



بسمه تعالی

مینی پروژه اول

یادگیری تقویتی

مهلت تحویل : ۱۴۰۲/۰۱/۲۷

نیم سال دوم ۱۴۰۱

استاد : دکتر هراتی

دانشگاه فردوسی مشهد - گروه مهندسی کامپیوتر

مقطع کارشناسی ارشد - هوش مصنوعی و رباتیک

هدف از این مینی پروژه پیاده سازی روش های پیمایش ارزش^۱ و پیمایش سیاست^۲ برای محیط^۳ Frozen Lake کتابخانه OpenAI Gym است. برای پیاده سازی هر کدام از روش ها کد پایتونی در فایل frozenlake.zip وجود دارد که شما باید کد خود را در قسمتهای # Implement policy iteration here و # Implement value iteration here بنویسید. توجه کنید که توابع شما باید ارزیابی تابع ارزش بهینه^۴ و همچنین سیاست بهینه^۵ برای هر حالت را بازگرداند. در نهایت برای درک بهتر میتوانید به کمک کتابخانه Matplotlib خروجی ها را به صورت یک جدول نمایش دهید.



محیط Frozen Lake:

عامل در این محیط از خانه (S) شروع کرده و سعی در عبور از یک دریاچه یخ زده (F) برای رسیدن به هدف (G) را دارد در این مسیر ممکن است در سوراخ ها (H) گیر بیافد که در این صورت باید از خانه شروع مجدد پیمایش خود را آغاز کند به دلیل لغزنده بودن دریاچه یخ زده، عامل ممکن است همیشه در جهت مورد نظر حرکت نکند. عامل در هر خانه چهار عمل ممکن دارد که شامل بالا، پایین، چپ و راست است. اعمالی که باعث شوند که عامل از محیط خارج شود موقعیت عامل را تغییر نمیدهند. رسیدن به هدف (G) پاداش +1 را برای عامل به همراه دارد.

پس از پیاده سازی روش های خواسته شده موارد زیر را بررسی کنید و نتایج آن را به صورت مناسب گزارش کنید.

- تغییر مقدار ضریب تنزیل به نحوی که تاثیر آن در پاسخ نهایی قابل مشاهده باشد.
 - حل مساله به صورت غیر قطعی (تغییر پارامتر is_slippery).
 - اضافه کردن جریمه 0.05- برای هر تغییر موقعیت عامل.
 - اضافه کردن جریمه 2- برای تغییراتی که منجر به قرار گرفتن عامل در سوراخ ها (H) شوند.
- راهنمایی : برای مقایسه پارامتر ها و همچنین حل مساله در محیط غیرقطعی نمودار پاداش تجمعی برحسب تعداد پیمایش^۶ را رسم کنید .

۱. Value iteration

۲. Policy iteration

۳. Frozen Lake

۴. value_function

۵. Policy

۶. iteration

توضیحات :

- این پروژه به صورت انفرادی است.
- فایل نهایی برای ارسال باید شامل پیاده سازی کد و گزارش باشد.