



آزمون میان ترم

۱. (۲۰ نمره) به سوالات زیر به طور مختصر پاسخ دهید:

- (آ) درست یا غلط بودن عبارت زیر را با ذکر دلیل مشخص کنید.
الگوریتم ژنتیک^۱ با اندازه‌ی جمعیت^۲ $N = ۱$ معادل Simulated Annealing با $T = \infty$ خواهد بود.
- (ب) آیا k -consistency، $(k-1)$ -consistency را نتیجه می‌دهد؟ در صورت مثبت بودن پاسخ، توضیح دهید در غیر این صورت مثال نقضی ارائه کنید.
- (ج) پیچیدگی زمانی Policy Iteration از $O(S^4 A)$ و پیچیدگی زمانی Policy Evaluation از $O(S^2)$ است.
- (د) یک صفحه به صورت زیر در نظر بگیرید. شما در ابتدا در خانه‌ی سمت بالا چپ که با حرف S مشخص شده است قرار دارید و قصد دارید به خانه‌ی G برسید. برای رسیدن به هدف مجاز به حرکت به هر چهار طرف هستید. البته نمی‌توانید وارد خانه‌هایی که در آنها دیوار قرار دارد، شوید. عده‌ای از دوستان شما می‌خواهند تونلی حفر کنند تا شما بتوانید از زیر دیوارها عبور کنید. ورودی تونل خانه‌ی سمت چپ پایین است که با حرف X مشخص شده است. مشکل اینجاست که شما اطلاعی از اینکه حفر تونل به پایان رسیده است یا نه ندارید. در صورتی که تونل ناقص حفر شده باشد از خانه‌ی X می‌توان با یک حرکت به Y_1 رفت. این در حالی است که اگر تونل به صورت کامل حفر شده باشد، از X می‌توان به Y_2 که همان خانه‌ی G است، رفت. دقت داشته باشید که هزینه‌ی عبور از تونل مثل بقیه‌ی حرکات برابر با یک است.
- حالا شما می‌خواهید یک نماینده از طرف خودتان را به صفحه‌ی داده شده بفرستید تا مسیر بهینه را به کمک الگوریتم A^* پیدا کند. برای هدایت او لازم است تا یک تابع مکاشفه^۳ در اختیار او بگذارید. یک تابع مکاشفه‌ی admissible به جز تابع $h(x) = ۰$ ارائه دهید.

S				
X	Y_1			Y_2

۲. (۲۰ نمره) الگوریتم ژنتیک را در نظر بگیرید که در آن از کروموزوم‌هایی به فرم $x = ABCDEFGH$ با طول ثابت ۸ استفاده می‌کنیم. هر ژن می‌تواند عددی بین ۰ تا ۹ را بگیرد. تابع fitness را برای هر کروموزوم x به صورت زیر تعریف می‌کنیم:

$$f(x) = A + (2 \times B) + (3 \times C) - D - E - (2 \times F) + G - H$$

جمعیت اولیه نیز از ۴ کروموزوم به شکل زیر تشکیل شده است:

$$x_1 = 24513892, x_2 = 13562893, x_3 = 43213205, x_4 = 12903621$$

- (آ) در این قسمت ابتدا fitness چهار نمونه زیر را به دست آورده و سپس احتمال انتخاب آن‌ها برای دور بعد را به دست آورید.
- (ب) حال عملیات crossover را بر روی دو نمونه با بیشترین fitness انجام دهید، به صورت تک نقطه‌ای که محل crossover نقطه‌ی میانی کروموزوم باشد. همچنین عملیات crossover را به صورت دو نقطه‌ای بر روی اولین و سومین کروموزوم از لحاظ بیشتر بودن fitness انجام دهید. نقاط C و F را به عنوان دو نقطه مورد نیاز در نظر بگیرید.
- (روش این نوع crossover دو نقطه‌ای به این صورت است که CDEF از یک کروموزوم و بقیه ژن‌ها از یک کروموزوم دیگر بدست می‌آید)
- (ج) نتایج قسمت قبل را به عنوان جمعیت جدیدمان در نظر بگیرید. حال دوباره fitness را برای این جمعیت جدید حساب کنید. آیا مجموع fitness افزایش پیدا کرده است؟

(د) بهترین کروموزوم مسئله (بیشترین fitness را داشته باشد) را بدست آورید. مقدار fitness آن را نیز حساب کنید.

(ه) طبق جمعیت فعلی که داریم آیا بدون انجام عملیات mutation امکان رسیدن به جواب بهینه وجود دارد؟ توضیح دهید.

۳. (۱۲ نمره) در رابطه با هر یک از مجموعه‌های زیر تعیین کنید که مجموعه موردنظر محدب می‌باشد یا خیر. توجه کنید که برای نشان دادن محدب بودن یک مجموعه بایستی اثبات کنید که مجموعه مفروض محدب است و برای نشان دادن عدم محدب بودن یک مجموعه کافی است که مثال نقضی برای آن بیاورید.

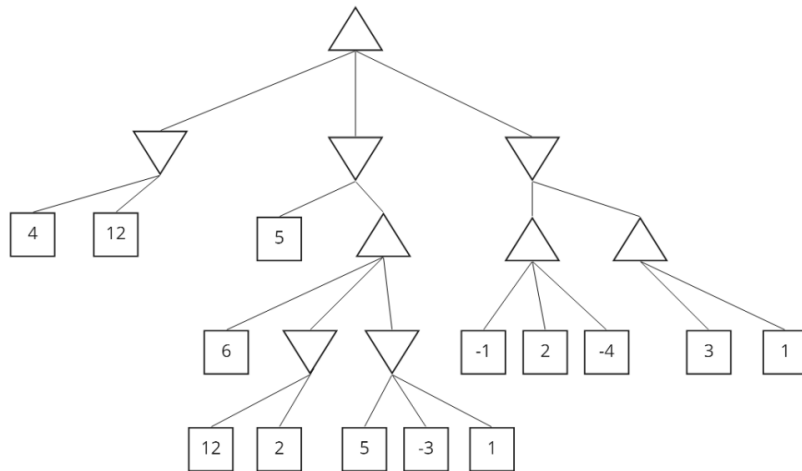
(آ) مجموعه C که برای زوج مرتب‌های شامل یک بردار x و عدد حقیقی t به شکل زیر تعریف می‌شود.

$$C = \{(x, t) \mid \|x\| \leq t\}$$

(ب) مجموعه C که شامل توابع احتمالاتی f می‌باشد. برای تعریف دقیق‌تر C خواهیم داشت:

$$C = \{f \mid f \geq 0, \int f = 1\}$$

۴. (۱۰ نمره) درخت minimax زیر را در نظر بگیرید و با توجه به آن به دو بخش زیر پاسخ دهید.



- مقادیر تمام گره‌های max و min را پر کنید. مثلث‌های رو به بالا گره‌های بیشینه هستند، در حالی که مثلث‌های رو به پایین گره‌های کمینه هستند.
- هرس آلفا-بتا را با فرض گسترش از چپ به راست انجام دهید.

۵. (۱۲ نمره) مسئله رمزگذاری زیر را در نظر بگیرید:

$$\begin{array}{r} T W O \\ + T W O \\ \hline F O U R \end{array}$$

هر حرف انگلیسی نمایانگر یک رقم متمایز است. می‌خواهیم مقداری به این حروف را به گونه‌ای انجام دهیم تا رابطه بالا برقرار باشد. به عبارتی حاصل جمعی که در بالا با جایگذاری اعداد بجای حروف بدست می‌آید، جایگذاری صحیحی باشد. همچنین استفاده از رقم ۰ در شرایطی که تعداد ارقام اعداد را کاهش دهد مجاز نیست. (T و F نمی‌توانند ۰ باشند).

(آ) مسئله را به صورت یک مسئله CSP مدل کنید و گراف قیود آن را رسم کنید.

(ب) با استفاده از هیوریستیک‌های MRV و LCV این مسئله را حل کنید و درخت نهایی را رسم کنید.

۶. (۱۵ نمره) فرض کنید حسین MDP ساده $M = (S, A, R, P, \gamma)$ با $|S| < \infty$ و $|A| < \infty$ و $0 < \gamma < 1$ را دارد.

حال یک MDP جدید $\hat{M} = (S, A, \hat{R}, P, \gamma)$ به طوریکه $\hat{R} - R = \epsilon$ را فرض کنید.

درواقع در این سوال می‌خواهیم بررسی کنیم آیا اگر به تمامی پاداش‌ها مقدار ϵ اضافه شود تابع ارزش و سیاست بهینه چه تغییری می‌کنند؟ راهنمایی: ثابت کنید که $\hat{V}^* - V^* = \frac{\epsilon}{(1-\gamma)}$

آیا سیاست بهینه تغییر می‌کند؟

۷. (۱۶ نمره) حسین از شاگردان خوب در رشته خودش است اما در تصمیم‌گیری برای آینده تحصیلی‌اش دچار سردرگمی شده و مرتباً در حال عوض کردن تصمیم‌هایش است. او با پریدن بین فیلدها به دنبال فیلد مورد علاقه خودش می‌گردد. فرض کنید حسین بین ۵ فیلد Hardware, RL, Information Theory, Robotics, Quantum در حال جابجایی است. با توجه به شناختی که از حسین داریم می‌دانیم که بعد از ورود به هر فیلد میزان جایزه‌ای که می‌گیرد (معلوم نیست از چه کسی) چقدر است. شما می‌توانید مقدار این جایزه‌ها را به ازای فیلدهای مختلف در جدول زیر مشاهده کنید:

Field	Reward
Hardware	-1
RL	+1
Robotics	+1
Quantum	+2

از آنجایی که حسین کمی هم حواس پرت است، بعضی وقت‌ها فیلدها را با هم قاطی می‌کند. جدول زیر نشان می‌دهد اگر حسین قصد تغییر فیلد بدهد با چه احتمالی به کدام فیلدها می‌پرد:

Start State	Action	Probability	End State
Hardware	ToML	0.6	Robotics
Hardware	ToML	0.4	RL
Hardware	ToPhysics	1	Quantum
RL	ToPhysics	0.2	Hardware
RL	ToPhysics	0.8	Robotics
Robotics	ToHardware	0.5	Hardware
Robotics	ToHardware	0.5	Robotics
Robotics	ToPhysics	0.8	Quantum
Robotics	ToPhysics	0.2	RL
Quantum	ToML	1	RL
Quantum	ToHardware	1	Hardware

حال فرض کنید حسین در یک سال اخیر با سیاست زیر قصد داشته بین فیلدها جابجا شود:

$$\begin{aligned}\pi_0(\text{Hardware}) &= \text{ToPhysics} \\ \pi_0(\text{Robotics}) &= \text{ToHardware} \\ \pi_0(\text{RL}) &= \text{ToPhysics} \\ \pi_0(\text{Quantum}) &= \text{ToML}\end{aligned}$$

(آ) تابع ارزش سیاست اخیر حسین را بدست آورید. لازم نیست مقادیر نهایی را بدست آورید و صرف نوشتن معادله‌ها کافی است. (بر حسب γ)

(ب) با استفاده از یک گام اجرای الگوریتم Policy Iteration با فرض

$$V_Q > V_{RL} > V_{Ro} > V_{Hw}$$

به حسین کمک کنید تا بتواند راحت‌تر فیلد خودش را پیدا کند.