

This file has been cleaned of potential threats.

If you confirm that the file is coming from a trusted source, you can send the following SHA-256 hash value to your admin for the original file.

06b3ba6bb41ab29e49129dea510ed7af0f238aea744602bd7f370967b6dd120c

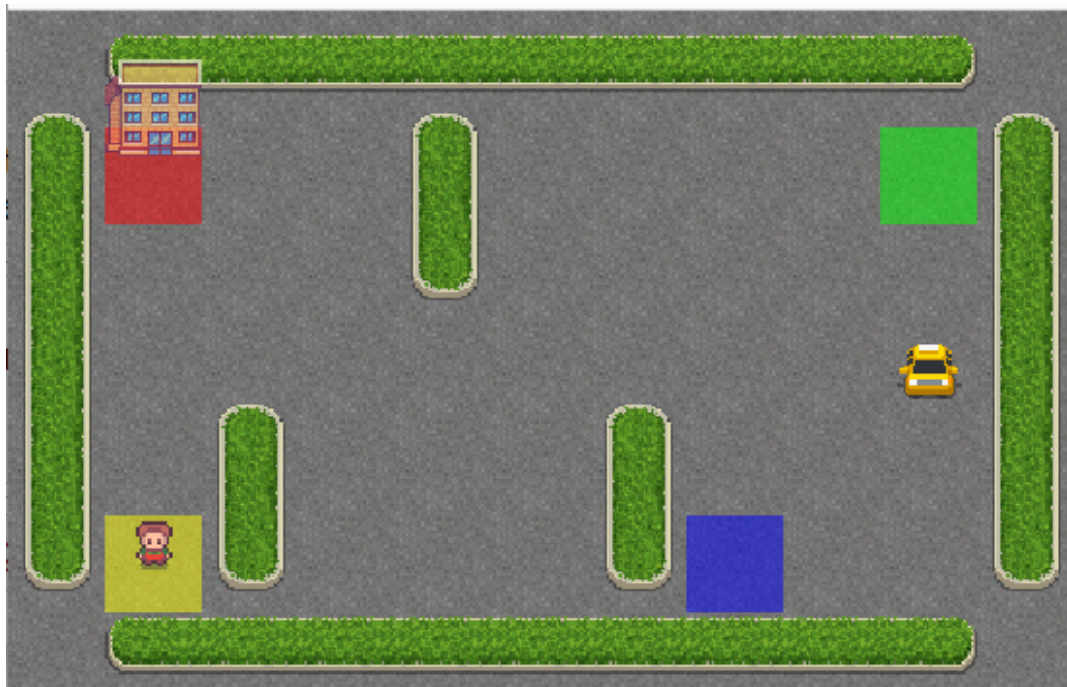
To view the reconstructed contents, please SCROLL DOWN to next page.

به نام خدا

پروژه دوم درس یادگیری تقویتی

در این پروژه قصد داریم الگوریتم ها Monte Carlo , Q-Learning و Sarsa را پیاده سازی کنیم .

محیطی که برای انجام این کار در نظر گرفتیم محیط Taxi در کتابخانه Gym است . محیط مورد نظر به شکل زیر است :



تاکسی مورد نظر و یک مسافر را مشاهده می کنید که هدف این است که تاکسی مسافر مورد نظر را به مقصدش برساند.

برای توضیحات بیش تر می توانید از لینک زیر استفاده کنید .

https://gymnasium.farama.org/environments/toy_text/taxi/

در این محیط اکشن های مورد نظر چهار جهت جغرافیایی و پیاده و سوار کردن مسافر است . ولی ما در این پروژه قصد داریم ۴ جهت مورب را هم اضافه کنیم.

۱. یک Wrapper بنویسید که اکشن های مورد نظر را پشتیبانی کند .

۲. الگوریتم MC First visit را پیاده سازی کنید و مقادیر ارزش را نمایش دهید . این نتایج را به γ های ۰.۹ و ۰ نمایش دهید . همچنین نمودار پاداش تجمعی را رسم کنید .

۳. الگوریتم Q-Learnin را پیاده سازی کنید و مقادیر ارزش را نمایش دهید . این نتایج را به γ های ۰.۹ و ۰ نمایش دهید . همچنین نمودار پاداش تجمعی ان را نیز نمایش دهید.

۴. الگوریتم Sarsa را پیاده سازی کنید و مقادیر ارزش را نمایش دهید . این نتایج را به γ های ۰.۹ و ۰ نمایش دهید . همچنین نمودار پاداش تجمعی انرا نیز رسم کنید .

۵. چه تفاوتی بین الگوریتم های Sarsa , Q-Learning وجود دارد .