number1：激活函数全用relu，hidden\_size=200，学习率1e-4

每隔5e3步评估一次，评估时探索率衰减0.95 batch\_size=256

number2：激活函数全用relu，hidden\_size=200，学习率1e-4

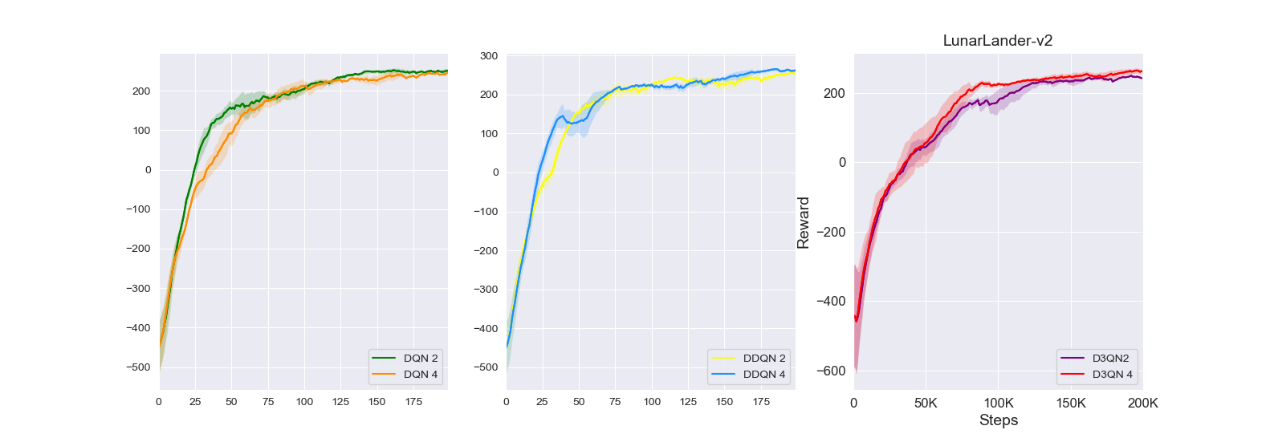
每隔1e3步评估一次，评估时探索率衰减0.99 batch\_size=256

number3：：激活函数全用relu，hidden\_size=200，学习率2e-4

每隔1e3步评估一次，评估时探索率衰减0.99 batch\_size=256

number4：激活函数全用relu，hidden\_size=200，学习率1e-4

每隔1e3步评估一次，评估时探索率衰减0.99 batch\_size=512



number5：激活函数全用relu，hidden\_size=128，学习率1e-4

每隔1e3步评估一次，评估时探索率衰减0.99 batch\_size=256

number6：激活函数全用relu，hidden\_size=200，学习率1e-4

每隔1e3步评估一次，评估时探索率衰减0.99 batch\_size=512，起始随机1e3步

number7：激活函数全用relu，hidden\_size=200，学习率1e-4

每隔1e3步评估一次，评估时探索率衰减0.99 batch\_size=256，起始随机256步

number8：激活函数全用relu，hidden\_size=200，学习率1e-4

每隔1e3步评估一次，探索率保持0.1不变，batch\_size=512

number9：激活函数全用relu，hidden\_size=200，学习率1e-4

每隔1e3步评估一次，评估时探索率衰减0.99 batch\_size=512 每走一步学习一次

number10：激活函数全用tanh，hidden\_size=200，学习率1e-4

每隔1e3步评估一次，评估时探索率衰减0.99 batch\_size=512

number10：激活函数全用tanh，hidden\_size=200，学习率1e-4

每隔1e3步评估一次，评估时探索率衰减0.99 batch\_size=256