
Cognitive Navigation using Value Iteration

Smruti Amarjyoti
Carnegie Mellon University
Pittsburgh, PA 15213
samarjyo@andrew.cmu.edu

Dinesh Narapureddy
Carnegie Mellon University
Pittsburgh, PA 15213
dnarapur@andrew.cmu.edu

George Tan
Carnegie Mellon University
Pittsburgh, PA 15213
georget1@andrew.cmu.edu

1 Introduction

1.1 Problem statement

Robotic navigation has been a domain previously untouched from the advancements in data-driven learning techniques. The problem of finding efficient paths to target goals is deeply ingrained in the framework of reinforcement learning. Moreover, the idea of integrating exploration, map building and optimal planning into a single architecture is a novel field that we will be exploring with this project. In this work we are attempting to combine the ideas of navigation and deep reinforcement learning in a coherent manner. Further, we will be addressing the prevalent problems of multi-target generalization and scene generalization. Whereas, earlier DRL models trained for game playing work towards maximizing rewards for a specific goal, our network focuses towards training the model to handle various targets by sharing parameters across the tasks [1].

1.2 Challenges

The deep siamese actor-critic networks used by the authors in [1] handles the problems of multiple targets and scene generalization. Our task at hand would be to analyze the performance of their implementation and coming up with architectural enhancements that bolster the navigation performance. We would be extrapolating ideas of cognitive mapping and value iteration networks from [2] and [4] to try and improve the performance. Our proposed network will first create a map via interacting with the environment and then find a optimal policy through value iteration networks.

2 Proposal

2.1 Datasets

As part of the the THOR Challenge, we select the AI2-THOR framework which simulates high-quality 3D scenes [1]. It creates a simulation of real-world appearance and physics alongside with physical interactions, such as applying a force, grasping, or object manipulations [1]. The agent uses the 2D observations of the environment taken by its first person RGB camera to reach a given target. Similar to AI2-THOR, UnrealCV [5] also builds simulated 3D scenes where the agent can the virtual world.

2.2 Method

We propose a map-based method where a top-view map of the scene is generated from the observations received by the agent[3]. Given the features of the top-view map, a value iteration network constructs a 2D value map [4] to generate the policy for the agent.

2.3 Evaluation Metrics

Both the networks will be evaluated using the number of action steps taken to reach the specified goal. The baseline comparison is to the siamese actor-critic network in [1] and the shortest path distance.

3 Related work

Navigation has been a widely open problem in robotics for decades. Most of the previous literature addressed navigation as a path planning algorithm. We attempt to learn the shortest path to the goal using deep reinforcement learning. This is motivated from the recent success of learning algorithms for playing games like Atari, Go, Doom [2] etc. We attempt to solve the navigation problem in uncontrolled environments as proposed by [1]. [1] propose a siamese network for learning a embedded mapping training a reinforcement network do predict actions to reach the goal. Recent literature show promising results for visual navigation using cognition mapping and planning by [3]. They use value iteration network proposed by [4] to learn the policy for navigation.

Milestone goals

For the second milestone, we intend to implement baseline model of [1] Fornamely: siamese layers and scene specific layers. The former is used for the current state and goal state's representation in a common embedded space. The later is used to represent distinct characteristics of every scene that will aid the agent in navigating. End result would be the model that generates policy and value outputs(actor-critic models). The weights in the network will be trained based on a belief 2D map based on [3]. Using the belief map, Value Iteration Networks[4], will predict the suitable next action based on reasoning of its current state and belief map. This will be our end-goal for this project.

Computation resources

The TAs stated that course resources may be allocated to this project. We estimate that we need four GPUs for ten days overall to train the networks.

Reference

- [1] Zhu, Yuke, et al. "Target-driven visual navigation in indoor scenes using deep reinforcement learning." arXiv:1609.05143 (2016)
- [2] Mirowski, Piotr, et al. "Learning to navigate in complex environments." arXiv preprint arXiv:1611.03673 (2016).
- [3] Gupta, Saurabh, et al. "Cognitive Mapping and Planning for Visual Navigation." arXiv preprint arXiv:1702.03920 (2017).
- [4] Tamar, Aviv, et al. "Value iteration networks." Advances in Neural Information Processing Systems. 2016.
- [5] Qiu, W., & Yuille, A., UnrealCV: Connecting Computer Vision to Unreal Engine. In Computer Vision–ECCV 2016 Workshops (pp. 909-916). Springer International Publishing. (2016)