

DSSE - Drone Swarm Search Environment



Alunos: Jorás Oliveira, Pedro Andrade, Ricardo Rodrigues, Renato Laffranchi.

Professor: Dr. Fabricio Barth.

Ambientes

- Ambiente feito para missões de **SAR marítimas**.
- Feito para uso de drones **multi-agente**.
- Rewards **esparsas** e **densas**.

Search Environment

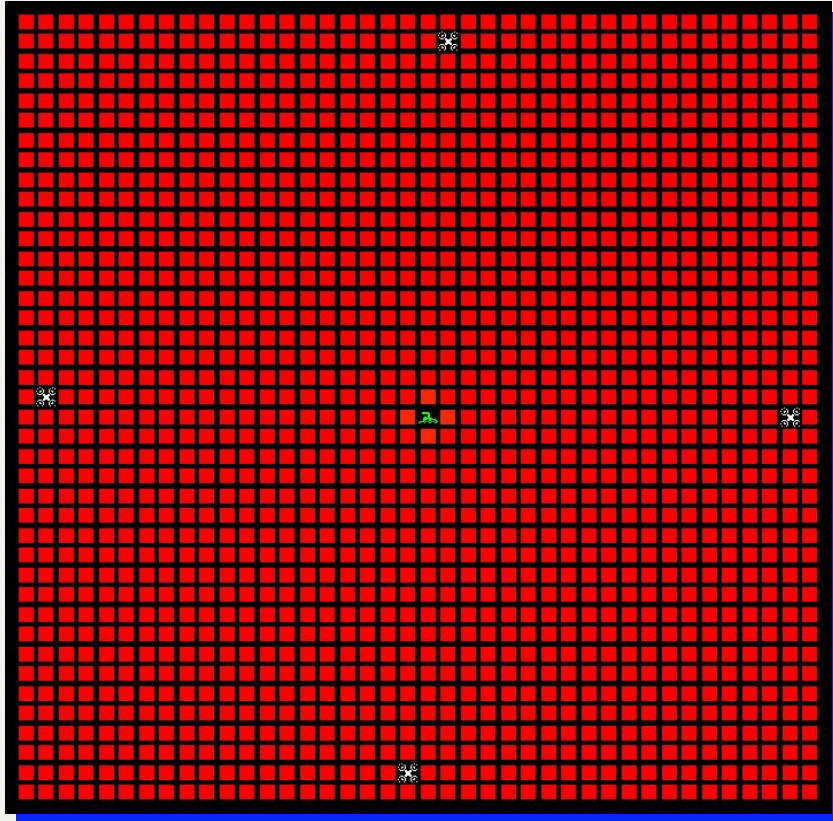
X

Coverage Environment

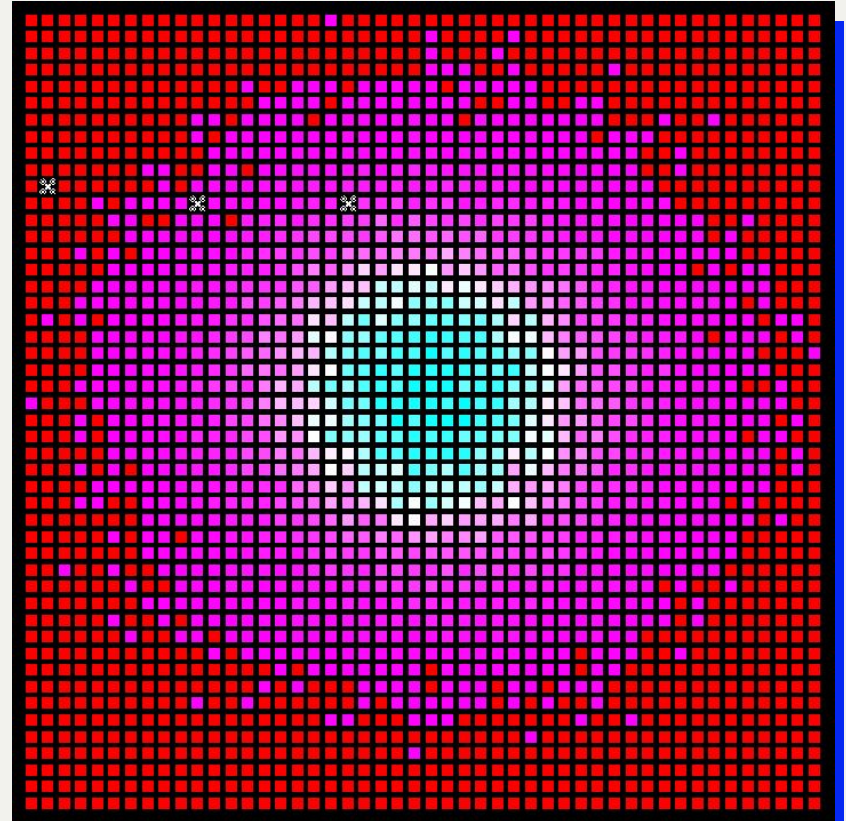
Action Space	Discrete (g)
Action Values	[0, 1, 2, 3, 4, 5, 6, 7, 8]
Observations	{droneN: ((x, y), probability_matrix)}

Ambiente

Search



Coverage



Hipóteses Formuladas

1

Algoritmos de RL podem **superar** a política greedy em ambientes mais complexos?

2

Algoritmos com redes independentes **convergem** mais rápido?

3

A **comunicação** da trajetória dos agentes **aumenta o sucesso** da missão de SAR?

4

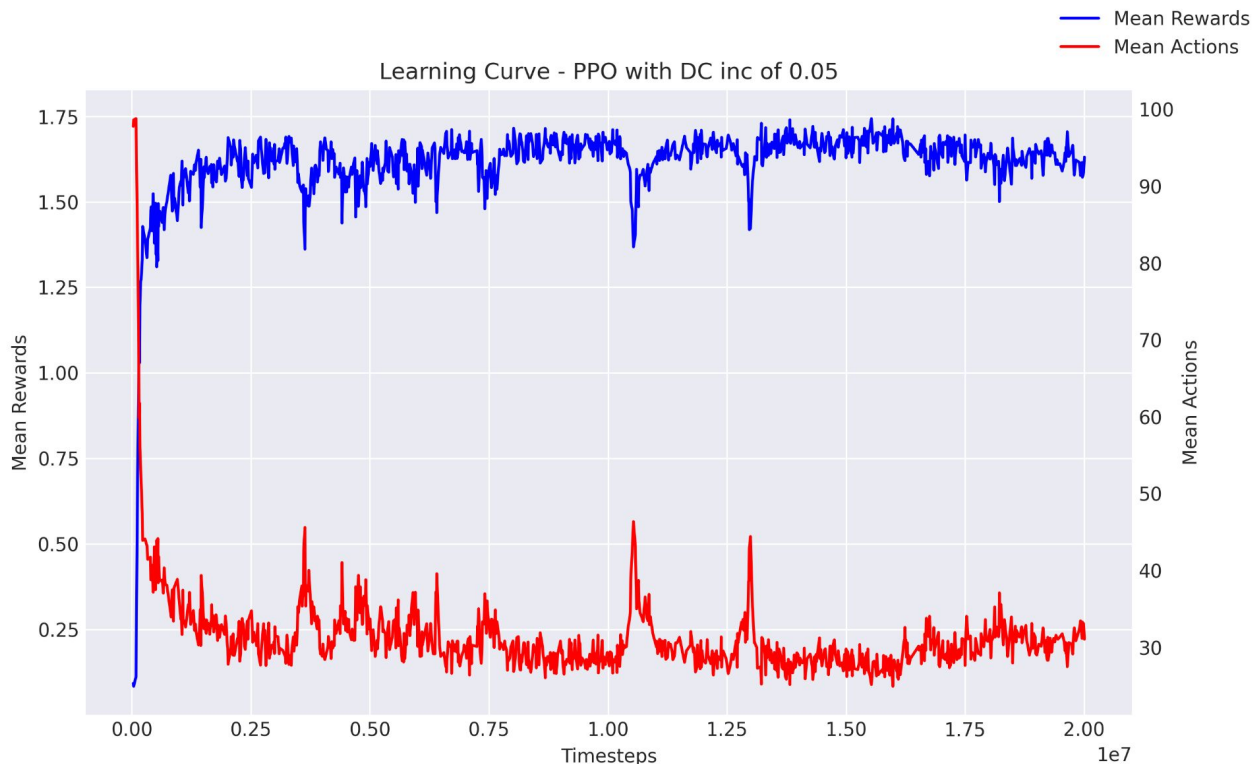
Agentes conseguem cobrir uma área no **menor tempo** e priorizar áreas de maior probabilidade?

5

Como o padrão de busca dos agentes **muda** quando há **mais de um naufrago**?

Resultados H1 - RL supera política greedy ?

PPO com incremento de dispersão 0.05

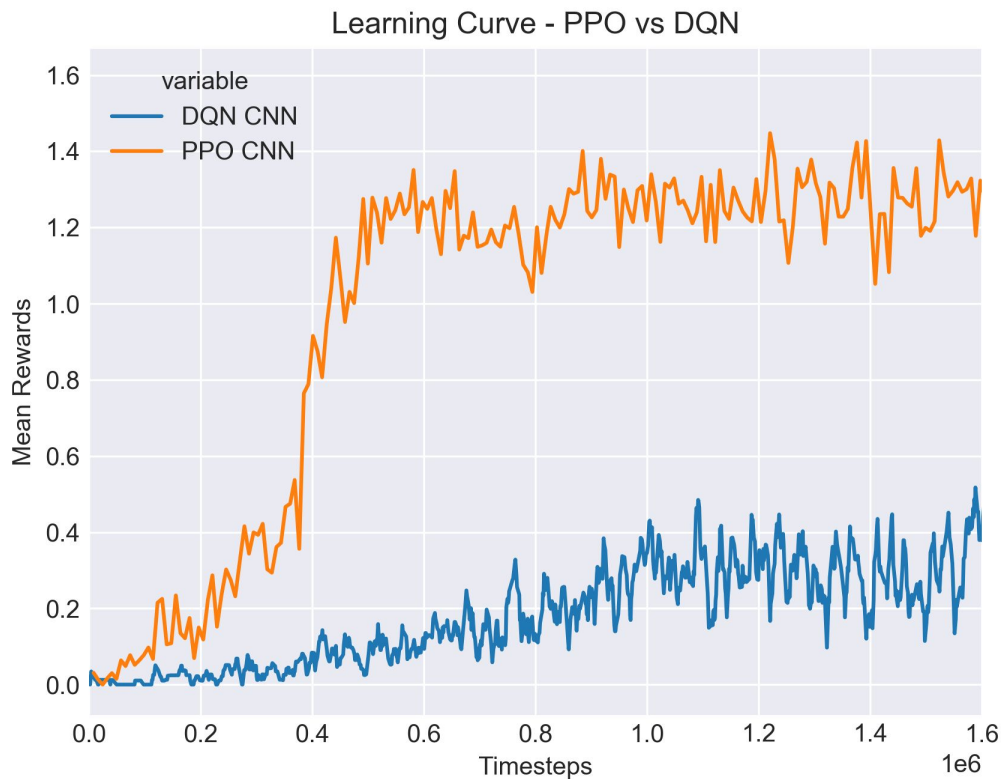


Resultados H1 - RL supera política greedy ?

PPO com incremento de dispersão 0.1



Resultados H1 - RL supera política greedy ?



Resultados H1 - RL supera política greedy ?

Incremento de dispersão de 0.1

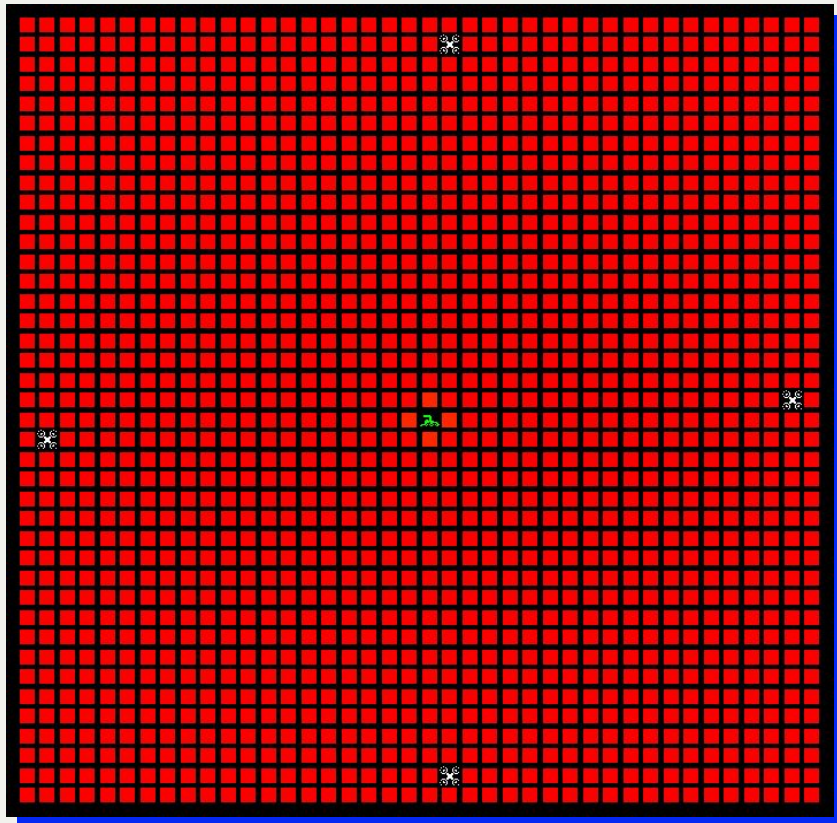
Algoritmo Métrica	PPO	Greedy Search
Taxa de sucesso (%)	75.44	35.84
Média do número de ações	42.47	77.48
Mediana do número de ações	23	100
Reward média	1.34	0.59

Incremento de dispersão de 0.05

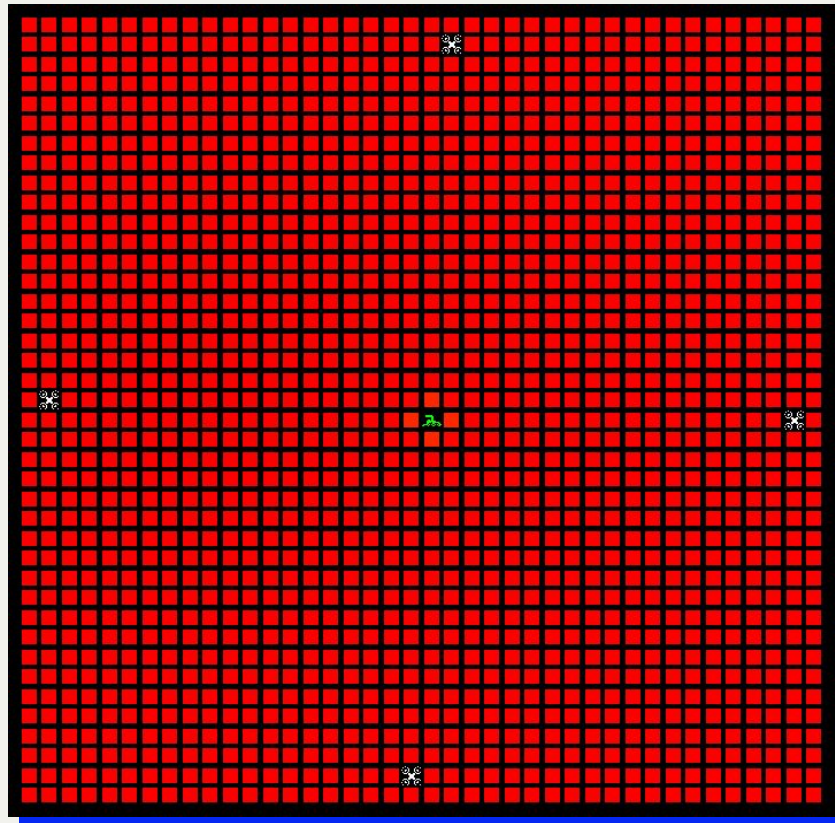
Algoritmo Métrica	PPO	Greedy Search
Taxa de sucesso (%)	83.0	50.18
Média do número de ações	35.91	65.07
Mediana do número de ações	22.0	94
Reward média	1.48	0.86

Resultados H1 - RL supera política greedy ?

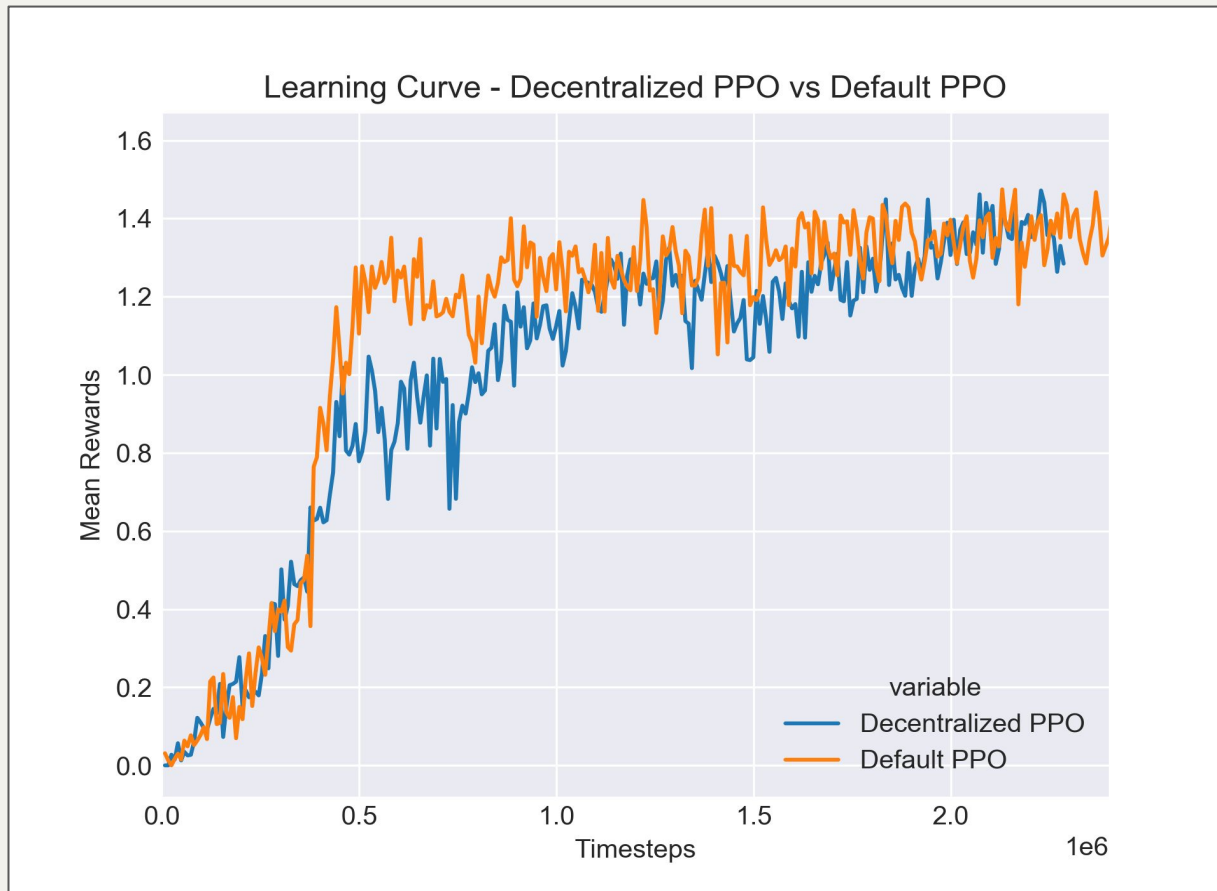
PPO



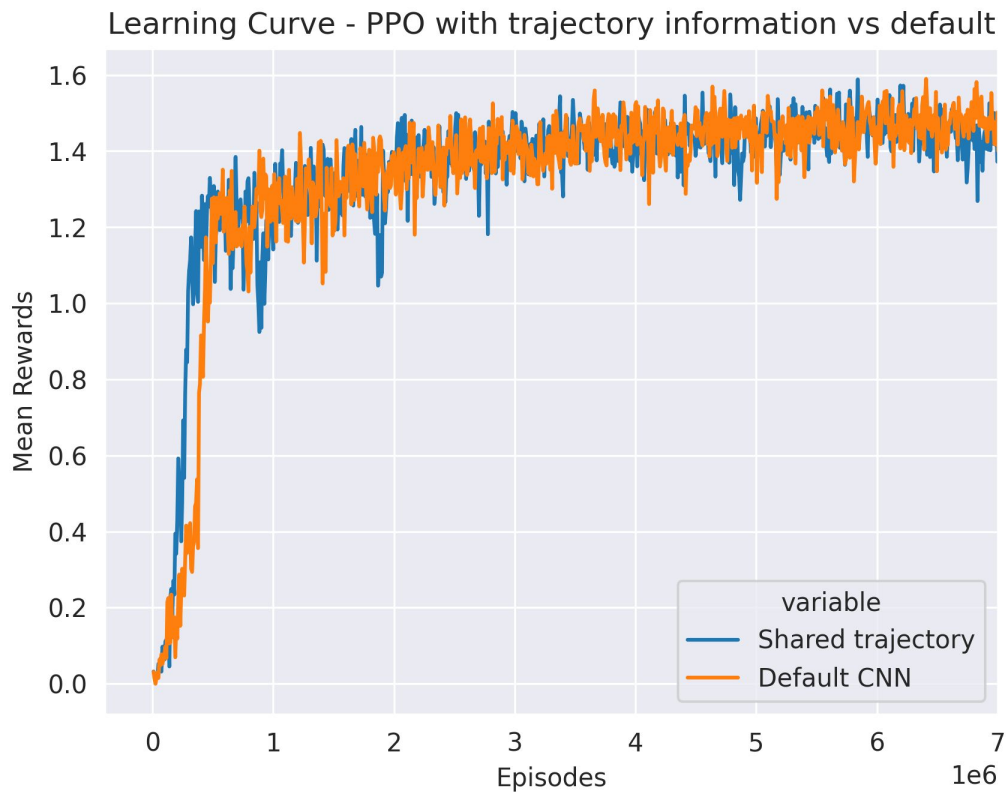
Greedy



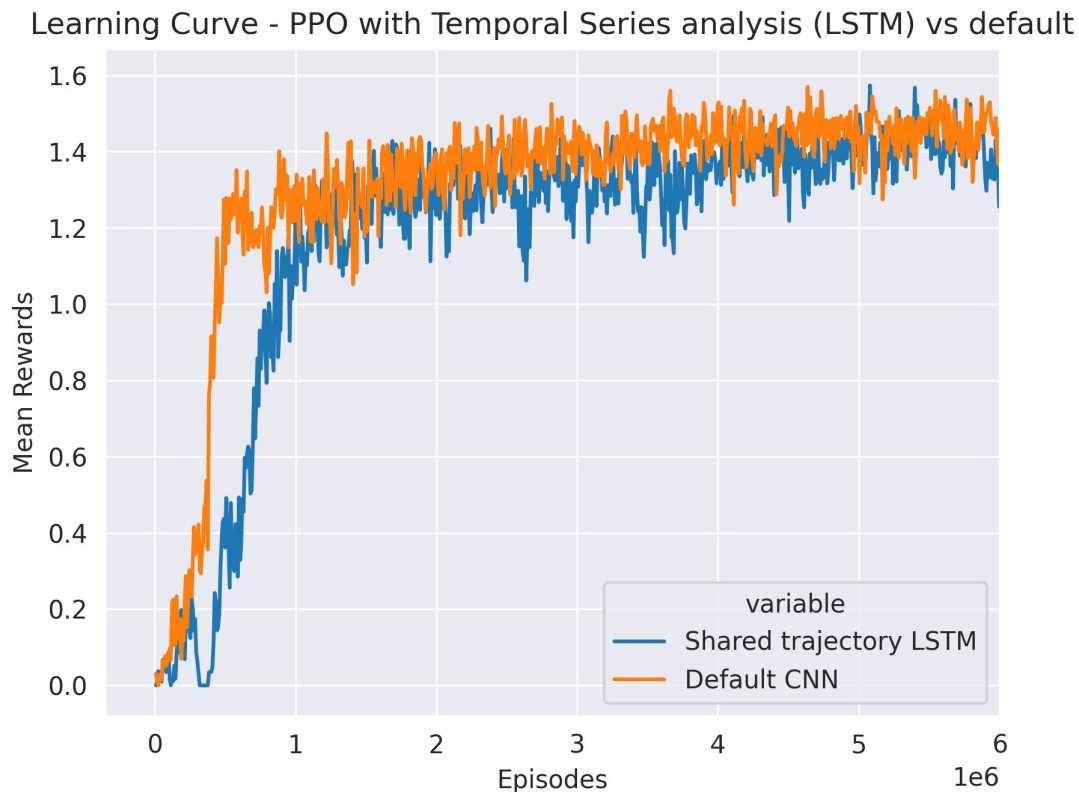
Resultados H2 - Redes independentes convergem mais rápido ?



Resultados H3 - Saber a trajetória impacta no sucesso ?



Resultados H3 - Saber a trajetória impacta no sucesso ?

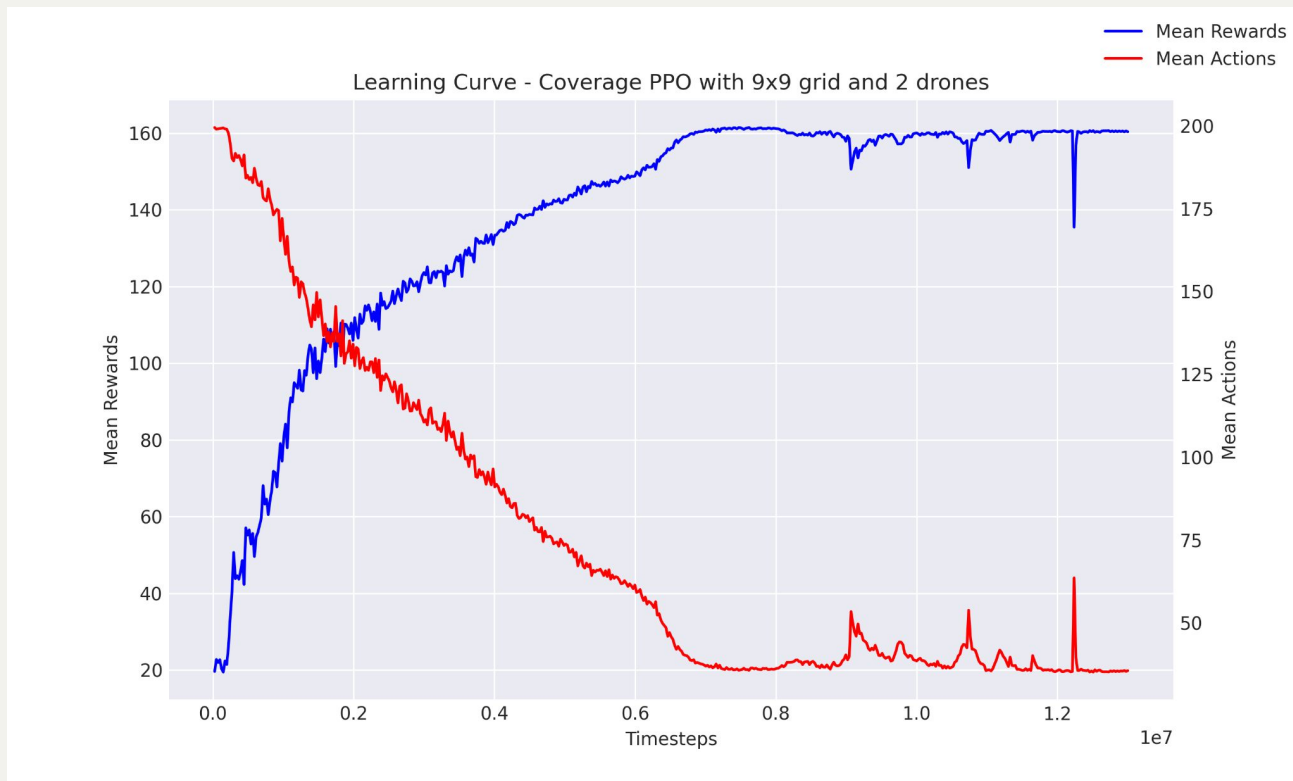


Resultados H3 - Saber a trajetória impacta no sucesso ?

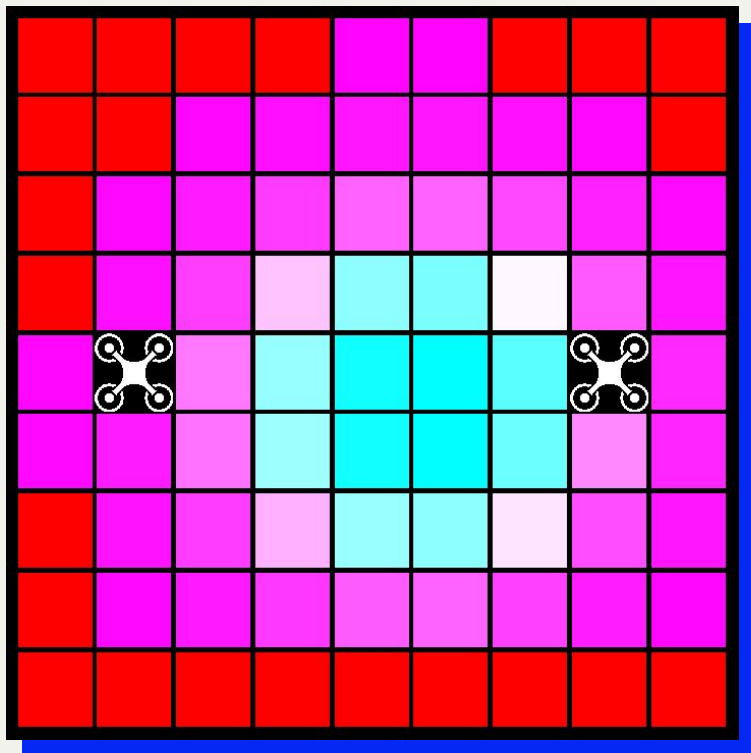
Comparação de PPOs com diferentes abordagens de comunicação da trajetória.

Algoritmo Métrica	PPO normal	PPO com trajetória na matriz	PPO com LSTM
Taxa de sucesso (%)	75.44	75.98	76.46
Média do número de ações	42.47	41.99	41.57
Mediana do número de ações	23	23	23
Reward média	1.34	1.35	1.36

Resultados H4 - Agentes podem cobrir áreas rápido e com prioridade ?



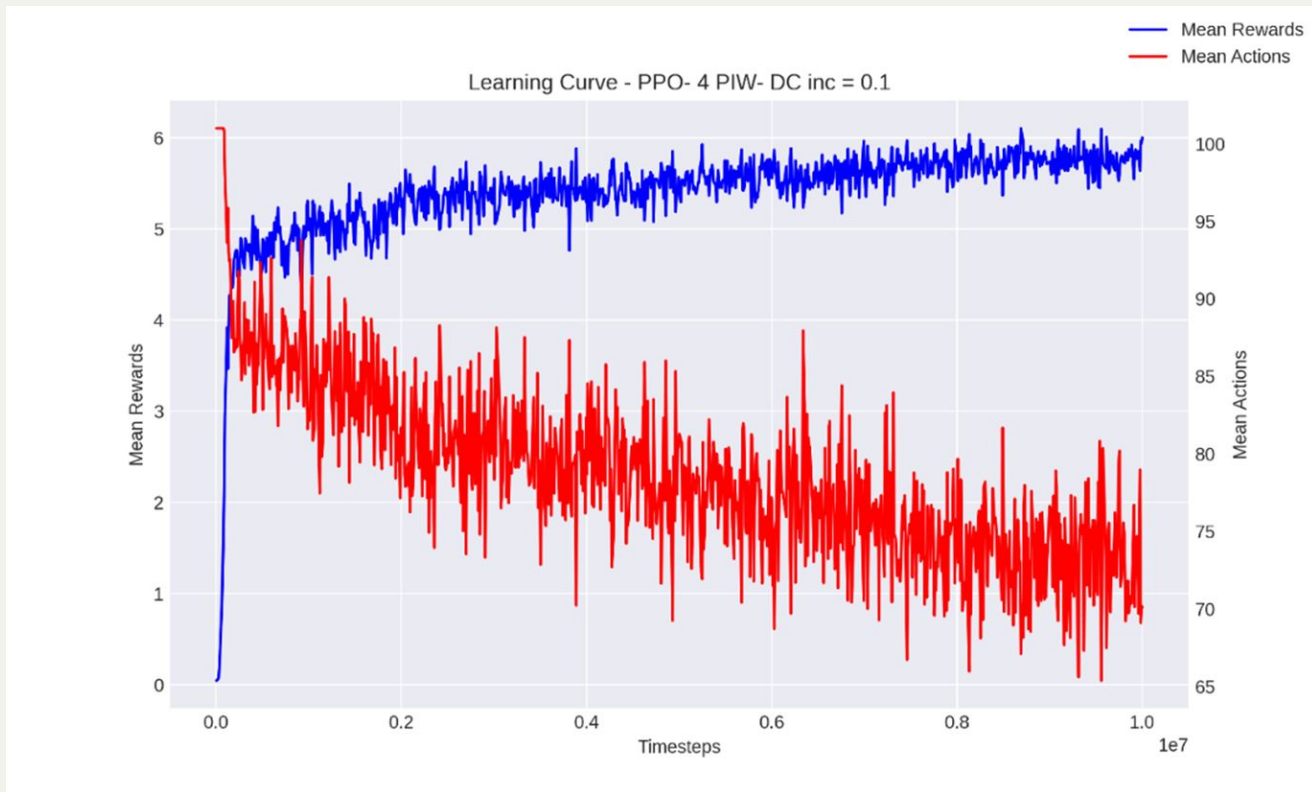
Resultados H4 - Agentes podem cobrir áreas rápido e com prioridade ?



Detalhes dos Resultados

- Busca feita com **2 agentes** (drones).
- Experimentos do estado da arte foram com **30^1** e **19^2** células.
- **58 células** foram distribuídas em 34 passos, equivalente a **58.6% do tempo necessário**.
- Priorização de células com **maior probabilidade**.

Resultados H5 - Como muda a busca com múltiplos naufragos ?

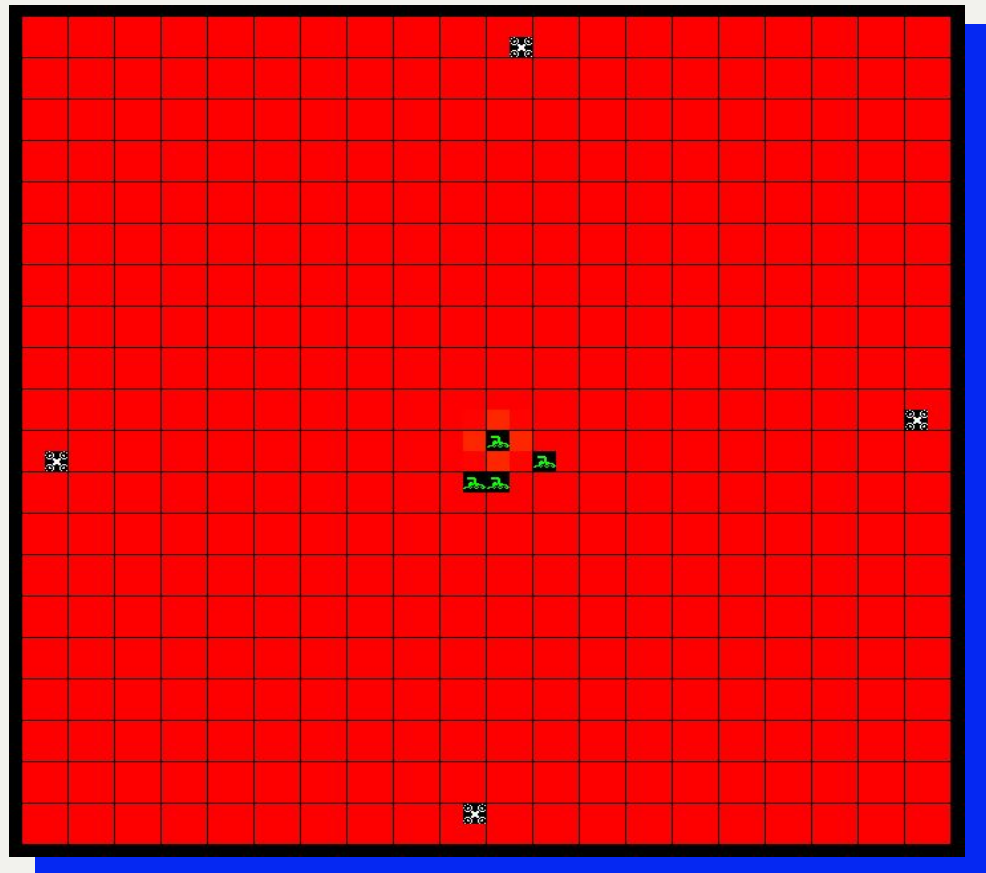


Resultados H5 - Como muda a busca com múltiplos naufragos ?

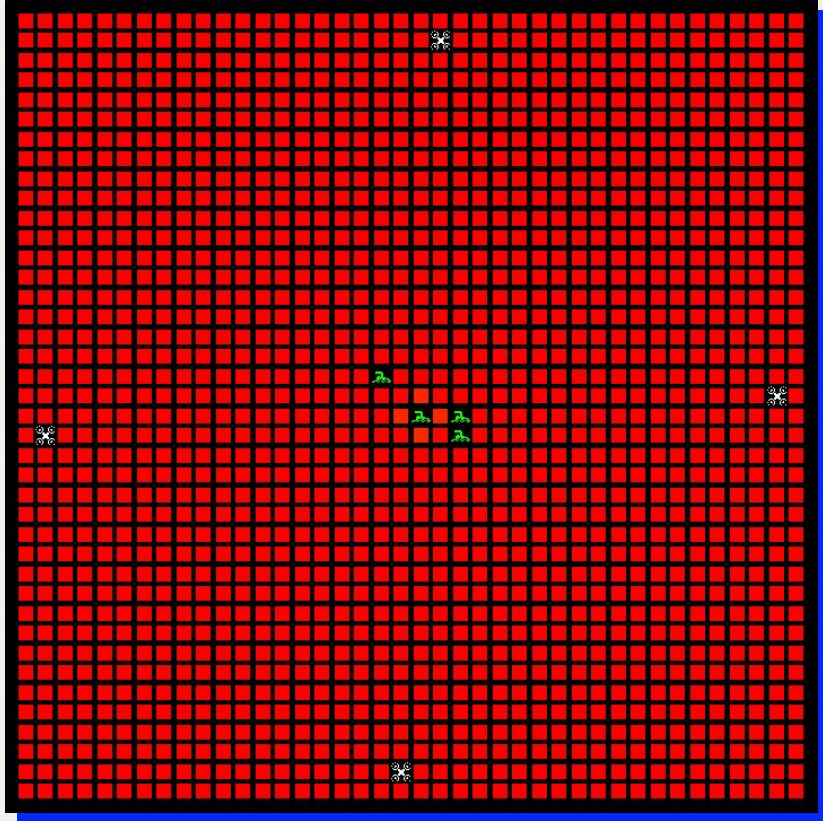
Avaliação dos resultados do PPO para 4 naufragos:

Métrica \ Algoritmo	PPO
Encontrou todos os naufragos (% dos testes)	21.54
Média do número de naufragos encontrados	2.3
Média do número de ações	86.81

Resultados H5



Resultados H5 - Como muda a busca com múltiplos naufragos ?



Descrição dos Resultados

- Cria uma estratégia de **cercar** os targets (náufrago)

Obrigado

Q & A

Appendix A: Configurações dos testes

Hyperparâmetros PPO

Parâmetro	Valor	Descrição
B	8192	Training batch size
Lr	10^{-5}	Learning rate
γ	0.99999999	Discount factor
M	300	Stochastic Gradient Descent (SGD) minibatch
K	10	Number of SGD iterations

Appendix A: Configurações dos testes

Hyperparâmetros DQN

Parâmetro	Valor	Descrição
B	8192	Training batch size
Lr	10^{-5}	Learning rate
γ	0.99999999	Discount factor
U	500	Update target network every U steps.
ε_0	1	Initial epsilon for ε -greedy.
ε_f	0.1	Final epsilon for ε -greedy.
T	400000	T timesteps for epsilon to decay from ε_0 to ε_f .

Appendix A: Configurações dos testes

Configurações do ambiente de search para os testes.

Parâmetro	Valor
Incremento de dispersão	0.1 (0.05 H1)
Número de PIW	1 (4 H5)
Tamanho do grid	40x40
Número de drones	4
Número máximo de timesteps por simulação	100

Appendix A: Configurações dos testes

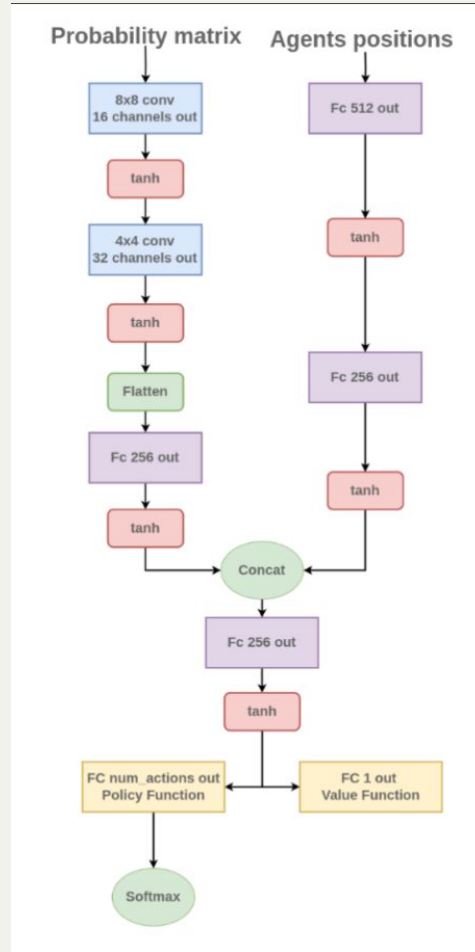
Configurações do ambiente de coverage para os testes.

Parâmetro	Valor
Horas da simulação de partículas	2
Posição naufrágio	-24.04 lat, -46.17 long (oceano próximo ao guarujá)
Tamanho do grid	9x9
Número de drones	4
Número máximo de timesteps por simulação	200

Appendix B:

Modelos - RL

PPO



DQN

