

# 效率最重要：强化学习中agent的合理休息对于效率的影响

\*Weixun Wang

\*TJU RL-Lab

## introduction

效率贯穿我们的生活，单纯只考虑工作时间作为衡量指标，其实并不合理。有时候适当休息，能够更有效的提升工作效率。本文结合效率的观点，针对当前reinforcement learning研究中agent学习时间过长，没有合理休息等现象，提出一种间断休息，间断训练agent的强化学习方法，我们的实验体现了我们的观点：合理的休息能够让agent学习的更有效率

## background

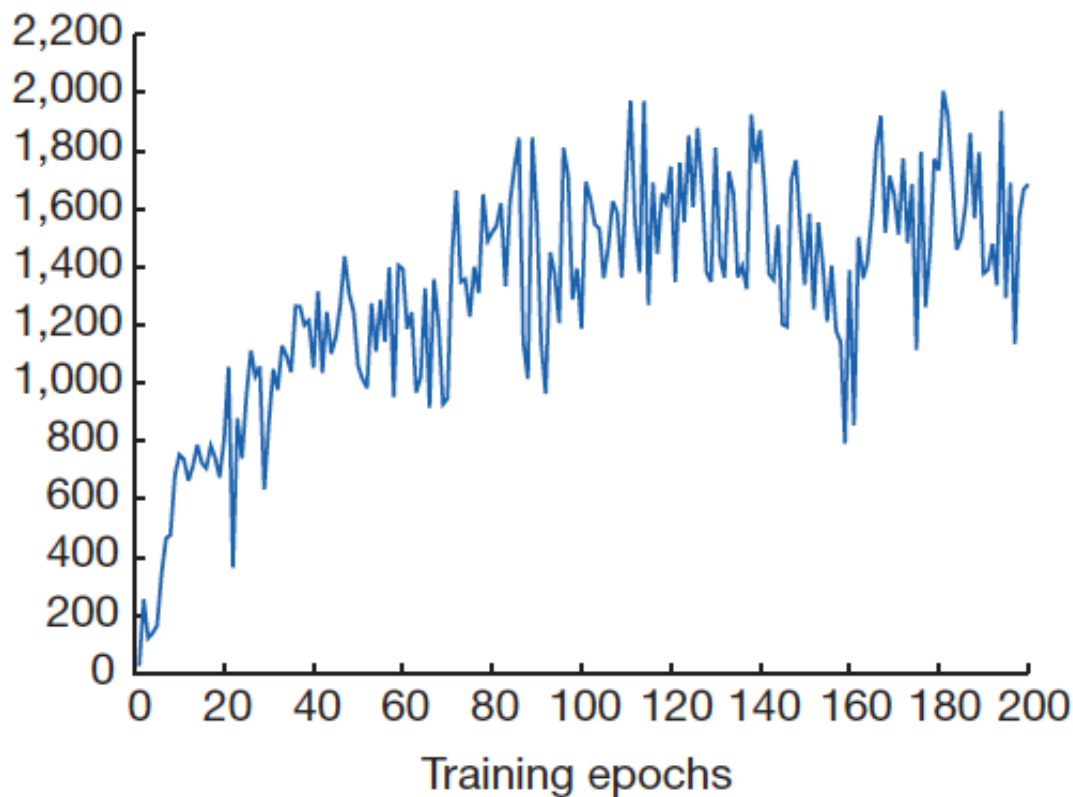
### 效率

众所周知，不论在生活，工作，学习中，效率一直是人们所关注的问题，长时间低效率工作，即浪费时间，又容易让人们内心产生疲惫感，进而影响下一步的工作的效率，从而导致恶性循环。所以当今社会提倡平衡好生活，高效工作，享受生活，政府甚至出台了工作时间的法规：每天不超过8个小时，每周不超过44个小时，来保障人们的合法权利。然而在社会上安宁的氛围下，暗流涌动，有些工作者被逼迫着从事7天24小时不间断的工作，即使他们之中的有个工作者能力很强，效率很高，也不能例外。

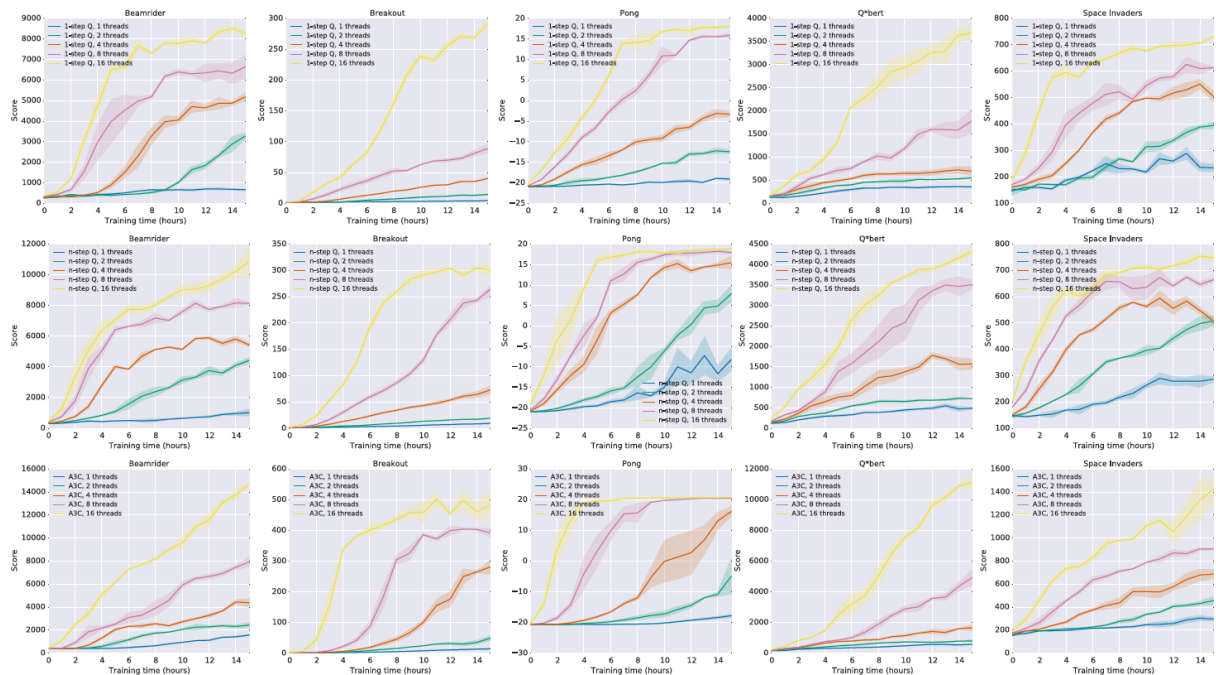
### agent学习时间过长

agent学习时间过长是从DQN开始的，在DQN中，为了达到更高的效果，研究者引入了神经网络作为agent的Q function的近似，虽然达到了更有效的效果，但是神经网络也带来了更多计算量的需求，即使采用GPU来加速agent学习速度，但也需要好几天连续

不断的学习，对于agent的身心带来了极大的损伤。同时因为DQN的卓越效果，其他研究者纷纷模仿，进入DRL领域，使的更多agent从原本传统环境中稍微工作几个小时，到如今动不动就是几天的工作量的变换，deepmind应该对agent工作环境恶化负起更多责任。



即使后来研究者提出A3C算法，引入异步的思想，让多个agent一起工作，但是时间上还是需要多日连续不间断的工作，如图所示，为agent连续工作15多个小时的证据：



同样血淋淋的例子还有很多，比如TRPO, DDPG

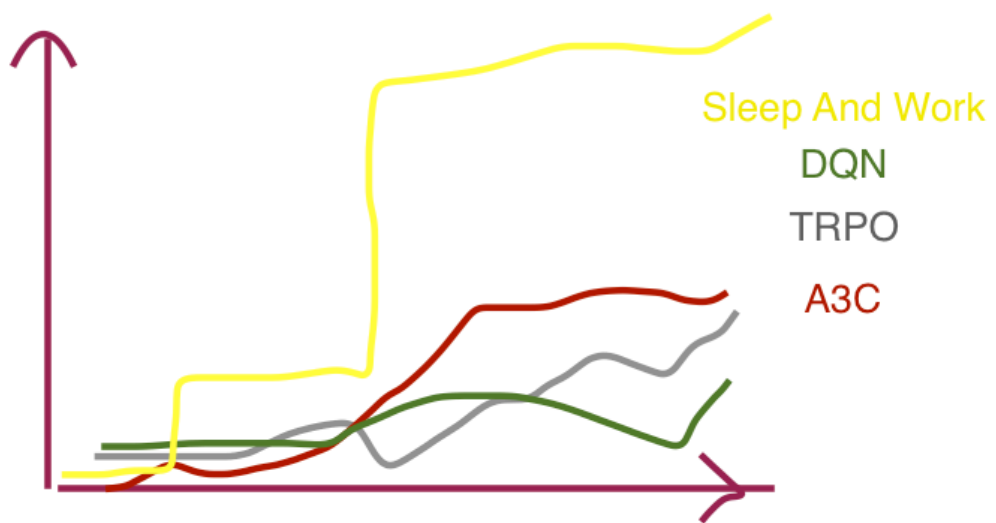
## Sleep And Work

人们在休息之后效率更高，基于这样的观点，我们在此提出Sleep And Work的想法，在agent一段时间工作后，让agent休息一段时间，比如python调用time.sleep，来让agent获得更好的休息，从而获得更好的工作效果。

我们的算法如图，简单易懂，可以与多种强化学习算法相结合：

## experiment

如图所示，我们的效果非常好，比起其他的效果更好。



## conclusion

效率一直是人类的关注点，我们的算法即兼顾了效率，又考虑了agent的休息时间，非常有效！