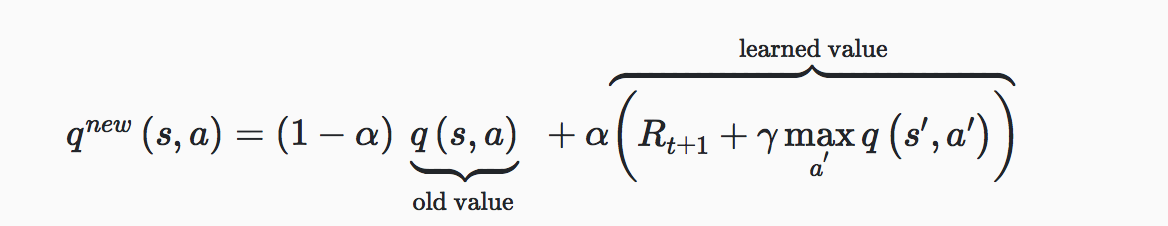
صالح زارع زاده

شماره دانشجویی : ۶۱۰۳۹۶۱۰۹

پروژه هوش مصنوعی :

1. بازی را برای ۱۰۰۰۰ بار اجرا کردم و یک table مناسب بدست آوردم که win rate خوبی دارد و آن را برای ۱۰۰۰ بار دیگر با استفاده از table به دست امده اجرا کردم تا win rate table را بدست آورم.
2. Win rate بدست آمده از کد رندوم بهتر است و برای ۴\*۴ حدود ۷۰ درصد و برای ۸\*۸ حدود ۳۰ درصد است.(برای اجراهای مختلف فرق میکند)
3. با استفاده از یک جدول که برای ۴\*۴ ؛ ۴\*۱۶ و برای ۸\*۸ ؛ ۴\*۶۴ است که درواقع همان تعداد خانه ها در تعداد عمل ها (راست و چپ و بالا و پایین) است که یا این عمل به صورت رندم انتخاب میشود(با استفاده از randome در کد) یا بهترین عملی که میتوان انجام داد را که با استفاده از ماکسیمم گیری بین خانه های هر استیت بدست می آید بدست می اوریم و ان کار را در آن استیتی که هستیم انجام میدهیم و هر خانه با استفاده از فرمول زیر آپدیت میشود(در قسمت اول کار که میخواهیم table را بسازیم) که درواقع از معادله ی بلمن استفاده کرده ایم : q∗(s,a)=E[Rt+1 + γ maxa′q∗(s′,a′) ] که میتوان آن را به فرمول زیر برای حل این مسيله تبدیل کرد :   
4. Q-value برای خانه های Hole و Goal هم در ابتدا و در انتها و در کل عملیات صفر است چون زمانی که وارد این خانه ها شویم باید از اول کار را شروع کنیم و نیازی به Q-value ای برای این خانه ها نداریم .
5. در فرمول qnew از reward استفاده شده است پس با تغییر reward ؛ table تغییر میکند در نتیجه ممکن است بهترین عملی که در ان خانه میتواند انتخاب کند تغییر کند بنابراین فرایند تغییر میکند.
6. کنار Hole عملی را انجام میدهد که ان را در Hole نیندازد زیرا ان عمل value کمتری دارد زیرا در فرمول حساب کردن Q-value که برای هر خانه table است از این فرمول استفاده میشود که در بخش ۳ به ان اشاره شد :

table[state, action] = q\_table[state, action] \* (1 - learning\_rate) + \

learning\_rate \* (reward + discount\_rate \* numpy.max(q\_table[new\_state, :])

و درواقع رفتن به خانه hole باعث میشود مقدار ان عمل ضرب در 0.9 شود(زیرا در کد مقدار Lrate ؛0.1 است) بنابراین ارزش عملی که بیشتر وارد Hole میشود کم میشود بنابراین احتمال این که آن عمل را انجام دهد کم میشود و در کنار خانه Goal عملی را انجام میدهد که آن را در Goal بیندازد زیرا عملی که آن را وارد Goal میکند را ضرب در 0.9 و به علاوه 0.1 میکند و عملی که ان را وارد خانه دیگری میکند را ضرب در 0.9 و به علاوه عددی کمتر از 0.1 میکند زیرا همه خانه های جدول از یک کمتر است و reward برای همه خانه ها به جز Goal صفر است بنابراین یک عدد کمتر از 0.1 بدست میاید .

هر اپیزود که به پایان میرسد باید exploration\_rate را آپدیت کنیم زیرا باید میزان اکتشاف آن متناسب با ارزش فعلی باشد درواقع یعنی احتمال استفاده از تابع رندم را در برابر استفاده از table متناسب میکنیم که بااستفاده از فرمول زیر این کار را انجام میدهیم :

exploration\_rate = min\_exploration\_rate +(max\_exploration\_rate - min\_exploration\_rate) \*numpy.exp(exploration\_decay\_rate\*episode)

حال این table و win ratio برای ۴\*۴ است:(یک مثال)

: Table

[[0.58394144 0.53637955 0.5377275 0.54381779]

[0.31730743 0.2845755 0.30635541 0.54049709]

[0.40174789 0.39702123 0.39892721 0.51620126]

[0.28354008 0.19987657 0.32289374 0.49334465]

[0.60744535 0.45241885 0.47673803 0.42517164]

[0. 0. 0. 0. ]

[0.18549026 0.16167132 0.32094455 0.08811163]

[0. 0. 0. 0. ]

[0.52914853 0.34292257 0.41807025 0.63350601]

[0.47355892 0.66021431 0.44913222 0.38728568]

[0.62563637 0.42190239 0.38928024 0.22014022]

[0. 0. 0. 0. ]

[0. 0. 0. 0. ]

[0.50609238 0.40839831 0.76580294 0.52863001]

[0.72738535 0.90284839 0.75838329 0.74768317]

[0. 0. 0. 0. ]]

win ratio for 4\*4: 0.719

حال این table و win ratio برای ۸\*۸ است:(یک مثال)

: Table

[[4.21719326e-02 4.26678176e-02 4.24221065e-02 4.26443635e-02]

[4.31755145e-02 4.41368721e-02 4.42090252e-02 4.36303971e-02]

[4.48688139e-02 4.62578562e-02 4.63698101e-02 4.69698805e-02]

[4.84169641e-02 4.91769673e-02 5.00185089e-02 5.06767073e-02]

[5.09090199e-02 5.47376090e-02 5.60185490e-02 5.70394277e-02]

[6.00240346e-02 6.56738984e-02 6.64350501e-02 6.37084193e-02]

[6.72667451e-02 6.83277391e-02 7.30718579e-02 6.83804114e-02]

[7.24071442e-02 7.45317571e-02 7.58136094e-02 7.12231766e-02]

[4.13560463e-02 4.14502864e-02 4.14109281e-02 4.23936611e-02]

[4.17521380e-02 4.26685753e-02 4.31440935e-02 4.31929415e-02]

[4.08211324e-02 4.06529111e-02 4.37608537e-02 4.48958047e-02]

[3.69192024e-02 3.61543389e-02 3.49610384e-02 4.86024742e-02]

[4.70871623e-02 4.85963627e-02 4.98889611e-02 5.44452826e-02]

[5.43531778e-02 5.76747086e-02 6.40181730e-02 6.50007603e-02]

[6.78606122e-02 7.16947337e-02 7.73651121e-02 7.28168160e-02]

[8.10890943e-02 8.38640251e-02 8.29770404e-02 7.62771416e-02]

[3.92023433e-02 3.81231017e-02 3.90901475e-02 4.01243262e-02]

[3.76639515e-02 3.63960159e-02 3.69723849e-02 3.93324718e-02]

[3.76408206e-02 2.00793982e-02 2.29543103e-02 3.04243985e-02]

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

[1.91323733e-02 2.53595935e-02 3.96310662e-02 3.50806741e-02]

[3.29525735e-02 2.03245455e-02 4.98817652e-02 6.09660666e-02]

[6.22068833e-02 7.49670819e-02 8.49954972e-02 7.72926046e-02]

[8.94173280e-02 1.02337794e-01 1.09750244e-01 8.18702531e-02]

[3.22945573e-02 3.26571241e-02 3.34218990e-02 3.61635328e-02]

[3.45922571e-02 2.87153308e-02 3.23151366e-02 3.46328037e-02]

[2.67765573e-02 2.18492684e-02 2.01642338e-02 2.80799694e-02]

[1.01904056e-02 1.88387873e-02 9.31934389e-03 9.65706332e-03]

[2.14715593e-02 5.07006795e-03 1.16806341e-02 1.31770580e-02]

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

[3.78285239e-02 7.51028763e-02 7.97402157e-02 4.72620203e-02]

[9.52151344e-02 1.29049135e-01 1.38183516e-01 9.81966680e-02]

[2.70557644e-02 2.07933436e-02 2.43610564e-02 2.86553547e-02]

[2.27165822e-02 1.21552686e-02 1.38777860e-02 2.48829835e-02]

[1.37295127e-02 8.17172633e-03 1.17903234e-02 1.89884382e-02]

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

[5.00309788e-03 5.85436463e-03 7.31723742e-03 4.65786003e-03]

[1.37363935e-03 2.24416873e-02 2.64503236e-02 1.47407912e-02]

[2.41220632e-02 4.91219648e-02 6.81905169e-02 3.46566921e-02]

[2.07237949e-01 1.79819170e-01 1.77776607e-01 9.59329287e-02]

[1.56394445e-02 7.82395612e-03 7.41225496e-03 1.11924514e-02]

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

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

[0.00000000e+00 2.28181409e-04 4.52174419e-05 0.00000000e+00]

[2.85148134e-03 1.10947109e-03 1.63305928e-03 1.45944735e-03]

[4.58439323e-03 1.70685927e-04 2.22490800e-03 1.22935744e-03]

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

[2.35791881e-01 2.22806118e-01 3.54349730e-01 9.62408227e-02]

[8.45012808e-03 2.67215491e-03 4.92813355e-03 9.05930396e-03]

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

[5.62122254e-07 2.92151769e-06 1.81926713e-05 8.95526825e-09]

[7.04171071e-10 0.00000000e+00 1.04576047e-05 5.81572215e-06]

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

[0.00000000e+00 0.00000000e+00 2.50617455e-04 9.90378087e-06]

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

[2.18252092e-01 3.24246330e-01 5.36547184e-01 1.66669257e-01]

[2.55385649e-03 1.83815762e-03 2.11994279e-03 3.39171118e-03]

[1.14114833e-03 2.92656034e-04 5.97204236e-05 4.16342403e-04]

[4.20765602e-05 1.37359209e-04 1.70127140e-05 5.52654121e-06]

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

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

[0.00000000e+00 0.00000000e+00 9.80474306e-07 0.00000000e+00]

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

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

win ratio for 8\*8: 0.266