```
    options, feed_dict, fetch_list, target_list, run_metadata)
  File "/home/ubuntu/.local/lib/python3.7/site-packages/tensorflow/python/client/session.py", line 1429, in _call_tf_se
ssionrun
    run_metadata)
KeyboardInterrupt
Traceback (most recent call last):
ubuntu@ip-172-31-9-175:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
17:40, 04/12; num of cores:16
median_u_O_u_D

Basic setting:[rep_times, sd_O, sd_D, sd_u_O, w_O, w_A, u_O_u_D, t_func] = [16, None, None, 20, 0.5, 1.5, 10, None]

[thre_range, sd_R_range, day_range, penalty_range]:  [[100, 105, 110, 115], [0, 15, 30], [7], [[0.0001, 5e-05], [0.0001
, 5e-05]]]

---------------------------------------
[pattern_seed, day, sd_R] = [2, 7, 0]

max(u_O) =  145.8
O_threshold = 100
number of reward locations:  9
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
O_threshold = 115
number of reward locations:  3
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!

---------------------------------------
Value of Behaviour policy:52.865
O_threshold = 100
MC for this TARGET:[58.154, 0.081]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.24, -0.33, -1.1]][[-1.31, -58.15, -5.29]]
std:[[0.6, 0.59, 0.41]][[0.32, 0.0, 0.23]]
MSE:[[0.65, 0.68, 1.17]][[1.35, 58.15, 5.29]]
MSE(-DR):[[0.0, 0.03, 0.52]][[0.7, 57.5, 4.64]]
***
==============
O_threshold = 105
MC for this TARGET:[57.708, 0.073]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.18, -2.24, -3.07]][[-4.14, -57.71, -4.84]]
std:[[0.63, 0.61, 0.4]][[0.32, 0.0, 0.23]]
MSE:[[2.27, 2.32, 3.1]][[4.15, 57.71, 4.85]]
MSE(-DR):[[0.0, 0.05, 0.83]][[1.88, 55.44, 2.58]]
***
==============
O_threshold = 110
MC for this TARGET:[56.697, 0.063]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.12, -2.15, -2.93]][[-5.27, -56.7, -3.83]]
std:[[0.52, 0.5, 0.41]][[0.33, 0.0, 0.23]]
MSE:[[2.18, 2.21, 2.96]][[5.28, 56.7, 3.84]]
MSE(-DR):[[0.0, 0.03, 0.78]][[3.1, 54.52, 1.66]]
***
==============
O_threshold = 115
MC for this TARGET:[58.647, 0.054]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.17, -7.14, -7.54]][[-13.2, -58.65, -5.78]]
std:[[0.74, 0.74, 0.5]][[0.32, 0.0, 0.23]]
MSE:[[7.21, 7.18, 7.56]][[13.2, 58.65, 5.78]]
MSE(-DR):[[0.0, -0.03, 0.35]][[5.99, 51.44, -1.43]]
***
==============
******************** THIS SETTING IS GOOD ********************
[[ 0.65  0.68  1.17  1.35 58.15  5.29]
 [ 2.27  2.32  3.1   4.15 57.71  4.85]
 [ 2.18  2.21  2.96  5.28 56.7   3.84]
 [ 7.21  7.18  7.56 13.2  58.65  5.78]]


time spent until now: 39.0 mins
```

18:19, 04/12

[pattern_seed, day, sd_R] = [2, 7, 15]

max(u_O) = 145.8
O_threshold = 100
number of reward locations: 9
O_threshold = 105
number of reward locations: 7
O_threshold = 110
number of reward locations: 6
O_threshold = 115
number of reward locations: 3
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!

Value of Behaviour policy:52.843
O_threshold = 100
MC for this TARGET:[58.178, 0.188]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.22, -0.31, -1.2]][[-1.37, -58.18, -5.33]]
std:[[0.74, 0.76, 0.56]][[0.31, 0.0, 0.21]]
MSE:[[0.77, 0.82, 1.32]][[1.4, 58.18, 5.33]]
MSE(-DR):[[0.0, 0.05, 0.55]][[0.63, 57.41, 4.56]]
***
==========
O_threshold = 105
MC for this TARGET:[57.732, 0.182]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.94, -2.04, -3.19]][[-4.24, -57.73, -4.89]]
std:[[0.92, 0.91, 0.57]][[0.32, 0.0, 0.21]]
MSE:[[2.15, 2.23, 3.24]][[4.25, 57.73, 4.89]]
MSE(-DR):[[0.0, 0.08, 1.09]][[2.1, 55.58, 2.74]]
***
==========
O_threshold = 110
MC for this TARGET:[56.721, 0.18]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-2.02, -2.08, -3.0]][[-5.33, -56.72, -3.88]]
std:[[0.73, 0.71, 0.67]][[0.37, 0.0, 0.21]]
MSE:[[2.15, 2.2, 3.07]][[5.34, 56.72, 3.89]]
MSE(-DR):[[0.0, 0.05, 0.92]][[3.19, 54.57, 1.74]]
***
==========
O_threshold = 115
MC for this TARGET:[58.671, 0.183]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.24, -7.23, -7.57]][[-13.25, -58.67, -5.83]]
std:[[1.0, 0.98, 0.7]][[0.38, 0.0, 0.21]]
MSE:[[7.31, 7.3, 7.6]][[13.26, 58.67, 5.83]]
MSE(-DR):[[0.0, -0.01, 0.29]][[5.95, 51.36, -1.48]]
***
==========
******************** THIS SETTING IS GOOD ********************
[[0.65 0.68 1.17 1.35 58.15 5.29]
 [2.27 2.32 3.1   4.15 57.71 4.85]
 [2.18 2.21 2.96 5.28 56.7   3.84]
 [7.21 7.18 7.56 13.2 58.65 5.78]]

[[0.77 0.82 1.32 1.4 58.18 5.33]
 [2.15 2.23 3.24 4.25 57.73 4.89]
 [2.15 2.2   3.07 5.34 56.72 3.89]
 [7.31 7.3   7.6 13.26 58.67 5.83]]

time spent until now: 78.0 mins

18:58, 04/12

[pattern_seed, day, sd_R] = [2, 7, 30]

```
max(u_O) =  145.8
O_threshold = 100
number of reward locations:  9
O_threshold = 105
number of reward locations:  7
O_threshold = 110
number of reward locations:  6
O_threshold = 115
number of reward locations:  3
target 1 in 4 DONE!
target 2 in 4 DONE!
target 3 in 4 DONE!
target 4 in 4 DONE!

---------------------------------------
Value of Behaviour policy:52.822
O_threshold = 100
MC for this TARGET:[58.202, 0.352]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-0.19, -0.31, -1.3]][[-1.43, -58.2, -5.38]]
std:[[1.33, 1.36, 0.89]][[0.54, 0.0, 0.3]]
MSE:[[1.34, 1.39, 1.58]][[1.53, 58.2, 5.39]]
MSE(-DR):[[0.0, 0.05, 0.24]][[0.19, 56.86, 4.05]]
***
==============
O_threshold = 105
MC for this TARGET:[57.755, 0.348]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.77, -1.88, -3.31]][[-4.36, -57.76, -4.93]]
std:[[1.65, 1.68, 0.86]][[0.55, 0.0, 0.3]]
MSE:[[2.42, 2.52, 3.42]][[4.39, 57.76, 4.94]]
MSE(-DR):[[0.0, 0.1, 1.0]][[1.97, 55.34, 2.52]]
***
==============
O_threshold = 110
MC for this TARGET:[56.744, 0.347]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-1.95, -2.03, -3.05]][[-5.39, -56.74, -3.92]]
std:[[1.32, 1.28, 0.99]][[0.63, 0.0, 0.3]]
MSE:[[2.35, 2.4, 3.21]][[5.43, 56.74, 3.93]]
MSE(-DR):[[0.0, 0.05, 0.86]][[3.08, 54.39, 1.58]]
***
==============
O_threshold = 115
MC for this TARGET:[58.695, 0.353]
   [DR/QV/IS]; [DR_NO_MARL, DR_NO_MF, V_behav]
bias:[[-7.27, -7.26, -7.62]][[-13.3, -58.7, -5.87]]
std:[[1.72, 1.7, 1.3]][[0.67, 0.0, 0.3]]
MSE:[[7.47, 7.46, 7.73]][[13.32, 58.7, 5.88]]
MSE(-DR):[[0.0, -0.01, 0.26]][[5.85, 51.23, -1.59]]
***
==============
******************** THIS SETTING IS GOOD ********************
[[ 0.65  0.68  1.17  1.35 58.15  5.29]
 [ 2.27  2.32  3.1   4.15 57.71  4.85]
 [ 2.18  2.21  2.96  5.28 56.7   3.84]
 [ 7.21  7.18  7.56 13.2  58.65  5.78]]


[[ 0.77  0.82  1.32  1.4  58.18  5.33]
 [ 2.15  2.23  3.24  4.25 57.73  4.89]
 [ 2.15  2.2   3.07  5.34 56.72  3.89]
 [ 7.31  7.3   7.6  13.26 58.67  5.83]]


[[ 1.34  1.39  1.58  1.53 58.2   5.39]
 [ 2.42  2.52  3.42  4.39 57.76  4.94]
 [ 2.35  2.4   3.21  5.43 56.74  3.93]
 [ 7.47  7.46  7.73 13.32 58.7   5.88]]


time spent until now: 117.2 mins

19:37, 04/12
ubuntu@ip-172-31-9-175:~$ export openblas_num_threads=1; export OMP_NUM_THREADS=1; python EC2.py
19:47, 04/12; num of cores:16
median_u_O_u_D_other_C
```