

强化学习基础篇（十）OpenAI Gym环境汇总

Gym中从简单到复杂，包含了许多经典的仿真环境和各种数据，主要包含了经典控制、算法、2D机器人，3D机器人，文字游戏，Atari视频游戏等等。接下来我们会简单看看主要的常用的环境。在Gym注册表中有着大量的其他环境，就没办法介绍了。

1、经典控制环境（Classic control）

经典的强化学习示例，方便入门，包括了：

- Acrobot-v1

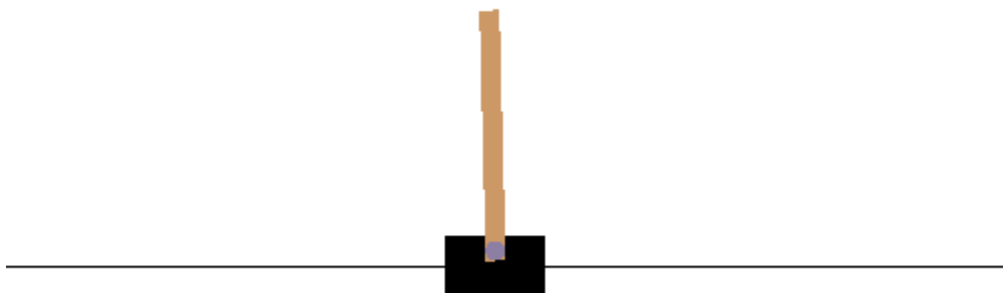
Acrobot机器人系统包括两个关节和两个连杆，其中两个连杆之间的关节可以被致动。最初，连杆是向下悬挂的，目标是将下部连杆的末端摆动到给定的高度。



Episode

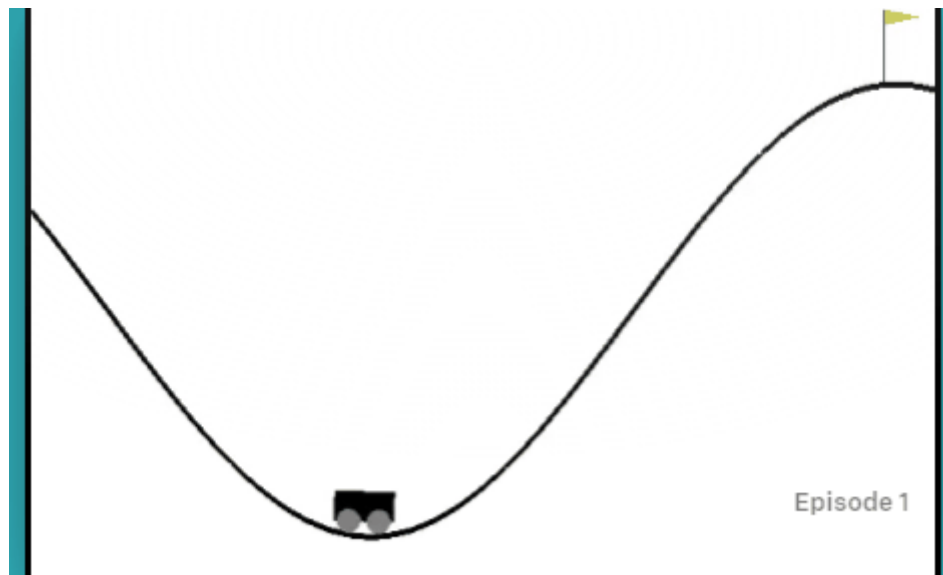
- CartPole-v1

CartPole-v1环境中，手推车上有一个杆，手推车沿着无摩擦的轨道移动。通过对推车施加+1或-1的力来控制系统。钟摆最开始为直立状态，训练的目的是防止其跌落。杆保持直立的每个时间步长都提供+1的奖励。当杆与垂直线的夹角超过15度时，或者推车从中心移出2.4个单位以上时，训练结束。



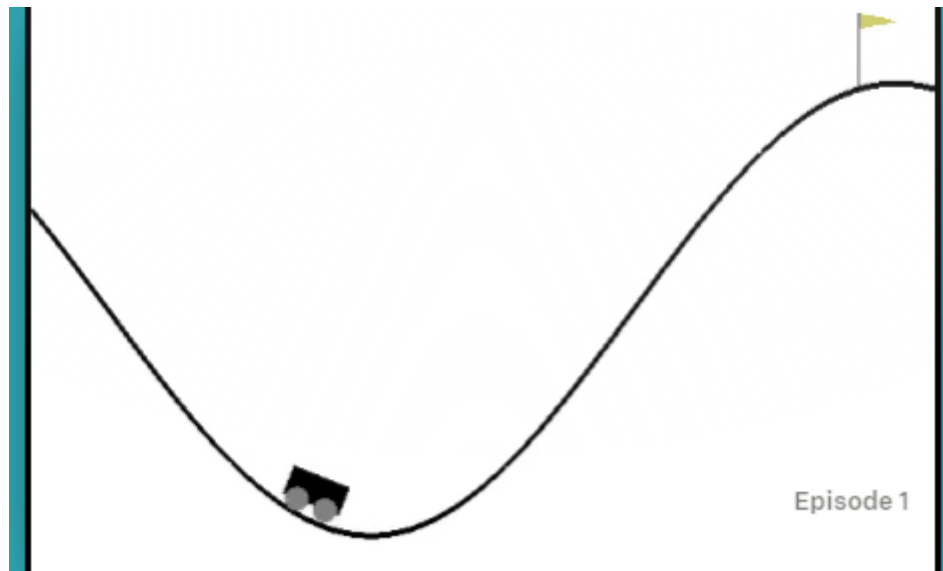
- MountainCar-v0

MountainCar-v0环境中汽车位于一维轨道上，轨道位于两个“山”之间。目标是让汽车驶向右侧的山峰；由于汽车的引擎强度不足以直接到达右侧山峰，所以，成功的唯一方法是来回驱动以产生动力。



- MountainCarContinuous-v0

MountainCarContinuous-v0环境与MountaiCar-v0环类似，动力不足的汽车必须爬上一维小山才能到达目标。与MountainCar-v0不同，动作（应用的引擎力）允许是连续值。目标位于汽车右侧的山顶上。如果汽车到达或超出，则剧集终止。



- Pendulum-v0

倒立摆摆动问题是控制文献中的经典问题。在此问题的版本中，摆锤开始于随机位置，目标是将其摆动以使其保持直立。

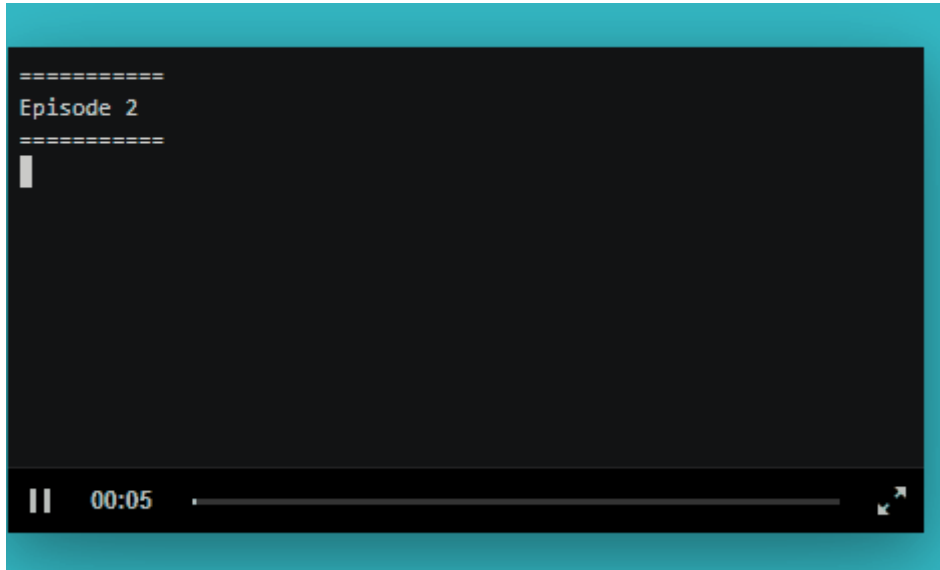


2 算法学习环境(Algorithms)

从例子中学习强化学习的相关算法，在Gym的仿真算法中，由易到难方便新手入坑，包括了：

- Copy-v0

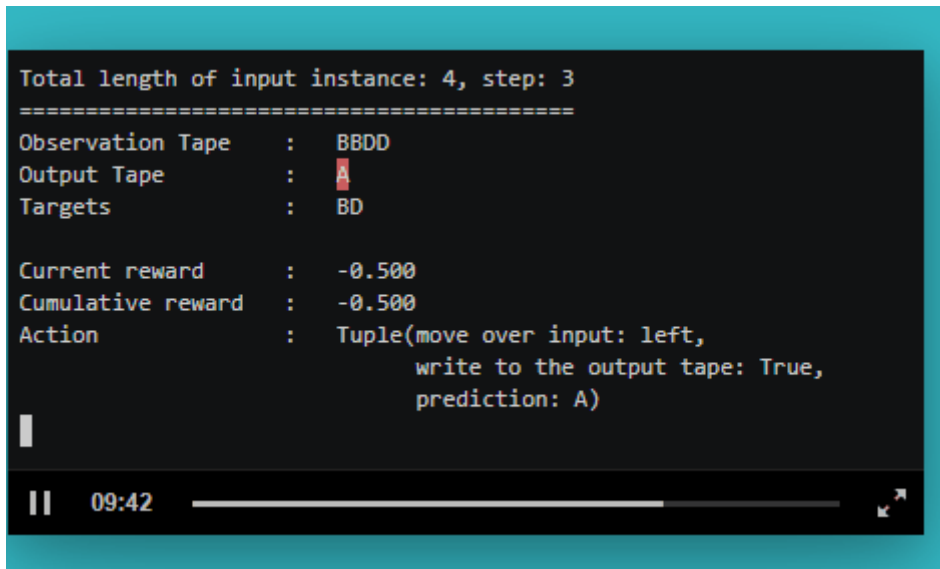
此任务就是让智能体学习对序列的拷贝操作。



- DuplicatedInput-v0

此任务让智能体学习从输入序列 $[x_1 x_2 \dots x_k]$ ，输出对每个元素的三次复制

$[x_1 x_1 x_1 x_2 x_2 x_2 \dots x_k x_k x_k]$



- RepeatCopy-v0

此任务让智能体学习对输入序列 $[mx_1 x_2 \dots x_k]$ ，输出对序列的m次复制

$[x_1 x_2 \dots x_k x_1 x_2 \dots x_k x_1 x_2 \dots x_k]$ 。

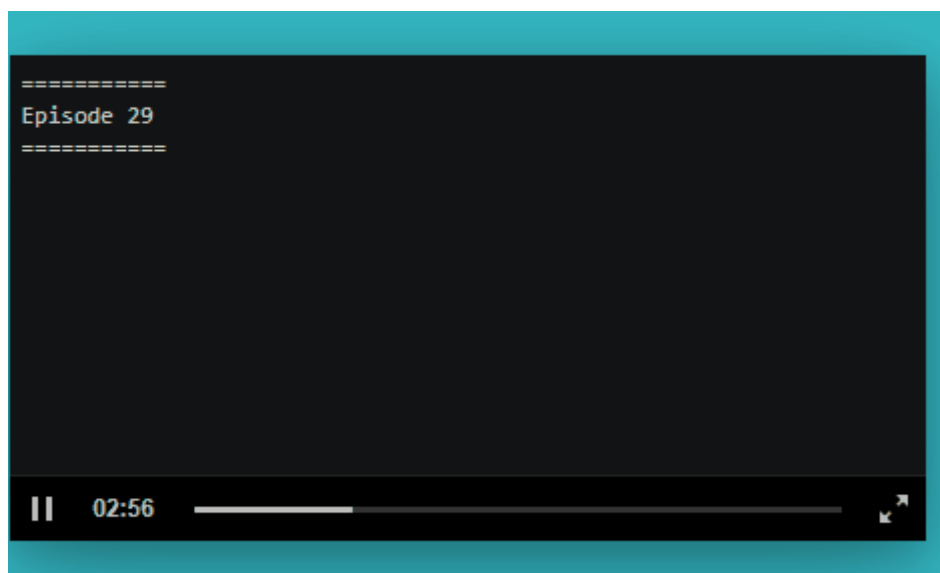
```
Total length of input instance: 3, step: 2
=====
Observation Tape : AEE
Output Tape      : B
Targets         : AEEEEAAEE

Current reward   : -0.500
Cumulative reward : -0.500
Action          : Tuple(move over input: left,
                        write to the output tape: True,
                        prediction: B)
```

10:21

- Reverse-v0

目的是反转输入序列



- ReversedAddition-v0

此任务让智能体学习在提供的网格数字序列之上增加数字序列。在这个任务中需要智能体学会 (i) 记住加法表; (ii) 学习如何在输入网格上移动, 以及 (iii) 发现进位的概念。

```
Total length of input instance: 4, step: 5
=====
Observation Grid :
                2122
                1122

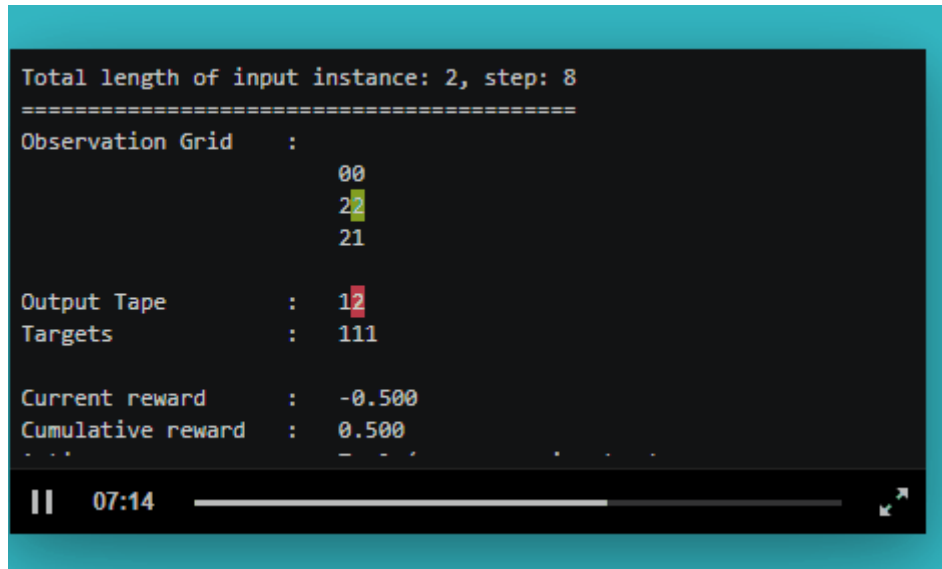
Output Tape      : 00
Targets         : 00221

Current reward   : 1.000
Cumulative reward : 2.000
Action          : Tuple(move over input: right,
```

01:32

- ReversedAddition3-v0

ReversedAddition3-v0与ReversedAddition-v0基本任务相同，但是现在要添加三个数字。由于奖励信号的更具有稀疏性，因此难度加大（因为必须先完成更多正确的操作，然后才能产生正确的输出结果）。

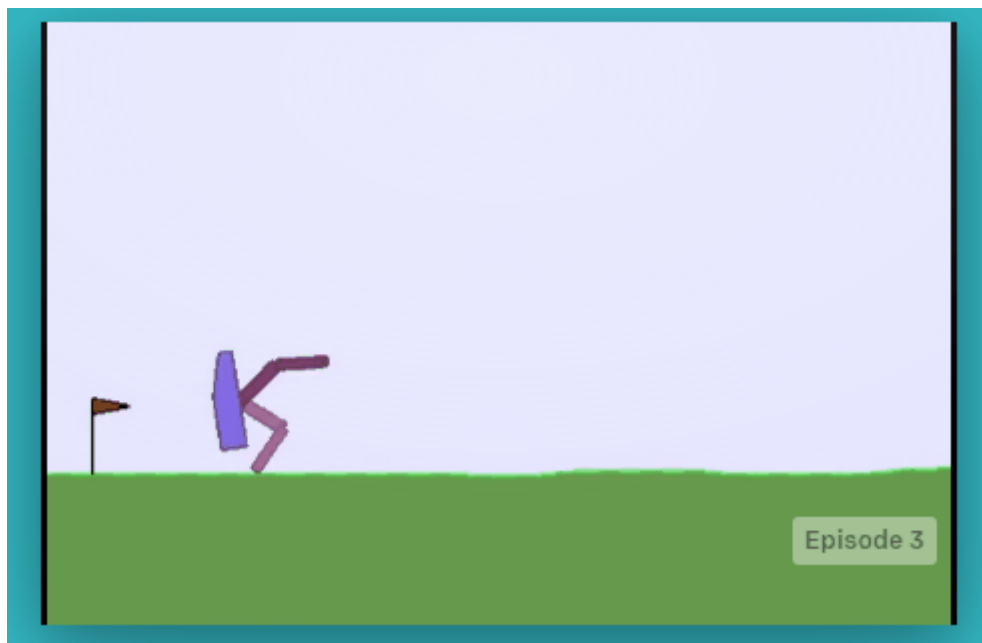


3、2D仿真环境（Box2D）

- BipedalWalker-v2

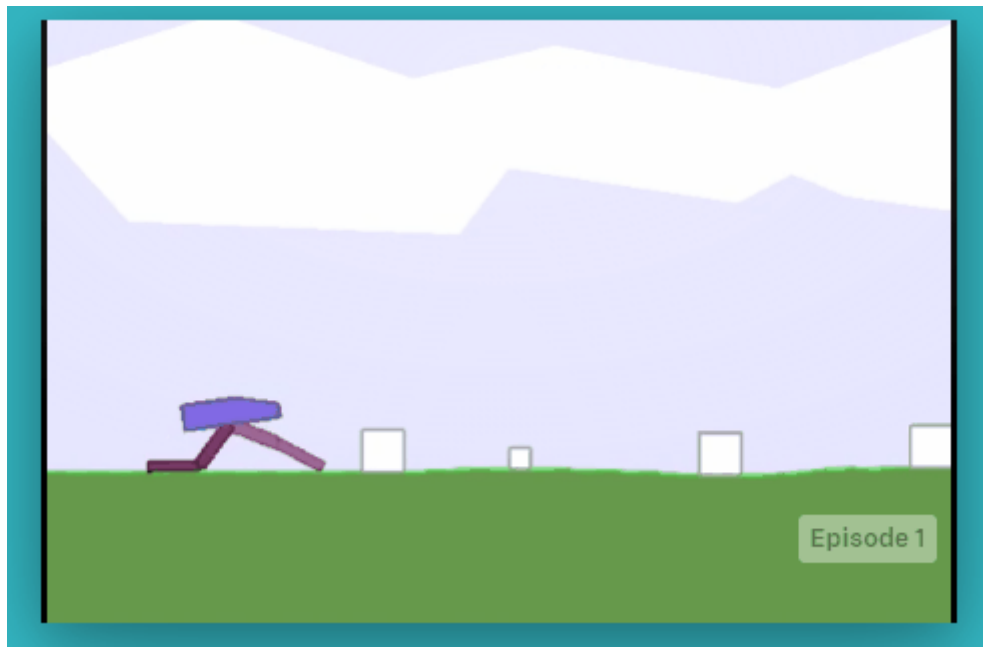
这个环境要训练机器人向前移动，走到最远的位置一共有奖励300+，如果摔倒奖励-100。

环境的状态包括了船体角速度，角速度，水平速度，垂直速度，关节位置和关节角速度，腿是否与地面的接触以及10个激光雷达测距仪的测量值。（hull angle speed, angular velocity, horizontal speed, vertical speed, position of joints and joints angular speed, legs contact with ground, and 10 lidar rangefinder measurements.）状态向量中没有坐标。



- BipedalWalkerHardcore-v2

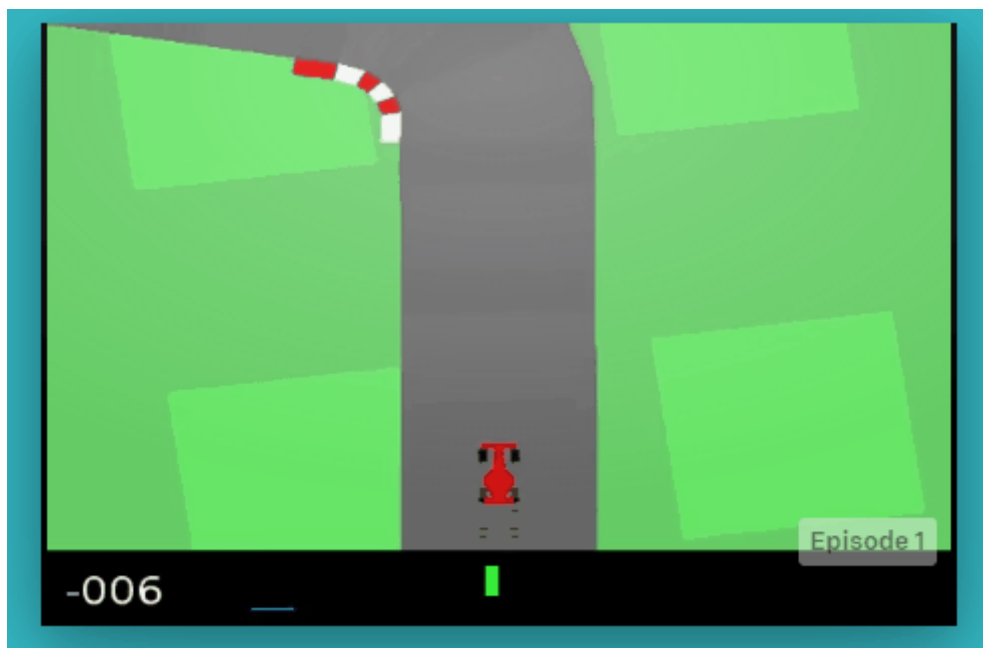
与BipedalWalker-v2相比，增加了一些障碍物。



- CarRacing-v0

CarRacing-v0是一个最简单的连续控制任务。智能体从环境的图片像素中学习。这个环境也可以进行离散动作控制，环境带有离散化的开关控制是否进行离散化的动作控制。状态是由96x96的像素组成。奖励是每帧-0.1。到达每个预定义位置会有+1000/N的奖励，所以环境里面有N个小目标。

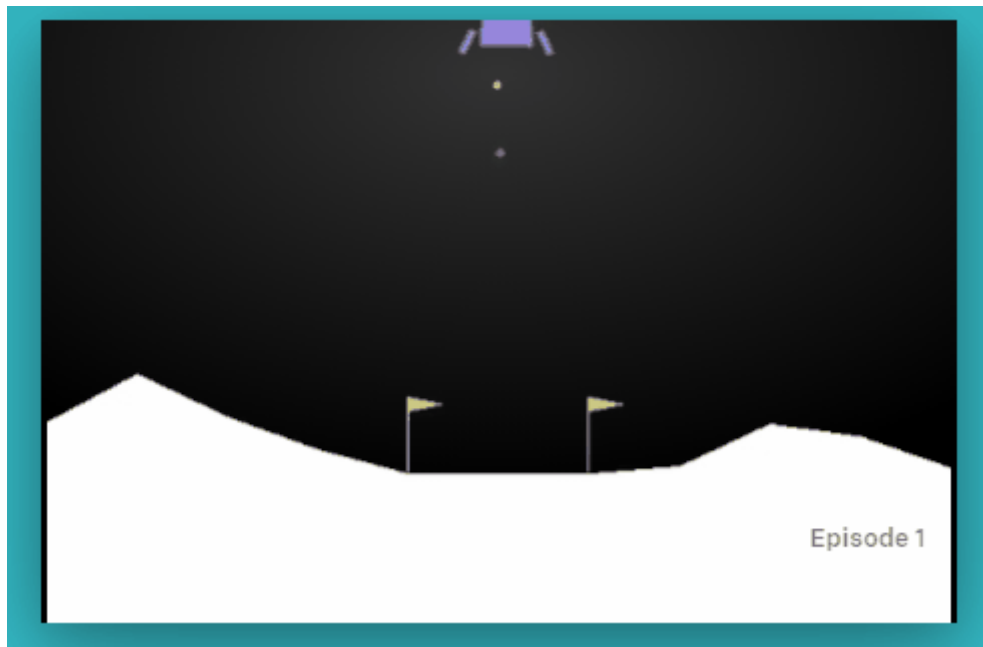
例如，智能体在732帧完成目标，则奖励为 $1000 - 0.1 * 732 = 926.8$ 。赛车碰到边界就结束。窗口底部会显示一些指标。从左到右为：真实速度，四个ABS传感器，方向盘位置，陀螺仪。



- LunarLander-v2

这是一个控制智能体着陆到指定目标的任务。

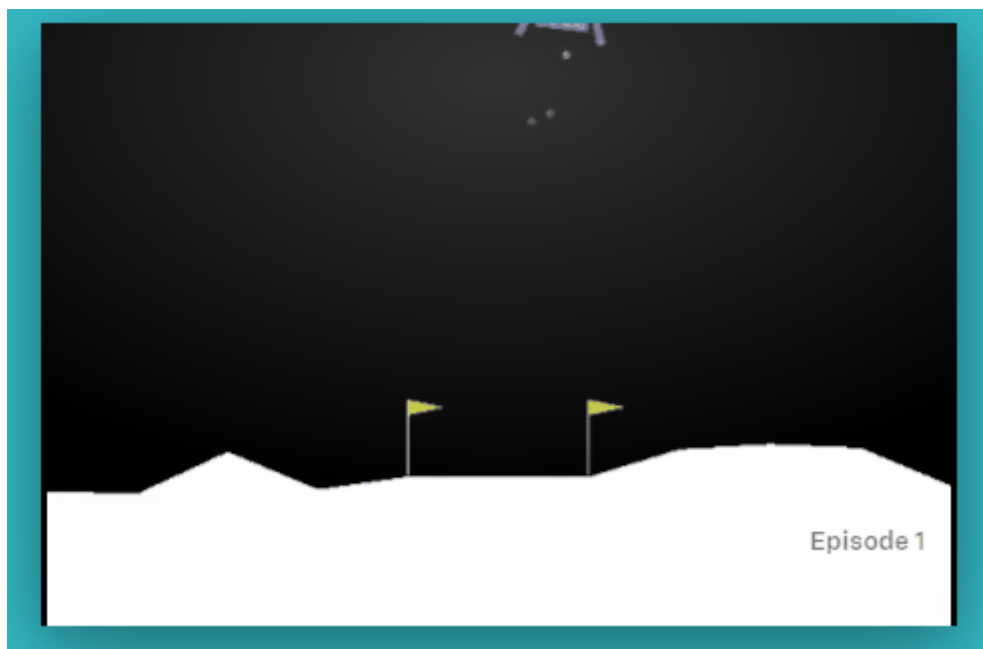
Landing pad is always at coordinates (0,0). Coordinates are the first two numbers in state vector. Reward for moving from the top of the screen to landing pad and zero speed is about 100..140 points. If lander moves away from landing pad it loses reward back. Episode finishes if the lander crashes or comes to rest, receiving additional -100 or +100 points. Each leg ground contact is +10. Firing main engine is -0.3 points each frame. Solved is 200 points. Landing outside landing pad is possible. Fuel is infinite, so an agent can learn to fly and then land on its first attempt. Four discrete actions available: do nothing, fire left orientation engine, fire main engine, fire right orientation engine.



- LunarLanderContinuous-v2

LunarLander-v2的连续控制版本

Landing pad is always at coordinates (0,0). Coordinates are the first two numbers in state vector. Reward for moving from the top of the screen to landing pad and zero speed is about 100..140 points. If lander moves away from landing pad it loses reward back. Episode finishes if the lander crashes or comes to rest, receiving additional -100 or +100 points. Each leg ground contact is +10. Firing main engine is -0.3 points each frame. Solved is 200 points. Landing outside landing pad is possible. Fuel is infinite, so an agent can learn to fly and then land on its first attempt. Action is two real values vector from -1 to +1. First controls main engine, -1..0 off, 0..+1 throttle from 50% to 100% power. Engine can't work with less than 50% power. Second value -1.0..-0.5 fire left engine, +0.5..+1.0 fire right engine, -0.5..0.5 off.

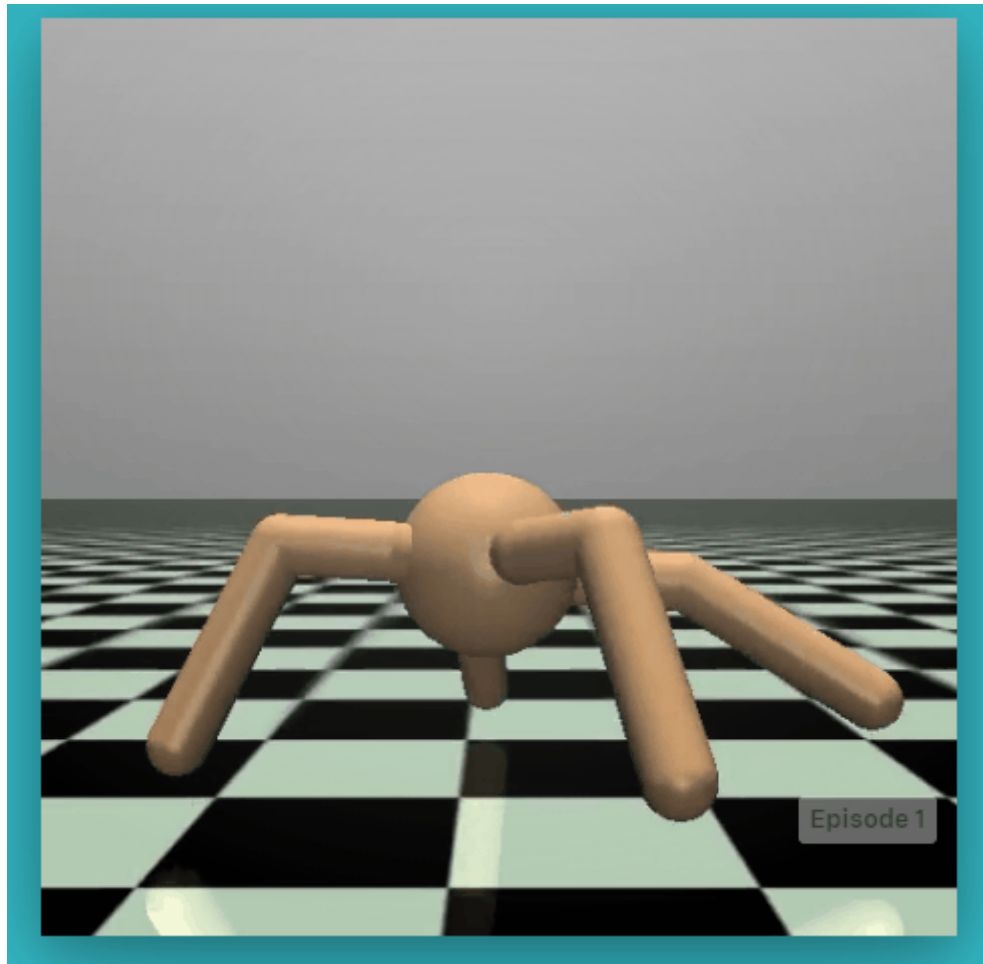


4、MuJoCo环境

MuJoCo (Multi-Joint dynamics with Contact) 是一个物理模拟器，可以用于机器人控制优化等研究。

- Ant-v2

需要训练一个四足的智能体学会行走。

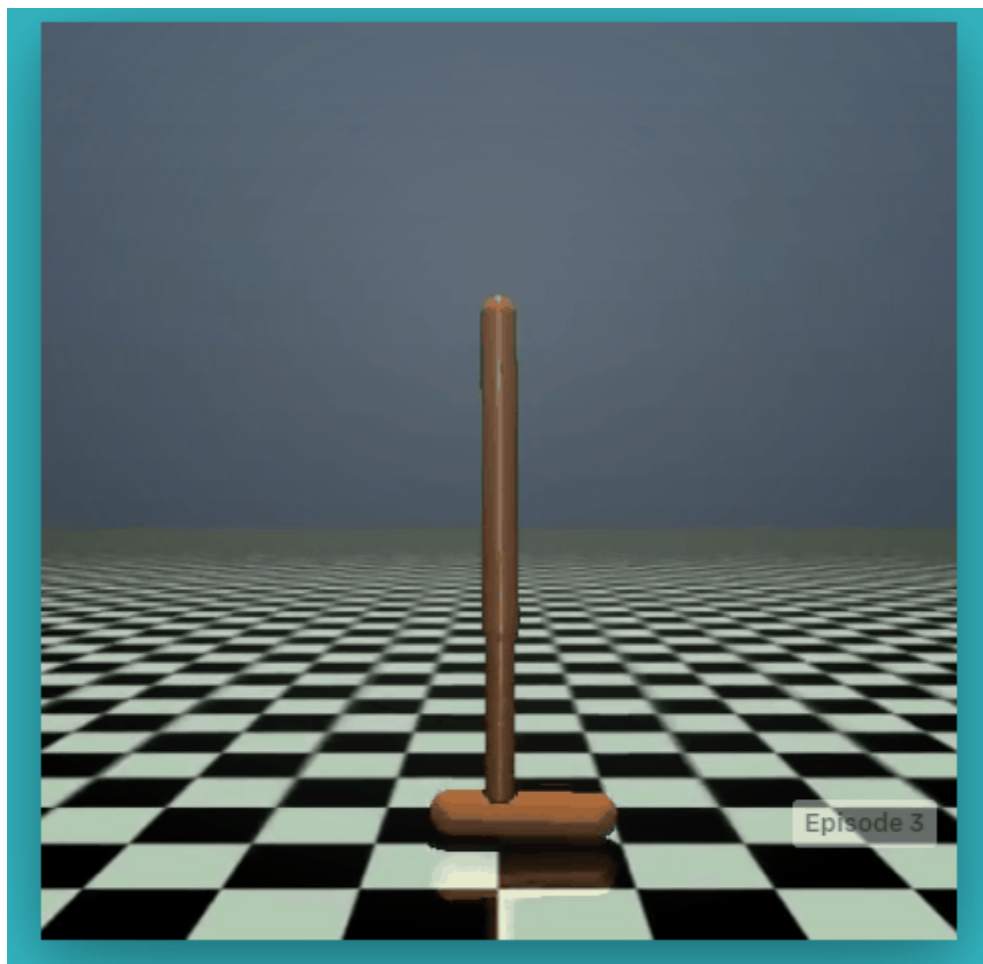


- HalfCheetah-v2

需要训练一个两足的智能体学会行走。

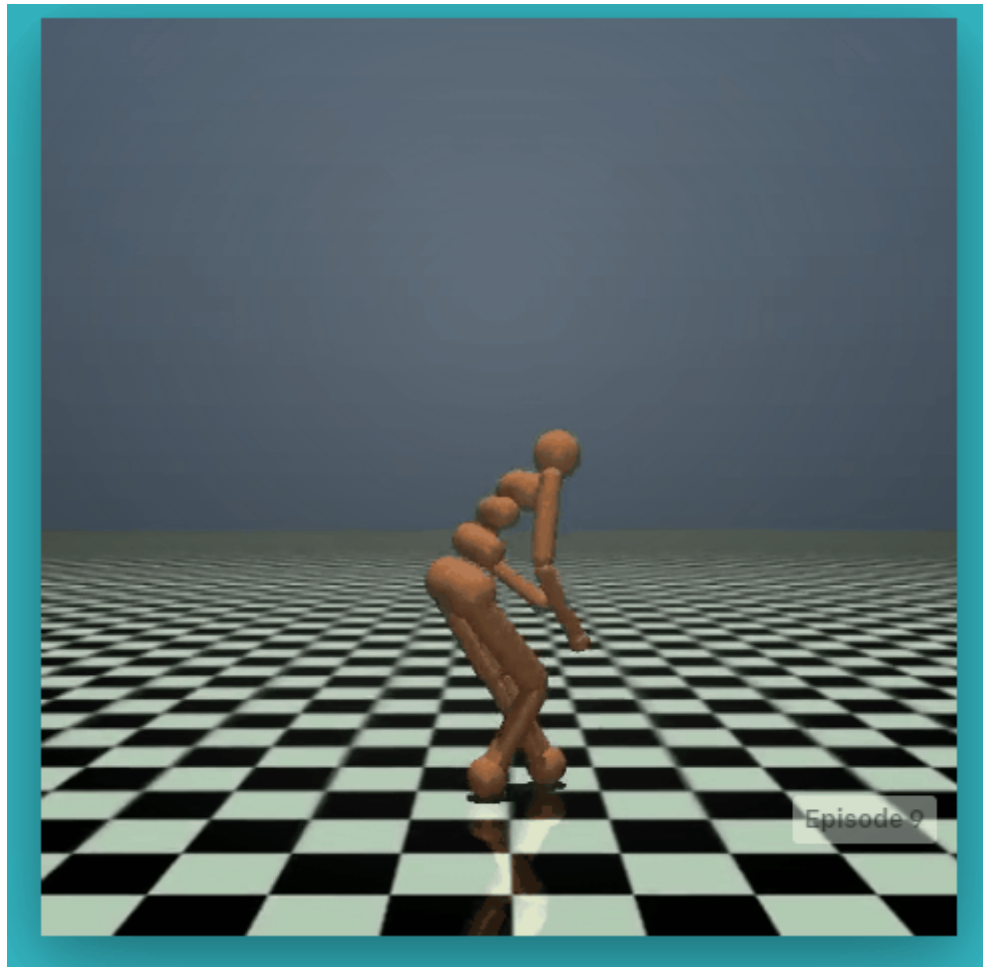


- Hopper-v2
需要训练一个单腿的智能体向前跳跃



- Humanoid-v2

需要训练三维双足智能体尽可能快地向前走，并且不会摔倒。

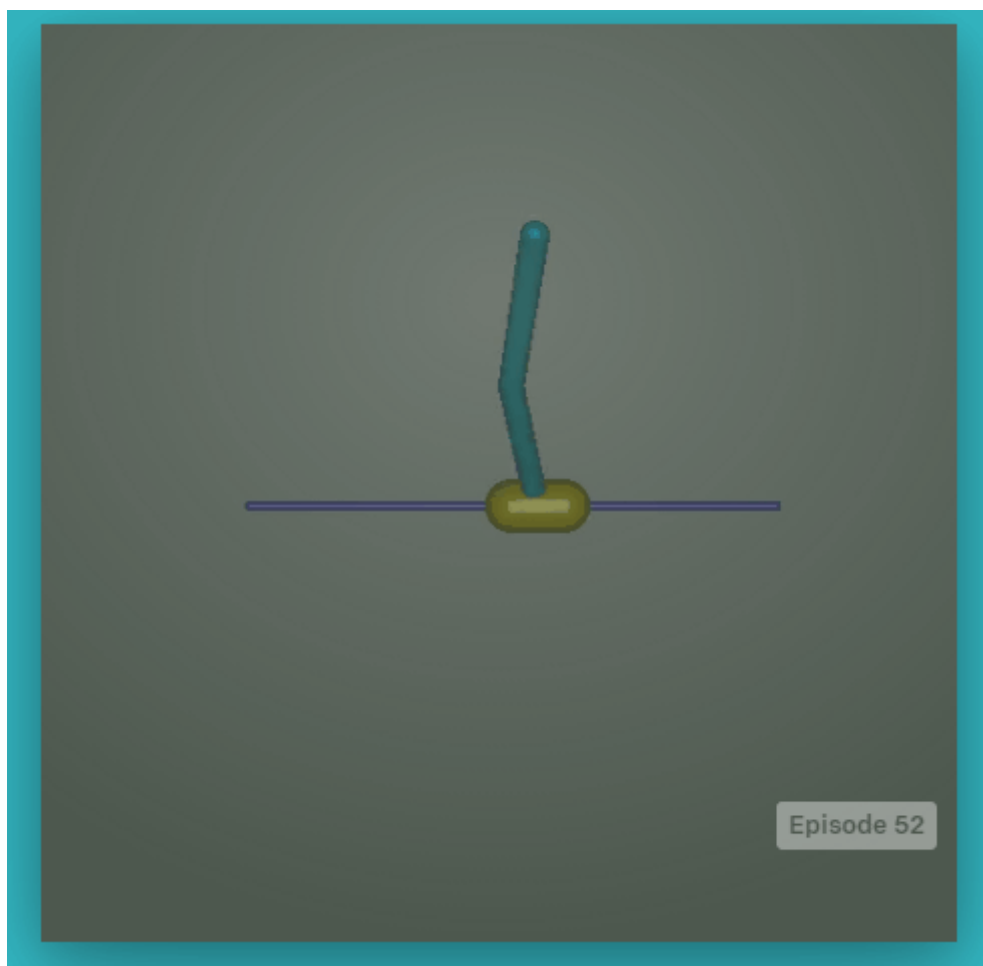


- HumanoidStandup-v2

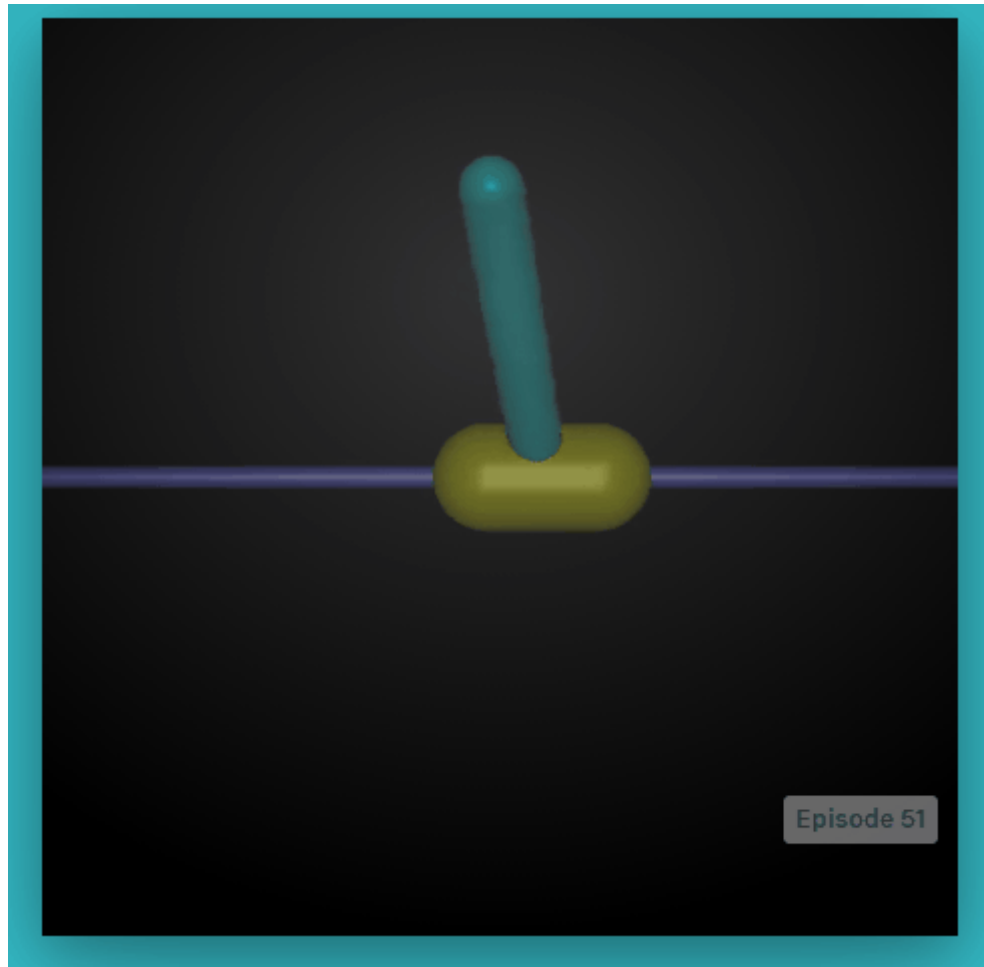
需要训练尽可能快地让使三维双足智能体学会站立。



- InvertedDoublePendulum-v2
需要智能体学会保持连杆不要掉下来

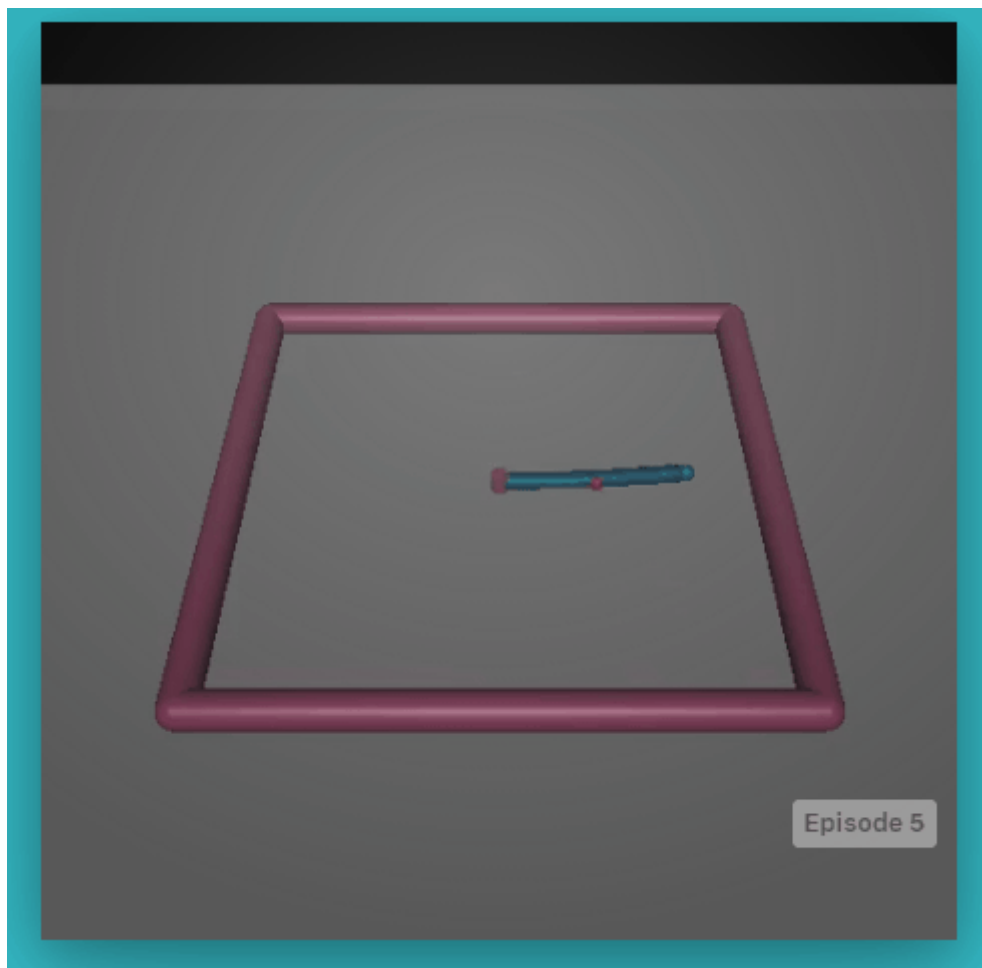


- InvertedPendulum-v2



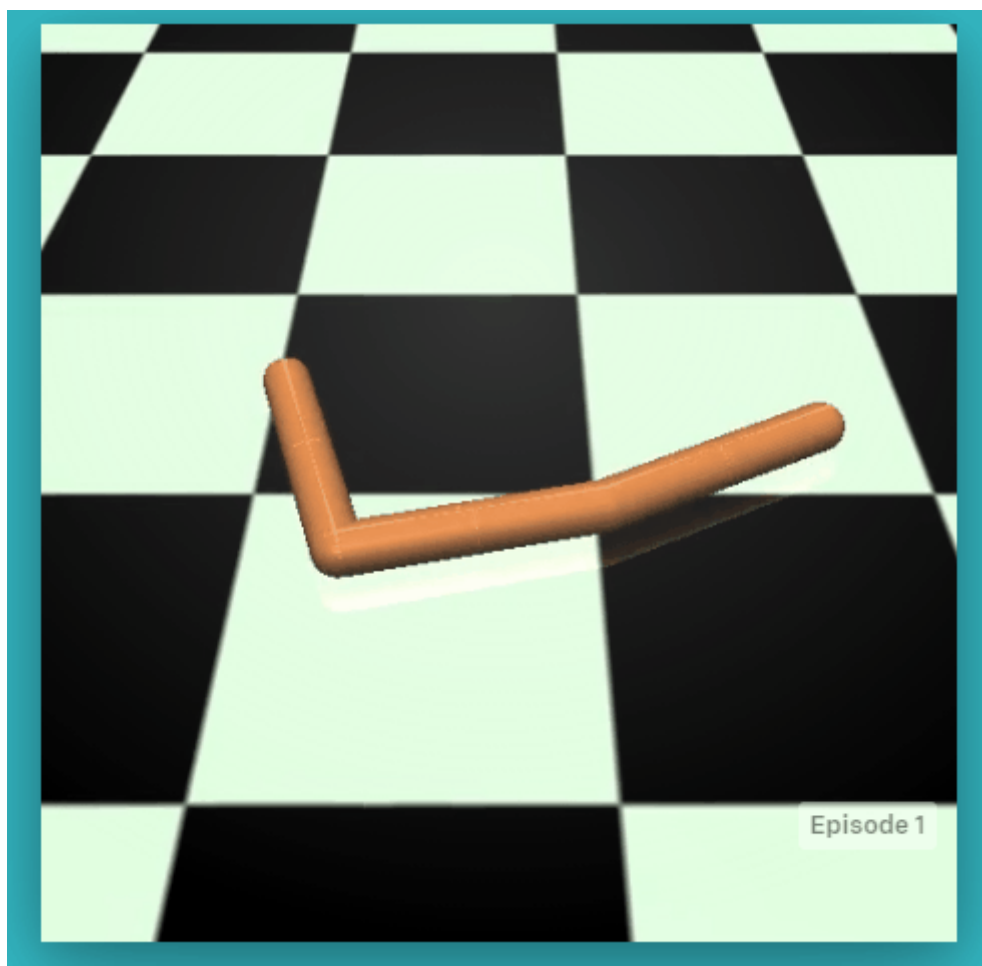
- Reacher-v2

训练智能体不断去接近一个目标。



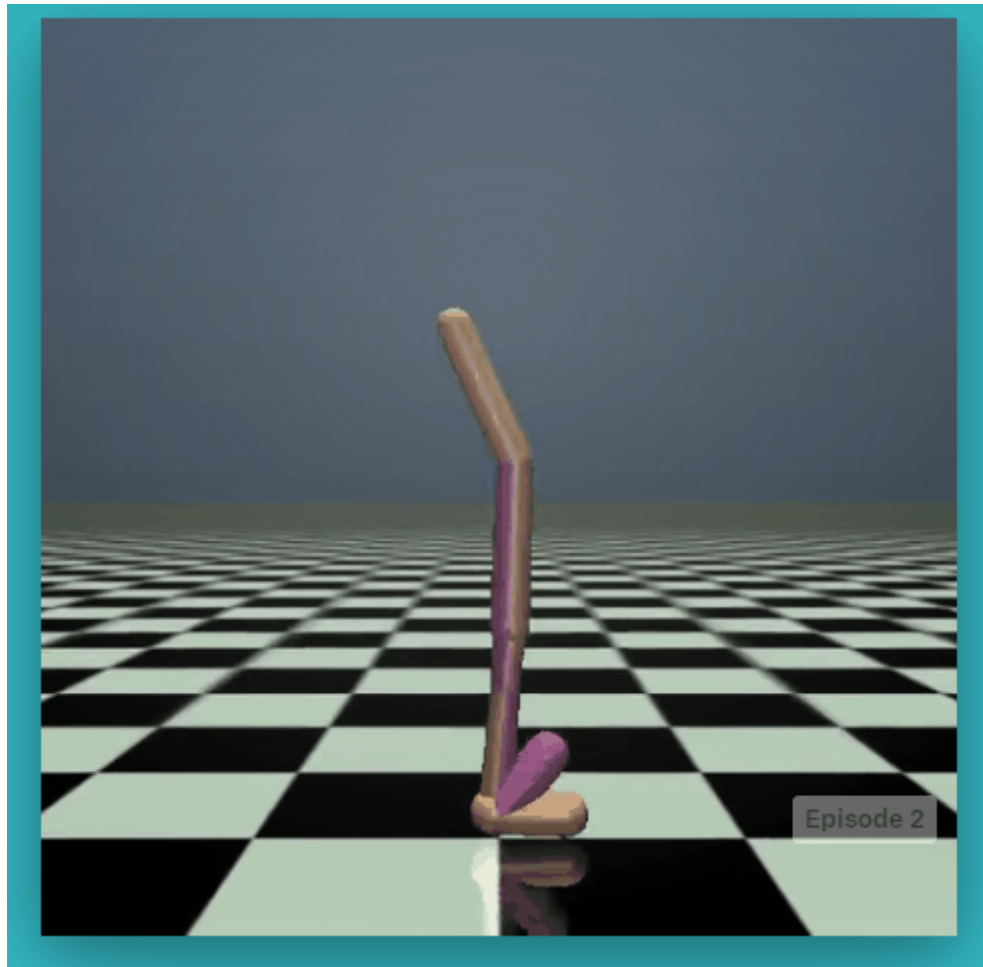
- Swimmer-v2

此任务涉及在粘性流体中的三连杆游泳智能体，其目的是通过控制两个关节，使其尽可能快地向前游泳。



- Walker2d-v2

需要训练使二维双足智能体尽可能快地向前走。



5、Atari视频游戏环境

利用强化学习来玩雅达利的游戏。Gym中集成了对强化学习有着重要影响的*Arcade Learning Environment*，并且方便用户安装；

游戏的目标都是为了在游戏中最大化游戏分数。但是他们的状态分为两类，一类是直接观测屏幕的像素输出，另一类是观测到RAM中的数据。所有的环境名称列在下表中：

游戏名称	环境名称（屏幕输出）	环境名称（RAM）
AirRaid	AirRaid-v0	AirRaid-ram-v0
Alien	Alien-v0	Alien-ram-v0
Amidar	Amidar-v0	Amidar-ram-v0
Assault	Assault-v0	Assault-ram-v0
Asterix	Asterix-v0	Asterix-ram-v0
Asteroids	Asteroids-v0	Asteroids-ram-v0
Atlantis	Atlantis-v0	Atlantis-ram-v0
BankHeist	BankHeist-v0	BankHeist-ram-v0
BattleZone	BattleZone-v0	BattleZone-ram-v0
BeamRider	BeamRider-v0	BeamRider-ram-v0
Berzerk	Berzerk-v0	Berzerk-ram-v0
Bowling	Bowling-v0	Bowling-ram-v0
Boxing	Boxing-v0	Boxing-ram-v0
Breakout	Breakout-v0	Breakout-ram-v0
Carnival	Carnival-v0	Carnival-ram-v0
Centipede	Centipede-v0	Centipede-ram-v0
ChopperCommand	ChopperCommand-v0	ChopperCommand-ram-v0
CrazyClimber	CrazyClimber-v0	CrazyClimber-ram-v0
DemonAttack	DemonAttack-v0	DemonAttack-ram-v0
DoubleDunk	DoubleDunk-v0	DoubleDunk-ram-v0
ElevatorAction	ElevatorAction-v0	ElevatorAction-ram-v0
Enduro	Enduro-v0	Enduro-ram-v0
FishingDerby	FishingDerby-v0	FishingDerby-ram-v0
Freeway	Freeway-v0	Freeway-ram-v0
Frostbite	Frostbite-v0	Frostbite-ram-v0
Gopher	Gopher-v0	Gopher-ram-v0
Gravitar	Gravitar-v0	Gravitar-ram-v0
IceHockey	IceHockey-v0	IceHockey-ram-v0
Jamesbond	Jamesbond-v0	Jamesbond-ram-v0
JourneyEscape	JourneyEscape-v0	JourneyEscape-ram-v0

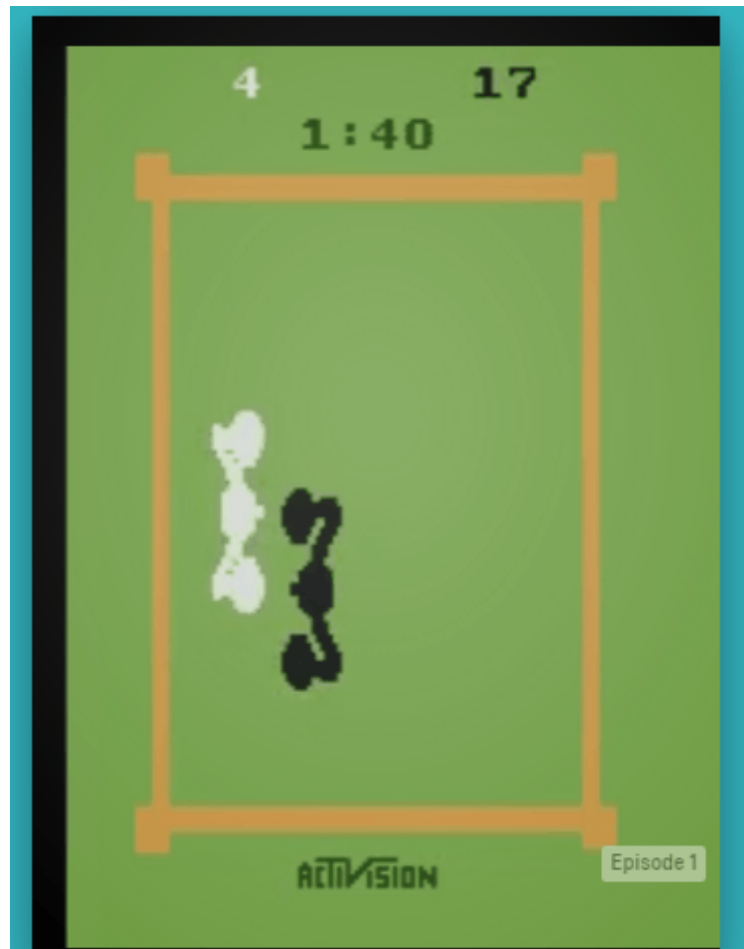
游戏名称	环境名称（屏幕输出）	环境名称（RAM）
Kangaroo	Kangaroo-v0	Kangaroo-ram-v0
Krull	Krull-v0	Krull-ram-v0
KungFuMaster	KungFuMaster-v0	KungFuMaster-ram-v0
MontezumaRevenge	MontezumaRevenge-v0	MontezumaRevenge-ram-v0
MsPacman	MsPacman-v0	MsPacman-ram-v0
NameThisGame	NameThisGame-v0	NameThisGame-ram-v0
Phoenix	Phoenix-v0	Phoenix-ram-v0
Pitfall	Pitfall-v0	Pitfall-ram-v0
Pong	Pong-v0	Pong-ram-v0
Pooyan	Pooyan-v0	Pooyan-ram-v0
PrivateEye	PrivateEye-v0	PrivateEye-ram-v0
Qbert	Qbert-v0	Qbert-ram-v0
Riverraid	Riverraid-v0	Riverraid-ram-v0
RoadRunner	RoadRunner-v0	RoadRunner-ram-v0
Robotank	Robotank-v0	Robotank-ram-v0
Seaquest	Seaquest-v0	Seaquest-ram-v0
Skiing	Skiing-v0	Skiing-ram-v0
Solaris	Solaris-v0	Solaris-ram-v0
SpaceInvaders	SpaceInvaders-v0	SpaceInvaders-ram-v0
StarGunner	StarGunner-v0	StarGunner-ram-v0
Tennis	Tennis-v0	Tennis-ram-v0
TimePilot	TimePilot-v0	TimePilot-ram-v0
Tutankham	Tutankham-v0	Tutankham-ram-v0
UpNDown	UpNDown-v0	UpNDown-ram-v0
Venture	Venture-v0	Venture-ram-v0
VideoPinball	VideoPinball-v0	VideoPinball-ram-v0
WizardOfWor	WizardOfWor-v0	WizardOfWor-ram-v0
YarsRevenge	YarsRevenge-v0	YarsRevenge-ram-v0
Zaxxon	Zaxxon-v0	Zaxxon-ram-v0

下面简单看几个在课程与论文中常常引用到的环境。

- Boxing-ram-v0

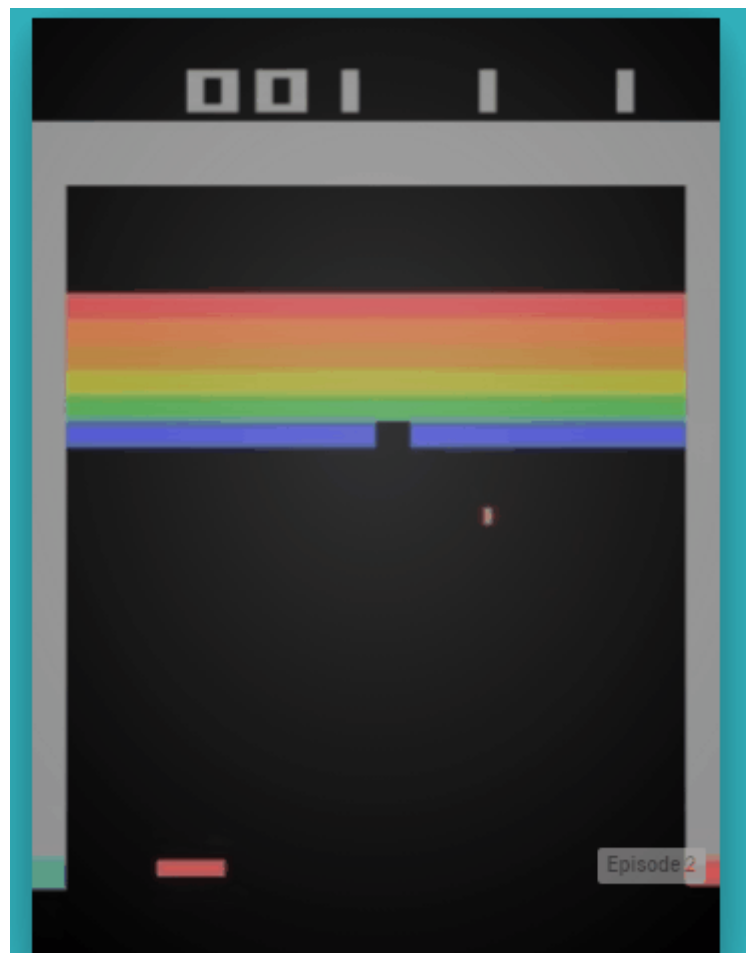
Atari 2600的拳击游戏。

拳击显示了两个拳击手的俯视图，一个白人和一个黑人。足够接近时，拳击手可以用拳打他的对手（通过按Atari操纵杆上的开火按钮执行）。这会使他的对手稍微退缩。长拳得一分，近拳（强力出拳，手动）得2分。没有击倒或回合。当一个玩家进行100次打球（“淘汰赛”）或经过两分钟（“决定”）时，比赛结束。在作出决定的情况下，拳头落地最多的玩家就是赢家。



- Breakout-ram-v0

游戏开始时，画面显示8排砖块，每隔两排，砖块的颜色就不同。由下至上的颜色排序为黄色、绿色、橙色和红色。游戏开始后，玩家必须控制一块平台左右移动以反弹一个球。当那个球碰到砖块时，砖块就会消失，而球就会反弹。当玩家未能用平台反弹球的话，那么玩家就输掉了那个回合。当玩家连续输掉3次后，玩家就会输掉整个游戏。玩家在游戏目的就是清除所有砖块。玩家破坏黄色砖块能得1分、绿色能得3分、橙色能得5分、而红色则能得7分。当球碰到画面顶部时，玩家所控制的平台长度就会减半。另外，球的移动速度会在接触砖块4次、接触砖块12次、接触橙色砖块和接触红色砖块后加速。玩家在此游戏中最高能获896分，即完成两个448分的关卡。当玩家获得896分后，游戏不会自动结束，只会待玩家输掉3次后才会结束。



- Enduro-ram-v0

Enduro includes playing a long-distance race National Enduro by controlling a race car. The goal of the game is to pass a certain number of cars each day. This will allow the player to continue the race the next day. The driver must avoid other race cars, pass 200 cars the first day, and 300 cars the second day.



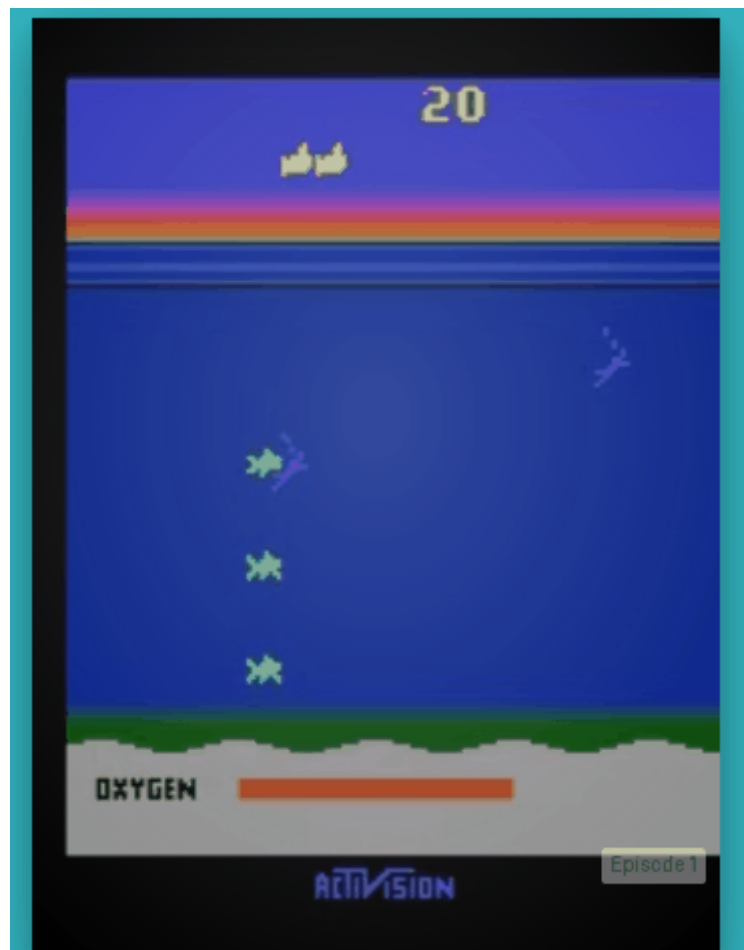
- Qbert-ram-v0

玩家通过操纵游戏角色，在躲避敌人与障碍物的同时，将方块改为目标颜色。



- Seaquest-ram-v0

水下射击游戏



- SpaceInvaders-ram-v0



6、机器人环境 (Robotics)

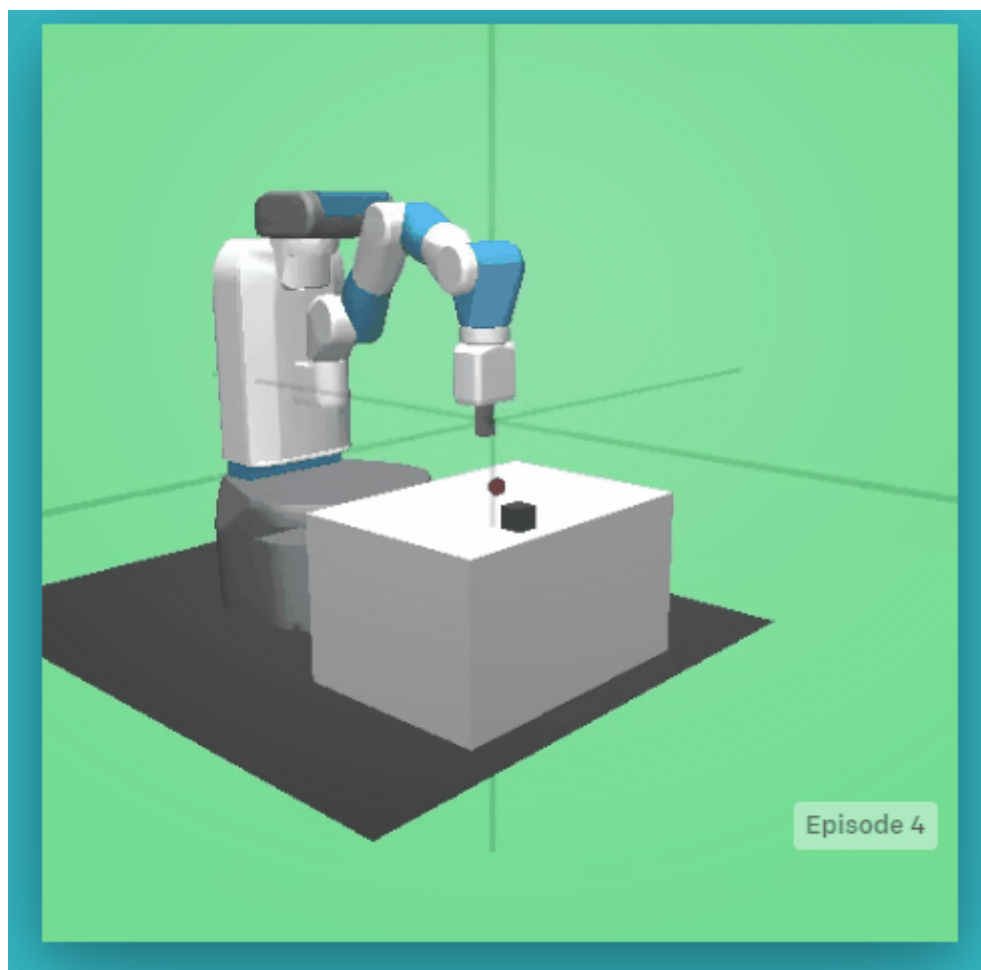
- FetchPickAndPlace-v1

训练智能体在一个3D环境中抓取一个随机选取的目标,然后把它放到。



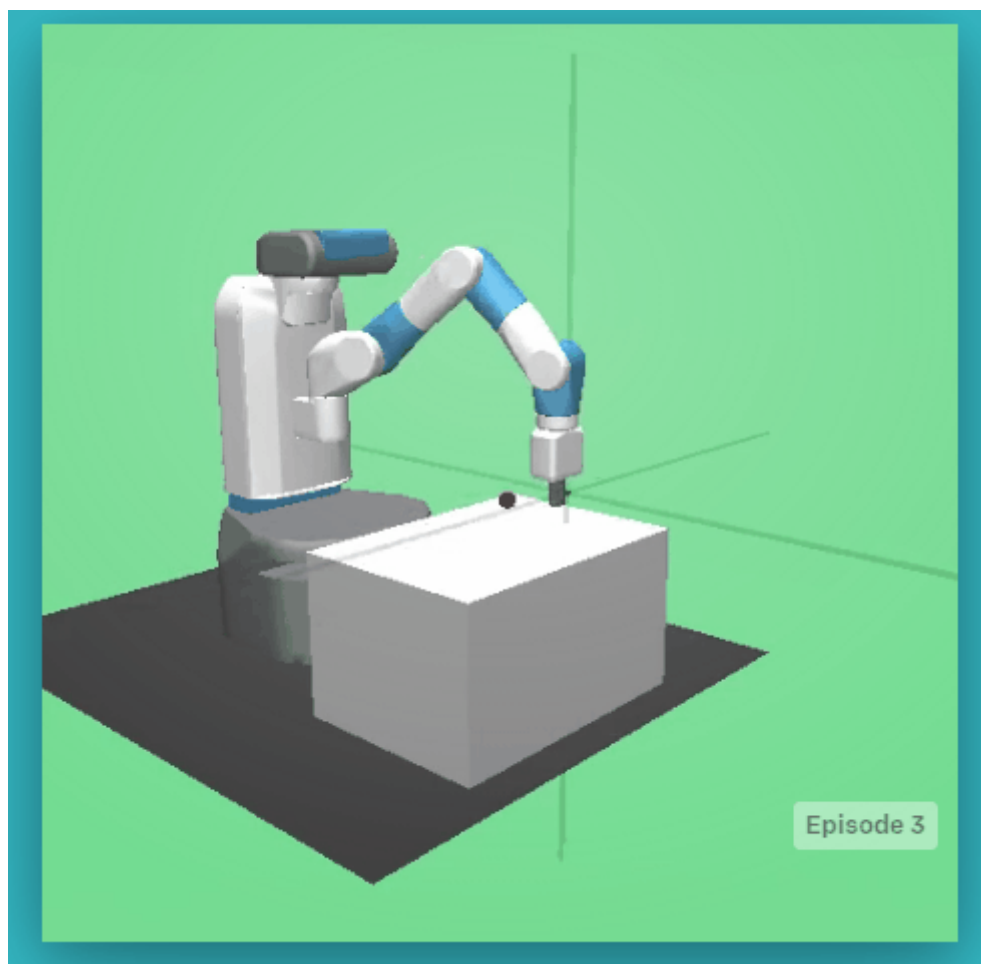
- FetchPush-v1

训练智能体在一个3D环境中将一个随机的目标推到指定位置。



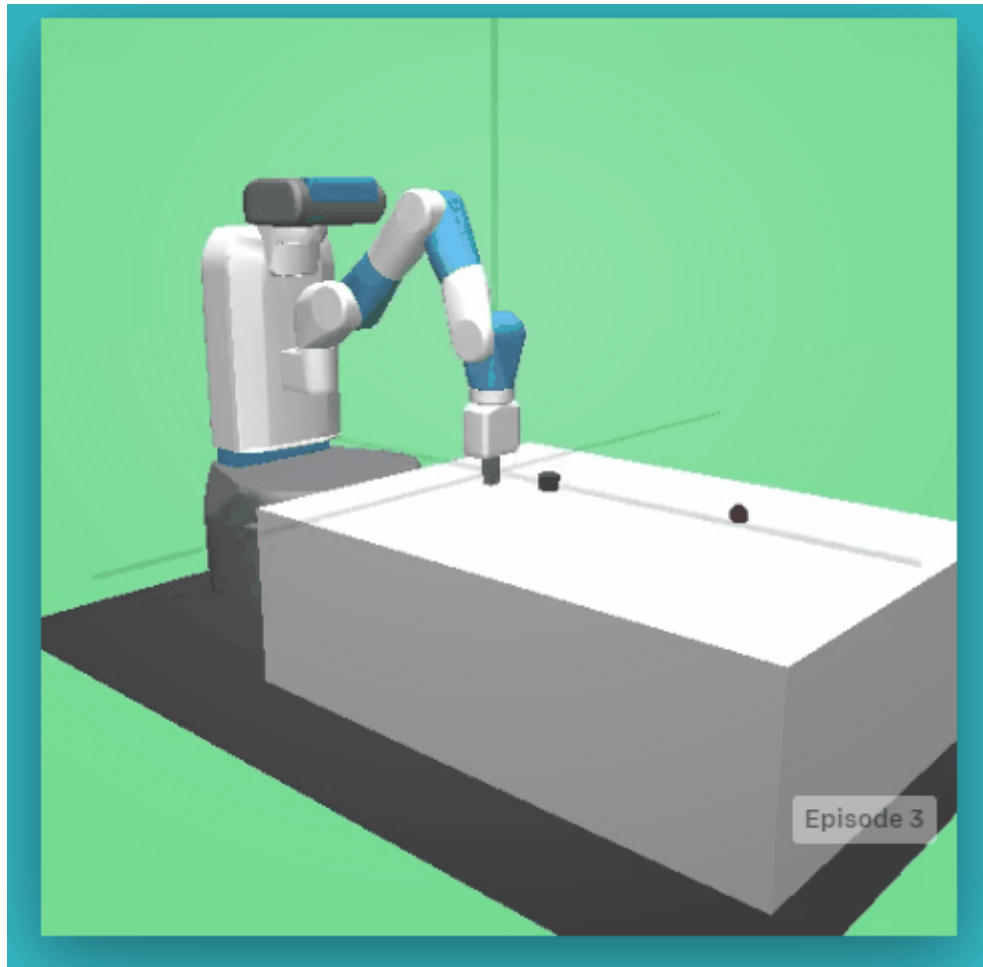
- FetchReach-v1

训练智能体，让他尽可能快的触碰到随机目标。



- FetchSlide-v1

这个环境中智能体需要学会把一个圆盘推到一个无法触碰到的随机目标旁边。



- HandManipulateBlock-v0

需要控制机械手臂把随机摆放的方块旋转为目标样式。



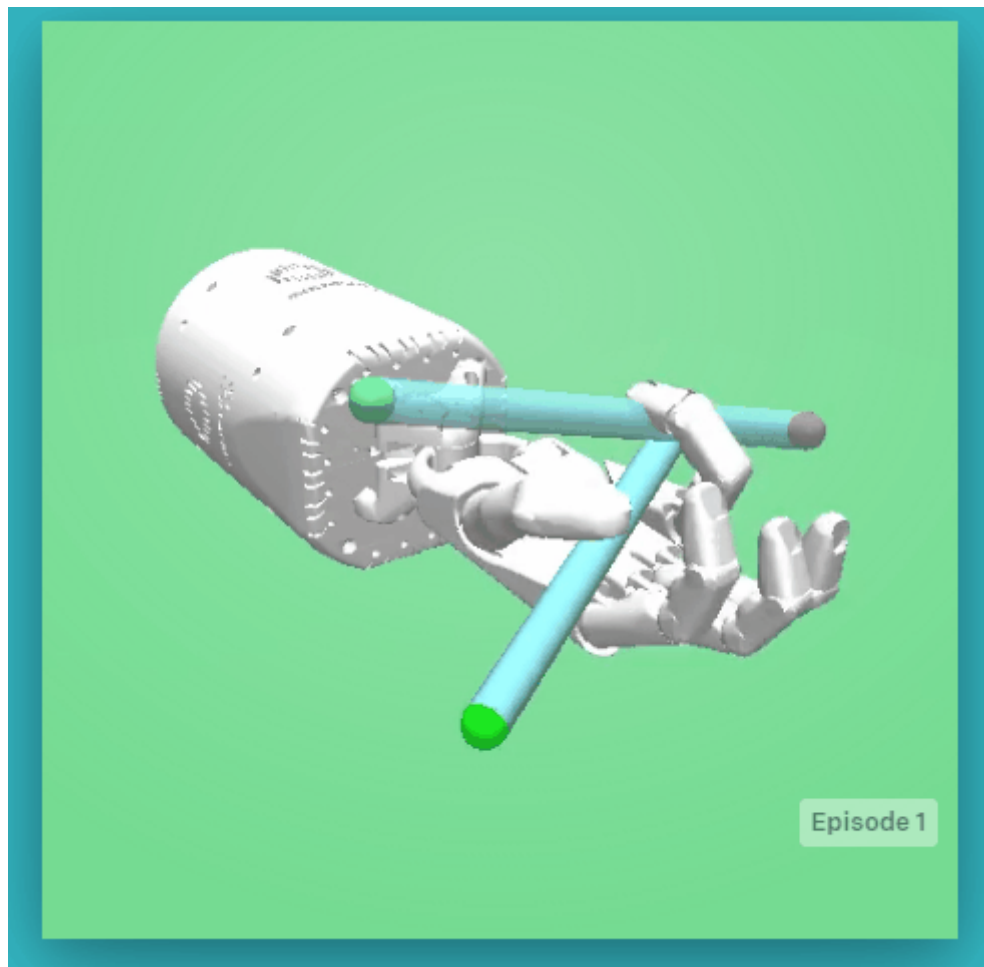
- HandManipulateEgg-v0

机械手臂上有个鸡蛋形状的物体，需要训练机械手臂学会旋转他到指定的方向。



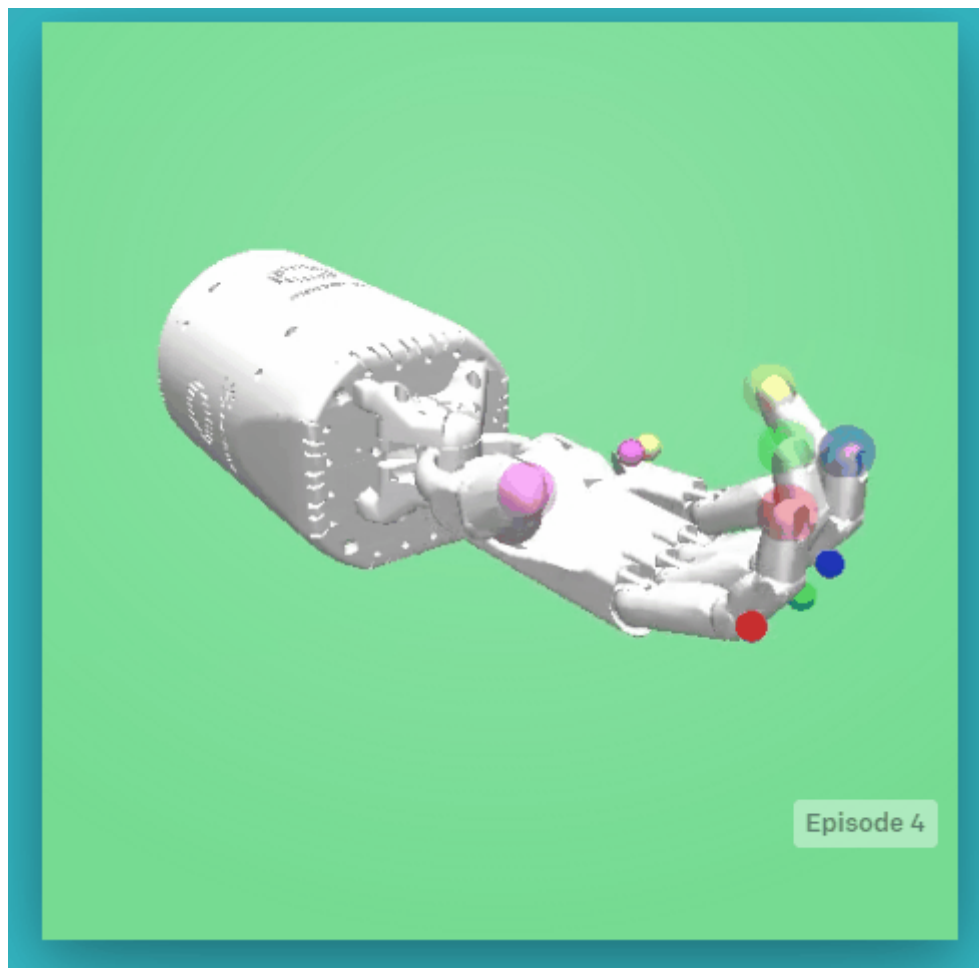
- HandManipulatePen-v0

训练机器臂将笔摆到指定的随机位置。



- HandReach-v0

训练智能体将污垢手指都摆放为目标姿势。



7、文本游戏环境 (Toy text)

- Blackjack-v0

游戏规则：

拥有最高点数的玩家获胜，其点数必须等于或低于21点；超过21点的玩家称为爆牌（Bust）。2点至10点的牌以牌面的点数计算，J、Q、K 每张为10点。A可记为1点或为11点，而2-10则按牌面点数算，若玩家会因A而爆牌则A可算为1点。当一手牌中的A算为11点时，这手牌便称为“软牌”（soft hand），因为除非玩者再拿另一张牌，否则不会出现爆牌。

庄家在取得17点之前必须要牌，因规则不同会有软17点或硬17点才停牌的具体区分。

每位玩家的目的是要取得最接近21点数的牌来击败庄家，但同时要避免爆牌。要注意的是，若玩家爆牌在先即为输，就算随后庄家爆牌也是如此。若玩家和庄家拥有同样点数，这样的状态称为“push”，玩家和庄家皆不算输赢。每位玩者和庄家之间的游戏都是独立的，因此在同一局内，庄家有可能会输给某些玩家，但也同时击败另一些玩家。

牌桌上通常会印有最小和最大的赌注，每一间赌场的每一张牌桌的限额都可能不同。在第一笔筹码下注后，庄家开始发牌，若是从一副或两副牌中发牌，称为“pitch”牌局；较常见的则是从四副牌中发牌。庄家会发给每位玩家和自己两张牌，庄家的两张牌中会有一张是点数朝上的“明牌”，所有玩家皆可看见，另一张则是点数朝下的“暗牌”。若是四副牌时，发牌时点数会朝上，若为“pitch”牌局则发牌时点数朝下。

- GuessingGame-v0

游戏的目标是在200个时间步内猜测一个随机数，并允许1%的误差。在每个步骤之后，为智能体提供四个可能的观察之一，该观察包含的信息为：0-尚未提交任何猜测（仅在重置后），1-猜测低于目标，2-猜测等于目标，3-猜测高于目标

奖励为：猜测结果在目标1%之内奖励1，否则奖励0。

- HotterColder-v0

这个游戏与GuessingGame-v0差不多，但是目标是让智能体可以有效利用奖励来采取最佳行动。

在每个步骤之后，智能体都会收到以下观察结果：0-尚未提交任何猜测（仅在重置后，1-猜测低于目标，2-猜测等于目标，3-猜测高于目标

奖励的计算公式为：

$$\left(\frac{\min(\text{action}, \text{self.number}) + \text{self.bounds}}{\max(\text{action}, \text{self.number}) + \text{self.bounds}} \right)^2$$

这样的奖励的设计方式实际上是对猜测结果的百分比做了一个平方，结果更能反馈猜测结果与目标的差距。

- NChain-v0

n-chain环境，智能体在一个n-chain的一维空间上前后移动，先前移动没有任何奖励，向后移动会有少量奖励。但是，链尾提供了很大的奖励，通过在链的末尾移动“前进”，可以重复此奖励。

- Roulette-v0

这是智能体玩一个有0-36数字的轮盘（赌场那种轮盘游戏）。每次旋转轮盘，智能体都下注一个数字（0-36），如果结果不是0，并且与下注部分匹配则得到正奖励。同时，智能体可以随之选择离场。

- FrozenLake-v0

FrozenLake环境就是一个GridWorld环境，名字是指在一块冰面上有四种state：

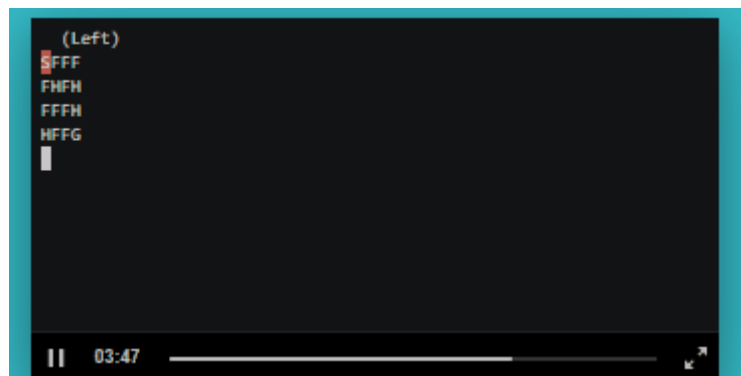
S: initial stat 起点

F: frozen lake 冰湖

H: hole 窟窿

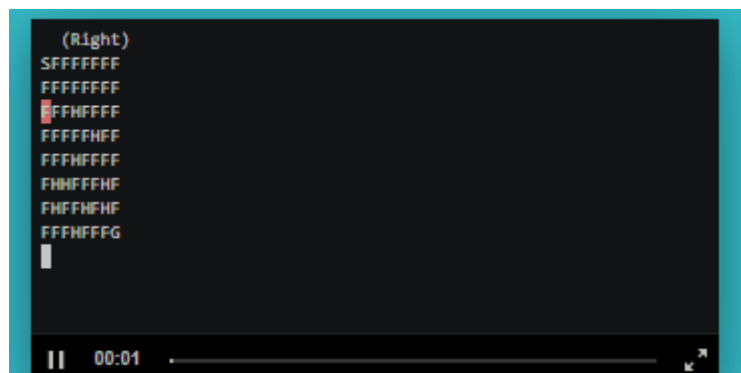
G: the goal 目的地

智能体要学会从起点走到目的地，并且不要掉进窟窿。



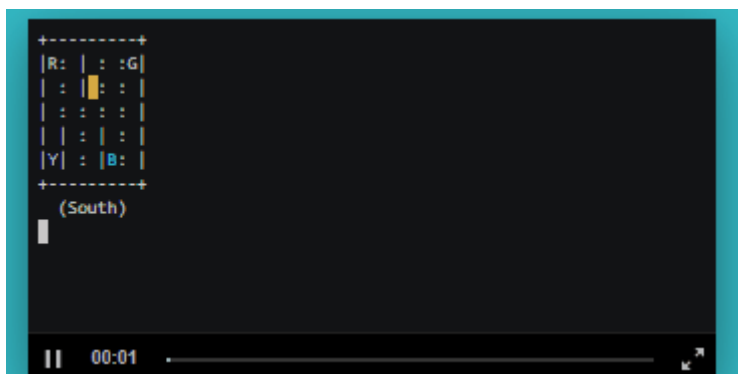
- FrozenLake8x8-v0

8*8的FrozenLake



- Taxi-v3

这个任务是在《Hierarchical Reinforcement Learning with the MAXQ Value Function Decomposition》这篇论文中引用，来论述分层强化学习中的问题。设定中有4个用不同字母标记的四个点（R,G,Y,B）。智能体作为出租车，在一个点接客，然后在另一个点下车。成功下客获得+20奖励，每个时间步奖励-1。如果接送行为错了-10分。



8、其他注册的环境

上面只能介绍到主要的环境，还有大量的注册环境，可以通过查询注册表找到：

```
1 from gym import envs
2 print(envs.registry.all())
```

这里列出来在0.17.3版本以查询到的所有环境：

RepeatCopy-v0	Enduro-ram-v0	BankHeistNoFrameskip-v4	Pong-v4
ReversedAddition-v0	Enduro-ram-v4	BankHeist-ram-v0	PongDeterministic-v0
ReversedAddition3-v0	Enduro-ramDeterministic-v0	BankHeist-ram-v4	PongDeterministic-v4
DuplicatedInput-v0	Enduro-ramDeterministic-v4	BankHeist-ramDeterministic-v0	PongNoFrameskip-v0
Reverse-v0	Enduro-ramNoFrameskip-v0	BankHeist-ramDeterministic-v4	PongNoFrameskip-v4
CartPole-v0	Enduro-ramNoFrameskip-v4	BankHeist-ramNoFrameskip-v0	Pong-ram-v0
CartPole-v1	FishingDerby-v0	BankHeist-ramNoFrameskip-v4	Pong-ram-v4
MountainCar-v0	FishingDerby-v4	BattleZone-v0	Pong-ramDeterministic-v0
MountainCarContinuous-v0	FishingDerbyDeterministic-v0	BattleZone-v4	Pong-ramDeterministic-v4
Pendulum-v0	FishingDerbyDeterministic-v4	BattleZoneDeterministic-v0	Pong-ramNoFrameskip-v0
Acrobot-v1	FishingDerbyNoFrameskip-v0	BattleZoneDeterministic-v4	Pong-ramNoFrameskip-v4
LunarLander-v2	FishingDerbyNoFrameskip-v4	BattleZoneNoFrameskip-v0	Pooyan-v0
LunarLanderContinuous-v2	FishingDerby-ram-v0	BattleZoneNoFrameskip-v4	Pooyan-v4
BipedalWalker-v3	FishingDerby-ram-v4	BattleZone-ram-v0	PooyanDeterministic-v0
BipedalWalkerHardcore-v3	FishingDerby-ramDeterministic-v0	BattleZone-ram-v4	PooyanDeterministic-v4
CarRacing-v0	FishingDerby-ramDeterministic-v4	BattleZone-ramDeterministic-v0	PooyanNoFrameskip-v0
Blackjack-v0	FishingDerby-ramNoFrameskip-v0	BattleZone-ramDeterministic-v4	PooyanNoFrameskip-v4
KellyCoinflip-v0	FishingDerby-ramNoFrameskip-v4	BattleZone-ramNoFrameskip-v0	Pooyan-ram-v0
KellyCoinflipGeneralized-v0	Freeway-v0	BattleZone-ramNoFrameskip-v4	Pooyan-ram-v4
FrozenLake-v0	Freeway-v4	BeamRider-v0	Pooyan-ramDeterministic-v0
FrozenLake8x8-v0	FreewayDeterministic-v0	BeamRider-v4	Pooyan-ramDeterministic-v4
CliffWalking-v0	FreewayDeterministic-v4	BeamRiderDeterministic-v0	Pooyan-ramNoFrameskip-v0
NChain-v0	FreewayNoFrameskip-v0	BeamRiderDeterministic-v4	Pooyan-ramNoFrameskip-v4
Roulette-v0	FreewayNoFrameskip-v4	BeamRiderNoFrameskip-v0	PrivateEye-v0
Taxi-v3	Freeway-ram-v0	BeamRiderNoFrameskip-v4	PrivateEye-v4
GuessingGame-v0	Freeway-ram-v4	BeamRider-ram-v0	PrivateEyeDeterministic-v0
HotterColder-v0	Freeway-ramDeterministic-v0	BeamRider-ram-v4	PrivateEyeDeterministic-v4
Reacher-v2	Freeway-ramDeterministic-v4	BeamRider-ramDeterministic-v0	PrivateEyeNoFrameskip-v0
Pusher-v2	Freeway-ramNoFrameskip-v0	BeamRider-ramDeterministic-v4	PrivateEyeNoFrameskip-v4
Thrower-v2	Freeway-ramNoFrameskip-v4	BeamRider-ramNoFrameskip-v0	PrivateEye-ram-v0
Striker-v2	Frostbite-v0	BeamRider-ramNoFrameskip-v4	PrivateEye-ram-v4
InvertedPendulum-v2	Frostbite-v4	Berzerk-v0	PrivateEye-ramDeterministic-v0
InvertedDoublePendulum-v2	FrostbiteDeterministic-v0	Berzerk-v4	PrivateEye-ramDeterministic-v4
HalfCheetah-v2	FrostbiteDeterministic-v4	BerzerkDeterministic-v0	PrivateEye-ramNoFrameskip-v0
HalfCheetah-v3	FrostbiteNoFrameskip-v0	BerzerkDeterministic-v4	PrivateEye-ramNoFrameskip-v4
Hopper-v2	FrostbiteNoFrameskip-v4	BerzerkNoFrameskip-v0	Qbert-v0
Hopper-v3	Frostbite-ram-v0	BerzerkNoFrameskip-v4	Qbert-v4
Swimmer-v2	Frostbite-ram-v4	Berzerk-ram-v0	QbertDeterministic-v0
Swimmer-v3	Frostbite-ramDeterministic-v0	Berzerk-ram-v4	QbertDeterministic-v4
Walker2d-v2	Frostbite-ramDeterministic-v4	Berzerk-ramDeterministic-v0	QbertNoFrameskip-v0
Walker2d-v3	Frostbite-ramNoFrameskip-v0	Berzerk-ramDeterministic-v4	QbertNoFrameskip-v4
Ant-v2	Frostbite-ramNoFrameskip-v4	Berzerk-ramNoFrameskip-v0	Qbert-ram-v0
Ant-v3	Gopher-v0	Berzerk-ramNoFrameskip-v4	Qbert-ram-v4
Humanoid-v2	Gopher-v4	Bowling-v0	Qbert-ramDeterministic-v0
Humanoid-v3	GopherDeterministic-v0	Bowling-v4	Qbert-ramDeterministic-v4
HumanoidStandup-v2	GopherDeterministic-v4	BowlingDeterministic-v0	Qbert-ramNoFrameskip-v0
FetchSlide-v1	GopherNoFrameskip-v0	BowlingDeterministic-v4	Qbert-ramNoFrameskip-v4
FetchPickAndPlace-v1	GopherNoFrameskip-v4	BowlingNoFrameskip-v0	Riverraid-v0
FetchReach-v1	Gopher-ram-v0	BowlingNoFrameskip-v4	Riverraid-v4
FetchPush-v1	Gopher-ram-v4	Bowling-ram-v0	RiverraidDeterministic-v0
HandReach-v0	Gopher-ramDeterministic-v0	Bowling-ram-v4	RiverraidDeterministic-v4
HandManipulateBlockRotateZ-v0	Gopher-ramDeterministic-v4	Bowling-ramDeterministic-v0	RiverraidNoFrameskip-v0
HandManipulateBlockRotateZTouchSensors-v0	Gopher-ramNoFrameskip-v0	Bowling-ramDeterministic-v4	RiverraidNoFrameskip-v4
HandManipulateBlockRotateZTouchSensors-v1	Gopher-ramNoFrameskip-v4	Bowling-ramNoFrameskip-v0	Riverraid-ram-v0
HandManipulateBlockRotateParallel-v0	Gravitar-v0	Bowling-ramNoFrameskip-v4	Riverraid-ram-v4
HandManipulateBlockRotateParallelTouchSensors-v0	Gravitar-v4	Boxing-v0	Riverraid-ramDeterministic-v0
HandManipulateBlockRotateParallelTouchSensors-v1	GravitarDeterministic-v0	Boxing-v4	Riverraid-ramDeterministic-v4
HandManipulateBlockRotateXYZ-v0	GravitarDeterministic-v4	BoxingDeterministic-v0	Riverraid-ramNoFrameskip-v0

RepeatCopy-v0	Enduro-ram-v0	BankHeistNoFrameskip-v4	Pong-v4
HandManipulateBlockRotateXYZTouchSensors-v0	GravitarNoFrameskip-v0	BoxingDeterministic-v4	Riverraid-ramNoFrameskip-v4
HandManipulateBlockRotateXYZTouchSensors-v1	GravitarNoFrameskip-v4	BoxingNoFrameskip-v0	RoadRunner-v0
HandManipulateBlockFull-v0	Gravitar-ram-v0	BoxingNoFrameskip-v4	RoadRunner-v4
HandManipulateBlock-v0	Gravitar-ram-v4	Boxing-ram-v0	RoadRunnerDeterministic-v0
HandManipulateBlockTouchSensors-v0	Gravitar-ramDeterministic-v0	Boxing-ram-v4	RoadRunnerDeterministic-v4
HandManipulateBlockTouchSensors-v1	Gravitar-ramDeterministic-v4	Boxing-ramDeterministic-v0	RoadRunnerNoFrameskip-v0
HandManipulateEggRotate-v0	Gravitar-ramNoFrameskip-v0	Boxing-ramDeterministic-v4	RoadRunnerNoFrameskip-v4
HandManipulateEggRotateTouchSensors-v0	Gravitar-ramNoFrameskip-v4	Boxing-ramNoFrameskip-v0	RoadRunner-ram-v0
HandManipulateEggRotateTouchSensors-v1	Hero-v0	Boxing-ramNoFrameskip-v4	RoadRunner-ram-v4
HandManipulateEggFull-v0	Hero-v4	Breakout-v0	RoadRunner-ramDeterministic-v0
HandManipulateEgg-v0	HeroDeterministic-v0	Breakout-v4	RoadRunner-ramDeterministic-v4
HandManipulateEggTouchSensors-v0	HeroDeterministic-v4	BreakoutDeterministic-v0	RoadRunner-ramNoFrameskip-v0
HandManipulateEggTouchSensors-v1	HeroNoFrameskip-v0	BreakoutDeterministic-v4	RoadRunner-ramNoFrameskip-v4
HandManipulatePenRotate-v0	HeroNoFrameskip-v4	BreakoutNoFrameskip-v0	Robotank-v0
HandManipulatePenRotateTouchSensors-v0	Hero-ram-v0	BreakoutNoFrameskip-v4	Robotank-v4
HandManipulatePenRotateTouchSensors-v1	Hero-ram-v4	Breakout-ram-v0	RobotankDeterministic-v0
HandManipulatePenFull-v0	Hero-ramDeterministic-v0	Breakout-ram-v4	RobotankDeterministic-v4
HandManipulatePen-v0	Hero-ramDeterministic-v4	Breakout-ramDeterministic-v0	RobotankNoFrameskip-v0
HandManipulatePenTouchSensors-v0	Hero-ramNoFrameskip-v0	Breakout-ramDeterministic-v4	RobotankNoFrameskip-v4
HandManipulatePenTouchSensors-v1	Hero-ramNoFrameskip-v4	Breakout-ramNoFrameskip-v0	Robotank-ram-v0
FetchSlideDense-v1	IceHockey-v0	Breakout-ramNoFrameskip-v4	Robotank-ram-v4
FetchPickAndPlaceDense-v1	IceHockey-v4	Carnival-v0	Robotank-ramDeterministic-v0
FetchReachDense-v1	IceHockeyDeterministic-v0	Carnival-v4	Robotank-ramDeterministic-v4
FetchPushDense-v1	IceHockeyDeterministic-v4	CarnivalDeterministic-v0	Robotank-ramNoFrameskip-v0
HandReachDense-v0	IceHockeyNoFrameskip-v0	CarnivalDeterministic-v4	Robotank-ramNoFrameskip-v4
HandManipulateBlockRotateZDense-v0	IceHockeyNoFrameskip-v4	CarnivalNoFrameskip-v0	Sequest-v0
HandManipulateBlockRotateZTouchSensorsDense-v0	IceHockey-ram-v0	CarnivalNoFrameskip-v4	Sequest-v4
HandManipulateBlockRotateZTouchSensorsDense-v1	IceHockey-ram-v4	Carnival-ram-v0	SequestDeterministic-v0
HandManipulateBlockRotateParallelDense-v0	IceHockey-ramDeterministic-v0	Carnival-ram-v4	SequestDeterministic-v4
HandManipulateBlockRotateParallelTouchSensorsDense-v0	IceHockey-ramDeterministic-v4	Carnival-ramDeterministic-v0	SequestNoFrameskip-v0
HandManipulateBlockRotateParallelTouchSensorsDense-v1	IceHockey-ramNoFrameskip-v0	Carnival-ramDeterministic-v4	SequestNoFrameskip-v4
HandManipulateBlockRotateXYZDense-v0	IceHockey-ramNoFrameskip-v4	Carnival-ramNoFrameskip-v0	Sequest-ram-v0
HandManipulateBlockRotateXYZTouchSensorsDense-v0	Jamesbond-v0	Carnival-ramNoFrameskip-v4	Sequest-ram-v4
HandManipulateBlockRotateXYZTouchSensorsDense-v1	Jamesbond-v4	Centipede-v0	Sequest-ramDeterministic-v0
HandManipulateBlockFullDense-v0	JamesbondDeterministic-v0	Centipede-v4	Sequest-ramDeterministic-v4
HandManipulateBlockDense-v0	JamesbondDeterministic-v4	CentipedeDeterministic-v0	Sequest-ramNoFrameskip-v0
HandManipulateBlockTouchSensorsDense-v0	JamesbondNoFrameskip-v0	CentipedeDeterministic-v4	Sequest-ramNoFrameskip-v4
HandManipulateBlockTouchSensorsDense-v1	JamesbondNoFrameskip-v4	CentipedeNoFrameskip-v0	Skiing-v0
HandManipulateEggRotateDense-v0	Jamesbond-ram-v0	CentipedeNoFrameskip-v4	Skiing-v4
HandManipulateEggRotateTouchSensorsDense-v0	Jamesbond-ram-v4	Centipede-ram-v0	SkiingDeterministic-v0
HandManipulateEggRotateTouchSensorsDense-v1	Jamesbond-ramDeterministic-v0	Centipede-ram-v4	SkiingDeterministic-v4
HandManipulateEggFullDense-v0	Jamesbond-ramDeterministic-v4	Centipede-ramDeterministic-v0	SkiingNoFrameskip-v0
HandManipulateEggDense-v0	Jamesbond-ramNoFrameskip-v0	Centipede-ramDeterministic-v4	SkiingNoFrameskip-v4
HandManipulateEggTouchSensorsDense-v0	Jamesbond-ramNoFrameskip-v4	Centipede-ramNoFrameskip-v0	Skiing-ram-v0
HandManipulateEggTouchSensorsDense-v1	JourneyEscape-v0	Centipede-ramNoFrameskip-v4	Skiing-ram-v4
HandManipulatePenRotateDense-v0	JourneyEscape-v4	ChopperCommand-v0	Skiing-ramDeterministic-v0
HandManipulatePenRotateTouchSensorsDense-v0	JourneyEscapeDeterministic-v0	ChopperCommand-v4	Skiing-ramDeterministic-v4
HandManipulatePenRotateTouchSensorsDense-v1	JourneyEscapeDeterministic-v4	ChopperCommandDeterministic-v0	Skiing-ramNoFrameskip-v0
HandManipulatePenFullDense-v0	JourneyEscapeNoFrameskip-v0	ChopperCommandDeterministic-v4	Skiing-ramNoFrameskip-v4
HandManipulatePenDense-v0	JourneyEscapeNoFrameskip-v4	ChopperCommandNoFrameskip-v0	Solaris-v0
HandManipulatePenTouchSensorsDense-v0	JourneyEscape-ram-v0	ChopperCommandNoFrameskip-v4	Solaris-v4
HandManipulatePenTouchSensorsDense-v1	JourneyEscape-ram-v4	ChopperCommand-ram-v0	SolarisDeterministic-v0
Adventure-v0	JourneyEscape-ramDeterministic-v0	ChopperCommand-ram-v4	SolarisDeterministic-v4

RepeatCopy-v0	Enduro-ram-v0	BankHeistNoFrameskip-v4	Pong-v4
Adventure-v4	JourneyEscape-ramDeterministic-v4	ChopperCommand-ramDeterministic-v0	SolarisNoFrameskip-v0
AdventureDeterministic-v0	JourneyEscape-ramNoFrameskip-v0	ChopperCommand-ramDeterministic-v4	SolarisNoFrameskip-v4
AdventureDeterministic-v4	JourneyEscape-ramNoFrameskip-v4	ChopperCommand-ramNoFrameskip-v0	Solaris-ram-v0
AdventureNoFrameskip-v0	Kangaroo-v0	ChopperCommand-ramNoFrameskip-v4	Solaris-ram-v4
AdventureNoFrameskip-v4	Kangaroo-v4	CrazyClimber-v0	Solaris-ramDeterministic-v0
Adventure-ram-v0	KangarooDeterministic-v0	CrazyClimber-v4	Solaris-ramDeterministic-v4
Adventure-ram-v4	KangarooDeterministic-v4	CrazyClimberDeterministic-v0	Solaris-ramNoFrameskip-v0
Adventure-ramDeterministic-v0	KangarooNoFrameskip-v0	CrazyClimberDeterministic-v4	Solaris-ramNoFrameskip-v4
Adventure-ramDeterministic-v4	KangarooNoFrameskip-v4	CrazyClimberNoFrameskip-v0	SpaceInvaders-v0
Adventure-ramNoFrameskip-v0	Kangaroo-ram-v0	CrazyClimberNoFrameskip-v4	SpaceInvaders-v4
Adventure-ramNoFrameskip-v4	Kangaroo-ram-v4	CrazyClimber-ram-v0	SpaceInvadersDeterministic-v0
AirRaid-v0	Kangaroo-ramDeterministic-v0	CrazyClimber-ram-v4	SpaceInvadersDeterministic-v4
AirRaid-v4	Kangaroo-ramDeterministic-v4	CrazyClimber-ramDeterministic-v0	SpaceInvadersNoFrameskip-v0
AirRaidDeterministic-v0	Kangaroo-ramNoFrameskip-v0	CrazyClimber-ramDeterministic-v4	SpaceInvadersNoFrameskip-v4
AirRaidDeterministic-v4	Kangaroo-ramNoFrameskip-v4	CrazyClimber-ramNoFrameskip-v0	SpaceInvaders-ram-v0
AirRaidNoFrameskip-v0	Krull-v0	CrazyClimber-ramNoFrameskip-v4	SpaceInvaders-ram-v4
AirRaidNoFrameskip-v4	Krull-v4	Defender-v0	SpaceInvaders-ramDeterministic-v0
AirRaid-ram-v0	KrullDeterministic-v0	Defender-v4	SpaceInvaders-ramDeterministic-v4
AirRaid-ram-v4	KrullDeterministic-v4	DefenderDeterministic-v0	SpaceInvaders-ramNoFrameskip-v0
AirRaid-ramDeterministic-v0	KrullNoFrameskip-v0	DefenderDeterministic-v4	SpaceInvaders-ramNoFrameskip-v4
AirRaid-ramDeterministic-v4	KrullNoFrameskip-v4	DefenderNoFrameskip-v0	StarGunner-v0
AirRaid-ramNoFrameskip-v0	Krull-ram-v0	DefenderNoFrameskip-v4	StarGunner-v4
AirRaid-ramNoFrameskip-v4	Krull-ram-v4	Defender-ram-v0	StarGunnerDeterministic-v0
Alien-v0	Krull-ramDeterministic-v0	Defender-ram-v4	StarGunnerDeterministic-v4
Alien-v4	Krull-ramDeterministic-v4	Defender-ramDeterministic-v0	StarGunnerNoFrameskip-v0
AlienDeterministic-v0	Krull-ramNoFrameskip-v0	Defender-ramDeterministic-v4	StarGunnerNoFrameskip-v4
AlienDeterministic-v4	Krull-ramNoFrameskip-v4	Defender-ramNoFrameskip-v0	StarGunner-ram-v0
AlienNoFrameskip-v0	KungFuMaster-v0	Defender-ramNoFrameskip-v4	StarGunner-ram-v4
AlienNoFrameskip-v4	KungFuMaster-v4	DemonAttack-v0	StarGunner-ramDeterministic-v0
Alien-ram-v0	KungFuMasterDeterministic-v0	DemonAttack-v4	StarGunner-ramDeterministic-v4
Alien-ram-v4	KungFuMasterDeterministic-v4	DemonAttackDeterministic-v0	StarGunner-ramNoFrameskip-v0
Alien-ramDeterministic-v0	KungFuMasterNoFrameskip-v0	DemonAttackDeterministic-v4	StarGunner-ramNoFrameskip-v4
Alien-ramDeterministic-v4	KungFuMasterNoFrameskip-v4	DemonAttackNoFrameskip-v0	Tennis-v0
Alien-ramNoFrameskip-v0	KungFuMaster-ram-v0	DemonAttackNoFrameskip-v4	Tennis-v4
Alien-ramNoFrameskip-v4	KungFuMaster-ram-v4	DemonAttack-ram-v0	TennisDeterministic-v0
Amidar-v0	KungFuMaster-ramDeterministic-v0	DemonAttack-ram-v4	TennisDeterministic-v4
Amidar-v4	KungFuMaster-ramDeterministic-v4	DemonAttack-ramDeterministic-v0	TennisNoFrameskip-v0
AmidarDeterministic-v0	KungFuMaster-ramNoFrameskip-v0	DemonAttack-ramDeterministic-v4	TennisNoFrameskip-v4
AmidarDeterministic-v4	KungFuMaster-ramNoFrameskip-v4	DemonAttack-ramNoFrameskip-v0	Tennis-ram-v0
AmidarNoFrameskip-v0	MontezumaRevenge-v0	DemonAttack-ramNoFrameskip-v4	Tennis-ram-v4
AmidarNoFrameskip-v4	MontezumaRevenge-v4	DoubleDunk-v0	Tennis-ramDeterministic-v0
Amidar-ram-v0	MontezumaRevengeDeterministic-v0	DoubleDunk-v4	Tennis-ramDeterministic-v4
Amidar-ram-v4	MontezumaRevengeDeterministic-v4	DoubleDunkDeterministic-v0	Tennis-ramNoFrameskip-v0
Amidar-ramDeterministic-v0	MontezumaRevengeNoFrameskip-v0	DoubleDunkDeterministic-v4	Tennis-ramNoFrameskip-v4
Amidar-ramDeterministic-v4	MontezumaRevengeNoFrameskip-v4	DoubleDunkNoFrameskip-v0	TimePilot-v0
Amidar-ramNoFrameskip-v0	MontezumaRevenge-ram-v0	DoubleDunkNoFrameskip-v4	TimePilot-v4
Amidar-ramNoFrameskip-v4	MontezumaRevenge-ram-v4	DoubleDunk-ram-v0	TimePilotDeterministic-v0
Assault-v0	MontezumaRevenge-ramDeterministic-v0	DoubleDunk-ram-v4	TimePilotDeterministic-v4
Assault-v4	MontezumaRevenge-ramDeterministic-v4	DoubleDunk-ramDeterministic-v0	TimePilotNoFrameskip-v0
AssaultDeterministic-v0	MontezumaRevenge-ramNoFrameskip-v0	DoubleDunk-ramDeterministic-v4	TimePilotNoFrameskip-v4

RepeatCopy-v0	Enduro-ram-v0	BankHeistNoFrameskip-v4	Pong-v4
AssaultDeterministic-v4	MontezumaRevenge-ramNoFrameskip-v4	DoubleDunk-ramNoFrameskip-v0	TimePilot-ram-v0
AssaultNoFrameskip-v0	MsPacman-v0	DoubleDunk-ramNoFrameskip-v4	TimePilot-ram-v4
AssaultNoFrameskip-v4	MsPacman-v4	ElevatorAction-v0	TimePilot-ramDeterministic-v0
Assault-ram-v0	MsPacmanDeterministic-v0	ElevatorAction-v4	TimePilot-ramDeterministic-v4
Assault-ram-v4	MsPacmanDeterministic-v4	ElevatorActionDeterministic-v0	TimePilot-ramNoFrameskip-v0
Assault-ramDeterministic-v0	MsPacmanNoFrameskip-v0	ElevatorActionDeterministic-v4	TimePilot-ramNoFrameskip-v4
Assault-ramDeterministic-v4	MsPacmanNoFrameskip-v4	ElevatorActionNoFrameskip-v0	Tutankham-v0
Assault-ramNoFrameskip-v0	MsPacman-ram-v0	ElevatorActionNoFrameskip-v4	Tutankham-v4
Assault-ramNoFrameskip-v4	MsPacman-ram-v4	ElevatorAction-ram-v0	TutankhamDeterministic-v0
Asterix-v0	MsPacman-ramDeterministic-v0	ElevatorAction-ram-v4	TutankhamDeterministic-v4
Asterix-v4	MsPacman-ramDeterministic-v4	ElevatorAction-ramDeterministic-v0	TutankhamNoFrameskip-v0
AsterixDeterministic-v0	MsPacman-ramNoFrameskip-v0	ElevatorAction-ramDeterministic-v4	TutankhamNoFrameskip-v4
AsterixDeterministic-v4	MsPacman-ramNoFrameskip-v4	ElevatorAction-ramNoFrameskip-v0	Tutankham-ram-v0
AsterixNoFrameskip-v0	NameThisGame-v0	ElevatorAction-ramNoFrameskip-v4	Tutankham-ram-v4
AsterixNoFrameskip-v4	NameThisGame-v4	Enduro-v0	Tutankham-ramDeterministic-v0
Asterix-ram-v0	NameThisGameDeterministic-v0	Enduro-v4	Tutankham-ramDeterministic-v4
Asterix-ram-v4	NameThisGameDeterministic-v4	EnduroDeterministic-v0	Tutankham-ramNoFrameskip-v0
Asterix-ramDeterministic-v0	NameThisGameNoFrameskip-v0	EnduroDeterministic-v4	Tutankham-ramNoFrameskip-v4
Asterix-ramDeterministic-v4	NameThisGameNoFrameskip-v4	EnduroNoFrameskip-v0	UpNDown-v0
Asterix-ramNoFrameskip-v0	NameThisGame-ram-v0	EnduroNoFrameskip-v4	UpNDown-v4
Asterix-ramNoFrameskip-v4	NameThisGame-ram-v4	Asteroids-ramNoFrameskip-v4	Phoenix-ram-v4
Asteroids-v0	NameThisGame-ramDeterministic-v0	Atlantis-v0	Phoenix-ramDeterministic-v0
Asteroids-v4	NameThisGame-ramDeterministic-v4	Atlantis-v4	Phoenix-ramDeterministic-v4
AsteroidsDeterministic-v0	NameThisGame-ramNoFrameskip-v0	AtlantisDeterministic-v0	Phoenix-ramNoFrameskip-v0
AsteroidsDeterministic-v4	NameThisGame-ramNoFrameskip-v4	AtlantisDeterministic-v4	Phoenix-ramNoFrameskip-v4
AsteroidsNoFrameskip-v0	Phoenix-v0	AtlantisNoFrameskip-v0	Pitfall-v0
AsteroidsNoFrameskip-v4	Phoenix-v4	AtlantisNoFrameskip-v4	Pitfall-v4
Asteroids-ram-v0	PhoenixDeterministic-v0	Atlantis-ram-v0	PitfallDeterministic-v0
Asteroids-ram-v4	PhoenixDeterministic-v4	Atlantis-ram-v4	PitfallDeterministic-v4
Asteroids-ramDeterministic-v0	PhoenixNoFrameskip-v0	Atlantis-ramDeterministic-v0	PitfallNoFrameskip-v0
Asteroids-ramDeterministic-v4	PhoenixNoFrameskip-v4	Atlantis-ramDeterministic-v4	PitfallNoFrameskip-v4
Asteroids-ramNoFrameskip-v0	Phoenix-ram-v0	Atlantis-ramNoFrameskip-v0	Pitfall-ram-v0
BankHeistDeterministic-v4	Pitfall-ramNoFrameskip-v4	Atlantis-ramNoFrameskip-v4	Pitfall-ram-v4
BankHeistNoFrameskip-v0	Pong-v0	BankHeist-v0	Pitfall-ramDeterministic-v0
BankHeistDeterministic-v0	Pitfall-ramNoFrameskip-v0	BankHeist-v4	Pitfall-ramDeterministic-v4