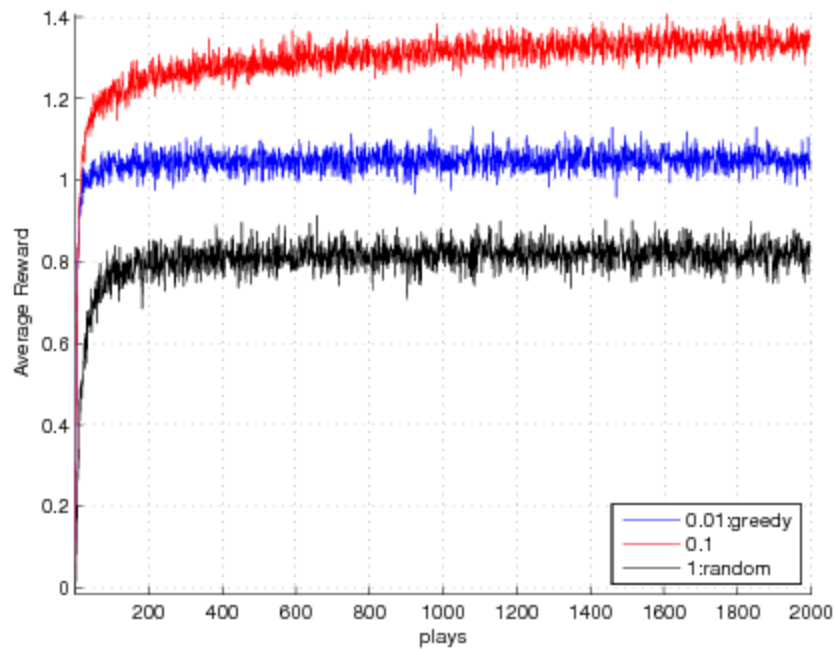
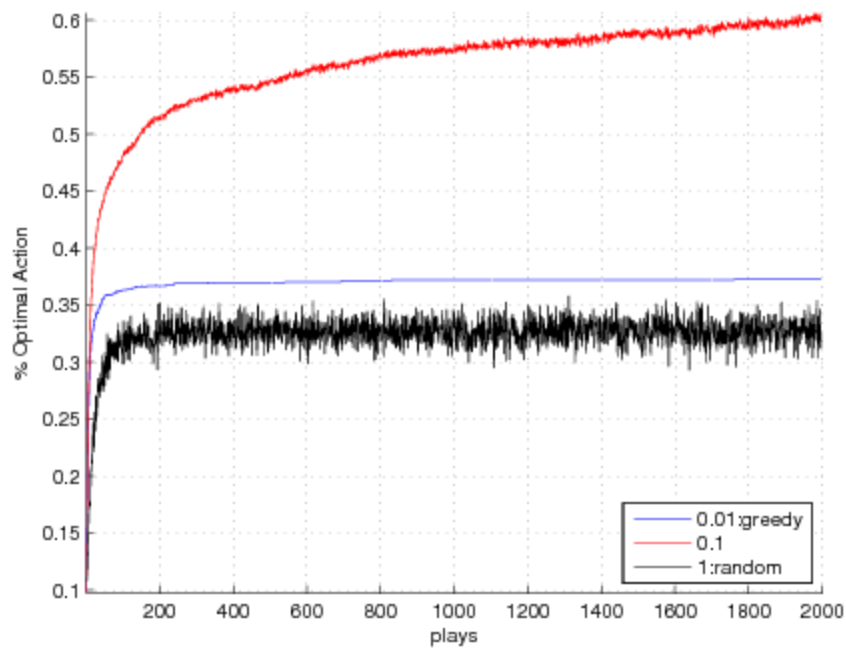
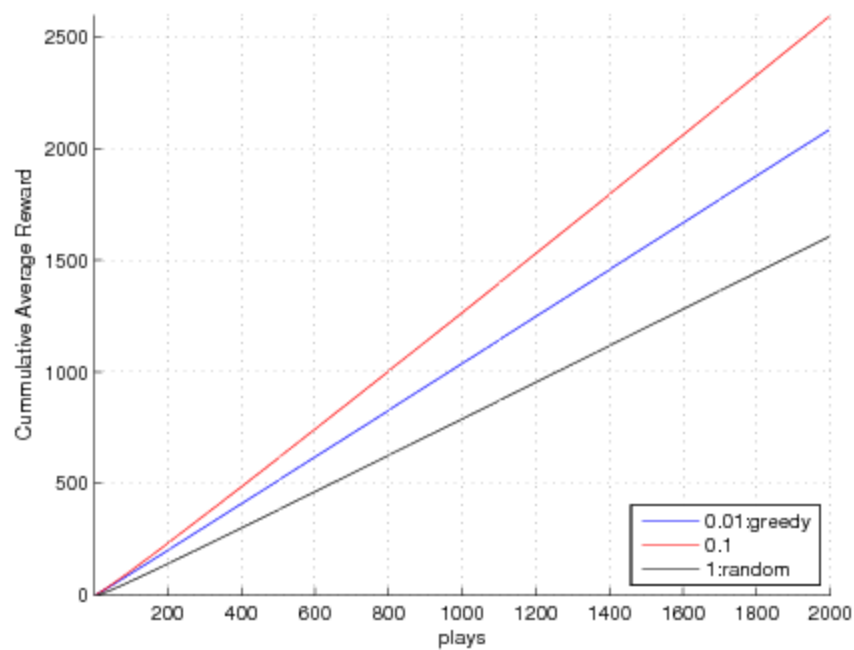




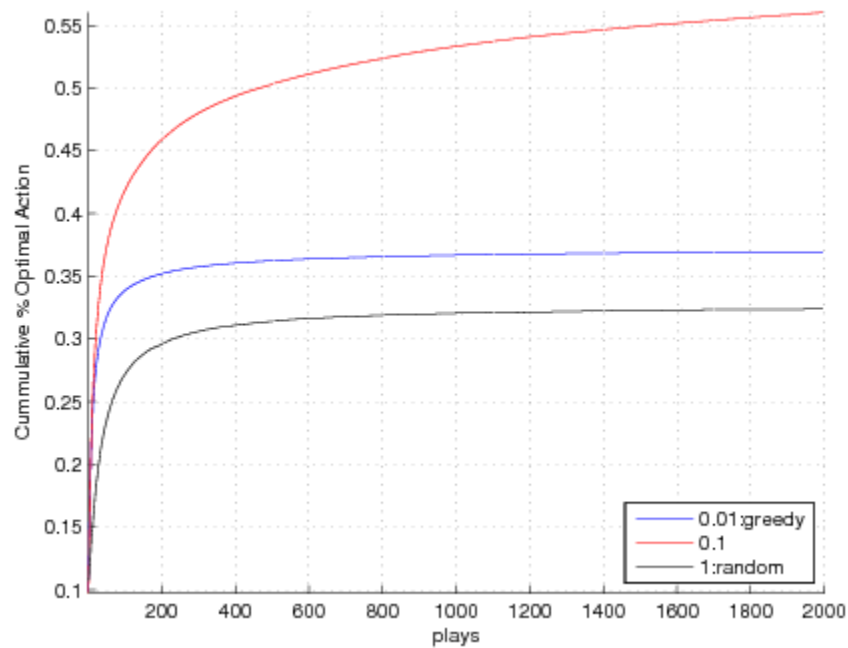
The average reward for several action value methods with standard deviation 1 and softmax action selection.



The percentage of times we select the optimal action when the standard deviation of rewards is 1 and softmax action selection



The cumulative average reward for several action value methods with standard deviation 1 and softmax action selection.



The cumulative percentage of times we select the optimal action when the standard deviation of rewards is 1 and softmax action selection.

[John Weatherwax](#)