2023 may 3

Wednesday, May 3, 2023 13:04

	1	2
1	-.04	-.04
2	-1	1
3	0	-.04

Value Iteration

$U_0(s) = 0$

$U_i(s) = R(s) + \gamma \max_a \sum_{s'} P(s'|s,a)\, U_{i-1}(s')$

$U_1(s) = R(s)$

$U_2((2,1)) = -1$

$U_2(2,2) = +1$

$U_2(3,2) = -0.04 + \max_a \sum_{s'} P(s'|s,a)\, U_1(s')$

$= -.04 + \max(\underbrace{-0.04}_{U_1(3,2)},$

$\underbrace{(0.8)\cdot 1 + (0.2)(-0.04)}_{0.8\, U_1(2,2) + 0.2\, U_1(3,2)},$

$(0.8)\cdot 0 + (0.2)(-0.04)$

$0.8\, U_1(3,1) + 0.2\, U_1(3,2)$

$$= -0.04 + 0.8 + -0.008 = U_2(3,2)$$

$$U_2(3,1) = 0 + \max_a \left(\sum_{s'} P(s' \mid s, a)\, U_1(s') \right)$$

$$= 0 + \sum_{s'} P(s' \mid s, \text{STAY})\, U_1(s')$$

$U_2(3,1) = 0$ IF MOVE TOWARD WALL OR STAY HERE ARE ALLOWED ACTIONS

...

$$U_2(1,1) = -0.04 + \max_a \left(\sum_{s'} P(s' \mid s, a)\, U_1(s') \right)$$

$$= -0.04 + (0.8)(-0.04) + (0.2)(-0.04)$$

$U_2(1,1) = -0.08$

$$U_2(1,2) = -0.04 + \max_a \left(\sum_{s'} P(s' \mid s, a)\, U_1(s') \right)$$

$U_2(1,2) = -0.04 + (0.8)\cdot 1 + (0.2)(-0.04)$

(b) 2^{nd} iteration: $U_2(3,2)$ becomes positive

3^{rd} iteration:

$$U_3(3,1) = R(3,1) + \max_a \left(\sum_{s'} P(s'|s,a)\, U_2(s') \right)$$

WILL CHOOSE a = RIGHT

& WILL OBTAIN POSITIVE $U_3(3,1)$

ALTERNATE EXPLANATION

$U_i(s)$ CAN INCLUDE TOTAL REWARDS FOR PATH OF

...   ...

$i-1$ STEPS

& SHORTEST PATH FROM (3,1) TO POS. REWARD IS 2 STEPS

$\Rightarrow U_3(s)$ CAN BE POSITIVE

REVIEW EXAM 1 Q18

$ax + by + cz = d$, $y = g$ ←

$ax + by + cz = e$, $y = g$ ←

PINHOLE CAMERA

$$\frac{x'}{f} = -\frac{x}{z}, \quad \frac{y'}{f} = -\frac{y}{z}$$

$$\frac{y'}{f} = -\frac{g}{z} \qquad \lim_{z\to\infty} \frac{y'}{f} = \lim_{z\to\infty} -\frac{g}{z} = 0$$

VANISHING POINT $y' = 0$

$$\frac{x'}{f} = -\frac{x}{z} = -\frac{(d - by - cz)/a}{z}$$

$$= -\frac{(d - bg - cz)/a}{z}$$

$$\lim_{z\to\infty} \frac{x'}{f} = \lim_{z\to\infty} -\left(\frac{d - bg - cz}{za}\right)$$

$$= \frac{c}{a}$$

$$x' = \frac{cf}{a}$$

VANISHING PT $(x', y') = \left(\frac{cf}{a}, 0\right)$

KALMAN FILTER PROBLEM 1

Δ = VELOCITY × 1 SECOND

$$\mu_\Delta = v, \quad \sigma^2_\Delta = p^2$$ } CHANGE IN STATE

$$\mu_x = x, \quad \sigma^2_x = e^2$$ } OBSERVATIONS

$$\mu_{t|t} = y, \quad \sigma^2_{t|t} = s^2$$ } STATE ESTIMATE AT t

$$\mu_{t+1|t} = \mu_{t|t} + \mu_\Delta = y + v$$

$$\sigma^2_{t+1|t} = \sigma^2_{t|t} + \sigma^2_\Delta = s^2 + p^2$$

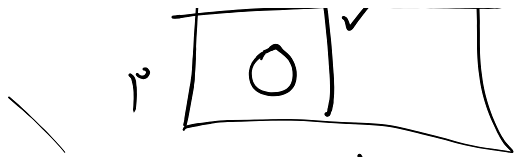
$$\sigma_{t+1|t} = \sqrt{s^2 + p^2}$$

REVIEW EXAM 3, Q19

Q S

$Q_0(s,a) = 0$

	1	2
W	✓	0



TIME STEP 1

$$Q_{local}(1, w) = r_0 + \max_{a'} \gamma \cdot \underbrace{Q_0(2, a)}_{\text{all } 0}$$

$$= r_0 = 2$$

TIME STEP 2

$$Q_{local}(\underset{\uparrow s_1}{2}, p) = r_1 + \max_{a'} \gamma \cdot \overbrace{Q_1(\underset{\uparrow s_2}{2}, a)}^{\text{all } 0}$$

$$Q_{l}(2, p) = 5$$

local

$$Q_2(s, a) = Q_1(s, a) + 0.05(Q_{local} - Q_1)$$

$$Q_2(1, w) = 0.05 \times 2$$

$$Q_2(2, p) = 0.05 \times 5$$

---

[b]

$$P(s' \mid s = 2, a = pull)$$

$$= \frac{\#times + k}{\#trials + k(\#poss)} \qquad s' = 1$$

$$= \begin{cases} \dfrac{\#\text{times} + k}{\#\text{trials} + k(\#\text{pos})} & s' = 2 \end{cases}$$

$$= \begin{cases} \dfrac{0+1}{1+2} & s' = 1 \\ \dfrac{1+1}{1+2} & s' = 2 \end{cases}$$

1c

$$U^{\pi}(1) = R(1) + \gamma \sum_{s'} P(s' \mid 1, p)\, U^{\pi}(1)$$

~~$$U^{\pi}(2) = R(2) + \gamma \sum_{s'} P(s' \mid 2, w)\, U^{\pi}(2)$$~~

$$U^{\pi}(1) = 2 + \left(\frac{3}{4}\right)\left(\frac{2}{3} U^{\pi}(1) + \frac{1}{3} U^{\pi}(2)\right)$$

$$U^{\pi}(2) = 5 + \left(\frac{3}{4}\right)\left(\frac{2}{3} U^{\pi}(1) + \frac{1}{3} U^{\pi}(2)\right)$$

1d

CRITIC NETWORK: s, a

s: 2D, a: 2D, INPUT = 4D

OUTPUT = 1D

ACTOR = PROBABILITY OF CHOOSING ACTION

OUTPUT: 2D (ACTIONS)

INPUT: 2D (STATE)