



### 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؟