

به نام خدا

تمرین سوم

درس یادگیری تعاملی

پاییز 99

محمد امین حق پناه (mdan.hagh@gmail.com) – شیما ربیعی (shimarabiee@ut.ac.ir)

سوال 1

هدف از این مساله آشنایی با الگوریتم Policy Iteration و یادگیری تقویتی در فضای چند حالت است.

دو شرکت ارائه دهنده ی خدمات انتقال دیتا به صورت ماهانه به مشتریان خود دیتا ارائه می دهند و به ازای هر مشتری مبلغ 10 را دریافت می کنند. بمنظور سادگی در حل مساله ظرفیت کل هر شرکت را 20 واحد در نظر می گیریم. این دو شرکت می توانند تا تعداد حداکثر 5 واحد ظرفیت را هر ماه از شرکت رقیب خود خریداری و مبلغ 2 را به ازای هر واحد به آن شرکت پرداخت نمایند. در هر ماه تعداد مشتریان متقاضی دریافت دیتا و تعداد مشتریانی که به قرارداد خود خاتمه می دهند، با یک متغیر تصادفی پواسن مدل می شود. به این معنا که احتمال تعداد n مشتری برابر با $\frac{\lambda^n}{n!} e^{-\lambda}$ می باشد که λ میانگین تابع توزیع است. اگر λ را برای مشتریان متقاضی در شرکت A و B به ترتیب 3 و 4 و برای مشتریانی که به قرارداد خود خاتمه می دهند به ترتیب 3 و 2 در نظر بگیریم و با فرض اینکه بیشترین ظرفیت هر شرکت 20 باشد، عاملی طراحی نمایید که با الگوریتم Policy Iteration سیاست بهینه ای برای تعداد انتقال ظرفیت بین دو شرکت با توجه به تقاضای مشتریان بدست آورد.

الف) ضریب تخفیف را یکبار 0.9 و بار دیگر 1 قرار دهید.

ب) مبلغ خرید ظرفیت از شرکت رقیب را 6 در نظر بگیرید و سیاست بهینه را تحلیل نمایید.

سوال ۲

هدف از این مساله طراحی یک مدل MDP است.

یک کارخانه‌ی تولید مواد غذایی قصد دارد تا با استفاده از روش‌های هوش مصنوعی تولیدش را بهینه‌تر کرده و به فروش و سود بیشتری دست پیدا کند. در نهایت هدف این است که یک عامل کامپیوتری مدیریت خرید مواد اولیه و تبدیل آن‌ها به مواد غذایی و فروش را انجام دهد.

این کارخانه قصد دارد تا تعدادی متخصص را گرد هم آورده و مسئله را به صورت حدودی و نه دقیق برای آن‌ها شرح دهد، چرا که تمام جزئیات مسئله برای خود آن‌ها هم هنوز مشخص نیست. شما نیز به عنوان یک متخصص در این گردهمایی حضور دارید و قصد دارید به صاحبان کارخانه نشان دهید که شایستگی بیشتری از سایر متخصصین برای حل این مسئله دارید!

فرض کنید این کارخانه توانایی تولید ۱۰۰ نوع ماده‌ی غذایی مختلف را داشته و برای تولید آن‌ها لازم است تا مجموعاً ۵۰۰ نوع ماده‌ی اولیه‌ی مختلف در اختیار داشته باشد. از ویژگی‌های مواد اولیه می‌توان به قیمت، ماندگاری، حجم، سهولت در تهیه (مثلاً تحریم باشد یا نه) و ... اشاره کرد. به کمک این مواد اولیه نیز می‌توان مواد غذایی متنوعی تولید کرد و به دست مشتری رساند. هر نوع ماده‌ی غذایی با استفاده از تعدادی ماده‌ی اولیه با حجم‌های متفاوت قابل تولید است و از ویژگی‌های آن می‌توان به قیمت و ماندگاری اشاره کرد. برای ساده‌سازی مسئله، زمان مورد نیاز برای تبدیل ماده‌ی اولیه به ماده‌ی غذایی نهایی را صفر در نظر بگیرید. بنابراین، کارخانه هر زمان که اراده کند، در صورت داشتن ماده‌ی اولیه به مقدار کافی می‌تواند یک ماده‌ی غذایی را در لحظه تولید کند.

کارخانه مواد اولیه را تنها در روز اول ماه می‌تواند بخرد و اگر وسط ماه متوجه شود که به یک ماده‌ی اولیه نیاز جدی دارد، مجبور است تا روز اول ماه بعدی صبر کند. این کارخانه محصولات خود را بدون واسطه و به صورت اینترنتی به مردم می‌فروشد. در هر لحظه مردم می‌توانند وارد سایت کارخانه شده و محصولات موجود را ببینند، سپس سفارش خود را ثبت کنند. برای سادگی، زمان رساندن محصول به مشتریان پس از ثبت سفارش را صفر در نظر بگیرید. همچنین فرض کنید تمامی محصولات کارخانه دارای کیفیت عالی بوده و عدم رضایت مشتری تنها مربوط به موجود بودن یا نبودن مواد غذایی مورد نیاز است.

درخواست‌ها نیز طبیعتاً در ماه‌های مختلف متفاوت بوده و بعضی از مواد غذایی تنها در بعضی از ماه‌های سال مشتری دارند. در واقع نیازهای مشتریان در طول زمان متغیر است.

برای حل این مسئله به کمک یادگیری تعاملی، لازم است تا ابتدا یک مدل برای آن ارائه دهیم، سپس با استفاده از این مدل مسئله را حل کنیم. یک مدل MDP برای این مسئله طراحی کنید و حالت‌ها، عمل‌ها، نحوه‌ی انتقال از یک حالت به حالت دیگر و پاداش دریافتی را مشخص کنید و در نهایت توضیح دهید چرا این مدل MDP است.

توجه: این سوال یک پاسخ مشخص ندارد و به صورت مبهم طراحی شده تا به مسائل موجود در دنیای واقعی نزدیک‌تر باشد. هر فرضی که فکر می‌کنید لازم است را ذکر کنید و با استفاده از آن مدل خود را پیشنهاد دهید.

ملاحظات:

- در صورت ابهام یا سوال در مورد تمرین، آنها را در فروم Q&A بنویسید. دستیاران آموزشی در اسرع وقت به سوالات شما پاسخ می‌دهند. در این صورت بقیه دانشجویان هم از این پرسش و پاسخ استفاده خواهند کرد.
- برای حل سوال اول باید از زبان Python 3.x استفاده کنید.