

به نام خدا

مدت زمان آزمون: ۱۲۰ دقیقه

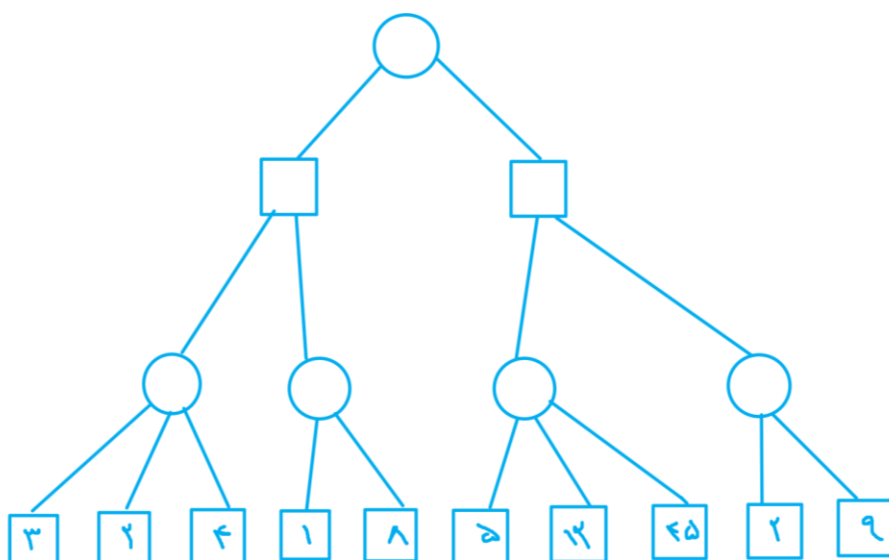
ترم اول ۱۴۰۰

آزمون پایانترم درس هوش مصنوعی

نام و نام خانوادگی:

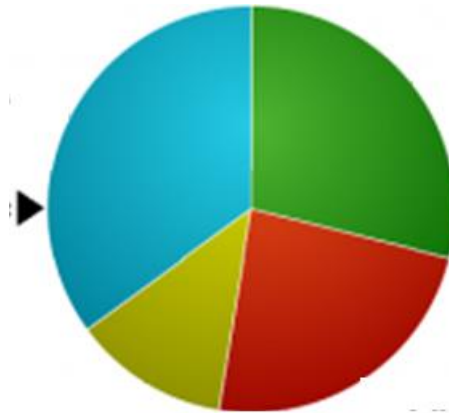
شماره دانشجویی:

سوال شماره یک) در درخت جستجوی زیر، گرههای دایره نشاندهنده بازیکن MIN و گرههای مربع بازیکن MAX هستند. ابتدا، مقادیر Minimax را برای هر یک از گرههای درخت بدست آورده و سپس این درخت را با استفاده از هرس آلفا-بتا، هرس کنید. (۵ نمره)



سوال شماره ۲: فرض کنید بازیکن حریف (MIN) در سوال شماره ۱، در ۲۰ درصد از مواقع، تصمیمات بهینه می گیرد و در ۸۰ درصد مواقع، از مکانیسم چرخ رولت استفاده می کند. در مکانیسم چرخ رولت، یک دایره به تعدادی ناحیه تقسیم می شود (تعداد نواحی برابر تعداد فرزندان گره است). ناحیه تخصیص یافته به هر فرزند، متناسب با معکوس اندازه آن است. به عبارت دیگر، هر چه مقدار minimax فرزند بزرگتر باشد، ناحیه کمتری به او تخصیص داده می شود. پس

از این تخصیص، چرخ رولت چرخانده می شود و مقدار متناظر با ناحیه ای که در مقابل هد قرار می گیرد، به عنوان مقدار گره انتخاب می شود. مقدار مورد انتظار هر یک از گرهها را مشخص کنید. (۴ نمره)



سوال شماره سه) شکل زیر نسخه ساده تری از مسئله GridWorld را نشان میدهد. این محیط دارای ۷ حالت (state) است که از ۱ تا ۷ شماره گذاری شده اند. حالت ۶ معادل حالت طلا بوده و در صورتی که عامل در آن قرار گیرد، از محیط خارج شده و جایزه +2 دریافت خواهد کرد. حالت ۷ نیز معادل حالت آتش بوده و در صورتی که عامل در آن قرار گیرد نیز از محیط خارج شده و جایزه -2 دریافت خواهد کرد. عامل می تواند در جهات بالا، پایین، چپ و راست حرکت کند. در صورتی که عامل تصمیم به حرکت در جهت خاصی داشته باشد، با احتمال ۷۰ درصد این حرکت موفقیت آمیز خواهد بود ولی با احتمال ۳۰ درصد عامل در جهت های قطری حرکت خواهد کرد. به عنوان مثال، فرض کنید عامل قصد حرکت از حالت ۴ به حالت ۵ (حرکت در جهت بالا) را داشته باشد. آنگاه عامل پس از انجام این عمل با احتمال ۷۰ درصد در حالت ۵، با احتمال ۱۵ درصد در حالت ۷ و با احتمال ۱۵ درصد در حالت ۴ (برخورد با دیوار و ماندن در حالت فعلی) خواهد بود. در صورتی که عامل در انجام عملی شکست بخورد، در همان حالت باقی خواهد ماند. به عنوان مثال، در صورتی که عامل قصد حرکت از حالت ۱ در جهت راست را داشته باشد، آنگاه با احتمال ۷۰ درصد به دیوار برخورد (شکست) و در حالت ۱ باقی خواهد ماند.

- با فرض استفاده از فاکتور تخفیف ۰.۷، مقادیر سودمندی هر یک از حالات هفتگانه را پس از ۳ مرحله اجرای الگوریتم value iteration محاسبه کرده و جهت های بهینه پس از این ۳ مرحله را برای هر حالت در شکل مشخص کنید. (living reward = -0.2) (۶ نمره)
- برنامه ای بنویسید که پاسخ نهایی این الگوریتم را محاسبه می کند. (برای ارسال کدهای برنامه، سه روز زمان خواهید داشت). (۵ نمره)

