

---

# Optimistic initialization of parameterized value functions using a neural network

---

Anonymous Authors<sup>1</sup>

## Abstract

Optimistic initialization of value functions is a popular approach to exploration in tabular reinforcement learning. However, it is rarely analyzed in deep reinforcement learning. We explore this problem through a parameterized value function using linear neural networks and compare our results to existing popular learning algorithms.

## 1. Introduction

Tradeoff between exploration and exploitation, is an eternal problem that any RL algorithms face. Since this challenge arises everywhere, the numerous amount of research centered to exploration/exploitation problem are published. Among those approaches, one of the most fundamental and scalable approach is an optimistic initialization of the value functions. By simply setting the initial values greater than the reward maxima, one can enforce the agents to explore every state(-action) pair at least once. Although this simple technique is proven to provide the tremendous advantages in tabular context, its effect on deep nonlinear function approximation is yet to be discovered. This project aims to 1. provide a basic method to optimistically initialize the deep network, 2. simple fix on the implementation so that optimistic values remain effective after few gradient steps and 3. empirical analysis.

## 2. Background

## 3. Research Question

## 4. Experimental Design

## 5. Contributions

---

<sup>1</sup>Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. **AUTHORERR: Missing \icmlcorrespondingauthor.**

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.