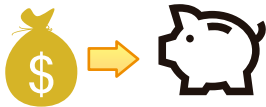


in the end...

take all rewards



goal: maximize the cumulative reward

exploitation v.s. exploration