



# The Effects of Applying Multiple Constraints in Reinforcement Learning



Serigne Toure, Stephen Scott, Ph.D., Eleanor Quint

Department of Computer Science and Engineering, University of Nebraska–Lincoln

## INTRODUCTION

- This research studies the relationship between a reinforcement learning (RL) agent and the constraints applied to it
- Constraints are very important because they add the safety factor on the agent. They can also potentially reduce learning time, save money and improve learning of the agents

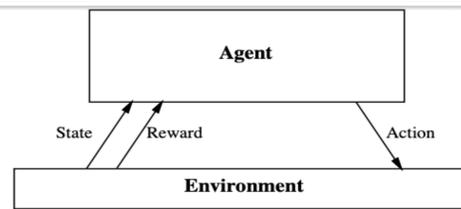
## PURPOSE

- Make RL agents learn better
- Save Graphics Processing Unit hours
- Develop the advancement of artificial intelligence (AI)
- Improve safety in AI controlled devices

**Motivation:** The biggest motivation for constraints and multiple constraints would be in cars equipped with AI. We want the car to know that an action is bad before it even starts learning. For example, we want to constrain a car from running over pedestrians not let it learn by experience.

## BACKGROUND

- Environment
  - Where the agent is placed to learn
- Agent
  - The learner (i.e.: Player, driving car etc.)
- Reward
  - The bonus it gets when it performs well (i.e.: Score in game)
- Constraints
  - Actions an agent cannot perform
- Augmentation
  - Process that helps the agent take constraints into account



$$s_0 \xrightarrow{a_0} r_0 \rightarrow s_1 \xrightarrow{a_1} r_1 \rightarrow s_2 \xrightarrow{a_2} r_2 \rightarrow \dots$$

Goal: Learn to choose actions that maximize

$$r_0 + \gamma r_1 + \gamma^2 r_2 + \dots, \text{ where } 0 \leq \gamma < 1$$

Figure 1: Interaction between agent and environment

## MATERIALS AND METHODS

- Resources
  - For this study we use the baseline library to run all the tests and algorithms necessary
- Environment
  - Two different environments are used, both Atari games: Breakout, and SeaQuest
- Hyperparameters
  - We use a hyper parameter called R-value to randomize the learning process of the agent
  - We use four different R-values: 0.0100, 0.0025, 0.0050, 0.0010
- Tests
  - For each environment, we test every R-value
  - For each of the R-values, 10 random seed agents are tested
  - Each agent is evaluated against the environment 10 separate times

## RESULTS

- Adding constraints adds on the safety factor of an RL agent
- Adding constraints improves the learning of the agent by saving both time and resources
- A combination of constraints that generally help the agent significantly improves the reward of an agent
- A combination of a constraint that improve and another that hurts the agent has no difference between it and the initial baseline results

## CONCLUSIONS

- Adding constraints helps in the learning process of the agent and adds on safety

## FUTURE WORK

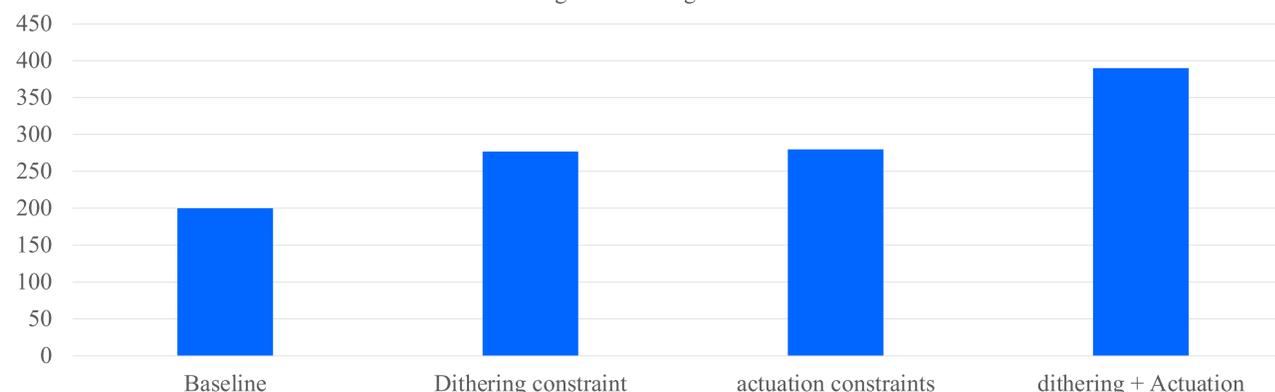
- Designing additional constraints
- Run tests in more environments
- Run test with more constraints

## ACKNOWLEDGMENTS

- Steven Scott, Ph.D., UNL faculty mentor
- Eleanor Quint, Graduate student
- UNL McNair Scholar Program

## RESULTS

Figure 2: Average reward



MCNAIR SCHOLARS PROGRAM