Instructions for ACL 2023 Proceedings

Anonymous ACL submission

1 Introduction

As technologies continue to advance, the integration of robots into every day life is continuing to increase. As humans rely mostly on speech for communication and instruction, it is essential that robots are also developed to understand and decipher language, executing commands via language effectively and efficiently.

In this project, we explore the application of reinforcement learning (RL) to train an agent capable of navigating a 2D grid environment and identifying geometric shapes with distinct colors. The agent moves one space at a time across the grid, using RL techniques to learn an optimal policy for efficiently locating and identifying shapes such as green triangles and red squares. We define the task as a partially observable Markov decision process (POMDP), where the agent's observations are limited to the grid space it occupies. Our approach involves implementing Reinforcement Learning to teach the agent to maximize rewards by minimizing the time and steps required to find and correctly identify a shape. The results demonstrate how reinforcement learning can be applied to shape recognition and navigation tasks, with potential applications in robotics, search-and-rescue missions, and autonomous systems. If time permits, we want to further explore what modifications need to be made in order to command a fleet of agents to accomplish tasks in an optimal manner.

2 Related Work

Most of the work in this project will be based on the work done by [1]. They utilize a combination of large language models and reinforcement learning to help an agent understand specific instructions to navigate in the game environment of doom. Others utlized the world of minecraft to help agents complete tasks, with less of a focus on instruction and more on generalized learning rather than interpreting language [4], [5]. Combining the implementation of [1] with the Minigrid environment [3], [2] we can further explore the challenging problem of robotic instruction with large language models.

3 Experimental Design

Using a similiar setup as noted in Figure 1, we will use the Minigrid environment in place of DOOM to train our agent to complete specified tasks.

References

[1] Devendra Singh Chaplot, Kanthashree Mysore Sathyendra, Rama Kumar Pasumarthi, Dheeraj Rajagopal, and Ruslan Salakhutdinov. Gated-attention architectures for task-oriented language grounding. 6 2017.

[2] Maxime Chevalier-Boisvert, Dzmitry Bahdanau, Salem Lahlou, Lucas Willems, Chitwan Saharia, Thien Huu Nguyen, and Yoshua Bengio. Babyai: A platform to study the sample efficiency of grounded language learning. *arXiv preprint arXiv:1810.08272*, 2018.

[3] Maxime Chevalier-Boisvert, Bolun Dai, Mark Towers, Rodrigo Perez-Vicente, Lucas Willems, Salem Lahlou, Suman Pal, Pablo Samuel Castro, and Jordan Terry. Minigrid & miniworld: Modular & customizable reinforcement learning environments for goal-oriented tasks. In *Advances in Neural Information Processing Systems 36, New Orleans, LA, USA*, December 2023.

[4] Junhyuk Oh, Satinder Singh, Honglak Lee, and Pushmeet Kohli. Zero-shot task generalization with multi-task deep reinforcement learning. 6 2017.

[5] Chen Tessler, Shahar Givony, Tom Zahavy, Daniel J Mankowitz, and Shie Mannor. A deep hierarchical approach to lifelong learning in minecraft.

A Example Appendix

This is a section in the appendix.

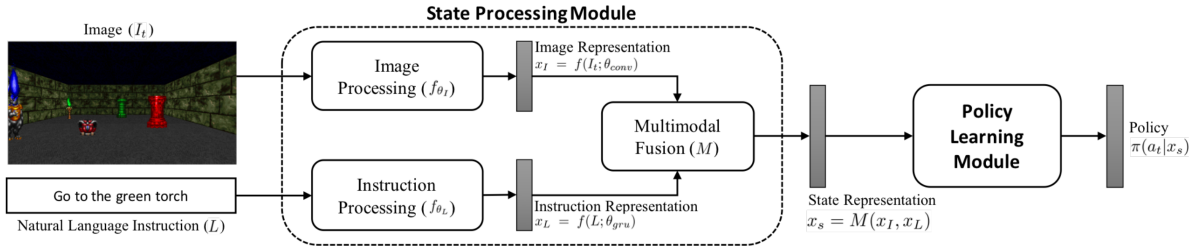

Figure 1: State processing method as developed by [1]