

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 بازي ريختن تاس شركت كرده ايد كه هزينه هر بار ريختن تاس در آن ۱ سكه است و احتمال آمدن تمام اعداد در تاس با يكديگر برابر است. شما پس از ريختن تاس به اندازه عدد روي تاس سكه دريافت مي كنيد. قانون بازي به اين شكل است كه شما موظف هستيد در بار اول يك تاس بريزيد، اما در ساير مراحل دو انتخاب داريد: \* اتمام بازي: با اين حركت شما به اندازه عدد روي تاس سكه دريافت مي كنيد. \* تاس ريختن: يك سكه هزينه مي كنيد و بار ديگر تاس مي ريزيد. لذا بازي را مي توان به اين صورت در نظر گرفت كه بازيكن در ابتداي بازي در حالت شروع قرار دارد و در حالت شروع فقط حركت ريختن تاس وجود دارد. در ساير حالات يك حركت اتمام بازي وجود دارد كه بازيكن را به حالت پاياني مي برد و نمايش داده مي شود كه بدین معني است كه عدد  $s_i$  در حالت پاياني حركتي وجود ندارد. هر حالت بين شروع و پايان با در تاس آمده است.  $i$  با توجه به توضيحات فوق به سوالات زير پاسخ دهيد:

۱. فرض كنيد  $Pi_i$  هاي زير در ابتدا وجود دارد، ردیف  $v^{pi_i}$  را كامل كنيد ( $y = 1$ )

حالت	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$	$s_6$
$\pi_i$	تاس ريختن	تاس ريختن	اتمام بازي	اتمام بازي	اتمام بازي	اتمام بازي
$v^{pi_i}$						

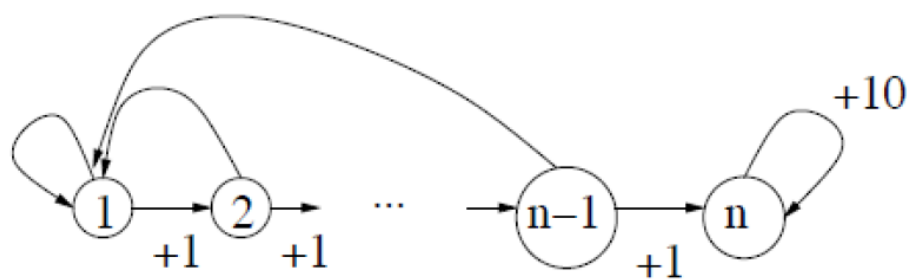
۲. با توجه به جدول فوق مقادير  $pi_i$  را برورساني كنيد و در جدول زير جايگذاري كنيد. اين مقادير مي تواند سه حالت تاس ريختن، اتمام بازي و تاس ريختن / اتمام بازي باشد. ( $y = 1$ )

حالت	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$	$s_6$
$\pi_i$	تاس ريختن	تاس ريختن	اتمام بازي	اتمام بازي	اتمام بازي	اتمام بازي
$\pi_{i+1}$						

با توجه به مقادير جدول فوق آيا مي توان نتيجه گرفت كه مقادير بدست آمده بهينه هستند و ديگر نياز به برورساني ندارند؟ توضيح دهيد.

## 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.