assignment13

May 13, 2021

```python
import numpy as np
import pandas as pd
import gym
```

```python
np.random.seed(42)
from gym import envs
print(envs.registry.all())
```

dict_values([EnvSpec(Copy-v0), EnvSpec(RepeatCopy-v0),
EnvSpec(ReversedAddition-v0), EnvSpec(ReversedAddition3-v0),
EnvSpec(DuplicatedInput-v0), EnvSpec(Reverse-v0), EnvSpec(CartPole-v0),
EnvSpec(CartPole-v1), EnvSpec(MountainCar-v0),
EnvSpec(MountainCarContinuous-v0), EnvSpec(Pendulum-v0), EnvSpec(Acrobot-v1),
EnvSpec(LunarLander-v2), EnvSpec(LunarLanderContinuous-v2),
EnvSpec(BipedalWalker-v3), EnvSpec(BipedalWalkerHardcore-v3),
EnvSpec(CarRacing-v0), EnvSpec(Blackjack-v0), EnvSpec(KellyCoinflip-v0),
EnvSpec(KellyCoinflipGeneralized-v0), EnvSpec(FrozenLake-v0),
EnvSpec(FrozenLake8x8-v0), EnvSpec(CliffWalking-v0), EnvSpec(NChain-v0),
EnvSpec(Roulette-v0), EnvSpec(Taxi-v3), EnvSpec(GuessingGame-v0),
EnvSpec(HotterColder-v0), EnvSpec(Reacher-v2), EnvSpec(Pusher-v2),
EnvSpec(Thrower-v2), EnvSpec(Striker-v2), EnvSpec(InvertedPendulum-v2),
EnvSpec(InvertedDoublePendulum-v2), EnvSpec(HalfCheetah-v2),
EnvSpec(HalfCheetah-v3), EnvSpec(Hopper-v2), EnvSpec(Hopper-v3),
EnvSpec(Swimmer-v2), EnvSpec(Swimmer-v3), EnvSpec(Walker2d-v2),
EnvSpec(Walker2d-v3), EnvSpec(Ant-v2), EnvSpec(Ant-v3), EnvSpec(Humanoid-v2),
EnvSpec(Humanoid-v3), EnvSpec(HumanoidStandup-v2), EnvSpec(FetchSlide-v1),
EnvSpec(FetchPickAndPlace-v1), EnvSpec(FetchReach-v1), EnvSpec(FetchPush-v1),
EnvSpec(HandReach-v0), EnvSpec(HandManipulateBlockRotateZ-v0),
EnvSpec(HandManipulateBlockRotateZTouchSensors-v0),
EnvSpec(HandManipulateBlockRotateZTouchSensors-v1),
EnvSpec(HandManipulateBlockRotateParallel-v0),
EnvSpec(HandManipulateBlockRotateParallelTouchSensors-v0),
EnvSpec(HandManipulateBlockRotateParallelTouchSensors-v1),
EnvSpec(HandManipulateBlockRotateXYZ-v0),
EnvSpec(HandManipulateBlockRotateXYZTouchSensors-v0),
EnvSpec(HandManipulateBlockRotateXYZTouchSensors-v1),
EnvSpec(HandManipulateBlockFull-v0), EnvSpec(HandManipulateBlock-v0),
EnvSpec(HandManipulateBlockTouchSensors-v0),

EnvSpec(HandManipulateBlockTouchSensors-v1),
EnvSpec(HandManipulateEggRotate-v0),
EnvSpec(HandManipulateEggRotateTouchSensors-v0),
EnvSpec(HandManipulateEggRotateTouchSensors-v1),
EnvSpec(HandManipulateEggFull-v0), EnvSpec(HandManipulateEgg-v0),
EnvSpec(HandManipulateEggTouchSensors-v0),
EnvSpec(HandManipulateEggTouchSensors-v1), EnvSpec(HandManipulatePenRotate-v0),
EnvSpec(HandManipulatePenRotateTouchSensors-v0),
EnvSpec(HandManipulatePenRotateTouchSensors-v1),
EnvSpec(HandManipulatePenFull-v0), EnvSpec(HandManipulatePen-v0),
EnvSpec(HandManipulatePenTouchSensors-v0),
EnvSpec(HandManipulatePenTouchSensors-v1), EnvSpec(FetchSlideDense-v1),
EnvSpec(FetchPickAndPlaceDense-v1), EnvSpec(FetchReachDense-v1),
EnvSpec(FetchPushDense-v1), EnvSpec(HandReachDense-v0),
EnvSpec(HandManipulateBlockRotateZDense-v0),
EnvSpec(HandManipulateBlockRotateZTouchSensorsDense-v0),
EnvSpec(HandManipulateBlockRotateZTouchSensorsDense-v1),
EnvSpec(HandManipulateBlockRotateParallelDense-v0),
EnvSpec(HandManipulateBlockRotateParallelTouchSensorsDense-v0),
EnvSpec(HandManipulateBlockRotateParallelTouchSensorsDense-v1),
EnvSpec(HandManipulateBlockRotateXYZDense-v0),
EnvSpec(HandManipulateBlockRotateXYZTouchSensorsDense-v0),
EnvSpec(HandManipulateBlockRotateXYZTouchSensorsDense-v1),
EnvSpec(HandManipulateBlockFullDense-v0), EnvSpec(HandManipulateBlockDense-v0),
EnvSpec(HandManipulateBlockTouchSensorsDense-v0),
EnvSpec(HandManipulateBlockTouchSensorsDense-v1),
EnvSpec(HandManipulateEggRotateDense-v0),
EnvSpec(HandManipulateEggRotateTouchSensorsDense-v0),
EnvSpec(HandManipulateEggRotateTouchSensorsDense-v1),
EnvSpec(HandManipulateEggFullDense-v0), EnvSpec(HandManipulateEggDense-v0),
EnvSpec(HandManipulateEggTouchSensorsDense-v0),
EnvSpec(HandManipulateEggTouchSensorsDense-v1),
EnvSpec(HandManipulatePenRotateDense-v0),
EnvSpec(HandManipulatePenRotateTouchSensorsDense-v0),
EnvSpec(HandManipulatePenRotateTouchSensorsDense-v1),
EnvSpec(HandManipulatePenFullDense-v0), EnvSpec(HandManipulatePenDense-v0),
EnvSpec(HandManipulatePenTouchSensorsDense-v0),
EnvSpec(HandManipulatePenTouchSensorsDense-v1), EnvSpec(Adventure-v0),
EnvSpec(Adventure-v4), EnvSpec(AdventureDeterministic-v0),
EnvSpec(AdventureDeterministic-v4), EnvSpec(AdventureNoFrameskip-v0),
EnvSpec(AdventureNoFrameskip-v4), EnvSpec(Adventure-ram-v0), EnvSpec(Adventure-
ram-v4), EnvSpec(Adventure-ramDeterministic-v0), EnvSpec(Adventure-
ramDeterministic-v4), EnvSpec(Adventure-ramNoFrameskip-v0), EnvSpec(Adventure-
ramNoFrameskip-v4), EnvSpec(AirRaid-v0), EnvSpec(AirRaid-v4),
EnvSpec(AirRaidDeterministic-v0), EnvSpec(AirRaidDeterministic-v4),
EnvSpec(AirRaidNoFrameskip-v0), EnvSpec(AirRaidNoFrameskip-v4), EnvSpec(AirRaid-
ram-v0), EnvSpec(AirRaid-ram-v4), EnvSpec(AirRaid-ramDeterministic-v0),
EnvSpec(AirRaid-ramDeterministic-v4), EnvSpec(AirRaid-ramNoFrameskip-v0),

EnvSpec(AirRaid-ramNoFrameskip-v4), EnvSpec(Alien-v0), EnvSpec(Alien-v4),
EnvSpec(AlienDeterministic-v0), EnvSpec(AlienDeterministic-v4),
EnvSpec(AlienNoFrameskip-v0), EnvSpec(AlienNoFrameskip-v4), EnvSpec(Alien-
ram-v0), EnvSpec(Alien-ram-v4), EnvSpec(Alien-ramDeterministic-v0),
EnvSpec(Alien-ramDeterministic-v4), EnvSpec(Alien-ramNoFrameskip-v0),
EnvSpec(Alien-ramNoFrameskip-v4), EnvSpec(Amidar-v0), EnvSpec(Amidar-v4),
EnvSpec(AmidarDeterministic-v0), EnvSpec(AmidarDeterministic-v4),
EnvSpec(AmidarNoFrameskip-v0), EnvSpec(AmidarNoFrameskip-v4), EnvSpec(Amidar-
ram-v0), EnvSpec(Amidar-ram-v4), EnvSpec(Amidar-ramDeterministic-v0),
EnvSpec(Amidar-ramDeterministic-v4), EnvSpec(Amidar-ramNoFrameskip-v0),
EnvSpec(Amidar-ramNoFrameskip-v4), EnvSpec(Assault-v0), EnvSpec(Assault-v4),
EnvSpec(AssaultDeterministic-v0), EnvSpec(AssaultDeterministic-v4),
EnvSpec(AssaultNoFrameskip-v0), EnvSpec(AssaultNoFrameskip-v4), EnvSpec(Assault-
ram-v0), EnvSpec(Assault-ram-v4), EnvSpec(Assault-ramDeterministic-v0),
EnvSpec(Assault-ramDeterministic-v4), EnvSpec(Assault-ramNoFrameskip-v0),
EnvSpec(Assault-ramNoFrameskip-v4), EnvSpec(Asterix-v0), EnvSpec(Asterix-v4),
EnvSpec(AsterixDeterministic-v0), EnvSpec(AsterixDeterministic-v4),
EnvSpec(AsterixNoFrameskip-v0), EnvSpec(AsterixNoFrameskip-v4), EnvSpec(Asterix-
ram-v0), EnvSpec(Asterix-ram-v4), EnvSpec(Asterix-ramDeterministic-v0),
EnvSpec(Asterix-ramDeterministic-v4), EnvSpec(Asterix-ramNoFrameskip-v0),
EnvSpec(Asterix-ramNoFrameskip-v4), EnvSpec(Asteroids-v0),
EnvSpec(Asteroids-v4), EnvSpec(AsteroidsDeterministic-v0),
EnvSpec(AsteroidsDeterministic-v4), EnvSpec(AsteroidsNoFrameskip-v0),
EnvSpec(AsteroidsNoFrameskip-v4), EnvSpec(Asteroids-ram-v0), EnvSpec(Asteroids-
ram-v4), EnvSpec(Asteroids-ramDeterministic-v0), EnvSpec(Asteroids-
ramDeterministic-v4), EnvSpec(Asteroids-ramNoFrameskip-v0), EnvSpec(Asteroids-
ramNoFrameskip-v4), EnvSpec(Atlantis-v0), EnvSpec(Atlantis-v4),
EnvSpec(AtlantisDeterministic-v0), EnvSpec(AtlantisDeterministic-v4),
EnvSpec(AtlantisNoFrameskip-v0), EnvSpec(AtlantisNoFrameskip-v4),
EnvSpec(Atlantis-ram-v0), EnvSpec(Atlantis-ram-v4), EnvSpec(Atlantis-
ramDeterministic-v0), EnvSpec(Atlantis-ramDeterministic-v4), EnvSpec(Atlantis-
ramNoFrameskip-v0), EnvSpec(Atlantis-ramNoFrameskip-v4), EnvSpec(BankHeist-v0),
EnvSpec(BankHeist-v4), EnvSpec(BankHeistDeterministic-v0),
EnvSpec(BankHeistDeterministic-v4), EnvSpec(BankHeistNoFrameskip-v0),
EnvSpec(BankHeistNoFrameskip-v4), EnvSpec(BankHeist-ram-v0), EnvSpec(BankHeist-
ram-v4), EnvSpec(BankHeist-ramDeterministic-v0), EnvSpec(BankHeist-
ramDeterministic-v4), EnvSpec(BankHeist-ramNoFrameskip-v0), EnvSpec(BankHeist-
ramNoFrameskip-v4), EnvSpec(BattleZone-v0), EnvSpec(BattleZone-v4),
EnvSpec(BattleZoneDeterministic-v0), EnvSpec(BattleZoneDeterministic-v4),
EnvSpec(BattleZoneNoFrameskip-v0), EnvSpec(BattleZoneNoFrameskip-v4),
EnvSpec(BattleZone-ram-v0), EnvSpec(BattleZone-ram-v4), EnvSpec(BattleZone-
ramDeterministic-v0), EnvSpec(BattleZone-ramDeterministic-v4),
EnvSpec(BattleZone-ramNoFrameskip-v0), EnvSpec(BattleZone-ramNoFrameskip-v4),
EnvSpec(BeamRider-v0), EnvSpec(BeamRider-v4),
EnvSpec(BeamRiderDeterministic-v0), EnvSpec(BeamRiderDeterministic-v4),
EnvSpec(BeamRiderNoFrameskip-v0), EnvSpec(BeamRiderNoFrameskip-v4),
EnvSpec(BeamRider-ram-v0), EnvSpec(BeamRider-ram-v4), EnvSpec(BeamRider-
ramDeterministic-v0), EnvSpec(BeamRider-ramDeterministic-v4), EnvSpec(BeamRider-

ramNoFrameskip-v0), EnvSpec(BeamRider-ramNoFrameskip-v4), EnvSpec(Berzerk-v0), EnvSpec(Berzerk-v4), EnvSpec(BerzerkDeterministic-v0), EnvSpec(BerzerkDeterministic-v4), EnvSpec(BerzerkNoFrameskip-v0), EnvSpec(BerzerkNoFrameskip-v4), EnvSpec(Berzerk-ram-v0), EnvSpec(Berzerk-ram-v4), EnvSpec(Berzerk-ramDeterministic-v0), EnvSpec(Berzerk-ramDeterministic-v4), EnvSpec(Berzerk-ramNoFrameskip-v0), EnvSpec(Berzerk-ramNoFrameskip-v4), EnvSpec(Bowling-v0), EnvSpec(Bowling-v4), EnvSpec(BowlingDeterministic-v0), EnvSpec(BowlingDeterministic-v4), EnvSpec(BowlingNoFrameskip-v0), EnvSpec(BowlingNoFrameskip-v4), EnvSpec(Bowling-ram-v0), EnvSpec(Bowling-ram-v4), EnvSpec(Bowling-ramDeterministic-v0), EnvSpec(Bowling-ramDeterministic-v4), EnvSpec(Bowling-ramNoFrameskip-v0), EnvSpec(Bowling-ramNoFrameskip-v4), EnvSpec(Boxing-v0), EnvSpec(Boxing-v4), EnvSpec(BoxingDeterministic-v0), EnvSpec(BoxingDeterministic-v4), EnvSpec(BoxingNoFrameskip-v0), EnvSpec(BoxingNoFrameskip-v4), EnvSpec(Boxing-ram-v0), EnvSpec(Boxing-ram-v4), EnvSpec(Boxing-ramDeterministic-v0), EnvSpec(Boxing-ramDeterministic-v4), EnvSpec(Boxing-ramNoFrameskip-v0), EnvSpec(Boxing-ramNoFrameskip-v4), EnvSpec(Breakout-v0), EnvSpec(Breakout-v4), EnvSpec(BreakoutDeterministic-v0), EnvSpec(BreakoutDeterministic-v4), EnvSpec(BreakoutNoFrameskip-v0), EnvSpec(BreakoutNoFrameskip-v4), EnvSpec(Breakout-ram-v0), EnvSpec(Breakout-ram-v4), EnvSpec(Breakout-ramDeterministic-v0), EnvSpec(Breakout-ramDeterministic-v4), EnvSpec(Breakout-ramNoFrameskip-v0), EnvSpec(Breakout-ramNoFrameskip-v4), EnvSpec(Carnival-v0), EnvSpec(Carnival-v4), EnvSpec(CarnivalDeterministic-v0), EnvSpec(CarnivalDeterministic-v4), EnvSpec(CarnivalNoFrameskip-v0), EnvSpec(CarnivalNoFrameskip-v4), EnvSpec(Carnival-ram-v0), EnvSpec(Carnival-ram-v4), EnvSpec(Carnival-ramDeterministic-v0), EnvSpec(Carnival-ramDeterministic-v4), EnvSpec(Carnival-ramNoFrameskip-v0), EnvSpec(Carnival-ramNoFrameskip-v4), EnvSpec(Centipede-v0), EnvSpec(Centipede-v4), EnvSpec(CentipedeDeterministic-v0), EnvSpec(CentipedeDeterministic-v4), EnvSpec(CentipedeNoFrameskip-v0), EnvSpec(CentipedeNoFrameskip-v4), EnvSpec(Centipede-ram-v0), EnvSpec(Centipede-ram-v4), EnvSpec(Centipede-ramDeterministic-v0), EnvSpec(Centipede-ramDeterministic-v4), EnvSpec(Centipede-ramNoFrameskip-v0), EnvSpec(Centipede-ramNoFrameskip-v4), EnvSpec(ChopperCommand-v0), EnvSpec(ChopperCommand-v4), EnvSpec(ChopperCommandDeterministic-v0), EnvSpec(ChopperCommandDeterministic-v4), EnvSpec(ChopperCommandNoFrameskip-v0), EnvSpec(ChopperCommandNoFrameskip-v4), EnvSpec(ChopperCommand-ram-v0), EnvSpec(ChopperCommand-ram-v4), EnvSpec(ChopperCommand-ramDeterministic-v0), EnvSpec(ChopperCommand-ramDeterministic-v4), EnvSpec(ChopperCommand-ramNoFrameskip-v0), EnvSpec(ChopperCommand-ramNoFrameskip-v4), EnvSpec(CrazyClimber-v0), EnvSpec(CrazyClimber-v4), EnvSpec(CrazyClimberDeterministic-v0), EnvSpec(CrazyClimberDeterministic-v4), EnvSpec(CrazyClimberNoFrameskip-v0), EnvSpec(CrazyClimberNoFrameskip-v4), EnvSpec(CrazyClimber-ram-v0), EnvSpec(CrazyClimber-ram-v4), EnvSpec(CrazyClimber-ramDeterministic-v0), EnvSpec(CrazyClimber-ramDeterministic-v4), EnvSpec(CrazyClimber-ramNoFrameskip-v0), EnvSpec(CrazyClimber-ramNoFrameskip-v4), EnvSpec(Defender-v0), EnvSpec(Defender-v4), EnvSpec(DefenderDeterministic-v0),

EnvSpec(DefenderDeterministic-v4), EnvSpec(DefenderNoFrameskip-v0),
EnvSpec(DefenderNoFrameskip-v4), EnvSpec(Defender-ram-v0), EnvSpec(Defender-
ram-v4), EnvSpec(Defender-ramDeterministic-v0), EnvSpec(Defender-
ramDeterministic-v4), EnvSpec(Defender-ramNoFrameskip-v0), EnvSpec(Defender-
ramNoFrameskip-v4), EnvSpec(DemonAttack-v0), EnvSpec(DemonAttack-v4),
EnvSpec(DemonAttackDeterministic-v0), EnvSpec(DemonAttackDeterministic-v4),
EnvSpec(DemonAttackNoFrameskip-v0), EnvSpec(DemonAttackNoFrameskip-v4),
EnvSpec(DemonAttack-ram-v0), EnvSpec(DemonAttack-ram-v4), EnvSpec(DemonAttack-
ramDeterministic-v0), EnvSpec(DemonAttack-ramDeterministic-v4),
EnvSpec(DemonAttack-ramNoFrameskip-v0), EnvSpec(DemonAttack-ramNoFrameskip-v4),
EnvSpec(DoubleDunk-v0), EnvSpec(DoubleDunk-v4),
EnvSpec(DoubleDunkDeterministic-v0), EnvSpec(DoubleDunkDeterministic-v4),
EnvSpec(DoubleDunkNoFrameskip-v0), EnvSpec(DoubleDunkNoFrameskip-v4),
EnvSpec(DoubleDunk-ram-v0), EnvSpec(DoubleDunk-ram-v4), EnvSpec(DoubleDunk-
ramDeterministic-v0), EnvSpec(DoubleDunk-ramDeterministic-v4),
EnvSpec(DoubleDunk-ramNoFrameskip-v0), EnvSpec(DoubleDunk-ramNoFrameskip-v4),
EnvSpec(ElevatorAction-v0), EnvSpec(ElevatorAction-v4),
EnvSpec(ElevatorActionDeterministic-v0),
EnvSpec(ElevatorActionDeterministic-v4), EnvSpec(ElevatorActionNoFrameskip-v0),
EnvSpec(ElevatorActionNoFrameskip-v4), EnvSpec(ElevatorAction-ram-v0),
EnvSpec(ElevatorAction-ram-v4), EnvSpec(ElevatorAction-ramDeterministic-v0),
EnvSpec(ElevatorAction-ramDeterministic-v4), EnvSpec(ElevatorAction-
ramNoFrameskip-v0), EnvSpec(ElevatorAction-ramNoFrameskip-v4),
EnvSpec(Enduro-v0), EnvSpec(Enduro-v4), EnvSpec(EnduroDeterministic-v0),
EnvSpec(EnduroDeterministic-v4), EnvSpec(EnduroNoFrameskip-v0),
EnvSpec(EnduroNoFrameskip-v4), EnvSpec(Enduro-ram-v0), EnvSpec(Enduro-ram-v4),
EnvSpec(Enduro-ramDeterministic-v0), EnvSpec(Enduro-ramDeterministic-v4),
EnvSpec(Enduro-ramNoFrameskip-v0), EnvSpec(Enduro-ramNoFrameskip-v4),
EnvSpec(FishingDerby-v0), EnvSpec(FishingDerby-v4),
EnvSpec(FishingDerbyDeterministic-v0), EnvSpec(FishingDerbyDeterministic-v4),
EnvSpec(FishingDerbyNoFrameskip-v0), EnvSpec(FishingDerbyNoFrameskip-v4),
EnvSpec(FishingDerby-ram-v0), EnvSpec(FishingDerby-ram-v4),
EnvSpec(FishingDerby-ramDeterministic-v0), EnvSpec(FishingDerby-
ramDeterministic-v4), EnvSpec(FishingDerby-ramNoFrameskip-v0),
EnvSpec(FishingDerby-ramNoFrameskip-v4), EnvSpec(Freeway-v0),
EnvSpec(Freeway-v4), EnvSpec(FreewayDeterministic-v0),
EnvSpec(FreewayDeterministic-v4), EnvSpec(FreewayNoFrameskip-v0),
EnvSpec(FreewayNoFrameskip-v4), EnvSpec(Freeway-ram-v0), EnvSpec(Freeway-
ram-v4), EnvSpec(Freeway-ramDeterministic-v0), EnvSpec(Freeway-
ramDeterministic-v4), EnvSpec(Freeway-ramNoFrameskip-v0), EnvSpec(Freeway-
ramNoFrameskip-v4), EnvSpec(Frostbite-v0), EnvSpec(Frostbite-v4),
EnvSpec(FrostbiteDeterministic-v0), EnvSpec(FrostbiteDeterministic-v4),
EnvSpec(FrostbiteNoFrameskip-v0), EnvSpec(FrostbiteNoFrameskip-v4),
EnvSpec(Frostbite-ram-v0), EnvSpec(Frostbite-ram-v4), EnvSpec(Frostbite-
ramDeterministic-v0), EnvSpec(Frostbite-ramDeterministic-v4), EnvSpec(Frostbite-
ramNoFrameskip-v0), EnvSpec(Frostbite-ramNoFrameskip-v4), EnvSpec(Gopher-v0),
EnvSpec(Gopher-v4), EnvSpec(GopherDeterministic-v0),
EnvSpec(GopherDeterministic-v4), EnvSpec(GopherNoFrameskip-v0),

```
EnvSpec(GopherNoFrameskip-v4), EnvSpec(Gopher-ram-v0), EnvSpec(Gopher-ram-v4),
EnvSpec(Gopher-ramDeterministic-v0), EnvSpec(Gopher-ramDeterministic-v4),
EnvSpec(Gopher-ramNoFrameskip-v0), EnvSpec(Gopher-ramNoFrameskip-v4),
EnvSpec(Gravitar-v0), EnvSpec(Gravitar-v4), EnvSpec(GravitarDeterministic-v0),
EnvSpec(GravitarDeterministic-v4), EnvSpec(GravitarNoFrameskip-v0),
EnvSpec(GravitarNoFrameskip-v4), EnvSpec(Gravitar-ram-v0), EnvSpec(Gravitar-
ram-v4), EnvSpec(Gravitar-ramDeterministic-v0), EnvSpec(Gravitar-
ramDeterministic-v4), EnvSpec(Gravitar-ramNoFrameskip-v0), EnvSpec(Gravitar-
ramNoFrameskip-v4), EnvSpec(Hero-v0), EnvSpec(Hero-v4),
EnvSpec(HeroDeterministic-v0), EnvSpec(HeroDeterministic-v4),
EnvSpec(HeroNoFrameskip-v0), EnvSpec(HeroNoFrameskip-v4), EnvSpec(Hero-ram-v0),
EnvSpec(Hero-ram-v4), EnvSpec(Hero-ramDeterministic-v0), EnvSpec(Hero-
ramDeterministic-v4), EnvSpec(Hero-ramNoFrameskip-v0), EnvSpec(Hero-
ramNoFrameskip-v4), EnvSpec(IceHockey-v0), EnvSpec(IceHockey-v4),
EnvSpec(IceHockeyDeterministic-v0), EnvSpec(IceHockeyDeterministic-v4),
EnvSpec(IceHockeyNoFrameskip-v0), EnvSpec(IceHockeyNoFrameskip-v4),
EnvSpec(IceHockey-ram-v0), EnvSpec(IceHockey-ram-v4), EnvSpec(IceHockey-
ramDeterministic-v0), EnvSpec(IceHockey-ramDeterministic-v4), EnvSpec(IceHockey-
ramNoFrameskip-v0), EnvSpec(IceHockey-ramNoFrameskip-v4), EnvSpec(Jamesbond-v0),
EnvSpec(Jamesbond-v4), EnvSpec(JamesbondDeterministic-v0),
EnvSpec(JamesbondDeterministic-v4), EnvSpec(JamesbondNoFrameskip-v0),
EnvSpec(JamesbondNoFrameskip-v4), EnvSpec(Jamesbond-ram-v0), EnvSpec(Jamesbond-
ram-v4), EnvSpec(Jamesbond-ramDeterministic-v0), EnvSpec(Jamesbond-
ramDeterministic-v4), EnvSpec(Jamesbond-ramNoFrameskip-v0), EnvSpec(Jamesbond-
ramNoFrameskip-v4), EnvSpec(JourneyEscape-v0), EnvSpec(JourneyEscape-v4),
EnvSpec(JourneyEscapeDeterministic-v0), EnvSpec(JourneyEscapeDeterministic-v4),
EnvSpec(JourneyEscapeNoFrameskip-v0), EnvSpec(JourneyEscapeNoFrameskip-v4),
EnvSpec(JourneyEscape-ram-v0), EnvSpec(JourneyEscape-ram-v4),
EnvSpec(JourneyEscape-ramDeterministic-v0), EnvSpec(JourneyEscape-
ramDeterministic-v4), EnvSpec(JourneyEscape-ramNoFrameskip-v0),
EnvSpec(JourneyEscape-ramNoFrameskip-v4), EnvSpec(Kangaroo-v0),
EnvSpec(Kangaroo-v4), EnvSpec(KangarooDeterministic-v0),
EnvSpec(KangarooDeterministic-v4), EnvSpec(KangarooNoFrameskip-v0),
EnvSpec(KangarooNoFrameskip-v4), EnvSpec(Kangaroo-ram-v0), EnvSpec(Kangaroo-
ram-v4), EnvSpec(Kangaroo-ramDeterministic-v0), EnvSpec(Kangaroo-
ramDeterministic-v4), EnvSpec(Kangaroo-ramNoFrameskip-v0), EnvSpec(Kangaroo-
ramNoFrameskip-v4), EnvSpec(Krull-v0), EnvSpec(Krull-v4),
EnvSpec(KrullDeterministic-v0), EnvSpec(KrullDeterministic-v4),
EnvSpec(KrullNoFrameskip-v0), EnvSpec(KrullNoFrameskip-v4), EnvSpec(Krull-
ram-v0), EnvSpec(Krull-ram-v4), EnvSpec(Krull-ramDeterministic-v0),
EnvSpec(Krull-ramDeterministic-v4), EnvSpec(Krull-ramNoFrameskip-v0),
EnvSpec(Krull-ramNoFrameskip-v4), EnvSpec(KungFuMaster-v0),
EnvSpec(KungFuMaster-v4), EnvSpec(KungFuMasterDeterministic-v0),
EnvSpec(KungFuMasterDeterministic-v4), EnvSpec(KungFuMasterNoFrameskip-v0),
EnvSpec(KungFuMasterNoFrameskip-v4), EnvSpec(KungFuMaster-ram-v0),
EnvSpec(KungFuMaster-ram-v4), EnvSpec(KungFuMaster-ramDeterministic-v0),
EnvSpec(KungFuMaster-ramDeterministic-v4), EnvSpec(KungFuMaster-
ramNoFrameskip-v0), EnvSpec(KungFuMaster-ramNoFrameskip-v4),
```

```
EnvSpec(MontezumaRevenge-v0), EnvSpec(MontezumaRevenge-v4),
EnvSpec(MontezumaRevengeDeterministic-v0),
EnvSpec(MontezumaRevengeDeterministic-v4),
EnvSpec(MontezumaRevengeNoFrameskip-v0),
EnvSpec(MontezumaRevengeNoFrameskip-v4), EnvSpec(MontezumaRevenge-ram-v0),
EnvSpec(MontezumaRevenge-ram-v4), EnvSpec(MontezumaRevenge-ramDeterministic-v0),
EnvSpec(MontezumaRevenge-ramDeterministic-v4), EnvSpec(MontezumaRevenge-
ramNoFrameskip-v0), EnvSpec(MontezumaRevenge-ramNoFrameskip-v4),
EnvSpec(MsPacman-v0), EnvSpec(MsPacman-v4), EnvSpec(MsPacmanDeterministic-v0),
EnvSpec(MsPacmanDeterministic-v4), EnvSpec(MsPacmanNoFrameskip-v0),
EnvSpec(MsPacmanNoFrameskip-v4), EnvSpec(MsPacman-ram-v0), EnvSpec(MsPacman-
ram-v4), EnvSpec(MsPacman-ramDeterministic-v0), EnvSpec(MsPacman-
ramDeterministic-v4), EnvSpec(MsPacman-ramNoFrameskip-v0), EnvSpec(MsPacman-
ramNoFrameskip-v4), EnvSpec(NameThisGame-v0), EnvSpec(NameThisGame-v4),
EnvSpec(NameThisGameDeterministic-v0), EnvSpec(NameThisGameDeterministic-v4),
EnvSpec(NameThisGameNoFrameskip-v0), EnvSpec(NameThisGameNoFrameskip-v4),
EnvSpec(NameThisGame-ram-v0), EnvSpec(NameThisGame-ram-v4),
EnvSpec(NameThisGame-ramDeterministic-v0), EnvSpec(NameThisGame-
ramDeterministic-v4), EnvSpec(NameThisGame-ramNoFrameskip-v0),
EnvSpec(NameThisGame-ramNoFrameskip-v4), EnvSpec(Phoenix-v0),
EnvSpec(Phoenix-v4), EnvSpec(PhoenixDeterministic-v0),
EnvSpec(PhoenixDeterministic-v4), EnvSpec(PhoenixNoFrameskip-v0),
EnvSpec(PhoenixNoFrameskip-v4), EnvSpec(Phoenix-ram-v0), EnvSpec(Phoenix-
ram-v4), EnvSpec(Phoenix-ramDeterministic-v0), EnvSpec(Phoenix-
ramDeterministic-v4), EnvSpec(Phoenix-ramNoFrameskip-v0), EnvSpec(Phoenix-
ramNoFrameskip-v4), EnvSpec(Pitfall-v0), EnvSpec(Pitfall-v4),
EnvSpec(PitfallDeterministic-v0), EnvSpec(PitfallDeterministic-v4),
EnvSpec(PitfallNoFrameskip-v0), EnvSpec(PitfallNoFrameskip-v4), EnvSpec(Pitfall-
ram-v0), EnvSpec(Pitfall-ram-v4), EnvSpec(Pitfall-ramDeterministic-v0),
EnvSpec(Pitfall-ramDeterministic-v4), EnvSpec(Pitfall-ramNoFrameskip-v0),
EnvSpec(Pitfall-ramNoFrameskip-v4), EnvSpec(Pong-v0), EnvSpec(Pong-v4),
EnvSpec(PongDeterministic-v0), EnvSpec(PongDeterministic-v4),
EnvSpec(PongNoFrameskip-v0), EnvSpec(PongNoFrameskip-v4), EnvSpec(Pong-ram-v0),
EnvSpec(Pong-ram-v4), EnvSpec(Pong-ramDeterministic-v0), EnvSpec(Pong-
ramDeterministic-v4), EnvSpec(Pong-ramNoFrameskip-v0), EnvSpec(Pong-
ramNoFrameskip-v4), EnvSpec(Pooyan-v0), EnvSpec(Pooyan-v4),
EnvSpec(PooyanDeterministic-v0), EnvSpec(PooyanDeterministic-v4),
EnvSpec(PooyanNoFrameskip-v0), EnvSpec(PooyanNoFrameskip-v4), EnvSpec(Pooyan-
ram-v0), EnvSpec(Pooyan-ram-v4), EnvSpec(Pooyan-ramDeterministic-v0),
EnvSpec(Pooyan-ramDeterministic-v4), EnvSpec(Pooyan-ramNoFrameskip-v0),
EnvSpec(Pooyan-ramNoFrameskip-v4), EnvSpec(PrivateEye-v0),
EnvSpec(PrivateEye-v4), EnvSpec(PrivateEyeDeterministic-v0),
EnvSpec(PrivateEyeDeterministic-v4), EnvSpec(PrivateEyeNoFrameskip-v0),
EnvSpec(PrivateEyeNoFrameskip-v4), EnvSpec(PrivateEye-ram-v0),
EnvSpec(PrivateEye-ram-v4), EnvSpec(PrivateEye-ramDeterministic-v0),
EnvSpec(PrivateEye-ramDeterministic-v4), EnvSpec(PrivateEye-ramNoFrameskip-v0),
EnvSpec(PrivateEye-ramNoFrameskip-v4), EnvSpec(Qbert-v0), EnvSpec(Qbert-v4),
EnvSpec(QbertDeterministic-v0), EnvSpec(QbertDeterministic-v4),
```

EnvSpec(QbertNoFrameskip-v0), EnvSpec(QbertNoFrameskip-v4), EnvSpec(Qbert-ram-v0), EnvSpec(Qbert-ram-v4), EnvSpec(Qbert-ramDeterministic-v0), EnvSpec(Qbert-ramDeterministic-v4), EnvSpec(Qbert-ramNoFrameskip-v0), EnvSpec(Qbert-ramNoFrameskip-v4), EnvSpec(Riverraid-v0), EnvSpec(Riverraid-v4), EnvSpec(RiverraidDeterministic-v0), EnvSpec(RiverraidDeterministic-v4), EnvSpec(RiverraidNoFrameskip-v0), EnvSpec(RiverraidNoFrameskip-v4), EnvSpec(Riverraid-ram-v0), EnvSpec(Riverraid-ram-v4), EnvSpec(Riverraid-ramDeterministic-v0), EnvSpec(Riverraid-ramDeterministic-v4), EnvSpec(Riverraid-ramNoFrameskip-v0), EnvSpec(Riverraid-ramNoFrameskip-v4), EnvSpec(RoadRunner-v0), EnvSpec(RoadRunner-v4), EnvSpec(RoadRunnerDeterministic-v0), EnvSpec(RoadRunnerDeterministic-v4), EnvSpec(RoadRunnerNoFrameskip-v0), EnvSpec(RoadRunnerNoFrameskip-v4), EnvSpec(RoadRunner-ram-v0), EnvSpec(RoadRunner-ram-v4), EnvSpec(RoadRunner-ramDeterministic-v0), EnvSpec(RoadRunner-ramDeterministic-v4), EnvSpec(RoadRunner-ramNoFrameskip-v0), EnvSpec(RoadRunner-ramNoFrameskip-v4), EnvSpec(Robotank-v0), EnvSpec(Robotank-v4), EnvSpec(RobotankDeterministic-v0), EnvSpec(RobotankDeterministic-v4), EnvSpec(RobotankNoFrameskip-v0), EnvSpec(RobotankNoFrameskip-v4), EnvSpec(Robotank-ram-v0), EnvSpec(Robotank-ram-v4), EnvSpec(Robotank-ramDeterministic-v0), EnvSpec(Robotank-ramDeterministic-v4), EnvSpec(Robotank-ramNoFrameskip-v0), EnvSpec(Robotank-ramNoFrameskip-v4), EnvSpec(Seaquest-v0), EnvSpec(Seaquest-v4), EnvSpec(SeaquestDeterministic-v0), EnvSpec(SeaquestDeterministic-v4), EnvSpec(SeaquestNoFrameskip-v0), EnvSpec(SeaquestNoFrameskip-v4), EnvSpec(Seaquest-ram-v0), EnvSpec(Seaquest-ram-v4), EnvSpec(Seaquest-ramDeterministic-v0), EnvSpec(Seaquest-ramDeterministic-v4), EnvSpec(Seaquest-ramNoFrameskip-v0), EnvSpec(Seaquest-ramNoFrameskip-v4), EnvSpec(Skiing-v0), EnvSpec(Skiing-v4), EnvSpec(SkiingDeterministic-v0), EnvSpec(SkiingDeterministic-v4), EnvSpec(SkiingNoFrameskip-v0), EnvSpec(SkiingNoFrameskip-v4), EnvSpec(Skiing-ram-v0), EnvSpec(Skiing-ram-v4), EnvSpec(Skiing-ramDeterministic-v0), EnvSpec(Skiing-ramDeterministic-v4), EnvSpec(Skiing-ramNoFrameskip-v0), EnvSpec(Skiing-ramNoFrameskip-v4), EnvSpec(Solaris-v0), EnvSpec(Solaris-v4), EnvSpec(SolarisDeterministic-v0), EnvSpec(SolarisDeterministic-v4), EnvSpec(SolarisNoFrameskip-v0), EnvSpec(SolarisNoFrameskip-v4), EnvSpec(Solaris-ram-v0), EnvSpec(Solaris-ram-v4), EnvSpec(Solaris-ramDeterministic-v0), EnvSpec(Solaris-ramDeterministic-v4), EnvSpec(Solaris-ramNoFrameskip-v0), EnvSpec(Solaris-ramNoFrameskip-v4), EnvSpec(SpaceInvaders-v0), EnvSpec(SpaceInvaders-v4), EnvSpec(SpaceInvadersDeterministic-v0), EnvSpec(SpaceInvadersDeterministic-v4), EnvSpec(SpaceInvadersNoFrameskip-v0), EnvSpec(SpaceInvadersNoFrameskip-v4), EnvSpec(SpaceInvaders-ram-v0), EnvSpec(SpaceInvaders-ram-v4), EnvSpec(SpaceInvaders-ramDeterministic-v0), EnvSpec(SpaceInvaders-ramDeterministic-v4), EnvSpec(SpaceInvaders-ramNoFrameskip-v0), EnvSpec(SpaceInvaders-ramNoFrameskip-v4), EnvSpec(StarGunner-v0), EnvSpec(StarGunner-v4), EnvSpec(StarGunnerDeterministic-v0), EnvSpec(StarGunnerDeterministic-v4), EnvSpec(StarGunnerNoFrameskip-v0), EnvSpec(StarGunnerNoFrameskip-v4), EnvSpec(StarGunner-ram-v0), EnvSpec(StarGunner-ram-v4), EnvSpec(StarGunner-ramDeterministic-v0), EnvSpec(StarGunner-ramDeterministic-v4), EnvSpec(StarGunner-ramNoFrameskip-v0),

EnvSpec(StarGunner-ramNoFrameskip-v4), EnvSpec(Tennis-v0), EnvSpec(Tennis-v4),
EnvSpec(TennisDeterministic-v0), EnvSpec(TennisDeterministic-v4),
EnvSpec(TennisNoFrameskip-v0), EnvSpec(TennisNoFrameskip-v4), EnvSpec(Tennis-
ram-v0), EnvSpec(Tennis-ram-v4), EnvSpec(Tennis-ramDeterministic-v0),
EnvSpec(Tennis-ramDeterministic-v4), EnvSpec(Tennis-ramNoFrameskip-v0),
EnvSpec(Tennis-ramNoFrameskip-v4), EnvSpec(TimePilot-v0), EnvSpec(TimePilot-v4),
EnvSpec(TimePilotDeterministic-v0), EnvSpec(TimePilotDeterministic-v4),
EnvSpec(TimePilotNoFrameskip-v0), EnvSpec(TimePilotNoFrameskip-v4),
EnvSpec(TimePilot-ram-v0), EnvSpec(TimePilot-ram-v4), EnvSpec(TimePilot-
ramDeterministic-v0), EnvSpec(TimePilot-ramDeterministic-v4), EnvSpec(TimePilot-
ramNoFrameskip-v0), EnvSpec(TimePilot-ramNoFrameskip-v4), EnvSpec(Tutankham-v0),
EnvSpec(Tutankham-v4), EnvSpec(TutankhamDeterministic-v0),
EnvSpec(TutankhamDeterministic-v4), EnvSpec(TutankhamNoFrameskip-v0),
EnvSpec(TutankhamNoFrameskip-v4), EnvSpec(Tutankham-ram-v0), EnvSpec(Tutankham-
ram-v4), EnvSpec(Tutankham-ramDeterministic-v0), EnvSpec(Tutankham-
ramDeterministic-v4), EnvSpec(Tutankham-ramNoFrameskip-v0), EnvSpec(Tutankham-
ramNoFrameskip-v4), EnvSpec(UpNDown-v0), EnvSpec(UpNDown-v4),
EnvSpec(UpNDownDeterministic-v0), EnvSpec(UpNDownDeterministic-v4),
EnvSpec(UpNDownNoFrameskip-v0), EnvSpec(UpNDownNoFrameskip-v4), EnvSpec(UpNDown-
ram-v0), EnvSpec(UpNDown-ram-v4), EnvSpec(UpNDown-ramDeterministic-v0),
EnvSpec(UpNDown-ramDeterministic-v4), EnvSpec(UpNDown-ramNoFrameskip-v0),
EnvSpec(UpNDown-ramNoFrameskip-v4), EnvSpec(Venture-v0), EnvSpec(Venture-v4),
EnvSpec(VentureDeterministic-v0), EnvSpec(VentureDeterministic-v4),
EnvSpec(VentureNoFrameskip-v0), EnvSpec(VentureNoFrameskip-v4), EnvSpec(Venture-
ram-v0), EnvSpec(Venture-ram-v4), EnvSpec(Venture-ramDeterministic-v0),
EnvSpec(Venture-ramDeterministic-v4), EnvSpec(Venture-ramNoFrameskip-v0),
EnvSpec(Venture-ramNoFrameskip-v4), EnvSpec(VideoPinball-v0),
EnvSpec(VideoPinball-v4), EnvSpec(VideoPinballDeterministic-v0),
EnvSpec(VideoPinballDeterministic-v4), EnvSpec(VideoPinballNoFrameskip-v0),
EnvSpec(VideoPinballNoFrameskip-v4), EnvSpec(VideoPinball-ram-v0),
EnvSpec(VideoPinball-ram-v4), EnvSpec(VideoPinball-ramDeterministic-v0),
EnvSpec(VideoPinball-ramDeterministic-v4), EnvSpec(VideoPinball-
ramNoFrameskip-v0), EnvSpec(VideoPinball-ramNoFrameskip-v4),
EnvSpec(WizardOfWor-v0), EnvSpec(WizardOfWor-v4),
EnvSpec(WizardOfWorDeterministic-v0), EnvSpec(WizardOfWorDeterministic-v4),
EnvSpec(WizardOfWorNoFrameskip-v0), EnvSpec(WizardOfWorNoFrameskip-v4),
EnvSpec(WizardOfWor-ram-v0), EnvSpec(WizardOfWor-ram-v4), EnvSpec(WizardOfWor-
ramDeterministic-v0), EnvSpec(WizardOfWor-ramDeterministic-v4),
EnvSpec(WizardOfWor-ramNoFrameskip-v0), EnvSpec(WizardOfWor-ramNoFrameskip-v4),
EnvSpec(YarsRevenge-v0), EnvSpec(YarsRevenge-v4),
EnvSpec(YarsRevengeDeterministic-v0), EnvSpec(YarsRevengeDeterministic-v4),
EnvSpec(YarsRevengeNoFrameskip-v0), EnvSpec(YarsRevengeNoFrameskip-v4),
EnvSpec(YarsRevenge-ram-v0), EnvSpec(YarsRevenge-ram-v4), EnvSpec(YarsRevenge-
ramDeterministic-v0), EnvSpec(YarsRevenge-ramDeterministic-v4),
EnvSpec(YarsRevenge-ramNoFrameskip-v0), EnvSpec(YarsRevenge-ramNoFrameskip-v4),
EnvSpec(Zaxxon-v0), EnvSpec(Zaxxon-v4), EnvSpec(ZaxxonDeterministic-v0),
EnvSpec(ZaxxonDeterministic-v4), EnvSpec(ZaxxonNoFrameskip-v0),
EnvSpec(ZaxxonNoFrameskip-v4), EnvSpec(Zaxxon-ram-v0), EnvSpec(Zaxxon-ram-v4),

```
EnvSpec(Zaxxon-ramDeterministic-v0), EnvSpec(Zaxxon-ramDeterministic-v4),
EnvSpec(Zaxxon-ramNoFrameskip-v0), EnvSpec(Zaxxon-ramNoFrameskip-v4),
EnvSpec(CubeCrash-v0), EnvSpec(CubeCrashSparse-v0),
EnvSpec(CubeCrashScreenBecomesBlack-v0), EnvSpec(MemorizeDigits-v0)])
```

```python
env = gym.make('FrozenLake-v0')

env.seed(42)
env.reset()
env.render()

print("Action Space {}".format(env.action_space))
print("State Space {}".format(env.observation_space))
```

```
SFFF
FHFH
FFFH
HFFG
Action Space Discrete(4)
State Space Discrete(16)
```

```python
# take some random steps
reward_total = 0
obs = env.reset()
env.render()
for i in range(5):
    action = env.action_space.sample() #take step using random action from
 →possible actions (action_space)
    obs, reward, done, info = env.step(action)
    reward_total += reward
    env.render()
#Print the reward of these random action
print("Total reward: %r" % reward_total)
```

```
SFFF
FHFH
FFFH
HFFG
  (Right)
SFFF
FHFH
FFFH
HFFG
  (Up)
SFFF
FHFH
```

```
FFFH
HFFG
  (Right)
SFFF
FHFH
FFFH
HFFG
  (Right)
SFFF
FHFH
FFFH
HFFG
  (Right)
SFFF
FHFH
FFFH
HFFG
Total reward: 0.0
```

1 Q-learning algorithm

```python
qtable = np.zeros([env.observation_space.n, env.action_space.n]) #You could
↪also make this dynamic if you don't know all games states upfront
discount = 0.9 # discount factor
learningrate = 0.9 # learning rate
for episode in range(1, 10001):
    done = False
    reward_total = 0
    state = env.reset()
    while done != True:
        action = np.argmax(qtable[state]) #choosing the action with the highest
↪Q value
        state_new, reward, done, info = env.step(action) #take the action
        qtable[state,action] += learningrate * (reward + discount * np.
↪max(qtable[state_new,:]) - qtable[state,action]) #Update Q-marix using
↪Bellman equation
            #Q[obs,action] = rew + gamma * np.max(Q[obs2]) # same equation but
↪with learning rate = 1 returns the basic Bellman equation
        reward_total = reward_total + reward
        state = state_new
    if episode % 50 == 0:
        print('Episode {} Total Reward: {}'.format(episode,rew_tot))
```

```
Episode 50 Total Reward: 0.0
Episode 100 Total Reward: 0.0
Episode 150 Total Reward: 0.0
```

```
Episode 200 Total Reward: 0.0
Episode 250 Total Reward: 0.0
Episode 300 Total Reward: 0.0
Episode 350 Total Reward: 0.0
Episode 400 Total Reward: 0.0
Episode 450 Total Reward: 0.0
Episode 500 Total Reward: 0.0
Episode 550 Total Reward: 0.0
Episode 600 Total Reward: 0.0
Episode 650 Total Reward: 0.0
Episode 700 Total Reward: 0.0
Episode 750 Total Reward: 0.0
Episode 800 Total Reward: 0.0
Episode 850 Total Reward: 0.0
Episode 900 Total Reward: 0.0
Episode 950 Total Reward: 0.0
Episode 1000 Total Reward: 0.0
Episode 1050 Total Reward: 0.0
Episode 1100 Total Reward: 0.0
Episode 1150 Total Reward: 0.0
Episode 1200 Total Reward: 0.0
Episode 1250 Total Reward: 0.0
Episode 1300 Total Reward: 0.0
Episode 1350 Total Reward: 0.0
Episode 1400 Total Reward: 0.0
Episode 1450 Total Reward: 0.0
Episode 1500 Total Reward: 0.0
Episode 1550 Total Reward: 0.0
Episode 1600 Total Reward: 0.0
Episode 1650 Total Reward: 0.0
Episode 1700 Total Reward: 0.0
Episode 1750 Total Reward: 0.0
Episode 1800 Total Reward: 0.0
Episode 1850 Total Reward: 0.0
Episode 1900 Total Reward: 0.0
Episode 1950 Total Reward: 0.0
Episode 2000 Total Reward: 0.0
Episode 2050 Total Reward: 0.0
Episode 2100 Total Reward: 0.0
Episode 2150 Total Reward: 0.0
Episode 2200 Total Reward: 0.0
Episode 2250 Total Reward: 0.0
Episode 2300 Total Reward: 0.0
Episode 2350 Total Reward: 0.0
Episode 2400 Total Reward: 0.0
Episode 2450 Total Reward: 0.0
Episode 2500 Total Reward: 0.0
Episode 2550 Total Reward: 0.0
```

```
Episode 2600 Total Reward: 0.0
Episode 2650 Total Reward: 0.0
Episode 2700 Total Reward: 0.0
Episode 2750 Total Reward: 0.0
Episode 2800 Total Reward: 0.0
Episode 2850 Total Reward: 0.0
Episode 2900 Total Reward: 0.0
Episode 2950 Total Reward: 0.0
Episode 3000 Total Reward: 0.0
Episode 3050 Total Reward: 0.0
Episode 3100 Total Reward: 0.0
Episode 3150 Total Reward: 0.0
Episode 3200 Total Reward: 0.0
Episode 3250 Total Reward: 0.0
Episode 3300 Total Reward: 0.0
Episode 3350 Total Reward: 0.0
Episode 3400 Total Reward: 0.0
Episode 3450 Total Reward: 0.0
Episode 3500 Total Reward: 0.0
Episode 3550 Total Reward: 0.0
Episode 3600 Total Reward: 0.0
Episode 3650 Total Reward: 0.0
Episode 3700 Total Reward: 0.0
Episode 3750 Total Reward: 0.0
Episode 3800 Total Reward: 0.0
Episode 3850 Total Reward: 0.0
Episode 3900 Total Reward: 0.0
Episode 3950 Total Reward: 0.0
Episode 4000 Total Reward: 0.0
Episode 4050 Total Reward: 0.0
Episode 4100 Total Reward: 0.0
Episode 4150 Total Reward: 0.0
Episode 4200 Total Reward: 0.0
Episode 4250 Total Reward: 0.0
Episode 4300 Total Reward: 0.0
Episode 4350 Total Reward: 0.0
Episode 4400 Total Reward: 0.0
Episode 4450 Total Reward: 0.0
Episode 4500 Total Reward: 0.0
Episode 4550 Total Reward: 0.0
Episode 4600 Total Reward: 0.0
Episode 4650 Total Reward: 0.0
Episode 4700 Total Reward: 0.0
Episode 4750 Total Reward: 0.0
Episode 4800 Total Reward: 0.0
Episode 4850 Total Reward: 0.0
Episode 4900 Total Reward: 0.0
Episode 4950 Total Reward: 0.0
```

```
Episode 5000 Total Reward: 0.0
Episode 5050 Total Reward: 0.0
Episode 5100 Total Reward: 0.0
Episode 5150 Total Reward: 0.0
Episode 5200 Total Reward: 0.0
Episode 5250 Total Reward: 0.0
Episode 5300 Total Reward: 0.0
Episode 5350 Total Reward: 0.0
Episode 5400 Total Reward: 0.0
Episode 5450 Total Reward: 0.0
Episode 5500 Total Reward: 0.0
Episode 5550 Total Reward: 0.0
Episode 5600 Total Reward: 0.0
Episode 5650 Total Reward: 0.0
Episode 5700 Total Reward: 0.0
Episode 5750 Total Reward: 0.0
Episode 5800 Total Reward: 0.0
Episode 5850 Total Reward: 0.0
Episode 5900 Total Reward: 0.0
Episode 5950 Total Reward: 0.0
Episode 6000 Total Reward: 0.0
Episode 6050 Total Reward: 0.0
Episode 6100 Total Reward: 0.0
Episode 6150 Total Reward: 0.0
Episode 6200 Total Reward: 0.0
Episode 6250 Total Reward: 0.0
Episode 6300 Total Reward: 0.0
Episode 6350 Total Reward: 0.0
Episode 6400 Total Reward: 0.0
Episode 6450 Total Reward: 0.0
Episode 6500 Total Reward: 0.0
Episode 6550 Total Reward: 0.0
Episode 6600 Total Reward: 0.0
Episode 6650 Total Reward: 0.0
Episode 6700 Total Reward: 0.0
Episode 6750 Total Reward: 0.0
Episode 6800 Total Reward: 0.0
Episode 6850 Total Reward: 0.0
Episode 6900 Total Reward: 0.0
Episode 6950 Total Reward: 0.0
Episode 7000 Total Reward: 0.0
Episode 7050 Total Reward: 0.0
Episode 7100 Total Reward: 0.0
Episode 7150 Total Reward: 0.0
Episode 7200 Total Reward: 0.0
Episode 7250 Total Reward: 0.0
Episode 7300 Total Reward: 0.0
Episode 7350 Total Reward: 0.0
```

```
Episode 7400 Total Reward: 0.0
Episode 7450 Total Reward: 0.0
Episode 7500 Total Reward: 0.0
Episode 7550 Total Reward: 0.0
Episode 7600 Total Reward: 0.0
Episode 7650 Total Reward: 0.0
Episode 7700 Total Reward: 0.0
Episode 7750 Total Reward: 0.0
Episode 7800 Total Reward: 0.0
Episode 7850 Total Reward: 0.0
Episode 7900 Total Reward: 0.0
Episode 7950 Total Reward: 0.0
Episode 8000 Total Reward: 0.0
Episode 8050 Total Reward: 0.0
Episode 8100 Total Reward: 0.0
Episode 8150 Total Reward: 0.0
Episode 8200 Total Reward: 0.0
Episode 8250 Total Reward: 0.0
Episode 8300 Total Reward: 0.0
Episode 8350 Total Reward: 0.0
Episode 8400 Total Reward: 0.0
Episode 8450 Total Reward: 0.0
Episode 8500 Total Reward: 0.0
Episode 8550 Total Reward: 0.0
Episode 8600 Total Reward: 0.0
Episode 8650 Total Reward: 0.0
Episode 8700 Total Reward: 0.0
Episode 8750 Total Reward: 0.0
Episode 8800 Total Reward: 0.0
Episode 8850 Total Reward: 0.0
Episode 8900 Total Reward: 0.0
Episode 8950 Total Reward: 0.0
Episode 9000 Total Reward: 0.0
Episode 9050 Total Reward: 0.0
Episode 9100 Total Reward: 0.0
Episode 9150 Total Reward: 0.0
Episode 9200 Total Reward: 0.0
Episode 9250 Total Reward: 0.0
Episode 9300 Total Reward: 0.0
Episode 9350 Total Reward: 0.0
Episode 9400 Total Reward: 0.0
Episode 9450 Total Reward: 0.0
Episode 9500 Total Reward: 0.0
Episode 9550 Total Reward: 0.0
Episode 9600 Total Reward: 0.0
Episode 9650 Total Reward: 0.0
Episode 9700 Total Reward: 0.0
Episode 9750 Total Reward: 0.0
```

```
Episode 9800 Total Reward: 0.0
Episode 9850 Total Reward: 0.0
Episode 9900 Total Reward: 0.0
Episode 9950 Total Reward: 0.0
Episode 10000 Total Reward: 0.0
```

```python
[ ]: # Let's see how the algorithm solves the taxi game by following the policy to␣
     ↪take actions delivering max value

     reward_total=0
     state= env.reset()
     env.render()
     done=False
     while done != True:
         action = np.argmax(qtable[state])
         state, reward, done, info = env.step(action) #take step using selected␣
     ↪action
         reward_total = reward_total + reward
         env.render()
     #Print the reward of these actions
     print("Total reward is %r" % reward_total)
```

```
SFFF
FHFH
FFFH
HFFG
  (Left)
SFFF
FHFH
FFFH
HFFG
  (Left)
SFFF
FHFH
FFFH
HFFG
  (Left)
SFFF
FHFH
FFFH
HFFG
  (Left)
SFFF
FHFH
FFFH
HFFG
Total reward is 0.0
```

16

```
import time
from IPython.display import clear_output

env.s = 166
reward_total=0
env.render()
done=False
while done != True:
    action = np.argmax(qtable[state])
    state, reward, done, info = env.step(action)
    reward_total = reward_total + reward
    time.sleep(1)
    clear_output(wait=True)
    env.render()
```

```
  (Left)
SFFF
FHFH
FFFH
HFFG
```

#Adding stochasticity to our Q-learning algorithm

```
qtable = np.zeros([env.observation_space.n, env.action_space.n]) #we could also
 ↪make this dynamic if we
#don't know all games states upfront
discount = 0.9 # discount factor
learningrate = 0.9 # learning rate
epsilon = 0.2 # threshold of stochasticity
for episode in range(1,10001):
    done = False
    reward_total = 0
    state = env.reset()
    while done != True:
        explore_eploit = np.random.uniform(0, 1)
        if explore_eploit < epsilon:
            action = env.action_space.sample() # explore action space
        else:
            action = np.argmax(qtable[state]) # exploit learned values

        state_new, reward, done, info = env.step(action) #take the action
        qtable[state,action] += learningrate * (reward + discount * np.
 ↪max(qtable[state_new,:]) - qtable[state,action]) #Update Q-marix using
 ↪Bellman equation
        reward_total = reward_total + reward
        state = state_new
    if episode % 50 == 0:
        print('Episode {} Total Reward: {}'.format(episode,reward_total))
```

```
Episode 50 Total Reward: 0.0
Episode 100 Total Reward: 0.0
Episode 150 Total Reward: 0.0
Episode 200 Total Reward: 0.0
Episode 250 Total Reward: 0.0
Episode 300 Total Reward: 0.0
Episode 350 Total Reward: 0.0
Episode 400 Total Reward: 0.0
Episode 450 Total Reward: 0.0
Episode 500 Total Reward: 0.0
Episode 550 Total Reward: 0.0
Episode 600 Total Reward: 0.0
Episode 650 Total Reward: 1.0
Episode 700 Total Reward: 1.0
Episode 750 Total Reward: 0.0
Episode 800 Total Reward: 0.0
Episode 850 Total Reward: 1.0
Episode 900 Total Reward: 0.0
Episode 950 Total Reward: 0.0
Episode 1000 Total Reward: 0.0
Episode 1050 Total Reward: 0.0
Episode 1100 Total Reward: 1.0
Episode 1150 Total Reward: 0.0
Episode 1200 Total Reward: 0.0
Episode 1250 Total Reward: 0.0
Episode 1300 Total Reward: 0.0
Episode 1350 Total Reward: 0.0
Episode 1400 Total Reward: 1.0
Episode 1450 Total Reward: 0.0
Episode 1500 Total Reward: 0.0
Episode 1550 Total Reward: 0.0
Episode 1600 Total Reward: 0.0
Episode 1650 Total Reward: 0.0
Episode 1700 Total Reward: 0.0
Episode 1750 Total Reward: 0.0
Episode 1800 Total Reward: 0.0
Episode 1850 Total Reward: 0.0
Episode 1900 Total Reward: 0.0
Episode 1950 Total Reward: 0.0
Episode 2000 Total Reward: 0.0
Episode 2050 Total Reward: 0.0
Episode 2100 Total Reward: 0.0
Episode 2150 Total Reward: 0.0
Episode 2200 Total Reward: 0.0
Episode 2250 Total Reward: 0.0
Episode 2300 Total Reward: 1.0
Episode 2350 Total Reward: 0.0
Episode 2400 Total Reward: 0.0
```

```
Episode 2450 Total Reward: 1.0
Episode 2500 Total Reward: 0.0
Episode 2550 Total Reward: 1.0
Episode 2600 Total Reward: 0.0
Episode 2650 Total Reward: 0.0
Episode 2700 Total Reward: 0.0
Episode 2750 Total Reward: 0.0
Episode 2800 Total Reward: 0.0
Episode 2850 Total Reward: 0.0
Episode 2900 Total Reward: 0.0
Episode 2950 Total Reward: 0.0
Episode 3000 Total Reward: 0.0
Episode 3050 Total Reward: 0.0
Episode 3100 Total Reward: 0.0
Episode 3150 Total Reward: 1.0
Episode 3200 Total Reward: 0.0
Episode 3250 Total Reward: 0.0
Episode 3300 Total Reward: 0.0
Episode 3350 Total Reward: 0.0
Episode 3400 Total Reward: 1.0
Episode 3450 Total Reward: 0.0
Episode 3500 Total Reward: 0.0
Episode 3550 Total Reward: 0.0
Episode 3600 Total Reward: 0.0
Episode 3650 Total Reward: 1.0
Episode 3700 Total Reward: 0.0
Episode 3750 Total Reward: 0.0
Episode 3800 Total Reward: 0.0
Episode 3850 Total Reward: 0.0
Episode 3900 Total Reward: 0.0
Episode 3950 Total Reward: 0.0
Episode 4000 Total Reward: 0.0
Episode 4050 Total Reward: 1.0
Episode 4100 Total Reward: 0.0
Episode 4150 Total Reward: 0.0
Episode 4200 Total Reward: 0.0
Episode 4250 Total Reward: 0.0
Episode 4300 Total Reward: 0.0
Episode 4350 Total Reward: 1.0
Episode 4400 Total Reward: 0.0
Episode 4450 Total Reward: 0.0
Episode 4500 Total Reward: 1.0
Episode 4550 Total Reward: 0.0
Episode 4600 Total Reward: 0.0
Episode 4650 Total Reward: 0.0
Episode 4700 Total Reward: 0.0
Episode 4750 Total Reward: 0.0
Episode 4800 Total Reward: 0.0
```

```
Episode 4850 Total Reward: 0.0
Episode 4900 Total Reward: 0.0
Episode 4950 Total Reward: 0.0
Episode 5000 Total Reward: 1.0
Episode 5050 Total Reward: 0.0
Episode 5100 Total Reward: 0.0
Episode 5150 Total Reward: 0.0
Episode 5200 Total Reward: 0.0
Episode 5250 Total Reward: 0.0
Episode 5300 Total Reward: 0.0
Episode 5350 Total Reward: 0.0
Episode 5400 Total Reward: 0.0
Episode 5450 Total Reward: 0.0
Episode 5500 Total Reward: 0.0
Episode 5550 Total Reward: 0.0
Episode 5600 Total Reward: 0.0
Episode 5650 Total Reward: 0.0
Episode 5700 Total Reward: 0.0
Episode 5750 Total Reward: 1.0
Episode 5800 Total Reward: 0.0
Episode 5850 Total Reward: 0.0
Episode 5900 Total Reward: 1.0
Episode 5950 Total Reward: 0.0
Episode 6000 Total Reward: 0.0
Episode 6050 Total Reward: 1.0
Episode 6100 Total Reward: 0.0
Episode 6150 Total Reward: 0.0
Episode 6200 Total Reward: 0.0
Episode 6250 Total Reward: 0.0
Episode 6300 Total Reward: 0.0
Episode 6350 Total Reward: 0.0
Episode 6400 Total Reward: 1.0
Episode 6450 Total Reward: 0.0
Episode 6500 Total Reward: 0.0
Episode 6550 Total Reward: 0.0
Episode 6600 Total Reward: 0.0
Episode 6650 Total Reward: 0.0
Episode 6700 Total Reward: 0.0
Episode 6750 Total Reward: 0.0
Episode 6800 Total Reward: 0.0
Episode 6850 Total Reward: 0.0
Episode 6900 Total Reward: 0.0
Episode 6950 Total Reward: 1.0
Episode 7000 Total Reward: 0.0
Episode 7050 Total Reward: 0.0
Episode 7100 Total Reward: 0.0
Episode 7150 Total Reward: 0.0
Episode 7200 Total Reward: 0.0
```

```
Episode 7250 Total Reward: 0.0
Episode 7300 Total Reward: 0.0
Episode 7350 Total Reward: 0.0
Episode 7400 Total Reward: 0.0
Episode 7450 Total Reward: 0.0
Episode 7500 Total Reward: 0.0
Episode 7550 Total Reward: 0.0
Episode 7600 Total Reward: 0.0
Episode 7650 Total Reward: 0.0
Episode 7700 Total Reward: 0.0
Episode 7750 Total Reward: 0.0
Episode 7800 Total Reward: 0.0
Episode 7850 Total Reward: 0.0
Episode 7900 Total Reward: 0.0
Episode 7950 Total Reward: 0.0
Episode 8000 Total Reward: 0.0
Episode 8050 Total Reward: 0.0
Episode 8100 Total Reward: 0.0
Episode 8150 Total Reward: 0.0
Episode 8200 Total Reward: 0.0
Episode 8250 Total Reward: 0.0
Episode 8300 Total Reward: 0.0
Episode 8350 Total Reward: 0.0
Episode 8400 Total Reward: 0.0
Episode 8450 Total Reward: 0.0
Episode 8500 Total Reward: 0.0
Episode 8550 Total Reward: 1.0
Episode 8600 Total Reward: 0.0
Episode 8650 Total Reward: 0.0
Episode 8700 Total Reward: 0.0
Episode 8750 Total Reward: 0.0
Episode 8800 Total Reward: 0.0
Episode 8850 Total Reward: 0.0
Episode 8900 Total Reward: 0.0
Episode 8950 Total Reward: 0.0
Episode 9000 Total Reward: 1.0
Episode 9050 Total Reward: 0.0
Episode 9100 Total Reward: 0.0
Episode 9150 Total Reward: 1.0
Episode 9200 Total Reward: 0.0
Episode 9250 Total Reward: 0.0
Episode 9300 Total Reward: 0.0
Episode 9350 Total Reward: 0.0
Episode 9400 Total Reward: 0.0
Episode 9450 Total Reward: 0.0
Episode 9500 Total Reward: 0.0
Episode 9550 Total Reward: 0.0
Episode 9600 Total Reward: 0.0
```

```
Episode 9650 Total Reward: 0.0
Episode 9700 Total Reward: 0.0
Episode 9750 Total Reward: 0.0
Episode 9800 Total Reward: 0.0
Episode 9850 Total Reward: 0.0
Episode 9900 Total Reward: 0.0
Episode 9950 Total Reward: 0.0
Episode 10000 Total Reward: 1.0
```

```python
# referenced from https://www.kaggle.com/charel/
 ↪learn-by-example-reinforcement-learning-with-gym


env = gym.make('FrozenLake-v0')
env.reset()

Q = np.zeros([env.observation_space.n, env.action_space.n]) #we could also make
 ↪this dynamic if we don't know all games states upfront
gamma = 0.95 # discount factor
alpha = 0.9 # learning rate
epsilon = 0.1 #
for episode in range(1,50001):
    done = False
    obs = env.reset()
    while done != True:
        if np.random.rand(1) < epsilon:
            # exploration with a new option with probability epsilon, the
 ↪epsilon greedy approach
            action = env.action_space.sample()
        else:
            # exploitation
            action = np.argmax(Q[obs])
        obs2, rew, done, info = env.step(action) #take the action
        Q[obs,action] += alpha * (rew + gamma * np.max(Q[obs2]) -
 ↪Q[obs,action]) #Update Q-marix using Bellman equation
        obs = obs2

    if episode % 5000 == 0:
        #report every 5000 steps, test 100 games to get avarage point score for
 ↪statistics and verify if it is solved
        rew_average = 0.
        for i in range(100):
            obs= env.reset()
            done=False
            while done != True:
                action = np.argmax(Q[obs])
```

```
                obs, rew, done, info = env.step(action) #take step using␣
    ↪selected action
                rew_average += rew
        rew_average=rew_average/100
        print('Episode {} avarage reward: {}'.format(episode, rew_average))

        if rew_average > 0.8:
            # FrozenLake-v0 defines "solving" as getting average reward of 0.78␣
    ↪over 100 consecutive trials.
            # Test it on 0.8 so it is not a one-off lucky shot solving it
            print("Frozen lake solved")
            break
```

```
Episode 5000 avarage reward: 0.0
Episode 10000 avarage reward: 0.28
Episode 15000 avarage reward: 0.45
Episode 20000 avarage reward: 0.37
Episode 25000 avarage reward: 0.54
Episode 30000 avarage reward: 0.68
Episode 35000 avarage reward: 0.21
Episode 40000 avarage reward: 0.0
Episode 45000 avarage reward: 0.17
Episode 50000 avarage reward: 0.44
```

```python
[ ]: # Let's see how the algorithm solves the frozen-lakes game

rew_tot=0.
obs= env.reset()
done=False
while done != True:
    action = np.argmax(Q[obs])
    obs, rew, done, info = env.step(action) #take step using selected action
    rew_tot += rew
    env.render()

print("Reward:", rew_tot)
```

```
  (Left)
SFFF
FHFH
FFFH
HFFG
  (Left)
SFFF
FHFH
FFFH
```

```
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Up)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
```

```
FHFH
FFFH
HFFG
   (Up)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Up)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Up)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
```

```
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Left)
SFFF
FHFH
FFFH
HFFG
   (Up)
SFFF
FHFH
FFFH
HFFG
   (Down)
SFFF
FHFH
FFFH
HFFG
   (Up)
SFFF
FHFH
FFFH
HFFG
   (Down)
SFFF
FHFH
FFFH
HFFG
   (Up)
SFFF
FHFH
```

```
FFFH
HFFG
   (Up)
SFFF
FHFH
FFFH
HFFG
   (Down)
SFFF
FHFH
FFFH
HFFG
   (Up)
SFFF
FHFH
FFFH
HFFG
   (Down)
SFFF
FHFH
FFFH
HFFG
Reward: 1.0
```

[ ]: