ELTE | IK
INFORMATIKAI KAR

Multi-Agent Simulation & Learning

Introduction to Machine Learning
Computer Science BSc Course, ELTE Faculty of Informatics

László Gulyás

Associate Professor
Department of Artificial Intelligence
lgulyas@inf.elte.hu

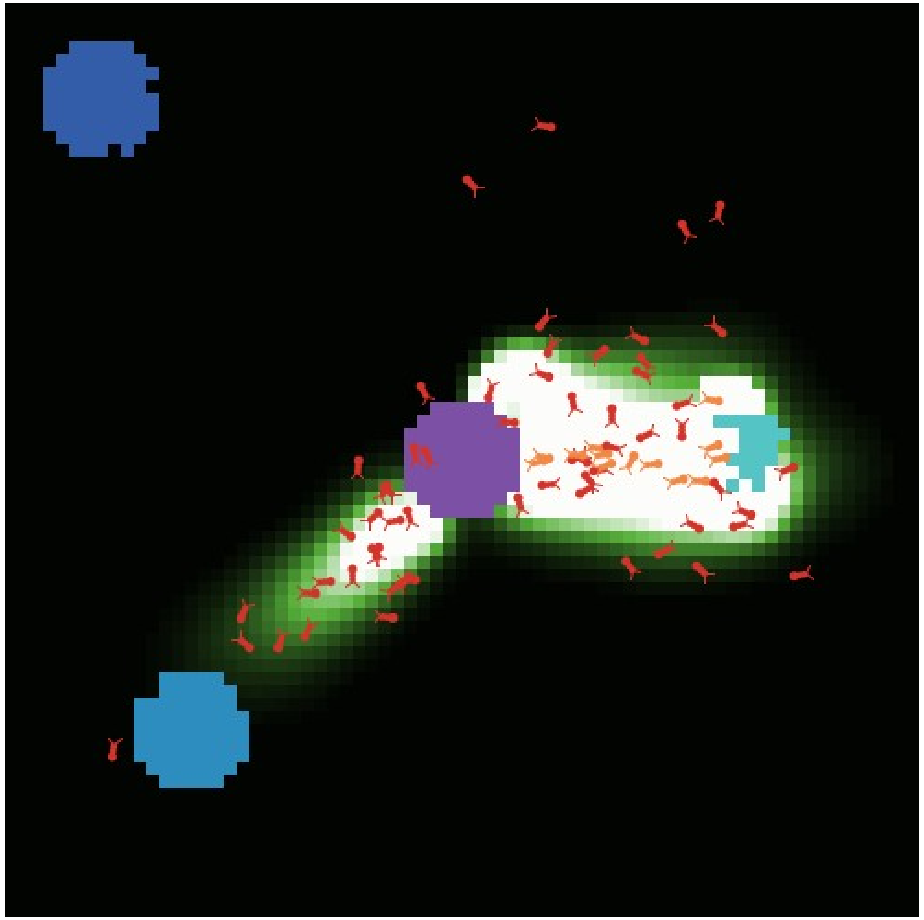
ELTE | IK
DEPARTMENT OF ARTIFICIAL INTELLIGENCE

Motivation

Multi-Agent Systems

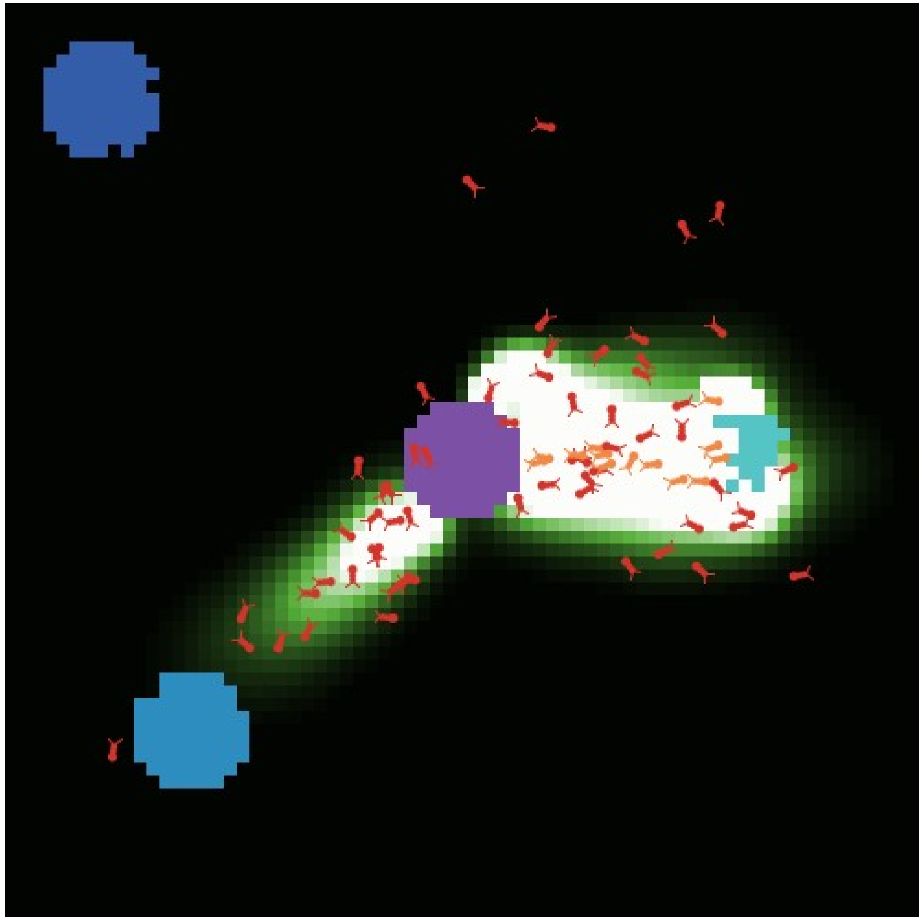
Who is Intelligent?

- Foraging Ants



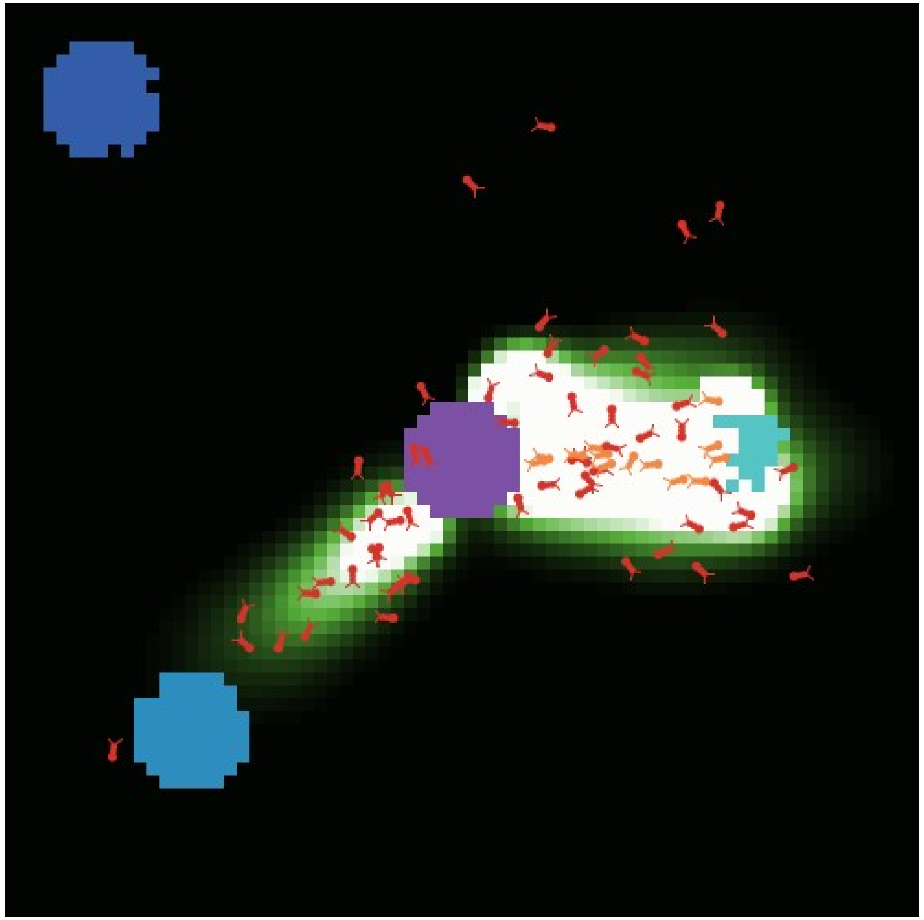
Who is Intelligent?

- Foraging Ants
  - Intelligent?
  - Solves the problem
  - Robust and cheap…



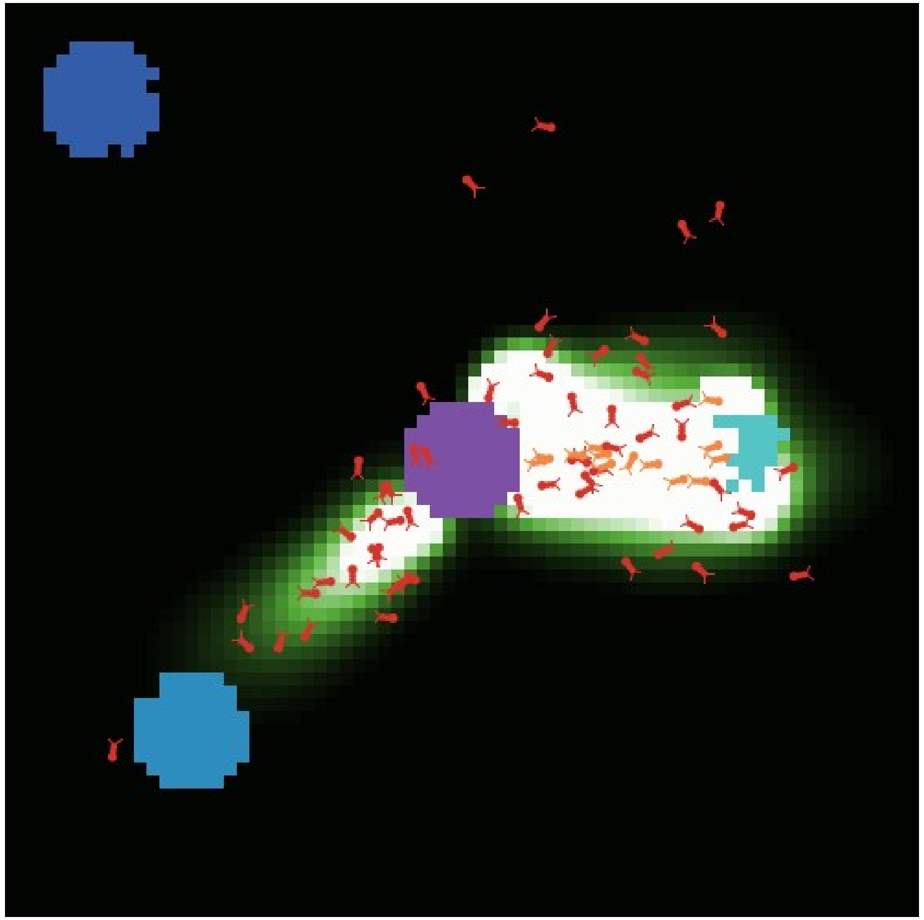
Who is Intelligent?

- Foraging Ants
  - Intelligent?
  - Solves the problem
  - Robust and cheap…
- In the eye of the beholder



Who is Intelligent?

- Foraging Ants
  - Intelligent?
  - Solves the problem
  - Robust and cheap…
- In the eye of the ~~beholder~~ observer
  - And there are different levels of observation (consideration)



Micro and Macro Behavior

- Building Systems
  - From top-down
  - From bottom-up
- When parts also have autonomy
  - Make decisions (have behavior)
  - Are potentially rational/intelligent
  - Come from different (perhaps unknown) provenance
  - Have potentially different goals
- When intelligence is (also) measured at system level
  - Collective Intelligence

Unintended consequences
Driving on the road

© CGP Grey https://www.cgpgrey.com/

The micro-macro dichotomy

How to work with it?

Basic Notion

Emergence

- Something observable at a ‚higher level'
  - That does not exist at ‚lower levels'
  - That is not expressable at ‚lower levels'

- This can be a property or a behavior

- Hard concept to define formally
  - Subject to discussions and interpretations

- **Yet, at the core of ‚Micro vs Macro'**

Two Related Topics

Analysis:
Agent-Based Simulation

Experimenting with

- Possible *micro* rules that
- *Generate*
- (observed) *macro* behavior

→ A tool of explanation

Construction:
Multi-Agent (Reinf.) Learning

- Learning to Cooperate
- Learning to Swarm

Agent-Based Simulation

An Example, A Tool

New Segregation: Races accept divide

[US News Excerpt]

Exclusive Report

New segregation: Races accept divide

But experts say financial, social costs of living apart are just as detrimental

By Ron French / *The Detroit News*

More than four decades after Americans fought and died to end segregation, many in Metro Detroit are comfortable living apart.

Why?

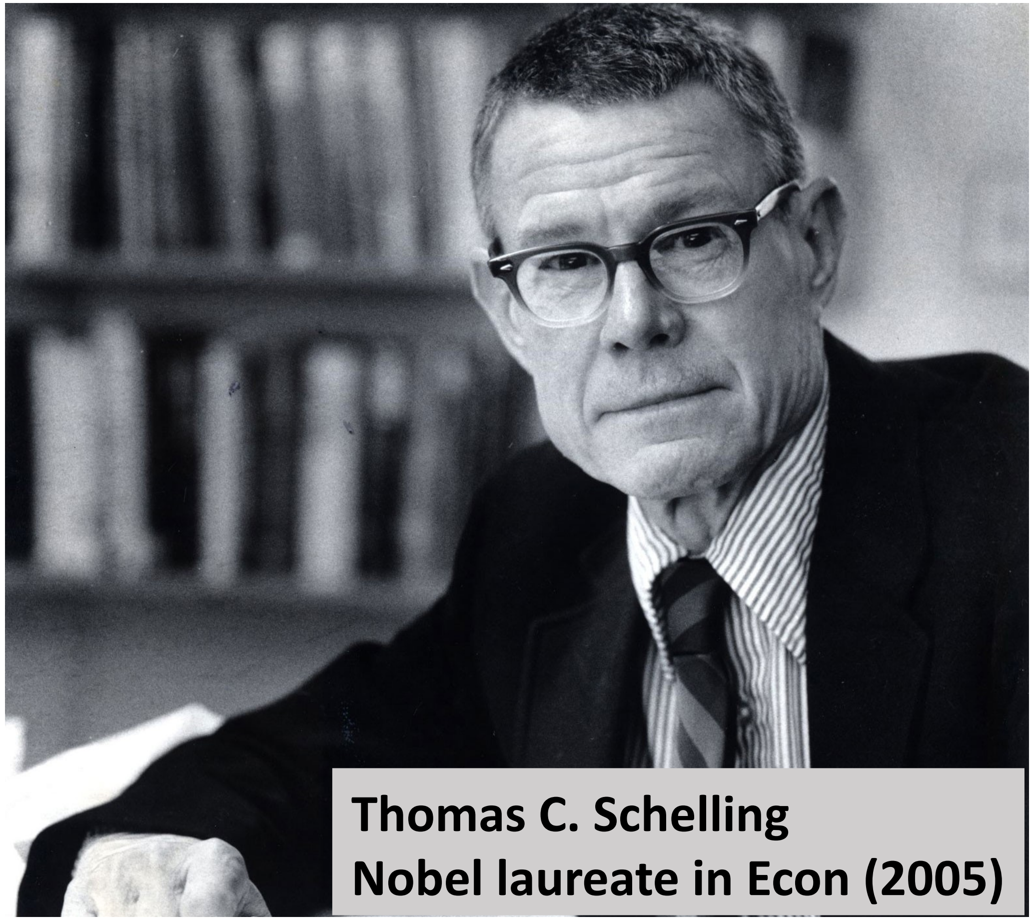
- Endless number of opinions
  - Economical reasons
  - The level of racism in the society
  - Etc.

- Untested opinions
  - To test: *simplifications* → *models*

A Model of Housing Segregation by *Thomas C. Schelling* (1978)

- Residential choices
  - Stylized 2D environment
  - “Oranges” and “Blues”
  - The level of personal tolerance
    (given as a %)

- **Where does a general 60% tolerance level lead us?**
  - **And 70%?**



Thomas C. Schelling
Nobel laureate in Econ (2005)

A Model of Housing Segregation by *Thomas C. Schelling* (1978)

Tolerance (int) in [0,100]: (%)

Neighbors (int) in [0,100]:
  #neighbors_of_other_color //
  #all_neighbors

Happy (boolean):

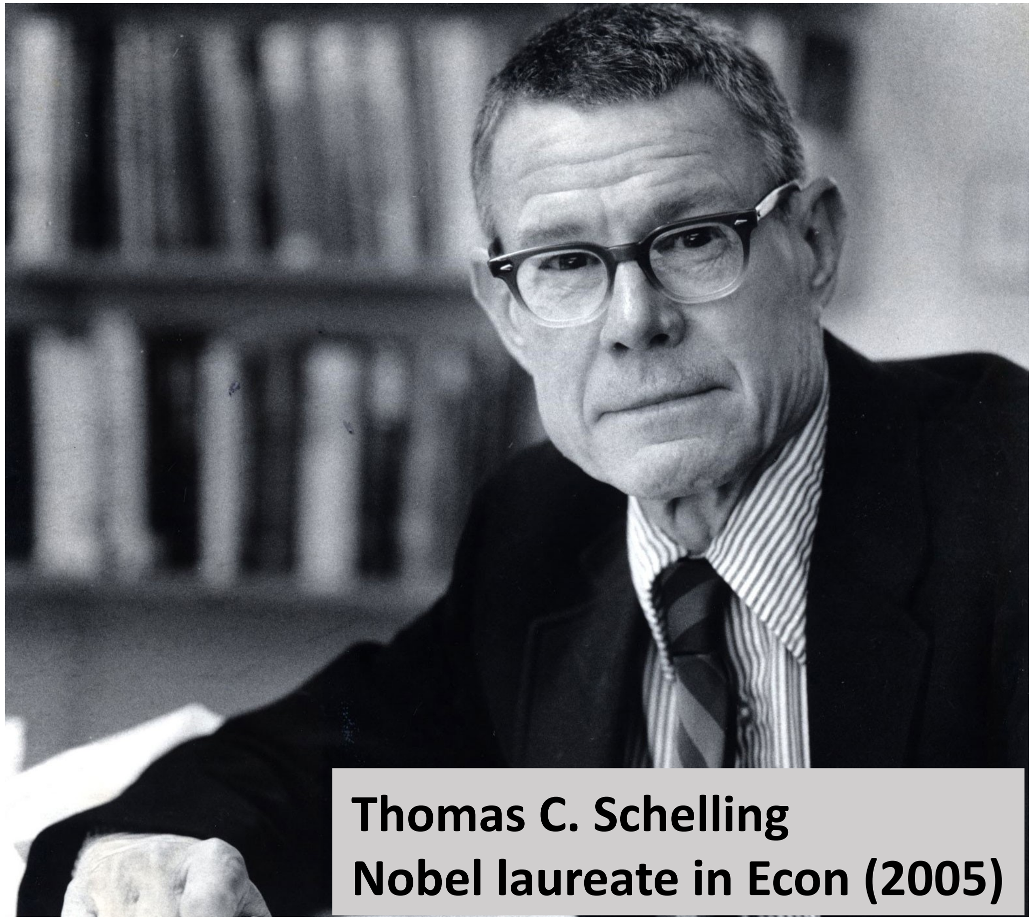
  Neighbors <= Tolerance

Movement Rule:

```
If (not Happy):
  Move_to_an_empty_location
```

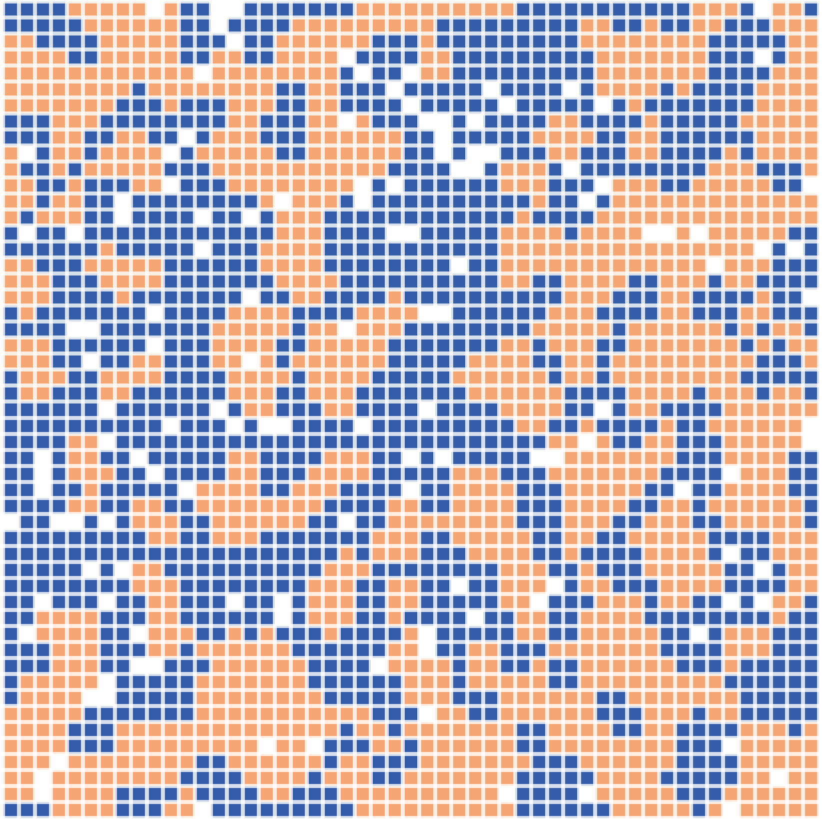
Variants:

- Move_to_random_empty_location_if_happy_there
- Move_to_closest_empty_location
- Etc.
- Various neighborhood definitions possible

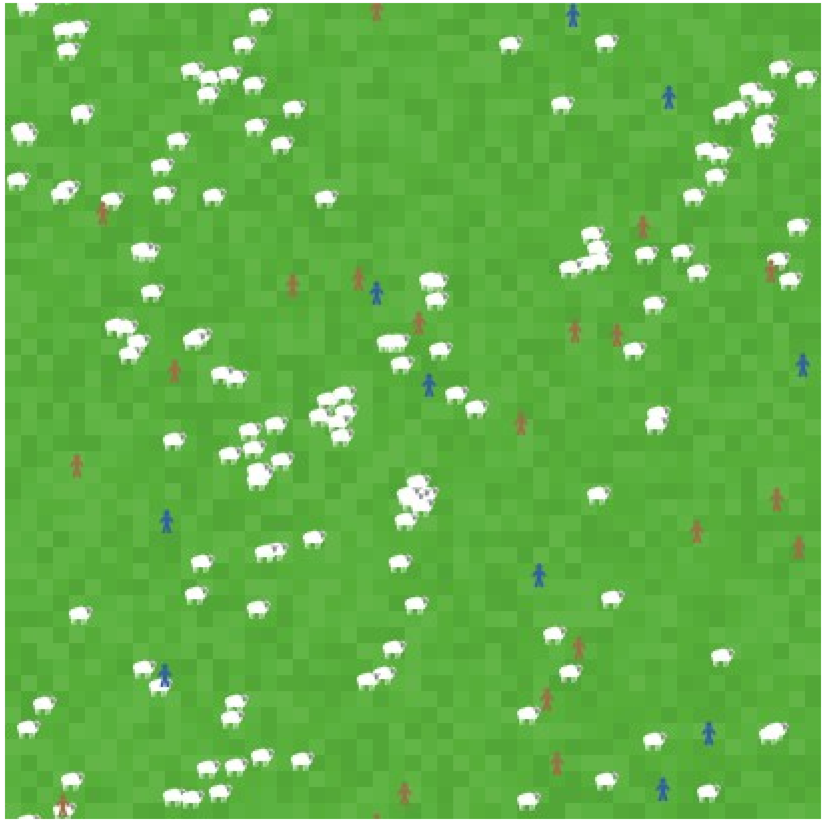


70 percent tolerance, random initialisation

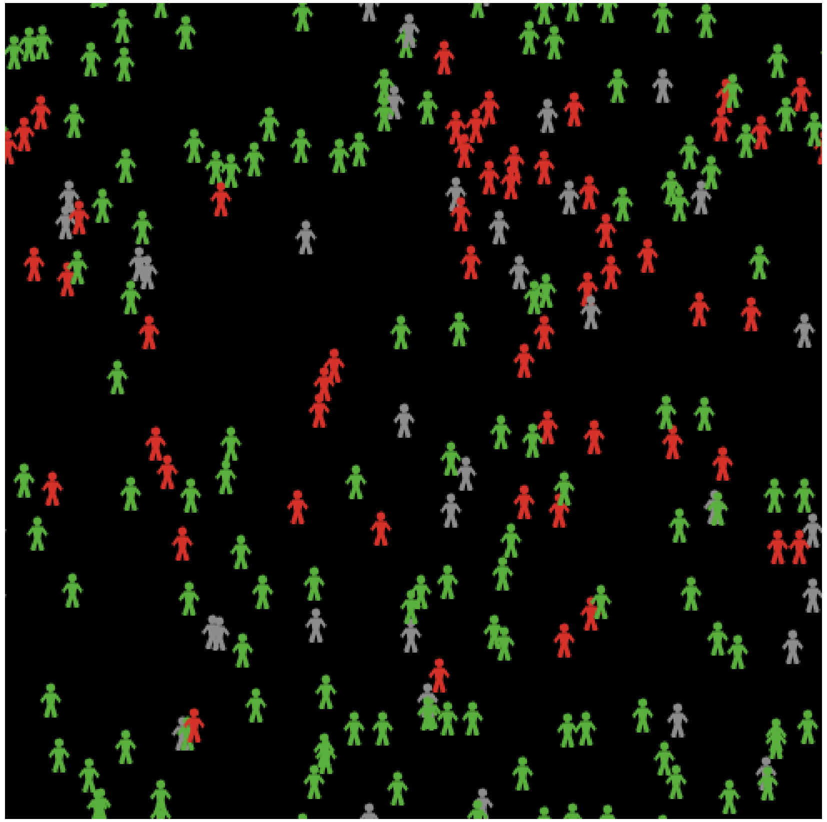
70 percent tolerance, random initialisation



NetLogo
https://ccl.northwestern.edu/netlogo/



Models library/Biology/Shepherds



Models library/Biology/Virus

Multi-Agent (Reinf.) Learning

Some Examples

Goal – „Learning to Swarm”

- Training teams of mobile robots
  - To solve a problem collectively
  - → Multi-Agent Reinforcement Learning (MARL)

- ‚Swarm robotics in spirit’:
  - No control hierarchy
  - No explicit, direct communication
  - No complex negotiation protocols
  - No explicit dependence on team size
    - Cf. Graceful degradation

Background of Examples

Technical Setup

- StableBaseline3
  - Multi layer perceptron (MLP)
  - LSTM
- Training
  - Proximal Policy Optimisation
  - Q-learning
  - Actor-Critic Model
- Negative reward for collision

Challenges Explored

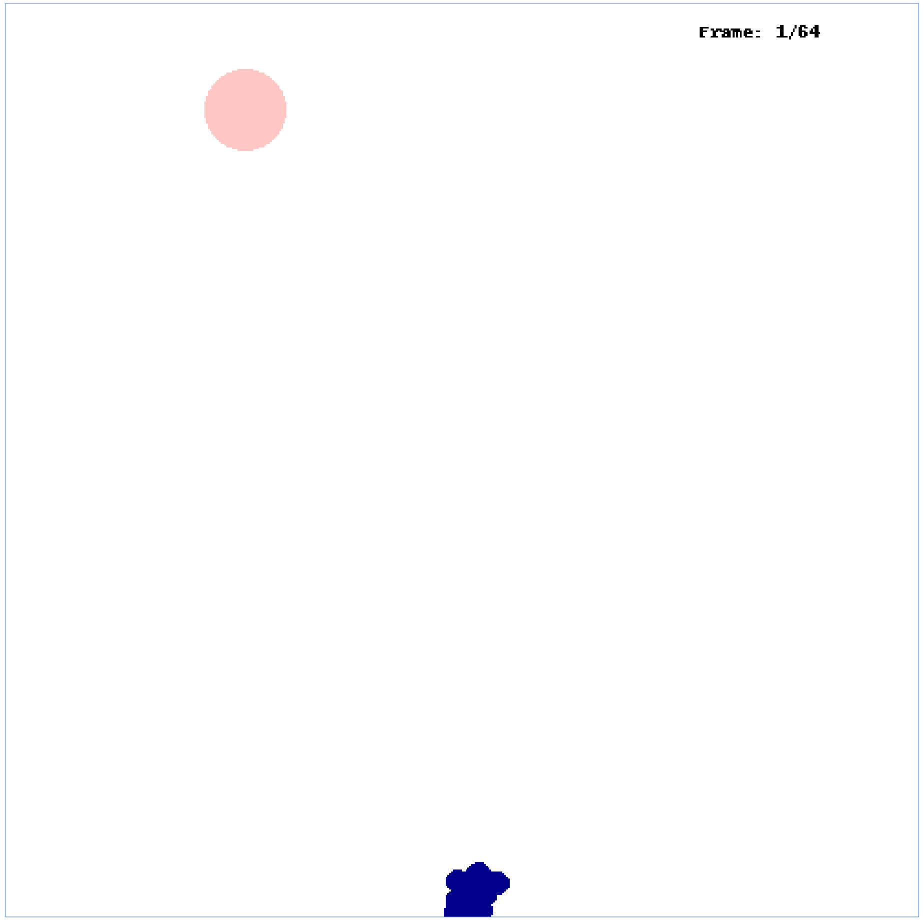
- Traditional models
  - Fixed input size (observation space)
- Varying #agents
  - Challenge →
  - Masking, based on distance
  - Pooling based on distance (k-Nearest Neighbors)

Simple Tag Game with Multi-Agent Actor Critic policy learning

- Normal approach
  - Single agent policy learning for each agent
- The Multi-Agent Actor Critic
  - Decentralized actors (policies)
  - Centralized critic
  - → Info sharing
- Better for multi-agent environments

Move the team to the target area

- Conditions
  - Agents have limited vision
    - N closest agent and
the goal state if in vision range
  - No direct communication
  - Grid environment
  - Terminate: >80% at target
- Approach
  - MLP model and PPO policy learning
- Behavior:
  - Agents spreading out as much as they can see each other
  - After finding the goal, they oscillate to propagate the information

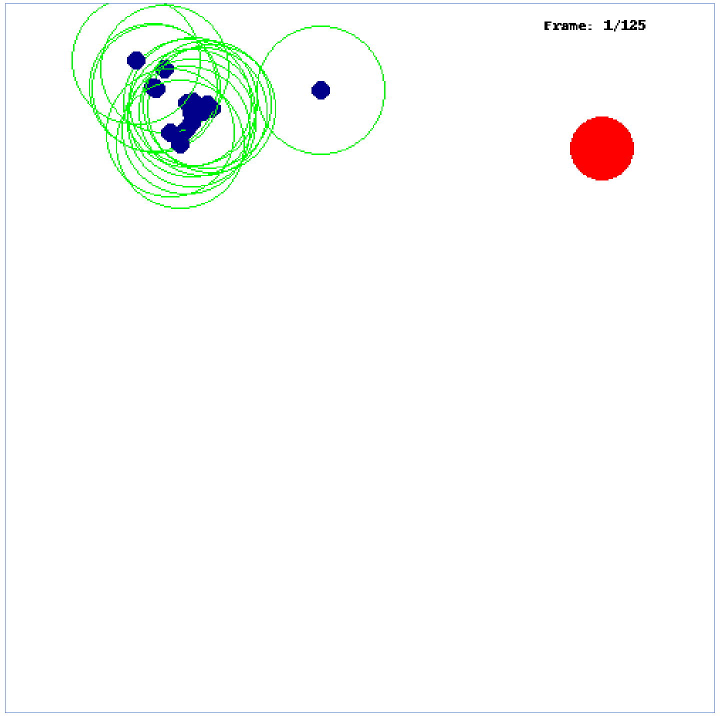


Move the team to the target area

2 distinct strategies

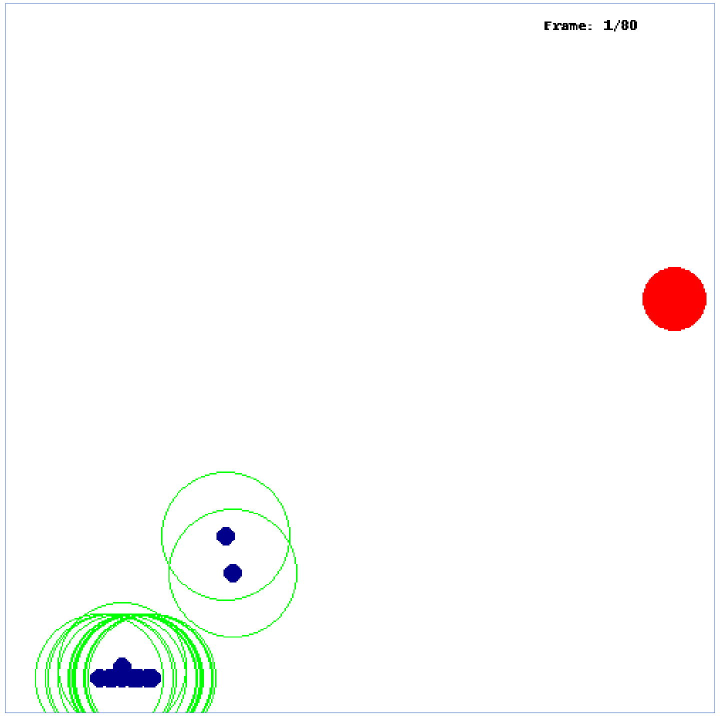
Fast Exploration

Spread wide, communicate later



Fast communication

Stick and explore together



Move to the Target
Multi-Agent Multi-Target Case

- Incremental and curriculum learning
  - Agents learn step by step from individual "walk" to move collectively

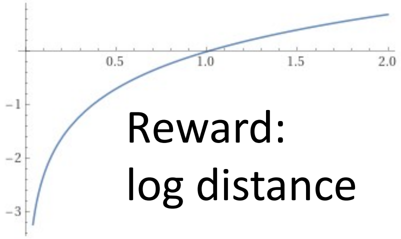
Punishment for colliding with obstacles

Red: target, **Blue:** agent, **Black:** obstacle

Move to the Target
Multi-Agent Multi-Target Case
w/o incremental learning

Red: target, **Blue:** agent, **Black:** obstacle

- Non-linear (log) reward



Reward:
log distance

  - Maintains local focus
  - Helps avoiding collision

Before

After

- Simplified observation space
  - K nearest objects
  - Same trained model for arbitrary #agents, #targets

Traffic Signal Control
with local communication

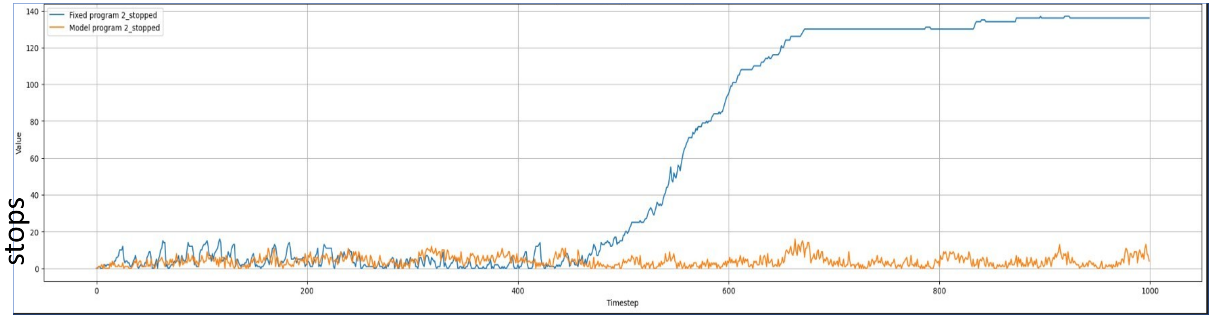
Classic approach

- Independent intersections, or
- Centralized control

Our approach

- Independent intersections (agents)
  - Communicating with nearest neighbors
- Agent classes
  - Based on #neighbors

Number of stops



Traffic increases by time

ELTE | IK
INFORMATIKAI KAR

Thank you!

László Gulyás

Associate Professor
Department of Artificial Intelligence
lgulyas@inf.elte.hu

ELTE | IK
DEPARTMENT OF ARTIFICIAL INTELLIGENCE