Reinforcement Learning

Q-Learning

Luccas Rojas 201923052 31/03/2023

1. Sobre el ambiente gridworld se realizó un algoritmo de Q-leaning. Para este caso se utilizaron un factor de descuento de 0.9, un factor de aprendizaje de 0.8, un ruido de 0.2 y un épsilon de 0.7. Estas variables algunas fueron de forma arbitrarias, otras fueron para que el agente tuviera la capacidad de estar en una etapa de exploración antes de pasar a explotar como tal el algoritmo. A continuación se presenta el resultado de la Q-tabla, junto con la política optima, luego de 139 iteraciones, para las cuales ya el algoritmo converge hacia una misma política.

Imagen de la pantalla de un celular de un mensaje en letras negras

Descripción generada automáticamente con confianza baja

A continuación se muestran los valores de la Q-Tabla para cada celda y cada acción posible.

[[{'up': 0.10443021205343916, 'down': 0.10972326457028987, 'left': 0.09414122558754313, 'right': 0.11787798334131512}, {'up': 0.15399302476148216, 'down': 0.07437899027427183, 'left': 0.09797696039153214, 'right': 0.13616453624969543}, {'up': 0.12697372413975183, 'down': 0.19894736169638635, 'left': 0.12044823105993946, 'right': 0.14932857160262752}, {'up': 0.12223564836419865, 'down': 0.1888913713770209, 'left': 0.14080880306861757, 'right': 0.23939454697194554}, {'up': 0.25001012413330426, 'down': 0.2837691760068538, 'left': 0.20948505260620898, 'right': 0.2814374520684926}, {'up': 0.2359833458664351, 'down': 0.33118489449119637, 'left': 0.23672788312450005, 'right': 0.24757595540107943}, {'up': 0.2552953744386597, 'down': 0.29576564503823605, 'left': 0.2736990980694063, 'right': 0.22990302033065613}, {'up': 0.2353660076519746, 'down': 0.25331208450731424, 'left': 0.2585414327354688, 'right': 0.22792841158551563}, {'up': 0.20796088386183292, 'down': 0.23644698988519103, 'left': 0.19456347454167908, 'right': 0.2574307176264662}, {'up': 0.2177314563445936, 'down': 0.24517286635815125, 'left': 0.23009552503023614, 'right': 0.213248296957124}], [{'up': 0.05520161549038879, 'down': 0.09660909613412023, 'left': 0.08412099033097681, 'right': 0.058849433305977174}, {'up': 0.1256389455778368, 'down': 0.07723192777766198, 'left': 0.08363866471900622, 'right': 0.10726656635786386}, {'up': 0.09186930441116226, 'down': 0.1259240523501707, 'left': 0.09223316727640099, 'right': 0.1860158988471393}, {'up': 0.27266504246145185, 'down': 0.24195001917898779, 'left': 0.196422354763098, 'right': 0.2872319991691176}, {'up': 0.24931613026543273, 'down': 0.2890021586895789, 'left': 0.2551664498686162, 'right': 0.3289377916299479}, {'up': 0.29802998781180845, 'down': 0.3679963050664923, 'left': 0.3302217258709293, 'right': 0.28687954212881256}, {'up': 0.23190334284414557, 'down': 0.2895811252283016, 'left': 0.329316570300268, 'right': 0.22817022672005025}, {'up': 0.23144625166196975, 'down': 0.28286782610390326, 'left': 0.2896093835686434, 'right': 0.25209880203115087}, {'up': 0.285582272402606, 'down': 0.24322901587796583, 'left': 0.2595859606155131, 'right': 0.27951108432563626}, {'up': 0.21281205443099852, 'down': 0.33357270295350694, 'left': 0.2563701686535534, 'right': 0.22079011647336652}], [{'up': 0.05847533002658692, 'down': 0.11772597844180782, 'left': 0.08161507655563645, 'right': 0.09378526081447917}, {}, {}, {}, {}, {'up': 0.329045164082724, 'down': 0.37784725031467004, 'left': 0.34725984575966184, 'right': 0.3206278655560593}, {}, {}, {}, {'up': 0.3055011584248673, 'down': 0.3502546699345659, 'left': 0.37068337673548224, 'right': 0.32828697510816623}], [{'up': 0.12382111264786042, 'down': 0.13269290661188454, 'left': 0.0930982179218074, 'right': 0.11424262166858157}, {'up': 0.140978150956583, 'down': 0.12468167888583061, 'left': 0.11745234108310469, 'right': 0.14115375219517978}, {'up': 0.13963588656326498, 'down': 0.14043807078446355, 'left': 0.15590859456106004, 'right': 0.15266365348930386}, {'up': 0.10897703112619171, 'down': 0.14715939236379327, 'left': 0.1415057134107545, 'right': 0.1314966887398518}, {}, {'up': 0.3514622623272563, 'down': 0.15390486811117024, 'left': -0.6463383240018575, 'right': 0.2080340655953424}, {'up': 0.44884223600638257, 'down': 0.5598850735155033, 'left': 0.32201744613800926, 'right': 0.463641908752249}, {'up': 0.42016497036211037, 'down': 0.5295441309183396, 'left': 0.4255096901408252, 'right': 0.43912758222151665}, {'up': 0.43878359521736193, 'down': 0.41958518586236015, 'left': 0.4610373760533973, 'right': 0.3924525729545364}, {'up': 0.3339043382212492, 'down': 0.4483163572929302, 'left': 0.43943286478674803, 'right': 0.3517866842217778}], [{'up': 0.11216385255776523, 'down': 0.13571179991536964, 'left': 0.13274258691214688, 'right': 0.1320369270972186}, {'up': 0.13209899630118985, 'down': 0.13920712614730196, 'left': 0.13316089984413823, 'right': 0.1327292260351121}, {'up': 0.15633678941086834, 'down': 0.16475211567753234, 'left': 0.14508242604763533, 'right': 0.15876362463122695}, {'up': 0.15366416126031718, 'down': 0.17044432619876004, 'left': 0.16111229152008721, 'right': 0.15251845800705036}, {}, {'exit': -1.0}, {'up': 0.5442158764120435, 'down': 0.6555122080334561, 'left': 0.466355701305744, 'right': 0.5419095589867925}, {'up': 0.47705201122674035, 'down': 0.5181247110913273, 'left': 0.5912884228932536, 'right': 0.48207452366994197}, {'up': 0.4390775196315042, 'down': 0.532319141640262, 'left': 0.53099763747438, 'right': 0.43551217054995556}, {'up': 0.4419066907146243, 'down': 0.4746496260735841, 'left': 0.48392803632722026, 'right': 0.4456327379795243}], [{'up': 0.1313772508040854, 'down': 0.16163385183710988, 'left': 0.14575074959931844, 'right': 0.1555160836123881}, {'up': 0.132330495973136, 'down': 0.17630185860372563, 'left': 0.14579954530774455, 'right': 0.16662241971292602}, {'up': 0.15481324864308207, 'down': 0.1814614174940541, 'left': 0.1471540470067454, 'right': 0.17210271682794392}, {'up': 0.14768779224596418, 'down': 0.20556935450648747, 'left': 0.17779533231775707, 'right': 0.17775421115907325}, {}, {'exit': 0.9999999999999979}, {'up': 0.5982363666018442, 'down': 0.5562610180761809, 'left': 0.899520645752095, 'right': 0.6623508061845974}, {'up': 0.5325051984888779, 'down': 0.5277149458426117, 'left': 0.7888659689155285, 'right': 0.5313922051051851}, {'up': 0.6353155009228633, 'down': 0.5117583336033789, 'left': 0.5186495924636002, 'right': 0.477337461302196}, {'up': 0.4571737407471878, 'down': 0.44435079641603104, 'left': 0.4796858842404013, 'right': 0.48035828220877286}], [{'up': 0.15012381711226402, 'down': 0.1728812733507698, 'left': 0.1650212568180672, 'right': 0.1794274490668496}, {'up': 0.15893245513894025, 'down': 0.19863210478928575, 'left': 0.19160600262329286, 'right': 0.17531271815670435}, {'up': 0.19899352945492635, 'down': 0.19524394104263532, 'left': 0.1708331102738873, 'right': 0.17543959700704087}, {'up': 0.17384533812367628, 'down': 0.16186624523108636, 'left': 0.16742928487773232, 'right': 0.1934080916425856}, {}, {'up': 0.8999991410404884, 'down': -0.8384665190399925, 'left': 0.8098809050244733, 'right': 0.6972907266806572}, {'up': 0.7559697098586853, 'down': 0.3447918536761424, 'left': 0.5363989020602488, 'right': 0.5501761769155653}, {'up': 0.6969135334789522, 'down': 0.5489715105202747, 'left': 0.6513848062176957, 'right': 0.5702834430999261}, {'up': 0.5308497163660996, 'down': 0.46426414503570107, 'left': 0.5039995408793913, 'right': 0.4572142825310899}, {'up': 0.47736696919362326, 'down': 0.43329087396086974, 'left': 0.4997856942551695, 'right': 0.4849744445922395}], [{'up': 0.1476614248593218, 'down': 0.19439995446384012, 'left': 0.06458064771144438, 'right': 0.19395025911225672}, {'up': 0.17902213477401696, 'down': 0.21925386365977093, 'left': 0.17738779064573257, 'right': 0.22008499552630886}, {'up': 0.17849157583591413, 'down': 0.20623725396409043, 'left': 0.19779633971348595, 'right': 0.20021716465814454}, {'up': 0.17331170987365335, 'down': 0.2551033871649184, 'left': 0.20532956808183878, 'right': 0.17600303123619676}, {}, {'exit': -1.0}, {'exit': -1.0}, {'up': 0.5932701469190468, 'down': 0.4030793852180531, 'left': 0.28613977172900973, 'right': 0.48682274524329966}, {'up': 0.44077912102925987, 'down': 0.5538186717401659, 'left': 0.5134770293359358, 'right': 0.4134790497463032}, {'up': 0.44980803207117825, 'down': 0.47863771766735896, 'left': 0.4827266234297786, 'right': 0.40293761914306264}], [{'up': 0.16719509562946283, 'down': 0.20092549969334447, 'left': 0.1754277058474269, 'right': 0.22161113619272532}, {'up': 0.19261762235882324, 'down': 0.21755979311983892, 'left': 0.19311956963534277, 'right': 0.21214694368016407}, {'up': 0.20855635782297482, 'down': 0.2049028694662335, 'left': 0.22478595985362573, 'right': 0.21248742701916895}, {'up': 0.1798907954223566, 'down': 0.24058627561904827, 'left': 0.22685425277368326, 'right': 0.23722110559819182}, {'up': 0.24873138334049374, 'down': 0.2552434518377479, 'left': 0.24913241964031144, 'right': 0.30572342253660645}, {'up': 0.1613759526375117, 'down': 0.3502203910149609, 'left': 0.24935553323417678, 'right': -0.650207433673542}, {'up': -0.899999999971745, 'down': 0.35060464634839555, 'left': 0.3377187307942344, 'right': 0.40074028465416356}, {'up': 0.44508911040454413, 'down': 0.3305643505387247, 'left': 0.35970826404445055, 'right': 0.3895994027150615}, {'up': 0.4858570042830874, 'down': 0.3459071290541572, 'left': 0.4165898274986145, 'right': 0.37486538908220063}, {'up': 0.4214822155729102, 'down': 0.29349450163490515, 'left': 0.36365458396943984, 'right': 0.3451390919690339}], [{'up': 0.21502988727085712, 'down': 0.19534689333894661, 'left': 0.19775028517217996, 'right': 0.21385186762405167}, {'up': 0.1967321739489135, 'down': 0.20190949848465525, 'left': 0.19999381819693865, 'right': 0.22014711927195954}, {'up': 0.20482473963733439, 'down': 0.21091233907481577, 'left': 0.19899397921359366, 'right': 0.23207603168606278}, {'up': 0.21202837187626628, 'down': 0.23921210354550404, 'left': 0.215497760916505, 'right': 0.2411502784861742}, {'up': 0.24972803499870452, 'down': 0.25685494976348083, 'left': 0.2288648309804465, 'right': 0.2789579550947551}, {'up': 0.2760122772253771, 'down': 0.3090441573214391, 'left': 0.2593937752538388, 'right': 0.3145588004466382}, {'up': 0.36019790405384006, 'down': 0.3309964072173548, 'left': 0.2873257330398573, 'right': 0.3320164878670236}, {'up': 0.36584244926184245, 'down': 0.34458658880345205, 'left': 0.33436944249846234, 'right': 0.3620794336703196}, {'up': 0.3911793879031276, 'down': 0.3422508827638395, 'left': 0.38943447302027284, 'right': 0.28896879502813216}, {'up': 0.37403226739854245, 'down': 0.2967585486065693, 'left': 0.38207560446412214, 'right': 0.24473196816420437}]]

Como se pudo ver en la política anterior, el agente aprende a llegar a la salida con la recompensa positiva, sin embargo, el agente también intenta evadir a toda costa acercarse a las recompensas negativas.

1. Para el ambiente del laberinto de cuartos tenemos que después de iterar frente a diferentes políticas, note que lo mejor era disminuir el factor de aprendizaje, ya que de esta forma los valores pasados tienen mas relevancia para el agente. Teniendo en cuenta esto, la política obtenida fue la siguiente para el caso del laberinto de cuartos:

Imagen que contiene marcador, texto, reloj

Descripción generada automáticamente

La siguiente tabla de q-valores fue obtenida:

[[{'up': -0.46862470639674736, 'down': -0.844893225175529, 'left': -0.44902984706792837, 'right': -0.32859297937955867}, {'up': -0.18172512615156464, 'down': -0.5201917196587559, 'left': -0.24400000000000002, 'right': 0.2904680426378505}, {'exit': 0.9999990192028541}, {'up': -0.45948199122430006, 'down': -0.6853781271393086, 'left': 0.22768088726344543, 'right': -1.2561891968043373}, {'up': -1.177034051296195, 'down': -1.4382493521281474, 'left': -0.6851202152196926, 'right': -1.3680219071651631}, {}, {'up': -3.3017304790131434, 'down': -3.0487440436869964, 'left': -3.25320868244344, 'right': -3.3572482559054087}, {'up': -3.381770038935203, 'down': -3.1903550068087565, 'left': -3.292378094458045, 'right': -3.463609568279277}, {'up': -3.503095552861007, 'down': -3.4487843507113474, 'left': -3.3801800910723396, 'right': -3.6734633346046284}, {'up': -3.709703158932819, 'down': -3.6026650782803293, 'left': -3.569036638768774, 'right': -3.7684114823876538}, {'up': -3.8204478611113633, 'down': -3.7178213693551525, 'left': -3.7236107851512927, 'right': -3.82039685348054}], [{'up': -0.797063322934863, 'down': -1.1503838141959095, 'left': -1.0730127798010152, 'right': -0.864641719545918}, {'up': -0.27655286681707225, 'down': -1.131204321143451, 'left': -0.9477456998737948, 'right': -0.488336299348449}, {'up': 0.23530912034429585, 'down': -0.9764847027420903, 'left': -0.834505585325, 'right': -0.9194239040366662}, {'up': -0.5193910089801388, 'down': -1.449214743771068, 'left': -0.3512111175005815, 'right': -1.4361202901261505}, {'up': -1.3559985952540665, 'down': -1.6040018206205273, 'left': -0.9735537822937828, 'right': -1.5713188918902847}, {}, {'up': -3.250154491709392, 'down': -2.768684240686733, 'left': -2.948723108310612, 'right': -2.947074690745029}, {'up': -3.34440019171461, 'down': -3.0310452332374553, 'left': -2.9786323923884037, 'right': -3.398372401789466}, {'up': -3.5228985686904735, 'down': -3.365576372361562, 'left': -3.1924864475171386, 'right': -3.4650531769518107}, {'up': -3.6855134837045638, 'down': -3.515565026078457, 'left': -3.447880033928119, 'right': -3.7079616091231475}, {'up': -3.8111784680344036, 'down': -3.6577772883682047, 'left': -3.623734099952966, 'right': -3.729669810450994}], [{'up': -1.1707267008304267, 'down': -1.6274216401169777, 'left': -1.4091520292995678, 'right': -1.1714441793988157}, {'up': -0.8445208371756732, 'down': -1.566018068932661, 'left': -1.2146057718668373, 'right': -1.072325893891997}, {'up': -0.4226435128215035, 'down': -1.7505812385829067, 'left': -1.1684703000900565, 'right': -1.5399328386252165}, {'up': -0.9658096959165997, 'down': -2.0005182298094324, 'left': -1.031186637127273, 'right': -1.871086224075031}, {'up': -1.498293072125779, 'down': -2.2091910523947487, 'left': -1.4705917880318498, 'right': -2.28415451163833}, {'up': -2.336384797013764, 'down': -2.337431945951371, 'left': -1.8830161238762806, 'right': -2.5114602671152126}, {'up': -2.940524585095154, 'down': -2.9443091490295137, 'left': -2.4550980805651594, 'right': -2.9788358905058194}, {'up': -3.0576499576228455, 'down': -3.239768046416308, 'left': -2.737063331634267, 'right': -3.2675445161674466}, {'up': -3.4016885307235913, 'down': -3.4381145768038426, 'left': -3.0295761014231157, 'right': -3.482014612205618}, {'up': -3.5721169351137885, 'down': -3.6332420786762776, 'left': -3.337691652103788, 'right': -3.7123622453220184}, {'up': -3.7172521510792205, 'down': -3.745404998485535, 'left': -3.611304194698342, 'right': -3.715849423572367}], [{'up': -1.5445511121554008, 'down': -1.719766848248947, 'left': -1.787091771343781, 'right': -1.6223120363914876}, {'up': -1.2681632323718708, 'down': -1.9869359599046745, 'left': -1.8259882355842876, 'right': -1.5711042821231382}, {'up': -1.6375394842490727, 'down': -2.040338399526002, 'left': -1.9046632371198176, 'right': -1.9256997297556102}, {'up': -1.6281425709654607, 'down': -2.2007803829548918, 'left': -1.8495041519737279, 'right': -2.1553064220374463}, {'up': -1.9515640057663184, 'down': -2.3010368701795616, 'left': -2.137417230654661, 'right': -2.2186399422609147}, {}, {'up': -2.781662023101854, 'down': -3.283409481883443, 'left': -2.9624285609149266, 'right': -3.2338534038515143}, {'up': -3.150830093533555, 'down': -3.4318772555834545, 'left': -3.060976241801075, 'right': -3.4123924624238535}, {'up': -3.3357031585655257, 'down': -3.5502172373873786, 'left': -3.229424999263627, 'right': -3.594324893544429}, {'up': -3.5963179946785586, 'down': -3.6843020975918885, 'left': -3.4556479134412035, 'right': -3.7298779286738313}, {'up': -3.7353955942327715, 'down': -3.8291882364782492, 'left': -3.621799872196002, 'right': -3.7638031982820723}], [{'up': -1.8387168142221808, 'down': -2.102306535631375, 'left': -2.10529052473329, 'right': -1.9884504468988062}, {'up': -1.8117832938097205, 'down': -2.0861054597473747, 'left': -2.095673895157067, 'right': -2.0061807807435614}, {'up': -1.853553288188387, 'down': -2.2780026755930205, 'left': -2.1144141160015586, 'right': -2.2941675986663603}, {'up': -2.1029735962029967, 'down': -2.2690823612118995, 'left': -2.036391648054736, 'right': -2.3944703556180436}, {'up': -2.2736191802808854, 'down': -2.454728941490873, 'left': -2.286492198919933, 'right': -2.5078423580354685}, {}, {'up': -3.0279966284623603, 'down': -3.28687699597293, 'left': -3.310772079099202, 'right': -3.428445621381008}, {'up': -3.34323033745413, 'down': -3.437929055233263, 'left': -3.3153259623744717, 'right': -3.5454074254864407}, {'up': -3.432261076245497, 'down': -3.7654528951504003, 'left': -3.442999013873897, 'right': -3.7341945973199895}, {'up': -3.6136325499672677, 'down': -3.7519246850663004, 'left': -3.5998063772765896, 'right': -3.8405003896463876}, {'up': -3.7553618309855294, 'down': -3.8252332852621116, 'left': -3.752780677091275, 'right': -3.843920584848739}], [{}, {}, {'up': -2.0322319440529353, 'down': -2.564311640951333, 'left': -2.2931718518976876, 'right': -2.163722393137181}, {}, {}, {}, {}, {}, {'up': -3.6287490355986676, 'down': -3.911222474935358, 'left': -3.7551801152502082, 'right': -3.8125645705293727}, {}, {}], [{'up': -3.0550406265582692, 'down': -3.07367672300882, 'left': -3.0049104881495703, 'right': -2.937558931339542}, {'up': -2.917961585811149, 'down': -3.128597407721736, 'left': -3.069820683055377, 'right': -2.6603741880864553}, {'up': -2.3425557342014995, 'down': -2.8812506154803965, 'left': -2.886036759862557, 'right': -2.855649070797051}, {'up': -2.899981865215975, 'down': -3.0584079881583452, 'left': -2.6258615239091343, 'right': -3.124009541371362}, {'up': -3.1066862691390966, 'down': -3.3057311982333366, 'left': -2.9831350979818625, 'right': -3.1384936436009037}, {}, {'up': -3.9564170680449307, 'down': -3.966771867922886, 'left': -3.9681932047803725, 'right': -3.9711444917671197}, {'up': -3.998799096861509, 'down': -4.017644282692596, 'left': -3.9883770469007827, 'right': -3.941366321065358}, {'up': -3.806012959140356, 'down': -4.028222586787807, 'left': -4.01452920571408, 'right': -4.048829872782761}, {'up': -4.050418880755728, 'down': -4.102603229516591, 'left': -3.972307717236418, 'right': -4.134657406060633}, {'up': -4.129638685792922, 'down': -4.143313816852647, 'left': -4.054415780522258, 'right': -4.181808135872686}], [{'up': -3.0516091703548973, 'down': -3.3350432106010173, 'left': -3.223970181900161, 'right': -3.1307148919794727}, {'up': -2.94425413269649, 'down': -3.2945863014345353, 'left': -3.2500941915283335, 'right': -3.041861225718249}, {'up': -2.751062159483647, 'down': -3.1767774312770483, 'left': -3.1330179728138523, 'right': -3.142590624125844}, {'up': -2.8915197537870903, 'down': -3.3512468614926383, 'left': -3.0733737634012455, 'right': -3.27280743219495}, {'up': -3.159472108140356, 'down': -3.5313838531422332, 'left': -3.3457906586849533, 'right': -3.3997246484256456}, {}, {'up': -3.9592359015671743, 'down': -3.9085416169092064, 'left': -3.9470714663168676, 'right': -4.0630638817246965}, {'up': -4.021621813417449, 'down': -4.004223983485565, 'left': -4.005843698471226, 'right': -3.9889066131867006}, {'up': -3.9348456781608245, 'down': -4.073836941259226, 'left': -4.042364834658156, 'right': -4.093242778907064}, {'up': -4.06938307767322, 'down': -4.1873097199795355, 'left': -4.001714036022847, 'right': -4.16930436968336}, {'up': -4.159281060938641, 'down': -4.21625104256761, 'left': -4.145341428244649, 'right': -4.210126313253179}], [{'up': -3.2548573200987194, 'down': -3.3778611889086245, 'left': -3.3685478688759827, 'right': -3.3313512193515478}, {'up': -3.172990953731231, 'down': -3.406872124529562, 'left': -3.327914027616426, 'right': -3.178670996189781}, {'up': -2.9574547453435005, 'down': -3.38477123668714, 'left': -3.3518333205913136, 'right': -3.3121080240880856}, {'up': -3.1479720698126803, 'down': -3.554075203684713, 'left': -3.248028999518436, 'right': -3.5274497583033195}, {'up': -3.437484919009725, 'down': -3.705030127425181, 'left': -3.3672644424931963, 'right': -3.6738690948059145}, {'up': -3.691520939502405, 'down': -3.7253212792606196, 'left': -3.550195918779547, 'right': -3.8101314346977624}, {'up': -3.96989264392514, 'down': -3.990737637203749, 'left': -3.8145786140576314, 'right': -4.019450144768035}, {'up': -4.061604201584155, 'down': -4.101723495655272, 'left': -3.970694570192032, 'right': -4.068338343134742}, {'up': -3.9970466091238896, 'down': -4.145764663385232, 'left': -4.065691351219114, 'right': -4.149198291709169}, {'up': -4.098183384621583, 'down': -4.196021914818727, 'left': -4.145185905858696, 'right': -4.215458711743994}, {'up': -4.2106481940053175, 'down': -4.281630129410907, 'left': -4.162540437491718, 'right': -4.2114114244686025}], [{'up': -3.3610615485173065, 'down': -3.5437346526557114, 'left': -3.45503476606189, 'right': -3.445814390209622}, {'up': -3.303773434615656, 'down': -3.555374417865224, 'left': -3.470113353285929, 'right': -3.4243022006330084}, {'up': -3.219753476535626, 'down': -3.5964623750038047, 'left': -3.4491514994402004, 'right': -3.5818764533723613}, {'up': -3.4336887566014522, 'down': -3.7141710917297988, 'left': -3.413540955904673, 'right': -3.7019355124790616}, {'up': -3.5833110993850545, 'down': -3.8042382729439703, 'left': -3.572098143882302, 'right': -3.669564268160273}, {}, {'up': -3.9180281132162276, 'down': -4.1150326194767075, 'left': -4.0272563693471195, 'right': -4.122339258948031}, {'up': -4.068267374719869, 'down': -4.165837817089396, 'left': -4.0269937724472005, 'right': -4.159433269496546}, {'up': -4.086526841993333, 'down': -4.229402050025682, 'left': -4.135402144561682, 'right': -4.236315624894557}, {'up': -4.222502421485717, 'down': -4.274843900465793, 'left': -4.1621482219861985, 'right': -4.301225615164361}, {'up': -4.250536131110444, 'down': -4.359890503206868, 'left': -4.2405174800265275, 'right': -4.305716018828871}], [{'up': -3.466012704588973, 'down': -3.5353016195742586, 'left': -3.4824394835412877, 'right': -3.5549687894892243}, {'up': -3.435516788141737, 'down': -3.575426558810938, 'left': -3.518088634430291, 'right': -3.604351100162877}, {'up': -3.506556562217729, 'down': -3.612302889626593, 'left': -3.562827278368031, 'right': -3.6830031200755657}, {'up': -3.6304720274172126, 'down': -3.72461700667048, 'left': -3.624234922814293, 'right': -3.826859289646354}, {'up': -3.728191869511382, 'down': -3.8136588285080952, 'left': -3.7185702746152223, 'right': -3.792939838410868}, {}, {'up': -4.012628379460892, 'down': -4.120583065127574, 'left': -4.107944094005566, 'right': -4.161220563242307}, {'up': -4.12625032071599, 'down': -4.185350999516087, 'left': -4.1164821568718075, 'right': -4.22489550155115}, {'up': -4.1969003862582515, 'down': -4.24553329118735, 'left': -4.202467638195727, 'right': -4.295206587221716}, {'up': -4.2599083890336615, 'down': -4.30595183576262, 'left': -4.263324430069981, 'right': -4.331723961732058}, {'up': -4.309567102207394, 'down': -4.365243228138205, 'left': -4.319145719840218, 'right': -4.3637730648222455}]

Para poder mejorar el comportamiento del agente y agilizar su llegada en la menor cantidad de movimientos posibles, se definió el agente con una recompensa por paso de -0.5 y una recompensa final de 10, de este modo el agente intenta salir en la menor cantidad de pasos desde cualquier posición del tablero.