

به نام خدا



درس هوش مصنوعی و سیستم‌های خبره

---

## تمرین یازدهم

---

مدرس درس:  
جناب آقای دکتر محمدی

طراحان:  
سهیل حمزه بیگی  
حامد فیض آبادی

مهلت ارسال: ۱۴۰۱/۱۰/۷

## سوال ۱

فرض کنید در يك بازی ریختن تاس شرکت کرده اید که هزینه هر بار ریختن تاس در آن ۱ سکه است و احتمال آمدن تمام اعداد در تاس با یکدیگر برابر است. شما پس از ریختن تاس به اندازه عدد روی تاس سکه دریافت می کنید. قانون بازی به این شکل است که شما موظف هستید در بار اول يك تاس بریزید، اما در سایر مراحل دو انتخاب دارید: \* اتمام بازی: با این حرکت شما به اندازه عدد روی تاس سکه دریافت می کنید. \* تاس ریختن: يك سکه هزینه می کنید و بار دیگر تاس می ریزید. لذا بازی را می توان به این صورت در نظر گرفت که بازیکن در ابتدای بازی در حالت شروع قرار دارد و در حالت شروع فقط حرکت ریختن تاس وجود دارد. در سایر حالات يك حرکت اتمام بازی وجود دارد که بازیکن را به حالت پایانی می برد و نمایش داده می شود که بدین معنی است که عدد  $s_i$  در حالت پایانی حرکتی وجود ندارد. هر حالت بین شروع و پایان با در تاس آمده است.  $i$  با توجه به توضیحات فوق به سوالات زیر پاسخ دهید:

۱. فرض کنید  $Pi_i$  های زیر در ابتدا وجود دارد، ردیف  $v^{pi_i}$  را کامل کنید ( $y = 1$ )

| حالت       | $s_1$     | $s_2$     | $s_3$      | $s_4$      | $s_5$      | $s_6$      |
|------------|-----------|-----------|------------|------------|------------|------------|
| $\pi_i$    | تاس ریختن | تاس ریختن | اتمام بازی | اتمام بازی | اتمام بازی | اتمام بازی |
| $v^{pi_i}$ |           |           |            |            |            |            |

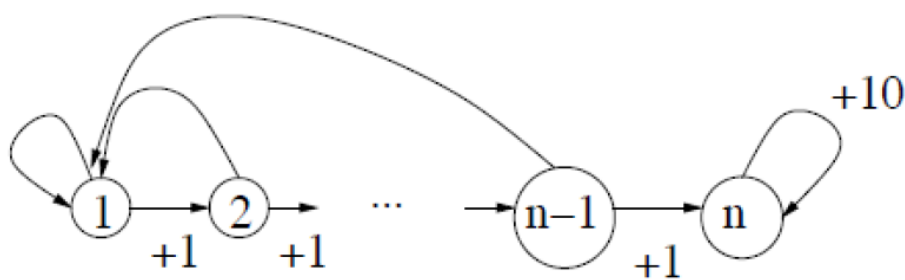
۲. با توجه به جدول فوق مقادیر  $pi_i$  را بروزرسانی کنید و در جدول زیر جایگذاری کنید. این مقادیر می تواند سه حالت تاس ریختن، اتمام بازی و تاس ریختن / اتمام بازی باشد. ( $y = 1$ )

| حالت        | $s_1$     | $s_2$     | $s_3$      | $s_4$      | $s_5$      | $s_6$      |
|-------------|-----------|-----------|------------|------------|------------|------------|
| $\pi_i$     | تاس ریختن | تاس ریختن | اتمام بازی | اتمام بازی | اتمام بازی | اتمام بازی |
| $\pi_{i+1}$ |           |           |            |            |            |            |

با توجه به مقادیر جدول فوق آیا می توان نتیجه گرفت که مقادیر بدست آمده بهینه هستند و دیگر نیاز به بروزرسانی ندارند؟ توضیح دهید.

## سوال ۲

تصویر زیر را در نظر بگیرید



در استیت  $n$  خارج می‌شویم و امتیاز  $+10$  را کسب می‌کنیم در بقیه ی استیت ها می‌توانیم عملیات زیر را انجام بدهیم:

- حرکت آزادانه که می‌تواند به یک حرکت به سمت راست برود و یا  $reset$  شود و به استیت شماره ی یک برود
- امتیاز حرکت کردن یک و امتیاز  $reset$  شدن صفر است
- $\gamma = 1/2$

حال به سوالات زیر پاسخ دهید:

۱. راهبرد ( $policy$ ) بهینه چیست؟
۲. مقدار بهینه  $V^*(n)$  را محاسبه کنید.
۳. مقدار بهینه برای  $V^*(k)$  را به ازای  $k = 1, \dots, n-1$  محاسبه کنید.
۴. فرض کنید که  $value iteration$  این سوال به دست شما بود و شما با مقادیر نزدیک صفر شروع می‌کردید. نشان دهید به ازای  $iteration$  های ۱ و ۲ چه مقادیری نمی‌توانند صفر باشند.

### قوانین:

۱. تمرین ها به صورت فردی انجام شوند و حل گروهی تمرین ها مجاز نیست.
۲. نمره شما بر اساس گزارش راه طی شده برای حل مسئله و پاسخ صحیح خواهد بود لذا از هرگونه اطناب در گزارش پرهیز و به موارد خواسته شده به صورت کامل پاسخ دهید.
۳. برای تحویل تمرین یک فایل zip شامل گزارش حل سوالات، با نام [HW11\_ID\_NAME] در سامانه LMS بارگذاری کنید.