

« به نام خدا »

امیررضا ولیپور

۹۹۵۱۲۲۲۱

تکلیف ۲ / ارزش



فروردین

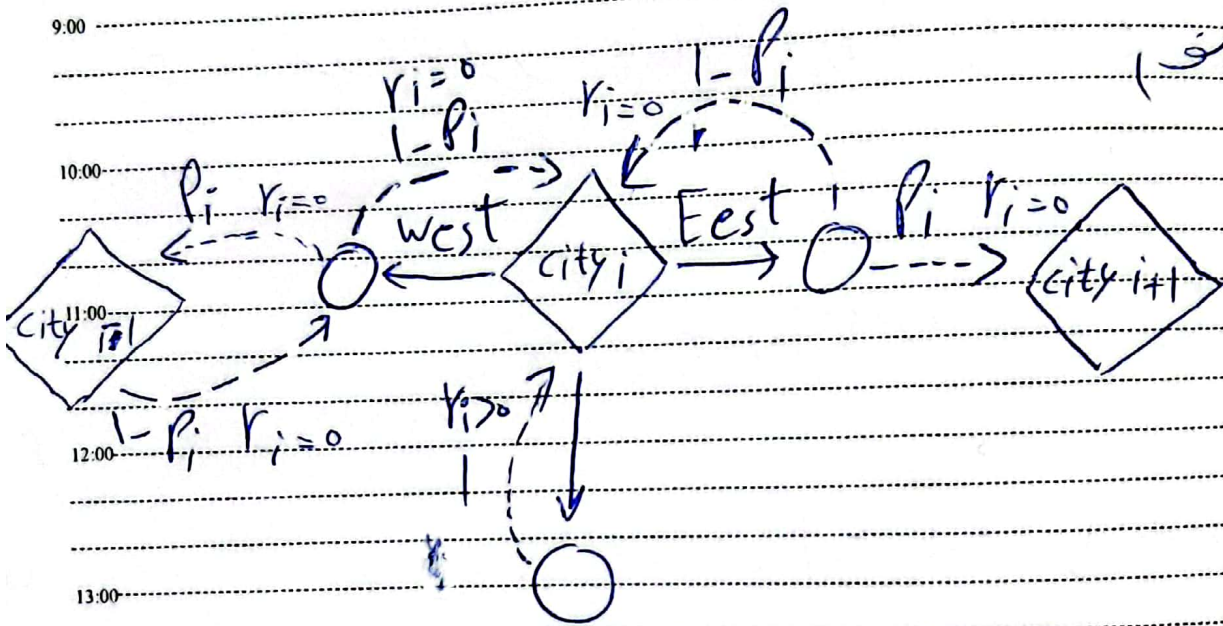
یکشنبه

۴ رمضان ۱۴۴۴

2023 Mar 26

(1)

الف)



$$V^{s, \pi^*}(i) = \max_a \sum_{s'} T(s, a, s') [R(s, a, s') + \gamma V^{s, \pi^*}(s')]$$

$$= 1 + \frac{1}{2} + \frac{1}{2} + \dots = \frac{1}{1 - \frac{1}{2}} = 2$$

روز دوشنبه

چو کمل پیش ما خاک آستان شاست کجا رویم بغیرا از این جناب کجا

۷  
۸.۰۰ (ب) بارش هر مثل سر در قیل بر ۲ هکتار است

۹.۰۰ (ث) برویم بر او state که ۲ اش از همه بزرگ است  
۱۰.۰۰ stay کنیم

۱۱.۰۰  
۱۲.۰۰  
(ث)  
$$V_{cs}^{\pi} = (1-\alpha) V_{cs}^{\pi} + \alpha [R_{cs}(\pi(s), s') + \gamma V_{s'}^{\pi}]$$

۱۳.۰۰  
sample (stay, ۱)  $\alpha = \frac{1}{T}$

۱۴.۰۰  
$$Q(1, stay) = (1-\alpha)0 + \alpha \left[ 5 + \frac{1}{T} \times 0 \right] = 2$$

۱۵.۰۰  
$$Q(1, East) \neq Q(2, East) \neq Q(1, East)$$

۱۶.۰۰  
$$\gamma Q(2, stay) = 0$$

۱۷.۰۰  
sample (East, ۲)

~~$$Q(1, stay) \neq Q(2, East) \neq Q(2, stay) = 0$$~~

$$Q(1, East) = (1-\alpha)0 + \frac{1}{T} \left[ 0 \times \frac{1}{T} + 0 \right] = 0$$

روز هنرهای نمایش

مبین به سبب زخمندان که چاه در راه است کجای روی ای دل بدین شتاب کجا  
$$Q(1, stay) = 2$$





سه شنبه

2023 Mar 28

$$\text{sample} = R(\text{stay}) + \gamma_m Q(\text{stay})$$

$$Q(\text{stay}) = (1-\alpha) Q(\text{stay}) + \alpha(\text{sample})$$

فرض

8:00

sample 3: ~~stay~~

9:00

$(R, \text{stay}, \gamma, \alpha) \rightarrow$

10:00

فرض

$$Q(R, \text{stay}) = (1-\alpha) \cdot 0 + \alpha \left( \gamma \times \frac{1}{\gamma} + 0 \right)$$

$$= 1$$

11:00

sample 4:  $(R, \text{west}, \gamma, \alpha)$

12:00

$$Q(R, \text{west}) = \left(1 - \frac{1}{\gamma}\right) \cdot 0 + \frac{1}{\gamma} \left(0 + \frac{1}{\gamma} \times 2\right) = \frac{1}{\gamma}$$

13:00

sample 5:  $(R, \text{stay}, \gamma, \alpha)$

14:00

$$\left(1 - \frac{1}{\gamma}\right) \times 2 + \frac{1}{\gamma} \left(1 + \frac{1}{\gamma} \times 2\right) = 1, 5$$

15:00

16:00

17:00

بشد که یاد خوش باد روزگار وصال خود آن کرشمه کبارفت و آن عتاب کبار

9

چهارشنبه

فروردین

2023 Mar 29

۷ رمضان ۱۴۴۴

	$Q_{(b, \text{stay})}$	$Q_{(b, \text{East})}$	$Q_{(b, \text{West})}$	$Q_{(b, \text{Stay})}$
Sample 1	۲	۰	۰	۰
Sample 2	۲	۰	۰	۰
Sample 3	۲	۰	۰	۳
Sample ۴	۲	۰	$\frac{1}{2}$	۳
Sample ۵	$\frac{۳}{۴}$	۰	$\frac{1}{2}$	۳

8.00

9.00

10.00

11.00

12.00

13.00

14.00

15.00

16.00

17.00

قرار و خواب ز عاظمه ای دوست قرار پست صودی کدام و خواب کجا



از یایی که ما

8:00

به این علت دقیقاً هر بار در لحظاتی که ما می‌خواهیم

10:00

که house دارد

در کل در طی انجام مراحل نقطه ایی مرور

11:00

و یادگیری سریع تر به مقصد می‌رساند و جمعاً با اینها

12:00

13:00

من کنه

در حین یادگیری و در  $\text{step delay}$

$\text{learning rate}$  سرعت یادگیری  
Learning rate  
که در حین یادگیری  
بالاتر نمونه‌های جدید برای ما مهم‌تر می‌شوند

اگر آن ترک شیرازی به دست آرد دل ما را به خال هندویش بخشم سمرقند و بخارا را

هر چه بیشتر باشد حرکت  $\rightarrow$  Epsilon  
random بیشتر می شود و هر چه کمتر باشد optime  
و بر اساس تجربه کار می کند

کسین گام ها را که Discount  
من نویسد Value ها را نزدیک ارزش بیشتر می دارند  
هر چه کمتر باشد و ربات ارزش را از حرکت می کند  
و هر چه بیشتر باشد با توجه به کارهای بزرگتری می تواند



AVERAGE RETURNS FROM START STATE: 0.31887829542225943

```

values here..."
__init__(self, **args)
__enter__()

```

```
action):
```

```

    Actions(state):
        action)]

```

```
(self, state):
```