

«پروژه پایانی»

درس یادگیری تعاملی - زمستان ۱۴۰۰

در این پروژه، هدف آشنایی با مباحث و چالش‌های جدید در حوزه یادگیری تعاملی است. برای انجام پروژه پایانی، شما می‌توانید یکی از دو گزینه زیر را انتخاب نمایید.

1. ارائه یک مسئله از دنیای واقعی، مدل‌سازی و حل آن با استفاده از استراتژی‌های یادگیری تعاملی.

2. ساخت یک عامل کارآمد برای محیط HardRacetrack با استفاده از روش‌های یادگیری تعاملی.

در صورت انتخاب گزینه اول، شما باید ابتدا در طی یک ایمیل به دکتر نیلی، موضوع خود را به همراه یک توضیح مختصر از صورت مسئله اعلام کنید و پس از تایید، اقدام به پیاده‌سازی راه‌حل خود کنید. در صورت انتخاب گزینه دوم، شما باید یک عامل برای محیط HardRacetrack توسعه دهید بصورتیکه بتواند بهترین عملکرد را از خود نشان بدهد. دقت کنید که در هر دو گزینه، کیفیت راه‌حل پیاده‌سازی شده، گزارش کار و ارائه نهایی شما هر کدام حائز اهمیت بوده و در میزان نمره شما از این پروژه تاثیرگذار هستند.

در ادامه توضیحاتی راجع به محیط HardRacetrack ارائه می‌شود.

محیط HardRacetrack

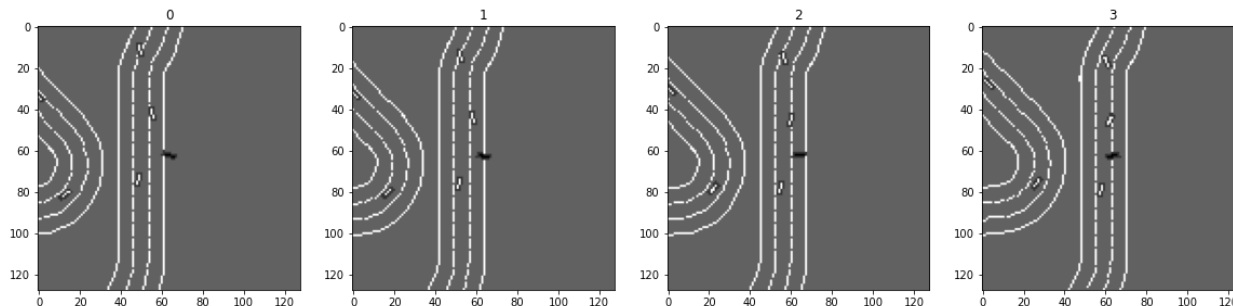
این محیط با استفاده از پکیج highway-env ساخته شده است.^۱ این محیط در واقع یک شبیه‌سازی رانندگی دوبعدی است. در این مسیر حلقه‌بسته، تعدادی ماشین در کنار ماشین شما در حال حرکت هستند. بطور کلی شما سه هدف اصلی دارید:

- ماندن در مسیر
- حرکت در میان خطوط
- پرهیز از برخورد با دیگر ماشین‌ها

Interface کار کردن با این محیط مشابه دیگر محیط‌های ارائه شده توسط پکیج gym هست. یعنی با فراخوانی متد step(action) شما ۴ خروجی observation، reward، done و info را بدست می‌آورید. Observation در این محیط بصورت یک دسته ۴ تایی از چهار مشاهده آخر است.^۲ هر مشاهده شامل یک تصویر Grayscale به اندازه 128*128 از بخشی محدود از مسیر مسابقه است. همانطور که در تصاویر زیر دیده می‌شود، عامل همواره در وسط تصویر قرار دارد اما ممکن است جهت ماشین به هر طرفی باشد.

^۱ برای راهنمای نصب این پکیج به آدرس <https://highway-env.readthedocs.io/en/latest/installation.html> مراجعه کنید.

^۲ در صورت لزوم، شما می‌توانید این مقدار را تغییر دهید. شما مجاز به تغییر بقیه پارامترهای محیط نیستید.



فضای action در این محیط شامل دو بعد است. بعد اول مشخص کننده شتاب ماشین (و نه سرعت) است که مقداری بین -1 و $+1$ دارد. همچنین بعد دوم مربوط به زاویه فرمان است که این مقدار نیز عددی بین -1 و $+1$ است.

به همراه این سند، یک پوشه شامل اسکریپت‌های مربوط به محیط و یک نوت‌بوک شامل کدی ساده برای راه‌اندازی محیط قرار داده شده است.

لطفاً به نکات زیر دقت فرمایید:

- **این پروژه کاملاً اختیاری است و انجام آن، نمره امتیازی برای شما خواهد داشت.** البته میزان نمره امتیازی به کیفیت راه‌حل، گزارش و ارائه شما بستگی دارد.
- حتماً برای ارائه راهکارهای خود به روش‌های بروز و جدید در حوزه یادگیری تعاملی رجوع کنید.
- شما می‌توانید بصورت تک‌نفره و یا حداکثر دونفره این پروژه را انجام بدهید.
- شما تا ۵ روز پس از آخرین روز آزمون‌های پایان‌ترم برای ارسال پاسخ‌ها فرصت دارید.
- پاسخ‌های شما باید شامل کدها و همچنین یک فایل PDF از گزارش کار شما باشد.
- به غیر پارامتر اندازه دسته مشاهدات در محیط، شما مجاز به تغییر هیچ‌کدام از پارامترهای محیط نیستید.
- با توجه به فشردگی کار، پیشنهاد می‌کنم هرچه زودتر دست بکار شوید.