1. Criação do Ambiente de aprendizagem no Unity
2. Criação do(s) Agente(s)
   1. Quais as observações?
   2. Quais as ações?
   3. Que recompensas extrínsecas?
   4. Validação do setup de treino usando o modo Heurístico
3. Treino do Agente usando algoritmo de aprendizagem por reforço
   1. Que algoritmo?

No Unity, o framework ML-Agents (Machine Learning Agents) permite treinar agentes usando algoritmos de aprendizagem por reforço. O ML-Agents inclui uma implementação do algoritmo de aprendizagem por reforço chamado Proximal Policy Optimization (PPO).

O Proximal Policy Optimization é um algoritmo de aprendizagem por reforço que pertence à família de algoritmos de gradiente de política. Ele é projetado para treinar agentes em ambientes complexos, onde a política do agente é atualizada iterativamente para melhorar o desempenho. O PPO utiliza otimização baseada em gradiente e busca maximizar a função de recompensa acumulada ao longo do tempo.

Portanto, se você estiver usando o Unity ML-Agents para treinar um agente em um ambiente, é provável que você esteja usando o algoritmo Proximal Policy Optimization (PPO) para atualizar a política do agente durante o treinamento.

* 1. Que parâmetros de treino?
  2. Como evolui o desempenho do agente ao longo do treino (por exemplo, usar os gráficos do tensorboard)

Uma imagem com parque infantil, captura de ecrã, verde, relva

Descrição gerada automaticamente

Uma imagem com texto, Tipo de letra, captura de ecrã

Descrição gerada automaticamente

1. Avaliação do desempenho do Agente, depois de treinado, ou seja, fazer inferência, usando a rede neuronal que resultou do treino, no Unity (no ambiente de aprendizagem ou noutro relacionado com aquele)



Uma imagem com Retângulo, quadrado, interior

Descrição gerada automaticamente

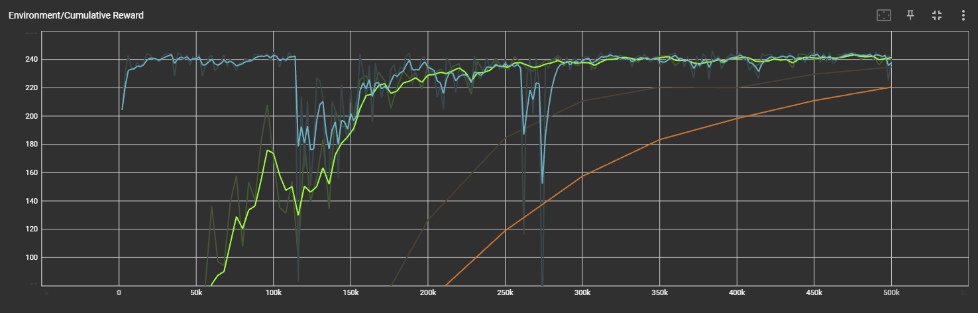
Uma imagem com Retângulo, quadrado, captura de ecrã, Saturação de cores

Descrição gerada automaticamente

File: RayAgent.yaml – configurações das rede e de treino

Uma imagem com texto, captura de ecrã, software, Sistema operativo

Descrição gerada automaticamente



Uma imagem com captura de ecrã, file, Gráfico

Descrição gerada automaticamente

Mudou-se de ambiente para um mais complicado:



Uma imagem com quadrado, captura de ecrã, Retângulo, verde

Descrição gerada automaticamente

Uma imagem com Retângulo, quadrado, captura de ecrã

Descrição gerada automaticamente

Mudança no campo:

Uma imagem com captura de ecrã, quadrado, Retângulo

Descrição gerada automaticamente

Problemas que o agente conseguiu contornar:

Uma imagem com captura de ecrã, quadrado, Retângulo, verde

Descrição gerada automaticamente

Criar outro mundo:

Uma imagem com captura de ecrã, verde, Retângulo, quadrado

Descrição gerada automaticamente

Criar novo mundo:

Uma imagem com Retângulo, quadrado, captura de ecrã, padrão

Descrição gerada automaticamente

Uma imagem com captura de ecrã, Retângulo, quadrado

Descrição gerada automaticamente

Steps a por ne relatório:

- introdução

- introdução a cena

- agentecom reforcement learning

- perceber resultados

Uma imagem com texto, Tipo de letra, captura de ecrã, logótipo

Descrição gerada automaticamente

Uma imagem com texto, captura de ecrã, software, Software de multimédia

Descrição gerada automaticamente

Uma imagem com captura de ecrã, texto

Descrição gerada automaticamente

Uma imagem com texto, captura de ecrã, Tipo de letra, software

Descrição gerada automaticamente

Uma imagem com texto, captura de ecrã, software

Descrição gerada automaticamente

Uma imagem com texto, captura de ecrã, Tipo de letra, logótipo

Descrição gerada automaticamente

Limitações do reinforcement learning

Normal run

Uma imagem com captura de ecrã, quadrado, Retângulo

Descrição gerada automaticamente

Uma imagem com captura de ecrã, texto, diagrama, design

Descrição gerada automaticamente

Uma imagem com captura de ecrã, diagrama, texto

Descrição gerada automaticamente

Metodo de imitation – heuristic only and demostration recorder e correr algumas vezes como jogador para a ia aprender por imitacao

Uma imagem com texto, captura de ecrã, software, Software de multimédia

Descrição gerada automaticamente

Copiar das demos o AgentDemo\_Imitation.demo para os assests:

Uma imagem com texto, captura de ecrã, software, Tipo de letra

Descrição gerada automaticamente

Para utilizar a demo:

Abrir o file config/MoveToTarget.yaml

Uma imagem com texto, captura de ecrã, software, Software de multimédia

Descrição gerada automaticamente

Run com os dados da demo no config file:

Uma imagem com texto, captura de ecrã, software, Software de multimédia

Descrição gerada automaticamente

No file:

Demos/AgentDemo.demo

Modar o nome no MoveToTarget.yaml ou no Demos/AgentDemo.demo



Uma imagem com texto, captura de ecrã, Tipo de letra

Descrição gerada automaticamente

Uma imagem com texto, captura de ecrã, documento, carta

Descrição gerada automaticamente