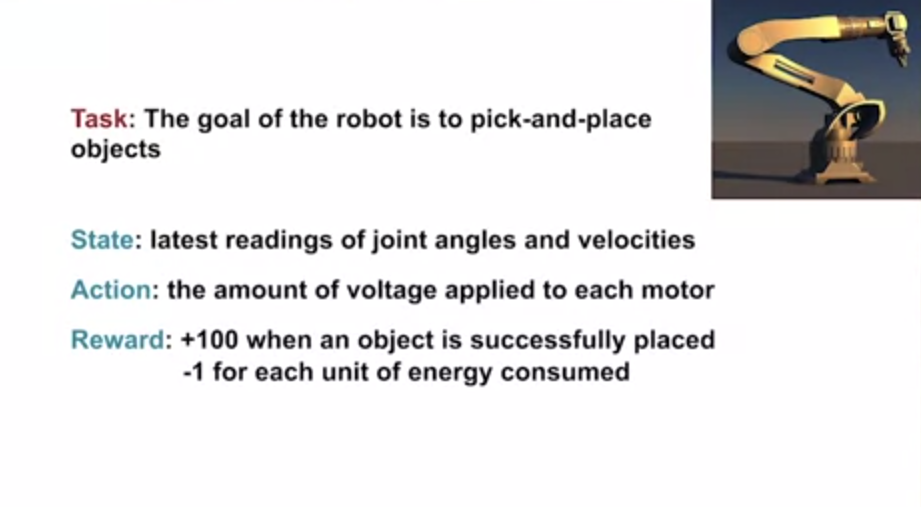
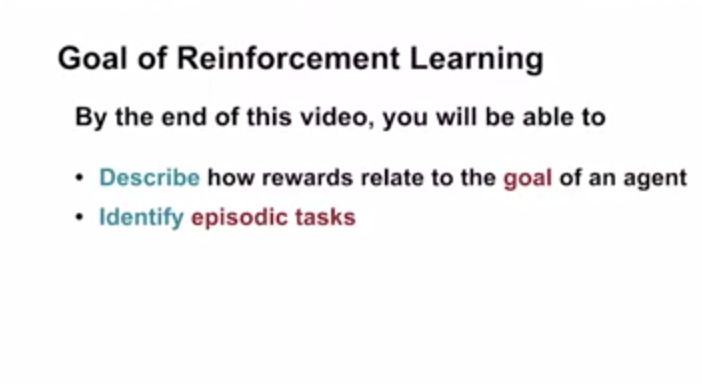


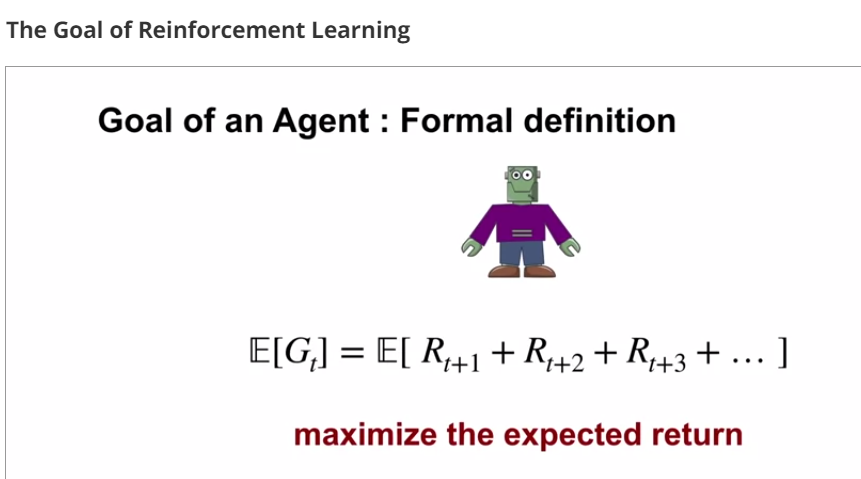
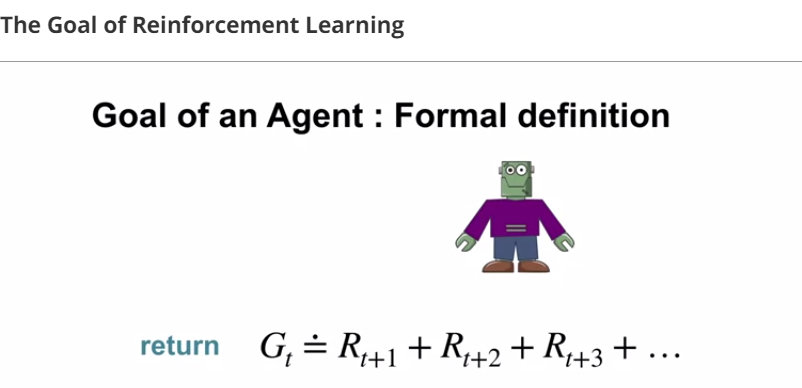
Time seto can be 1 msec or 1 month



Note:RL use for Goal driven problems.



How to maximise Future rewards ??

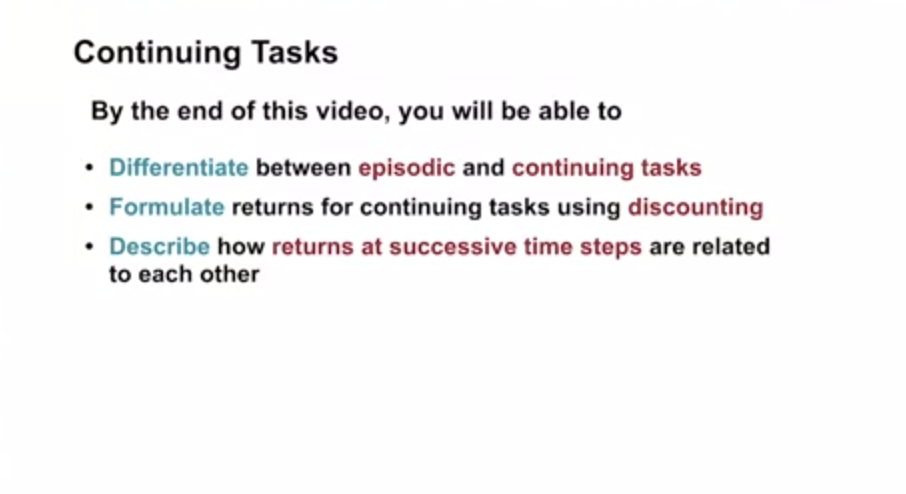
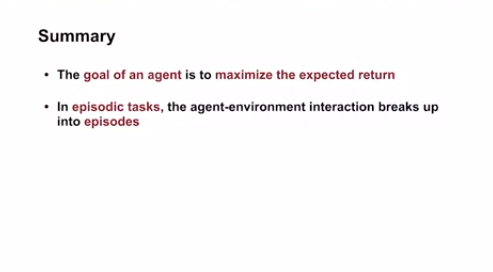
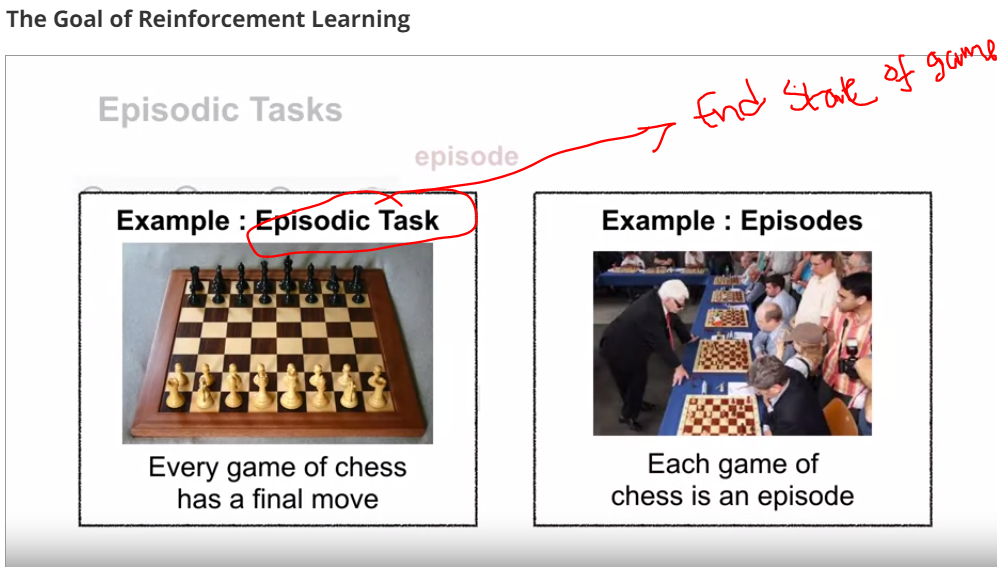
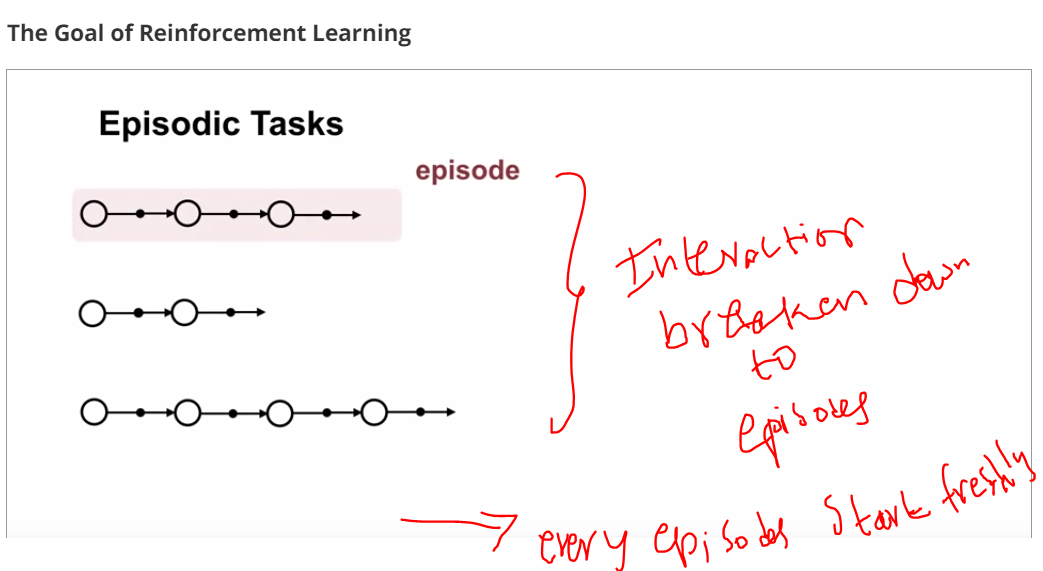


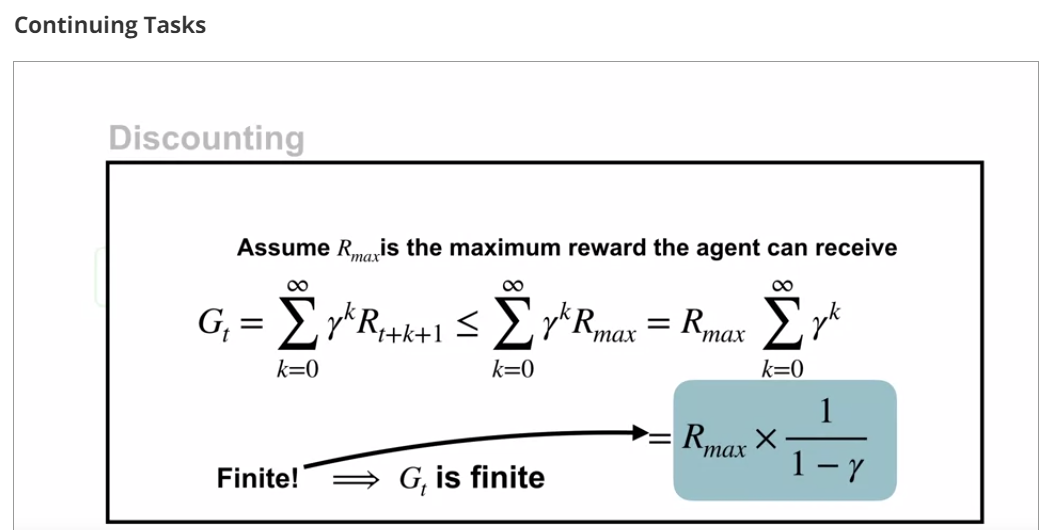
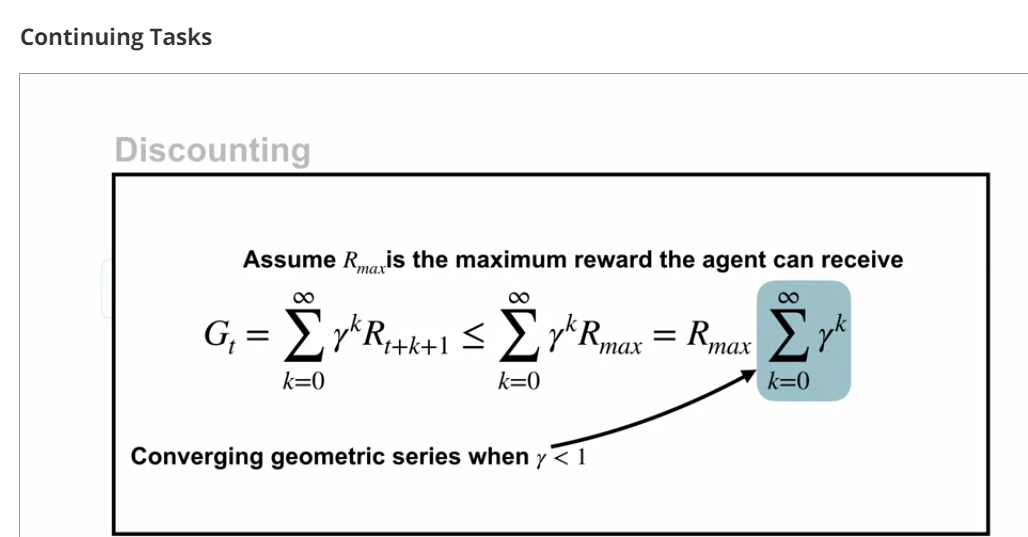
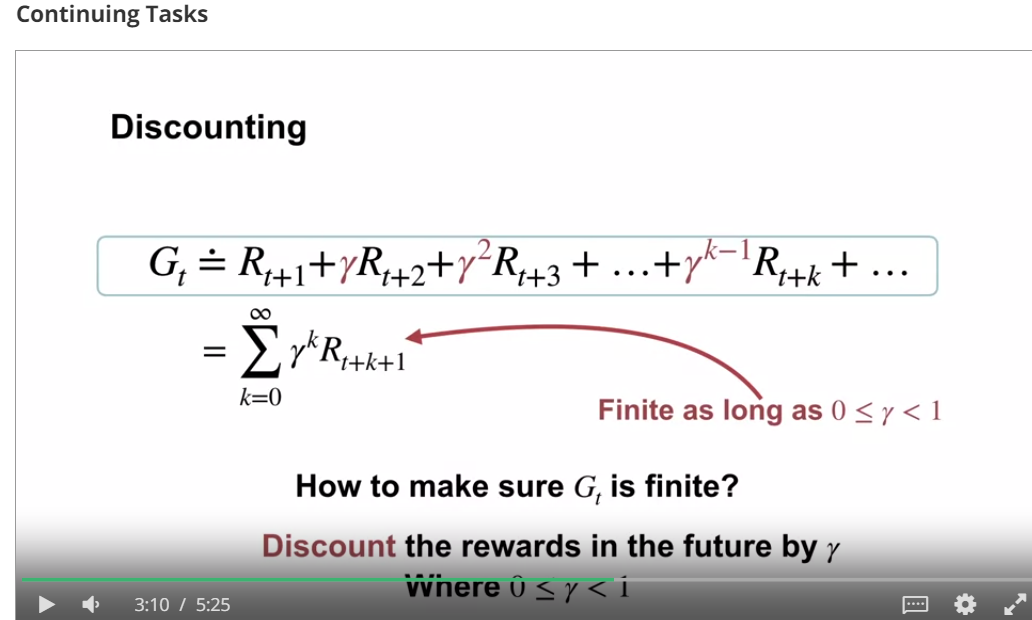
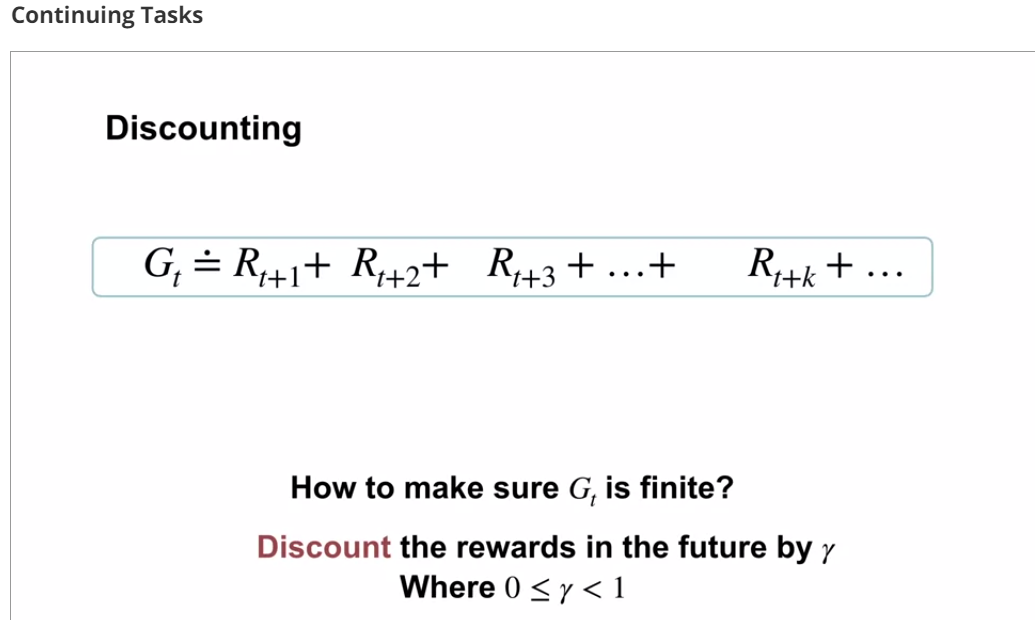
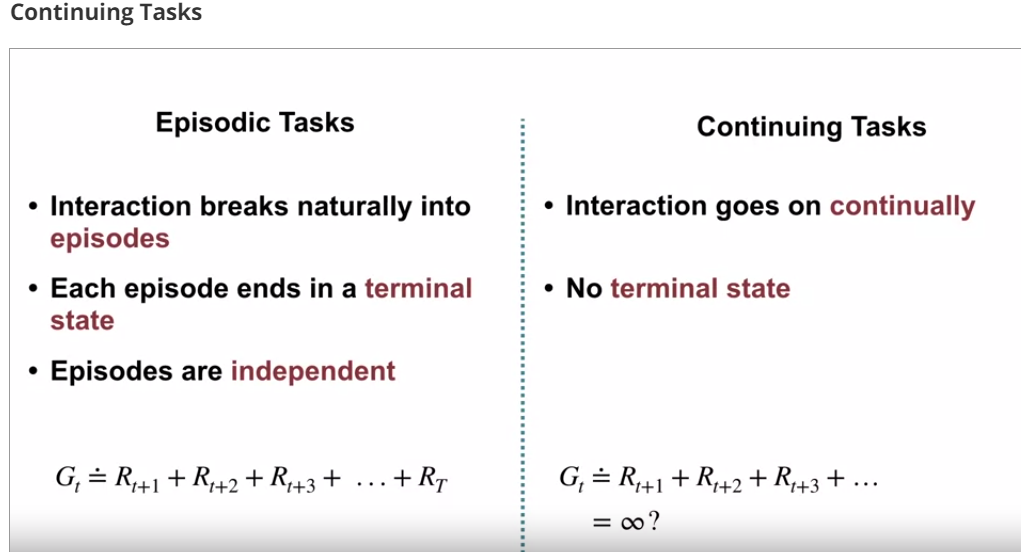
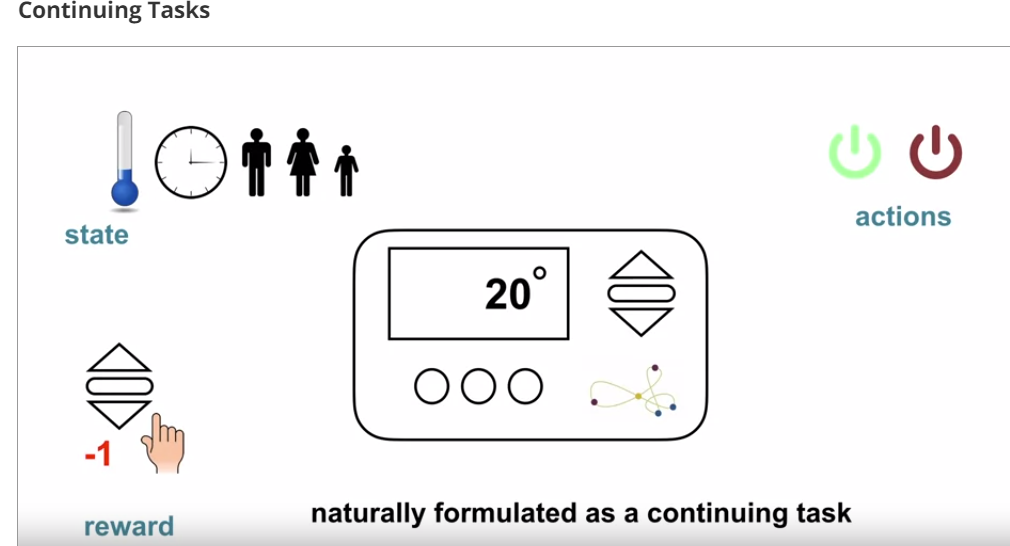
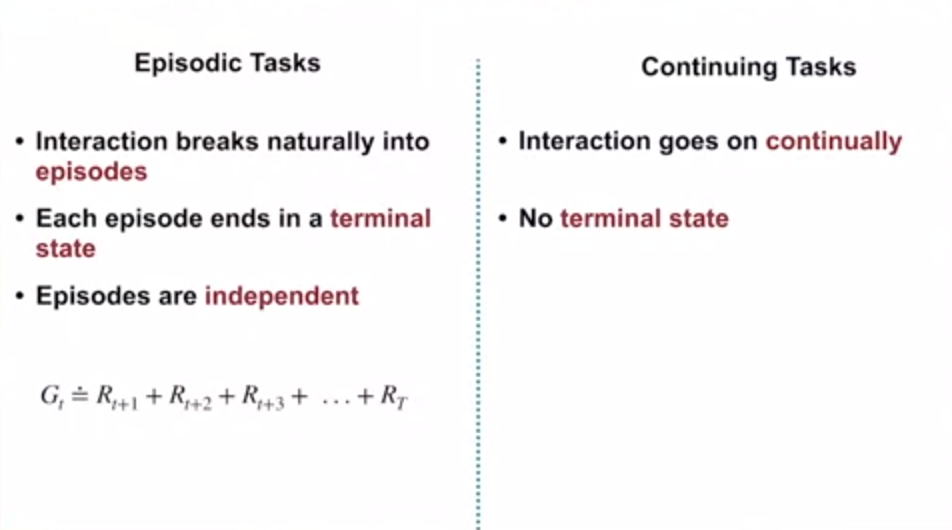
WARNING: In general many different Trajectories from same state possible .(eg:Different rewards for same transitions based on randomness)

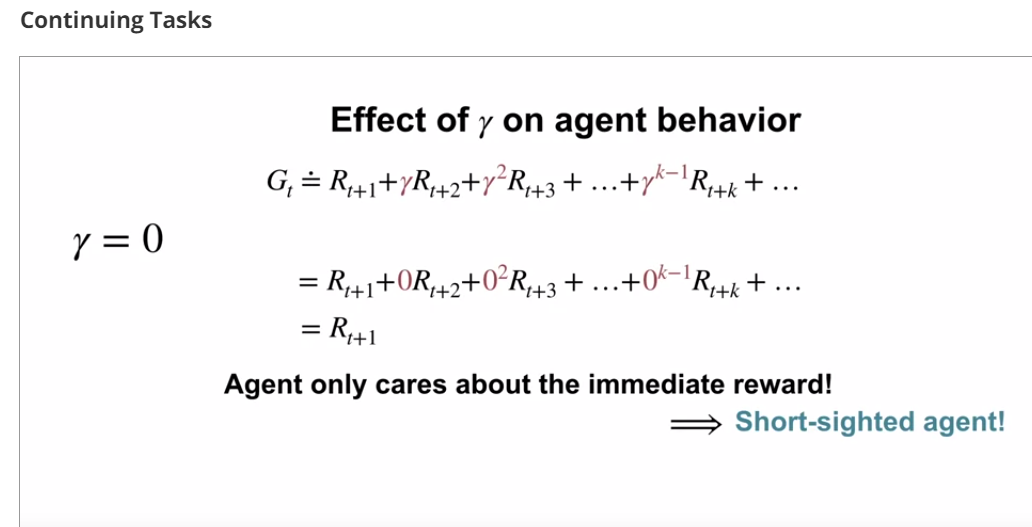
So Let take Finite time T where Agent Environment interacting Ends.

This is Called one EPISODE

RT can be Terminating state







Rt+1 = reward of next time step

