Moving drones for Wireless Coverage in a three-dimensional grid Analyzed via Game Theory

Elena Camuffo, Luca Gorghetto and Leonardo Badia

Outline

Introduction

- Drone's coordination
- Why Game Theory?
- State of the Art : Two-Dimensional Grid Games

Innovative scenario

- "Three-Dimensional Grid Game"

Game and Algorithm explanation

- Game-theoretical Analysis
- Nash Q - Algorithm

Results

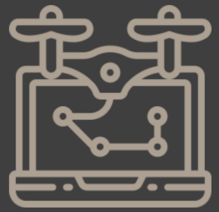
- Convergence and Stationarity
- Average path length

Conclusions

Introduction

- Unmanned Aerial Vehicles (UAVs):
  aircraft without any human pilot onboard.

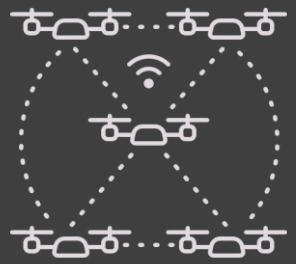
logistics



production

wireless communication

- Future uses of UAVs.

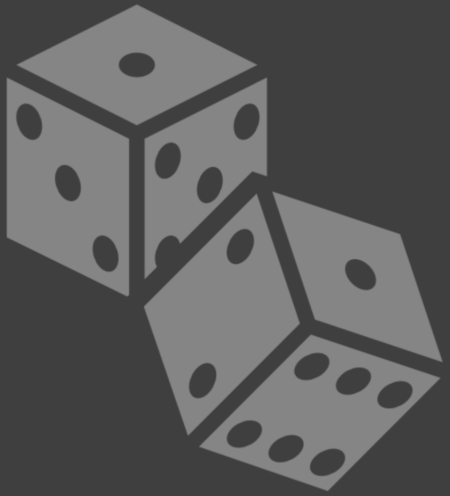


- Problem statement: drones need coordination.

Università degli Studi di Padova

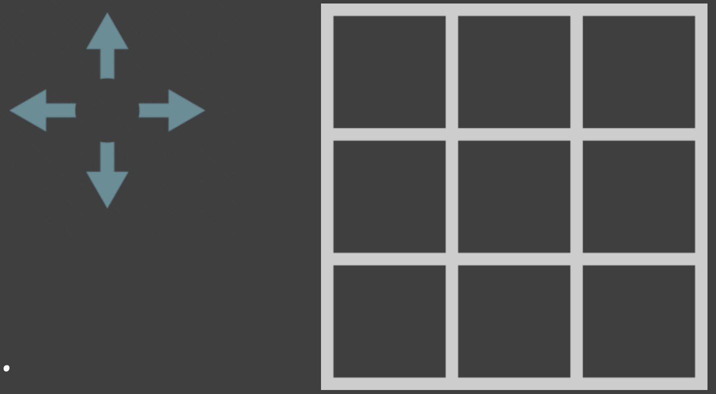
Why Game Theory?

- Motivation for the use of Game Theory
- Stochastic games
  - main idea: allowing to model drone's interactions
- Reinforcement learning (model-free)
  - used when learning problems arise
  - equilibrium learning vs adaptive learning algorithms



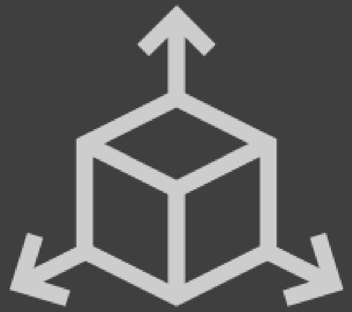
State of the Art

- **Two-dimensional** Grid games.



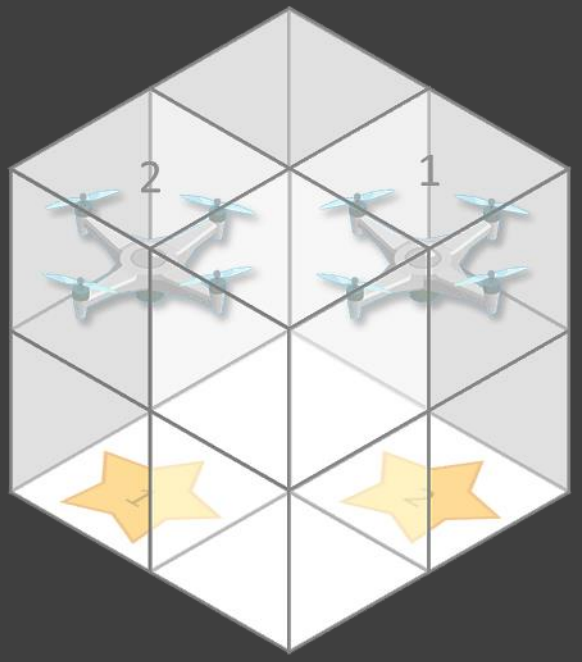
- Modeling problem for drones: only 2 dimensions.

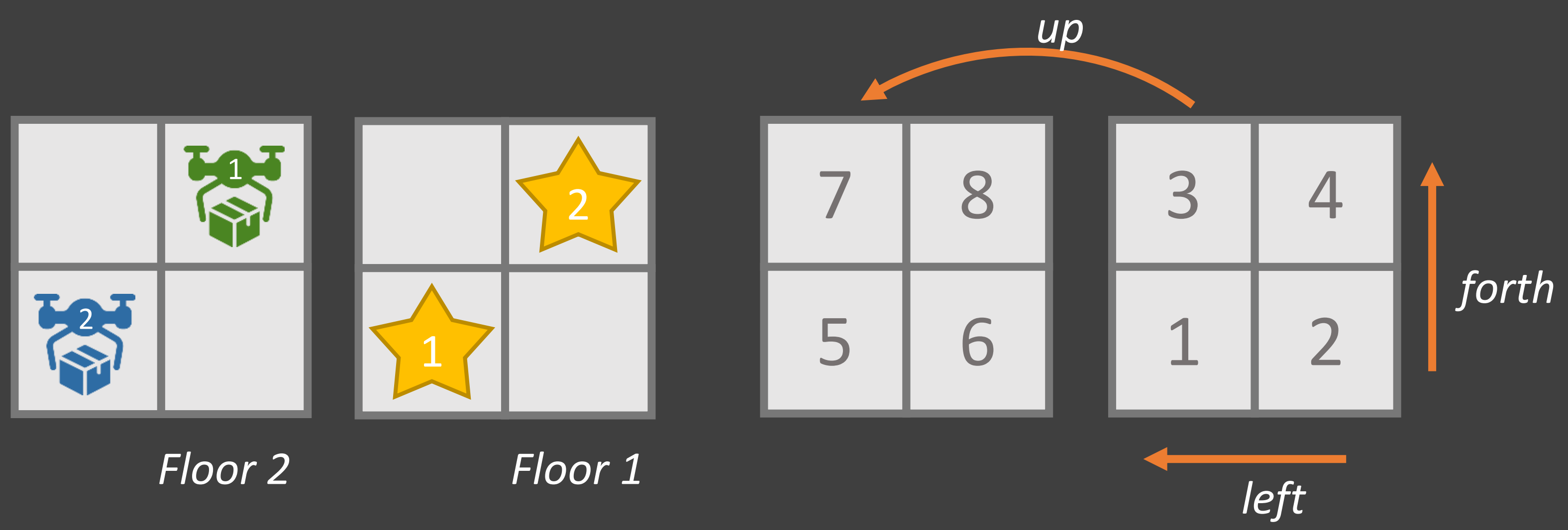
- Main idea of the contribution of our work:
  - expansion towards the **third dimension**



Innovative Scenario (1/2)

- Extended structure: the Three-Dimensional Grid Game
  - New variety of possible settings
- Possible moves (*Up, Down, Left, Right, Forth, Back*)

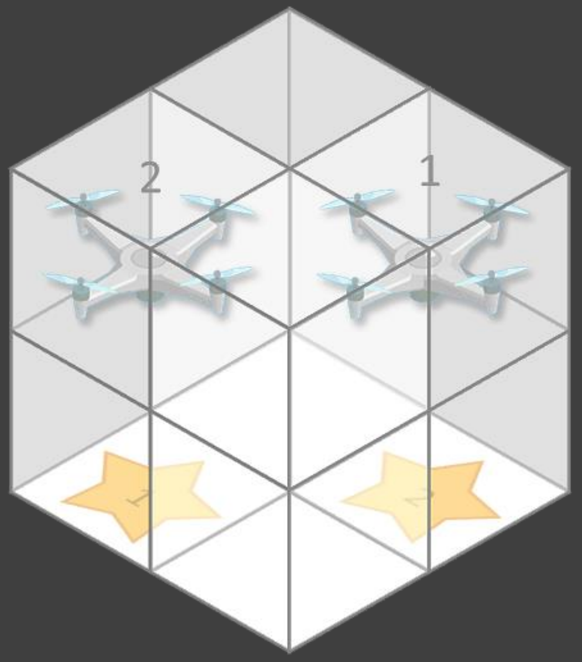




Innovative Scenario (2/2)

- Game structure/**rules**: drones

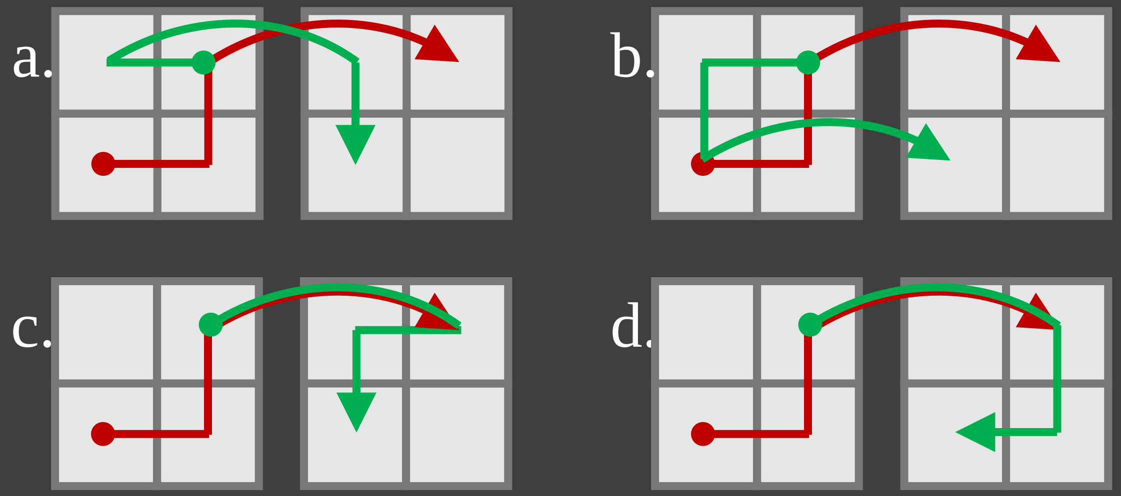
1. Choose their actions **simultaneously**.
2. When reaching the goal earn a **positive reward**.
3. **Game ends** as soon as a drone reaches its goal.
4. When moving into **the same cell** are **bounced back** to the previous.

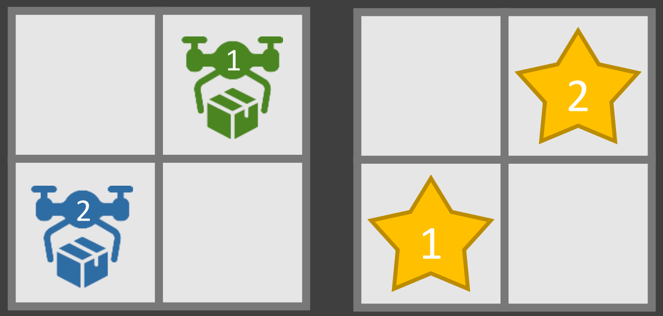


Game Analysis

- Assumptions:

  1. **Rewards** that each drone can receive:
     - 100 points if it reaches the goal position.
     - -1 points if it collides with the other drone.
     - 0 points otherwise.





Floor 2 *Floor 1*

  2. State transitions are **deterministic.**

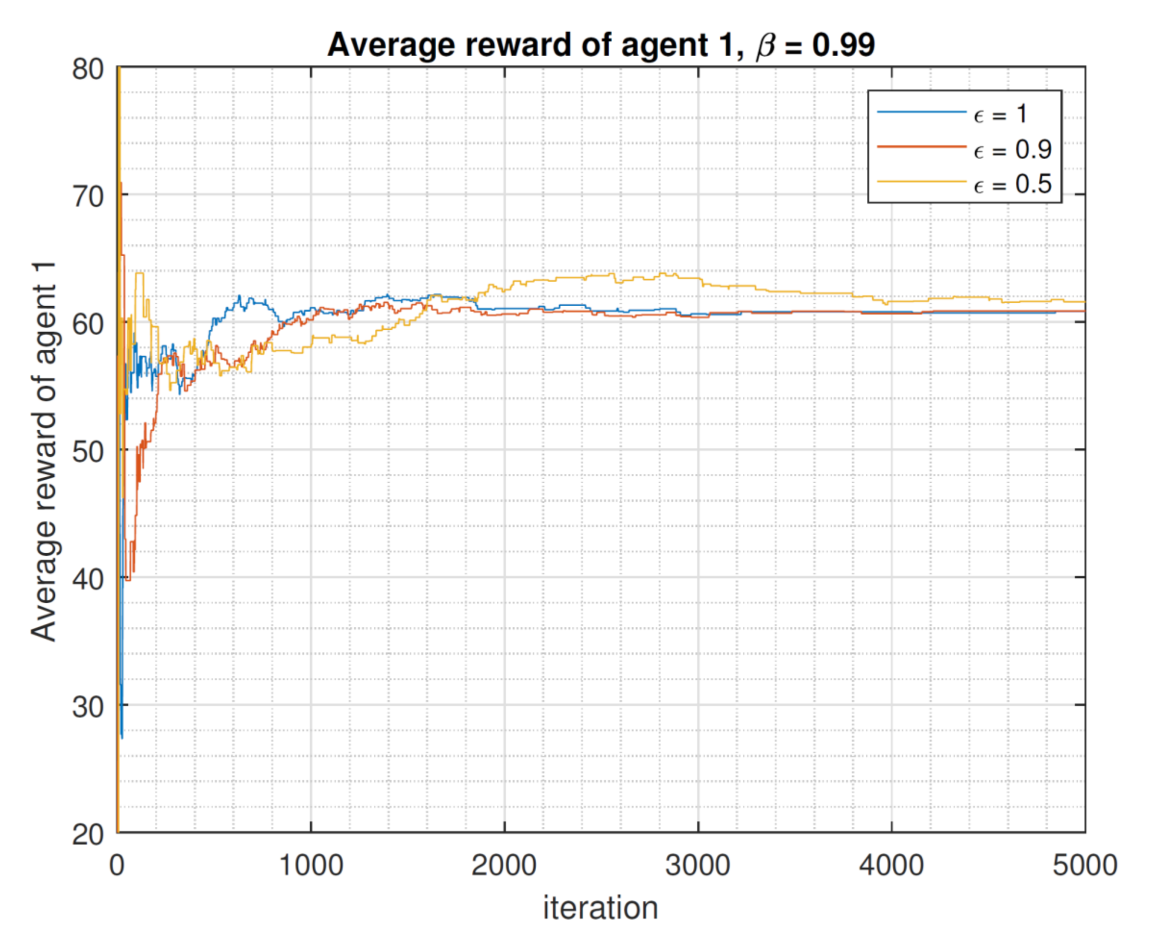
- Game Theoretical Analysis:

  7 possible **Nash Equilibria** obtained.

Nash-Q Algorithm

- Nash-Q Learning
  - Our version of Nash – Q algorithm by Hu & Wellman developed in Matlab.
  - Convergence in *self-play.*
  - Takes advantage of the Lemke Howson algorithm to find the NEs.
- Exploits ε-greedy exploration strategy
  - strategies to adopt can be explore, exploit, or explore and exploit.

Results (1/3)

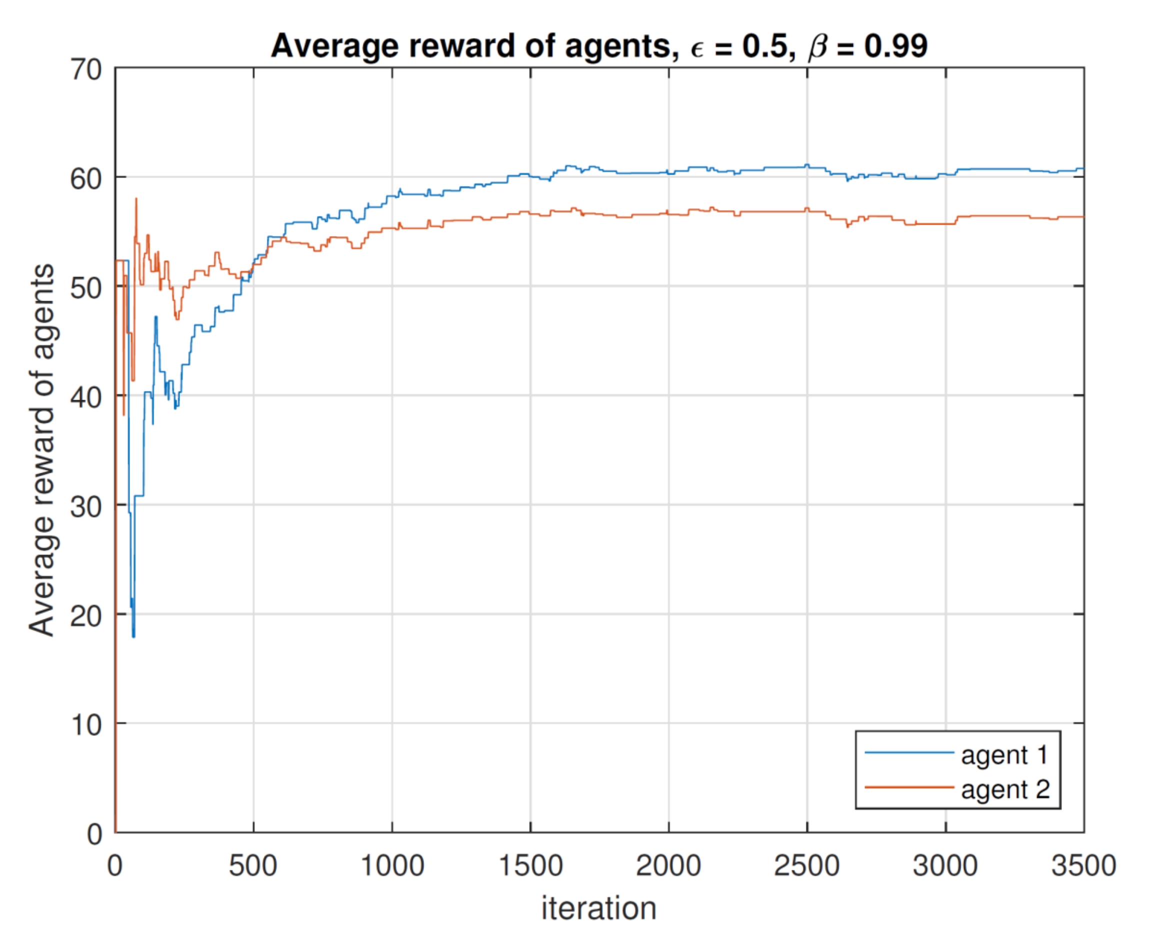


- *Convergence* of the algorithm for **both players** …
- … but with **different timing.**

- What does ε represent?
  - controls the probability of choosing the exploit strategy
  - How does the variation of ε change the final outcome?

Università degli Studi di Padova

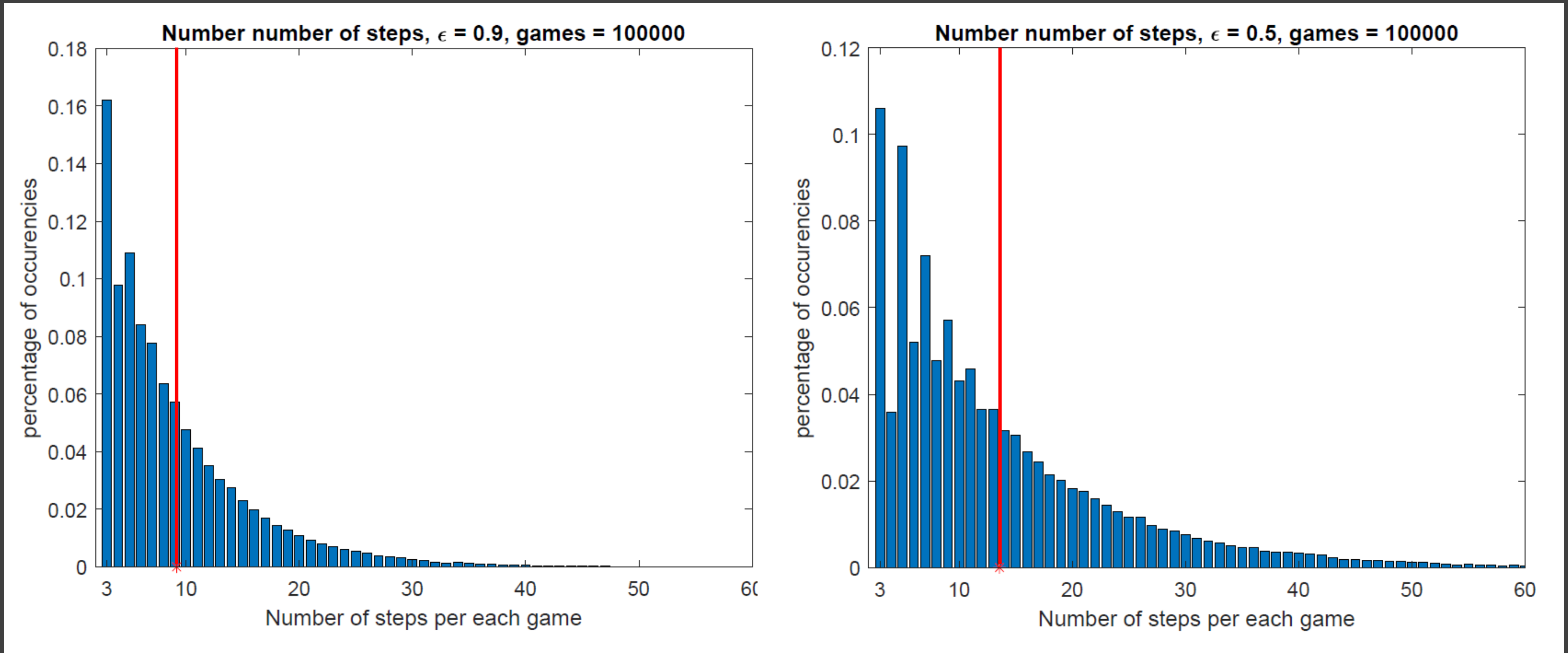
Results (2/3)



- What does β represent?
  - Discount factor.
- Average reward: similar behavior.
- **Stationarity** of the algorithm.

Results (3/3)

- New Metric introduced: **Average path length.**
- Number of steps per path distributed **geometrically.**



Conclusions

✓ Nash Equilibria verified.

✓ Convergence of Nash-Q algorithm.

✓ Average Path length evaluated.

- Possible extensions and Future works:
  - Include **more players/obstacles.**
  - Verify the solution with **other learning algorithms.**

Thank you!

Università degli Studi di Padova