5-step methods

- $\sigma = 0$, greedy
- $\sigma = 0$, epsilon_greedy
- $\sigma = 0$, boltzmann_exploration
- $\sigma = 0.25$, greedy
- $\sigma = 0.25$, epsilon_greedy
- $\sigma = 0.25$, boltzmann_exploration
- $\sigma = 0.5$, greedy
- $\sigma = 0.5$, epsilon_greedy
- $\sigma = 0.5$, boltzmann_exploration
- $\sigma = 0.75$, greedy
- $\sigma = 0.75$, epsilon_greedy
- $\sigma = 0.75$, boltzmann_exploration
- $\sigma = 1$, greedy
- $\sigma = 1$, epsilon_greedy
- $\sigma = 1$, boltzmann_exploration