دانشکده مهندسی کامپیوتر

هوش مصنوعی و سیستم‌های خبره

تمرین تشریحی چهارم

نام و نام خانوادگی . . . . . .

شماره دانشجویی . . . . . .

مدرس . . . . . محمدطاهر پیله‌ور - سید صالح اعتمادی

طراحی و تدوین . . . . . سپهر باباپور (@Spr_Bpr)

تاریخ انتشار . . . . . ۱۵ آبان ۱۳۹۹

تاریخ تحویل . . . . . ۲۹ آبان ۱۳۹۹

فهرست مطالب

۱ سوالات بخش تئوری

** در این بخش به سوالاتی که دارای ⋆ هستند پاسخ دهید **

۱.۱ سوال ۱

توضیح دهید چرا در MDPها نمی‌توان از روش planning استفاده کرد. راه‌حل جایگزین را توضیح دهید.

پاسخ:

.....

.....

.....

.....

.....

.....

۲.۱ سوال ۲

فرایندهای مارکوفی را تعریف کرده و بگویید کدام یک از فرایندهای زیر مارکوفی هستند.

- بازی سودوکو - بازی اسم - فامیل - بازی بی‌سوالی

پاسخ:

.....

.....

.....

.....

.....

۳.۱ ⋆ سوال ۳ (۲۰ نمره)

اگر به جای ضریب تخفیف γ^t از توابع زیر استفاده شود:

* $|sin(t)|$ * e^{-t} * $log(t)$

به سوالات زیر پاسخ دهید.

۱- کدام یک مشکل نامحدود شدن بازی را برطرف می‌کنند؟ توضیح دهید.

۲- برای تابع پاسخ قسمت اول، با فرض پاداش یک واحد در هر لحظه کوچک زمانی (dt) پاداش کل را محاسبه کنید.

پاسخ:

.....

.....

.....

.....

۲ مسائل محاسباتی

** در این بخش به سوالاتی که دارای ⋆ هستند پاسخ دهید **

۱.۲ سوال ۱: کارت بردار!!

فرض کنید شما در یک مسابقه کارت بازی شرکت کرده‌اید که در آن ۳ نوع کارت با شماره‌های ۲، ۳، ۴ وجود دارد. شما در هر مرحله از بازی تا زمانی که به مجموع امتیاز ۶ نرسیده‌اید می‌توانید یا یک کارت بردارید یا بازی را به اتمام برسانید. احتمال آمدن هر کارت با هم برابر است. زمانی که مجموع امتیازات شما ۶ یا بیشتر شود امتیازات شما صفر می‌شود و بازی تمام می‌شود و زمانی که خودتان بازی را تمام کرده باشید امتیازتان برابر مجموع کارت‌هایی که کسب کرده‌اید می‌شود. همچنین برداشتن کارت را بدون هزینه در نظر بگیرید.
در این سوال از شما خواسته شده است که بازی فوق را به صورت یک مدل مارکوفی در نظر بگیرید و به سوالات زیر پاسخ دهید.

۱. ابتدا تابع انتقال (transition function) و تابع پاداش (reward function) را برای این مدل محاسبه کنید.

۲. سپس جدول زیر را کامل کنید.

حالت	۰	۲	۳	۴	۵
π_i	برداشتن کارت	اتمام بازی	برداشتن کارت	اتمام بازی	برداشتن کارت
v^{π_i}					
π_{i+1}					

شکل ۱: جدول سوال کارت بردار!!

پاسخ:

۲.۲ ⋆ سوال ۲: تاس بریز!! (۴۵ نمره)

فرض کنید در یک بازی ریختن تاس شرکت کرده‌اید 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 می‌کنید.
قانون بازی به این شکل است که شما موظف هستید در بار اول یک تاس بریزید، اما در سایر مراحل دو انتخاب دارید:

- اتمام بازی: با این حرکت شما به اندازه عدد روی تاس سکه دریافت می‌کنید.
- تاس ریختن: یک سکه هزینه می‌کنید و بار دگیر تاس می‌ریزید.

لذا بازی را می‌توان 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 یک حرکت اتمام بازی وجود دارد که بازیکن را به حالت پایانی می‌برد و در حالت پایانی حرکتی وجود ندارد. هر حالت بین شروع و پایان با s_i نمایش داده می‌شود که بدین معنی است که عدد i در تاس آمده است.
با توجه به توضیحات فوق به سوالات زیر پاسخ دهید:

۱. فرض کنید π_iهای زیر در ابتدا وجود دارد، ردیف v^{π_i} را کامل کنید. $(\gamma = 1)$

حالت	s_1	s_2	s_3	s_4	s_5	s_6
π_i	تاس ریختن	تاس ریختن	اتمام بازی	اتمام بازی	اتمام بازی	اتمام بازی
v^{π_i}						

شکل ۲: جدول قسمت اول سوال تاس بریز!!

۲. باتوجه به جدول فوق مقادیر π_i را بروزرسانی کنید و در جدول زیر جایگذاری کنید. این مقادیر می‌تواند سه حالت تاس ریختن، اتمام بازی و تاس ریختن / اتمام بازی باشد. $(\gamma = 1)$

حالت	s_1	s_2	s_3	s_4	s_5	s_6
π_i	تاس ریختن	تاس ریختن	اتمام بازی	اتمام بازی	اتمام بازی	اتمام بازی
π_{i+1}						

شکل ۳: جدول قسمت دوم سوال تاس بریز!!

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

پاسخ:

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

۳.۲ ⋆ سوال ۳: یک **MDP** ساده که دیگر مثل قبل نیست؟ (۳۵ نمره)

یک مسئله MDP را تصور کنید که در آن تابع پاداش به جای $R(s)$، $\eta R(s)$ باشد که در آن η یک ثابت مثبت است.
سایر خصوصیات این مسئله MDP تغییر نکرده است.
ثابت کنید راهبرد (policy) بهینه در مسئله MDP جدید مشابه راهبرد (policy) در مسئله اولیه است.

پاسخ: