Deep Q Learning: From Paper to Code

Agents, Environments, and Actions

Fundamental Concepts

A Simple Example

GOOD

BAD

???

Environment Definition

What changes when the robot acts?

Position of the box & rewards!

Environment is  + rewards

Classification (action) causes state transition

State vs. Environment

- State is reading of weight sensor
- Actions cause new box to load → new weight
- Set of all possible states → **state space**
- Human checker is irrelevant
- Time delay doesn’t matter either

Supervised Learning?

Good

Bad

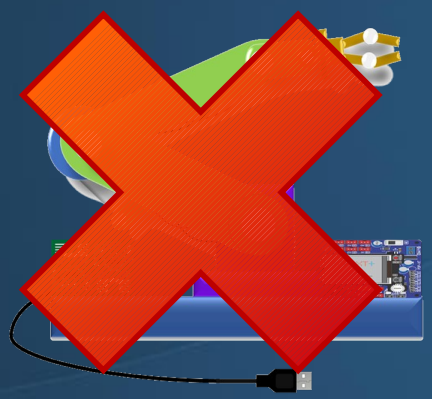
Ugly

???

Need more information for log. reg.

Classification strategy through RL

Agent Definition

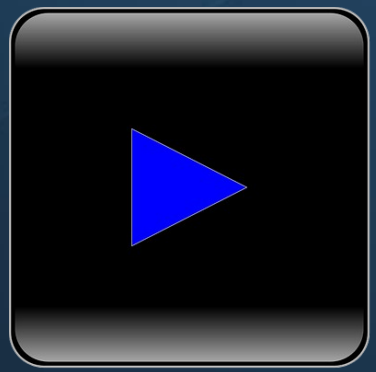


- Agent is software, not hardware!
- Memory of states, actions, and rewards
- Decision making process
- Don’t anthropomorphize
- Software and hardware not colocated
- Rewards constitute reinforcement
- Algorithm (Q learning) to max rewards

Actions

- Discrete (good bin / bad bin / unknown)
- Set of all possible actions → **action space**
- Some problems have continuous actions
- Q learning deals with discrete actions

First Exercise



- Frozen Lake environment
- 0 reward per step; +1 for escaping
- Agent slides!
- Holes (H) terminate episode
- Check out docs at gym.openai.com
- Random agent; 1000 games
- Plot win % over trailing 10 games

Summary

- Interactions of agent and environment
- Agent learns and makes decisions
- Environment is what changes & its representation
- Rewards almost always part of environment
- Set of all possible states → **state space**
- Set of all possible actions → **action space**

Up Next