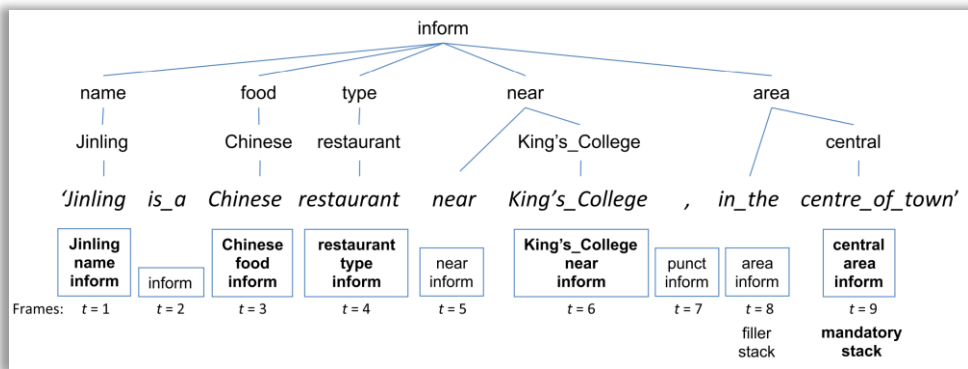
有指导学习

强化学习

策略：追问、澄清、给出结果

强化学习

# 对话生成技术的发展过程



基于模板

语言模型

序列到序列

基于模板

深度学习



# 任务型对话技术的最新进展

自然语言理解 (NLU)	对话状态跟踪 (DST)	对话策略学习 (PL)	自然语言生成 (NLG)
多任务学习	神经网络DST	模仿学习 (初始化)	序列到序列
融合外部知识	基于NBT的DST	BBQ-networks	有监督端到端生成
结合上下文推断		跨任务策略学习	PL与NLG联合建模
NLU与DST联合建模		Deep Dyna-Q数据扩充	

Lee and Derroncourt (2016) NAACL

Chen et al. (2016) ISCA

Hakkani-Tur et al., (2016) INTERSPEECH

Liu and Lane, (2016) INTERSPEECH

Bapna et al., (2017) INTERSPEECH

Lee and Jha, (2018) Arxiv

Mrkšić et al. (2015) ACL

Mrkšić et al. (2017) ACL

Wen et al. (2017) EACL

Lei et al. (2018) ACL

Dhingra et al., (2017) ACL

Lipton et al. (2018) AAAI

Chen et al. (2018) COLING

Peng et al. (2017) EMNLP

Tang et al. (2018) EMNLP

Peng et al. (2018) ACL

Dhingra et al., (2017) ACL

Bordes et al., (2017) ICLR

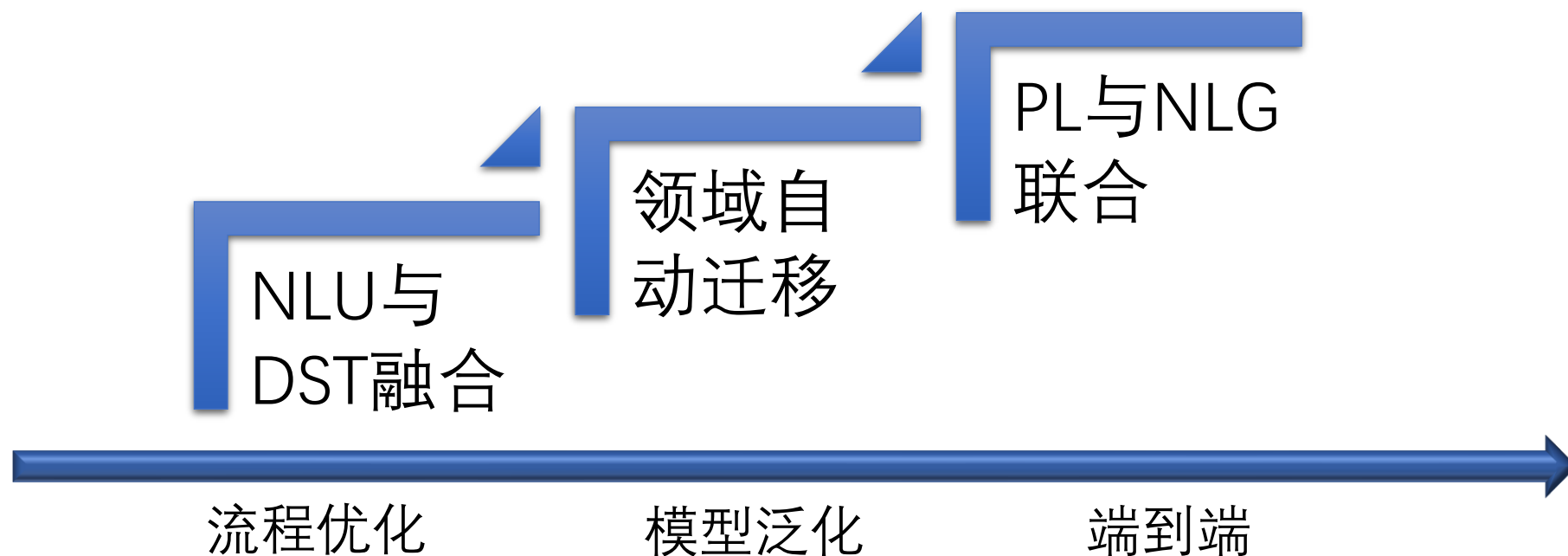
Eric et al. (2017) SIGDIAL

Li et al. (2017) IJCNLP

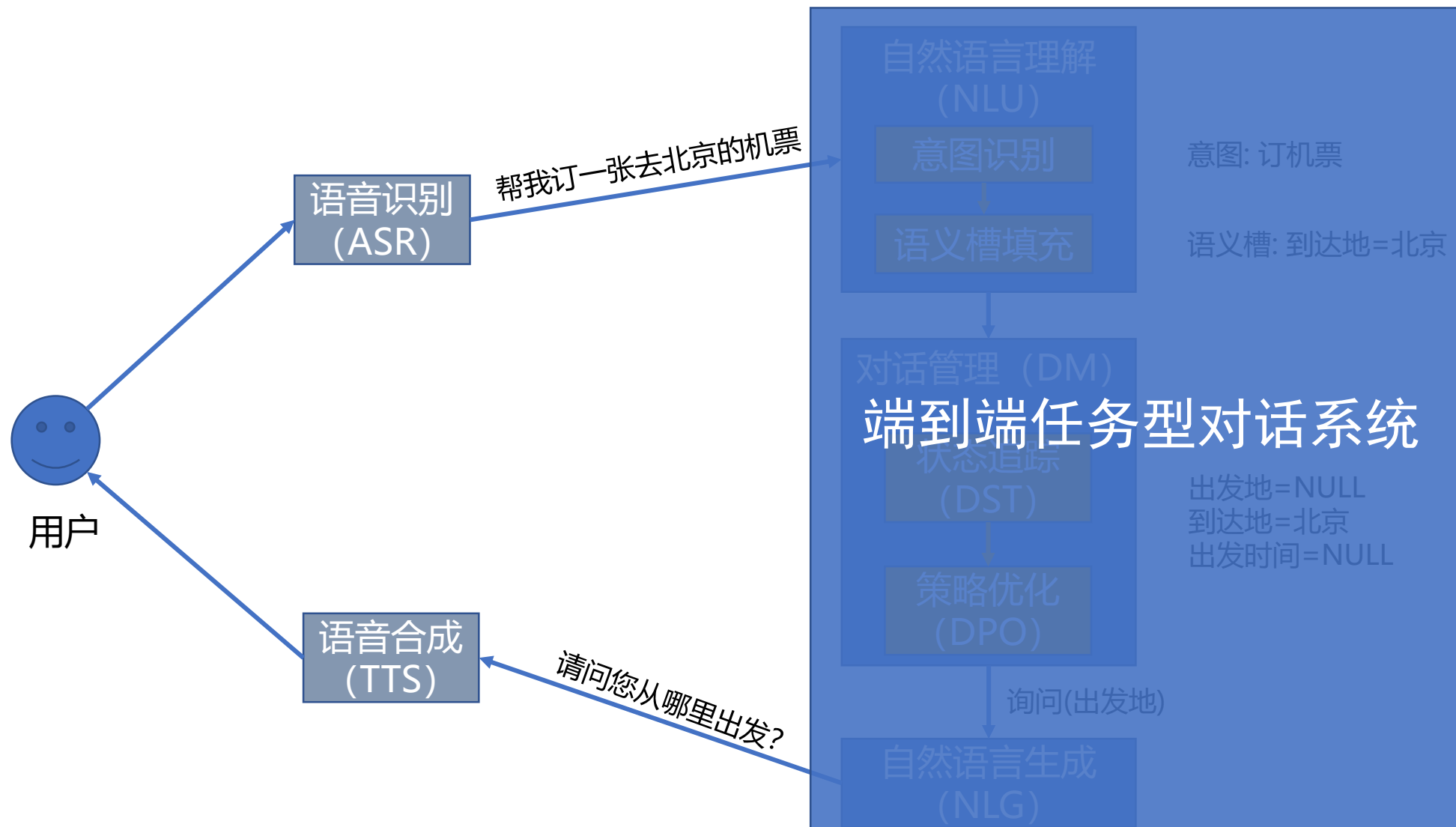
Williams et al., (2017) ACL

# 任务型对话技术的最新进展

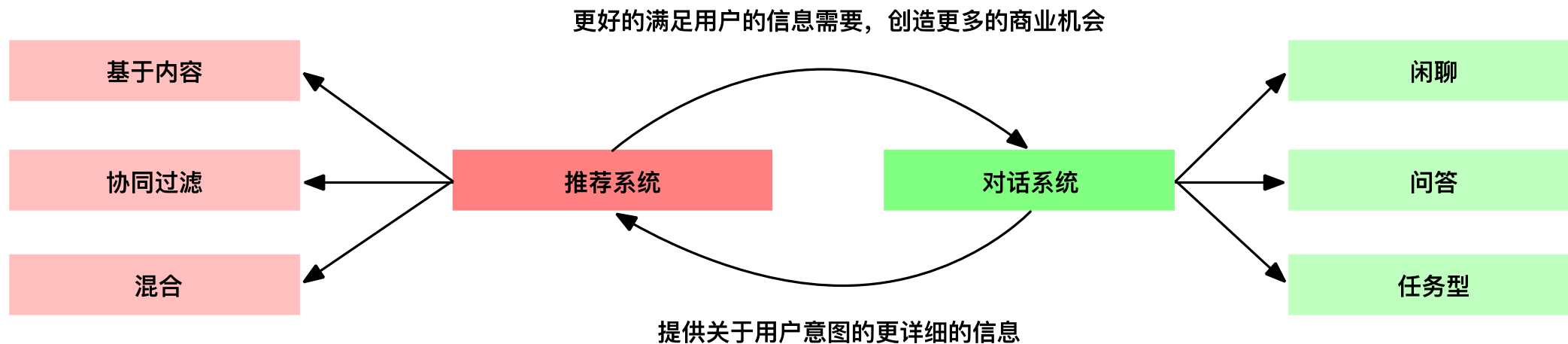
自然语言理解 (NLU)	对话状态跟踪 (DST)	对话策略学习 (PL)	自然语言生成 (NLG)
多任务学习	SLU及DST联合建模	模仿学习 (初始化)	序列到序列
融合外部知识	神经网络DST	BBQ-networks	有监督端到端生成
结合上下文推断	基于NBT的DST	跨任务策略学习	PL与NLG联合建模
NLU与DST联合建模		Deep Dyna-Q数据扩充	



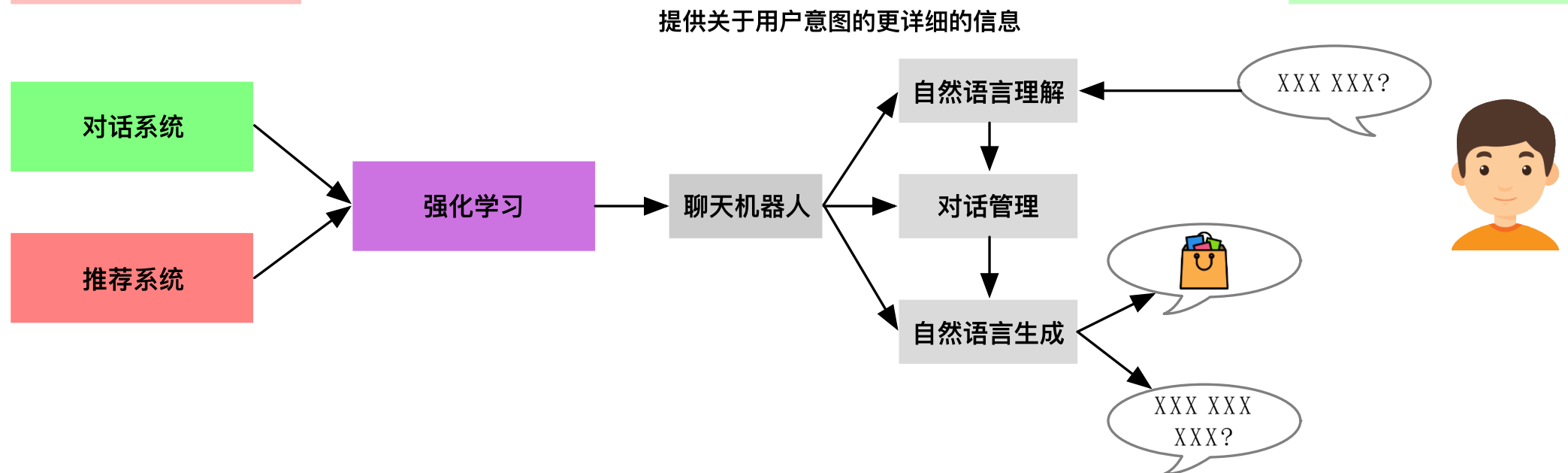
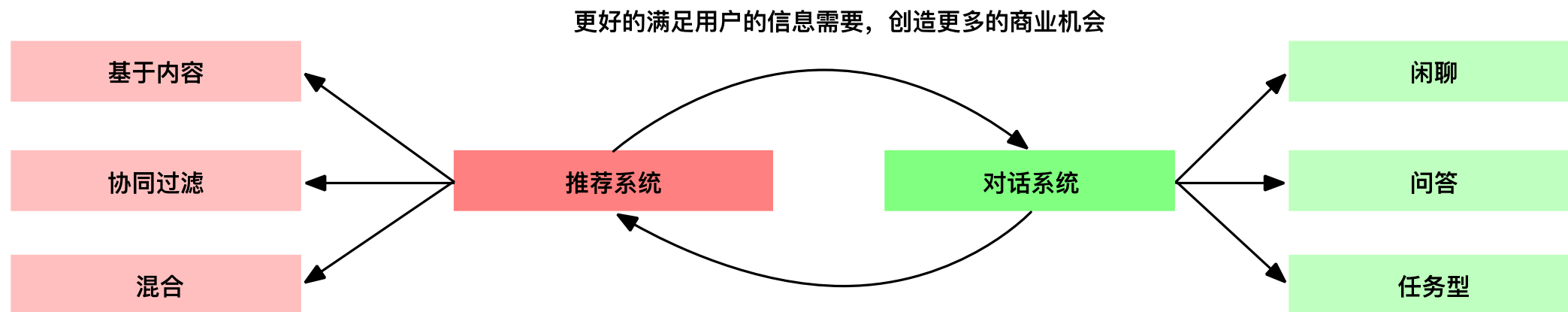
# 任务型对话系统的发展趋势



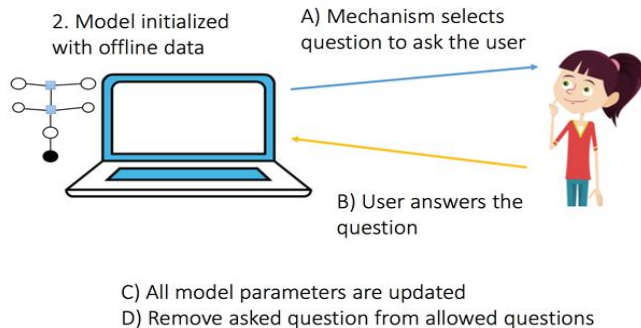
# 对话式推荐动机及方案



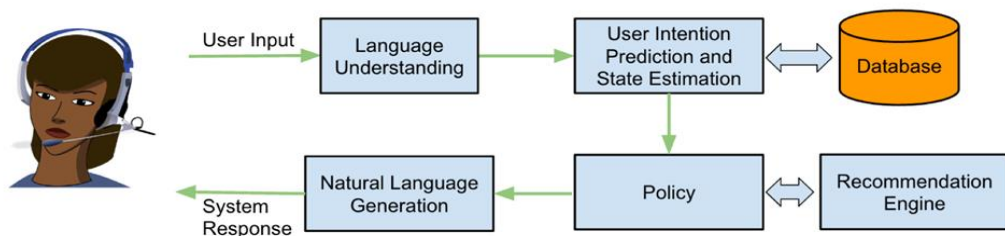
# 对话式推荐动机及方案



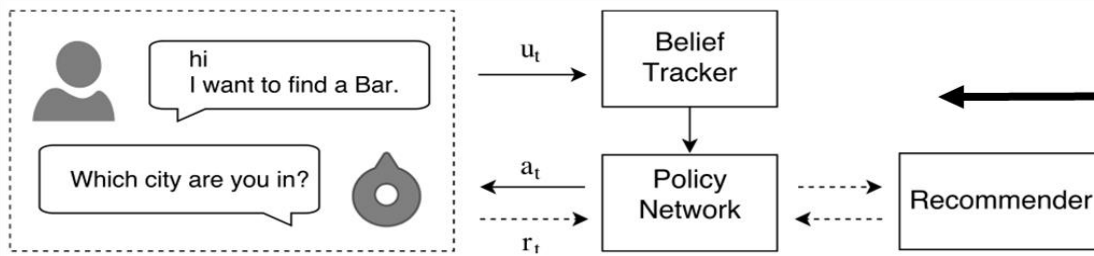
1. Pick a model and a preference elicitation mechanism



Konstantina Christakopoulou et al, Towards Conversational Recommender Systems, KDD, 2016

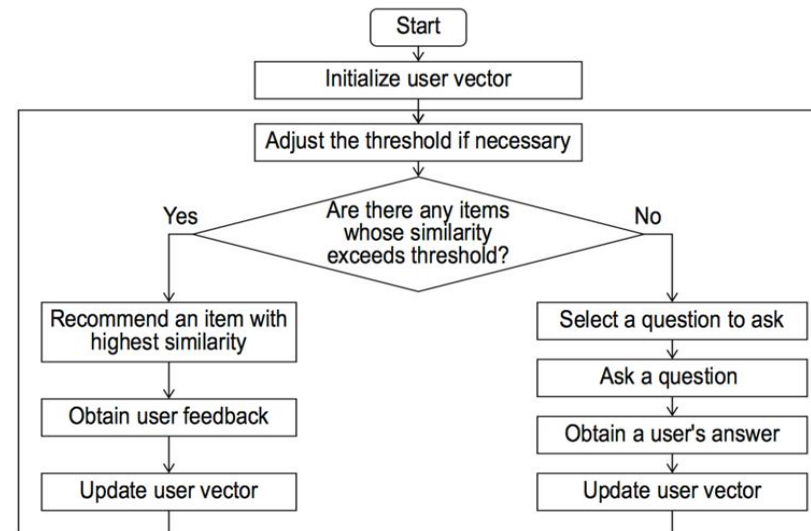


Yueming Sun et al, Conversational Recommendation System with Unsupervised Learning, RecSys, 2016

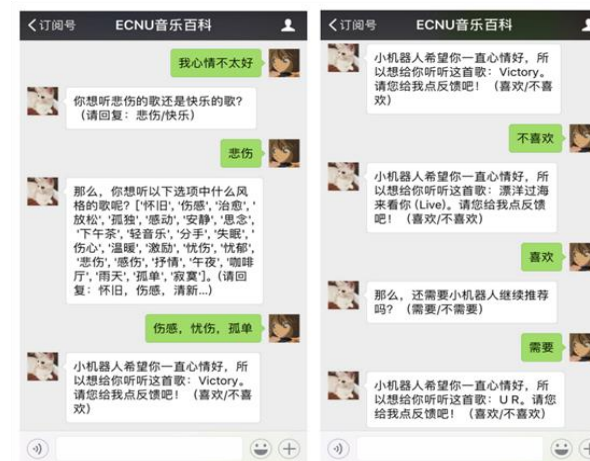


Yueming Sun et al, Conversational Recommender System, SIGIR, 2018

2016



Yuichiro Ikemoto et al, Conversation Strategy of a Chatbot for Interactive Recommendations, Springer, 2018



Chunyi Zhou et al, MusicRoBot: Towards Conversational Context-Aware Music Recommender System, Springer, 2018

2018

# 大纲

- 热点及趋势分析
- 技术及任务介绍
- 数据及评测汇总

# 开放域对话数据集

语言	任务类型	数据名称	数据类型	发布机构	语料规模	语料地址
英文	聊天	Ubuntu	论坛对话	麦吉尔大学	692,321	<a href="http://dataset.cs.mcgill.ca/ubuntu-corpus-1.0/">http://dataset.cs.mcgill.ca/ubuntu-corpus-1.0/</a>
		Opensubtitles	影视脚本	格罗宁根大学	40,000,000+	<a href="http://opus.nlpl.eu/OpenSubtitles.php">http://opus.nlpl.eu/OpenSubtitles.php</a>
		Twitter	Tweets	MSR	12,696	<a href="https://www.microsoft.com/en-us/download/details.aspx?id=52375">https://www.microsoft.com/en-us/download/details.aspx?id=52375</a>
		Cornell Movie-Dialogs Corpus	影视脚本	康奈尔大学	304,713	<a href="https://www.cs.cornell.edu/~cristian/Cornell_Movie-Dialogs_Corpus.html">https://www.cs.cornell.edu/~cristian/Cornell_Movie-Dialogs_Corpus.html</a>
		DailyDialog	日常会话	香港理工大学/中科院/萨兰大学	103,632	<a href="http://yanran.li/dailydialog">http://yanran.li/dailydialog</a>
	对话问答	CoQA	新闻/试卷/儿童故事	斯坦福大学	127,000	<a href="https://stanfordnlp.github.io/coqa/">https://stanfordnlp.github.io/coqa/</a>
		QuAC	教师-学生问答对话	AI2/华盛顿大学/斯坦福大学/Umass Amherst	98,407	<a href="http://quac.ai/">http://quac.ai/</a>
中文	聊天	STC	微博对话	NTCIR12/13/14	5,000,000+	<a href="http://ntcir12.noahlab.com.hk/stc.htm">http://ntcir12.noahlab.com.hk/stc.htm</a>
		CECG/STC3	微博文本	清华大学/NTCIR14	1,115,000	<a href="http://www.aihuang.org/p/challenge.html">http://www.aihuang.org/p/challenge.html</a>
		豆瓣	论坛对话	北航/MSRA	1,060,000+	<a href="https://github.com/MarkWuNLP/MultiTurnResponseSelection">https://github.com/MarkWuNLP/MultiTurnResponseSelection</a>



# 任务型对话数据集

语言	任务类型	数据名称	数据类型	发布机构	语料规模	语料地址
英文	DM/端到端	DSTC (2-6)	多垂类	MSR/剑桥大学 /COLIPS	1,612	<a href="https://www.microsoft.com/en-us/research/event/dialog-state-tracking-challenge/">https://www.microsoft.com/en-us/research/event/dialog-state-tracking-challenge/</a>
	对话帧检测/端到端	FRAMES	旅行	Maluuba	1,369	<a href="https://datasets.maluuba.com/Frames">https://datasets.maluuba.com/Frames</a>
	DM/端到端	KVRET	日历/导航/天气	斯坦福大学	3000	<a href="https://nlp.stanford.edu/blog/a-new-multi-turn-multi-domain-task-oriented-dialogue-dataset/">https://nlp.stanford.edu/blog/a-new-multi-turn-multi-domain-task-oriented-dialogue-dataset/</a>
	DM/端到端	MultiWoZ	多垂类	剑桥大学	8,438	<a href="http://dialogue.mi.eng.cam.ac.uk/index.php/corpus/">http://dialogue.mi.eng.cam.ac.uk/index.php/corpus/</a>
	端到端	Microsoft Dialogue Challenge	订电影票/订餐/ 订出租车	Microsoft	10,087	<a href="https://www.microsoft.com/en-us/research/publication/microsoft-dialogue-challenge-building-end-to-end-task-completion-dialogue-systems/">https://www.microsoft.com/en-us/research/publication/microsoft-dialogue-challenge-building-end-to-end-task-completion-dialogue-systems/</a>
	意图理解	ATIS	订飞机票	BBN/CMU/MIT/SRI	7,300	<a href="https://www.kaggle.com/siddhadev/atis-dataset">https://www.kaggle.com/siddhadev/atis-dataset</a>
中文	意图分类	CCL18	电信	中移在线	100,000	<a href="http://www.cips-cl.org/static/CCL2018/call-evaluation.html">http://www.cips-cl.org/static/CCL2018/call-evaluation.html</a>
	意图分类	ECDT-Task1	多垂直领域	哈工大/讯飞	3,832	<a href="http://smp2018.cips-smp.org/ecdt_index.html">http://smp2018.cips-smp.org/ecdt_index.html</a>
	在线对话	EDDT-Task2	机票/火车票/酒店	哈工大/讯飞	-	

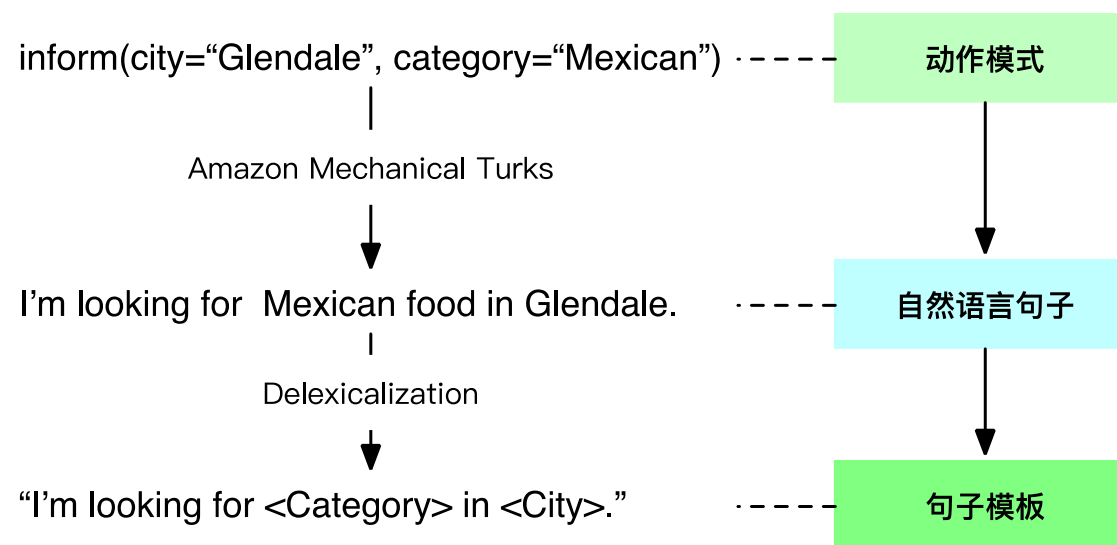
# 对话式推荐技术数据集

## Yelp

以 **Yelp** 为基础数据集，推荐数据集中的餐馆和食品信息

人工构建人人对话

**Yelp** 推荐数据集包含丰富的评价信息，但是它不包括任何人人对话



Yueming Sun et al, Conversational Recommender System, SIGIR, 2018

# 人机对话技术评测

	开放域对话		任务型对话						对话式问答
组织机构	NIPS ConvAI2	Amazon Alexa Prize	DSTC 1-6	DSTC 7	Microsoft Dialog Challenge	JD DC	SMP-ECDT	CCL	CoQA
	MIPT、FAIR、UM、McGill、CMU	Amazon	MSR/Cambridge University/COLIPS MERL/Naver Labs/NAIST/Adobe等		MSR	京东	哈工大+讯飞	中移在线	Stanford
任务简述	角色对话	闲聊	端到端任务型对话、闲聊、对话终止识别	句子选择、生成、多模态对话	端到端任务型对话	客服对话	意图分类、任务型对话	意图分类	对话式问答
评测方式	客观+人工	人工	客观+人工	客观+人工	客观+人工	客观+人工	客观+人工	客观	客观+人工
评测时间	2017/2018	2017/2018	2012-2017	2018	2018	2018	2017/2018	2017/2018	2018

# 可带走的干货：待研究的问题

- 开放域对话有待研究的问题
  - 开放域用户意图、情感建模、用户画像
  - 回复一致性、多样化、个性化、风格化
  - 机器人主导话题、机器人画像
  - 上下文建模、隐式反馈
- 任务型对话有待研究的问题
  - 多任务切换问题、领域自适应问题
  - 端到端任务型对话系统
- 对话式推荐有待研究的问题
  - 对话管理与推荐“行为”的融合问题
  - 通过对话的方式获取个性化的用户表示

# 总结

- 问题依旧存在
- 挑战只增不减
- 研究方兴未艾
- 应用层出不穷



CCL 2018

谢谢!

