



هوش مصنوعی

پاییز ۱۴۰۱

استاد: دکتر سلیمانی

دانشگاه صنعتی شریف

دانشکده مهندسی کامپیوتر

گردآورندگان: آرين احدی نیا، سعید هدایتیان

جستجوی تخصصی، فرآیند تصمیم‌گیری مارکوف

تمرین سوم

مهلت ارسال: ۲۷ آبان

- مهلت ارسال پاسخ تا ساعت ۲۳:۵۹ روز مشخص شده است.
- امکان ارسال با تأخیر پاسخ این تمرین تا سقف ۷ روز وجود دارد. پس از گذشت این مدت، پاسخ‌های ارسال شده پذیرفته نخواهند بود. همچنین، به ازای هر روز تأخیر غیر مجاز ۱۰ درصد از نمره تمرین به صورت ساعتی کسر خواهد شد.
- همکاری و هم‌فکری شما در انجام تمرین مانعی ندارد اما پاسخ ارسالی هر کس حتماً باید توسط خود او نوشته شده باشد.
- در صورت هم‌فکری و یا استفاده از هر منابع خارج درسی، نام هم‌فکران و آدرس منابع مورد استفاده برای حل سوال مورد نظر را ذکر کنید.
- لطفاً تصویری واضح از پاسخ سوالات نظری بارگذاری کنید. در غیر این صورت پاسخ شما تصحیح نخواهد شد.

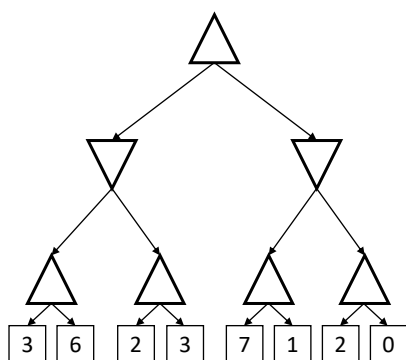
سوالات نظری (۶۵ نمره)

۱. (۱۲ نمره) آرين و سعید یک بازی بسیار خلاقانه کشف کرده‌اند و اکنون می‌خواهند آن را با یکدیگر بازی کنند. بازی به این صورت است که ابتدا یک عدد مثبت طبیعی بزرگتر از صفر انتخاب می‌کنند و سپس در هر مرحله، هر بازیکن می‌تواند عدد را یک یا دو واحد تقلیل دهد. بازیکنی که به صفر برسد برنده است.
 - (آ) (۴ نمره) برای مقدار اولیه ۵ درخت جستجوی تخصصی را رسم کنید و با حل آن مشخص کنید چه کسی در بازی برنده می‌شود.
 - (ب) (۴ نمره) برای مقدار اولیه ۶ درخت جستجوی تخصصی را رسم کنید و با حل آن مشخص کنید چه کسی در بازی برنده می‌شود.
 - (ج) (۴ نمره) نتیجه‌گیری کنید که در این بازی استراتژی برد چگونه است.
۲. (۱۲ نمره) در یک فرآیند تصمیم‌گیری مارکوف، n حالت داریم. در هر یک از حالات اول تا $n-1$ ام دو حرکت می‌توانیم انجام دهیم.

- اینکه با دریافت صفر امتیاز به حالت اول برگردیم.
 - اینکه با دریافت یک امتیاز، از حالت i ام به حالت $i+1$ ام برویم.
- در حالت n ام تنها حرکت مجاز این است که با دریافت ۱۰ امتیاز دوباره به همان حالت برگردیم. البته در این حالت در هر زمانی می‌توانیم فرآیند را به پایان برسانیم.

- (آ) (۶ نمره) سیاست بهینه در هر حالت را بیابید و بهینه بودن آن را اثبات کنید.
- (ب) (۶ نمره) مقدار V^* هر حالت را با ضریب تخفیف $\frac{1}{4}$ بدست آورید.

۳. (۱۲ نمره) آراین و سعید می‌خواهند یک بازی متفکرانه انجام دهند. آراین این بازی را فوت آب است (!) اما سعید این بازی را به تازگی یاد گرفته است برای همین آراین به سعید پیشنهاد می‌دهد تا در این بازی از کمک محمدرضا استفاده کند. سعید نیز (از خدا خواسته) سریعاً پیشنهاد آراین را قبول می‌کند غافل از اینکه محمدرضا عامل نفوذی آراین است! بازی به این صورت است که هر یک از دو طرف به نوبت بازی می‌کنند و در نهایت برنده به اندازه امتیازش دلار برره از بازنده دریافت می‌کند. توجه کنید که بازی از آراین شروع می‌شود. آنچه عامل فساد این بازی است این است که زمانی که نوبت سعید است، آراین می‌تواند d دلار برره به محمدرضا بدهد تا محمدرضا به دستور آراین بازی کند. فرض کنید که اگر بازی به صورت سالم انجام شود، درخت جستجوی تخصصی آن به صورت زیر باشد.

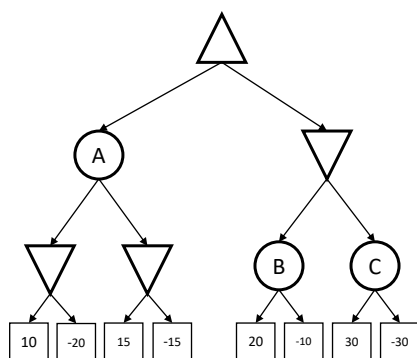


- (آ) (۴ نمره) اگر محمدرضا در ازای ۲ دلار برره در هر مرحله قبول کند که حرکت مورد درخواست آراین را انجام دهد، بیشینه مقداری که آراین می‌تواند برنده شود چقدر است.
- (ب) (۴ نمره) اگر محمدرضا در ازای ۵ دلار برره در هر مرحله قبول کند که حرکت مورد درخواست آراین را انجام دهد، بیشینه مقداری که آراین می‌تواند برنده شود چقدر است.
- (ج) (۴ نمره) استدلال کنید که چرا نمی‌توان از حرص آلفا-بتا برای این بازی فاسد استفاده کرد.

۴. (۱۱ نمره) در محیط زیر که به ابعاد 100×100 است، بازکن در ابتدا در خانه S قرار دارد. بازیکن در حرکت اول می‌تواند انتخاب کند که به سمت چپ یا پایین برود ولی در ادامه همان روند را باید تا انتها طی کند.

S	۱۰۰	-۱	-۱	-۱	...	-۱	-۱	-۱
۵۰								
۱								
۱								
۱								
⋮								
۱								
۱								
۱								

- (آ) (۵ نمره) مقدار Q را به ازای هر دو حرکت از خانه S بر حسب ضریب تخفیف γ محاسبه کنید.
- (ب) (۶ نمره) به صورت تقریبی محاسبه کنید که به ازای چه مقادیری از ضریب تخفیف، حرکت اول رو به پایین خواهد بود.
۵. (۱۸ نمره) در یک بازی جمع صفر، گره‌های احتمالاتی وجود دارند که یک سکه، نتیجه عبور از آنها را مشخص می‌کند. درخت این بازی به صورت زیر است.



پیش از شروع بازی سه سکه A و B و C را پرتاب می‌کنیم و روی آن را می‌پوشانیم. هر زمان که به هر یک از آنها رسیدیم، آن سکه را مشاهده می‌کنیم.

(آ) (۶ نمره) مقدار میانگین ریشه را بیابید.

(ب) (۶ نمره) فرض کنید که بازیگر اول می‌تواند تقلب کند و نتیجه یکی از سکه‌ها را ببیند. دیدن کدام سکه بیشترین منفعت را برای او دارد؟

(ج) (۶ نمره) فرض کنید که بازیگر اول می‌تواند تقلب کند و نتیجه یک سکه را مشخص کند. مشخص کردن نتیجه کدام سکه بیشترین منفعت را برای او دارد؟

سوالات عملی (۵۵ نمره)

۱. (۲۵ نمره) برای حل سوال عملی اول به نوت‌بوک مربوطه مراجعه کنید.

۲. (۳۰ نمره) برای حل سوال عملی اول به نوت‌بوک مربوطه مراجعه کنید.