



دانشکده مهندسی کامپیوتر
دانشگاه اصفهان

مستند پروژه یادگیری تقویتی آزمایشگاه رباتیک

دکتر حمیدرضا برادران

تیر ۱۴۰۱

مقدمه:

در این پروژه قصد داریم در ابتدا توانایی‌های کلی یادگیری تقویتی نظیر شناخت محیط و مشخص کردن کنش‌ها و عامل و ... را مورد سوال قرار دهیم؛ سپس به سراغ پیاده‌سازی یک محیط بازی در یادگیری تقویتی برویم و در انتها با کتابخانه OpenAI GYM آشنا شده و یکی از محیط‌های آن را بررسی نماییم.

نمره کلی این پروژه ۹ نمره می‌باشد که به تفضیل هر کدام از بخش‌های آن در ادامه بررسی می‌گردد.

قسمت اول:

فرض کنید به شما یک تسک برای انجام می‌دهند که به صورت زیر می‌باشد:

از شما خواسته شده که دو ربات موبایل را به سمت هدف‌اش راهنمایی نمایید. این دو ربات به صورت همکار می‌باشند و رقابتی با یکدیگر ندارند اما امکان دارد با هم برخورد نمایند. در راه مانع‌هایی وجود دارد که ممکن است به آن برخورد نماید و هدف آن کاملاً مشخص و از پیش تعیین شده است. نکته‌ای که وجود دارد این است که با توجه به نوع موتور آن حرکات این ربات به صورت پیوسته می‌باشد (دامنه کنش‌های گسسته نیست). در صورتی که بخواهید به کمک یادگیری تقویتی این تسک را حل نمایید چه راه‌حلی در پیش خواهید گرفت؟

لطفاً به سوالات زیر پاسخ دهید.

- در این جا عامل شما چه چیزی می‌باشد؟
- با توجه به اینکه مجموعه کنش‌های عامل پیوسته می‌باشد چه راه‌حلی برای گسسته کردن آن دارید؟
- پس از گسسته‌سازی مجموعه کنش‌ها پیرامون علت انتخاب این مجموعه توضیح داده و براساس این انتخاب دامنه کنش خود را تعریف نمایید.
- بخش بسیار مهم این تسک انتخاب تابع مناسب برای پاداش می‌باشد؛ تابع پاداش خود را مشخص کرده و پیرامون منطق آن توضیح دهید.
- در صورتی که این دو ربات به جای همکاری رقابت می‌کردند چه تغییری در تابع پاداش و همینطور مجموعه کنش‌ها ایجاد می‌گشت؟

قسمت دوم:

در این تسک هدف ما این است که یک محیط مبتنی بر یادگیری تقویتی داشته باشیم که در آن بتوانیم عامل‌های خود را آزمایش نماییم. با توجه به اینکه قصد نداریم پروژه را سنگین کنیم از قرار دادن عامل و تست آن در محیط صرف نظر می‌کنیم. برای آشنایی بیشتر با این کار بیایید ویژگی‌های محیط را بررسی کنیم؛

(۱) هر محیطی چند ویژگی دارد که به آن‌ها پرداخته‌ایم؛ فرضا باید ما بتوانیم به کار اجازه انجام کنش‌ها را بدهیم.

(۲) به عنوان مثالی دیگر باید بتوانیم دامنه اکشن‌ها را در این فضا مشخص کنیم.

و موارد دیگری که می‌توان به اشاره کرد حرکت کاربر در صفحه بازی می‌باشد.

نکته قابل توجه این است که برای این پیاده‌سازی صرفا شناخت محیط کفایت می‌کند و می‌توانید برای ساده‌سازی از عدم قطعیت کنش‌ها در محیط صرف نظر کنید. هرچه جزئیات بیشتری در محیط پیاده‌سازی گردد، به دنبال آن نمره گرفته شده از این تسک بیشتر خواهد شد.

برای پیاده‌سازی هیچ نیازی به کتابخانه‌های خاص نداشته و صرفا هدف آشنایی هرچه بیشتر با محیط‌های موجود می‌باشد.

سورس کد پروژه را به همراه فایل‌های دیگر ارسال فرمایید.

قسمت سوم:

در این تسک هدف ما این است که با ابزارهای یادگیری تقویتی آشنا شویم؛ برای آغاز این کار بهتر است به سراغ بررسی gym رفته و آن‌ها را بررسی نماییم.

در ابتدا قصد داریم محیط Frozen Lake را بررسی نموده و جزئیات آن را بررسی کنیم و سپس به سراغ محیط کارت پل می‌رویم (Cart pole). به عنوان تسک اول ابتدا این دو محیط را بررسی کرده و جزئیات آن را گزارش دهید.

به عنوان تسک دوم محیط OpenAI GYM را نصب کرده و به کمک داکيومنت سایت

OpenAI محیط‌های دو مورد بالا را ساخته و با جزئیات آن را از روی کد توضیح دهید. توضیحات ارائه شده در داک باید شامل نحوه ساخت محیط و همین‌طور تسک و هدف محیط و همین‌طور فانکشن‌های آن باشد.

قسمت چهارم:

با توجه به توضیحات دو جلسه آخر، الگوریتم Policy iteration را پیاده‌سازی نمایید. نکته قابل توجه این است که باید در ابتدا کارکرد الگوریتم را توضیح داده و سپس اقدام به پیاده‌سازی آن نمایید.

این مورد را توجه داشته باشید که هدف از این تسک صرفاً پیاده‌سازی یک کد نمی‌باشد و هدف تسلط شما بر این الگوریتم است.

سورس کد پیاده‌سازی شده را به همراه فایل ارسالی تحویل دهید.

قسمت پنجم (اختیاری):

در این بخش قصد داریم که به عنوان یک تمرین اختیاری یک الگوریتم یادگیری تقویتی را روی محیط واقعی تست نمایید. برای این کار در ابتدا الگوریتم خود را (نوع آن اهمیتی ندارد) بر روی محیط Cart pole که در تمرین‌های قبلی بررسی کرده‌ایم پیاده‌سازی نموده و خروجی را گزارش دهید. در این قسمت هم هدف و بارم اصلی به فهمیدن درس کد و اجرای صحیح آن اختصاص دارد و مسلماً کد صرف نمره چندانی نخواهد داشت.

بارم‌بندی هر کدام از بخش‌های تمرین:

بخش اول) یک و نیم

بخش دوم) چهار نمره

بخش سوم) یک نمره

بخش چهارم) دو و نیم نمره

بخش اختیاری) یک نمره

نکات مربوط به تحویل پروژه:

- تحویل پروژه به صورت ارائه‌های افلاین می‌باشد که هرتیمی باید به صورت کامل بخش‌های مختلف پروژه را توضیح داده و ویدئو آن را به همراه موارد دیگر خواسته شده ارسال نمایند.

- آخرین مهلت ارسال پروژه تا ساعت ۲۲ پنجشنبه نهم تیر ماه می باشد و این زمان به هیچ وجه تمدید نخواهد شد.
- پروژه را می توانید در قالب تیم های دو نفره یا فردی انجام دهید. در صورتی که قصد دارید به صورت گروهی این پروژه را تکمیل نمایید لطفا یکی از اعضا فایل را ارسال کرده و نام هردونفر در داک پروژه باشد.
- مواردی که باید به عنوان فایل ارسالی تحویل گردد:

- داکيومنت پروژه

- سورس کدهای هربخش

- ویدئوهای ارسالی