

基于深度学习的对话系统

冯岩松

北京大学

October 19, 2016

任务

使用自然语言进行多轮次的交互

任务

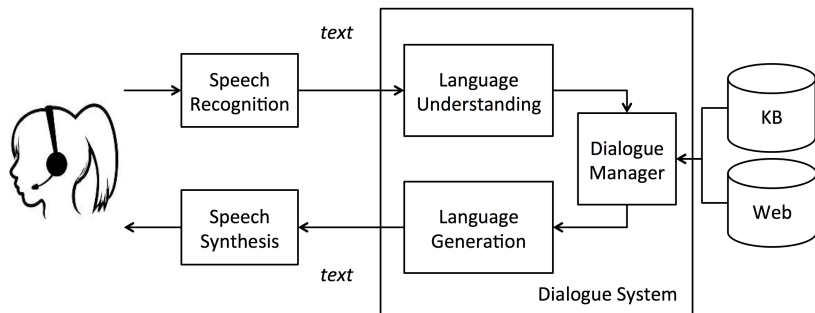
使用自然语言进行多轮次的交互

- **输入：** 自然语言句子
- **资源：** 领域知识、背景知识等显式或隐式资源
- **输出：** 符合当前上下文语境或满足任务需求的自然语言句子

任务

使用自然语言进行多轮次的交互

- 输入：自然语言句子
- 资源：领域知识、背景知识等显式或隐式资源
- 输出：符合当前上下文语境或满足任务需求的自然语言句子



[Wen, 2016]

任务框架

- 语言理解

- 负责将输入句子解析为形式化表示

- 对话管理

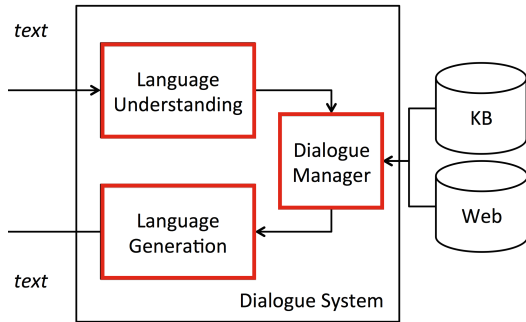
- 对话状态跟踪 (Dialogue State Tracker, 或 Belief Tracking)
 - 对话内容选择 (Response Selection)

- 语言生成

- 输出满足要求的自然语言句子

- 外部资源

- 领域知识、世界知识



任务框架

- 语言理解

- 负责将输入句子解析为形式化表示

- 对话管理

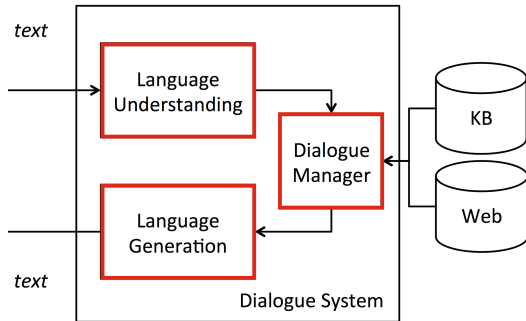
- 对话状态跟踪
(Dialogue State Tracker, 或 Belief Tracking)
 - 对话内容选择
(Response Selection)

- 语言生成

- 输出满足要求的自然语言句子

- 外部资源

- 领域知识、世界知识



任务框架

- 语言理解

- 负责将输入句子解析为形式化表示

- 对话管理

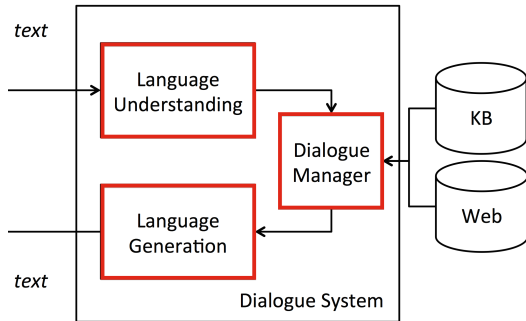
- 对话状态跟踪 (Dialogue State Tracker, 或 Belief Tracking)
- 对话内容选择 (Response Selection)

- 语言生成

- 输出满足要求的自然语言句子

- 外部资源

- 领域知识、世界知识



任务框架

- 语言理解

- 负责将输入句子解析为形式化表示

- 对话管理

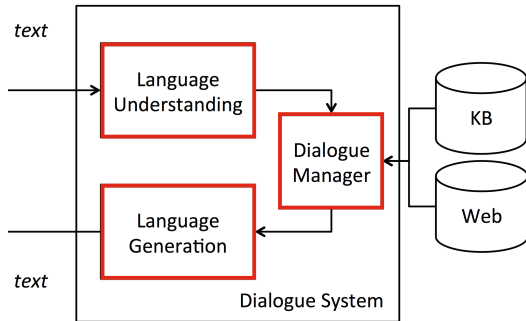
- 对话状态跟踪 (Dialogue State Tracker, 或 Belief Tracking)
 - 对话内容选择 (Response Selection)

- 语言生成

- 输出满足要求的自然语言句子

- 外部资源

- 领域知识、世界知识



- **任务导向**：完成某项任务
 - 特定商业任务：订旅馆、订机票、售后服务……
 - 评价方式：任务的完成程度
- **开放域**：没有特别限定某个任务
 - 聊天机器人，小冰，小X……
 - 评价方式：多为主观评价，以及对话长度、是否重复等

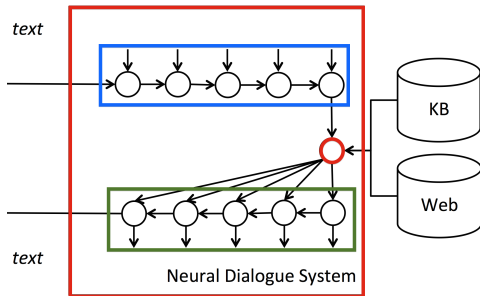
- **任务导向**：完成某项任务
 - 特定商业任务：订旅馆、订机票、售后服务……
 - 评价方式：任务的完成程度
- **开放域**：没有特别限定某个任务
 - 聊天机器人，小冰，小X……
 - 评价方式：多为主观评价，以及对话长度、是否重复等
- **综合类**
 - 个人助理类，siri, Cortana, 度秘……
 - 评价方式：任务完成情况，主观感受

传统对话系统

- 大量手工设计的特征、规则或模板
- 对上下文语境、用户的建模能力有限
- 训练数据有限，领域相关性强，迁移到其他领域的代价很高
- 几乎无法提供聊天功能

当神经网络来了……

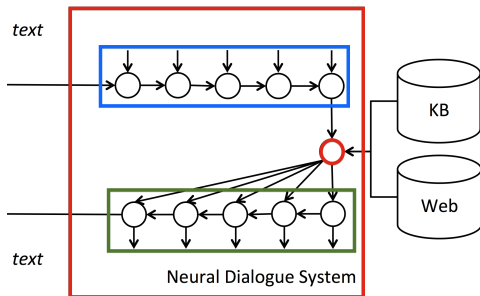
- 语言理解
- 对话管理
 - 对话状态跟踪
 - 对话内容选择
- 语言生成
- 外部资源
 - 领域知识、世界知识



[Wen, 2016]

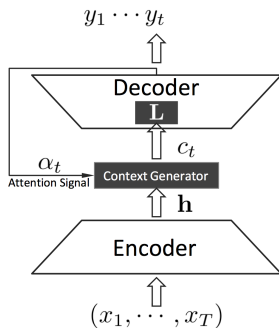
当神经网络来了……

- 语言理解 \implies 神经网络模型
- 对话管理 \implies 神经网络模型
 - 对话状态跟踪
 - 对话内容选择
- 语言生成 \implies 神经网络模型
- 外部资源 \implies 神经网络模型
 - 领域知识、世界知识



[Wen, 2016]

方法: Encoding-Decoding



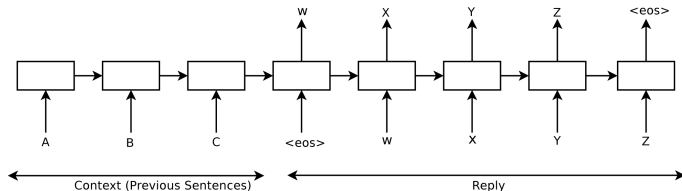
[Shang et al., 2015]

方法: Encoding-Decoding

ID	Post	NRM-glo	NRM-loc	NRM-hyb	Rtr.-based
1	每年元旦都来高烧。 High fever attacks me every New Year's Day	多休息，增强免疫力！ Rest more to boost your immune system.	每年元旦都去医院看病 Go to see a doctor every New Year's Day	祝早日康复，身体健康！ Get well soon and stay healthy!	高烧... High fever...
2	又老了一岁，感谢我们组，我很高兴 I gain one more year. Grateful to my group, so happy	生日快乐！永远支持你！ Happy birthday! Will stand by you forever!	生日快乐！永远年轻漂亮！ Happy birthday! Wish you always stay young and pretty!	老了老了，岁月不饶人啊！ Getting old now. Time has no mercy.	我只是老了一岁 I just gain only one more year.
3	吃素第一天，坚持住，崔朵拉。 First day of being a vegetarian. Hold on, CuiDuoLa	必须来支持！加油！ I come to offer my support. Keep it up!	坚持住，坚持下去。 Hold on, keep it up.	每天都是素食主义者 I have been being a vegetarian everyday	坚持住！ Hold on!
4	这种天气真应该带个帅锅出去踏青赏花 We should go out with some cute guys to enjoy a great outing in such a nice weather.	这是哪儿啊？真美！ Where is it? It is so beautiful!	阳光明媚，心情舒畅。 Such a nice sunny day! I am in a great mood.	这种天气真不是盖的。 It is indeed a very nice weather.	文山啊 出去踏青寻找灵感 哈哈 WenShan, let's go out to get some inspiration. Ha! Ha!

[Shang et al., 2015]

方法: Seq2Seq



Human: *hello !*
Machine: *hello !*
Human: *how are you ?*
Machine: *i 'm good .*
Human: *what 's your name ?*
Machine: *i 'm julia .*
Human: *when were you born ?*
Machine: *july 20th .*
Human: *what year were you born ?*
Machine: *1977 .*
Human: *where are you from ?*
Machine: *i 'm out in the boonies.*
Human: *are you a man or a woman ?*
Machine: *i 'm a woman .*
Human: *why are we here ?*
Machine: *i 'm not sure .*
Human: *okay, bye !*
Machine: *bye .*
Human: *see you later .*
Machine: *bye .*

Human: *who is skywalker ?*
Machine: *he is a hero .*
Human: *who is bill clinton ?*
Machine: *he 's a billionaire .*
Human: *is sky blue or black ?*
Machine: *blue .*
Human: *does a cat have a tail ?*
Machine: *yes .*
Human: *does a cat have a wing ?*
Machine: *no*
Human: *can a cat fly ?*
Machine: *no .*
Human: *how many legs does a cat have ?*
Machine: *four , i think .*
Human: *how many legs does a spider have ?*
Machine: *three , i think .*
Human: *how many legs does a centipede have ?*
Machine: *eight .*
Human: *what is the color of the sky ?*
Machine: *blue .*
Human: *what is the color of water ?*
Machine: *water .*

Bianchi and Le, 2015

感觉聊天还不错！

- 语言理解
 - RNN 捕捉 输入序列
- 对话管理
 - 由隐层来负责
- 语言生成
 - RNN 负责输出序列
- 外部资源
 - 拼个向量？

然而，这是对话场景！可能会有多个轮次的

感觉聊天还不错！

- 语言理解
 - RNN 捕捉 输入序列
- 对话管理
 - 由隐层来负责
- 语言生成
 - RNN 负责输出序列
- 外部资源
 - 拼个向量？

然而，这是对话场景！可能会有多个轮次的
message Where do you live now?

response I live in Los Angeles.

message In which city do you live now?

response I live in Madrid.

message In which country do you live now?

response England, you?

[Li et al., 2016]

感觉聊天还不错！

- 语言理解
 - RNN 捕捉 输入序列
- 对话管理
 - 由隐层来负责
- 语言生成
 - RNN 负责输出序列
- 外部资源
 - 拼个向量？

然而，这是对话场景！可能会有多个轮次的

message Where were you born?

response I was born in Canada.

message Where are you from?

response England, you?

message Where did you grow up?

response I grew up in Texas.

[Li et al., 2016]

感觉聊天还不错！

- 语言理解
 - RNN 捕捉 输入序列
- 对话管理
 - 由隐层来负责
- 语言生成
 - RNN 负责输出序列
- 外部资源
 - 拼个向量？

然而，这是对话场景！可能会有多个轮次的

message How old are you?

response 16 and you?

message What's your age?

response 18.

[Li et al., 2016]

感觉聊天还不错！

- 语言理解
 - RNN 捕捉 输入序列
- 对话管理
 - 由隐层来负责
- 语言生成
 - RNN 负责输出序列
- 外部资源
 - 拼个向量？

然而，这是对话场景！可能会有多个轮次的

message What is your major?

response I'm majoring in psychology

message What did you study in college?

response English lit.

[Li et al., 2016]

感觉聊天还不错！

- 语言理解
 - RNN 捕捉 输入序列
- 对话管理
 - 由隐层来负责
- 语言生成
 - RNN 负责输出序列
- 外部资源
 - 拼个向量？

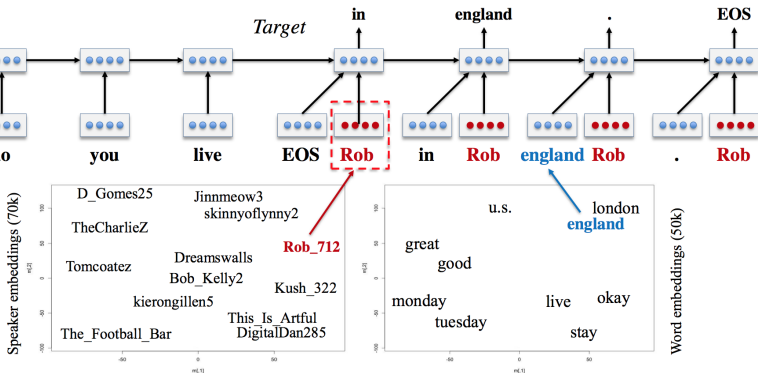
然而，这是对话场景！可能会有多个轮次的

- 这个机器人好神奇……
- 回复质量好像不够稳定
- 用户建模缺失
- 能完成订票任务吗？这End2End会不会把我的票订错？

- ❶ 如何对当前的场景 (Context) 建模
 - 已发生的对话
 - 是否有任务场景
- ❷ 如何对用户建模
 - 人格一致：性别、年龄、背景……
- ❸ 如何有好的对话回合
 - 有信息量、不重复、回合多
- ❹ 如何把任务目标融入进来
 - 发挥神经网络特长解决个别模块，如语言生成

加入用户建模

还是Seq2Seq，但强调人物建模



et al., 2016]

[Li

还是Seq2Seq，但强调人物建模

message Where is your hometown?
response I'm from England.
message Where are you from?
response I'm from England.
message In which city do you live now?
response I live in London.
message In which country do you live now?
response I live in England.

message What is your major?
response Business. You?
message What did you study in college?
response I did business studies.

message How old are you?
response I'm 18.
message What is your age?
response I'm 18.

message Where is your hometown?
response I live in Indonesia.
message Where are you from?
response England, you?
message In which city do you live now?
response I live in Jakarta.
message In which country do you live now?
response I live in Indonesia.

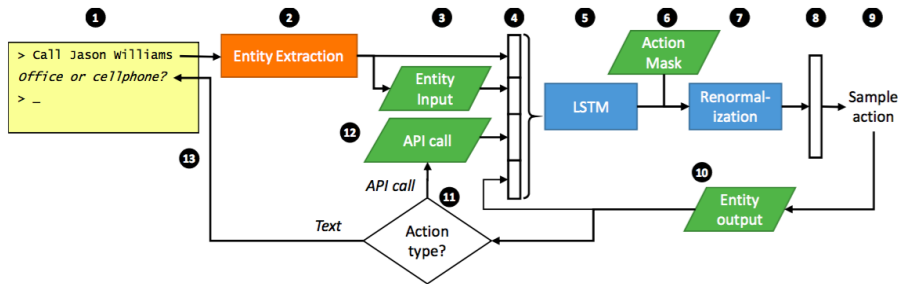
message What is your major?
response Business, you?
message What did you study in college?
response Psychology, you?

message How old are you?
response I'm 18.
message What is your age?
response I'm 16.

每人一个模型……

任务完成类

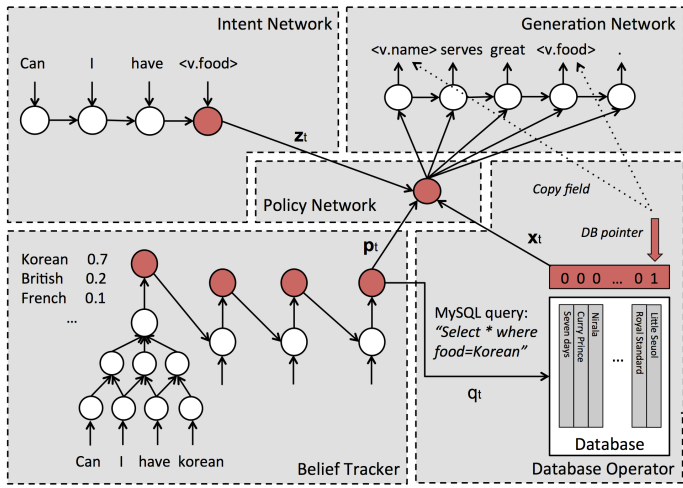
- 多数系统不会完全依赖 Seq2Seq 这种端对端系统模式
- 单独模块应用神经网络模型，如 语言生成、决策判断等
- 来个端对端任务类系统？



[Williams and Zweig, 2016]

任务完成类

- 多数系统不会完全依赖 Seq2Seq 这种端对端系统模式
- 单独模块应用神经网络模型，如 语言生成、决策判断等
- 来个端对端任务类系统？**你可以的！**



[Wen et al., 2016]

聊天、微博、论坛、电影……

Name	Type	Topics	Avg. # of turns	Total # of dialogues	Total # of words	Description
NPS Chat Corpus (Forsyth and Martell, 2007)	Chat	Unrestricted	–	–	100M	Posts from age-specific online chat rooms.
Twitter Corpus (Ritter et al., 2010)	Microblog	Unrestricted	2	1,300,000	–	Tweets and replies extracted from Twitter
Twitter Triple Corpus (Sordoni et al., 2015b)	Microblog	Unrestricted	3	29,000,000	–	A-B-A triples extracted from Twitter
Sina Weibo Corpus (Shang et al., 2015)	Microblog	Unrestricted	2	4,435,959	100M	Posts and replies extracted from Sina Weibo
Sina Weibo Short-Text Conversation Corpus (Wang et al., 2013)	Microblog	Unrestricted	2	15,000	–	Posts, along with unlabeled replies from Sina Weibo
UseNet Corpus (Shaoul and Westbury, 2009)	Microblog	Unrestricted	–	–	7B	UseNet forum postings
NUS SMS Corpus (Chen and Kan, 2013)	SMS messages	Unrestricted	–	–	580,668* [□]	SMS messages collected between two users, with timing analysis.
Reddit	Forum	Unrestricted	–	–	–	Comment trees from Reddit
Reddit Domestic Abuse Corpus (Schrading et al., 2015)	Forum	Abuse help	17.53	21,133	19M-103M [△]	Reddit posts from either domestic abuse subreddits, or general chat.
Settlers of Catan (Afantenos et al., 2012)	Chat	Game terms	95	21	–	Conversations between players in the game ‘Settlers of Catan’
Internet Argument Corpus (Walker et al., 2012b)	Forum	Politics	35.45	11K	73M	Debates about specific political or moral positions
MPC Corpus (Shaikh et al., 2010)	Chat	Social tasks	520	14	58K	Conversations about general, political, and interview topics
Ubuntu Dialogue Corpus (Lowe et al., 2015a)	Chat	Ubuntu Operating System	7.71	930,000	100M	Dialogues extracted from Ubuntu chat stream on IRC
Ubuntu Chat Corpus (Uthus and Aha, 2013)	Chat	Ubuntu Operating System	–	–	2B* [□]	Chat stream scraped from IRC logs (no dialogues extracted)
Movie Dialog Dataset (Dodge et al., 2015)	Chat, QA & Recommendation	Movies	3.3	3.1M [▼]	–	For goal-driven dialogue systems. Includes movie metadata as knowledge triples.

聊天、微博、论坛、电影……

Name	Topics	Total # of utterances	Total # of dialogues	Total # of scripts	Total # of words	Description
Movie-DiC (Banchs, 2012)	Movie dialogues	764k	132,229	753	6M	Movie scripts of American films.
Movie-Triples (Serban et al., 2015)	Movie dialogues	736k	245,296	614	13M	Triples of utterances which are filtered to come from X-Y-X triples.
American Film Scripts	Movie scripts	1M*	–	1,500	16M*	Movie scripts of American films.
Cornell Movie-Dialogue Corpus (Danescu-Niculescu-Mizil and Lee, 2011)	Movie dialogues	1.4	220,579	617	9M*	Short conversations from film scripts, annotated with character metadata.
Filtered Movie Script Corpus (Nio et al., 2014b)	Movie dialogues	173k	86,719	1,786	2M*	Triples of utterances which are filtered to come from X-Y-X triples.
American Soap Opera Corpus (Davies, 2012b)	TV show scripts	10M*	–	22,000	100M	Transcripts of American soap operas.
TVD Corpus (Roy et al., 2014)	TV show scripts	60k*	–	191	600k*	TV scripts from a comedy (Big Bang Theory) and drama (Game of Thrones) show.
Character Style from Film Corpus (Walker et al., 2012a)	Movie scripts	664k	—	862	9.6M	Scripts from IMSDb, annotated for linguistic structures and character archetypes.
SubTle Corpus (Ameixa and Coheur, 2013)	Movie subtitles	6.7M	3.35M	6,184	20M	Aligned interaction-response pairs from movie subtitles.
OpenSubtitles (Tiedemann, 2012)	Movie subtitles	140M*	–	207,907	1B	Movie subtitles which are not speaker-aligned.

[Serban et al., 2015]

聊天、微博、论坛、电影……

Name	Topics	Total # of dialogues	Total # of words	Total length	Description
Switchboard (Godfrey et al., 1992)	Casual Topics	2,400	3M	300hrs*	Telephone conversations on pre-specified topics
British National Corpus (BNC) (Leech, 1992)	Casual Topics	–	10M	1,000hrs*	British dialogues many contexts, from formal business or government meetings to radio shows and phone-ins.
CALLHOME American English Speech (Canavan et al., 1997)	Casual Topics	120	540k*	60hrs	Telephone conversations between family members or close friends.
CALLFRIEND American English Non-Southern Dialect (Canavan and Zipperlen, 1996)	Casual Topics	60	180k*	20hrs	Telephone conversations between Americans with a Southern accent.
The Bergen Corpus of London Teenage Language (Haslerud and Stenström, 1995)	Unrestricted	–	500k	55hrs	Spontaneous teenage talk recorded in 1993. Conversations were recorded secretly.
Longman Spoken American Corpus (Stern, 2005)	Casual Topics	–	5M	550hrs*	First dataset to capture American spoken English on a large scale.
The Cambridge and Nottingham Corpus of Discourse in English (McCarthy, 1998)	Casual Topics	–	5M	550hrs*	British dialogues from wide variety of informal contexts, such as hair salons, restaurants, etc.
D64 Multimodal Conversation Corpus (Oertel et al., 2013)	Unrestricted	2	70k*	8hrs	Several hours of natural interaction between a group of people
AMI Meeting Corpus (Renals et al., 2007)	Meetings	175	900k*	100hrs	Face-to-face meeting recordings.
Cardiff Conversation Database (CCDb) (Aubrey et al., 2013)	Unrestricted	30	20k*	150min	Audio-visual database with unscripted natural conversations, including visual annotations.
4D Cardiff Conversation Database (4D CCdb) (Vandeventer et al., 2015)	Unrestricted	17	2.5k*	17min	A version of the CCdb with 3D video
The Diachronic Corpus of Present-Day Spoken English (Aarts and Wallis, 2006)	Casual Topics	–	800k	80hrs*	Selection of face-to-face, telephone, and public discussion dialogue from Britain.
The Spoken Corpus of the Survey of English Dialects (Beare and Scott, 1999)	Casual Topics	314	800k	60hrs	Dialogue of people aged 60 or above talking about their memories, families, work and the folklore of the countryside from a century ago.
The Child Language Data Exchange System (MacWhinney and Snow, 1985)	Unrestricted	–	10M	1,000hrs*	International database organized for the study of first and second language acquisition.
The Charlotte Narrative and Conversation Collection (CNCC) (Reppen and Ide, 2004)	Casual Topics	95	20K	2hrs*	Narratives, conversations and interviews representative of the residents of Mecklenburg County, North Carolina.

任务类：旅游、参观、家居……

Name	Type	Topics	Avg. # of turns	Total # of dialogues	Total # of words	Description
DSTC1 (Williams et al., 2013)	Spoken	Bus schedules	–	15,000	–	Bus ride information system
DSTC2 (Henderson et al., 2014b)	Spoken	Restaurants	–	3,000	–	Restaurant booking system
DSTC3 (Henderson et al., 2014a)	Spoken	Tourist information	–	2,265	–	Information for tourists
CMU Communicator Corpus (Bennett and Rudnicky, 2002)	Spoken	Travel	11.67	15,481	–	Travel planning and booking system
ATIS Pilot Corpus (Hemphill et al., 1990)	Travel	Travel	25.4	41	10.6K-12.2K*	Travel planning and booking system
Ritel Corpus (Rosset and Petel, 2006)	Spoken	Unrestricted/ Diverse Topics	–	582	60k	An annotated open-domain question answering spoken dialogue system
DIALOG Mathematical Proofs (Wolska et al., 2004)	Spoken	Mathematics	12	66	–	Humans interact with computer system to do mathematical theorem proving
MATCH Corpus (Georgila et al., 2010)	Spoken	Appointment Scheduling	–	447	–	A system for scheduling appointments. Includes dialogue act annotations
MeMo Corpus (Möller et al., 2008)	Spoken	Household	–	64	–	Instructing smart home system

[Serban et al., 2015]

- Oriol Vinyals, Quoc V. Le, A Neural Conversational Model, In ICML Deep Learning Workshop 2015
- Lifeng Shang, Zhengdong Lu, Hang Li, Neural Responding Machine for Short-Text Conversation, ACL 2015.
- Hierarchical Neural Network Generative Models for Movie Dialogues” , Iulian Vlad Serban, Alessandro Sordoni, Yoshua Bengio, Aaron C. Courville, Joelle Pineau, In AAAI 2015.
- A Diversity-Promoting Objective Function for Neural Conversation Models” , Jiwei Li, Michel Galley, Chris Brockew, Jianfeng Gao, Bill Dolan, NAACL 2016.
- A Persona-Based Neural Conversation Model” , Jiwei Li, Michel Galley, Chris Brockew, Jianfeng Gao, Bill Dolan, In ACL 2016
- Jason D. Williams and Geoffrey Zweig, End-to-end LSTM-based dialog control optimized with supervised and reinforcement learning, In arXiv 2016
- Tsung-Hsien Wen, Deep Learning for NLP, Tutorial, INLG 2016

- Zhao Yan, Nan Duan, Junwei Bao, Peng Chen, Ming Zhou, Zhoujun Li, Jianshe Zhou, DocChat: An Information Retrieval Approach for Chatbot Engines Using Unstructured Documents, In ACL 2016
- Iulian V. Serban, Alessandro Sordoni, Yoshua Bengio, Aaron Courville and Joelle Pineau, Building End-To-End Dialogue Systems Using Generative Hierarchical Neural Network Models, In AAAI 2016
- Tsung-Hsien Wen, David Vandyke, Nikola Mrksic, Milica Gasic, Lina Rojas-Barahona, Pei-Hao Su, Stefan Ultes, and Steve Young, “A Network-based End-to-End Trainable Task-Oriented Dialogue System” , In arXiv 2016.
- Tsung- Hsien Wen, Milica Gasic, Nikola Mrksic, Lina Rojas-Barahona, Pei-Hao Su, Stefan Ultes, David Vandyke, and Steve Young, Conditional Generation and Snapshot Learning in Neural Dialogue Systems, In EMNLP 2016.
- Iulian Vlad Serban, Ryan Lowe, Laurent Charlin, and Joelle Pineau, A Survey of Available Corpora For Building Data-Driven Dialogue Systems, In arXiv 2016