Semi-Supervised Knowledge-Grounded Pre-training for Task-Oriented Dialog Systems

Weihao Zeng[1*], **Keqing He**[2*], **Zechen Wang**[1*], **Dayuan Fu**[1], **Guanting Dong**[1]
Ruotong Geng[1], **Pei Wang**[1], **Jingang Wang**[2], **Chaobo Sun**[2], **Wei Wu**[2], **Weiran Xu**[1*]
[1]Beijing University of Posts and Telecommunications, Beijing, China
[2]Meituan, Beijing, China
`{zengwh,zechen_wang,fdy,dongguanting,ruotonggeng,wangpei,xuweiran}@bupt.edu.cn`
`{hekeqing,wangjingang,sunchaobo,wuwei}@meituan.com`

Abstract

Recent advances in neural approaches greatly improve task-oriented dialogue (TOD) systems which assist users to accomplish their goals. However, such systems rely on costly manually labeled dialogs which are not available in practical scenarios. In this paper, we present our models for Track 2 of the SereTOD 2022 challenge, which is the first challenge of building semi-supervised and reinforced TOD systems on a large-scale real-world Chinese TOD dataset MobileCS. We build a knowledge-grounded dialog model to formulate dialog history and local KB as input and predict the system response. And we perform semi-supervised pre-training both on the labeled and unlabeled data. Our system achieves the first place both in the automatic evaluation and human interaction, especially with higher BLEU (+7.64) and Success (+13.6%) than the second place.[1]

1 Introduction

Task-oriented dialogue (TOD) systems assist users to accomplish their goals like booking a ticket and make an effect on everyone's lives with recent advances in neural approaches (Gao et al., 2018). A typical TOD system consists of three sub-modules: (1) natural language understanding (NLU) for recognizing the user's intent and slots (Goo et al., 2018; Qin et al., 2019; He et al., 2020a; Xu et al., 2020; He et al., 2020b); (2) dialog management (DM) for tracking dialog states (Wu et al., 2019; Gao et al., 2019) and deciding which system action to take (Peng et al., 2018; Liu et al., 2021); (3) natural language generation (NLG) for generating dialogue response corresponding to the predicted system action (Peng et al., 2020). Traditional modular methods (Goo et al., 2018; Wu et al., 2019; Peng et al., 2020) and recent end-to-end modeling methods (Peng et al., 2021; Su et al., 2022; Liu et al., 2022a) achieve decent performance in several or all modules. However, such systems rely on costly manually labeled dialogs which are not available in practical scenarios. It's valuable to explore semi-supervised learning (SSL) (Zhu, 2005) for TOD, which aims to leverage both labeled and unlabeled data.

To facilitate relevant research, SereTOD 2022 Workshop [2] proposes the first challenge of building semi-supervised and reinforced TOD systems by releasing a large-scale Chinese TOD dataset MobileCS from real-world dialog transcripts between real users and customer-service staffs from China Mobile. MobileCS contains 10,000 labeled dialogs and 90,000 unlabeled dialogs. There are two tracks: (1) Information extraction (Track 1) aims to extract entities together with their slot values. (2) Task-oriented dialog system (Track 2) aims to build a complete TOD system, including predicting the user intent, querying the local KB, and generating appropriate system intent and response according to the given dialog history. The core challenge is how to combine a small labeled dataset and a large unlabeled dataset.

In this paper, we present our system for Track 2 of the SereTOD 2022 challenge. The main intuition behind our system comes from semi-supervised knowledge-grounded pre-training on both labeled and unlabeled datasets. We divide Track 2 into two task groups, classification (user intent and system intent) and generation (system response). For the classification tasks, we employ Roberta-large [3] and build two separate classification models. We also perform continual pre-training on all the dialog data. For the generation task, we build a knowledge-grounded dialog model, which is the key point of this paper. Specifically, we firstly use

---

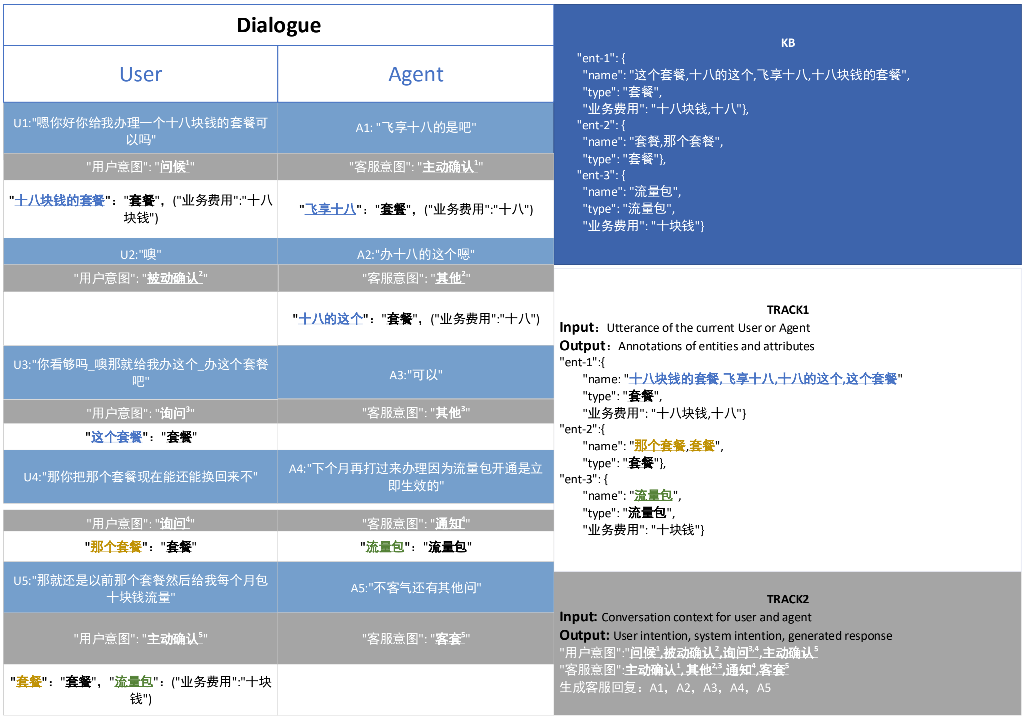
[*] The first three authors contribute equally. Weiran Xu is the corresponding author.

[1] Our code, models and other related resources are publicly available at https://github.com/Zeng-WH/S2KG.

[2] http://seretod.org/

[3] https://huggingface.co/hfl/chinese-roberta-wwm-ext-large

Proceedings of the Towards Semi-Supervised and Reinforced Task-Oriented Dialog Systems (SereTOD), pages 39 - 47
December 7, 2022 ©2022 Association for Computational Linguistics



Dialogue

User	Agent
U1:"嗯你好你给我办理一个十八块钱的套餐可以吗"	A1: "飞享十八的是吧"
"用户意图": "**问候^{1}**"	"客服意图": "**主动确认^{1}**"
"**十八块钱的套餐**"："**套餐**"，("业务费用":"十八块钱")	"**飞享十八**"："**套餐**"，("业务费用":"十八")
U2:"噢"	A2:"办十八的这个嗯"
"用户意图": "**被动确认^{2}**"	"客服意图": "**其他^{2}**"
	"**十八的这个**"："**套餐**"，("业务费用":"十八")
U3:"你看够吗_噢那就给我办这个_办这个套餐吧"	A3:"可以"
"用户意图": "**询问^{3}**"	"客服意图": "**其他^{3}**"
"**这个套餐**"："**套餐**"	
U4:"那你把那个套餐现在能还能换回来不"	A4:"下个月再打过来办理因为流量包开通是立即生效的"
"用户意图": "**询问^{4}**"	"客服意图": "**通知^{4}**"
"**那个套餐**"："**套餐**"	"**流量包**"："**流量包**"
U5:"那就还是以前那个套餐然后给我每个月包十块钱流量"	A5:"不客气还有其他问"
"用户意图": "**主动确认^{5}**"	"客服意图": "**客套^{5}**"
"**套餐**"："**套餐**"，"**流量包**"：("业务费用":"十块钱")	

KB

```
"ent-1": {
  "name": "这个套餐,十八的这个,飞享十八,十八块钱的套餐",
  "type": "套餐",
  "业务费用": "十八块钱,十八"},
"ent-2": {
  "name": "套餐,那个套餐",
  "type": "套餐"},
"ent-3": {
  "name": "流量包",
  "type": "流量包",
  "业务费用": "十块钱"}
```

TRACK1

Input：Utterance of the current User or Agent

Output：Annotations of entities and attributes

```
"ent-1":{
    "name: "十八块钱的套餐,飞享十八,十八的这个,这个套餐"
    "type": "套餐",
    "业务费用": "十八块钱,十八"}
"ent-2":{
    "name": "那个套餐,套餐",
    "type": "套餐"},
"ent-3": {
    "name": "流量包",
    "type": "流量包",
    "业务费用": "十块钱"}
```

TRACK2

Input: Conversation context for user and agent

Output: User intention, system intention, generated response

"用户意图":"**问候^{1},被动确认^{2},询问^{3,4},主动确认^{5}**

"客服意图":**主动确认^{1},其他^{2,3},通知^{4},客套^{5}**

生成客服回复：A1，A2，A3，A4，A5

Figure 1: An example from MobileCS.

pre-trained language models (e.g. T5 [4] and UFA (He et al., 2022)) as our backbone. Then, we take the dialog history and serialized local KB [5] as input and output system response. Here, we simply concatenate each key-value pair in the local KB as *key: value* to build a string input. We only use response generation as the learning objective. For the labeled dataset, we use the golden KB annotations as our input. For the unlabeled dataset, we obtain the predicted KB results using our model in Track 1. Finally, we mix up all the data to train a knowledge-grounded dialog model.

We summarize the main contributions of our system S2KG (**S**emi **S**upervised **K**nowledge-**G**rounded pre-training) as follows:

- We build a knowledge-grounded dialog model to formulate dialog history and local KB as input and predict the system response.
- We perform semi-supervised pre-training both on the labeled and unlabeled data.

Our system achieves the first place both in the automatic evaluation and human interaction, especially with higher BLEU (+7.64) and Success (+13.6%) than the second place.

Metric	labeled	unlabeled
Dialogs	8,975	87,933
Turns	100,139	972,573
Tokens	3,991,197	39,491,883
Avg.turns per dialog	11.16	11.06
Avg.tokens per turn	39.86	40.61
Slots	26	-
Values	14,623	-

Table 1: Training dataset statistics of MobileCS. The challenge also provides another 1,000 labeled dialogs as evaluation data (dev set).

2 Task Description

MobileCS is a large Chinese TOD dataset collected from real-world dialog transcripts between real users and customer-service staffs. Different from the simulated MultiWOZ dataset (Budzianowski et al., 2018), it consists of real-life data and large unlabeled dialogs. Specifically, MobileCS contains 10,000 labeled dialogs and 90,000 unlabeled dialogs. The full data statistics are shown in Table 1. The challenge has two tracks. Track 1 (information extraction) aims to extract entities and attributes to

---

[4] https://github.com/ZhuiyiTechnology/t5-pegasus

[5] A local KB for a dialog could be viewed as being composed of the relevant snapshots from the global KB. Please see more details in Ou et al. (2022).

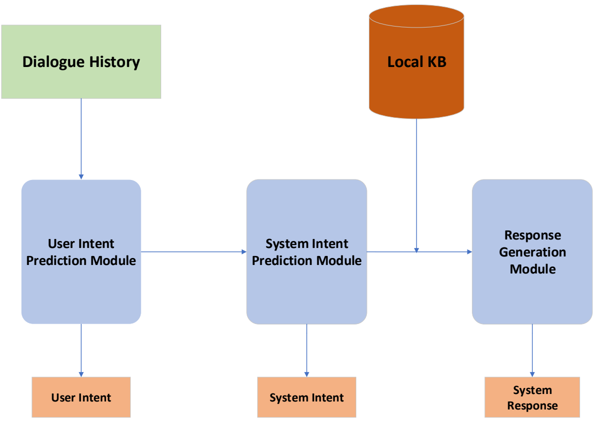

Dialogue History
Local KB
User Intent Prediction Module
System Intent Prediction Module
Response Generation Module
User Intent
System Intent
System Response

Figure 2: Overall architecture of our knowledge-grounded task-oriented dialogue system.

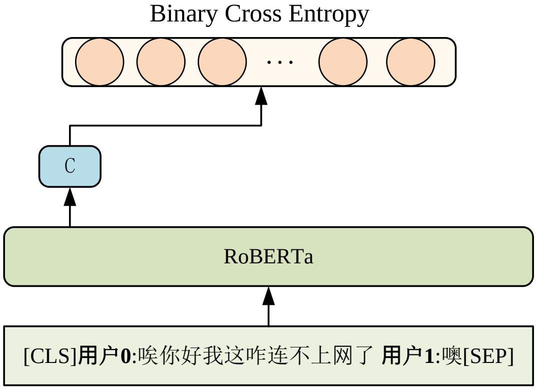


Figure 3: The architecture of the classification models.

build a local knowledge base (KB). And Track 2 uses the KB and raw dialogs to train a complete TOD system. We provide a real annotated dialog in Figure 1. In this paper, we focus on Track 2. Here, we elaborate on the task details. Track 2 for the TOD system is, for each dialog turn, given the dialog history, the user utterance and the local KB, to predict the user intent, query the local KB and generate appropriate system intent and response according to the queried information. For every labeled dialog, the annotations consist of user intents, system intents and a local KB. The local KB is obtained by collecting the entities and triples annotated for Track 1. For unlabeled dialogs, there are no such annotations.

To measure the performance of TOD systems, both automatic evaluation and human evaluation will be conducted. For automatic evaluation, metrics include Precision/Recall/F1 score, Success rate and BLEU (Papineni et al., 2002) score. P/R/F1 are calculated for both predicted user intents and system intents. Success rate is the percentage of generated dialogs that achieve user goals. BLEU score evaluates the fluency of generated responses[6]. For human evaluation for different TOD systems, real users will interact with those systems according to randomly given goals. For each dialog, the user will score the system on a 5-point scale (1-5) by the following 3 metrics. 5 denotes the best and 1 denotes the worst, respectively.

- **Success**. This metric measures if the system successfully completes the user goal by interacting with the user;
- **Coherency**. This metric measures whether the system's response is logically coherent with the dialogue context;
- **Fluency**. The metric measures the fluency of the system's response.

The average score from automatic evaluation and human evaluation is the main ranking basis on the leaderboard.

3 Methodology

3.1 Overall Architecture

Figure 2 shows the overall system architecture for Track 2. Track 2 contains three tasks: user intent, system intent, and system response. For the user intent and system intent tasks, we use Roberta-large and build two separate classification models. For the system response task, we build a knowledge-grounded dialog model and perform semi-supervised pre-training both on the labeled and unlabeled data.

3.2 Subtask 1: Classification

Given a dialog history, the user intent and system intent tasks aim to predict the user intent or system intent(act) respectively. Considering both the tasks are multi-label, we formulate the tasks as multi-label text classification questions. As Figure 3 displays, we adopt Roberta as our backbone and use the dialog history as input. For the user intent task, we concatenate two user utterances as input. We find too many turns bring no further improvements and introducing system responses has a side effect. We suppose the gap between training and prediction affects the model performance. For the system intent task, we concatenate three

---

[6] The challenge adopts BLEU-4.

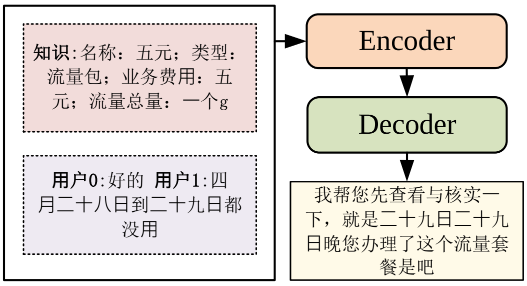

知识:名称：五元；类型：流量包；业务费用：五元；流量总量：一个g
用户0:好的 用户1:四月二十八日到二十九日都没用
Encoder
Decoder
我帮您先查看与核实一下，就是二十九日二十九日晚您办理了这个流量套餐是吧

Figure 4: The architecture of the generation model.

user utterances as input.[7] Then, we use the hidden state of the [CLS] token to predict the results. Binary cross entropy is the learning objective. Section 4.3 proves classification models outperform GPT-based end-to-end models. Besides, we also introduce some augmentation strategies as follows:

- **Continual Pre-training**. We pre-train Roberta on the labeled and unlabeled dialogs using MLM objective like BERT (Devlin et al., 2019). We pre-train 20 epochs using a learning rate of 5e-4 and 15% mask rate. MLM continual pre-training brings large improvements of 1.68% on User Intent F1 and 1.86% on System Intent F1.

- **Class-wise Threshold**. We adaptively select the best threshold for each intent type based on the performance on the dev set. This strategy brings improvements of 1.23% on User Intent F1 and 1.48% on System Intent F1.

- **Adversarial Training**. We adopt FGM (Goodfellow et al., 2015) as our adversarial training strategy. This strategy brings improvements of 0.64% on User Intent F1 and 0.46% on System Intent F1.

3.3 Subtask 2: Generation

For the generation task, we build a knowledge-grounded dialog model, S2KG in Figure 4. Specifically, we firstly use pre-trained language models (e.g. T5 [8] and UFA (He et al., 2022)) as our backbone. Then, we take the dialog history and serialized local KB as input and output system response. Here, we simply concatenate each key-value pair in the local KB as *key: value* to build a string input. We only use response generation as the learning objective. We find KB grounding has a large improvement over baselines (see Section 4.4).

The SereTOD challenge gives a large-scale (90,000) unlabeled dataset that doesn't contain KB annotations and a relatively small (10,000) labeled dataset. So we perform semi-supervised pre-training to utilize all these dialogs. For the labeled dataset, we use the golden KB annotations as our input. For the unlabeled dataset, we obtain the predicted KB results using our model in Track 1. We implement our system of Track 1 mainly based on the official baseline (Liu et al., 2022b). Finally, we mix up all the data to train a knowledge-grounded dialog model. We find only using unsupervised pre-training gets an improvement of 1.91 BLEU, but drops by 14.6 on Success, because raw response generation pre-training makes the model memorize similar dialogs but predict unfaithful responses without grounding ability. Therefore, it's necessary to obtain pseudo KB annotations to perform pre-training. Note that the performance of the Track 1 system is relatively poor so we argue the quality of pseudo KB makes no significant effect on the final results. We leave more discussion to future work.

4 Experiment

4.1 Setup

We train our models on the training set and report our results on the dev set. The final leaderboard results are evaluated on the test set. Since the test set is not released until the end of the challenge, we perform ablation studies only on the dev set. We conduct our experiments using Huggingface[9] and computation platform JiuTian[10].

4.2 Main Results

Table 2 shows the final automatic results on the test set of the top 5 teams[11]. Our system (Team 11) achieves the state-of-the-art on all metrics, especially for generation task, demonstrating the effectiveness of our proposed S2KG. Specifically, our method outperforms the second place (Team 5) by 1.4% on User Intent F1 and 0.6% on System Intent F1. The improvements mainly come from better pre-trained LM, continual pre-training, class-wise threshold, and adversarial training. We will dive

---

[7] The system intent task also requires intent arguments. We use heuristic rules based on the local KB to match the entities.

[8] https://github.com/ZhuiyiTechnology/t5-pegasus

[9] https://huggingface.co/

[10] https://jiutian.10086.cn/edu/#/home

[11] See all the results in the official leaderboard.

Team ID	Automatic Evaluation				
	User Intent F1	System Intent F1	BLEU	Success	Combined
Team-11 (Ours)	**0.728**	**0.595**	**14.430**	**0.780**	**2.392**
Team-5	0.714	0.589	6.790	0.432	1.871
Team-13	0.706	0.587	5.526	0.251	1.655
Team-10	0.664	0.504	3.629	0.217	1.458
Team-8	0.699	0.550	6.440	0.644	2.022
official baseline	0.644	0.394	4.170	0.315	1.436

Table 2: Final automatic results on the test set of the top 5 teams released by the officials. User Intent F1 denotes the performance of classifying the input user query and System Intent F1 denotes the predicted system acts. Success rate is the percentage of generated dialogs that achieve user goals. Combined score is the overall result which is calculated as follows: Combined score = User intent F1 + System intent F1 + Success + BLEU/50.

Methods	User Intent F1	System Intent F1
GPT-2 (baseline)	0.6488	0.4012
Roberta	0.7448	0.5158
Roberta+FGM	0.7512	0.5204
Roberta+FGM+Threshold	0.7635	0.5352
Roberta+FGM+Threshold+MLM	**0.7803**	**0.5538**

Table 3: Comparison of different user intent and system intent models on the dev set.

into details in Section 4.3. For generation metrics, our S2KG model significantly outperforms the second place with a large margin of 7.640 on BLEU and 34.8% on Success. The improvements are mainly attributed to knowledge-grounded dialog model and semi-supervised pre-training, which are the key points of this paper. We leave the discussion to Section 4.4.

4.3 Classification

To verify the effect of our proposed models, we perform an ablation study of different user intent and system intent models on the dev set in Table 3. GPT-2 is the official baseline (Liu et al., 2022a) which is an end-to-end generative model based on Chinese GPT-2[12]. For pre-trained language models, we find Roberta-based classification models get better performance with improvements of 9.60% on User Intent F1 and 11.46% on System Intent F1. Based on Roberta, we also introduce some training or inference strategies, including adversarial training FGM, class-wise threshold, and MLM continual pre-training. All the strategies show advantages. MLM continual pre-training brings the largest improvements of 1.68% on User Intent F1 and 1.86% on System Intent F1, demonstrating the effectiveness of pre-training on domain corpus. Other strategies also get 0.5-1% improvements.

4.4 Generation

Table 5 displays the ablation study of our S2KG system for the response generation task. We analyze the results from the following perspectives.

Knowledge Grounding GPT2-FT (finetune) denotes the official baseline. GPT2-KGFT is our proposed knowledge grounding finetuning method which uses the serialized KB as knowledge. The first two lines in Table 5 show GPT2-KGFT significantly outperforms GPT2-FT by 3.09 BLEU and 34.8% Success, demonstrating the effectiveness of knowledge grounding based on local KB. We also find knowledge grounding improves the factual consistency of generated responses. We give examples in Section 5.1.

Semi-Supervised Pre-training The SereTOD challenge gives a large-scale unlabeled dataset that doesn't contain KB annotations. So we perform different pre-training settings to utilize these unlabeled dialogs. T5-KGFT is our proposed knowledge grounding model which replaces GPT2 with T5. Based on T5-KGFT, T5-Unsup-KGFT first performs an unsupervised response generation pre-training without KB input and then adopts knowledge grounding finetuning. Results show unsupervised pre-training gets an improvement of 1.91 BLEU, but drops by 14.6 on Success. We argue it's because raw response generation pre-training makes the model memorize similar dialogs but predict unfaithful responses without grounding ability. T5-Semi replaces unsupervised pre-training with semi-supervised pre-training which uses Track 1 system to generate the pseudo local KB for these unlabeled dialogs. T5-Semi outperforms T5-Unsup-KGFT by 1.16 BLEU and 17.2% Success, demonstrating the effectiveness of semi-supervised pre-training. We also find continual knowledge grounding finetuning on labeled data (T5-Semi-

---

[12] https://huggingface.co/uer/gpt2-chinese-cluecorpussmall

Team ID	Human Evaluation				Final Score
	Fluency	Coherency	Success	Average	
Team-11 (Ours)	**4.23**	**3.73**	**3.47**	**3.81**	**3.10**
Team-5	4.06	3.14	3.40	3.53	2.70
Team-13	3.55	3.03	2.77	3.12	2.39
Team-10	3.20	2.98	3.11	3.10	2.28
Team-8	2.39	2.29	2.03	2.24	2.13

Table 4: Final human evaluation results on the test set of the top 5 teams released by the officials. Final score is the average of Combined score from automatic evaluation and averaged human evaluation score. It's the main ranking basis on the Track 2 leaderboard.

Methods	BLEU	Success
GPT2-FT (baseline)	4.39	0.344
GPT2-KGFT	7.48	0.692
T5-KGFT	11.32	0.741
T5-Unsup-KGFT	13.23	0.595
T5-Semi	14.39	0.767
T5-Semi-KGFT	12.30	0.761
UFA-Semi	14.51	0.789

Table 5: Comparison of different system response generation models on the dev set.

KGFT) can't bring further improvements upon T5-Semi because of knowledge forgetting.

Pre-trained Language Model We also compare different PLMs. We find that T5 consistently achieves better results than GPT-2. Besides, we experiment with a large PLM specified for customer service, UFA-large (He et al., 2022), which has 1.2B parameters compared to 220M T5 and 117M GPT-2. UFA-large further outperforms T5 by 0.12 BLEU and 2.2% Success.[13]

4.5 Human Evaluation

SereTOD performs human evaluation for different TOD systems, where real users interact with those systems according to randomly given goals. Table 4 shows the results of human evaluation and final scores. Our system also achieves state-of-the-art on all the metrics. Specifically, our method outperforms the second place (Team 5) by 0.17 on Fluency, 0.59 on Coherency, and 0.07 on Success.

5 Analysis

5.1 Case Study

Figure 5 shows three examples from the baseline model and the S2KG model, respectively, proving the advantages of S2KG model from the three dimensions of Success, Fluency, and Coherency.

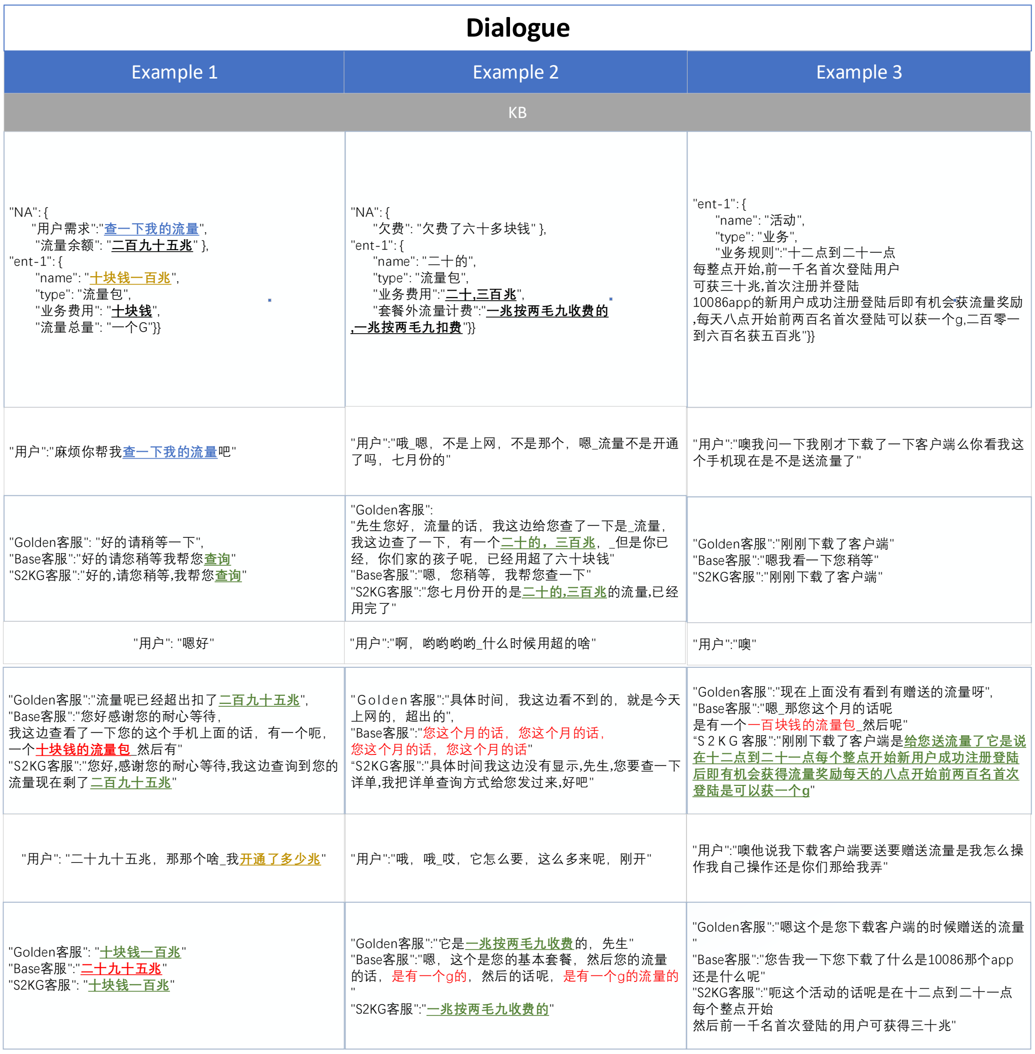
Success In example one, the local KB includes the user's mobile package balance and information about the data package plan currently held by the user. The user's utterance is "Could you please check my data package for me?", which means the user asks the system to query the mobile package balance. The baseline system misidentified the user's intent and mistakenly believed that the user was querying the information of the data package plan, so it retrieved the wrong knowledge "ten yuan data package plan", thereby generating a reply wrongly. The S2KG model correctly identified the user's request, retrieved the correct result based on local KB, and successfully answered that the current mobile package balance was 295M in the reply. It proves that knowledge-grounded semi-supervised pre-training can greatly improve the accuracy of knowledge selection.

Fluency In example two, the user's second round of utterance is intended to query the date when the data package cap is exceeded. Since there is no corresponding information in the current local KB, the system cannot retrieve the knowledge. In this scenario, the baseline system repeated meaninglessly and failed to generate fluent responses. Due to the large-scale pre-training, the S2KG model can explain the situation to the user, guide the user correctly, and provide the user with a reasonable solution.

Consistency In example three, the user's dialogue history is mainly related to the data package given by the activity, while the baseline model mainly answers the user's current data package in the second round of replies, which cannot be consistent with the dialogue history. The S2KG model has good modeling of the dialogue history through pre-training, so it can explain the activity rules to the

---

[13] Considering the inference efficiency and hardware limit, we submit our final results on T5.



Dialogue

Example 1	Example 2	Example 3
KB		

"NA": {
"用户需求":"查一下我的流量",
"流量余额": "二百九十五兆" },
"ent-1": {
"name": "十块钱一百兆",
"type": "流量包",
"业务费用": "十块钱",
"流量总量": "一个G"}}
	"NA": {
	"欠费": "欠费了六十多块钱" },
	"ent-1": {
	"name": "二十的",
	"type": "流量包",
	"业务费用":"二十,三百兆",
	"套餐外流量计费":"一兆按两毛九收费的,一兆按两毛九扣费"}}
		"ent-1": {
		"name": "活动",
		"type": "业务",
		"业务规则":"十二点到二十一点每整点开始,前一千名首次登陆用户可获三十兆,首次注册并登陆10086app的新用户成功注册登陆后即有机会获流量奖励,每天八点开始前两百名首次登陆可以获一个g,二百零一到六百名获五百兆"}}

"用户":"麻烦你帮我查一下我的流量吧"	"用户":"哦_嗯，不是上网，不是那个，嗯_流量不是开通了吗，七月份的"	"用户":"噢我问一下我刚才下载了一下客户端么你看我这个手机现在是不是送流量了"
"Golden客服": "好的请稍等一下", "Base客服":"好的请您稍等我帮您查询" "S2KG客服":"好的,请您稍等,我帮您查询"	"Golden客服": "先生您好，流量的话，我这边给您查了一下是_流量，我这边查了一下，有一个二十的，三百兆，_但是你已经，你们家的孩子呢，已经用超了六十块钱" "Base客服":"嗯，您稍等，我帮您查一下" "S2KG客服":"您七月份开的是二十的,三百兆的流量,已经用完了"	"Golden客服":"刚刚下载了客户端" "Base客服":"嗯我看一下您稍等" "S2KG客服":"刚刚下载了客户端"
"用户": "嗯好"	"用户":"啊，哟哟哟哟_什么时候用超的啥"	"用户":"噢"
"Golden客服":"流量呢已经超出扣了二百九十五兆", "Base客服":"您好感谢您的耐心等待，我这边查看了一下您的这个手机上面的话，有一个呃，一个十块钱的流量包_然后有" "S2KG客服":"您好,感谢您的耐心等待,我这边查询到您的流量现在剩了二百九十五兆"	"Golden客服":"具体时间，我这边看不到的，就是今天上网的，超出的", "Base客服":"您这个月的话，您这个月的话，您这个月的话，您这个月的话" “S2KG客服":"具体时间我这边没有显示,先生,您要查一下详单,我把详单查询方式给您发过来,好吧"	"Golden客服":"现在上面没有看到有赠送的流量呀", "Base客服":"嗯_那您这个月的话呢是有一个一百块钱的流量包_然后呢" “S2KG客服":"刚刚下载了客户端是给您送流量了它是说在十二点到二十一点每个整点开始新用户成功注册登陆后即有机会获得流量奖励每天的八点开始前两百名首次登陆是可以获一个g"
"用户": "二十九十五兆，那那个啥_我开通了多少兆"	"用户":"哦，哦_哎，它怎么要，这么多来呢，刚开"	"用户":"噢他说我下载客户端要送要赠送流量是我怎么操作我自己操作还是你们那给我弄"
"Golden客服": "十块钱一百兆" "Base客服":"二十九十五兆" "S2KG客服": "十块钱一百兆"	"Golden客服":"它是一兆按两毛九收费的，先生" "Base客服":"嗯，这个是您的基本套餐，然后您的流量的话，是有一个g的，然后的话呢，是有一个g的流量的" "S2KG客服":"一兆按两毛九收费的"	"Golden客服":"嗯这个是您下载客户端的时候赠送的流量" "Base客服":"您告我一下您下载了什么是10086那个app还是什么呢" "S2KG客服":"呃这个活动的话呢是在十二点到二十一点每个整点开始然后前一千名首次登陆的用户可获得三十兆"

Figure 5: Case study for three examples from Baseline and S2KG system. We present the local KB, the user utterances, golden response, baseline predictions, and prediction of S2KG system.

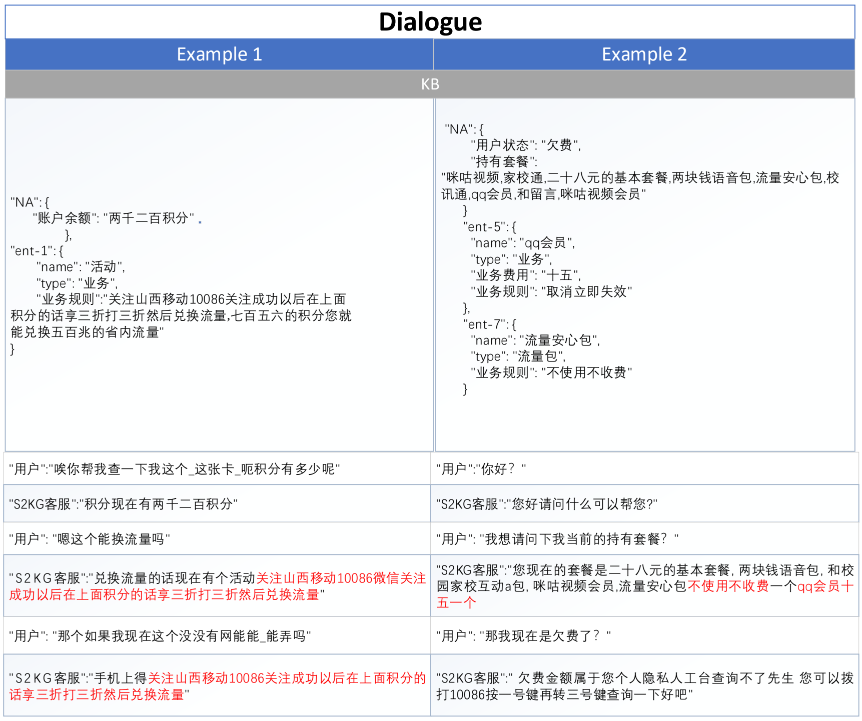
user in detail, to meet the user's intent, and the generated replies are consistent with the dialogue history.

5.2 Challenge

Although the S2KG model has achieved SOTA in the three dimensions of Success, Fluency and Coherency, there are still issues unresolved as showed in Figure 6: (1) **Response Diversity**: As shown in example 1, the user wants to figure out the rules for the user points redemption activity. Although the system provides the retrieved plan rules, the user still cannot understand these rules. So the user asks a question again, then the system repeats the business rules in the KB, resulting in a decrease in the diversity of the response and the user's engagement. (2) **Knowledge Redundancy**: As shown in example 2, the user asks the system about the data packages the user hold, the S2KG model not only provides the name of the packages in the reply, but also retrieves other information associated with the packages, such as fees, etc. As a result, there is knowledge redundancy in the system's reply, which is not conducive to the user's ability to grasp the key points.

6 Conclusion

In this paper, we present our models for Track 2 of the SereTOD 2022 challenge aiming to build semi-supervised and reinforced TOD systems. We divide Track 2 into two task groups, classification (user intent and system intent) and generation (system response). For the classification tasks, we employ Roberta-large and build two separate classification models. We also perform continual pre-training, class-wise threshold, and adversarial training strate-



Dialogue

Example 1	Example 2
KB	

"NA": {
"账户余额": "两千二百积分" ,
},
"ent-1": {
"name": "活动",
"type": "业务",
"业务规则":"关注山西移动10086关注成功以后在上面积分的话享三折打三折然后兑换流量,七百五六的积分您就能兑换五百兆的省内流量"
}
	"NA": {
	"用户状态": "欠费",
	"持有套餐": "咪咕视频,家校通,二十八元的基本套餐,两块钱语音包,流量安心包,校讯通,qq会员,和留言,咪咕视频会员"
	}
	"ent-5": {
	"name": "qq会员",
	"type": "业务",
	"业务费用": "十五",
	"业务规则": "取消立即失效"
	},
	"ent-7": {
	"name": "流量安心包",
	"type": "流量包",
	"业务规则": "不使用不收费"
	}

"用户":"唉你帮我查一下我这个_这张卡_呃积分有多少呢"	"用户":"你好？"
"S2KG客服":"积分现在有两千二百积分"	"S2KG客服":"您好请问什么可以帮您?"
"用户": "嗯这个能换流量吗"	"用户": "我想请问下我当前的持有套餐？"
"S2KG客服":"兑换流量的话现在有个活动关注山西移动10086微信关注成功以后在上面积分的话享三折打三折然后兑换流量"	"S2KG客服":"您现在的套餐是二十八元的基本套餐, 两块钱语音包, 和校园家校互动a包, 咪咕视频会员,流量安心包不使用不收费一个qq会员十五一个
"用户": "那个如果我现在这个没没有网能能_能弄吗"	"用户": "那我现在是欠费了？"
"S2KG客服":"手机上得关注山西移动10086关注成功以后在上面积分的话享三折打三折然后兑换流量"	"S2KG客服":" 欠费金额属于您个人隐私人工台查询不了先生 您可以拨打10086按一号键再转三号键查询一下好吧"

Figure 6: Case study for two examples from S2KG system. We present the local KB, the user utterances and prediction of S2KG system.

gies. For the generation task, we build a knowledge-grounded dialog model S2KG and perform semi-supervised pre-training both on labeled data and unlabeled data. Our system achieves first place both in the automatic evaluation and human interaction. We also discuss the advantages and challenges of our system to provide a guideline for future work.

Acknowledgements

We thank all anonymous reviewers for their helpful comments and suggestions. We are also grateful to the track organizers for their valuable work. This work was partially supported by National Key R&D Program of China No. 2019YFF0303300 and Subject II No. 2019YFF0303302, DOCOMO Beijing Communications Laboratories Co., Ltd, MoE-CMCC "Artifical Intelligence" Project No. MCM20190701.

References

Paweł Budzianowski, Tsung-Hsien Wen, Bo-Hsiang Tseng, Iñigo Casanueva, Stefan Ultes, Osman Ramadan, and Milica Gasic. 2018. Multiwoz - a large-scale multi-domain wizard-of-oz dataset for task-oriented dialogue modelling. In *EMNLP*.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Jianfeng Gao, Michel Galley, and Lihong Li. 2018. Neural approaches to conversational AI. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts*, pages 2–7, Melbourne, Australia. Association for Computational Linguistics.

Shuyang Gao, Abhishek Sethi, Sanchit Agarwal, Tagyoung Chung, and Dilek Z. Hakkani-Tür. 2019. Dialog state tracking: A neural reading comprehension approach. In *SIGdial*.

Chih-Wen Goo, Guang-Lai Gao, Yun-Kai Hsu, Chih-Li Huo, Tsung-Chieh Chen, Keng-Wei Hsu, and Yun-Nung (Vivian) Chen. 2018. Slot-gated modeling for joint slot filling and intent prediction. In *NAACL*.

Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy. 2015. Explaining and harnessing adversarial examples. *CoRR*, abs/1412.6572.

Keqing He, Jingang Wang, Chaobo Sun, and Wei Wu. 2022. Unified knowledge prompt pre-training for customer service dialogues. *ArXiv*, abs/2208.14652.

Keqing He, Yuanmeng Yan, and Weiran Xu. 2020a. Learning to tag oov tokens by integrating contextual representation and background knowledge. In *ACL*.

Keqing He, Jinchao Zhang, Yuanmeng Yan, Weiran Xu, Cheng Niu, and Jie Zhou. 2020b. Contrastive zero-shot learning for cross-domain slot filling with adversarial attack. In *COLING*.

Hong Liu, Yucheng Cai, Zhijian Ou, Yi Huang, and Junlan Feng. 2022a. Revisiting markovian generative architectures for efficient task-oriented dialog systems. *ArXiv*, abs/2204.06452.

Hong Liu, Hao Peng, Zhijian Ou, Juan-Zi Li, Yi Huang, and Junlan Feng. 2022b. Information extraction and human-robot dialogue towards real-life tasks: A baseline study with the mobilecs dataset. *ArXiv*, abs/2209.13464.

Sihong Liu, Jinchao Zhang, Keqing He, Weiran Xu, and Jie Zhou. 2021. Scheduled dialog policy learning: An automatic curriculum learning framework for task-oriented dialog system. In *FINDINGS*.

Zhijian Ou, Junlan Feng, Juan-Zi Li, Yakun Li, Hong Liu, Hao Peng, Yi Huang, and Jiangjiang Zhao. 2022. A challenge on semi-supervised and reinforced task-oriented dialog systems. *ArXiv*, abs/2207.02657.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *ACL*.

Baolin Peng, Chunyuan Li, Jinchao Li, Shahin Shayandeh, Lars Lidén, and Jianfeng Gao. 2021. Soloist: Building task bots at scale with transfer learning and machine teaching. *Transactions of the Association for Computational Linguistics*, 9:807–824.

Baolin Peng, Xiujun Li, Jianfeng Gao, Jingjing Liu, Kam-Fai Wong, and Shang-Yu Su. 2018. Deep dyna-q: Integrating planning for task-completion dialogue policy learning. In *ACL*.

Baolin Peng, Chenguang Zhu, Chunyuan Li, Xiujun Li, Jinchao Li, Michael Zeng, and Jianfeng Gao. 2020. Few-shot natural language generation for task-oriented dialog. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 172–182, Online. Association for Computational Linguistics.

Libo Qin, Wanxiang Che, Yangming Li, Haoyang Wen, and Ting Liu. 2019. A stack-propagation framework with token-level intent detection for spoken language understanding. In *EMNLP*.

Yixuan Su, Lei Shu, Elman Mansimov, Arshit Gupta, Deng Cai, Yi-An Lai, and Yi Zhang. 2022. Multi-task pre-training for plug-and-play task-oriented dialogue system. In *ACL*.

Chien-Sheng Wu, Andrea Madotto, Ehsan Hosseini-Asl, Caiming Xiong, Richard Socher, and Pascale Fung. 2019. Transferable multi-domain state generator for task-oriented dialogue systems. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics.

Hong Xu, Keqing He, Yuanmeng Yan, Sihong Liu, Zijun Liu, and Weiran Xu. 2020. A deep generative distance-based classifier for out-of-domain detection with mahalanobis space. In *COLING*.

Xiaojin Zhu. 2005. Semi-supervised learning literature survey.