Adaptive Agents on Evolving Networks (Extended Abstract)

Ardeshir Kianercy
USC Information Sciences Institute
Marina del Rey, CA, USA
kianercy@usc.edu

Aram Galstyan
USC Information Sciences Institute
Marina del Rey, CA, USA
galstyan@usc.edu

Armen Allahverdyan
Yerevan Physics Institute
Yerevan, Armenia
aarmen@usc.edu

ABSTRACT

We propose a model of strategic network formation in repeated games where players adopt actions and connections simultaneously using a simple reinforcement learning scheme. We demonstrate that under certain plausible assumptions the dynamics of such systems can be described by so called replicator equations that characterize the co-evolution of agent strategies and network topology. Within this framework, the network structures emerging as a result of the game-dynamical interactions are described by the stable rest points of the replicator dynamics. In particular, we show using both simulations and analytical methods that for certain N-agent games the stable equilibria consist of star motifs as the main building blocks of the network.

Categories and Subject Descriptors

I.2 [**Artificial intelligence**]: Distributed Artificial Intelligence

General Terms

Theory, Algorithms

Keywords

Strategic network formation, Q-learning, Evolutionary game theory

1. INTRODUCTION

Many complex systems can be represented as networks where nodes correspond to entities and links encode interdependencies between them. Generally, statistical models of networks can be classified into two different approaches. In the first approach, networks are modeled via active nodes with a given distribution of links, where each node of the network represents a dynamical system. In this setting, one usually studies problems related to epidemic spreading, opinion formation, signaling and synchronization and so on. In the second approach, which is grounded mainly in a graph-theoretical approach, nodes are treated as passive elements. Instead, the main focus is on dynamics of link formation and network growth. Specifically, one is interested in algorithmic methods to build graphs formed by passive elements (nodes) and links, which evolve according to pre-specified, often local rules. This approach has produced important results on topological features of social, technological and biological networks.

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), June, 4–8, 2012, Valencia, Spain.

More recently, however, it has been realized that modeling individual and network dynamics separately is too limited to capture realistic behavior of networks. Indeed, most real–world networks are inherently complex dynamical systems, where both attributes of individuals (nodes) and topology of the network (links) can have inter–coupled dynamics. For instance, it is known that in social networks, nodes tend to divide into groups, or communities, of like-minded individuals. One can ask whether individuals become like-minded because they are connected via the network, or whether they form network connections because they are like-minded. Clearly, the distinction between the two scenarios is not clear-cut. Rather, the real world self-organizes by a combination of the two, the network changing in response to opinion and opinion changing in response to the network. Recent research has focused on the interplay between attribute and link dynamics (e.g., see [2, 4, 1] for a recent survey of the literature).

Here we suggest a simple model of co–evolving network that is based on the notion of interacting adaptive agents. Specifically, we consider network–augmented multi–agent systems where agents play repeated game with their neighbors, and adapt both their behaviors and the network ties depending on the outcome of their interactions. To adapt, agents use a simple learning mechanism to reinforce (punish) behaviors and network links that produce favorable (unfavorable) outcomes. Thus, the agent strategies and network topology evolve in tandem. We show that the collective evolution of such a system can be described by appropriately defined replicator dynamics equations. Originally suggested in the context of evolutionary game theory (e.g., see [3]), replicator equations have been used to model collective learning and adaptation in systems of interacting self–interested agents [5].

2. MODEL

Let us consider a set of agents that play repeated games with each other. We differentiate agents by indices $x, y, z, \ldots$. The time–dependent mixed strategies of agents can be characterized by a probability distribution over the choice of the neighbors and the actions. For instance, $p^i_{xy}(t)$ is the probability that the agent x will choose to play with agent y and perform action i at time t.

Furthermore, we assume that the agents adapt to their environment through a simple reinforcement mechanism. Among different reinforcement schemes, here we focus on (stateless) Q-learning [6]. In this case, it is known that the evolution of the agent strategies is governed by so called replicator equation [5]:

$$\frac{\dot{p}^i_{xy}}{p^i_{xy}} = \sum_j A^{ij}_{xy} p^j_{yx} - \sum_{i,j,\tilde{y}} A^{ij}_{x\tilde{y}} p^i_{x\tilde{y}} p^j_{\tilde{y}x} + T \sum_{\tilde{y},j} p^j_{x\tilde{y}} \ln \frac{p^j_{x\tilde{y}}}{p^i_{xy}} \quad (1)$$

We now make the assumption that the agents' strategies can be factorized as follows:

$$p^i_{xy} = c_{xy} p^i_x \, , \; \sum\nolimits_y c_{xy} = 1, \; \sum\nolimits_i p^i_x = 1. \quad (2)$$

Here p^i_x is the probability for agent x to select action i, whereas c_{xy} characterizes the strength of the directed link $x \to y$. Note that generally links are asymmetric, $c_{xy} \neq c_{yx}$.

Substituting 2 in 1, then taking summation of both sides in the resulting equation, once over y and then over i, we obtain

$$\begin{aligned} \frac{\dot{p}^i_x}{p^i_x} &= \sum_{\tilde{y},j} A^{ij}_{x\tilde{y}} c_{x\tilde{y}} c_{\tilde{y}x} p^j_{\tilde{y}} - \sum_{i,j,\tilde{y}} A^{ij}_{x\tilde{y}} c_{x\tilde{y}} c_{\tilde{y}x} p^i_x p^j_{\tilde{y}} \\ &+ T \sum_j p^j_x \ln(p^j_x / p^i_x) \qquad (3) \\ \frac{\dot{c}_{xy}}{c_{xy}} &= c_{yx} \sum_{i,j} A^{ij}_{xy} p^i_x p^j_y - \sum_{i,j,\tilde{y}} A^{ij}_{x\tilde{y}} c_{x\tilde{y}} c_{\tilde{y}x} p^i_x p^j_{\tilde{y}} \\ &+ T \sum_{\tilde{y}} c_{x\tilde{y}} \ln(c_{x\tilde{y}} / c_{xy}) \qquad (4) \end{aligned}$$

Equations 3, 4 describe the mutual evolution of the agents' strategies and the network structure. Here we focus on the case $T = 0$.

We should note that generally, the replicator dynamics (and Nash equilibria) in matrix games are invariant with respect to adding any column vector to the payoff matrix. However, this invariance does not hold in the present networked game. The reason for this is the following: if an agent does not have any incoming links (i.e., no other agent plays with him/her), then he always gets a zero reward. This poses a certain problem. For instance, consider the game of Prisoner's Dilemma where the payoff for mutual defection is P: In general, the outcome of the game should not depend on P as long as the structural properties of the payoff matrix is the same. However, in our case the situation is different. Indeed, if $P < 0$, an agent might decide to avoid the game by isolating himself (i.e., linking to agents that do not reciprocate), whereas for $P > 0$ the agent might be better of participating in a game.

To resolve this issue, we assume that every time a partner of agent x refuses to play, x receives a negative payoff $-c_p < 0$, which can be viewed as a *cost of isolation*. The introduction of this cost merely means adding a constant to the reward matrix. The adjusted reward matrix elements a_{ij} are given by $a_{ij} = b_{ij} + c_p$, where B is the game reward matrix and similar for all agents.

3. REST-POINTS AND LOCAL STABILITY

To examine the emergent network structures, we need to study the stable rest points of the replicator equations. Those rest points can be found by nullifying the right hand sides of Equations 3 and 4. Furthermore, the stability of those rest points are characterized by the eigenvalues of the corresponding Jacobian matrix

$$J = \begin{pmatrix} \frac{\partial \dot{c}_{ij}}{\partial c_{mn}} & \frac{\partial \dot{c}_{ij}}{\partial p_m} \\ \\ \frac{\partial \dot{p}_m}{\partial c_{ij}} & \frac{\partial \dot{p}_m}{\partial p_n} \end{pmatrix} = \begin{pmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{pmatrix} \quad (5)$$

For two-action games, the diagonal blocks J_{11} and J_{22} are $L \times L$ and $N \times N$ square matrices, respectively, where $L = N(N-2)$. Similarly, J_{12} and J_{21} are $L \times N$ and $N \times L$ matrices, respectively.

We have performed thorough analytical characterization of the above system for three-player two-action games, which is the minimal system that exhibits non-trivial structural dynamics. In particular, we have demonstrated that for this class of games it is possible to characterize all the rest-points of the learning dynamics and examine their stability properties analytically.

We have also examined the behavior of the co-evolving system for large number of agents using both simulations and analytical techniques. We found that in the asymptotic limit, the networks formed by the reciprocated links (i.e., $c_{xy} c_{yx} \neq 0$) consists of $star$ motifs. A star graph S_n is a graph with n nodes and $n-1$ links, connecting one central node with the other $n-1$ nodes. We further observed that the basin of attraction of motifs shrinks as the motif size grows, so that smaller motifs are more prevalent.

As an example, we performed simulations for 100 agents interacting via the following Prisoner's Dilemma (PD) reward matrix:

$$B = \begin{pmatrix} (3,3) & (0,5) \\ (5,0) & (1,1) \end{pmatrix}$$

We run the simulation for 5000 different random initializations. When the cost of isolation is sufficiently large, $c_p \geq -b_{22}$, then the network breaks down into isolated star motifs. We observed that 91.26% of the motifs are S_2, 8.41% are S_3, 0.32% S_4, and 0.02% of S_5 star motifs. Within those stable networks, all the players choose the second action (*defect*). Furthermore, we have shown analytically that in a system with N-agents the star network S_N is in fact a stable rest point of the learning dynamics. Finally, when $c_p < -b_{22}$, the network structure is such that links are not reciprocated, so that the agents effectively do not interact.

4. DISCUSSION

In conclusion, we have presented a replicator–dynamics based framework for studying mutual evolution of network topology and agent behavior in a network–augmented system of interacting adaptive agents. By assuming that the agents' strategies allow appropriate factorization, we derived a system of a coupled replicator equations that describe the mutual evolution of agent behavior and network link structure. For N-player games, we reported both simulation and analytical results, which suggests that star-like structures are the most prevalent motifs emerging in our game-dynamical network formation. As future work, we plan to perform a more thorough analysis of the N-agent systems. Finally, we note that the main premise behind our model is that the strategies can be factorized according to Equations 2. While this assumption seems to be justified for certain games, its limitations need to be studied further.

5. ACKNOWLEDGMENTS

This research was supported in part by the National Science Foundation under grant No. 0916534 and the US AFOSR MURI grant No. FA9550-10-1-0569.

6. REFERENCES

[1] S. Castellano, C. Fortunato and V. Loreto. Statistical physics of social dynamics. *Rev. Mod. Phys.*, 81(2):591, 2009.

[2] T. Gross and B. Blasius. Adaptive coevolutionary networks: a review. *J. of the Royal Soc. Interface*, 5(20):259, 2008.

[3] J. Hofbauer and K. Sigmund. *Evolutionary games and Population dynamics*. Cambridge University Press, 1998.

[4] M. Perc and A. Szolnoki. Coevolutionary games–a mini review. *BioSystems*, 99(2):109–125, 2010.

[5] Y. Sato and J.P. Crutchfield. Coupled replicator equations for the dynamics of learning in multiagent systems. *Physical Review E*, 67(1), 2003.

[6] C.J.C.H. Watkins and P. Dayan. Technical note: Q-learning. *Machine learning*, 8(3):279–292, 1992.