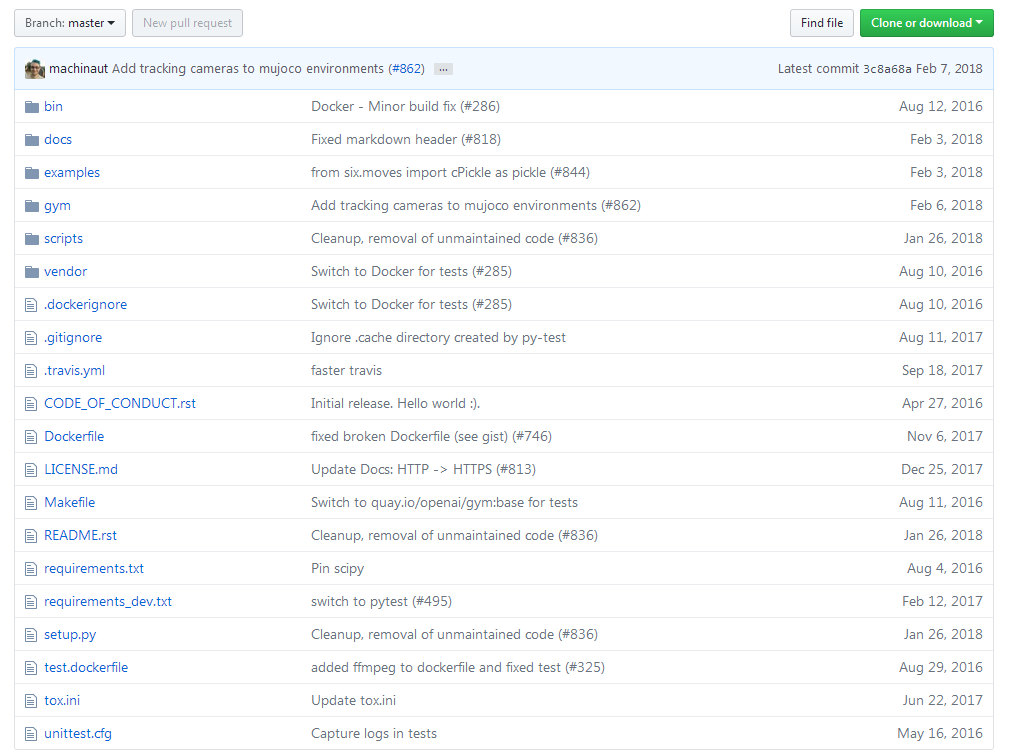
**OpenAI Gym安装**

**由来**

2015年12月，特斯拉CEO埃隆·马斯克（Elon Musk）和创业孵化器Y Combinator总裁山姆·奥特曼（Sam Altman）创建了人工智能公司OpenAI，并将其研究成果开源分享。

2016年4月，Open AI对外发布了人工智能一款用于研发强化学习算法的工具包OpenAI Gym。GitHub上的OpenAI Gym源码：



**OpenAI**

OpenAI是着眼于研究非监督式学习、强化学习、深度强化学习的人工智能研究团队。

**OpenAI Gym**

OpenAI Gym是OpenAI推出的一款用于机器学习算法的工具包，主要用于非监督式学习、强化学习、深度强化学习。它与很多数值计算库兼容，比如tensorflow和theano。目前支持的开发语言是Python，原生支持Linux和Mac，刚支持Windows不久。

**强化学习**

强化学习是机器学习的一个分支，它考虑的是做出一系列的决策。它假定有一个智能体（agent）存在于环境中。在每一步中，智能体（agent）采取一个行动，随后从环境中收到观察与回报。一个强化学习算法寻求的是，在一个原先毫无了解的环境中通过一段学习过程——通常包括许多试错——让智能体（agent）收到的总体回报最大化。

强化学习关注的是做出好决策，与监督式学习和非监督式学习之间有大量相通之处，除了不同的侧重点之外。在强化学习中，探索成为必须——智能体（agent）需要不停地进入可能会收获大量回报的未知领域。

**神经网络+强化学习**

一般来说强化学习算法关注的是学习一个策略（policy）。鉴于我们通常都无法查知世界的整体状态（full state），我们一般会使用一个或是更多过去的观察来作为替代。以Q函数为例，Q函数（估值函数的一种）衡量的是状态-行动组(s, a)的好坏，也就是说，Q(s, a)能告诉你“如果我处于状态s中并选择行动a，我能获得多少回报”。有了这个Q函数以后，你就能简单地选择出带来最高预期回报的行动。这也就是说，Q函数定义了策略。

我们可以用神经网络来表示策略和Q函数。例如，当玩游戏的时候，向这些网络输入的是屏幕上的一个图像，同时有一组离散的行动。你可以用一个卷积神经网络将屏幕图像作为输入并输出一个代表四种行动之一的数字，表示出行动的好坏，作为这个任务的Q函数，并用一个结构相似、输出每种行动可能性的卷积神经网络作为策略。

**Ubuntu安装**

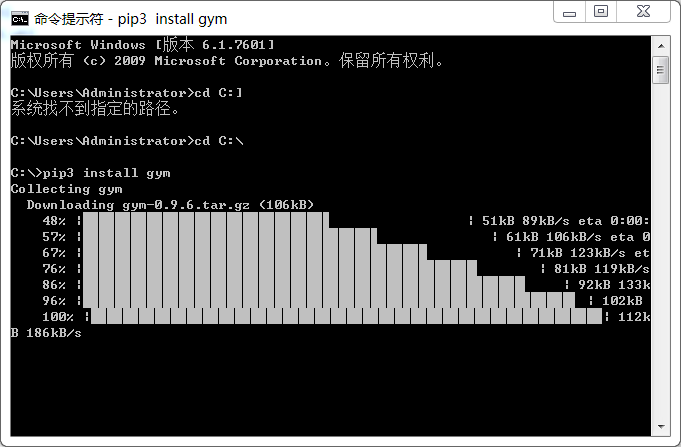
apt-get install -y python-numpy python-dev cmake zlib1g-dev libjpeg-dev xvfb libav-tools xorg-dev python-opengl libboost-all-dev libsdl2-dev swig

**Windows安装**

（1）使用pip安装：（需联网）

pip3 install gym（5分钟）

pip3 install gym[all]（报错，需make支持、需Visual Studio 2015(VC14)支持）



还有一些常用的库numpy库、pyglet库如果未装或旧的需要装一下

（2）使用源码安装：

https://github.com/openai/gym

cd gym

pip install -e . # minimal install

pip install -e .[all] # full install **(This requires cmake and a recent pip version)**

**Python安装**

目前只支持Python 2.7 or Python 3.5