Quickly Training a Reinforcement Learning Agent for Task-Based Dialogue Systems with User Simulation and Transfer Learning

Dan Iter [1] **Allan Jiang** [1] **Li Deng** [1]

Abstract

With the proliferation of speech based assistants and chat bots there is a growing demand for natural language interfaces to a wide variety of applications. Reinforcement learning has shown to be a powerful method of modeling a task-based dialogue. However, there are two major challenges to widely deploying reinforcement learning agents in these settings; training an agent can require a large amount of data and a model for one task does not generalize well to other tasks. We present a system for transferring models between tasks using transfer learning and word embeddings, and we show to efficiently retrain and fine-tine these models using user simulation to generate artificial dialogues. We evaluate our system on a number of HTML form filling tasks, for which we have no real dialogue data. We will show both the empirical success rate in our simulated environment and a user study of how well our agent can interact with real users.

1. Introduction

With the growing popularity of speech based assistants such as Siri, Google Home and Amazon Echo and the growing trend of chat bots being integrated into our every day applications like Slack and Facebook M, there is a growing demand for natural language interfaces to a wide variety of applications. Reinforcement learning has shown to be a powerful method of modeling a dialogue[()], in particular for specific tasks. In this setting a user interacts with a natural language interface to an agent that must take the correct actions to complete the task. However, there are two major challenges to widely deploying reinforcement learning agents in these settings; training an agent can require a large amount of data and a model for one task does not generalize well to other tasks. Collecting a large amount of data can be challenging because it is based on real human interactions and requires manual labeling. Since the action space for different tasks vary, it is unclear how to generalize one model for many tasks. This paper presents a system to mitigate both challenges. We build a user simulation to efficiently simulate many conversations for training. Then we use transfer learning to initialize the weights of the reinforcement learning model and use the simulated conversations to fine tune the model. This significantly decreases training time and makes the agent applicable to most slot filling applications.

*Equal contribution [1]Stanford University, Stanford, USA.

One of the most fundamental tasks for task-based dialogue systems is to act as an interface for slot-filling systems. In a slot-filling system, the agent must collect specific data from a user and fill the relevant slots to be able to complete a tasks. For example, when buying movie tickets, the agent must find out details such as movie title, theater, time, number of tickets, etc. We use HTML forms as a underlying tasks that the agent must complete and submit. We hope to show that given an HTML form we can quickly build a natural language interface with a reinforcement learning agent. We accomplish this by initializing the RL model with weights from another pretrained model for a different task. We then retrain and fine-tune the model using generated dialogues using our user simulation which generation user goals and actions based on hallucinated data for the HTML form.

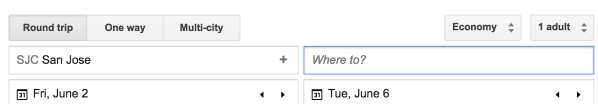

Figure 1. A sample HTML form.

Since we have a fully simulated environment where we have both a programmatic user and a model driven agent, we can measure the success rate, average reward and average dialogue length in the simulated environment to evaluate the quality of our model. However, there will necessarily be some divergence between the behavior of our user simulator and a real user. Therefore, we will also conduct a small user study and error analysis to show how effective the final RL agent is at interacting with real users.

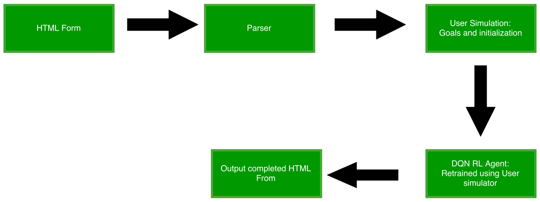

Figure 2. The end to end system.

2. Related Work

Abundant research has been conducted into task-oriented conversation agent. Related work may be categorized into the following three types:

Dialog Generation Deep Reinforcement Learning for Dialog Generation (Li et al., 2016) represents a state by previous two turns. Reward is defined as a weighted sum of ease of answering, variety of information flow and semantic coherence. The system is task-free, and is evaluated against length of dialog, diversity of content, as well as human judgment.

Task-oriented Conversation Xinjun(Li et al., 2017) proposed an end-to-end learning framework backed by a structured database to provide assistance for booking movie tickets. There are three components in their framework: A Natural Language Generator(NLG) module generate texts corresponding to user dialog actions; a Language Understanding(LU) module parses the user input into a semantic frame; and a Dialog Management(DL) module(including a state tracker and a policy learner), accumulates semantics, tracks the dialog states, and produces the next dialog action.

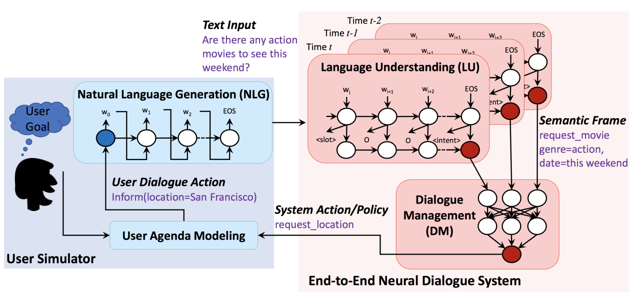

Figure 3. Framework for Xinjun E2E Task Completion Agent

Deep RL for Task-completion Dialog System Combined with task-completion and Deep Reinforcement Learning, Tiancheng at CMU (Zhao & Eskenazi, 2016) built a model for 20Q game, where the agent ask binary questions to the user, get answers *yes/no/unknown* and then guesses a famous person in knowledge base. A correct guess results in a positive reward, while incorrect guess results in a nega-

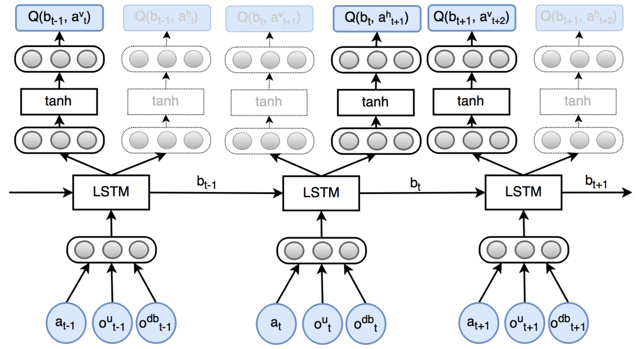

Figure 4. LSTM used to track dialog state

tive reward. LSTM is used to track conversation states and capture essential information in latent dialog states.

So based on the previous work above, we propose a more general task-completion agent trained with deep RL, which can be applied to multiple domains with natural language generation and interpretation.

3. Datasets

4. Approach

- mention user simulation code base and prebuilt NLU and NLG subsystems

5. Next Milestones

References

Li, Jiwei, Monroe, Will, Ritter, Alan, Galley, Michel, Gao, Jianfeng, and Jurafsky, Dan. Deep reinforcement learning for dialogue generation. *arXiv preprint arXiv:1606.01541*, 2016.

Li, Xuijun, Chen, Yun-Nung, Li, Lihong, and Gao, Jianfeng. End-to-end task-completion neural dialogue systems. *arXiv preprint arXiv:1703.01008*, 2017.

Zhao, Tiancheng and Eskenazi, Maxine. Towards end-to-end learning for dialog state tracking and management using deep reinforcement learning. *arXiv preprint arXiv:1606.02560*, 2016.