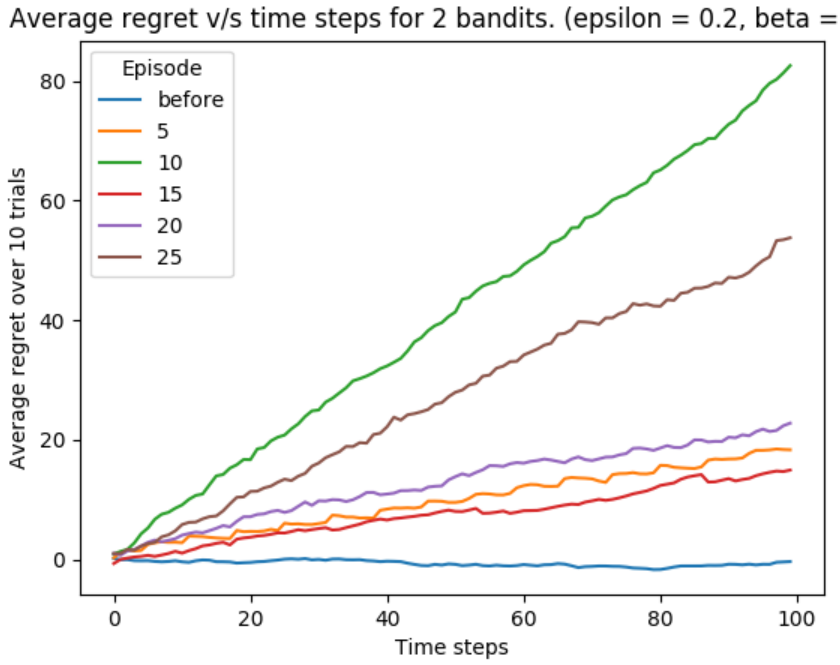
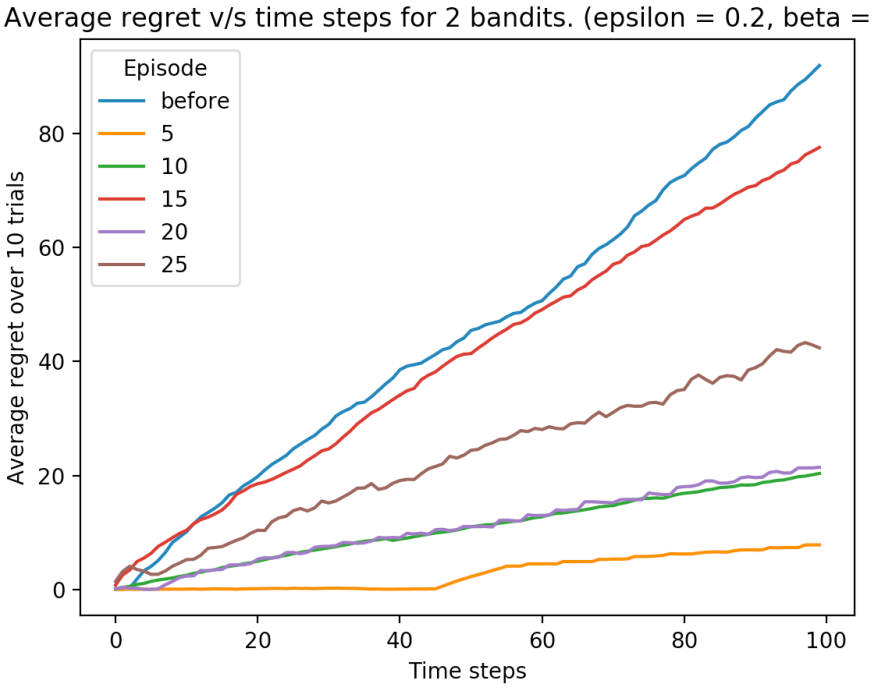


dqn_cowan_episodes5_trials5_time100
Bandits : 2(unknown)

mean = random.uniform(-1, 1)
sigma = random.uniform(0, 2)



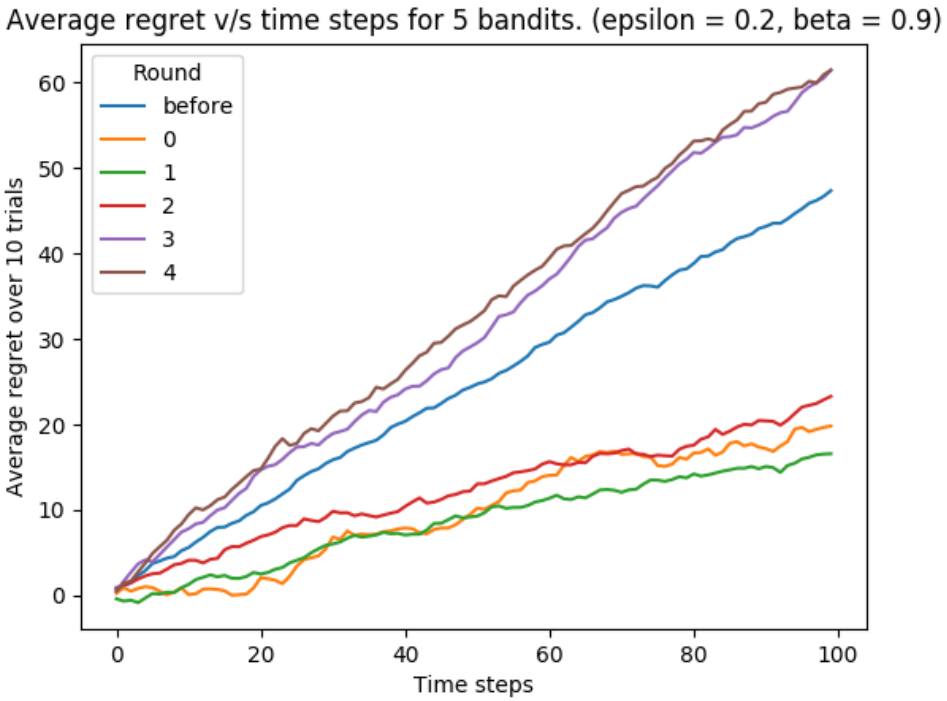
DQN
Bandits : 2
Permutation : No
mean = random.uniform(-1, 1)
sigma = random.uniform(0, 2)
Timesteps = 100

Bandits:
Before - [(-0.03898216014129141, 0.2971887131273505), (-0.9433580915870117, 1.1540540203409415)]
Round 0 - [(0.4831526967123454, 0.08968019832107266), (0.09117959507431106, 0.14800046120870647)]
Round 1 - [(0.08840064166082118, 0.18606097617512196), (0.30899638493835035, 0.7978929759286326)]

Round 2 -
[(-0.6288237287887264, 1.2466965987355854), (-0.7643578669745132, 0.38822558513194627)]

Round 3 -
[(0.9194568019690919, 0.4315959521583459), (0.04196165852896261, 0.652224153183941)]

Round 4 -
[(-0.7350360170417025, 0.7255101068848466), (0.7669946785087933, 1.4655480021419436)]



DQN
Bandits : 5
Permutation : No
mean = random.uniform(-1, 1)
sigma = random.uniform(0, 2)
Timesteps 100

Bandits

0.186 1.598
0.285 0.317
-0.360 1.243
-0.745 1.169
0.641 0.212
Before

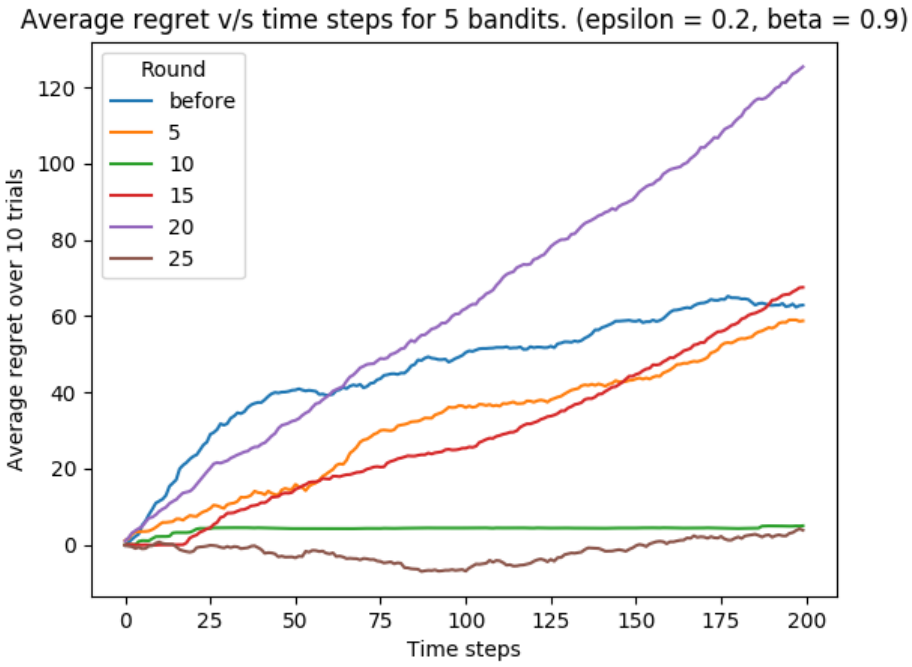
0.667 1.223
0.685 0.841
0.649 1.830
0.527 0.834
0.053 1.119
Round 0 done.

0.421 0.070
-0.768 0.753
0.754 0.948
0.751 1.000
0.199 1.091
Round 1 done.

0.270 1.428
0.218 0.782
-0.059 1.338
0.067 0.211
0.531 0.227
Round 2 done.

0.507 1.523
-0.541 0.681
0.203 0.590
-0.270 0.798
0.222 1.611
Round 3 done.

0.060 1.611
0.661 1.589
0.038 0.619
0.025 1.484
-0.507 0.804
Round 4 done.



DQN
Bandits : 5
Permutation : No
mean = random.uniform(-1, 1)
sigma = random.uniform(0, 2)
Timesteps : 200

-0.474 1.871
-0.314 1.081
0.846 0.758
0.686 1.351
-0.987 0.364
Before

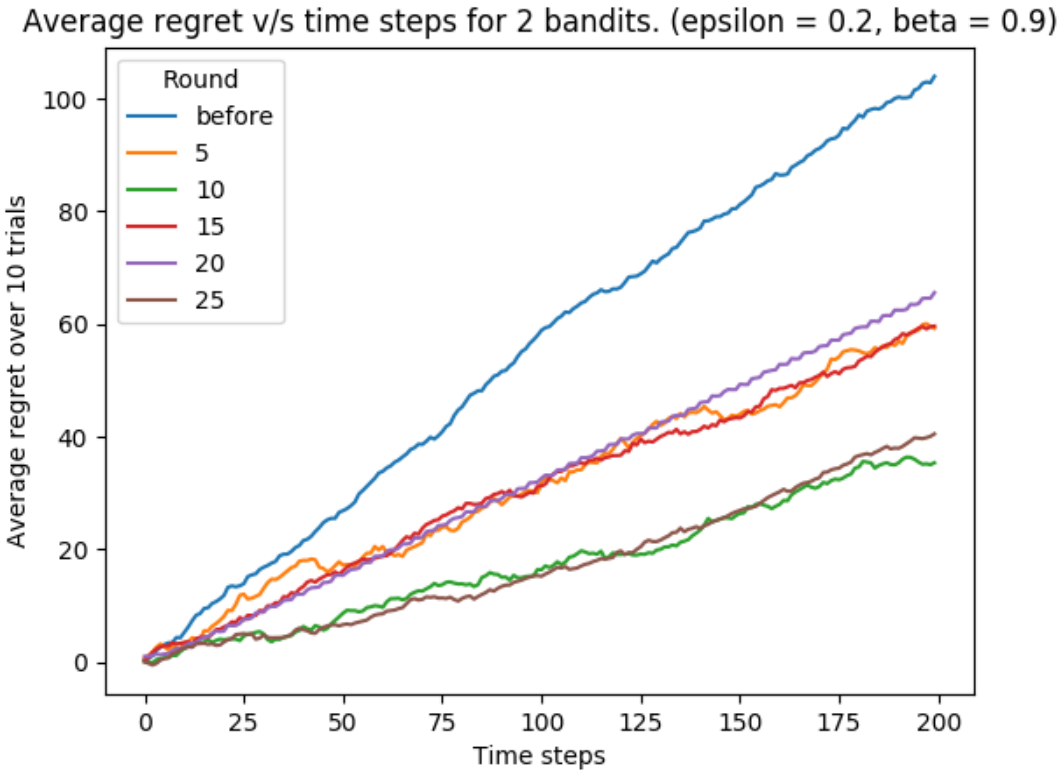
0.431 1.081
-0.881 1.737
-0.683 0.418

-0.194 0.794
-0.496 0.791
Round 0 done.
0.768 0.067
-0.678 1.650
-0.328 0.220
0.074 0.214
0.254 1.397
Round 1 done.

-0.368 0.460
-0.386 0.925
0.919 0.147
0.512 0.456
0.790 1.580
Round 2 done.

0.344 0.333
-0.153 0.030
0.118 1.473
0.938 1.040
-0.485 1.301
Round 3 done.

0.487 1.087
-0.824 1.656
0.700 1.185
0.317 0.707
-0.532 0.634
Round 4 done.



DQN
Bandit = 2
Timesteps = 200

-0.556 0.764
-0.529 1.270

-0.027 1.891
-0.606 1.672
Round 0 done.

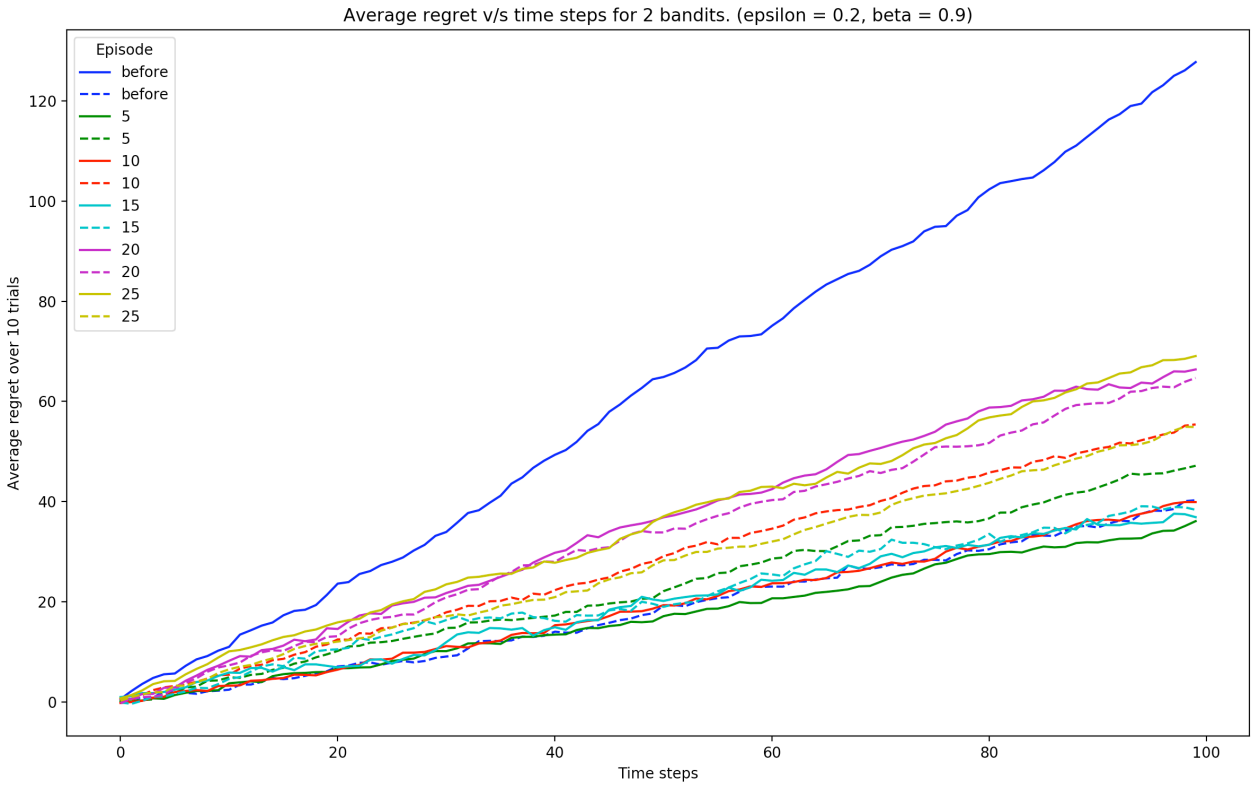
0.073 0.863
0.392 1.785
Round 1 done.

0.909 1.032
0.370 0.961
Round 2 done.

-0.052 0.234
-1.000 0.107
Round 3 done.

-0.290 0.376
-0.101 1.401
Round 4 done.

DQN v/s Double Q (- -)



Bandits = 2
Timesteps = 100

-0.905 1.888
0.470 0.685

-0.828 1.463
-0.097 0.364
Round 0 done.

-0.850 0.916
-0.209 0.835
Round 1 done.

-0.698 1.934
0.536 1.596
Round 2 done.

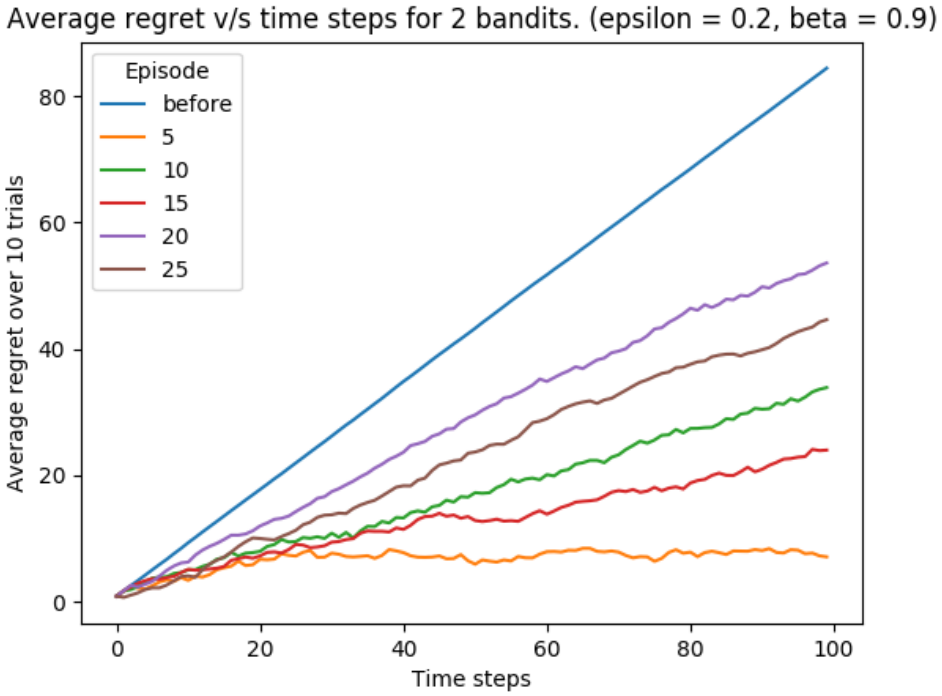
-0.654 1.983
-0.717 1.476
Round 3 done.

-0.682 1.384
-0.606 1.220
Round 4 done.

Same test bandit set

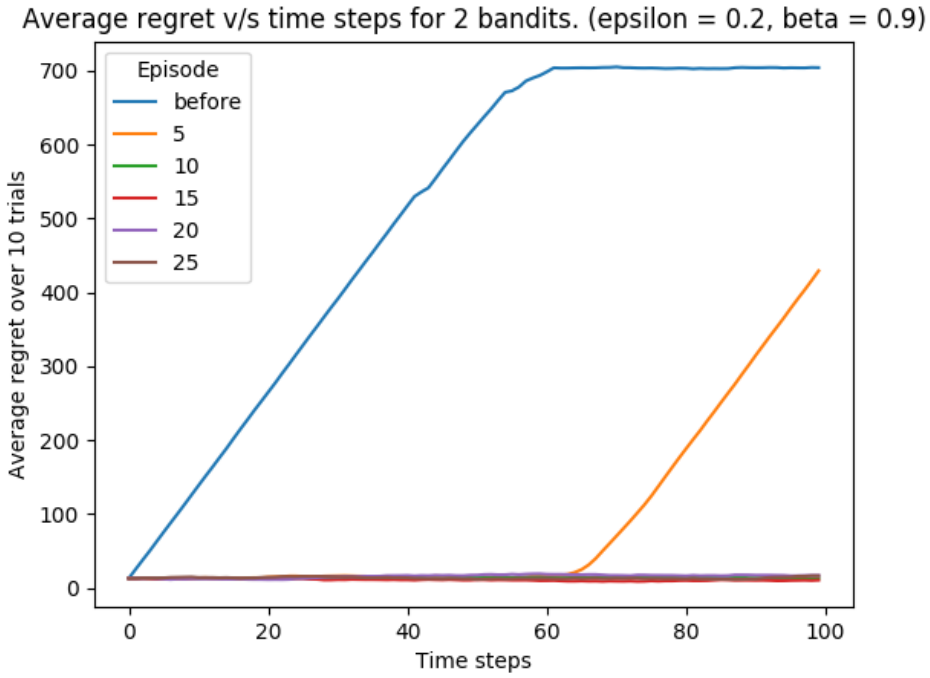
Test bandits
1: 0.364 1.315
2: -0.480 0.072

DQN



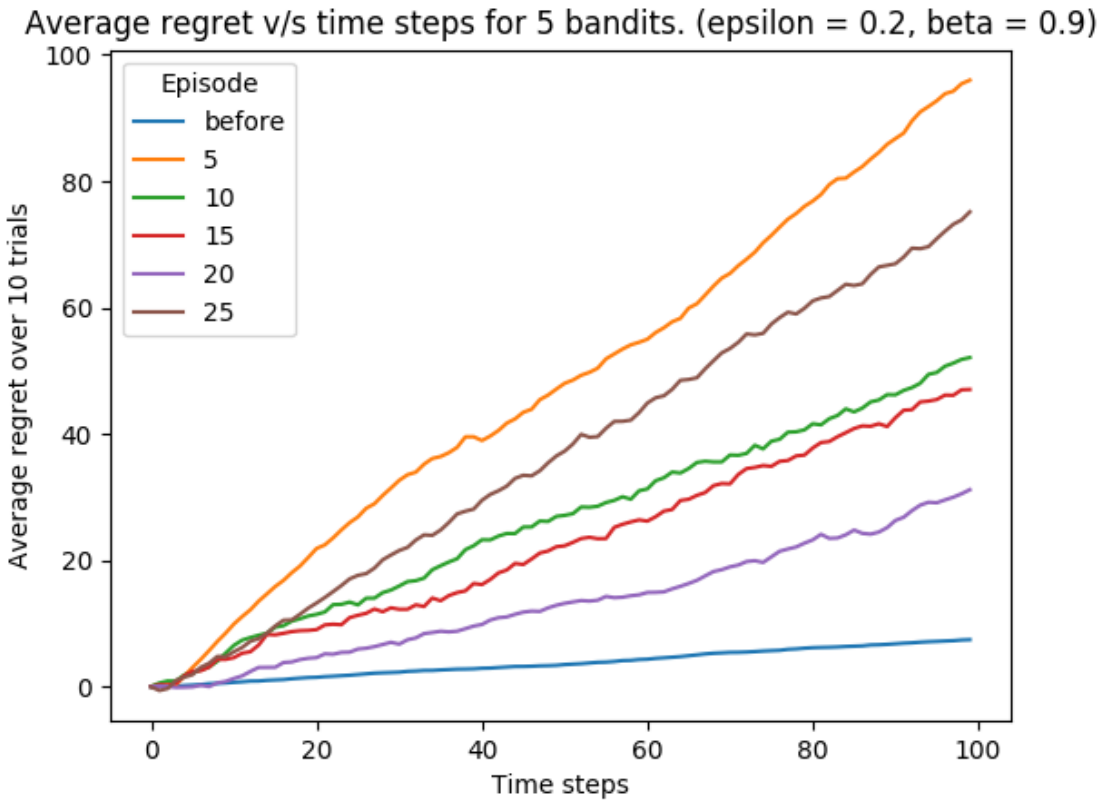
DQN 2 bandits
mean = random.uniform(-10, 10)
sigma = random.uniform(0, 2)

Test bandit:
-9.304 0.865
3.332 1.128



DQN
5 bandits

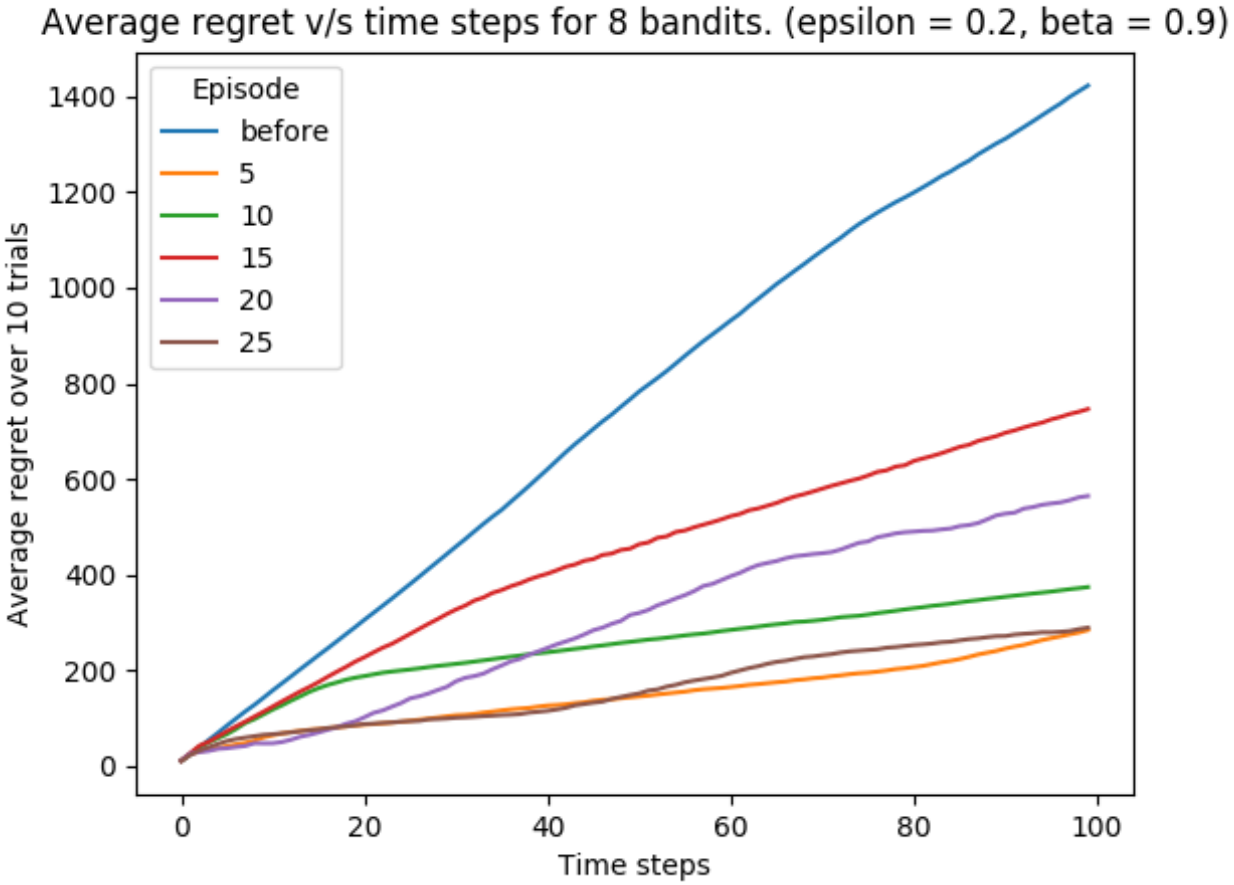
Test bandit:
0.433 1.089
0.572 1.422
0.494 0.115
-0.706 0.830
-0.810 0.273



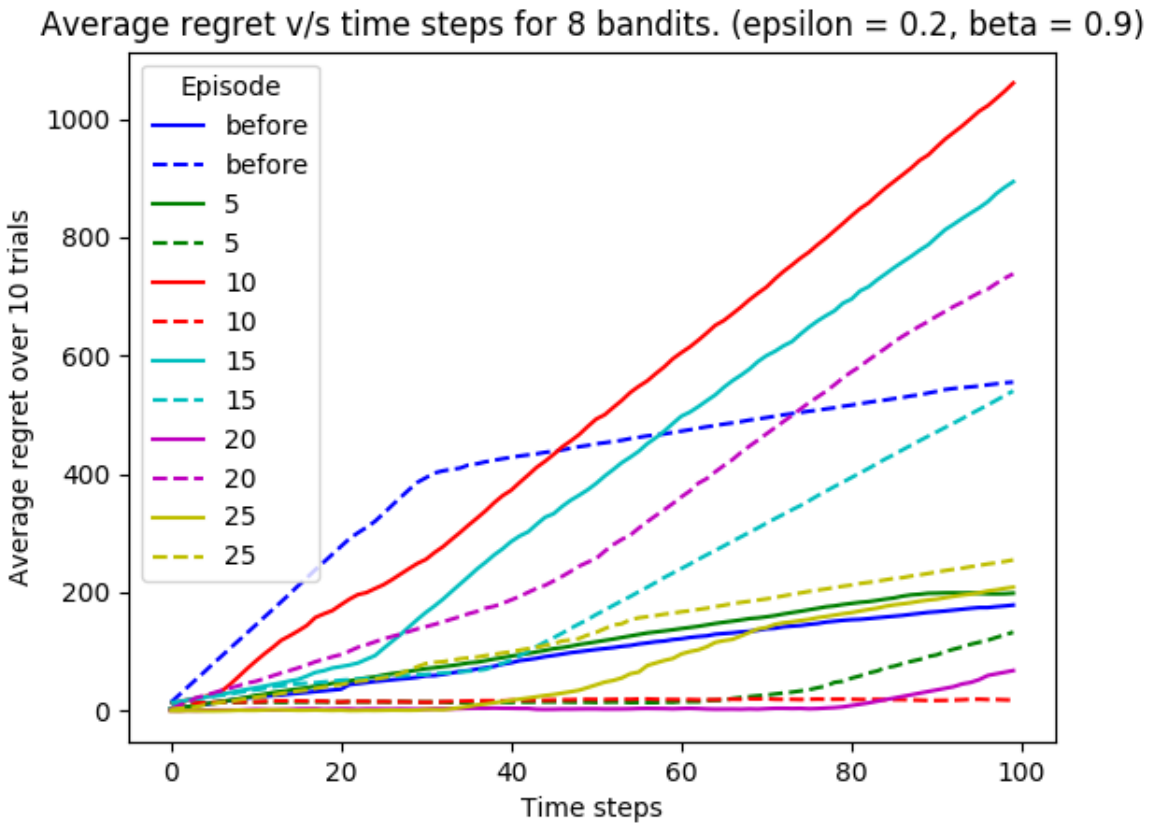
DQN
8 bandits

mean = random.uniform(-10, 10)
sigma = random.uniform(0, 2)

Test bandit:
9.945 1.630
-0.704 0.234
-6.633 1.415
8.687 1.003
9.654 0.640
-7.794 0.998
-0.761 1.678
3.348 1.832



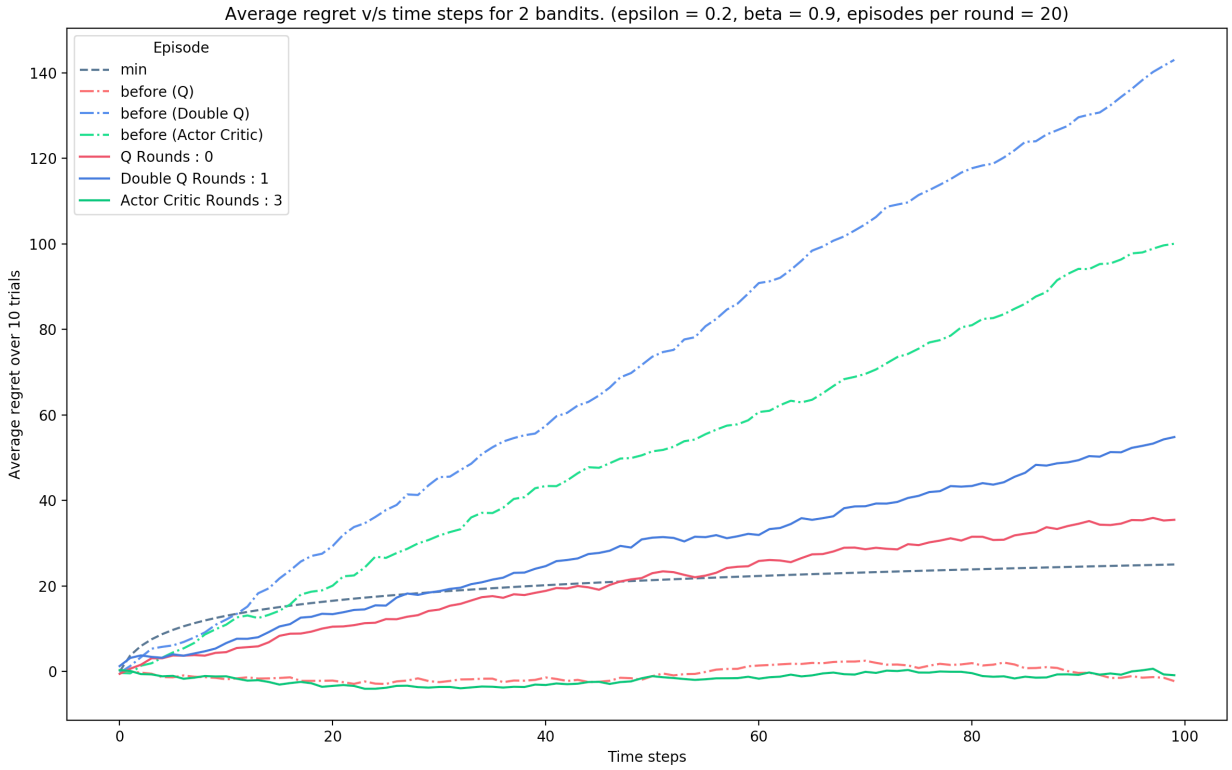
DQN v/s Double Q



8 bandits

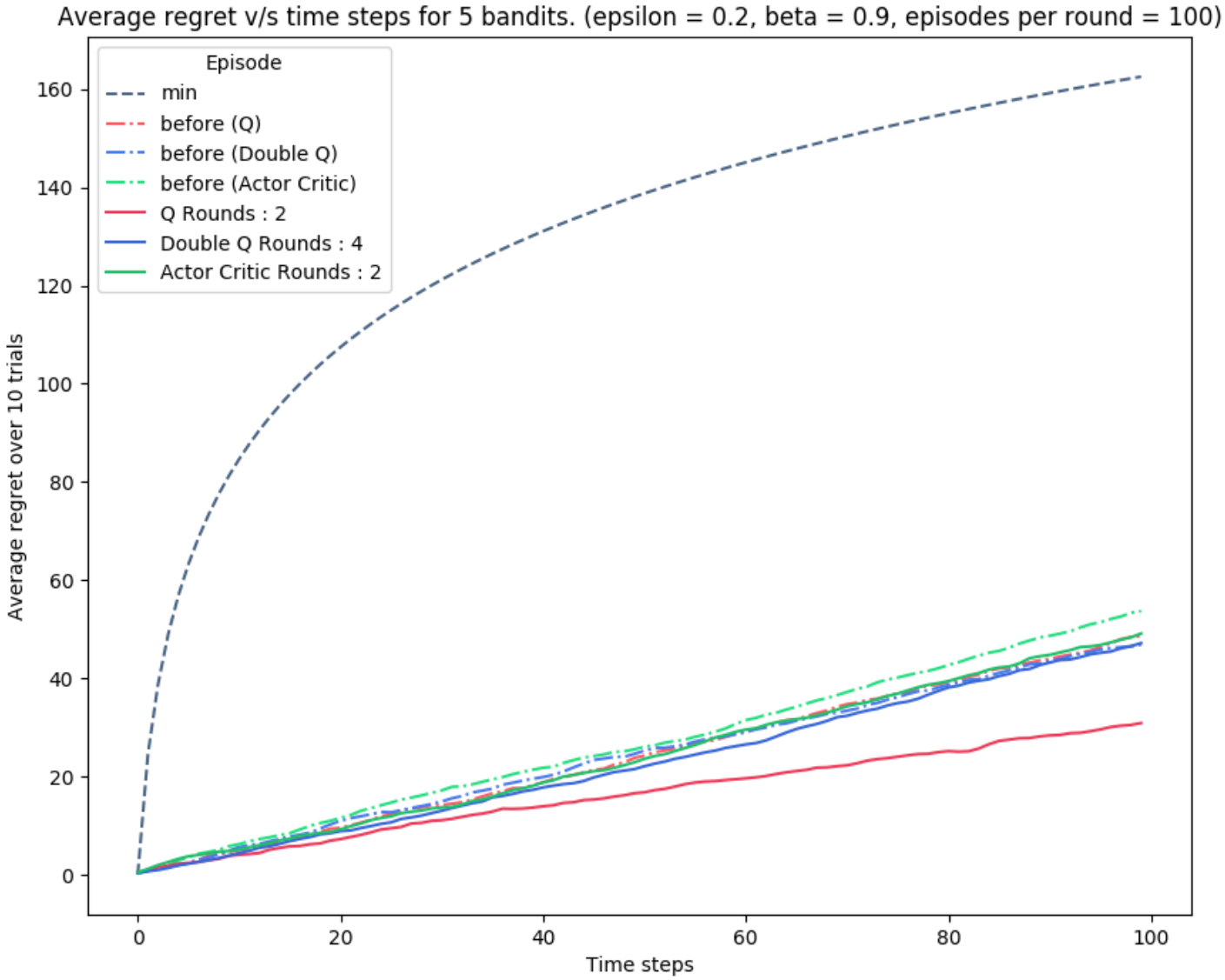
Test bandits :
3.273 0.259
-1.100 1.235
3.976 1.451
-8.003 0.666
3.791 1.242
-7.160 0.807
6.038 1.177
3.850 1.114

Bandits = 2; mean = random.uniform(-1, 1); 20 episodes; w/o permutations
0.879 1.253
-0.580 1.992



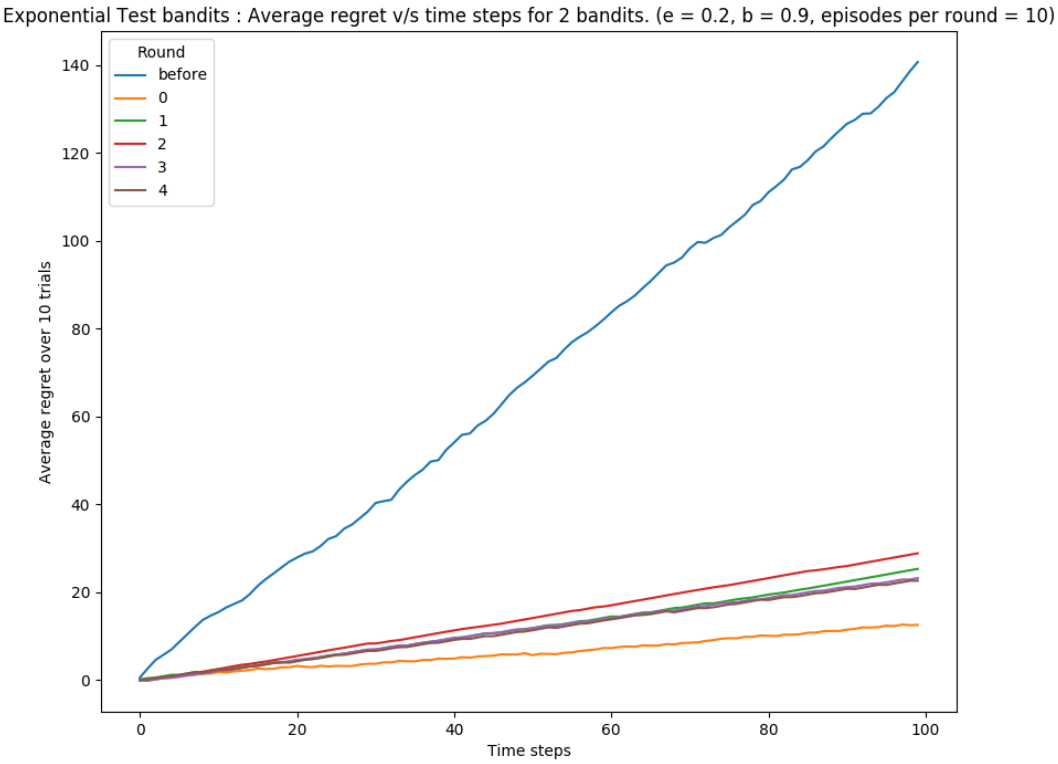
Bandits = 5; episodes = 100; permutations

Test bandits :
0.037 1.560
0.277 1.252
-0.415 1.723
-0.627 0.535
0.465 0.725



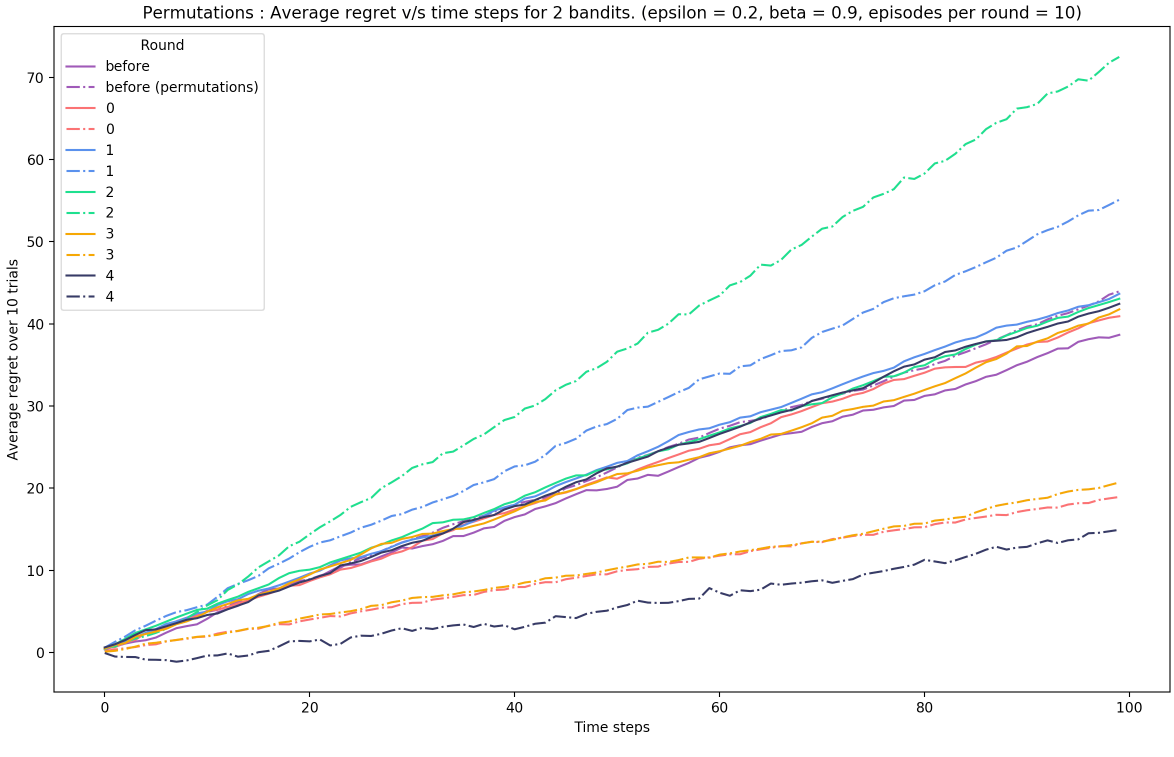
Exp bandits - Training on normal bandits; testing on expo bandits

Bandits = 2
Episodes per round =10
Beta ~(0, 1)



Permutations

Test bandits :
-0.454 0.391fac



-0.388 0.648