|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Reward | | | | | | | | |
| Combination | Mean | S.T.D | Median | I.Q.R | Max | Min | Failed | Difference |
| Single Q-Learning | | | | | | | | |
| - | 293.87 | 36.64 | 286.5 | 43.0 | 414 | 252 | 0 | 0 |
| - |  |  |  |  |  |  |  |  |
| L | 268.47 | 29.96 | 260.0 | 23.0 | 359 | 237 | 0 | 4.73 |
| L |  |  |  |  |  |  |  |  |
| L + D | 266.87 | 29.70 | 259.5 | 20.0 | 384 | 235 | 0 | 6.40 |
| L + D |  |  |  |  |  |  |  |  |
| L + D + R | 274.57 | 41.31 | 263.5 | 29.0 | 443 | 243 | 0 | 0 |
| L + D + R |  |  |  |  |  |  |  |  |
| L + O + D + R | - | - | - | - | - | - | - | - |
| L + O + D + R |  |  |  |  |  |  |  |  |
| Double Q-Learning | | | | | | | | |
| - | 258.03 | 16.76 | 259.0 | 21.0 | 302 | 233 | 2 | 18.24 |
| - |  |  |  |  |  |  |  |  |
| L | 243.83 | 17.35 | 245.0 | 26.0 | 293 | 214 | 1 | 32.33 |
| L |  |  |  |  |  |  |  |  |
| L + D | 250.43 | 22.90 | 245.5 | 20.0 | 322 | 219 | 0 | 24.57 |
| L + D |  |  |  |  |  |  |  |  |
| L + D + R | 270.80 | 67.56 | 247.0 | 42.0 | 500 | 225 | 0 | 0 |
| L + D + R |  |  |  |  |  |  |  |  |
| L + O + D + R | - | - | - | - | - | - | - | - |
| L + O + D + R |  |  |  |  |  |  |  |  |
| Triple Q-Learning | | | | | | | | |
| - | 286.97 | 67.94 | 262.0 | 68.0 | 496 | 226 | 3 | 0 |
| - |  |  |  |  |  |  |  |  |
| R | 263.93 | 37.28 | 251.5 | 37.0 | 414 | 227 | 2 | 7.11 |
| R |  |  |  |  |  |  |  |  |
| L + D | 245.97 | 39.84 | 234.0 | 22.0 | 417 | 213 | 1 | 24.26 |
| L + D |  |  |  |  |  |  |  |  |
| L + D + R | 236.90 | 17.40 | 234.5 | 17.0 | 291 | 216 | 1 | 39.25 |
| L + D + R |  |  |  |  |  |  |  |  |
| L + O + D + R | - | - | - | - | - | - | - | - |
| L + O + D + R |  |  |  |  |  |  |  |  |
| Quadruple Q-Learning | | | | | | | | |
| - | 263.23 | 51.20 | 250.0 | 53.0 | 491 | 216 | 4 | 3.14 |
| - |  |  |  |  |  |  |  |  |
| L | 233.00 | 38.63 | 226.0 | 18.0 | 428 | 207 | 0 | 37.61 |
| L |  |  |  |  |  |  |  |  |
| L + D | 232.97 | 23.27 | 223.5 | 23.0 | 308 | 209 | 1 | 41.94 |
| L + D |  |  |  |  |  |  |  |  |
| L + D + R | 241.37 | 41.27 | 232.0 | 23.0 | 423 | 210 | 0 | 28.39 |
| L + D + R |  |  |  |  |  |  |  |  |
| L + O + D + R | - | - | - | - | - | - | - | - |
| L + O + D + R |  |  |  |  |  |  |  |  |
| Quintuple Q-Learning | | | | | | | | |
| - | 251.47 | 34.26 | 243.5 | 40.0 | 343 | 213 | 4 | 20.49 |
| - |  |  |  |  |  |  |  |  |
| L | 222.37 | 10.50 | 221.5 | 12.0 | 261 | 208 | 0 | 54.85 |
| L |  |  |  |  |  |  |  |  |
| L + D | 228.77 | 19.21 | 225.5 | 18.0 | 285 | 202 | 1 | 47.03 |
| L + D |  |  |  |  |  |  |  |  |
| L + D + R | 222.67 | 10.28 | 223.0 | 14.0 | 250 | 206 | 0 | 54.58 |
| L + D + R |  |  |  |  |  |  |  |  |
| L + O + D + R | - | - | - | - | - | - | - | - |
| L + O + D + R |  |  |  |  |  |  |  |  |
| Reward | | | | | | | | |
| Combination | Mean | S.T.D | Median | I.Q.R | Max | Min | Failed | Difference |
| Single Q-Learning | | | | | | | | |
| - | 293.87 | 36.64 | 286.5 | 43.0 | 414 | 252 | 0 | 0 |
| - |  |  |  |  |  |  |  |  |
| L | 268.47 | 29.96 | 260.0 | 23.0 | 359 | 237 | 0 | 4.73 |
| L |  |  |  |  |  |  |  |  |
| L + D | 266.87 | 29.70 | 259.5 | 20.0 | 384 | 235 | 0 | 6.40 |
| L + D |  |  |  |  |  |  |  |  |
| L + D + R | 274.57 | 41.31 | 263.5 | 29.0 | 443 | 243 | 0 | 0 |
| L + D + R |  |  |  |  |  |  |  |  |
| L + O + D + R | - | - | - | - | - | - | - | - |
| L + O + D + R |  |  |  |  |  |  |  |  |
| Double Q-Learning | | | | | | | | |
| - | 258.03 | 16.76 | 259.0 | 21.0 | 302 | 233 | 2 | 18.24 |
| - |  |  |  |  |  |  |  |  |
| L | 243.83 | 17.35 | 245.0 | 26.0 | 293 | 214 | 1 | 32.33 |
| L |  |  |  |  |  |  |  |  |
| L + D | 250.43 | 22.90 | 245.5 | 20.0 | 322 | 219 | 0 | 24.57 |
| L + D |  |  |  |  |  |  |  |  |
| L + D + R | 270.80 | 67.56 | 247.0 | 42.0 | 500 | 225 | 0 | 0 |
| L + D + R |  |  |  |  |  |  |  |  |
| L + O + D + R | - | - | - | - | - | - | - | - |
| L + O + D + R |  |  |  |  |  |  |  |  |
| Triple Q-Learning | | | | | | | | |
| - | 286.97 | 67.94 | 262.0 | 68.0 | 496 | 226 | 3 | 0 |
| - |  |  |  |  |  |  |  |  |
| R | 263.93 | 37.28 | 251.5 | 37.0 | 414 | 227 | 2 | 7.11 |
| R |  |  |  |  |  |  |  |  |
| L + D | 245.97 | 39.84 | 234.0 | 22.0 | 417 | 213 | 1 | 24.26 |
| L + D |  |  |  |  |  |  |  |  |
| L + D + R | 236.90 | 17.40 | 234.5 | 17.0 | 291 | 216 | 1 | 39.25 |
| L + D + R |  |  |  |  |  |  |  |  |
| L + O + D + R | - | - | - | - | - | - | - | - |
| L + O + D + R |  |  |  |  |  |  |  |  |
| Quadruple Q-Learning | | | | | | | | |
| - | 263.23 | 51.20 | 250.0 | 53.0 | 491 | 216 | 4 | 3.14 |
| - |  |  |  |  |  |  |  |  |
| L | 233.00 | 38.63 | 226.0 | 18.0 | 428 | 207 | 0 | 37.61 |
| L |  |  |  |  |  |  |  |  |
| L + D | 232.97 | 23.27 | 223.5 | 23.0 | 308 | 209 | 1 | 41.94 |
| L + D |  |  |  |  |  |  |  |  |
| L + D + R | 241.37 | 41.27 | 232.0 | 23.0 | 423 | 210 | 0 | 28.39 |
| L + D + R |  |  |  |  |  |  |  |  |
| L + O + D + R | - | - | - | - | - | - | - | - |
| L + O + D + R |  |  |  |  |  |  |  |  |
| Quintuple Q-Learning | | | | | | | | |
| - | 251.47 | 34.26 | 243.5 | 40.0 | 343 | 213 | 4 | 20.49 |
| - |  |  |  |  |  |  |  |  |
| L | 222.37 | 10.50 | 221.5 | 12.0 | 261 | 208 | 0 | 54.85 |
| L |  |  |  |  |  |  |  |  |
| L + D | 228.77 | 19.21 | 225.5 | 18.0 | 285 | 202 | 1 | 47.03 |
| L + D |  |  |  |  |  |  |  |  |
| L + D + R | 222.67 | 10.28 | 223.0 | 14.0 | 250 | 206 | 0 | 54.58 |
| L + D + R |  |  |  |  |  |  |  |  |
| L + O + D + R | - | - | - | - | - | - | - | - |
| L + O + D + R |  |  |  |  |  |  |  |  |