Reinforcment Learning : A Survey

L.P. Kaelbling, M. L. Littman, A.W. Moore

L.P. Kaelbling, M. L. Littman, A.W. Moore, « Reinforcement Learning : A Survey », Journal of Artificial Intelligence Research 4, 1996

Key words :

- Reinforcment Learning
- Q-Learning
- Exploitation vs Exploration
- Applications (Game – Robotics and Control)

Abstract :

This article speak about reinforcment learning. It descibed theorical operation of algorithms like TD(λ) or Q-Learning. It descibed also reinforcement learning applications in Games and Robotics control.

Q-Learning

Watkins' algorithm, Q-Learning, is a ameliration of the TD(λ) algorithm. It use a state/action function whitch will found the optimal policy π(s) for each state : Q(s,a).

Q(s,a) = ... [see on paper]

This equation come from the Bellman equation which use a Markov Decision Process (MDP)

It represent the sum of the hoped-for rewards from a state “s” if I choose to do the action “a”

Exploration vs Exploitation

One major difference between reinforcment learning and supervised learning is that a reinforcement learner must explore its environment.

If the intelligent agent only choose the action with the hiest Q-Value, it will pose a lot of problem:

- Tge sequence of actions may be not the optimal one
- If a add a new goal with a biggest reward, the agent may never found it

Use :

It will be use to make Q-Learning part and the Exploration vs exploitation part