ULiège
Prof. Damien Ernst

INFO8003-1
Optimal decision making
for complex problems

Project

Catcher : Searching High-Quality Policies to Solve a Video Game.

1 DOMAIN

The Catcher video game is illustrated by Figure 1. The agent interacts with the environment by applying an horizontal force on a bar while a fruit is falling at a constant rate from above. The agent continuously receives a positive feedback during the interaction, and a higher one when catching the fruit. However, it receives a negative feedback if the fruit hits the ground. In both cases, a new fruit randomly appears somewhere in the grid. The game ends with no feedback when a given number of fruits have hit the ground.

As first task, provide a formalization of the game based on the *Python class*.

2 POLICY SEARCH TECHNIQUES

Design your own policy search technique based on the reinforcement learning literature. Provide references upon which your approach is based. Your algorithm has to be
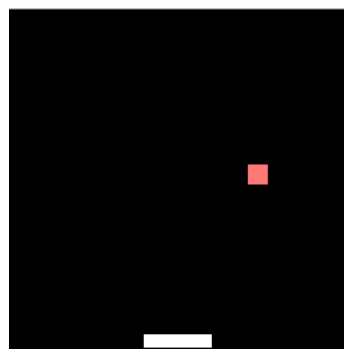

Figure 1: Catcher environment.

able to learn both discrete and continuous policies. Design an experimental protocol which assesses the performance of your algorithm with both policies, compared to a classical algorithm seen during the lectures (e.g., FQI with ensemble of trees). Display the performance of your policies at the end of each episode in terms of expected discounted cumulative reward. You may propose any other metrics you want that are relevant with your approach.

3 IMPLEMENTATION AND DELIVERABLES

Implement your experimental protocol along with your policy search technique. You need to deliver (i) the cleaned and well documented source code and (ii) a report which is outlined accordingly of the project description. Your report also need to explain possible improvements of your approach.