

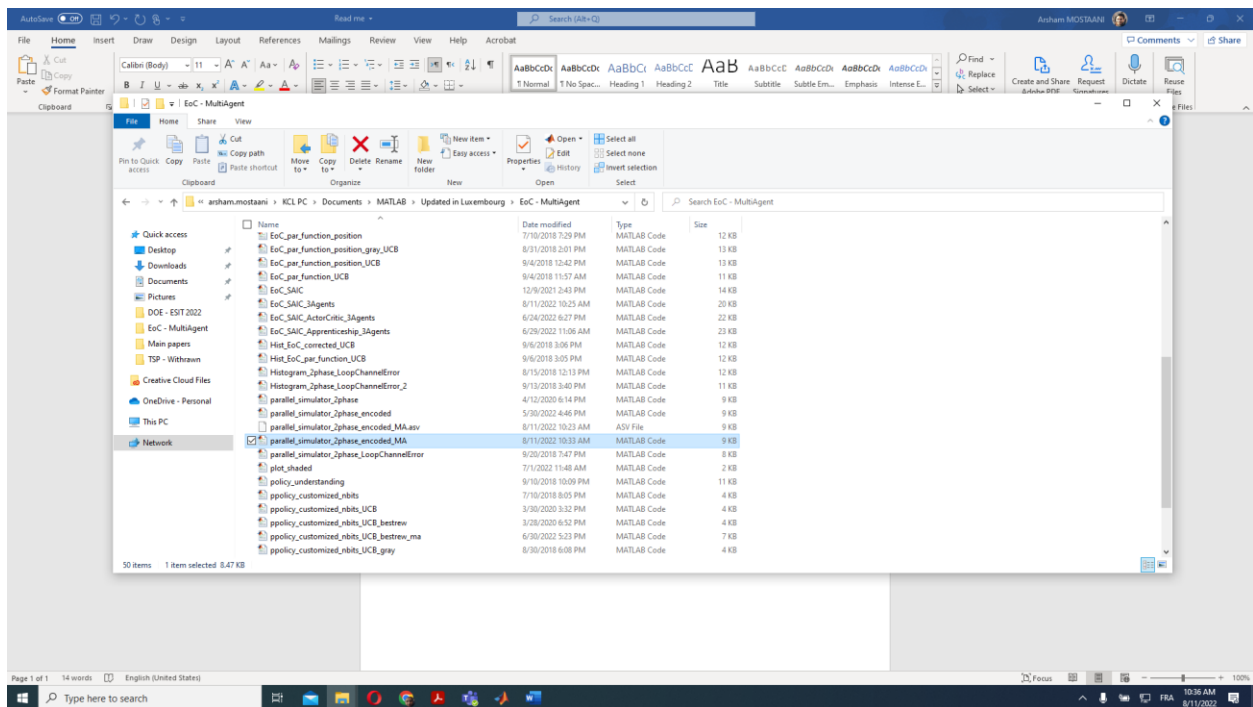
How to adjust the setting for multiagent simulations

By Arsham Mostaani – Affiliated with SnT - Supervised by Bjorn Ottersten – Date : 11/08/2022

ESAIC – SAIC

Finding the right functions and scripts

- 1- The parameter setup is “Parallel_simulator_2phase_encoded_MA ”

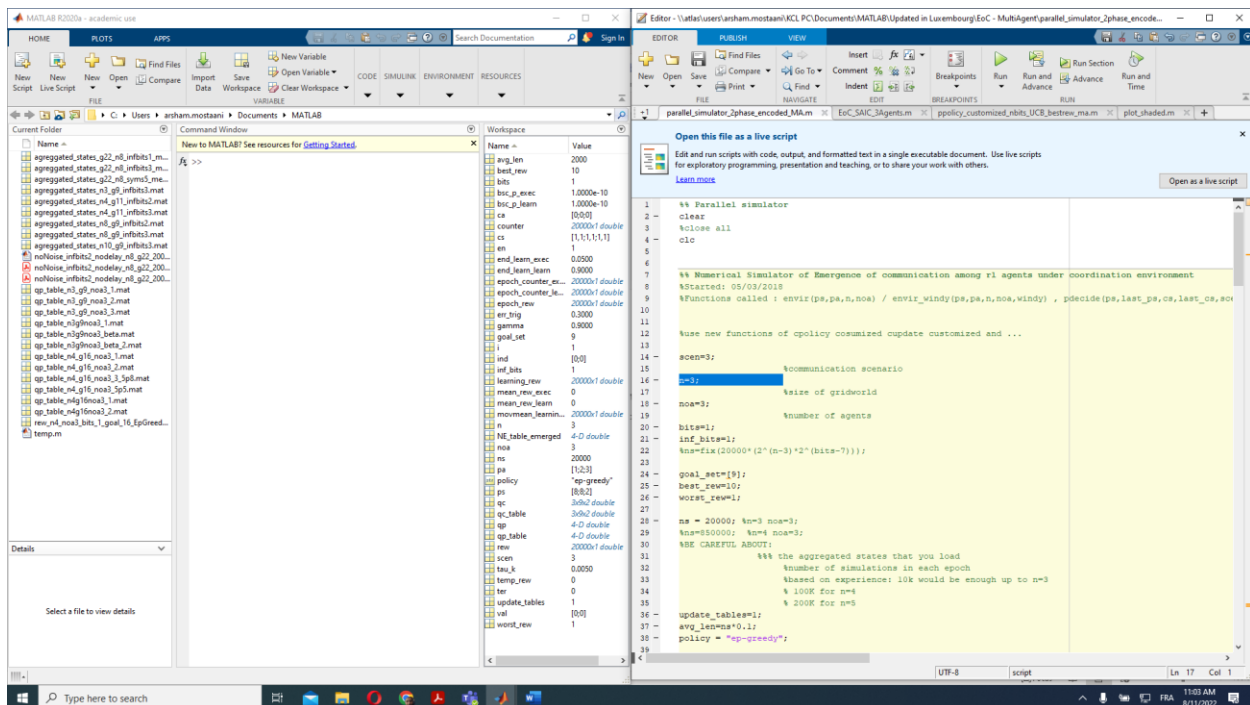


- The parameter setup is control panel that gives the user all sorts of control facilities to adjust the simulations to their needs.
- 2- “EoC_SAIC_3Agents ” is a function wrapper that encapsulates all functions that interact with each other to completely simulate one episode of training.

- 4- All these scripts and functions are available in the current repository:

Adjustment of the parameter setup

- 1- Select the number of grids in each direction e.g., $n = 3$ is selected when you want a 3x3 grid world



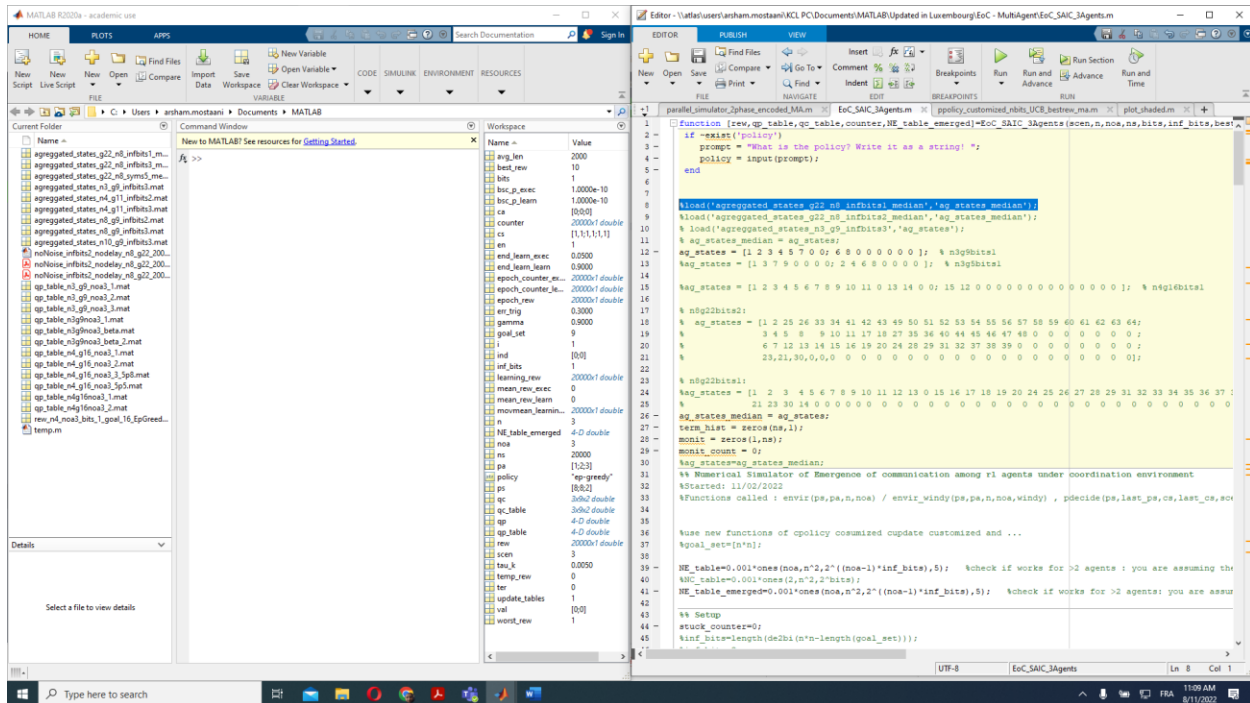
- 2- Select the number of agents with the parameter noa (standing for the Number of Agents)

noa=3;

- 3- Select the number of communication bits – in the case of task-oriented quantization bits and inf_bits should have similar values

```
bits=1;
inf_bits=1;
```

- 4- According to the size of the grid-world and the number of quantized bits, load the right quantizer inside the wrapper, or obtain your quantizer by running the centralized training phase



5- Select the right goal set/location
`goal_set=[9] ;`

6- Update tables = 1 – This parameter should only be zero when using Apprenticeship learning wrapper.
`update_tables=1;`

7- Choose the right policy – we have obtained very good results with decaying epsilon greedy. Decaying epsilon greedy or Linearly anealing epsilon, has been detailed also in David Silver's seminal work "Mnih, Volodymyr, et al. "Human-level control through deep reinforcement learning." *nature* 518.7540 (2015): 529-533." :

`policy = "ep-greedy"`

8- Check your rewarding scheme – the default setting is adjusted according to our paper: "Task-Oriented Data Compression for Multi-Agent Communications Over Bit-Budgeted Channels". Accordingly, the common reward signal increases exponentially as the number of agents arrived at the goal point are increased.