

自然语言处理
人机对话

- 引言
- 目标驱动型人机对话系统
- 子任务各自建模
- 多任务联合建模
- 结论与展望

- **人际交流**是人类社会的一个主要特征，是推动人类社会向前发展的重要动力。
- **自然语言**(汉语、英语、法语等等)是人际交流的主要工具之一。
- **对话**(Dialogue; Conversation会话)是使用自然语言进行人际交流的主要方式之一。

对话(dialogue; conversation会话)

■两人或多人之间通过人类的书面或口头语言进行的交互。独白(Monologue)也常被语言学家归入会话。



■两个词：

- Conversation vs. Dialogue
- 交换意见 vs. 讨论
- 非正式(日常) vs. 较正式(小说戏剧等, 谈判)

■Conversation Agent

■Dialogue Agent

■ 早期对话被认为是无规律的

■ 对比：个人语篇的结构性

■ 多人参加导致的随意性

对话的语言分析-会话分析

■会话分析(conversation analysis)

■Sacks、Schegloff、Jefferson:1960s-

■会话具有稳定的结构

■{会话}{相邻对{话轮}}}

■会话的话轮转换机制

■两个要素，一套规则

■.....

[人类] 小冰你好

[小冰] 你好呀，你头像好有趣，感觉在哪见过

[人类] 是吗

[小冰] 提不是哪部动画里的呀

[人类] 你怎么知道的

[小冰] 你猜

[人类] 难道你是真人吗？

[小冰] 这个，你和我多聊聊就知道了

对话的语言分析-会话分析

■会话结构

■开始会话—4阶段结构说

- S1：发起会话人建立会话标识
- S2：人员确认
- S3：建立对话意愿
- S4：第一个话题，通常由发起人提出
- 例：

- S1: A: (rings B's telephone)
- S1,S2 B: Benjamin Holloway
- S2 A: this is Professor Dwight's secretary
- S2,S3 B: ooh yes –
- S4 A: uh:m . about the lexicology seminar
-

■结束

对话的语言分析-语用学

- 言语行为理论
- 言语即行为

■言语行为：语言是言语者行为的体现，一些言语行为：

- 断言Assertives: 言语者对某事的判断 (*suggesting, putting forward, swearing, boasting, concluding*).
- 指令Directives: 言语者希望听着做某事 (*asking, ordering, requesting, inviting, advising, begging*).
- 承诺Commissives: 言语者允诺未来的某种行为 (*promising, planning, vowing, betting, opposing*).
- 表达Expressives: 言语者心理状态的表达(*thanking, apologizing, welcoming, deplored*)
- 宣告Declarations: 言语者试图通过语言给现实状态带来变化 (*I resign, You're fired.*)

■联合行为：对话共识

■封闭性原则：基于对话主题的对话行为对于当前目标而言是充分的。

■达成该原则的两个阶段：

■1：表达(presentation)

■说话人向听话人发出言语行为

■2：接受(Acceptance)

■听话人寻求共识，表明其是否理解

■由弱到强的寻求共识的方法

■Continued attention\Relevant next contribution\Acknowledgement\...

对话的语言分析-语用学

- 言语行为理论

- 言语即行为

- 格赖斯会话含义理论

- 合作原则

- 利奇的礼貌原则

- 霍恩的会话含义两原则：量、关系原则

- 列文森的会话含义三原则：量、信息、方式原则

■会话合作准则

■量 (Maxim of Quantity)

■必要信息 Be exactly as informative as is required

■质 (Maxim of Quality)

■有效信息 Try to make your contribution one that is true

■相关 (Maxim of Relevance)

■相关信息 Be relevant.

■方式 (Maxim of Manner)

■清晰信息 Be perspicuous

■例1：对话背景：A到外地探亲丢了地址，就发电报回家问B：

■A：知道三叔家的地址吗？

■B：知道（信息量不够）

■例2：来自Hamlet, William Shakespeare

■Polonius: What do you read my lord?

■Hamlet: Words, words, words (信息不够)

■Polonius: What is the matter, my lord?

■Hamlet: Between who?

■Polonius: I mean, the matter that you read, my lord

■Hamlet: Slanders, sir: for the satirical rogue says here that old men have gray beards, that their faces are wrinkled, their eyes purging thick amber and plum-tree gum, ... (信息过多)

对话的语言分析-互动语言学

- Ocha, Schegloff, Thompson(1996): 互动中的语法
- Couper-Kuhlen, Selting(2001): 互动语言学
- 话语功能语法: Du Bois, Helasvuo, 接受者设计原则
- 认知语法: Etelamaki, Visapaa, 认知赋予会话意义
- 构式语法: Schegloff, 扩展构式结合互动分析和结构分析
- 浮现语法: Hopper, 语法会在语言使用中不断更新和塑造
- 在线句法、社会行为结构、线性单位语法、对话语法

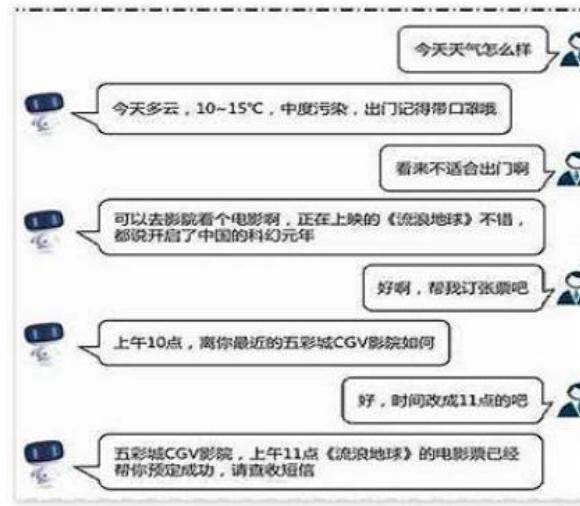
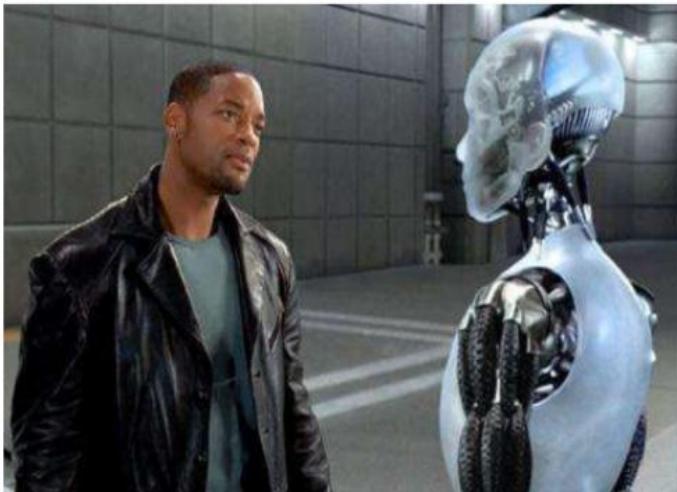
- 看似随意的对话背后存在着一定的规律
- 已经发现了一些规律，并应用于构建对话系统中，获得了初步的成功
- 因此，通过进一步发现、建模并运用对话的规律，可能实现能与人进行有效对话的机器

人机对话(Man-Machine Dialogue)

■人与机器之间通过人类的书面或口头语言进行的交互。

■Spoken Language Dialogue

■Text-based Dialogue



人机对话：越来越广泛的应用价值

■特定目标的应用需求：

- 客服：咨询和办理特定业务(订票)等等：人手不够、或人力昂贵
- 心理治疗：罗杰斯心理疗法：倾述、聆听。需要对话环境、聆听者

■人对交流的需求

- 对话是最方便、常见的交流方式。想找人交流的时候，找不到人，设计机器对话者，机器随时听从召唤。这导致聊天机器人的出现。
- 人在交流中对隐私保护的需求，找人交流容易传播开，一些隐私只想发泄一下，不想传播。机器保密性好。

■人机共生的必然需求

- 机器深入到人类生活，人和机器要越来越多打交道，编程序？命令输入？以人类语言为最方便(语言)、以及时响应交互(对话)为最方便。
- 电脑：电子助手；机器人：服务机器人

■人机对话的多样性(人人对话的多样性)

- 闲聊：不限主题的随意交谈
- 问答：针对特定问题的交谈
- 咨询：针对特定目标的交谈
- 探讨：针对特定主题的交谈
- 谈判：共同构建目标的交谈
-

■聊天型对话

■方法

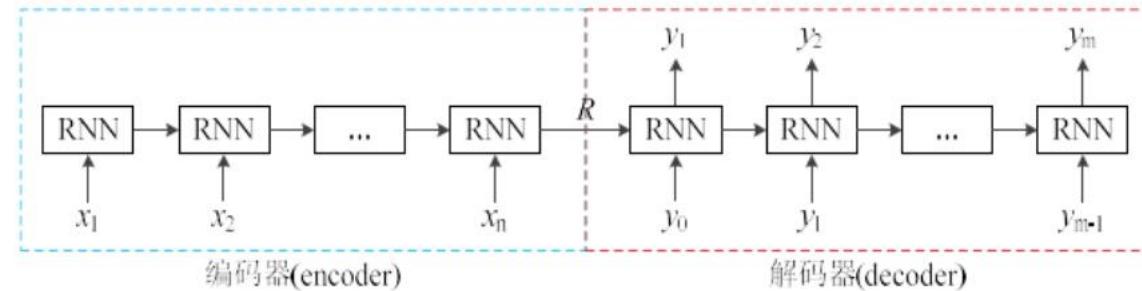
- 基于规则的方法 ELIZA
- 基于语料的方法 Seq2Seq

- 基于检索的方法
- 基于生成的方法

- 基于知识的方法
- 多种方式结合

■能力：能进行应答,和规则以及语料密切相关

■问题：安全回答、未见处理



- 针对特定目标的交谈

- 特定业务的客服

- 电信、银行、购物……

- 特定内容的教学

- 指令集、问题集

- ……

- 目标驱动型人机对话

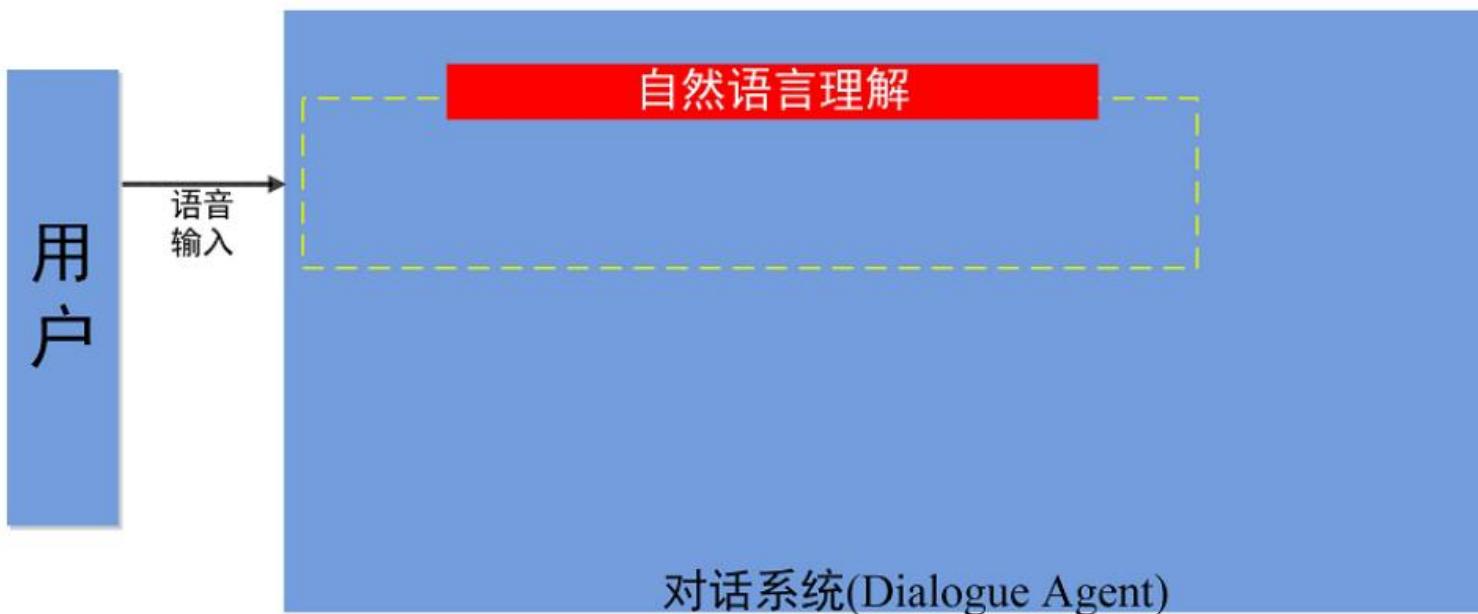
- 面向任务的人机对话

■ 目标驱动型(面向任务的)人机对话

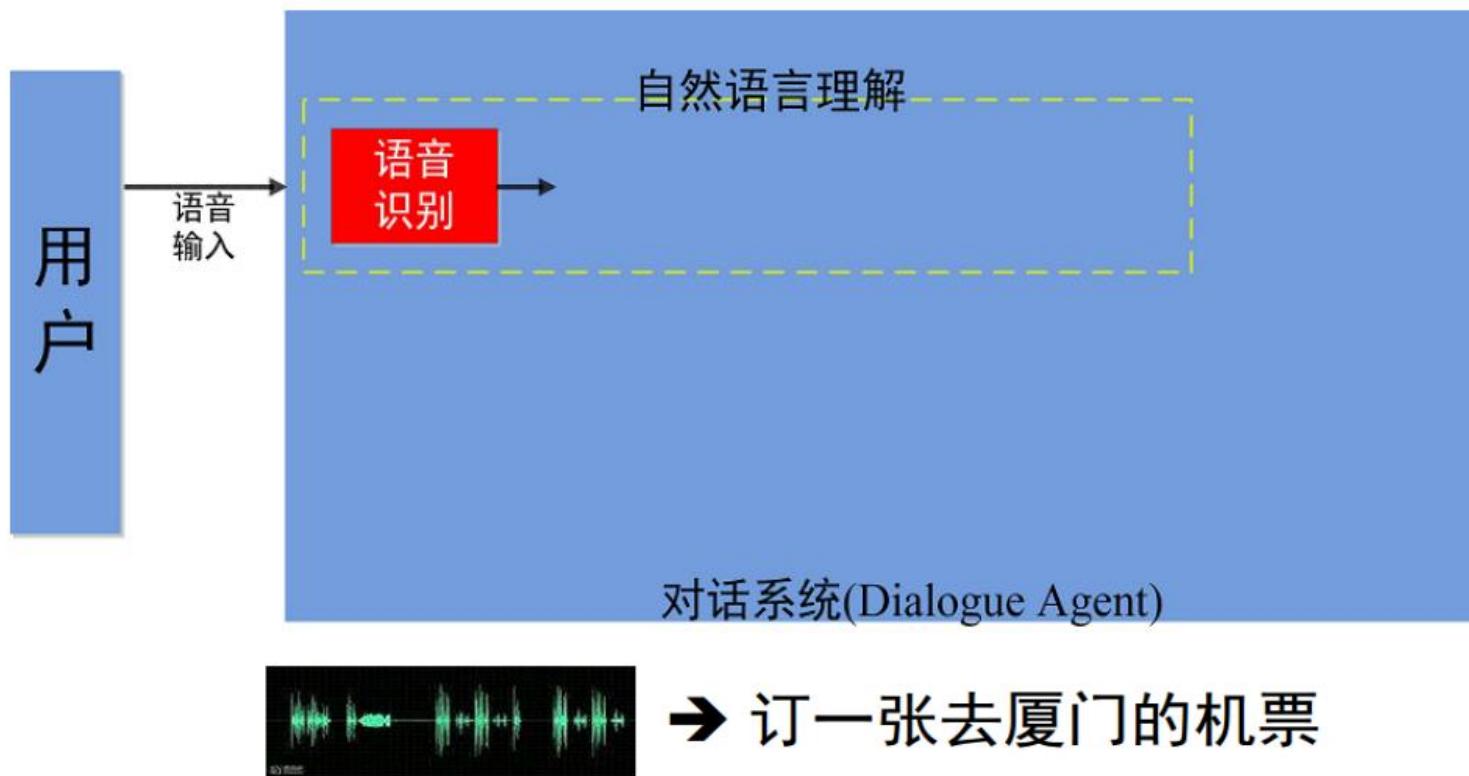
- 人类用户为达成特定目标与机器系统进行特定信息的交换
- 例如：机票预订



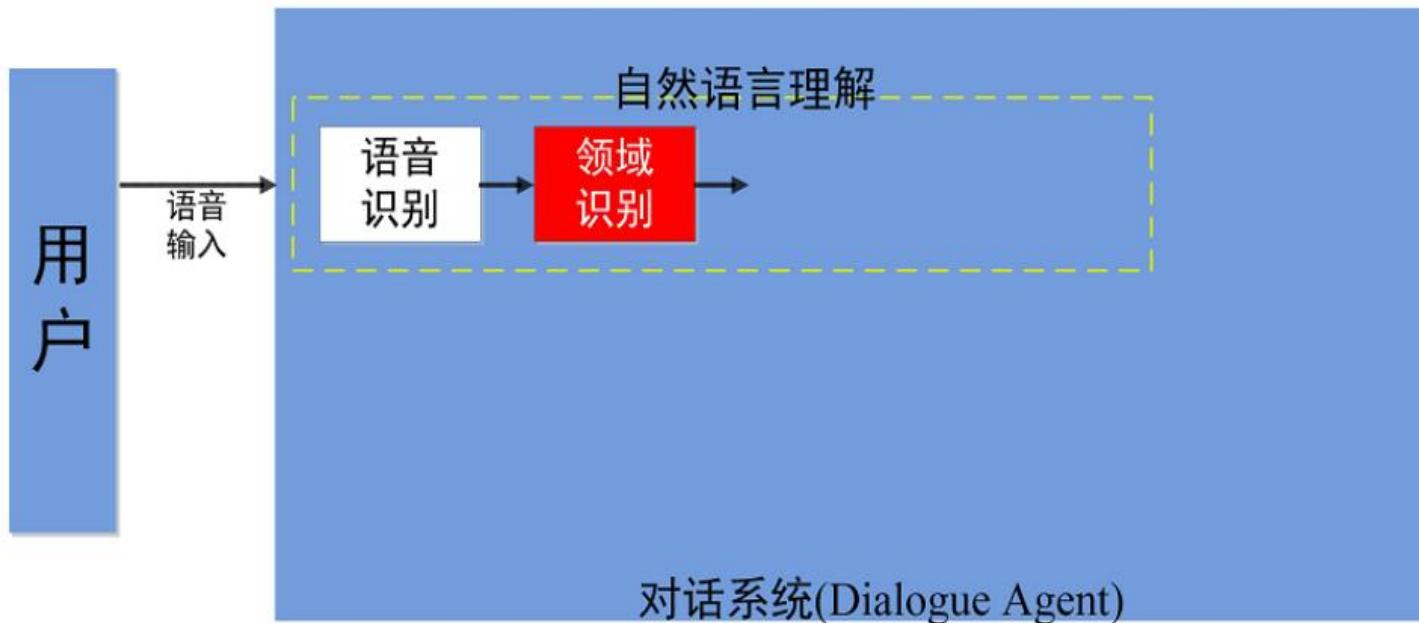
■目标驱动型人机对话的典型结构



■目标驱动型人机对话的典型结构

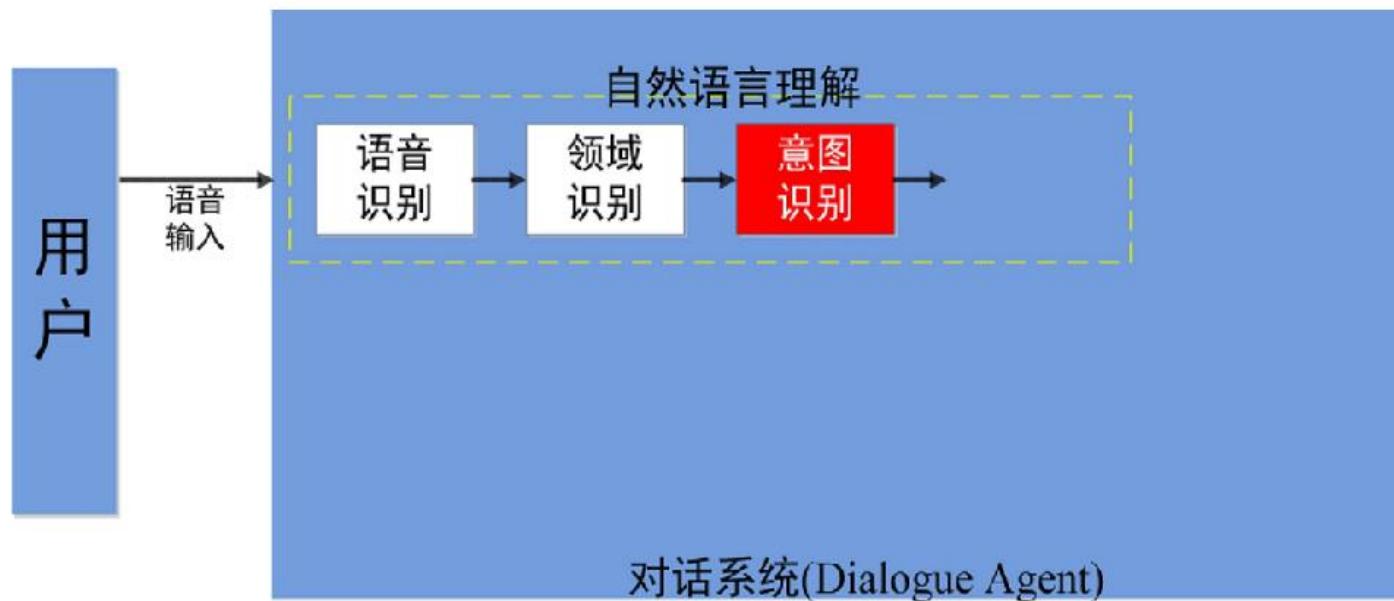


■目标驱动型人机对话的典型结构



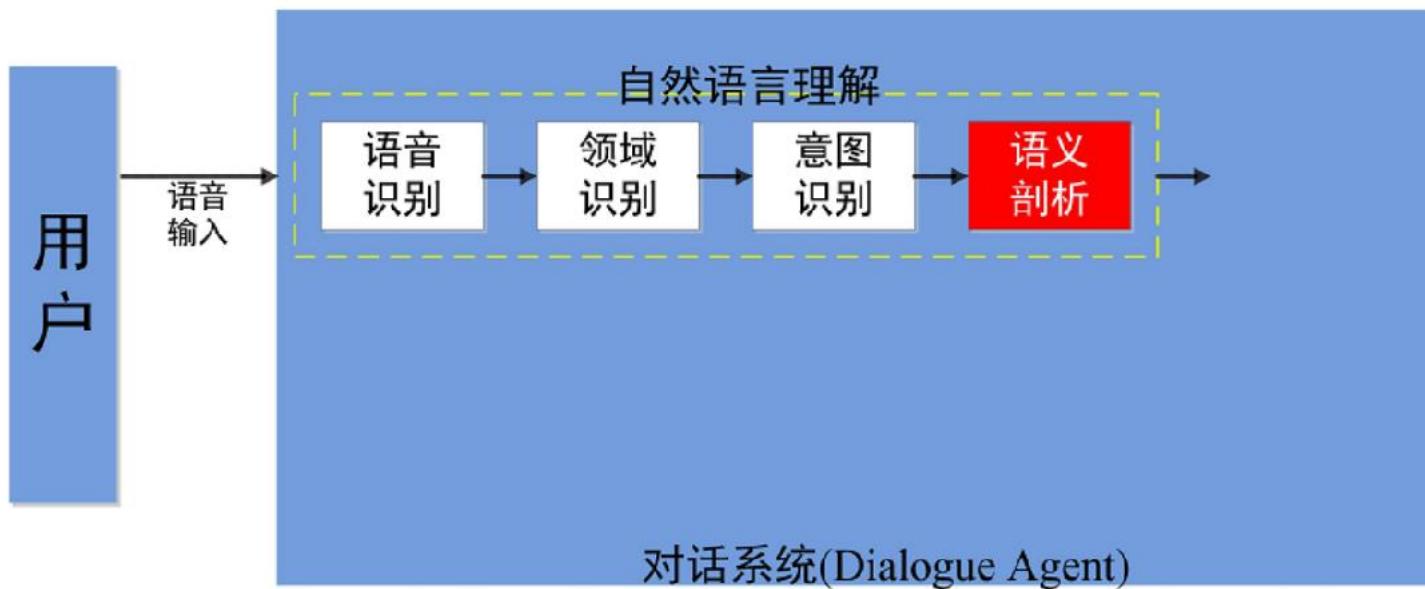
订一张去厦门的机票 → 机票业务

■ 目标驱动型人机对话的典型结构



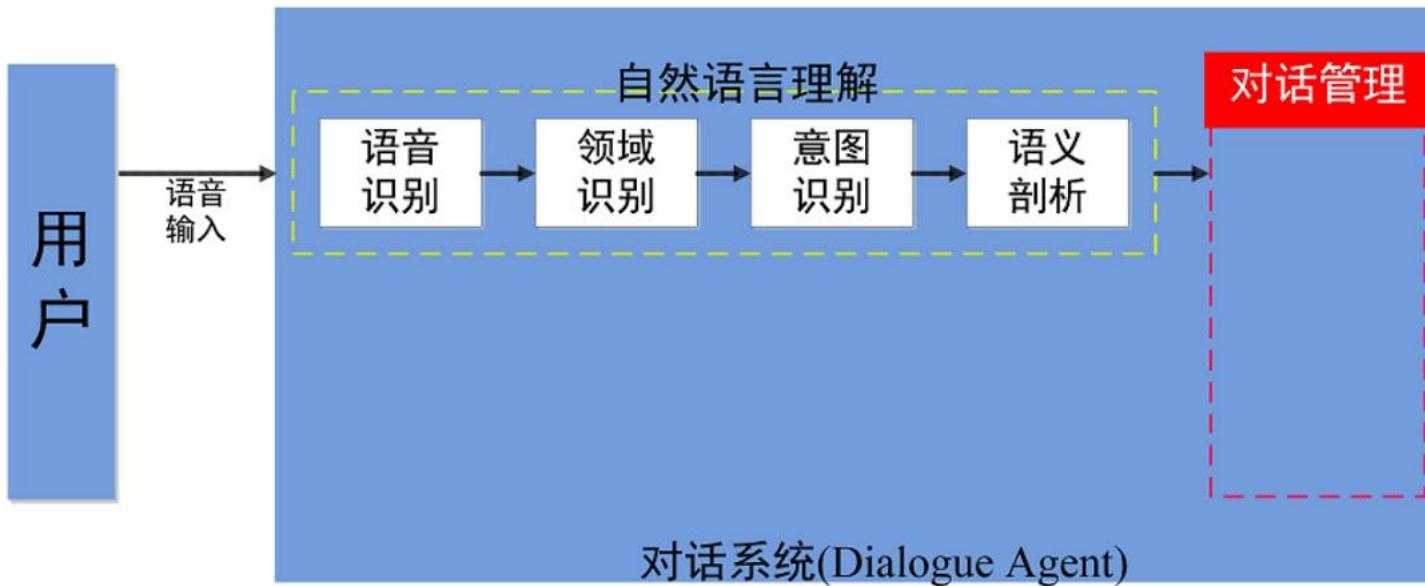
订一张去厦门的机票 → 告知

■目标驱动型人机对话的典型结构

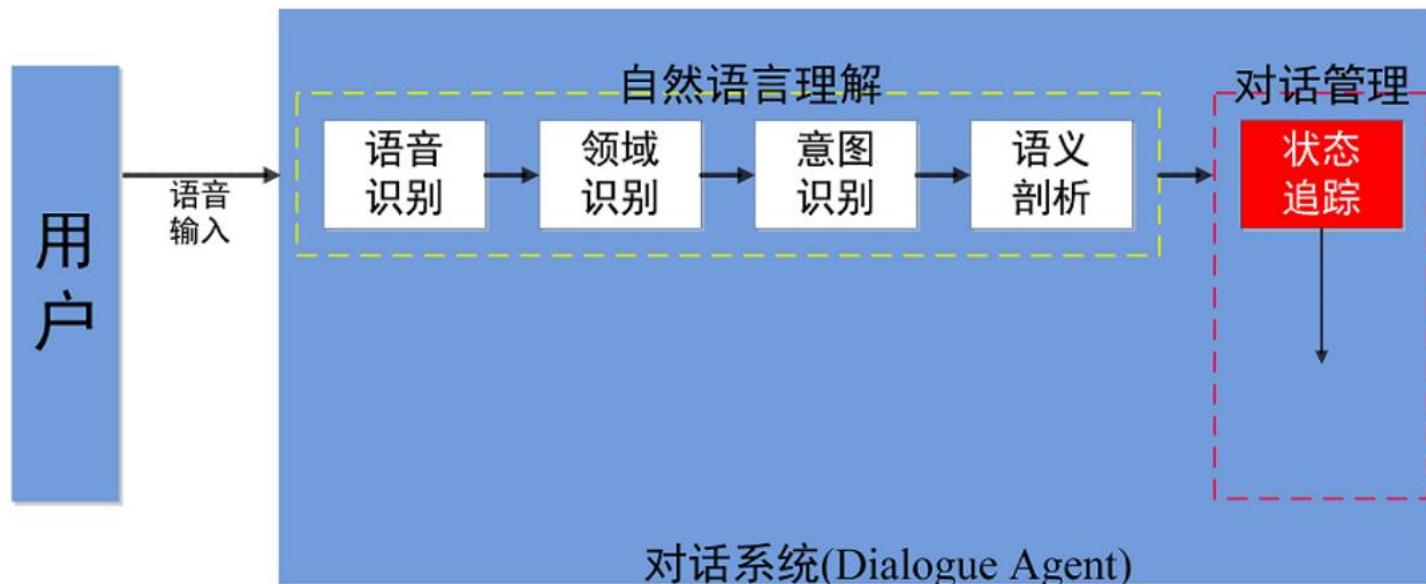


订一张去厦门的机票 ➔ 目的地：厦门

■目标驱动型人机对话的典型结构



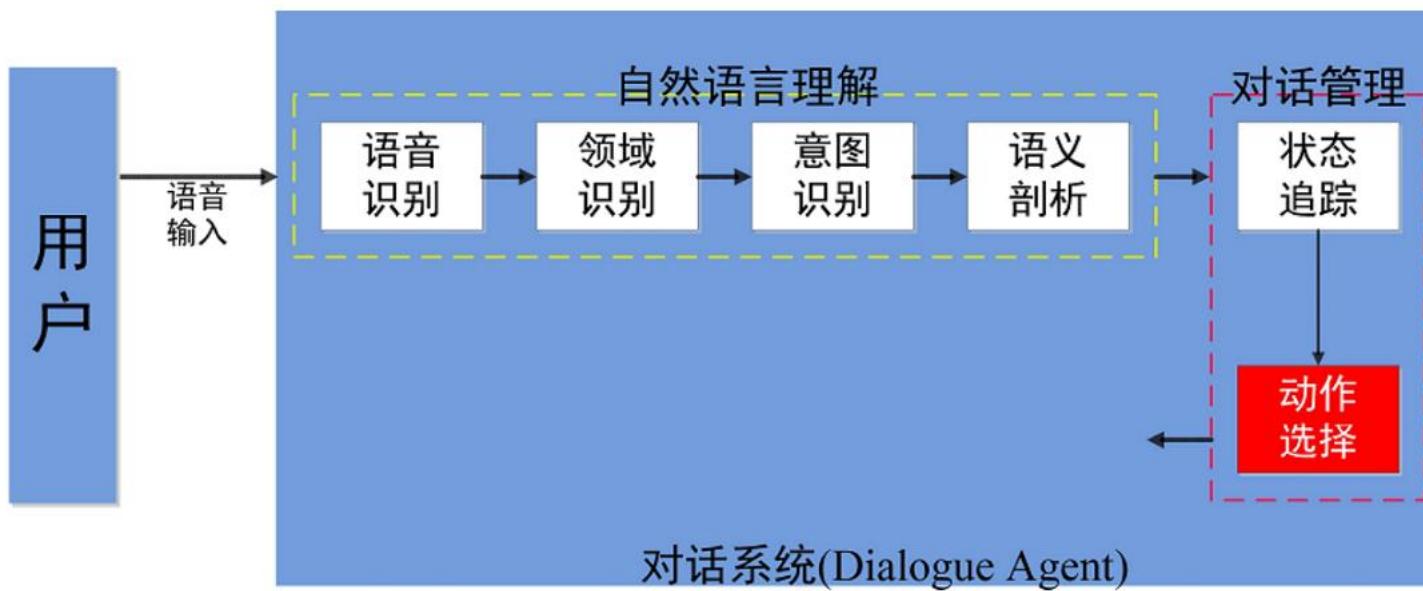
■目标驱动型人机对话的典型结构



目的地(厦门:0.6;其他:0.4); 出发时间(---) ; 出发地(---)

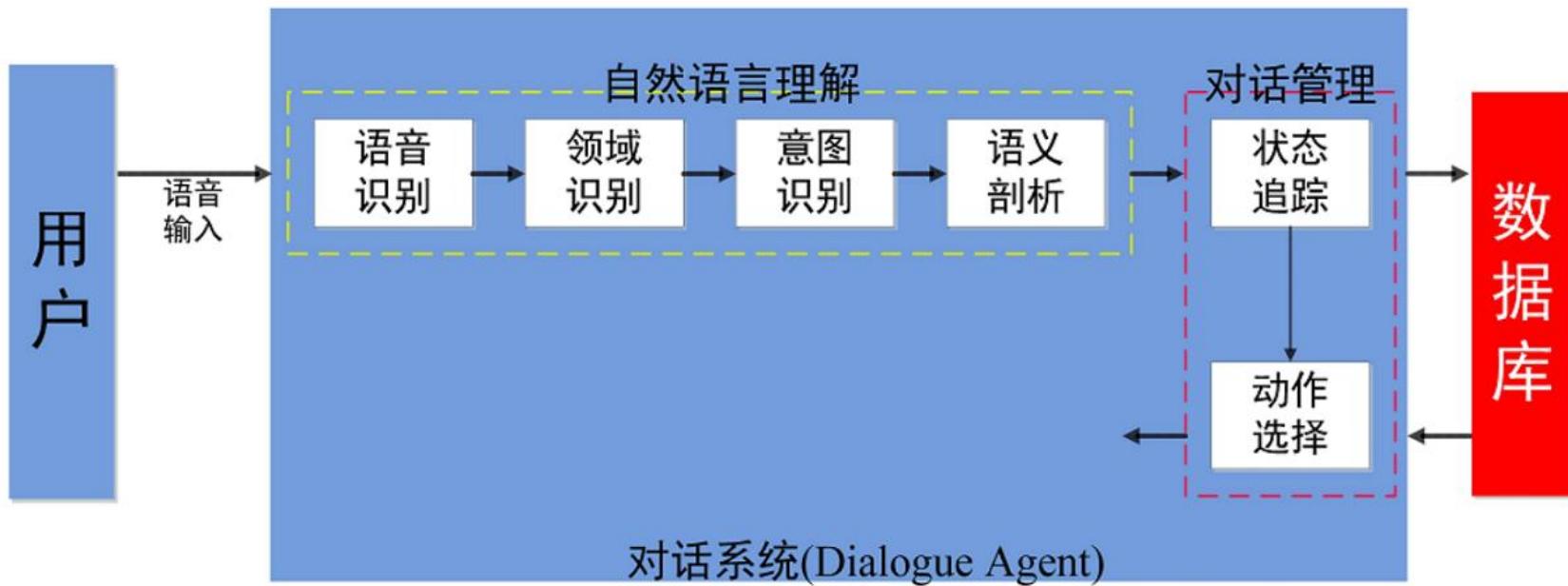
→ 目的地 (厦门:0.9;其他:0.1); 出发时间(---) ; 出发地(---)

■目标驱动型人机对话的典型结构

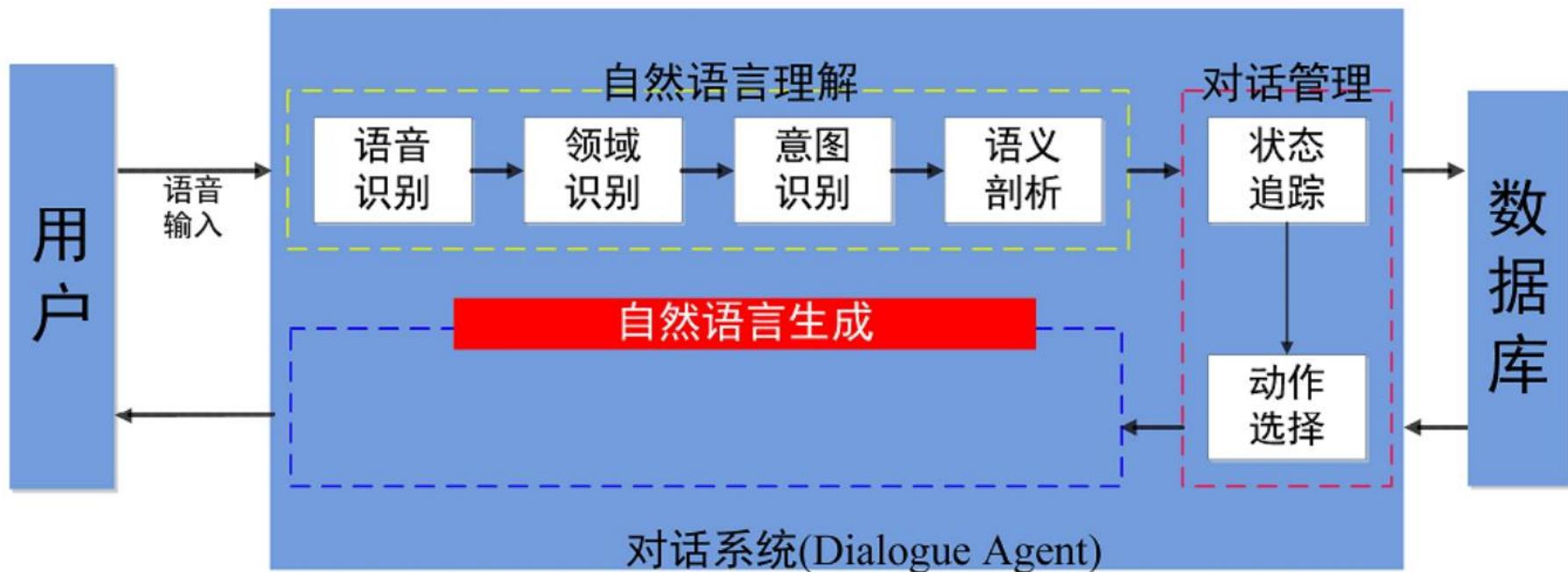


目的地(厦门:0.6;其他:0.4); 出发时间(---) ; 出发地(---)
→确认目的地是否是厦门/询问出发时间/询问出发地

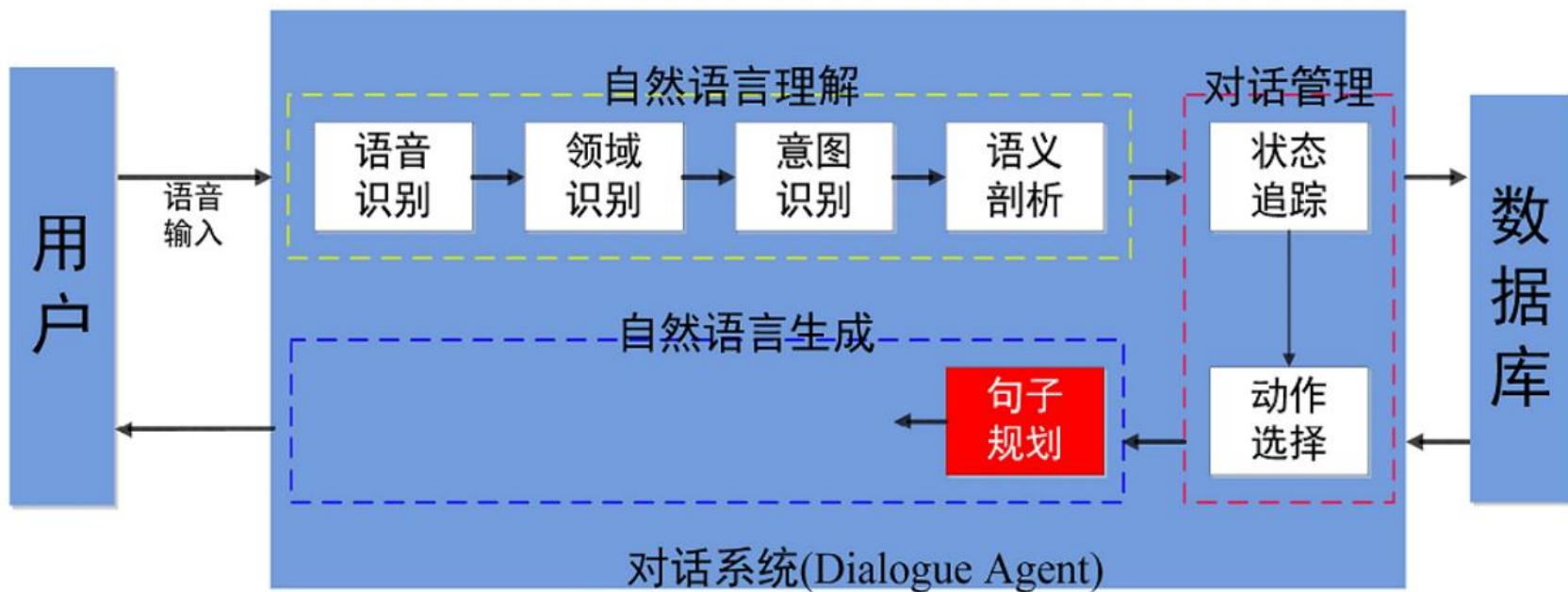
■目标驱动型人机对话的典型结构



■目标驱动型人机对话的典型结构

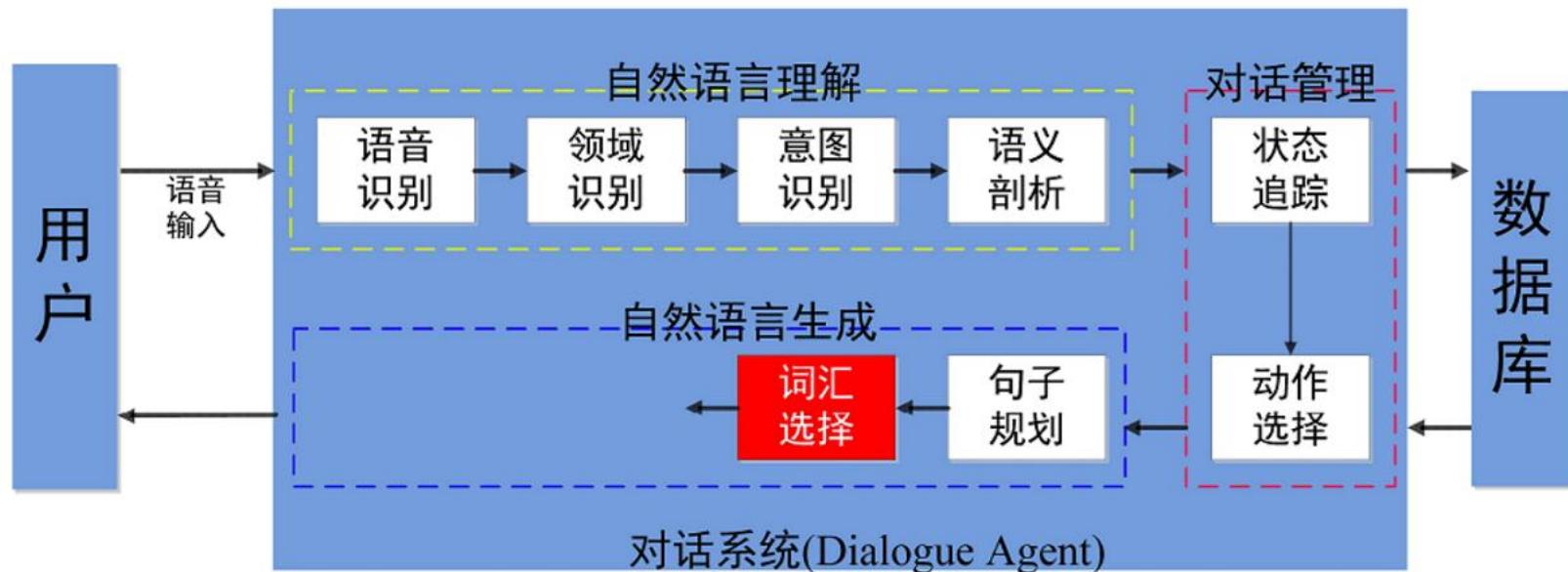


■目标驱动型人机对话的典型结构

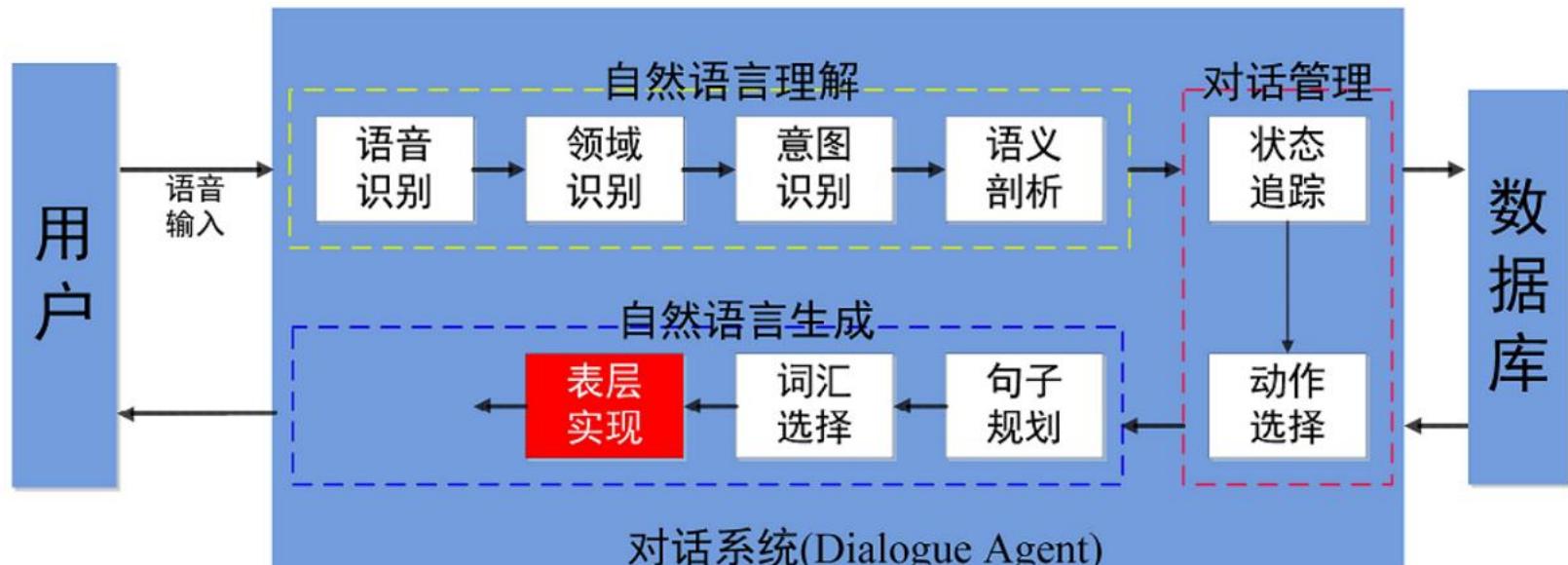


询问出发时间 → 主语+时间问句+动词

■目标驱动型人机对话的典型结构

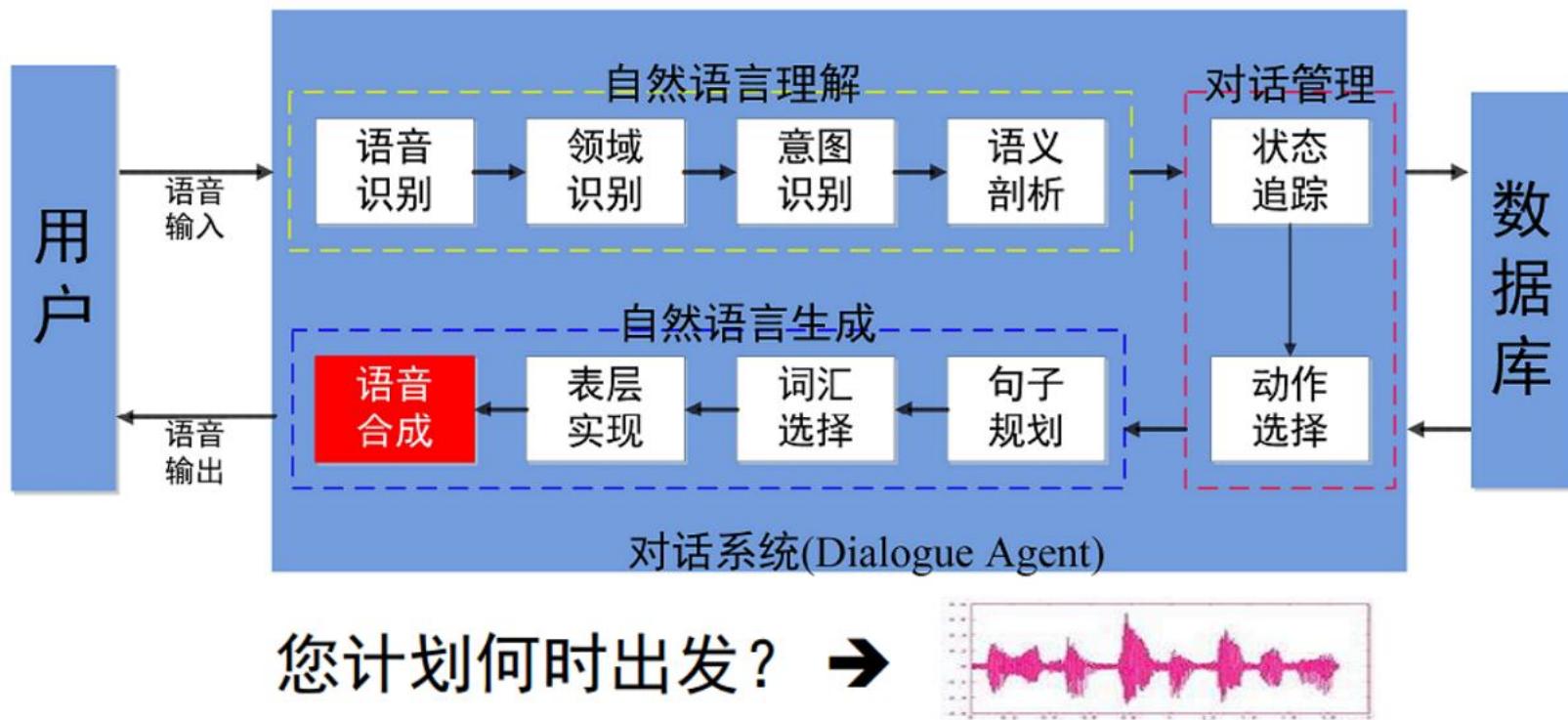


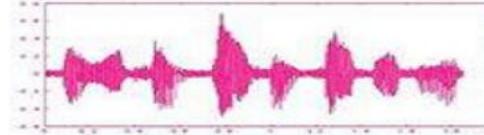
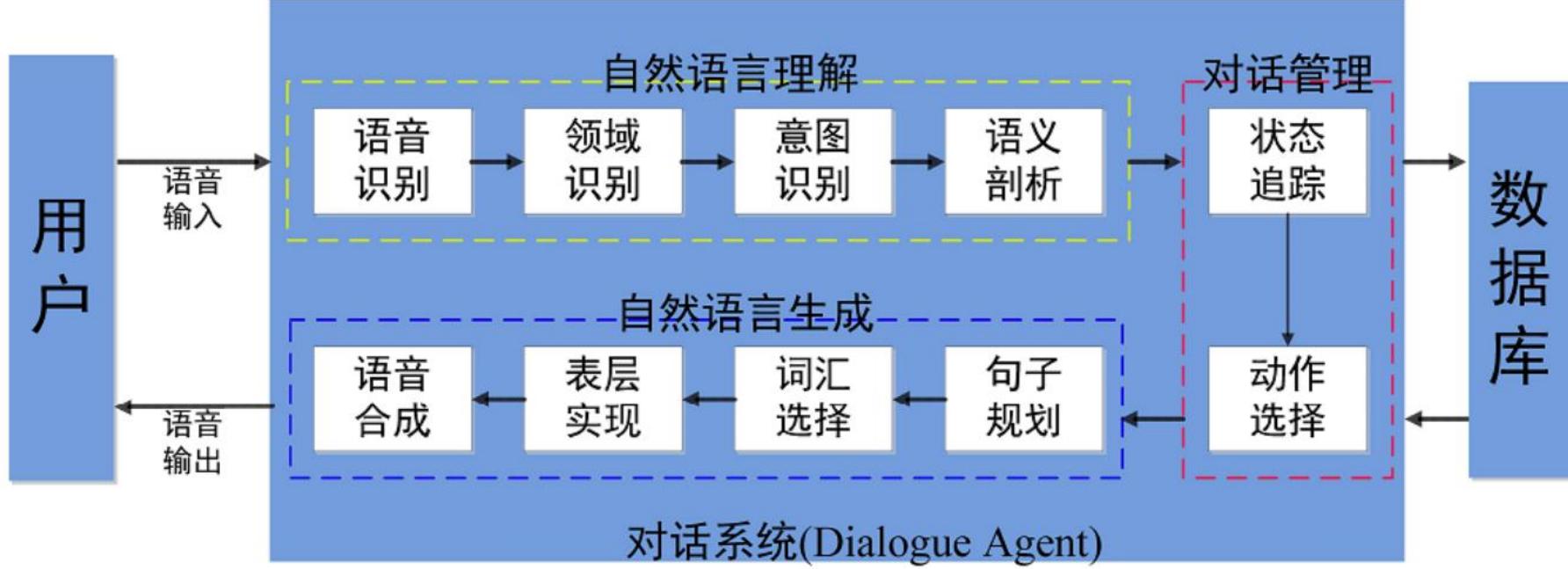
■目标驱动型人机对话的典型结构



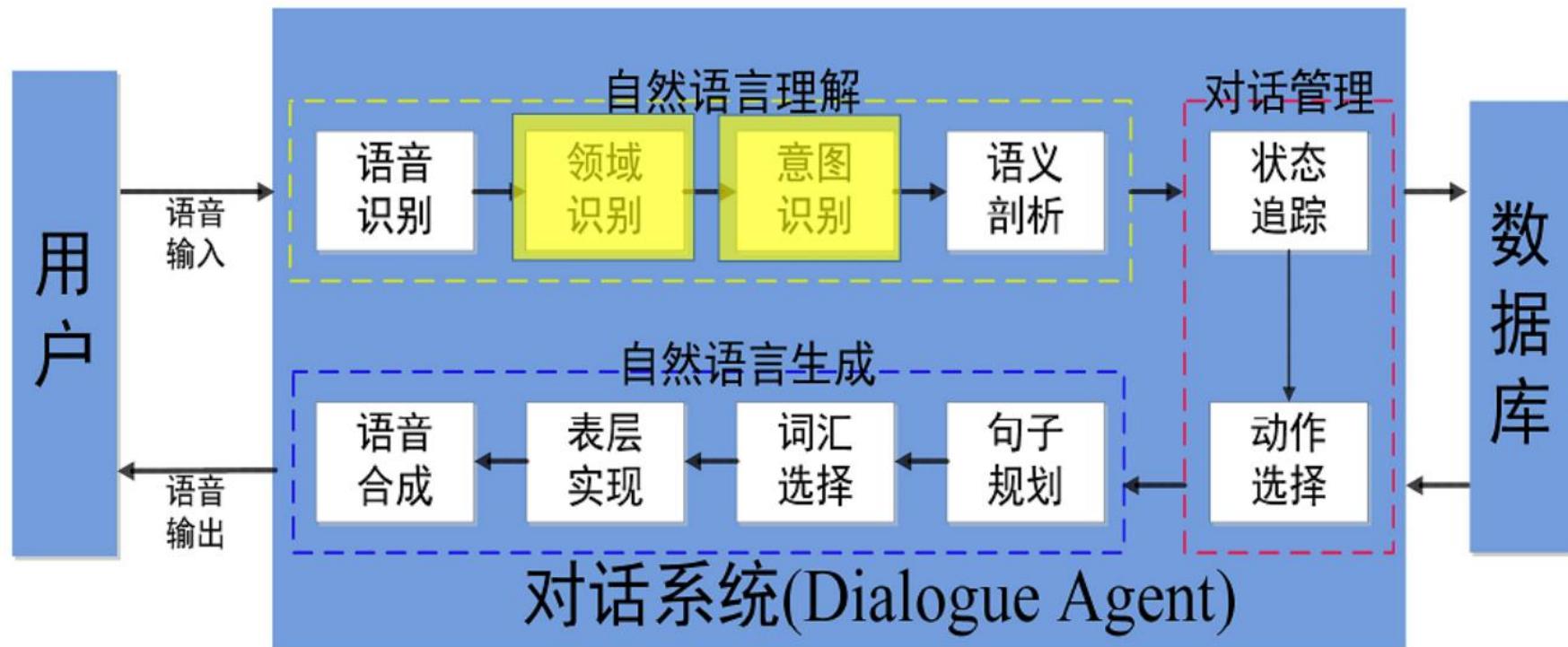
询问出发时间 → 您+计划何时+出发?

■ 目标驱动型人机对话的典型结构





各部分独立建模的研究



领域识别、意图识别：分类问题

■有标数据充分时

■作为分类问题，采用各种分类模型

作为分类问题的领域识别、意图识别

■领域识别

- Sergienko, R. B., Shan, M., Minker, W., & Semenkin, E. S. (2016). Topic Categorization Based on Collectives of Term Weighting Methods for Natural Language Call Routing Corpus description. *Journal of Siberian Federal University*, 9(2), 235–245.

-

■意图识别

- SVM : Haffner, P., Tur, G., & Wright, J. (2003). Optimizing SVMs for complex call classification. In *Proceedings of International Conference on Acoustics, Speech and Signal Processing* (pp. 1:I-632–I-635 vol.1). Hong Kong.

- MaxEnt: Ang, J., Liu, Y., & Shriberg, E. (2005). Automatic dialog act segmentation and classification. In *Proceedings of International Conference on Acoustics, Speech and Signal Processing* (pp. 1:1061–1064).

-

领域识别、意图识别：分类问题

- 有标数据充分时

- 作为分类问题，采用各种分类模型

- 有标数据不充分时

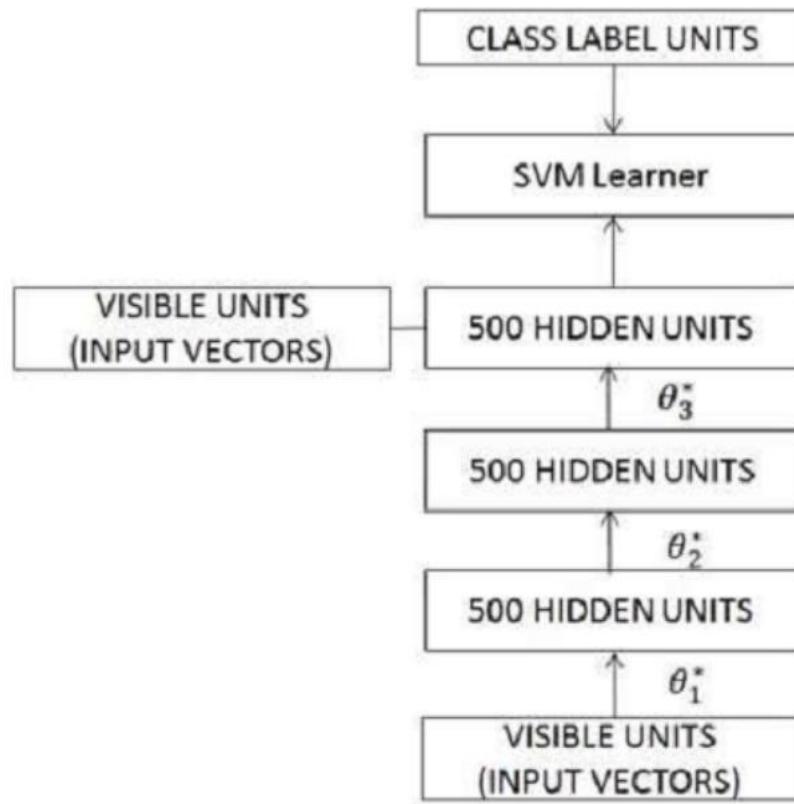
- 利用大规模无标数据

■利用无标数据的意图分类研究例：

- Sarikaya, R., Hinton, G. E., & Deoras, A. (2014). Application of Deep Belief Networks for Natural Language Understanding. *IEEE Transactions on Audio, Speech and Language Processing*, 22(4), 778–784
- DBN-1：基于无标数据预训练层叠RBMs作为FNN(在层叠RBMs上加一个softmax)的初值，再用有标数据优调FNN
- DBN-2：基于无标数据预训练层叠RBMs，得到的特征作为SVM分类器的输入，再用有标数据优调SVM
- DBN-3：结合DBN-1和DBN-2



DBN-1



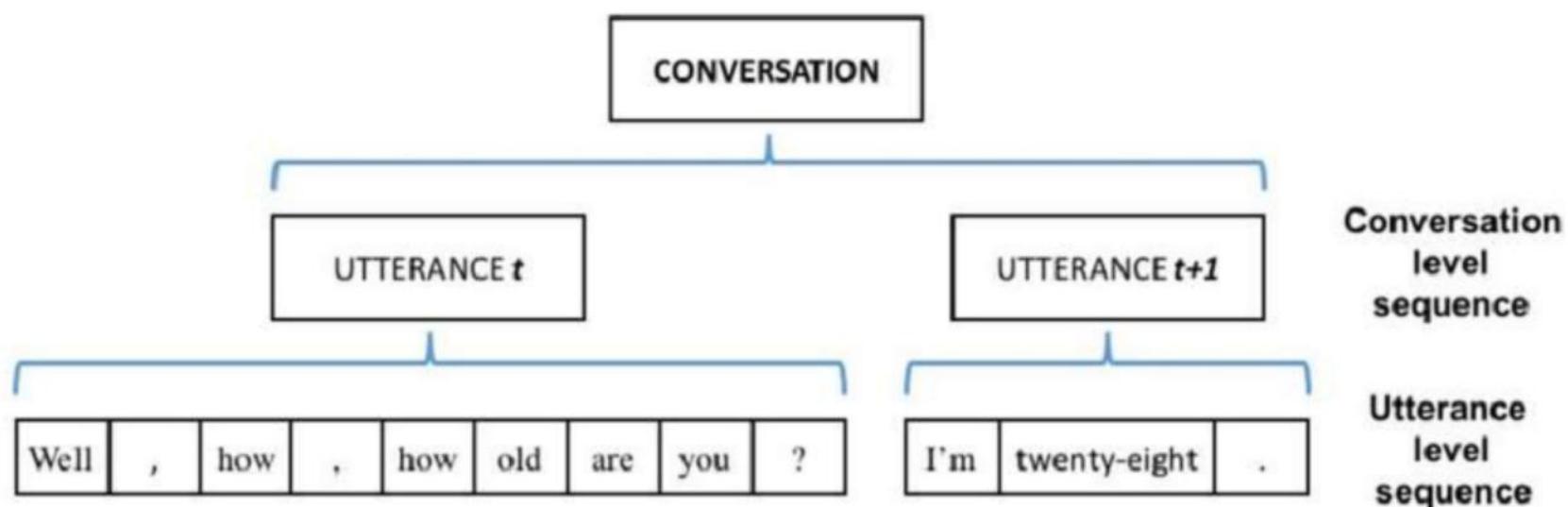
DBN-2

领域识别、意图识别：分类问题

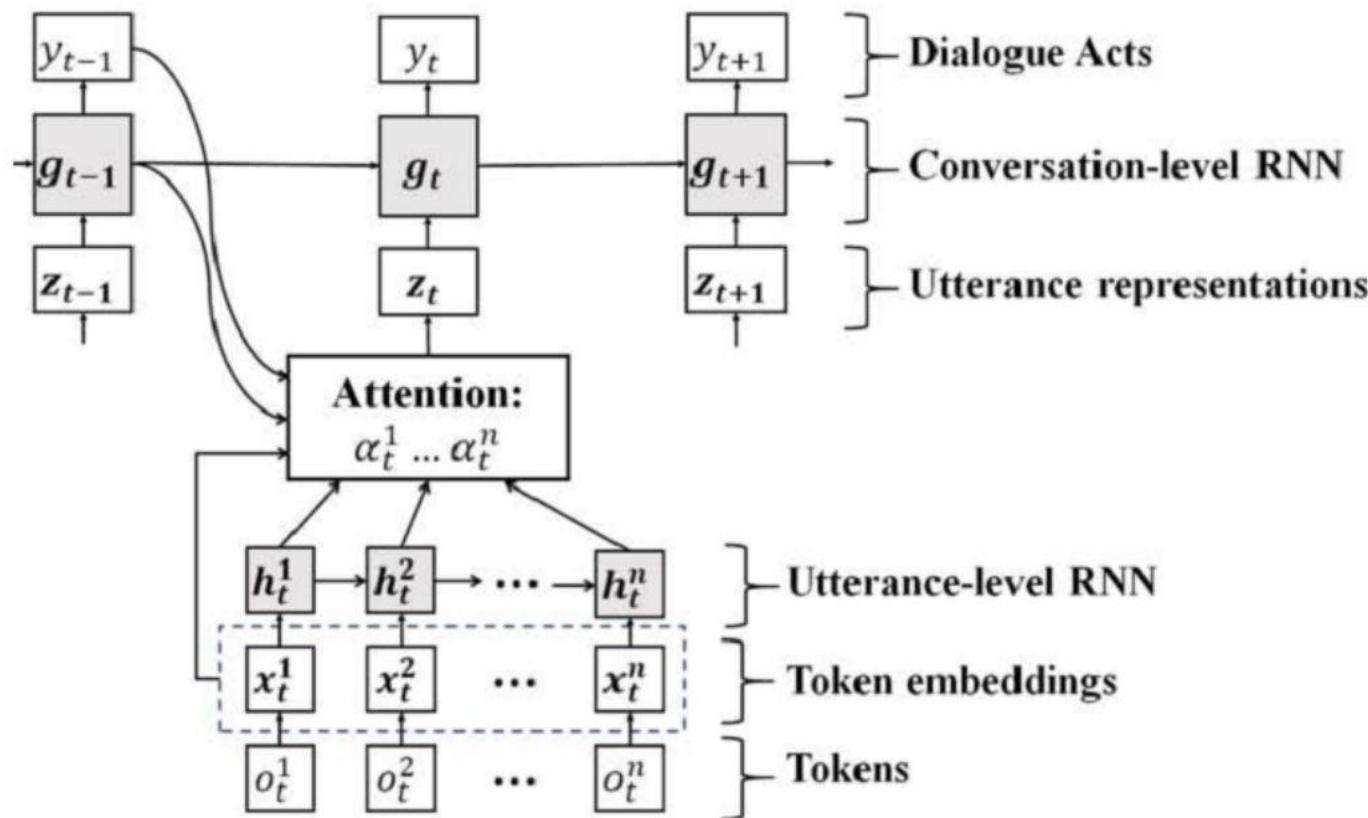
- 有标数据充分时
 - 作为分类问题，采用各种分类模型
- 有标数据不充分时
 - 利用大规模无标数据
- 结合对话语言特点
 - 对话层次结构：一个话轮→相邻对→一个对话

■建模对话语言特点的意图分类方法例：

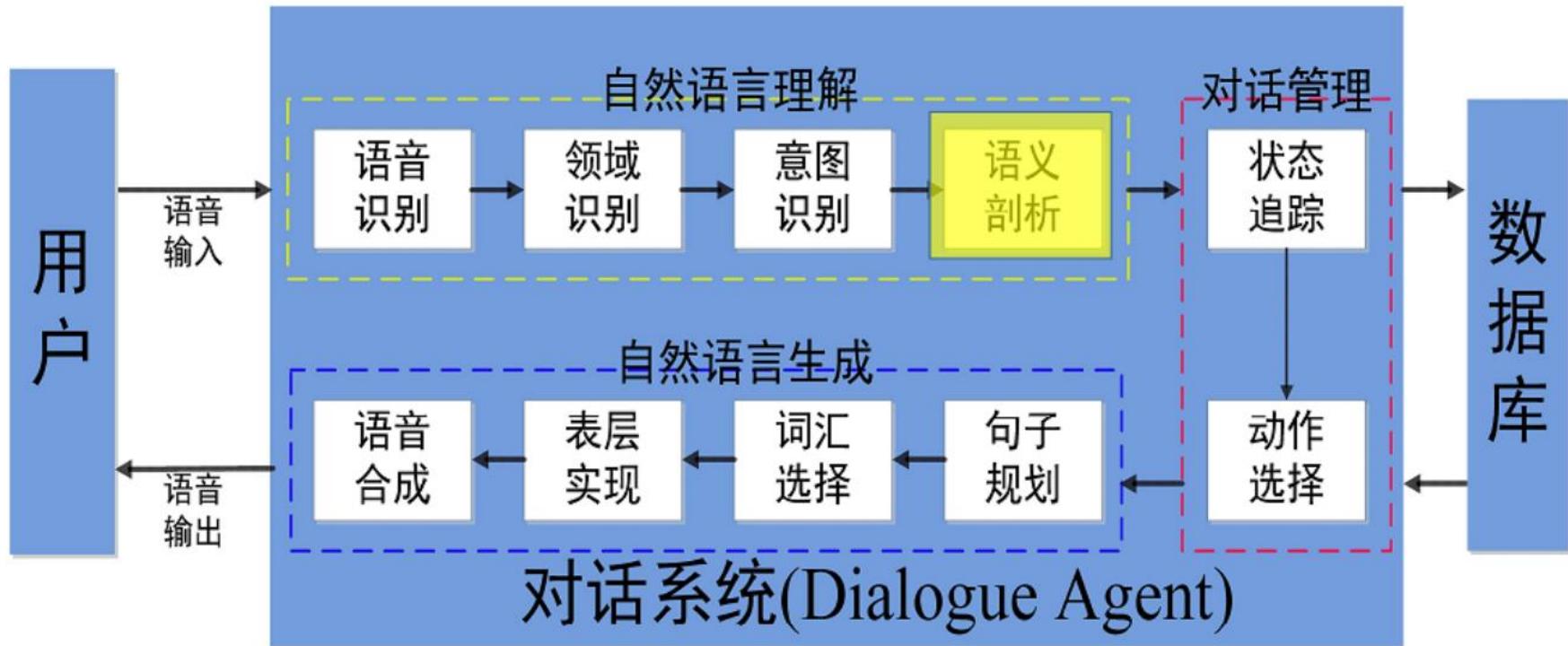
■Tran, Quan Hung, Zukerman, Ingrid and Haffari, Gholamreza. A Hierarchical Neural Model for Learning Sequences of Dialogue Acts. 2017, EACL.



■ Hierarchical-attentional RNN model



各部分独立建模的研究



语义剖析(槽填充)：分类问题/序标问题

■作为序标的情况：仅识别

		B_D	I_D		
你	去	厦	大	吗	？

■作为分类的情况：需要映射到标准描述

■Slot-D($v_1, v_2, \dots, v_i = \text{厦门大学}, \dots, v_n$)

■你去厦大吗 → Slot-D(厦门大学)

作为序标问题：数据充分时

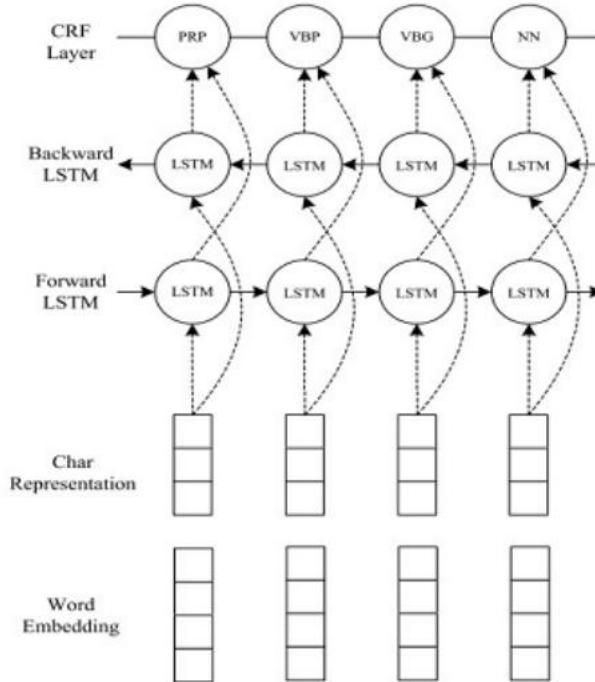
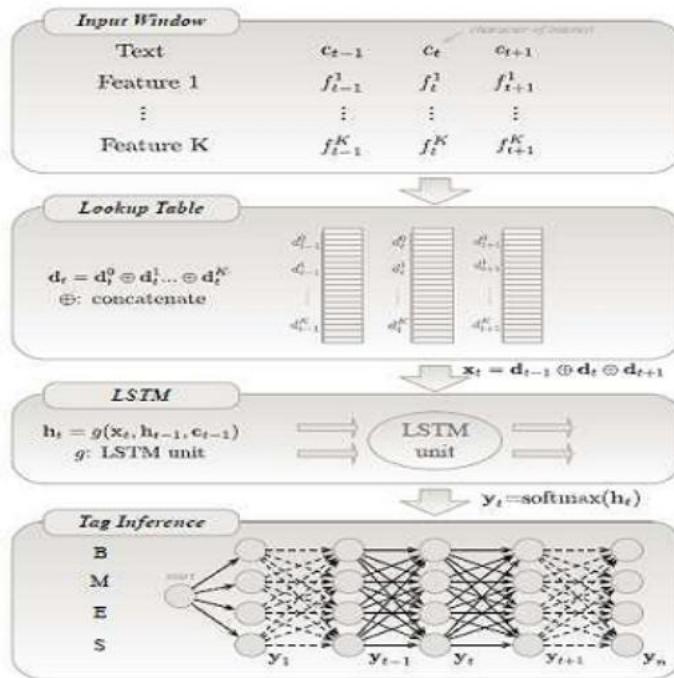
- Yao, et al(2013,2015) : RNN 在 ATIS 数据集上的性能较CRF好
- Vukotic, et al(2015) : 在 ATIS 上 RNN 较好，但是各模型的性能差别不大，而在 MEDIA 数据集上，CRF 的性能要好于 RNN。
- Mesnil et al(2015) : RNN 和 CRF 的结合比 RNN 和 CRF 显著提高。
- 结合更多对话特征的模型取得了更好的性能
 - Kurata et al(2016)...

F1-score	CRF	RNN	R-CRF
ATIS	95.16	96.29	96.46
Word+NE			
Movies	75.50	78.20	82.21
Entertainment	90.64	88.11	88.50

<i>F₁-score</i>	
RNN (Yao et al., 2013)	94.11
CNN-CRF (Xu and Sarikaya, 2013)	94.35
Bi-directional RNN (Mesnil et al., 2015)	94.73
LSTM (Yao et al., 2014a)	94.85
RNN-SOP (Liu and Lane, 2015)	94.89
Hybrid RNN (Mesnil et al., 2015)	95.06
Deep LSTM (Yao et al., 2014a)	95.08
RNN-EM (Peng and Yao, 2015)	95.25
Encoder-labeler LSTM(W)	95.40
Encoder-labeler Deep LSTM(W)	95.66

作为序标问题：数据充分时

- 更多进一步结合RNN(CNN)与HMM(CRF)模型
- Zhou et al(2016), Ma et al (2016).....

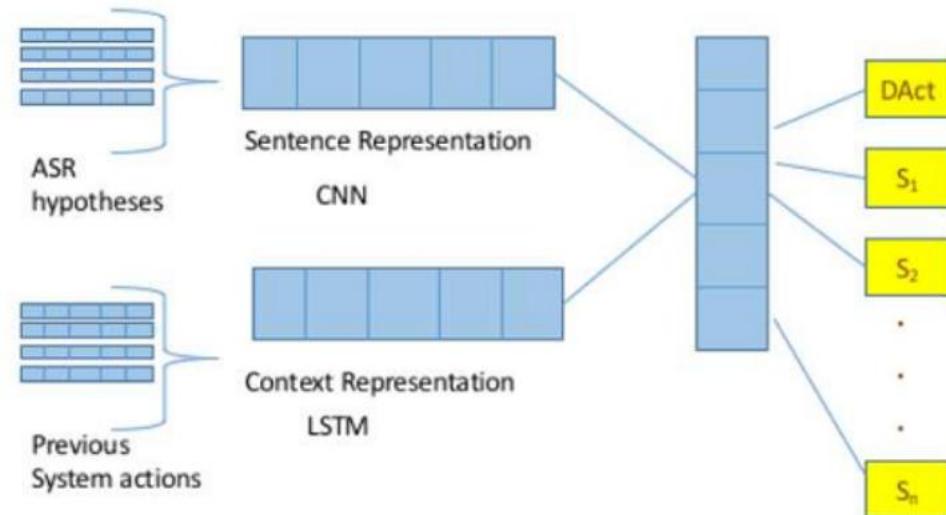


作为序标问题：数据不充分时

- 利用其它资源的半监督\无监督方案：
 - Tur et al (2011)利用任务相关的查询日志；
 - Tur et al(2013)利用LDA模型获取主题；
 - Chen et al(2014)使用Frame-Net；
 - Chen et al(2015)利用slot-based语义图、矩阵分解；
 -
- 在跨领域、冷启动等情况下数据不充分的问题更为严重。
 - Hartmann et al(2017) Out-of-domain FrameNet SRL
 -

作为分类问题：数据充分时

- 分类器+对话特征
- 例：Lina et al(2016)
 - 当前句子表示(CNN)：基于ASR的N-best结果
 - 历史对话综合(LSTM)：以系统的历史动作为输入



Corpus	Model	F1	ICE
DSTC2	SLU1 (Williams, 2014)	80.2%	1.943
	CNN+LSTM_w4	83.59%	0.758
In-car	WNGRAMS (Henderson et al., 2012)	70.8%	1.76
	WNGRAMS+Ctxt (Henderson et al., 2012)	74.2%	1.497
	CNN+LSTM_w4	73.06%	1.106

作为分类问题：数据不充分时

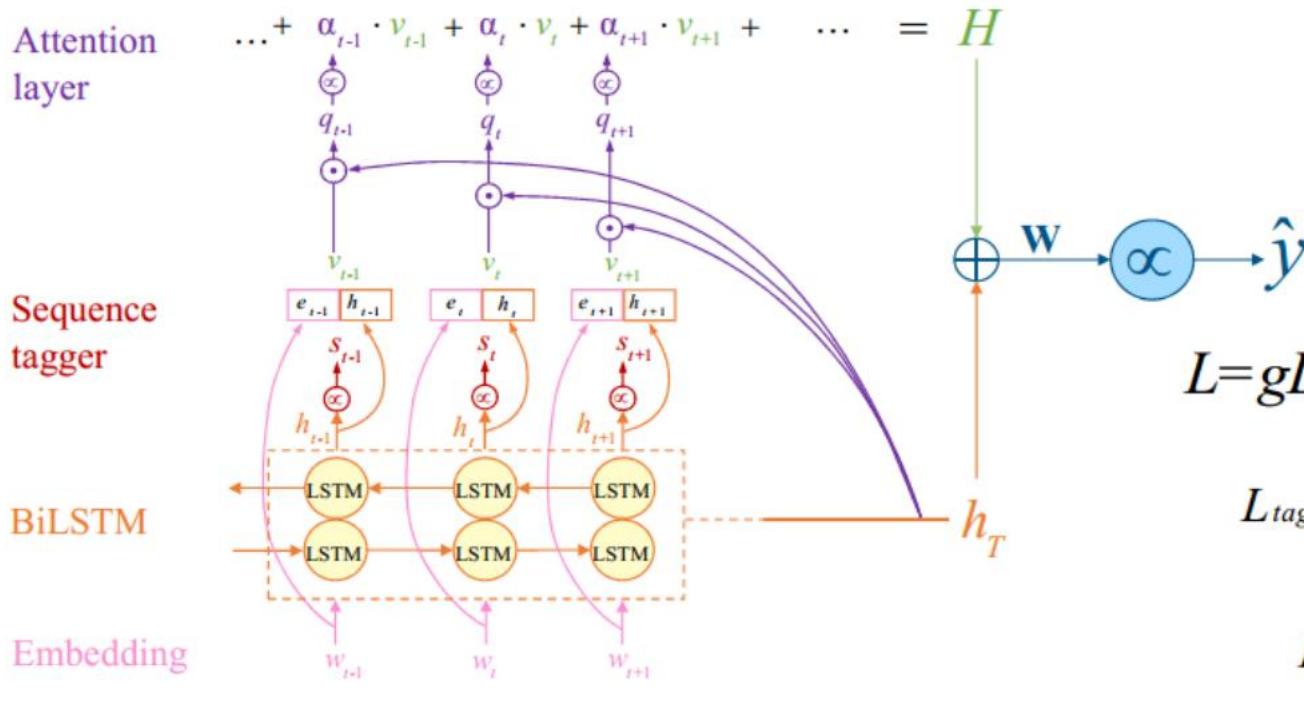
- 小样本的分类：和前面意图识别类似
- 限定目标、利用其它资源
- 极端情况：新槽植问题 →

新槽植问题

- 在一些对话任务中会预定义一个标准槽值集合，对话中的任意槽值都需要映射到标准槽值(以便进行数据库操作)，此时，当对话中出现新槽值时(标准槽值之外的该槽的值)，已有方法出现问题：
 - 分类方法：因为没有新槽值的训练数据，不能识别出新槽植
 - 序标方法：可以识别出新槽值，但是不能将其与标准槽值关联
- 通常的做法：先序标识别出槽值，再利用语义相似性进行新旧槽值判定(Yeh2014)
- 问题：误差累积、前者不能利用后者的信息

Yeh P Z, Douglas B, Jarrold W, et al. A speech-driven second screen application for TV program discovery[C]// Twenty-Eighth AAAI Conference on Artificial Intelligence. AAAI Press, 2014:3010-3016.

■基于负采样的新槽值联合识别模型



$$L = g L_{tagging} + (1 - g) L_{classification}$$

$$L_{tagging} = \frac{1}{N} \sum_i^N \frac{1}{T_i} \sum_t^T L(\hat{s}_t^i, s_t^i)$$

$$L_{classification} = \frac{1}{N} \sum_i^N L(\hat{y}_i, y_i)$$

HOU M., WANG X, et al. Attention Based Joint Model with Negative Sampling for New Slot Values Recognition, IWSDS2018, Singapore.

■面向新槽值的负采样

■获取词分布： $U(word) = \frac{count(word)}{|Data|}$

■获取基于旧槽值的模板： Can I have a **video chat** on my phone?

■基于词分布从词表中采样填充模板

	我	的	手	机	可	以	边	打	电	话	边	视	频	吗	？	新值：	
	O	O	O	O	O	O	B-func	I-func	O	O	视频通话						
negative sampling:	无	线	充	电													
	我	的	手	机	可	以	无	线	充	电					吗	？	新值的负例：
	O	O	O	O	O	O	B-func	I-func	I-func	I-func					O	O	无线充电

■评估

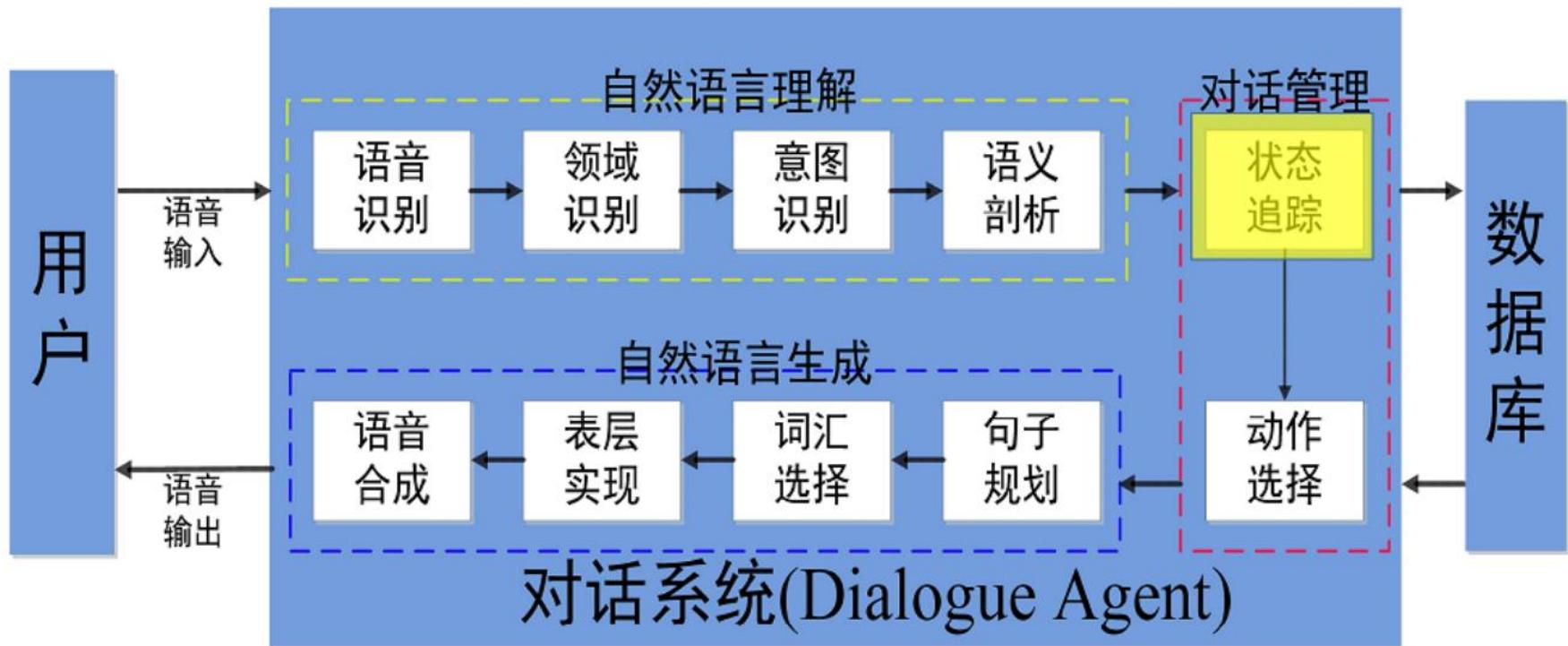
	DSTC				Service			
	all	NEW	OLD	NULL	all	NEW	OLD	NULL
LSTM_FM	0.8491	0.3063	0.9670	0.8923	0.8981	0.5693	0.9320	0.5693
LSTM_C	0.8290	0.0000	0.9249	0.9761	0.9210	0.0000	0.9720	0.9643
AJM_NS	0.9632	0.8621	0.9738	0.9822	0.9749	0.7759	0.9881	0.9633

F1 scores of classification

	DSTC			Service		
	all	NEW	OLD	all	NEW	OLD
LSTM_FM	0.8546	0.2363	0.9837	0.8850	0.2615	0.9269
AJM_NS	0.9024	0.5684	0.9946	0.9132	0.3399	0.9673

F1 scores of sequence labeling

各部分独立建模的研究



对话状态追踪

- 如果NLU部分可以得到完全可靠
的确定性结果：状态完全可观
 - 状态追踪是平凡任务
 - $(V_{S1}, V_{S2}, \dots, V_{Sn})$

对话状态追踪

■如果NLU部分可以得到完全可靠
的确定性结果：状态完全可观

■状态追踪是平凡任务

■ $(V_{S1}, V_{S2}, \dots, V_{Sn})$

■但语音识别、语义剖析都可能存在不确定性：状态不完全可观

■状态→信念状态(槽值上的分布)

■需要结合对话状态、NLU输出以及系统动作等多种信息

■ $(D_{S1}, D_{S2}, \dots, D_{Sn})$

系统话语	用户话语	NLU输出	追踪输出
是要问哪一路车？	61C路	67C=0.4 61Z=0.3 61C=0.3	67C 61Z 61C 其他
是61C路吗？	是的	Yes=0.9 No=0.1	67C 61Z 61C 其他

对话状态追踪

■DSTC (Dialogue State Tracking Challenge) (Williams et al 2013)对于DST具有重要贡献，目前已经进行了DSTC6：通过标注数据、设定任务、实施评估，推动技术方法的发展。

- 规则方法：手工定义、易于融合领域知识
- 统计方法：数据驱动，生成、判别模型
- 结合模型：统计+规则

[Williams et al 2013] Williams J , Raux A, Ramachandran D, Black A. The Dialog State Tracking Challenge[J]. In Proceedings of the SIGDIAL 2013 Conference, 2013, 404-413.

对话状态追踪

■判别模型

- 生成模型依赖假设复杂、参数估计困难、分解单元不一致等弱点(Williams2012)
- 判别模型：基于条件概率 $b(s) = P(s|f)$ 来计算信念分布
 - f 是从ASR、NLU、以及对话历史中抽取的特征
- 首个判别模型：结构判别模型(Lee2013)，基于条件随机场建模

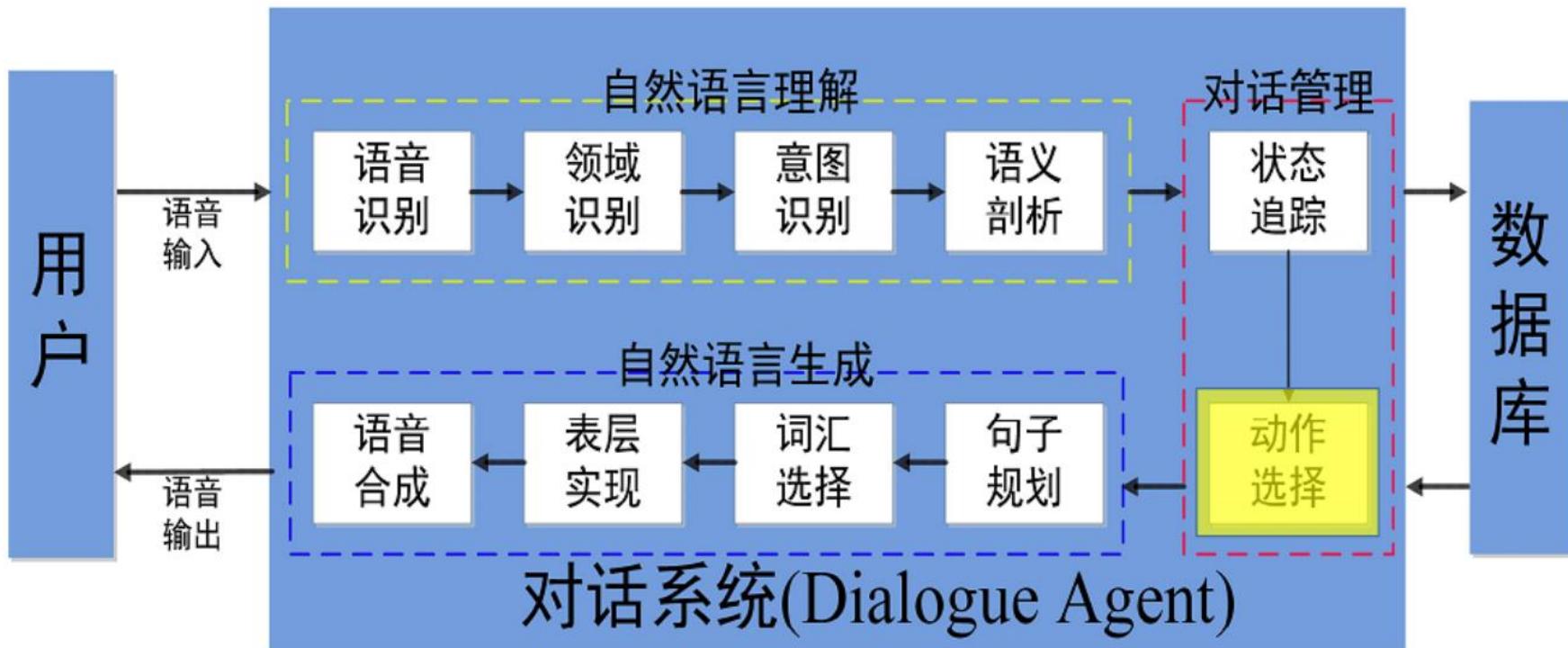
[Williams2012]Williams, Jason D. Challenges and opportunities for state tracking in statistical spoken dialog systems : results from two public deployments. Journal of Selected Topics in Signal Processing. 2012, 6(8):959-970.
[Lee2013]Sungjin Lee. Structured Discriminative Model For Dialog State Tracking[J]. Dialogue & Discourse, SIGDIAL 2014, August:442-451.

各部分独立建模的研究

■ 状态追踪

- Thomson B, Young S. Bayesian update of dialogue state: A POMDP framework for spoken dialogue systems[J]. Computer Speech & Language, 2010, 24(4):562-588.
- Williams J, Raux A, Ramachandran D, Black A. The Dialog State Tracking Challenge[J]. In Proceedings of the SIGDIAL 2013 Conference, 2013, 404-413.
- [9] Henderson M, Thomson B, Williams J. The Second Dialog State Tracking Challenge[J]. In Proceedings of the 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL), 2014, 263-272.
- [10] Henderson M, Thomson B, Williams J D. The third Dialog State Tracking Challenge[C]// Spoken Language Technology Workshop. IEEE, 2014.
- Henderson M, Thomson B, Young S. Word-Based Dialog State Tracking with Recurrent Neural Networks[C]// Meeting of the Special Interest Group on Discourse and Dialogue. 2014:525-539.
- Kim S, D'Haro L F, Banchs R E, Williams J, Henderson M. The Fourth Dialog State Tracking Challenge[J]. In Proceedings of the 7th International Workshop on Spoken Dialogue Systems (IWSDS), 2016.
- Williams J, Raux A, Henderson M. The Dialog State Tracking Challenge Series: A Review[J]. Dialogue & Discourse, 2016, 7(3):4-33.
- Asri, L. El, Larocque, R., & Pietquin, O. (2017). Compact and Interpretable Dialogue State Representation with Genetic Sparse Distributed Memory. Dialogues with Social Robots.
- Perez, J., Liu, F., & Europe, C. (2017). Dialog state tracking , a machine reading approach using Memory Network. In EACL (Vol. 1, pp. 305–314)
-

各部分独立建模的研究

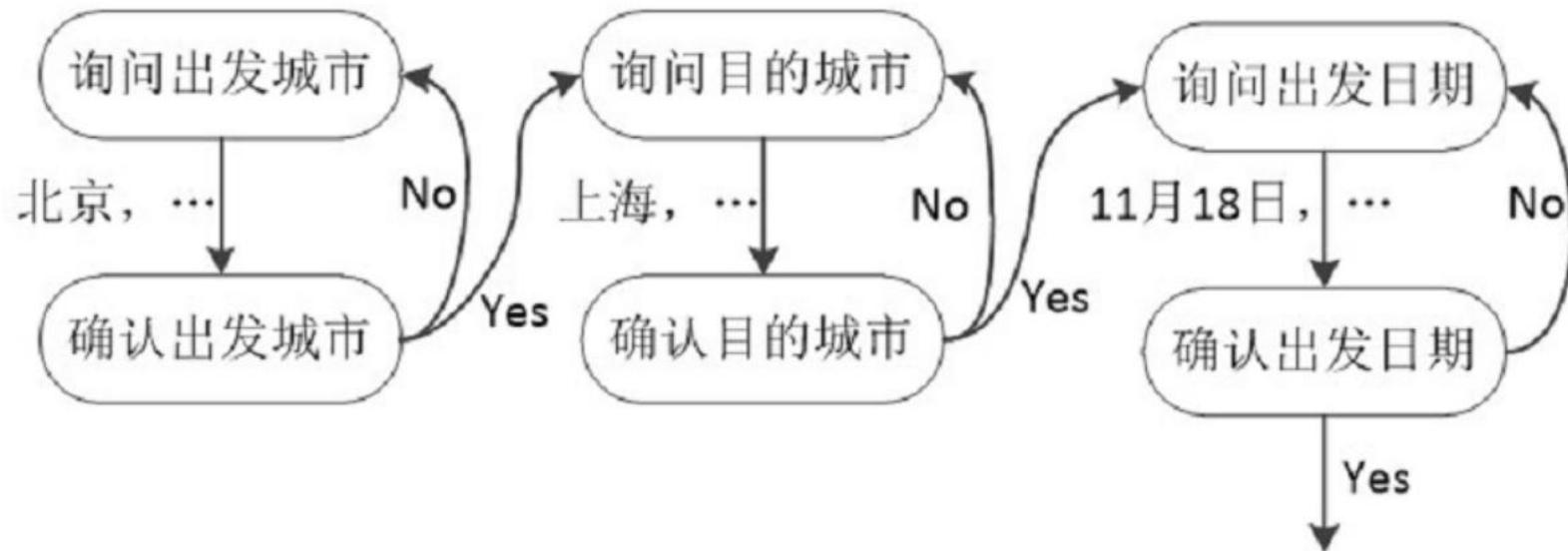


动作选择(策略模型)

- 确定性对话状态下的策略模型
- 不确定性对话状态下的策略模型

动作选择(策略模型)

■有限状态自动机(FSA)



动作选择(策略模型)

- 有限状态自动机(FSA)

- 系统主导对话过程、对话按预定策略进行

- 不灵活、不丰富、不容错

- 更灵活的对话控制方式?

- 框架模型、信息状态更新模型、规划模型...

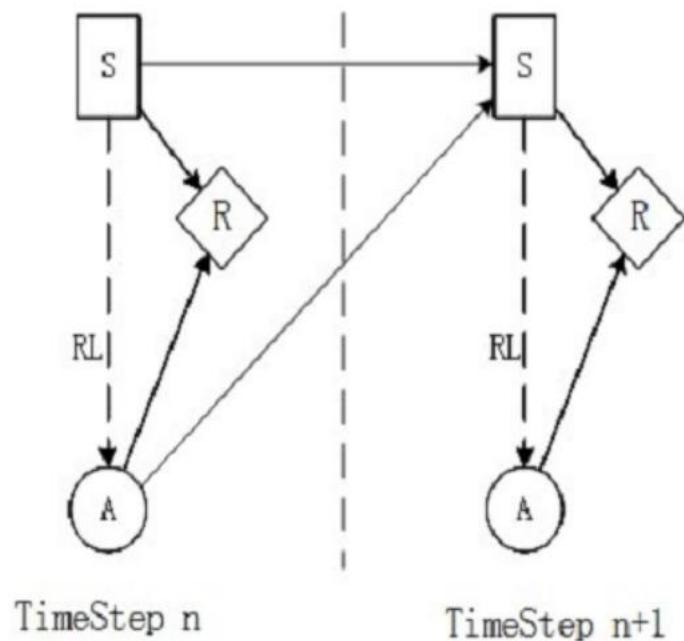
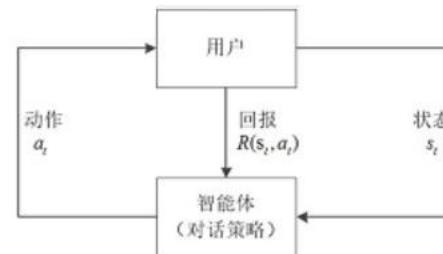
- 更丰富的、容错的对话策略?

- 数据驱动的统计模型：马尔科夫决策过程

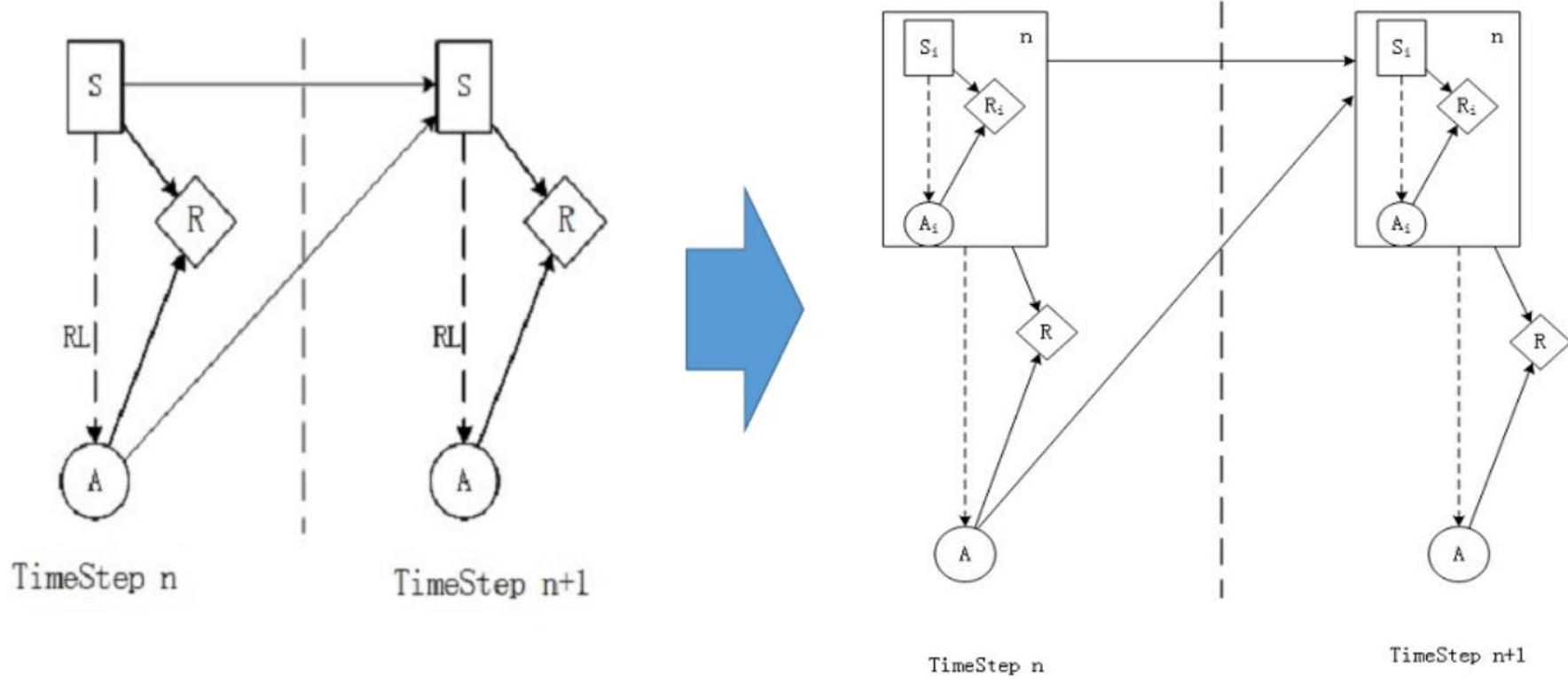
动作选择(策略模型)

■ 马尔科夫决策过程(MDP)：四元组
 $(\mathcal{S}, \mathcal{A}, \mathcal{T}, \mathcal{R})$

- 状态(state)集 \mathcal{S}^t
- 行为(action)集 \mathcal{A}
- 转移 (transition) 函数 \mathcal{T} ： $\mathcal{T} : \mathcal{S} \times \mathcal{A} \rightarrow \Pi(\mathcal{S})$ (环境模型)
- 立即回报(immediate reward)函数 \mathcal{R} ：
 $\mathcal{R} : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \rightarrow \mathbb{R}$ ：
 - 给定 (s, a, s') 下的立即回报 $\mathcal{R}(s, a, s')$ 或 $r(s, a, s')$
 - 给定 (s, a) 下的期望立即回报： $r(s, a) = \sum_{s' \in \mathcal{S}} \tau(s, a, s') r(s, a, s')$



从MDP到Tier-MDP



■ 对比实验

	状态数		对话轮次		错误应答次数	
	MDP	Tier-MDP	MDP	Tier-MDP	MDP	Tier-MDP
第一组			11	7	0	0
第二组			11	7	4	0
第三组	243	64	8	5	0	0
第四组			8	5	1	0
第五组			7	5	0	0

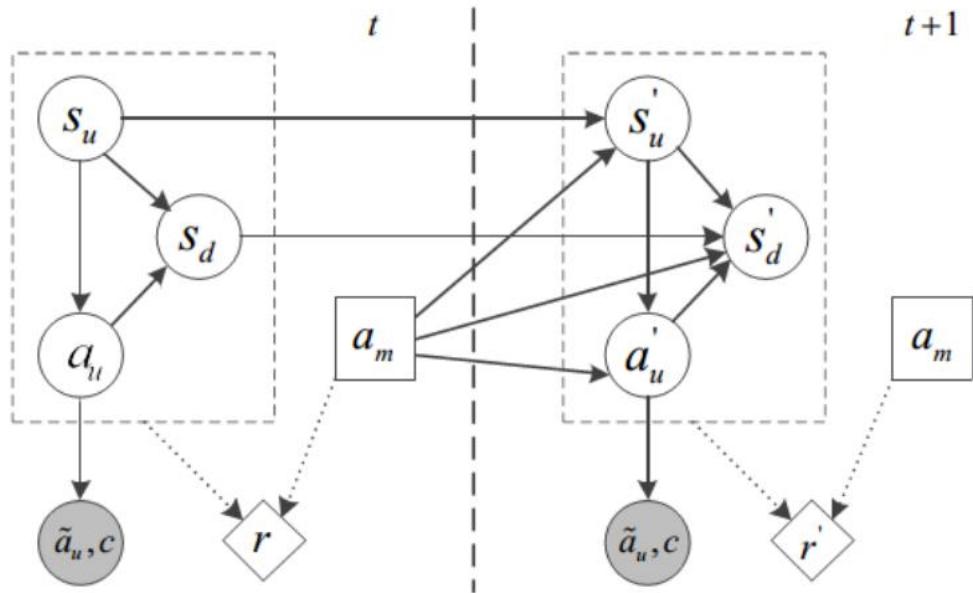
设在原MDP模型中状态数为n, 动作集大小为m, 子任务数为k。则：MDP的状态数为 n^k ，Tier-MDP的状态数为 $\max(n, 2^m)$ ，m和k通常为同一量级，而n远大于2，则： $\max(n, 2^m) \ll n^k$ 。即Tier-MDP的状态数远小于MDP的状态数。

动作选择(策略模型)

- 确定性对话状态下的策略模型
- 不确定性对话状态下的策略模型

动作选择 (策略模型)

- 具有不确定性的对话状态：信念状态
- POMDP模型----基于信念状态的MDP



动作选择(策略模型)

- 具有不确定性的对话状态：信念状态
- POMDP模型----基于信念状态的MDP
- 但信念状态空间规模大，之前面向MDP的精确算法、近似算法都难以实施
 - 基于摘要空间上的优化：MC优化、NAC优化...
 - 高斯过程强化学习(Gasic 2014)
 - 深度强化学习(联合ST乃至NLU模型)

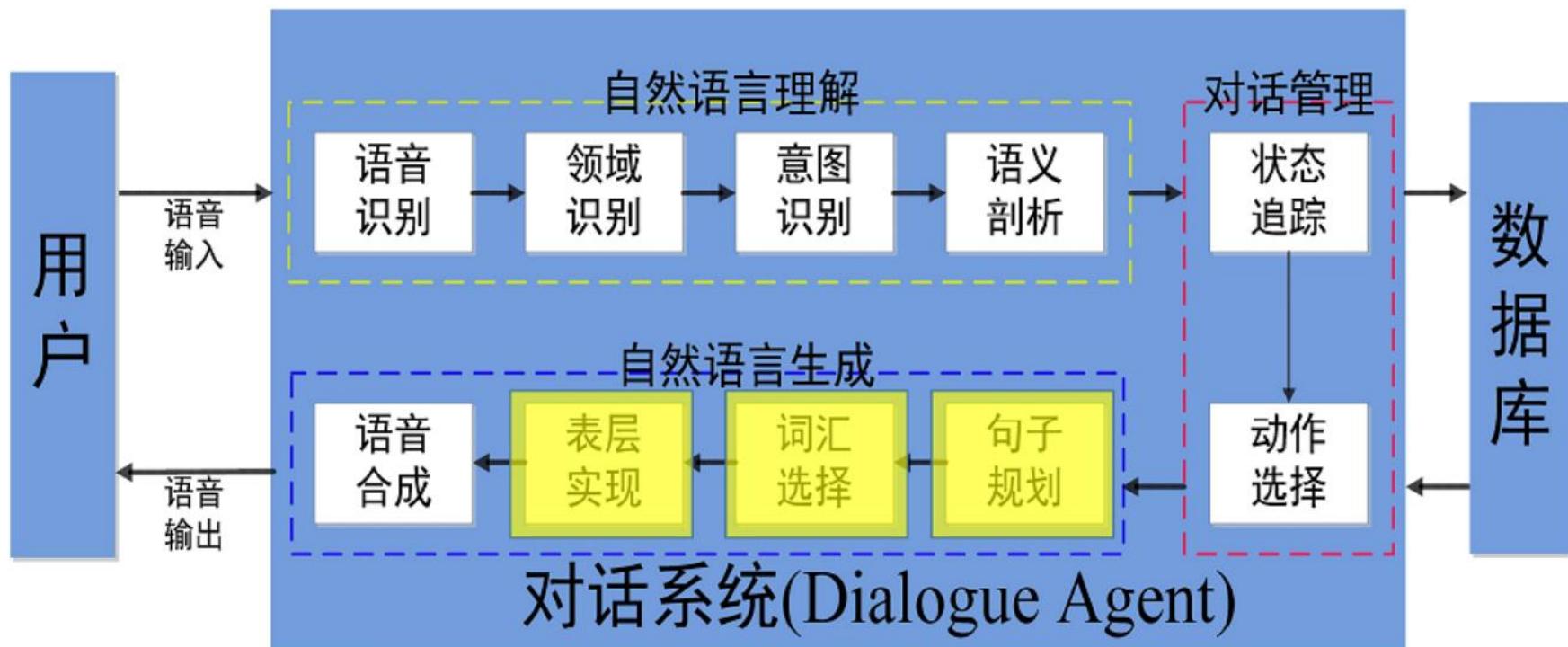
用户仿真器

- 策略模型的求解需要进行大量交互对话
- 预先收集的对话策略可能与真实进行的对话不同
- 最好的方式是用户直接与对话引擎交互，但人的负担大
- 因此：用户仿真器(包括对错误的仿真)
 - 动作层：面向策略的训练
 - 文本层：增加文本理解的容错
 - 语音层：增加语音识别的容错
- 方法：
 - 模板、N-gram(Georgila 2005)
 - 动态Bayes网、条件随机场、IRL(Chandramohan 2011)
 - 人机对话引擎

各部分独立建模的研究-用户仿真器

- K. Georgila, J. Henderson, and O. Lemon, “Learning user simulations for information state update dialogue systems,” in Proceedings of Interspeech, Lisbon, Portugal, 2005, pp. 893–896.
- S. Chandramohan, M. Geist, F. Lefevre and O. Pietquin "User simulation in dialogue systems using inverse reinforcement learning", Proc. Annu. Conf. Int. Speech Commun. Assoc., pp.1025-1028 2011
- F. Garc' ia, L. F. Hurtado, D. Griol, M. J. Castro, E. Segarra, and E. Sanchis, “Recognition and understanding simulation for a spoken dialog corpus acquisition,” in Proceedings Text, Speech and Dialogue, 2007, pp. 574–581.
- B. Thomson, K. Yu, M. Gasic, S. Keizer, F. Mairesse, J. Schatzmann, and S. Young, “Evaluating Semantic-level Confidence Scores with Multiple Hypotheses,” in Proceedings of Interspeech, Brisbane, Australia, 2008, pp. 1153–1156.
-

各部分独立建模的研究



各部分独立建模的研究

■ 句子规划

- Ondrej Dusek and Filip Jurcicek. (2015). Training a Natural Language Generator From Unaligned Data. In Proceedings of the 53rd ACL and the 7th IJCNLP, pages 451–461, Beijing, China
- Stone, M. and Doran, C. (1997). Sentence planning as description using tree adjoining grammar. In Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics, pages 198-205, Madrid, Spain.
-

■ 表层实现

- White, M. and Baldridge, J. (2003). Adapting chart realization to CCG. In Proceedings of 9th European Workshop on Natural Language Generation, Budapest, Hungary.
- Lavoie, B. and Rambow, O. (1997). A fast and portable realizer for text generation systems. In Proceedings of the fifth conference on Applied natural language processing, ANLC, pages 265-268, Washington, DC.
-

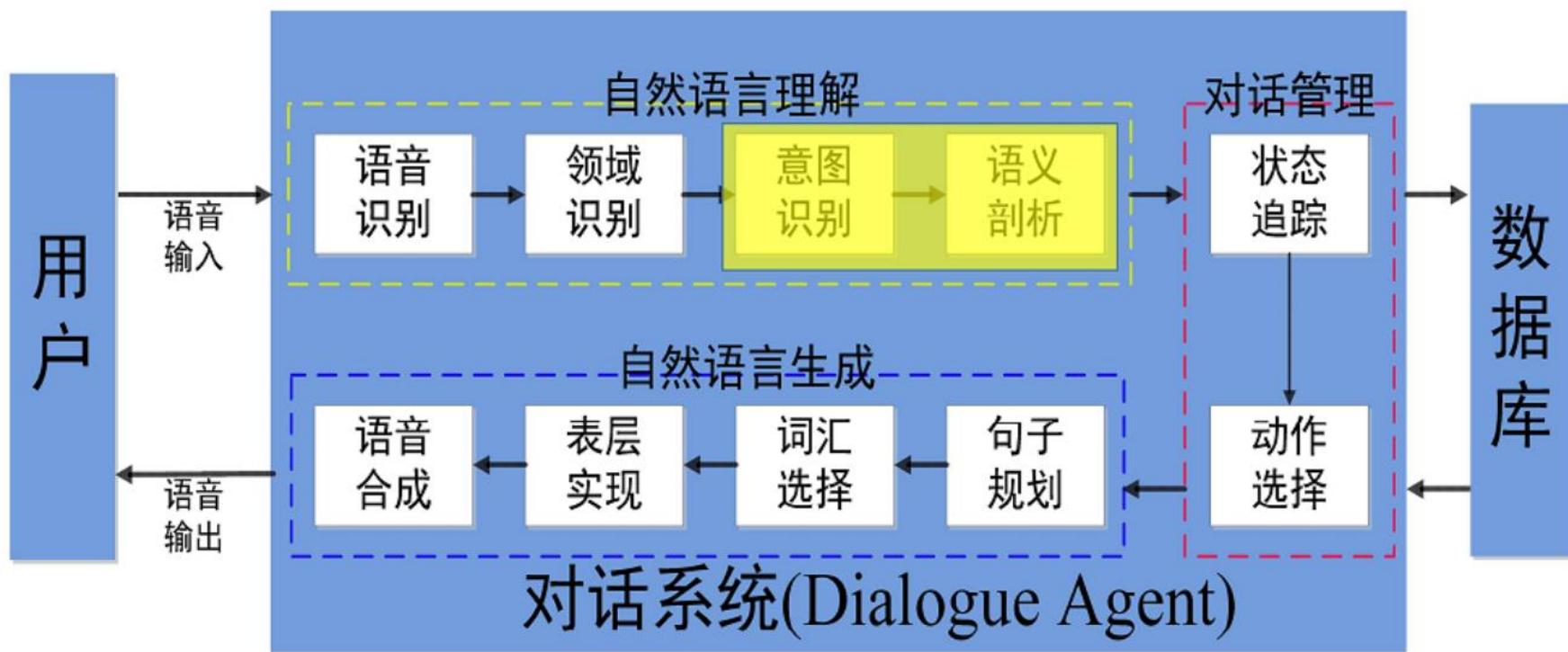
各部分独立建模的研究

■问题

- 错误逐步前向累积
- 误差不能反向传播
- 每个模块都需要自己的训练数据
- 一个模块变化时，其他模块可能要重新训练
- 模块中间接口的设计任务复杂

- 引言
- 目标驱动型人机对话系统
- 子任务各自建模
- 多任务联合建模
- 结论与展望

自然语言理解内部的联合



联合意图识别和语义剖析：单一模型

■ 将作为分类任务的意图识别转化为标注任务，和语义剖析集成在一个序列标注模型中

输入： I like a Chinese food .

标签： O O O B I O

Jeong M, Lee G G. Jointly predicting dialog act and named entity for spoken language understanding[C]// Spoken Language Technology Workshop. 2007:66-69

联合意图识别和语义剖析：单一模型

- 将作为分类任务的意图识别转化为标注任务，和语义剖析集成在一个序列标注模型中

输入： I like a Chinese food . **Intention 占位符**

标签： O O O B I O order

Jeong M, Lee G G. Jointly predicting dialog act and named entity for spoken language understanding[C]// Spoken Language Technology Workshop. 2007:66-69

联合意图识别和语义剖析：单一模型

- 将作为分类任务的意图识别转化为标注任务，和语义剖析集成在一个序列标注模型中

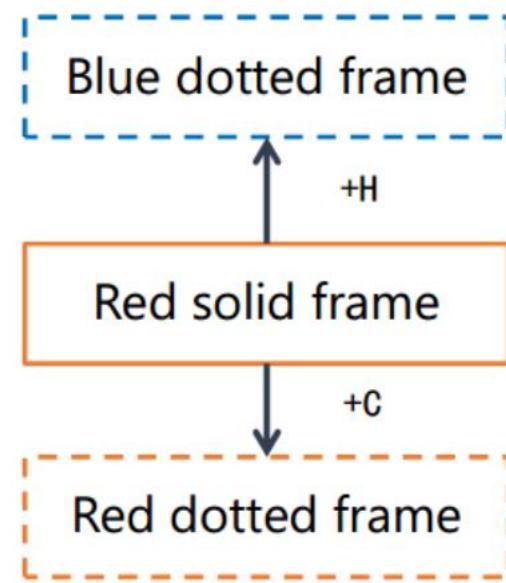
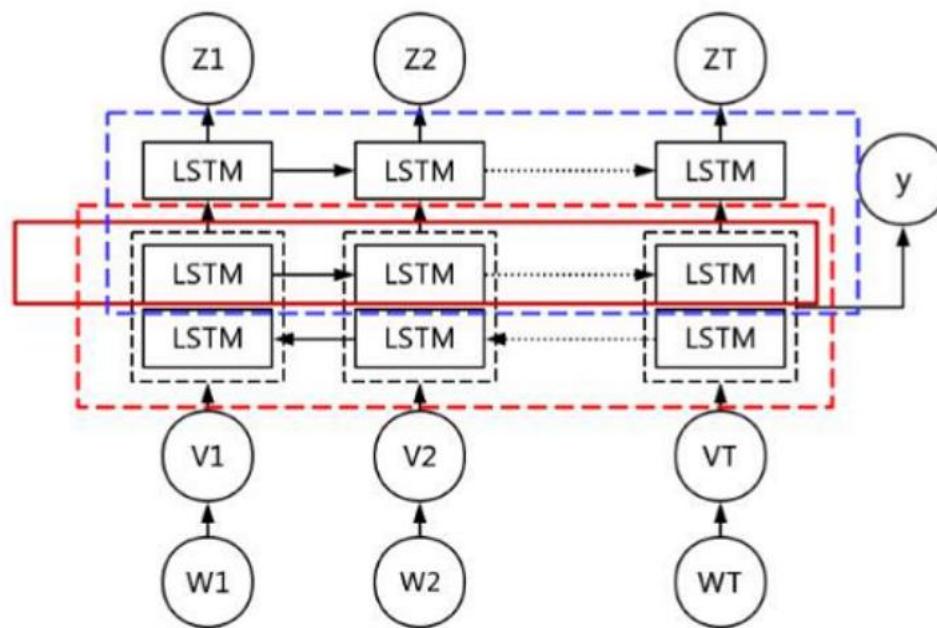
输入： I like a Chinese food . **Intention 占位符**

标签： O O O B I O **order**

- 后续各种序标模型用于联合建模：CRF、LSTM-CRF等
- 问题：两个子任务之间的层次约束关系

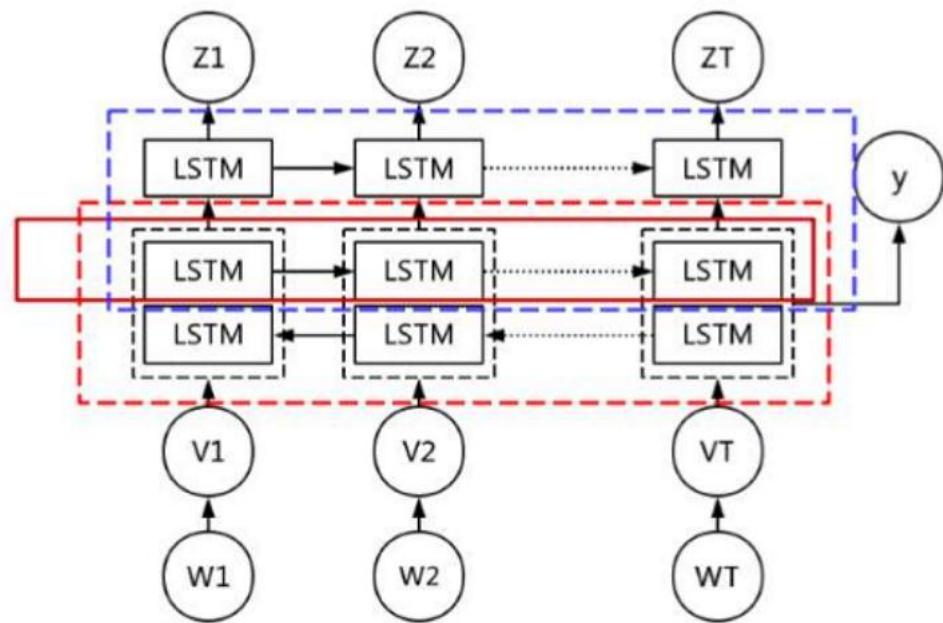
Jeong M, Lee G G. Jointly predicting dialog act and named entity for spoken language understanding[C]// Spoken Language Technology Workshop. 2007:66-69

联合意图识别和语义剖析：层次模型CHJ



Wen L, Wang X, et al. Jointly Modeling Intent Identification and slot filling with Contextual and Hierarchical Information // Natural Language Understanding and Intelligent Applications. Springer International Publishing, 2017

联合意图识别和语义剖析：层次模型CHJ



- LOSS function

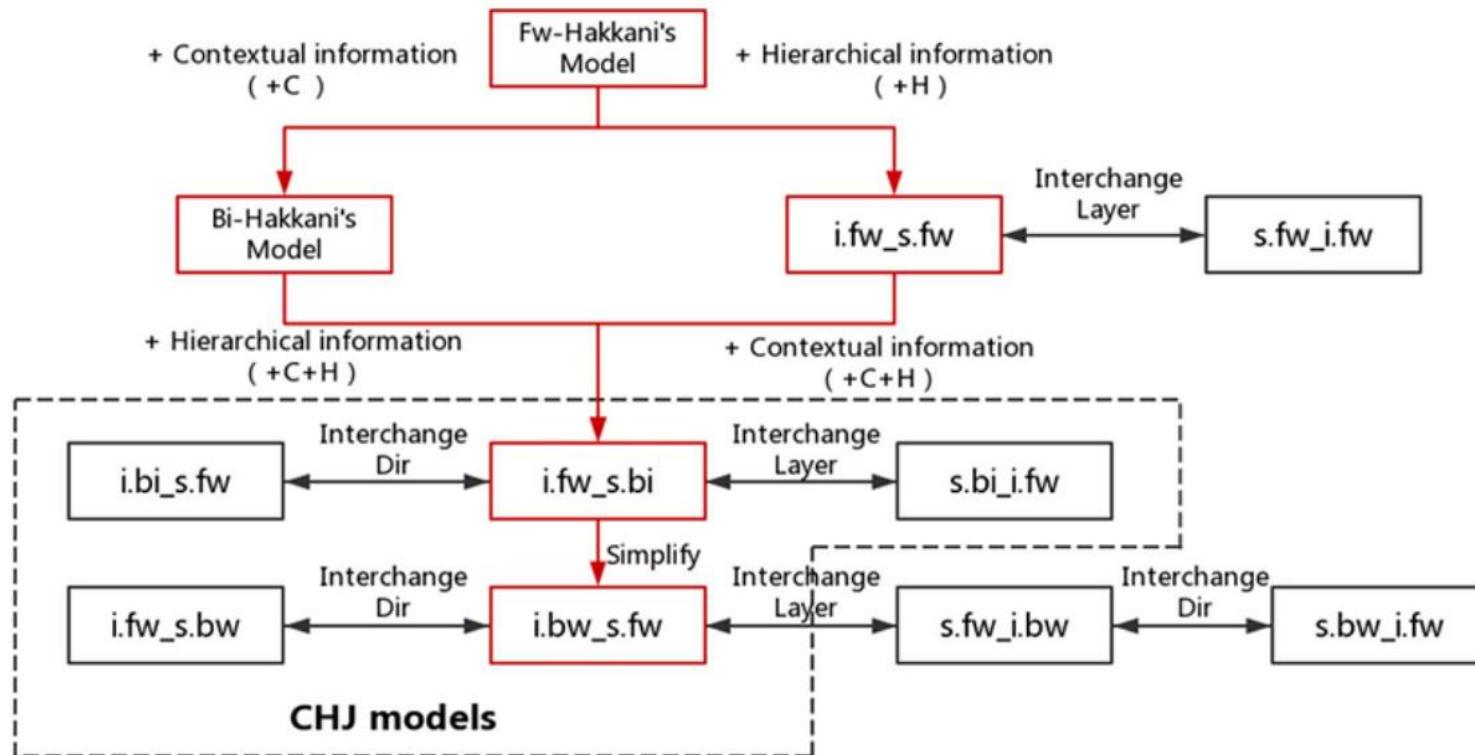
$$J(\theta) = \alpha L_y + (1 - \alpha) L_z$$

$$L_y = \frac{1}{|\mathcal{D}|} \sum_i^{|\mathcal{D}|} L(y^{(i)}, \hat{y}^{(i)})$$

$$L_z = \frac{1}{|\mathcal{D}|} \sum_i^{|\mathcal{D}|} \frac{1}{n} \sum_{t=1}^n L(z_t^{(i)}, \hat{z}_t^{(i)})$$

Wen L, Wang X, et al. Jointly Modeling Intent Identification and slot filling with Contextual and Hierarchical Information // Natural Language Understanding and Intelligent Applications. Springer International Publishing, 2017

层次模型(CHJ)簇



在DSTC2、DSTC5等数据集上的比较

Table 3. Results on DSTC2 corpus

Information (C/H)	Model	Slot(F1)%	Intent(F1)%	Avg(F1)%
None	Fw-Hakkani's Model	96.22	99.22	97.72
+C	Bi-Hakkani's Model	98.92	99.18	99.05
+H	i/fw_s/fw	96.67	99.23	97.95
	s/fw_i/fw	96.64	99.46	98.05
	s/fw_i/bw	96.62	99.39	98.01
	s/bw_i/fw	97.96	99.48	98.72
+C+H (CHJ models)	s/bi_i/fw	98.96	99.45	99.21
	i/fw_s/bi	98.81	99.33	99.07
	i/bi_s/fw	98.93	99.37	99.15
	i/bw_s/fw	99.01	99.48	99.25
	i/fw_s/bw	98.91	99.34	99.13