

به نام خدا



دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکده برق و کامپیوتر



درس سیستم‌های هوشمند

پروژه‌ی پایانی

بهمن ۱۴۰۰

فهرست سوالات

۳ مقدمه
۳ مدل سازی
۴ مسیریابی ربات ها
۵ پیاده سازی مسئله
۶ نکات:

مقدمه

در این پروژه سعی بر این داریم که ابتدا برای یک سری مسائل در دنیای واقعی، یک مدل یادگیری تقویتی طراحی کنیم. در قسمت بعد، به بررسی و پیاده‌سازی یک مسئله یادگیری تقویتی برای یک ربات بپردازیم، که کاربردهای مختلفی در زمینه‌های مخابراتی جهت طراحی UAV و ... کاربرد دارد.

مدل‌سازی

برای هر کدام از مسائل زیر مدلی برای حل ارائه دهید. لازم به ذکر است برای مدل نیاز دارید استیته‌ها، اکشن‌های انتقال بین استیته‌ها به صورت احتمالی و پاداش را تعیین کنید.

۱. در یک دهکده قصد داریم در ابتدای هر ماه تصمیم بگیریم که فروش خرچنگ مجاز است یا خیر. هر سری که تصمیم به فروش خرچنگ گرفته شود مقداری از خرچنگ‌ها کاسته شده و مقداری سود از فروش خرچنگ‌ها به دست ما خواهد رسید. لازم به ذکر است که اگر جمعیت خرچنگ‌ها از حدی کمتر شوند، نیازمند هزینه زیادی برای جبران جمعیت آن‌ها هستیم، وگرنه کلاً صنعت فروش خرچنگ در این شهر ورشکسته می‌شود.
۲. در خرید سهام در آخر هر دوره باید تصمیم بگیریم که سهام خود را نگه داریم و یا آن را بفروشیم. فرض کنید که خرید و یا فروش سهام در هر دوره به طور ۷ بار انجام می‌شود و برای سادگی reset صورت می‌گیرد. همچنین در انتهای هر تصمیم با احتمالی به سود فرد اضافه شده و با احتمالی درصد از ارزش سهام خود را از دست می‌دهد.
۳. در یک کارخانه قصد داریم با توجه به تقاضای سال گذشته میزان تولید سال آینده را تعیین کنیم. توجه کنید که هر مشتری که درخواست کالا داشته باشد و کالا تولید کافی نداشته باشد، علاوه بر ضرر، آن مشتری را کامل از دست می‌دهیم که برای ما ضرری اضافه‌تر نیز دارد. همچنین، کالاهایی که تولید شدند، اما خریداری نشد نیز ضرر دارند.
۴. در یک آتش‌نشانی برای هر هشدار آتش که از هر منطقه (مسکونی، اداری، تجاری و ...) می‌بایست تصمیم بگیریم از تعداد ماشین‌های آتش‌نشانی موجود چند ماشین ارسال شوند. با توجه به نوع آن منطقه و بلندی و کوتاهی آژیر، تعداد ماشین نیازمند به ارسال را تعیین کنید. در نظر داشته باشید که ارزش انسان‌ها از ساختمان‌ها بیشتر است و در تعیین ارزش منطقه‌ها باید پیش فرض جمعیتی آن منطقه را بیان کنید. همچنین، اگر آژیر زده شود و ماشینی در آتش‌نشانی نباشد، ضرر زیادی خورده و برای مناطق پرجمعیت می‌بایست تعدادی بیشتر ماشین برای نجات انسان‌ها ارسال شوند.

مسیریابی ربات‌ها

مسئله مسیریابی برای ربات‌ها از مسائل مهم و اساسی برای ربات‌های امروزی می‌باشد. مواردی مانند عدم برخورد با موانع موجود و پیدا کردن بهترین مسیر در محیط داده شده از این قبیل مسائل می‌باشند. این مسائل را می‌توان در ترکیب با مسائل MDP حل کرد. محیط زیر را طبق توضیحات داده شده در نظر بگیرید و پس از پیاده‌سازی موارد خواسته شده، به سوالات هر بخش پاسخ داده و تحلیل خود را ارائه دهید.

یک محیط مشبک طبق شکل ۱-۲ زیر با ابعاد ۱۵ در ۱۵ در نظر بگیرید. ربات ما در ابتدا در نقطه (۱،۱)، قرار گرفته است. هدف ربات این است که به خانه (۱۰،۱) برود. ربات مورد نظر ما در هر استیت قادر به انجام ۹ عمل مختلف می‌باشد. ۸ عمل برای جابجایی در جهت‌های ۸ گانه و یک عمل برای باقی‌ماند در نقطه فعلی. برای ربات دو سری مجموعه استیت داریم. مجموعه استیت‌های قابل دسترسی و مجموعه استیت‌های ممکن. همسایه‌ی ممکن همسایه‌ای است که خارج از محدوده محیط و مانع نباشد. همسایه‌ی در دسترس همسایه‌ای است که با یکی از اکشن‌های ممکنه بتوان به آن رسید.

	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵
۱															
۲															
۳															
۴															
۵															
۶															
۷															
۸															
۹															
۱۰															
۱۱															
۱۲															
۱۳															
۱۴															
۱۵															

شکل ۱-۲: محیط سوال

در انجام هر اکشن، ربات با احتمال p به جهت انتخابی می‌رود، و در غیر اینصورت به یکی از همسایگان ممکن و در دسترس، لیز می‌خورد. دقت شود که احتمال لیز خوردن به همه خانه‌ها به صورت یکنواخت و یکسان است. برای انجام هر حرکت به خاطر وجود انرژی مصرف شده و زمان تلف شده پاداش منفی در نظر

گرفته شده است. همچنین، هنگام برخورد با مانع، هزینهی برخورد با مانع نیز برای ربات در نظر گرفته شده است. همچنین، در هنگام رسیدن به هدف ربات پاداش دریافت می‌کند.

حالت‌های زیر حالت‌های ممکن در محیط می‌باشند:

- احتمال انجام اکشن و رفتن به استیت بعدی برابر 0.8
- احتمال لیز خوردن ربات و رفتن به یک خانه ممکن و در دسترس یا ماندن در استیت فعلی برابر 0.2
- هزینه برخورد با مانع 1 -
- برای انجام هر حرکت به خاطر از دست رفتن وقت و انرژی، پاداش منفی برابر با -0.01
- پاداش رسیدن به خانه هدف برابر با 1000

حالت حرکت بدون هزینه: همانند حالت پایه می‌باشد، با این فرق که هزینه حرکت برابر با 0 در نظر گرفته شود. در صورت برخورد با مانع پاداش منفی 0.01 برای ربات در نظر گرفته شود.

حالت حرکت با هزینه زیاد: همانند حالت پایه می‌باشد، با این تفاوت که هزینه هر حرکت برابر با 1 -) برای همه اکشن‌ها به جز اکشن ماندن در خانه)، هزینه برخورد با مانع برابر با 10 - و پاداش رسیدن به خانه هدف را برابر با 100 در نظر بگیرید.

پیاده‌سازی مسئله

۱. شبه کدی برای روش مونت کارلو جهت به دست آوردن ارزش استیت‌های محیط، پیاده‌سازی کنید.
۲. سیاست بهینه را برای حالت پایه با استفاده از روش تکرار تصمیم (سیاست)^۱ به دست آورید.
۳. سیاست بهینه را با روش تکرار تصمیم برای حالت بدون اصطکاک به دست آورده و با نتایج قسمت دو مقایسه کنید.
۴. حالت با اصطکاک زیاد را در نظر گرفته و سیاست بهینه را با استفاده از روش تکرار تصمیم به دست آورده و با دو حالت قبل مقایسه کنید.
۵. با توجه به نتایج قسمت قبل، مقدار بهینه برای پاداش محیط را در نظر گرفته و نقش تفاوت مقدارهای مختلف ضریب کاهشی^۲ را برای مقادیر مختلف بررسی کنید. آینده نگری ربات را بررسی کنید.
۶. الگوریتم تکرار ارزش^۳ را برای محیط داده شده اجرا کنید و نتایج بدست آمده را با قسمت قبل مقایسه کنید.
۷. دلیل تفاوت در بخش‌های ۲ و ۳ و ۴ را بررسی کنید و راه حلی برای آن ارائه دهید.

^۱ policy iteration

^۲ discount factor

^۳ value iteration

نکات:

- مهلت تحویل گزارش کتبی این پروژه، پنج‌شنبه ۱۴ بهمن است.
- انجام این پروژه به صورت یک نفره یا گروه‌های دو الی سه نفره مجاز است.
- برای انجام پروژه فقط مجاز به استفاده از زبان های برنامه نویسی Python و MATLAB خواهید بود. در سؤالاتی که از شما خواسته شده است یک الگوریتم را پیاده سازی کنید **مجاز** به استفاده از توابع آماده **نمی‌باشید** مگر اینکه در صورت سوال مجاز بودن استفاده از این توابع یا کتابخانه‌ها صریح ذکر شده باشد.
- کدهای مربوط به هر بخش می‌بایست در پوشه‌ای با نام Codes در کنار گزارش کار شما موجود باشد. این کدها باید خوانا و به صورت مرتبط نام گذاری شده باشند، لذا توضیحات لازم را به صورت یادداشت^۱ در کدهای خود قرار دهید.
- لطفاً تمامی نکات و مفروضاتی که برای پیاده‌سازی‌ها و محاسبات خود در نظر می‌گیرید را در گزارش ذکر کنید. همچنین به رعایت موازین نگارشی در گزارش خود توجه ویژه‌ای داشته باشید (به طور مثال استفاده از زیرنویس برای تصاویر و بالانویس برای جداول).
- برای پروژه هر گروه (یا فرد) علاوه بر گزارش کتبی ملزم به ارائه گزارشی در قالب ارائه علمی خواهد بود، این ارائه از اهمیت ویژه‌ای در مراحل نمره‌دهی برخوردار است و تمام افراد گروه باید به تمام مباحث پروژه اشراف داشته باشند و به سوالات مطرح شده در ارائه پاسخ دهند.
- لطفاً گزارش، فایل کدها و سایر ضمائم مورد نیاز را با ترتیب نام‌گذاری زیر در صفحه درس در سامانه یادگیری الکترونیکی بارگذاری نمایید.

FinalPtoject_[StudentNumber(s)].zip

- در صورت وجود هرگونه ابهام یا مشکل می‌توانید از طریق رایانامه‌های زیر با دستیاران آموزشی مربوطه در تماس باشید:

جناب آقای نوری - رایانامه { salar.nouri@ut.ac.ir }

^۱ comment