**Ex1\_model\_free\_rl**

SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.0002 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.0373 | Mean Test Reward: 0.04 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.0384 | Mean Test Reward: 0.0386 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0325 | Mean Test Reward: 0.1056 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.0994 | Mean Test Reward: 0.1026 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.0384 | Mean Test Reward: 0.038 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.073 | Mean Test Reward: 0.0868 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0245 | Mean Test Reward: 0.0654 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.038 | Mean Test Reward: 0.0404 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.0609 | Mean Test Reward: 0.0672 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.4307 | Mean Test Reward: 0.5726 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.062 | Mean Test Reward: 0.0808 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0392 | Mean Test Reward: 0.3416 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.0806 | Mean Test Reward: 0.0832 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.5593 | Mean Test Reward: 0.7766 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.3984 | Mean Test Reward: 0.7834 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0478 | Mean Test Reward: 0.4662 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.0001 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.2005 | Mean Test Reward: 0.2032 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.5553 | Mean Test Reward: 0.7716 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.3572 | Mean Test Reward: 0.7746 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0452 | Mean Test Reward: 0.1952 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.304 | Mean Test Reward: 0.339 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.702 | Mean Test Reward: 0.7796 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.4958 | Mean Test Reward: 0.8184 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.2502 | Mean Test Reward: 0.6194 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0283 | Mean Test Reward: 0.1586 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.0002 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.6984 | Mean Test Reward: 0.7316 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.4549 | Mean Test Reward: 0.7306 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.1903 | Mean Test Reward: 0.4642 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0239 | Mean Test Reward: 0.065 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.0001 | Mean Test Reward: 0.0 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.0003 | Mean Test Reward: 0.0 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.0269 | Mean Test Reward: 0.0342 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.0331 | Mean Test Reward: 0.043 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0336 | Mean Test Reward: 0.1464 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.0529 | Mean Test Reward: 0.0584 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.0633 | Mean Test Reward: 0.0804 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0258 | Mean Test Reward: 0.0748 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.0166 | Mean Test Reward: 0.0152 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.1084 | Mean Test Reward: 0.1074 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.0654 | Mean Test Reward: 0.0664 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.039 | Mean Test Reward: 0.0428 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0518 | Mean Test Reward: 0.6802 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.0464 | Mean Test Reward: 0.0462 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.0211 | Mean Test Reward: 0.025 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.5431 | Mean Test Reward: 0.735 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.4012 | Mean Test Reward: 0.771 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0474 | Mean Test Reward: 0.7772 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.6136 | Mean Test Reward: 0.6346 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.5882 | Mean Test Reward: 0.7786 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.4096 | Mean Test Reward: 0.821 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0461 | Mean Test Reward: 0.216 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.6812 | Mean Test Reward: 0.7394 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.4615 | Mean Test Reward: 0.0 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.2684 | Mean Test Reward: 0.6702 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0261 | Mean Test Reward: 0.2936 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.1102 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.3758 | Mean Test Reward: 0.7336 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.1953 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.026 | Mean Test Reward: 0.0342 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.2419 | Mean Test Reward: 0.2418 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.0714 | Mean Test Reward: 0.0754 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.0466 | Mean Test Reward: 0.0506 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.0771 | Mean Test Reward: 0.1 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0276 | Mean Test Reward: 0.0632 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.0382 | Mean Test Reward: 0.043 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.0338 | Mean Test Reward: 0.0358 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0218 | Mean Test Reward: 0.0336 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.0002 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.0385 | Mean Test Reward: 0.045 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.0609 | Mean Test Reward: 0.0752 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0465 | Mean Test Reward: 0.6816 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.4473 | Mean Test Reward: 0.4806 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.5656 | Mean Test Reward: 0.7782 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.3864 | Mean Test Reward: 0.7682 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0562 | Mean Test Reward: 0.6804 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.1367 | Mean Test Reward: 0.1456 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.723 | Mean Test Reward: 0.7658 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.5629 | Mean Test Reward: 0.786 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.3624 | Mean Test Reward: 0.7752 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0526 | Mean Test Reward: 0.6258 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.7272 | Mean Test Reward: 0.771 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.7162 | Mean Test Reward: 0.8198 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.523 | Mean Test Reward: 0.816 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.2776 | Mean Test Reward: 0.3702 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0368 | Mean Test Reward: 0.5264 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.7351 | Mean Test Reward: 0.8134 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.4759 | Mean Test Reward: 0.8078 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.22 | Mean Test Reward: 0.7786 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0249 | Mean Test Reward: 0.0244 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.0001 | Mean Test Reward: 0.0 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.038 | Mean Test Reward: 0.0324 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.0892 | Mean Test Reward: 0.096 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.0697 | Mean Test Reward: 0.095 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0238 | Mean Test Reward: 0.0452 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.0753 | Mean Test Reward: 0.0788 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.0762 | Mean Test Reward: 0.0788 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.0638 | Mean Test Reward: 0.0768 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.0439 | Mean Test Reward: 0.056 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0482 | Mean Test Reward: 0.3274 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.0348 | Mean Test Reward: 0.0374 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.0872 | Mean Test Reward: 0.089 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.0539 | Mean Test Reward: 0.0746 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0426 | Mean Test Reward: 0.6176 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.087 | Mean Test Reward: 0.0882 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.13 | Mean Test Reward: 0.1306 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.5613 | Mean Test Reward: 0.7598 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.1755 | Mean Test Reward: 0.0 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0476 | Mean Test Reward: 0.7306 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.1352 | Mean Test Reward: 0.1312 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.0781 | Mean Test Reward: 0.0 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.0951 | Mean Test Reward: 0.0 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0419 | Mean Test Reward: 0.0 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.7472 | Mean Test Reward: 0.7796 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.0203 | Mean Test Reward: 0.0 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.0329 | Mean Test Reward: 0.0 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.056 | Mean Test Reward: 0.0 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0325 | Mean Test Reward: 0.0 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.0018 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.0067 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.0267 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.0507 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0255 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.1145 | Mean Test Reward: 0.1138 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.029 | Mean Test Reward: 0.0324 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0923 | Mean Test Reward: 0.1238 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0334 | Mean Test Reward: 0.112 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.0406 | Mean Test Reward: 0.0416 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0924 | Mean Test Reward: 0.1138 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.045 | Mean Test Reward: 0.0582 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0261 | Mean Test Reward: 0.065 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0194 | Mean Test Reward: 0.0214 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0711 | Mean Test Reward: 0.0832 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0367 | Mean Test Reward: 0.1918 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.0465 | Mean Test Reward: 0.0478 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.5505 | Mean Test Reward: 0.774 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.3856 | Mean Test Reward: 0.7798 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0478 | Mean Test Reward: 0.4516 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.589 | Mean Test Reward: 0.608 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.5574 | Mean Test Reward: 0.774 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.3671 | Mean Test Reward: 0.7432 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0464 | Mean Test Reward: 0.278 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.0001 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.5026 | Mean Test Reward: 0.8156 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.2571 | Mean Test Reward: 0.8156 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0321 | Mean Test Reward: 0.2508 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.6862 | Mean Test Reward: 0.7284 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.7311 | Mean Test Reward: 0.8142 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.4699 | Mean Test Reward: 0.808 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.2258 | Mean Test Reward: 0.5916 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0245 | Mean Test Reward: 0.0288 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.0435 | Mean Test Reward: 0.044 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.1513 | Mean Test Reward: 0.1552 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0279 | Mean Test Reward: 0.0294 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.034 | Mean Test Reward: 0.0384 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0275 | Mean Test Reward: 0.0652 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.048 | Mean Test Reward: 0.054 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.0481 | Mean Test Reward: 0.0558 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0412 | Mean Test Reward: 0.0448 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0442 | Mean Test Reward: 0.0548 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0296 | Mean Test Reward: 0.079 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.0817 | Mean Test Reward: 0.0814 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.0782 | Mean Test Reward: 0.0854 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0312 | Mean Test Reward: 0.0304 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0536 | Mean Test Reward: 0.067 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0333 | Mean Test Reward: 0.138 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.0001 | Mean Test Reward: 0.0 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.4976 | Mean Test Reward: 0.515 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.5538 | Mean Test Reward: 0.7804 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.4111 | Mean Test Reward: 0.7822 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0529 | Mean Test Reward: 0.7752 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.0001 | Mean Test Reward: 0.0 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.1282 | Mean Test Reward: 0.1468 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.5886 | Mean Test Reward: 0.7776 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.4214 | Mean Test Reward: 0.822 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0527 | Mean Test Reward: 0.4966 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.0001 | Mean Test Reward: 0.0 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.7321 | Mean Test Reward: 0.7818 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.4068 | Mean Test Reward: 0.0 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.331 | Mean Test Reward: 0.7316 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.032 | Mean Test Reward: 0.6488 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.0205 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.042 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.2677 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.2092 | Mean Test Reward: 0.24 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0276 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.9999