Learning a Robot to Score a Penalty: Minimal Reward Reinforcement Learning
C.G. Lagrand

Abstract

Reinforcement Learning is applied to learn a robot how to score a penalty in simulation. Minimal rewards are given to let the robot learn as much as possible by itself. Intermediate rewards are applied to reduce the amount of trials needed to learn a good action policy up to 75% without effecting the accuracy. Furthermore, different methods of combining object states (multiplying and additive) are examined to determine the method that uses less trials to learn. The additive method results in a decrease of 58% in the amount of trials. A robot is able to learn how to score a penalty against a standing goalkeeper within an average of 170 trials using all intermediate rewards, the additive state representation method and sixteen different states for both the ball and the goalkeeper, with an average accuracy of 69%.