

Communication in Collaborative Multi-Agent Reinforcement Learning

Élie Michel

January 18, 2017

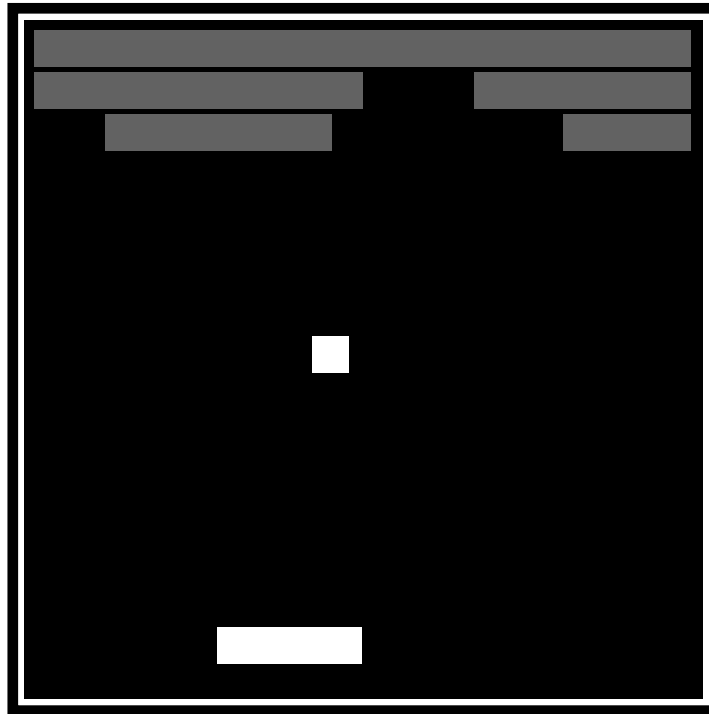
Outline

1. Non-stationarity and multiple agents
2. Challenges of Multi-Agent RL
3. Communication
4. Conclusion

1. Non-stationarity and multiple agents

1. Non-stationarity and multiple agents

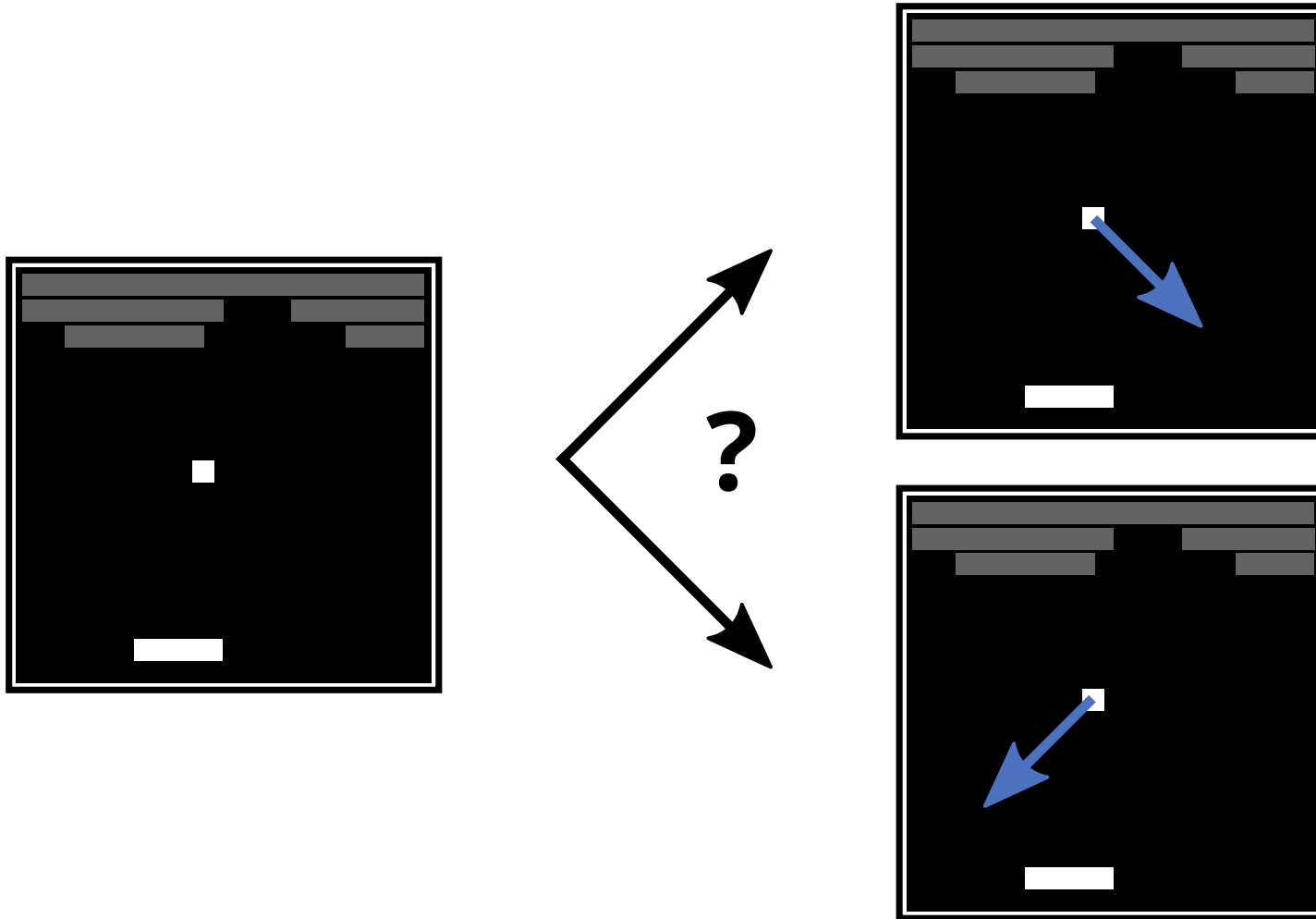
1.1. World dynamics



How to perceive the dynamics?

1. Non-stationarity and multiple agents

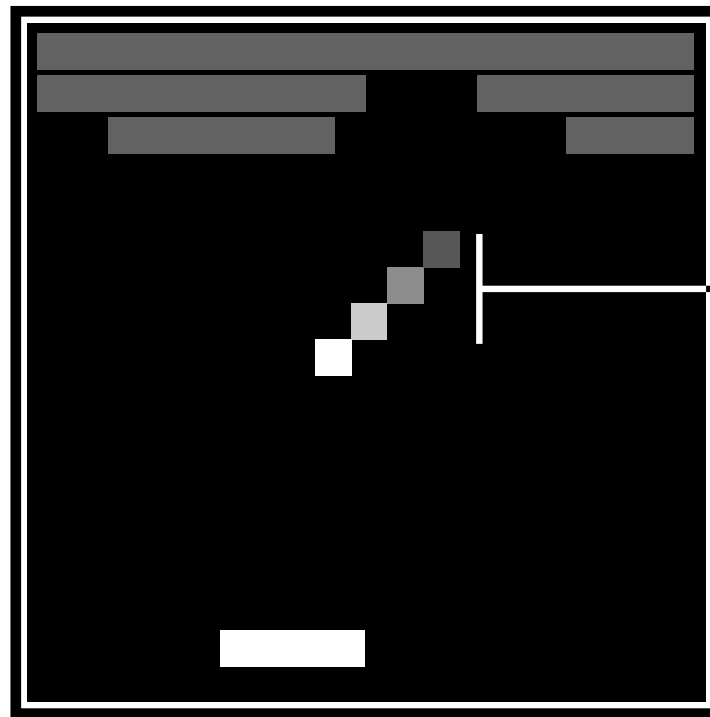
1.1. World dynamics



How to perceive the dynamics?

1. Non-stationarity and multiple agents

1.1. World dynamics



Short term
Memory

1. Non-stationarity and multiple agents

1.1. World dynamics

$$\mathcal{S} = \square_t$$

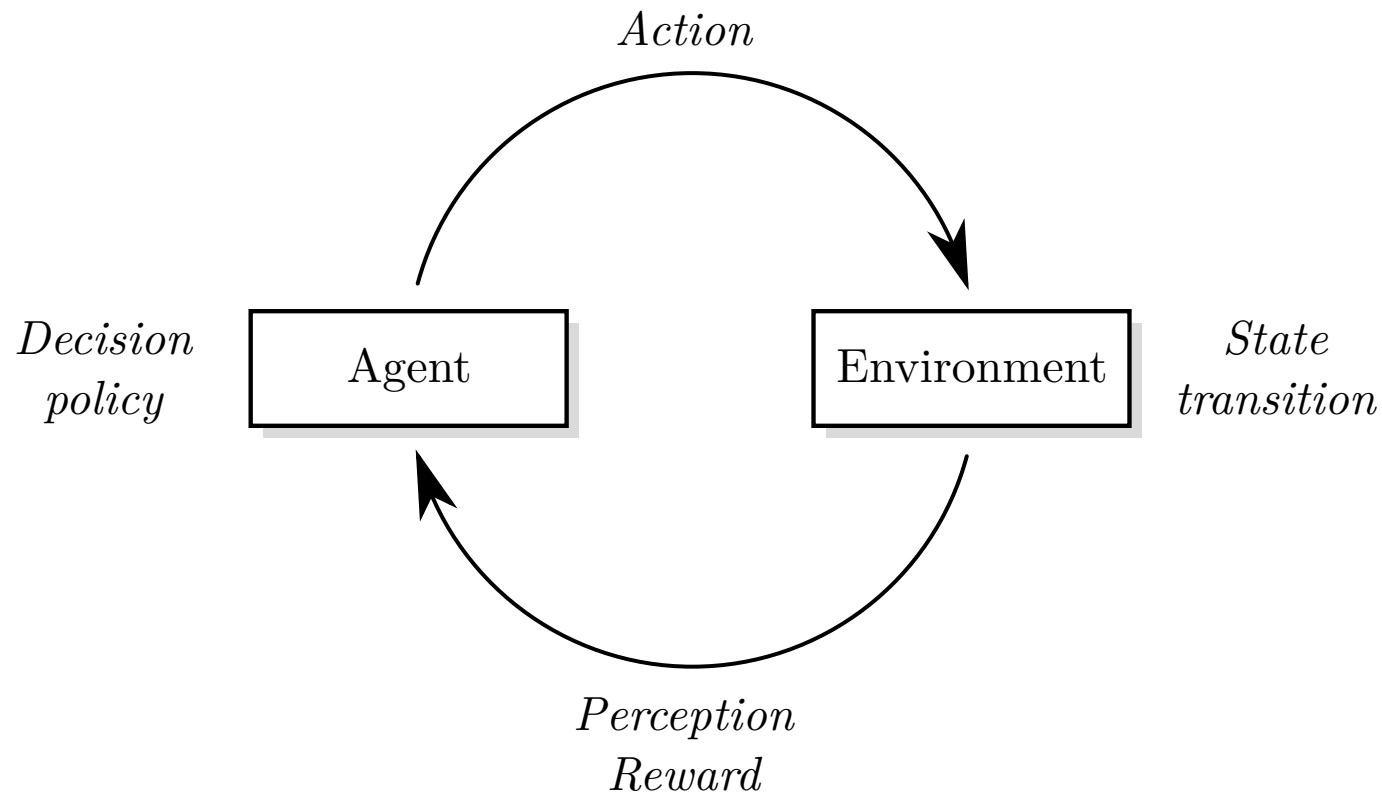
↓

$$\mathcal{S} = \square_t \times \square_{t-1} \times \square_{t-2} \times \square_{t-3}$$

Augmented state space

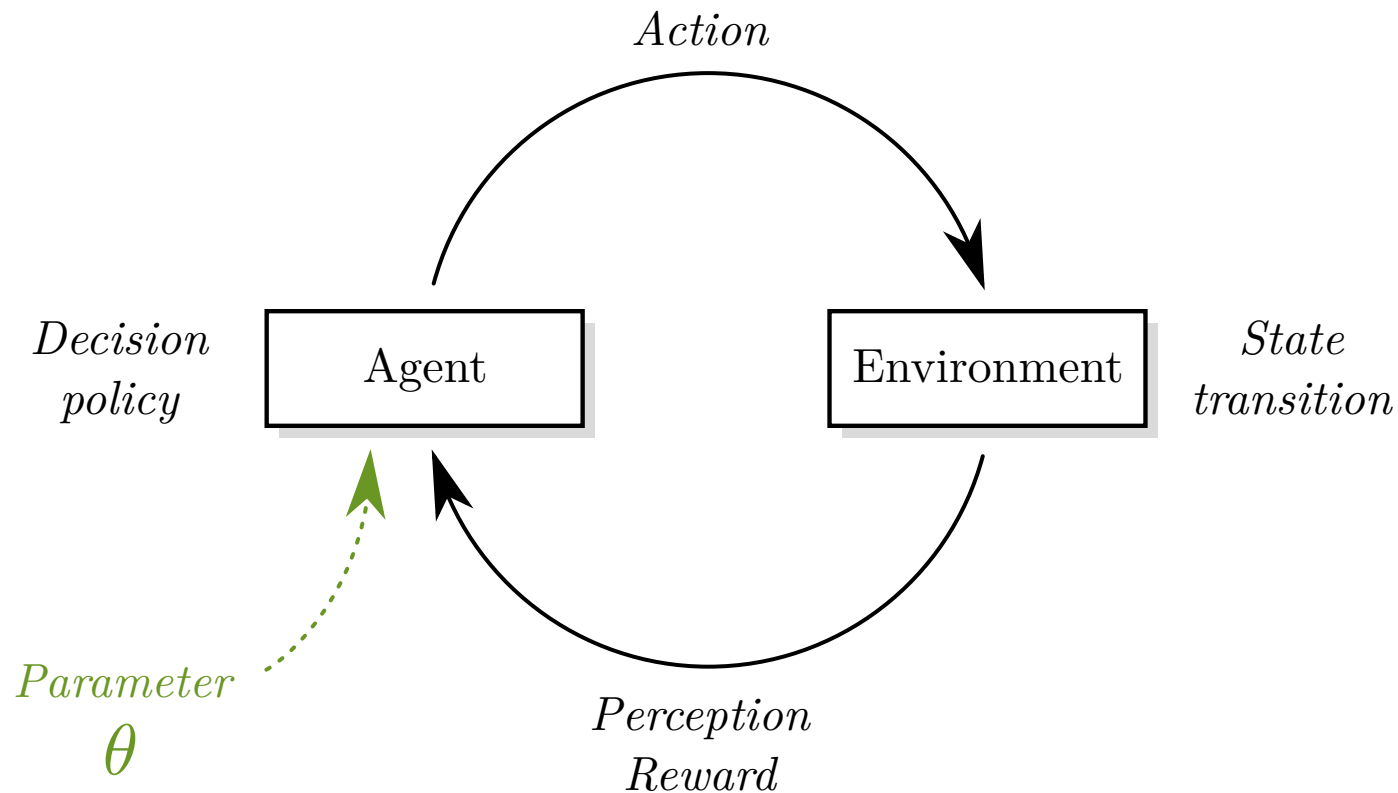
1. Non-stationarity and multiple agents

1.2. Non-stationary policy



1. Non-stationarity and multiple agents

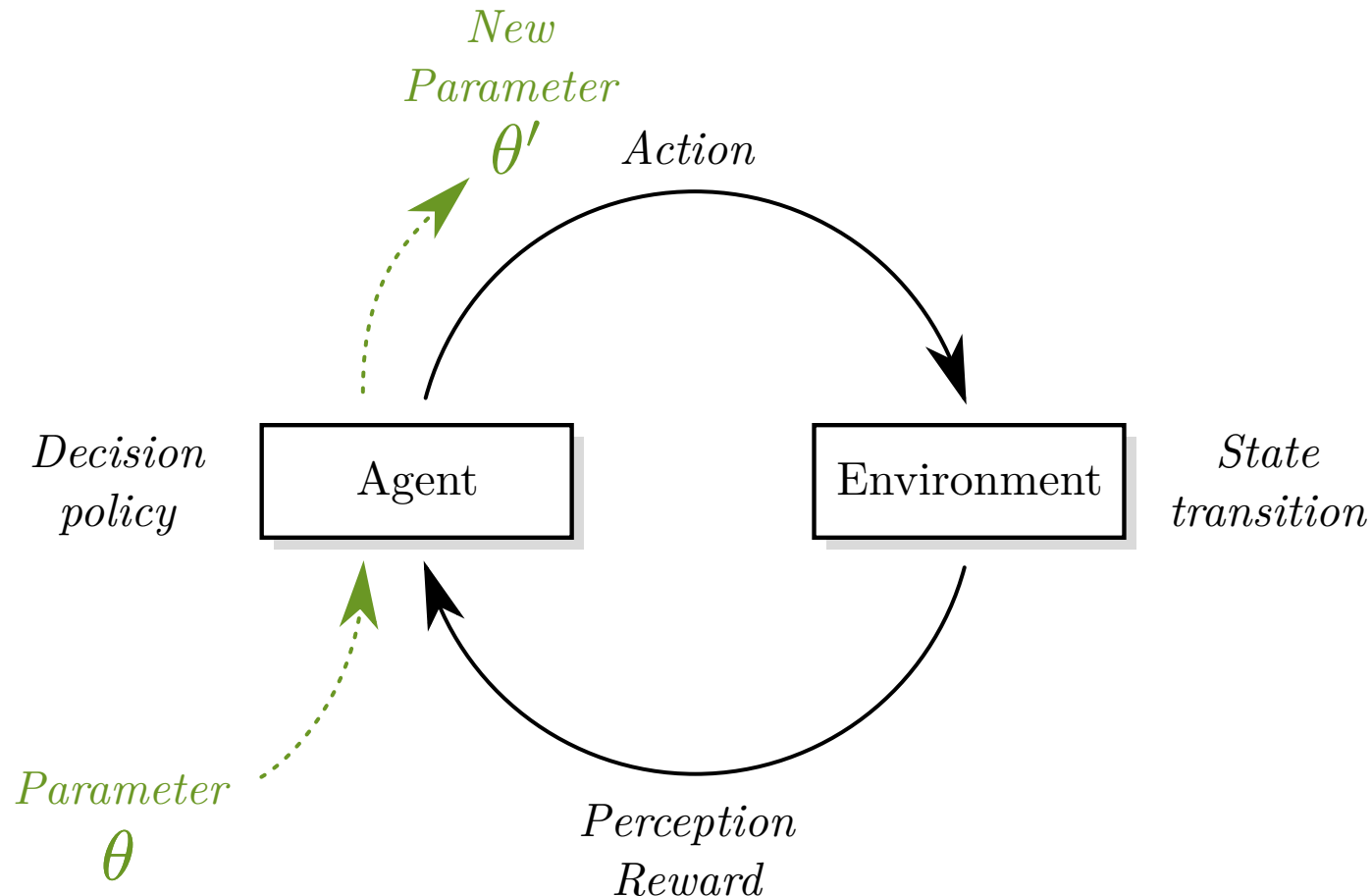
1.2. Non-stationary policy



Parametric policy

1. Non-stationarity and multiple agents

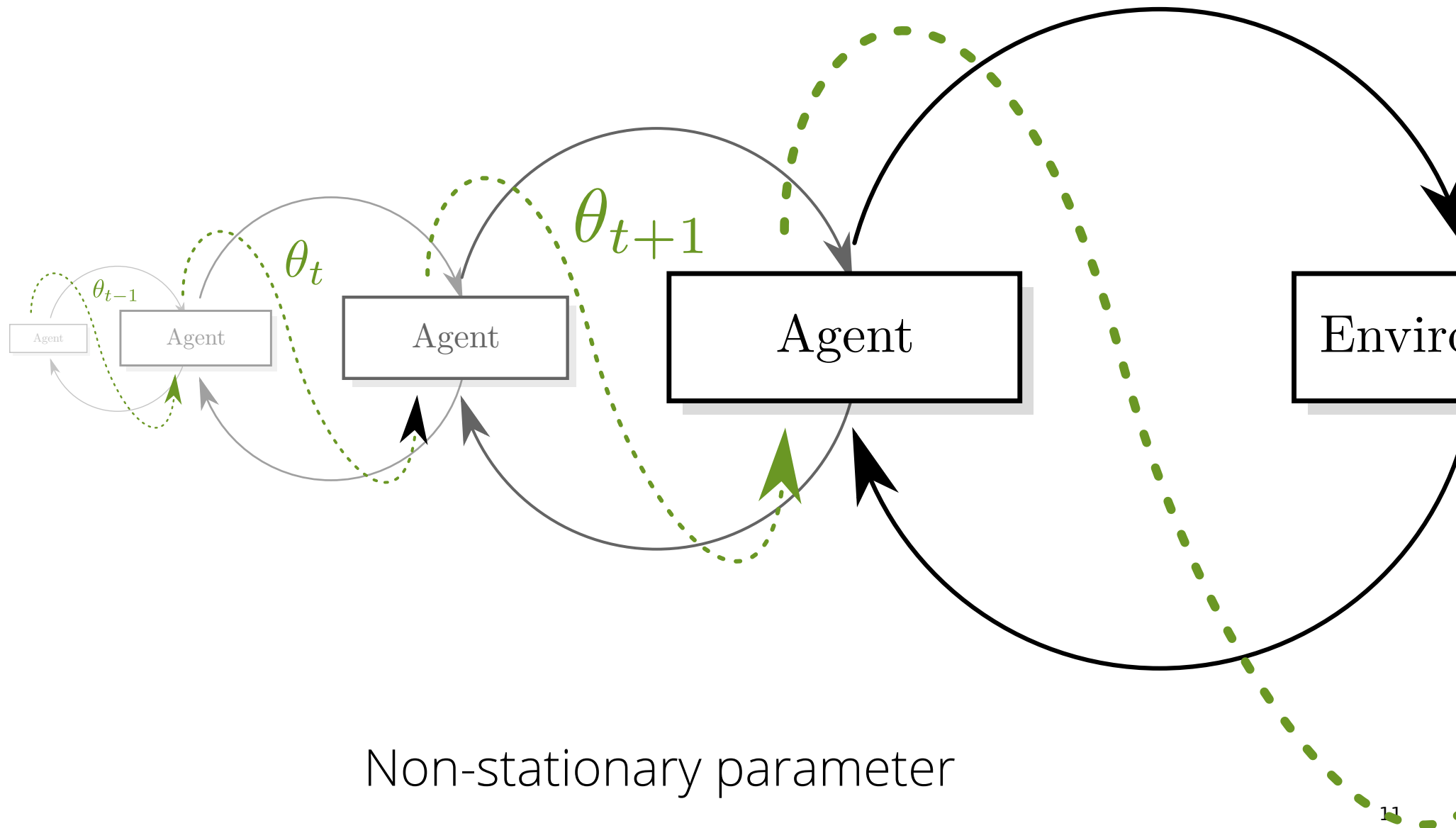
1.2. Non-stationary policy



Non-stationary parameter

1. Non-stationarity and multiple agents

1.2. Non-stationary policy



1. Non-stationarity and multiple agents

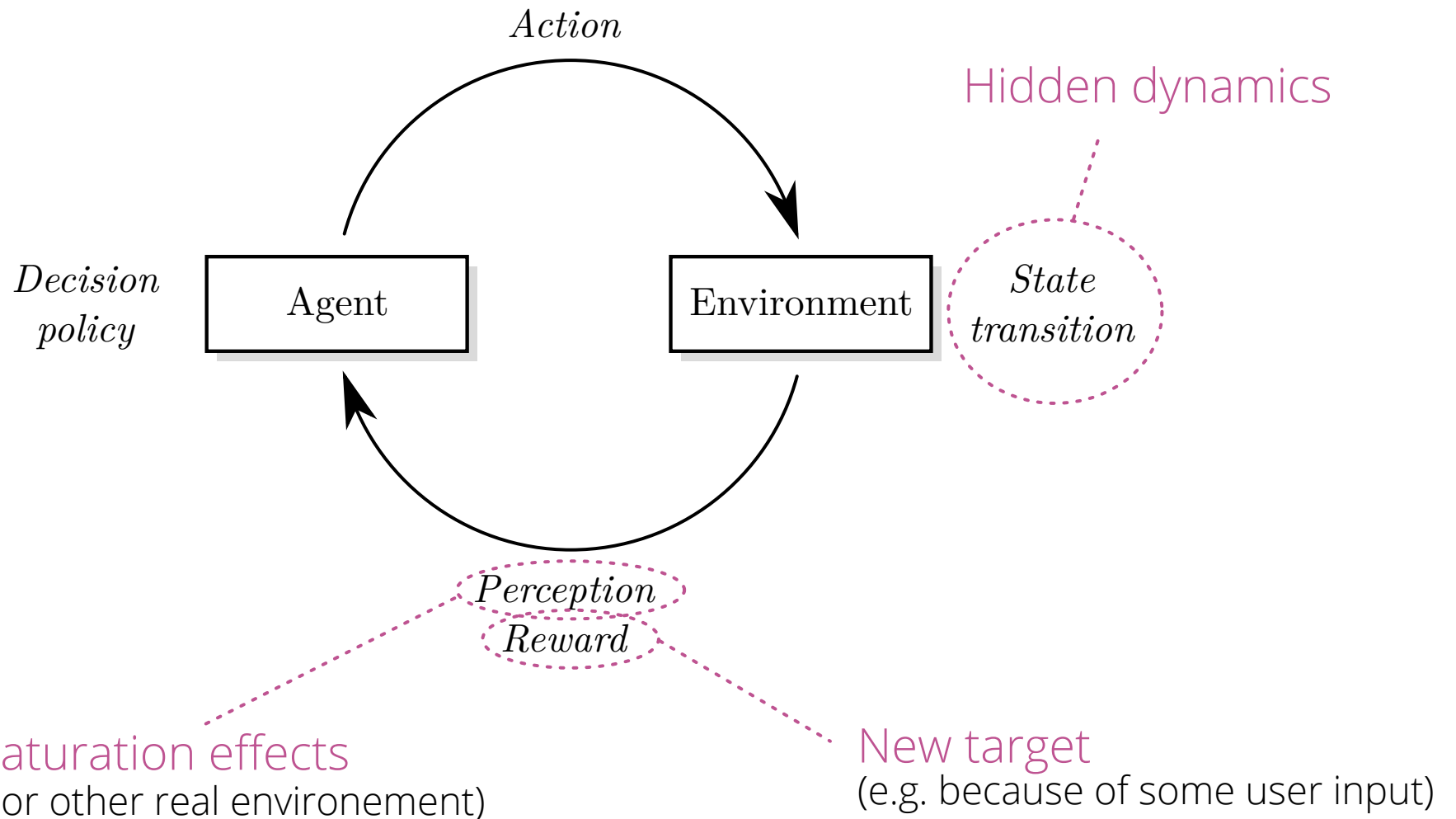
1.2. Non-stationary policy

Parameter θ may encode for:

- ▶ Short term memory e.g. $\theta \in \blacksquare \times \blacksquare \times \blacksquare$
- ▶ Q-Learning's Q matrix
- ▶ Learned Arbitrary Memory Representation
as Recurrent Neural Networks do
- ▶ ...

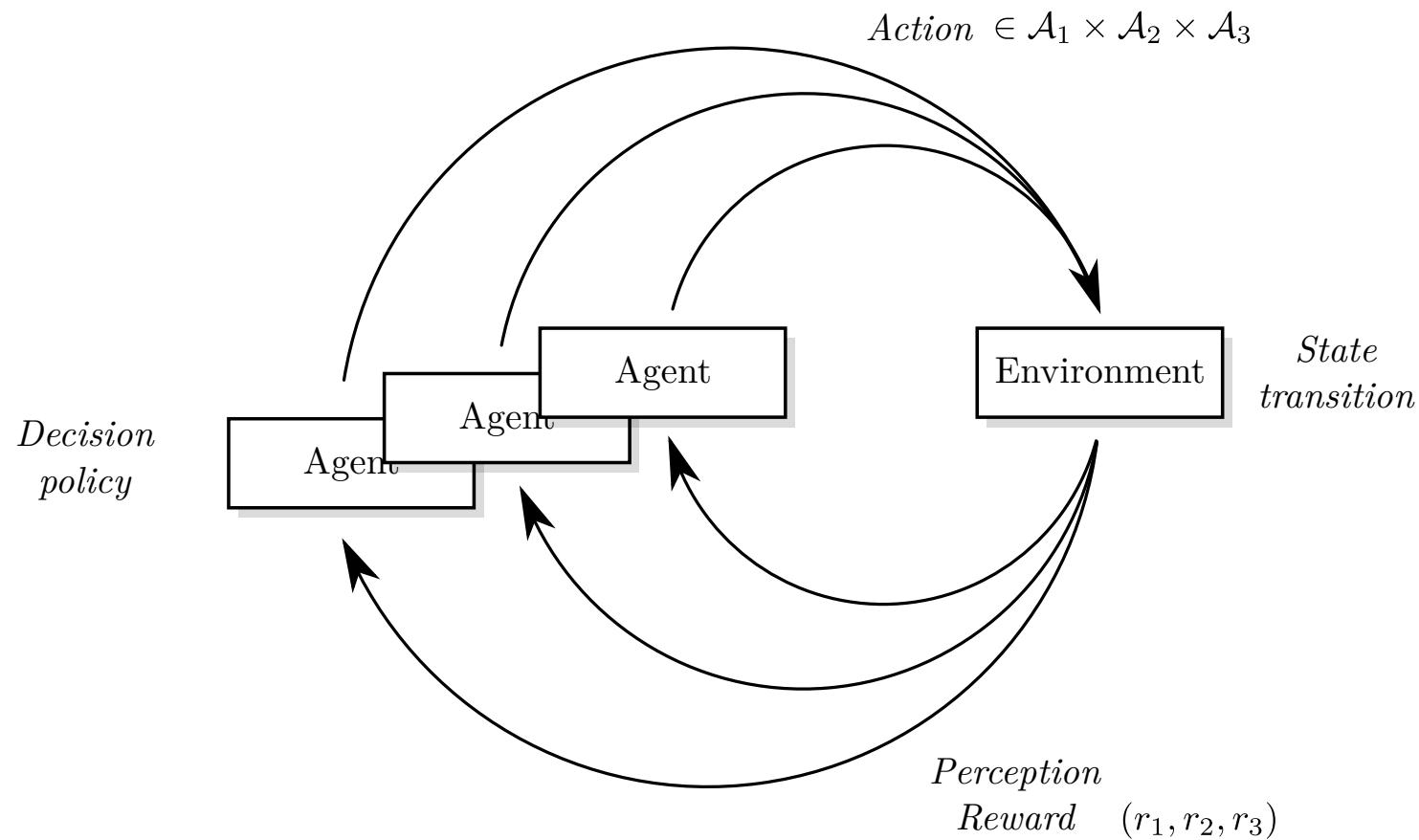
1. Non-stationarity and multiple agents

1.3. Other non-stationarities



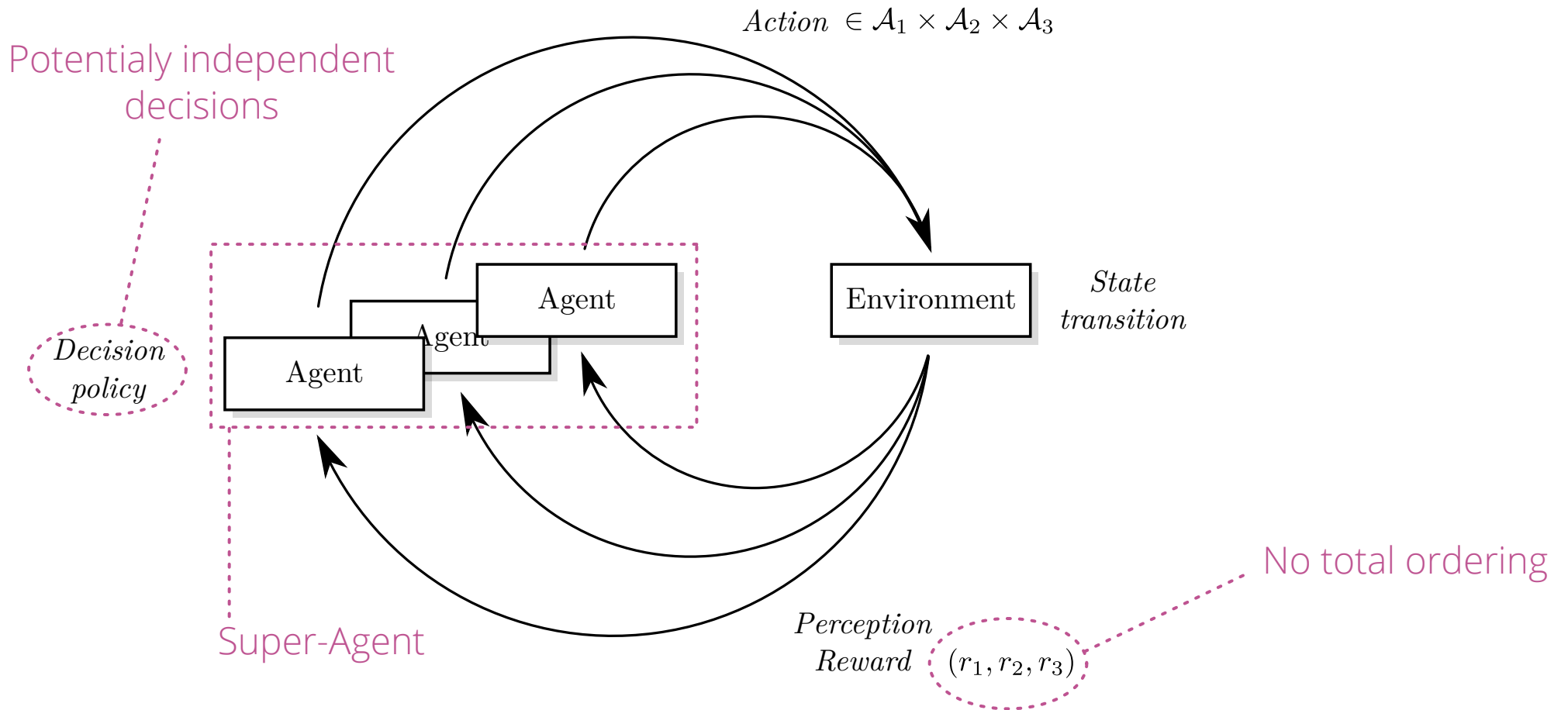
1. Non-stationarity and multiple agents

1.4. Multiple agents



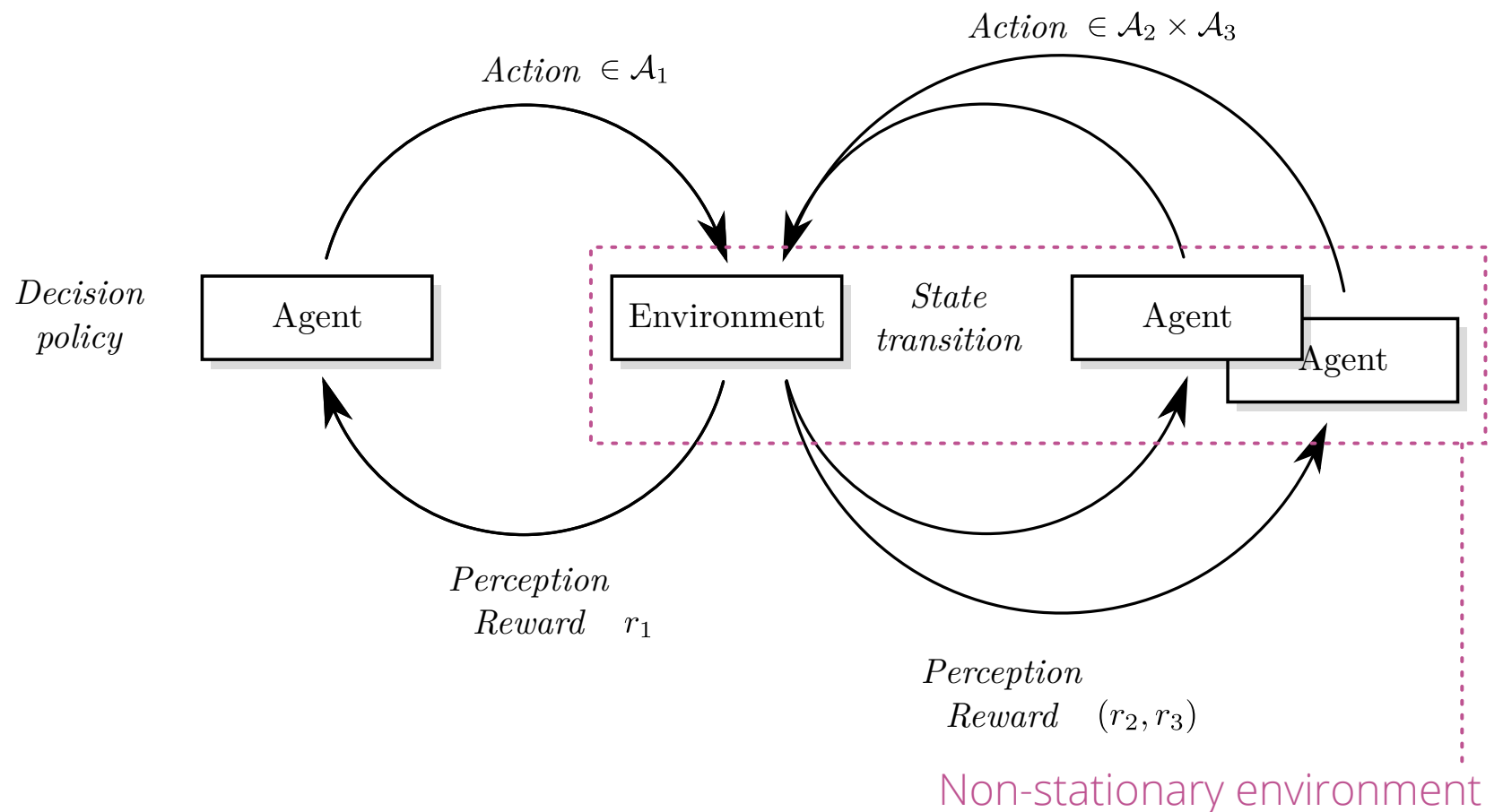
1. Non-stationarity and multiple agents

1.4. Multiple agents



1. Non-stationarity and multiple agents

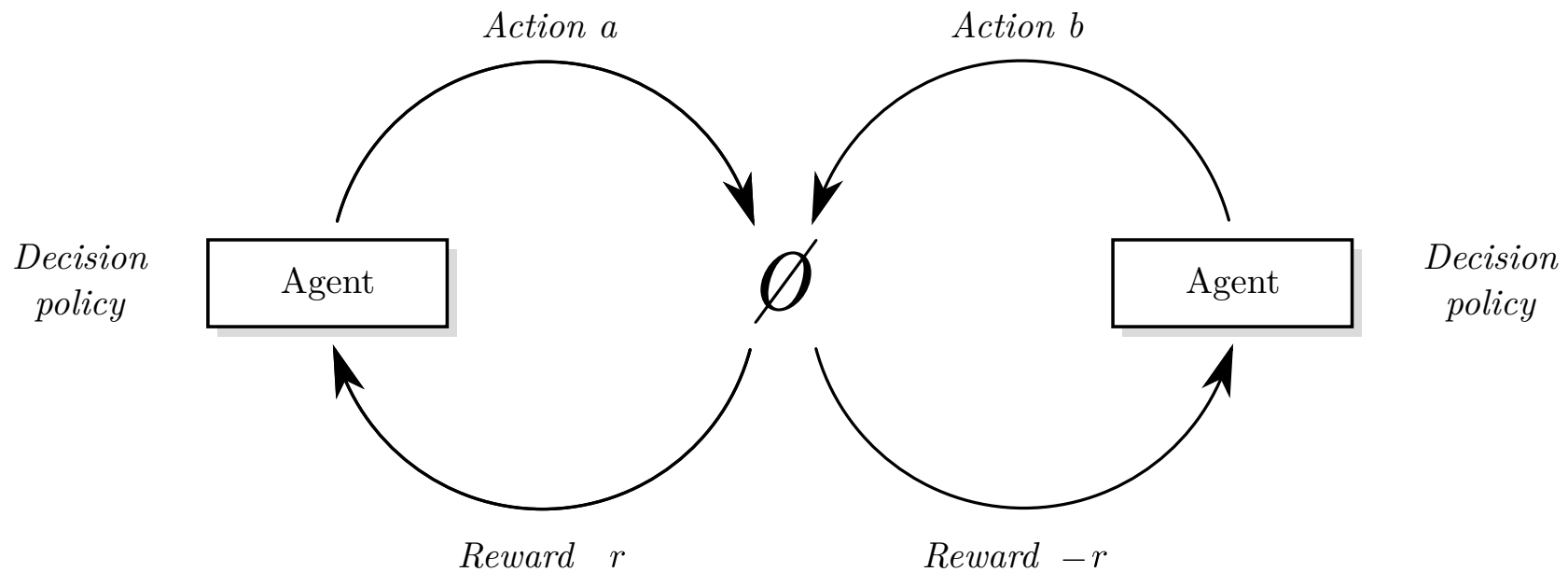
1.4. Multiple agents



From the point of view of an agent

1. Non-stationarity and multiple agents

1.4. Multiple agents

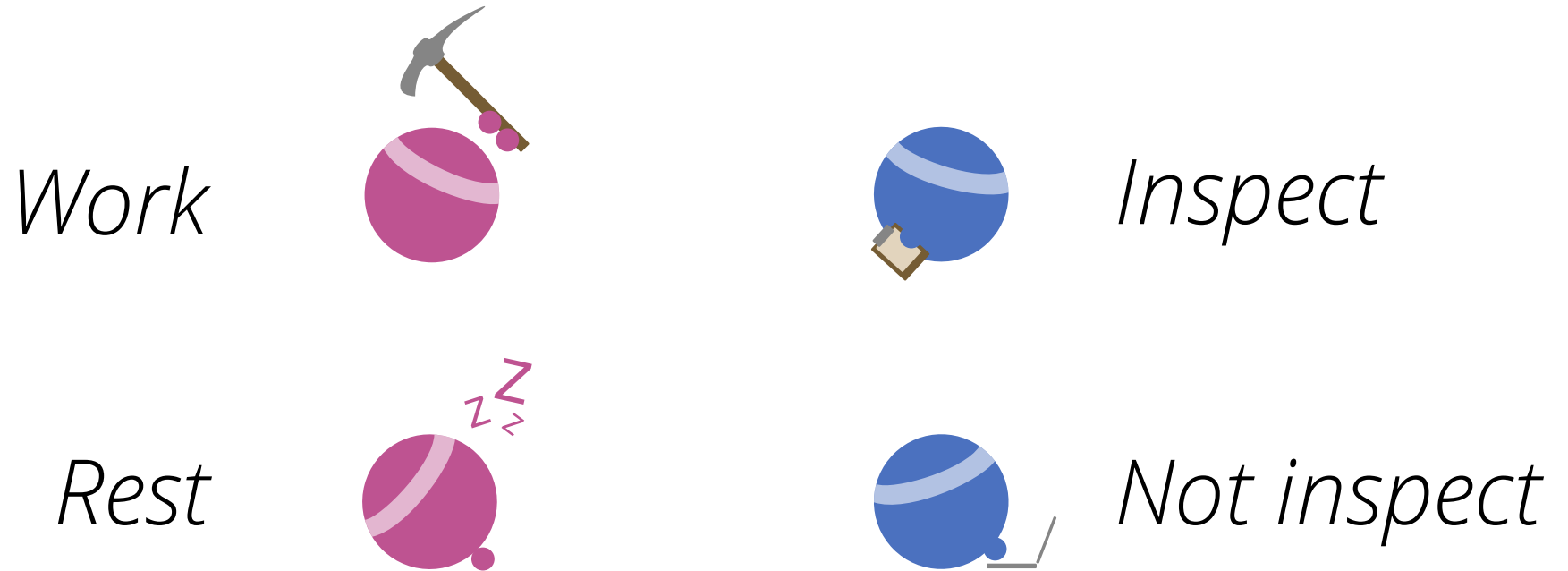




example:
Zero-sum stateless game

2. Challenges of Multi-Agent RL

2. Challenges of Multi-Agent RL

2.1. Need for stochasticity



 
Worker-Inspector

2. Challenges of Multi-Agent RL

2.2. Global and local goal

Reward at time t : (r_1, r_2, \dots, r_n)
one per agent



No ordering over rewards!

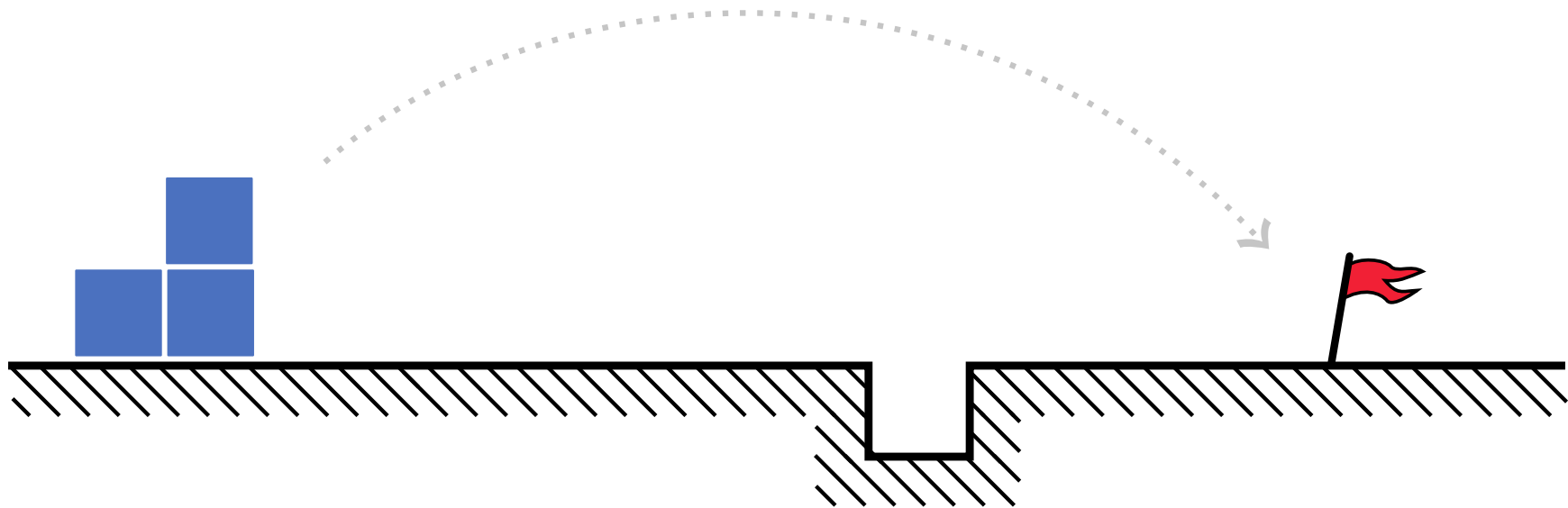
Particular case: $r_1 = r_2 = \dots = r_n$



Collaborative Agents

2. Challenges of Multi-Agent RL

2.2. Global and local goal

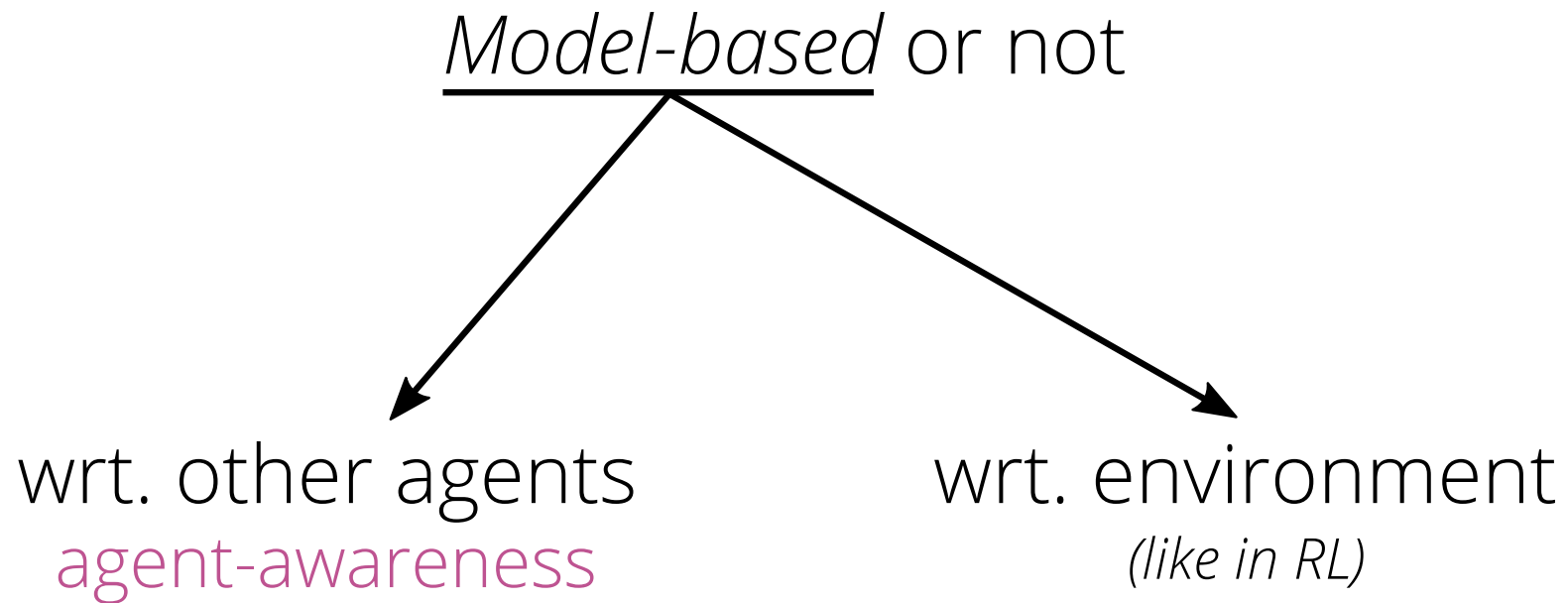


Even in collaborative setup, an agent may locally lower its reward for others to win

Learn to sacrifice

2. Challenges of Multi-Agent RL

2.3. Agent-awareness



2. Challenges of Multi-Agent RL

2.3. Agent-awareness

Different levels of agent-awareness:

- ▶ Other agent's existence

- ▶ Static modeling

- ▶ Dynamic modeling

e.g. Opponent modeling

need for

adaptation/stability trade-off

2. Challenges of Multi-Agent RL

2.3. Agent-awareness

agent-awareness

is a fundamental requirement for

communication

3. Communication

3. Communication

3.1. Use case



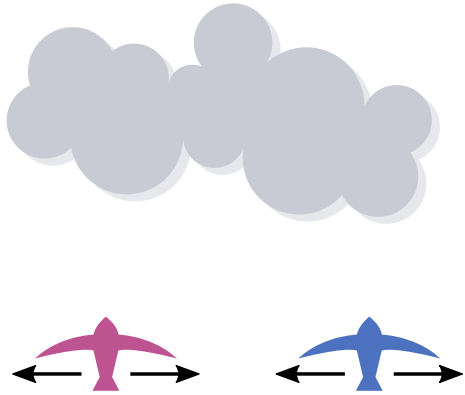
```
Target_  
> avoid cloud  
> keep formation
```



need for
Coordination

3. Communication

3.1. Use case



Different modes of communication:

- ▶ Passive

 looks at 

requires full agent model

- ▶ Convention-based

 knows what  will do

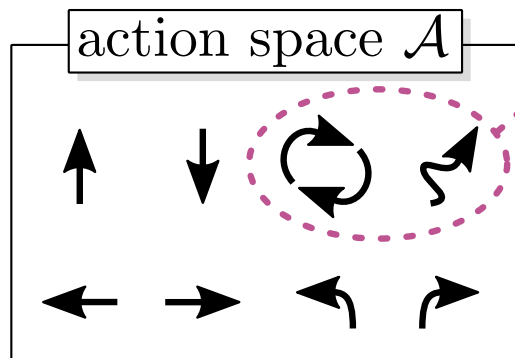
- ▶ Active

 and  decide together

requires priority ordering

3. Communication

3.2. Symbol grounding



Leverage on "rewardless" actions to **mean** sth to other agents

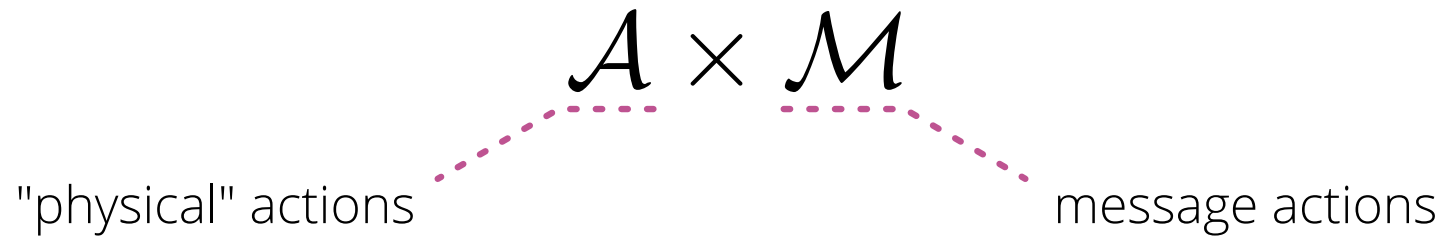
What meaning?

Symbol Grounding Problem

3. Communication

3.3. Message space

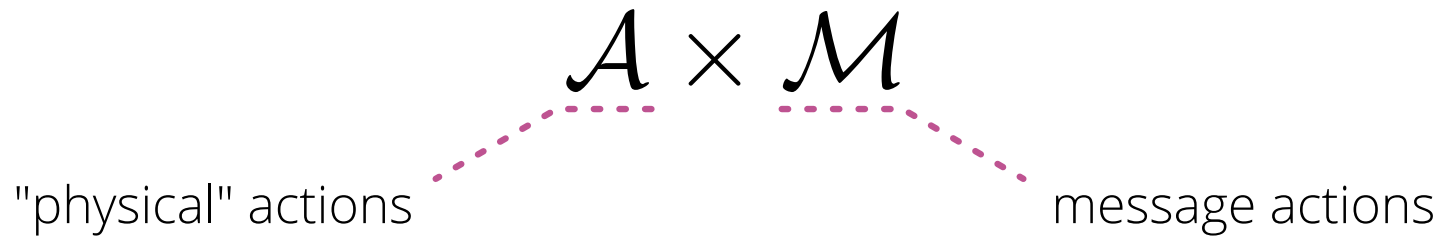
Effective action space:



3. Communication

3.3. Message space

Effective action space:

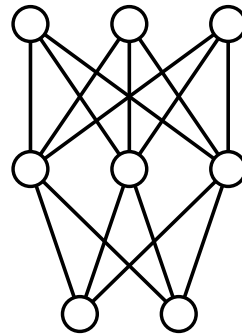


May be a continuous vector
Deep Learning

3. Communication

3.4. Deep communication

observation

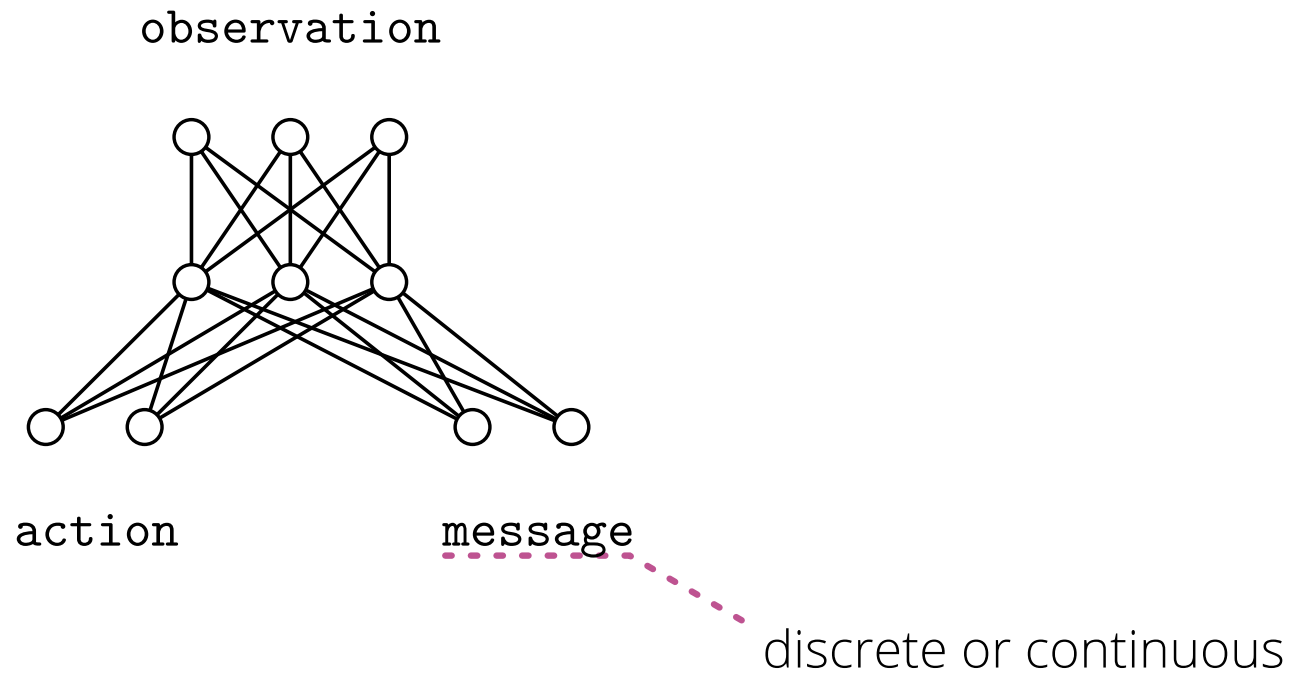


action

Deep Agent

3. Communication

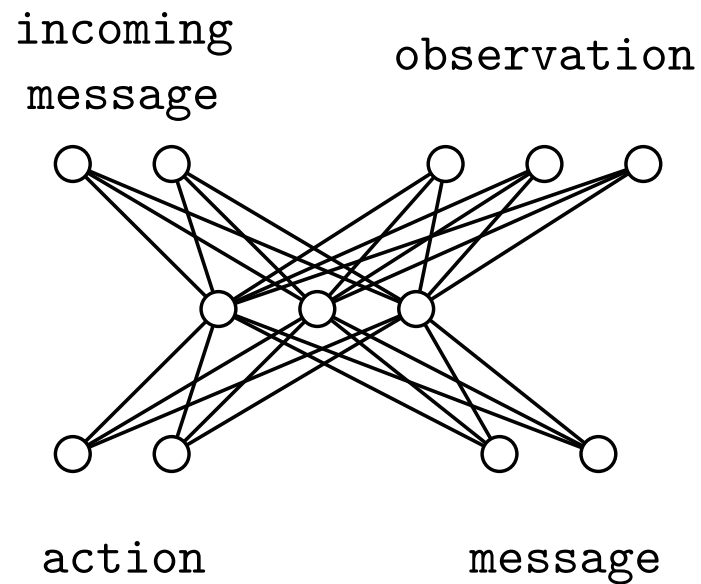
3.4. Deep communication



Deep Agent

3. Communication

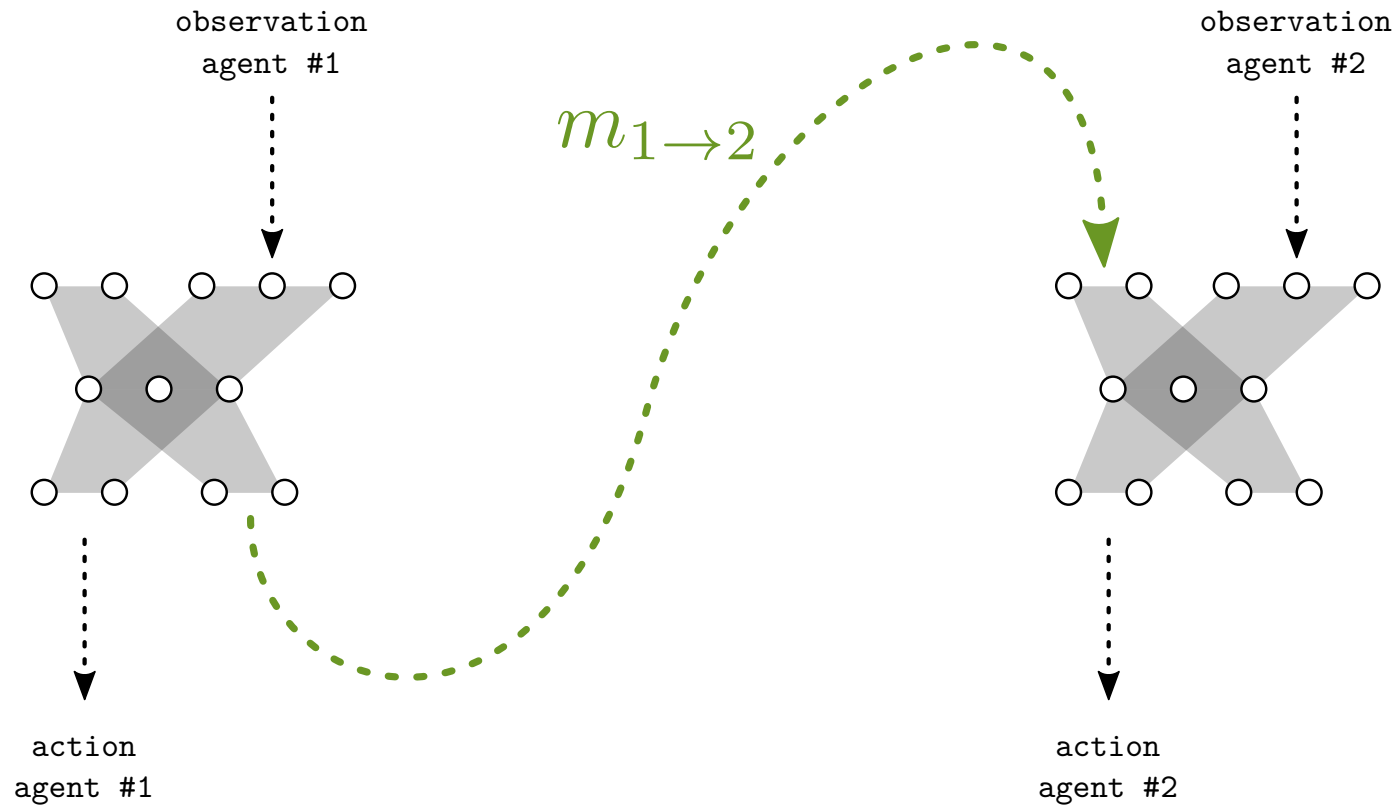
3.4. Deep communication



Deep Agent

3. Communication

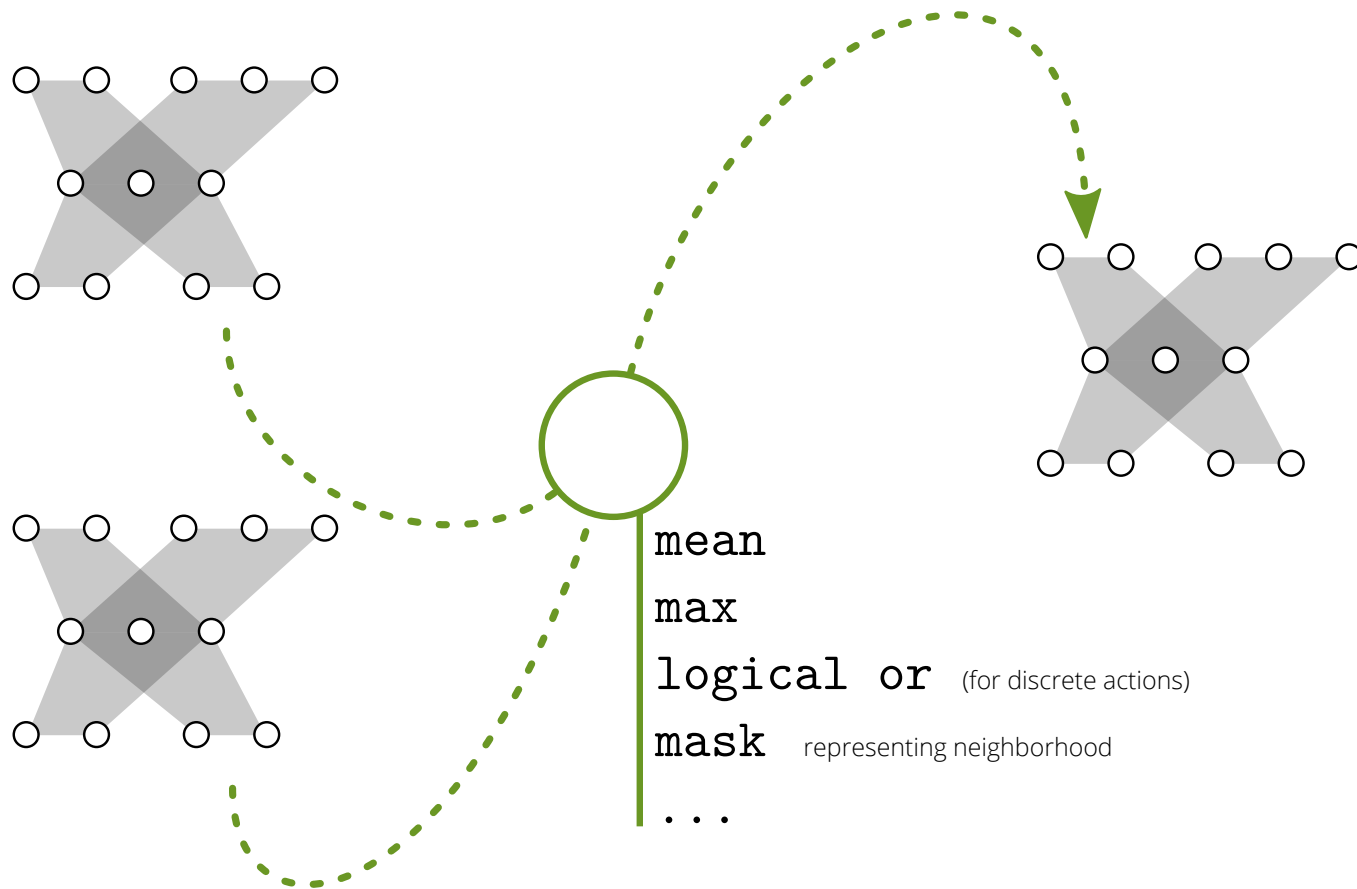
3.4. Deep communication



Deep Agent

3. Communication

3.4. Deep communication

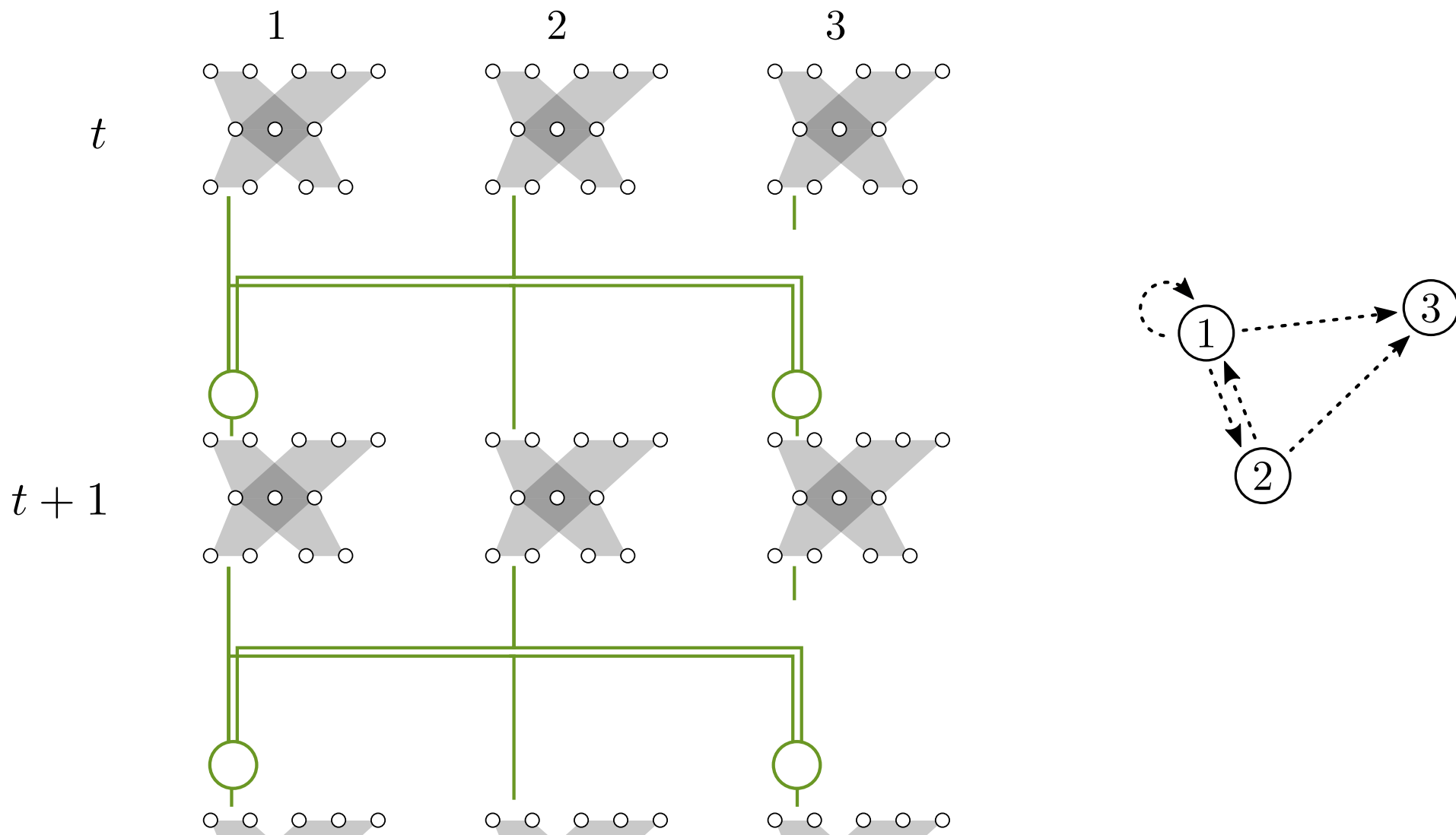


Combination

3. Communication

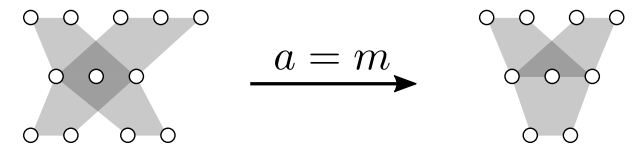
3.4. Deep communication

Communication Tree \rightarrow Communication Graph + Time

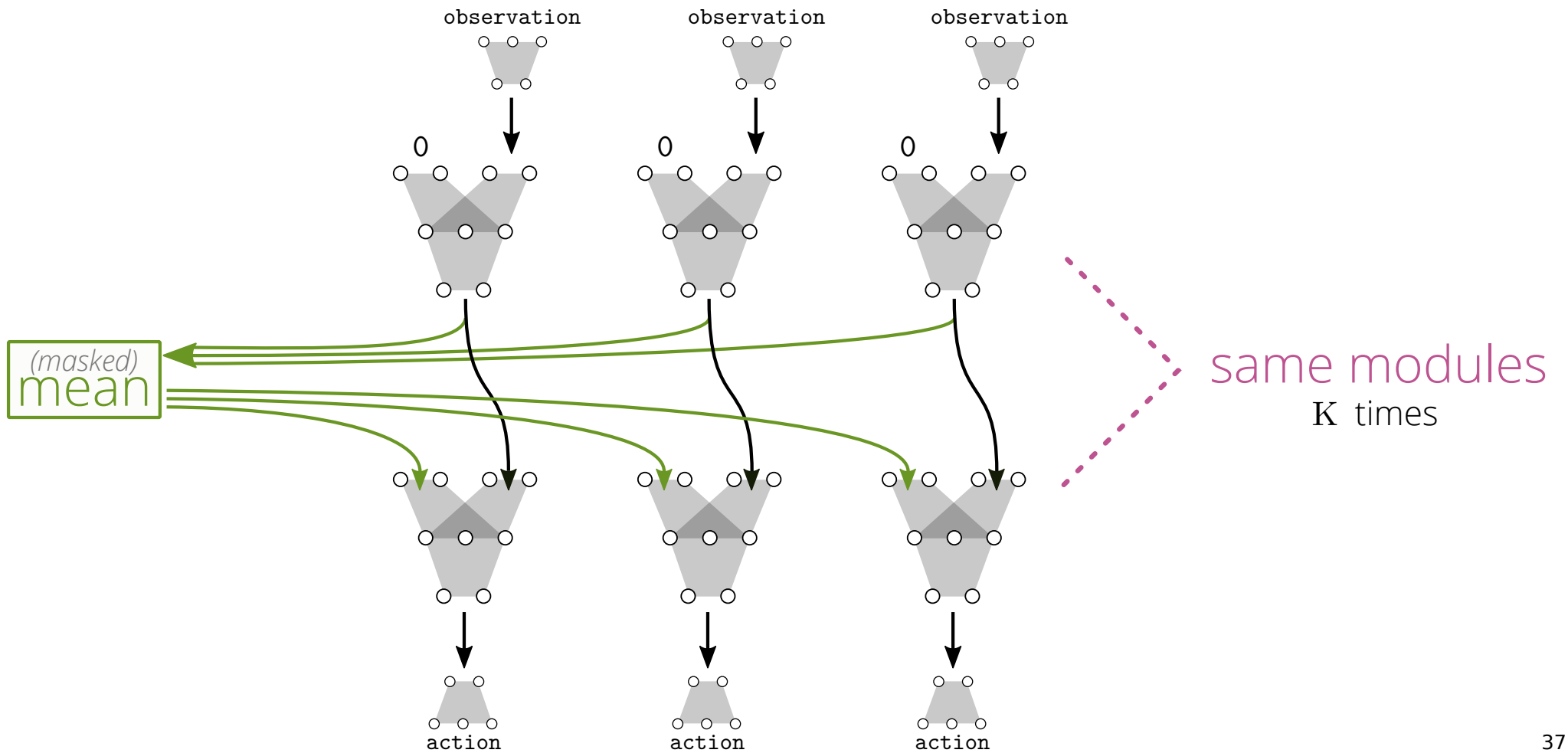


3. Communication

3.4. Deep communication

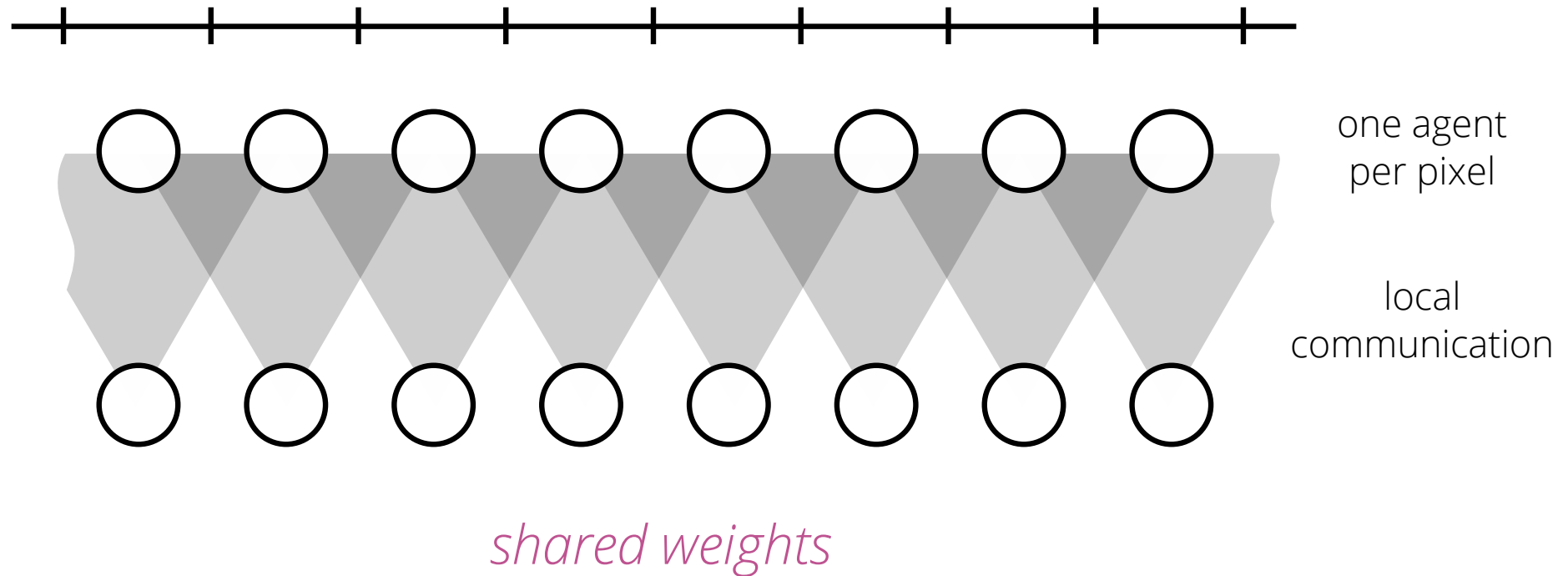


Example: [Sukhbaatar16]



3. Communication

3.4. Generalized communication



CNN seen as Multi-Agent Communication

4. Conclusion

Thanks
Questions?