



基于策略建模的任务型对话预训练研究与应用

开题汇报

汇报人： 钟鹭岑

指导老师：王小捷

汇报时间：2022年11月18日

目 录

1. 选题背景和意义
2. 研究现状和问题
3. 研究内容和目标
4. 前期工作和未来规划

Part 01

选题背景和意义

■ 选题背景（一）：预训练模型

- 2017年Transformer被提出后，基于该架构的预训练模型广泛应用于许多NLP应用中
 - 常用的预训练语言模型结构：Encoder-only（BERT）、Decoder-only（GPT）、Encoder-Decoder（T5、BART）
 - 这些模型利用大规模的文本数据进行自监督的预训练，通过进一步微调在下游任务上取得很好的性能

- 随着预训练模型的发展，对话领域的研究也逐渐开始关注基于预训练的端到端对话系统，2020-2022这三年的时间涌现出很多关于开放域对话系统预训练的相关研究
 - 基于英文的包括google在2020年1月发表的Meena、Facebook在4月发表的Blender
 - 基于中文的主要以百度PLATO系列模型为代表
 - 这些模型的成功一定程度上表明海量数据和大模型能为对话系统带来很好的收益

■ 选题背景（二）：任务型对话系统

- 随着大规模预训练模型的发展，越来越多的工作将预训练模型用于任务型对话系统，将对话管理和回复生成建模在统一的预训练模型下，如：UBAR、MMTOD和BOTR等
- 这些工作都直接在文本预训练模型上进行微调，通过设计辅助任务来提高预训练模型在任务型对话上的性能

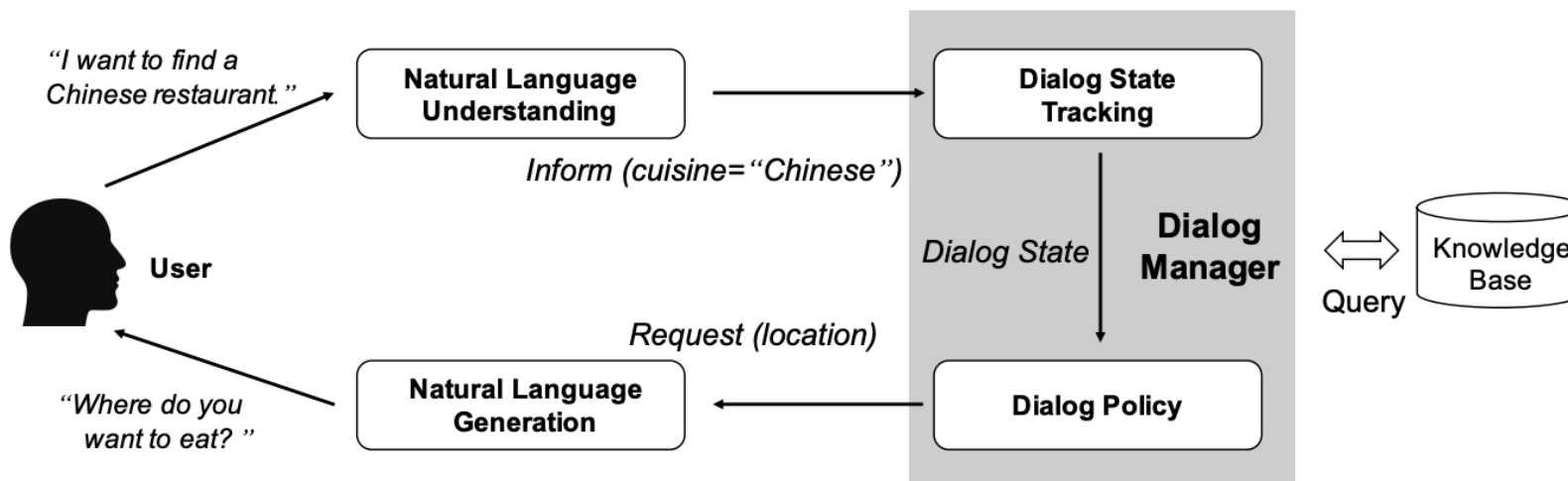


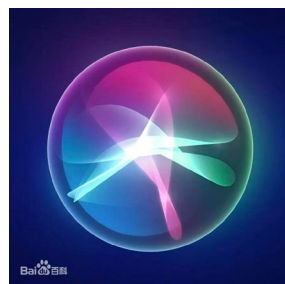
图1. 基于Pipeline方法的任务型对话系统的总体框架

■ 选题背景（三）：任务型对话预训练



- 文本预训练模型在对话数据上直接微调效果不好：
 - 数据分布差异：文本数据和对话数据的语言模式不同
 - 角色差异：文本数据来自同一个作者，而对话至少存在两个说话人
- 闲聊型对话预训练模型在任务型对话数据集上直接微调效果不好：
 - 对话目标差异：闲聊型对话没有具体的目标、主题；任务型对话通常有一个特定的对话目标
 - 数据形式差异：闲聊型对话一般数据量大、轮次少、噪音较大；任务型对话数据通常包含多轮交互、且数据量较小
- 越来越多的工作开始关注针对任务型对话的预训练模型，在预训练阶段设计相关的预训练任务来挖掘任务型对话中的任务信息，如GALAXY、SPACE3.0等

- 任务型对话系统具有明确的用户目标和系统行为，对其语言理解和策略学习比闲聊型对话更为重要。任务型对话预训练让模型提前学习一些通用的任务相关知识，从而在下游任务上取得更好的性能，具有较高的学术价值。
- 任务型对话系统可以帮助用户便捷处理复杂任务，减轻人工客服的负担，例如：Apple Siri、阿里小蜜等，具有较高的应用价值。



难点一

数据量小且标注困难

- 预训练模型参数量巨大，通常需要非常多的数据才能使模型学到相应的知识
- 任务型对话数据标注费时费力，高质量的有标数据集非常少

难点二

对话中任务信息的建模

任务型对话中包含很多任务相关的信息，如：对话结构、对话策略等。如果可以通过**设计特定的预训练任务**使得模型可以在预训练阶段学习到这些信息，就可以对下游任务起到辅助作用

难点三

预训练模型的跨任务迁移

- 任务型对话所涉及的任务多种多样，预训练过程中没有办法完全覆盖下游的所有任务
- 少样本设定下任务型对话预训练模型如何取得好的性能

Part 02

研究现状和问题

■ 现有研究（一）

TOD-BERT: 面向对话理解和回复
选择任务，联合9个任务型对话数
据集对BERT进行进一步的预训练

2020.04

SOLOIST: 提出了通过预训练一
个通用的端到端（end-to-end）任
务型对话系统，并通过迁移学习
和机器学习使其适应新的任务

2020.05

2020.02

SC-GPT: 针对基于管道方法
（pipeline method）的任务型
对话系统中的自然语言生成
（NLG）模块进行预训练

2020.04

PRAL: 处理了13个对话数据集，
基于GPT的双编码器模型进行预
训练，并在回复生成任务上进行
进行评测

■ 现有研究 (二)

Galaxy: 基于UniLM的显式建模
对话策略的半监督任务型对话生成式预训练模型

2021.11

SPACE 3.0: 基于UniLM的统一
对话理解、对话策略和对话生成
的端到端任务型对话系统, 且模型维持了对话中的任务流信息

2022.06

2021.09

PPTOD: 以T5模型为基础, 采用
prompt方法构建了一个通用的任务型对话预训练模型

2022.03

SPACE 2.0: 基于UniLM的建模
语义树结构信息的对话理解的预训练模型

2022.09

OPAL: 利用了一个外部的工具
OpenIE来对Wikipedia corpus进行了
实体的抽取, 用来模拟对话状态。
分别利用chit-chat、TOD数据进行了两阶段的预训练

现有研究发展趋势

➤ pipeline-based（管道方法） → end-to-end（端到端方法）

2020年主要基于pipeline系统中的某个模块进行预训练，2021年之后都转为端到端的预训练。

➤ 专用预训练模型 → 通用预训练模型

大部分模型只能完成TOD中某一个特定的任务，而最新的PPTOD、SPACE3.0可以完成多个下游任务

➤ 单一数据预训练 → 混合数据预训练 → 外部知识库/工具

最开始只使用TOD数据，之后很多工作倾向于使用闲聊+TOD数据同时预训练，最新的OPAL更进一步用外部工具对闲聊数据进行了“改造”，使其更像任务型数据

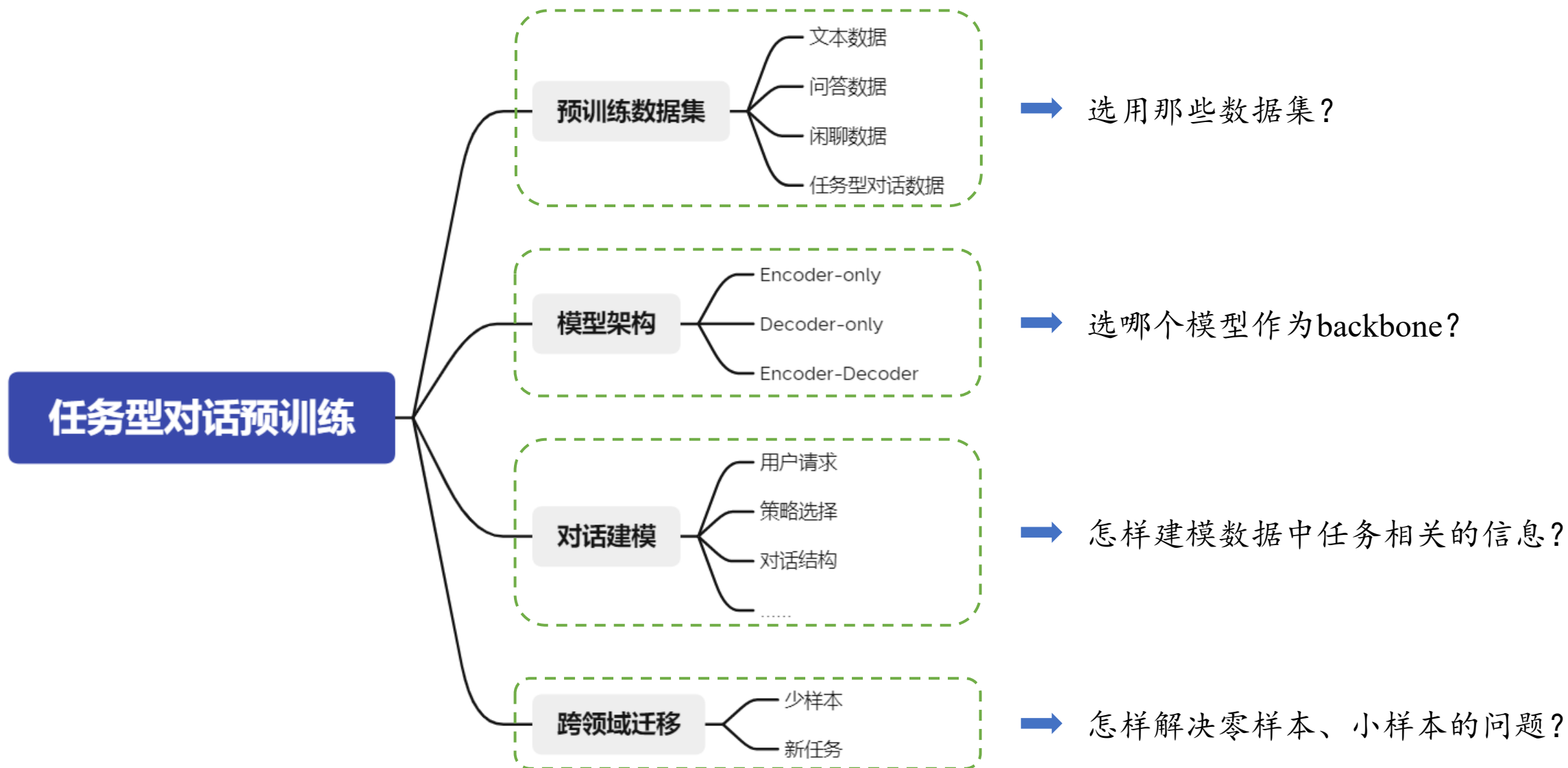
未来可能的研究方向

➤ 数据方面：通过外部知识库引入知识，通过外部工具构建更多的训练数据

➤ 任务方面：设计预训练任务建模任务型对话的任务信息，如：对话策略、对话结构等

➤ 领域方面：模型在跨领域/跨任务迁移上的研究

■ 任务型对话研究归类



■ 预训练数据集

模型	预训练数据	贡献点
SC-GPT	4个有对话策略标注的TOD数据集	• 首次提出用TOD数据集对NLG模型进行预训练
TOD-BERT	9个TOD数据集（不需要标注信息）	• 联合多个TOD数据集进行自监督预训练
PRAL	13个高质量的ChitChat、TOD数据集	• 首次同时利用ChitChat、TOD数据集进行自监督预训练
SOLOIST	2个TOD数据集（需要对话状态标注）	• /
PPTOD	11个部分标注的TOD数据集	• 首个提出使用部分标注数据进行预训练
GALAXY	8个TOD数据集（有统一的DA标注） + 14个ChitChat数据集	• 提出了统一的DA标注方法，构建了一个有统一DA标注的数据集（UniDA） • 收集了一个大规模无标注闲聊对话数据集（UnDial）
SPACE 2.0	32个TOD数据集（带有结构语义标注信息） + 21个ChitChat数据集	• 构建了一个带有结构语义标注的数据集（AnPreDial） • 构建了一个大规模的无标签对话语料（UnPreDial）
SPACE 3.0	同SPACE 2.0	• /
OPAL	5个TOD数据集+Wikipedia corpus	• 利用外部工具抽取闲聊型对话中的ontology-like知识，用这种知识来模仿任务型对话中的对话状态

■ 模型架构（一）：编码器

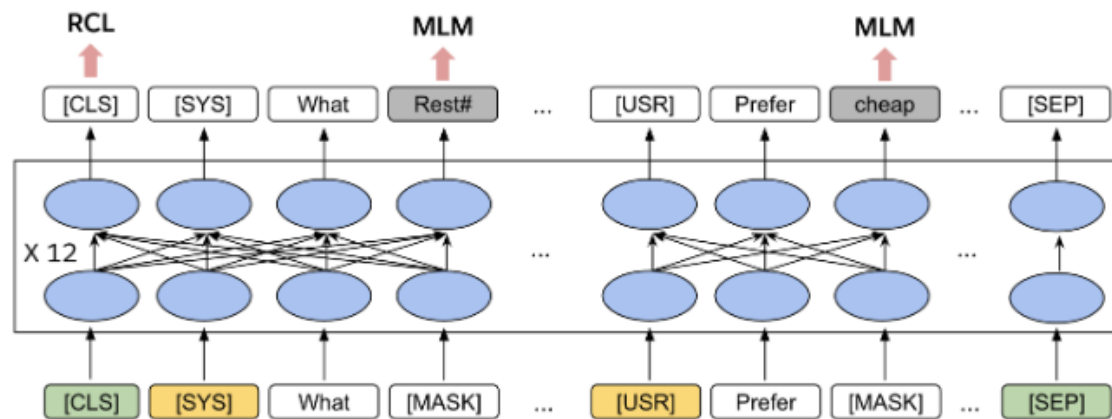
Encoder-only

相关工作

TOD-BERT

架构分析

- 基于Transformer的双向编码器结构
- 只能做对话理解类任务，如：意图识别、对话状态追踪
- 缺点：限制性比较强，不通用



■ 模型架构 (二) : 解码器

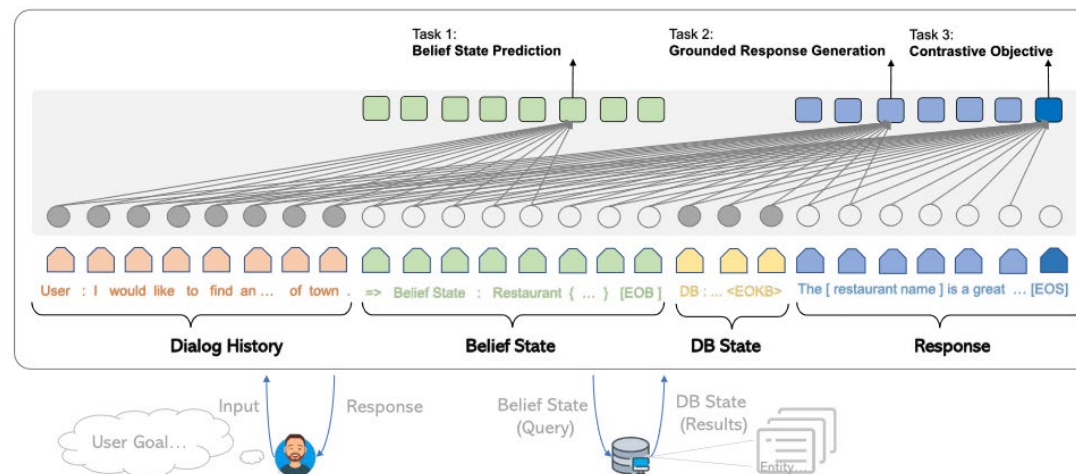
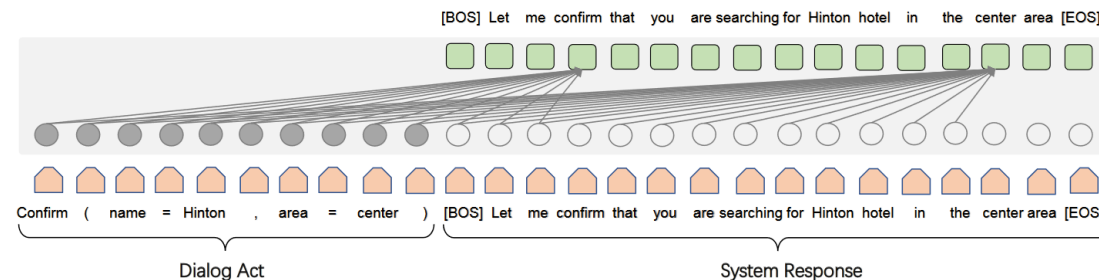
Decoder-only

相关工作

SC-GPT、PRAL、SOLOIST

架构分析

- 基于Transformer的自回归结构，天然地适用于回复生成任务
- 对于对话理解任务：一方面可以将其建模为生成式任务；另一方面可以利用最后一个token的表示作为全局表示进行后续任务
- 缺点：对话历史的编码是单向的



■ 模型架构（三）：编码器-解码器

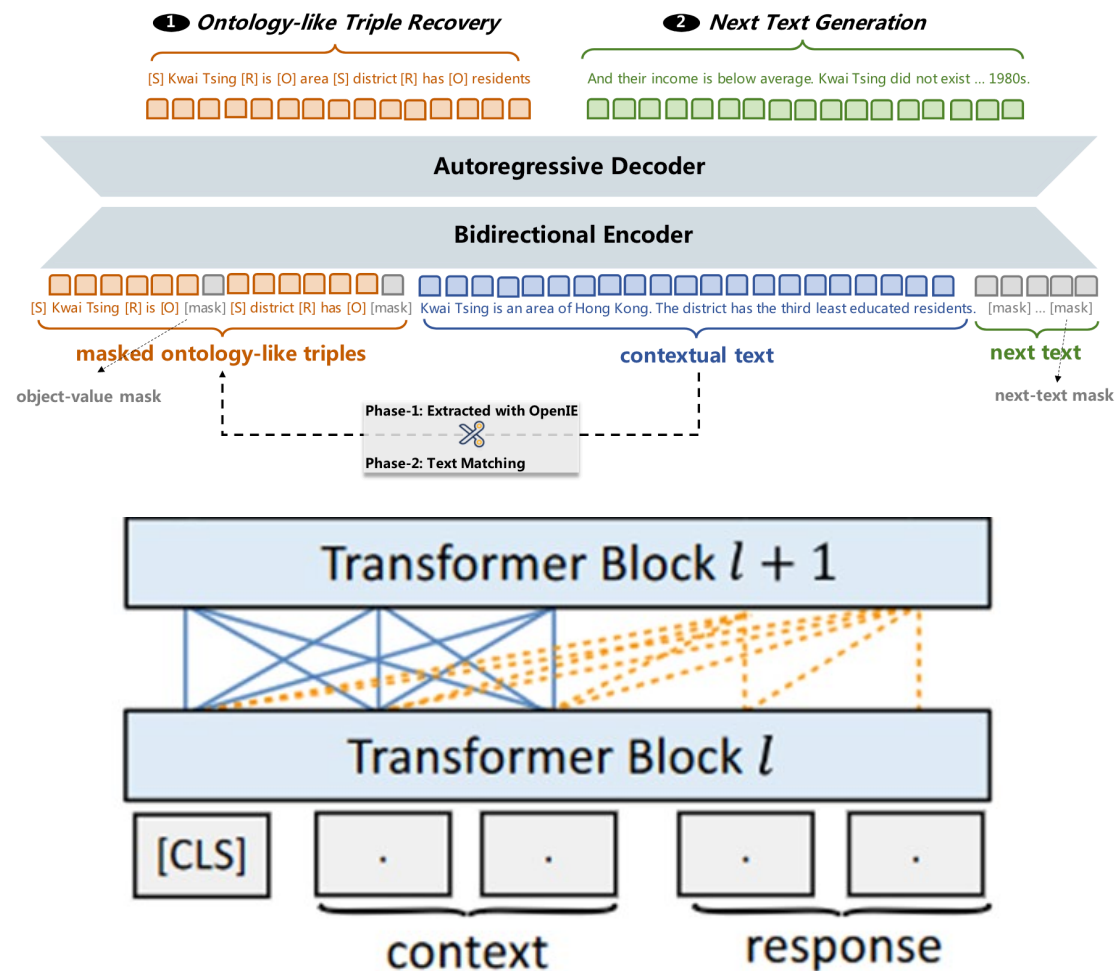
Encoder-Decoder

相关工作

PPTOD、SPACE 1.0、SPACE 3.0、OPAL

架构分析

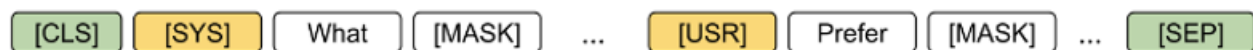
- T5和BART是基于Transformer原始的encoder-decoder架构；UniLM基于不同的attention mask机制实现了encoder-decoder架构
- 适用于任务型对话的各种下游任务
- 目前最新的工作均采用这种架构，结构上更合理



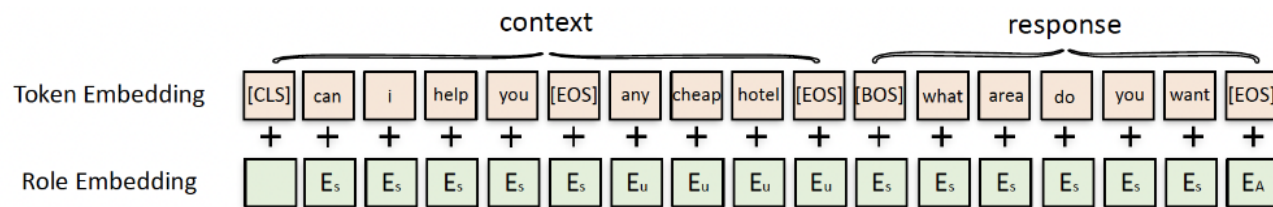
■ 对话建模（一）：角色信息建模

角色信息建模

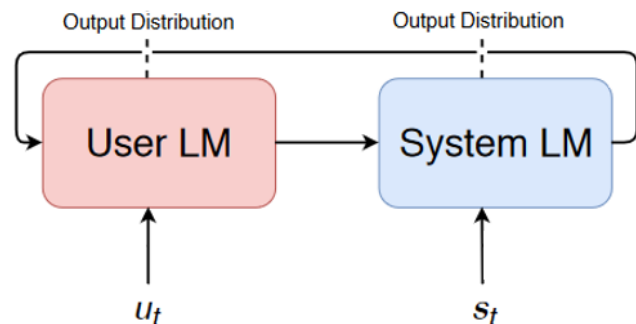
- 在不同角色的话语前添加特殊token，如：[SYS]、[USER]



- 添加角色编码信息，即role embedding



- 用两个编码器分别编码系统和用户的话语



TOD-BERT
SOLOIST
PPTOD
OPAL



GALAXY
SPACE 2.0
SPACE 3.0

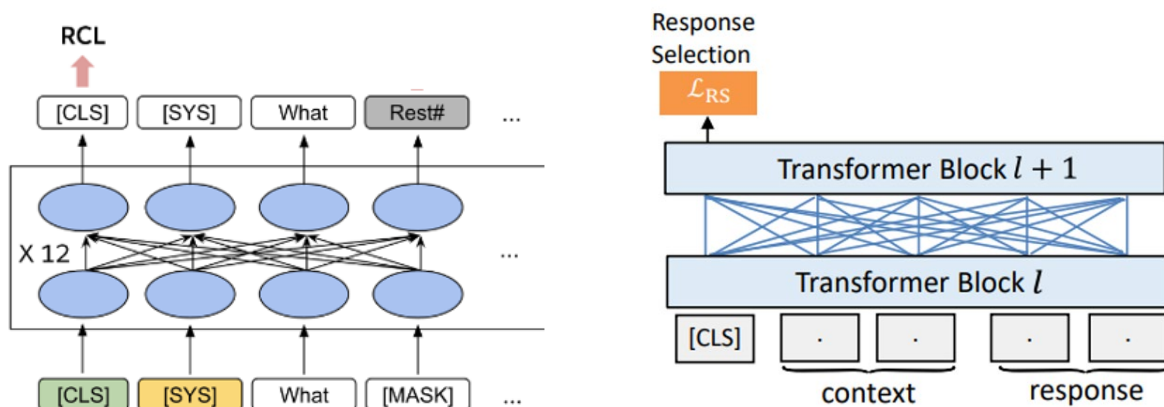


PRAL

■ 对话建模（二）：预训练任务

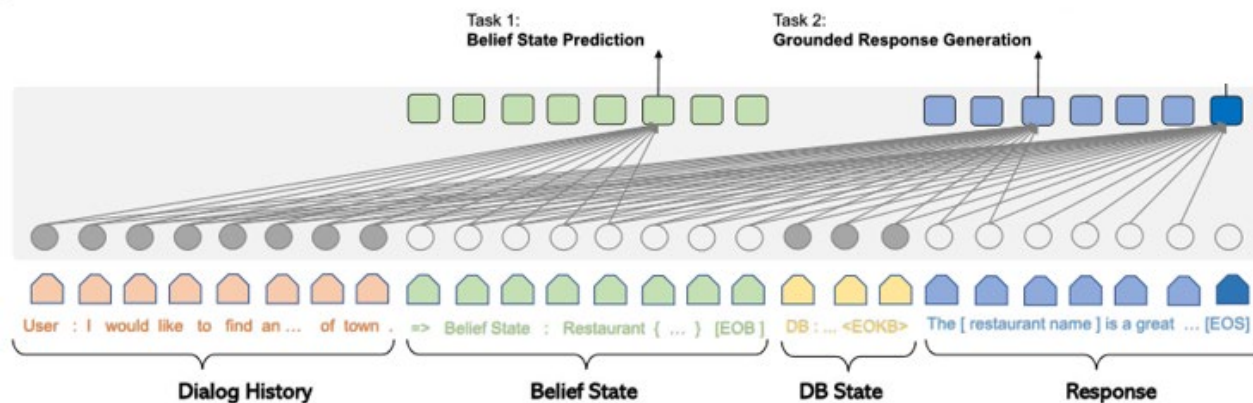
预训练任务——面向下游任务设计

- Context-Response匹配损失函数模拟回复选择任务



TOD-BERT
GALAXY

- 序列生成任务模型回复生成、对话状态追踪等任务



SC-GPT
SOLOIST
GALAXY
SPACE 3.0
OPAL

■ 对话建模（三）：预训练任务

预训练任务——对话理解

➤ MLM及其变种

- TOD-BERT中使用原始的MLM来学习更好的表示

$$\mathcal{L}_{mlm} = - \sum_{m=1}^M \log P(x_m)$$

- SPACE 3.0进一步对于有标注数据Mask掉槽值信息，对于无标注数据随机Mask span进行预训练，希望模型可以学习更多的任务相关的信息

$$\mathcal{L}_{slm} = - \sum_{\hat{c} \in m(c)} \log p(\hat{c} | c_{\setminus m(c)})$$

➤ Bag-of-words Loss: 丢弃单词的顺序，并以非自回归的方式预测无序的系统回复

$$\mathcal{L}_{\text{bow}}^r = -\log p(r_{\text{bow}} | c, r, p^u) = -\log \prod_{n=1}^N H''_{\text{index}(r_n)}$$
$$H'' = \text{softmax}(W_2 h^r + b_2) \in \mathbb{R}^{|V|}$$

■ 对话建模（四）：预训练任务

预训练任务——对话理解

➤ 对比学习：学习更好的上下文表示

■ SOLOIST：学习对话历史与对话状态、系统回复之间的对应关系

□ 具体而言，对于正样本 x ，从数据集中随机采样对话状态、系统回复作为负样本 x' ，对正负样本进行0-1分类，采用交叉熵损失函数进行约束：

$$\mathcal{L}_c = y \log((p_\theta x) + (1 - y) \log(1 - p_\theta(x')))$$

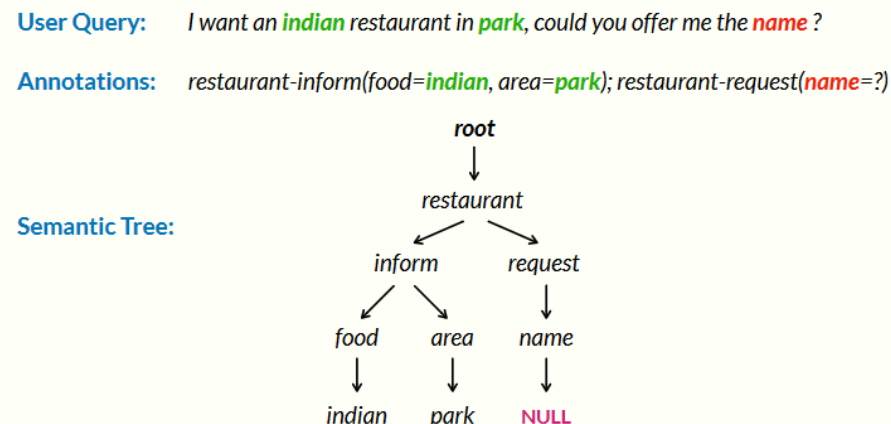
■ SPACE 2.0/3.0：学习对话历史之间的差异性

□ 有标注数据：tree-induced supervised contrastive objective

$$\mathcal{L}_{\text{sup}}^q = - \sum_{i \in I} \sum_{j \in I} \frac{f_{i,j}^q}{\sum_{v \in I} f_{i,v}^q} \log \frac{\exp(\sigma(h_i^q) \cdot \sigma(h_j^q) / \tau)}{\sum_{l \in I, l \neq i} \exp(\sigma(h_i^q) \cdot \sigma(h_l^q) / \tau)}$$

□ 无标注数据：vanilla self-supervised contrastive objective

$$\mathcal{L}_{\text{self}}^q = - \sum_{i \in I} \sum_{j \in I} \mathbf{1}(j = i^+) \log \frac{\exp(\sigma(h_i^q) \cdot \sigma(h_j^q) / \tau)}{\sum_{l \in I, l \neq i} \exp(\sigma(h_i^q) \cdot \sigma(h_l^q) / \tau)}$$



■ 对话建模（五）：对话策略

预训练任务——对话策略

➤ GALAXY：有监督的系统动作预测

- 对于有对话动作标注的数据，根据对话历史 c 来预测系统的动作 a ，是一个多标签分类任务

➤ GALAXY：KL散度损失，约束对话策略的分布

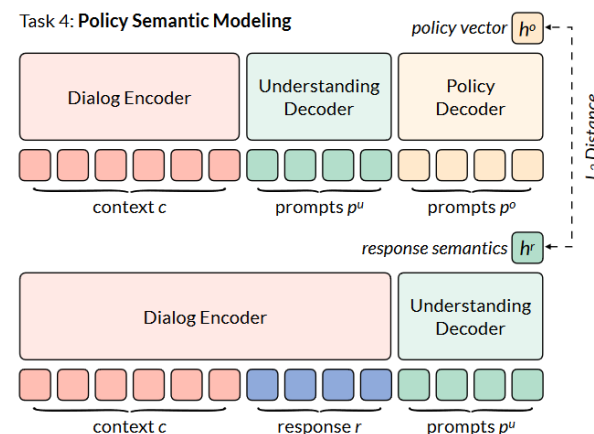
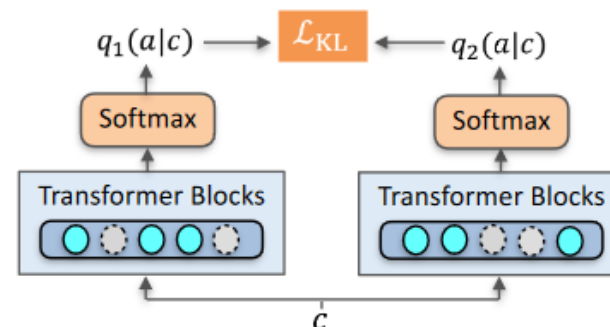
- 具体而言，将一段对话历史两次输入编码器，由于dropout扰动会得到两个不同的分布，采用KL loss来最小化这两个分布之间的距离：

$$\mathcal{L}_{KL} = \frac{1}{2} (\mathcal{D}_{KL}(q_1 || q_2) + \mathcal{D}_{KL}(q_2 || q_1))$$

➤ SPACE 3.0：对话策略的先验概率应该和后验概率接近

- 最小化先验和后验之间的L2距离

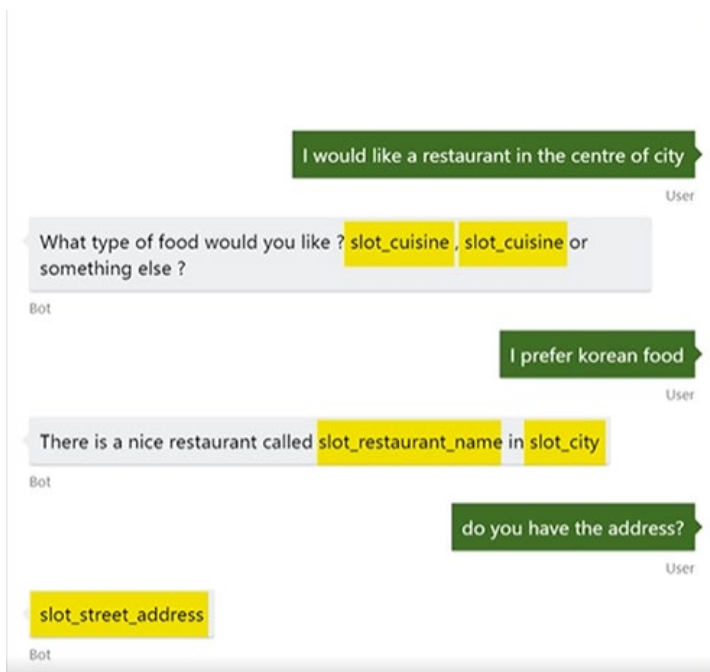
$$\mathcal{L}_{psm} = \|h^o - h^r\|_2^2$$



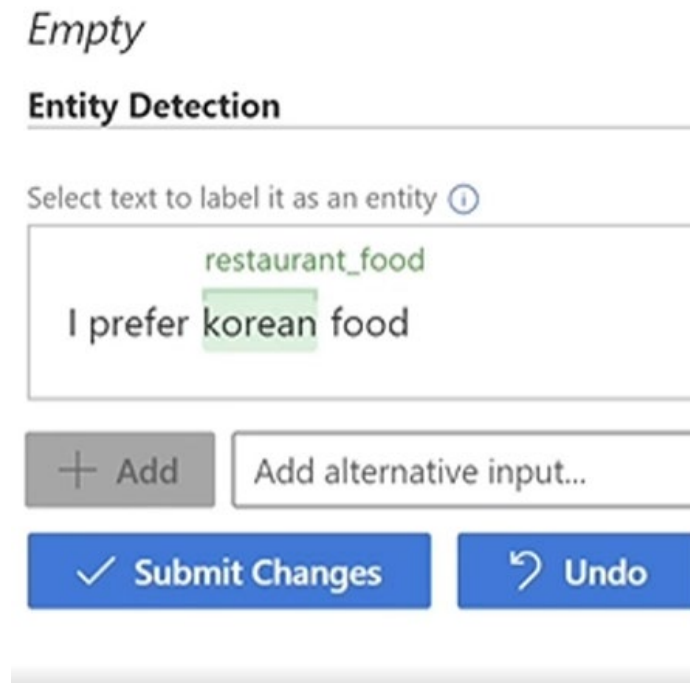
- 预训练的优势更加体现在少样本设定下，经过大规模对话数据集进行预训练后的模型性能相较于没有预训练过的模型性能有非常明显的优势

	Model	5%	10%	20%	50%	100%
UnPre-trained	MinTL	82.20	87.93	88.52	102.25	97.79
	UBAR	82.69	90.4	92.27	101.90	101.50
Pre-trained	PPTOD	86.56	91.96	95.33	/	102.92
	SPACE 1.0	91.40	98.65	100.24	106.22	110.35

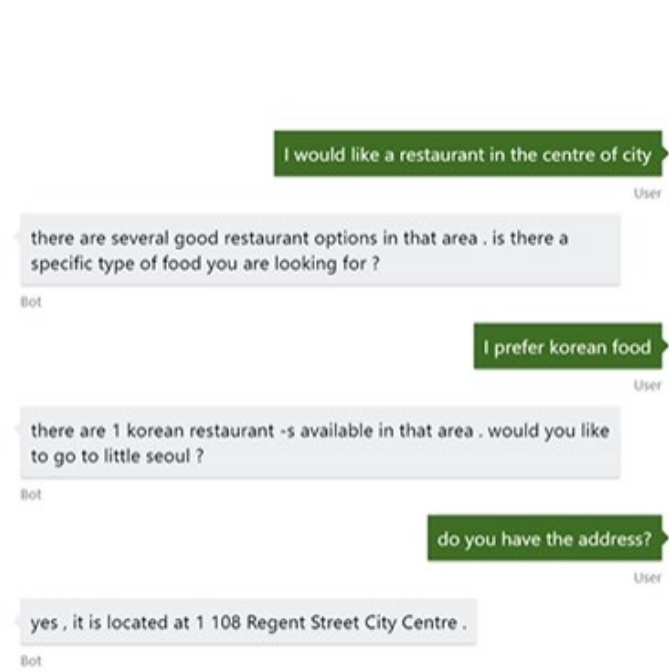
- SOLOIST中采用Machine Teaching^[1]的方式来将模型迁移到新的任务上



(a) human-bot dialog



(b) belief-state annotation



(c) new human-bot dialog

[1]. <https://www.microsoft.com/en-us/research/project/conversation-learner/>

- 从预训练数据的角度看，目前TOD预训练模型主要关注的是如何在预训练过程中怎样合理利用数据来辅助下游任务，这其中包括：
 - 如何利用有标注（完全标准 & 部分标注）任务型对话数据（TOD-BERT、SPACE、PPTOD）
 - 如何利用无标注任务型对话数据（SPACE）
 - 如何利用大规模的闲聊数据，以及怎样让闲聊数据更接近任务型对话数据（SPACE、OPAL）
- 从模型架构的角度看，从最开始只能完成特定的下游任务，到目前开始出现统一了对话理解和对话生成的架构，TOD预训练模型在向着更通用的方向发展（encoder-decoder架构是目前的主流）
- 从预训练任务角度看，除了传统的学习文本表层语义的任务（eg. MLM），TOD预训练目前更关注如何根据TOD特点来设计相关预训练任务，以此来更好的学习对话数据中的任务相关的知识
- 从各个模型的实验结果可以看出，利用大规模数据集进行预训练后的模型性能在少样本的设定下，可以有更显著的性能提升，因此研究TOD预训练可以一定程度上缓解下游任务数据不足的问题

Part 03

研究内容和目标

- 问题一：任务型对话数据的标注费时费力，现有的工作不断增加预训练的数据量，预训练的成本非常高昂，且大部分工作对数据标注的利用不充分，又较大的改进空间
- 问题二：目前模型在对话策略预测的准确率上依然很低（完全正确的比例不足50%），现有的策略建模方式比较单一，且对话策略本身的特点还未被完全探索，策略建模还有很大的提升空间
- 问题三：现有的模型只在公共数据集上进行离线的评测，并没有在实际业务中进行线上评测，也没有考虑实际业务中系统回复可能存在的问题

■ 研究目标

目标一

基于现有的任务型对话数据的标注信息，设计新的预训练方式来**提高标注数据的利用效率**，充分建模标注数据之间的逻辑关系

目标二

分析对话策略本身的特点，设计相关的预训练任务来**从各个维度建模对话策略**，以提高预训练模型的策略预测准确率，进而提高回复生成的质量。

目标三

基于所提出的模型，在**现实场景下进行测试**，分析在真实对话业务中存在的问题，并针对具体的业务对进行模型的改进

Part 04

前期工作和未来规划

- 提出了task-progressive预训练框架，使模型可以循序渐进地学习各个子任务，避免了之前工作中同时学习三个任务难度较大的问题
- 设计了两个基于策略的预训练任务：Global Policy Consistent Task 和 Act-based Contrastive Learning Task
- 在MultiWOZ和In-Car两个任务型对话数据的端到端评测上取得了SOTA性能

Model	Inform	Success	BLEU	Comb
<i>w/o pre-train</i>				
UBAR	83.4	70.3	17.6	94.4
BORT	85.5	77.4	17.9	99.4
MMTOD	85.9	76.5	19.0	100.2
<i>w/ pre-train</i>				
SOLOIST	82.3	72.4	13.6	90.9
PPTOD	83.1	72.7	18.2	96.1
GALAXY	85.4	75.7	19.64	100.2
TPPA	89.5	77.6	18.68	102.23

Table 2: E2E performances on MultiWOZ. All results are adopted from the official MultiWOZ benchmark ².

Model	Match	SuccF1	BLEU	Comb
<i>w/o pre-train</i>				
SEDST	84.5	82.9	19.3	103.00
TSCP	84.5	81.1	21.9	104.70
LABES	85.8	77.0	22.8	104.20
FSDM	84.8	82.1	21.5	104.95
<i>w/ pre-train</i>				
GALAXY	85.3	83.6	23.0	107.45
SPACE-3	85.2	83.1	22.9	107.13
TPPA	86.2	83.9	23.6	108.69

Table 3: E2E performances on In-car. All results are from original papers.

■ 2022.11-2022.12

完善后续实验：Few-shot Setting、Visualization、Case Study

■ 2023.01-2023.03

了解美团的业务场景，尝试在业务数据上对现有的方法进行实验，分析存在的问题并改进
目前了解到最大的问题是轮次间的策略重复问题，这类问题原因是session-level的知识没有办法进行建模，比如：对话本身更远轮次的长期回报、对话在session-level的分类标签

■ 2023.04-2023.06

从业务场景中抽象出具体的科研问题，进行开放数据集的评测
例如：如何解决任务型对话系统中的重复生成问题

1. Zhang Z, Takanobu R, Zhu Q, et al. Recent advances and challenges in task-oriented dialog systems[J]. Science China Technological Sciences, 2020, 63(10): 2011-2027.
2. Yang Y, Li Y, Quan X. Ubar: Towards fully end-to-end task-oriented dialog system with gpt-2[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2021, 35(16): 14230-14238.
3. Lee Y. Improving End-to-End Task-Oriented Dialog System with A Simple Auxiliary Task[C]//Findings of the Association for Computational Linguistics: EMNLP 2021. 2021: 1296-1303.
4. Sun H, Bao J, Wu Y, et al. BORT: Back and Denoising Reconstruction for End-to-End Task-Oriented Dialog[J]. arXiv preprint arXiv:2205.02471, 2022.
5. Aron J. How innovative is Apple's new voice assistant, Siri?[J]. 2011.
6. Li F L, Qiu M, Chen H, et al. Alime assist: An intelligent assistant for creating an innovative e-commerce experience [C]//Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. 2017: 2495-2498.
7. Peng B, Zhu C, Li C, et al. Few-shot Natural Language Generation for Task-Oriented Dialog[C]//Findings of the Association for Computational Linguistics: EMNLP 2020. 2020: 172-182.
8. Chen Z, Liu Y, Chen L, et al. OPAL: Ontology-Aware Pretrained Language Model for End-to-End Task-Oriented Dialogue[J]. arXiv preprint arXiv:2209.04595, 2022.

9. Wu C S, Hoi S C H, Socher R, et al. TOD-BERT: Pre-trained Natural Language Understanding for Task-Oriented Dialogue[C]//Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). 2020: 917-929.
10. Gu J, Wu Q, Wu C, et al. A tailored pre-training model for task-oriented dialog generation[J]. arXiv preprint arXiv:2004.13835, 2020.
11. Peng B, Li C, Li J, et al. SOLOIST: Few-shot Task-Oriented Dialog with A Single Pre-trained Auto-regressive Model[J].
12. Su Y, Shu L, Mansimov E, et al. Multi-Task Pre-Training for Plug-and-Play Task-Oriented Dialogue System[C]//Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2022: 4661-4676.
13. He W, Dai Y, Zheng Y, et al. Galaxy: A generative pre-trained model for task-oriented dialog with semi-supervised learning and explicit policy injection[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2022, 36(10): 10749-10757.
14. He W, Dai Y, Yang M, et al. Unified Dialog Model Pre-training for Task-Oriented Dialog Understanding and Generation[C]//Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2022: 187-200.

感谢聆听，期待批评指正

Thanks for Listening

CIST of BUPT

汇报人： 钟鹭岑

汇报时间：2021年11月18日

问题：

- PPT字太多，语速太快，不需要所有说出来的内容都要在PPT上
- 未来规划只写到了6月，应该规划到中期答辩、大论文撰写等
- 大论文中还需要说明一下增大模型参数量和数据量为什么不能继续提升性能，是否是任务型对话本身的原因？