SCALING MULTI-AGENT REINFORCEMENT LEARNING WITH SELECTIVE PARAMETER SHARING

Filippos Christianos, Georgios Papoudakis, Arrasy Rahman, Stefano Albrecht

In 38th International Conference on Machine Learning 2021

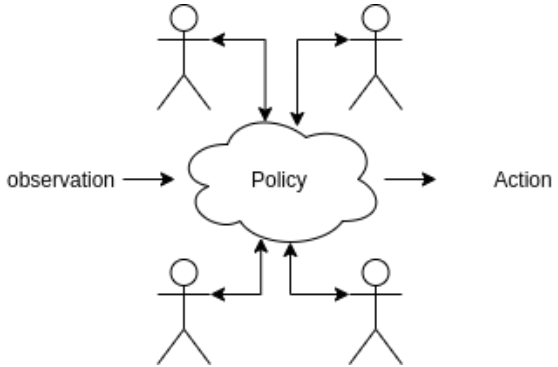
SCALING MARL TO MANY AGENTS

Scaling MARL to many agents?

Typical numbers in works like MADDPG range from 2 to 10 agents (e.g. centralised critics = large inputs scaling with the number of agents).

PARAMETER SHARING

Parameter sharing: agents share parameters in their policy or critic networks.



In the literature, parameter sharing is typically applied indiscriminately across all agents, which we call *naive.*

PARAMETER SHARING

Can *naive* parameter sharing work when applied across heterogeneous agents?

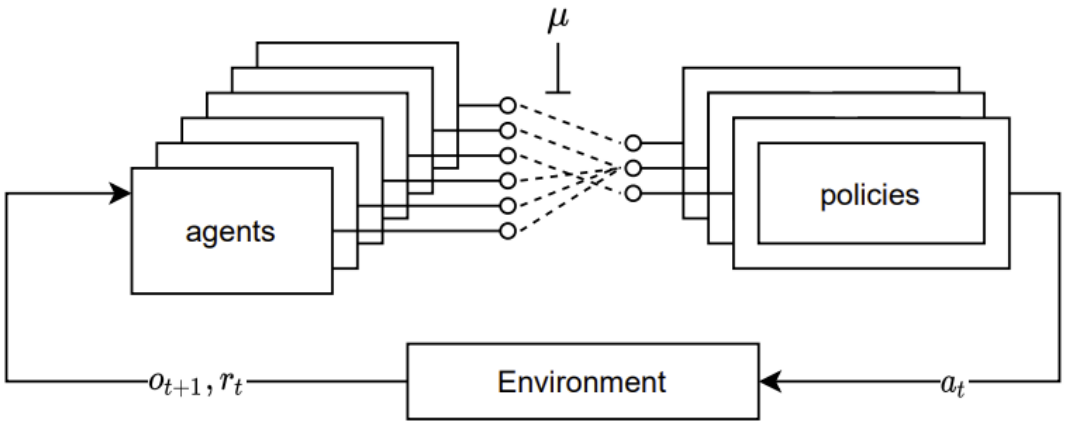
PARAMETER SHARING

Can *naive* parameter sharing work when applied across heterogeneous agents?

… not really.

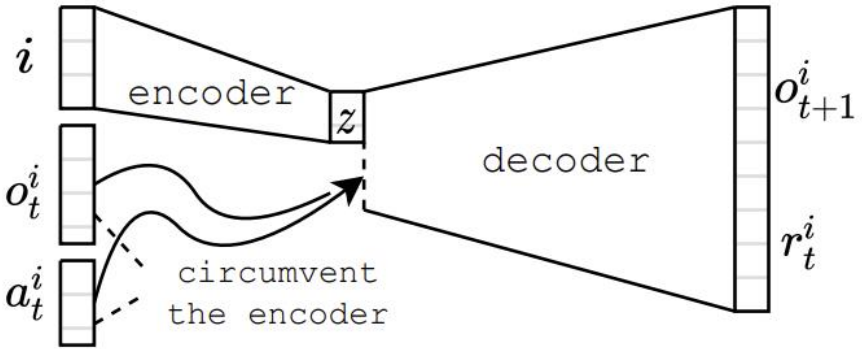
SELECTIVE PARAMETER SHARING

But we can apply it selectively.

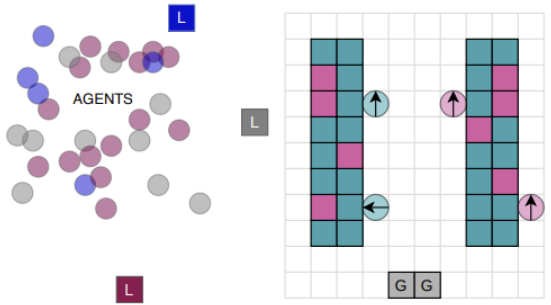


SELECTIVE PARAMETER SHARING

We identify agents with similar reward and observation transition functions and have them share parameters.

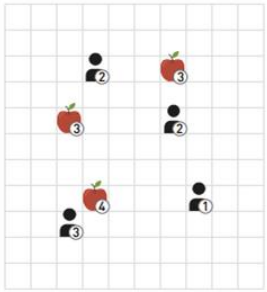


EXPERIMENTS: ENVIRONMENTS



(a) Blind-Particle Spread

(b) Coloured Multi-Robot Warehouse



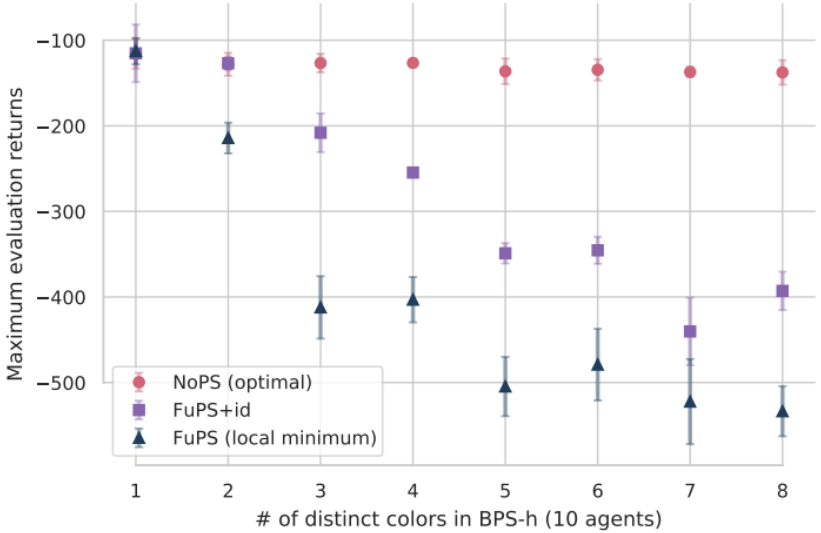
(c) Level-based Foraging

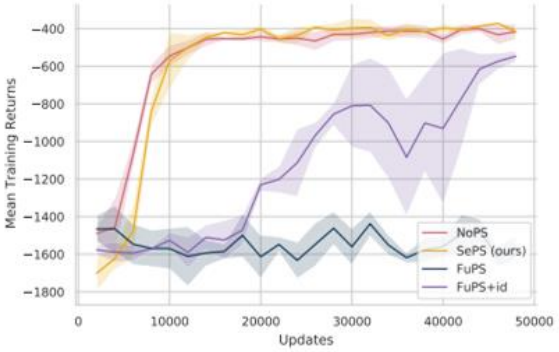
(d) SMAC

	# Agents	# Types	Type Distribution
BPS (1)	15	3	5–5–5
BPS (2)	30	3	10–10–10
BPS (3)	30	5	6–6–6–6–6
BPS (4)	30	5	2–2–2–15–9
BPS-h (1)	15	3†	5–5–5
BPS-h (2)	30	5†	6–6–6–6–6
BPS-h (3)	200	4†	50–50–50–50
C-RWARE (1)	4	2‡	2–2
C-RWARE (2)	8	2‡	4–4
C-RWARE (3)	16	2‡	8–8
LBF	12	3	4–4–4–4
MMM2	10	3§	7–2–1

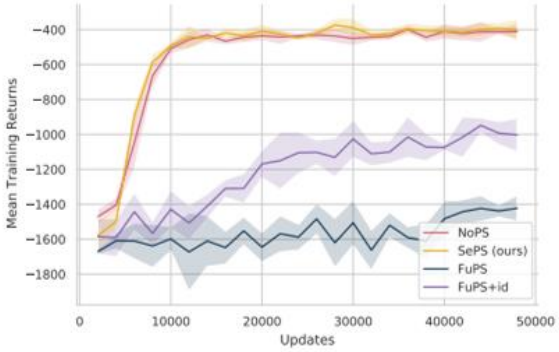
EXPERIMENTS: RESULTS

Can naive parameter sharing work?

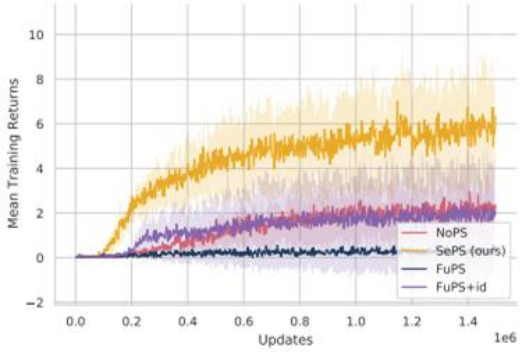




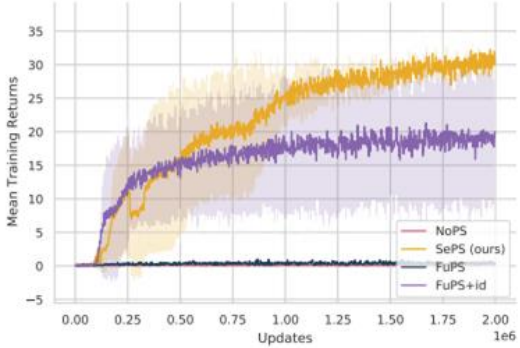
(a) BPS (3)



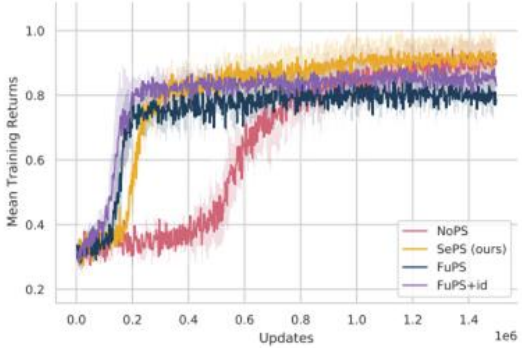
(b) BPS-h (2)



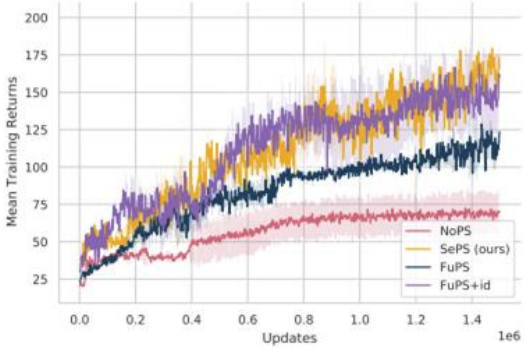
(c) C-RWARE (1)



(d) C-RWARE (3)

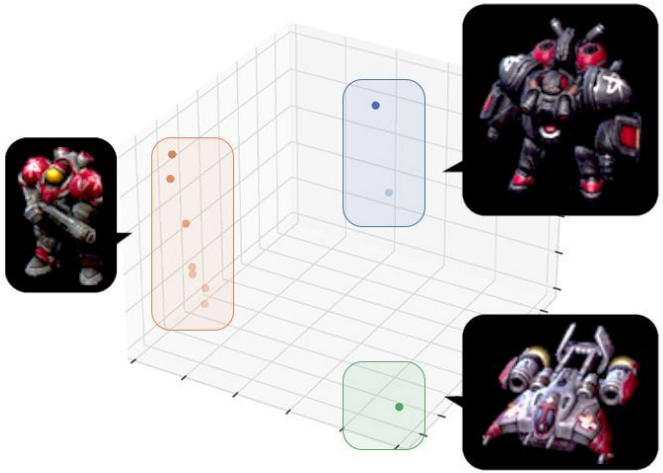


(e) LBF



(f) SMAC (MMM2)

VISUALIZING THE EMBEDDING SPACE

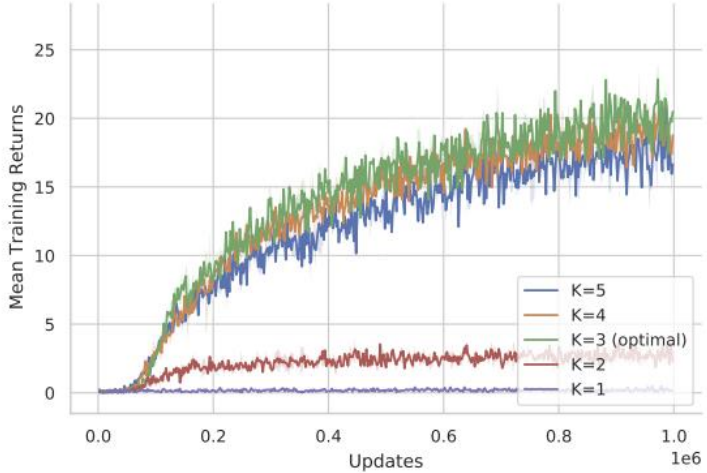


HOW DO WE DECIDE THE NUMBER OF CLUSTERS?

1. Manually
2. As a hyperparameter
3. Well-studied heuristics (e.g. DB-index)

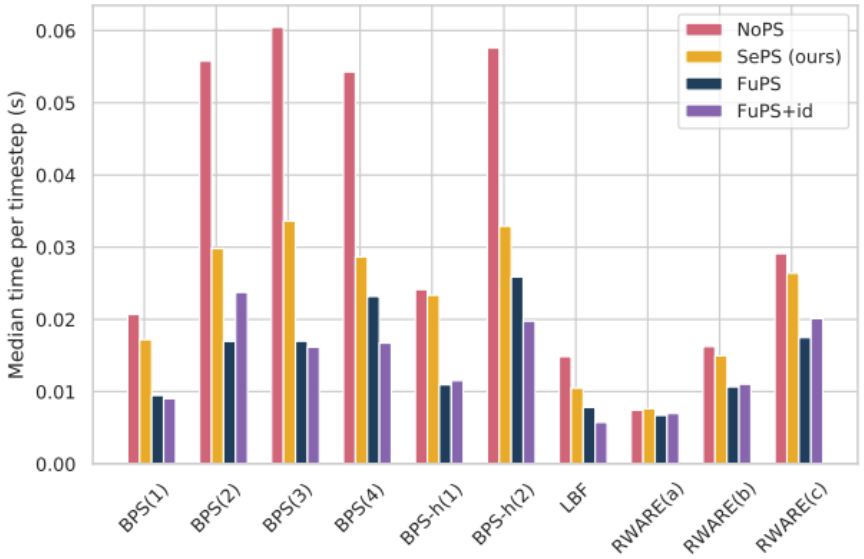
…

But does it matter?



COMPUTATIONAL BENEFITS

of parameters scale with the number of clusters (not the number of agents).



Scaling Multi-Agent Reinforcement Learning with Selective Parameter Sharing

https://arxiv.org/abs/2102.07475

Contributions:

1. We demonstrate the impact of parameter sharing methods (converged returns and training speed)
2. We propose a method to automatically identify agents that benefit from parameter sharing.

In 38th International Conference on Machine Learning 2021

THE UNIVERSITY of EDINBURGH informatics
Autonomous Agents Research Group