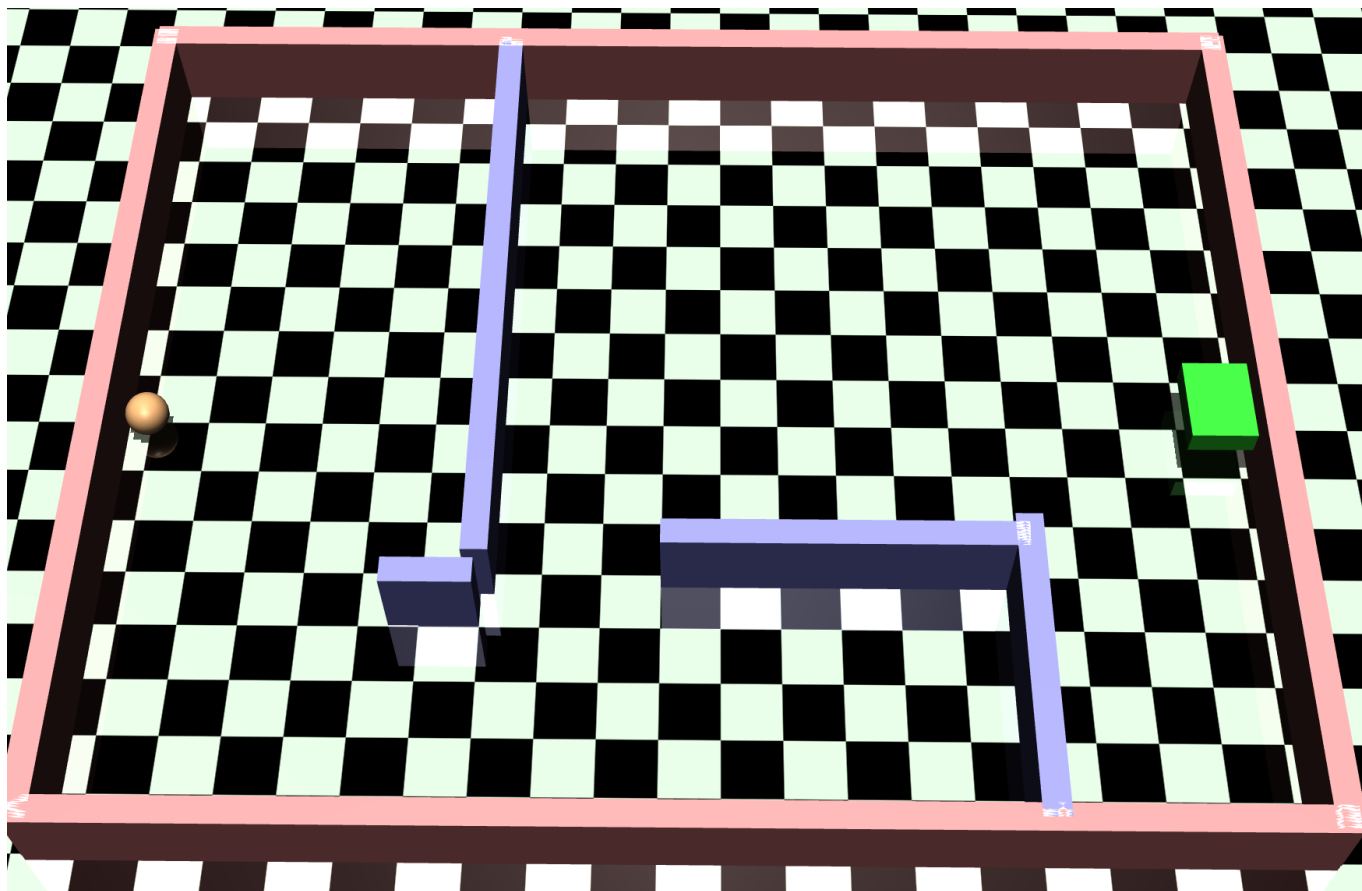


SAC - tuning on layers size

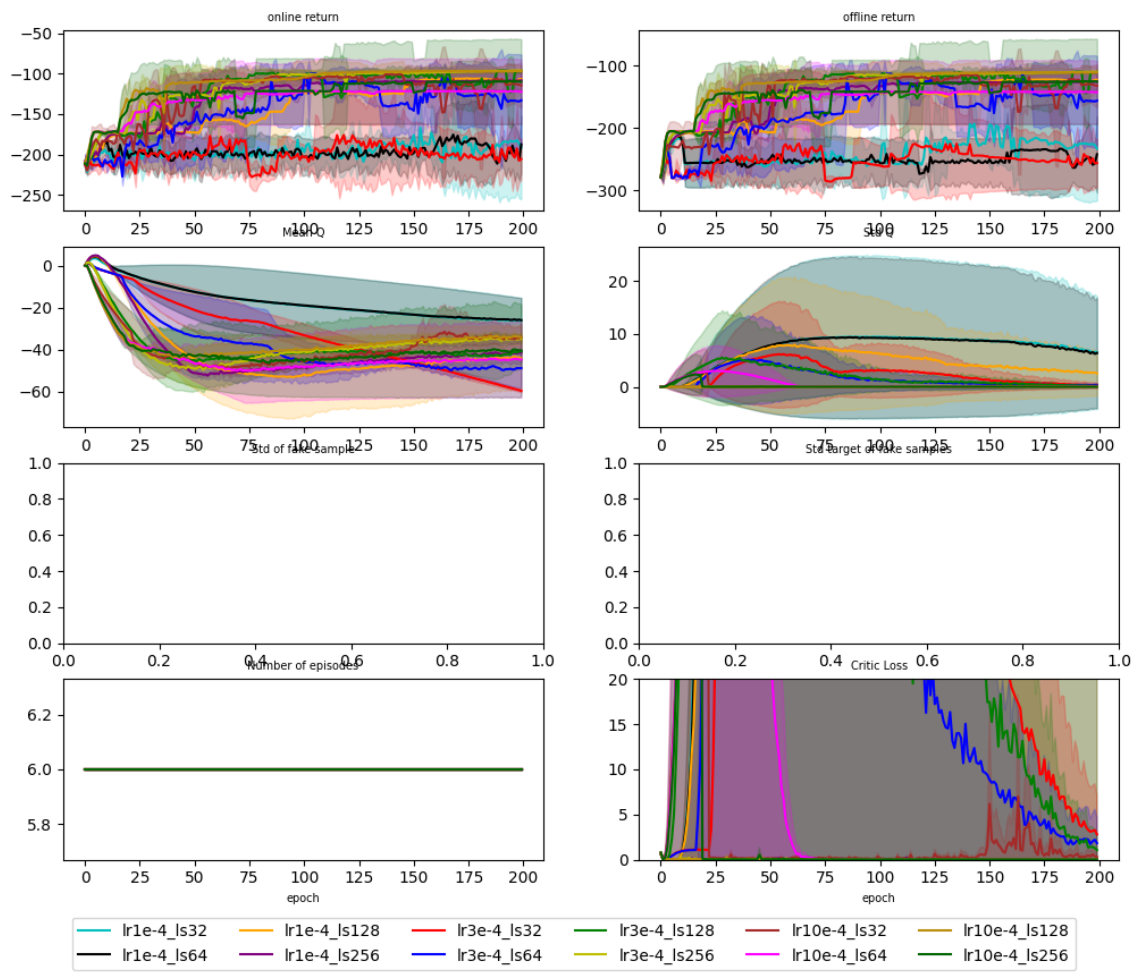


```
lr_s=(1e-4 3e-4 10e-4)  
ls_s=(32 64 128 256)
```

Ho runnato SAC sia con che senza stati terminali su la griglia di valori riportata sopra

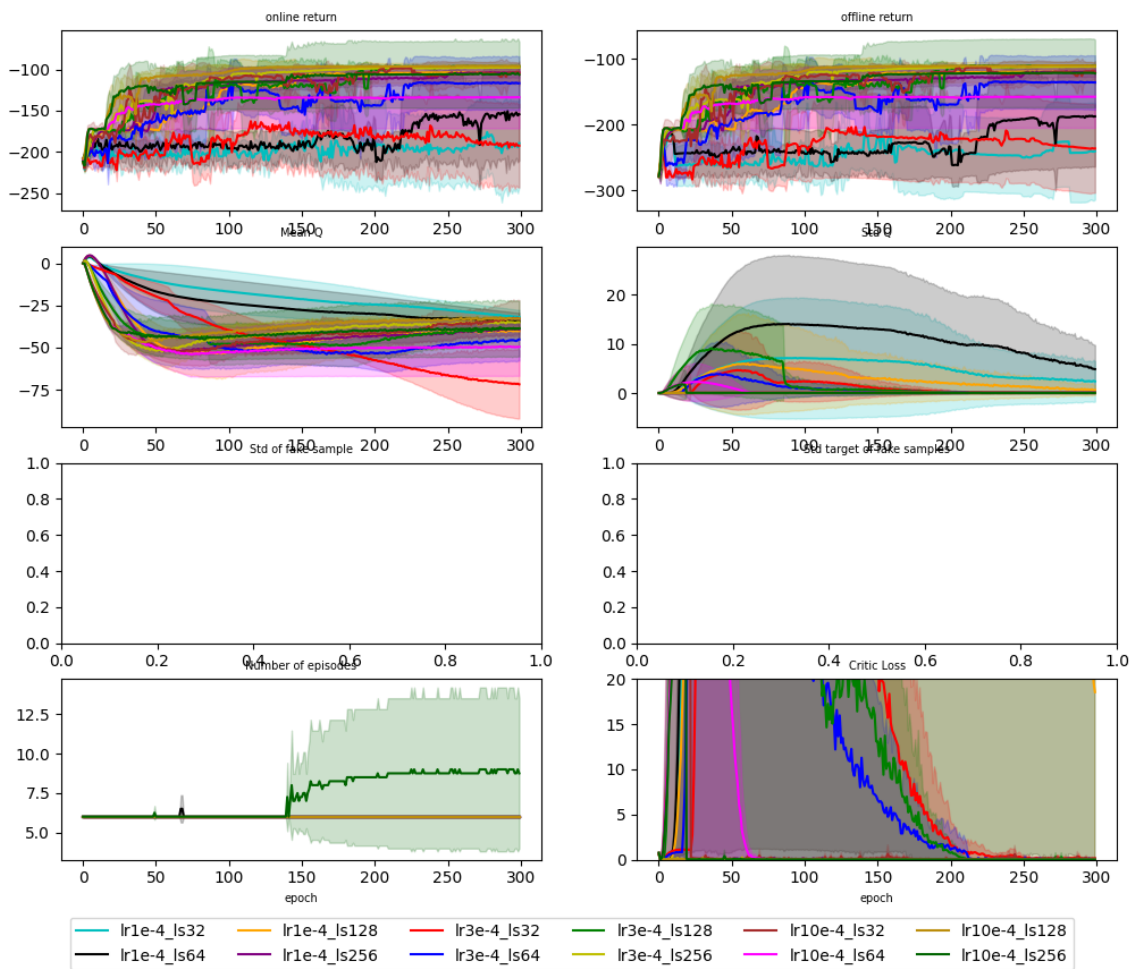
Senza stati terminali:

point/f003



Con stati terminali:

point/i001

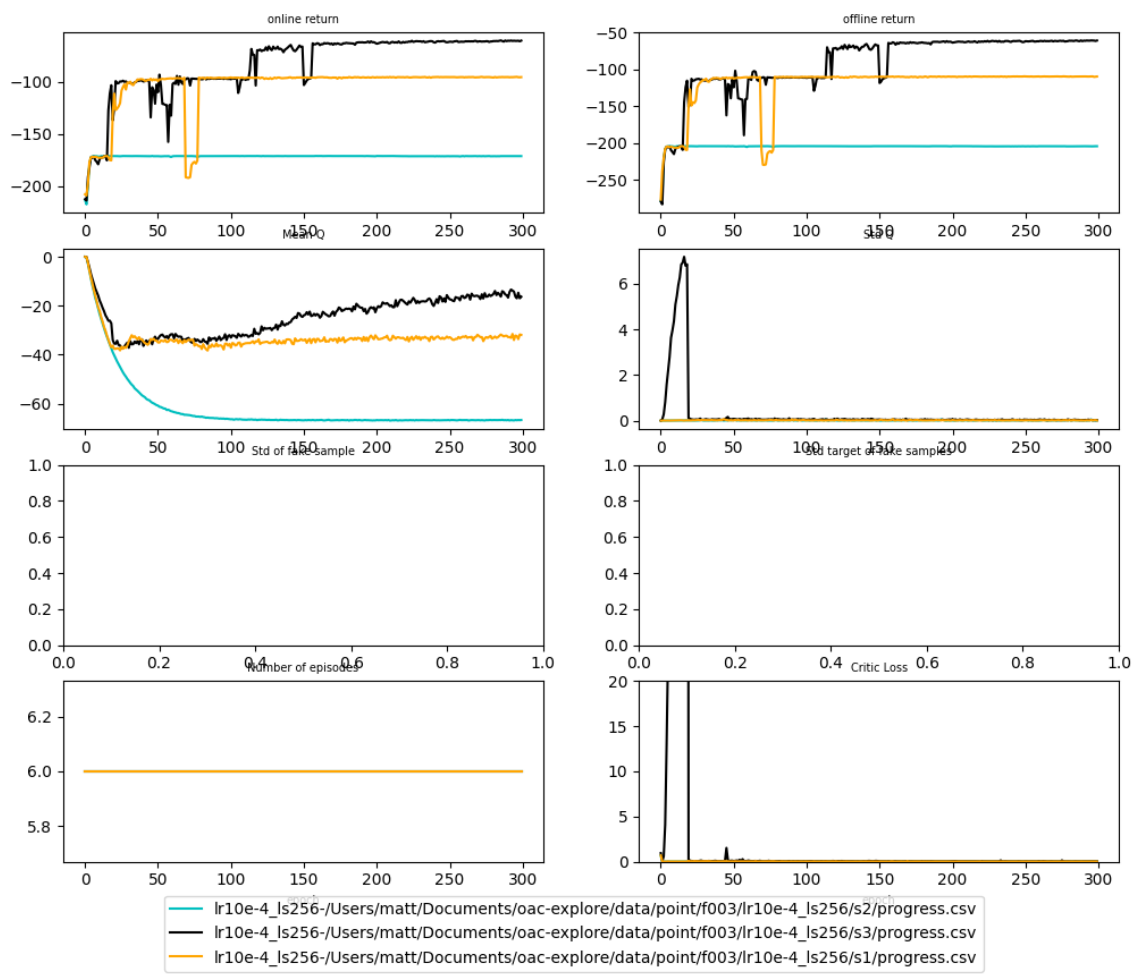


In entrambi i casi l'unica run ad arrivare infondo è con 256 neuroni e lr 1e-3 (10e-4).

Ma soltanto una delle 3 (delle 4 nel caso con stati terminali)

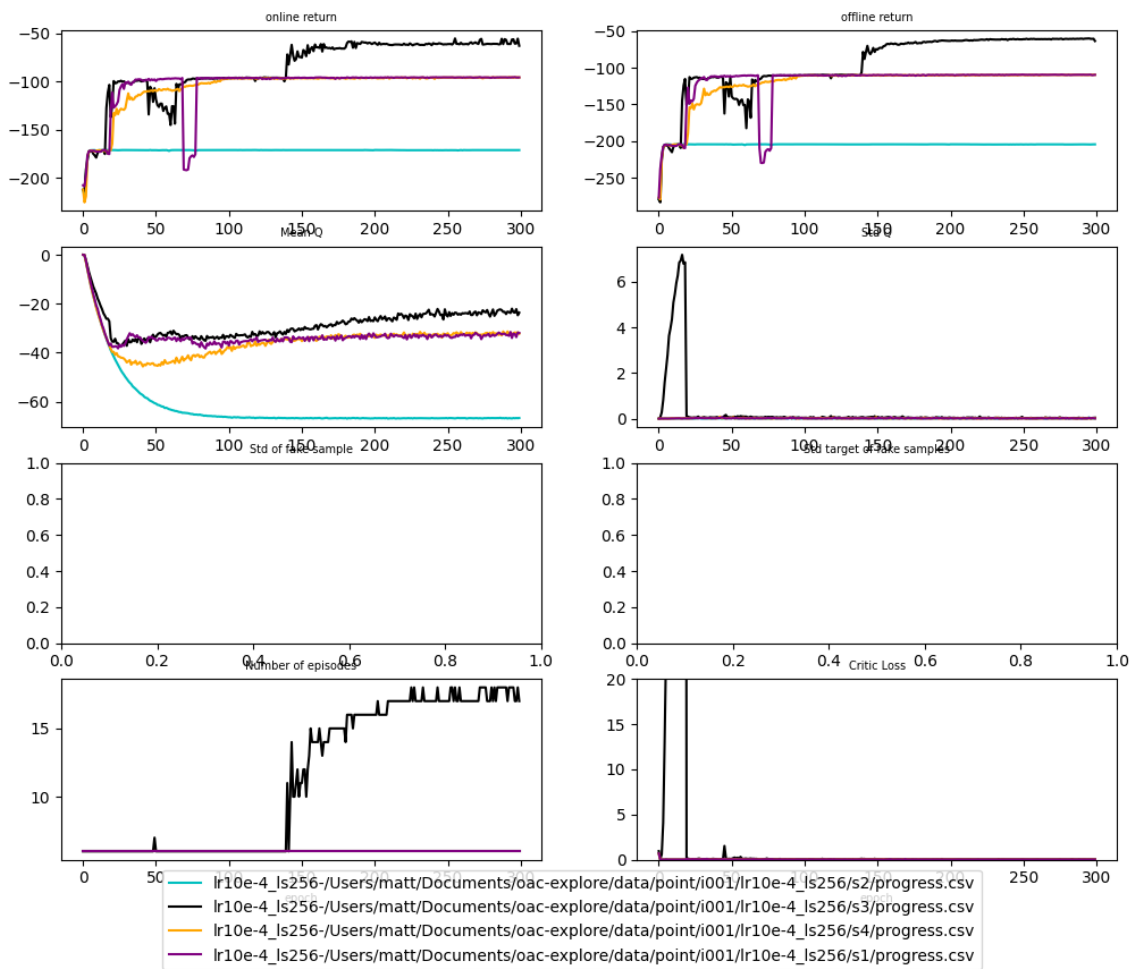
Senza stati terminali:

point/f003



Con stati terminali:

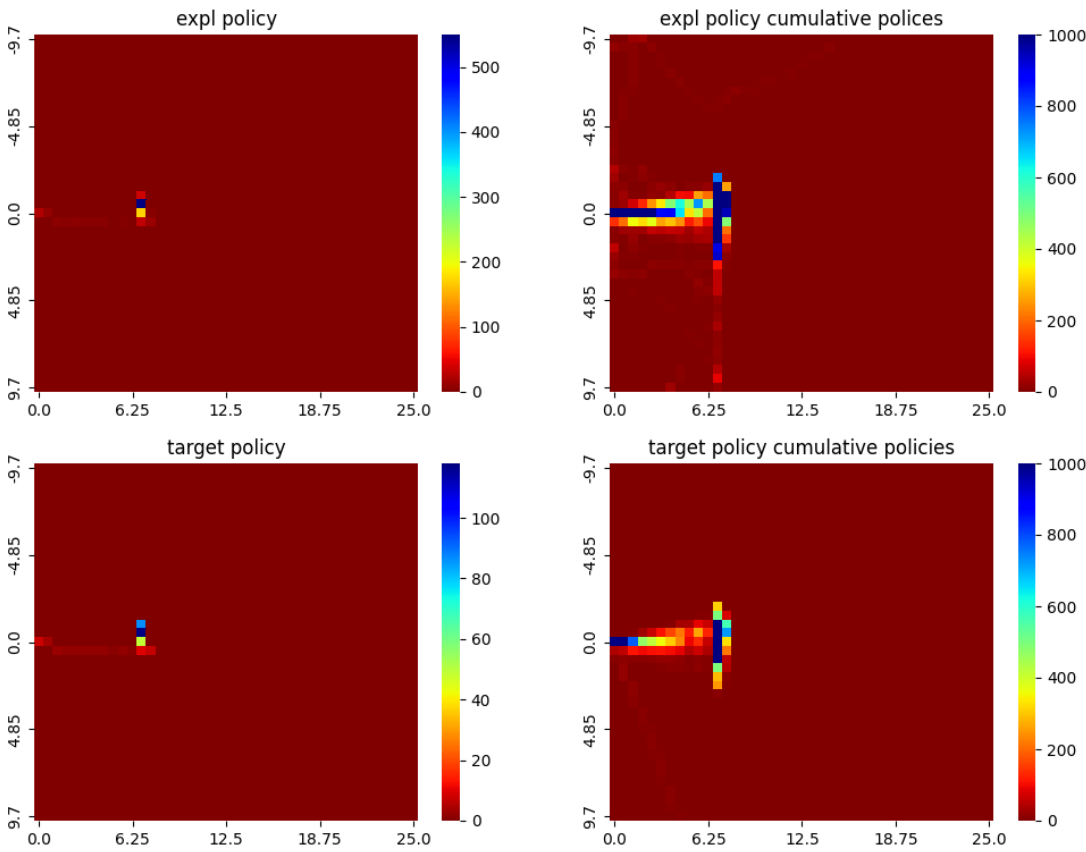
point/i001



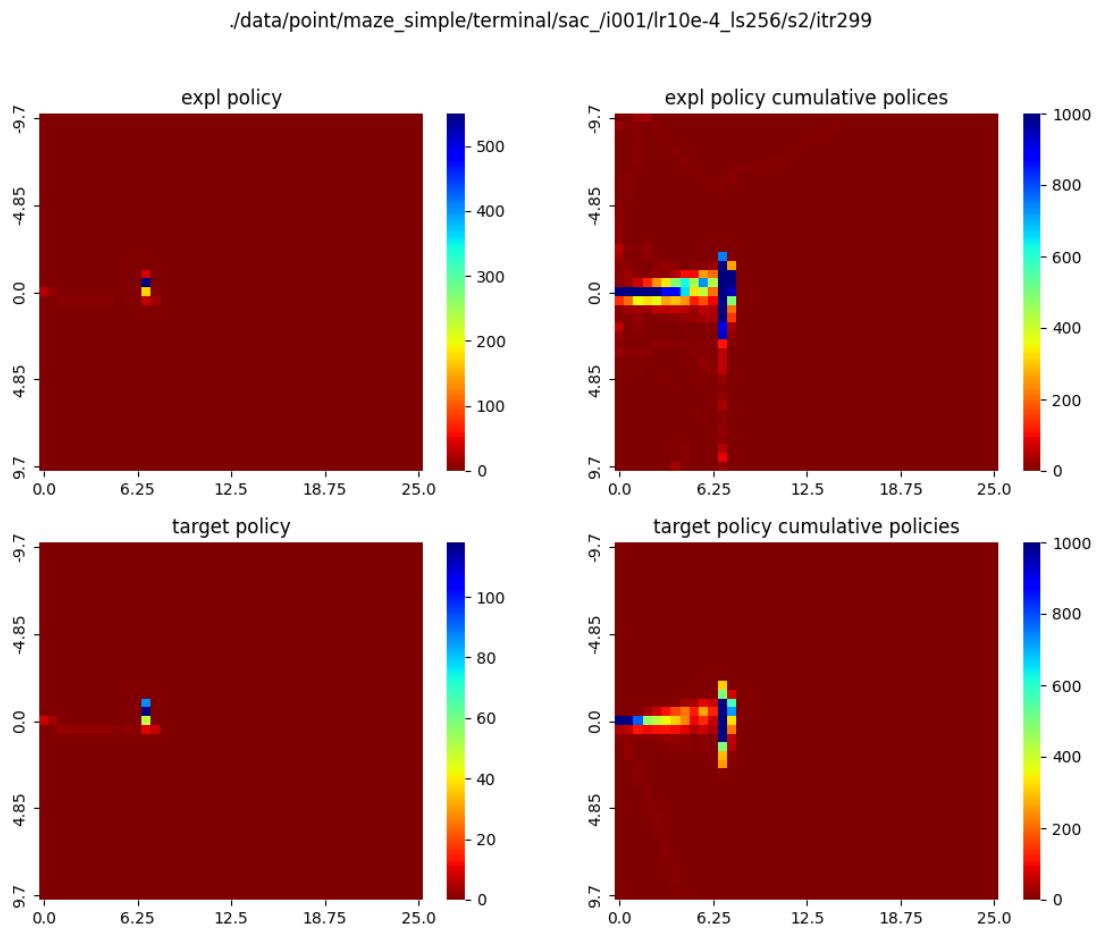
In entrambi i casi una delle 4 run non impara quasi niente e si blocca sulla politica che va semplicemente a destra (entrambi i casi seed=2)

Senza stati terminali:

./data/point/maze_simple/terminal/sac_f003/lr10e-4_ls256/s2/itr299



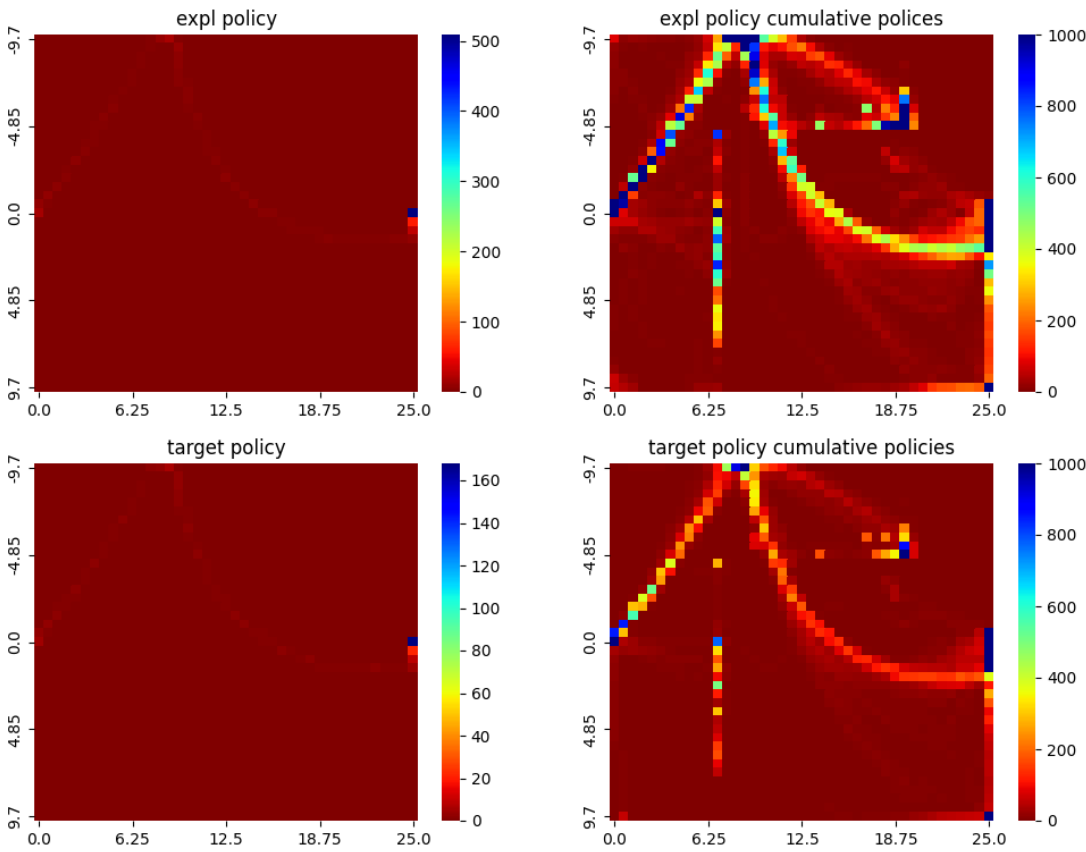
Con stati terminali:



Run che sono arrivati al goal (seed=3)

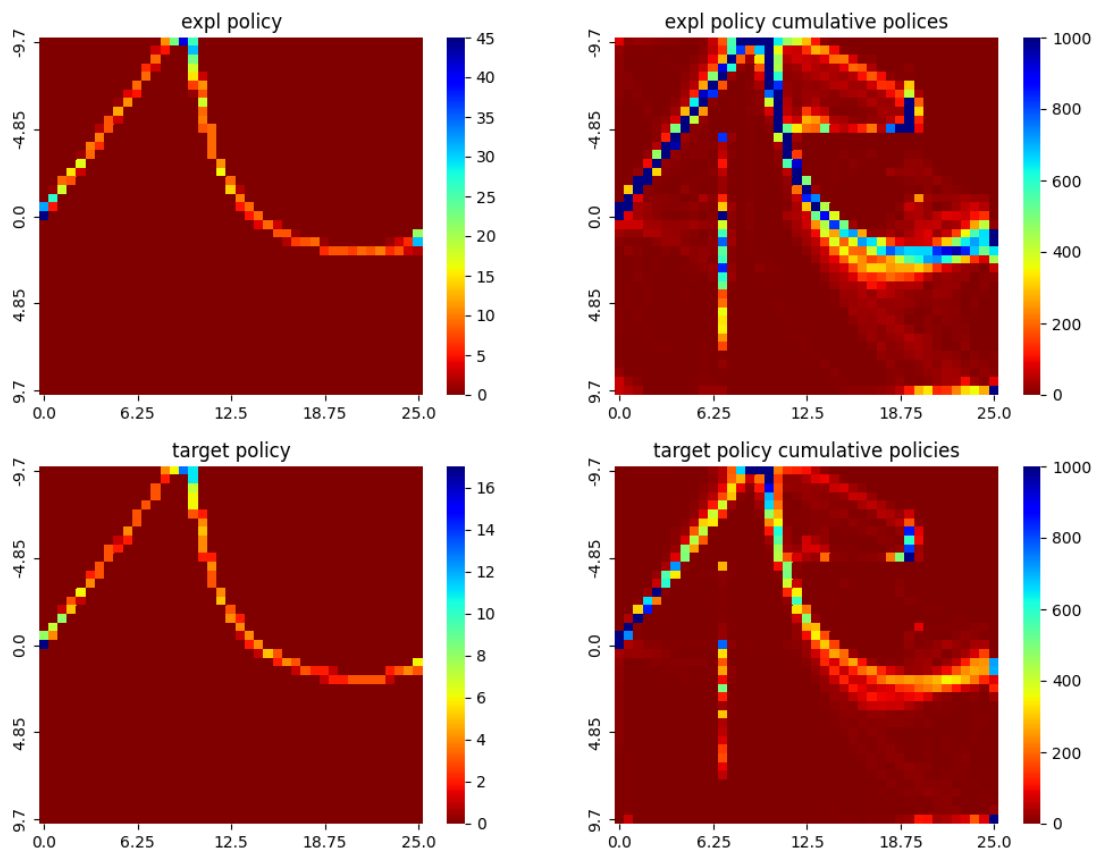
Senza stati terminali:

./data/point/maze_simple/terminal/sac_f003/lr10e-4_ls256/s3/itr299



Con stati terminali:

./data/point/maze_simple/terminal/sac_i001/lr10e-4_ls256/s3/itr299

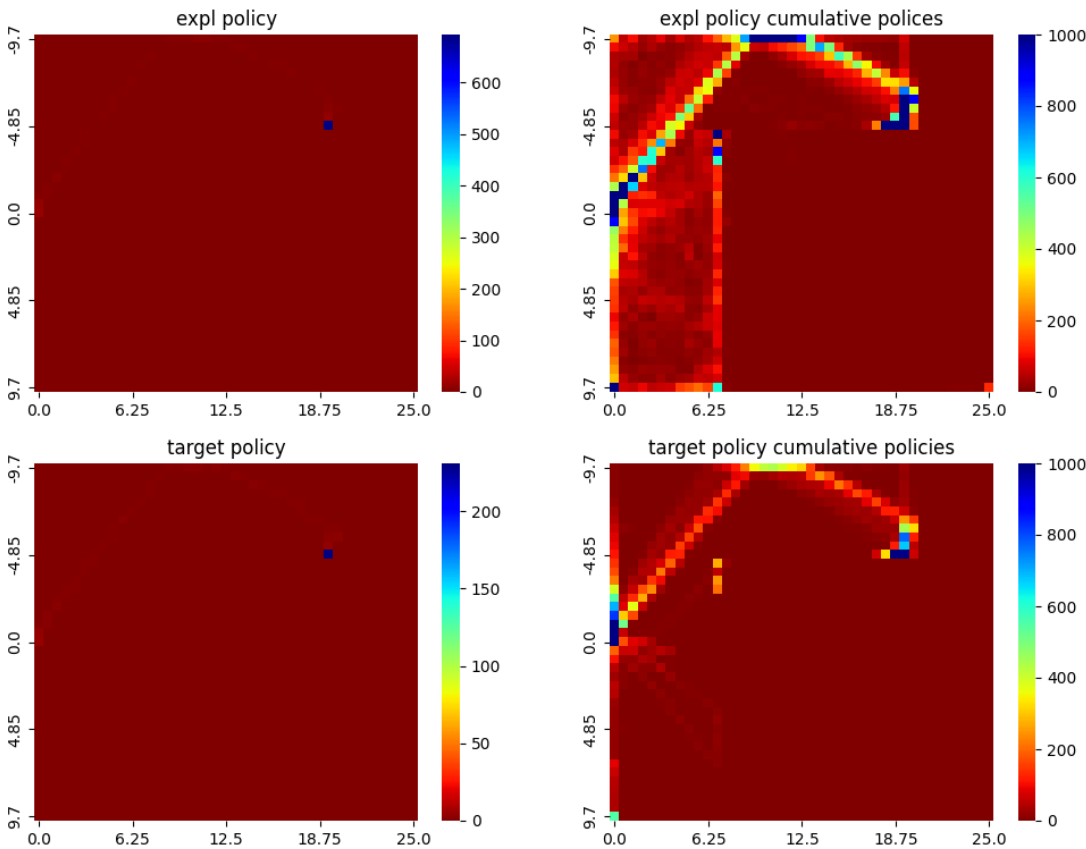


Una run con pochi neuroni riesce ad entrare nel buco ma fatica a rimanerci e ad avere in comportamento "stabile" una volta entrata

```
ls=32  
lr=10e-4
```

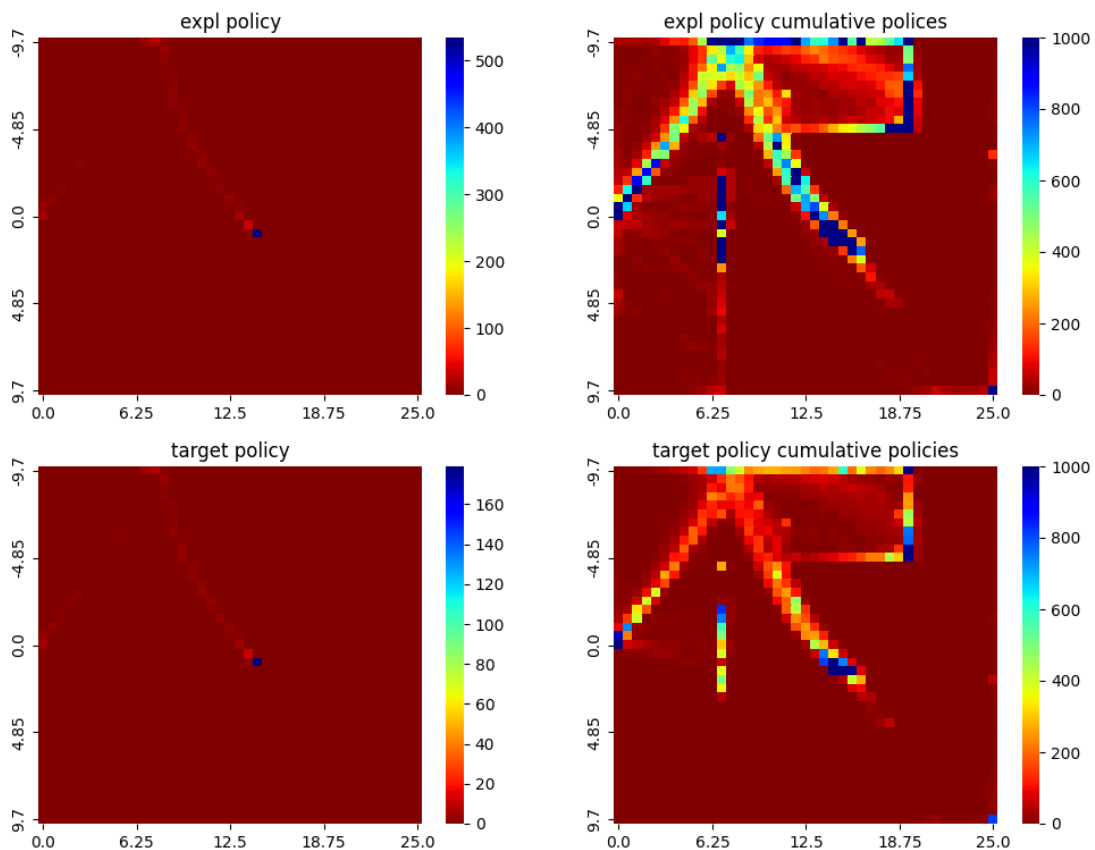
Senza stati terminali:

./data/point/maze_simple/terminal/sac_f003/lr10e-4_ls32/s3/itr299



Con stati terminali:

./data/point/maze_simple/terminal/sac_/i001/lr10e-4_ls32/s1/itr299



Se hai altre domande sulle run chiedimi pure

Io a questo punto opterei per reti più grandi con 128 o 256 neuroni, per il lr invece sono più indeciso, nel nostro caso l'esplorazione dovrebbe venire meno penalizzata dal lr qui secondo me entra in gioco una qualche dipendenza con l'euristica dell'entropia che fa sì che l'entropia converga a valori bassi presto. Lo stesso probabilmente succede con un numero di neuroni bassi con la quale l'entropia tende a rimanere molto alta.

L'altra domanda è se prendere per buoni questi hyper-par di SAC o continuare con il tuning. I parametri potrebbero essere i seguenti. Sono già stati tutti testati ma con reti da 32 o 64 neuroni

- `num_expl_steps_per_train_loop` tested =(500 1000 2000) made a big difference
- `num_trains_per_train_loop` didn't make much of a difference
- `batch_size` 256 which is SAC default showed the best performance on lower layer sizes
- `policy_lr` different from `qf_lr`: no appreciable difference on lower layers size
- `replay_buffer_size` tried 1e6 (SAC default) and 1e5 and didn't make much of difference
- `reward_scale` should make a big difference but we could decide never to tune it