

A Bayesian Approach for Learning and Tracking Switching, Non-Stationary Opponents (Extended Abstract)

Pablo Hernandez-Leal

Instituto Nacional de Astrofísica, Óptica y Electrónica Puebla, México
pablohl@ccc.inaoep.mx

Benjamin Rosman

Council for Scientific and Industrial Research, and the University of the Witwatersrand, South Africa
brosman@csir.co.za

Matthew E. Taylor

Washington State University, Pullman, Washington, USA
taylorm@eecs.wsu.edu

L. Enrique Sucar

Instituto Nacional de Astrofísica, Óptica y Electrónica Puebla, México
esucar@inaoep.mx

Enrique Munoz de Cote

Instituto Nacional de Astrofísica, Óptica y Electrónica Puebla, México
jemc@inaoep.mx

Abstract

In many situations, agents are required to use a set of strategies (behaviors) and switch among them during the course of an interaction. This work focuses on the problem of recognizing the strategy used by an agent within a small number of interactions. We propose using a Bayesian framework to address this problem. Bayesian policy reuse (BPR) has been empirically shown to be efficient at correctly detecting the best policy to use from a library in sequential decision tasks. In this paper we extend BPR to adversarial settings, in particular, to opponents that switch from one stationary strategy to another. Our proposed extension enables learning new models in an online fashion when the learning agent detects that the current policies are not performing optimally. Experiments presented in repeated games show that our approach is capable of efficiently detecting opponent strategies and reacting quickly to behavior switches, thereby yielding better performance than state-of-the-art approaches in terms of average rewards.