Results

SARSA:

10000 average reward = 8.52 best avr reward = 9.1

20000 average reward = 8.54 best avr reward = 9.16

30000 average reward = 8.92 best avr reward = 9.16

ended in 36131

SARSA max:

10000 average reward = 8.29 best avr reward = 9.26

20000 average reward = 8.4 best avr reward = 9.26

ended in 21867

expected SARSA:

5000 average reward = 8.16 best avr reward = 9.16

10000 average reward = 8.49 best avr reward = 9.22

ended in 12231