

Trabalho 4 Aprendizado por Reforço

1. Introdução

O objetivo do trabalho é explorar a implementação de algoritmos de aprendizado por reforço e análise da influência de alguns parâmetros.

2. Instruções

As implementações devem ser feitas em Python 3, preenchendo as funções em arquivos específicos do kit do trabalho. Não é necessário instalar bibliotecas adicionais.

O kit do trabalho está disponível no GitHub: <https://github.com/andertavares/berkeleyRL3>. Clone-o para começar o trabalho.

O `Readme.md` do repositório contém uma série de 8 exercícios para serem feitos com o kit. Note que todos os exercícios possuem um corretor automatizado de forma que você pode verificar sua implementação. A pontuação será a seguinte:

Exercício	Pontos
1	3
2	0.5
3	2.25
4	2.5
5	0.5
6	0.25
7	0.5
8	0.5
Total	10

3. Entrega

Você deverá enviar no moodle um arquivo `.zip` contendo 4 arquivos em sua raiz:

1. `valueIterationAgents.py`, com sua implementação do algoritmo de Iteração de Valor;
2. `qlearningAgents.py`, com sua implementação do Q-learning;
3. `analysis.py`, com respostas às questões pedidas;
4. `group.md`, um arquivo de texto sem formatação ou Markdown com os nomes, cartões de matrícula e turma dos integrantes do grupo (não é um `Readme.md` como nos outros trabalhos para evitar confusão com o `Readme.md` do repositório com o kit do trabalho).

Note que os 3 primeiros arquivos estão no kit do trabalho. Você deverá preenchê-los, rodar o autograder pra ver se sua implementação está correta, juntá-los com seu `group.md` no seu `.zip` e enviar esse `.zip` no link de entrega do moodle.

IMPORTANTE: todos os 4 arquivos devem estar na raiz do seu `.zip`!

4. Observações gerais

- O trabalho deve ser feito em trios.
- Fiquem atentos à política de plágio!
- É a corretude da sua implementação - e não o julgamento do verificador automatizado (autograder) - que dará o veredito final da sua nota. Se necessário, revisaremos os envios manualmente para garantir que seu trio receba a pontuação apropriada para o trabalho. Porém, note que é importante não haver erros provocados por “desobediência” à especificação do trabalho (e.g. coisas que “quebram” o autograder, como trocar o tipo dos argumentos ou do retorno das funções). Esse tipo de erro é considerado como implementação incorreta.

Política de Plágio

Trios poderão apenas discutir questões de alto nível relativas a resolução do problema em questão. Poderão discutir, por exemplo, questões sobre as estruturas de dados utilizadas, técnicas para visualizar dados, etc. Não é permitido que os trios utilizem quaisquer códigos fonte provido por outros trios, ou encontrados na internet. Em especial, é proibido o uso do ChatGPT e similares para solução do problema. Além disso, recomenda-se a desativação do Github Copilot e similares pois essas ferramentas reproduzem implementações em repositórios online e isso é entendido como plágio.

Pode-se fazer consultas na internet ou em livros apenas para estudar o modo de funcionamento das técnicas de IA, e para analisar o pseudo-código que as implementa. Não é permitida a análise ou cópia de implementações concretas (em quaisquer linguagens de programação) da técnica escolhida. O objetivo deste trabalho é justamente implementar as técnicas do zero e descobrir as dificuldades envolvidas na sua utilização para resolução do problema em questão. Toda e qualquer

fonte consultada pelo trio (tanto para estudar os métodos a serem utilizadas, quanto para verificar a estruturação da técnica em termos de pseudo-código) precisa obrigatoriamente ser citada no Readme.md.

Usamos rotineiramente um sistema anti-plágio que compara o código-fonte desenvolvido pelos trios com soluções enviadas em edições passadas da disciplina, e também com implementações disponíveis online.

Qualquer nível de plágio (ou seja, utilização de implementações que não tenham sido 100% desenvolvidas pelo trio) poderá resultar em nota zero no trabalho. Caso a cópia tenha sido feita de outro trio da disciplina, todos os envolvidos (não apenas os que copiaram) serão penalizados. Esta política de avaliação não é aberta a debate. Se você tiver quaisquer dúvidas se uma determinada prática pode ou não, ser considerada plágio, não assuma nada: pergunte ao professor e aos monitores.

Note que, considerando-se os pesos das avaliações desta disciplina (especificados e descritos no Plano de Ensino), nota zero em qualquer um dos trabalhos de implementação abaixa muito a média final dos projetos práticos, e se ela for menor que 6, não é permitida prova de recuperação. Ou seja: caso seja detectado plágio, há o risco direto de reprovação. Os trios deverão desenvolver o trabalho sozinhos.