Context-Aware-VR

Mathias Ooms

University of Antwerp

7 september 2020

context-extraction

Context characteristics:

- Eye tracking
- Position tracking

Why reinforcement learning

Supervised learning

- needs data to increase performance
- can never become better than the quality of the data

Why reinforcement learning

Reinforcement learning

- provides a way to have an agent learn completely by itself (without requiring data)
- agents can become better than the 'best' players

Reinforcement learning

Policy Network

- transforms the input to the possible actions

Policy Gradients

- method to train a policy network

Policy Gradients

- starts off from a completely random network

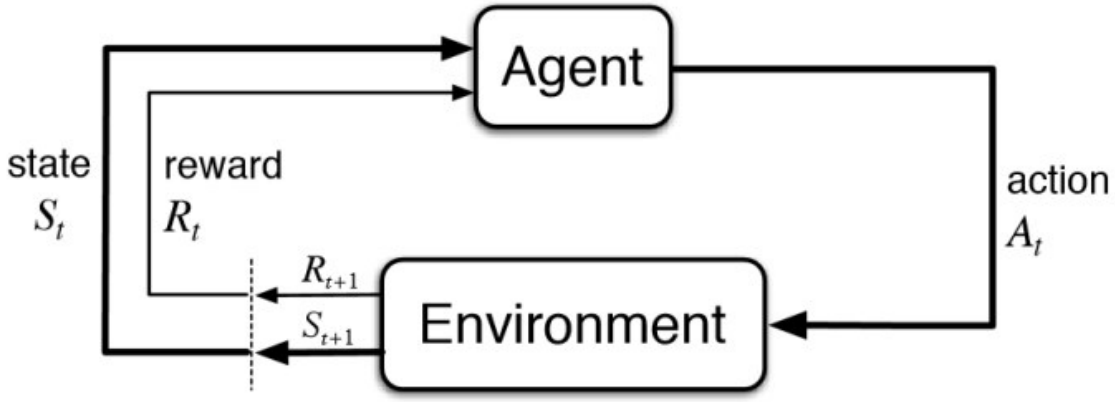

Figure 1: The principle of RL

- calculate for each possible action a probability
- agent will choose an action and get feedback (reward)
- the goal of the agent is to receive as much reward as possible

Policy Gradients

How does it improve the Policy network?

- In the first generations it will miserably fail, until the agent gets lucky (received a good reward)
- this successful series of actions now need to be more preferred over other ones
- when receiving a good reward we multiply the probability with a positive gradient
- when receiving a bad reward we multiply the probability with the same negative gradient

"Troubles in paradise"

Unfortunately there are also downsides:

- *Credit assignment problem*: punished hard for small mistakes
- Not appropriate in complex environments