Multi-Domain Task-Oriented Dialogue Policy Optimization through Sigmoidal Discrete-SAC (DPO-SSAC)

This repository contains an implementation of DPO-SSAC, as described in the corresponding paper, integrated with Convlab-3.

DPO-SSAC is an off-policy deep reinforcement learning method designed for task-oriented dialogue policy learning. It aims to enhance the stability, exploration, robustness, and learning speed of dialogue policies in discrete task-oriented dialogue environments.

One of the key features of DPO-SSAC is the introduction of the sigmoidal discrete-SAC (SSAC) approach, which addresses underestimation bias that may lead to pessimistic exploration in discrete-SAC, particularly in the context of dialogue systems.

Features

- Implementation of DPO-SSAC for task-oriented dialogue policy learning.
- Integration with Convlab-3 for experimentation and evaluation.
- Utilization of the SSAC approach to alleviate underestimation bias and enhance exploration.

Requirements

- Convlab-3

Usage

1. Clone this repository.
2. Install the required dependencies of Covlab-3
3. Follow the instructions provided in the documentation or examples to train and evaluate the DPO-SSAC model.

Citation

If you find this work useful in your research, please consider citing the corresponding paper: [Insert citation details here]