User Intent Modeling for Dialog Systems
对话系统用户意图建模

LIU Qun 刘群

Huawei Noah's Ark Lab 华为诺亚方舟实验室

A talk at 对话分级启智会
2022-08-21

NOAH'S ARK LAB

HUAWEI

Content

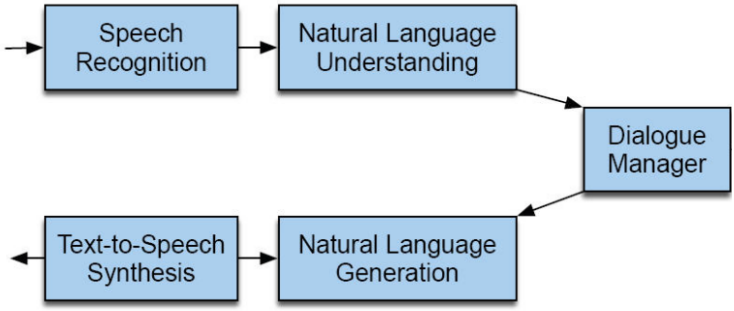
Content

Task-oriented vs. chitchat dialog systems

Two main categories of dialog systems:

- Task-oriented (or goal-oriented) dialog systems
  - Help users to complete certain types of tasks.
  - Tasks (or domain knowledge) should be given in advance, usually as a set of pre-defined intentions and slots.
  - Dialog sessions: the shorter is the better.
- Chitchat dialog systems (or social chatbots, social bots).
  - Chitchat with users on unrestricted topics.
  - Maximize user engagement by generating enjoyable and more human-like conversations.
  - Emotional conversation and personality is welcome.
  - Dialog sessions: the longer is the better.

Task-oriented Dialog Systems



Chitchat Dialog Systems

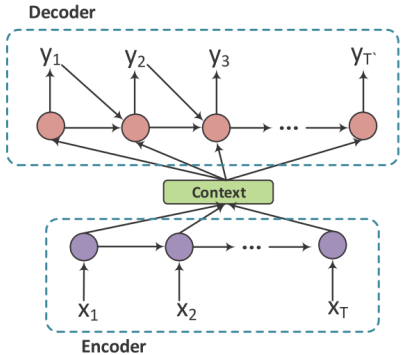


Figure 2: An Illustration of the Encoder-Decoder Model.

Hongshen Chen et al., A Survey on Dialogue Systems: Recent Advances and New Frontiers,

ACM SIGKDD Explorations Newsletter 19.2 (2017)

Multi-domain Task-Completion Dialog Systems

Results of the Multi-Domain Task-Completion Dialog Challenge

Jinchao Li[1], **Baolin Peng**[1], **Sungjin Lee**[1*], **Jianfeng Gao**[1]
Ryuichi Takanobu[2], **Qi Zhu**[2], **Minlie Huang**[2]
Hannes Schulz[3], **Adam Atkinson**[3], **Mahmoud Adada**[3]
[1] Microsoft Research, Redmond, WA, US
{jincli, bapeng, jfgao}@microsoft.com
[2]Tsinghua University, Beijing, China
{gxly19,zhu-q18}@mails.tsinghua.edu.cn aihuang@tsinghua.edu.cn
[3] Microsoft Research, Montréal, Canada
{haschulz,adatkins,maadada}@microsoft.com

Abstract

The paper provides an overview of the "Multi-domain Task Completion" track (Track 1) at the 8th Dialog System Technology Challenge (DSTC-8). There are two tasks in this track. The first task is end-to-end multi-domain task-completion, which aims to build end-to-end task completion dialog systems based on ConvLab. The second task is fast domain adaptation, seeking to develop models that predict user responses when only limited in-domain data is available. We describe the submissions for both tasks, automatic evaluation and human evaluation procedures, and discuss the outcomes of these two evaluations.

Multi-Woz Dataset

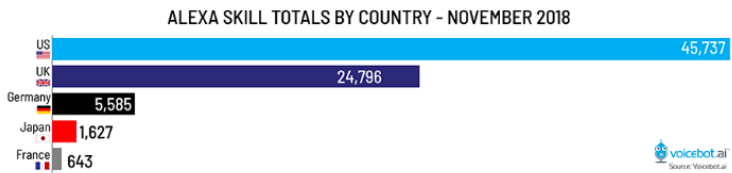
MultiWOZ - A Large-Scale Multi-Domain Wizard-of-Oz Dataset for Task-Oriented Dialogue Modelling

Paweł Budzianowski[1], **Tsung-Hsien Wen**[2*], **Bo-Hsiang Tseng**[1],
Iñigo Casanueva[2*], **Stefan Ultes**[1], **Osman Ramadan**[1] **and Milica Gašić**[1]
[1]Department of Engineering, University of Cambridge, UK,
[2]PolyAI, London, UK
{pfb30,mg436}@cam.ac.uk

Table 2: Full ontology for all domains in our data-set. The upper script indicates which domains it belongs to. *: universal, 1: restaurant, 2: hotel, 3: attraction, 4: taxi, 5: train, 6: hospital, 7: police.

act type	inform* / request* / select[123] / recommend/[123] / not found[123] request booking info[123] / offer booking[1235] / inform booked[1235] / decline booking[1235] welcome* /greet* / bye* / reqmore*
slots	address* / postcode* / phone* / name[1234] / no of choices[1235] / area[123] / pricerange[123] / type[123] / internet[2] / parking[2] / stars[2] / open hours[3] / departure[45] destination[45] / leave after[45] / arrive by[45] / no of people[1235] / reference no.[1235] / trainID[5] / ticket price[5] / travel time[5] / department[7] / day[1235] / no of days[123]

Amazon Alexa: Multi-skill Digital Assistant



https://insidethecask.com/2019/01/01/the-top-5-rated-alexa-skills-in-drinks/

Microsoft XiaoIce: Multi-skill Social Chatbot

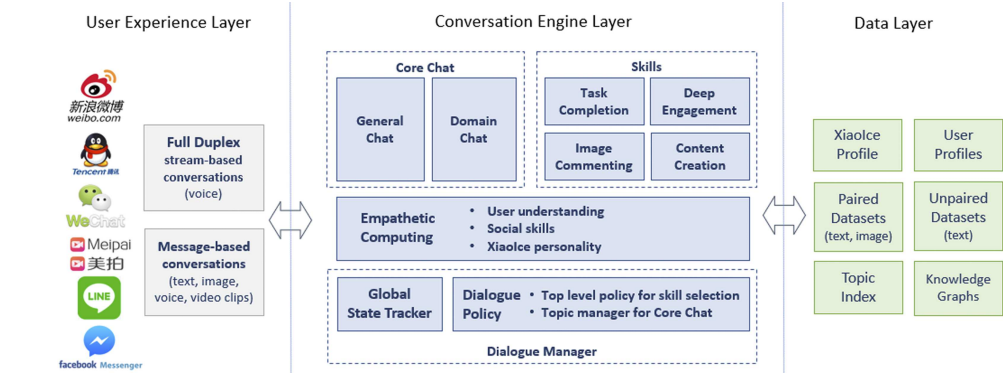
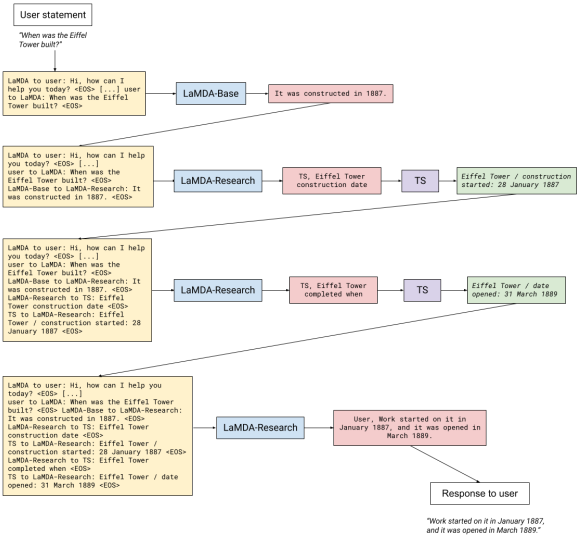


Figure 4
XiaoIce system architecture.

Li et al.,The Design and Implementation of XiaoIce, an Empathetic Social Chatbot, Computational Linguistics, 2020

Google LaMDA: Language Models for Dialog Applications



TS: Toolset

- Information Retrieval System
- Language Translator
- Calculator

Content

User Intent Modeling in Dialog Systems

- Multi-Domain Task Competion System (MultiWOZ, Amazon Alexa):
  - an unified domain-intent-slot ontology
- Multi-Skill Social Chatbot (XiaoIce, LaMDA):
  - no general intent modeling for the whole system
  - intent modeling may exist in the skills as plug-in components

Why User Intent Modeling is Important for Dialog Systems

- Users are engaged in a conversation to fullfil their certain requirements
- The user requirements can be represented as an hierarchy of intents
- Without understanding the user intents, a dialog agent is hard to meet the user requirements
- The user intents could be regarded as extensions of Dialog Acts or Speech Acts which has been researched a lot in theory

Speech Acts: categories

- By John Langshaw Austin:
  - 判定语（verdictives）
  - 裁定语（exercitives）
  - 承诺语（commissives）
  - 阐述语（expositives）
  - 行为语（behabitives）
- By John Rogers Searle:
  - 断言（assertives）
  - 指令（directives）
  - 承诺语（commissives）
  - 表情语（expressives）
  - 宣布（declarations）
- By MultiWOZ:

Table 2: Full ontology for all domains in our data-set. The upper script indicates which domains it belongs to. *: universal, 1: restaurant, 2: hotel, 3: attraction, 4: taxi, 5: train, 6: hospital, 7: police.

act type	inform* / request* / select^{123} / recommend/^{123} / not found^{123} request booking info^{123} / offer booking^{1235} / inform booked^{1235} / decline booking^{1235} welcome* /greet* / bye* / reqmore*
slots	address* / postcode* / phone* / name^{1234} / no of choices^{1235} / area^{123} / pricerange^{123} / type^{123} / internet^{2} / parking^{2} / stars^{2} / open hours^{3} / departure^{45} destination^{45} / leave after^{45} / arrive by^{45} / no of people^{1235} / reference no.^{1235} / trainID^{5} / ticket price^{5} / travel time^{5} / department^{7} / day^{1235} / no of days^{123}

Speech Acts: Statistics in LDC online chat corpus

Table 2: Speech act taxonomy and frequencies in the LDC online chat corpus

Classification	Percent	Example
Statement	34.50%	10-19-40sUser11...some people have a lot of blank pages
System	17.02%	JOIN
Greet	13.40%	Hey You
Emotion	11.52%	lmao
Wh-Question	5.33%	where from@11-09-adultsUser12
Yes/No Question	5.22%	wisconsin?
Continuer	3.48%	but i didnt chance it
Accept	2.45%	ok
Reject	2.14%	I can't do newspaper.. I can't throw that far and stairs give me problems
Bye	1.57%	goodnite
Yes Answer	1.17%	yeah
No Answer	0.94%	nope 11-09-adultsUser27
Emphasis	0.48%	Ok I'm gonna put it up ONE MORE TIME 10-19-30sUser37
Other	0.43%	0
Clarify	0.34%	i mean the pepper steak lol

Moldovan et al., Automated Speech Act Classification For Online Chat, MAICS, 2011, 710: 23-29.

User Intent Modeling is far from satisfactory in Dialog Systems

- In task-oriented dialog systems:
  - Although User Intent is explicitly modeled, there is still far to go...
  - A disastrous example: 山东大汉大战语音导航
- In chit-chat dialog systems:
  - There is not user intent modeling at all.

Content

UniDS: A Unified Dialogue System for Chit-Chat and Task-oriented Dialogues

	Unified dialogue data schema	Chit-chat example	Task-oriented example
User input	Tokenized utterance	does money buy happiness ?	i am looking for a cheap hotel .
Belief state	<domain> slot [value]	<chit> money happiness	<hotel> price cheap
DB result	A token indicated the number of candidate entities	<db_nore>	<db_2>
Act	<domain> <act> [slot]	<chit> <chit_act>	<hotel> <request> area
Response	Tokenized utterance	depends on how much money you spend on it .	do you have a specific area you want to stay in ?

Table 2: Unified dialogue data schema (where tokens inside the square bracket are optional) and examples.

Published in DialDoc 2022: https://aclanthology.org/2022.dialdoc-1.2

UniDS: A Unified Dialogue System for Chit-Chat and Task-oriented Dialogues

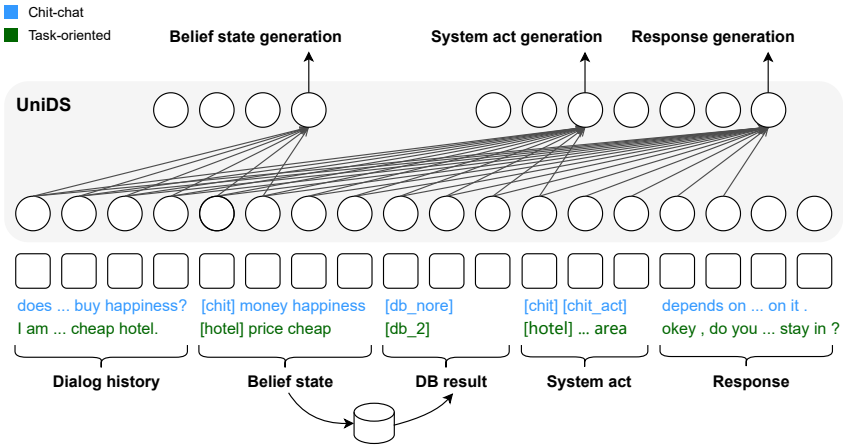


Figure 2: The architecture of UniDS.

Published in DialDoc 2022: https://aclanthology.org/2022.dialdoc-1.2

UniDS: A Unified Dialogue System for Chit-Chat and Task-oriented Dialogues

Model	# of para.	Task-oriented Dialogue				Chit-chat			
		Inform	Success	BLEU	Combined	BLEU	Dist-1	Dist-2	AvgLen
UBAR*	82M	91.5	77.4	17.0	101.5	-	-	-	-
PPTOD	∼220M	89.20	79.40	18.62	102.92	-	-	-	-
UBAR-12L	117M	**89.40**	75.10	16.93	99.18	-	-	-	-
DialoGPT-12L	117M	-	-	-	-	0.27	**6**	**32**	14.00
UniDS-12L	117M	87.10	**77.00**	**18.01**	**100.06**	**0.35**	**6**	30	12.00
UBAR-24L	345M	89.40	75.50	16.86	99.31	-	-	-	-
DialoGPT-24L	345M	-	-	-	-	0.43	**7**	**36**	12.28
UniDS-24L	345M	**90.30**	**80.50**	**18.72**	**104.12**	**0.45**	6	35	14.62

Table 3: Automatic evaluations of UniDS with two model sizes over two types of dialogue datasets. All results are reported in percentage, except Combined and AvgLen. Best results are in **bold**. *: Results reported in original paper (Yang et al., 2021) is not obtained by end-to-end evaluation. This result is reported by authors of UBAR in https://github.com/TonyNemo/UBAR-MultiWOZ/issues/3.

Published in DialDoc 2022: https://aclanthology.org/2022.dialdoc-1.2

UniDS: A Unified Dialogue System for Chit-Chat and Task-oriented Dialogues

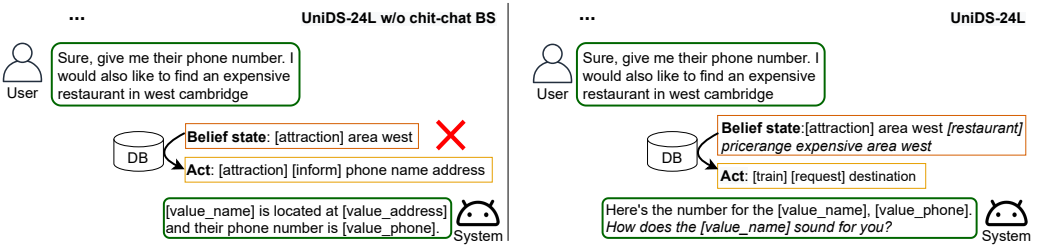


Figure 4: TOD examples from UniDS w/o chit-chat BS and UniDS. UniDS w/o chit-chat BS does not extract the user intent of searching restaurants, but UniDS extracts this intent successfully (highlighted in italics).

Published in DialDoc 2022: https://aclanthology.org/2022.dialdoc-1.2

Content

Some Thoughts of Future Directions

- Unified Ontology of Domains-Intents-slots
- Intent modeling should be also applied to chit-chat dialog
- The definition of intents should be more fine-grained and has higher coverage.
- Data annotation will be a problem.

Content

Summary

Thank you!

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。

Bring digital to every person, home and organization for a fully connected, intelligent world.

HUAWEI