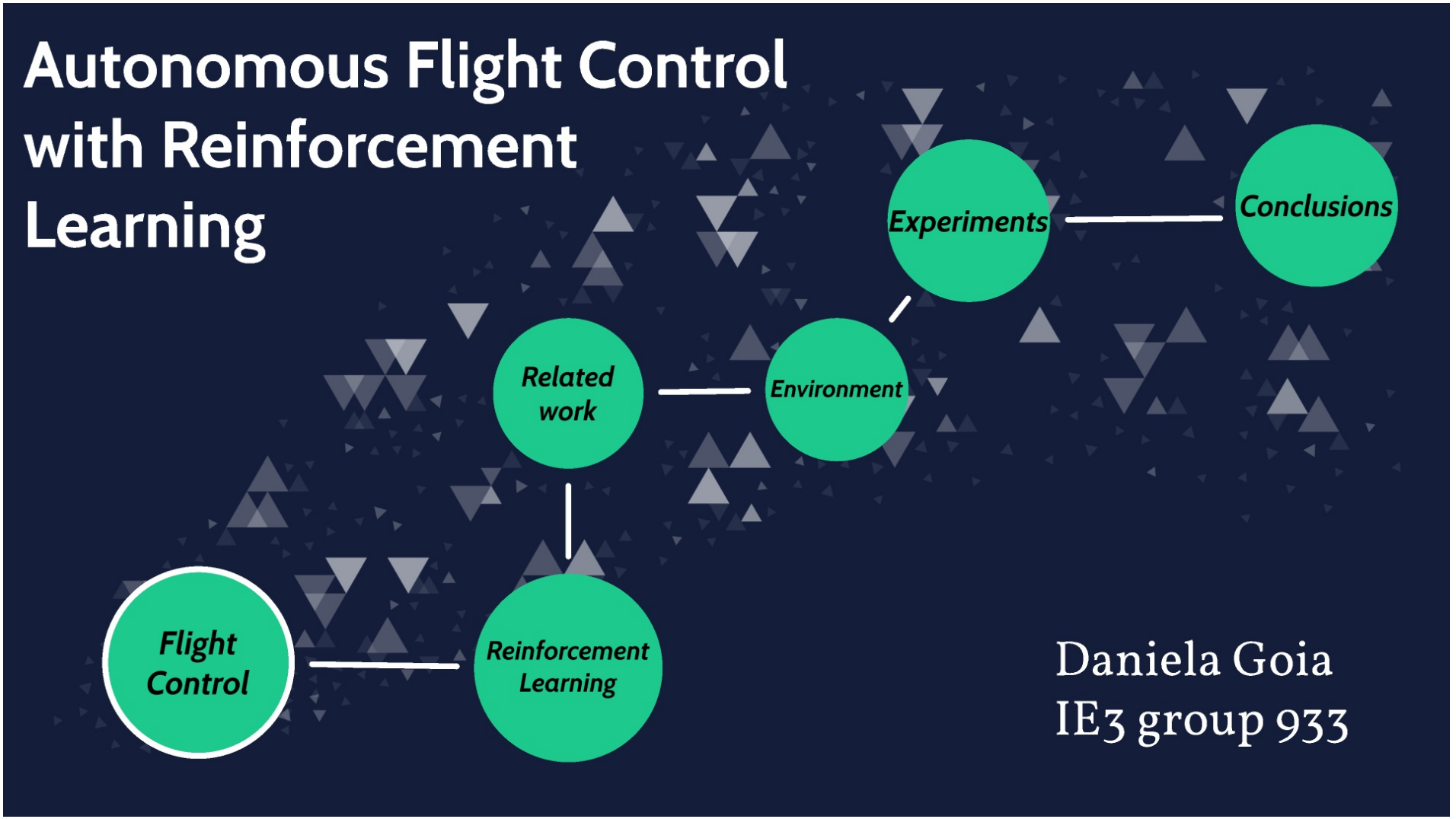
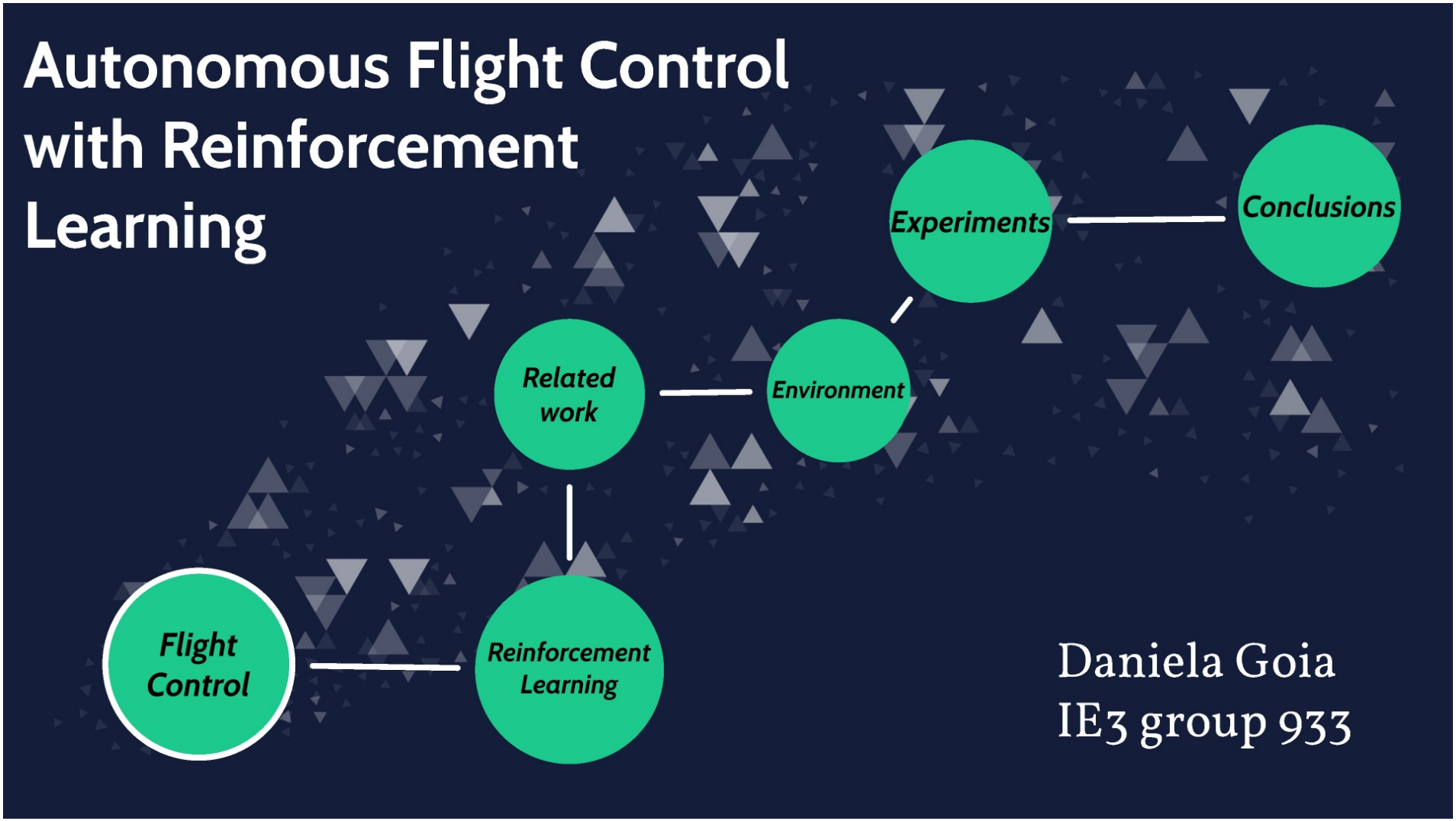
Autonomous Flight Control with Reinforcement Learning

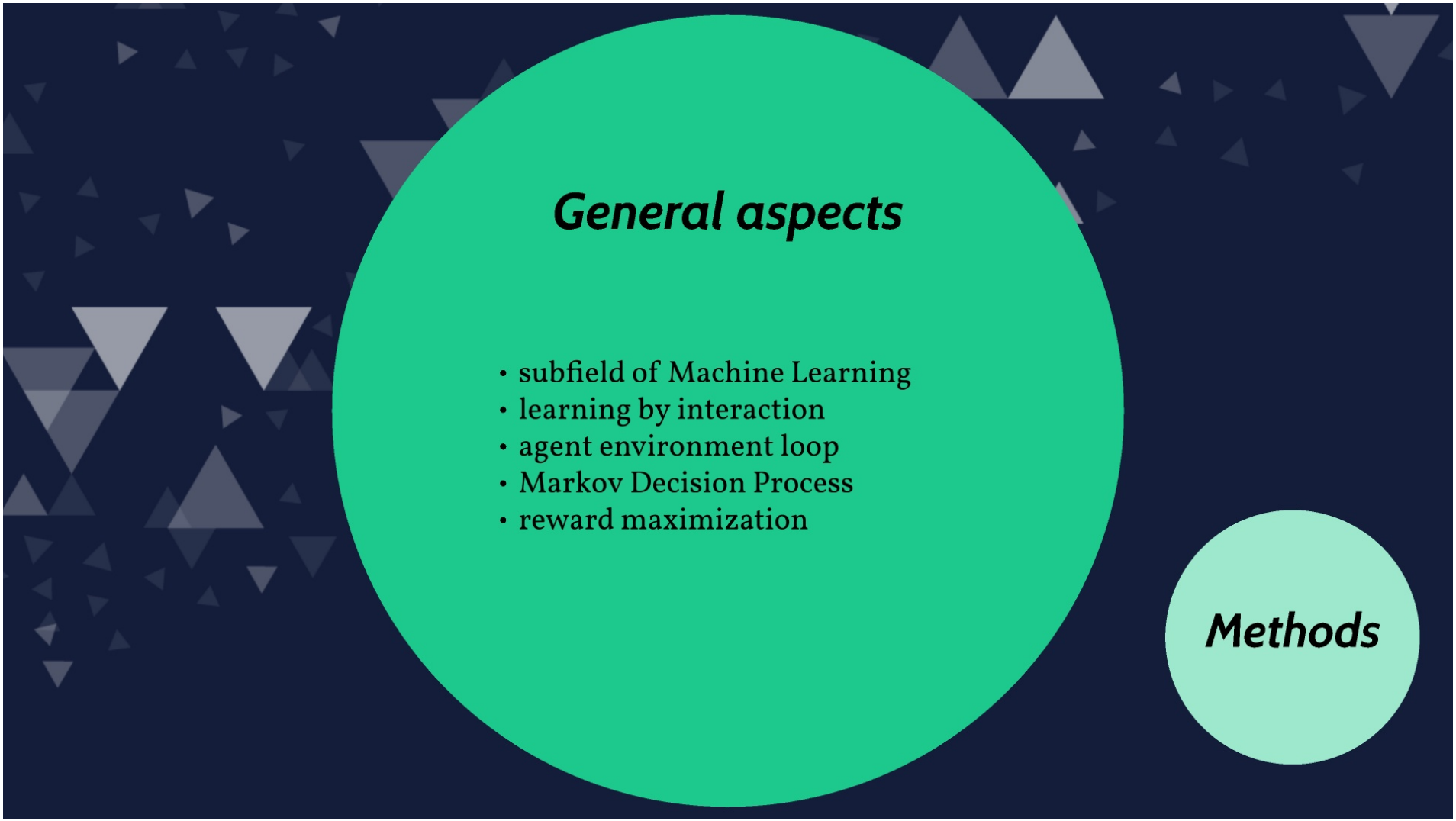


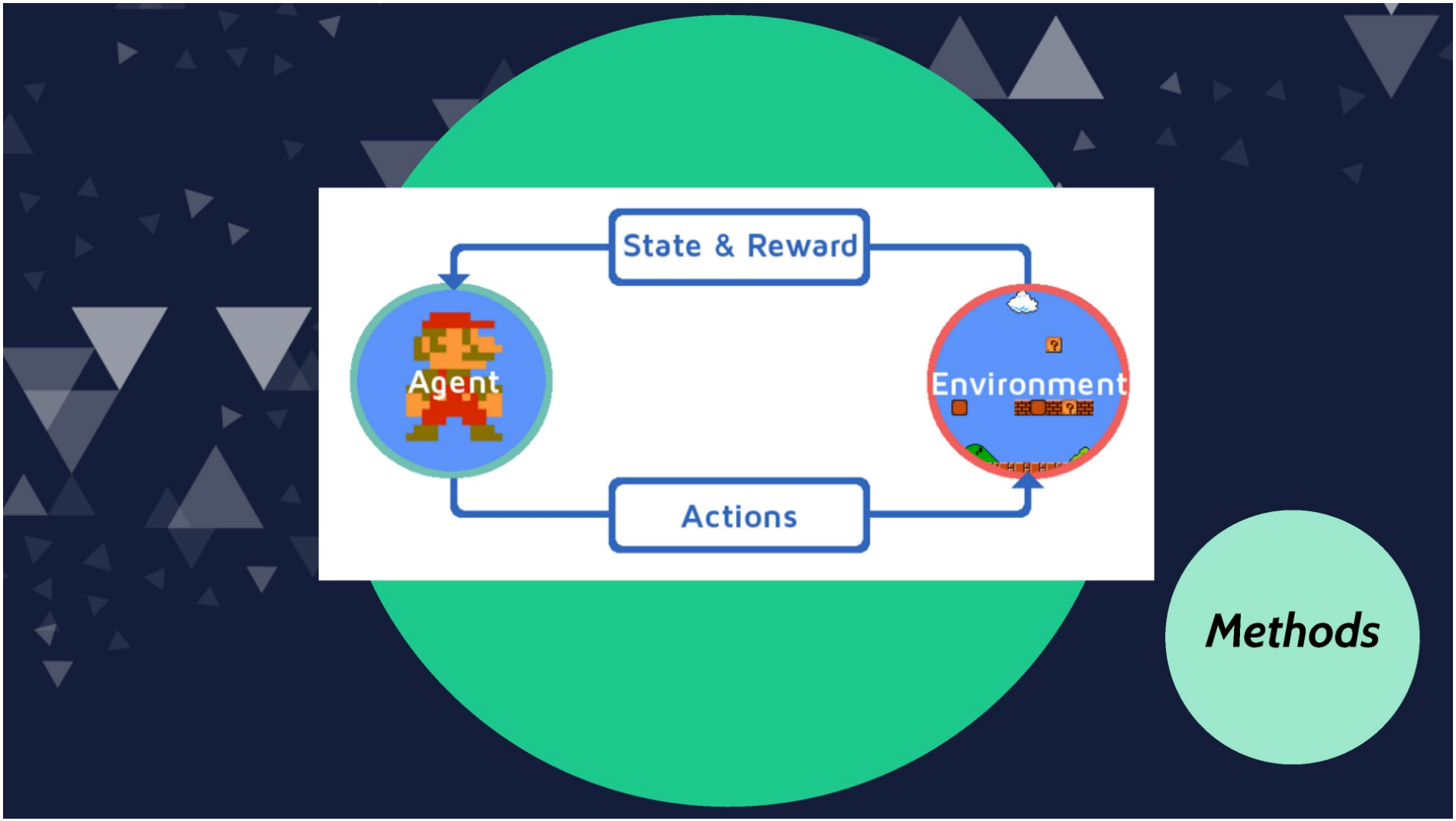
Daniela Goia
IE3 group 933

Motivation

- Launch type vehicles
- Autonomous decision making
- Descent trajectory optimization

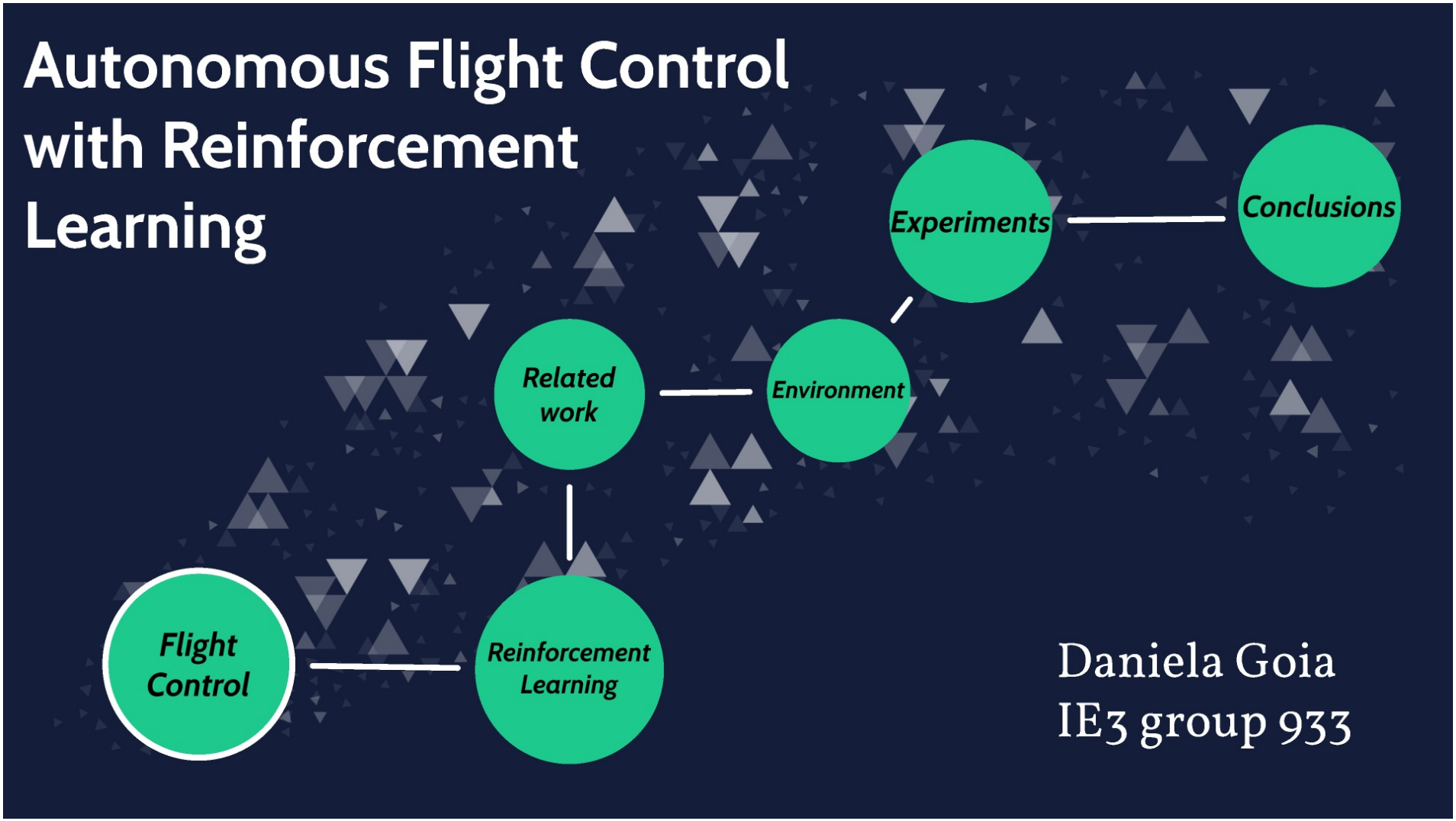

Autonomous Flight Control with Reinforcement Learning
Experiments
Conclusions
Related work
Environment
Flight Control
Reinforcement Learning
Daniela Goia
IE3 group 933


General aspects
• subfield of Machine Learning
• learning by interaction
• agent environment loop
• Markov Decision Process
• reward maximization
Methods


State & Reward
Agent
Environment
Actions
Methods

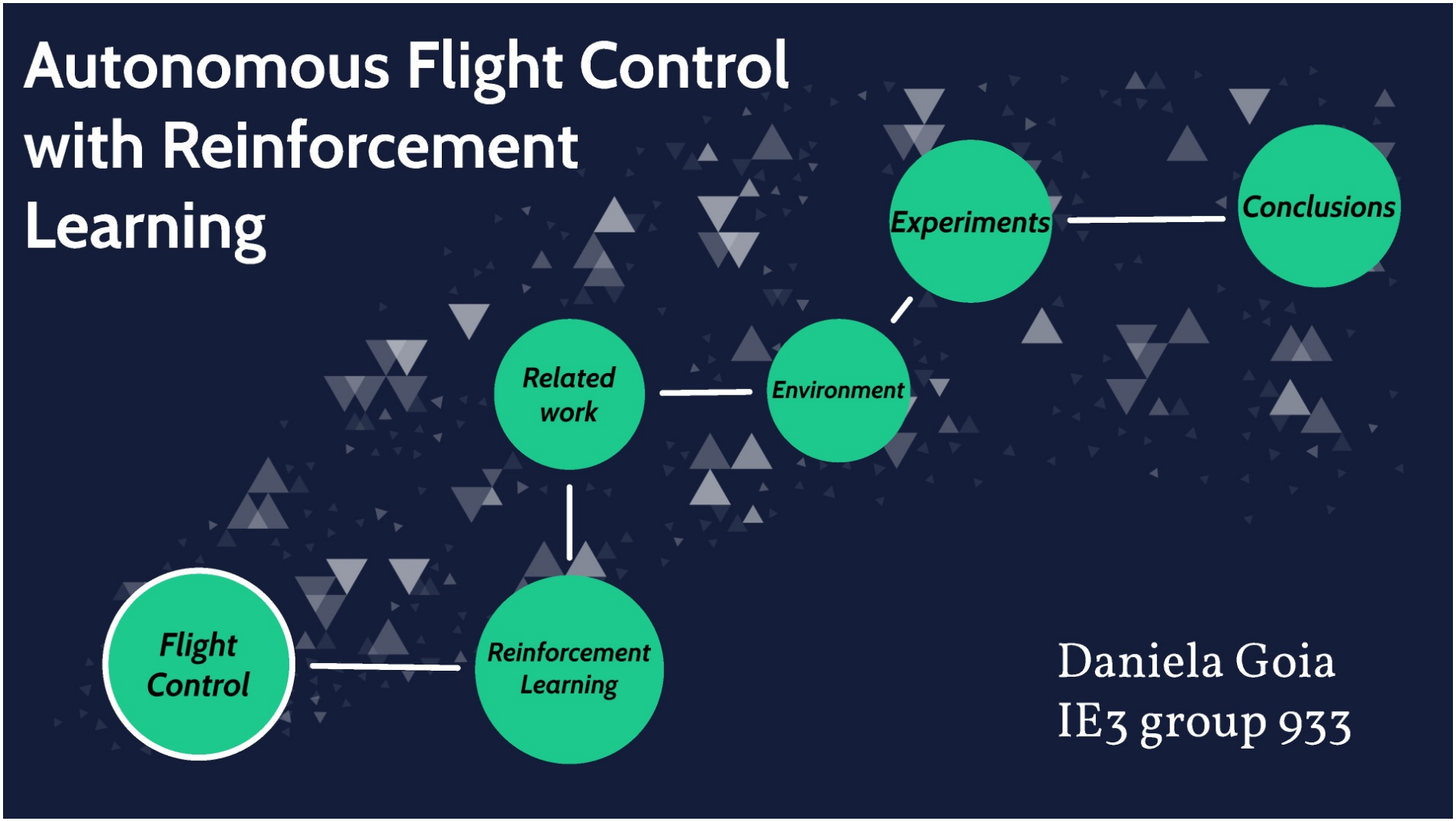
Methods

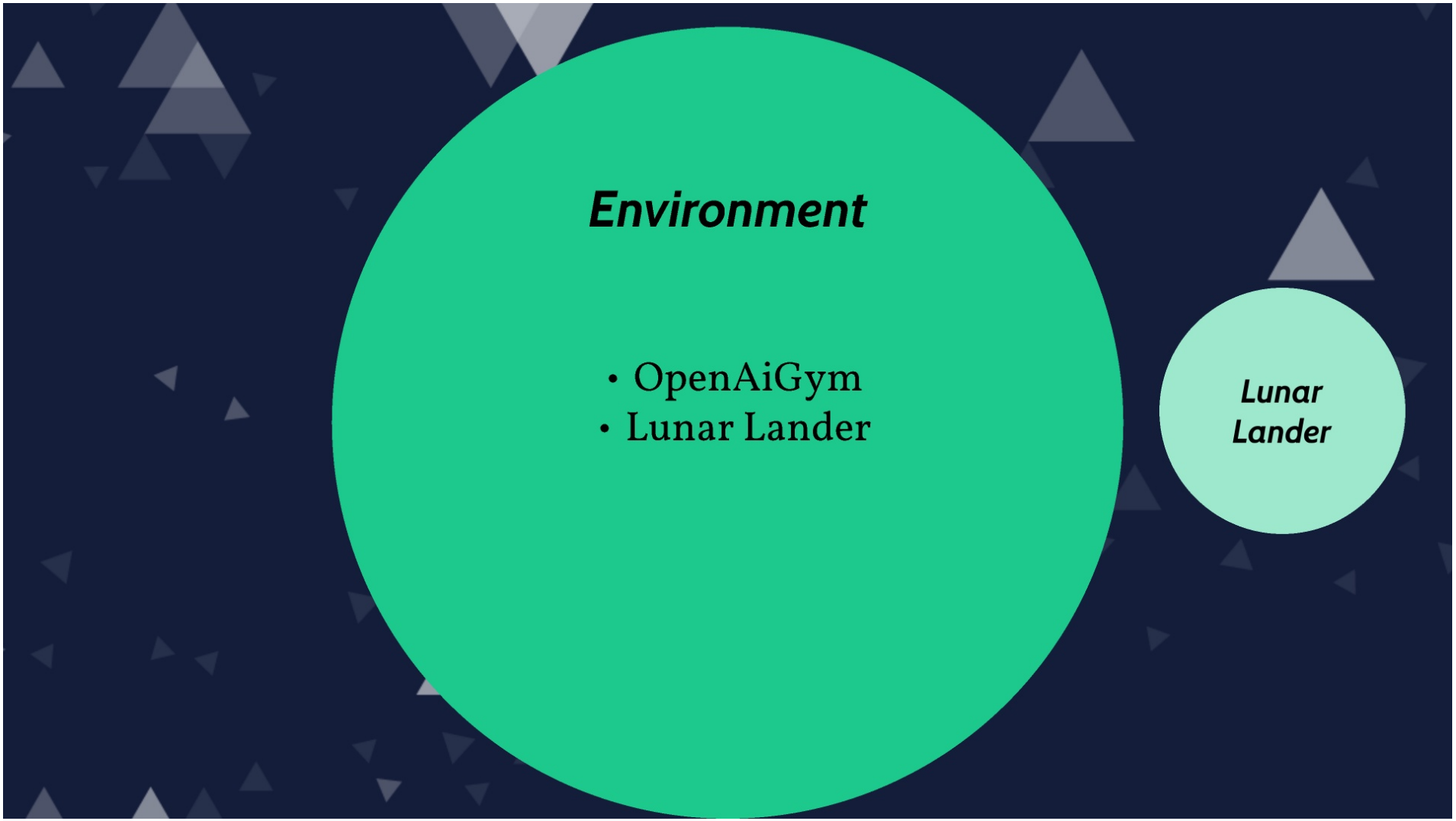
- goal
- Monte-Carlo
- Temporal Difference Learning
    - Q-Learning
- ANN


Autonomous Flight Control with Reinforcement Learning
Experiments
Conclusions
Related work
Environment
Flight Control
Reinforcement Learning
Daniela Goia
IE3 group 933

Related work

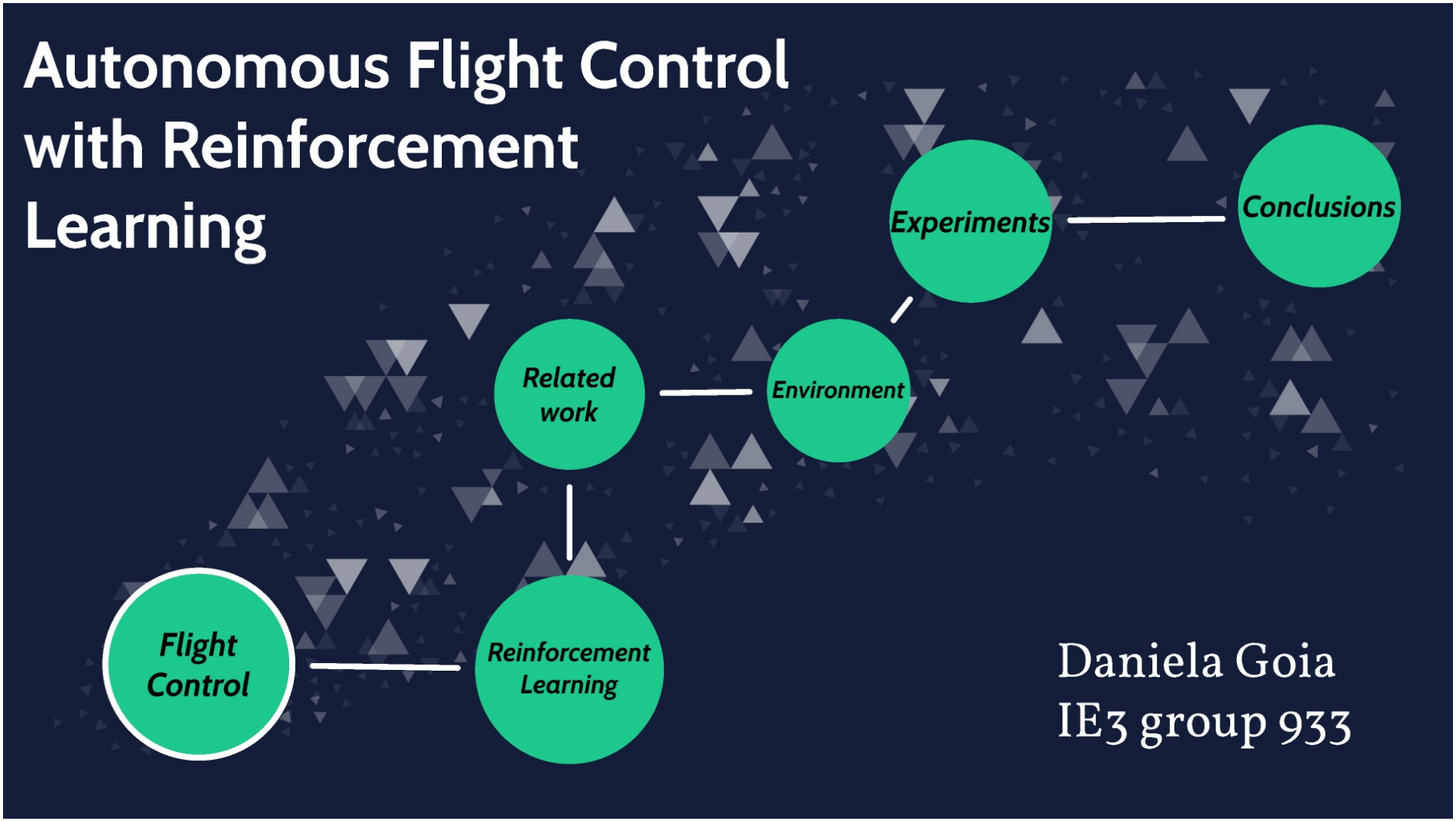
- Towards Intelligent Aircraft Through Deep Reinforcement Learning, S. Morrison*, A. Fisher, and F. Zambetta
- Deep Learning Quadcopter Control via Risk-Aware Active Learning, O. Andersson, M. Wzorek, P. Doherty
- Reinforcement Learning for UAV Attitude Control, W. Koch, R. Mancuso, R. West, A. Bestavros

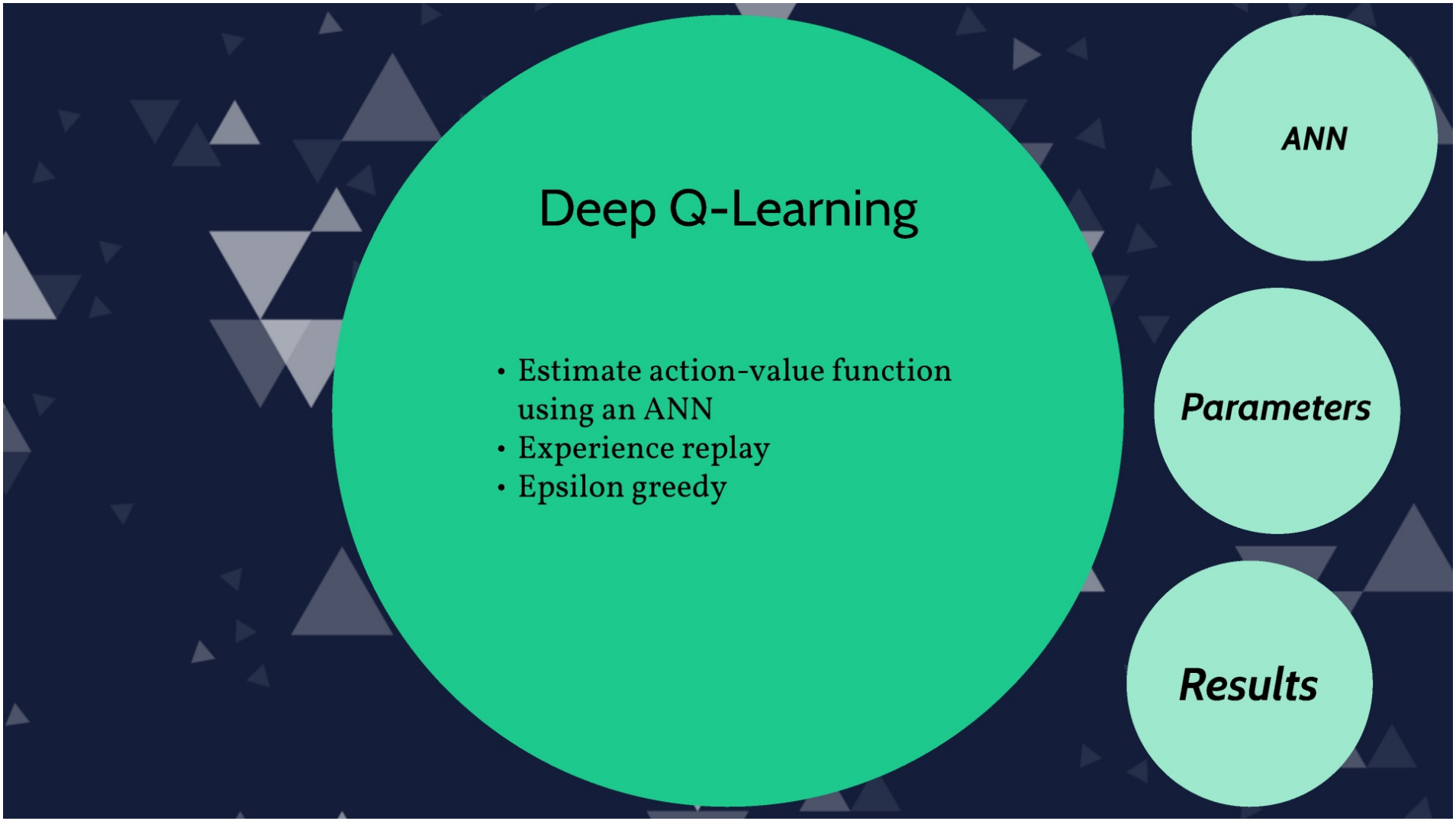

Autonomous Flight Control with Reinforcement Learning
Experiments
Conclusions
Related work
Environment
Flight Control
Reinforcement Learning
Daniela Goia
IE3 group 933


Environment
• OpenAiGym
• Lunar Lander
Lunar Lander

Rules

- state: position, velocity, angular velocity, contact with ground
- actions: fire one of 3 engines or do nothing
- rewards:
    - successful landing +100
    - crash -100
    - firing an engine -0.3


Autonomous Flight Control with Reinforcement Learning
Experiments
Conclusions
Related work
Environment
Flight Control
Reinforcement Learning
Daniela Goia
IE3 group 933

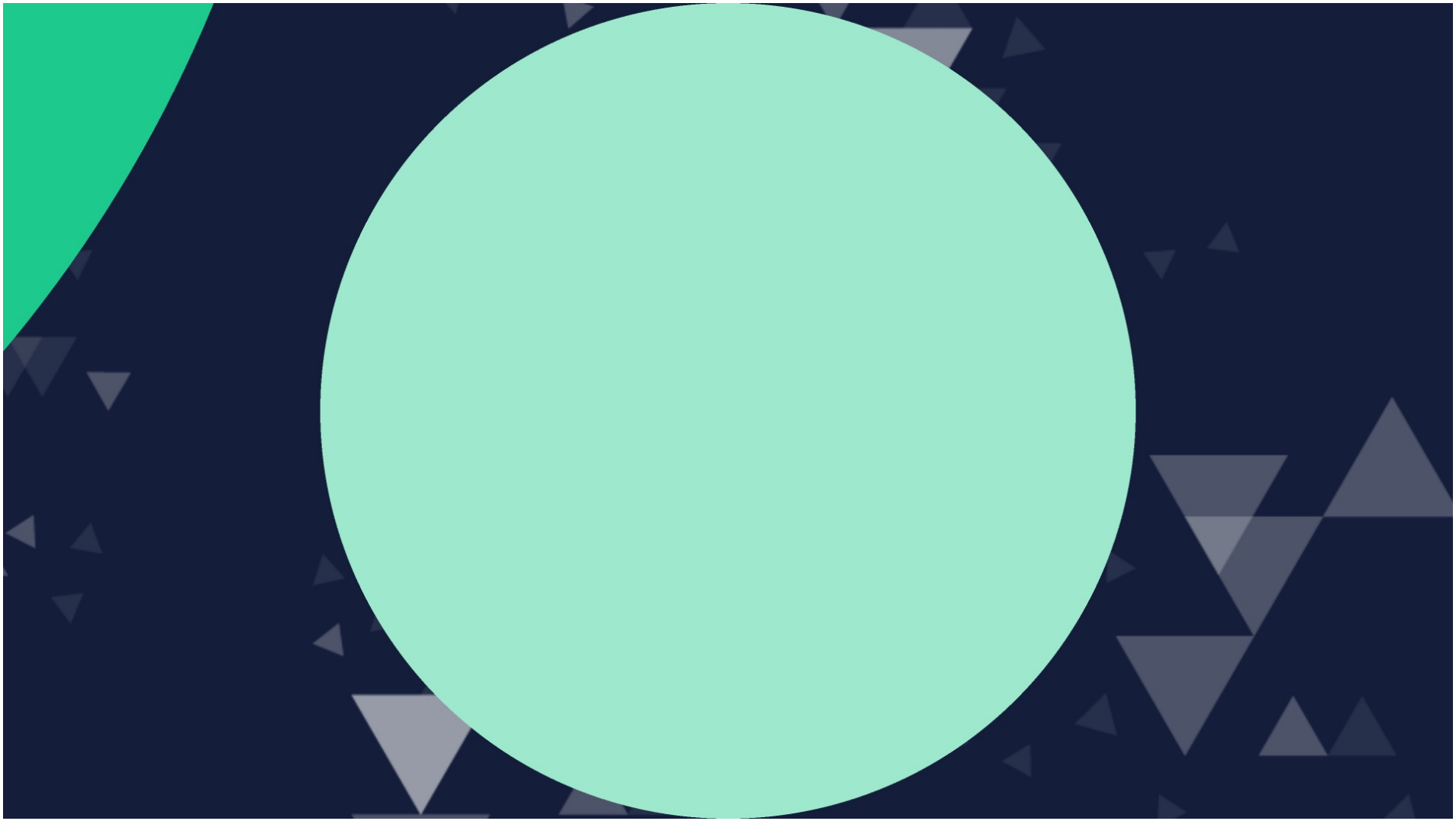

Deep Q-Learning
• Estimate action-value function using an ANN
• Experience replay
• Epsilon greedy
ANN
Parameters
Results

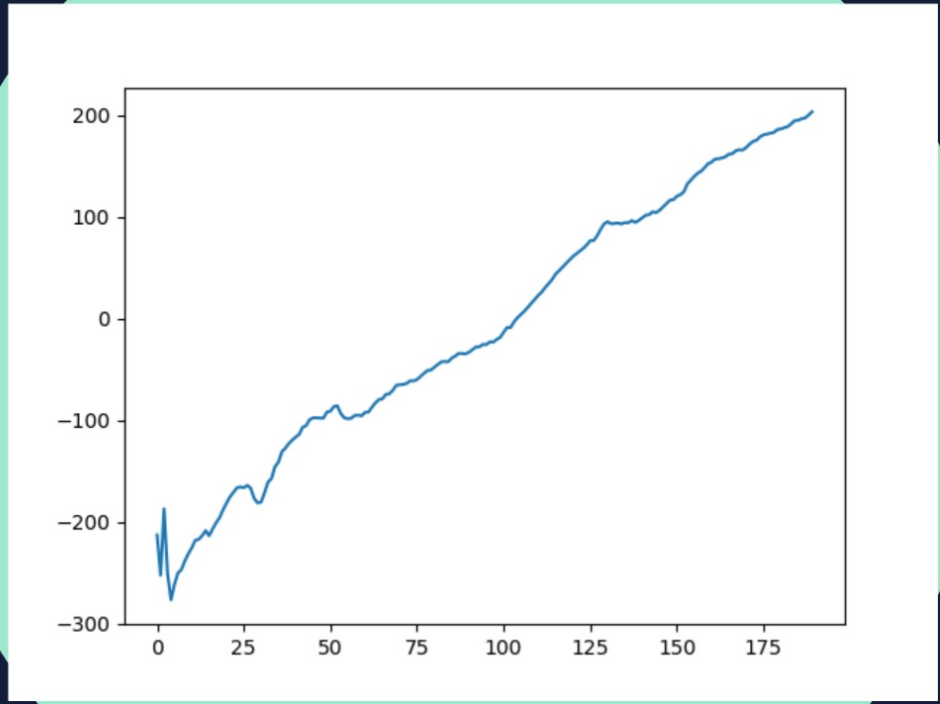
ANN Architecture

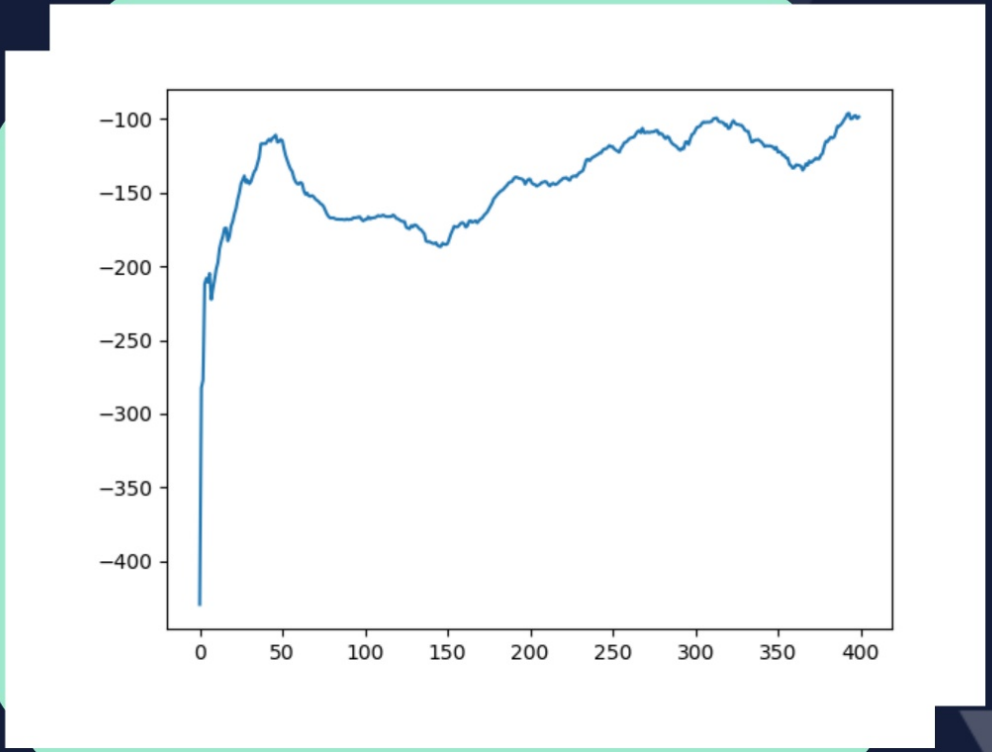
- input layer of length 8
- 2 hidden layers of 150 and 120 neurons
- relu activation between hidden layers
- linear activation at output layer
- one output neuron for each action

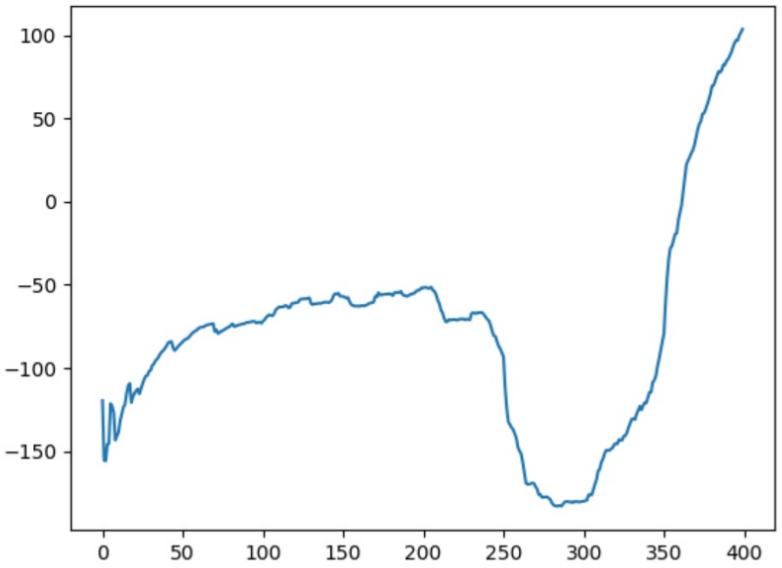
Parameters

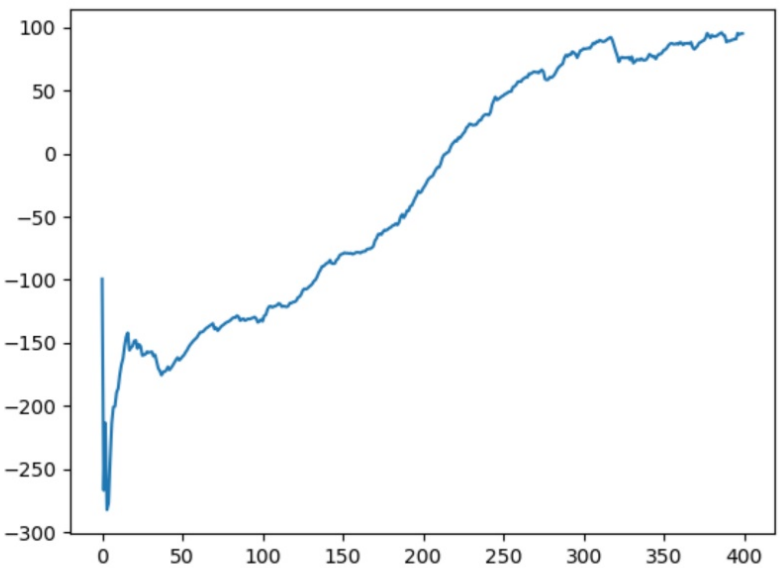
- initial epsilon: 1
- epsilon: 0.1
- epsilon decay:0.996
- alpha: 0.01
- gamma: 0,99
- batch size: 64
- max replay memory size:1M

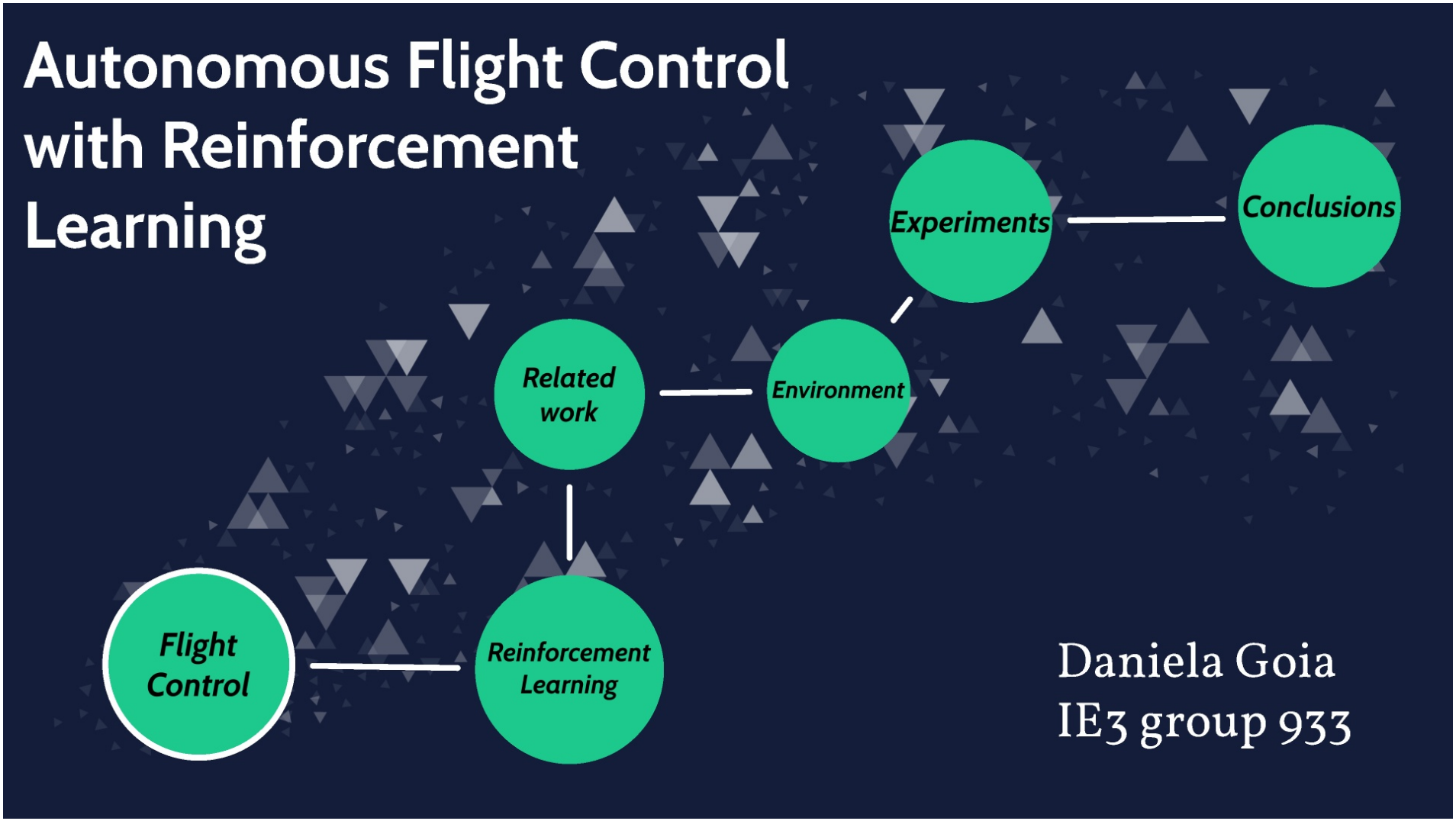



200
100
0
−100
−200
−300
0
25
50
75
100
125
150
175


−100
−150
−200
−250
−300
−350
−400
0
50
100
150
200
250
300
350
400

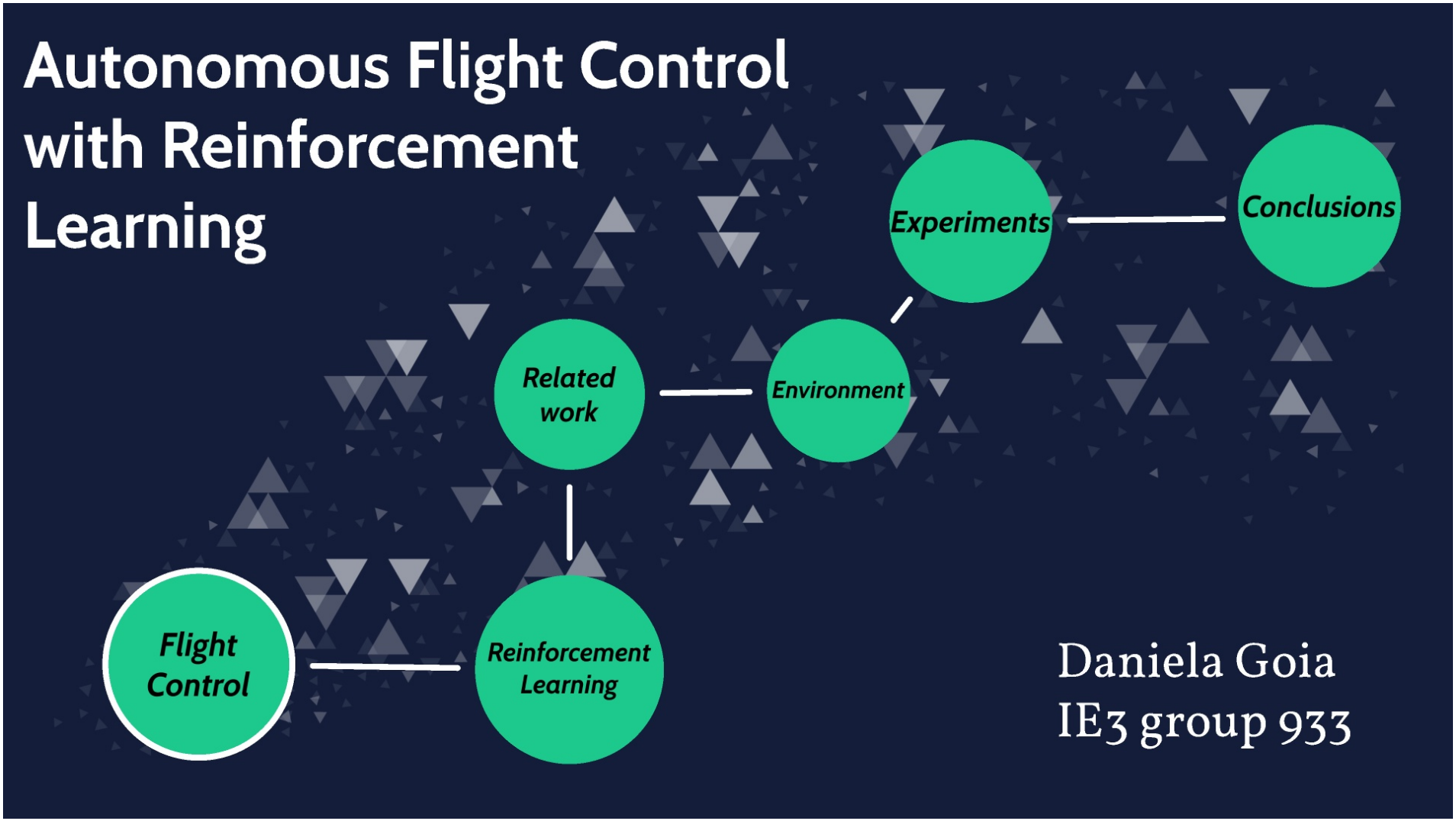

100
50
0
−50
−100
−150
0
50
100
150
200
250
300
350
400


100
50
0
−50
−100
−150
−200
−250
−300
0
50
100
150
200
250
300
350
400


Autonomous Flight Control with Reinforcement Learning
Experiments
Conclusions
Related work
Environment
Flight Control
Reinforcement Learning
Daniela Goia
IE3 group 933

Conclusions

- Transferability
- Small replay size -> weaker performance
- Lack of anticipation-> unsatisfactory running average
- Exploration matters


Autonomous Flight Control with Reinforcement Learning
Experiments
Conclusions
Related work
Environment
Flight Control
Reinforcement Learning
Daniela Goia
IE3 group 933