|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Reward | | | | | | | | |
| Combination | Mean | S.T.D | Median | I.Q.R | Max | Min | Failed | Difference |
| Single Q-Learning | | | | | | | | |
| - | 257.27 | 14.94 | 256.0 | 25.0 | 288 | 235 | 0 | 0 |
| - |  |  |  |  |  |  |  |  |
| G | 237.03 | 12.74 | 236.0 | 17.0 | 269 | 217 | 0 | 11.66 |
| G |  |  |  |  |  |  |  |  |
| G + R | 231.10 | 9.32 | 231.5 | 12.0 | 250 | 216 | 0 | 18.48 |
| G + R |  |  |  |  |  |  |  |  |
| G + D + R | 234.57 | 13.34 | 232.0 | 17.0 | 268 | 216 | 0 | 13.95 |
| G + D + R |  |  |  |  |  |  |  |  |
| G + O + D + R | 247.80 | 7.06 | 247.0 | 12.0 | 261 | 237 | 0 | 2.25 |
| G + O + D + R |  |  |  |  |  |  |  |  |
| Double Q-Learning | | | | | | | | |
| - | 254.90 | 19.99 | 252.5 | 29.0 | 303 | 223 | 1 | 0 |
| - |  |  |  |  |  |  |  |  |
| G | 230.03 | 16.98 | 224.5 | 23.0 | 275 | 206 | 1 | 17.36 |
| G |  |  |  |  |  |  |  |  |
| G + D | 229.50 | 19.48 | 223.5 | 16.0 | 283 | 207 | 0 | 17.05 |
| G + D |  |  |  |  |  |  |  |  |
| G + D + R | 232.57 | 13.84 | 230.5 | 19.0 | 265 | 205 | 0 | 15.80 |
| G + D + R |  |  |  |  |  |  |  |  |
| G + O + D + R | 245.97 | 13.43 | 244.5 | 18.0 | 285 | 228 | 0 | 2.53 |
| G + O + D + R |  |  |  |  |  |  |  |  |
| Triple Q-Learning | | | | | | | | |
| - | 242.50 | 18.16 | 237.5 | 15.0 | 304 | 220 | 3 | 4.50 |
| - |  |  |  |  |  |  |  |  |
| G | 217.13 | 10.23 | 215.5 | 10.0 | 248 | 205 | 0 | 32.23 |
| G |  |  |  |  |  |  |  |  |
| G + R | 218.13 | 14.51 | 215.0 | 16.0 | 255 | 199 | 1 | 30.04 |
| G + R |  |  |  |  |  |  |  |  |
| G + O + D | 221.77 | 5.95 | 221.5 | 7.0 | 236 | 209 | 0 | 28.48 |
| G + O + D |  |  |  |  |  |  |  |  |
| G + O + D + R | 224.40 | 9.05 | 227.0 | 14.0 | 238 | 206 | 0 | 25.24 |
| G + O + D + R |  |  |  |  |  |  |  |  |
| Quadruple Q-Learning | | | | | | | | |
| - | 233.60 | 15.50 | 231.0 | 21.0 | 275 | 211 | 13 | 14.27 |
| - |  |  |  |  |  |  |  |  |
| G | 223.77 | 20.00 | 219.0 | 17.0 | 299 | 194 | 13 | 22.60 |
| G |  |  |  |  |  |  |  |  |
| G + O | 217.23 | 7.31 | 216.0 | 12.0 | 234 | 207 | 0 | 32.78 |
| G + O |  |  |  |  |  |  |  |  |
| G + O + D | 218.20 | 6.28 | 217.5 | 7.0 | 233 | 204 | 0 | 31.99 |
| G + O + D |  |  |  |  |  |  |  |  |
| G + O + D + R | 219.03 | 10.13 | 218.0 | 12.0 | 249 | 203 | 0 | 30.36 |
| G + O + D + R |  |  |  |  |  |  |  |  |
| Quintuple Q-Learning | | | | | | | | |
| - | - | - | - | - | - | - | - | - |
| - |  |  |  |  |  |  |  |  |
| O | 263.10 | 14.37 | 260.5 | 12.0 | 298 | 239 | 0 | 0 |
| O |  |  |  |  |  |  |  |  |
| G + O | 215.53 | 7.88 | 214.0 | 11.0 | 231 | 199 | 0 | 34.36 |
| G + O |  |  |  |  |  |  |  |  |
| G + O + R | 214.07 | 9.53 | 214.5 | 12.0 | 231 | 196 | 0 | 35.46 |
| G + O + R |  |  |  |  |  |  |  |  |
| G + O + D + R | 217.07 | 6.33 | 216.5 | 9.0 | 231 | 208 | 0 | 33.11 |
| G + O + D + R |  |  |  |  |  |  |  |  |
| Reward | | | | | | | | |
| Combination | Mean | S.T.D | Median | I.Q.R | Max | Min | Failed | Difference |
| Single Q-Learning | | | | | | | | |
| - | 257.27 | 14.94 | 256.0 | 25.0 | 288 | 235 | 0 | 0 |
| - |  |  |  |  |  |  |  |  |
| G | 237.03 | 12.74 | 236.0 | 17.0 | 269 | 217 | 0 | 11.66 |
| G |  |  |  |  |  |  |  |  |
| G + R | 231.10 | 9.32 | 231.5 | 12.0 | 250 | 216 | 0 | 18.48 |
| G + R |  |  |  |  |  |  |  |  |
| G + D + R | 234.57 | 13.34 | 232.0 | 17.0 | 268 | 216 | 0 | 13.95 |
| G + D + R |  |  |  |  |  |  |  |  |
| G + O + D + R | 247.80 | 7.06 | 247.0 | 12.0 | 261 | 237 | 0 | 2.25 |
| G + O + D + R |  |  |  |  |  |  |  |  |
| Double Q-Learning | | | | | | | | |
| - | 254.90 | 19.99 | 252.5 | 29.0 | 303 | 223 | 1 | 0 |
| - |  |  |  |  |  |  |  |  |
| G | 230.03 | 16.98 | 224.5 | 23.0 | 275 | 206 | 1 | 17.36 |
| G |  |  |  |  |  |  |  |  |
| G + D | 229.50 | 19.48 | 223.5 | 16.0 | 283 | 207 | 0 | 17.05 |
| G + D |  |  |  |  |  |  |  |  |
| G + D + R | 232.57 | 13.84 | 230.5 | 19.0 | 265 | 205 | 0 | 15.80 |
| G + D + R |  |  |  |  |  |  |  |  |
| G + O + D + R | 245.97 | 13.43 | 244.5 | 18.0 | 285 | 228 | 0 | 2.53 |
| G + O + D + R |  |  |  |  |  |  |  |  |
| Triple Q-Learning | | | | | | | | |
| - | 242.50 | 18.16 | 237.5 | 15.0 | 304 | 220 | 3 | 4.50 |
| - |  |  |  |  |  |  |  |  |
| G | 217.13 | 10.23 | 215.5 | 10.0 | 248 | 205 | 0 | 32.23 |
| G |  |  |  |  |  |  |  |  |
| G + R | 218.13 | 14.51 | 215.0 | 16.0 | 255 | 199 | 1 | 30.04 |
| G + R |  |  |  |  |  |  |  |  |
| G + O + D | 221.77 | 5.95 | 221.5 | 7.0 | 236 | 209 | 0 | 28.48 |
| G + O + D |  |  |  |  |  |  |  |  |
| G + O + D + R | 224.40 | 9.05 | 227.0 | 14.0 | 238 | 206 | 0 | 25.24 |
| G + O + D + R |  |  |  |  |  |  |  |  |
| Quadruple Q-Learning | | | | | | | | |
| - | 233.60 | 15.50 | 231.0 | 21.0 | 275 | 211 | 13 | 14.27 |
| - |  |  |  |  |  |  |  |  |
| G | 223.77 | 20.00 | 219.0 | 17.0 | 299 | 194 | 13 | 22.60 |
| G |  |  |  |  |  |  |  |  |
| G + O | 217.23 | 7.31 | 216.0 | 12.0 | 234 | 207 | 0 | 32.78 |
| G + O |  |  |  |  |  |  |  |  |
| G + O + D | 218.20 | 6.28 | 217.5 | 7.0 | 233 | 204 | 0 | 31.99 |
| G + O + D |  |  |  |  |  |  |  |  |
| G + O + D + R | 219.03 | 10.13 | 218.0 | 12.0 | 249 | 203 | 0 | 30.36 |
| G + O + D + R |  |  |  |  |  |  |  |  |
| Quintuple Q-Learning | | | | | | | | |
| - | - | - | - | - | - | - | - | - |
| - |  |  |  |  |  |  |  |  |
| O | 263.10 | 14.37 | 260.5 | 12.0 | 298 | 239 | 0 | 0 |
| O |  |  |  |  |  |  |  |  |
| G + O | 215.53 | 7.88 | 214.0 | 11.0 | 231 | 199 | 0 | 34.36 |
| G + O |  |  |  |  |  |  |  |  |
| G + O + R | 214.07 | 9.53 | 214.5 | 12.0 | 231 | 196 | 0 | 35.46 |
| G + O + R |  |  |  |  |  |  |  |  |
| G + O + D + R | 217.07 | 6.33 | 216.5 | 9.0 | 231 | 208 | 0 | 33.11 |
| G + O + D + R |  |  |  |  |  |  |  |  |