



### سوال ۱. مسئله‌ی مسیریابی کارمندطور

هدف از این مسئله آشنایی با الگوریتم‌های Q-Learning و SARSA در فضای چندحالتی است.

کارمند یک شرکت خصوصی هر روز صبح با ماشین شخصی از خانه‌ی خود (میدان آزادی) تا محل کار خود (میدان تجریش) رانندگی می‌کند. با توجه به ترافیک صبحگاهی و محدوده‌ی طرح ترافیک و زوج و فرد، تعداد مسیرهایی که می‌تواند تا محل کار انتخاب کند محدود است. مسیرهای قابل انتخاب را می‌توانید در نقشه (که در بخش نکات لینک آن آمده است) مشاهده کنید. میزان سوخت مصرفی و زمان صرف‌شده برای طی یک مسیر، به عوامل مختلفی از جمله طول مسیر، میزان شیب مسیر، میزان ترافیک و... بستگی دارد. میزان سوخت مصرفی، میزان زمان لازم برای طی یک مسیر و مقصدهایی که از یک محل می‌توانید به آن بروید، همچنین حالت ابتدایی و حالت هدف، در یک کتابخانه پیاده‌سازی و در اختیار شما قرار داده شده است.

**الف)** در روزهای شنبه، فرد باید بچه‌های خود را به مدرسه‌ای که در نزدیکی منزلش است برساند (برای رساندن بچه‌ها از ماشین استفاده نمی‌کند). پیاده به مدرسه رفته، به خانه برمی‌گردد و با ماشین به محل کار می‌رود. و سپس به سمت شرکت حرکت کند. به همین دلیل زمان کمتری برای رسیدن به شرکت دارد. مسیری که در کمترین زمان ممکن می‌توان در روزهای شنبه از منزل به شرکت رسید را مشخص کنید.

**ب)** شرکت در انتهای ماه، به کارمندی که در ۱۰ روز انتهای ماه، کمترین مصرف بنزین را داشته باشند، پاداش می‌دهد. در نتیجه در این ۱۰ روز این کارمند سعی می‌کند مصرف بنزین خود را کمینه کند. مسیر بهینه برای رسیدن از خانه به شرکت که کمترین میزان بنزین در آن مصرف می‌شود را مشخص کنید.

**پ)** کارمند می‌خواهد مسیری را پیدا کند که تابع زیر را کمینه کند. این مسیر را پیدا کنید.

$$C = Fuel + Time^2$$

- این مساله را در هر سه حالت با استفاده از الگوریتم‌های SARSA و Q-Learning حل کنید.
- برای هر قسمت، مسیر نهایی را گزارش کنید. آیا مسیر به دست آمده مسیر بهینه است؟
- به نظر شما چه تفاوتی بین این دو الگوریتم از منظر میزان پشیمانی و تعداد اپیزودها تا همگرایی وجود دارد؟

## نکات:

1. نقشه‌ی مسیرها در این [لینک](#) آمده است.
2. برای استفاده از کدهای آماده که در اختیار شما قرار داده شده است، نیاز به کتابخانه‌ی [networkx](#) دارید.
3. در کلاس Mapbuilder تعدادی تابع تعریف شده اند که به صورت زیر هستند. برای پیاده سازی الگوریتم خود، از این توابع استفاده کنید. در این کلاس، نقشه پیاده سازی شده و هر مسیر با نقاط ابتدایی و انتهایی آن مسیر مشخص می‌شود.

<b>next_State(state)</b>	یک حالت را به عنوان ورودی گرفته و حالات بعدی این حالت را باز می‌گرداند.
<b>get_Reward(start,end)</b>	دو حالت ابتدایی و انتهایی را دریافت کرده و لیستی شامل دو عنصر باز می‌گرداند که به ترتیب میزان بنزین مصرفی و زمان صرف شده برای طی مسیر در آن روز است. این مقادیر وابسته به یک توزیع نرمال با واریانس ثابت برای همه‌ی مسیرها هستند. اما میانگین این توزیع به مسیر طی شده وابسته است.
<b>initial_state()</b>	حالت ابتدایی را باز می‌گرداند.
<b>terminal_state()</b>	حالت هدف را باز می‌گرداند.

لطفا به نکات زیر توجه کنید:

- ✓ حجم گزارش شما به هیچ وجه معیار نمره‌دهی نیست، پس لطفا در حد نیاز توضیح دهید.
- ✓ تایپ کردن تمرین‌ها اجباری نیست ولی در صورتی که روی کاغذ می‌نویسید علاوه بر آپلود اسکن در صفحه‌ی درس، برگه‌ی خود را در اولین کلاس درس پس از ددلاین به استاد تحویل دهید.
- ✓ سعی کنید از پاسخ‌های روشن در گزارش خود استفاده کنید و اگر پیش‌فرضی در حل سوال در ذهن خود دارید، حتما در گزارش خود آن را ذکر کنید.
- ✓ از نمودارهای واضح در گزارش خود استفاده کنید، نمودارهایی که دارای لیبل‌گذاری روشن روی هر محور و همین‌طور توضیح مناسب باشد.
- ✓ کدهایی که به همراه گزارش تحویل می‌دهید باید قابل اجرا باشد. همچنین توجه کنید که به تمرین بدون گزارش نمره‌ای تعلق نمی‌گیرد.
- ✓ لطفا در گزارش و کدهای خود از تمرین دیگران استفاده نکنید، مشورت و همفکری در مورد سوال‌ها اشکالی ندارد اما اگر شباهت بیش از اندازه در تمرین‌ها دیده شود منجر به صفر شدن نمره خواهد شد.
- ✓ تمام فایل‌ها را در قالب یک فایل zip یا rar در سایت درس بارگذاری کنید.
- ✓ برای پیاده‌سازی تمرین فقط از زبان‌های MATLAB و یا Python می‌توانید استفاده کنید.

موفق و سلامت باشید. (: