Cops and Thieves: a Multi-Agent Game

Matteo Nestola
Simona Scala

The game

- The game dynamics unfold through the random generation of teams comprising cops and thieves, the strategic placement of safe zones and coins, and the dynamic determination of the playing area size based on the number of players involved. Thieves embark on a mission to navigate the playing area, including safe zones, to collect coins, while cops diligently patrol, aiming to apprehend the elusive criminals. The game concludes with either the triumph of the thieves, achieved by collecting all coins, or the victory of the cops, attained by apprehending all the cunning thieves.

The tools behind the project

- **Jason** (a robust platform for multi-agent system development)
- **Q-learning** (a reinforcement learning algorithm)
- **Gradle** (build automation tool designed for multi-language software development)
- **Java** (programming language)

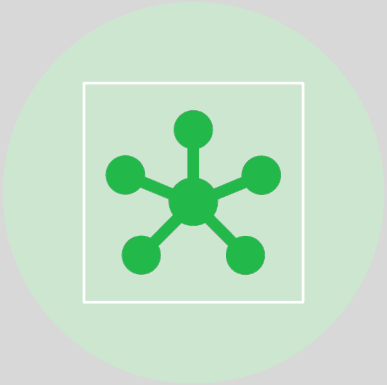
Requirements

- Players
- Safe Zones
- Playing Area
- Coins
- Starting Positions
- Movements Rules
- Ending the Game

Design

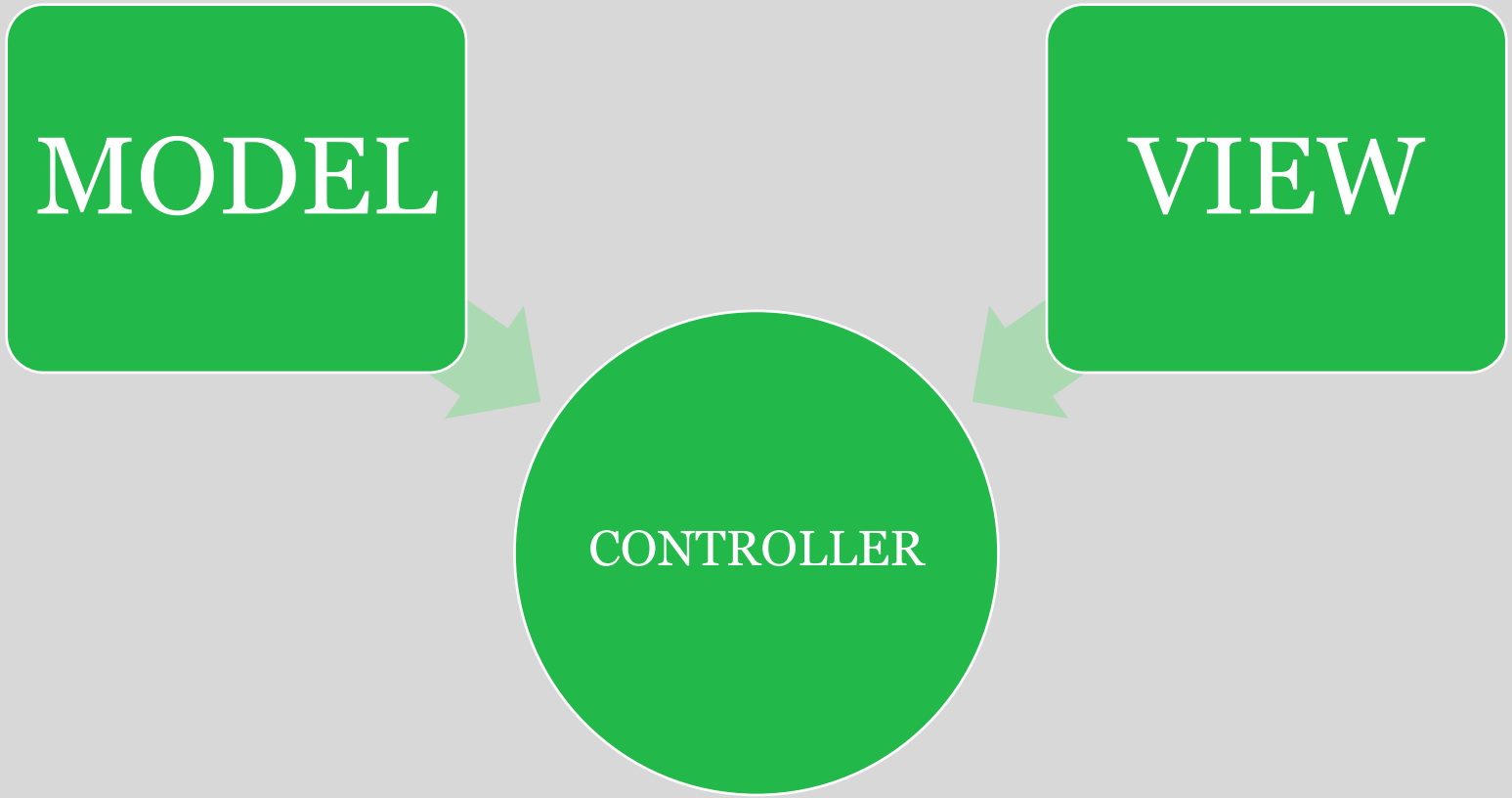
STRUCTURE

BEHAVIOUR

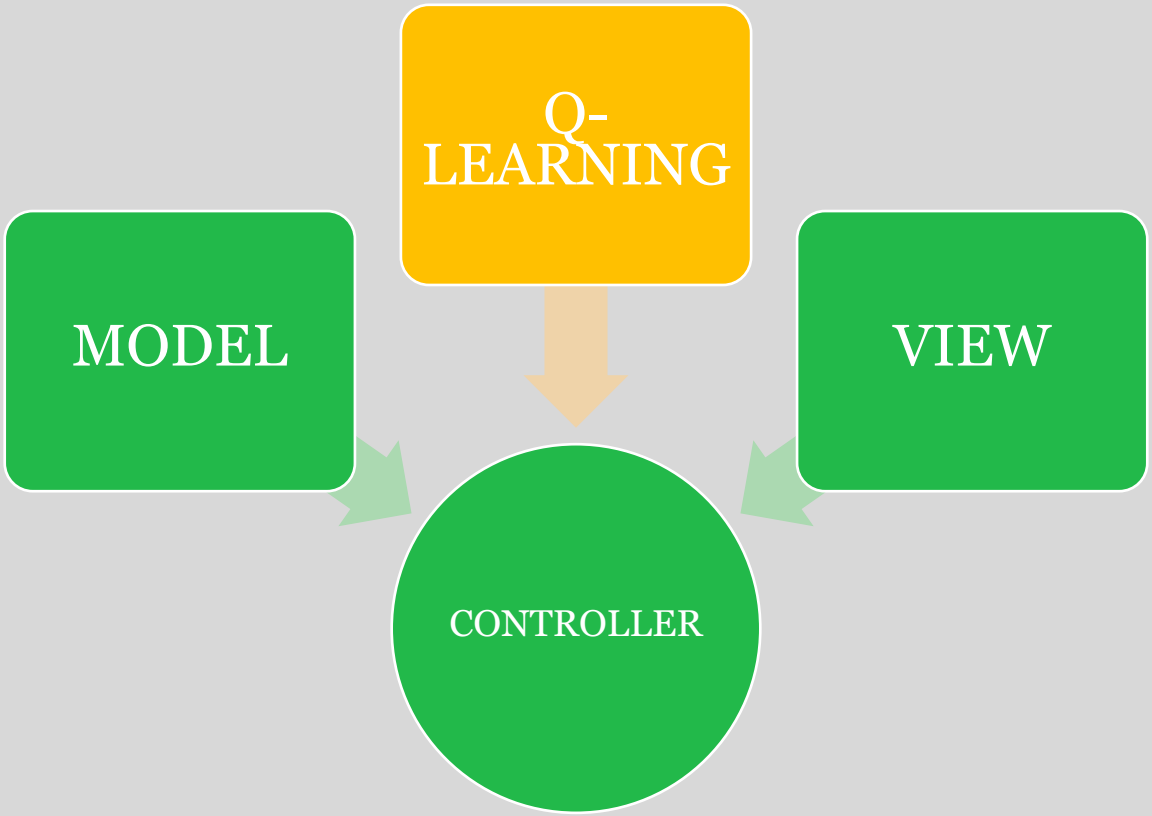


INTERACTION

Structure



Structure



Behaviour

Thieves

- Look for coins
- Q-learning policies
- Change goal (coin)
- Hide in safe zones

Cops

- Look for thieves
- No Q-learning policies
- Broadcast message when thief detected

Common

- Move in four directions

Interactions: Q-learning

$$New\ Q(s,a) = Q(s,a) + \alpha[R(s,a)\ + \gamma\ maxQ'(s',a') - Q(s,a)]$$

Interactions: Q-learning

$$New\ Q(s,a) = Q(s,a) + \alpha[R(s,a)\ + \gamma\ maxQ'(s',a') - Q(s,a)]$$

New Q value for the state and action

Current Q values

Current Q values

Interactions: Q-learning

$$New\ Q(s,a) = Q(s,a) + \alpha[R(s,a) + \gamma\ maxQ'(s',a') - Q(s,a)]$$

Reward for taking an action in a state

Interactions: Q-learning

$$New\ Q(s,a) = Q(s,a) + \alpha[R(s,a) + \gamma\ maxQ'(s',a') - Q(s,a)]$$

Learning rate

Discount rate

Interactions: Q-learning

$$New\ Q(s,a) = Q(s,a) + \alpha[R(s,a) + \gamma\ maxQ'(s',a') - Q(s,a)]$$

Maximum expected future reward

Final results