Simple Decentralized Algorithm for Coordination Games

Mihail Mihaylov MMIHAYLO@VUB.AC.BE
Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

Karl Tuyls K.TUYLS@MAASTRICHTUNIVERSITY.NL
Maastricht University, Sint Servaasklooster 39, 6200 MD Maastricht, The Netherlands

Ann Nowé ANN.NOWE@VUB.AC.BE
Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium

Keywords: decentralized coordination, learning algorithm, pure coordination games

Many biological or computer systems are comprised of intelligent, but highly constrained agents with common objectives that are beyond the capabilities of the individual. Often such multi-agent systems (MASs) are inherently decentralized and therefore agents need to coordinate their behavior in the absence of central control to achieve their design objectives. Wireless sensor networks (WSNs) are an example of a decentralized MAS where sensor nodes gather environmental data and collectively forward it towards the base station of the observer. The limited resources of such sensor nodes and the lack of global knowledge make the design of a WSN application challenging. The main question we are concerned with is the following: based only on local interactions and incomplete knowledge how can the designer of a decentralized system make agents achieve good collective performance imposing minimal system requirements and overhead?

In our work in progress we examine the decentralized coordination problem through repeated local interactions between agents arranged in different topologies. Pairs of neighboring agents are randomly assigned to play 2-by-2 pure coordination games, as done by Shoham and Tennenholtz (1997). The only information agents receive from each interaction is a payoff of 1 if they select the same action and 0 if their actions differ. Inspired by the challenges and constraints of WSNs, we would like to make memory-constrained agents learn to coordinate under limited (binary) feedback in as few interactions as possible.

We propose a simple decentralized algorithm that when adopted by individual agents leads them to global successful coordination. Our algorithm, called Win-Stay Lose-probabilistic-Shift (WSLpS) guides agents in selecting their actions in order to reach online a mutually beneficial outcome. Intuitively, if an agent receives payoff of 1 from an interaction ("win"), it means that its interaction partner has selected the same action. In that case it is reasonable that the agent will select the same action in the next time period ("stay"). A payoff of 0, on the other hand ("lose"), indicates that its neighbor has picked a different action. To avoid having both agents swap their actions upon conflict, each agent should change with a certain probability $\alpha \in (0, 1)$ ("probabilistic shift").

Our action selection algorithm is unique in that it is *very light*, requires *no memory* of previous interactions, given the current one, and quickly drives agents to *full coordination*. Moreover, coordination outcomes are *absorbing states* of the system and therefore agents will never change their actions once converged.

We carried out a Markov chain analysis of WSLpS and compared it to other algorithms applied in this domain in terms of convergence time. Initial experiments suggest that, besides being a much simpler algorithm and imposing virtually no system requirements, WSLpS outperforms Q-learning, highest cumulative reward rule (Shoham & Tennenholtz, 1997) and Win-Stay Lose-Randomize (Barrett & Zollman, 2009) in ring, scale-free and fully connected topologies.

References

Barrett, J., & Zollman, K. J. S. (2009). The role of forgetting in the evolution and learning of language. *Journal of Experimental & Theoretical Artificial Intelligence*, *21*, 293–309.

Shoham, Y., & Tennenholtz, M. (1997). On the emergence of social conventions : modeling, analysis, and simulations. *Artificial Intelligence*, *94*, 139–166.