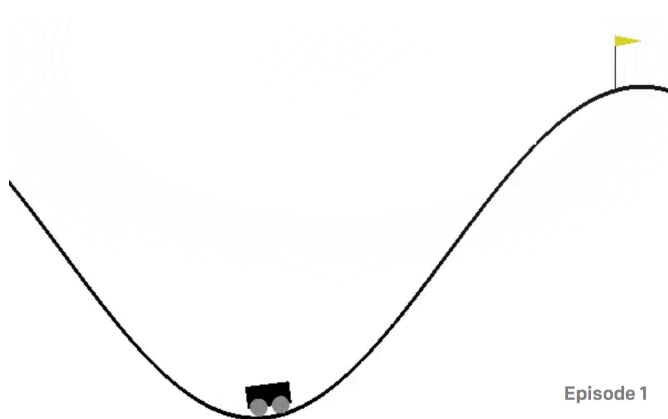




سوال ۱. بالا رفتن از تپه

یک ارابه میان دو کوه گیر افتاده است و هدف آن این است که به قله‌ی کوه سمت راست برسد. موتور این ارابه این قدرت را ندارد که به صورت مستقیم از این تپه بالا برود و باید با انجام رفت و برگشت این کار را انجام دهد. در صورتی که ارابه با صرف انرژی کمتر به بالای تپه برسد، پاداش بیشتری دریافت خواهد کرد. همان‌طور که مشاهده می‌شود محیط این مسئله یک محیط پیوسته است و لازم است که برای انجام یادگیری، تخمین ارزش حالت‌ها را با استفاده از تعمیم^۱ انجام دهیم.

- یادگیری این مسئله را به صورت Value-Based با کدینگ الف) ترکیب RBF و خطی (۴۰ نمره) ب) Fuzzy (۷۰ نمره) مدل کنید.
- در قسمت الف در صورتی که فقط از RBF استفاده کنیم از منظر پشیمانی^۲ چه تغییری می‌کند؟ (۱۰ نمره)
- قسمت الف را با یکی از روش‌های Policy Gradient و با در نظر گرفتن Baseline مناسب پیاده‌سازی کنید. (۸۰ نمره)



برای این تمرین نیازی به تهیه‌ی گزارش نیست و زمانی متعاقبا به شما اعلام خواهد شد که این تمرین را حضوری به دستیاران آموزشی تحویل دهید. این تمرین دو برابر سایر تمرین‌ها نمره خواهد داشت.

توضیحات پیاده‌سازی:

برای پیاده‌سازی محیط این مسئله می‌توانید از ابزار Gym که شرکت OpenAI آن را توسعه داده است، استفاده کنید. جهت مشاهده‌ی مستندات این ابزار می‌توانید به لینک زیر مراجعه کنید:

<https://gym.openai.com/docs>

در این تمرین مجاز هستید از کتابخانه‌های موجود استفاده کنید.

^۱ Generalization

^۲ Regret