
Estratégias de Busca em IA (INF 628)

Informações Gerais

Neste trabalho você irá implementar um algoritmo genético para o problema de controle do robô da OpenAI. O trabalho deve ser desenvolvido individualmente. A colaboração entre colegas de classe é permitida, mas sempre através da troca de ideias, nunca através da troca de código fonte.

Entrega do trabalho Enviar no PVANet um relatório explicando o método desenvolvido e os resultados obtidos. No relatório do projeto você irá colocar um link para a sua conta no Github com o código do projeto. Não irei aceitar entregas de trabalho por email e o relatório deve ser entregue em pdf. A forma de escrita do relatório é livre.

Política de Atraso Cada aluno poderá atrasar um total de 7 dias a entrega de trabalhos durante o semestre. Uma vez esgotados os dias de atraso, a nota do trabalho atrasado será zero. Por exemplo, ao entregar o primeiro trabalho com 2 dias de atraso, a nota do aluno não sofrerá nenhuma penalização. No entanto, se no segundo trabalho o aluno atrasar 6 ou mais dias, a nota do segundo trabalho será zero.

Código do Trabalho

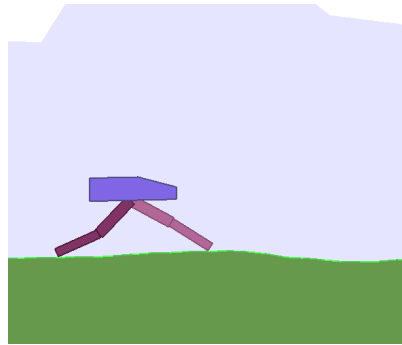
Veja abaixo um 'Hello World' do ambiente do robô da OpenAI. Para rodar o código, instale via pip3 as seguintes bibliotecas: gym e box2d-py.

```
import gym
env = gym.make('BipedalWalker-v2')
env.reset()
for _ in range(500):
    env.render()
    action = env.action_space.sample()
    print('Action: ', action)
    observation, reward, done, info = env.step(action) # take a random action
    print('Reward: ', reward)
env.close()
```

O Problema

O robô bípede exibido na figura abaixo deve ser controlado com ações para os dois joelhos e as duas juntas no quadril. As ações aceitam valores entre -1 e 1. Para os propósitos do trabalho não é necessário saber

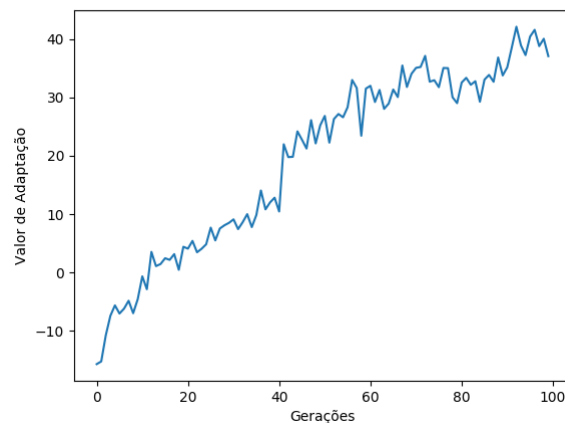
o significado dos números. O ambiente também oferece observações feitas pelos sensores do robô, que não necessariamente serão utilizadas nesse trabalho.



O robô começa na posição vertical e tem como objetivo caminhar até ao final da fase, à direita.

A Solução

Nesse trabalho você deve utilizar um algoritmo genético para criar uma sequência de ações que faça o robô caminhar. O gráfico abaixo mostra o valor de recompensa (utilizado como função de adaptação na minha implementação) obtido pelo melhor indivíduo de várias gerações de um algoritmo genético.



O valor de adaptação varia com o tamanho do episódio. O gráfico foi gerado com episódios de tamanho 500. Episódios maiores retornarão valores maiores de adaptação, assim como episódios menores retornarão valores menores de adaptação.