علیرضا قنبری

۶۱۰۳۹۶۱۳۰

گزارش هوش مصنوعی:

اول، ما یک لیست را ایجاد می کنیم تا تمامی پاداش هایی که ما از هر قسمت دریافت می کنیم را نگه داریم. این امر می تواند نشان دهد که چگونه نمره بازی ما در طول زمان تغییر می کند.

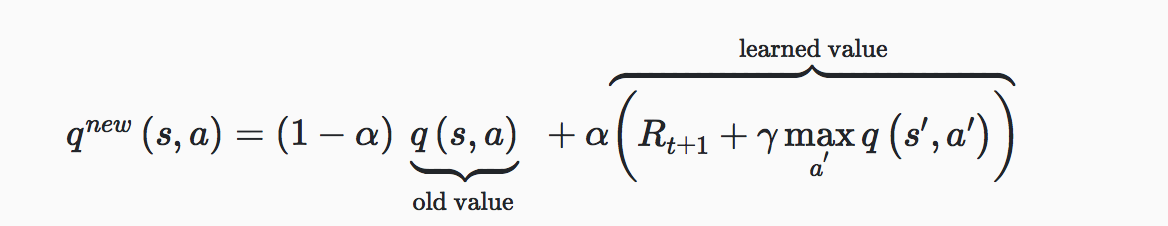
**for** episode **in** range(n):  
 state = env.reset()  
 done = **False** Rnow = 0

متغیر "done" فقط پیگیری می کند که قسمت ما به پایان رسیده است یا خیر، بنابراین ما ابتدا آن را به False قرار می دهیم و بعد زمانی که قسمت اول را شروع می کنیم، آن را دریافت می کنیم و زمانی که eposide پایان می یابد به ما اطلاع می دهد.

سپس، ما باید پاداش ها در قسمت فعلی نگه داریم، بنابراین ما Rnow را به 0 اضافه می کنیم، زیرا ما در ابتدای هر قسمت هیچ پاداشی شروع نداریم.

به روز رسانی Q-value

پس از مشاهده پاداشهایی که ما از از state قبلی دریافت کردیم، می توانیم Q-value را برای جفت state-actionدر جدول Q به روز کنیم. این کار با استفاده از فرمول زیر انجام شده است:



نکته:Q-value برای خانه های hole و goal همیشه صفر است زیرا وقتی به این خانه برسیم عملیات باید دوباره از سر گرفت

طبق این فرمول در خانه های کنار hole عملی انجام می شود که در خانه نیوفتد و در خانه های کنار goal برعکس .

زیرا زیرا طبق این فرمول وقتی به خانه hole بیفتیم مقدار value در q-table کاهش می یابد بنابراین احتمال این که این کار را انجام بدهد کم می شود برای goal عکس این عمل اتفاق می افتد.

بعد، ما وضعیت کنونی خود را به new\_state تنظیم می کنیم که بعد از انجام آخرین فعالیت ما به ما بازگردانده شد و پس از آن پاداش هایی را که ما برای اقدام قبلی خود دریافت کرده ایم به روز می کنیم.

سپس ما بررسی می کنیم که آیا آخرین action ما state ما را به پایان رسانده است یا خیر، آیا state به سوراخ یا به هدف رسیده است ؟

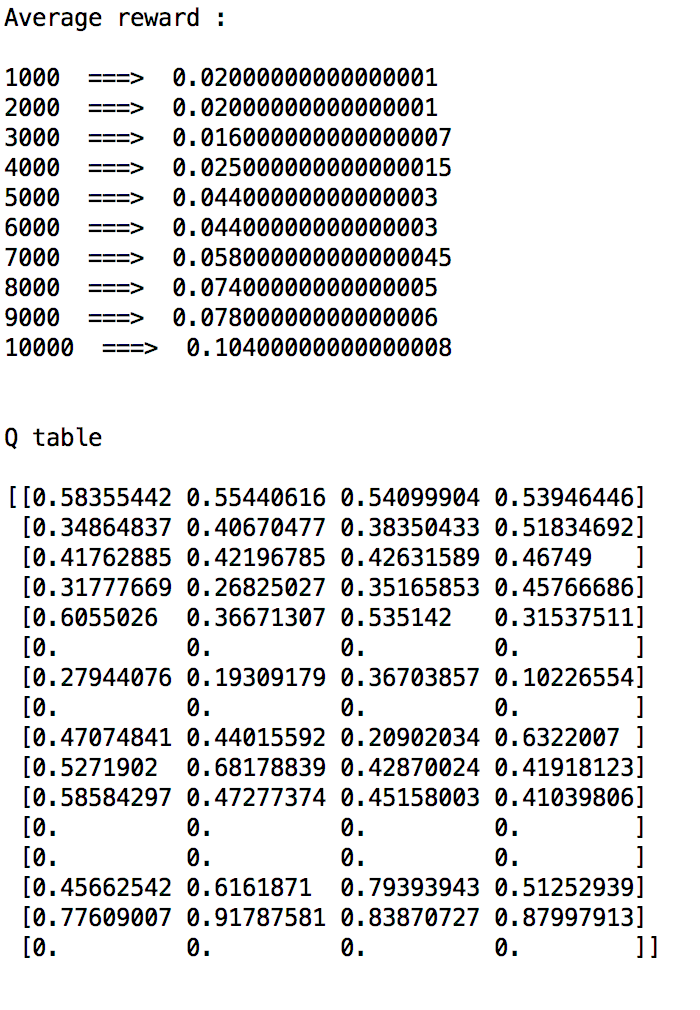
# Exploration rate decay

exploration\_rate = min\_exploration\_rate + \

(max\_exploration\_rate - min\_exploration\_rate) \* np.exp(-exploration\_decay\_rate\*episode)

هنگامی که یک قسمت به پایان رسید، ما نیاز به به روز رسانی exploration\_rate با استفاده از واپاشی نمایی، که بدان معناست که میزان کاهش اکتشاف یا واپاشی اش متناسب با ارزش فعلی آن است. ما می توانیم exploration\_rate را با استفاده از فرمول بالا بالا ببریم، که از تمام پارامترهای نرخ اکتشاف که ما در آخرین زمان تعریف کرد استفاده می کنیم.

پس از ان که تمام episode ها تمام شد ، ما اکنون فقط پاداش متوسط ​​هر هزار قسمت از لیستمان را محاسبه می کنیم که شامل پاداش برای تمام قسمت ها است تا ما بتوانیم آن را چاپ کنیم و ببینیم که پاداش ها در طول زمان تغییر کرده است.



4\*4

8\*8

Average reward :

1000 ===> 0.001

2000 ===> 0.002

3000 ===> 0.002

4000 ===> 0.014000000000000005

5000 ===> 0.017000000000000008

6000 ===> 0.02900000000000002

7000 ===> 0.04500000000000003

8000 ===> 0.06300000000000004

9000 ===> 0.06400000000000004

10000 ===> 0.08900000000000007

Q table

[[4.31926428e-01 4.32660694e-01 4.32805885e-01 4.33887823e-01]

[4.40386403e-01 4.42780244e-01 4.51238609e-01 4.42136463e-01]

[4.55525265e-01 4.61933777e-01 4.71717604e-01 4.59100615e-01]

[4.79340340e-01 4.96535693e-01 5.02767827e-01 4.85849277e-01]

[5.01712763e-01 5.07094676e-01 5.16887337e-01 5.16926035e-01]

[5.29342419e-01 5.38953209e-01 5.41803277e-01 5.35986553e-01]

[5.54846312e-01 5.58021830e-01 5.64957474e-01 5.51400457e-01]

[5.69866473e-01 5.64758470e-01 5.70717206e-01 5.60248288e-01]

[4.27197471e-01 4.30458848e-01 4.29889609e-01 4.34172019e-01]

[4.34872939e-01 4.34157733e-01 4.39342680e-01 4.44314959e-01]

[4.34340016e-01 4.26159630e-01 4.48049208e-01 4.62781347e-01]

[3.05002028e-01 3.35898255e-01 3.28870828e-01 4.88843173e-01]

[4.62312626e-01 4.75684433e-01 4.82640279e-01 5.04514043e-01]

[5.13089495e-01 5.28778834e-01 5.39671440e-01 5.27487915e-01]

[5.58978250e-01 5.68057674e-01 5.74764711e-01 5.56955021e-01]

[5.76659646e-01 5.87444930e-01 5.78501245e-01 5.70532311e-01]

[4.03751827e-01 3.94585711e-01 4.08443988e-01 4.19782307e-01]

[4.09235404e-01 3.80543400e-01 3.93542464e-01 4.18159623e-01]

[3.78016658e-01 2.54627327e-01 2.26731795e-01 3.78703492e-01]

[0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00]

[2.64878673e-01 2.94387080e-01 4.65056615e-01 3.63859469e-01]

[3.09969901e-01 3.78769820e-01 4.14958122e-01 5.14339877e-01]

[5.53787142e-01 5.69189135e-01 5.94596345e-01 5.58092954e-01]

[6.01887692e-01 6.08674802e-01 6.04433419e-01 5.94389942e-01]

[3.57712357e-01 3.48803246e-01 3.69091263e-01 3.79123141e-01]

[3.38352682e-01 3.33940699e-01 3.36331065e-01 3.75648248e-01]

[2.79913686e-01 2.42432496e-01 2.32099816e-01 3.15626812e-01]

[1.47429661e-01 2.44381595e-01 6.00409228e-02 1.51281138e-01]

[3.15731352e-01 1.96388919e-01 1.65471568e-01 2.52633318e-01]

[0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00]

[3.96462831e-01 4.65329247e-01 6.07857878e-01 3.52595356e-01]

[6.32828585e-01 6.37264575e-01 6.38310596e-01 6.18313981e-01]

[3.03844950e-01 2.63495465e-01 2.80116453e-01 3.42400100e-01]

[1.95452249e-01 1.75476518e-01 1.88707527e-01 3.22182981e-01]

[1.80684030e-01 8.26055634e-02 9.77237673e-02 1.31814371e-01]

[0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00]

[1.83735657e-01 1.84268949e-01 3.03605494e-01 1.84975512e-01]

[1.31047988e-01 3.50644958e-01 1.80872604e-01 2.47775312e-01]

[4.27826160e-01 2.60365657e-01 4.55458387e-01 5.36932629e-01]

[6.10852258e-01 6.63554960e-01 6.81947304e-01 6.26060591e-01]

[2.45865916e-01 5.95985954e-02 1.07189207e-01 1.12228214e-01]

[0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00]

[0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00]

[6.97488896e-09 3.00266093e-03 2.59512460e-02 4.34901402e-02]

[8.58599250e-02 4.42515228e-02 1.21483589e-01 2.09026253e-01]

[2.35146434e-01 7.39136431e-02 6.24018763e-02 1.25054739e-01]

[0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00]

[5.50349383e-01 5.76519650e-01 7.75982901e-01 4.53292524e-01]

[1.29856101e-01 2.07543267e-02 3.30824274e-02 3.98013764e-02]

[0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00]

[3.71251183e-07 0.00000000e+00 3.54399130e-07 3.52366814e-09]

[2.14833030e-08 4.94363450e-08 0.00000000e+00 1.94074958e-03]

[0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00]

[1.78886056e-02 1.53287165e-03 3.34365394e-02 5.40177996e-02]

[0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00]

[5.96375308e-01 6.53371214e-01 8.86960014e-01 7.13755054e-01]

[7.82242558e-02 4.04306687e-03 4.22545244e-03 2.64981622e-02]

[2.28796917e-08 1.01871893e-02 4.98267995e-05 3.60976837e-04]

[2.15048079e-08 4.63548906e-05 3.86699455e-08 0.00000000e+00]

[0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00]

[0.00000000e+00 0.00000000e+00 8.81207910e-04 8.90804264e-04]

[5.61419824e-03 4.04725072e-02 3.04971748e-03 0.00000000e+00]

[0.00000000e+00 0.00000000e+00 1.71000000e-01 0.00000000e+00]

[0.00000000e+00 0.00000000e+00 0.00000000e+00 0.00000000e+00]]