***DRL***

***参考：***

***Deep Reinforcement Learning Based Green Resource Allocation Mechanism in Mobile Edge Network for Ubiquitous Power IoT***

基于Mobile edge computing 的 Green resource allocation资源消耗过大，DRL可以更好节约资源消耗和更好地分配绿色资源。

以最小化能源效率为目标，同时满足每个移动用户的需求，DRL定义了状态空间、行动空间和奖励功能

状态空间、行动空间、奖励功能具体定义

算法1

对比了在三种策略（DA、UC、DRL）下的能源消耗

Summary: Minimizing energy efficiency while ensuring the needs of each user. A well-trained network can achieve the goal of green energy-saving, also solve the effective resource allocation problem, and the results achieved the lowest energy efficiency compared with the other two strategies.

***参考：***

***Efficient Adaptive Resource Provisioning for Cloud Applications using Reinforcement Learning***

Reinforcement learning (RL) is a natural solution for adaptive decision making problems such as the one we have described.

强化学习(RL)是适应性决策问题的自然解决方案

RL based solutions are application agnostic which means that different applications can use the same algorithm without customized settings or parameters.

基于RL的解决方案与应用程序无关，这意味着不同的应用程序可以使用相同的算法，而无需定制设置或参数

While several such solutions exist already [1], they suffer from issues such as slow convergence, non scalability and poor performance during the learning period.

虽然已经存在几个这样的解决方案（RL)，但它们存在诸如收敛速度慢、不可扩展性和学习期间性能不佳等问题。

Another issue is the large amount of time taken for training and poor performance of the algorithm during the training phase.

另一个问题是训练时间长，算法在训练阶段表现不佳。

**DRL是解决以上DL问题的方法之一**

Deep Reinforcement learning (DRL) is an emerging area of research that approximates the Q-values using a deep neural network.

深度强化学习(DRL)是一个新兴的研究领域，使用深度神经网络近似q值。

**Q学习还分为快速Q学习和常规Q学习**

Speedy Q-learning converged faster, but the final cost obtained was slightly higher than what regular Q-learning gave.

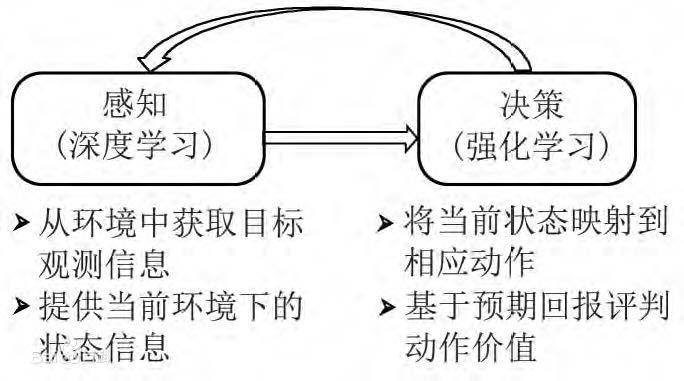
**快速q学习**收敛更快，但最终获得的成本略高于**常规q学习**给出的成本。

***为什么使用DRL***

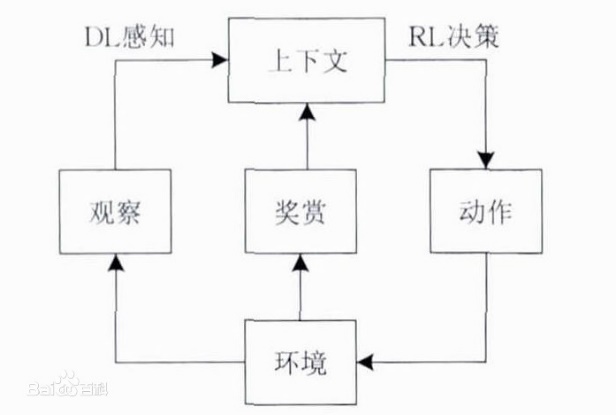
***参考：***

***赵冬斌,邵坤,朱圆恒,李栋,陈亚冉,王海涛,刘德荣,周彤,王成红.深度强化学习综述:兼论计算机围棋的发展[J].控制理论与应用,2016,33(06):701-717.***

***刘全,翟建伟,章宗长,钟珊,周倩,章鹏,徐进.深度强化学习综述[J].计算机学报,2018,41(01):1-27.***



深度学习具有较强的感知能力，但是缺乏一定的决策能力；而强化学习具有决策能力，对感知问题束手无策。因此，将两者结合起来，优势互补，为复杂系统的感知决策问题提供了解决思路。即深度强化学习以一种通用的形式将深度学习的感知能力与强化学习的决策能力相结合,并能够通过端对端的学习方式实现从原始输入到输出的直接控制。



DRL原理框图

DRL是一种端对端（end-to-end）的感知与控制系统，具有很强的通用性．其学习过程可以描述为：

(1)在每个时刻agent与环境交互得到一个高维度的观察，并利用DL方法来感知观察，以得到具体的状态特征表示；

(2)基于预期回报来评价各动作的价值函数，并通过某种策略将当前状态映射为相应的动作；

(3)环境对此动作做出反应，并得到下一个观察．通过不断循环以上过程，最终可以得到实现目标的最优策略。

***Q-Learning（DL中的一种）***

***简要思路：***

Q-Learning是强化学习算法中value-based的算法，Q即为Q（s,a）就是在某一时刻的 s 状态下(s∈S)，采取 动作a (a∈A)动作能够获得收益的期望，环境会根据agent的动作反馈相应的回报reward r，所以算法的主要思想就是将State与Action构建成一张Q-table来存储Q值，然后根据Q值来选取能够获得最大的收益的动作。