

# 基于值函数的深度强化学习算法研究进展

**摘要** 强化学习是当前机器学习的一个比较新的领域。跟监督学习和无监督学习相比，强化学习不是通过大量的静态数据学习一种映射关系，而是通过不断与环境进行交互，不断“试错”，动态的学习环境状态到所采取策略的映射关系。本文首先从强化学习的基本概率出发，主要以基于值函数的强化学习算法为主线，阐述深度强化学习的研究发展状况，包括各种深度强化学习算法的演进，以及应用前景，最后简要介绍其发展方向。

**关键词** 强化学习，Value-base，Q-learning

## 引言

受行为驱动学科的启发，强化学习是关于智能体与环境交互的一种学习方式。环境处于一定的状态下，智能体通过一种行为与环境产生互动，然后环境给与智能体奖励，如何最大化这种奖励就是强化学习要解决的问题。20世纪末，伴随着机器人应用的研究发展，强化学习也逐渐称为热门的研究领域。特别是近十几年来随着深度学习的研究，越来越多的研究与深度学习相结合取得了突破性的进展，如[hoho：Hinton的Reducing the dimensionality of data with neural networks]用RBM深度神经网络实现对图像的降维，取得比PCA[hoho: PCA引用]更好的效果，被誉为首次深度学习兴起的成功实践。强化学习也不例外，通过与深度学习的结合，解决了很多复杂的连续状态空间、连续动作空间的问题，为机器人、金融、生物工程等领域开创了新局面。

## 强化学习基本概念回顾

## 基于值函数的强化学习算法

## 应用与前景

## 总结