




بسم الله الرحمن الرحيم
تمرین نظری سری سوم هوش مصنوعی

پاییز ۱۳۹۶

سوال اول

یک **gridworld** را در نظر بگیرید که در آن حرکت های **Left** و **Right** با موفقیت ۱۰۰ درصد انجام می شوند. تنها حرکت های مجاز در هر **state** حرکت به خانه ی همسایه ی خود می باشد. در **state a** حرکت **exit** نیز وجود دارد که باعث می شود به **terminal state** برویم و **reward** آن (۱۰ امتیاز) را دریافت کنیم. این امکان نیز در **state e** وجود دارد. حرکت **exit** نیز با احتمال ۱۰۰ درصد انجام می شود.
الف) فرض کنید مقدار **discount** برابر ۱ باشد (یعنی $\lambda = 1$). با توجه به شکل مقادیر زیر را به دست آورید.

10				1
a	b	c	d	e

$$V_0(d) =$$

$$V_1(d) =$$

$$V_2(d) =$$

$$V_3(d) =$$

$$V_4(d) =$$

$$V_5(d) =$$



بسم الله الرحمن الرحيم

تمرین نظری سری سوم هوش مصنوعی

پاییزه ۱۳۹۶

ب) فرض کنید مقدار discount برابر ۰,۲ باشد (یعنی $\lambda = 0,2$). با توجه به شکل مقادیر زیر را به دست آورید.

10				1
a	b	c	d	e

$$V^*(a) = V_\infty(a) =$$

$$V^*(b) = V_\infty(b) =$$

$$V^*(c) = V_\infty(c) =$$

$$V^*(d) = V_\infty(d) =$$

$$V^*(e) = V_\infty(e) =$$

ج) فرض کنید مقدار discount برابر ۱ باشد (یعنی $\lambda = 1$).

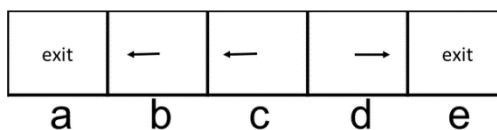
10				1
a	b	c	d	e



بسم الله الرحمن الرحيم
تمرین نظری سری سوم هوش مصنوعی

پاییز ۱۳۹۶

فرض کنید π_1 به صورت زیر باشد، مقادیر زیر را برای این policy ارزیابی کنید.



$$V^{\pi_1}(a) =$$

$$V^{\pi_1}(b) =$$

$$V^{\pi_1}(c) =$$

$$V^{\pi_1}(d) =$$

$$V^{\pi_1}(e) =$$

(د) فرض کنید مقدار discount برابر ۰,۹ باشد (یعنی $\gamma = ۰,۹$).



بسم الله الرحمن الرحيم
تمرین نظری سری سوم هوش مصنوعی

پاییز ۱۳۹۶

10				1
a	b	c	d	e

فرض کنید π_i به صورت زیر باشد، مقادیر زیر را برای این policy ارزیابی کنید.

exit	←	→	←	exit
a	b	c	d	e

$$V^{\pi_i}(a) =$$

$$V^{\pi_i}(b) =$$

$$V^{\pi_i}(c) =$$

$$V^{\pi_i}(d) =$$

$$V^{\pi_i}(e) =$$



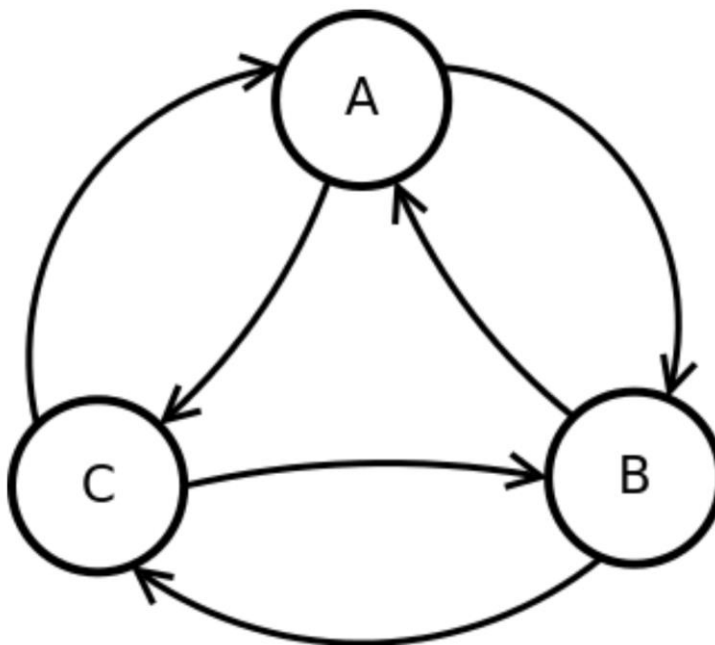
بسم الله الرحمن الرحيم
تمرین نظری سری سوم هوش مصنوعی

پاییز ۱۳۹۶

سوال دوم

دیاگرام انتقال، تابع انتقال و تابع **reward** به صورت زیر مشخص شده است.

Discount Factor, $\gamma = 0.5$



s	a	s'	T(s,a,s')	R(s,a,s')
A	Clockwise	B	0.8	-1.0
A	Clockwise	C	0.2	-2.0
A	Counterclockwise	B	0.2	-2.0
A	Counterclockwise	C	0.8	-2.0
B	Clockwise	A	0.2	2.0
B	Clockwise	C	0.8	0.0
B	Counterclockwise	A	1.0	-1.0
C	Clockwise	A	1.0	-2.0
C	Counterclockwise	A	0.2	0.0
C	Counterclockwise	B	0.8	2.0

الف) فرض کنید بعد از k iteration؛ V_k value iteration مقادیر برای V_k به صورت زیر باشد.

$V_k(A)$	$V_k(B)$	$V_k(C)$
-0.880	0.920	1.640

مقدار $V_{k+1}(C)$ چقدر می باشد؟

ب) حال فرض کنید که ما V_k value iteration را برای completion انجام دادیم و مقادیر زیر را به دست آورده ایم.

$V^*(A)$	$V^*(B)$	$V^*(C)$
-0.540	1.148	2.005



بسم الله الرحمن الرحيم
تمرین نظری سری سوم هوش مصنوعی

پاییز ۱۳۹۶

مقدار $Q^*(C, \text{clockwise})$ چیست؟

مقدار $Q^*(C, \text{counterclockwise})$ چیست؟

سوال سوم

کدام یک از عبارات زیر درباره ی value iteration درست می باشد؟ فرض می کنیم که MDP ما تعدادی محدودی action و state دارد و مقدار discount بین ۰ تا ۱ می باشد.

الف) value iteration تضمین می شود که همگرا (converge) می باشد.

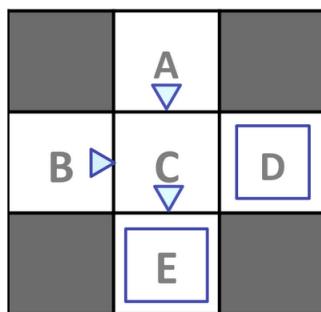
ب) value iteration به برداری یکسان از value ها (V^*) همگرا می باشد. مهم نیست که چه مقادیری را برای مقدار دهی اولیه ی V استفاده کرده ایم.

ج) هیچکدام

سوال چهارم

شکل و فرضیات زیر را در نظر بگیرید.

Input Policy π



Assume: $\gamma = 1$

Observed Episodes (Training)

Episode 1

A, south, C, -1
C, south, E, -1
E, exit, x, +10

Episode 2

B, east, C, -1
C, south, D, -1
D, exit, x, -10

Episode 3

B, east, C, -1
C, south, E, -1
E, exit, x, +10

Episode 4

A, south, C, -1
C, south, E, -1
E, exit, x, +10



بسم الله الرحمن الرحيم
تمرین نظری سری سوم هوش مصنوعی

پاییز ۱۳۹۶

چه مدلی از episode های مشاهده شده در بالا یاد گرفته خواهند شد؟

$T(A, \text{south}, C) =$

$T(B, \text{east}, C) =$

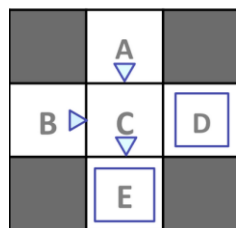
$T(C, \text{south}, E) =$

$T(C, \text{south}, D) =$

سوال پنجم

شکل و فرضیات زیر را در نظر بگیرید.

Input Policy π



Assume: $\gamma = 1$

Observed Episodes (Training)

Episode 1

A, south, C, -1
C, south, E, -1
E, exit, x, +10

Episode 2

B, east, C, -1
C, south, D, -1
D, exit, x, -10

Episode 3

B, east, C, -1
C, south, E, -1
E, exit, x, +10

Episode 4

A, south, C, -1
C, south, E, -1
E, exit, x, +10



بسم الله الرحمن الرحيم
تمرین نظری سری سوم هوش مصنوعی

پاییز ۱۳۹۶

برای مقادیر زیر چه تخمین هایی از direct evaluation به دست می آید؟

$$\hat{V}^{\pi}(A) =$$

$$\hat{V}^{\pi}(B) =$$

$$\hat{V}^{\pi}(C) =$$

$$\hat{V}^{\pi}(D) =$$

$$\hat{V}^{\pi}(E) =$$

- تمرینها از طریق ایمیل تحویل داده شود. موضوع ایمیل ارسالی حتما به فرم مثال زیر باشد:

<StudentNo>-<FirstName>-<LastName>-HW<No>-Kharazmi-AI-Fall96

83405307-Hadi-Asheri-HW1-Kharazmi-AI-Fall96

- در صورت تاخیر در تحویل پاسخ تمرین بعد از ۴۸ ساعت ۲۰ درصد نمره، از ۴۸ ساعت تا یک هفته ۵۰ درصد نمره و در صورت تاخیر بیشتر هیچ نمره‌ای تعلق نخواهد گرفت.
- تمرین ها به آدرس ai.kharazmi.fall96@gmail.com ارسال شود.
- مهلت ارسال: ۱۳۹۵/۹/۵ - ساعت ۱۸