

## A Novel Exploration/Exploitation Policy Accelerating Learning In Both Stationary And Non Stationary Environment Navigation Tasks

**Authors :** Wiem Zemzem, Moncef Tagina

**Abstract :** In this work, we are addressing the problem of an autonomous mobile robot navigating in a large, unknown and dynamic environment using reinforcement learning abilities. This problem is principally related to the exploration/exploitation dilemma, especially the need to find a solution letting the robot detect the environmental change and also learn in order to adapt to the new environmental form without ignoring knowledge already acquired. Firstly, a new action selection strategy, called  $\epsilon$ -greedy-MPA (the  $\epsilon$ -greedy policy favoring the most promising actions) is proposed. Unlike existing exploration/exploitation policies (EEPs) such as  $\epsilon$ -greedy and Boltzmann, the new EEP doesn't only rely on the information of the actual state but also uses those of the eventual next states. Secondly, as the environment is large, an exploration favoring least recently visited states is added to the proposed EEP in order to accelerate learning. Finally, various simulations with ball-catching problem have been conducted to evaluate the  $\epsilon$ -greedy-MPA policy. The results of simulated experiments show that combining this policy with the Qlearning method is more effective and efficient compared with the  $\epsilon$ -greedy policy in stationary environments and the utility-based reinforcement learning approach in non stationary environments.

**Keywords :** autonomous mobile robot, exploration/ exploitation policy, large, dynamic environment, reinforcement learning

**Conference Title :** ICMLA 2015 : International Conference on Machine Learning and Applications

**Conference Location :** Copenhagen, Denmark

**Conference Dates :** June 11-12, 2015