Subject:

یاسمین مدنی ۹۷۵۳۲۲۶۵

سؤال ۱:

آ) به فرآیندهایی که حالت آینده تنها به حالت کنونی وابسته باشد و به حالت های پیش از این نیاز نباشد فرآیندهای مارکوفی گفته می شود.

ب) planing در شرایطی بود که با قصد انجام یک حرکت می توانستیم آن حرکت را انجام دهیم، در فرآیندهای مارکوفی انتخاب یک حرکت (action) حتماً به معنای انجام دادن آن نیست و در حرکت ها uncertainty وجود دارد بنابراین روش planing کاربرد ندارد.

راه حل مسائل مارکوف محاسبه بهترین (policy) برای حالت های مختلف است. در این مسائل باید حرکتی که بهترین نتیجه را می دهد سود تحقق فرا گیرد.

ج) با توجه به تعریف فرآیند مارکوفی که در قسمت الف ارائه شد سودوکو یک فرآیند مارکوفی است چرا که state بعدی تنها به وضعیت کنونی بستگی دارد.

د) بازی بیست سؤالی یک فرآیند مارکوفی نیست چرا که سؤالی که در هر مرحله جدید پرسیده می شود به تمامی سؤالات قبلی تا مرحله ی کنونی مربوط می شود.

سؤال ۲:

آ) برای جلوگیری از نامحدود شدن بازی باید ضریب تخفیف در طول زمان کوچک شده و به صفر میل کند. از میان توابع صورت سؤال تابع $e^{-t}$ تابع اکیداً نزولی است و به مرور زمان در حد به صفر میل می کند.

ب) پاداش کل مساحت زیر نمودار تابع تخفیف می شود که داریم:

$$\int_0^\infty e^{-t}dt = -(e^{-t})_0^\infty = -(0-1) = 1$$

Year: Month: Date: Subject:

سؤال ۳:

الف)

$$V^*(D) = x$$

$$V^*(A) = \max(\overset{A\to D}{1 + 0.5x}, \overset{loop(\infty)}{2})$$

چون اگر از A به B برویم سپس به A برگردیم و به D برویم واضح است که از حالت مستقیم A به D امتیاز کمتری می‌گیریم از این حالت در نوشتن در تابع max صرف نظر می‌کنیم.

برای چرخش بی‌نهایت در map و خروج در حلقه‌ی بی‌نهایت داریم

$$reward(\infty) = \frac{1}{1-\frac{1}{2}} = 2$$

چون که یک دنباله هندسی به صورت $(\frac{1}{2})^n$ تشکیل می‌دهد.

$B \to C \to D$ or $B \to A \to D$

$$V^*(B) = \max(1 + 0.5(1 + 0.5x), 2)$$

مشابه بالا که توضیح داده شد.

$$V^*(C) = \max(\overset{C\to D}{1 + 0.5x}, \overset{loop(\infty)}{2})$$

$C \to B \to A \to B \to C$

C مشابه A است و توضیحات برای A برای C نیز صدق می‌کند.

ب)

$$Q^*(A, Down) = Q^*(A, Right) = V^*(A)$$

$$V^*(A) = Q^*(A, Down) = \frac{1}{2}\left(0 + \frac{1}{2}V^*(A)\right) + \frac{1}{2}\left(1 + \frac{1}{2}x\right) =$$

$$\frac{1}{2} + \frac{1}{4}x + \frac{1}{4}(V^*(A)) = V^*(A)$$

$$V^*(A) = \frac{2}{3} + \frac{1}{3}x \quad \text{I}$$

$$V^*(A) = Q^*(A, Right) = \frac{1}{2}\left(0 + \frac{1}{2}V^*(A)\right) + \frac{1}{2}\left(1 + \frac{1}{2}V^*(B)\right) =$$

$$\frac{1}{2} + \frac{1}{4}V^*(A) + \frac{1}{4}V^*(B)$$

$$V^*(A) = \frac{2}{3} + \frac{1}{3}V^*(B) \quad \text{II}$$

Year: Month: Date: Subject:

از آنجا که حرکت از B به A و C قرینه است.

$$V^*(B) = Q^*(B, Left)$$

$$V^*(B) = \frac{1}{2}\left(0 + \frac{1}{2}V^*(B)\right) + \frac{1}{2}\left(1 + \frac{1}{2}V^*(A)\right) = \frac{1}{2} + \frac{1}{4}V^*(B) + \frac{1}{4}V^*(A)$$

$$V^*(B) = \frac{2}{3} + \frac{1}{3}V^*(A) \quad \text{III}$$

$$\text{I}, \text{II}, \text{III} \longrightarrow x = 1$$

راه حل جالبه $x$ :

$$\left.\begin{array}{ll} \text{I} & V^*(A) = \frac{2}{3} + \frac{1}{3}x \\ \text{II} & V^*(A) = \frac{2}{3} + \frac{1}{3}V^*(B) \\ \text{III} & V^*(B) = \frac{2}{3} + \frac{1}{3}V^*(A) \end{array}\right\}$$

$$V^*(A) = \frac{2}{3} + \frac{1}{3}\left(\frac{2}{3} + \frac{1}{3}V^*(A)\right) = \frac{2}{3} + \frac{2}{9} + \frac{1}{9}V^*(A)$$

$$\cancel{\frac{2}{3}} + \frac{1}{3}x = \cancel{\frac{2}{3}} + \frac{2}{9} + \frac{1}{9}V^*(A) \rightarrow x = \frac{2}{3} + \frac{1}{3}V^*(A) \rightarrow \text{۳}x = 2 + V^*(A)$$

$$x = 3V^*(A) - 2$$

$$\left\{\begin{array}{ll} \text{۴}x = 4V^*(A) & \rightarrow x = V^*(A) \\ \text{۲}x = -2V^*(A) + 4 & \rightarrow x = \text{۲} - V^*(A) \end{array}\right\} \quad \text{۲}x = \text{۲} \rightarrow \boxed{x = 1}$$

(2.

$$V_{k+1}(s) = \max_a \sum_{s'} T(s,a,s') \cdot \left[\frac{1}{6}\sum_{i=1}^{6} \left(R(s,a,s') + i\right) + \gamma V_k(s')\right]$$

$$= \max_a \sum_{s'} T(s,a,s')\left(R(s,a,s') + \cancel{\frac{21}{6}}^{\,3.5} + \gamma V_k(s')\right)$$

Year: Month: Date: Subject:

سؤال ۴:

الف)

| حالت | $S_1$ | $S_2$ | $S_3$ | $S_4$ | $S_5$ | $S_6$ |
|---|---|---|---|---|---|---|
| $\pi i$ | تاس ریختن | تاس ریختن | اتمام بازی | اتمام بازی | اتمام بازی | اتمام بازی |
| $V^{\pi i}$ | 3 | 3 | 3 | 4 | 5 | 6 |

با اتمام بازی داریم: $V_j^{\pi i} = j$

برای ریختن تاس داریم: $\frac{1}{4}\sum_{j=1}^{4} V_j^{\pi i}$

$$V_1^{\pi i} = \frac{1}{6}\left(V_1^{\pi i} + V_2^{\pi i} + V_3^{\pi i} + V_4^{\pi i} + V_5^{\pi i} + V_6^{\pi i}\right)$$

$$V_1^{\pi i} = \frac{1}{6}\left(V_1^{\pi i} + V_2^{\pi i} + 18\right)$$ رابطه ①

$$V_2^{\pi i} = \frac{1}{6}\left(V_1^{\pi i} + V_2^{\pi i} + V_3^{\pi i} + V_4^{\pi i} + V_5^{\pi i} + V_6^{\pi i}\right)$$

$$V_2^{\pi i} = \frac{1}{6}\left(V_1^{\pi i} + V_2^{\pi i} + 18\right)$$ رابطه ②

از رابطه ① و رابطه ② $\Longrightarrow V_1^{\pi i} = V_2^{\pi i} = 3$

ب)

اگر تاس پرتاب کنیم: $V_5^{\pi i} = -1 + \frac{1}{6}(3+3+3+4+5+6) = 3$

اگر بازی تمام شود: $V_5^{\pi i} = 5$

پس در حالت های $S_4$ و $S_5$ و $S_6$ که بهترین راه حل اتمام بازی است. امتیاز در حالت اتمام بازی بیشتر از ریختن تاس می‌شود.

در حالت $S_3$ دو امتیاز یک شده برابر می‌شود پس دو حالت داریم. $\pi_3 =$ اتمام بازی / پرتاب تاس

در حالت $S_1$ و $S_2$ امتیاز پرتاب تاس بیشتر از امتیاز پایان بازی می‌شود پس

$\pi_1 = \pi_2 =$ پرتاب تاس

Year: Month: Date: Subject:

ج) از آنجا که در قسمت ب $\Pi_{i+1}$ محاسبه شده و policy ارائه شده مانند مرحله $\Pi_i$ است
می‌توان نشان داد که مرحله همگرایی رخ داده و به policy بهینه رسیده‌ایم.

Year: Month: Date: Subject:

سؤال ۵:

الف)
$$V_{(k+1)}(S) = \max_a \left( R(S,a,S') + \gamma V_k(S') \right)$$

$V_0(d) = 0$

| | a | b | c | d | e |
|---|---|---|---|---|---|
| $V_0$ | 0 | 0 | 0 | 0 | 0 |
| $V_1$ | 10 | 0 | 0 | 0 | 1 |
| $V_2$ | 10 | 10 | 0 | 1 | 1 |
| $V_3$ | 10 | 10 | 10 | 1 | 1 |
| $V_4$ | 10 | 10 | 10 | 10 | 1 |
| $V_5$ | 10 | 10 | 10 | 10 | 10 |

$V_5(a) = 10$

$V_5(b) = \max(V_4(a), V_4(c)) = 10$

$V_5(c) = \max(V_4(b), V_4(d)) = 10$

$V_5(d) = \max(V_4(c), V_4(e)) = 10$

$V_5(e) = 10$

$V_1(a) = R(a, exit) + V_0(t) = 10 + 0 = 10$

$V_1(b) = \max(V_0(a), V_0(c)) = 0$

$V_1(c) = \max(V_0(b), V_0(d)) = 0$

$V_1(d) = \max(V_0(c), V_0(e)) = 0$

$V_1(e) = R(e, exite) + V_0(t) = 1$

$V_3(a) = 10$

$V_3(b) = \max(V_2(a), V_2(c)) = 10$

$V_3(c) = \max(V_2(b), V_2(d)) = 10$

$V_3(d) = \max(V_2(c), V_2(e)) = 1$

$V_3(e) = 1$

$V_2(a) = 10$

$V_2(b) = \max(V_1(a), V_1(c)) = 10$

$V_2(c) = \max(V_1(b), V_1(d)) = 0$

$V_2(d) = \max(V_1(c), V_1(e)) = 1$

$V_2(e) = 1$

$V_4(a) = 10$

$V_4(b) = \max(V_3(a), V_3(c)) = 10$

$V_4(c) = \max(V_3(b), V_3(d)) = 10$

$V_4(d) = \max(V_3(c), V_3(e)) = 10$

$V_4(e) = 1$

۵)

ب) $\max\left[R(s,a,s') + 0/2\, V_k(s')\right]$

| | a | b | c | d | e |
|---|---|---|---|---|---|
| $V_0$ | 0 | 0 | 0 | 0 | 0 |
| $V_1$ | 10 | 0 | 0 | 0 | 1 |
| $V_2$ | 10 | 2 | 0 | 0/2 | 1 |
| $V_3$ | 10 | 2 | 0/4 | 0/2 | 1 |
| $V_4$ | 10 | 2 | 0/4 | 0/2 | 1 |
| $V_5$ | 10 | 2 | 0/4 | 0/2 | 1 |

$V_1(a) = 10$

$V_1(b) = \max(V_0(b) + 0/2(0),\ V_0(b) + 0/2(0)) = 0$

$V_1(c) = \max(V_0(c) + 0/2(0),\ V_0(c) = 0$

$V_1(d) = \max(0/2\, V_0(c),\ 0/2\, V_0(e)) = 0$

$V_1(e) = 1$

$V_2(a) = 10$

$V_2(b) = \max(0/2\, V_1(a),\ 0/2\, V_1(c)) = \max(2, 0) = 2$

$V_2(c) = \max(0/2\, V_1(b),\ 0/2\, V_1(d)) = 0$

$V_2(d) = \max(0/2\, V_1(c),\ 0/2\, V_1(e)) = 0/2$

$V_2(e) = \max(0/2\, V_1(d),\ 1) = 1$

$V_3(a) = 10$

$V_3(b) = \max(0/2\, V_2(a),\ 0/2\, V_2(c)) = 2$

$V_3(c) = \max(0/2\, V_2(b),\ 0/2\, V_2(d)) = 0/4$

$V_3(d) = \max(0/2\, V_2(c),\ 0/2\, V_2(d)) = 0/2$

$V_3(e) = 1$

$V_4(a) = 10$

$V_4(b) = \max(0/2\, V_3(a),\ 0/2\, V_3(c))$

$V_4(c) = \max(0/2\, V_3(b),\ 0/2\, V_3(d))$

$V_4(d) = \max(0/2\, V_3(c),\ 0/2\, V_3(e))$

$V_4(e) = 1$

$V_5$ مشابه حالت چهارم است تغییری نداریم.

سوال ۴:
$V_k(A) = 0/4 \qquad V_k(B) = 1,4 \qquad V_k(C) = 2,16$

الف)

$$V_{k+1}(s) = \max_a \sum T(s,a,s')\left[R(s,a,s') + 0/5\,V_k(s')\right]$$

$$V_{k+1}(A) = \max\Big(1(0 + 0/5\,V_k(B)),\ 1(-2 + 0/5\,V_k(C))\Big) = 0/7$$

$$V_{k+1}(B) = \max\Big(\big(0/4(-1 + 0/5\,V_k(A)) + 0/6(2 + 0/5\,V_k(C))\big),$$
$$\big(0/6(2 + 0/5\,V_k(A)) + 0/6(-1 + 0/5\,V_k(C))\big)\Big) = 1,44$$

$$V_{k+1}(C) = \max\Big(\big(0/6(2 + 0/5\,V_k(A)) + 0/4(2 + 0/5\,V_k(B))\big),$$
$$\big(0/4(2 + 0/5\,V_k(A)) + 0/6(0 + 0/5\,V_k(B))\big)\Big) = 2,4$$

$V^*(A) = 0/881 \qquad V^*(B) = 1,761 \qquad V^*(C) = 2,616$

ب)

$$Q^*(A, \text{clockwise}) = 1(0 + 0/5\,V^*(B)) = 0/88$$

$$Q^*(A, \text{counterclockwise}) = 1(-2 + 0/5\,V^*(C)) = -0/49$$

ج)

$$\pi^*(s) = \arg\max Q^*(s,a) = \arg\max\big(Q^*(A,CW),\ Q^*(A,CCW)\big)$$

$= \text{clockwise (CW)}$