**Ex1\_model\_free\_rl**

SARSA Mean Training Reward: 0.0616 | Mean Test Reward: 0.0576 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.0716 | Mean Test Reward: 0.0692 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.0332 | Mean Test Reward: 0.0298 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.0399 | Mean Test Reward: 0.051 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0235 | Mean Test Reward: 0.0422 | 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: 0.2994 | Mean Test Reward: 0.312 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.0201 | Mean Test Reward: 0.0168 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.0789 | Mean Test Reward: 0.0994 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0379 | Mean Test Reward: 0.1744 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.9999

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

SARSA Mean Training Reward: 0.0393 | Mean Test Reward: 0.0464 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.0344 | Mean Test Reward: 0.0334 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.0728 | Mean Test Reward: 0.0972 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0449 | Mean Test Reward: 0.3244 | 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: 0.0997 | Mean Test Reward: 0.1036 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.5758 | Mean Test Reward: 0.7704 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.3904 | Mean Test Reward: 0.7746 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0491 | Mean Test Reward: 0.7674 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.9999

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

SARSA Mean Training Reward: 0.2907 | Mean Test Reward: 0.2944 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.52 | Mean Test Reward: 0.7716 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.3588 | Mean Test Reward: 0.6806 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0419 | Mean Test Reward: 0.6456 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.9999

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

SARSA Mean Training Reward: 0.7241 | Mean Test Reward: 0.7674 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.4527 | Mean Test Reward: 0.8272 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.2395 | Mean Test Reward: 0.8084 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0267 | Mean Test Reward: 0.322 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.0001 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.4274 | Mean Test Reward: 0.8202 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.1976 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0226 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.0419 | Mean Test Reward: 0.0366 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.1598 | Mean Test Reward: 0.1626 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.0378 | Mean Test Reward: 0.0394 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.0333 | Mean Test Reward: 0.0374 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0259 | Mean Test Reward: 0.0658 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.0457 | Mean Test Reward: 0.0452 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.0782 | Mean Test Reward: 0.0762 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.0479 | Mean Test Reward: 0.0516 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.0406 | Mean Test Reward: 0.0562 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0312 | Mean Test Reward: 0.1116 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.0001 | Mean Test Reward: 0.0 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.0745 | Mean Test Reward: 0.0794 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.0719 | Mean Test Reward: 0.076 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.0663 | Mean Test Reward: 0.0828 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0418 | Mean Test Reward: 0.4562 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.9999

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

SARSA Mean Training Reward: 0.1315 | Mean Test Reward: 0.1436 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.3907 | Mean Test Reward: 0.6442 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.4122 | Mean Test Reward: 0.7818 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0494 | Mean Test Reward: 0.7784 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.1593 | Mean Test Reward: 0.1558 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.3808 | Mean Test Reward: 0.413 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.5867 | Mean Test Reward: 0.781 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.4056 | Mean Test Reward: 0.7742 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0487 | Mean Test Reward: 0.624 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.9999

SARSA Mean Training Reward: 0.3982 | Mean Test Reward: 0.4334 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.99

SARSA Mean Training Reward: 0.1094 | Mean Test Reward: 0.0 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.4804 | Mean Test Reward: 0.0 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.2688 | Mean Test Reward: 0.8116 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.03 | Mean Test Reward: 0.0862 | 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: 0.6614 | Mean Test Reward: 0.7364 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.995

SARSA Mean Training Reward: 0.3848 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.999

SARSA Mean Training Reward: 0.1809 | Mean Test Reward: 0.2872 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.9995

SARSA Mean Training Reward: 0.0251 | Mean Test Reward: 0.0504 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.0442 | Mean Test Reward: 0.0524 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.051 | Mean Test Reward: 0.0554 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.0174 | Mean Test Reward: 0.0172 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.0771 | Mean Test Reward: 0.1036 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0284 | Mean Test Reward: 0.1106 | 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: 0.096 | Mean Test Reward: 0.099 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.0683 | Mean Test Reward: 0.084 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.0686 | Mean Test Reward: 0.0896 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0254 | Mean Test Reward: 0.086 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.0618 | Mean Test Reward: 0.0614 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.99

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

Q-Learning Mean Training Reward: 0.0624 | Mean Test Reward: 0.0826 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.055 | Mean Test Reward: 0.0684 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0459 | Mean Test Reward: 0.3326 | 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: 0.0607 | Mean Test Reward: 0.0676 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.5745 | Mean Test Reward: 0.7788 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.3996 | Mean Test Reward: 0.6242 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0528 | Mean Test Reward: 0.6968 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.2612 | Mean Test Reward: 0.2642 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.6123 | Mean Test Reward: 0.774 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.5578 | Mean Test Reward: 0.7802 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.3637 | Mean Test Reward: 0.7832 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0477 | Mean Test Reward: 0.511 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.7134 | Mean Test Reward: 0.8206 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.4641 | Mean Test Reward: 0.8196 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.2743 | Mean Test Reward: 0.7806 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0375 | Mean Test Reward: 0.2066 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.7439 | Mean Test Reward: 0.7232 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.6737 | Mean Test Reward: 0.7282 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.4704 | Mean Test Reward: 0.7338 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.2192 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0286 | Mean Test Reward: 0.0552 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.0396 | Mean Test Reward: 0.044 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.0426 | Mean Test Reward: 0.046 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.075 | Mean Test Reward: 0.1104 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0341 | Mean Test Reward: 0.1744 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.281 | Mean Test Reward: 0.2938 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.02 | Mean Test Reward: 0.0228 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.0607 | Mean Test Reward: 0.0698 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.0934 | Mean Test Reward: 0.1376 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0271 | Mean Test Reward: 0.1182 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.9999

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

Q-Learning Mean Training Reward: 0.0636 | Mean Test Reward: 0.063 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.1217 | Mean Test Reward: 0.1452 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.1466 | Mean Test Reward: 0.2372 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0381 | Mean Test Reward: 0.2082 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.085 | Mean Test Reward: 0.0864 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.591 | Mean Test Reward: 0.772 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.1509 | Mean Test Reward: 0.0 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0544 | Mean Test Reward: 0.8222 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.9999

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

Q-Learning Mean Training Reward: 0.5641 | Mean Test Reward: 0.6188 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.0872 | Mean Test Reward: 0.0 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.0898 | Mean Test Reward: 0.0 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0398 | Mean Test Reward: 0.0 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.9999

Q-Learning Mean Training Reward: 0.7449 | Mean Test Reward: 0.765 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.99

Q-Learning Mean Training Reward: 0.0096 | Mean Test Reward: 0.0 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.995

Q-Learning Mean Training Reward: 0.0313 | Mean Test Reward: 0.0 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.062 | 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: 0.0032 | 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: 0.0241 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.999

Q-Learning Mean Training Reward: 0.0483 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.9995

Q-Learning Mean Training Reward: 0.0276 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.0001 | Mean Test Reward: 0.0 | 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: 0.0885 | Mean Test Reward: 0.0958 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.081 | Mean Test Reward: 0.11 | gamma=0.95, alpha=0.001, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0279 | Mean Test Reward: 0.0442 | 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: 0.2767 | Mean Test Reward: 0.3038 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0122 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0075 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.005, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0122 | Mean Test Reward: 0.0 | 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: 0.0842 | Mean Test Reward: 0.0768 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0045 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0082 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0133 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.01, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.0006 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.0043 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0036 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0062 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.05, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0117 | Mean Test Reward: 0.0 | 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: 0.0959 | Mean Test Reward: 0.099 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0048 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0046 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0105 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.1, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.052 | Mean Test Reward: 0.0634 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.0002 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0017 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0026 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0092 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.5, eps=1, eps\_decay=0.9999

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

Expected SARSA Mean Training Reward: 0.0 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0006 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0023 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0075 | Mean Test Reward: 0.0 | gamma=0.95, alpha=0.9, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.0004 | Mean Test Reward: 0.0 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.0332 | Mean Test Reward: 0.03 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.1129 | Mean Test Reward: 0.1282 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0255 | Mean Test Reward: 0.0282 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.019 | Mean Test Reward: 0.0318 | gamma=1, alpha=0.001, eps=1, eps\_decay=0.9999

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

Expected SARSA Mean Training Reward: 0.0417 | Mean Test Reward: 0.0402 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0638 | Mean Test Reward: 0.0754 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.066 | Mean Test Reward: 0.0888 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0103 | Mean Test Reward: 0.0 | gamma=1, alpha=0.005, eps=1, eps\_decay=0.9999

Expected SARSA Mean Training Reward: 0.0003 | Mean Test Reward: 0.0 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.99

Expected SARSA Mean Training Reward: 0.0005 | Mean Test Reward: 0.0 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0056 | Mean Test Reward: 0.0 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0093 | Mean Test Reward: 0.0 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0106 | Mean Test Reward: 0.0 | gamma=1, alpha=0.01, eps=1, eps\_decay=0.9999

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

Expected SARSA Mean Training Reward: 0.0012 | Mean Test Reward: 0.0 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0012 | Mean Test Reward: 0.0 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0056 | Mean Test Reward: 0.0 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0099 | Mean Test Reward: 0.0 | gamma=1, alpha=0.05, eps=1, eps\_decay=0.9999

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

Expected SARSA Mean Training Reward: 0.0064 | Mean Test Reward: 0.0 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0013 | Mean Test Reward: 0.0 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0037 | Mean Test Reward: 0.0 | gamma=1, alpha=0.1, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0099 | Mean Test Reward: 0.0 | 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: 0.0144 | Mean Test Reward: 0.0116 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.003 | Mean Test Reward: 0.0 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0027 | Mean Test Reward: 0.0 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.0099 | Mean Test Reward: 0.0 | gamma=1, alpha=0.5, eps=1, eps\_decay=0.9999

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

Expected SARSA Mean Training Reward: 0.0015 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.995

Expected SARSA Mean Training Reward: 0.0018 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.999

Expected SARSA Mean Training Reward: 0.0021 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.9995

Expected SARSA Mean Training Reward: 0.006 | Mean Test Reward: 0.0 | gamma=1, alpha=0.9, eps=1, eps\_decay=0.9999