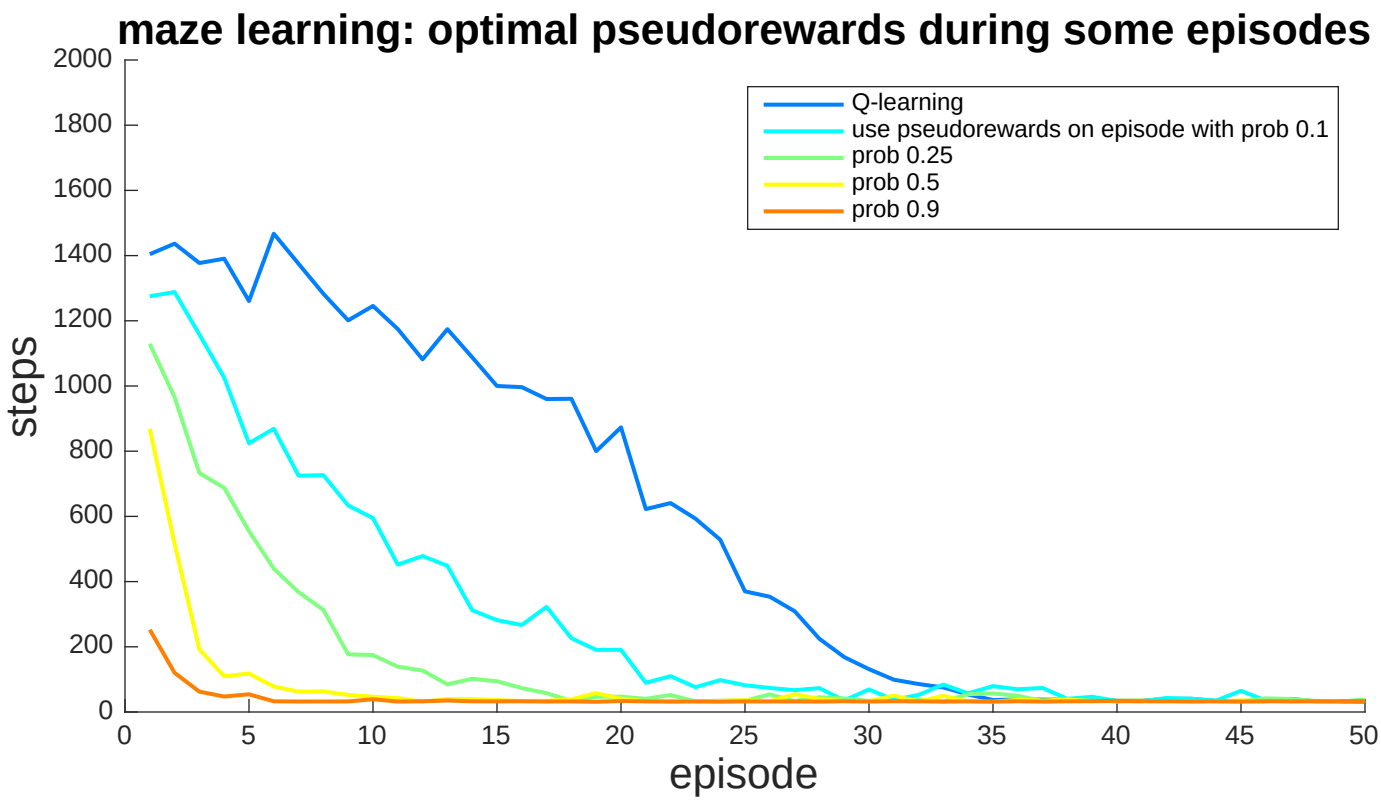
maze learning: optimal pseudorewards during some episodes
steps
episode
Q-learning
use pseudorewards on episode with prob 0.1
prob 0.25
prob 0.5
prob 0.9
0
200
400
600
800
1000
1200
1400
1600
1800
2000
0
5
10
15
20
25
30
35
40
45
50